Measurement of the CP violation phase $\phi_s$ in $B_s^0 \rightarrow J/\psi \phi$ decays in ATLAS at 13 TeV

The ATLAS Collaboration

A measurement of the $B_s^0 \rightarrow J/\psi \phi$ decay parameters using 80.5 fb$^{-1}$ of integrated luminosity collected with the ATLAS detector from 13 TeV $pp$ collisions at the LHC is presented. The measured parameters include the CP-violating phase $\phi_s$, the width difference $\Delta \Gamma_s$ between the $B_s^0$ 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 TeV and 8 TeV data, leading to the following:

$$\phi_s = -0.076 \pm 0.034 \, \text{(stat.)} \pm 0.019 \, \text{(syst.) \, rad}$$
$$\Delta \Gamma_s = 0.068 \pm 0.004 \, \text{(stat.)} \pm 0.003 \, \text{(syst.) \, ps}^{-1}$$
$$\Gamma_s = 0.669 \pm 0.001 \, \text{(stat.)} \pm 0.001 \, \text{(syst.) \, ps}^{-1}$$

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

© 2019 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.
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 $B^0 \to J/\psi \phi$ decay, CP violation occurs due to interference between a direct decay and a decay with $B^0_s \to B_s^0$ 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 $b \to c\bar{t}s$ decay amplitude. In the SM the phase $\phi_s$ is small and is related to Cabibbo–Kobayashi–Maskawa (CKM) quark mixing matrix elements via the relation $\phi_s \approx -2\beta_s$, with $\beta_s = \arg[-(V_{ts}V_{cb}^*)/(V_{cs}V_{cb}^*)]$; assuming no NP contributions to $B^0_s$ mixing and decays, a value of $-2\beta_s = -0.0363^{+0.0016}_{-0.0015}$ rad can be predicted by combining beauty and kaon physics observables [2]. While large NP enhancements of the mixing amplitude have been excluded by the precise measurement of the oscillation frequency [3], 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 $\Gamma_s = (\Gamma_L + \Gamma_H)/2$ 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. In the SM the width difference is predicted to be $\Delta \Gamma_s = 0.087 \pm 0.021$ ps$^{-1}$ [4]. A potential NP enhancement of $\phi_s$ would also decrease the size of $\Delta \Gamma_s$, however it is not expected to be affected as significantly as $\phi_s$ [5]. Nevertheless, extracting $\Delta \Gamma_s$ from data is interesting as it allows theoretical predictions to be tested [5].

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 [6–10]. Additional improvement in measuring $\phi_s$ from $B^0_s$ decays to $\psi(2S)\phi$ and to $D^*_sD^*_s$ has been achieved by the LHCb collaboration [11, 12].

The analysis presented here introduces a measurement of the $B^0_s \to J/\psi \phi$ decay parameters using 80.5 fb$^{-1}$ of LHC $pp$ 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 [9] that was performed using 19.2 fb$^{-1}$ of data collected at 7 TeV and 8 TeV and introduces more precise models for both signal and backgrounds.

2 ATLAS detector and Monte Carlo simulation

The ATLAS detector* 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 micro-strip, 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 end-cap and forward regions are equipped with LAr calorimeters for electromagnetic and hadronic measurements. The MS surrounds the calorimeters and

* 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.
provides a system of tracking chambers and detectors for triggering. A full description can be found in Refs. [13–15].

The muon and tracking systems are of particular importance in the reconstruction of $B$ meson candidates. Only data collected when both these systems were operating correctly and when the LHC beams were declared to be stable are used in the analysis. The data were collected during periods with different instantaneous luminosity; therefore several triggers were used in the analysis. All of them 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.

The measurement uses $80.5 \text{ fb}^{-1}$ of $pp$ collision data. The uncertainty in the combined 2015–2017 integrated luminosity is 2.0%. It is derived, following a methodology similar to that detailed in Ref. [16], and using the LUCID-2 detector for the baseline luminosity measurements [17], from calibration of the luminosity scale using $x$–$y$ beam-separation scans.

To study the detector response, estimate backgrounds and model systematic effects, 100 million Monte Carlo (MC) simulated $B^0 \rightarrow J/\psi \phi$ events were generated using PYTHIA 8.210 [18] tuned with ATLAS data, using the A14 set of parameters [19] together with the CTEQ6L1 set [20]. The detector response was simulated using the ATLAS simulation framework based on GEANT4 [21, 22]. 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 data. Additional samples of both exclusive ($B^0_d \rightarrow J/\psi K^{0*}$ and $\Lambda_b \rightarrow J/\psi p K^-$) and inclusive ($b\bar{b} \rightarrow J/\psi X$ and $pp \rightarrow J/\psi X$) decays, which are backgrounds for this analysis, were simulated, using the same simulation tools as in case of the signal events. For validation studies related to flavour tagging, events of $B^\pm \rightarrow J/\psi K^\mp$ exclusive decays were also simulated.

3 Reconstruction and candidate selection

The reconstruction and candidate selection for the decay $B^0 \rightarrow J/\psi(\mu^+ \mu^-)\phi(K^+ K^-)$ proceeds as follows. Events must pass the trigger selections described in Section 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 muon track parameters used in this analysis 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 refitted to a common vertex and the pair is accepted for further consideration if the quality of the fit meets the requirement $\chi^2/n.d.o.f. < 10$. 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 and 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
\[ B^0_s \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-) \] 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 calculated from the two muon tracks to the \( J/\psi \) mass [23]. A quadruplet of tracks is accepted for further analysis if the vertex fit has a \( \chi^2/\text{n.d.o.f.} < 3 \). For the \( \phi \rightarrow K^+K^- \) candidate the invariant mass of the track pairs (using a 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 \( \chi^2/\text{n.d.o.f.} \) is selected in cases where more than one candidate passes all selections. In total 3 210 429 \( B^0_s \) candidates are collected within the mass range of \( 5.150–5.650 \text{ GeV} \). This range is chosen to give enough background events in side bands to allow a high precision determination of background events properties. The mass window choice has been varied and found to have a negligible systematic effect on the results.

The mean number of interactions per crossing is 30, necessitating a choice of the best candidate for the primary vertex at which the \( B^0_s \) meson is produced. The variable used is the three-dimensional impact parameter \( a_0 \), which is calculated as the minimum distance between the line extrapolated from the reconstructed \( B^0_s \) meson vertex in the direction of the \( B^0_s \) momentum, and each primary vertex candidate. The chosen primary vertex is the one with the smallest \( a_0 \).

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

\[
t = \frac{L_{xy} m_B}{p_{TB}},
\]

where \( p_{TB} \) 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. [23]. The transverse decay length, \( L_{xy} \), is the displacement in the transverse plane of the \( B^0_s \) meson decay vertex with respect to the primary vertex, projected onto the direction of the \( B^0_s \) transverse momentum. The primary vertex position is recalculated after removing any tracks used in the \( B^0_s \) meson candidate to avoid biasing \( L_{xy} \).

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 approach is called opposite-side tagging (OST).

The OST algorithms each define a discriminating variable, based on charge information, which is sensitive to the flavour of the opposite-side \( b \)-hadron, and are calibrated to provide a probability that a signal \( B \) meson in a given event is produced in a given flavour (ie. containing a \( b \)- or \( \bar{b} \)-quark). 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 \( B^0_s \) meson state, which is used in the maximum likelihood fit (described in Section 5). Section 4.1 describes the reconstruction of the \( B^\pm \rightarrow J/\psi K^\pm \) candidates, followed by a description of the OST methods in Section 4.2. The performance of the OST algorithms on the calibration sample is given in Section 4.3, and details on how the probabilities from the OST algorithms are used in the maximum likelihood fit, including the determination of the distributions of these probabilities in signal and background, discussed in Section 4.4.
Figure 1: The invariant mass distribution for selected $B^\pm \rightarrow J/\psi K^\pm$ candidates. Data are shown as points, and the overall result of the fit is given by the blue curve. The contributions from the combinatorial background component is indicated by the red dotted line, partially reconstructed $b$-hadron decays by the purple shaded area, and decays of $B^\pm \rightarrow J/\psi \pi^\pm$, where the pion is misassigned as a kaon, by the green dashed line.

