technische universität
dortmund

Dissertation zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

im Fach Physik

Measurements of the CKM parameter $\sin(2\beta)$ in $B^0 \to \psi K^0_S$ decays with the LHCb experiment

---

vorgelegt von

Vukan Jevtić

aus Dortmund

Fakultät Physik

Technische Universität Dortmund

Dortmund, Juli 2024

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

Betreuer:
Prof. Dr. Bernhard Spaan†
Prof. Dr. Johannes Albrecht

Gutachter:
Prof. Dr. Johannes Albrecht
Prof. Dr. Kevin Kröninger

Vorsitzender der Prüfungskommission:
Prof. Dr. Emmanuel Stamou

Vertreter der wissenschaftlichen Mitarbeiter*innen:
Jun.-Prof. Dr. Benedikt Fauseweh

Datum des Einreichens der Dissertation:
17. Juli 2024

Datum der mündlichen Prüfung:
7. Oktober 2024

Abstract

This thesis presents *CP*-violation measurements conducted with data recorded by the LHCb detector during the Run 2 data-taking phase of the LHC. The size of the data set corresponds to an integrated luminosity of $6\,\text{fb}^{-1}$. A time-dependent measurement of the *CP* asymmetry arising in decays of neutral B mesons to *CP*-invariant final states is performed in three decay channels: $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$, $B^0 \to J/\psi(\to e^+e^-)K^0_S$, and $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$. The K^0_S mesons are reconstructed in decays to two oppositely charged pions. A combined fit to the Run 2 data set of LHCb is performed in the three analyzed channels and gives the result

$$S_{\psi K^0_S} = 0.717 \pm 0.013\,(\text{stat}) \pm 0.008\,(\text{syst}),$$
$$C_{\psi K^0_S} = 0.008 \pm 0.012\,(\text{stat}) \pm 0.003\,(\text{syst}).$$

This result is more precise than the previous world averages of these parameters, making this measurement the most precise single measurement of the quantity $S_{\psi K^0_S}$ to date. The parameter $S_{\psi K^0_S}$ is related to the CKM angle β, which parametrizes *CP* violation in the Standard Model of particle physics. The result is consistent with previous measurements from the LHCb collaboration and the B-factories Belle and BaBar.

Kurzfassung

In dieser Arbeit werden Messungen von *CP*-Verletzung vorgestellt. Die Messungen wurden mit Daten durchgeführt, die vom LHCb-Detektor während der Run 2-Datennahme des LHC aufgezeichnet wurden. Die Größe des Datensatzes entspricht einer integrierten Luminosität von $6\,\text{fb}^{-1}$. Eine zeitabhängige Messung der *CP*-Asymmetrie, die bei Zerfällen von neutralen B-Mesonen in *CP*-invariante Eigenzustände entsteht, wird in drei Zerfallskanälen durchgeführt: $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$, $B^0 \to J/\psi(\to e^+e^-)K^0_S$, und $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$. Die K^0_S-Mesonen werden in Zerfällen zu zwei entgegengesetzt geladenen Pionen rekonstruiert. Eine kombinierte Messung des LHCb Run 2-Datensatzes wird in den drei gemessenen Endzuständen durchgeführt und liefert das Ergebnis

$$S_{\psi K^0_S} = 0.717 \pm 0.013\,(\text{stat}) \pm 0.008\,(\text{syst}),$$
$$C_{\psi K^0_S} = 0.008 \pm 0.012\,(\text{stat}) \pm 0.003\,(\text{syst}).$$

Dieses Ergebnis ist präziser als die bisherigen Weltmittelwerte dieser Parameter. Dadurch ist diese Messung zugleich die bisher genaueste Einzelmessung der Größe $S_{\psi K^0_S}$. Der Parameter $S_{\psi K^0_S}$ hängt mit dem CKM-Winkel β zusammen, der die *CP*-Verletzung im Standardmodell der Teilchenphysik parametrisiert. Das Ergebnis ist konsistent mit früheren Messungen der LHCb-Kollaboration und der B-Fabriken Belle und BaBar.

Contents

1 Introduction

The primary aim of physics is to advance our knowledge of the natural world by developing explanations of natural phenomena that withstand empirical testing and are predictive and generalizable. In elementary particle physics, processes are studied at the smallest distances at the subatomic level, at which the fundamentally quantized natures of matter and forces dominate. The fields of theoretical and experimental physics are inseparable and part of the same scientific endeavor. Significant scientific progress is only conceivable with advances in both disciplines: the underlying mechanisms need to be understood in a theoretical framework, and its complete range of predictions must be reflected in reality. Through this interconnectedness, knowledge is advanced either through discoveries of novel or predicted phenomena, or by imposing constraints on previous assumptions by narrowing down the range of valid explanations, through which more robust explanations can prevail. The latter approach is enabled by precision tests, of which the measurement presented here is an example.

The most successful theory of elementary particle physics is a quantum field theory, referred to as the Standard Model (SM) of particle physics. The SM is a description of the fundamental forces and the elementary matter particles, which are described as interacting quantum fields. The structure of the SM is determined by the symmetries that the fundamental forces appear to obey and that determine which elementary interactions can occur. The symmetries of charge and parity are of particular importance for the purposes of this work. Charge symmetry (C) is given if a process is unaffected by an inversion of all charge-like properties, such as the electromagnetic charge, while parity symmetry (P) holds if a process is invariant under spatial inversion. A discovery of non-invariance under such basic symmetries reveals the character of the fundamental forces that mediate that process.

Contrary to the expectation informed by classical physics, both symmetries are individually violated in processes involving the weak nuclear force. Even the combination of both symmetries, charge-parity (CP), can be violated by the weak force despite forming a mirror image of physics that is seemingly immune to the non-invariance of only charge or parity symmetry. Charge-Parity violation was first discovered in 1964 by Cronin, Fitch, et al. at the Brookhaven Alternating Gradient Synchrotron [1], for which Cronin and Fitch were awarded the Nobel Prize in Physics in 1980. As the CP operation relates matter and antimatter, the discovery of CP violation shows a slight yet fundamental difference between the two. Today, the CP-violation mechanism is an integral part of the SM and remains an active part of research, not least because our universe is matter-dominated and the absence of antimatter is still not entirely understood.

The SM is known to be incomplete because it does not explain certain striking cosmological observations, such as the observed matter-antimatter asymmetry in the universe. Furthermore, the SM contains several free parameters that can only be measured; mathematical physics cannot predict them since the underlying mechanism that relates them is unproven, if not unknown. Any observation that contradicts the SM implies the existence of new phenomena and would hint at tangible new discoveries. Experimental tests of the SM can be performed in two main varieties: the measured energy scale can be increased, which can cause previously unknown, heavier elementary particles to be formed or their influence to manifest more strongly. In addition, precision tests can be performed that reduce the SM parameter uncertainties and thereby uncover inaccuracies in our theories, pointing just as strongly to new phenomena.

At the European Organization for Nuclear Research (CERN), a significant part of the scientific program is dedicated to the acceleration and collision of particles and the measurement of these collisions. The higher the energy scale of the collision, the higher the resolution of structures at subatomic scales that can be probed. CERN's particle acceleration complex consists of multiple synchrotrons, of which the Large Hadron Collider (LHC) is the largest in the world, accelerating protons and heavy ions to energies of up to 14 TeV. At four interaction points, particle detectors like the Large Hadron Collider beauty (LHCb) detector measure the collision events and the particle interactions and decays created therein. The LHCb detector is specialized for detecting and reconstructing b and c-hadron decays, which offer an especially advantageous environment for high-precision measurements, including measurements of CP violation.

Charge-parity violation in B-meson decays, specifically in the decay mode $B^0 \to \psi K^0$, offers a theoretically clean opportunity to measure CP violation and the natural constant $\sin(2\beta)$. This parameter is an indicator of CP violation in the SM, and its relation to this decay is described in Section 2.4. The occurrence of CP violation in neutral B-meson systems and the possibility for accessing the natural constants relevant for its parametrization in the SM has been suggested long before its discovery (see for example Refs. [2–4]). Measurements were proposed in which B^0 and $\bar{B}^0$ mesons decay to the same final state, as potentially large CP-asymmetries can result from an interference of decay amplitudes. In particular, the decay $B^0 \to \psi K^0_S$ has been identified as a good candidate for a search for CP violation in the B system, as it offers a theoretically clean access to CP asymmetries [4]. Early measurements of the natural constant $\sin(2\beta)$ were conducted by the ALEPH, OPAL, and CDF collaborations [5–7], which aimed at measuring a CP-asymmetry in $B^0 \to J/\psi K^0_S$ decays and collectively established a confidence level for non-zero $\sin(2\beta)$ at 98 % [5]. CP violation in the neutral B-meson system was discovered simultaneously in 2001 at the electron-positron detectors Belle (KEK) and BaBar (SLAC) [8, 9] in the decay mode $B^0 \to J/\psi K^0$. These measurements have since been further improved by Belle and BaBar [10, 11]. The LHCb experiment has confirmed these findings in more recent measurements [12–14] in multiple decay modes.

The measurement described in this thesis details the measurement of CP violation in three B^0 decay modes with the Run 2 data set of LHCb: the mode of its discovery, $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$, and additional modes exhibiting the same CP-violating mechanism, $B^0 \to J/\psi(\to e^+e^-)K^0_S$ and $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$. The K^0_S mesons are always reconstructed in decays to two oppositely charged pions, and ψ denotes the charmonium states J/ψ or $\psi(2S)$. The measurement has been published in Ref. [15] as

Measurement of CP Violation in $B^0 \to \psi(\to \ell^+\ell^-)K^0_S(\to \pi^+\pi^-)$ Decays
Phys. Rev. Lett. 132 (2024) 021801

and was highlighted by the editors as an Editors' Suggestion. I presented the results publicly in an LHC seminar talk for the first time on June 13$^{\text{th}}$ 2023 [16], a special honor that is rarely granted to students. A combined fit of the full LHCb Run 2 data set is performed, and the following result is obtained

$$S_{\psi K^0_S} = 0.717 \pm 0.013\,(\text{stat}) \pm 0.008\,(\text{syst}),$$
$$C_{\psi K^0_S} = 0.008 \pm 0.012\,(\text{stat}) \pm 0.003\,(\text{syst}).$$

This is the most precise single measurement of the CP-violation amplitude $S_{\psi K^0_S}$. As shown in the following, the amplitude $S_{\psi K^0_S}$ corresponds to the natural constant $\sin(2\beta)$ to a high degree of precision due to the small theoretical impurities in this decay mode. The CP-violation parameter $C_{\psi K^0_S}$ is found to be consistent with zero, which is compatible with the SM expectation for this level of measurement precision. This result is well compatible with previous measurements by the experiments Belle, BaBar, and LHCb and is consistent with theoretical expectations for these parameters in the SM Cabibbo, Kobayashi, and Maskawa formalism.

2 CP violation in the Standard Model of particle physics

In this chapter, the Standard Model of particle physics is introduced (Section 2.1) with an emphasis on the phenomenon of CP violation (Section 2.2). In Section 2.3, the ways in which CP violation manifests itself in different decay modes are discussed, and the relevant formalism is presented. The $B^0 \to \psi K^0_S$ decay is introduced in Section 2.4, and it is explained how CP-violation parameters can be experimentally accessed in this decay mode. Finally, in Section 2.5, the influence of CP violation in the kaon system on the measurement is estimated.

2.1 The Standard Model of particle physics

The Standard Model of particle physics is a quantum field theory describing three of the four fundamental forces in nature, and it is the most successful theory in predicting interactions on subatomic scales [17–20]. In the SM, elementary particles and fundamental forces are described as interacting quantum fields. The characteristic properties and underlying symmetries of these fields restrict the types of fundamental interactions that occur. The SM is defined as a locally gauge-invariant theory, meaning that its Lagrangian density is invariant under a gauge transformation of its fields. As a consequence of imposing gauge invariance, the addition of massless gauge fields, described by integer spin particles, is required, which are referred to as *gauge bosons*. More specifically, for each generator of the underlying symmetry group of a given field, one massless gauge boson is introduced. In quantum electrodynamics, described in the SM in $U(1)$-symmetry, the photon can be identified as the only associated massless gauge boson of the electromagnetic field. In this way, imposing gauge invariance leads to a description of matter particles and fundamental forces that has proven to be remarkably consistent with experimental observations. Analogous implications also extend to the weak and strong interactions, including predictions of gauge-boson self-interactions.

The SM has been significantly developed over the past century to account for all experimental observations, and its predictions have been validated numerous times. While the SM predicts processes on subatomic scales with high accuracy, the consequences of which extend to macroscopic, even cosmological scales, it is incomplete as it does not provide a definitive explanation for all observed phenomena. As an example, after the discovery of CP violation and the discovery of a third particle generation, the SM is able to accurately predict an asymmetry between matter and antimatter through

the mechanism of *CP* violation. Yet, the measured size of this asymmetry is too small to explain the observed baryon asymmetry in the universe [21, 22]. Furthermore, the SM does not include a quantum theory of gravity. Finally, the SM contains 19 independent fundamental parameters, which may be related by a more fundamental mechanism to a fewer number of parameters, for example, through a potential unification of all fundamental forces [23]. At the same time, there is currently no observation that is in direct, unambiguous contradiction to a SM prediction. By means of precision measurements at high energy colliders, as presented in this thesis, the allowed parameter space of the SM is reduced, which can contribute to uncovering tensions observed between SM predictions and experimental observations that hint at novel phenomena in the future.

The SM Lagrangian, from which explicit field dynamics are derived, defines the types of possible interactions and the kinematics of quantum fields. It contains a kinetic term for fermions and their gauge-boson couplings, $\mathcal{L}_{\text{Fermion}}$, a term for gauge-bosons kinematics and gauge-boson self-coupling $\mathcal{L}_{\text{Gauge}}$, and two additional terms describing the dynamics of the Higgs field, $\mathcal{L}_{\text{Higgs}}$, and its coupling to fermions which are realized by the $\mathcal{L}_{\text{Yukawa}}$ term. The total SM Lagrangian, $\mathcal{L}_{\text{SM}}$, can be written as a sum of these terms:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{Fermion}} + \mathcal{L}_{\text{Gauge}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \tag{2.1}$$

Special attention is given in the following to the Yukawa term in Section 2.2, as it gives rise to the flavor structure of the SM and *CP* violation and is therefore central to the measurement presented in this thesis.

The matter particles contained in the SM are *fermions* with half-integer spin and *gauge bosons* with integer spin, which characterize the fundamental forces. A Higgs boson with zero spin, corresponding to a scalar Higgs field, which gives rise to spontaneous symmetry breaking (SSB), causes $W^{\pm}$ and Z bosons to acquire masses. Quarks and leptons acquire masses as a consequence of SSB. In Figure 2.1, the particle content of the SM is shown. The fermions are grouped into *quarks* and *leptons*, and both these families contain three generations of particles with increasing mass but otherwise identical properties. For each SM fermion, there exists a corresponding antifermion with inverse charge-like quantum numbers. The charged leptons are the electron (e), the muon (μ), and the tauon (τ), each of which carries one unit of the elementary charge. The electron is the only stable, charged lepton. For each charged lepton, there is a corresponding chargeless neutrino ν. Neutrinos have been shown to have non-zero mass because of their ability to oscillate between generations with time [25]. Multiple mechanisms for the generation of neutrino masses have been proposed, including SSB [20].

There are six different types of quarks: The up-type quarks *up*, (u), *charm*, (c) and *top*, (t) with electromagnetic charge $Q = +\frac{2}{3}$ in units of the elementary charge, and the down-type quarks *down*, (d), *strange*, (s) and *bottom* or *beauty*, (b) with $Q = -\frac{1}{3}$. Apart from electromagnetic charge, quarks carry one of three kinds of *color-charge*, *red*, *green*, or *blue*. The force-carrier of the strong (nuclear) force is the gluon which carries superpositions of color charge and couples to all color-charged particles, including to itself. The symmetry group $SU(3)_C$ of quantum chromodynamics (QCD), which

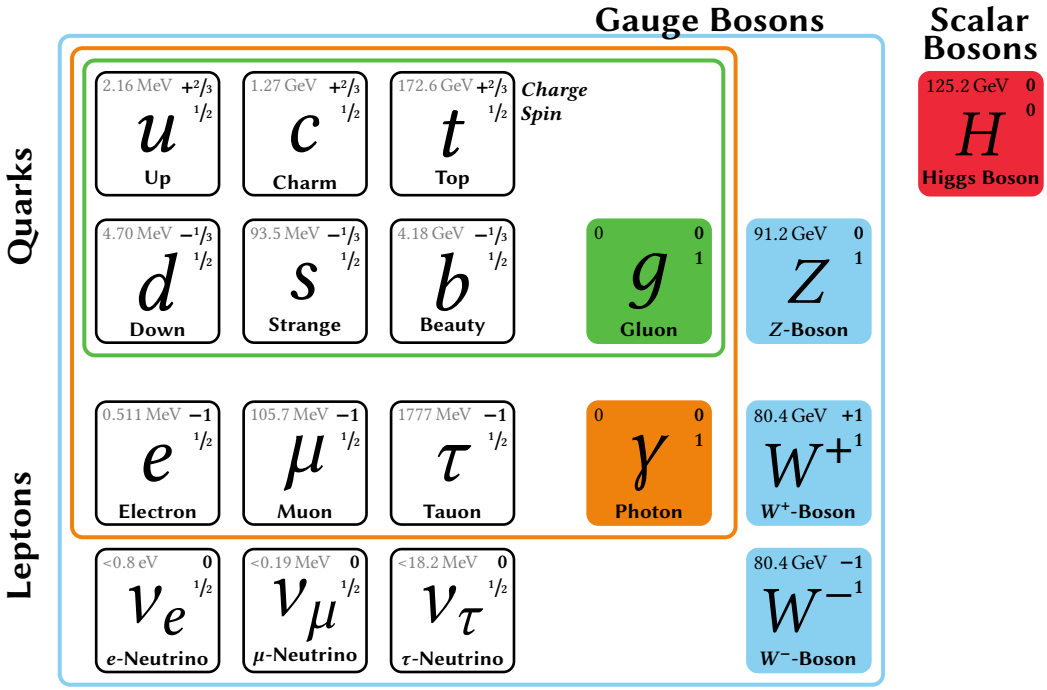

Figure 2.1 – The particle content of the Standard Model. Particle masses are given in natural units, and electromagnetic charges are given in units of the elementary charge. The known interactions of matter particles are symbolized by the colored outlined boxes they are contained in. Particle data taken from Ref. [24].

characterizes the nature of the strong force, implies eight massless gluons of distinct color-charge superpositions. Quarks and gluons do not occur in isolated form, and quarks form bound states of multiple quarks called *hadrons* which are always net-color-neutral. This phenomenon is referred to as *color confinement*. Because of confinement, when quarks in a common bound state are driven apart, new pairs of quarks and antiquarks are created from the gained potential energy and recombine into new hadrons. The process by which quarks interact with the vacuum to form hadrons is referred to as *hadronization*. Bound states of a quark and an antiquark are called *mesons*, while *baryons* are states of three quarks or three antiquarks. In recent years, hadrons with four (*tetraquark*) and five quarks (*pentaquark*) have been discovered at the LHCb experiment [26–28].

The weak (nuclear) force is responsible for quark and lepton-flavor-changing transitions, including radioactive beta decay, and is mediated by three massive gauge bosons. Beta decay was first described by Fermi as an interaction of four fermions through which neutrons are converted into protons with an associated emission of an electron and a neutrino [29]. Intermediate charged bosons with mass have been proposed by Klein for the first time as a mediator of the Fermi interaction [178]. The fact that gauge bosons have mass is seemingly in contradiction to the requirement of gauge invariance, and its presence therefore requires a breakdown of gauge symmetry. This problem was overcome in the 1960s after recognizing that the symmetry of the SM Lagrangian

and the symmetry of its vacuum state do not need to be identical: SM symmetries can be *spontaneously* broken [19, 30–34]. Spontaneous symmetry breaking can be accomplished by introducing an additional scalar field, the Higgs field, which does not alter the global SM symmetry but causes it to break down in its vacuum state. In this case, however, the Goldstone theorem states that for each spontaneously broken symmetry, an additional massless Goldstone-boson should exist [32–34], but such bosons have not been observed. A solution to this problem was found independently by Englert, Brout, and Higgs, who realized that this requirement is more nuanced in the case of broken gauge symmetries and that the Higgs-field-boson effectively absorbs its Goldstone boson and thereby acquires mass [30, 31]. Around the same time, in 1961, the weak and the electromagnetic force were unified by Weinberg, Glashow, and Salam under a common symmetry group $SU(2) \times U(1)$ which, when accounting for SSB, implies three massive gauge bosons as weak force carriers [18, 19, 35]. A neutral current with mass, called Z boson, is thereby predicted as an additional weak force carrier, and by 1983, the W^+, W^- and Z bosons have been discovered experimentally [36, 37]. The photon remains massless after SSB and is charge neutral and couples to all particles with electromagnetic charge. In 2012, the ATLAS and CMS collaborations simultaneously announced the observation of a new particle consistent with the Higgs boson [38, 39], and it remains an active field of research to measure its properties.

Motivated by the fact that parity conservation has not been explicitly tested, and its violation could be a solution to the τ^+/θ^+-puzzle, Lee and Yang proposed a set of experimental tests in 1956 to test parity conservation [40, 41]. Shortly after, parity violation was discovered by Wu et al. in a measurement of the radioactive β-decay of Co_{60} nuclei [42]. The Co_{60} nucleus emits an electron and a neutrino with spins both oriented in parallel to the one of the decaying nucleus. Wu demonstrated that the direction of emitted electrons depends on the direction of the nuclear spin of Co_{60} nuclei that are polarized by an external magnetic field. This asymmetry was further found to be more pronounced the higher the degree of polarization of the sample. The lepton momenta are therefore found to be correlated to their spin direction. Under parity transformation, P, spin transforms as an axial vector, while momentum changes its sign, and therefore this finding can only be explained by P violation [41]. The weak interaction therefore differentiates between particles of different handedness, defined by the projection of the particle's spin onto its momentum vector. In addition to violating parity, the weak force simultaneously violates C symmetry. Experimentally, it can be shown that the decay $\pi^+ \longrightarrow \mu^+\nu_\mu$ has no immediate mirror image under C transformation [43] as it would involve a right-handed neutrino and would thus be most strongly suppressed. To explain these findings, Feynman and Gell-Mann described the weak force as acting exclusively on the left-handed spinor or right-handed antispinor component of particles [44].

The combined symmetry of charge and parity, CP, appears to be more robust, as it provides a symmetry under which the mentioned experiments are indeed invariant. Nevertheless, in 1964, CP violation was discovered neutral kaon decays by Cronin, Fitch et al. [1]. The implementation of CP violation in the SM and the ways in which CP violation manifests itself in measurements are described in the following sections.

2.2 Quark-flavor structure and CP violation

Flavor is a quantum number of SM fermions that is specific to the particle generation and type, and in weak interactions, flavor can change. The selection rules under which flavor changes determine the flavor structure of the SM and, thus, the nature of the weak force. The origin of these selection rules are the interactions of the fermions with the Higgs field after SSB. The unbroken Yukawa SM Lagrangian describing the Higgs-fermion interactions is given by [20]

$$-\mathcal{L}_{\text{Yukawa}} = \overline{Q^I}\Phi Y^D d_R^I + \overline{Q^I}\tilde{\Phi} Y^U u_R^I + \overline{L^I}\Phi Y^E e_R^I + h.c.\,. \tag{2.2}$$

Here, the left-handed quark fields are expressed by Q, and left-handed lepton fields are expressed by L, while right-handed quark and lepton fields are marked with an R index. The index I expresses the fact that the fields are given in the weak interaction basis. The 3×3 complex Yukawa matrices that quantify the coupling of SM fermions to the Higgs field Φ are denoted by $Y^{D,U,E}$.

After SSB, the vacuum expectation value v takes the place of the Higgs field and the Yukawa matrices can be diagonalized by appropriate rotations in flavor space. The unitary transformations that accomplish the diagonalizations, define the relationship between the weak eigenstates and the mass eigenstates of the quark fields, and are parametrized by four unitary 3×3 matrices:

$$u_L^I = V_L^u u_L \qquad d_L^I = V_L^d d_L \qquad u_R^I = V_R^u u_R \qquad d_R^I = V_R^d d_R. \tag{2.3}$$

It is not possible to diagonalize the two quark-related Yukawa matrices simultaneously. By convention, Y^U is diagonalized by the given transformations and the weak and mass eigenstates of up-type quarks are therefore identical. The term Y^D can only be diagonalized in connection with rotated down-quark fields, expressed by an irreducible combination of tranforms, which is identified as the CKM matrix [45, 46]. The CKM matrix quantifies the difference of mass and weak eigenstates of down-quarks, $d_L^I = V_{\text{CKM}} \cdot d_L$. When the charged current interaction of quarks and $W^\pm$ bosons is transformed into the mass basis, the CKM matrix parametrizes the coupling strength [47]:

$$\begin{aligned}\mathcal{L}_{W^\pm}^{\text{Quarks}} &= \frac{g}{\sqrt{2}}\left\{\overline{u_L^I}\gamma_\mu W^{-\mu} d_L^I + \overline{d_L^I}\gamma_\mu W^{+\mu} u_L^I\right\} \\ &\overset{\text{mass basis}}{=} \frac{g}{\sqrt{2}}\left\{\overline{u}_L \underbrace{V_{\text{CKM}}}_{(V_L^u)^\dagger V_L^d}\gamma_\mu W^{-\mu} d_L + \overline{d}_L V_{\text{CKM}}^\dagger \gamma_\mu W^{+\mu} u_L\right\}\end{aligned} \tag{2.4}$$

The CKM matrix appears therefore in the charged-current interaction of quarks and $W^\pm$ bosons and determines the probabilities with which flavor-changing transitions occur between up- and down-type quarks, within and between generations. The elements of the CKM matrix are named as follows:

$$\begin{pmatrix} d^I \\ s^I \\ b^I \end{pmatrix}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L . \tag{2.5}$$

The CKM matrix is a complex and unitary matrix parametrized by three mixing angles and one complex phase δ and can be defined as [48]

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.6}$$

where $s_{ij} \equiv \sin\theta_{ij}$ and $c_{ij} \equiv \cos\theta_{ij}$ and θ_{ij} are the three Euler mixing-angles. Unitarity ensures that the net quark-current is preserved in weak interactions.

Under CP transformation, the charged current interaction is no longer invariant in the quark sector as $V_{\mathrm{CKM}} \neq V^*_{\mathrm{CKM}}$ due to the non-zero, irreducible complex CKM phase δ. A phase-convention-independent indicator of CP violation caused by this non-invariance for a given CKM matrix is the Jarlskog determinant [49], $\mathcal{J} = c_{12}s_{12}s_{13}c_{13}^2 s_{23}c_{23}\sin(\delta)$. The Jarlskog determinant measures the area of unitarity triangles and its non-zero value implies that a given CKM matrix cannot be reduced to basis-dependent mixing angles alone. The specific form of the Jarlskog determinant implies further that all three quark generations need to take part in a process in order for CP-violating effects to occur, as otherwise, any such effect can be nullified with an alternative choice of mixing angles, rendering it unphysical. From the extensive available experimental evidence, it follows that the complex phase δ remains the only source of CP-violation in the SM needed to explain all measurements of CP violation.

The CKM matrix is almost diagonal, from which follows that flavor transitions between generations are far less likely than transitions in the same generations. An approximation of V_{CKM} as an expansion in orders of the Wolfenstein-parameter $\lambda = V_{us}$ [50], with $\lambda = 0.2243 \pm 0.0008$ [51] up to order λ^3 gives the useful representation

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.7}$$

From the unitarity property of the CKM matrix, follow six interdependent unitarity conditions for the CKM elements. The unitarity condition that is of most relevance in the following is

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0. \tag{2.8}$$

As a sum of three complex numbers, which sum to zero, it can be interpreted as a closed triangle in the complex plane. This triangle is shown in Figure 2.2, and this representation is commonly scaled so that its base length is unity. Measurements of flavor-changing processes in the quark sector therefore always constrain the CKM unitarity conditions and serve as a test of the SM. From simple geometric considerations, the angles of this triangle can be defined as

$$\alpha \equiv \arg\left(-\frac{V_{td}V^*_{tb}}{V_{ud}V^*_{ub}}\right) \qquad \beta \equiv \arg\left(-\frac{V_{cd}V^*_{cb}}{V_{td}V^*_{tb}}\right) \qquad \gamma \equiv \arg\left(-\frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}}\right). \tag{2.9}$$

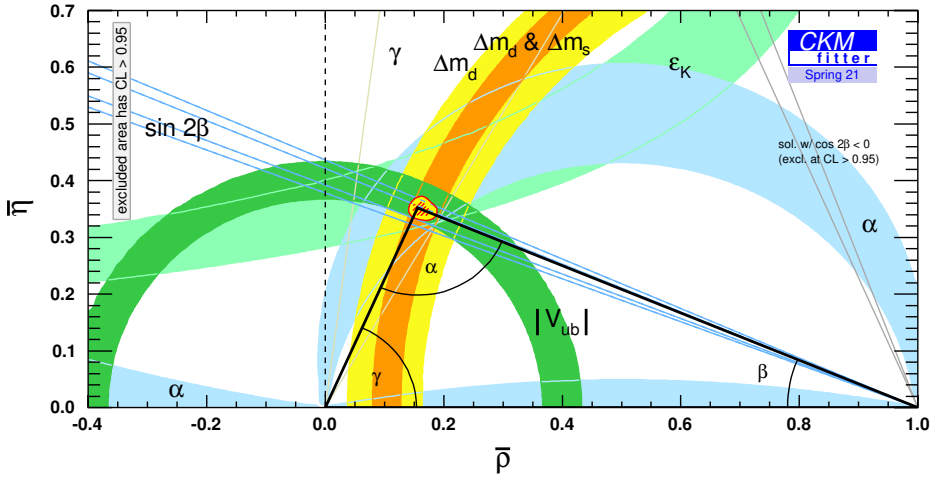

Figure 2.2 – CKM triangle fit for the unitarity condition in Equation (2.8) showing the global constraints on the CKM angles and the triangle apex for the 2021 averages [52].

The measurement of the CKM angle β is the main subject of the measurement presented in this thesis. This angle is geometrically related to the unitarity triangle apex coordinates, which are in turn related to δ in the Wolfenstein parametrization:

$$A\lambda^3(\rho + i\eta) = s_{13}e^{i\delta} \quad \text{and} \quad \beta = \arctan\left(\frac{\overline{\eta}}{1-\overline{\rho}}\right). \tag{2.10}$$

The parameters ρ and η define the triangle apex coordinates $\overline{\rho} = \rho(1 - \lambda^2/2)$ and $\overline{\eta} = \eta(1 - \lambda^2/2)$. This establishes the relationship of the CKM phase δ and the CKM angle β. Consequently, vanishing β would imply $\overline{\eta} = 0$ and finally $\delta = 0$, demonstrating that the angle β is an indicator of CP violation in the SM.

2.3 Classification of CP-violating phenomena

All observations of CP violation can be attributed to the non-zero value of the complex phase δ of the CKM matrix in Equation (2.6). Despite this singular origin, in a measurement of a physical system, there are multiple ways in which CP violation can occur. The distinction is made between *direct* and *indirect* CP violation, and the latter is closely related to *neutral meson mixing* or *oscillation*. Mixing refers to the transition of a neutral meson to its antiparticle state before its decay, and it is mediated by the weak force. Direct CP violation does not require mixing and can therefore also occur in charged meson decays, where it is the only possible CP-violating mechanism.

The hypothesis that neutral mesons are particle mixtures was introduced by Pais and Gell-Mann to reconcile the transformation properties of neutral kaon and antikaon production and decay under charge conjugation [53]. In a later publication [54], Gell-Mann and Pais propose a method to demonstrate the existence of mixing through

a measurement of neutral-kaon absorption in matter, which should cause the short and long-lived kaon mass eigenstates to be converted into one another. Six years later, kaon regeneration and the mass difference of kaon mass eigenstates are measured for the first time [55], supporting this theory. Today, neutral meson mixing and *CP* violation are well-established phenomena and have been measured in multiple neutral meson systems to a high degree of precision [56]. The phenomenon of neutral-kaon regeneration is detailed in Section 2.5 and its possible implications for *CP*-violation observables in B^0 decays to neutral kaons, primarily in the decay $B^0 \to \psi K^0_S$, are discussed.

The two leading-order Feynman-diagrams representing neutral B-meson mixing are shown in Figure 2.3. As with neutral kaons, the neutral B meson state should be regarded as a two-state system of flavor eigenstates B^0 and $\overline{B}^0$. The time evolution of a general neutral meson system, consisting of meson and antimeson $\Psi(t) = (X^0(t), \overline{X}^0(t))^\top$ can be described by a Schrödinger-like equation of the form

$$i\frac{\partial}{\partial t}\begin{pmatrix} X^0(t) \\ \overline{X}^0(t) \end{pmatrix} = \mathcal{H}\begin{pmatrix} X^0(t) \\ \overline{X}^0(t) \end{pmatrix}. \tag{2.11}$$

An effective, non-Hermitian, Hamiltonian operator $\mathcal{H}$ enables transitions between flavor eigenstates and transitions to the final state, whereby the latter is not accounted for by the state vector. As a consequence of *CPT* invariance, the effective Hamiltonian operator $\mathcal{H}$ has symmetric properties because of identical masses and decay widths of the flavor eigenstates of X. The Hamiltonian is defined as a sum of two Hermitian operators, $\mathbf{M}$ and $\mathbf{\Gamma}$, which parametrize the transitions between flavor eigenstates, including those involving virtual intermediate final states and the irreversible transitions to the final state:

$$\mathcal{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} m & m_{12} \\ m^*_{12} & m \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma^*_{12} & \Gamma \end{pmatrix}. \tag{2.12}$$

Oscillations are parametrized by off-diagonal matrix elements, and decays to the final state f are introduced by the $\mathbf{\Gamma}$ term. The mass and decay width of the flavor eigenstates are m and Γ. The parameter m_{12} parametrizes contributions from the box diagrams quantifying the off-shell transitions, which are shown for the leading-order terms in Figure 2.3 for the B^0 system. Oscillation through any intermediate decayed on-shell state of the form $B^0 \to f \to \bar{B}^0$ are parametrized by Γ_{12}. This contribution is negligible in the B system [57], therefore $|\Gamma_{12}| \ll |m_{12}|$ can be assumed. An explicit representation of the effective Hamiltonian in terms of quark-field couplings to the $W^\pm$ bosons via the transitions represented in Figure 2.3 is given in Ref. [58]. In Ref. [59] this result is reused to derive further B mixing parameters and in Ref. [57] the results of these calculations are summarized for m_{12} as

$$m_{12} = -\frac{G_F^2 m_W^2 \eta_B m_B B_B f_B^2}{12\pi^2} \cdot S_0\left(\frac{m_t^2}{m_W^2}\right) \cdot |V^*_{td} V_{tb}|^2, \tag{2.13}$$

where G_F and f_B are the Fermi constant and the B-decay constant; m_B, m_W and

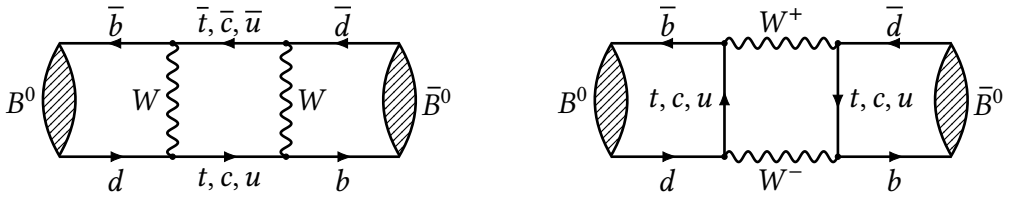

Figure 2.3 – Leading-order Feynman diagrams representing B^0-$\bar{B}^0$ oscillation. Transitions with intermediate t-quarks are favored.

m_t are the B-meson, $W^\pm$ boson and top-quark masses, respectively. The function $S_0(x)$ is the Inami-Lim function [60] parametrizing the contribution from the internal quark propagators as a function of the quark mass and can be approximated as $S_0(x) = 0.784x^{0.76}$ [57]. Perturbative QCD contributions are given by η_B, while non-perturbative contributions are parametrized by the bag parameter B_B. The Inami-Lim function of the argument m_t^2/m_W^2 is of order 1, while contributions from c-quarks are suppressed by four orders of magnitude. In summary, the internal quark lines in the B mixing box diagrams are strongly dominated by t-quark contributions and proceed to first order as depicted in Figure 2.3. While the contributing Inami-Lim factor in the neutral kaon system is the same, the t-quark loop is strongly CKM-suppressed, and thus the c-quark loop dominates instead [61].

To proceed with solving the two-state system explicitly, the effective Hamiltonian operator $\mathcal{H}$ is diagonalized so that Equation (2.11) decouples to an orthogonal basis. In the basis of orthogonal mass eigenstates X_H and X_L one obtains

$$i\frac{\partial}{\partial t}\begin{pmatrix} X_\mathrm{L}(t) \\ X_\mathrm{H}(t) \end{pmatrix} = \begin{pmatrix} \mu_\mathrm{L} & 0 \\ 0 & \mu_\mathrm{H} \end{pmatrix}\begin{pmatrix} X_\mathrm{L}(t) \\ X_\mathrm{H}(t) \end{pmatrix}, \tag{2.14}$$

with eigenvalues

$$\mu_\mathrm{L} = m - \Re F - \frac{i}{2}(\Gamma + 2\Im F) \equiv m_\mathrm{L} - \frac{i}{2}\Gamma_\mathrm{L} \tag{2.15}$$

$$\mu_\mathrm{H} = m + \Re F - \frac{i}{2}(\Gamma - 2\Im F) \equiv m_\mathrm{H} - \frac{i}{2}\Gamma_\mathrm{H} \tag{2.16}$$

and

$$F = \sqrt{\left(m_{12} - \frac{i}{2}\Gamma_{12}\right)\left(m_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}. \tag{2.17}$$

The masses and decay widths of the mass eigenstates X_H and X_L are denoted as $m_\mathrm{H,L}$ and $\Gamma_\mathrm{H,L}$. The mass difference of the mass eigenstates, Δm, and their decay width difference $\Delta\Gamma$ characterize the mixing dynamics of the neutral meson system and are defined by [56]

$$\Delta m \equiv m_\mathrm{H} - m_\mathrm{L} = 2\Re F, \tag{2.18}$$

$$\Delta\Gamma \equiv \Gamma_\mathrm{H} - \Gamma_\mathrm{L} = -4\Im F. \tag{2.19}$$

In the B system, from $|\Gamma_{12}| \ll |m_{12}|$ follows that $\Delta\Gamma$ is a small parameter, which has so far not been measured to deviate significantly from zero [62]. Therefore, the off-shell contribution m_{12} is the main component of the parameter Δm. It is demonstrated in the following that Δm is the frequency of the time-dependent mixing asymmetry.

