First measurement of the CP-violating phase in $B_{s}^{0} \rightarrow J/\psi(\rightarrow e^{+}e^{-})\phi$ decays

LHCb collaboration†

Abstract
A flavour-tagged time-dependent angular analysis of $B_{s}^{0} \rightarrow J/\psi\phi$ decays is presented where the $J/\psi$ meson is reconstructed through its decay to an $e^{+}e^{-}$ pair. The analysis uses a sample of $pp$ collision data recorded with the LHCb experiment at centre-of-mass energies of 7 and 8 TeV, corresponding to an integrated luminosity of 3 fb$^{-1}$. The CP-violating phase and lifetime parameters of the $B_{s}^{0}$ system are measured to be $\phi_{s} = 0.00 \pm 0.28 \pm 0.05$ rad, $\Delta \Gamma_{s} = 0.115 \pm 0.045 \pm 0.011$ ps$^{-1}$ and $\Gamma_{s} = 0.608 \pm 0.018 \pm 0.011$ ps$^{-1}$ where the first uncertainty is statistical and the second systematic. This is the first time that CP-violating parameters are measured in the $B_{s}^{0} \rightarrow J/\psi\phi$ decay with an $e^{+}e^{-}$ pair in the final state. The results are consistent with previous measurements in other channels and with the Standard Model predictions.

Submitted to Eur. Phys. J. C
© 2021 CERN for the benefit of the LHCb collaboration. [CC BY 4.0 licence]

†Authors are listed at the end of this paper.
1 Introduction

The phase difference $\phi_s$ between direct decays and decays through mixing of $B_s^0$ mesons to Charge-Parity ($CP$) eigenstates is a $CP$-violating observable. In the Standard Model (SM), considering $b \rightarrow (c\bar{s})s$ transitions and neglecting 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 the elements of the CKM quark-flavour mixing matrix [1].

The precise measurement of the $\phi_s$ phase is potentially sensitive to new physics (NP) processes. The measured phase could be modified if new particles were to contribute to the $B_s^0$–$\bar{B}_s^0$ mixing amplitudes [2]. Measurements of $\phi_s$ using different decay channels with muons in the final state, namely $B_s^0 \rightarrow J/\psi K^+K^-$ [3,4], $B_s^0 \rightarrow J/\psi\pi^+\pi^-$ [5], $B_s^0 \rightarrow \psi(2S)\phi$ [6], and a channel with open charm mesons, $B_s^0 \rightarrow D_s^+D_s^-$ [7], have been reported previously by the LHCb collaboration. Measurements of $\phi_s$ in $B_s^0 \rightarrow J/\psi\phi$ decays with $J/\psi \rightarrow \mu^+\mu^-$ have also been performed by the ATLAS [8,9], CMS [10], CDF [11] and D0 [12] collaborations. The world-average value of these measurements is $\phi_s = -0.051 \pm 0.023\,\text{rad}$ [13]. A precise prediction of the $\phi_s$ phase value is available from global fits of the CKM matrix within the SM. The CKMfitter group result is $\phi_s = -0.0365 \pm 0.0013\,\text{rad}$ [14] while the UTfit collaboration result is $\phi_s = -0.0370 \pm 0.0010\,\text{rad}$ [15].

This paper presents a measurement of $\phi_s$ using a flavour-tagged time-dependent angular analysis of the $B_s^0 \rightarrow J/\psi\phi$ mode with $J/\psi \rightarrow e^+e^-$ and $\phi \rightarrow K^+K^-$ decays. This is the first time that the $B_s^0 \rightarrow J/\psi(e^+e^-)\phi$ decay is used to measure $CP$-violating observables, and in particular the phase $\phi_s$. The analysis is based on a data set corresponding to an integrated luminosity of 3 fb$^{-1}$ collected at the LHC in proton-proton ($pp$) collisions at centre-of-mass energies of 7 and 8 TeV by the LHCb experiment. The yield of the $B_s^0 \rightarrow J/\psi(e^+e^-)\phi(K^+K^-)$ sample amounts to about 10% of that of the previously analysed $B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ mode using the same data set [16]. The analysis follows closely that of the two muons decay mode, reported in Refs. [3,5]. Relevant changes are described in more detail in this paper.

A comparison of the two results is of interest given the different main sources of systematic uncertainties induced by the markedly different reconstruction of decays with muons in the final state compared to decays with electrons. These differences arise from the significant bremsstrahlung emission of the electrons and the different signatures exploited in the online trigger selection [17,19].

The article is structured in the following way. The phenomenological description of the $B_s^0 \rightarrow J/\psi(e^+e^-)\phi(K^+K^-)$ decay and the relevant physics observables are described in Sec. 2. A brief description of the LHCb detector, the candidates selection and the background subtraction are outlined in Sec. 3. The relevant inputs to the analysis, namely the resolution, efficiency and the flavour tagging, are detailed in Sec. 4 and 5. The maximum-likelihood fit procedure used to determine the physics parameters and the results of the fit are described in Sec. 6, while the evaluation of the systematic uncertainties is discussed in Sec. 7. Finally, conclusions are presented in Sec. 8.

\footnote{The inclusion of charge-conjugate processes is implied throughout this paper, unless otherwise noted. For simplicity, the resonance $\phi(1020)$ is referred to as $\phi$ here and in the following.}
Figure 1: Definition of the angles in the helicity basis. The polar angle \( \theta_K (\theta_e) \) is the angle between the \( K^+ (e^+) \) momentum and the direction opposite to the \( B_{s0}^0 \) momentum in the \( K^+K^- (e^+e^-) \) centre-of-mass system, and the \( \phi_h \) is the azimuthal angle between the \( K^+K^- \) and \( e^+e^- \) decay planes.

2 Phenomenology

The phenomenological aspects of the analysis are presented in Ref. [20]. This formalism also holds for the \( B_{s0}^0 \to J/\psi(e^+e^-)\phi(K^+K^-) \) decay. Angular momentum conservation in the \( B_{s0}^0 \to J/\psi\phi \) decay implies that the final state is an admixture of two \( CP \)-even and one \( CP \)-odd components, with orbital angular momentum of 0 or 2, and 1, respectively. Moreover, along with the three P-wave states of the \( \phi \to K^+K^- \) transition, there is also a \( CP \)-odd \( K^+K^- \) component in an S-wave state [21]. The \( CP \)-even and \( CP \)-odd components are disentangled by a time-dependent angular analysis, where the angular observables \( \Omega = \{\cos \theta_e, \cos \theta_K, \phi_h\} \) are defined in the helicity basis as shown in Fig. 1. The polar angle \( \theta_K \) (\( \theta_e \)) is the angle between the \( K^+ \) (\( e^+ \)) momentum and the direction opposite to the \( B_{s0}^0 \) momentum in the \( K^+K^- (e^+e^-) \) centre-of-mass system. The azimuthal angle between the \( K^+K^- \) and \( e^+e^- \) decay planes is \( \phi_h \). A definition of the angles in terms of the particles momenta can be found in Ref. [20].

