

# Search for $CP$ violation in $B^0 \rightarrow J/\psi K_S^0$ decays with first LHCb data

The LHCb Collaboration <sup>1</sup>

## Abstract

We report on the measurement of the well-established  $CP$  violation in  $B^0 \rightarrow J/\psi K_S^0$  decays. We perform a time-dependent analysis of the decays reconstructed in  $35 \text{ pb}^{-1}$  of LHCb data that was taken in 2010. We measure the  $CP$  asymmetry parameter  $S_{J/\psi K_S^0}$ , which is connected to the CKM angle  $\beta$  through  $S_{J/\psi K_S^0} = \sin 2\beta$ , neglecting  $CP$  violation in  $B^0-\bar{B}^0$  mixing and decay. We find  $S_{J/\psi K_S^0} = 0.53^{+0.28}_{-0.29}(\text{stat}) \pm 0.05(\text{syst})$ .

---

<sup>1</sup>Conference report prepared for *Beauty*, Amsterdam, 4–11 April 2011; contact authors: Till Moritz Karbach and Bruno Souza de Paula



# 1 Introduction

The decay  $B^0 \rightarrow J/\psi K_S^0$  is well known as the gold-plated mode for the study of  $CP$  violation in the  $B^0$  meson system. Here, the  $B^0$  meson decays to a  $CP$  eigenstate common to both  $B^0$  and  $\bar{B}^0$ , allowing for interference through oscillation. Therefore, the measurement of the decay  $B^0 \rightarrow J/\psi K_S^0$  has a good sensitivity to the CKM angle  $\sin 2\beta$ . One typically measures the parameter  $S_{J/\psi K_S^0}$ , which is equal to  $\sin 2\beta$  up to a small correction. During the last decade the  $B$ -factories BABAR and Belle reached outstanding precision in the measurement of  $S_{J/\psi K_S^0}$ . The most recent BABAR measurement in  $B^0 \rightarrow J/\psi K_S^0$  ( $K_S^0 \rightarrow \pi^+\pi^-$ ) reports  $S_{J/\psi K_S^0} = 0.663 \pm 0.039(\text{stat}) \pm 0.012(\text{syst})$  [1]. The most recent Belle result in  $B^0 \rightarrow J/\psi K_S^0$  ( $K_S^0 \rightarrow \pi^+\pi^-, \pi^0\pi^0$ ) is  $S_{J/\psi K_S^0} = 0.642 \pm 0.031(\text{stat}) \pm 0.017(\text{syst})$  [2]. Besides  $J/\psi K_S^0$ , many other charmonium  $B^0$  final states such as  $J/\psi K_L^0$ ,  $\psi(2S)K_S^0$ ,  $\chi_{c1}K_S^0$ , or  $\eta_c K_S^0$ , give sensitivity to  $\sin 2\beta$ , too. Currently, the world average [3] of  $\sin 2\beta$  only has 3% uncertainty:  $\sin 2\beta = 0.673 \pm 0.023$ . Using the first LHCb data it is clear that only little additional knowledge on  $\sin 2\beta$  can be gained. But still the measurement of  $\sin 2\beta$  is an important step in order to demonstrate the possibilities of the LHCb experiment. It demonstrates that the essential flavour tagging algorithms are under control.

The CKM angle  $\sin 2\beta$  is connected to the parameters  $S_{J/\psi K_S^0}$  and  $C_{J/\psi K_S^0}$ , which govern the time-dependent decay rate asymmetry of the  $B^0 \rightarrow J/\psi K_S^0$  decays [4, 5],

$$\begin{aligned} \mathcal{A}_{J/\psi K_S^0}(t) &\equiv \frac{\Gamma(\bar{B}^0(t) \rightarrow J/\psi K_S^0) - \Gamma(B^0(t) \rightarrow J/\psi K_S^0)}{\Gamma(\bar{B}^0(t) \rightarrow J/\psi K_S^0) + \Gamma(B^0(t) \rightarrow J/\psi K_S^0)} \\ &= S_{J/\psi K_S^0} \sin(\Delta m_d t) - C_{J/\psi K_S^0} \cos(\Delta m_d t) , \end{aligned} \quad (1)$$

where  $\Delta m_d = m_{B_H^0} - m_{B_L^0}$  is the mass difference between the mass eigenstates. We have neglected the decay width difference between the mass eigenstates of the  $B^0$  system,  $\Delta\Gamma \approx 0$ . The sine term results from the interference between direct decay and decay after  $B^0$ - $\bar{B}^0$  mixing. A non-zero cosine term arises from the interference between decay amplitudes with different weak and strong phases (direct  $CP$  violation), or from  $CP$  violation in  $B^0$ - $\bar{B}^0$  mixing. The connection [6] of the  $S_{J/\psi K_S^0}$  and  $C_{J/\psi K_S^0}$  parameters to the CKM angle is

$$S_{J/\psi K_S^0} \simeq \sqrt{1 - C_{J/\psi K_S^0}^2} \sin 2\beta . \quad (2)$$

In the Standard Model, both  $CP$  violation in mixing and direct  $CP$  violation are negligible in  $b \rightarrow c\bar{c}s$  decays. As a consequence the cosine term vanishes, implying

$$S_{J/\psi K_S^0} \simeq \sin 2\beta . \quad (3)$$

In this note, we measure  $S_{J/\psi K_S^0}$  under the assumption that  $C_{J/\psi K_S^0} = 0$ . In addition, we quote the resulting value of  $C_{J/\psi K_S^0}$  if the Standard Model constraint is relaxed.