The mass eigenstates can now be written as an admixture of flavor eigenstates

$$|X_\mathrm{L}\rangle \equiv p\,|X^0\rangle + q\,|\overline{X}^0\rangle \qquad \text{and} \qquad |X_\mathrm{H}\rangle \equiv p\,|X^0\rangle - q\,|\overline{X}^0\rangle \tag{2.20}$$

$$\tag{2.21}$$

with complex coefficients p and q satisfying $|p|^2 + |q|^2 = 1$ that fulfill

$$\frac{p}{q} = \sqrt{\frac{m_{12} - \frac{i}{2}\Gamma_{12}}{m_{12}^* - \frac{i}{2}\Gamma_{12}^*}}. \tag{2.22}$$

A trivial time evolution can then be applied to the mass eigenstates to yield

$$|X_\mathrm{L}(t)\rangle = e^{-i\mu_\mathrm{L}t}(p\,|X^0\rangle + q\,|\overline{X}^0\rangle) \quad \text{and} \quad |X_\mathrm{H}(t)\rangle = e^{-i\mu_\mathrm{H}t}(p\,|X^0\rangle - q\,|\overline{X}^0\rangle). \tag{2.23}$$

As mesons are initially produced as pure flavor eigenstates and experimental techniques are able to measure these initial states, it is more practical to now solve for the time dependencies of these initial states:

$$\begin{aligned} |X^0(t)\rangle &= \frac{1}{2p}(|X_\mathrm{L}(t)\rangle + |X_\mathrm{H}(t)\rangle) \\ &= g_+(t)\,|X^0\rangle + \frac{q}{p} g_-(t)\,|\overline{X}^0\rangle\,, \end{aligned} \tag{2.24}$$

$$|\overline{X}^0(t)\rangle = g_+(t)\,|\overline{X}^0\rangle + \frac{p}{q} g_-(t)\,|X^0\rangle\,, \tag{2.25}$$

with

$$g_\pm(t) = \frac{1}{2}\left(e^{-im_\mathrm{H}t - \Gamma_\mathrm{H}t/2} \pm e^{-im_\mathrm{L}t - \Gamma_\mathrm{L}t/2}\right). \tag{2.26}$$

Here, the states $|X^0\rangle$ and $|\overline{X}^0\rangle$ denote the meson states at the time of production, while $g_\pm(t)$ describe their relative decay probabilities with time. To obtain decay rates which are the subject of measurement, the amplitudes $A_f = \langle f|H|X^0\rangle$ and $\overline{A}_f = \langle f|H|\overline{X}^0\rangle$ of the decay process are squared. As the final state f is CP invariant and therefore $f = \overline{f}$ in this measurement, amplitudes of the CP-conjugated final states involving $\overline{f}$ states are omitted here, but their derivation proceeds analogously.

$$\begin{aligned} \Gamma(X^0(t) \to f) &= \left| \underbrace{\langle f|H|X^0\rangle}_{A_f} g_+(t) + \frac{q}{p} \underbrace{\langle f|H|\overline{X}^0\rangle}_{\overline{A}_f} g_-(t) \right|^2 \\ &= |A_f|^2 \left|g_+(t) + \lambda_f g_-(t)\right|^2 = \frac{1}{2} e^{-\Gamma t} \frac{|A_f|^2}{1 + |\lambda_f|^2} \times \\ &\left\{ \cosh\left(\frac{1}{2}\Delta\Gamma t\right) + \frac{2\mathfrak{R}\lambda_f}{1 + |\lambda_f|^2} \sinh\left(\frac{1}{2}\Delta\Gamma t\right) + \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2} \cos(\Delta m t) - \frac{2\mathfrak{I}\lambda_f}{1 + |\lambda_f|^2} \sin(\Delta m t) \right\}, \end{aligned} \tag{2.27}$$

and similary, for the $\overline{B}^0$ initial state one finds

$$\Gamma(\overline{X}^0(t) \to f) = \frac{1}{2} e^{-\Gamma t} \left|\frac{p}{q}\right|^2 \frac{|A_f|^2}{1+|\lambda_f|^2} \times \left\{ \cosh\left(\frac{1}{2}\Delta\Gamma t\right) + \frac{2\Re\lambda_f}{1+|\lambda_f|^2} \sinh\left(\frac{1}{2}\Delta\Gamma t\right) - \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2} \cos(\Delta m t) + \frac{2\Im\lambda_f}{1+|\lambda_f|^2} \sin(\Delta m t) \right\}. \tag{2.28}$$

A complex parameter λ_f, defined as

$$\lambda_f \equiv \frac{q}{p} \cdot \frac{\overline{A}_f}{A_f} = \left|\frac{q}{p}\right| \cdot \left|\frac{\overline{A}_f}{A_f}\right| \cdot e^{2i\phi} \tag{2.29}$$

is introduced here and defines the oscillating behavior of the decay rate caused by possible *CP*-violating effects that are specific to the neutral meson system and decay in question, whereby ϕ denotes a weak, *CP*-violating phase. More specifically, if $\lambda_f = 1$, there is no difference between the decay rates of B^0 and $\overline{B}^0$, and otherwise, *CP* violation is present. In the following Sections 2.3.1 to 2.3.3, the possible types of observed *CP*-violating phenomena and the way in which they individually contribute to the parameter λ_f are explained.

2.3.1 *CP* violation in the mixing

CP violation in the mixing occurs if the transition probabilities of the flavor eigenstates are not equal and therefore

$$\Gamma(X^0 \to \overline{X}^0) \neq \Gamma(\overline{X}^0 \to X^0) \qquad \text{as a consequence of} \qquad |q/p| \neq 1. \tag{2.30}$$

The neutral-kaon system is an example of a neutral meson system with that property. As historically, *CP* violation was discovered in the neutral-kaon system, its *CP*-violating nature is explained in more detail in this section.

The kaon flavor eigenstates $|K^0\rangle = |d\overline{s}\rangle$ and $|\overline{K}^0\rangle = |\overline{d}s\rangle$ mix and the two possible *CP*-eigenstates can be written as

$$|K_1\rangle \equiv \frac{1}{\sqrt{2}}(|K^0\rangle + |\overline{K}^0\rangle) \qquad \text{and} \qquad |K_2\rangle \equiv \frac{1}{\sqrt{2}}(|K^0\rangle - |\overline{K}^0\rangle). \tag{2.31}$$

Under *CP* transformation, these states are eigenvectors:

$$CP\,|K_1\rangle = +\,|K_1\rangle\,, \qquad CP\,|K_2\rangle = -\,|K_2\rangle\,. \tag{2.32}$$

The K_1 state is therefore *CP-even* and K_2 is *CP-odd*. Experimentally, two kinds of neutral kaons are observed with drastically different lifetimes: K^0_S and K^0_L. The short-lived state decays predominantly as $K^0_\mathrm{S} \to \pi\pi$ and the long lived state decays predominantly as $K^0_\mathrm{L} \to \pi\pi\pi$. If one identifies the kaon mass eigenstates K^0_S and K^0_L with the admixtures K_1 and K_2, then the short-lived state K^0_S should always decay into a *CP*-even

configuration ($\pi\pi$), whereas the long-lived state K_L^0 should always decay into a CP-odd configuration ($\pi\pi\pi$) if CP symmetry was conserved. As the lifetimes of K_L^0 and K_S^0 differ substantially, it is possible to study their decays almost independently.

In 1964, the decay $K_\mathrm{L}^0 \to \pi\pi$ was observed at the Brookhaven laboratory in an experiment conducted by Cronin, Fitch, and others [1]. The measured final state $|\pi\pi\rangle$ is CP-even, unlike K_2, and therefore, this observation demonstrated for the first time that CP symmetry can be violated in weak interactions. For this discovery, Cronin and Fitch were awarded the Nobel Prize in physics in 1980. Therefore, one concludes that the kaon mass-eigenstates K_S^0 and K_L^0 cannot correspond to the CP eigenstates K_1 and K_2. The mass eigenstates can instead be defined such that they contain a small admixture of the respective other CP-eigenstate:

$$|K_\mathrm{S}^0\rangle \equiv \frac{|K_1\rangle + \epsilon_K |K_2\rangle}{\sqrt{1+\epsilon_K^2}} \qquad |K_\mathrm{L}^0\rangle \equiv \frac{|K_2\rangle + \epsilon_K |K_1\rangle}{\sqrt{1+\epsilon_K^2}} \tag{2.33}$$

$$\equiv p_K |K^0\rangle + q_K |\overline{K}^0\rangle\,, \qquad \equiv p_K |K^0\rangle - q_K |\overline{K}^0\rangle\,, \tag{2.34}$$

whereby ϵ_K is a CP-violating parameter that is known today to a high degree of precision, $|\epsilon_K| = (2.228 \pm 0.011) \times 10^{-3}$ [56], and p_K, q_K are the mixing coefficients in the kaon system. At the same time, $\epsilon_K \neq 0$ implies that the mass eigenstates are not symmetric combinations of the flavor eigenstates K^0 and $\overline{K}^0$, which implies that the mixing process is asymmetric in time, fulfilling the rate condition in Equation (2.30). Finally, as generally the relation of ϵ_K and the mixing coefficients is $\frac{q}{p} = (1-\epsilon_K)/(1+\epsilon_K)$, $|\epsilon_K| > 0$ implies $|q/p| \neq 1$ with the previous definitions.

2.3.2 *CP* violation in the decay

CP violation in the decay, or *direct* CP violation, occurs if the CP-conjugated decay rate of a given decay differs from the original rate, hence

$$\Gamma(X \to f) \neq \Gamma(\overline{X} \to \overline{f}) \quad \text{or} \quad \left|\frac{\overline{A}_{\overline{f}}}{A_f}\right| \neq 1. \tag{2.35}$$

This type of CP violation can occur both in charged and neutral meson systems, and it is the only possible type of CP violation in charged meson systems. The precondition for CP violation in the decay is that each decay amplitude can be expressed as the sum of at least two contributions with generally different weak and strong phases, ϕ_j and θ_j and associated partial amplitudes a_j: $A_f = \sum_j a_j e^{i\phi_j} e^{i\theta_j}$. If the decay amplitude is in this form, the absolute amplitude ratio in Equation (2.35) cannot cancel to 1. For CP violation to occur, the weak and strong phases each need to differ, and their respective partial amplitudes a_j need to differ or at least interfere. Experimentally, CP violation in the decay was observed for the first time at the NA31, NA48, and kTeV experiments in measurements of the double ratio of K_S^0 and K_L^0 mesons decay rates to either $\pi^0\pi^0$ or $\pi^+\pi^-$ final states [63, 64]. This finding implies that even the K_2 state can decay to

two pions and hence *directly* violate *CP*. Yet, in the neutral kaon system, direct *CP* violation occurs far less frequently than *CP* violation in the mixing [65] and is ignored in the following. The occurrence of direct *CP* violation can be explained by the fact that the $K \to \pi\pi$ decay can result in different isospin configurations for the final state, and its amplitude is thus composite. Direct *CP* violation in neutral and charged B meson decays has since been found in multiple decay modes [56]. Experimentally, in decay modes with charge-asymmetric final states, the *CP* asymmetry can be measured by counting the normalized, time-integrated difference of final state frequencies f^+ and f^- [56]:

$$A_{f^\pm} = \frac{\Gamma(\overline{X} \to f^-) - \Gamma(X \to f^+)}{\Gamma(\overline{X} \to f^-) + \Gamma(X \to f^+)} = \frac{|\overline{A}_{f^-}/A_{f^+}|^2 - 1}{|\overline{A}_{f^-}/A_{f^+}|^2 + 1}, \tag{2.36}$$

whereby the initial meson can be charged or neutral.

2.3.3 *CP* violation in the interference of mixing and decay

Neutral meson decays to *CP*-invariant final states can exhibit a third kind of *CP* violation that occurs due to interfering amplitudes of the direct decay and the decay of the mixed meson, *i.e.* $X^0 \to f_{CP}$ and $X^0 \to \overline{X}^0 \to f_{CP}$. The consequence of this kind of *CP* violation is that the time-dependent decay rates of initial X^0 and $\overline{X}^0$ mesons are not identical functions of time:

$$\Gamma(X^0(t) \to f_{CP}) \neq \Gamma(\overline{X}^0(t) \to f_{CP}) \qquad \text{due to} \qquad \Im\lambda_f \neq 0. \tag{2.37}$$

The time evolutions of the mixing of the initial flavor eigenstates are $X^0(t)$ and $\overline{X}^0(t)$ for decays to the same final state that have been derived in Equations (2.27) and (2.28), allow for *CP* violation through the complex parameter λ_f. A measurement of the associated, time-dependent decay rate asymmetry, or *CP asymmetry*

$$\mathcal{A}^{CP}(t) = \frac{\Gamma(\overline{X}^0(t) \to f_{CP}) - \Gamma(X^0(t) \to f_{CP})}{\Gamma(\overline{X}^0(t) \to f_{CP}) + \Gamma(X^0(t) \to f_{CP})} \tag{2.38}$$

gives access to the observables parametrizing *CP violation in the interference*:

$$\mathcal{A}^{CP}(t) = \frac{S_f \sin(\Delta m_d t) - C_f \cos(\Delta m_d t)}{\cosh\left(\frac{1}{2}\Delta\Gamma_d t\right) + A_f^{\Delta\Gamma} \sinh\left(\frac{1}{2}\Delta\Gamma_d t\right)}. \tag{2.39}$$

The following *CP*-violation parameters are introduced from a comparison with Equations (2.27) and (2.28):

$$S_f \equiv \frac{2\Im\lambda_f}{1 + |\lambda_f|^2} \qquad C_f \equiv \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2} \qquad A_f^{\Delta\Gamma} \equiv \frac{2\Re\lambda_f}{1 + |\lambda_f|^2}. \tag{2.40}$$

These parameters are directly accessible to measurements as they modify the exponential B decay rates and parametrize the amplitude and phase of the measured *CP* asymmetry in Equation (2.39). By construction, the normalization condition

$$S_f^2 + C_f^2 + \left(A_f^{\Delta\Gamma}\right)^2 = 1 \tag{2.41}$$

holds. In the case $S_f = C_f = 0$, no CP violation of any kind is observed. The case $C_f = 0$ is hereby fulfilled if CP violation in the mixing (Section 2.3.1) and in the decay (Section 2.3.2) are not present or cancel, which is equivalent to $|p/q| = 1$ and $|A_f/\overline{A}_f| = 1$ and thus $|\lambda_f| = 1$ according to Equation (2.29). Importantly, this condition does not imply $\arg(\lambda_f) = 0$, which then allows for the possibility

$$\Im(\lambda_f) = S_f = \sin(2\phi) \neq 0 \qquad \text{with} \qquad \phi = \frac{1}{2}\arg\left(\frac{q}{p}\frac{\overline{A}_f}{A_f}\right). \tag{2.42}$$

The requirement that the imaginary part of λ_f is non-vanishing defines the presence of CP violation in the interference of decays with and without mixing.

In general, all three types of CP violation can occur simultaneously, and if the aim is to precisely measure the weak phase ϕ, all sources of CP violation need to be accounted for. Nevertheless, even without or with negligible CP violation in the mixing and decay, interference CP violation is possible and can be a large effect, as is the case for the CP-violating decay $B^0 \to \psi K^0_S$ which is introduced in Section 2.4.

It should be noted that if the weak phase ϕ is obtained from measurements, it can include additional weak phases, as well as CP-conserving contributions which can manifest as $C_f \neq 0$. These contributions can be present due to higher-order decay transitions. For this reason, in generality, a measurement of the CP-violation amplitude S does not give access to the pure angle ϕ, but rather an effective phase $\phi_{\text{eff}} = \phi + \Delta\phi$ instead. The properties of the decay in question determine the relative size of CP-violating and CP-conserving phases, and the latter can contribute to $\Delta\phi_d$. In addition, there are trigonometric ambiguities associated with the approach of measuring β through a measurement of $\sin(2\beta)$. These details are explored in the next section for the CP-violating decay $B^0 \to \psi K^0_S$.

2.4 The CP-violating decay $B^0 \to \psi K^0_S$

The properties of the decay $B^0 \to \psi K^0_S$ make it particularly well suited for precise measurements of the CKM angle β. It is shown in the following that the tree level $b \to c\bar{c}s$ transition is the leading contribution to the decay and that partially CP-conserving phases, induced by higher order decay transitions, either have the same weak phase as the tree-process or are at least doubly Cabibbo-suppressed otherwise. Furthermore, the parameter $S_{\psi K^0_S}$ was measured to be large, and $C_{\psi K^0_S}$ is compatible with zero to high accuracy [62]. Finally, the decay topology of two high-transverse-momentum leptons and two charged pions with significantly separated origin vertices allow for a good separation of backgrounds in a hadronic collision environment.

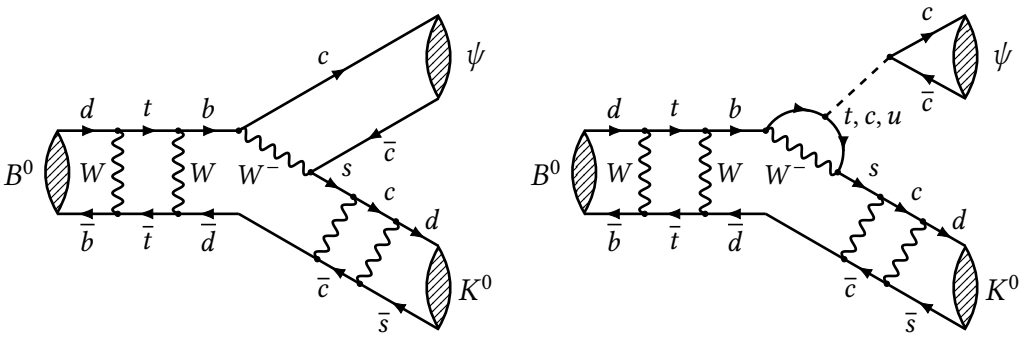

Figure 2.4 – Left: Feynman diagram of the $B^0 \to \psi K^0$ decay with neutral meson oscillations in the initial and final state. Right: Feynman diagram of the contribution of the color-singlet exchanges to the same decay.

CP violation in the B^0 system was first discovered at the B-factories Belle [8] and BaBar [9] in multiple $B^0 \to \psi K^0_S$ and $B^0 \to \psi K^0_L$ decay channels. This finding confirmed the predictions for large CP violation in this mode [4]. At the B factories, electrons and positrons are brought to collision at a center-of-mass energy corresponding to the rest mass of the $\Upsilon(4S)$ state, which predominantly decays to a pair of either two charged or two neutral B mesons, which are Lorentz-boosted due to the momentum-asymmetric collision. CP violation in the mixing has not been observed in the B^0 system [56], therefore $|q/p| = 1$ is assumed from here on. Feynman diagrams of the leading-order tree-level decay and the penguin decay are shown in Figure 2.4, where neutral meson oscillation in the initial and final states is represented by box-diagrams.

The initial-state B meson can decay directly into the final state ψK^0_S or undergo mixing before its decay to the same final state. While B^0 mesons decay to ψK^0, $\overline{B}^0$ mesons decay to the CP-conjugated final state $\psi \overline{K}^0$, but as neutral kaons oscillate and are exclusively measured in the physical states K^0_S or K^0_L, the final state is CP-invariant. For this reason, the direct decay can therefore interferes with the mixed decay, which gives rise to interference CP violation. The main contribution to the decay amplitude comes from the tree-level transition $b \to c\bar{c}s$ and in sub-leading-order, *penguin*-type

diagrams involving color-singlet exchange with three possible internal quarks u, c and t contribute.

The total decay amplitude, not considering B or K-mixing, can be written as a sum of these contributions, which are proportional to the relevant CKM factors $\lambda_q = V_{qs}V_{qb}^*$ [66],

$$A_{\psi K^0} = \lambda_c(A^c_{\text{Tree}} + A^c_{\text{Pen}}) + \lambda_u A^u_{\text{Pen}} + \lambda_t A^t_{\text{Pen}}, \tag{2.43}$$

with the tree A^c_{Tree} and penguin amplitudes A^q_{Pen} for an internal quark q. From CKM unitarity, it follows $\lambda_t = -\lambda_c - \lambda_u$, which is applied to yield

$$= \lambda_c(A^c_{\text{Tree}} + A^c_{\text{Pen}} - A^t_{\text{Pen}}) + \lambda_u(A^u_{\text{Pen}} - A^t_{\text{Pen}}) \tag{2.44}$$

$$= \lambda_c(A^c_{\text{Tree}} + A^c_{\text{Pen}} - A^t_{\text{Pen}})\left\{1 + \frac{\lambda_u}{\lambda_c} \cdot \frac{A^u_{\text{Pen}} - A^t_{\text{Pen}}}{A^c_{\text{Tree}} + A^c_{\text{Pen}} - A^t_{\text{Pen}}}\right\}, \tag{2.45}$$

The amplitude difference between t and u loop diagrams is therefore relevant for corrections to the tree-level process. After substituting the amplitude proportional to $A^u_{\text{Pen}} - A^t_{\text{Pen}}$ with the notation used in the literature [66], namely

$$a' e^{i\theta'} \equiv \frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right|\left(1 - \frac{\lambda^2}{2}\right)\left(\frac{A^u_{\text{Pen}} - A^t_{\text{Pen}}}{A^c_{\text{Tree}} + A^c_{\text{Pen}} - A^t_{\text{Pen}}}\right) \tag{2.46}$$

the following expression is obtained when explicit λ^4 terms are omitted:

$$A_{\psi K^0} = \lambda_c(A^c_{\text{Tree}} + A^c_{\text{Pen}} - A^t_{\text{Pen}})\left\{1 + \frac{\lambda^2}{1-\lambda^2} a' e^{i\theta'} \exp\left(i \arg \frac{\lambda_u}{\lambda_c}\right)\right\}. \tag{2.47}$$

The same CKM factors of $\lambda_c = V_{cs}V_{cb}^*$ therefore contribute to the tree-level and to the amplitude difference of the c and t-penguin decay, which enters in leading order. The term $a'e^{i\theta'}$ includes a CP-conserving strong phase θ' representing the relative phase difference between the tree-level process and the penguin decays, with an associated amplitude of a'. Contributions from penguin topologies that contribute with an additional strong and a weak phase can be viewed as being doubly Cabibbo-suppressed by a factor of $\tilde{\epsilon} = \lambda^2/(1-\lambda^2)$ in excellent approximation. This suppressed, additional weak phase can be identified with the CKM angle γ by rewriting the complex argument of λ_u/λ_c

$$\arg\left(\frac{\lambda_u}{\lambda_c}\right) = \arg\left(\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*}\right) = \arg\left(\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) + \arg\left(\frac{V_{us}V_{cd}}{V_{cs}V_{ud}}\right) = -\gamma + \zeta', \tag{2.48}$$

whereby ζ' is zero in the leading-order Wolfenstein expansion of the CKM matrix. Using an $\mathcal{O}(\lambda^6)$ expansion of the CKM matrix [67] and latest Wolfenstein parameters [51], the CKM suppression factor $\tilde{\epsilon} \approx 0.053$ is obtained with an insignificant correction $\zeta' \approx 6.7\times10^{-4}$ to the angle γ which can be safely ignored in the following. In conclusion, the $B^0 \to \psi K^0$ decay amplitude can be written as

$$A_{\psi K^0} = V_{cs}V_{cb}^*(A^c_{\text{Tree}} + A^c_{\text{Pen}} - A^t_{\text{Pen}})\left\{1 + \tilde{\epsilon} a' e^{i\theta'} e^{i\gamma}\right\}. \tag{2.49}$$

The most significant, *i.e.* Cabibbo-allowed, penguin transitions contribute with the same CKM factors as the tree-level process and therefore share its weak phase and do not cause any corrections to ϕ to occur. As penguin topologies that introduce corrections to the measurement objective $\sin(2\beta)$ are at least suppressed by $\tilde{\epsilon}$ and penguin amplitudes are loop-suppressed, the influence of penguin-decays is rather small in this decay mode.

To find the relation between the CKM angle β and the amplitude $S_{\psi K^0_S}$, the *CP* violation parameter $\lambda_{\psi K^0_S}$ is derived for the $B^0 \to \psi K^0_S$ decay. Doubly Cabibbo-suppressed hadronic corrections from Equation (2.49) are omitted, and these corrections can be determined from additional decay modes in separate measurements as discussed below. As explained in Sections 2.3 and 2.3.1, the neutral-kaon system exhibits *CP* violation in the mixing, and the main contribution to the mixing is due to internal c-quarks. *CP* violation in the neutral-kaon system is a small effect as estimated by the parameter ϵ_K, and so is its influence on the measurement of β, which is demonstrated in Section 2.5. From these assumptions follows, that $|\lambda_{\psi K^0_S}| = 1$. The *CP*-violation parameter $\lambda_{\psi K^0_S}$ then becomes

$$\lambda_{\psi K^0_S} = \frac{q}{p} \cdot \frac{\overline{A}_{\psi K^0_S}}{A_{\psi K^0_S}} = \frac{q}{p} \cdot \frac{\overline{A}_{\psi \bar{K}^0}}{A_{\psi K^0}} \cdot \frac{q_K}{p_K} = - \underbrace{\frac{V^*_{tb} V_{td}}{V_{tb} V^*_{td}}}_{B^0 \text{ mixing}} \cdot \underbrace{\frac{V^*_{cs} V_{cb}}{V_{cs} V^*_{cb}}}_{B^0 \text{ decay}} \cdot \underbrace{\frac{V^*_{cd} V_{cs}}{V_{cd} V^*_{cs}}}_{K^0 \text{ mixing}} \tag{2.50}$$

$$= - \frac{V^*_{tb} V_{td}}{V_{cd} V^*_{cb}} \cdot \frac{V^*_{cd} V_{cb}}{V_{tb} V^*_{td}}, \tag{2.51}$$

and the *CP*-violation amplitude $S_{\psi K^0_S}$ becomes by the previous definition in Equation (2.42)

$$S_{\psi K^0_S} = \Im \lambda_{\psi K^0_S} = - \sin\left(\arg\left(\frac{V^*_{tb} V_{td}}{V_{cd} V^*_{cb}} \cdot \frac{V^*_{cd} V_{cb}}{V_{tb} V^*_{td}} \right) \right) = - \sin\left(2 \arg\left(\frac{V^*_{tb} V_{td}}{V_{cd} V^*_{cb}} \right) \right), \tag{2.52}$$

whereby a geometric property of the complex argument is used in the last step[1]. Once the defnition of β in Equation (2.9) is substituted one obtains the amplitudes

$$S_{\psi K^0_S} = \sin(2\beta) \qquad \text{and} \qquad C_{\psi K^0_S} = 0 \tag{2.53}$$

for the tree-level relation between the measurable amplitude $S_{\psi K^0_S}$ and the CKM angle β. With the previous assumptions, *i.e.* in absence of direct *CP*-violation and *CP*-violation in the mixing, $|\lambda_f| = 1$ and $C_{\psi K^0_S} = 0$. It is interesting to note that the contribution of CKM elements from kaon mixing in Equation (2.50) enables a clean cancellation of terms leading to an exact representation of the phase 2β, but even in the hypothetical absence of those terms the deviation to the measured phase due to imperfect cancellation is one permille of the value of $\sin(2\beta)$. The large *CP*-violating effect in this mode is therefore attributed to the interference of B mixing and decay.

Penguin contributions to the decay amplitude add an additional phase shift $\Delta\phi_d$ to the value of 2β, and therefore the amplitude $S_{\psi K^0_S}$ is expressed more precisely as

[1] $\arg(z/z^*) = 2 \arg z$.

$S_{\psi K^0_S} = \sin(2\beta + \Delta\phi_d + \Delta\phi^{\mathrm{NP}})$, whereby ${\Delta\phi_d}^{\mathrm{NP}}$ is a potential New Physics contribution. The phase shift $\Delta\phi_d$ is expected to be small, $\Delta\phi_d \approx -1^\circ$ [68–73], which is not fully in reach of current experimental precision. In future measurements of β at LHCb and the B factories, the determination of $\Delta\phi_d$ is thus going to be more relevant to disentangle β from $\Delta\phi_d$. The value of $\Delta\phi_d$ can be estimated in measurements of decays with similar decay topology like $B^0_s \to \psi K^0$ or $B^0 \to J/\psi\{\pi^0, \phi, \rho\}$ [68]. In these decays, using similar derivations as above, one finds that the suppression factor $\tilde{\epsilon}$ is absent and that the penguin contributions are therefore Cabibbo-allowed. As these channels share several CKM elements, corrections can be found for the $B^0 \to \psi K^0_S$ mode. The $B^0_s \to \psi K^0$ decay is related to $B^0 \to \psi K^0$ through U-Spin symmetry whereby the decay modes share the same strong phase, and for this reason B^0_s mode offers an ideal channel for the determination of $\Delta\phi_d$. It is therefore possible to correct any measurement of $S_{\psi K^0_S}$ by such a future measurement [68].

Finally, from a measurement of $S_{\psi K^0_S}$ alone, β cannot be extracted without ambiguity as the sine function is cyclic, and from a measurement of just the amplitude $S_{\psi K^0_S}$, it is ambiguous whether the argument is 2β or $\pi - 2\beta$ [74]. To resolve the angular ambiguity, it is possible to measure CP violation in modes where the relevant CP observable happens to be $\cos(2\beta)$ as was done in Ref. [75]. In addition, it is necessary to determine $\mathrm{sgn}(\sin(\beta))$, which relies on input from theory [76]. From the available data and present theoretic assumptions, it follows that the CKM apex lies in the first quadrant of the complex $\overline{\rho}$-$\overline{\eta}$ plane [62, 77].

2.5 Kaon regeneration and CP-violation corrections

A neutral-kaon flavor eigenstate K^0 or $\overline{K}^0$ converts to a K^0_S state in exactly 50 % of cases through mixing and becomes a K^0_L state otherwise. Therefore, half of all neutral-kaon decays are not selected a priori as K^0_L mesons cannot be reconstructed at LHCb. The final state kaon in the signal decay mode is reconstructed from decays into two charged pions of short-lived neutral-kaon states, and due to its short lifetime, the K^0_S state is almost exclusively selected.

Nevertheless, pollution of CP-violating $K^0_L \to \pi^+\pi^-$ decays can influence measurements of CP-violation observables $S_{\psi K^0_S}$ and $C_{\psi K^0_S}$ in the B^0 system. That is because $B^0 \to \psi K^0_L$ decays have a CP eigenvalue of the opposite sign, and therefore, the sign of the *observed* asymmetry appears smaller on average if such decays are present in the data. The opposite sign of the associated K^0_L amplitude has also been demonstrated in measurements [8]. One can write

$$S^{\mathrm{obs}}_{\psi K^0_S} = -S^{\mathrm{obs}}_{\psi K^0_L}. \tag{2.54}$$

A fully inclusive measurement of CP-violation of $B^0 \to \psi K^0_{S,L}$ modes, even results in a total CP-asymmetry of zero [4].

The admixture of K^0_L decays is further amplified by the mechanism of kaon regeneration, which was first proposed by Pais and Piccioni [54]. Accordingly, a K^0_L meson can

turn into a K^0_S meson that decays to two pions, in which case the effectively selected decay is $B^0 \to \psi K^0_L$ despite the correct decay signature. This effect was experimentally confirmed for the first time in 1961 [55]. In a vacuum, the mass eigenstates K^0_S and K^0_L do not mix as they are orthogonal states with respect to the Hamiltonian operator in Equation (2.11). When these states propagate through the detector material, a very significant difference between K^0 and $\overline{K}^0$ interactions becomes relevant as the absorption of $\overline{K}^0$ is far more likely through interactions of the kind

$$\overline{K}^0 + p \to \Lambda + \pi^+ \qquad \text{or} \qquad \overline{K}^0 + n \to \Lambda + \pi^0, \tag{2.55}$$

while interactions like $K^0 + p \to K^+ n$ are less likely [54]. In addition, elastic scattering of K^0 and $\overline{K}^0$ with matter can further contribute to a restoration of a K^0 component [54]. Kaon regeneration therefore causes a loss of the $\overline{K}^0$ component in a beam of K^0_L mesons, through which the kaon states are brought back to K^0-enhanced states, which can acquire the K^0_S state again. K^0_S mesons are thus said to *regenerate*.

In the following, the corrections to the amplitudes $S_{\psi K^0_S}$ and $C_{\psi K^0_S}$ caused by these effects are derived with the regeneration formalism developed in Ref. [78, 79]. Such corrections have been determined for the predecessor analyses [13, 14], and the corrections were found to be very small given the sparse material distribution of the LHCb detector, low densities of utilized materials and high momenta of produced kaons. Nevertheless, the effect is large enough to change the reported values and is comparable in size to some systematic uncertainties in this analysis, and for this reason, it is estimated.

From Equation (2.34) follows that the mass eigenstates are composed in terms of flavor eigenstates as follows

$$|K^0_L\rangle = \frac{1}{\sqrt{2(1+|\epsilon_K|^2)}}\left[(1+\epsilon_K)\,|K^0\rangle + (1-\epsilon_K)\,|\overline{K}^0\rangle\right] \tag{2.56}$$

$$|K^0_S\rangle = \frac{1}{\sqrt{2(1+|\epsilon_K|^2)}}\left[(1+\epsilon_K)\,|K^0\rangle - (1-\epsilon_K)\,|\overline{K}^0\rangle\right]. \tag{2.57}$$

A general, time-dependent kaon state $|\Psi\rangle$ can then either be expressed in the basis $(K^0, \overline{K}^0)$ or (K^0_S, K^0_L) [78],

$$|\Psi(t)\rangle = \alpha_L(t)\,|K^0_L\rangle + \alpha_S(t)\,|K^0_S\rangle = \alpha(t)\,|K^0\rangle + \overline{\alpha}(t)\,|\overline{K}^0\rangle\,, \tag{2.58}$$

whereby the basis change is accomplished by

$$\alpha = \frac{1+\epsilon_K}{\sqrt{1+|\epsilon_K|^2}}\frac{\alpha_L+\alpha_S}{\sqrt{2}} \qquad \text{and} \qquad \overline{\alpha} = \frac{1-\epsilon_K}{\sqrt{1+|\epsilon_K|^2}}\frac{\alpha_L-\alpha_S}{\sqrt{2}}, \tag{2.59}$$

with implied time-dependencies. In a vacuum, there are no transitions from K^0_L to K^0_S and the vacuum time evolution of $|\Psi\rangle$ states in proper time, t, can be expressed with a diagonal, effective Hamiltonian operator:

$$i\left(\frac{d\Psi}{dt}\right)_{\text{Vac}} = i\frac{d}{dt}\begin{pmatrix}\alpha_L(t)\,|K^0_L\rangle\\ \alpha_S(t)\,|K^0_S\rangle\end{pmatrix} = \begin{pmatrix}\lambda_L & 0\\ 0 & \lambda_S\end{pmatrix}\begin{pmatrix}\alpha_L(t)\,|K^0_L\rangle\\ \alpha_S(t)\,|K^0_S\rangle\end{pmatrix}, \tag{2.60}$$

with $\lambda_{\mathrm{L,S}} = m_{\mathrm{L,S}} - \frac{i}{2}\Gamma_{\mathrm{L,S}}$ in analogy to the B^0 system. The time evolution in matter is expressed in the basis of $(K^0, \overline{K}^0)$ instead, as these components have different matter cross-sections:

$$i\left(\frac{d\Psi}{dt}\right)_{\mathrm{Mat}} = i\frac{d}{dt}\begin{pmatrix}\alpha(t)\,|K^0\rangle\\ \overline{\alpha}(t)\,|\overline{K}^0\rangle\end{pmatrix} = \begin{pmatrix}\chi & 0\\ 0 & \overline{\chi}\end{pmatrix}\begin{pmatrix}\alpha(t)\,|K^0\rangle\\ \overline{\alpha}(t)\,|\overline{K}^0\rangle\end{pmatrix}, \tag{2.61}$$

whereby the factors χ and $\overline{\chi}$ are complex parameters describing the attenuation and phase shift of K^0 and $\overline{K}^0$ amplitudes when propagating through a matter phase. The total time evolution in the basis of experimental relevance $(K^0_{\mathrm{S}}, K^0_{\mathrm{L}})$ is then

$$i\frac{d\Psi}{dt} = \begin{pmatrix}\lambda_{\mathrm{L}} + \frac{\chi+\overline{\chi}}{2} & \frac{1}{2}\Delta\chi\\ \frac{1}{2}\Delta\chi & \lambda_{\mathrm{S}} + \frac{\chi+\overline{\chi}}{2}\end{pmatrix}\begin{pmatrix}\alpha_{\mathrm{L}}(t)\,|K^0_{\mathrm{L}}\rangle\\ \alpha_{\mathrm{S}}(t)\,|K^0_{\mathrm{S}}\rangle\end{pmatrix}. \tag{2.62}$$

The effect of kaon regeneration is therefore characterized by the parameter $\Delta\chi \equiv \chi - \overline{\chi}$. The solution of this Hamiltonian system is given by [78, 79]

$$\alpha_{\mathrm{L}}(t) = e^{i\Sigma t}\left[\alpha^0_{\mathrm{L}}\cos(\Omega t) - \frac{i}{2\Omega}(\Delta\lambda\alpha^0_{\mathrm{L}} + \Delta\chi\alpha^0_{\mathrm{S}})\sin(\Omega t)\right], \tag{2.63}$$

$$\alpha_{\mathrm{S}}(t) = e^{i\Sigma t}\left[\alpha^0_{\mathrm{S}}\cos(\Omega t) + \frac{i}{2\Omega}(\Delta\lambda\alpha^0_{\mathrm{S}} - \Delta\chi\alpha^0_{\mathrm{L}})\sin(\Omega t)\right], \tag{2.64}$$

where the abbreviations

$$\Omega \equiv \frac{1}{2}\sqrt{\Delta\lambda^2 + \Delta\chi^2}, \quad \Sigma \equiv \frac{1}{2}(\lambda_{\mathrm{L}} + \lambda_{\mathrm{S}} + \chi + \overline{\chi}), \quad \Delta\lambda \equiv \lambda_{\mathrm{L}} - \lambda_{\mathrm{S}} = \Delta m - \frac{i}{2}\Delta\Gamma \tag{2.65}$$

have been used, with $\Omega \in \mathbb{C}$. The kaon amplitudes at $t = 0$ are α^0_{L} and α^0_{S} and are given by the inverse of Equation (2.59), assuming that K^0 and $\overline{K}^0$ become K^0_{S} and K^0_{L} with equal probability. These amplitudes can now be inserted into the decay rate in Equation (2.27) to redefine the amplitudes $S_{\psi K^0_{\mathrm{S}}}$ and $C_{\psi K^0_{\mathrm{S}}}$:

$$\begin{aligned}\langle\psi K_{\mathrm{L,S}}|H|B^0(t)\rangle = \Big(&g_+(t)\left[\alpha_{\mathrm{L}}(t_K)\,\langle\psi K^0_{\mathrm{L}}|H|B^0\rangle + \alpha_{\mathrm{S}}(t_K)\,\langle\psi K^0_{\mathrm{S}}|H|B^0\rangle\right]\\ &+\frac{q}{p}g_-(t)\left[\overline{\alpha_{\mathrm{L}}}(t_K)\,\langle\psi K^0_{\mathrm{L}}|H|\overline{B}^0\rangle + \overline{\alpha_{\mathrm{S}}}(t_K)\,\langle\psi K^0_{\mathrm{S}}|H|\overline{B}^0\rangle\right]\Big),\end{aligned} \tag{2.66}$$

where the amplitudes $\alpha_{\mathrm{S,L}}$ and $\overline{\alpha_{\mathrm{S,L}}}$ belong to initial K^0 or $\overline{K}^0$ states from the flavor-specific B^0 and $\overline{B}^0$ decays and t_K is the decay time of the neutral kaon, whereas t is the B decay time. As only $K \longrightarrow \pi^+\pi^-$ reconstructed decays are of interest, one can substitute quantitatively $\langle\pi^+\pi^-|K^0_{\mathrm{L}}\rangle \mapsto \epsilon_K\,\langle\pi^+\pi^-|K^0_{\mathrm{S}}\rangle$ in good approximation. This substitution reintroduces CP violation in the kaon system through the CP-violation parameter ϵ_K:

$$\begin{aligned}\langle\psi K^0_{\mathrm{S}}(\longrightarrow \pi^+\pi^-)|H|B^0(t)\rangle = |A_{\psi K^0_{\mathrm{S}}}|\Big\{&g_+(t)\left[\alpha_{\mathrm{L}}(t_K)\epsilon_K + \alpha_{\mathrm{S}}(t_K)\right]\\ &+\lambda_{\psi K^0_{\mathrm{S}}}g_-(t)\left[\overline{\alpha_{\mathrm{L}}}(t_K)\epsilon_K + \overline{\alpha_{\mathrm{S}}}(t_K)\right]\Big\}.\end{aligned} \tag{2.67}$$

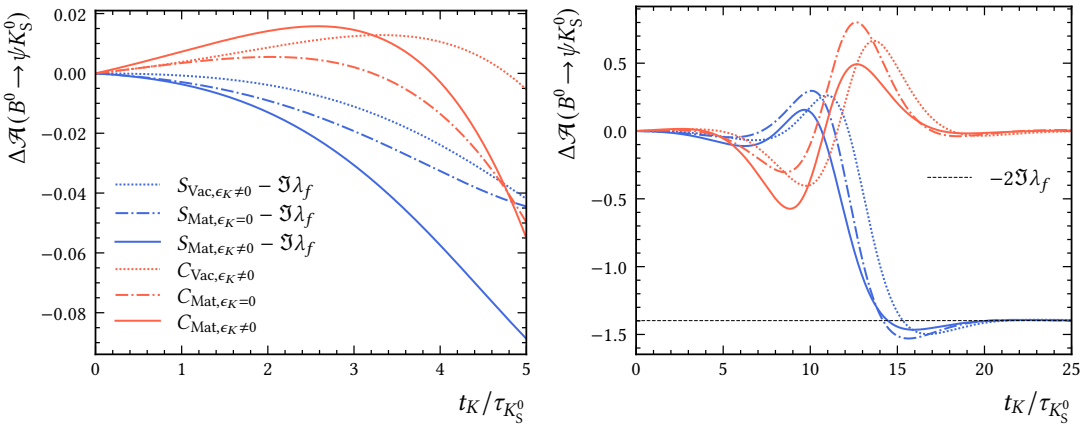

Figure 2.5 – Dependence of the correction for the *CP*-violation observables in the $B^0 \to \psi K^0_S$ channel on the kaon decay time in units of the K^0_S mean lifetime. Three scenarios are considered: Vacuum propagation with *CP* violation (Vac, $\epsilon \neq 0$), and material interaction (Mat) with and without *CP* violation. The plots show two different decay time ranges.

