Not reviewed, for internal circulation only

ATLAS Paper Draft

BPHY-2018-01

Version 2.0

Comments are due by: 20th October 2019

Supporting internal notes

Internal documentation: https://cds.cern.ch/record/2628521
Supersede ATLAS-CONF-2019-009. List of changes : https://cds.cern.ch/record/2693198

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

A measurement of the $B_s^0 \to J/\psi\phi$ decay parameters using $80.5\ \text{fb}^{-1}$ 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 ϕ_s, the width difference $\Delta\Gamma_s$ between the B_s^0 meson mass eigenstates and the average decay width Γ_s. The values measured for the physical parameters are combined with those from $19.2\ \text{fb}^{-1}$ of 7 TeV and 8 TeV data, leading to the following:

$$\begin{aligned}\phi_s &= -0.096 \pm 0.036\ (\text{stat.}) \pm 0.024\ (\text{syst.})\ \text{rad}\\ \Delta\Gamma_s &= 0.0696 \pm 0.0042\ (\text{stat.}) \pm 0.0029\ (\text{syst.})\ \text{ps}^{-1}\\ \Gamma_s &= 0.6684 \pm 0.0014\ (\text{stat.}) \pm 0.0018\ (\text{syst.})\ \text{ps}^{-1}\end{aligned}$$

Results for ϕ_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.

To be submitted to: EPJC

Document created on 11th October 2019 using ATLAS LaTeX Version 05-03-00.

Analysis Team

[*email:* atlas-BPHY-2018-01-editors@cern.ch]

Adam Barton, Alex Cerri, Tomas Jakoubek, Roger Jones, Artem Maevskiy, Oleg Meshkov, Vladimir Nikolaenko, Lukas Novotny, Radek Novotny, Sandro Palestini, Pavel Reznicek, Sinem Simsek, Maria Smizanska, Emin Tagiev, James Walder

Language editor: Lily Asquith lily.asquith@cern.ch

Not reviewed, for internal circulation only

ATLAS Paper

BPHY-2018-01

11th October 2019

Draft version 2.0

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

The ATLAS Collaboration

A measurement of the $B_s^0 \to J/\psi\phi$ 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 ϕ_s, the width difference $\Delta\Gamma_s$ between the B_s^0 meson mass eigenstates and the average decay width Γ_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:

$$\begin{aligned}
\phi_s &= -0.096 \pm 0.036\ \text{(stat.)} \pm 0.024\ \text{(syst.) rad} \\
\Delta\Gamma_s &= 0.0696 \pm 0.0042\ \text{(stat.)} \pm 0.0029\ \text{(syst.) ps}^{-1} \\
\Gamma_s &= 0.6684 \pm 0.0014\ \text{(stat.)} \pm 0.0018\ \text{(syst.) ps}^{-1}
\end{aligned}$$

Results for ϕ_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.

To be submitted to: *EPJC*

Not reviewed, for internal circulation only

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 $B_s^0 \to J/\psi\phi$ decay, CP violation occurs due to interference between a direct decay and a decay with $B_s^0 - \bar{B}_s^0$ mixing. The oscillation frequency of B_s^0 meson mixing is characterised by the mass difference Δm_s of the heavy (B_{H}) and light (B_{L}) mass eigenstates. The CP violating phase ϕ_s is defined as the weak phase difference between the $B_s^0 - \bar{B}_s^0$ mixing amplitude and the $b \to c\bar{c}s$ decay amplitude. In the SM the phase ϕ_s is small and is related to Cabibbo–Kobayashi–Maskawa (CKM) quark mixing matrix elements via the relation $\phi_s \simeq -2\beta_s$, with $\beta_s = \arg[-(V_{ts}V_{tb}^*)/(V_{cs}V_{cb}^*)]$; assuming no NP 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 [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 ϕ_s with respect to the SM value.

Other physical quantities involved in $B_s^0 - \bar{B}_s^0$ mixing are the decay width $\Gamma_s = (\Gamma_{\mathrm{L}} + \Gamma_{\mathrm{H}})/2$ and the width difference $\Delta\Gamma_s = \Gamma_{\mathrm{L}} - \Gamma_{\mathrm{H}}$, where Γ_{L} and Γ_{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 ϕ_s would also decrease the size of $\Delta\Gamma_s$, however it is not expected to be affected as significantly as ϕ_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 $B_s^0 \to J/\psi\phi$ decay provides the most precise determination of ϕ_s and $\Delta\Gamma_s$. Previous measurements of these quantities have been reported by the D0 , CDF, LHCb, ATLAS and CMS collaborations [**Aad:2016tdj**, 6–10]. Additional improvements in measuring ϕ_s from B_s^0 decays to $\psi(2S)\phi$ and to $D_s^+D_s^-$ have been achieved by the LHCb collaboration [**Aaij:2016psitwoS**, **Aaij:2014Ds**].

The analysis presented here introduces a measurement of the $B_s^0 \to J/\psi\phi$ 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 [**Aad:2016tdj**] 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

* 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, ϕ) are used in the transverse plane, r being the distance from the origin and ϕ being the azimuthal angle around the beam pipe. The pseudorapidity η is defined as $\eta = -\ln[\tan(\theta/2)]$ where θ is the polar angle.

calorimeters for electromagnetic and hadronic measurements. The MS surrounds the calorimeters and provides a system of tracking chambers and detectors for triggering. A full description can be found in Refs. [11–13].

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 triggers were based on the identification of a $J/\psi \to \mu^+\mu^-$ decay, with transverse momentum (p_{T}) thresholds of either 4 GeV or 6 GeV for the muons.

The measurement uses 80.5 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. [14], and using the LUCID-2 detector for the baseline luminosity measurements [15], from calibration of the luminosity scale using x–y beam-separation scans.

To study the detector response, estimate backgrounds, and model systematic effects, 100M Monte Carlo (MC) simulated $B_s^0 \to J/\psi\phi$ events were generated using Pythia 8.210 [16] tuned with ATLAS data, using the A14 set of parameters [17] together with the CTEQ6L1 set [18]. The detector response was simulated using the ATLAS simulation framework based on GEANT4 [19, 20]. 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. Additionally, the background samples of both exclusive ($B_d^0 \to J/\psi K^{0*}$ and $\Lambda_b \to J/\psi pK^-$) and inclusive ($b\bar{b} \to J/\psi X$ and $pp \to J/\psi X$) decays were simulated, using the same simulation tools as in case of the signal events. For validation studies related to *flavour tagging*, detailed in 4, events of $B^\pm \to J/\psi K^+$ exclusive decays were also simulated.

3 Reconstruction and candidate selection

The reconstruction and candidate selection for the decay $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ is described here. 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 χ^2/n.d.o.f. < 10. In order to account for varying mass resolution in different parts of the detector, the J/ψ candidates are divided into three subsets according to the pseudorapidity η 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/ψ mass and the corresponding mass resolution for these three subsets, and in each case the signal region is defined symmetrically around the fitted mass, so as to retain 99.7% of the J/ψ candidates identified in the fits.

The candidates for the decay $\phi \to K^+K^-$ are reconstructed from all pairs of oppositely charged tracks, with $p_{\mathrm{T}} > 1$ GeV and $|\eta| < 2.5$, that are not identified as muons. Candidate events for $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ decays are selected by fitting the tracks for each combination of $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$ to a common vertex. The fit is also constrained by fixing the invariant mass calculated from the two muon tracks to the J/ψ mass [21]. A quadruplet of tracks is accepted for further analysis if the vertex fit has χ^2/n.d.o.f. < 3. For the $\phi \to K^+K^-$ candidate, the invariant mass of the track pairs (using a kaon mass hypothesis) must fall within the interval 1.0085 GeV $< m(K^+K^-) < 1.0305$ GeV. The interval, chosen using MC simulation, is selected to retain 98% of true $\phi \to K^+K^-$ decays. The B_s^0 candidate with the lowest χ^2/n.d.o.f. is selected in cases where more than one candidate passes all selections. In total, 2 977 526 B_s^0 candidates are collected within the mass range of 5.150–5.650 GeV. This range is chosen to give enough background events in the side bands of the mass distributions to allow a high precision determination of the properties of background events. The mass window choice has been varied and found to have a negligible systematic effect on the results.

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

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

$$t = \frac{L_{xy}\, m_B}{p_{\mathrm{T}_B}},$$

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 Ref. [21]. 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 primary vertex position is recalculated after removing any tracks used in the B_s^0 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 (ie. b- or $\bar{b}$-quark) of the opposite-side b-hadron. The algorithms thus provide a probability that a signal B meson in a given event is produced in a given flavour. The calibration of the OST algorithms proceeds using data containing $B^\pm \to J/\psi K^\pm$ 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 $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ candidate events to provide a probability that each candidate was produced in a B_s^0 or $\bar{B}_s^0$ meson state, which is used in the maximum likelihood fit (described in Section 5). Section 4.1 describes the reconstruction of the $B^\pm \to J/\psi K^\pm$ 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, are discussed in Section 4.4.

