Measurement of the CP-violating phase $\phi_s$ in $B_0^s \rightarrow J/\psi \phi$ decays in ATLAS at 13 TeV

ATLAS Collaboration

DOI
10.1140/epjc/s10052-021-09011-0

Publication date
2021

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):
ATLAS Collaboration (2021). Measurement of the CP-violating phase $\phi_s$ in $B_0^s \rightarrow J/\psi \phi$ decays in ATLAS at 13 TeV. European Physical Journal C, 81(4), [342].
https://doi.org/10.1140/epjc/s10052-021-09011-0
Measurement of the CP-violating phase $\phi_s$ in $B^0_s \to J/\psi \phi$ decays in ATLAS at 13 TeV

ATLAS Collaboration*

CERN, 1211 Geneva 23, Switzerland

Received: 21 January 2020 / Accepted: 26 February 2021 / Published online: 21 April 2021 © CERN for the benefit of the ATLAS collaboration 2021

Abstract A measurement of the $B^0_s \to J/\psi \phi$ decay parameters using 80.5 fb$^{-1}$ of integrated luminosity collected with the ATLAS detector from 13 TeV proton–proton collisions at the LHC is presented. The measured parameters include the CP-violating phase $\phi_s$, the width difference $\Delta\Gamma_s$ between the $B^0_s$ meson mass eigenstates and the average decay width $\Gamma_s$. The values measured for the physical parameters are combined with those from 19.2 fb$^{-1}$ of 7 and 8 TeV data, leading to the following:

$\phi_s = -0.087 \pm 0.036$ (stat.) $\pm 0.021$ (syst.) rad
$\Delta\Gamma_s = 0.0657 \pm 0.0043$ (stat.) $\pm 0.0037$ (syst.) ps$^{-1}$
$\Gamma_s = 0.6703 \pm 0.0014$ (stat.) $\pm 0.0018$ (syst.) ps$^{-1}$

Results for $\phi_s$ and $\Delta\Gamma_s$ are also presented as 68% confidence level contours in the $\phi_s$-$\Delta\Gamma_s$ plane. Furthermore, the transversity amplitudes and corresponding strong phases are measured. $\phi_s$ and $\Delta\Gamma_s$ measurements are in agreement with the Standard Model predictions.

1 Introduction

In the presence of new physics (NP) phenomena, sources of CP violation in $b$-hadron decays can arise in addition to those predicted by the Standard Model (SM) [1]. In the $B^0_s \to J/\psi \phi$ decay, CP violation occurs due to interference between a direct decay and a decay with $B^0_s$-$\bar{B}^0_s$ mixing. The oscillation frequency of $B^0_s$ meson mixing is characterised by the mass difference, $\Delta m_s$, of the heavy ($B_H$) and light ($B_L$) mass eigenstates. The CP-violating phase $\phi_s$ is defined as the weak phase difference between the $B^0_s$-$\bar{B}^0_s$ mixing amplitude and the $b \to c\bar{s}\phi$ decay amplitude. In the SM the phase $\phi_s$ is small and is related to the Cabibbo–Kobayashi–Maskawa (CKM) quark mixing matrix elements via the relation $\phi_s \simeq -2\beta_s$, with $\beta_s = \arg[-(V_{ts}^* V_{tb})/(V_{cd} V_{cb}^*)]$. By combining beauty and kaon physics observables, and assuming no NP contributions to $B^0_s$ mixing and decays, a value of $-2\beta_s = -0.03696 \pm 0.00082$ rad was predicted by the CKMFitter group [2] and $-2\beta_s = -0.03700 \pm 0.00104$ rad according to the UTfit Collaboration [3]. While large NP enhancements of the mixing amplitude have been excluded by the precise measurement of the oscillation frequency [4], the NP couplings involved in the mixing may still increase the size of the observed CP violation by enhancing the mixing phase $\phi_s$ with respect to the SM value.

Other physical quantities involved in $B^0_s$-$\bar{B}^0_s$ mixing are the decay width $\Gamma_s$ and the width difference $\Delta\Gamma_s = \Gamma_L - \Gamma_H$, where $\Gamma_L$ and $\Gamma_H$ are the decay widths of the light and heavy mass eigenstates, respectively. The latest predictions for the width difference in the SM are $\Delta\Gamma_s = 0.091 \pm 0.013$ ps$^{-1}$ [5] and $\Delta\Gamma_s = 0.092 \pm 0.014$ ps$^{-1}$ [6]. A potential NP enhancement of $\phi_s$ would also decrease the size of $\Delta\Gamma_s$, but it is not expected to be affected as significantly as $\phi_s$ [7]. Nevertheless, extracting $\Delta\Gamma_s$ from the data is an important test of theoretical predictions [7].

Theory predictions have been made for the lifetime ratios $\tau(B^0_s)/\tau(B_d)$ and $\tau(B^0_s)/\tau(B^+)$, with the latest update Ref. [8]. The lifetime $\tau(B^0_s)$ has not been calculated in theory yet at a precision comparable with those obtained by experiments. The current world combined value of the decay width, $\Gamma_s$, obtained from experimental results is $\Gamma_s = 0.6600 \pm 0.0016$ ps$^{-1}$ [9].

The analysis of the time evolution of the $B^0_s \to J/\psi \phi$ decay provides the most precise determination of $\phi_s$ and $\Delta\Gamma_s$. Previous measurements of these quantities have been reported by the D0, CDF, LHCb, ATLAS and CMS Collaborations [10–17]. Additional improvements in measuring $\phi_s$ from $B^0_s \to \psi(2S)\phi$, $B^0_s \to D^+_s D^-_s$ and $B^0_s \to J/\psi \pi^+\pi^-$ decays have been achieved by the LHCb Collaboration [18–21].

The analysis presented here introduces a measurement of the $B^0_s \to J/\psi \phi$ decay parameters using 80.5 fb$^{-1}$ of the LHC proton–proton (pp) data collected by the ATLAS detector during 2015–2017, at a centre-of-mass energy, $\sqrt{s}$, equal to 13 TeV. The analysis closely follows a previous ATLAS measurement [13] that was performed using 19.2 fb$^{-1}$ of the
data collected at 7 and 8 TeV, and introduces more precise signal and background models.

2 ATLAS detector and Monte Carlo simulation

The ATLAS detector\(^1\) consists of three main components: an inner detector (ID) tracking system immersed in a 2 T axial magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer (MS). The inner tracking detector covers the pseudorapidity range \(|\eta| < 2.5\), and consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. The ID is surrounded by a high-granularity liquid-argon (LAr) sampling electromagnetic calorimeter. A steel/scintillator tile calorimeter provides hadronic coverage in the central rapidity range. The endcap and forward regions are equipped with LAr calorimeters for electromagnetic and hadronic measurements. The MS surrounds the calorimeters and provides a system of tracking chambers and detectors for triggering. A full description can be found in Refs. [22–24].

The data were collected during periods with different instantaneous luminosity, so several triggers were used in the analysis. All triggers were based on the identification of a \(J/\psi \rightarrow \mu^+\mu^-\) decay, with transverse momentum (\(p_T\)) thresholds of either 4 GeV or 6 GeV for the muons. Data quality requirements are imposed on the data, notably on the performance of the MS, ID and calorimeter systems. The measurement uses 80.5 fb\(^{-1}\) of \(pp\) collision data. The uncertainty in the combined 2015–2017 integrated luminosity is 2.0% [25], obtained using the LUCID-2 detector [26] for the primary luminosity measurements.

To study the detector response, estimate backgrounds, and model systematic effects, 100 M Monte Carlo (MC) simulated \(B^0 \rightarrow J/\psi \phi\) events were generated using PYTHIA 8.210 [27] tuned with ATLAS data, using the A14 set of parameter values [28] together with the CTEQ6L1 set of parton distribution functions [29]. The detector response was simulated using the ATLAS simulation framework based on GEANT4 [30,31]. In order to account for the varying number of proton–proton interactions per bunch crossing (pile-up) and trigger configurations during data-taking, the MC events were weighted to reproduce the same pile-up and trigger conditions as in the data. Additionally, background samples of both the exclusive (\(B^0 \rightarrow J/\psi K^0_s\) and \(\Lambda_b \rightarrow J/\psi p K^-\)) and inclusive \((b\bar{b} \rightarrow J/\psi X\) and \(p p \rightarrow J/\psi X\)) decays were simulated, using the same simulation tools as in the case of the signal events. For validation studies related to flavour tagging, detailed in Sect. 4, events with \(B^\pm \rightarrow J/\psi K^\pm\) exclusive decays were also simulated.

3 Reconstruction and candidate selection

The reconstruction and candidate selection for the decay \(B^0 \rightarrow J/\psi (\mu^+\mu^-)\phi (K^+K^-)\) is described here. Events must pass the trigger selections described in Sect. 2. In addition, each event must contain at least one reconstructed primary vertex, formed from at least four ID tracks, and at least one pair of oppositely charged muon candidates that are reconstructed using information from the MS and the ID. The muons used in the analysis are required to meet the Tight\(^2\) or Low-\(p_T\)\(^3\) working point identification criteria. The muon track parameters are determined from the ID measurement alone, since the precision of the measured track parameters is dominated by the ID track reconstruction in the \(p_T\) range of interest for this analysis. Pairs of oppositely charged muon tracks are re-fitted to a common vertex and the pair is accepted if the quality of the fit meets the requirement \(\chi^2/\text{ndof} < 10\). In order to account for varying mass resolution in different parts of the detector, the \(J/\psi\) candidates are divided into three subsets according to the pseudorapidity \(\eta\) of the muons. In the first subset, both muons have \(|\eta| < 1.05\), where the values \(\eta = \pm 1.05\) correspond to the edges of the barrel part of the MS. In the second subset, one muon has \(1.05 < |\eta| < 2.5\) and the other muon \(|\eta| < 1.05\). The third subset contains candidates where both muons have \(1.05 < |\eta| < 2.5\). A maximum likelihood fit is used to extract the \(J/\psi\) mass and the corresponding mass resolution for these three subsets, and in each case the signal region is defined symmetrically around the fitted mass, so as to retain 99.7% of the \(J/\psi\) candidates identified in the fits.

The candidates for the decay \(\phi \rightarrow K^+K^-\) are reconstructed from all pairs of oppositely charged tracks, with \(p_T > 1\) GeV and \(|\eta| < 2.5\), that are not identified as muons. Candidate events for \(B_s^0 \rightarrow J/\psi (\mu^+\mu^-)\phi (K^+K^-)\) decays are selected by fitting the tracks for each combination of \(J/\psi \rightarrow \mu^+\mu^-\) and \(\phi \rightarrow K^+K^-\) to a common vertex. The fit is also constrained by fixing the invariant mass of the hadronic jets.

