Measurement of the CP-violating phase $\phi_s$ and the $B_s^0$ meson decay width difference with $B_s^0 \rightarrow J/\psi \phi$ decays in ATLAS

The ATLAS collaboration

E-mail: atlas.publications@cern.ch

Abstract: A measurement of the $B_s^0$ decay parameters in the $B_s^0 \rightarrow J/\psi \phi$ channel using an integrated luminosity of 14.3 fb$^{-1}$ collected by the ATLAS detector from 8 TeV $pp$ collisions at the LHC is presented. The measured parameters include the CP-violating phase $\phi_s$, the decay width $\Gamma_s$ and the width difference between the mass eigenstates $\Delta \Gamma_s$. The values measured for the physical parameters are statistically combined with those from 4.9 fb$^{-1}$ of 7 TeV data, leading to the following:

$$\begin{align*}
\phi_s &= -0.090 \pm 0.078 \text{ (stat.)} \pm 0.041 \text{ (syst.)} \text{ rad} \\
\Delta \Gamma_s &= 0.085 \pm 0.011 \text{ (stat.)} \pm 0.007 \text{ (syst.)} \text{ ps}^{-1} \\
\Gamma_s &= 0.675 \pm 0.003 \text{ (stat.)} \pm 0.003 \text{ (syst.)} \text{ ps}^{-1}.
\end{align*}$$

In the analysis the parameter $\Delta \Gamma_s$ is constrained to be positive. Results for $\phi_s$ and $\Delta \Gamma_s$ are also presented as 68% and 95% likelihood contours in the $\phi_s$-$\Delta \Gamma_s$ plane. Also measured in this decay channel are the transversity amplitudes and corresponding strong phases. All measurements are in agreement with the Standard Model predictions.

Keywords: B physics, CP violation, Flavor physics, Hadron-Hadron scattering (experiments)

ArXiv ePrint: 1601.03297
1 Introduction

New phenomena beyond the predictions of the Standard Model (SM) may alter CP violation in b-hadron decays. A channel that is expected to be sensitive to new physics contributions is the decay $B_s^0 \rightarrow J/\psi \phi$. CP violation in the $B_s^0 \rightarrow J/\psi \phi$ decay occurs due to interference between direct decays and decays with $B_s^0$-$\bar{B}_s^0$ mixing. The oscillation frequency of $B_s^0$ meson mixing is characterized 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 $B_s^0$-$\bar{B}_s^0$ mixing amplitude and the $b \rightarrow c\bar{s}s$ decay amplitude. In the absence of CP violation, the $B_H$ state would correspond to the CP-odd state and the $B_L$ to the CP-even state. In the SM the phase $\phi_s$ is small and can be related to Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix elements via the relation $\phi_s \sim -2\beta_s$, etc.
with \( \beta_s = \arg[-(V_{ts}V_{tb}^*)/(V_{cs}V_{cb}^*)] \); assuming no physics beyond the SM contributions to \( B_s^0 \) 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 [1].

Other physical quantities involved in \( B_s^0-\bar{B}_s^0 \) 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 different eigenstates. The width difference is predicted to be \( \Delta\Gamma_s = 0.087 \pm 0.021 \) ps\(^{-1} \) [2]. Physics beyond the SM is not expected to affect \( \Delta\Gamma_s \) as significantly as \( \phi_s \) [3]. However, extracting \( \Delta\Gamma_s \) from data is interesting as it allows theoretical predictions to be tested [3]. Previous measurements of these quantities have been reported by the DØ, CDF, LHCb, ATLAS and CMS collaborations [4–9].

The decay of the pseudoscalar \( B_s^0 \) to the vector-vector \( J/\psi(\mu^+\mu^-)\phi(K^+K^-) \) final state results in an admixture of \( CP \)-odd and \( CP \)-even states, with orbital angular momentum \( L = 0, 1 \) or \( 2 \). The final states with orbital angular momentum \( L = 0 \) or \( 2 \) are \( CP \)-even, while the state with \( L = 1 \) is \( CP \)-odd. The same final state can also be produced with \( K^+K^- \) pairs in an \( S \)-wave configuration [10]. This \( S \)-wave final state is \( CP \)-odd. The \( CP \) states are separated statistically using an angular analysis of the final-state particles. Flavour tagging is used to distinguish between the initial \( B_s^0 \) and \( \bar{B}_s^0 \) states.

The analysis presented here provides a measurement of the \( B_s^0 \rightarrow J/\psi\phi \) decay parameters using 14.3 fb\(^{-1} \) of LHC \( pp \) data collected by the ATLAS detector during 2012 at a centre-of-mass energy of 8 TeV. This is an update of the previous flavour-tagged time-dependent angular analysis of \( B_s^0 \rightarrow J/\psi\phi \) [8] that was performed using 4.9 fb\(^{-1} \) of data collected at 7 TeV. Electrons are now included, in addition to final-state muons, for the flavour tagging using leptons.

## 2 ATLAS detector and Monte Carlo simulation

The ATLAS detector [11] is a multi-purpose particle physics detector with a forward-backward symmetric cylindrical geometry and nearly 4\( \pi \) coverage in solid angle.\(^1 \) The inner tracking detector (ID) consists of a silicon pixel detector, a silicon microstrip detector and a transition radiation tracker. The ID is surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, and 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 instrumented with LAr calorimeters for electromagnetic and hadronic measurements. The muon spectrometer (MS) surrounds the calorimeters and consists of three large superconducting toroids with eight coils each, a system of tracking chambers, and detectors for triggering.

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

\(^1\) ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the centre of the LHC ring, and the y-axis points upward. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \( \phi \) being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle \( \theta \) as \( \eta = -\ln \tan(\theta/2) \).
correctly and when the LHC beams were declared to be stable are used in the analysis. The data were collected during a period of rising instantaneous luminosity, and the trigger conditions varied over this time. The triggers used to select events for this analysis are based on 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 14.3 fb$^{-1}$ of $pp$ collision data collected with the ATLAS detector at a centre-of-mass energy of 8 TeV. Data collected at the beginning of the 8 TeV data-taking period are not included in the analysis due to a problem with the trigger tracking algorithm. The trigger was subsequently changed to use a different tracking algorithm that did not have this problem.

To study the detector response, estimate backgrounds and model systematic effects, 12 million Monte Carlo (MC) simulated $B_s^0 \rightarrow J/\psi\phi$ events were generated using PYTHIA 8 [12, 13] tuned with ATLAS data [14]. No $p_T$ cuts were applied at the generator level. The detector response was simulated using the ATLAS simulation framework based on GEANT4 [15, 16]. In order to take into account 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 in data. Additional samples of the background decay $B_{d}^0 \rightarrow J/\psi K^0_{s}$, as well as the more general $b\bar{b} \rightarrow J/\psi X$ and $pp \rightarrow J/\psi X$ backgrounds were also simulated using PYTHIA 8.

3 Reconstruction and candidate selection

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 [17]. A muon identified using a combination of MS and ID track parameters is referred to as a combined-muon. A muon formed from a MS track segment that is not associated with a MS track but is matched to an ID track extrapolated to the MS is referred to as a segment-tagged muon. 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 refitted to a common vertex and the pair is accepted for further consideration if the quality of the fit meets the requirement $\chi^2$/d.o.f. $< 10$. The invariant mass of the muon pair is calculated from the refitted track parameters. 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. A maximum-likelihood fit is used to extract the $J/\psi$ mass and the corresponding mass resolution for these three subsets. When both muons have $|\eta| < 1.05$, the dimuon invariant mass must fall in the range 2.959–3.229 GeV to be accepted as a $J/\psi$ candidate. When one muon has $1.05 < |\eta| < 2.5$ and the other muon $|\eta| < 1.05$, the corresponding signal region is 2.913–3.273 GeV. For the third subset, where both muons have $1.05 < |\eta| < 2.5$, the signal region is 2.852–3.332 GeV. In each case the signal region is defined so as to retain 99.8% 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 particles 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. Each of the four tracks is required to have at least one hit in the pixel detector and at least four hits in the silicon microstrip detector. The fit is further constrained by fixing the invariant mass calculated from the two muon tracks to the $J/\psi$ mass [18]. A quadruplet of tracks is accepted for further analysis if the vertex fit has a $\chi^2$/d.o.f. $< 3$, the fitted $p_T$ of each track from $\phi \rightarrow K^+K^-$ is greater than 1 GeV and the invariant mass of the track pairs (assuming that they are kaons) falls within the interval $1.0085 \text{ GeV} < m(K^+K^-) < 1.0305 \text{ GeV}$. If there is more than one accepted candidate in the event, the candidate with the lowest $\chi^2$/d.o.f. is selected. In total, 375,987 $B_s^0$ candidates are collected within a mass range of 5.150–5.650 GeV.