The differential decay rate for \( B_{s0}^0 \to J/\psi\phi \) decay as a function of the decay time and angles can be expressed as a sum of polarisation amplitudes and their interference terms. Each of these can be factorised into a part dependent on the decay time \( t \) and a part dependent on the set of angular variables \( \Omega \), as

\[
G(t, \Omega) \equiv \frac{d^4\Gamma(B_{s0}^0 \to J/\psi\phi)}{dt\,d\Omega} \propto \sum_{k=1}^{10} h_k(t) f_k(\Omega). \tag{1}
\]

The time-dependent functions \( h_k(t) \) are given as

\[
h_k(t|B_{s0}^0) = N_k e^{-\Gamma_s t} \left[ a_k \cosh \frac{\Delta \Gamma_s t}{2} + b_k \sinh \frac{\Delta \Gamma_s t}{2} + c_k \cos(\Delta m_s t) + d_k \sin(\Delta m_s t) \right], \tag{2}
\]

\[
h_k(t|\bar{B}_{s0}^0) = \bar{N}_k e^{-\Gamma_s t} \left[ a_k \cosh \frac{\Delta \Gamma_s t}{2} + b_k \sinh \frac{\Delta \Gamma_s t}{2} - c_k \cos(\Delta m_s t) - d_k \sin(\Delta m_s t) \right], \tag{3}
\]

where \( \Delta \Gamma_s \equiv \Gamma_L - \Gamma_H \) is the decay width difference between the light and the heavy \( B_s \) mass eigenstates, \( \Delta m_s \equiv m_H - m_L \) is their mass difference, and \( \Gamma_s \equiv (\Gamma_L + \Gamma_H)/2 \) is their average width. The coefficients \( N_k \) (\( \bar{N}_k \)) and \( a_k, b_k, c_k, d_k \) can be expressed in terms of \( \phi_s \) and four complex transversity amplitudes \( A_i \) (\( \bar{A}_i \)) at \( t = 0 \), as detailed in Table 1. The label \( i \) takes the values \( \{\perp, \parallel, 0\} \) for the three P-wave amplitudes and \( S \) for the S-wave amplitude. The amplitudes are parameterised by \( |A_i|e^{i\delta_i} \) with the
Table 1: Definition of angular and time-dependent functions for $B^0_s$ and $\bar{B}^0_s$ mesons.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$f_k(\theta_{K}, \theta_{s}, \phi_{s})$</th>
<th>$N_k$</th>
<th>$\bar{N}_k$</th>
<th>$a_k$</th>
<th>$b_k$</th>
<th>$c_k$</th>
<th>$d_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$2 \cos^{2} \theta_{K} \sin^{2} \theta_{s}$</td>
<td>$</td>
<td>A_{0}</td>
<td>^2$</td>
<td>$</td>
<td>A_{0}</td>
<td>^2$</td>
</tr>
<tr>
<td>2</td>
<td>$\sin^{2} \theta_{K} (1 - \sin^{2} \theta_{s} \cos^{2} \phi_{s})$</td>
<td>$</td>
<td>A_{1}</td>
<td>^2$</td>
<td>$</td>
<td>A_{1}</td>
<td>^2$</td>
</tr>
<tr>
<td>3</td>
<td>$\sin^{2} \theta_{K} (1 + \sin^{2} \theta_{s} \cos^{2} \phi_{s})$</td>
<td>$</td>
<td>A_{1}</td>
<td>^2$</td>
<td>$</td>
<td>A_{1}</td>
<td>^2$</td>
</tr>
<tr>
<td>4</td>
<td>$\sin^{2} \theta_{K} \sin^{2} \theta_{s} \sin 2\phi_{s}$</td>
<td>$</td>
<td>A_{01}</td>
<td>^2$</td>
<td>$</td>
<td>A_{01}</td>
<td>^2$</td>
</tr>
<tr>
<td>5</td>
<td>$\frac{1}{2} \sqrt{2} \sin 2\theta_{K} \sin 2\theta_{s} \cos \phi_{s}$</td>
<td>$</td>
<td>A_{01}</td>
<td>$</td>
<td>$</td>
<td>A_{01}</td>
<td>$</td>
</tr>
<tr>
<td>6</td>
<td>$-\frac{1}{2} \sqrt{2} \sin 2\theta_{K} \sin 2\theta_{s} \sin \phi_{s}$</td>
<td>$</td>
<td>A_{01}</td>
<td>$</td>
<td>$</td>
<td>A_{01}</td>
<td>$</td>
</tr>
<tr>
<td>7</td>
<td>$\frac{1}{2} \sqrt{2} \sin^{2} \theta_{s}$</td>
<td>$</td>
<td>A_{0}</td>
<td>^2$</td>
<td>$</td>
<td>A_{0}</td>
<td>^2$</td>
</tr>
<tr>
<td>8</td>
<td>$\frac{1}{2} \sqrt{3} \sin \theta_{K} \sin 2\theta_{s} \cos \phi_{s}$</td>
<td>$</td>
<td>A_{01}</td>
<td>$</td>
<td>$</td>
<td>A_{01}</td>
<td>$</td>
</tr>
<tr>
<td>9</td>
<td>$\frac{1}{2} \sqrt{6} \sin \theta_{K} \sin 2\theta_{s} \sin \phi_{s}$</td>
<td>$</td>
<td>A_{01}</td>
<td>$</td>
<td>$</td>
<td>A_{01}</td>
<td>$</td>
</tr>
<tr>
<td>10</td>
<td>$\frac{1}{2} \sqrt{3} \cos \theta_{K} \sin^{2} \theta_{s}$</td>
<td>$</td>
<td>A_{0}</td>
<td>$</td>
<td>$</td>
<td>A_{0}</td>
<td>$</td>
</tr>
</tbody>
</table>

conventions $\delta_0 = 0$ and $|A_{1}|^2 + |A_{0}|^2 + |A||^2 = 1$. The S-wave fraction is defined as $F_S = |A_{0}|^2 / (|A_{1}|^2 + |A_{0}|^2 + |A||^2)$. In contrast to Ref. [3], the S-wave parameters are measured in a single range of $m(K^+ K^-)$ within $\pm 30$ MeV/c$^2$ of the known $\phi$ mass [13]. For a particles produced in a $B^0_s$ and $\bar{B}^0_s$ flavour eigenstates, the coefficients in Eqs. (2) and (3), respectively are given in Table 1, together with the angular functions $f_k(\Omega)$, where the $S$, $D$, $C$ coefficients are defined as

$$S = -\frac{2|\lambda|}{1 + |\lambda|^2} \sin(\phi_s), \quad D = -\frac{2|\lambda|}{1 + |\lambda|^2} \cos(\phi_s) \quad \text{and} \quad C = \frac{1 - |\lambda|^2}{1 + |\lambda|^2}. \quad (4)$$