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

Candidate $B^\pm \rightarrow J/\psi K^\pm$ decays are identified as follows. First, $J/\psi$ candidates are selected from oppositely charged muon pairs forming a good vertex, as described in Section 3. Each muon is required to have a transverse momentum of at least 4 GeV and pseudorapidity $|\eta| < 2.5$. Dimuon candidates with invariant mass $2.8 < m(\mu^+\mu^-) < 3.4$ GeV, as determined from the refitted track parameters of the vertex, are retained for further analysis. To form the $B^\pm$ candidate, an additional track is required. 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 [23]. Prompt background contributions are suppressed with the requirement on the proper decay time of the $B^\pm$ candidate of $t > 0.2$ ps.

The tag 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 parametrize 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 constrains the shapes of the signal and backgrounds, excluding the slope of the exponential function. Subsequent fits are performed in the intervals of the tagging discriminating variables, separately for $B^+$ and $B^-$ candidate events, with the normalisation (and exponential slope) parameters left free. The $B^+$ and $B^-$ signal yields are extracted from these fits. Figure 1 shows the invariant mass distribution of $B^\pm$ candidates overlaid with a fit to all selected candidates, and including the individual fit components for the signal and backgrounds.
4.2 Flavour tagging methods

The flavour of the signal $B$ meson at production is inferred using several methods, differing in efficiency and discrimination. The measured charge of a lepton (electron or muon) from the semileptonic decay of a $B$ meson provides strong discrimination; however, the $b \to \ell$ transitions are diluted through processes that can change the charge of the observed lepton, such as through neutral $B$ meson oscillations, or through cascade decays $b \to c \to \ell$. The separation power of lepton tagging is enhanced by considering a weighted sum of the charge of the charged-particle tracks in a cone around the lepton, with parameters determined separately for each tagging method based on optimisation of the tagging performance. If no lepton is present, a weighted sum of the charge of the charged-particle tracks in a jet associated with the opposite-side $b$-hadron decay is used to provide discrimination. This weighted sum, or cone charge, is defined as:

$$Q_x = \frac{\sum_i^N |q_i| \cdot (p_{Ti})^x}{\sum_i^N (p_{Ti})^x},$$

where $x = \{\mu, e, \text{jet}\}$ refers to muon, electron, or jet charge, respectively, and the summation is made over a selected set of charged-particle tracks (including the lepton) in a cone, $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, around the lepton or jet direction. The requirements on the charged-particle tracks and $\Delta R$ are described below, dependent on the OST method. Two sub-categories of $Q_x$ are considered, the first called discrete, in the case where only one charged-particle track is used in the cone charge, or more than one charged-particle track of the same charge, resulting in a charge of $Q_x = \pm 1$. The second category is called continuous, where more than one charged-particle track is considered, and the sum contains charged-particle tracks of both negative and positive charge. In the continuous case, $Q_x$ is divided into intervals within the range $-1 < Q_x < 1$ for each OST algorithm.

A probability, $P(B|Q_x)$ ($P(\bar{B}|Q_x)$), is constructed, which is defined as the probability that a $B$ meson is produced in a state containing a $b$-quark ($\bar{b}$-quark) given the value of the cone charge $Q_x$. Using the $B^\pm$ calibration samples, $P(Q_x|B^\pm)$ for each tagging method used can be defined. The probability to tag a $B^\pm_s$ meson as containing a $b$-quark is therefore given as $P(B|Q_x) = P(Q_x|B^\pm)(P(B_s|B^\pm) + P(B_s|B^\mp))$, and correspondingly $P(\bar{B}|Q_x) = 1 - P(B|Q_x)$. If there is no OST information available for a given $B^\pm_s$ meson, a probability of 0.5 is assigned to that candidate.

The OST algorithms used in the analysis are described below, noting that the same algorithms are used for the calibrations using $B^\pm$ mesons, and as applied to $B^\pm_s$ meson candidates to infer the initial flavour.

Muon tagging

For muon-based tagging, at least one additional muon is required in the event, with $p_T > 2.5$ GeV, $|\eta| < 2.5$ and with $|\Delta z|$, the difference in $z$ between the primary vertex and the longitudinal impact parameter of the ID track associated to the muon, less than 5 mm. Muons are classified and kept if their identification quality selection working point is either Tight or Low-$p_T$. The two muon categories are subsequently

---

1 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. [24].

2 This working point is optimized to provide good muon reconstruction efficiency down to a $p_T$ of $\approx 3$ GeV, while controlling the fake rate. It allows $\geq 1$ (at least 2) MDT station tracks up to $|\eta| < 1.3$ (1.3 < $|\eta| < 1.55$) for candidates reconstructed by algorithms utilizing inside-out combined reconstruction [24]. 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 fakes.
treated as distinct flavour tagging methods. For muons with $p_T > 4$ GeV Tight muons are the dominant category, with Low-$p_T$ muons typically identifying muons of $p_T < 4$ GeV. In the case of multiple muons in the event that pass the selections, Tight muons are chosen over Low-$p_T$ muons. Within the same muon category, the muon with the highest transverse momentum that passes the selections is used.

A muon cone charge variable, $Q_\mu$, is constructed, where $\kappa = 1.1$ and the sum is performed over the reconstructed ID charged-particle tracks, including the muon, within a cone, $\Delta R = 0.5$, around the muon direction. These charged-particle tracks must have $p_T > 0.5$ GeV, $|\eta| < 2.5$, and $|\Delta z| < 5$ mm. Charged-particle tracks associated with the decay of $B$ 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 fitted number of signal events extracted. The fit model used is described in Section 4.1. Figures 2 and 3 show the distributions of the muon cone charge using $B^*$ signal candidates, for Tight and Low-$p_T$ muons, respectively, and including the tag probability as a function of the cone charge variable.

![Figure 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). Superimposed is the distribution of the tag-probability, $P(B|Q_\mu)$, as a function of the cone charge, derived from a data sample of $B^\pm \rightarrow J/\psi K^\pm$, and defined as the probability to have a $B^*$ meson (on the signal-side) given a particular cone charge $Q_\mu$. The fitted parametrization, shown in black, is used as the calibration curve to infer the probability to have a $B_s^0$ or $\bar{B}_s^0$ meson at production in the decays to $J/\psi \phi$.](image)

**Electron tagging**

Studies of parameter distributions of the electron tagging are performed using sideband subtraction on $m(J/\psi K^\pm)$. Electrons are identified using inner detector and calorimeter information, which satisfy the *Medium* electron quality criteria [25]. The inner detector 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 opening angle between the $B$ meson candidate and electron momenta, $\zeta_b$, of $\cos(\zeta_b) > 0.93$ are not considered. In the case of more than one electron passing the selection, the electron with the highest transverse momentum is chosen. As in the case of muon tagging, additional charged-particle tracks within a cone of size $\Delta R = 0.5$ are used to form the electron cone charge $Q_e$ with $\kappa = 1.0$. If there are no additional charged-particle tracks within the cone, only the charge of the electron is used. The resulting electron cone charge distributions are shown in Figure 4, together with the corresponding tag probability.