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

Candidate $B^\pm \to J/\psi K^\pm$ decays are identified in a series of steps described here. First, J/ψ candidates are selected from oppositely charged muon pairs forming a good vertex, as described in Section 3. Each muon is required to have $p_{\mathrm{T}} > 4$ GeV and $|\eta| < 2.5$. 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/ψ mass [21]. Prompt background contributions are suppressed with the requirement on the proper decay time of the $B^\pm$ candidate of $t > 0.2$ ps.

The tagging probabilities are determined from B^+ and B^- signal events. These signal yields are derived from fits to the invariant mass distribution, $m(J/\psi K^\pm)$, and performed in intervals of the discriminating variables. To describe the $B^\pm \to J/\psi K^\pm$ 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 parametrise the low-mass contribution from incorrectly or partially reconstructed b-hadron decays. A Gaussian function is used to describe the $B^\pm \to J/\psi \pi^\pm$ 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 the point of production is inferred using several methods, which differ in their efficiency and discrimination power. 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 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 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:

$$Q_x = \frac{\sum_i^{N \text{ tracks}} q_i \cdot (p_{\mathrm{T}i})^\kappa}{\sum_i^{N \text{ tracks}} (p_{\mathrm{T}i})^\kappa}, \tag{1}$$

where $x = \{\mu, e, \text{jet}\}$ refers to muon, electron, or jet charge, respectively, and the summation is made over a selected set of tracks — including the lepton — in a cone, $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, around the lepton or

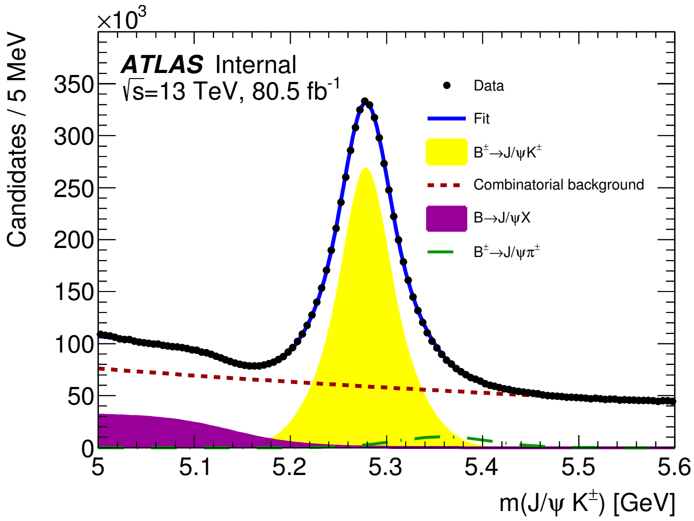

Figure 1: The invariant mass distribution for selected $B^{\pm} \to J/\psi K^{\pm}$ candidates. Data are shown as points, and the overall result of the fit is given by the blue curve. The contributions from the combinatorial background component are indicated by the red dotted line, partially reconstructed b-hadron decays by the purple shaded area, and decays of $B^{\pm} \to J/\psi\pi^{\pm}$, where the pion is misassigned as a kaon, by the green dashed line.

jet direction. The requirements on the tracks and ΔR are described below, dependent on the OST method. Two sub-categories of Q_x are considered: the first *discrete* category is used in the case where the cone charge is formed either from only one track or from more than one track of the same charge; this results in a cone charge of $Q_x = \pm 1$. The second *continuous* category is used when more than one track is considered, and the sum contains 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 $P(B|Q_x)$ is constructed, which is defined as the probability that a B meson is produced in a state containing a $\bar{b}$-quark, given the value of the cone charge Q_x. An equivalent probability for the b-quark case is defined as $P(\bar{B}|Q_x)$. 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_s^0 meson as containing a $\bar{b}$-quark is therefore given as $P(B|Q_x) = P(Q_x|B^+)/(P(Q_x|B^+) + P(Q_x|B^-))$, and correspondingly $P(\bar{B}|Q_x) = 1 - P(B|Q_x)$. If there is no OST information available for a given B_s^0 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_s^0 meson candidates to infer the initial flavour.

Muon tagging

For muon-based tagging, at least one additional muon is required in the event, with $p_{\mathrm{T}} > 2.5$ GeV, $|\eta| < 2.5$ and with $|\Delta z| < 5$ mm, where $|\Delta z|$ is the difference in z between the primary vertex and the longitudinal impact parameter of the ID track associated to the muon. Muons are classified and kept if their identification quality selection working point is either *Tight*[†] or *Low-p_T*;[‡] these categories are subsequently treated as distinct flavour tagging methods. For muons with $p_T > 4$ GeV, Tight muons are the dominant category, with the Low-p_T requirement typically identifying muons of $p_T < 4$ GeV. In the case of multiple muons passing selection criteria in one event, Tight muons are chosen over Low-p_T muons. Within the same muon category, the muon with the highest p_T that passes the selections is used.

A muon cone charge variable, Q_μ, is constructed in the same way as Eq. 1, with $\kappa = 1.1$ and the sum over the reconstructed ID tracks within a cone, $\Delta R = 0.5$, around the muon direction. These tracks must have $p_T > 0.5$ GeV, $|\eta| < 2.5$, and $|\Delta z| < 5$ mm. Tracks associated with the decay of a B meson signal candidate are excluded from the sum. In each interval of Q_μ, a fit to the $J/\psi K^\pm$ invariant mass spectrum is performed and the number of signal events extracted. The fit model used is described in Section 4.1. Figure 2 shows the distributions of the muon cone charge using $B^\pm$ signal candidates for Tight muons, and includes the tagging probability as a function of the cone charge variable. The corresponding distributions for Low-p_T muons are shown in Figure 3.

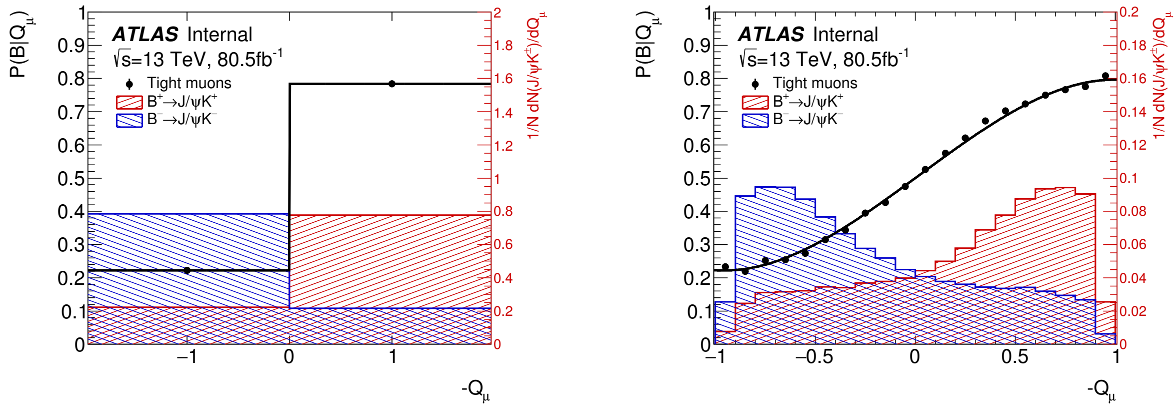

Figure 2: Cone charge distributions, Q_μ, for Tight muons, shown for cases of discrete charge (left), and for the continuous distribution (right). For each plot, in red (blue), the normalised B^+ (B^-) cone charge distribution is shown (corresponding to the right axis scale). Superimposed is the distribution of the tagging probability, $P(B|Q_\mu)$, as a function of the cone charge, derived from a data sample of $B^\pm \to J/\psi K^\pm$, and defined as the probability to have a B^+ meson (on the signal-side) given a particular cone charge Q_μ. 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 $J/\psi\phi$.

Electron tagging

Electrons are identified using inner detector and calorimeter information, which satisfy the *Medium* electron quality criteria [23]. 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, ζ_b, of $\cos(\zeta_b) > 0.93$ are not

[†] 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. [22].

[‡] This working point is optimized to provide good muon reconstruction efficiency down to a p_T of ≈ 3 GeV, while controlling the fake rate. It allows ≥ 1 (≥ 2) MDT station tracks up to $|\eta| < 1.3$ ($1.3 < |\eta| < 1.55$) for candidates reconstructed by algorithms utilizing inside-out combined reconstruction [22]. 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.

