

Measurements of CP Violation in $B^0 \rightarrow [c\bar{c}]K_S^0$ Transitions at LHCb

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

vorgelegt von

Ramon Niet

geboren am 15.03.1988 in Dinslaken

Fakultät Physik

Technische Universität Dortmund

Dortmund, im Mai 2018

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

Datum des Einreichens der Arbeit: 30. Mai 2018

Datum der mündlichen Prüfung: 10. September 2018

1. Gutachter: Prof. Dr. Bernhard Spaan
2. Gutachter: Prof. Dr. Kevin Kröninger

Vorsitzender der Prüfungskommission: Prof. Dr. Roland Böhmer

Vertreter der wissenschaftlichen Mitarbeiter: Dr. Tim Ruhe

Vorsitzender des Promotionsausschusses: Prof. Dr. Thomas Weis

Abstract

This thesis presents a measurement of decay-time-dependent CP violation in the decays $B^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$, where the J/ψ is reconstructed from two electrons and the $\psi(2S)$ from two muons. The used data sample corresponds to an integrated luminosity of 3 fb^{-1} of pp collisions and has been recorded by the LHCb experiment at centre-of-mass energies of 7 and 8 TeV. The CP violation observables are measured to be

$$\begin{aligned} C(B^0 \rightarrow J/\psi K_S^0) &= 0.12 \pm 0.07 \text{ (stat.)} \pm 0.02 \text{ (syst.)}, \\ S(B^0 \rightarrow J/\psi K_S^0) &= 0.83 \pm 0.08 \text{ (stat.)} \pm 0.01 \text{ (syst.)}, \\ C(B^0 \rightarrow \psi(2S)K_S^0) &= -0.05 \pm 0.10 \text{ (stat.)} \pm 0.01 \text{ (syst.)}, \\ S(B^0 \rightarrow \psi(2S)K_S^0) &= 0.84 \pm 0.10 \text{ (stat.)} \pm 0.01 \text{ (syst.)}, \end{aligned}$$

where C describes CP violation in the direct decay, and S describes CP violation in the interference between the amplitudes for the direct decay and for the decay after B^0 - \bar{B}^0 oscillation. The results are compatible with the previous LHCb measurement that used $B^0 \rightarrow J/\psi K_S^0$ decays, where the J/ψ was reconstructed from two muons. A combination of these three measurements results in

$$\begin{aligned} C(B^0 \rightarrow [c\bar{c}]K_S^0) &= -0.017 \pm 0.029 \text{ (stat. + syst.)}, \\ S(B^0 \rightarrow [c\bar{c}]K_S^0) &= 0.760 \pm 0.034 \text{ (stat. + syst.)}, \end{aligned}$$

assuming that higher-order contributions can be neglected. The results are compatible with previous measurements and improve the precision of S at LHCb by 20 %.

Zusammenfassung

Diese Arbeit beschreibt eine Messung von zerfallszeitabhängiger CP -Verletzung in den Zerfällen $B^0 \rightarrow J/\psi K_S^0$ und $B^0 \rightarrow \psi(2S)K_S^0$, in welchen das J/ψ aus zwei Elektronen und das $\psi(2S)$ aus zwei Muonen rekonstruiert wird. Der verwendete Datensatz entspricht einer integrierten Luminosität von 3 fb^{-1} pp -Kollisionen und wurde mit dem LHCb-Experiment bei Schwerpunktsenergien von 7 und 8 TeV aufgenommen. Die gemessenen CP -Observablen sind

$$\begin{aligned} C(B^0 \rightarrow J/\psi K_S^0) &= 0.12 \pm 0.07 \text{ (stat.)} \pm 0.02 \text{ (syst.)}, \\ S(B^0 \rightarrow J/\psi K_S^0) &= 0.83 \pm 0.08 \text{ (stat.)} \pm 0.01 \text{ (syst.)}, \\ C(B^0 \rightarrow \psi(2S)K_S^0) &= -0.05 \pm 0.10 \text{ (stat.)} \pm 0.01 \text{ (syst.)}, \\ S(B^0 \rightarrow \psi(2S)K_S^0) &= 0.84 \pm 0.10 \text{ (stat.)} \pm 0.01 \text{ (syst.)}, \end{aligned}$$

wobei C die direkte CP -Verletzung im Zerfall, und S die CP -Verletzung in der Interferenz zwischen den Amplituden des direkten Zerfalls und des Zerfalls nach B^0 - \bar{B}^0 -Oszillation darstellt. Die Ergebnisse sind mit der vorherigen Messung von LHCb kompatibel, die $B^0 \rightarrow J/\psi K_S^0$ Zerfälle mit Rekonstruktion des J/ψ aus zwei Muonen verwendete. Eine Kombination dieser drei Messungen resultiert in

$$\begin{aligned} C(B^0 \rightarrow [c\bar{c}]K_S^0) &= -0.017 \pm 0.029 \text{ (stat. + syst.)}, \\ S(B^0 \rightarrow [c\bar{c}]K_S^0) &= 0.760 \pm 0.034 \text{ (stat. + syst.)}, \end{aligned}$$

unter der Annahme, dass Beiträge höherer Ordnung vernachlässigbar sind. Die Ergebnisse sind kompatibel mit vorherigen Messungen und verbessern die Präzision von S bei LHCb um 20 %.

Contents

1	Introduction	1
2	Flavour theory	5
2.1	Standard Model	5
2.2	CP violation	6
2.3	Neutral-meson system	9
2.4	Types of CP violation in neutral-meson systems	14
2.4.1	Direct CP violation	14
2.4.2	Indirect CP violation	15
2.4.3	Interference CP violation	16
2.5	$B^0 \rightarrow [c\bar{c}]K_S^0$ Decays	17
2.5.1	B^0 - \bar{B}^0 mixing phenomenology	17
2.5.2	Decay amplitudes and $\sin(2\beta)$	19
2.5.3	Status in experiment and theory	21
3	The LHCb experiment	23
3.1	The LHC and its experiments	23
3.2	The LHCb detector	25
3.2.1	Tracking system	25
3.2.2	Particle identification system	28
3.3	The LHCb trigger	29
3.3.1	Level-zero trigger	29
3.3.2	High level trigger	30
3.4	Software	31
3.4.1	Track reconstruction	31
3.4.2	Bremsstrahlung recovery	31
3.4.3	Particle identification	33
3.4.4	Simulation	34
4	Analysis tools and ingredients	37
4.1	Flavour tagging	37
4.1.1	Opposite-side flavour tagging	38
4.1.2	Same-side flavour tagging	39
4.1.3	Flavour-tagging combination	39
4.2	Maximum-likelihood formalism	40
4.3	Background-subtraction technique	41

4.4	Sample preparation	41
4.4.1	Subsamples	42
4.4.2	Decay-tree fit	42
4.4.3	Selection strategy	43
4.4.4	Figure of merit	44
5	$B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ preparations	47
5.1	Selection	47
5.1.1	Stripping selection	47
5.1.2	Loose preselection	53
5.1.3	Multivariate selection	54
5.1.4	Exclusive backgrounds – $\Lambda_b^0 \rightarrow J/\psi(e^+e^-)\Lambda(p\pi^-)$ veto	56
5.1.5	Modelling of the reconstructed B^0 mass	58
5.1.6	BDT-selection optimization	59
5.1.7	Fit windows and multiple candidates	60
5.1.8	Trigger structure of selected events	61
5.2	Results of the reconstructed-mass fit	61
5.3	Production asymmetry	62
5.4	Correction of detector effects	64
5.4.1	Decay-time resolution	64
5.4.2	Decay-time acceptance	65
5.5	Flavour tagging	67
5.5.1	Control-mode preparation	68
5.5.2	Calibration strategy	70
5.5.3	Evaluation of systematic uncertainties and results	72
5.5.4	Performance	74
6	$B^0 \rightarrow \psi(2S)K_S^0$ preparations	77
6.1	Stripping and trigger requirements	77
6.2	Multivariate selection and optimization	78
6.3	Exclusive backgrounds and multiple candidates	80
6.4	Results of the reconstructed-mass fit	81
6.5	Production asymmetry	82
6.6	Correction of detector effects	82
6.7	Flavour tagging	85
7	Measurement of $\sin(2\beta)$	89
7.1	CP -fit model	89
7.2	Observables and parameters	91
7.3	Blinding	91
7.4	Results	92
7.5	Validation and systematic uncertainties	96
7.5.1	Bias recognition and treatment	96
7.5.2	Model for validation and systematics	97

7.5.3	Validation of the fit	98
7.5.4	Systematic uncertainties	99
7.5.5	Other cross-checks	102
7.6	Correction for kaon regeneration and CP violation	104
8	Combinations	107
8.1	Likelihood function	107
8.2	Combination of $B^0 \rightarrow J/\psi K_S^0$ modes	108
8.3	Combination of $B^0 \rightarrow c\bar{c}K_S^0$ modes	108
9	Conclusion and Outlook	111
	Bibliography	115

1 Introduction

Why do we live in a matter dominated universe, although we observe that matter and anti-matter are created in equal parts? — This question represents a puzzling mystery that physicists struggle to solve, while their theory describing the fundamental interactions and constituents of nature, the Standard Model (SM) [1–3], has been overwhelmingly successful to date. Experiments reach higher energies and collect larger datasets than ever before, looking yet unsuccessfully for contradictions that could hint at solutions to some of the questions that the SM leaves unanswered: What constitutes dark matter? Why are there three generations of particles? How is gravity connected to the other forces that we know, and why is it so weak? But instead of finding contradictions, the Standard Model is further established and complemented with the last predicted pieces. What we know about the problem of the matter anti-matter imbalance mentioned earlier is that there must have been three conditions in the early universe [4]: a departure of the universe from thermal equilibrium; a violation of baryon number conservation; and a violation of C (charge) and CP (charge-parity) symmetry, which represents a violation of the symmetry between particles and anti-particles. Although CP asymmetry is incorporated into the SM and has been experimentally observed, it cannot fully account for the observed imbalance. Thus, a key task of modern particle physics experiments is to search for new sources of this symmetry violation.

Nature’s preference for symmetries, i.e. its tendency to behave invariant under transformations, has influenced the building of physical theories. For a long time physics was believed to be invariant under a reflection of spatial coordinates, that parity was conserved in all processes, until its violation was discovered in 1957 [5]. However, the desire for a universal, symmetric coupling of the weak interaction led to the introduction of the Cabibbo mixing matrix in 1963 [6]. The violation of parity has led to the belief that physics would at least behave CP invariantly, but the phenomenon of CP violation was discovered in 1964 in the decay of neutral kaons [7]. There was no room in the theory to allow for this effect with the 2×2 Cabibbo mixing matrix, which is why a third quark-family [8] was postulated. Finally, the discovery of the bottom or beauty quark in 1977 [9], and of the top quark in 1995 [10, 11] confirmed this theoretical extension and illustrated the success of this procedural scientific method. While the originally discovered type of CP violation originated from the mixing — transitions from a neutral meson into its anti-meson — CP violation in the direct decay of particles was discovered decades later in 1999 [12, 13]. Mixing in the b -meson system was established in 1987 [14] and first measurements of CP violation in this system were performed in the clean environment of e^+e^- collisions by the B -factories in 2001 [15, 16]. These

1 Introduction

experiments used decays of the *golden mode* $B^0 \rightarrow J/\psi K_S^0$, where a substantial amount of CP violation has been predicted to originate from the interference between the direct decay and the decay after mixing into a final state common to B^0 and \bar{B}^0 mesons (charge-conjugate states are implied throughout this thesis if not stated otherwise) [17, 18]. The measured quantity $\sin(2\beta)$ is associated to CP violation in this decay and can be visualized as the interior angle, β , of a triangle which comprises constraints of various quantities in the flavour sector. Precise direct and indirect measurements of $\sin(2\beta)$ have ruled out that large New Physics effects can be observed in $B^0 \rightarrow J/\psi K_S^0$, but have also established its role as a standard candle for CP -violation measurements.

Today, the Large Hadron Collider (LHC) in Geneva is leading the high-energy frontier, inducing proton-proton collisions at centre-of-mass energies of presently $\sqrt{s} = 13$ TeV and planned $\sqrt{s} = 14$ TeV. The Large Hadron Collider beauty (LHCb) experiment exploits the characteristic angular correlation in the production of hadrons containing b quarks or c quarks. It benefits from an unprecedented number of these heavy flavoured hadrons, but at the same time faces a challenging hadronic high-multiplicity environment. During the first running period of the experiment, data corresponding to 3 fb^{-1} of proton-proton collisions was collected. Its analysis has both led to results which were not accessible to the B -factories, like tests for CP violation in the system of neutral B_s^0 mesons, and precision measurements of previously measured quantities, like the aforementioned CP violation in the decay $B^0 \rightarrow J/\psi K_S^0$. Using a reconstruction of the experimentally most promising final state, $J/\psi \rightarrow \mu^+ \mu^-$ and $K_S^0 \rightarrow \pi^+ \pi^-$, has led to LHCb results which are compatible and competitive with the previous measurements of the B -factories [19]. While the perspective is to produce the world's most precise measurement repeating the analysis on the full dataset that will be available at the end of the second running period this year (2018), the potential of already recorded data is not yet fully utilized.

This thesis presents a decay-time-dependent measurement of CP violation using $B^0 \rightarrow J/\psi K_S^0$ decays, where the J/ψ meson is reconstructed from an electron pair, and $B^0 \rightarrow \psi(2S) K_S^0$ decays, where the $\psi(2S)$ meson is reconstructed from two muons. The inclusion of these additional final states does not only extend the scope of $\sin(2\beta)$ measurements at LHCb. Due to their light mass, electrons suffer radiative energy losses and, unlike muons, do not penetrate the full detector, so that they can be easily confused with other particles. Consequently, the reconstruction of final states involving electrons is accompanied by losses in efficiencies and resolutions, representing the main reasons why $B^0 \rightarrow J/\psi(e^+e^-) K_S^0$ decays have not yet been analyzed at LHCb. Being the first decay-time-dependent measurement at LHCb that uses electrons, the analysis described in this thesis demonstrates the experiment's capabilities to cope with these particles when performing precision measurements. Furthermore, the analysis of $B^0 \rightarrow \psi(2S) K_S^0$ decays can be regarded as an independent test of CP violation, as it might be affected differently by higher-order corrections.

Collaboration and Publication

Together with two other contact authors from the Dortmund LHCb group, Vanessa Müller and Patrick Mackowiak, the analysis has been published as:

R. Aaij *et al.*, *Measurement of CP violation in $B^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ decays*, JHEP **11** (2017) 170, DOI: 10.1007/JHEP11(2017)170, arXiv:1709.03944 [hep-ex]

Because of the complexity and the large scope of the analysis, it has been performed in close collaboration and under constant review by other members of the LHCb collaboration, most notably Vanessa Müller and Dr. Julian Wishahi. In particular, Vanessa has performed the studies on the decay-time resolution and acceptance for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode on simulated samples (see Sec. 5.4), she prepared the full $B^0 \rightarrow \psi(2S)K_S^0$ dataset (see Ch. 6), and she reviewed and cross-checked the decay-time-dependent fit (see Ch. 7). She too will publish these results within her own Ph.D. thesis:

V. Müller, *Improvement of the LHCb measurement of $\sin(2\beta)$ in additional B^0 decay modes*, Ph.D. thesis, TU Dortmund University, in preparation

Furthermore, Alex Seuthe has prepared the flavour-tagging calibration for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode (see Sec. 5.5) as part of his master thesis [22], which has been supervised by the author. Alex Birnkraut and Frank Meier contributed by providing cross-checks of the result through their own independent fitter implementations (see Sec. 7.5.5). Corrections for effects from the K^0 - \bar{K}^0 system (see Sec. 7.6) have been prepared with the help of Jeroen van Tilburg and Julian Wishahi. The bachelor theses of Konrad Mielke [23] and Lukas Nickel [24], and the master thesis of Christopher Hasenberg [25] have contributed to various aspects of the analysis. Although emphasis is put on the parts that reflect the author's own work, a comprehensive picture can only be provided by presenting all contributions.

Outline

The theoretical foundation of the analysis is laid in Ch. 2 and followed by a detailed presentation of the LHCb experiment in Ch. 3. While most of the analysis tools and techniques are discussed at the appropriate place and as soon as they are needed, Ch. 4 discusses those which are used in a wider context. Chs. 5 and 6 elaborate on the preparation of the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ datasets, which is followed by the presentation of the decay-time-dependent fit used to measure CP violation, the validation of the fit, and the assignment of systematic uncertainties in Ch. 7. Finally, Ch. 8 presents combinations with the previous LHCb results and the thesis closes with a conclusion and an outlook in Ch. 9.

2 Flavour theory

The Standard Model (SM) is the relativistic quantum field theory describing the fundamental constituents of nature and their interactions [1–3]. Lorentz invariance of the theory implies that it needs to fulfill the CPT theorem [26, 27], stating the requirement of invariance under the combined discrete symmetry operations of C , P , and T . Here, C inverts all charges, P inverts spatial coordinates $\vec{r} \rightarrow -\vec{r}$, and T inverts time $t \rightarrow -t$. Despite the invariance under this combined transformation, the SM allows for violation of the individual symmetries and thus also for CP violation. In systems of neutral mesons, like the B^0 - \bar{B}^0 system, oscillations between particle and antiparticle occur, allowing to assess different manifestations of CP violation. If a final state can be reached by both mesons, CP violation can be measured in the interference between the direct decay and the decay after oscillation to the common final state. For this type of CP violation the decays $B^0 \rightarrow [c\bar{c}]K_S^0$ are considered ideal and referred to as *golden modes* of the B^0 - \bar{B}^0 system [17, 18], because other sources of CP violation are strongly suppressed. These decays provide precise experimental access to the SM parameter $\sin(2\beta)$.

The formalism in this chapter follows similar presentations in Ref. [28–31].

2.1 Standard Model

The constituents of the SM are referred to as elementary particles and are shown in Fig. 2.1. They are categorized as fermions, carrying spin $1/2$, and bosons, carrying integer spins of 0 (scalar boson) or 1 (vector boson).

Fermions, which constitute matter, can be categorized as six leptons (ℓ) and six quarks (q), and their antiparticles (\bar{q} , $\bar{\ell}$), which constitute anti-matter, and whose fields are obtained by CP transformation of the corresponding particle fields. The six leptons are divided into three generations. Each consists of a charged lepton of unitary charge (e , μ , τ) and an electrically neutral lepton called neutrino (ν_e , ν_μ , ν_τ), which hence does not interact electromagnetically. Also the quarks are divided into three generations, and carry fractional electrical charges of $+2/3$ in the case of up-type quarks (u , c , t) and $-1/3$ in the case of down-type quarks (d , s , b). Besides interacting electroweakly, quarks couple to the strong interaction through a property referred to as color charge.

The fundamental bosons are the spin-1 force mediators (W^\pm , Z^0) and the spin-0 Higgs particle (H), which has recently been discovered [33]. The fields of the force mediators are derived by imposing invariance of the theory under local $SU(3) \otimes SU(2) \otimes U(1)$ gauge transformations. Here, $SU(3)$ is the gauge group of quantum chromodynamics (QCD), associated with eight gluons (g), and

2 Flavour theory

Mass	2.2 MeV/c ²	1.28 GeV/c ²	173.1 GeV/c ²	0	125.1 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	4.7 MeV/c ²	96 MeV/c ²	4.18 GeV/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	
511 keV/c ²	105.7 MeV/c ²	1.777 GeV/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		
Leptons	<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 electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	W W boson	
				Gauge bosons	

Figure 2.1 – Overview of the fundamental particles and their properties as described by the Standard Model, taken from Ref. [32] and updated according to Ref. [31].

$SU(2) \otimes U(1)$ is the gauge group of the unified electroweak interaction, associated with the massless photon (γ), mediating the electromagnetic interaction, and the massive W^+ , W^- , and Z^0 , mediating the weak interaction. The Higgs particle is a consequence of the Higgs mechanism [34, 35], using the formalism of spontaneous symmetry breaking to provide masses to the exchange bosons of the weak interaction.

Composite particles of quarks, called hadrons, are formed under the influence of the strong interaction: Mesons ($q\bar{q}$), baryons ($qq\bar{q}$ or $\bar{q}q\bar{q}$), and recently, also first exotic hadrons, pentaquarks ($qqqq\bar{q}$) [36] and tetraquarks ($q\bar{q}q\bar{q}$) [37, 38] have been observed by LHCb.

2.2 CP violation

The origin of CP violation in the SM lies in the mechanism of fermion mass acquisition: Couplings of the left-handed doublets and the right-handed singlets to the Higgs field, the Yukawa couplings, are used to maintain local gauge invariance under $SU(2) \otimes U(1)$ transformations. After spontaneous symmetry breaking, the fermion mass terms arise through couplings of their fields to the vacuum expectation value of the Higgs field. When considering multiple generations, the coupling constants become arbitrary matrices in generation space, which are diagonalized in

order to change to the basis of the fermion mass eigenstates. Due to the massless neutrinos, the diagonalization can be performed in a way so that no mixing matrix occurs in the leptonic sector. However, this is not the case for the quark sector. Here, the Lagrangian of the charged current, describing the coupling of the quark mass eigenstates to the charged force carriers of the weak interactions, reads

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

The diagonalization matrices have been merged into the unitary CKM matrix, \mathbf{V}_{CKM} , which now transforms the down-type-quark mass eigenstates to their weak eigenstates as

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

It becomes evident that V_{ij} and $|V_{ij}|^2$ are proportional to the amplitude and probability of $i \rightarrow j$ quark transitions, respectively. By definition the CKM matrix is complex and unitary. Of the nine parameters that each unitary matrix has, only three parameters are really free. An exact parametrization, the Chau-Keung parametrization [39], shows that \mathbf{V}_{CKM} can be interpreted as a complex rotation matrix with three angles, $\theta_{ij} \in [0, \frac{\pi}{2}]$, and one phase, δ . It is given by

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

where s_{ij} and c_{ij} are shorthands for $\sin \theta_{ij}$ and $\cos \theta_{ij}$, respectively. The phase δ violates the CP symmetry of the SM Lagrangian and is therefore the only cause of CP violating effects in the theory.

A prominent parametrization of the CKM matrix is the Wolfenstein parametrization [40]. The exact version of this parametrization [41] is obtained by defining the four parameters $\lambda \approx 0.23$, $A \approx 0.83$, $\rho \approx 0.16$ and $\eta \approx 0.35$ (values from Ref. [42]) as

$$\lambda = s_{12}, \quad (2.4)$$

$$A\lambda^2 = s_{23}, \quad (2.5)$$

$$A\lambda^3(\rho - i\eta) = s_{13}e^{-i\delta}. \quad (2.6)$$

In order to expose the hierarchy of the CKM matrix elements, the Wolfenstein parametrization is most commonly used as a series expansion in the parameter λ , which leads to

$$\begin{aligned}
 \mathbf{V}_{\text{CKM}} = & \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \\
 & + \begin{pmatrix} -\frac{1}{8}\lambda^4 & 0 & 0 \\ \frac{1}{2}A^2\lambda^5[1 - 2(\rho + i\eta)] & -\frac{1}{8}\lambda^4(1 + 4A^2) & 0 \\ \frac{1}{2}A\lambda^5(\rho + i\eta) & \frac{1}{2}A\lambda^4[1 - 2(\rho + i\eta)] & -\frac{1}{2}A^2\lambda^4 \end{pmatrix} \\
 & + O(\lambda^6),
 \end{aligned} \tag{2.7}$$

where, when truncating this expansion, the unitarity of \mathbf{V}_{CKM} is only approximately satisfied. The magnitude for the probabilities of the quark transitions can now be easily identified from Eq. (2.7): Transitions within a generation are most likely, because they are of $\mathcal{O}(1)$. In contrast, transitions to other generations are suppressed by $\mathcal{O}(\lambda)$, in the case of 1st \leftrightarrow 2nd generation transitions, $\mathcal{O}(\lambda^2)$, in the case of 2nd \leftrightarrow 3rd generation transitions, and even $\mathcal{O}(\lambda^3)$, for 1st \leftrightarrow 3rd generation transitions. It becomes also evident that only $b \leftrightarrow u$ and $t \leftrightarrow d$ transitions involve complex matrix elements in first order of the expansion, which is the primary reason why the B^0 meson with a valence quark content of $|\bar{b}d\rangle$ serves especially well for studies of CP violation.

The unitarity of the CKM matrix leads to six relations among its elements which equal to zero. These sums of complex numbers can be identified as triangles in the complex plane, so called unitarity triangles. The triangles show identical areas, corresponding to half of the Jarlskog's parameter [43],

$$J = \pm \text{Im} \left(V_{ik} V_{jl} V_{il}^* V_{jk}^* \right) \quad \text{with} \quad i \neq j, l \neq k, \tag{2.8}$$

which can be interpreted as a measure of the amount of CP violation in the SM. However, as all relations for the unitarity triangles involve different matrix elements, the length of the sides of the triangles can be very different. A triangle that shows side lengths of the same order in λ , and which hence possesses interior angles that are of $\mathcal{O}(1)$, is defined by

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{2.9}$$

It is usually referred to as *the* unitarity triangle (UT), or the (bd) -triangle, because any occurring matrix element is associated with the transition of a b or d quark. By dividing Eq. (2.9) by $V_{cd}V_{cb}^*$ the triangle is aligned in the complex $\bar{\rho}$ - $i\bar{\eta}$ -plane, as shown in Fig. 2.2. This triangle has one side of unit length along the vertical axis of the coordinate system, and an apex, defined by

$$\bar{\rho} + i\bar{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}. \tag{2.10}$$

The three interior angles of the triangle are

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \quad (2.11)$$

while its side lengths are defined by

$$R_t = \left|\frac{V_{td}V_{tb}^*}{V_{cb}V_{cb}^*}\right|, \quad R_u = \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right|, \quad R_c = \left|\frac{V_{cd}V_{cb}^*}{V_{cd}V_{cb}^*}\right|. \quad (2.12)$$

The side lengths and angles of the triangle are observables in flavour-physics processes that involve the corresponding matrix elements. The general strategy of indirect searches for New Physics is to find contradictions in the measurements, that would lead to a nonclosing triangle or an interior angle exceeding π . For convenience, the triangle apex is overconstrained through a combined fit of the measurements as illustrated in Fig. 2.2. Another triangle is obtained from the unitarity relation

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0. \quad (2.13)$$

While the triangle itself is degenerated due to the different side lengths, its small interior angle

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) \quad (2.14)$$

is of $\mathcal{O}(\lambda^2)$ and particularly interesting for studies of CP violation in the B_s^0 system.

2.3 Neutral-meson system

The following describes the formalism for transitions of neutral mesons and is based on similar presentations in Refs. [29, 31].

The decay amplitudes of a neutral meson, P^0 , and its CP conjugate anti-meson, \bar{P}^0 , to a final state, f , and the CP conjugate final state, \bar{f} , are defined via the matrix elements

$$\begin{aligned} A_f &= \langle f|H|P^0\rangle, & \bar{A}_f &= \langle f|H|\bar{P}^0\rangle, \\ A_{\bar{f}} &= \langle \bar{f}|H|P^0\rangle, & \bar{A}_{\bar{f}} &= \langle \bar{f}|H|\bar{P}^0\rangle, \end{aligned} \quad (2.15)$$

where the Hamiltonian, H , governs weak interactions. The CP conjugation of the initial and final states, considering $(CP)^2 = 1$, allows to define two arbitrary and unobservable phases, ξ_{P^0} and ξ_f , so that

$$\begin{aligned} CP|P^0\rangle &= e^{+i\xi_{P^0}}|\bar{P}^0\rangle, & CP|f\rangle &= e^{+i\xi_f}|\bar{f}\rangle, \\ CP|\bar{P}^0\rangle &= e^{-i\xi_{P^0}}|P^0\rangle, & CP|\bar{f}\rangle &= e^{-i\xi_f}|f\rangle. \end{aligned} \quad (2.16)$$

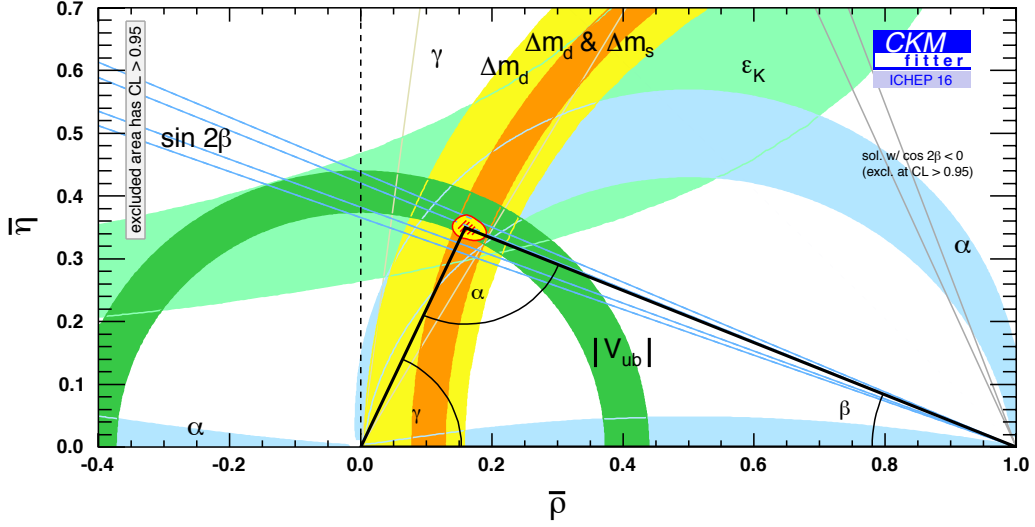


Figure 2.2 – The CKM triangle in the complex $\bar{\rho}$ - $\bar{\eta}$ -plane [44]. The colored bands represent confidence regions of various measurements in the flavour sector that help to constrain the triangle apex. The red hatched region corresponds to the 68% confidence level of a fit to these measurements and is in no tension with any of the direct measurements.

If the initial state is composed of a superposition of P^0 and \bar{P}^0 , as

$$|\psi(t)\rangle = \psi_1(t)|P^0\rangle + \psi_2(t)|\bar{P}^0\rangle, \quad (2.17)$$

it will evolve in time, allowing for a change in the compositions of P^0 and \bar{P}^0 , as well as for the acquisition of further decay final states, $\{f_1, f_2, \dots\}$, so that

$$|\psi(t)\rangle = a(t)|P^0\rangle + b(t)|\bar{P}^0\rangle + c_1(t)|f_1\rangle + c_2(t)|f_2\rangle + \dots \quad (2.18)$$

In order to compute $a(t)$ and $b(t)$ the Weisskopf-Wigner approximation [45, 46] can be used to describe the meson system using an effective 2×2 Hamiltonian, \mathbf{H} , in a Schrödinger-like equation

$$i \frac{d}{dt} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \mathbf{H} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \left(\mathbf{M} - \frac{i}{2} \mathbf{\Gamma} \right) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \quad (2.19)$$

While \mathbf{H} is composed of two Hermitian matrices \mathbf{M} and $\mathbf{\Gamma}$, \mathbf{H} itself is not Hermitian to allow for the decay of the P^0 and \bar{P}^0 states. Invariance under CPT transformation dictates that these two states have equal masses, m , and decay widths, Γ , and hence that \mathbf{H} has identical diagonal elements so that it can be written as

$$\mathbf{H} = \begin{pmatrix} m - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & m - \frac{i}{2}\Gamma \end{pmatrix}. \quad (2.20)$$

The off-diagonal elements of \mathbf{H} are connected to transitions from one flavour eigenstate into its conjugate. In the SM these flavour changing transitions can be calculated from two types of box diagrams (see Sec. 2.5.1): Short-distance contributions via off-shell states constitute M_{12} , and virtual intermediate decays to on-shell states contribute to Γ_{12} . The relative phase difference that arises between these two kinds of contributions is defined as

$$\phi = \arg(M_{12}\Gamma_{12}^*) = \arg(e^{i\phi_M}e^{-i\phi_\Gamma}) = \phi_M - \phi_\Gamma, \quad (2.21)$$

and is connected to CP violation in the mixing (see Sec. 2.4.2). Diagonalization of \mathbf{H} results in two complex eigenvalues of defined mass and decay width, given as

$$\begin{aligned} \mu_{1,2} &= m - \frac{i}{2}\Gamma \pm \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) \left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)} \\ &= m \pm \text{Re}(F) - \frac{i}{2}(\Gamma \mp 2\text{Im}(F)), \end{aligned} \quad (2.22)$$

where the term F is defined

$$F = \sqrt{\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) \left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}. \quad (2.23)$$

The corresponding eigenstates of \mathbf{H} are commonly distinguished by their well defined masses, in the case of the B^0 - \bar{B}^0 or B_s^0 - \bar{B}_s^0 system, or by their well defined decay widths, in the case of the K^0 - \bar{K}^0 system. Following the convention for neutral B meson systems, a light mass eigenstate, $|P_L^0\rangle$, and a heavy mass eigenstate $|P_H^0\rangle$, with corresponding eigenvalues, μ_L and μ_H , are obtained as

$$\begin{aligned} |P_L^0\rangle &= p|P^0\rangle + q|\bar{P}^0\rangle, & \mu_L &= m_L - \frac{i}{2}\Gamma_L, \\ |P_H^0\rangle &= p|P^0\rangle - q|\bar{P}^0\rangle, & \mu_H &= m_H - \frac{i}{2}\Gamma_H, \end{aligned} \quad (2.24)$$

and characterized by their mass difference, Δm , and their decay width difference, $\Delta\Gamma$, which are given as

$$\begin{aligned} \Delta m &= m_H - m_L = \text{Re}(\mu_H - \mu_L) = 2\text{Re}(F), \\ \Delta\Gamma &= \Gamma_H - \Gamma_L = -2\text{Im}(\mu_H - \mu_L) = -4\text{Im}(F). \end{aligned} \quad (2.25)$$

The sign of Δm is positive by definition, whereas the sign of $\Delta\Gamma$ requires experimental determination. The parameters p and q fulfill the normalization condition $|p|^2 + |q|^2 = 1$ and are related to the off-diagonal elements of \mathbf{H} by

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} = \frac{\Delta m - \frac{i}{2}\Delta\Gamma}{2(M_{12} - \frac{i}{2}\Gamma_{12})}. \quad (2.26)$$

2 Flavour theory

The time evolution of the two states is given by

$$\begin{aligned} |P_L^0(t)\rangle &= e^{-im_L t} e^{-\frac{\Gamma_L}{2} t} |P_L\rangle, \\ |P_H^0(t)\rangle &= e^{-im_H t} e^{-\frac{\Gamma_H}{2} t} |P_H\rangle. \end{aligned} \quad (2.27)$$

Using Eqs. (2.24) and (2.27), it is possible to examine the time evolution of initial states with well-defined quark flavour, obtaining

$$\begin{aligned} |P^0(t)\rangle &= g_+(t)|P^0\rangle - \frac{q}{p}g_-(t)|\bar{P}^0\rangle, \\ |\bar{P}^0(t)\rangle &= g_+(t)|\bar{P}^0\rangle - \frac{q}{p}g_-(t)|P^0\rangle, \end{aligned} \quad (2.28)$$

where g_{\pm} are defined as

$$g_{\pm}(t) = \frac{1}{2} \left(e^{-im_H t} e^{-\frac{\Gamma_H}{2} t} \pm e^{-im_L t} e^{-\frac{\Gamma_L}{2} t} \right). \quad (2.29)$$

The flavour content of an initially pure state is not constant in time, but changes according to $|\langle P^0|P^0(t)\rangle| \sim |g_+(t)|^2$ and $|\langle \bar{P}^0|P^0(t)\rangle| \sim |g_-(t)|^2$ for initial P^0 states, and analogously for initial \bar{P}^0 states. This change in flavour content is referred to as flavour oscillation because of the trigonometric dependence in

$$|g_{\pm}(t)|^2 = \frac{e^{-\Gamma t}}{2} \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) \pm \cos(\Delta m t) \right], \quad (2.30)$$

causing an oscillation of frequency Δm . These flavour oscillations can be directly measured in flavour-specific decays, i.e. decays that can only be reached by either the P^0 or the \bar{P}^0 state.

For the study of CP violation in decay processes involving the final state f and its conjugate \bar{f} , the differential decay rates

$$\begin{aligned} \Gamma(P^0(t) \rightarrow f) &= |\langle f|H|P^0(t)\rangle|^2, & \Gamma(\bar{P}^0(t) \rightarrow f) &= |\langle f|H|\bar{P}^0(t)\rangle|^2, \\ \Gamma(P^0(t) \rightarrow \bar{f}) &= |\langle \bar{f}|H|P^0(t)\rangle|^2, & \Gamma(\bar{P}^0(t) \rightarrow \bar{f}) &= |\langle \bar{f}|H|\bar{P}^0(t)\rangle|^2, \end{aligned} \quad (2.31)$$

are of interest. Defining

$$\lambda_f = \frac{1}{\bar{\lambda}_f} = \frac{q \bar{A}_f}{p A_f} \quad \text{and} \quad \bar{\lambda}_{\bar{f}} = \frac{1}{\lambda_{\bar{f}}} = \frac{p A_{\bar{f}}}{q \bar{A}_{\bar{f}}}, \quad (2.32)$$

these decay rates are obtained as

$$\begin{aligned}
 \Gamma(P^0(t) \rightarrow f) &= |A_f|^2 \left(|g_+(t)|^2 + |\lambda_f|^2 |g_-(t)|^2 - 2\text{Re}(\lambda_f g_+^*(t) g_-(t)) \right), \\
 \Gamma(\bar{P}^0(t) \rightarrow f) &= |A_f|^2 \left| \frac{p}{q} \right|^2 \left(|g_-(t)|^2 + |\lambda_f|^2 |g_+(t)|^2 - 2\text{Re}(\lambda_f g_+(t) g_-^*(t)) \right), \\
 \Gamma(P^0(t) \rightarrow \bar{f}) &= |\bar{A}_{\bar{f}}|^2 \left| \frac{q}{p} \right|^2 \left(|g_-(t)|^2 + |\bar{\lambda}_{\bar{f}}|^2 |g_+(t)|^2 - 2\text{Re}(\bar{\lambda}_{\bar{f}} g_+(t) g_-^*(t)) \right), \\
 \Gamma(\bar{P}^0(t) \rightarrow \bar{f}) &= |\bar{A}_{\bar{f}}|^2 \left(|g_+(t)|^2 + |\bar{\lambda}_{\bar{f}}|^2 |g_-(t)|^2 - 2\text{Re}(\bar{\lambda}_{\bar{f}} g_+^*(t) g_-(t)) \right),
 \end{aligned} \tag{2.33}$$

where

$$g_+^*(t) g_-(t) = -\frac{e^{-\Gamma t}}{2} \left[\sinh\left(\frac{\Delta\Gamma t}{2}\right) + i \sin(\Delta m t) \right]. \tag{2.34}$$

The terms in Eq. (2.33) proportional to $g_+(t)^2$ and A^2 are associated with decays after no net oscillation. In contrast, terms proportional to $g_-(t)^2$ and $\frac{p}{q}A^2$ or $\frac{q}{p}A^2$ are linked to decays after a net oscillation to the opposite flavour meson. Terms proportional to $[g_+^*(t)g_-(t)]^{(*)}$ originate from the interference of direct decays and decays after oscillation. A substantial simplification of the decay rates can be achieved by introducing the CP observables

$$D_f = -\frac{2\text{Re}\lambda_f}{1 + |\lambda_f|^2}, C_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, S_f = \frac{2\text{Im}\lambda_f}{1 + |\lambda_f|^2}, \tag{2.35}$$

$$D_{\bar{f}} = -\frac{2\text{Re}\bar{\lambda}_{\bar{f}}}{1 + |\bar{\lambda}_{\bar{f}}|^2}, C_{\bar{f}} = \frac{1 - |\bar{\lambda}_{\bar{f}}|^2}{1 + |\bar{\lambda}_{\bar{f}}|^2}, S_{\bar{f}} = \frac{2\text{Im}\bar{\lambda}_{\bar{f}}}{1 + |\bar{\lambda}_{\bar{f}}|^2}, \tag{2.36}$$

which fulfill the relations $(D_f)^2 + (C_f)^2 + (S_f)^2 = 1$ and $(D_{\bar{f}})^2 + (C_{\bar{f}})^2 + (S_{\bar{f}})^2 = 1$. Using these definitions leads to the most common parametrization of the decay-

time-dependent decay rates as

$$\frac{\Gamma(P^0(t) \rightarrow f)}{e^{-\Gamma t}} = \frac{1}{2} |A_f|^2 \left(1 + |\lambda_f|^2\right) \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D_f \sinh\left(\frac{\Delta\Gamma}{2}t\right) + C_f \cos(\Delta mt) - S_f \sin(\Delta mt) \right], \quad (2.37)$$

$$\frac{\Gamma(\bar{P}^0(t) \rightarrow f)}{e^{-\Gamma t}} = \frac{1}{2} |A_f|^2 \left|\frac{p}{q}\right|^2 \left(1 + |\lambda_f|^2\right) \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D_f \sinh\left(\frac{\Delta\Gamma}{2}t\right) - C_f \cos(\Delta mt) + S_f \sin(\Delta mt) \right], \quad (2.38)$$

$$\frac{\Gamma(P^0(t) \rightarrow \bar{f})}{e^{-\Gamma t}} = \frac{1}{2} |\bar{A}_{\bar{f}}|^2 \left|\frac{q}{p}\right|^2 \left(1 + |\bar{\lambda}_{\bar{f}}|^2\right) \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma}{2}t\right) - C_{\bar{f}} \cos(\Delta mt) + S_{\bar{f}} \sin(\Delta mt) \right], \quad (2.39)$$

$$\frac{\Gamma(\bar{P}^0(t) \rightarrow \bar{f})}{e^{-\Gamma t}} = \frac{1}{2} |\bar{A}_{\bar{f}}|^2 \left(1 + |\bar{\lambda}_{\bar{f}}|^2\right) \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma}{2}t\right) + C_{\bar{f}} \cos(\Delta mt) - S_{\bar{f}} \sin(\Delta mt) \right]. \quad (2.40)$$

2.4 Types of CP violation in neutral-meson systems

In general there are three types of phases involved in transition amplitudes: Weak phases are CP -odd and originate from the complex couplings in the SM Lagrangian, which themselves are linked to the single phase of the CKM quark mixing matrix; Strong phases originate for example from final-state-interaction scatterings of on-shell states through the electromagnetic or strong interaction. As these interactions conserve CP , strong phases are CP -even and hence do not change sign under the CP transformation; Spurious phases are global phases in the transition amplitudes that are convention-dependent, like seen in Eq. (2.16). They do not enter any dynamics and cannot be measured. Neutral-meson decays allow for three different manifestations of CP violation: CP violation in the decay (direct CP violation), CP violation in the mixing (indirect CP violation), and CP violation in the interference between the direct decay and the decay after mixing (interference CP violation).