![Figure 3: 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). Superimposed is the distribution of the tag-probability, $P(B|Q_\mu)$, as a function of the cone charge, derived from a data sample of $B^\pm \rightarrow J/\psi K^\pm$, and defined as the probability to have a $B^*$ meson (on the signal-side) given a particular cone charge $Q_\mu$. The fitted parametrization, shown in black, is used as the calibration curve to infer the probability to have a $B_s^0$ or $\bar{B}_s^0$ meson at production in the decays to $J/\psi \phi$.](image)
In the absence of a muon or electron, a jet identified as containing a $b$-hadron is required. Jets are reconstructed from calorimetric information using the anti-$k_t$ algorithm [26, 27] with a radius parameter $R = 0.4$. The identification of a $b$-tagged jet uses a multivariate algorithm $MV2c10$ [28], utilising boosted decision trees (BDT), which output a classifier value. Jets are selected that exceed the BDT classifier output value of 0.4. This value is optimized to maximize the tagging power of the calibration sample. In the case of multiple selected jets, the jet with the highest value of the multivariate output classifier is used. Jets associated to the signal decay are not considered in this selection.

Charged-particle tracks within a cone of size $\Delta R = 0.5$ of the jet axis are used to define a jet cone charge, $Q_{\text{jet}}$, where $\kappa = 1.1$ and the sum is over the charged-particle tracks associated with the jet, with $|\Delta \eta| < 5$ mm, and excluding charged-particle tracks from the decay of the signal $B$ meson candidate. The signal yields are extracted from the data from fits to the $J/\psi K^\pm$ invariant mass spectrum, using the same

**Jet tagging**

Figure 3: Normalised cone charge distributions, $Q_\mu$, 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 tag-probability, $P(B|Q_\mu)$.

Figure 4: Normalised cone charge distributions, $Q_e$, 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 tag probability, $P(B|Q_e)$. 
procedure described for muon tagging. Figure 5 shows the distribution of the opposite side jet cone charge for $B^\pm$ signal candidates.

**Figure 5**: Normalised cone charge distributions, $Q_{\text{jet}}$, for $B^+$ ($B^-$) events shown in red (blue) for jets, for cases of discrete charge (left), and the continuous distribution (right). Superimposed is the distribution of the tag probability, $P(B|Q_{\text{jet}})$.

### 4.3 Flavour tagging performance

It is possible to define a quantity that represents the strength of a particular flavour tagging method, called the dilution $D(Q_x) = 2P(B|Q_x) - 1$. The efficiency, $\epsilon_x$, of an individual tagging method is defined as the fraction of signal events tagged by that method compared to the total number of signal events in the sample. The tagging power of a particular tagging method is then defined as $T_x = \sum_i \epsilon_{x,i} \cdot (2P(B|Q_{x,i}) - 1)^2$, 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 both a muon (either Tight or Low-$p_T$) and electron, corresponds to around 0.6% of all tagged events. In the case of multiply-tagged events, the OST method is selected in 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 $B_s^0$ fit

For the maximum likelihood fit performed on the $B_s^0$ data, and described in detail in Section 5, the per-candidate probability, $P(B|Q_x)$, that the $B$ meson candidate was produced in a state $B_s^0$ is provided using the calibrations derived from the $B^\pm \rightarrow J/\psi K^\mp$ sample, described above, and shown in Figures 2–5. As the distributions of $P(B|Q_x)$ can be expected to be different between signal $B_s^0$ mesons and background, additional PDFs are necessary to describe these distributions in the likelihood function. These PDFs are defined as $P_s(P(B|Q_x))$ and $P_b(P(B|Q_x))$, describing the probability distributions for signal and background, respectively, and are extracted using the sample of $B_s^0$ candidates. The PDFs $P_{s,b}(P(B|Q_x))$
Table 1: Summary of tagging performance for the different flavour tagging methods on the sample of $B^*$ signal candidates, as described in the text. Uncertainties shown are statistical only. The efficiency and tagging power are each determined by summing over the individual bins of the cone charge distribution. The effective dilution 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>Efficiency [%]</th>
<th>Effective Dilution [%]</th>
<th>Tagging Power [%]</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>5.54 ± 0.01</td>
<td>20.4 ± 0.1</td>
<td>0.231 ± 0.005</td>
</tr>
<tr>
<td>Total</td>
<td>14.74 ± 0.02</td>
<td>33.4 ± 0.1</td>
<td>1.65 ± 0.01</td>
</tr>
</tbody>
</table>

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.

For the continuous distributions the parametrizations for each OST method are defined as follows. For events in the $m(J/\psi KK)$ sideband region: $5.150 < m(J/\psi KK) < 5.317$ GeV and $5.417 < m(J/\psi KK) < 5.650$ GeV, unbinned maximum likelihood fits are performed to $P(B|Q_s)$ distributions (for each OST method) to extract the background continuous category PDFs for $P_b(P(B|Q_s))$. For the Tight muon and electron methods, the parametrization has the form of the sum of a second-order polynomial and two exponential functions. A 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 $P(B|Q_s)$ distributions, using all events in the $m(J/\psi KK)$ distributions to extract the continuous category signal PDFs for $P_s(P(B|Q_s))$. In these fits, the previously extracted parameters to describe the background continuous PDFs are fixed, as is the relative normalisation of signal and background, extracted from a fit to the $m(J/\psi KK)$ 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 electrons, the signal function has the form of the sum of a second-order polynomial and two exponential functions, and for Low-$p_T$ muon and jet tagging methods a Gaussian function is used.

In the case where the cone charge is discrete, the fractions of events $f_{+1}$ ($f_{-1}$) with cone charges +1 (-1) are determined separately for signal and background using events from the same $B^0_s$ mass signal and sideband regions. The remaining fraction of events, $1 - f_{+1} - f_{-1}$, constitutes the continuous parts of the distributions. 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 summarizes for the different tag methods, the fractions $f_{+1}$ and $f_{-1}$ obtained for signal and background events.

The relative fractions of signal and background events tagged using the different OST methods are found using a similar sideband-subtraction method, and are summarized in Table 3. To account for possible deviations between data and the selected fit models, variations of the procedure are used to determine systematic uncertainties, as described in Section 6.
Table 2: Fraction of events $f_{+1}$ and $f_{-1}$ with cone charges of +1 and −1, respectively, for signal and background events and for the different tag methods. The remaining fraction, $1 - f_{+1} - f_{-1}$, is the fraction of events from the continuous part of the distributions, and not explicitly shown in the table. Only statistical uncertainties are quoted.