A mandatory step in measuring the time-dependent asymmetry is to identify (tag) the initial  $B$  flavour. This is done by identifying a flavour specific decay of the other  $b$  hadron produced in the initial  $pp$  collision. Alternatively, the initial  $B$  flavour is tagged by

identifying a charged pion that accompanies the signal  $B$ , compensating the  $d$  quark from the hadronization of the signal  $b$  quark. Because of the imperfect tagging performance and limited proper time resolution the observable asymmetry is diluted:

$$\mathcal{A}_{J/\psi K_S^0}^{\text{meas}}(t) = (1 - 2\omega)\mathcal{A}_{J/\psi K_S^0}(t) \otimes \mathcal{R}(t) . \quad (4)$$

Here,  $\omega$  is the average wrong tag probability (mistag), and  $\mathcal{R}(t)$  is the proper time resolution function. We check the performance of the flavour tagging algorithms on the self-tagging channel  $B^0 \rightarrow J/\psi K^{*0}$ .

We extract the  $S_{J/\psi K_S^0}$  parameter by means of a simultaneous, extended unbinned maximum likelihood fit to the proper time and the invariant mass distributions. The analysis was performed using a “blind” analysis technique: to minimize unconscious experimenter bias, the parameter of interest,  $S_{J/\psi K_S^0}$ , was encrypted in the likelihood fit. Only after the full analysis strategy was developed and proved to be stable, the encryption was removed, unblinding the true result.

This note is organized as follows: Section 2 introduces the LHCb detector and the data set, Section 3 deals with the reconstruction and selection of  $B^0$  candidates, Section 4 gives details on the flavour tagging. The likelihood fit method is described in Section 5, and Section 6 discusses systematic uncertainties and presents the final result.

## 2 The Data Set And LHCb Detector

The results presented in this report are based on data collected with the LHCb detector at the LHC collider at a center-of-mass energy of  $\sqrt{s} = 7 \text{ TeV}$ . Details on the LHCb experiment can be found in Ref. [7]. Designed for precise measurements in the  $B$  system, the LHCb detector is built as a forward spectrometer covering a range of large forward rapidities. The LHCb tracking system is composed of the silicon Vertex Locator (VELO) surrounding the proton-proton interaction point, a tracking station in front of, and the main tracking station behind the LHCb spectrometer magnet with an integrated field of 3.7 Tm. Further, LHCb is equipped with a dedicated particle identification (PID) system, based on two Ring Imaging Cherenkov detectors. One of these is installed in front and a second one is installed behind the magnet. The LHCb trigger system consists of a hardware based L0 trigger and a two staged high level trigger (HLT1 and HLT2) implemented in software. The HLT1 performs a partial event reconstruction to confirm the L0 trigger decision and imposes first selection on track candidates and combined candidates. The second stage, HLT2, performs a full event reconstruction and applies selection criteria close to offline criteria. The HLT contains trigger lines which bias the proper time distribution of selected events as well as lines which do not. We have classified the triggered events into two groups, that have been treated separately in the rest of the analysis. The first group (unbiased) contains events that were triggered by trigger decisions that do not bias the proper time distribution. The second group (biased) contains the remaining events. The data set analyzed has an integrated luminosity of  $\approx 35 \text{ pb}^{-1}$ . We use Monte Carlo (MC) simulated samples that are based on the PYTHIA 6.4 generator [8]. The

EVTGEN package [9] was used to generate hadron decays and the GEANT4 package [10] for detector simulation.

### 3 Reconstruction and Selection of $B$ Candidates

We reconstruct  $J/\psi$  candidates in the decay mode  $J/\psi \rightarrow \mu^+\mu^-$ . We form  $J/\psi$  candidates from pairs of opposite sign tracks that have a transverse momentum of  $p_T > 500$  MeV/ $c$  each and PID signatures consistent with those of muons. The invariant mass of the pair must be compatible with the known  $J/\psi$  mass [11]. We reconstruct  $K_S^0$  candidates through their decay into  $\pi^+\pi^-$ . Here we take pairs of oppositely charged tracks, where both tracks either left hits in all tracking stations, or only in the downstream stations other than the VELO. An additional requirement of  $L/\sigma_L > 5$  is made on the  $K_S^0$  candidate, where  $L$  is the distance between the  $K_S^0$  decay vertex and the  $B$  decay vertex and  $\sigma_L$  is the uncertainty of  $L$ . We constrain the invariant masses of the reconstructed  $J/\psi$  and  $K_S^0$  candidates to their known masses [11]. We reconstruct  $K^{*0}$  candidates through the decay  $K^{*0} \rightarrow K^+\pi^-$ . We require that  $K^+$  candidate tracks have PID signatures consistent with those of kaons. We impose additional constraints on the qualities of all tracks, the vertex quality and the distance of closest approach of the reconstructed  $B$  track with respect to the primary vertex. We also require the  $K^{*0}$  or  $K_S^0$  candidates to have a  $p_T$  greater than 1 GeV/ $c$ . Finally, we form  $B^0$  candidates by combining  $J/\psi$  candidates with either  $K^{*0}$  or  $K_S^0$  candidates. We compute the proper decay time  $t$  of the  $B^0$  candidates from their decay length. For this we use a dedicated vertex fit with the additional constraint that the  $B^0$  momentum points back to the primary vertex.

