Measurement of the time-dependent CP asymmetry in $B^0 \to J/\psi K_S^0$ decays

LHCb Collaboration

1. Introduction

The source of CP violation in the electroweak sector of the Standard Model (SM) is the single irreducible complex phase of the Cabibbo–Kobayashi–Maskawa (CKM) quark mixing matrix [1,2]. The decay $B^0 \to J/\psi K_S^0$ is one of the theoretically cleanest modes for the study of CP violation in the $B^0$ meson system. Here, the $B^0$ and $B^0$ mesons decay to a common CP-odd eigenstate allowing for interference through $B^0\to B^0$ mixing.

In the $B^0$ system the decay width difference $\Delta \Gamma$ between the heavy and light mass eigenstates is negligible. Therefore, the time-dependent rate asymmetry can be written as [3,4]

$$A_{J/\psi K_S^0}(t) \equiv \frac{\Gamma(B^0(t) \to J/\psi K_S^0) - \Gamma(B^0(t) \to J/\psi K_S^0)}{\Gamma(B^0(t) \to J/\psi K_S^0) + \Gamma(B^0(t) \to J/\psi K_S^0)} = S_{J/\psi K_S^0} \sin(\Delta m t) - C_{J/\psi K_S^0} \cos(\Delta m t).$$

(1)

Here $B^0(t)$ and $B^0(t)$ are the states into which particles produced at $t = 0$ as $B^0$ and $B^0$ respectively have evolved, when decaying at time $t$. The parameter $\Delta m_d$ is the mass difference between the two $B^0$ mass eigenstates. The sine term results from the interference between direct decay and decay after $B^0\to B^0$ mixing. The cosine term arises either from the interference between decay amplitude with different weak and strong phases (direct CP violation) or from CP violation in $B^0\to B^0$ mixing.

In the SM, CP violation in mixing and direct CP violation are both negligible in $B^0 \to J/\psi K_S^0$ decays, hence $C_{J/\psi K_S^0} \approx 0$, while $S_{J/\psi K_S^0} \approx \sin 2\beta$, where the CKM angle $\beta$ can be expressed in terms of the CKM matrix elements as $\arg[-V_{ub}V_{cb}^\ast/V_{td}V_{ts}^\ast]$. It can also be measured in other $B^0$ decays to final states including charmonium such as $J/\psi K_L^0$, $J/\psi K^{*0}$, $\psi(2S)K^{*0}$, which have been used in measurements by the BaBar and Belle Collaborations [5,6]. Currently, the world averages are $S_{J/\psi K_S^0} = 0.679 \pm 0.020$ and $C_{J/\psi K_S^0} = 0.005 \pm 0.017$ [7].

The time-dependent measurement of the CP parameters $S_{J/\psi K_S^0}$ and $C_{J/\psi K_S^0}$ requires flavour tagging, i.e. the knowledge whether the decaying particle was produced as a $B^0$ or a $B^0$ meson. If a fraction $\omega$ of candidates is tagged incorrectly, the accessible time-dependent asymmetry $A_{J/\psi K_S^0}(t)$ is diluted by a factor $(1-2\omega)$. Hence, a measurement of the CP parameters requires precise knowledge of the wrong tag fraction. Additionally, the asymmetry between the production rates of $B^0$ and $B^0$ has to be determined as it affects the observed asymmetries.

In this Letter, the most precise measurement of $S_{J/\psi K_S^0}$ and $C_{J/\psi K_S^0}$ to date at a hadron collider is presented using approximately 8200 flavour-tagged $B^0 \to J/\psi K_S^0$ decays.

