

First measurement of CP violation in $B_S^0 \rightarrow J/\psi K_S^0$ decays with the LHCb experiment

Dissertation zur Erlangung des akademischen Grades
Dr. rer. nat.

vorgelegt von
Florian Kruse
geboren am 01.09.1982 in Wesel

Fakultät Physik
Technische Universität Dortmund

Dortmund, im September 2015

Der Fakultät Physik der Technischen Universität Dortmund zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften vorgelegte Dissertation.

1. Gutachter: Prof. Dr. Bernhard Spaan

2. Gutachter: Priv.-Doz. Dr. Reiner Klingenberg

Vorsitzender der Prüfungskommission: Prof. Dr. Heinz Hövel

Vertreter der wissenschaftlichen Mitarbeiter: Dr. Bärbel Siegmann

Vorsitzender des Promotionsausschusses: Prof. Dr. Thomas Weis

Abstract

This thesis presents the first measurement of CP violation in decays of B_s^0 and \bar{B}_s^0 mesons into the final state $J/\psi K_S^0$, where the CP observables allow to constrain the contribution of penguin topologies in $B^0 \rightarrow J/\psi K_S^0$ decays. The measurement is performed on a dataset corresponding to an integrated luminosity of 3 fb^{-1} of proton-proton collisions, which were recorded by the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. Using an unbinned maximum-likelihood fit, the CP observables are measured as

$$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} = 0.49 \pm_{0.65}^{0.77} \text{ (stat)} \pm 0.06 \text{ (syst)},$$

$$C_{B_s^0 \rightarrow J/\psi K_S^0} = -0.28 \pm 0.41 \text{ (stat)} \pm 0.08 \text{ (syst)},$$

$$S_{B_s^0 \rightarrow J/\psi K_S^0} = -0.08 \pm 0.40 \text{ (stat)} \pm 0.08 \text{ (syst)}.$$

Furthermore, the ratio of branching fractions $\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0) / \mathcal{B}(B^0 \rightarrow J/\psi K_S^0)$ is determined as

$$0.0431 \pm 0.0017 \text{ (stat)} \pm 0.0012 \text{ (syst)} \pm 0.0025 (f_s/f_d),$$

where the last uncertainty is due to the limited knowledge of the ratio f_s/f_d of B_s^0 to B^0 meson hadronisation fractions.

Zusammenfassung

Diese Arbeit beschreibt die erste Messung von CP -Verletzung in Zerfällen von B_s^0 - und \bar{B}_s^0 -Mesonen in den Endzustand $J/\psi K_S^0$, wobei die CP -Observablen eine Bestimmung der Penguinbeiträge im Zerfallskanal $B^0 \rightarrow J/\psi K_S^0$ erlauben. Die Messungen werden auf einem vom LHCb-Experiment bei Schwerpunktsenergien von 7 und 8 TeV aufgenommenen Datensatz von Proton-Proton-Kollisionen durchgeführt, der einer integrierten Luminosität von 3 fb^{-1} entspricht. Mithilfe eines ungebinnten Maximum-Likelihood-Fits werden die folgenden CP -Observablen gemessen:

$$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} = 0.49 \pm_{0.65}^{0.77} \text{ (stat)} \pm 0.06 \text{ (syst)},$$

$$C_{B_s^0 \rightarrow J/\psi K_S^0} = -0.28 \pm 0.41 \text{ (stat)} \pm 0.08 \text{ (syst)},$$

$$S_{B_s^0 \rightarrow J/\psi K_S^0} = -0.08 \pm 0.40 \text{ (stat)} \pm 0.08 \text{ (syst)}.$$

Weiterhin ergibt die Bestimmung des Quotienten der Verzweungsverhältnisse $\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0) / \mathcal{B}(B^0 \rightarrow J/\psi K_S^0)$:

$$0.0431 \pm 0.0017 \text{ (stat)} \pm 0.0012 \text{ (syst)} \pm 0.0025 (f_s/f_d),$$

wobei sich die letzte Unsicherheit auf die begrenzte Kenntnis des Produktionsverhältnisses zwischen B_s^0 - und B^0 -Mesonen bezieht.

Contents

1	Introduction	1
2	CP violation in $B \rightarrow J/\psi K_S^0$ decays	5
2.1	Fundamentals of the Standard Model of particle physics	5
2.2	CKM mechanism	7
2.3	Decay and mixing formalism of neutral mesons	10
2.4	CP violation in neutral meson decays	13
2.4.1	CP violation in the decay	14
2.4.2	CP violation in the mixing	14
2.4.3	CP violation in the interference of mixing and decay	15
2.5	The decays $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$	15
2.5.1	Mixing in the $B^0-\bar{B}^0$ and $B_s^0-\bar{B}_s^0$ meson systems	16
2.5.2	Decay amplitudes and CP violation in $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$	17
2.5.3	Current status in experiment and theory	21
3	The LHCb experiment	23
3.1	The Large Hadron Collider	23
3.1.1	Accelerator and collider	23
3.1.2	LHC experiments	25
3.2	The LHCb detector	25
3.2.1	Tracking system	27
3.2.2	Particle identification system	29
3.3	Data processing at LHCb	31
3.3.1	Trigger	31
3.3.2	Reconstruction and analysis	33
3.3.3	Simulation	34
3.4	Conditions of the data-taking in Run I	34
3.5	Upgrade of the LHCb detector	35
4	Analysis ingredients and preparatory studies	37
4.1	Data sample preparation	37
4.1.1	Event samples	37
4.1.2	Candidate reconstruction	39
4.1.3	Stripping selection	41

4.1.4	Trigger requirements	42
4.1.5	Backgrounds	45
4.1.6	Further preparatory selection requirements	47
4.1.7	Removal of $B^0 \rightarrow J/\psi K^{*0}$ background	49
4.1.8	Neural network selection	51
4.1.9	Nominal data sample	52
4.2	Flavour tagging studies	52
4.2.1	Principles of the flavour tagging at LHCb	54
4.2.2	Tagging quantities	56
4.2.3	Combination and calibration of tagging algorithms	57
4.2.4	Flavour tagging in $B_s^0 \rightarrow J/\psi K_S^0$	58
4.2.5	Calibration of the opposite-side taggers	59
4.2.6	Calibration of the same-side kaon tagger	60
4.2.7	Studies on the same-side kaon tagger on B^0 candidates	61
4.2.8	Tagging performance	63
4.3	Cubic splines as a tool for empirical parameterisation	65
4.4	Decay-time resolution	66
4.4.1	Parameterisation of the resolution	66
4.4.2	Calibration of the resolution parameters	67
4.4.3	Validation of the resolution parameterisation	71
4.5	Decay-time acceptance	72
4.6	Production asymmetry	72
4.7	Branching ratio measurement	75
5	Measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$	77
5.1	Maximum-likelihood fit	77
5.1.1	Observables	79
5.1.2	Parameterisation	79
5.1.3	External inputs	83
5.2	Preliminary results	83
5.3	Fit validation	85
5.3.1	Studies using pseudo-experiments	87
5.3.2	Comparisons between independent fitter implementations and subsamples	88
5.3.3	sFit implementation	91
5.3.4	Fits on simulated samples	93
5.3.5	Likelihood profile scans	96
5.3.6	Scan of the CP parameter space	96
5.3.7	Feldman–Cousins method	100
5.4	Systematic uncertainties	101
5.4.1	Method to evaluate systematic uncertainties	102

5.4.2	Choice of parameterisation	103
5.4.3	Ignored contributions from non-prompt B_s^0 production . . .	106
5.4.4	Mass resolution differences	107
5.4.5	Momentum and decay-length scale	108
5.4.6	Correlation between mass and decay-time resolution	108
5.4.7	Combined systematic uncertainty	109
5.5	Final results	110
5.5.1	Confidence intervals from the Feldman–Cousins method . .	111
6	Conclusion and outlook	113
Appendices		
A	Supplementary material	117
A.1	Tagging performance of the neural-net same-side kaon tagger . . .	117
A.2	Distributions and parameterisations of conditional observables . .	119
A.3	Detailed fit results	125
A.4	Detailed results of fits on simulated samples	128
A.4.1	Results for the decay-time resolution and acceptance	128
A.4.2	Results for the flavour tagging calibration	130
A.4.3	Results for the nuisance parameters on simulated samples .	131
	Bibliography	135
	Acknowledgements (Danksagung)	145

1 Introduction

The field of modern particle physics is a relatively young discipline, where most discoveries were made in the last hundred years. This has led to the Standard Model of particle physics (SM), a comprehensive and very well-tested theory with extensive predictive power, which describes elementary particles and their interactions. As a relativistic quantum field theory, it is invariant under the continuous symmetries describing the interactions. However, this does not apply to discrete symmetries. While nature is supposed to be invariant under CPT , i.e. the combined operation of charge conjugation C , parity P , and time reversal T , violation of the individual discrete symmetries and the combined CP symmetry are allowed in the SM. Parity violation in weak interactions was first discussed by Lee and Yang in 1956 [1] and observed shortly after in 1957 by Wu et al. in β^- decays [2]. CP violation was then first discovered in neutral kaon decays by Christenson, Cronin, Fitch, and Turlay in 1964 [3]. The first observations of CP violation in the system of B mesons were made by the B factories BABAR and Belle in $B^0 \rightarrow J/\psi K_S^0$ decays in 2001 [4, 5]. Through the unification of the electromagnetic and weak interaction utilising spontaneous symmetry breaking [6, 7], CP violation is incorporated into the SM by postulating a third quark generation, which introduces the Cabibbo-Kobayashi-Maskawa (CKM) matrix with the necessary complex phase [8, 9]. This matrix describes flavour-changing quark transitions in charged weak currents.

Nevertheless, the SM has flaws, where it fails to describe particular observations. For instance, it does not incorporate gravity and is unable to provide explanations for dark matter and dark energy, which are known to dominate the energy content in the universe. Another shortcoming, which is relevant in the scope of this thesis, is the matter-antimatter asymmetry in today's universe. Astronomical observations suggest that the universe is dominated by matter. However, it is assumed that matter and antimatter have been produced in identical amounts in the Big Bang. Sakharov postulated three conditions which would allow for the emergence of this asymmetry [10]: Baryon number violation, C and CP violation, and thermal non-equilibrium. Although CP violation is incorporated in the SM, its amount is too small so that the mechanism cannot account for the observed asymmetry. Therefore, the search for new physics beyond the SM and especially for other sources of CP violation is an important goal in particle physics. All measurements of CP violation performed so far are in very good agreement with the SM.

The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) is constructed to put the SM to stringent tests and to search for new physics. One LHC experiment is the Large Hadron Collider beauty (LHCb) experiment, which focuses on precision measurements of CP violating parameters and branching ratios of rare decays in the beauty and charm sector. This allows to check the SM for consistency by improving the precision of already measured quantities, where a significant contradiction between two quantities would indicate new physics. Additionally, LHCb is able to explore new areas of decays, like those in the B_s^0 meson system, which have not been properly accessible before. Being a dedicated flavour physics detector located at a hadron collider, the detector design differs both from the previous B factories and from other general purpose detectors at the LHC. The hadron collider comes with the benefit of a high rate of b and c quark production at the cost of a very harsh environment with high track multiplicities in each event. This poses several challenges and makes high demands on the detector and the subsequent data processing. After the first run of data taking, LHCb successfully performed many key measurements [11].

In the aforementioned decay channel $B^0 \rightarrow J/\psi K_S^0$, CP violation has already been measured with very high precision. It is additionally only affected by small theoretical uncertainties, making it the so-called gold-plated decay channel to study CP violation in the $B^0-\bar{B}^0$ meson system. The measurement allows access to β , one of the parameters describing the CKM matrix. However, this usually requires to neglect suppressed higher-order contributions from penguin topologies, which is only justified to a certain precision. The latest LHCb measurement in $B^0 \rightarrow J/\psi K_S^0$ [12] already comes close to the precision achieved at the B factories and future measurements will allow to substantially improve the experimental sensitivity. Therefore, it will become mandatory to assess the magnitude of the penguin contributions in order to properly extract β and thus test the SM.

One possible tool to extract the contribution from penguin topologies is the similar decay channel $B_s^0 \rightarrow J/\psi K_S^0$, which is related to $B^0 \rightarrow J/\psi K_S^0$ through flavour symmetries. Although the decays occur two orders of magnitude less frequently, their experimental signature is identical to the B^0 counterpart. As penguin contributions are not suppressed in $B_s^0 \rightarrow J/\psi K_S^0$, their magnitude can be extracted through a CP violation measurement. Flavour symmetries can then be utilised to relate the penguin contributions in $B_s^0 \rightarrow J/\psi K_S^0$ to $B^0 \rightarrow J/\psi K_S^0$.

This thesis presents the first measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ with the subsequent decays $J/\psi \rightarrow \mu^+ \mu^-$ and $K_S^0 \rightarrow \pi^+ \pi^-$. The measurement is performed on the full dataset of the first run of data taking at LHCb, which corresponds to an integrated luminosity of 3 fb^{-1} at centre-of-mass energies of 7 and 8 TeV. A time-dependent measurement of CP violation in the interference of mixing and decay is performed by measuring the time-dependent decay rates of B_s^0 and \bar{B}_s^0 mesons into the common final state.

The analysis is performed as a collaboration between the LHCb groups at the Technische Universität Dortmund, Germany and at the Nationaal instituut voor subatomaire fysica (NIKHEF), Amsterdam, the Netherlands. It is published as

R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetries in $B_s^0 \rightarrow J/\psi K_S^0$* , JHEP **2015.6**, 131 (2015), DOI: [10.1007/JHEP06\(2015\)131](https://doi.org/10.1007/JHEP06(2015)131), [arXiv:1503.07055 \[hep-ex\]](https://arxiv.org/abs/1503.07055).

Apart from the author's own work, the measurement benefits from contributions of other group members. Furthermore, theses which directly or indirectly contribute to the analysis and which have been supervised by the author are the bachelor theses of Titus Mombächer [14], Vanessa Müller [15], Stefanie Roeser [16], Timon Schmelzer [17], and Tobias Tekampe [18]. Finally, input from other groups of the LHCb collaboration is used in the measurement.

This thesis is structured into four major parts. Chapter 2 covers the theoretical formalisms. It describes how CP violation is incorporated into the SM through the CKM mechanism, presents how the mixing and decay of neutral mesons is parameterised, and discusses different types of CP violation. Furthermore, it is shown how the discussed formalisms apply to the decay channels $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$. The influence of penguin contributions is illustrated and strategies how to determine these are mentioned. Finally, the current status of both experiment and theory is presented.

Chapter 3 describes the LHCb experiment at the LHC. It covers the design of the collider and the LHCb detector including the different sub-detectors. This is followed by a presentation of the data processing in the trigger and the subsequent software stages. A description of the data taking in the first run of LHCb and an outline of the planned detector upgrade conclude this part.

Chapter 4 covers necessary ingredients for the measurement and preparatory studies, which have to be conducted previous to performing the fit to extract the CP parameters. This includes the selection of suitable candidates, studies on the flavour tagging, the determination of the decay-time resolution and acceptance, the handling of the production asymmetry, and details about the branching ratio measurement, which is performed in addition to the CP violation measurement.

Finally, Chapter 5 describes the measurement itself, where the maximum-likelihood fit to extract the CP parameters is described in detail, followed by preliminary results and various studies to validate the fit. The latter include, amongst others, studies with pseudo-experiments, an sFit implementation, and fits on simulated samples. This is followed by the determination of systematic uncertainties. The chapter closes with the presentation of the final results together with confidence intervals obtained with the Feldman–Cousins method.

The thesis concludes with prospects for the measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ with future LHCb data.

2 CP violation in $B \rightarrow J/\psi K_S^0$ decays

The Standard Model of particle physics (SM) describes elementary particles and their interactions through fundamental forces. Using the formalisms establishing particle masses and the unification of interactions, the Cabibbo-Kobayashi-Maskawa (CKM) matrix is introduced, which describes transitions between different quark generations. This also establishes the possibility of CP violation, which amongst others can be studied in decays of ground-state flavoured neutral mesons. The latter can oscillate between particle and antiparticle, which introduces CP violation in the interference of mixing and decay. This is studied in decays of $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$, where the latter channel can be used to determine effects of higher order in the first.

The formalisms and conventions used in this chapter follow the description used in Refs. [19–22].

2.1 Fundamentals of the Standard Model of particle physics

The Standard Model is established as a relativistic Lorentz-invariant quantum field theory. It describes the dynamics of elementary particles and their interactions. Three discrete symmetries are important in the SM. The time reversal T transforms positive into negative times $t \rightarrow -t$, the parity transformation P reflects spatial coordinates at the origin $r \rightarrow -r$, and the charge conjugation C inverts all additive quantum numbers, like the electrical charge. A fundamental assumption for the SM is the CPT theorem, which states that every local Lorentz-invariant field theory is invariant under the combined CPT transformation. However, this does not imply that the individual symmetries are conserved as well. In fact, the SM allows for C , P , and T violation and consequently also for CP violation.

The twelve fermions, which represent the building blocks of all known matter form one part of the fundamental particles in the SM. These are further divided into six quarks and six leptons, where both kinds are arranged in three generations. Applying the CP operation on the fermions transforms these into antiquarks and antileptons, from which antimatter is constructed. The twelve gauge bosons, eight gluons g , the photon γ , the Z^0 boson, and the W^\pm bosons, form the second part and are the force carriers of the different interactions discussed below. Figure 2.1 presents an overview of the fundamental particles and their respective properties.

2 CP violation in $B \rightarrow J/\psi K_S^0$ decays

	1.	2.	3.		
Mass	2.3 MeV/c ²	1275 MeV/c ²	173.2 GeV/c ²	0	126 GeV/c ²
Charge	$+\frac{2}{3}$	$+\frac{2}{3}$	$+\frac{2}{3}$	0	0
Spin	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	0
	u up	c charm	t top	g gluon	H Higgs boson
Quarks					
	4.8 MeV/c ²	95 MeV/c ²	4660 MeV/c ²	0	
	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	d down	s strange	b bottom	γ photon	
Leptons					
	0.511 MeV/c ²	105.7 MeV/c ²	1777 MeV/c ²	91.2 GeV/c ²	
	-1	-1	-1	0	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	e electron	μ muon	τ tau	Z^0 Z boson	
	< 2 eV/c ²	< 0.19 MeV/c ²	< 18.2 MeV/c ²	80.4 GeV/c ²	
	0	0	0	± 1	
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
	ν_e e neutrino	ν_μ μ neutrino	ν_τ τ neutrino	W W boson	
					Gauge bosons

Figure 2.1 – Fundamental particles of the Standard Model and their respective properties. The three generations of quarks and leptons are the fundamental fermions, of which all matter is composed. Fundamental interactions are mediated through the gauge bosons. The Higgs boson is a consequence of the Higgs mechanism. Illustration based on Ref. [23] with properties from Ref. [19].

The fundamental interactions described in the SM are parameterised using the Lagrangian formalism, where the Lagrangian densities are required to be invariant under local gauge transformations. Based on Maxwell’s theory of electromagnetism, the electromagnetic force is described as Quantum Electrodynamics (QED), characterised by local U(1) invariance. This symmetry results in the photon γ as gauge boson, which couples to the electric charge of particles. The weak interaction is unified with the electromagnetic interaction into the electroweak interaction with a formalism involving spontaneous symmetry breaking. This is used to introduce mass terms for the fermions and bosons and results in the massless photon as gauge boson of the QED and the massive gauge bosons Z^0 and W^\pm of the weak interaction. As a consequence, the electromagnetic and weak interaction have different coupling strengths at low momentum scales. Another effect of the broken symmetry is the existence of the Higgs boson H . The resulting gauge symmetry is $SU(2) \times U(1)$ and the interaction couples to the weak isospin and the electric charge. Finally, the strong interaction, called Quantum Chromodynamics (QCD), involves the gauge group $SU(3)$. The carriers are the eight massless gluons g , coupling to the colour charge. Quarks are the only fermions carrying colour and strongly interacting particles can solely appear as hadrons, which are bound states with vanishing total colour charge. Two types of hadrons are well established: mesons consisting of a quark-antiquark pair and baryons being comprised of three quarks or three antiquarks. Recently, LHCb has observed resonances being consistent with a pentaquark bound state [24].

2.2 CKM mechanism

The electroweak unification introduces charged currents describing the interaction of left-handed fermions with the massive W^\pm gauge bosons. In the SM, these currents are the sole origin of CP violation. Yukawa couplings between the Higgs field and both the left-handed and right-handed quark fields are introduced to generate quark masses. These couplings induce transitions between the different quark generations. Furthermore, they feature eigenstates of the electroweak interaction and non-diagonal mass matrices for the up-type and down-type quarks. The latter are diagonalised by unitary matrices which transform the electroweak eigenstates into physically observable mass eigenstates. In the kinetic terms and in the interaction of fermions with neutral gauge bosons, the diagonalisation matrices cancel. However, this is not the case in the charged currents, where the CKM matrix V_{CKM} is introduced as [8, 9]

$$\mathcal{L}_{\text{cc}} = \frac{g}{\sqrt{2}} W_\mu^+ (\bar{u}, \bar{c}, \bar{t})_L \gamma^\mu V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L + \frac{g}{\sqrt{2}} W_\mu^- (\bar{d}, \bar{s}, \bar{b})_L \gamma^\mu V_{\text{CKM}} \begin{pmatrix} u \\ c \\ t \end{pmatrix}_L, \quad (2.1)$$

with the left-handed mass eigenstates of the up-type quarks $(u, c, t)_L$ and down-type quarks $(d, s, b)_L$. The CKM matrix is chosen such that it transforms the down-type mass eigenstates into the respective electroweak eigenstates $(d', s', b')_L$ as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_L = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L. \quad (2.2)$$

Consequently, the CKM matrix element V_{ij} relates the up-type quark i to the down-type quark j between identical and different generations, with the transition probability in a charged-current interaction being proportional to $|V_{ij}|^2$.

The CKM matrix is a unitary matrix, which in general results in nine free parameters to parameterise V_{CKM} . These contain three angles and six complex phases. Five of the latter can be eliminated through the freedom of choosing relative quark phases arbitrarily. Therefore, the CKM matrix is parameterised by the three mixing angles and one complex phase, which introduces CP violation as Equation (2.1) is not CP -invariant in presence of complex CKM elements. A straightforward parameterisation of V_{CKM} is chosen as [25]

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (2.3)$$

with $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$, the mixing angles $\theta_{ij} \in [0, \pi/2]$, and the phase δ . The CKM matrix shows a clear hierarchical structure, which makes it convenient to

2 CP violation in $B \rightarrow J/\psi K_S^0$ decays

choose a parameterisation reflecting this hierarchy. Defining [26, 27]

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \quad s_{23} = A\lambda^2 = \lambda \left| \frac{V_{cb}}{V_{us}} \right|, \quad s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) = V_{ub}^*, \quad (2.4)$$

the CKM matrix in the Wolfenstein parameterisation in terms of λ , A , ρ , and η up to order $\mathcal{O}(\lambda^6)$ becomes

$$\begin{aligned} V_{\text{CKM}} = & \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \\ & + \begin{pmatrix} -1/8\lambda^4 & 0 & 0 \\ 1/2A^2\lambda^5[1 - 2(\rho + i\eta)] & -1/8\lambda^4(1 + 4A^2) & 0 \\ 1/2A\lambda^5(\rho + i\eta) & 1/2A\lambda^4[1 - 2(\rho + i\eta)] & -1/2A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6). \end{aligned} \quad (2.5)$$

The chosen parameterisation satisfies unitarity in all orders in λ . With [28]

$$\lambda = 0.22543 \pm \frac{0.00042}{0.00031}, \quad A = 0.8227 \pm \frac{0.0066}{0.0136}, \quad \rho \approx 0.15, \quad \eta \approx 0.35, \quad (2.6)$$

the hierarchy of the CKM matrix becomes apparent. Furthermore, in this approximation, V_{ub} and V_{td} are the only complex elements up to $\mathcal{O}(\lambda^4)$, making $u \leftrightarrow b$ and $t \leftrightarrow d$ quark transitions especially suited to study CP violation. The matrix element V_{ts} responsible for $t \leftrightarrow s$ transitions is complex in higher orders and relevant for CP violation in B_s^0 decays.

The unitarity relations of the CKM matrix can be expressed as $\sum_i V_{ij}V_{ik}^* = \delta_{jk}$ and $\sum_j V_{ij}V_{kj}^* = \delta_{ik}$. Six of the nine sums vanish, which allows to represent these as triangles in the complex plane. As the CKM matrix contains one irreducible phase, all triangles have the same area. While most of these are nearly degenerate with sides of different order in λ , one that has sides of comparable length and is commonly studied is

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (2.7)$$

By dividing each side by $V_{cd}V_{cb}^*$ the triangle's vertices in the complex plane become $(0, 0)$, $(1, 0)$, and $(\bar{\rho}, \bar{\eta})$, with the apex

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}. \quad (2.8)$$

The side lengths of the triangle are represented by the moduli of the rescaled terms in the unitarity relation. Furthermore, the three angles of the triangle are

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \quad (2.9)$$

Measurements of CP violation involve phases of the CKM elements and are therefore well-suited to gain access to the parameters of the unitarity triangle. For instance, due to the specific CKM matrix elements in Equation (2.7), studies of processes involving decays of B^0 mesons allow to constrain angles and sides of this triangle and therefore the position of its apex. Combining all measurements over-constrains the triangle, which in turn allows to test the SM. Significant discrepancies between measurements of a single parameter with the corresponding indirect prediction from other measurements assuming CKM unitarity would lead to physics beyond the SM. The current experimental situation of the unitarity triangle is shown in Figure 2.2. No evidence of non-SM effects is visible.

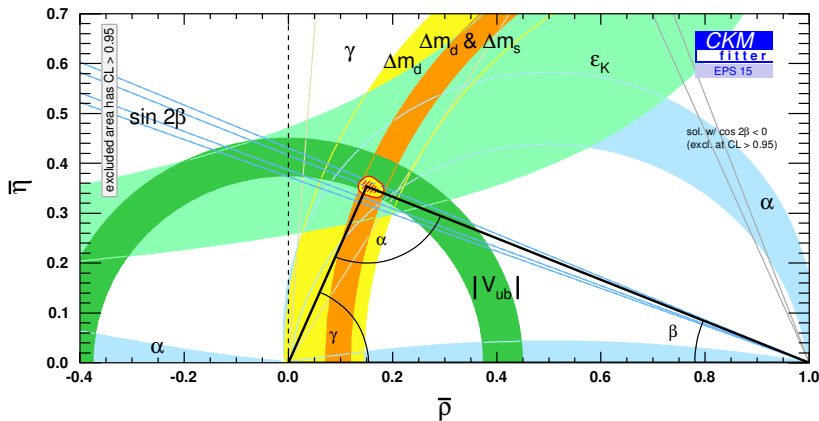


Figure 2.2 – Unitarity triangle in the complex $(\bar{\rho}, \bar{\eta})$ plane [28, 29]. Coloured areas show the experimental constraints on single parameters, with the regions outside having $1 - p > 95.45\%$. The yellow area around the apex is the result of a global fit with $1 - p < 95.45\%$, where the red shaded area represents $1 - p < 68.3\%$.

Another unitarity relation of interest leading to a nearly degenerate triangle is

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0, \quad (2.10)$$

where the corresponding angle β_s is of specific interest. It is accessible through measurements of B_s^0 meson decays and is being defined as

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right). \quad (2.11)$$

As mentioned before, the area of all CKM triangles is identical, being half of the Jarlskog invariant J , which is defined as [30]

$$J = \text{Im } V_{ij}V_{kl}V_{il}^*V_{kj}^* \quad \text{with } i \neq k, j \neq l. \quad (2.12)$$

It can be interpreted as a phase-convention-independent measure of CP violation in the SM, being $J \approx 3 \cdot 10^{-5}$.

2.3 Decay and mixing formalism of neutral mesons

As previously noted, *CP* violation occurs in charged currents of the electroweak interaction. A ground-state flavoured neutral meson M carries non-zero flavour quantum numbers, consisting of a quark and an anti-quark, where both partners originate from different generations. The according *CP* conjugate \bar{M} contains the respective opposite quarks. Therefore, M and \bar{M} cannot decay via the electromagnetic or strong interaction, but instead require charged currents for transitions. The latter can contain states common to both M and \bar{M} . As a consequence, flavoured neutral mesons can mix between particle and antiparticle state and allow for *CP* violation. The following considerations are particularly interesting for B^0 and B_s^0 mesons, but are in principle applicable to all flavoured neutral mesons.

The flavour eigenstates $|M\rangle$ and $|\bar{M}\rangle$ are examined. For instance, these could be $|B_s^0\rangle = |\bar{b}s\rangle$ and $|\bar{B}_s^0\rangle = |b\bar{s}\rangle$. The decay amplitudes to the final state f or its *CP* conjugate \bar{f} are defined as

$$\begin{aligned} A_f &= \langle f|T|M\rangle, & \bar{A}_f &= \langle f|T|\bar{M}\rangle, \\ A_{\bar{f}} &= \langle \bar{f}|T|M\rangle, & \bar{A}_{\bar{f}} &= \langle \bar{f}|T|\bar{M}\rangle, \end{aligned} \quad (2.13)$$

with the Hamiltonian T of the weak interaction. Applying *CP* on initial and final states with the convention $(CP)^2 = 1$ results in

$$\begin{aligned} CP|M\rangle &= e^{+i\xi_M}|\bar{M}\rangle, & CP|f\rangle &= e^{+i\xi_f}|\bar{f}\rangle, \\ CP|\bar{M}\rangle &= e^{-i\xi_M}|M\rangle, & CP|\bar{f}\rangle &= e^{-i\xi_f}|f\rangle, \end{aligned} \quad (2.14)$$

with the so-called spurious phases ξ_M and ξ_f . These are convention-dependent and not observable.

The time-dependent wave function of a superposition of M and \bar{M} is examined,

$$|\Psi(t)\rangle = a(t)|M\rangle + b(t)|\bar{M}\rangle, \quad (2.15)$$

where the values of $a(t)$ and $b(t)$ can be described by the time-dependent Schrödinger-like equation

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{H}\begin{pmatrix} a \\ b \end{pmatrix} = (\mathbf{M} - i/2\mathbf{\Gamma})\begin{pmatrix} a \\ b \end{pmatrix}, \quad (2.16)$$

with the effective non-Hermitian Hamiltonian \mathbf{H} and the Hermitian 2×2 matrices \mathbf{M} and $\mathbf{\Gamma}$. Assuming *CPT* invariance, M and \bar{M} have identical mass m and decay width Γ , resulting in identical diagonal elements of \mathbf{H} . Therefore, \mathbf{H} is identified as

$$\mathbf{H} = \begin{pmatrix} m - i/2\Gamma & M_{12} - i/2\Gamma_{12} \\ M_{12}^* - i/2\Gamma_{12}^* & m - i/2\Gamma \end{pmatrix}. \quad (2.17)$$

2.3 Decay and mixing formalism of neutral mesons

Flavour-changing $M \leftrightarrow \bar{M}$ transitions are described by the off-diagonal elements, where M_{12} represents dispersive short-distance off-shell transitions and Γ_{12} represents absorptive virtual intermediate decays to on-shell final states common to M and \bar{M} .

The eigenvectors of \mathbf{H} are typically chosen as the light and heavy mass eigenstates M_L and M_H ,

$$|M_L\rangle = p|M\rangle + q|\bar{M}\rangle, \quad (2.18a)$$

$$|M_H\rangle = p|M\rangle - q|\bar{M}\rangle, \quad (2.18b)$$

with the complex parameters p and q and the normalization condition $|p|^2 + |q|^2 = 1$. These states have well-defined masses m_L, m_H and decay widths Γ_L, Γ_H . The mass and decay-width differences are defined as

$$\Delta m = m_H - m_L, \quad \Delta\Gamma = \Gamma_L - \Gamma_H, \quad (2.19)$$

where $\Delta\Gamma$ is defined so that $\Delta\Gamma > 0$ for B_s^0 mesons. Consequently, the average mass $m = (m_H + m_L)/2$ and decay width $\Gamma = (\Gamma_L + \Gamma_H)/2$ characterise the meson system. The ratio of mixing parameters q/p then follows as

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^* - i\frac{1}{2}\Gamma_{12}^*}{M_{12} - i\frac{1}{2}\Gamma_{12}}, \quad \frac{q}{p} = -\frac{\Delta m + i\frac{1}{2}\Delta\Gamma}{2(M_{12} - i\frac{1}{2}\Gamma_{12})}. \quad (2.20)$$

The time-dependent evolution of M_L and M_H is given by

$$|M_L(t)\rangle = e^{-im_L t} e^{-\frac{\Gamma_L}{2}t} |M_L\rangle, \quad (2.21a)$$

$$|M_H(t)\rangle = e^{-im_H t} e^{-\frac{\Gamma_H}{2}t} |M_H\rangle. \quad (2.21b)$$

Combining these relations with Equation (2.18), the evolution of states produced as a pure M or \bar{M} at $t = 0$ follows as

$$|M(t)\rangle = g_+(t)|M\rangle - \frac{q}{p}g_-(t)|\bar{M}\rangle, \quad (2.22a)$$

$$|\bar{M}(t)\rangle = g_+(t)|\bar{M}\rangle - \frac{p}{q}g_-(t)|M\rangle, \quad (2.22b)$$

where

$$g_{\pm}(t) = \frac{1}{2} \left(e^{-im_H t} e^{-\frac{\Gamma_H}{2}t} \pm e^{-im_L t} e^{-\frac{\Gamma_L}{2}t} \right). \quad (2.23)$$

Pure flavour eigenstates M and \bar{M} can be produced via the strong or electromagnetic interaction. Their decay into the final state f or its CP conjugate \bar{f} can then be observed. Therefore, the differential time-dependent decay rates are of interest:

$$\begin{aligned} \Gamma(M(t) \rightarrow f) &= |\langle f|T|M(t)\rangle|^2, & \Gamma(\bar{M}(t) \rightarrow f) &= |\langle f|T|\bar{M}(t)\rangle|^2, \\ \Gamma(M(t) \rightarrow \bar{f}) &= |\langle \bar{f}|T|M(t)\rangle|^2, & \Gamma(\bar{M}(t) \rightarrow \bar{f}) &= |\langle \bar{f}|T|\bar{M}(t)\rangle|^2. \end{aligned} \quad (2.24)$$

2 CP violation in $B \rightarrow J/\psi K_S^0$ decays

Introducing the quantities λ_f and $\bar{\lambda}_{\bar{f}}$ as

$$\lambda_f = \frac{q \bar{A}_f}{p A_f}, \quad \bar{\lambda}_{\bar{f}} = \frac{p A_{\bar{f}}}{q \bar{A}_{\bar{f}}}, \quad (2.25)$$

the decay rates can be expressed in terms of phase-convention-independent quantities as

$$\begin{aligned} \Gamma(M(t) \rightarrow f) &= |A_f|^2 \left(|g_+(t)|^2 + |\lambda_f|^2 |g_-(t)|^2 - 2 \operatorname{Re} (\lambda_f g_+^*(t) g_-(t)) \right), \\ \Gamma(\bar{M}(t) \rightarrow f) &= |A_f|^2 \left| \frac{p}{q} \right|^2 \left(|g_-(t)|^2 + |\lambda_f|^2 |g_+(t)|^2 - 2 \operatorname{Re} (\lambda_f g_+(t) g_-^*(t)) \right), \\ \Gamma(M(t) \rightarrow \bar{f}) &= |\bar{A}_{\bar{f}}|^2 \left| \frac{q}{p} \right|^2 \left(|g_-(t)|^2 + |\bar{\lambda}_{\bar{f}}|^2 |g_+(t)|^2 - 2 \operatorname{Re} (\bar{\lambda}_{\bar{f}} g_+(t) g_-^*(t)) \right), \\ \Gamma(\bar{M}(t) \rightarrow \bar{f}) &= |\bar{A}_{\bar{f}}|^2 \left(|g_+(t)|^2 + |\bar{\lambda}_{\bar{f}}|^2 |g_-(t)|^2 - 2 \operatorname{Re} (\bar{\lambda}_{\bar{f}} g_-^*(t) g_-(t)) \right). \end{aligned} \quad (2.26)$$

Decays without net flavour change where an initially produced M (\bar{M}) decays as M (\bar{M}) are represented by terms proportional to $|A|^2$. Furthermore, the terms proportional to $|A|^2 |p/q|^2$ or $|A|^2 |q/p|^2$ result from decays where the meson decays after net oscillation in the state opposite to its production. Finally, the terms proportional to $g_{\pm}^*(t) g_{\mp}(t)$ denote interference of direct decay and decay after mixing.

By expressing $|g_{\pm}(t)|^2$ and $g_{\pm}^*(t) g_{\mp}(t)$ in terms of trigonometric and hyperbolic functions, the decay rates can be rewritten as

$$\Gamma(M(t) \rightarrow f) = \frac{1}{2} |A_f|^2 \left(1 + |\lambda_f|^2 \right) e^{-\Gamma t} \left[\cosh \left(\frac{\Delta\Gamma}{2} t \right) + A_f^{\Delta\Gamma} \sinh \left(\frac{\Delta\Gamma}{2} t \right) + C_f \cos(\Delta m t) - S_f \sin(\Delta m t) \right], \quad (2.27a)$$

$$\Gamma(\bar{M}(t) \rightarrow f) = \frac{1}{2} |A_f|^2 \left(1 + |\lambda_f|^2 \right) \left| \frac{p}{q} \right|^2 e^{-\Gamma t} \left[\cosh \left(\frac{\Delta\Gamma}{2} t \right) + A_f^{\Delta\Gamma} \sinh \left(\frac{\Delta\Gamma}{2} t \right) - C_f \cos(\Delta m t) + S_f \sin(\Delta m t) \right], \quad (2.27b)$$

$$\Gamma(M(t) \rightarrow \bar{f}) = \frac{1}{2} |\bar{A}_{\bar{f}}|^2 \left(1 + |\bar{\lambda}_{\bar{f}}|^2\right) \left| \frac{q}{p} \right|^2 e^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + A_{\bar{f}}^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma}{2}t\right) - C_{\bar{f}} \cos(\Delta mt) + S_{\bar{f}} \sin(\Delta mt) \right], \quad (2.27c)$$

$$\Gamma(\bar{M}(t) \rightarrow \bar{f}) = \frac{1}{2} |\bar{A}_{\bar{f}}|^2 \left(1 + |\bar{\lambda}_{\bar{f}}|^2\right) e^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + A_{\bar{f}}^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma}{2}t\right) + C_{\bar{f}} \cos(\Delta mt) - S_{\bar{f}} \sin(\Delta mt) \right], \quad (2.27d)$$

where the following CP observables are introduced as

$$\begin{aligned} A_f^{\Delta\Gamma} &= -\frac{2 \operatorname{Re} \lambda_f}{1 + |\lambda_f|^2}, & C_f &= \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, & S_f &= \frac{2 \operatorname{Im} \lambda_f}{1 + |\lambda_f|^2}, \\ A_{\bar{f}}^{\Delta\Gamma} &= -\frac{2 \operatorname{Re} \bar{\lambda}_{\bar{f}}}{1 + |\bar{\lambda}_{\bar{f}}|^2}, & C_{\bar{f}} &= \frac{1 - |\bar{\lambda}_{\bar{f}}|^2}{1 + |\bar{\lambda}_{\bar{f}}|^2}, & S_{\bar{f}} &= \frac{2 \operatorname{Im} \bar{\lambda}_{\bar{f}}}{1 + |\bar{\lambda}_{\bar{f}}|^2}. \end{aligned} \quad (2.28)$$

The latter fulfil the relations $(A_f^{\Delta\Gamma})^2 + (C_f)^2 + (S_f)^2 = 1$ and $(A_{\bar{f}}^{\Delta\Gamma})^2 + (C_{\bar{f}})^2 + (S_{\bar{f}})^2 = 1$. Terms proportional to $\cos(\Delta mt)$ and $\cosh(\Delta\Gamma/2t)$ in Equation (2.27) represent decays with and without net oscillation of the mesons, while the terms with $\sin(\Delta mt)$ and $\sinh(\Delta\Gamma/2t)$ result from the interference of direct decay and decay after mixing.

2.4 CP violation in neutral meson decays

Three different types of CP violation can be studied in decays of neutral mesons, which all involve complex phases. As quantum mechanics allows to rephase the states $|M\rangle$ or $|f\rangle$, absolute phases of transitions are arbitrary and therefore not physically meaningful. Only rephasing-invariant quantities can be observed, which typically are relative phases of different contributions to a transition amplitude. Three types of phases are considered. The first being weak phases, which are CP -odd and thus change sign under CP operation. These arise from complex couplings in the Lagrangian, which in the SM become manifest in the phases of the CKM matrix elements. Strong phases are CP -even and thus unaffected by the CP conjugation. For instance, these originate in final-state-interaction scattering from on-shell states through the electromagnetic or strong interaction. Finally, spurious phases as introduced in Equation (2.14) are arbitrary phases between a state and its CP conjugate.

The following examination is restricted to decays of the meson M or \bar{M} into a final state f or \bar{f} . In this case, only the quantities $|A_f|^2$, $|A_{\bar{f}}|^2$, $|\bar{A}_f|^2$, $|\bar{A}_{\bar{f}}|^2$, $|q/p|^2$, λ_f ,

2 CP violation in $B \rightarrow J/\psi K_S^0$ decays

and $\bar{\lambda}_{\bar{f}}$ are invariant with respect to rephasing and thus observable. Three different kinds of CP violation are considered: CP violation in the decay (also referred to as direct CP violation), CP violation in the mixing (indirect CP violation), and CP violation in the interference between mixing and decay.

2.4.1 CP violation in the decay

A requirement for CP violation in the decay is that two or more interfering amplitudes with different weak and different strong phases contribute to the considered decays. For instance, the decay amplitudes could be parameterised as

$$\begin{aligned} A_f &= A_1 e^{i(\delta_1 + \phi_1)} + A_2 e^{i(\delta_2 + \phi_2)}, \\ \bar{A}_{\bar{f}} &= A_1 e^{i(\delta_1 - \phi_1)} + A_2 e^{i(\delta_2 - \phi_2)}, \end{aligned} \quad (2.29)$$

with the two moduli A_1, A_2 , the two strong phases δ_1, δ_2 , and the two weak phases ϕ_1, ϕ_2 . The two amplitudes could, for instance, arise from two different Feynman diagrams of the decay process. As a direct consequence, the moduli $|A_f|$ and $|\bar{A}_{\bar{f}}|$ will be different, thus implying CP violation in the decay:

$$\left| \frac{A_f}{\bar{A}_{\bar{f}}} \right| \neq 1. \quad (2.30)$$

This type of CP violation can be observed in decays of B^0 and B_s^0 mesons. It is furthermore the only type of CP violation which can occur in decays of charged mesons.

2.4.2 CP violation in the mixing

CP violation in the mixing of neutral mesons occurs if the mixing rates between particle and antiparticle differ, $\Gamma(M \rightarrow \bar{M}) \neq \Gamma(\bar{M} \rightarrow M)$. Following from Equations (2.15) and (2.17), this is the case if M_{12} and Γ_{12} have different phases, for example if

$$M_{12} = |M_{12}| e^{i\phi_M}, \quad \Gamma_{12} = |\Gamma_{12}| e^{i\phi_\Gamma}. \quad (2.31)$$

A non-vanishing phase difference $\phi = \phi_M - \phi_\Gamma \neq 0$ and a non-vanishing ratio of moduli $|\Gamma_{12}/M_{12}| \neq 0$ directly results in

$$\left| \frac{q}{p} \right| \neq 1. \quad (2.32)$$

This type of CP violation can be studied with semileptonic decays, where the final states are exclusively accessible by either M or \bar{M} , e.g. $M \rightarrow l^+ X$ and $\bar{M} \rightarrow l^- X$.

In both the B^0 and B_s^0 system, measurements show that CP violation in the mixing is negligible [31, 32]. This is in accordance with SM expectations as $|\Gamma_{12}| \ll |M_{12}|$ is assumed due to suppressed on-shell transitions in mixing diagrams.

2.4.3 CP violation in the interference of mixing and decay

The third kind of CP violation occurs in the interference of the direct decay $M \rightarrow f$ and the decay after mixing $M \rightarrow \bar{M} \rightarrow f$. Interference CP violation can appear in decays where no direct or indirect CP violation is present. For simplicity, it is assumed that the decay amplitude can be parameterised as $A_f = Ae^{i(\delta_D + \phi_D)}$ with the strong phase δ_D and the weak phase ϕ_D . Furthermore, CP violation in the mixing is assumed to be negligible due to $|\Gamma_{12}/M_{12}| \approx 0$, so that $q/p = e^{-i\phi_M}$ is a pure phase. This results in $|\lambda_f| = 1$, while the phase difference between mixing and decay allows for

$$\text{Im } \lambda_f \neq 0. \quad (2.33)$$

In the case of the decay into a single CP eigenstate f , into which both M and \bar{M} can decay, the time-dependent asymmetry $\mathcal{A}(t)$ is considered, which is defined as

$$\begin{aligned} \mathcal{A}(t) &= \frac{\Gamma(\bar{M}(t) \rightarrow f) - \Gamma(M(t) \rightarrow f)}{\Gamma(\bar{M}(t) \rightarrow f) + \Gamma(M(t) \rightarrow f)} \\ &= \frac{S_f \sin(\Delta mt) - C_f \cos(\Delta mt)}{\cosh\left(\frac{\Delta\Gamma}{2}t\right) + A_f^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma}{2}t\right)}. \end{aligned} \quad (2.34)$$

If direct and indirect CP violation are negligible, C_f follows as $C_f = 0$. In that case, the resulting parameter of interest is $S_f = \eta_{CP} \sin(\phi_M + 2\phi_D)$ with the CP eigenvalue $\eta_{CP} = \pm 1$ of f . The presence of additional direct or indirect CP violation results in $|\lambda_f| \neq 1$ and thus in $C_f \neq 0$ as well as $A_f^{\Delta\Gamma} \neq \pm\sqrt{1 - S_f^2}$.

2.5 The decays $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$

Two decay channels are relevant for this thesis. The first is $B^0 \rightarrow J/\psi K_S^0$, which is often referred to as the gold-plated channel to measure CP violation in the B^0 system and as such is a key channel to test the quark-flavour sector of the SM. With $J/\psi K_S^0$ as final state accessible to both B^0 and \bar{B}^0 , CP violation in the interference of mixing and decay can be studied. In the SM, the decays are dominated by a single nearly clean weak phase difference, which can be measured without the necessity of external input and which allows to determine the CKM angle β . From an experimental point of view, $B^0 \rightarrow J/\psi K_S^0$ is advantageous with charged muons and pions in the final state, which are convenient to detect. However, the formalism usually neglects contributions from sub-leading penguin diagrams, which can in principle bias the determination of β . While this approximation is justified for the experimental precision achieved so far at the B factories BABAR and Belle, the so-called penguin pollution needs to be accounted for with the sensitivity accessible to LHCb and Belle II. Furthermore, a possible contribution of physics

beyond the SM, which in $B^0 \rightarrow J/\psi K_S^0$ is inevitably small, can only be determined if the penguin contribution is precisely measured.

Therefore, the second channel of interest is $B_s^0 \rightarrow J/\psi K_S^0$, which is studied in detail in this thesis. It is topologically similar to $B^0 \rightarrow J/\psi K_S^0$, while the tree diagram is suppressed so that it is of similar order as the penguin diagram. A measurement of CP violation thus allows to determine the magnitude of the penguin pollution by relating the result to the tree-only expectation. In general, the branching ratio of $B_s^0 \rightarrow J/\psi K_S^0$ is smaller than its B^0 counterpart, which imposes different experimental challenges compared to a measurement of CP violation in $B^0 \rightarrow J/\psi K_S^0$. The strategy is to determine the magnitude of the penguin contribution in $B_s^0 \rightarrow J/\psi K_S^0$ and then relate it to $B^0 \rightarrow J/\psi K_S^0$ to control the impact of the penguin contribution on the determination of β .

To achieve this, the quantity λ_f (see Equation (2.25)), which determines the CP observables, needs to be developed for both channels by discussing mixing and decays. This shows the similarities and differences of the channels and points to strategies to determine the penguin parameters contributing to $B^0 \rightarrow J/\psi K_S^0$. Finally, the current experimental and theoretical status is discussed.