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

$$t = \frac{L_{xy} m_B}{p_T B},$$

where $p_T B$ is the reconstructed transverse momentum of the $B_s^0$ meson candidate and $m_B$ denotes the mass of the $B_s^0$ meson, taken from [18]. The transverse decay length, $L_{xy}$, is the displacement in the transverse plane of the $B_s^0$ meson decay vertex with respect to the primary vertex, projected onto the direction of the $B_s^0$ transverse momentum. The position of the primary vertex used to calculate this quantity is determined from a refit following the removal of the tracks used to reconstruct the $B_s^0$ meson candidate.

For the selected events the average number of pile-up proton-proton interactions is 21, necessitating a choice of the best candidate for the primary vertex at which the $B_s^0$ meson is produced. The variable used is the three-dimensional impact parameter $d_0$, which is calculated as the distance between the line extrapolated from the reconstructed $B_s^0$ meson vertex in the direction of the $B_s^0$ momentum, and each primary vertex candidate. The chosen primary vertex is the one with the smallest $d_0$.

A study [19] made using a MC simulated dataset has shown that the precision of the reconstructed $B_s^0$ proper decay time remains stable over the range of pile-up encountered during 2012 data-taking. No $B_s^0$ meson decay-time cut is applied in this analysis.

### 4 Flavour tagging

The initial flavour of a neutral $B$ meson can be inferred using information from the opposite-side $B$ meson that contains the other pair-produced $b$-quark in the event [20, 21]. This is referred to as opposite-side tagging (OST).

To study and calibrate the OST methods, events containing $B^\pm \rightarrow J/\psi K^\pm$ decays are used, where the flavour of the $B^\pm$-meson is provided by the kaon charge. A sample of $B^\pm \rightarrow J/\psi K^\pm$ candidates is selected from the entire 2012 dataset satisfying the data-quality selection described in section 2. Since the OST calibration is not affected by the trigger problem at the start of the 8 TeV data-taking period, the tagging measurement uses 19.5 fb$^{-1}$ of integrated luminosity of $pp$ collision data.
4.1 $B^\pm \to J/\psi K^\pm$ event selection

In order to select candidate $B^\pm \to J/\psi K^\pm$ decays, firstly $J/\psi$ candidates are selected from pairs of oppositely charged combined-muons forming a good vertex, following the criteria described in section 3. Each muon is required to have a transverse momentum of at least 4 GeV and pseudorapidity within $|\eta| < 2.5$. The invariant mass of the dimuon candidate is required to satisfy $2.8 \text{ GeV} < m(\mu^+\mu^-) < 3.4 \text{ GeV}$. To form the $B$ candidate, an additional track, satisfying the same quality requirements described for tracks in section 3, is combined with the dimuon candidate using the charged kaon mass hypothesis, and a vertex fit is performed with the mass of the dimuon pair constrained to the known value of the $J/\psi$ mass. To reduce the prompt component of the combinatorial background, a requirement is applied to the transverse decay length of the $B$ candidate of $L_{xy} > 0.1 \text{ mm}$.

A sideband subtraction method is used in order to study parameter distributions corresponding to the $B^\pm$ signal processes with the background component subtracted. Events are divided into sub-sets into five intervals in the pseudorapidity of the $B$ candidate and three mass regions. The mass regions are defined as a signal region around the fitted peak signal mass position $\mu \pm 2\sigma$ and the sideband regions are defined as $[\mu - 5\sigma, \mu - 3\sigma]$ and $[\mu + 3\sigma, \mu + 5\sigma]$, where $\mu$ and $\sigma$ are the mean and width of the Gaussian function describing the $B$ signal mass. Separate binned extended maximum-likelihood fits are performed to the invariant mass distribution in each region of pseudorapidity.

An exponential function is used to model the combinatorial background and a hyperbolic tangent function to parameterize the low-mass contribution from incorrectly or partially reconstructed $B$ decays. A Gaussian function is used to model the $B^\pm \to J/\psi \pi^\pm$ contribution. The contribution from non-combinatorial background is found to have a negligible effect on the tagging procedure. Figure 1 shows the invariant mass distribution of $B$ candidates for all rapidity regions overlaid with the fit result for the combined data.

4.2 Flavour tagging methods

Several methods that differ in efficiency and discriminating power are available to infer the flavour of the opposite-side $b$-quark. The measured charge of a muon or electron from a semileptonic decay of the $B$ meson provides strong separation power; however, the $b \to \ell$ transitions are diluted through neutral $B$ meson oscillations, as well as by cascade decays $b \to c \to \ell$, which can alter the charge of the lepton relative to those from direct $b \to \ell$ decays. The separation power of lepton tagging is enhanced by considering a weighted sum of the charge of the tracks in a cone around the lepton, where the weighting function is determined separately for each tagging method by optimizing the tagging performance. If no lepton is present, a weighted sum of the charge of tracks in a jet associated with the opposite-side $B$ meson decay provides some separation. The flavour tagging methods are described in detail below.

For muon-based tagging, an additional muon is required in the event, with $p_T > 2.5$ GeV, $|\eta| < 2.5$ and with $|\Delta z| < 5 \text{ mm}$ from the primary vertex. Muons are classified according to their reconstruction class, combined or segment-tagged, and subsequently treated as
Figure 1. The invariant mass distribution for $B^\pm \to J/\psi K^\pm$ candidates satisfying the selection criteria, used to study the flavour tagging. Data are shown as points, and the overall result of the fit is given by the blue curve. The contribution from the combinatorial background component is indicated by the red dotted line, partially reconstructed $B$ decays by the green shaded area, and decays of $B^\pm \to J/\psi \pi^\pm$, where the pion is mis-assigned a kaon mass, by the purple dashed line.

Figure 2. The opposite-side muon cone charge distribution for $B^\pm$ signal candidates for segment-tagged (left) and combined (right) muons. The $B^\pm$ charge is determined from the kaon charge.

distinct flavour tagging methods. In the case of multiple muons, the muon with the highest transverse momentum is selected.

A muon cone charge variable is constructed, defined as

$$Q_\mu = \frac{\sum_i^N q_i \cdot (p_T^i)^\kappa}{\sum_i^N (p_T^i)^\kappa},$$

where $q$ is the charge of the track, $\kappa = 1.1$ and the sum is performed over the reconstructed ID tracks within a cone, $\Delta R = \sqrt{\Delta \phi^2 + (\Delta \eta)^2} < 0.5$, around the muon direction. The reconstructed ID tracks must have $p_T > 0.5$ GeV and $|\eta| < 2.5$. Tracks associated with the $B^\pm$ signal decay are excluded from the sum. In figure 2 the opposite-side muon cone charge distributions are shown for candidates from $B^\pm$ signal decays.
For electron-based tagging, an electron is identified using information from the inner detector and calorimeter and is required to satisfy the tight electron quality criteria [22]. The inner detector track associated with the electron is required to have $p_T > 0.5$ GeV and $|\eta| < 2.5$. It is required to pass within $|\Delta z| < 5$ mm of the primary vertex to remove electrons from non-signal interactions. To exclude electrons associated with the signal-side of the decay, electrons are rejected that have momenta within a cone of size $\Delta R = 0.4$ around the signal $B$ candidate direction in the laboratory frame and opening angle between the $B$ candidate and electron momenta, $\cos(\zeta_b) > 0.98$. 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 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 tracks within the cone, the charge of the electron is used. The resulting opposite-side electron cone charge distribution is shown in figure 3 for $B^+$ and $B^-$ signal events.

In the absence of a muon or electron, $b$-tagged jets (i.e. jets that are the product of a $b$-quark) are identified using a multivariate tagging algorithm [23], which is a combination of several $b$-tagging algorithms using an artificial neural network and outputs a $b$-tag weight classifier. Jets are selected that exceed a $b$-tag weight of 0.7. This value is optimized to maximize the tagging power of the calibration sample. Jets are reconstructed from track information using the anti-$k_t$ algorithm [24] with a radius parameter $R = 0.8$. In the case of multiple jets, the jet with the highest value of the $b$-tag weight is used.

