technische universität
dortmund

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

Improvement of the LHCb Measurement of $\sin(2\beta)$ in Additional B^0 Decay Modes

Vanessa Müller
geboren in Gelsenkirchen
am 08. September 1989

Lehrstuhl für Experimentelle Physik V
Fakultät Physik
Technische Universität Dortmund
Dortmund, im Mai 2019

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

Datum des Einreichens der Arbeit:	06. Mai 2019
Datum der mündlichen Prüfung:	04. September 2019

Erstgutachter:	Prof. Dr. Bernhard Spaan
Zweitgutachter:	Prof. Dr. Kevin Kröninger

Vorsitzender der Prüfungskommission:	Prof. Dr. Mirco Cinchetti
Vertreter der wissenschaftlichen Mitarbeiter:	Dr. Michael Paulus

Abstract

This thesis presents the measurement of decay-time-dependent *CP* violation in the decays of $B^0 \to \psi(2S) K^0_S$ and $B^0 \to J/\psi\, K^0_S$ using data collected with the LHCb experiment in pp collisions at centre-of-mass energies of 7 TeV and 8 TeV, corresponding to a total integrated luminosity of $3\,\text{fb}^{-1}$. In both considered decay modes, the K^0_S meson is reconstructed from two charged pions, while the $\psi(2S)$ meson is reconstructed from two muons, whereas the J/ψ meson is reconstructed from two electrons. This kind of decays makes it possible to measure the *CP*-violation observables S and C to be

$$\begin{aligned} S\big(B^0 \to \psi(2S) K^0_S\big) &= 0.84 \pm 0.10 \pm 0.01\,, \\ C\big(B^0 \to \psi(2S) K^0_S\big) &= -0.05 \pm 0.10 \pm 0.01\,, \\ S\big(B^0 \to J/\psi\, K^0_S\big) &= 0.83 \pm 0.08 \pm 0.01\,, \\ C\big(B^0 \to J/\psi\, K^0_S\big) &= 0.12 \pm 0.07 \pm 0.02\,, \end{aligned}$$

where the first uncertainties are statistical and the second are systematic. The results for both decay modes are compatible with each other and with the previous LHCb measurement using $B^0 \to J/\psi\, K^0_S$ decays, where the J/ψ meson is reconstructed from two muons. An average of all LHCb measurements of $B^0 \to [c\bar{c}] K^0_S$ decays, under the assumption, that higher-order contributions to the decay amplitudes are negligible, results in

$$\begin{aligned} S(B^0 \to [c\bar{c}] K^0_S) &= 0.760 \pm 0.034\,, \\ C(B^0 \to [c\bar{c}] K^0_S) &= -0.017 \pm 0.029\,, \end{aligned}$$

where the uncertainties include statistical and systematic contributions. The presented measurement enhances the precision of S at LHCb by 20 % and also improves the world average.

Kurzfassung

Diese Arbeit stellt die Messung der zerfallszeitabhängigen *CP*-Verletzung in den Zerfällen $B^0 \to \psi(2S) K^0_S$ und $B^0 \to J/\psi\, K^0_S$ vor, unter Verwendung von Daten des LHCb-Experiments, die in pp-Kollisionen bei Schwerpunktsenergien von 7 TeV und 8 TeV aufgenommen wurden, was einer integrierten Luminosität von $3\,\text{fb}^{-1}$ entspricht. In den beiden betrachteten Zerfallskanälen wird das K^0_S-Meson aus zwei geladenen Pionen rekonstruiert, während das $\psi(2S)$-Meson aus zwei Myonen rekonstruiert wird, wohingegen das J/ψ-Meson aus zwei Elektronen rekonstruiert wird. Diese Art von Zerfällen ermöglicht es die *CP*-Verletzungs-

observablen S und C sehr präzise zu folgenden Werten

$$\begin{aligned} S\left(B^0 \to \psi(2S)K^0_S\right) &= \quad 0.84 \pm 0.10 \pm 0.01\,, \\ C\left(B^0 \to \psi(2S)K^0_S\right) &= -0.05 \pm 0.10 \pm 0.01\,, \\ S\left(B^0 \to J/\psi\, K^0_S\right) &= \quad 0.83 \pm 0.08 \pm 0.01\,, \\ C\left(B^0 \to J/\psi\, K^0_S\right) &= \quad 0.12 \pm 0.07 \pm 0.02\,, \end{aligned}$$

zu bestimmen. Dabei sind die ersten Unsicherheiten statistisch und die zweiten systematisch. Die Ergebnisse beider Zerfallskanäle sind sowohl untereinander kompatibel, als auch mit dem Ergebnis einer früheren LHCb-Messung in dem Zerfallskanal $B^0 \to J/\psi\, K^0_S$, bei dem das J/ψ-Meson aus zwei Myonen rekonstruiert wurde. Eine Kombination aller LHCb-Messungen, die $B^0 \to [c\bar{c}]K^0_S$-Zerfälle betrachtet haben, wird unter der Annahme ermittelt, dass Beiträge höherer Ordnung zu den Zerfallsamplituden zu vernachlässigen sind. Das Ergebnis ist

$$\begin{aligned} S(B^0 \to [c\bar{c}]K^0_S) &= \quad 0.760 \pm 0.034\,, \\ C(B^0 \to [c\bar{c}]K^0_S) &= -0.017 \pm 0.029\,, \end{aligned}$$

dabei beinhalten die Unsicherheiten sowohl statistische als auch systematische Beiträge. Die vorgestellte Messungen verbessert die Präzision auf den Parameter S bei LHCb um 20 % und verringert die Unsicherheit des Weltmittelwerts.

Contents

1 Introduction

The Standard Model of particle physics, SM, tries to explain the structure of the universe on the smallest scale. Although it is well tested and successful, it is still incomplete and cannot give a comprehensive explanation of the universe. At the origin of the universe, matter and antimatter must have been produced in equal parts, according to the Big Bang theory. But today's universe is matter dominated without any indication of antimatter present. The theory of Baryogenesis tries to find an explanation for this baryon asymmetry. The physicist Andrei Sakharov proposed three necessary conditions [1] for the observed asymmetry in 1967; first, the violation of the Baryon number conservation, second, the violation of the C and CP symmetry, and third the deviation from the thermal equilibrium at some point in the history of the universe.

Symmetries play a distinctive role in physics and especially in the SM. According to the theorem by Emmy Noether [2], any continuous symmetry operation infers a conservation law. Additional to these continuous symmetries, there exist three discrete symmetries in the SM: parity, P, charge conjugation, C, and time-reversal, T. Nature is supposed to be invariant under the combination of all three transformations, CPT [2], while a system is allowed to behave differently under the individual transformations. In 1956 T. D. Lee and C. N. Yang stated, that the weak interaction violates parity to its maximum [3], which was experimentally confirmed later by C. S. Wu [4]. During the same period it was discovered, that C parity is violated in the weak interaction as well. In 1964, the violation of the combination of C and P, CP, could likewise be confirmed in neutral kaon decays [5] by Christenson, Cronin, Fitch, and Turlay. The CP violation in the neutral B-meson system had not been observed until 2001 by the B factories, BaBar and Belle, in $B^0 \to J/\psi K^0_S$ decays [6, 7]. It arises in the SM from a single weak phase introduced in the quark-mixing matrix, the Cabbibo-Kobayashi-Maskawa, CKM, matrix. This matrix is complex, unitary and the resulting unitary conditions can be represented as triangles in the complex plane.

As the CP operation transforms particles into antiparticles, a violation of this symmetry is of great interest when considering Baryogenesis. But the observed amount of CP violation described in the SM is not enough to explain the matter-antimatter asymmetry at this point. Thus, it is vital to search for New Physics beyond the SM and to test the SM itself by conducting precision measurements; for example by testing the consistency of the unitarity of the CKM matrix and thereby measuring its properties very precisely. If one was to find inconsistencies in the CKM matrix, which is a fundamental building block of the SM, it would hint to New Physics and an incomplete and inconsistent SM.

The Large Hadron Collider beauty, LHCb, experiment, located at the Large Hadron Collider, LHC, at the European Organization for Nuclear Research, CERN, focuses on

precision measurements of *CP*-violation observables and searches for (very) rare decays in the beauty and charm sector. In contrast to the *B* factories, which were operating at electron-positron colliders, the LHCb experiment uses the very high production rates of b and c quarks produced in the hadronic environment of the pp collider, but also has to manage the higher track multiplicities in this environment.

The "golden" modes for measuring *CP* violation are $b \to [c\bar{c}]s$ decays, as higher-order contributions, that could introduce additional strong and weak phases in the considered decay amplitudes, are expected to be small. This allows to measure $\sin(2\beta)$, which is a parameter derived from the CKM matrix, very precisely. The parameter β can also be identified as one of the angles of the most prominent CKM triangle. The parameter $\sin(2\beta)$ can be related to the decay-time-dependent *CP* asymmetry given by difference of the decay-time-dependent decay widths of a meson produced as a B^0 meson to a *CP*-final state and of a meson produced as a $\overline{B}^0$ to the same final state, normalized to the sum of both decay widths. This asymmetry can be determined by an unbinned maximum-likelihood fit to the decay time of the *B* meson, including the necessary information of the production flavour, which is given by the flavour tagging and considering the experimental conditions. The LHCb experiment already measured $\sin(2\beta)$ in $B^0 \to J/\psi K^0_S$ decays to a very high precision using data collected in the LHC's run I [8]. To exploit this existing dataset further and to increase the precision even more on the *CP* observables, it is of interest to consider additional charmonium resonances and different final states.

In this thesis the measurement of the decay-time-dependent *CP* violation in the decays $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi K^0_S$ using data collected with the LHCb experiment is presented. In both considered decay modes, the K^0_S meson is reconstructed from two charged pions, the J/ψ is reconstructed from two electrons, whereas the $\psi(2S)$ is reconstructed from two muons. This analysis is the first decay-time-dependent measurement at a hadron collider of the decay $B^0 \to \psi(2S)K^0_S$ and the first that uses electrons in the final state. The used data was collected during run I of the LHC during the years 2011 and 2012 in pp collisions at centre-of-mass energies of 7 TeV and 8 TeV, corresponding to a total integrated luminosity of $3\,\text{fb}^{-1}$. Besides the *CP*-violation measurement, an additional study is performed, where a preliminary and unofficial value for the ratio of branching fractions of the decays $B^0_s \to \psi(2S)K^0_S$ and $B^0 \to \psi(2S)K^0_S$, $\mathscr{B}(B^0_s)/\mathscr{B}(B^0)$, is determined.

Chapter 2 presents the theoretical basics of the analysis and motivates the measurement. It is followed by the presentation of the LHCb experiment in Ch. 3, which introduces in particular the components that are important for the final states reconstructed in this analysis. In the following chapters, the data preparation for the $B^0 \to \psi(2S)K^0_S$ decay is introduced in detail, then the preparation in the case of the $B^0 \to J/\psi K^0_S$ channel is briefly outlined in Sec. 4.4. In Sec. 4.2.1 the preliminary result for the ratio of branching fractions is presented. It is followed by the introduction of the measurement of the *CP*-violation observables in Sec. 4.5, the determination of the systematic uncertainties in Sec. 4.7, and the combination of the determined results with a previous LHCb measurement in Sec. 4.8. Finally, the results are summarised in Ch. 5 and an outlook is given.

Collaboration

The work presented in this thesis was published in 2017 with two more contact authors, Ramon Niet and Patrick Mackowiak, by the LHCb collaboration as

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

This was only possible to achieve in close collaboration with other members of the LHCb collaboration, especially with Ramon Niet and Julian Wishahi. Ramon Niet was responsible for the full preparation of the $B^0 \to J/\psi K^0_S$ dataset. He also implemented the frameworks for the decay-time-dependent *CP* fit, the evaluation of the systematic uncertainties, and the combination of the results with the *GammaCombo* framework [10], which were developed in close collaboration with and also cross-checked by me. He published the results within his thesis

R. Niet, *Measurement of CP Violation in* $B^0 \to [c\bar{c}]K^0_S$ *Transitions at LHCb*, Ph. D. thesis, May 2018, DOI: 10.17877/DE290R-19128

at TU Dortmund University. Besides this, more people helped to make this analysis possible. Alex Seuthe did the studies on flavour tagging in case of the $B^0 \to J/\psi K^0_S$ decay channel as part of his master thesis [12], Alex Birnkraut and Frank Meier provided cross-checks of the implementation of the decay-time-dependent fit, and the kaon-regeneration effects were prepared with the help of Jeroen van Tilburg and Julian Wishahi. Furthermore, Patrick Mackowiak supported the studies of the physics backgrounds [13] in the reconstructed mass distribution of the B meson. Moreover, the supervised studies done by bachelor students, Konrad Mielke [14] and Lukas Nickel [15], as well as by master student Christopher Hasenberg [16], provided useful contributions on the Bremsstrahlung corrections, the development of the figure of merit, and on incorrectly associated PVs, respectively.

2 *CP* Violation in the Standard Model

This chapter gives an overview of *CP* violation in the Standard Model of particle physics. It begins by introducing the general theory of this model in Sec. 2.1, then expands on *CP* violation in Sec. 2.2 and specifically on *CP* violation in the class $B^0 \to [c\bar{c}]K^0_S$ decay modes[1], which are essential for this thesis in Sec. 2.2.4. The following sections are based on Refs. [2], [17], [18].

2.1 Standard Model of Particle Physics

Besides the elementary constituents of matter and antimatter the Standard Model of particle physics describes also the fundamental interaction between those. It comprises 12 fermions, which are divided into six quarks and six leptons, as well as their antiparticles[2], that have opposite charge-related quantum numbers. Fermions are elementary particles with a spin of half-integer [3] and are divided into three generations or families in ascending order of mass.

The quark families contain an up-type and a down-type quark, which are categorized by their electric charge. The up-type quarks have an electric charge of $+2/3$ of the elementary charge, while the down-type quarks have an electric charge of $-1/3$ of the elementary charge. Besides the electric charge, quarks also carry the so-called *colour charge*, which can occur in six different states, red, green, and blue and in the anticolours. Due to the so-called *confinement* quarks cannot propagate freely and occur only in bound states, which are by construction colourless, as a combination of all three colours or all three anticolours as well as a combination of an anticolour and a colour sum up to be colourless. These composite particles are called *hadrons* and are sub-divided into *baryons*, which have a half-integer spin, *mesons*, which have an integer spin, and into the recently discovered spin-0 *tetraquarks* [19, 20] and the half-integer-spin *pentaquarks* [21]. Mesons contain a quark and an antiquark, while the baryons are made of three quarks. The tetra- and pentaquarks are the four- and five-quark bound states, respectively.

The first quark family contains the up, u, and the down quark, d. The common matter is mostly made of protons, uud, and neutrons, udd, which are thus baryons made of quarks from this first family. The second family contains the charm, c, and the strange quark, s, and the last family contains the top, t, and beauty quark, b, also named bottom quark.

[1] Throughout this thesis the inclusion of charge-conjugate processes is implied, unless otherwise noted. The notation B^0 refers to a neutral B meson containing a $\bar{b}$ and a d including the charge- conjugate state as well.

[2] Antiparticles are symbolized either by a bar over the symbol or by an opposite electric charge at the symbol.

[3] Throughout the thesis natural units are used, so that the Planck constant, $\hbar$, and the speed of light, c, are set to one, if not stated otherwise.

The lepton families contain a negatively charged particle, carrying the negative single elementary charge, and a neutral, massless, neutrino. The charged particles are the electron, e, the muon, μ, and the tau, τ, with their corresponding neutrinos, the electron neutrino, ν_e, the muon neutrino, ν_μ, and the tau neutrino, ν_τ.

Furthermore the SM describes three of the fundamental forces and its mediators; the so-called *gauge bosons*, which have a spin of 1. The mediators of the strong interaction are eight so-called *gluons*, g, and couple to the color charge. They carry different colour charge configurations. The massive mediators of the weak force are the two charged $W^\pm$ bosons and the neutral Z boson and couple to the weak and the hyper charge. The mediator of the electromagnetic interaction is the massless photon, γ, and it couples to the electric charge. The strong and the weak interaction have a rather short range of around $1 \cdot 10^{-13}$ m [2], while the electromagnetic interaction has an infinite range. At this point gravity is not included in the SM.

The Higgs boson, H^0, completes the SM. It was proposed in 1964 by François Englert, Robert Brout, and Peter Higgs [22–24] and discovered as the last missing piece of the SM in 2012 at CERN [25]. Through the Brout-Englert-Higgs mechanism, which describes the interaction with the Higgs field, the gauge bosons obtain their masses. The symmetry breaking has its origin in the choice of the vacuum expectation value. Through the Yukawa coupling to the Higgs field the fermions get their masses [26]. According to the *Goldstone theorem* the spontaneous breaking of a continuous, global symmetry results in one or more scalar particles with a spin of 0. With a suitable calibration these so-called *Goldstone bosons* can be eliminated and a massive, scalar particle remains, the Higgs boson. The fermions, the gauge bosons, as well as the Higgs boson are listed in Fig. 2.1.

According to a theorem by the mathematician Emmy Noether [2], conserved quantities, such as charge or energy, result from continuous symmetries of a physical system. Symmetry means the invariance of a system under a certain transformation. Additionally to these continuous symmetries there exist three discrete symmetries in the SM: the P, C, and T parity. The parity transformation, P, changes the sign of all spatial coordinates, which corresponds to the inversion of the three coordinate axes through the origin, this changes the handedness of a system's axes. For example it thus transforms a left-handed into a right-handed fermion. The weak interaction violates P to a maximum, *i.e.* only left-handed neutrinos and right-handed antineutrinos exist. The charge conjugation, C, reverses all charge-related quantum numbers and in combination with P exchanges a particle with its antiparticle. It replaces a particle by its antiparticle and reverses its momentum and its *helicity*, which is the sign of the projection of the spin vector onto the momentum vector. Time reversal, T, reverses the time coordinate. The consecutive application of all the operations, CPT, needs to be preserved, otherwise the quantum field theory would not be Lorentz invariant. This invariance of a system under a combination of all three transformations is captured by the CPT theorem. A system under the individually transformations, C, P, and T, can behave differently and thus the symmetry can be violated. The theoretical physicists T. D. Lee and C. N. Yang won the Nobel price in 1957 for the idea that parity must be violated in the weak interaction [3]. In

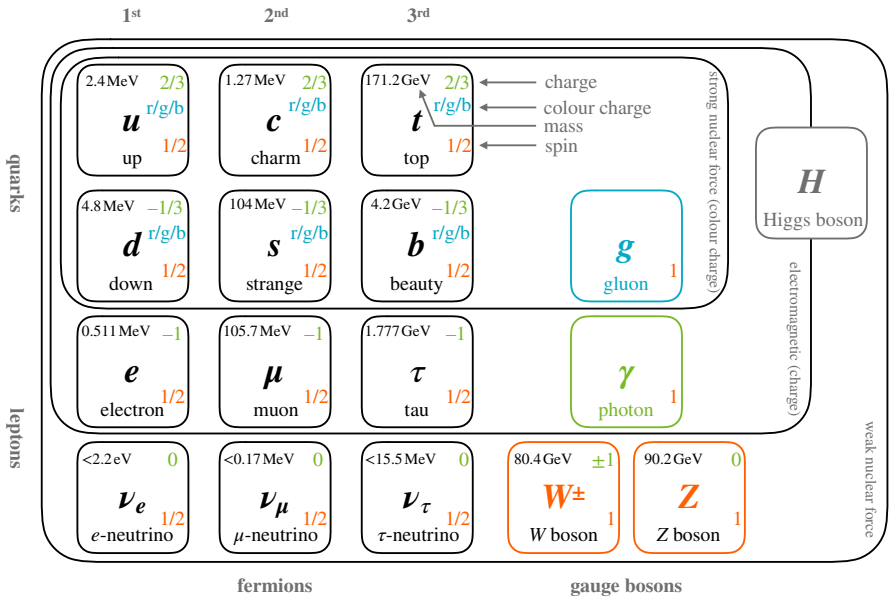

Figure 2.1: An overview of the fundamental particles in the Standard Model and their properties. The fermions, quarks and leptons, appear in three generations. The properties are taken from Ref. [27]

the year 1956 C. S. Wu could confirm this prediction experimentally [4] . During the same period it was also discovered that *C* parity is also violated in the weak interaction. In 1964 the violation of the combination of *C* and *P*, *CP*, could also be confirmed in neutral kaon decays [5].

Although the SM is a very well tested theory and seems to be consistent with experimental findings, there are still some open questions for which the SM has no real answers. Until this point it does not include gravity or it cannot explain why neutrinos in fact are not massless, and why exactly three different generations exist. Also, the imbalance of matter and antimatter in the observed universe cannot be explained by the experimentally measured amount of *CP* violation as of this point. Furthermore the SM only describes around 5 % of the universe's energy, which are the ordinary matter and energy. Around 24 % should be dark matter and the remaining 71 % are an unknown form of energy known as *dark energy*.

2.2 *CP* Violation in the Standard Model

In nature, all fermions, *i.e.* quarks and leptons, have a different *flavour*, which is characterized by the flavour-quantum numbers. These numbers differ for the various fermions. Flavour physics refers to interactions that distinguish between the different flavours in the SM; these comprise weak and Yukawa interactions.

2.2.1 Quark-Mixing Matrix

The *CP* violation in the SM has its origin in a single phase of the quark-mixing matrix considering three fermion families. The following section closely follows Refs. [27, 28].

The masses and mixing of the quarks originate from the Yukawa interaction with the Higgs field. The Lagrangian, $\mathscr{L}_Y$, can be written as

$$\mathscr{L}_Y = -\boldsymbol{Y}^d_{ij}\overline{Q^I_{\mathrm{L}i}}\phi d^I_{\mathrm{R}i} - \boldsymbol{Y}^u_{ij}\overline{Q^I_{\mathrm{L}i}}\epsilon\phi^* u^I_{\mathrm{R}j} + \mathrm{h.c.}, \tag{2.1}$$

when introducing a scalar Higgs field, ϕ. In equation Eq. (2.1) $\boldsymbol{Y}^{\{u,d\}}$ are 3×3 complex matrices, which comprise the Yukawa couplings, i and j are generation labels and ϵ is the 2×2 antisymmetric tensor. Furthermore Q^I_{L} denotes the left-handed quark doublets, and $\{d,u\}^I_{\mathrm{R}}$ are the right-handed down- and up-type quark singlets in the weak-eigenstate basis. Due to the spontaneous symmetry breaking and when ϕ achieves a vacuum expectation value of $\langle\phi\rangle = (0, v/\sqrt{2})$ the quarks, the $W^\pm$ bosons, and Z bosons obtain their masses.

The Yukawa matrices are non-diagonal in the weak interaction basis and the physical states can be obtained by diagonalizing $\boldsymbol{Y}^q$ by four unitary matrices, $\boldsymbol{V}^q_{\{\mathrm{L,R}\}}$, as

$$\boldsymbol{M}^q_{\mathrm{diag}} = \boldsymbol{V}^q_{\mathrm{L}}\boldsymbol{Y}^q\boldsymbol{V}^{q\dagger}_{\mathrm{R}}(v/\sqrt{2}) \quad \text{with } q \in \{u,d\}. \tag{2.2}$$

As a direct consequence of this, the quark fields are also transformed into their mass-eigenstate basis and thus the charged-current interactions couple to the physical states of the quarks. The mass eigenstates q can be obtained by

$$q_{\{\mathrm{L,R}\}} = \boldsymbol{V}^q_{\{\mathrm{L,R}\}}q_{\{\mathrm{L,R}\}} \quad \text{with } q \in \{u,d\}. \tag{2.3}$$

And the couplings are given by

$$-\frac{g}{\sqrt{2}}\left[(\overline{u},\overline{c},\overline{t})_{\mathrm{L}}\gamma^\mu W^+_\mu \boldsymbol{V}_{\mathrm{CKM}}\begin{pmatrix}d\\s\\b\end{pmatrix}_{\mathrm{L}} + (\overline{d},\overline{s},\overline{b})_{\mathrm{L}}\gamma^\mu W^-_\mu \boldsymbol{V}^\dagger_{\mathrm{CKM}}\begin{pmatrix}u\\c\\t\end{pmatrix}_{\mathrm{L}}\right], \tag{2.4}$$

with the weak charge, g, and the Cabbibo-Kobayashi-Maskawa matrix, $\boldsymbol{V}_{\mathrm{CKM}}$, as the transformation matrix. This matrix is a 3×3 complex and unitary matrix, $V_{\mathrm{CKM}}V^\dagger_{\mathrm{CKM}} = \mathbb{1}$,

$$\begin{pmatrix}d'\\s'\\b'\end{pmatrix}_{\mathrm{L}} = \boldsymbol{V}_{\mathrm{CKM}}\cdot\begin{pmatrix}d\\s\\b\end{pmatrix}_{\mathrm{L}} = \begin{pmatrix}V_{ud} & V_{us} & V_{ub}\\ V_{cd} & V_{cs} & V_{cb}\\ V_{td} & V_{ts} & V_{tb}\end{pmatrix}\cdot\begin{pmatrix}d\\s\\b\end{pmatrix}_{\mathrm{L}}, \tag{2.5}$$

and transforms the mass eigenstates of the down-type quarks, q, into the eigenstates of the weak interaction, q'. The squared matrix elements, $|V_{ij}|^2$, represent the transition

probabilities of a quark i into a quark j. The CKM matrix can be parametrized by nine free parameters, which can be reduced using the unitarity conditions,

$$\sum\nolimits_i V_{ij} V_{ik}^* = \delta_{jk} \quad \text{and} \quad \sum\nolimits_j V_{ij} V_{kj}^* = \delta_{ik}, \tag{2.6}$$

and a redefinition of the quark phases to three real mixing angles and one complex phase. This complex phase is responsible for *CP* violation in the SM. The standard parametrization uses three Euler angles, $\theta_{\{12,23,13\}}$, and a *CP*-violating phase, δ_{13}. The trigonometric functions of these angles are denoted $c_{ij} = \cos(\theta_{ij})$ in case of cosines and in case of sines as $s_{ij} = \sin(\theta_{ij})$, which leads to the following parametrization:

$$\boldsymbol{V}_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta 13} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta 13} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}. \tag{2.7}$$

This representation clearly shows a hierarchy of the off-diagonal elements and is an exact parametrization of the matrix. Another common representation of the CKM matrix is the Wolfenstein parametrization [29] in which the matrix can be described in the third order of $\lambda = 0.23$ as

$$\boldsymbol{V}_{\text{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{2.8}$$

with

$$\lambda = s_{12} = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \; A\lambda^2 = s_{23} = \lambda \left| \frac{V_{cb}}{V_{us}} \right|, \; V_{ub}^* = s_{13}e^{i\delta} = A\lambda^3(\rho - i\eta), \tag{2.9}$$

where the three real parameters are A, λ, ρ and the complex phase is identified by η. In contrast to the standard parametrization, this is only an approximation. The parameters are measured to be [30] around:

$$\lambda \approx 0.23, \; A \approx 0.82, \; \rho \approx 0.13, \; \eta \approx 0.26. \tag{2.10}$$

Due to the unitarity of the CKM matrix, nine conditions can be formulated as given in Eq. (2.6). The six vanishing combinations of the unitarity conditions can be represented as triangles in the complex $(\overline{\rho}, \overline{\eta})$ plane. The measured uncertainties of the parameters describing the CKM matrix, can lead to degenerated triangles, but the area of all triangles is equal to half of the Jarlskog invariant [31], J, nonetheless. It is a measure of the amount of *CP* violation and that is given by

$$J = \pm \text{Im}\left(V_{ij} V_{kl} V_{il}^* V_{kj}^* \right) \quad \text{for } i \neq j, k \neq l. \tag{2.11}$$

It is determined to be $|J| \approx \lambda^6 A^2 \eta \approx 3 \cdot 10^{-5}$ in the Wolfenstein parametrization. In this parametrization it also becomes obvious, that transitions within one quark family are more likely than between the families, as the absolute values of the diagonal matrix elements are larger compared to the off-diagonal elements.

One of the most common unitarity triangles results from the following condition

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{2.12}$$

which contains matrix elements that are important for b-meson decays. A more detailed description of *CP* violation in b-meson decays follows in Sec. 2.2.4. Normalization of the unitarity condition, *i.e.* by dividing each side by the best-known one, here $V_{cd}V_{cb}^*$, results in the triangle seen in Fig. 2.2. Consequently, the determination of the triangle depends on the sole determination of the free apex. The angles of the unitarity triangle are defined as

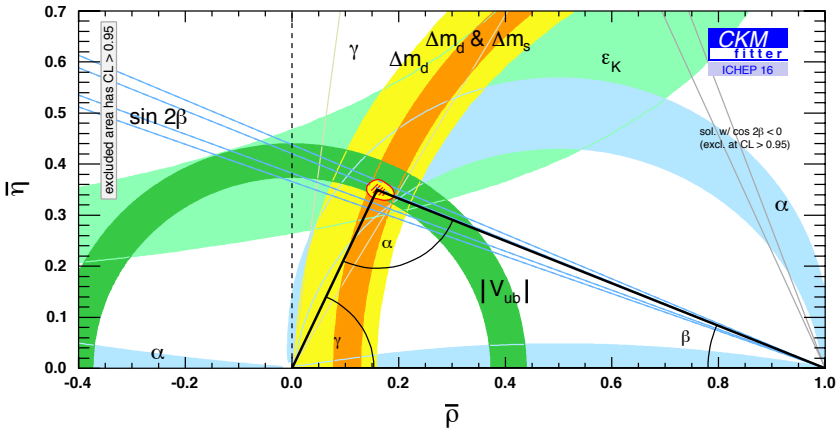

Figure 2.2: The most prominent CKM triangle (see Eq. (2.12)) including constraints represented by the coloured bands zoomed in the complex $(\bar{\rho}, \bar{\eta})$ plane. The red hashed region represents the global combination of all measurements and corresponds to 68 % CL as determined by the CKM fitter group [32].

$$\beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \; \alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \; \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{2.13}$$

Measurement of *CP* violation in b-meson decays gives direct access to these angles. The over-constraining of the triangle by various measurements serves to improve the determination of the CKM elements and to reveal effects of New Physics beyond the SM.

2.2.2 Neutral b-Meson Decays and Mixing

An important aspect in the measurement of *CP* violation in the sector of b mesons is that neutral mesons can oscillate between particle, M^0, and antiparticle, $\bar{M}^0$, before they decay

due to the weak interaction and the flavour-changing charged currents. The flavour eigenstates $|B^0\rangle = |\bar{b}d\rangle$ and $|\overline{B}^0\rangle = |b\bar{d}\rangle$ are considered. The following section closely follows Refs. [18] and [28].

The decay amplitudes of a B^0 meson and its *CP* conjugate, $\overline{B}^0$, into a final state, $|f\rangle$, or its *CP* conjugate, $|\bar{f}\rangle$, are defined by

$$\begin{aligned} A_f &= \langle f|H|B^0\rangle, & \overline{A}_f &= \langle f|H|\overline{B}^0\rangle, \\ A_{\bar{f}} &= \langle \bar{f}|H|B^0\rangle, & \overline{A}_{\bar{f}} &= \langle \bar{f}|H|\overline{B}^0\rangle, \end{aligned} \tag{2.14}$$

where H is the Hamiltonian of the weak interaction. There exist two flavour eigenstates with a specific quark content and there are mass eigenstates, *i.e.* the eigenstates of the Hamiltonian, with a specific mass and lifetime. As the mass and flavour eigenstates differ, the flavour eigenstates are mixed with each other when propagating through space.

The flavour eigenstates, B^0 and $\overline{B}^0$, and the final states, f and $\bar{f}$, are connected via *CP* transformations,

$$\begin{aligned} CP|B^0\rangle &= e^{i\xi_B}|\overline{B}^0\rangle \quad \text{and} \quad CP|\overline{B}^0\rangle = e^{-i\xi_B}|B^0\rangle, \\ CP|f\rangle &= e^{i\xi_f}|\bar{f}\rangle \quad \text{and} \quad CP|\bar{f}\rangle = e^{-i\xi_f}|f\rangle, \end{aligned} \tag{2.15}$$

where ξ_B and ξ_f are arbitrary phases, under the assumption of $(CP)^2 = \mathbb{1}$. In case *CP* symmetry is conserved, $\overline{A}_{\bar{f}} = e^{i(\xi_f - \xi_B)}A_f$ follows, which means that $|\overline{A}_{\bar{f}}| = |A_f|$.

In the SM flavour-changing neutral currents are forbidden at tree level, so that the transition of a B^0 into a $\overline{B}^0$ can be described by loops involving quarks and $W^\pm$ bosons at the lowest order. The corresponding Feynman diagrams are shown in Fig. 2.3.

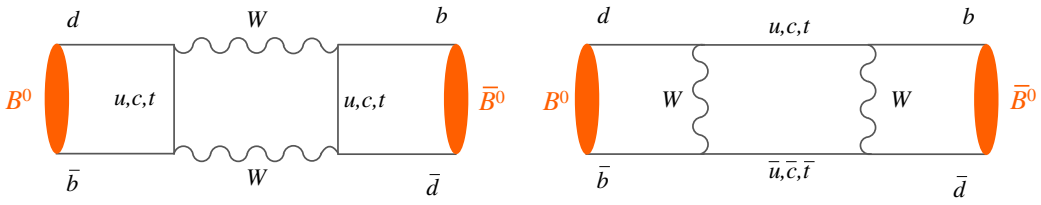

Figure 2.3: Feynman diagrams describing the B^0-$\overline{B}^0$ oscillation at the lowest order. Two charged currents and two up-type quarks are involved in this process, although it is dominated by a top or an antitop quark, respectively.