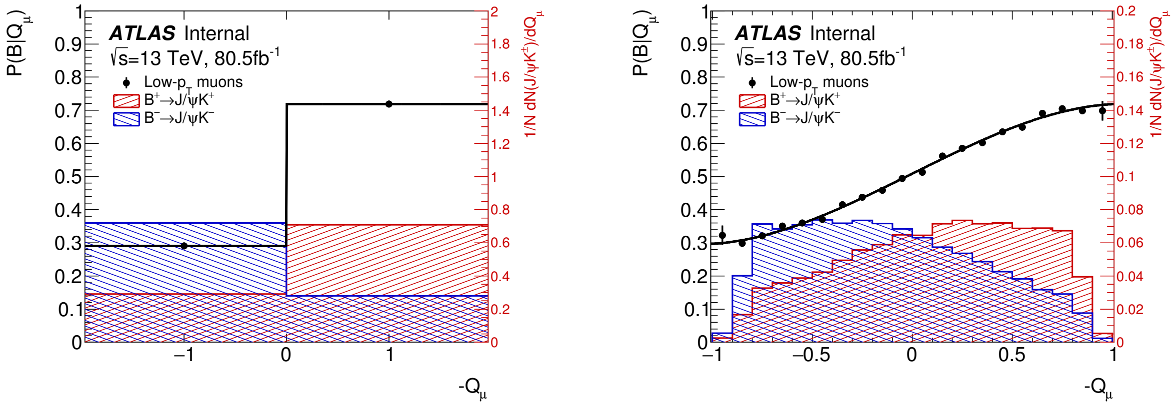

Figure 3: Normalised cone charge distributions, Q_μ, for B^+ (B^-) events shown in red (blue) for Low-p_T muons, for cases of discrete charge (left), and for the continuous distribution (right). Superimposed is the distribution of the tagging probability, $P(B|Q_\mu)$.

considered. In the case of more than one electron passing the selection, the electron with the highest p_T is chosen. Charged-particle tracks within a cone of size $\Delta R = 0.5$ are used to form the electron cone charge Q_e, constructed in the same way as Eq. 1, with $\kappa = 1.0$. The resulting electron cone charge distributions are shown in Figure 4, together with the corresponding tagging probability.

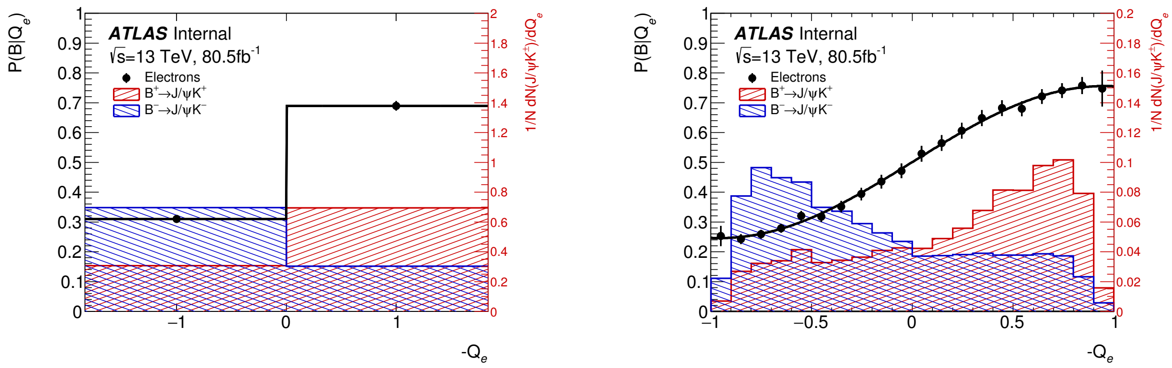

Figure 4: Normalised cone charge distributions, Q_e, for B^+ (B^-) events shown in red (blue) for electrons, for cases of discrete charge (left), and the continuous distribution (right). Superimposed is the distribution of the tagging probabilities, $P(B|Q_e)$.

Jet tagging

In the absence of a muon or electron, a jet identified as containing a b-hadron is required. Jets are reconstructed from calorimetric information using the anti-k_t algorithm [24, 25] with a radius parameter $R = 0.4$. The identification of a b-tagged jet uses a multivariate algorithm *MV2c10* [26], utilising boosted decision trees (BDT), which output a classifier value. Jets are selected that exceed the BDT classifier output value of 0.56. This value is optimised to maximise 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.

Tracks within a cone of size $\Delta R = 0.5$ of the jet axis are used to define a jet cone charge, Q_{jet},constructed in the same way as Eq. 1, where $\kappa = 1.1$ and the sum is over the tracks associated with the jet, with $|\Delta z| < 5$ mm, and excluding tracks from the decay of the signal B meson candidate. The signal yields are extracted from the data from fits to the $J/\psi K^{\pm}$ invariant mass spectrum, using the same procedure described for muon tagging. Figure 5 shows the distribution of the opposite side jet cone charge for $B^{\pm}$ signal candidates.

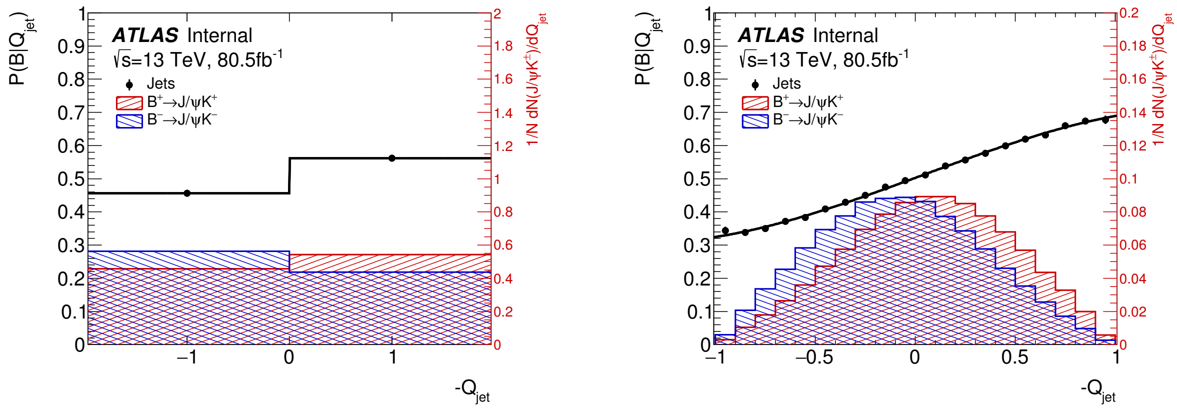

Figure 5: Normalised cone charge distributions, Q_{jet}, for B^+ (B^-) events shown in red (blue) for jets, for cases of discrete charge (left), and the continuous distribution (right). Superimposed is the distribution of the tag probability, $P(B|Q_{\text{jet}})$.

4.3 Flavour tagging performance

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

By definition, there is no overlap between lepton-tagged and jet-charge-tagged events. The overlap between events with 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.

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

Tag method	ϵ_x [%]	D_x [%]	T_x [%]
Tight muon	4.50 ± 0.01	43.8 ± 0.2	0.862 ± 0.009
Electron	1.57 ± 0.01	41.8 ± 0.2	0.274 ± 0.004
Low-p_{T} muon	3.12 ± 0.01	29.9 ± 0.2	0.278 ± 0.006
Jet	12.04 ± 0.02	16.6 ± 0.1	0.334 ± 0.006
Total	21.23 ± 0.03	28.7 ± 0.1	1.75 ± 0.01

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 (versus a $\bar{B}_s^0$) is provided using the calibrations derived from the $B^{\pm} \to J/\psi K^{\pm}$ sample, described above, and shown in Figures 2–5. As the distributions of $P(B|Q_x)$ from signal B_s^0 mesons and background data can be expected to be different, separate probability density functions (PDFs) are necessary to describe these distributions in the likelihood function. These PDFs are defined as $P_{\mathrm{s}}(P(B|Q_x))$ and $P_{\mathrm{b}}(P(B|Q_x))$, describing the probability distributions for signal and background, respectively, and are extracted using the sample of B_s^0 candidates. The PDFs consist of the fraction of events that are tagged with a particular method (or are untagged), the fractions of those events categorised as discrete or continuous, and for those that are continuous, a PDF of the corresponding probability distribution.

Continuous PDF

The parametrisations of the continuous PDF components of $P_{\mathrm{s,b}}(P(B|Q_x))$ for each OST method are defined as follows: In the sideband regions, $5.150 < m(J/\psi KK) < 5.317$ GeV and $5.417 < m(J/\psi KK) < 5.650$ GeV, unbinned maximum likelihood fits are performed to $P(B|Q_x)$ distributions to extract the background (continuous category) PDFs for $P_{\mathrm{b}}(P(B|Q_x))$. For the Tight muon and electron methods, the parametrisation 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 $P(B|Q_x)$ distributions, using all events in the $m(J/\psi KK)$ distributions to extract the signal (continuous category) PDFs for $P_{\mathrm{s}}(P(B|Q_x))$. In these fits, the previously extracted parameters to describe the background 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 the electron tagging method, the signal function has the form of the sum of a second-order polynomial and two exponential functions, and for the Low-p_{T} muon and jet tagging methods a Gaussian function is used.

Discrete PDF

In the case where the cone charge is discrete, the fractions of events f_{+1} (f_{-1}) with cone charges $+1$ (-1) are determined separately for signal and background using events from the signal and sideband regions of the B_s^0 mass distribution (as defined in Section 3). The remaining fraction of events, $1 - f_{+1} - f_{-1}$, corresponds to the continuous parts of the distribution. Positive and negative charges are equally probable for background candidates formed from a random combination of a J/ψ and a pair of tracks, but this is not necessarily the case for background candidates formed from a partially reconstructed b-hadron. Table 2 summarises for the different tagg methods, the fractions f_{+1} and f_{-1} obtained for signal and background events.