<table>
<thead>
<tr>
<th>Tag method</th>
<th>Signal $f_{+1}$ [%]</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>
</tr>
<tr>
<td></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>
</tr>
<tr>
<td></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>11.7 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>7.0 ± 0.1</td>
<td>7.6 ± 0.1</td>
</tr>
<tr>
<td>Jet</td>
<td>4.51 ± 0.15</td>
<td>4.58 ± 0.16</td>
</tr>
<tr>
<td></td>
<td>3.76 ± 0.03</td>
<td>3.86 ± 0.03</td>
</tr>
</tbody>
</table>

Table 3: Relative fractions of signal and background events tagged using the different tag 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>4.00 ± 0.06</td>
<td>3.16 ± 0.01</td>
</tr>
<tr>
<td>Electron</td>
<td>1.87 ± 0.04</td>
<td>1.48 ± 0.01</td>
</tr>
<tr>
<td>Low-$p_T$ muon</td>
<td>2.91 ± 0.05</td>
<td>2.64 ± 0.01</td>
</tr>
<tr>
<td>Jet</td>
<td>14.4 ± 0.1</td>
<td>11.96 ± 0.02</td>
</tr>
<tr>
<td>Un tagged</td>
<td>76.7 ± 0.3</td>
<td>80.77 ± 0.05</td>
</tr>
</tbody>
</table>

5 Maximum likelihood fit

An unbinned maximum-likelihood fit is performed on the selected events to extract the parameter values of the $B^0_s \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ decay. The fit uses information about the reconstructed mass $m$, the measured proper decay time $t$, the measured proper decay time uncertainty $\sigma_t$, the tagging probability, and the transversity angles $\Omega$ of each $B_s^0 \rightarrow J/\psi\phi$ decay candidate. The measured proper decay time uncertainty $\sigma_t$ is calculated from the covariance matrix associated with the vertex fit of each candidate event. The transversity angles $\Omega = (\theta_T, \psi_T, \phi_T)$ are defined in Section 5.1. The likelihood is independent of the $K^+K^-$ mass distribution. The likelihood function is defined as a combination of the signal and background probability density functions as follows:

$$
\ln L = \sum_{i=1}^{N} w_i \cdot \ln \left[ f_s \cdot \mathcal{F}_s(m_i, t_i, \sigma_{t_i}, \sigma_{t}, \Omega_i, P_t(B|Q_i), p_{T_t}) + f_{b^0} \cdot \mathcal{F}_{b^0}(m_i, t_i, \sigma_{t_i}, \sigma_{t}, \Omega_i, P_t(B|Q_i), p_{T_t}) + f_{\Lambda_b} \cdot \mathcal{F}_{\Lambda_b}(m_i, t_i, \sigma_{t_i}, \sigma_{t}, \Omega_i, P_t(B|Q_i), p_{T_t}) + (1 - f_s \cdot (1 + f_{b^0} + f_{\Lambda_b})) \mathcal{F}_{b^k g}(m_i, t_i, \sigma_{t_i}, \sigma_{t}, \Omega_i, P_t(B|Q_i), p_{T_t}) \right],
$$

where $N$ is the number of selected candidates, $w_i$ is a weighting factor to account for the trigger efficiency (described in Section 5.3), $\mathcal{F}_s$, $\mathcal{F}_{b^0}$, $\mathcal{F}_{\Lambda_b}$ and $\mathcal{F}_{b^k g}$ are the probability density functions (PDF) 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 $f_{b^0}$ and $f_{\Lambda_b}$ are the background fractions of $B^0$ mesons.
and \(\Lambda_b\) baryons misidentified as \(B_s^0\) candidates calculated relative to the number of signal events. These background fractions are fixed to their expectation from the MC simulation and varied as part of the systematic uncertainties. The mass \(m_i\), the proper decay time \(t_i\) 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. (1) is given in Section 5.1. The three background functions are described in Section 5.2.

5.1 Signal PDF

The PDF used to describe the signal events, \(\mathcal{F}_s\), has the following composition:

\[
\mathcal{F}_s(m_i, t_i, \sigma_{m_i}, \sigma_{t_i}, \Omega_i, P_1(B|Q_s), p_{T_1}) = P_s(m_i, \sigma_{m_i}) \cdot P_s(\sigma_{t_i}) \cdot P_s(\Omega_i, t_i, P_1(B|Q_s), \sigma_{t_i}) \cdot P_s(p_{T_1}) \cdot P_s(\Omega_i, p_{T_1}) \cdot P_s(p_{T_1}).
\]

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

\[
P_s(m_i, \sigma_{m_i}) \equiv \frac{1}{\sqrt{2\pi}S_m\sigma_{m_i}} e^{-\frac{(m_i - m_{B_s})^2}{2S_m\sigma_{m_i}^2}}.
\]

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 4-track vertex fit. Each measured candidate mass is convoluted by 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 \(m_{B_s}\), which is the \(B_s^0\) meson mass, are free parameters determined in the fit.

The probability terms \(P_s(\sigma_{m_i}), P_s(\sigma_{t_i})\) and \(P_s(p_{T_1})\) are introduced to account for differences between signal and background events for the values of the per-candidate mass error, time error and \(p_{T_1}\) values, respectively. Distributions of these variables for signal and background are described by gamma functions and the method is unchanged from the analysis explained in Ref. [29]. The tagging probability term for signal \(P_s(P_1(B|Q_s))\) is described in Section 4.4.

The term \(P_s(\Omega_i, t_i, P_1(B|Q_s), \sigma_{t_i})\) is a joint PDF for the decay time \(t\) and the transversity angles \(\Omega\) for the \(B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-)\) decay. Ignoring detector effects, the distribution for the time \(t\) and the angles \(\Omega\) is given by the differential decay rate [30]:

\[
\frac{d^4\Gamma}{dt d\Omega} = \sum_{k=1}^{10} O^{(k)}(t) g^{(k)}(\theta_T, \phi_T, \phi_T),
\]

where \(O^{(k)}(t)\) are the time-dependent functions corresponding to the contributions of the four different amplitudes \((A_0, A_||, A_\perp, A_S)\) and their interference terms, and \(g^{(k)}(\theta_T, \phi_T, \phi_T)\) 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 \(\bar{B}_s^0\) but with a sign reversal in the terms containing \(\Delta m_s\), which is a fixed parameter of the fit (using Ref. [23]). The formalism used throughout this analysis assumes no direct CP violation. In Table 4, the parameter \(A_{\perp}(t)\) is the time-dependent amplitude for the \(CP\)-odd final-state configuration while \(A_0(t)\) and \(A_||(t)\) correspond to \(CP\)-even final-state configurations. The amplitude \(A_S(t)\) gives the contribution from the \(CP\)-odd non-resonant \(B_s^0 \rightarrow J/\psi K^+K^-\) \(S\)-wave state (which includes the \(f_0\)). The corresponding functions are given in the last four lines of Table 4 \((k = 7\text{--}10)\). The amplitudes are parameterized by \(|A_i|e^{i\delta_i}\), where
\[ i = \{0, ||, \perp, S\}, \text{with } \delta_0 = 0 \text{ and are normalized such that } |A_0(0)|^2 + |A_{\perp}(0)|^2 + |A_{\parallel}(0)|^2 = 1. |A_{\perp}(0)| \text{ is determined according to this condition, while the remaining three amplitudes are parameters of the fit. The phase } \delta_S \text{ is the phase difference between } A_S(0) \text{ and } A_0(0) \text{ at the resonance peak. } |A_S|^2 \text{ gives the ratio of non-resonant over resonant yield in the interval of } m(K^+K^-) \text{ 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 = 0.51 \pm 0.02 \text{ that takes into account the mass–dependent differences in absolute amplitude and phase between the resonant and the } S-\text{wave amplitudes. The correction is based on the Breit-Wigner description of the resonance and on the assumption of uniform } A_S. \text{ The uncertainty on the value of } \alpha \text{ has been calculated based on the Flatté parameterization [31] and the corresponding systematic uncertainty is explained in Section 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 PDF term \(P_b(\Omega_t, t_i, B|Q_x, \sigma_{t_i})\) takes into account the lifetime resolution, so each time element in Table 4 is smeared with a Gaussian function. This smearing is performed numerically on an event-by-event basis where the width of the Gaussian function is the proper decay time uncertainty, measured for each event, multiplied by a scale factor to account for any mismeasurements. The average value of this uncertainty for signal events is 69 fs.