2.4.1 Direct CP violation

Direct CP violation is present when the absolute value of the decay amplitude of a meson to a given final state differs from the absolute value of the decay amplitude for the CP conjugate process. This means that

$$\left| \frac{A_f}{\bar{A}_{\bar{f}}} \right| \neq 1, \quad (2.41)$$

allowing both of the corresponding direct decays to occur at different rates. For this to be the case it requires the interference of direct decay amplitudes with moduli $|A_i|$, strong phases δ_i and weak phases ϕ_i . A parametrization of a process with two contributing Feynman diagrams could then be

$$\begin{aligned} A_f &= A_1 e^{i(\delta_1 + \phi_1)} + A_2 e^{i(\delta_2 + \phi_2)}, \\ \bar{A}_{\bar{f}} &= A_1 e^{i(\delta_1 - \phi_1)} + A_2 e^{i(\delta_2 - \phi_2)}, \end{aligned} \quad (2.42)$$

which satisfies Eq. (2.41) for $\delta_1 \neq \delta_2$ and $\phi_1 \neq \phi_2$: This implies that a weak as well as a strong phase difference are required between underlying processes, in order for direct CP violation to occur.

Due to the abundance of oscillations for charged mesons, this is the only type of CP violation that can occur in such systems. Concerning systems of neutral B mesons, direct CP violation can contribute and the effect has been significantly observed in decays of the B^0 and B_s^0 to the $K^+ \pi^-$ final state and its charge conjugate at LHCb [47]. However, direct CP violation is strongly suppressed whenever a single weak phase dominates the transition amplitude for the direct decay, as for example in $B^0 \rightarrow [c\bar{c}]K_S^0$ decays (see Sec. 2.5).

2.4.2 Indirect CP violation

Indirect CP violation is characterized by different transition rates between a meson and its anti-meson in neutral-meson systems, and thus by $\Gamma(P^0 \rightarrow \bar{P}^0) \neq \Gamma(\bar{P}^0 \rightarrow P^0)$. These transitions are described by the off-diagonal elements of the effective Hamiltonian defined in Eq. (2.20), which can be parametrized by

$$M_{12} = |M_{12}| e^{i\phi_M}, \quad \Gamma_{12} = |\Gamma_{12}| e^{i\phi_\Gamma}. \quad (2.43)$$

Different transition rates between P^0 and \bar{P}^0 occur if $|\Gamma_{12}/M_{12}| \neq 0$ and $\phi = \phi_M - \phi_\Gamma \neq 0$, which, following Eq. (2.26), implies

$$\left| \frac{q}{p} \right| \neq 1. \quad (2.44)$$

Measurements of this type of CP violation use flavour-specific decays, i.e. decays that can only be reached by either P^0 or \bar{P}^0 .

At LHCb measurements using muonic semileptonic B meson decays were performed, where the $\mu^+ X$ and the $\mu^- X$ final states can only be reached directly from the B_q and \bar{B}_q , respectively ($q = d$ for B^0 , $q = s$ for B_s^0). Hence, the asymmetry

$$a_{\text{sl}}^q(t) = \frac{\Gamma(\bar{B}_q^0 \rightarrow \mu^+ X) - \Gamma(B_q \rightarrow \mu^- X)}{\Gamma(\bar{B}_q^0 \rightarrow \mu^+ X) + \Gamma(B_q \rightarrow \mu^- X)} = \frac{1 - \left(\frac{q}{p}\right)^2}{1 + \left(\frac{q}{p}\right)^2}, \quad (2.45)$$

is used, where the final state in each individually measured decay rate can only be reached through mixing. SM predictions state that this type of CP violation is

2 Flavour theory

small for the B^0 and B_s^0 system, as $a_{\text{sl}}^q = \mathcal{O}(10^{-4})$ [48, 49]. Experiments confirm these predictions and the results of the mentioned LHCb measurements [50, 51],

$$\begin{aligned} a_{\text{sl}}^d &= (-0.02 \pm 0.019 \pm 0.30) \% , \\ a_{\text{sl}}^s &= (+0.39 \pm 0.26 \pm 0.20) \% , \end{aligned} \quad (2.46)$$

show that indirect CP violation in neutral B meson systems can be neglected at the current level of experimental precision.

2.4.3 Interference CP violation

Interference CP violation can occur whenever a final state f can be reached by the meson as well as by its anti-meson. Then, the decay amplitudes for the direct decay, $P^0 \rightarrow f$, and the decay after mixing, $P^0 \rightarrow \bar{P}^0 \rightarrow f$, interfere. Even in the absence of direct and indirect CP violation, i.e. if $|\lambda_f| = 1$, it can be present and is characterized by

$$\text{Im } \lambda_f \neq 0. \quad (2.47)$$

When considering the case of decays into a CP eigenstate, f_{CP} , Eqs. (2.37) to (2.40) reduce to only two independent decay rates, as $A_{\bar{f}} = A_f$ and $\bar{A}_{\bar{f}} = \bar{A}_f$ and therefore also $\lambda_{\bar{f}} = \lambda_f = 1/\bar{\lambda}_{\bar{f}} = 1/\bar{\lambda}_f$. Furthermore, if only a single weak phase, ϕ_A , and strong phase, δ_A , contribute, the decay amplitude of the meson can be written as $A_f = Ae^{i(\phi_A + \delta_A)}$, which corresponds to the absence of direct CP violation as discussed in Sec. 2.4.1. The amplitude can then be connected to the decay amplitude of the anti-meson through CP transformation, $A_{\bar{f}} = \eta_f Ae^{i(-\phi_A + \delta_A)}$, where η_f is the CP eigenvalue of the final state f_{CP} . Moreover, in the limit of negligible CP violation in the mixing, the term $q/p = e^{i\phi_M}$ is a pure phase, so that following Eq. (2.35), also $\lambda_f = \eta_f e^{i(\phi_M + 2\phi_A)}$ is a pure phase. For decays of initial P^0 and \bar{P}^0 into a CP eigenstate, f_{CP} , it is common to define the decay-time-dependent asymmetry as

$$\begin{aligned} \mathcal{A}_{f_{CP}}(t) &= \frac{\Gamma(P^0(t) \rightarrow f_{CP}) - \Gamma(\bar{P}^0(t) \rightarrow f_{CP})}{\Gamma(P^0(t) \rightarrow f_{CP}) + \Gamma(\bar{P}^0(t) \rightarrow f_{CP})} \\ &= \frac{S_{f_{CP}} \sin(\Delta m t) - C_{f_{CP}} \cos(\Delta m t)}{\cosh(\Delta \Gamma t/2) + D_{f_{CP}} \sinh(\Delta \Gamma t/2)}, \end{aligned} \quad (2.48)$$

where $S_{f_{CP}} = \eta_f \sin(\phi_M + 2\phi_A)$, $C_{f_{CP}} = 0$ in and $D_{f_{CP}} = \pm \sqrt{1 - (S_{f_{CP}})^2}$ in the limit of no direct and indirect CP violation.

Preferred modes for the measurement of this type of CP violation are decays that are dominated by a single weak phase, like e.g. $B^0 \rightarrow J/\psi K_S^0$ in the B^0 system and $B_s^0 \rightarrow J/\psi \phi$ in the B_s^0 system. The phenomenology of $B^0 \rightarrow [c\bar{c}]K_S^0$ that includes the first of these decays is discussed in the next section. The latter of these modes requires the disentanglement of different CP eigenvalues through an angular analysis, due to the vector-vector final state and the associated multiple possibilities of orbital angular momenta.

2.5 $B^0 \rightarrow [c\bar{c}]K_S^0$ decays

The mode $B^0 \rightarrow J/\psi K_S^0$ is considered ideal for studies of CP violation in the interference for multiple reasons: It provides a clean measurement, because it is dominated by a single weak phase, while higher order contributions are suppressed, which could contribute additional strong and weak phases. Furthermore, the vector-pseudoscalar final state has a definite CP eigenvalue of $\eta_{J/\psi K_S^0} = -1$ and thus no angular analysis is required to disentangle contributions of different eigenvalues. Lastly, it is experimentally well reconstructible in the $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0(\pi^+\pi^-)$ final state due its high branching ratio and reconstruction efficiencies. A natural extension are other final state reconstructions, like $J/\psi \rightarrow e^+e^-$, and the inclusion of different charmonium resonances $[c\bar{c}]$, as in the decay $B^0 \rightarrow \psi(2S)K_S^0$. The common phenomenology of all these $B^0 \rightarrow [c\bar{c}]K_S^0$ decays allows to measure the UT angle $\sin(2\beta)$ defined in Eq. (2.11). Following a discussion of the mixing and decay phenomenology that is strongly inspired by Ref. [29], this section ends with an overview of the experimental status on the parameter.

2.5.1 B^0 - \bar{B}^0 mixing phenomenology

As derived in Sec. 2.3, neutral-meson mixing enters the time-dependent decay rates of neutral-meson decays through the complex term q/p . These parameters are related to the off-diagonal elements of the effective Hamiltonian describing the neutral-meson system that is given in Eq. (2.26). Therefore, relating the mixing part that enters the decay to parameters of the SM requires examination of the elements M_{12} and Γ_{12} .

The matrix element M_{12} is mainly given by the short-distance contributions coming from the box diagrams in Fig. 2.3 that correspond to B^0 - \bar{B}^0 transitions. A computation of these diagrams results in

$$M_{12} = -\frac{G_F^2 m_W^2}{12\pi^2} f_{B^0}^2 m_{B^0} B_{B^0} \mathcal{F}^*, \quad (2.49)$$

where G_F denotes the Fermi constant, m_W the W boson mass, m_{B^0} the B^0 mass, f_{B^0} the weak decay constant and B_{B^0} is the bag parameter, describing

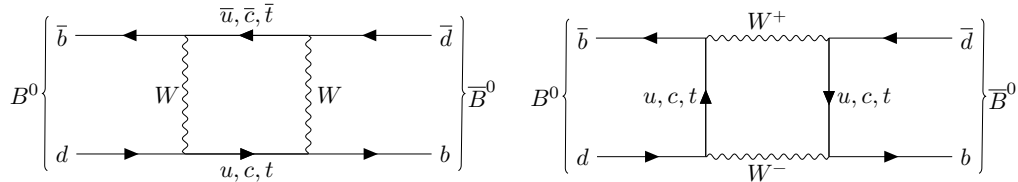


Figure 2.3 – Dominant box diagrams that correspond to mixing transitions in the B^0 - \bar{B}^0 system.

2 Flavour theory

non-perturbative QCD corrections. The function

$$\mathcal{F} = \eta_1 \lambda_c^2 S_0 \left(\frac{m_c^2}{m_W^2} \right) + \eta_2 \lambda_t^2 S_0 \left(\frac{m_t^2}{m_W^2} \right) + 2\eta_3 \lambda_c \lambda_t S_0 \left(\frac{m_c^2}{m_W^2}, \frac{m_t^2}{m_W^2} \right) \quad (2.50)$$

sums the box diagrams that contribute to the mixing, considering their CKM couplings $\lambda_\alpha = V_{\alpha b}^* V_{\alpha d}$, and perturbative QCD-correction coefficients η_i . The Inami-Lim-function [52], S_0 , satisfies

$$S_0(x_t) \gg S_0(x_c, x_t) > S_0(x_c), \quad (2.51)$$

which allows to approximate Eq. (2.49) as

$$M_{12} = -\frac{G_F^2 m_W^2}{12\pi^2} f_{B^0}^2 m_{B^0} B_{B^0} \eta_2 (V_{tb} V_{td}^*)^2 S_0 \left(\frac{m_t^2}{m_W^2} \right). \quad (2.52)$$

Contributions to Γ_{12} arise from physical states, f , to which both mesons, B^0 and \bar{B}^0 , can decay and can be calculated as

$$\Gamma_{12} = \sum_f \langle f | T | B^0 \rangle^* \langle f | T | \bar{B}^0 \rangle. \quad (2.53)$$

However, due to the high top quark mass, such transitions to physical states are only possible to c or u flavoured hadrons. These transitions are either Cabibbo-suppressed for both mesons, B^0 and \bar{B}^0 , or Cabibbo-allowed for one of them while doubly Cabibbo-suppressed for the other. Calculations find that

$$\left| \frac{\Gamma_{12}}{M_{12}} \right| \sim \mathcal{O} \left(\frac{m_b^2}{m_t^2} \right) \approx 10^{-3} \quad (2.54)$$

so that using Eqs. (2.23) and (2.25) the expressions for Δm and $\Delta\Gamma$ can be approximated as

$$\Delta m \approx 2 |M_{12}|, \quad (2.55)$$

$$\Delta\Gamma \approx 2 |\Gamma_{12}| \cos(\phi), \quad (2.56)$$

and hence, $\Delta\Gamma$ is expected to be small. Predictions of $\Delta\Gamma/\Gamma = \mathcal{O}(10^{-3})$ [48] are confirmed by the current experimental average of $\Delta\Gamma/\Gamma = -0.002 \pm 0.010$ [53]. The finding that Γ_{12} is small compared to M_{12} allows to approximate Eq. (2.26) as

$$\frac{q}{p} \approx \sqrt{\frac{M_{12}^*}{M_{12}}} = \frac{V_{tb}^* V_{td}}{V_{td}^* V_{tb}}. \quad (2.57)$$

This approximation corresponds to the absence of indirect CP violation, which is experimentally well supported, as discussed in Sec. 2.4.2.

2.5.2 Decay amplitudes and $\sin(2\beta)$

The decay amplitudes of B^0 and \bar{B}^0 into the common $[c\bar{c}]K_S^0$ final state are defined as the matrix elements

$$\bar{A}_{[c\bar{c}]K_S^0} = \langle [c\bar{c}]K_S^0 | T | \bar{B}^0 \rangle, \quad A_{[c\bar{c}]K_S^0} = \langle [c\bar{c}]K_S^0 | T | B^0 \rangle. \quad (2.58)$$

The K_S^0 in the final state is not a flavour eigenstate, but the short lived mass eigenstate of the K^0 - \bar{K}^0 system, given as the superposition

$$|K_S^0\rangle = p_K |K^0\rangle - q_K |\bar{K}^0\rangle, \quad (2.59)$$

where q_K/p_K can be calculated similar to the way described in the previous section to obtain

$$\frac{q_K}{p_K} = -\frac{V_{cs}^* V_{cd}}{V_{cs} V_{cd}^*}. \quad (2.60)$$

Consequently, the common $[c\bar{c}]K_S^0$ final state can only be reached through the $B^0 \rightarrow [c\bar{c}]K^0$ and $\bar{B}^0 \rightarrow [c\bar{c}]\bar{K}^0$ decays with a subsequent transition of the K^0/\bar{K}^0 into its mass eigenstate. The matrix elements defined in Eq. (2.58) can then be expressed as

$$\langle [c\bar{c}]K_S^0 | T | B^0 \rangle = +\frac{1}{2p_K} \langle [c\bar{c}]K^0 | T | B^0 \rangle, \quad (2.61)$$

$$\langle [c\bar{c}]K_S^0 | T | \bar{B}^0 \rangle = +\frac{1}{2q_K} \langle [c\bar{c}]\bar{K}^0 | T | \bar{B}^0 \rangle. \quad (2.62)$$

The leading order contributions to these transitions are given through a tree-level $b \rightarrow c\bar{c}s$ transition as well as a penguin diagram with an internal up-type quark loop, as shown in Fig. 2.4. Following Ref. [54], the amplitude for the decay in terms of the relevant CKM couplings can be written as

$$A(B^0 \rightarrow [c\bar{c}]K^0) = V_{cs}V_{cb}^*T_c + V_{us}V_{ub}^*P_u + V_{cs}V_{cb}^*P_c + V_{ts}V_{tb}^*P_t, \quad (2.63)$$

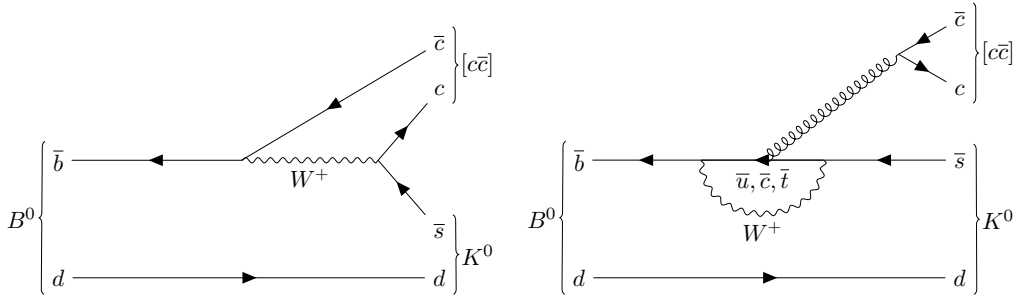


Figure 2.4 – Feynman diagrams for the leading-order contribution (left) and the dominant penguin contribution (right) of $B^0 \rightarrow [c\bar{c}]K_S^0$ decays.

2 Flavour theory

where T_c denotes the strong amplitude that is associated with the tree-level diagram, and P_q denotes the strong amplitude that is related to the penguin topology with an internal quark of flavour q . It is useful to reorder the terms in this expression to identify weak phases. Using CKM unitarity and the definitions of the UT angle γ and the UT side R_u from Eq. (2.11) and Eq. (2.12), one obtains

$$A(B^0 \rightarrow [c\bar{c}]K^0) = V_{cs}V_{cb}^* [T_c + P_c - P_t] \left(1 - \frac{V_{us}V_{cd}}{V_{cs}V_{ud}} R_u e^{i\gamma} \left[\frac{P_u - P_t}{T_c + P_c - P_t} \right] \right). \quad (2.64)$$

By using the Wolfenstein parametrization up to terms of $\mathcal{O}(\lambda^2)$ and introducing the definitions

$$\mathcal{A} = A\lambda^2 [T_c + P_c - P_t], \quad \epsilon = \frac{\lambda^2}{1 - \lambda^2}, \quad ae^{i\theta} = R_u \left[\frac{P_u - P_t}{T_c + P_c - P_t} \right], \quad (2.65)$$

the amplitude can be approximated as

$$A(B^0 \rightarrow [c\bar{c}]K^0) = \left(1 - \frac{\lambda^2}{2} \right) \mathcal{A} (1 + \epsilon ae^{i\theta} e^{+i\gamma}). \quad (2.66)$$

The only weak phase at this level of approximation is γ , which enters with a suppression of $\epsilon \approx 0.05$ [44]. The amplitude for the CP conjugate decay is obtained via CP transformation, which only changes the sign of the weak phase γ but not of the strong phase θ , so that

$$A(\bar{B}^0 \rightarrow [c\bar{c}]\bar{K}^0) = \left(1 - \frac{\lambda^2}{2} \right) \mathcal{A} (1 + \epsilon ae^{i\theta} e^{-i\gamma}). \quad (2.67)$$

When relating the decay amplitudes for the decay into the $[c\bar{c}]K_S^0$ final state to each other, it is important to consider the CP eigenvalue of the final state, which enters the equation as a relative sign. Since the scalar B^0 decays into a $[c\bar{c}]$ resonance of spin 1 (considering the J/ψ and the $\psi(2S)$) and a K_S^0 of spin 0, a relative angular momentum of $l = 1$ between both is needed for conservation. When reconstructing the K_S^0 from two π , both final state particles are CP -even, causing a negative CP eigenvalue of the system. Using the results on the mixing in the B^0 - \bar{B}^0 and the K^0 - \bar{K}^0 system from Eqs. (2.57) and (2.60), and the amplitudes of the direct decay to the $[c\bar{c}]K^0$ and $[c\bar{c}]\bar{K}^0$ final states from Eqs. (2.66) and (2.67), the parameter $\lambda_{[c\bar{c}]K_S^0}$ entering the time-dependent decay rates can be computed as

$$\begin{aligned} \lambda_{[c\bar{c}]K_S^0} &= \frac{q \bar{A}_f}{p A_f} = - \frac{V_{tb}^* V_{td} V_{cd}^* V_{cs}}{V_{tb} V_{td}^* V_{cd} V_{cs}^*} \frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*} \frac{1 + \epsilon ae^{i\theta} e^{-i\gamma}}{1 + \epsilon ae^{i\theta} e^{+i\gamma}} \\ &= -e^{-2i\beta} \frac{1 + \epsilon ae^{i\theta} e^{-i\gamma}}{1 + \epsilon ae^{i\theta} e^{+i\gamma}}, \end{aligned} \quad (2.68)$$

where in the last equation the CKM angle β is used, as defined in Eq. (2.11). In the limit of neglected penguin contributions, $\epsilon = 0$, the parameter $\lambda_{[c\bar{c}]K_S^0}$ is a pure phase corresponding to 2β , so that the CP coefficients in the time-dependent decay rates are given by

$$S_{[c\bar{c}]K_S^0} = \sin(2\beta), \quad C_{[c\bar{c}]K_S^0} = 0, \quad D_{[c\bar{c}]K_S^0} = \cos(2\beta). \quad (2.69)$$

As $\Delta\Gamma = 0$ can be assumed at the current experimental precision, the time-dependent asymmetry in Eq. (2.48) simplifies substantially to

$$\mathcal{A}(t) = C_{[c\bar{c}]K_S^0} \cos(\Delta mt) + S_{[c\bar{c}]K_S^0} \sin(\Delta mt) \approx \sin(2\beta) \sin(\Delta mt), \quad (2.70)$$

where the last approximation is valid for the case of no direct CP violation.

2.5.3 Status in experiment and theory

Since the discovery of CP violation in b hadrons by the B factories [15, 16] in the $B^0 \rightarrow J/\psi K_S^0$ mode the precision on the angle $\sin(2\beta)$ has been further increased, both by analysing more data and reconstructing further final states. Besides the $J/\psi K_S^0$ final state, other final states with $b \rightarrow c\bar{c}s$ transitions have been used by the B factories, like $J/\psi K_L^0$, $J/\psi K^*(K^0\pi^0)$, $\psi(2S)K_S^0$, $\eta_c K_S^0$, $\chi_{c1} K_S^0$, and $\chi_{c0} K_S^0$, where it is taken into account that $J/\psi K_L^0$ and $\chi_{c1} K_S^0$ are CP -even, $J/\psi K^*(K^0\pi^0)$ is an admixture of CP -even and CP -odd parts, while all other are purely CP -odd [55–58]. Earlier measurements using the $J/\psi K_S^0$ final state from ALEPH [59], OPAL [60], and CDF [61], as well as a measurement from Belle using the $Y(5S)$ data [62] are combined with the recent LHCb measurement [19]. This combination results in the world average of the parameter S , provided by the Heavy Flavour Averaging Group [53],

$$S_{c\bar{c}s}^{\text{HF\text{LAV}}} = 0.691 \pm 0.017. \quad (2.71)$$

The main contributions to this world average are provided by the $B^0 \rightarrow J/\psi K_S^0$ measurements at the B factories, BaBar [55] and Belle [56], and the LHCb measurement [19], which resulted in

$$S_{J/\psi K_S^0}^{\text{Belle}} = 0.670 \pm 0.029 \text{ (stat.)} \pm 0.013 \text{ (syst.)}, \quad (2.72)$$

$$S_{J/\psi K_S^0}^{\text{BaBar}} = 0.657 \pm 0.036 \text{ (stat.)} \pm 0.012 \text{ (syst.)}, \quad (2.73)$$

$$S_{J/\psi K_S^0}^{\text{LHCb}} = 0.731 \pm 0.035 \text{ (stat.)} \pm 0.020 \text{ (syst.)}. \quad (2.74)$$

Here, LHCb only considered the $J/\psi \rightarrow \mu^+\mu^-$ final state, while BaBar and Belle additionally included a reconstruction of $J/\psi \rightarrow e^+e^-$. The $J/\psi K_L^0$ final state, which yields the second best sensitivity of all measurements, is not reconstructible at LHCb, as the K_L^0 decays outside the LHCb acceptance. The next best sensitivity is obtained from the $B^0 \rightarrow \psi(2S)K_S^0$ decays, which yield [55, 56]

$$S_{\psi(2S)K_S^0}^{\text{BaBar}} = 0.897 \pm 0.100 \text{ (stat.)} \pm 0.036 \text{ (syst.)}, \quad (2.75)$$

$$S_{\psi(2S)K_S^0}^{\text{Belle}} = 0.738 \pm 0.079 \text{ (stat.)} \pm 0.036 \text{ (syst.)}. \quad (2.76)$$

2 Flavour theory

The currently computed average on C , again provided by the Heavy Flavour Averaging Group, is given as [53]

$$C_{c\bar{c}s}^{\text{HF LAV}} = -0.004 \pm 0.015. \quad (2.77)$$

The individual contributions to this average provided by the B factories as well as the LHCb collaboration obtained from measurements in the $B^0 \rightarrow J/\psi K_S^0$ mode are [19, 55, 56]

$$C_{J/\psi K_S^0}^{\text{Belle}} = 0.015 \pm 0.021 \text{ (stat.) } {}^{+0.023}_{-0.045} \text{ (syst.)}, \quad (2.78)$$

$$C_{J/\psi K_S^0}^{\text{BaBar}} = 0.026 \pm 0.025 \text{ (stat.) } \pm 0.016 \text{ (syst.)}, \quad (2.79)$$

$$C_{J/\psi K_S^0}^{\text{LHCb}} = -0.038 \pm 0.032 \text{ (stat.) } \pm 0.005 \text{ (syst.)}. \quad (2.80)$$

The corresponding results in the $B^0 \rightarrow \psi(2S)K_S^0$ mode are [55, 56]

$$C_{\psi(2S)K_S^0}^{\text{BaBar}} = 0.089 \pm 0.076 \text{ (stat.) } \pm 0.020 \text{ (syst.)}, \quad (2.81)$$

$$C_{\psi(2S)K_S^0}^{\text{Belle}} = -0.104 \pm 0.055 \text{ (stat.) } {}^{+0.027}_{-0.047} \text{ (syst.)}. \quad (2.82)$$

Within the current experimental precision all results on C are compatible with zero. However, with increasingly precise measurements, the penguin contributions mentioned in Sec. 2.5.2 will need to be considered. These contributions can cause penguin shifts $\Delta\phi_d$ in the measured phase,

$$S = \sin(2\beta + \Delta\phi_d + \Delta\phi_{\text{NP}}), \quad (2.83)$$

which need to be distinguished from a phase difference caused by potential New Physics effects, $\Delta\phi_{\text{NP}}$. From a theoretical side penguin shifts are assessed using different formalisms: While some rely on U-spin symmetry [54, 63, 64] or estimate the effect using a full SU(3) analysis [65], others use an operator product expansion [66] to estimate the magnitude of the shifts. The latter approach yields $|\Delta\phi_d| < 0.68^\circ$ ($|\Delta S_f| < 0.0086$) for the $B^0 \rightarrow J/\psi K_S^0$ mode and $|\Delta\phi_d| < 0.74^\circ$ ($|\Delta S_f| < 0.0094$) for the $B^0 \rightarrow \psi(2S)K_S^0$ mode. Based on this, the penguin shifts will need to be considered as soon as the experimental precision of S reaches 0.01. Hence, the assumption of identical (negligible) penguin contributions when combining S measured in different final states is well motivated for the $J/\psi K_S^0$ and $\psi(2S)K_S^0$ states. Furthermore the LHCb measurement is still statistically limited, which motivates the inclusion of two obvious additional final states, $J/\psi(e^+e^-)K_S^0$ and $\psi(2S)K_S^0$.

The current value of S is perfectly consistent with other measurements in the flavour sector, as can be seen from indirect measurements performed by the CKMfitter group [42] and the UTfit collaboration [67] group. As no effects of physics beyond the SM have been observed in any direct measurements either, this motivates even more precise measurements, to look for tiny deviations and inconsistencies. The measurement of $\sin(2\beta)$ also serves as an important standard measurement, to prove LHCb's capability of performing tagged, time-dependent precision measurements in the harsh hadronic environment of the LHC.

3 The LHCb experiment

The Large Hadron Collider beauty experiment is one of four big experiments located at the Large Hadron Collider at CERN, the European Organization for Nuclear Research, near Geneva, Switzerland. The LHCb research programme is focused on indirect searches for physics beyond the SM in processes involving b and c hadrons. While the LHC can provide particle collision data of unprecedented size for the studies of rare decays and CP violation, the hadronic environment of pp collisions is particularly challenging, due to the high particle multiplicities in the interactions. The LHCb flavour physics programme requires the identification of final state particles and often also the measurement of particles' decay times. Hence, the detector is dedicated to provide an excellent particle identification and vertex reconstruction. A high-performance online and offline system reconstructs the data and distinguishes physically interesting signatures from backgrounds. Many parts rely on accurate simulations of the pp collision and the detector interaction of final state particles. The information provided in this chapter is based on Refs. [68, 69].

3.1 The LHC and its experiments

An overview of the CERN accelerator complex, consisting of the LHC and its preaccelerators, is shown in Fig. 3.1. The collider itself has a circumference of 26.7 km and lies in the tunnel formerly used by the Large Electron Positron Collider (LEP). It is located 50 to 175 m beneath the ground, at the France-Switzerland border near Geneva. Protons are preaccelerated up to an energy of 450 GeV until they are injected into the LHC, where they are further accelerated to their nominal energies. Each of the two LHC beams contains up to 2808 bunches of up to $1.15 \cdot 10^{11}$ protons. In individual beam pipes they circulate parallel and in opposite directions, so that a bunch spacing of 25 ns can lead to a collision rate of 40 MHz and a design luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ at four interaction points, where the beams are crossed and the experiments are located. The magnetic field strength necessary to keep the beams on track is above 8 T. It is produced by 1232 NbTi superconducting dipole magnets that are cooled by a large cryogenic system using liquid Helium. An advanced protection system against magnet quenches is necessary to avoid damages to the machine. Additional support magnets like quadrupoles prevent the beams from diverging.

During the first running period (Run I), lasting from March 2010 until early 2013, centre-of-mass energies $\sqrt{s} = 7 \text{ TeV}$ (2010/2011) and $\sqrt{s} = 8 \text{ TeV}$ (2012) were used at a bunch spacing of 50 ns. The analysis presented in this thesis is based on data recorded during this period. After an upgrade phase, the beam

3 The LHCb experiment

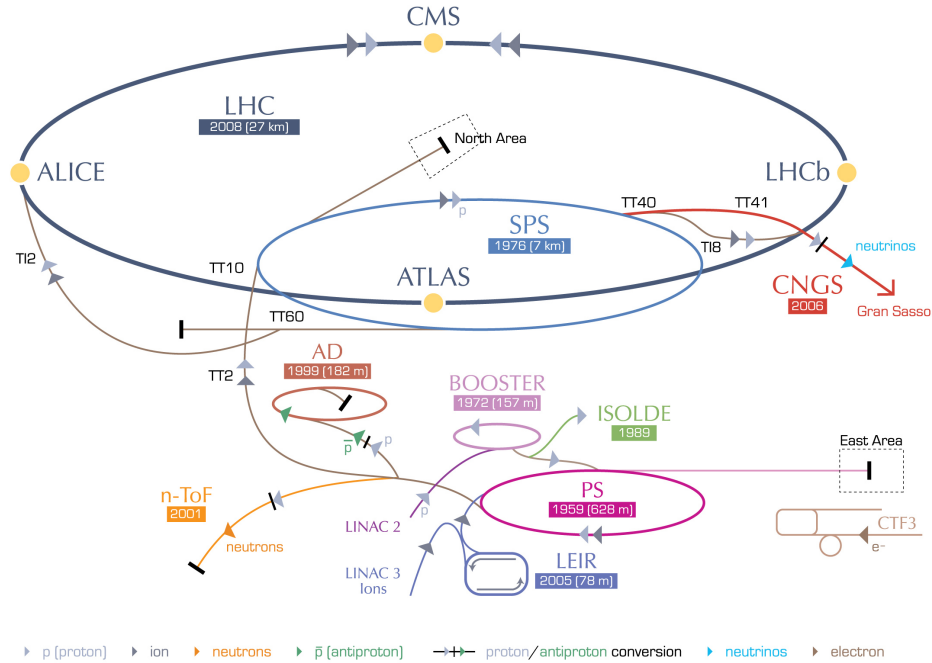


Figure 3.1 – Scheme of the CERN accelerator complex [70]. Multiple preaccelerators are passed by the protons until their nominal energy is reached: Linear Accelerator 2 (LINAC2), the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS). If the machine is fed with lead ions instead of protons, Linear Accelerator 3 (LINAC3) and the Low Energy Ion Ring (LEIR) are used instead of LINAC2 and BOOSTER.

energies were increased, so that during the second running period (Run II), which started in early 2015, the collider is operated at $\sqrt{s} = 13$ TeV.

Besides LHCb there are three other big experiments, ATLAS, CMS and ALICE, and three small experiments, TOTEM, MoEDAL and LHCf.

ATLAS (A Toroidal LHC Apparatus) [71] and CMS (Compact Muon Solenoid) [72] are general purpose-detectors (GPDs) and have become most popular for the discovery of the Higgs particle [33]. Their focus lies on the direct search for New Physics like hints for supersymmetry, dark matter candidates, and leptoquarks, but also more precise measurements involving the top quark and other particles like the W boson are performed.

ALICE (A Large Ion Collider Experiment) [73] is constructed with a focus on the research of lead-ion collisions. In these collisions of highly dense nuclei, physicists are creating a state that is high in temperature and density called quark-gluon plasma to better understand its properties and its phase transition from a confined, bound state.

TOTEM (TOTAl cross section, Elastic scattering and diffraction dissociation Measurement at the LHC) [74] is dedicated to the measurement of the proton-

proton interaction cross section, and also performs measurements to understand the proton structure, being installed at the CMS interaction point. MoEDAL (Monopole and Exotics Detector at the LHC) [75] shares an interaction point with LHCb and is constructed as a passive detector using nuclear track detectors, which are analyzed offline to search for magnetic monopoles and other highly ionizing stable massive particles. Until now MoEDAL did not find hints for magnetic monopoles or comparable exotic particles, but set limits on their production cross section. LHCf (Large Hadron Collider forward) [76] is the smallest of the LHC experiments and shares an interaction point with the ALICE experiment. It contributes to the understanding of high energy cosmic rays by analysing neutral pions in the extreme forward region.

3.2 The LHCb detector

The LHCb detector is built to maximize its efficiency to reconstruct processes of b and c hadrons, exploiting their characteristic production mechanism: These quarks are dominantly produced through gluon fusion, $gg \rightarrow q\bar{q}$, or quark-antiquark annihilation, $q\bar{q} \rightarrow q\bar{q}$, in pairs together with their antiquarks. The kinetic energy of the proton is shared among its partons, which interact in the proton collision, making it likely that two partons with very different momenta interact, and thereby cause a high boost of their centre-of-mass system along the beam axis. Accordingly, a characteristic forward correlation for $b\bar{b}$ pairs in terms of their production angles with respect to the beam pipe can be observed, which can be seen in Fig. 3.2. As also the formed hadrons are produced in this region, this presents the main motivation why the LHCb detector is built to cover the forward region of the pp interaction. The detector acceptance with a pseudo-rapidity range of $2 < \eta < 5$ covers 25% of the produced $b\bar{b}$ pairs.

A scheme of the LHCb detector design is given in Fig. 3.3. Comparing this design to that of GPDs like ATLAS or CMS, a striking difference is of course the forward orientation of the detector that covers an acceptance of ± 250 mrad vertically and ± 300 mrad horizontally. But also the used subdetectors differ from the ones of GPDs and are especially suited to the measurement of flight distances and the identification of final state particles: A retractable vertex locator (VELO) is able to resolve secondary decay vertices with high precision and supports the tracking system, while velocity information from two ring-imaging Cherenkov (RICH) detectors complements the particle identification information from the calorimeter and the muon systems. The detector components are described in the following sections in more detail.

3.2.1 Tracking system

The tracking system of the detector allows the measurement of charged particle trajectories, used to determine their origins in subsequent reconstructions. The warm dipole magnet possesses an integrated magnetic field of 4 Tm and causes

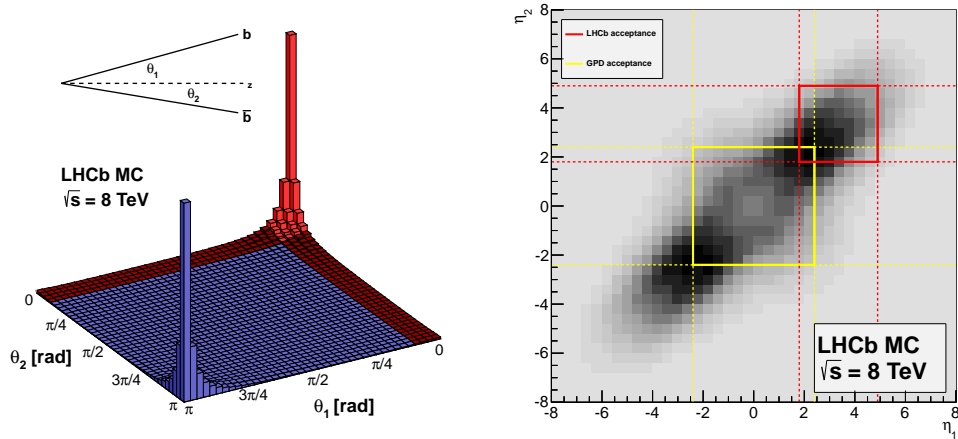


Figure 3.2 – The production cross section of b and \bar{b} quarks in terms of their angles with respect to the beam pipe $\theta_{1,2}$ (left) and their pseudo rapidity $\eta_{1,2}$ (right) from simulated events [77]. The red region in the left plot marks the LHCb acceptance, as does the red rectangle in the right plot. The yellow rectangle visualizes the acceptance of GPDs like ATLAS and CMS.

a bend of these trajectories, enabling the momentum measurement of charged particles. In order to limit systematic effects, the polarity of the magnetic field can be reversed (MagUp vs. MagDown polarity). The tracking subdetectors consist of the VELO, surrounding the pp interaction region, the tracker turicensis (TT) right before the magnet, and the three tracking stations (T1-T3) right behind the magnet. While the VELO and the TT use silicon microstrip detectors, this technology is only used in the inner region of T1-T3, the inner tracker (IT), whereas straw-tubes are used in the outer region, the outer tracker (OT).

The momentum resolution of the tracking system ranges from $\Delta p/p = 0.5\%$ at low momentum to 1% at 200 GeV , while it can provide an impact parameter resolution of $(15 + 29/p_T[\text{GeV}])\ \mu\text{m}$. These result in B mass-resolutions of $\approx 8\text{ MeV}$ and decay-time resolutions of $\approx 45\text{ fs}$ for modes like $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi \phi$, where $J/\psi \rightarrow \mu\mu$.

Vertex locator

The weak decays of b and c hadrons are associated with comparably long lifetimes that cause decays at displaced vertices. The VELO is capable of resolving these vertices by measuring cylindrical track coordinates in the direct vicinity of the interaction region. It is built of 21 half-disc shaped silicon strip R - ϕ -modules of 42 mm -diameter, located inside a vacuum vessel, where each of the modules measures the azimuthal angle, ϕ , and the radial distance from the beam pipe, R , of traversing particles. The z position of the trajectory is obtained through the

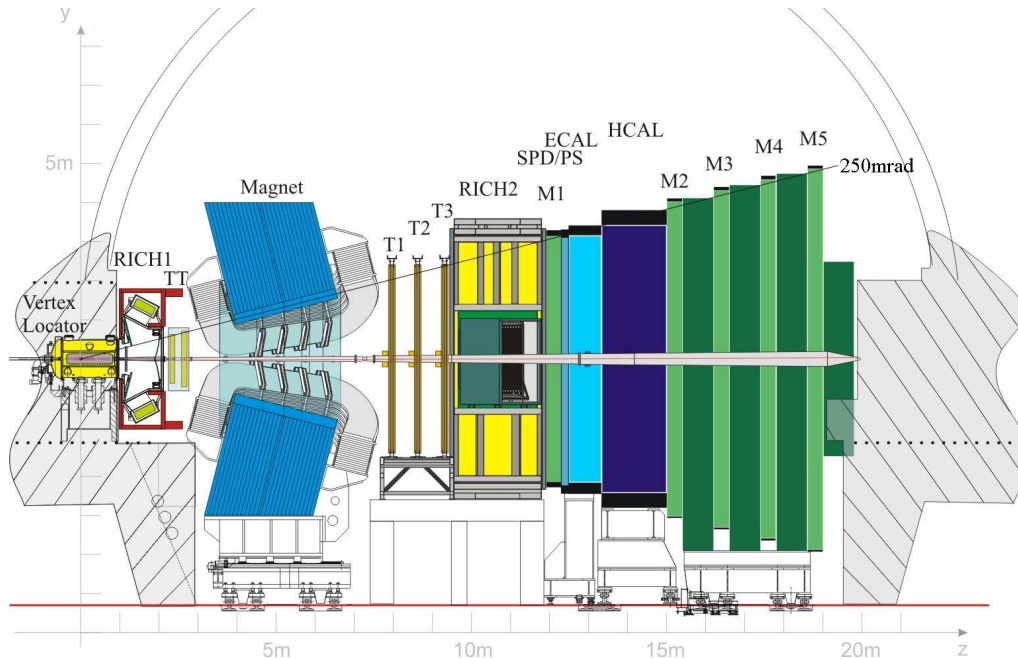


Figure 3.3 – Schematic view of the LHCb detector [68]. From left to right the visible components are the Vertex Locator (VELO), the first ring-imaging Cherenkov detector (RICH1), the tracker turicensis (TT), the magnet, the three tracking stations (T1-T3), the second RICH (RICH2), the first muon chamber (M1), the scintillator pad and preshower detectors (SPD/PS), the electromagnetic calorimeter (ECAL), the hadron calorimeter (HCAL) and the last four muon chambers (M2-M5).

position of the module along the beam pipe. The strip pitch of the individual modules ranges from $38\ \mu\text{m}$ to $102\ \mu\text{m}$. During stable running conditions the VELO discs are as close as $8\ \text{mm}$ to the interaction region. As the aperture of the beam is larger during the injection phase, the discs can be retracted by $3\ \text{cm}$ to move into the shadow of the $54\ \text{mm}$ diameter LHCb beam pipe.