The time evolution of an arbitrary linear combination of the flavour eigenstates,

$$a(t)|B^0\rangle + b(t)|\overline{B}^0\rangle, \tag{2.16}$$

is described by the time-dependent Schrödinger equation in the Wigner-Weisskopf approxi-

mation [33, 34]

$$
\begin{aligned}
i\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \boldsymbol{H}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} &\equiv \left(\boldsymbol{M} - \frac{i}{2}\boldsymbol{\Gamma}\right)\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \\
&= \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix},
\end{aligned}
\tag{2.17}
$$

where the Hamiltonian, $\boldsymbol{H}$, is a 2×2 non-Hermitian matrix, otherwise the mesons would just oscillate and not decay. Whereas, $\boldsymbol{M}$ and $\boldsymbol{\Gamma}$ are Hermitian, $\boldsymbol{M} = \boldsymbol{M}^\dagger$ and $\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^\dagger$. Due to the invariance of CPT $H_{11} = H_{22}$, thus $M_{11} = M_{22} \equiv m$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$, and $M_{12} = M_{21}^*$ as well as $\Gamma_{12} = \Gamma_{21}^*$. The flavour-changing transitions $B^0 \leftrightarrow \overline{B}^0$ are described by the off-diagonal elements in the Hamiltonian. The matrix element M_{12} represents transitions via dispersive, off-shell, intermediate states, whereas the element Γ_{12} represents transitions via absorptive, on-shell, intermediate states. The diagonal elements of $\boldsymbol{M}$ and $\boldsymbol{\Gamma}$ are associated with flavour-conserving transitions. Taking into account the introduced identities, the Hamiltonian simplifies to

$$
\boldsymbol{H} = \left(\boldsymbol{M} - \frac{i}{2}\boldsymbol{\Gamma}\right) = \begin{pmatrix} m - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & m - \frac{i}{2}\Gamma \end{pmatrix}. \tag{2.18}
$$

The eigenvalues, μ_L and μ_H, of this matrix are calculated via $\det(\boldsymbol{H} - \lambda\mathbb{1}) = 0$ and are given by

$$
\mu_{H,L} = m - \frac{i}{2}\Gamma \pm \sqrt{\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)}, \tag{2.19}
$$

sorting by imaginary and real part, and with

$$
S \equiv \sqrt{\left(M_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)}, \tag{2.20}
$$

the eigenvalues can also be defined as

$$
\begin{aligned}
\mu_H \equiv m_H - \frac{i}{2}\Gamma_H &= m + \mathrm{Re}(S) - \frac{i}{2}\left(\Gamma + \mathrm{Im}(S)\right), \\
\mu_L \equiv m_L - \frac{i}{2}\Gamma_L &= m - \mathrm{Re}(S) - \frac{i}{2}\left(\Gamma - \mathrm{Im}(S)\right),
\end{aligned}
\tag{2.21}
$$

The mass eigenstates in the B system, B_L and B_H, where L stands for *light* and H for *heavy*, are given by a linear combination of the flavour eigenstates, due to the non-zero off-diagonal matrix elements in $\boldsymbol{H}$ as

$$
\begin{aligned}
\left|B_H\right\rangle &= p\left|B^0\right\rangle - q\left|\overline{B}^0\right\rangle, \\
\left|B_L\right\rangle &= p\left|B^0\right\rangle + q\left|\overline{B}^0\right\rangle,
\end{aligned}
\tag{2.22}
$$

here q and p are complex coefficients and follow the normalization condition, $|q|^2 + |p|^2 = 1$, and their ratio can be defined using the matrix elements of $\boldsymbol{H}$ by

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \tag{2.23}$$

The time evolution of an initially, *i.e.* at $t = 0$, produced B^0 and of an initially produced $\overline{B}^0$ are given by

$$\begin{aligned} \left|B^0(t)\right\rangle &= g_+(t)\left|B^0\right\rangle - \frac{q}{p} g_-(t)\left|\overline{B}^0\right\rangle, \\ \left|\overline{B}^0(t)\right\rangle &= g_+(t)\left|\overline{B}^0\right\rangle - \frac{p}{q} g_-(t)\left|B^0\right\rangle, \end{aligned} \tag{2.24}$$

with

$$\begin{aligned} g_\pm(t) &= \frac{1}{2}\left(e^{-i\mu_H t} \pm e^{-i\mu_L t}\right), \\ &= \frac{1}{2}\left(e^{-im_H t} e^{-\frac{1}{2}\Gamma_H t} \pm e^{-im_L t} e^{-\frac{1}{2}\Gamma_L t}\right), \end{aligned} \tag{2.25}$$

using Eq. (2.21). Furthermore, using Eq. (2.22) results in the following definitions

$$\begin{aligned} \left|B_L\right\rangle + \left|B_H\right\rangle &= \left(p\left|B^0\right\rangle + q\left|\overline{B}^0\right\rangle\right) + \left(pB^0 - q\left|\overline{B}^0\right\rangle\right) = 2p\left|B^0\right\rangle \\ \Rightarrow \left|B^0\right\rangle &= \frac{1}{2p}\left(\left|B_H\right\rangle + \left|B_L\right\rangle\right), \\ \left|B_L\right\rangle - \left|B_H\right\rangle &= \left(p\left|B^0\right\rangle + q\left|\overline{B}^0\right\rangle\right) - \left(p\left|B^0\right\rangle + q\left|\overline{B}^0\right\rangle\right) = 2q\left|\overline{B}^0\right\rangle \\ \Rightarrow \left|\overline{B}^0\right\rangle &= \frac{1}{2q}\left(\left|B_L\right\rangle - \left|B_H\right\rangle\right). \end{aligned} \tag{2.26}$$

The mass difference, Δm, and the decay-width difference, $\Delta\Gamma$, of the mass eigenstates are given by

$$\Delta m \equiv m_H - m_L = 2 \cdot \mathrm{Re}(S) \quad \text{and} \quad \Delta\Gamma \equiv \Gamma_H - \Gamma_L = 4 \cdot \mathrm{Im}(S). \tag{2.27}$$

using the definition from Eq. (2.20). The sign of the mass difference is positive by definition, while the sign of $\Delta\Gamma$ needs to be determined experimentally. Using Eq. (2.36) results in

$$\Delta m \approx 2\left|M_{12}\right|, \tag{2.28}$$

$$\Delta\Gamma \approx 2\left|\Gamma_{12}\right|\cos(\phi), \tag{2.29}$$

which allows to assume that $\Delta\Gamma$ is small. The average mass and decay width are defined as the arithmetic means as

$$m \equiv \frac{m_H + m_L}{2} \quad \text{and} \quad \Gamma \equiv \frac{\Gamma_H + \Gamma_L}{2}. \tag{2.30}$$

Computing the matrix-mixing element M_{12} for the neutral b-meson system, where the amplitudes are mainly dominated by the short-distance contribution arising from the box diagrams shown in Fig. 2.3, leads to

$$M_{12} = -\frac{G_F^2 m_W^2}{12\pi^2} f_{B_q}^2 m_{B_q} B_{B_q} \mathscr{F}^*. \tag{2.31}$$

Here G_F is the Fermi constant, m_W is the mass of the W boson; and f_{B_q}, m_{B_q}, B_{B_q} are the weak decay constant, the mass of the b meson and the bag parameter, respectively. The weak decay constant and the bag parameter describe corrections from non-perturbative QCD. The function $\mathscr{F}$ is defined as

$$\mathscr{F} = \eta_1 \lambda_c^2 S_0(x_c) + \eta_2 \lambda_t S_0(x_t) + 2\eta_3 \lambda_c \lambda_t S_0(x_c, x_t), \tag{2.32}$$

with combinations of matrix elements, $\lambda_\alpha = V_{\alpha b}^* V_{\alpha q}$, the ratio of masses, $x_\alpha = m_\alpha^2/m_W^2$, and perturbative QCD corrections, η_i. If $q = d$, λ_c and λ_t are of the same order of magnitude as λ^3 and if $q = s$, they both are of the same order of magnitude as λ^2. The function S_0 is the Inami-Lim function [35] and depends on ratios of the quark masses to the W-boson mass and shows a clear hierarchy

$$S_0(x_t) >> S_0(x_c, x_t) > S_0(x_c). \tag{2.33}$$

The top and charm quark enter at a comparable magnitude of around λ^3, while the up-quark loop is suppressed. This simplifies Eq. (2.32) to $\mathscr{F} \approx \eta_2 \lambda_t^2 S_0(x_t)$ and thus Eq. (2.31) to

$$M_{12} \approx -\frac{G_F^2 m_W^2}{12\pi^2} f_{B_q}^2 m_{B_q} B_{B_q} \eta_2 \left(V_{tb}^* V_{tq}\right)^2 S_0\left(\frac{m_t^2}{m_W^2}\right). \tag{2.34}$$

The matrix element Γ_{12} is given by

$$\Gamma_{12} = \sum_f \langle f \big| H \big| B_q^0 \rangle^* \langle f \big| H \big| \bar{B}_q^0 \rangle, \tag{2.35}$$

where f denotes the physical states to which b mesons can decay. It can be interpreted as the absorptive part of the box diagrams with intermediate charm or up quarks. As the top mass is much higher compared to the mass of the B_q^0 meson, neutral b mesons cannot decay into top-flavoured hadrons. Thus, the box diagrams including intermediate top quarks have a vanishing absorptive part. Resulting from that, the absorptive part of the box diagram contribution to Γ_{12} is dominated by the mass available in the decays of the neutral b mesons, *i.e.* $m_{B_q^0} \approx m_b$. Considering $M_{12} \approx S_0(x_t) \approx x_t \approx m_t^2$ results in the prediction of

$$\left|\frac{\Gamma_{12}}{M_{12}}\right| \approx \frac{m_b^2}{m_t^2} \approx 10^{-3}. \tag{2.36}$$

This means that *CP* violation in the mixing can be assumed as negligible.

The time evolution of the mass eigenstates can be written shortly using the eigenvalues like

$$\begin{aligned} \left|B_H\right\rangle(t) &= e^{-i\mu_H t}\left|B_H\right\rangle, \\ \left|B_L\right\rangle(t) &= e^{-i\mu_L t}\left|B_L\right\rangle. \end{aligned} \tag{2.37}$$

The relevant differential decay rates, *i.e.* the squared decay amplitudes, are given by

$$\begin{aligned} \Gamma(B^0(t) \to f) &= \left|\langle f\left|H\right|B^0\rangle\right|^2, \\ \Gamma(\overline{B}^0(t) \to f) &= \left|\langle f\left|H\right|\overline{B}^0\rangle\right|^2. \end{aligned} \tag{2.38}$$

Using the previously defined coefficients, $g_{\pm}(t)$, in Eq. (2.25) the decay rates can further be expressed by

$$\begin{aligned} \Gamma(B^0(t) \to f) = \left|\langle f\left|H\right|B^0\rangle\right|^2 &= \left|g_+(t)\langle f\left|H\right|B^0\rangle - \frac{q}{p}g_-(t)\langle f\left|H\right|\overline{B}^0\rangle\right|^2 \\ &= \left|A_f\right|^2 \cdot \left|g_+(t) - \frac{q}{p}\frac{\overline{A}_f}{A_f}g_-(t)\right|^2 = \left|A_f\right|^2 \cdot \left|g_+(t) - \lambda_f g_-(t)\right|^2 \\ &= \left|A_f\right|^2 \cdot \left(\left|g_+(t)\right|^2 + \left|\lambda_f\right|^2\left|g_-(t)\right|^2 + 2\mathrm{Re}\left[\lambda_f g_+^*(t)g_-(t)\right]\right), \end{aligned} \tag{2.39}$$

with

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f}, \tag{2.40}$$

for a B^0 meson produced at $t = 0$ to a final state, f. And analogously for a $\overline{B}^0$ meson produced at $t = 0$ to the same final state, f, the decay rate is given by

$$\begin{aligned} \Gamma(\overline{B}^0(t) \to f) = \left|\langle f\left|H\right|\overline{B}^0\rangle\right|^2 &= \left|g_+(t)\langle f\left|H\right|\overline{B}^0\rangle - \frac{q}{p}g_-(t)\langle f\left|H\right|B^0\rangle\right|^2 \\ &= \left|A_f\right|^2 \cdot \left|\frac{p}{q}\right|^2 \left|\lambda_f g_+(t) - g_-(t)\right|^2 \\ &= \left|A_f\right|^2 \left|\frac{p}{q}\right|^2 \cdot \left(\left|g_-(t)\right|^2 + \left|\lambda_f\right|^2\left|g_+(t)\right|^2 + 2\mathrm{Re}\left[\lambda_f g_+(t)g_-^*(t)\right]\right). \end{aligned} \tag{2.41}$$

In Eqs. (2.39) and (2.41) for the given decay rates terms proportional to $\left|A_f\right|^2$ represent decays without a flavour change, where an initially produced B^0 or $\overline{B}^0$ decays as a B^0 or $\overline{B}^0$, respectively. Terms proportional to $\left|A_f\right|^2 |p/q|^2$ represent decays, where the neutral meson decays after mixing into the opposite state as the initial state and terms proportional

to $g_{\pm}^{*}(t)g_{\mp}(t)$ describe the interference of the direct decay and the decay after mixing. These decay rates can be written in terms of hyperbolic and trigonometric functions as

$$\begin{aligned}\Gamma(B^0(t) \to f) &= \frac{1}{2}\left|A_f\right|^2 \cdot e^{-\Gamma t} \cdot \left(1 + \left|\lambda_f\right|^2\right) \\ &\quad \cdot \left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D\sinh\left(\frac{\Delta\Gamma}{2}t\right) + C\cos\left(\Delta m t\right) - S\sin\left(\Delta m t\right)\right), \\ \Gamma(\overline{B}^0(t) \to f) &= \frac{1}{2}\left|A_f\right|^2 \cdot \left|\frac{p}{q}\right|^2 \cdot e^{-\Gamma t} \cdot \left(1 + \left|\lambda_f\right|^2\right) \\ &\quad \cdot \left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D\sinh\left(\frac{\Delta\Gamma}{2}t\right) - C\cos\left(\Delta m t\right) + S\sin\left(\Delta m t\right)\right).\end{aligned} \tag{2.42}$$

The so-called *CP coefficients* are defined as

$$D = \frac{2\mathrm{Re}\left(\left|\lambda_f\right|\right)}{1 + \left|\lambda_f\right|^2}, \quad C = \frac{1 - \left|\lambda_f\right|^2}{1 + \left|\lambda_f\right|^2}, \quad \text{and} \quad S = \frac{2\mathrm{Im}\left(\left|\lambda_f\right|\right)}{1 + \left|\lambda_f\right|^2}, \tag{2.43}$$

and satisfy $D^2 + C^2 + S^2 = 1$.

2.2.3 Classification of *CP*-violating Effects

In neutral meson systems three types of *CP*-violating effects can be distinguished: direct *CP* violation, indirect *CP* violation, and *CP* violation in the direct decay and the decay after mixing.

Weak and strong phases are the only two types of phases that can enter in the decay amplitudes. The weak phases originate from the CKM matrix and have opposite sign for A_f and $\overline{A}_{\bar{f}}$. Strong phases occur in scattering processes and have the same sign for A_f and $\overline{A}_{\bar{f}}$. As the SM is a gauge-invariant theory, only differences of phases have a physical meaning as absolute phases can be removed by rotations.

Direct *CP* Violation

Direct CP violation or *CP violation in the decay* means that the decay amplitudes into the final state, f, and the *CP*-conjugate final state, $\bar{f}$, are different for the meson and the antimeson, respectively. It can occur in charged as well as in neutral decays. In charged decays, where no mixing can happen, it is the only possible type of *CP* violation. The amplitudes can be written using weak, $e^{i\delta_i}$, and strong phases, $e^{i\phi_i}$, like

$$A_f = \sum_i A_i e^{i\left(\delta_i + \phi_i\right)}, \tag{2.44}$$

$$\overline{A}_{\bar{f}} = e^{2i\left(\xi_f - \xi_B\right)} \sum_i A_i e^{i\left(\delta_i - \phi_i\right)}, \tag{2.45}$$

with $\xi_{\{f,B\}}$ from Eq. (2.15). If f is a *CP* eigenstate, then $e^{2i\xi_f} = \pm 1$ is its *CP* eigenvalue. If *CP* is conserved, the weak phases, ϕ_i, are all equal. Direct *CP* violation is results in

$$\left|\frac{A_f}{\overline{A}_{\bar{f}}}\right| = \left|\frac{\sum_i A_i e^{i(\delta_i+\phi_i)}}{\sum_i A_i e^{i(\delta_i-\phi_i)}}\right| \neq 1. \tag{2.46}$$

Indirect *CP* Violation

Indirect CP violation or *CP violation in mixing* occurs only in neutral mesons and results from the difference between the mass and the *CP* eigenstates, thus that from Eq. (2.25) follows

$$\left|\frac{q}{p}\right| \neq 1. \tag{2.47}$$

It can be studied in flavour-specific decays, where the final state can be reached from either a meson or an antimeson, but not from both, *e.g.* in semi-leptonic decays, where the meson only decays into l^+X, while the antimeson decays only into l^-X. Thus an initial meson or anitimeson can only reach the l^-X or the l^+X final state only via mixing, respectively.

CP Violation in the Interference of Decay and Decay after Mixing

The third class of *CP* violation occurs if the neutral meson and the antimeson decay into a common final state. It arises from the interference between the direct decay without mixing, $B^0 \to f$, and the decay including mixing, $B^0 \leftrightarrow \overline{B}^0 \to f$. It is also evident even if no direct and indirect *CP* violation is present. The definition from Eq. (2.40) changes if the final state is a *CP* eigenstate, f_{CP}, to

$$\lambda_{f_{CP}} = \eta_{CP} \cdot \frac{q}{p}\frac{\overline{A}_f}{A_f}, \tag{2.48}$$

with the *CP* eigenvalue, η_{CP}, of the final state with

$$CP\left|f_{CP}\right\rangle = \left|\bar{f}_{CP}\right\rangle = \eta_{CP}\left|f_{CP}\right\rangle = \pm\left|f_{CP}\right\rangle. \tag{2.49}$$

The decay amplitude can be parametrized as $A_f = Ae^{i(\delta_D+\phi_D)}$ (see Eq. (2.44)). Moreover, *CP* violation in the mixing can be assumed to be negligible (as $\left|\Gamma_{12}/M_{12}\right| \approx 0$, see Eq. (2.36)), thus the mixing parameter, $q/p = e^{-i\phi_M}$, can be expressed by one phase. Due to that $|\lambda_f| = 1$, but the phase difference between the direct decay and the decay after mixing allows for

$$\mathrm{Im}(\lambda_f) \neq 0. \tag{2.50}$$

The decay-time-dependent *CP* asymmetry in the case of the decay to a *CP* eigenstate is given by

$$\begin{aligned}\mathcal{A}_{CP}(t) &\equiv \frac{\Gamma(\overline{B}^0(t) \to f_{CP}) - \Gamma(B^0(t) \to f_{CP})}{\Gamma(\overline{B}^0(t) \to f_{CP}) + \Gamma(B^0(t) \to f_{CP})} \\ &= \frac{S \cdot \sin(\Delta m t) - C \cdot \cos(\Delta m t)}{\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D \cdot \sinh\left(\frac{\Delta\Gamma}{2}t\right)},\end{aligned} \tag{2.51}$$

when also assuming $|q/p| = 1$ and where B^0 and $\overline{B}^0$ denote the production flavor of the corresponding meson at $t = 0$. The coefficients S, C and D are the previously defined *CP* observables from Eq. (2.43).

2.2.4 *CP* Violation in $B^0 \to [c\bar{c}]K^0_S$ Decays

In general Cabibbo-favoured $b \to [c\bar{c}]s$ transitions are called *golden modes* to measure decay-time-dependent *CP* violation in the interference of decays with and without mixing. In this thesis, especially the group of $B^0 \to [c\bar{c}]K^0_S$ decays are considered, where $[c\bar{c}]$ denotes a charmonium resonance, like the J/ψ or the $\psi(2S)$ meson. In such decays higher-order contributions that could introduce additional strong and weak phases in the decay amplitudes are expected to be small [36–38] and *CP* violation in the mixing is negligible at the current experimental precision [39].

In the decays considered in this thesis, $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi K^0_S$, the B^0 candidates are reconstructed by combining the charmonium and the K^0_S candidates that form a common vertex. The charmonium candidates are reconstructed from two muons in the case of the $\psi(2S)$ meson and from two electrons in the case of the J/ψ meson. In both cases the K^0_S candidates are formed from two oppositely charged pions. The values in the following section are taken from Ref. [27].

The branching ratios for the decays of $B^0 \to \psi(2S)K^0$ and of $B^0 \to J/\psi K^0$ are given by

$$\begin{aligned}\mathcal{B}(B^0 \to \psi(2S)K^0) &= (8.73 \pm 0.32) \cdot 10^{-4}, \\ \mathcal{B}(B^0 \to J/\psi K^0) &= (5.8 \pm 0.5) \cdot 10^{-4},\end{aligned} \tag{2.52}$$

and are of the same order of magnitude. The K^0_S meson consists of a down quark and an antistrange quark, has a mass of $(497.611 \pm 0.013)\,\mathrm{MeV}/c^2$, and a considerably long lifetime of about (89.54 ± 0.04) ps, which means that it has a non-vanishing flight distance in the detector in the order of 1 m. The $B^0 \to [c\bar{c}]K^0_S$ decay proceeds via the decay of the $K^0/\overline{K}^0$ flavour eigenstates to the K^0_S mass eigenstate with a branching ratio of about 50 %. The K^0_S further decays with around 70 % to two oppositely charged pions. The J/ψ and $\psi(2S)$ meson consist both of a charm quark and an anticharm quark, and have a mass of $(3096.900 \pm 0.006)\,\mathrm{MeV}/c^2$ and $(3686.097 \pm 0.025)\,\mathrm{MeV}/c^2$, respectively. They both

decay almost instantaneously and leave no visible track in the detector. While the J/ψ meson decays with around 88 % via the strong and electromagnetic interaction to hadrons and with around 6 % each to two electrons or two muons [27], the $\psi(2S)$ meson decays about 98 % to hadrons and with around 0.8 % each to two electrons or to two muons. Taking everything into account the overall branching ratios for the considered decays into the given final states are

$$\begin{aligned}\mathscr{B}(B^0 \to \psi(2S)(\to \mu^+\mu^-)K_S^0(\to \pi^+\pi^-)) &= (1.61 \pm 0.20) \cdot 10^{-6},\\ \mathscr{B}(B^0 \to J/\psi(\to e^+e^-)K_S^0(\to \pi^+\pi^-)) &= (18.4 \pm 0.7) \cdot 10^{-6}.\end{aligned} \tag{2.53}$$

The important decay amplitudes for such decays are

$$\begin{aligned}A([c\bar{c}]K_S^0) &= \langle [c\bar{c}]K_S^0 \big| H \big| B^0\rangle,\\ \overline{A}([c\bar{c}]K_S^0) &= \langle [c\bar{c}]K_S^0 \big| H \big| \bar{B}^0\rangle.\end{aligned} \tag{2.54}$$

As mentioned before the K_S^0 is not a flavour eigenstate, but a mass eigenstate of the neutral kaon system. The two mass eigenstates, K_S^0 and K_L^0, can be written as a linear combination of the flavour eigenstates, K^0 and $\bar{K}^0$, as

$$\begin{aligned}\big|K_S^0\rangle &= p_K \big|K^0\rangle - q_K \big|\bar{K}^0\rangle,\\ \big|K_L^0\rangle &= p_K \big|K^0\rangle + q_K \big|\bar{K}^0\rangle,\end{aligned} \tag{2.55}$$

analogously to the neutral b-meson system (see Eq. (2.22)). The K_L^0 is much heavier than the K_S^0 and has a longer lifetime. For the ratio of the mixing parameters follows

$$\frac{q_K}{p_K} = -\frac{V_{cs}^* V_{cd}}{V_{cs} V_{cd}^*}, \tag{2.56}$$

when neglecting CP violation in the mixing. As a B^0 and a $\bar{B}^0$ meson can only decay into the flavour eigenstates K^0 and $\bar{K}^0$, respectively, the decay amplitudes change to

$$\begin{aligned}A([c\bar{c}]K_S^0) &= \langle [c\bar{c}]K_S^0 \big| H \big| B^0\rangle = +\frac{1}{2p_K}\langle [c\bar{c}]K^0 \big| H \big| B^0\rangle,\\ \overline{A}([c\bar{c}]K_S^0) &= \langle [c\bar{c}]K_S^0 \big| H \big| \bar{B}^0\rangle = -\frac{1}{2q_K}\langle [c\bar{c}]\bar{K}^0 \big| H \big| \bar{B}^0\rangle.\end{aligned} \tag{2.57}$$

The dominantly contributing Feynman diagrams are shown in Fig. 2.4. The tree as well as the penguin diagram with an internal up-type quark loop contribute to the decay amplitudes. Following Refs. [27, 40] the amplitude can be expressed using the CKM-matrix elements as

$$A(B^0 \to [c\bar{c}]K_S^0) = V_{cs}V_{cb}^* T_c \; + V_{us}V_{ub}^* P_u \; + V_{cs}V_{cb}^* P_c \; + V_{ts}V_{tb}^* P_t \;, \tag{2.58}$$

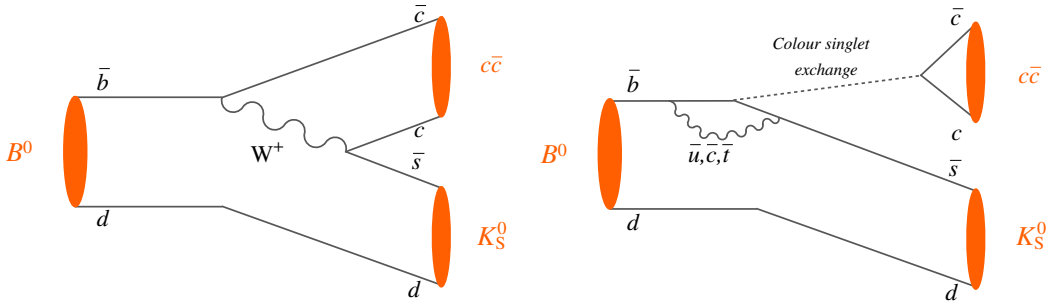

Figure 2.4: Relevant Feynman diagrams of the $B^0 \to \psi(2S)K^0_S$ decay at the lowest order, here more generalized for the decay including a charmonium state. The (left) dominant, color-supressed tree diagram and the (right) loop-supressed penguin diagram.

where T_c represents the strong contribution form the tree diagram, and P_q with $q \in \{u, c, t\}$ represents strong penguin contributions, from the inside up-type quark-loop. All these contributions are conserved under *CP* transformation. Considering the CKM unitarity and using the Wolfenstein parametrization up to terms of $\mathcal{O}(\lambda^2)$ the amplitudes can be approximated as

$$\begin{aligned} A(B^0 \to [c\bar{c}]K^0) &= \left(1 - \frac{\lambda^2}{2}\right) \mathcal{A}\left(1 + \epsilon a e^{i\theta} e^{+i\gamma}\right), \\ A(B^0 \to [c\bar{c}]\bar{K}^0) &= \left(1 - \frac{\lambda^2}{2}\right) \mathcal{A}\left(1 + \epsilon a e^{i\theta} e^{-i\gamma}\right), \end{aligned} \tag{2.59}$$

with

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

with the side, $R_u = |V_{ud}V^*_{ub}/V_{cd}V^*_{cb}|$ of the unitarity triangle. Furthermore, γ is the only weak phase and while it changes its sign under *CP* transformation as the difference between $A(B^0 \to [c\bar{c}]K^0)$ and $A(B^0 \to [c\bar{c}]\bar{K}^0)$ shows, the sign in front of the strong phase, θ, remains unchanged.

Taking everything into account, the B^0-$\bar{B}^0$ mixing, the terms from the kaon system, and the decay amplitude, the mixing parameter λ_f from Eq. (2.40) can be expressed by

$$\begin{aligned} \lambda_{[c\bar{c}]K^0_S} = \frac{q}{p}\frac{\overline{A}([c\bar{c}]K^0_S)}{A([c\bar{c}]K^0_S)} &= -\left(\frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}\right)\left(\frac{V^*_{cs}V_{cb}}{V_{cs}V^*_{cb}}\right)\left(\frac{V^*_{cd}V_{cs}}{V_{cd}V^*_{cs}}\right) \cdot \frac{1 + \epsilon a e^{i\theta} e^{+i\gamma}}{1 + \epsilon a e^{i\theta} e^{-i\gamma}} \\ &= -e^{-2i\beta} \cdot \frac{1 + \epsilon a e^{i\theta} e^{+i\gamma}}{1 + \epsilon a e^{i\theta} e^{-i\gamma}}, \end{aligned} \tag{2.60}$$

where the first term comes from the B^0-$\bar{B}^0$ mixing, the second term arises from the ratio $\overline{A}_f/A_f$, the third from K^0-$\bar{K}^0$ mixing, and the last one represents the penguin contributions.

Here, β and γ are angles of the CKM triangle as defined in Eq. (2.13). The relative negative sign between the amplitudes arises from the fact, that when the scalar B^0 meson decays to a charmonium resonance of spin 1, when considering the J/ψ or $\psi(2S)$ meson, and to a K_S^0 of spin 0, a relative angular momentum between the charmonium and the K_S^0 is needed (a relative $l = 1$ state [18]). And as the K_S^0 is reconstructed from two pions, which are both *CP*-even, the whole kaon system ends up with a negative *CP* eigenvalue and is *CP*-odd.

Neglecting penguin contributions, for $\epsilon = 0$, assuming that higher-order contributions are small, the parameter $\lambda_{[c\bar{c}]K_S^0}$ in Eq. (2.60) is only proportional to phase of 2β and thus the *CP* coefficients simplify to

$$S([c\bar{c}]K_S^0) = \sin(2\beta), \quad C([c\bar{c}]K_S^0) = 0, \quad \text{and} \quad D([c\bar{c}]K_S^0) = \cos(2\beta). \tag{2.61}$$

When considering the vanishing decay-width difference of the mass eigenstates, $\Delta\Gamma/\Gamma = -0.002 \pm 0.010$ [39], the decay-time-dependent *CP* asymmetry in the case of the considered decays in Eq. (2.51) simplifies to

$$\mathcal{A}_{CP}(t) = S \cdot \sin(\Delta m t) - C \cdot \cos(\Delta m t), \tag{2.62}$$

As *CP* violation in the mixing is negligible and *CP* violation in the decay is expected to be small, the *CP* coefficient S can be identified as $\sin(2\beta)$.

As these types of decays are considered the most suitable for measuring *CP* violation in the interference, $B^0 \to [c\bar{c}]K_S^0$ decays have been studied extensively by LHCb, BaBar and Belle before. While LHCb operates at the LHC in a hadronic environment, the two *B* factories BaBar and Belle operated at electron-positron colliders. They run at the centre-of-mass energies of the $\Upsilon(4S)$ bottonium resonance, producing very high rates of $B^0\overline{B}^0$ pairs. Belle was operating at KEKB in Tsukuba in Japan and BaBar at PEP-II in Stanford in the US. The world averages [39] for S and C obtained in various $b \to [c\bar{c}]s$ transitions are

$$S(b \to [c\bar{c}]s) = 0.699 \pm 0.017, \quad C(b \to [c\bar{c}]s) = -0.005 \pm 0.015.$$

When considering only the specific decay channels of $B^0 \to \psi(2S)K_S^0$ and $B^0 \to J/\psi K_S^0$, the averages [39] are

$$\begin{aligned} S(B^0 \to \psi(2S)K_S^0) &= 0.817 \pm 0.056, & C(B^0 \to \psi(2S)K_S^0) &= -0.019 \pm 0.048, \\ S(B^0 \to J/\psi K_S^0) &= 0.695 \pm 0.019, & C(B^0 \to J/\psi K_S^0) &= \quad 0.000 \pm 0.020. \end{aligned}$$

When exclusively looking at single experiments, the Belle and BaBar (see Refs. [41, 42]) collaborations have measured $\sin(2\beta)$ in the $B^0 \to J/\psi K_S^0$ mode to a very high precision and they measured $\sin(2\beta)$ exclusively in $B^0 \to \psi(2S)K_S^0$ decays as well. The results are

$$\begin{aligned} S(B^0 \to J/\psi K_S^0, \text{Belle}) &= 0.670 \pm 0.032, \\ S(B^0 \to J/\psi K_S^0, \text{BaBar}) &= 0.657 \pm 0.038, \\ S(B^0 \to \psi(2S)K_S^0, \text{Belle}) &= 0.738 \pm 0.079 \pm 0.036, \\ S(B^0 \to \psi(2S)K_S^0, \text{BaBar}) &= 0.897 \pm 0.100 \pm 0.036. \end{aligned}$$

A measurement of $\sin(2\beta)$ using $B^0 \to J/\psi K^0_S$ decays, where the J/ψ meson was reconstructed from two muons, by the LHCb collaboration obtained a value of $S = 0.73 \pm 0.04$ [8] using $3\,\text{fb}^{-1}$.

All of the quoted measurements are well in agreement and show no real tensions among each other and also between the SM prediction of $\sin(2\beta)$ using CKM unitarity and measurements of other CKM-related quantities, which leads to $\sin(2\beta) = 0.740^{+0.020}_{-0.025}$ by the CKMfitter group [32] and to $\sin(2\beta) = 0.724 \pm 0.028$ by the UTfit collaboration [43].

3 The LHC and the LHCb Experiment

The Large Hadron Collider is a circular proton-proton, *pp*, collider at the European Organistaion for Nuclear Research, CERN, in the vicinity of Geneva. The accelerator is installed in a tunnel with a circumference of 26.7 km which is located 45 to 175 m below ground near the Swiss-French border. The tunnel was constructed between 1984 and 1989 for the Large Electron-Positron Collider, LEP. Two transfer tunnels, each around 2.5 km long, connect the LHC to the CERN accelerator complex.

In the LHC, two proton beams are accelerated in opposite directions to a centre-of-mass energy of 7 and 8 TeV in the run I of the LHC in the years 2011 and 2012, respectively. The centre-of-mass energy has been increased since then for run II of the LHC to 13 TeV in the years 2015 until 2018. The LHC is supplied by protons from an injection chain consisting of the Linear Accelerator 2, LINAC 2, the Proton Synchrotron Booster, BOOSTER, the Proton Synchrotron, PS, and the Super Proton Synchrotron, SPS. The protons are produced by ionising hydrogen atoms, because the remaining nuclei only consist of protons. These protons are then accelerated by the LINAC 2 to an energy of 50 MeV. After that, the BOOSTER increases the energy to 1.4 GeV and sends the protons to the PS, there they are accelerated to 25 GeV. After that they enter the SPS which provides the LHC at the nominal running conditions proton beams at energies of 450 GeV. In the near future LINAC 2 will be superseded by LINAC 4, which will be 86 m long and will be able to take the hydrogen ions to an energy of 160 MeV. In Fig. 3.1 a schematic depiction of the accelerator complex and a selection of experiments is shown.

The collider has an instant peak luminosity of $\mathscr{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is the ratio of the number of detected events, N, at a given time, t, and the corresponding cross section, σ, given as

$$L = \frac{1}{\sigma}\frac{\mathrm{d}N}{\mathrm{d}t}. \tag{3.1}$$

An integration over time leads to the integrated luminosity

$$L_{\mathrm{int}} = \int L\mathrm{d}t. \tag{3.2}$$

At this design peak luminosity 2808 bunches of protons with a bunch spacing of 25 ns are stored in each proton beam. Each bunch contains up to $1.15 \cdot 10^{11}$ protons, which leads to a total beam energy at 7 TeV of 362 MJ per beam [44].

The two proton beams are collided at four main points along the accelerator, where the four big experiments ATLAS (A Toroidal LHC Apparatus) [45], CMS (Compact Muon Solenoid) [46], ALICE (A Large Ion Collider Experiment) [47], and LHCb (Large Hadron

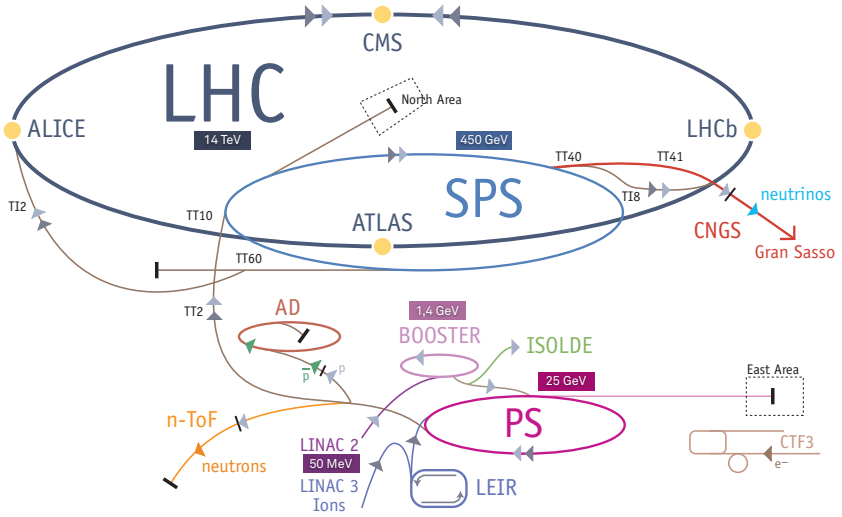

Figure 3.1: The LHC with a selection of experiments and its pre-accelerators is shown. The protons are produced via ionisation of hydrogen atoms and are then pre–accelerated by the (pruple) Linear Accelerator 2, LINAC 2, and the (rose) Proton Synchrotron Booster, BOOSTER, to 1.4 GeV before they are filled into the (pink) Proton Synchrotron, PS, and there they are accelerated to 25 GeV. After that they enter the (light blue) Super Proton Synchrotron, SPS, and finally they are accelerated in the (dark blue) LHC.

Collider beauty) [48] are located. Besides these experiments there exists a variety of smaller experiments at CERN, *e.g.* fixed target experiments in the north area. ATLAS and CMS are the two high luminosity experiments both aiming at the designed peak luminosity of the LHC, whereas LHCb aims at $\mathscr{L} = 2 \cdot 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. For heavy ion operations, especially lead ions, ALICE is aiming at a peak luminosity of $\mathscr{L} = 10^{27}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The two general-purpose detectors, ATLAS and CMS, are build as so-called 4π *detectors*, which surround the interaction points symmetrically. Both experiments are especially designed for direct searches for New Physics effects. Besides this, they discovered the Higgs boson in 2012 [25], which was proposed in 1964 and completes the SM. The ALICE experiment focuses on the physics of strongly interacting matter at high energy densities. At such extreme densities a phase of matter called *quark-gluon plasma* forms. For a more detailed description see Ref. [49].

3.1 The LHCb Experiment and Detector

The Large Hadron Collider beauty, LHCb, experiment is designed to study New Physics effects through indirect searches, *i.e.* high precision measurements. The main focus of LHCb is the measurement of *CP* violation, and of rare and very rare decays in decays of b and c hadrons. In such modes small deviations from the expectations can point to significant deviations from the SM and can therefore provide a valuable test of the SM. To be able to conduct measurements with the desired precision it is necessary to have data with a high amount of b and c hadrons provided by the LHC. Furthermore, the detector needs to be designed to be able to reconstruct such decays including the accompanying decay products, especially kaons, pions, muons, and electrons. In addition, the knowledge of the decay time of each reconstructed b- and c-hadron candidate is important to measure *CP* violation in the interference of decay and decay after mixing. For a detailed description see [48].

The production of heavy quarks in pp collisions is dominated by the production of a quark pair via gluon-gluon fusion, *i.e.* $gg \to q\bar{q}$, here $q \in \{c, b\}$. The high beam energies, that can be reached at the LHC, can lead to a high momentum asymmetry between the two gluons, which again leads to a large momentum of the quark pair and this causes a strong boost along the beam axis. Thus, both quarks have a high probability to be produced along the beam axis. This can be seen in Fig. 3.2, where simulated data at a centre-of-mass energy of 7 TeV shows the production angles of $b\bar{b}$ pairs. The red region includes around 25 % of all produced $b\bar{b}$ pairs and represents the geometrical acceptance of the LHCb detector. The production cross-section of $b\bar{b}$ pairs in the covered pseudorapidity, η, range from 2.0 to 5.0 was measured by LHCb to $\sigma(pp \to b\bar{b}X) = 72.0 \pm 0.3 \pm 6.8\,\mu\mathrm{b}$ [51] at a centre-of-mass energy of 7 TeV. The pseudorapidity is defined using the polar angle, θ, as $\eta = -\ln(\tan(\theta/2))$. The extrapolated $b\bar{b}$ cross-section over the full η range was measured to be $\sigma(pp \to b\bar{b}X) = 284 \pm 20 \pm 49\,\mu\mathrm{b}$ [52] at 7 TeV.