The rest of the derivation proceeds analogously as in Section 2.3, and the modified *CP*-violation parameters are found by comparing the amplitudes of time-dependent terms with and without neutral-kaon influences resulting in

$$\tilde{S}_{\psi K^0_S} = \frac{2\Im\left((\epsilon_K\alpha_L + \alpha_S)^*(\epsilon_K\overline{\alpha_L} + \overline{\alpha_S})\lambda_{\psi K^0_S}\right)}{|\epsilon_K\alpha_L + \alpha_S|^2 + |\epsilon_K\overline{\alpha_L} + \overline{\alpha_S}|^2|\lambda_{\psi K^0_S}|^2}, \tag{2.68}$$

$$\tilde{C}_{\psi K^0_S} = \frac{|\epsilon_K\alpha_L + \alpha_S|^2 - |\epsilon_K\overline{\alpha_L} + \overline{\alpha_S}|^2|\lambda_{\psi K^0_S}|^2}{|\epsilon_K\alpha_L + \alpha_S|^2 + |\epsilon_K\overline{\alpha_L} + \overline{\alpha_S}|^2|\lambda_{\psi K^0_S}|^2}, \tag{2.69}$$

whereby $\lambda_{\psi K^0_S} = \exp(i2\beta)$ can be assumed with satisfactory accuracy for the purpose of determining corrections. To correct for these effects in the present measurement, the corrections $\Delta S \equiv \tilde{S}_{\psi K^0_S} - \Im\lambda_{\psi K^0_S}$ and $\Delta C \equiv \tilde{C}_{\psi K^0_S}$ are defined and subtracted from the measured values and the 2023 PDG central value for $\Im\lambda_{[c\bar{c}]X} = 0.699$ is used that does not include our result [62].

It is worth pointing out that *CP* violation in the kaon system *and* kaon regeneration cause the same qualitative effect for large kaon decay times and are not conditional on each other, as illustrated in Figure 2.5. With $\epsilon = 0$, regeneration alone causes an enhancement of $K_1 \to \pi^+\pi^-$ decays from K_2 states, without additional *CP* violating $K^0_L \to \pi^+\pi^-$ contributions. The strength of regeneration is determined by the parameter $\Delta\chi$, which is proportional to the cross-section difference of K^0 and $\overline{K}^0$ in matter. This parameter can be measured as it is related to the forward scattering amplitude of kaon beams in matter through the optical theorem [78]; a relationship that is used in Section 6.4 to estimate the size of the needed corrections. Figure 2.5 shows the size of the corrections to the *CP*-violation parameters for a homogenous medium of aluminium with a density of 2 % of its usual value, which approximates the LHCb

detector material. For very long decay times, the correction for $S_{\psi K^0_S}$ approaches $-2\Im\lambda$ (compare to Equation (2.54)), which implies that for very long decay times K^0_L decays dominate. This is caused by the fact that K^0_S mesons from $B^0 \to \psi K^0_S$ decays happen after a comparably short time, after which only the rare decays

$$K^0_L \xrightarrow{\epsilon_K} \pi^+\pi^- \qquad \text{or} \qquad K^0_L \xrightarrow{\text{Reg}} K^0_S \to \pi^+\pi^-, \tag{2.70}$$

can result in two-pion decays, the latter being enabled by regeneration. The inverse regeneration effect $K^0_S \xrightarrow{\text{Reg}} K^0_L \xrightarrow{\epsilon_K} \pi^+\pi^-$ is allowed and accounted for in Equations (2.68) and (2.69), but heavily suppressed in addition by ϵ_K.

Authors contribution

In experimental high-energy particle physics, measurements are often conducted by large teams of researchers and take considerable time due to the multitude of standard and measurement-specific validations performed to ensure the quality of the obtained results. This is especially true in the case of measurements of fundamental parameters of the SM. Regarding analysis team size, the measurements described in this thesis differ from the norm, as they have been performed by only three doctoral students, Patrick Mackowiak, Gerwin Meier, and myself from 2017 to 2023, under the supervision of Prof. Dr. Bernhard Spaan† and later by Prof. Dr. Johannes Albrecht. The analysis started as my master's thesis and was supervised by P. Mackowiak and the contact authors of Ref. [14], V. Müller and R. Niet. The measurement strategy was outlined during this time for the $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ decay channel and was later extended to include the $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$ channel.

In 2019, the analysis was joined by G. Meier, who implemented the analysis of the $B^0 \to J/\psi(\to e^+e^-)K^0_S$ mode and worked together with me on validating, optimizing, and documenting the complete analysis of all measurement channels until its completion and approval. G. Meier continued and extended the time-dependent studies of P. Mackowiak, took responsibility for calculating systematic uncertainties and conducted numerous validation studies. Prof. Dr. Wouter Hulsbergen (Nikhef) contributed to the end phase of the analysis with valuable advice and an independent measurement approach that reaffirmed the suitability of the chosen method.

I was responsible for implementing the analysis framework from the very beginning, and I contributed to all major parts of the analysis. The implementation of the flavor tagging procedure for this measurement, the calculation of CP-violation corrections due to kaon physics, and the combination of results was performed exclusively by myself. Furthermore, the BDT-based candidate selection was implemented by myself early on and was later adjusted by my colleagues. The definition of background vetos and the implementation of the initial CP-fitter was done by P. Mackowiak and myself in equal parts, and this model was significantly improved in collaboration with G. Meier. The computation of systematic uncertainties related to flavor tagging were conceptualized by G. Meier and myself and were implemented in collaboration. Most consistency tests mentioned in Section 7.2 were performed by myself. During the final stages of its development, I remained the only contact author of this measurement and defended the measurement in the final part of internal LHCb review. In addition, I have conducted the journal correspondence with Physical Review Letters in coordination with the LHCb collaboration. Finally, I have contributed to the development of a flavor tagging algorithm that uses full-event-interpretation techniques. This work is not detailed in this thesis as it is still ongoing.

3 The LHCb detector at the LHC

The Large Hadron Collider (LHC) [80] is a particle accelerator at the European Organization for Nuclear Research (CERN). The LHC is a proton synchrotron that brings oppositely directed particle beams to collision at four fixed interaction points. Particle detectors surround the collision positions and measure the interactions of particles created in proton collisions. Before proton beams are injected into the LHC in order to be accelerated to their terminal collision energy, multiple pre-accelerators increase the proton beam energy in succession. The accelerators comprising today's accelerator complex have been constructed over the course of CERN's history. Each has been used to collide particles before being superseded by a larger accelerator and converted into a specialized accelerator and injector. A schematic of CERN's experimental facilities as of 2018 is shown in Figure 3.1. In the following, the accelerator setup for the second data-taking phase of the LHC, referred to as *Run 2* (2015-2018), is summarized.

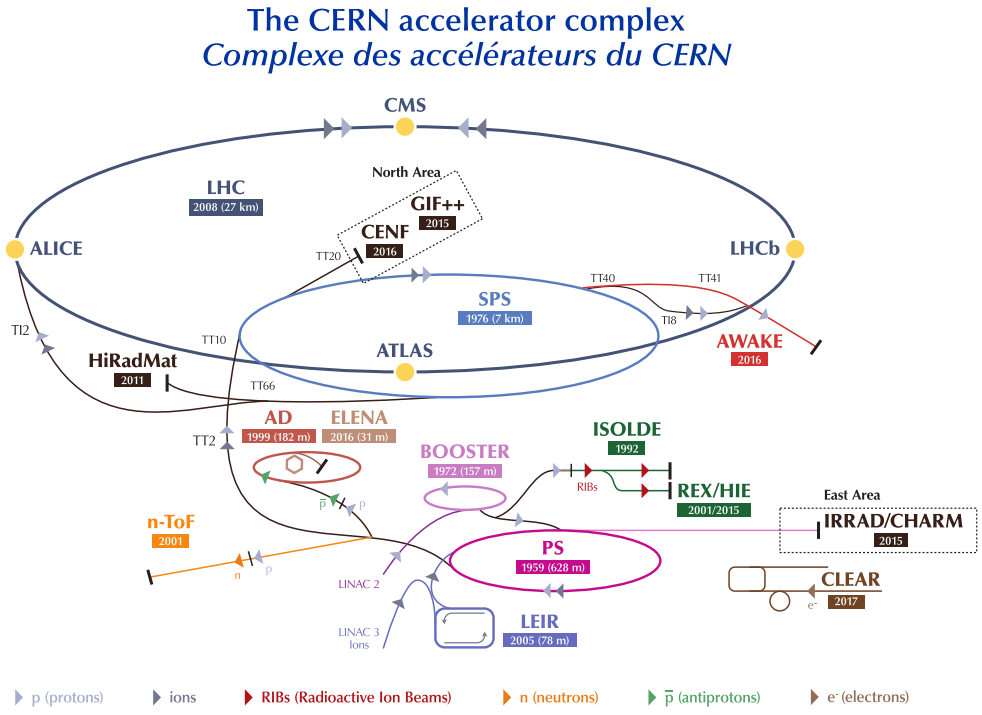

Figure 3.1 – CERN accelerator complex and dimensions as of 2018 [81] (modified).

Protons are obtained by ionization of hydrogen gas and are then accelerated from rest in the Linac2 linear accelerator [82], which injects the beam into a sequence of

proton synchrotrons. The Proton Synchrotron Booster is the first circular accelerator in this sequence and increases the beam energy from the final Linac2 beam energy of 160 MeV to 2 GeV. The Proton Synchrotron then accelerates the beams to 26 GeV and injects the beams into the Super Proton Synchrotron, which increases the proton beam energy to 450 GeV. Finally, the proton beams enter the LHC, which has a circumference of 26.7 km and increases the beam energy to 6.5 TeV per beam during the Run 2 data taking phase [83], corresponding to a proton-proton (pp) collision energy of 13 TeV. A proton beam consists of discrete proton packets, or *bunches*, each containing $O(10^{11})$ protons [80]. Proton collisions are induced at the beam crossing positions through an adjustment of LHC control magnets, after which proton bunches collide with a rate of about 40 MHz [80]. In each bunch collision, multiple protons can collide or scatter off another, and the totality of particles created in these collisions constitutes an *event*.

Four major particle detectors with different technical designs and capabilities are located at the LHC accelerator ring and measure the particles created in high-energetic bunch-collisions. The ATLAS [84] and CMS [85] detectors are general-purpose detectors designed to measure pp collisions at high rates and high transverse momenta and approximately cover the full solid angle around their respective interaction points. The ALICE [86] detector is specialized for measuring collisions of heavy ions in order to gain a better understanding of the properties of the quark-gluon plasma. The LHCb detector [87] is specialized for precision measurements of bottom and charm-hadron decays, which are predominantly created in momentum-asymmetric parton collisions. For this reason, the created particles are significantly Lorentz-boosted along the beam axis. The analyzed data for this measurement is recorded and reconstructed with the LHCb-detector, which is introduced in the following section.

3.1 The LHCb detector

The LHCb detector [87, 88] is specialized for precision measurements of bottom and charm-hadron decays. Bottom and charm-mesons are mainly created from collisions of two incoming gluons or quark-antiquark pairs from each of the colliding protons. As the momentum difference of these partons is more likely to be asymmetric, and their momentum is considerably smaller in comparison to the total center-of-mass energy, the created $b\bar{b}$ or $c\bar{c}$ pair is strongly boosted along the proton beam axis. For this reason, the LHCb detector is instrumented in the forward pseudorapidity (η) range $2 < \eta < 5$ and covers one principal direction. Pseudorapidity is a non-linear function of the angle θ between a track and the beam direction, $\eta \equiv -\log(\tan(\theta/2))$, and can be related to *rapidity* (y), an additive measure of velocity under Lorentz transformations. The proton beam parameters are controlled to limit the instantaneous luminosity in the LHCb detector to correspond to an average of 1.1 proton collisions per bunch crossing during Run 2 [89]. This proton collision rate allows for full online reconstruction and identification of individual tracks at a bunch-crossing frequency of 40 MHz [88] and provides a constant collision environment over the duration of an LHC fill. A schematic of the LHCb detector with labeled subdetector components is shown in Figure 3.2. After

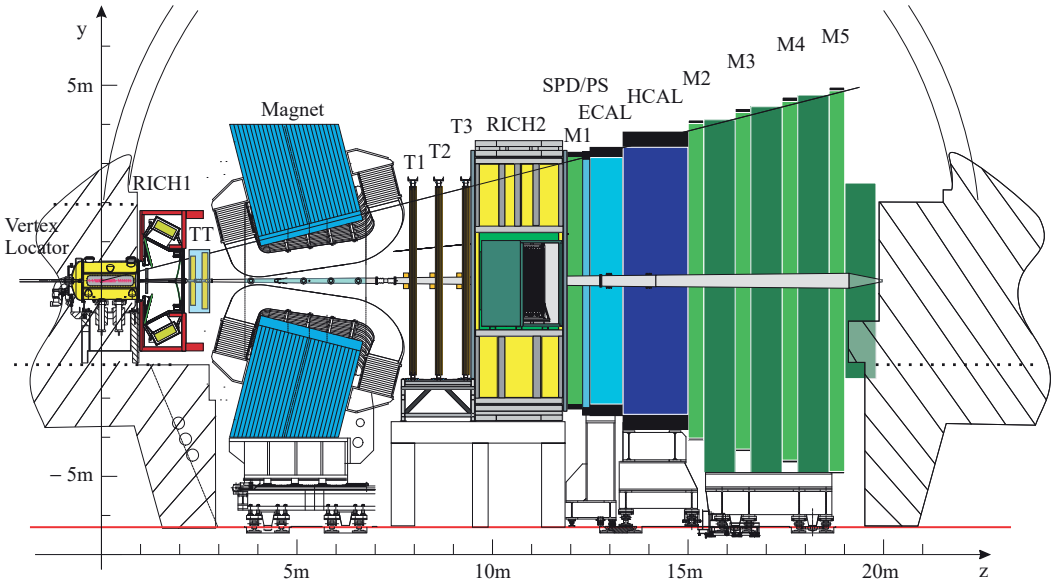

Figure 3.2 – Side-view of the LHCb detector with labelled subdetector components [91].

the successful Run 2 data-taking phase (2015-2018), the LHCb detector underwent an extensive hardware and software upgrade. Currently, the upgraded LHCb detector is taking data for the Run 3 data taking phase, which is presently ongoing and is scheduled to last until the end of 2025 [90]. In the following, the focus is on the Run 2 LHCb detector, as this is the detector that recorded the analyzed data. The LHCb detector consists of several subdetectors, which can be grouped into tracking and particle identification detectors. Nevertheless, the overlap of contexts in which individual subdetector-readouts are used is quite significant, and one should therefore keep in mind that the reconstruction of events always relies on the combined interpretation of information contributed by individual sudetectors.

3.1.1 Vertex Locator

The oppositely directed proton beams of the LHC are brought to collision inside the Vertex Locator (VELO), which measures the created primary charged tracks and reconstructs the primary interaction vertices (PV) from these measurements. A detailed description of the VELO detector and its design constraints are given in Ref. [92]. Due to the long lifetimes of several b and d hadrons, and their significant Lorentz boosts, these particles can traverse several millimeters in the VELO before decaying at secondary vertices (SV). The VELO reconstructs and matches the SV to one of the measured PVs with high precision. The accurate positional measurements of the PV and SV allow for precise lifetime reconstructions, which are necessary for all time-dependent measurements, including measurements of time-dependent CP violation. Secondary vertex identification is a precondition for further analysis of an event, and the VELO is therefore a crucial component of the LHCb trigger system (Section 3.1.5).

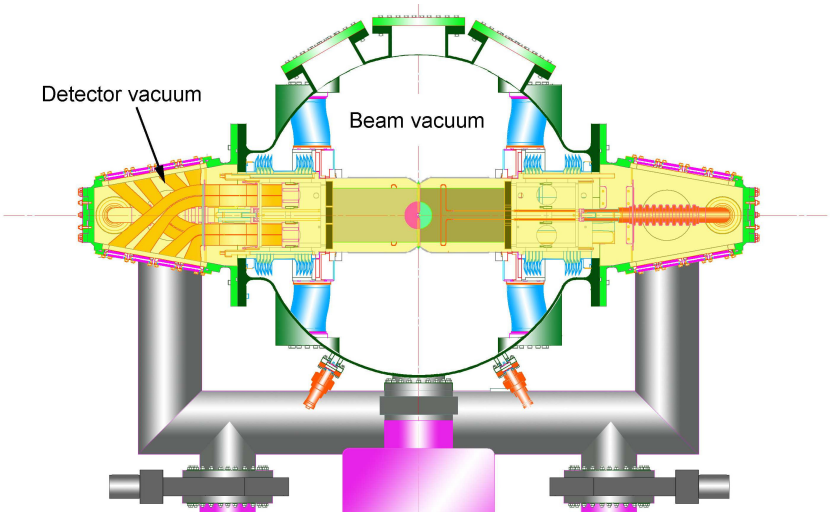

Figure 3.3 – Cross-section of the Vertex Locator orthogonal to the beam direction showing its inner and outer vacuum vessels (Modified image from Ref. [87]). The VELO is shown in the closed configuration. The left and right detector halves that are part of the inner assembly (shaded in yellow) can be separated and moved away from the beam line.

The VELO is a tracking detector that consists of two symmetric, movable halves of arrays of silicon pixel detectors. Each array consists of 21 modules of silicon strip detectors providing measurements in polar coordinates r and φ around the beam direction. The inner VELO detector is contained in an aluminium pressure vessel, referred to as radio-frequency-shielding foil (RF-foil), which shields the silicon detectors and electronics from the LHC RF and separates the inner beam-vacuum chamber from the vacuum of the outer VELO detector vessel. The distance of VELO modules to the beam can be controlled, and during data taking, the VELO operates in a closed configuration, where the distance of the pixel detector layers to the beam is at a minimum of 7 mm [93]. In Figure 3.3, a VELO cross-section is shown to illustrate its construction. To protect the VELO, this configuration is set up only after the LHC beams are reported to be stable. The stability of the proton beams is continuously monitored by LHCb's *Beam Conditions Monitor* [94], which initiates a beam dump in case of adverse beam conditions or a progressive loss of beam stability. For the reconstruction of tracks, the VELO plays a crucial role, as it reconstructs track seeds, which are extrapolated through the detector in subsequent reconstruction steps to match the measured event. A track seed thereby needs to be measured with a sufficient amount of hits. Due to the VELO detector geometry, tracks with either very high or very low pseudorapidity tracks are more likely not to be identified by the VELO but can be reconstructed by other tracking detectors. The VELO has a track reconstruction efficiency of at least 98 % inside the LHCb acceptance, the relevant momentum range and track multiplicity [93]. The high segmentation density of the silicon-microstrip detectors allows for a PV reconstruction with a vertex resolution as small as 71 µm in the beam direction and 13 µm in the direction transversal to it. A precise measurement

of the PV and SV is an essential requirement for time-dependent measurements of B and D mesons. In this measurement, a mean decay-time resolution of 60 fs is found for B^0 mesons, which is two orders of magnitude smaller than the B^0 oscillation period.

3.1.2 Tracking system

Apart from the VELO, the LHCb detector is equipped with further dedicated tracking detectors that measure the hit positions of charged particle tracks as they propagate through the detector layers. Four silicon-microstrip trackers of similar design are in use and instrument the region near the beam pipe. The tracker turicensis (TT) [95] is located downstream of RICH1, and three silicon-microstrip detectors comprise the inner tracker (IT) [95]. The IT stations are located downstream of the magnet and instrument the region close to the beam pipe where the particle flux is high. To ensure a low detector occupancy at a particle flux of up to $5 \times 10^5\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [95], the silicon trackers have high spatial hit resolutions ranging from 48 to 55 µm [88] as measured with Run 1 data. The outer tracker (OT) [96] is a straw-tube detector and consists of three stations, each of which is made up of four layers of straw-tube modules. The hit position is calculated from coincidental straw-tube signals and is corrected by taking the electron drift time into account. The OT covers the remaining 98 % of the combined OT and IT tracker surface area [96], and its hit resolution has been improved for the Run 2 data-taking phase and is at 171 µm [97]. Collectively, the three tracking units of one IT and OT module each are referred to as tracking stations T1-T3. The individual hit efficiencies of TT and the IT stations each exceed 99 % in all scenarios, and the overall track-finding efficiency of all tracking detectors, including the muon chambers, is measured to be above 96 % in LHCb's phase space.

Tracks of charged particles are reconstructed in LHCb's high-level trigger stage (Section 3.1.5), which utilizes hit information of all tracking detectors by matching tracks reconstructed in T1-T3 with track fragments from the VELO, taking into account the intermediate magnetic field of known spatial field density. From the reconstructed track curvature and the known magnetic field, the particle momentum is calculated and the sign of the track curvature determines the particle's charge sign. Charged particles are deflected by LHCb's dipole magnet [98], which has an integrated magnetic field strength along the center line of 10 T m. The momentum resolution provided by the tracking system is estimated to range from $\delta p/p \approx 0.35$ to 0.55 % in simulation [87, 97], depending on the total track momentum and reconstruction circumstances for the pion tracks of the decay $B^0 \to J/\psi K^0_S (\to \pi^+\pi^-)$ in particular.

The muon identification system consists of five straw-tube muon chambers (M1-M5) [99]. Muon chambers M2-M5 are located downstream of the calorimeters, while M1 is placed upstream of the calorimeters. The downstream muon chambers are separated by thick iron absorber layers to reduce the background from charged particles that escape the HCAL. In the collision environment at LHCb, muons are minimum ionizing particles and therefore do not lose significant fractions of their initial kinetic energy through detector material interactions. A muon candidate is required to produce a signal in a sufficient number of muon stations, for which a momentum of at least

3 GeV/c is required, as otherwise its path would be too deflected by the magnet [89]. Muons are identified with an efficiency exceeding 95 % for muon momenta exceeding 3 GeV/c [88]. The identification of muons is of central importance for the measurement presented here, as for the selection of the main-statistics channel, and for a precise measurement of the B^0 decay time, the precise reconstruction of $J/\psi \to \mu^+\mu^-$ candidates is required. For the decay $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ in particular, the muon reconstruction efficiency is measured to be 94 % [87].

3.1.3 Ring Imaging Cherenkov detectors

The LHCb detector is equipped with two Ring Imaging Cherenkov detectors (RICH) which measure the velocities of charged particles. A detailed description of the RICH system and its performance can be found in Refs. [100, 101]. The functioning principle of the RICH system is the Čerenkov effect [102] which occurs when a charged particle traverses a medium at a higher velocity than the phase velocity of light in that medium, $c_n = c/n$, whereby n is the medium-specific refractive index. The Čerenkov effect causes electromagnetic radiation to be emitted in a cone with an opening angle θ, which is determined by the particle velocity in the direction of the particle motion. The relation of particle velocity v, refractive index, and cone opening angle is given by $\cos\theta = c_n/v$.

To maximize the covered phase space of tracks, defined by the track angle to the beam and its momentum, two RICH systems are in place at different positions and make use of different fill media. In the RICH1 detector, a mixture of C_4F_{10} gas and aerogel is utilized, and the detector is located immediately downstream of the VELO and is particularly sensitive to particles in the momentum range $1 < p < 60\,\text{GeV}/c$. The RICH2 detector identifies tracks in the momentum range $15 < p < 100\,\text{GeV}/c$ and beyond [87]. It is filled with CF_4 and is located downstream of the tracking stations. The light produced by charged particles traversing the fill media is reflected by internal spherical mirrors onto two arrays of photomultipliers. The detection patterns projected onto the photomultipliers are circular due to the velocity-dependent opening angle, θ, of the Čerenkov light cone. From a reconstruction of overlapping circles, each track's Čerenkov angle is reconstructed, and with an independent momentum measurement supplied by the tracking system, the particle velocity can be determined for a particle mass hypothesis that best fits the data. The relation between particle momentum and Čerenkov radiation angle for particles of different masses is shown in Figure 3.4 from measurements of the RICH system. Considered particle identities are pions, kaons, protons, and muons, as these particles rarely decay in the LHCb detector acceptance due to their high Lorentz boosts and long lifetimes.

As the identification of particle mass determines the particle identity, the RICH system provides crucial particle identification information for most LHCb measurements, including the measurement presented here. From the information provided to a large extent by the RICH system, particle identification probabilities are calculated with multivariate techniques in addition. These features are frequently used in signal candidate selections (see Sections 4.3.1 and 4.3.2). While the RICH system can identify muons,

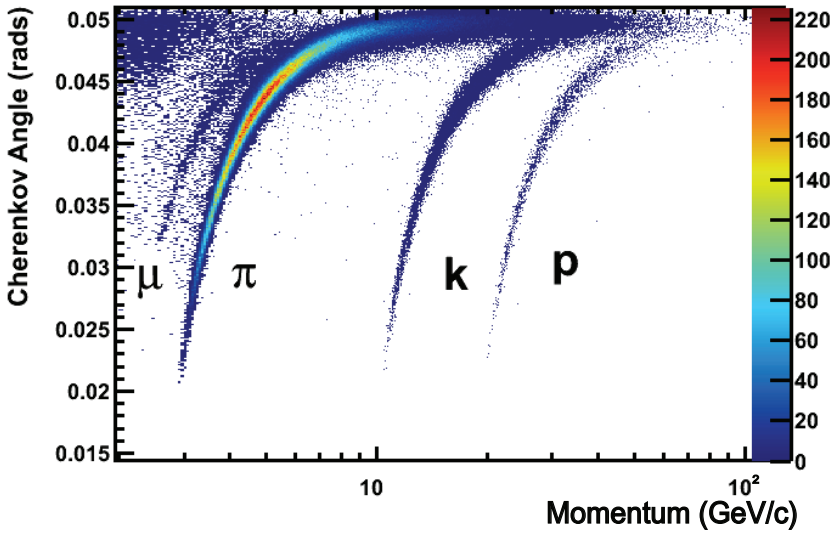

Figure 3.4 – Relation of the particle momentum and Čerenkov-radiation angle as measured by the LHCb RICH system for charged particles of different mass [101]. Charged particles of different masses are clearly separable in this plane.

the muon particle identity is primarily determined from a suitable track measurement in the muon chambers.

3.1.4 Calorimetry system

A calorimetry system is used to measure the particle energy. This is achieved by causing incoming particles to deposit their total kinetic energy in the dense calorimeter material. Incoming particles cause the formation of particle cascades, called *showers*, which are ideally fully contained in the calorimeter volume. The initial particle energy is then inferred from the properties of the measured particle shower. Electromagnetic showers are characterized by pair production and bremsstrahlung processes, while hadronic showers include intermediate hadrons and their interactions with the detector material. At LHCb, an electromagnetic calorimeter (ECAL) is followed by a hadronic calorimeter (HCAL) further downstream [103] and in addition, upstream of the ECAL, a scintillating pad detector, and a preshower (SPD/PS) system is in place to improve the particle identification information obtained with the calorimeter system. The ECAL measures the energy of electrons and photons, and the HCAL measures the energy of hadrons.

The same detector design principle is in use for the ECAL and HCAL: the calorimeter volumes are segmented into independent cells, which are made up of alternating layers of material with high mass-density and intermediate scintillation layers. In the ECAL, lead is used as an absorber, while in the HCAL, iron is used. An energy calibration allows inferring the particle energy from the integrated light yield of each calorimeter cell.

The SPD and PS are two similarly constructed detector layers of scintillating fibers separated by a lead plate. The SPD/PS system helps with separating photons from electrons as only electrons cause the emission of scintillation light in the SPD while photons are only detected in the PS layer as the separating lead plate causes an electromagnetic shower to form.

Since particle momenta and velocities are measured independently by dedicated LHCb subdetector systems with high precision, the final-state particle energies can essentially be derived without an independent energy measurement from a calorimeter. At LHCb, the calorimeter system is used in multiple other ways in addition: The energy measurement of the calorimeter is used for the hardware trigger stage for the initial selection of high-E_{T} events. Furthermore, it improves particle identification; most notably, it allows better differentiation between electrons, neutral pions, and photons. Finally, the measured shower position is utilized by tracking algorithms to improve track finding. In the measurement described here, the $B^0 \to J/\psi(\to e^+e^-)K^0_{\mathrm{S}}$ decay is analyzed for which the identification of high-energetic electrons is required. For this reason, the calorimeter plays an important role in this measurement, as it allows for the reconstruction of electrons and can distinguish these candidates from π^0's and photons. In Ref. [87], the electron identification efficiency for $J/\psi \to e^+e^-$ candidates is measured to be as high as 95 % in the decay $B^0 \to J/\psi(\to ee)K^0_{\mathrm{S}}$. Furthermore, as electrons are deflected by the magnet, a significant amount of kinetic energy can be lost through bremsstrahlung, which is detected in the ECAL in the form of electromagnetic showers, which are located in the bending plane of the electron track. The measured energy of these photon candidates is added to the energy of the initial electron. The read-out of all subdetectors previously mentioned is processed in several stages and combined to complete B^0 decays with LHCb's trigger system, which is explained in the following section.

3.1.5 Trigger system and event reconstruction

The bunch-crossing rate at the LHC is 40 MHz, and the rate of inelastic proton collisions of interest with sufficiently high transverse momentum is low in comparison. Furthermore, the final detector readout frequency is limited by the hardware and at 1.1 MHz [89]. To only select events of interest and thus reduce the measurement rate to a level at which events can be read out, analyzed, and recorded, a three-stage trigger system is in place during the Run 2 data-taking phase. This system accepts only a fraction of collision events for permanent storage. The trigger system is able to mirror the offline event reconstruction and thereby identify events of interest with detailed selections. Events not rejected by the trigger system are permanently saved to disk in the form of the raw detector output for several classes of events, including the $\psi \to \ell^+\ell^-$ data sets analyzed here. Identical trigger algorithms are used for the offline reconstruction, which is performed at a later point in time from the recorded detector readout. The benefit of storing raw detector read-out is that reconstruction algorithms can be optimized and applied to the same data. The initial trigger decisions

are, however, permanent. The following explanation of the LHCb trigger is based on Ref. [89].

The first trigger stage (L0) is a fast hardware trigger that selects events with high transverse energy by summing the transverse energies deposited in the ECAL and HCAL cells. Alternatively, muon candidates reconstructed in the muon system alone can cause the L0 trigger to accept an event if the transverse momenta of the two highest-p_{T} muons are sufficiently high. Events with exceedingly high detector occupancy are rejected by L0 based on hit information in the SPD and PS. The rate of events accepted by L0 is 1 MHz.

The two subsequent *high-level-trigger* (HLT) stages HLT1 and HLT2 are software triggers and apply a full event reconstruction for each event accepted by L0. The HLT1 stage performs a partial event reconstruction in which single tracks are reconstructed, a set of PVs is identified, and a displaced vertex is searched for. The track reconstruction begins with single reconstructed tracks starting in the VELO from which a set of PVs is reconstructed. Track measurements of the VELO and TT are then combined for an initial track momentum estimate used for the full track fit for tracks with transverse momentum exceeding 500 MeV/c. Tracks below this momentum threshold are considered as additional muon candidates and can be reconstructed with p_{T} as low as 80 MeV/c. The momentum muons need to have in order to reach the muon stations is at least 3 GeV/c for which hits in M1 and M2 are required. With increasing momentum,

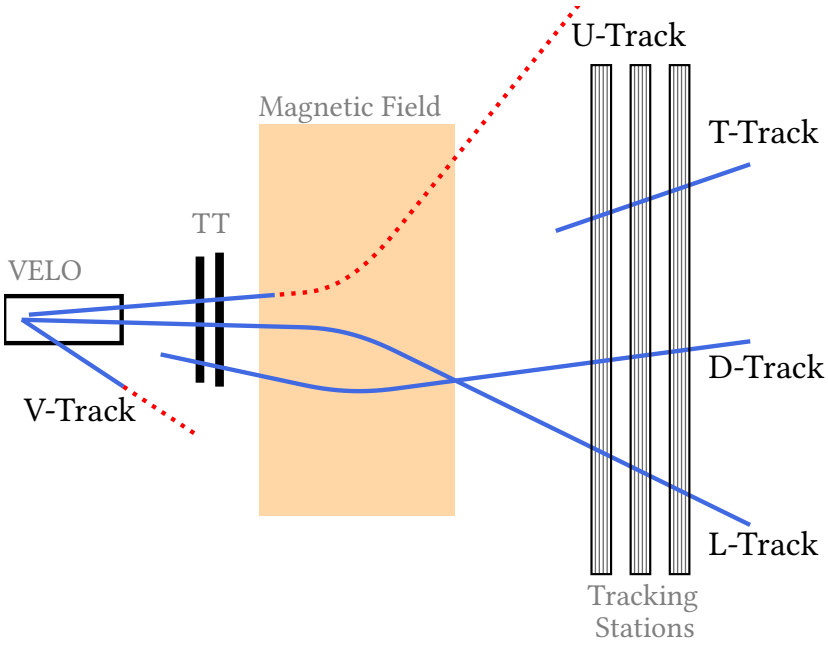

Figure 3.5 – Illustration of track reconstruction categories that are assigned to tracks which are measured by different LHCb tracking-subdetectors. Only the subdetectors relevant to the classification of tracks are shown. The bending plane of the magnet is in the horizontal direction, and the dotted lines mark missing track segments.

hits in muon chambers further downstream are required in addition. Events containing either one track with displaced origin vertex or two tracks originating at the same vertex displaced from the PV are accepted by HLT1 as those are possible B or D meson candidates. The rate of events passed to the HLT2 that fulfill this requirement and have a successful HLT1 reconstruction is 110 kHz. Information gathered from the HLT1 stage is used to align the subdetector components in real-time. The real-time alignment is performed in frequent intervals using recorded buffered data and is especially important for the VELO as its components that are closest to the beam move for each proton fill [104]. Furthermore, the downstream-tracker positioning is affected by the magnetic field polarity, and the RICH detectors are calibrated to account for changes in the refractive index of the fill media due to changes in atmospheric conditions.

The HLT2 trigger stage repeats the PV identification and the single track reconstruction using lower momentum thresholds before proceeding with the selections of a large number of decays and has an output rate of 12.5 kHz, which is permanently recorded. Reconstructed particle tracks are classified into different categories, depending on which sub-detectors measured it. In Figure 3.5, the mentioned track type categories of relevance are illustrated. A reconstruction of *long tracks* (L) is performed, which are measured by all tracking subdetectors, the VELO, TT, and the tracking stations. To that end, VELO tracks (V) are matched with hits in the TT and extrapolated to the tracking stations in a similar procedure as in the HLT1 stage, whereby additional time is spent to refine the search window of the track finder, and thereby the rate of fake tracks is reduced further. Tracks without a match in the T stations are classified as *upstream tracks* (U) and usually correspond to low-momentum particles that are deflected out of the T station acceptance by the magnet. A similar track fit is performed by matching the track seeds of the T stations (T-tracks) with V-tracks, and the overlap of both procedures results in an additional improvement of the tracking efficiency. Long-lived particles like K_S^0 or Λ traverse a significant distance before decaying, and in a majority of cases, they decay outside of the VELO acceptance, resulting in downstream tracks (D), which are measured by the TT and the T stations. In these cases, their decay products do not produce hits in the VELO and are reconstructed from matching track segments in the T stations and the TT. Muon candidate selections are further improved by utilizing the RICH and calorimeter detector information. Fake tracks are then further reduced by applying a limit on the track reconstruction χ^2 value and by the use of a multivariate classifier with high signal efficiency and a background rejection rate of 60 %.

The reconstructed tracks are subsequently assembled in stages to complete events through multiple selections, called *trigger lines*. A trigger line selects a set of tracks and can apply selections to these tracks. At first, pairs of tracks emerging from the same decay vertex that commonly occur are considered. Thresholds are applied on the track distance and reconstruction precision before these intermediate decays, and hypothesized origin particles are reused in trigger lines that select complete decays of B and D mesons.

In the centralized offline selection, which mirrors the HLT reconstructions, events are classified based on distinctive signatures, like the presence of two high momentum

muons, and the trigger decisions for each event are saved. Trigger line decisions are used in physics analyses for the candidate event assembly, which simplifies comparing different measurements. Selection requirements that trigger decisions are based on can be tightened and, in some cases, relaxed so far as the recorded data permits it.

3.2 Simulation of collision events at LHCb

For the development and validation of most LHCb measurements, simulated data, or *Monte Carlo* data of specific decays, are generated and reconstructed in the same way as real data. The simulation of events encompasses all processes beginning at the initial parton interaction and ends with the simulation of the LHCb-detector response, after which the simulated data is selected with identical algorithms as real data. At the LHCb experiment, the Gauss software project [105] sequences all generation steps, which are briefly mentioned here.

The initial hard parton interactions and subsequent hadronizations are simulated with the event-generator Pythia [106, 107], for a collision energy of $\sqrt{s} = 13\,\text{TeV}$ and with an LHCb-specific configuration [108]. The momenta of incoming partons are sampled from specific parton distribution functions, which are widely used at the LHC experiments [109]. The initial interaction results in a set of high-energetic outgoing partons whose strong-force interactions are simulated while preserving the color charge. After a sufficient number of interactions, which result in a certain number of hadronic jets containing any number of partons in their substructure, hadronization is simulated. During hadronization, quarks recombine to form colorless hadrons. The momentum fraction each hadron receives is thereby determined by the Lund-String hadronization model [110].

Decays of initial hadrons are handled with EvtGen [111], which includes the decay amplitudes for a wide variety of particles that account for angular correlations of spin and momenta. EvtGen can be configured to decay certain particles to chosen final states, and furthermore, neutral meson mixing and *CP*-violation are simulated in this step. *CP* violation in the neutral kaon system is usually omitted in the simulation due to its small size. Final state radiation is simulated with Photos [112] and in the sub-decay $\psi \to \ell^+\ell^-$, such corrections are applied to the lepton candidates.

All generated tracks are propagated through a model of the LHCb-detector volumes of known material compositions, and the detector material interactions are simulated with the Geant4 [113, 114] toolkit. The final step of the event generation is the simulation of the detector response and signal digitization, which is handled with the Boole [115] software package. The output of Boole is a set of electronic signals corresponding to the readout of each subdetector, as it would be measured in real collision events. This signal is thereafter passed through to the simulated L0 trigger and reconstructed in the high-level-trigger software. As the data-taking conditions were not constant during Run 2 due to smaller detector adjustments, simulated data sets are generally produced for each year of data-taking.