The jet charge is defined as

$$Q_{\text{jet}} = \sum_{i}^{N \text{ tracks}} q_i \cdot (p_{Ti})^\kappa / \sum_{i}^{N \text{ tracks}} (p_{Ti})^\kappa,$$

where $\kappa = 1.1$ and the sum is over the tracks associated with the jet, excluding those tracks associated with a primary vertex other than that of the signal decay and tracks from the signal candidate. Figure 4 shows the distribution of the opposite-side jet-charge for $B^\pm$ signal candidates.

The efficiency, $\epsilon$, of an individual tagging method is defined as the ratio of the number of events tagged by that method to the total number of candidates. A probability $P(B|Q)$ ($P(B|Q)$) that a specific event has a signal decay containing a $\bar{b}$-quark ($b$-quark).
Figure 4. Opposite-side jet-charge distribution for $B^\pm$ signal candidates.

given the value of the discriminating variable is constructed from the calibration samples for each of the $B^+$ and $B^-$ samples, which defines $P(Q|B^+)$ and $P(Q|B^-)$, respectively. The probability to tag a signal event as containing a $b$-quark is therefore $P(B|Q) = P(Q|B^+)/[P(Q|B^+) + P(Q|B^-)]$, and correspondingly $P(\bar{B}|Q) = 1 - P(B|Q)$. It is possible to define a quantity called the dilution $D = P(B|Q) - P(\bar{B}|Q) = 2P(B|Q) - 1$, which represents the strength of a particular flavour tagging method. The tagging power of a particular tagging method is defined as $T = \epsilon D^2 = \sum_i \epsilon_i \cdot (2P_i(B|Q_i) - 1)^2$, where the sum is over the bins of the probability distribution as a function of the charge variable. An effective dilution, $D_{\text{eff}} = \sqrt{T/\epsilon}$, is calculated from the measured tagging power and efficiency.

The flavour tagging method applied to each $B^0_s$ candidate event is taken from the information contained in a given event. By definition there is no overlap between lepton-tagged and jet-charge-tagged events. The overlap between muon- and electron-tagged events, corresponding to 0.4% of all tagged events, is negligibly small. In the case of doubly tagged events, the tagger with the highest tagging power is selected; however, the choice of hierarchy between muon- and electron-tagged events is shown to have negligible impact on the final fit results. If it is not possible to provide a tagging response for the event, then a probability of 0.5 is assigned. A summary of the tagging performance is given in table 1.

4.3 Using tag information in the $B^0_s$ fit

The tag-probability for each $B^0_s$ candidate is determined from calibrations derived from a sample of $B^\pm \to J/\psi K^\pm$ candidates, as described in section 4.2. The distributions of tag-probabilities for the signal and background are different and since the background cannot be factorized out, additional probability terms, $P_s(P(B|Q))$ and $P_b(P(B|Q))$ for signal and background, respectively, are included in the fit. The distributions of tag-probabilities for the $B^0_s$ candidates consist of continuous and discrete parts (events with a tag charge of $\pm 1$); these are treated separately as described below.

To describe the continuous part, a fit is first performed to the sideband data, i.e., $5.150 \text{ GeV} < m(B^0_s) < 5.317 \text{ GeV}$ or $5.417 \text{ GeV} < m(B^0_s) < 5.650 \text{ GeV}$, where $m(B^0_s)$ is the mass of the $B^0_s$ candidate. Different functions are used for the different tagging methods. For the combined-muon tagging method, the function has the form of the sum of a
fourth-order polynomial and two exponential functions. A second-order polynomial and two exponential functions are applied for the electron tagging algorithm. A sum of three Gaussian functions is used for the segment-tagged muons. For the jet-charge tagging algorithm an eighth-order polynomial is used. In all four cases unbinned maximum-likelihood fits to data are used. In the next step, the same function as applied to the sidebands is used to describe the distributions for events in the signal region: the background parameters are fixed to the values obtained from the fits to the sidebands while the signal parameters are free in this step. The ratio of background to signal (obtained from a simultaneous mass-lifetime fit) is fixed as well. The results of the fits projected onto histograms of $B_s^0$ tag-probability for the different tagging methods are shown in figure 5.

To account for possible deviations between data and the selected fit models a number of alternative fit functions are used to determine systematic uncertainties in the $B_s^0$ fit. These fit variations are described in section 7.

The discrete components of the tag-probability distribution originate from cases where the tag is derived from a single track, giving a tag charge of exactly +1 or −1. The fractions of events $f_{+1}$ and $f_{-1}$ with charges +1 and −1, respectively, are determined separately for signal and background using events from the same $B_s^0$ mass signal and sideband regions. 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 the case for background candidates formed from a partially reconstructed $b$-hadron. For signal and background contributions, similar fractions of events that are tagged with +1 or −1 tagging charge are observed for each of the tagging methods. The remaining fraction of events, $1 - f_{+1} - f_{-1}$, constitute the continuous part of the distributions. Table 2 summarizes the fractions $f_{+1}$ and $f_{-1}$ obtained for signal and background events and for the different tag methods.

To estimate the fractions of signal and background events which have tagging, a similar sideband-subtraction method is used to determine the relative fraction of signal and background events tagged using the different methods. These fractions are also included in the maximum-likelihood fit, described in section 5. The results are summarized in table 3.

<table>
<thead>
<tr>
<th>Tagger</th>
<th>Efficiency [%]</th>
<th>Dilution [%]</th>
<th>Tagging Power [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined $\mu$</td>
<td>4.12 ± 0.02</td>
<td>47.4 ± 0.2</td>
<td>0.92 ± 0.02</td>
</tr>
<tr>
<td>Electron</td>
<td>1.19 ± 0.01</td>
<td>49.2 ± 0.3</td>
<td>0.29 ± 0.01</td>
</tr>
<tr>
<td>Segment-tagged $\mu$</td>
<td>1.20 ± 0.01</td>
<td>28.6 ± 0.2</td>
<td>0.10 ± 0.01</td>
</tr>
<tr>
<td>Jet-charge</td>
<td>13.15 ± 0.03</td>
<td>11.85 ± 0.03</td>
<td>0.19 ± 0.01</td>
</tr>
<tr>
<td>Total</td>
<td>19.66 ± 0.04</td>
<td>27.56 ± 0.06</td>
<td>1.49 ± 0.02</td>
</tr>
</tbody>
</table>

Table 1. Summary of tagging performance for the different flavour tagging methods 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 charge distribution. The effective dilution is obtained from the measured efficiency and tagging power. For the efficiency, dilution, and tagging power, the corresponding uncertainty is determined by combining the appropriate uncertainties in the individual bins of each charge distribution.
Figure 5. The continuous part of tag-probability for tagging using combined-muons (top-left), electrons (top-right), segment-tagged muons (bottom-left) and jet-charge (bottom-right). Black dots are data, blue is a fit to the sidebands, purple to the signal and red is a sum of both fits.

<table>
<thead>
<tr>
<th>Tag method</th>
<th>Signal</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$f_{+1}$</td>
<td>$f_{-1}$</td>
</tr>
<tr>
<td>Combined $\mu$</td>
<td>0.124 ± 0.012</td>
<td>0.127 ± 0.012</td>
</tr>
<tr>
<td>Electron</td>
<td>0.105 ± 0.020</td>
<td>0.139 ± 0.021</td>
</tr>
<tr>
<td>Segment-tagged $\mu$</td>
<td>0.147 ± 0.024</td>
<td>0.118 ± 0.023</td>
</tr>
<tr>
<td>Jet-charge</td>
<td>0.071 ± 0.005</td>
<td>0.069 ± 0.005</td>
</tr>
</tbody>
</table>

Table 2. Table summarizing the fraction of events $f_{+1}$ and $f_{-1}$ with tag charges of +1 and −1, respectively for signal and background events and for the different tag methods. Only statistical errors are quoted.
<table>
<thead>
<tr>
<th>Tag method</th>
<th>Signal</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined $\mu$</td>
<td>0.047 ± 0.003</td>
<td>0.038 ± 0.001</td>
</tr>
<tr>
<td>Electron</td>
<td>0.012 ± 0.001</td>
<td>0.008 ± 0.001</td>
</tr>
<tr>
<td>Segment-tagged $\mu$</td>
<td>0.013 ± 0.001</td>
<td>0.015 ± 0.001</td>
</tr>
<tr>
<td>Jet-charge</td>
<td>0.135 ± 0.003</td>
<td>0.100 ± 0.001</td>
</tr>
<tr>
<td>Untagged</td>
<td>0.793 ± 0.002</td>
<td>0.839 ± 0.002</td>
</tr>
</tbody>
</table>

Table 3. Table summarizing the relative fractions of signal and background events tagged using the different tag methods. The fractions include both the continuous and discrete contributions. Only statistical errors are quoted.