### 4 Flavour Tagging

The initial  $B$  flavour is determined by the combination of various tagging algorithms. These either determine the flavour of the non-signal  $b$  hadron produced in the event (*opposite side*, OS), or they search for an additional pion accompanying the signal  $B^0$  or  $\bar{B}^0$  (*same side*, SS $\pi$ ). There are four OS taggers. They use the charge of the lepton ( $\mu$ ,  $e$ ) from semileptonic  $B$  decays, or that of the kaon from the  $b \rightarrow c \rightarrow s$  decay chain, or the charge of the inclusive secondary vertex reconstructed from  $b$  decay products. All of these algorithms have an intrinsic mistag rate, due to picking up tracks from the underlying event, or due to flavour oscillations of neutral tag  $B$  mesons. For each signal  $B^0$  candidate the tagging algorithms also predict the mistag probability  $\omega$ . For this various kinematic variables such as momenta and angles of the tagging particles are combined into neural networks. The neural networks are trained on MC simulated events.

The flavour asymmetry that is accessible in  $B^0 \rightarrow J/\psi K_S^0$  decays directly depends on the dilution  $D$  due to the mistag probability,  $D = 1 - 2\omega$ . Its statistical precision is proportional to the inverse square root of the effective tagging efficiency  $\varepsilon_{\text{eff}}$ ,

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}(1 - 2\omega)^2, \quad (5)$$

where  $\varepsilon_{\text{tag}}$  is the probability that a tagging decision is found. The tagging algorithms are optimized for highest  $\varepsilon_{\text{eff}}$  on data, using the self-tagging decays  $B^+ \rightarrow J/\psi K^+$  and  $B^0 \rightarrow D^{*-} \mu^+ \nu$ . The estimated mistag probability is calibrated on these same channels.

The mistag probability depends on the final state considered. Therefore it is crucial to validate the predicted mistag probability on a self-tagging control channel that is similar to the signal decay in terms of trigger, reconstruction and kinematics. For  $B^0 \rightarrow J/\psi K_S^0$  such a control channel is  $B^0 \rightarrow J/\psi K^{*0}$ ,  $K^{*0} \rightarrow K^+ \pi^-$ , where the charge of the kaon indicates the  $B^0$  flavour at decay time. In MC simulated events the differences between the mistag distributions in  $B^0 \rightarrow J/\psi K^{*0}$  and  $B^0 \rightarrow J/\psi K_S^0$  are negligible. Therefore it is sufficient to check on data whether the calibrated mistag probability is valid for the  $B^0 \rightarrow J/\psi K^{*0}$  channel. Because in the  $B^0 \rightarrow J/\psi K^{*0}$  channel the  $B^0$  meson oscillates, the mistag probability can only be determined through a fit to the time-dependent flavour asymmetry (described in the following section). In this fit we include a linear calibration function, and we fit its coefficients  $p_0$  and  $p_1$ ,

$$\omega = p_0 + p_1 \cdot (\eta - \langle \eta \rangle) . \quad (6)$$

Here,  $\eta$  is the calibrated mistag probability as estimated by the tagging algorithms, and  $\langle \eta \rangle$  is its average value. This parameterization minimizes the correlation between  $p_0$  and  $p_1$ . If the calibrated mistag probability is valid for  $B^0 \rightarrow J/\psi K^{*0}$  we expect  $p_0 = \langle \eta \rangle = 0.35$  and  $p_1 = 1$  [12]. We find good agreement with these values,  $p_0 = 0.333 \pm 0.025$  and  $p_1 = 0.71 \pm 0.36$ . Therefore we conclude that the established flavour tagging algorithms are suitable to our measurement. We measure the effective tagging efficiency to be  $\varepsilon_{\text{eff}} = (2.82 \pm 0.87)\%$  [12].