Of a produced $b\bar{b}$ pair each quark hadronises separately and incoherently from the other

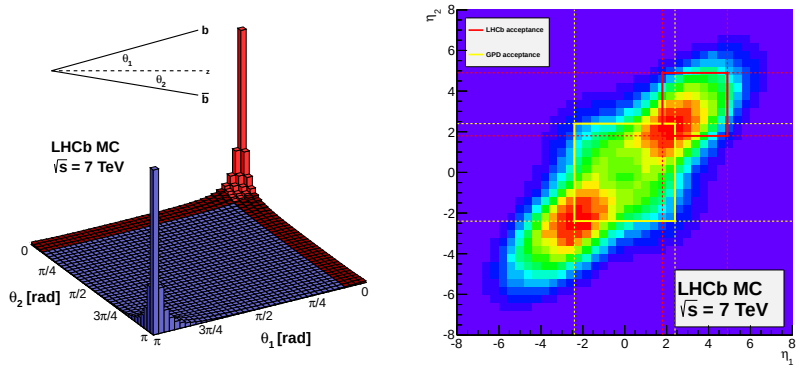

Figure 3.2: The production number of $b\bar{b}$ pairs as a function of (left) the azimuthal angles, θ_1 and θ_2, and of (right) the corresponding pseudorapidity from simulation at a centre-of-mass energy of 7 TeV. The (left) red shaded region and the (right) red box represent the geometrical acceptance of the LHCb detector, which corresponds to around 25 % of all produced $b\bar{b}$ pairs. The (right) yellow box shows the acceptance of the general purpose detectors, ATLAS and CMS. [50]

in hadron collisions. Nonetheless, it is possible to obtain a probability of the charge of the produced quark, the so-called *flavour tag*, d, from the accompanying particles produced in the hadronisation or from the fragmentation process of the other quark of the pair (see Sec. 4.3 for more details). The b quarks hadronise predominantly into lighter charged mesons, *i.e.* B^+ and B^-, with around 40 %, into neutrally charged mesons, *i.e.* B^0 and $\bar{B}^0$ also with around 40 %, and with around 10 % into B_s^0 and $\bar{B}_s^0$, and with the last 10 % into heavier mesons and baryons [27].

Due to the production of b and c quarks predominantly along the beam axis, the LHCb detector is built as a single-arm forward spectrometer covering an angular range from about 10 mrad to 300 mrad in the horizontal and to 250 mrad in the vertical plane. A schematic view of the LHCb detector is depicted in Fig. 3.3. Around the interaction point the Vertex Locator, VELO, is stationed. Further downstream the first Ring Imaging Cherenkov Detector, RICH1, followed by the Tracker Turicensis, TT, the dipole magnet, and the tracking stations, T1-T3, followed by the second Ring Imaging Cherenkov Detector, RICH2, are located. The three tracking stations consist each of an inner silicon tracker, IT, and an outer drift-tube detector, OT. After those, the first muon station, M1, follows, superseded by the calorimeter system consisting of a Scintillating Pad Detector, SPD, a Preshower detector, PS, an Electromagnetic Calorimeter, ECAL, and a Hadronic Calorimeter, HCAL. The detector is completed by additional muon stations, M2-M5. The detector components can be subdivided in two categories according to their purpose, first in sub-detectors designed for finding particle tracks and second in components designed for particle identification, PID. The VELO, the tracking stations, T1-T3 including the TT, the magnet, as well as the muon chambers belong

to the tracking system. Although, the latter can be assigned to both categories. The two RICH detectors, as well as the calorimeter system and also the muon chambers form the particle identification system. As the detector can be positioned very close to the beam (more

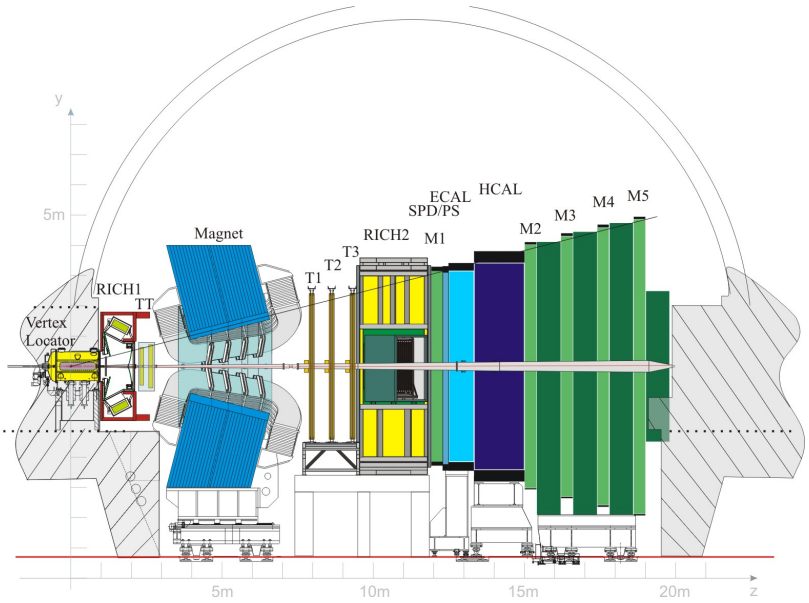

Figure 3.3: Schematic view of the LHCb detector. The beam pipe is located horizontally along the z-axis. Around the interaction point the Vertex Locator, VELO, is stationed. Further downstream the first Ring Imaging Cherenkov Detector, RICH1, followed by the Tracker Turicensis, TT, the dipole magnet, and the tracking stations, T1-T3, followed by the second Ring Imaging Cherenkov Detector, RICH2, are located. The three tracking stations consist each of an inner silicon tracker, IT, and an outer drift-tube detector, OT. After those, the first muon station, M1, superseded by the calorimeter system consisting of a Scintillating Pad Detector, SPD, a Preshower detector, PS, an Electromagnetic Calorimeter, ECAL, and a Hadronic Calorimeter, HCAL. The detector is completed by additional muon stations, M2-M5. [48]

details in Sec. 3.1.1), equipped with several sensitive and fragile sub-detectors, and due to the unique position of LHCb in direction of the injection line of the counter clockwise proton beam, there exist a variety of safety systems for monitoring the LHC's beam conditions. One of those is the Beam Conditions Monitor, BCM [53, 54]. The BCM continuously measures the particle flux at two points symmetrical to the interaction point in the detector and consists of eight diamond sensors at each of the two points. If it measures a current above a certain threshold, the BCM is able to initiate a controlled beam dump within two turns of the beam.

3.1.1 Tracking System

Due to its close position to the interaction point in LHCb, the VELO provides precise measurements of track coordinates, especially of the position of the primary (PVs) and the distinct secondary vertices (SVs). It consists of silicon modules arranged along the beam axis. They measure the cylindrical r and ϕ coordinates. The VELO is schematically shown in Fig. 3.4. The modules can be positioned radially with a distance of 8 mm to the beam. Due to this very small distance, the BCM is necessary to protect these fragile silicon detectors. During injection the sensors can be retracted, due to their composition of two halves. The sensors are built in their own vacuum vessel, which is separated from the beam vacuum by an RF-foil. The VELO covers the whole LHCb acceptance from $1.6 < \eta < 4.9$.

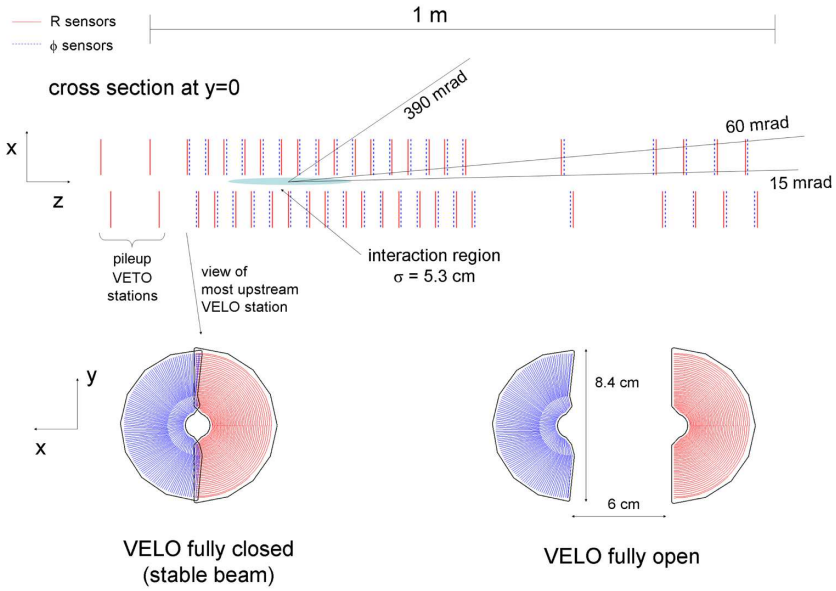

Figure 3.4: Schematic view of the VELO silicon sensors in the (x, z) plane. And the closed and open position of the sensors from the front are shown. [48]

The TT upstream of the magnet and the inner part of the tracking stations downstream of the magnet, IT, form the Silicon Tracker, ST. Both trackers are made of silicon microstrip sensors. The TT as a planar tracking station covers the whole LHCb acceptance region and is 150 cm wide and 130 cm high, while the IT covers a 120 cm wide and 40 cm high cross - shaped region in the inner part of the three tracking stations. Each of this four ST stations has a (x-u-v-x) arrangement of the four detection layers within, which means in the first and last layer the strips are arranged vertically and in the two inner layers the strips are rotated by $-5°$ and $+5°$ around the z-axis, respectively. The TT station consists of two halves, one half on each side of the beam pipe. Each of the three IT stations comprises four

single detector boxes, that are arranged symmetrically around the beam pipe. Each of the boxes contains the four detection layers, in the (*x*-*u*-*v*-*x*) design, which leads to a spatial resolution of about 50 µm.

The OT is a gaseous straw-tube detector [55]. It can detect charged particles and measure their momenta very precisely with a resolution of about $\delta p/p \approx 0.4\,\%$, where p is the particle's momentum. It is also arranged in three stations with four layers each in the same pattern, in the (*x*-*u*-*v*-*x*) design which leads to a spatial resolution of about 200 µm in the OT. The OT is an array of airtight tubes, which are filled with a mixture of argon (70 %), carbon dioxide (28 %) and oxygen (2 %). This assures drift times lower than 50 ns. The tubes build the cathode surface and a wire in the middle acts as the anode. If a charged particle traverses the tube, it ionizes the gas atoms and the resulting electrons and ions are accelerated by the electric field that prevails between the tubes and the wire. This causes a cascade of ionization also called *Townsend avalanche*, which results in a charge proportional to the ionization effect when arriving at the wire (see Fig. 3.5).

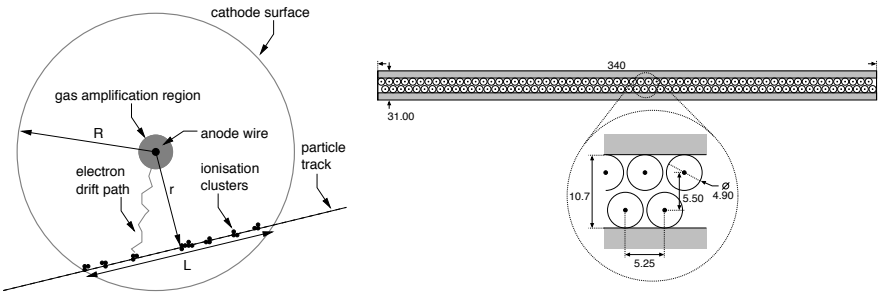

Figure 3.5: (Left) Schematic picture of a particle traversing a tube of the OT and (right) the cross section of a straw-tubes OT module. [48, 55]

The dipole magnet can provide a magnetic field of 4 T m. It is divided in two identical trapezodial coils bent at 45° on the two transverse sides, which are placed symmetrically above and below the beam pipe in an iron yoke. The coils consist each of fifteen pancakes arranged in five triplets and are made of pure aluminium, Al-99.7, hollow conductors in an annealed state, with a 25 mm diameter bore for water cooling. The polarization of the magnetic field alternates throughout a run of LHCb, to prevent a systematic influence of the magnetic field, thus the data can be classified in magnet up, MU, and magnet down, MD, which should not differ significantly or otherwise would point to a problem during data taking.

Taking the information from the VELO, the TT, and the tracking stations and including the curvature of the particles trajectory induced by the dipole magnet into account, the trajectory can be reconstructed, and the momentum of the particle can be determined.

Depending on where in the tracking sub-detectors a particle is detected, the tracks are classified differently (see Fig. 3.6). Tracks, that are detected in the VELO and in all other

tracking sub-detector are called *long tracks*, L, and tracks, which do not show a signature in the VELO but in all of the other tracking stations are called *downstream tracks*, D. Other distinctions are made between tracks, that are only detected in the VELO, *VELO tracks*, or only in the tracking stations downstream of the magnet, *T tracks*, and tracks that are visible in the VELO and in the TT, *upstream tracks*. Interesting for this analysis, especially for the analysis of the $B^0 \to \psi(2S)K^0_S$ channel, are the long and downstream tracks, due to the finite lifetime and thus, finite flight distance of the K^0_S mesons. They can decay early enough for their daughter pions to be reconstructed from hits in the VELO and in the following tracking detectors and are characterised as long track, LL candidates, as both pions are reconstructed as long track candidates. Due to the long lifetime of the K^0_S meson, it is also possible that the K^0_S decays outside of the VELO and therefore the tracks are reconstructed from hits from all tracking detectors but the VELO. Those candidates are characterised as *downstream* candidates, DD, because both pions are reconstructed as downstream candidates. Both track types differ in momentum and vertex resolution of the tracks.

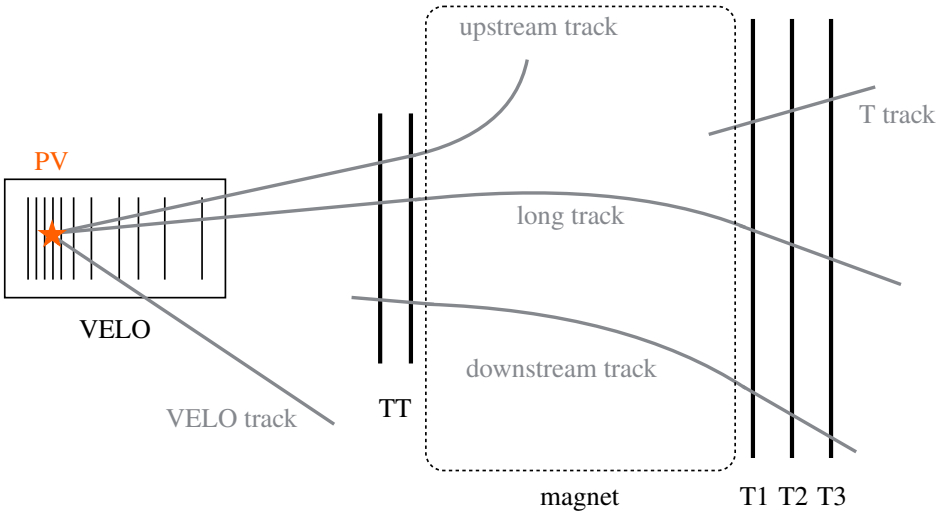

Figure 3.6: Schematic view of the different track types at LHCb.

3.1.2 Particle Identification System

The particle identification system is important for LHCb to be able to distinguish between kaons, pions, protons and electrons. It comprises the two RICH detectors, the calorimeter system, and the muon system. The upstream detector, RICH1, covers the low momentum range from 1 GeV/c to 60 GeV/c, while the downstream detector, RICH2, covers the high momentum range from 15 GeV/c to 100 GeV/c. These detectors exploit Cherenkov radiation of charged particles, which traverse faster through a medium than light in the same medium.

The velocity of the particles can be determined from the opening angle of the Cherenkov light cone, via $\cos(\theta) = c'/v = 1/(n\beta)$, with the speed of light in the medium, $c' = c/n$, and the refractive index of the medium, n. The radiators used are aerogel and fluorobutane, C_4F_{10}, for RICH1 and CF_4 for RICH2. Both detectors use a combination of spherical and flat mirrors to reflect the image of the Cherenkov light out of the acceptance. Whereas the optical layout of RICH1 is vertical, the one of RICH2 is horizontal. To detect Cherenkov photons Hybrid Photon Detectors, HPDs, are used. These HPDs are surrounded by iron shields to prevent the magnetic field of the dipole magnet to interfere with the measurement.

The task of the calorimeter system is to determine the PID of electrons, photons, and hadrons as well as the energies and positions of the particles. The ECAL is responsible for identifying electrons and photons and the HCAL is designed to identify hadrons, such as pions and kaons. All calorimeter sub-detectors use scintillating material alternating with absorption material, *i.e.* lead or iron, to measure the energy deposition. To suppress background from charged pions and from neutral pions with high E_T the PS is stationed in front of the ECAL and the SPD is stationed in front of the PS, respectively. Both sub-systems are separated by a thin lead layer. The SPD is sensitive to charged particles and allows for a distinction between electrons and photons, as electrons shower in the SPD and the ECAL, whereas photons only shower in the ECAL. The ECAL uses scintillator plates alternating with lead plates, while the HCAL uses iron as absorber material. Thus, a distinction between electrons and charged hadrons is possible, as the probability that the electrons shower in lead plates is high and the electrons will then deposit their energy in the ECAL, while the charged hadrons are able to reach the HCAL.

The muon system provides information for muons with a high transverse momentum w.r.t. the beam pipe, p_T, and muon identification for the trigger system. It is composed of five stations, one upstream and four downstream of the calorimeter system. Between each of the stations M2-M5 are iron absorbers, which are 80 cm in thickness. The transverse dimensions of all five stations scale with the distance from the interaction point. The muon system is built of 1380 multi-wire proportional chambers, MWPCs, except for the inner region of M1, where triple Gas Electron Multiplier, GEM, detectors are used. The MWPCs allow for a $> 95\,\%$ detection efficiency in a 20 ns window. Each station has an increasing readout granularity towards the beam axis.

3.2 Recording Data at LHCb

Due to the very high event rate and the large amount of data associated with it, a trigger system is necessary to be able to record interesting physics data. The trigger system comprises requirements on particle properties, but is also able to do a full event reconstruction.

3.2.1 Trigger System

The trigger system comprises two stages: the level 0 trigger, L0, which is implemented in hardware, and a high level trigger, HLT, implemented in software. It reduces the 40 MHz bunch-crossing rate to 5 kHz that can be written to storage. An overview of the LHCb-trigger layout in the LHC's run I can be seen in Fig. 3.7.

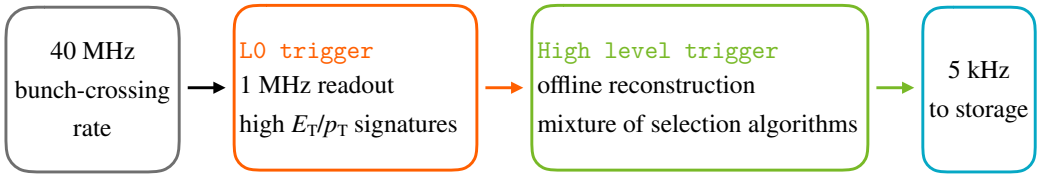

Figure 3.7: Layout of the LHCb-trigger system in run I. The figure is based on Ref. [56].

The bunch-crossing frequency of 40 MHz is first reduced by L0 to 1 MHz. The L0 works synchronous to the bunch-crossing rate. It relies only on information from the calorimeter and muon system due to the fast availability of this information. The L0 can be divided into three independent triggers, depending on detector components from which the decision is obtained: L0 calorimeter, L0 muon, and L0 pile-up trigger. The L0 calorimeter comprises specialized triggers: `L0Hadron`, `L0Electron`, and `L0Photon`, and it depends on information of the calorimeter system including the PS and SPD. The L0 muon consists of the `L0Muon` and `L0DiMuon` triggers and depends on information from the muon system. It tries to identify candidates with the highest and second highest p_T and searches for hits, that build a straight line through the five muon stations and that points back to the interaction point. The L0 pile-up trigger is used for the determination of the luminosity and to veto high pile-up events.

The software trigger reduces the rate provided by the hardware stage of 1 MHz further to 5 kHz. It can be divided into two stages. The first stage, HLT1, uses a partial reconstruction of the event, whereas the second stage, HLT2, performs a complete event reconstruction. For this event reconstruction the HLT uses data from the VELO, the tracking stations, the RICH detectors, and the muon stations.

The three stages, L0, HLT1, and HLT2, comprise various so-called *trigger lines*, which consist of a sequence of reconstruction algorithms and selection requirements and are specialized for types of particles or combinations of particles.

Depending on whether the reconstructed signal event is sufficient to trigger a decision by the system an event is classified as *triggered on signal*, TOS. If the decision would have been triggered without the signal event present, the event is classified as *triggered independent of signal*, TIS. An event can also be classified as both or none. A more detailed description can be found in [48].

3.2.2 Data Processing and Software

The basis of the LHCb software framework is Gaudi [57]. It meets all the requirements of a physics analysis, such as simulation, trigger, reconstruction, and visualization. A schematic view of the data flow from the experiment to the data used for analyses as well as the processing of simulated data can be seen in Fig. 3.8.

For the simulation Gauss [58] operates as a controlling software for the various sub-programs or generators, which handle the *pp* collisions in the case of Pythia [59, 60], the decays of heavy particles in the case of EvtGen [61], where radiation processes are simulated by Photos [62], and the particle interactions in the case of Geant4 [63, 64]. The detector response is simulated by the LHCb digitization application, Boole [65], which also accounts for spillover events and LHC background, *i.e.* particle fluxes from the accelerator tunnel. After the simulation step, recorded and simulated data follow the same path. First the data passes Moore [66], the software of the HLT, which is described in more detail in Sec. 3.2.1. The triggered, raw data has to be transformed from the detector hits to actual tracks and clusters. This is done by the reconstruction software Brunel [67]. This software builds so-called *proto-particles*, which are containers, that include information from the tracking and PID system. They do not represent actual particles, yet. The output of Brunel are *Data Summary Tape* files, DSTs. These are Root [68] files that contain the full event information. The data is further filtered and analysed by DaVinci [69], which takes the *proto-particles* and reconstructs the actual particles. Particles are combined and very loose selections are applied in this step. This process is also called *stripping* and is organized like the trigger in lines according to the decay or particles of interest. DaVinci provides the algorithms and tools for the physics analyses, *e.g.* the tool that summarizes all information that is needed to perform a flavour-tagged analysis (see Sec. 4.3). For a more detailed description see Refs. [48, 70].

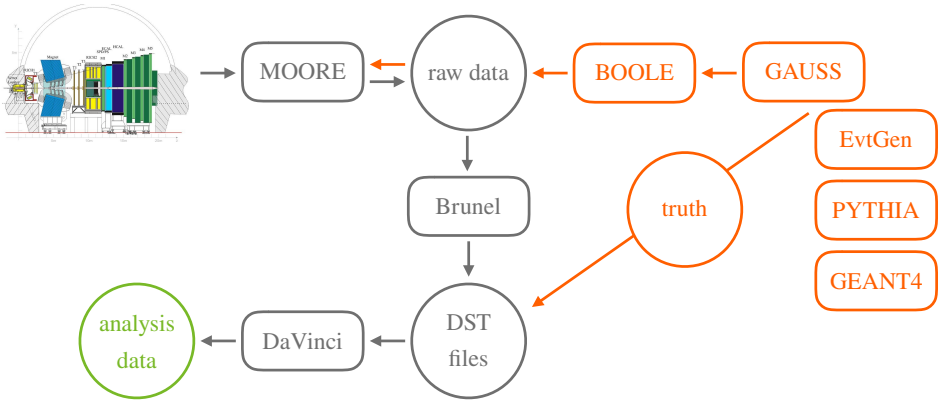

Figure 3.8: Data flow at LHCb of (grey) recorded, and (orange) simulated data and the used software. Figure based on Ref. [70].

4 Measurement of sin(2β)

The main aim of this measurement is the determination of the *CP* observables S and C through a fit to the reconstructed decay time of the $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi\, K^0_S$ signal candidates, extracted from data collected in the LHC's run I with the LHCb experiment. The parameter S can be related to the decay-time-dependent *CP* asymmetry (Eq. (2.62)). This asymmetry can be determined by an unbinned maximum-likelihood fit to the decay time of the B meson, including the necessary information of the production flavour and experimental properties. In this analysis the K^0_S meson is reconstructed from two charged pions in both decay channels, while the $\psi(2S)$ is reconstructed through its decay into two muons and the J/ψ is reconstructed through its decay into two electrons. This represents the first decay-time-dependent measurement of *CP* violation at a hadron collider in the decay channel $B^0 \to \psi(2S)K^0_S$ and that uses electrons in the final state.

First, the procedure of the data preparation (Sec. 4.1), the extraction of the signal candidates (Sec. 4.2), and the flavour-tagging studies (Sec. 4.3) in the case of the $B^0 \to \psi(2S)K^0_S$ mode will be presented. Further, a first estimation of the branching fraction of $B^0_s \to \psi(2S)K^0_S$ decays compared to branching fraction of $B^0 \to \psi(2S)K^0_S$ decays is presented in Sec. 4.2.1. Then a brief overview of the same preparations in case of the $B^0 \to J/\psi\, K^0_S$ decay channel is presented in Sec. 4.4. In Sec. 4.5 the model to extract the *CP* observables is presented including the consideration of experimental resolution effects. The measurement is completed by a brief summary of the estimation of the systematic uncertainties (Sec. 4.7), the consideration of kaon regeneration (Sec. 4.6), and finishing with the presentation of the results and the combination of the presented results with a previous measurement by LHCb [8] in Sec. 4.8.

4.1 Selection

Due to the huge amount of data it is necessary to develop a suitable selection to suppress as much background as possible while retaining a satisfying amount of signal candidates. The prevailing hadronic conditions lead to high track multiplicities, that in turn lead to events which are dominated by a flat combinatorial background in the invariant mass spectrum of the B^0 meson. This combinatorial background consists of combinations of particles, which are incorrectly associated to the reconstructed signal decay. Whereas exclusive physics backgrounds lead to secondary maxima in the invariant mass spectrum. Both types of backgrounds need to be suppressed and thus several selection steps are applied. If a complete reduction is not possible, the backgrounds need to be taken into account by the fit model.

The first step consists of a trigger system (see Secs. 3.2.1 and 4.1.2), then a centralised loose and very general selection, the *stripping*, is applied to the triggered data (see Secs. 3.2.2 and 4.1.3). The next step is the *offline selection* (see Sec. 4.1.4) of the remaining candidates to further extract the specific decays. This selection is optimised for the considered measurement. In this analysis the offline selection consists of a multivariate classifier, *i.e.* a Boosted Decision Tree, BDT. A suitable figure of merit for the optimisation of the selection is developed, which considers the features of the considered decays and the characteristics of the measurement approach (see Sec. 4.1.4).

After the selection, the fit model for obtaining the signal-candidate weights (see Sec. 4.1.4) through the sPlot technique [71], is introduced in Sec. 4.1.4. These weights are then used in the fit to the decay-time distribution to measure the *CP* parameters. Additionally, to suppress the $\Lambda_b^0 \to \psi(2S)\Lambda^0$ background a veto is applied after the selection. After the preceded selection steps there may still be more than only one primary vertex associated to one B^0 candidate. A special requirement is developed to just obtain one PV per B^0 candidate (see Sec. 4.1.6). An event can also contain multiple B^0 candidates, although due to the branching ratios of the analysed channels not more than one candidate is expected per event. Therefore one of the remaining candidates after all preceding selection steps have been applied is chosen randomly, as it is assumed that they are all equally likely to be real signal candidates (see Sec. 4.1.6).

4.1.1 Data Samples, Categories and Observables

The data samples used in this analysis were recorded in 2011 and 2012 and correspond to an overall integrated luminosity of $\mathscr{L}_{\text{int}} = 3\,\text{fb}^{-1}$. In 2011 the centre-of-mass energy of the pp collisions was $\sqrt{s} = 7\,\text{TeV}$, while in 2012 it was increased to $\sqrt{s} = 8\,\text{TeV}$. To consider the different running conditions in the two years of data taking, the simulated data is based on the data-taking conditions from the given periods.

The datasets used for evaluating the signal efficiencies of the different selection steps and for training, testing, and optimizing the BDT have to meet the following requirements (if not stated otherwise): a B^0 candidate needs to be associated with at least one PV, here the best PV, which is the PV with the highest impact parameter, IP, significance. In the case of simulated data it is required that the candidates match generated signal decays, this requirement will be called *truth matching* throughout the thesis. The efficiencies for each selection step quoted in the respective section are calculated with respect to all preceding selection steps if not stated otherwise.

In the $B^0 \to \psi(2S)K_S^0$ decays the K_S^0 is reconstructed through its decay in two oppositely charged pions and the $\psi(2S)$ through its decay into two muons. Due to very different momentum resolutions, the dataset in the case of the $\psi(2S)$ mode is divided according to the track type of the K_S^0 meson (see Sec. 3.1.1). Besides this distinction in track types no further split in other categories is necessary. In the case of the J/ψ channel, the necessity to split into Bremsstrahlung categories was also investigated and was not found to be necessary.

In this analysis two methods are used to reconstruct the decay chain of the B-meson candidate and to obtain the considered observables. The first is the `OfflineVertexFitter`, OVF. It reconstructs the decay chain using a step-by-step bottom-to-top approach, *i.e.* starting with combining the final state particles to intermediate particles and determining the decay vertices and kinematic properties until the whole decay chain up to the B-meson candidate is reconstructed. The second method is the `DecayTreeFitter` [72], DTF. It reconstructs the whole decay chain in one single fit and can thus obtain the correct correlations and uncertainties on the vertex positions, particle momenta, flight distances, decay times, and invariant masses. The invariant mass of the B-meson candidates is determined by the DTF where the masses of the daughter particles are constrained to their PDG [27] values if not stated otherwise. The decay time and the decay-time error-estimates are also obtained by the DTF in which the position of the primary vertex is constrained. The observables used in the first loose selection step are determined by the OVF.

4.1.2 Trigger Selection

Due to the geometry of the detector only around 18 % of all produced $B^0 \to \psi(2S)K^0_S$ decays lie within the acceptance of the detector as depicted in Fig. 3.2. The exact numbers can be found in Tab. 4.1.

Table 4.1: Geometrical signal efficiencies for the different years of data taking and the weighted arithmetic mean of these two numbers as the efficiency of run I.

2011	2012	run I
$(18.117 \pm 0.033)\,\%$	$(18.457 \pm 0.032)\,\%$	$(18.339 \pm 0.024)\,\%$

The decays, that lie within the detector acceptance, go through the online event selection, which is performed by the trigger system, for more detail on the trigger itself see Ref. Sec. 3.2.1. At the trigger stage not actual particles are used, as only proto-particles exist. The trigger stages use just the information of particular detector components which can be put into conjunction to distinctive particles, such as the information of the muon chambers or the calorimeter components.

In this analysis the two muons in the final state in the case of the $\psi(2S)$ mode can be exploited due to their clean signature in the detector, as they are expected to also reach beyond the muon chambers at the end of the detector at the considered energies and thus, have an expected flight distance of greater than 20 m. At the hardware stage a TOS decision of the `L0Global` trigger line is required. This requirement can be split into various single decisions based on the considered particle or combination of particles: `L0Muon`, `L0DiMuon`, `L0Hadron`, `L0Electron`, and `L0Photon`. The exact requirements of the hardware stage are listed in Tab. 4.2. The decisions of the `L0Muon` and `L0DiMuon` make up about 99.89 % of the candidates that pass the `L0Global` line in this analysis. The remaining candidates

stem from the L0Hadron, whereas the L0Electron and L0Photon do not have a significant contribution.

In the following software stage a TOS decision of one of the HLT1 lines, HLT1DiMuonHighMass or HLT1TrackMuon, and of the HLT2DiMuonDetachedHeavy line is required. The events are required to contain at least one muon with a transverse momentum $p_T(\mu) > 1.0\,\text{GeV}/c$ and an impact parameter $\text{IP}(\mu) > 100\,\mu\text{m}$ with respect to all PVs in the event, *i.e.* the muon should not originate from the PV. Additionally, this impact parameter needs to be of satisfying quality, that is why the χ^2 of the impact parameter of the muon tracks, $\chi^2_{\text{IP}}(\mu)$, is required to be greater than 16. Furthermore, the events need to contain two oppositely charged muons with a combined mass of $m(\mu^+\mu^-) > 2.95\,\text{GeV}$, which enables of the muons to form a $\psi(2S)$ meson. The requirement on the distance of closest approach of the dimuon system, $\text{DOCA}(\mu^+\mu^-) < 0.2\,\text{mm}$, ensures that the muons from a common vertex. This vertex needs to be of satisfying quality, which is guaranteed by requiring the χ^2 of the dimuon vertex fit, $\chi^2_{\text{vtx}}(\mu^+\mu^-)$, to be smaller than 25. All those requirements ensure that the selected events contain well reconstructed pseudo-particles, which can be identified as muons in the reconstruction. A detailed list of the requirements can be found in Tab. 4.3 and 4.4.

Table 4.2: Thresholds of the L0Global trigger in run I. The variables are the transverse momenta, p_T, of the muon, the hadron, the electron, and the photon and the product of the transverse momenta of the two muons.

	2011	2012
$p_T(\mu)$	$> 1.48\,\text{GeV}/c$	$> 1.76\,\text{GeV}/c$
$p_T(\mu_1) \cdot p_T(\mu_2)$	$> (1.296\,\text{GeV}/c)^2$	$> (1.6\,\text{GeV}/c)^2$
$p_T(\text{hadron})$	$> 3.5\,\text{GeV}/c$	$> 3.7\,\text{GeV}/c$
$p_T(e)$	$> 2.5\,\text{GeV}/c$	$> 3\,\text{GeV}/c$
$p_T(\gamma)$	$> 2.5\,\text{GeV}/c$	$> 3\,\text{GeV}/c$

Due to the production of simulated data the signal efficiency of the trigger requirements is computed by the number of candidates that pass the trigger stages compared to the number of reconstructed candidates that pass also the loose general selection (see Sec. 4.1.3) and is $\varepsilon_{\text{trig}} \approx 64.9\,\%$.

4.1.3 Preselection

A very loose and centralized preselection is applied when processing the triggered data, the so-called *Stripping*. The requirements on the reconstructed $\psi(2S)$ and its daughter particles are listed in Tab. 4.5 and the ones on the reconstructed K^0_S and its decay products can be found in Tab. 4.6. The cuts on the properties of the reconstructed B^0 meson are summarised in Tab. 4.7.

Table 4.3: Requirements of the used HLT1 lines that trigger on one single track and on the dimuon system. The variables are the transverse momentum, $p_T(\mu)$, the momentum, $p(\mu)$, the χ^2 over the number of degrees of freedom of the muon track fit, $\chi^2_{\text{track}}/\text{ndf}(\mu)$, the distance of closest approach in the dimuon system, $\text{DOCA}(\mu^+\mu^-)$, the χ^2 of the dimuon-vertex fit, $\chi^2_{\text{vtx}}(\mu)$, the invariant dimuon mass, $m(\mu^+\mu^-)$, the impact parameter, $\text{IP}(\mu)$ and the χ^2 of the impact parameter, $\chi^2_{\text{IP}}(\mu)$, of the muon tracks.

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

Table 4.4: Requirements of the used HLT2 line which bases its decision on the dimuon system. The variables are the χ^2 over the number of degrees of freedom of the dimuon-track fit, $\chi^2_{\text{track}}/\text{ndf}(\mu^+\mu^-)$, the χ^2 of the dimuon-vertex fit, $\chi^2_{\text{vtx}}(\mu^+\mu^-)$, the χ^2 of the flight distance, FD, of the dimuon system, $\chi^2_{\text{FD}}(\mu^+\mu^-)$, and its mass computed by the OVF, $m(\mu^+\mu^-)$.

	Hlt2DiMuonDetachedHeavy
$\chi^2_{\text{track}}/\text{ndf}(\mu^+\mu^-)$	< 5
$\chi^2_{\text{vtx}}(\mu^+\mu^-)$	< 25
$\chi^2_{\text{FD}}(\mu^+\mu^-)$	> 9
$m(\mu^+\mu^-)$	$> 2.95\,\text{GeV}/c^2$

In the case of the $\psi(2S)$ meson it is required that the daughter particles satisfy the muon hypothesis and that the invariant mass of the dimuon system corresponds to the nominal mass of the $\psi(2S)$ meson. Besides this it is necessary that the two tracks form a common vertex, that is displaced from the primary vertex. The requirements in the case of the K_S^0 meson are chosen accordingly. The pions should satisfy the pion hypothesis and their mass should correspond to the nominal mass of the K_S^0 meson. The tracks should also form a common vertex which is significantly displaced from the primary interaction point. It is also required that the K_S^0 meson has a non-vanishing decay time to suppress backgrounds coming from misidentified decays including the short-lived K^{*0} meson. Furthermore, the invariant mass of the four final state particles should correspond to the nominal mass of the B^0 meson and the vertex fit should be of good quality.