5 Maximum likelihood fit

An unbinned maximum-likelihood fit is performed on the selected events to extract the parameter values of the $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ decay. The fit uses information about the reconstructed mass $m$, the measured proper decay time $t$, the measured proper decay time uncertainty $\sigma_t$, the tagging probability, and the transversity angles $\Omega$ of each $B_s^0 \to J/\psi\phi$ 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:

$$
\ln L = \sum_{i=1}^{N} \left\{ w_i \cdot \ln(f_s \cdot F_s(m_i, t_i, \sigma_t, \Omega_i, P(B|Q), p_{T_i}) \\
+ f_{B^0} \cdot F_{B^0}(m_i, t_i, \sigma_t, \Omega_i, P(B|Q), p_{T_i}) \\
+ f_{\Lambda_b} \cdot F_{\Lambda_b}(m_i, t_i, \sigma_t, \Omega_i, P(B|Q), p_{T_i}) \\
+ (1 - f_s \cdot (1 + f_{B^0} + f_{\Lambda_b})) F_{bkg}(m_i, t_i, \sigma_t, \Omega_i, P(B|Q), p_{T_i}) \right\},
$$

(5.1)

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), and $f_s$ is the fraction of signal candidates. The background fractions $f_{B^0}$ and $f_{\Lambda_b}$ are the fractions of $B^0$ mesons and $\Lambda_b$ baryons mis-identified as $B_s^0$ candidates calculated relative to the number of signal events; these parameters are fixed to their MC values 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 $i$. $F_s$, $F_{B^0}$, $F_{\Lambda_b}$ and $F_{bkg}$ are the probability density functions (PDF) modelling the signal, $B^0$ background, $\Lambda_b$ background, and the other background distributions, respectively. A detailed description of the signal PDF terms in equation (5.1) 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:

$$\mathcal{F}_s(m_i, t_i, \Omega_i, P(B|Q), p_{T_i}) = P_s(m_i) \cdot P_s(\Omega_i, t_i, P(B|Q), \sigma_{t_i}) \cdot P_s(\sigma_{t_i}) \cdot P_s(P(B|Q)) \cdot A(\Omega_i, p_{T_i}) \cdot P_s(p_{T_i}).$$

(5.2)

The mass function $P_s(m_i)$ is modelled by a sum of three Gaussian distributions. The probability terms $P_s(\sigma_{t_i})$ and $P_s(p_{T_i})$ are described by gamma functions and are unchanged from the analysis described in ref. [25]. The tagging probability term for signal $P_s(P(B|Q))$ is described in section 4.3.

The term $P_s(\Omega_i, t_i, P(B|Q), \sigma_{t_i})$ is a joint PDF for the decay time $t$ and the transversity angles $\Omega$ 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 $\Omega$ is given by the differential decay rate [26]:

$$\frac{\mathrm{d}^4\Gamma}{\mathrm{d}t \, \mathrm{d}\Omega} = \sum_{k=1}^{10} \mathcal{O}^{(k)}(t) g^{(k)}(\theta_T, \psi_T, \phi_T),$$

where $\mathcal{O}^{(k)}(t)$ are the time-dependent functions corresponding to the contributions of the four different amplitudes ($A_0$, $A_\parallel$, $A_\perp$, and $A_S$) and their interference terms, and $g^{(k)}(\theta_T, \psi_T, \phi_T)$ are the angular functions. Table 4 shows these time-dependent functions and the angular functions of the transversity angles. The formulae for the time-dependent functions have the same structure for $B_s^0$ and $B_d^0$ but with a sign reversal in the terms containing $\Delta m_s$. 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 $A_\parallel(t)$ 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 (which includes the $f_0$). The corresponding functions are given in the last four lines of table 4 ($k = 7-10$). The amplitudes are parameterized by $|A_i|e^{i\delta_i}$, where $i = \{0, \parallel, \perp, S\}$, with $\delta_0 = 0$ and are normalized such that $|A_0(0)|^2 + |A_\perp(0)|^2 + |A_\parallel(0)|^2 = 1$. $|A_\perp(0)|$ is determined according to this condition, while the remaining three amplitudes are parameters of the fit. The formalism used throughout this analysis assumes no direct $CP$ violation.

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 $\vec{p}(K^+) = -\vec{p}(J/\psi)$ in the $\phi$ meson rest frame.

The PDF term $P_s(\Omega_i, t_i, P(B|Q), \sigma_{t_i})$ 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 mis-measurements. The proper decay time uncertainty distribution for data, including the fits to the background and the signal contributions is shown in figure 6. The average value of this uncertainty for signal events is 97 fs.

The angular acceptance of the detector and kinematic cuts on the angular distributions are included in the likelihood function through $A(\Omega_i, p_{T_i})$. 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 $B^0_s \to J/\psi\phi$ MC events. Taking the small discrepancies between data and MC events into account have negligible effect on the fit results. 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 $B^0_s \to J/\psi(K^+K^-)$ 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:

$$F_{bkg}(m_i, t_i, \sigma_{t_i}, \Omega_i, P(B|Q), p_{T_i}) = P_b(m_i) \cdot P_b(t_i|\sigma_{t_i}) \cdot P_b(P(B|Q)) \cdot P_b(\Omega_i) \cdot P_b(\sigma_{t_i}) \cdot P_b(p_{T_i}).$$  

The proper decay time function $P_b(t_i|\sigma_{t_i})$ is parameterized 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 around zero. The two positive exponential functions represent a fraction of longer-lived backgrounds with non-prompt
Table 4. Table showing the ten time-dependent functions, \( \mathcal{O}^{(k)}(t) \) and the functions of the transversity angles \( g^{(k)}(\theta_T, \psi_T, \phi_T) \). The amplitudes \( |A_0(0)|^2 \) and \( |A_\parallel(0)|^2 \) are for the CP-even components of the \( B_0^\pm \to J/\psi \phi \) decay, \( |A_\perp(0)|^2 \) is the CP-odd amplitude; they have corresponding strong phases \( \delta_0, \delta_\parallel \) and \( \delta_\perp \). By convention \( \delta_0 \) is set to be zero. The S-wave amplitude \( |A_S(0)|^2 \) gives the fraction of \( B_0^\pm \to J/\psi K^+K^- (f_0) \) and has a related strong phase \( \delta_S \). The \( \pm \) and \( \mp \) terms denote two cases: the upper sign describes the decay of a meson that was initially a \( B_0^0 \) meson, while the lower sign describes the decays of a meson that was initially \( \bar{B}_0^0 \).
\(J/\psi\), 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 \(P_b(p_{T\ell})\) are described by gamma functions. They are unchanged from the analysis described in ref. [25] and explained in detail there. The tagging probability term for background \(P_b(P(B|Q))\) is described in section 4.3.

The shape of the background angular distribution, \(P_b(\Omega_i)\) arises primarily from detector and kinematic acceptance effects. These are described by Legendre polynomial functions:

\[
Y^m_l(\theta_T) = \frac{\sqrt{(2l+1)/(4\pi)}}{(l-m)!/(l+m)!} P^m_l(\cos \theta_T)
\]

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

\[
\mathcal{P}_b(\theta_T, \psi_T, \phi_T) = \sum_{k=0}^{6} \sum_{l=0}^{6} \sum_{m=-l}^{l} \left\{ \begin{array}{ll}
ak_{k,l,m} \sqrt{2} Y^m_l(\theta_T) \cos(m \phi_T) P_k(\cos \psi_T) & \text{where } m > 0 \\
ak_{k,l,m} \sqrt{2} Y^{-m}_l(\theta_T) \sin(m \phi_T) P_k(\cos \psi_T) & \text{where } m < 0 \\
ak_{k,l,m} \sqrt{2} Y^0_l(\theta_T) P_k(\cos \psi_T) & \text{where } m = 0
\end{array} \right.
\]

where the coefficients \(a_{k,l,m}\) are adjusted to give the best fit to the angular distributions for events in the \(B^0_s\) mass sidebands. These parameters are then fixed in the main fit. The \(B^0_s\) mass interval used for the background fit is between 5.150 and 5.650 GeV excluding the signal mass region \(|m(B^0_s) - 5.366\text{ GeV}| < 0.110\text{ GeV}\). The background mass model, \(P_b(m_i)\) is an exponential function with a constant term added.