The angular acceptance of the detector and kinematic cuts on the angular distributions are included in the likelihood function through \(A(\Omega_t, p_{T_i})\). This is calculated using a 4D 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 \(B^0_s\) candidate. The acceptance is calculated from the \(B^0_s \rightarrow J/\psi \phi\) MC events with additional weighting for \(p_T\) and \(\eta\) 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 \(B^0_s \rightarrow J/\psi (\mu^+ \mu^-) \phi(K^+K^-)\) decays. As both the acceptance and time- and angle-dependent decay PDFs depend on the transversity angles they must be normalized together. This normalization is done numerically during the likelihood fit. The PDF is normalized over the entire \(B^0_s\) mass range 5.150–5.650 GeV.

### 5.2 Background PDF

The background PDF has the following composition:

\[
T_{\text{bkg}}(m_i, t_i, \sigma_{t_i}, \Omega_t, P_b(B|Q_x), p_{T_i}) = P_b(m_i) \cdot P_b(t_i|\sigma_{t_i}) \cdot P_b(P_b(B|Q_x)) \cdot P_b(\Omega_t) \cdot P_b(\sigma_{t_i}) \cdot P_b(p_{T_i}).
\]

The proper decay time function \(P_b(t_i|\sigma_{t_i})\) is parametrized as a prompt peak modelled by a Gaussian distribution, two positive exponential functions and a negative exponential function. These functions are smeared with the same resolution function as the signal decay time-dependence. The prompt peak models the combinatorial background events, which are expected to have reconstructed lifetimes distributed...
Table 4: The ten time-dependent functions, $O^{(k)}(t)$ and the functions of the transversity angles $g^{(k)}(\theta_T, \phi_T, \phi_T)$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$O^{(k)}(t)$</th>
<th>$g^{(k)}(\theta_T, \psi_T, \phi_T)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{1}{2}</td>
<td>A_0(0)</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{2}</td>
<td>A_1(0)</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{1}{2}</td>
<td>A_{\perp}(0)</td>
</tr>
<tr>
<td>4</td>
<td>$\frac{1}{2}</td>
<td>A_0(0)</td>
</tr>
<tr>
<td>5</td>
<td>$</td>
<td>A_0(0)</td>
</tr>
<tr>
<td>6</td>
<td>$</td>
<td>A_0(0)</td>
</tr>
<tr>
<td>7</td>
<td>$\frac{1}{2}</td>
<td>A_0(0)</td>
</tr>
<tr>
<td>8</td>
<td>$\alpha</td>
<td>A_0(0)</td>
</tr>
<tr>
<td>9</td>
<td>$\frac{1}{2}\alpha</td>
<td>A_0(0)</td>
</tr>
<tr>
<td>10</td>
<td>$\alpha</td>
<td>A_0(0)</td>
</tr>
</tbody>
</table>
around zero. The two positive 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 negative exponential function takes into account events with poor vertex resolution. The probability terms $P_b(\sigma_i)$ and $P_b(p_T)$ are described by gamma functions. They are unchanged from the analysis described in Ref. [29] and explained in detail there. The tagging probability term for background events $P_b(P_t(B|Q_s))$ is described in Section 4.4.

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

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

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

$$P_b(\theta_T, \psi_T, \phi_T) = \sum_{k=0}^{14} \sum_{|l|=0}^{14} \sum_{m=-l}^{l} \left\{ \begin{array}{ll}
  a_{k,l,m} \sqrt{2} Y_l^m(\theta_T) \cos(m\phi_T) P_k(\cos \psi_T) & \text{where } m > 0 \\
  a_{k,l,m} \sqrt{2} Y_l^{-m}(\theta_T) \sin(m\phi_T) P_k(\cos \psi_T) & \text{where } m < 0 \\
  a_{k,l,m} \sqrt{2} Y_l^0(\theta_T) P_k(\cos \psi_T) & \text{where } m = 0
\end{array} \right.$$

where the coefficients $a_{k,l,m}$ are adjusted to give the best fit to the angular distributions for events in the $B_s^0$ mass sidebands. These parameters are then fixed in the main fit, defined by Eq. (1). The $B_s^0$ mass interval used for the background fit is between 5.150 and 5.650 GeV excluding the signal mass region $|m(B_s^0) - 5.366| < 0.110$ GeV. Higher order Legendre polynomial functions were tested as a systematic check, described in Section 6. The background mass model, $P_b(m_i)$ is an exponential function with a constant term added.

Contamination from $B_d \to J/\psi K^{0*}$ and $\Lambda_b \to J/\psi pK^-$ events mis-reconstructed as $B_s^0 \to J/\psi \phi$ are accounted for in the fit through the $\mathcal{F}_{B_s^0}$ and $\mathcal{F}_{\Lambda_b}$ terms in the PDF function described in Eq. (1). 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 $B_s^0 \to J/\psi \phi$ signal events and are evaluated from MC simulation using production and branching fractions from Refs. [23, 32–36]. MC simulated events are also used to determine the shape of the mass and transversity angle distributions. The 3D angular distributions of $B_s^0 \to J/\psi K^{0*}$ and of the conjugate decay are modelled using input from Ref. [37], while angular distributions for $\Lambda_b \to J/\psi pK^-$ and the conjugate decay are considered flat. These distributions are sculpted for detector acceptance effects and then described by Legendre polynomial functions, Eq. (3). These shapes are fixed 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 $B_d/B_s^0$ or $\Lambda_b/B_s^0$ masses as appropriate, where the lifetimes and masses are taken from Ref. [23]. Systematic uncertainties due to the background from $B_d \to J/\psi K^{0*}$ and $\Lambda_b \to J/\psi pK^-$ decays are described in Section 6. The contribution of the $S$–wave $B_d \to J/\psi K\pi$ decays as well as their interference with the $P$–wave $B_d \to J/\psi K^{0*}$ decays are included in the PDF of the fit, using the parameters measured in Ref. [37].

### 5.3 Muon trigger proper decay time-dependent efficiency

A limited acceptance of the trigger muon tracking at high values of transverse impact parameters of muons, results in inefficiency at large values of the proper decay time. This inefficiency is measured using MC simulated events, by comparing the $B_s^0$ 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 $w$:

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

(4)
where $p_0$, $p_1$, $p_2$ and $p_3$ are parameters determined in the fit to MC events. No significant bias or inefficiency due to offline track reconstruction, vertex reconstruction, or track quality selection criteria is observed.

### 6 Systematic uncertainties

Systematic uncertainties are evaluated for effects that are not accounted for in the likelihood fit. These are described below.

- **Flavour tagging**: There are two contributions to the uncertainties in the fit parameters due to the flavour tagging procedure, the statistical and systematic components. The statistical uncertainty due to the size of the sample of $B^+ \rightarrow J/\psi K^+$ decays is included in the overall statistical error. The systematic uncertainty arising from the precision of the OST calibration, described in Section 4.2, is estimated by changing the models used to parametrize the probability distribution, $P(B|Q_x)$, as a function of the cone charge from the third-order polynomial (or sinusoidal for electrons) function used by default to one of several alternative functions. The alternative functions are: a linear function; a fifth-order polynomial; or two third-order polynomials, which describe the positive and negative regions and have common constant and linear terms, but independent quadratic and cubic terms. The $B_s^0$ fit is repeated using the alternative models and the largest difference with respect to the nominal fit is assigned as the systematic uncertainty. To verify the calibration procedure, calibration curves are derived from simulated samples of $B^+$ and $B^0_s$ signals. The variations between the curves from these two samples are propagated to the calibration curves derived from data. The difference in the fitted parameters between the nominal fit, and the fits performed with the variations of calibration curves are included in the systematic uncertainty. An additional systematic uncertainty is assigned due to any dependency on the pile-up distribution. The calibration data are split into subsets of approximately equal yields, separated according to the pile-up profile 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 pile-up profile 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 $P_0(P(B|Q_x))$ and $P_0(P(B|Q_x))$, variations of the parametrisation are considered, including using histograms, rather than a parametrisation. The resulting changes in the parameters of the $B^0_s$ fit are similarly included in the systematic uncertainties.