---

\(^1\) ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point. The \(z\)-axis is along the beam pipe, the \(x\)-axis points to the centre of the LHC ring and the \(y\)-axis points upward. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \(r\) being the distance from the origin and \(\phi\) being the azimuthal angle around the beam pipe. The pseudorapidity \(\eta\) is defined as \(\eta = -\ln[\tan(\theta/2)]\) where \(\theta\) is the polar angle.

\(^2\) Tight muon reconstruction is optimised to maximise the purity of muons at the cost of some efficiency, requiring combined muons with hits in at least two stations of the MS and additional criteria, described in Ref. [32].

\(^3\) This working point is optimised to provide good muon reconstruction efficiency down to a \(p_T\) of \(\approx 3\) GeV, while controlling the fake-muon rate. It allows \(1 \gtrsim 1.2\) (\(\gtrsim 1.3\) \(\gtrsim 1.55\)) for candidates reconstructed by algorithms utilising inside-out combined reconstruction [32]. Additional cuts on the number of precision stations and on variables very sensitive to the decays in flight of hadrons are also applied to suppress fake muons.
culated from the two muon tracks to the $J/\psi$ mass [33]. A quadruplet of tracks is accepted for further analysis if the vertex fit has $\chi^2$/ndof $< 3$. For the $\phi \rightarrow K^+K^-$ candidate, the invariant mass of the track pairs (using a charged kaon mass hypothesis) must fall within the interval $1.0085 \text{ GeV} < m(K^+K^-) < 1.0305 \text{ GeV}$. The interval, chosen using MC simulation, is selected to retain 98% of true $\phi \rightarrow K^+K^-$ decays. The $B^0_s$ candidate with the lowest $\chi^2$/ndof is selected in events where more than one candidate passes all selections. In total, 2,977,526 $B^0_s$ candidates are collected within the mass range of 5.150–5.650 GeV. This range is chosen to give enough background events in the sidebands of the mass distributions to allow precise determination of the properties of the background events.

The mean number of interactions per bunch crossing is 30, necessitating a choice of the best candidate for the primary vertex at which the $B^0_s$ meson is produced. Primary vertex positions are recalculated after removing any tracks used in the $B^0_s$ meson reconstruction. The variable used to select the best candidate for the primary vertex is the three-dimensional impact parameter, $\alpha_0$, which is calculated as the minimum distance between each primary vertex candidate and the line extrapolated from the reconstructed $B^0_s$ meson vertex in the direction of the $B^0_s$ momentum. The chosen primary vertex is the one with the smallest $\alpha_0$. A simulated dataset is used to estimate the fraction of $B^0_s$ candidates where the incorrect production vertex is selected (12%) and demonstrates that the mis-selection of reconstructed primary vertex does not bias the reconstructed proper decay time.

For each $B^0_s$ meson candidate the proper decay time $t$ is estimated using:

$$ t = \frac{L_{xy} m_{B}}{p_{T_B}}, $$

where $p_{T_B}$ is the reconstructed transverse momentum of the $B^0_s$ meson candidate and $m_{B}$ denotes the mass of the $B^0_s$ meson, taken from Ref. [33]. The transverse decay length, $L_{xy}$, is the displacement in the transverse plane of the $B^0_s$ meson decay vertex relative to the primary vertex, projected onto the direction of the $B^0_s$ transverse momentum.

4 Flavour tagging

To identify, or tag, the flavour of a neutral $B$ meson at the point of production, information is extracted using the decay of the other (or opposite) $b$-hadron that is produced from the pair production of $b$ and $\bar{b}$ quarks. This method is called opposite-side tagging (OST).

The OST algorithms each define a discriminating variable, based on charge information, which is sensitive to the flavour (i.e. $b$- or $b$-quark) of the opposite-side $b$-hadron. The algorithms thus provide a probability that a signal $B$ meson in a given event is produced in a given flavour. The calibration of the OST algorithms proceeds using data containing $B^{\pm} \rightarrow J/\psi K^{\pm}$ candidate decays, where the charge of the kaon determines the flavour of the $B$ meson, providing a self-tagging sample of events. These OST algorithms are calibrated as a function of the discriminating variable, using yields of signal $B^{\pm}$ mesons extracted from fits to the data. Once calibrated, the OST algorithms are applied to $B^0_s \rightarrow J/\psi (\mu^+\mu^-) \phi (K^+K^-)$ candidate events to provide a probability that each candidate was produced in a $B^0_s$ or $\bar{B}^0_s$ meson state, which is used in the maximum likelihood fit (described in Sect. 5). This approach assumes invariance of the OST algorithm with respect to the specific signal $b$-hadron type (i.e $B^{\pm}$ meson or $B^0_s$ meson), which is tested and the difference is considered as a systematic uncertainty.

4.1 $B^{\pm} \rightarrow J/\psi K^{\pm}$ event selection

Candidate $B^{\pm} \rightarrow J/\psi K^{\pm}$ decays are identified in a series of steps. First, $J/\psi$ candidates are selected from oppositely charged muon pairs forming a good vertex, as described in Sect. 3. Each muon is required to have $p_T > 4 \text{ GeV}$ and $|\eta| < 2.5$. Dimuon candidates with invariant mass $2.8 < m(\mu^+\mu^-) < 3.4 \text{ GeV}$, as determined from the re-fitted track parameters of the vertex, are retained for further analysis. To form the $B^{\pm}$ candidate, an additional track is required, which is not identified as an electron or muon. The track is assigned the charged-kaon mass hypothesis and combined with the dimuon candidate using a vertex fit, performed with the mass of the dimuon pair constrained to the $J/\psi$ mass. Prompt background contributions are suppressed by a requirement on the proper decay time of the $B^{\pm}$ candidate of $t > 0.2 \text{ ps}$.

The tagging probabilities are determined from $B^+$ and $B^-$ signal events. These signal yields are derived from fits to the invariant mass distribution, $m(J/\psi K^{\pm})$, and performed in intervals of the discriminating variables. To describe the $B^{\pm} \rightarrow J/\psi K^{\pm}$ signal, two Gaussian functions with a common mean are used. An exponential function is used to describe the combinatorial background and a hyperbolic tangent function to parameterise the low-mass contribution from incorrectly or partially reconstructed $b$-hadron decays. A Gaussian function is used to describe the $B^{\pm} \rightarrow J/\psi \pi^{\pm}$ contribution, with fixed parameters taken from simulation except for the normalisation, which is a free parameter. A fit to the overall mass distribution is used to define the shapes of signal and backgrounds. Subsequent fits are performed in the intervals of the tagging discriminating variables, separately for $B^+$ and $B^-$ candidate events, with the normalisations and also the slope of the exponential function left free. The $B^+$ and $B^-$ signal yields are extracted from these fits. Figure 1 shows the invariant mass distribution of $B^{\pm}$ candidates over-
such as through neutral processes that can change the charge of the observed lepton, cascade decays of a lepton (electron or muon) from the semileptonic decay of a meson provides strong discrimination; however, the ATLAS sensitivity to the flavour of the signal meson at the point of production is inferred using several methods, which differ in their selection criteria in one event, Tight muons are chosen according to Eq. (1), with $\kappa = 1.1$ and the sum over the reconstructed ID tracks within a cone of size $\Delta R = 0.5$ around the muon direction. These tracks must have $p_T > 0.5$ GeV, $|\eta| < 2.5$, and $|\Delta z| < 5$ mm. Tracks associated with the decay of a $B$ meson signal candidate are excluded from the sum. In each interval of $Q_\mu$, a fit to the $J/\psi K^\pm$ invariant mass spectrum is performed and the number of signal events extracted. The fit model used is described in Sect. 4.1. Figure 2 shows the distributions of the muon cone charge using $B^\pm$ signal candidates for Tight muons, and includes the tagging power (defined in Sect. 4.3).

\[ Q_\mu = \frac{\sum_i^{N_{\text{tracks}}} q_i \cdot (p_{T_i})^x}{\sum_i^{N_{\text{tracks}}} (p_{T_i})^x}, \]
gging probability as a function of the cone charge variable. The corresponding distributions for Low-$p_T$ muons are shown in Fig. 3.

**Electron tagging**

Electrons are identified using ID and calorimeter information, and must satisfy the Medium electron quality criteria [34]. The ID track associated with the electron is required to have $p_T > 0.5$ GeV, $|\eta| < 2.5$, and $|\Delta z| < 5$ mm. To reject electrons from the signal-side of the decay, electrons with $\cos(\zeta_b) > 0.93$, where $\zeta_b$ is the opening angle between the momentum of the signal $B$ meson candidate and the electron momentum, are not considered. In the case of more than one electron passing the selection, the electron with the highest $p_T$ is chosen. Charged particle tracks within a cone of size $\Delta R = 0.5$ are used to form the electron cone charge $Q_e$, constructed according to Eq. (1), with $\kappa = 1.0$. The resulting electron cone charge distributions are shown in Fig. 4, together with the corresponding tagging probability.

**Jet tagging**

In the absence of a muon or electron, a jet identified as containing a $b$-hadron is required. Jets are reconstructed from calorimetric information [35] using the anti-$k_t$ algorithm [36,37] with a radius parameter $R = 0.4$. The identification of a $b$-tagged jet uses a multivariate algorithm MV2c10 [38], utilising boosted decision trees (BDT), which output a classifier value. Jets are selected if this value exceeds 0.56. This value is chosen to maximise the tagging power of the calibration sample. In the case of multiple selected jets, the jet with the highest value of the BDT output classifier is used. Jets associated with the signal decay are not considered in this selection.

Tracks within a cone of size $\Delta R = 0.5$ around the jet axis are used to define a jet cone charge, $Q_{\text{jett}}$, constructed according to Eq. (1), where $\kappa = 1.1$ and the sum is over the tracks associated with the jet, with $|\Delta z| < 5$ mm, and excluding tracks from the decay of the signal $B$ meson candidate. Figure 5 shows the distribution of the opposite-side jet cone charge for $B^{\pm}$ signal candidates.

4.3 Flavour tagging performance

In order to quantify and compare the performance of the various tagging methods, three figure-of-merit terms are constructed, which describe: the fraction of events used by a given tagging method, the purity of the method, and the overall power of the tagging method in the sample. The efficiency, $\epsilon_x$, of an individual tagging method is defined as the number of signal events tagged by that method divided by the total number of signal events in the sample. The purity of a particular flavour tagging method, called the dilution, is defined as $D(Q_x) = 2P(B|Q_x) - 1$. The tagging power of a particular tagging method is then defined as $T_x = \sum_i \epsilon_{x_i} \cdot D^2(Q_{x_i})$, where the sum is over the probability distribution in intervals of the cone charge variable. An effective dilution, $D_x = \sqrt{T_x/\epsilon_x}$, is calculated from the measured tagging power and efficiency.