Contamination from \(B_d \rightarrow J/\psi K^{0*}\) and \(\Lambda_b \rightarrow J/\psi p K^-\) events mis-reconstructed as \(B^0_s \rightarrow J/\psi \phi\) are accounted for in the fit through the \(\mathcal{F}_{B^0_s}\) and \(\mathcal{F}_{\Lambda_b}\) terms in the PDF function described in equation (5.1). The fraction of these contributions, \(f_{B^0_s} = (3.3 \pm 0.5)\%\) and \(f_{\Lambda_b} = (1.8 \pm 0.6)\%\), are evaluated from MC simulation using production and branching fractions from refs. [18, 27–31]. MC simulated events are also used to determine the shape of the mass and transversity angle distributions. The 3D angular distributions of \(B^0_d \rightarrow J/\psi K^{*0}\) and of the conjugate decay are modelled using input from ref. [32], while angular distributions for \(\Lambda_b \rightarrow J/\psi p K^-\) and the conjugate decay are modelled as flat. These distributions are sculpted for detector acceptance effects and then described by Legendre polynomial functions, equation (5.4), as in the case of the background described by equation (5.3). 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 \(B_d/B^{0}_s\) or \(\Lambda_b/B^{0}_s\) masses as appropriate, where the lifetimes and masses are taken from ref. [18]. Systematic uncertainties due to the background from \(B_d \rightarrow J/\psi K^{0*}\) and \(\Lambda_b \rightarrow J/\psi p K^-\) decays are described in section 7. The contribution of \(B_d \rightarrow J/\psi K \pi\) events as well as their interference with \(B_d \rightarrow J/\psi K^{0*}\) events is not included in the fit and is instead assigned as a systematic uncertainty.

To account for possible deviations between data and the selected fit models a number of alternative fit functions and mass selection criteria are used to determine systematic uncertainties in the \(B^0_s\) fit. These fit variations are described in section 7.
5.3 Muon trigger proper time-dependent efficiency

It was observed that the muon trigger biases the transverse impact parameter parameter of muons, resulting in a minor 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 of an unbiased sample with the distribution obtained including the trigger. To account for this inefficiency in the fit, the events are re-weighted by a factor $w$:

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

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 off-line track reconstruction, vertex reconstruction, or track quality selection criteria is observed.

6 Results

The full simultaneous unbinned maximum-likelihood fit contains nine physical parameters: $\Delta \Gamma_s, \phi_s, \Gamma_s, |A_0(0)|^2, |A_1(0)|^2, \delta_0, |A_S(0)|^2$ and $\delta_s$. The other parameters in the likelihood function are the $B_s^0$ signal fraction $f_s$, parameters describing the $J/\psi\phi$ mass distribution, parameters describing the $B_s^0$ meson 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 353 nuisance parameters describing the background and acceptance functions that are fixed at the time of the fit. The fit model is tested using pseudo-experiments as described in section 7. These tests show no significant bias, as well as no systematic underestimation of the statistical errors reported from the fit to data.

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 $B_s^0$ meson candidates of 74900 $\pm$ 400. The results and correlations of the physics parameters obtained from the fit are given in tables 5 and 6. Fit projections of the mass, proper decay time and angles are given in figures 7 and 8, respectively.

7 Systematic uncertainties

Systematic uncertainties are assigned by considering 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 $B^{\pm} \rightarrow J/\psi K^{\pm}$ decays is included in the overall statistical error. The systematic uncertainty arising from the precision of the tagging calibration is estimated by changing the model used to parameterize the probability distribution, $P(B|Q)$, as a function of tag charge from the third-order polynomial function used by default to one of several alternative
Table 5. 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>-0.110</td>
<td>0.082</td>
<td>0.042</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>0.101</td>
<td>0.013</td>
<td>0.007</td>
</tr>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>0.676</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>$</td>
<td>A_\parallel(0)</td>
<td>^2$</td>
<td>0.230</td>
</tr>
<tr>
<td>$</td>
<td>A_0(0)</td>
<td>^2$</td>
<td>0.520</td>
</tr>
<tr>
<td>$</td>
<td>A_S(0)</td>
<td>^2$</td>
<td>0.097</td>
</tr>
<tr>
<td>$\delta_\perp$ [rad]</td>
<td>4.50</td>
<td>0.45</td>
<td>0.30</td>
</tr>
<tr>
<td>$\delta_\parallel$ [rad]</td>
<td>3.15</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>$\delta_\perp - \delta_S$ [rad]</td>
<td>-0.08</td>
<td>0.03</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 6. Fit correlations between the physical parameters of interest.

|          | $\Delta \Gamma$ | $\Gamma_s$ | $|A_\parallel(0)|^2$ | $|A_0(0)|^2$ | $|A_S(0)|^2$ | $\delta_\parallel$ | $\delta_\perp$ | $\delta_\perp - \delta_S$ |
|----------|-----------------|------------|---------------------|-------------|-----------------|----------------|----------------|------------------|
| $\phi_s$ | 0.097           | -0.085     | 0.030               | 0.029       | 0.048           | 0.067          | 0.035          | -0.008           |
| $\Delta \Gamma$ | 1            | -0.414     | 0.098               | 0.136       | 0.045           | 0.009          | 0.008          | -0.011           |
| $\Gamma_s$ | 1              | -0.119     | -0.042              | 0.167       | -0.027         | -0.009         | 0.018          | 0.018            |
| $|A_\parallel(0)|^2$ | 1             | -0.330     | 0.072               | 0.105       | 0.025           | 0.007          | 0.001          | 0.018            |
| $|A_0(0)|^2$ | 1              | 0.234      | -0.011              | -0.011      | 0.004           | 0.052          | 0.007          | 0.014            |
| $|A_S(0)|^2$ | 1              | -0.046     | 0.046               | 0.158       | -0.006         | 1              | 0.018          |                  |
| $\delta_\parallel$ | 1              | 1          | 0.158               | -0.006      | 1              |                |                |                  |
| $\delta_\perp$ |                |            |                     |             |                |                |                |                  |

functions. The alternatives used are: a linear function; a fifth-order polynomial; or two third-order polynomials describing the positive and negative regions that share the constant and linear terms but have independent quadratic and cubic terms. For the combined-muon tagging, an additional model consisting of two third-order polynomials sharing the constant term but with independent linear, quadratic and cubic terms is also used. The $B^0_s$ fit is repeated using the alternative models and the largest difference is assigned as the systematic uncertainty.

- **Angular acceptance method:** the angular acceptance (from the detector and kinematic effects mentioned 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. These effects are found to be negligible.
Figure 7. (Left) Mass fit projection for the $B^0_s \rightarrow J/\psi \phi$ sample. The red line shows the total fit, the dashed purple line shows the signal component, the long-dashed dark blue line shows the $B^0_d \rightarrow J/\psi K^0$ component, while the solid light blue line shows the contribution from $\Lambda_b \rightarrow J/\psi p K$ events. (Right) Proper decay time fit projection for the $B^0_s \rightarrow J/\psi \phi$ sample. The red line shows the total fit while the purple 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 uncertainty ($\sigma$) of that point.

- **Inner detector alignment**: residual misalignments of the ID affect the impact parameter, $d_0$, distribution with respect to the primary vertex. The effect of a radial expansion on the measured $d_0$ is determined from data collected at 8 TeV, with a trigger requirement of at least one muon with a transverse momentum greater than or equal to 4 GeV. The radial expansion uncertainties determined in this way are 0.14% for $|\eta| < 1.5$ and 0.55% for $1.5 < |\eta| < 2.5$. These values are used to estimate the effect on the fitted $B^0_s$ parameter values. Small deviations are seen in some parameters, and these are included as systematic uncertainties.