## 5 Likelihood Fit to $B^0 \rightarrow J/\psi K_S^0$

The  $CP$  violation parameters  $S_{J/\psi K_S^0}$  and  $C_{J/\psi K_S^0}$  are extracted through a simultaneous multidimensional unbinned extended maximum likelihood fit. We define the extended likelihood function for  $n$  observed events as

$$\mathcal{L}(\vec{\lambda}) = \frac{e^{-N} N^n}{n!} \prod_s \prod_{i=1}^{N^s} \mathcal{P}^s(\vec{x}_i; \vec{\lambda}_s) . \quad (7)$$

We minimize  $-\ln \mathcal{L}$  to find optimal values for the fit parameters  $\vec{\lambda}$ . The fit is simultaneous in four subsamples  $s$ . These subsamples are defined by whether the candidates were triggered by the lifetime unbiased (“U”) or the lifetime biased (“B”) trigger lines; and by whether or not a tagging decision is available (“u” for untagged, “t” for tagged). Each subsample contains  $N^s$  events, and  $n = \sum_s N^s$ . The likelihood contains a Poisson term that reflects the probability that  $n$  events are observed, when  $N$  are expected. The expected number of events  $N$  is estimated by the fit. We consider four observables,  $\vec{x}^T = (m, t, d, \omega)^T$ . These are the reconstructed mass  $m$  of the  $B^0$  candidate ( $5.15 \text{ GeV}/c^2 < m < 5.4 \text{ GeV}/c^2$ ), its proper time  $t$  ( $-1 \text{ ps} < t < 4 \text{ ps}$ ), the flavour tag decision  $d$ , and the

combined per-event mistag prediction  $\omega$ . The flavour tag  $d$  can take the discrete values  $d = 1$  if tagged as initial  $B^0$  and  $d = -1$  if tagged as initial  $\bar{B}^0$ . The probability density functions (p.d.f.s)  $\mathcal{P}^s$  consist of three components, signal (S), prompt background (P), and long lived background (L):

$$y^s \mathcal{P}^s = N_S^s \mathcal{P}_S^s + N_P^s \mathcal{P}_P^s + N_L^s \mathcal{P}_L^s . \quad (8)$$

The total yields in the subsamples fulfill the normalization conditions  $y^s = N_S^s + N_P^s + N_L^s$  and  $N = \sum_s y^s$ . We assume negligible correlations across the two continuous observables, thus each component p.d.f. factorizes into independent mass and proper time terms. The mass term is independent from the tagging decision  $d$  and from the per-event mistag prediction  $\omega$ .

The mass p.d.f. of the signal component consists of a single Gaussian. We assume both background components have similar mass distributions, and we use the same parameterization for both. It is modeled as an exponential with a single shape parameter. We validate this assumption using an sPlot method [13] with the proper time distribution as discriminating variable.

The proper time p.d.f. of the signal component can be written as  $\mathcal{P}_S(t, d, \omega) = \mathcal{P}_S(t, d|\omega) \cdot \mathcal{P}_S(\omega)$ . The first term is a conditional p.d.f. as it depends on the value of  $\omega$ , the second term describes the distribution of  $\omega$ . The background parameterization of the proper time factorizes,  $\mathcal{P}_B(t, d, \omega) = \mathcal{P}_B(t, d) \cdot \mathcal{P}_B(\omega)$ , where  $B = P, L$ .

The conditional proper time p.d.f. of the  $B^0 \rightarrow J/\psi K_S^0$  signal is given by

$$\mathcal{P}_S(t, d|\omega) = \epsilon_S(t) \cdot \left( \mathcal{P}_{S,CP}(t, d|\omega) \otimes \mathcal{R}(t) \right) . \quad (9)$$

Here, the p.d.f.  $\mathcal{P}_{S,CP}(t, d|\omega)$  describes decay, mixing and  $CP$  violation in the  $B^0$  system, and thus depends on the  $CP$  violation parameters  $S_{J/\psi K_S^0}$  and  $C_{J/\psi K_S^0}$ . It is convoluted with a triple Gaussian resolution function  $\mathcal{R}(t)$  which is taken to be the same between all subsamples of the simultaneous fit. The efficiency function  $\epsilon_S(t)$  describes acceptance effects observed in the biased subsample at low proper times. We describe it empirically by

$$\epsilon_S(t) = \frac{(a_a t)^{c_a}}{1 + (a_a t)^{c_a}} . \quad (10)$$

No signal candidates in the biased sample are expected to have negative proper times, so  $\epsilon_S(t < 0) \equiv 0$ . We extract the parameters  $a_a$  and  $c_a$  from the data, using an sPlot method and events that were triggered by both the biased and unbiased lines:  $a_a = (3.7 \pm 1.0) \text{ ps}^{-1}$ ,  $c_a = 1.9 \pm 1.0$ . For candidates in the unbiased subsamples, the efficiency is by definition uniform.

We also observe an acceptance effect at high lifetimes [14], affecting both biased and unbiased candidates. To account for this effect, we apply an efficiency function of the form  $e^{\beta_a t}$ . We determine  $\beta_a = (-0.0037 \pm 0.0033) \text{ ps}^{-1}$  using MC. Hence the correction

needs to be applied before convolving with  $\mathcal{R}(t)$ . This is done by transforming the fitted lifetime:

$$\exp(\beta_a t) \exp\left(-\frac{t}{\tau}\right) = \exp\left(\beta_a t - \frac{t}{\tau}\right) = \exp\left(-\frac{t}{\tau'}\right) \quad (11)$$

with

$$\tau' = \frac{\tau}{1 - \beta_a \tau}. \quad (12)$$

In the tagged samples the following p.d.f. is used to describe decay, mixing and  $CP$  violation of the  $B^0$  mesons:

$$\begin{aligned} \mathcal{P}_{S,CP}^t(t, d|\omega) \propto e^{-t/\tau'} & (1 - d(1 - 2\omega) S_{J/\psi K_S^0} \sin \Delta m_d t \\ & + d(1 - 2\omega) C_{J/\psi K_S^0} \cos \Delta m_d t). \end{aligned} \quad (13)$$

In the case of the untagged subsamples, it reduces to  $\mathcal{P}_{S,CP}^u(t) \propto e^{-t/\tau'}$ .