2.5.1 Mixing in the B^0 - \bar{B}^0 and B_s^0 - \bar{B}_s^0 meson systems

The formalism discussed in Section 2.3 will be applied to B_q mesons, where q denotes either d or s and thus reflects a B^0 or B_s^0 meson. For both mesons it can be shown that $|\Gamma_{12}| \ll |M_{12}|$, so that Equation (2.20) can be simplified to

$$\left(\frac{q}{p}\right)^2 \approx \frac{M_{12}^*}{M_{12}} = \exp(2i \arg(M_{12}^*)). \quad (2.35)$$

Consequently, q/p is a pure phase with $|q/p| \approx 1$, which allows to neglect CP violation in the mixing. There are different arguments justifying $|\Gamma_{12}| \ll |M_{12}|$. For instance, Γ_{12} is given as the sum over all final states f accessible to B_q and \bar{B}_q ,

$$\Gamma_{12} = \sum_f \langle f | T | B_q \rangle^* \langle f | T | \bar{B}_q \rangle. \quad (2.36)$$

In the SM, the dominant contribution to the absorptive transitions comes from tree-level diagrams. Their branching ratios are small as these are either Cabibbo-suppressed decays for both B_q and \bar{B}_q or decays that are Cabibbo-favoured for one meson, but doubly Cabibbo-suppressed for the other meson. Thus Γ_{12} is small compared to the overall decay width, $\Gamma_{12} \ll \Gamma$. As M_{12} can be related to Δm as $M_{12} \approx \Delta m/2$ and $x_d = \Delta m_d/\Gamma_d \approx 0.77$ (for B^0) and $x_s = \Delta m_s/\Gamma_s \approx 27$ (for B_s^0) are measured [19], the claim is justified. Another argument shows that $|\Gamma_{12}/M_{12}| \propto m_b^2/m_t^2 \approx 0$ with the b and t quark masses m_b and m_t [19].

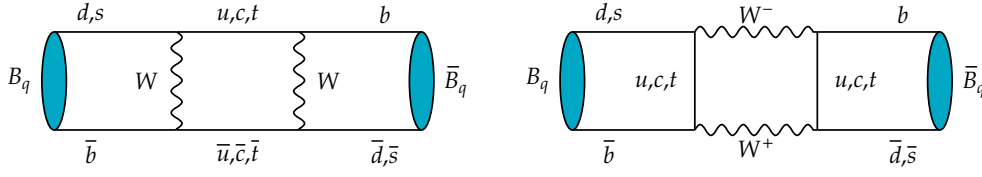


Figure 2.3 – Dominant box diagram contributions to the B_q - \bar{B}_q mixing.

The dominant mixing diagrams contributing to M_{12} are shown in Figure 2.3. In the SM, long-distance contributions to M_{12} are negligible as the B_q mass is large compared to relevant hadronic resonances. Omitting spurious phases, M_{12} is given by

$$M_{12} = -\frac{G_F^2 m_W^2}{12\pi} m_{B_q} f_{B_q}^2 B_{B_q} \left[\eta_1 (V_{cq}^* V_{cb})^2 S_0 \left(\frac{m_c^2}{m_W^2} \right) + \eta_2 (V_{tq}^* V_{tb})^2 S_0 \left(\frac{m_t^2}{m_W^2} \right) + 2\eta_3 (V_{cq}^* V_{cb}) (V_{tq}^* V_{tb}) S_0 \left(\frac{m_c^2}{m_W^2}, \frac{m_t^2}{m_W^2} \right) \right], \quad (2.37)$$

where G_F is the Fermi constant, m_W the W boson mass, m_{B_q} the B_q meson mass, and m_q the mass of quark q . The weak decay constant f_{B_q} and the bag parameter B_{B_q} describe corrections from non-perturbative QCD related to the transition from bound to free quarks. Finally, η_i are perturbative QCD corrections and S_0 are the Inami-Lim functions, which depend on the ratio of quark masses to the W boson mass. The latter show the clear hierarchy

$$S_0 \left(\frac{m_t^2}{m_W^2} \right) \gg S_0 \left(\frac{m_c^2}{m_W^2}, \frac{m_t^2}{m_W^2} \right) \gg S_0 \left(\frac{m_c^2}{m_W^2} \right), \quad (2.38)$$

which allows to simplify M_{12} to

$$M_{12} = -\frac{G_F^2 m_W^2}{12\pi} m_{B_q} f_{B_q}^2 B_{B_q} \eta_2 (V_{tq}^* V_{tb})^2 S_0 \left(\frac{m_t^2}{m_W^2} \right). \quad (2.39)$$

For the B_q - \bar{B}_q meson system the mixing parameter q/p then follows as

$$\frac{q}{p} = \frac{V_{tb}^* V_{tq}}{V_{tq}^* V_{tb}}. \quad (2.40)$$

2.5.2 Decay amplitudes and CP violation in $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$

The relevant decay amplitudes for the $B_q \rightarrow J/\psi K_S^0$ decays are given as

$$\bar{A}_f = \langle J/\psi K_S^0 | T | \bar{B}_q \rangle, \quad A_f = \langle J/\psi K_S^0 | T | B_q \rangle, \quad (2.41)$$

2 CP violation in $B \rightarrow J/\psi K_S^0$ decays

with the transition matrix T . In the $K^0-\bar{K}^0$ system, a similar formalism for the mixing as in Section 2.3 is necessary. The K_S^0 meson as mass eigenstate is parameterised as a superposition of the flavour eigenstates K^0 and \bar{K}^0 ,

$$|K_S^0\rangle = p_K |K^0\rangle - q_K |\bar{K}^0\rangle. \quad (2.42)$$

Performing the according calculation as in Section 2.5.1, the ratio of the mixing parameters then follows as

$$\frac{q_K}{p_K} = -\frac{V_{cs}^* V_{cd}}{V_{cs} V_{cd}^*}, \quad (2.43)$$

again neglecting CP violation in the mixing. A B^0 (\bar{B}^0) meson can only decay into the flavour eigenstate K^0 (\bar{K}^0), while contrarily a B_s^0 (\bar{B}_s^0) meson can only decay into a \bar{K}^0 (K^0) meson. Therefore, the decay amplitudes for \bar{B}^0 and B^0 follow as

$$\bar{A}_f = -\frac{1}{2q_K} \langle J/\psi \bar{K}^0 | T | \bar{B}^0 \rangle, \quad A_f = \frac{1}{2p_K} \langle J/\psi K^0 | T | B^0 \rangle, \quad (2.44)$$

while the respective amplitudes for \bar{B}_s^0 and B_s^0 are

$$\bar{A}_f = \frac{1}{2p_K} \langle J/\psi K^0 | T | \bar{B}_s^0 \rangle, \quad A_f = -\frac{1}{2q_K} \langle J/\psi \bar{K}^0 | T | B_s^0 \rangle. \quad (2.45)$$

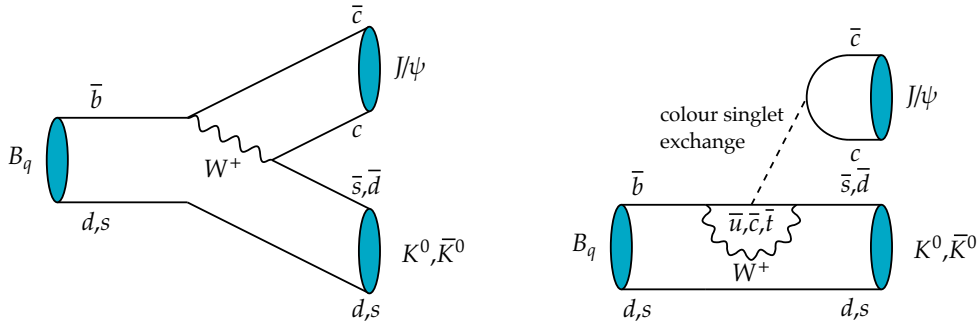


Figure 2.4 – Decay topologies of $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ decays. The left diagram depicts the leading tree topology, while the right diagram shows the main penguin topology.

The leading Feynman diagrams are shown in Figure 2.4. Both tree and penguin topologies contribute to the decays. In case of $B^0 \rightarrow J/\psi K_S^0$ decays, the penguin contribution can often be neglected compared to the singly Cabibbo-suppressed tree topology, which will be shown below. For $B_s^0 \rightarrow J/\psi K_S^0$ decays, the tree topology is doubly Cabibbo-suppressed, which relatively enhances the penguin contribution in this decay channel.

In case penguin contributions are neglected, the decay amplitudes for both decay channels follow as

$$A(B^0 \rightarrow J/\psi K^0) = T_c V_{cs} V_{cb}^*, \quad A(B_s^0 \rightarrow J/\psi \bar{K}^0) = T'_c V_{cd} V_{cb}^*, \quad (2.46)$$

with the CP -conserving amplitudes T_c and T'_c of the tree process. In order to calculate λ_f , the CP eigenvalue of the $J/\psi K_S^0$ final state, $\eta_{J/\psi K_S^0} = -1$ is required, which follows from the necessity of an angular momentum of $l = 1$ between the CP -even J/ψ and the almost CP -even K_S^0 . For $B^0 \rightarrow J/\psi K_S^0$, λ_f then follows as

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \left(-\frac{V_{cd}^* V_{cs}}{V_{cd} V_{cs}^*} \right) \frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*} = -\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \frac{V_{cd}^* V_{cb}}{V_{cd} V_{cb}^*} = -e^{-2i\beta} = -e^{-i\phi_d}, \quad (2.47)$$

with the CKM angle β and the $B^0-\bar{B}^0$ mixing phase $\phi_d = 2\beta$. Analogously, for $B_s^0 \rightarrow J/\psi K_S^0$ λ_f is determined as

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f} = \frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*} \left(-\frac{V_{cd}^* V_{cs}}{V_{cd} V_{cs}^*} \right) \frac{V_{cd}^* V_{cb}}{V_{cd} V_{cb}^*} = -\frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*} \frac{V_{cd}^* V_{cb}}{V_{cd} V_{cb}^*} = -e^{2i\beta_s} = -e^{-i\phi_s}, \quad (2.48)$$

with the angle β_s and the $B_s^0-\bar{B}_s^0$ mixing phase $\phi_s = -2\beta_s$. Therefore, in the absence of penguin contributions, the CP observables follow as

$$S_{B^0 \rightarrow J/\psi K_S^0} = \sin(\phi_d), \quad C_{B^0 \rightarrow J/\psi K_S^0} = 0, \quad A_{B^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} = \cos(\phi_d), \quad (2.49)$$

$$S_{B_s^0 \rightarrow J/\psi K_S^0} = \sin(\phi_s), \quad C_{B_s^0 \rightarrow J/\psi K_S^0} = 0, \quad A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} = \cos(\phi_s). \quad (2.50)$$

Under these circumstances, $B^0 \rightarrow J/\psi K_S^0$ serves as a clean decay channel to extract the CKM angle β . As will be shown further below, the approximation to neglect the penguin topologies is justified up to a high precision. Similarly, $B_s^0 \rightarrow J/\psi K_S^0$ would serve as a channel to extract the $B_s^0-\bar{B}_s^0$ mixing phase ϕ_s . However, even in case the penguin contribution was negligible, other channels like $B_s^0 \rightarrow J/\psi K^+ K^-$ and $B_s^0 \rightarrow J/\psi \pi^+ \pi^-$ with considerably larger branching ratios are better suited for this purpose.

On the contrary, the asset of $B_s^0 \rightarrow J/\psi K_S^0$ is that it allows to determine the penguin contribution in $B^0 \rightarrow J/\psi K_S^0$, as will be shown in the following. Taking penguin topologies into account and utilising CKM unitarity, the decay amplitudes become

$$\begin{aligned} A(B^0 \rightarrow J/\psi K^0) &= T_c V_{cs} V_{cb}^* + P_u V_{us} V_{ub}^* + P_c V_{cs} V_{cb}^* + P_t V_{ts} V_{tb}^* \\ &= V_{cs} V_{cb}^* (T_c + P_c - P_t) + V_{us} V_{ub}^* (P_u - P_t), \end{aligned} \quad (2.51a)$$

$$\begin{aligned} A(B_s^0 \rightarrow J/\psi \bar{K}^0) &= T'_c V_{cd} V_{cb}^* + P'_u V_{ud} V_{ub}^* + P'_c V_{cd} V_{cb}^* + P'_t V_{td} V_{tb}^* \\ &= V_{cd} V_{cb}^* (T'_c + P'_c - P'_t) + V_{ud} V_{ub}^* (P'_u - P'_t), \end{aligned} \quad (2.51b)$$

introducing the CP -conserving penguin amplitudes P_q and P'_q for the loop with quark q . Using the Wolfenstein parameterisation, the decay amplitudes are trans-

2 CP violation in $B \rightarrow J/\psi K_S^0$ decays

formed into

$$A(B^0 \rightarrow J/\psi K^0) = \left(1 - \frac{\lambda^2}{2}\right) \mathcal{A} (1 + \epsilon a e^{i\theta} e^{i\gamma}), \quad (2.52a)$$

$$A(B_s^0 \rightarrow J/\psi \bar{K}^0) = -\lambda \mathcal{A}' (1 - a' e^{i\theta'} e^{i\gamma}), \quad (2.52b)$$

with

$$\mathcal{A} = A\lambda^2 (T_c + P_c - P_t), \quad a e^{i\theta} = R_u \frac{P_u - P_t}{T_c + P_c - P_t}, \quad \epsilon = \frac{\lambda^2}{1 - \lambda^2}, \quad (2.53a)$$

$$\mathcal{A}' = A\lambda^2 (T'_c + P'_c - P'_t), \quad a' e^{i\theta'} = R_u \frac{P'_u - P'_t}{T'_c + P'_c - P'_t}, \quad (2.53b)$$

and $R_u = |V_{ud}V_{ub}^*/V_{cd}V_{cb}^*|$, one of the sides of the unitarity triangle. The terms $a e^{i\theta}$ and $a' e^{i\theta'}$ are CP-conserving with the strong phases θ and θ' , while γ is one of the CKM angles and a weak phase, allowing for direct CP violation. The factor $\epsilon \approx 0.05$ only appears for $B^0 \rightarrow J/\psi K_S^0$ and represents the relative suppression of penguin contributions with respect to the tree topology in this decay channel [28].

Analogously to Equations (2.47) and (2.48), λ_f in $B^0 \rightarrow J/\psi K_S^0$ follows as

$$\lambda_f = -e^{-i\phi_d} \frac{1 + \epsilon a e^{i\theta} e^{-i\gamma}}{1 + \epsilon a e^{i\theta} e^{+i\gamma}}, \quad (2.54)$$

while in $B_s^0 \rightarrow J/\psi K_S^0$ it can be expressed as

$$\lambda_f = -e^{-i\phi_s} \frac{1 - a' e^{i\theta'} e^{-i\gamma}}{1 - a' e^{i\theta'} e^{+i\gamma}}. \quad (2.55)$$

Taking the additional terms into account to evaluate the CP observables leads to complex expressions. A more convenient term can be found by introducing the experimentally observed effective phases ϕ_d^{eff} and ϕ_s^{eff} and penguin-induced phase shifts $\Delta\phi_d$ and $\Delta\phi_s$ for which the following expressions are derived [33–36]:

$$\frac{S_{B^0 \rightarrow J/\psi K_S^0}}{\sqrt{1 - C_{B^0 \rightarrow J/\psi K_S^0}^2}} = \sin(\phi_d^{\text{eff}}) = \sin(\phi_d + \Delta\phi_d), \quad (2.56)$$

$$\frac{S_{B_s^0 \rightarrow J/\psi K_S^0}}{\sqrt{1 - C_{B_s^0 \rightarrow J/\psi K_S^0}^2}} = \sin(\phi_s^{\text{eff}}) = \sin(\phi_s + \Delta\phi_s), \quad (2.57)$$

with

$$\tan(\Delta\phi_d) = \frac{2\epsilon a \cos(\theta) \sin(\gamma) + \epsilon^2 a^2 \sin(2\gamma)}{1 + 2\epsilon a \cos(\theta) \cos(\gamma) + \epsilon^2 a^2 \cos(2\gamma)}, \quad (2.58)$$

$$\tan(\Delta\phi_s) = \frac{-2a' \cos(\theta') \sin(\gamma) + a'^2 \sin(2\gamma)}{1 - 2a' \cos(\theta') \cos(\gamma) + a'^2 \cos(2\gamma)}. \quad (2.59)$$

In $B^0 \rightarrow J/\psi K_S^0$, the phase shift $\Delta\phi_d$ is expected to be small due to the suppression with ϵ . This justifies to neglect penguin contributions with the experimental precision achieved in measurements of $B^0 \rightarrow J/\psi K_S^0$ decays so far. However, in the future a careful determination of the penguin contributions becomes mandatory to thoroughly test the SM.

The hadronic parameters a , θ , a' , and θ' cannot reliably be calculated using perturbative QCD and thus suffer from large theoretical uncertainties. However, using ϕ_s and γ from other measurements as external input, a measurement of $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$, $C_{B_s^0 \rightarrow J/\psi K_S^0}$, and $S_{B_s^0 \rightarrow J/\psi K_S^0}$ with $B_s^0 \rightarrow J/\psi K_S^0$ decays allows to determine a' and θ' .

The approach to subsequently determine $\Delta\phi_d$ is to relate a' and θ' in $B_s^0 \rightarrow J/\psi K_S^0$ to a and θ in $B^0 \rightarrow J/\psi K_S^0$. A key feature of both decay channels is that one can be transformed into the other by interchanging all s and d quarks, which is called U-spin flavour symmetry, where U-spin is a subgroup of the flavour SU(3). This symmetry implies [33, 35, 36]

$$a = a', \quad \theta = \theta', \quad \mathcal{A} = \mathcal{A}'. \quad (2.60)$$

In this case, a measurement of time-dependent CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ allows to directly determine the penguin-induced phase shift in $B^0 \rightarrow J/\psi K_S^0$.

However, U-spin symmetry is broken and different strategies exist on how to handle this drawback. Refs. [33, 35, 36] introduce additional parameters accounting for U-spin-breaking effects with associated uncertainties. Ref. [37] proposes a full SU(3) analysis investigating up to six decay channels simultaneously. Finally, Ref. [38] suggests to avoid relying on SU(3) symmetry altogether and uses an operator product expansion approach for the calculation of the penguin contribution.

2.5.3 Current status in experiment and theory

As stated before, $B^0 \rightarrow J/\psi K_S^0$ is the gold-plated decay channel to measure the CKM angle β . It has a sufficiently large branching ratio of $\mathcal{B}(B^0 \rightarrow J/\psi K_S^0) \approx 4.4 \cdot 10^{-4}$ [19] and the mass difference amounts to $\Delta m_d = (0.510 \pm 0.003) \text{ ps}^{-1}$ [39]. Furthermore, the decay width difference $\Delta\Gamma_d$ is usually neglected, as $\Delta\Gamma_d/\Gamma_d = \pm(0.001 \pm 0.010)$ [39]. Thus, the time-dependent asymmetry $\mathcal{A}(t)$ simplifies to

$$\mathcal{A}(t) = S_{B^0 \rightarrow J/\psi K_S^0} \sin(\Delta m t) - C_{B^0 \rightarrow J/\psi K_S^0} \cos(\Delta m t). \quad (2.61)$$

The current world averages for $S_{B^0 \rightarrow J/\psi K_S^0}$ and $C_{B^0 \rightarrow J/\psi K_S^0}$ are [39]

$$S_{B^0 \rightarrow J/\psi K_S^0}^{\text{WA}} = 0.665 \pm 0.024, \quad C_{B^0 \rightarrow J/\psi K_S^0}^{\text{WA}} = 0.024 \pm 0.026,$$

where the latest LHCb measurement [12]

$$\begin{aligned} S_{B^0 \rightarrow J/\psi K_S^0}^{\text{LHCb}} &= 0.731 \pm 0.035 (\text{stat}) \pm 0.020 (\text{syst}), \\ C_{B^0 \rightarrow J/\psi K_S^0}^{\text{LHCb}} &= -0.038 \pm 0.032 (\text{stat}) \pm 0.005 (\text{syst}), \end{aligned}$$

2 CP violation in $B \rightarrow J/\psi K_S^0$ decays

is not yet included. Consequently, the precision of the world average translates into an uncertainty of $\approx 2^\circ$ on ϕ_d .

There are no previous measurements of the CP observables in $B_s^0 \rightarrow J/\psi K_S^0$ available to extract $\Delta\phi_d$. However, there are estimations using theoretical arguments and data from other decay channels. Ref. [36] estimates $\Delta\phi_d = (-1.1_{-0.85}^{+0.70})^\circ$, while Ref. [38] argues $\Delta\phi_d < 0.68^\circ$ claiming to utilise a conservative approach. Finally, Ref. [37] estimates $|\Delta S_{B^0 \rightarrow J/\psi K_S^0}| \lesssim 0.01$, translating into $|\Delta\phi_d| \lesssim 0.8^\circ$. In conclusion, for the forthcoming precision accessible with future LHCb and Belle II data, the penguin-induced phase shift $\Delta\phi_d$ should be determined using experimental data from $B_s^0 \rightarrow J/\psi K_S^0$ to improve these estimations and thoroughly test the SM.

Experimentally, the $B_s^0 \rightarrow J/\psi K_S^0$ decay channel introduces additional complications compared to its B^0 counterpart. The branching ratio of $\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0) = (1.87 \pm 0.17) \cdot 10^{-5}$ [19] is considerably smaller, requiring a very stringent selection to isolate the small signal from background. Furthermore, the mass difference amounts to $\Delta m_s = (17.757 \pm 0.021) \text{ ps}^{-1}$ [39], which results in the $B_s^0 - \bar{B}_s^0$ oscillation to occur at a ~ 35 times higher frequency compared to $B^0 - \bar{B}^0$ oscillations. This fact makes high demands on the decay-time resolution in order to extract the oscillation-dependent CP observables with sufficient precision. Finally, the decay width difference is non-vanishing with $\Delta\Gamma_s = (0.081 \pm 0.006) \text{ ps}^{-1}$ [39].

Although there are no experimental results on the CP observables available, Ref. [36] gives a prediction for $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$, $C_{B_s^0 \rightarrow J/\psi K_S^0}$, and $S_{B_s^0 \rightarrow J/\psi K_S^0}$ using experimental input from other decay channels:

$$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} = 0.957 \pm 0.061, \quad C_{B_s^0 \rightarrow J/\psi K_S^0} = -0.003 \pm 0.021, \quad S_{B_s^0 \rightarrow J/\psi K_S^0} = -0.29 \pm 0.20.$$

This prediction can be compared to the results of a measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ decays.

3 The LHCb experiment

The Large Hadron Collider beauty (LHCb) experiment is being operated at the LHC at the European Organization for Nuclear Research (CERN) near Geneva, Switzerland. Its main focus are precision measurements of decays of hadrons containing b or c quarks, which places high demands on the detector. A large sample of b or c hadron decays is required with an excellent resolution of the decay vertices to properly measure the decay time of particles. Furthermore, to identify various final states, the particle identification (PID) system needs to perform equally well. The subsequent data processing involves selecting and reconstructing the small fraction of decays of interest from the vast amount of pp collisions.

3.1 The Large Hadron Collider

With 26.7 km circumference and a centre-of-mass energy of up to 14 TeV, the LHC is the world's largest particle accelerator and collider, reaching unprecedented energies. Beside smaller experiments, the four large experiments ATLAS, ALICE, CMS, and LHCb cover a wide range of particle physics topics.

3.1.1 Accelerator and collider

The LHC accelerates two beams of protons in opposite directions in a circular collider of 26.7 km circumference [40]. It is located underground in a tunnel between 50 m and 175 m below the surface near Geneva on both French and Swiss territory and is being operated by CERN. It has been constructed in the same tunnel that was previously used for the Large Electron–Positron Collider (LEP). Protons are accelerated in up to 2808 bunches per beam consisting of up to $1.15 \cdot 10^{11}$ protons per bunch. Both beams traverse the accelerator in individual beam pipes and are collided at four points of the accelerator, where the large experiments are situated. This translates into a minimum spacing of 25 ns between two bunch crossings or a collision rate of up to 40 MHz. The maximum centre-of-mass energy \sqrt{s} of the LHC is 14 TeV. Furthermore, the collider is designed to provide an instantaneous luminosity of up to $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. In 2011 and 2012, when the analysed data sample was taken, the LHC was operated at a minimum bunch spacing of 50 ns.

3 The LHCb experiment

Previous to being injected into the LHC, the proton bunches have to be pre-accelerated by a chain of different accelerators. The protons pass through—in that order—the Linear Accelerator 2 (LINAC2), the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS), where a beam energy of 450 GeV is reached. Afterwards, the bunches are injected into the LHC, where these are accelerated to the nominal energy. The LHC consists of eight arcs, in which the beams are bend into a circular path, and eight straight sections between the arcs. In the latter, the experiments or facilities for beam acceleration, beam cleaning, and beam dump are located. The arcs in total contain 1232 superconducting NbTi dipole magnets, which are cooled by large amounts of liquid ^4He to an operating temperature of below 1.9 K and which provide more than 8 T nominal field strength at $\sqrt{s} = 14 \text{ TeV}$. This places high demands on the magnet quench protection systems. Additionally, there are support magnets such as quadrupoles. Figure 3.1 gives an overview of the accelerator complex at CERN.

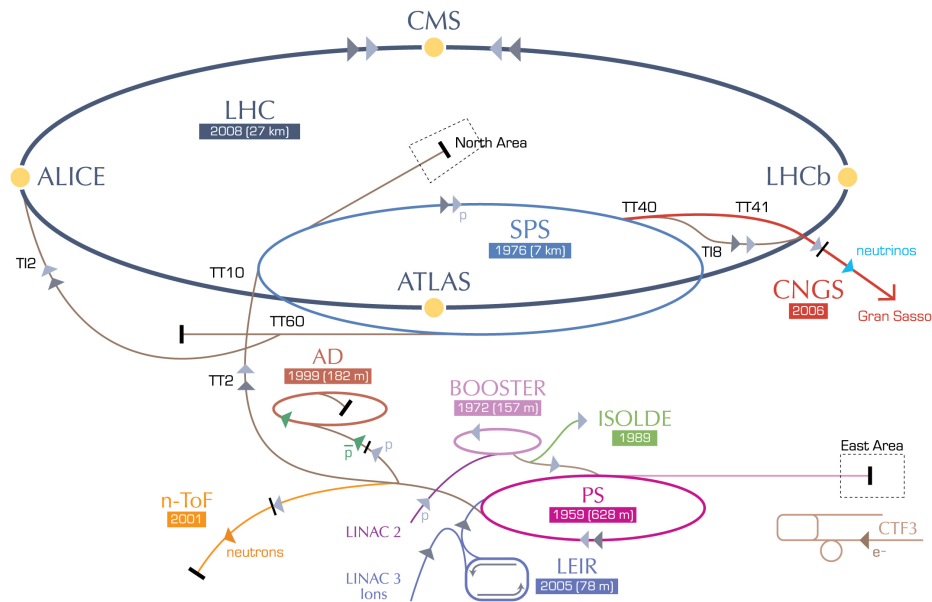


Figure 3.1 – Overview of the CERN accelerator complex [41]. Protons for the LHC are accelerated through the Linear Accelerator 2 (LINAC2), the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS) before these enter the LHC and are accelerated to the nominal energy. Lead ions do not use LINAC2 and BOOSTER, but the Linear Accelerator 3 (LINAC3) and Low Energy Ion Ring (LEIR) instead.

Apart from being operated with proton beams, the LHC can as well accelerate lead ions. This is especially important for the ALICE experiment and not relevant for this analysis. Therefore, this mode of operation will not be covered further.

3.1.2 LHC experiments

Apart from the LHCb experiment, there are six further experiments being operated at the LHC. The two large general-purpose detectors (GPDs) ATLAS [42] and CMS [43] cover a wide area of particle physics research, where both physics programmes overlap to allow for cross-checking each other's results. Both experiments are most notably known for the discovery of the Higgs Boson [44], but cover other areas of research as well. These involve the search for super-symmetric particles, investigation of CP violation, physics of the top quark, as well as searches for dark matter candidates. ALICE [45] mainly investigates collisions of lead ions with other lead ions. Its main focus is the study of the quark-gluon plasma, which is produced in the ion-ion collisions at extreme energies.

The smaller LHC experiments are TOTEM [46], Large Hadron Collider forward (LHCf) [47], and MoEDAL [48]. TOTEM aims at measuring the pp interaction cross-section and is installed in the forward region near CMS. LHCf is also installed in the forward region, near ATLAS, and investigates neutral pion decays to gain a better understanding of cosmic rays. Finally, MoEDAL is located near LHCb and searches for magnetic monopoles and other exotic particles.

3.2 The LHCb detector

The LHCb detector observes decays of hadrons containing b or c quarks. It is composed of several sub-detectors, which either aim at tracking the particles being produced in the pp interaction or at identifying these, i.e. distinguishing between the different types of final-state particles. The layout of the detector is driven by the production mechanism of b or c hadrons in pp collisions at LHC energies. In the following, b quarks are discussed, although the situation for c quarks is similar.

At the LHC, b quarks are predominantly produced as $b\bar{b}$ quark pairs. In leading order, these are produced in either gluon-gluon fusion or quark-antiquark annihilation [49]. The production cross-section compared to other heavy particles is relatively high, so that in approximately 3 % to 5 % of all pp interactions a $b\bar{b}$ pair is produced. A high momentum asymmetry of the proton constituents is likely at the LHC energies. This results in a large boost in the direction of one of the proton beams for the $b\bar{b}$ pair. Thus, both b quarks are predominantly produced in either the forward or the backward direction with small polar angles θ with respect to the beam axis. Figure 3.2 shows the according distributions of $b\bar{b}$ pairs being produced at the LHC. Each quark of the pair hadronises independently into b hadrons, where $\approx 34\%$ form a B^+/B^- meson, $\approx 34\%$ form a B^0/\bar{B}^0 meson, $\approx 11\%$ form a B_s^0/\bar{B}_s^0 meson, and $\approx 21\%$ form a b baryon or heavier mesons [19].

Because b hadrons are predominantly produced at small angles θ and with large boosts, the LHCb detector is designed as a single-arm forward spectrometer,

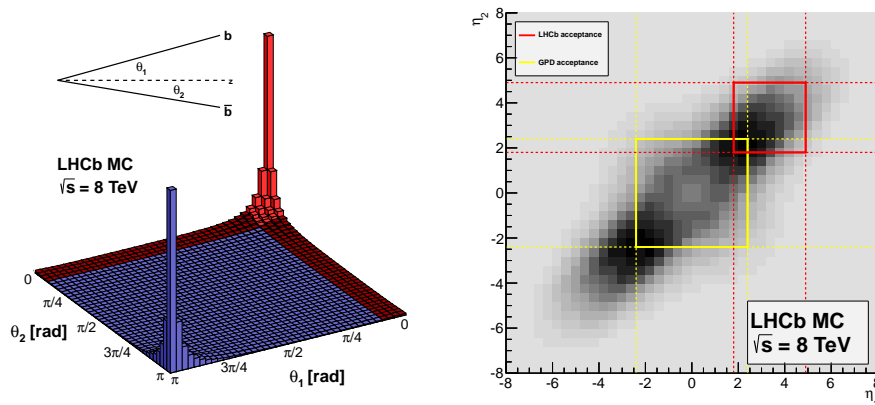


Figure 3.2 – Distributions of b and \bar{b} quarks being produced in pp collisions at $\sqrt{s} = 8$ TeV as a function of the polar angle θ with respect to the beam axis (left) and pseudorapidity η (right) obtained from simulation [50]. In the left figure, the height of the bars represents the frequency of $b\bar{b}$ quark pairs being produced in the respective bin. Clear peaks are visible close to the beam axis. The red area shows the angular acceptance of the LHCb detector. In the right figure, darker bins represent a higher frequency of $b\bar{b}$ pair production in the respective bin. The yellow area shows the acceptance of a GPD, while the red area shows the LHCb acceptance.

covering the range of high rapidity. The detector design is shown in Figure 3.3. A detailed description of the LHCb detector can be found in Ref. [51].

The sub-detectors are aligned next to each other along the beam axis, covering an angular acceptance from 10 mrad to 300 mrad (250 mrad) in the horizontal (vertical) plane. Therefore, LHCb approximately covers the pseudorapidity range of $2 < \eta < 5$. This design results in approximately 25 % of all $b\bar{b}$ quark pairs being produced inside the detector acceptance. The right-handed coordinate system has its origin inside the vertex locator, where the pp interactions occur. The z -axis is oriented along the beam pipe and the y -axis vertically points upwards.

At the maximum instantaneous design luminosity of the LHC of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, an average bunch crossing contains many pp interactions, each producing a high multiplicity of tracks. The planned precision measurements at LHCb require an efficient reconstruction, which is only possible up to a limited level of track multiplicity. Therefore, the LHCb experiment is designed to run at a lower luminosity of $2 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, which is realised by changing the beam focus through displacing the colliding beams. On average, this results in less than one pp interaction per bunch crossing in the nominal set-up and has the additional advantage of limiting the radiation damage in the detector. The beam displacement is constantly adjusted over a fill of the LHC, so that the luminosity is kept constant although

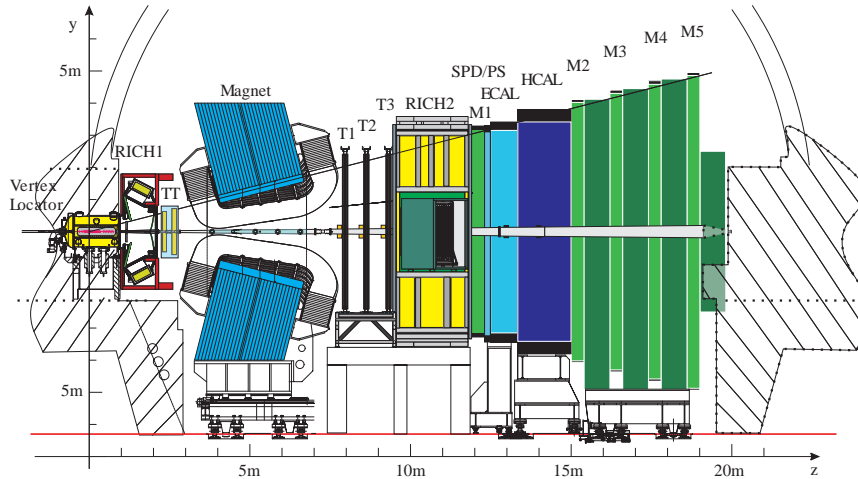


Figure 3.3 – Schematic side view of the LHCb detector [51]. The pp interactions occur in the origin of the coordinate system inside the vertex locator. Both proton beams traverse the detector through the beam pipe that spans through the detector at $x = y = 0$. The sub-detectors are labelled in the figure.

the beam intensities decrease over time. This method is unique among the LHC experiments and called luminosity levelling.

The possible interaction of decay products with the detector material can produce secondary particles and thus complicate the reconstruction. Therefore, a key design aspect is to minimise the material budget up to the end of all tracking sub-detectors. This is also reflected in the beam pipe, which is partly constructed from beryllium and therefore offers a high transparency towards the high-rapidity particles.

The protective systems Beam Conditions Monitor (BCM) and Beam Loss Scintillator (BLS) monitor the particle flux near the beam pipe [52, 53]. Their task is to detect possible adverse beam conditions like misaligned beams, which need to be detected as early as possible to avoid damaging the detector components. The BCM is installed at two positions opposite to the interaction point around the beam pipe. It monitors the beam quality through chemical-vapour deposition diamond sensors and is designed to immediately and autonomously request a beam dump in case certain radiation thresholds in the sensors are exceeded. The BLS offers a high time resolution and is more sensitive than the BCM. It therefore offers a complementary approach to monitor radiation background.

3.2.1 Tracking system

The tracking system at LHCb contains the vertex locator (VELO), the tracker Turicensis (TT), the inner tracker (IT), and the outer tracker (OT), where the latter

two form the three tracking stations T1, T2, and T3. It is responsible for the efficient and precise vertex and momentum determination mandatory for the LHCb physics programme despite the high multiplicity of the hadronic environment. All tracking detectors measure hits produced by charged particles traversing the sensitive material. The subsequent reconstruction algorithms combine these hits into tracks. An essential component for the momentum determination is the warm dipole magnet with an integrated magnetic field of 4 Tm. It bends charged particles on a curvature from which the momentum can be deduced. Furthermore, the magnet polarity can be flipped to reduce systematic uncertainties.

In the LHCb Run I data taking, the overall track reconstruction efficiency is measured to be $> 96\%$ for tracks with hits in the VELO, TT, and tracking stations T1–T3 in the momentum range of 5 GeV/ c to 200 GeV/ c . Furthermore, the relative momentum resolution $\Delta p/p$ is found to range from 0.5% for 20 GeV/ c tracks to 0.8% for 100 GeV/ c tracks [11].

Vertex locator

The VELO is a tracking detector based on silicon strips, which is closely built around the pp interaction region. Its task is precise tracking around the collision in order to reconstruct the vertices of both the primary pp interactions, called the primary vertices (PVs), and of the decays of long-lived particles, for instance of b and c hadrons. Due to the finite lifetime and the large boost, the latter are significantly displaced from the PVs with flight distances of up to $\mathcal{O}(\text{mm})$. A precise measurement of these vertices and the momenta directly allows for precisely measured decay times of the particles, which is a key requirement for time-dependent measurements of CP violation.

The VELO consists of 21 stations arranged along the beam axis, with each station consisting of two modules. The modules are installed horizontally opposite to each other and each one contains two half-disc-shaped sensors with a radius of 42 mm, one with quasi-radial silicon strips to achieve polar sensitivity (called ϕ sensors) and one with concentric semicircular strips to determine radial coordinates of the tracks of charged particles (called r sensors). The pitch of the silicon strips varies between approximately 40 μm and 100 μm . In the nominal position for data taking, the half discs slightly overlap and the innermost radius of the sensitive area is as low as 8 mm, which is smaller than the LHC beam aperture at beam injection. Therefore, the modules are retractable in horizontal direction into a parking position, where the modules are protected from radiation damage. Only after obtaining stable beams, the modules are moved into the nominal position. The VELO modules are housed in a vacuum vessel. This allows to reduce the beam pipe to a thin corrugated aluminium foil, which permits to move the modules very close to the beam, minimises the material budget traversed by the particles,

and furthermore acts as radio-frequency shielding.

In Run I, the achieved impact parameter (IP) resolution is determined to be $< 35 \mu\text{m}$ for tracks with a transverse momentum of $p_T > 1 \text{ GeV}/c$, while the average decay-time resolution of B mesons is measured as $\approx 50 \text{ fs}$ [54].

Silicon tracker and outer tracker

The remaining tracking detectors are TT, IT and OT. TT and IT are both based on silicon strip detectors with a pitch of about $200 \mu\text{m}$ and commonly both called silicon tracker (ST). After the RICH1 detector and right before the magnet, the TT is installed. It consists of four layers and spans an area of 150 cm in width and 130 cm in height. The tracking stations T1–T3 are installed directly after the magnet, where the IT spans the inner high-multiplicity area of 120 cm in width and 40 cm in height, where a good spatial resolution and fast detector response are necessary to distinguish between tracks. In the outer area, the OT covers the remaining part of the LHCb acceptance. It is based on straw drift tubes with a diameter of $\approx 5 \text{ mm}$. Both the IT and OT are composed of four layers in each station of T1–T3.

All three detectors follow a so-called x - u - v - x configuration, where a layer of vertically aligned strips or drift tubes (x) is followed by two layers being rotated by a stereo angle of -5° (u) and $+5^\circ$ (v). The fourth layer is again vertically aligned. This provides little sensitivity of the tracking detectors in the y -direction.

3.2.2 Particle identification system

As b and c hadrons decay into many different final states relevant for the physics programme of LHCb and as the hadronic environment results in a high multiplicity of diverse particle tracks, it is essential to correctly identify the final state particles with small misidentification rates. Therefore, a set of different detectors is designed for particle identification (PID) purposes. These are especially important as many decays of interest are rare compared to similar decays that occur at higher frequencies. A sizeable misidentification rate would thus pollute the samples. The PID detectors at LHCb are the two ring-imaging Cherenkov detectors (RICH), the calorimeters, and the muon system.

Ring-imaging Cherenkov detectors

The two RICH detectors utilise the Cherenkov effect to mainly distinguish between pions, kaons and protons. Charged particles travelling through the radiator materials emit a cone of Cherenkov light. Combined with the momentum information, the opening angle of this cone allows to draw conclusions about the particle's identity. In order to minimise the material budget inside the angular acceptance,

both RICH detectors use thin spherical and flat mirrors to guide the light outside the detector's acceptance, where the projected light circles are measured by magnetically shielded hybrid photon detectors (HPDs).

The RICH1 detector is located between VELO and TT, uses the two radiators aerogel and C_4F_{10} , and covers the low momentum range from 1 GeV/c to 60 GeV/c. It covers the whole angular acceptance. RICH2 is installed after the tracking stations T1–T3 and uses CF_{10} as radiator. It is designed for particles with high momentum from 15 GeV/c to 100 GeV/c and therefore only covers a limited angular acceptance.

The kaon identification efficiency in Run I is measured as 95 % at a $K-\pi$ misidentification rate of 10 % (i.e. 10 % of the pions are identified as kaons). At a stricter PID requirement the kaon identification efficiency is 85 % at a $K-\pi$ misidentification rate of 3 % [11].

Calorimeter system

The calorimeter system is composed of the electromagnetic calorimeter (ECAL) and hadron calorimeter (HCAL), where the ECAL is additionally supported by the scintillator pad detector (SPD) and preshower detector (PS). These detectors provide PID for electrons, photons and hadrons, measure the energy deposition of particles stopped in the calorimeter's absorbers, and additionally measure particle positions through the position of electromagnetic showers. Furthermore, the calorimeters measure transverse energy deposition, which is important for the decision of the first level trigger.

All calorimeters are based on scintillating materials which measure the deposited energy from the showers. These are segmented in the xy -plane with smaller segments close to the beam. The SPD is the first calorimeter system and constructed from scintillators to distinguish electrons from photons. The PS follows after a lead plate and is mainly used for $e-\pi$ separation. Afterwards, the following ECAL is constructed as a so-called shashlik-type calorimeter from alternating tiles of lead plates to generate electromagnetic showers and scintillators to measure the energy. It is designed so that a complete high-energy photon shower is fully contained within. The HCAL is constructed similarly, but uses iron plates as stopping material.

Muon system

Many final states analysed at LHCb contain muons. These do not produce showers or otherwise deposit considerable amounts of energy in the calorimeters, so that they can traverse the calorimeters without being stopped. Therefore, muon detectors, which are typically installed after the calorimeter system, allow to use the

clear signature for muon identification. Furthermore, the system is also involved in the momentum measurement. All this results in very high trigger and muon identification efficiencies for muonic final states.

LHCb uses five muon stations, called M1–M5. M1 is installed in front of the calorimeters and M2–M5 are located after the HCAL. The latter are interleaved with iron plates as absorbers. All stations are segmented in the xy -plane with smaller segments close to the beam. The inner part of M1 is instrumented with triple-GEM detectors, the rest of the muon system uses multi-wire proportional chambers to measure the muon tracks.

In the LHCb Run I the muon identification efficiency is measured as $>95\%$, with a misidentification rate for pions, protons and kaons of $<1\%$ [11].

3.3 Data processing at LHCb

The data from the detector is processed in several stages, starting with the trigger, which itself is comprised of three different parts. Triggered events are then processed by the reconstruction software and the following analysis software applying further requirements to suppress background. Additionally, the simulation allows to generate samples to validate analyses. All LHCb software packages are based on the GAUDI framework [55].

3.3.1 Trigger

The harsh hadronic environment of the LHC requires a highly efficient and flexible trigger system to reject the large rate of background events while retaining as many decays of interest as possible. At LHCb, the trigger is divided into three stages: The first is the hardware trigger (L0), which is able to process events at the LHC proton bunch crossing rate of 40 MHz and uses the transverse energy deposited in the calorimeters to identify hadron, photon and electron candidates or hits in the muon system to identify single muon or dimuon candidates. It reduces the rate of events with inelastic pp interactions to 1 MHz.

Events triggered by the L0 are then processed by the high level trigger (HLT), which is a software trigger and organised in the MOORE package [56]. It is again divided into HLT1 and HLT2. Different so-called lines in both stages aim at identifying different exclusive and inclusive decays for various physics purposes. HLT1 performs a partial event reconstruction with a simplified tracking algorithm, further reducing the rate. A fast muon identification for (di)muon triggers searches for hits in the muon system being compatible with extrapolated VELO tracks. Finally, HLT2 performs a more complex event reconstruction, allowing more sophisticated decisions. It features trigger lines for decays including (di)muon final states, for c hadron decays, and topological lines that aim at identifying

partially reconstructed b hadron decays. More details about the trigger lines relevant for this analysis can be found in Section 4.1.4. The trigger configuration used in the data taking of 2012 is shown in Figure 3.4.

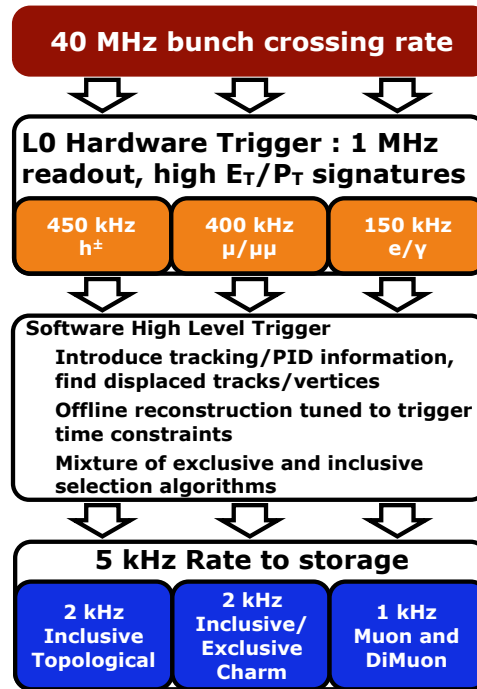


Figure 3.4 – Scheme of the LHCb trigger architecture for the data taking of 2012 [57].

Because of inter-fill preparations, machine development phases, and technical difficulties, the LHC on average delivers stable pp collisions in only 30 % of the time. The LHCb trigger in Run I uses the so-called deferred trigger to maximise the utilisation of the HLT computing resources and therefore optimise the quality of the triggered physics data [58]. During pp collisions, about 20 % of the L0-triggered events are temporarily written to the local hard disks of the HLT computing nodes. These are then processed afterwards when no pp collisions are performed. This allows to use less strict requirements in the trigger lines and utilise better reconstruction algorithms.

In the data taking of 2011, the HLT output rate was 3 kHz, while in 2012 it was increased to 5 kHz. A thorough description of the LHCb trigger and its performance in the Run I data taking can be found in Refs. [59, 60]. The trigger efficiency for dimuon triggers in Run I is determined as $\approx 90\%$ [61].

3.3.2 Reconstruction and analysis

The offline reconstruction is performed on the triggered events by the BRUNEL package [62], which contains algorithms performing tracking, PID tasks, and the reconstruction of neutral particles from clusters in the calorimeters. The algorithms provide so-called protoparticles, which are used in the subsequent analyses.

Different methods are utilised for the tracking. One starts by reconstructing tracks from hits in the VELO. These are then extrapolated and combined with hits or already identified tracks in the tracking stations T1–T3. Afterwards, the algorithm searches for matching hits in the TT. The tracks determined that way span all tracking sub-detectors and are therefore called long tracks. An alternative algorithm identifies tracks in T1–T3 and combines these with matching hits in the TT, which results in so-called downstream tracks. These do not include any VELO hits and are important for the reconstruction of long-lived particles like K_S^0 mesons which can decay outside the VELO acceptance. Further algorithms reconstruct different track types which are not relevant for this analysis. Following the different reconstruction approaches, duplicate tracks are removed and a track fit is performed that precisely determines the trajectory and particle momentum.

The PID algorithms process and combine information from the RICH detectors, the muon system, and the calorimeters. In case of the RICH detectors, a global likelihood is calculated from all tracks and the hits in the HPDs, successively varying the particle hypotheses of single tracks [63]. The muon PID calculates likelihoods for muon and non-muon hypotheses from extrapolated tracks and hits in the muon system in the vicinity of these tracks [64]. Furthermore, the calorimeter information is mainly used for the identification of electrons, photons, and hadrons. It is additionally utilised to recover bremsstrahlung photons and match calorimeter clusters to extrapolated tracks. Finally, to improve the overall PID performance, the information of all PID systems is added to combined likelihoods for each hypothesis. An improved approach uses multivariate techniques taking correlations between the detector systems and additional information into account to derive combined probabilities for each particle hypothesis [11].

The software package used for the subsequent analysis is called DAVINCI [65]. It combines the protoparticles from BRUNEL into intermediate particles to build the complete decay chain of interest and is used to apply selection requirements. A large set of algorithms is available to produce datasets with various observables of the particles. These algorithms provide kinematic, geometric, and PID information; analyse trigger responses; and match reconstructed particles to generated particles in case of simulated samples. More details about the construction of the decay chain and the selection relevant for this analysis are given in Section 4.1.

The stripping selection as a part of DAVINCI is the centralised pre-selection of LHCb data. It applies very loose requirements to reduce the size of the data sample

and save computing resources. For more details about the stripping selection used in this measurement, see Section 4.1.3.