The requirements in the case of the K_S^0 meson and its decay products can differ for the two track type categories. In Tab. 4.6 the values in parentheses represent requirements on K_S^0 mesons which are reconstructed as long track candidates and the nominal values represent requirements on K_S^0 mesons which are reconstructed as downstream candidates. If just one value is given the requisite is the same for both track types.

Table 4.5: Stripping requirements on the reconstructed final state of the $\psi(2S)$ meson. The used variables are the difference of the logarithmic likelihood values of the muon and pion hypotheses, $\Delta \log \mathscr{L}_{\mu\pi}$, the mass computed by the OVF, $m(\mu^+\mu^-)$ compared to the nominal $\psi(2S)$ mass as given in the PDG, $m(\psi(2S), \mathrm{PDG})$ [73], the χ^2 of the distance of closest approach between the two tracks, $\chi^2_{\mathrm{DOCA}}(\mu^+\mu^-)$, and the χ^2 over the number of degrees of freedom of the $\psi(2S)$-vertex fit, $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}(\psi(2S))$.

$\Delta \ln \mathscr{L}(\mu\pi)$	> 0
$\lvert m(\mu^+\mu^-) - m(\psi(2S), \mathrm{PDG})\rvert$	$< 60\,\mathrm{MeV}/c^2$
$\chi^2_{\mathrm{DOCA}}(\mu^+\mu^-)$	< 30
$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}(\psi(2S))$	< 16

The signal efficiency of the combined reconstruction and loose centralised pre-selection, determined as the number of truth-matched candidates with best PV requirements compared to the number of reconstructed candidates that pass the stripping requirements on simulated data compared to the number of generated events that lie in the detector's acceptance, is $\varepsilon_{\mathrm{reco,strip}} \approx 9.7\,\%$.

4.1.4 Multivariate Selection

Multivariate methods are very popular in high energy physics analyses as the size of the datasets grows significantly and new challenges arise. The following paragraphs are loosely based on Refs. [74], [75], [76], and [77]. In this analysis a Boosted Decision Tree, BDT, is used for classification, which is a non-parametric supervised learning method. In this case

Table 4.6: Stripping requirements on the K^0_S meson and on its decay products. The values in parentheses represent requirements on K^0_S mesons, which are reconstructed as long track candidates, and the nominal values represent requirements on K^0_S mesons which are reconstructed as downstream candidates. If only one value is given, the requisite applies for both track types. The used variables are the K^0_S momentum, $p(K^0_S)$, the minimum impact parameter with respect to any other PV, min χ^2_{IP} w.r.t. any PV, the χ^2 of the distance of closest approach between the two tracks, $\chi^2_{DOCA}(\pi^+\pi^-)$, the mass computed by the OVF, $m(\pi^+\pi^-)$, compared to the nominal K^0_S mass as given in the PDG, $m(K^0_S, \text{PDG})$ [73], the χ^2 over the number of degrees of freedom of the K^0_S-vertex fit, $\chi^2_{vtx}/\text{ndf}(K^0_S)$, and the decay time significance, t/σ_t, w.r.t. the associated best PV of the K^0_S-meson candidate.

$P(\pi)$	$> 2\,\text{GeV}/c$	
min χ^2_{IP} w.r.t. any PV (π)	> 4	(> 9)
$\chi^2_{DOCA}(\pi^+\pi^-)$	< 25	
$\lvert m(\pi^+\pi^-) - m(K^0_S, \text{PDG})\rvert$	$< 64\,\text{MeV}/c^2$	$(< 35\,\text{MeV}/c^2)$
$\chi^2_{vtx}/\text{ndf}(K^0_S)$	< 20	
$t/\sigma_t(K^0_S)$	> 5	

Table 4.7: Stripping requirements on the B^0-meson candidate. The variables are the mass computed by the OVF, $m(\mu^+\mu^-\pi^+\pi^-)$, and the χ^2 over the number of degrees of freedom of the B^0-vertex fit, $\chi^2_{vtx}/\text{ndf}(B^0)$.

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

"supervised" means that the training of the classifier is based on labeled input data for which the classification result is already known.

For the training of the BDT, simulated signal data is used as a signal proxy and the upper mass sideband from data for $m(\psi(2S)K^0_S) > 5450\,\text{MeV}/c^2$ is used as a background proxy. These datasets have to meet the trigger and stripping requirements introduced in Sec. 4.1.2 and Sec. 4.1.3, the candidates with the smallest χ^2_{IP} are chosen, and also the DTFs need to have converged. Besides this the minimum χ^2_{IP} of the tracks needs to be greater than zero. The signal efficiency for these prerequisites is $(99.94 \pm 0.01)\,\%$. Additionally, a *train-test split* is used split the data in half; one half for training the machine-learning model and one half for validation.

Boosted Decision Trees consist of a number of binary decision trees [78], where at each node a requirement greater or smaller than a given threshold for a particular feature is met. A schematic view of a binary decision tree is shown in figure Fig. 4.1. Each branch represents the result of this inquiry. The classification rule is then summarised along the path from the root to the leaves at the end. Boosting implies the idea of taking multiple *weak learners*, *i.e.* hypotheses which perform slightly better than random guessing, and combining them

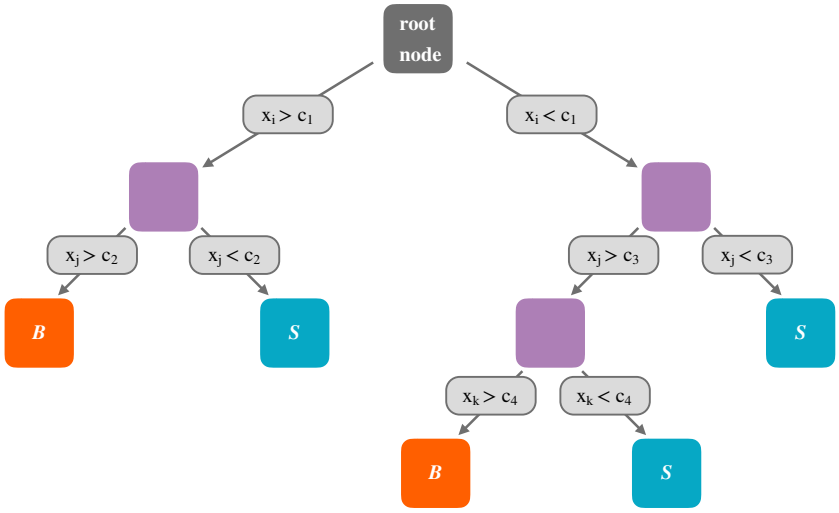

Figure 4.1: A schematic view of a binary decision tree. The variables are represented by $x_{\{i,j,k\}}$ and the requirements by c_i. In the end the input candidates are divided into two classes, signal, *S*, and background, *B*.

into a single *strong learner*, *i.e.* into one accurate hypothesis. Thus, the decision is extended from only one decision tree to several trees which form a *forest*. As a boosting algorithm the *AdaBoost* [75], short for adaptive boosting algorithm, is used, which basically gives candidates that were misclassified in a decision tree a higher candidate weight in the training of the following decision tree. The weights are initialized with $w_i^1 = 1/N$ and the following tree is trained using a modified dataset, where the weights of misclassified candidates from the tree before are multiplied by boosting weights, α_i, which are given by

$$\alpha_i^t = \frac{1-\mathrm{err}_i^t}{\mathrm{err}_i^t}, \tag{4.1}$$

where err_i^t is the misclassification rate for the tth tree, denoted by the superscript. This rate is defined by

$$\mathrm{err}_i^t = \frac{\sum_i \left(w_i^t \,:\, h_i^t(x_i) \neq y_i\right)}{\sum_i w_i^t}, \tag{4.2}$$

where $h_i^t(x_i)$ is the hypothesis given by the tth tree for the ith candidate and input variable x_i with the true label y_i. The resulting weights

$$w_i^{t+1} = w_i^t \cdot (\alpha_i^t \,:\, h_i^t(x_i) \neq y_i) \tag{4.3}$$

are then re-normalized, so that the sum of weights remains constant. In the end, after all

trees are trained, the boosted candidate classification is given by

$$y_{\text{boosted}}(\boldsymbol{x}) = \frac{1}{N_{\text{trees}}} \cdot \sum_{i}^{N_{\text{trees}}} \ln(\alpha_i) \cdot h_i(\boldsymbol{x}), \tag{4.4}$$

where $\boldsymbol{x}$ is the tuple of input variables.

In this analysis the BDT should distinguish between signal and background and while small values for $y_{\text{boosted}}(\boldsymbol{x})$ indicate background-like candidates, large values represent signal-like candidates.

This analysis uses the *AdaBoostClassifier* from the scikit-learn [76] framework. The BDT comprises 50 estimators and is trained with a learning rate of 1.0. The choice of input features for the BDT is based on the experience of previous analyses such as the run I measurement of $\sin(2\beta)$ in the decay channel of $B^0 \to J/\psi\, K^0_S$ [8]. The BDT features comprise the symmetrised χ^2_{IP} for the tracks, *i.e.* the logarithm of the maximum and minimum of the minimum IP χ^2 for the muons and the pions, the lifetime significance of the K^0_S meson, $t/\sigma_t(K^0_S)$, the transverse momenta of the K^0_S and the $\psi(2S)$ candidates, as well as the χ^2 over the number of degrees of freedom of the DTF with constraints on the $\psi(2S)$ and K^0_S-meson mass, and with a constraint on the PV, and the χ^2 of the K^0_S-vertex fit.

A list of the input features is summarised in Fig. 4.2 ranked by their importance in the training of the classifier. The feature importance is defined as the *Gini Importance* or the *Mean Decrease in Impurity* and is calculated as the total decreases in impurity of a node, which is weighted by the probability to reach this node, and then averaged over all trees. The ranking of the input features according to their importance in the training of the multivariate classifier is also confirmed by the comparison of the input features for the signal and background proxy, which are shown in in Fig. 4.3 and Fig. 4.4. The distributions which show a distinctive difference in their distribution between signal and background are generally also the ones with a higher importance for the classifier.

The response of the multivariate classifier is shown in Fig. 4.5 for the signal and background proxy as well as split into train and test sample and the *receiver operating characteristic*, ROC, curve. The ROC curve is determined by plotting the *true-positive rate* against the *false-positive rate* at various cut points of the classifier response. The true-positive rate is identified as the *sensitivity*, *i.e.* is the probability of classifying an actual signal candidate as signal, whereas the false-positive rate represents the *fall-out*, *i.e.* the probability of classifying an actual signal candidate as background. The *area under the curve*, AUC, when using normalised units, represents the probability that a classifier will rank a randomly chosen signal candidate higher than a randomly chosen background candidate. An AUC value of 0.5 means random guessing, whereas a value of one represents perfect performance. In Fig. 4.5 a separation between the signal distribution in blue and the background distribution can be seen, but no significant difference between the respective training (coloured bars) and test sample (data points) is visible. This gives confidence, that the classifier shows no *overtraining*, *i.e.* learning the specific characteristics of the training samples. The mean

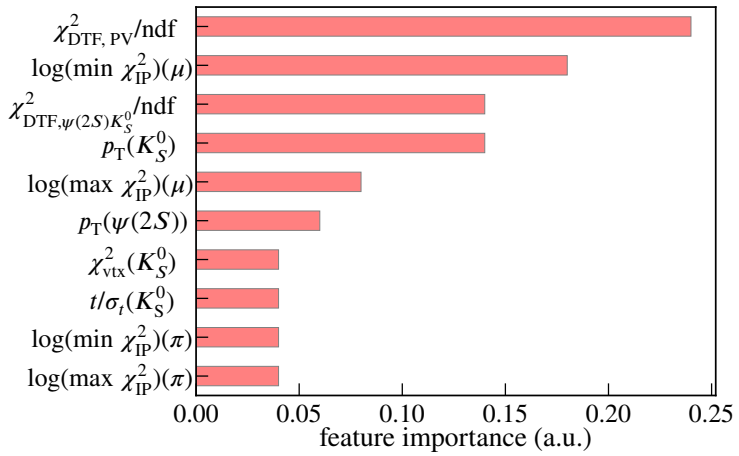

Figure 4.2: A list of the input features used in the BDT in case of the $B^0 \to \psi(2S)K^0_S$ mode ranked by their importance in the training of the classifier, where the x-axis is arbitrarily chosen.

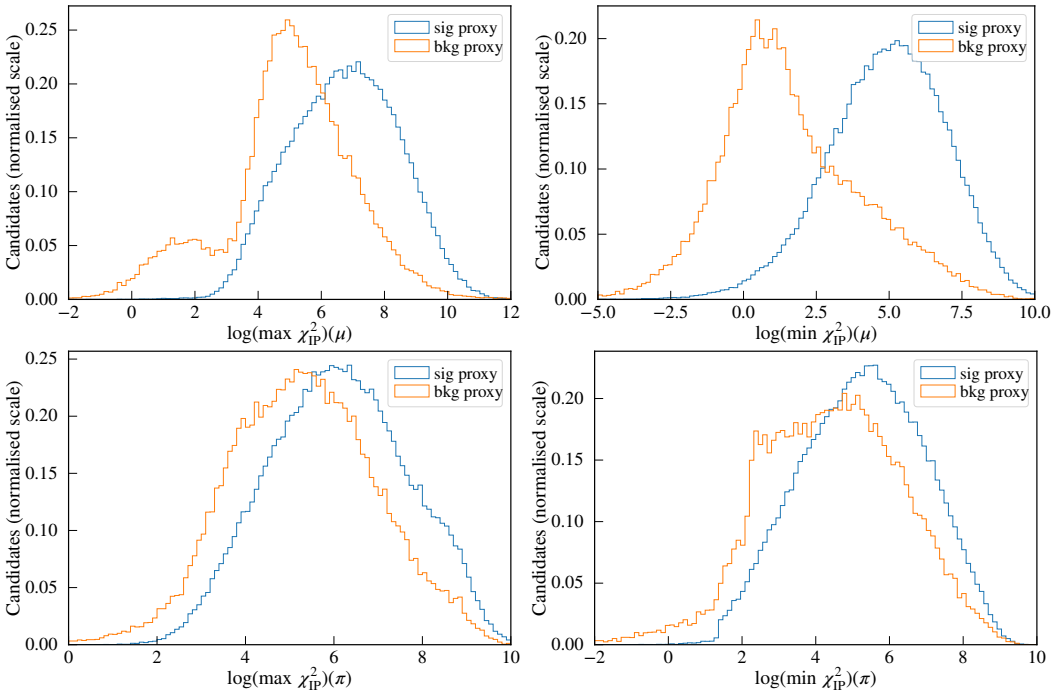

Figure 4.3: The symmetrised χ^2_{IP} for the tracks, *i.e.* the logarithm of the maximum and minimum of the minimum IP χ^2 for the muons and the pions, for the (blue) signal and (orange) background proxy used in the training of the multivariate classifier.

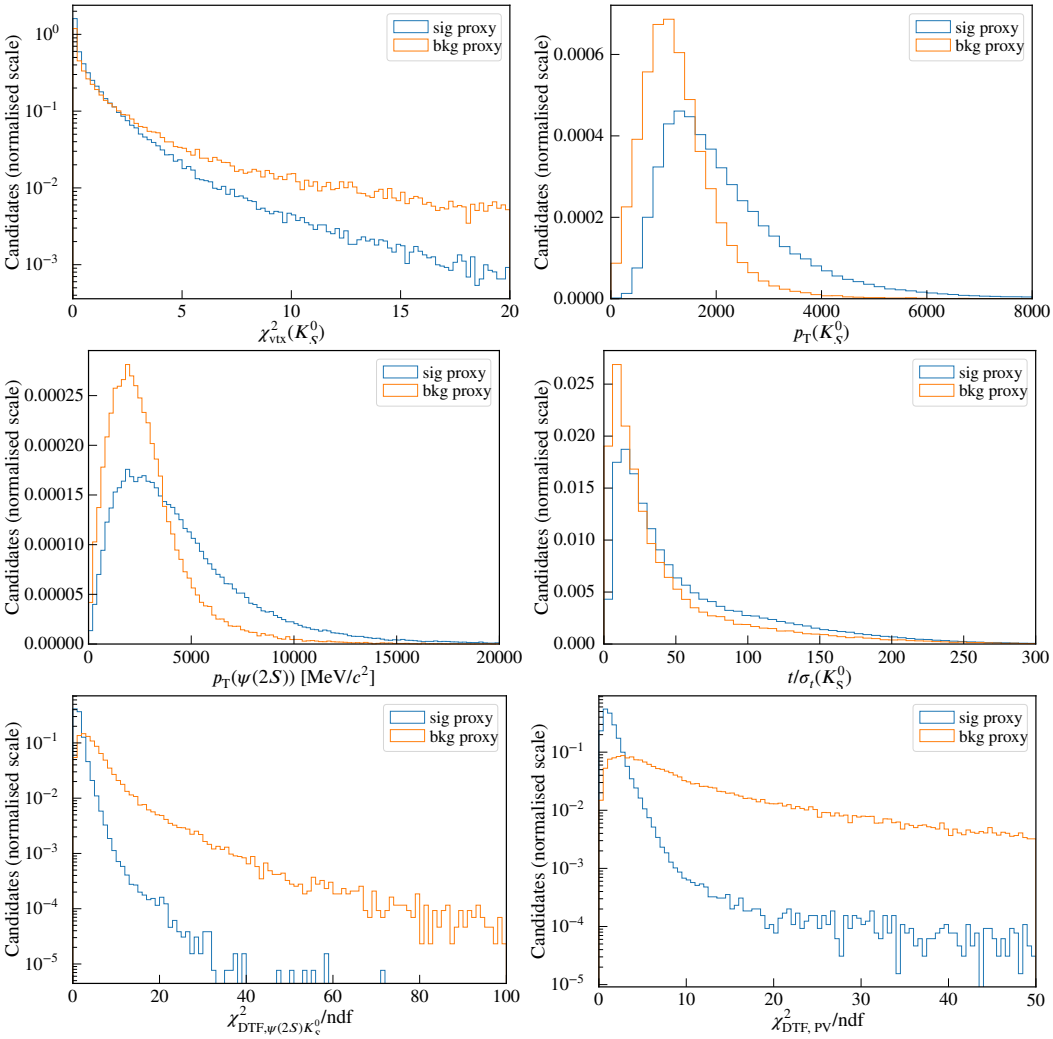

Figure 4.4: Input features of the multivariate classifier related to properties of the daughter particles, the $\psi(2S)$ and the K^0_S meson, as well as to the quality of the DTF with a constraint on the daughter masses and a constraint on the PV for the (blue) signal and (orange) background proxy.

accuracy on the given test dataset around 95 %. The accuracy is given by

$$\text{accuracy}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} \delta_{\hat{y}_i, y_i}. \tag{4.5}$$

It is defined from zero, where none of the predicted values, $\hat{y}_i$, of the ith sample match the true value, y_i, to one, where the predictions strictly match the true values. The given accuracy of 95 % is satisfying. In addition, the AUC value equals 0.9885 and is close to 1. It hints at a well performing classifier.

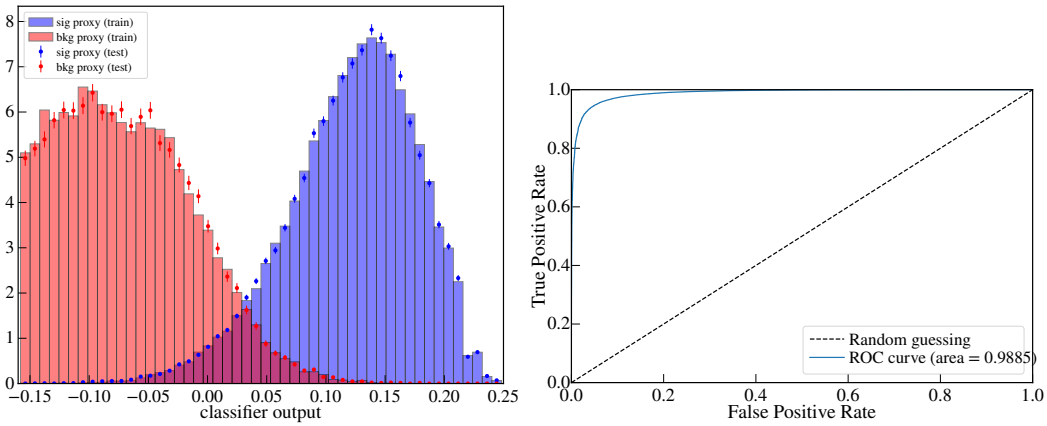

Figure 4.5: The (left) response of the multivariate classifier for the (blue) signal and (red) background proxy for the (coloured bars) train and the (coloured points) test sample and (right) the corresponding ROC curve.

Figure of Merit

It is essential to develop a suitable figure of merit to find the most appropriate threshold of the classifier response. The figure of merit, FOM, used in this thesis considers effects that directly influence the decay-time dependent and flavour-tagged measurement of the *CP* parameter sin(2β) and thus it describes the expected sensitivity on the observable S. A high effective signal size, low decay-time error-estimates, and a low mistag probability will lead to a more precise measurement. As the determination of S is also decay-time dependent, candidates at low decay times around $t' = 0$ ps do not contribute to the measurement as no interference and thus no effect of *CP* violation can have occurred at this time. These events are not that valuable as candidates at decay times of around $t' = 3$ ps, when the sin-function is at its maximum in the B^0 system and the analysis is most sensitive. The used FOM needs to take all these effects into account and is based on Ref. [79]. It is modified to use signal candidate weights, s_w, to determine the effective signal size (see Sec. 4.1.4). This aspect is

developed and implemented in the analysis of the decay of $B^0 \to J/\psi(e^+e^-)K^0_S$. The FOM is defined as

$$\mathrm{FOM} \equiv Q_{\mathrm{mod}} = \frac{\left(\sum_i s_{w_i}\right)^2}{\sum_i {s_{w_i}}^2} \bar{D}^{\mathrm{mod}}, \tag{4.6}$$

where $\bar{D}^{\mathrm{mod}}$ is given by

$$\bar{D}^{\mathrm{mod}} = \frac{1}{\sum_i s_{w_i}} \sum_i (1-2\omega_i)^2 e^{-\left(\Delta m \sigma(t_i)\right)^2} \cdot X_i \cdot s_{w_i}, \tag{4.7}$$

with X_i being

$$X_i = \left[\frac{\sin\left(\Delta m \cdot t_i\right)}{1 + d_i(1-2\omega_i) \cdot S \cdot e^{-\left(\Delta m \sigma(t_i)\right)^2/2} \cdot \sin\left(\Delta m \cdot t_i\right)} \right]^2. \tag{4.8}$$

The per-event decay-time resolution is described by $\sigma(t_i)$ and d_i denotes the tag decision. For determining the optimal cut point on the classifier output, S is fixed to the current world average of $S = 0.691$ [39]. The FOM can either be calculated completely including all effects or separately for the different effects as can be seen in detail in Fig. 4.6. The influence of flavour tagging (see Sec. 4.3) is given by

$$\mathrm{FOM}_{\varepsilon D^2} \equiv \frac{1}{\sum_i s_{w_i}} \sum_i (1-2\omega_i)^2 \cdot s_{w_i}, \tag{4.9}$$

and the effect of the decay-time resolution (see Sec. 4.5.1) is

$$\mathrm{FOM}_{\sigma_t} \equiv \frac{1}{\sum_i s_{w_i}} \sum_i e^{-\left(\Delta m \sigma(t_l)\right)^2} \cdot s_{w_i}. \tag{4.10}$$

The effective signal size, which is comparable to the simple signal significance,

$$\mathrm{FOM}_{\mathrm{sig}} = \frac{N_{\mathrm{S}}}{\sqrt{N_{\mathrm{S}}+N_{\mathrm{B}}}},$$

where N_{S} is the number of signal candidates and N_{B} represents the number of background candidates, is given as

$$\mathrm{FOM}_{S_{\mathrm{eff}}} \equiv \frac{\left(\sum_i s_{w_i}\right)^2}{\sum_i {s_{w_i}}^2}. \tag{4.11}$$

And

$$\mathrm{FOM}_{B^0} \equiv \frac{1}{\sum_i s_{w_i}} \sum_i X_i \cdot s_{w_i} \tag{4.12}$$

represents the $\sin(2\beta)$ term and takes the flavour tagging as well as experimental conditions, and the value of $\sin(2\beta)$ itself into account. The $\mathrm{FOM}_{B^0_s}$, which can be seen in Fig. 4.6, is defined as Q_{mod} with X_i set to 1. The BDT response is scanned and at each cut point of the BDT classifier the FOM is newly calculated.

Maximum-Likelihood Method

Due to the unavoidable measurement uncertainties, parameters and their uncertainties should be estimated as accurately as possible. The following paragraphs are based on Ref. [80]. There may be n measurements of the random variable $\boldsymbol{x}$, which is either a single variable or a vector of variables. The underlying probability density function, PDF, $\mathscr{P}(\boldsymbol{x}|\boldsymbol{a})$, of the values x_1, x_2, ..., x_n should be known, where $\boldsymbol{a}$ stands for one or more unknown parameters on which the PDF is dependent on. The task of the *maximum- likelihood method* is to obtain the best estimate, $\hat{\boldsymbol{a}}$, of the parameters, $\boldsymbol{a}$, from the available data. This method uses the *likelihood function*,

$$L(\boldsymbol{a}|\boldsymbol{x}) = \mathscr{P}(x_1|\boldsymbol{a}) \cdot \mathscr{P}(x_2|\boldsymbol{a}) \cdots \mathscr{P}(x_n|\boldsymbol{a}) = \prod_{i=1}^{n} \mathscr{P}(x_i|\boldsymbol{a}), \tag{4.13}$$

which specifies the probability of obtaining the measured values, $\boldsymbol{x}$, for given parameters, $\boldsymbol{a}$. According to the maximum-likelihood principle the best estimate of the parameters, $\boldsymbol{a}$, is the value, $\hat{\boldsymbol{a}}$, which makes $L(\boldsymbol{a})$ a maximum, *i.e.* which maximizes the probability of obtaining the observed set of x_i values. It is essential, that $\mathscr{P}(\boldsymbol{x}|\boldsymbol{a})$ for any given value of $\boldsymbol{a}$ needs to be normalized. Numerically it is simpler to maximize a sum rather than a product, thus the logarithm of the likelihood function is formed. Furthermore it is more common to solve minimization than maximization problems, which leads to the following adaptation of the likelihood function from Eq. (4.13) to

$$L(\boldsymbol{a}) = -\sum_{i=1}^{n} \ln\left(\mathscr{P}(x_i|\boldsymbol{a})\right). \tag{4.14}$$

In the case of the $B^0 \to \psi(2S)K^0_S$ decay, in the fit to the invariant mass of the B^0 meson, the maximum-likelihood method is performed simultaneously in different subsets of the data, here in categories according to the track types of the K^0_S meson. Additionally, the used PDF comprises the description of different components, here signal and background and therefore a modification is used, the *unbinned, extended maximum-likelihood method*. In this case the likelihood includes the number of candidates, the so-called yield, N_j, for component j. As the number of measured candidates itself follows a Poisson distribution, it is necessary to introduce a factor, which considers this effect. For each category the PDFs for the individual components are summed, and in the end the product of the PDFs for each category is built, which leads to the modified likelihood function

$$L(\boldsymbol{a}|\boldsymbol{x}) = \prod_k \frac{e^{-\sum_j N_j^k}}{n!} \prod_i^{n_t} \sum_j N_j^k \mathscr{P}_j^k(x_i^k|\boldsymbol{a}_j^k), \tag{4.15}$$

where the index $k \in \{\text{DD,LL}\}$ represents the track types, the index $j \in \{\text{sig, bkg}, B^0_s\}$ specifies the components, here the $B^0 \to \psi(2S)K^0_S$ signal, the combinatorial background, and the $B^0_s \to \psi(2S)K^0_S$ background, and N_j^k describes the expected number of candidates of component j in category k. For further details on the mass model, see Sec. 4.1.4.

sPlot Technique and Signal-Candidate Weights

Signal-candidate weights are used in two major steps of this analysis. First, they are determined to calculate the FOM to find the most suitable requirement on the BDT response. Second, afterwards they are determined on the selected dataset in order to separate signal from background candidates as the data can still include background candidates due to imperfections of the selection. This is done because the PDF used to determine $\sin(2\beta)$ needs to describe the signal as well as the background component, if the dataset would still include background candidates. However, the description of background in the decay-time distribution of the B^0 poses difficulties. Thus, the signal is first extracted from the data via the sPlot method [71] and the signal-candidate weights are used in the final fit to the decay-time distribution of the B^0 meson.

This method takes advantage of the fact that the distribution of an observable is known for the single components and that the individual PDFs differ as much as possible from each other. A suitable variable is *e.g.* the invariant mass of the B^0 meson as the signature is very clear. Here, the signal component is described by a peaking Gaussian-like distribution, whereas the combinatorial background is a flat distribution. The physics background is well separated from the signal component (for more details see Sec. 4.1.4). This model is fitted to the data via an extended maximum-likelihood fit (for the likelihood see Eq. (4.15)). For every candidate a signal-candidate weight, sWeight, s_w, is determined. They are defined as follows

$$s_{w,l}(x_i) = \frac{\sum_{j=1}^{N_c} \boldsymbol{V}_{lj} \mathscr{P}_j(x_i)}{\sum_{k=1}^{N_c} N_k \mathscr{P}_k(x_i)} \tag{4.16}$$

for $B^0 \to \psi(2S)K^0_S$ signal, combinatorial background, and $B^0_s \to \psi(2S)K^0_S$ background with $l \in \{\text{sig, bkg, } B^0_s\}$, and $N_c = 3$ due to the three components. The matrix $\boldsymbol{V}_{lj}$ is the covariance matrix, which is determined by the second derivative of the negative likelihood function, $-L(\boldsymbol{a})$.

Model of the Invariant Mass of the B^0 Meson

In order to calculate the signal-candidate weights for computing the FOM and in preparation of the extraction of the *CP* parameters an unbinned, extended maximum-likelihood fit to the reconstructed B^0-meson mass is performed. The description contains specific parts for the single components. The signal component is described by a modified Hypatia PDF [81]:

$$\mathscr{P}(m;\mu,\sigma,\lambda,\zeta,\beta,a_1,n_1,a_2,n_2) \propto \begin{cases} G(m;\mu,\sigma,\lambda,\zeta,\beta) & ,-a_1 < \frac{m-\mu}{\sigma} < a_2 \\ \frac{G(\mu-a_1\sigma,\mu,\sigma,\lambda,\zeta,\beta)}{\left(1-m/\left(n_1 H(\mu-a_1\sigma,\mu,\sigma,\lambda,\zeta,\beta)-a_1\sigma\right)\right)^{n_1}} & ,-a_1 > \frac{m-\mu}{\sigma} \\ \frac{G(\mu-a_2\sigma,\mu,\sigma,\lambda,\zeta,\beta)}{\left(1-m/\left(n_2 H(\mu-a_2\sigma,\mu,\sigma,\lambda,\zeta,\beta)-a_2\sigma\right)\right)^{n_2}} & , \quad a_2 < \frac{m-\mu}{\sigma}, \end{cases} \tag{4.17}$$

with the generalized hyperbolic distribution

$$G(m;\mu,\sigma,\lambda,\zeta,\beta) = \left((m-\mu)^2 + A_\lambda^2(\zeta)\sigma^2\right)^{\frac{1}{2}\lambda-\frac{1}{4}} e^{\beta(m-\mu)} K_{\lambda-\frac{1}{2}}\left(\zeta\sqrt{1+\left(\frac{m-\mu}{A_\lambda(\zeta)\sigma}\right)^2}\right), \tag{4.18}$$

representing the core, with

$$H(m;\mu,\sigma,\lambda,\zeta,\beta) = \frac{G(m,\mu,\sigma,\lambda,\zeta,\beta)}{G'(m,\mu,\sigma,\lambda,\zeta,\beta)}, \tag{4.19}$$

where G' is the derivative of G, and with

$$A_\lambda^2(\zeta) = \frac{\zeta K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}. \tag{4.20}$$

Here, K_λ are cylindrical harmonics. The Hpyatia function has a hyperbolic core and power-law tails. When setting the parameter ζ to zero, the core can be described similarly to a Gaussian distribution by a mean, μ, and a width, σ, while the tails are described by the parameters a_1, a_2 and n_1, n_2. The parameters, a_i, n_i, are determined by a fit to simulated data and are fixed in the fit to data. The parameters ζ and β are set to zero in all fits to the reconstructed B^0-meson mass.

The background component can be divided into the description of the physics background, *i.e.* $B_s^0 \to \psi(2S)K_S^0$, whose signature is very similar to the signal channel, and into the description of the combinatorial background. While the physics background is also described by an Hypatia function, where the parameters of the tails are the same as for the signal component assuming that the shapes do not differ, the combinatorial background is described by an exponential function. The tail parameters are shared between the signal description and the description of the $B_s^0 \to \psi(2S)K_S^0$ component, but the mean is described by the difference between the B_s^0-meson and B^0-meson mass, $\mu_{B_s^0} - \mu_{B^0}$, and is fixed in all fits to the known mass difference as given by the PDG [27]. The values of all constant and floating parameters in the mass fit are summarised in Sec. 4.2.

Result of the Scan of the Multivariate-Classifier Response

The scan of the BDT response is done separately for the two track-type categories. The scans through the classifier response can be found in Fig. 4.6, split for the various parts of the figure of merit (see Sec. 4.1.4), and separately for track-type categories. It is performed from -0.05 to 0.15 and in steps of 0.01 in the classifier response. The parameters of the signal mass model are determined by a scan of the classifier on simulated data. As the parametrisation should not be dependent on the threshold required of the classifier, the values of the tail parameters are determined in the fit, where the BDT classifier is required to be greater than zero and are then fixed for the scan on data. As an example the fits on

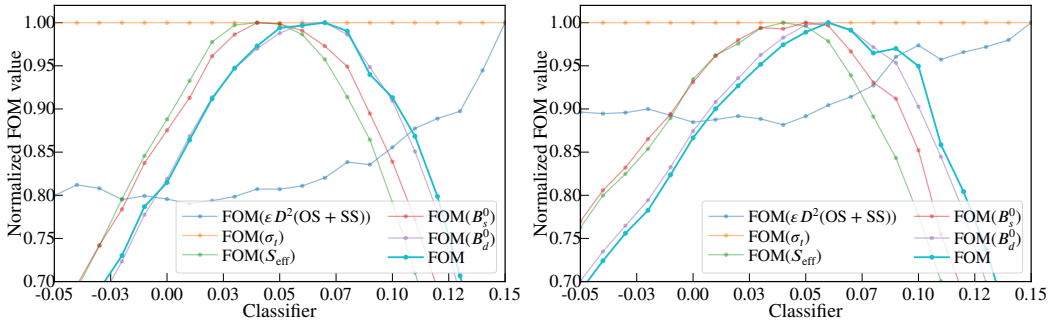

Figure 4.6: Scan through the output of the multivariate classifier split for the two track types, (left) downstream and (right) long track reconstructed candidates. At each cut point the overall figure of merit, as well as the figure of merits considering the different effects on the measurement, are determined.

simulated data and real data to the invariant mass of the B^0 meson are shown in Fig. 4.7 at BDT classifier > 0.0 split for the two track types. The most suitable cut points are chosen to be at the maximum of the FOM and are determined to be

$$\text{classifier response} > \begin{cases} 0.07, & \text{for DD candidates,} \\ 0.06, & \text{for LL candidates.} \end{cases}$$

The combined signal efficiency of the multivariate classifier, again with respect to all previous selection steps as computed on simulated data, is around $\varepsilon_{\text{BDT}} \approx 86.2\,\%$.

4.1.5 Reduction of Background from $\Lambda_b^0 \to \psi(2S)\Lambda^0$ Decays

Various physics backgrounds can influence the mass fit and thus the measurement of $\sin(2\beta)$ severely, but no peaking structures besides the signal channel are observed in the reconstructed B^0-mass distribution, except for the component due to the $B_s^0 \to \psi(2S)K_S^0$ decay channel. This decay is very similar to the signal decay and is irreducible. Due to the proximity of the mass of the B^0 and B_s^0 meson, a high probability of also suppressing signal candidates, when trying to suppress this background follows. Consequently, it is parametrized in the fit to the reconstructed B^0 mass.