- **Trigger efficiency**: to correct for the trigger lifetime bias the events are re-weighted according to equation (5.5). The uncertainty of the parameters $p_0, p_1, p_2$ and $p_3$ are used to estimate the systematic uncertainty due to the time efficiency correction. These uncertainties originate from the following sources: the limited size of the MC simulated dataset, the choice of bin-size for the proper decay time distributions and variations between different triggers. The systematic effects are found to be negligible.

- **Background angles model, choice of $p_T$ bins**: the shape of the background angular distribution, $P_b(\theta_T, \phi_T, \psi_T)$, is described by the Legendre polynomial functions given in equation (5.4). The shapes arise primarily from 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 four $p_T$ intervals: 0–13 GeV, 13–18 GeV, 18–25 GeV and >25 GeV. The system-
Figure 8. Fit projections for the transversity angles of events with $5.317 \text{ GeV} < m(J/\psi KK) < 5.417 \text{ GeV}$ for $\phi_T$ (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 CP-odd and CP-even signal components are shown by the red dot-dashed and orange dashed lines respectively, the S-wave component is given by the green dashed line and the blue dotted line shows the background contribution. The contributions of the interference terms are negligible in these projections and are not shown.

\begin{itemize}
\item **Background angles model, choice of mass sidebands:** the parameters of the Legendre polynomial functions given in equation (5.4) 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 signal mass regions: $|m(B_0^0) - 5.366| > 0.085 \text{ GeV}$ and $|m(B_0^0) - 5.366| > 0.160 \text{ GeV}$ (instead of $|m(B_0^0) - 5.366| > 0.110 \text{ GeV}$). The differences in the fit results are assigned as systematic uncertainties.

\item **$B_d$ contribution:** the contamination from $B_d \rightarrow J/\psi K^{0*}$ events mis-reconstructed as $B_0^0 \rightarrow J/\psi \phi$ 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 mass and transversity angles distribution. In the MC events the angular distribution of the $B_d \rightarrow J/\psi K^{0*}$ decay is modelled using parameters taken from ref. [32]. The uncertainties of these parameters are taken into account in the estimation of systematic uncertainties.
\end{itemize}
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 tests. The impact of all these uncertainties is found to have a negligible effect on the $B_s^0$ fit results. The contribution of $B_d \to J/\psi K\pi$ events as well as their interference with $B_d \to J/\psi K^{0*}$ events is not included in the fit and is instead assigned as a systematic uncertainty. To evaluate this uncertainty, the MC background events are modelled using both the P-wave $B_d \to J/\psi K^{0*}$ and S-wave $B_d \to J/\psi K\pi$ decays and their interference, using the input parameters taken from ref. [32]. The $B_s^0$ fit using this input was compared to the default fit, and differences are included in table 7.

- **$\Lambda_b$ contribution:** the contamination from $\Lambda_b \to J/\psi pK^-$ events mis-reconstructed as $B_s^0 \to J/\psi \phi$ is accounted for in the final fit. Studies are performed to evaluate the effect of the uncertainties in the $\Lambda_b \to J/\psi pK^-$ fraction $f_{\Lambda_b}$, and the shapes of the mass, transversity angles, and lifetime distributions. Additional studies are performed to determine the effect of the uncertainties in the $\Lambda_b \to J/\psi \Lambda^*$ branching ratios used to reweight the generated MC. These are uncertainties are included in table 7.

- **Fit model variations:** to estimate the systematic uncertainties due to the fit model, variations of the model are tested in pseudo-experiments. A set of $\approx 2500$ pseudo-experiments is generated for each variation considered, and fitted with the default model. The systematic error quoted for each effect is the difference between the mean shift of the fitted value of each parameter from its input value for the pseudo-experiments altered for each source of systematic uncertainty. In the first variation tested, the signal mass is generated using the fitted $B_s^0$ mass convolved with a Gaussian function using the measured per-candidate mass errors. In another test, the background mass is generated from an exponential function with the addition of a first-degree polynomial function instead of an exponential function plus a constant term. The time resolution model was varied by using two different scale factors to generate the lifetime uncertainty, instead of the single scale factor used in the default model. The non-negligible uncertainties derived from these tests are included in the systematic uncertainties shown in table 7. To determine the possible systematics effects of mis-modelling of the background events by the fitted background model, as seen in the low mass side-band region (5.150–5.210 GeV) of figure 7, left, alternative mass selection cuts are used with the default fit model. The effect of these changes on the fit results are found to be negligible.

- **Default fit model:** due to its complexity, the fit model is less sensitive to some nuisance parameters. This limited sensitivity could potentially lead to a bias in the measured physics parameters, even when the model perfectly describes the fitted data. To estimate the systematic uncertainty due to the choice of default fit model, a set of pseudo-experiments were conducted using the default model in both the generation and fit. The systematic uncertainties are determined from the mean of
Table 7. Summary of systematic uncertainties assigned to the physical parameters of interest.

|                          | $\phi_s$ | $\Delta \Gamma_s$ | $\Gamma_s$ | $|A_1(0)|^2$ | $|A_0(0)|^2$ | $|A_S(0)|^2$ | $\delta_{\perp}$ | $\delta_{\parallel}$ | $\delta_{\perp} - \delta_S$ |
|--------------------------|----------|--------------------|------------|--------------|--------------|--------------|----------------|----------------|----------------|
| Tagging                  | 0.025    | 0.003              | $< 10^{-3}$| $< 10^{-3}$  | $< 10^{-3}$  | 0.001        | 0.236          | 0.014          | 0.004          |
| Acceptance               | $< 10^{-3}$| $< 10^{-3}$       | 0.003      | $< 10^{-3}$  | 0.001        | 0.004        | 0.008          | 0.008          | $< 10^{-3}$  |
| Inner detector alignment | 0.005    | 0.002              | $< 10^{-3}$| $< 10^{-3}$  | $< 10^{-3}$  | 0.134        | 0.007          | 0.008          | $< 10^{-3}$  |
| Background angles model: | Choice of $p_T$ bins | 0.020 | 0.006 | 0.003         | 0.003     | 0.008        | 0.004         | 0.006          | 0.008          |
|                         | Choice of mass interval | 0.008 | 0.001 | 0.001         | $< 10^{-3}$ | 0.002        | 0.021         | 0.005          | 0.003          |
| $B^0_d$ background model | 0.023    | 0.001              | $< 10^{-3}$| 0.002        | 0.002        | 0.017        | 0.090         | 0.011          | 0.009          |
| $\Lambda_b$ background model | 0.011 | 0.002 | 0.001 | 0.001        | 0.007        | 0.009        | 0.045         | 0.006          | 0.007          |
| Fit model:               |          |                    |            |              |              |              |                |                |                |
| Mass signal model        | 0.004    | $< 10^{-3}$        | $< 10^{-3}$| 0.002        | $< 10^{-3}$  | 0.001        | 0.015         | 0.017          | $< 10^{-3}$  |
| Mass background model    | $< 10^{-3}$| 0.002              | $< 10^{-3}$| 0.002        | $< 10^{-3}$  | 0.002        | 0.027         | 0.038          | $< 10^{-3}$  |
| Time resolution model    | 0.003    | $< 10^{-3}$        | 0.001      | 0.002        | $< 10^{-3}$  | 0.002        | 0.057         | 0.011          | 0.001          |
| Default fit model        | 0.001    | 0.002              | $< 10^{-3}$| 0.002        | $< 10^{-3}$  | 0.002        | 0.025         | 0.015          | 0.005          |
| Total                    | 0.042    | 0.007              | 0.004      | 0.006        | 0.007        | 0.022        | 0.30          | 0.05           | 0.01           |

The systematic uncertainties are listed in table 7. For each parameter, the total systematic error is obtained by adding all of the contributions in quadrature.