3.3.3 Simulation

The generation of simulated samples at LHCb is handled by the software packages GAUSS [66] and BOOLE [67]. Simulated samples describe LHCb events from the pp interaction to the detector response and allow to gain a better understanding of the various influences on the data, test different physics assumptions, and validate analysis strategies. The GAUSS package simulates the bunch crossing, the pp interactions, and the particles produced therein through the PYTHIA6 and PYTHIA8 generators [68]. The decay of these particles is then simulated with the generator EVTGEN [69], which utilises the PHOTOS package for radiative corrections [70]. The interaction of the particles with the detector material, like charge deposition and production of secondary particles and showers, is described using GEANT4 [71]. In the following step, the digitisation package BOOLE uses this particle-material interaction to simulate the detector response including the L0 decision. After the generation stage is complete, the simulated events are processed with the same tool chain as data events. This includes MOORE, BRUNEL, and DAVINCI, where in DAVINCI properties of the generated particles are accessible for the purpose of comparing these to the properties of the reconstructed particles.

3.4 Conditions of the data-taking in Run I

The Run I of data taking at LHCb mainly comprises the years 2011 and 2012. In 2011, a data sample corresponding to an integrated luminosity of 1.11 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$ was recorded, while in 2012 the integrated luminosity amounts to 2.08 fb^{-1} at $\sqrt{s} = 8 \text{ TeV}$ [11]. The evolution of the integrated luminosity in the different years of data taking is shown in Figure 3.5.

Although the LHC is intended to be operated at a proton bunch spacing of 25 ns, the minimum bunch spacing in Run I was 50 ns, which directly halves the amount of b and c hadron decays with respect to the original design in case no other adjustments are made. Furthermore, LHCb is designed to be operated at an average number μ of visible pp interactions per bunch crossing of $\mu \approx 0.4$. The loss in the number of bunch crossings was compensated by increasing μ to on average 1.4 in 2011 and 1.6 in 2012 [59, 60]. It is observed that this increase does not degrade the detector, trigger, and reconstruction performance.

In total, LHCb was successfully operated at an instantaneous luminosity of $3.5 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ in 2011 and $4 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ in 2012, which is both significantly larger than the design specification of $2 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ [11]. The data-taking efficiency, defined as the ratio of recorded luminosity to delivered luminosity by the LHC, was on average 93 % in 2011 and 95 % in 2012.

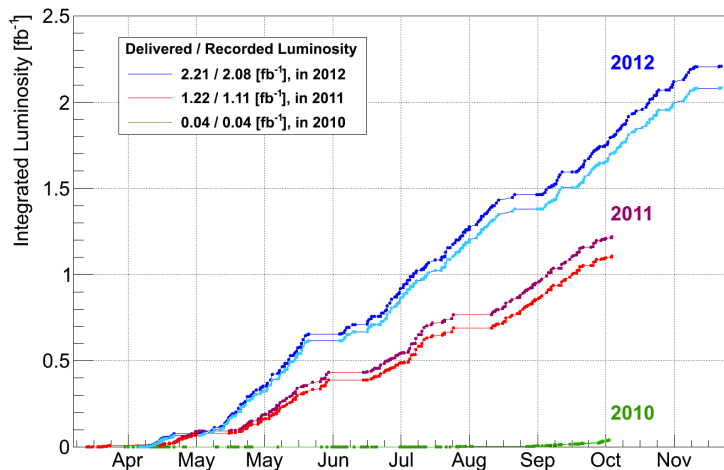


Figure 3.5 – Evolution of the integrated luminosity at LHCb in 2010–2012 [72].

3.5 Upgrade of the LHCb detector

After the successful Run I and the following two-year shutdown of the LHC, the next data-taking period at LHCb, Run II, started in 2015. The centre-of-mass energy at the LHC is increased to $\sqrt{s} = 13$ TeV, which results in the expected production cross-section for b and c quarks to be a factor of approximately 1.6 to 1.9 larger than in Run I. Furthermore, the proton bunch spacing is reduced from 50 ns to 25 ns. As a result, the average number of visible pp interactions per bunch crossing is reduced to $\mu = 1$, while operating LHCb at the same instantaneous luminosity of $4 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ as in 2012 [73]. The main improvements at LHCb with respect to Run I concern the trigger. The deferral strategy is changed so that the first software-trigger stage HLT1 processes all L0-triggered events. Afterwards, all HLT1-triggered events are deferred and successively and asynchronously processed by HLT2. Together with doubled computing resources for the HLT and an increased deferral capacity, a higher HLT output rate of 12.5 kHz is achieved and more sophisticated reconstruction and selection algorithms are performed. Additionally, the detector calibration and alignment is no longer conducted offline, but directly in the trigger. Finally, a part of the trigger output is stored in a reduced data format, which can directly be used in analyses without the need for further offline processing. In total, the data sample to be recorded in Run II is expected to correspond to an integrated luminosity of at least 5 fb^{-1} .

Run II is scheduled to end in 2018, followed by a two-year LHC shutdown in which the LHCb detector will undergo a major upgrade. In the following Run III, the instantaneous luminosity will be increased to $2 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$, resulting

in $\mu = 5.2$ [74]. Both the current detector and trigger are not able to handle the expected high detector occupancy and track multiplicity. Therefore, the readout electronics of all sub-detectors are upgraded to operate at the nominal LHC proton bunch crossing frequency of 40 MHz [75]. Furthermore, several sub-detectors are exchanged by improved technology.

For the VELO, the sensors based on silicon strips will be replaced by hybrid-pixel-based sensors with a pixel size of $55 \times 55 \mu\text{m}^2$ with the closest pixels being separated only 5.1 mm from the proton beams, directly providing three-dimensional tracking information [76]. Cooling will be integrated directly into the modules, which itself will be L-shaped. The thickness of both the detector modules and the foil are reduced.

The TT will become the upstream tracker (UT) with finer granularity in the high-occupancy region and improved radiation hardness [74]. Where for the TT, the signal processing is performed distant from the sensors, the UT will perform this task in close vicinity with reduced electronic noise. The IT and OT stations will be replaced by a unified tracking detector, which will cover the complete LHCb acceptance [74]. This tracker will be comprised of scintillating fibres with a diameter of $250 \mu\text{m}$, which will be incorporated into 2.5 m long modules and read out by cooled silicon photomultipliers. The scintillating fibre (SciFi) tracker features a high hit efficiency, excellent spatial resolution of up to $60 \mu\text{m}$, low noise, low material budget, and sufficient radiation hardness for the complete upgrade detector runtime. Both the UT and the SciFi tracker will use the identical x - u - v - x geometry as the former tracking detectors.

In case of the RICH detectors, the aerogel radiator will be removed and the HPDs, which do not allow a 40 MHz readout, will be replaced by multi-anode photomultipliers [77]. Apart from upgrading the readout electronics, the most significant change for the muon system is the removal of the first station M1 [77].

The LHCb trigger architecture will undergo crucial changes. At the planned luminosity and high rate of pp interactions, more complex trigger decisions are required and the responsibility of the trigger shifts from rejecting background to categorising different signal decays. As the L0 cannot satisfy these requirements, it will be completely removed [78]. Therefore, LHCb will feature a trigger-less readout and processing of all events in software by the HLT. This upgraded trigger will feature unprecedented flexibility and reconstruction algorithms that will be of identical quality as the offline reconstruction.

The total dataset recorded with the upgraded LHCb detector is planned to correspond to an integrated luminosity of 50 fb^{-1} [74]. The implications of the future LHCb data-taking runs for the measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ decays are discussed in Chapter 6.

4 Analysis ingredients and preparatory studies

The measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ decays comprises several preparatory studies which need to be conducted before the actual measurement can be performed. Additionally, it is based on techniques that require introduction and examination themselves.

Initially, the data sample itself needs to be refined in several steps in order to increase the signal purity. Connected to that procedure is the study of undesirable background contributions in the sample. Careful studies of the flavour tagging algorithms, which allow for deducing the production flavour of the B mesons, are essential for a measurement of time-dependent CP asymmetries. Due to the high oscillation frequency of B_s^0 mesons, the decay-time resolution needs to be precisely determined and modelled in the description of the time-dependent decay rates. Finally, the various selection stages introduce a decay-time inefficiency, which is empirically described.

4.1 Data sample preparation

The measurement uses the full Run I dataset of pp collisions recorded by the LHCb detector in 2011 and 2012. The sample preparation consists of both constructing of suitable $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ candidates from the tracks in the events, as well as multi-stage processing of the sample to reduce the background contribution.

4.1.1 Event samples

The samples used in this analysis include both the data sample, which is used in different variations for the nominal CP violation measurement, preparatory studies, and cross-checks, as well as simulated samples, which are required for preparatory studies and to validate certain analysis steps.

LHCb data

The full Run I data sample contains pp collisions collected at a centre-of-mass energy of $\sqrt{s} = 7$ TeV in 2011 corresponding to an integrated luminosity of 1 fb^{-1}

and at $\sqrt{s} = 8$ TeV in 2012, with an integrated luminosity of 2 fb^{-1} . Only a small fraction of these collision events is of interest for the analysis, which is extracted with the steps discussed in the following sections.

The raw data is processed with BRUNEL v43r2p6 (Reco14) to reconstruct particle tracks and with DAVINCI v32r2p1 for the stripping selection (Stripping20). Tuples to be used in the analysis are produced with DAVINCI v33r9 using revision r167275 of the FLAVOURTAGGING package.

The data sample inevitably contains both $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ decays. This is firstly due to the mass difference of B_s^0 and B^0 mesons, $m_{B_s^0} - m_{B^0} \approx 87 \text{ MeV}/c^2$ [19]. Secondly, as the branching ratio $\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0) / \mathcal{B}(B^0 \rightarrow J/\psi K_S^0) \approx 0.044$ [79] and as the ratio of B_s^0 to B^0 meson hadronisation fractions is $f_s/f_d \approx 0.26$ [80, 81], the sample contains around 100 times more $B^0 \rightarrow J/\psi K_S^0$ than $B_s^0 \rightarrow J/\psi K_S^0$ decays. Therefore, the contribution of the former cannot be eliminated with a requirement on the reconstructed mass or other criteria. The implication is that the complete analysis is simultaneously performed for both channels, which makes the analysis more complex, but also has the advantage of being able to utilise the larger $B^0 \rightarrow J/\psi K_S^0$ contribution as a representative to determine properties for the $B_s^0 \rightarrow J/\psi K_S^0$ component. This is for instance used in the selection and for the decay-time acceptance (see Section 4.5). In view of possible future analyses aiming to measure CP violation in both decay channels simultaneously, this measurement can provide valuable information. In the following, $B \rightarrow J/\psi K_S^0$ is used to denote both decays.

Simulated samples

Two types of simulated samples are used in this analysis. The first are fully simulated Monte Carlo (MC) samples (see Section 3.3.3), which simulate the pp collision and detector response and are afterwards processed with the same software tool chain as the data sample, where a specific MC configuration is used for the FLAVOURTAGGING package. Properties of the proton beams, like energy and crossing angle, and of the detector configuration are set up as close as possible to the data-taking conditions. As the simulation is exceptionally time-consuming, only limited amounts of simulated data are available. Fully simulated MC samples allow to associate reconstructed candidates and their properties to generated particles, which is called truth-matching. This makes them ideal to study effects which cannot be examined in data or to validate analysis techniques.

The second type of simulated samples are Toy Monte Carlo samples, for which the properties of the candidates are generated from probability density functions (PDFs) which describe the according distributions. That way, large samples of these pseudo-experiments can be generated with little computing resources. However, as no underlying particle interactions are simulated, assumptions on the

distributions are required either from phenomenological observations or theory. Toy MC samples are used to validate the fitter, study systematic uncertainties and determine statistical properties of the results. The properties of the specific samples are described in the according sections.

Fully simulated samples are produced with GAUSS v45r2 to v45r7 in the Sim08 configuration. The events are generated using the generators PYTHIA6 and PYTHIA8, the data-taking conditions of 2011 and 2012, and both magnet polarities. The fully simulated MC samples used in the analysis are:

$B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ signal MC Samples of 4 million generated $B_s^0 \rightarrow J/\psi K_S^0$ and 8 million generated $B^0 \rightarrow J/\psi K_S^0$ decays are used to validate analysis techniques, study systematic uncertainties, and tune parts of the selection.

$B^0 \rightarrow J/\psi K^{*0}$ signal MC In total 10 million generated $B^0 \rightarrow J/\psi K^{*0}$ decays are available and used to tune parts of the selection and study the flavour tagging.

Inclusive $J/\psi \rightarrow \mu^+ \mu^-$ MC A sample with 40 million generated inclusive $J/\psi \rightarrow \mu^+ \mu^-$ decays is used to study the decay-time resolution and possible background contributions.

Inclusive $B \rightarrow J/\psi X$ MC Two samples with each 6 million generated $B^0 \rightarrow J/\psi X$ and $B^+ \rightarrow J/\psi X$ decays and one sample with 2 million generated $B_s^0 \rightarrow J/\psi X$ decays are used to study possible background contributions in the analysed sample. In contrast to the other MC samples, these samples are generated in the Sim05 configuration with data-taking conditions of 2011.

4.1.2 Candidate reconstruction

The $B \rightarrow J/\psi K_S^0$ candidates used in the analysis are reconstructed through the subsequent decays $J/\psi \rightarrow \mu^+ \mu^-$ and $K_S^0 \rightarrow \pi^+ \pi^-$. Other final states that J/ψ and K_S^0 can decay into are neglected. In case of the J/ψ , the $\mu^+ \mu^-$ final state leaves an especially clean signature and both the detector and trigger are designed for optimal muon reconstruction efficiency which justifies its choice despite the small branching fraction of $\approx 6\%$. The additional J/ψ decays into $e^+ e^-$ and distinct hadronic final states are significantly more difficult to reconstruct and would each require specific treatment so that their limited statistical power does not justify their inclusion in the scope of this analysis. Analogously, $K_S^0 \rightarrow \pi^0 \pi^0$ decays are not beneficial because of the subsequent $\pi^0 \rightarrow \gamma \gamma$ decays, which are complicated to reconstruct.

Figure 4.1 shows the schematic decay topology of $B \rightarrow J/\psi K_S^0$ decays. The B meson is produced at the primary vertex (PV) from where it traverses through the detector until it decays into $J/\psi K_S^0$ forming the secondary vertex (SV). As the lifetime of the J/ψ is too short to decay at another displaced vertex, both $\mu^+ \mu^-$

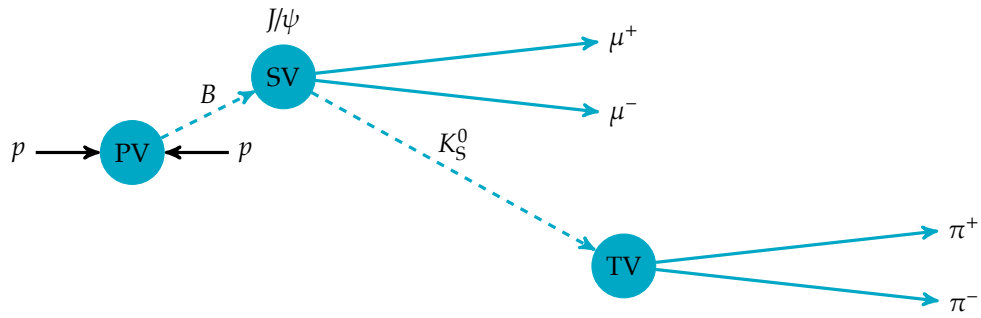


Figure 4.1 – Schematic decay topology of $B \rightarrow J/\psi K_S^0$ decays. Solid lines represent the tracks of final state particles that are reconstructed. Dashed lines show the trajectories of intermediate particles. Sizes and lengths are not to scale.

particle tracks will originate from the SV. The K_S^0 , however, traverses through the detector to decay at the tertiary vertex (TV) into a $\pi^+\pi^-$ pair.

Accordingly, B meson candidates are reconstructed from each two tracks of oppositely charged muon and pion candidates. Due to the high lifetime, approximately only one third of all K_S^0 mesons decay inside the VELO detector acceptance, so that both two long (LL) or two downstream (DD) tracks (see Section 3.3.2) can be used to form a K_S^0 candidate. In general, downstream tracks are less precise than long tracks, leading to less precisely reconstructed K_S^0 and B candidates. This will influence other properties as well so that the LL and DD subsamples are treated separately throughout the analysis.

Two approaches to reconstruct the decay chain of a B candidate are used in this analysis. The first is the default bottom-up method used by DAVINCI, which starts to combine final state particles into intermediate particles using χ^2 fits that determine the decay vertex and the kinematic properties of the intermediate particles. These particles are then combined likewise until the complete decay chain is reconstructed. This approach is straightforward and efficient as it already allows to neglect candidates at the first stages of the combination without building the complete decay chain. However, constraints used in one fit cannot be propagated downstream to preceding fits.

The decay tree fitter (DTF, often also referred to as decay chain fitter) offers a more sophisticated approach by fitting the complete decay chain in a single fit [82] which simultaneously determines all decay vertices as well as the particles' kinematic properties and their uncertainties. This method has several advantages over the standard step-by-step procedure. For instance, it takes correlations among the parameters into account and conveniently allows to integrate external constraints into the fit, which will be propagated to all particles of the decay chain properly. Commonly, the reconstructed mass of an intermediate particle can be constrained

to its known mass or it can be required that the momentum of the B candidate points back to the PV. As the DTF is more complex and computing intensive, it is performed after initially applying loose selection requirements to suppress obvious background contributions.

A single LHCb event is likely to contain multiple pp interactions and thus multiple reconstructed PVs. It is not a priori obvious in which PV a B meson candidate was produced so that in principle each candidate can be associated with any PV. Furthermore, one event can contain multiple $B \rightarrow J/\psi K_S^0$ candidates. It is highly unlikely that more than one genuine $B \rightarrow J/\psi K_S^0$ decay is produced in a single pp collision and in fact these multiple candidates regularly share parts of their daughter candidates. Evidently, only one of the PV and candidate combinations should be considered for the analysis. A common strategy is to select a single ‘best PV’ and ‘best candidate’ based on previously defined criteria for each event and neglect all others. However, this approach can bias the measurement in case the best PV/candidate criteria are correlated with any of the observables used in the analysis [83]. Therefore, all PVs and candidates in each event are treated equally throughout all steps of the selection. The latter implicitly and explicitly uses criteria to neglect incorrectly associated PVs and background candidates, which directly reduces the rate of multiple PVs and candidates. Afterwards, for each event with multiple candidate-PV combinations, one is chosen at random to avoid the introduction of biases.

4.1.3 Stripping selection

The size of the LHCb data sample, the small frequency in which decays of interest occur, and the extensive range of decays studied at LHCb make it infeasible for each analysis to process the whole dataset. The stripping is the centralised processing of the full data sample using simple and efficient reconstruction and selection criteria in order to provide a set of manageable samples, so-called stripped data, which can be processed in each analysis. These samples are grouped between analyses with similar decay modes, e.g. all analyses investigating $\mu^+\mu^-$ final states, to avoid computational overhead. The selection requirements applied in the stripping are generally very loose to avoid rejecting signal decays and to keep a reasonable amount of background for studies.

Two samples from the stripping are used in this analysis. The first, referred to as the ‘detached’ sample, is the nominal sample upon which most parts of the analysis rely. It is based on the default step-by-step reconstruction to build $B \rightarrow J/\psi K_S^0$ candidates. Table 4.1 summarises the applied selection requirements. The selection is designed to efficiently consider only suitable final state particles by applying requirements on these before combining them to intermediate particle candidates. A common criterion is $\chi_{\text{DOCA}}^2/\text{ndf}$ which approximately denotes how

significant the distance of closest approach (DOCA) of two tracks is different from zero with respect to its uncertainty. It is used to reject tracks clearly not originating from a common vertex before performing a time-consuming vertex fit. The quantity $\chi_{\text{IP,PV}}^2$ approximately describes the χ^2 of the impact parameter (IP) of the particle's track with respect to the PV and thus helps to reject tracks originating from the latter. Afterwards, J/ψ and K_S^0 candidates are constructed using a vertex fit. Its fit quality and requirements on the candidate's mass are used to further dismiss background. The decay length significance requirement on the K_S^0 candidates ensures a significantly displaced TV, which rejects background from misidentified instantaneous K^{*0} decays. Finally, the B candidate is constructed from the J/ψ and K_S^0 candidates. Additionally to the vertex fit quality and B candidate mass, the requirement on the decay time $t > 0.2$ ps demands the B candidate to be detached from the PV, which rejects background from prompt tracks originating in the PV.

The requirement on the decay time makes the detached sample unsuitable for studies that require prompt candidates with $t = 0$. Therefore, a second sample, called the 'prescaled' sample, is used in the analysis, which applies the identical criteria as the detached sample with the exception of the decay-time requirement. The inclusion of candidates with vanishing decay time adds an extensive amount of prompt background to the sample, which also increases the required computing resources to run the stripping. In order to meet computing constraints, the prescaled stripping selection processes only a random subset of 30 % of the whole data sample.

4.1.4 Trigger requirements

Additionally to the stripping, specific requirements regarding the trigger are applied to the sample. Specifically dimuon triggers, i.e. trigger algorithms aiming at selecting $\mu^+\mu^-$ final states, show an excellent signal efficiency. Requesting a positive decision of these algorithms therefore helps to further reject background. Moreover, the J/ψ decay vertex strongly influences the decay-time resolution of the B candidate. By restricting the sample to events specifically triggered by the according algorithms, possible biases on the decay-time resolution can be minimised. For both the HLT1 and HLT2 triggers it is required that the reconstructed $J/\psi \rightarrow \mu^+\mu^-$ candidate alone fulfils all requirements.

No explicit requirements are applied with respect to the L0 trigger. In HLT1, any of the following trigger algorithms is required to provide a positive decision:

- TrackMuon
- DiMuonHighMass
- TrackAllL0

Table 4.1 – Requirements applied for the stripping of $B \rightarrow J/\psi K_S^0$ candidates. Utilised variables are the track fit quality of the final state particle $\chi_{\text{track}}^2/\text{ndf}$, the transverse momentum p_T , the difference in log-likelihoods for the π and μ hypotheses of the PID system $\Delta \ln \mathcal{L}_{\mu\pi}$, the reduced χ^2 of the distance of closest approach (DOCA) fit $\chi_{\text{DOCA}}^2/\text{ndf}$, the vertex fit quality $\chi_{\text{vtx}}^2/\text{ndf}$, the candidate masses $m_{\mu^+\mu^-}$, $m_{\pi^+\pi^-}$ and $m_{\mu\mu\pi\pi}$, the momentum p , the decay length significance DLS, and the decay time t . Furthermore, $\chi_{\text{IP,PV}}^2$ describes the difference of the χ^2 of the PV fit with and without the inclusion of the particle's tracks. Consequently, $\min(\chi_{\text{IP,PV}}^2)$ denotes the minimum of $\chi_{\text{IP,PV}}^2$ with respect to any PV.

Selection step	Candidate	Requirement	Unit
$J/\psi \rightarrow \mu^+\mu^-$	μ^\pm	$\chi_{\text{track}}^2/\text{ndf}$	< 5
		p_T	> 500 MeV/c
		$\Delta \ln \mathcal{L}_{\mu\pi}$	> 0
		$\chi_{\text{DOCA}}^2/\text{ndf}$	< 20
	J/ψ	$\chi_{\text{vtx}}^2/\text{ndf}$	< 16
		$m_{\mu^+\mu^-}$	> 3017 MeV/c ²
			< 3177 MeV/c ²
$K_S^0 \rightarrow \pi^+\pi^-$ (LL)	π^\pm	$\chi_{\text{track}}^2/\text{ndf}$	< 5
		p_T	> 250 MeV/c
		p	> 2000 MeV/c
		$\min(\chi_{\text{IP,PV}}^2)$	> 9
		$\chi_{\text{DOCA}}^2/\text{ndf}$	< 25
	K_S^0	$\chi_{\text{vtx}}^2/\text{ndf}$	< 20
		DLS	> 5
		$m_{\pi^+\pi^-}$	> 463 MeV/c ²
			< 533 MeV/c ²
$K_S^0 \rightarrow \pi^+\pi^-$ (DD)	π^\pm	$\chi_{\text{track}}^2/\text{ndf}$	< 10
		p	> 2000 MeV/c
		$\min(\chi_{\text{IP,PV}}^2)$	> 4
		$\chi_{\text{DOCA}}^2/\text{ndf}$	< 25
	K_S^0	$\chi_{\text{vtx}}^2/\text{ndf}$	< 20
		DLS	> 5
		$m_{\pi^+\pi^-}$	> 434 MeV/c ²
			< 562 MeV/c ²
$B \rightarrow J/\psi K_S^0$	B	$\chi_{\text{vtx}}^2/\text{ndf}$	< 10
		$m_{\mu\mu\pi\pi}$	> 5150 MeV/c ²
			< 5550 MeV/c ²
		t	> 0.2 ps

The requirements for these trigger algorithms are summarised in Table 4.2. Both TrackMuon and DiMuonHighMass require a positive decision of the L0 triggers LOMuon or LODiMuon and the tracks to be validated as muon candidates. The TrackAllL0 algorithm requires a positive decision of any L0 trigger. Further requirements among others aim at track quality, displacement of the tracks from the PV and, in case of the DiMuonHighMass trigger, demand two muon candidates originating from a common vertex.

Table 4.2 – Requirements applied by the HLT1 trigger algorithms used in the analysis. Criteria are the triggered particle’s track fit quality $\chi^2_{\text{track}}/\text{ndf}$, the minimum IP with respect to any reconstructed PV, the minimum of $\chi^2_{\text{IP,PV}}$ with respect to any PV, the transverse momentum p_T , the momentum p , the number of hits of the track in a sub-detector $N_{\text{hits},x'}$ and the number of missed hits of the track in the VELO $N_{\text{miss,VELO}}$. In case of the TrackMuon and TrackAllL0 triggers, criteria apply to a single candidate, which for TrackMuon has to be successfully identified as a muon. The DiMuonHighMass trigger criteria must be met by both muon candidates. Consequently, the distance of closest approach DOCA, the vertex fit quality $\chi^2_{\text{vtx}}/\text{ndf}$, and the candidate mass $m_{\mu^+\mu^-}$ apply to the combined $\mu^+\mu^-$ candidate. Numbers in parentheses indicate the requirement in the data taking of 2012, where this differs from 2011.

Requirement	TrackMuon	DiMuonHighMass	TrackAllL0	Unit
$\chi^2_{\text{track}}/\text{ndf}$	< 2	< 4	< 2.5	
$\min(\text{IP}_{\text{PV}})$	> 0.1	—	> 0.1	mm
$\min(\chi^2_{\text{IP,PV}})$	> 16	—	> 16	
p_T	> 1000	> 500	> 1700 (1600)	MeV/c
p	> 8000	> 6000	> 10000	MeV/c
$N_{\text{hits,VELO}}$	—	—	> 9	
$N_{\text{hits,OT}} + 2N_{\text{hits,IT}}$	—	—	> 16	
$N_{\text{miss,VELO}}$	—	—	< 3	
DOCA	—	< 0.2	—	mm
$\chi^2_{\text{vtx}}/\text{ndf}$	—	< 25	—	
$m_{\mu^+\mu^-}$	—	> 2700 (2500)	—	MeV/c ²

Additionally, any of the following HLT2 algorithms is required to provide a positive decision:

- DiMuonJPsi
- DiMuonJPsiHighPT
- DiMuonDetachedJPsi
- TopoMu2BodyBBDT

The requirements for the HLT2 trigger algorithms are summarised in Table 4.3. All three dimuon trigger algorithms require two candidates being identified as muons which originate from a common vertex with the candidate mass being compatible with a J/ψ . Additionally, the DiMuonDetachedJPsi trigger requires this vertex to be significantly displaced from the PV. The topological trigger TopoMu2BodyBBDT targets partially reconstructed b -hadron decays at displaced decay vertices with at least one track being identified as a muon candidate and additional requirements on the utilised tracks. Two tracks are then combined into a combined decay vertex and a boosted decision tree using different kinematic variables is employed to select suitable candidates.

Table 4.3 – Requirements applied by the HLT2 trigger algorithms used in the analysis. Criteria are the track fit quality $\chi^2_{\text{track}}/\text{ndf}$ of both muon candidates, the vertex fit quality $\chi^2_{\text{vtx}}/\text{ndf}$, the transverse momentum p_T , the candidate mass $m_{\mu^+\mu^-}$, and the flight distance with respect to its uncertainty $\chi^2_{\text{FD}}/\text{ndf}$. All criteria except $\chi^2_{\text{track}}/\text{ndf}$ apply to the J/ψ candidate. Numbers in parentheses indicate the requirement in the data taking of 2012, where this differs from 2011.

Requirement	DiMuonJPsi	DiMuonJPsiHighPT	DiMuonDetachedJPsi	Unit
$\chi^2_{\text{track}}/\text{ndf}$	< 5	< 5	< 5	
$\chi^2_{\text{vtx}}/\text{ndf}$	< 25	< 25	< 25	
$p_{T,\mu^+\mu^-}$	—	> 2000	—	MeV/c
$m_{\mu^+\mu^-}$	> 2977 (2997)	> 2977 (2997)	> 2977 (2997)	MeV/c ²
	< 3217 (3197)	< 3217 (3197)	< 3217 (3197)	MeV/c ²
$\chi^2_{\text{FD}}/\text{ndf}$	—	—	> 9	

More details of the trigger algorithms and their performance can be found in Refs. [59, 60].

4.1.5 Backgrounds

The data sample after applying stripping and trigger requirements still contains many background candidates which the following selection steps aim to reduce. It is essential to understand their nature and assess whether and how these influence the measurement of the CP asymmetries.

The largest background contribution comes from combinatorial background, where random tracks or J/ψ or K_S^0 decays originating in the PV are reconstructed as a $B \rightarrow J/\psi K_S^0$ candidate. Even after applying tight selection requirements, a significant number of non-signal combinations is still present in the sample due to the high track multiplicity of pp interactions. Thus, combinatorial background can only be reduced but never completely be eliminated. Therefore, it has to be modelled in the fit for the CP asymmetries. The mass distribution of combinatorial

background candidates is generally flat so that it can be separated from the signal using a mass fit.

Misreconstructed decays, also referred to as exclusive background, are the second class of background contributions, which typically consist of a specific b -hadron decay where the daughter J/ψ is combined with a misreconstructed K_S^0 candidate. These candidates usually show very distinct mass and decay-time distributions, thus making a careful study essential. Preferably, exclusive background is eliminated using suitable selection criteria. If that is not feasible, its contribution needs to be modelled.

Simulated samples are examined to determine possible exclusive background contributions. For that purpose, the inclusive $J/\psi \rightarrow \mu^+ \mu^-$ and $B \rightarrow J/\psi X$ samples are processed with the stripping selection requirements and the resulting candidates are identified with generated particles wherever possible. This allows to identify which generated decays are accidentally reconstructed as $B \rightarrow J/\psi K_S^0$ candidates. Table 4.4 summarises the observed contributions.

Table 4.4 – Number of occurrences in which generated decays are reconstructed as $B \rightarrow J/\psi K_S^0$ candidates and associated with their respective simulated particles after the stripping and additional preparatory loose requirements (see Section 4.1.6) are applied to the respective samples. Signal decays are listed in bold.

Sample	Size	Decay	#Stripping	#Loose Cuts
$B^0 \rightarrow J/\psi X$	6 M	$B^0 \rightarrow J/\psi K_S^0$	43837	28249
		$B^0 \rightarrow J/\psi K^{*0}$	2948	365
$B_s^0 \rightarrow J/\psi X$	2 M	$B_s^0 \rightarrow J/\psi K_S^0$	696	460
		$B_s^0 \rightarrow J/\psi \phi$	59	9
$B^+ \rightarrow J/\psi X$	6 M	$B^+ \rightarrow J/\psi K^{*+}$	22	—
incl. $J/\psi \rightarrow \mu^+ \mu^-$	20 M	$B^0 \rightarrow J/\psi K_S^0$	3262	2237
		$B^0 \rightarrow J/\psi K^{*0}$	120	11
		$\Lambda_b^0 \rightarrow J/\psi \Lambda$	57	—
		$B_s^0 \rightarrow J/\psi K_S^0$	46	37

The study shows that only few decays contribute to the misreconstructed background. Most notably, $B^0 \rightarrow J/\psi K^{*0}$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays are incorrectly reconstructed as $B \rightarrow J/\psi K_S^0$, where the kaon (proton) of the subsequent $K^{*0} \rightarrow K^+ \pi^-$ ($\Lambda \rightarrow p \pi^-$) decay is misreconstructed as a π^+ . The contribution from $B^0 \rightarrow J/\psi K^{*0}$ decays is eliminated with the help of a neural net, see Section 4.1.7, while the contamination arising from $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays is suppressed by requirements described in Section 4.1.6. Afterwards, both contributions can be neglected in the further analysis. The remaining observed exclusive background decays are rare

compared to $B_s^0 \rightarrow J/\psi K_S^0$. As the subsequent selection steps contain requirements improving the $K_S^0 \rightarrow \pi^+ \pi^-$ decay reconstruction, the relative contribution of misreconstructed background will diminish even further. Therefore, the remaining background decays do not require special treatment. Finally, when comparing the results on the $B \rightarrow J/\psi X$ samples it should be noted that the relative sample size between $B_s^0 \rightarrow J/\psi X$ and $B^0 \rightarrow J/\psi X$ does not properly reflect the relative production of B_s^0 with respect to B^0 mesons at LHCb of $f_s/f_d \approx 0.25$ [81].

4.1.6 Further preparatory selection requirements

In preparation of the multivariate selection steps discussed in the following sections, loose criteria are applied to the reconstructed B candidates after the stripping and trigger requirements. These are summarised in Table 4.5 and help to remove obvious background as well as to restrict several quantities to reasonable boundaries, for instance by neglecting candidates with very large mass and decay-time uncertainty estimates.

Multiple decay tree fits are performed for all candidates. The default fit constrains the J/ψ and K_S^0 candidate masses to their known values [19] and additionally constrains the B candidate momentum to point to the PV (PV constraint). Quantities from this fit are used for all requirements except those on the K_S^0 candidate. Additionally to the listed criteria, this decay tree fit is demanded to be properly converged. Requirements on the K_S^0 candidate are extracted from a fit with only the PV constraint. Contributions from misreconstructed $\Lambda_b^0 \rightarrow J/\psi \Lambda$ decays are suppressed by excluding all candidates for which the K_S^0 daughter mass under the $p\pi^-$ or $\pi^+\bar{p}$ hypothesis is compatible with the Λ mass. Additional decay tree fits with the alternating mass hypotheses and the PV constraint are performed for that task.

Misreconstructed $B^0 \rightarrow J/\psi K^{*0}$ background is suppressed by requiring a positive K_S^0 candidate decay time. As the K^* decays instantaneously, its decay-time distribution will be a smeared peak around $t \approx 0$ due to the finite decay-time resolution. Neglecting all negative values for t reduces roughly half of the according background.

As the B candidate mass from the DTF differs from the variable used in the stripping, the additional requirement on $m_{J/\psi K_S^0}$ removes tails resulting from these differences. Neglecting PVs with large z -position removes events where a pp collision occurred outside the designated interaction region due to narrow beam crossing angles during the data taking of 2011.

In principle, for all events with multiple PVs, each B candidate will be associated with every PV while of course only a single PV can be correct. Incorrectly associated candidates are not necessarily a problem. In most cases these candidates will be poor with respect to criteria used in the selection and eventually be neglected.

Table 4.5 – Additional selection criteria applied to reconstructed $B \rightarrow J/\psi K_S^0$ candidates after stripping and trigger requirements. All quantities are obtained through decay tree fits. Used variables are the smallest χ_{IP}^2 of the J/ψ candidate to all *other* PVs, the K_S^0 candidate mass under the $p\pi^-$ or $\pi^+\bar{p}$ hypothesis $m_{p\pi^-}/m_{\pi^+\bar{p}}$, the mass error estimate of the K_S^0 candidate $\sigma_{m_{\pi^+\pi^-}}$, the respective decay time of the K_S^0 and B candidate t , the DTF fit quality $\chi_{\text{DTF}}^2/\text{ndf}$, the B candidate mass $m_{J/\psi K_S^0}$, the mass error estimate of the B candidate $\sigma_{m_{J/\psi K_S^0}}$, the decay-time error estimate of the B candidate σ_t , and the z -position of the PV z_{PV} .

Candidate	Requirement	Unit
J/ψ (LL)	$\min(\chi_{\text{IP,other PV}}^2) > 66$	
J/ψ (DD)	$\min(\chi_{\text{IP,other PV}}^2) > 200$	
K_S^0 (LL)	$m_{p\pi^-}/m_{\pi^+\bar{p}}$	$> 1122 \vee < 1110$
	$\sigma_{m_{\pi^+\pi^-}}$	< 15
	t	> 0
		< 100
K_S^0 (DD)	$m_{p\pi^-}/m_{\pi^+\bar{p}}$	$> 1126 \vee < 1106$
	$\sigma_{m_{\pi^+\pi^-}}$	< 15
	t	> 0
B	$\chi_{\text{DTF}}^2/\text{ndf}$	< 9
	$m_{J/\psi K_S^0}$	> 5180
		< 5520
	$\sigma_{m_{J/\psi K_S^0}}$	< 30
	t	> 0.2
		< 15
	σ_t	< 0.2
Other	z_{PV}	< 250

However, a small fraction can remain in the sample even after the selection. Being associated to a displaced PV, these candidates usually show a very broad decay-time distribution and thus populate high decay-time regions where only a small amount of signal is expected. Therefore, the incorrectly associated candidates can in turn affect the determination of the decay-time resolution and the CP asymmetry measurement. For each candidate-PV pair the χ_{IP}^2 of the J/ψ candidate to all *other* PVs $\chi_{\text{IP,other PV}}^2$ is evaluated. In case the candidate is incorrectly associated, it is likely to show a significantly small $\chi_{\text{IP,other PV}}^2$. Thus, if any $\chi_{\text{IP,other PV}}^2$ is below a threshold, the candidate is considered to be incorrectly associated and neglected. The chosen criterion has the advantage of being independent of any PV ranking, thus allowing to still treat all PVs equally.

Table 4.4 also shows the observed exclusive background contributions on the simulated samples after applying the preparatory requirements. The only remaining visible contribution arises from misreconstructed $B^0 \rightarrow J/\psi K^{*0}$ decays.

4.1.7 Removal of $B^0 \rightarrow J/\psi K^{*0}$ background

Misreconstructed decays from $B^0 \rightarrow J/\psi K^{*0}$ are the dominant non-combinatorial background contribution. As the $K^{*0} \rightarrow K^+ \pi^-$ decay is instantaneous, only the LL sample is affected. Furthermore, the K - π misidentification leads to a reconstructed B candidate mass below the nominal B^0 mass. The mass distribution of B candidates after stripping and preparatory requirements including a fit is shown in Figure 4.2. Misreconstructed $B^0 \rightarrow J/\psi K^{*0}$ decays can approximately be parameterised as a Gaussian distribution. All other components rely on the parameterisation as discussed in Section 5.1.2.

Different strategies how to treat this contribution are assessed. For instance, the misreconstructed decays could be parameterised. This is feasible for the reconstructed mass, but difficult for the other dimensions essential to the analysis without affecting the fit stability. Besides, this parameterisation introduces additional parameters which complicate the measurement and possibly even affect the fit stability. Therefore, the strategy to eliminate the misreconstructed background contribution from the data sample to a level which can be neglected in the parameterisation is chosen instead.

An artificial neural network is utilised as a classifier to identify misreconstructed $B^0 \rightarrow J/\psi K^{*0}$ decays. It is trained using simulated samples of $B^0 \rightarrow J/\psi K_S^0$ decays as signal and $B^0 \rightarrow J/\psi K^{*0}$ decays reconstructed as $B \rightarrow J/\psi K_S^0$ as background representatives. As most misreconstructed $B^0 \rightarrow J/\psi K^{*0}$ decays will be assigned a reconstructed mass to a value below the lower mass requirement of $5180 \text{ MeV}/c^2$ of the stripping, the mass requirement has been weakened to $4000 \text{ MeV}/c^2$ to $5550 \text{ MeV}/c^2$ for the background sample.

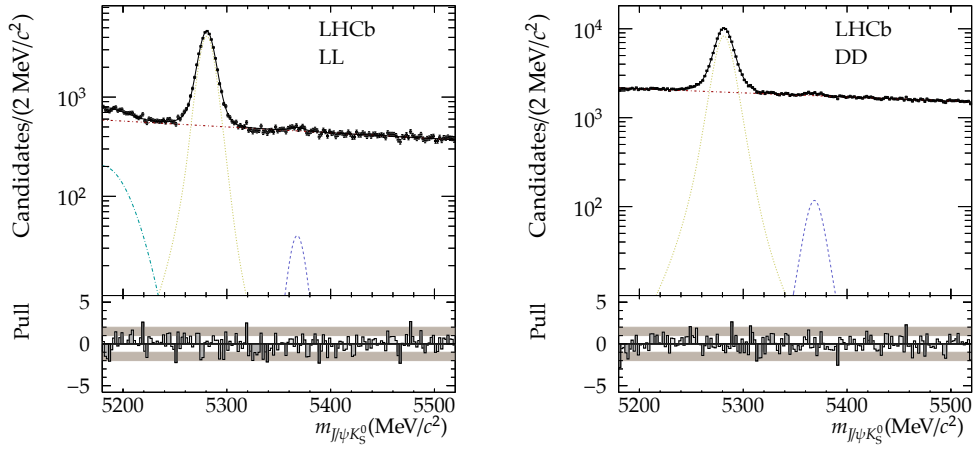


Figure 4.2 – Mass distribution of reconstructed $B \rightarrow J/\psi K_S^0$ candidates after applying stripping and preparatory selection requirements in the LL (left) and DD (right) subsample. The fit (black, solid) shows the $B_s^0 \rightarrow J/\psi K_S^0$ (blue, dashed), the $B^0 \rightarrow J/\psi K_S^0$ (yellow, dotted), and the $B^0 \rightarrow J/\psi K^{*0}$ (turquoise, long-dash-dotted) component as well as the combinatorial background (red, short-dash-dotted).

The input variables used in the neural net exclusively contain properties of the $K_S^0 \rightarrow \pi^+ \pi^-$ decay. These include for the K_S^0 candidate the decay time $t_{K_S^0}$, the reconstructed mass $m_{\pi^+ \pi^-}$, the decay-time significance $t_{K_S^0}/\sigma_{t_{K_S^0}}$, the minimum of $\chi_{\text{IP,PV}}^2$ with respect to any PV, the mass uncertainty estimate $\sigma_{m_{\pi^+ \pi^-}}$, the momentum p , and the transverse momentum p_T . For all pion properties only the smaller or larger value of both pion candidates is used for the respective quantity. These are the larger probability of the candidate to be a kaon provided by the PID system $\mathcal{P}_{\text{NN}}(K)$, the smaller momentum p , the smaller transverse momentum p_T , the smaller minimum of $\chi_{\text{IP,PV}}^2$ with respect to any PV, the smaller minimum IP with respect to any PV IP_{PV} , and the larger track fit quality $\chi_{\text{track}}^2/\text{ndf}$.

The neural network provides a single variable to reject the background, for which the requirement is chosen such that 99 % of the simulated $B^0 \rightarrow J/\psi K_S^0$ decays are retained. Of all set-ups tested for the network training, the one rejecting most of the background at that requirement is used. This network rejects (99.55 ± 0.12) % of the simulated $B^0 \rightarrow J/\psi K^{*0}$ decays. Therefore, this excellent performance allows to suppress the misreconstructed $B^0 \rightarrow J/\psi K^{*0}$ contribution to a level so low it can be neglected (see Figure 4.3). Any $B^0 \rightarrow J/\psi K^{*0}$ decays still present in the sample will be absorbed in the combinatorial background component. More detailed information of the neural network can be found in Ref. [84].

4.1.8 Neural network selection

After applying the stripping, trigger, and preparatory requirements as well as the neural network to suppress misreconstructed background decays, the data sample is still dominated by combinatorial background candidates making a measurement of the CP asymmetries in $B_s^0 \rightarrow J/\psi K_S^0$ infeasible. Therefore, two secondary neural networks, one each for the LL and DD subsample, are used to drastically reduce the background contribution. These networks are trained using the data sample after applying all previous selection requirements. An unbinned maximum likelihood fit to the mass distribution of the B candidates is performed to extract $sWeights$ for the network training [85]. Due to the kinematic similarities between both the $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ decays, the $sWeighted$ $B^0 \rightarrow J/\psi K_S^0$ component is used as signal representative for both $B \rightarrow J/\psi K_S^0$ decays and the combinatorial background component is used as background representative. Decays from the $B_s^0 \rightarrow J/\psi K_S^0$ component are not used in the training by excluding all candidates with a reconstructed mass from $5340 \text{ MeV}/c^2$ to $5390 \text{ MeV}/c^2$. Figure 4.3 shows the distribution of the reconstructed mass before the secondary neural network training and the according fit.

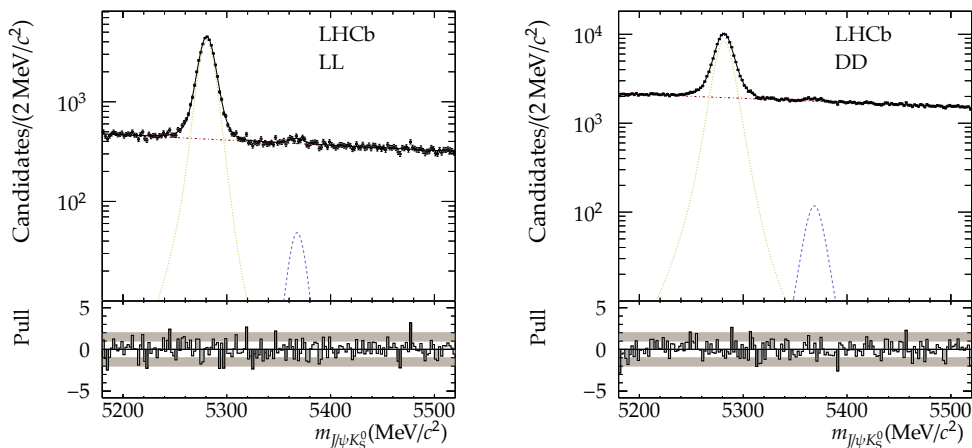


Figure 4.3 – Mass distribution of reconstructed $B \rightarrow J/\psi K_S^0$ candidates after applying the requirement on the neural network to reject misreconstructed $B^0 \rightarrow J/\psi K^{*0}$ decays in the LL (left) and DD (right) subsample. The fit (black, solid) shows the $B_s^0 \rightarrow J/\psi K_S^0$ (blue, dashed) and the $B^0 \rightarrow J/\psi K_S^0$ (yellow, dotted) component as well as the combinatorial background (red, dash-dotted).

Both networks use different sets of input variables, in total 31 for LL and 35 for DD. Naturally, many variables are used in both networks. These sets contain kinematic properties of the B candidate and its daughter candidates from the performed DTF, properties of the underlying event, and information provided by

the PID system. The exact list of variables and more detailed information on the neural network training can be found in Ref. [84].

Like the first neural network, the secondary networks each provide a single output variable. A more stringent requirement on this variable leads to a higher level of background rejection at the cost of an inevitable loss of signal decays. It is not a priori clear which value provides optimal sensitivity for the CP asymmetry measurement. A detailed study shows that in a wide range of the network output no significant dependence of the achievable sensitivity on the output can be found. Therefore, the requirements which maximise $N_S/\sqrt{N_S + N_B}$ are chosen for the measurement, where N_S is the number of $B_s^0 \rightarrow J/\psi K_S^0$ decays present in the selected sample and N_B is the number of combinatorial background candidates observed in a $60 \text{ MeV}/c^2$ wide mass window around the B_s^0 peak. Additional studies are performed to confirm the independence of the neural networks on the reconstructed mass of the B candidates using a simulated sample of inclusive $J/\psi \rightarrow \mu^+ \mu^-$ decays.

4.1.9 Nominal data sample

After applying the stripping, trigger, preparatory, and both neural network requirements, the data sample still contains single events with multiple B candidates and/or multiple PVs, of which only one (B, PV) pair should be used in the measurement. In the LL (DD) sample only 1.5% (0.7%) of the events contain at least two (B, PV) pairs. A detailed examination counting unique and multiple occurrences is shown in Table 4.6. Events with multiple B candidates passing the selection requirements are very rare, most multiple (B, PV) occurrences arise from events with multiple PVs still present in the sample. As mentioned earlier, for all multiple occurrences, a single (B, PV) pair is chosen randomly, leading to the nominal data sample for the measurement.