The parameter $\lambda$ is related to CP violation in the interference between mixing and decay, and is defined by $\lambda = \eta_i(q/p)(\bar{A}_i/A_i)$ where the polarisation states $i$ have the CP eigenvalue $\eta_i = +1$ for $i \in \{0, \parallel\}$ and $\eta_i = -1$ for $i \in \{\perp, S\}$. The complex parameters $p$ and $q$ relate the mass eigenstates to the flavour eigenstates, $|B_{L,H}| = p|B^0_s| \pm q|\bar{B}^0_s|$. The CP-violating phase is defined by $\phi_s \equiv -\arg(\lambda)$ and is assumed here to be the same for all polarisation states. The value of $|\lambda|$ equals unity in the absence of CP violation in decay [22, 24]. In this paper, the CP violation in $B_s$ meson mixing is assumed to be negligible, following the measurements in Refs. [25, 26].

3 Detector, data set and selection

The LHCb detector [27] 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 $pp$ collision vertex (PV), the impact parameter (IP), 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 (RICH). Photons, electrons,
and hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter (ECAL), and a hadronic calorimeter. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

Samples of simulated events are used to optimise the signal selection, to derive the angular efficiency and to correct the decay-time efficiency. The simulated pp collisions are generated using PYTHIA [28] with a specific LHCb configuration [29]. The decays of hadronic particles are described by EvtGen [30], in which final-state radiation is generated using PHOTOS [31]. The interaction of the generated particles with the detector and its response are implemented using GEANT4 toolkit [32], as described in Ref. [33].

The online candidate selection is performed by a trigger [34], which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full decay reconstruction. At the hardware stage, events are required to have a hadron or electron with a high transverse-energy deposit in the calorimeters, $E_T > 3$ GeV and $E_T > 3.68$ GeV, respectively. The subsequent software trigger is implemented as two separate levels that further reduce the event rate. The first level is designed to select decays which are displaced from all PVs. At the second level, $B^0_s \rightarrow J/\psi \phi$ candidates are selected by identifying events containing a pair of oppositely charged kaons with an invariant mass within $\pm 30$ MeV/c$^2$ of the known $\phi$-meson mass [13] or by using topological $b$-hadron triggers. These topological triggers require a two-, three- or four-track secondary vertex with a large sum of the $p_T$ of the charged particles and significant displacement from all PVs. A multivariate algorithm [35] is used for the identification of secondary vertices consistent with the decay of a $b$ hadron. The trigger signals are associated with reconstructed particles in the offline selection. The candidate selection is devised in order to minimise the impact on the decay-time efficiency.

Electrons radiate bremsstrahlung photons when travelling through the detector material. For events where the photons are emitted upstream of the spectrometer magnet, the photon and the electron deposit their energy in different ECAL cells, and the electron momentum measured by the tracking system is underestimated. Neutral energy deposits in the ECAL compatible with being emitted by the electron are used to correct for this effect. The limitations of the recovery technique degrade the resolution of the reconstructed invariant masses of both the di-electron pair and the $B^0_s$ candidate [17].

In the offline selection, $J/\psi$ candidates are formed from two oppositely charged tracks identified as electrons, and $\phi$ candidates from pairs of oppositely charged tracks identified as kaons. The pairs of tracks need to form a good quality vertex. The electron candidates are required to have $p_T > 0.5$ GeV/c and di-electron invariant mass $m(e^+e^-) \in [2.5, 3.3]$ GeV/c$^2$, where a wider range compared to the dimuon mode analysis is chosen to account for the radiative tail arising due to bremsstrahlung. The $p_T$ of the $\phi$ candidate is required to be larger than 1 GeV/c.

The $J/\psi$ and $\phi$ candidates that are consistent with originating from a common vertex are combined to form $B^0_s$ candidates. The mass of the $B^0_s$ candidates is required to be in the range $m(e^+e^-K^+K^-) \in [4.7, 5.6]$ GeV/c$^2$. The reconstructed decay time of the $B^0_s$ candidate, $t$, is obtained from a kinematic fit with the $J/\psi$ mass constrained to its known value [13] and the $B^0_s$ candidate constrained to originate from the associated PV. Each $B^0_s$ candidate is associated with the PV that yields the smallest $\chi^2_{IP}$, where $\chi^2_{IP}$ is defined as the difference in the vertex-fit $\chi^2$ of a given PV reconstructed with and without the particle under consideration. The $B^0_s$ candidates are selected if they have decay times in
the range $0.3 < t < 14$ ps and decay-time uncertainty estimates $\sigma_t < 0.12$ ps. The fraction of events containing more than one $B_s^0$ candidate within the $m(e^+e^-K^+K^-)$ range is 2.6%. All candidates are retained in the subsequent analysis. The impact of allowing multiple candidates per event is negligible.

The main sources of background are partially reconstructed $b$-hadron decays and combinatorial background. The first of these arises from the $B_s^0 \rightarrow \chi_{c1}(1P)(\rightarrow J/\psi\gamma)\phi$ and $B_s^0 \rightarrow \psi(2S)(\rightarrow J/\psi X)\phi$ decay.\footnote{The symbol $X$ stands for unreconstructed particles.} The combinatorial background is due to random combination of tracks in the event that pass the candidate selection. In addition, possible background contributions to the signal region originate from $A_0^0 \rightarrow J/\psi pK^-$ and $B^0 \rightarrow J/\psi K^*(892)^0$ decays, where the proton or the $\pi^-$ meson from the $K^*(892) \rightarrow K^+\pi^-$ decay is misidentified as a $K^+$ or $K^-$ meson, respectively.

The combinatorial background is suppressed using a boosted decision tree (BDT)\footnote{The symbol $X$ stands for unreconstructed particles.} analysis, trained using the TMVA toolkit\cite{37}. The BDT discriminant is trained using a signal sample of simulated $B_s^0 \rightarrow J/\psi\phi$ decays, and a sample of background from data. For the background same-sign combinations of electron and/or kaon pairs are chosen with the same selection criteria as for signal. The simulation is corrected to match the distributions observed in data for variables used in the identification of electrons and kaons. The eight variables used for the training of the BDT discriminant are the transverse momenta of the $J/\psi$ and $\phi$ candidates, the vertex $\chi^2$ of the $B_s^0$ candidate, the $\chi^2$ of the kinematic fit of the $B_s^0$ candidate with the $J/\psi$ mass constrained to its known value and the electron and kaon identification probability as provided mainly from the RICH and calorimeter systems. The optimal working point for the BDT discriminant 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 the number of signal and background candidates into account\cite{38}.