Table 2: Fractions f_{+1} and f_{-1} of events with cone charges of $+1$ and -1 respectively, for signal and background events and for the different tagging methods. Only statistical uncertainties are given.

Tag method	Signal		Background	
	f_{+1} [%]	f_{-1} [%]	f_{+1} [%]	f_{-1} [%]
Tight muon	6.9 ± 0.3	7.5 ± 0.3	4.7 ± 0.1	4.9 ± 0.1
Electron	20 ± 1	19 ± 1	16.8 ± 0.2	17.3 ± 0.2
Low-p_{T} muon	10.9 ± 0.5	11.6 ± 0.5	7.0 ± 0.1	7.5 ± 0.1
Jet	3.60 ± 0.15	3.54 ± 0.15	3.05 ± 0.03	3.17 ± 0.03

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

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

Tag method	Signal efficiency [%]	Background efficiency [%]
Tight muon	4.06 ± 0.06	3.21 ± 0.01
Electron	1.86 ± 0.04	1.48 ± 0.01
Low-p_{T} muon	2.95 ± 0.05	2.70 ± 0.01
Jet	12.1 ± 0.1	9.41 ± 0.02
Untagged	79.1 ± 0.3	83.20 ± 0.05

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

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 σ_t, the tagging probability, $P(B|Q_x)$, and the transversity angles Ω of each $B_s^0 \to J/\psi\phi$ decay candidate. The measured proper decay time uncertainty σ_t is calculated from the covariance matrix associated with the vertex fit for each candidate event. The transversity angles $\Omega = (\theta_T, \psi_T, \phi_T)$ are defined in Section 5.1. The likelihood function is defined as a combination of the signal and background PDFs as follows:

$$
\begin{aligned}
\ln \mathcal{L} = \sum_{i=1}^{N} w_i \cdot \ln[f_{\mathrm{s}} \cdot \mathcal{F}_{\mathrm{s}}(m_i, t_i, \sigma_{m_i}, \sigma_{t_i}, \Omega_i, P_i(B|Q_x), p_{\mathrm{T}_i}) + f_{\mathrm{s}} \cdot f_{B^0} \cdot \mathcal{F}_{B^0}(m_i, t_i, \sigma_{m_i}, \sigma_{t_i}, \Omega_i, P_i(B|Q_x), p_{\mathrm{T}_i}) \\
+ f_{\mathrm{s}} \cdot f_{\Lambda_b} \cdot \mathcal{F}_{\Lambda_b}(m_i, t_i, \sigma_{m_i}, \sigma_{t_i}, \Omega_i, P_i(B|Q_x), p_{\mathrm{T}_i}) \\
+ (1 - f_{\mathrm{s}} \cdot (1 + f_{B^0} + f_{\Lambda_b}))\mathcal{F}_{\mathrm{bkg}}(m_i, t_i, \sigma_{m_i}, \sigma_{t_i}, \Omega_i, P_i(B|Q_x), p_{\mathrm{T}_i})],
\end{aligned}
\tag{2}
$$

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). $\mathcal{F}_{\mathrm{s}}$, $\mathcal{F}_{B^0}$, $\mathcal{F}_{\Lambda_b}$ and $\mathcal{F}_{\mathrm{bkg}}$ are the PDFs modelling the signal, B^0 background, Λ_b background, and the other background distributions, respectively. The term f_{s} is the fraction of signal candidates and f_{B^0} and f_{Λ_b} are the background fractions of B^0 mesons and Λ_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 variations are applied as part of the evaluation of the effects of systematic uncertainties. The mass m_i, the proper decay time t_i and the decay angles Ω_i are the values measured from the data for each event i. A detailed description of the signal PDF terms in Eq. (2) is given in 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, $\mathcal{F}_{\mathrm{s}}$, has the following composition:

$$
\begin{aligned}
\mathcal{F}_{\mathrm{s}}(m_i, t_i, \sigma_{m_i}, \sigma_{t_i}, \Omega_i, P_i(B|Q_x), p_{\mathrm{T}_i}) &= P_{\mathrm{s}}(m_i, \sigma_{m_i}) \cdot P_{\mathrm{s}}(\sigma_{m_i}) \cdot P_{\mathrm{s}}(\Omega_i, t_i, P_i(B|Q_x), \sigma_{t_i}) \\
&\cdot P_{\mathrm{s}}(\sigma_{t_i}) \cdot P_{\mathrm{s}}(P_i(B|Q_x)) \cdot A(\Omega_i, p_{\mathrm{T}_i}) \cdot P_{\mathrm{s}}(p_{\mathrm{T}_i}).
\end{aligned}
$$

The mass term $P_{\mathrm{s}}(m_i, \sigma_{m_i})$ is modelled in the following way:

$$
P_{\mathrm{s}}(m_i, \sigma_{m_i}) \equiv \frac{1}{\sqrt{2\pi} S_m \sigma_{m_i}} \cdot \mathrm{e}^{\frac{-(m_i - m_{B_s})^2}{2(S_m \sigma_{m_i})^2}}. \tag{3}
$$

The term $P_{\mathrm{s}}(m_i, \sigma_{m_i})$ uses per-candidate mass errors, σ_{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 σ_{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 $P_{\mathrm{s}}(\sigma_{m_i})$, $P_{\mathrm{s}}(\sigma_{t_i})$ and $P_{\mathrm{s}}(p_{\mathrm{T}i})$ 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_i} 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. [27]. The tagging probability term for signal $P_{\mathrm{s}}(P_i(B|Q_x))$ is described in Section 4.4. The term $A(\Omega_i, p_{\mathrm{T}_i})$ is the acceptance function, described at the end of the current Section 5.1.

The term $P_s(\Omega_i, t_i, P_i(B|Q_x), \sigma_{t_i})$ is a joint PDF for the decay time t and the transversity angles Ω 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 Ω is given by the differential decay rate [28]:

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

where $O^{(k)}(t)$ are the time-dependent functions corresponding to the contributions of the four different amplitudes (A_0, $A_{||}$, $A_\perp$, and A_S) and their interference terms, and $g^{(k)}(\theta_T, \psi_T, \phi_T)$ are the angular functions. Table 4 shows 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 Δm_s, which is a fixed parameter of the fit (using Ref. [21]). 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 $B_s^0 \to J/\psi K^+K^-$ 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 parametrised by $|A_i|e^{i\delta_i}$, where $i = \{0, ||, \perp, S\}$, with $\delta_0 = 0$ and are normalised such that $|A_0(0)|^2 + |A_\perp(0)|^2 + |A_\|(0)|^2 = 1$. $|A_\perp(0)|$ is determined according to this condition, while the remaining three amplitudes are parameters of the fit. The phase δ_S is the phase difference between $A_S(0)$ and $A_0(0)$ at the resonance peak. $|A_S|^2$ gives the ratio of non-resonant over resonant yield in the interval of $m(K^+K^-)$ used in the analysis. In the sum over the mass interval, the interference terms (lines 8 – 10 in Table 4) are corrected by a factor $\alpha = 0.51 \pm 0.02$ that takes into account the mass–dependent differences in absolute amplitude and phase between the resonant and the S–wave amplitudes. The correction is based on the Breit-Wigner description of the resonance and on the assumption of uniform A_S. The uncertainty on the value of α has been calculated based on the Flatté parametrisation [**Flatte:1976xu**] 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 ϕ meson in the J/ψ rest frame, and the K^+K^- system defines the x–y plane, where $p_y(K^+) > 0$. The three angles are defined as:

- θ_T, the angle between $\vec{p}(\mu^+)$ and the normal to the x–y plane, in the J/ψ meson rest frame,
- ϕ_T, the angle between the x-axis and $\vec{p}_{xy}(\mu^+)$, the projection of the μ^+ momentum in the x–y plane, in the J/ψ meson rest frame,
- ψ_T, the angle between $\vec{p}(K^+)$ and $-\vec{p}(J/\psi)$ in the ϕ meson rest frame.

The PDF term $P_s(\Omega_i, t_i, P_i(B|Q_x), \sigma_{t_i})$ 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 the kinematic cuts on the angular distributions are included in the likelihood function through $A(\Omega_i, p_{Ti})$. 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_s^0

candidate. The acceptance is calculated from the $B_s^0 \to J/\psi\phi$ MC events with additional weighting for p_{T} and η 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 $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ decays. As both the acceptance and time- and angle-dependent decay PDFs depend on the transversity angles they must be normalised together. This normalisation is done numerically during the likelihood fit. The PDF is normalised over the entire B_s^0 mass range, 5.150–5.650 GeV.

5.2 Background PDF

The background PDF has the following composition:

$$\begin{aligned}\mathcal{F}_{\mathrm{bkg}}(m_i, t_i, \sigma_{t_i}, \Omega_i, P_i(B|Q_x), p_{\mathrm{T}_i}) &= P_{\mathrm{b}}(m_i) \cdot P_{\mathrm{b}}(t_i|\sigma_{t_i}) \cdot P_{\mathrm{b}}(P_i(B|Q_x)) \\ &\quad \cdot P_{\mathrm{b}}(\Omega_i) \cdot P_{\mathrm{b}}(\sigma_{t_i}) \cdot P_{\mathrm{b}}(p_{\mathrm{T}i}).\end{aligned}$$

The proper decay time function $P_{\mathrm{b}}(t_i|\sigma_{t_i})$ 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 around zero. The two positive exponential functions represent a fraction of longer-lived backgrounds with non-prompt J/ψ, 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, $P_{\mathrm{b}}(\sigma_{t_i})$ and $P_{\mathrm{b}}(p_{\mathrm{T}i})$, are described by gamma functions. They are unchanged from the analysis described in Ref. [27] and explained in detail there. The tagging probability term for background events $P_{\mathrm{b}}(P_i(B|Q_x))$ is described in Section 4.4.