By definition, there is no overlap between lepton-tagged and jet-charge-tagged events. The overlap between events with a muon (either Tight or Low-$p_T$) and events with an electron corresponds to around 0.6% of all tagged events. In the case of multiply tagged events, the OST method is selected in the following order: Tight muon, electron, Low-$p_T$ muon, jet. However, the ordering of muon- and electron-tagged events is shown to have negligible impact on the final results. A summary of the tagging performance for each method and the overall performance on the $B^{\pm}$ sample is given in Table 1.

4.4 Using tag information in the $B^0$ fit

For the maximum likelihood fit performed on the $B^0$ data, and described in detail in Sect. 5, the per-candidate probability, $P(B|Q_x)$, that the $B$ meson candidate was produced in a state $B^0$ (versus a $\bar{B}^0$) is provided by the calibrations derived from the $B^\pm \to J/\psi K^\pm$ sample, described above, and shown in Figs. 2, 3, 4 and 5. Since the distributions of $P(B|Q_x)$ from signal $B^0$ mesons and backgrounds can be expected to be different, separate probability density functions (PDFs) are necessary to describe these distributions in the likelihood function. These PDFs are defined as $P_a(P(B|Q_x))$ and $P_b(P(B|Q_x))$, describing the probability distributions for signal and background, respectively, and are derived from the sample of $B^0$ candidates. For the exclusive decays $B_d \to J/\psi K^0$ and $\Lambda_b \to J/\psi pK^-$ that are present in the sample of $B^0$ candidates, $P_a(P(B|Q_x))$ is used to model the probability distributions for these contributions (described further in Sect. 5.2). The PDFs consist of the fraction of events that are tagged with a particular method (or are untagged), the fractions of those events categorised as discrete or continuous, and for those that are continuous, a PDF of the corresponding probability distribution.

**Continuous PDF**

The parameterisations of the continuous PDF components of $P_{ab}(P(B|Q_x))$ for each OST method are defined as follows. In the sideband regions, $5.150 < m(J/\psi KK) < 5.317$ GeV and $5.417 < m(J/\psi KK) < 5.650$ GeV, unbinned maximum likelihood fits to the $P(B|Q_x)$ distributions are performed to extract the background (continuous category) PDFs for $P_a(P(B|Q_x))$. For the Tight muon and electron methods, the parameterisation has the form of the sum of a second-order polynomial and two exponential functions. A
Fig. 2 Cone charge distributions, $-Q_{\mu}$, for Tight muons, shown for cases of discrete charge (left), and for the continuous distribution (right). For each plot, in red (blue), the normalised $B^+(B^-)$ cone charge distribution is shown (corresponding to the right axis scale). A $B^+$ ($B^-$) candidate is more likely to have a large negative (positive) value of $Q_{\mu}$. Superimposed is the distribution of the tagging probability, $P(B|Q_{\mu})$, as a function of the cone charge, derived from a data sample of $B^{\pm}\rightarrow J/\psi K^\pm$ decays, and defined as the probability to have a $B^+$ meson (on the signal-side) given a particular cone charge $Q_{\mu}$. The fitted parameterisation, shown in black, is used as the calibration curve to infer the probability to have a $B^+_s$ or $B^-_s$ meson at production in the decays to $J/\psi \phi$.

Fig. 3 Normalised cone charge distributions (shown against the right axis scale), $-Q_{\mu}$, for $B^+(B^-)$ events shown in red (blue) for Low-$p_T$ muons, for cases of discrete charge (left), and for the continuous distribution (right). Superimposed is the distribution of the tagging probability, $P(B|Q_{\mu})$.

Fig. 4 Normalised cone charge distributions (shown against the right axis scale), $-Q_{e}$, for $B^+(B^-)$ events shown in red (blue) for electrons, for cases of discrete charge (left), and the continuous distribution (right). Superimposed is the distribution of the tagging probabilities, $P(B|Q_{e})$. 
Jets and two exponential functions, and for the Low-

tion has the form of the sum of a second-order polynomial

of events, 1 - f_{+1} - f_{-1}, corresponds to the continuous parts of the distribution. Positive and negative charges are equally probable for background candidates formed from a random combination of a $J/\psi$ and a pair of tracks, but this is not necessarily the case for background candidates formed from a partially reconstructed $b$-hadron. Table 2 summarises the fractions $f_{+1}$ and $f_{-1}$ obtained from each tagging method for signal and background events.

The fractions of signal and background events tagged using the different OST methods are found using a similar sideband-subtraction method, and are summarised in Table 3.

To account for possible deviations of the data from the selected fit models, variations of the procedure described here are used to determine systematic uncertainties, as described in Sect. 6.

5 Maximum likelihood fit

An unbinned maximum likelihood fit is performed on the selected events to extract the parameter values of the $B^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ decay. The fit uses information about the reconstructed mass, $m$, the measured proper decay time, $t$, the measured mass uncertainty, $\sigma_m$, the measured proper decay time uncertainty, $\sigma_t$, the measured transverse momentum, $p_T$, the tagging probability, $P(B|Q_{jet})$, and the transversity angles, $\Omega$, of each $B^0 \to J/\psi$ decay candidate. The measured value of the proper decay time uncertainty, $\sigma_t$, is calculated from the covariance matrix associated with the vertex fit for each candidate event. The transversity angles $\Omega = (\theta_T, \psi_T, \phi_T)$ are defined in Sect. 5.1. The like-

Higer-order polynomial is used.

For the signal, fits are performed to the $P(B|Q_{jet})$ distributions, using all events in the $m(J/\psi K K)$ distributions to extract the signal (continuous category) PDFs for $P_s(P(B|Q_{jet}))$. In these fits, the parameters describing the background PDFs are fixed to their previously extracted values, as is the relative normalisation of signal and background, extracted from a fit to the $m(J/\psi K K)$ distribution. For the signal PDFs, the Tight muon tagging method uses the sum of two exponential functions and a constant function to describe the signal. For the electron tagging method, the signal function has the form of the sum of a second-order polynomial and two exponential functions, and for the Low-$p_T$ muon and jet tagging methods a Gaussian function is used.

Discrete PDF

In the case where the cone charge is discrete, the fractions of events $f_{+1}$ ($f_{-1}$) with cone charges +1 (-1) are deter-

mined separately for signal and background using events from the signal and sideband regions of the $B^0$ mass distribution (as defined in Sect. 3). The remaining fraction of events, $1 - f_{+1} - f_{-1}$, corresponds to the continuous parts of the distribution. Positive and negative charges are equally probable for background candidates formed from a random combination of a $J/\psi$ and a pair of tracks, but this is not necessarily the case for background candidates formed from a partially reconstructed $b$-hadron. Table 2 summarises the fractions $f_{+1}$ and $f_{-1}$ obtained from each tagging method for signal and background events.

The fractions of signal and background events tagged using the different OST methods are found using a similar sideband-subtraction method, and are summarised in Table 3.

To account for possible deviations of the data from the selected fit models, variations of the procedure described here are used to determine systematic uncertainties, as described in Sect. 6.

Table 1 Summary of tagging performances for the different flavour tagging methods on the sample of $B^\pm$ signal candidates, as described in the text. Uncertainties shown are statistical only. The efficiency ($\epsilon_x$) and tagging power ($D_x$) are each determined by summing over the individual bins of the cone charge distribution. The effective dilution ($D_x$) is obtained from the measured efficiency and tagging power. For the efficiency, effective dilution, and tagging power, the corresponding uncertainty is determined by combining the appropriate uncertainties in the individual bins of each charge distribution.

<table>
<thead>
<tr>
<th>Tag method</th>
<th>$\epsilon_x$ (%)</th>
<th>$D_x$ (%)</th>
<th>$T_x$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tight muon</td>
<td>4.50 ± 0.01</td>
<td>43.8 ± 0.2</td>
<td>0.862 ± 0.009</td>
</tr>
<tr>
<td>Electron</td>
<td>1.57 ± 0.01</td>
<td>41.8 ± 0.2</td>
<td>0.274 ± 0.004</td>
</tr>
<tr>
<td>Low-$p_T$ muon</td>
<td>3.12 ± 0.01</td>
<td>29.9 ± 0.2</td>
<td>0.278 ± 0.006</td>
</tr>
<tr>
<td>Jet</td>
<td>12.04 ± 0.02</td>
<td>16.6 ± 0.1</td>
<td>0.334 ± 0.006</td>
</tr>
<tr>
<td>Total</td>
<td>21.23 ± 0.03</td>
<td>28.7 ± 0.1</td>
<td>1.75 ± 0.01</td>
</tr>
</tbody>
</table>

Gaussian function is used for the Low-$p_T$ muons. For the jet tagging algorithm an eighth-order polynomial is used.

For the signal, fits are performed to the $P(B|Q_{jet})$ distributions, using all events in the $m(J/\psi K K)$ distributions to extract the signal (continuous category) PDFs for $P_s(P(B|Q_{jet}))$. In these fits, the parameters describing the background PDFs are fixed to their previously extracted values, as is the relative normalisation of signal and background, extracted from a fit to the $m(J/\psi K K)$ distribution. For the signal PDFs, the Tight muon tagging method uses the sum of two exponential functions and a constant function to describe the signal. For the electron tagging method, the signal function has the form of the sum of a second-order polynomial and two exponential functions, and for the Low-$p_T$ muon and jet tagging methods a Gaussian function is used.

Discrete PDF

In the case where the cone charge is discrete, the fractions of events $f_{+1}$ ($f_{-1}$) with cone charges +1 (-1) are deter-

mined separately for signal and background using events from the signal and sideband regions of the $B^0$ mass distribution (as defined in Sect. 3). The remaining fraction of events, $1 - f_{+1} - f_{-1}$, corresponds to the continuous parts of the distribution. Positive and negative charges are equally probable for background candidates formed from a random combination of a $J/\psi$ and a pair of tracks, but this is not necessarily the case for background candidates formed from a partially reconstructed $b$-hadron. Table 2 summarises the fractions $f_{+1}$ and $f_{-1}$ obtained from each tagging method for signal and background events.

The fractions of signal and background events tagged using the different OST methods are found using a similar sideband-subtraction method, and are summarised in Table 3.

To account for possible deviations of the data from the selected fit models, variations of the procedure described here are used to determine systematic uncertainties, as described in Sect. 6.

5 Maximum likelihood fit

An unbinned maximum likelihood fit is performed on the selected events to extract the parameter values of the $B^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ decay. The fit uses information about the reconstructed mass, $m$, the measured proper decay time, $t$, the measured mass uncertainty, $\sigma_m$, the measured proper decay time uncertainty, $\sigma_t$, the measured transverse momentum, $p_T$, the tagging probability, $P(B|Q_{jet})$, and the transversity angles, $\Omega$, of each $B^0 \to J/\psi$ decay candidate. The measured value of the proper decay time uncertainty, $\sigma_t$, is calculated from the covariance matrix associated with the vertex fit for each candidate event. The transversity angles $\Omega = (\theta_T, \psi_T, \phi_T)$ are defined in Sect. 5.1. The like-
Table 2 Fractions $f_{+1}$ and $f_{-1}$ of events with cone charges of $+1$ and $-1$, respectively, for signal and background events and for the different tagging methods. Only statistical uncertainties are given.