2. Data samples and selection requirements

The data sample consists of 1 fb$^{-1}$ of pp collisions at $\sqrt{s} = 7$ TeV collected by the LHCb experiment at CERN. The detector [8] is a single-arm forward spectrometer covering the pseudorapidity range 2 to 5, designed for the study of particles containing $b$ or $c$ quarks. It includes a high precision tracking system consisting of a silicon-strip vertex detector surrounding the interaction region, a large-area silicon-strip detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three stations of silicon-strip detectors and straw drift-tubes placed downstream. The combined tracking system has a momentum resolution $\Delta p/p$ that varies from 0.4% at 5 GeV/c to 0.6% at 100 GeV/c, and an impact parameter resolution of 20 µm for tracks with high transverse momentum. Charged hadrons are identified using two ring-imaging Cherenkov detectors. Photon, electron and hadron candidates are identified by a...
The analysis is performed on events reconstructed $B^0 \rightarrow J/\psi K^0_S$ candidates with subsequent $J/\psi \rightarrow \mu^+\mu^-$ and $K^0_S \rightarrow \pi^+\pi^-$ decays. Events are selected by the trigger consisting of hardware and software stages. The hardware stage accepts events if muon or hadron candidates with high transverse momentum ($p_T$) with respect to the beam axis are detected. In the software stage, events are required to contain two oppositely-charged particles, both compatible with a muon hypothesis, that form an invariant mass greater than 2.7 GeV/c$^2$. The resulting $J/\psi$ candidate has to be clearly separated (decay length significance greater than 3) from the production vertex (PV) with which it is associated on the basis of the impact parameter. The overall signal efficiency of these triggers is found to be 64%.

Further selection criteria are applied offline to decrease the number of background candidates. The $J/\psi$ candidates are reconstructed from two oppositely-charged, well identified muons with $p_T > 500$ MeV/c that form a common vertex with a fit $\chi^2$/ndf of less than 11, where ndf is the number of degrees of freedom, and with an invariant mass in the range 3035–3160 MeV/c$^2$. It is required that the $J/\psi$ candidate fulfills the trigger requirements described above. The $K^0_S$ candidates are formed from two oppositely-charged pions, both with (long $K^0_S$ candidate) or without (downstream $K^0_S$ candidate) hits in the vertex detector. Any $K^0_S$ candidates where both pion tracks have hits in the tracking stations but only one has additional hits in the vertex detector are ignored, as they would only contribute to $<2\%$ of the events. Each pion must have $p > 2$ GeV/c and a clear separation from any PV. Furthermore, they must form a common vertex with a fit $\chi^2$/ndf of less than 20 and an invariant mass within the range 485.6–509.6 MeV/c$^2$ (long $K^0_S$ candidates) or 476.6–518.6 MeV/c$^2$ (downstream $K^0_S$ candidates). Different mass windows are chosen to account for different mass resolutions for long and downstream $K^0_S$ candidates. The $K^0_S$ candidate’s decay vertex is required to be significantly displaced with respect to the associated PV.

The $B^0$ candidates are constructed from combinations of $J/\psi$ and $K^0_S$ candidates that form a vertex with a reconstructed mass $m_{J/\psi K^0_S}$ in the range 5230–5330 MeV/c$^2$. The value of $m_{J/\psi K^0_S}$ is computed constraining the invariant masses of the $\mu^+\mu^-$ and $\pi^+\pi^-$ to the known $J/\psi$ and $K^0_S$ masses [9], respectively. As most events involve more than one reconstructed PV, $B^0$ candidates are required to be associated to one PV only and are therefore omitted if their impact parameter significance with respect to other PVs in the event is too small. Additionally, the $K^0_S$ candidate’s decay vertex is required to be separated from the $B^0$ decay vertex by a decay time significance of the $K^0_S$ greater than 5.

The decay time $t$ of the $B^0$ candidates is determined from a vertex fit to the whole decay chain under the constraint that the $B^0$ candidate originates from the associated PV [10]. Only candidates with a good quality vertex fit and with $0.3 < t < 183.5$ ps are retained. In case more than one candidate is selected in an event, that with the best vertex fit quality is chosen. The fit uncertainty on $t$ is used as an estimate of the decay time resolution $\sigma_t$, which is required to be less than 0.2 ps. Finally, candidates are only retained if the flavour tagging algorithms provide a prediction for the production flavour of the candidate, as discussed in Section 3.