The distribution of the reconstructed mass of B candidates in the nominal data sample is shown in Figure 4.4 along with a fit to this distribution describing the $B_s^0 \rightarrow J/\psi K_S^0$, $B^0 \rightarrow J/\psi K_S^0$, and combinatorial background components using the parameterisation as discussed in Section 5.1.2. Especially the secondary neural networks provide a crucial level of background rejection. In both the LL and DD samples, the $B_s^0 \rightarrow J/\psi K_S^0$ component is clearly discernible. The fitted numbers of candidates in each component are specified in Table 4.7.

4.2 Flavour tagging studies

The time-dependent CP asymmetry Equation (2.34) in Section 2.4.3 separately requires the time-dependent decay rates of particles produced as B and \bar{B} . Methods

Table 4.6 – Overview of unique and multiple occurrences of (B, PV) pairs, B candidates, and PVs in the LL and DD data samples after applying all selection requirements. In the first row in each group the total number of pairs, candidates, and PVs in the sample is specified, followed by how often these are unique in a given event. Multiple pairs, candidates, and PVs specify the total number of occurrences, not the number of events with multiple occurrences. The last row in each group counts the occurrences with more than two pairs, candidates, and PVs in a given event.

	LL		DD	
total (B, PV) pairs	29250		55239	
unique (B, PV) pairs	28324	96.8 %	54412	98.5 %
multiple (B, PV) pairs	926	3.2 %	827	1.5 %
multiple (B, PV) pairs (> 2 pairs)	56	0.2 %	15	0.0 %
total B candidates	28806		54934	
unique B candidates	28748	99.8 %	54712	99.6 %
multiple B candidates	58	0.2 %	222	0.4 %
multiple B candidates (> 2 candidates)	—		—	
total PVs	29221		55128	
unique PVs	28353	97.0 %	54523	98.9 %
multiple PVs	868	3.0 %	605	1.1 %
multiple PVs (> 2 PVs)	56	0.2 %	15	0.0 %

Table 4.7 – Fitted number of candidates of the $B_s^0 \rightarrow J/\psi K_S^0$, $B^0 \rightarrow J/\psi K_S^0$, and combinatorial background component in the nominal data sample.

	LL	DD
$N_{B_s^0 \rightarrow J/\psi K_S^0}$	309 ± 20	602 ± 31
$N_{B^0 \rightarrow J/\psi K_S^0}$	$27\,790 \pm 170$	$51\,320 \pm 230$
N_{comb}	670 ± 40	2880 ± 80

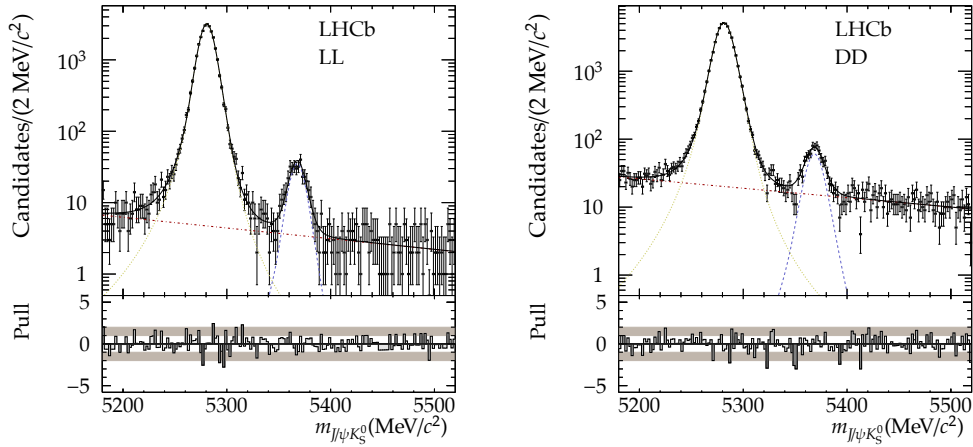


Figure 4.4 – Mass distribution of reconstructed $B \rightarrow J/\psi K_S^0$ candidates after applying all selection requirements and removing multiple candidates and PVs in the LL (left) and DD (right) subsample. The fit (black, solid) shows the $B_s^0 \rightarrow J/\psi K_S^0$ (blue, dashed) and the $B^0 \rightarrow J/\psi K_S^0$ (yellow, dotted) component as well as the combinatorial background (red, dash-dotted).

to infer this production flavour for each candidate are called flavour tagging algorithms.

4.2.1 Principles of the flavour tagging at LHCb

Figure 4.5 shows the tagging algorithms employed for this measurement. For every reconstructed B meson, each algorithm evaluates the event and determines an individual tag decision. These decisions are then transformed into a single combined decision by a method which is described later on. The taggers are divided into two classes: opposite-side (OS) and same-side (SS) taggers.

The OS taggers exploit the fact that b quarks are predominantly produced in pairs as $b\bar{b}$ at the LHC [86]. While one b quark (signal b) forms the reconstructed signal B meson (here: B_s^0 or B^0), the other b quark hadronises independently into a hadron h_b (e.g. B^\pm , B^0/\bar{B}^0 , $\Lambda_b^0/\bar{\Lambda}_b^0$). The OS taggers search for decay particles of h_b to deduce its production flavour, which then in turn allows to identify the signal B production state. Four distinct OS taggers are available: The OS kaon tagger identifies charged kaons from subsequent $b \rightarrow c \rightarrow s$ decays, while the OS muon and electron taggers search for μ^\pm and e^\pm from semileptonic h_b decays. While the former algorithms are single-particle taggers, the OS vertex charge tagger aims at determining the inclusive charge of the h_b decay vertex. To achieve that, the tagger selects two tracks with the highest probability of originating from the h_b vertex as a seed. Additional tracks being compatible with originating from the

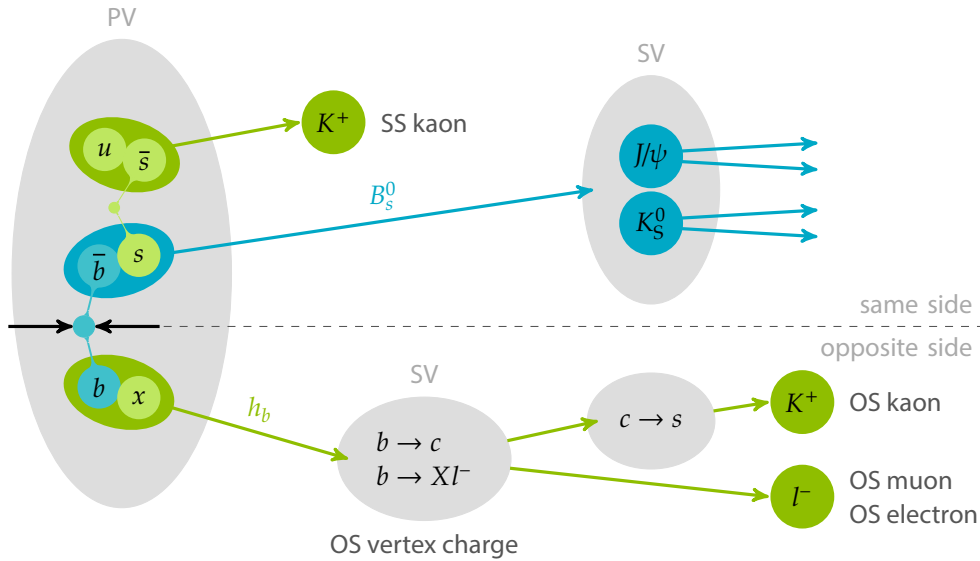


Figure 4.5 – Overview of flavour tagging algorithms for $B_s^0 \rightarrow J/\psi K_S^0$ at LHCb. The same-side kaon tagger (top) searches for remnants of the signal B_s^0 hadronisation, while the opposite-side taggers (bottom) infer the B_s^0 production flavour from the decay of the accompanying h_b hadron.

same vertex but not the PV are added to this and finally the inclusive secondary vertex charge is evaluated to determine the tag decision.

The criteria to select tagging particles for the OS taggers contain, amongst others, track fit quality, polar angle with respect to the beam axis, momentum, transverse momentum, and PID requirements. Additionally, the particle tracks are required to be well separated from other PVs and to be located outside a cone of 5 mrad around any of the daughter tracks of the signal decay to avoid mistaking signal particles for OS tagging particles. The OS taggers are developed and studied utilising simulated events and re-tuned on the control channels $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$ with data samples.

For the OS kaon tagger, two variants are available, which differ in the selection of tagging particles. This analysis uses the cut-based variant with a selection based on rectangular requirements. The more recent neural-net OS kaon tagger uses a neural network for that purpose [87, 88]. Recently, the OS charm tagger has been developed, which was not ready at the time this analysis was performed [89].

The second class of tagging algorithms are the SS taggers. While the OS taggers are independent of the type of signal B meson, the SS algorithms search for remnants of the B meson hadronisation and are specific for the type of B meson of interest. For a B_s^0 (\bar{B}_s^0) meson, the additional \bar{s} (s) quark produced in the fragmentation often forms a K^+ (K^-) meson. By searching for such kaons, the SSK tagger aims at identifying the B_s^0 meson production flavour. Two variants of the SSK

tagger are implemented at LHCb, which differ in the way the tagging particle is selected. The cut-based SSK tagger uses a selection based on rectangular cuts [90], while the more recent neural-net SSK tagger features a neural net for this task [87, 88, 91]. Selection requirements for both variants contain similar variables as for the OS taggers and additionally requirements to separate K^\pm candidates from kaons of the opposite-side or the underlying event. Similar to the OS taggers, the SSK taggers are developed utilising simulated samples and re-tuned and optimised with data samples using $B_s^0 \rightarrow D_s^- \pi^+$ decays.

In case of B^0 mesons, two SS taggers can be used which work in analogy to the SSK tagger. The same-side pion tagger searches for π^\pm mesons correlated with the B^0 production, while the same-side proton tries to identify p or \bar{p} being produced alongside the B^0 .

As mentioned above, each tagging algorithm provides a tag decision d for the initial flavour with

$$d = \begin{cases} +1 & \text{for } B_s^0/B^0, \\ -1 & \text{for } \bar{B}_s^0/\bar{B}^0, \\ 0 & \text{untagged.} \end{cases} \quad (4.1)$$

If the B meson production flavour cannot be inferred or if the determination is considered to be too unreliable, the tagger will output a tag of zero. As all taggers are imperfect, a fraction of non-zero tag decisions will be incorrect. Reasons for these so-called mistags can be misreconstruction effects like the selection of wrongly assigned tracks or flavour oscillations of an OS B meson. In order to assess the confidence in the tag decision, each algorithm also provides an estimate η for the probability of the tag to be incorrect. This prediction is derived by means of neural networks.

4.2.2 Tagging quantities

A set of quantities defines the performance of a flavour tagging algorithm. The tagging efficiency ε_{tag} is defined as the ratio of tagged candidates N_{tag} to all reconstructed candidates N ,

$$\varepsilon_{\text{tag}} = \frac{N_{\text{tag}}}{N} = \frac{N_{\text{R}} + N_{\text{W}}}{N_{\text{R}} + N_{\text{W}} + N_{\text{U}}}. \quad (4.2)$$

Here, N_{R} is the number of correctly tagged candidates, N_{W} the number of mis-tagged candidates, and N_{U} the number of candidates without a tag decision. The mistag fraction ω is defined as

$$\omega = \frac{N_{\text{W}}}{N_{\text{tag}}} \quad (4.3)$$

and describes the fraction of incorrectly tagged candidates. It can also be interpreted on a per-candidate basis as the probability of the tag decision to be incorrect.

Both ε_{tag} and ω affect the numbers $N_{\bar{B}}(t)$ and $N_B(t)$ of reconstructed candidates which are tagged as \bar{B} and B as

$$N_{\bar{B}}(t) = \varepsilon_{\text{tag}}(1 - \omega)N_B^{\text{true}}(t) + \varepsilon_{\text{tag}}\omega N_{\bar{B}}^{\text{true}}(t), \quad (4.4a)$$

$$N_B(t) = \varepsilon_{\text{tag}}(1 - \omega)N_{\bar{B}}^{\text{true}}(t) + \varepsilon_{\text{tag}}\omega N_B^{\text{true}}(t), \quad (4.4b)$$

where $N_{\bar{B}}^{\text{true}}(t)$ and $N_B^{\text{true}}(t)$ are the true time-dependent rates of mesons being produced as \bar{B} and B . This results in the observed CP asymmetry $\mathcal{A}_{\text{obs}}(t)$ being diluted by the factor $\mathcal{D}_{\text{tag}} = (1 - 2\omega)$ with respect to the true asymmetry $\mathcal{A}(t)$:

$$\mathcal{A}_{\text{obs}}(t) = \mathcal{D}_{\text{tag}}\mathcal{A}(t). \quad (4.5)$$

With these considerations it can be shown that the figure of merit for evaluating the performance of a flavour tagging algorithm is the effective tagging efficiency or tagging power

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}\mathcal{D}_{\text{tag}}^2 = \varepsilon_{\text{tag}}(1 - 2\omega)^2. \quad (4.6)$$

The statistical uncertainty on $\mathcal{A}_{\text{obs}}(t)$ directly scales with $1/\sqrt{\varepsilon_{\text{eff}}}$. In other words, the factor ε_{eff} denotes the fraction of perfectly tagged (i.e. $\omega = 0$) candidates which yield the same statistical uncertainty on the CP asymmetries as the available imperfectly tagged sample.

Hence, two things are crucial for a time-dependent measurement of CP violation: The flavour tagging algorithms should be optimised with respect to enlarging the tagging power ε_{eff} . Furthermore, as can be seen in Equation (4.5), the mistag fraction ω needs to be precisely known to separate the tagging dilution from the physical quantities of interest. Ideally, for each reconstructed candidate the mistag prediction η reflects the true mistag probability ω . As this is not necessarily the case, the taggers need to be calibrated. This procedure is described in the following section.

4.2.3 Combination and calibration of tagging algorithms

As explained above, for a given candidate each individual tagging algorithm i provides an individual tag decision d_i and mistag prediction η_i . It is therefore convenient to combine the taggers to profit from smaller mistag fractions and thus improved statistical sensitivity on the CP asymmetries. For this combination, the combined probability $P(b)$ ($P(\bar{b})$) of the meson candidate to contain a b (\bar{b}) quark is calculated as

$$P(b) = \frac{p(b)}{p(b) + p(\bar{b})} \quad \text{and} \quad P(\bar{b}) = 1 - P(b), \quad (4.7)$$

with

$$p(b) = \prod_i \left(\frac{1 + d_i}{2} - d_i(1 - \eta_i) \right) \quad \text{and} \quad p(\bar{b}) = \prod_i \left(\frac{1 - d_i}{2} + d_i(1 - \eta_i) \right). \quad (4.8)$$

The combined tag d and mistag prediction η is then derived as

$$d = \begin{cases} +1 & \text{for } P(b) \leq P(\bar{b}), \\ -1 & \text{for } P(b) > P(\bar{b}), \end{cases} \quad \text{and} \quad \eta = \begin{cases} 1 - P(\bar{b}) & \text{for } P(b) \leq P(\bar{b}), \\ 1 - P(b) & \text{for } P(b) > P(\bar{b}). \end{cases} \quad (4.9)$$

For both individual taggers as well as for combined tagging decisions, the mistag probability prediction η is evaluated on a candidate-by-candidate basis. This is not necessarily equivalent to the per-candidate mistag probability ω . Consequently, a calibration function $\omega(\eta)$ needs to be chosen to correct for this for single taggers as well as combined taggers alike. Usually, a linear dependence is found and thus $\omega(\eta)$ is parameterised as

$$\omega(\eta) = p_0 + p_1(\eta - \langle\eta\rangle), \quad (4.10)$$

where p_0 and p_1 are the calibration parameters and $\langle\eta\rangle$ is the average mistag prediction of all signal candidates. The parameters p_0 and p_1 are determined in calibration fits, which are performed on samples of suitable decay channels for the specific tagger or combination. In these control channels the charges of the final state particles allow to identify the B meson flavour and thus verify the tagging decision and mistag prediction. One exemplary control channel is $B^+ \rightarrow J/\psi K^+$ where the charge of the daughter K identifies the B^\pm flavour. In case of $p_0 = \langle\eta\rangle$ and $p_1 = 1$, the mistag prediction does not require calibration, i.e. $\omega = \eta$.

The procedure for the OS taggers is as follows: Each single tagger is first calibrated with decays of $B^+ \rightarrow J/\psi K^+$ using an unbinned maximum likelihood fit to the candidate mass m , tagging decision d , and mistag prediction η [86]. The calibration functions for the OS taggers show only little deviation from the expectation of $p_0 = \langle\eta\rangle$ and $p_1 = 1$. In the next step, all OS taggers are merged into the OS combination. To limit the contribution of poor tags, the mistag probabilities of the OS kaon and vertex charge tagger are required to be less than 0.46. As the combination neglects the correlation between the taggers, the combined η is expected to be slightly biased. To account for that, the combined OS tagging output is again calibrated on samples of the control channels $B^+ \rightarrow J/\psi K^+$, $B^+ \rightarrow D^0 \pi^+$, $B^0 \rightarrow J/\psi K^{*0}$, $B^0 \rightarrow D^{*-} \mu^+ \nu_\mu$, and $B_s^0 \rightarrow D_s^- \pi^+$.

The two variants of the SSK taggers are calibrated on a sample of $B_s^0 \rightarrow D_s^- \pi^+$ decays [90]. As there are no other feasible control channels available, the calibration is verified on simulated samples of $B_s^0 \rightarrow D_s^- \pi^+$, $B_s^0 \rightarrow J/\psi \phi$, and $B_s^0 \rightarrow K^+ K^-$. As only one of the two variants is used in a physics analysis, no combination is being performed.

4.2.4 Flavour tagging in $B_s^0 \rightarrow J/\psi K_s^0$

In order to achieve the optimal sensitivity on the CP asymmetries in the $B_s^0 \rightarrow J/\psi K_s^0$ channel, this analysis utilises both the standard OS tagging combination as well as

the cut-based SSK tagger variant. On the opposite-side this involves the cut-based OS kaon, the OS muon and electron as well as the OS vertex charge taggers.

The analysed sample contains a large quantity of $B^0 \rightarrow J/\psi K_S^0$ decays. As the SSK tagger is not designed for B^0 decays, its response for these candidates has to be properly understood in order to model it in the measurement. In fact, on average it shows a non-random tag response for B^0 candidates with a preference to yield the incorrect tag decision. That means it can be exploited as a B^0 tagger by inverting the tag d and using dedicated calibration parameters for the $B^0 \rightarrow J/\psi K_S^0$ component of the measurement. Studies on the non-negligible tagging power of the SSK on B^0 candidates are presented in Section 4.2.7.

The sample is divided into each three subsamples for LL and DD. Candidates with only OS tagging response are called exclusively OS tagged. For these, the mistag predictions η^{OS} are calibrated in the fit and the OS tag decisions d^{OS} and mistag probabilities $\omega^{\text{OS}}(\eta^{\text{OS}})$ are used for the measurement. The second subsample contains all SSK tagged candidates. These are composed of both the exclusively SSK tagged candidates and all candidates with both an OS and SSK response (overlap subsample). The SSK mistag predictions η^{SSK} are calibrated into $\omega^{\text{SSK}, B_S^0}(\eta^{\text{SSK}})$ and $\omega^{\text{SSK}, B^0}(\eta^{\text{SSK}})$ with the respective B_S^0 and B^0 calibration functions, while the SSK tag decisions for the B^0 component are inverted in the fit. Where applicable, the OS and SSK tagging responses are combined following Equations (4.7) to (4.9) after applying the respective calibration function. The last subsample consists of the untagged events without any tagging response.

4.2.5 Calibration of the opposite-side taggers

The calibration of the combined output of the OS tagging algorithms is determined using different control channels by combining the individual calibration parameters into one universally valid set. The applied calibration function is

$$\omega^{\text{OS}}(\eta^{\text{OS}}) = p_0^{\text{OS}} + p_1^{\text{OS}}(\eta^{\text{OS}} - \langle \eta^{\text{OS}} \rangle). \quad (4.11)$$

Due to the chosen method, p_0^{OS} is determined using

$$p_0^{\text{OS}} - \langle \eta^{\text{OS}} \rangle = 0.0062 \pm 0.0019 \text{ (stat)} \pm 0.0040 \text{ (syst)} \quad (4.12)$$

from the combined calibration and $\langle \eta^{\text{OS}} \rangle$ as measured on an sWeighted sample of $B^0 \rightarrow J/\psi K_S^0$ events. The resulting calibration parameters are

$$\begin{aligned} p_0^{\text{OS,LL}} &= 0.3811 \pm 0.0019 \text{ (stat)} \pm 0.0040 \text{ (syst)}, \\ p_0^{\text{OS,DD}} &= 0.3801 \pm 0.0019 \text{ (stat)} \pm 0.0040 \text{ (syst)}, \\ p_1^{\text{OS}} &= 0.982 \pm 0.007 \text{ (stat)} \pm 0.034 \text{ (syst)}, \\ \langle \eta^{\text{OS,LL}} \rangle &= 0.3749, \\ \langle \eta^{\text{OS,DD}} \rangle &= 0.3739. \end{aligned} \quad (4.13)$$

4 Analysis ingredients and preparatory studies

As $\langle \eta^{\text{OS}} \rangle$ is determined separately for the LL and DD samples, individual calibration functions are assigned accordingly. The statistical uncertainties result from the combination, while the systematic uncertainties contain both systematic uncertainties of the individual calibration measurements on a single control channel as well as those accounting for differences between the control channels and physics channels of interest [86].

Although the tagging algorithms are designed to perform equally well for B and \bar{B} , differences in the tagging efficiency and the calibration parameters can occur. For the OS combination, these are determined as

$$\begin{aligned}\Delta \varepsilon_{\text{tag}}^{\text{OS}} &= \varepsilon_{\text{tag}}^{\text{OS},\bar{B}} - \varepsilon_{\text{tag}}^{\text{OS},B} = -0.0008 \pm 0.0007, \\ \frac{\Delta p_0^{\text{OS}}}{2} &= \frac{1}{2}(p_0^{\text{OS},B} - p_0^{\text{OS},\bar{B}}) = +0.0070 \pm 0.0006, \\ \frac{\Delta p_1^{\text{OS}}}{2} &= \frac{1}{2}(p_1^{\text{OS},B} - p_1^{\text{OS},\bar{B}}) = +0.033 \pm 0.006,\end{aligned}\tag{4.14}$$

with $\varepsilon_{\text{tag}}^{\text{OS},\bar{B}}$, $p_0^{\text{OS},\bar{B}}$ and $p_1^{\text{OS},\bar{B}}$ ($\varepsilon_{\text{tag}}^{\text{OS},B}$, $p_0^{\text{OS},B}$ and $p_1^{\text{OS},B}$) being the tagging efficiency and calibration parameters for particles being produced as \bar{B} (B).

The OS calibration parameters as quoted in Equation (4.13) are used for the measurement of the CP observables to transform the mistag estimates η^{OS} into mistag probabilities ω^{OS} . The uncertainties on p_0^{OS} and p_1^{OS} and the asymmetries $\Delta p_0^{\text{OS}}/2$ and $\Delta p_1^{\text{OS}}/2$ are treated as a systematic uncertainty for the measurement of CP violation (see Section 5.4.2). As the tagging efficiency difference $\Delta \varepsilon_{\text{tag}}^{\text{OS}}$ is not significant and its influence on the CP violation measurement deemed to be marginal, this parameter is neglected altogether.

4.2.6 Calibration of the same-side kaon tagger

The calibration of the SSK tagger is performed similarly to the OS combination. For B_s^0 events, the control channel $B_s^0 \rightarrow D_s^- \pi^+$ is used to determine the calibration with the function

$$\omega^{\text{SSK},B_s^0}(\eta^{\text{SSK}}) = p_0^{\text{SSK},B_s^0} + p_1^{\text{SSK},B_s^0}(\eta^{\text{SSK}} - \langle \eta^{\text{SSK},B_s^0} \rangle).\tag{4.15}$$

The extracted parameters are:

$$\begin{aligned}p_0^{\text{SSK},B_s^0} &= 0.347 \pm 0.009 \text{ (stat)} \pm 0.005 \text{ (syst)}, \\ p_1^{\text{SSK},B_s^0} &= 1.715 \pm 0.020 \text{ (stat)} \pm 0.057 \text{ (syst)}, \\ \langle \eta^{\text{SSK},B_s^0} \rangle &= 0.3423.\end{aligned}\tag{4.16}$$

In contrast to the OS combination, the SSK calibration parameters show a clear deviation from the expectation of $p_1^{\text{SSK},B_s^0} = 1$. Nevertheless, a linear dependence

between $\omega^{\text{SSK},B_s^0}$ and η^{SSK} is found justifying the usage of the cut-based SSK tagger. Due to p_1^{SSK,B_s^0} not being consistent with one, the value of $\langle \eta^{\text{SSK},B_s^0} \rangle$ from the calibration channel is used. Systematic uncertainties on p_0^{SSK,B_s^0} and p_1^{SSK,B_s^0} are determined analogously as for the OS combination [90].

At the time this analysis is performed, no tagging-related asymmetries are available for the SSK tagger. The calibration parameters in Equation (4.16) are used to determine $\omega^{\text{SSK},B_s^0}$ for the description of the SSK-tagged events in the $B_s^0 \rightarrow J/\psi K_S^0$ component. Like for the OS combination, the uncertainties on p_0^{SSK,B_s^0} and p_1^{SSK,B_s^0} are treated as systematic uncertainty (see Section 5.4.2).

As mentioned before, the SSK tagger shows a non-negligible tagging power for B^0 decays if the tagging decision is inverted, i.e. $d^{\text{SSK},B^0} = -1 \cdot d^{\text{SSK}}$. Therefore, a dedicated calibration is being performed on $B^0 \rightarrow J/\psi K^{*0}$ decays with this tag flip. The data is divided into equally populated bins of η^{SSK} and the respective ω^{SSK,B^0} is determined for each bin by fitting the decay-time distribution with a model describing the $B^0-\bar{B}^0$ flavour oscillation. Afterwards, the linear function

$$\omega^{\text{SSK},B^0}(\eta^{\text{SSK}}) = p_0^{\text{SSK},B^0} + p_1^{\text{SSK},B^0}(\eta^{\text{SSK}} - \langle \eta^{\text{SSK},B^0} \rangle) \quad (4.17)$$

is fitted to the $(\eta^{\text{SSK},B^0}, \omega^{\text{SSK},B^0})$ pairs resulting in the calibration parameters

$$\begin{aligned} p_0^{\text{SSK},B^0} &= 0.459 \pm 0.003 \text{ (stat)}, \\ p_1^{\text{SSK},B^0} &= 0.493 \pm 0.082 \text{ (stat)}, \\ \langle \eta^{\text{SSK},B^0} \rangle &= 0.350. \end{aligned} \quad (4.18)$$

Figure 4.6 shows the data points and calibration function. Additionally, a quadratic fit is performed as a cross-check. The resulting $\omega^{\text{SSK},B^0}(\eta^{\text{SSK}})$ clearly shows the weak tagging performance. Nevertheless, the linear function describes the data well and thus it is used to parameterise the mistag probability for the $B^0 \rightarrow J/\psi K_S^0$ component in the fit. As the $B^0 \rightarrow J/\psi K_S^0$ decays mainly serve as a cross-check for the CP violation measurement in $B_s^0 \rightarrow J/\psi K_S^0$ and as the contribution of the SSK tagging for the $B^0 \rightarrow J/\psi K_S^0$ component is minor, no systematic uncertainties are determined for this calibration.

4.2.7 Studies on the same-side kaon tagger on B^0 candidates

In order to investigate which effects contribute to the SSK tagging power for B^0 candidates, a dedicated study on a sample of six million simulated $B^0 \rightarrow J/\psi K^{*0}$ events is performed [14]. For this study, the identity of the generated particle which was used as tagging candidate by the SSK and—where available—the identities of all its ancestor particles are examined in case the ancestors contain the SS or

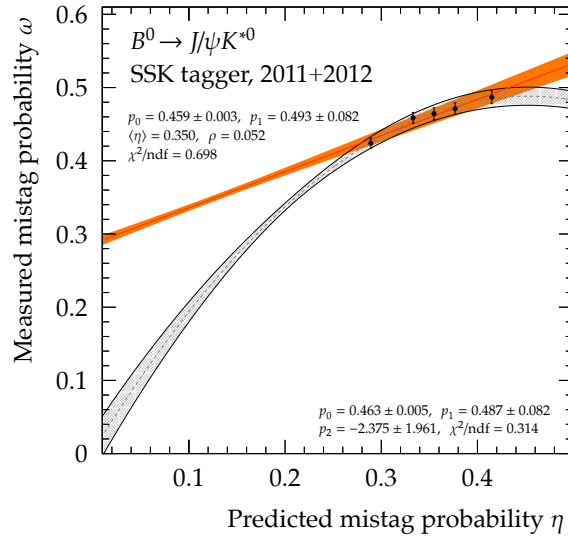


Figure 4.6 – Calibration of the cut-based SSK tagger using $B^0 \rightarrow J/\psi K^{*0}$ decays after applying a flipped tag decision. Results of the linear (orange, solid) and quadratic (grey, dashed) fits are shown together with the respective error contour.

OS b quark. In the Monte Carlo terminology this means that the particle is being produced in direct relation with the hadronisation process of the b quark.

The mistag probability ω is evaluated as defined in Equation (4.3) for each type of characteristic decay chain for particles being related with the SS or OS b . Additionally, it is relevant how often these decay chains are selected by the SSK tagger. Thus, the fraction of these events to the number $N_{\text{SSK, assoc.}}$ of all SSK-tagged events with an association of the tagging candidate to a generated particle is defined as

$$\varepsilon'_{\text{tag}} = \frac{N_{\text{R}} + N_{\text{W}}}{N_{\text{SSK, assoc.}}} . \quad (4.19)$$

The three main sources of B^0 tagging power of the SSK tagger are identified as:

1. Misidentified π^\pm from the SS. These carry the same charge as the K^\pm from a corresponding B_s^0 and thus no tag flip is required. The analysed sample contains 2356 π^\pm from the SS with a mistag of $\omega = (33.2 \pm 1.0)\%$ and $\varepsilon'_{\text{tag}} = 5.3\%$.
2. Correctly identified K^\pm from the SS. These originate from subsequent decays like $\bar{K}^{*0} \rightarrow K^- \pi^+$ decays, where the \bar{K}^{*0} is produced in correlation with the hadronisation of a B^0 meson. Consequently, the K charge is opposite to the expectation for the B_s^0 case and thus a tag flip is required in these cases. The analysed sample contains 9722 K^\pm from the SS with a mistag of $\omega = (61.1 \pm 0.5)\%$ and $\varepsilon'_{\text{tag}} = 21.7\%$.

3. Misidentified protons from the SS. Analogously, the proton charge is opposite to the charge of an SS K for a B_s^0 , requiring a tag flip in these cases as well. The analysed sample contains 2438 protons from the SS with a mistag of $\omega = (59.5 \pm 1.0) \%$ and $\varepsilon'_{\text{tag}} = 5.4 \%$.

Additionally, K^\pm from the OS are occasionally utilised by the SSK tagger, which on average require a flip of the tag decision. In the analysed sample, 3711 of OS K^\pm are used with $\omega = (57.5 \pm 0.8) \%$ and $\varepsilon'_{\text{tag}} = 8.3 \%$. The remaining MC-associated particle sources show no significant non-random tagging response.

The effects resulting in the correct tag decision compete with those requiring a tag flip. As the contribution to the overall tagging power of the latter dominates, in total a tag flip is required to utilise the SSK tagger for B^0 mesons.

4.2.8 Tagging performance

The tagging performance is evaluated individually for both the $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ component using sWeights samples. To obtain these weights, a fit to the mass distribution is performed using the nominal parameterisation as described in Section 5.1.2. The effective tagging efficiency ε_{eff} is then determined using the per-candidate mistag $\omega_i = \omega(\eta_i)$ as

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}(1 - 2\omega)^2 = \frac{N_{\text{tag}}}{N} \sum_i \frac{s_i}{N_{\text{tag}}} (1 - 2\omega_i)^2 = \frac{1}{N} \sum_i s_i (1 - 2\omega_i)^2 \quad (4.20)$$

with the sWeights s_i and sum of numbers of both tagged and untagged signal candidates N . Table 4.8 lists the tagging power for both components in the LL and DD subsamples. To break the individual contributions of both tagging classes and their respective overlap down, ε_{eff} is evaluated for exclusively tagged candidates as well as for those with both an OS and SSK response. Uncertainties are obtained by propagating the statistical uncertainties on N and the calibration parameters p_0 and p_1 into the uncertainty on ε_{eff} . The weak tagging power of the SSK tagger on $B^0 \rightarrow J/\psi K_S^0$ candidates is evident.

The tagging efficiency ε_{tag} listed in Table 4.9 shows the relative size of the tagging categories. As the OS response is a combination of multiple individual algorithms, it forms the largest fraction, while only few events share an OS and SSK response. The average mistag probability ω in the various subsamples is determined using sWeights as

$$\omega = \frac{1}{N_{\text{tag}}} \sum_i s_i \omega_i. \quad (4.21)$$

Results for the OS combination and SSK tagger are quoted in Table 4.10 for both exclusively tagged candidates and candidates in the overlap subsample. Again,

4 Analysis ingredients and preparatory studies

Table 4.8 – Effective tagging efficiencies $\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}(1 - 2\omega)^2$ in the subsamples used in the analysis. Uncertainties are statistical.

Sample	$B_s^0 \rightarrow J/\psi K_S^0$		$B^0 \rightarrow J/\psi K_S^0$	
	LL (%)	DD (%)	LL (%)	DD (%)
OS exclusive	2.51 ± 0.15	2.48 ± 0.11	2.28 ± 0.05	2.23 ± 0.05
SSK exclusive	0.27 ± 0.04	1.08 ± 0.11	0.042 ± 0.006	0.064 ± 0.009
OS+SSK overlap	1.02 ± 0.10	0.472 ± 0.035	0.274 ± 0.008	0.327 ± 0.011
Total	3.80 ± 0.18	4.03 ± 0.16	2.60 ± 0.05	2.63 ± 0.05

Table 4.9 – Tagging efficiencies ε_{tag} on the subsamples used in the analysis. Uncertainties are statistical.

Sample	$B_s^0 \rightarrow J/\psi K_S^0$		$B^0 \rightarrow J/\psi K_S^0$	
	LL (%)	DD (%)	LL (%)	DD (%)
OS exclusive	23.8 ± 2.4	27.6 ± 1.8	26.43 ± 0.26	25.56 ± 0.19
SSK exclusive	4.1 ± 1.1	8.5 ± 1.1	5.67 ± 0.14	6.92 ± 0.11
OS+SSK overlap	6.8 ± 1.4	3.3 ± 0.7	3.17 ± 0.11	3.78 ± 0.08

the weak tagging performance of the SSK tagger on B^0 candidates can clearly be observed.

As mentioned above, the analysis utilises the cut-based SSK tagger, while a more recent version based on a neural net for the tagging particle selection is available. For comparison, the tagging performance of the neural-net SSK tagger is evaluated in Appendix A.1. No significant increase in effective tagging efficiency is found with respect to the cut-based SSK tagger.

Table 4.10 – Average mistag probability ω for both the OS and cut-based SSK tagger for both the respective exclusively tagged subsamples as well as the overlap sample.

Sample	$B_s^0 \rightarrow J/\psi K_S^0$		$B^0 \rightarrow J/\psi K_S^0$	
	LL (%)	DD (%)	LL (%)	DD (%)
OS exclusive	37.1 ± 1.1	37.8 ± 0.7	37.97 ± 0.10	37.86 ± 0.07
OS overlap	37.2 ± 1.8	38.8 ± 1.7	38.63 ± 0.29	38.85 ± 0.19
SSK exclusive	38.1 ± 1.4	33.5 ± 0.9	46.13 ± 0.05	45.651 ± 0.035
SSK overlap	35.8 ± 1.6	34.2 ± 1.3	46.17 ± 0.06	45.77 ± 0.05

4.3 Cubic splines as a tool for empirical parameterisation

The description of distributions should rely on physically motivated parameterisations wherever possible. However, some effects are not straight-forward to describe and thus require empirical models. For this task, cubic splines are a convenient solution. These are composed of piecewise cubic polynomials to a function which is continuously differentiable up to the second derivative. Thus, cubic splines are able to provide a smooth approximation for almost any line shape. Cubic B-splines, a special representation of cubic splines, are used in this measurement.

The model relies on a set of interval boundaries or knots which are not necessarily uniformly distributed. Through these, the B-splines $b_i(x)$ are uniquely defined so that between each two boundaries four B-splines are non-zero. All $b_i(x)$ form a partition of unity:

$$\sum_i b_i(x) = 1. \quad (4.22)$$

By assigning each $b_i(x)$ a different weight c_i , an efficiency function $\varepsilon(x)$ can be modelled as

$$\varepsilon(x) = \sum_i c_i b_i(x). \quad (4.23)$$

In case of using n knots, $n + 2$ base splines and thus coefficients c_i are defined. The latter can then be varied and determined in a fit. Typically, the efficiency is multiplied with a given PDF so that its shape can be manipulated. However, cubic splines can as well be used to model distributions directly. An example function for $\varepsilon(x)$ along with the individual base splines is shown in Figure 4.7.

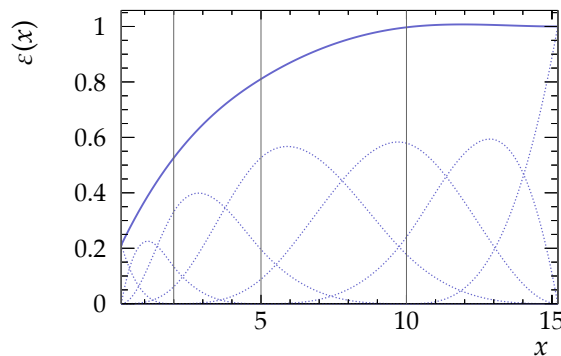


Figure 4.7 – Efficiency function $\varepsilon(x)$ (blue, solid) using cubic B-splines. Boundaries are defined at 0.2, 2.0, 5.0, 10.0 and 15.2 (grey vertical lines). Base splines $b_i(x)$ (blue, dotted) are already scaled by their individual weight c_i so that $\varepsilon(x)$ results as the sum of all individual base splines.

The choice of interval boundaries is usually arbitrary and no trivial strategy to find the optimal configuration with the lowest degrees of freedom exists. Therefore, the positioning of knots itself is subject to empirical studies as well.

4.4 Decay-time resolution

The LHCb detector offers a finite vertex and momentum resolution. This translates directly into a finite decay-time resolution, which is approximately 45 fs at LHCb. As the oscillation period of a B_s^0 meson is $2\pi/\Delta m_s \approx 350$ fs, this resolution will dilute the observed oscillation similar to the dilution from imperfect flavour tagging. Therefore, to measure the CP violating asymmetries, it is essential for the analysis to precisely determine the decay-time resolution and incorporate the description into the fit.

4.4.1 Parameterisation of the resolution

The resolution model $R(t - t')$ describes the deviation of the measured decay time t from the true decay time t' . This model is convolved with the PDF of the true decay time $P(t')$ to determine the PDF of the measured decay time $P(t)$. The DTF used to determine t provides a per-candidate estimate σ_t of the decay time uncertainty which is incorporated in $R(t - t')$. This improves the description and thus increases the sensitivity on the CP asymmetries.

It can be shown that the finite decay-time resolution is comprised of different effects making a non-trivial description necessary. A sum of three Gaussian distributions with common mean μ_r and different widths $\delta_i(\sigma_t)$ is chosen in this analysis as

$$R(t - t' | \sigma_t) \propto \sum_{i=1}^3 f_{r,i} \frac{1}{\sqrt{2\pi}\delta_i(\sigma_t)} \exp\left(-\frac{(t - t' - \mu_r)^2}{2\delta_i^2(\sigma_t)}\right), \quad (4.24)$$

where $f_{r,i}$ denote the fraction of each Gaussian function, where the second is chosen as $f_{r,2} = 1 - f_{r,1} - f_{r,3}$. The widths $\delta_i(\sigma_t)$ account for different resolution effects and are parameterised using scale factors as

$$\delta_1(\sigma_t) = s_r \sigma_t, \quad \delta_2(\sigma_t) = r_{r,2} s_r \sigma_t, \quad \delta_3(\sigma_t) = r_{r,3} s_r \sigma_t, \quad (4.25)$$

where s_r is an absolute scale factor and $r_{r,2}$ and $r_{r,3}$ are relative scale factors for the wider Gaussian functions. To allow for possible differences in the resolution of the LL and DD candidates, all parameters are determined individually for both subsamples.

Due to a technical limitation it is not possible to use a non-zero resolution mean μ_r in a set-up where cubic splines are used as decay-time acceptance function. As

the nominal fit requires the usage of spline functions, the resolution mean will be fixed to zero in the fit for the CP asymmetries.

4.4.2 Calibration of the resolution parameters

The decay-time resolution of $B \rightarrow J/\psi K_S^0$ candidates is dominated by the $J/\psi \rightarrow \mu^+ \mu^-$ decay as the muon tracks define the B decay vertex while the K_S^0 decay vertex is displaced from the B decay due to the longevity of the K_S^0 . Therefore, a sample of non-signal B candidates containing true J/ψ decays shows the same resolution as $B \rightarrow J/\psi K_S^0$ signal decays. Accordingly, prompt background candidates with $J/\psi \rightarrow \mu^+ \mu^-$ decays can be utilised to analyse the decay-time resolution. These candidates are constructed from tracks originating directly from the PV, i.e. from the primary pp interaction or instantly decaying resonances. Consequently, the true decay time of the according B candidate will be $t' = 0$ so that the measured decay-time distribution will form a smeared peak around $t = 0$, reflecting the resolution function. A fit of $R(t - t')$ to this peak can thus determine the resolution parameters.

The nominal data sample does not contain prompt candidates due to the requirement on the decay time of $t > 0.2$ ps. Therefore, the prescaled dataset is used, which does not impose this restriction. As this sample is selected with the same efficient selection criteria as the nominal detached sample, it is both very pure with respect to the J/ψ daughter candidates and additionally only contains signal-like $\pi^+ \pi^-$ background candidates. Thus, this sample is ideal to study the decay-time resolution.

Figure 4.8 shows the mass distribution of the B meson candidates of the selected prescaled sample. A fit to this distribution is performed to extract an sWeighted sample of the combinatorial background, which is used to study the resolution. The parameterisation is identical to the one discussed in Section 5.1.2 and the result of this fit is shown in Figure 4.8 as well. Figure 4.9 shows the mass distribution of J/ψ candidates in the sWeighted sample. No contamination from non- J/ψ decays is apparent.

The background sample does not only consist of prompt background, but also contains a long-lived contribution of partially reconstructed B decays and other non-prompt decays. This contribution is parameterised as an exponential decay, which is corrected by a cubic spline function $\varepsilon_{\text{sp}}(t)$ allowing for an empirical description of this component. The lifetime of the exponential decay is fixed to the B^0 lifetime $\tau_{B^0} = 1.520$ ps [39].

Knots for the spline function are chosen at $-2.00, 0.00, 0.25, 0.50, 1.00, 2.50$ and 5.00 ps, while for $t < 0$ the spline function is defined to be $\varepsilon_{\text{sp}}(t) = 1$ by fixing the first four base spline coefficients to $c_{r,1} = c_{r,2} = c_{r,3} = c_{r,4} = 1$. Additionally, the spline function is defined to be flat for $t > 5$ ps, which constrains the last two

4 Analysis ingredients and preparatory studies

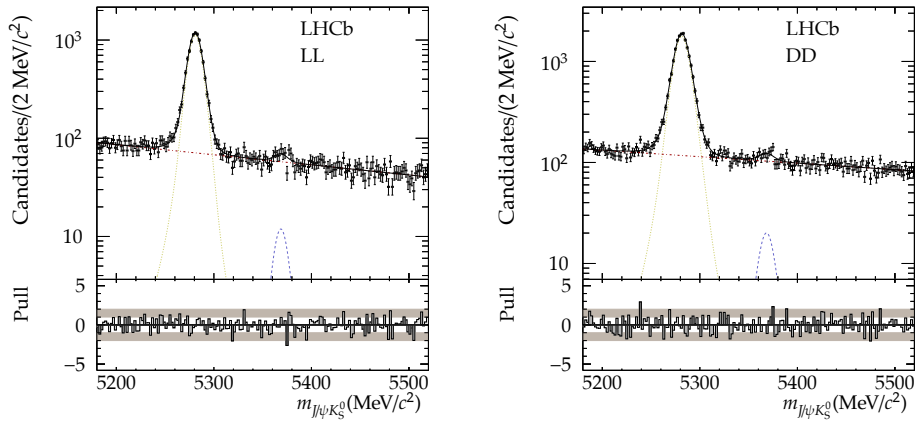


Figure 4.8 – Mass distribution of $B \rightarrow J/\psi K_S^0$ candidates in the prescaled sample for the LL (left) and DD (right) sample. The fit to extract sWeights for the resolution calibration is shown (black, solid) with the $B_s^0 \rightarrow J/\psi K_S^0$ signal component (blue, dashed), the $B^0 \rightarrow J/\psi K_S^0$ signal component (yellow, dotted), and the combinatorial background (red, dash-dotted).

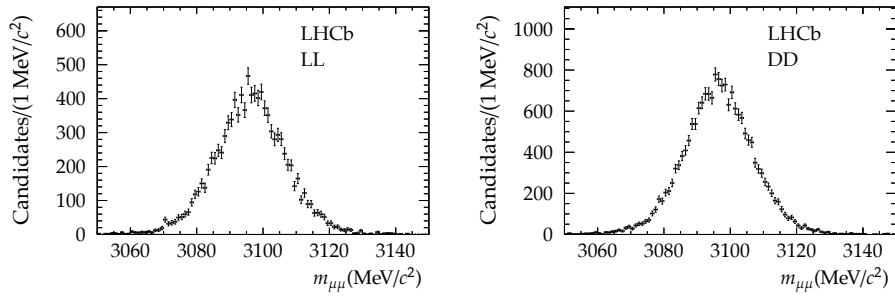


Figure 4.9 – Mass distribution of J/ψ candidates in the sWeighted background sample for LL (left) and DD (right). The stringent selection ensures a very pure sample.

coefficients to be equal, i.e. $c_{r,8} = c_{r,9}$.

The resolution parameters are determined by the means of an unbinned extended maximum likelihood fit in the range -2 ps to 5 ps to the sWeighted background decay-time distribution using the model

$$P(t | \sigma_t) \propto R(t - t' | \sigma_t) \otimes N_p \delta(t') + \left(R(t - t' | \sigma_t) \otimes N_L \exp(-t'/\tau_{pp}) \right) \varepsilon_{sp}(t), \quad (4.26)$$

with the yields N_p and N_L of the prompt and long-lived component. Figure 4.10 shows both the prescaled background decay-time distribution as well as the fitted model. The peak of prompt candidates is clearly visible and allows to determine the resolution parameters. Results for these are given in Table 4.11 for the resolution itself and Tables 4.12 and 4.13 for the remaining parameters. As mentioned before, the resolution mean μ_r is not used in the nominal analysis. The results show only

a very small deviation, justifying this simplification. A sophisticated study (see Section 5.4.6) investigates the systematic uncertainty resulting from fixing μ_r .

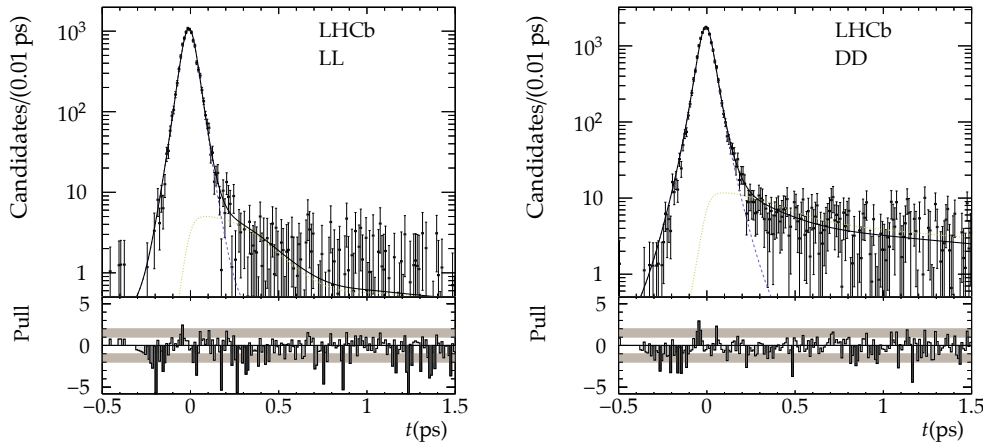


Figure 4.10 – Decay-time distribution of $B \rightarrow J/\psi K_S^0$ candidates in the sWeighted background sample for LL (left) and DD (right). The fit to determine the resolution (black, solid) shows the prompt component (blue, dashed) which is described using $R(t - t' | \sigma_t)$ in Equation (4.24) and the long-lived component (yellow, dotted) utilising cubic splines.