<table>
<thead>
<tr>
<th>Tag method</th>
<th>Signal $f_{+1}$ (%)</th>
<th>Signal $f_{-1}$ (%)</th>
<th>Background $f_{+1}$ (%)</th>
<th>Background $f_{-1}$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tight muon</td>
<td>6.9 ± 0.3</td>
<td>7.5 ± 0.3</td>
<td>4.7 ± 0.1</td>
<td>4.9 ± 0.1</td>
</tr>
<tr>
<td>Electron</td>
<td>20 ± 1</td>
<td>19 ± 1</td>
<td>16.8 ± 0.2</td>
<td>17.3 ± 0.2</td>
</tr>
<tr>
<td>Low-$p_T$ muon</td>
<td>10.9 ± 0.5</td>
<td>11.6 ± 0.5</td>
<td>7.0 ± 0.1</td>
<td>7.5 ± 0.1</td>
</tr>
<tr>
<td>Jet</td>
<td>3.60 ± 0.15</td>
<td>3.54 ± 0.15</td>
<td>3.05 ± 0.03</td>
<td>3.17 ± 0.03</td>
</tr>
</tbody>
</table>

Table 3 Fractions of signal and background events tagged using the different methods. The efficiencies include both the continuous and discrete contributions. Only statistical uncertainties are quoted.

<table>
<thead>
<tr>
<th>Tag method</th>
<th>Signal efficiency (%)</th>
<th>Background efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tight muon</td>
<td>4.06 ± 0.06</td>
<td>3.21 ± 0.01</td>
</tr>
<tr>
<td>Electron</td>
<td>1.86 ± 0.04</td>
<td>1.48 ± 0.01</td>
</tr>
<tr>
<td>Low-$p_T$ muon</td>
<td>2.95 ± 0.05</td>
<td>2.70 ± 0.01</td>
</tr>
<tr>
<td>Jet</td>
<td>12.1 ± 0.1</td>
<td>9.41 ± 0.02</td>
</tr>
<tr>
<td>Untagged</td>
<td>79.1 ± 0.3</td>
<td>83.20 ± 0.05</td>
</tr>
</tbody>
</table>

The likelihood function is defined as a combination of the signal and background PDFs as follows:

$$
\ln L = \sum_{i=1}^{N} w_i \cdot \ln[f_{s} \cdot F_{s}(m_i, \theta_i, \sigma_{m_i}, \Omega_i, P_i(B|Q_1), p_{rT}) + f_{b0} \cdot F_{b0}(m_i, \theta_i, \sigma_{m_i}, \Omega_i, P_i(B|Q_1), p_{rT}) + f_{Lb} \cdot F_{Lb}(m_i, \theta_i, \sigma_{m_i}, \Omega_i, P_i(B|Q_1), p_{rT}) + (1 - f_{s} - (1 + f_{b0} + f_{Lb}) F_{bkg}(m_i, \theta_i, \sigma_{m_i}, \Omega_i, P_i(B|Q_1), p_{rT})),]
$$  \tag{2}

where $N$ is the number of selected candidates, $w_i$ is a weighting factor to account for the trigger efficiency (described in Sect. 5.3). The terms $F_{s}$, $F_{b0}$, $F_{Lb}$ and $F_{bkg}$ are the PDFs modelling the signal, $B^0$ background, $\Lambda_b$ background, and the other background distributions, respectively. The term $f_s$ is the fraction of signal candidates and $f_{b0}$ and $f_{Lb}$ are the background fractions of $B^0$ mesons and $\Lambda_b$ baryons misidentified as $B^0$ candidates, calculated relative to the number of signal events. These background fractions are fixed to their expectation from the MC simulation, and variations are applied as part of the evaluation of the effects of systematic uncertainties. The mass $m_i$, the proper decay time $\theta_i$ and the decay angles $\Omega_i$ are the values measured from the data for each event $i$. A detailed description of the signal PDF terms in Eq. (2) is given in Sect. 5.1. The three background functions are described in Sect. 5.2.

5.1 Signal PDF

The PDF used to describe the signal events, $F_{s}$, has the following composition:

$$
F_{s}(m_i, \theta_i, \sigma_{m_i}, \sigma_{\theta_i}, \Omega_i, P_i(B|Q_1), p_{rT}) = P_s(m_i|\sigma_{m_i}) \cdot P_s(\sigma_{m_i}|p_{rT}) \cdot P_s(\sigma_{\theta_i}|p_{rT}) \cdot P_s(\Omega_i|p_{rT}) \cdot P_s(p_{rT}).
$$

The mass term $P_s(m_i|\sigma_{m_i})$ is modelled in the following way:

$$
P_s(m_i|\sigma_{m_i}) \equiv \frac{1}{\sqrt{2\pi \sigma_{m_i}}} \cdot e^{-\frac{(m_i - m_{B_i})^2}{2\sigma_{m_i}^2}}. \tag{3}
$$

The term $P_s(m_i|\sigma_{m_i})$ uses per-candidate mass errors, $\sigma_{m_i}$, calculated for each $J/\psi \phi$ candidate from the covariance matrix associated with the four-track vertex fit. Each measured candidate mass is convolved with a Gaussian function with a width equal to $\sigma_{m_i}$ multiplied by a scale factor $S_m$, introduced to account for any mismeasurements. Both $S_m$ and the mean value $m_{B_i}$, which is the $B^0$ meson mass, are free parameters determined in the fit.

The PDF term $P_s(\sigma_{\theta_i}|p_{rT})$ takes into account the lifetime resolution, so each time element in Table 4 is convolved with a Gaussian function defined as:

$$
R(\tau' - \tau, \sigma_{\tau}) \equiv \frac{1}{\sqrt{2\pi} S_{\tau}} \cdot e^{-\frac{(\tau' - \tau)^2}{2S_{\tau}^2 \sigma_{\tau}^2}}. \tag{4}
$$

$S_{\tau}$ is a scale factor (a parameter of the fit) and $\sigma_{\tau}$ is the per-candidate uncertainty on proper decay time $\tau$. This convolution is performed numerically on an event-by-event basis and the value $\sigma_{\tau}$ is measured for each $B^0$ candidate, based on the tracking error matrix of the four final state particles. The probability term $P_s(\sigma_{\theta_i}|p_{rT})$ is introduced to account for differences between signal and background events for the values of the per-candidate time errors. Distributions of this variable for signal and background described by gamma functions are shown in Fig. 6. The average value of the time error for signal events is 69 fs.
Table 4 The ten time-dependent functions, \(O^{(k)}(t)\) and the functions of the transversity angles \(g^{(k)}(\theta_T, \psi_T, \phi_T)\). The amplitudes \(|A_0(0)|^2\) and \(|A_1(0)|^2\) are for the CP-even components of the \(B_s^0 \to J/\psi K^+ K^-\) decay, \(|A_0(0)|^2\) is the CP-odd amplitude; they have corresponding strong phases \(\delta_0, \delta_3\) and \(\delta_L\). By convention, \(\delta_0\) is set to be zero. The S-wave amplitude \(|A_S(0)|^2\) gives the fraction of \(B_s^0 \to J/\psi K^+ K^-\) and has a related strong phase \(\delta_S\). The factor \(\alpha\) is described in the text of Sect. 5.1. The \(\pm\) and \(\mp\) terms denote two cases: the upper sign describes the decay of a meson that was initially a \(B_s^0\) meson, while the lower sign describes the decays of a meson that was initially \(B_s^0\)

\[
\begin{array}{ll}
|A_0(0)|^2 & = (1 + \cos \phi_T) e^{-\tau_{1/2}} + (1 - \cos \phi_T) e^{-\tau_{1/2}} + 2e^{-\tau_{1/2}} \sin (\Delta m t) \sin \phi_T \\
|A_1(0)|^2 & = (1 + \cos \phi_T) e^{-\tau_{1/2}} + (1 - \cos \phi_T) e^{-\tau_{1/2}} \mp 2e^{-\tau_{1/2}} \sin (\Delta m t) \sin \phi_T \\
|A_S(0)|^2 & = 1 - 2\cos^2 \phi_T \\
\end{array}
\]

The proper decay time uncertainty distribution for data (black), and the fits to the background (blue) and the signal (purple) contributions. The total fit is shown as a red curve.

![Fig. 6](image)

The same approach was applied for the probability terms \(P_s(\sigma_m | p_T)\) and \(P_s(p_T)\) accounting for differences between signal and background events for the values of the per-candidate mass error and \(p_T\) values, respectively. The tagging probability term for signal \(P_s(P_t | Q_s)\) is described in Sect. 4.4.

The term \(P_s(p_t, \Omega_s, |\sigma_t|, P_t | Q_s)\) is a joint PDF for the decay time \(t\) and the transversity angles \(\Omega_s\) for the \(B_s^0 \to J/\psi (\mu^+ \mu^-) \phi (K^+ K^-)\) decay. Ignoring detector effects, the distribution for the time \(t\) and the angles \(\Omega_s\) is given by the differential decay rate [39]:

\[
\frac{d^2 \Gamma}{dt d\Omega} = \sum_{k=1}^{10} O^{(k)}(t) g^{(k)}(\theta_T, \psi_T, \phi_T),
\]

where \(O^{(k)}(t)\) are the time-dependent functions corresponding to the contributions of the four different amplitudes \((A_0, A_1, A_L, A_S)\) and their interference terms, and \(g^{(k)}(\theta_T, \psi_T, \phi_T)\) are the angular functions. Table 4 shows the time-dependent and the angular functions of the transversity angles. The formulae for the time-dependent functions have the same structure for \(B_s^0\) and \(B_s^0\) but with a sign reversal in the terms containing \(\Delta m t\), which is a fixed parameter of the fit (using Ref. [33]). The formalism used throughout this analysis assumes no direct CP violation, i.e. \(\lambda = (q/p)(\bar{A}/A) = 1\). In Table 4, the parameter \(A_L(t)\) is the time-dependent amplitude for the CP-odd final-state configuration while \(A_0(t)\) and \(A_1(t)\) correspond to CP-even final-state configurations. The amplitude \(A_S(t)\) gives the contribution from the CP-odd non-resonant \(B_s^0 \to J/\psi K^+ K^-\) S-wave state [40] (which includes the \(f_0\) meson). The corresponding functions are given in the last four lines of Table 4 (\(k = 7\)–10). The amplitudes are parameterised by \(|A_i|e^{i\theta_i}\), where \(i = \{0, 1, L, S\}, \delta_0 = 0\) and are normalised such that \(|A_0(0)|^2 + |A_1(0)|^2 + |A_0(0)|^2 = 1\). The amplitude \(|A_S(0)|^2\) is determined according to this condition, while the remaining three amplitudes are parameters of the fit. The value \(|A_s|^2\) is the ratio of the S-wave yield to the \(\phi \to K^+ K^-\) yield in the interval of \(m(K^+ K^-)\) used in the analysis. In the sum over the mass interval, the interference terms (lines 8–10 in
Table 4) are corrected by a factor $\alpha$ that takes into account the mass-dependent differences in absolute amplitude and phase between the $\phi \to K^+K^-$ and $S$-wave amplitudes. The correction is based on the Breit–Wigner description of the $\phi$ and on model assumptions for the shape and the phase variations of the $S$-wave amplitude. The phase $\delta_S$ is the phase difference between $A_0(0)$ and the $S$-wave amplitudes at the $\phi \to K^+K^-$ peak. The values of $\alpha$ and the related systematic uncertainty are discussed in Sect. 6.