Simulated samples are used for cross-checks and studies of decay time distributions. For the simulation, $p\bar{p}$ collisions are generated using PYTHIA 6.4 [11] with a specific LHCb configuration [12]. Decays of hadronic particles are described by EVTGEN [13] in which final state radiation is generated using PHOTOS [14]. The interaction of the generated particles with the detector is implemented using the GEANT4 toolkit [15] as described in Ref. [16].

3. Flavour tagging

A mandatory step for the study of $CP$ violating quantities is to tag the initial, i.e. production, flavour of the decaying $B^0$ meson. Since $b$ quarks are predominantly produced in $b\bar{b}$ pairs in LHCb, the flavour tagging algorithms used in this analysis [17] reconstruct the flavour of the non-signal $b$ hadron. The flavour of the non-signal $b$ hadron is determined by identifying the charge of its decay products, such as that of an electron or a muon from a semileptonic $b$ decay, a kaon from a $b \rightarrow c \rightarrow s$ decay chain, or the charge of its inclusively reconstructed decay vertex. The algorithms use this information to provide a tag $d$ that takes the value $+1 (-1)$ in the case where the signal candidate is tagged as an initial $B^0$ ($\bar{B}^0$) meson.

A careful study of the fraction of candidates that are wrongly tagged (mistag fraction) is necessary as the measured asymmetry is diluted due to the imperfect tagging performance. The mistag fraction ($\omega$) is extracted on an event-by-event basis from the combined per-event mistag probability prediction $\rho$ of the tagging algorithms. On average, the mistag fraction is found to depend linearly on $\eta$ and is parameterised as

$$\omega(\eta) = p_1 \cdot (\eta - \langle \eta \rangle) + p_0.$$

Using events from the self-tagging control channel $B^+ \rightarrow J/\psi K^+$, the parameters are determined to be $p_1 = 1.035 \pm 0.021$ (stat) $\pm 0.012$ (syst), $p_0 = 0.392 \pm 0.002$ (stat) $\pm 0.009$ (syst) and $\langle \eta \rangle = 0.391$ [18]. The systematic uncertainties on the tagging calibration parameters are estimated by comparing the tagging performance obtained in different decay channels such as $B^0 \rightarrow J/\psi K^0$, in $B^+$ and $B^-$ subsamples separately, and in different data taking periods.

The difference in tagging response between $B^0$ and $\bar{B}^0$ is parameterised by using

$$\omega = \omega(\eta) = \frac{\Delta p_0}{2},$$

where the $+ (-)$ is used for a $B^0$ ($\bar{B}^0$) meson at production and $\Delta p_0$ is the mistag fraction asymmetry parameter, which is the difference of $p_0$ for $B^0$ and $\bar{B}^0$ mesons. It is measured as $\Delta p_0 = 0.011 \pm 0.003$ using events from the control channel $B^+ \rightarrow J/\psi K^+$. By using $\Delta p_0$ in the analysis, the systematic uncertainty on the $p_0$ parameter is reduced to 0.008. The difference of tagging efficiency for $b^0$ and $\bar{b}$ mesons is measured in the same control channel as $\Delta p_0 = 0.000 \pm 0.001$ and is therefore negligible. Thus, it is only used to estimate possible systematic uncertainties in the analysis.

The effect of imperfect tagging is the reduction of the statistical power by a factor $\varepsilon_{\text{tag}} D^2$, where $\varepsilon_{\text{tag}}$ is the tagging efficiency and $D = 1 - 2\omega$ is the dilution factor. The effective $\varepsilon_{\text{tag}}$ and $D$ values are measured as $\varepsilon_{\text{tag}} = (32.65 \pm 0.31\%)$ and $D = 0.270 \pm 0.015$, resulting in $\varepsilon_{\text{tag}} D^2 = (2.38 \pm 0.27\%)$, where combined systematic and statistical uncertainties are quoted. The measured dilution corresponds to a mistag fraction of $\omega = 0.365 \pm 0.008$.