The candidates are rejected if the $K^+$ candidate can also be identified as a proton by a dedicated neural network\cite{39} to suppress any possible contamination from $A_0^0 \rightarrow J/\psi pK^-$ decays. The remaining misidentified background contribution is estimated using simulated samples and amounts to 1% of the expected signal yield for $A_0^0$ decays and is negligible for $B_s^0$ decays.

Figure\footnote{The symbol $X$ stands for unreconstructed particles.} 2 shows the distribution of $m(e^+e^-K^+K^-)$ for the selected $B_s^0 \rightarrow J/\psi\phi$ candidates. In order to describe better the left tail of the $m(e^+e^-K^+K^-)$ distribution, the sample is split into three categories by the number of electron candidates: zero, one or both electrons of the pair that received bremsstrahlung corrections. An extended maximum-likelihood fit is made to the unbinned $m(e^+e^-K^+K^-)$ distribution.

In the fit the signal component is described by the sum of two Crystal Ball (CB) functions\cite{40} and the combinatorial background by an exponential function. The partially reconstructed background components from $B_s^0 \rightarrow \chi_{c1}(1P)\phi$ and $B_s^0 \rightarrow \psi(2S)\phi$ decays are modelled using a Gaussian function and the sum of two Gaussian functions, respectively. The parameters that describe the shape of the signal candidates and the partially reconstructed background are fixed to values obtained from simulation. The core widths and the common mean of the CB functions are left free in the fit. The fit to the three categories gives a yield of $(1.27 \pm 0.05) \times 10^4$ signal candidates where the uncertainty is statistical only.

The fit results are used to assign per-candidate weights via the sPlot technique with $m(e^+e^-K^+K^-)$ as the discriminating variable\cite{41}. This is used to subtract the background
Figure 2: Distribution of $m(e^+e^-K^+K^-)$ for selected $B^0_s \to J/\psi \phi$ candidates divided into three categories: (a) zero, (b) one and (c) both electrons with bremsstrahlung correction. The blue solid line shows the total fit which is composed of (red short-dashed line) the signal and the background contributions. The combinatorial background is indicated by the green long-dashed line while the partially reconstructed background from the $B^0_s \to \psi(2S)\phi$ and $B^0_s \to \chi_{c1}(1P)\phi$ decays are indicated by pink and purple dash-dotted lines, respectively.

contribution in the maximum-likelihood fit described in Sec. 6. As the three categories are statistically independent further steps of the analysis are performed on the combined sample.

4 Detector resolution and efficiency

The finite decay-time resolution is a diluting factor that will affect the relative precision of $\phi_s$ and has to be accounted for. The way this is introduced into the analysis is described in Sec. 6. The assumed decay-time resolution model, $R$, consists of a sum of two Gaussian distributions with their widths depending on the per-candidate decay-time uncertainty determined by the vertex fit as detailed in Ref. [16]. The parameters of this model are loosely constrained in the fit of the $B^0_s \to J/\psi(e^+e^-)K^+K^-$ decay to the values determined using an identical model from a sample of $J/\psi \to \mu^+\mu^-$ candidates produced at the PV. They are allowed to vary within a Gaussian constraint of twice the difference of their values between the electron and muon modes as extracted from simulation. The parameters are determined from the unbinned maximum-likelihood fit, as described in Sec. 6. Taking into account the $\sigma_t$ distribution of the $B^0_s$ signal, the resulting effective resolution is $45.6 \pm 0.5$ fs.

Due to the displacement requirements made on signal tracks in the trigger and offline selections, the reconstruction efficiency depends on the decay time of the $B^0_s$ candidate.
The efficiency is determined with the same method as described in Ref. [6], by using the control channel $B^0 \to J/\psi K^*(892)^0$, with $J/\psi \to e^+e^-$ and $K^*(892)^0 \to K^+\pi^-$ decays.

The decay-time dependence of the signal efficiency is determined as

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

where $\varepsilon_{\text{data}}(t)$ is the efficiency of the control channel, determined on data, and $\varepsilon_{\text{sim}}^B(t)/\varepsilon_{\text{sim}}^B(t)$ is the ratio of efficiencies of the simulated signal and control modes after the selection. The efficiencies are extracted by normalisation to the known lifetimes of $\tau_{B^0} = 1.527 \pm 0.011$ ps and $\tau_{B^0} = 1.520 \pm 0.004$ ps [13]. The second term accounts for the small differences in the decay time and kinematics between the signal and the control modes. The control channel efficiency is defined as $\varepsilon_{\text{data}}(t) = N_{\text{data}}^B(t)/N_{\text{gen}}^B(t)$ where $N_{\text{data}}^B(t)$ is the number of the $B^0 \to J/\psi K^*(892)^0$ decays in a given time bin as determined using sPlot technique [41] with $m(e^+e^-K^+\pi^-)$ as discriminating variable. The $N_{\text{gen}}^B(t)$ is the number of events generated from an exponential distribution with lifetime $\tau_{B^0}$ [13]. The analysis is not sensitive to the absolute scale of the efficiency.

The $B^0 \to J/\psi K^*(892)^0$ decay is selected using trigger, selection and BDT requirements similar to those used for the signal, adapted to the different final states. The background contribution to the control sample from the misidentification of final-state particles from...
the \( \Lambda^0_b \to J/\psi\pi^- \) decay is estimated to be 0.06% of the expected signal yield, while the background contribution from \( B^0_s \to J/\psi\phi \) decays is negligible.

The \( m(e^+e^-K^+\pi^-) \) invariant-mass distribution is shown in Fig. 3 divided into the three bremsstrahlung categories, as for the signal sample. The contribution from \( B^0 \to J/\psi K^*(892)^0 \) decays is described by the sum of two CB functions while an exponential function is used to describe the combinatorial background. Similarly to the signal sample, partially reconstructed background arises from \( B^0 \to J/\psi K^*(892)^0 \) decays where one or more particles are not reconstructed; background components stemming from \( B^0 \to \chi_{c1}(1P)(\to J/\psi \gamma)K^*(892)^0 \), \( B^0 \to \psi(2S)(\to J/\psi X)K^*(892)^0 \) and \( B^0 \to J/\psi K_1(1270)^0(\to K^*(892)^0\pi^0) \) decays\(^2\) are described using a single Gaussian function, the sum of two Gaussian functions and the sum of two CB functions, respectively. The \( B^0 \to J/\psi K^*(892)^0 \) yield is found to be \((5.45 \pm 0.05) \times 10^4\) signal candidates.

The decay-time efficiency for the \( B^0_s \to J/\psi\phi \) signal is shown in Fig. 4. The efficiency is relatively uniform at high values of decay time but decreases at low decay times due to the selection criteria that require displaced tracks.