4 Selection of $B^0 \to \psi K^0_S$ decays

4.1 Measurement strategy

The aim of this measurement is to determine the CP-violation parameters $S_{\psi K^0_S}$ and $C_{\psi K^0_S}$ in the three decay channels $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$, $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$ and $B^0 \to J/\psi(\to e^+e^-)K^0_S$. The neutral meson K^0_S is reconstructed from decay to two oppositely charged pions in all decays. The parameters $S_{\psi K^0_S}$ and $C_{\psi K^0_S}$ are the amplitudes of the CP asymmetry, which is the periodic change of the asymmetry of B^0 and $\bar{B}^0$ decays to the same final state, and it is defined as follows:

$$\begin{aligned}\mathcal{A}^{CP}(t) &= \frac{\Gamma(\bar{B}^0(t) \to \psi K^0_S) - \Gamma(B^0(t) \to \psi K^0_S)}{\Gamma(\bar{B}^0(t) \to \psi K^0_S) + \Gamma(B^0(t) \to \psi K^0_S)} \\ &= \frac{S_{\psi K^0_S}\sin(\Delta m_d t) - C_{\psi K^0_S}\cos(\Delta m_d t)}{\cosh\left(\frac{1}{2}\Delta\Gamma_d t\right) + A^{\Delta\Gamma}\sinh\left(\frac{1}{2}\Delta\Gamma_d t\right)}.\end{aligned} \tag{4.1}$$

The notation $\Gamma(B^0(t) \to \psi K^0_S)$ describes the time-dependent decay rate of B mesons produced in the state B^0 to the final state ψK^0_S with time. This time dependence is caused by the varying relative concentrations of B^0 and $\bar{B}^0$ decays caused by mixing, which is a time-dependent process. An asymmetry of decay rates is caused by CP-violation in this decay mode. In the baseline measurement, the assumption $\Delta\Gamma_d = 0$ is made as the parameter is compatible with zero to our current knowledge [62], which implies that this measurement is not sensitive to $A^{\Delta\Gamma}$. The asymmetry then simplifies to

$$\mathcal{A}^{CP}(t) = S_{\psi K^0_S}\sin(\Delta m_d t) - C_{\psi K^0_S}\cos(\Delta m_d t), \tag{4.2}$$

with $\Delta m_d = (0.5065 \pm 0.0019)\,\mathrm{ps}^{-1}$ [62]. As motivated in Section 2.4, the CP-violation parameter $C_{\psi K^0_S}$ is small, and with the available sample size, no significant deviation from zero should be expected, which is verified by measuring it. It is evident from Equation (4.1) that the number of reconstructed decays needs to be as large as possible to minimize the statistical uncertainty and that the production flavor of B mesons needs to be identified, which is accomplished by the flavor tagging technique (Chapter 5).

Each step of the analysis is associated with a finite measurement precision, which affects the observed amplitude in various ways. Any measurement uncertainty that leads to a worse separation of B^0 and $\bar{B}^0$ mesons causes a decrease of the observed CP-asymmetry amplitude, which is quantified by dilution factors $\mathcal{D}_j$. To achieve an unbiased measurement of $S_{\psi K^0_S}$, the dilution factors need to be understood and measured

accurately. The total experimental influences on the measurement can be summarized in a simplified manner:

$$\begin{aligned} A_{CP}^{\text{obs}}(t) &= \mathcal{D}_{\text{FT}} \times \mathcal{D}_t \times \mathcal{A}^{CP}(t) + \mathcal{I}_{\Delta\text{FT}} \\ &\approx (1-2\omega) \times \exp\left(-\frac{1}{2}\Delta m_d^2 \sigma_t^2\right) \times \mathcal{A}^{CP}(t) + \mathcal{I}_{\Delta\text{FT}}. \end{aligned} \tag{4.3}$$

Here, $A_{CP}^{\text{obs}}(t)$ is the observed CP asymmetry, and $\mathcal{A}^{CP}(t)$ is the original asymmetry that is independent of experimental influences. The flavor tagging and decay-time dilutions are $\mathcal{D}_{\text{FT}}$ and $\mathcal{D}_t$, respectively. The mean *mistag*, ω, is a measure of the error probability of the production-flavor assignment, and σ_t is the mean decay-time resolution of the LHCb detector. The offset $\mathcal{I}_{\Delta\text{FT}}$, which is at the permille level, can be caused by efficiency biases in the assignment of production flavors and by a production asymmetry and does not need to be estimated in this measurement as the chosen fit model implicitly includes a description of this offset. Nevertheless, such biases can have an additional small effect on the flavor tagging dilution. This factor and its event-wise representation that takes flavor asymmetries into account is derived in Chapter 5.

The leading-order effect on the asymmetry amplitude is given by the flavor tagging dilution factor $(1-2\omega)$, while decay-time-resolution effects are almost negligible due to the high decay-time resolution of the LHCb detector, which is two orders of magnitude smaller than the B oscillation period. The data-selection part of the measurement is largely independent of the assigned B production flavors B^0 and $\bar{B}^0$ as their final states are identical. Signal B mesons whose production flavor has not been identified do not influence the measurements. For this reason, if not explicitly stated, candidate yields and efficiencies reported for the signal modes are related to the subset of data for which estimated production flavors (*tagging decisions*) are available. Data with available tagging decisions is referred to as *tagged* in the following, and otherwise, it is referred to as *untagged*. The results are obtained from a maximum-likelihood (ML) fit to the B decay-time distribution with a model that is derived in Chapters 5 and 6.

4.2 Reconstruction of $B^0 \to \psi(\to \ell^+\ell^-)K^0_S(\to \pi^+\pi^-)$ decays

For the reconstruction of the $B^0 \to \psi K^0_S$ modes, trigger line selections are chosen that select $J/\psi \to \mu^+\mu^-$ or $\psi(2S) \to \mu^+\mu^-$ decays. These selections require the difference of the reconstructed dimuon mass to the known values of the parent particles J/ψ and $\psi(2S)$ [116] not to exceed $100\,\text{MeV}/c^2$, and they require a good-quality common vertex. The selection of $J/\psi \to e^+e^-$ candidates is more involved as the potential for particle mismatches between pions, photons, and electrons is significantly higher. The strategy used for the selection of these J/ψ candidates is to find a set of trigger lines targeting electrons that result in the highest possible electron track selection efficiency at the cost of an increased level of background contamination in the first selection steps [1].

[1] see Ref. [117] for more details on trigger line selections

Table 4.1 – Fraction of events containing each considered K^0_S track type in each decay mode in a simulated data set for 2016 running conditions.

Candidate fractions [%]			
	$B^0 \to J/\psi K^0_S$	$B^0 \to J/\psi K^0_S$	$B^0 \to \psi(2S)K^0_S$
Track type	$J/\psi \to \mu^+\mu^-$	$J/\psi \to e^+e^-$	$\psi(2S) \to \mu^+\mu^-$
LL	26.83 ± 0.04	26.72 ± 0.05	28.88 ± 0.11
DD	58.80 ± 0.05	58.91 ± 0.05	53.39 ± 0.12
LD	5.215 ± 0.021	5.227 ± 0.024	6.08 ± 0.06
UL	7.821 ± 0.025	7.788 ± 0.029	9.77 ± 0.07
UU	0.581 ± 0.007	0.594 ± 0.008	0.894 ± 0.023
UD	0.753 ± 0.008	0.761 ± 0.010	0.988 ± 0.025

The reconstructed J/ψ mass is required to be in the range $2.3 \leq m(e^+e^-) < 3.3\,\text{GeV}/c^2$ in the dielectron mode.

Multiple trigger lines for the selection of neutral kaons have been developed for the predecessor analysis in Ref. [13]. In this measurement, instead of using existing trigger lines to select kaons from pairs of two L or two D tracks (LL and DD combinations), which is the standard procedure, new selections are defined to maximize the number of kaon candidates. Upstream pion tracks are matched to L tracks to form UL-K^0_S candidates. For this reason, selected pion tracks are assembled to kaons with adjusted requirements that largely replicate the existing selections. For upstream tracks, the minimum momentum requirement is lowered to $1\,\text{GeV}/c$, while the requirement for L and D-track pions is left unchanged at $2\,\text{GeV}/c$. In addition to LL, DD, and UL-K^0_S candidates, the viability of the combination LD is tested, and it is likewise included as a separate category. The pion track combinations UL and LD are used for the first time in a time-dependent measurement at LHCb. These additional combinations are only included for the $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ mode as their total contribution in the other two measurement modes is globally insignificant, and independent validation of measurements is impractical on such limited data subsets. Likewise, the combinations UU and UD are too small for independent validations in any subset and cannot significantly improve this measurement. For this reason, these combinations are discarded. Table 4.1 lists the observed frequencies of K^0_S reconstructions in each decay mode in a simulated data set. The LD and UL categories increase the sample size significantly, and their fraction in the combined $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ sample is 13 %.

Pairs of reconstructed pions are required to have a good-quality common vertex with a reconstructed-mass difference to the known K^0_S mass [116] of at most $80\,\text{MeV}/c^2$ for the dimuon modes, and to be significantly separated from the B^0 decay vertex. For the dielectron mode, this requirement is stricter: $25\,\text{MeV}/c^2$ for the LL and $33\,\text{MeV}/c^2$ for the DD combination. Pion tracks are required to be significantly separated from the PV, which is similar to imposing a flight-distance-significance requirement on

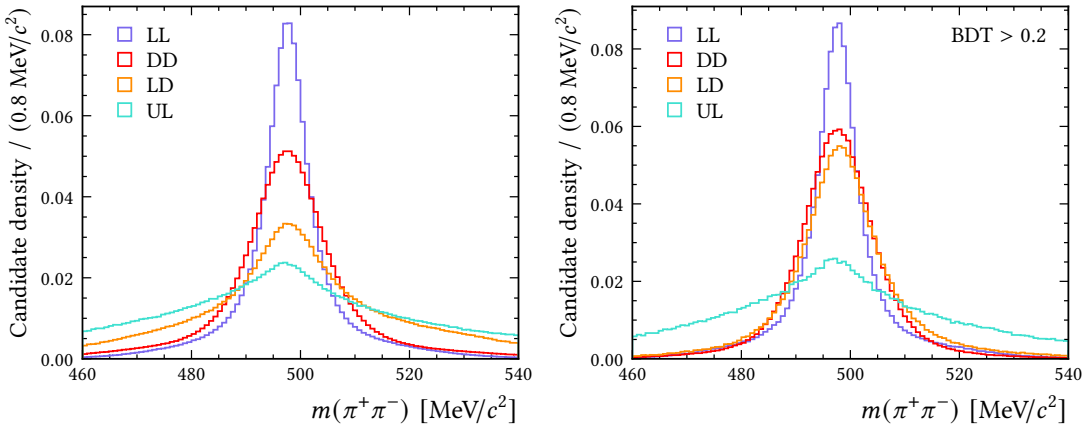

Figure 4.1 – Reconstructed mass of two reconstructed pion tracks after the centralized preselection (left) and with a soft BDT cut (right) in the K^0_S mass region. The K^0_S width is larger for tracks with missing momentum information. After the removal of most of the background (right), the LD and DD categories become comparable in reconstruction quality.

reconstructed kaons. The reconstructed B^0-meson mass is required to be in the range $5 \leq m(\mu^+\mu^-\pi^+\pi^-) < 6\,\text{GeV}/c^2$ for the dimuon, or $4.6 \leq m(e^+e^-\pi^+\pi^-) < 5.9\,\text{GeV}/c^2$ for the dielectron modes. The B^0 decay vertex must be compatible with being the origin of the reconstructed ψ and K^0_S trajectories. The B^0 origin vertex is required to be reconstructed with good precision, and the reconstructed decay time of the B^0 candidate needs to exceed 0.2 ps. Identical candidate selections are applied to all available simulated data sets. Backgrounds are not simulated as part of the generation of $B^0 \to \psi K^0_S$ events. Simulated physical backgrounds from $B^0 \to J/\psi K^*$ and $\Lambda^0_b \to J/\psi \Lambda$ decays are processed separately using the selection requirements meant for the signal channels in order to better understand the shapes of peaking backgrounds in the signal modes. Additional simulations of physical backgrounds are performed with phase-space-only simulation [118]. Radiative energy losses due to bremsstrahlung are taken into account in the simulation.

Long tracks are measured by all components of the tracking system and thus have the smallest momentum uncertainty, while D tracks are missing the fine-grained VELO information. Upstream tracks are only measured upstream of the magnet, and the momentum resolution is thus significantly decreased for these tracks. In Figure 4.1, a comparison of the invariant-mass distributions $m(\pi^+\pi^-)$ is shown for each reconstruction category before and after subsequent multivariate selections. The lack of momentum resolution for D and U tracks significantly increases the mass width, which is especially noticeable in the UL category. For LD candidates, this effect mainly extends to the background retained by the selection, as the width is more similar to the DD category after most of the combinatorial background is removed. A possible explanation for the presence of LD candidates is that one of the tracks is not reconstructed as an L track due to insufficient hits in the VELO or because of a failed track match, which causes the mass to be reconstructed with less hit information,

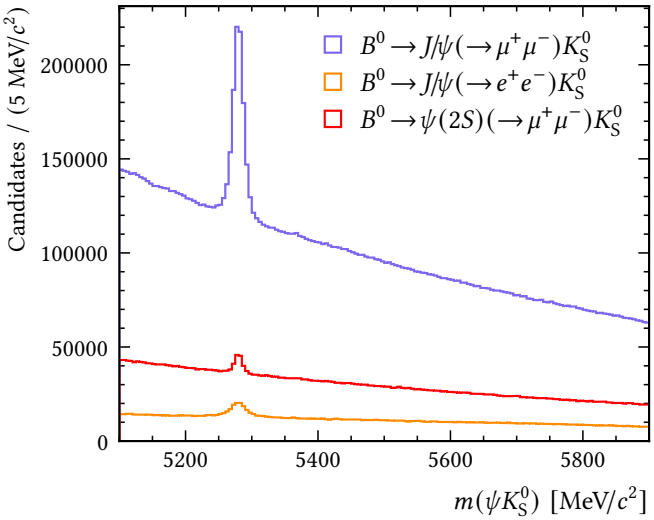

Figure 4.2 – Initial distributions of the reconstructed B mass in each decay channel when the initial candidate selection is applied.

increasing the width. For measurements of time-dependent CP violation, the B^0 decay-time reconstruction is of the highest importance, and as the contribution of the K^0_S vertex to the B^0 decay-time width is nearly insignificant, the decay-time width is determined by the ψ decay vertex. This vertex coincides with the B^0 decay vertex due to the short ψ lifetime. For this reason, if LD and UL candidates are selected and the signal-over-background ratio is increased to an acceptable level, these categories are as well suited for CP-violation measurements as LL and DD.

For the $B^0 \to J/\psi(\to e^+e^-)K^0_S$ decay, additional selection cuts are applied to reduce the significant surplus of combinatorial background caused by the inclusive trigger selection. Most importantly, requiring the angle between the reconstructed B-meson momentum and the direction between the PV and reconstructed B decay vertex to be small has a high impact on reducing combinatorial background in this mode. In addition, multiple selection requirements are applied to kaon and pion candidates to select more significant kaon flight distances.

The masses of muon, electron and pion candidates are set to their known values by the reconstruction software a priori, using particle identification information. The reconstruction of the B^0 decays is optimized by additionally setting the K^0_S, J/ψ, and $\psi(2S)$ masses to their known values [116] after which a fit of the fully reconstructed decay chain, referred to as *decay tree fit*, is performed with these constraints [119]. This procedure improves the resolution of the reconstructed B^0 mass and helps with the identification of backgrounds. In addition, the obtained fit χ^2 data that results from these fits is used to suppress combinatorial background. The selections mentioned so far are applied to the recorded LHCb data, and the initial mass distributions are shown in Figure 4.2.

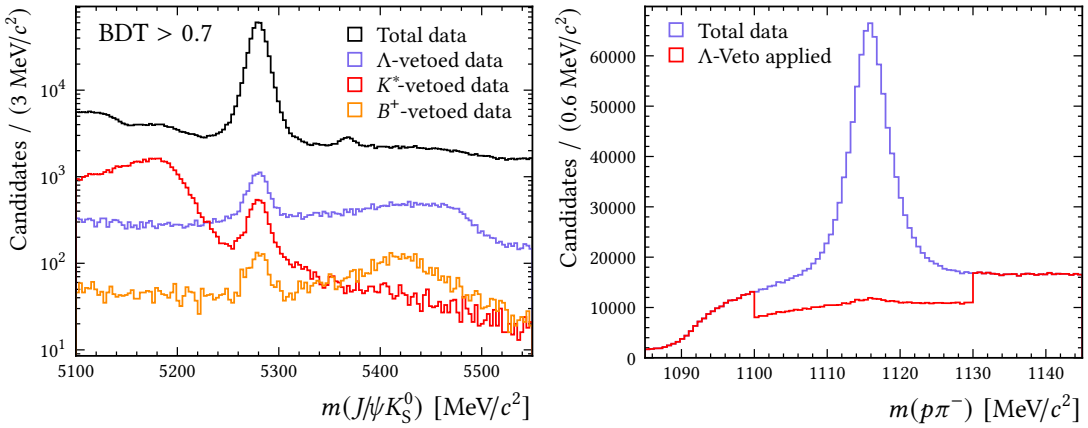

Figure 4.3 – Invariant mass of $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ candidates after the initial preselection described in Section 4.2, shown as a black histogram with distributions of vetoed data (left). Combinatorial background is reduced with a BDT for demonstration purposes. On the right, the invariant-mass distribution $m(p\pi^-)$ mass distribution is shown before (blue) and after (red) applying the Λ veto to the preselected data.

In the following, an invariant-mass observable like $m(J/\psi K^0_S)$ refers to the reconstructed B^0 mass with mass constraints on the B^0 descendants J/ψ and K^0_S, while $m(\pi^+\pi^-\mu^+\mu^-)$ does not include these constraints.

4.3 Background reduction

Multiple sources of background contributions can be identified after the preselection described in Section 4.2, and additional background sources become noticeable in the $m(\psi K^0_S)$ spectrum if constraints on the reconstruction quality are tightened. These sources include backgrounds from misidentified physical backgrounds like $\Lambda^0_b \to \psi\Lambda(\to p\pi^-)$ and $B^0 \to \psi K^*(892)(\to K^+\pi^-)$, and in some $B^+ \to \psi K^+$ decays, additional random pions are matched to the decay and mimic the signal. In addition, a component of irreducible background from partially reconstructed decays is present, which cannot be efficiently reduced with selections. To reduce background contributions of physical processes, selection requirements are applied on estimated particle identification probabilities, which are computed by multivariate techniques and rely largely on the RICH system and calorimeter information.

The combinatorial background is reduced with a boosted-decision-tree classifier (BDT) [120], which is a supervised learning algorithm. The BDT is trained on background in the upper mass sideband in data and simulated B^0 signal. During the training, a Λ veto (Section 4.3.1) is applied to remove Λ^0_b candidates from the upper mass sideband. The training of BDT classifiers is explained in Section 4.3.3. The invariant-mass distribution $m(J/\psi K^0_S)$ after the preselection and the contribution of each identified physical background is shown in Figure 4.3 on the left. In this plot, the reconstructed data removed with each veto is shown in addition.

4.3.1 $\Lambda_b^0 \to \psi\Lambda$ background

The decay $\Lambda_b^0 \to \psi\Lambda(\to p\pi^-)$ is a source of background in all three final states, caused by the misidentification of protons as pions, and the decay $\Lambda_b^0 \to \psi(2S)\Lambda$ is less frequent by a factor of approximately 1/2 compared to the $J/\psi\Lambda$ final state [116]. As a result, when the mass hypothesis of one pion is substituted with that of a proton, the invariant $m(p\pi^-)$ distribution has a clear resonance at the known Λ mass (Figure 4.3 right). In the reconstructed B^0 mass distribution, the Λ_b^0 contribution has a maximum above the B_s^0 mean mass with a long lower-mass tail extending into the signal region (Figure 4.3 left). The shape of this distribution is validated with $\Lambda_b^0 \to J/\psi\Lambda$ simulation that is reconstructed with the baseline selection requirements listed before. To reduce this contribution, pion candidates are required to have a low pion-proton misidentification probability in the mass window $1100 \leq m(p\pi^-) < 1130\,\mathrm{MeV}/c^2$ of the parent particle. This requirement is applied if the mass substitution of either pion causes the $m(p\pi^-)$ mass to be in the Λ mass window. The invariant-mass distribution $m(p\pi^-)$ before and after this requirement is shown in Figure 4.3 on the right. The signal efficiency measured in a simulated data set is $(98.510 \pm 0.010)\,\%$ for $B^0 \to J/\psi(\to \mu^+\mu^-)K_S^0$, $(98.532 \pm 0.013)\,\%$ for $B^0 \to \psi(2S)K_S^0$ and $(98.602 \pm 0.027)\,\%$ for $B^0 \to J/\psi(\to e^+e^-)K_S^0$ when applied after the preselection step. The efficiency in $\Lambda_b^0 \to J/\psi\Lambda$ simulation is $(2.33 \pm 0.07)\,\%$. This efficiency is reduced further to about 1.6 % by the BDT selection and other background vetoes. A further lowering of the misidentification probability threshold causes a disproportionate loss of signal efficiency. With these selections, a contribution from Λ_b^0 decays cannot be identified from a fit to the reconstructed B^0 mass in data, and therefore, this contribution is assumed to be well described by the combinatorial background description.

4.3.2 Misidentified and random-pion backgrounds

The decay $B^0 \to \psi K^*(892)(\to K^+\pi^-)$ is another source of background that is retained due to a misidentification of charged kaons as pions. Due to the short lifetime of the $K^*(892)$ state, which is abbreviated with K^* from here on, these decay happen inside the VELO and therefore only affect the reconstruction categories containing L track pions. This background is reduced efficiently by a lower cut on the neutral-kaon lifetime of $t_K > 0.5\,\mathrm{ps}$ and is further reduced by removing candidates with high pion-kaon misidentification probability; the veto applied in Section 4.3.1. The signal efficiency of the kaon lifetime requirement is 98 % for the LL category and 99.9 % for the LD and UL reconstructions as determined in simulated data sets. In a simulated data set of $B^0 \to J/\psi K^*$ decays, the lifetime requirement has a signal efficiency of $(1.127 \pm 0.007)\,\%$. The remaining contribution is negligible and thus described by the combinatorial-background model.

An additional background contribution from $B^+ \to J/\psi K^+$ decays is identified where the K^+ candidate is reconstructed as an L track and misidentified as a pion and an additional random upstream pion is matched to that track to form a fake K_S^0 candidate. Due to the additional particle, the reconstructed B^0 mass peaks at an offset of the pion

mass just above the B^0_s resonance (Figure 4.3 left). This contribution is quite small, only affects the UL category, and is removed by a loose pion-kaon misidentification probability threshold. This requirement lowers the rate of simulated K^* backgrounds from (1.127 ± 0.007) % to (0.328 ± 0.005) % in addition.

4.3.3 Combinatorial background

Combinatorial background is accumulated through the combination of unrelated tracks that resemble the signal process within the selection requirements of the analysis, and is the largest source of background in this measurement. The combinatorial background is reduced by a set of decision trees [121]. A decision tree is a multivariate, supervised learning algorithm that applies a sequence of binary decisions to a data point in a certain, optimized order. As such, its structure resembles that of a binary tree, with nodes representing the binary decisions and leaves representing the end of a given selection sequence in the tree. The values contained in the data point hereby define the path that is taken through this binary tree. Each leaf corresponds to one of a fixed number of classes that the data point may belong to. The order of selection requirements and the applied thresholds are optimized in the training process, which minimizes the classification error by comparing the current classification to the provided true training data labels. A BDT [120] is an ensemble of individual decision trees that are sequentially trained on the same data and whose predictions are combined. Boosting refers to the metric used to train each new decision tree: each decision tree is optimized to minimize the classification error of the previous set of decision trees. This procedure ensures faster convergence and model structure stability and, at the same time, provides a transparent hierarchy of selections from coarse cuts, typically applied by the first few trees, to intricate combinations of requirements that target smaller regions in the training parameter set.

In this measurement, the reduction of combinatorial background can be formulated as a binary classification problem of signal and background. Exemplary representations of signal and background, or *proxies*, are simulated B^0 decays for signal, and the upper mass sideband $m(\psi K^0_S) > 5450\,\mathrm{MeV}/c^2$ in data is used as a background proxy. The Λ veto is applied for the BDT training. BDTs are trained for each final state and each kaon reconstruction category, resulting in eight classifiers. As a cross-check, the BDTs are additionally trained for each year of data taking to verify the consistency between these subsets.

To prevent the classifiers from learning statistical fluctuations, and thus losing their generalizability and classification performance, some precautions are taken to minimize this effect. The main incentive for minimizing overtraining effects is to maximize the model performance, as biases due to the BDT selection can be neglected. This was verified by performing the measurement in bins of the BDT output and in bins of multiple training features of the BDT models, where the central values were observed to be unaffected by these setups beyond an expected statistical level. A natural countermeasure against overtraining is to stop the addition of decision trees as soon as the classification error stops decreasing significantly. Furthermore, the depth of the

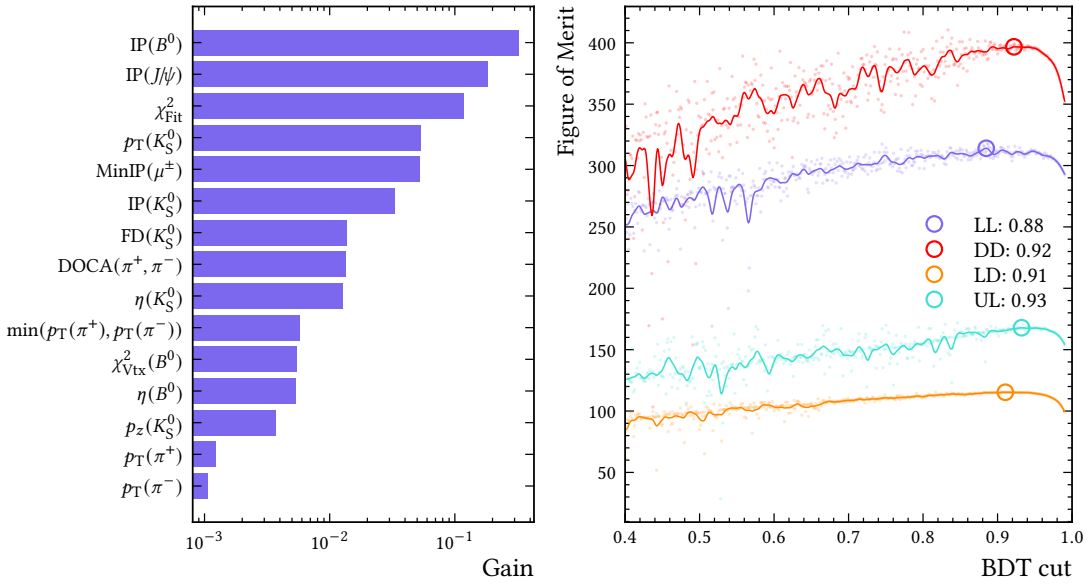

Figure 4.4 – Feature importance metric (logarithmic gain) of the set of training features (left) and figures of merit in dependence of the BDT cut (right).

decision trees is limited to four. Finally, in order to prevent overfitting and a possible modification of the combinatorial-background shape at the 5450 MeV/c^2 threshold, the k-folding technique [122] is used: In each instance, five BDT models instead of one are trained on 4/5 of the total training data, and the model performance is evaluated on test data sets comprising 1/5 of the data. The test data sets are statistically independent. The predictions of these classifiers are averaged for each candidate, and when applying the BDT classifier to all candidates, it is ensured that for the data points used during the training, BDTs are applied that were not trained on that specific candidate. These measures allow the upper mass sideband to be kept for the measurement and used for better constraining the combinatorial-background shape.

The training features used to characterize each candidate need to fulfill multiple criteria. Since simulated data is used to model the signal, it needs to be ensured that the simulated training feature is well described and that its shape matches the signal in the data well. To obtain the shapes of the signal distribution in data, which is a mixture of signal and background, the *sPlot* technique is used [123]. The *sPlot* technique assigns each data point a signal weight $w_s(m)$, or *sWeight*, which is a function of a chosen discriminating variable m. The discriminating variable is always the reconstructed B mass in the following. When the data set is weighted by $w_s(m)$, the signal component shape is obtained for all features independent of the discriminating variable. The functions $w_s(m)$ and $w_b(m)$ for the background are determined analytically using the fitted line shapes describing the signal and background components, and the set of weights is determined so that the variance is minimized. The derivation of sWeights is best explained in Ref. [124]. The signal distributions can further be used to identify features whose shapes differ between signal and background components. Using this

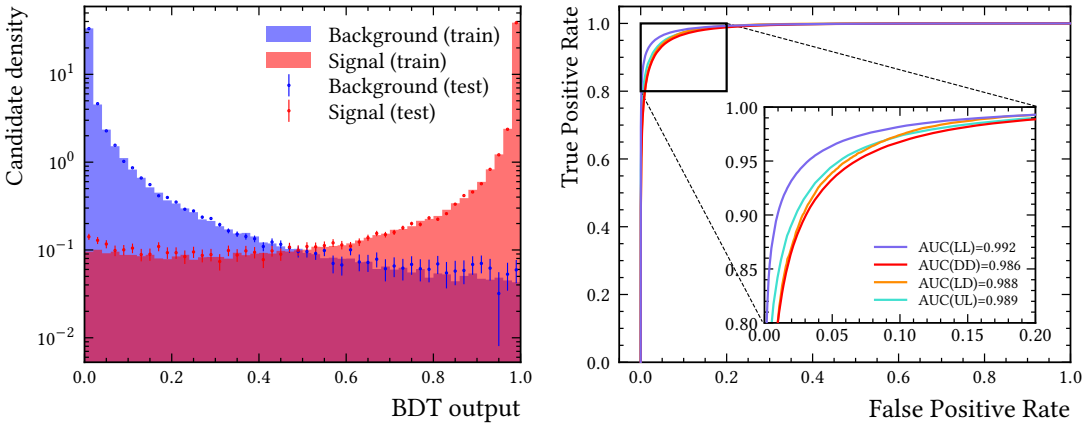

Figure 4.5 – Comparison of BDT output distributions for the training and test proxies (left), and ROC curves of the trained BDT models for the final state $J/\psi(\to\mu^+\mu^-)K^0_S$ for a balanced input of signal and background (right).

procedure, a set of 15 common features is identified for the kaon reconstruction classes of the $B^0 \to J/\psi(\to\mu^+\mu^-)K^0_S$ and $B^0 \to \psi(2S)(\to\mu^+\mu^-)K^0_S$ decay modes and 19 common features are found for the $B^0 \to J/\psi(\to e^+e^-)K^0_S$ mode. As there are small differences in the simulation quality of certain features between track reconstruction categories, additional features are selected that are well described in each category, and these features are listed in Appendix T.1. As these features do not add much separation power globally, their description is omitted here, and only the overlap of training features used in every BDT is shown.

The feature importances of the chosen training variables are shown in Figure 4.4. The selected features are as follows: The impact parameters (IP) measure the shortest distances between the reconstructed track and the PV and help to constrain the correct PV association of a track. The minimum impact parameter of a particle (MinIP) refers to the smallest IP with respect to any reconstructed PV. From the decay tree fit, in which particle mass constraints are applied, a χ^2 fit value is obtained (χ^2_{Fit}), which is a useful variable for reducing combinatorial background and can be understood as a constraint on the signal decay kinematics. The distance of closest approach of two reconstructed tracks (DOCA) is used as an additional parent particle vertex constraint. As shown in Figure 4.4, the B^0 and J/ψ-IP, the decay tree fit χ^2, track DOCA of the B^0 descendants J/ψ and K^0_S and the K^0_S transverse momentum (p_T) features contribute most to the separation power between signal and background. Additional features which are found to further improve the separation power, albeit to a smaller extent, and are well described in simulation, are included in the training. These features include the K^0_S flight distance (FD) in the laboratory frame of reference, the track pseudorapidity (η), and the projection of the K^0_S momentum onto the z axis (p_z). A momentum symmetrization variable for pions is added $\min(p_T(\pi^+), p_T(\pi^-))$, which selects the smaller transverse momentum of the two selected pions irrespective of track charge.

Table 4.2 – Signal efficiencies in simulated data sets of the trained BDT models.

Track type	$B^0 \to J/\psi K^0_S$ $J/\psi \to \mu^+\mu^-$	$B^0 \to J/\psi K^0_S$ $J/\psi \to e^+e^-$	$B^0 \to \psi(2S) K^0_S$ $\psi(2S) \to \mu^+\mu^-$
LL	87.94 ± 0.05	94.54 ± 0.10	85.02 ± 0.07
DD	79.37 ± 0.04	76.91 ± 0.11	82.84 ± 0.05
LD	80.01 ± 0.16		
UL	79.87 ± 0.11		

In Figure 4.5, the comparison of BDT outputs for the labeled test and training data sets is shown on the left, and in the right plot, the receiver operating characteristic (ROC) of the trained models is shown. The output distributions of the test and training data sets match well across the full output range, indicating that overtraining effects have been reduced to a reasonable level. The ROC curves of the trained classifiers have high area under curve (AUC) scores for a test setup with equal amounts of signal and background, which implies a strong separation power for these classes. In Table 4.2, the signal efficiencies of the trained BDT models are listed per final state and track category.

To understand which relationships the BDT models learn, a single decision tree with a comparable accuracy metric is trained. The signal selection requirements learned by this simplified model can be characterized by the following selections for the reconstruction category LL

$$(\mathrm{IP}(B^0) < 0.043) \wedge (\mathrm{IP}(\psi) > 0.061) \wedge (\mathrm{MinIP}(K^0_S) > 0.077) \quad (79\,\%)$$
$$(\mathrm{IP}(B^0) < 0.043) \wedge (\mathrm{IP}(\psi) < 0.061) \wedge (\chi^2_{\mathrm{Fit}} < 12.14) \wedge (\mathrm{MinIP}(K^0_S) > 0.088) \quad (14\,\%)$$

and the selections for the DD category can be summarized as

$$(\mathrm{IP}(B^0) < 0.053) \wedge (\mathrm{IP}(\psi) > 0.061) \quad (79\,\%)$$
$$(\mathrm{IP}(B^0) < 0.03) \wedge (\mathrm{IP}(\psi) < 0.061) \wedge (\chi^2_{\mathrm{Fit}} < 13.588) \quad (16\,\%)$$

where the two selections per track category define the two most frequent paths through the decision tree. The given percentages are the fractions of signal events taking the shown path through the decision tree, irrespective of the associated amount of background events taking the same path. Impact parameter requirements are given in millimeters. In summary, the relationship of impact parameters of the B^0, ψ, and K^0_S particles is the main source of separation power between signal and background. In addition, a cut on the fit χ^2_{Fit} value helps to select additional signal events if the impact parameter of the ψ candidate is small.

The multivariate selection is finalized by a lower cut on the BDT output threshold, which is chosen such that it maximizes the signal sensitivity, defined as $S/\sigma(S)$, which is the ratio of signal yield and its uncertainty that is determined from an ML fit of the reconstructed B mass. The signal sensitivity is evaluated over a range of cuts and the cut which maximizes the signal sensitivity is chosen. Figure 4.4 shows the signal sensitivity

in dependence of the BDT cut. A cubic smoothing spline [125, 126] is fitted to this data to avoid choosing the optimum on the basis of a statistical fluctuation, but in practice, the choice of the optimizing metric was seen to have no relevant influence on the precision of the final result. An alternative metric $S/\sqrt{S+B}$, *i.e.* the signal sensitivity expressed as a signal yield over a background fluctuation, was seen to be a viable alternative choice. The mass model used to integrate the signal yield is documented in Section 4.4.

In the final selection step, B-meson candidates to which multiple candidate decays are matched are removed. A set of such candidates shares the same event and running conditions index and can therefore be easily filtered. For these groupings, the candidate with the highest B-meson p_T is chosen while the other reconstructions are removed. The fraction of multiple candidates after the BDT selection was found to be $(0.719 \pm 0.012)\,\%$, $(0.533 \pm 0.021)\,\%$, and $(0.657 \pm 0.035)\,\%$ for the final states $J/\psi(\to \mu^+\mu^-)K^0_S$, $J/\psi(\to e^+e^-)K^0_S$, and $\psi(2S)(\to \mu^+\mu^-)K^0_S$, respectively.

4.4 Description of the reconstructed B mass

The reconstructed B-mass distribution consists of one signal component and multiple background components. Each final state and each considered track type category in the data is described with the same mass model, which is defined as a probability distribution function (pdf). This model is optimized separately for each data subset due to the kinematic differences between final states and reconstruction classes. The signal is described by a two-tailed Hypatia distribution [127], which is a generalized Gaussian peak with exponential tails towards lower and higher masses. For this measurement, the Hypatia parameters β and ζ are set to zero as they do not improve the description of the B-peak. The Hypatia distribution is defined here as

$$H_{\text{Signal}}(m, \mu, \sigma, n_1, n_2, \alpha_1, \alpha_2, \lambda) \propto \begin{cases} c_1(c_2 - (m-\mu))^{-n_1} & \text{if } m-\mu < -\alpha_1\sigma \\ c_3(c_4 + (m-\mu))^{-n_2} & \text{if } m-\mu > \alpha_2\sigma \\ \left(1 - \frac{(m-\mu)^2}{2\sigma^2(1+\lambda)}\right)^{\lambda-\frac{1}{2}} & \text{otherwise} \end{cases}, \tag{4.4}$$

and the constants c_1, c_2, c_3, and c_4 are uniquely determined by the free parameters and ensure the continuity and differentiability of H_{Signal} and are given in Equation (E.1). The mean and width of this distribution are parametrized by μ and σ, and the shape of the exponential tails is determined by $n_{1,2}$. The parameter λ modifies the shape of the peak and influences the width of the distribution. It can have any value outside the interval $[-1, \frac{1}{2}]$ and $\lambda < -1$ is chosen for all mass fits. To limit the number of free parameters in mass fits to data, the Hypatia parameters $\alpha_{1,2}$ and $n_{1,2}$ are determined in simulated data sets and set constant in fits to data. The $B^0_s \to \psi K^0_S$ component is modeled with the identical distribution and shares all free parameters with the signal component, and a constant offset is added to its mean, which equals the B^0_s-B^0 mass difference [116]. The combinatorial background is modeled with a single exponential function with one free parameter. The irreducible background component is modeled with an exponential

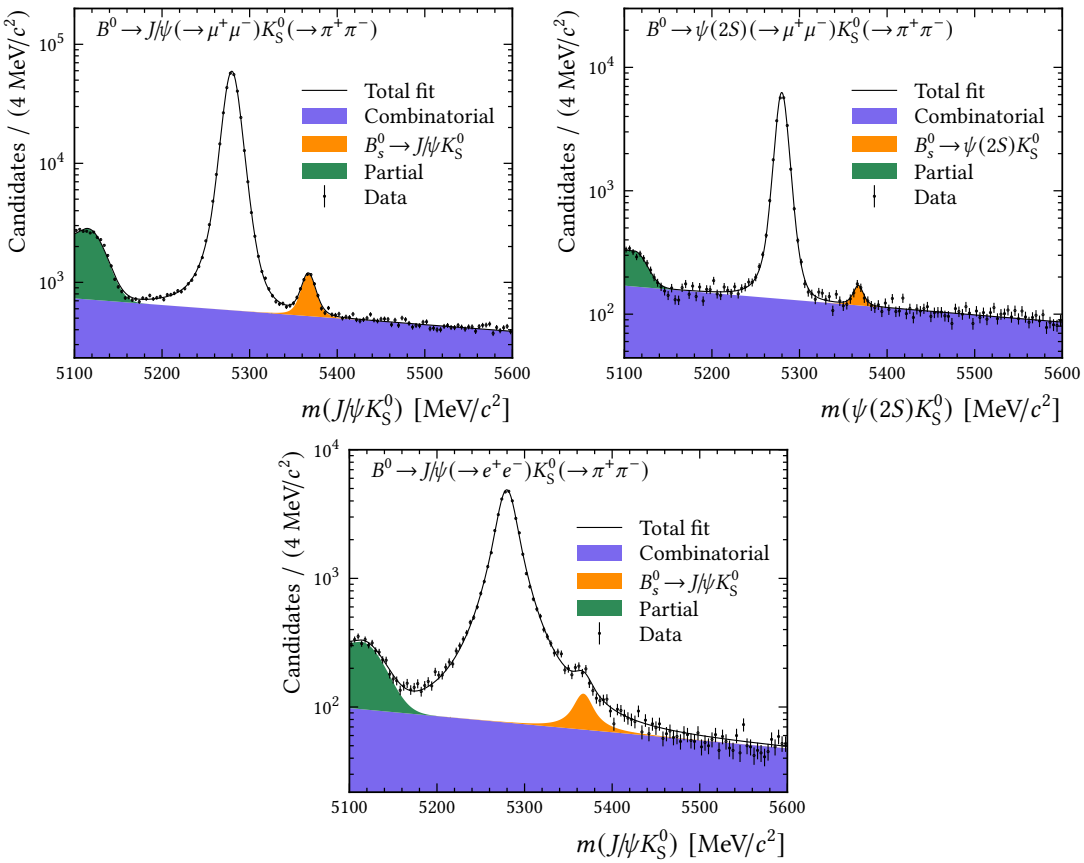

Figure 4.6 – Mass fits of the analyzed decay modes of the full Run 2 data set. The high upper mass range is not shown to better emphasize the signal region. The lower mass range $5100 < m(\psi K^0_S) \leq 5150\,\text{MeV}/c^2$ is included in fits of the reconstructed B mass and is statistically removed for the final fit.

function that smoothly transitions to a Gaussian distribution at a certain mass threshold. This model was seen to be a satisfactory description of the reconstructed B mass of the multi-body decay $B^0 \to J/\psi(\to \mu^+\mu^-)K^*(892)(\to K^0_S(\to \pi^+\pi^-)\pi^0)$ if it is reconstructed without the neutral-pion four-vector, which is a likely physical background that contributes to the irreducible background component. This decay is simulated in phase-space-only simulation with a simplified LHCb detector geometry model [118]. As the lower mass sideband $m < 5150\,\text{MeV}/c^2$ is excluded from the data with which the *CP*-violation parameters are determined, the irreducible background model can be understood as a single Gaussian distribution in the chosen B mass range.