4. Decay time acceptance and resolution

The bias on the decay time distribution due to the trigger is estimated by comparing candidates selected using different trigger requirements. In the selection, the reconstructed decay times of the $B^0 \rightarrow J/\psi K^0_S$ candidates are required to be greater than 0.3 ps. This requirement makes the acceptance effects of the trigger nearly negligible. However, some small efficiency loss remains
for small decay times. Neglecting this efficiency loss is treated as a source of systematic uncertainty.

A decrease of efficiency is also observed at large decay times, mostly affecting the candidates in the long \(K_S^0\) subsample. This can be described with a linear efficiency function with parameters determined from simulated data for the downstream and long \(K_S^0\) subsamples separately. The efficiency function is then used to correct the description of the decay time distribution.

The finite decay time resolution of the detector leads to an additional dilution of the experimentally accessible asymmetry. It is modelled event-by-event with a triple Gaussian function,

\[
\mathcal{R}(t - t'|\sigma_t) = \sum_{i=1}^{3} f_i \frac{1}{\sqrt{2\pi} s_i} \exp\left(-\frac{(t - t' - b_{\sigma_t})^2}{2(s_{\sigma_t}^2)}\right),
\]

where \(t\) is the reconstructed decay time, \(t'\) is the true decay time, and \(\sigma_t\) is the per-event decay time resolution estimate. The parameters are: the three fractions \(f_i\), which sum to unity, the three scale factors \(s_i\), and a relative bias \(b\), which is found to be small. They are determined from a fit to the \(t\) and \(\sigma_t\) distributions of prompt \(J/\psi\) events that pass the selection and trigger criteria for \(B^0 \rightarrow J/\psi K_S^0\), except for decay time biasing requirements. The parameters are determined separately for the subsamples formed from downstream and long \(K_S^0\) candidates. This results in an average effective decay time resolution of 55.6 fs (65.6 fs) for candidates with long (downstream) \(K_S^0\).

5. Measurement of \(S_{J/\psi K_S^0}\) and \(C_{J/\psi K_S^0}\)

The analysis is performed using the following set of observables: the reconstructed mass \(m_{J/\psi K_S^0}\), the decay time \(t\), the estimated decay time resolution \(\sigma_t\), the flavour tag \(d\), and the per-event mistag probability \(\eta\). The \(CP\) observables \(S_{J/\psi K_S^0}\) and \(C_{J/\psi K_S^0}\) are determined as parameters in an unbinned extended maximum likelihood fit to the data.

Due to different resolution and acceptance effects for the downstream and long \(K_S^0\) subsamples, a simultaneous fit to both subsamples is performed. In each subsample, the probability density function (PDF) is defined as the sum of two individual PDFs, one for each of the components of the fit: the \(B^0\) signal and the background. The latter component contains both combinatorial background and mis-reconstructed \(b\)-hadron decays.

The reconstructed mass distribution of the signal is described by the sum of two Gaussian PDFs with common mean but different widths. Only the mean is shared between the two subsamples. The background component is parameterised as an exponential function, different for each subsample.

The signal and background distributions of the per-event mistag probability \(\eta\) are modelled with PDFs formed from histograms obtained with the sPlot technique [19] on the reconstructed mass distribution. In both subsamples the same signal and background models are used.

The distributions of the estimated decay time resolution \(\sigma_t\) are different in each component and each subsample. Hence, no parameters are shared between subsamples or components. All \(\sigma_t\) PDFs are modelled with lognormal functions

\[
\ln(\sigma_t; M_{\sigma_t}, k) = \frac{1}{\sqrt{2\pi} \sigma_t} \ln k \exp\left(-\frac{\ln^2(\sigma_t/M_{\sigma_t})}{2\ln^2(k)}\right),
\]

where \(M_{\sigma_t}\) is the median and \(k\) the tail parameter. The background components in both subsamples are parameterised by single lognormal functions. For the signal a sum of two lognormals with common (different) median parameter(s) is chosen for the long \(K_S^0\) (downstream) \(K_S^0\) subsample.