The efficiency as a function of the \( B^0_s \to J/\psi\phi \) helicity angles is not uniform due to the forward geometry of the LHCb detector and the requirements imposed on the final-state particle momenta. Projections of the three-dimensional efficiency, \( \varepsilon(\Omega) \), to the three helicity angles are shown in Fig. 5. The angular efficiency correction is introduced in the analysis through normalisation integrals in the probability density function describing the signal decays in the fit described in Sec. 6. The integrals given in Table 2 are calculated using simulated candidates that are subject to the same trigger and selection criteria as the data, following the same technique as in Ref. [20]. The relative efficiency is constant for the azimuthal angle \( \phi_h \). A dependence of up to 15% is observed for \( \cos \theta_e \) and \( \cos \theta_K \). The finite angular resolution has small impact on the results of the analysis and is neglected. A systematic uncertainty is assigned to account for this effect.

5 Flavour tagging

The \( B_s \) candidate flavour at production is determined by two independent categories of flavour tagging algorithms, the opposite-side (OS) taggers [42] and the same-side kaon (SSK) tagger [43], which exploit specific features of the production of \( b\bar{b} \) quark pairs in
Table 2: Angular acceptance integrals for the simulated sample. The \( I_k \) integrals are normalised with respect to the \( I_0 \) integral.

<table>
<thead>
<tr>
<th>( k )</th>
<th>( I_k/I_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (00)</td>
<td>0.9801 ± 0.0014</td>
</tr>
<tr>
<td>2 (∥∥)</td>
<td>1.0200 ± 0.0017</td>
</tr>
<tr>
<td>3 (⊥⊥)</td>
<td>1.0209 ± 0.0016</td>
</tr>
<tr>
<td>4 (∥⊥)</td>
<td>0.0003 ± 0.0018</td>
</tr>
<tr>
<td>5 (0</td>
<td></td>
</tr>
<tr>
<td>6 (0⊥)</td>
<td>0.0015 ± 0.0012</td>
</tr>
<tr>
<td>7 (SS)</td>
<td>0.9983 ± 0.0011</td>
</tr>
<tr>
<td>8 (S∥)</td>
<td>0.0004 ± 0.0016</td>
</tr>
<tr>
<td>9 (S⊥)</td>
<td>0.0012 ± 0.0016</td>
</tr>
<tr>
<td>10 (S0)</td>
<td>−0.0007 ± 0.0006</td>
</tr>
</tbody>
</table>

Figure 5: Efficiency projected onto (left) \( \cos \theta_K \), (middle) \( \cos \theta_e \) and (right) \( \phi_h \) obtained from a simulated \( B_0^s \to J/\psi \phi \) sample, scaled by the average efficiency.

The \( \eta \) and \( \omega \) are estimated event-by-event, and represent the probability that the algorithm assigns a wrong tag decision. The effective tagging power is given by \( \varepsilon_{\text{tag}}(1 - 2\omega)^2 \) and for the combined taggers in the \( B_0^s \to J/\psi(e^+e^-)\phi \) signal sample a value of \( (5.07 \pm 0.16)\% \) is obtained.

6 Maximum-likelihood fit and results

The \( CP \) observables are determined by an unbinned maximum-likelihood fit to the background-subtracted candidates in four-dimensions, namely the \( B_0^s \) decay time and...
the three helicity angles, with a probability density function (PDF) describing $B^0_s \to J/\psi(e^+e^-)\phi$ signal decay. The negative log-likelihood function to be minimised is given by

$$-\ln L = -\alpha \sum_{i=1}^{N} w_i \ln P,$$

(6)

where $N$ is the total number of candidates. The $w_i$ coefficients are the sPlot weights [41] computed using $m(e^+e^-K^+K^-)$ as discriminating variable, and the factor $\alpha = \sum w_i / \sum w^2_i$ is used to account for the correct signal yield in the sample. The PDF, $P = S/\int Sdt d\Omega$, is normalised over the four-dimensional space where

$$S(t, \Omega, q^{OS}, q^{SSK}|\eta^{OS}, \eta^{SSK}) = T(t', \Omega, q^{OS}, q^{SSK}|\eta^{OS}, \eta^{SSK}) \otimes R(t-t'|\sigma_t) \times \epsilon_{data}(t),$$

(7)

with the decay-time resolution function, $R$, defined in Sec. 4 and $T(t', \Omega, q^{OS}, q^{SSK}|\eta^{OS}, \eta^{SSK}) = (1 + q^{OS} (1 - 2\omega^{OS})) (1 + q^{SSK} (1 - 2\omega^{SSK})) G(t, \Omega) + (1 - q^{OS} (1 - 2\bar{\omega}^{OS})) (1 - q^{SSK} (1 - 2\bar{\omega}^{SSK})) \bar{G}(t, \Omega),$ (8)

which allows for the inclusion of the information from both tagging algorithms in the computation of the decay rate. The function $G(t, \Omega)$ is defined in Eq. (1) and $\bar{G}(t, \Omega)$ is the corresponding function for $B^0_s$ decays. The angular efficiency is included in the normalisation of the PDF via the ten integrals, $I_k = \int d\Omega \epsilon(\Omega) f_k(\Omega)$. The integrals are pre-calculated using simulation as described in Sec. 4.

When using weights from the sPlot method, the standard uncertainty estimate based on the Hessian matrix will generally not give asymptotically correct confidence intervals [44]. A bootstrap method [45] is used to obtain a correct estimate of the statistical uncertainty. The weights are recalculated for each bootstrap sample. In the fit, Gaussian constraints are included for certain nuisance parameters, namely the mixing frequency $\Delta m_s = 17.757 \pm 0.021 \text{ps}^{-1}$ [13], the tagging calibration parameters, and the time resolution parameters. The fitting procedure is validated using pseudoexperiments and simulated $B^0_s \to J/\psi(e^+e^-)\phi$ decays.

The results of the fit to the data are shown in Table 3 while the projections of the fit results on the decay time and helicity-angle distributions are reported in Fig. 6. The correlation matrix of statistical uncertainties is reported in Table 5 of Appendix A. The results are consistent with previous measurements of these parameters [3,8 –12], and the SM predictions for $\phi_s$ [22–24]. They show no evidence of $CP$ violation in the interference between $B^0_s$ meson mixing and decay, nor for direct $CP$ violation in $B^0_s \to J/\psi(e^+e^-)\phi$ decays, as the parameter $|\lambda|$ is consistent with unity within uncertainties.

7 Systematic uncertainties

Systematic uncertainties for each of the measured parameters are reported in Table 4. They are evaluated by observing the change in the physics parameters after repeating the likelihood fit with a modified model assumption, or through pseudoexperiments, in case of uncertainties originating from the limited size of calibration samples.