Table 4.11 – Results for parameters describing the decay-time resolution as determined on the sWeighted background sample.

Parameter	LL	DD	Unit
s_r	1.09 ± 0.09	1.04 ± 0.04	
$r_{r,2}$	1.7 ± 10.7	1.87 ± 0.08	
$r_{r,3}$	19 ± 13	10.0 ± 1.7	
$f_{r,1}$	0.61 ± 0.05	0.75 ± 0.06	
$f_{r,3}$	0.004 ± 0.025	0.0015 ± 0.0008	
μ_r	-0.009 ± 0.006	-0.00551 ± 0.00034	ps

The resolution parameters of the LL sample show large uncertainties for the higher-order Gaussian functions. The reason is the neural network selection being very efficient at rejecting background, which affects the number of prompt candidates available for the determination of the resolution. The amount of background could be increased with a looser cut on the neural net. However, the resolution is not necessarily independent of the neural-network requirement so that this would make additional studies on a possible dependence necessary, which may entail further corrections. As this would likely trade increased statistical power

Table 4.12 – Results for spline coefficients describing the long-lived component in the sWeighted background sample.

Parameter	LL	DD
$c_{r,5}$	0.16 ± 0.29	0.57 ± 0.20
$c_{r,6}$	0.24 ± 0.35	0.57 ± 0.17
$c_{r,7}$	0.3 ± 0.7	0.87 ± 0.33
$c_{r,8} = c_{r,9}$	0.6 ± 0.4	0.70 ± 0.22

Table 4.13 – Results for the yields of the prompt and long-lived component in the sWeighted background sample.

Parameter	LL	DD
N_P	$10\,180 \pm 140$	$17\,050 \pm 150$
N_L	350 ± 50	1140 ± 60

for additional systematic uncertainties, the unbiased determination is preferred.

Other LHCb measurements parameterise $\delta_i(\sigma_t)$ including a constant offset as this provides a more precise calibration of the decay-time uncertainty estimates [12, 91]. Due to the limited statistics, a comparable study is not feasible for this analysis. However, for future measurements of $B_s^0 \rightarrow J/\psi K_S^0$ decays the improved statistical precision on the CP asymmetries will make further effort to improve the understanding of the resolution mandatory.

Analogously to \mathcal{D}_{tag} in Equation (4.5), the dilution \mathcal{D}_{res} of the observed CP asymmetry $\mathcal{A}_{\text{obs}}(t)$ due to the finite decay-time resolution can be defined. The resolution power \mathcal{P}_{res} is analogous to the tagging power and describes the fraction of candidates with perfect decay-time resolution yielding the same statistical precision on the CP asymmetries. The respective per-candidate quantities for $B_s^0 \rightarrow J/\psi K_S^0$ decays can be quantified as [92]

$$\mathcal{D}_{\text{res}}(\sigma_{t,i}) = \sum_{j=1}^3 f_{r,j} \exp\left(-\frac{1}{2} \Delta m_s^2 \delta_j^2(\sigma_{t,i})\right), \quad \mathcal{P}_{\text{res}}(\sigma_{t,i}) = \mathcal{D}_{\text{res}}^2(\sigma_{t,i}), \quad (4.27)$$

where Δm_s is the B_s^0 mass difference. From this follow the average sample dilution and resolution power for the sWeighted sample with sWeights s_i and number of candidates N_{Bkg} in the prescaled sample:

$$\langle \mathcal{D}_{\text{res}} \rangle = \sum_i \frac{s_i}{N_{\text{Bkg}}} \mathcal{D}_{\text{res}}(\sigma_{t,i}), \quad \langle \mathcal{P}_{\text{res}} \rangle = \sum_i \frac{s_i}{N_{\text{Bkg}}} \mathcal{P}_{\text{res}}(\sigma_{t,i}). \quad (4.28)$$

Table 4.14 summarises the results for both quantities for the LL and DD subsamples. The respective terms for the $B^0 \rightarrow J/\psi K_S^0$ component can be calculated accordingly.

However, due to the small Δm_d and high resolution, the average sample dilution follows as $\langle \mathcal{D}_{\text{res}} \rangle \approx 1$.

Table 4.14 – Average decay-time resolution dilution and power as obtained on the prompt background sample.

	LL	DD
$\langle \mathcal{D}_{\text{res}} \rangle$	0.7344 ± 0.0027	0.7205 ± 0.0021
$\langle \mathcal{P}_{\text{res}} \rangle$	0.553 ± 0.004	0.5419 ± 0.0034

4.4.3 Validation of the resolution parameterisation

For the decay-time resolution it is assumed that the parameterisation determined on the prompt background sample is applicable to the signal $B \rightarrow J/\psi K_S^0$ samples. In general this is not necessarily the case. Furthermore, other effects, like correlations with other observables, may influence the resolution. In this case, these effects must either be corrected for or treated as a systematic uncertainty.

To examine whether the resolution parameterisation of one (sub-)sample is compatible with another (sub-)sample, the average resolution power $\langle \mathcal{P}_{\text{res}} \rangle$ or equivalent quantities are often compared. However, in general this is not a good figure of merit. The resolution Equation (4.24) determines the per-candidate resolution based on the decay-time uncertainty estimates σ_t . Two samples can share the same resolution parameters while having different σ_t distributions. Therefore, both samples will show significantly different $\langle \mathcal{P}_{\text{res}} \rangle$ results, while the parameterisation obtained on one sample is perfectly applicable to the other. Consequently, in order to compare two samples, one has to compare the compatibility of the parameterisation itself, which is not trivial due to the correlations between the individual resolution parameters.

Nevertheless, the resolution power $\langle \mathcal{P}_{\text{res}} \rangle$ is compared between simulated samples of $B_s^0 \rightarrow J/\psi K_S^0$ decays and inclusive J/ψ decays as a function of the neural network requirement and is found to be in good agreement over a wide range, especially in the region of the nominal requirement. As the σ_t distributions are compatible between signal and background candidates (see Appendix A.2), this approach is justified despite the statement above and shows that prompt background candidates are suitable to determine the decay-time resolution.

Furthermore, a study investigating a possible correlation between the decay-time resolution and the decay time itself using a simulated sample of $B^0 \rightarrow J/\psi K_S^0$ decays shows no such dependence. However, a dependence of the resolution offset μ_r and the candidate mass m is found. This effect is studied as a systematic uncertainty in Section 5.4.6.

4.5 Decay-time acceptance

The decay-time acceptance describes inefficiencies in the decay-time distribution of the signal candidates, which are introduced by the candidate reconstruction, the trigger, and the selection. The multivariate selection, for example, utilises variables which measure how well the daughter tracks are separated from the PV in order to suppress prompt background tracks. As daughter tracks of B candidates with small decay times t only show little separation from the PV, these candidates will less likely pass the neural network with respect to candidates with high t . Another effect concerns the track reconstruction, which is known to be less efficient for candidates with high decay times where the daughter tracks only leave few hits in the VELO [93, 94]. Altogether, the efficiency $\varepsilon_{\text{acc}}(t)$ describing the acceptance is a non-trivial function which cannot be derived analytically. Therefore, the acceptance is determined empirically using cubic splines.

The kinematic similarities between $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ allow the usage of a common acceptance function for both samples simultaneously. In the nominal fit, the spline parameters are varied, where the statistical power to determine these stems from the much larger $B^0 \rightarrow J/\psi K_S^0$ component. Knots at 0.2, 0.5, 1.0, 2.0, 8.0 and 15.0 ps are found to be well-suited to describe the decay-time acceptance. For $t > 15$ ps the function is designed to be constant, which fixes the last two coefficients to $c_{\text{acc},7} = c_{\text{acc},8} = 1$.

Although the spline coefficients are determined in the nominal multidimensional fit for the CP observables, an sWeighted sample of $B^0 \rightarrow J/\psi K_S^0$ decays is used to validate the parameterisation. The sWeights are determined using a fit to the mass m with the identical parameterisation as described in Section 5.1.2. A fit to the sWeighted decay-time distribution of the $B^0 \rightarrow J/\psi K_S^0$ signal candidates allows to determine the spline coefficients using the PDF

$$P(t | \sigma_t) \propto \varepsilon_{\text{acc}}(t) \left(R(t - t' | \sigma_t) \otimes \exp(-t'/\tau_{B^0}) \right), \quad (4.29)$$

where σ_t is the decay-time uncertainty estimate, $R(t - t' | \sigma_t)$ the resolution as in Equation (4.24), and τ_{B^0} the B^0 lifetime, which is fixed to $\tau_{B^0} = 1.520$ ps [39]. Figure 4.11 shows the decay-time distribution of the $B^0 \rightarrow J/\psi K_S^0$ signal candidates, the PDF to determine the spline coefficients, and the acceptance function $\varepsilon_{\text{acc}}(t)$. Results for the coefficients are listed in Table 4.15.

4.6 Production asymmetry

Although the production ratio of b and \bar{b} quarks is symmetric at the LHC, the same is not necessarily true for the respective b hadrons. The production asymmetry A_P

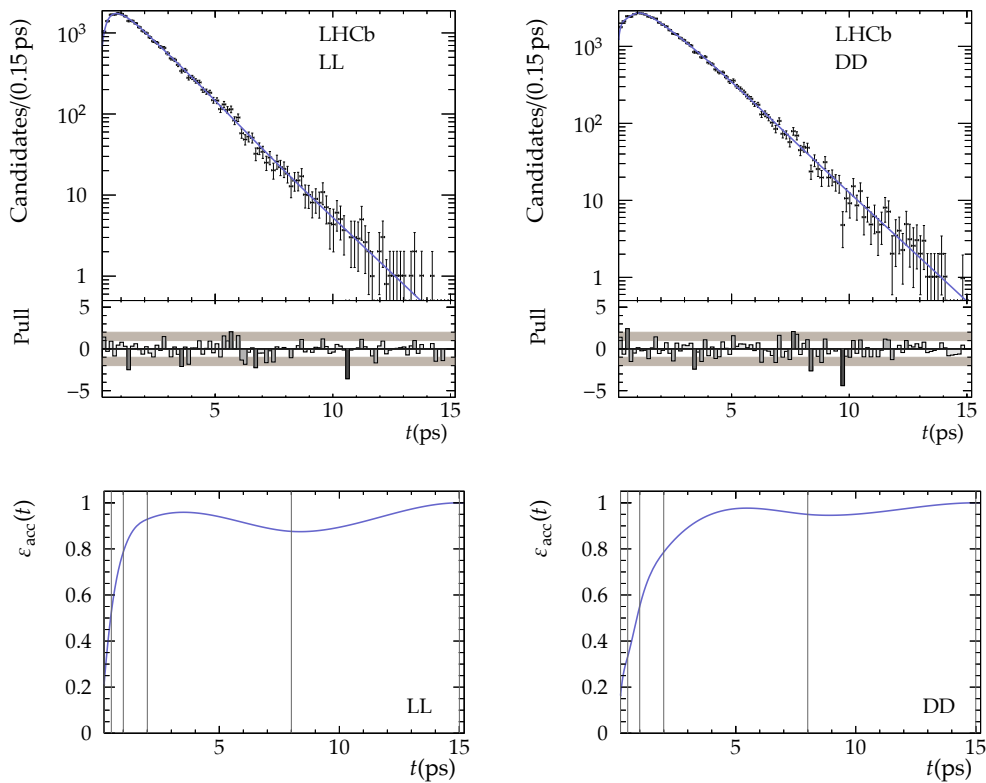


Figure 4.11 – Decay-time distribution of $B^0 \rightarrow J/\psi K_S^0$ signal candidates obtained using sWeights and fitted PDF to determine spline parameters (top), and efficiency function $\varepsilon_{\text{acc}}(t)$ describing the decay-time acceptance (bottom) for the LL (left) and DD (right) subsample.

describes the difference in B vs. \bar{B} production rates and is defined as

$$A_P(B) = \frac{\sigma(\bar{B}) - \sigma(B)}{\sigma(\bar{B}) + \sigma(B)}, \quad (4.30)$$

where B (\bar{B}) represents a B^0 or B_s^0 (\bar{B}^0 or \bar{B}_s^0) meson and σ denotes the respective production cross-section. There are different underlying effects which can lead to a production asymmetry, the most notable connected to the fact that the b and \bar{b} quarks can form hadrons with the u or d valence quarks of the beam remnant. This will lead to a slightly increased production rate of B^0 and B^+ mesons with respect to \bar{B}^0 and B^- mesons. Consequently, this effect will in return lower the production rate of B_s^0 over \bar{B}_s^0 mesons.

The knowledge of the production asymmetry is an essential ingredient for the measurement of CP violation as it affects the observable asymmetry $\mathcal{A}_{\text{obs}}(t)$ with

Table 4.15 – Results for the spline coefficients describing the decay-time acceptance as obtained on an sWeighted sample of $B^0 \rightarrow J/\psi K_S^0$ decays.

Parameter	LL	DD
$c_{\text{acc},1}$	0.21 ± 0.04	0.167 ± 0.018
$c_{\text{acc},2}$	0.36 ± 0.05	0.269 ± 0.025
$c_{\text{acc},3}$	0.63 ± 0.08	0.350 ± 0.027
$c_{\text{acc},4}$	0.91 ± 0.12	0.67 ± 0.05
$c_{\text{acc},5}$	1.05 ± 0.13	1.08 ± 0.08
$c_{\text{acc},6}$	0.81 ± 0.18	0.91 ± 0.11

respect to the true CP asymmetry $\mathcal{A}(t)$ by introducing an intrinsic asymmetry:

$$\mathcal{A}_{\text{obs}}(t) = \mathcal{A}(t) + A_P. \quad (4.31)$$

LHCb has measured the production asymmetry for B^0 and B_s^0 mesons in decays of $B^0 \rightarrow J/\psi K^{*0}$, $B^0 \rightarrow D^- \pi^+$, and $B_s^0 \rightarrow D_s^- \pi^+$ by performing a time-dependent analysis, which distinguishes between the flavour-specific final states of the decays and thus allows to determine A_P [95]. The measurement uses a data sample corresponding to an integrated luminosity of 1 fb^{-1} of pp collisions at a centre-of-mass energy of 7 TeV and extracts A_P as a function of transverse momentum p_T and pseudorapidity η . The overall value, integrated over p_T and η in the analysed samples, is $\mathcal{O}(1\%)$ for B^0 and B_s^0 and in agreement with the theoretical predictions.

As A_P depends on p_T and η , the effective production asymmetry $A_{P,\text{eff}}$ is used for this analysis, which is valid for the p_T and η distributions of the $B \rightarrow J/\psi K_S^0$ signal candidates. It is determined by reweighing the binned results $A_{P,ij}$ from Ref. [95] with the observed fraction f_{ij} of signal candidates in the respective (p_T, η) bins as

$$A_{P,\text{eff}} = \sum_i \sum_j f_{ij} A_{P,ij}, \quad f_{ij} = \frac{N_{ij}}{N_{B^0}}, \quad (4.32)$$

where the indices i and j represent the respective p_T and η bins. The number of signal candidates N_{ij} in the bin is determined separately for LL and DD on an sWeighted sample of $B^0 \rightarrow J/\psi K_S^0$ decays, where the sWeights are determined using a fit to the mass distribution of the candidates. The same weighted sample is also used for $A_{P,\text{eff}}(B_s^0)$ as the $B_s^0 \rightarrow J/\psi K_S^0$ sample contains too few events to determine two-dimensional fractions. Finally, N_{B^0} is the number of all $B^0 \rightarrow J/\psi K_S^0$ decays in the sample. Statistical uncertainties on $A_{P,\text{eff}}$ are calculated accordingly and the systematic uncertainties take the uncertainties on f_{ij} as well as the correlated and uncorrelated systematic uncertainties on $A_{P,ij}$ into account. Details of the calculation of the uncertainties as well as correlated and uncorrelated systematic uncertainties can be found in Ref. [96]. The resulting effective production

asymmetries for the analysed samples are:

$$\begin{aligned}
A_{\text{P,eff}}(B^0, \text{LL}) &= -0.0117 \pm 0.0057 (\text{stat}) \pm 0.0013 (\text{syst}), \\
A_{\text{P,eff}}(B^0, \text{DD}) &= -0.0095 \pm 0.0051 (\text{stat}) \pm 0.0013 (\text{syst}), \\
A_{\text{P,eff}}(B_s^0, \text{LL}) &= -0.0415 \pm 0.0316 (\text{stat}) \pm 0.0030 (\text{syst}), \\
A_{\text{P,eff}}(B_s^0, \text{DD}) &= -0.0219 \pm 0.0246 (\text{stat}) \pm 0.0030 (\text{syst}).
\end{aligned} \tag{4.33}$$

4.7 Branching ratio measurement

The ratio of $B_s^0 \rightarrow J/\psi K_S^0$ to $B^0 \rightarrow J/\psi K_S^0$ branching fractions can naively be calculated by dividing the ratio $R_B = N_{B_s^0}/N_{B^0}$ of observed yields by the ratio f_s/f_d of B_s^0 to B^0 meson hadronisation fractions. However, this is incorrect in case the overall selection efficiencies are different for the two channels. Therefore, a correction factor f_B is introduced and the branching ratio is defined as

$$\frac{\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0)}{\mathcal{B}(B^0 \rightarrow J/\psi K_S^0)} = R_B f_B \frac{f_d}{f_s}. \tag{4.34}$$

The correction factor is comprised of the ratio f_{sel} of selection efficiencies ε_{sel} determined on simulated samples and an additional correction factor f_{corr} which accounts for differences in the selection efficiencies between data and simulation:

$$f_B = f_{\text{sel}} f_{\text{corr}} = \frac{\varepsilon_{\text{sel}, B_s^0} I_{B^0}^{\text{data}} I_{B_s^0}^{\text{MC}}}{\varepsilon_{\text{sel}, B^0} I_{B^0}^{\text{MC}} I_{B_s^0}^{\text{data}}}. \tag{4.35}$$

The selection efficiencies ε_{sel} consist of the geometrical acceptance of the detector and the efficiencies of the reconstruction, stripping, trigger, preparatory selection requirements, and the neural networks. These do not necessarily treat $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ decays identically and therefore, f_{sel} accounts for these differences. Additionally, the assumed physical properties, especially the lifetime τ and decay-width difference $\Delta\Gamma$, of the B mesons in the simulation and the shapes of the decay-time acceptance in the samples affect the observed yields. Consequently, differences between data and simulation have to be accounted for. This is done by integrating the acceptance-corrected theoretical decay rates over the decay time, resulting in the factors I . These factors are calculated for the simulated samples using the identical physical properties as in the sample generation, while for the data the world averages (see Table 5.3) are used.

The resulting selection efficiency ratios for the LL and DD subsamples are

$$f_{\text{sel}}^{\text{LL}} = 0.973 \pm 0.010, \quad f_{\text{sel}}^{\text{DD}} = 0.989 \pm 0.007. \tag{4.36}$$

4 Analysis ingredients and preparatory studies

Differences between data and simulation largely cancel between B_s^0 and B^0 so that the correction factors f_{corr} are found to be

$$f_{\text{corr}}^{\text{LL}} = 0.999 \pm 0.028, \quad f_{\text{corr}}^{\text{DD}} = 1.00 \pm 0.04. \quad (4.37)$$

Finally, the overall correction factors f_B are determined as

$$f_B^{\text{LL}} = 0.972 \pm 0.029, \quad f_B^{\text{DD}} = 0.99 \pm 0.04. \quad (4.38)$$

5 Measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$

The time-dependent measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ mainly consists of a multidimensional maximum-likelihood fit that extracts the CP observables. Extensive studies are conducted to validate the fit and confirm that it is able to correctly estimate the parameters. This includes studies with large samples of simulated pseudo-experiments, comparisons of different fitter implementations, fits on fully simulated Monte Carlo samples, and studies to investigate the correctness of the extracted uncertainties.

As the results from the maximum-likelihood fit only account for statistical uncertainties, more studies need to be performed to determine possible sources of systematic uncertainties. The latter arise from the choice of fit parameterisation and from higher-order effects which are neglected in the nominal fit. Finally, the results including statistical and systematic uncertainties as well as a detailed investigation of the confidence intervals are presented.

5.1 Maximum-likelihood fit

In high energy physics, the unbinned maximum-likelihood method is a popular tool for parameter estimation. The multidimensional case uses a data sample of n reconstructed $B \rightarrow J/\psi K_S^0$ candidates with the observables x of each candidate as

$$\mathbf{x} = (m_{J/\psi K_S^0}, t, d^{\text{OS}}, d^{\text{SSK}}, \eta^{\text{OS}}, \eta^{\text{SSK}}, \sigma_t), \quad (5.1)$$

where x represents the sample of all $B \rightarrow J/\psi K_S^0$ candidates and x_i the set of observables for a single candidate in event i . The individual observables will be explained further below. It is assumed that the multidimensional distribution of the candidates can be modelled with a probability density function (PDF) $P(\mathbf{x}|\lambda)$, which depends on a set of parameters λ . The likelihood $L(\lambda|\mathbf{x})$ is then defined as

$$L(\lambda|\mathbf{x}) = \prod_{i=1}^n P(x_i|\lambda). \quad (5.2)$$

This represents the probability to observe the given sample under the assumption that $P(\mathbf{x}|\lambda)$ describes the underlying distribution. Consequently, the parameters λ

can be varied to find the global maximum of $L(\lambda | \mathbf{x})$ resulting in the best estimate of the parameters λ_{best} . The maximum-likelihood method is consistent and unbiased in the asymptotic limit of $n \rightarrow \infty$. A maximum-likelihood fit implements the optimisation to find λ_{best} .

In this analysis, the maximum-likelihood method is performed simultaneously on multiple subsamples, which are identified by categories. Furthermore, the PDF describes different components, thus a common variation, the extended unbinned maximum-likelihood method, is used. This incorporates the expected number of candidates or yields N_j for a component j into the likelihood by adding an according Poisson term. In each category, the PDF is a sum of PDFs for the individual components. Thus, the likelihood is defined as

$$L(\lambda | \mathbf{x}) = \prod_k \prod_l \frac{\exp\left(-\sum_j N_j^{kl}\right)}{n!} \prod_i \sum_j^{n^{kl}} N_j^{kl} P_j^{kl}(\mathbf{x}_i^{kl} | \lambda_j^{kl}). \quad (5.3)$$

The indices k , l , and j describe the division of the sample and PDF into simultaneously fitted categories and components:

- The index $k = \{\text{LL}, \text{DD}\}$ specifies the track type of the daughter $\pi^+ \pi^-$ of the K_S^0 candidates, i.e. whether both are long (LL) or downstream (DD) tracks.
- The index $l = \{\text{OS}, \text{SSK}, \text{U}\}$ identifies the division into exclusively OS tagged candidates (OS), all candidates with an SSK tagging response (SSK), and untagged candidates (U).
- The index $j = \{B_s^0, B^0, \text{Bg}\}$ separates between the components for $B_s^0 \rightarrow J/\psi K_S^0$ decays, $B^0 \rightarrow J/\psi K_S^0$ decays, and combinatorial background (Bg). In all cases where the description applies to both B_s^0 and B^0 , the symbol B will be used.

Consequently, N_j^{kl} describes the expected number of candidates of component j in the category kl , while n^{kl} is the actual total number of candidates in the according category. For each category and component, an individual PDF $P_j^{kl}(\mathbf{x}_i^{kl} | \lambda_j^{kl})$ is defined for the candidate's observables in that category \mathbf{x}_i^{kl} with an individual set of parameters λ_j^{kl} . The latter can contain parameters being shared between different PDFs.

In most implementations the negative logarithm of the likelihood $-\ln L(\lambda | \mathbf{x})$ is minimised instead of maximising $L(\lambda | \mathbf{x})$. This does not affect the resulting λ_{best} and is more convenient as usually numerical iterative minimisation procedures are used for maximum-likelihood fits. This analysis uses the MINUIT2 minimisation package for that purpose, which also determines uncertainties for all optimised parameters [97]. The MINOS algorithm allows to extract asymmetric uncertainties by analysing the shape of the log-likelihood minimum. It is utilised for the uncertainties of all CP observable estimates.

5.1.1 Observables

This analysis uses kinematic and flavour tagging observables for the fit, where the former are extracted using a decay tree fit which constrains the J/ψ and K_S^0 candidate masses to their known values [19] and additionally constrains the B candidate momentum to point to the PV. The utilised observables are the reconstructed mass of the $B \rightarrow J/\psi K_S^0$ candidates $m_{J/\psi K_S^0}$ (for simplicity also referred to as m) mainly used to separate both signal components from each other and from the background, the decay time t of the B candidate, the OS and SSK tagging responses d^{OS} and d^{SSK} , the OS and SSK mistag probability estimates η^{OS} and η^{SSK} , and the decay-time uncertainty estimates σ_t . Table 5.1 shows the ranges for all observables.

Table 5.1 – Ranges of the observables used in the analysis.

Observable	Range
$m_{J/\psi K_S^0}$	5180–5520 MeV/ c^2
t	0.2–15.2 ps
d^{OS}	{-1, +1}
d^{SSK}	{-1, +1}
η^{OS}	0.0–0.5
η^{SSK}	0.0–0.5
σ_t	0.0–0.2 ps

5.1.2 Parameterisation

Although the PDF in each simultaneously fit category kl can in principle be modelled independently, most categories share the same parameterisation which only differs for certain aspects where necessary. As mentioned before, the PDF for category kl is the sum over PDFs for the $B_s^0 \rightarrow J/\psi K_S^0$, $B^0 \rightarrow J/\psi K_S^0$ and background components

$$\sum_j N_j^{kl} P_j^{kl}(\mathbf{x}^{kl}) = N_{B_s^0}^{kl} P_{B_s^0}^{kl}(\mathbf{x}^{kl}) + N_{B^0}^{kl} P_{B^0}^{kl}(\mathbf{x}^{kl}) + N_{B_g}^{kl} P_{B_g}^{kl}(\mathbf{x}^{kl}). \quad (5.4)$$

Parameters λ^{kl} are omitted for clarity and will only be specified where explicitly needed. Both signal PDFs $P_{B_s^0}^{kl}(\mathbf{x}^{kl})$ and $P_{B^0}^{kl}(\mathbf{x}^{kl})$ are parameterised similarly with partly shared and partly individual parameters. Furthermore, each PDF $P_j^{kl}(\mathbf{x}^{kl})$ is modelled to factorise into a PDF describing the reconstructed mass $P_j^{kl}(m)$ and a PDF describing the decay time and other observables $P_j^{kl}(t, d^{\text{OS}}, d^{\text{SSK}} | \eta^{\text{OS}}, \eta^{\text{SSK}}, \sigma_t)$. The individual parameterisations will be explained in the following section.

Mass parameterisation

The signal PDF describing the reconstructed mass of signal components commonly relies on the Crystal Ball (CB) function [98]. This combines a Gaussian core with a power-law tail, where the former describes the detector resolution and the latter is used to model radiative tails or other detector-specific effects. Typically there are multiple effects leading to different uncertainties on the reconstructed mass so that a single Gaussian core is no longer sufficient to describe the observed distribution. Therefore, the double-sided Hypatia distribution is utilised, which is a modification of the CB function resulting from marginalising over the per-candidate mass error uncertainty estimates and which features tails to both sides of the mass peak at μ [99]. A special case with the core distribution being symmetric around this peak is chosen. The according PDF is defined as

$$P_{\Upsilon}(m \mid \mu, \sigma, \lambda, \zeta, a_1, n_1, a_2, n_2) \propto \begin{cases} G(m, \mu, \sigma, \lambda, \zeta) & \text{for } -a_1 < \frac{m-\mu}{\sigma} < a_2, \\ \frac{G(\mu+a_1\sigma, \mu, \sigma, \lambda, \zeta)}{(1-m/(n_1 H(\mu+a_1\sigma, \mu, \sigma, \lambda, \zeta)+a_1\sigma))^{n_1}} & \text{for } -a_1 > \frac{m-\mu}{\sigma}, \\ \frac{G(\mu-a_2\sigma, \mu, \sigma, \lambda, \zeta)}{(1-m/(n_2 H(\mu-a_2\sigma, \mu, \sigma, \lambda, \zeta)-a_2\sigma))^{n_2}} & \text{for } a_2 < \frac{m-\mu}{\sigma}, \end{cases} \quad (5.5)$$

with the generalised hyperbolic distribution $G(x, \mu, \sigma, \lambda, \zeta)$ as the core of a CB-like function and $H(x, \mu, \sigma, \lambda, \zeta)$ defined as

$$G(x, \mu, \sigma, \lambda, \zeta) = \left((x - \mu)^2 + A_\lambda^2(\zeta)\sigma^2 \right)^{1/2\lambda - 1/4} K_{\lambda - 1/2} \left(\zeta \sqrt{1 + \left(\frac{x - \mu}{A_\lambda(\zeta)\sigma} \right)^2} \right), \quad (5.6)$$

$$H(x, \mu, \sigma, \lambda, \zeta) = \frac{G(x, \mu, \sigma, \lambda, \zeta)}{G'(x, \mu, \sigma, \lambda, \zeta)}, \quad A_\lambda^2(\zeta) = \frac{\zeta K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}.$$

In these equations, $K_\lambda(\zeta)$ are modified Bessel functions of the second kind. The parameter ζ is fixed to zero, which allows σ to be interpreted as the width of the core distribution. Furthermore, $\lambda < 0$ is a shape parameter and a_1/a_2 and n_1/n_2 determine the transition points between core and tails as well as the shape of the tails. Due to the combinatorial background, it is not feasible to determine these four tail parameters on data. Therefore, the Hypatia distribution is fit to the selected simulated sample of $B^0 \rightarrow J/\psi K_S^0$ decays. Figure 5.1 shows the mass distribution and according fit. The determined tail parameters, which are fixed in the fits to the data sample are summarised in Table 5.2.

The differently tagged categories share all Hypatia parameters, while for the LL and DD categories completely individual parameters are used. Finally, the

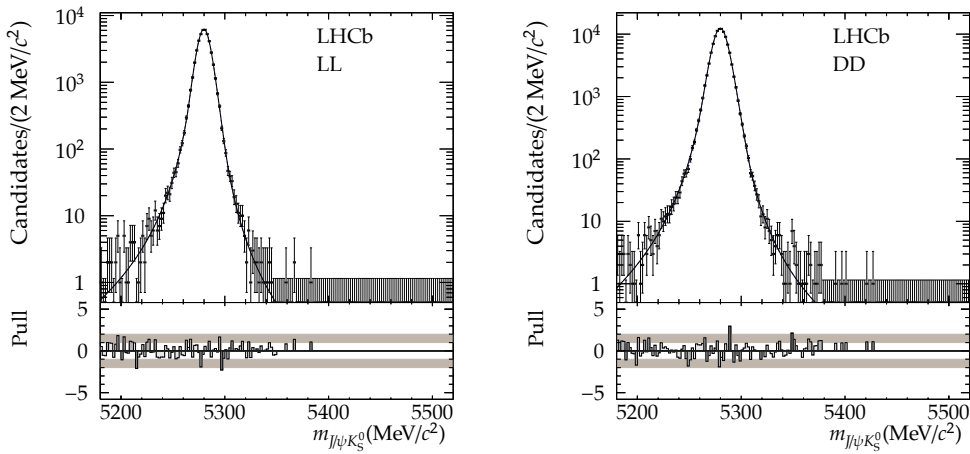


Figure 5.1 – Mass distribution of candidates in the simulated $B^0 \rightarrow J/\psi K_S^0$ sample after applying all selection requirements for the LL (left) and DD (right) category. The fit with the Hypatia distribution (black, solid) is used to determine the tail parameters.

Table 5.2 – Resulting Hypatia tail parameters as obtained on the simulated sample of $B^0 \rightarrow J/\psi K_S^0$ decays.

Parameter	LL	DD
a_1	2.05 ± 0.14	2.22 ± 0.12
a_2	3.0 ± 0.5	3.28 ± 0.30
n_1	3.4 ± 0.4	3.9 ± 0.4
n_2	3.6 ± 0.8	3.4 ± 0.5

descriptions of both the $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ component share all parameters with the exception of μ . For the B_s^0 component, the latter is parameterised as $\mu_{B_s^0} = \mu_{B^0} + s_m$ with the shift s_m being a free parameter of the fit which is shared between all categories.

The background mass PDF is parameterised as an exponential distribution with the slope c_{Bg}^k :

$$P_{Bg}^{kl}(m | c_{Bg}^k) \propto \exp(c_{Bg}^k m). \quad (5.7)$$

Decay-time parameterisation

The decay-time PDF describing both the B_s^0 and B^0 components is a conditional PDF which describes the observables t and d given the conditional observables η and σ_t . It is defined as

$$P_B^{kl}(t, d | \eta, \sigma_t) \propto \varepsilon_{acc}^k(t) \left(R^k(t - t' | \sigma_t) \otimes P_B^{kl}(t', d | \eta) \right), \quad (5.8)$$

where $\varepsilon_{\text{acc}}^k(t)$ describes the decay-time acceptance (see Section 4.5), $R^k(t - t' | \sigma_t)$ the decay-time resolution (see Section 4.4.1 and Equation (4.24)), and $P_B^{kl}(t', d | \eta)$ the distribution of the true decay time t' . For simplicity, only one tag decision d and mistag estimate η is given. These have to be substituted for the OS, SSK or combined tagging quantities where applicable. In case the production flavour of the B mesons is known, the true decay-time distribution using the true tag d' is given as

$$P_B^{kl}(t', d') \propto \exp(-\Gamma t') \left[\cosh\left(\frac{\Delta\Gamma}{2} t'\right) + A^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma}{2} t'\right) + C d' \cos(\Delta m t') - S d' \sin(\Delta m t') \right], \quad (5.9)$$

with the decay width of the B meson Γ , the decay-width difference $\Delta\Gamma$, the B meson oscillation frequency Δm , and the three CP observables S , C , and $A^{\Delta\Gamma}$. Alternatively to Γ , the lifetime $\tau = 1/\Gamma$ can be used. With imperfect tagging, the calibrated mistag probability $\omega(\eta)$, and a production asymmetry $A_{P,B}^k$ between \bar{B} and B , the PDF transforms into

$$P_B^{kl}(t', d | \eta) \propto \exp(-\Gamma t') \left[(1 - d A_{P,B}^k (1 - 2\omega(\eta))) \cosh\left(\frac{\Delta\Gamma}{2} t'\right) + A^{\Delta\Gamma} (1 - d A_{P,B}^k (1 - 2\omega(\eta))) \sinh\left(\frac{\Delta\Gamma}{2} t'\right) + C (d (1 - 2\omega(\eta)) - A_{P,B}^k) \cos(\Delta m t') - S (d (1 - 2\omega(\eta)) - A_{P,B}^k) \sin(\Delta m t') \right]. \quad (5.10)$$

The PDF for $B^0 \rightarrow J/\psi K_S^0$ decays is simplified by assuming $\Delta\Gamma_d = 0$, which eliminates $A^{\Delta\Gamma}$ in the B^0 system.

The background decay-time PDF does not depend on tagging quantities. The PDF for the tag d is assumed to be uniformly distributed, $P_{B_g}^{kl}(d) = 1$. For the decay time itself, a sum of two exponential distributions is found to describe the data. Thus the PDF is defined as

$$P_{B_g}^{kl}(t | \sigma_t) \propto R^k(t - t' | \sigma_t) \otimes \left(\frac{f_{B_g}^k}{N_{1,B_g}^k} \exp\left(\frac{-t'}{\tau_{B_g}^k}\right) + \frac{1 - f_{B_g}^k}{N_{2,B_g}^k} \exp\left(\frac{-t'}{s_{t,B_g}^k \tau_{B_g}^k}\right) \right), \quad (5.11)$$

with the fraction $f_{B_g}^k$, pseudo-lifetime $\tau_{B_g}^k$, relative scale factor s_{t,B_g}^k for the second pseudo-lifetime, and appropriately chosen normalisation factors N_{i,B_g}^k . The parameters are not shared between the LL and DD category.

No explicit parameterisations for the conditional observables η^{OS} , η^{SSK} , and σ_t are used for the maximum-likelihood fit. This implies uniform PDFs for these

observables. This simplification does not affect the result of the likelihood fit in case the underlying distributions are sufficiently similar for the different components. However, if pronounced differences are neglected, the parameter estimation can be biased [100]. The effect of not accounting for these differences in the distributions of the conditional observables is implicitly studied with pseudo-experiments (see Section 5.3.1) and found to be negligible. The distributions of the conditional observables and possible parameterisations are presented in Appendix A.2.

5.1.3 External inputs

Additionally to the five CP observables, the likelihood $L(\lambda | x)$ depends on parameters which are not of direct interest for this measurement. Most of these so-called nuisance parameters are determined in the fit. However, especially for some physical properties it is sensible to provide these externally. Therefore, for the B_s^0 decay width Γ_s , the B^0 lifetime τ_{B^0} , the B_s^0 and B^0 oscillation frequencies Δm_s and Δm_d , and the B_s^0 decay-width difference $\Delta\Gamma_s$, the likelihood $L(\lambda | x)$ is multiplied by Gaussian distributions which constrain these parameters to the respective world averages within their uncertainties. The correlation coefficient between Γ_s and $\Delta\Gamma_s$ of $\rho(\Gamma_s, \Delta\Gamma_s) = -0.271$ is taken into account in the constraining functions. All values are based on Ref. [39] and are listed in Table 5.3. Additionally, the effective production asymmetries are constrained to the values in Equation (4.33).

Table 5.3 – Values for physics parameters which are used in the measurement of the CP observables. All values are taken from Ref. [39].

Parameter	Value	Unit
Γ_s	0.6628 ± 0.0019	ps^{-1}
Δm_s	17.757 ± 0.021	ps^{-1}
$\Delta\Gamma_s$	0.081 ± 0.006	ps^{-1}
τ_{B^0}	1.520 ± 0.004	ps
Δm_d	0.510 ± 0.003	ps^{-1}

5.2 Preliminary results

In the maximum-likelihood fit on the nominal data sample results for 59 parameters are determined, including the 5 CP observables, 18 yields, and 9 constrained externally provided parameters. The results on the CP parameters in the $B_s^0 \rightarrow J/\psi K_S^0$

5 Measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$

decay channel are measured as

$$\begin{aligned} A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} &= 0.49 \pm 0.77, \\ C_{B_s^0 \rightarrow J/\psi K_S^0} &= -0.28 \pm 0.41, \\ S_{B_s^0 \rightarrow J/\psi K_S^0} &= -0.08 \pm 0.40, \end{aligned}$$

where all uncertainties are purely statistical. This is the first measurement of these parameters. All values are in agreement with the theoretical predictions (cf. Section 2.5.3). Figure 5.2 shows the mass and decay-time distribution of the reconstructed candidates as well as the according fit projections.

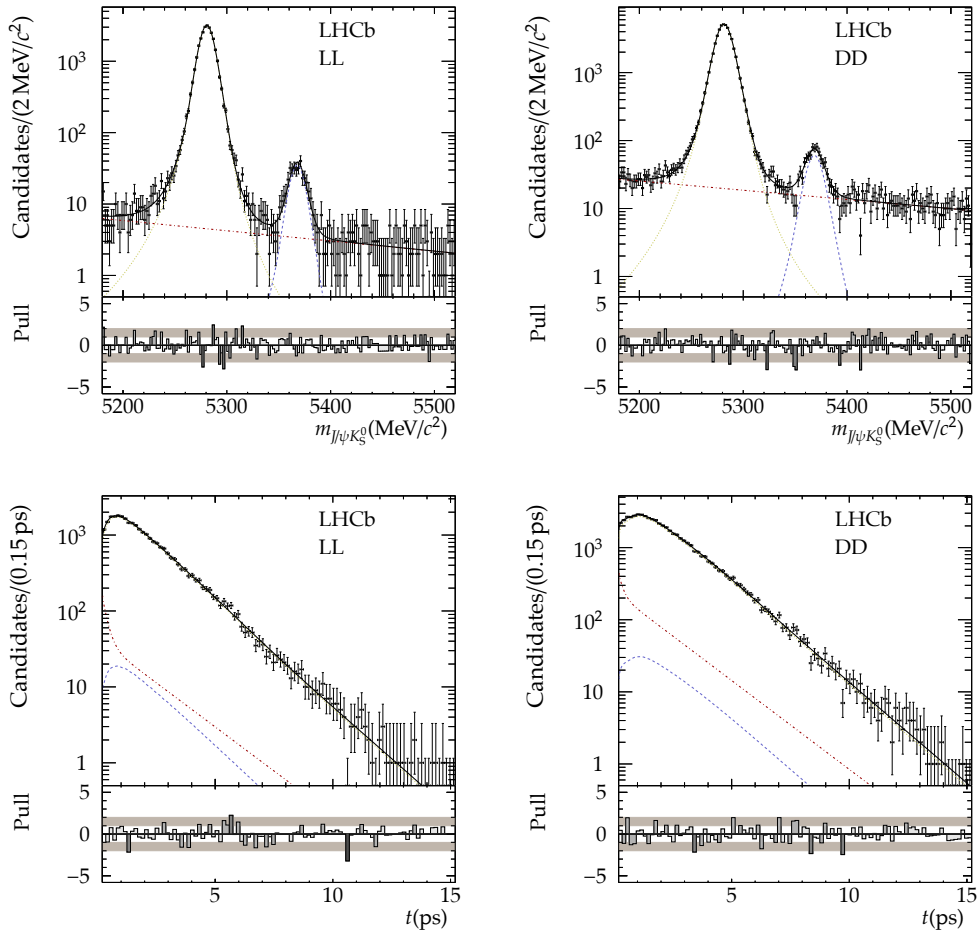


Figure 5.2 – Mass (top) and decay-time (bottom) distribution of the nominal data sample for the LL (left) and DD (right) subsample. The projection of the fitted PDF (black, solid) shows the $B_s^0 \rightarrow J/\psi K_S^0$ (blue, dashed) and the $B^0 \rightarrow J/\psi K_S^0$ (yellow, dotted) component as well as the combinatorial background (red, dash-dotted).

In the $B^0 \rightarrow J/\psi K_S^0$ decay channel, the CP parameters are found as

$$\begin{aligned} C_{B^0 \rightarrow J/\psi K_S^0} &= -0.028 \pm 0.035, \\ S_{B^0 \rightarrow J/\psi K_S^0} &= 0.719 \pm 0.034, \end{aligned}$$

which are in very good agreement with former measurements and the latest dedicated LHCb measurement of CP violation in $B^0 \rightarrow J/\psi K_S^0$ (cf. Section 2.5.3). The measurement of the branching ratio of $B_s^0 \rightarrow J/\psi K_S^0$ to $B^0 \rightarrow J/\psi K_S^0$ decays requires knowledge of the overall yields of the according components. These are determined in the LL and DD subsamples as

$$\begin{aligned} N_{B_s^0}^{\text{LL}} &= 307 \pm 20, & N_{B^0}^{\text{LL}} &= 27800 \pm 170, \\ N_{B_s^0}^{\text{DD}} &= 600 \pm 30, & N_{B^0}^{\text{DD}} &= 51350 \pm 230. \end{aligned}$$

Using Equation (4.34) and determining the weighted average for LL and DD, the resulting overall branching ratio is measured as

$$\frac{\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0)}{\mathcal{B}(B^0 \rightarrow J/\psi K_S^0)} = 0.0431 \pm 0.0017,$$

where $f_s/f_d = 0.259 \pm 0.015$ [80, 81] is used as the value of B_s^0 to B^0 meson hadronisation fractions. This result is in excellent agreement with former measurements of this quantity [79].

The correlation matrix of the CP parameters and constrained parameters is shown in Figure 5.3. No major correlation between the CP parameters and other parameters is observed, except for an intrinsic correlation between $S_{B^0 \rightarrow J/\psi K_S^0}$ and $C_{B^0 \rightarrow J/\psi K_S^0}$. For completeness, the detailed list of results for all nuisance parameters and the full correlation matrix are given in Appendix A.3.

5.3 Fit validation

The maximum-likelihood fit is extensively validated to verify the convergence and correctness of the parameter estimation. First of all, the maximum-likelihood implementation is verified with large samples of simulated pseudo-experiments to investigate whether the parameter or uncertainty estimation in the implementation itself is biased. Additionally, the results of two independent fitter implementations are compared and the fit is performed on independent subsamples of the data. The sFit technique is a complementary method which allows to estimate the parameters without the need to parameterise non-signal components in the decay time. As such it is an ideal cross-check. Furthermore, fully simulated samples allow to

5 Measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$

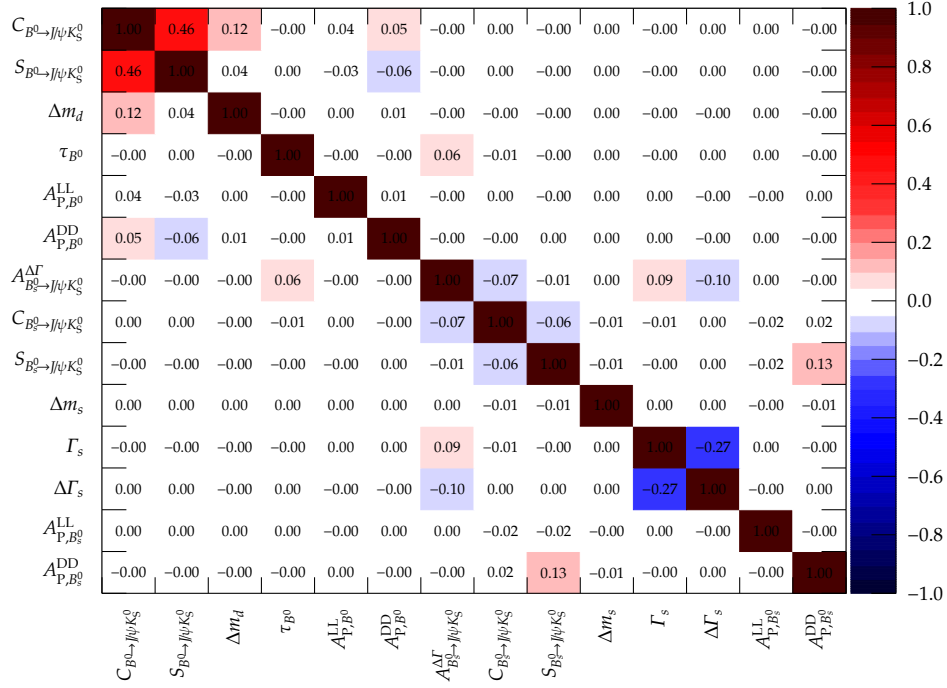


Figure 5.3 – Correlation matrix of the fit parameters in the nominal fit, in which all nuisance parameters are varied. The size of the correlation coefficients is represented by red and blue colours, where a correlation coefficient below 5% is displayed as white. Only the CP asymmetries and all parameters constrained in the fit are shown. See Figure A.5 in Appendix A.3 for the complete correlation matrix.

test multiple aspects of the analysis by comparing results with generation values. Likelihood scans show the shape of the log-likelihood function and are used to validate the minimisation process. A scan of the allowed parameter space for the CP asymmetries investigates whether the fit can correctly determine these independently of the true values. Finally, the Feldman–Cousins method allows to determine frequentist confidence intervals. It is used to validate the uncertainties of the maximum-likelihood fit.

5.3.1 Studies using pseudo-experiments

Studies which repeatedly generate and fit simulated pseudo-experiments, also called toy datasets, are a common tool to validate a fitter implementation. Useful quantities are the residual $r_i(\lambda)$ and pull $p_i(\lambda)$, which are defined for the parameter λ in toy dataset i as

$$r_i(\lambda) = \lambda_i - \lambda_{\text{gen},i}, \quad p_i(\lambda) = \frac{\lambda_i - \lambda_{\text{gen},i}}{\sigma_{\lambda_i}}, \quad (5.12)$$

where λ_i is the parameter result and σ_{λ_i} its estimated uncertainty in that dataset. The generation value is specified by $\lambda_{\text{gen},i}$. In case of asymmetric uncertainties for x , the lower uncertainty $\sigma_{\lambda_i}^{\text{low}}$ is used in the denominator of the pull if $\lambda_i > \lambda_{\text{gen},i}$ while for $\lambda_i \leq \lambda_{\text{gen},i}$ the upper uncertainty $\sigma_{\lambda_i}^{\text{high}}$ is used. Both $\sigma_{\lambda_i}^{\text{low}}$ and $\sigma_{\lambda_i}^{\text{high}}$ are positive in this convention.

If the PDF describes the examined sample, which is true by construction in case the identical PDF is used for the generation and the fit, and the likelihood implementation correctly estimates both λ_i and σ_{λ_i} , the pulls follow a standard normal distribution. By fitting a Gaussian function to the residual and pull distribution, it can be examined whether the parameter estimation is biased, which would be reflected by the fitted mean μ being significantly different from zero. Additionally, uncertainty underestimation is signalled by the fitted pull width s being significantly larger than one, $s > 1$, while uncertainty overestimation results in $s < 1$. The fitted width of the residual distribution ideally reflects the statistical uncertainty of the parameter.