The  $B^0$  lifetime  $\tau$  is shared between the fits of all four subsamples, while the other  $B^0$  system physics parameters  $\Delta m_d$ ,  $S_{J/\psi K_S^0}$  and  $C_{J/\psi K_S^0}$  are shared between the tagged subsamples.

The proper time p.d.f. of the prompt background is just  $\mathcal{R}(t)$ . That of the long lived background in the unbiased samples is the sum of two exponentials with different pseudo lifetimes  $\tau_1^L$  and  $\tau_2^L$  and a relative fraction  $f_t^L$ . In the biased samples we use a single exponential. These exponentials are also convoluted with the proper time resolution  $\mathcal{R}(t)$ , in order to properly describe the drop-off at  $t < 0$ .

The per-event mistag probability p.d.f. is extracted from the tagged sample in data by using an sPlot technique. An unbinned likelihood fit on the reconstructed mass with a single Gaussian parameterization for the signal and an exponential function for the background description are used to disentangle the per-event mistag probability distribution for signal and background. These distributions are then used to create histogram p.d.f.s for signal and background, both using eleven equidistant bins.

To extract the tagging calibration parameters  $p_0$  and  $p_1$  from the  $B^0 \rightarrow J/\psi K^{*0}$  control channel we use nearly the same parameterizations for the mass and the proper time background distributions. However, for the signal proper time of the tagged candidates a different p.d.f. has to be chosen. This p.d.f. depends on the mix state  $q$ , which is determined by comparing the tagging decision and the reconstructed flavour. If the reconstructed flavour and the tagging decision agree, it is assumed that the meson has not mixed ( $q = 1$ ). If they disagree it is assumed that the meson has mixed ( $q = -1$ ). The resulting signal proper time p.d.f. of the  $B^0 \rightarrow J/\psi K^{*0}$  signal is given as

$$\mathcal{P}_{S,t}^t(t, q|\omega) \propto e^{-t/\tau'} (1 + q(1 - 2\omega) \cos(\Delta m_d t)) \otimes \mathcal{R}(t). \quad (14)$$

Also the larger sample sizes in this channel require a quadruple Gaussian proper time resolution, and the acceptance parameters  $\beta_a$ ,  $a_a$  and  $c_a$  differ from those for  $B^0 \rightarrow J/\psi K_S^0$ .

In the fit to the  $B^0 \rightarrow J/\psi K_S^0$  channel we fix the mixing frequency  $\Delta m_d$  to its nominal value of  $\Delta m_d = (0.507 \pm 0.005) \cdot 10^{12} \text{ } \hbar s^{-1}$  [11], and  $C_{J/\psi K_S^0} = 0$ . In total, there are 27 floating parameters: the  $CP$  parameter  $S_{J/\psi K_S^0}$ , the  $B^0$  lifetime  $\tau$ , the  $B^0$  mass  $m_0^S$ , twelve event yields, four parameters of the long-lived proper time background, five parameters of the time resolution, the mass signal resolution  $\sigma_m^S$ , and two parameters of the mass background shape.

We have checked the fit implementation on a large sample of MC generated signal events, where we find good agreement with the generated values. To validate the fit further, we generate a large number of toy data samples according to the nominal p.d.f. The generated numbers of events in the signal and background components correspond to the yields found in the fit to the nominal sample. We fit the toy samples using the nominal p.d.f., and calculate for each parameter  $\lambda$  the pull, defined as  $p = (\lambda_{\text{fit}} - \lambda_{\text{gen}})/\sigma_{\text{fit}}$ , where  $\lambda_{\text{fit}}$  is the fitted value of  $\lambda$ ,  $\lambda_{\text{gen}}$  is the value used in the generation of the toy sample, and  $\sigma_{\text{fit}}$  is the error of  $\lambda_{\text{fit}}$  as reported by the fit. The pulls are expected to follow a Gaussian distribution with zero mean and unit width. We find the pull distributions for all fit parameters to be well compatible with the expectation. For the parameter of interest,  $S_{J/\psi K_S^0}$ , we find the mean  $\mu_p$  and width  $\sigma_p$  of the pull distribution to be  $\mu_p = -0.052 \pm 0.023$  and  $\sigma_p = 1.016 \pm 0.016$ .