Tracker turicensis and tracking stations

The remaining part of the tracking system can be divided into the silicon trackers, which consists of the TT and the IT, and the OT, which is a drift-gas detector. Their detection layers use an $(x-u-v-x)$ alignment, where the strips (for the TT and the IT) and tubes (for the OT) are vertically aligned with respect to the y -axis in the first and last layer, while they are rotated around the beam axis by $+5^\circ$ and -5° in the second and third layer, respectively. Through this tilt of two layers, sensitivity to the y -coordinate of the hit is obtained.

Being installed in regions of high particle multiplicities, TT and IT benefit from the higher spatial resolution of the silicon strip technology, using strip pitches

of $\approx 200 \mu\text{m}$. While the IT is located inside T1-T3 directly behind the magnet, the TT is installed right in front of the magnet. This way also low-momentum particles, which are bent out of the detector acceptance by the magnet, can be measured.

The OT forms the outer part of the three tracking stations, where particle multiplicities decrease to a level, at which a lower spatial resolution can be tolerated. It is built of about 55000 straw-tubes, filled with a mixture of 70% Argon, 28.5% CO₂ and 1.5% O₂, able to provide fast drift times of $<50 \text{ ns}$.

3.2.2 Particle identification system

A well performing identification of final state particles is essential to the analysis of flavour physics processes, as many rare processes could be confused with processes of similar topologies and therefore impede the measurement. The particle identification (PID) system of LHCb consists of the two RICH detectors, the calorimeter system, and the muon system. The combined PID information allows to assign likelihoods for particle hypotheses.

Ring-imaging Cherenkov detectors

The RICH detectors exploit the Cherenkov effect, i.e. the emittance of light by a particle, when it traverses a medium with a velocity higher than the speed of light in the medium. The opening angle under which photons are emitted, θ , can be used to measure the particles velocity, v , using the refraction index, n , through the relation $\cos \theta = c/(nv)$, where c denotes the speed of light in vacuum.

RICH1 uses a mixture of aerogel and C₄F₁₀ to cover a momentum range of $\approx 1\text{--}60 \text{ GeV}/c$. RICH2 employs a CF₄ radiator and covers $\approx 15\text{--}200 \text{ GeV}/c$ and beyond. Flat and spherical mirrors are used in both detectors to reflect photons out of the detector acceptance and focus their image on hybrid photo detectors.

Calorimeter system

The calorimeter system of the detector is mainly used for trigger decisions based on the presence of electrons, photons or hadrons. Furthermore, it provides energy measurements of particles and serves as the only way to detect neutral particles. The system consists of the scintillator pad detector (SPD), the preshower detector (PS), the electromagnetic calorimeter (ECAL), and the hadron calorimeter (HCAL).

The SPD allows to distinguish electrons from photons, because the former cause showers, while the latter do not. Following, a 2.5 mm lead layer separates the PS, which is used for a better separation of electrons from charged hadrons. The next component, the ECAL, is based on the so called shashlik calorimeter technology, where layers of lead, which cause particles to shower, are alternated with layers of scintillator material for the energy measurement. The last component, the HCAL, uses an analogue technique where iron plates are used to create the showers.

Muon chambers

Muons are present in the final state of many flavour physics processes of interest. Their ability to penetrate even thick layers of iron or steel makes their experimental signature very clean, and is also the reason why the muon system is placed in the back of the detector. The system can provide fast information for the trigger system as well as muon identification and momentum measurements for the offline analysis. It is composed of five stations, the first, M1, being placed in front of the SPD, while the following four, M2-M5, are placed behind the HCAL. This placement helps to measure momenta more accurately, as interactions with the calorimeter system disturb the measurement in the last stations. The stations consist of multi-wire-proportional chambers (MWPCs) except for an inner region of M1 where triple-gas-electron multipliers (TripleGEMs) are used, because of the better ageing properties in this area of higher rates.

3.3 The LHCb trigger

Only a tiny fraction of all bunch crossings contain pp interactions that show processes which are of interest to physicists. The task of the trigger system is to filter the events online so that the bunch collision rate of 40 MHz can be reduced to a level that can be stored. A step-wise system consisting of a hardware (level zero) trigger and two software (high level) triggers lowers the rate to 5 kHz.

3.3.1 Level-zero trigger

The level-zero (L0) trigger is a hardware trigger that works synchronously to the bunch crossing rate. Its decisions are based on detector systems that offer fast responses at these high rates: the calorimeter and the muon system. The highest transverse energy, E_T , photon candidate (L0Photon), electron candidate (L0Electron), and hadron candidate (L0Hadron) are searched for, using information provided by the ECAL and HCAL in association with that from the PS and SPD. Muon candidates with high transverse momenta, p_T , are searched for in the muon system. For candidates to pass the selection, a minimum threshold for the p_T of one muon in the event (L0Muon) or for the product of the p_T of two muons (L0DiMuon) is required. In addition to these specific candidate triggers, a part of the VELO system is used to identify and discard collisions that exceed a threshold of the primary-vertex multiplicity, as such events are disadvantageous to analyses involving b hadron decays. Given the reduced rate of 1 MHz after the L0 selection, the next trigger stages can use more complex and time-consuming algorithms for more accurate reconstructions of the events.

3.3.2 High level trigger

The high level triggers consist of a software application whose algorithms run on a computing farm, the event filter farm (EFF). They work in two stages: high level trigger 1 (HLT1) confirms decisions made by the L0 trigger, and lowers the data rate to 10 kHz and the subsequent HLT2 further reduces the rate to 5 kHz which can be recorded. The requirements that are made in the software triggers are organized in so called trigger lines whose configuration can be adjusted through trigger configuration keys (TCKs). The latter allow the use of different thresholds for the requirements or to prescale the decision of certain lines.

HLT1 can already use information from more subdetectors than L0. It uses information from the VELO and the tracking stations to reconstruct the event partially and confirm the presence of high p_T candidates found at the L0 stage. For this, a tracking algorithm based on the extrapolation of VELO tracks is used, also allowing for the determination of track impact parameters with respect to the primary vertices (PVs). The latter are of particular interest to select tracks from the displaced vertices of b and c hadron decays.

The input rate of HLT2 is sufficiently low, so that a full reconstruction of the events triggered by HLT1 can be performed. The selections are both, exclusive and inclusive in nature. While exclusive selections aim for the complete reconstruction of interesting b and c final states, inclusive selections try to select a broad spectrum of final states with the occurrence of resonances that are common to the decays of interest. For the latter a partial reconstruction of b hadron decays is often performed in topological lines, designed to look for B decays with at least two charged daughters [78]. The strategy here is to start from 2-body proto-candidates, which are built from particles associated with the same primary vertex and which have a small distance of closest approach. These candidates are passed to a boosted decision tree (BDT) selection or, using further tracks and corresponding filters, they are combined to 3- or 4-body proto-candidates which are then passed to a BDT selection (see Sec. 5.1.3 for details on BDTs). The majority of the events containing $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays can only be triggered through these multibody topological trigger lines (see. Sec. 5.1.8). More details on the configuration and the performance of the LHCb trigger can be found in Refs. [79–81].

After a signal candidate has been selected from the event, it is possible to associate the trigger response of the individual L0 triggers and HLT lines to the candidate. If the signal candidate alone can cause a positive trigger decision, the event is referred to as triggered on signal (TOS). If the part of the event excluding the signal candidate can cause a positive decision, it is referred to as triggered independent of signal (TIS). It is also possible that both the signal and the rest of the event are required to cause a positive decision (TOB).

3.4 Software

The LHCb software is based on the GAUDI [82, 83] framework, providing interfaces to all other software applications and services specialized to satisfy the demands of high energy physics. The HLT application running on the EFF is MOORE [84]. It is also used offline to emulate the trigger decision for simulated events. The reconstruction of the event as performed by the BRUNEL [85] package, i.e. the track-fitting, the bremsstrahlung recovery, and the identification of particles, is described in Secs. 3.4.1 to 3.4.3. For the selection of candidates from tracks and the fit of decay chains, DAVINCI [86] is used. Details on the production of simulated samples using GAUSS [87, 88] and BOOLE [89] are given in Sec. 3.4.4.

3.4.1 Track reconstruction

The reconstruction of tracks is achieved by different algorithms that use information from the tracking system. Particles that traverse the complete tracking system can be reconstructed by extrapolating VELO tracks to hits in the T stations (forward tracking) or by matching tracks in the T stations (T tracks) to VELO tracks (track matching), while taking the magnetic field into account. Matching hits in the TT are added to the particle track and the resulting tracks are referred to as long tracks (L tracks), which show the best resolution of all track types that are reconstructed. Another important tracking algorithm finds tracks in the T stations and directly matches these to hits in the TT, resulting in so called downstream tracks (D tracks) of worse resolution, caused when short-lived particles decay outside the VELO, like the K_S^0 . Upstream tracks (U tracks) are tracks that are reconstructed from hits in the VELO and the TT, which can originate from particles of low momentum, that are bend out of the detector acceptance. After a Clone Killer algorithm [90] removes duplicate tracks, the trajectory and the momentum of the track is determined using a fit that takes interactions with the detector material (like multiple scattering and energy losses) into account.

3.4.2 Bremsstrahlung recovery

When charged particles are accelerated perpendicular to their direction of movement, e.g. through a magnetic field, or parallel to their direction of movement, e.g. deceleration through interaction with material that they traverse, they radiate energy through emission of photons. The effect is referred to as bremsstrahlung and the radiated power is found to be proportional to $1/m^4$ and $1/m^6$ for parallel and perpendicular accelerations, respectively, where m is the mass of the particle. Comparing the electron and the muon mass with each other, $m_e/m_\mu \approx 207$, electrons are found to be affected 10^9 or 10^{14} as strong as muons by these energy losses, depending on the direction of the acceleration. Furthermore, for electrons which exceed a few tens of MeV it is the dominant effect by which they lose energy [31]. The radiation length X_0 is a characteristic property of the material,

3 The LHCb experiment

and corresponds to the length that a charged particle traverses the given material until it loses all but $1/e$ of its energy through radiation. In order to relate the detector geometry and material to the loss that charged particles experience in it, the material budget is commonly expressed in fractions of X_0 .

It is found that the most significant energy losses due to bremsstrahlung occur in the detector region in front of and behind the magnet, whereas losses inside the magnet through synchrotron radiation in the detector are at the level of 0.1% [91]. Fig. 3.4 shows results of simulation studies on the origin of bremsstrahlung photons and the material budget of the LHCb detector integrated until the magnet. Photons that are emitted after the magnet cannot be measured separately, as they share calorimeter cells with the particle itself. However, they are also unproblematic as they do not have a significant impact on the flight direction. In contrast, bremsstrahlung that occurs before the magnet is important to consider, as it reduces the momentum before the track's curvature is measured to determine the momentum, thereby reducing its measured momentum.

The approach to correct this loss in momentum is to search for photon clusters in a calorimeter region that is linearly extrapolated from the electron track at its origin and at the TT, as illustrated in Fig. 3.5. Under the assumption that the photon originates from the primary vertex and its momentum points to the 3D barycentre of the calorimeter cluster, the photon 4-momentum is added to the electron candidates 4-momentum [92].

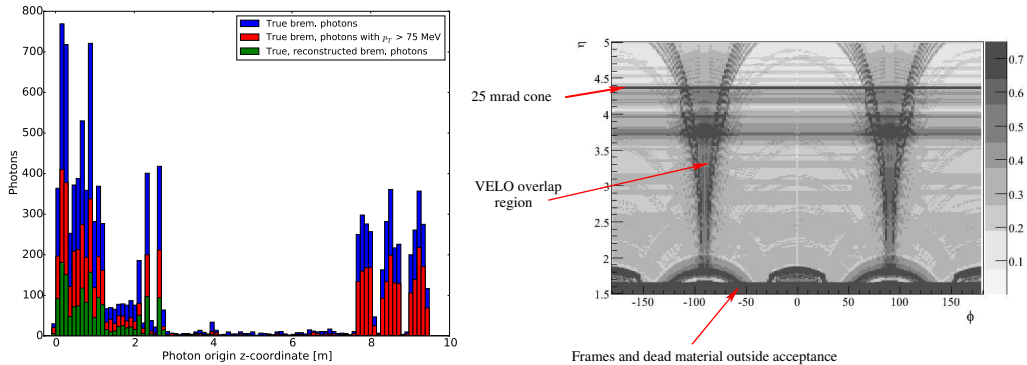


Figure 3.4 – Studies on the energy loss of electrons inside LHCb. Left: Emitted bremsstrahlung photons (blue), bremsstrahlung photons with $p_T > 75$ MeV (red) and reconstructed bremsstrahlung photons in the ECAL (green) with respect to their origin z -coordinate along the beam axis, based on simulations from Ref. [91]. Right: Fraction of the radiation length X_0 (truncated) with respect to the pseudo rapidity η and the polar angle ϕ , computed by tracing test particles until the magnet from Ref. [93]. The red arrows mark regions of comparably high material budget, the 25 mrad beam pipe cone in RICH1, the VELO overlap region, and material outside the detector acceptance.

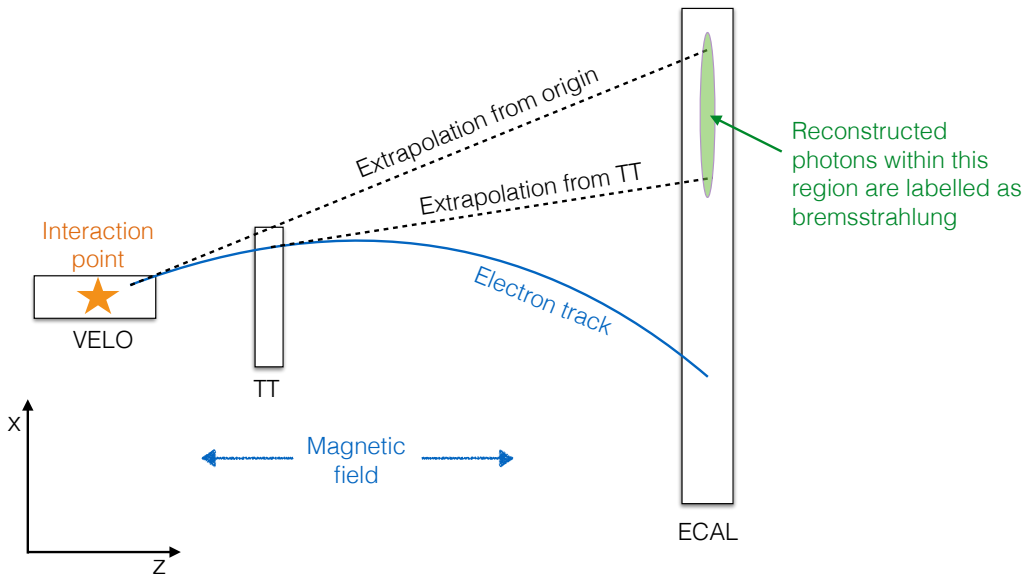


Figure 3.5 – Scheme of the bremsstrahlung correction, as implemented in the `BremAdder` algorithm from Ref. [94]. The electron track is linearly extrapolated from the origin and from the TT. The intersections of these extrapolations with the ECAL mark the zone, that is searched for reconstructed photons, which are then labelled as bremsstrahlung.

3.4.3 Particle identification

The identification of particles is performed by the subdetectors of the PID systems: the calorimeters, the RICH detectors, and the muon chambers.

The calorimeter system mainly identifies photons, electrons and π^0 candidates. Charged particles are distinguished from neutral particles, by searching for tracks in front of the calorimeter clusters. For charged particles the electron hypothesis is used and tested using the information from the ECAL, the PS, and the HCAL. For each of these systems, the likelihoods $\mathcal{L}(e - h)$ are constructed and summed to obtain the combined total likelihood of the calorimeter system, $\mathcal{L}^{\text{CALO}}(e - h)$. Photons and π^0 decaying to pairs of photons, are mainly distinguished by their cluster shape. The π^0 candidates can be reconstructed as resolved or merged, depending on whether it is possible to reconstruct the two photons from independent calorimeter clusters or not, respectively. The conversion of photons into electron-positron pairs through interaction with the detector material is taken into account during their identification.

The RICH systems serve mainly to distinguish the charged hadrons (π , K , p) from each other, but also contribute to the identification of the charged leptons (e and μ). The systems operate at extremely high occupancies, which is why the reconstruction is performed simultaneously in RICH1 and RICH2 under consideration of all tracks in the event, using an event log-likelihood algorithm [95, 96]. In the beginning

the pion hypothesis is assumed for all tracks. Then, sequentially the hypothesis is interchanged for each track individually, so that the event likelihood is minimized. The output of the procedure for each of the tracks after convergence are the differences in the log-likelihood, $\Delta\mathcal{L}$, that occur, when replacing the pion with each of the other particle hypotheses.

The reconstruction of the muon chambers' response aims to distinguish muon candidates from other charged particles that penetrate the detector. For this, in a first stage a binary decision is met, based on the presence of hits in a defined field of interest within the muon chambers. Both, the dimensions of the search field, as well as the number of chambers that show hits are dependent on the track momentum. In a second stage the identification of muon candidates is further improved by computing likelihood values for the muon and the non-muon hypothesis, based on the average squared distance significance of hits in the muon chambers with respect to the track's extrapolation from the tracking stations [97].

Combined likelihood values are obtained by summing the likelihood values of the subsystems for the given hypothesis, resulting in combined likelihoods for the particle hypothesis X against the pion hypothesis, $\Delta \log \mathcal{L}_{\text{Comb}}(X - \pi)$. Another approach uses multivariate techniques, artificial neural nets (ANNs), to obtain single particle probabilities, ($\text{ProbNN}\mu, \text{ProbNN}\pi, \dots$), while considering potential correlations between the responses of the PID subsystems [69]. All these measures are referred to as the PID of a particle within this thesis.

3.4.4 Simulation

Many parts of high energy physics experiments rely on the use of simulated datasets. The analysis of such datasets is already essential in the early stages of an experiment, e.g. to make detector design choices, but stays important later on, allowing to verify analysis steps and algorithms, as well as to estimate effects that are not or at least much harder accessible on real data. An accurate simulation of a collision experiment as LHCb involves various simulation phases that require precise modelling, so that reconstructed quantities can be reproduced.

The application `GAUSS` [87, 88] serves as a steering software controlling an event generation phase in which the pp interactions are simulated, and a detector simulation phase, in which the interactions of particles with the detector material are simulated. Besides serving as an interface for these two phases, the application also provides the running conditions, like beam energies and crossing angles, to the underlying generation software. The generation of the pp interaction is simulated using `PYTHIA` [98, 99] in the configuration for nominal productions, but also other event generators [100] like, `SHERPA` [101, 102] or `HERWIG++` [103] can be used. For historical reasons the decay of heavy hadrons is simulated using `EVTGEN` [104], as it comprises many established models of CP violation and mixing, whose simulation is crucial to a flavour physics experiment. Radiative corrections of processes are generated through `PHOTOS` [105, 106]. The particle interaction with the detector is simulated using `GEANT4` [107, 108]. The digitalization of the energy depositions

of traversing particles, i.e. the detector response, is generated with `BOOLE` [89], which also emulates the `L0` decision. After this final step, the simulated data is reconstructed and processed as real data.

Simulated samples can be categorized based on the process that is generated. In Signal Monte Carlo (MC) samples the presence of a certain signal process is required in each event. Hence, they can be used to study the signal signature in the detector, e.g. to obtain parametrizations of observables or extract efficiencies. In inclusive MC samples the occurrence of intermediate particles is required, so that families of processes can be simulated, like for example $pp \rightarrow J/\psi X$. Minimum bias MC samples aim for a full simulation of the pp collision without any restrictions, although a minimum requirement, like e.g. a minimum momentum transfer or alike, needs to be defined. An exemplary application of the latter would be the exploration of the average particle production that is seen by the detector.

4 Analysis tools and ingredients

The general strategy of the analysis is to use samples of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ decays, containing information on the B^0 decay time and its initial flavour, for the determination of the CP asymmetry. The heart of this determination is a simultaneous weighted decay-time-dependent fit (CP fit) using a model derived from the theoretical decay rates defined in Eqs. (2.37) to (2.40), which is presented in detail in Sec. 7.1. Important ingredients that are needed for preparatory studies of the two modes are discussed in this chapter. They comprise the experimental determination of the initial flavour through flavour tagging (FT), which is introduced in Sec. 4.1. It is followed by a short introduction to the maximum likelihood formalism in Sec. 4.2, needed for the signal-unfolding technique outlined in Sec. 4.3, but also for the CP fit. The datasets and the strategy to filter (select) signal decays from the data are presented in Sec. 4.4.

4.1 Flavour tagging

In decay-time-dependent measurements of CP violation it is essential to use information on the initial flavour of the neutral b meson. This information is provided by various flavour-tagging algorithms, so called taggers. At LHCb these taggers can be categorized as shown in Fig. 4.1: algorithms which reconstruct particles from the flavour-specific decay of the partner b hadron are referred to as opposite-side (OS) taggers; such that use hadronization fragments from the signal b meson are referred to as same-side (SS) taggers.

Each tagger provides an estimate for the initial flavour (tag), d' . The tag corresponds to -1 , for a prediction of an initial \bar{B}^0 , $+1$ for a prediction of an initial B^0 , and 0 if the algorithm cannot identify the initial flavour, due to e.g. poor reconstruction properties. In addition to the tag, each algorithm provides an estimate, η , corresponding to the probability of a wrong tag. These mistag estimates are the output of multivariate classifiers, that use information on the kinematic properties of the tagging particles and the event. A value of $\eta = 0$ corresponds to a tag with no associated uncertainty, while $\eta = 0.5$ corresponds to random guessing and therefore $d' = 0$.

The tagging algorithms are developed and trained using samples of flavour-specific decays. However, as the performance of the taggers is mode-dependent, it is necessary to recalibrate the algorithms using modes that have similar decay and event topologies to the decays that are used for the CP measurement. Using such control modes, the estimated mistag rate η is calibrated, most commonly under the assumption of a linear dependence between η and the true mistag rate, ω . This

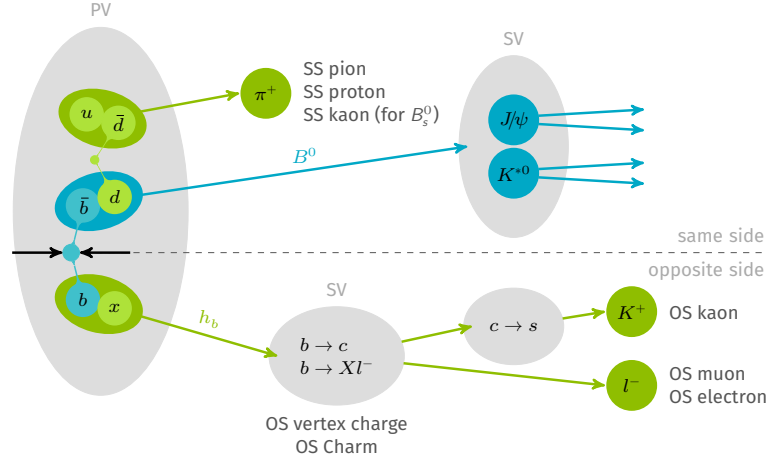


Figure 4.1 – Schematic overview of the underlying principles of LHCb’s flavour tagging algorithms from Ref. [109]. At LHCb taggers are divided into opposite side, where flavour-specific decays of the partner b hadron are exploited, and same side, where particles correlated to the hadronization process of the signal b meson are searched for.

way a parametrization, $\omega(\eta)$, can be determined and used to compute calibrated mistags for the decays of interest.

Wrong tags affect the measurement through a reduction of the CP asymmetry, that can be measured. The factor describing this asymmetry reduction compared to a scenario with perfect tags is referred to as dilution, $D = 1 - 2\omega$. The dilution causes an FT-induced loss in statistical power, when comparing the actual sample containing wrong tags to a (hypothetically) perfectly tagged sample. Thus, the statistical power of the FT is measured through the so called effective tagging efficiency, which can be evaluated on a per-event basis, as

$$\varepsilon_{\text{eff}} = \frac{1}{N} \sum_{i=1}^N (1 - 2\omega(\eta_i))^2, \quad (4.1)$$

where the index i sums over N events. Values on ε_{eff} can therefore be used to directly compare the FT performance in different modes and also between different experiments.

4.1.1 Opposite-side flavour tagging

The opposite-side FT algorithms [110] try to infer the initial flavour of the signal meson from the partner b hadron, which is produced in the $b\bar{b}$ -quark-pair production. This inference can be affected by errors through identifying wrong particles, i.e. particles that do not stem from the partner hadron, but also by neutral partner b mesons that oscillated to the opposite flavour until their decay.

Three taggers are used on the opposite-side, which base their decision on single particles. Two of them use semileptonic decays of the partner b hadron, the

OS electron and the OS muon tagger, and the third, the OS kaon tagger, identifies kaons that originate from decays following a $b \rightarrow c \rightarrow s$ chain.

Two OS taggers derive their decision from multiple particles in the event. The OS vertex charge tagger inspects the decay vertex of the partner b hadron and computes a momentum-weighted average charge of the originating tracks. The OS charm tagger [111] performs inclusive and exclusive reconstructions of decays of the partner b hadron that originate from $b \rightarrow c$ transitions.

In the context of this analysis all single OS taggers are combined into the OS combination tagger (see Sec. 4.1.3).

4.1.2 Same-side flavour tagging

The same-side FT algorithms look for particles correlated to the hadronization process of the signal b quark: In the case of a hadronization into a B^0 (\bar{B}^0) meson, the needed \bar{d} (d) quark originates from a $d\bar{d}$ -quark pair. The SS algorithms exploits the fact, that the left-over d (\bar{d}) quark from this pair can hadronize into a charged particle.

The two employed algorithms are the SS pion and the SS proton taggers [112], which look for a positively (negatively) charged pion and proton, respectively, which are associated with the initial production of B^0 (\bar{B}^0) mesons.

4.1.3 Flavour-tagging combination

After the single OS and SS taggers are calibrated individually using flavour specific control modes, they are combined, forming an OS combination tagger and an SS combination tagger. For this combination the tags, d'_i , and mistag estimates, η_i , of the single taggers indexed by i are used to calculate probabilities p_a (p_b) for a tag of -1 ($+1$) to be correct, assuming no correlations between the decisions, as

$$p_a = \prod_i \left(\frac{1 - d'_i}{2} + d'_i \cdot (1 - \eta_i) \right), \quad (4.2)$$

$$p_b = \prod_i \left(\frac{1 + d'_i}{2} - d'_i \cdot (1 - \eta_i) \right). \quad (4.3)$$

The combined tag, d' , equals $+1$ for $p_b > p_a$ and -1 otherwise. The associated mistag probability is then given as

$$\eta = 1 - \frac{\max(p_a, p_b)}{p_a + p_b}. \quad (4.4)$$

If the individual single taggers were perfectly calibrated and uncorrelated, there would be no need for further calibrations of the combined taggers. However, as correlations between the individual taggers are present, it is common to recalibrate OS and SS combinations individually again. These final calibrations, which are used in the CP fit, are presented in the corresponding sections for the two modes (see Secs. 5.5 and 6.7).

4.2 Maximum-likelihood formalism

The maximum-likelihood formalism provides a way to estimate parameters from observed data, through the fit of an assumed underlying model that depends on these parameters. The principle of the formalism lies in maximizing the likelihood \mathcal{L} for the set of model parameters $\vec{\theta}$, given a set of observed data \mathbf{X} . Here, $\mathbf{X} = [\vec{x}_1, \dots, \vec{x}_m]^T$ is an $m \times n$ matrix consisting of m observations, each comprising n observables. A model is needed that describes the probability densities to obtain a single observation \vec{x} , given a set of parameters $\vec{\theta}$. The corresponding probability density function, the PDF, is written as $\mathcal{P}(\vec{x}|\vec{\theta})$.

In general, there can be different species (classes) that need to be modelled, so that \mathcal{P} itself is composed of various PDFs describing the species, \mathcal{P}_j , each carrying a yield N_j . The resulting PDF is then given as the superposition

$$\mathcal{P}(\vec{x}|\vec{\theta}) = \frac{1}{\sum_j N_j} \sum_j N_j \mathcal{P}_j(\vec{x}|\vec{\theta}_j). \quad (4.5)$$

It may also be necessary or advantageous to find PDF descriptions for subsets (categories) of the data. The PDFs, as well as the parameters and yields, then obtain an additional index s , so that individual descriptions for the subsets can be used. Given statistical independence of the single observations, the extended maximum likelihood function is then defined as the probability product over all observations, including a Poisson-distribution term for all subsets, each containing n_s observations, as

$$\begin{aligned} \mathcal{L}(\vec{\theta}|\mathbf{X}) &= \prod_s \frac{e^{-\sum_j N_j^s}}{n_s!} \left(\sum_j N_j^s \right)^{n_s} \prod_i \frac{\sum_j N_j^s \mathcal{P}_j^s(\vec{x}_i|\vec{\theta}_j^s)}{\sum_j N_j^s} \\ &= \prod_s \frac{e^{-\sum_j N_j^s}}{n_s!} \prod_i \sum_j N_j^s \mathcal{P}_j^s(\vec{x}_i|\vec{\theta}_j^s). \end{aligned} \quad (4.6)$$

In the case of only one species that is obtained through weighting the data with weights, w_j , the likelihood function simplifies significantly to [113]

$$\mathcal{L}(\vec{\theta}|\mathbf{X}) = \prod_s \prod_i \left[\mathcal{P}^s(\vec{x}_i|\vec{\theta}^s) \right]^{w_i}. \quad (4.7)$$

The extended likelihood function in Eq. (4.6) is used in the fits to the reconstructed B^0 mass, which are performed to subtract backgrounds that remain in the dataset after the full selection, as described in the next section. In contrast, the weighted likelihood defined in Eq. (4.7) is used for the CP fit of this analysis.

For computational efficiency the logarithm of the likelihood function is considered and parameters are estimated by scanning for a minimum in the negative log-likelihood function. The parameter estimation is technically performed using the Minuit function minimization via the RooFit framework [114], which also provides

parameter uncertainties. To provide accurate uncertainties for the CP parameters, their uncertainties are obtained from the Minos algorithm, which analyses the function shape around the found minimum.

4.3 Background-subtraction technique

After a full selection of the candidates, the remaining combinatorial background components are reduced through statistical unfolding using the sPlot technique [115]. The idea behind this technique is to find a discriminating observable, x , that can be used to separate signal and background distributions through different PDFs. A popular choice for the discriminating observable is the reconstructed B^0 mass, which often has a distinct shape that depends on the source of the species. In order to determine these so-called sWeights, an extended maximum likelihood fit has to be performed, fixing all parameters to previously determined best-fit values, but leaving the component yields floating. From the result of this fit, sWeights that can be used to unfold the n -th species can be calculated for each candidate. Using its associated discriminating observable value, x_i , the weights are obtained as

$$w_n(x_i) = \frac{\sum_{j=1} V_{nj} \mathcal{P}_j(x_i)}{\sum_{k=1} N_k \mathcal{P}_k(x_i)}. \quad (4.8)$$

Here, the species are indexed by j , and k , and the covariances between the yields of the species, N_k , are contained in the covariance matrix, \mathbf{V} . To visualize background subtracted distributions, the sWeighted candidates are displayed in histograms, where the uncertainty on each bin is given as

$$\sigma_{\text{bin}} = \sqrt{\sum_{i \in \text{bin}} w_i^2}. \quad (4.9)$$

In order to obtain correct uncertainty estimates from sWeighted maximum likelihood fits (sFits), the individual weights in Eq. (4.7) are multiplied with the global correction factor $\sum_i w_i / (\sum_i w_i)^2$.

4.4 Sample preparation

The data samples of reconstructed $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ decays correspond to an integrated luminosity of 3 fb^{-1} , recorded at center-of-mass energies of $\sqrt{s} = 7 \text{ TeV}$ and $\sqrt{s} = 8 \text{ TeV}$, during the 2011 and 2012 running period of the LHCb experiment, respectively. Simulated samples have been generated using the default LHCb MC production conditions, i.e. using PYTHIA, EVTGEN and GEANT4, and using running conditions that resemble the 2011/2012 data-taking period. The design of the selection is based on simulated signal samples (signal MC samples) of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ decays, that each contain 4 million generated decays. Besides these samples, inclusive MC samples are used to study possible backgrounds, which are described in place.

4.4.1 Subsamples

Subsamples of the data can be defined based on discrete observables, like e.g. the type of track reconstruction or the type of trigger response that candidates show. Different splits have been considered in this analysis, but only two have been found to be of use at different stages of the analysis:

Track type While the J/ψ and the $\psi(2S)$ candidates are always built from two long tracks, the K_S^0 candidates are built from either two long (LL) or two downstream tracks (DD). During the selection of candidates, a split in the K_S^0 reconstruction type (DD vs. LL) is performed to develop tailored selections that take into account the different resolutions in the case of the $B^0 \rightarrow \psi(2S)K_S^0$ mode. In this mode the split is also considered in the reconstructed-mass fit to unfold the signal contribution. For the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode dedicated criteria are only applied in the first, cut-based stage of the selection, whereas no split is performed in the further selection and unfolding steps.

Year Production asymmetry values are computed separately for the 2011 and 2012 data-taking periods. As these values are used in the CP fit, the corresponding fit model is split with respect to the year of data-taking, sharing all but the production asymmetry parameters.

Bremsstrahlung and trigger Simulated samples have been checked to test, whether a split in the number of photons used for the bremsstrahlung correction, or in the type of the L0-trigger response are of any use for reconstructed $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ candidates. No benefits were found when performing the associated splits (see Sec. 5.1.1).

4.4.2 Decay-tree fit

The initial process to find signal candidates follows a traditional bottom-to-top, leaf-by-leaf fitting of the decay chain. This approach is referred to as `LokiVertexFitter`, which is the name of its implementation within the LHCb software framework. In terms of computation resources this approach is very efficient and allows to reduce combinatorics without a full build of the whole decay chain. This is the reason, for it being the standard approach that is used in the centralized reconstruction of candidates within the stripping selection.

In contrast, a simultaneous, more time-consuming fit, which is referred to as decay-tree fit (DTF) [116], offers the advantage to use correlations between particle candidates' properties, as well as constraints on their origin vertex and mass, for a more accurate computation of the observables. This analysis uses two versions of the DTF: The first constrains the B^0 candidate to originate from a PV, so that corresponding observables are appended by DTF^{PV} . The other constrains the J/ψ

($\psi(2S)$) and K_S^0 particles to their known masses and does not use a PV constraint. Its corresponding observables are therefore denoted $\text{DTF}_{J/\psi K_S^0}$ ($\text{DTF}_{\psi(2S) K_S^0}$).

Previous analyses found, that constraints on the daughter particles' masses lead to a biased measurement of the B^0 decay time. At the same time, a constraint to the PV leads to a more accurate decay-time measurement, which is why these two versions are used, and a simultaneous constraint of the daughter particles and the B^0 origin is avoided. These findings are supported by results in a bachelor thesis, supervised by the author [23].

As the reconstructed decay time and the reconstructed mass of B^0 mesons are frequently used throughout the thesis, the shorthand notations $t' \equiv t' \text{DTF}_{\text{PV}}$ and $m \equiv m \text{DTF}_{\{J/\psi, \psi(2S)\} K_S^0}$ are introduced. If there is any chance of ambiguity, “(DTF)” or “(LVF)” will be added to the observable name, to clarify whether the decay-tree fit or the `LokiVertexFitter` has been used.

4.4.3 Selection strategy

The selection of candidates is performed in a sequence of steps that start from the reconstructed quantities provided by the PID and tracking systems of the detector. During the reconstruction, calorimeter clusters and tracks are combined with the PID information to form ProtoParticles, which get assigned particle hypotheses in the stripping. This latter selection step also builds the full decay chain by combining final state particles to intermediate particles, i.e. the $K_S^0 \rightarrow \pi^+ \pi^-$ and $J/\psi \rightarrow e^+ e^-$ ($\psi(2S) \rightarrow \mu^+ \mu^-$) candidates, and lastly the B^0 candidate, successively applying requirements to filter the data, and thereby reducing the number of combinations that can be made. Reconstruction and stripping run in a centralized fashion, which is why the selection is still referred to as being carried out online. All vertex-fit related quantities in this step are obtained from the `LokiVertexFitter`.

After decay candidates have been built, a refined selection, referred to as off-line selection, is applied to further enhance the signal purity of the sample. In contrast to the online selection, where the goal is to retain candidates at high signal efficiencies while reducing the background to a manageable level, this offline selection aims for an optimization of the signal to background ratio in order to maximize the sensitivity of the measurement to the parameters of interest, i.e. the CP parameters. The first, cut-based stage of the offline selection consists of trigger requirements and DTF convergence criteria in the case of the $\psi(2S)$ mode, and a loose cut-based preselection in the case of the J/ψ mode.

A multivariate selection using BDTs is used to carry out the main reduction in combinatorial background for both modes (see Sec. 5.1.3). The exact requirement on the BDT response is chosen using a figure of merit (FOM) that is specifically dedicated to the study of CP violation (see Sec. 4.4.4).

Following the multivariate selection, veto requirements for possible peaking A_b^0 background contributions are applied and multiple candidates are removed through random candidate selections. For the $\psi(2S)$ mode, a dedicated handling of possible wrong PV associations is carried out.

The last step of the candidate selection consists of fits to the reconstructed-mass distributions of selected B^0 candidates, which are used to extract sWeights, as described in Sec. 4.3.

4.4.4 Figure of merit

The goal of the selection optimization is to maximize the sensitivity of the measurement to the parameters of interest, i.e. to minimize their expected uncertainty. Frequently, analyses use signal-significance-like FOMs (e.g. $N_S/\sqrt{N_S + N_B}$) to judge the statistical power of the sample. The approach in this analysis is different and considers both, dilutions coming from the flavour tagging and the decay-time resolution, as well as the fact that the sensitivity to interference CP violation itself depends on the reconstructed decay time, t' . The FOM is derived from the inverse variance of the CP parameter S and has been firstly presented in Ref. [117]. It has been slightly modified to use sWeights for the analysis presented in Ref. [118], resulting in

$$\text{FOM} \equiv Q_{\text{mod}} = \frac{\left(\sum_i w_i\right)^2}{\sum_i w_i^2} \bar{D}^{\text{mod}}, \quad (4.10)$$

where the w_i are the sWeights, and \bar{D}^{mod} is a dilution term, given as

$$\bar{D}^{\text{mod}} = \frac{1}{\sum_i w_i} \sum_i (1 - 2\omega_i)^2 e^{-(\Delta m \sigma_{t,i})^2} \cdot X_i \cdot w_i. \quad (4.11)$$

Its role is to average the dilution effects coming from the mistag probabilities ω_i and the decay-time-uncertainty estimates, $\sigma_{t,i}$ over the signal candidates, considering the mass difference in the B^0 - \bar{B}^0 system, Δm . Furthermore, it considers the decay-time-dependent sensitivity to S through

$$X_i = \left[\frac{\sin(\Delta m \cdot t'_i)}{1 + d_i(1 - 2\omega_i) \cdot S \cdot e^{-(\Delta m \sigma_{t,i})^2/2} \cdot \sin \Delta m t'_i} \right]^2. \quad (4.12)$$

In the scope of a bachelor thesis under supervision of the author (see Ref. [24]), the suitability of the FOM has been validated in pseudo-experiments (see Fig. 4.2). To understand what happens during the selection optimization, the FOM can be decomposed into the following components that can be inspected:

Effective tagging efficiency

$$\text{FOM}_{\varepsilon D^2} \equiv \frac{1}{\sum_i w_i} \sum_i (1 - 2\omega_i)^2 \cdot w_i \quad (4.13)$$

Effective decay-time-resolution efficiency

$$\text{FOM}_{\sigma_t} \equiv \frac{1}{\sum_i w_i} \sum_i e^{-(\Delta m \sigma_{t,i})^2} \cdot w_i \quad (4.14)$$

Effective signal size

$$\text{FOM}_{S_{\text{eff}}} \equiv \frac{\left(\sum_i w_i\right)^2}{\sum_i w_i^2} \quad (4.15)$$

Two other components are inspected that probe different scenarios. For the first, the effect of the weighted X -term is tested, which corresponds to a scenario of perfectly tagged and resolved B^0 mesons. For the second, the X -term is set constant, which is appropriate for measurements using B_s^0 mesons, as their high oscillation frequency ($\Delta m_s \gg \Delta m_d$) prevents X from affecting the optimization, due to its rapid changes with respect to the reconstructed decay-time. Still, $\Delta m = \Delta m_d$ is used in all evaluations, so that the latter can be interpreted as a scenario, in which the original FOM is insensitive to the reconstructed decay-time.