- **Angular acceptance method**: The angular acceptance (arising from the detector sculpting and kinematic selection cuts, described in Section 5.1) is calculated from a binned fit to MC simulated data. In order to estimate the size of the systematic uncertainty introduced from the choice of binning, different acceptance functions are calculated using different bin widths and central values.

- **Inner detector alignment**: Residual misalignments of the ID affect the impact parameter, $d_0$, distribution with respect to the primary vertex. The effects on the fit parameters have been studied and observed deviations are included in the systematics uncertainties.

- **Trigger efficiency**: To correct for the proper decay time dependent inefficiencies due to the triggers, the events are re-weighted according to Eq. (4). An alternative fit has been performed using different sets of binning in the MC sample that was used to determine the efficiency. The systematic effects are found to be negligible.
• **S-wave phase:** As explained in Section 5.1, the model for the interference between the $B_s^0 \rightarrow J/\psi (K^+K^-)$ and the $S$–wave $B_s^0 \rightarrow J/\psi K^+K^-$ is corrected by a factor $\alpha = 0.51 \pm 0.02$ to account for the mass–dependent differences in absolute amplitude and phase between the resonant and $S$–wave amplitudes. To account for uncertainty of $\alpha$, the fit was repeated with $\alpha = 0.51 \pm 0.02$ and $\alpha = 0.51 - 0.02$ values. The variations of the parameter values relative to those from the default fit using the central value $\alpha = 0.51$ are included in the systematic uncertainties.

• **Background angles model:** The shape of the background angular distribution, $P_b(\theta_T, \varphi_T, \psi_T)$, is described by the Legendre polynomial functions of 14th degree, given in Eq. (3). Alternatively, higher order Legendre polynomial functions were tested, and the differences in fit parameters 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 $B^0_s$ meson candidate. For this reason, the parameterization using the Legendre polynomial functions is performed in six $p_T$ intervals: 10–15 GeV, 15–20 GeV, 20–25 GeV, 25–30 GeV, 30–55 GeV and >55 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 biggest deviations observed in the fit results were taken to represent the systematic uncertainties.

The parameters of the Legendre polynomial functions given in Eq. (3) are adjusted to give the best fit to the angular distributions for events in the $B^0_s$ 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 intervals: $|m(B^0_s) - 5.366\text{ GeV}| > 0.085\text{ GeV}$ and $|m(B^0_s) - 5.366\text{ GeV}| > 0.160\text{ GeV}$ (instead of $|m(B^0_s) - 5.366\text{ GeV}| > 0.110\text{ GeV}$). The differences in the fit results are assigned as systematic uncertainties.

• **$B_d$ contribution:** The contamination from $B_d \rightarrow J/\psi K^{0*}$ events mis-reconstructed as $B^0_s \rightarrow J/\psi \phi$ is accounted for in the final fit. Studies are performed to evaluate the effect of the uncertainties in the $B_d \rightarrow J/\psi K^{0*}$ fraction and the shapes of the distributions of the mass, transversity angles, and lifetime. In the MC events the angular distribution of the $B_d \rightarrow J/\psi K^{0*}$ decay is modelled using parameters taken from Ref. [37]. The contribution of the $S$–wave $B_d \rightarrow J/\psi K\pi$ decays as well as its interference with the $P$–wave $B_d \rightarrow J/\psi K^{0*}$ decays are also included in the PDF of the fit, following the parameters measured in Ref. [37]. The uncertainties of these parameters are taken into account in the estimation of 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 within the systematic uncertainty.

• **$\Lambda_b$ contribution:** The contamination from $\Lambda_b \rightarrow J/\psi pK^-$ events mis-reconstructed as $B^0_s \rightarrow J/\psi \phi$ is accounted for in the final fit. Studies are performed to evaluate the effect of the uncertainties in the $\Lambda_b \rightarrow J/\psi pK^-$ 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.

• **Signal fit model:** To estimate the systematic uncertainties due to the signal $B^0_s$ mass model, the default model has been altered by adding a second Gaussian function in Eq. (2), which has the same structure as the first Gaussian but a different scale factor, $S_{m_i}^1$, which is an additional free parameter of the fit. Respective changes in fit parameters are found negligible. To test the sensitivity of the part of the fit model describing the lifetime, the following two systematic tests have been
performed. The determination of signal and background lifetime errors is sensitive to the choice of $p_T$ bins, in which the relative contributions of these two components are evaluated. To estimate the systematic uncertainty, the fit has been repeated varying the intervals of the default $p_T$ binning.

The determination of signal and background lifetime errors is also sensitive to the determination of the signal fraction. The fit has been repeated by varying this fraction within one standard deviation of its uncertainty and differences are included in the systematic uncertainty.

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.

### Table 5: Summary of systematic uncertainties assigned to the physical parameters of interest.

|                      | $\phi_s$   | $\Delta \Gamma_s$ | $\Gamma_s$ | $|A_0(0)|^2$ | $|A_{ll}(0)|^2$ | $\delta_{\perp}$ | $\delta_{||}$ | $\delta_{\perp} - \delta_S$ |
|----------------------|------------|--------------------|-----------|-------------|-----------------|----------------|-------------|-----------------|
| Tagging              | $1.7 \times 10^{-2}$ | $0.4 \times 10^{-3}$ | $0.3 \times 10^{-3}$ | $0.2 \times 10^{-3}$ | $2.3 \times 10^{-3}$ | $1.9 \times 10^{-2}$ | $2.2 \times 10^{-2}$ | $2.2 \times 10^{-3}$ |
| Acceptance           | $0.7 \times 10^{-3}$ | $<10^{-4}$ | $<10^{-4}$ | $0.8 \times 10^{-3}$ | $0.7 \times 10^{-3}$ | $2.4 \times 10^{-3}$ | $3.3 \times 10^{-2}$ | $1.4 \times 10^{-2}$ | $2.6 \times 10^{-3}$ |
| ID alignment         | $0.7 \times 10^{-3}$ | $0.1 \times 10^{-3}$ | $0.5 \times 10^{-3}$ | $<10^{-4}$ | $<10^{-4}$ | $<10^{-4}$ | $1.0 \times 10^{-2}$ | $7.2 \times 10^{-3}$ | $<10^{-4}$ |
| $S$–wave phase       | $0.2 \times 10^{-3}$ | $<10^{-4}$ | $<10^{-4}$ | $0.3 \times 10^{-3}$ | $<10^{-4}$ | $0.3 \times 10^{-3}$ | $1.1 \times 10^{-2}$ | $2.1 \times 10^{-2}$ | $8.3 \times 10^{-3}$ |
| Background angles model: |                  |                  |               |              |                   |                  |                  |                  |
| Choice of fit function | $1.8 \times 10^{-3}$ | $0.8 \times 10^{-3}$ | $<10^{-4}$ | $1.4 \times 10^{-3}$ | $0.7 \times 10^{-3}$ | $0.2 \times 10^{-3}$ | $8.5 \times 10^{-2}$ | $1.9 \times 10^{-1}$ | $1.8 \times 10^{-3}$ |
| Choice of $p_T$ bins | $1.3 \times 10^{-3}$ | $0.5 \times 10^{-3}$ | $<10^{-4}$ | $0.4 \times 10^{-3}$ | $0.5 \times 10^{-3}$ | $1.2 \times 10^{-3}$ | $1.5 \times 10^{-3}$ | $7.2 \times 10^{-3}$ | $1.0 \times 10^{-3}$ |
| Choice of mass interval | $0.4 \times 10^{-3}$ | $0.1 \times 10^{-3}$ | $0.1 \times 10^{-3}$ | $0.3 \times 10^{-3}$ | $0.3 \times 10^{-3}$ | $1.3 \times 10^{-3}$ | $4.4 \times 10^{-3}$ | $7.4 \times 10^{-3}$ | $2.3 \times 10^{-3}$ |
| Dedicated backgrounds: |            |                  |               |              |                   |                  |                  |                  |
| $B_s^\phi$           | $2.3 \times 10^{-3}$ | $1.1 \times 10^{-3}$ | $<10^{-4}$ | $0.2 \times 10^{-3}$ | $3.1 \times 10^{-3}$ | $1.4 \times 10^{-3}$ | $1.0 \times 10^{-2}$ | $2.3 \times 10^{-2}$ | $2.1 \times 10^{-3}$ |
| $\Lambda_{Bs}$       | $1.6 \times 10^{-3}$ | $0.4 \times 10^{-3}$ | $0.2 \times 10^{-3}$ | $0.5 \times 10^{-3}$ | $1.2 \times 10^{-3}$ | $1.8 \times 10^{-3}$ | $1.4 \times 10^{-2}$ | $2.9 \times 10^{-2}$ | $0.8 \times 10^{-3}$ |
| Fit model:           |                  |                  |               |              |                   |                  |                  |                  |
| Time res. sig frac   | $1.4 \times 10^{-3}$ | $1.1 \times 10^{-3}$ | $<10^{-4}$ | $0.5 \times 10^{-3}$ | $0.6 \times 10^{-3}$ | $0.6 \times 10^{-3}$ | $1.2 \times 10^{-2}$ | $3.0 \times 10^{-2}$ | $0.4 \times 10^{-3}$ |
| Time res. $p_T$ bins | $3.3 \times 10^{-3}$ | $1.4 \times 10^{-3}$ | $0.1 \times 10^{-2}$ | $<10^{-4}$ | $<10^{-4}$ | $0.5 \times 10^{-3}$ | $6.2 \times 10^{-3}$ | $5.2 \times 10^{-3}$ | $1.1 \times 10^{-3}$ |
| **Total**            | $1.8 \times 10^{-2}$ | $0.2 \times 10^{-2}$ | $0.1 \times 10^{-2}$ | $0.2 \times 10^{-2}$ | $0.4 \times 10^{-2}$ | $0.4 \times 10^{-2}$ | $9.7 \times 10^{-2}$ | $2.0 \times 10^{-1}$ | $0.1 \times 10^{-1}$ |