The reconstructed B^0 decay time is reduced to the range $0.2 \leq t < 15\,\text{ps}$ in each channel for consistency. As is discussed in Chapter 5, candidates for which the B production flavor is not known do not add any sensitivity to *CP*-violating effects. For this reason, mass fits are performed, and resulting yields are reported on the data subset for which likely production flavors have been assigned. This ensures proper sWeight normalization in the final fit, in which only tagged events are considered. Table 4.3 lists

Table 4.3 – Candidate yields of the three decay channels determined via ML fit in the full B^0 mass range $5100 \leq m(\psi K^0_S) < 5900\,\mathrm{MeV}/c^2$ for the full sample and the tagged subsample.

Component	$B^0 \to J/\psi K^0_S$ $J/\psi \to \mu^+\mu^-$	$B^0 \to J/\psi K^0_S$ $J/\psi \to e^+e^-$	$B^0 \to \psi(2S) K^0_S$ $\psi(2S) \to \mu^+\mu^-$
Total yields			
B^0 signal	358 630 ± 669	47 152 ± 283	27 489 ± 179
B^0_s signal	4040 ± 107	582 ± 61	252 ± 35
Combinatorial bkg.	107 901 ± 542	12 183 ± 288	24 809 ± 191
Partial bkg.	22 122 219	2598 111	1238 67
Tagged yields			
B^0 signal	306 322 ± 621	42 870 ± 271	23 570 ± 166
B^0_s signal	3410 ± 99	538 ± 59	200 ± 31
Combinatorial bkg.	91 851 ± 507	11 647 ± 279	20 736 ± 175
Partial bkg.	18 798 204	2364 108	1092 62

the obtained component yields when untagged yields are included for completeness. Mass fits are performed separately in each channel and track-type category, and it is ensured, that the yields and shape parameters are compatible between years of data taking to justify the approach of merging the data sets of individual years. The signal and background yields are normalized after the mass fit to the final mass range $5150 \leq m(\psi K^0_S) < 5900\,\mathrm{MeV}/c^2$ in which the measurement is performed as there is no benefit of including the partial background. sWeights are then computed using these normalized yields and the line shapes obtained in the mass fit. Mass fits are performed on the full mass ranges $5100 < m(\psi K^0_S) \leq 5900\,\mathrm{MeV}/c^2$ and include the partial background in order to increase the precision of the combinatorial background description. In Figure 4.6, the mass fits of the three analyzed final states are shown, and the model is seen to be a good description of the data. In Figure F.1, the same mass fits are shown in linear y-axis scale and in Figure F.2 mass fits of the new kaon reconstruction categories LD and UL are shown. The B^0 resonance is narrower in the $\psi(2S)K^0_S$ final state because of the smaller available phase space for the B^0 meson, and in the $J/\psi(\to e^+e^-)K^0_S$ mode, the B^0 peak is wider due to the worse momentum resolution for electrons compared to muons.

In total, 4874 ± 128 $B^0_s \to \psi K^0_S$ decays are reconstructed. From the number of reconstructed B^0 decays, the expected $B^0_s \to \psi K^0_S$ yield can be calculated with the

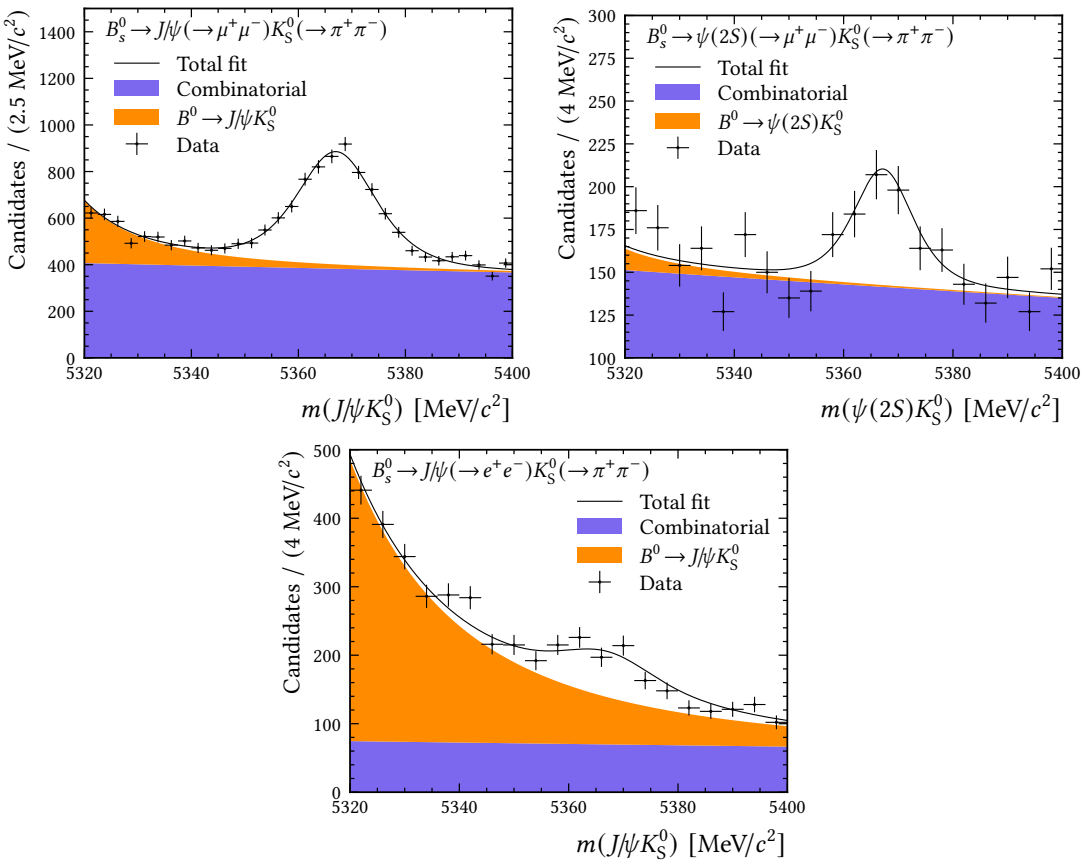

Figure 4.7 – Mass fits of the analyzed decay modes of the full Run 2 data set in the B_s^0 mass range.

known branching fractions and hadronization fraction f_s/f_d listed in Table T.2.1:

$$\begin{aligned} N_{B_s^0} &= \frac{f_s}{f_d} \times \sum_f \frac{\mathcal{B}(B_s^0 \to f)}{\mathcal{B}(B^0 \to f)} \times \frac{\epsilon_{B^0 \to f}}{\epsilon_{B_s^0 \to f}} \times N(B^0 \to f) \\ &= 4669 \pm 346(\text{stat}) \pm 146(f_s/f_d). \end{aligned} \tag{4.5}$$

The measured yield matches the expectation well given a total of 433 271±748 $B^0 \to \psi K_S^0$ decays, which includes untagged events. In Figure 4.7, the mass fits are shown for each final state in the region near the invariant B_s^0 mass.

The $B_s^0 \to \psi(2S)(\to \mu^+\mu^-)K_S^0$ channel has recently been discovered by the CMS collaboration with a measured relative branching ratio of [128]

$$\left(\frac{\mathcal{B}(B_s^0 \to \psi(2S)K_S^0)}{\mathcal{B}(B^0 \to \psi(2S)K_S^0)}\right)_{\text{CMS}} = (3.33 \pm 0.69(\text{stat}) \pm 0.11(\text{syst}) \pm 0.34(f_s/f_d)) \times 10^{-2}, \tag{4.6}$$

Table 4.4 – Summary of selection efficiencies in percent determined in simulation and weighted by the yields of the reconstruction decay channels in data.

Selection	$B^0 \to J/\psi K^0_S$ $J/\psi \to \mu^+\mu^-$	$B^0 \to J/\psi K^0_S$ $J/\psi \to e^+e^-$	$B^0 \to \psi(2S) K^0_S$ $\psi(2S) \to \mu^+\mu^-$
LHCb geometry	19.897 ± 0.083	19.75 ± 0.50	19.57 ± 0.05
Reconstruction	2.648 ± 0.003	0.826 ± 0.002	3.479 ± 0.005
Λ-veto	98.563 ± 0.017	98.660 ± 0.034	98.517 ± 0.019
BDT selection	82.21 ± 0.05	82.54 ± 0.11	83.67 ± 0.06
K^*-veto	99.264 ± 0.011	99.964 ± 0.005	99.365 ± 0.011
B^+-veto	99.820 ± 0.006	99.846 ± 0.011	99.817 ± 0.006
Decay-time range	98.738 ± 0.017	99.996 ± 0.002	97.873 ± 0.026
Multiple candidates	99.590 ± 0.012	99.819 ± 0.014	99.837 ± 0.006
Tagging efficiency	85.57 ± 0.06	90.01 ± 0.11	86.14 ± 0.06
Total	0.354 ± 0.002	0.115 ± 0.001	0.470 ± 0.002

with a significance of 5.2σ. From the $B^0_s \to \psi(2S) K^0_S$ yield in this analysis, the same branching ratio can be estimated to be

$$\begin{aligned}\frac{\mathcal{B}(B^0_s \to \psi(2S) K^0_S)}{\mathcal{B}(B^0 \to \psi(2S) K^0_S)} &= \frac{N(B^0_s \to \psi(2S) K^0_S)}{N(B^0 \to \psi(2S) K^0_S)} \times \frac{f_d}{f_s} \times \frac{\epsilon_{B^0 \to \psi(2S) K^0_S}}{\epsilon_{B^0_s \to \psi(2S) K^0_S}} \\ &= (3.36 \pm 0.50(\text{stat}) \pm 0.11(f_s/f_d)) \times 10^{-2}. \end{aligned} \tag{4.7}$$

This value is remarkably close to the CMS result and has higher statistical sensitivity despite the BDT cut not being optimized for the selection of B^0_s candidates. An estimated efficiency correction $\epsilon_{B^0 \to \psi(2S) K^0_S} / \epsilon_{B^0_s \to \psi(2S) K^0_S}$ that measures efficiency differences between B^0 and B^0_s modes is very close to unity and thus has negligible influence on this estimate.

The selection efficiencies are measured for each analysis step with corresponding simulated data sets and in each reconstruction category, and their track-category and final-state averages are summarized in Table 4.4. The efficiencies obtained in each reconstruction category are weight-averaged by the corresponding signal yield of each reconstruction category obtained in the data. In Tables T.2.2 to T.2.4, more detailed summaries of selection efficiencies in each track type category and decay channel are listed. Between reconstruction categories, only minor efficiency differences are observed for selection steps applied in the main analysis, and larger variations are seen in the reconstruction steps. Because of the different decay topologies and differences between pion track categories, such differences are expected. Furthermore, the BDT efficiency for LL candidates is consistently higher by a significant amount as these candidates are reconstructed with all tracking subdetectors and are thus better separable

from combinatorial backgrounds. In the end, only candidates with tagging information are kept, and tagging performances quoted in the following chapters depend on the tagging efficiency of the underlying sample shown here. The tagging efficiency of the $B^0 \to J/\psi(\to e^+e^-)K^0_S$ sample is increased compared to the dimuon modes because of stricter kinematic selection requirements applied to the electron candidates in the candidate selection.

5 Flavor Tagging at LHCb

B_d mesons are produced in one of two possible flavor eigenstates B^0 or $\bar{B}^0$ and can decay either in the state of their production or as the opposite flavor eigenstate through mixing. The time-dependent difference of these decay rates is at the core of the measurement described here, as it is sensitive to CP-violating effects and the mixing parameters in general, as motivated in Section 2.4. In some cases, the initial B-meson flavor can be deduced by the charge and particle identity configuration of its final state, as is the case in $B^+ \to J/\psi K^+$ decays. In neutral decays like $B^0 \to J/\psi K^*(\to K^+\pi^-)$, a good estimate is possible. In the decay $B^0 \to \psi K^0_S$ in contrast, the final state is always fully charge symmetric, and a production flavor assignment based on the final state charge is therefore ruled out. Despite the charge-symmetric final state, it is

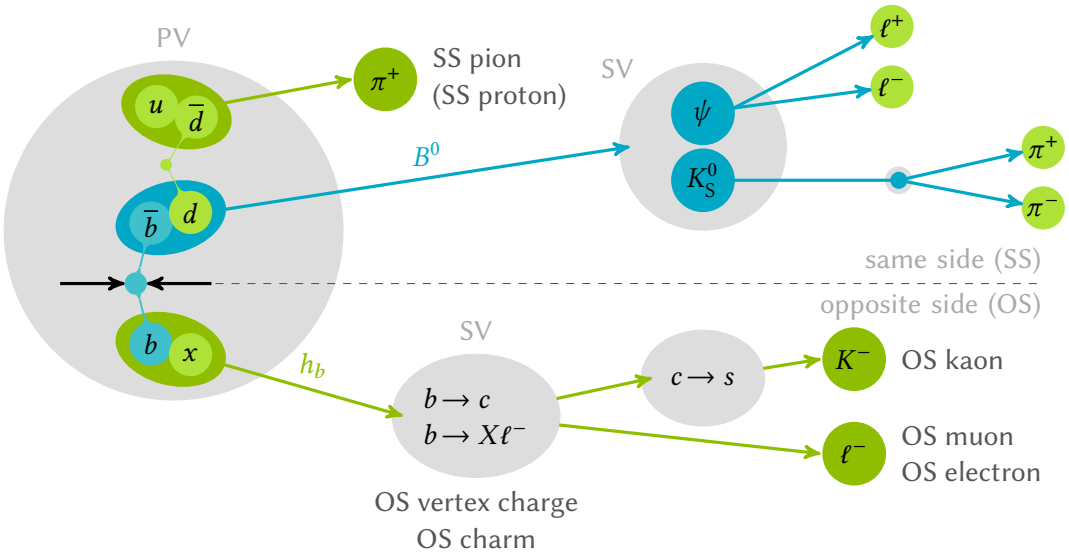

Figure 5.1 – LHCb flavor tagging technique illustrating the $b\bar{b}$ production and subsequent hadronizations and decays in which tagging particles are created. The tagging algorithms associated to these particles are annotated [129].

possible to reconstruct the B production flavor for a significant fraction of candidates by analyzing the particles created in the hadronization processes in which the B meson is created. Furthermore, as b-quarks are produced in pairs and are of opposite flavor, associating charged final-state particles to the decay of the other b-hadron can be used to reconstruct the flavor of the signal B meson. These associated witness particles, referred to as *tagging particles* in the following, can carry a charge that is correlated to

the B production process and from which the production flavor can be inferred with the help of multivariate techniques.

At electron-positron colliders, B mesons are created as quantum-entangled pairs in $\Upsilon(4S) \rightarrow B^0\bar{B}^0$ decays instead, and their flavor identity is determined from a reconstruction of the B meson decay, which is not the signal. In all cases, these methods are referred to as *flavor tagging*, and in the following, the hadronization analysis approach used at LHCb [130, 131] is explained.

Figure 5.1 illustrates the $b\bar{b}$ production and subsequent hadronization and decays of the created b-hadrons. It needs to be emphasized that this illustration is idealized and that the amount of hadronization processes in actual proton-proton collisions is considerably higher. The hadronization processes are interconnected so that colorless hadrons are formed, conserving color. Furthermore, the illustration does not depict all the numerous ways in which a correct reconstruction of tagging particle charges can indicate the wrong production flavor correlation, nor does it show all the known cases in which the production flavor is fundamentally unobservable. A common example of a way in which a charge identification can be wrong is that generally, in $b \rightarrow c \rightarrow s$ transitions two leptons of opposite charge can be created in $W^{\pm}$ boson decays [74]. Nevertheless, it is a useful illustration for the purpose of understanding the flavor tagging approach at LHCb and for introducing the usual flavor tagging algorithms that are in use.

Hadronization particles created by the b-quark that becomes the signal B meson through hadronization are called *same-side* (SS) particles, while particles created by the partner b-quark are called *opposite-side* (OS) particles. For each of these processes, dedicated tagging algorithms (*taggers*) exist that are optimized for their reconstruction and production flavor assignment. In the case of B^0 mesons, the charged SS particles of interest can be charged pions or protons. On the OS, the partner b hadron decays via $b \rightarrow c \rightarrow s$ transitions in which typically high momentum muons or electrons are formed in the $b \rightarrow c$ transition, while charged kaons can be created in the final state. Furthermore, the reconstruction of D meson decays on the OS can contribute with information and enhance the production flavor assignment. Finally, a momentum-weighted analysis of all track charges associated with the OS-secondary vertex can add valuable information as it improves the classifications performed by single-particle taggers. The flavor tagging algorithms used in this measurement have been first developed for the Run 1 data taking phase [130, 132, 133] and have been optimized and reimplemented based on Run 1, as well as 2015 and 2016 data from Run 2 [134].

Flavor tagging algorithms compute two numbers for each reconstructed B candidate, an integer tagging decision, and a mistag probability. The tagging decision, q, can have three possible values, $q = -1$ for $\bar{B}^0$, $q = 1$ for B^0, and $q = 0$ is assigned if the candidate is not tagged by the algorithm. The mistag estimate η is the probability of the tag decision being incorrect, and it is usually the output of a multivariate model like a neural network classifier or a BDT. This estimate is calibrated with calibration data sets, as explained in Section 5.5, to represent the mistag rates accurately and is

denoted by $\omega(\eta)$. A calibration function of the general form

$$\omega(\eta, q, \boldsymbol{p}) = \frac{(1-q)}{2}\omega^{-}(\eta, \boldsymbol{p}) + \frac{(1+q)}{2}\omega^{+}(\eta, \boldsymbol{p}) \tag{5.1}$$

for a set of calibration parameters $\boldsymbol{p}$ is determined on a suitable calibration data set, where the tagging algorithm response distribution is measured on data subsets of true initial B^0 with $\omega^{+}(\eta, \boldsymbol{p})$ and true initial $\bar{B}^0$ with $\omega^{-}(\eta, \boldsymbol{p})$. These subsets are generally not well identifiable in the data, and for this reason, an averaged representation of calibration parameters is introduced in Section 5.1. Calibration functions are not evaluated for untagged events.

The fraction of tagged events, *i.e.* events with assigned production flavor, is given by the tagging efficiency ϵ_{tag}

$$\epsilon_{\text{tag}} \equiv \frac{N_{\text{tag}}}{N} = \frac{\sum_j^n \kappa_j |q_j|}{\sum_j^n \kappa_j}, \tag{5.2}$$

whereby the right-hand side shows how ϵ_{tag} is computed for a given set of event weights κ_j. The mean tagging dilution $\mathcal{D}_{\text{FT}}$ measures to first order the decrease of the time-dependent *CP*-violation amplitude $\mathcal{A}^{CP}(t)$ in Equation (2.39), as derived in Equation (5.10), due to the effect of tagging with large production flavor uncertainties, and its measurement is therefore of high relevance. Powers of the mean tagging dilution are defined as

$$\mathcal{D}_{\text{FT}}^m \equiv \frac{1}{\sum_j^n \kappa_j |q_j|} \sum_j^n \kappa_j |q_j| (1 - 2\omega_j)^m \tag{5.3}$$

and $\mathcal{D}_{\text{FT}}$ can have a value in the range from zero to one, representing an average of zero and full production flavor knowledge for the data sample.

5.1 Derivation of decay rate coefficients

In the following, the B decay rates from Equations (2.27) and (2.28) are further developed to account for different kinds of flavor tagging asymmetries and, most significantly, imperfect flavor tagging. These decay rates can be summarized in the form

$$\begin{aligned} P_{B^0}(t) &\propto e^{-\Gamma t/\hbar} \left[H + C\cos(\Delta m_d t) - S\sin(\Delta m_d t)\right], \\ P_{\bar{B}^0}(t) &\propto e^{-\Gamma t/\hbar} \left[H - C\cos(\Delta m_d t) + S\sin(\Delta m_d t)\right], \end{aligned} \tag{5.4}$$

where $\Delta\Gamma_d = 0$ has been assumed, and here, the parameter H is the factor of the hyperbolic cosine component, which is defined as being one given how B^0 decay rates are usually normalized according to the definitions in Equation (2.40). In the following, it is shown that all *CP* asymmetry amplitudes, including H, effectively deviate from their assumed values if imperfect and potentially biased flavor tagging information and the production asymmetry are taken into account.

A priori, the amount of B^0 and $\bar{B}^0$ mesons produced in proton collisions in hadronizations of b-quarks can be asymmetric to a small degree [135], and this effect is measured with the production asymmetry parameter

$$\alpha \equiv \frac{N^{\text{true}}(\bar{B}^0) - N^{\text{true}}(B^0)}{N^{\text{true}}(\bar{B}^0) + N^{\text{true}}(B^0)}, \tag{5.5}$$

and to our current knowledge, this parameter is compatible with zero for B meson production at the LHC. This parameter needs to be measured in a suitable control mode, which is detailed in Section 5.6. The two B decay rates in the total pdf describing the data set are therefore rebalanced to account for α:

$$P_{CP}(t) \propto (1-\alpha)P_{B^0}(t) + (1+\alpha)P_{\bar{B}^0}(t). \tag{5.6}$$

The calibrated mistag estimate is integrated into this expression to interpolate between the description of B^0 and $\bar{B}^0$ decays as follows

$$\begin{aligned} P_{CP}(t) \propto &\left[\epsilon^+(\delta_{q,1}(1-\omega^+) + \delta_{q,-1}\omega^+) + \delta_{q,0}(1-\epsilon^+)\right](1-\alpha)P_{B^0}(t) \\ &+ \left[\epsilon^-(\delta_{q,-1}(1-\omega^-) + \delta_{q,1}\omega^-) + \delta_{q,0}(1-\epsilon^-)\right](1+\alpha)P_{\bar{B}^0}(t). \end{aligned} \tag{5.7}$$

Here, ω^+ (ω^-) and ϵ^+ (ϵ^-) are the mistags and tagging efficiencies assigned to *true* B^0 ($\bar{B}^0$); data subsets that are not obviously identifiable without flavor tagging techniques in the first place. The measured mistags and tagging efficiencies are therefore reparametrized to the more useful form

$$\omega \equiv \frac{\omega^+ + \omega^-}{2} \qquad \Delta\omega \equiv \omega^+ - \omega^- \tag{5.8}$$

$$\epsilon_{\text{tag}} \equiv \frac{\epsilon^- + \epsilon^+}{2} \qquad \Delta\epsilon \equiv \epsilon^- - \epsilon^+ \tag{5.9}$$

whereby ω and ϵ_{tag} now represent the measured mean quantities and the $\Delta\omega$ distribution is known from the calibration function in Equation (5.1) as it is determined in a separate decay mode. The flavor tagging efficiency asymmetry $\Delta\epsilon$ is likewise determined in the calibration mode $B^0 \to J/\psi K^*$ as is described in Section 5.6. In the following, ω, $\Delta\omega$, ω^+, and ω^- are to be understood as calibration functions of the raw mistag η and share the same calibration parameters, which are implied.

It follows from Equation (5.7), that in the case of perfect tagging, when $\omega^\pm = 0$ and $\epsilon^\pm = 1$ and the production asymmetry is zero, no modification of the decay rate is needed, which remains in its original form Equation (5.4). Consequently, non-zero production asymmetry, mistag asymmetry, or efficiency asymmetry all directly influence the determination of the CP-violation observables. A sizable production asymmetry value would manifest as an oscillating component in the decay-time distribution, even in the absence of flavor tagging information. Finally, to express the size of these contributions concisely, Equation (5.7) is expanded and rearranged so that all flavor tagging parameters are absorbed by common coefficients $\tilde{S}(q,\eta)$, $\tilde{C}(q,\eta)$ and $\tilde{H}(q,\eta)$. These are functions of the flavor tagging calibrations and take the place of the theoretical

parameters S, C, and H in Equation (5.4) and thereby express all effects of flavor tagging on each $\mathcal{A}^{CP}$-amplitude. The resulting expressions for the case $q \neq 0$ are

$$\tilde{S}(q,\eta) = S\left\{-qD\left(\frac{\alpha\Delta\epsilon}{2} + \epsilon_{\text{tag}}\right) - \frac{1}{2}q\Delta\epsilon\Delta\omega - q\alpha\epsilon_{\text{tag}}\Delta\omega + \alpha\epsilon_{\text{tag}} + \frac{\Delta\epsilon}{2}\right\} \tag{5.10}$$

$$\tilde{C}(q,\eta) = -\frac{\tilde{S}(q,\eta)}{S}C \tag{5.11}$$

$$\tilde{H}(q,\eta) = \epsilon_{\text{tag}} - qD\left(\frac{\Delta\epsilon}{2} + \alpha\epsilon_{\text{tag}}\right) - q\epsilon_{\text{tag}}\Delta\omega - \frac{1}{2}q\alpha\Delta\epsilon\Delta\omega + \frac{\alpha\Delta\epsilon}{2} \tag{5.12}$$

with the tagging dilution $D = (1 - 2\omega)$. When merely tagging asymmetries $\Delta\epsilon$ and $\Delta\omega$ are ignored one obtains the simplified expressions

$$\tilde{S}(q,\eta) = [qD + \alpha] \cdot S, \qquad \tilde{C}(q,\eta) = -[qD + \alpha] \cdot C, \qquad \tilde{H}(q,\eta) = 1 - q\alpha D, \tag{5.13}$$

which identifies the tagging dilution as the most significant influence on the value of $\sin(2\beta)$ in this measurement. The role of ϵ_{tag} in Equations (5.10) to (5.12) is to account for an amplification of the effect of tagging asymmetries, whereby the selection of a tagged subsample can skew the data toward selecting one production flavor with higher probability if there is an underlying tagging or production asymmetry. The effects of $\Delta\epsilon$, $\Delta\omega$, or α on the asymmetry amplitude are therefore scaled by ϵ_{tag}, and perhaps surprisingly, the selection of a tagged subset can reveal an oscillation in the data if there is an underlying flavor tagging bias, even if the mistag values if this data are not directly used. In the absence of tagging and production asymmetries, ϵ_{tag} becomes an isolated global factor and is thus absorbed by the normalization factor. The tagging asymmetry $\Delta\epsilon$ and production asymmetry α are found to be small but are directly added to the value of S. Terms involving any product of two or three asymmetries can be considered negligible, and no sensitivity to their influence can be expected.

The influence of the flavor tagging dilution factor D on the uncertainty of the CP-violation parameter S can be derived by calculating the expectation value of the log-likelihood curvature with respect to the probability distribution given in Equation (5.4):

$$\sigma_S^{-2} = -E\left[\frac{\partial^2}{\partial S^2}\log(P_{CP}(t))\right] \approx \int_0^\infty e^{-2\Gamma t/\hbar}\frac{D^2\sin^2(\Delta m_d t)}{1 + qSD\sin(\Delta m_d t)}dt \;\propto\; D^2 \tag{5.14}$$

if D is assumed to be a sample average and $S \cdot D$ is approximated to be small. This demonstrates that the measurement uncertainty of S depends, besides the data set size $\epsilon \cdot N$, with $\epsilon = \epsilon_{\text{sel}}\epsilon_{\text{tag}}$ being the product of flavor tagging and selection efficiencies, on the flavor tagging quality as measured by the dilution factor in the following way

$$\sigma_S \propto \frac{1}{\sqrt{\epsilon_{\text{tag}}\mathcal{D}_{\text{FT}}^2\epsilon_{\text{sel}}N}}. \tag{5.15}$$

The term $\epsilon_{\text{tag}}\mathcal{D}_{\text{FT}}^2$ is referred to as the *tagging power*, and therefore measures the decrease in effective sample size caused by a flavor tagging efficiency and corresponding tagging

quality. The uncertainty on the parameter C has the same proportionality to the tagging power, but due to a different time dependence in the integrand in Equation (5.14), the sensitivity for the parameter C can be expected to be increased compared to the sensitivity on S, which is confirmed by the obtained result. An additional key influence on the relation between the relative sizes of σ_S and σ_C that is ignored in Equation (5.14) for simplicity is the shape of the time-dependent selection efficiency, which is modeled as documented in Section 6.3.

The tagging decisions and mistags of multiple taggers are statistically combined into combined tagging decisions and mistag estimates. The combined probability of simultaneously arriving at a set of tagging decisions $\boldsymbol{q}$ and mistags $\boldsymbol{\omega}$ given that the B meson truly contains a b or a $\overline{b}$ quark is given by the combined quark probabilities Ω_b and $\Omega_{\overline{b}}$ of N_T different taggers

$$\Omega_b(\boldsymbol{\omega}, \boldsymbol{q}) = \prod_j^{N_T} \begin{cases} \omega_j & \text{if } q_j = 1 \\ 1-\omega_j & \text{if } q_j = -1 \\ \frac{1}{2} & \text{else} \end{cases} \quad \text{and} \quad \Omega_{\overline{b}}(\boldsymbol{\omega}, \boldsymbol{q}) = \prod_j^{N_T} \begin{cases} 1-\omega_j & \text{if } q_j = 1 \\ \omega_j & \text{if } q_j = -1 \\ \frac{1}{2} & \text{else} \end{cases}. \tag{5.16}$$

The combined tagging decision and mistag is then determined by the larger of the two tagging probabilities

$$q_{\text{comb}} = \text{sign}(\Omega_b - \Omega_{\overline{b}}) \qquad \omega_{\text{comb}} = 1 - \frac{\max(\Omega_b, \Omega_{\overline{b}})}{\Omega_b + \Omega_{\overline{b}}}. \tag{5.17}$$

In the measurement described here, the SS and OS taggers are each calibrated and combined into two sets of tagging decisions and mistags for the combined SS and combined OS single taggers. Some fractions of data are exclusively tagged by the SS or the OS tagger, and for a fraction of candidates, decisions of both tagging sides are available. Candidates without an assigned tagging decision are not included as they do not add sensitivity to the measurement. Equations (5.10) to (5.12) are therefore only used in the fit in case either the SS or OS tagger has a decision, but not both. If both taggers have a decision, coefficients $\tilde{S}$, $\tilde{C}$, and $\tilde{H}$ for this case can be derived analogously and result in expressions with very similar properties as Equations (5.10) to (5.12). The derivation of the decay rate coefficients for this case is shown in Appendix E.2. With these definitions, the core of the pdf with which the CP-violation parameters are measured via ML fit is obtained from Equation (5.6) with the definitions from Equations (5.10) to (5.12) and (E.10) to (E.12) and can be written as

$$\mathcal{P}_{CP}(t, \boldsymbol{q}_j, \boldsymbol{\eta}_j) \propto e^{-\Gamma t/\hbar} \left[\tilde{H}(\boldsymbol{q}_j, \boldsymbol{\eta}_j) + \tilde{S}(\boldsymbol{q}_j, \boldsymbol{\eta}_j) \sin(\Delta m_d t) + \tilde{C}(\boldsymbol{q}_j, \boldsymbol{\eta}_j) \cos(\Delta m_d t)\right], \tag{5.18}$$

for the j-th event, whereby $\boldsymbol{q}_j = (q_j^{\text{SS}}, q_j^{\text{OS}})$ and $\boldsymbol{\eta}_j = (\eta_j^{\text{SS}}, \eta_j^{\text{OS}})$ are the two possible tagging decisions and mistag estimates of both tagging sides.

5.2 Selection of calibration modes

The outputs of individual taggers are a tagging decision q and a raw mistag estimate η for each individual candidate in the data set. From Equation (5.7) it follows that the mistag estimate needs to behave like a frequentist probability, in the sense that in a small mistag bin $[\eta_0 - \zeta, \eta_0 + \zeta]$ the rate of wrong tagging decisions should converge to η_0 as $\zeta \to 0^+$ in the high statistics limit. As the output of multivariate techniques does not necessarily have this property, a calibration function needs to be determined with appropriate calibration data sets. Furthermore, a calibration function can be determined on a weighted data set in which the kinematic distributions more closely resemble the target signal kinematics. Calibration data sets that are selected for this purpose have the property that the neutral or charged B flavor at the time of production is known with high certainty. The false classification rate of a tagger can then be compared with its actual false classification rate, and a functional relation between η and the mistag rates on the calibration sample, denoted as ω^{true} in the following can be determined.

For the calibration of OS taggers in the signal decay modes $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ and $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$ the calibration mode $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ is used, and the SS taggers are calibrated with the calibration mode $B^0 \to J/\psi(\to \mu^+\mu^-)K^*(\to K^+\pi^-)$. These specific channels are chosen for their kinematic similarity to the signal. Additionally, it is beneficial to calibrate the SS of the signal mode with a neutral B calibration channel. On the opposite side, there are no constraints for the type of created b-hadron and its specific decay, and for this reason, B decays have similar OS characteristics. While the OS can thus be calibrated with almost any suitable B decay, the $B^+ \to J/\psi K^+$ channel is chosen for its high available statistics. The taggers of the $B^0 \to J/\psi(\to e^+e^-)K^0_S$ mode are calibrated with similar calibration modes, with the difference that the J/ψ candidates are reconstructed from two electrons instead. The decays of the calibration modes are referred to as *flavor-specific* because their final state is charge asymmetric and is fully correlated to the B decay flavor. A complication arises thus in the $B^0 \to J/\psi K^*$ mode, as B^0 mesons can decay in the oscillated state, and it needs to be taken into account that their likely production flavor is diluted with time.

The selection requirements of the calibration channels $B^+ \to J/\psi(\to \ell^+\ell^-)K^+$ and $B^0 \to J/\psi(\to \ell^+\ell^-)K^*(\to K^+\pi^-)$ are as follows. Initial J/ψ candidates are selected by choosing trigger lines that target dimuon pairs with high reconstructed mass near the J/ψ mass, which originate from a vertex that is separated from the PV. For the selection of $J/\psi(\to e^+e^-)$ candidates, a more inclusive trigger selection is applied, similar to the signal selection described in Section 4.2. Two oppositely charged electrons or muons with transverse momentum $p_{\text{T}} > 500\,\text{MeV}/c$ are combined to form J/ψ candidates with a good quality common vertex and a reconstructed mass in a mass window $\pm 60\,\text{MeV}/c^2$ around the known J/ψ mass. Charged kaons and pions with $p_{\text{T}} > 500\,\text{MeV}/c$ with well-reconstructed tracks are required to form a good quality common vertex, and this K^* candidate is required to have transverse momentum above $1300\,\text{MeV}/c$ with a reconstructed mass in the range $826 \leq m(K^+\pi^-) < 966\,\text{MeV}/c^2$. For the single K^+ track of the $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ channel, a minimum transverse momentum $p_{\text{T}} > 1\,\text{GeV}/c$ and total momentum $p > 10\,\text{GeV}/c$ is required. The requirement for the

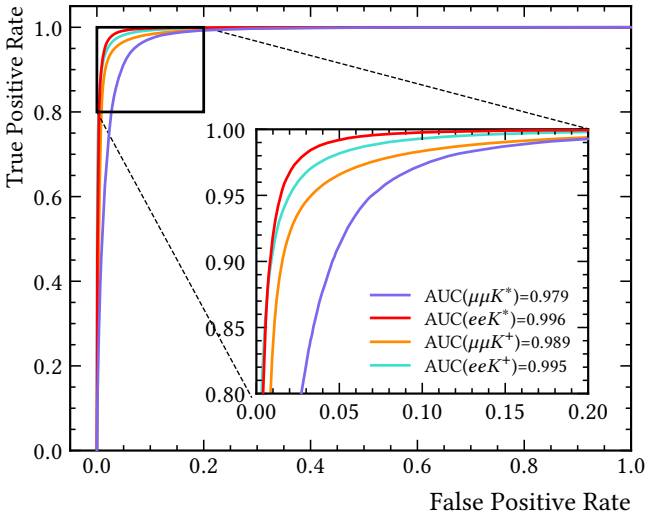

Figure 5.2 – Receiver operating characteristics for BDT classifiers trained to reduce the combinatorial background in four calibration data sets.

reconstruction of the B^0 is a good vertex fit, and the reconstructed B mass is required to be in the range $5150 \leq m < 5450\,\mathrm{MeV}/c^2$. In addition, a minimum lifetime cut of 0.2 ps is applied to B meson candidates to ensure that the tracks do not originate directly from the PV and an upper bound for the J/ψ-mass-constrained decay-chain fit is applied.

The selection requirements for the dielectron modes are similar to the dimuon modes, with looser mass range requirements for J/ψ ($2300 \leq m(e^+e^-) < 3300\,\mathrm{MeV}/c^2$) and B^0 ($4400 \leq m(e^+e^-\pi^+\pi^-) < 6000\,\mathrm{MeV}/c^2$) candidates. Finally, the only momentum requirement for the single K^+ track is $p_\mathrm{T} > 800\,\mathrm{MeV}/c$ and the p_T threshold for the K^* candidate is at $p_\mathrm{T} > 1500\,\mathrm{MeV}/c^2$. As is the case in the signal modes, the background level in the dielectron modes is significantly higher in comparison to the dimuon modes due to the more inclusive trigger selection. For this reason, the angle between the B decay vertex in relation to the PV and the reconstructed B track is required to be low in the neutral mode. In the charged B mode, this requirement is applied for the J/ψ candidate instead.

Apart from a partial background component and a $B_s^0 \to J/\psi K^*$ component, no additional background components are identified in the calibration data sets. To reduce combinatorial background, four BDT classifiers are trained, one for each calibration channel. The approach of training BDTs resembles the setup used for the signal modes closely, regarding the measures for reducing overtraining and maximizing generalizability (see Section 4.3). The upper mass sideband serves as a background proxy, while simulated calibration modes, which are selected with the same requirements as in data, serve as the signal proxy.

Training features for the BDT classifier training are chosen with similar criteria as in the signal modes: in simulation, the feature needs to be well-modeled and its distribution is required to differ in signal versus background-enhanced sample, which is verified with the help of sWeights. As is also seen in the signal modes, IP variables are

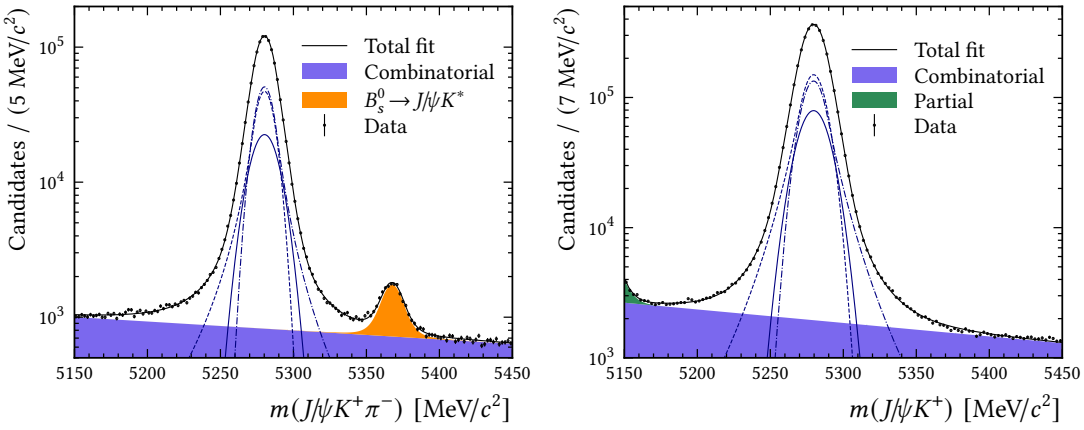

Figure 5.3 – Mass fits of the flavor tagging calibration channels $B^0 \to J/\psi K^*$ (left) and $B^+ \to J/\psi K^+$ (right). The three signal components are shown in dark blue: Left-tailed Crystal Ball function (dashed), right-tailed Crystal Ball (dashed-dotted), and a central Gaussian component (line).

seen to contribute most significantly to the separation between signal and background. In total, 12 training features are selected for the neutral calibration mode, and 18 features are chosen for the charged mode. In figure Figure 5.2 the ROC curves of each classifier are shown and the ROC curve areas imply good separation power between signal and background in all calibration data sets. In the dimuon calibration samples, the selection cut on the BDT output is optimized by choosing a cut for which the signal purity metric $N_{\text{Sig}}/(N_{\text{Sig}} + N_{\text{Bkg}})$ reaches 99.9 % (99 %) for the SS (OS) channel and a purity metric of 95 % is required for the dielectron calibration modes.