 X -term component (decay-time dependence)

$$\text{FOM}_X \equiv \text{FOM}_{S_{\text{eff}}} \cdot \underbrace{\frac{1}{\sum_i w_i} \sum_i X_i \cdot w_i}_{\text{“average } X\text{”}} \quad (4.16)$$

 X -term set constant

$$\text{FOM}_{B_s^0} \equiv \text{FOM}_{S_{\text{eff}}} \cdot \frac{1}{\sum_i w_i} \sum_i (1 - 2w_i)^2 e^{-(\Delta m \sigma_{t,i})^2} \cdot w_i \quad (4.17)$$

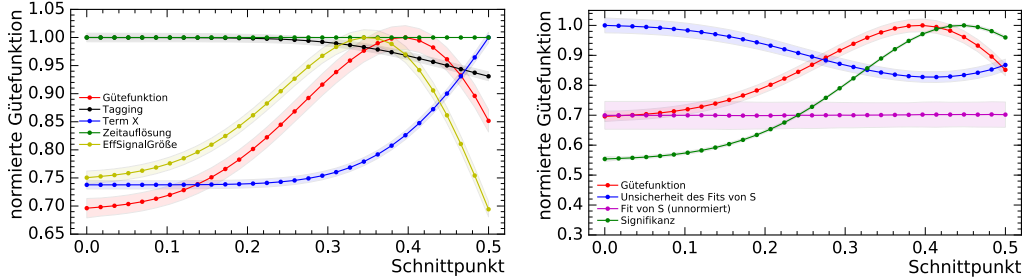


Figure 4.2 – Example FOM evaluations in studies of pseudo-datasets from Ref. [24]. All curves are normalized to their respective maximum value and plotted with respect to a requirement that is correlated to the decay-time and the distribution of mistags. Left: Effective signal size (yellow), effective decay-time-resolution efficiency (green) and effective tagging efficiency (black), average X (blue), overall FOM (red). Right: Fit value of the CP parameter S (purple), its uncertainty (blue), the signal significance (green), and the overall FOM (red). The colored bands correspond to one standard deviation of the distribution and not of the mean value. The studies show that the highest FOM value coincides with the smallest uncertainty for S .

5 $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ preparations

Preparatory studies for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decay mode are presented in this chapter. The selection of candidates including the selection optimization following the FOM is presented in detail. As an important intermediate result, the fit to the reconstructed B^0 mass is presented in Sec. 5.2. The part on the selection is followed by the results on the production asymmetry in Sec. 5.3 and the presentation of studies on mode-specific detector effects, i.e. the decay-time resolution and acceptance, in Sec. 5.4. The presented preparatory studies conclude with the presentation of the FT calibration in Sec. 5.5.

5.1 Selection

In general, the number of possible combinations that can be built from all tracks in the event is reduced by requiring that intermediate particle candidates form common vertices and are compatible with the corresponding meson masses and lifetimes, given in Tab. 5.1. Details on the selection criteria that are applied in the various selection steps are provided in the next sections. Signal efficiencies are estimated using the signal MC sample of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays, and are compared to the efficiencies from the previous $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ analysis [19] taken from Ref. [119].

Table 5.1 – Nominal values used for the masses and lifetimes of candidates in the selection of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ decays from Ref. [31]. Vertices for the decay of the J/ψ and the $\psi(2S)$ cannot be distinguished from the secondary B^0 decay vertex, due to their instantaneous decays. The flight distances of the B^0 meson and the K_S^0 are $\mathcal{O}(1 \text{ cm})$ and $\mathcal{O}(1 \text{ m})$, respectively.

Candidate	Mass [MeV/ c^2]	Lifetime [ps]	
J/ψ	3096.900 ± 0.006	$\mathcal{O}(10^{-9})$	($\Gamma = 92.9 \pm 2.8 \text{ keV}$)
$\psi(2S)$	3686.097 ± 0.025	$\mathcal{O}(10^{-9})$	($\Gamma = 296 \pm 8 \text{ keV}$)
K_S^0	497.611 ± 0.013	89.54 ± 0.04	
B^0	5279.63 ± 0.15	1.520 ± 0.004	

5.1.1 Stripping selection

The stripping selection to build $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ candidates has been developed by building upon the selection of K_S^0 candidates used in the previous analysis of

Table 5.2 – Requirements on the charged tracks as applied in the stripping. The used variables are the total momentum, p , the transverse momentum, p_T , the difference of logarithmic likelihood values of e and π hypotheses, $\Delta \ln \mathcal{L}_{e\pi}$, the trackfit’s χ^2 value, χ_{track}^2 , the distance of closest approach χ^2 , χ_{DOCA}^2 , and the smallest change in vertex χ^2 when including a track into the vertex fit with respect to any PV, MINIP χ^2 . Statements in parentheses refer to a different cut for downstream-reconstructed pions. The loose cut on the electron particle identification value, e ProbNNe, is applied in the processing of the datasets and will be moved to the stripping selection in the future.

Observable	Requirement
$e p_T$	$> 500 \text{ MeV}/c$
$e \Delta \ln \mathcal{L}_{e\pi}$	> 0
$e \chi_{\text{track}}^2$	< 5
$e^+ e^- \chi_{\text{DOCA}}^2$	< 30
πp	$> 2 \text{ GeV}/c$
π MINIP χ^2	> 9 (> 4)
$\pi^+ \pi^- \chi_{\text{DOCA}}^2$	< 25
e ProbNNe	> 0.01

Table 5.3 – Requirements on the J/ψ as applied in the stripping. The used variables are the reconstructed mass of the J/ψ , $m(e^+e^-)$, and the χ^2 value of its vertex fit, χ_{vtx}^2 .

Observable	Requirement
$m(e^+e^-)$	$2300 \text{ MeV}/c^2 < m(e^+ e^-) < 4000 \text{ MeV}/c^2$
χ_{vtx}^2	< 15

$B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ [19] and the selection of J/ψ candidates used in the analysis of $B_s^0 \rightarrow J/\psi(e^+e^-)\phi$, which has not yet been published, but which is being performed and reviewed in parallel to this analysis. The selection is performed through v21r1p1 and v21r0p1 of the `Bd2JpsiieKSFromTracksDetachedLine` for 2011 and 2012 data, respectively. Tab. 5.2 to 5.5 provide detailed listings of the particle candidates’ selection criteria applied in the stripping step.

In order to increase the chance that tracks originate from a common vertex, requirements on their distance of closest approach, DOCA, and the χ^2 of subsequent vertex fits, χ_{vtx}^2 , are made. A very loose requirement on the electron particle identification value, e ProbNNe, removes mis-identified tracks from kaons and pions at a high signal efficiency. Additionally, requirements on the total momentum of particles, p , and on the transverse momentum, p_T , increase the chance that these particles originate from the decay of a heavy hadron. Furthermore, requirements on the minimal impact parameter with respect to any PV, MINIP, aim to reduce π tracks that originate from PVs and not from decays of K_S^0 mesons. The significant K_S^0 decay time is also exploited through a cut on its decay-length significance, DLS. Asymmetric mass windows are used for the combined e^+e^- mass, as well as

Table 5.4 – Requirements on the K_S^0 as applied in the stripping. The used variables are the decay length significance, DLS, the χ^2 value of the vertex fit, χ_{vtx}^2 , and the reconstructed mass of the K_S^0 , $m(\pi^+\pi^-)$. Statements in parentheses refer to different cuts for DD reconstructed K_S^0 mesons.

Observable	Requirement
χ_{vtx}^2	< 20
DLS	> 5
$ m(\pi^+\pi^-) - m_{K_S^0} $	$< 50 \text{ MeV}/c^2$ ($< 80 \text{ MeV}/c^2$)

Table 5.5 – Requirements on the B^0 as applied in the stripping. The used variables are the reconstructed decay time, t' , the reconstructed mass, $m(e^+e^-\pi^+\pi^-)$, and the χ^2 value of the vertex fit, χ_{vtx}^2 .

Observable	Requirement
t' (LVF)	$> 0.2 \text{ ps}$
$m(e^+e^-\pi^+\pi^-)$	$4400 \text{ MeV}/c^2 < m(e^+e^-\pi^+\pi^-) < 6000 \text{ MeV}/c^2$
χ_{vtx}^2	< 10

the combined $e^+e^-\pi^+\pi^-$ mass, to account for energy losses, which cause a large tail to lower reconstructed masses. Finally, contributions from promptly produced J/ψ and K_S^0 mesons are suppressed by requiring a reconstructed B^0 decay time of $t' > 0.2 \text{ ps}$.

The signal efficiencies of the reconstruction and of the stripping selection are evaluated using the signal MC sample. For the reconstruction efficiency only successful reconstructions as long tracks in the case of the electrons, and as either two downstream tracks or two long tracks in the case of the pions are considered, as only these track-type combinations are used. The resulting reconstruction efficiency with respect to candidates inside the LHCb detector acceptance, corresponds to $\varepsilon_{\text{reco}}^{e^+e^-} \approx 10.9\%$. The stripping efficiency, with respect to the number of successfully reconstructed candidates, is obtained as $\varepsilon_{\text{strip}}^{e^+e^-} \approx 35.4\%$, which results in the combined efficiency with respect to candidates inside the LHCb acceptance of $\varepsilon_{\text{reco+strip}}^{e^+e^-} \approx 3.9\%$. Compared to the $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ analysis [19], where the combined efficiency for stripping and reconstruction was $\varepsilon_{\text{reco+strip}}^{\mu^+\mu^-} \approx 9.2\%$, the value is substantially lower. This relative decrease in performance can be traced to a loss in reconstruction efficiency, $\varepsilon_{\text{reco}}^{e^+e^-}/\varepsilon_{\text{reco}}^{\mu^+\mu^-} \approx 67\%$, as well as in stripping efficiency, $\varepsilon_{\text{strip}}^{e^+e^-}/\varepsilon_{\text{strip}}^{\mu^+\mu^-} \approx 62\%$. The loss in reconstruction efficiency for the electron mode can be further divided into a general loss in electron track reconstruction efficiency, $\varepsilon_{\text{track}}^{e^+e^-}/\varepsilon_{\text{track}}^{\mu^+\mu^-} \approx 77\%$, and a loss by requiring long track reconstructed electrons, $\varepsilon_{\text{long}}^{e^+e^-}/\varepsilon_{\text{long}}^{\mu^+\mu^-} \approx 86\%$. The latter can be explained by the higher energy loss of the electrons making them prone to get bent out of the detector acceptance

by the magnetic field before the T stations. This is also the reason why the number of events containing upstream reconstructed lepton tracks is enhanced by a factor of 3.5 compared to the dimuon mode.

Subsamples

The dependence of reconstructed quantities on discrete observables (see Sec. 4.4.1) has been studied in the stripping-selected signal MC sample, prior to further selections. Concerning the number of candidates in the track-type subsets, 27% (73%) are found to be reconstructed from long track (downstream track) pions. For the bremsstrahlung correction, a total of 30% / 49% / 21% of the candidates have been corrected using 0 / 1 / 2 or more photons, respectively. For the level zero trigger decision, 51% (22%) of candidates are TOS (TIS) exclusive and 27% are TIS and TOS with respect to any level zero trigger. Furthermore, 71% (18%) of candidates are TOS by the L0Electron (L0Hadron).

Comparisons between the reconstructed mass distributions of the intermediate particles in the studied subsets are shown in Fig. 5.1, where candidates are categorized into inclusive TIS and exclusive TOS, for convenience. Regarding the type of the pion track reconstruction, a worse mass resolution (higher width) is obtained for DD reconstructed K_S^0 compared to LL reconstructed K_S^0 . Obviously the reconstructed J/ψ mass of the candidates is unaffected from a split in the pion track type. Concerning the number of photons used in the bremsstrahlung correction, it can be seen that the reconstructed J/ψ mass has a large tail in the uncorrected subset of the sample, whose extension is clearly reduced for candidates in the corrected subsets. The reconstructed K_S^0 mass of the candidates is of course unaffected with respect to a split in the order of the bremsstrahlung correction. Lastly, the type of the level zero trigger response also has a slight influence on the reconstructed J/ψ mass, as the candidates in the exclusive TOS subset are of higher chance to be of better reconstruction quality, compared to candidates in the inclusive TIS subset.

From the corresponding comparisons of the reconstructed B^0 mass distributions in Fig. 5.2 it can be seen that the mass resolution (peak width) of B^0 candidates is largely dominated by the radiative tail caused by bremsstrahlung. When using the DTF with constraints for the masses of the J/ψ and K_S^0 meson the peak resolution of the reconstructed B^0 mass is much better and only slight differences between the subsamples can be observed. However, due to the large background contribution in data, these minor differences cannot be observed. Furthermore, no significant benefits in terms of the FOM (see Sec. 4.4.4) were found in studies including a split in the categories. For the preselection that is presented in the next section, it has been decided to use K_S^0 candidate requirements that can depend on the track-reconstruction type. For the multivariate selection the dataset is not split, since this does not yield any additional benefits. Finally, an averaged description of the presented subsamples is used in the fits to the reconstructed mass and decay-time of B^0 candidates.

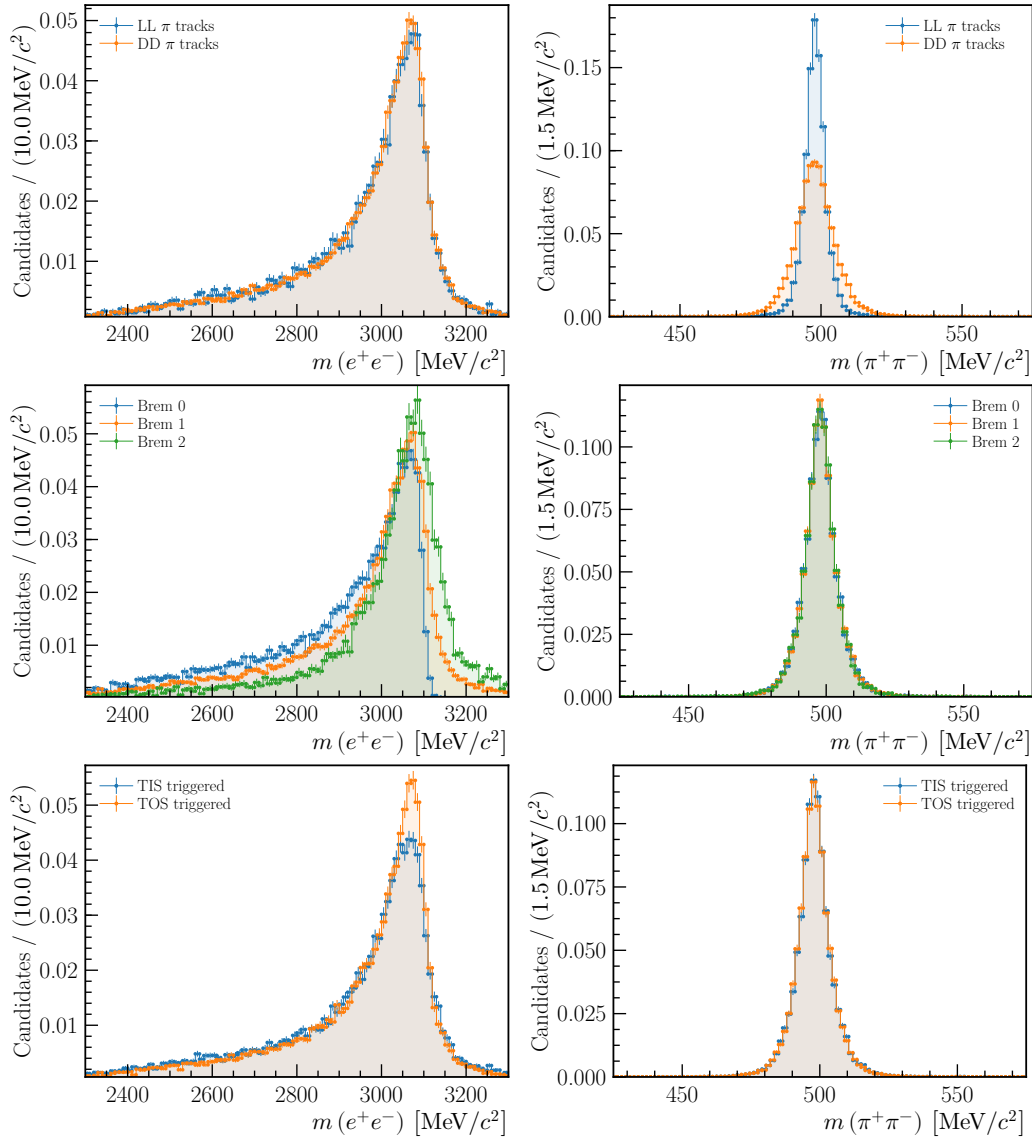


Figure 5.1 – Reconstructed invariant-mass distributions of stripping-selected J/ψ candidates (left) and K_S^0 candidates (right) in the signal MC sample. The data is split into candidates built from K_S^0 candidates reconstructed from either two long track pions or two downstream track pions (top), into candidates that have been corrected using zero, one or two and more photons (middle), and into candidates that are inclusively TIS or exclusively TOS (bottom). The total number of candidates is normalized to one for all distributions.

5 $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ preparations

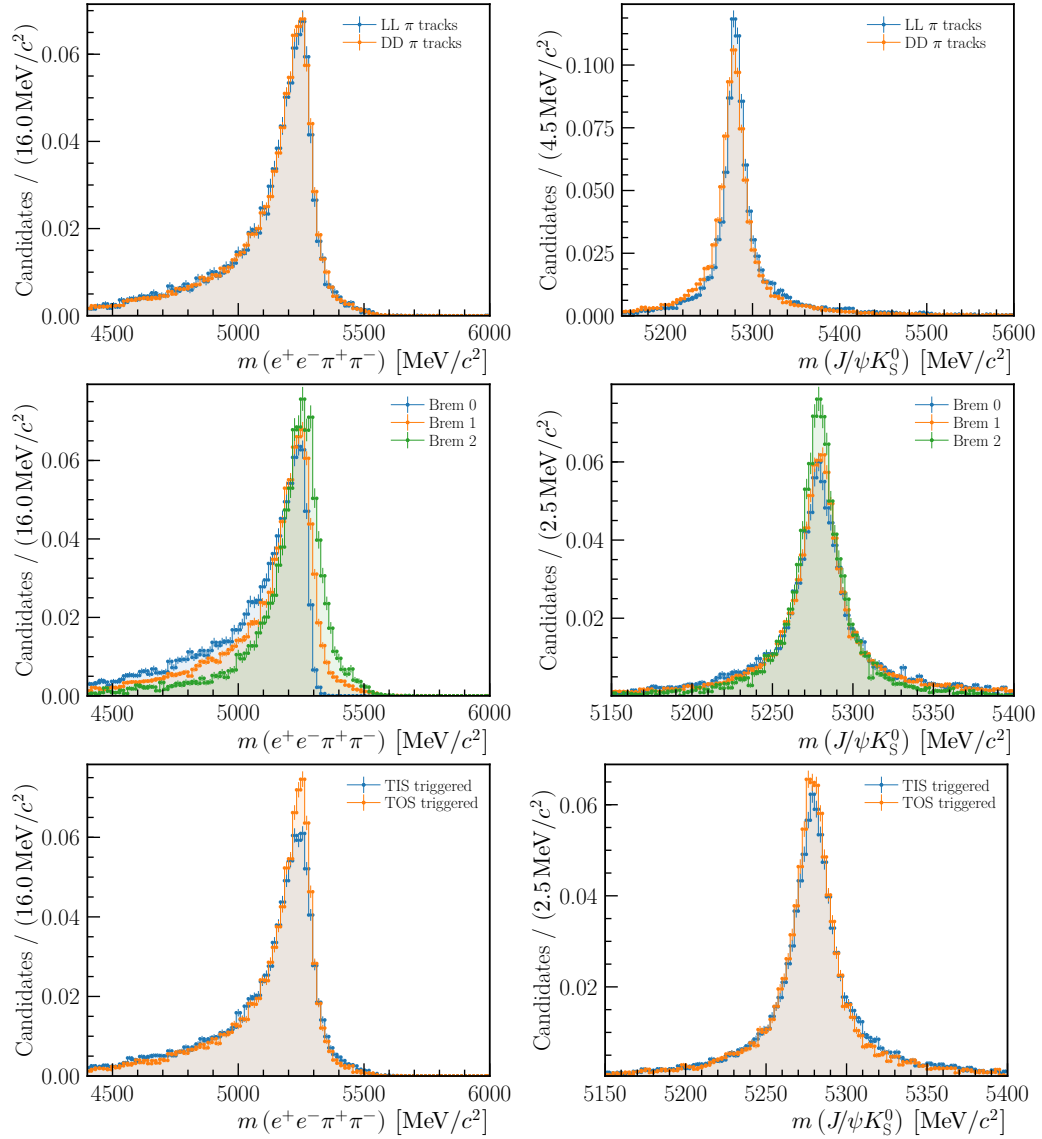


Figure 5.2 – Reconstructed invariant-mass distributions of stripping-selected B^0 candidates in the signal MC sample, without mass constraints for the intermediate particles (left), and with mass constraints for the intermediate particles (right). The data is split into candidates built from K_S^0 candidates reconstructed from either two long track pions or two downstream track pions (top), into candidates that have been corrected using zero, one or two and more photons (middle), and into candidates that are inclusively TIS or exclusively TOS (bottom). The total number of candidates is normalized to one for all distributions.

5.1.2 Loose preselection

The data after the stripping selection consists of approximately 50 million candidates, while $\mathcal{O}(10^4)$ signal decays can be expected. Fig. 5.3 shows the reconstructed mass distributions for the intermediate particles, i.e. the J/ψ and the K_S^0 meson candidates, as well as for the B^0 meson candidates with and without mass constraints for the intermediate particles. While a significant K_S^0 contribution is found among the candidates, the J/ψ candidates are still contaminated with a huge combinatorial background, resulting in no observable $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays after the stripping. The main background contribution at this stage cannot be accounted to specific decays, but is caused by high particle multiplicities. In contrast to the signal part of an event, such combinatorial backgrounds are poorly described in simulated samples. Therefore, parts of the data sample itself are used as a background sample in further selection steps, which allows to find observables that clearly separate the signal decays from other candidates. Requirements made in the preselection are used to further suppress obvious backgrounds and bring the data to a more manageable level at high signal efficiencies.

The stripped data is used to find observables that provide a high discriminating power compared to the distributions in simulated signal decays. Distributions

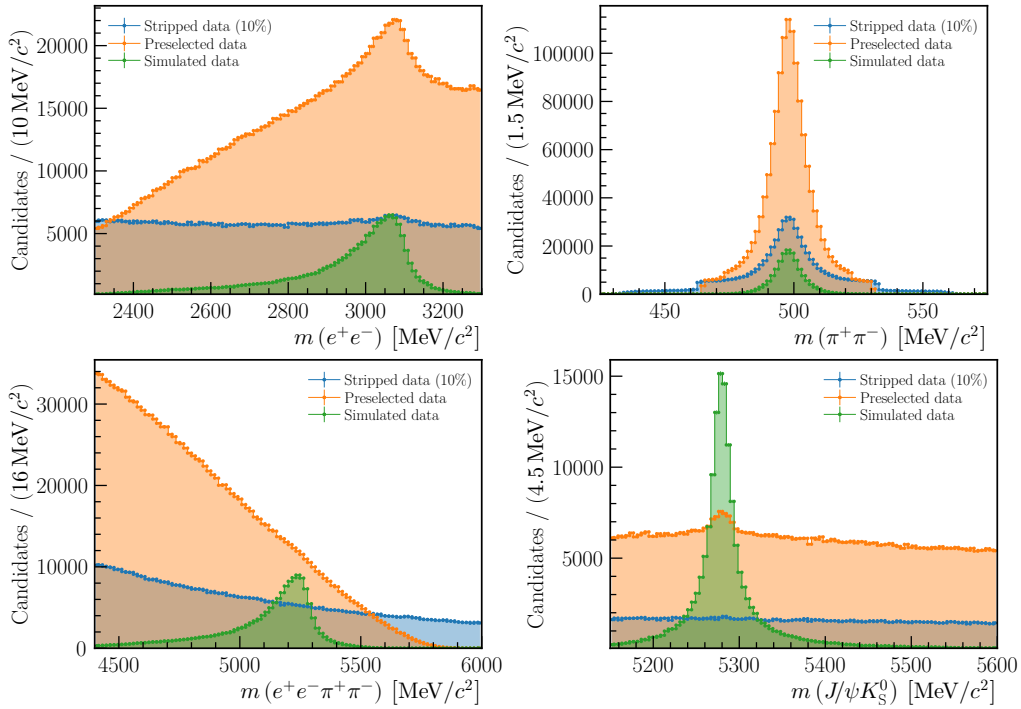


Figure 5.3 – Reconstructed invariant-mass distributions of intermediate particles and B^0 candidates in data after the stripping selection (blue), after the preselection (orange), and in simulated $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays after the stripping (green). Only 10 % of the data passing the stripping is shown, as $> 50 \cdot 10^6$ candidates pass this selection step.

5 $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ preparations

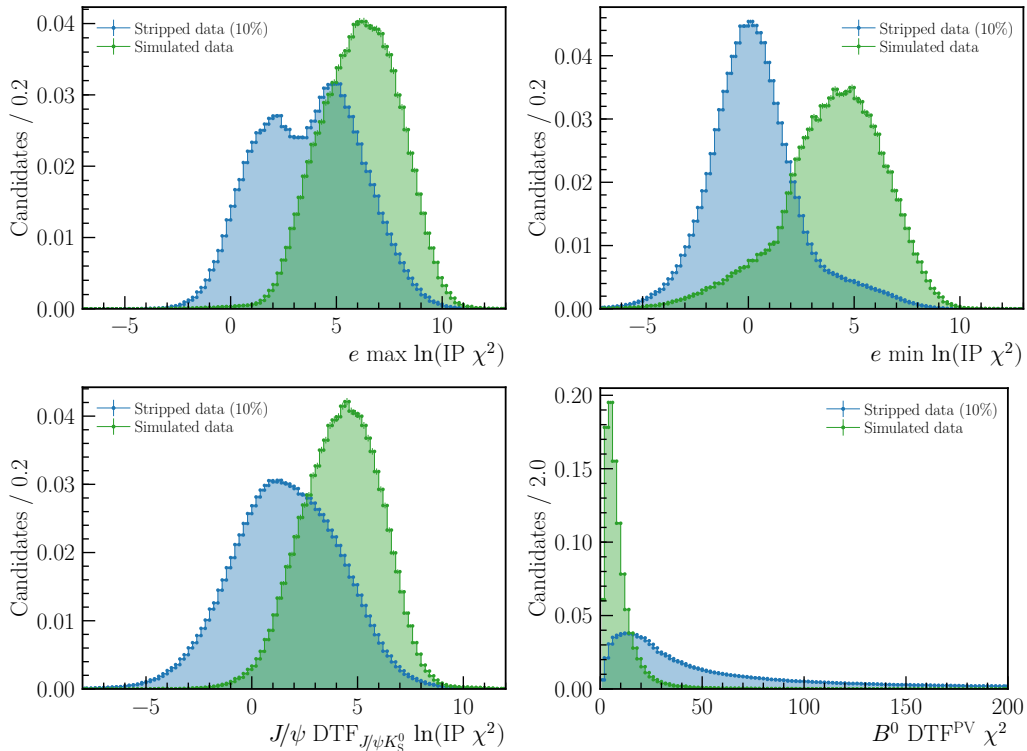


Figure 5.4 – Distributions of observables providing the most separation between signal decays in the stripped signal MC sample (green) and the stripped data (blue): the change in vertex χ^2 when including the electron / positron track into the PV fit (top left and right), the IP χ^2 of the J/ψ with respect to the PV (bottom left), and the χ^2 of the DTF with PV constraint (bottom right). Logarithms are used for convenience and a min / max operation is used to symmetrize observables for the tracks.

for the four most powerful observables, the minimum and maximum value of the two electron’s IP χ^2 , the J/ψ IP χ^2 , as well as the quality of the DTF with PV constraint, i.e. its overall χ^2 , are shown in Fig. 5.4. The full list of criteria is given in Tab. 5.6, where the single cuts have been tuned by hand to retain at least 97% of the signal decays in the simulated sample.

The overall signal efficiency of the loose preselection with respect to the number of candidates passing the stripping selection is estimated on the simulated sample to $\varepsilon_{\text{preselect}}^{e^+e^-} \approx 91\%$, while the number of candidates in data can be reduced by 83%.

5.1.3 Multivariate selection

Further selection of signal decays is carried out through boosted decision trees (BDTs). Decision trees are predictive models that use a set of input variables of an item (features) to predict its target class by applying binary decision rules. In the context of particle physics analyses the most common application is a binary

Table 5.6 – Preselection requirements applied to reconstructed $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ candidates. The used variables are the χ^2 value of the vertex fit, χ_{vtx}^2 , the logarithm of the change in χ_{vtx}^2 when including a particle into the vertex fit, $\ln(\text{IP } \chi^2)$, the K_S^0 flight distance χ^2 , FD χ^2 , the reconstructed K_S^0 mass, $m(\pi^+\pi^-)$, and the χ^2 of the decay-tree fit, $\chi^2 \text{ DTF}^{\text{PV}}$. Statements in parentheses correspond to different requirements for candidates with DD reconstructed K_S^0 candidates.

Observable	Requirement
DTF^{PV}	converged
$\text{DTF}_{J/\psi K_S^0}$	converged
$\pi \chi_{\text{track}}^2$	< 2.4
$\pi \max \ln(\text{IP } \chi^2)$	> 4 ($> -\infty$)
$e \max \ln(\text{IP } \chi^2)$	> 2.5
$e \min \ln(\text{IP } \chi^2)$	> -0.5
$J/\psi \text{ DTF}_{J/\psi K_S^0} \ln(\text{IP } \chi^2)$	> 0.5
$K_S^0 \text{ FD } \chi^2$	> 40
$K_S^0 m(\pi^+\pi^-) - m_{K_S^0} $	$< 25 \text{ MeV}/c^2$ ($< 33 \text{ MeV}/c^2$)
$B^0 \chi_{\text{vtx}}^2$	< 20
$B^0 \text{ DTF}^{\text{PV}} \chi^2$	< 120

classification of events into signal or background.

The rules of a single decision tree are derived through a supervised learning algorithm, which the class-labelled data is provided to. An individual tree is then built by partitioning the input items in each node to minimize the Gini index. This measure of impurity is defined as $G = 1 - \sum_i p_i^2$, where p_i is the probability that an item at a given node belongs to the class i .

In the case of *boosted* decision trees the ensemble is built using algorithms which iteratively learn individual decision trees. In each iteration another decision tree is added that helps to classify items which were misclassified by previous trees. The variety of existing boosting algorithms mainly differs in the method that is used to improve the model. For example, the GradientBoost method [120] is based on minimizing an error function that measures the classification error of previous trees. In contrast, the Adaboost algorithm [121] identifies misclassified items, and assigns weights that allow to consider these items more strongly and correctly classified items less strongly in the next iteration. The resulting BDT prediction is obtained as a weighted average of the individual decision trees and can be interpreted as the signal likelihood of the observation.

The multivariate selection of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays employs BDTs that are boosted using the Adaboost algorithm. The classifier is trained using simulated and preselected $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ signal events that serve as a signal proxy, and candidates from data whose reconstructed B^0 mass is far off the signal region ($m > 5600 \text{ MeV}/c^2$), which serve as a background sample. The full training data is

Table 5.7 – Normalized feature importance (Gini importance), averaged over the individual decision trees. The used features have been chosen in an iterative fashion, so that the importance of the least important feature is $> 1\%$.

Feature	Importance [%]	Feature	Importance [%]
$e \min \ln(\text{IP } \chi^2)$	26	$\pi \max \ln(\text{IP } \chi^2)$	5
$B^0 \text{ DTF}^{\text{PV}} \chi^2$	23	$K_S^0 \text{ DTF}_{J/\psi K_S^0} \text{ IP} \chi^2$	4
$B^0 \text{ DTF}_{J/\psi K_S^0} p_{\text{T}}$	12	$B^0 \chi_{\text{vtx}}^2$	3
$e \max \ln(\text{IP } \chi^2)$	10	$J/\psi \text{ DTF}_{J/\psi K_S^0} p_{\text{T}}$	3
$K_S^0 \text{ DTF}_{J/\psi K_S^0} p_{\text{T}}$	7	$J/\psi \text{ DTF}_{J/\psi K_S^0} \text{ IP} \chi^2$	2
$B^0 \text{ DTF}_{J/\psi K_S^0} \text{ IP} \chi^2$	6		

split into a training and a test sample. The first one is used to build the decision trees in the described way, while the latter is retained during the learning phase so that the built model can be validated on this unseen data. In particular, the test sample serves to check if the model has picked up and learned statistical fluctuations in the training sample, which is an undesired effect referred to as overtraining. For the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode the k -folding cross validation technique is used to make the prediction even more robust: The sample is split into six folds of equal size. For each fold a classifier is trained and validated on the five other folds. The final classifier is obtained as the average of all six individual classifiers.

The BDT features are determined in an iterative fashion, starting from a large selection of observables that provide the lowest overlap between the signal and the background proxy sample. After a training the features are ranked by their importance, which is computed through the overall decrease in Gini index from each individual feature. This way the feature list is truncated, so that observables which contribute only $< 1\%$ in terms of their importance are dropped. The final list of features ranked by importance is given in Tab. 5.7.

The performance of a BDT is usually judged by receiver-operating-characteristic (ROC) curves, where the true positive rate is plotted against the false positive rate. Hence, it corresponds to a plot of the signal efficiency against the background retention. Fig. 5.5 shows the classifier output distribution on the training and test samples for one of the folds and the ROC curve for all six folds. It can be seen that the background is expected to be suppressed to less than 2% when retaining 90% of the signal. The optimization of the working point on the ROC curve will be described in Sec. 5.1.6.

5.1.4 Exclusive backgrounds – $\Lambda_b^0 \rightarrow J/\psi(e^+e^-)\Lambda(p\pi^-)$ veto

Various background sources can contribute to the reconstructed and selected candidates. No peaking structures are observed in the reconstructed B^0 mass distribution

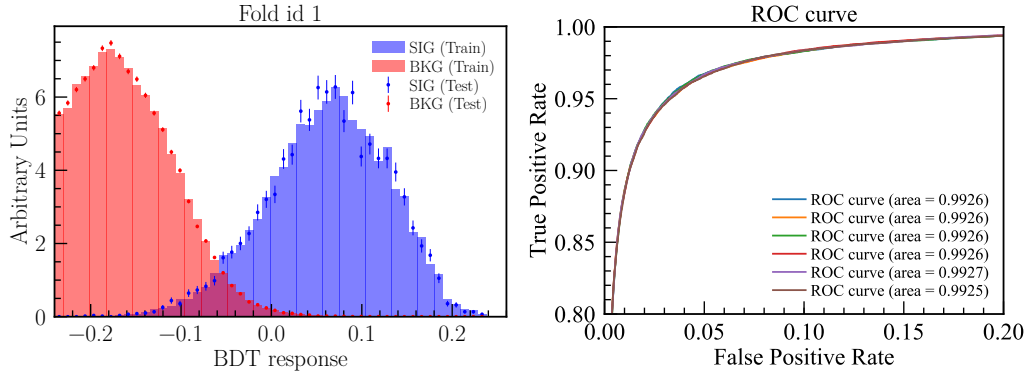


Figure 5.5 – Left: Classifier response for signal (blue) and background (red) for the training and test data in one of the folds. Right: Virtually identical ROC curves for all six folds in a zoomed region to high signal efficiencies and low background rates. The cited area corresponds to the area under the curves and could take a hypothetical value of 1 for a perfect-classification scenario.

except for the signal $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays. An irreducible component of $B_s^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays is not regarded as a background component, but is parametrized in the fit to the reconstructed B^0 mass, as described in the next section. Despite of the absence of other peaking structures, simulated samples have been checked to test whether possible background decays are efficiently vetoed through the selection chain.

Inclusive samples of b hadron decays to $J/\psi(ee) + X$ final states are used to study possible left-over backgrounds. A total of 8 million events has been generated for each of the decays $B^0 \rightarrow J/\psi(e^+e^-)X$, $B_s^0 \rightarrow J/\psi(e^+e^-)X$, $B^+ \rightarrow J/\psi(e^+e^-)X$, and $\Lambda_b^0 \rightarrow J/\psi(e^+e^-)X$. After a candidate selection as described in the previous section, including a preliminary version of the BDT, only the $\Lambda_b^0 \rightarrow J/\psi(e^+e^-)X$ sample has non-negligible left-over events which contribute to the signal region. Here, $\Lambda_b^0 \rightarrow J/\psi(e^+e^-)\Lambda(p\pi^-)$ decays with a $p \rightarrow \pi$ mis-identification constitute a source of background, which is particularly cumbersome as its branching fraction is poorly known. To be sure that these backgrounds are rejected to negligible levels a dedicated veto is required: The reconstructed K_S^0 mass is recalculated under the exchange of the π -mass hypothesis with a p -mass hypothesis, for each of the pions. If for either π the recomputed K_S^0 mass falls within a region of 7 MeV (5 MeV) within the known Λ^0 mass [31] for DD (LL) reconstructed candidates, the candidate is discarded. For an improved signal efficiency of this veto, candidates are only rejected when the exchanged π candidate is more likely to be p than a π according to the PID information, i.e. if $\Delta \ln \mathcal{L}_{p\pi} > 0$. These requirements result in a signal efficiency of $\epsilon_{\text{veto}}^{e^+e^-} \approx 99.6\%$ at a Λ_b^0 background rejection of $\approx 96.2\%$. The number of Λ_b^0 decays in the final data sample is estimated to be 37 ± 8 and therefore negligible.

5.1.5 Modelling of the reconstructed B^0 mass

It has become a standard to use a quite complicated looking PDF to parametrize the reconstructed mass distribution of the signal B^0 component. The Hypatia PDF, which in Ref. [122] is derived from unknown per-event mass resolutions that are marginalized over, and is defined as

$$\mathcal{J}(m|\mu, \sigma, \lambda, \zeta, \beta, a_1, n_1, a_2, n_2) \propto \begin{cases} ((m - \mu)^2 + A_\lambda^2(\zeta)\sigma^2)^{\frac{1}{2}\lambda - \frac{1}{4}} e^{\beta(m - \mu)} K_{\lambda - \frac{1}{2}} \left(\zeta \sqrt{1 + \left(\frac{m - \mu}{A_\lambda(\zeta)\sigma} \right)^2} \right) & , -a_1 < \frac{m - \mu}{\sigma} < a_2 \\ \frac{G(\mu - a_1\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{\left(1 - m / \left(n \frac{G(\mu - a_1\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{G'(\mu - a_1\sigma, \mu, \sigma, \lambda, \zeta, \beta)} - a_1\sigma \right)\right)^{n_1}} & , -a_1 > \frac{m - \mu}{\sigma} \\ \frac{G(\mu - a_2\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{\left(1 - m / \left(n \frac{G(\mu - a_2\sigma, \mu, \sigma, \lambda, \zeta, \beta)}{G'(\mu - a_2\sigma, \mu, \sigma, \lambda, \zeta, \beta)} - a_2\sigma \right)\right)^{n_2}} & , a_2 < \frac{m - \mu}{\sigma}, \end{cases} \quad (5.1)$$

with

$$G(m|\mu, \sigma, \lambda, \zeta, \beta) = ((m - \mu)^2 + A_\lambda^2(\zeta)\sigma^2)^{\frac{1}{2}\lambda - \frac{1}{4}} e^{\beta(m - \mu)} K_{\lambda - \frac{1}{2}} \left(\zeta \sqrt{1 + \left(\frac{m - \mu}{A_\lambda(\zeta)\sigma} \right)^2} \right),$$

and

$$A_\lambda^2(\zeta) = \frac{\zeta K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)},$$

where G is the generalized hyperbolic distribution, G' is the derivative of G and K_λ are cylindrical harmonics. The hyperbolic core is described by a mean, μ , a width, σ , and another parameter, λ . It is extended by power-law-tails to both sides, which depend on the parameters a_1 , a_2 , n_1 , and n_2 . The parameters ζ and β are set to zero. All shape parameters are determined using simulated signal decays. They are fixed in the fits to data, except for μ and σ , which are left floating to account for differences between data and simulation in terms of momentum scaling and resolution. The Hypatia PDF is used for the signal $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ component, as well as for the irreducible $B_s^0 \rightarrow J/\psi(e^+e^-)K_S^0$ component, using the assumption that B^0 and B_s^0 candidates share the same shape parameters and only differ in terms of their mean masses, μ . The combinatorial-background component is described using an exponential function of the form $\propto \exp(a_{\text{exp}} m)$ with a negative slope parameter, a_{exp} .

Fig. 5.6 shows the fit to simulated $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays used to extract the Hypatia shape parameters, as well as a fit to the data at an arbitrary requirement on the BDT prediction, showing the suitability of the description in both samples. Furthermore, all shape parameters except σ (which is floating in the fit to data) have shown to be independent of the BDT working point, which is why an arbitrary BDT requirement of > 0 is used to extract the shape from the simulated sample.

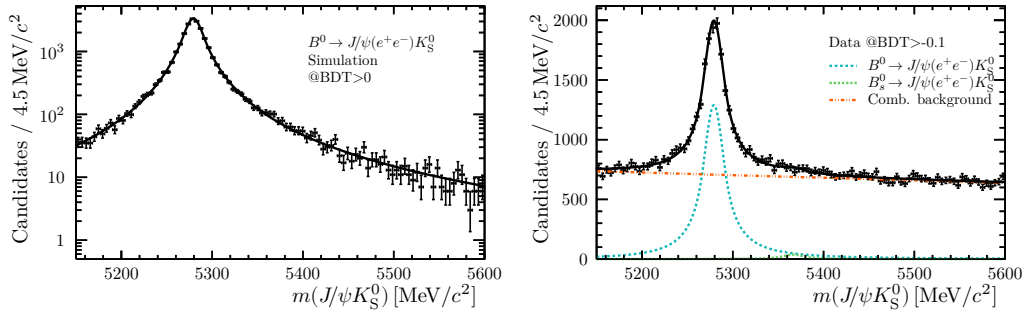


Figure 5.6 – Fit to simulated $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ signal decays to extract the mass shape for the optimization of the BDT cut point (left) and fit to data using an arbitrary BDT requirement of > -0.1 (right).