## 6 Results

The result of the maximum likelihood fit to the full data sample is summarized in Tables 1 and 2. The mass and proper time distributions and the fit projections are shown in Figure 1. Figure 2 shows the resulting time dependent raw asymmetry, which contains all fit components. The asymmetry in the lowest proper time bins is therefore dominated by the backgrounds, whereas the measured asymmetry in the high proper time bins is dominated by signal events, see Eq. 4. The measured value of  $S_{J/\psi K_S^0}$  is

$$S_{J/\psi K_S^0} = 0.53_{-0.29}^{+0.28} , \quad (15)$$

where the error is statistical only. We find the global correlation coefficient of  $S_{J/\psi K_S^0}$  to be  $\rho(S_{J/\psi K_S^0}) = 0.016$ . We also perform the nominal fit without the Standard Model constraint  $C_{J/\psi K_S^0} = 0$ . In this case, we find

$$C_{J/\psi K_S^0} = 0.28 \pm 0.32 , \quad (16)$$

$$S_{J/\psi K_S^0} = 0.38 \pm 0.35 , \quad (17)$$

again quoting statistical errors only. The correlation coefficient between the parameters is  $\rho(S_{J/\psi K_S^0}, C_{J/\psi K_S^0}) = 0.53$ . Their correlations to other parameters are negligible.

Table 1: Fit result of the nominal fit to the full  $B^0 \rightarrow J/\psi K_S^0$  data sample. The parameters are explained in Section 5.

Parameter	Unit	Fitted Value
$S_{J/\psi K_S^0}$		$0.53^{+0.28}_{-0.29}$
$m_S$	MeV/ $c^2$	$5278.13 \pm 0.29$
$\sigma_{S,m}$	MeV/ $c^2$	$8.82 \pm 0.24$
$\tau$	ps	$1.517 \pm 0.046$
$\alpha_m^B$	(MeV/ $c^2$ ) $^{-1}$	$-8.71 \pm 3.8 \cdot 10^{-4}$
$\alpha_m^U$	(MeV/ $c^2$ ) $^{-1}$	$-5.86 \pm 0.87 \cdot 10^{-4}$
$f_{L,t}^U$		$0.836 \pm 0.054$
$\tau_{L,1}^U$	ps	$0.221 \pm 0.036$
$\tau_{L,2}^U$	ps	$1.04 \pm 0.24$
$\tau_L^B$	ps	$0.482 \pm 0.029$
$f_{\mathcal{R},1}$		$0.500 \pm 0.019$
$f_{\mathcal{R},2}$		$0.477 \pm 0.017$
$\sigma_{\mathcal{R},1}$	ps	$0.02522 \pm 0.00066$
$\sigma_{\mathcal{R},2}$	ps	$0.0685 \pm 0.0016$
$\sigma_{\mathcal{R},3}$	ps	$0.293 \pm 0.019$

Table 2: Fitted event yields in the full  $B^0 \rightarrow J/\psi K_S^0$  data sample.

Sample	Parameter	Fitted Value
U,t	$N_S^{U,t}$	$221 \pm 17$
	$N_P^{U,t}$	$3218 \pm 62$
	$N_L^{U,t}$	$309 \pm 33$
B,t	$N_S^{B,t}$	$59.8 \pm 8.7$
	$N_P^{B,t}$	$164 \pm 14$
	$N_L^{B,t}$	$102 \pm 12$
U,u	$N_S^{U,u}$	$767 \pm 32$
	$N_P^{U,u}$	$21134 \pm 161$
	$N_L^{U,u}$	$807 \pm 79$
B,u	$N_S^{B,u}$	$279 \pm 18$
	$N_P^{B,u}$	$747 \pm 30$
	$N_L^{B,u}$	$339 \pm 23$

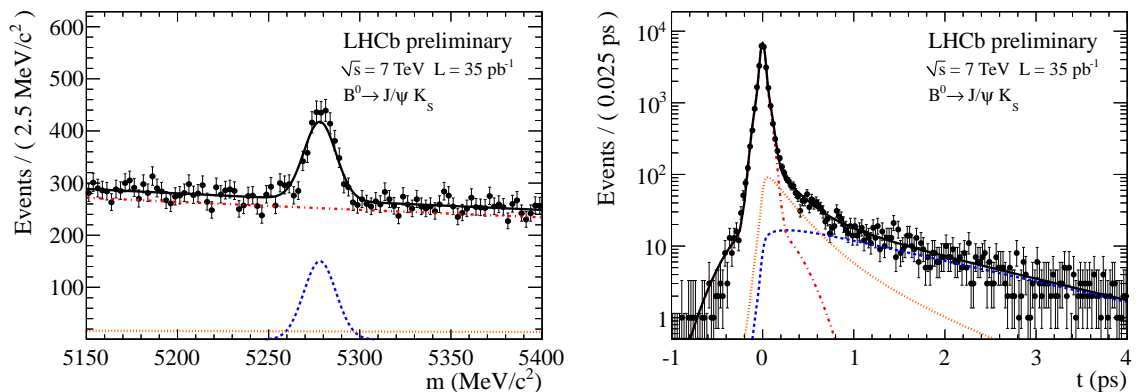


Figure 1: Reconstructed mass (left) and proper time (right) distributions of  $B^0 \rightarrow J/\psi K_S^0$  candidates. Overlaid are projections of the component p.d.f.s used in the fit: full p.d.f. (solid black), signal (dashed blue), prompt background (dash-dotted red), long lived background (dotted orange).