The angles $(\theta_T, \psi_T, \phi_T)$, are defined in the rest frames of the final-state particles. The $x$-axis is determined by the direction of the $\phi$ meson in the $J/\psi$ rest frame, and the $K^+K^-$ system defines the $x$–$y$ plane, where $p_y(K^+) > 0$. The three angles are defined as:

- $\theta_T$, the angle between $\vec{p}(\mu^+)$ and the normal to the $x$–$y$ plane, in the $J/\psi$ meson rest frame,
- $\phi_T$, the angle between the $x$-axis and $\vec{p}_{xy}(\mu^+)$, the projection of the $\mu^+$ momentum in the $x$–$y$ plane, in the $J/\psi$ meson rest frame,
- $\psi_T$, the angle between $\vec{p}(K^+)$ and $-\vec{p}(J/\psi)$ in the $\phi$ meson rest frame.

The angular acceptance of the detector and the kinematic cuts on the angular distributions are included in the likelihood function through $A(\Omega_i, p_T)$. This is calculated using a four-dimensional binned acceptance method, applying an event-by-event efficiency according to the transversity angles $(\theta_T, \psi_T, \phi_T)$ and the $p_T$ of the candidate. The $p_T$ binning is necessary, because the angular acceptance is influenced by the $p_T$ of the $B^0$ candidate. The acceptance is calculated from the $B^0 \to J/\psi \phi$ MC events with additional weighting for $p_T$ and $q$ distributions. In the likelihood function, the acceptance is treated as an angular acceptance PDF, which is multiplied with the time- and angle-dependent PDF describing the $B^0_s \to J/\psi (\mu^+ \mu^-) \phi(K^+K^-)$ decays. As both the acceptance and time- and angle-dependent PDFs depend on the transversity angles they must be normalised together. This normalisation is done numerically during the likelihood fit. The PDF is normalised over the entire $B^0_s$ mass range, 5.150–5.650 GeV.

5.2 Background PDF

The background PDF has the following composition:

$$
\mathcal{F}_{\text{bkg}}(m_t, t_i, \sigma_t, \Omega_i, p_T(B|Q_x), p_T) = P_b(m_t) \cdot P_b(t_i|\sigma_t, p_T) \cdot P_b(p_T(B|Q_x)) \cdot P_b(\Omega_i) \cdot P_b(\sigma_m(p_T) \cdot P_b(\sigma_n(p_T)) \cdot P_b(p_T).
$$

The proper decay time function $P_b(t_i|\sigma_t)$ is parameterised as a peak modelled by a Gaussian distribution, two positive-time exponential functions and two negative-time exponential functions. These functions are convolved with the same resolution function, defined in Eq. (4) as the signal decay time-dependence. The prompt peak models the combinatorial background events, which are expected to have reconstructed lifetimes distributed around zero. The two positive-time exponential functions represent a fraction of longer-lived backgrounds with non-prompt $J/\psi$, combined with hadrons from the primary vertex or from a $B/D$ meson in the same event. The two negative-time exponential functions take into account events with poor vertex resolution. The probability terms $P_b(\sigma_m(p_T)), P_b(\sigma_n(p_T))$ and $P_b(p_T)$ are described by gamma functions. They are unchanged from the analysis described in Ref. [41] and explained in detail there. The tagging probability term for background events $P_b(p_T(B|Q_x))$ is described in Sect. 4.4.

The shape of the background angular distribution, $P_b(\Omega_i)$ arises primarily from detector and kinematic acceptance effects. The best description is achieved by Legendre polynomial functions:

$$
Y^m_l(\theta_T) = \sqrt{(2l+1)/(4\pi)}(l-m)!/(l+m)! P_l^{|m|}(\cos \theta_T),
$$

$$
P_k(x) = \frac{1}{2^{k!}k!}(x^2 - 1)^k,
$$

$$
P_l^{|m|}(x) = (-1)^{|m|}1/(2l+1)[(1 - x^2)^{|m|/2}d^{|m|}/dx^{|m|}(P_l(x))],
$$

where $l_{\text{max}} = 14$, $k_{\text{max}} = 14$ and the coefficients $a_{k,l,m}$ are adjusted to give the best fit to the angular distributions for events in the sidebands of the $B^0_s$ mass distribution. These parameters are then fixed in the main fit, defined by Eq. (2). The $B^0_s$ mass interval used for the background fit is between 5.150 and 5.650 GeV excluding the signal mass region $(m(B^0_s) - 3.566) < 0.110$ GeV. Higher-order Legendre polynomial functions were tested as a systematic check, described in Sect. 6.

The background mass model, $P_b(m_t)$ is the sum of an exponential and a constant, with the exponential slope and the relative normalisation left as free parameters of the fit.

Contamination from $B_d \to J/\psi K^{0*}$ and $\Lambda_{b*} \to J/\psi K^-$ events misreconstructed as $B^0_s \to J/\psi \phi$ is accounted for in the fit through the $\mathcal{F}_{B^0_s}$ and $\mathcal{F}_{\Lambda_b}$ terms in the PDF described in Eq. (2). The PDFs are determined using MC simulation of these decays. The fractions of these contributions, $f_{B^0} = (4.3 \pm 0.5\%)$ and $f_{\Lambda_b} = (2.1 \pm 0.6\%)$, are defined relative to the number of the $B^0_s \to J/\psi \phi$ signal events and are evaluated from MC simulation using production cross sections and branching fractions from Refs. [33,42–46].

© Springer
MC simulated events are also used to determine the shape of the mass and transversity angle distributions. The 3D angular distributions of $B_d^0 \to J/\psi K_{s0}$ and of the conjugate decay are modelled using input from Ref. [47], while angular distributions for $B_d \to J/\psi K^-$ and the conjugate decay are considered flat. These distributions are sculpted for detector acceptance effects and then described by Legendre polynomial functions with $l_{\text{max}} = 10$ and $k_{\text{max}} = 10$, Eq. (5). These shapes are used as templates in the fit. The $B_d$ and $\Lambda_b$ lifetimes are accounted for in the fit by adding additional exponential terms, scaled by the ratio of $B_d/B_0^0$ or $\Lambda_b/B_0^0$ masses as appropriate, where the lifetimes and masses are taken from Ref. [33]. The PDF terms that describe each of the tagging, mass, decay time and $p_T$ probability distributions are taken from the same PDFs used to describe the $B_d^0 \to J/\psi K_{s0}$ signal (described in Sects. 4.4 and 5.1) to account for the fact that these dedicated background events are fully reconstructed $b$-hadron decays. Systematic uncertainties due to the background from $B_d \to J/\psi K_{s0}$ and $\Lambda_b \to J/\psi pK^-$ decays are described in Sect. 6. The contribution of the $S$-wave $B_d \to J/\psi K\pi$ decays as well as their interference with the $P$-wave $B_d \to J/\psi K_{s0}$ decays are included in the PDF of the fit, using the parameters measured in Ref. [47].

5.3 Proper decay time dependence of muon trigger efficiency

In the triggers used in this analysis, there is no minimum cut applied on the transverse impact parameter $d_0$ of muons. On the other side, trigger muons with values of $d_0 > 10$ mm are not accepted. This results in inefficiency at large values of the proper decay time. This inefficiency is estimated using MC simulated events, by comparing the $B_d^0$ proper decay time distribution obtained before and after applying the trigger selection. To account for this inefficiency in the fit, the events are reweighted by a factor $w$, inversely proportional to the trigger efficiency:

$$1/w = p_0 \cdot [1 - p_1 \cdot (\text{Erf}((t - p_3)/p_2) + 1)],$$

where Erf denotes the error function and $p_0$, $p_1$, $p_2$ and $p_3$ are parameters determined in the fit to MC events. For more than 99% of the $B_d^0$ candidates the inefficiency is below 2%. For the most affected candidates the inefficiency reaches up to 50% at high decay time. No significant bias or inefficiency due to offline track reconstruction, vertex reconstruction, or track quality selection criteria is observed.

5.4 Summary of the fit parameters

The joint PDF of proper decay time and decay angles includes the main physics parameters of interest:

- the $CP$-violating phase $\phi_3$,
- the average decay width $\Gamma$ and the decay width difference $\Delta \Gamma$,
- the size of the $CP$-state amplitudes at $t = 0$: $|A_1(0)|^2$, $|A_0(0)|^2$ and their corresponding strong phases $\delta_1$ and $\delta_0$,
- and the size of the $S$-wave amplitude at $t = 0$: $|A_S(0)|^2$ and corresponding strong phase $\delta_S$.

The size of the remaining amplitude $|A_0(0)|^2$ is constrained by the normalisation condition, phase $\delta_0$ is set to zero and $\Delta m_s$ is fixed as mentioned above.

The likelihood function also includes other parameters referred to as “nuisance parameters” such as: the $B_d^0$ signal fraction $f_s$, parameters describing the invariant mass and decay time-angular distributions of combinatorial background events and scale factors of mass and decay time uncertainties. In addition, there are also other nuisance parameters describing: acceptance functions, parametrisations of the angles of dedicated backgrounds $B_d \to J/\psi K_{s0}$ and $\Lambda_b \to J/\psi pK^-$ and their fractions $f_{B_d}$ and $f_{\Lambda_b}$, the probability density functions of time error distributions $P(|\sigma_1|/p_T)$, mass error distributions $P(|\sigma_m|/p_T)$, $p_T$ distributions $P(p_T)$ and tagging parameters and calibrations. These parameter values are mainly fixed in the fit to the values extracted from the $B_d^0$ mass signal and sideband regions or from MC simulations.

6 Systematic uncertainties

Systematic uncertainties are evaluated for effects that are described below.