5.1.6 BDT-selection optimization

To optimize the selection, scans in the BDT prediction are performed on the selected dataset. For each scanned value the resulting reconstructed invariant B^0 mass distribution is fit to determine sWeights for all components in the fit. The weights for the signal B^0 component are used to calculate the FOM as given in Eq. (4.10), by using the per-candidate decay times, t' , as well as their mistag and decay-time-uncertainty estimates, ω and σ_t , respectively. The mistag calibrations are obtained from signal-reweighted control modes (see Sec. 5.5), so that the selection optimization needs to be performed in an iterative fashion: The first iteration uses the uncalibrated OS mistags, due to the lack of a proper calibration for the SS and OS tagging algorithms. The control modes are then prepared as described in Sec. 5.5, to obtain a preliminary calibration for both tagging algorithms, which is used in a second iteration of the optimization scan. This second scan is presented here, as further reweighting and reoptimization does not result in any change.

Fig. 5.7 shows how the FOM and its components, defined in Eqs. (4.13) to (4.17), depend on the BDT prediction value that is required. The maximum in the full FOM is found at a prediction value of 0.01. The effect of the decay-time resolution is constant throughout the whole scan range, so that it does not influence the working point determination. While the tagging performance slightly decreases towards higher values of the BDT requirement, this effect is compensated by the decay-time dependent term (FOM_X). This causes the full FOM to be shifted towards the right with respect to the effective signal size component (S_{eff}). The expression in which the decay-time-dependent term is set constant ($FOM_{B_S^0}$) peaks to the left of the effective signal size, as it would avoid the loss in tagging efficiency towards higher values. As only a very loose requirement on the electron PID has been used in the loose preselection, it has been checked whether additional PID requirements are beneficial. Therefore, a two-dimensional scan is performed, using

5 $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ preparations

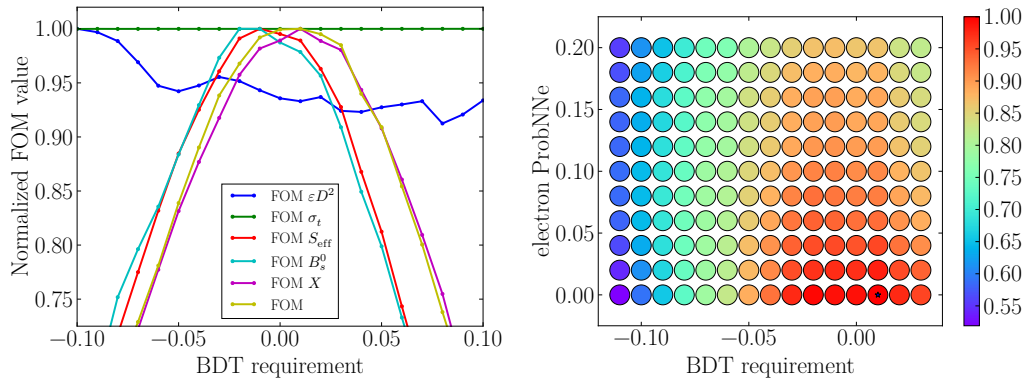


Figure 5.7 – Left: Scan of the FOM and its components as defined in Eqs. (4.13) to (4.17) with respect to the BDT requirement. All graphs are normalized to their respective maximum value. Right: Two-dimensional scan of the full FOM with respect to the BDT requirement and the electron PID value, ProbNNe. The color scale ranges from low normalized FOM values in blue to high values in red. The star marks the optimum, which coincides with the maximum of the one-dimensional scan, thus no additional PID requirement needs to be applied.

a grid of the BDT prediction values and of minimum thresholds for the electron ProbNNe. The results of this scan for the full FOM are shown in Fig. 5.7, where it can be seen an additional requirement cannot increase the quality of the signal compared to the standalone BDT.

5.1.7 Fit windows and multiple candidates

The fit observables are restricted to ranges that do not cause a loss in signal efficiency: A requirement of $5150 < m(J/\psi K_S^0) < 5600 \text{ MeV}/c^2$ helps to fully eliminate partially reconstructed decays at lower masses. Requirements on the decay time and its uncertainty estimate of $0.2 < t' < 15 \text{ ps}$ and $0 < \sigma_t < 0.4 \text{ ps}$ are made, respectively.

Finally, not all remaining candidates in the data sample are unique within an event. In fact, 0.4% of the events count two candidates. Considering the small branching fraction of the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decay, not more than one signal decay can be expected per event. Consequently, in a final selection step a random candidate is chosen for each event containing multiple candidates, reducing the number of candidates to 13395.

The combined signal-selection efficiency of the offline selection with respect to candidates that pass the stripping selection is $\varepsilon_{\text{offsel}}^{e^+e^-} \approx 69\%$. Compared to the previous analysis of $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$, which obtained $\varepsilon_{\text{offsel}}^{\mu^+\mu^-} \approx 90\%$, the value is substantially lower, due to the challenging electron reconstruction.

5.1.8 Trigger structure of selected events

No specific trigger requirements are made in the selection of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays. Instead, any candidate that passes the triggers is accepted, no matter if the candidate itself caused a positive trigger decision or the remaining event. Nevertheless, the dominant HLT2 lines have been checked to test whether the trigger response for the selected candidates is comparable to other analyses with electrons in the final state. The corresponding list, given in Tab. 5.8, confirms the assumption that selected events are triggered by topological trigger lines to a large extent. As described in Sec. 3.3.2, these topological trigger lines perform an inclusive selection of a variety of n -body B decays. The letter “E” (e.g. in `TopoE2BodyBBDT`) refers to lines which require a positive decision of the `LOElectron`, and which therefore allow a looser cut on the response of the BDT used within these lines.

Using the signal MC sample the total trigger efficiency is roughly estimated with respect to candidates that pass the stripping selection as $\varepsilon_{\text{trigger}}^{e^+e^-} \approx 22\%$. The different trigger stages contribute $\varepsilon_{\text{LO}}^{e^+e^-} \approx 50\%$, $\varepsilon_{\text{HLT1}}^{e^+e^-} \approx 63\%$, and $\varepsilon_{\text{HLT2}}^{e^+e^-} \approx 70\%$, where candidates that pass the respective previous trigger stage are chosen for the normalization. These trigger efficiency values are significantly below the trigger efficiencies of $J/\psi(\mu^+\mu^-)$ modes like $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$, which can be triggered at $\varepsilon_{\text{trigger}}^{\mu^+\mu^+} \approx 90\%$, due to the characteristic experimental signature of the muons.

Table 5.8 – Top ten HLT2 lines and rates at which they accept candidates after the full selection. All lines are topological trigger lines, which perform an inclusive selection of a variety of multi-body B decays. Details on the selection strategy of these lines and their requirements can be found in Refs. [78, 79].

HLT2 line	Rate [%]	HLT2 line	Rate [%]
<code>Topo2BodyBBDT</code>	66	<code>TopoE4BodyBBDT</code>	17
<code>TopoE2BodyBBDT</code>	65	<code>Topo2BodySimple</code>	16
<code>Topo3BodyBBDT</code>	44	<code>TopoRad2BodyBBDT</code>	10
<code>TopoE3BodyBBDT</code>	34	<code>Topo3BodySimple</code>	8
<code>Topo4BodyBBDT</code>	20	<code>TopoRad2plus1BodyBBDT</code>	8

5.2 Results of the reconstructed-mass fit

After the full candidate selection a fit to the reconstructed B^0 mass is performed for the unfolding of the signal component, using the parametrization discussed in Sec. 5.1.5. The detailed fit results are given in Tab. 5.9 and the projection is given in Fig. 5.8. In total 10629 ± 139 $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and 70 ± 41 $B_s^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays are observed in the fit. From the result of the fit `sWeights` are determined following the procedure described in Sec. 4.3. These weights are used in the decay-time dependent CP fit described in Ch. 7. The previous LHCb analysis observed a total of 114 000 $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ decays [19]. Compared to this result, the

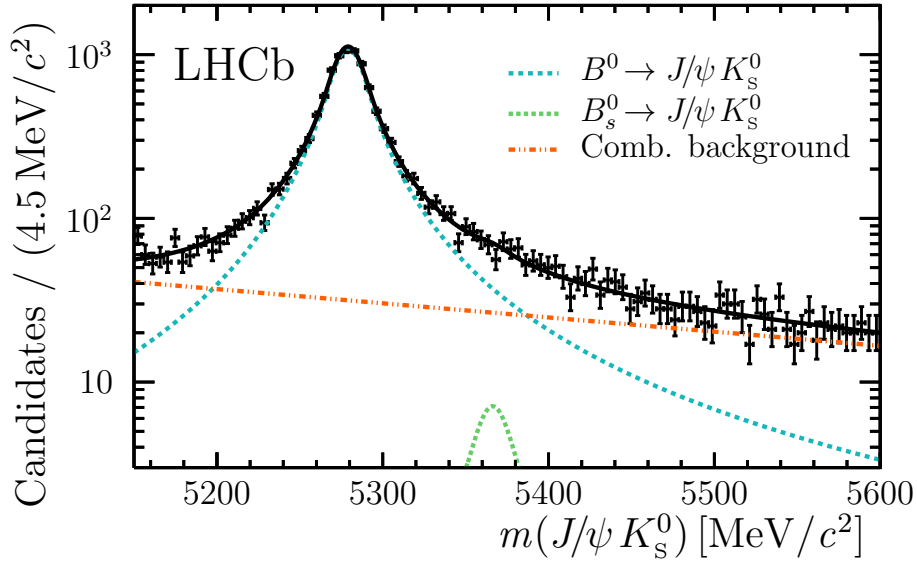


Figure 5.8 – Invariant mass of the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ candidates [20]. The lines represent the result of the fit for the extraction of signal-decay weights that are used in the sFit for the CP -violation measurement.

number of observed $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays is consistent when considering the previously discussed efficiency losses that lead to a relative signal selection efficiency of $\approx 9\%$ with respect to the efficiency of the previous analysis. As the number of $B_s^0 \rightarrow J/\psi(e^+e^-)K_S^0$ candidates is compatible with zero, fits have been performed without the corresponding component, and also constraining the component using the known $B_s^0 \rightarrow J/\psi(e^+e^-)K_S^0$ branching ratio from Ref. [31] and the ratio of the hadronization fractions of B_s^0 and B^0 mesons, f_s/f_d , from Ref. [123]. These checks have resulted in fully compatible values in the mass fit and the CP fit.

5.3 Production asymmetry

Despite the dominant pairwise production of b and \bar{b} at the LHC, their hadronizations to B^0 (B^+) and \bar{B}^0 (B^-) occur at different rates. The effect is caused through the interaction of these quarks with the beam remnants, consisting of the proton valence quarks [124–126]. The production asymmetry is defined using the production cross sections, σ , for \bar{B}^0 and B^0 as

$$A_P(B^0) = \frac{\sigma(\bar{B}^0) - \sigma(B^0)}{\sigma(\bar{B}^0) + \sigma(B^0)}. \quad (5.2)$$

Table 5.9 – Results of the fit to the reconstructed mass of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ candidates after the full selection. The mass difference between the B^0 and the B_s^0 meson, $\mu_{B_s^0} - \mu_{B^0}$, is fixed to its known value [31]. The other parameters are the Hypatia shape parameters, $\alpha_1, \alpha_2, n_1, n_2, \beta, \zeta, \lambda$, and σ , the slope parameter of the exponential function for the combinatorial-background, $a_{\text{exp.}}$, and the yields of the B^0 component, n_{B^0} , of the B_s^0 component, $n_{B_s^0}$, and of the combinatorial-background component, $n_{\text{comb.}}$.

Parameter	Value	Constraint
α_1	0.21	const.
α_2	0.22	const.
n_1	2.27	const.
n_2	1.99	const.
β	0.0	const.
ζ	0.0	const.
λ	-1.05	const.
$\mu_{B_s^0} - \mu_{B^0}$ [MeV/ c^2]	87.2	const.
σ [MeV/ c^2]	57.4 \pm 1.2	floating
μ_B^0 [MeV/ c^2]	5279.1 \pm 0.2	floating
$a_{\text{exp.}}$ [1/ (MeV/ c^2)]	-0.00199 \pm 0.00021	floating
n_{B^0}	10629 \pm 139	floating
$n_{B_s^0}$	70 \pm 41	floating
$n_{\text{comb.}}$	2696 \pm 108	floating

Using the measurement of LHCb [127] in bins of transverse momentum, p_T , and rapidity, y , the values of the production asymmetry for an arbitrary decay mode can be computed as

$$A_P = \sum_i \epsilon_i A_{P,i}, \quad (5.3)$$

where ϵ_i is the fraction of signal candidates in bin i , and $A_{P,i}$ denotes the value of the production asymmetry in bin i taken from Ref. [127]. As distinct measurements have been published for the two years of data-taking (2011, 2012) and the respective center-of-mass energies (7 TeV, 8 TeV), two values are computed, yielding

$$A_P^{11} = -0.0100 \pm 0.0084 \text{ (stat.)} \pm 0.0005 \text{ (syst.)} \quad (5.4)$$

$$A_P^{12} = -0.0077 \pm 0.0054 \text{ (stat.)} \pm 0.0004 \text{ (syst.)}. \quad (5.5)$$

These values are well in accordance with the expected magnitude of the effect and show no significant production asymmetry. To propagate the uncertainty to the measurement of the CP parameters, the production-asymmetry values are Gaussian constrained using their statistical uncertainty in the CP fit (see Sec. 7.2). The systematic uncertainty is covered in dedicated studies (see Sec. 7.5.4).

5.4 Correction of detector effects

The theoretical decay-time distributions defined in Eqs. (2.37) to (2.40) need to be corrected for detector effects, in order to obtain a PDF, which can describe the distribution of the reconstructed decay time, t' , appropriately. Two effects are considered, which are assessed using simulated samples.

The first effect is the finite decay-time resolution of the detector, which causes a deviation of the theoretical decay time, t , from the reconstructed decay time, t' . It can be considered by convolving the theoretical decay-time PDF with a resolution model, \mathcal{R} , describing the probability density of the deviations.

The second effect is the decay-time acceptance, introduced by a decay-time-dependent efficiency coming from the reconstruction of the candidates. The corresponding efficiency function, $\varepsilon(t')$, is multiplied with the PDF after the convolution with the resolution model.

5.4.1 Decay-time resolution

The main impacts on the decay-time resolution of the candidates are the momentum, impact-parameter and vertex resolution of the reconstructed particles. The DTF's propagated uncertainty of the decay-time measurement, σ_t , is obtained for each candidate, but the true standard deviation of the effect is known to differ, which cannot only be observed on simulated events, but also on data: Often analyses build pseudo-signal candidates from tracks originating from the primary vertex. This pseudo-signal then allows to study candidates with negative reconstructed decay times, which make it possible to assess the magnitude of the resolution effect. However, this usually introduces the need for corrections of differences between the kinematic properties of such pseudo-candidates and the actual signal kinematics. For this reason, and as the decay-time resolution in the B^0 system is of minor impact to the CP measurement, the resolution model is fully obtained from simulated signal samples. To account for possible deviations from the true resolution, the effects of such deviations are conservatively estimated in dedicated systematic uncertainty studies (see Sec. 7.5.4).

The parametrization of the resolution model is chosen as a sum of three Gaussians, carrying widths that have a linear dependence on the decay-time uncertainty estimate, σ_t . The third Gaussian describes candidates that are associated to the wrong PV, and therefore does not depend on the σ_t . The resolution model is then given as

$$\mathcal{R}(t' - t | \sigma_t) = \frac{1}{\sqrt{2\pi}} \left(\sum_{i=1}^2 \frac{f_i}{a_i \cdot \sigma_t + b_i} e^{-\frac{(t' - t - \mu)^2}{2(a_i \cdot \sigma_t + b_i)^2}} + \frac{f_{\text{wPV}}}{\sigma_{\text{wPV}}} e^{-\frac{(t' - t - \mu)^2}{2\sigma_{\text{wPV}}^2}} \right), \quad (5.6)$$

where a_i and b_i are the parameters of the linear calibrations, σ_{wPV} is the width of the Gaussian describing the wrong PV association and the parameters f denote fractions of the different Gaussians. The linear σ_t -dependence that is employed

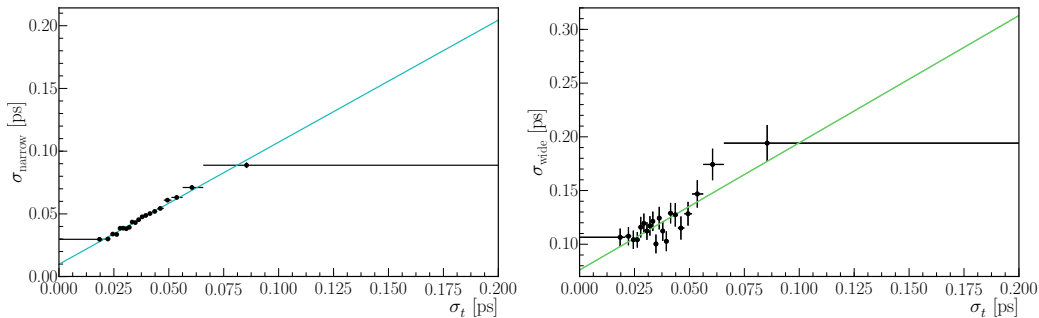


Figure 5.9 – Linear calibration functions for the scaling of the decay-time uncertainty estimate σ_t from a simultaneous fit to the observed $t' - t$ distribution in bins of σ_t using simulated signal decays. Horizontal errorbars indicate the chosen bin widths in σ_t , while the vertical errorbars are the fit uncertainties for the narrow (left) and the wide (right) Gaussian component of the resolution model.

for the first two Gaussians is established studying the $t' - t$ distribution using 20 equally filled bins in σ_t . The sub samples are fitted simultaneously using a discrete version of the resolution model in Eq. (5.6), composed of two Gaussians that carry widths that are individual to each σ_t bin, and a third Gaussian, whose width is shared among all σ_t bins. Fig. 5.9 shows linear minimum χ^2 fits to $\sigma_{\text{narrow}}(\sigma_t)$ and $\sigma_{\text{wide}}(\sigma_t)$, which confirm that a linear dependence can be assumed.

Finally, the parameters of the resolution model are determined through a fit to the simulated signal sample, which yields the results given in Tab. 5.10 and the fit projections in Fig. 5.10. The resulting effective single Gaussian resolution of correctly associated $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays is 67 fs. In the CP fit to data, the resolution parameters are fixed and μ is set to zero. Neglecting the finite decay-time bias and potential deviations of the factors that scale the decay-time-uncertainty estimates is studied as a source of systematic uncertainty (see Sec. 7.5.4).

5.4.2 Decay-time acceptance

The decay-time acceptance originates from different effects that cause a decay-time-dependent reconstruction inefficiency of the B^0 candidates. At low decay times candidates are less likely to be reconstructed, as the selection uses variables that are sensitive to the particles' separation from the primary vertex, like requirements on impact parameters and flight distances. For higher B^0 decay times, candidates leave less hits in the VELO, which also decreases the reconstruction probability and therefore causes a loss in efficiency [128, 129].

The decay-time acceptance function is by construction defined to describe any deviation of the tag-integrated decay-time distribution from the exponential shape. In general, there is no particular physically motivated mathematical function to describe the inefficiency. The only requirements are smoothness, and, if possible, efficient computational evaluation of integrals [130]. A flexible function that satisfies

Table 5.10 – Parameters of the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decay-time-resolution model determined from a fit to simulated decays. The fraction of the narrow Gaussian is implicitly defined through $f_{\text{narrow}} = 1 - f_{\text{wpv}} - f_{\text{wide}}$. All parameters are fixed during the CP fit to data and μ is set to zero.

Parameter	Fit result
μ [fs]	-2.90 ± 0.29
f_{wide}	0.151 ± 0.009
f_{wpv}	0.0033 ± 0.0007
a_{narrow}	0.957 ± 0.022
b_{narrow} [ps]	0.0090 ± 0.0008
a_{wide}	1.363 ± 0.129
b_{wide} [ps]	0.071 ± 0.005
σ_{wpv} [ps]	0.885 ± 0.095

these two needs are cubic splines, which are differentiable up to the second order at any point, and whose integral can be calculated analytically. Here used is their representation in form of B-Splines, defined through a set of n interval boundaries (knot positions), which uniquely define the $n + 2$ base functions b_i . While the base functions form a partition of unity, a superposition with coefficients c_i , can be used to obtain the efficiency parametrization

$$\epsilon(t) = \sum_i^n c_i b_i(t) \quad \text{with} \quad \sum_i b_i(t) = 1. \quad (5.7)$$

The more dense the interval boundaries are chosen, the more base splines are available in a given region, allowing for an accurate interpolation of the efficiency, whose functional shape can be altered by varying the coefficients. Furthermore, setting the last two coefficients equal ($c_{n+2} = c_{n+1}$) ensures a constant efficiency value at the end of the decay-time range and ensures the stability of the parametrization in this low-statistics region. As the absolute scale of the efficiency function can be absorbed into the overall PDF normalization and only relative values of the function are meaningful, one of the parameters is set to an arbitrary constant value of 1, defining the scale and preventing ambiguities in the fitting procedure.

The knot positions themselves are optimized empirically on simulated signal samples, where the B^0 lifetime can be fixed to the value that was used in the generation of the samples. The boundaries are placed at $\{0.2, 0.5, 1.0, 4.0, 15.0\}$ ps, allowing for a more granular description in the low decay-time region. In the sFit used to extract the CP coefficients these positions are fixed and the coefficients are left floating, while the B^0 lifetime is constrained to the world average value and the full fit model described in Ch. 7 is used. Projections of the decay-time fit for the simulated signal sample and data are provided in Fig. 5.11, together with the respective spline acceptances. The corresponding coefficients are given in Tab. 5.11 and point to a slight non-agreement between data and simulation

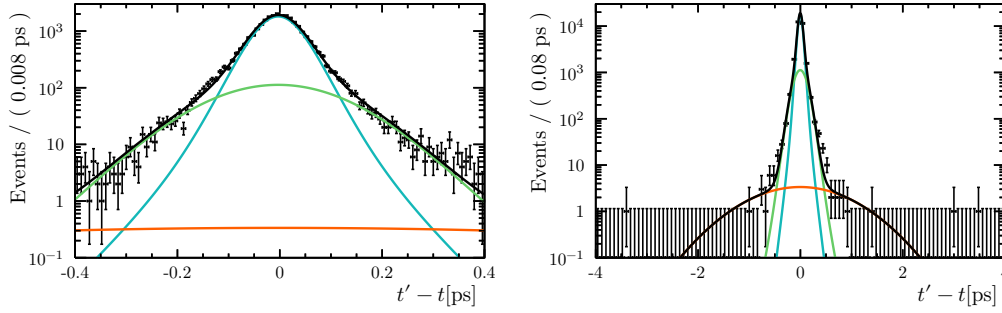


Figure 5.10 – Projection of the resolution model that is fit to simulated $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays in a narrow (left) and wide (range) of the decay-time residual, $t' - t$. The total model (black) consists of a narrow (blue) and a wide (green) Gaussian component that use linear functions of the decay-time uncertainty estimate as widths, and a wide average Gaussian component that accounts for wrong PV associations (orange).

in the low acceptance region. As the data is well described, possible deviations are covered in the systematic studies, where a more granular efficiency function including more interval boundaries is used.

Table 5.11 – Decay-time-acceptance parameters for $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays in the simulated signal sample and data. The coefficient c_3 is set to a constant value of 1 to fix the arbitrary scale of the efficiency function.

Parameter	Simulation	Data
c_0	0.021 ± 0.007	0.015 ± 0.010
c_1	0.128 ± 0.011	0.007 ± 0.017
c_2	0.526 ± 0.017	0.502 ± 0.032
c_3	1	1
c_4	0.67 ± 0.06	0.74 ± 0.12
c_5	0.76 ± 0.09	0.64 ± 0.18

5.5 Flavour tagging

The flavour-tagging calibration has been performed using a tool developed by collaborators of the flavour-tagging group at the LHCb experiment, the Espresso Performance Monitor [131]. It can perform flavour-tagging calibrations based on binomial regression fits, given a signal-weighted dataset of a flavour-specific decay. For the calibration using control modes of flavour-specific decays of charged b mesons (self-tagging decays), only an identification of the final-state flavour is

5 $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ preparations

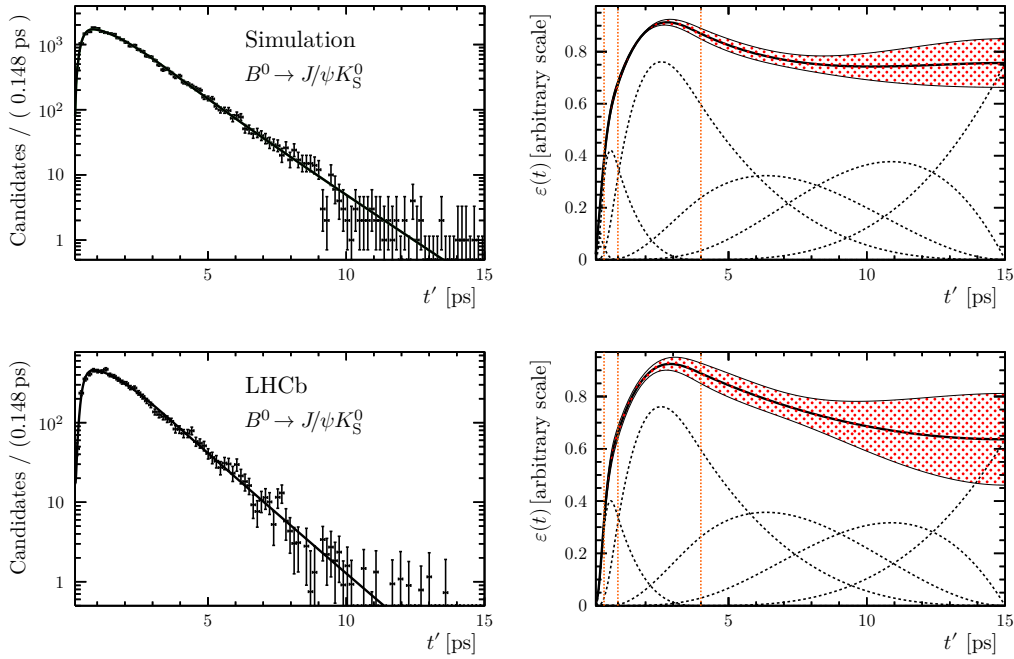


Figure 5.11 – Projections of the decay-time fit (left) and the respective acceptance function (right) for the simulated signal sample (top) and data (bottom) of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays. The red-dotted region corresponds to the uncertainty of the acceptance description coming from the fit, the dashed vertical lines represent the inner interval boundaries, and the dashed curves the base splines.

needed, which is why such modes are also referred to as self-tagging decays. In contrast, for flavour-specific decay modes of neutral b mesons, which can oscillate until their decay, a fit to the decay-time is performed during the calibration. This fit allows to determine the true mistag rate, ω , through a measurement of the mixing asymmetry, defined as

$$\mathcal{A}_{\text{mix}}(t) = \frac{N_{\text{unmixed}}(t) - N_{\text{mixed}}(t)}{N_{\text{unmixed}}(t) + N_{\text{mixed}}(t)} = (1 - 2\omega) \cos(\Delta mt), \quad (5.8)$$

where N_{unmixed} and N_{mixed} denote the number of mixed and unmixed neutral b -mesons for a given decay time, t , respectively.

5.5.1 Control-mode preparation

Two control modes are chosen, having similar decay topologies and kinematic properties to the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode: the SS tagging is calibrated using flavour-specific $B^0 \rightarrow J/\psi(e^+e^-)K^{*0}$ decays, while for the OS tagging self-tagging $B^+ \rightarrow J/\psi(e^+e^-)K^+$ decays are studied. These two control modes are selected in

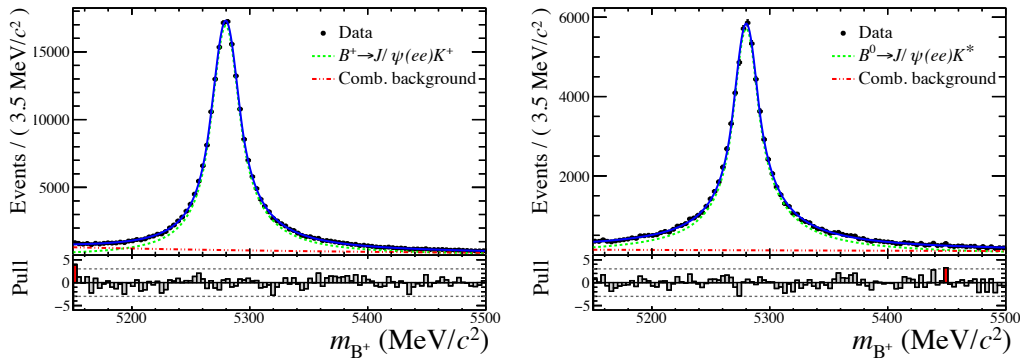


Figure 5.12 – Fits to the reconstructed mass of $B^+ \rightarrow J/\psi(e^+e^-)K^+$ candidates (left) and $B^0 \rightarrow J/\psi(e^+e^-)K^{*0}$ candidates (right). The data is described using a Hypatia function for the signal contribution (green) and an exponential function for the combinatorial background (red).

a similar fashion to the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode, i.e. through a sequence of a cut-based and a multivariate BDT selection. However, the multivariate selection is not optimized using the full FOM presented in Sec. 4.4.4, but through optimizing the effective signal size of the signal component, defined in Eq. (4.15). Fits to the invariant mass distributions, after a full selection of the candidates, are shown in Fig. 5.12. In total $226\,757 \pm 611$ $B^+ \rightarrow J/\psi(e^+e^-)K^+$ decays and $82\,709 \pm 430$ $B^0 \rightarrow J/\psi(e^+e^-)K^{*0}$ decays are observed in the fit, which are used for the taggers' calibrations.

Despite the similarity of the calibration modes to the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode, there are still slight differences found in distributions of observables, which the flavour-tagging calibration is known to depend on: the number of primary vertices in the event, nPV, the number of tracks in the event, nTracks, and the b -meson's pseudo-rapidity, η_p , its angle, ϕ , and its transverse momentum, p_T . These differences are caused by the different requirements used to select the control modes. For example, in the case of the $B^+ \rightarrow J/\psi(e^+e^-)K^+$ mode only a single charged kaon track is required to originate from the same vertex as the two electrons. As this allows for stricter requirements on the quality of the B^+ decay vertex, the requirements on the p_T of the particles can be loosened, without a sacrifice in background rejection. This in turn also results in a lower p_T spectrum and higher average pseudo-rapidity values for $B^+ \rightarrow J/\psi(e^+e^-)K^+$ candidates compared to $B^0 \rightarrow J/\psi K_S^0$ decays. The same is true for $B^0 \rightarrow J/\psi(e^+e^-)K^{*0}$ decays, where the electron pair and the $K^+ \pi^-$ pair originate from the same vertex. Likewise, a higher number of PVs and tracks can be tolerated in the event selection of the calibration modes, affecting the corresponding average numbers. As different distributions of the mentioned observables could affect the portability of the calibration, a reweighting procedure based on BDTs, Gradient Boosting Reweighting [132], is used to compute candidate weights, whose application compensates the differences

in the corresponding distributions. After the reweighting of the control modes, the samples are used to calibrate the single OS and SS taggers, introduced in Sec. 4.1.

5.5.2 Calibration strategy

A linear function mapping the predicted mistag, η , to the true mistag, ω , has been found to yield an adequate description for the calibration of the individual taggers. In general this linear function can be parametrized as

$$\omega = p_0 + p_1 (\eta - \langle \eta \rangle) \quad (5.9)$$

with the calibration parameters p_0 , and p_1 which would equal to the average mistag rate, $\langle \eta \rangle$, and 1, respectively, if the tagger was perfectly calibrated. However, to take into account differences in the tagging responses for initial B^0 and \bar{B}^0 mesons, separate functions are used for both initial states:

$$\omega^{B^0} = p_0^{B^0} + p_1^{B^0} (\eta - \langle \eta \rangle) , \quad (5.10)$$

$$\omega^{\bar{B}^0} = p_0^{\bar{B}^0} + p_1^{\bar{B}^0} (\eta - \langle \eta \rangle) . \quad (5.11)$$

Furthermore, the calibration parameters can be expressed through the tagging asymmetry parameters Δp_0 and Δp_1 as

$$p_i^{B^0} = p_i + \frac{\Delta p_i}{2} , \quad (5.12)$$

$$p_i^{\bar{B}^0} = p_i - \frac{\Delta p_i}{2} . \quad (5.13)$$

Treatment of near-limit mistags

A general problem with this choice of parametrization is that it does not necessarily yield values in a sane range of $[0, 0.5]$. Therefore, the question arises, how calibrated mistags that are smaller than 0 or greater than 0.5 are treated. Due to the distribution of η for the individual taggers, calibrated mistags smaller than 0 occur rarely. When they do occur, they have to be interpreted as extremely good tags, so that ω is set to 0. In contrast, mistags greater than 0.5 correspond to extremely bad tags, which should therefore be set to 0.5. A frequently discussed point is, whether it is legitimate to flip the tag decision and assign a new mistag corresponding to the inverted calibrated mistag decision, $1 - \omega$. However, recent studies of a collaborator working on the analysis of decay-time-dependent CP violation in $B^0 \rightarrow D^\mp \pi^\pm$ decays [133] have shown that such a flip can lead to biased results for the CP parameters. For this reason, it is preferred to set the events untagged. While this works for the calibration of the single taggers, another challenge arises in the CP fit, in which Gaussian constraints on the flavour-tagging calibration parameters are used to propagate their statistical uncertainties into the result (see Sec. 7.2). In these constraints, it can happen that tagged events with high (combined OS or SS) η values are shifted above the meaningful threshold of $\omega > 0.5$.

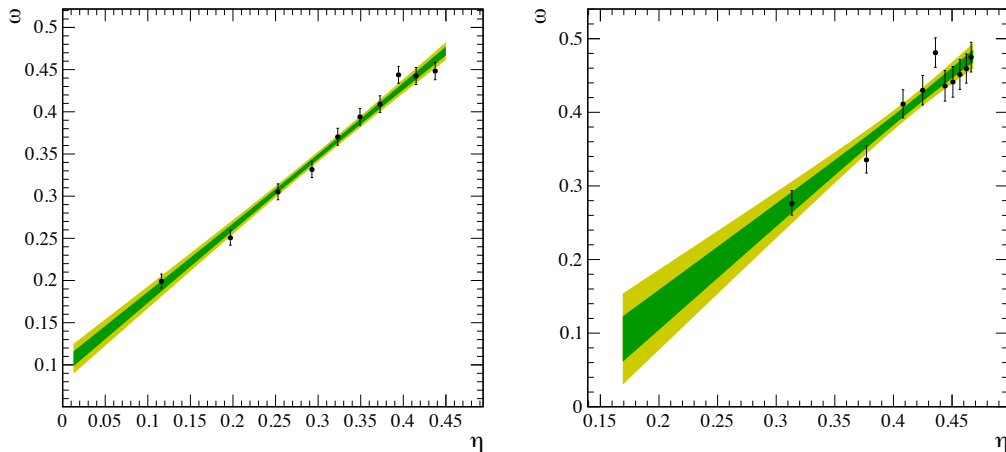


Figure 5.13 – Flavour-tagging calibration using reweighted $B^+ \rightarrow J/\psi(e^+e^-)K^+$ decays for the OS-combination tagger (left) and reweighted $B^0 \rightarrow J/\psi(e^+e^-)K^{*0}$ decays for the SS-combination tagger (right).

It is found, that these shifts cause discontinuities in the likelihood function, which lead to instabilities during the minimization. Therefore, the strategy is to limit the combined mistags beforehand by truncating the OS and SS mistag distributions. The thresholds are chosen so that no calibrated mistags above 0.5 can occur, when varying the calibration parameters by one standard deviation. This results in the additional requirements

$$\omega_{\text{OS}} \leq 0.450, \quad (5.14)$$

$$\omega_{\text{SS}} \leq 0.468, \quad (5.15)$$

for tagged candidates. Candidates which lie above these thresholds are not completely rejected, but set to an untagged state, i.e. $d'_{\text{OS}} = d'_{\text{SS}} = 0$, and $\omega_{\text{OS}} = \omega_{\text{SS}} = 0.5$, which causes a negligible loss in tagging power. Following this step, final recalibrations of the OS-combination and the SS-combination tagger need to be performed, which are shown in Fig. 5.13.

5.5.3 Evaluation of systematic uncertainties and results

Two types of systematic uncertainties are considered for the flavour-tagging calibration. The first type covers uncertainties that stem from the calibration method itself and account for a potential systematic variation of the calibration within the control channels alone. The second type considers a systematic effect that arises from using calibration parameters from the control modes for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode, caused by e.g. kinematic differences between the two.

Validity of the calibration method

The systematic uncertainty that arises within the control channels themselves is assessed by performing various splits of the calibration samples. For each subsample the flavour-tagging calibration is performed and each flavour-tagging parameter is compared to its corresponding parameter obtained in the statistically independent subset. Here, obvious categories of the LHCb data are used to create these splits: the years of data taking (2011 vs. 2012) and the magnet polarity used during data taking (MagUp vs. MagDown). For none of the calibration parameters a significant deviation is found between subsets, which is why no systematic uncertainty is assigned for the validity of the calibration method.

Portability of the control channel

The reweighting of the control channels' kinematics to the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode should ensure a good portability of the flavour-tagging calibration. Still, differences can be present, that are assessed by studying the impact of the reweighting procedure on the flavour-tagging-calibration parameters, which can be done for both sorts of tagging algorithms, OS and SS. Moreover, it is possible to perform the calibration of the OS algorithms on both employed reweighted calibration modes, which offers two statistically independent calibrations that can be compared with each other.

All these checks yield calibration values that are compatible with each other. Nevertheless, a conservative approach to cover possible systematic effects is chosen: Half of the computed differences of the parameters obtained from calibrations with and without reweighting are assigned as systematic uncertainties on the flavour-tagging calibration. Even with this conservative approach the systematic uncertainties are far below the statistical uncertainties of the calibration, as can be seen in Eqs. (5.16) and (5.17).

Results of the flavour-tagging calibration

The analysis of $226\,757 \pm 611$ $B^+ \rightarrow J/\psi(e^+e^-)K^+$ decays results in flavour-tagging calibration parameters of

$$\begin{aligned}
p_{0,\text{OS}}^{J/\psi} &= 0.3603 \pm 0.0031 \text{ (stat.)} \pm 0.0006 \text{ (syst.)}, \\
p_{1,\text{OS}}^{J/\psi} &= 0.834 \pm 0.029 \text{ (stat.)} \pm 0.002 \text{ (syst.)}, \\
\Delta p_{0,\text{OS}}^{J/\psi} &= 0.0162 \pm 0.0061 \text{ (stat.)} \pm 0.0004 \text{ (syst.)}, \\
\Delta p_{1,\text{OS}}^{J/\psi} &= -0.024 \pm 0.057 \text{ (stat.)} \pm 0.004 \text{ (syst.)}, \\
\langle \eta_{\text{OS}}^{J/\psi} \rangle &= 0.316,
\end{aligned} \tag{5.16}$$

for the combined tagging decision of the OS tagging algorithms. Furthermore, the analysis of $82\,709 \pm 430$ $B^0 \rightarrow J/\psi(e^+e^-)K^{*0}$ decays results in

$$\begin{aligned}
p_{0,\text{SS}}^{J/\psi} &= 0.4213 \pm 0.0063 \text{ (stat.)} \pm 0.0009 \text{ (syst.)}, \\
p_{1,\text{SS}}^{J/\psi} &= 1.29 \pm 0.12 \text{ (stat.)} \pm 0.01 \text{ (syst.)}, \\
\Delta p_{0,\text{SS}}^{J/\psi} &= 0.0026 \pm 0.0086 \text{ (stat.)} \pm 0.0009 \text{ (syst.)}, \\
\Delta p_{1,\text{SS}}^{J/\psi} &= 0.036 \pm 0.175 \text{ (stat.)} \pm 0.016 \text{ (syst.)}, \\
\langle \eta_{\text{SS}}^{J/\psi} \rangle &= 0.425,
\end{aligned} \tag{5.17}$$

for the combined tagging decision of the SS algorithms. These calibration parameters are used in the decay-time-dependent CP fit, where their statistical uncertainties, as well as the linear correlation coefficients between the parameters, as given in Tab. 5.12, are used as multidimensional Gaussian constraints.

Table 5.12 – Linear correlation coefficients between the flavour-tagging parameters for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode.

	OS correlations				SS correlations			
	p_0	p_1	Δp_0	Δp_1	p_0	p_1	Δp_0	Δp_1
p_0	1.000	0.139	0.010	-0.023	1.000	0.122	-0.028	0.019
p_1	0.139	1.000	-0.023	0.068	0.122	1.000	0.020	-0.082
Δp_0	0.010	-0.023	1.000	0.139	-0.028	0.020	1.000	0.077
Δp_1	-0.023	0.068	0.139	1.000	0.019	-0.082	0.077	1.000

5.5.4 Performance

The effective tagging efficiency for the combination of the OS and SS taggers in the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ data sample is calculated, following Eq. (4.1), as

$$\varepsilon_{\text{eff}}^{B^0 \rightarrow J/\psi(e^+e^-)K_S^0} = (5.93 \pm 0.29) \%, \quad (5.18)$$

where a sole use of either OS and SS algorithms would contribute

$$\varepsilon_{\text{eff, OS}}^{B^0 \rightarrow J/\psi(e^+e^-)K_S^0} = (3.60 \pm 0.13) \%, \quad (5.19)$$

$$\varepsilon_{\text{eff, SS}}^{B^0 \rightarrow J/\psi(e^+e^-)K_S^0} = (2.40 \pm 0.28) \%, \quad (5.20)$$

on the full dataset.

The performance of the single taggers has been evaluated using the control modes. For the OS algorithms the highest tagging power is obtained from the OS vertex charge (1.37%), the OS muon (1.23%) and the OS kaon (1.13%) taggers. Lower tagging powers are obtained from the OS charm (0.47%) and the OS electron (0.38%) taggers. For the SS algorithms, the SS pion tagger contributes significantly more (1.56%) than the SS proton tagger (0.28%).