The background PDFs of the decay time are modelled in each subsample by the sum of two exponential functions. These are convolved with the corresponding resolution function \(R(t - t'|\sigma_t)\). The parameters are not shared between the two subsamples. The background distribution of tags \(d\) is described as a uniform distribution.

The signal PDF for the decay time simultaneously describes the distribution of tags \(d\), and is given by

\[
\mathcal{P}(t, d|\sigma_t, \eta) = \epsilon(t) \cdot \mathcal{P}_{CP}(t', d|\sigma_t, \eta) \oplus \mathcal{R}(t - t'|\sigma_t),
\]

with

\[
\mathcal{P}_{CP}(t', d|\sigma_t, \eta) \propto e^{-\frac{1}{\tau}(1 - d\Delta p_0 - d A_p (1 - 2\omega(\eta)))} - (d(1 - 2\omega(\eta)) - A_p (1 - d\Delta p_0)) S_{J/\psi K_S^0} \sin \Delta m_d t' \\
+ (d(1 - 2\omega(\eta)) - A_p (1 - d\Delta p_0)) C_{J/\psi K_S^0} \cos \Delta m_d t'.
\]

This PDF description exploits time-dependent asymmetries, while its normalisation adds sensitivity by accessing time-integrated asymmetries. The lifetime \(\tau\), the mass difference \(\Delta m_d\), and the \(CP\) parameters \(S_{J/\psi K_S^0}\) and \(C_{J/\psi K_S^0}\) are shared in the PDFs of the downstream and long \(K_S^0\) subsamples, as well as the asymmetry \(A_p = (R_F - R_G)/(R_F + R_G)\) of the production rates \(R\) for \(B^0\) and \(B^0\) mesons in pp collisions at LHCb. The latter value has been measured in Refs. [20,21] to be \(A_p = -0.015 \pm 0.013\).

In the fit all parameters related to decay time resolution and acceptance are fixed. The tagging parameters and the production asymmetry parameter are constrained within their statistical uncertainties by Gaussian constraints in the likelihood. The fit yields \(S_{J/\psi K_S^0} = 0.73 \pm 0.07\), \(C_{J/\psi K_S^0} = 0.03 \pm 0.09\), with a correlation coefficient \(\rho(S_{J/\psi K_S^0}, C_{J/\psi K_S^0}) = 0.42\). Both of the uncertainties and the correlation are statistical only. The lifetime is fitted as \(\tau = 1.496 \pm 0.018\) ps and the oscillation frequency as \(\Delta m_d = 0.53 \pm 0.05\) ps\(^{-1}\), both in good agreement with the world averages [7,72]. The mass and decay time distributions are shown in Fig. 1. The measured signal asymmetry and the projection of the signal PDF are shown in Fig. 2.

6. Systematic uncertainties

Most systematic uncertainties are estimated by generating a large number of pseudo-experiments from a modified PDF and fitting each sample with the nominal PDF. The PDF used in the generation is chosen according to the source of systematic uncertainty that is being investigated. The variation of the fitted values of the \(CP\) parameters is used to estimate systematic effects on the measurement.

The largest systematic uncertainty arises from the limited knowledge of the accuracy of the tagging calibration. It is estimated by varying the calibration parameters within their systematic uncertainties in the pseudo-experiments. Another minor systematic uncertainty related to tagging emerges from ignoring a possible difference of tagging efficiencies of \(B^0\) and \(\bar{B}^0\).