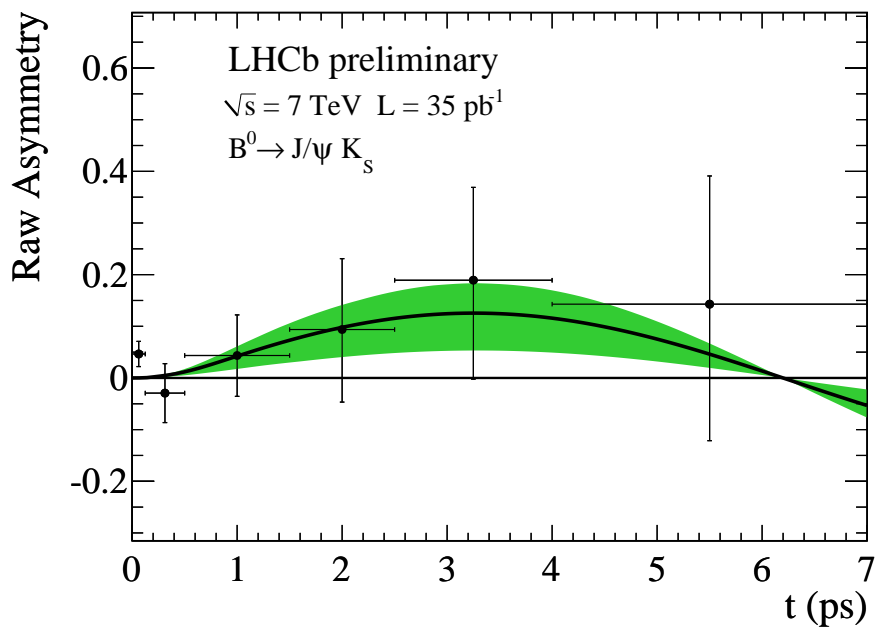


Figure 2: Time dependent raw  $CP$  asymmetry in  $B^0 \rightarrow J/\psi K_S^0$ . The solid curve is the full p.d.f. (signal and background) overlaid onto the data points. The green band corresponds to the one standard deviation statistical error.

## 7 Systematic Untertainties

We consider eight sources of systematic uncertainty, possibly affecting the measurement of  $S_{J/\psi K_S^0}$ . They are summarized in Table 3 and described in turn. This measurement is clearly statistically limited. We therefore estimate the systematic errors in a conservative fashion.

Table 3: Systematic uncertainties to  $S_{J/\psi K_S^0}$  in absolute terms.

Source	uncertainty
tagger calibration	0.044
per-event mistags p.d.f.	0.016
$\Delta m_d$ uncertainty, $z$ scale	0.0017
proper time resolution	0.0085
high proper time acceptance	0.0018
biased events acceptance	0.0039
biased TIS events acceptance	0.0063
production asymmetry	0.024
total (sum in squares)	0.054

The leading contribution arises from the uncertainty in the tagger calibration. This is expressed through the uncertainty of the calibration parameters  $p_k$ ,  $k = 0, 1$ . We propagate their uncertainties as estimated on the  $B^+ \rightarrow J/\psi K^+$  sample into the effective dilution  $D_{\text{eff}}^2 = 1/N \sum_i^N D_i^2$ . Here  $D_i$  is the dilution of the  $i^{\text{th}}$  event of a high signal purity sample of size  $N$ . This results in a 6.5% relative error. The signal p.d.f. (see Eq. 13) contains the term  $DS_{J/\psi K_S^0}$ , where  $D = (1 - 2\omega)$  is evaluated on a per-event basis. Therefore we assign a 6.5% relative systematic error to  $S_{J/\psi K_S^0}$ , where we conservatively assume the world average for the central value. We note that this source of uncertainty will decrease as the size of the control samples increases.

The nominal p.d.f. contains histogram descriptions of the distributions of the per-event mistag probabilities. We estimate the systematic error due to this choice by varying the number of bins and bin borders. We perform two test fits. At first, we use a different binning for signal and background: 50 and 9 equidistant bins, respectively. The second test fit uses a binning defined by the four tagging categories used in Ref. [15], both for signal and background. We sum the deviations of the  $S_{J/\psi K_S^0}$  central values of the test fits in squares and take this as the systematic uncertainty on  $S_{J/\psi K_S^0}$ .

The  $B^0$  mixing frequency  $\Delta m_d$  is fixed to its PDG value of  $(0.507 \pm 0.005)\hbar \text{ps}^{-1}$  [11] in the nominal fit. We propagate its uncertainty into  $S_{J/\psi K_S^0}$  by means of two test fits, in which we increase and decrease the fixed value of  $\Delta m_d$  by its error. Both deviations go in the same direction. We take their squared sum to be the systematic uncertainty. This also accounts for a 0.1% uncertainty on the  $z$  scale.

We study the effect of the proper time resolution model on the measurement by performing two test fits. We fix the width parameters of the three Gaussians to 80% and

120% of their nominal values. Both deviations go in the same direction. We take their squared sum to be the systematic uncertainty.

The nominal p.d.f. contains an acceptance correction at high proper times. To account for systematic uncertainties related to this, we perform two test fits, in which we vary the shape parameter  $\beta_{s,a} = -0.0037 \pm 0.0033 \text{ ps}^{-1}$  up and down by its error. Both deviations go in the same direction. We take their squared sum to be the systematic uncertainty.