All three values in Eqs. (5.18) to (5.20) individually exceed the tagging power of the previous LHCb analysis which used $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ decays [19]. Using only the SS pion tagger for the same side combination and not including the OS charm tagger, this previous analysis obtained [119]

$$\varepsilon_{\text{eff}}^{B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0} = (3.02 \pm 0.05) \%, \quad (5.21)$$

$$\varepsilon_{\text{eff, OS}}^{B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0} = (2.63 \pm 0.04) \%, \quad (5.22)$$

$$\varepsilon_{\text{eff, SS}}^{B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0} = (0.38 \pm 0.02) \%. \quad (5.23)$$

Given the low tagging powers of the OS charm and the SS proton tagger, the difference cannot be explained by the inclusion of these new taggers alone. This is also evident when comparing the tagging powers in $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays to the lower values obtained in the $B^0 \rightarrow \psi(2S)K_S^0$ mode, which uses the same combination of taggers (see Sec. 6.7).

Concerning the better OS-tagging performance of the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode compared to the dimuon mode, the explanation is found in the dependence of the OS tagging on the type of the trigger response. For this purpose a distinction for the L0 trigger is made between events that are triggered independent of the signal candidate (TIS), and events that are triggered on the signal candidate (TOS). From Tab. 5.13 it can be seen that the tagging power of TIS candidates is three times as high as for TOS candidates. In contrast to TOS events, where the part of the event that is used for the OS tagging does not need to fulfil any trigger requirements, TIS events by definition contain other particles than the signal that are sufficient to cause a positive trigger decision. Therefore, the posterior probability to find an

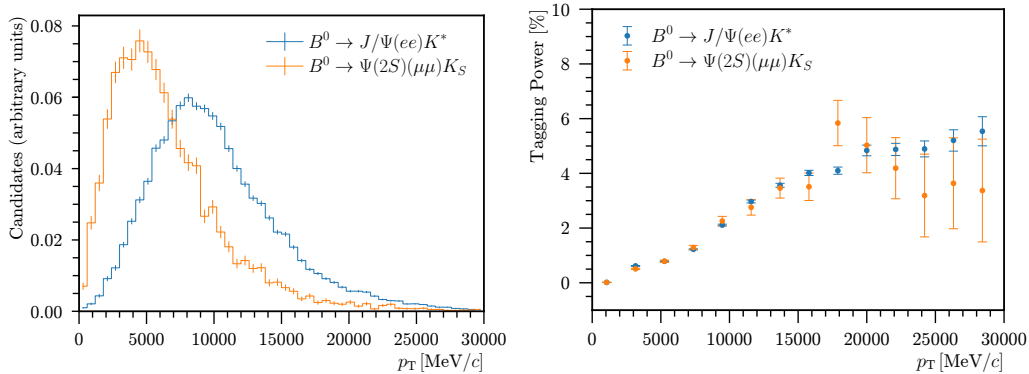
OS b -hadron of good reconstruction quality is higher for the latter type of events, which directly translates into higher tagging efficiencies, ε , and lower mistag rates, ω , for such events. As the fraction of inclusively L0-TIS triggered events in the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ dataset (50%) is significantly higher compared to the fraction in the $B^0 \rightarrow \psi(2S)K_S^0$ dataset (24%), this results in the higher overall performance of the OS algorithms for the former.

On the other hand, the increased performance of the SS taggers cannot be explained from the different population in trigger categories, as can be seen in Tab. 5.13. Instead, the effective tagging efficiency of the SS tagging is observed to increase with the B^0 p_T , which, together with the harder momentum spectrum of the signal candidates in the electron mode, causes the higher value compared to the dimuon modes (see Fig. 5.14).

Table 5.13 – Tagging-performance measures for $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays in L0-trigger subsets for both tagging algorithms.

	L0 response	ε [%]	ω	ε_{eff} [%]
OS	TIS excl.	44.43 ± 0.38	32.86 ± 0.50	5.22 ± 0.31
	TOS excl.	28.95 ± 0.22	37.61 ± 0.42	1.78 ± 0.12
	TIS and TOS	46.66 ± 0.35	33.35 ± 0.45	5.17 ± 0.28
SS	TIS excl.	73.09 ± 0.63	43.16 ± 1.17	1.37 ± 0.47
	TOS excl.	74.29 ± 0.39	41.07 ± 0.71	2.37 ± 0.37
	TIS and TOS	75.90 ± 0.53	40.46 ± 0.96	2.76 ± 0.56

Figure 5.14 – Distribution of the B^0 transverse momentum for signal-reweighted $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ candidates and for $B^0 \rightarrow \psi(2S)K_S^0$ candidates (left) and dependence of effective tagging performance in bins of the B^0 transverse momentum (right).



6 $B^0 \rightarrow \psi(2S)K_S^0$ preparations

For the preparation of the $B^0 \rightarrow \psi(2S)K_S^0$ dataset, a similar strategy compared to the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ dataset has been pursued, which helps to keep the overall picture of the analysis consistent. Candidates are selected, aiming for maximum sensitivity to the CP parameter S , and fits to the invariant reconstructed B^0 mass are performed to unfold the signal $B^0 \rightarrow \psi(2S)K_S^0$ component and to optimize the selection. Finally, decay-channel-dependent detector effects are analyzed and the calibration of the flavour-tagging response is performed using signal-reweighted control modes. This chapter presents the indispensable aspects of these preparations to obtain a full picture of the analysis - a more detailed presentation will be given in Ref. [21].

6.1 Stripping and trigger requirements

The initial $B^0 \rightarrow \psi(2S)K_S^0$ candidate selection is performed through version 20r1 and 20r0 of the `Bd2Psi2SKsMuMuPrescaled` line for 2011 and 2012 data, respectively. The requirements for the $\psi(2S)$, the K_S^0 , and the B^0 candidate are listed in Tab. 6.1 to 6.3. The combined reconstruction and stripping efficiency of these requirements is $\epsilon_{\text{reco+strip}}^{\psi(2S)} \approx 9.7\%$, with respect to candidates lying inside the LHCb acceptance. In contrast to the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode, decays of $B^0 \rightarrow \psi(2S)K_S^0$ can be very efficiently triggered due to the characteristic signature of the $\psi(2S) \rightarrow \mu^+\mu^-$ final state. Accordingly, a TOS decision is required at HLT1 level of the `DiMuonHighMass` or `TrackMuon` line with respect to the $\psi(2S)$ candidate. Both of these lines require a positive decision of the level zero triggers

Table 6.1 – Stripping requirements for the $\psi(2S)$ and its decay products. The used variables are the difference of the logarithmic likelihood values of the muon and pion hypotheses, $\Delta \log \mathcal{L}_{\mu\pi}$, the reconstructed mass of the $\psi(2S)$ candidate, $m(\mu^+\mu^-)$, the χ^2 value of the distance of closest approach, χ_{DOCA}^2 , and the χ^2 value over the number of degrees of freedom of the $\psi(2S)$ vertex fit, $\chi_{\text{vtx}}^2/\text{ndf}$.

Observable	Requirement
$\Delta \ln \mathcal{L}_{\mu\pi}$	> 0
$ m(\mu^+\mu^-) - m_{\psi(2S)} $	$< 60 \text{ MeV}/c^2$
$\mu^+\mu^- \chi_{\text{DOCA}}^2$	< 30
$\psi(2S) \chi_{\text{vtx}}^2/\text{ndf}$	< 16

Table 6.2 – Stripping requirements for the K_S^0 and its decay products. The used variables are the momentum, p , the minimum impact parameter with respect to any other PV, $\text{MINIP}\chi^2$, the χ^2 value of the distance of closest approach, χ_{DOCA}^2 , the reconstructed mass of the K_S^0 candidate, $m(\pi^+\pi^-)$, the χ^2 value over the number of degrees of freedom of the vertex fit, $\chi_{\text{vtx}}^2/\text{ndf}$, and the decay time significance with respect to the associated best PV of the K_S^0 candidate t'/σ_t . Statements in parentheses refer to different requirements for DD reconstructed K_S^0 mesons.

Observable	Requirement
πp	$> 2 \text{ GeV}/c$
$\pi \text{ MINIP } \chi^2$	$> 9 (> 4)$
$\pi^+ \pi^- \chi_{\text{DOCA}}^2$	< 25
$ m(\pi^+\pi^-) - m_{K_S^0} $	$< 35 \text{ MeV}/c^2 \quad (< 64 \text{ MeV}/c^2)$
$K_S^0 \chi_{\text{vtx}}^2/\text{ndf}$	< 20
$K_S^0 t'/\sigma_t$	> 5

Table 6.3 – Stripping requirements for B^0 candidates. The variables are the reconstructed B^0 mass, $m(\mu^+\mu^-\pi^+\pi^-)$, and the χ^2 value over the number of degrees of freedom of the B^0 vertex fit, $\chi_{\text{vtx}}^2/\text{ndf}$.

Observable	Requirement
$m(\mu^+\mu^-\pi^+\pi^-)$	$5150 \text{ MeV}/c^2 < m(\mu^+\mu^-\pi^+\pi^-) < 5500 \text{ MeV}/c^2$
$B^0 \chi_{\text{vtx}}^2/\text{ndf}$	< 10

L0Muon or L0DiMuon, and the according tracks to be established as muon tracks through the requirements given in Tab. 6.4.

For the HLT2 it is required that candidates are TOS by the `DiMuonDetachedHeavy` line, which establishes a detached decay vertex of the dimuon system through the selection criteria given in Tab. 6.5. The signal efficiency of the trigger selection, evaluated with respect to the candidates that passed the stripping stage, is $\epsilon_{\text{trigger}}^{\psi(2S)} \approx 64.9\%$.

6.2 Multivariate selection and optimization

The stripping and trigger selection is followed by a multivariate selection, which uses a BDT to transform multiple observables into a single BDT prediction value. This prediction value is scanned to find the optimal working point for the measurement of the CP parameter S .

The BDT is trained using a signal sample of simulated $B^0 \rightarrow \psi(2S)K_S^0$ decays and a background sample consisting of candidates that are located in the upper mass-sideband of the stripping-selected data, defined by $m(\psi(2S)K_S^0) > 5450 \text{ MeV}/c^2$. In addition to the stripping and trigger requirements, the best PV association

Table 6.4 – Requirements of the used HLT1 lines that trigger on the dimuon system. The variables are the transverse momentum, p_T , the momentum, p , the χ^2 value over the number of degrees of freedom of the muon track fit, $\chi_{\text{track}}^2/\text{ndf}$, the distance of closest approach, DOCA, the χ^2 value of the dimuon vertex fit, χ_{vtx}^2 , the invariant dimuon mass, $m(\mu^+\mu^-)$, the impact parameter, IP, and the χ^2 value of the impact parameter, $\text{IP}\chi^2$.

Observable	Hlt1DiMuonHighMass	Hlt1TrackMuon
μp_T	$> 0.5 \text{ GeV}/c$	$> 1 \text{ GeV}/c$
μp	$> 6 \text{ GeV}/c$	$> 8 \text{ GeV}/c$
$\mu \chi_{\text{track}}^2/\text{ndf}$	< 4	< 2
$\mu \text{ DOCA}$	$< 0.2 \text{ mm}$	-
$\mu^+ \mu^- \chi_{\text{vtx}}^2$	< 25	-
$m(\mu^+\mu^-)$	$> 2.7 \text{ GeV}/c^2$	-
$\mu \text{ IP}$	-	$> 0.1 \text{ mm}$
$\mu \text{ IP } \chi^2$	-	> 16

Table 6.5 – Requirements of the used HLT2 line that triggers on the dimuon system. The variables are the χ^2 value over the number of degrees of freedom of the muon track fit, $\chi_{\text{track}}^2/\text{ndf}$, the χ^2 value of the dimuon vertex fit, χ_{vtx}^2 , the χ^2 of the flight distance of the dimuon system with respect to the PV, FD χ^2 , and the dimuon mass, $m(\mu^+\mu^-)$.

Observable	Hlt2DiMuonDetachedHeavy
$\mu \chi_{\text{track}}^2/\text{ndf}$	< 5
$\mu^+ \mu^- \chi_{\text{vtx}}^2$	< 25
$\mu^+ \mu^- \text{ FD } \chi^2$	> 25
$m(\mu^+\mu^-)$	$> 2.95 \text{ GeV}/c^2$

(smallest $\text{IP}\chi^2$ PV) is chosen for each event with multiple PVs for these training samples. The BDT input features comprise kinematic observables, and variables that are sensitive to the detached decay topology and the reconstruction quality: the minimal muon and pion impact parameters with respect to any PV, the $\psi(2S)$ and K_S^0 transverse momentum, the K_S^0 decay vertex χ^2 and decay-time significance, and the B^0 decay vertex χ^2 . It has been studied whether separate BDTs for DD and LL reconstructed K_S^0 candidates are needed, but no gain in performance could be found compared to a single BDT, whose working point is optimized individually on the two subsamples.

The working-point optimization is performed through fits to the reconstructed B^0 mass distribution in the stripping-and-trigger-selected data with a (temporary) best PV selection. The same model as in the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ case is used, consisting of an exponential function for the combinatorial background, and an Hypatia function for the signal B^0 and the irreducible B_s^0 component, as defined in Eq. (5.1). The signal shape is determined from simulated $B^0 \rightarrow \psi(2S)K_S^0$ signal decays, and has been observed to be widely independent of the BDT requirement.

6 $B^0 \rightarrow \psi(2S)K_S^0$ preparations

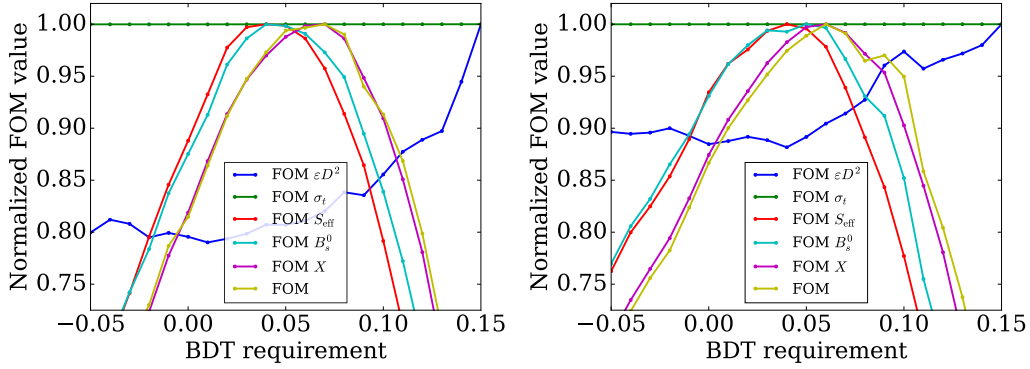


Figure 6.1 – Scan of the FOM normalized to the respective maximum value with respect to the BDT requirement for DD (left) and LL (right) reconstructed candidates. Different graphs correspond to the different components of the FOM, as defined in Eqs. (4.13) to (4.17).

Therefore, the tail parameters α_1 , α_2 , n_1 , n_2 , and $\beta = 0$ and $\zeta = 0$ are set constant in the scans of the FOM shown in Fig. 6.1. These scans result in an identical BDT prediction requirement of > 0.05 for DD and LL reconstructed candidates. The combined signal efficiency of the BDT selection is evaluated in the signal MC sample as $\epsilon_{\text{BDT}}^{\psi(2S)} \approx 86.2\%$, with respect to the number of candidates that pass the previous selection steps.

6.3 Exclusive backgrounds and multiple candidates

Based on studies in the previous $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ analysis by LHCb, the only exclusive background that is considered is coming from $\Lambda_b^0 \rightarrow \psi(2S)\Lambda$ decays, with a $p \rightarrow \pi$ mis-identification of the $\Lambda \rightarrow p\pi$ final state. This background is efficiently suppressed by using veto requirements for which the K_S^0 mass is recomputed under exchange of the π hypothesis with the proton hypothesis. If the recomputed mass lies within $10 \text{ MeV}/c^2$ ($6 \text{ MeV}/c^2$) around the known Λ mass [31] for DD (LL) reconstructed candidates, the candidate is rejected. This choice of requirements has a signal efficiency of $\epsilon_{\text{veto}}^{\psi(2S)} \approx 95.2\%$, while rejecting $\approx 93\%$ of Λ_b^0 decays, with respect to the candidates after the BDT selection. After the veto application the number of remaining $\Lambda_b^0 \rightarrow \psi(2S)\Lambda$ decays in the sample is reduced to a negligible amount of 46 ± 10 .

During the centralized stripping, the $B^0 \rightarrow \psi(2S)K_S^0$ data is saved to FullDST. This means that not only the reconstructed signal candidate with the best PV is available, but also all other reconstructed tracks and PVs in the event are stored. Accordingly it is possible to study the PV association and to discard events in which candidates are likely to originate from a different PV than the one they have been associated to. In simulated $B^0 \rightarrow \psi(2S)K_S^0$ signal decays it is observed that

candidates associated to the wrong PV cause a wide tail in the reconstructed decay-time distribution. A veto criterion has been developed which makes a requirement on the minimal IP χ^2 with respect to all other PVs within the event. If this value is below a certain threshold the candidate-PV pair is discarded. It is found that requiring IP $\chi^2 > 5$ is able to suppress the wrong PV component to low levels at a signal efficiency of $\epsilon_{\text{wPV}}^{\psi(2S)} \approx 98.5\%$.

After these requirements 0.4% of the events contain multiple candidates, i.e. candidates which are not unique within the event. One of these candidates is selected randomly, so that 10 844 candidates remain in the final data sample.

6.4 Results of the reconstructed-mass fit

After the full selection of the DD and LL reconstructed candidates, two separate fits to the reconstructed B^0 mass are performed, using the same model that has already been used in the selection optimization described in Sec. 6.2. In total, $7968 \pm 97 B^0$ decays and $79 \pm 19 B_s^0$ decays to the $\psi(2S) K_S^0$ final state are observed in the fits. A merged projection of the fit is shown in Fig. 6.2, and the detailed values can be taken from Tab. 6.6. The two fits are used to extract sWeights, which are used in the weighted CP fit, described in Ch. 7.

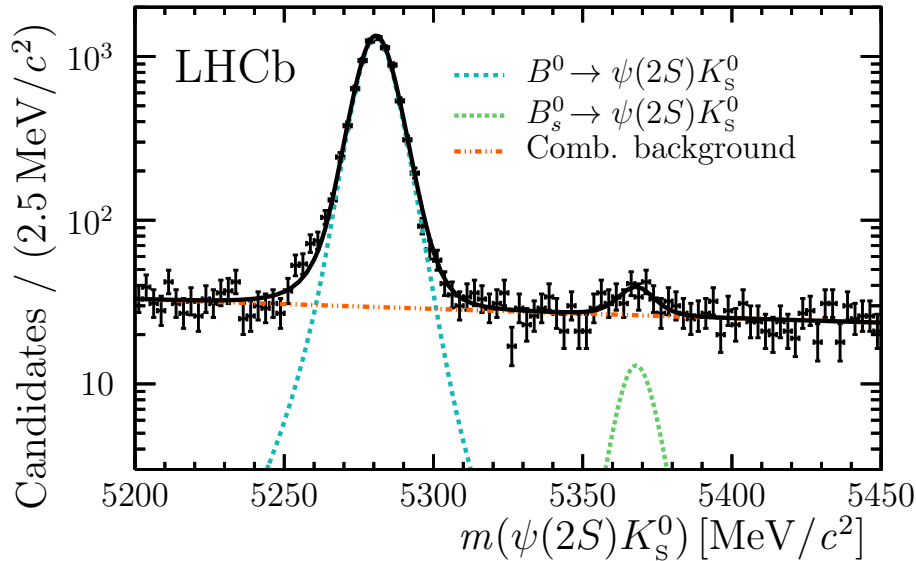


Figure 6.2 – Invariant mass of the $B^0 \rightarrow \psi(2S)K_S^0$ candidates [20]. The lines represent the result of the fit for the extraction of signal-decay weights that are used in the sFit for the CP -violation measurement.

Table 6.6 – Results of the fit to the reconstructed mass of $B^0 \rightarrow \psi(2S)K_S^0$ candidates after the full selection. The mass difference between the B^0 and the B_s^0 meson, $\mu_{B_s^0} - \mu_{B^0}$, is fixed to its known value [31]. The other parameters are the Hypatia shape parameters, $\alpha_1, \alpha_2, n_1, n_2, \beta, \zeta, \lambda$, and σ , the slope parameter of the exponential function for the combinatorial-background, $a_{\text{exp.}}$, and the yields of the B^0 component, n_{B^0} , of the B_s^0 component, $n_{B_s^0}$, and of the combinatorial-background component, $n_{\text{comb.}}$.

Parameters	DD	LL	
α_1	2.17	2.51	const.
α_2	3.09	4.43	const.
n_1	3.83	3.97	const.
n_2	3.73	2.52	const.
β	0.0	0.0	const.
ζ	0.0	0.0	const.
$\mu_{B_s^0} - \mu_{B^0}$ [MeV/ c^2]	87.2	87.2	const.
λ	-3.5 ± 0.7	-2.4 ± 0.5	floating
σ [MeV/ c^2]	7.36 ± 0.24	6.6 ± 0.4	floating
μ_{B^0} [MeV/ c^2]	5281.00 ± 0.10	5280.40 ± 0.12	floating
$a_{\text{exp.}}$ [1/(MeV/ c^2)]	-0.00086 ± 0.00034	-0.00256 ± 0.00054	floating
n_{B^0}	5342 ± 80	2626 ± 55	floating
$n_{B_s^0}$	56 ± 17	23 ± 9	floating
$n_{\text{comb.}}$	2035 ± 59	762 ± 36	floating

6.5 Production asymmetry

Values for the production asymmetry are calculated in a fully analogous way to the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode (see Sec. 5.3). Using the LHCb measurement in Ref. [127] the following values are computed for the $B^0 \rightarrow \psi(2S)K_S^0$ mode:

$$A_{\text{P}}^{11} = -0.0143 \pm 0.0077 \text{ (stat.)} \pm 0.0005 \text{ (syst.)}, \quad (6.1)$$

$$A_{\text{P}}^{12} = -0.0138 \pm 0.0051 \text{ (stat.)} \pm 0.0003 \text{ (syst.)}. \quad (6.2)$$

These values are used in the CP fit, where the statistical uncertainties are propagated using Gaussian constraints (see Sec. 7.2) and the systematic uncertainties are covered in the evaluation of the systematic uncertainties (see Sec. 7.5.4).

6.6 Correction of detector effects

Studies on the decay-time resolution and acceptance have been performed using simulated $B^0 \rightarrow \psi(2S)K_S^0$ decays, analogously to the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ approach, presented in Sec. 5.4. The decay-time resolution is described using the model, defined in Eq. (5.6). Again, the deviation of the decay-time from its true value, $t' - t$, is studied in the simulated signal sample. As a first step, the linear relation between the $t' - t$ width and the σ_t prediction is established, using a discrete version

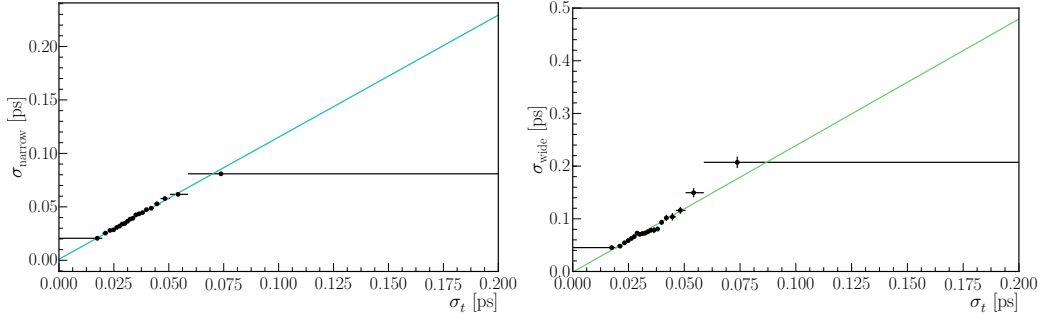


Figure 6.3 – Linear calibration functions for the scaling of the decay-time uncertainty estimate σ_t from a simultaneous fit to the observed $t' - t$ distribution in bins of σ_t using simulated signal decays. Horizontal errorbars indicate the chosen bin widths in σ_t , while the vertical errorbars are the fit uncertainties for the narrow (left) and the wide (right) Gaussian component of the resolution model.

Table 6.7 – Parameters of the $B^0 \rightarrow \psi(2S)K_S^0$ decay-time-resolution model determined from a fit to simulated decays. The fraction of the narrow Gaussian is implicitly defined through $f_{\text{narrow}} = 1 - f_{\text{wpv}} - f_{\text{wide}}$. The parameters μ and b_{wide} were found to be compatible with zero and have been fixed.

Parameter	Fit result
μ [fs]	0.0
f_{wide}	0.058 ± 0.006
f_{wpv}	0.0039 ± 0.0002
a_{narrow}	1.162 ± 0.011
b_{narrow} [ps]	0.0007 ± 0.0003
a_{wide}	2.481 ± 0.079
b_{wide} [ps]	0.0
σ_{wpv} [ps]	10.044 ± 0.363

of the resolution model in bins of σ_t , as presented in Fig. 6.3. The full calibration of the resolution model is obtained from fits to the simulated signal sample, whose projections are shown in Fig. 6.4 and whose values are given in Tab. 6.7. The effective single Gaussian resolution for correctly associated $B^0 \rightarrow \psi(2S)K_S^0$ decays is 48 fs. For the decay-time dependent efficiency function cubic B-splines are used. Studies using the simulated signal sample have shown that interval boundaries placed at $\{0.2, 0.5, 1.0, 5.0, 15.0\}$ ps yield a satisfying description while minimizing the number of coefficients needed. Projections of fits to the decay-time distribution in the simulated sample and of the CP fit to data (see Ch. 7) are given in Fig. 6.5, together with the respective acceptance functions. For the fits to the simulated sample the B^0 lifetime is constrained to the generation value. The corresponding values for the coefficients are given in Tab. 6.8.

6 $B^0 \rightarrow \psi(2S)K_S^0$ preparations

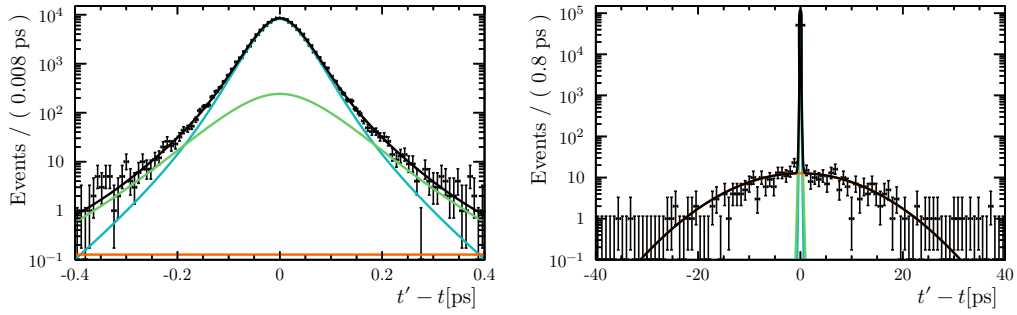


Figure 6.4 – Projection of the resolution model that is fit to simulated $B^0 \rightarrow \psi(2S)K_S^0$ decays in a narrow (left) and wide (range) of the decay-time residual, $t' - t$. The total model (black) consists of a narrow (blue) and a wide (green) Gaussian component that use linear functions of the decay-time uncertainty estimate as widths, and a wide average Gaussian component that accounts for wrong PV associations (orange).

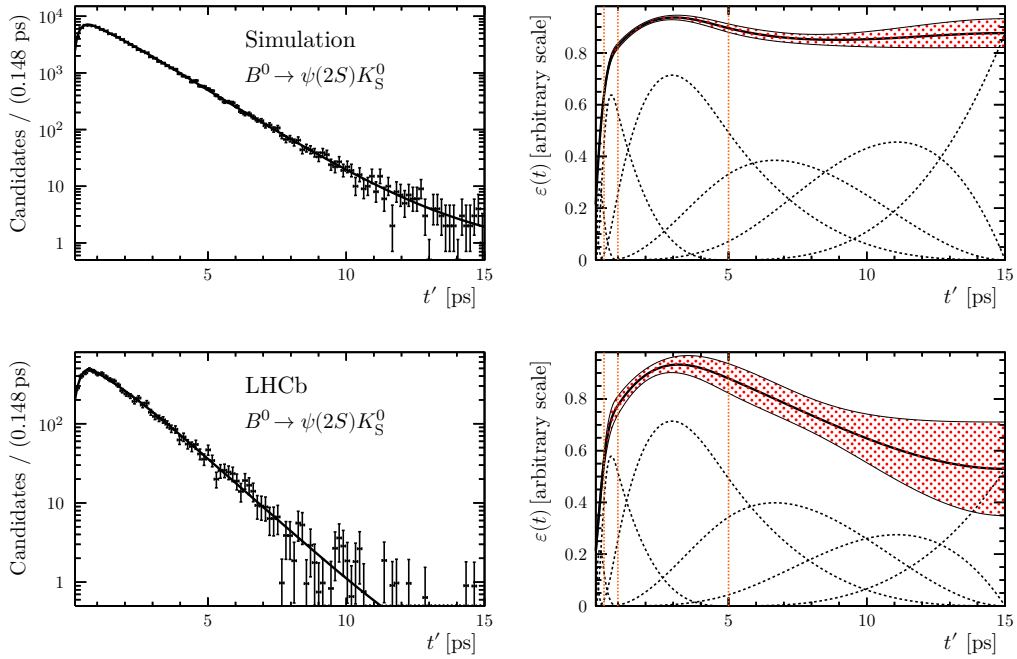


Figure 6.5 – Projections of the decay-time fit (left) and the respective acceptance function (right) for the simulated signal sample (top) and data (bottom) of $B^0 \rightarrow \psi(2S)K_S^0$ decays. The red-dotted region corresponds to the uncertainty of the acceptance description coming from the fit, the dashed vertical lines represent the inner interval boundaries, and the dashed curves the base splines.

Table 6.8 – Decay-time-acceptance parameters for $B^0 \rightarrow \psi(2S)K_S^0$ decays in the simulated signal sample and data. The coefficient c_3 is set to a constant value of 1 to fix the arbitrary scale of the efficiency function.

Parameter	Simulation	Data
c_0	0.215 ± 0.011	0.21 ± 0.04
c_1	0.422 ± 0.010	0.334 ± 0.034
c_2	0.768 ± 0.014	0.70 ± 0.05
c_3	1	1
c_4	0.78 ± 0.04	0.81 ± 0.15
c_5	0.88 ± 0.06	0.53 ± 0.18

6.7 Flavour tagging

For the flavour-tagging calibration of the $B^0 \rightarrow \psi(2S)K_S^0$ mode, flavour-specific decays of $B^0 \rightarrow J/\psi(\mu^+\mu^-)K^*$ and self-tagging decays of $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ are studied for the SS and OS algorithms, respectively. These two modes are preferred instead of the corresponding $\psi(2S)$ control modes $B^0 \rightarrow \psi(2S)(\mu^+\mu^-)K^*$ and $B^+ \rightarrow \psi(2S)(\mu^+\mu^-)K^+$, due to the higher branching fraction and associated number of candidates available, which directly result in smaller calibration uncertainties.

For both modes a cut-based selection is used, which makes requirements on the reconstruction quality and exploits the detached decay topology. After the selection, the samples are reweighted to the signal kinematics using Gradient Boosting Reweighting [132]. A preliminary calibration leads to thresholds for the combined OS and SS mistag predictions of

$$\omega_{\text{OS}} \leq 0.476, \quad (6.3)$$

$$\omega_{\text{SS}} \leq 0.472, \quad (6.4)$$

which are required to avoid fit stability problems (see Sec. 5.5.2). Candidates that do not satisfy these requirements are set to an untagged state, i.e. $d'_{\text{OS}} = d'_{\text{SS}} = 0$, and $\omega_{\text{OS}} = \omega_{\text{SS}} = 0.5$. Lastly, the final recalibrations of the combined OS and SS taggers are performed, which are shown in Fig. 6.6. The analysis of 1231000 ± 2000 $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ decays results in flavour-tagging calibration parameters of

$$\begin{aligned}
p_{0,\text{OS}}^{\psi(2S)} &= 0.3725 \pm 0.0012 \text{ (stat.)} \pm 0.00005 \text{ (syst.)}, \\
p_{1,\text{OS}}^{\psi(2S)} &= 0.815 \pm 0.011 \text{ (stat.)} \pm 0.005 \text{ (syst.)}, \\
\Delta p_{0,\text{OS}}^{\psi(2S)} &= 0.0093 \pm 0.0024 \text{ (stat.)} \pm 0.00007 \text{ (syst.)}, \\
\Delta p_{1,\text{OS}}^{\psi(2S)} &= 0.022 \pm 0.023 \text{ (stat.)} \pm 0.002 \text{ (syst.)}, \\
\langle \eta_{\text{OS}}^{\psi(2S)} \rangle &= 0.3387,
\end{aligned} \quad (6.5)$$

6 $B^0 \rightarrow \psi(2S)K_S^0$ preparations

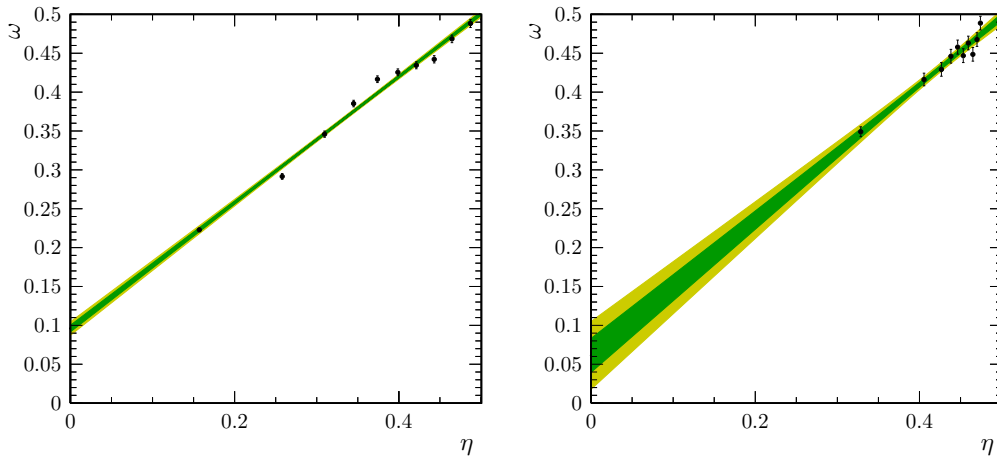


Figure 6.6 – Flavour-tagging calibration using reweighted $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ decays for the OS-combination tagger (left) and reweighted $B^0 \rightarrow J/\psi(\mu^+\mu^-)K^*$ decays for the SS-combination tagger (right).

for the combined tagging decision of the OS tagging algorithms. Furthermore, the analysis of $430800 \pm 800 B^0 \rightarrow J/\psi(\mu^+\mu^-)K^*$ decays results in the calibration parameters

$$\begin{aligned}
 p_{0,SS}^{\psi(2S)} &= 0.439 \pm 0.003 \text{ (stat.)} \pm 0.002 \text{ (syst.)}, \\
 p_{1,SS}^{\psi(2S)} &= 0.870 \pm 0.051 \text{ (stat.)} \pm 0.002 \text{ (syst.)}, \\
 \Delta p_{0,SS}^{\psi(2S)} &= 0.0123 \pm 0.0040 \text{ (stat.)} \pm 0.0007 \text{ (syst.)}, \\
 \Delta p_{1,SS}^{\psi(2S)} &= 0.054 \pm 0.074 \text{ (stat.)} \pm 0.012 \text{ (syst.)}. \\
 \langle \eta_{SS}^{\psi(2S)} \rangle &= 0.4343,
 \end{aligned} \tag{6.6}$$

for the combined tagging decision of the SS algorithms. These calibration parameters are used in the decay-time-dependent CP fit, where their statistical uncertainties, as well as the linear correlation coefficients that are given in Tab. 6.9 are used as multidimensional Gaussian constraints.

The effective tagging efficiency for the combination of the OS and SS taggers in the $B^0 \rightarrow \psi(2S)K_S^0$ data sample is calculated, following Eq. (4.1), as

$$\varepsilon_{\text{eff}}^{B^0 \rightarrow \psi(2S)K_S^0} = (3.42 \pm 0.09) \%, \tag{6.7}$$

where a sole use of either OS or SS algorithms would contribute with

$$\varepsilon_{\text{eff, OS}}^{B^0 \rightarrow \psi(2S)K_S^0} = (2.46 \pm 0.05) \%, \tag{6.8}$$

$$\varepsilon_{\text{eff, SS}}^{B^0 \rightarrow \psi(2S)K_S^0} = (1.07 \pm 0.08) \%, \tag{6.9}$$

Table 6.9 – Linear correlation coefficients between the flavour-tagging parameters for the $B^0 \rightarrow \psi(2S)K_S^0$ mode.

	OS correlations				SS correlations			
	p_0	p_1	Δp_0	Δp_1	p_0	p_1	Δp_0	Δp_1
p_0	1.000	0.121	-0.003	-0.005	1.000	0.163	-0.009	-0.005
p_1	0.121	1.000	-0.005	0.006	0.163	1.000	-0.005	-0.015
Δp_0	-0.003	-0.005	1.000	0.121	-0.009	-0.005	1.000	0.122
Δp_1	-0.005	0.006	0.121	1.000	-0.005	-0.015	0.122	1.000

on the full data set.

The performance of the single taggers has been evaluated on the corresponding control modes. The highest contribution for the OS algorithms is obtained from the OS vertex charge (1.094%), the OS muon (1.048%), and the OS kaon (1.002%) taggers. Lower tagging powers are obtained from the OS charm (0.375%) and the OS electron (0.339%) taggers. For the SS algorithms, the SS pion tagger contributes significantly more (0.957%) than the SS proton tagger (0.339%).

Compared to the tagging power of the previous LHCb analysis (see Eq. (5.21)), a relative increase in total tagging power of 13% is present, thanks to improvements made for the SS pion tagger that is used, and the additional use of the SS proton tagger.

7 Measurement of $\sin(2\beta)$

This chapter presents the technique and the results of the weighted, decay-time-dependent CP fit to the prepared $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ datasets. Furthermore, the model validation is elaborated on, followed by an estimation of systematic uncertainties coming from model assumptions, as well as systematic uncertainties on input parameters. The chapter closes with a section on corrections that need to be made considering the effect of K^0 - \bar{K}^0 regeneration.

7.1 CP -fit model

A model is derived from Eqs. (2.37) and (2.38), which can describe the relative likelihoods of candidates' tags, \vec{d}' , and reconstructed decay-times, t' , given their mistag probabilities, $\vec{\omega}$, and their decay-time-uncertainty predictions, σ_t . This model is used to measure the CP parameters in the context of a simultaneous maximum-likelihood parameter estimation, while taking into account the production asymmetry between B^0 and \bar{B}^0 mesons.