The maximum-likelihood fit is validated with a toy study which reflects the nominal fit. For the generation, the nuisance parameters are set to the values observed in the nominal fit, while the CP parameters are set to the theory predictions (see Section 2.5.3) for the $B_s^0 \rightarrow J/\psi K_S^0$ decay channel and the former world averages for the $B^0 \rightarrow J/\psi K_S^0$ decay channel [39]. Constrained parameters are set to the respective mean of the Gaussian constraint in the generation. Before the subsequent fit, a new mean for the constraint is drawn from the Gaussian distribution and applied in the fit. This procedure follows Ref. [101] and allows the following interpretation: The true value of parameter x is the same for all pseudo-experiments and

thus used in the generation. Each pseudo-experiment is based on an individual externally provided input for x , for instance a separately measured world average. Decay-time uncertainty estimates σ_t and mistag probability estimates η^{OS} and η^{SSK} are drawn from the parameterisations shown in Appendix A.2. Individual distributions are used to sample the conditional observables for all components. As the subsequent fit neglects differences between these distributions, the toy validation study implicitly examines whether this simplification leads to biased parameter estimation. In the fit, all CP parameters and additionally the nuisance parameters are varied like in the nominal fit.

For the nominal toy validation study, 1500 pseudo-experiments are generated and fitted, of which 9 fail to converge. Resulting residual and pull distributions for the CP parameters are shown in Figure 5.4. All parameters and uncertainties are estimated correctly. The width of the residual distributions is in agreement with the statistical uncertainties found in the nominal fit on the data sample.

Additionally to generating pseudo-experiments with the identical PDF which is used to fit the samples, an external generator is employed, which is completely independent of the fit implementation [102]. It is able to generate signal and background candidates from theoretical and empirical parameterisations and, among other features, supports the convolution of the decay-time distribution with a per-candidate resolution model and is able to generate per-candidate mistag estimates. Not using an independent generator can mask an incorrectly implemented PDF as the identical incorrect PDF would be used to fit the samples. The externally generated pseudo-experiments do not use constrained parameters and assume a perfect decay-time acceptance. Apart from that the parameterisation is identical to the nominal fit. All CP and nuisance parameters are varied in the fits. For this toy validation study, 4000 pseudo-experiments are generated and fitted, of which 54 fail to converge. Figure 5.5 shows resulting residual and pull distributions. All parameters and uncertainties are estimated correctly.

5.3.2 Comparisons between independent fitter implementations and subsamples

As a further cross-check, the nominal maximum-likelihood fit is compared to an independent alternative implementation from another LHCb group on both the data sample and a set of simulated pseudo-experiments. The latter are designed to specifically test particular aspects of the analysis. Additionally, the results on different data subsamples are compared with each other and with the nominal result. This includes separating between LL and DD, the LHCb magnet polarities, and years of data taking. Furthermore, the sample is divided into random mutually exclusive subsamples and the fit is performed exclusively on OS and SSK tagged candidates. Comparing the difference between the results of the two

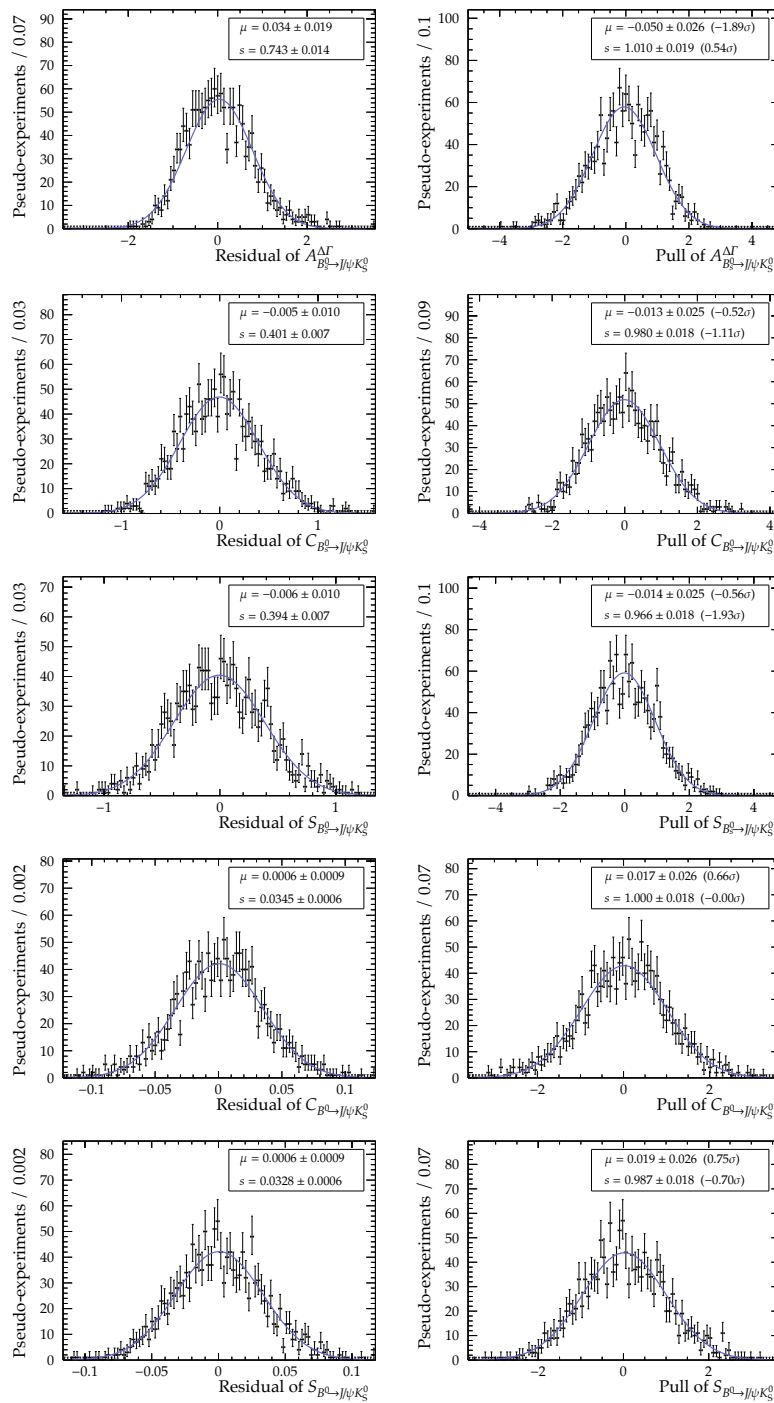


Figure 5.4 – Residual and pull distributions in the nominal toy validation study. A Gaussian function is fitted to all distributions with results for the mean μ and width s shown in each sub-figure. For the pull distributions, the deviations of μ and s from the expectation with respect to their according uncertainty are indicated as well.

5 Measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$

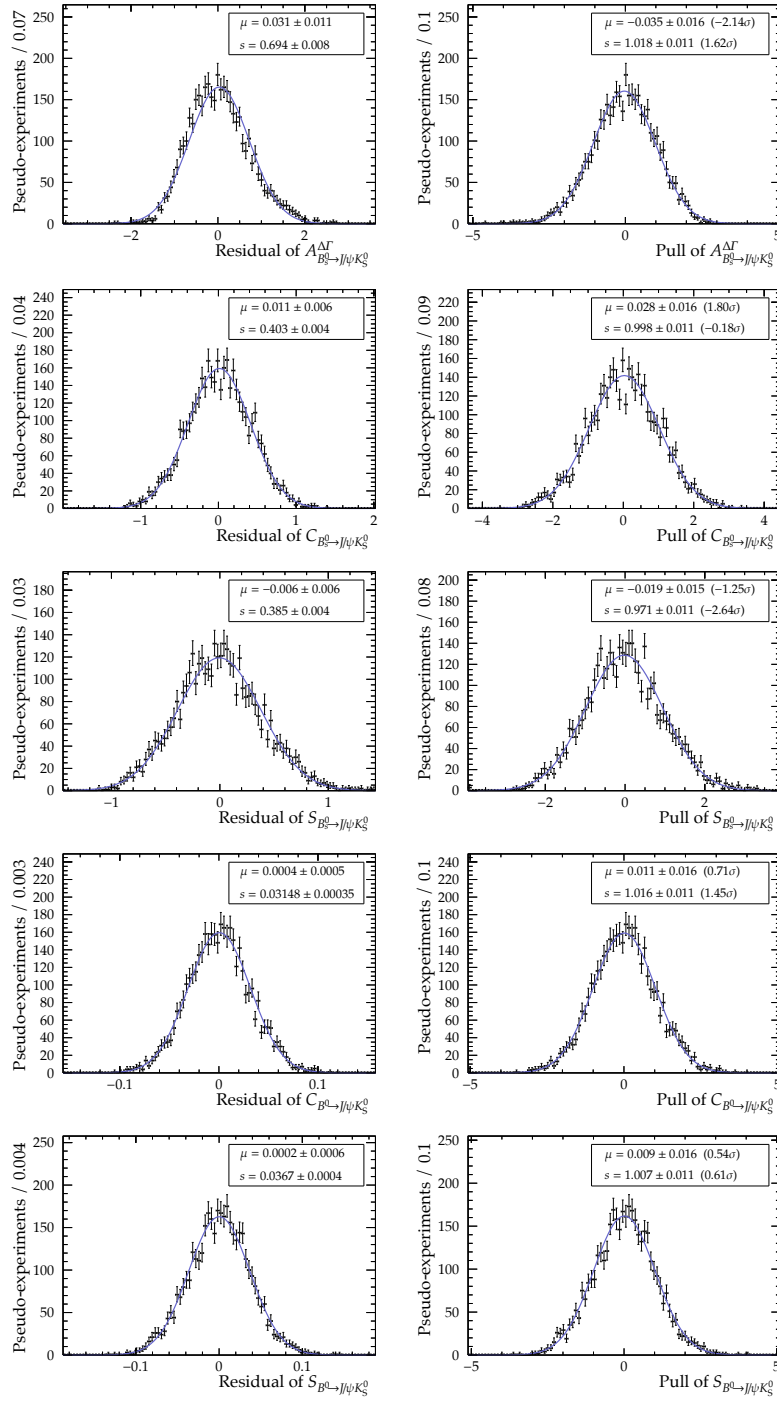


Figure 5.5 – Residual and pull distributions in the toy validation study using the external generator. A Gaussian function is fitted to all distributions with results for the mean μ and width s shown in each sub-figure. For the pull distributions, the deviations of μ and s from the expectation with respect to their according uncertainty are indicated as well.

implementations on a given sample, the full correlation of these is taken into account [103]. Both fit implementations show excellent agreement on all considered samples. The agreement of the results on different subsamples is very well within the expectation of purely statistical fluctuations.

5.3.3 sFit implementation

The $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ signal samples are additionally fitted with the sFit technique. Using a fit of the mass distribution, sWeighted samples for both components are extracted, which in turn are then fitted with the decay-time signal PDFs [85]. sFits offer the advantage of avoiding the necessity of describing the decay-time distributions of non-signal components. Apart from that the procedure follows that of the nominal fit. All nuisance parameters are varied in the fits and Gaussian constraints are applied. The results for the CP parameters are measured as

$$\begin{aligned} A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} &= 0.43 \pm_{0.51}^{0.61}, \\ C_{B_s^0 \rightarrow J/\psi K_S^0} &= -0.38 \pm_{0.34}^{0.33}, & C_{B^0 \rightarrow J/\psi K_S^0} &= -0.030 \pm 0.034, \\ S_{B_s^0 \rightarrow J/\psi K_S^0} &= -0.22 \pm_{0.31}^{0.32}, & S_{B^0 \rightarrow J/\psi K_S^0} &= 0.715 \pm_{0.034}^{0.033}. \end{aligned}$$

Compared to the nominal results (cf. Section 5.2), it is apparent that both values and uncertainties especially for the $B_s^0 \rightarrow J/\psi K_S^0$ component differ. To investigate whether this deviation is within expectation, a set of pseudo-experiments is generated and each sample is fitted once with the nominal fit method and the sFit afterwards. For each toy dataset i the difference $\delta_i(\lambda)$ between the two results for parameter λ is calculated. Figure 5.6 shows the according δ distributions for the CP parameters. The fitted mean of these distributions is consistent with being zero, which shows that both nominal fit method and sFit on average yield identical results. On the other hand, the width of the δ distributions reflects how much a single sFit result can differ from the respective nominal fit result. The fitted widths are in excellent agreement with the observed differences on the nominal data sample. Altogether, this shows that the results for the CP parameters as obtained with the sFits are in very good agreement with the nominal fit.

The uncertainties for the $B_s^0 \rightarrow J/\psi K_S^0$ parameters obtained with the sFits are considerably smaller than those of the nominal fit. In fact, the pull distributions for the sFits on pseudo-experiments, which are also shown in Figure 5.6, confirm that the sFit underestimates the uncertainties of these parameters. Correcting the sFit uncertainties with the fitted pull widths results in excellent agreement with the uncertainties of the nominal fit. The reason for the uncertainty underestimation is that the sFit neglects the dilution from the background candidates. As the signal

5 Measurement of CP violation in $B_S^0 \rightarrow J/\psi K_S^0$

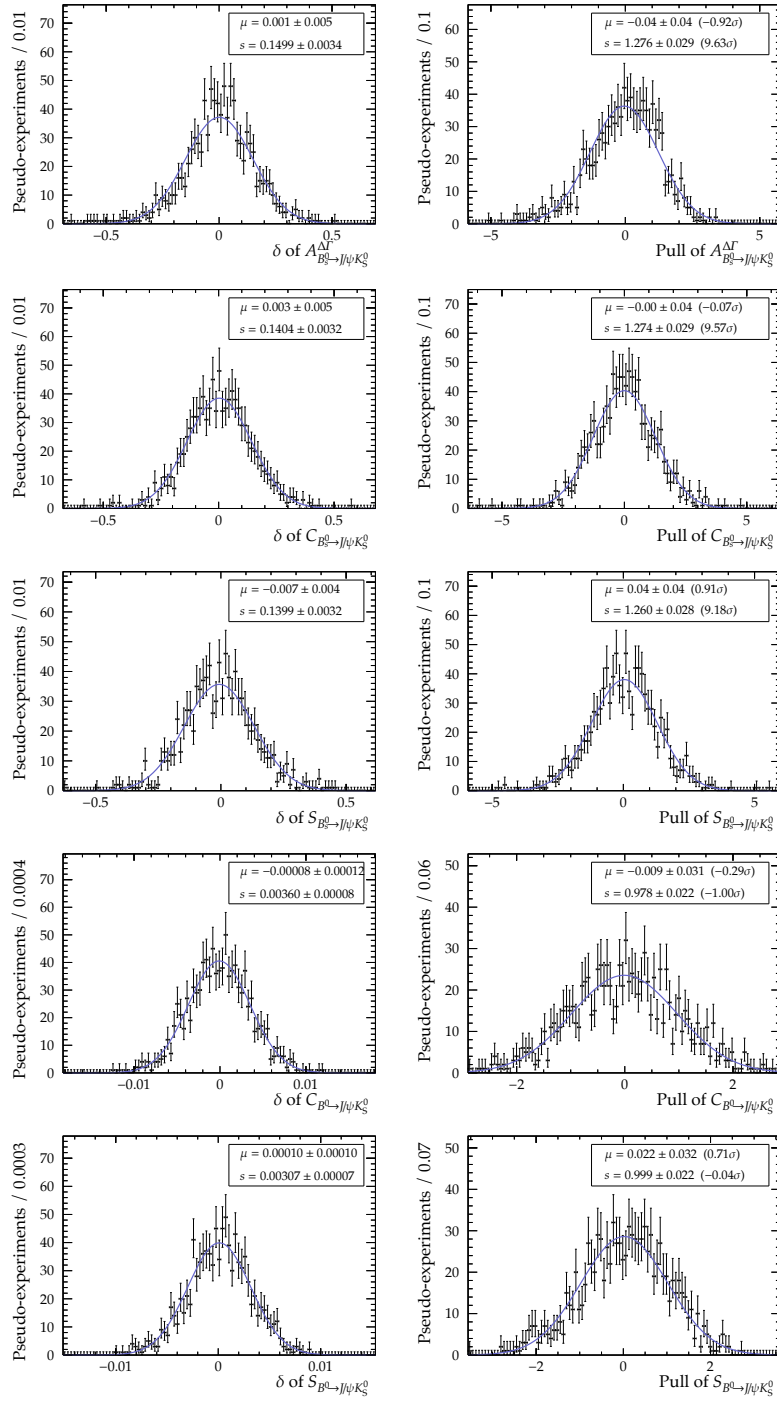


Figure 5.6 – Differences δ between nominal fit method and sFit (left) and pull distributions for the sFit toy study (right) for the CP parameters. All distributions are fitted with a Gaussian function and results for the respective mean μ and width s are shown in each sub-figure.

to background ratio for $B^0 \rightarrow J/\psi K_S^0$ candidates is much higher, this effect is only visible for the $B_s^0 \rightarrow J/\psi K_S^0$ CP parameters.

More precise uncertainty estimates can be determined with the bootstrapping method. For this procedure, new equally large datasets are sampled randomly from the original dataset with replacement. Afterwards, the sFit is performed on these resampled datasets, which leads to a distribution for the respective parameter. Uncertainties can then be determined by the distances between the 15.87th and 84.13th percentiles and the median of the distribution. The bootstrapping procedure is performed both for the nominal fit and the sFit on the data sample and yields compatible uncertainties which agree with the results of the nominal fit.

5.3.4 Fits on simulated samples

Fully simulated samples of each 4 million $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ decays are treated in a similar way as the nominal data samples to eventually extract the CP parameters. This allows to further validate the maximum-likelihood fit with another sample type. Additionally, this procedure serves to verify other aspects, such as the portability of the decay-time resolution and acceptance parameterisation between $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ samples and the portability of the tagging calibration between calibration and signal channels.

Table 5.4 lists the values of physics parameters used in the generation. For the CP observables, these are based on an older theory prediction. All selection requirements are applied to the samples. Additionally, the surviving candidates are required to be truth-matched to assure only signal decays remain in the sample. In terms of the flavour tagging algorithms, the respective MC tuning has been used.

Decay-time resolution and acceptance on simulated samples

In contrast to the procedure described in Section 4.4.2, the decay-time resolution is determined on the $B_s^0 \rightarrow J/\psi K_S^0$ sample by directly fitting the distribution of the decay-time residuals $t - t'$, where t' is the generated true decay time. The identical resolution model as on the data sample is used to describe the residuals. As the sample only contains true $B_s^0 \rightarrow J/\psi K_S^0$ events, no additional long-lived component to the core resolution function is necessary.

The decay-time acceptance is determined on the sample of simulated $B^0 \rightarrow J/\psi K_S^0$ decays using the identical parameterisation as in Section 4.5. In the fit, the B^0 lifetime τ_{B^0} is fixed to the generation value. Results for the resolution and acceptance parameters as well as the excellent agreement of the resolution parameterisation between $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ decays are discussed in Appendix A.4.1.

Table 5.4 – Values of parameters used in the generation of the fully simulated samples.

Parameter	Value	Unit
Γ_s	0.6623	ps ⁻¹
$\Delta\Gamma_s$	0.0938	ps ⁻¹
Δm_s	17.761	ps ⁻¹
τ_{B^0}	1.5191	ps
Δm_d	0.502	ps ⁻¹
$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$	0.9437	
$C_{B_s^0 \rightarrow J/\psi K_S^0}$	-0.0118	
$S_{B_s^0 \rightarrow J/\psi K_S^0}$	-0.3306	
$C_{B^0 \rightarrow J/\psi K_S^0}$	0.0	
$S_{B^0 \rightarrow J/\psi K_S^0}$	0.6997	

Flavour tagging calibration on simulated samples

As for the analysis performed on the data sample, the mistag probability predictions η_i of the flavour tagging algorithms in the simulated sample need to be calibrated into the measured mistag probabilities ω_i . The calibration parameters for simulated samples are a priori not identical to those on data.

In case of the OS combination, the parameters are determined on a simulated sample of $B^+ \rightarrow J/\psi K^+$. For the $B_s^0 \rightarrow J/\psi K_S^0$ sample, the SSK tagger is calibrated directly on this sample by comparing the tagging response with the particle's production flavour. To achieve this, the sample is divided into equally filled bins of η^{SSK} . For each bin, ω^{SSK} is determined by the fraction of incorrectly tagged events over all tagged events. The calibration function is then fitted to these data points. In case of the $B^0 \rightarrow J/\psi K_S^0$ decays, the same calibration approach is followed for the SSK tagger with a simulated sample of $B^0 \rightarrow J/\psi K^{*0}$ decays. A flipped tagging response is assumed here (see Section 4.2.6). The resulting calibration parameters are discussed in Appendix A.4.2.

Results on the simulated $B_s^0 \rightarrow J/\psi K_S^0$ signal sample

To determine the CP parameters, the maximum-likelihood fit is performed on the simulated $B_s^0 \rightarrow J/\psi K_S^0$ signal sample. The procedure is identical to the one on the data sample with the exception of not fitting the reconstructed mass distribution, fixing all acceptance parameters to the previously extracted values, and not including Gaussian constraints. Alternatively, the externally provided parameters are fixed to the respective generation value. Table 5.5 lists the results for the CP parameters, showing excellent agreement with the generation values. Figure 5.7 shows

the fitted decay-time distribution. Furthermore, Table A.13 in Appendix A.4.3 lists the results for the fitted event yields.

Table 5.5 – Resulting CP parameters for the fit to the simulated $B_s^0 \rightarrow J/\psi K_S^0$ signal sample. The deviation from the generation value in units of the fit uncertainty is shown as p .

Parameter	Result	p
$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$	0.92 ± 0.04	-0.49
$C_{B_s^0 \rightarrow J/\psi K_S^0}$	-0.021 ± 0.022	-0.40
$S_{B_s^0 \rightarrow J/\psi K_S^0}$	-0.327 ± 0.022	0.14

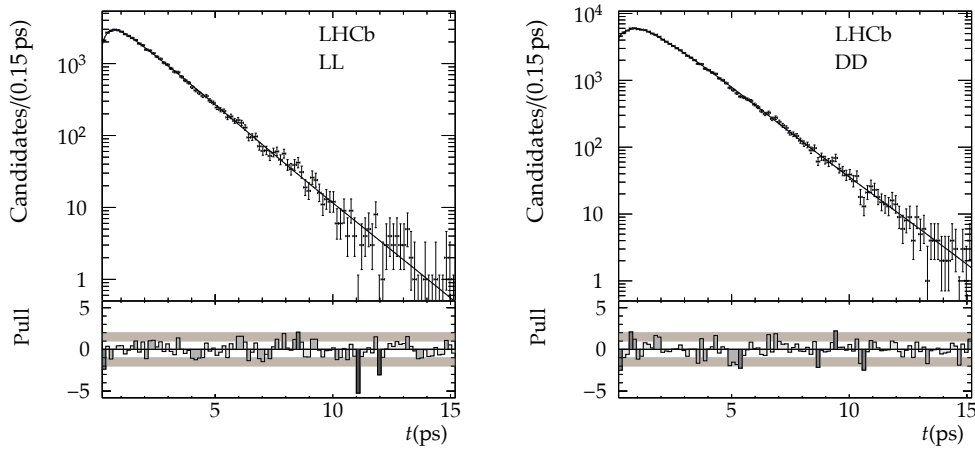


Figure 5.7 – Decay-time distribution in the simulated $B_s^0 \rightarrow J/\psi K_S^0$ signal sample for the LL (left) and DD (right) category. The fitted PDF is shown as black solid line.

Results on a simulated compound sample

Additionally to the simulated $B_s^0 \rightarrow J/\psi K_S^0$ signal sample, a fit to a compound sample is performed. This so-called cocktail MC sample is designed to come close to the nominal data sample, consisting of both the simulated $B_s^0 \rightarrow J/\psi K_S^0$ and simulated $B^0 \rightarrow J/\psi K_S^0$ samples and additional events representing combinatorial background, with the latter being generated from the background PDF as a pseudo-experiment. The $B_s^0 \rightarrow J/\psi K_S^0$ sample has been scaled so that the ratio of the numbers of B_s^0 to B^0 decays corresponds to the one observed in the data sample. Likewise, the combinatorial sample is constructed to maintain the observed signal to background ratio.

The fit to the simulated compound sample is performed like the nominal fit

on data with the exception of not including Gaussian constraints. Again, the externally provided parameters are fixed to the generation values. Table 5.6 lists the results for the CP parameters, again showing excellent agreement with the generation values. Figure 5.8 shows the mass and decay-time distribution with the fitted PDF. Results for the nuisance parameters are given in Tables A.14 and A.15 in Appendix A.4.3.

Table 5.6 – Resulting CP parameters for the fit to the simulated compound sample. The deviation from the generation value in units of the fit uncertainty is shown as p .

Parameter	Result	p
$C_{B_s^0 \rightarrow J/\psi K_S^0}$	-0.022 ± 0.022	1.01
$S_{B_s^0 \rightarrow J/\psi K_S^0}$	0.711 ± 0.022	-0.51
$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$	0.3 \pm 0.4	1.58
$C_{B_s^0 \rightarrow J/\psi K_S^0}$	-0.13 ± 0.25	0.49
$S_{B_s^0 \rightarrow J/\psi K_S^0}$	-0.41 ± 0.24	0.31

5.3.5 Likelihood profile scans

Supplementary to the other validation procedures which aim at repeating the fit with different configurations, one can also directly investigate the shape of the negative log-likelihood as a function of the CP parameters. So-called likelihood profile plots scan one or more parameters in the region of interest by fixing these parameters to the respective values and repeating the likelihood fit. The difference $\Delta \ln L$ between the log-likelihoods at the scan point and of the nominal fit can then be evaluated for all scan points and visualised. This helps to gain confidence that no local minima exist near the obtained minimum and also provides information about the estimated uncertainties by determining confidence intervals at certain levels of $\Delta \ln L$. Furthermore, multidimensional likelihood profiles can visualise possible correlations of the parameters. Figure 5.9 shows the two-dimensional likelihood profile scans for the CP parameters of the $B_s^0 \rightarrow J/\psi K_S^0$ decays. These show no conspicuous features and no evident large correlation of the parameters (cf. Figure 5.3). Additionally, the profiles confirm that the uncertainties for $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$ are asymmetric around the determined fit result.

5.3.6 Scan of the CP parameter space

The previously described studies with pseudo-experiments and fully simulated samples show that the likelihood fit can correctly estimate the CP parameters in

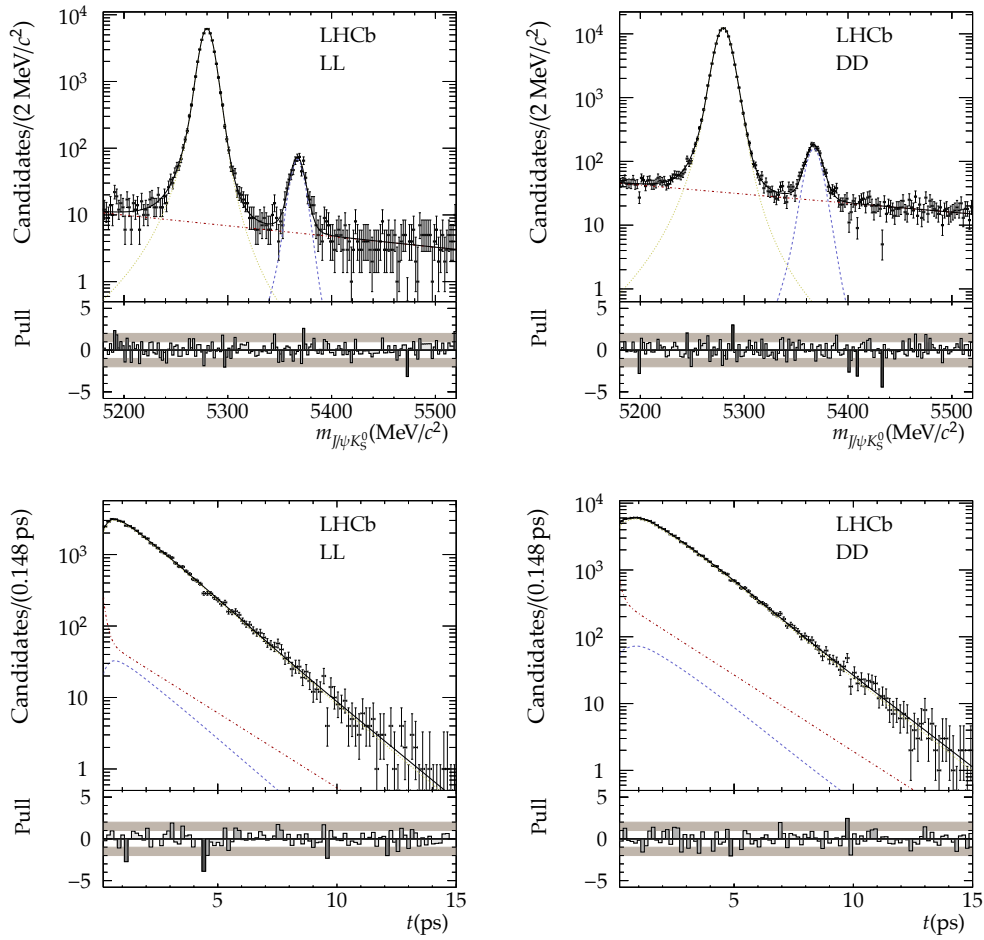


Figure 5.8 – Mass (top) and decay-time (bottom) distributions of the simulated compound sample for the LL (left) and DD (right) subsample. The projection of the fitted PDF (black, solid) shows the $B_s^0 \rightarrow J/\psi K_S^0$ (blue, dashed) and the $B^0 \rightarrow J/\psi K_S^0$ (yellow, dotted) component as well as the combinatorial background (red, dash-dotted).

5 Measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$

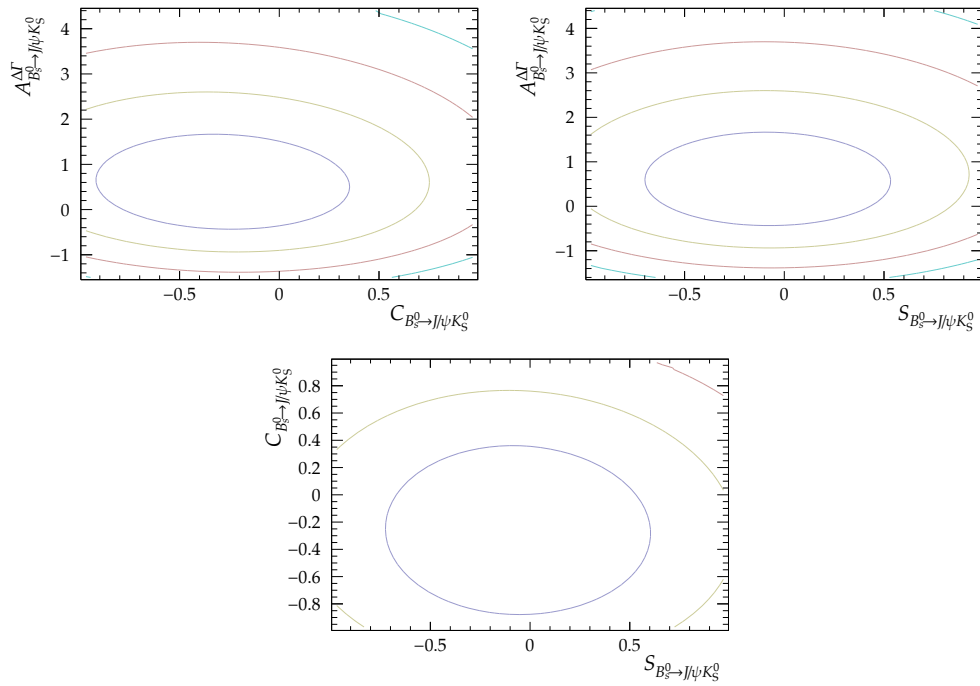


Figure 5.9 – Two-dimensional likelihood profiles for the $B_s^0 \rightarrow J/\psi K_S^0$ CP parameters obtained on the nominal data sample. The coloured contours correspond to the 68.27 % (blue), 95.45 % (yellow), 99.73 % (red) and 99.9 % (cyan) confidence levels.

case the values from theory predictions are assumed in the generation. However, as this analysis is the first measurement of the CP parameters in $B_s^0 \rightarrow J/\psi K_S^0$, the likelihood fit has to be able to correctly estimate the parameters independently of their true value. To investigate this, more studies based on pseudo-experiments are conducted, each with a different set of true CP parameters, where the assumed values are randomly chosen in the allowed region of ± 1 . The samples are fit with the identical configuration as in Section 5.3.1.

To evaluate the parameter and uncertainty estimation, the pull distributions of the CP parameters are fitted with a Gaussian function. In all studies, the resulting fitted means are consistent with zero, confirming unbiased parameter estimation. However, several fitted pull widths, especially for $C_{B_s^0 \rightarrow J/\psi K_S^0}$ and $S_{B_s^0 \rightarrow J/\psi K_S^0}$ in case the generation values are close to the boundaries of ± 1 , are significantly smaller than the expectation of one. This indicates an overestimation of the uncertainties or so-called over-coverage. Table 5.7 shows the assumed CP parameters in each set-up and the resulting pull widths. In case of significant over-coverage, specialised techniques can be used to determine confidence intervals with correct coverage.

Table 5.7 – Fitted pull widths s_{pull} of the $B_s^0 \rightarrow J/\psi K_S^0$ CP parameters from toy MC studies with different generation values for the CP parameters. Deviations from the expectation for s_{pull} are highlighted in case the deviation in units of the respective uncertainty is larger than 3σ (dark grey) or 5σ (red).

Generation value			Pull width s_{pull}		
$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$	$C_{B_s^0 \rightarrow J/\psi K_S^0}$	$S_{B_s^0 \rightarrow J/\psi K_S^0}$	$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$	$C_{B_s^0 \rightarrow J/\psi K_S^0}$	$S_{B_s^0 \rightarrow J/\psi K_S^0}$
-0.766	0.432	0.477	0.897 \pm 0.028	0.967 \pm 0.031	0.920 \pm 0.029
0.208	-0.332	0.920	0.919 \pm 0.029	0.934 \pm 0.030	0.877 \pm 0.028
-0.522	0.570	0.634	0.972 \pm 0.031	0.915 \pm 0.029	0.898 \pm 0.029
0.762	0.642	-0.084	1.065 \pm 0.034	0.978 \pm 0.031	1.021 \pm 0.032
-0.201	-0.668	0.717	0.980 \pm 0.031	0.840 \pm 0.027	0.831 \pm 0.027
0.088	0.483	-0.871	0.910 \pm 0.029	0.840 \pm 0.027	0.869 \pm 0.028
-0.225	-0.959	-0.174	0.940 \pm 0.030	0.826 \pm 0.026	0.900 \pm 0.029
0.827	-0.393	0.402	1.040 \pm 0.033	0.957 \pm 0.030	1.013 \pm 0.032
0.304	-0.922	0.238	0.980 \pm 0.031	0.839 \pm 0.027	0.923 \pm 0.029
-0.145	-0.795	-0.589	0.953 \pm 0.030	0.894 \pm 0.028	0.868 \pm 0.028
0.081	-0.824	0.417	0.912 \pm 0.029	0.914 \pm 0.029	0.922 \pm 0.029
-0.985	0.230	-0.549	0.865 \pm 0.028	0.937 \pm 0.030	0.870 \pm 0.028
-0.266	-0.520	0.758	0.897 \pm 0.028	0.906 \pm 0.029	0.825 \pm 0.026

5.3.7 Feldman–Cousins method

The Feldman–Cousins (FC) method allows to avoid the previously described over-coverage by determining confidence intervals with correct coverage for the CP parameters [104]. This method uses the Neyman construction of confidence belts to derive confidence intervals $[\lambda_1, \lambda_2]$ at a given confidence level α for the parameter of interest λ by utilising likelihood ratio ordering. This allows the interpretation that in a fraction α of experiments the unknown true value λ_{true} is contained inside the interval $[\lambda_1, \lambda_2]$, which itself is different in each experiment. The method guarantees non-empty frequentist confidence intervals with correct coverage even for parameters near boundaries.

In practice it is useful to consider so-called 1 – CL plots which allow the determination of $[\lambda_1, \lambda_2]$ in dependence of the confidence level α . Therefore, the parameter λ is scanned in the region of interest and at each point λ_0 1 – CL is determined. The procedure starts with the reference fit on data in which λ is varied, resulting in the log-likelihood $\ln L_{\text{best}}$. For each value λ_0 the procedure is then [105]:

1. The fit is repeated with fixed $\lambda = \lambda_0$, yielding the log-likelihood $\ln L_{\text{fixed}}$, and $\Delta\chi_{\text{data}}^2 = 2 \ln L_{\text{best}} - 2 \ln L_{\text{fixed}}$ is computed.
2. A sample of N_{toy} pseudo-experiments is generated with $\lambda = \lambda_0$. For the generation, all nuisance parameters are set to the values obtained in the fit to the data sample with fixed $\lambda = \lambda_0$. Each toy MC sample is fitted once with varied λ and once with fixed $\lambda = \lambda_0$. Consequently, these fits yield $\ln L_{\text{float}}$ and $\ln L_{\text{fixed}}$. Analogous to the previous step, $\Delta\chi_{\text{toy}}^2 = 2 \ln L_{\text{float}} - 2 \ln L_{\text{fixed}}$ is calculated for each sample.
3. The 1 – CL = $1 - \alpha(\lambda_0)$ is calculated as the fraction of toy MC fits in which $\Delta\chi_{\text{toy}}^2$ exceeds $\Delta\chi_{\text{data}}^2$:

$$1 - \alpha(\lambda_0) = \frac{N(\Delta\chi_{\text{data}}^2 < \Delta\chi_{\text{toy}}^2)}{N_{\text{toy}}}. \quad (5.13)$$

This method obviously requires large amounts of time-consuming fits to pseudo-experiments. Alternatively, Wilks' Theorem allows to determine confidence intervals by avoiding pseudo-experiments altogether. Here, 1 – CL can be determined from the cumulative distribution function of the χ^2 distribution $P_{\chi^2}(t)$ as

$$1 - \alpha(\lambda_0) = \int_{\Delta\chi_{\text{data}}^2}^{\infty} P_{\chi^2}(t) dt. \quad (5.14)$$

The underlying assumption is that $\Delta\chi_{\text{data}}^2$ asymptotically follows a χ^2 distribution, which is not justified in the presence of boundaries. Furthermore, the MINOS

algorithm of the MINUIT2 minimisation package by construction yields identical results for the 68.3 % confidence intervals. As the uncertainty overestimation observed in Section 5.3.6 is present although MINOS was used to determine asymmetric uncertainties, this simplification is not sufficient to treat the over-coverage.

In order to perform the FC method for all $B_s^0 \rightarrow J/\psi K_S^0$ CP parameters with limited computing resources, a different fit model is used. Only the three $B_s^0 \rightarrow J/\psi K_S^0$ CP parameters, Gaussian constrained parameters describing the $B_s^0 \rightarrow J/\psi K_S^0$ decays, and nuisance parameters showing a sizeable correlation with the CP parameters are varied in the fits. This reduces the fit to nine varied parameters: $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$, $C_{B_s^0 \rightarrow J/\psi K_S^0}$, $S_{B_s^0 \rightarrow J/\psi K_S^0}$, Γ_s , $\Delta\Gamma_s$, Δm_s , τ_{B^0} , $A_{P,B_s^0}^{\text{LL}}$, and $A_{P,B_s^0}^{\text{DD}}$. Furthermore, systematic uncertainties can be incorporated into the confidence intervals by adding these as additional offsets to the CP parameters, where the offsets themselves are constrained with Gaussian functions with a mean of zero and the width representing the respective systematic uncertainty. The additional Gaussian function will enlarge the resulting statistical uncertainty accordingly.

The FC method is successfully verified on a single pseudo-experiment for CP parameters affected by over-coverage. In this set-up the widths of the obtained 68.3 % FC confidence intervals with respect to the intervals defined by the fit uncertainties are in excellent agreement with the observed uncertainty overestimation, confirming that the FC method is able to provide confidence intervals with correct coverage. Furthermore, the 68.3 % confidence intervals obtained from Wilks' Theorem are in perfect agreement with the MINOS uncertainties.

Results of the FC method for the nominal measurement including the systematic uncertainties are discussed in Section 5.5.1.

5.4 Systematic uncertainties

The likelihood fit is designed to properly describe the data and take uncertainties on all parameters correctly into account, including the limited knowledge about externally provided parameters through Gaussian constraints. However, not all effects possibly influencing the measurement can be accounted for directly in the fit so that systematic uncertainties need to be determined. For instance, the uncertainty whether the chosen parameterisation in the likelihood fit is correct should be reflected in a systematic uncertainty. Related to this is the limited knowledge about fixed parameters. A different source of systematic uncertainties are higher-order effects which are not described by the fit. Effects studied include a mass-dependent mass resolution, a correlation of mass and decay time, a momentum and decay-length scale uncertainty, and a contribution from $B_c^+ \rightarrow B_s^0 \pi^+$ decays. The ignorance of these effects could bias the measurement, so that their influence needs to be evaluated and if necessary be accounted for in additional

systematic uncertainties.

For each evaluated effect, the systematic uncertainties are simultaneously determined for all five CP parameters as well as for the LL and DD ratios of $B_s^0 \rightarrow J/\psi K_S^0$ to $B^0 \rightarrow J/\psi K_S^0$ event yields. The latter are referred to as

$$R^k = \frac{N_{B_s^0}^k}{N_{B^0}^k}. \quad (5.15)$$

5.4.1 Method to evaluate systematic uncertainties

Unless noted otherwise, all systematic uncertainties are evaluated with the following procedure: A sample of 1000 pseudo-experiments is generated with the nominal set-up as described in Section 5.3.1. To test the influence of the alternative model A, each sample is first fitted with the nominal likelihood fit model N and subsequently with model A. For the parameter λ , this yields two results, λ_N and λ_A . The distribution of differences $\delta_\lambda = \lambda_N - \lambda_A$ is then fitted with a Gaussian function and the systematic uncertainty $\sigma_{\text{syst},\lambda}$ is determined as

$$\sigma_{\text{syst},\lambda} = \sqrt{\mu_{\delta_\lambda}^2 + s_{\delta_\lambda}^2}, \quad (5.16)$$

where μ_{δ_λ} is the respective fitted mean and s_{δ_λ} the fitted width of the Gaussian function. The mean μ_{δ_λ} quantifies the bias on λ from choosing model N over A, while the width s_{δ_λ} represents the change in the uncertainty of λ between both models. Consequently, this is a conservative approach. However, as the statistical precision of the $B_s^0 \rightarrow J/\psi K_S^0$ CP parameters is clearly limited, this cautious approach is justified. For illustration, Figure 5.10 shows an example δ distribution.

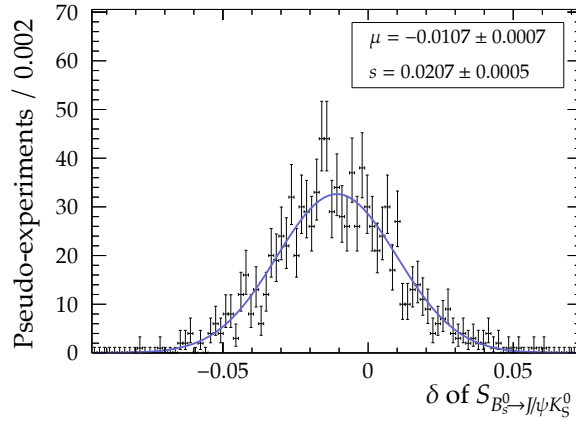


Figure 5.10 – Example of a δ distribution for CP parameter $S_{B_s^0 \rightarrow J/\psi K_S^0}$ in a scenario where different tagging parameters are compared to the nominal fit model.

The performed likelihood fits to the pseudo-experiments are simplified with respect to the nominal fit in order to reduce the amount of necessary computing resources. Unless otherwise specified, the acceptance coefficients are fixed in the fits and no Gaussian constraints are used. Instead, the externally provided parameters are fixed to the values in Table 5.3. All other nuisance parameters are varied in the fits.

5.4.2 Choice of parameterisation

The parameterisation of the likelihood fit depends on the chosen PDFs as well as on a set of parameters being fixed in the fit. While the choice of PDFs is motivated and justified, it is unknown whether alternative models describe the data even better. By testing a set of different PDFs with the previously described approach, systematic uncertainties can be derived for the choice of parameterisation. The situation is similar for fixed parameters which are only known with limited precision. Therefore, systematic uncertainties are determined by varying the parameters within their respective uncertainties.

Uncertainties are evaluated for different aspects of the analysis, where different variations are tested and for each parameter λ the largest occurring deviation $\sigma_{\text{sys},\lambda}$ is selected as systematic uncertainty.

Parameterisation of the reconstructed mass

To study systematic uncertainties resulting from the parameterisation of the reconstructed mass, the following variations are studied: The fixed parameters describing the tails of the Hypatia distribution are varied within their uncertainties so that the largest and smallest tails with respect to distribution width are achieved. Furthermore, the Hypatia distribution is exchanged by a sum of two Crystal Ball (CB) functions for which the tail parameters are likewise determined on simulated samples. Finally, the mass PDF of the combinatorial background is exchanged by a second-order polynomial function.

The $B^0 \rightarrow J/\psi K_S^0$ CP parameters are nearly unaffected by any variation; the largest effect comes from the variation using the double CB model. For both the $B_S^0 \rightarrow J/\psi K_S^0$ CP parameters and the yield ratios R^k , the systematic uncertainty results from exchanging the background PDF with a polynomial function.

Parameterisation of the decay-time resolution

For the decay-time resolution, the fixed parameters are varied within their respective uncertainties so that the smallest and largest possible sample resolution dilution $\langle \mathcal{D}_{\text{res}} \rangle$ (see Equation (4.28)) is obtained. Furthermore, other LHCb measurements observe a constant offset in the calibration functions $\delta_i(\sigma_i)$ (see Equation (4.25)) [12,

91]. Due to limited statistics, this offset is not parameterised in this measurement. To assess its influence, the offset observed in the measurement CP violation in $B^0 \rightarrow J/\psi K_S^0$ is adopted into the resolution model and evaluated [106].

For all parameters, the systematic uncertainty results from the variation with the smallest resolution dilution. Both the $B^0 \rightarrow J/\psi K_S^0$ CP parameters as well as the yield ratios R^k are virtually not affected by any variation, which is expected as the $B^0-\bar{B}^0$ oscillation period is large compared to the decay-time resolution and the yield estimation does not depend on the resolution at all. For the oscillation-dependent parameters $C_{B_s^0 \rightarrow J/\psi K_S^0}$ and $S_{B_s^0 \rightarrow J/\psi K_S^0}$, the uncertainty on the decay-time resolution parameterisation results in the largest systematic uncertainty. This is within expectation and results from the high oscillation frequency in the $B_s^0-\bar{B}_s^0$ meson system.

Parameterisation of the decay-time acceptance

The position and number of cubic spline knots used in the parameterisation of the decay-time acceptance is chosen empirically to describe the observed shape. However, a different configuration might be equally well suited to describe the distribution. To study the systematic uncertainty arising from this arbitrary choice, both the number and position of the knots are varied. Four alternative scenarios are evaluated, being summarised in Table 5.8 and visualised in Figure 5.11. In all fits, the acceptance coefficients $c_{\text{acc},i}$ are varied. The largest systematic influence on all parameters comes from the first variation with the smallest number of spline knots. Only $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$ shows a sizeable systematic uncertainty.

Table 5.8 – Position of knots used to describe the alternative decay-time acceptance parameterisations. Variations 1 and 2 change the number of knots with respect to the nominal configuration, while variations 3 and 4 change the position of the knots. The configuration of the nominal fit is given for comparison.