The decay-time and angular efficiencies obtained independently in the three bremsstrahlung categories are compatible within statistical uncertainties. While the
Table 3: Results of the maximum-likelihood fit, described in Sec. [6] to the $B^+_s \to J/\psi(e^+e^-)\phi$ decays including all acceptance and resolution effects. The first uncertainty is statistical and the second is systematic.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fit result and uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>$0.608 \pm 0.018 \pm 0.011$</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>$0.115 \pm 0.045 \pm 0.011$</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>$3.11^{+0.08}_{-0.07} \pm 0.06$</td>
</tr>
<tr>
<td>$\delta_\perp$ [rad]</td>
<td>$2.41^{+0.43}_{-0.43} \pm 0.10$</td>
</tr>
<tr>
<td>$\phi_s$ [rad]</td>
<td>$0.00 \pm 0.28 \pm 0.05$</td>
</tr>
<tr>
<td>$</td>
<td>\lambda</td>
</tr>
<tr>
<td>$F_S$</td>
<td>$0.062^{+0.042}_{-0.051} \pm 0.022$</td>
</tr>
<tr>
<td>$\delta_S$ [rad]</td>
<td>$0.01^{+0.25}_{-0.27} \pm 0.04$</td>
</tr>
</tbody>
</table>

Figure 6: Decay time and helicity-angle distributions for (data points) $B^+_s \to J/\psi(e^+e^-)\phi$ decays with the one-dimensional projections of the PDF extracted in the maximum-likelihood fit. The solid blue line shows the total signal contribution, which is composed of (long-dashed red) CP-even, (short-dashed green) CP-odd and (dash-dotted purple) S-wave contributions.

The effective decay-time resolution differs for the three categories, it was verified with simulations that the result of a weighted average of three independent maximum-likelihood fits is consistent with the default one.

Repeating the mass fit in bins of the decay time and helicity angles shows that the mass resolution depends on $\cos \theta_e$ and $\cos \theta_K$. As the sPlot technique assumes that the discriminating variable is independent of the observables of interest, the effect of this
Table 4: Statistical and systematic uncertainties. A dash corresponds to systematic uncertainties that are negligible. Systematic uncertainties from different sources are added in quadrature.

| Source                  | $\Gamma_s$ [ps$^{-1}$] | $\Delta \Gamma_s$ [ps$^{-1}$] | $A^2_s$ | $A^0_s$ | $\delta_\parallel$ [rad] | $\delta_\perp$ [rad] | $\phi_s$ | $|\lambda|$ | $F_S$ | $\delta_S$ [rad] |
|-------------------------|------------------------|-------------------------------|---------|---------|--------------------------|----------------------|---------|------------|-------|------------------|
| Stat. uncertainty       | 0.018                  | 0.045                         | 0.034   | 0.029   | +0.08                    | +0.43                | 0.011   | 0.017      | 0.01  | +0.25           |
| Mass factorisation      | 0.003                  | 0.003                         | 0.005   | 0.007   | +0.06                    | +0.43                | 0.01    | 0.007      | 0.03  | +0.25           |
| Mass model              | 0.009                  | 0.005                         | 0.004   | 0.005   | +0.01                    | +0.09                | 0.03    | 0.011      | 0.03  | +0.25           |
| Ang. acceptance         | –                     | –                             | –       | –       | –                        | –                    | –       | –          | –    | –                |
| Time resolution         | 0.002                  | 0.008                         | 0.004   | 0.002   | +0.06                    | +0.02                | 0.01    | 0.005      | 0.03  | +0.25           |
| Time acceptance         | 0.003                  | 0.003                         | 0.001   | 0.001   | –                        | –                    | –       | 0.001      | –    | –                |
| MC (time acc.)          | 0.001                  | 0.001                         | 0.001   | –       | –                        | –                    | –       | –          | –    | –                |
| MC (ang. acc.)          | –                     | –                             | 0.001   | 0.001   | +0.01                    | +0.01                | 0.02    | 0.017      | 0.03  | –                |
| $A^0_s$ background      | 0.001                  | 0.001                         | 0.001   | 0.001   | +0.01                    | +0.01                | 0.01    | 0.005      | 0.01  | –                |
| Ang. resolution         | –                     | 0.002                         | 0.002   | 0.003   | –                        | –                    | –       | 0.005      | –    | –                |
| $B^+_s$ background      | 0.003                  | –                             | –       | –       | –                        | –                    | –       | –          | –    | –                |
| Fit bias                | –                     | –                             | 0.009   | –       | –                        | –                    | –       | 0.020      | –    | –                |
| Syst. uncertainty       | 0.011                  | 0.011                         | 0.008   | 0.013   | 0.06                      | 0.10                 | 0.05    | 0.031      | 0.022 | 0.04            |
| Total uncertainty       | 0.021                  | 0.046                         | 0.035   | 0.032   | 0.10                      | 0.44                 | 0.28    | +0.117     | +0.047| +0.25           |

correlation is quantified. The data sample is divided in intervals of $\cos \theta_e$ and $\cos \theta_K$ and new weights are computed with fits to $m(e^+e^-K^+K^-)$. The four-dimensional likelihood fit is evaluated with modified weights. The variation of each physics parameter is assigned as a systematic uncertainty. For the decay time and azimuthal $\phi_h$ angle the effect is negligible.

The mass model is tested in two ways. For the first test a new set of weights using a sum of two Ipatia functions $[46]$ is computed to describe the signal component of the $m(e^+e^-K^+K^-)$ distribution. For the second test a set of pseudoexperiments is used by fluctuating the default mass model parameters within their uncertainties (accounting for correlations), providing a new set of weights. The width of the obtained physics parameters distributions from the pseudoexperiments or the difference between the default and Ipatia based result is assigned as systematic uncertainty, whichever is larger.

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 matrix. The width of the obtained distributions for each physics parameter is taken as the systematic uncertainty. The angular resolution is neglected in the maximum-likelihood fit. The effect of this assumption is studied using pseudoexperiments, where the helicity angles are smeared according to the experimental resolution. There is a small effect on the polarisation amplitudes, strong phase and decay width difference while all other parameters are unaffected.

A systematic contribution is evaluated to take into account the effect of the finite decay-time resolution by comparing pseudoexperiments with fixed and constrained decay-time resolution parameters. A sample of pseudoexperiments with the four-dimensional $B^0_s \rightarrow J/\psi(e^+e^-)\phi$ PDF including time and angular efficiencies is used. The procedure is evaluated for two scenarios: the former with decay-time resolution parameters fixed...
to generated values, and the latter with parameters constrained to twice the difference between values obtained from signal simulation with $J/\psi \rightarrow e^+e^-$ and $J/\psi \rightarrow \mu^+\mu^-$ decays. The quadratic difference between the uncertainties of pseudoexperiments with fixed and constrained parameters is assigned as a systematic uncertainty. In addition tests with decay-time resolution parameters fixed in the fit to the data sample are performed. The parameters are fixed to values obtained from the time angle fit at $\phi_s$ value fixed to 0 or $\pi/2$, or to values from a sample of $J/\psi \rightarrow \mu^+\mu^-$ candidates produced at the PV corrected for the difference between $e^+e^-$ and $\mu^+\mu^-$ simulation samples. The test results are compatible within statistical uncertainties to the default fit results.