Signal weights are determined with the *sPlot* technique using the reconstructed B^0 and B^+ mass as discriminating variables. The B mass peaks are described in each case by the sum of a Gaussian distribution and two Crystal Ball distributions [136]. One distribution is defined as having an exponential tail towards lower masses and the other as a tail towards higher masses. The three signal components share a common mean parameter, and the combinatorial background is described by a single exponential function. The $B_s^0 \to J/\psi K^*$ component is modeled in an analogous way as in the signal mode $B^0_{(s)} \to \psi K^0_S$, by constraining the shape to the B^0 peak. The partial background is modeled with a single Gaussian distribution. In Figure 5.3, the mass fits of the control modes $B_s^0 \to J/\psi(\to \mu^+\mu^-)K^*$ and $B^+ \to J/\psi K^+$ are shown and from the result of this mass fit signal weights are computed with the *sPlot* method.

5.3 Kinematic reweighting of calibration channels

Each flavor tagging algorithm is tested on the flavor-specific calibration channels $B^0 \to J/\psi K^*$ and $B^+ \to J/\psi K^+$ and in a later step, these calibration parameters serve as inputs to the pdf in Equation (5.18) with which the CP-violation parameters are measured in the signal modes (Section 6.5). To improve the transferability of flavor

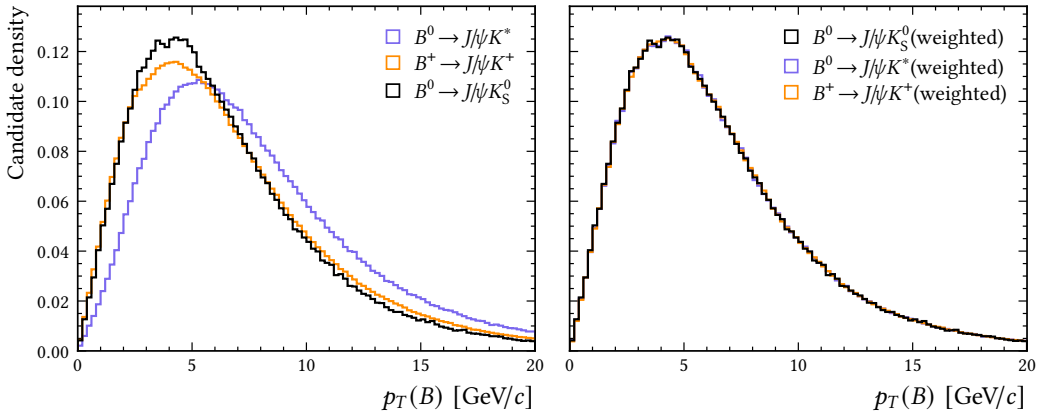

Figure 5.4 – Comparison of the transverse momentum distribution with (right) and without reweighting (left). The reweighted distributions are found to match the signal distribution (black) well when kinematic weights are applied.

tagging responses between modes, the calibration mode kinematics need to match the signal mode kinematics closely. A multivariate, BDT-based reweighting approach [137] is chosen to compute a set of event weights for each calibration data sample which aligns the data sets for a given set of features. The chosen set of features is the B transverse momentum, B-pseudorapidity, the number of reconstructed PVs and tracks per event, and the B-meson ϕ-angle around the beam axis, as flavor tagging algorithms are particularly sensitive to these variables. Figure 5.4 shows the good agreement of the B-p_T distribution when kinematic weights are applied. The reweighting is configured such that the signal weighted distribution in the calibration modes is reweighted to the background subtracted signal mode. This way, the reweighting is particularly optimized to reweight kinematic differences of the decay modes. Therefore, the product of sWeights and reweighting weights is the total candidate weight applied to the calibration channels for the following calibration steps.

The flavor tagging algorithms illustrated in Figure 5.1 are used: On the SS, a proton (SS-p) and a pion (SS-π) tagger and on the OS, a muon (OS-μ), electron (OS-e) and a kaon tagger (OS-K) are used which target high momentum particles created in weak $b \to c \to s$ transitions. In addition, a charm tagger selects flavor-specific D meson decays (OS-c) on the OS, and a vertex charge tagger (VtxCh) computes a momentum-weighted charge average of tracks originating from the secondary vertex on the OS. The SS taggers are calibrated with the calibration mode $B^0 \to J/\psi K^*$ since this is a B^0 decay with similar decay kinematics, and thus similar to the signal mode. The taggers of the electron signal modes are calibrated with corresponding calibration data sets of $J/\psi \to e^+e^-$ decays.

A tagger output is considered to be calibrated if its mistag distribution maximizes the tagging probability likelihood for a given set of calibration parameters $\boldsymbol{p}$. The tagging

likelihood is defined as

$$\mathcal{L}_{\mathrm{FT}}(\boldsymbol{p}) = \prod_{j}^{N} \begin{cases} 1 - \omega^{(f_j)}(\eta_j, \boldsymbol{p}) & \text{if } \; q_j = f_j \\ \omega^{(f_j)}(\eta_j, \boldsymbol{p}) & \text{else} \end{cases} \tag{5.19}$$

and assigns correctly tagged events, *i.e.* events for which the tagging decision coincides with the likely production flavor f_j, the tagging probability $1 - \omega$ and ω otherwise. The notation $\omega^{(f_j)}$ refers to the functions ω^+ or ω^-, depending on the likely production flavor, which describes the calibration function for samples of true B^0 or $\bar{B}^0$ mesons.

The production flavor in the case of the $B^+ \to J/\psi K^+$ mode coincides with its decay flavor as B^+ mesons do not mix, and the kaon charge is measured with high precision. In the calibration mode $B^0 \to J/\psi K^*$ the B mixing probability needs to be taken into account, which can be derived from the time-dependent mixing asymmetry $A_{\mathrm{mix}}(t)$, which equals $\cos(\Delta m_d t)$ to a high degree of precision. Calibration events for which $A_{\mathrm{mix}}(t) < 0$ are more likely to have undergone mixing than not, and their likely production flavor is assumed to be the inverse of their decay flavor: $f_j = \mathrm{sgn}(\cos(\Delta m_d t)) \cdot f_{\mathrm{decay,\,i}}$. The residual probability that the likely production flavor is wrong for this assumption is given by the quantity

$$\gamma(t) \equiv \frac{1}{2}\left(1 - |\cos(\Delta m_d t)|\right). \tag{5.20}$$

and the time-dependent tagging likelihood is then defined as

$$\mathcal{L}_{\mathrm{FT}}(\boldsymbol{p}) = \prod_{j}^{N} \begin{cases} (1-\gamma_j)(1-\omega^{(f_j)}(\eta_j, \boldsymbol{p})) + \gamma_j \omega^{(-f_j)}(\eta_j, \boldsymbol{p}) & \text{if } \; q_j = f_j \\ (1-\gamma_j)\omega^{(f_j)}(\eta_j, \boldsymbol{p}) + \gamma_j (1-\omega^{(-f_j)}(\eta_j, \boldsymbol{p})) & \text{if } \; q_j \neq f_j \end{cases} \tag{5.21}$$

with $\gamma_j \equiv \gamma(t_j)$ for brevity. The factor γ_j interpolates between tagging probabilities of flavor-specific calibration functions, and for $\gamma_j = 0$, Equation (5.21) is equal to the time-independent case in Equation (5.19). At the zeros of $A_{\mathrm{mix}}(t)$, the initial B flavor is fully diluted, and a reliable tag is not possible on the calibration sample. For these decay times, $\gamma_j = 1/2$ weighs both flavor specific calibrations equally.

5.4 Development of flavor tagging calibration software

The software framework "EspressoPerformanceMonitor" (EPM) [138], that was used for flavor tagging response calibrations by the LHCb collaboration in previous measurements of *CP* violation has been fully validated and rewritten in Python [139] and its functionality has been extended. The reimplementation effort is motivated by the EPM's rigid software design, which prohibits the implementation of certain new functionalities. Furthermore, software design choices severely impact its runtime performance, which extended the runtime of this measurement by multiple days. The primary reasons for this are its heavily object-oriented data storage approach and high

level of computational redundancy: the EPM tool needs to be executed up to six times to complete a full calibration of taggers. In addition, a separate configuration file is required for each run, in which mostly redundant information needs to be specified. In previous iterations of this measurement, thousands of calibration files needed to be generated with dedicated scripts to enable all cross-checks performed for flavor tagging for data and simulation. For each incremental execution step, the same data is read from disk and identical statistics are computed, despite only minimal changes to the computation task. In the course of the new implementation effort, it was possible to validate the output of the EPM tool for the first time by comparing its results with an alternative implementation. It can be confirmed that the EPM tool is essentially bug-free despite its considerable complexity and that results obtained with the EPM tool are valid if the framework is used exactly in the way it was intended.

The new software package `lhcb-ftcalib` [140], or *FTcalib*, is implemented in Python using only the basic, widely used, and well-tested numerical computation libraries numpy [141], pandas [142], scipy [143] and the tagging likelihood is optimized by iMinuit [144, 145]. A data-oriented design for data storage is chosen for FTcalib, and it is able to calibrate single taggers and their combinations in a single run. Despite being written in an interpreted language, it is faster by a factor of up to 50 compared to the EPM for a full calibration run of a single typical data set containing multiple tagging outputs. This time factor enabled the feasibility of multiple cross-checks and calibrations on all data and simulated subsamples, as the total flavor tagging calibration runtime was reduced from multiple days, despite parallel computation, to only a few hours, most of which is spent reading and writing to disk. An extensive unit testing framework ensures that all computations are valid and yield expected results, even in edge cases. Compatibility with the EPM is ensured by compiling and running the EPM as part of automated unit testing and comparing the calibrated data and the reported performance values for multiple scenarios. In collaboration with my colleague Quentin Führing, it was possible to implement exact error propagation across calibrations and tagging response combinations for the first time, which enables new types of systematic studies in some measurements.

5.5 Tagging response calibration

The individual tagging algorithm outputs are calibrated with a generalized linear model, as implemented in the EPM and in FTcalib, which is defined up to order m as

$$\omega^{\pm}(\eta, p_0^{\pm}, \ldots, p_m^{\pm}) = g\left(g^{-1}(\eta) + \sum_{i=0}^{m} p_i^{\pm} P_i(\eta)\right) \tag{5.22}$$

whereby $P_i(\eta)$ are orthogonal basis polynomials of the transformed mistag distribution

$$P_i(\eta) = \sum_{k=0}^{m} B_{ik} g^{-1}(\eta)^{m-k-1}, \tag{5.23}$$

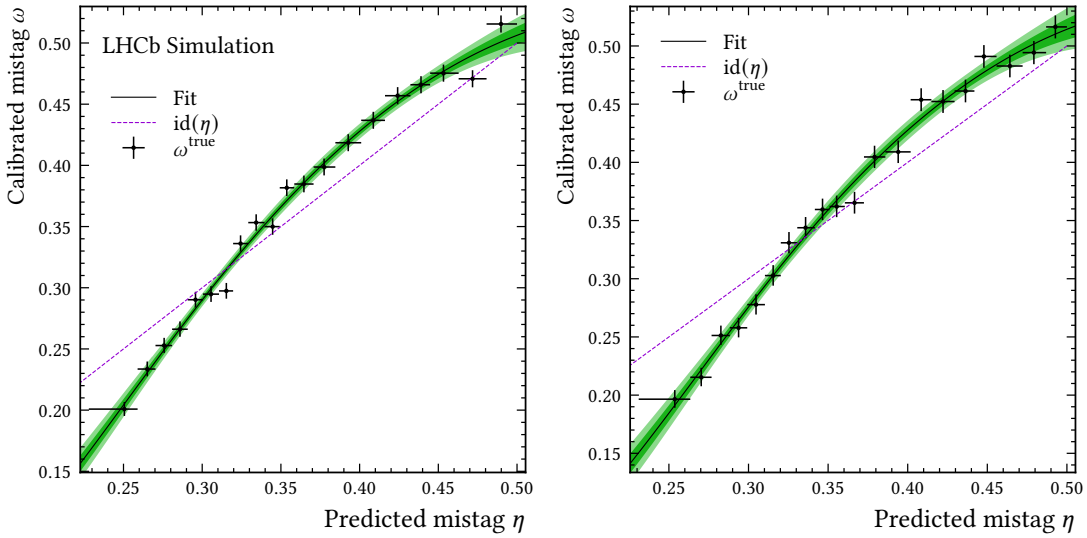

Figure 5.5 – OS-e tagger calibrations in $B^+ \to J/\psi K^+$ simulation (left) and in the $B^+ \to J/\psi K^+$ channel in data (right).

with basis matrix $\boldsymbol{B}$ and linear calibration parameters $p_i^{\pm}$. The link function $g(\eta)$ is a nonlinear transformation of the mistag and can be used to model the true mistag rate in a transformed coordinate system. The choice of orthogonal basis polynomials has the advantage that the calibration parameters are stable against an increase of the order of the function. Furthermore, due to the linearity of ω with respect to the calibration parameters, the previous definitions Equations (5.8) and (5.9) hold for arbitrary $g(\eta)$ and expansion orders m. In analogy to Equation (5.8), a useful equivalent set of calibration parameters can be defined

$$p_i = \frac{p_i^+ + p_i^-}{2} \qquad \Delta p_i = p_i^+ - p_i^- \tag{5.24}$$

to describe the calibration functions of the mean of B flavor subsets and their difference. In addition, $p_i = \Delta p_i = 0$ can always be interpreted as being the identity calibration such that each p_i accurately measures the contributions of powers of η. More details concerning the chosen formalism and its derivation are given in Ref. [146].

For the calibration of single taggers $m = 2$ and $g(\eta) = (1 + e^{\eta})^{-1}$ is chosen as the link function, and the candidates are weighted by the product of the sWeight and kinematic weight. Figure 5.5 show the calibration result of the OS-e tagger, where the nonlinearity of $\omega(\eta)$ is apparent. The agreement between the tagging response in simulation is seen to be remarkably similar to the tagging response on data. After each tagger is calibrated, the OS-taggers (OS-e, OS-μ, OS-K, OS-c, and VtxCh) and SS-taggers (SS-π and SS-p) are each combined into an SS and an OS combination. These two combinations are calibrated again using a calibration function that is linear in η. Calibrations of higher order were found not to be advantageous. The chosen calibration function is

$$\omega^{\pm}(\eta, p_0^{\pm}, p_1^{\pm}) = \eta + p_0^{\pm} + p_1^{\pm}(\eta - \langle\eta\rangle), \tag{5.25}$$

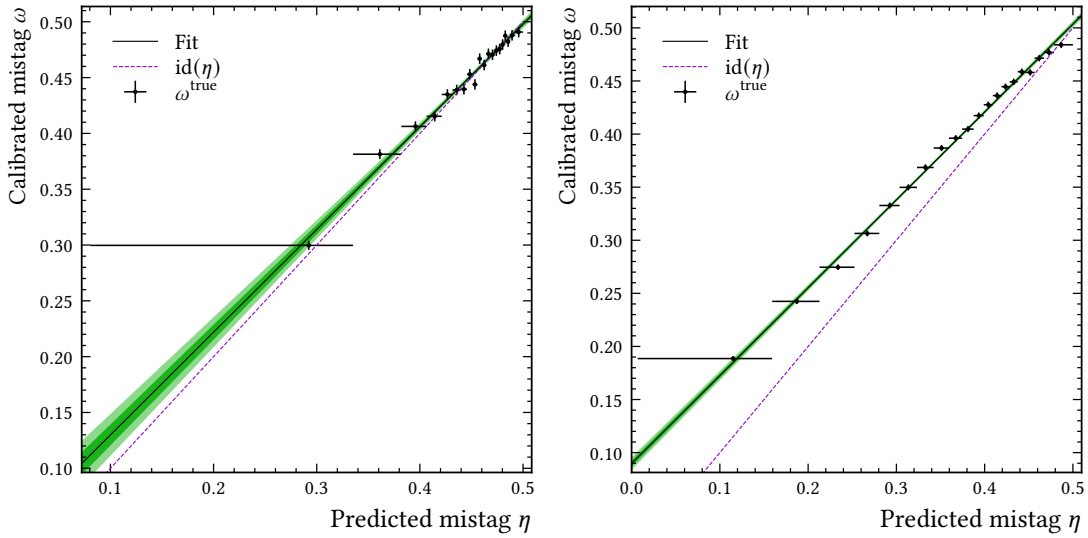

Figure 5.6 – Calibration curves of the SS and OS tagger combinations in the calibration channels $B^0 \to J/\psi K^*$ (left) and $B^+ \to J/\psi K^+$ (right).

which corresponds to $g(\eta) = \eta$ and $m = 1$ and $\langle\eta\rangle$ is the weighted mean mistag. The linear fit showing the calibration of the combined SS and OS combination is shown in Figure 5.6. As is explained in the next section, the flavor tagging calibrations described here are an intermediate step towards a final set of calibration parameters, and therefore, calibration results and detailed performance numbers are given in Section 5.6. The tagging performances found for the single tagging algorithms are listed in Table 5.1. Tagging performance numbers in the simulation are not shown here and are considerably higher for each single tagging algorithm, which is a well-known shortcoming of simulation. A possible cause of increased performance in simulation is that an accurate replication of the track multiplicity remains a challenging task, and it is usually underestimated in simulation. Furthermore, while the simulation of the signal decay is quite precise, the accurate simulation of the multitude of interactions that define an event and the associated detector response is challenging. In addition, the particle identification performance is known to be significantly overestimated in simulation for the raw, neural-network-based particle identification probabilities, which flavor tagging relies on heavily. For these reasons, while agreement with simulation is usually tested, potential corrections for CP-violation measurements are determined on data or from probability distributions sampled from real data.

The total tagging power of the merged signal sample of all three decays is found to be $\epsilon_{\text{tag}}\mathcal{D}^2_{\text{FT}} = (3.978 \pm 0.011)\,\%$ for the subset of data containing at least one tagging decision from the SS or OS. With this result, the expected statistical sensitivity for the CP-violation parameter S in the decay mode $B^0 \to J/\psi(\to \mu\mu)K^0_S$ in Run 2 can be predicted on the basis of the reported Run 1 result for the same mode:

$$\sigma_S^{\text{Run 2}} \approx \sigma_S^{\text{Run 1}} \sqrt{\frac{[\epsilon_{\text{tag}} \cdot \mathcal{D}^2_{\text{FT}} \cdot N_B]_{\text{Run 1}}}{[\epsilon_{\text{tag}} \cdot \mathcal{D}^2_{\text{FT}} \cdot N_B]_{\text{Run 2}}}} = 0.017\,17 \pm 0.000\,16. \tag{5.26}$$

Table 5.1 – Tagging performances of individual tagging algorithms as measured in the three signal modes after the calibrations have been transferred from the calibration channels and applied. The first uncertainty of the tagging power is statistical, and the second uncertainty quantifies the uncertainty of the calibration function.

Tagger	$\epsilon_{\text{tag}} \cdot \mathcal{D}^2_{\text{FT}}$ [%]	ϵ_{tag} [%]
	$B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$	
OS-c	$0.345 \pm 0.003 \pm 0.009$	4.34 ± 0.035
OS-e	$0.313 \pm 0.004 \pm 0.007$	3.10 ± 0.030
OS-K	$0.993 \pm 0.006 \pm 0.014$	18.77 ± 0.07
OS-μ	$0.805 \pm 0.007 \pm 0.012$	5.59 ± 0.04
VtxCharge	$0.992 \pm 0.005 \pm 0.014$	18.54 ± 0.07
SS-π	$0.905 \pm 0.004 \pm 0.029$	73.91 ± 0.08
SS-p	$0.374 \pm 0.003 \pm 0.02$	38.29 ± 0.08
	$B^0 \to J/\psi(\to e^+e^-)K^0_S$	
OS-c	$0.395 \pm 0.011 \pm 0.045$	4.98 ± 0.11
OS-e	$0.348 \pm 0.013 \pm 0.043$	4.21 ± 0.10
OS-K	$1.319 \pm 0.019 \pm 0.083$	18.42 ± 0.19
OS-μ	$1.082 \pm 0.024 \pm 0.07$	7.33 ± 0.13
VtxCharge	$1.342 \pm 0.019 \pm 0.083$	19.04 ± 0.19
SS-π	$1.492 \pm 0.014 \pm 0.155$	82.24 ± 0.19
SS-p	$0.808 \pm 0.011 \pm 0.123$	38.86 ± 0.24
	$B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$	
OS-c	$0.341 \pm 0.012 \pm 0.008$	4.33 ± 0.13
OS-e	$0.293 \pm 0.015 \pm 0.007$	2.94 ± 0.11
OS-K	$0.974 \pm 0.02 \quad \pm 0.014$	18.60 ± 0.24
OS-μ	$0.767 \pm 0.026 \pm 0.011$	5.44 ± 0.14
VtxCharge	$0.985 \pm 0.019 \pm 0.014$	18.52 ± 0.24
SS-π	$0.937 \pm 0.015 \pm 0.029$	74.37 ± 0.27
SS-p	$0.386 \pm 0.009 \pm 0.02$	38.23 ± 0.31

For this estimate, the values $\epsilon_{\text{tag}} = (36.54 \pm 0.16)\,\%$, $\mathcal{D}_{\text{FT}} = (28.75 \pm 0.24)\,\%$ and $\sum_j |q_j| = 41\,560 \pm 270$ for the combined 2011 and 2012 analysis [13], implying $N_{B,\text{Run 1}} = 113\,738 \pm 891$, are used. A reported squared flavor tagging dilution for Run 2, $\mathcal{D}^2_{\text{FT}} = (4.661 \pm 0.013)\,\%$, and a tagged signal yield of $\epsilon_{\text{tag}} \cdot N_B = 306\,090 \pm 570$ is inserted [15], which is the sum of sWeights in the mass-range used in the fit. This sensitivity estimate is exceeded in this measurement: $\sigma_S^{\text{Run 2}} = 0.015$ (see Chapter 8). The achieved tagging power for Run 2 is higher than the flavor tagging performance reported for Run 1, primarily for the reason that the Run 2 tagging algorithms have been better optimized. Furthermore, SS taggers have not been available during the 2011 data-taking phase. As this sensitivity estimate already takes these improvements into account, it implies a significant mismatch of fit quality. One can conclude that the measurement performed in this Run 2 LHCb analysis is comparable, if not higher, in precision. This improvement could be attributed, in part, to the improved background rejection achieved with the BDT selection approach. In the previous measurement, the background distributions in the selected data set are explicitly modeled as part of the decay time description. Furthermore, as the center-of-mass energy of the LHC has been increased for the Run 2 data-taking phase, differences in the effective detector acceptance could have contributed to more precise measurements of particle momenta and identity. Finally, advances in event reconstruction and particle identification achieved by the LHCb collaboration after Run 1 have likely contributed to this improvement.

5.6 Determination of calibration parameters and measurement of flavor asymmetries

With the calibration results of the SS and OS combinations, it is conceptually possible to measure $\mathcal{A}^{CP}(t)$ in a time-dependent ML fit by applying these calibrations to the signal modes with a suitable decay time model (Equation (5.18), which is extended in Sections 6.1 and 6.3). This analysis approach has been refined in order to be able to consider a wider range of parameter correlations, primarily the correlation between Δp_i and α, as both are measured on the same data set. Reassuringly, the impact on the final result was seen to be minimal, and the approach was adopted for its improved consistency.

In this refined approach, the final set of calibration parameters is determined as follows: the calibration parameters for the SS tagger combination are determined in a time-dependent tagged fit of the mode $B^0 \to J/\psi(\to \ell^+\ell^-)K^*$ and the combined OS calibrations of this mode are Gaussian constrained to the results obtained previously with the calibration channel $B^+ \to J/\psi(\to \ell^+\ell^-)K^+$ to maximize the statistical precision. This fit is repeated for each final state and associated candidate weights. For the decay $B^0 \to J/\psi(\to e^+e^-)K^0_S$, the calibration parameters are determined with a time-dependent fit to the $B^0 \to J/\psi(\to e^+e^-)K^*$ calibration data set. In these time-dependent fits, the SS and OS calibration parameters, the production asymmetry, and flavor tagging efficiency

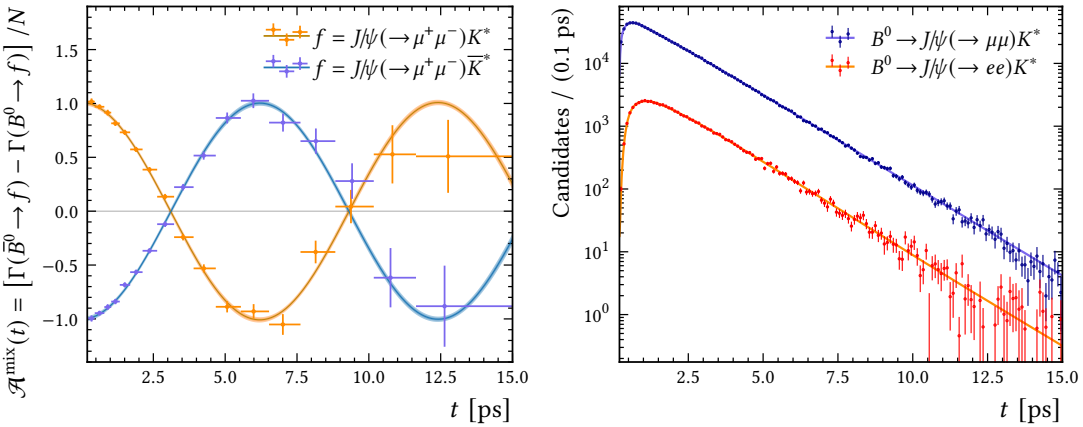

Figure 5.7 – Normalized, time-dependent final-state charge asymmetry of the decay $B^0 \to J/\psi K^*(\to K^+\pi^-)$ (left) and decay time projection of the time-dependent fits to the two calibration data sets (right).

asymmetries are determined simultaneously so that the correlations between all these parameters can be taken into account in the final fits.

In the decay mode $B^0 \to J/\psi K^*$, two non-interfering decay amplitudes lead to two different final state configurations depending on the B decay flavor. CP violation has not been discovered in this mode [56], and it can therefore be assumed here that the asymmetry of B^0 and $\overline{B}^0$ decays is zero for any time interval and thus $S = \overline{S} = 0$. Because of the flavor-specific final states, the mixing asymmetry should be maximal $C = -\overline{C} = 1$. Four decay rates contribute to this decay:

$$\Gamma(B^0(t) \to J/\psi K^*) \propto e^{-\Gamma t}(1-\alpha)(1+A_{\mathrm{d}})[H + C\cos(\Delta m_d t) - S\sin(\Delta m_d t)], \quad (5.27)$$

$$\Gamma(B^0(t) \to J/\psi \overline{K}^*) \propto e^{-\Gamma t}(1-\alpha)(1-A_{\mathrm{d}})[\overline{H} + \overline{C}\cos(\Delta m_d t) - \overline{S}\sin(\Delta m_d t)], \quad (5.28)$$

$$\Gamma(\overline{B}^0(t) \to J/\psi K^*) \propto e^{-\Gamma t}(1+\alpha)(1+A_{\mathrm{d}})[H - C\cos(\Delta m_d t) + S\sin(\Delta m_d t)], \quad (5.29)$$

$$\Gamma(\overline{B}^0(t) \to J/\psi \overline{K}^*) \propto e^{-\Gamma t}(1+\alpha)(1-A_{\mathrm{d}})[\overline{H} - \overline{C}\cos(\Delta m_d t) + \overline{S}\sin(\Delta m_d t)], \quad (5.30)$$

with A_{d} being the detection asymmetry and $H = \overline{H} = 1$. As is expected in this decay mode, and according to these decay rates, the main sensitivity to the amplitudes S, C, $\overline{S}$ and $\overline{C}$ stems from the correct charge identification of the final state: in the subsamples of decays to $J/\psi K^*$ and $J/\psi \overline{K}^*$, statistically independent mixing asymmetries can be measured, which differ by a sign. As demonstrated in Section 5.1, flavor tagging performance and all kinds of flavor asymmetries can be accounted for by deriving common decay rate coefficients $\tilde{S}(q,\eta)$, $\tilde{C}(q,\eta)$ and $\tilde{H}(q,\eta)$. In the decay mode $B^0 \to J/\psi K^*$, three coefficients per final state f; $\tilde{H}_f(q,\eta)$, $\tilde{S}_f(q,\eta)$ and $\tilde{C}_f(q,\eta)$ are needed. To that end, both possible decay amplitudes for B^0 (Equations (5.27) and (5.28)) and $\overline{B}^0$ (Equations (5.29) and (5.30)) decays are substituted into Equation (5.7) and rearranged to get

$$\tilde{C}_f(q,\eta) = -(1+fA_{\mathrm{d}})C_f \begin{cases} \tilde{C}(q,\eta) & \text{if } |q| = 1 \\ \Delta\epsilon - 2\alpha(1-\epsilon_{\mathrm{tag}}) & \text{if } q = 0 \end{cases}, \tag{5.31}$$

$$\tilde{H}_f(q,\eta) = \quad (1+fA_{\mathrm{d}}) \quad \begin{cases} \tilde{H}(q,\eta) & \text{if } |q| = 1 \\ 2(1-\epsilon_{\mathrm{tag}}) - \alpha\Delta\epsilon & \text{if } q = 0 \end{cases}, \tag{5.32}$$

$$\tilde{S}_f(q,\eta) = -\frac{\tilde{C}_f(q,\eta)}{C_f} S_f. \tag{5.33}$$

The final state factors $f = 1$ and $f = -1$ corresponds to the final states $J/\psi K^*$ and $J/\psi \overline{K}^*$, and to the parameter sets (C, S) and $(\overline{C}, \overline{S})$, respectively. The influence of flavor tagging information merely serves as a correction to the underlying production and detection asymmetries.

When constraining the OS calibration parameters to the result obtained previously in the decay mode $B^+ \to J/\psi K^+$, sufficient sensitivity is achieved to determine the SS calibration parameters with the same precision as by optimizing the flavor tagging likelihood. The inclusion of untagged data provides sensitivity to measure all flavor asymmetries $\Delta\epsilon$, $\Delta\omega$ [1], α and A_{d} simultaneously. Figure 5.7 shows the mixing asymmetry in the $B^0 \to J/\psi K^*$ mode if the amplitudes S and C, $\overline{S}$ and $\overline{C}$ are allowed to vary in the fit, and in the right plot, the decay-time projections of the two neutral calibration modes are shown. The averaged amplitudes are indeed found to be compatible with zero:

$$\frac{S+\overline{S}}{2} = -0.001 \pm 0.011 \qquad \frac{C+\overline{C}}{2} = -0.003 \pm 0.009, \tag{5.34}$$

and the baseline assumptions $\langle S \rangle = \langle C \rangle = 0$ thus seem well justified.

Table 5.2 – Tagging performance summary in each analyzed signal channel as given in Ref. [15]. The given tagging efficiencies differ from the ones given in Table 4.4 to a small degree as they are computed directly in the *CP*-violation fit routine on the basis of the fully selected data.

Channel	ϵ_{tag} [%]	$\mathcal{D}^2_{\mathrm{FT}}$ [%]
$B^0 \to J/\psi(\to \mu^+\mu^-)K^0_{\mathrm{S}}$	85.34 ± 0.05	4.661 ± 0.013
$B^0 \to J/\psi(\to e^+e^-)K^0_{\mathrm{S}}$	92.20 ± 0.08	6.462 ± 0.032
$B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_{\mathrm{S}}$	84.81 ± 0.15	4.59 ± 0.04

This result simultaneously serves as a useful validation of the baseline fit model because the model used here almost exactly resembles it. The baseline fit model is introduced in Section 6.5 in more detail. Furthermore, it serves as a good validation of flavor tagging, as it is directly performed on data.

In these fits, the OS parameters are Gaussian constrained to their central values with their common covariance matrix, as determined by FTcalib. The final set of calibration parameters and tagging asymmetries is summarized in Table 5.3. The production and

[1] The parameter $\Delta\omega$ is a function of the calibration parameters Δp_i and the mistag η.

Table 5.3 – Flavor tagging calibrations and tagging asymmetries resulting from time-dependent fits to the $B^0 \to J/\psi(\to \ell^+\ell^-)K^*$ calibration mode using candidate weights for each signal mode.

Parameter	$B^0 \to J/\psi K^0_S$ $J/\psi \to \mu^+\mu^-$	$B^0 \to J/\psi K^0_S$ $J/\psi \to e^+e^-$	$B^0 \to \psi(2S) K^0_S$ $\psi(2S) \to \mu^+\mu^-$
p_0^{SS}	0.0010 ± 0.0008	0.0024 ± 0.0030	0.0009 ± 0.0008
p_1^{SS}	-0.060 ± 0.016	-0.08 ± 0.05	-0.060 ± 0.016
Δp_0^{SS}	0.0041 ± 0.0012	0.006 ± 0.004	0.0040 ± 0.0011
Δp_1^{SS}	-0.052 ± 0.025	-0.06 ± 0.07	-0.047 ± 0.024
$\langle\eta\rangle^{SS}$	0.4499	0.4351	0.4491
$\Delta\epsilon^{SS}$	-0.0002 ± 0.0011	0.003 ± 0.004	-0.0001 ± 0.0011
p_0^{OS}	0.0272 ± 0.0006	0.0305 ± 0.0026	0.0273 ± 0.0006
p_1^{OS}	-0.172 ± 0.005	-0.187 ± 0.024	-0.173 ± 0.005
Δp_0^{OS}	0.0125 ± 0.0010	0.007 ± 0.005	0.0125 ± 0.0010
Δp_1^{OS}	0.035 ± 0.010	0.02 ± 0.04	0.034 ± 0.010
$\langle\eta\rangle^{OS}$	0.3616	0.3447	0.3616
$\Delta\epsilon^{OS}$	-0.0011 ± 0.0013	0.006 ± 0.005	-0.0010 ± 0.0013

detection asymmetries are determined in the same fit:

$$\alpha = -0.0026 \pm 0.0019 \qquad A_d = 0.011 \pm 0.001. \tag{5.35}$$

Correlation coefficients between calibration parameters and flavor asymmetries are observed to be small, except for the correlation coefficient between α and A_d, which is near 50 %. The tagging performances achieved for each final state are listed in Table 5.2.

6 Time-dependent fit description

At the core of the decay-time model describing the time-dependent influence of *CP* violation on the $B^0 \to \psi K^0_S$ decays is the model shown in Equation (5.18). This model is extended in the following to account for additional experimental effects. Most significantly, the reconstruction efficiency is time-dependent, and for very short decay times, it approaches zero due to selection requirements applied in the centralized preselections. Furthermore, the precision of the decay-time reconstruction is determined by the energy and vertex resolution of the LHCb detector and can be described by a Gaussian resolution model centered around $t = 0$. To account for the time-dependent reconstruction efficiency, an acceptance model is multiplied to the final model. As B^0 and $\bar{B}^0$ mesons have the same efficiency dependence due to their charge-symmetric final state, no net effect on *CP*-violation observables can be expected. Finally, external parameters like the production asymmetry and flavor tagging calibrations have uncertainties that are propagated to the final fit through Gaussian constraints.

6.1 Decay-time resolution modelling and calibration

The decay-time uncertainty of each candidate is determined in the reconstruction step, and it results from the variable precision of the decay chain fits. The precision of these fits is, in turn, determined by the finite resolution of the LHCb detector. To correct the measured observables according to the expected influence of the decay-time resolution of each individual candidate, a decay-time resolution model $R(\Delta t, \sigma_t)$ is determined. To that end, the true decay time of each candidate, t_{true}, in the simulated data set is subtracted from the reconstructed decay time t_{reco} after the detector simulation and event reconstruction steps have been performed. The resulting difference-distribution $\Delta t = t_{\text{true}} - t_{\text{reco}}$ is modeled with a weighted sum of Gaussian distributions and is a function of the measured decay-time uncertainty σ_t:

$$R(\Delta t, \sigma_t) = \sum_{j=1}^{3} f_j \frac{1}{\sqrt{2\pi}\sigma_j(\sigma_t, \boldsymbol{c}_j)} \exp\left(-\frac{\Delta t^2}{2\sigma_j^2(\sigma_t, \boldsymbol{c}_j)}\right). \tag{6.1}$$

The width-function of component j is given by $\sigma_j(\sigma_t, \boldsymbol{c}_j) = c_{1,j}\sigma_t + c_{2,j}$, with two calibration parameters $c_{1,j}$ and $c_{2,j}$ per component. The fractions f_j weigh each Gaussian component and ensure proper normalization at the same time. The decay-time-resolution model therefore has a total of eight free parameters, all of which are determined in fits to simulated data. In fits to data, these parameters are set constant to their central values. As the influence of the decay-time resolution on the result is so small, no

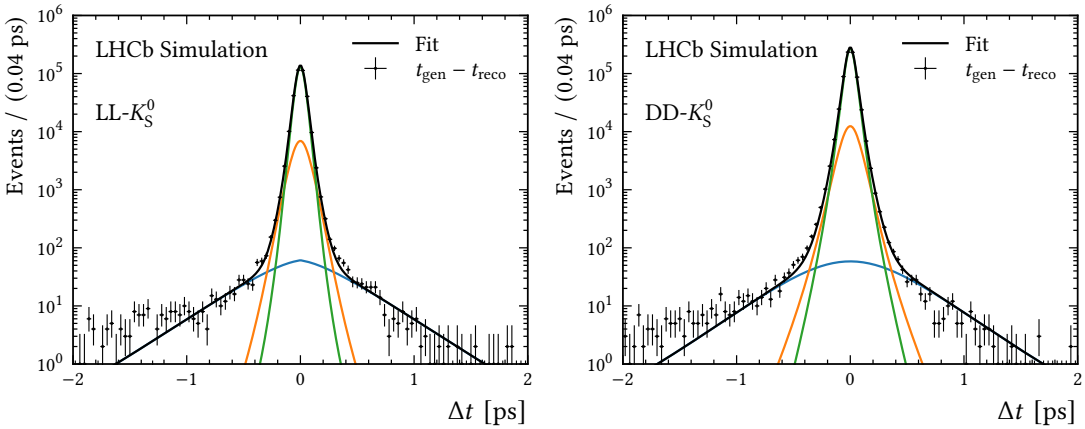

Figure 6.1 – Fits of the difference between generated and reconstructed decay time for the kaon reconstructions LL and DD in $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ decays.

Table 6.1 – Effective decay-time dilution $\mathcal{D}_t$ as determined in each simulated analysis channel and track category in percentages.

Track type	$B^0 \to J/\psi K^0_S$ $J/\psi \to \mu^+\mu^-$	$B^0 \to J/\psi K^0_S$ $J/\psi \to e^+e^-$	$B^0 \to \psi(2S) K^0_S$ $\psi(2S) \to \mu^+\mu^-$
LL	99.9397 ± 0.0016	99.956 ± 0.005	99.9501 ± 0.0020
DD	99.9607 ± 0.0010	99.9656 ± 0.0017	99.9639 ± 0.0014
LD	99.9370 ± 0.0026		
UL	99.922 ± 0.006		

systematic uncertainties associated to its parameter uncertainties are assigned. In Figure 6.1, fits of this model are shown for the track type reconstructions LL and DD in $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ simulation.