The nominal p.d.f. contains a proper time acceptance correction for events triggered by the biased lines. We propagate the uncertainty of the parameters of the acceptance function,  $a_a$  and  $c_a$  by means of four test fits, each parameter varying by its error. We sum the same-sign deviations in squares and assign the larger of both summed deviations as the systematic uncertainty.

The biased sample contains a fraction of events, that were triggered independently from the signal (TIS). Their proper time acceptance is therefore not properly described by the acceptance function. To estimate the resulting uncertainty, we perform a test fit to a sample without these events. We use half of the deviation in  $S_{J/\psi K_s^0}$  as the systematic uncertainty.

Any  $CP$  violation measurement can possibly be altered by a non-zero production asymmetry. In a recently presented LHCb measurement of the  $CP$  violation in  $B^0 \rightarrow K^+\pi^-$  decays [16] the production asymmetry was estimated to be compatible with zero,  $|A_p| = 0.024 \pm 0.016$ . To estimate the influence to our measurement, we generate toy experiments with a  $\pm 2.4\%$  asymmetry in the  $B^0$  and  $\bar{B}^0$  tagging efficiency. We fit back with the nominal p.d.f. We observe a symmetric shift of  $\pm 0.024$  in  $S_{J/\psi K_s^0}$  and assign it as systematic uncertainty.

We implemented an alternative version of the likelihood fitter to strengthen the confidence in the nominal result. Here we use a different background model for the invariant mass: a second order polynomial rather than an exponential. Other differences are that we require the candidates in the biased samples to have a proper time greater than  $t > 0.05 \text{ ps}$ , and that the alternative likelihood doesn't contain a description of the  $\omega$  probabilities. The difference in  $S_{J/\psi K_s^0}$  to the nominal result is 0.0016.

## 8 Conclusion

The final result on the  $CP$  violation parameter  $S_{J/\psi K_s^0}$  is

$$S_{J/\psi K_s^0} = 0.53_{-0.29}^{+0.28}(\text{stat}) \pm 0.05(\text{syst}) . \quad (18)$$

This result is compatible with the World Average, and dominated by the statistical uncertainty. We calculate the statistical significance of a non-zero  $CP$  violation from the likelihood ratio of a test fit, in which we fix  $S_{J/\psi K_s^0} = 0$ , to be

$$S = \sqrt{2 \ln(\mathcal{L}_{\text{nom}}/\mathcal{L}_{\text{null}})} = 1.8 . \quad (19)$$

This is the first  $CP$  violation result in the golden channel  $B^0 \rightarrow J/\psi K_s^0$  in LHCb.

## References

- [1] The BABAR Collaboration, B. Aubert *et al.*, Phys. Rev. D **79**, 072009 (2009).
- [2] The Belle Collaboration, K.-F. Chen *et al.*, Phys. Rev. Lett. **98**, 031802 (2007).
- [3] The HFAG Collaboration, E. Barbiero *et al.*, arXiv:0808.1297, online update at <http://www.slac.stanford.edu/xorg/hfag/triangle/latest/#sin2b>.
- [4] I. Bigi and A. Sanda, CP violation, Cambridge Univ. Press, 2000.
- [5] P. Ball *et al.*, CERN-TH/2000-101, arXiv:hep-ph/0003238.
- [6] S. Faller, M. Jung, R. Fleischer, and T. Mannel, Phys.Rev. **D79**, 014030 (2009), arXiv:0809.0842.
- [7] The LHCb Collaboration, A. A. Alves *et al.*, **3**, S08005 (2008).
- [8] T. Sjöstrand *et al.*, JHEP **05**, 26 (2006).
- [9] D. J. Lange, Nucl. Instrum. Methods A **462** (2001).
- [10] The GEANT4 Collaboration, S. Agostinelli *et al.*, Nucl. Instrum. Methods A **506** (2003).
- [11] Particle Data Group, C. Amsler *et al.*, Physics Letters B **667**, 1 (2008).
- [12] The LHCb Collaboration, R. Aaij *et al.*, Optimization and calibration of the tagging performances and measurement of  $\Delta m_d$ , LHCb-CONF-2011-003, 2011.
- [13] M. Pivk and F. L. Diberder, sPlot: A Statistical Tool to Unfold Data Distributions, arXiv:physics/04020834.
- [14] The LHCb Collaboration, R. Aaij *et al.*, Selections, relative branching fractions and lifetime measurements for exclusive  $B \rightarrow J/\psi X$  decays with  $J/\psi \rightarrow \mu^+ \mu^-$  with the 2010 data, LHCb-CONF-2011-001, 2011.
- [15] M. Calvi *et al.*, Calibration of Flavour Tagging with  $B^+ \rightarrow J/\psi K^+$  and  $B^0 \rightarrow J/\psi K^*$  control channels at LHCb, CERN-LHCb-2009-020, 2009.
- [16] The LHCb Collaboration, R. Aaij *et al.*, Measurement of direct CP violation in charmless charged two-body B decays at LHCb, LHCb-CONF-2011-011, 2011.