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

$$\begin{aligned}Y_l^m(\theta_T) &= \sqrt{(2l+1)/(4\pi)}\sqrt{(l-m)!/(l+m)!}P_l^{|m|}(\cos\theta_T) \\ P_k(x) &= \frac{1}{2^k k!}\frac{\mathrm{d}^k}{\mathrm{d}x^k}(x^2-1)^k \\ \mathscr{P}_{\mathrm{b}}(\theta_T, \psi_T, \phi_T) &= \sum_{k=0}^{14}\sum_{l=0}^{14}\sum_{m=-l}^{l}\begin{cases} a_{k,l,m}\sqrt{2}Y_l^m(\theta_T)\cos(m\phi_T)P_k(\cos\psi_T) & \text{where } m>0 \\ a_{k,l,m}\sqrt{2}Y_l^{-m}(\theta_T)\sin(m\phi_T)P_k(\cos\psi_T) & \text{where } m<0 \\ a_{k,l,m}\sqrt{2}Y_l^0(\theta_T)P_k(\cos\psi_T) & \text{where } m=0\end{cases}\end{aligned} \tag{4}$$

where the coefficients $a_{k,l,m}$ are adjusted to give the best fit to the angular distributions for events in the sidebands of the B_s^0 mass distribution. These parameters are then fixed in the main fit, defined by Eq. (2). 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, $P_{\mathrm{b}}(m_i)$ is an exponential function with a constant term added.

Contamination from $B_d \to J/\psi K^{0*}$ and $\Lambda_b \to J/\psi p K^-$ events mis-reconstructed as $B_s^0 \to J/\psi\phi$ is accounted for in the fit through the $\mathcal{F}_{B^0}$ and $\mathcal{F}_{\Lambda_b}$ terms in the PDF function described in Eq. (2). 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 cross

Table 4: The ten time-dependent functions, $O^{(k)}(t)$ and the functions of the transversity angles $g^{(k)}(\theta_T, \psi_T, \phi_T)$. The amplitudes $|A_0(0)|^2$ and $|A_\parallel(0)|^2$ are for the CP-even components of the $B_s^0 \to J/\psi\phi$ decay, $|A_\perp(0)|^2$ is the CP-odd amplitude; they have corresponding strong phases δ_0, $\delta_\parallel$ and $\delta_\perp$. By convention, δ_0 is set to be zero. The S–wave amplitude $|A_S(0)|^2$ gives the fraction of $B_s^0 \to J/\psi K^+K^-(f_0)$ and has a related strong phase δ_S. The factor α is described in the text of Section 5.1. The $\pm$ and $\mp$ terms denote two cases: the upper sign describes the decay of a meson that was initially a B_s^0 meson, while the lower sign describes the decays of a meson that was initially $\bar{B}_s^0$.

k	$O^{(k)}(t)$	$g^{(k)}(\theta_T, \psi_T, \phi_T)$
1	$\frac{1}{2}\|A_0(0)\|^2 \left[(1+\cos\phi_s)\, e^{-\Gamma_L^{(s)}t} + (1-\cos\phi_s)\, e^{-\Gamma_H^{(s)}t} \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s\right]$	$2\cos^2\psi_T(1-\sin^2\theta_T\cos^2\phi_T)$
2	$\frac{1}{2}\|A_\parallel(0)\|^2 \left[(1+\cos\phi_s)\, e^{-\Gamma_L^{(s)}t} + (1-\cos\phi_s)\, e^{-\Gamma_H^{(s)}t} \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s\right]$	$\sin^2\psi_T(1-\sin^2\theta_T\sin^2\phi_T)$
3	$\frac{1}{2}\|A_\perp(0)\|^2 \left[(1-\cos\phi_s)\, e^{-\Gamma_L^{(s)}t} + (1+\cos\phi_s)\, e^{-\Gamma_H^{(s)}t} \mp 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s\right]$	$\sin^2\psi_T\sin^2\theta_T$
4	$\frac{1}{2}\|A_0(0)\|\|A_\parallel(0)\|\cos\delta_\parallel \left[(1+\cos\phi_s)\, e^{-\Gamma_L^{(s)}t} + (1-\cos\phi_s)\, e^{-\Gamma_H^{(s)}t} \pm 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s\right]$	$\frac{1}{\sqrt{2}}\sin 2\psi_T\sin^2\theta_T\sin 2\phi_T$
5	$\|A_\parallel(0)\|\|A_\perp(0)\| \left[\frac{1}{2}(e^{-\Gamma_L^{(s)}t} - e^{-\Gamma_H^{(s)}t})\cos(\delta_\perp - \delta_{\|\|})\sin\phi_s \pm e^{-\Gamma_s t}(\sin(\delta_\perp-\delta_\parallel)\cos(\Delta m_s t) - \cos(\delta_\perp - \delta_\parallel)\cos\phi_s\sin(\Delta m_s t))\right]$	$-\sin^2\psi_T\sin 2\theta_T\sin\phi_T$
6	$\|A_0(0)\|\|A_\perp(0)\| \left[\frac{1}{2}(e^{-\Gamma_L^{(s)}t} - e^{-\Gamma_H^{(s)}t})\cos\delta_\perp\sin\phi_s \pm e^{-\Gamma_s t}(\sin\delta_\perp\cos(\Delta m_s t) - \cos\delta_\perp\cos\phi_s\sin(\Delta m_s t))\right]$	$\frac{1}{\sqrt{2}}\sin 2\psi_T\sin 2\theta_T\cos\phi_T$
7	$\frac{1}{2}\|A_S(0)\|^2 \left[(1-\cos\phi_s)\, e^{-\Gamma_L^{(s)}t} + (1+\cos\phi_s)\, e^{-\Gamma_H^{(s)}t} \mp 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s\right]$	$\frac{2}{3}\left(1-\sin^2\theta_T\cos^2\phi_T\right)$
8	$\alpha\|A_S(0)\|\|A_\parallel(0)\| \left[\frac{1}{2}(e^{-\Gamma_L^{(s)}t} - e^{-\Gamma_H^{(s)}t})\sin(\delta_\parallel - \delta_S)\sin\phi_s \pm e^{-\Gamma_s t}(\cos(\delta_\parallel - \delta_S)\cos(\Delta m_s t) - \sin(\delta_\parallel - \delta_S)\cos\phi_s\sin(\Delta m_s t))\right]$	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin^2\theta_T\sin 2\phi_T$
9	$\frac{1}{2}\alpha\|A_S(0)\|\|A_\perp(0)\|\sin(\delta_\perp - \delta_S) \left[(1-\cos\phi_s)\, e^{-\Gamma_L^{(s)}t} + (1+\cos\phi_s)\, e^{-\Gamma_H^{(s)}t} \mp 2e^{-\Gamma_s t}\sin(\Delta m_s t)\sin\phi_s\right]$	$\frac{1}{3}\sqrt{6}\sin\psi_T\sin 2\theta_T\cos\phi_T$
10	$\alpha\|A_0(0)\|\|A_S(0)\| \left[\frac{1}{2}(e^{-\Gamma_H^{(s)}t} - e^{-\Gamma_L^{(s)}t})\sin\delta_S\sin\phi_s \pm e^{-\Gamma_s t}(\cos\delta_S\cos(\Delta m_s t) + \sin\delta_S\cos\phi_s\sin(\Delta m_s t))\right]$	$\frac{4}{3}\sqrt{3}\cos\psi_T\left(1-\sin^2\theta_T\cos^2\phi_T\right)$

sections and branching fractions from Refs. [21, 29–33]. MC simulated events are also used to determine the shape of the mass and transversity angle distributions. The 3D angular distributions of $B_d^0 \to J/\psi K^{0*}$ and of the conjugate decay are modelled using input from Ref. [34], while angular distributions for $\Lambda_b \to J/\psi p K^-$ and the conjugate decay are considered flat. These distributions are sculpted for detector acceptance effects and then described by Legendre polynomial functions, Eq. (4). These shapes are used as templates in the fit. The B_d and Λ_b lifetimes are accounted for in the fit by adding additional exponential terms, scaled by the ratio of B_d/B_s^0 or Λ_b/B_s^0 masses as appropriate, where the lifetimes and masses are taken from Ref. [21]. Systematic uncertainties due to the background from $B_d \to J/\psi K^{0*}$ and $\Lambda_b \to J/\psi p K^-$ 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. [34].