Nominal	Variation 1	Variation 2	Variation 3	Variation 4
0.2	0.2	0.2	0.2	0.2
0.5	1.0	0.5	0.75	0.35
1.0	5.0	0.8	1.25	0.75
2.0	11.5	1.5	2.5	1.75
8.0	15.0	2.0	9.0	7.0
15.0	—	6.0	15.0	15.0
—	—	12.0	—	—
—	—	15.0	—	—

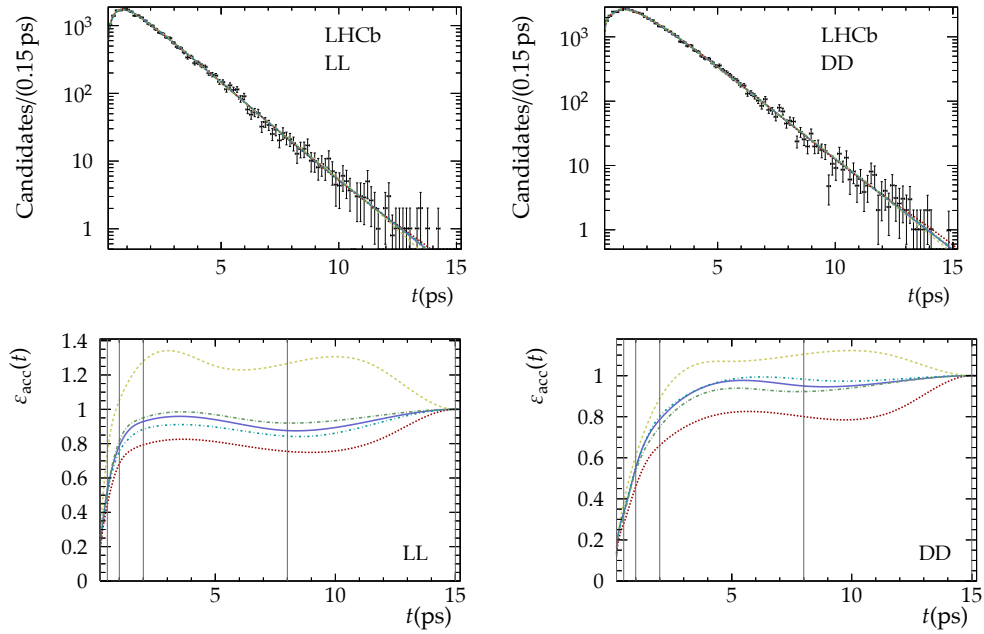


Figure 5.11 – Decay-time distribution of $B^0 \rightarrow J/\psi K_S^0$ signal candidates obtained using sWeights and PDF with spline acceptance (top), and efficiency function $\varepsilon_{\text{acc}}(t)$ describing the decay-time acceptance (bottom) for the LL (left) and DD (right) subsample. Shown are the nominal acceptance function (blue, solid), variation 1 (yellow, dashed), variation 2 (red, dotted), variation 3 (cyan, dash-dotted), and variation 4 (green, long-dash-dotted).

Parameterisation of the flavour tagging calibration

Both the flavour tagging asymmetries Δp_0 and Δp_1 , and the uncertainties associated with the flavour tagging calibration parameters are ignored in the nominal fit. The choice of these parameters determines the tagging dilution \mathcal{D}_{tag} of the sample and therefore influences the CP parameters S and C . To determine a systematic uncertainty associated with the choice of the tagging calibration parameters, first all parameters, including Δp_0 and Δp_1 , are varied randomly within their uncertainties. Afterwards, the sample tagging dilution \mathcal{D}_{tag} is recalculated for the given choice of parameters. This is repeated until the sets of parameters are found which result in the smallest and largest possible \mathcal{D}_{tag} . For these two parameter sets, the systematic uncertainties are evaluated with the toy-based approach. The resulting dilutions are specified in Table 5.9.

Table 5.9 – Resulting sample flavour tagging dilutions \mathcal{D}_{tag} for the nominal tagging calibration parameters and the variations resulting in the smallest and largest possible \mathcal{D}_{tag} after varying the parameters within their uncertainties.

	$B_s^0 \rightarrow J/\psi K_S^0$		$B^0 \rightarrow J/\psi K_S^0$	
	LL	DD	LL	DD
Nominal	0.26	0.27	0.214	0.212
Minimal	0.25	0.26	0.209	0.207
Maximal	0.27	0.28	0.226	0.224

In case of the $B^0 \rightarrow J/\psi K_S^0$ component, the recalculation of \mathcal{D}_{tag} is performed on both the nominal data sample as well as on pseudo-experiments. Both approaches result in identical sets of parameters and yield compatible dilutions. For the $B_s^0 \rightarrow J/\psi K_S^0$ component, the data sample does not contain enough statistics to conclusively compare the \mathcal{D}_{tag} results. Therefore, the approach based on pseudo-experiments is chosen, where the number of $B_s^0 \rightarrow J/\psi K_S^0$ decays is increased by a factor of 100.

As expected, the yield ratios R^k and $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$ are practically not affected by the variations. For the $B^0 \rightarrow J/\psi K_S^0$ CP parameters the changes of the flavour tagging calibration form the dominant systematic uncertainty. As expected, both $C_{B_s^0 \rightarrow J/\psi K_S^0}$ and $S_{B_s^0 \rightarrow J/\psi K_S^0}$ are considerably affected by the variations.

5.4.3 Ignored contributions from non-prompt B_s^0 production

The reconstruction of $B \rightarrow J/\psi K_S^0$ candidates assumes that the B mesons originate directly from the PV. However, the B_s^0 candidates could also be produced in $B_c^+ \rightarrow B_s^0 X$ decays, where the finite B_c^+ lifetime of (0.507 ± 0.009) ps [39] could lead

to a displaced B_s^0 production vertex with respect to the PV. This would affect the B_s^0 decay time and thus bias the measurement of the CP parameters. With the recent observation of the $B_c^+ \rightarrow B_s^0 \pi^+$ decay [107], the rate f of $B_s^0 \rightarrow J/\psi K_S^0$ decays from B_c^+ decays can be estimated. The calculation follows the one made for the measurement of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$ decays and yields $f = 0.8\%$ [108]. This translates into an expected number of at most 3 (LL) and 5 (DD) $B_s^0 \rightarrow J/\psi K_S^0$ decays in the nominal data sample originating from B_c^+ decays. As these numbers are small compared to the statistical uncertainties of the measured $B_s^0 \rightarrow J/\psi K_S^0$ event yields, no additional systematic uncertainties are assigned due to non-prompt B_s^0 production.

5.4.4 Mass resolution differences

The likelihood fit parameterises the $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ mass shapes identically. However, other LHCb measurements show that the width σ of the mass distribution scales with the reconstructed mass. To study the systematic uncertainty associated with this effect, a scale factor between the widths is introduced in a fit to the reconstructed mass as

$$\sigma_{B_s^0} = s_\sigma \sigma_{B^0}. \quad (5.17)$$

The scale factor s_σ is determined both on simulated signal samples of $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ decays and on the data sample, with the results summarised in Table 5.10. The results on data and simulation are compatible, although the uncertainties on the data sample are too large to show a significant difference between $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$. Therefore, for the study of an associated systematic uncertainty the result obtained on the simulated samples are used. The method using pseudo-experiments is performed, where the fit with the alternative model includes the fixed scale parameter s_σ . Neglecting s_σ mainly affects the yield ratios R^k and $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$.

Table 5.10 – Results for the relative scale factor s_σ between the mass distribution widths of the $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ components as determined on the data sample and fully simulated samples.

	LL	DD
Data	1.087 ± 0.071	1.000 ± 0.014
Simulation	1.0515 ± 0.0042	1.0212 ± 0.0026

5.4.5 Momentum and decay-length scale

The momentum scale uncertainty at LHCb is determined to be at most 0.15% [109]. This affects both the candidate's momenta $|\mathbf{p}|$ and reconstructed mass m so that the uncertainty cancels in the reconstructed decay time $t \propto m/|\mathbf{p}|$ in first order. Higher-order effects, though, lead to a relative shift on t of $\mathcal{O}(10^{-4})$ [23]. Furthermore, the uncertainty on the decay-length or z-scale is estimated as 0.022% [109]. Both effects only influence the determination of Γ_s , $\Delta\Gamma_s$, Δm_s , τ_{B^0} , and Δm_d . As all of these parameters are constrained to the respective world averages in the likelihood fit, no systematic uncertainties are assigned.

5.4.6 Correlation between mass and decay-time resolution

Simulated signal samples of $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ decays show a correlation of the decay-time resolution mean μ_r (see Equation (4.24)) and the reconstructed candidate mass m . This effect can be parameterised as

$$\mu_r(m) = c_0 + c_1 (m - m_{\min}) , \quad (5.18)$$

where $m_{\min} = 5180 \text{ MeV}/c^2$. Results for c_0 and c_1 from fits to the simulated samples are listed in Tables 5.11 and 5.12.

Table 5.11 – Parameters describing the correlation between the resolution mean μ_r and the reconstructed mass for $B_s^0 \rightarrow J/\psi K_S^0$ decays as obtained on a simulated sample.

Parameter	LL	DD	Unit
c_0	-0.206 ± 0.085	0.021 ± 0.079	ps
c_1	0.00111 ± 0.00046	-0.00011 ± 0.00043	ps/(MeV/ c^2)

Table 5.12 – Parameters describing the correlation between the resolution mean μ_r and the reconstructed mass for $B^0 \rightarrow J/\psi K_S^0$ decays as obtained on a simulated sample.

Parameter	LL	DD	Unit
c_0	-0.109 ± 0.043	0.010 ± 0.044	ps
c_1	0.00110 ± 0.00042	-0.00011 ± 0.00044	ps/(MeV/ c^2)

The nominal likelihood fit not only neglects this correlation, but also assumes $\mu_r = 0$ due to technical limitations (see Section 4.4.1). Therefore, systematic uncertainties are determined for both this simplification and for neglecting the previously described correlation. Samples of 1000 pseudo-experiments are generated with the correlation using the parameters obtained on simulated samples. Afterwards, these samples are fitted once with the nominal likelihood fit with $\mu_r = 0$ and

once with a fit incorporating the correlation. The method described in Section 5.4.1 is used to extract systematic uncertainties. Only the CP parameters $C_{B_s^0 \rightarrow J/\psi K_S^0}$ and $S_{B_s^0 \rightarrow J/\psi K_S^0}$ show a considerable systematic uncertainty, which is expected as the resolution parameterisation is especially crucial for the rapidly oscillating B_s^0 mesons.

5.4.7 Combined systematic uncertainty

All individual contributions to the systematic uncertainties and the combined systematic uncertainties for the CP parameters of the $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K_S^0$ decays as well as for the yield ratios R^k are summarised in Table 5.13.

Table 5.13 – Individual systematic uncertainties for each studied effect and combined systematic uncertainties of the five CP parameters and the yield ratios R^k .

Source	$A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$	$C_{B_s^0 \rightarrow J/\psi K_S^0}$	$S_{B_s^0 \rightarrow J/\psi K_S^0}$	$C_{B^0 \rightarrow J/\psi K_S^0}$	$S_{B^0 \rightarrow J/\psi K_S^0}$	R^{LL}	R^{DD}
Mass	0.045	0.009	0.009	—	0.001	$16 \cdot 10^{-5}$	$17 \cdot 10^{-5}$
Resolution	0.038	0.066	0.070	0.001	0.002	$1 \cdot 10^{-5}$	—
Acceptance	0.022	0.004	0.004	—	—	$1 \cdot 10^{-5}$	$1 \cdot 10^{-5}$
Tagging calibration	0.002	0.021	0.023	0.006	0.036	—	—
Mass width scale	0.010	0.005	0.006	—	—	$13 \cdot 10^{-5}$	$8 \cdot 10^{-5}$
Correlated $\mu_r(m)$	0.003	0.037	0.036	—	—	—	—
Combined	0.064	0.079	0.083	0.006	0.036	$20 \cdot 10^{-5}$	$19 \cdot 10^{-5}$

Except for $S_{B^0 \rightarrow J/\psi K_S^0}$, all combined systematic uncertainties are small compared to the respective statistical uncertainties. The largest systematic contributions to $C_{B_s^0 \rightarrow J/\psi K_S^0}$ and $S_{B_s^0 \rightarrow J/\psi K_S^0}$ emerge from the decay-time resolution, the correlation between mass and decay-time resolution, and the flavour tagging. All of these can be decreased in the future with more data and increased statistics. For instance, the decay-time resolution parameters suffer from limited statistics. With future data, a better determination and more sophisticated resolution models will be feasible, which in turn will improve the systematic uncertainty. The correlation of the mass and decay-time resolution requires more studies with larger simulated samples. If necessary, the effect can be accounted for in the fit model, allowing to avoid the systematic uncertainty altogether. In case of the flavour tagging, the according uncertainties are expected to reduce with more statistics in the calibration and control samples. Furthermore, higher-order effects like tagging asymmetries can be accounted for in the fit. For $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$, the leading systematic uncertainties result from the mass modelling and the decay-time acceptance, where for both effects additional data are likely to improve the uncertainties. The yield ratios R^{LL} and R^{DD} are mainly affected by the mass modelling and the mass resolution differences, where the latter can be better studied with more data and—if necessary—be incorporated into the fit model. Finally, $C_{B^0 \rightarrow J/\psi K_S^0}$ and $S_{B^0 \rightarrow J/\psi K_S^0}$ are mainly affected by the systematic uncertainty from the flavour tagging.

5.5 Final results

The measurement of CP violation in the $B_s^0 \rightarrow J/\psi K_S^0$ decay channel is performed on a dataset corresponding to 3 fb^{-1} of pp collisions collected by the LHCb experiment at centre-of-mass energies of $\sqrt{s} = 7 \text{ TeV}$ and $\sqrt{s} = 8 \text{ TeV}$. A total of around 900 $B_s^0 \rightarrow J/\psi K_S^0$ and nearly 80 000 $B^0 \rightarrow J/\psi K_S^0$ decays are observed. For the $B_s^0 \rightarrow J/\psi K_S^0$ decays, a resolution power of $\langle \mathcal{P}_{\text{res}} \rangle \approx 0.54$ and a total effective tagging efficiency of $\varepsilon_{\text{eff}} \approx 4\%$ are measured. The final results for the CP parameters in $B_s^0 \rightarrow J/\psi K_S^0$ are determined as

$$\begin{aligned} A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} &= 0.49 \pm_{0.65}^{0.77} (\text{stat}) \pm 0.06 (\text{syst}), \\ C_{B_s^0 \rightarrow J/\psi K_S^0} &= -0.28 \pm 0.41 (\text{stat}) \pm 0.08 (\text{syst}), \\ S_{B_s^0 \rightarrow J/\psi K_S^0} &= -0.08 \pm 0.40 (\text{stat}) \pm 0.08 (\text{syst}), \end{aligned}$$

with the statistical correlations between these parameters being

$$\begin{aligned} \rho(A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}, C_{B_s^0 \rightarrow J/\psi K_S^0}) &= -0.07, \\ \rho(A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}, S_{B_s^0 \rightarrow J/\psi K_S^0}) &= -0.01, \\ \rho(C_{B_s^0 \rightarrow J/\psi K_S^0}, S_{B_s^0 \rightarrow J/\psi K_S^0}) &= -0.06. \end{aligned}$$

Although the uncertainties do not yet allow to draw conclusions about the possible penguin diagram contributions in $B^0 \rightarrow J/\psi K_S^0$ decays, these results represent the first measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ decays. The CP parameters in $B^0 \rightarrow J/\psi K_S^0$ decays are simultaneously determined as

$$\begin{aligned} C_{B^0 \rightarrow J/\psi K_S^0} &= -0.028 \pm 0.035 (\text{stat}) \pm 0.006 (\text{syst}), \\ S_{B^0 \rightarrow J/\psi K_S^0} &= 0.719 \pm 0.034 (\text{stat}) \pm 0.036 (\text{syst}), \end{aligned}$$

being in excellent agreement with the latest dedicated LHCb measurement of CP violation in $B^0 \rightarrow J/\psi K_S^0$ [12]. Compared to that measurement, these results show larger statistical and systematic uncertainties. The main reason for this are the stringent selection requirements needed to isolate the small $B_s^0 \rightarrow J/\psi K_S^0$ component, which lead to a significant reduction of observed $B^0 \rightarrow J/\psi K_S^0$ decays with respect to the dedicated analysis.

Furthermore, the branching ratio of $B_s^0 \rightarrow J/\psi K_S^0$ to $B^0 \rightarrow J/\psi K_S^0$ decays is determined. The associated yield ratios R^k are measured as

$$\begin{aligned} R^{\text{LL}} &= 0.01104 \pm 0.00072 (\text{stat}) \pm 0.00020 (\text{syst}), \\ R^{\text{DD}} &= 0.01170 \pm 0.00059 (\text{stat}) \pm 0.00019 (\text{syst}). \end{aligned}$$

With the weighted average for LL and DD, this results in the branching ratio being determined as

$$\frac{\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0)}{\mathcal{B}(B^0 \rightarrow J/\psi K_S^0)} = 0.0431 \pm 0.0017 (\text{stat}) \pm 0.0012 (\text{syst}) \pm 0.0025 (f_s/f_d),$$

where the last uncertainty is due to the limited knowledge of the ratio f_s/f_d of B_s^0 to B^0 meson hadronisation fractions. This result is the most precise measurement of this quantity. The known branching fraction $\mathcal{B}(B^0 \rightarrow J/\psi K^0) = (8.97 \pm 0.35) \cdot 10^{-4}$ [19] allows to determine the $B_s^0 \rightarrow J/\psi K_S^0$ branching fraction as

$$\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0) = \left[1.93 \pm 0.08 (\text{stat}) \pm 0.05 (\text{syst}) \pm 0.11 (f_s/f_d) \pm 0.07 (\mathcal{B}(B^0 \rightarrow J/\psi K^0)) \right] \cdot 10^{-5},$$

where the last uncertainty results from the uncertainty on the $B^0 \rightarrow J/\psi K^0$ branching fraction.

5.5.1 Confidence intervals from the Feldman–Cousins method

As discussed in Section 5.3.7, the Feldman–Cousins method is performed to determine confidence intervals for the three CP parameters in $B_s^0 \rightarrow J/\psi K_S^0$ decays. The respective confidence contours are shown in Figure 5.12 and the according 68.3 % confidence intervals are determined as

$$\begin{aligned} A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} &\in [-0.16, 1.26], \\ C_{B_s^0 \rightarrow J/\psi K_S^0} &\in [-0.68, 0.14], \\ S_{B_s^0 \rightarrow J/\psi K_S^0} &\in [-0.48, 0.32]. \end{aligned}$$

As discussed before, these confidence intervals contain the systematic uncertainties. All results are in very good agreement with those obtained in the nominal likelihood fit, which is reflected in Figure 5.12 by the FC results agreeing with the profile likelihood result from Wilks' Theorem at 68.3 % confidence level. Parameter uncertainty overestimation of the fit as discussed in Section 5.3.6 becomes evident at higher confidence levels for $C_{B_s^0 \rightarrow J/\psi K_S^0}$ and $S_{B_s^0 \rightarrow J/\psi K_S^0}$. The according 95.5 % confidence intervals are determined as

$$\begin{aligned} A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} &\in [-0.68, 2.06], \\ C_{B_s^0 \rightarrow J/\psi K_S^0} &\in [-1.03, 0.52], \\ S_{B_s^0 \rightarrow J/\psi K_S^0} &\in [-0.83, 0.68]. \end{aligned}$$

5 Measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$

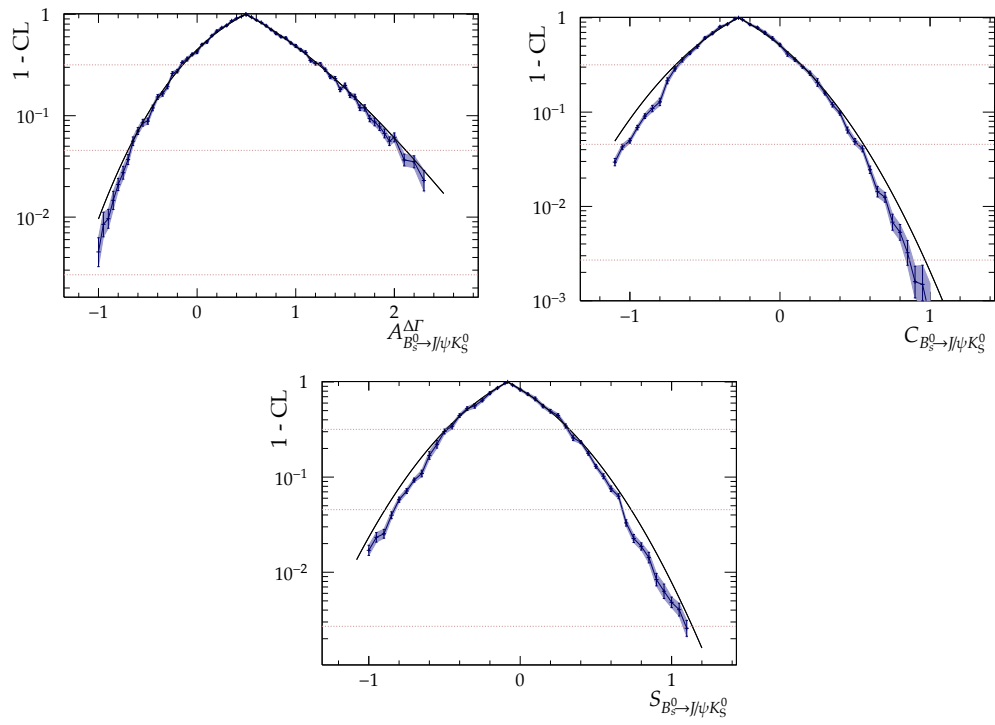


Figure 5.12 – Confidence level contours ($1 - \text{CL}$) for the CP parameters obtained with the FC method (blue). The results from applying Wilks' Theorem are shown in black.

6 Conclusion and outlook

Our understanding of fundamental particles and their interactions has seen tremendous advancements in the past decades. In the field of flavour physics, the B factories BABAR and Belle played a key role in this progress. LHCb ties up to that success by further improving the precision and establishing new fields of research. So far, the Standard Model of particle physics (SM) is a very successful theory being able to describe many observations in particle physics. However, as it has its shortcomings in some details, the search for new physics beyond the SM continues.

This work presents the first measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ decays, where the J/ψ and K_S^0 subsequently decay into $\mu^+ \mu^-$ and $\pi^+ \pi^-$. The time-dependent measurement studies CP violation in the interference of $B_s^0 - \bar{B}_s^0$ mixing and the decay. With the help of external input, the resulting CP observables can be utilised to extract the amount of contributions from penguin topologies with respect to the tree topology in $B_s^0 \rightarrow J/\psi K_S^0$ decays. These contributions can then be related to the according penguin contributions in the very similar decay channel $B^0 \rightarrow J/\psi K_S^0$. The latter is the so-called gold-plated channel to extract the CKM angle β . With the forthcoming precision achievable with LHCb and Belle II, the penguin contributions can no longer be neglected. Especially with regard to necessarily small new physics contributions in $B^0 \rightarrow J/\psi K_S^0$, it is essential to isolate the penguin contributions which could otherwise mask the contribution from new physics. Currently, the precision of the CP violating phase ϕ_d in the $B^0 - \bar{B}^0$ system is $\approx 2^\circ$, while estimations on the penguin-induced shift $\Delta\phi_d$ on ϕ_d are $\Delta\phi_d \lesssim \mathcal{O}(1^\circ)$. For the coming LHCb data-taking periods Run II and Run III, the sensitivities on the CP observables $S_{B^0 \rightarrow J/\psi K_S^0}$ and $C_{B^0 \rightarrow J/\psi K_S^0}$ are estimated to be $\lesssim 0.02$ and $\lesssim 0.006$, respectively [110]. This translates into uncertainties on ϕ_d of $\approx 1.6^\circ$ for Run II and $\approx 0.5^\circ$ for Run III.

The presented measurement comprises the sample preparation to isolate signal candidates from background, studies on the flavour tagging to correctly identify the production state of the B mesons, an examination of the decay-time resolution, the parameterisation of the decay-time acceptance, the multidimensional unbinned maximum-likelihood fit to determine the CP observables, various studies to validate the former, and finally the determination of systematic uncertainties. With approximately 900 $B_s^0 \rightarrow J/\psi K_S^0$ signal decays, the CP observables are determined

6 Conclusion and outlook

as

$$\begin{aligned} A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma} &= 0.49 \pm_{0.65}^{0.77} \text{ (stat)} \pm 0.06 \text{ (syst)}, \\ C_{B_s^0 \rightarrow J/\psi K_S^0} &= -0.28 \pm 0.41 \text{ (stat)} \pm 0.08 \text{ (syst)}, \\ S_{B_s^0 \rightarrow J/\psi K_S^0} &= -0.08 \pm 0.40 \text{ (stat)} \pm 0.08 \text{ (syst)}, \end{aligned}$$

with small correlations coefficients of $\rho \lesssim \mathcal{O}(0.07)$ between the parameters. The leading systematic uncertainty on $A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}$ results from the parameterisation of the mass distribution, while for $C_{B_s^0 \rightarrow J/\psi K_S^0}$ and $S_{B_s^0 \rightarrow J/\psi K_S^0}$ it arises from the decay-time resolution.

Furthermore, the branching ratio of $B_s^0 \rightarrow J/\psi K_S^0$ to $B^0 \rightarrow J/\psi K_S^0$ decays is measured as

$$\frac{\mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0)}{\mathcal{B}(B^0 \rightarrow J/\psi K_S^0)} = 0.0431 \pm 0.0017 \text{ (stat)} \pm 0.0012 \text{ (syst)} \pm 0.0025 (f_s/f_d),$$

where the last uncertainty is due to the limited knowledge of the ratio f_s/f_d of B_s^0 to B^0 meson hadronisation fractions. This allows to extract the $B_s^0 \rightarrow J/\psi K_S^0$ branching fraction as

$$\begin{aligned} \mathcal{B}(B_s^0 \rightarrow J/\psi K_S^0) &= \\ &\left[1.93 \pm 0.08 \text{ (stat)} \pm 0.05 \text{ (syst)} \pm 0.11 (f_s/f_d) \pm 0.07 (\mathcal{B}(B^0 \rightarrow J/\psi K^0)) \right] \cdot 10^{-5}, \end{aligned}$$

where the last uncertainty results from the uncertainty on the $B^0 \rightarrow J/\psi K^0$ branching fraction.

The results on the CP parameters are clearly statistically limited. Their precision does not yet allow sensible conclusions about the size of the penguin contributions in either $B_s^0 \rightarrow J/\psi K_S^0$ or $B^0 \rightarrow J/\psi K_S^0$ to be drawn. Therefore, the sensitivity of a future LHCb measurement of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ is estimated in a study with pseudo-experiments, where the candidate yields are scaled accordingly with the aspired integrated luminosities and expectations for the $b\bar{b}$ quark pair production cross-section. For Run II and Run III the integrated luminosities are estimated as at least 5 fb^{-1} and 50 fb^{-1} [74]. The $b\bar{b}$ production cross-sections are linearly scaled with the centre-of-mass energy from previous measurements [111, 112]. No assumptions are made about the changes of performance figures, for example the flavour tagging performance, as these are difficult to predict. Thus, the identical performance is assumed in the studies. This results in the estimated uncertainties for the respective parameters for Run II and Run III as

$$\begin{aligned} \text{Run II:} \quad & \sigma_{A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}} = 0.41, & \sigma_{C_{B_s^0 \rightarrow J/\psi K_S^0}} &= 0.22, & \sigma_{S_{B_s^0 \rightarrow J/\psi K_S^0}} &= 0.21, \\ \text{Run III:} \quad & \sigma_{A_{B_s^0 \rightarrow J/\psi K_S^0}^{\Delta\Gamma}} = 0.18, & \sigma_{C_{B_s^0 \rightarrow J/\psi K_S^0}} &= 0.07, & \sigma_{S_{B_s^0 \rightarrow J/\psi K_S^0}} &= 0.07. \end{aligned}$$

The current systematic uncertainties are dominated by limited statistics, which prevent the usage of more sophisticated models or limit the determination of fixed parameters. Most of these uncertainties are expected to decrease with more data from coming LHCb runs. Thus, it will most likely be possible to confine the systematic uncertainties to a level well below the statistical uncertainties.

It is finally evaluated how the uncertainties on the CP observables translate into uncertainties on the penguin-induced phase shift $\Delta\phi_d$ in $B^0 \rightarrow J/\psi K_S^0$. Therefore, samples for the CP parameters are drawn for the Run I, Run II, and Run III scenarios from normal distributions with the predictions from Ref. [36] as means and the measured and estimated future uncertainties as widths. These predictions correspond to an assumed shift of $\Delta\phi_d = -1.1^\circ$. Consequently, samples for ϕ_s and γ , which are required as input for the determination of $\Delta\phi_d$ (see Section 2.5.2), are drawn using the current world averages, where for ϕ_s Ref. [39] and for γ Ref. [28] is used. It is assumed that the respective uncertainties will decrease like the predicted uncertainties stated above. The samples are used to construct distributions of resulting $\Delta\phi_d$ from which uncertainties are determined. Influences from U-spin-symmetry breaking are neglected. This results in the expected Run I, Run II, and Run III uncertainties on $\Delta\phi_d$ as

$$\text{Run I: } \sigma_{\Delta\phi_d} = \begin{matrix} +1.6^\circ \\ -1.9^\circ \end{matrix}, \quad \text{Run II: } \sigma_{\Delta\phi_d} = \begin{matrix} +1.4^\circ \\ -1.1^\circ \end{matrix}, \quad \text{Run III: } \sigma_{\Delta\phi_d} = \begin{matrix} +0.28^\circ \\ -0.35^\circ \end{matrix}.$$

Assuming these, the experimental precision after Run II should allow a first estimation of the penguin contribution in $B^0 \rightarrow J/\psi K_S^0$. However, a more accurate determination will require sufficient data from Run III of LHCb. Moreover, improvements on the theory side, for example enhanced methods taking additional experimental data from other SU(3)-related decays into account, can further increase the sensitivity.

A Supplementary material

A.1 Tagging performance of the neural-net same-side kaon tagger

The neural-net SSK tagger is supposed to provide a higher tagging power than the cut-based variant. For reference, Table A.1 lists the tagging power for the $B_s^0 \rightarrow J/\psi K_S^0$ component in case the neural-net SSK tagger would be used. Specific calibration parameters p_0^{SSK, B_s^0} and p_1^{SSK, B_s^0} for the neural-net variant are applied. Both the exclusively SSK tagged and overlap category show an increase in tagging power with respect to the cut-based variant. This, however, comes at the cost of a diminished tagging power in the exclusive OS subsample so that the summed total effective tagging efficiency is not significantly increased. As the cut-based SSK tagging response on B^0 candidates is better understood, this variant is chosen for the analysis.

Tables A.2 and A.3 show the respective tagging efficiencies and average mistag probabilities for a setup involving the neural-net SSK tagger. In comparison to Tables 4.9 and 4.10 it becomes apparent that the neural-net SSK tagger shows a considerably higher ϵ_{tag} than the cut-based variant. This enlarges the SSK exclusive and overlap subsample which also shows that the decrease in OS exclusive tagging power in this scenario is simply due to fewer candidates being exclusively OS tagged. In contrast, the neural-net SSK tagger shows a significantly worse average mistag probability. Both effects combined do not result in an increase in overall tagging power.

Finally, Table A.4 shows the fractions of candidates which are exclusively tagged by one of the two SSK variants or by both simultaneously. Practically all candidates tagged by the cut-based variant are also tagged by the neural-net version.

A Supplementary material

Table A.1 – Effective tagging efficiencies $\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}(1 - 2\omega)^2$ for the $B_s^0 \rightarrow J/\psi K_S^0$ sample when exchanging the cut-based SSK tagger with the neural-net variant (for comparison only). Uncertainties are statistical.

Sample	$B_s^0 \rightarrow J/\psi K_S^0$	
	LL (%)	DD (%)
OS exclusive	1.39 ± 0.08	1.18 ± 0.05
SSK exclusive	0.41 ± 0.10	1.20 ± 0.17
OS+SSK overlap	2.00 ± 0.17	1.77 ± 0.24
Total	3.80 ± 0.21	4.15 ± 0.30

Table A.2 – Tagging efficiencies ε_{tag} for the $B_s^0 \rightarrow J/\psi K_S^0$ sample when exchanging the cut-based SSK tagger with the neural-net variant (for comparison only). Uncertainties are statistical.

Sample	$B_s^0 \rightarrow J/\psi K_S^0$	
	LL (%)	DD (%)
OS exclusive	11.5 ± 1.8	13.3 ± 1.4
SSK exclusive	29.3 ± 2.6	36.3 ± 2.0
OS+SSK overlap	17.1 ± 2.1	16.5 ± 1.5

Table A.3 – Average mistag probability ω for the OS and the neural-net SSK tagger for both the respective exclusively tagged subsamples as well as the overlap sample.

Sample	$B_s^0 \rightarrow J/\psi K_S^0$	
	LL (%)	DD (%)
OS exclusive	35.8 ± 1.7	37.5 ± 0.9
OS overlap	37.4 ± 1.3	38.4 ± 0.9
SSK exclusive	45.6 ± 0.4	43.8 ± 0.5
SSK overlap	42.1 ± 0.8	45.1 ± 0.5

Table A.4 – Fractions of $B_s^0 \rightarrow J/\psi K_S^0$ candidates tagged either exclusively by one of the two SSK taggers or by both simultaneously.

Sample	$B_s^0 \rightarrow J/\psi K_S^0$	
	LL (%)	DD (%)
neural-net SSK exclusive	36.1 ± 2.7	40.8 ± 2.0
cut-based SSK exclusive	0.5 ± 0.4	0.49 ± 0.28
neural-net + cut-based SSK overlap	10.5 ± 1.7	11.3 ± 1.3

A.2 Distributions and parameterisations of conditional observables

The distributions of the decay-time uncertainty estimates σ_t of the different components are studied by extracting sWeighted samples using a fit to the reconstructed mass with the parameterisation as discussed in Section 5.1.2. A sum of two log-normal distributions is found to describe the σ_t distributions, with the exception of the background sample in the LL category, where a single log-normal distribution is sufficient. Both the $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ signal component use the same parameterisation. The log-normal distribution with the median m_{σ_t} and the shape parameter k_{σ_t} , is defined as

$$P_{\text{Ln}}(\sigma_t | m_{\sigma_t}, k_{\sigma_t}) = \frac{1}{\sqrt{2\pi}\sigma_t \ln k_{\sigma_t}} \exp\left(-\frac{\ln^2(\sigma_t/m_{\sigma_t})}{2 \ln^2 k_{\sigma_t}}\right). \quad (\text{A.1})$$

In the sum of two log-normal distributions, f_{σ_t} denotes the fraction between both distributions. Resulting parameters for the fits to the sWeighted distributions are given in Tables A.5 to A.7. Figures A.1 and A.2 show the extracted sWeighted distributions for the $B^0 \rightarrow J/\psi K_S^0$ and the combinatorial background components along with fitted PDFs. This parameterisation is used for the generation of the σ_t observable in pseudo-experiments.

The distributions of the mistag probability estimates η^{OS} and η^{SSK} are shown in Figures A.3 and A.4. No physically motivated parameterisation is found for these distributions. Therefore, both η^{OS} and η^{SSK} are parameterised using histogram PDFs which are determined using sWeighted samples. These PDFs are shown in Figures A.3 and A.4 as well and are used solely for the generation of toy MC samples. Similar to the decay-time uncertainties, the $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ components use the same histograms.

Table A.5 – Values for the parameters describing the distribution of the per-candidate decay-time uncertainty estimates σ_t of the $B^0 \rightarrow J/\psi K_S^0$ candidates in the LL sample for the different tagging categories.

Parameter	U	OS	SSK
$m_{\sigma_t,1}$	$0.030\,76 \pm 0.000\,10$	$0.030\,48 \pm 0.000\,35$	0.0285 ± 0.0004
$m_{\sigma_t,2}$	0.0326 ± 0.0004	$0.029\,95 \pm 0.000\,22$	0.0273 ± 0.0005
$k_{\sigma_t,1}$	1.276 ± 0.007	1.228 ± 0.022	1.228 ± 0.029
$k_{\sigma_t,2}$	1.487 ± 0.022	1.390 ± 0.018	1.366 ± 0.027
f_{σ_t}	0.77 ± 0.05	0.39 ± 0.12	0.45 ± 0.19

Table A.6 – Values for the parameters describing the distribution of the per-candidate decay-time uncertainty estimates σ_t of the $B^0 \rightarrow J/\psi K_S^0$ candidates in the DD sample for the different tagging categories.

Parameter	U	OS	SSK
$m_{\sigma_t,1}$	$0.031\,62 \pm 0.000\,23$	$0.030\,71 \pm 0.000\,26$	0.0289 ± 0.0004
$m_{\sigma_t,2}$	0.0419 ± 0.0029	0.0388 ± 0.0021	0.039 ± 0.012
$k_{\sigma_t,2}$	1.358 ± 0.012	1.350 ± 0.016	1.344 ± 0.027
$k_{\sigma_t,2}$	0.659 ± 0.007	0.644 ± 0.007	0.663 ± 0.033
f_{σ_t}	0.73 ± 0.08	0.70 ± 0.08	0.83 ± 0.20

Table A.7 – Values for the parameters describing the distribution of the per-candidate decay-time uncertainty estimates σ_t of the combinatorial background candidates.

Parameter	LL	DD
$m_{\sigma_t,1}$	0.0292 ± 0.0005	0.0330 ± 0.0006
$m_{\sigma_t,2}$	—	0.051 ± 0.008
$k_{\sigma_t,1}$	0.720 ± 0.009	0.720 ± 0.011
$k_{\sigma_t,2}$	—	0.62 ± 0.04
f_{σ_t}	—	0.79 ± 0.09

A.2 Distributions and parameterisations of conditional observables

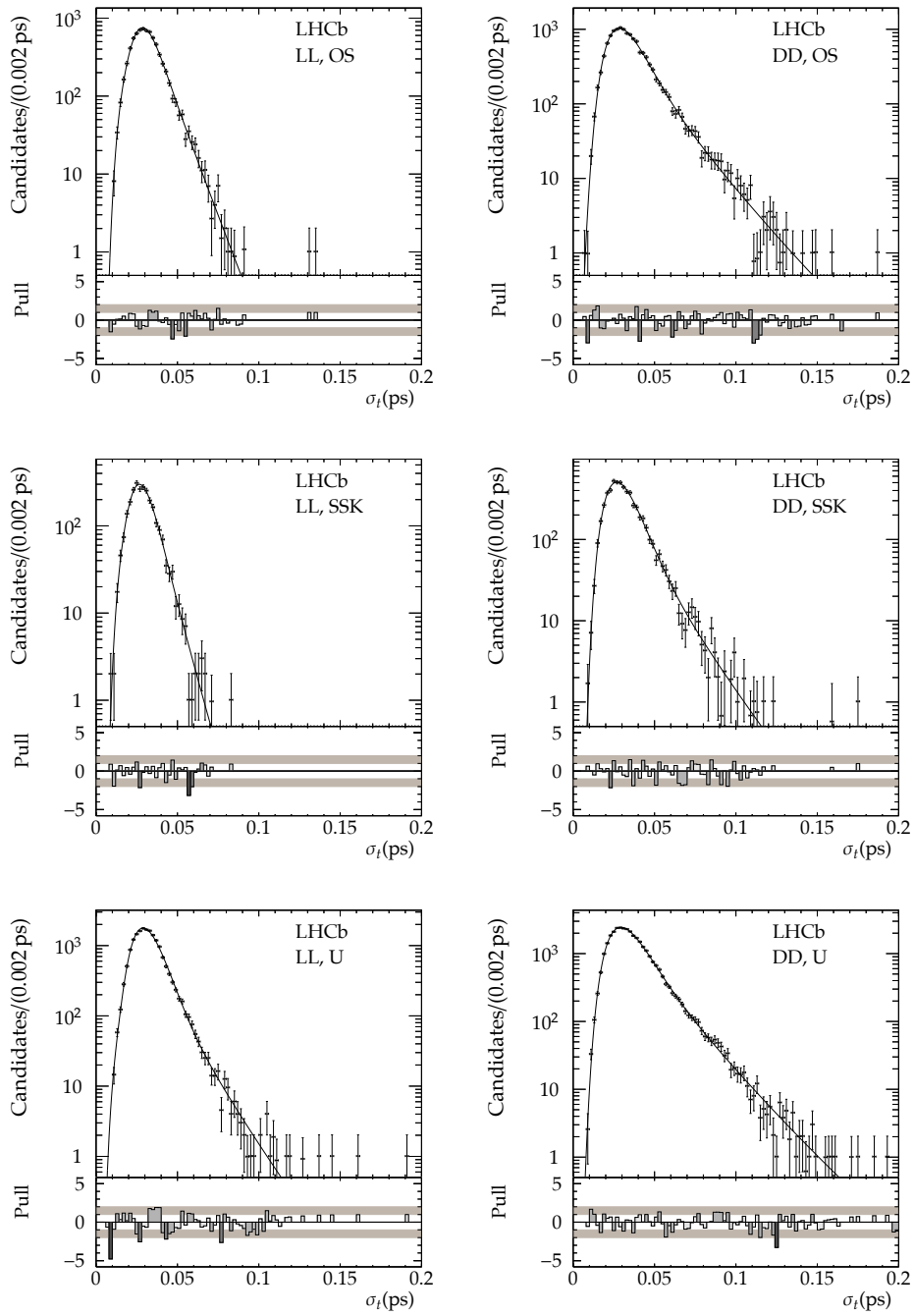


Figure A.1 – Distribution of decay-time uncertainty estimates σ_t for $B^0 \rightarrow J/\psi K_S^0$ candidates in the LL (left) and DD (right) sample, separated for the exclusively OS tagged (top), all SSK tagged (middle), and untagged (bottom) categories. Fitted parameterisations (black, solid) are based on log-normal distributions.

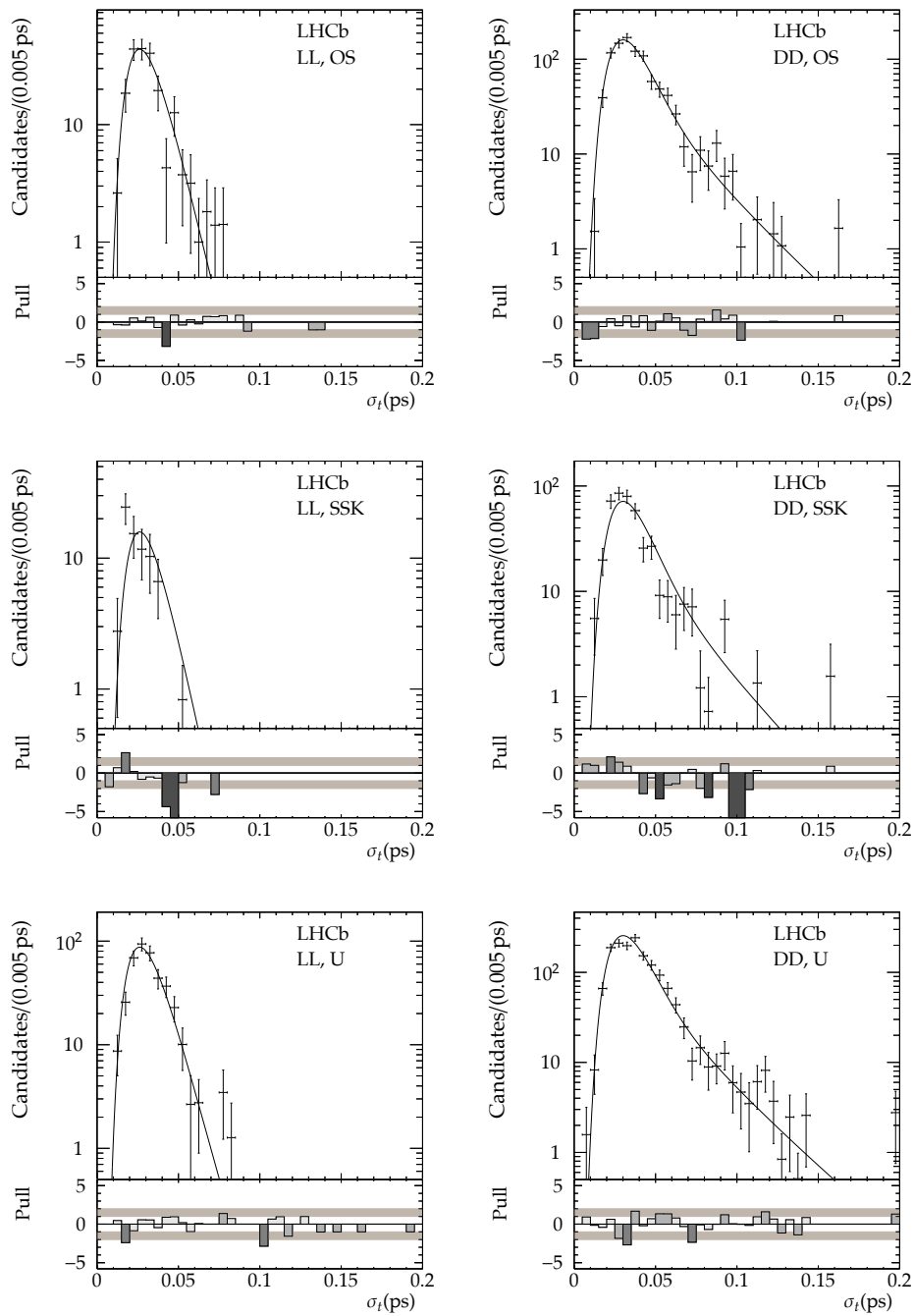


Figure A.2 – Distribution of decay-time uncertainty estimates σ_t for background candidates in the LL (left) and DD (right) sample, separated for the exclusively OS tagged (top), all SSK tagged (middle), and untagged (bottom) categories. Fitted parameterisations (black, solid) are based on log-normal distributions.

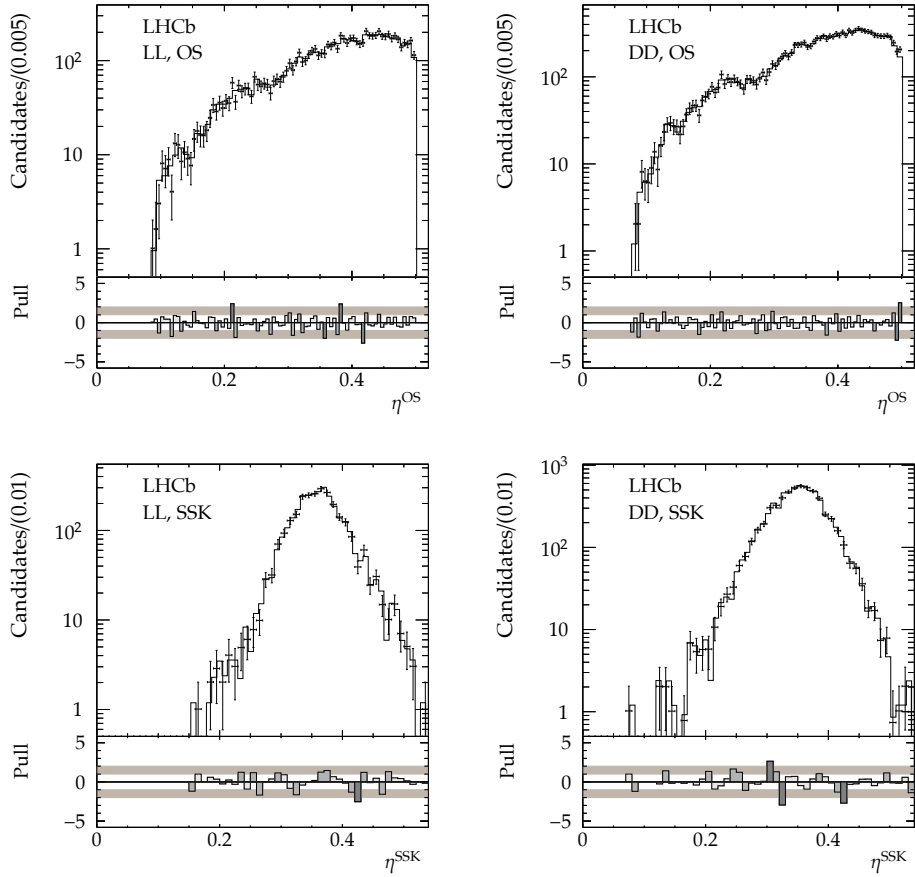


Figure A.3 – Distribution of mistag probability estimates η^{OS} (top) and η^{SSK} (bottom) for $B^0 \rightarrow J/\psi K_S^0$ candidates in the LL (left) and DD (right) sample. Histogram PDFs (black, solid) are shown as well.