8 Discussion

The PDF describing the $B^0_s \rightarrow J/\psi \phi$ decay is invariant under the following simultaneous transformations:

$$\{ \phi_s, \Delta \Gamma_s, \delta_{\perp}, \delta_{\parallel} \} \rightarrow \{ \pi - \phi_s, - \Delta \Gamma_s, \pi - \delta_{\perp}, 2\pi - \delta_{\parallel} \}.$$  

Since $\Delta \Gamma_s$ was determined to be positive [33], there is a unique solution. Figure 9 shows the 1D log-likelihood scans of $\phi_s$, $\Delta \Gamma_s$ and of the three measured strong phases $\delta_{\parallel}$, $\delta_{\perp}$ and $\delta_{\perp} - \delta_S$. The variable on vertical axis, $2\Delta \ln(L) \equiv 2(\ln(L^G) - \ln(L^i))$, is a difference between the likelihood values of a default fit, $(L^G)$, and of the fit in which the physical parameter is fixed to a value shown on horizontal axis, $(L^i)$. $2\Delta \ln(L) = 1$ corresponds to the estimated 1$\sigma$ confidence level. There are a small asymmetries in the likelihood curves, however at the level of one statistical $\sigma$ these are small compared to the corresponding statistical uncertainties of the physical variables, for which the scan is done. Therefore symmetric statistical uncertainties are quoted. Figure 10 shows the likelihood contours in the $\phi_s - \Delta \Gamma_s$ plane. The region predicted by the Standard Model is also shown.
Figure 9. 1D likelihood contours (statistical errors only) for $\phi_s$ (top left), $\Delta \Gamma_s$ (top centre), $\delta_\parallel$ (top right), $\delta_\perp$ (bottom left) and $\delta_\perp - \delta_S$ (bottom right).

Figure 10. Likelihood contours in the $\phi_s$–$\Delta \Gamma_s$ plane for 8 TeV data. The blue line shows the 68% likelihood contour, while the red dotted line shows the 95% likelihood contour (statistical errors only). The SM prediction is taken from ref. [1], at this scale the uncertainty on $\phi_s$ is not visible on the figure.
Figure 11. Likelihood contours in the $s - \Delta \Gamma_s$ plane for individual results from 7 TeV and 8 TeV data (left) and a final statistical combination of the results from 7 TeV and 8 TeV data (right). The blue line shows the 68% likelihood contour, while the red dotted line shows the 95% likelihood contour (statistical errors only). The SM prediction is taken from ref. [1], at this scale the uncertainty on $s$ is not visible on the figure.

9 Combination of 7 TeV and 8 TeV results

The measured values are consistent with those obtained in a previous analysis [8], using ATLAS data collected in 2011 at a centre-of-mass energy of 7 TeV. This consistency is also clear from a comparison of the likelihood contours in the $s - \Delta \Gamma_s$ projection shown in figure 11. A Best Linear Unbiased Estimate (BLUE) combination [34] is used to combine the 7 TeV and 8 TeV measurements to give an overall result for Run 1. In ref. [8] the strong phases $\delta_{\|}$ and $\delta_{\perp}$ were given as 1σ confidence intervals. These are not considered in the combination and the 8 TeV result is taken as the Run 1 result.

The BLUE combination requires the measured values and uncertainties of the parameters in question as well as the correlations between them. These are provided by the fits separately in the 7 TeV and 8 TeV measurements. The statistical correlation between these two measurements is zero as the events are different. The correlations of the systematic uncertainties between the two measurements are estimated by splitting the uncertainty into several categories.

The trigger efficiency is included as a systematic uncertainty only in the 7 TeV measurement, so there is no correlation with the 8 TeV measurement. Similarly, the systematic uncertainties arising from the $\Lambda_b \rightarrow J/\psi pK^-$ background, and the choice of $p_T$ bins and mass sidebands in the modelling of background angles, are included as systematic uncertainties only in the 8 TeV measurement so there is no correlation with the 7 TeV measurement. In both the 7 TeV and 8 TeV results, a systematic uncertainty is assigned to the inner detector alignment and $B_d$ contribution. The inner detector alignment systematic uncertainties are highly correlated and small. The assumed correlation between these systematics made no difference to the final combined result and was set to 100%. For the $B_d$
contribution, while the systematic uncertainty tests are different, they are both performed to account for an imprecise knowledge of the $B_d$ contribution and are therefore assumed to be 100%. The tagging, acceptance and fit model uncertainties are quoted for both 7 TeV and 8 TeV. For the fit model, there are several different model variations each with their own uncertainty. For each year, these are summed in quadrature to produce a single fit model systematic uncertainty.

The tagging, acceptance and fit model systematic uncertainties are each assigned a variable ($\rho_i$, where $i =$ tag, acc, mod) corresponding to the correlation between the 7 TeV and 8 TeV results. Several different combinations were tried with different values of $\rho_i = 0, 0.25, 0.5, 0.75, 1.0$. The acceptance systematic uncertainty is small and therefore regardless of what value of $\rho_{acc}$ is chosen the combination stays the same. For the 8 TeV measurement, electron tagging is added, therefore the systematic uncertainty is not 100% correlated. For $\rho_{tag} = 0.25, 0.5, 0.75$ there is negligible difference between the results. The fit model was changed between the 7 TeV and 8 TeV measurement, the most significant change is that the mass uncertainty modelling was removed and the event-by-event Gaussian error distribution was replaced with a sum of three Gaussian distributions. It would be incorrect to estimate the correlation as 100% and there is negligible difference between the results for $\rho_{mod} = 0.25, 0.5, 0.75$.

The combined results for the fit parameters and their uncertainties for Run 1 are given in table 8. Due to the negative correlation between $\Gamma_s$ and $\Delta \Gamma_s$, and the change in the value of $\Delta \Gamma_s$ between the 7 TeV and 8 TeV results, the combined value of $\Gamma_s$ is less than either individual result. The Run 1 likelihood contours in the $\phi_s$--$\Delta \Gamma_s$ plane are shown in figure 11. They agree with the Standard Model predictions.

<table>
<thead>
<tr>
<th>Par</th>
<th>8 TeV data</th>
<th>7 TeV data</th>
<th>Run1 combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td>Stat</td>
<td>Syst</td>
</tr>
<tr>
<td>$\phi_s$ [rad]</td>
<td>-0.110</td>
<td>0.082</td>
<td>0.042</td>
</tr>
<tr>
<td>$\Delta \Gamma_s$ [ps$^{-1}$]</td>
<td>0.101</td>
<td>0.013</td>
<td>0.007</td>
</tr>
<tr>
<td>$\Gamma_s$ [ps$^{-1}$]</td>
<td>0.676</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>$</td>
<td>A_| (0)</td>
<td>^2$</td>
<td>0.230</td>
</tr>
<tr>
<td>$</td>
<td>A_| (0)</td>
<td>^2$</td>
<td>0.520</td>
</tr>
<tr>
<td>$</td>
<td>A_S</td>
<td>^2$</td>
<td>0.097</td>
</tr>
<tr>
<td>$\delta_\perp$ [rad]</td>
<td>4.500</td>
<td>0.450</td>
<td>0.300</td>
</tr>
<tr>
<td>$\delta_\parallel$ [rad]</td>
<td>3.150</td>
<td>0.100</td>
<td>0.050</td>
</tr>
<tr>
<td>$\delta_| - \delta_S$ [rad]</td>
<td>-0.080</td>
<td>0.030</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 8. Current measurement using data from 8 TeV $pp$ collisions, the previous measurement using data taken at centre of mass energy of 7 TeV and the values for the parameters of the two measurements, statistically combined.
10 Summary

A measurement of the time-dependent $CP$ asymmetry parameters in $B^0_s \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ decays from a 14.3 fb$^{-1}$ data sample of $pp$ collisions collected with the ATLAS detector during the 8 TeV LHC run is presented. The values from the 8 TeV analysis are consistent with those obtained in the previous analysis using 7 TeV ATLAS data [8]. The two measurements are statistically combined leading to the following results:

\begin{align*}
\phi_s &= -0.090 \pm 0.078 \text{ (stat.)} \pm 0.041 \text{ (syst.)} \text{ rad} \\
\Delta \Gamma_s &= 0.085 \pm 0.011 \text{ (stat.)} \pm 0.007 \text{ (syst.)} \text{ ps}^{-1} \\
\Gamma_s &= 0.675 \pm 0.003 \text{ (stat.)} \pm 0.003 \text{ (syst.)} \text{ ps}^{-1} \\
|A_\parallel(0)|^2 &= 0.227 \pm 0.004 \text{ (stat.)} \pm 0.006 \text{ (syst.)} \\
|A_0(0)|^2 &= 0.522 \pm 0.003 \text{ (stat.)} \pm 0.007 \text{ (syst.)} \\
|A_S(0)|^2 &= 0.072 \pm 0.007 \text{ (stat.)} \pm 0.018 \text{ (syst.)} \\
\delta_\perp &= 4.15 \pm 0.32 \text{ (stat.)} \pm 0.16 \text{ (syst.)} \text{ rad} \\
\delta_\parallel &= 3.15 \pm 0.10 \text{ (stat.)} \pm 0.05 \text{ (syst.)} \text{ rad} \\
\delta_\perp - \delta_\parallel &= -0.08 \pm 0.03 \text{ (stat.)} \pm 0.01 \text{ (syst.)} \text{ rad}.
\end{align*}

The ATLAS Run 1 results for the $B^0_s \to J/\psi\phi$ decay are consistent with the SM.