As no further peaking structures are observed and based on the background studies of the run I analysis of the $B^0 \to J/\psi K_S^0$ decay channel (see Ref. [8]), only one further physics background is studied, the contribution from $\Lambda_b^0 \to \psi(2S)\Lambda^0$ [13]. Its decay topology is also very similar to the one of the signal channel, as the $\psi(2S)$ meson is also reconstructed from two muons and the Λ^0 is reconstructed from its decay into a pion and a proton. The only distinction in the final state between this background and the signal decay is the exchange of a pion by a proton.

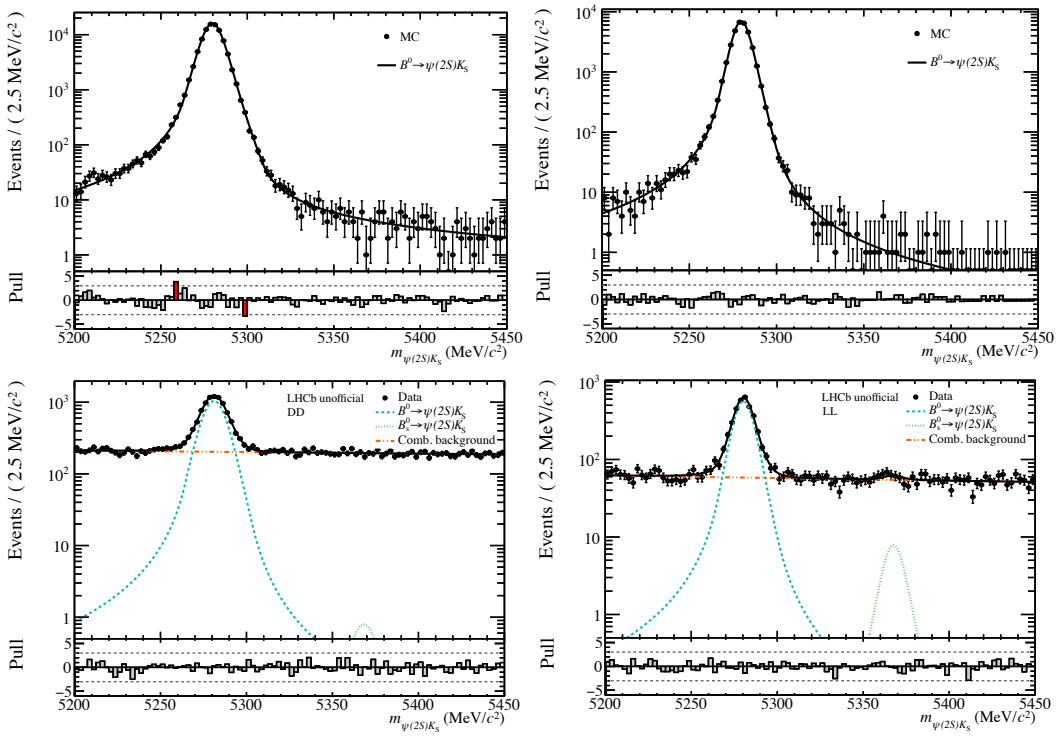

Figure 4.7: Invariant mass distribution of the B^0 meson for (left) downstream and (right) long-track reconstructed candidates and for (top) simulated data and for (bottom) real data at BDT classifier > 0.0. The data points are shown as black points and the projection of the complete PDF is shown as a black line. In the bottom plots, the orange dotted line represents the projection of the component describing the combinatorial background, the blue dotted line represents the signal component, and the green dotted line represents the $B_s^0 \to \psi(2S)K_S^0$ component. The y-axis is in logarithmic scale.

To reduce this background, the mass of the K_S^0 meson is recomputed under the exchange of the pion with the proton hypothesis, such that one of the pions is misidentified as a proton. Candidates that lie within a certain window of this recomputed mass are rejected. The windows are chosen individually for the two track type categories and are

$$\left|m(\pi p) - m(\Lambda^0, \text{PDG})\right| > \begin{cases} 10\,\text{MeV}/c^2, & \text{for DD candidates,} \\ 6\,\text{MeV}/c^2, & \text{for LL candidates.} \end{cases}$$

The veto of the $\Lambda_b^0 \to \psi(2S)\Lambda^0$ background has a signal efficiency computed on top of the previous requirements of $\varepsilon_{\text{veto}} \approx 95.2\,\%$ at a background retention of about 93 %. To compute an estimate of how many $\Lambda_b^0 \to \psi(2S)\Lambda^0$ candidates remain after the veto, the total efficiency of the selection, including the veto is determined on simulated $\Lambda_b^0 \to \psi(2S)\Lambda^0$ data to be $1.5 \cdot 10^{-4}$. Including this calculated efficiency, using the measured branching fractions for $\Lambda_b^0 \to \psi(2S)\Lambda^0$ decays [27, 82], and including the number of produced Λ_b^0 mesons at LHCb, the number of remaining background candidates is estimated to be around 46 ± 10. This low number suggests that the remaining candidates do not pose a problem to the measurement.

4.1.6 Incorrectly Associated PVs and Multiple Candidates

Considering the running conditions at LHCb on average more than one pp collision can be visible per bunch crossing leading to more than one primary vertex per event. Hence, each B^0-meson candidate can be combined with more than one PV, leading to multiple (B, PV) pairs per event. If these combinations pass all selection steps required before, they can lead to large tails in the decay-time distribution of the B^0-meson candidate and thus can bias the determination of the decay-time model. It is therefore necessary to study the impact of incorrectly associated PVs. In the considered dataset that meets all requirements of the introduced selection steps, around 1.0 % are candidates that are possibly associated to the incorrect PV. If more than one PV exists, then on average 2 PVs are reconstructed in the considered dataset.

The usual procedure at LHCb to reduce the amount of multiple (B, PV) pairs and in the end to get unique combinations, is to keep only candidates with the *best* associated PV, *i.e.* the PV with the smallest χ^2_{IP}. The χ^2_{IP} is defined as the difference in χ^2 between the PV fit with and without inclusion of the tracks of the signal daughters. A selection based on the best PV can however lead to biases in the reconstructed decay time and can further lead to a bias of the measurement of decay-time-dependent *CP* violation [83]. Thus, a different approach is developed based on studies in the LHCb analysis of $B_s^0 \to J/\psi K_S^0$ decays [84]. The associated PV is chosen according to the smallest χ^2_{IP} with respect to any other PV in the event for a given (B, PV) pair.

As the true information is given in simulation, a control variable is developed, which is

defined as

$$z_{\text{pull}} = \frac{z_{\text{PVassoc.}} - z_{\text{PVtrue}}}{\sigma_{z_{\text{PVassoc.}}}}, \tag{4.21}$$

where z_{PVtrue} is the true, known z-position of the PV, $z_{\text{PVassoc.}}$ is the z-position of the reconstructed and associated PV, and $\sigma_{z_{\text{PVassoc.}}}$ represents the uncertainty of the determination of the associated PV. The distribution of this variable for simulated data is shown in Fig. 4.8, with stripping and trigger requirements applied for a larger and a zoomed in range. The shape of the distribution close to zero is very Gaussian-like with a width of about 9 and it is assumed that the Gaussian-core represents (B, PV) pairs with correctly assigned PVs. Thus, very conservatively those candidates with $\left|z_{\text{pull}}\right| < 30$ are defined as candidates with correctly associated PVs and all other candidates are classified as candidates with incorrectly associated PVs. This classification results in the decay-time distributions given in Fig. 4.9,

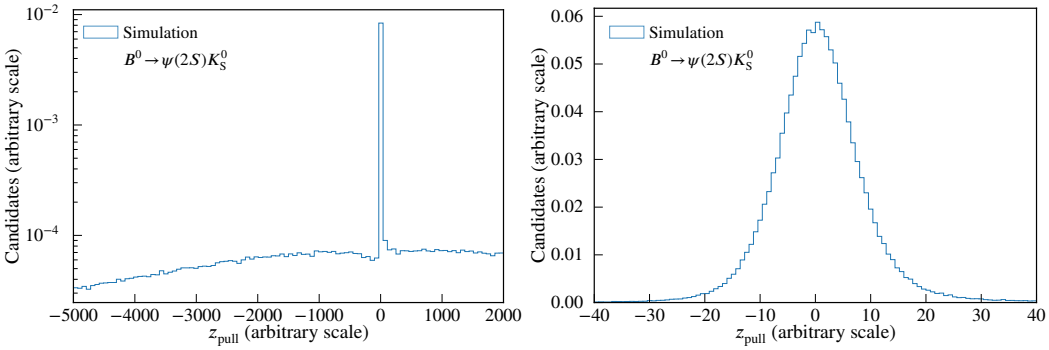

Figure 4.8: The distribution of the difference of the true PV position compared to the reconstructed position normalized to the uncertainty of the reconstruction on simulated data (left) for a wide range from −5000 to 2000 in logarithmic scale and (right) zoomed in on the range from −40 to 40 in arbitrary units.

where it is obvious, that the incorrectly associated PVs are responsible for the large tails in the distribution.

The distribution of the developed variable, the smallest χ^2_{IP} with respect to any other PV in the event for a given (B, PV) pair, on simulated data is shown in Fig. 4.10, again split according to z_{pull} from Eq. (4.21) in correctly and incorrectly associated PV candidates. Based on these distributions a loose cut on the smallest χ^2_{IP} with respect to any other PV greater than 5 is chosen. The resulting decay-time distributions on simulated data are shown in Fig. 4.10, where the large tails in the distribution for incorrectly associated PVs at higher decay times is significantly reduced. Furthermore, fits of the decay-time distribution to simulated data result in the correct input values of the lifetime of the B^0 meson in the simulation, after handling of the incorrectly associated PVs, while before the fit results differed from the input parameters of the simulation significantly. That gives confidence in the used method.

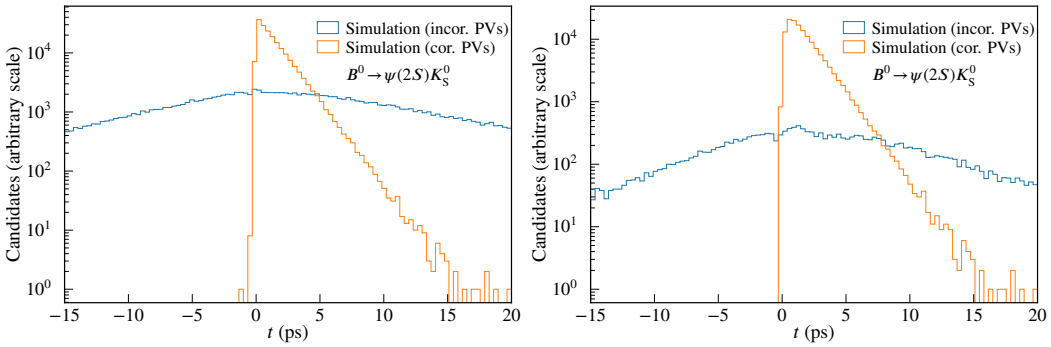

Figure 4.9: Decay-time distribution of reconstructed B^0 candidates on simulated data classified according to the difference of the true PV position compared to the reconstructed position in (orange) correctly and (blue) incorrectly associated PVs (left) after the applied preselection requirements and (right) after the applied preselection and the multivariate selection excluding the requirement on the decay-time distribution in the range of 0 ps to 15 ps for better visualization.

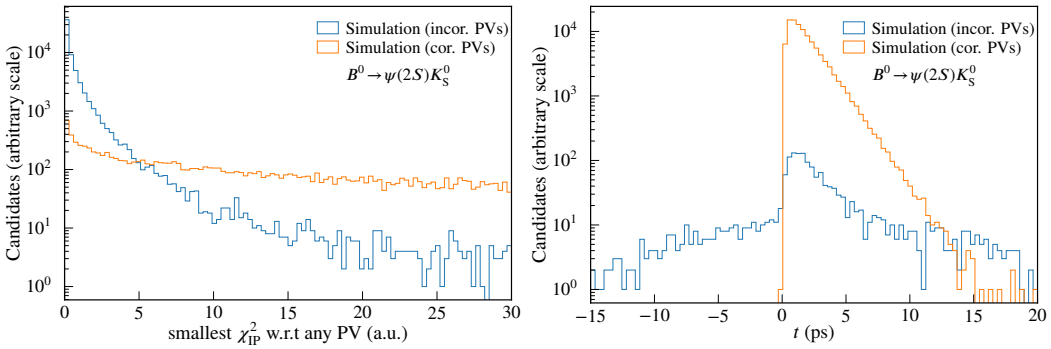

Figure 4.10: The (left) distribution of the smallest χ^2_{IP} with respect to any other PV in the event for (orange) correctly and (blue) incorrectly associated PVs on simulated data after all preselection requirements are applied. And (right) the decay-time distribution of the reconstructed B^0 candidates on simulated data classified according to the difference of the true PV position compared to the reconstructed position in (orange) correctly and (blue) incorrectly associated PVs after the full selection including the requirement on the smallest χ^2_{IP} w.r.t. all other PVs in the event is applied.

Applying the criterion to suppress incorrectly associated PVs has a signal efficiency of $\varepsilon_{\text{wPV}} \approx 98.5\,\%$. After this approximately 0.17 % of candidates are still associated to the incorrect PV.

Furthermore, not only multiple PVs per event can occur, but also multiple B^0-meson candidates. For more information see Ref. [83]. After all preceding selection steps, stripping, trigger, multivariate classifier, and the suppression of incorrectly associated PVs not all of the candidates are unique ones within an event, more precisely around 0.4 % of all candidates are multiple candidates. As the presence of more than one B^0 candidate within an event is highly unlikely, a candidate is chosen randomly in the end, as all the candidates that passed the selection steps before are equally likely to be real signal candidates. After this random selection the number of candidates in the data sample equals the number of events in the sample. In the final dataset 10 844 possible B^0 candidates remain. Only the mass fit will be able to show, how many of these candidates are real signal and how many are background.

4.2 Mass Fit

An unbinned, extended maximum-likelihood fit to the invariant mass distribution is performed to determine signal-candidate weights. These weights are used to statistically subtract the background in the *CP* fit (Sec. 4.5) afterwards. The fit is done in the range of 5200 MeV/c^2 to 5450 MeV/c^2. The used model is described in Sec. 4.1.4. Assuming that the shape does not differ between the $B^0 \to \psi(2S)K^0_S$ and $B^0_s \to \psi(2S)K^0_S$, the descriptions share the parameters describing the tails on both sides. They are extracted from simulation and fixed in the fit to data. The widths for both components and the mean of the B^0 component are allowed to vary in the fit, while the mean of the B^0_s component is defined as the offset from the B^0 mean by the known $B^0_s - B^0$ mass difference. The parameters ζ and β are set to zero in the fit to simulated and to real data. Setting β to zero ensures the symmetry of the hyperbolic core of the Hypatia. The mass fit is done simultaneously in the two track-type categories and the results are listed in Tab. 4.8 and the fit is shown in Fig. 4.11.

4.2.1 Estimation of the $B^0_s \to \psi(2S)K^0_S$ Branching Ratio

An additional study is performed, where a preliminary and unofficial value for the ratio of branching fractions, $\mathcal{B}$, of the decays $B^0_s \to \psi(2S)K^0_S$ and $B^0 \to \psi(2S)K^0_S$, $\mathcal{B}(B^0_s)/\mathcal{B}(B^0)$, is determined. The number of measured B^0 or B^0_s candidates, $n(\{B^0, B^0_s\})$ at LHCb is given as

$$n_{\{B^0,B^0_s\}} = \mathcal{L}_{\text{int}} \cdot 2 f_{\{d,s\}} \cdot \sigma(b\bar{b}) \cdot \epsilon \cdot \mathcal{B}(\{B^0, B^0_s\}), \tag{4.22}$$

where $\mathcal{L}_{\text{int}}$ is the integrated luminosity, $f_{\{d,s\}}$ the hadronization factor, *i.e.* the probability that a b quark hadronizes with a d or s quark, $\sigma(b\bar{b})$ is the cross section for a $b\bar{b}$ pair to be produced at LHCb for the given running conditions, and ϵ is the total efficiency, including reconstruction and selection steps. It is convenient to measure ratios of branching fractions,

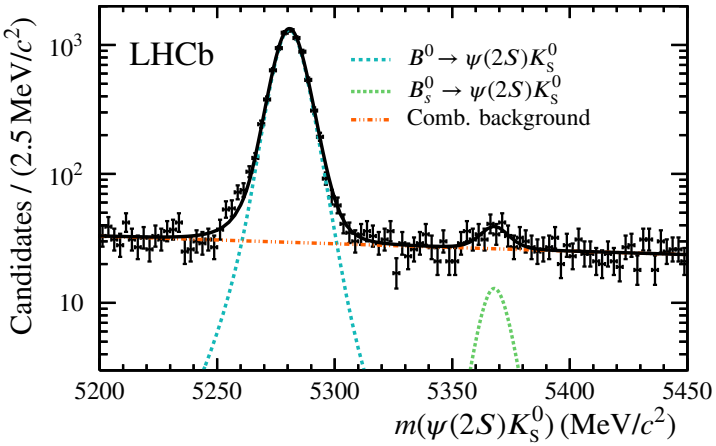

Figure 4.11: Nominal fit to the Invariant mass distribution of the B^0 meson. The data points are shown in black and the projection of the fitted PDF is also shown in black. The orange dashed line represents the projection of the exponential function describing the combinatorial background component, the blue dashed line shows the signal component, $B^0 \to \psi(2S)K^0_s$, and the green dashed line the $B^0_s \to \psi(2S)K^0_s$ component.

Table 4.8: Results of the mass fit to the invariant mass distribution of the B^0 meson in $B^0 \to \psi(2S)K^0_s$ split in track-type categories on run I data.

parameters	downstream	long track	
α_1	2.1669	2.5134	const.
α_2	3.0886	4.4273	const.
n_1	3.8276	3.9685	const.
n_2	3.7279	2.5165	const.
β	0.0	0.0	const.
ζ	0.0	0.0	const.
$\mu\left(B^0_s\right) - \mu\left(B^0\right)$ [MeV/c^2]	87.200	87.200	const.
λ	-3.5 ± 0.7	-2.4 ± 0.5	floating
σ [MeV/c^2]	7.36 ± 0.24	6.6 ± 0.4	floating
$\mu\left(B^0\right)$ [MeV/c^2]	5281.00 ± 0.10	5280.40 ± 0.12	floating
$n\left(B^0\right)$	5342 ± 80	2626 ± 55	floating
$n\left(B^0_s\right)$	56 ± 17	23 ± 9	floating
a(exp.) [(MeV/c^2)$^{-1}$]	$-0.000\,86 \pm 0.000\,34$	$-0.002\,56 \pm 0.000\,54$	floating
n(comb. bkg)	2035 ± 59	762 ± 36	floating

as most of the factors, *i.e.* uncertainties on the luminosity and the cross section, cancel in the ratio. Additionally, when assuming that the efficiencies are the same for both considered decay channels, results in

$$\frac{\mathscr{B}(B_s^0)}{\mathscr{B}(B^0)} = \frac{n(B_s^0)}{n(B^0)} \cdot \frac{f_d}{f_s}, \quad \Rightarrow n(B_s^0) = \frac{\mathscr{B}(B_s^0)}{\mathscr{B}(B^0)} \cdot \frac{f_s}{f_d} \cdot n(B^0). \tag{4.23}$$

Thus, the yield of the $B_s^0 \to \psi(2S)K_S^0$ decay, $n(B_s^0)$, in the mass model with the likelihood from Eq. (4.15) can be described by the number of $B^0 \to \psi(2S)K_S^0$ candidates, the ratio of hadronization factors, and the ratio of branching fractions which needs to be determined. Another fit to the reconstructed B^0-mass distribution is performed, including the modifications to $n(B_s^0)$, to be able to measure the ratio of branching fractions directly. This extended maximum-likelihood fit is done simultaneously for the two track-type categories, where $\mathscr{B}(B_s^0)/\mathscr{B}(B^0)$ is shared for the categories. Besides this the term f_d/f_s [85] and the difference between the mean of the B_s^0 and B^0 component [27], $\mu\left(B_s^0\right) - \mu\left(B^0\right)$, are Gaussian constrained in the fit and are allowed to vary within their uncertainty. This is done by including the uncertainties of external and fixed inputs in the fit. A Gaussian PDF

$$\mathscr{G}(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{4.24}$$

is used, described by a mean, μ, and a width, σ. Here, the parameter, x, is the variable to constrain and the mean and the width are fixed to the parameter's value and uncertainty, respectively. The exact values are given in Tab. 4.9. The likelihood is then multiplied by the Gaussian PDFs. The same technique is later on also used in the *CP* fit Sec. 4.5.

Table 4.9: Constrained parameters in the fit to determine the ratio of branching fractions of the decays $B_s^0 \to \psi(2S)K_S^0$ and $B^0 \to \psi(2S)K_S^0$, $\mathscr{B}(B_s^0)/\mathscr{B}(B^0)$.

parameters	values	refs.
f_d/f_s	0.259 ± 0.015	[85]
$\mu\left(B_s^0\right) - \mu\left(B^0\right)$ [MeV/c^2]	87.29 ± 0.26	[27]

The ratio of branching fractions is determined to be

$$\frac{\mathscr{B}\left(B_s^0\right)}{\mathscr{B}\left(B^0\right)} = (3.8 \pm 0.9)\,\%,$$

which differs with more than four standard deviations from the null hypothesis, *i.e.* not observing the decay channel, and is comparable to the ratio of branching fractions from the decays $B_s^0 \to J/\psi K_S^0$ and $B^0 \to J/\psi K_S^0$, which is $(4.31 \pm 0.17 \pm 0.12)\%$ [27].

4.3 Flavour Tagging

The precise knowledge of the production flavour of each B^0-meson candidate is essential in a decay-time-dependent analysis of *CP* violation. At LHCb various algorithms were developed that make use of other tracks and decays besides the signal decay in an event to infer from those the initial b flavour. The tagging algorithms, *taggers*, are subdivided into same-side, SS, and opposite-side, OS, taggers. The SS taggers exploit the information of particles created in the fragmentation process of the signal b meson, whereas the OS taggers use decay products of the other, accompanying b hadron that is produced in association with the signal b meson. Besides a decision, *tag*, $d_i' \in \{-1, 0, 1\}$, for each reconstructed candidate, each tagging algorithm provides a mistag-probability estimate, *mistag probability*, $\eta_i \in [0, 0.5]$, *i.e.* the probability that this decision is incorrect. The OS tagging decisions are determined from the charges of muons, *OS muon*, electrons, *OS electron*, or kaons, *OS kaon*; a weighted average of the charges of all tracks, *OS vertex charge*; and the decay products of charm decays possibly originating from the other b hadron in the event, *OS charm*. In the case of the SS taggers the decisions are based on the charges of pions and protons originating from the fragmentation process of the signal b meson,*SS pion* and *SS proton*. The relevant taggers and how they work is schematically shown in Fig. 4.12. The tagging algorithms

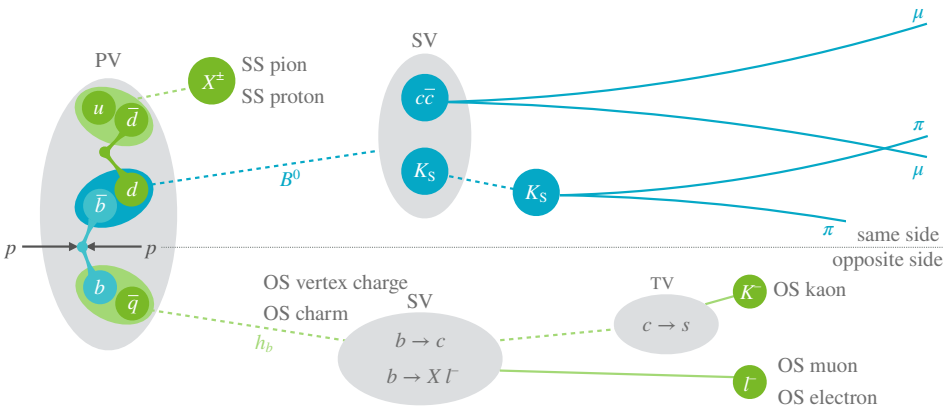

Figure 4.12: A schematic overview of the operating principle of the (green shades) flavour-tagging algorithms at LHCb, including the decay topology of the (blue shades) $B^0 \to [c\bar{c}]K^0_S$ decays.

are in general built very similar. They contain a selection of the relevant tagging particles, *e.g.* muons or pions, and use multivariate classifiers to determine the mistag probability.

The tagging algorithms are not always able to assign a decision. This allows to define a tagging efficiency, ε_{tag}, which is the number of candidates with a decision assigned over the number of all candidates,

$$\varepsilon_{\text{tag}} = \frac{N_{\text{tagged}}}{N_{\text{tagged}} + N_{\text{untagged}}}. \tag{4.25}$$

Besides this, the decisions of the tagging algorithms can sometimes be incorrect, which defines the true average mistag fraction, *mistag*, $\langle\omega\rangle$, as the number of incorrectly tagged candidates over all tagged candidates

$$\langle\omega\rangle = \frac{N_{\text{incorrect}}}{N_{\text{tagged}}}. \tag{4.26}$$

The true average mistag is directly connected to the *dilution*, D, (see also Sec. 4.1.4) given by the flavour tagging,

$$\langle D\rangle = 1 - 2\langle\omega\rangle, \tag{4.27}$$

which is 1 for perfectly tagged candidates and 0 when the tagging algorithms choose the decision randomly. When combining the two properties, the tagging efficiency and the dilution, one can define the effective tagging efficiency, also known as the *tagging power*, as

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}(1 - 2\langle\omega\rangle)^2 = \varepsilon_{\text{tag}}\langle D^2\rangle, \tag{4.28}$$

which represents the statistical reduction factor of a data sample in a tagged analysis and therefore is the figure of merit, that needs to be maximized when developing tagging algorithms.

The mistag probability given by the tagging algorithms is the response of multivariate classifiers, which are trained using datasets of flavour-specific or self-tagging decay channels. This output is then translated via regression into a probability. The training and test samples are very likely to differ in the variables important for the flavour tagging compared to the signal channel, due to different selections and different kinematic properties. Nonetheless, decay channels that are very similar to the signal channel and that enable to infer the b-meson flavour are used to determine a calibration function, $\omega(\eta)$, that transforms the mistag probabilities into the true mistag measured in such a control sample. As a linear dependence between the mistag-probability estimate of the tagging algorithms and the true mistag is observed in most LHCb analyses, the aforementioned calibration functions can be parametrized by

$$\omega(\eta_i) = p_0 + p_1(\eta_i - \langle\eta\rangle), \tag{4.29}$$

where $\langle\eta\rangle$ is the arithmetic mean of the mistag probabilities, and p_0 and p_1 are the parameters of the linear function. The introduction of the average mistag probability allows to decorrelate the parameters p_0 and p_1. A perfect calibration of a tagging algorithm would result in $p_0 = \langle\eta\rangle$ and $p_1 = 1$.

The output and performance of the tagging algorithms is not necessarily independent of the initial b-meson flavour. The selected tagging particles can have different interaction rates with the detector material and this can lead to different tagging efficiencies and true mistags for initial B^0 and $\overline{B}^0$ mesons. This can be considered in two independent calibration

functions for B^0 and $\overline{B}^0$ mesons

$$\begin{aligned}\omega^{B^0}(\eta_i) &= p_0^{B^0} + p_1^{B^0}(\eta_i - \langle\eta\rangle),\\ \omega^{\bar{B}^0}(\eta_i) &= p_0^{\bar{B}^0} + p_1^{\bar{B}^0}(\eta_i - \langle\eta\rangle).\end{aligned} \tag{4.30}$$

The calibration parameters $p_i^{B^0}$ and $p_i^{\bar{B}^0}$, here $i \in 0, 1$ can be written as

$$\begin{aligned}p_i^{B^0} &= p_i + \frac{\Delta p_i}{2},\\ p_i^{\bar{B}^0} &= p_i - \frac{\Delta p_i}{2}.\end{aligned} \tag{4.31}$$

With these definitions Eq. (4.30) can be written as

$$\begin{aligned}\omega^{B^0}(\eta_i) &= p_0 + p_1(\eta_i - \langle\eta\rangle) + \frac{\Delta p_0}{2} + \frac{\Delta p_1}{2}(\eta_i - \langle\eta\rangle),\\ \omega^{B^0}(\eta_i) &= p_0 + p_1(\eta_i - \langle\eta\rangle) - \frac{\Delta p_0}{2} - \frac{\Delta p_1}{2}(\eta_i - \langle\eta\rangle),\end{aligned} \tag{4.32}$$

with the tagging-asymmetry parameters, Δp_i, and thus for the difference of a B^0 meson mistag compared to one of a $\overline{B}^0$ meson follows

$$\begin{aligned}\Delta\omega(\eta_i) &= \omega^{B^0}(\eta_i) - \omega^{\bar{B}^0}(\eta_i)\\ &= \Delta p_0 + \Delta p_1(\eta_i - \langle\eta\rangle).\end{aligned} \tag{4.33}$$

To profit from the performance of the single taggers, the single decisions and mistag probabilities are combined for each of the two types of tagging algorithms for each b-meson candidate, *i.e.* d'_{OS} and d'_{SS}, and η_{OS} and η_{SS}. The true mistag can be translated into a probability, that the flavour d_i is observed given the candidate's initial flavour d'

$$\begin{aligned}p_i(d_i|d') &= \delta_{d_i,d'}(1 - \omega(\eta_i)) + \delta_{d_i,-d'}\omega(\eta_i),\\ &= \frac{1 - d'd_i}{2} + d'd_i(1 - 2\omega(\eta_i)),\\ &= \frac{1 + d'd_i(1 - 2\omega(\eta_i))}{2}.\end{aligned} \tag{4.34}$$

The probabilities, that the candidate is of flavour d' given the observed tagging decision d_i, are given by

$$\begin{aligned}p(\overline{B}^0) &= \prod_i \left(\frac{1 + d_i}{2} - d_i(1 - \omega(\eta_i))\right),\\ p(B^0) &= \prod_i \left(\frac{1 - d_i}{2} + d_i(1 - \omega(\eta_i))\right).\end{aligned} \tag{4.35}$$

The combined probabilities that the signal candidate contained a beauty quark or an antibeauty quark at production, are given by

$$P(\overline{B}^0) = \frac{\max\left(p(\overline{B}^0), p(B^0)\right)}{p(\overline{B}^0) + p(B^0)}, \quad P(B^0) = 1 - P(\overline{B}^0), \tag{4.36}$$

respectively. If $P(B^0) > P(\overline{B}^0)$ the tag decision is $d = 1$ and the mistag is $\omega = 1 - P(B^0)$, otherwise the combined decision is $d = -1$ with $\omega = 1 - P(\overline{B}^0)$. Correlations among the tagging algorithms are not considered, which can lead to an overestimation or an incorrect estimation of the combined probabilities, and need to be corrected. Thus, the combinations as well as the single tagging algorithms need to be calibrated.

Including the calibration (see Eq. (4.29)) of the mistag probabilities an improved tagging power can be obtained when using not only the true average mistag and thus an average dilution as in Eqs. (4.27) and (4.28), but also the per-event properties including the signal-candidate weights, $s_{w,i}$, obtained from the fit to the invariant mass distribution of the B^0 meson and is given by

$$\varepsilon_{\text{eff}} = \frac{\sum_{i=1}^{N} s_{w,i} D_i^2}{\sum_{i=1}^{N} s_{w,i}} = \frac{\sum_{i=1}^{N} s_{w,i}(1 - 2\omega(\eta_i))^2}{\sum_{i=1}^{N} s_{w,i}}, \tag{4.37}$$

where the sum iterates over all candidates including the untagged candidates with $\omega = 0.5$ and $d = 0$.

4.3.1 Flavour-Tagging Strategy

In the case of the $B^0 \to \psi(2S)K^0_S$ mode the calibration functions are determined using the self-tagged decay channel $B^+ \to J/\psi K^+$ for the OS single taggers as well as their combination and using the flavour-specific decay channel $B^0 \to J/\psi K^{*0}$ for the SS single taggers and also for their combination. The J/ψ is reconstructed from two muons for both calibration channels. The K^{*0} is reconstructed from a charged kaon and charged pion.

As the OS single taggers were developed on $B^+ \to J/\psi K^+$ data from run I and possible correlations need to be avoided, the data in this analysis is also split into subsamples as was done for training, testing, and for the combination of the OS single taggers by the flavour-tagging working group when developing the tagging algorithms. The data was split into three independent subsamples according to the event number and the year of data taking. In this analysis only two of these subsets are used. While one dataset is used to calibrate the OS single taggers, the other one is used to calibrate the OS combination and to obtain the performance numbers on the calibration channels. As the SS taggers were not developed on the considered calibration channels, the whole run I $B^0 \to J/\psi K^{*0}$ sample is divided into two parts. The SS single taggers are calibrated on one half of the run I $B^0 \to J/\psi K^{*0}$ sample,

while the SS combination is calibrated and the performance numbers are determined on the other half of the sample.

The strategy is similar to the strategy of the nominal channel. First, a selection for the calibration or *control* modes is developed, which is a simple cut-based selection without using a multivariate classifier. After that a fit to the invariant mass distribution of the b-meson is performed to extract signal candidate weights. Furthermore, the relevant variables for flavour tagging are reweighted to account for kinematic differences between the signal channel and the control channels. Subsequently, first the single taggers are calibrated, then they are combined per side and after that the OS and SS combination are calibrated.

The used datasets were recorded in 2011 and 2012 and correspond to an overall integrated luminosity of $\mathscr{L}_{\text{int}} = 3\,\text{fb}^{-1}$ at centre-of-mass energies of 7 TeV and 8 TeV in 2011 and 2012, respectively. For both control modes the invariant mass of the B^0 meson is, if not stated otherwise, obtained by the OVF, where the mass of the J/ψ candidate is constrained to the known mass of the J/ψ meson given by the PDG [73].

4.3.2 Selection

There are no specific trigger requirements enforced, but again the charmonium mesons can be exploited due to the clean signature of the muons in the detector. The output of the `Bu2JpsiKDetached` and `Bd2JpsiKstarDetached` lines is used. The lines are called *detached*, because candidates with a decay time below 0.2 ps are removed. The requirements in the following tables are summarised for the different stripping conditions. The requirements on the J/ψ meson and its daughters are given in Tab. 4.10 summarised for both stripping lines. The criteria for the K^+ and B^+ meson are listed in Tab. 4.11 and Tab. 4.12. The criteria for the K^{*0} and the B^0 meson are given in Tab. 4.13 and Tab. 4.14.

Table 4.10: Stripping requirements on the reconstructed final state of the J/ψ meson for $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$ summarised for both stripping conditions. The used variables are the difference of the logarithmic likelihood values of the muon and pion hypotheses, $\Delta \log \mathscr{L}_{\mu\pi}$, the transverse momentum of the muons, $p_{\text{T}}(\mu)$, the mass computed by the `OfflineVertexFitter`, $m(\mu^+\mu^-)$, compared to the nominal J/ψ mass as given by the PDG, $m(J/\psi, \text{PDG})$ [73], the χ^2 of the distance of closest approach between the two tracks, $\chi^2_{\text{DOCA}}(\mu^+\mu^-)$, and the χ^2 over the number of degrees of freedom of the J/ψ vertex fit, $\chi^2_{\text{vtx}}/\text{ndf}(J/\psi)$.

$\Delta \log \mathscr{L}_{\mu\pi}$	> 0
$p_{\text{T}}(\mu)$	$< 0.5\,\text{GeV}/c$
$\lvert m(\mu^+\mu^-) - m(J/\psi, \text{PDG})\rvert$	$< 150\,\text{MeV}/c^2$
$\chi^2_{\text{DOCA}}(\mu^+\mu^-)$	< 20
$\chi^2_{\text{vtx}}/\text{ndf}(J/\psi)$	< 16

Table 4.11: Stripping requirements on the K-meson candidates in the case of the $B^+ \to J/\psi K^+$ decay channel summarised for both stripping conditions. The used variables are the difference of the logarithmic likelihood values of the kaon and pion hypotheses, $\Delta \log \mathscr{L}_{K\pi}$, the χ^2 over the number of degrees of freedom of the K-track fit, $\chi^2_{\text{track}}/\text{ndf}(K)$, and the transverse momentum of the kaon, $p_{\text{T}}(K)$.

$\Delta \log \mathscr{L}_{K\pi}$	> 0
$\chi^2_{\text{track}}/\text{ndf}(K)$	< 5
$p_{\text{T}}(K)$	$> 0.5\,\text{GeV}/c$

Table 4.12: Requirements on B^+-meson candidates in the case of the $B^+ \to J/\psi K^+$ decay channel applied in the stripping summarised for both stripping conditions. The used variables are the reconstructed decay time of the B meson, t, the B^+-meson mass computed by the `OfflineVertexFitter`, $m(\mu^+\mu^- K)$, and the χ^2 over the number of degrees of freedom of the B^+-vertex fit, $\chi^2_{\text{vtx}}/\text{ndf}(B^+)$.