- **Flavour tagging:** The effects on the main physics parameters from the fit, due to uncertainties introduced by the flavour tagging procedure, are assessed as follows: The statistical uncertainty due to the size of the sample of $B^\pm \to J/\psi K^\pm$ decays is included in the overall statistical uncertainty. The systematic uncertainty arising from the precision of the OST calibration, described in Sect. 4.2, is estimated by changing the models used to parameterise the probability distribution, $P(B|Q_s)$, as a function of the cone charge from the function used by default (a third-order polynomial for muons and a sinusoid for electrons) to one of several alternative functions: a linear function; a fifth-order polynomial; or two third-order polynomials that describe the positive and negative regions and have common constant and linear terms, but independent quadratic and cubic terms. The $B_d^0$ fit is repeated using the alternative models and the largest deviation from the nominal fit is assigned as the systematic uncertainty. To validate the calibration procedure,
calibration curves are derived from simulated samples of $B^\pm$ and $B_s^0$ signals. The variations between the curves from these two samples are propagated to the calibration curves derived from data. The differences in the parameter values between the nominal fit and that with the varied calibration curves are included in the systematic uncertainty.

An additional systematic uncertainty is assigned to account for potential dependencies on the pile-up distribution. The calibration data are split into subsets of approximately equal size, separated according to the estimated pile-up of the event, and separate calibrations are made for each subset. For the $B_s^0$ fit, the fit is repeated using the calibrations corresponding to the estimated pile-up of that event. Differences between the nominal and the modified fit for the parameters of interest are taken as the systematic uncertainty. For the terms $P_0(B(Q_i))$ and $P_s(B(Q_s))$, variations of the parameterisation are considered (including using histograms in place of a parameterisation). The resulting changes in the parameter values of the $B_s^0$ fit are similarly included in the systematic uncertainties.

- **Time efficiency:** To correct for the proper decay time dependence of trigger inefficiencies, the events are reweighted according to Eq. (5). To estimate systematic uncertainties connected with this procedure, several alternative fits are performed. Firstly, using different sets of $p_T$ binning in the MC sample used to determine the efficiency. Secondly, to assess the effects of mis-modelling of the $B_s^0$ vertex $x^2/\text{ndof}$ in simulated data, an alternative fit is done with MC $x^2/\text{ndof}$ reweighted according to data. Finally, the groups of similar triggers used to determine the efficiencies were subdivided further by their common features. These smaller groups allowed to address even more details of trigger varieties in time-efficiency determination. Deviations from the default fit result are included in the systematic uncertainties of the measurement.

- **Angular acceptance model:** The angular acceptance of the detector and the kinematic cuts, $A(Q_1, p_T)$, described in Sect. 5.1, is calculated from a binned fit to MC simulated data. In order to estimate the systematic uncertainty introduced by the choice of binning, different acceptance functions are calculated using different numbers of $p_T$ bins as well as different widths and central values of the bins.

- **ID alignment:** The changes of the fit parameters due to residual misalignments of the ID were studied and the observed deviations are included in the systematic uncertainties.

- **Angular acceptance method:** The angular acceptance of the detector and the kinematic cuts, $A(Q_1, p_T)$, described in Sect. 5.1, is calculated from a binned fit to MC simulated data. In order to estimate the systematic uncertainty introduced by the choice of binning, different acceptance functions are calculated using different numbers of $p_T$ bins as well as different widths and central values of the bins.

- **Background angles model:** The shape of the background angular distribution, $P_b(\theta_r, \varphi_r, \psi_r)$, is described by the fourteenth-order Legendre polynomial functions, given in Eq. (5). Alternatively, higher-order Legendre polynomial functions with $l_{\text{max}} = 16$ and $k_{\text{max}} = 16$ were tested, and the changes in the fit parameter values relative to the default fit are taken as systematic uncertainties.

The shapes are primarily determined by detector and kinematic acceptance effects and are sensitive to the $p_T$ of the $B_s^0$ meson candidate. For this reason, the parameterisation using the Legendre polynomial functions is performed in six $p_T$ intervals: 10–15 GeV, 15–20 GeV, 20–25 GeV, 25–30 GeV, 30–35 GeV and >35 GeV. The systematic uncertainties due to the choice of $p_T$ intervals are estimated by repeating the fit, with these intervals enlarged and reduced by 1 GeV and by 2 GeV. The largest changes in the fit results are taken to represent the systematic uncertainties.

The sensitivity of the fit to the choice of the invariant mass window is tested by varying its size. The difference to the default fit results is assigned as a systematic uncertainty.

The parameters of the Legendre polynomial functions given in Eq. (5) are adjusted to give the best fit to the angular distributions for events in the $B_s^0$ mass sidebands. To test the sensitivity of the fit results to the choice of sideband regions, the fit is repeated with alternative choices for the excluded signal mass regions: $|m(B_s^0) - 5.366\text{ GeV}| > 0.085\text{ GeV}$ and $|m(B_s^0) - 5.366\text{ GeV}| > 0.160\text{ GeV}$ (instead of the default $|m(B_s^0) - 5.366\text{ GeV}| > 0.110\text{ GeV}$). The changes in the fit results are assigned as systematic uncertainties.

- **$B_d$ contribution:** The contamination from $B_d \to J/\psi K^{0*}$ events misreconstructed as $B_s^0 \to J/\psi \phi$ is accounted for in the final fit. Studies are performed to evaluate the systematic uncertainties due to trigger effects, the $B_d \to J/\psi K^{0*}$ fraction, and the distributions of the mass, transversity angles, and lifetime PDFs. In the MC events the angular distribution of the $B_d \to J/\psi K^{0*}$ decay is modelled using parameters taken from Ref. [47]. The contribution of the $S$-wave $B_d \to J/\psi K\pi$ decays as well as its interference with the $P$-wave $B_d \to J/\psi K^{0*}$ decays are also included in the PDF of the fit, following the parameters measured in Ref. [47]. The uncertainties of these parameters are taken into account in the estima-
tion of the systematic uncertainty. After applying the $B_s^0$ signal selection cuts, the angular distributions are fitted using Legendre polynomial functions. The uncertainties of this fit are included in the systematic uncertainty.

- **$\Lambda_b$ contribution**: The contamination from $\Lambda_b \rightarrow J/\psi p K^-$ events misreconstructed as $B_s^0 \rightarrow J/\psi \phi$ is accounted for in the final fit. Studies are performed to evaluate the effect of the uncertainties in the $\Lambda_b \rightarrow J/\psi p K^-$ fraction $f_{\Lambda_b}$, and the shapes of the distributions of the mass, transversity angles, and lifetime. Additional studies are performed to determine the effect of the uncertainties in the $\Lambda_b \rightarrow J/\psi \Lambda^*$ branching ratios used to reweight the generated MC sample.

- **Alternate $\Delta m_s$**: The systematics due to fixing the parameter $\Delta m_s$ to the PDG value were estimated by running an alternative fit where the default model was altered by releasing $\Delta m_s$ within a Gaussian constraint of which the width was taken from uncertainties assigned to $\Delta m_s$ in [33]. The resulting changes to all the fit parameter values are found to be negligible except for $\delta_\perp$. The effects on the other variables are small for the parameters $\phi_s$ and $\delta_\parallel$.

- **Fit model mass and lifetime**: To estimate the systematic uncertainties due to the signal $B_s^0$ mass model, the default model was altered by adding a second Gaussian function in Eq. (3), which has the same structure as the first Gaussian function but a different scale factor, $S^1_m$, which is an additional free parameter of the fit. The resulting changes to other fit parameter values are found to be negligible. To test the sensitivity of the part of the fit model describing the lifetime, two systematic tests are performed. The determination of signal and background lifetime errors is sensitive to the choice of $t_\tau$ bins, in which the relative contributions of these two components are evaluated. To estimate the systematic uncertainty, the fit is repeated varying the intervals of the default $t_\tau$ binning. The determination of signal and background lifetime errors is also sensitive to the determination of the signal fraction. The fit is repeated by varying this fraction within one standard deviation of its uncertainty and differences are included in the systematic uncertainty.

- **Fit model $S$-wave phase**: As explained in Sect. 5.1, the model for the interference between $B_s^0 \rightarrow J/\psi \phi (K^+ K^-)$ and $S$-wave $B_s^0 \rightarrow J/\psi K^+ K^-$ is corrected by a factor $\alpha$ to account for the mass-dependent variations in absolute amplitude and phase of the two amplitudes within the interval $1.0085$–$1.0305$ GeV in $m(K^+ K^-)$. The value of $\alpha$ is $0.51 \pm 0.08$. The central value is obtained under the hypothesis of uniformity of the $S$-wave amplitude. The uncertainty is systematic and it is due to: detector mass resolution and mass scale uncertainties, uncertainties in the description of the $\phi$ resonance (mass, width, and shape descriptions as relativistic Breit–Wigner or Flatté parameterisation [48]), and to uncertainties in the description of the shape and phase variation of the $S$-wave amplitude. For the last effect, which is dominant, the $S$-wave amplitude is assumed to be due to the $f_0(980)$ resonance and it is described as in Ref. [49], similarly to what was done in Ref. [16]. The variation from the value of $\alpha$ obtained with the default assumption is taken as a systematic uncertainty. This procedure uses descriptions of the $f_0(980)$ based on other measurements [33]. To account for the uncertainty in $\alpha$, the fit was repeated with $\alpha = 0.51 + 0.08$ and $\alpha = 0.51 - 0.08$ values. The variations of the parameter values relative to those from the default fit using the central value of $\alpha$ are included in the systematic uncertainties.

- **Fit bias**: Due to its complexity, the fit model can be sensitive to some nuisance parameters. This limited sensitivity could potentially lead to a bias in the measured physics parameters, even when the model describes the fitted data well. To test the stability of the results obtained from the chosen default fit model, a set of pseudo-experiments is conducted using the default model in both the generation and fit. The systematic uncertainties are determined from the mean of the pull distributions of the pseudo-experiments scaled according to the statistical uncertainty of that parameter in the fit to data. The observed deviations are included in the systematic uncertainties.

The systematic uncertainties are listed in Table 5. For each parameter, the total systematic uncertainty is obtained by adding all of the contributions in quadrature.

### 7 Results

#### 7.1 Fit results

The results of the likelihood fit are shown in Table 6. The total number of $B_s^0$ meson candidates is $453\,570 \pm 740$. The fitted value of the $B_s^0$ mass agrees well with the world average value [33]. Fit projections, including ratio plots, are shown in Fig. 7 for the mass and proper decay time and in Fig. 8 for the angles. The ratio plots show the difference between each data point and the total fit line divided by the statistical and systematic uncertainties summed in quadrature ($\sigma$) for that point. The deviations of ratio plots are within $2\sigma$, which shows that the total uncertainties cover any discrepancy between data and fit model.