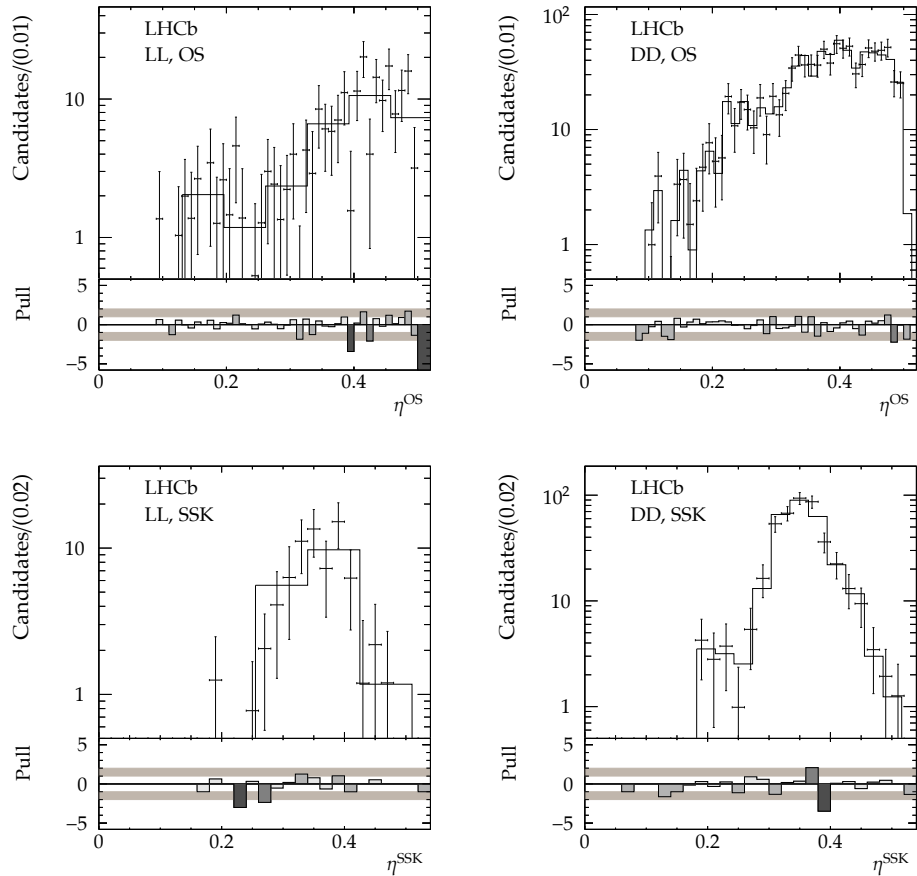


Figure A.4 – Distribution of mistag probability estimates η^{OS} (top) and η^{SSK} (bottom) for background candidates in the LL (left) and DD (right) sample. Histogram PDFs (black, solid) are shown as well.

A.3 Detailed fit results

Besides the CP parameters, the nominal maximum-likelihood fit also determines the nuisance parameters. Table A.8 lists results for the event yields for the three components, which are an ingredient for the branching ratio measurement. Tables A.9 and A.10 list the results for all remaining parameters. These also include the constrained parameters, which are in excellent agreement with the externally provided world averages. Finally, Figure A.5 shows the full correlation matrix for all varied parameters. Pronounced correlation is evident for all parameters determining the acceptance, for the parameters describing the background decay-time distribution, and the signal mass shape parameters. Furthermore, there is a smaller correlation between the $B_s^0 \rightarrow J/\psi K_S^0$ component and the combinatorial background, visible in yields and shape parameters.

Table A.8 – Fit results of the nominal fit for the event yields for the $B_s^0 \rightarrow J/\psi K_S^0$ component, the $B^0 \rightarrow J/\psi K_S^0$ component, and the combinatorial background. As the fit is performed simultaneously on the exclusively OS tagged (OS), all SSK tagged (SSK), and untagged (U) candidates, separate yields are determined for these categories. The summed yields for a given component are specified as well.

Parameter	LL	DD
$N_{B_s^0}^{\text{OS}}$	80 ± 10	171 ± 16
$N_{B_s^0}^{\text{SSK}}$	33 ± 6	70 ± 11
$N_{B_s^0}^{\text{U}}$	195 ± 16	359 ± 23
$N_{B_s^0}$	307 ± 20	600 ± 30
$N_{B^0}^{\text{OS}}$	7400 ± 90	$13\,220 \pm 120$
$N_{B^0}^{\text{SSK}}$	2610 ± 50	5680 ± 80
$N_{B^0}^{\text{U}}$	$17\,800 \pm 140$	$32\,450 \pm 180$
N_{B^0}	$27\,800 \pm 170$	$51\,350 \pm 230$
$N_{\text{Bg}}^{\text{OS}}$	192 ± 19	930 ± 40
$N_{\text{Bg}}^{\text{SSK}}$	69 ± 11	436 ± 26
N_{Bg}^{U}	397 ± 28	1490 ± 50
N_{Bg}	660 ± 40	2850 ± 70

Table A.9 – Fit results of the nominal fit for nuisance parameters in the LL and DD subsamples. Listed are mass shape parameters, the acceptance coefficients $c_{\text{acc},i}$, the background decay-time parameters, and results for the constrained effective production asymmetries $A_{P,j}$. The mass shape parameters include the mass peak μ , the shape parameter λ , the core distribution width σ , and the background slope c_{Bg}^k . The background decay-time parameters are the fraction f_{Bg}^k between the two exponential distributions, the pseudo-lifetime τ_{Bg}^k and the relative scale $s_{t,\text{Bg}}^k$.

Parameter	LL		DD		Unit
μ	5280.67	± 0.05	5281.43	± 0.04	MeV/c^2
λ	-3.17	± 0.19	-2.83	± 0.12	
σ	8.07	± 0.08	9.40	± 0.08	MeV/c^2
c_{Bg}^k	-0.0033	± 0.0005	-0.003 13	$\pm 0.000 21$	$(\text{MeV}/c^2)^{-1}$
$c_{\text{acc},1}$	0.21	± 0.04	0.164	± 0.023	
$c_{\text{acc},2}$	0.34	± 0.07	0.27	± 0.04	
$c_{\text{acc},3}$	0.61	± 0.11	0.34	± 0.04	
$c_{\text{acc},4}$	0.89	± 0.17	0.66	± 0.09	
$c_{\text{acc},5}$	1.04	± 0.18	1.07	± 0.13	
$c_{\text{acc},6}$	0.74	± 0.22	0.87	± 0.16	
f_{Bg}^k	0.27	± 0.06	0.103	± 0.019	
τ_{Bg}^k	0.20	± 0.06	0.174	± 0.035	ps
$s_{t,\text{Bg}}^k$	9.2	± 2.3	10.1	± 2.0	
A_{P,B_s^0}	-0.041	± 0.032	-0.021	± 0.025	
A_{P,B^0}	-0.011	± 0.006	-0.010	± 0.005	

Table A.10 – Fit results of the nominal fit for nuisance parameters being shared between LL and DD. Listed are the external physical parameters being constrained in the fit and the shift between the $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ mass peaks, s_m .

Parameter	Result		Unit
Δm_s	17.757	± 0.021	ps^{-1}
Γ_s	0.6628	± 0.0019	ps^{-1}
$\Delta \Gamma_s$	0.081	± 0.006	ps^{-1}
Δm_d	0.5102	± 0.0030	ps^{-1}
τ_{B^0}	1.520	± 0.004	ps
s_m	87.13	± 0.34	MeV/c^2

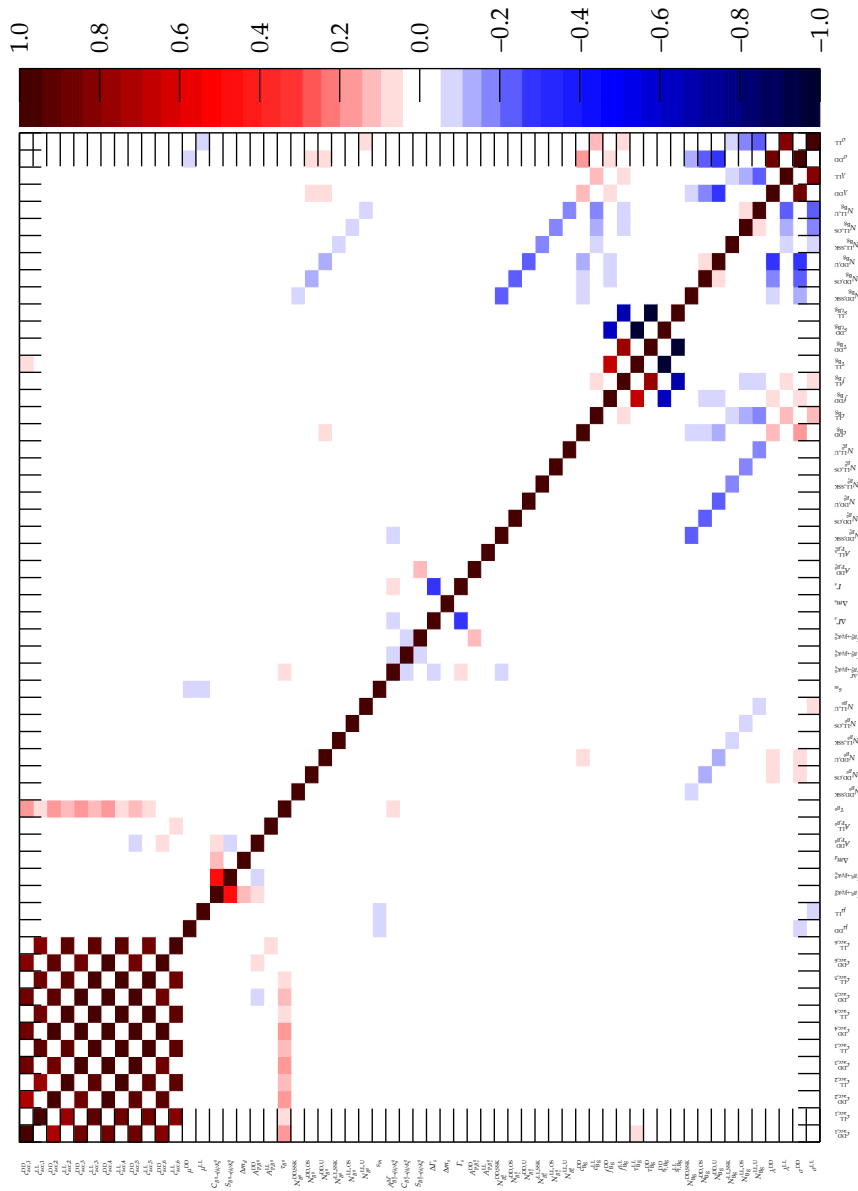


Figure A.5 – Detailed correlation matrix of the fit parameters in the nominal fit, in which all nuisance parameters are varied. The size of the correlation coefficients is represented by red and blue colours, where a correlation coefficient below 5 % is displayed as white.

A.4 Detailed results of fits on simulated samples

As for the data sample, the fits on simulated samples require preparatory studies for the decay-time resolution, decay-time acceptance, and the flavour tagging. Furthermore, the fits determine nuisance parameters.

A.4.1 Results for the decay-time resolution and acceptance

Table A.11 lists the determined decay-time resolution parameters, while Figure A.6 shows the according distributions as well as the fitted PDF. The $t - t'$ distribution on the $B^0 \rightarrow J/\psi K_S^0$ sample is displayed with the identical PDF as well. An excellent agreement between the resolution of both decay channels can be observed. Results for the coefficients describing the decay-time acceptance are listed in Table A.12.

Table A.11 – Results for parameters describing the decay-time resolution on simulated samples. See Section 4.4.1 for details of the parameterisation.

Parameter	LL		DD		Unit
s_r	1.178	± 0.007	1.176	± 0.004	
$r_{r,2}$	1.76	± 0.04	1.854	± 0.031	
$r_{r,3}$	16.0	± 0.7	15.8	± 0.7	
$f_{r,1}$	0.884	± 0.008	0.920	± 0.004	
$f_{r,3}$	0.106	± 0.008	0.075	± 0.004	
μ_r	-0.00049 ± 0.00017		-0.00079 ± 0.00011		ps

Table A.12 – Results for the spline coefficients describing the decay-time acceptance on simulated samples as obtained with simulated $B^0 \rightarrow J/\psi K_S^0$ decays.

Parameter	LL	DD
$c_{\text{acc},1}$	0.37 ± 0.06	0.259 ± 0.024
$c_{\text{acc},2}$	0.48 ± 0.08	0.316 ± 0.030
$c_{\text{acc},3}$	0.75 ± 0.12	0.42 ± 0.04
$c_{\text{acc},4}$	0.95 ± 0.16	0.67 ± 0.06
$c_{\text{acc},5}$	1.08 ± 0.17	1.01 ± 0.08
$c_{\text{acc},6}$	0.88 ± 0.22	0.81 ± 0.11

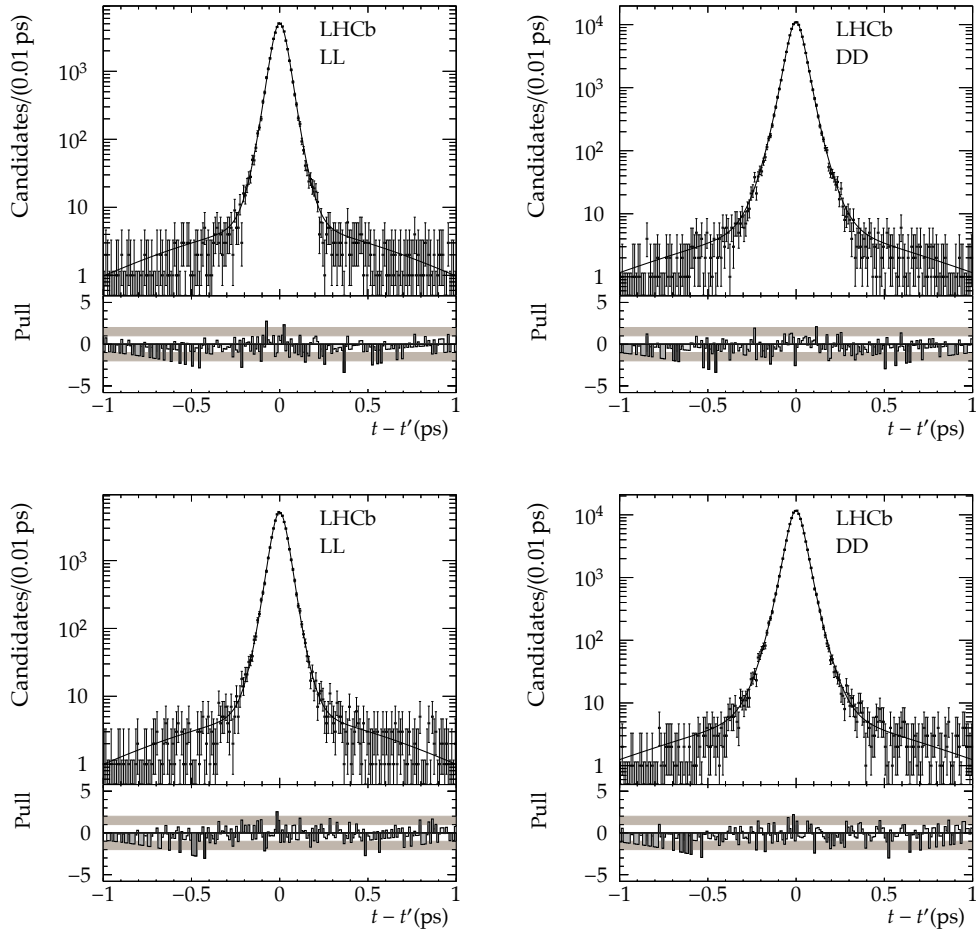


Figure A.6 – Distribution of decay-time residuals $t - t'$ in the simulated $B_s^0 \rightarrow J/\psi K_S^0$ (top) and $B^0 \rightarrow J/\psi K_S^0$ (bottom) samples for the LL (left) and DD (right) category. The fitted resolution function $R(t - t' | \sigma_t)$ (black, solid) is determined on the $B_s^0 \rightarrow J/\psi K_S^0$ sample.

A.4.2 Results for the flavour tagging calibration

The calibration parameters for the OS combination are

$$\begin{aligned} p_0^{\text{OS}} &= 0.3527 \pm 0.0008 \text{ (stat)}, \\ p_1^{\text{OS}} &= 0.962 \pm 0.009 \text{ (stat)}, \\ \langle \eta^{\text{OS}} \rangle &= 0.34941. \end{aligned} \quad (\text{A.2})$$

For the SSK tagger on $B_s^0 \rightarrow J/\psi K_S^0$ decays, the calibration parameters are determined on the simulated $B_s^0 \rightarrow J/\psi K_S^0$ sample as

$$\begin{aligned} p_0^{\text{SSK}, B_s^0} &= 0.300 \pm 0.004 \text{ (stat)}, \\ p_1^{\text{SSK}, B_s^0} &= 1.80 \pm 0.07 \text{ (stat)}, \\ \langle \eta^{\text{SSK}, B_s^0} \rangle &= 0.338. \end{aligned} \quad (\text{A.3})$$

In case of the $B^0 \rightarrow J/\psi K_S^0$ decays, the calibration parameters are determined on the simulated $B^0 \rightarrow J/\psi K^{*0}$ sample as

$$\begin{aligned} p_0^{\text{SSK}, B^0} &= 0.471 \pm 0.002 \text{ (stat)}, \\ p_1^{\text{SSK}, B^0} &= 0.182 \pm 0.030 \text{ (stat)}, \\ \langle \eta^{\text{SSK}, B^0} \rangle &= 0.358, \end{aligned} \quad (\text{A.4})$$

where a flipped tagging response is assumed (see Section 4.2.6).

Figure A.7 shows both calibration fits for the SSK tagger for the $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K^{*0}$ samples.

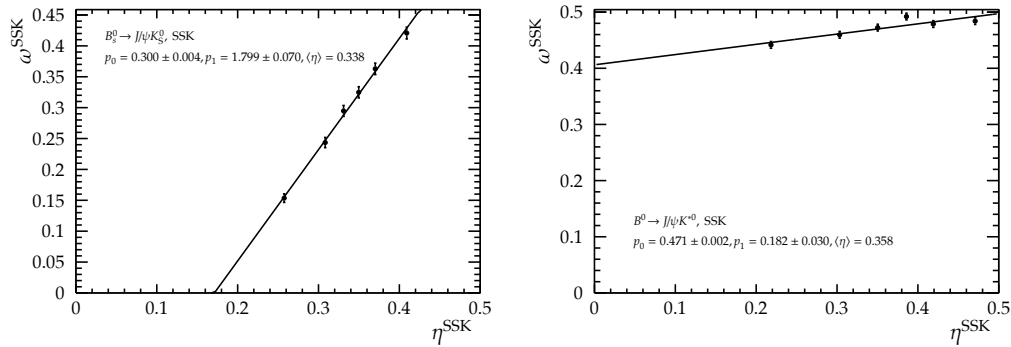


Figure A.7 – Calibration of the cut-based SSK tagger using simulated signal samples of $B_s^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow J/\psi K^{*0}$ decays, where in the latter case a flipped tag decision is applied.

A.4.3 Results for the nuisance parameters on simulated samples

Resulting event yields for the fit on the simulated $B_s^0 \rightarrow J/\psi K_S^0$ signal sample are listed in Table A.13. Results for the event yields and other nuisance parameters for the fit on the simulated compound sample are given in Tables A.14 and A.15.

Table A.13 – Fit results for the event yields for the fit to the simulated $B_s^0 \rightarrow J/\psi K_S^0$ signal sample. As the fit is performed simultaneously on the exclusively OS tagged (OS), all SSK tagged (SSK), and untagged (U) candidates, separate yields are determined for these categories. The summed yields for a given category are specified as well.

Parameter	LL	DD
$N_{B_s^0}^{\text{OS}}$	$12\,960 \pm 110$	$29\,860 \pm 170$
$N_{B_s^0}^{\text{SSK}}$	4080 ± 60	$11\,870 \pm 110$
$N_{B_s^0}^{\text{U}}$	$30\,790 \pm 180$	$73\,090 \pm 270$
$N_{B_s^0}$	$47\,820 \pm 220$	$114\,820 \pm 340$

Table A.14 – Fit results of the fit to the simulated compound sample for the event yields of the $B_s^0 \rightarrow J/\psi K_S^0$ component, the $B^0 \rightarrow J/\psi K_S^0$ component, and the combinatorial background. As the fit is performed simultaneously on the exclusively OS tagged (OS), all SSK tagged (SSK), and untagged (U) candidates, separate yields are determined for these categories. The summed yields for a given component are specified as well.

Parameter	LL	DD
$N_{B_s^0}^{\text{OS}}$	126 ± 13	353 ± 23
$N_{B_s^0}^{\text{SSK}}$	50 ± 8	132 ± 14
$N_{B_s^0}^{\text{U}}$	338 ± 20	892 ± 34
$N_{B_s^0}$	515 ± 25	1380 ± 40
$N_{B^0}^{\text{OS}}$	$13\,090 \pm 120$	$29\,500 \pm 170$
$N_{B^0}^{\text{SSK}}$	2760 ± 50	7480 ± 90
$N_{B^0}^{\text{U}}$	$31\,100 \pm 180$	$71\,120 \pm 270$
N_{B^0}	$46\,960 \pm 220$	$108\,100 \pm 330$
$N_{\text{Bg}}^{\text{OS}}$	341 ± 24	1530 ± 50
$N_{\text{Bg}}^{\text{SSK}}$	106 ± 13	774 ± 34
N_{Bg}^{U}	636 ± 34	2410 ± 70
N_{Bg}	1080 ± 40	4710 ± 90

Table A.15 – Fit results of the fit to the simulated compound sample for nuisance parameters in the LL and DD subsamples. Listed are mass shape parameters, the acceptance coefficients $c_{\text{acc},i}$, and the background decay-time parameters. The mass shape parameters include the mass peak μ , the shape parameter λ , the core distribution width σ , the background slope c_{Bg}^k , and the shift between the $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi K_S^0$ mass peaks s_m which is shared between the LL and DD categories. The background decay-time parameters are the fraction f_{Bg}^k between the two exponential distributions, the pseudo-lifetime τ_{Bg}^k , and the relative scale $s_{t,\text{Bg}}^k$.

Parameter	LL		DD		Unit
μ	5279.879	± 0.030	5279.888	± 0.023	MeV/ c^2
λ	-3.63	± 0.18	-3.55	± 0.11	
σ	6.79	± 0.05	7.883	± 0.035	MeV/ c^2
c_{Bg}^k	-0.00370 ± 0.00035		-0.00327 ± 0.00016		(MeV/ c^2) $^{-1}$
s_m	87.14 ± 0.20				MeV/ c^2
$c_{\text{acc},1}$	0.37	± 0.06	0.265	± 0.026	
$c_{\text{acc},2}$	0.47	± 0.08	0.328	± 0.031	
$c_{\text{acc},3}$	0.76	± 0.12	0.42	± 0.04	
$c_{\text{acc},4}$	0.94	± 0.16	0.68	± 0.06	
$c_{\text{acc},5}$	1.08	± 0.17	1.03	± 0.09	
$c_{\text{acc},6}$	0.86	± 0.21	0.83	± 0.11	
f_{Bg}^k	0.206	± 0.027	0.089	± 0.013	
τ_{Bg}^k	0.136	± 0.024	0.153	± 0.027	ps
$s_{t,\text{Bg}}^k$	15.2	± 2.5	12.4	± 2.1	

Bibliography

- [1] T. Lee and C.-N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1956) 254–258, DOI: [10.1103/PhysRev.104.254](https://doi.org/10.1103/PhysRev.104.254).
- [2] C. Wu *et al.*, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413–1414, DOI: [10.1103/PhysRev.105.1413](https://doi.org/10.1103/PhysRev.105.1413).
- [3] J. Christenson *et al.*, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138–140, DOI: [10.1103/PhysRevLett.13.138](https://doi.org/10.1103/PhysRevLett.13.138).
- [4] B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, DOI: [10.1103/PhysRevLett.87.091801](https://doi.org/10.1103/PhysRevLett.87.091801), [arXiv:hep-ex/0107013](https://arxiv.org/abs/hep-ex/0107013) [hep-ex].
- [5] K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, DOI: [10.1103/PhysRevLett.87.091802](https://doi.org/10.1103/PhysRevLett.87.091802), [arXiv:hep-ex/0107061](https://arxiv.org/abs/hep-ex/0107061) [hep-ex].
- [6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509, DOI: [10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508); F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323, DOI: [10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321).
- [7] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588, DOI: [10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2); S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266, DOI: [10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264); A. Salam, *Weak and Electromagnetic Interactions*, 8th Nobel Symposium, *Elementary Particle Theory*, ed. by N. Svartholm, **C68-05-19**, Lerum, Sweden, 1968.
- [8] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533, DOI: [10.1103/PhysRevLett.10.531](https://doi.org/10.1103/PhysRevLett.10.531).
- [9] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657, DOI: [10.1143/PTP.49.652](https://doi.org/10.1143/PTP.49.652).
- [10] A. Sakharov, *Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32–35, DOI: [10.1070/PU1991v034n05ABEH002497](https://doi.org/10.1070/PU1991v034n05ABEH002497).

- [11] R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, DOI: 10.1142/S0217751X15300227, arXiv:1412.6352 [hep-ex].
- [12] R. Aaij *et al.*, *Measurement of CP violation in $B^0 \rightarrow J/\psi K_S^0$ decays*, Phys. Rev. Lett. **115** (2015) 031601, DOI: 10.1103/PhysRevLett.115.031601, arXiv:1503.07089 [hep-ex].
- [13] R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetries in $B_s^0 \rightarrow J/\psi K_S^0$* , JHEP **2015.6**, 131 (2015), DOI: 10.1007/JHEP06(2015)131, arXiv:1503.07055 [hep-ex].
- [14] T. Mombächer, *Untersuchung des Same Side Kaon Taggings auf den Kanälen $B^0 \rightarrow J/\psi K^{*0}$ und $B^0 \rightarrow J/\psi K_S^0$ am LHCb-Experiment*, B.Sc. thesis, TU Dortmund University, 2014.
- [15] V. Müller, *Untersuchung multivariater Methoden zur Selektion von B-Meson-Zerfällen am LHCb-Experiment*, B.Sc. thesis, TU Dortmund University, 2012.
- [16] S. Roese, *Implementierung der Feldman-Cousins-Methode zur Bestimmung von Konfidenzintervallen für CP-Analysen am LHCb-Experiment*, B.Sc. thesis, TU Dortmund University, 2012.
- [17] T. Schmelzer, *Untersuchungen und Vergleich verschiedener Algorithmen zur Bestimmung von Zerfallszeiten von B-Mesonen am LHCb-Experiment*, B.Sc. thesis, TU Dortmund University, 2012.
- [18] T. Tekampe, *Goodness-of-Fit-Tests für die $\sin(2\beta)$ -Analyse in $B^0 \rightarrow J/\psi K_S^0$ am LHCb-Experiment*, B.Sc. thesis, TU Dortmund University, 2012.
- [19] K. Olive *et al.*, *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001, DOI: 10.1088/1674-1137/38/9/090001.
- [20] D. Boutigny *et al.*, *The Babar physics book: Physics at an asymmetric B factory* (1998).
- [21] G. C. Branco, L. Lavoura, and J. P. Silva, *CP Violation*, Int. Ser. Monogr. Phys. **103** (1999).
- [22] I. I. Bigi and A. Sanda, *CP violation*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **9** (2000).
- [23] J. T. Wishahi, *Measurement of CP Violation in $B^0 \rightarrow J/\psi K_S^0$ Decays with the LHCb Experiment*, Ph.D. thesis, TU Dortmund University, 2013, DOI: 10.17877/DE290R-13417.
- [24] R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi K^- p$ decays*, Phys. Rev. Lett. **115** (2015) 072001, DOI: 10.1103/PhysRevLett.115.072001, arXiv:1507.03414 [hep-ex].

- [25] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802, doi: [10.1103/PhysRevLett.53.1802](https://doi.org/10.1103/PhysRevLett.53.1802).
- [26] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945, doi: [10.1103/PhysRevLett.51.1945](https://doi.org/10.1103/PhysRevLett.51.1945).
- [27] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, *Waiting for the top quark mass, $K^+ \rightarrow \pi^+ \nu \bar{\nu}$, $B_s^0 - \bar{B}_s^0$ mixing and CP asymmetries in B decays*, Phys. Rev. **D50** (1994) 3433–3446, doi: [10.1103/PhysRevD.50.3433](https://doi.org/10.1103/PhysRevD.50.3433), arXiv: [hep-ph/9403384](https://arxiv.org/abs/hep-ph/9403384) [hep-ph].
- [28] J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1–131, doi: [10.1140/epjc/s2005-02169-1](https://doi.org/10.1140/epjc/s2005-02169-1), arXiv: [hep-ph/0406184](https://arxiv.org/abs/hep-ph/0406184) [hep-ph], updated results and plots available at: <http://ckmfitter.in2p3.fr>.
- [29] J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, Phys. Rev. **D91.7** (2015) 073007, doi: [10.1103/PhysRevD.91.073007](https://doi.org/10.1103/PhysRevD.91.073007), arXiv: [1501.05013](https://arxiv.org/abs/1501.05013) [hep-ph].
- [30] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Violation*, Phys. Rev. Lett. **55** (1985) 1039, doi: [10.1103/PhysRevLett.55.1039](https://doi.org/10.1103/PhysRevLett.55.1039).
- [31] R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in $B^0 - \bar{B}^0$ mixing*, Phys. Rev. Lett. **114** (2015) 041601, doi: [10.1103/PhysRevLett.114.041601](https://doi.org/10.1103/PhysRevLett.114.041601), arXiv: [1409.8586](https://arxiv.org/abs/1409.8586) [hep-ex].
- [32] R. Aaij *et al.*, *Measurement of the flavour-specific CP-violating asymmetry a_{sl}^s in B_s^0 decays*, Phys. Lett. **B728** (2014) 607–615, doi: [10.1016/j.physletb.2013.12.030](https://doi.org/10.1016/j.physletb.2013.12.030), arXiv: [1308.1048](https://arxiv.org/abs/1308.1048) [hep-ex].
- [33] R. Fleischer, *Extracting γ from $B_{s(d)} \rightarrow J/\psi K_S^0$ and $B_{d(s)} \rightarrow D_{d(s)}^+ D_{d(s)}^-$* , Eur. Phys. J. **C10** (1999) 299–306, doi: [10.1007/s100529900099](https://doi.org/10.1007/s100529900099), arXiv: [hep-ph/9903455](https://arxiv.org/abs/hep-ph/9903455) [hep-ph].
- [34] S. Faller *et al.*, *The Golden Modes $B^0 \rightarrow J/\psi K_{S,L}^0$ in the Era of Precision Flavour Physics*, Phys. Rev. **D79** (2009) 014030, doi: [10.1103/PhysRevD.79.014030](https://doi.org/10.1103/PhysRevD.79.014030), arXiv: [0809.0842](https://arxiv.org/abs/0809.0842) [hep-ph].
- [35] K. De Bruyn, R. Fleischer, and P. Koppenburg, *Extracting γ and Penguin Topologies through CP Violation in $B_s^0 \rightarrow J/\psi K_S^0$* , Eur. Phys. J. **C70** (2010) 1025–1035, doi: [10.1140/epjc/s10052-010-1495-z](https://doi.org/10.1140/epjc/s10052-010-1495-z), arXiv: [1010.0089](https://arxiv.org/abs/1010.0089) [hep-ph].
- [36] K. De Bruyn and R. Fleischer, *A Roadmap to Control Penguin Effects in $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi \phi$* , JHEP **1503** (2015) 145, doi: [10.1007/JHEP03\(2015\)145](https://doi.org/10.1007/JHEP03(2015)145), arXiv: [1412.6834](https://arxiv.org/abs/1412.6834) [hep-ph].

- [37] M. Jung, *Determining weak phases from $\rightarrow J/\psi P$ decays*, Phys. Rev. **D86** (2012) 053008, DOI: 10.1103/PhysRevD.86.053008, arXiv:1206.2050 [hep-ph].
- [38] P. Frings, U. Nierste, and M. Wiebusch, *Penguin contributions to CP phases in $B_{d,s}$ decays to charmonium* (2015), arXiv:1503.00859 [hep-ph].
- [39] Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ -lepton properties as of summer 2014* (2014), arXiv:1412.7515 [hep-ex], updated results and plots available at: <http://www.slac.stanford.edu/xorg/hfag/>.
- [40] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008), ed. by L. Evans S08001, DOI: 10.1088/1748-0221/3/08/S08001.
- [41] C. Lefèvre, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, 2008, <https://cds.cern.ch/record/1260465/files/0812015.pdf> (visited on 2015-07-21).
- [42] G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003, DOI: 10.1088/1748-0221/3/08/S08003.
- [43] S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004, DOI: 10.1088/1748-0221/3/08/S08004.
- [44] G. Aad *et al.*, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803, DOI: 10.1103/PhysRevLett.114.191803, arXiv:1503.07589 [hep-ex].
- [45] K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002, DOI: 10.1088/1748-0221/3/08/S08002.
- [46] G. Anelli *et al.*, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007, DOI: 10.1088/1748-0221/3/08/S08007.
- [47] O. Adriani *et al.*, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006, DOI: 10.1088/1748-0221/3/08/S08006.
- [48] B. Acharya *et al.*, *The Physics Programme Of The MoEDAL Experiment At The LHC*, Int. J. Mod. Phys. **A29** (2014) 1430050, DOI: 10.1142/S0217751X14300506, arXiv:1405.7662 [hep-ph].
- [49] J. M. Campbell, J. Huston, and W. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, DOI: 10.1088/0034-4885/70/1/R02, arXiv:hep-ph/0611148 [hep-ph].
- [50] *$b\bar{b}$ production angle plots*, http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html (visited on 2015-07-21).
- [51] J. Alves A. Augusto *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005, DOI: 10.1088/1748-0221/3/08/S08005.

- [52] C. Ilgner *et al.*, *The Beam Conditions Monitor of the LHCb Experiment* (2010), [arXiv:1001.2487](https://arxiv.org/abs/1001.2487) [physics.ins-det].
- [53] F. Alessio, Z. Guzik, and R. Jacobsson, *An LHCb general-purpose acquisition board for beam and background monitoring at the LHC*, JINST **6** (2011) C01001, doi: 10.1088/1748-0221/6/01/C01001.
- [54] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, doi: 10.1088/1748-0221/9/09/P09007, [arXiv:1405.7808](https://arxiv.org/abs/1405.7808) [physics.ins-det].
- [55] M. Clemencic *et al.*, *Recent developments in the LHCb software framework Gaudi*, *17th International Conference on Computing in High Energy and Nuclear Physics*, ed. by J. Gruntorad, **219**, 2010 042006, doi: 10.1088/1742-6596/219/4/042006; G. Barrand *et al.*, *GAUDI - The software architecture and framework for building LHCb data processing applications*, *Comput. Phys. Commun.* **140**.1-2 (2001), ed. by M. Mazzucato 92–95.
- [56] R. Aaij *et al.*, *The Moore Project*, ed. by G. Raven, 2015, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/>.
- [57] *LHCb trigger schemes*, <http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html> (visited on 2015-07-24).
- [58] M. Frank *et al.*, *Deferred High Level Trigger in LHCb: A Boost to CPU Resource Utilization*, *J. Phys. Conf. Ser.* **513** (2014) 012006, doi: 10.1088/1742-6596/513/1/012006.
- [59] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, doi: 10.1088/1748-0221/8/04/P04022, [arXiv:1211.3055](https://arxiv.org/abs/1211.3055) [hep-ex].
- [60] J. Albrecht *et al.*, *Performance of the LHCb High Level Trigger in 2012*, *J. Phys. Conf. Ser.* **513** (2014) 012001, doi: 10.1088/1742-6596/513/1/012001, [arXiv:1310.8544](https://arxiv.org/abs/1310.8544) [hep-ex].
- [61] *LHCb performance numbers*, <http://lhcb.web.cern.ch/lhcb/speakersbureau/html/PerformanceNumbers.html> (visited on 2015-07-21).
- [62] R. Aaij *et al.*, *The Brunel Project*, ed. by M. Cattaneo and R. Lambert, 2015, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/>.
- [63] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, *Eur. Phys. J.* **C73** (2013) 2431, doi: 10.1140/epjc/s10052-013-2431-9, [arXiv:1211.6759](https://arxiv.org/abs/1211.6759) [physics.ins-det].
- [64] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, doi: 10.1088/1748-0221/8/10/P10020, [arXiv:1306.0249](https://arxiv.org/abs/1306.0249) [physics.ins-det].

- [65] R. Aaij *et al.*, *The DaVinci Project*, ed. by E. Rodrigues, 2015, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/>.
- [66] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011), ed. by S. C. Lin 032023, doi: 10.1088/1742-6596/331/3/032023; I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011), ed. by S. C. Lin 032047, doi: 10.1088/1742-6596/331/3/032047.
- [67] R. Aaij *et al.*, *The Boole Project*, ed. by M. Cattaneo, 2015, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/>.
- [68] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, doi: 10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175 [hep-ph]; T. Sjöstrand *et al.*, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, doi: 10.1016/j.cpc.2015.01.024, arXiv:1410.3012 [hep-ph].
- [69] D. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155, doi: 10.1016/S0168-9002(01)00089-4.
- [70] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97–107, doi: 10.1140/epjc/s2005-02396-4, arXiv:hep-ph/0506026 [hep-ph]; N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS Interface in C++: Technical and Physics Documentation* (2010), arXiv:1011.0937 [hep-ph].
- [71] S. Agostinelli *et al.*, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303, doi: 10.1016/S0168-9002(03)01368-8; J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270, doi: 10.1109/TNS.2006.869826.
- [72] *Delivered/recorded LHCb luminosities in Run I*, <http://lhcb-operation-splots.web.cern.ch/lhcb-operationsplots/index.htm> (visited on 2015-07-21).
- [73] J. Albrecht, *The LHCb Trigger System: Present and Future*, J. Phys. Conf. Ser. **623.1** (2015) 012003, doi: 10.1088/1742-6596/623/1/012003.
- [74] R. Aaij *et al.*, *LHCb Tracker Upgrade Technical Design Report* (2014), <http://cds.cern.ch/record/1647400>.
- [75] R. Aaij *et al.*, *Framework TDR for the LHCb Upgrade: Technical Design Report* (2012), <http://cds.cern.ch/record/1443882>.
- [76] R. Aaij *et al.*, *LHCb VELO Upgrade Technical Design Report* (2013), <http://cds.cern.ch/record/1624070>.

- [77] R. Aaij *et al.*, *LHCb PID Upgrade Technical Design Report* (2013), <http://cds.cern.ch/record/1624074>.
- [78] *LHCb Trigger and Online Upgrade Technical Design Report* (2014), <http://cds.cern.ch/record/1701361>.
- [79] R. Aaij *et al.*, *Measurement of the effective $B_s^0 \rightarrow J/\psi K_S^0$ lifetime*, *Nucl.Phys.* **B873** (2013) 275–292, DOI: 10.1016/j.nuclphysb.2013.04.021, arXiv:1304.4500 [hep-ex].
- [80] R. Aaij *et al.*, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions*, LHCb conference note LHCb-CONF-2013-011, 2013.
- [81] R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, *JHEP* **1304** (2013) 001, DOI: 10.1007/JHEP04(2013)001, arXiv:1301.5286 [hep-ex].
- [82] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, *Nucl. Instrum. Meth.* **A552** (2005) 566–575, DOI: 10.1016/j.nima.2005.06.078, arXiv:physics/0503191 [physics].
- [83] P. Koppenburg, *Dealing with Multiple Candidates*, LHCb-INT-2011-009, 2011, <https://cds.cern.ch/record/1340942>, internal note.
- [84] K. De Bruyn *et al.*, *Measurement of the time-dependent CP asymmetries in $B_s^0 \rightarrow J/\psi K_S^0$* (2013), LHCb internal note, <https://cds.cern.ch/record/1507215>, Published as R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetries in $B_s^0 \rightarrow J/\psi K_S^0$* , *JHEP* **2015.6**, 131 (2015), DOI: 10.1007/JHEP06(2015)131, arXiv:1503.07055 [hep-ex].
- [85] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, *Nucl. Instrum. Meth.* **A555** (2005) 356–369, DOI: 10.1016/j.nima.2005.08.106, arXiv:physics/0402083 [physics.data-an].
- [86] R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, *Eur. Phys. J.* **C72** (2012) 2022, DOI: 10.1140/epjc/s10052-012-2022-1, arXiv:1202.4979 [hep-ex].
- [87] R. Aaij *et al.*, *Neural-network-based kaon tagging: development of new same side and opposite side tagging algorithms* (2015), in preparation.
- [88] K. Kreplin *et al.*, *Neural-network-based kaon tagging: development of new same side and opposite side tagging algorithms*, LHCb internal note, 2015, Published as R. Aaij *et al.*, *Neural-network-based kaon tagging: development of new same side and opposite side tagging algorithms* (2015), in preparation.
- [89] R. Aaij *et al.*, *B flavour tagging using charm decays at the LHCb experiment*, *JINST* **10.10** (2015) P10005, DOI: 10.1088/1748-0221/10/10/P10005, arXiv:1507.07892 [hep-ex].

- [90] R. Aaij *et al.*, *Optimization and calibration of the same-side kaon tagging algorithm using hadronic B_s^0 decays in 2011 data*, LHCb conference note CERN-LHCb-CONF-2012-033, 2012.
- [91] R. Aaij *et al.*, *Precision measurement of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$ decays*, Phys. Rev. Lett. **114.4** (2015) 041801, doi: 10.1103/PhysRevLett.114.041801, arXiv:1411.3104 [hep-ex].
- [92] R. Aaij *et al.*, *Selections and lifetime measurements for exclusive $b \rightarrow J/\psi X$ decays with $J/\psi \rightarrow \mu^+ \mu^-$ with 2010 data*, LHCb internal note, 2011, <https://cds.cern.ch/record/1308189>, Published as *b-hadron lifetime measurements with exclusive $b \rightarrow J/\psi X$ decays reconstructed in the 2010 data*, LHCb conference note CERN-LHCb-CONF-2011-001, 2011.
- [93] R. Aaij *et al.*, *Measurements of the B^+ , B^0 , B_s^0 meson and Λ_b^0 baryon lifetimes*, JHEP **04** (2014) 114, doi: 10.1007/JHEP04(2014)114, arXiv:1402.2554 [hep-ex].
- [94] M. Schellenberg, *Analyse der VELO-Rekonstruktionseffizienz im Zerfall $B^0 \rightarrow J/\psi K_S^0$ am LHCb-Experiment*, B.Sc. thesis, TU Dortmund University, 2013.
- [95] R. Aaij *et al.*, *Measurement of the \bar{B}^0 - B^0 and \bar{B}_s^0 - B_s^0 production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218–228, doi: 10.1016/j.physletb.2014.10.005, arXiv:1408.0275 [hep-ex].
- [96] A. Carbone *et al.*, *Measurement of the \bar{B}^0 - B^0 and \bar{B}_s^0 - B_s^0 production asymmetries in 7 TeV pp collisions*, LHCb internal note, 2013, <https://cds.cern.ch/record/1605075>, Published as R. Aaij *et al.*, *Measurement of the \bar{B}^0 - B^0 and \bar{B}_s^0 - B_s^0 production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218–228, doi: 10.1016/j.physletb.2014.10.005, arXiv:1408.0275 [hep-ex].
- [97] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343–367, doi: 10.1016/0010-4655(75)90039-9.
- [98] M. Oreglia, *A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$* , Ph.D. thesis, Stanford University, 1980; J. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the J/ψ and ψ'* , Ph.D. thesis, Stanford University, 1982; T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, Ph.D. thesis, Cracow, INP, 1986.
- [99] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150–155, doi: 10.1016/j.nima.2014.06.081, arXiv:1312.5000 [hep-ex].
- [100] G. Punzi, *Comments on likelihood fits with variable resolution*, eConf **C030908** (2003) WELT002, arXiv:physics/0401045 [physics.data-an].

- [101] T. M. Karbach and M. Schlupp, *Constraints on Yield Parameters in Extended Maximum Likelihood Fits* (2012), [arXiv:1210.7141 \[physics.data-an\]](#).
- [102] J. Wishahi *et al.*, *cp toy mc: A generator of pseudo-datasets of decays with time-dependent CP violation*, version 0.0, 2015, [doi: 10.5281/zenodo.20356](#).
- [103] R. Barlow, *Systematic errors: Facts and fictions* (2002) 134–144, [arXiv:hep-ex/0207026 \[hep-ex\]](#).
- [104] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, *Phys.Rev.* **D57** (1998) 3873–3889, [doi: 10.1103/PhysRevD.57.3873](#), [arXiv:physics/9711021 \[physics.data-an\]](#).
- [105] T. M. Karbach, *Feldman-Cousins Confidence Levels - Toy MC Method* (2011), [arXiv:1109.0714 \[physics.data-an\]](#).
- [106] C. Cauet *et al.*, *Measurement of CP violation in the decay $B^0 \rightarrow J/\psi K_S^0$* , LHCb internal note, 2013, <https://cds.cern.ch/record/1548396>, Published as R. Aaij *et al.*, *Measurement of CP violation in $B^0 \rightarrow J/\psi K_S^0$ decays*, *Phys. Rev. Lett.* **115** (2015) 031601, [doi: 10.1103/PhysRevLett.115.031601](#), [arXiv:1503.07089 \[hep-ex\]](#).
- [107] R. Aaij *et al.*, *Observation of the Decay $B_c^+ \rightarrow B_s^0 \pi^+$* , *Phys.Rev.Lett.* **111**.18 (2013) 181801, [doi: 10.1103/PhysRevLett.111.181801](#), [arXiv:1308.4544 \[hep-ex\]](#).
- [108] R. Aaij *et al.*, *Flavour tagged time-dependent angular analysis of $B_s^0 \rightarrow J/\psi K^+ K^-$ decays in the low $K^+ K^-$ mass range*, LHCb internal note, 2014, <https://cds.cern.ch/record/1690588>, Published as R. Aaij *et al.*, *Precision measurement of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$ decays*, *Phys. Rev. Lett.* **114**.4 (2015) 041801, [doi: 10.1103/PhysRevLett.114.041801](#), [arXiv:1411.3104 \[hep-ex\]](#).
- [109] R. Aaij *et al.*, *Precision measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency with the decay $B_s^0 \rightarrow D_s^- \pi^+$* , *New J.Phys.* **15** (2013) 053021, [doi: 10.1088/1367-2630/15/5/053021](#), [arXiv:1304.4741 \[hep-ex\]](#).
- [110] A. Mödden, *Sensitivitätsstudien für die Messung von $\sin(2\beta)$ im Zerfallskanal $B^0 \rightarrow J/\psi K_S^0$ am LHCb Experiment*, B.Sc. thesis, TU Dortmund University, 2015.
- [111] R. Aaij *et al.*, *Measurement of $\sigma(pp \rightarrow b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, *Phys. Lett.* **B694** (2010) 209–216, [doi: 10.1016/j.physletb.2010.10.010](#), [arXiv:1009.2731 \[hep-ex\]](#).
- [112] R. Aaij *et al.*, *Production of J/ψ and Υ mesons in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **06** (2013) 064, [doi: 10.1007/JHEP06\(2013\)064](#), [arXiv:1304.6977 \[hep-ex\]](#).

Acknowledgements (Danksagung)

‘If I have seen further, it is by standing on the shoulders of giants.’
— Isaac Newton standing on the shoulders
of Bernard of Chartres

Finishing a work like this thesis is not possible without the tremendous support of many people. These last pages are dedicated to thanking everyone helping me in different ways along the way.

First and foremost I want to express my gratitude to my supervisor Prof. Dr. Bernhard Spaan for allowing me to pursue my research and for creating and maintaining the open atmosphere of our working group. Despite the cluttered slides, it was you who convinced me with your particle physics introduction lecture to focus on this field of research. I also want to thank Priv.-Doz. Dr. Reiner Klingenberg for unhesitatingly reviewing my thesis, and Prof. Dr. Heinz Hövel and Dr. Bärbel Siegmann for being part of the board of examiners.

As the measurement presented in my thesis was conducted as part of the LHCb collaboration, it is built on the work and effort of countless other people. Thus, I want to thank everyone within LHCb who directly or indirectly contributed to my work. I am especially grateful to Kristof De Bruyn with whom I worked closely on the analysis. Although we needed some time to warm to each other, I enjoyed our ensuing fruitful collaboration and especially the meetings in person. Being invited to act as paronymf at your own PhD defence was a great honour. Furthermore, I want to extend my gratitude to Ramon Niet and Julian Wishahi, who also contributed to the analysis, as well as to Mirco Dorigo, Alessia Satta, Ronan McNulty, and all others helping to improve the analysis and publication.

My work would not have been half the fun were it not for the working group of Experimentelle Physik 5 at TU Dortmund. I have been able to work with people many of whom I regard as good friends. I consider this a rare privilege. Representative for many others I want to mention my office mates Christophe, Julian, Tobias, Ramon, and Max; the other B2CC fellows Frank and Vanessa; and my proofreaders Ulrich, Ramon, and Julian.

Weiterhin gebührt mein tiefer Dank meiner Familie. Meinen Eltern und Schwestern mit Partnern und Kindern für Unterstützung, zurückhaltende Geduld und den willkommenen Gegenpol zur Arbeit. Zu guter Letzt danke ich meiner Frau

Acknowledgements (Danksagung)

Sonja. Deine Liebe, der bedingungslose Rückhalt, das Zurückstellen eigener Ziele und Wünsche und das gelegentlich nötige Auffangen und Wiederaufbauen in Zeiten von Frustration und Stress haben einen erheblichen Anteil am Gelingen dieser Arbeit.