The decay-time resolution model acts as a Gaussian convolution kernel in the time-dependent fits, and its effect on a time-dependent CP asymmetry can be estimated analytically by convolving the B decay rate function with a single Gaussian kernel $\mathcal{N}(\Delta t, \sigma_t)$ of width σ_t:

$$\begin{aligned}\left[\Gamma(\overset{(-)}{B}(t') \to f) * \mathcal{N}(t', \sigma_t)\right](t) \propto \exp\left(-\frac{2t\tau + \sigma_t^2(\Delta m_d^2 \tau^2 - 1)}{2\tau^2}\right) \times \\ \left[\exp\left(\frac{\Delta m_d^2 \sigma_t^2}{2}\right) \mp S \sin(\Delta m_d (t - \delta_t)) \pm C \sin(\Delta m_d (t - \delta_t))\right].\end{aligned} \tag{6.2}$$

For simplicity, the convolution is performed without setting the decay rate to zero in the $t < 0$ region. This modification of the decay rates causes the time-dependent CP asymmetry in Equation (4.1) to take on the form

$$\mathcal{A}^{CP}(t) = e^{-\frac{1}{2}\Delta m_d^2 \sigma_t^2} \left[S \sin(\Delta m_d (t - \delta_t)) - C \cos(\Delta m_d (t - \delta_t))\right]. \tag{6.3}$$

Table 6.2 – Effective decay-time resolution width in femtoseconds in each analysis channel and track category.

Track Type	$B^0 \to J/\psi K^0_S$ $J/\psi \to \mu^+\mu^-$	$B^0 \to J/\psi K^0_S$ $J/\psi \to e^+e^-$	$B^0 \to \psi(2S) K^0_S$ $\psi(2S) \to \mu^+\mu^-$
LL	63.6 ± 1.5	59.4 ± 2.9	61.4 ± 2.0
DD	63.5 ± 0.9	58.9 ± 2.1	57.4 ± 1.0
LD	67.4 ± 2.2		
UL	71.1 ± 3.3		

The effect of a finite decay-time resolution is therefore two-fold: the amplitude of the *CP*-asymmetry is attenuated by a constant factor which can be approximated by the decay-time dilution factor $D_t = \exp(-\frac{1}{2}\Delta m_d^2 \sigma_t^2)$ and the *CP*-asymmetry receives an insignificant phase shift of $\Delta m_d \delta_t = \Delta m_d \sigma_t^2 / \tau$. Due to the non-linearity of the decay-time dilution with respect to the resolution width, wider Gaussian components of the resolution model shown in Equation (6.1) have a disproportionally larger influence on the result which is not quantified by the component fractions f_i alone. The effective dilution factor $\mathcal{D}_t$ is therefore computed in a resolution-model independent way through the use of Monte Carlo integration[1]. With an estimate of the effective decay-time dilution, an estimate of the detector resolution can be computed if it was described by a single Gaussian component:

$$\langle \sigma_t \rangle = \frac{1}{\Delta m_d} \sqrt{-\log \mathcal{D}_t^2}. \tag{6.4}$$

The effective dilution factors of each resolution model are listed in Table 6.1 and are very close to unity in this measurement due to the good decay-time resolution of the LHCb detector and a B oscillation period which is larger by two orders of magnitude. In Table 6.2, the mean resolution widths computed per Equation (6.4) are listed per final state and track category for all fitted models. The effective time-resolution of the detector ranges from $57 < \langle \sigma_t \rangle \leq 71\,\text{ps}$, depending on the final state and track category. The category UL has the largest resolution width of all data subsets, which is due to the reduced kaon momentum information of upstream pions. Overall, the decay-time resolution widths do not differ much between data subsets, which should be expected as the reconstruction of the dimuon or dielectron vertex is the most important contribution to the precise measurement of the B decay vertex. In addition, as the impact of the decay-time resolution on the amplitude increases exponentially with the resolution width, the widest components of the models have a disproportionally large influence on the observed *CP*-asymmetry amplitude. For this reason, the effective LL and DD resolution widths are almost identical, even though the LL candidates are reconstructed with higher precision on average (compare Figure 6.1).

[1] See Ref. [147] for more details

A phase shift of the *CP*-asymmetry of $\Delta m_d \sigma_t^2/\tau \approx 2\,\text{fs} \cdot \Delta m_d$ is caused by a mean resolution width of $\langle \sigma_t \rangle \approx 60\,\text{fs}$. This offset is smaller by one order of magnitude than the bias caused by small VELO misalignments, and its description is accounted for by the same model (see Section 6.2). Due to its small size, this time shift can only have a negligible effect on the measured amplitudes and a test in which a free time offset parameter is added to the decay time produced results in perfect agreement with the baseline results.

6.2 Modelling of VELO misalignment bias

The reconstructed decay time can be further systematically biased for measurements of reconstructed data due to small detector misalignments, especially in the VELO. As the inner silicon-microstrip-detector arrays of the VELO are movable and their distance to the beam is changed frequently, their positions can become offset by up to a few micrometers in each spatial direction due to limitations of mechanical accuracy and detector-alignment precision [104]. More specifically, a VELO misalignment in the beam direction (z-axis) leads to a decay-time reconstruction bias.

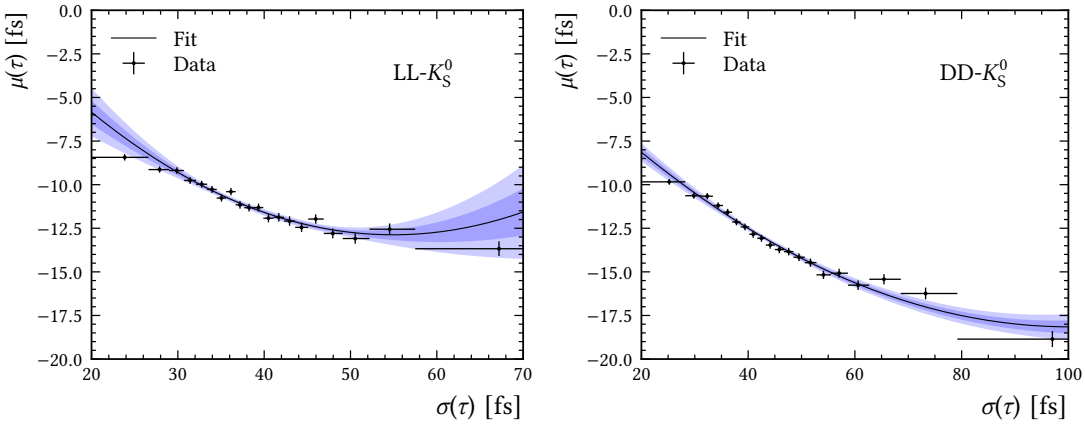

Figure 6.2 – Decay-time reconstruction bias in bins of the decay-time uncertainty, determined with a prompt $J/\psi(\to \mu^+\mu^-)K^0_S$ data sample for the track reconstruction categories LL (left) and DD (right).

This bias can be measured from fits to data of the *prompt* decay-time distribution, which includes J/ψ and K^0_S candidates emerging directly from the PV without an associated, genuine B origin particle. This data set is selected using the selection requirements described earlier, whereby direct requirements for the reconstructed B flight distance are omitted. The prompt decay-time distribution consists of several exponential components describing decays of real B mesons and various backgrounds. For genuine prompt candidates that do not originate from B mesons, a decay time of zero should be expected. Details regarding the model definition and fit results of the prompt decay-time fits can be found in Ref. [147]. Because of the finite detector resolution, this component contributes as a time-symmetric distribution near $t = 0$. A fit of this

distribution is therefore useful for describing the properties of the detector resolution and can be performed directly with reconstructed data. A decay-time reconstruction bias causes this distribution to be offset from $t = 0$ in addition, and this bias can then be accounted for by correcting it in the nominal data sets.

The decay-time reconstruction bias in prompt data is seen to vary in different bins of the decay-time uncertainty, so the chosen model describing this bias is defined as dependent on the decay-time uncertainty. A second-degree polynomial is therefore fitted to the dependence of the measured time bias μ_t in equally filled bins of the decay-time uncertainty $\sigma_{t_{\text{Bin}}}$. The fit results for prompt $J/\psi(\to \mu^+\mu^-)K^0_S$ data set are shown in Figure 6.2 for the LL and DD track categories. It can be seen that both calibration curves significantly deviate from a zero-bias assumption and that the bias ranges from approximately 10 to 15 fs. This calibration is performed individually for each of the four track-reconstruction categories and for each considered final state. This model is then used to correct the decay-time distributions in the data by applying it to the decay time of each reconstructed candidate. The decay-time resolution widths obtained in the fits to prompt data are found to be similar to widths present in the simulation, and the baseline approach of using simulation to describe the time resolution is therefore validated.

6.3 Time-dependent efficiency modelling

The B reconstruction efficiency is influenced by decay-time-biasing cuts in the event selection, where particularly short-lived B candidates are frequently discarded due to their proximity to the PV and thus higher rates of descendant-particle mismatches. The B decay-time distribution can therefore be expressed as a product of the exponential decay with a threshold function $\epsilon_{\text{acc}}(t)$, referred to as *decay-time acceptance* in the following. The decay-time acceptance is described by basic, one-dimensional cubic splines (B-splines) [148], which is a sum of a chosen number of third-degree spline basis polynomials, which define a continuous and smooth acceptance function in the total fit range. In a fit to the B decay-time distribution, the y-positions of this model are optimized, while the associated fixed time-positions remain constant. The time positions are chosen with the help of simulated data sets, and a set of seven time positions at $[0.2; 0.4; 0.9; 1.6; 5; 10; 15]$ ps was found, which is seen to be appropriate for every analyzed final state. The last spline y-position is set constant as its value is a normalization constant of the acceptance model and does not affect the result. Figure 6.3 shows a fit of the decay-time distribution of $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ candidates and a plot of the optimized decay-time models, illustrating that $\epsilon_{\text{acc}}(t)$ accurately describes the ratio of reconstructed decay time and a purely exponential distribution. This description is further seen to be accurate in individual years of data-taking.

For the time-dependent fits of the calibration channel $B^0 \to J/\psi(\to \ell^+\ell^-)K^*$, which are described in Section 5.6, the same approach is chosen for the modeling of the time-dependent acceptance. A total of ten spline components are required for an accurate fit due to the large sample sizes of the calibration data sets.

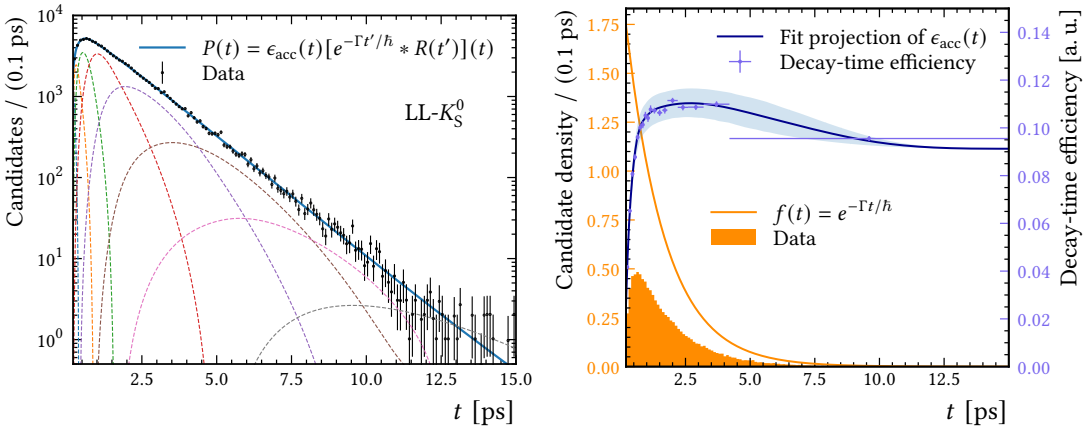

Figure 6.3 – Left: Decay time fit of the decay $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ in the LL category. The seven spline basis functions multiplied with the decay model are shown as dotted lines and sum up to the total acceptance model. Right: Ratio (blue fit curve) of reconstructed data (orange histogram) and an exponential function (orange line) of the same fit.

6.4 Estimation of kaon regeneration corrections

To account for the effect of kaon regeneration and CP violation in the neutral-kaon system, for each reconstructed K^0_S-track in the data set, the detector materials encountered by that particle are determined with a detailed material description of the Run 2 LHCb detector. This model also serves as a detector description for standard LHCb simulation and contains precise descriptions of the atomic composition of the installed materials. The kaon amplitudes given in Equations (2.63) and (2.64) are computed at the kaon decay vertices in an iterative procedure for each encountered detector material, whereby the tracks are propagated in discrete time steps for each time interval spent in a different material. The influence of matter interactions is thereby proportional to the thickness and composition of the material layer encountered by the track. This procedure is described in Ref. [79], whereby the formalism introduced in Ref. [78] is used here instead. Accordingly, the kaon amplitudes from Equations (2.63) and (2.64) are first rewritten as

$$\alpha_L(t) = X_L(t)\alpha^0_L + x(t)\alpha^0_S \qquad \alpha_S(t) = X_S(t)\alpha^0_S + x(t)\alpha^0_L, \tag{6.5}$$

whereby the common time dependencies are identified as

$$X_L(t) = \cos(\Omega t) - \frac{i\Delta\lambda}{2\Omega}\sin(\Omega t), \qquad X_S(t) = \cos(\Omega t) + \frac{i\Delta\lambda}{2\Omega}\sin(\Omega t), \tag{6.6}$$

$$x(t) = -\frac{i\Delta\chi}{2\Omega}\sin(\Omega t), \tag{6.7}$$

Table 6.3 – Input constants for the kaon regeneration effect calculation.

Parameter	Value	Source
Δm_K	$(0.5293 \pm 0.0009) \times 10^{-2}\ \hbar/\mathrm{ps}$	[65]
τ_S	(89.54 ± 0.04) ps	[116]
τ_L	$(51\,160 \pm 210)$ ps	[116]
m_K	$0.497\ \mathrm{GeV}/c^2$	[116]
ϕ	(-124.7 ± 0.8) deg	[151]
$\lvert\epsilon_K\rvert$	$(2.228 \pm 0.011) \times 10^{-3}$	[56]
$\phi_{+-} = \arg \epsilon_K$	(43.51 ± 0.05) deg	[56]

and the definitions introduced in Section 2.5 are used. The kaon amplitude arriving at the j-th detector layer can then be computed in an iterative approach [79]

$$\alpha_\mathrm{L}(t_j) = X_\mathrm{L}(t_j - t_{j-1})\alpha_\mathrm{L}(t_{j-1}) + x(t_j - t_{j-1})\alpha_\mathrm{S}(t_{j-1}), \tag{6.8}$$

$$\alpha_\mathrm{S}(t_j) = X_\mathrm{S}(t_j - t_{j-1})\alpha_\mathrm{S}(t_{j-1}) + x(t_j - t_{j-1})\alpha_\mathrm{L}(t_{j-1}), \tag{6.9}$$

with $\alpha_{\mathrm{L,S}}(t_0 = 0) = \alpha^0_{\mathrm{L,S}}$. The parameters $\Delta\lambda$ are defined by the known K^0-$\overline{K}^0$ oscillation parameters, but the $\Delta\chi$ parameter, that is also needed to compute Ω needs to be derived through experimental methods. As motivated in Ref. [78], the attenuation difference $\Delta\chi$ is proportional to the difference of forward scattering amplitudes Δf, which are in turn related to the cross-section difference $\Delta\sigma$ of K^0 and $\overline{K}^0$ in matter through the optical theorem, and this difference is accessible to measurements [149–151], which leads to the relation

$$\Delta\sigma = \frac{4\pi}{p}\Im\Delta f, \quad \text{with} \quad \Delta\chi = -\frac{2\pi\mathcal{N}}{m_K}\Delta f. \tag{6.10}$$

Here, p is the total kaon momentum, m_K is the kaon mass, and $\mathcal{N}$ is a density factor of the medium. Furthermore, the phase, ϕ, of Δf can be measured [151], and with this information, $\Delta\chi$ can be estimated

$$\begin{aligned}\Delta\chi &= -\frac{2\pi\mathcal{N}}{m_K}\underbrace{e^{i\phi}\Im(\Delta f)\sqrt{1+\cot^2\phi}}_{\Delta f} \\ &= -\frac{\rho N_A c}{2M m_K} p\Delta\sigma e^{i\phi}\sqrt{1+\cot^2\phi},\end{aligned} \tag{6.11}$$

whereby ρ is the medium density, N_A is the Avogadro constant, c is the speed of light and M is the molar mass of the medium. The empiric relationship between the kaon cross-section difference, kaon momentum, and atomic number A is measured in Ref. [150] and is given by

$$\Delta\sigma = 23.2 \cdot A^{0.758} p^{-0.614}\ \mathrm{mb}, \tag{6.12}$$

Table 6.4 – Kaon regeneration and *CP*-violation corrections that are subtracted from the final result weighted by the flavor tagging dilution. A separate correction is defined for the combined fit, including J/ψ decays to the final states $\mu^+\mu^-$ or e^+e^-.

Data set	$\Delta S \times 10^3$	$\Delta C \times 10^3$
$B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$	-1.573 ± 0.006	3.417 ± 0.008
$B^0 \to J/\psi(\to e^+e^-)K^0_S$	-1.721 ± 0.017	3.734 ± 0.023
$B^0 \to J/\psi(\to \ell^+\ell^-)K^0_S$	-1.594 ± 0.006	3.462 ± 0.008
$B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$	-1.996 ± 0.028	4.042 ± 0.034
Combined fit	-1.619 ± 0.006	3.498 ± 0.008

whereby $[p] = 1\,\mathrm{GeV}/c$. With these relations and values from the literature listed in Table 6.3, the kaon regeneration and *CP*-violation corrections in Equations (2.68) and (2.69) can be evaluated for each kaon track.

The most significant detector interaction for kaon tracks is seen to occur in the form of kaon-aluminium interactions in the VELO RF-foil. A map of kaon intersection frequencies in the detector is shown in Figure F.3. Figure 6.4 shows a scatter plot of the distribution of corrections applied to each reconstructed kaon track in the data set, weighted by the flavor tagging dilution and signal weight of the candidate. Table 6.4 lists the mean corrections that are subtracted from the final results for each fit category. The larger correction applied to the $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$ decay can be attributed to the lower mean kaon momentum, and thus larger cross-section difference $\Delta\sigma$. The advantage of subtracting corrections from the final result, instead of accounting for the effect on a per-event basis, is that future improvements of measurements of kaon-matter interactions can be used to reevaluate such corrections without affecting the observed amplitudes of $\mathcal{A}^{CP}(t)$. A candidate-wise correction is tested in the ML fit and causes the same mean offsets of the result as when averaging the corrections. Therefore, a subtraction of corrections does not compromise the overall precision.

6.5 Definition of the *CP*-fit model and fits to simulation

The B decay rates are derived in Equations (2.27) and (2.28) and are extended in Section 5.1 to account for imperfect and possibly biased flavor tagging. In Sections 6.1 and 6.3 this model is extended further to account for the measured decay-time resolution and decay-time acceptance. With the model components derived so far, the *CP*-violation likelihood can be written as

$$\mathcal{L}_{CP}(S, C, \ldots) = \mathcal{N}(\xi) \prod_{j=1}^{N} \epsilon_{\mathrm{acc}}(t_j) \int_{-\infty}^{\infty} e^{-\Gamma t_j/\hbar} P_{CP}(t_j, \boldsymbol{\eta}_j, \boldsymbol{q}_j) \cdot R(t_j - t' | \sigma(\sigma_{t_j}, \boldsymbol{c})) dt' \Big|_{t'=t_j}, \tag{6.13}$$

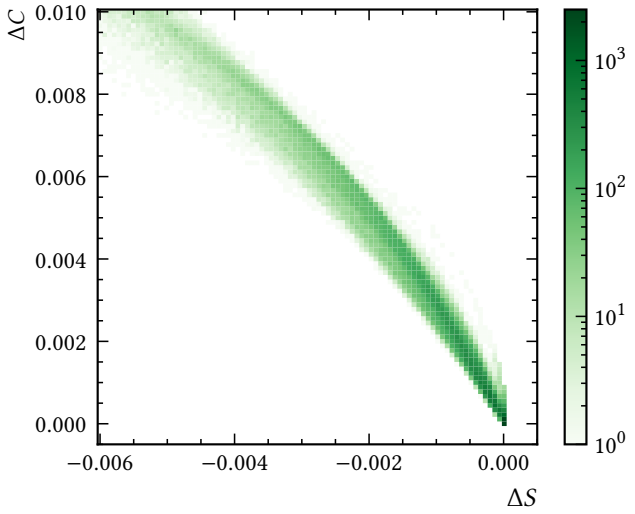

Figure 6.4 – Scatter plot of the corrections applied to the parameters S and C due to CP-violation in the kaon system and kaon regeneration for each reconstructed kaon track, weighted by the flavor tagging dilution.

and is referred to as CP-fit. At the core of this likelihood are the time-dependent decay rates $P_{CP}(t_j, \boldsymbol{\eta}_j, \boldsymbol{q}_j)$, which assign initial B^0 and $\bar{B}^0$ different time dependencies, which do not cancel because of CP violation. This component is formulated in Equation (5.18). The decay-time resolution function is convolved with the decay-time model to account for uncertainties in the decay-time reconstruction. Here, $\mathcal{N}(\boldsymbol{\xi})$ symbolizes a set of multidimensional Gaussian constraints for multiple, partially correlated input parameters $\boldsymbol{\xi}$ and thus propagates the uncertainties of external parameters to the final result. The input parameters $\boldsymbol{\eta} = (\eta^{\mathrm{SS}}, \eta^{\mathrm{OS}})$ and $\boldsymbol{q} = (q^{\mathrm{SS}}, q^{\mathrm{OS}})$ are the uncalibrated mistag combinations of calibrated single taggers and tagging decisions of each tagging side. The decay-time uncertainty associated with each candidate is denoted by σ_t. The time variable t refers to its reconstruction-bias-corrected decay time, as documented in Section 6.1. The parameters of the time-dependent efficiency function $\epsilon_{\mathrm{acc}}(t)$ are initialized with previous fit results of the decay-time distributions to achieve faster convergence in the CP-violation fit. To that end, pre-fits of the decay-time distribution are performed for each track-type reconstruction class and final state.

In total, seven spline parameters per decay mode and track type reconstruction, eight flavor tagging calibration parameters per decay mode, the B^0 oscillation frequency Δm_d, and two CP-violation parameters $S_{\psi K^0_{\mathrm{S}}}$ and $C_{\psi K^0_{\mathrm{S}}}$ are optimized in the fit. Flavor tagging efficiency asymmetries are determined in the calibration channels $B^0 \to J/\psi(\to \mu^+\mu^-)K^*$ and $B^0 \to J/\psi(\to e^+e^-)K^*$, and set constant in the fit to data, for the reason that a CP fit to $B^0 \to \psi K^0_{\mathrm{S}}$ cannot have sensitivity to α, Δp_i, and $\Delta\epsilon$ simultaneously in absence of a flavor-tagging-independent production flavor estimate. A scaling of the α uncertainty to incorporate $\Delta\epsilon$ uncertainties is tested and found to be negligible and is therefore not implemented. The fit approach is replicated on simulated data sets using inputs from simulated calibration modes.

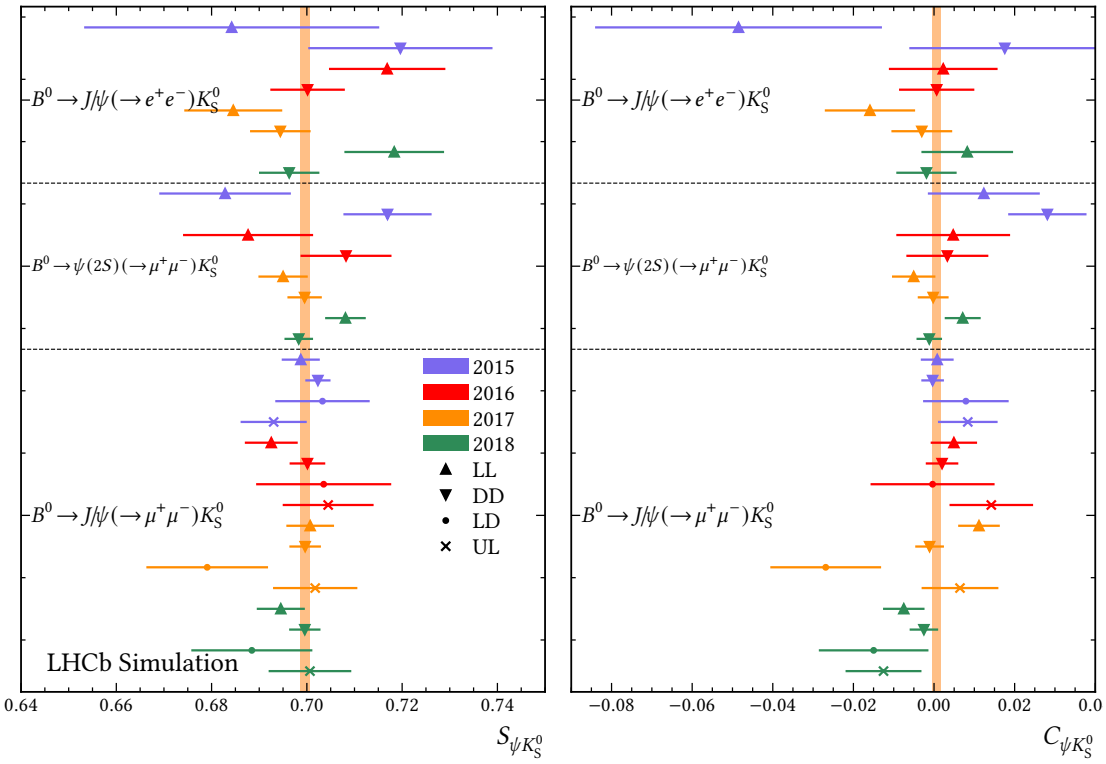

Figure 6.5 – Fit results of simulated data sets where the true production flavor is used for tagging instead of flavor tagging. The fit results are close to the generator values.

Several external input parameters ξ and their uncertainties are included in the fit. The B^0 oscillation frequency Δm_d and the B^0 decay time $\tau_{B^0} = \hbar/\Gamma_d$ are set to the current world averages [62]. A determination of the oscillation frequency with the available $B^0 \to J/\psi K^*$ data sets yields a result that is in excellent agreement with the world average and has a comparable statistical uncertainty. As this result alone would not significantly improve the measurement presented here, the world average value is chosen instead. At the same time, this finding implies that a precise determination of the B^0 oscillation frequency can be performed with the available LHCb Run 2 data in multiple decay modes, including $B^0 \to \psi K^0_S$ and $B^0 \to \psi K^*$, and such a measurement would be likely to improve the world average of Δm_d significantly. The flavor tagging calibration parameters and the production asymmetry are obtained from the time-dependent fit to the $B^0 \to J/\psi K^*$ calibration data set and are subject to a Gaussian constraint to their central values and uncertainties. One multidimensional Gaussian constraint per decay channel accounts for all correlations between the parameters of the SS and the OS calibrations and the production asymmetry. The input correlation

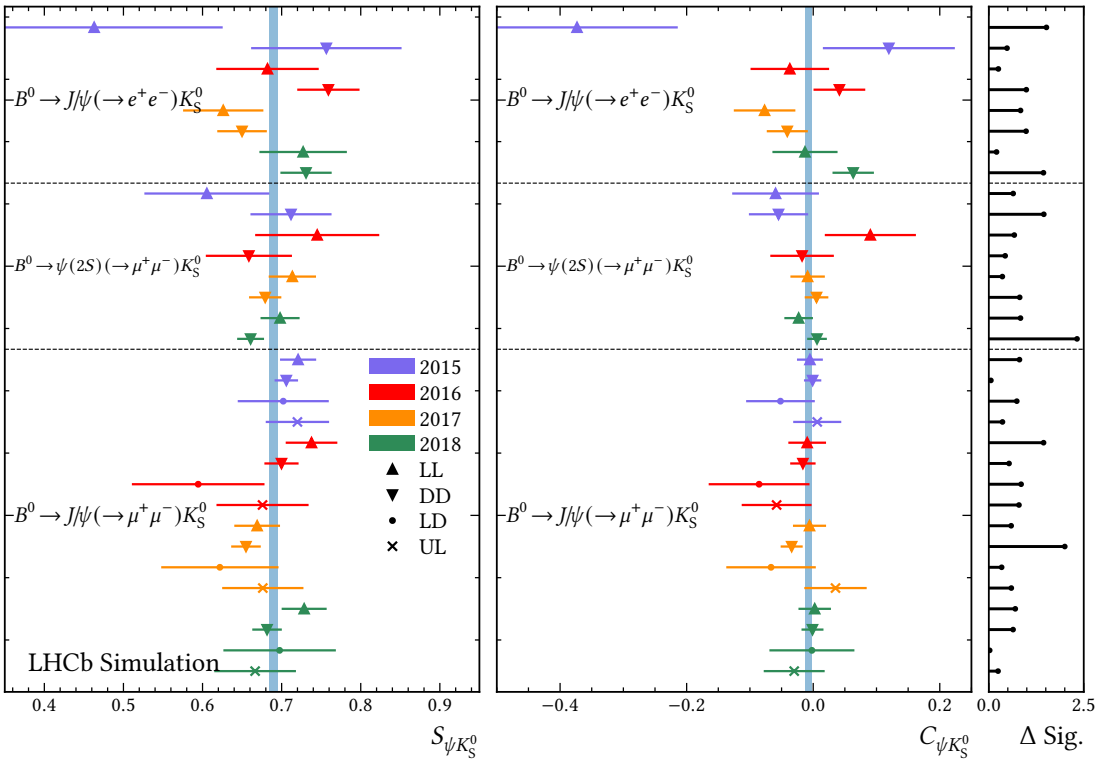

Figure 6.6 – Fit results of each simulated data set in which a simulated tagging response is used. The fit results are well compatible with the full-statistics fit, in which every event is perfectly tagged.

matrix that is applied as a constraint for the $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ mode is

	p_0^{SS}	p_1^{SS}	Δp_0^{SS}	Δp_1^{SS}	p_0^{OS}	p_1^{OS}	Δp_0^{OS}	Δp_1^{OS}	α
p_0^{SS}	1	0.036	−0.011	−0.004	0	0.001	0	0	0
p_1^{SS}		1	−0.005	0.005	0.001	0	0	0	0.001
Δp_0^{SS}			1	0.016	0	0	0.007	−0.005	−0.069
Δp_1^{SS}				1	0	0	−0.007	0.004	0.057
p_0^{OS}					1	0.115	−0.008	−0.007	0.001
p_1^{OS}						1	−0.007	−0.001	0.001
Δp_0^{OS}							1	0.092	−0.059
Δp_1^{OS}								1	0.036
α									1

and shows that FT-parameter correlations are generally small by definition and do not exceed 11.5 % for any two parameters. In the other two modes, similar correlations are observed, and these are listed in Equations (T.13) and (T.14). The fit model is implemented with the ROOT software framework [152, 153]. In Figure 6.7, the projections of the decay-time fits from which the final results have been extracted are shown.

The first direct validation step of the fit model is to apply it to simulated signal data sets and to use the true production flavor as B-tag instead of flavor tagging. This

ensures that the coefficients $S(q,\eta)$, $C(q,\eta)$ and $H(q,\eta)$ use the proper definitions and that the decay-time description is satisfactory on a large sample. As every single event is perfectly tagged in these fits, errors in the fit setup would be strongly amplified by the statistics of the available samples. The fit results should therefore result in values of S and C that are very close to the input parameters $S^{\mathrm{MC}} = 0.7$ and $C^{\mathrm{MC}} = 0$ and serve as good estimators of the CP asymmetry in the generated sample. Figure 6.5 lists the results of the true tag fits, and the results are in excellent agreement with the generator values. Vertical bands represent a simple weighted average of individual parameters. These fits are repeated on simulated data sets, where flavor tagging information is used instead of the true flavor, which reduces the statistics effectively to a fraction of ca. 4 %. For this setup, the nominal analysis strategy is followed as closely as possible, whereby the calibration is fully replicated on simulated calibration channels, including the time-dependent fit on the $B^0 \to J/\psi(\to \mu^+\mu^-)K^*$ and $B^0 \to J/\psi(\to e^+e^-)K^*$ samples. As the simulation does not include background components, the *sPlot* method is not needed to subtract background.

In Figure 6.6, the fit results of the fit setup are listed where simulated flavor tagging information is used, and it can be seen that the measurements are well compatible with the results obtained in the full-statistics fits. The given pulls measure the significance of the difference between the full-statistics fit and the flavor tagging fit and thus probe the flavor tagging response. The significance is computed from the squared Mahalanobis-distance [154] between measured (S, C) pairs and takes the sizeable correlations of both parameters into account. Furthermore, for simplicity, the total sample is assumed to be effectively uncorrelated to the smaller tagged subset, which is justified by the small effective tagging efficiency. The validity of results in data subsets showing deviations of approximately 2σ are verified with least-squares fits of the binned, time-integrated CP-asymmetry.

Finally, fits are performed in each data subset, *i.e.* for each year of data taking and for each track category. Due to the blinding method used during the development of the measurement, the agreement of fits to data was studied in great detail, and a selection of these studies is listed in Section 7.2. The fit results are seen to be well compatible with one another. A multitude of checks and alternative measurements did not indicate that the measurement strategy was biased. After the finalization of validation studies, the measurement was approved for unblinding, and the measurement strategy has not been altered since. The reported results are determined in fits, where the CP-violation amplitudes S and C are shared between the subsets of individual years and track categories.

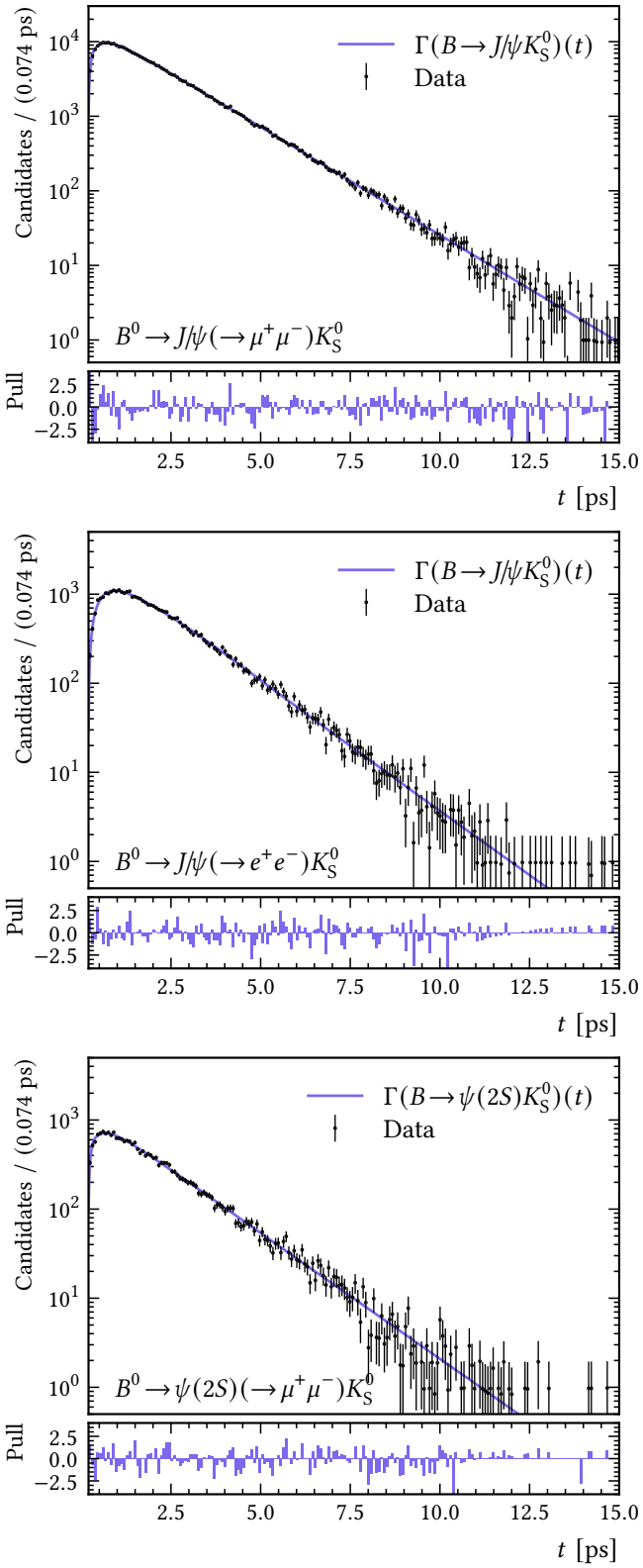

Figure 6.7 – Projection of data fits of the three analysis channels onto the time axis for the full Run 2 data set of LHCb.

7 Systematic uncertainties

7.1 Assigned systematic uncertainties

Systematic uncertainties are assigned to account for influences on the result, which are due to the choice of measurement strategy and assumptions made in this respect or due to statistical and systematic uncertainties of certain external input parameters. External input parameters that are part of the fit model, are included as Gaussian constraints to propagate their total uncertainty and thus contribute to the statistical uncertainty of the result. The primary method for estimating systematic uncertainties are *toy*-fits for which data sets are randomly sampled from the probability distributions of key observables in the data.

Table 7.1 – Summary of systematic uncertainties assigned to the result of the simultaneous fit to all three decay channels.

Source	$\sigma(S) \times 10^3$	$\sigma(C) \times 10^3$
Fit validation	0.388	0.597
$\Delta\Gamma_d$ uncertainty	5.492	1.682
Time bias correction	0.749	0.744
Time bias model uncertainty	0.659	1.294
FT calibration transfer	5.322	0.112
FT $\Delta\epsilon$ transfer	1.376	1.742
Total	7.844	2.908

During the development of the analysis, the assignment of a multitude of systematic uncertainties was considered, most of which were found to be insignificant. A selection of these studies is mentioned in Section 7.2. We assign systematic uncertainties for flavor tagging, specifically, the transfer of calibration parameters from the calibration modes to the signal modes and for the model correcting the VELO decay-time bias. The parameter $\Delta\Gamma_d$ is not part of the fit model, and the propagation of its uncertainty to the result is accomplished through toy studies. Finally, a fit validation systematic uncertainty is assigned, measuring the bias of the fit model itself when applying it to baseline toy data sets.

The generation of toy data is performed as follows: the reconstructed B mass, the raw mistag distributions of the SS and the OS combinations, and the B decay time uncertainty are sampled from the distributions found in the data. For the sampling

Table 7.2 – Listing of systematic uncertainties assigned to each measured decay channel for each contribution to the total uncertainty, scaled by a factor of 10^3.

	$B^0 \to J/\psi K^0_S$ $J/\psi \to \mu^+\mu^-$		$B^0 \to J/\psi K^0_S$ $J/\psi \to e^+e^-$		$B^0 \to \psi(2S) K^0_S$ $\psi(2S) \to \mu^+\mu^-$	
Source	$\frac{\sigma(S)}{10^{-3}}$	$\frac{\sigma(C)}{10^{-3}}$	$\frac{\sigma(S)}{10^{-3}}$	$\frac{\sigma(C)}{10^{-3}}$	$\frac{\sigma(S)}{10^{-3}}$	$\frac{\sigma(C)}{10^{-3}}$
Fit validation	0.37	0.01	0.3	0.2	0.5	2.0
$\Delta\Gamma_d$ uncertainty	5.3	1.7	7.7	4.0	5.9	1.4
Time bias correction	0.7	0.8	1.5	1.2	0.1	0.8
Time bias model uncertainty	0.4	0.3	1.3	6.2	2.5	3.6
FT calibration transfer	4.97	0.07	2.1	0.38	16.47	0.1
FT $\Delta\epsilon$ transfer	1.3	1.6	1.7	2.0	2.9	3.0
Total	7.4	2.5	8.4	7.7	17.9	5.3

of the reconstructed mass, the data are generated for each background and the signal component of the baseline mass model that was optimized on data. Generated mass distributions are fitted with the baseline mass model again, and sWeights are generated from the fit result. Finally, the fully assembled *CP*-pdf in Equation (6.13) is used to generate a decay-time distribution, for which the values of physics and model parameters can be chosen. The values of the *CP*-violation amplitudes, the oscillation frequency Δm_d, and the B decay time are set according to each contribution in the generated data. The decay time of the partially reconstructed background and signal is set to the known value for B^0 mesons, and for the B^0_s component, the lifetime is set accordingly [116]. The combinatorial background is assigned a lifetime of 0.8 ps, consistent with results from background-only fits. The distributions of the fit results for the parameters S and C are then subtracted from the generated values 0.7 and 0, respectively, to yield residual distributions. The mean of the residual distribution is interpreted as the systematic uncertainty associated with this variation.