### 7 Results

The full simultaneous unbinned maximum-likelihood fit contains nine physical parameters: $\Delta \Gamma_s$, $\phi_s$, $\Gamma_s$, $|A_0(0)|^2$, $|A_{ll}(0)|^2$, $\delta_{||}$, $|A_S(0)|^2$ and $\delta_S$. The other parameters in the likelihood function are the $B_s^\phi$ signal fraction $f_s$, parameters describing the $J/\psi \phi$ mass distribution, parameters describing the decay time plus angular distributions of background events, parameters used to describe the estimated decay time uncertainty distributions for signal and background events, and scale factors between the estimated decay time uncertainties and their true uncertainties. In addition there are also nuisance parameters describing the background and acceptance functions that are fixed at the time of the fit.

Multiplying the total number of events supplied to the fit with the extracted signal fraction and its statistical uncertainty provides an estimate for the total number of $B_s^\phi$ meson candidates of $477 \pm 760$. The results and correlations of the physics parameters obtained from the fit are given in Tables 6 and 7. Fit projections of the mass, proper decay time and angles are given in Figures 6 and 7, respectively.
Table 6: Fitted values for the physical parameters of interest with their statistical and systematic uncertainties.

<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>-0.068</td>
<td>0.038</td>
<td>0.018</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>0.067</td>
<td>0.005</td>
<td>0.002</td>
</tr>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>0.669</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$</td>
<td>A_{\parallel}(0)</td>
<td>^2$</td>
<td>0.219</td>
</tr>
<tr>
<td>$</td>
<td>A_{0}(0)</td>
<td>^2$</td>
<td>0.517</td>
</tr>
<tr>
<td>$</td>
<td>A_{S}(0)</td>
<td>^2$</td>
<td>0.046</td>
</tr>
<tr>
<td>$\delta_\perp$ [rad]</td>
<td>2.946</td>
<td>0.101</td>
<td>0.097</td>
</tr>
<tr>
<td>$\delta_\parallel$ [rad]</td>
<td>3.267</td>
<td>0.082</td>
<td>0.201</td>
</tr>
<tr>
<td>$\delta_\perp - \delta_S$ [rad]</td>
<td>-0.220</td>
<td>0.037</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Table 7: Fit correlations between the physical parameters of interest.

| $\phi_s$ | $\Delta \Gamma_s$ | $|A_{\parallel}(0)|^2$ | $|A_{0}(0)|^2$ | $|A_{S}(0)|^2$ | $\delta_\parallel$ | $\delta_\perp$ | $\delta_\perp - \delta_S$ |
|----------|------------------|-----------------|-----------------|-----------------|----------------|----------------|------------------|
| -0.111   | 1                | -0.008          | -0.008          | -0.015          | 0.019          | -0.001         | -0.011           |
| 0.038    | -0.563           | 0.092           | 0.097           | 0.042           | 0.036          | 0.011          | 0.009            |
| 0.02       | -0.139           | -0.040          | 0.103           | -0.105          | -0.041         | 0.016          |                  |
| 0.02       | -0.349           | -0.216          | 0.571           | 0.223           | -0.035         |                  |                  |
| 0.02       | 1                | 0.299           | -0.129          | -0.056          | 0.051          |                  |                  |
| 0.02       | 1                | -0.408          | -0.175          | 0.164           | 0.051          |                  |                  |
| 0.02       | 1                | 0.392           | -0.041          |                  |                  |                  |                  |
| 0.02       | 1                | 0.052           |                  |                  |                  |                  |                  |

8 Combination with 7 TeV and 8 TeV results

The measured values are consistent with those obtained in a previous analysis [9], using ATLAS 19.2 fb$^{-1}$ of data collected at $\sqrt{s}$ of 7 TeV and 8 TeV. A Best Linear Unbiased Estimator (BLUE) combination [38] is used to perform a combination of the current measurements with those from 19.2 fb$^{-1}$ of 7 TeV and 8 TeV data. The measured values, uncertainties and correlations are taken from the measurements performed at each centre-of-mass energy. The statistical correlation between these three measurements is zero as the events are different. The correlations of the systematic uncertainties between the three measurements are estimated and tested in several categories depending of whether the given systematic changed significantly between the measurements. The systematic uncertainties such as the background angles model and the specific $\Lambda_b$ and $B_d$ backgrounds are treated as 100% correlated. For the ID alignment, the fit model, the acceptance and the tagging systematic uncertainty, the estimated value of the correlation is 50%. The obtained combined results for the fit parameters and their uncertainties are given in Table 8.

The two dimensional likelihood contours in the $\phi_s - \Delta \Gamma_s$ plane for the ATLAS results based on 7 TeV and 8 TeV data, the result from 13 TeV, and the combined results from 7 TeV, 8 TeV and 13 TeV are shown in Figure 8. The statistical and systematic uncertainties are combined in quadrature and correlations are taken into account in the construction of Gaussian contours. The correlation between the $\phi_s$ and $\Delta \Gamma_s$ values determined in combination is $-0.07$. 19
Figure 6: (Left) Mass fit projection for the $B_s^0 \rightarrow J/\psi \phi$ sample. The red line shows the total fit, the dashed magenta line shows the $B_s^0 \rightarrow J/\psi \phi$ signal component, the blue line shows the $B_s^0 \rightarrow J/\psi K^0$ component, while the green line shows the contribution from $\Lambda_b \rightarrow J/\psi p K^-$ events. (Right) Proper decay time fit projection for the $B_s^0 \rightarrow J/\psi \phi$ sample. The red line shows the total fit while the magenta dashed line shows the total signal. The total background is shown as a blue dashed line with a long-dashed grey line showing the prompt $J/\psi$ background. 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.