$t(B^+)$	$> 0.2\,\text{ps}$
$m(\mu^+\mu^- K)$	$5150\,\text{MeV}/c^2 < m(\mu^+\mu^- K) < 5450\,\text{MeV}/c^2$
$\chi^2_{\text{vtx}}/\text{ndf}(B^+)$	< 10

Table 4.13: Requirements on K^{*0}-meson candidates and on its reconstructed final state in the case of the $B^0 \to J/\psi K^{*0}$ decay channel applied in the stripping summarised for both stripping conditions. The used variables are the χ^2 value of the K^{*0}-vertex fit, $\chi^2_{\text{vtx}}(K^{*0})$, and the transverse momentum of the K^{*0}, $p_{\text{T}}(K^{*0})$, the χ^2 of the distance of closest approach between the two tracks, $\chi^2_{\text{DOCA}}(K,\pi)$, the mass computed by the `OfflineVertexFitter`, $m(K\pi)$, compared to the nominal K^{*0} mass as given by the PDG, $m(K^{*0}, \text{PDG})$ [73], the χ^2 over the number of degrees of freedom of the kaon- and pion-track fit, $\chi^2_{\text{track}}/\text{ndf}(K,\pi)$, respectively, and the difference of the logarithmic likelihood values of the kaon and pion hypotheses $\Delta \log \mathscr{L}_{K\pi}$ for the K.

$\chi^2_{\text{vtx}}(K^{*0})$	< 25
$p_{\text{T}}(K^{*0})$	$> 1300\,\text{MeV}/c$
$m(K\pi)$	$826\,\text{MeV}/c^2 < m(K\pi) < 966\,\text{MeV}/c^2$
$\chi^2_{\text{DOCA}}(K,\pi)$	< 30
$\chi^2_{\text{track}}/\text{ndf}(K)$	< 5
$\chi^2_{\text{track}}/\text{ndf}(\pi)$	< 5
$\Delta \log \mathscr{L}_{K\pi}(K)$	> 0

Table 4.14: Requirements on B^0-meson candidates in the case of the $B^0 \to J/\psi K^{*0}$ decay channel applied in the stripping summarised for both stripping conditions. The used variables are the decay time of the B^0 meson, t, the B^0-meson mass computed by the `OfflineVertexFitter`, $m(J/\psi K\pi)$, and the χ^2 over the number of degrees of freedom of B^0-vertex fit, $\chi^2_{\text{vtx}}/\text{ndf}(B^0)$.

t	$> 0.2\,\text{ps}$
$m(J/\psi K\pi)$	$5150\,\text{MeV}/c^2 < m(J/\psi K\pi) < 5450\,\text{MeV}/c^2$
$\chi^2_{\text{vtx}}/\text{ndf}(B^0)$	< 20

After the stripping selection a basic cut-based selection follows, which tightens some of the stripping requirements. Only the relevant, tightened requirements are summarised in Tab. 4.15 for $B^+ \to J/\psi K^+$s and in Tab. 4.16 for $B^0 \to J/\psi K^{*0}$.

Table 4.15: Cut-based selection in the case of the $B^+ \to J/\psi K^+$ calibration channel. The used variables are the reconstructed J/ψ mass computed by the `OfflineVertexFitter`, $m(\mu^+\mu^-)$, and the transverse momentum of the K meson, $p_{\text{T}}(K)$.

$m(\mu^+\mu^-)$	$3036\,\text{MeV}/c^2 < m(\mu^+\mu^-) < 3156\,\text{MeV}/c^2$
$p(K)$	$> 1000\,\text{MeV}/c$

Table 4.16: Cut-based selection in the case of the $B^0 \to J/\psi K^{*0}$ calibration channel. The used variables are the reconstructed J/ψ mass computed by the `OfflineVertexFitter`, $m(\mu^+\mu^-)$, the minimum impact parameter, with respect to any other PV, for all final state particles, min χ^2_{IP} w.r.t. any PV ($\{\mu, K, \pi\}$), and the transverse momentum of the pion, $p_{\text{T}}(\pi)$.

$m(\mu^+\mu^-)$	$3036\,\text{MeV}/c^2 < m(\mu^+\mu^-) < 3156\,\text{MeV}/c^2$
min χ^2_{IP} w.r.t. any PV (μ)	> 16
min χ^2_{IP} w.r.t. any PV (K)	> 2
min χ^2_{IP} w.r.t. any PV (π)	> 2
$p_{\text{T}}(\pi)$	$> 500\,\text{MeV}/c$

4.3.3 Mass Fits

The invariant mass models for describing the signal and combinatorial background component are taken from the nominal model of the signal decay mode. The signal is modelled by a Hypatia function, whereas the combinatorial background is described by an exponential function. The tail parameters are taken from the nominal mass fit to the signal-decay channel and are fixed in the fit of the calibration channels. Due to the large data sets, the fits are

done separately in years of data taking and magnet polarities, magnet up, MU, and magnet down, MD. The mass observable used in the fits is computed by the DTF in the case of the $B^+ \to J/\psi K^+$ mode, whereas in the case of the $B^0 \to J/\psi K^{*0}$ the mass observable is obtained by the OVF, as the DTF variables were not present in the datasets for the latter decay mode. The results can be found in Tab. 4.18 and Tab. 4.19 for 2011 and 2012, respectively. The parameters that are set constant during the fits are listed in Tab. 4.17. The calculation of the signal-candidate weights is done separately for the subsamples and in the end they are combined to one single signal-candidate weight. The corresponding plots of the mass fits exemplary for the 2012 MU datasets for both control channels can be found in Fig. 4.13. The number of signal candidates obtained in the fits is $N_{B^+ \to J/\psi K^+} = 1\,231\,000 \pm 2300$ in the case of the $B^+ \to J/\psi K^+$ channel and $N_{B^0 \to J/\psi K^{*0}} = 43\,100 \pm 1900$ in the case of the $B^0 \to J/\psi K^{*0}$ mode.

Table 4.17: The constant parameters for the mass fits on calibration channels $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$ are fixed to the same values as in the fit to the signal channel $B^0 \to \psi(2S)K^0_S$.

parameter	value
α_1	2.5134
α_2	4.4273
n_1	3.9685
n_2	2.5165
β	0.0
ζ	0.0

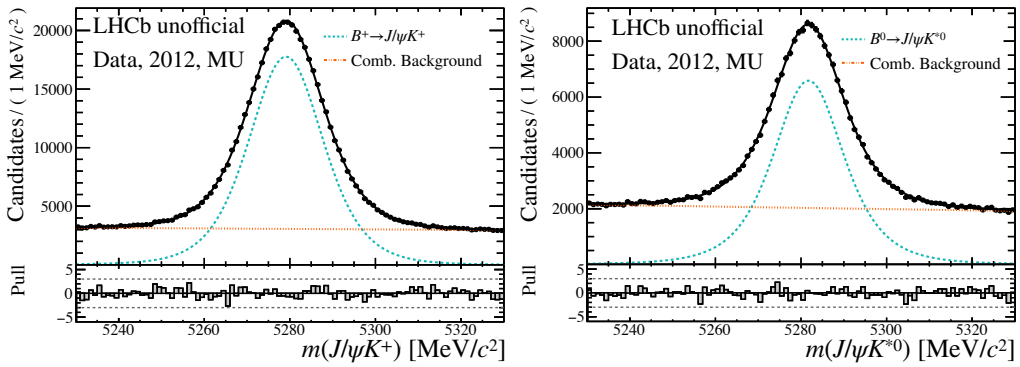

Figure 4.13: Mass fits to the invariant b mass of (left) $B^+ \to J/\psi K^+$ and (right) $B^0 \to J/\psi K^{*0}$ exemplary for the year of data taking of 2012 and magnet-up polarity. The data points are shown in black as well as the complete fit model. The signal component is depicted in blue, while the combinatorial-background component is shown in orange.

Table 4.18: Results of the mass fits for the calibration channels $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$ for 2011 split in magnet polarities, magnet up, MU, and magnet down, MD.

$B^+ \to J/\psi K^+$		
	MD	MU
λ	-3.1 ± 0.2	-3.3 ± 0.2
σ $\left[\text{MeV}/c^2\right]$	11.1 ± 0.1	10.9 ± 0.1
$\mu(B^+)$ $\left[\text{MeV}/c^2\right]$	5279.10 ± 0.03	5279.10 ± 0.03
$n(B^+)$	$225\,837 \pm 960$	$158\,411 \pm 760$
n(comb.bkg)	$128\,470 \pm 908$	$86\,429 \pm 711$
a(exp.) $\left[(\text{MeV}/c^2)^{-1}\right]$	-0.0009 ± 0.0001	-0.0008 ± 0.0001
$B^0 \to J/\psi K^{*0}$		
	MD	MU
λ	-2.0 ± 0.2	-2.7 ± 0.3
σ $\left[\text{MeV}/c^2\right]$	12.2 ± 0.5	10.8 ± 0.4
$\mu(B^0)$ $\left[\text{MeV}/c^2\right]$	5281.40 ± 0.05	5281.70 ± 0.06
$n(B^0)$	$77\,578 \pm 738$	$55\,364 \pm 545$
n(comb.bkg)	$72\,610 \pm 735$	$51\,048 \pm 541$
a(exp.) $\left[(\text{MeV}/c^2)^{-1}\right]$	-0.0007 ± 0.0001	-0.0008 ± 0.0002

Table 4.19: Results of the mass fits for the calibration channels $B^+ \to J/\psi K^+$ and $B^0 \to J/\psi K^{*0}$ for 2012 split in magnet polarities, magnet up, MU, and magnet down, MD.

$B^+ \to J/\psi K^+$		
	MD	MU
λ	-3.2 ± 0.1	-2.9 ± 0.1
σ $\left[\text{MeV}/c^2\right]$	10.8 ± 0.1	10.9 ± 0.1
$\mu(B^+)$ $\left[\text{MeV}/c^2\right]$	5279.10 ± 0.02	5279.10 ± 0.02
$n(B^+)$	$432\,274 \pm 1374$	$414\,865 \pm 1384$
n(comb.bkg)	$319\,670 \pm 1333$	$306\,533 \pm 1345$
a(exp.) $\left[(\text{MeV}/c^2)^{-1}\right]$	-0.0007 ± 0.0001	-0.0007 ± 0.0001
$B^0 \to J/\psi K^{*0}$		
	MD	MU
λ	-2.3 ± 0.2	-1.9 ± 0.1
σ $\left[\text{MeV}/c^2\right]$	11.4 ± 0.4	12.6 ± 0.5
$\mu(B^0)$ $\left[\text{MeV}/c^2\right]$	5281.40 ± 0.04	5281.80 ± 0.04
$n(B^0)$	$146\,500 \pm 1151$	$151\,383 \pm 1238$
n(comb.bkg)	$198\,220 \pm 1174$	$203\,052 \pm 1259$
a(exp.) $\left[(\text{MeV}/c^2)^{-1}\right]$	-0.0009 ± 0.0001	-0.0011 ± 0.0001

4.3.4 Reweighting

Considering the kinematic differences of the calibration channels compared to the signal mode, the selected signal-candidate weights, *sWeighting*, of the calibration channels are weighted to correct for these differences. The observables that influence the decision and response of the tagging algorithms are the pseudorapidity, η, the angle, ϕ, the number of primary vertices in the event, nPV, the number of tracks in the event, nTracks, and the transverse momentum of the B^0-meson candidate, p_{T}.

The multidimensional reweighting uses an algorithm including BDTs called Gradient Boosting Reweighting [86], *GBReweighter*. The standard and typical way to reweight a distribution to another one is to split the variable space in bins and then calculate a multiplier in each bin like

$$m_{\text{bin}} = \frac{w_{\text{bin, target}}}{w_{\text{bin, original}}}, \tag{4.38}$$

which can then be multiplied per candidate to the original distribution to compensate for the difference between the original and the target distribution. Here $w_{\text{bin, target}}$ and $w_{\text{bin, original}}$ are the total weight of candidates in a bin for the target and original distribution, respectively. This procedure is comparable to a histogram division. The advantage of this approach is that it is easy to implement and understand, and it works well in one dimension. But when a reweighting in more than one dimension at a time is needed, this approach can lead to empty bins due to a low number of candidates. The method of the GBReweighter attempts to find the optimal binning scheme, *i.e.* regions or leaves, in the multidimensional variable space using BDTs with a minimum number of bins. It maximizes the symmetrised $\chi^2_{\text{original-target}}$, *i.e.* the difference between the original and target distributions normalized to the sum, which is given by

$$\chi^2_{\text{original-target}} = \sum_{\text{region}} \frac{\left(w_{\text{region, original}} - w_{\text{region, target}}\right)^2}{w_{\text{region, original}} + w_{\text{region, target}}}. \tag{4.39}$$

A high value in $\chi^2_{\text{original-target}}$ suggests that the difference is more significant than at lower values and that the original distribution needs to be reweighted. The GBReweighter algorithm uses a number of trees, which are trained repeating the following steps many times. First, build a tree to maximize the $\chi^2_{\text{original-target}}$ and thus find the most suitable regions, then compute the predictions,

$$p_{\text{region}} = \ln\left(\frac{w_{\text{region, original}}}{w_{\text{region, target}}}\right), \tag{4.40}$$

for each region and at last, reweight the original distribution per candidate using the computed predictions as candidate weights

$$w = \begin{cases} w, & \text{if event from target distribution} \\ w \cdot e^{p_{\text{region}}}, & \text{if event from original distribution.} \end{cases} \tag{4.41}$$

Within each tree the predictions are summed up and the final weight is computed as the product of the contributions of the various trees. Here, the original distributions are the selected calibration channels containing signal-candidate weights and the target distribution is the signal-decay channel, also containing signal-candidate weights. The GBReweighter computes new weights for each candidate of the calibration channels, which is basically the product of the signal-candidate weights and the predictions of the GBReweighter.

Most of the hyperparameters of the GBReweighter are set to their default values, except for the maximum depth, which is set to 3, the learning rate, which is set to 0.01, and the number of estimators, that is set to 500. The hyperparameters are set before the training and define the architecture of the machine-learning algorithm. The run I dataset of the $B^+ \to J/\psi K^+$ mode is split into years of data taking and those subsamples are reweighted separately to the whole run I data set of the signal channel, as the calibration is done for subsamples of the data (see Sec. 4.3.1). The distributions of the reweighted and not- reweighted distributions of the control channels and the target distributions of the $B^0 \to \psi(2S)K^0_S$ channel are shown in Figs. 4.14 and 4.15, exemplary for 2012 in the case of the $B^+ \to J/\psi K^+$ channel and for the whole run I data in the case of the $B^0 \to J/\psi K^{*0}$ channel.

4.3.5 Flavour-Tagging Calibration

The flavour-tagging calibration is done with a tool, provided by the LHCb flavour-tagging working group, the *EspressoPerformanceMonitor* [87], EPM. The calibration parameters of the OS and SS combinations are listed in Tab. 4.20. The corresponding plots are shown in Fig. 4.16. The performances of the combinations are summarised in Tab. 4.21. The OS combination is performing significantly better (2.7 %) than the SS combination (1.1 %). All performance numbers to this point are determined on the reweighted control channels.

Table 4.20: Flavour-tagging calibration parameters of the OS and SS combinations determined on the respective reweighted calibration samples.

Tagger	p_0	p_1	Δp_0	Δp_1	$\overline{\eta}$
OS combination	0.372 ± 0.001	0.815 ± 0.011	0.009 ± 0.002	0.022 ± 0.023	0.3387
SS combination	0.439 ± 0.003	0.870 ± 0.051	0.012 ± 0.004	0.054 ± 0.073	0.4343

Table 4.21: Flavour-tagging performances the OS and SS combinations on the respective reweighted data samples.

Tagger	ε_{tag}	$\varepsilon_{\text{tag}}\langle D^2\rangle = \varepsilon_{\text{tag}}(1-2\omega)^2$
OS Combination	$(29.981 \pm 0.063)\,\%$	$(2.727 \pm 0.007(\text{stat}) \pm 0.041(\text{cal}))\,\%$
SS Combination	$(49.880 \pm 0.153)\,\%$	$(1.074 \pm 0.005(\text{stat}) \pm 0.075(\text{cal}))\,\%$

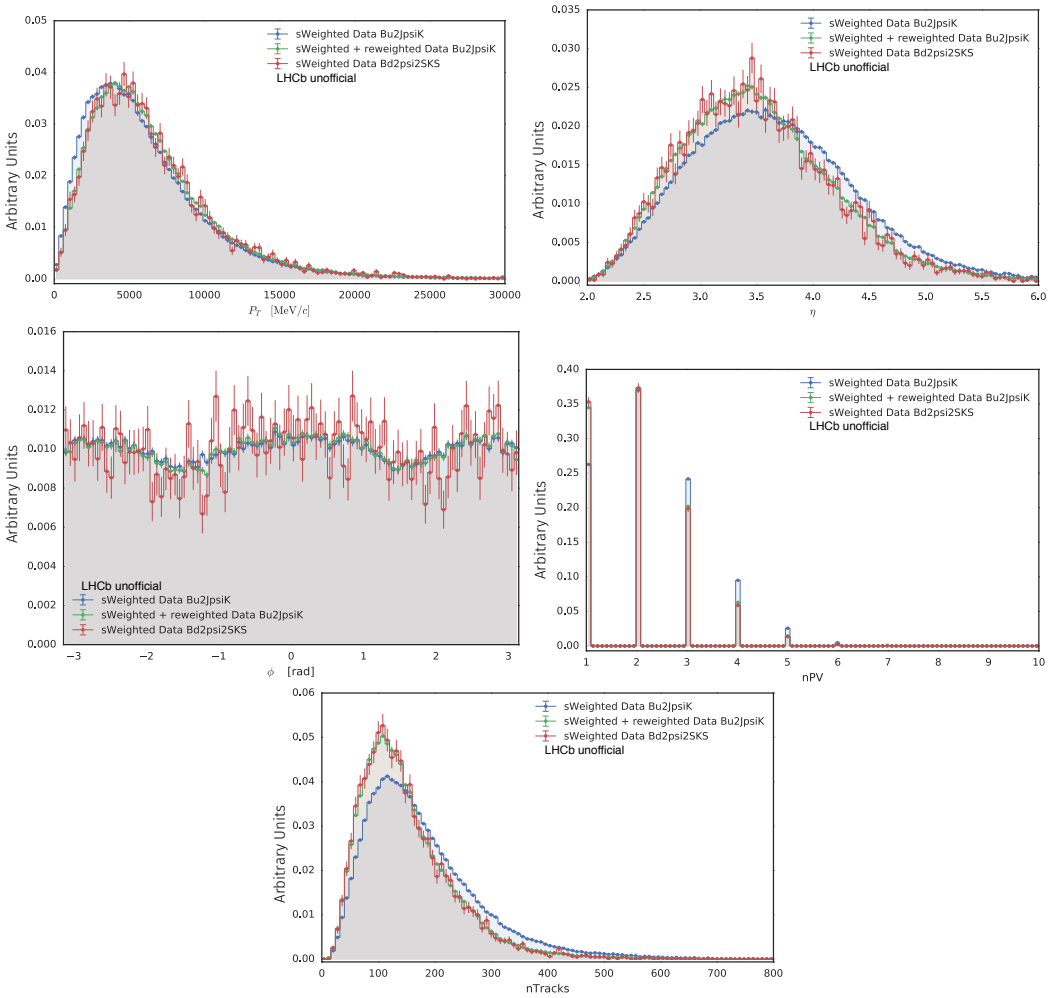

Figure 4.14: The (blue) original distributions of the flavour-tagging-related observables in the case of the sWeighted $B^+ \to J/\psi K^+$ channel, the (green) sWeighted and reweighted distributions also for the $B^+ \to J/\psi K^+$ control channel, and the (red) target distributions in the case of the sWeighted $B^0 \to \psi(2S)K^0_S$ signal channel are shown exemplary for 2012 data.

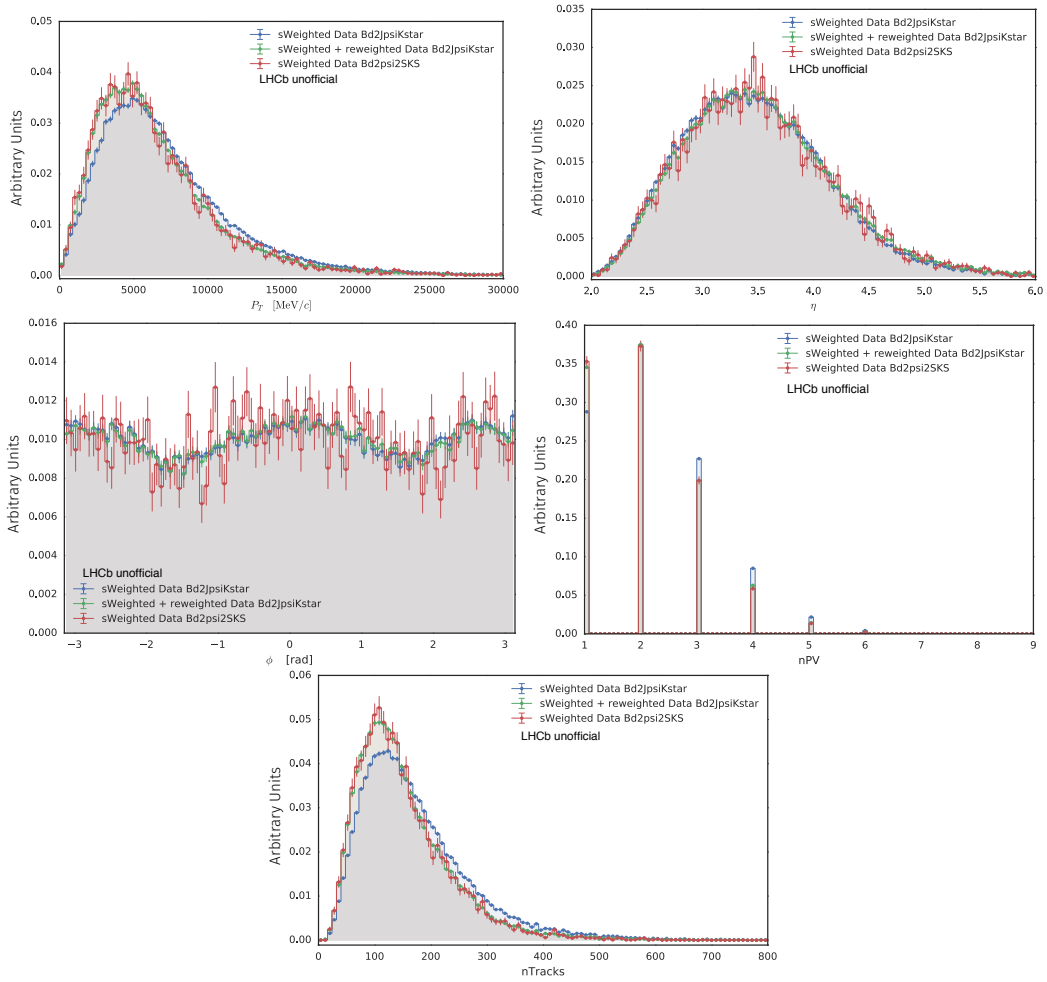

Figure 4.15: The (blue) original distributions of the flavour-tagging-related observables in the case of the sWeighted $B^0 \to J/\psi K^{*0}$ channel, the (green) sWeighted and reweighted distributions also for the $B^0 \to J/\psi K^{*0}$ control channel, and the (red) target distributions in the case of the sWeighted $B^0 \to \psi(2S)K^0_S$ signal channel are shown for run I data.

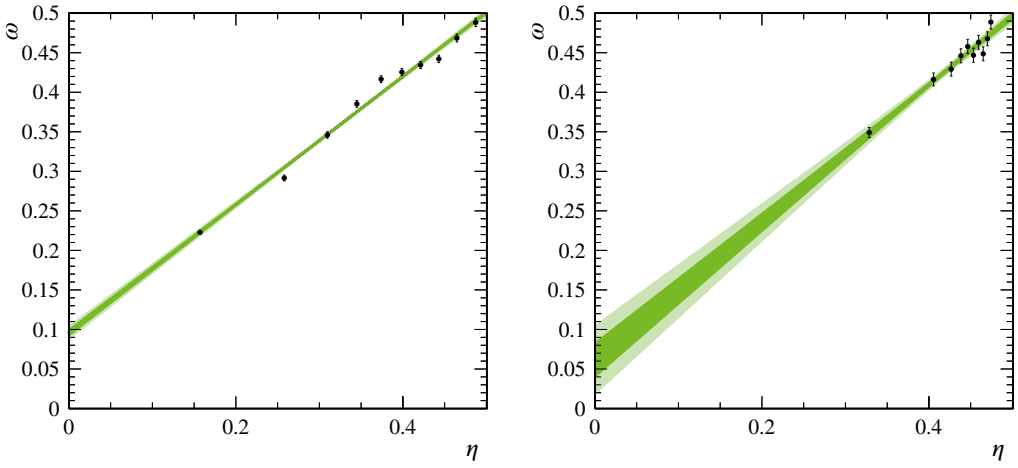

Figure 4.16: Flavour-tagging calibration on the (left) reweighted $B^+ \to J/\psi K^+$ sample for the OS combination and (right) on the reweighted $B^0 \to J/\psi K^{*0}$ sample for the SS combination. The light and dark green shaded areas correspond to the 68 % and 95 % confidence-level regions of the calibrations function.

4.3.6 Systematic Uncertainties

The systematic uncertainties arising from the flavour-tagging calibration comprise two different sources of uncertainties. The first type of uncertainties comes from the method of determining the calibration parameters in the control channels, and the second type results from using these parameters determined on control channels in the fit to the signal channel.

Calibration Method (Type I)

To determine the systematic uncertainty arising from the calibration method itself, the calibration parameters are determined on different subsets of the data sample split according to years of data taking and magnet polarities. The differences of the absolute calibration values for the different categories. The significances of those differences being compatible with zero are determined. As all determined significances are all below one standard deviation, no systematic uncertainty is assigned arising from the calibration method itself.

Portability of the Calibration Parameters (Type II)

A slightly different approach for the SS and OS taggers is used to compute the systematic uncertainties arising form the portability of the calibration parameters. The calibration parameters for the SS taggers can only be determined on the $B^0 \to J/\psi K^{*0}$ channel, while the parameters for the OS taggers can be determined on both control channels. The uncertainties

on the calibration parameters of the SS combination are evaluated by comparing the parameters determined on a sample with and without reweighting of the flavour- tagging-related observables. The same method is also used for determining the systematic uncertainty on the parameters of the OS combination. Additionally, the results of the calibration on both control channels are compared.

When first comparing the results for the calibration of the OS combination on the nominal reweighted calibration channel, $B^+ \to J/\psi K^+$, to the calibration on the other reweighted calibration channel, $B^0 \to J/\psi K^{*0}$, it becomes obvious that the reweighting method is basically working, as no significant differences are observed and the parameters for both channels are compatible. Nonetheless systematic uncertainties are assigned very conservatively, when comparing the results for reweighting and not reweighting for both combinations.

Summary

In the following the calibration parameters for the OS and SS combination are listed with their statistical and systematic uncertainties for $B^0 \to \psi(2S)K^0_S$

$$\begin{aligned} p_0^{\mathrm{OS}} &= 0.372 \pm 0.001\,(\mathrm{stat}) \pm 0.00005\,(\mathrm{syst})\,, \\ p_1^{\mathrm{OS}} &= 0.815 \pm 0.011\,(\mathrm{stat}) \pm 0.005\,(\mathrm{syst})\,, \\ \langle \eta^{\mathrm{OS}} \rangle &= 0.339\,, \\ \Delta p_0^{\mathrm{OS}} &= 0.009 \pm 0.002\,(\mathrm{stat}) \pm 0.00007\,(\mathrm{syst})\,, \\ \Delta p_1^{\mathrm{OS}} &= 0.022 \pm 0.023\,(\mathrm{stat}) \pm 0.002\,(\mathrm{syst})\,, \end{aligned} \tag{4.42}$$

$$\begin{aligned} p_0^{\mathrm{SS}} &= 0.439 \pm 0.003\,(\mathrm{stat}) \pm 0.002\,(\mathrm{syst})\,, \\ p_1^{\mathrm{SS}} &= 0.870 \pm 0.051\,(\mathrm{stat}) \pm 0.002\,(\mathrm{syst})\,, \\ \langle \eta^{\mathrm{SS}} \rangle &= 0.434\,, \\ \Delta p_0^{\mathrm{SS}} &= 0.012 \pm 0.004\,(\mathrm{stat}) \pm 0.001\,(\mathrm{syst})\,, \\ \Delta p_1^{\mathrm{SS}} &= 0.054 \pm 0.073\,(\mathrm{stat}) \pm 0.012\,(\mathrm{syst})\,. \end{aligned} \tag{4.43}$$

These calibration parameters are used in the final fit to extract the *CP* parameters (see Sec. 4.5).

4.3.7 Performance on the Signal Channel

The effective tagging efficiencies for the combination of the OS and SS combination on the complete signal $B^0 \to \psi(2S)K^0_S$ sample, as well as the influence of individual OS and SS combinations on the full dataset are listed in Tab. 4.22 as well as the effective tagging efficiencies determined in the analysis of $B^0 \to J/\psi K^0_S$ [8] for comparison purposes. The total effective tagging efficiency is slightly higher compared to the one from the LHCb run I analysis of $B^0 \to J/\psi K^0_S$ decays, where the effective tagging efficiency is 3.02 % [8]. The

Table 4.22: Effective tagging efficiencies of the single combinations and of their combination on the complete $B^0 \to \psi(2S)K^0_S$ sample. And the effective tagging efficiencies determined in the analysis of $B^0 \to J/\psi K^0_S$ [8].

Tagging algorithm	$\varepsilon_{\text{eff}}^{B^0 \to \psi(2S)K^0_S}$	$\varepsilon_{\text{eff}}^{B^0 \to J/\psi K^0_S}$
OS	$(2.46 \pm 0.05)\,\%$	$(2.63 \pm 0.04)\,\%$
SS	$(1.07 \pm 0.08)\,\%$	$(0.38 \pm 0.02)\,\%$
total (OS + SS)	$(3.42 \pm 0.09)\,\%$	$(3.02 \pm 0.05)\,\%$

increase in effective tagging efficiency results from the inclusion of newly developed tagging algorithms, the OS charm and the SS proton, as well as the usage of an updated SS pion tagger. The slight decrease in effective tagging efficiency for the OS combination arises probably due to the portability of the calibration parameters from the control channels to the signal channel.

4.4 Data Preparation in $B^0 \to J/\psi K^0_S$

The in parallel conducted measurement of $\sin(2\beta)$ in $B^0 \to J/\psi K^0_S$, where the J/ψ meson is reconstructed from two electrons the has a similar selection and flavour tagging strategy to the strategy in the analysis of the $B^0 \to \psi(2S)K^0_S$ decay. Here, also the full run I dataset of LHCb is used for the studies. The selection also consists of trigger and stripping requirements, as well as of a multivariate classifier, where the same FOM from Sec. 4.1.4 is used to determine the appropriate cut point on the classifier output, and of a veto of the physics background coming from $\Lambda^0_b \to J/\psi \Lambda^0$ decays, where the J/ψ meson is reconstructed from two electrons. It is checked, whether the dataset needs to be split according to track-type categories of the K^0_S meson or to the number of reconstructed Bremsstrahlung photons of the electrons. It is found that no split is necessary. In the case of the $B^0 \to J/\psi K^0_S$ mode no dedicated handling of incorrectly associated PVs can be performed, as due to technical reasons, the dataset includes only candidates with the best associated PV, *i.e.* the PV with the smallest χ^2_{IP}, are used. Thus, the effect of incorrectly assigned PVs cannot be studied and it is possible, that this requirement discards potential correctly associated (B, PV) pairs. In the end, if multiple candidates remain after all previous selection steps, one is chosen randomly.

The stripping selection for this specific decay channel requires a decay time of the B^0-meson candidates to be $> 0.2\,\text{ps}$. Furthermore, it requires that the tracks are well reconstructed and form a common vertex in case of the electrons and the pions, respectively. The electrons have also PID requirements imposed. Besides this, the reconstructed masses should be in reasonable ranges.

No specific restrictions on the trigger are required, due to the fact that for such measurements the exact number of single efficiencies is not relevant. The candidates only need to

pass all three trigger stages.

Nevertheless, the data sample contains about 50 M candidates after the trigger and stripping requirements and to be able to handle this huge amount of data, a dedicated pre-selection is applied, before the training of the multivariate classifier. These requirements comprise cuts on the quality of the tracks' vertices, PID and flight distances, as well as some tightening of the stripping cuts. In total, this pre-selection reduces the number of candidates in the data sample by 80 % to 90 %, while preserving around 90 % of signal candidates as computed on the simulated signal sample.

For the multivariate classifier a six-fold k-Folding is used, which means that the BDT is trained and validated on six folds of the data, where for each fold the training is performed on five-sixth of the data and the validation is done on the remaining one-sixth of the data. The number of folds is chosen arbitrarily. For computing the FOM and extracting the signal-candidate weights the same mass model as for the $B^0 \to \psi(2S)K^0_S$ decay channel is used, including a Hypatia function for the signal description and one for the B^0_s background component, as well as an exponential function to model the combinatorial background. Furthermore, as no significant PID requirements are used throughout the selection up to this point, it is checked whether it is necessary to optimize the cut point on the multivariate-classifier response not only based on the FOM but on the PID variable, *i.e.* the probability that the electrons are actually reconstructed as electrons. Thus, a two dimensional scan for through the FOM an dthe PID variable is conducted. But it is found that it has no impact on the choice of the cut point.

The veto of the physics background of $\Lambda^0_b \to J/\psi\,\Lambda^0$ decays is done similarly as in the analysis of $B^0 \to \psi(2S)K^0_S$. The mass of the K^0_S meson is recomputed under the exchange of the pion hypothesis with a proton hypothesis and to improve the signal efficiency, candidates are only vetoed when the pion candidate, whose mass hypothesis is exchanged, is more likely to be a proton than a pion according to PID. In the end, the background is estimated to be below 37 ± 8 candidates.

After this selection chain the signal candidate weights are extracted by an extend-maximum-likelihood fit to the reconstructed B^0 mass in the range of 5150 MeV/c^2 to 5600 MeV/c^2. The result is shown in Fig. 4.17 and the values of the parameters are listed in Tab. 4.23.

The flavour-tagging strategy is similar to the one of $B^0 \to \psi(2S)K^0_S$. The calibration channels are $B^+ \to J/\psi\,K^+$ and $B^0 \to J/\psi\,K^{*0}$, where the J/ψ is also reconstructed from two electrons. While the OS single taggers and their combination are calibrated on the charged decay channel, the SS single taggers and their combination are calibrated on the neutral calibration channel. The selection consists of dedicated stripping and trigger requirements, of a multivariate classifier, which is trained on simulated data as signal proxy and the upper mass side band from data as background proxy. Furthermore, to also account for kinematic differences between the calibration channels and the signal channel, the GBReweigher is used, before determining the calibration parameters. The calibration plots for the SS and OS combination are shown in Fig. 4.18, the corresponding parameters are listed in Tab. 4.24.

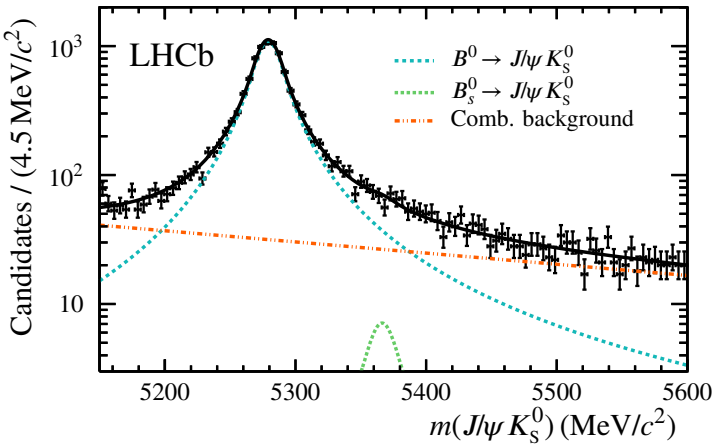

Figure 4.17: Nominal fit to the Invariant mass distribution of the B^0 meson. The data points are shown in black and the projection of the fitted PDF is also shown in black. The orange dashed line represents the projection of the exponential function describing the combinatorial background component, the blue dashed line shows the signal component, $B^0 \to J/\psi K^0_S$, and the green dashed line the $B^0 \to J/\psi K^0_S$ component. The y-axis is in logarithmic scale.