Acknowledgments

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; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, 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, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes cofinanced by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway;
Generalitat de Catalunya, Generalitat Valenciana, Spain; 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 (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [35].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[4] D0 collaboration, V.M. Abazov et al., Measurement of the CP-violating phase $\phi_+^{J/\psi\phi}$ using the flavor-tagged decay $B_0^+ \to J/\psi\phi$ in 8 fb$^{-1}$ of $p\bar{p}$ collisions, Phys. Rev. D 85 (2012) 032006 [arXiv:1109.3166] [inSPIRE].
[6] LHCb collaboration, Measurement of CP-violation and the $B_s$ meson decay width difference with $B_0^+ \to J/\psi K^+ K^-$ and $B_0^0 \to J/\psi \pi^+ \pi^-$ decays, Phys. Rev. D 87 (2013) 112010 [arXiv:1304.2600] [inSPIRE].
[8] ATLAS collaboration, Flavor tagged time-dependent angular analysis of the $B_s \to J/\psi\phi$ decay and extraction of $\Delta\Gamma_s$ and the weak phase $\phi_s$ in ATLAS, Phys. Rev. D 90 (2014) 052007 [arXiv:1407.1796] [inSPIRE].
[9] CMS collaboration, Measurement of the CP-violating weak phase $\phi_s$ and the decay width difference $\Delta\Gamma_s$ using the $B_0^0 \to J/\psi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV, Phys. Lett. B 757 (2016) 97 [arXiv:1507.07527] [inSPIRE].


[23] ATLAS collaboration, *Study of the kinematic dependences of \Lambda^0_b production in pp collisions and a measurement of the \Lambda^0_b \to \Lambda^+_c \pi^- branching fraction*, JHEP 08 (2014) 143 [arXiv:1405.6842] [inSPIRE].
[30] LHCb collaboration, Study of the production of $\Lambda_b^0$ and $B^0$ hadrons in pp collisions and first measurement of the $\Lambda_b^0 \to J/\psi pK^-$ branching fraction, Chin. Phys. C 40 (2016) 011001 [arXiv:1509.00292] [INSPIRE].


The ATLAS collaboration

Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro; (b) Electrical Circuits Department, Federal University of Juiz de Fora (UFJF), Juiz de Fora; (c) Federal University of Sao Joao del Rei (UFSJ), Sao Joao del Rei; (d) Instituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil

Physics Department, Brookhaven National Laboratory, Upton NY, United States of America

(a) National Institute of Physics and Nuclear Engineering, Bucharest; (b) National Institute for Research and Development of Isotropic and Molecular Technologies, Physics Department, Cluj Napoca; (c) University Politehnica Bucharest, Bucharest; (d) West University in Timisoara, Timisoara, Romania

Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina

Cavendish Laboratory, University of Cambridge, Cambridge, United Kingdom

Department of Physics, Carleton University, Ottawa ON, Canada

CERN, Geneva, Switzerland

Eurico Fermi Institute, University of Chicago, Chicago IL, United States of America

(a) Departamento de Física, Pontificia Universidad Católica de Chile, Santiago; (b) Departamento de Física, Universidad Técnica Federico Santa María, Valparaíso, Chile

Institute of High Energy Physics, Chinese Academy of Sciences, Beijing; (b) Department of Modern Physics, University of Science and Technology of China, Anhui; (c) Department of Physics, Nanjing University, Jiangsu; (d) School of Physics, Shandong University, Shandong; (e) Department of Physics and Astronomy, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University, Shanghai; (f) Physics Department, Tsinghua University, Beijing 100084, China

Laboratoire de Physique Corpusculaire, Clermont Université and Université Blaise Pascal and CNRS/IN2P3, Clermont-Ferrand, France

Nevis Laboratory, Columbia University, Irvington NY, United States of America

Niels Bohr Institute, University of Copenhagen, København, Denmark

(a) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati; (b) Dipartimento di Fisica, Università della Calabria, Rende, Italy

(a) AGH University of Science and Technology, Faculty of Physics and Applied Computer Science, Krakow; (b) Marian Smoluchowski Institute of Physics, Jagiellonian University, Krakow, Poland

Institute of Nuclear Physics Polish Academy of Sciences, Krakow, Poland

Physics Department, Southern Methodist University, Dallas TX, United States of America

Physics Department, University of Texas at Dallas, Richardson TX, United States of America

DESY, Hamburg and Zeuthen, Germany

Institut für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany

Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany

Department of Physics, Duke University, Durham NC, United States of America

SUPA - School of Physics and Astronomy, University of Edinburgh, Edinburgh, United Kingdom

INFN Laboratori Nazionali di Frascati, Frascati, Italy

Fakultät für Mathematik und Physik, Albert-Ludwigs-Universität, Freiburg, Germany

Section de Physique, Université de Genève, Geneva, Switzerland

(a) INFN Sezione di Genova; (b) Dipartimento di Fisica, Università di Genova, Genova, Italy

(a) E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi; (b) High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia

II Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany

SUPA - School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom

II Physikalisches Institut, Georg-August-Universität, Göttingen, Germany

Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, Grenoble, France

Department of Physics, Hampton University, Hampton VA, United States of America

Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge MA, United States of America
Department of Physics, University of Wisconsin, Madison WI, United States of America
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität, Würzburg, Germany
Fachbereich C Physik, Bergische Universität Wuppertal, Wuppertal, Germany
Department of Physics, Yale University, New Haven CT, United States of America
Yerevan Physics Institute, Yerevan, Armenia
Centre de Calcul de l’Institut National de Physique Nucléaire et de Physique des Particules (IN2P3), Villeurbanne, France

a Also at Department of Physics, King’s College London, London, United Kingdom
b Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
c Also at Novosibirsk State University, Novosibirsk, Russia
d Also at TRIUMF, Vancouver BC, Canada
e Also at Department of Physics, California State University, Fresno CA, United States of America
f Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
g Also at Departamento de Fisica e Astronomia, Faculdade de Ciencias, Universidade do Porto, Portugal
h Also at Tomsk State University, Tomsk, Russia
i Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
j Also at Università di Napoli Parthenope, Napoli, Italy
k Also at Institute of Particle Physics (IPP), Canada
l Also at Particle Physics Department, Rutherford Appleton Laboratory, Didcot, United Kingdom
m Also at Department of Physics, St. Petersburg State Polytechnical University, St. Petersburg, Russia
n Also at Louisiana Tech University, Ruston LA, United States of America
o Also at Institutio Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain
p Also at Graduate School of Science, Osaka University, Osaka, Japan
q Also at Department of Physics, National Tsing Hua University, Taiwan
r Also at Department of Physics, The University of Texas at Austin, Austin TX, United States of America
s Also at Institute of Theoretical Physics, Ilia State University, Tbilisi, Georgia
t Also at CERN, Geneva, Switzerland
u Also at Georgian Technical University (GTU), Tbilisi, Georgia
v Also at Manhattan College, New York NY, United States of America
w Also at Hellenic Open University, Patras, Greece
x Also at Institute of Physics, Academia Sinica, Taipei, Taiwan
y Also at LAL, Université Paris-Sud and CNRS/IN2P3, Orsay, France
z Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan
aa Also at School of Physics, Shandong University, Shandong, China
ab Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
ac Also at Section de Physique, Université de Genève, Geneva, Switzerland
ad Also at International School for Advanced Studies (SISSA), Trieste, Italy
ae Also at Department of Physics and Astronomy, University of South Carolina, Columbia SC, United States of America
af Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China
ag Also at Faculty of Physics, M.V.Lomonosov Moscow State University, Moscow, Russia
ah Also at National Research Nuclear University MEPhI, Moscow, Russia
ai Also at Department of Physics, Stanford University, Stanford CA, United States of America
aj Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
ak Also at Department of Physics, The University of Michigan, Ann Arbor MI, United States of America
al Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia
* Deceased