The decay-time efficiency introduces a systematic uncertainty from three different sources. First, the contribution due to the statistical uncertainty on the determination of the decay-time efficiency from the control channel is obtained by evaluating 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 $B^0 \rightarrow J/\psi K^{*}(892)^0$ control sample. Second, a sum of two Ipatia functions is used as an alternative mass model for the $m( e^+e^- K^+\pi^- )$ distribution and a new decay-time efficiency function is produced. Finally, the efficiency function is computed with the $B^0$ lifetime modified by $\pm 1\sigma$. In all cases the difference in the 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 for the signal channel where the difference of the number of signal candidates is within 10% between the default and varied BDT criteria. The effect of applying the modified BDT requirement in the likelihood fit is studied using pseudoexperiments. The mass model parameters for each BDT requirement are varied within their uncertainties (accounting for correlations) and the weights are re-evaluated based on the alternative model. The fit is repeated using a new set of weights and a new efficiency function. The observed variations in the physics parameters are compatible with statistical fluctuations. This is verified by pseudoexperiments with 10% of candidates removed at random.

A systematic uncertainty is assigned to account for the differences in the final-state kinematics between data and simulated samples. The simulated signal events are weighted using a multidimensional BDT-based algorithm\footnote{Reference number}. The procedure is repeated for the control sample $B^0 \rightarrow J/\psi(e^+e^-)K^{*}(892)^0$. The reweighted simulation samples of both channels are used to obtain new angular and decay-time acceptances. The difference with the default fit result is assigned as a systematic uncertainty.

The fraction of $A^0_b \rightarrow J/\psi pK^-$ candidates contributing to the signal sample is estimated to be 1% using simulation. The impact of neglecting this contribution is evaluated for the data sample by fitting the $m( e^+e^- K^+K^- )$ distribution with an additional component to account for, namely the sum of two CB functions, the shape of which is fixed to a fit to simulated $A^0_b \rightarrow J/\psi pK^-$ candidates. In addition, the decay-time efficiency is redetermined including a component for background from $A^0_b \rightarrow J/\psi p\pi^-$ decays. This component is modelled by the sum of two CB functions, the shape of which is fixed to a fit to simulated $A^0_b \rightarrow J/\psi p\pi^-$ candidates. The fraction of the $A^0_b \rightarrow J/\psi p\pi^-$ decays is estimated from the simulation to be at most 0.06%.\footnote{Reference number}. The differences of physics parameters obtained from the fit with modified weights and efficiency function is assigned as a systematic uncertainty.

A small fraction of $B^0 \rightarrow J/\psi \phi$ decays comes from the decays of $B^{+}_c$ mesons. The
fraction is estimated as 0.8% in Ref. [49] and pseudoexperiments are used to assess the impact of ignoring such a contribution on the extraction of the physics parameters. Only $\Gamma_s$ is observed to be affected, with a bias on its central value corresponding to 20% of the statistical uncertainty, which is assigned as a systematic uncertainty.

A possible bias in the fitting procedure is investigated through many pseudoexperiments of equivalent size to the data sample. For each pseudoexperiment the physics parameters are fluctuated in the underlying PDF and then compared to the obtained fit results. The resulting deviations are small and those that are not compatible with zero within three standard deviations are quoted as systematic uncertainties.

Inclusion of a result with a constraint on the $\Delta m_s$ into a global analysis leads to troublesome treatment of systematic effects introduced by choice of the constraint. Therefore we provide a result with the mixing frequency fixed to the PDG value, $\Delta m_s = 17.757 \text{ ps}^{-1}$ [13], as reported in Appendix B. No significant difference is observed with respect to the default result.

The systematic uncertainties associated to the mass model and mass factorisation can be treated as uncorrelated between this result and that of Ref. [16]. More details on the systematic effects for the studied channel are given in Ref. [50].

### 8 Conclusion

Using a data set corresponding to an integrated luminosity of 3 fb$^{-1}$ collected by the LHCb experiment in $pp$ collisions at centre-of-mass energies of 7 and 8 TeV, a flavour-tagged decay-time-dependent angular analysis of $(1.27 \pm 0.05) \times 10^4 B^0_s \rightarrow J/\psi(e^+e^-)\phi$ decays is performed. A number of physics parameters including the $CP$-violating phase $\phi_s$, average decay width $\Gamma_s$ and decay width difference $\Delta \Gamma_s$ as well as the polarisation amplitudes and strong phases of the decay are determined. The effective decay-time resolution and effective tagging power are 45.6 $\pm$ 0.1 fs and (5.07 $\pm$ 0.16)$\%$, respectively. The $CP$ parameters are measured to be

\[
\begin{align*}
\phi_s &= 0.00 \pm 0.28 \pm 0.05 \text{ rad}, \\
\Delta \Gamma_s &= 0.115 \pm 0.045 \pm 0.011 \text{ ps}^{-1}, \\
\Gamma_s &= 0.608 \pm 0.018 \pm 0.011 \text{ ps}^{-1}
\end{align*}
\]

where the first uncertainty is statistical and the second is systematic. The dominant sources of the systematic uncertainty are the imperfect mass and decay-time resolution models. This is the first measurement of the $CP$ content of the $B^0_s \rightarrow J/\psi(e^+e^-)\phi$ decay and first time that $\phi_s$ has been measured in the final state containing electrons. These results constitute an important check for the results with muons in the final state because the systematic uncertainties of the measurements are independent, while the studied mechanism of the $CP$ violation is the same. The results are consistent with previous measurements [3, 8–12], the SM predictions [22–24], and show no evidence of $CP$ violation in the interference between $B^0_s$ meson mixing and decay. In addition, no evidence for direct $CP$ violation in $B^0_s \rightarrow J/\psi(e^+e^-)\phi$ decays is observed.
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); MOST and NSFC (China); CNRS/IN2P3 (France); BMBF, DFG and MPG (Germany); INFN (Italy); NWO (Netherlands); MNiSW and NCN (Poland); MEN/IFA (Romania); MSHE (Russia); MICINN (Spain); SNSF and SER (Switzerland); NASU (Ukraine); STFC (United Kingdom); DOE NP and NSF (USA). We acknowledge the computing resources that are provided by CERN, IN2P3 (France), KIT and DESY (Germany), INFN (Italy), SURF (Netherlands), PIC (Spain), GridPP (United Kingdom), RRCKI and Yandex LLC (Russia), CSCS (Switzerland), IFIN-HH (Romania), CBPF (Brazil), PL-GRID (Poland) and NERSC (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 ARC and ARDC (Australia); AvH Foundation (Germany); EPLANET, Marie Skłodowska-Curie Actions and ERC (European Union); A*MIDEX, ANR, Labex P2IO and OCEVU, and Région Auvergne-Rhône-Alpes (France); Key Research Program of Frontier Sciences of CAS, CAS PIFI, CAS CCEPP, Fundamental Research Funds for the Central Universities, and Sci. & Tech. Program of Guangzhou (China); RFBR, RSF and Yandex LLC (Russia); GVA, XuntaGal and GENCAT (Spain); the Leverhulme Trust, the Royal Society and UKRI (United Kingdom).