Table 4.23: Results of the mass fit to the reconstructed B^0 mass in $B^0 \to J/\psi K^0_S$ on run I data.

parameter	value	
α_1	0.214 27	const.
α_2	0.221 24	const.
n_1	2.2720	const.
n_2	1.9901	const.
β	0.0	const.
ζ	0.0	const.
λ	-1.05	const.
$\mu(B^0_s) - \mu(B^0)$ $\left[\text{MeV}/c^2\right]$	87.2	const.
σ	57.4 ± 1.2	floating
$\mu(B^0)$ $\left[\text{MeV}/c^2\right]$	5279.1 ± 0.2	floating
$n(B^0)$	$10\,629 \pm 139$	floating
$n(B^0_s)$	70 ± 41	floating
n(comb.bkg)	2696 ± 108	floating
a(exp.) $\left[(\text{MeV}/c^2)^{-1}\right]$	$-0.001\,99 \pm 0.000\,21$	floating

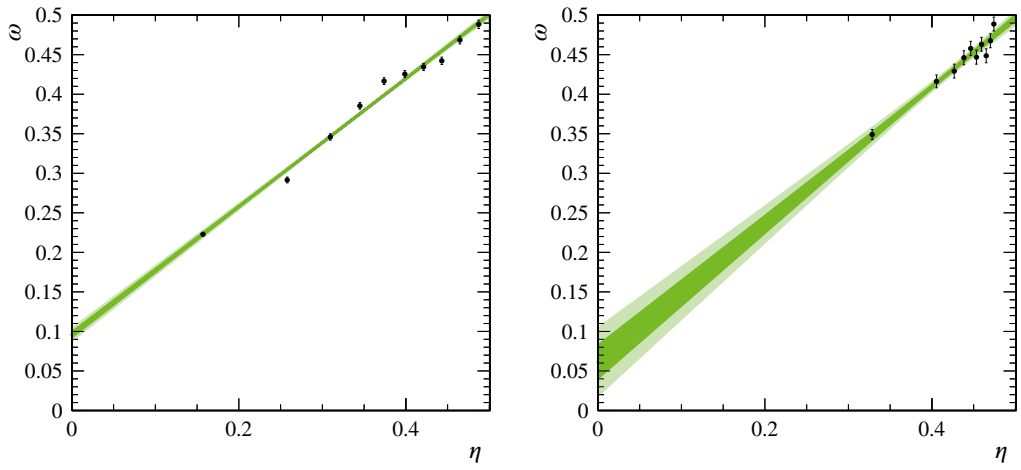

Figure 4.18: Flavour-tagging calibration on the (left) reweighted $B^+ \to J/\psi K^+$ sample for the OS combination and (right) on the reweighted $B^0 \to J/\psi K^{*0}$ sample for the SS combination. The light and dark green shaded areas correspond to the 68 % and 95 % confidence-level regions of the calibrations function, determined for the $B^0 \to J/\psi K^0_S$ decay channel

Table 4.24: Flavour-tagging calibration parameters of the OS and SS combinations determined on the respective reweighted calibration samples.

	p_0	p_1	Δp_0	Δp_1	$\overline{\eta}$
OS combination	0.3603 ± 0.0031	0.834 ± 0.029	0.0162 ± 0.0061	-0.024 ± 0.057	0.316
SS combination	0.4213 ± 0.0063	1.29 ± 0.12	0.0026 ± 0.0086	0.04 ± 0.18	0.425

Determining the systematic uncertainties coming from the calibration method itself and from the portability of the calibration parameters to the signal channel results in the following calibration parameters including statistical and systematic uncertainties

$$\begin{aligned} p_0^{\mathrm{OS}} &= 0.360 \pm 0.003\,(\mathrm{stat}) \pm 0.0006\,(\mathrm{syst})\,, \\ p_1^{\mathrm{OS}} &= 0.029 \pm 0.316\,(\mathrm{stat}) \pm 0.0019\,(\mathrm{syst})\,, \\ \langle \eta^{\mathrm{OS}} \rangle &= 0.316\,, \\ \Delta p_0^{\mathrm{OS}} &= 0.016 \pm 0.006\,(\mathrm{stat}) \pm 0.0003\,(\mathrm{syst})\,, \\ \Delta p_1^{\mathrm{OS}} &= -0.024 \pm 0.057\,(\mathrm{stat}) \pm 0.0038\,(\mathrm{syst})\,, \end{aligned} \tag{4.44}$$

$$\begin{aligned} p_0^{\mathrm{SS}} &= 1.287 \pm 0.121\,(\mathrm{stat}) \pm 0.0150\,(\mathrm{syst})\,, \\ p_1^{\mathrm{SS}} &= 0.421 \pm 0.006\,(\mathrm{stat}) \pm 0.0009\,(\mathrm{syst})\,, \\ \langle \eta^{\mathrm{SS}} \rangle &= 0.425\,, \\ \Delta p_0^{\mathrm{SS}} &= 0.036 \pm 0.175\,(\mathrm{stat}) \pm 0.0157\,(\mathrm{syst})\,, \\ \Delta p_1^{\mathrm{SS}} &= 0.003 \pm 0.009\,(\mathrm{stat}) \pm 0.0009\,(\mathrm{syst})\,. \end{aligned} \tag{4.45}$$

The performance evaluated on the signal channel $B^0 \to J/\psi K_S^0$ yields an effective tagging efficiency of the two different tagging algorithms on the complete data sample; and split for OS exclusively tagged events, SS exclusively tagged events, and events that are tagged by both algorithms (overlap); and the combination on the complete data sample as given in Tab. 4.25.

Table 4.25: Effective tagging efficiency of the single combinations and of their combination on the complete $B^0 \to J/\psi K_S^0$ sample.

Algorithm	$\varepsilon_{\mathrm{eff}}^{B^0 \to J/\psi K_S^0}$
OS	$(3.60 \pm 0.12)\,\%$
SS	$(2.40 \pm 0.28)\,\%$
total (OS + SS)	$(5.93 \pm 0.29)\,\%$

Here, all three determined effective tagging efficiencies are significantly higher than the ones determined in the run I analysis of $B^0 \to J/\psi K_S^0$ [8] (see Tab. 4.22). This cannot be explained only by the inclusion of the newly developed and updated tagging algorithms, as the difference can also be seen when comparing to the performance determined on the $B^0 \to \psi(2S) K_S^0$ channel (see Tab. 4.22), where the combination also includes the new tagging algorithms.

The higher tagging performance of the OS taggers in the $B^0 \to J/\psi K_S^0$ channel compared to the $B^0 \to \psi(2S) K_S^0$ mode is due to the dependence on the considered trigger requirements.

In the case of the $B^0 \to J/\psi K^0_S$ channel no specific restrictions on the trigger are required, which leads to a higher fraction of TIS candidates, compared to the $B^0 \to \psi(2S)K^0_S$ mode. The tagging power of TIS candidates is around three times higher for these candidates than the one from TOS candidates, because an event, that is triggered independent of the signal, contains other particles than just the signal candidate and thus the probability to find an OS b hadron is higher than for TOS candidates. The increase in tagging power for the SS tagging algorithms, comes from the increase in B^0 transverse momentum and the harder momentum spectrum in the electron mode compared to the $B^0 \to \psi(2S)K^0_S$ mode.

4.5 *CP*-Asymmetry Fit

The *CP* observables are determined through a simultaneous, unbinned maximum-likelihood fit to the reconstructed decay-time distribution of the B^0 meson for all $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi K^0_S$ candidates. The signal-candidate weights are extracted from an unbinned, extended maximum-likelihood fit to the invariant mass of the B^0 meson. The negative logarithm of the likelihood is minimized using MINUIT [88], that is implemented in the RooFit package [89].

4.5.1 *CP* Model

The decay-time PDF,

$$\mathscr{P}\left(t', d'_{\mathrm{OS}}, d'_{\mathrm{SS}} | \sigma_t, \eta_{\mathrm{OS}}, \eta_{\mathrm{SS}}\right) = \left[\mathscr{P}_{CP}\left(t, d'_{\mathrm{OS}}, d'_{\mathrm{SS}} | \eta_{\mathrm{OS}}, \eta_{\mathrm{SS}}\right) \otimes \mathscr{R}\left(t' - t | \sigma_t\right)\right] \cdot \epsilon\left(t'\right), \tag{4.46}$$

describes the measured decay time of the B^0-meson candidates, t', and the observed tags, $\vec{d}' = (d'_{\mathrm{OS}}, d'_{\mathrm{SS}})$. It depends further on the per-candidate decay-time-uncertainty estimates, σ_t, and mistag-probability estimates, $\vec{\eta} = (\eta_{\mathrm{OS}}, \eta_{\mathrm{SS}})$, and thus, it is called *conditional*. Here, t represents the proper decay time of the B^0-meson candidates. Experimental effects like the decay-time resolution are accounted for by convolving the PDF, that contains the *CP* observables, $\mathscr{P}_{CP}$, with the function $\mathscr{R}(t' - t|\sigma_t)$ and to consider a decay-time-dependent reconstruction-efficiency function, $\epsilon(t')$, which takes distortions coming from reconstruction and selection into account, the PDF is multiplied by this function. For more information see Sec. 4.5.1.

In the following paragraph the PDF, which contains the *CP* observables to measure, $\mathscr{P}_{CP}$, is derived from Eq. (2.42) and considers additionally the production asymmetry between B^0 and $\overline{B}^0$ mesons, A_{P}. The PDF for the distribution of the true tag, $d \in \{-1, +1\}$ for a $\overline{B}^0$ and B^0 meson, respectively, and the true decay time, t, is given by

$$\mathscr{P}(t, d) = \mathscr{N}(t, d)\left(1 - d A_{\mathrm{P}}\right) e^{-t/\tau} \cdot \left(H(t) + d T(t)\right), \tag{4.47}$$

with the hyperbolic and trigonometrical terms

$$\begin{aligned} H(t) &= \cosh\left(\frac{\Delta\Gamma}{2}t\right) + D\sinh\left(\frac{\Delta\Gamma}{2}t\right), \\ T(t) &= -S\sin\left(\Delta m t\right) + C\cos\left(\Delta m t\right), \end{aligned} \tag{4.48}$$

and the normalization factor, $\mathcal{N}(t,d)$, chosen that

$$\sum_d \int_{t_{\min}}^{t_{\max}} \left(1 - d\,A_{\mathrm{P}}\right) e^{-t/\tau} \cdot \left(H(t) + dT(t)\right) \mathrm{d}t = 1. \tag{4.49}$$

In contrast to the true distribution $\mathscr{P}(t,d)$ from Eq. (4.47), the PDF used in the likelihood fit, $\mathscr{P}(t,\vec{d}')$, needs to consider observed tags, $\vec{d}'$.

For a single tagger, i, the probability for observing a specific tag, d_i', given the true flavour, d, is given by Eq. (4.34), *i.e.* a correct decision , $d_i' = d$, is obtained with a probability of $1-\omega_i$ and an incorrect decision, $d_i' \neq d$, is obtained with a probability of ω_i. Besides this, a tag decision can be assigned with a probability of the tagging efficiency, $\varepsilon_{\mathrm{tag},i}$, and no decision can be assigned with the probability of $1-\varepsilon_{\mathrm{tag},i}$. Thus, when assuming the different tagging algorithms to be uncorrelated, the probability for observing tags given the true flavour of multiple taggers, here {OS,SS}, can be described by

$$\mathscr{P}\left(\vec{d}'|d\right) = \prod_{i\in\{\mathrm{OS,SS}\}} \varepsilon_{\mathrm{tag},i}\delta_{|d_i'|,1}\left(\frac{1+d\,d_i'(1-2\omega_i)}{2}\right) + \delta_{d_i',0}\left(1-\varepsilon_{\mathrm{tag},i}\right). \tag{4.50}$$

The PDF $\mathscr{P}(t,\vec{d}')$ can now be expressed using Eq. (4.47) by

$$\mathscr{P}(t,\vec{d}') = \sum_{d\in\{-1,+1\}} \mathscr{P}(\vec{d}'|d)\mathscr{P}(t,d) = \mathscr{P}(\vec{d}'|\overline{B}^0)\mathscr{P}(t,\overline{B}^0) + \mathscr{P}(\vec{d}'|B^0)\mathscr{P}(t,B^0), \tag{4.51}$$

where again the true tag, $d\in\{-1,+1\}$ represents a $\overline{B}^0$ and B^0 meson, respectively. Using the following definitions for the tag decisions

$$\begin{aligned} \Sigma_{\mathrm{tag}}(\vec{d}') &= \mathscr{P}\left(\vec{d}'|B^0\right) + \mathscr{P}\left(\vec{d}'|\overline{B}^0\right), \\ \Delta_{\mathrm{tag}}(\vec{d}') &= \mathscr{P}\left(\vec{d}'|B^0\right) - \mathscr{P}\left(\vec{d}'|\overline{B}^0\right), \end{aligned} \tag{4.52}$$

the PDF from Eq. (4.47) can now be rewritten considering also the impact of the imperfect flavour tagging as

$$\begin{aligned} \mathscr{P}(t,\vec{d}'|\vec{\eta}) = \mathscr{N}(t,d)e^{-t/\tau}\cdot\Big(&\left[\Sigma_{\mathrm{tag}}(\vec{d}'|\vec{\eta}) - A_{\mathrm{P}}\Delta_{\mathrm{tag}}(\vec{d}'|\vec{\eta})\right]H(t) \\ &+ \left[\Delta_{\mathrm{tag}}(\vec{d}'|\vec{\eta}) - A_{\mathrm{P}}\Sigma_{\mathrm{tag}}(\vec{d}'|\vec{\eta})\right]T(t)\Big). \end{aligned} \tag{4.53}$$

When assuming that the lifetime difference is negligible, $\Delta\Gamma = 0$, $H(t) = 1$ follows and thus the PDF simplifies to

$$\mathscr{P}_{CP}(t, \vec{d}'|\vec{\eta}) \propto e^{-t/\tau} \cdot \left(\left[\Sigma_{\text{tag}} - A_{\text{P}} \Delta_{\text{tag}} \right] + \left[\Delta_{\text{tag}} - A_{\text{P}} \Sigma_{\text{tag}} \right] \cdot (-S \sin(\Delta m t) + C \cos(\Delta m t)) \right), \tag{4.54}$$

where the functions Σ_{tag} and Δ_{tag} depend on the tag decisions and also include information on the true mistag, as it contains the calibration function from Eq. (4.29) via the formula for multiple tagging algorithms Eq. (4.50).

Decay-Time Resolution

The finite decay-time resolution results from the finite vertex resolution of the LHCb detector. This can affect the precision of the measurement of the *CP* observables, as it can dilute the amplitudes of the oscillation terms. But in the case of the B^0-meson system the oscillation frequency is low and thus the resolution has a minor influence on the measurement in contrast to the B_s^0-meson system, where the oscillation frequency is much higher and the resolution has greater impact.

The resolution function is modelled by three Gaussian functions which describe the difference between the measured and the actual decay time of the B^0-meson candidates, t' and t, respectively. The widths of two of these Gaussian functions depend linearly on the per-candidate decay-time-uncertainty estimates and thus vary for each individual candidate, while the width of the remaining Gaussian function is not dependent on σ_t. The means of the three Gaussian functions are shared. The third Gaussian function describes the decay-time resolution of the remaining candidates that are associated with an incorrect PV. Consequently, the resolution model is

$$\mathscr{R}(t' - t|\sigma_t) = \frac{1}{\sqrt{2\pi}} \left[\sum_{i=1}^{2} \frac{c_i}{a_i \cdot \sigma_t + b_i} \cdot e^{-\frac{\left(t'-t-\mu\right)^2}{2(a_i \cdot \sigma_t + b_i)^2}} + \frac{c_{\text{wPV}}}{\sigma_{\text{wPV}}} \cdot e^{-\frac{\left(t'-t-\mu\right)^2}{2\sigma_{\text{wPV}}^2}} \right], \tag{4.55}$$

where a_i ans b_i are the parameters of the linear calibration function of the decay-time-uncertainty estimates, $c_{\{i,\text{wPV}\}}$ are the fractions of the ith Gaussian component and of the component which considers the incorrectly assigned PVs. For these fractions it is valid, that $\sum_{i=1}^{3} c_i = 1$.

To proof the linear relation between the decay-time resolution and the decay-time-uncertainty estimates for the first two Gaussian components, a minimum χ^2-fit to simulated data is performed. For this the $(t' - t)$ distribution is split into 20 equally filled bins in σ_t. The 20 subsets are fitted simultaneously using a modified, discrete version of the introduced resolution model. The modified model consists of two Gaussian components with a width for each bin and a third Gaussian component, whose width is shared between all bins. The determined widths, $\sigma_{\text{narrow}}(\sigma_t)$ and $\sigma_{\text{wide}}(\sigma_t)$, are shown in Fig. 4.19 and Fig. 4.20 for the $B^0 \to \psi(2S) K_{\text{S}}^0$ and $B^0 \to J/\psi\, K_{\text{S}}^0$ mode, respectively. The linear dependence can be seen

in the plots and is thus assumed in the fit. The parameters of the resolution model are determined on selected, simulated data individually for both decay modes, but shared for the centre-of-mass energies and fixed in the fit to data. The results are given in Tab. 4.26 for both decay modes, and the corresponding fit projections are shown in Fig. 4.21 and in Fig. 4.22.

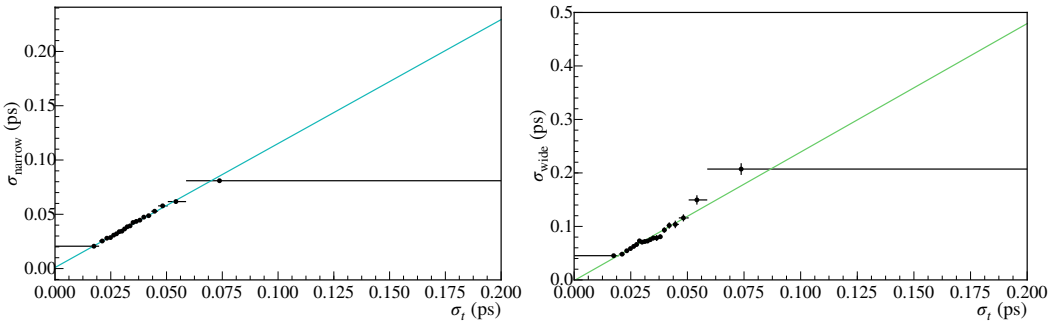

Figure 4.19: Linear calibration function of the decay-time-uncertainty estimates from simulated data. Horizontal errorbars represent the widths of the bins, while vertical errorbars represent the uncertainties of the fit for the (left) narrow and the (right) Gaussian component of the decay-time resolution model for the $B^0 \to \psi(2S)K^0_S$ mode.

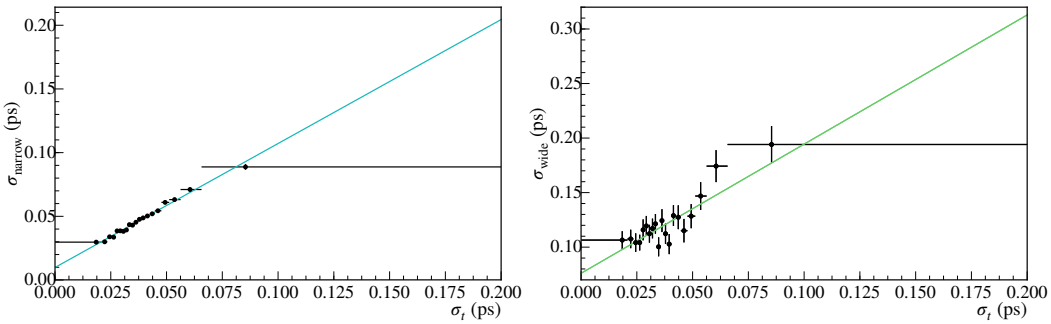

Figure 4.20: Linear calibration function of the decay-time-uncertainty estimates from simulated data. Horizontal errorbars represent the widths of the bins, while vertical errorbars represent the uncertainties of the fit for the (left) narrow and the (right) Gaussian component of the decay-time resolution model for the $B^0 \to J/\psi K^0_S$ mode.

The effective single Gaussian resolution is 48 fs in the case of the $B^0 \to \psi(2S)K^0_S$ and 67 fs in the case of the $B^0 \to J/\psi K^0_S$ mode, determined for correctly associated B^0 candidates.

Decay-Time Acceptance

The reconstruction method and the various selection steps can lead to decay-time dependent inefficiencies, the so-called decay-time *acceptance*. It can cause a deficit in the number of

Table 4.26: Results of the fit of the resolution parameters on simulated data for the $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi K^0_S$ mode. Every parameter is allowed to vary in the fit on data, except for the mean, μ, and b_{wide}, which are set to zero in the fit to data and in the fit to simulated data in the case of the $B^0 \to \psi(2S)K^0_S$. In the case of the $B^0 \to J/\psi K^0_S$ only the the mean, μ is set to zero in the fit to data and in the fit to simulated data.

parameters	$B^0 \to \psi(2S)K^0_S$	$B^0 \to J/\psi K^0_S$
μ [ps]	0.0	-2.90 ± 0.29
c_{wide}	0.058 ± 0.006	0.151 ± 0.009
c_{wPV}	0.0039 ± 0.0002	0.0033 ± 0.0007
a_{narrow}	1.162 ± 0.011	0.957 ± 0.022
b_{narrow} [ps]	0.0007 ± 0.0003	0.0090 ± 0.0008
a_{wide}	2.481 ± 0.079	1.363 ± 0.129
b_{wide} [ps]	0.0	0.071 ± 0.005
σb_{wPV} [ps]	10.044 ± 0.363	0.885 ± 0.095

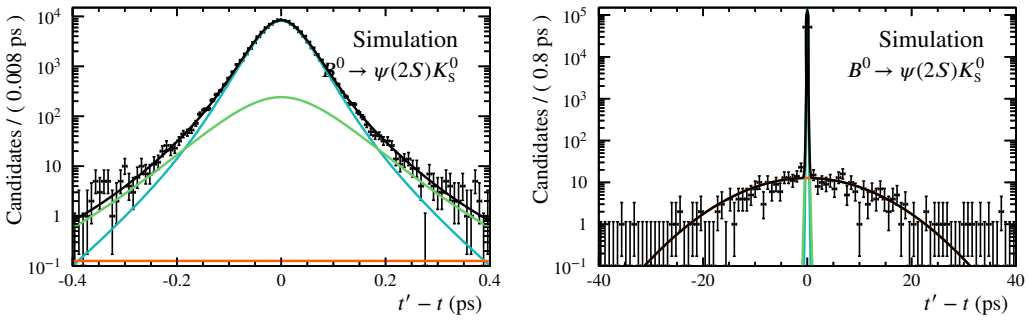

Figure 4.21: Projection of the decay-time resolution model on simulated data in case of the $B^0 \to \psi(2S)K^0_S$ mode in a (left) small and (right) wide range. The data is represented by the black points. The (black) complete model comprises a (blue) narrow and a (green) wide Gaussian component and a (orange) component that considers incorrectly associated PVs. While the width of the latter component is independent from the decay-time-uncertainty estimates, the widths of the first two components depends linearly on these estimates.

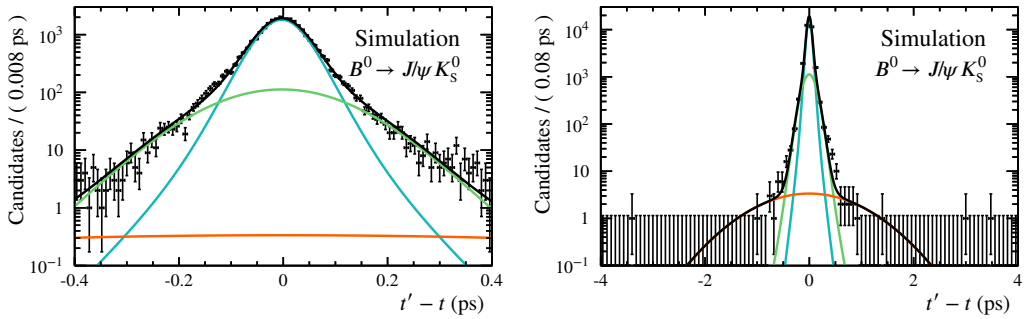

Figure 4.22: Projection of the decay-time resolution model on simulated data in case of the $B^0 \to J/\psi K^0_S$ mode in a (left) small and (right) wide range. The data is represented by the black points. The (black) complete model comprises a (blue) narrow and a (green) wide Gaussian component and a (orange) component that considers incorrectly associated PVs. While the width of the latter component is independent from the decay-time-uncertainty estimates, the widths of the first two components depends linearly on these estimates.

candidates at low decay times and an excess at higher decay times. The acceptance function is parametrized using cubic b-splines. The positions of the supporting points, *knots*, are determined on simulated data and then fixed in the fit to real data, whereas the parameters are left floating. The positions of the knots are chosen to be at 0.2, 0.5, 1.0, 5.0, and 15.0 ps in the case of the $B^0 \to \psi(2S)K^0_S$ mode and to be at 0.2, 1.0, 4.0, and 15.0 ps in the case of the $B^0 \to J/\psi K^0_S$ mode. The projections of the decay-time fit and the acceptance function to simulated data are shown in Fig. 4.23.

4.5.2 Strategy of the *CP*-Asymmetry Fit

The fit is performed simultaneously in subsets of the data split according to the decay channel and to the year of data taking, as the tagging-calibration parameters vary between the two channels, and the production asymmetries vary for both channels and the centre-of-mass energies. The used observables are the reconstructed decay time of the B^0 meson, t', in the range from 0.2 ps to 15 ps, the observed tags, $\vec{d}' = (d'_{\mathrm{OS}}, d'_{\mathrm{SS}}) \in \{-1, 0, 1\}$, the per-candidate decay-time-uncertainty estimates, σ_t ,in the range from 0 ps to 400 ps, and the mistag-probability estimates, $\vec{\eta} = (\eta_{\mathrm{OS}}, \eta_{\mathrm{SS}})$, in the range from 0 ps to 0.5 ps.

The production asymmetries for both channels and centre-of-mass energies are evaluated using the recent LHCb measurement [90] depending on the transverse momentum and the rapidity of the B^0 candidates. The values and uncertainties for the production asymmetries and the other external inputs in the fit, that describe the B^0 system, the lifetime of the B^0 meson, τ, and the mass difference, Δm, are given in Tab. 4.27.

To propagate the uncertainties of these external inputs correctly into the determination of the *CP* observables, the external inputs as well as the tagging-calibration parameters are

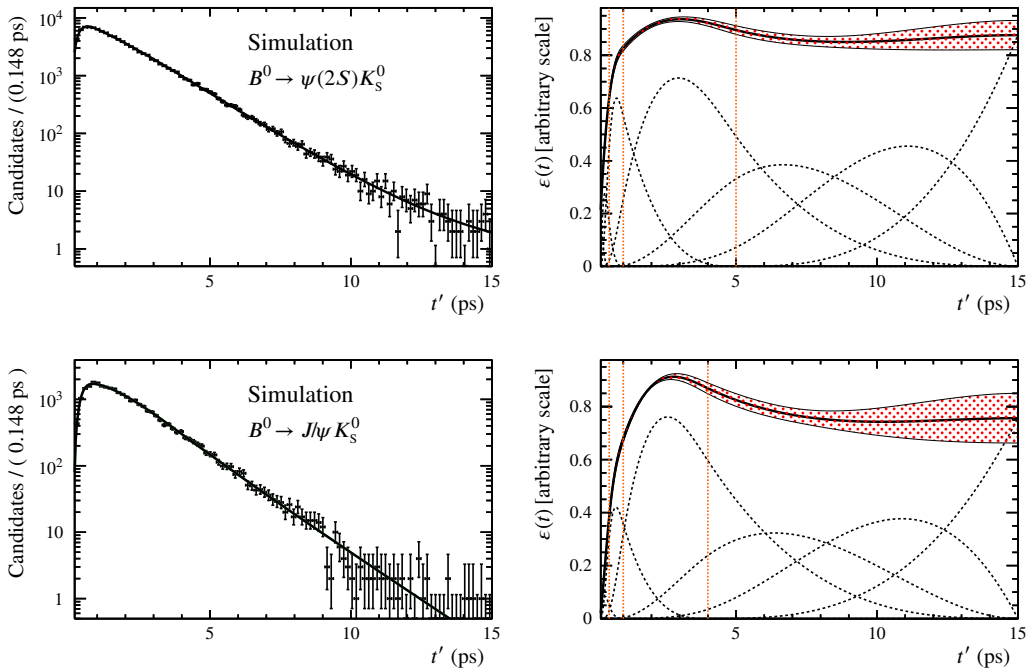

Figure 4.23: The projections of (left) the decay-time fit and (right) the acceptance function on simulated data in (top) the case of the $B^0 \to \psi(2S)K^0_S$ mode and in (bottom) the case of the $B^0 \to J/\psi\, K^0_S$ mode, where the lifetime of the B^0 mesons is fixed to the value from simulation($\tau = 1.520\,\text{ps}$). The (right) data points are shown as black points and the projection as a black line. The (left) projection of the acceptance function is also shown as a black line. The red-dotted area represents the uncertainty of the model coming from the fit. The black dashed lines are the single components of the b-splines and the vertical orange-dashed lines represent the position of the knots. [11]

Table 4.27: External inputs in the decay-time-dependent fit are the production asymmetries, A_P, the mixing frequency, Δm, and the lifetime of the B^0 meson, τ. The production asymmetries, A_P, are evaluated individually for both decay modes and for the different centre-of-mass energies of 7 and 8 TeV. If two uncertainties are given, the first is statistical and the second systematic. If one uncertainty is given, it includes statistical and systematic contributions.

Parameter	Value and uncertainty	Source
$A_{P,7\,TeV}(J/\psi)$	$-0.0100 \pm 0.0084 \pm 0.0005$	[90]
$A_{P,8\,TeV}(J/\psi)$	$-0.0077 \pm 0.0054 \pm 0.0004$	[90]
$A_{P,7\,TeV}(\psi(2S))$	$-0.0143 \pm 0.0077 \pm 0.0005$	[90]
$A_{P,8\,TeV}(\psi(2S))$	$-0.0138 \pm 0.0051 \pm 0.0003$	[90]
$\Delta m \left[\text{ps}^{-1}\right]$	$0.5065 \pm 0.0016 \pm 0.0011$	[39]
τ[ps]	1.520 ± 0.004	[39]

Gaussian constrained (see Sec. 4.2.1) using their statistical experimental uncertainties. The systematic uncertainties and the uncertainty coming from the assumption of $\Delta\Gamma = 0$ are considered in the systematic studies. When constraining the tagging-calibration parameters, also their correlations are taken into account. In addition to this, the parameters of the decay-time resolution are fixed in the fit.

Due to the constraints on the tagging-calibration parameters, it is possible that the mistag probabilities for candidates which had an assigned tag before are shifted to larger mistag values and thus these candidates end up untagged. This leads to a non-smooth likelihood and thus makes the fit instable. To avoid such an effect, it is necessary to set limits on the mistag probabilities of the tagger combinations. Therefore, if a candidate is above a certain threshold it gets a tag of zero and a mistag of 0.5 assigned. The thresholds are determined by varying the calibration parameters within one standard deviation in the *CP* fit, so that no candidate will result with a mistag greater than 0.5. The thresholds are

$$\begin{aligned}
\eta_{OS}\left(B^0 \to \psi(2S)K^0_S\right) &\leq 0.476,\\
\eta_{SS}\left(B^0 \to \psi(2S)K^0_S\right) &\leq 0.472,\\
\eta_{OS}\left(B^0 \to J/\psi\, K^0_S\right) &\leq 0.450,\\
\eta_{SS}\left(B^0 \to J/\psi\, K^0_S\right) &\leq 0.468.
\end{aligned}$$

In order to not bias the measurement by the analysts, a blinding transformation is applied on the *CP* observables to measure. The transformation is implemented in the `RooUnblindUniform` method from RooFit's `RooBlindTools` [89]. This transformation adds an offset to the fitted results, while the uncertainty of the extracted parameter stays the same and can be used to tune the analysis. The offset is determined from a uniform distribution in the range $\pm s$, using a random generator with a known seed. This seed is

calculated from the *blinding string*. Choosing the same scale and blinding string in different measurements of the same physics observables enables to compare the blinded results based on the uncertainties and the order of magnitude without unblinding them. In this analysis the range is chosen to be $[-2, 2]$. The used blinding strings are listed in Tab. 4.28.

Table 4.28: Blinding strings for `RooUnblindUniform` for the measured *CP* observables, S and C, for both decay modes.

Parameter	Blinding string
$S\left(J/\psi K^0_S\right)$	SJpsieePsi2SKSRunI
$S\left(\psi(2S)K^0_S\right)$	SJpsieePsi2SKSRunI
$C\left(J/\psi K^0_S\right)$	CJpsieePsi2SKSRunI
$C\left(\psi(2S)K^0_S\right)$	CJpsieePsi2SKSRunI

In the end the fit results are validated, by performing the fit on various subsets of the data and using different implementations of the framework. These studies show no bias from the nominal fit.

The fit results of the *CP* observables from the fit to the reconstructed decay time of the B^0 meson are

$$\begin{aligned} S\left(\psi(2S)K^0_S\right) &= 0.842 \pm 0.103, \\ C\left(\psi(2S)K^0_S\right) &= -0.044 \pm 0.096, \\ \rho\left(S\left(\psi(2S)K^0_S\right), C\left(\psi(2S)K^0_S\right)\right) &= 0.48, \\ S\left(J/\psi K^0_S\right) &= 0.824 \pm 0.076, \\ C\left(J/\psi K^0_S\right) &= 0.125 \pm 0.075, \\ \rho\left(S\left(J/\psi K^0_S\right), C\left(J/\psi K^0_S\right)\right) &= 0.46, \end{aligned}$$

where $\rho(S, C)$ is the linear statistical correlation coefficient between S and C. To determine the final results a correction is needed in order to take the K^0-$\overline{K}^0$ regeneration into account (see Sec. 4.6) and the systematic uncertainties need to be calculated in Sec. 4.7. The corresponding plots can be seen in Fig. 4.24.

4.6 Correction for Kaon Regeneration

The results from the fit to the reconstructed decay time of the B^0 meson need to be corrected for the *CP* violation occurring in K^0-$\overline{K}^0$ mixing and for the difference in the nuclear cross sections in material between K^0 and $\overline{K}^0$ interactions. The cross section depends on the momentum of the kaon and on the number of nuclei in the material. A pure K^0_L state can change into a superposition of K^0_S and K^0_L states when interacting with the detector

material [91]. This can be evaluated by using the formalism used in Refs. [92–94] to update the time-evolution of B mesons introduced in Eq. (2.24). To determine the corrections on the CP observables,

$$\begin{aligned} S &= S(\text{fit}) + \Delta S, \\ C &= C(\text{fit}) + \Delta C, \end{aligned} \tag{4.56}$$

for each candidate the path of the K^0_S meson through the LHCb detector is divided into small steps. At each step the amplitudes are updated, now including the description of the K^0_L and K^0_S regeneration, as well as CP violation in the kaon system. From this the corrections can be determined. The inaccurate knowledge of the material budget of the LHCb detector has the biggest impact on the uncertainty coming from this method. This uncertainty is estimated to be around 10 %. The corrections are determined to be

$$\begin{aligned} \Delta S(B^0 \to \psi(2S)K^0_S) &= (-0.18 \pm 0.02)\,\%, \\ \Delta C(B^0 \to \psi(2S)K^0_S) &= (0.36 \pm 0.04)\,\%, \\ \Delta S(B^0 \to J/\psi K^0_S) &= (-0.16 \pm 0.02)\,\%, \\ \Delta C(B^0 \to J/\psi K^0_S) &= (0.32 \pm 0.03)\,\%, \end{aligned} \tag{4.57}$$

as an average over all long track and downstream reconstructed candidates.

4.7 Systematic Uncertainties

Systematic uncertainties can originate from mismodelling the PDFs, from not correctly estimating the experimental uncertainties, or from uncertainties of the external input parameters. The effects are studied by using pseudoexperiments in which distributions are generated according to PDFs, where the parameters differ from the parameters used in the nominal fit model. These generated *pseudo*-datasets are then fitted again with the nominal model and the results can be compared to the generation values to test if biases arise for the parameters of interest. This is done multiple times in order to get a robust evaluation of the model.

The effect of setting the decay-time-width difference to zero in the fit, as well as effects arising from the constrained parameters from Tab. 4.27 are studied by varying their values within one standard deviation of their current experimental uncertainties. The impact of the systematic uncertainties of the flavour-tagging-calibration parameters are studied first by varying p_0 and Δp_0 up and p_1 and Δp_1 down by one systematic uncertainty and second the other way round as previous CP-violation [8] measurements of LHCb have shown, that the highest impact on the CP-violation parameters results from varying p_0 and p_1 in opposite directions. The impact of the decay-time bias from the resolution model is studied using pseudoexperiments, where a shift of 3 fs is assumed. In order to evaluate the systematic uncertainties arising from the linear calibration of the decay-time-uncertainty estimates, the

Table 4.29: Systematic uncertainties for the *CP*-violation observables S and C.