5.3 Muon trigger proper decay time-dependent efficiency

Trigger muons with high values of transverse impact parameter are especially effected by the limited tracking acceptance; this results in inefficiency at large values of the proper decay time. This inefficiency is estimated 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:

$$w = p_0 \cdot [1 - p_1 \cdot (\mathrm{Erf}((t - p_3)/p_2) + 1)], \tag{5}$$

where Erf denotes the error function and p_0, p_1, p_2 and p_3 are parameters determined in the fit to MC events. 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 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 $B^\pm \to J/\psi K^\pm$ 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 parametrise the probability distribution, $P(B|Q_x)$, as a function of the cone charge from the function used by default (a third-order polynomial for muons and a sinusoidal for electrons) to one of several alternative functions. The alternative functions are: a linear function; a fifth-order polynomial; or two third-order polynomials that describe the positive and negative regions and have common constant and linear terms, but independent quadratic and cubic terms. The B_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 $B^\pm$ and B_s^0 signals. The variations between the curves from these two samples are propagated to the calibration curves derived from data. The differences in the parameters between the nominal fit and that with the variations of calibration curves are included in the systematic uncertainty. An additional systematic uncertainty is assigned due to potential 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 $P_b(P(B|Q_x))$ and $P_s(P(B|Q_x))$, variations of the parametrisation are considered (including using histograms in place of a parametrisation). The resulting changes in the parameters of the B_s^0 fit are similarly included in the systematic uncertainties.

- **Angular acceptance method:** The angular acceptance of the detector and the kinematic cuts, $A(\Omega_i, p_{\mathrm{T}_i})$, described in Section 5.1, is calculated from a binned fit to MC simulated data. In order to estimate 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. (5). An alternative fit is performed using different sets of binning in the MC sample used to determine the efficiency. The systematic effects are found to be negligible.

- **Best candidate selection:** The systematic uncertainty of the B_s^0 fit from the selection of the candidate with the best quality in the $\approx 5\%$ of events that are found to contain multiple candidates after cuts is estimated. In the default fit, the B_s^0 candidate with the lowest χ^2/n.d.o.f. is selected. An equivalent sample is created where the candidate with the highest p_T is selected instead. Deviations from the default fit are included in the systematics of the measurement.

- **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. (4). 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_s^0 meson candidate. For this reason, the parametrisation 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. (4) are adjusted to give the best fit to the angular distributions for events in the B_s^0 mass sidebands. To test the sensitivity of the fit results to the choice of sideband regions, the fit is repeated with alternative choices for the excluded signal mass regions: $|(m(B_s^0) - 5.366\ \mathrm{GeV}| > 0.085$ GeV and $|(m(B_s^0) - 5.366\ \mathrm{GeV}| > 0.160$ GeV (instead of the default $|(m(B_s^0) - 5.366\ \mathrm{GeV}| > 0.110$ 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 $B_s^0 \rightarrow J/\psi\phi$ is accounted for in the final fit. Studies are performed to evaluate the effect of the uncertainties in the $B_d \rightarrow J/\psi K^{0*}$ 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. [34]. 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. [34]. 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.

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

- **Fit model mass and lifetime:** To estimate the systematic uncertainties due to the signal B_s^0 mass model, the default model has been altered by adding a second Gaussian function in Eq. (3), which has the same structure as the first Gaussian but a different scale factor, S_m^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, two systematic tests are 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 is 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 is repeated by varying this fraction within one standard deviation of its uncertainty and differences are included in the systematic uncertainty.

- **Fit model S–wave phase**: As explained in Section 5.1, the model for the interference between the $B_s^0 \rightarrow J/\psi\phi(K^+K^-)$ 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 in α, the fit was repeated with $\alpha = 0.51 + 0.02$ 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.

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

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.

	ϕ_s [10^{-3} rad]	$\Delta\Gamma_s$ [10^{-3} ps^{-1}]	Γ_s [10^{-3} ps^{-1}]	$\lvert A_\parallel(0)\rvert^2$ [10^{-3}]	$\lvert A_0(0)\rvert^2$ [10^{-3}]	$\lvert A_S(0)\rvert^2$ [10^{-3}]	$\delta_\perp$ [10^{-3} rad]	$\delta_\parallel$ [10^{-3} rad]	$\delta_\perp - \delta_S$ [10^{-3} rad]
Tagging	19	0.4	0.3	0.2	0.2	1.1	17	19	2.3
Acceptance	0.7	< 0.1	< 0.1	0.8	0.7	2.4	33	11	2.6
ID alignment	0.8	0.2	0.5	< 0.1	< 0.1	< 0.1	11	7.2	< 0.1
Best cand. sel.	0.5	0.4	0.7	0.5	0.2	0.2	12	17	7.5
Background angles model:									
Choice of fit function	2.5	< 0.1	0.3	1.1	< 0.1	0.6	12	0.9	1.1
Choice of p_T bins	1.3	0.5	< 0.1	0.4	0.5	1.2	1.5	7.2	1.0
Choice of mass interval	0.4	0.1	0.1	0.3	0.3	1.3	4.4	7.4	2.3
Dedicated backgrounds:									
B_d^0	2.3	1.1	< 0.1	0.2	3.0	1.5	10	23	2.1
Λ_b	1.6	0.3	0.2	0.5	1.2	1.8	14	30	0.8
Fit model:									
Time res. sig frac	1.4	1.1	0.5	0.5	0.6	0.8	12	30	0.4
Time res. p_T bins	0.7	0.5	0.8	0.1	0.1	0.1	2.2	14	0.7
S–wave phase	0.2	< 0.1	< 0.1	0.3	< 0.1	0.3	11	21	8.4
Model limitation	4.1	1.7	0.9	1.4	< 0.1	1.5	19	0.9	7.0
Total	20	2.5	1.7	2.2	3.4	4.4	52	62	14

7 Results

The full simultaneous unbinned maximum-likelihood fit contains nine physical parameters: $\Delta\Gamma_s$, ϕ_s, Γ_s, $|A_0(0)|^2$, $|A_\parallel(0)|^2$, $\delta_{||}$, $\delta_\perp$, $|A_S(0)|^2$ and δ_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 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^0 meson candidates of $457\,720 \pm 750$. 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.

8 Combination with 7 TeV and 8 TeV results

The measured values are consistent with those obtained in a previous analysis [**Aad:2016tdj**], 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 [**Nisius:2014wua**] 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 effect changed significantly between the measurements. The combined results for the fit parameters and their uncertainties are given in Table 8.

Not reviewed, for internal circulation only

Table 6: Fitted values for the physical parameters of interest with their statistical and systematic uncertainties.

Parameter	Value	Statistical uncertainty	Systematic uncertainty
ϕ_s[rad]	-0.093	0.041	0.020
$\Delta\Gamma_s$[ps^{-1}]	0.0668	0.0046	0.0025
Γ_s[ps^{-1}]	0.6662	0.0014	0.0017
$\lvert A_{\parallel}(0)\rvert^2$	0.2195	0.0021	0.0022
$\lvert A_0(0)\rvert^2$	0.5156	0.0013	0.0034
$\lvert A_S(0)\rvert^2$	0.0378	0.0035	0.0044
$\delta_\perp$ [rad]	3.07	0.11	0.05
$\delta_\parallel$ [rad]	3.31	0.06	0.06
$\delta_\perp - \delta_S$ [rad]	-0.23	0.04	0.01

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

	$\Delta\Gamma$	Γ_s	$\lvert A_{\vert\vert}(0)\rvert^2$	$\lvert A_0(0)\rvert^2$	$\lvert A_S(0)\rvert^2$	$\delta_\parallel$	$\delta_\perp$	$\delta_\perp - \delta_S$
ϕ_s	−0.10	0.02	0.00	−0.01	−0.01	0.02	0.01	−0.01
$\Delta\Gamma$	1	−0.57	0.10	0.09	0.04	0.04	0.00	0.01
Γ_s		1	−0.14	−0.04	0.10	−0.11	−0.03	0.02
$\lvert A_{\vert\vert}(0)\rvert^2$			1	−0.35	−0.21	0.56	0.17	−0.05
$\lvert A_0(0)\rvert^2$				1	0.28	−0.12	−0.05	0.06
$\lvert A_S(0)\rvert^2$					1	−0.40	−0.17	0.20
$\delta_\parallel$						1	0.32	−0.07
$\delta_\perp$							1	0.03

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 ϕ_s and $\Delta\Gamma_s$values determined in combination is −0.04.

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 [9] using the $B_s^0 \to J/\psi\phi$ decay, and a combination of three LHCb measurements [**Aaij:2016psitwoS**, **Aaij:2014Ds**, 8] using $B_s^0 \to J/\psi\phi$, and B_s^0 decays to $\psi(2S)\phi$ and to $D_s^+ D_s^-$, 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 one other and with the SM [2, 4].