While for most of the physics parameters, including $\phi_s$, $\Delta \Gamma_s$, and $\Gamma_s$, the fit determines a single solution with Gaussian behaviour of the projection of the log-likelihood (see Fig. 10 in Sect. 7.2), for the strong-phases $\delta_\parallel$ and $\delta_\perp$ two well separated local maxima of the likelihood are found, and shown as solution (a) and (b) in Table 6. The difference in
$-2\Delta \ln(L)$ between the two solutions is 0.03. As discussed in Sect. 7.2, the two-fold behaviour of the likelihood in the strong phases is the result of an approximate symmetry of the signal PDF. The effect is completely negligible for all other variables, for which the fit values and uncertainty ranges overlap accurately.

The correlation between statistical uncertainties of the parameters are also shown in Table 7 for solution (a) and in Table 8 for solution (b). The correlations do not change significantly between the physics variables which remain stable between the solutions (a) and (b). The correlations with the two strong phases: $\delta_\parallel$ and $\delta_\perp$, flip the sign, as expected, except for the correlations that are smaller than 0.01. The correlations between $\delta_\parallel$ and $\delta_\perp$ themselves keep the same sign in both solutions (a) and (b).

Releasing the direct $CP$ related $\lambda$ parameter (see Sect. 5.1) in the fit results in a $\lambda$ value compatible within the statistical precision with unity as well as with results of other experiments. The effect on the values of the other parameters of interest is negligible, due to small correlations with $\lambda$.

7.2 Fit to strong phases

As shown in Table 6, the likelihood fit has determined two solutions with well separated values for the strong phases $\delta_\parallel$ and $\delta_\perp$. Figure 9 shows results of the 2D log-likelihood scan in the $\delta_\parallel$, $\delta_\perp$ plane, revealing two minima, identified at $(\delta_\parallel = 3.35, \delta_\perp = 3.12)$ and $(\delta_\parallel = 2.94, \delta_\perp = 2.91)$. These minima are represented by two-dimensional contours at the level of $-2\Delta \ln(L) = 2.30, 6.18$, and 11.83, where $-2\Delta \ln(L) = 2(\ln(L^1) - \ln(L^0))$ is the difference between the likelihood values ($L^1$) of the fit in which the two strong phases are fixed to the values shown on the horizontal and vertical axis, and $L^0$ which is the likelihood value for solution (a) of the fit.

An approximate symmetry in the signal PDF is at the origin of this duality. The strong phases are determined by the six interference terms (4) – (6) and 8) – 10) in Table 4. Four of them, namely terms 4), 5), 8) and 9), are invariant under the transformation

$$\{\delta_\parallel, \delta_\perp, \delta_3\} \rightarrow \{2\pi - \delta_\parallel, \delta_\perp + 2(\pi - \delta_\parallel), \delta_3 + 2(\pi - \delta_\parallel)\}. \ (6)$$

This transformation is proportional to $\pi - \delta_\parallel$, which from our data is equal to 0.21. The two local maxima of the likelihood determined in this analysis satisfy the relation of Eq. (6) very accurately for $\delta_\parallel$ and $(\delta_\perp - \delta_3)$, and within the statistical uncertainty in the transformation for $\delta_\perp$. The difference in the log-likelihoods, $-2\Delta \ln(L)$, between the two solutions is equal to 0.03, favouring (a) but without ruling out (b).

As discussed in Sect. 7.1, the two-fold nature of the likelihood maxima for the strong phases has a negligible effect on
Table 6 Fitted values for the physical parameters of interest with their statistical and systematic uncertainties. For variables $\delta_\perp$ and $\delta_\parallel$ the values are given for the two solutions (a) and (b). The difference in $-2\Delta\ln(L)$ between solutions (b) and (a) is 0.03. For the rest of the variables the values for the two minima are consistent. The same is true for the statistical and systematic uncertainties of all the variables.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Statistical uncertainty</th>
<th>Systematic uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_s$ [rad]</td>
<td>$-0.081$</td>
<td>$0.041$</td>
<td>$0.022$</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>$0.0607$</td>
<td>$0.0047$</td>
<td>$0.0043$</td>
</tr>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>$0.6687$</td>
<td>$0.0015$</td>
<td>$0.0022$</td>
</tr>
<tr>
<td>$</td>
<td>A_1(0)</td>
<td>^2$</td>
<td>$0.2213$</td>
</tr>
<tr>
<td>$</td>
<td>A_0(0)</td>
<td>^2$</td>
<td>$0.5131$</td>
</tr>
<tr>
<td>$</td>
<td>A_S(0)</td>
<td>^2$</td>
<td>$0.0321$</td>
</tr>
<tr>
<td>$\delta_\perp - \delta_\parallel$ [rad]</td>
<td>$-0.25$</td>
<td>$0.05$</td>
<td>$0.04$</td>
</tr>
<tr>
<td>Solution (a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_\perp$ [rad]</td>
<td>$3.12$</td>
<td>$0.11$</td>
<td>$0.06$</td>
</tr>
<tr>
<td>$\delta_\parallel$ [rad]</td>
<td>$3.35$</td>
<td>$0.05$</td>
<td>$0.09$</td>
</tr>
<tr>
<td>Solution (b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_\perp$ [rad]</td>
<td>$2.91$</td>
<td>$0.11$</td>
<td>$0.06$</td>
</tr>
<tr>
<td>$\delta_\parallel$ [rad]</td>
<td>$2.94$</td>
<td>$0.05$</td>
<td>$0.09$</td>
</tr>
</tbody>
</table>

Fig. 7 (Left) Mass fit projection for the $B_0^0 \to J/\psi \phi$ sample. The red line shows the total fit, the short-dashed magenta line shows the $B_0^0 \to J/\psi \phi$ signal component, the combinatorial background is shown as a blue dotted line, the orange dash-dotted line shows the $B_0^0 \to J/\psi K^0_s$ component, and the green dash-dot-dot line shows the contribution from $\Lambda_{b}\to J/\psi p K^-$ events. (Right) Proper decay time fit projection for the $B_0^0 \to J/\psi \phi$ sample. The red line shows the total fit while the short-dashed magenta line shows the total signal. The total background is shown as a blue dotted line, and a long-dashed grey line shows the prompt $J/\psi$ background component. Below each figure is a ratio plot that shows the difference between each data point and the total fit line divided by the statistical and systematic uncertainties summed in quadrature ($\sigma$) of that point.

8 Combination with 7 TeV and 8 TeV results

The measured values of solution (a) and solution (b) are consistent with those obtained in the previous analysis [13] using 19.2 fb$^{-1}$ of data collected at $\sqrt{s} = 7$ TeV and 8 TeV. A best linear unbiased estimator (BLUE) [50,51] is used to combine the current results with those from the previous analysis. The measured values, uncertainties, and correlations are...
Fig. 8 Fit projections for the transversity angles $\phi_T$ (top left), $\cos(\theta_T)$ (top right), and $\cos(\psi_T)$ (bottom). In all three plots the red solid line shows the total fit, the $B_0^s \rightarrow J/\psi \phi$ signal component is shown by the magenta dashed line and the blue dotted line shows the contribution of all background components. Below each figure is a ratio plot that shows the difference between each data point and the total fit line divided by the statistical and systematic uncertainties summed in quadrature ($\sigma$) of that point.

The two-dimensional likelihood contours in the $\phi_T - \Delta \Gamma_\gamma$ plane for the ATLAS result based on 7 and 8 TeV data, the solution (a) of the 13 TeV measurement, and the combined result for the solution (a) are shown in Fig. 11. The statistical and systematic uncertainties are combined in quadrature and correlations are taken into account in the construction of Gaussian contours. Because there is no significant difference in the $\phi_T$ and $\Delta \Gamma_\gamma$ values between solution (a) and solution (b), only contours for solution (a) are shown.

Two-dimensional likelihood contours in the $\phi_T - \Delta \Gamma_\gamma$ plane are shown in Fig. 12 for this ATLAS result, the result from CMS $[17]$ using the $B_0^s \rightarrow J/\psi \phi$ decay, the result from
Table 7  Fit correlations between the physical parameters of interest, obtained from the fit for solution (a)

| $\Delta \Gamma$ | $\Gamma_s$ | $|A_1(0)|^2$ | $|A_0(0)|^2$ | $|A_S(0)|^2$ | $\delta_1$ | $\delta_\perp$ | $\delta_\perp - \delta_1$ |
|----------------|-----------|-------------|-------------|-------------|-----------|----------|-----------------|
| $\phi_s$       | -0.080    | 0.017       | -0.003      | -0.004      | -0.007    | 0.007    | 0.004           |
| $\Delta \Gamma'$ | 1         | -0.586      | 0.090       | 0.095       | 0.051     | 0.032    | 0.005           |
| $\Gamma_s$     | 1         | -0.125      | -0.045      | -0.172      | 0.522     | 0.133    | -0.052          |
| $|A_{||}(0)|^2$ | 1         | -0.341      | -0.172      | 0.522       | 0.133     | -0.052  | 0.015           |
| $|A_{0}(0)|^2$  | 1         | 0.276       | -0.103      | -0.034      | 0.070     |         |                 |
| $|A_S(0)|^2$    | 1         | -0.362      | -0.118      | 0.244       |          |         |                 |
| $\delta_1$     | 1         |             |             |             | 0.254     | -0.085  |                 |
| $\delta_\perp$ |           |             |             |             | 1         | 0.001   |                 |

Table 8  Fit correlations between the physical parameters of interest, obtained from the fit for solution (b)

| $\Delta \Gamma$ | $\Gamma_s$ | $|A_1(0)|^2$ | $|A_0(0)|^2$ | $|A_S(0)|^2$ | $\delta_1$ | $\delta_\perp$ | $\delta_\perp - \delta_1$ |
|----------------|-----------|-------------|-------------|-------------|-----------|----------|-----------------|
| $\phi_s$       | -0.084    | 0.019       | -0.011      | -0.003      | -0.006    | 0.007    | 0.005           |
| $\Delta \Gamma'$ | 1         | -0.586      | 0.090       | 0.096       | 0.057     | -0.029   | -0.010          |
| $\Gamma_s$     | 1         | -0.116      | -0.048      | -0.110      | 0.444     | -0.106   | -0.052          |
| $|A_{||}(0)|^2$ | 1         | -0.338      | -0.110      | -0.444      | -0.106    | -0.052  |                 |
| $|A_{0}(0)|^2$  | 1         | 0.269       | 0.080       | 0.017       | 0.017     |         |                 |
| $|A_S(0)|^2$    | 1         | 0.291       | 0.060       | 0.025       | 0.025     |         |                 |
| $\delta_1$     | 1         |             |             |             | 0.235     | 0.097   |                 |
| $\delta_\perp$ |           |             |             |             | 1         | 0.056   |                 |

![Fig. 9](image)

Fig. 9  Two-dimensional constraints on the values of $\delta_1$ and $\delta_\perp$ for solutions (a) and (b) at the level of $-2\Delta ln(L) = 2.30, 6.18, \text{ and } 11.83$ respectively created using a full 2D scan. The minimum of the solution (b) is $-2\Delta ln(L) = 0.03$ higher than the minimum of the solution (a).