Source	$B^0 \to J/\psi K^0_S$ $\sigma(S)$	$\sigma(C)$	$B^0 \to \psi(2S) K^0_S$ $\sigma(S)$	$\sigma(C)$
$\Delta\Gamma$	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

scale parameters of the linear function are varied within $\pm 30\,\%$. The effect coming from the choice of the model describing the decay-time acceptance is evaluated by using a different model with more supporting points compared to the nominal fit model. In Tab. 4.29 the results are summarised separately for the parameters S and C and for the two decay channels. To obtain the overall systematic uncertainties, the individual ones are added in quadrature. The systematic uncertainties are all small compared to the statistical uncertainties, thus the measurement is still statistically limited and the systematic uncertainties are understood and well under control.

4.8 Results and Combination

The reconstructed decay-time distributions and the corresponding projections of the fit, as well as the projections of the decay-time acceptance from data are shown in Fig. 4.24 for both modes. While the theoretical *CP* asymmetry is defined by the decay-time dependent decay rates (see Eq. (2.51)), the observed asymmetry is given by the corresponding signal-yield asymmetries,

$$\mathcal{A}_{\text{yield}} = \frac{N_{\bar{B}^0}(t) - N_{B^0}(t)}{N_{\bar{B}^0}(t) + N_{B^0}(t)}, \tag{4.58}$$

where $N_{\bar{B}^0}$ is the number of candidates with a combined $\overline{B}^0$-tag decision assigned and N_{B^0} is the number of candidates with a B^0-tag decision assigned. The observed asymmetry contains by definition the imperfections of the experimental conditions, such as the imperfect flavour tagging. These effects result in a dilution of the theoretical asymmetry as can be seen in Eq. (4.54), thus the amplitude of the oscillation is damped and reaches not one.

The signal-yield asymmetries are shown in Fig. 4.25, where the solid black curves are the projections of the PDF for the combined flavour-tagging decision.

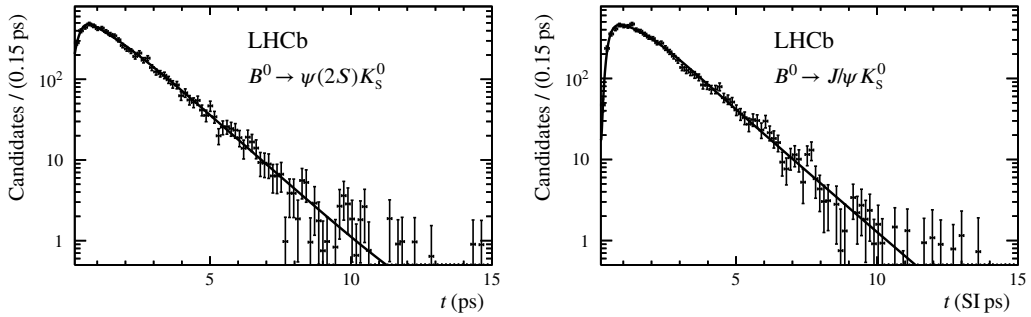

Figure 4.24: Projections of the decay-time fit to the weighted (left) $B^0 \to \psi(2S)K^0_S$ and (right) $B^0 \to J/\psi K^0_S$ candidates for run I of the LHCb detector.

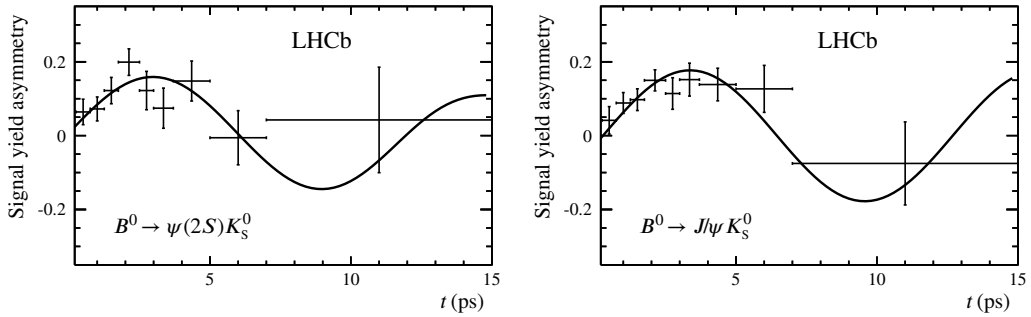

Figure 4.25: Signal-yield asymmetries, $(N_{\bar{B}^0}(t) - N_{B^0}(t))/(N_{\bar{B}^0}(t) + N_{B^0}(t))$, as a function of the reconstructed decay time for (left) $B^0 \to \psi(2S)K^0_S$ and (right) $B^0 \to J/\psi K^0_S$ decays.Here $N_j(t)$, with $j \in \{\bar{B}^0, B^0\}$, is the number of candidates that have a combined $\{\bar{B}^0, B^0\}$ tag decision assigned. The solid curves are the projections of the PDF for the combined flavor-tagging decision.

The values of the parameters of the decay-time-acceptance model are listed in Tab. 4.30 for both decay channels.

Table 4.30: Results of the decay-time-acceptance parameters in the nominal decay-time fit for both channels.

	$B^0 \to \psi(2S)K^0_S$	$B^0 \to J/\psi K^0_S$
h_0	0.21 ± 0.04	0.015 ± 0.010
h_1	0.334 ± 0.034	0.007 ± 0.017
h_2	0.70 ± 0.05	0.502 ± 0.032
h_3	1 (fixed)	1 (fixed)
h_4	0.81 ± 0.15	0.74 ± 0.12
h_5	0.53 ± 0.18	0.64 ± 0.18

The results for the *CP* observables are

$$\begin{aligned}
S\left(\psi(2\mathrm{S})K^0_S\right) &= 0.84 \pm 0.10\,(\mathrm{stat.}) \pm 0.01\,(\mathrm{syst.})\,,\\
C\left(\psi(2\mathrm{S})K^0_S\right) &= -0.05 \pm 0.10\,(\mathrm{stat.}) \pm 0.01\,(\mathrm{syst.})\,,\\
\rho\left(S\left(\psi(2\mathrm{S})K^0_S\right), C\left(\psi(2\mathrm{S})K^0_S\right)\right) &= 0.4814\,,\\
S\left(J/\psi K^0_S\right) &= 0.83 \pm 0.08\,(\mathrm{stat.}) \pm 0.01\,(\mathrm{syst.})\,,\\
C\left(J/\psi K^0_S\right) &= 0.12 \pm 0.07\,(\mathrm{stat.}) \pm 0.02\,(\mathrm{syst.})\,,\\
\rho\left(S\left(J/\psi K^0_S\right), C\left(J/\psi K^0_S\right)\right) &= 0.4565\,.
\end{aligned}$$

where the values have been corrected for the effect of Kaon regeneration as discussed in Sec. 4.6.

A combination of the two measurements introduced in this thesis with the previous from LHCb in the "golden" decay channel of $B^0 \to J/\psi K^0_S$ [8], where the J/ψ is reconstructed from two muons, and a combination of just the two J/ψ modes are determined using the *GammaCombo* framework. The combinations are performed using two- dimensional likelihood scans, which are shown in Fig. 4.26, taking into account the correlations between the single measurements. The quoted uncertainties include statistical and systematic contributions.

The combination of both J/ψ modes leads to

$$\begin{aligned}
S(B^0 \to J/\psi K^0_S) &= 0.75 \pm 0.04,\\
C(B^0 \to J/\psi K^0_S) &= -0.014 \pm 0.030,
\end{aligned}$$

with a correlation coefficient of 0.42. The combination of the results from all the $B^0 \to [c\bar{c}]K^0_S$ decay modes, *i.e.* $B^0 \to J/\psi K^0_S$, where the J/ψ is reconstructed either from two muons or

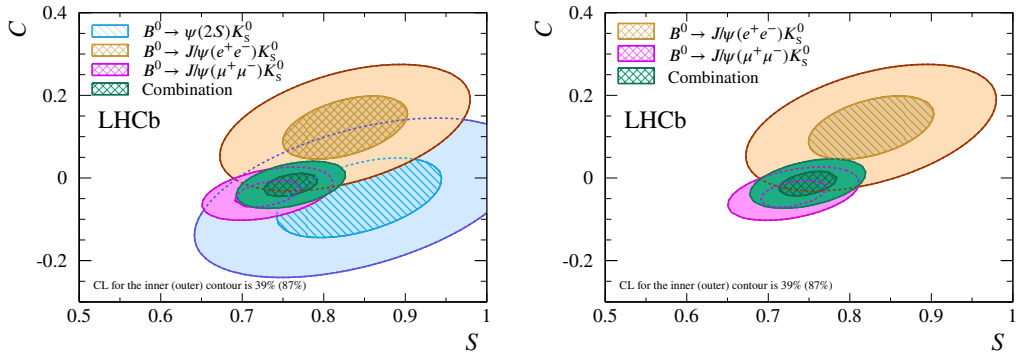

Figure 4.26: Two-dimensional likelihood scans to obtain the combination of the results of the *CP* parameters (top) in the case of the $B^0 \to J/\psi K^0_S$ modes, where the J/ψ is reconstructed either from muons or electrons, and (bottom) in the case of the $B^0 \to J/\psi K^0_S$ modes and the $B^0 \to \psi(2S)K^0_S$ modes. The confidence level for the inner and outer contour is 39 % and 87 %, respectively.

electrons, and $B^0 \to \psi(2S)K^0_S$, yields

$$S(B^0 \to [c\bar{c}]K^0_S) = \quad 0.760 \pm 0.034\,,$$
$$C(B^0 \to [c\bar{c}]K^0_S) = -0.017 \pm 0.029\,,$$

with a correlation coefficient of 0.42.

5 Conclusion and Outlook

During the first run of the Large Hadron Collider, LHC, from 2010 until 2012, the Higgs boson, the last missing proposed particle of the Standard Model of particle physics, SM, was discovered. This discovery strengthens the confidence in the theory. It is a very well-tested theory, but the SM seems to be incomplete, nonetheless, and a few open questions remain. Up to this point, gravity is the only one of the fundamental forces, which is not described by the SM. Direct searches could not find a candidate for dark matter, and the asymmetry between matter and anti-matter in the universe is still not completely understood. To be able to answer some of these questions, indirect searches for New Physics beyond the SM seem to be the key, as direct searches for new heavy particles are limited by the available energies of the particle colliders. Whereas, the higher-order contributions in indirect searches are sensitive to New Physics effects even at lower energies.

The Large Hadron Collider beauty, LHCb, experiment is designed for precision measurements in decays of B and D mesons, where the theoretical predictions of the observables to be measured are very precise. Hence, LHCb offers an excellent test of the SM and can hint to New Physics. An improved understanding of *CP* violation processes and the quark-mixing sector could benefit the understanding of the mechanism that results in the matter and anti-matter asymmetry in the universe. As *CP* violation arises in the SM from the Cabbibo-Kobayashi-Maskawa, CKM, matrix, probing its characteristics, *e.g.* the unitarity, presents an exemplary test of the SM. Predestined modes for measuring *CP* violation are $b \to [c\bar{c}]s$ decays, because higher-order contributions, that could introduce additional strong and weak phases in the considered decay amplitudes, are expected to be small.

This thesis presents the decay-time-dependent measurement of the *CP* parameter $\sin(2\beta)$ in the decays of $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi K^0_S$ using data collected by the LHCb experiment at centre-of-mass energies of 7 TeV and 8 TeV in pp collisions corresponding to a total integrated luminosity of $3\,\text{fb}^{-1}$. As *CP* violation in the mixing is negligible and *CP* violation in these decays is expected to be small, the *CP* coefficient S can be identified as $\sin(2\beta)$. Here, the $\psi(2S)$ meson is reconstructed from two muons, while the J/ψ meson is reconstructed from two electrons. In total 7970 ± 100 $B^0 \to \psi(2S)K^0_S$ decays and $10\,630 \pm 140$ $B^0 \to J/\psi K^0_S$ decays are used to determine the *CP* observables via an unbinned maximum-likelihood fit to the reconstructed decay-time distribution of the B^0

meson resulting in

$$\begin{aligned} S\left(B^0 \to \psi(2S)K^0_S\right) &= \quad 0.84 \pm 0.10 \pm 0.01\,, \\ C\left(B^0 \to \psi(2S)K^0_S\right) &= -0.05 \pm 0.10 \pm 0.01\,, \\ S\left(B^0 \to J/\psi\, K^0_S\right) &= \quad 0.83 \pm 0.08 \pm 0.01\,, \\ C\left(B^0 \to J/\psi\, K^0_S\right) &= \quad 0.12 \pm 0.07 \pm 0.02\,, \end{aligned}$$

where the first uncertainties are statistical and the second are systematic. The correlation coefficients between S and C are 0.48 and 0.46 in the case of the $\psi(2S)$ and in the case of the J/ψ mode, respectively. This analysis represents the first decay-time-dependent CP violation measurement in the decay of $B^0 \to \psi(2S)K^0_S$ and in a decay that uses electrons in the final state at a hadron collider. Considering the single measurements, the results of the electron and muon mode are compatible with each other and also with the previous measurement by LHCb using $B^0 \to J/\psi\, K^0_S$ decays of $S = 0.73 \pm 0.04$ and $C = -0.038 \pm 0.032$ [8], where the J/ψ is reconstructed from two muons using also $3\,\text{fb}^{-1}$.

Taking into account only the specific decay channels of $B^0 \to \psi(2S)K^0_S$ and $B^0 \to J/\psi\, K^0_S$, the averages [39] by the Heavy Flavor Averaging Group are also in agreement to the presented results in this thesis.

The B factories, the Belle and BaBar collaborations, have measured $\sin(2\beta)$ in the $B^0 \to J/\psi\, K^0_S$ and $B^0 \to \psi(2S)K^0_S$ mode to a very high precision. The results exclusively obtained in $B^0 \to \psi(2S)K^0_S$ decays are [41, 42]

$$\begin{aligned} S_{\text{Belle}}\left(B^0 \to \psi(2S)K^0_S\right) &= \quad 0.738 \pm 0.079 \pm 0.036\,, \\ C_{\text{Belle}}\left(B^0 \to \psi(2S)K^0_S\right) &= -0.104 \pm 0.055^{+0.027}_{-0.047}\,, \\ S_{\text{BaBar}}\left(B^0 \to \psi(2S)K^0_S\right) &= \quad 0.897 \pm 0.100 \pm 0.036\,, \\ C_{\text{BaBar}}\left(B^0 \to \psi(2S)K^0_S\right) &= \quad 0.089 \pm 0.076 \pm 0.0202\,, \end{aligned}$$

where the first uncertainties are statistical and the second are systematic. These values are in good agreement with the presented results in this thesis and the uncertainties are of the same order of magnitude.

Besides the single measurements, a combination of the at LHCb considered $B^0 \to [c\bar{c}]K^0_S$ decays is determined to be

$$\begin{aligned} S\left(B^0 \to [c\bar{c}]K^0_S\right) &= \quad 0.760 \pm 0.034\,, \\ C\left(B^0 \to [c\bar{c}]K^0_S\right) &= -0.017 \pm 0.029\,, \end{aligned}$$

where the uncertainties include statistical and systematic contributions. The average is determined under the assumption that higher-order contributions to the decay amplitudes are negligible. The presented measurement improves the precision of $\sin(2\beta)$ at LHCb by

20 %. Above that, it also improves the world averages [39] for S and C obtained in various $b \to [c\bar{c}]s$ transitions,

$$S\left(b \to [c\bar{c}]s\right) = \quad 0.699 \pm 0.017,$$
$$C\left(b \to [c\bar{c}]s\right) = -0.005 \pm 0.015,$$

including the results presented in this thesis, where the uncertainties include statistical and systematic contributions.

Further, this thesis presents the first preliminary and unofficial measurement for the ratio of branching fractions of the decays $B_s^0 \to \psi(2S)K_S^0$ and $B^0 \to \psi(2S)K_S^0$, $\mathscr{B}\left(B_s^0\right)/\mathscr{B}\left(B^0\right)$. It is determined to be

$$\frac{\mathscr{B}\left(B_s^0\right)}{\mathscr{B}\left(B^0\right)} = (3.8 \pm 0.9)\,\%,$$

where the uncertainty is only statistical. The result is comparable to the ratio of branching fractions obtained from the decays $B_s^0 \to J/\psi K_S^0$ and $B^0 \to J/\psi K_S^0$, which is determined to be $(4.31 \pm 0.17 \pm 0.12)\%$ [27].

During the run II of the LHC from 2015 until November 2018, the LHCb experiment collected around $6\,\text{fb}^{-1}$ of integrated luminosity at a centre-of-mass energy of 13 TeV. The estimated sensitivity on $\sin(2\beta)$ in $B^0 \to J/\psi K_S^0$ decays is 0.018 [95], which is of the same order of magnitude as the sensitivity on the current world average in all $b \to [c\bar{c}]s$ decays combined. Using all data from run I and run II will result in the world's most precise measurement by a single experiment of the CP observable, $\sin(2\beta)$.

In the long shutdown of the LHC, which started at the end of 2018, the LHCb detector receives a significant upgrade of the detector components and the computing farm, before restarting in 2020 and running until 2022 for run III of the LHC. The detector needs to be able to cope with a higher instantaneous luminosity of $L = 2 \cdot 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$ and a higher production rate of $b\bar{b}$-pairs of 10^6 per second. Due to these running conditions, almost all sub-detectors will be improved. During the run III the estimated sensitivity on $\sin(2\beta)$ in the decay channel $B^0 \to J/\psi K_S^0$ is 0.006 [95] using the expected $50\,\text{fb}^{-1}$.

The main competitor of the LHCb experiment will be the Belle II experiment, which is scheduled to start in 2019 and is expected to collect around $50\,\text{ab}^{-1}$ during its lifetime. Thus, it is expected to significantly decrease the uncertainty on $\sin(2\beta)$ to 0.010 [96]. Despite the competition, both physics programs complement each other perfectly, as a control instance as well as in the realm of $B_s^0 \to [c\bar{c}]K_S^0$ decays, where only the LHCb collaboration will be able to measure higher-order contributions from loop processes, which will have an increased impact when more data will be available.

Bibliography

[1] 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..

[2] D. Griffiths, *Introduction to Elementary Particles*, Wiley-VCH, 2011, isbn: 3527406018.

[3] T. D. Lee and C. N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1 1956) 254–258, doi: 10.1103/PhysRev.104.254.

[4] C. Wu *et al.*, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413–1415, doi: 10.1103/PhysRev.105.1413.

[5] J. Christenson *et al.*, *Evidence for the 2π Decay of the K20 Meson*, Phys. Rev. Lett. **13** (1964) 138–140, doi: 10.1103/PhysRevLett.13.138.

[6] B. Aubert *et al.*, *Observation of CP Violation in the B^0 Meson System*, Phys. Rev. Lett. **87** (9 2001) 091801, doi: 10.1103/PhysRevLett.87.091801.

[7] K. Abe *et al.*, *Observation of Large CP Violation in the Neutral B Meson System*, Phys. Rev. Lett. **87** (9 2001) 091802, doi: 10.1103/PhysRevLett.87.091802.

[8] R. Aaij *et al.*, *Measurement of CP violation in $B^0 \to J/\psi K^0_S$ decays*, Phys. Rev. Lett. **115** (2015) 031601, doi: 10.1103/PhysRevLett.115.031601, arXiv:1503.07089 [hep-ex].

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

[10] 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].

[11] R. Niet, *Measurement of CP Violation in $B^0 \to [c\bar{c}]K^0_S$ Transitions at LHCb*, Ph. D. thesis, 2018, doi: 10.17877/DE290R-19128.

[12] A. Seuthe and J. Albrecht, *Flavour-tagging and particle-identification calibration in key measurements with electrons at LHCb*, Presented 29 Sep 2017, 2017, https://cds.cern.ch/record/2293431.

[13] R. Aaij *et al.*, *Measurement of the ratio of branching fractions of the decays Λb0 ψ(2S)Λ and Λb0 J/ψΛ*, Journal of High Energy Physics **2019** (2019), issn: 1029-8479, doi: 10.1007/JHEP03(2019)126.

[14] K. Mielke, *Experimentelle Besonderheiten einer getaggten zeitabhängigen CP-Verletzungsmessung in $B^0 \to J/\psi(e^+e^-)K^0_S$ am LHCb-Experiment*, unpublished, B. Sc. thesis, TU Dortmund University, 2015.

[15] L. Nickel, *Studien zu einer Gütfunktion für CP-Verletzungsmessungen bei LHCb*, unpublished, B. Sc. thesis , TU Dortmund University, 2017.

[16] C. Hasenberg, *Sensivity study of the CPV-Paramter sin(2β) in the decay channel $B^0 \to J/\psi K^0_S$*, unpublished, M. Sc. thesis, TU Dortmund University, 2017.

[17] D. Perkins and P. H, *Introduction to High Energy Physics*, Cambridge University Press, 2000, isbn: 9780521621960.

[18] G. C. Branco, L. Lavoura, and J. P. Silva, *CP Violation*, Int. Ser. Monogr. Phys. **103** (1999) 1–536.

[19] R. Aaij *et al.*, *Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, Phys. Rev. Lett. **118**.2 (2017) 022003, doi: 10.1103/PhysRevLett.118.022003, arXiv:1606.07895 [hep-ex].

[20] R. Aaij *et al.*, *Amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, Phys. Rev. **D95**.1 (2017) 012002, doi: 10.1103/PhysRevD.95.012002, arXiv:1606.07898 [hep-ex].

[21] R. Aaij *et al.*, *Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda^0_b \to J/\psi K^- p$ Decays*, Phys. Rev. Lett. **115** (2015) 072001, doi: 10.1103/PhysRevLett.115.072001, arXiv:1507.03414 [hep-ex].

[22] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321–323, doi: 10.1103/PhysRevLett.13.321.

[23] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508–509, doi: 10.1103/PhysRevLett.13.508.

[24] G. S. Guralnik *et al.*, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (20 1964) 585–587, doi: 10.1103/PhysRevLett.13.585.

[25] M. Aaboud *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (20 2012) 1–29, doi: 10.1016/j.physletb.2012.08.020, arXiv:1207.7214 [hep-ex].

[26] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*, Includes exercises, Boulder, CO: Westview, 1995, https://cds.cern.ch/record/257493.

[27] C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40**.10 (2016) 100001, doi: 10.1088/1674-1137/40/10/100001.

[28] D. Boutigny *et al.*, *The BABAR physics book: Physics at an asymmetric B factory*, 1998, http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=SLAC-R-504.

[29] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945, doi: 10.1103/PhysRevLett.51.1945.

[30] J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41**.1 (2005) 1–131, doi: 10.1140/epjc/s2005-02169-1, arXiv:hep-ph/0406184 [hep-ph].

[31] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal* CP *Nonconservation*, Phys. Rev. Lett. **55** (10 1985) 1039–1042, doi: 10.1103/PhysRevLett.55.1039, https://link.aps.org/doi/10.1103/PhysRevLett.55.1039.

[32] J. Charles *et al.*, *CP violation and the CKM matrix: assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005), updated results and plots available at http://ckmfitter.in2p3.fr 1–131, doi: 10.1140/epjc/s2005-02169-1.

[33] V. Weisskopf and E. Wigner, *Berechnung der natürlichen Linienbreite auf Grund der Diracschen Lichttheorie*, Zeitschrift für Physik **63**.1 (1930) 54–73, doi: 10.1007/BF01336768.

[34] V. Weisskopf and E. Wigner, *Über die natürliche Linienbreite in der Strahlung des harmonischen Oszillators*, Zeitschrift für Physik **65**.1 (1930) 18–29, doi: 10.1007/BF01397406.

[35] T. Inami and C. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes* $K_L^0 \to \mu\bar{\mu}$, $K^+ \to \pi^+\nu\bar{\nu}$, $K^0 \leftrightarrow \overline{K}^0$, Progress of Theoretical Physics **65** (1981) 297–314, doi: 10.1143/PTP.65.297.

[36] M. Jung, *Determining weak phases from* $B \to J/\psi P$ *decays*, Phys. Rev. **D86** (2012) 053008, doi: 10.1103/PhysRevD.86.053008, arXiv:1206.2050 [hep-ph].

[37] S. Faller *et al.*, *The Golden Modes* $B^0 \to J/\psi K_{L,S}$ *in the Era of Precision Flavour Physics*, Phys. Rev. **D79** (2009) 014030, doi: 10.1103/PhysRevD.79.014030, arXiv:0809.0842 [hep-ph].

[38] P. Frings, U. Nierste, and M. Wiebusch, *Penguin contributions to CP phases in* $B_{d,s}$ *decays to charmonium*, Phys. Rev. Lett. **115** (2015) 061802, doi: 10.1103/PhysRevLett.115.061802, arXiv:1503.00859 [hep-ph].

[39] Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and* τ*-lepton properties as of summer 2016* (2016), arXiv:1612.07233 [hep-ex].

[40] R. Fleischer, *Extracting* γ *from* $B(s/d) \to J/\psi K_S$ *and* $B(d/s) \to D^+(d/s)D^-(d/s)$, Eur. Phys. J. **C10** (1999) 299–306, doi: 10.1007/s100529900099, arXiv:hep-ph/9903455 [hep-ph].

[41] I. Adachi *et al.*, *Precise measurement of the CP violation parameter* $\sin 2\phi_1$ *in* $B^0 \to (c\bar{c})K^0$ *decays*, Phys.Rev.Lett. **108** (2012) 171802, doi: 10.1103/PhysRevLett.108.171802, arXiv:1201.4643 [hep-ex].

[42] B. Aubert *et al.*, *Measurement of Time-Dependent CP Asymmetry in $B^0 \to c\bar{c}K^{(*)0}$ Decays*, Phys.Rev. **D79** (2009) 072009, doi: 10.1103/PhysRevD.79.072009, arXiv:0902.1708 [hep-ex].

[43] M. Bona *et al.*, *The unitarity triangle fit in the standard model and hadronic parameters from lattice QCD: a reappraisal after thm Measurements of Δm_s and BR($B \to \tau\nu$)*, JHEP **10** (2006), updated results and plots available at http://www.utfit.org/UTfit/, doi: 10.1088/1126-6708/2006/10/081, arXiv:0606167 [hep-ph].

[44] M. Lamont, *Status of the LHC*, Journal of Physics: Conference Series **455**.1 (2013) 012001, http://stacks.iop.org/1742-6596/455/i=1/a=012001.

[45] G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008), doi: 10.1088/1748-0221/3/08/S08003..

[46] S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008), doi: 10.1088/1748-0221/3/08/S08004.

[47] K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3**.08 (2008) S08002, doi: 10.1088/1748-0221/3/08/S08002, http://stacks.iop.org/1748-0221/3/i=08/a=S08002.

[48] A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005, doi: 10.1088/1748-0221/3/08/S08005.

[49] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001, doi: 10.1088/1748-0221/3/08/S08001.

[50] R. Aaij *et al.*, *$b\bar{b}$ production angle plots* (25, 2018), http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[51] R. Aaij *et al.*, *Measurement of the b-Quark Production Cross Section in 7 and 13 TeV pp Collisions*, Phys. Rev. Lett. **118** (5 2017) 052002, doi: 10.1103/PhysRevLett.118.052002.

[52] R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $s = 7$ TeV in the forward region*, Physics Letters B **694** (2010) 209–216, doi: https://doi.org/10.1016/j.physletb.2010.10.010.

[53] F. Alessio *et al.*, *The LHCb Online Framework for Experiment Protection, and Global Operational Control and Monitoring*, Journal of Physics: Conference Series **331** (2011) 022002.

[54] C. Ilgner *et al.*, *The Beam Conditions Monitor of the LHCb Experiment* (2010), arXiv:1001.2487 [physics.ins-det].

[55] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, doi: 10.1088/1748-0221/9/01/P01002, arXiv:1311.3893 [physics.ins-det].

[56] 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].

[57] M. Clemencic *et al.*, *Recent developments in the LHCb software framework gaudi*, Journal of Physics: Conference Series **219** (2010) 042006.

[58] M. Clemencic *et al.*, *The LHCb Simulation Application, Gauss: Design, Evolution and Experience*, Journal of Physics: Conference Series **331**.3 (2011) 032023, http://stacks.iop.org/1742-6596/331/i=3/a=032023.

[59] T. Sjöstrand and others", *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, doi: 10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175 [hep-ph].

[60] 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].

[61] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155, doi: 10.1016/S0168-9002(01)00089-4.

[62] 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].

[63] S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250, doi: 10.1016/S0168-9002(03)01368-8.

[64] J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans.Nucl.Sci. **53** (2006) 270, doi: 10.1109/TNS.2006.869826.

[65] R. Aaij *et al.*, *The Boole project* (2018), ed. by M.Cattaneo, http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[66] R. Aaij *et al.*, *The Moore project* (2018), http://lhcbdoc.web.cern.ch/lhcbdoc/moore/.

[67] R. Aaij *et al.*, *The Brunel project* (2018), http://lhcbdoc.web.cern.ch/lhcbdoc/brunel/.

[68] I. Antcheva *et al.*, *ROOT - A C++ framework for petabyte data storage, statistical analysis and visualization*, Computer Physics Communications **182**.6 (2011) 1384–1385, issn: 0010-4655, doi: https://doi.org/10.1016/j.cpc.2011.02.008.

[69] R. Aaij *et al.*, *The DaVinci project* (2018), ed. by E. Rodrigues, http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/.

[70] R. Aaij *et al.*, *The LHCb data flow* (2018), https://lhcb.github.io/starterkit-lessons/first-analysis-steps/dataflow.html.

[71] 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].

[72] 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].

[73] K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, doi: 10.1088/1674-1137/38/9/090001.

[74] A. Hocker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039 [PHYSICS].

[75] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning and an application to boosting*, Jour. Comp. and Syst. Sc. **55** (1997) 119, doi: 10.1006/jcss.1997.1504.

[76] F. Pedregosa *et al.*, *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research **12** (2011) 2825–2830.

[77] J. Zhu *et al.*, *Multi-class AdaBoost*, Technical Report 430, Departement of Statistics, University of Michigan (2006).

[78] L. Breiman *et al.*, *Classification and Regression Trees*, The Wadsworth and Brooks-Cole statistics-probability series, Taylor & Francis, 1984, isbn: 9780412048418, https://books.google.de/books?id=JwQx-WOmSyQC.

[79] Y. Xie, *Principles to optimize event selections for measurements of CP asymmetries*, LHCb-INT-2009-012, 2009, https://cds.cern.ch/record/%201187570.

[80] V. Blobel and E. Lohrmann, *Statistische und numerische Methoden der Datenanalyse*, Teubner Studienbücher Physik, Stuttgart: Teubner, 1998, https://cds.cern.ch/record/437773.

[81] D. M. Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl.Instrum.Meth. **A764** (2014) 150, doi: 10.1016/j.nima.2014.06.081, arXiv:1312.5000 [hep-ex].

[82] G. Aad *et al.*, *Measurement of the branching ratio* $\Gamma(\Lambda_b^0 \to \psi(2S)\Lambda^0)/\Gamma(\Lambda_b^0 \to J/\psi\Lambda^0)$ *with the ATLAS detector*, Phys. Lett. **B751** (2015) 63–80, doi: 10.1016/j.physletb.2015.10.009, arXiv:1507.08202 [hep-ex].

[83] P. Koppenburg, *Statistical biases in measurements with multiple candidates* (2017), arXiv:1703.01128 [hep-ex].

[84] R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetries in* $B_s^0 \to J/\psi K_S^0$, JHEP **06** (2015) 131, doi: 10.1007/JHEP06(2015)131, arXiv:1503.07055 [hep-ex].

[85] B. Storaci, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions* (2013), LHCb-CONF-2013-011, https://cds.cern.ch/record/1559262?ln=de.

[86] 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].

[87] J. Wimberley, *Calibrating flavor tagging algorithms with binomial regression*, 2016, in preparation.

[88] 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.

[89] W. Verkerke and D. Kirkby, *RooFit*, http://roofit.sourceforge.net.

[90] 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), in preparation.

[91] A. Pais and O. Piccioni, *Note on the decay and absorption of the θ^0*, Phys.Rev. **100** (1955) 1487–1489, doi: 10.1103/PhysRev.100.1487.

[92] R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, JHEP **07** (2014) 041, doi: 10.1007/JHEP07(2014)041, arXiv:1405.2797 [hep-ex].

[93] 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].

[94] W. Fetscher *et al.*, *Regeneration of arbitrary coherent neutral kaon states: A new method for measuring the $K^0 - \overline{K}^0$ forward scattering amplitude*, Z.Phys. **C72** (1996) 543–547, doi: 10.1007/s002880050277.

[95] A. Mödden, *Sensitivitätsstudien für die Messung von sin(2β) im Zerfall $B^0 \to J/\psi K_S^0$ am LHCb-Experiment*, unpublished, B. Sc. thesis, TU Dortmund University, 2015.

[96] T. Aushev *et al.*, *Physics at Super B Factory* (2010), arXiv:1002.5012 [hep-ex].

Danksagung – Acknowledgement

Ich möchte mich gerne bei einigen Menschen bedanken, ohne die es nicht möglich gewesen wäre diese Arbeit zu schreiben, dabei so viel zu lernen und ohne die es auch nur halb so viel Spaß gemacht hätte.

Zu allererst bedanke ich mich herzlich bei Prof. Dr. Bernhard Spaan, der mich seit der Bachelorarbeit begleitet und mir es ermöglicht hat in einer tollen Atmosphäre zu lernen und zu arbeiten. Außerdem gab er mir Möglichkeit in einem internationalen Umfeld am CERN zu arbeiten und somit konnte ich mir einen *kleinen* Traum erfüllen.

Außerdem möchte ich Prof. Dr. Kevin Kröninger danken, dass er die Zeit gefunden hat als Zweitgutachter meiner Arbeit zu fungieren. Weiterhin möchte ich Prof. Dr. Mirko Cinchetti und Dr. Michael Paulus danken, die die Prüfungskomission vervollständigen.

Zudem danke ich dem Bundesministerium für Bildung und Forschung sowie der Deutschen Forschungsgemeinschaft für die Unterstützung während der Promotion.

Ich möchte mich ganz herzlich bei Ramon Niet bedanken, der sich mit mir durch den ganzen Reviewprozess gekämpft hat, um am Ende diese Analyse gemeinsam veröffentlichen zu können. Außerdem warst du der beste FFW-Trainingspartner und Bürokollege. Vielen Dank, lieber Ramon.

Außerdem möchte ich mich bei Julian Wishahi bedanken. Ohne ihn wäre der Lehrstuhl nicht derselbe gewesen. Du hast mir nicht nur fachlich viel beigebracht, sondern hast mir geholfen mich weiter zu entwicklen und ich danke dir sehr dafür, dass du mir immer noch mit Rat und Tat zur Seite stehst. Des Weiteren wäre ich wahrscheinlich ohne Florian Kruse, Christophe Cauet und Tobias Brambach nicht bei E5 gelandet und ich möchte euch danken, dass ihr mich während meiner Bachelor- und Masterarbeit betreut habt und nie genervt wart, wenn ich zum tausendsten Mal mit ausgedruckten Plots zu euch gekommen bin und euch um Rat gebeten habe.

Lieber Timon, lieber Moritz, ihr vervollständigt, die beste Bürokonstellation und ich bin sehr froh, dass ich in euch so gute Freunde gefunden habe. Ich möchte mich noch mal speziell bei Timon bedanken, der mit mir seit dem ersten Semester Übungsgruppen und Chemiepratika überstanden hat. Außerdem möchte ich mich auch bei Mathis bedanken, der mit mir von Anfang an das Physikpraktikum gemeistert hat und mir in allen Phasen der Promotion beigestanden hat.

Weiterhin danke ich Alex B. , Frank, Ulrich, Kevin und Patrick für ihre Unterstützung und die Diskussionen zu Analysethemen wie Flavour Tagging und physikalischen Untergründen, von denen diese Arbeit sehr profitiert hat.

I also would like to thank the LHCb B2CC and Flavour Tagging working groups at CERN for their input and the fruitful discussion during the meetings. I would like to thank Jeroen for

assisting with the analysis and additionally the analysis reviewers Sevda, Liming, Yasmine, Lars, Michael and all other people at CERN that helped publishing this analysis. I would like to express my gratitude to Markus Frank, who accepted my CERN summer student application and gave me the opportunity to do research at CERN for one summer.

Außerdem möchte ich dir, Britta, danken, weil du mir immer in bürokratischen Sachen den Rücken freigehalten hast. Zudem möchte ich mich bei dem gesamten Lehrstuhl E5 für die entspannte und freundschaftliche Arbeitsatmosphäre bedanken. Mein großer Dank gilt Margarete und Janine für die Mädelsabende und den sportlichen Ausgleich neben der Uni, sowie Robert für seine guten Ratschläge und die Fotografie-Workshops. Außerdem danke ich Matthias unter Anderem für seine Unterstützung bei der Instandhaltung des LHCb-Tanks, bei der ich sehr viel gelernt habe.

Außerdem möchte ich ganz herzlich meiner Familie danken, die immer an mich geglaubt und mich bei allem unterstützt hat. Außerdem danke ich meinen Mädels, Anna, Jeannine, Katha und Laura. Ihr seid einfach die Besten. Zu guter Letzt, danke ich Dennis für seine Unterstützung und seine Nachsicht, wenn ich doch mal wieder abends oder am Wochenende arbeiten musste.