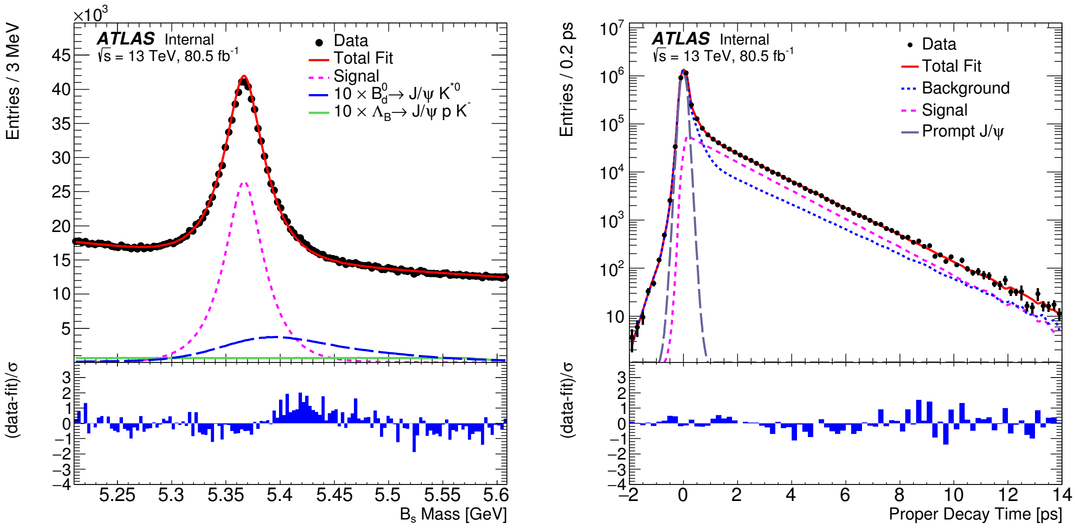

Figure 6: (Left) Mass fit projection for the $B_s^0 \to J/\psi\phi$ sample. The red line shows the total fit, the short-dashed magenta line shows the $B_s^0 \to J/\psi\phi$ signal component, the long-dashed blue line shows the $B_d^0 \to J/\psi K^{0*}$ component, and the solid green line shows the contribution from $\Lambda_b \to J/\psi p K^-$ events. (Right) Proper decay time fit projection for the $B_s^0 \to J/\psi\phi$ sample. The red line shows the total fit while the short-dashed magenta line shows the total signal. The total background is shown as a blue dotted line, and a long-dashed grey line shows the prompt J/ψ background component. Below each figure is a ratio plot that shows the difference between each data point and the total fit line divided by the statistical and systematic uncertainties summed in quadrature (σ) 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.

Parameter	Value	Statistical uncertainty	Systematic uncertainty
ϕ_s[rad]	−0.096	0.036	0.024
$\Delta\Gamma_s$[ps^{-1}]	0.0696	0.0042	0.0029
Γ_s[ps^{-1}]	0.6684	0.0014	0.0018
$\lvert A_\parallel(0)\rvert^2$	0.2210	0.0019	0.0026
$\lvert A_0(0)\rvert^2$	0.5178	0.0014	0.0040
$\lvert A_S\rvert^2$	0.0407	0.0032	0.0057
$\delta_\perp$ [rad]	3.19	0.11	0.07
$\delta_\parallel$ [rad]	3.32	0.06	0.09
$\delta_\perp - \delta_S$ [rad]	−0.23	0.04	0.02

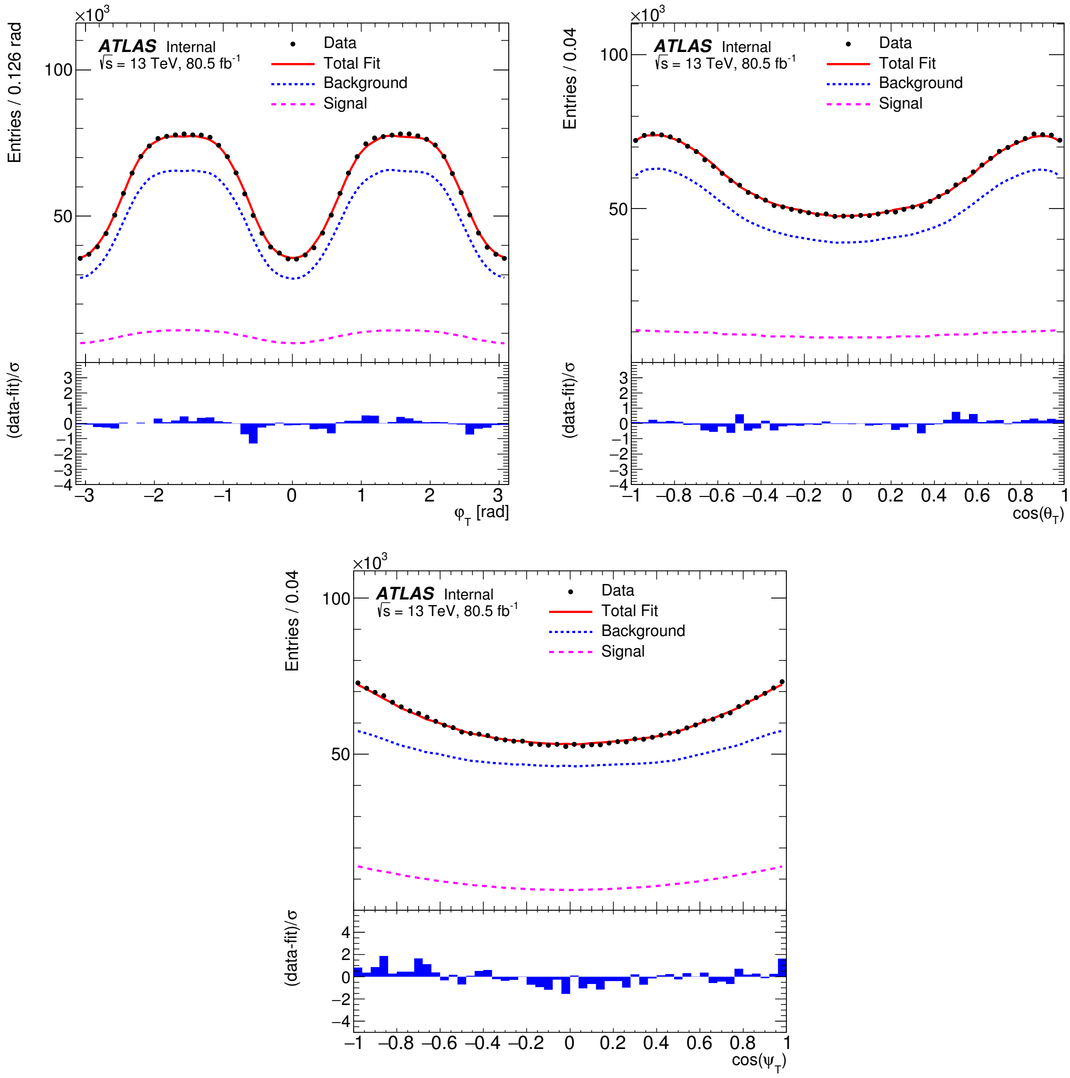

Figure 7: Fit projections for the transversity angles ϕ_T (top left), $\cos(\theta_T)$ (top right), and $\cos(\psi_T)$ (bottom). In all three plots the red solid line shows the total fit, the $B_s^0 \to J/\psi\phi$ signal component is shown by the magenta dashed line and the blue dotted line shows the 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 (σ) of that point.

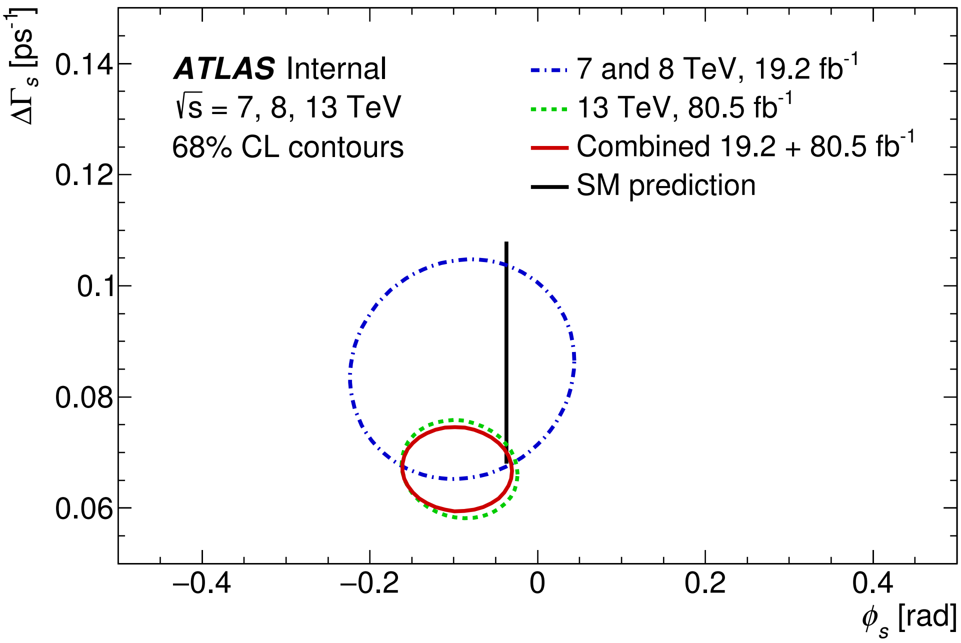

Figure 8: Likelihood 68% confidence level contours in the $\phi_s - \Delta\Gamma_s$ plane, showing ATLAS results for 7 TeV and 8 TeV data (blue dashed-dotted curve), for 13 TeV data (green dashed curve) and for 13 TeV data combined with 7 TeV and 8 TeV (red solid curve) data. In all contours the statistical and systematic uncertainties are combined in quadrature and correlations are taken into account.