LHCb [16] using the $B_s^0 \rightarrow J/\psi K^+ K^-$ decay and finally the LHCb result including all $B_s^0$ channels [16,18–21]. The contours are obtained by interpreting each result as a two-dimensional Gaussian probability distribution in the $\phi_s-\Delta \Gamma_s$ plane. All results are consistent with each other and with the SM [5,52].

9 Summary

This paper presents a measurement of the time-dependent $CP$ asymmetry parameters in $B_s^0 \rightarrow J/\psi \phi$ decays from an 80.5 fb$^{-1}$ data sample of $pp$ collisions collected with the ATLAS detector during the 13 TeV LHC run. The values from the 13 TeV analysis are consistent with those obtained in the previous ATLAS analysis using 7 TeV and 8 TeV data. The two measurements are statistically combined.

The $CP$-violating phase $\phi_s$ is measured to be $-0.087 \pm 0.036 \text{ (stat.)} \pm 0.021 \text{ (syst.)} \text{ rad}$, the decay width difference between heavy and light $B_s^0$ mass eigenstates, $\Delta \Gamma_s = 0.0657 \pm 0.0043 \text{ (stat.)} \pm 0.0037 \text{ (syst.)} \text{ ps}^{-1}$ and the measurement for their average decay width, $\Gamma_s = 0.6703 \pm 0.0014 \text{ (stat.)} \pm 0.0018 \text{ (syst.)} \text{ ps}^{-1}$. The measurement of the $CP$-violating phase $\phi_s$ is consistent with the Standard Model prediction, and it improves on the precision of previous ATLAS measurements, while for the average decay width, $\Gamma_s$, the comparison of the 13 TeV analysis with the current world combined value reveals a tension at the level of $3\sigma$. 
Fig. 10 1D log-likelihood scans of all other variables of the fit for the solution (a) (blue) and the solution (b) (dashed red). The variable on the vertical axis, $-2\Delta\ln(L) = 2(\ln(L^a) - \ln(L'))$, is the difference between the likelihood values of the default fit, $(L^a)$, and of the fit in which the physical parameter is fixed to values shown on the horizontal axis.

Table 9 Values of the physical parameters extracted in the combination of solution (a) and solution (b) of 13 TeV results with those obtained from 7 and 8 TeV data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Solution (a)</th>
<th>Solution (b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>Statistical uncertainty</td>
<td>Systematic uncertainty</td>
</tr>
<tr>
<td>$\phi_b$ [rad]</td>
<td>-0.087</td>
<td>0.036</td>
</tr>
<tr>
<td>$\Delta\Gamma_\gamma$ [ps$^{-1}$]</td>
<td>0.0657</td>
<td>0.0043</td>
</tr>
<tr>
<td>$\Gamma_\gamma$ [ps$^{-1}$]</td>
<td>0.6703</td>
<td>0.0014</td>
</tr>
<tr>
<td>$</td>
<td>A_4(0)</td>
<td>^2$</td>
</tr>
<tr>
<td>$</td>
<td>A_0(0)</td>
<td>^2$</td>
</tr>
<tr>
<td>$</td>
<td>A_S</td>
<td>^2$</td>
</tr>
<tr>
<td>$\delta_\perp$ [rad]</td>
<td>3.22</td>
<td>0.10</td>
</tr>
<tr>
<td>$\delta_\parallel$ [rad]</td>
<td>3.36</td>
<td>0.05</td>
</tr>
<tr>
<td>$\delta_\perp - \delta_S$ [rad]</td>
<td>-0.24</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Table 10  Correlation matrix of the BLUE combination of the 7 and 8 TeV results and the solution (a) of the 13 TeV result

|   | \(\Delta \Gamma\) | \(\Gamma_s\) | \(|A_s(0)|^2\) | \(|A_0(0)|^2\) | \(|A_\chi(0)|^2\) | \(\delta_\perp\) | \(\delta_\parallel\) | \(\delta_\perp - \delta_\parallel\) |
|---|--------|--------|-------------|-------------|-------------|----------|----------|----------------|
| \(\phi_s\) | -0.025 | 0.001  | 0.001       | 0.000       | -0.001     | 0.003    | 0.003    | -0.004      |
| \(\Delta \Gamma\) | 1      | -0.259 | 0.043       | 0.027       | 0.020      | 0.003    | 0.011    | 0.010        |
| \(\Gamma_s\) | 1      | -0.045 | -0.009      | 0.025       | -0.010     | -0.021   | 0.006    |              |
| \(|A_s(0)|^2\) | 1      | -0.071 | -0.056      | 0.068       | 0.155      | -0.024   |          |              |
| \(|A_0(0)|^2\) | 1      | 0.048  | -0.009      | -0.015      | 0.016      |          |          |              |
| \(|A_\chi(0)|^2\) | 1      | -0.056 | -0.099      | 0.108       |            |          |          |              |
| \(\delta_\perp\) | 1      |        | 0.106       | 0.001       |            |          |          |              |
| \(\delta_\parallel\) |        | 1      | -0.032      |            |            |          |          |              |

Table 11  Correlation matrix of the BLUE combination of the 7 and 8 TeV results and the solution (b) of the 13 TeV result

|   | \(\Delta \Gamma\) | \(\Gamma_s\) | \(|A_s(0)|^2\) | \(|A_0(0)|^2\) | \(|A_\chi(0)|^2\) | \(\delta_\perp\) | \(\delta_\parallel\) | \(\delta_\perp - \delta_\parallel\) |
|---|--------|--------|-------------|-------------|-------------|----------|----------|----------------|
| \(\phi_s\) | -0.027 | 0.002  | -0.003      | 0.001       | -0.001     | 0.004    | 0.002    | -0.003      |
| \(\Delta \Gamma\) | 1      | -0.259 | 0.043       | 0.027       | 0.023      | -0.006   | -0.010   | 0.011        |
| \(\Gamma_s\) | 1      | -0.042 | -0.010      | 0.023       | 0.007      | 0.018    | 0.006    |              |
| \(|A_s(0)|^2\) | 1      | -0.071 | -0.035      | -0.132      | 0.024      |          |          |              |
| \(|A_0(0)|^2\) | 1      | 0.046  | 0.004       | 0.012       | 0.016      |          |          |              |
| \(|A_\chi(0)|^2\) | 1      | 0.027  | 0.080       | 0.111       |            |          |          |              |
| \(\delta_\perp\) | 1      |        | 0.098       | 0.037       |            |          |          |              |
| \(\delta_\parallel\) |        | 1      | 0.036       |            |            |          |          |              |

Fig. 11  Contours of 68% confidence level in the \(\phi_s - \Delta \Gamma_s\) plane, showing ATLAS results for 7 and 8 TeV data (blue dashed-dotted curve), for 13 TeV data (green dashed curve) and for 13 TeV data combined with 7 and 8 TeV (red solid curve) data. The Standard Model prediction \([5,52]\) is shown as a very thin black rectangle. In all contours the statistical and systematic uncertainties are combined in quadrature and correlations are taken into account.

Fig. 12  Contours of 68% confidence level in the \(\phi_s - \Delta \Gamma_s\) plane, including results from CMS (orange) and LHCb (green) using the \(B_s^0 \to J/\psi K^+ K^-\) decay only and LHCb (red) for all the channels. The blue contour shows the ATLAS result for 13 TeV combined with 7 and 8 TeV. The Standard Model prediction \([5,52]\) is shown as a very thin black rectangle. In all contours the statistical and systematic uncertainties are combined in quadrature.
Acknowledgements

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; STFC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MESMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMWF, HGF and MPG, Germany; GSRT, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; MSSR, Slovakia; ARRS and MIŠZ, Slovenia; DST/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada, CRC and IVADO, Canada; Beijing Municipal Science & Technology Commission, China; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSSN and GIF, Israel; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafsson Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [53].

Data Availability Statement

This manuscript has no associated data or the data will not be deposited. [Authors’ comment: All ATLAS scientific output is published in journals, and preliminary results are made available in Conference Notes. All are openly available, without restriction on use by external parties beyond copyright law and the standard conditions agreed by CERN. Data associated with journal publications are also made available: tables and data from plots (e.g. cross section values, likelihood profiles, selection efficiencies, cross section limits, ...) are stored in appropriate repositories such as HEPDATA (http://hepdata.cedar.ac.uk/). ATLAS also strives to make additional material related to the paper available that allows a reinterpretation of the data in the context of new theoretical models. For example, an extended encapsulation of the analysis is often provided for measurements in the framework of RIVET (http://rivet.hepforge.org/). This information is taken from the ATLAS Data Access Policy, which is a public document that can be downloaded from http://opendata.cern.ch/record/413/opendata.cern.ch/].

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP3.

References

3. UTfit Collaboration, M. bona et al., the unitarity triangle fit in the standard model and hadronic parameters from lattice QCD: a reappraisal after the measurements of $\Delta m_s$ and $B(B \to \tau \nu)$.
4. LHCb Collaboration, Precision measurement of the $B^0_s - \bar{B}^0_s$ oscillation frequency with the decay $B^0_s \to D^+_s \pi^-$. New J. Phys. 15, 053021 (2013). arXiv:1304.4741 [hep-ex].
10. D0 Collaboration, Measurement of the CP-violating phase $\phi_{CP}^{J/\psi\phi}$ using the flavor-tagged decay $B_s^0 \to J/\psi\phi$ in 8 $fb^{-1}$ of $p\bar{p}$ collisions. Phys. Rev. D 85, 032006 (2012). arXiv:1109.3166 [hep-ex].
13. ATLAS Collaboration, Measurement of the CP-violating phase $\phi_{BS}$ and the $B_s^0$ meson decay width $\Delta m_s$ with decays in ATLAS. JHEP 08, 147 (2016). arXiv:1601.03297 [hep-ex].
15. LHCb Collaboration, Resonances and CP violation in $B_s^0$ and $B_s^0 \to J/\psi K^* K^-$ decays in the mass region above the $\phi(1020)$. JHEP 08, 037 (2017). arXiv:1704.08217 [hep-ex].


Also at Louisiana Tech University, Ruston, LA, USA
Also at LPNHE, Sorbonne Université, Université de Paris, CNRS/IN2P3, Paris, France
Also at Manhattan College, New York, NY, USA
Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
Also at National Research Nuclear University MEPhI, Moscow, Russia
Also at Physics Department, An-Najah National University, Nablus, Palestine
Also at Physics Dept, University of South Africa, Pretoria, South Africa
Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
Also at School of Physics, Sun Yat-sen University, Guangzhou, China
Also at The City College of New York, New York, NY, USA
Also at The Collaborative Innovation Center of Quantum Matter (CICQM), Beijing, China
Also at TRIUMF, Vancouver, BC, Canada
Also at Universita di Napoli Parthenope, Naples, Italy
*Deceased