The production asymmetry can be implemented easily by using factors, $1 \pm A_P$, which scale the decay rate for initial B^0 , $\mathcal{P}(t, B^0)$, compared to the decay rate for initial \bar{B}^0 , $\mathcal{P}(t, \bar{B}^0)$. Assuming no CP violation in the mixing ($|q/p| = 1$), and omitting identical constant terms the theoretical decay rates are then rewritten to

$$\mathcal{P}(t, B^0) = e^{-\Gamma t} (1 - A_P) \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D \sinh\left(\frac{\Delta\Gamma}{2}t\right) + C \cos(\Delta mt) - S \sin(\Delta mt) \right], \quad (7.1)$$

$$\mathcal{P}(t, \bar{B}^0) = e^{-\Gamma t} (1 + A_P) \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D \sinh\left(\frac{\Delta\Gamma}{2}t\right) - C \cos(\Delta mt) + S \sin(\Delta mt) \right]. \quad (7.2)$$

For the distribution of tagging decisions, a Bernoulli probability process for the individual taggers is considered: By definition, a correct decision of the individual tagger i ($d'_i = d$) is obtained at probability $1 - \omega_i$, whereas a wrong decision ($d'_i = -d$) is obtained at probability ω_i . Furthermore, a tag decision is only obtained at a probability corresponding to the individual tagging efficiency $\varepsilon_{\text{tag},i}$, whereas no decision can be made at $(1 - \varepsilon_{\text{tag},i})$. Assuming no correlations between the N taggers, the individual probabilities factorize. Given a true production flavour d , this results in the probability to observe the tagging decisions,

$$\mathcal{P}(\vec{d}'|d) = \prod_{i=1}^N \delta_{|d'_i|,1} \varepsilon_{\text{tag},i} \frac{1 + dd'_i(1 - 2\omega_i)}{2} + \delta_{d'_i,0} (1 - \varepsilon_{\text{tag},i}). \quad (7.3)$$

7 Measurement of $\sin(2\beta)$

Given an initially produced B^0 meson, the tags for an OS and an SS tagger, $\vec{d}' = [d'_{\text{OS}}, d'_{\text{SS}}]$, are then distributed according to

$$\begin{aligned}
\mathcal{P}(\vec{d}'|B^0) &= \delta_{|d'_{\text{OS}}|,1} \delta_{|d'_{\text{SS}}|,1} \varepsilon_{\text{tag,OS}} \varepsilon_{\text{tag,SS}} \frac{1 + d'_{\text{OS}}(1 - 2\omega_{\text{OS}}^{B^0})}{2} \frac{1 + d'_{\text{SS}}(1 - 2\omega_{\text{SS}}^{B^0})}{2} \\
&+ \delta_{|d'_{\text{OS}}|,1} \delta_{d'_{\text{SS}},0} \varepsilon_{\text{tag,OS}} (1 - \varepsilon_{\text{tag,SS}}) \frac{1 + d'_{\text{OS}}(1 - 2\omega_{\text{OS}}^{B^0})}{2} \\
&+ \delta_{|d'_{\text{SS}}|,1} \delta_{d'_{\text{OS}},0} \varepsilon_{\text{tag,SS}} (1 - \varepsilon_{\text{tag,OS}}) \frac{1 + d'_{\text{SS}}(1 - 2\omega_{\text{SS}}^{B^0})}{2} \\
&+ \delta_{d'_{\text{OS}},0} \delta_{d'_{\text{SS}},0} (1 - \varepsilon_{\text{tag,OS}}) (1 - \varepsilon_{\text{tag,SS}}). \tag{7.4}
\end{aligned}$$

The corresponding distribution for an initially produced \bar{B}^0 meson, $\mathcal{P}(\vec{d}'|\bar{B}^0)$, is obtained by replacing $\omega_{\text{OS}}^{B^0} \rightarrow \omega_{\text{OS}}^{\bar{B}^0}$ and $\omega_{\text{SS}}^{B^0} \rightarrow \omega_{\text{SS}}^{\bar{B}^0}$ in this expression. When defining the sum and difference of these probabilities for the tagging decisions as

$$\Delta_{\text{tag}} = \mathcal{P}(\vec{d}'|B^0) - \mathcal{P}(\vec{d}'|\bar{B}^0), \tag{7.5}$$

$$\Sigma_{\text{tag}} = \mathcal{P}(\vec{d}'|B^0) + \mathcal{P}(\vec{d}'|\bar{B}^0), \tag{7.6}$$

the full PDF describing the tagging decisions and the theoretical (true) decay times of signal mesons can be written as

$$\begin{aligned}
\mathcal{P}(t, \vec{d}') &= \mathcal{P}(\vec{d}'|B^0) \mathcal{P}(t, B^0) + \mathcal{P}(\vec{d}'|\bar{B}^0) \mathcal{P}(t, \bar{B}^0) \\
&= e^{-\Gamma t} \left[(\Sigma_{\text{tag}} - A_{\text{P}} \Delta_{\text{tag}}) \cosh\left(\frac{\Delta\Gamma}{2} t\right) + (\Sigma_{\text{tag}} - A_{\text{P}} \Delta_{\text{tag}}) D \sinh\left(\frac{\Delta\Gamma}{2} t\right) \right. \\
&\quad \left. + (\Delta_{\text{tag}} - A_{\text{P}} \Sigma_{\text{tag}}) C \cos(\Delta m t) - (\Delta_{\text{tag}} - A_{\text{P}} \Sigma_{\text{tag}}) S \sin(\Delta m t) \right]. \tag{7.7}
\end{aligned}$$

As the assumption that $\Delta\Gamma = 0$ is used, the model simplifies further to

$$\mathcal{P}(t, \vec{d}'|\vec{\omega}) = e^{-\Gamma t} \left[(\Sigma_{\text{tag}} - A_{\text{P}} \Delta_{\text{tag}}) + (\Delta_{\text{tag}} - A_{\text{P}} \Sigma_{\text{tag}}) (C \cos(\Delta m t) - S \sin(\Delta m t)) \right]. \tag{7.8}$$

The functions Δ_{tag} and Σ_{tag} depend on the tagging-calibration parameters, as the mistag probabilities $\vec{\omega}$ are obtained from the observed estimates $\vec{\eta}$, so that $\vec{\omega} = \vec{\omega}(\vec{\eta}|\vec{p}_0, \vec{p}_1)$ (see Secs. 5.5 and 6.7). As different flavour-tagging calibrations need to be performed for the two decay modes, distinct model parameters need to be employed. The same is also valid for the production-asymmetry parameters A_{P} , which do not only depend on the decay channel, but also on the year of data-taking period (see Secs. 5.3 and 6.5). Therefore, the fit is performed simultaneously to the decay mode and year subsets, denoted by $\text{ch} \in \{J/\psi, \psi(2S)\}$ and $\text{yr} \in \{11, 12\}$. The connection between the true decay-time, t , and the reconstructed decay-time, t' , is made using a convolution with the resolution model, and subsequent multiplication with the acceptance function, as discussed in Sec. 5.4. As both decay modes show

different decay-time resolutions and acceptances, the full simultaneous PDF is given by

$$\mathcal{P}^{\text{ch,yr}}(t', \vec{d}' | \sigma_t, \vec{\omega}) = \epsilon^{\text{ch}}(t') \cdot (\mathcal{P}^{\text{ch,yr}}(t, \vec{d}' | \vec{\omega}) \otimes \mathcal{R}^{\text{ch}}(t' - t | \sigma_t)) . \quad (7.9)$$

7.2 Observables and parameters

The observables that are used in the fit are listed in Tab. 7.1. Limited ranges of these observables ensure the validity of the fit model over the full interval.

Table 7.1 – Observables used in the CP fit.

Observable	Range
Decay time t'	[0.2, 15] ps
Decay time uncertainty estimate σ_t	[0, 400] fs
Flavour tag OS d'_{OS}	{-1, 0, 1}
Flavour tag SS d'_{SS}	{-1, 0, 1}
Mistag estimate OS η_{OS}	[0.0, 0.5]
Mistag estimate SS η_{SS}	[0.0, 0.5]

Parameters are distinguished by whether they are shared or split between the year of data-taking and decay-mode subsets, and whether they are free, fixed or constrained in the fit. In order to constrain a parameter to externally measured values, the likelihood function defined in Eq. (4.7) is multiplied by a multivariate normal distribution, whose mean vector is fixed to the measured values, and whose covariance matrix contains the uncertainty and correlation information.

All parameters except for the parameters of the B^0 - \bar{B}^0 system, i.e. the mass difference, Δm , and the lifetime, τ , are split between the two decay modes. Only the production-asymmetry parameters are additionally split between the year of data-taking.

The only fixed parameters in the fit are the ones from the resolution models. The production-asymmetry and tagging parameters are constrained, except for the individual tagging efficiencies, which are free in the fit. The acceptance parameters, as well as the CP coefficients are also free fit parameters.

7.3 Blinding

To guarantee that the analysis is performed without a bias from the experimenter, a self-imposed blinding transformation is applied to the CP parameters, adding a hidden offset to these parameters. The uncertainty on the extracted parameters remains unchanged and thus can still be used for inspecting the validity and sensitivity of the fit. The offset to obscure is sampled from a uniform distribution

with range $\pm s$, using a random number generator that has a fixed seed. The latter in turn is generated from a so-called blinding string which can be passed to analysts measuring the same physics parameters, allowing for a comparison of blinded results. The range of the uniform distribution is $[-2, 2]$, which, given the physical range of $\sin(2\beta)$ of $[-1, 1]$ ensures good opacity and provides sufficient ambiguity. As the results of the analysis have been published, all results shown in this thesis are unblinded.

7.4 Results

The simultaneous maximum-likelihood sFit of the weighted $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ datasets results in the CP parameters

$$\begin{aligned} C_{\text{fit}}^{J/\psi} &= 0.125 \pm 0.075, \\ S_{\text{fit}}^{J/\psi} &= 0.824 \pm 0.076, \\ C_{\text{fit}}^{\psi(2S)} &= -0.044 \pm 0.096, \\ S_{\text{fit}}^{\psi(2S)} &= 0.842 \pm 0.103, \end{aligned} \tag{7.10}$$

with linear statistical correlation coefficients between S and C of 0.46 and 0.48 for the J/ψ and the $\psi(2S)$ mode, respectively. The subscript "fit" is used in Eq. (7.10) as the fitted parameters are corrected for the effect of K^0 - \bar{K}^0 regeneration in Sec. 7.6. The signal-yield asymmetries $(N_{\bar{B}^0} - N_{B^0}) / (N_{\bar{B}^0} + N_{B^0})$ as a function of the decay time are shown in Fig. 7.1. The detailed results of the fit parameters are given in Tab. 7.2 and 7.3, where it can be seen that all final values for the constrained parameters are fully compatible with the central value of the constraint within one standard deviation.

The full correlation matrix of the fit is given in Fig. 7.2. Besides the large correlation intrinsic correlation between S and C that can be derived from first principles [134], smaller correlations up to $\approx 10\%$ – 20% are present between the two CP parameters, and the flavour-tagging-calibration parameters for each mode. The correlations between the production-asymmetry parameters and the CP parameters, as well as between Δm and the CP parameters, are below 5%. The parameters of the acceptance description are highly correlated among each other, but not to the parameters of interest.

The uncertainties in Eq. (7.10) are fit uncertainties and therefore of statistical origin. Studies on the systematic uncertainties are presented in Sec. 7.5.4.

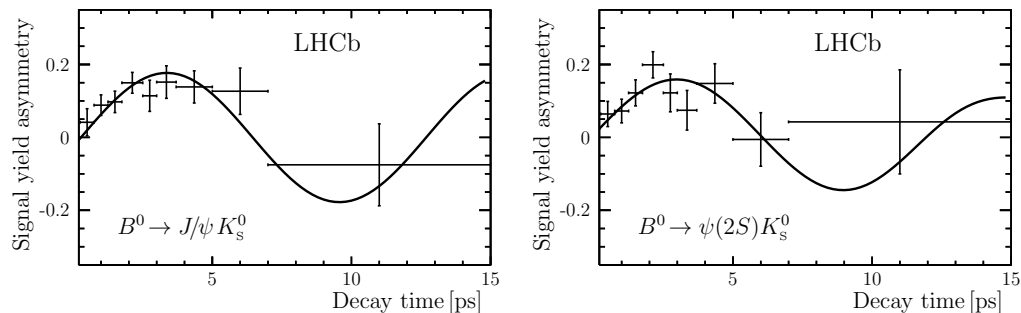


Figure 7.1 – Signal yield asymmetries $(N_{\bar{B}^0} - N_{B^0}) / (N_{\bar{B}^0} + N_{B^0})$ as a function of decay time for $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays (left) and $B^0 \rightarrow \psi(2S)K_S^0$ decays (right) [20]. The symbol N_{B^0} ($N_{\bar{B}^0}$) denotes the number of decays with a B^0 (\bar{B}^0) flavour tag. The solid curves are the projections of a PDF that uses a combined flavour tagging decision.

Table 7.2 – Results of the CP parameters, C and S , for each decay mode, of the constrained production-asymmetry parameters, A_P , for each decay mode and year of data-taking, and of the shared and constrained parameters of the B^0 - \bar{B}^0 system, Δm and τ . The decay-width difference, $\Delta\Gamma$, is fixed to zero. The results of the constrained parameters are fully compatible with the central values of the constraints.

Parameter	Fit result	Constraint
$C^{J/\psi}$	0.125 ± 0.075	floating
$C^{\psi(2S)}$	-0.044 ± 0.096	floating
$S^{J/\psi}$	0.824 ± 0.076	floating
$S^{\psi(2S)}$	0.842 ± 0.103	floating
$A_P^{J/\psi,11}$	-0.0102 ± 0.0083	-0.0101 ± 0.0084
$A_P^{\psi(2S),11}$	-0.0142 ± 0.0076	-0.0143 ± 0.0077
$A_P^{J/\psi,12}$	-0.0078 ± 0.0054	-0.0077 ± 0.0054
$A_P^{\psi(2S),12}$	-0.0140 ± 0.0051	-0.0138 ± 0.0051
Δm [ps^{-1}]	0.5065 ± 0.0016	0.5065 ± 0.0016
$\Delta\Gamma$ [keV]	0	0
τ [ps]	1.5199 ± 0.0040	1.520 ± 0.004

Table 7.3 – Results of the constrained flavour-tagging parameters, p_0 , p_1 , Δp_0 , Δp_1 , as well as of the free flavour-tagging efficiencies, ε_{tag} , for each tagging algorithm and decay mode. The average η values are fixed to the values obtained from the calibration modes. The results of the constrained parameters are fully compatible with the central values of the constraints.

Parameter	Fit result	Constraint
$p_{0,\text{OS}}^{J/\psi}$	0.3605 ± 0.0030	0.3603 ± 0.0031
$p_{1,\text{OS}}^{J/\psi}$	0.838 ± 0.028	0.834 ± 0.029
$\Delta p_{0,\text{OS}}^{J/\psi}$	0.0169 ± 0.0059	0.0162 ± 0.0061
$\Delta p_{1,\text{OS}}^{J/\psi}$	-0.033 ± 0.055	-0.024 ± 0.057
$\langle \eta_{\text{OS}}^{J/\psi} \rangle$	0.316	0.316
$\varepsilon_{\text{tag,OS}}^{J/\psi}$	0.3329 ± 0.0054	floating
$p_{0,\text{SS}}^{J/\psi}$	0.4203 ± 0.0059	0.4213 ± 0.0063
$p_{1,\text{SS}}^{J/\psi}$	1.33 ± 0.11	1.29 ± 0.12
$\Delta p_{0,\text{SS}}^{J/\psi}$	-0.0011 ± 0.0077	0.0026 ± 0.0086
$\Delta p_{1,\text{SS}}^{J/\psi}$	-0.06 ± 0.16	0.04 ± 0.18
$\langle \eta_{\text{SS}}^{J/\psi} \rangle$	0.425	0.425
$\varepsilon_{\text{tag,SS}}^{J/\psi}$	0.5782 ± 0.0057	floating
$p_{0,\text{OS}}^{\psi(2S)}$	0.3725 ± 0.0012	0.3725 ± 0.0012
$p_{1,\text{OS}}^{\psi(2S)}$	0.816 ± 0.011	0.815 ± 0.011
$\Delta p_{0,\text{OS}}^{\psi(2S)}$	0.0092 ± 0.0024	0.0093 ± 0.0024
$\Delta p_{1,\text{OS}}^{\psi(2S)}$	0.021 ± 0.023	0.022 ± 0.023
$\langle \eta_{\text{OS}}^{\psi(2S)} \rangle$	0.3387	0.3387
$\varepsilon_{\text{tag,OS}}^{\psi(2S)}$	0.2979 ± 0.0053	floating
$p_{0,\text{SS}}^{\psi(2S)}$	0.4390 ± 0.0028	0.4392 ± 0.0029
$p_{1,\text{SS}}^{\psi(2S)}$	0.862 ± 0.050	0.870 ± 0.051
$\Delta p_{0,\text{SS}}^{\psi(2S)}$	0.0129 ± 0.0039	0.0123 ± 0.0040
$\Delta p_{1,\text{SS}}^{\psi(2S)}$	0.057 ± 0.072	0.054 ± 0.074
$\langle \eta_{\text{SS}}^{\psi(2S)} \rangle$	0.4343	0.4343
$\varepsilon_{\text{tag,SS}}^{\psi(2S)}$	0.4904 ± 0.0058	floating

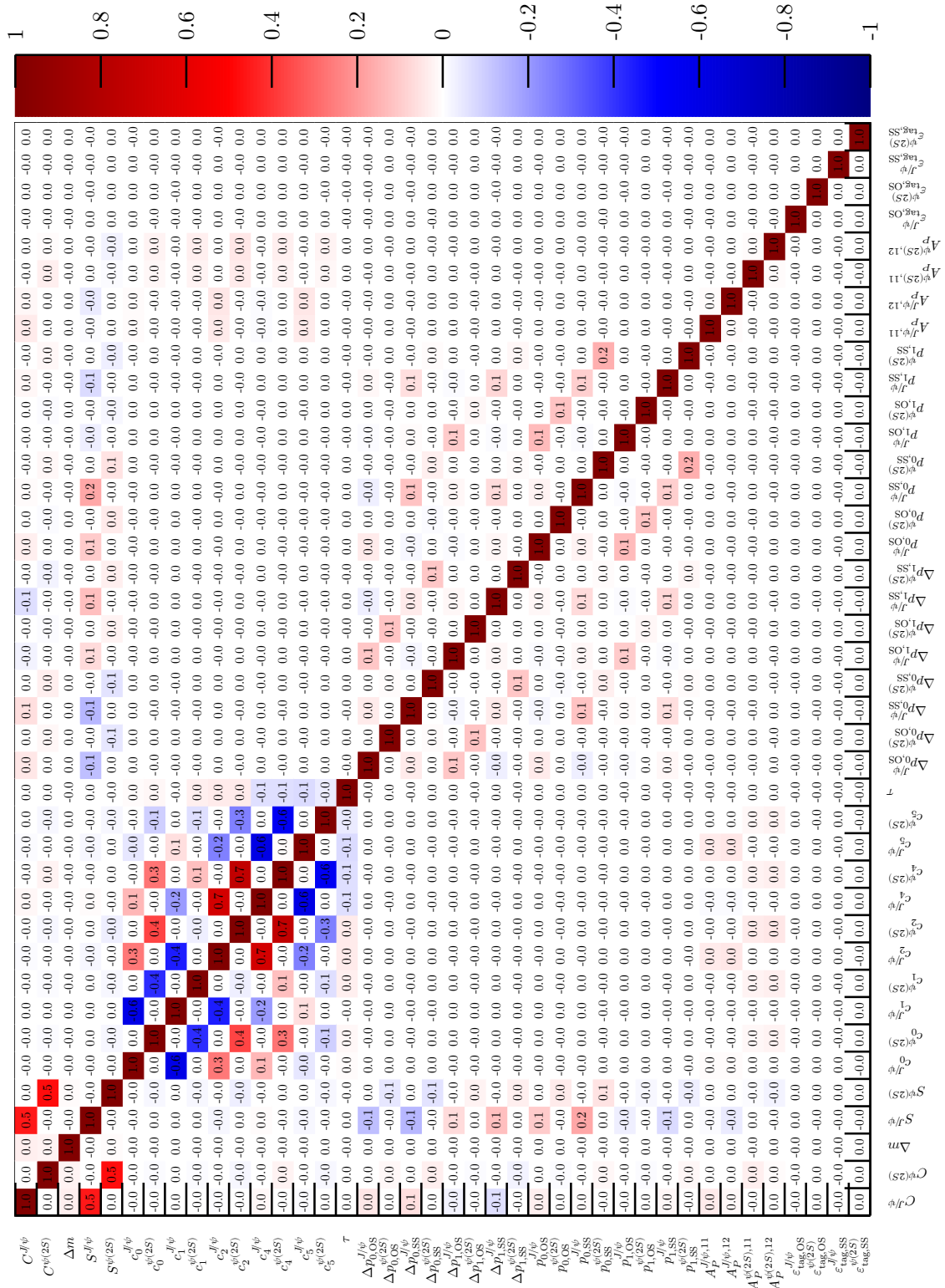


Figure 7.2 – Linear correlation coefficients between the fitted parameters. The color scale ranges from negative correlations of -1 in blue to positive correlations of $+1$ in red.

7.5 Validation and systematic uncertainties

The validation of the CP fit and the assessment of systematic uncertainties are both performed using pseudo-datasets. In contrast to a full physics simulation, where the sample preparation takes large computing resources, the observables contained in pseudo-datasets are sampled directly from PDFs. This allows to study $> \mathcal{O}(10^3)$ of these datasets in a variety of scenarios, to test whether the fitted parameters on average result in the expected values and uncertainties. While Sec. 7.5.1 elaborates on how biases are recognized and treated to derive systematic uncertainties, Sec. 7.5.2 presents the model, that is used to generate the pseudo-datasets. The results of the fit validation are then presented in Sec. 7.5.3 and the studies of systematic uncertainties are elaborated on in Sec. 7.5.4. Finally, other cross-checks of the CP fit are presented in Sec. 7.5.5.

7.5.1 Bias recognition and treatment

The difference between the observed and the expected parameter value, relative to its uncertainty, is defined as the pull,

$$p(S) = \frac{S_{\text{obs}} - S_{\text{exp}}}{\sigma_{S_{\text{obs}}}}. \quad (7.11)$$

If the measurement of S is unbiased, the pull distribution, which would be obtained after repeating the experiment N times, will show a mean value, $\mu(p)$, that is compatible with zero. If furthermore the uncertainty of S is correctly estimated, the standard deviation of the pull distribution, $\sigma(p)$, will be compatible with one. Therefore, a deviation of $\mu(p)$ from zero reflects a biased measurement, and a deviation of $\sigma(p)$ from one an incorrect uncertainty estimation of S , where both these deviations are given in units of the average uncertainty estimate of S .

Both points, deviations of $\mu(p)$ from zero and/or deviations of $\sigma(p)$ from one, are important to consider for the assignment of systematic uncertainties. However, a conservative approach is chosen in that only a possible underestimation of the uncertainty of the parameter is considered, since overestimation would require to decrease the overall uncertainty. Thus, a systematic uncertainty in units of the fit uncertainty is obtained as

$$s = \begin{cases} \mu(p) & \text{for } \sigma(p) \leq 1 \\ \sqrt{\mu(p)^2 + [\sigma(p) - 1]^2} & \text{for } \sigma(p) > 1 \end{cases}, \quad (7.12)$$

where the effects are summed in square for cases that show a bias as well as an uncertainty underestimation. Both, $\mu(p)$ and $\sigma(p)$, can only be estimated, due to the limited number of pseudo-datasets that is studied. For all studies using pseudo-datasets, the number of these datasets is chosen so that the uncertainties on $\mu(p)$ and $\sigma(p)$ are 2%–3%. Correspondingly, the systematic deviations according

to Eq. (7.12), can be evaluated at a precision of 3%–4% with respect to the statistical uncertainty of the fit.

For the fitter validation (see Sec. 7.5.3) it has been decided a priori to only assign a fit bias uncertainty, if deviations of $\mu(p(C)) / \mu(p(S))$ or $\sigma(p(C)) / \sigma(p(S))$ from 0 or 1 occur at a significance greater than two standard deviations. In the case of the systematic uncertainties (see Sec. 7.5.4) this criterion has not led to any systematics that need to be assigned. Therefore, it has been decided to drop the significance criterion and consider any deviation from the expected values as a systematic uncertainty. The fact that this might artificially enhance the magnitude of the systematic uncertainty is accepted, as the statistical uncertainty dominates the total uncertainty of the measurement, and therefore the overall coverage of the uncertainty is fine. To derive the absolute systematic uncertainty, Eq. (7.12) is scaled with the statistical uncertainty that is observed in the fit for each studied effect:

$$\sigma_{\text{sys}} = s \cdot \sigma_{\text{stat, fit}} \cdot \quad (7.13)$$

7.5.2 Model for validation and systematics

The fitting procedure to extract the CP parameters relies on a mass fit considering multiple components to extract signal sWeights, and a subsequent, weighted decay-time fit. To replicate this sequence in the validation and systematic studies, pseudo-datasets are generated for all individual components. The observables comprise the reconstructed B^0 mass, m , the reconstructed decay time, t' , its uncertainty prediction, σ_t , the tag decisions, \vec{d}' , and their mistag estimates, $\vec{\eta}$. Furthermore, factorization is assumed between the individual decay-time PDFs, $\mathcal{P}^{\text{ch,yr}}(t', \vec{d}' | \vec{\eta}, \sigma_t)$, the decay-time-uncertainty PDFs, $\mathcal{P}^{\text{ch}}(\sigma_t)$, the reconstructed-mass PDF, $\mathcal{P}^{\text{ch}}(m)$, and the mistag PDFs, $\mathcal{P}^{\text{ch}}(\eta_{\text{OS}})$ and $\mathcal{P}^{\text{ch}}(\eta_{\text{SS}})$. This assumption is well motivated, as correlations between the observables have been studied and found to be negligible in simulated signal samples, as well as in pure background samples. Consequently, the PDF generating the full set of observables for each component is written as

$$\mathcal{P}^{\text{ch,yr}}(m, t', \vec{d}', \vec{\eta}, \sigma_t) = \mathcal{P}^{\text{ch}}(m) \cdot \mathcal{P}^{\text{ch,yr}}(t', \vec{d}' | \vec{\eta}, \sigma_t) \cdot \mathcal{P}^{\text{ch}}(\eta_{\text{OS}}) \cdot \mathcal{P}^{\text{ch}}(\eta_{\text{SS}}) \cdot \mathcal{P}^{\text{ch}}(\sigma_t). \quad (7.14)$$

The considered components for both decay modes are the B^0 signal component, the irreducible B_s^0 background component, and the combinatorial background component. The next sections elaborate on the parametrization of the corresponding component PDFs in Eq. (7.14).

Reconstructed mass PDFs

For the parametrization of the reconstructed mass, $\mathcal{P}^{\text{ch}}(m)$, the PDFs used in the mass fits after the full candidate selection (see Secs. 5.2 and 6.4) are employed. All

shape parameters are fixed to the values obtained from these fits.

Decay-time PDF

The decay-time PDFs $\mathcal{P}(t', \vec{d}')$ simultaneously generate pairs of tag decisions \vec{d}' and reconstructed decay time t' .

Signal component For the signal component, the CP -fit model (see Sec. 7.1) is used. The CP parameters are set to $S = 0.7$ and $C = 0.0$.

B_s^0 component The B_s^0 component is parametrized with the nominal fit model (see Sec. 7.1) as well. However, the parameters of the B^0 - \bar{B}^0 system are replaced with those of the B_s^0 - \bar{B}_s^0 system ($\tau_{B^0} \rightarrow \tau_{B_s^0}$, $\Delta\Gamma_d \rightarrow \Delta\Gamma_s$, $\Delta m_d \rightarrow \Delta m_s$), and the CP coefficients S and C are set to zero, as no significant CP violation has been observed in the $B_s^0 \rightarrow J/\psi K_S^0$ mode yet [135].

Combinatorial background component A parametrization for the decay-time distribution of the combinatorial background component is derived from the selected datasets that are weighted with the background sWeights. Here, it is found that the decay-time distributions can be approximated by one or two exponential functions for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and the $B^0 \rightarrow \psi(2S)K_S^0$ mode, respectively, as shown in Fig. 7.3. As not only the decay-time, but also the tag decisions need to be generated for the combinatorial background components, a CP PDF is used (see Sec. 7.1), whose CP coefficients, as well as $\Delta\Gamma$ and Δm are set to zero. Furthermore, perfect decay-time resolutions ($\mathcal{R}(t' - t) = \delta(t' - t)$) and perfect acceptances ($\varepsilon(t') = 1$) are assumed, as only an effective description of the decay-time is needed. The slope parameters of the exponential terms in the PDF are determined from fits to the background sWeighted data.

Mistag-estimate and decay-time-uncertainty PDFs

Histogram PDFs are used to generate the OS and SS per-event mistag-probability estimates η_{OS} and η_{SS} , as well as the decay-time uncertainty estimates σ_t . The corresponding histograms are extracted from sWeighted data for the signal and combinatorial background components. As the yields of the sWeighted B_s^0 component are very low, a signal-like behaviour is assumed and the signal (B^0) sWeighted histograms are used for the generation of the observables.

7.5.3 Validation of the fit

For the validation of the fit, 2000 pseudo-datasets are generated and fitted using the nominal strategy, consisting of the separate, independent mass fits to compute sWeights, and the subsequent simultaneous and weighted decay-time fit to both decay modes.

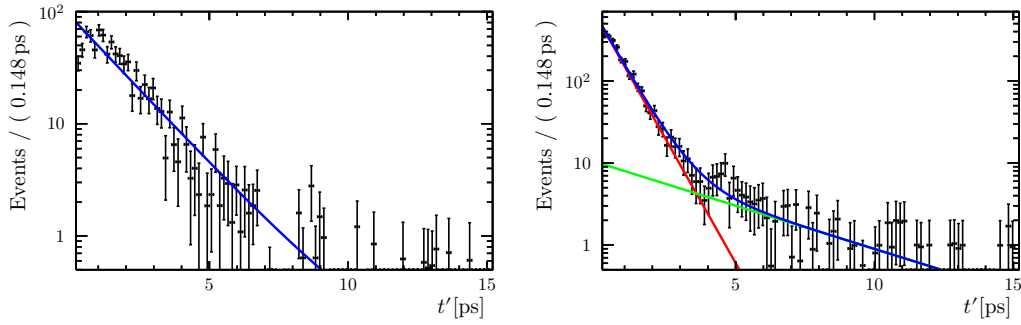


Figure 7.3 – Reconstructed decay-time distribution in background sWeighted $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ (left) and $B^0 \rightarrow \psi(2S)K_S^0$ candidates (right). A single CP PDF is used for the description of the former and the sum of two CP PDFs is used for the latter, with all CP coefficients set to zero. The acceptance effect at low decay times for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ events is consciously ignored to keep the model simple.

Parameters in the generation

While free parameters of the fit are fixed to their respective expected values in the generation of the pseudo-datasets, this approach would result in underestimated uncertainties for parameters which are constrained. For these parameters, for each generated dataset a set of values is sampled from the multivariate normal distribution that is used for the constraint. The sampled values are used for the generation of the individual dataset. The yield of the individual components is sampled from Poisson distributions, whose mean is fixed to the observed number of events in data.

Results of the validation

Fig. 7.4 shows that no significant biases are present in the pull distributions for the individual CP parameters that are determined for the two decay modes. Also, the uncertainties of the parameters are estimated correctly, which is why no systematic uncertainty is assigned for a potential fit bias.

7.5.4 Systematic uncertainties

The systematic uncertainties that are considered arise from mis-modelling of the PDFs used in the fit as well as from uncertainties on external input parameters which are not propagated through constraints into the statistical uncertainty of the results. For both cases the effect on the CP parameters is estimated by studying pseudo-experiments in which pseudo-data is generated using PDFs based on the model described in Sec. 7.5.2. Depending on the effect that is studied, parameters

7 Measurement of $\sin(2\beta)$

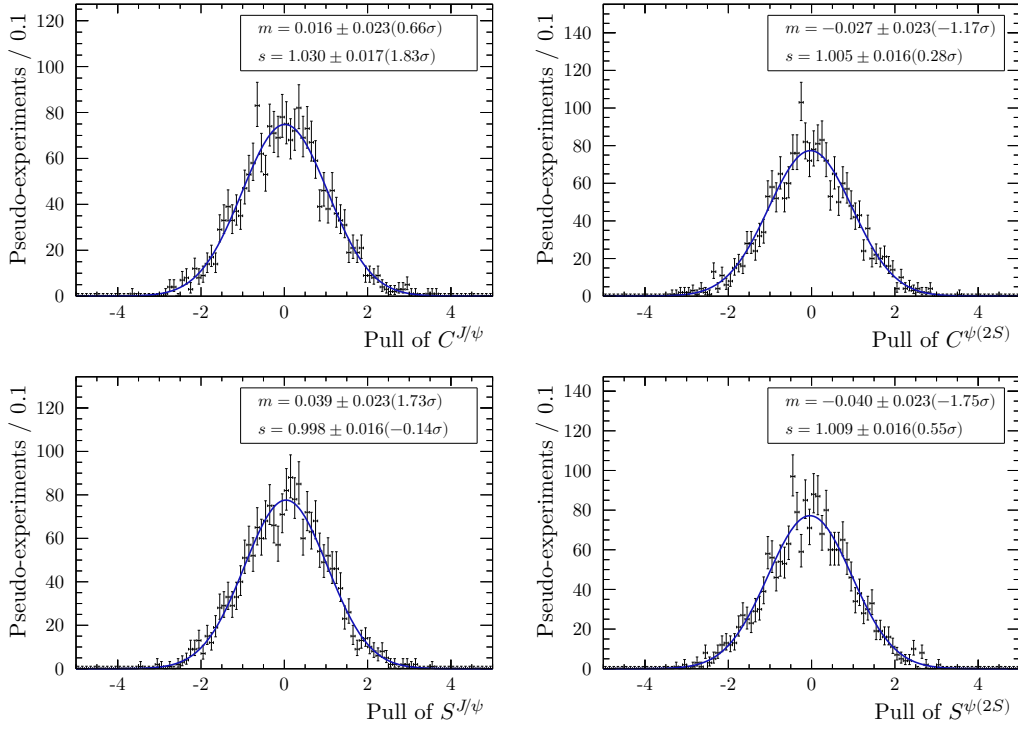


Figure 7.4 – Pull distributions for the CP parameters C (top) and S (bottom) in decays of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ (left) and $B^0 \rightarrow \psi(2S)K_S^0$ (right), obtained from studying 2000 pseudo-datasets in the fitter validation. The parameters are unbiased and uncertainties are correctly estimated.

that differ from the ones used in the nominal fit, or alternate shapes are used to generate the datasets, which are then fitted with the nominal model to test for biases. The probed scenarios are presented in the following section, which closes with a presentation of the derived systematic uncertainties.

Neglect of $\Delta\Gamma$

In the nominal CP fit it is assumed that $\Delta\Gamma = 0$. While, as mentioned before, this assumption is experimentally and theoretically well justified, a systematic uncertainty has to be assigned to cover a potential deviation of this parameter from zero. For this purpose $\Delta\Gamma$ is varied up and down, using its current experimental uncertainty from Ref [53] for the generation of pseudo-datasets. In these experiments the coefficient of the sinh-term, $D = \cos(2\beta)$, is set to 0.72. The 2×1000 pseudo-datasets are fitted with the nominal model under the $\Delta\Gamma = 0$ assumption.

Decay-time resolution

The decay-time-resolution model is entirely determined on simulated data for both modes, as the high period of B^0 - \bar{B}^0 oscillations means that the influence of the decay-time resolution on the measurement is minor. Nonetheless, two possible scenarios are considered in the systematic uncertainties of the measurement: a bias in the decay-time measurement and deviations of the factors which scale the per-event decay-time uncertainties obtained from the decay-tree fit.

Decay-time bias A decay-time bias is expressed as a discrepancy between the mean measured and the mean true decay time. Without corrections of the decay-time measurement, it therefore results in a non-zero mean of the resolution model. Based on the results for the decay-time resolution of the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode presented in Sec. 5.4.1, a small value of $\mu \pm 3$ fs is assumed for both modes in the generation of 2×1000 pseudo-datasets, which are fitted under the $\mu = 0$ hypothesis to test for biases.

Decay-time uncertainty scaling The hybrid approach for the decay-time resolution consists of using the per-event decay-time-uncertainty estimates, σ_t , from data, but determining scaling factors for these estimates on simulated data. This means that the actual scaling needed for the uncertainties on data could differ. A generous variation of the scaling factors a_{wide} and a_{narrow} by $\pm 30\%$ is used to test whether such deviations could cause biases in the measurement of the CP parameters. This alternative assumption is used for the generation of 2×1000 pseudo-datasets, which are fit with the nominal model, using the original scaling factors.

Tagging calibration

The tagging calibration parameters are constrained using their statistical uncertainties and also considering their statistical linear correlation coefficients. However, the systematic uncertainties on the calibration parameters (see Sec. 5.5.3) have not yet been taken into account. As there are four calibration parameters for each tagging algorithm and each decay mode, a completely general study would need to cover 16 possible combinations of varying the parameters individually in different directions. However, previous LHCb analyses of decay-time-dependent CP violation at LHCb [19, 136] found that the impact on the CP parameters is highest, when varying p_0 and p_1 in opposite directions. Therefore, two scenarios are probed for each tagger and decay mode, each using 1000 pseudo-experiments:

- p_0 and Δp_0 are varied up, while p_1 and Δp_1 are varied down by one systematic uncertainty.
- p_0 and Δp_0 are varied down, while p_1 and Δp_1 are varied up by one systematic uncertainty.

The fit to the generated data is then performed with the nominal fit model that uses the original (constrained) calibration parameters.

Production asymmetry and Δm

Only the statistical uncertainty is used to constrain the production-asymmetry values and Δm in the nominal fit. The systematic uncertainties on these external inputs are propagated to the systematic uncertainty of the CP parameters by studying pseudo-datasets, which are generated by varying the parameters by \pm their systematic uncertainty. In contrast, the fit to the generated data is performed with the nominal fit model and constraints. For the study 2×1000 pseudo-datasets are analyzed.

Decay-time acceptance

The knot positions of the decay-time-acceptance functions are fully optimized using simulated events, while the associated coefficients are left floating in the fit to data. In principle, the decay-time acceptance in data could be different from the one seen in simulated events. Therefore, the nominal fit on data is repeated using a more granular acceptance model which contains two knots in addition for each decay mode. The additional knots are inserted at 1.5 ps^{-1} and 3.0 ps^{-1} , because the region at these comparably low decay times is affected strongest by the acceptance effects and contains the most sensitive data. A number of 1000 pseudo-datasets is generated using the granular model and fitted with the nominal fit model to test whether the parameters of interest are affected.

Total systematic uncertainties

The individual systematic uncertainties are obtained by using the mean and the standard deviation of the pull distributions of S and C in Eqs. (7.12) and (7.13). The full systematic uncertainties for the CP parameters are obtained by summing the single uncertainties in quadrature and are presented in Tab. 7.4. All systematic uncertainties as well as the total systematic uncertainty are small compared to the statistical uncertainty of the measurement, and no dominant source is identified. Thus, the measurement is clearly limited by its statistical precision.

7.5.5 Other cross-checks

Additional cross-checks are presented in this section. The fit is repeated on various subsamples of the data, using different fitter implementations, and leaving the mass difference in the B^0 - \bar{B}^0 system free.

Cross-checks on subsamples

A common cross-check is to test whether the fit returns correct results on various subsamples of the data. The incompatibility of results obtained for the CP para-

Table 7.4 – Systematic uncertainties for the CP -violation observables S and C .

Source	$B^0 \rightarrow J/\psi K_S^0$		$B^0 \rightarrow \psi(2S)K_S^0$	
	σ_S	σ_C	σ_S	σ_C
$\Delta\Gamma$	0.003	0.007	0.007	0.003
Δm	0.004	0.004	0.004	0.004
Production asymmetry	0.004	0.009	0.007	0.005
Tagging calibration	0.002	0.005	0.005	0.002
Decay-time bias	0.006	0.006	0.006	0.004
σ_t scaling	0.003	0.005	0.002	0.002
Decay-time efficiency	0.006	0.004	0.006	0.004
Total	0.011	0.016	0.014	0.010

ometers on subsets of the data could hint at the necessity of a subset-dependent parametrization or at problems with the dataset. Therefore, besides the obvious subsets, i.e. the decay modes ($B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ vs. $B^0 \rightarrow \psi(2S)K_S^0$) and years of data taking (2011 vs. 2012), the data is further divided into magnet polarity orientation (MagUp vs. MagDown), and the tagging algorithm that is used (using only OS vs. only SS tagging). The fit is also performed using a setup in which the CP parameters are shared between the two modes. Summarizing, it is found that the received results are compatible for all subsets, as can be seen in Fig. 7.5.

Cross-checks using different fitter implementations

Independent fitter implementations developed by colleagues in Dortmund have been used in order to confirm a correct implementation of the PDF and cross-check the blinded and unblinded results. The two CP fitters used in the recent LHCb analyses of $B^0 \rightarrow D^+D^-$ decays [136] and $B^0 \rightarrow D^\mp\pi^\pm$ decays [133] have yielded fully compatible values and uncertainties for the measured CP parameters.

Cross-check using free Δm

Similar to the decay-time-dependent measurement of CP violation, also the measurement of the oscillation frequency Δm requires the proper reconstruction of the decay-time and tagging information. Therefore, an additional cross-check is performed, which is to not constrain but leave the parameter free and split between

7 Measurement of $\sin(2\beta)$

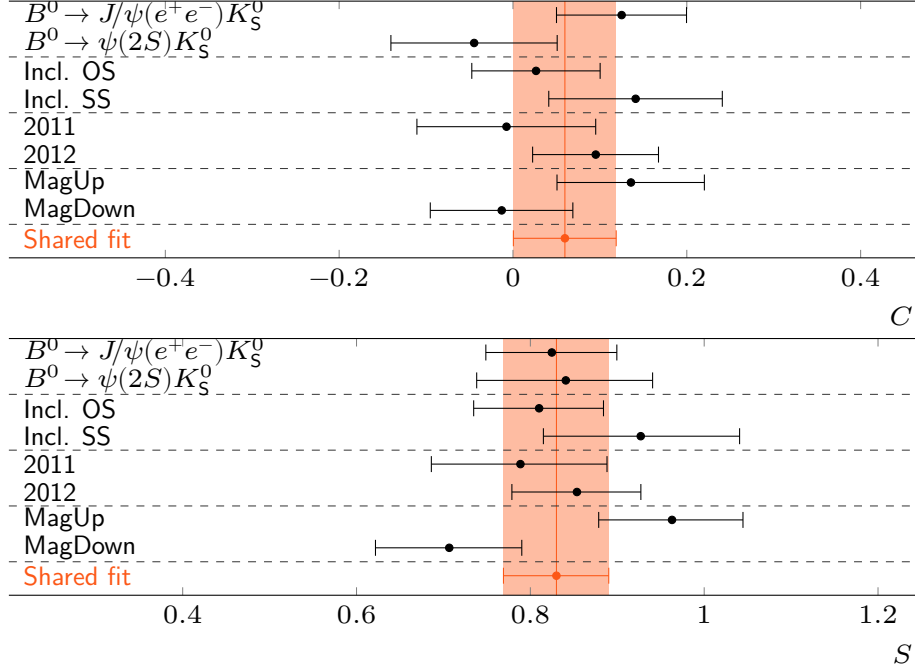


Figure 7.5 – Comparison of fit results for S and C for fits on various subsamples. The red region corresponds to the uncertainty obtained in a fit on the full sample that uses shared CP coefficients between the two modes.

both decay modes in the fit. The obtained results for the CP parameters

$$\begin{aligned}
 C_{\Delta m \text{ free}}^{J/\psi K_S^0} &= 0.107 \pm 0.107, \\
 C_{\Delta m \text{ free}}^{\psi(2S)} &= -0.018 \pm 0.143, \\
 S_{\Delta m \text{ free}}^{J/\psi K_S^0} &= 0.825 \pm 0.077, \\
 S_{\Delta m \text{ free}}^{\psi(2S)} &= 0.849 \pm 0.109,
 \end{aligned} \tag{7.15}$$

are fully compatible with the nominal results. Furthermore, the values on the mass difference determined in the fit are compatible with the current world average, $\Delta m = 0.5065 \pm 0.0016$ [53], for each individual decay mode as

$$\begin{aligned}
 \Delta m_{\Delta m \text{ free}}^{J/\psi K_S^0} &= (0.49 \pm 0.05) \text{ ps}^{-1}, \\
 \Delta m_{\Delta m \text{ free}}^{\psi(2S)} &= (0.52 \pm 0.07) \text{ ps}^{-1}.
 \end{aligned} \tag{7.16}$$

7.6 Correction for kaon regeneration and CP violation

During the reconstruction, the $c\bar{c}K_S^0$ final state must not be confused with the $c\bar{c}K_L^0$ final state, as the latter would lead to a changed sign in the measurement of S , because it has the exact opposite CP eigenvalue (see Sec. 2.5.2). In the context

of this analysis it is assumed that the K_S^0 state can unambiguously be reconstructed from the $\pi^+\pi^-$ final state. There are two problems with this assumption: Firstly, also the K_L^0 eigenstate can decay into the $\pi^+\pi^-$ final state, and secondly a pure K_L^0 state can change back into a superposition of K_L^0 and K_S^0 states, due to different material interaction rates with the detector material of the K^0 and the \bar{K}^0 [137].