Table 8: Values of the physical parameters extracted in the combination of 13 TeV results with those obtained from 7 TeV and 8 TeV data.

<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>-0.076</td>
<td>0.034</td>
<td>0.019</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>0.068</td>
<td>0.004</td>
<td>0.003</td>
</tr>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>0.669</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>$</td>
<td>A_{\parallel}(0)</td>
<td>^2$</td>
<td>0.220</td>
</tr>
<tr>
<td>$</td>
<td>A_{0}(0)</td>
<td>^2$</td>
<td>0.517</td>
</tr>
<tr>
<td>$</td>
<td>A_S</td>
<td>^2$</td>
<td>0.043</td>
</tr>
<tr>
<td>$\delta_{\perp}$ [rad]</td>
<td>3.075</td>
<td>0.096</td>
<td>0.091</td>
</tr>
<tr>
<td>$\delta_{\parallel}$ [rad]</td>
<td>3.295</td>
<td>0.079</td>
<td>0.202</td>
</tr>
<tr>
<td>$\delta_{\perp} - \delta_S$ [rad]</td>
<td>-0.216</td>
<td>0.037</td>
<td>0.010</td>
</tr>
</tbody>
</table>
Figure 7: 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 $B^0_s \to J/\psi\phi$ 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.

Two dimensional likelihood contours in the $\phi_s - \Delta \Gamma_s$ plane are shown in Figure 9 for this ATLAS result, the result of CMS [10] using the $B^0_s \to J/\psi\phi$ decay, and a combination of three LHCb measurements [8, 11, 12] using $B^0_s \to J/\psi\phi$, and $B^0_s$ decays to $\psi(2S)\phi$ and to $D^+_sD^-_s$, respectively. The contours are obtained interpreting each result as Gaussian 2D contour in the $\phi_s - \Delta \Gamma_s$ plane. All results are consistent between each other and with the SM [2, 4].
Figure 8: Likelihood 68% confidence level contours in the $\phi_s - \Delta \Gamma_s$ plane, showing ATLAS results for 7 TeV and 8 TeV data (blue dashed-dotted curve), for 13 TeV data (green dashed curve) and for 13 TeV data combined with 7 TeV and 8 TeV (red solid curve) data. In all contours the statistical and systematic uncertainties are combined in quadrature and correlations are taken into account.

Figure 9: Likelihood 68% confidence level contours in the $\phi_s - \Delta \Gamma_s$ plane, including results from LHCb (green) and CMS (red) using 7 TeV and 8 TeV data. The brown contour shows the ATLAS result for 13 TeV combined with 7 TeV and 8 TeV. In all contours the statistical and systematic uncertainties are combined in quadrature.
9 Summary

A measurement of the time-dependent CP asymmetry parameters in $B_0^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ decays from a 80.5 fb$^{-1}$ data sample of $pp$ collisions collected with the ATLAS detector during the 13 TeV LHC run is presented. The values from the 13 TeV analysis are consistent with those obtained in the previous analysis using 7 TeV and 8 TeV ATLAS data [9]. The two measurements are statistically combined leading to the following results:

\[
\begin{align*}
\phi_s &= -0.076 \pm 0.034 \text{ (stat.)} \pm 0.019 \text{ (syst.)} \text{ rad} \\
\Delta \Gamma_s &= 0.068 \pm 0.004 \text{ (stat.)} \pm 0.003 \text{ (syst.) ps}^{-1} \\
\Gamma_s &= 0.669 \pm 0.001 \text{ (stat.)} \pm 0.001 \text{ (syst.) ps}^{-1} \\
|A_{\parallel}(0)|^2 &= 0.220 \pm 0.002 \text{ (stat.)} \pm 0.002 \text{ (syst.)} \\
|A_{0}(0)|^2 &= 0.517 \pm 0.001 \text{ (stat.)} \pm 0.004 \text{ (syst.)} \\
|A_{S}(0)|^2 &= 0.043 \pm 0.004 \text{ (stat.)} \pm 0.004 \text{ (syst.)} \\
\delta_\perp &= 3.075 \pm 0.096 \text{ (stat.)} \pm 0.091 \text{ (syst.)} \text{ rad} \\
\delta_\parallel &= 3.295 \pm 0.079 \text{ (stat.)} \pm 0.202 \text{ (syst.)} \text{ rad} \\
\delta_\perp - \delta_S &= -0.216 \pm 0.037 \text{ (stat.)} \pm 0.010 \text{ (syst.)} \text{ rad}
\end{align*}
\]

The new ATLAS result is consistent with previous Run-1 results from LHCb [8] and CMS [10], using the $B_s^0 \to J/\psi\phi$ decay, and with the SM. The ATLAS result presented in this paper gives the most stringent measurement on parameters $\phi_s$, $\Delta \Gamma_s$, $\Gamma_s$ and the helicity functions parameters of the $B_s^0 \to J/\psi\phi$ decay from a single measurement.

References

[1] F. J. Botella, G. C. Branco and M. Nebot, 
*CP violation and limits on New Physics including recent $B_s$ measurements*, 

*Predictions of selected flavour observables within the standard model*, 

[3] LHCb Collaboration, 
*Precision measurement of the $B_s^0\to B_s^0$ oscillation frequency with the decay $B_s^0 \to D^-\pi^+$*, 

[4] A. Lenz and U. Nierste, 
*Numerical updates of lifetimes and mixing parameters of $B$ mesons*, 

[5] A. Lenz and U. Nierste, 
*Theoretical update of $B_s - \bar{B}_s$ mixing*, 

[6] D0 Collaboration, 
*Measurement of the CP-violating phase $\phi^{J/\psi\phi}_s$ using the flavor-tagged decay $B_s^0 \to J/\psi\phi$ in 8 $fb^{-1}$ of $p\bar{p}$ collisions*, 

[7] CDF Collaboration, 
*Measurement of the Bottom-Strange Meson Mixing Phase in the Full CDF Data Set*, 


25


[29] ATLAS Collaboration, Time-dependent angular analysis of the decay $B^0_s \to J/\psi\phi$ and extraction of $\Delta\Gamma_s$ and the CP-violating weak phase $\phi_s$ by ATLAS, JHEP 12 (2012) 072, arXiv: 1208.0572 [hep-ex].


[31] S. M. Flatté, Coupled - Channel Analysis of the $\pi\eta$ and $K\bar{K}$ Systems Near $K\bar{K}$ Threshold, Phys. Lett. B 63 (1976) 224.


[34] LHCb Collaboration, Study of the kinematic dependences of $\Lambda_b^0$ production in pp collisions and a measurement of the $\Lambda_b^0 \to \Lambda_c^- \pi^-$ branching fraction, JHEP 08 (2014) 143, arXiv: 1405.6842 [hep-ex].

[35] LHCb Collaboration, Study of the production of $\Lambda_b^0$ and $\bar{B}^0$ hadrons in pp collisions and first measurement of the $\Lambda_b^0 \to J/\psi pK^-\pi^-$ branching fraction, Chin. Phys. C 40 (2016) 011001, arXiv: 1509.00292 [hep-ex].