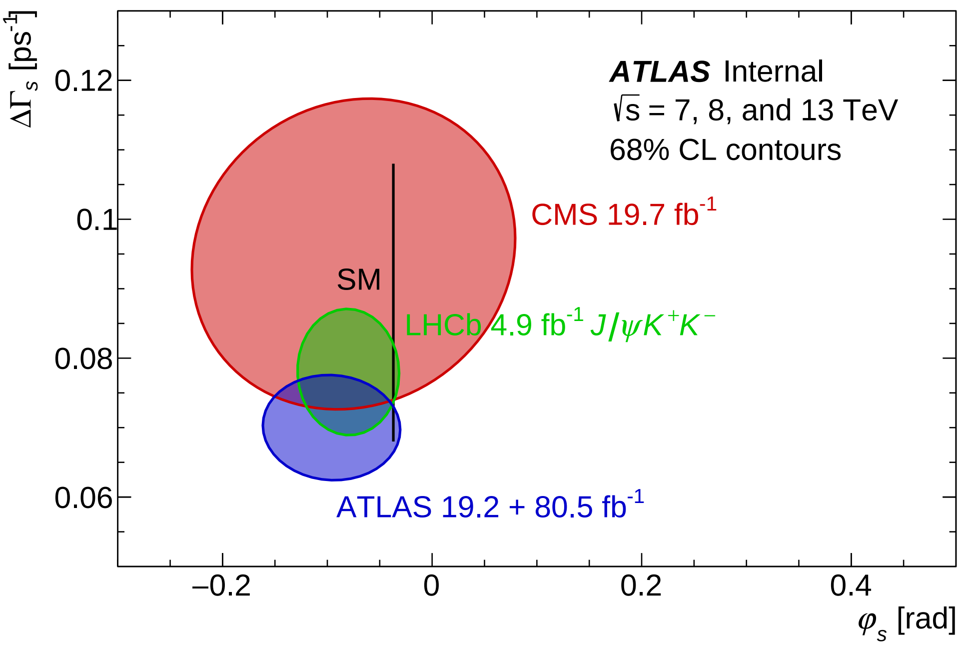

Figure 9: Likelihood 68% confidence level contours in the $\phi_s - \Delta\Gamma_s$plane, including results from LHCb (green) and CMS (red) using 7 TeV and 8 TeV data. The brown contour shows the ATLAS result for 13 TeV combined with 7 TeV and 8 TeV. In all contours the statistical and systematic uncertainties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

9 Summary

A measurement of the time-dependent CP asymmetry parameters in $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ 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 [**Aad:2016tdj**]. The two measurements are statistically combined, leading to the following results:

$$
\begin{aligned}
\phi_s &= -0.096 \pm 0.036\,(\text{stat.}) \pm 0.024\,(\text{syst.})\ \text{rad} \\
\Delta\Gamma_s &= 0.0696 \pm 0.0042\,(\text{stat.}) \pm 0.0029\,(\text{syst.})\ \text{ps}^{-1} \\
\Gamma_s &= 0.6684 \pm 0.0014\,(\text{stat.}) \pm 0.0018\,(\text{syst.})\ \text{ps}^{-1} \\
|A_{\parallel}(0)|^2 &= 0.2210 \pm 0.0019\,(\text{stat.}) \pm 0.0026\,(\text{syst.}) \\
|A_0(0)|^2 &= 0.5178 \pm 0.0014\,(\text{stat.}) \pm 0.0040\,(\text{syst.}) \\
|A_S(0)|^2 &= 0.0407 \pm 0.0032\,(\text{stat.}) \pm 0.0057\,(\text{syst.}) \\
\delta_{\perp} &= 3.19 \pm 0.11\,(\text{stat.}) \pm 0.07\,(\text{syst.})\ \text{rad} \\
\delta_{\parallel} &= 3.32 \pm 0.06\,(\text{stat.}) \pm 0.09\,(\text{syst.})\ \text{rad} \\
\delta_{\perp} - \delta_S &= -0.23 \pm 0.04\,(\text{stat.}) \pm 0.02\,(\text{syst.})\ \text{rad}
\end{aligned}
$$

The new ATLAS measurement on the CP violation phase ϕ_s increased a precision of the previous ATLAS measurement using 7 TeV and 8 TeV ATLAS data [**Aad:2016tdj**], and the new result is still consistent with the Standard Model prediction.

Acknowledgements

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; 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-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; 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, CANARIE, CRC and Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Skłodowska-Curie Actions, European Union; Investissements d' Avenir Labex and Idex, ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; CERCA Programme Generalitat de Catalunya, 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 (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [35].

References

[1] F. J. Botella, G. C. Branco and M. Nebot, *CP violation and limits on New Physics including recent B_s measurements*, Nuclear Physics B **768** (2007) 1.

[2] J. Charles et al., *Predictions of selected flavour observables within the standard model*, Phys. Rev. D **84** (2011) 033005, arXiv: `1106.4041 [hep-ph]`.

[3] LHCb Collaboration, *Precision measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency with the decay $B_s^0 \to D_s^- \pi^+$*, New Journal of Physics **15** (2013) 053021, arXiv: `1304.4741 [hep-ex]`.

[4] A. Lenz and U. Nierste, *Numerical updates of lifetimes and mixing parameters of B mesons*, (2011), arXiv: `1102.4274 [hep-ph]`.

[5] A. Lenz and U. Nierste, *Theoretical update of $B_s - \bar{B}_s$ mixing*, JHEP **06** (2007) 072, arXiv: `hep-ph/0612167`.

[6] D0 Collaboration, *Measurement of the CP-violating phase $\phi_s^{J/\psi\phi}$ using the flavor-tagged decay $B_s^0 \to J/\psi\phi$ in 8 fb^{-1} of $p\bar{p}$ collisions*, Phys. Rev. D **85** (2012) 032006, arXiv: 1109.3166 [hep-ex].

[7] CDF Collaboration, *Measurement of the Bottom-Strange Meson Mixing Phase in the Full CDF Data Set*, Phys. Rev. Lett. **109** (2012) 171802, arXiv: 1208.2967 [hep-ex].

[8] LHCb Collaboration, *Precision Measurement of CP violation in $B_s^0 \to J/\psi K^+K^-$ Decays*, Phys. Rev. Lett. **114** (2015) 041801, arXiv: 1411.3104 [hep-ex].

[9] CMS Collaboration, *Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the $B_s^0 \to J/\psi\phi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **757** (2016) 97, arXiv: 1507.07527 [hep-ex].

[10] LHCb Collaboration, *Updated measurement of time-dependent CP-violating observables in $B_s^0 \to J/\psi K^+K^-$ decays*, The European Physical Journal C **79** (2019) 706, URL: https://doi.org/10.1140/epjc/s10052-019-7159-8.

[11] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[12] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, JINST **13** (2018) T05008, arXiv: 1803.00844 [physics.ins-det].

[13] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19, 2010, URL: https://cds.cern.ch/record/1291633, *ATLAS Insertable B-Layer Technical Design Report Addendum*, ATLAS-TDR-19-ADD-1, 2012, URL: https://cds.cern.ch/record/1451888.

[14] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653, arXiv: 1608.03953 [hep-ex].

[15] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017.

[16] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph].

[17] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

[18] J. Pumplin et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **07** (2002) 012, arXiv: hep-ph/0201195.

[19] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].

[20] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[21] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (3 2018) 030001.

[22] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex].

[23] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV*, Submitted to: Eur. Phys. J. C (2019), arXiv: 1902.04655 [physics.ins-det].

[24] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[25] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph].

[26] ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, JHEP **08** (2018) 089, arXiv: 1805.01845 [hep-ex].

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

[28] A. S. Dighe, I. Dunietz and R. Fleischer, *Extracting CKM phases and $B_s - \bar{B}_s$ mixing parameters from angular distributions of non-leptonic B decays*, Eur. Phys. J. C **6** (1999) 647, arXiv: hep-ph/9804253.

[29] LHCb Collaboration, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions*, LHCb-CONF-2013-011, 2013, URL: https://cds.cern.ch/record/1559262.

[30] B. Aubert et al., *Search for the $Z(4430)^-$ at BABAR*, Phys. Rev. D **79** (2009) 112001, arXiv: 0811.0564 [hep-ex].

[31] LHCb Collaboration, *Study of the kinematic dependences of Λ_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].

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

[33] LHCb Collaboration, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv: 1507.03414 [hep-ex].

[34] LHCb Collaboration, *Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays*, Phys. Rev. D **88** (2013) 052002, arXiv: 1307.2782 [hep-ex].

[35] ATLAS Collaboration, *ATLAS Computing Acknowledgements*, ATL-GEN-PUB-2016-002, URL: https://cds.cern.ch/record/2202407.