The time evolution of arbitrary neutral kaon states presented in Ref. [138] takes into account both $K_L^0 \rightarrow K_S^0$ regeneration effects, as well as CP violation in the K^0 - \bar{K}^0 system through $K_L^0 \rightarrow \pi^+\pi^-$ decays. Predictions of the regeneration effect alone lead to corrections of the raw asymmetry at the level of 0.3% [139]. The effect has been confirmed and corrected for by LHCb in the past, e.g. in Ref. [140], where the combined effect of CP violation and mixing was found to be of the same order and same sign as the regenerations' effects, both contributing to the observed raw asymmetry at the sub-percent level. Consequently, and analogously to the previous LHCb analysis of $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ [19], a corresponding correction is applied here, according to

$$S = S_{\text{fit}} + \Delta S, \quad (7.17)$$

$$C = C_{\text{fit}} + \Delta C, \quad (7.18)$$

where C_{fit} and S_{fit} are the fitted parameters (see. Eq. (7.10)) and ΔC and ΔS are the correction terms. As the difference in the material interaction cross section of \bar{K}^0 and K^0 depends on the K momentum as well as the number of nucleons in the detector material [141], the full LHCb detector material description is used to estimate the effect.

For every single candidate the K_S^0 -candidate's momentum and trajectory are used to evaluate the corrections terms, by dividing the path of the K_S^0 candidate in small steps using the material description of the LHCb detector. The obtained corrections are

$$\Delta C^{J/\psi} = -(0.32 \pm 0.03) \times 10^{-2} \quad (7.19)$$

$$\Delta S^{J/\psi} = +(0.16 \pm 0.02) \times 10^{-2} \quad (7.20)$$

for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode and

$$\Delta C^{\psi(2S)} = -(0.36 \pm 0.04) \times 10^{-2} \quad (7.21)$$

$$\Delta S^{\psi(2S)} = +(0.18 \pm 0.02) \times 10^{-2} \quad (7.22)$$

for the $B^0 \rightarrow \psi(2S)K_S^0$ mode, where the uncertainty is derived from the uncertainty on the detector-material budget that is assumed to be 10%. The computed corrections are of comparable magnitude for both decay modes, and confirm a sub-percent contribution, which has also been observed by the previous analysis in Ref. [140].

8 Combinations

The results of the analysis of $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ decays are combined with the results of the previous LHCb measurement which used $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ decays [19]. For this purpose, the `GammaCombo` framework is used, which has already been used to obtain a combined value of the CKM angle γ from various LHCb results [142]. The likelihood function that is used for the combination is elaborated on in Sec. 8.1. It is followed by a presentation of the results for the combinations of all $B^0 \rightarrow J/\psi K_S^0$ modes in Sec. 8.2, and of all $B^0 \rightarrow [c\bar{c}]K_S^0$ modes in Sec. 8.3.

8.1 Likelihood function

For the combination of the CP parameters, S and C , a multivariate Gaussian likelihood is constructed as,

$$\mathcal{L}(S, C) = \frac{1}{\sqrt{\det \mathbf{V}(2\pi)^n}} \exp\left(-\frac{1}{2}(\vec{x} - \vec{x}')^T \mathbf{V}^{-1}(\vec{x} - \vec{x}')\right). \quad (8.1)$$

This likelihood relates the n measurements, through a vector of their observations, $\vec{x} = (S_1, C_1, \dots, S_n, C_n)^T$, to the two true CP coefficients, encoded in a vector of equal dimensionality, $\vec{x}' = (S, C, \dots, S, C)^T$. The full covariance matrix is composed of the sum of the measurements' statistical and systematic covariances, $\mathbf{V} = \mathbf{V}_{\text{stat}} + \mathbf{V}_{\text{syst}}$.

The statistical covariance matrix is built from the measurement uncertainties, $\vec{\sigma}_{\text{stat}}$, according to

$$\mathbf{V}_{\text{stat}} = \vec{\sigma}_{\text{stat}}^T \mathbf{C}_{\text{stat}} \vec{\sigma}_{\text{stat}}. \quad (8.2)$$

Here, the statistical correlation matrix, \mathbf{C}_{stat} , holds the statistical correlation coefficients for the CP coefficients of each individual measurement and assumes that different measurements are statistically uncorrelated.

To account for correlated systematic uncertainties between the measurements, the various sources of systematic effects (Δm , A_P , ... see Sec. 7.5.4) are accounted for individually, so that

$$\mathbf{V}_{\text{syst}} = \mathbf{V}_{\text{syst},1} + \dots + \mathbf{V}_{\text{syst},m}. \quad (8.3)$$

Here, the systematic covariances, $\mathbf{V}_{\text{syst},i}$, are built from the systematic uncertainties of the respective effects, $\vec{\sigma}_{\text{syst},i}$, according to

$$\mathbf{V}_{\text{syst},i} = \vec{\sigma}_{\text{syst},i}^T \mathbf{C}_{\text{syst},i} \vec{\sigma}_{\text{syst},i}. \quad (8.4)$$

For the correlation matrices, $\mathbf{C}_{\text{syst},i}$, only fully positive correlations (+1) or zero correlations are used, depending on whether the systematic source is correlated or not, respectively. The fully correlated uncertainties are those coming from the systematic uncertainties on Δm and A_p , the $\Delta\Gamma = 0$ assumption, the resolution-model bias and the resolution-model scaling. All other sources of systematic uncertainties are assumed to be uncorrelated.

In order to compute the combined values for the CP coefficients, a χ^2 function is defined as

$$\chi^2(S, C) = -2 \ln \mathcal{L}(S, C) \quad (8.5)$$

and scanned for a global minimum. Confidence contours are obtained by scanning the likelihood function around the minimum and finding points that take a certain difference from the minimum χ^2 value.

8.2 Combination of $B^0 \rightarrow J/\psi K_S^0$ modes

A combination of the LHCb results for the $B^0 \rightarrow J/\psi K_S^0$ modes is performed and published. The main motivation for preparing this combination is the fact, that (hadronic) penguin contributions could affect the J/ψ and $\psi(2S)$ mode differently, which would in turn lead to different effective CP parameters that are measured. Combining the results for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ mode of this analysis, with the results for the $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ mode of the previous LHCb measurement which obtained $S = 0.73 \pm 0.04$ and $C = -0.038 \pm 0.032$ [19], results in

$$C(B^0 \rightarrow J/\psi K_S^0) = -0.014 \pm 0.030 \text{ (stat. + syst.)}, \quad (8.6)$$

$$S(B^0 \rightarrow J/\psi K_S^0) = 0.75 \pm 0.04 \text{ (stat. + syst.)}, \quad (8.7)$$

with a linear correlation coefficient of $\rho = 0.42$. Fig. 8.1 shows the corresponding two-dimensional likelihood scan. This result is compatible within 1.9 standard deviations with the $B^0 \rightarrow J/\psi K_S^0$ average of the B -factories [53], and agrees with predictions from indirect measurements by the CKMfitter group [42] and the UTfit collaboration [67].

8.3 Combination of $B^0 \rightarrow c\bar{c}K_S^0$ modes

At the current level of experimental precision, the aforementioned penguin contributions are still small. Since this is true for both studied modes, a combination of measurements of all $c\bar{c}K_S^0$ modes is prepared and published. The obtained results are

$$C(B^0 \rightarrow [c\bar{c}]K_S^0) = -0.017 \pm 0.029 \text{ (stat. + syst.)}, \quad (8.8)$$

$$S(B^0 \rightarrow [c\bar{c}]K_S^0) = 0.760 \pm 0.034 \text{ (stat. + syst.)}, \quad (8.9)$$

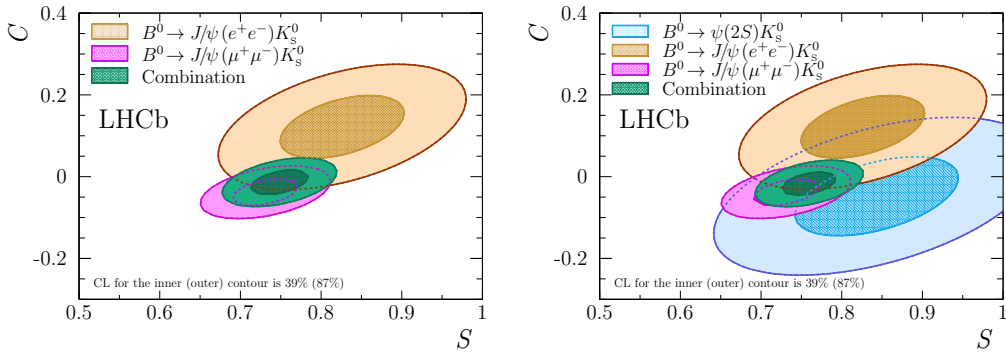


Figure 8.1 – Two-dimensional likelihood scans for the combination of the $B^0 \rightarrow J/\psi K_S^0$ modes (left) and all $B^0 \rightarrow [c\bar{c}]K_S^0$ modes (right) [20]. The coloured regions correspond to 39% and 87% confidence levels.

with a linear correlation coefficient of $\rho = 0.42$ between both measurements. Fig. 8.1 shows the corresponding two-dimensional likelihood scan. This combination agrees with predictions by the CKMfitter group [42] and the UTfit collaboration [67] and improves the sensitivity of $\sin(2\beta)$ at LHCb by 20 %.

9 Conclusion and Outlook

Analyses performed during the first and the current running period of the LHC confirm the tremendous success of the Standard Model of particle physics. But even though none of its predictions could be falsified so far, the existence of physics beyond the SM is indisputable. One important fact is that our current understanding of CP violation and its observed magnitude are inconsistent with the omni-present dominance of matter in the universe. Since the discovery of CP violation, research has made overwhelming improvements with regards to the accuracy that can be reached in the measurements of this effect. Precise analyses at BaBar and Belle were able to observe CP violation of substantial magnitude in the decay $B^0 \rightarrow J/\psi K_S^0$, which complements our picture of the CKM sector. These analyses performed in the clean environments of the e^+e^- machines lay the foundation in an era of precision measurements of CP violation. The lack of observations of New Physics effects in all direct and indirect searches shows that its manifestations must be small, which is why high-precision measurements become increasingly important. Situated in a challenging but also high-statistics hadronic environment, the LHCb experiment joins this venture of precision measurements.

Already the full analysis of 114000 $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ decays, performed using data corresponding to 3fb^{-1} of proton-proton collisions collected during the first running period of the experiment at centre-of-mass energies of $\sqrt{s} = 7\text{TeV}$ and $\sqrt{s} = 8\text{TeV}$, is competitive and compatible with the previous measurements of the B -factories [55, 56], and results in the CP observables [19]

$$\begin{aligned} C(B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0) &= -0.038 \pm 0.032 (\text{stat.}) \pm 0.005 (\text{syst.}), \\ S(B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0) &= 0.731 \pm 0.035 (\text{stat.}) \pm 0.020 (\text{syst.}). \end{aligned}$$

This thesis presents an analysis of decay-time-dependent CP violation using $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ decays, which has been published in Ref. [20]. It is the first decay-time-dependent measurement at a hadron collider that uses electrons in the final state. Compared to the previous analysis, the measurement benefits from the additional inclusion of the OS charm and the SS proton flavour-tagging algorithms. The more challenging reconstruction of the $J/\psi \rightarrow e^+e^-$ final state leads to an efficiency-loss and thus to a signal yield of 10630 $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ decays. This is partially compensated by the increase in tagging power, which is twice as high with respect to the former analysis, so that nearly half of its precision can be reached, resulting in measurements of

$$\begin{aligned} C(B^0 \rightarrow J/\psi(e^+e^-)K_S^0) &= 0.12 \pm 0.07 (\text{stat.}) \pm 0.02 (\text{syst.}), \\ S(B^0 \rightarrow J/\psi(e^+e^-)K_S^0) &= 0.83 \pm 0.08 (\text{stat.}) \pm 0.01 (\text{syst.}), \end{aligned}$$

9 Conclusion and Outlook

with a linear statistical correlation coefficient of 0.46. The lower branching ratio compared to the $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ mode leads to a yield of 7970 $B^0 \rightarrow \psi(2S)K_S^0$ decays. With a relative increase in tagging power of 13% compared to the $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ mode, the CP parameters are measured at a third of its precision, resulting in

$$\begin{aligned} C(B^0 \rightarrow \psi(2S)K_S^0) &= -0.05 \pm 0.10 (\text{stat.}) \pm 0.01 (\text{syst.}), \\ S(B^0 \rightarrow \psi(2S)K_S^0) &= 0.84 \pm 0.10 (\text{stat.}) \pm 0.01 (\text{syst.}), \end{aligned}$$

with a linear statistical correlation coefficient of 0.42. Both individual results are compatible with each other, with the previous measurements, and with predictions from indirect measurements by the CKMfitter group [42] and the UTfit collaboration [67]. The theoretical limits on higher-order corrections that could cause a deviation between the measured values of S and $\sin(2\beta)$ are $|\Delta S| < 0.0086$ and $|\Delta S| < 0.0094$ for $B^0 \rightarrow J/\psi K_S^0$ decays and $B^0 \rightarrow \psi(2S)K_S^0$ decays, respectively [66]. These imply that the corresponding contributions can be neglected at the current level of experimental precision. It is also not possible to identify potential different higher-order contributions in $B^0 \rightarrow \psi(2S)K_S^0$ decays compared to $B^0 \rightarrow J/\psi K_S^0$ decays at the given sensitivity.

As both modes could in principle be affected differently by higher-order corrections, two separate combinations are performed with the previous LHCb result from Ref. [19], using scans of a combined likelihood function. In the first combination the results for the $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ and the $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ mode are combined, resulting in

$$\begin{aligned} C(B^0 \rightarrow J/\psi K_S^0) &= -0.014 \pm 0.030 (\text{stat.} + \text{syst.}), \\ S(B^0 \rightarrow J/\psi K_S^0) &= 0.75 \pm 0.04 (\text{stat.} + \text{syst.}), \end{aligned}$$

with a linear correlation coefficient of 0.42. These results agree within 1.9 standard deviations with the $B^0 \rightarrow J/\psi K_S^0$ average of the B -factories [53], while the values for the $\psi(2S)$ mode are compatible within 0.3 standard deviations with the $B^0 \rightarrow \psi(2S)K_S^0$ average of the B -factories [53]. Finally a combination of all three modes, $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$, $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$, results in

$$\begin{aligned} C(B^0 \rightarrow [c\bar{c}]K_S^0) &= -0.017 \pm 0.029 (\text{stat.} + \text{syst.}), \\ S(B^0 \rightarrow [c\bar{c}]K_S^0) &= 0.760 \pm 0.034 (\text{stat.} + \text{syst.}), \end{aligned}$$

with a linear correlation coefficient of 0.42. These combined values agree with indirect measurements by the CKMfitter group [42] and the UTfit collaboration [67]. Furthermore, the precision of $\sin(2\beta)$ at LHCb is improved by 20%, which will directly result in an improvement of the world average.

While a net discrepancy of two standard deviations between the LHCb result and the average of the B -factory measurements is nothing to trigger concern, it will remain interesting to see whether this increasing trend continues with future updates of $\sin(2\beta)$ measurements by LHCb. These updates will benefit from an

increased size of the data sample, as well as developments in the flavour tagging and the trigger. At the end of the current (second) running period by the end of 2018, approximately 8 fb^{-1} of proton-proton collision data will be available. The higher centre-of-mass energy of $\sqrt{s} = 13 \text{ TeV}$ that the LHC is operated at during this current run, leads to a higher $b\bar{b}$ production cross section, as it scales approximately linearly with \sqrt{s} . Assuming constant efficiencies, this will allow to produce the single most precise measurement of $\sin(2\beta)$ in the world with an expected uncertainty of ≈ 0.011 using $B^0 \rightarrow J/\psi(\mu^+\mu^-)K_S^0$ decays alone. During LHC's Long Shutdown 2 (LS2), LHCb will be upgraded with improved subdetectors and the trigger system will be replaced by a full software trigger, allowing for a trigger-less readout at 40 MHz [143]. After LS2, the data-sample size is expected to increase to 30 fb^{-1} , during the third running period between 2021–2023 at a centre-of-mass energy of $\sqrt{s} = 14 \text{ TeV}$.

The biggest competitor of LHCb in this field is the successor of the Belle experiment, Belle II, located at the SuperKEKB collider. It has recently started to record e^+e^- collisions at a centre-of-mass energy close to the mass of the $\Upsilon(4S)$ resonance. Working at a target luminosity of $8 \cdot 10^{35} \text{ cm}^{-2}\text{s}^{-1}$, it will be the leading experiment in terms of collision intensity, and plans to record 5 ab^{-1} and 50 ab^{-1} of collisions by 2020 and 2024, respectively, where the latter represents a seventy-fold increase compared to the predecessor experiment [144, 145]. This should allow Belle II to catch up in precision by as soon as 2020. Shortly after, the uncertainty of the CP observables will reach < 0.009 , which is when theoretical uncertainties need to be controlled. The analysis of related decay modes, e.g. future updates of the study of CP violation in $B_s^0 \rightarrow J/\psi K_S^0$ at LHCb [135], can help to control these uncertainties.

Bibliography

- [1] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588, DOI: 10.1016/0029-5582(61)90469-2.
- [2] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168–171, DOI: 10.1016/0031-9163(64)90711-5.
- [3] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266, DOI: 10.1103/PhysRevLett.19.1264.
- [4] 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.
- [5] C. Wu *et al.*, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413–1414, DOI: 10.1103/PhysRev.105.1413.
- [6] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533, DOI: 10.1103/PhysRevLett.10.531.
- [7] 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.
- [8] 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.
- [9] S. W. Herb *et al.*, *Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252–255, DOI: 10.1103/PhysRevLett.39.252.
- [10] F. Abe *et al.*, *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626–2631, DOI: 10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002 [hep-ex].
- [11] S. Abachi *et al.*, *Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **74** (1995) 2422–2426, DOI: 10.1103/PhysRevLett.74.2422, arXiv:hep-ex/9411001 [hep-ex].
- [12] V. Fanti *et al.*, *A New measurement of direct CP violation in two pion decays of the neutral kaon*, Phys. Lett. **B465** (1999) 335–348, DOI: 10.1016/S0370-2693(99)01030-8, arXiv:hep-ex/9909022 [hep-ex].
- [13] A. Alavi-Harati *et al.*, *Observation of direct CP violation in $K_{S,L} \rightarrow \pi\pi$ decays*, Phys. Rev. Lett. **83** (1999) 22–27, DOI: 10.1103/PhysRevLett.83.22, arXiv:hep-ex/9905060 [hep-ex].

Bibliography

- [14] H. Albrecht *et al.*, *Observation of B^0 - \bar{B}^0 Mixing*, Phys. Lett. **B192** (1987) 245–252, DOI: 10.1016/0370-2693(87)91177-4.
- [15] 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, arXiv:hep-ex/0107013 [hep-ex].
- [16] 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, arXiv:hep-ex/0107061 [hep-ex].
- [17] I. I. Bigi and A. Sanda, *Notes on the Observability of CP Violations in B Decays*, Nucl. Phys. **B193** (1981) 85, DOI: 10.1016/0550-3213(81)90519-8.
- [18] A. B. Carter and A. Sanda, *CP Violation in B Meson Decays*, Phys. Rev. **D23** (1981) 1567, DOI: 10.1103/PhysRevD.23.1567.
- [19] 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].
- [20] R. Aaij *et al.*, *Measurement of CP violation in $B^0 \rightarrow J/\psi K_S^0$ and $B^0 \rightarrow \psi(2S)K_S^0$ decays*, JHEP **11** (2017) 170, DOI: 10.1007/JHEP11(2017)170, arXiv:1709.03944 [hep-ex].
- [21] V. Müller, *Improvement of the LHCb measurement of $\sin(2\beta)$ in additional B^0 decay modes*, Ph.D. thesis, TU Dortmund University, in preparation.
- [22] A. Seuthe and J. Albrecht, *Flavour-tagging and particle-identification calibration in key measurements with electrons at LHCb*, M.Sc. thesis, TU Dortmund University, 2017, <https://cds.cern.ch/record/2293431>.
- [23] K. Mielke, *Experimentelle Besonderheiten einer getaggtten zeitabhängigen CP-Verletzungsmessung in $B^0 \rightarrow J/\psi(e^+e^-)K_S^0$ am LHCb Experiment*, unpublished, B.Sc. thesis, TU Dortmund University, 2015.
- [24] L. Nickel, *Studien zu einer Gütefunktion für CP-Verletzungsmessungen bei LHCb*, unpublished, B.Sc. thesis, TU Dortmund University, 2017.
- [25] C. Hasenberg, *Sensitivity study of the CPV-Parameter $\sin(2\beta)$ in the decay channel $B^0 \rightarrow J/\psi K_S^0$* , unpublished, M.Sc. thesis, TU Dortmund University, 2017.
- [26] G. Lüders, *On the equivalence of invariance under time reversal and under particle-antiparticle conjugation for relativistic field theories*, Dan. Mat. Fys. Medd. **28.5** (1954) 1–17.
- [27] W. Pauli and L. F. Curtiss, *Niels Bohr and the Development of Physics*, Am. J. Phys. **24** (1956) 292, DOI: 10.1119/1.1934212.
- [28] D. Boutigny *et al.*, *The BABAR physics book: Physics at an asymmetric B factory*, SLAC-R-504 (1998), <http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=SLAC-R-504>.

- [29] G. C. Branco, L. Lavoura, and J. P. Silva, *CP Violation*, Int. Ser. Monogr. Phys. **103** (1999).
- [30] I. I. Bigi and A. Sanda, *CP violation*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **9** (2000).
- [31] C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40**.10 (2016) 100001 and 2017 update, DOI: 10.1088/1674-1137/40/10/100001.
- [32] 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.
- [33] 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].
- [34] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509, DOI: 10.1103/PhysRevLett.13.508.
- [35] 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.
- [36] R. Aaij *et al.*, *Observation of J/ψ 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].
- [37] R. Aaij *et al.*, *Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \rightarrow J/\psi\phi K^+$ decays*, Phys. Rev. Lett. **118**.2 (2017) 022003, DOI: 10.1103/PhysRevLett.118.022003, arXiv:1606.07895 [hep-ex].
- [38] R. Aaij *et al.*, *Amplitude analysis of $B^+ \rightarrow J/\psi\phi K^+$ decays*, Phys. Rev. **D95**.1 (2017) 012002, DOI: 10.1103/PhysRevD.95.012002, arXiv:1606.07898 [hep-ex].
- [39] 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.
- [40] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945, DOI: 10.1103/PhysRevLett.51.1945.
- [41] 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, arXiv:hep-ph/9403384 [hep-ph].

- [42] J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ new physics*, Phys. Rev. **D91** (2015), updated results and plots available at <http://ckmfitter.in2p3.fr/073007>, DOI: 10.1103/PhysRevD.91.073007, arXiv:1501.05013 [hep-ph].
- [43] 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.
- [44] 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, arXiv:hep-ph/0406184 [hep-ph], updated results and plots available at: <http://ckmfitter.in2p3.fr>.
- [45] V. Weisskopf and E. P. Wigner, *Calculation of the natural brightness of spectral lines on the basis of Dirac's theory*, Z. Phys. **63** (1930) 54–73, DOI: 10.1007/BF01336768.
- [46] V. Weisskopf and E. Wigner, *Over the natural line width in the radiation of the harmonius oscillator*, Z. Phys. **65** (1930) 18–29, DOI: 10.1007/BF01397406.
- [47] R. Aaij *et al.*, *First observation of CP violation in the decays of B_s^0 mesons*, Phys. Rev. Lett. **110**.22 (2013) 221601, DOI: 10.1103/PhysRevLett.110.221601, arXiv:1304.6173 [hep-ex].
- [48] A. Lenz and U. Nierste, *Numerical Updates of Lifetimes and Mixing Parameters of B Mesons, CKM unitarity triangle. Proceedings, 6th International Workshop, CKM 2010, Warwick, UK, September 6-10, 2010*, 2011, arXiv:1102.4274 [hep-ph].
- [49] M. Artuso, G. Borissov, and A. Lenz, *CP violation in the B_s^0 system*, Rev. Mod. Phys. **88**.4 (2016) 045002, DOI: 10.1103/RevModPhys.88.045002, arXiv:1511.09466 [hep-ph].
- [50] 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, arXiv:1409.8586 [hep-ex].
- [51] R. Aaij *et al.*, *Measurement of the CP asymmetry in B_s^0 - \bar{B}_s^0 mixing*, Phys. Rev. Lett. **117** (2016) 061803, DOI: 10.1103/PhysRevLett.117.061803, arXiv:1605.09768 [hep-ex].
- [52] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L^0 \rightarrow \mu\bar{\mu}$, $K^+ \rightarrow \pi^+\nu\bar{\nu}$ and $K^0 \leftrightarrow \bar{K}^0$* , Prog. Theor. Phys. **65** (1981), [Erratum: Prog. Theor. Phys.65,1772(1981)] 297, DOI: 10.1143/PTP.65.297.
- [53] Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ -lepton properties as of summer 2016* (2016), updated results and plots available at <http://www.slac.stanford.edu/xorg/hflav/>, arXiv:1612.07233 [hep-ex].

- [54] 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, arXiv: hep-ph/9903455 [hep-ph].
- [55] B. Aubert *et al.*, *Measurement of Time-Dependent CP Asymmetry in $B^0 \rightarrow c\bar{c}K^{(*)0}$ Decays*, Phys. Rev. **D79** (2009) 072009, DOI: 10.1103/PhysRevD.79.072009, arXiv:0902.1708 [hep-ex].
- [56] I. Adachi *et al.*, *Precise measurement of the CP violation parameter $\sin 2\phi_1$ in $B^0 \rightarrow (c\bar{c})K^0$ decays*, Phys. Rev. Lett. **108** (2012) 171802, DOI: 10.1103/PhysRevLett.108.171802, arXiv:1201.4643 [hep-ex].
- [57] B. Aubert *et al.*, *Time-dependent amplitude analysis of $B^0 \rightarrow K_S^0 \pi^+ \pi^-$* , Phys. Rev. **D80** (2009) 112001, DOI: 10.1103/PhysRevD.80.112001, arXiv:0905.3615 [hep-ex].
- [58] B. Aubert *et al.*, *Measurement of $\sin(2\beta)$ using hadronic J/ψ decays*, Phys. Rev. **D69** (2004) 052001, DOI: 10.1103/PhysRevD.69.052001, arXiv:hep-ex/0309039 [hep-ex].
- [59] R. Barate *et al.*, *Study of the CP asymmetry of $B^0 \rightarrow J/\psi K_S^0$ decays in ALEPH*, Phys. Lett. **B492** (2000) 259–274, DOI: 10.1016/S0370-2693(00)01091-1, arXiv:hep-ex/0009058 [hep-ex].
- [60] K. Ackerstaff *et al.*, *Investigation of CP violation in $B^0 \rightarrow J/\psi K_S^0$ decays at LEP*, Eur. Phys. J. **C5** (1998) 379–388, DOI: 10.1007/s100520050284, arXiv:hep-ex/9801022 [hep-ex].
- [61] T. Affolder *et al.*, *A measurement of $\sin(2\beta)$ from $B^0 \rightarrow J/\psi K_S^0$ with the CDF detector*, Phys. Rev. **D61** (2000) 072005, DOI: 10.1103/PhysRevD.61.072005, arXiv:hep-ex/9909003 [hep-ex].
- [62] Y. Sato *et al.*, *Measurement of the CP-violation Parameter $\sin 2\phi_1$ with a New Tagging Method at the $\Upsilon(5S)$ Resonance*, Phys. Rev. Lett. **108** (2012) 171801, DOI: 10.1103/PhysRevLett.108.171801, arXiv:1201.3502 [hep-ex].
- [63] 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, arXiv:1010.0089 [hep-ph].
- [64] 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, arXiv:1412.6834 [hep-ph].
- [65] M. Jung, *Determining weak phases from $B \rightarrow J/\psi P$ decays*, Phys. Rev. **D86** (2012) 053008, DOI: 10.1103/PhysRevD.86.053008, arXiv:1206.2050 [hep-ph].
- [66] 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].

- [67] M. Bona *et al.*, *The unitarity triangle fit in the standard model and hadronic parameters from lattice QCD: A reappraisal after the measurements of Δm_s and $BR(B \rightarrow \tau \nu_\tau)$* , JHEP **10** (2006), updated results and plots available at <http://www.utfit.org/> 081, DOI: 10.1088/1126-6708/2006/10/081, arXiv:hep-ph/0606167 [hep-ph].
- [68] J. Alves A. Augusto *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005, DOI: 10.1088/1748-0221/3/08/S08005.
- [69] R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, DOI: 10.1142/S0217751X15300227, arXiv:1412.6352 [hep-ex].
- [70] C. Lefèvre, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, 2008, <https://cds.cern.ch/record/1260465/files/0812015.pdf> (visited on 2018-05-08).
- [71] 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.
- [72] S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004, DOI: 10.1088/1748-0221/3/08/S08004.
- [73] K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002, DOI: 10.1088/1748-0221/3/08/S08002.
- [74] 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.
- [75] 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].
- [76] 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.
- [77] *Angle correlation in $b\bar{b}$ production at the LHC*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html (visited on 2018-02-01).
- [78] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, LHCb-PUB-2011-016 (2011), <https://cds.cern.ch/record/1384380/files/LHCb-PUB-2011-016.pdf>.
- [79] 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 [hep-ex].
- [80] 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 [hep-ex].
- [81] 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.

- [82] 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.
- [83] 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.
- [84] R. Aaij *et al.*, *The Moore Project*, ed. by G. Raven, 2015, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/>.
- [85] 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/>.
- [86] R. Aaij *et al.*, *The DaVinci Project*, ed. by E. Rodrigues, 2015, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/>.
- [87] 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.
- [88] 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.
- [89] R. Aaij *et al.*, *The Boole Project*, ed. by M. Cattaneo, 2015, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/>.
- [90] M. Needham, *Clone track identification using the Kullback-Liebler distance*, CERN-LHCb-2008-002 (2008), <http://cds.cern.ch/record/1082460>.
- [91] D. A. Berninghoff, J. Albrecht, and V. Gligorov, *Bremsstrahlung Recovery of Electrons using Multivariate Methods*, LHCb-INT-2016-018 (2016), <https://cds.cern.ch/record/2146447>.
- [92] O. Deschamps *et al.*, *Photon and neutral pion reconstruction*, LHCb-2003-091 (2003), <https://cds.cern.ch/record/691634>.
- [93] M. Needham and T. Ruf, *Estimation of the material budget of the LHCb detector*, LHCb-2007-025 (2007), <https://cds.cern.ch/record/1023537>.
- [94] D. Berninghoff, *Search for Lepton Flavour Violation and Optimisation of Electron Reconstruction at LHCb*, unpublished, M.Sc. thesis, TU Dortmund University, 2016.
- [95] 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 [physics.ins-det].
- [96] R. W. Forty, *Ring-Imaging Cherenkov Detectors for LHCb*, CERN-PPE-96-176 (1996) 16 p, <http://cds.cern.ch/record/317233>.

Bibliography

- [97] 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 [physics.ins-det].
- [98] 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].
- [99] 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].
- [100] A. Buckley *et al.*, *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145–233, DOI: 10.1016/j.physrep.2011.03.005, arXiv:1101.2599 [hep-ph].
- [101] T. Gleisberg *et al.*, *SHERPA 1. alpha: A Proof of concept version*, JHEP **02** (2004) 056, DOI: 10.1088/1126-6708/2004/02/056, arXiv:hep-ph/0311263 [hep-ph].
- [102] T. Gleisberg *et al.*, *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, DOI: 10.1088/1126-6708/2009/02/007, arXiv:0811.4622 [hep-ph].
- [103] M. Bahr *et al.*, *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707, DOI: 10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883 [hep-ph].
- [104] D. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155, DOI: 10.1016/S0168-9002(01)00089-4.
- [105] 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].
- [106] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS Interface in C++: Technical and Physics Documentation* (2010), arXiv:1011.0937 [hep-ph].
- [107] S. Agostinelli *et al.*, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303, DOI: 10.1016/S0168-9002(03)01368-8.
- [108] J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270, DOI: 10.1109/TNS.2006.869826.
- [109] *Schematic overview of the underlying principles of LHCb’s flavour tagging algorithms*, <https://twiki.cern.ch/twiki/bin/view/LHCb/FlavourTaggingConferencePlots> (visited on 2018-05-14).
- [110] 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].
- [111] 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].

- [112] R. Aaij *et al.*, *New algorithms for identifying the flavour of B^0 mesons using pions and protons*, Eur. Phys. J. **C77.4** (2017) 238, DOI: 10.1140/epjc/s10052-017-4731-y, arXiv:1610.06019 [hep-ex].
- [113] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit* (2009), arXiv:0905.0724 [physics.data-an].
- [114] 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.
- [115] 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].
- [116] 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].
- [117] Y. Xie, *Principles to optimize event selections for measurements of CP asymmetries*, LHCb-INT-2009-012 (2009), <https://cds.cern.ch/record/1187570>.
- [118] R. Aaij *et al.*, *First study of the CP-violating phase and decay-width difference in $B_s^0 \rightarrow \psi(2S)\phi$ decays*, Phys. Lett. **B762** (2016) 253–262, DOI: 10.1016/j.physletb.2016.09.028, arXiv:1608.04855 [hep-ex].
- [119] C. Cauet, *Precision measurement of the CKM parameter $\sin(2\beta)$ with the LHCb experiment*, Ph.D. thesis, TU Dortmund University, 2015.
- [120] J. H. Friedman, *Greedy Function Approximation: A Gradient Boosting Machine*, Annals of Statistics **29** (2000) 1189–1232.
- [121] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119, DOI: 10.1006/jcss.1997.1504.
- [122] 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].
- [123] 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].
- [124] M. Chaichian and A. Fridman, *On a possibility for measuring effects of CP violation at pp colliders*, Phys. Lett. **B298** (1993) 218–223, DOI: 10.1016/0370-2693(93)91733-4.
- [125] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000) 137–161, DOI: 10.1007/s100520000460, arXiv:hep-ph/0005110 [hep-ph].

Bibliography

- [126] E. Norrbin and R. Vogt, *Bottom production asymmetries at the LHC* (2000), arXiv:hep-ph/0003056 [hep-ph].
- [127] R. Aaij *et al.*, *Measurement of B^0 , B_s^0 , B^+ and Λ_b^0 production asymmetries in 7 and 8 TeV pp collisions* (2017), submitted to Phys. Lett. B, arXiv:1703.08464 [hep-ex].
- [128] 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].
- [129] M. Schellenberg, *Analyse der VELO-Rekonstruktionseffizienz im Zerfall $B^0 \rightarrow J/\psi K_S^0$ am LHCb-Experiment*, unpublished, B.Sc. thesis, TU Dortmund University, 2013.
- [130] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation* (2014), arXiv:1407.0748 [physics.data-an].
- [131] J. Wimberley, *EspressoPerformanceMonitor*, 2018, <https://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor> (visited on 2018-05-16).
- [132] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762**.1 (2016) 012036, DOI: 10.1088/1742-6596/762/1/012036, arXiv:1608.05806 [physics.data-an].
- [133] R. Aaij *et al.*, *Measurement of CP violation in $B^0 \rightarrow D^\mp \pi^\pm$ decays* (2018), submitted to JHEP, arXiv:1805.03448 [hep-ex].
- [134] T. Brambach *et al.*, *Measurement of CP violation in the time-dependent analysis of $B^0 \rightarrow J/\psi K_S^0$ decays with the 2010 data* (2011), LHCb-ANA-2011-004, <https://cds.cern.ch/record/1309564>.
- [135] 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].
- [136] R. Aaij *et al.*, *Measurement of CP violation in $B^0 \rightarrow D^+ D^-$ decays*, Phys. Rev. Lett. **117**.26 (2016) 261801, DOI: 10.1103/PhysRevLett.117.261801, arXiv:1608.06620 [hep-ex].
- [137] A. Pais and O. Piccioni, *Note on the Decay and Absorption of the θ^0* , Phys. Rev. **100** (5 1955) 1487–1489, DOI: 10.1103/PhysRev.100.1487, <https://link.aps.org/doi/10.1103/PhysRev.100.1487>.
- [138] W. Fetscher *et al.*, *Regeneration of arbitrary coherent neutral kaon states: A new method for measuring the K^0 - \bar{K}^0 forward scattering amplitude*, Z. Phys. **C72** (1996) 543–547, DOI: 10.1007/s002880050277.
- [139] B. Ko *et al.*, *Effect of nuclear interactions of neutral kaons on CP asymmetry measurements*, Phys. Rev. **D84** (2011) 111501, DOI: 10.1103/PhysRevD.84.111501, arXiv:1006.1938 [hep-ex].

- [140] R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \rightarrow K^- K^+$ and $D^0 \rightarrow \pi^- \pi^+$ decays*, JHEP **1407** (2014) 041, DOI: 10.1007/JHEP07(2014)041, arXiv:1405.2797 [hep-ex].
- [141] A. Gsponer *et al.*, *Precise Coherent K_S Regeneration Amplitudes for C, Al, Cu, Sn, and Pb Nuclei from 20 to 140 GeV/c and Their Interpretation*, Phys. Rev. Lett. **42** (1 1979) 13–16, DOI: 10.1103/PhysRevLett.42.13, <https://link.aps.org/doi/10.1103/PhysRevLett.42.13>.
- [142] R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of LHCb results*, JHEP **12** (2016) 087, DOI: 10.1007/JHEP12(2016)087, arXiv:1611.03076 [hep-ex].
- [143] *LHCb Trigger and Online Upgrade Technical Design Report*, CERN-LHCC-2014-016, LHCb-TDR-016 (2014), <http://cds.cern.ch/record/1701361>.
- [144] E. Kou, P. Urquijo, The Belle II collaboration, and The B2TiP theory community, *The Belle II Physics Book* (in preparation).
- [145] J. Albrecht *et al.*, *Future prospects for exploring present day anomalies in flavour physics measurements with Belle II and LHCb* (2017), arXiv:1709.10308 [hep-ph].

Acknowledgement / Danksagung

Lieber Prof. Dr. Bernhard Spaan, seit meiner Masterarbeit bis zum Ende der Promotion hatte ich die Möglichkeit an Ihrem Lehrstuhl in einem spannenden und internationalen Umfeld forschen und arbeiten zu dürfen. Vielen Dank für diese wertvolle Zeit.

Des Weiteren danke ich Prof. Dr. Kevin Kröniger, der freundlicherweise das Zweitgutachten der Arbeit übernimmt, sowie Prof. Dr. Roland Böhmer und Dr. Tim Ruhe, die die Prüfungskommission vervollständigen.

Ich danke dem Bundesministerium für Bildung und Forschung und der Deutschen Forschungsgemeinschaft für die Unterstützung meiner Promotion, durch die ich mich voll und ganz auf diese Arbeit konzentrieren konnte.

Vielen Dank Vanessa, dass du zusammen mit mir am Ball geblieben bist und dich mit mir durch Dick und Dünn gekämpft hast. Trotz vieler anstrengender, arbeitsintensiver und sicherlich manchmal auch schwieriger Zeiten, konnten wir dennoch oft zusammen lachen und unsere Analyse nun gemeinsam erfolgreich abschließen. Lieber Julian, du hast nicht nur mit deinem fachlichen Rat einen riesigen Anteil an meiner Arbeit, dieser Analyse, und meinem persönlichen Werdegang, sondern wusstest auch als Freund in allen weiteren Lebenslagen immer zu helfen. Dafür danke ich dir von Herzen. Lieber Moritz, ich danke dir für eine lange Freundschaft und die vielen gemeinsamen Gespräche diverser Art. Auf dich konnte ich im Studium und anschließend auch während der Promotion immer zählen.

Lieber Alex B., lieber Frank, euch möchte ich für zahlreiche ertragreiche und hoffentlich erträgliche Diskussionen mit mir danken, die mich häufig das entscheidende Stück weiter ans Ziel brachten. Lieber Alex S., ich danke dir für die super Zusammenarbeit während und nach deiner Masterarbeit.

Danke an meine Bürokollegen Timon, Patrick, Kevin, Christophe und Florian für eine stets angenehme Arbeits- und Pausen-Atmosphäre. Insbesondere möchte ich Christophe und Florian für eine wunderbare, lehrreiche Zeit danken, in der wir gemeinsam an zeitabhängiger *CP*-Verletzung arbeiten durften und damit den Grundstein dieser Arbeit legen konnten.

I would like to express my gratitude to the B2CC physics working group at CERN for the fruitful discussions upon presentations that helped to polish our results. I would like to thank Jeroen for assisting us and furthermore, our reviewers Sevda, Liming, Yasmine, Lars, Michael, and all other LHCb colleagues who helped to get these results ready for publication.

Vielen Dank Britta, dass du alles Organisatorische am Lehrstuhl im Griff hast und auch manchmal die ein oder andere Lebensweisheit weitergibst. Ebenso danke ich Johannes A. und Robert für ihre kompetente Unterstützung und die vielen hilfreichen Tipps organisatorischer und fachlicher Art. Dem gesamten Lehrstuhl Experimentelle Physik V möchte ich rückblickend für einen tollen Zusammenhalt, eine hervorragende Arbeitsatmosphäre und eine alles in Allem spaßige, bereichernde Zeit danken.

Zu guter Letzt danke ich in besonderem Maße meiner Familie, sowie Hannah, die während der ersten Hälfte für mich da war, und Laura, die mir in der zweiten Hälfte den Rücken freihielt.