The effect of an incorrect description of the decay time resolution model is derived from pseudo-experiments in which the scale factors of the resolution model are multiplied by a factor of either 0.5 or 2 in the generation. As the mean decay time resolution of LHCb is much smaller than the oscillation period of the \(B^0\) system this variation leads only to a small systematic uncertainty. The omission of acceptance effects for low decay times is estimated
from pseudo-experiments where the time-dependent efficiencies measured from data are used in the generation but omitted in the fits. Additionally, a possible inaccuracy in the description of the efficiency decrease at large decay times is checked by varying the parameters within their errors, but is found to be negligible.

The uncertainty induced by the limited knowledge of the background distributions is evaluated from a fit method based on the sPlot technique. A fit with the PDFs for the reconstructed mass is performed to extract signal weights for the distributions in the other observable dimensions. These weights are then used to perform a fit with the PDF of the signal component only. The difference in fit results is treated as an estimate of the systematic uncertainty.

To estimate the influence of possible biases in the CR parameters emerging from the fit method itself, the method is probed with a large set of pseudo-experiments. Systematic uncertainties of 0.004 for $S_{J/\psi K^0_S}$ and 0.005 for $C_{J/\psi K^0_S}$ are assigned based on the biases observed in different fit settings.

The uncertainty on the scale of the longitudinal axis and on the scale of the momentum [23] sum to a total uncertainty of < 0.1% on the decay time. This has a negligible effect on the CR parameters. Likewise, potential biases from a non-random choice of the $B^0$ candidate in events with multiple candidates are found to be negligible.

The sources of systematic effects and the resulting systematic uncertainties on the CR parameters are quoted in Table 1 where the total systematic uncertainty is calculated by summing the individual uncertainties in quadrature.

The analysis strategy makes use of the time-integrated and time-dependent decay rates of $B^0 \to J/\psi K^0_S$ decays that are tagged as $B^0/\bar{B}^0$ meson. Cross-check analyses exploiting only the time-integrated or only the time-dependent information show that both give results that are in good agreement and contribute to the full analysis with comparable statistical power.

### 7. Conclusion

In a dataset of 1.0 fb$^{-1}$ collected with the LHCb detector, approximately 8200 flavour tagged decays of $B^0 \to J/\psi K^0_S$ are selected to measure the CP observables $S_{J/\psi K^0_S}$ and $C_{J/\psi K^0_S}$, which are related to the CKM angle $\beta$. A fit to the time-dependent decay rates of $B^0$ and $\bar{B}^0$ decays yields

$$S_{J/\psi K^0_S} = 0.73 \pm 0.07\text{(stat)} \pm 0.04\text{(syst)},$$

$$C_{J/\psi K^0_S} = 0.03 \pm 0.09\text{(stat)} \pm 0.01\text{(syst)},$$

with a statistical correlation coefficient of $\rho(S_{J/\psi K^0_S}, C_{J/\psi K^0_S}) = 0.42$. This is the first significant measurement of CP violation in $B^0 \to J/\psi K^0_S$ decays at a hadron collider [24]. The measured values are in agreement with previous measurements performed at the $B$ factories [5,6] and with the world averages [7].

### Acknowledgements

We express our gratitude to our colleagues in the CERN accelerator departments for the excellent performance of the LHC. We thank the technical and administrative staff at the LHCb institutes. We acknowledge support from CERN and from the national agencies: CAPES, CNPq, FAPERJ and FINEP (Brazil); NSFC (China); CNRS/IN2P3 and Region Auvergne (France); BMBF, DFG, HGF and MPG (Germany); SFN (Indiana); INFN (Italy); FOM and NWO (Netherlands); SNSF (Switzerland); NAS (Ukraine).
This article is published Open Access at sciedirect.com. It is distributed under the terms of the Creative Commons Attribution License 3.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

References


[22] LHCb Collaboration, R. Aaij, et al., Measurement of the $B^0 - \bar{B}^0$ oscillation frequency $\Delta m_b$ with the decays $B^0 \to D^* \pi^-$ and $B^0 \to J/\psi K^0$, arXiv: 1210.6750.


LHCb Collaboration