Appendices

A Correlation matrix

The $CP$ observables are determined by an unbinned maximum-likelihood fit to the background-subtracted candidates with a probability density function (PDF) describing $B_s^0 \rightarrow J/\psi (e^+e^-)\phi$ signal decay. The correlation matrix of their statistical uncertainties is presented in Table 5. It is obtained using the bootstrap method.

B Fit results with fixed $\Delta m_s$

The fit is repeated with a fixed value of the mixing frequency $\Delta m_s = 17.757 \text{ps}^{-1}$ instead of a Gaussian constraint. The fit results are presented in Table 6 and corresponding correlation matrix in Table 7.

References

Table 6: Results of the maximum-likelihood fit described in Section 6 to the $B$ systematic, which is discussed in Section 7. the PDG value, $\Delta \frac{\text{decays}}{\text{including all acceptance and resolution effects}}$ and with the mixing frequency fixed to $\phi$.

Table 5: Correlation matrix of statistical uncertainties.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fit result and uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Gamma_s \ [\text{ps}^{-1}]$</td>
<td>$0.608 \pm 0.018 \pm 0.011$</td>
</tr>
<tr>
<td>$\Delta \Gamma_s \ [\text{ps}^{-1}]$</td>
<td>$0.115 \pm 0.043 \pm 0.011$</td>
</tr>
<tr>
<td>$</td>
<td>A_\perp</td>
</tr>
<tr>
<td>$</td>
<td>A_0</td>
</tr>
<tr>
<td>$\delta_\parallel \ [\text{rad}]$</td>
<td>$3.11^{+0.07}_{-0.08} \pm 0.06$</td>
</tr>
<tr>
<td>$\delta_\perp \ [\text{rad}]$</td>
<td>$2.41^{+0.46}_{-0.46} \pm 0.10$</td>
</tr>
<tr>
<td>$\phi_s \ [\text{rad}]$</td>
<td>$0.00 \pm 0.30 \pm 0.05$</td>
</tr>
<tr>
<td>$</td>
<td>\lambda</td>
</tr>
<tr>
<td>$F_S$</td>
<td>$0.062^{+0.045}_{-0.052} \pm 0.022$</td>
</tr>
<tr>
<td>$\delta_S \ [\text{rad}]$</td>
<td>$0.01 \pm 0.29 \pm 0.04$</td>
</tr>
</tbody>
</table>


Table 7: Correlation matrix of statistical uncertainties for a fit with fixed $\Delta m_s$.

|          | $\Gamma_s$ | $\Delta \Gamma_s$ | $|A_\parallel|^2$ | $|A_\perp|^2$ | $\delta_\parallel$ | $\delta_\perp$ | $\phi_s$ | $|\lambda|$ | $F_S$ | $\delta_S$ |
|----------|------------|-------------------|-------------------|--------------|-------------------|---------------|-----------|-------------|--------|-----------|
| $\Gamma_s$ | 1.00       | -0.25             | 0.37              | -0.32        | -0.01             | -0.04         | 0.01      | -0.04      | -0.05  | -0.03     |
| $\Delta \Gamma_s$ | 1.00       | -0.65             | 0.60              | 0.04         | -0.05             | 0.06          | -0.01     | -0.09      | 0.05   |           |
| $|A_\parallel|^2$ | 1.00       | -0.61             | -0.13             | 0.09         | -0.09             | -0.17         | -0.25     | 0.01       |        |           |
| $|A_\perp|^2$ | 1.00       | -0.05             | 0.14              | 0.07         | 0.17              | 0.31          | 0.0       |            |        |           |
| $\delta_\parallel$ | 1.00       | -0.51             | 0.10              | -0.01        | 0.17              | -0.22         |            |            |        |           |
| $\delta_\perp$ | 1.00       | -0.05             | 0.20              | -0.07        | -0.26             | -0.10         |            |            |        |           |
| $\phi_s$ | 1.00       | -0.05             | 0.20              | -0.07        | 0.20              | -0.07         |            |            |        |           |
| $|\lambda|$ | 1.00       | -0.05             | 0.51              | -0.03        |                   |               |            |            |        |           |
| $F_S$ |            |                   |                   |              |                   |               |            |            | 1.00   |           |
| $\delta_S$ |            |                   |                   |              |                   |               |            |            |        | 1.00      |


[17] LHCb collaboration, R. Aaij et al., Test of lepton universality with $B^0 \to K^{*0} \ell^+\ell^-$ decays, JHEP 08 (2017) 055, arXiv:1705.05802


[19] LHCb collaboration, R. Aaij et al., Test of lepton universality using $\Lambda^0_b \to pK^-\ell^+\ell^-$ decays, JHEP 05 (2020) 040, arXiv:1912.08139

[20] LHCb collaboration, R. Aaij et al., Measurement of CP violation and the $B^0_s$ meson decay width difference with $B^0_s \to J/\psi K^+K^-$ and $B^0_s \to J/\psi \pi^+\pi^-$ decays, Phys. Rev. D87 (2013) 112010, arXiv:1304.2600


LHCb collaboration

Università di Cagliari, Cagliari, Italy
Università di Ferrara, Ferrara, Italy
Università di Firenze, Firenze, Italy
Università di Genova, Genova, Italy
Università degli Studi di Milano, Milano, Italy
Università di Milano Bicocca, Milano, Italy
Università di Modena e Reggio Emilia, Modena, Italy
Università di Padova, Padova, Italy
Scuola Normale Superiore, Pisa, Italy
Università di Pisa, Pisa, Italy
Università della Basilicata, Potenza, Italy
Università di Roma Tor Vergata, Roma, Italy
Università di Siena, Siena, Italy
Università di Urbino, Urbino, Italy
MSU - Iligan Institute of Technology (MSU-IIT), Iligan, Philippines
AGH - University of Science and Technology, Faculty of Computer Science, Electronics and Telecommunications, Kraków, Poland
P.N. Lebedev Physical Institute, Russian Academy of Science (LPI RAS), Moscow, Russia
Novosibirsk State University, Novosibirsk, Russia
Hanoi University of Science, Hanoi, Vietnam