Systematic uncertainties are determined for each of the three decay channels analyzed using independent toy studies. The systematic uncertainty assigned to the result of the simultaneous fit to all channels is a weighted average of those uncertainties, for which the inverse variances of S and C are used as averaging weights. The systematic uncertainty measuring the fit validation itself is assumed to be fully correlated between channels instead. For the assignment of a total systematic uncertainty per channel and the simultaneous fits, individual contributions are added in quadrature, and the obtained value is rounded to match the precision of the measurement. Even though the fit model of each channel is functionally the same, due to differences in the time-dependent selection efficiency, the effect of a variation of external parameters like $\Delta\Gamma_d$ can affect the three channels to different degrees. For this reason, this systematic uncertainty is assumed to be uncorrelated for simplicity, which is justified, also considering that the

systematic uncertainty assigned to the simultaneous fit is dominated by the precision of the main statistics mode $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$. A direct determination of systematic uncertainties of the merged data set is prohibitively time and resource-expensive, and no improvement over the chosen approach can be expected in this measurement, whose precision is limited by the statistical uncertainty.

Table 7.2 lists the systematic uncertainties assigned for each of the three decay modes. These contributions are summed in quadrature to the total assigned systematic uncertainty. The systematic uncertainties assigned to the simultaneous fit to all decay modes are given in Table 7.1. The assigned total systematic uncertainty is significantly smaller than the statistical uncertainty for S and C and can be expected to become somewhat smaller in view of planned future measurements. In particular, an improvement of the measurement of $\Delta\Gamma_d$ should have a comparatively high impact in this regard.

The bias of the fit model and measurement approach is measured by producing toy data sets of the baseline setup and subtracting the fit results from the true simulation inputs. This bias is subtracted from other toy-based estimates of systematic uncertainties as the fit procedure and model are identical in each mode. The fit validation bias is the smallest systematic uncertainty, and as it includes the procedure of determining sWeights through a mass fit, the influence of the chosen background subtraction method on the result is thereby shown to be minimal. In particular, since a $B^0_s \to \psi K^0_S$ component with other associated CP-violation parameters is generated, it is shown that its influence on the result is negligible and that the statistical subtraction of this background component is effective. Because of the fast oscillation of B^0_s mesons, the associated time-integrated CP-asymmetries are small. Furthermore, no SS-kaon tagger, the most significant source of tagging information in B^0_s decays, is utilized for in the signal modes, and therefore, the baseline fit model cannot be expected to be sensitive to CP violation in the $B^0_s \to \psi K^0_S$ decay in the first place.

The largest source of systematic uncertainty is due to the current uncertainty of $\Delta\Gamma_d$, which is estimated with toy fits similarly as in the predecessor measurements in Refs. [13, 14]: the value of $\Delta\Gamma_d$ is varied by the sum of its current standard deviation and absolute central value and the variation observed for the parameters S and C is the assigned systematic uncertainty. Given that the sign of $A^{\Delta\Gamma}$ is unknown, and this measurement is not sensitive to its value, two relevant scenarios are considered: $A^{\Delta\Gamma}$ and $\Delta\Gamma_d$ can have the same or opposite signs. The value of $A^{\Delta\Gamma}$ is set to $\pm\sqrt{1-S^2}$, according to the normalization condition in Equation (2.41). Only the larger of the two resulting systematic uncertainties is assigned for each observable. Systematic uncertainties of ± 0.0055 for S and ± 0.0017 are assigned for C, equivalent to fractions of 41.21 % and 13.71 % of the statistical uncertainty of the simultaneous fit. For the other two modes, variations of similar size are observed.

The parameter uncertainties of the flavor tagging calibration are sufficiently small, such that their Gaussian constraint has little influence on the measurement precision. Yet, transferring calibrations from the calibration modes to the signal modes could potentially have more considerable associated systematic uncertainties, which need to be estimated. Due to the lack of an appropriate self-tagging control channel that is more similar to the signal than $B^0 \to \psi K^*$, the choice is made to use a simulation-only

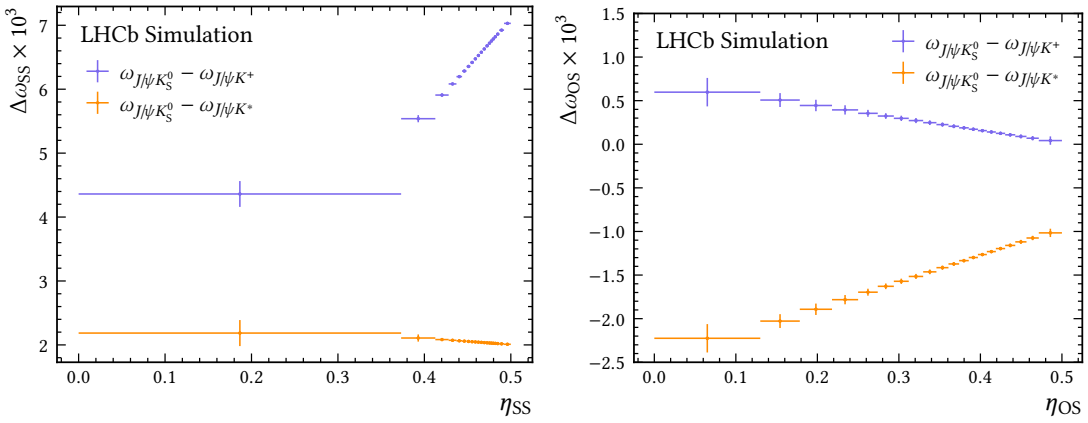

Figure 7.1 – Difference calibrations of the SS (left) and OS taggers (right) determined on simulation. The difference in calibrations is shown for true tag calibrations determined on signal and in the calibration modes.

approach, which is justified due to the comparably small systematic uncertainties in this measurement. The tagging algorithms in the simulated signal and kinematically weighted calibration modes are individually calibrated using the true production flavor in the flavor tagging likelihood given in Equation (5.19). The calibrations from the control mode are then applied to the signal mode in addition and compared to the calibration obtained directly on the signal. Ideally, the calibration parameters should be identical if all properties of tracks that are not part of the signal decay are identical between modes. Since small deviations are observed, a mistag-difference distribution is determined by subtracting the two calibrated mistag histograms per tagging side. In Figure 7.1, the obtained difference distributions are shown. These differences are used in the toy-generation step to generate mistag distributions, from which this difference is subtracted, and the effect on the measurement is determined. It is seen that the parameter C is almost unaffected by calibration-transfer systematic uncertainty in the main statistics mode. In contrast, a sizeable, relative effect on S of 33.2 % of its statistical uncertainty is observed.

While the value of tagging efficiency asymmetries $\Delta\epsilon$ are measured to be small, as shown in Table 5.3, and their effect on the result must therefore be small (see Equations (5.10) to (5.12)) compared to other influences, the question remains how well measurements of $\Delta\epsilon$ can be transferred from fits of the calibration data set $B^0 \to J/\psi K^*$ to the signal modes. In the predecessor measurements, $\Delta\epsilon \equiv 0$ has been set in all fits, and no systematic uncertainty has been assigned. A purely data-driven transferability estimate is not possible due to a lack of a suitable self-tagging control mode and the inherently indeterminate production flavor information in $B^0 \to J/\psi K^0_S$ signal, for which flavor tagging is needed in the first place. In this measurement, it was decided to therefore assign a systematic uncertainty based only on simulation. The chosen approach is to determine $\Delta\epsilon$ in the simulated control and signal mode by comparing

counts of B^0 and $\bar{B}^0$ tagging decisions:

$$\langle \Delta\epsilon \rangle = \frac{N_{\text{tag}}(\bar{B}^0) - N_{\text{tag}}(B^0)}{N_{\text{tag}}(\bar{B}^0) + N_{\text{tag}}(B^0)}, \quad \text{with} \quad N_{\text{tag}}(f^{\text{true}}) = \sum_{\substack{j=1 \\ f_j^{\text{true}}=\pm 1}}^{N} \lambda_j |q_j|, \tag{7.1}$$

where the true production flavor f_j^{true} is 1 for true B^0 and -1 for true $\bar{B}^0$ and q_j are the tagging decisions. In a sample without CP violation, the estimate of $\langle \Delta\epsilon \rangle$ is reliable, but in the signal mode, events need to be weighted by the inverse oscillation to effectively remove the CP-violation influence on the estimate:

$$\lambda_j = \frac{1}{1 - \text{sgn}(f_j^{\text{true}}) S^{\text{true}} \sin(\Delta m_d t_j)}, \tag{7.2}$$

whilst in the calibration sample $\lambda_j \equiv 1$ is chosen. To determine systematic uncertainties, the differences between the $\Delta\epsilon$ values obtained on signal and $B^0 \to J/\psi K^*$ simulation for each tagging side are chosen as variations for the $\Delta\epsilon^{\text{SS}}$ and $\Delta\epsilon^{\text{OS}}$ inputs of the fit. The obtained differences in the fits are comparable in size to the central values of $\Delta\epsilon$ obtained in the time-dependent fit on the control mode. For those values for which the uncertainty obtained by event counting is larger than the central value, the respective parameters are varied by one unit of its uncertainty instead. For the simultaneous fit, systematic uncertainties of ± 0.0013 for S and ± 0.0016 for C are assigned.

Systematic uncertainties associated with the decay time bias correction model are smaller in comparison and are determined for two setups: a variation of the fit-validation systematic uncertainty is determined by creating a decay-time bias in toy data sets according to the nominal bias-correction model, and this bias is corrected in fits to toy data using the baseline approach ("Time bias correction"). In addition, the parameters of the decay time bias model are varied within their uncertainties and the effect on the result is observed and added in quadrature to other systematic uncertainties. The assigned systematic uncertainties are of different sizes for S and C in individual channels, and for the total model, uncertainties of ± 0.0010 and ± 0.0015 are assigned for S and C.

The usage of a more detailed decay-time-efficiency model consisting of 20 cubic splines has a negligible effect on the result. Likewise, the effect of alternative flavor tagging calibrations, which are determined per track category, and the usage of alternative decay-time resolution models have negligible effects. The contribution of the world-average systematic uncertainty of Δm_d is included in the Gaussian constraint of this parameter. Expressed as systematic uncertainties for the target observables, the Δm_d uncertainty contributes by $\sigma_{\Delta m_d}(S) = 0.0004$ and $\sigma_{\Delta m_d}(C) = 0.0023$ in the simultaneous fit, which is included in the reported statistical uncertainties. The influence of the Gaussian constraint accounting for the uncertainty of the B lifetime is negligible.

In summary, systematic uncertainties of ± 0.008 and ± 0.003 are assigned for the parameters S and C for the simultaneous fit result, and further contributions to the total systematic uncertainty are assumed to be negligible. The systematic uncertainties

determined in each individual channel are listed in Table 7.2. Between individual channels, assigned systematic uncertainties are of comparable size for each source of systematic uncertainty, except for the flavor tagging-related uncertainties assigned to the $B^0 \to \psi(2S)K^0_S$ mode and its fit bias, which is elevated in comparison. These findings could be related to the fact that no specialized calibration channels $B^0 \to \psi(2S)K^*$ and $B^+ \to \psi(2S)K^+$ have been used for the calibration of this signal mode. As the measurement of CP-violation in the $B^0 \to \psi(2S)K^0_S$ decay is limited by the statistical uncertainty, this fact does not negatively influence the result obtained in this mode.

7.2 Consistency studies

During the development of the measurement, the result was blinded through the addition of two different random and unknown offsets to the final values of S and C in fits to data. This method allowed us to test the consistency of the data fits and measure the effect of various assumptions on the central values and the measurement sensitivity. This offset was removed only after the measurement passed an extensive list of consistency checks and after the measurement was approved for publication by the LHCb collaboration.

Numerous studies have been performed to ensure that results are consistent between years of data taking and in data subsets, like the kaon reconstruction category or between bins of kinematic observables before the fit results were unblinded on data. Simulated data sets that are generated with known values of S and C served as validation modes in many instances. Three measurement channels, four years of data taking, and up to four reconstruction categories result in 32 data subsets which are executed as parallel and almost independent sub-analyses. The automation of the main analysis workflow and the individual sub-analyses was managed using the workflow automation library “Snakemake” [155]. In conclusion, the results obtained in the sub-analyses are seen to be consistent. From the partial results obtained in these measurements, it was possible to identify and handle potential deviations early. The tests mentioned in the following were completed before the unblinding, which further increased our confidence in the chosen fit approach. The observed variations of the results caused by these tests were found to be negligible and did not result in a change of measurement strategy or an assignment of a systematic uncertainty.

Because of the fully charge-symmetric final state, experimental influences on the central values must be small. Generally, to influence the result of this measurement, the effect in question needs to bias the selection or tagging performance of one of the B production flavors and not be already accounted for by α, $\Delta\omega$, and $\Delta\epsilon$ parameters. A general time-dependence of a relevant observable alone, like the BDT output, is thus insufficient to cause a bias as it would affect the selection efficiencies of B^0 and $\overline{B}^0$ in the same way. These influences are accounted for by the decay-time acceptance model, which was verified to contain a sufficient number of degrees of freedom to describe the data accurately. The result is found to be consistent in bins of the BDT output and

transverse momentum, which demonstrates that the candidate selection is sufficiently production-flavor unbiased.

Generally, it is studied whether central values move correlatedly if different assumptions are made in the *CP* fit and whether such movements should be expected given the alternative assumptions. A wide range of tests is chosen for this precision measurement to understand whether deviations in subsets of the data should be attributed to a more systematic error instead of a statistical fluctuation and subsequently corrected. As a first cross-check, *CP* fits are performed in bins of the kinematic variables p, p_{T}, the pseudorapidity, and the BDT output, and it is tested whether a shift can be identified beyond a statistical effect, which is not the case. As the BDT applies time-biasing cuts, their effect on the selection of B^0 and $\bar{B}^0$ candidates is thus shown to be balanced.

To rule out the influence of the kaon decay vertex on the B decay time and thus possible reconstruction biases for long-lived K^0_{S} candidates, the B decay time is computed only from the dimuon vertex. No effect on the result is observed when using the modified decay time. Furthermore, *CP* fits are performed in bins of the kaon lifetime, flight distance, and pseudorapidity.

To test the effect of each set of complications implemented in the *CP* fit, results are compared between the simplest model and with each additional complication. These simplified assumptions include fixing the production asymmetry to zero, setting the B lifetime and oscillation frequency constant, not applying the decay time bias corrections, and finally fixing the flavor tagging constraints. While the result is not unaffected, the observed movements are expected and do not hint at problems. It is seen that Δp parameters, which measure the asymmetry of B^0 and $\bar{B}^0$ calibrations, are highly relevant if calibrating the taggers with the previous definitions, as when setting them to zero, significant biases can be identified on MC.

The B-transverse momentum differs between the four track reconstruction categories, with category DD having a significantly higher transverse momentum. As flavor tagging performance is known to depend on B kinematics, independent flavor tagging calibrations are determined for each reconstruction class to test whether this approach affects the result. The difference to the baseline result is measured and found to be negligible, which justifies the simplified approach of determining one calibration per channel.

A binned, time-dependent fit of the *CP* asymmetry of the upper mass-sideband, defined by $m(J/\psi K^0_{\mathrm{S}}) > 5400\,\mathrm{MeV}/c^2$, is performed and shows no signs of a *CP* asymmetry or other possible issues in the combinatorial background. The result of this fit is shown in Figure 7.2.

As the tagging efficiency is seen to be time-dependent for taggers of the SS and drop with increasing B decay time, toy studies are performed in which the generated time-dependent efficiency distributions are modified by selective removal of candidates to match this dependence[1], and the result is seen to be unaffected. Similarly, a time dependence of the mistag distribution cannot be ruled out on the basis of the available statistics, and analogous toy studies show no effect on the result. As the assumption is

[1] see Ref. [117] for more details.

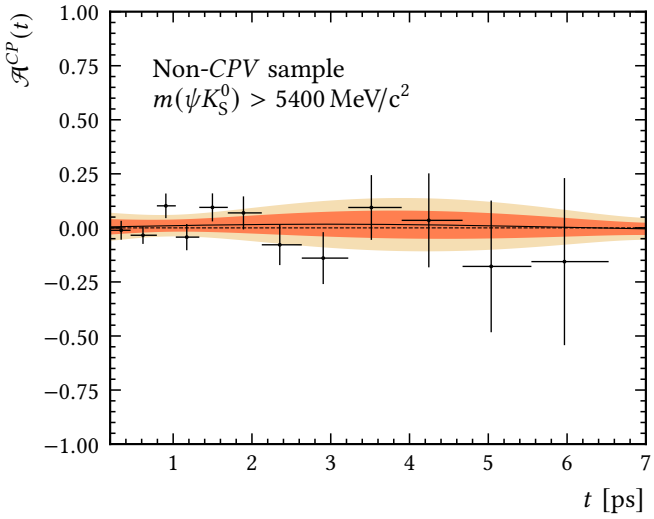

Figure 7.2 – Time-dependent asymmetry of the upper mass-sideband (background-only). No production-flavor asymmetry can be identified.

made that the calibrated SS and OS tagging responses are uncorrelated, two tests are performed: the correlation coefficient of mistag distributions is computed, and secondly, the calibrated SS and OS tagging responses are calibrated once again. The correlation coefficients are at the 1 % level for all final states, and the additional calibrations are compatible with the identity calibration, which demonstrates that the SS and OS are sufficiently uncorrelated and that no further calibration is required.

The correlation between reconstructed B mass and decay time is investigated using simulation, where a small but significant dependence is identified[1]. This dependence is precisely measured by determining the decay-time resolution in bins of the reconstructed B mass, and it is used to modify the generated toy data by selective candidate removal according to this dependence.

In the baseline model, as shown in Equation (5.18), the distributions describing initial B^0 and $\bar{B}^0$ mesons are optimized simultaneously, and their sum is thus globally normalized, which introduces a constraint on the yields of these two components. Using an alternative fit model [156] which does not contain a normalization constraint, does not require a functional description of the time-dependent efficiency, and which fits the datasets of the $B^0 \to J/\psi K^*$ calibration channel and the signal modes simultaneously, compatible results are found. Using this setup, two additional fits are performed: a fit of a model including an additional time-offset parameter, which is optimized in the fit, and a second fit in which the oscillation frequency is unconstrained and optimized by the fit. The results are in good agreement with the baseline result and have the same statistical uncertainty.

In addition, a potential VELO misalignment bias in the direction transversal to the beam direction is investigated[1] by generating ten simulated data sets of full-LHCb

[1] see Ref. [117] for more details.

Monte Carlo data in which one VELO half is offset by up to 10 µm in steps of 1 µm. The effect of a transversal VELO misalignment of this size is found to be negligible here, and no systematic uncertainties have been assigned. Finally, *CP* fits are performed in different B mass ranges and show perfect agreement with the nominal results, further ruling out *CP*-violating effects in the upper mass-sideband.

8 Results and combination of measurements

8.1 Fit results

The final results of this measurement are determined by running time-dependent, maximum-likelihood fits of the physics model derived previously, which can be compactly expressed as given in our Ref. [15]

$$\begin{aligned}\mathcal{P}(t,q,\eta) \propto e^{-\Gamma_d t/\hbar}\Big\{&\big[1+q(1-2\omega^{+}(\eta))\big]P_{B^0}(t)\\ &+\big[1+q(1-2\omega^{-}(\eta))\big]P_{\bar{B}^0}(t)\Big\},\end{aligned} \tag{8.1}$$

with

$$P_{B^0,\bar{B}^0}(t) \propto (1\mp\alpha)(1\mp\Delta\epsilon)[1\mp S\sin(\Delta m_d t)\pm C\cos(\Delta m_d t)]. \tag{8.2}$$

Here, q and η are the tagging decisions and mistag estimates, $\omega^{+}(\eta)$ and $\omega^{-}(\eta)$ are the flavor tagging mistag calibrations for initial B^0 and $\overline{B}^0$ mesons, α is the production asymmetry and $\Delta\epsilon$ is the flavor tagging efficiency asymmetry. This model is given in full detail in Equations (5.18) and (6.13) and Appendix E.2 to include the necessary descriptions of experimental effects including flavor tagging.

The fit is performed in each of the three decay samples with the full Run 2 LHCb data set and a simultaneous fit is performed with all three decay modes in which the physical parameters S and C are shared between data sets. The results have been verified to be compatible between years of data taking and track reconstruction categories within the methods' framework. Furthermore, all results obtained in this measurement are well-compatible. The CP-violation mechanism in all three analyzed final states is identical, and within the current experimental precision, it is therefore well justified to measure CP violation in a merged data set of all three modes. More specifically, higher-order corrections to the production of J/ψ versus $\psi(2S)$ in the final state could potentially modify $\Delta\phi_d$ to different degrees, but so far, such differences have not been observed and are likely to be small [157]. The influence of kaon regeneration and CP-violation in the neutral-kaon system on the neutral B system is estimated in Section 6.4 and subtracted from the result. The parameters S and C are expected to be significantly correlated since decay times are only measured in forward time at LHCb. For this reason, the correlation coefficient $\rho(S,C)$ is not averaged to zero in the total decay time range, as is the case in B-factory-measurements [8, 9]. Five main results are reported: fit results of the three individual channels, a simultaneous fit of the $J/\psi(\to\ell^+\ell^-)K^0_S$-modes, and a simultaneous fit of all three decay modes. In addition, a combination with LHCb-Run 1

results is performed in Section 8.2, and the result is interpreted in the current theoretical and empirical context.

The *CP*-violation parameters in the individual decay modes are measured to be

$$\begin{aligned}
S_{J/\psi(\to\mu^+\mu^-)K^0_S} &= 0.716 \pm 0.015\,(\text{stat}) \pm 0.007\,(\text{syst}),\\
C_{J/\psi(\to\mu^+\mu^-)K^0_S} &= 0.010 \pm 0.014\,(\text{stat}) \pm 0.003\,(\text{syst}),\\
S_{J/\psi(\to e^+e^-)K^0_S} &= 0.754 \pm 0.037\,(\text{stat}) \pm 0.008\,(\text{syst}),\\
C_{J/\psi(\to e^+e^-)K^0_S} &= 0.042 \pm 0.034\,(\text{stat}) \pm 0.008\,(\text{syst}),\\
S_{\psi(2S)(\to\mu^+\mu^-)K^0_S} &= 0.649 \pm 0.053\,(\text{stat}) \pm 0.018\,(\text{syst}),\\
C_{\psi(2S)(\to\mu^+\mu^-)K^0_S} &= -0.087 \pm 0.048\,(\text{stat}) \pm 0.005\,(\text{syst}),
\end{aligned}$$

with statistical correlation coefficients between S and C of 0.446, 0.374, and 0.503 for the final states $J/\psi(\to\mu^+\mu^-)K^0_S$, $J/\psi(\to e^+e^-)K^0_S$, and $\psi(2S)(\to\mu^+\mu^-)K^0_S$, respectively. These results are shown in Figure 8.1 in the S-C plane and are in good agreement. Differences between statistical correlation coefficients are primarily determined by different time-dependent selection efficiencies, and more specifically, stronger suppression of shorter-lived B mesons numerically leads to lower correlation coefficients. A simultaneous fit of all three modes yields

$$\begin{aligned}
S_{\psi K^0_S} &= 0.717 \pm 0.013\,(\text{stat}) \pm 0.008\,(\text{syst}),\\
C_{\psi K^0_S} &= 0.008 \pm 0.012\,(\text{stat}) \pm 0.003\,(\text{syst}),
\end{aligned}$$

with a correlation coefficient of 0.441 for the full Run 2 data set of LHCb and smaller uncertainties than the most recent 2021 world averages [62]. Additionally, a fit is performed that only includes the final states $J/\psi(\to\ell^+\ell^-)K^0_S$ and results in

$$\begin{aligned}
S_{J/\psi K^0_S} &= 0.722 \pm 0.014\,(\text{stat}) \pm 0.007\,(\text{syst}),\\
C_{J/\psi K^0_S} &= 0.015 \pm 0.013\,(\text{stat}) \pm 0.003\,(\text{syst}),
\end{aligned}$$

with a correlation coefficient of 0.437.

Figures 8.2 and 8.3 show the time-dependent yield asymmetry of B^0 and $\bar{B}^0$ decays to the *CP*-invariant final states of the simultaneous fit and each individual decay mode. The data points show the ML estimator of the time-integrated *CP*-asymmetry [156],

$$\mathcal{A}_{\text{int}} = -\frac{\sum_j q_j \kappa_j D_j}{\sum_j \kappa_j D_j^2}, \tag{8.3}$$

in each time bin, where κ_j is the sWeight, q_j is the tagging decision and D_j is the flavor tagging dilution of each candidate. The curve is a fit projection of the *CP*-asymmetry from the total model. The data points and the curve are corrected for the influence of the flavor tagging dilution, including tagging asymmetries and decay time dilution, and show the underlying unbiased *CP*-asymmetry of the data sample.

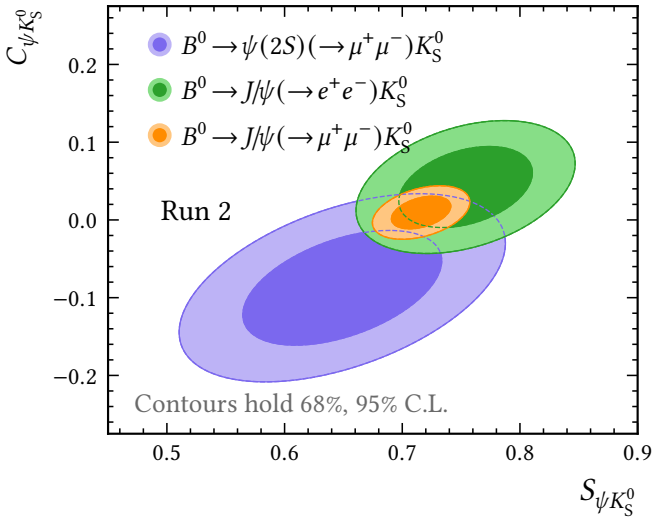

Figure 8.1 – Measurement results of the three analysis channels in the *S*-*C* plane.

8.2 Combination of measurements

The *CP*-violation measurement described here is combined with two previous LHCb measurements presented in Refs. [13, 14] where the *CP*-violation parameters S and C are measured with comparable methods with LHCb Run 1 data. In Ref. [13] a measurement of the decay channel $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S(\to \pi^+\pi^-)$ is performed, and Ref. [14] presents a measurement of *CP*-violation in the decay channels $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S(\to \pi^+\pi^-)$ and $B^0 \to J/\psi(\to e^+e^-)K^0_S(\to \pi^+\pi^-)$. Measurement [13] is performed on a data set that includes the data set used in Ref. [12] and therefore fully supersedes it. When combining parameters determined in multiple, uncorrelated measurements with uncertainties reported as standard deviations, a combination of measurements merely entails the computation of an uncertainty-weighted average $\hat{\mu}$ of central values and a combined uncertainty $\hat{\sigma}$. These are given by the ML estimators of the product likelihood of one-dimensional Gaussian likelihoods:

$$\hat{\mu} = \frac{\sum_j \mu_j \sigma_j^{-2}}{\sum_j \sigma_j^{-2}} \qquad \hat{\sigma} = \sqrt{\frac{1}{\sum_j \sigma_j^{-2}}}. \tag{8.4}$$

In the measurement described here and in Refs. [13, 14], the parameters S and C are measured simultaneously as one correlated pair $\boldsymbol{\mu}_i = (S_i, C_i)^\top$ per measurement i, with associated covariance matrix $\boldsymbol{V}_i$. Correlations of parameters, which are quantified by $\boldsymbol{V}_i$, need to be considered when averaging results. Individual measurements $\boldsymbol{\mu}_i$ are assumed to be statistically uncorrelated as the underlying data sets have no overlap. The results of the four measurements, which are combined, are shown in the left plot in Figure 8.4 in the *S*-*C*-plane. In addition, sources of systematic uncertainties that are caused by the uncertainties of external parameters, like Δm_d, affect all measurements in the same way and, therefore, need to be treated as being fully correlated. To account

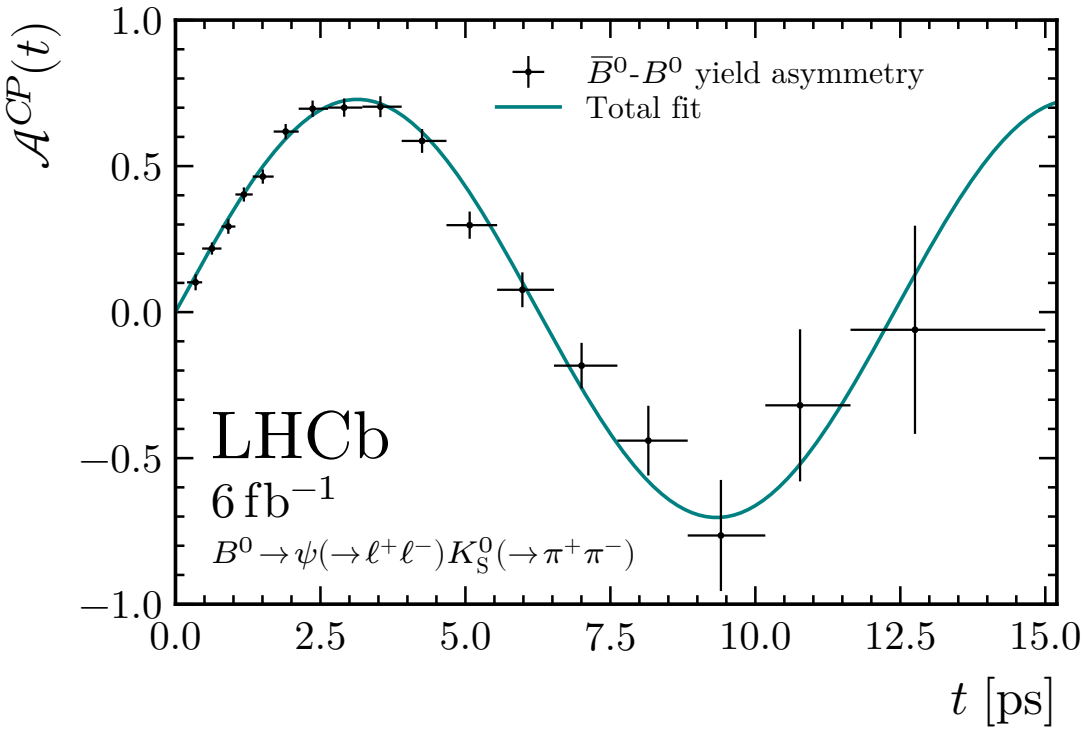

Figure 8.2 – Time dependence of the asymmetry of $\bar{B}^0$ and B^0 decays for the simultaneous fit [15]. The data points show the time-integrated *CP* asymmetry estimator per time bin, and the fit curve is a projection of the total decay rate model.

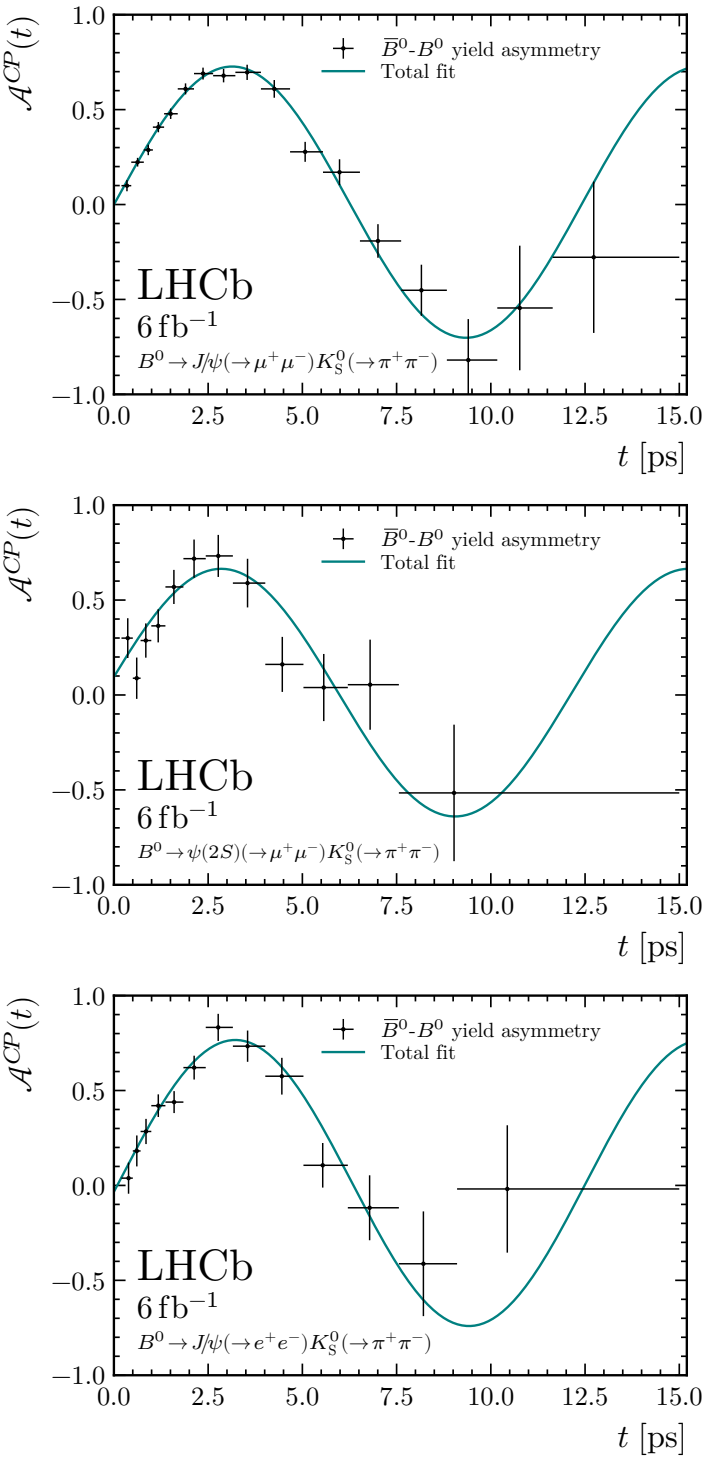

Figure 8.3 – Time dependence of the asymmetry of $\bar{B}^0$ and B^0 decays for the three individual decay modes, as annotated in the figures [15]. The data points show the time-integrated *CP* asymmetry estimator per time bin, and the fit curve is a projection of the total decay rate model.

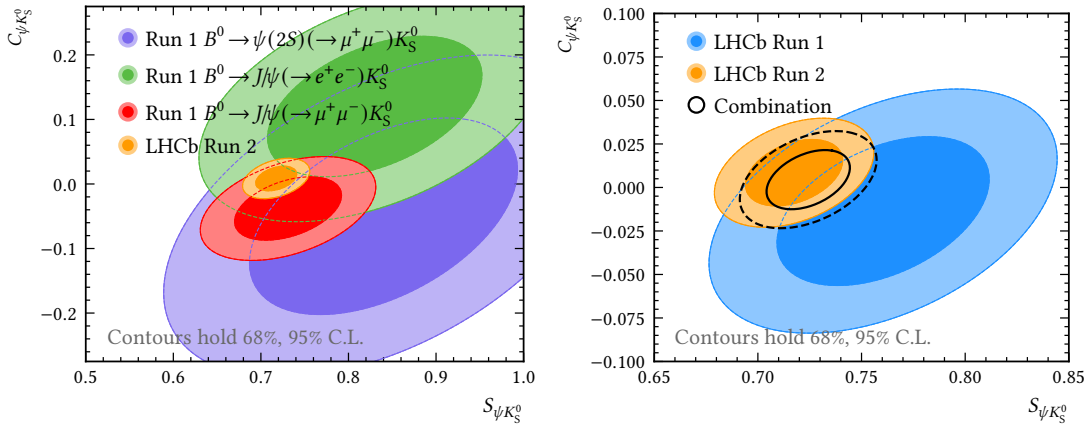

Figure 8.4 – Run 1 and Run 2 measurements which are combined to yield new LHCb combination (left) and comparison of LHCb Run 1 and Run 2 results and combined result in the S-C-plane (right).

for individual correlations, measurements can generally be combined by determining the likelihood product of each individual measurement and by reevaluating the ML estimator of this product.

The uncertainties of the reported measurements to be combined are given as symmetric standard deviations, as these amplitude measurements are not influenced by the nonlinearity of the sine function, given by $S \approx \sin(2\beta)$. The likelihood approximating the combination of individual measurements is therefore characterized by a product of two-dimensional Gaussian distributions

$$\mathcal{L}(\boldsymbol{x}) = \prod_i \frac{1}{2\pi\sqrt{|\boldsymbol{V}_i|}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_i)^\top \boldsymbol{V}_i^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_i)\right), \tag{8.5}$$

which is maximized by the ML estimator $\hat{\boldsymbol{x}}$ representing the combined central value of the measurements. In the case of Gaussian likelihoods, the extremal point and covariance matrix can be derived analytically, as the product of Gaussian distributions is essentially Gaussian with an additional constant factor, but this calculation is omitted here. The ML estimate is computed with the software tool GammaCombo [158], which is intended to compute the ML estimator of arbitrary product likelihoods and calculates the resulting parameter uncertainties with numerical methods involving a multidimensional scan of the product likelihood.

The treatment of systematic uncertainties differs between measurements in Refs. [13] and [14], and in the measurement described here, alternative choices are made yet again. More specifically, the treatment of statistical and systematic uncertainties of external parameters and the contribution of these uncertainties to either the statistical or systematic uncertainty of the result differ. To properly combine measurements, each uncertainty contributing to the total reported systematic or statistical uncertainty therefore needs to be isolated before a combination strategy can be pursued. Three sources of systematic uncertainty are identified, which are assumed to be fully correlated between measurements, which are the external parameter uncertainties of

Δm_d, $\Delta\Gamma_d$, and α. The treatment of these parameter uncertainties is described in the following. The B lifetime parameter constraint has negligible influence on the result and is therefore treated as if it was uncorrelated between measurements.

The parameter $\Delta\Gamma_d$ is treated the same in all measurements: its central value is fixed to zero in the baseline fit, and a systematic uncertainty is estimated from variations of the parameter within the sum of its uncertainty and absolute central value of $0.007\,\mathrm{ps}^{-1}$ [62].

In the measurement in Ref. [13], the B-oscillation frequency is constrained to the value $\Delta m_d^{2015} = 0.510 \pm 0.003(\mathrm{stat}) \pm 0.002(\mathrm{syst})\,\mathrm{ps}^{-1}$, whereby the parameter is Gaussian constrained to its central value and statistical uncertainty in the fit. A systematic uncertainty is assigned by varying the parameter by its systematic uncertainty alone. In the measurement in Ref. [14] the same technique is used with the value $\Delta m_d^{2017} = 0.5064 \pm 0.0016(\mathrm{stat}) \pm 0.0011(\mathrm{syst})\,\mathrm{ps}^{-1}$ instead. In the measurement presented in this thesis, the combined statistical and systematic uncertainty is propagated with one Gaussian constraint using the value $\Delta m_d = 0.5065 \pm 0.0019(\mathrm{stat+syst})\,\mathrm{ps}^{-1}$ [62] and no distinct systematic uncertainty is assigned.

Between 2011 and 2012, the center-of-mass energy of the LHC increased from 7 TeV to 8 TeV and therefore two different but correlated production asymmetry values are used in Ref. [13]. As the production asymmetry influence on the result is very small, these parameters are averaged here to $\alpha = -0.0108 \pm 0.0052(\mathrm{stat}) \pm 0.0014(\mathrm{syst})$. In the measurement described in Ref. [14] four different production asymmetry parameters values are used, which we average to $\alpha = X \pm 0.0065(\mathrm{stat}) \pm 0.00045(\mathrm{syst})$ and its central value can be ignored in this context. This parameter is constrained to its central value and statistical sensitivity in the fit, and a systematic uncertainty is assigned from variations by its systematic uncertainty. As described in Section 5.6, we measure α ourselves with a $B^0 \to J/\psi K^*$ data set and apply its statistical uncertainty as a Gaussian constraint in the signal fit.

The results of each measurement are reported in the form

$$P = \hat{P} \pm \sigma_P \pm \delta_P, \tag{8.6}$$

where statistical and systematic uncertainties are denoted as σ and δ in the following, and external input parameters are denoted in analogy in the form $E^{(i)} = \hat{E}^{(i)} \pm \sigma_E^{(i)} \pm \delta_E^{(i)}$. To make all measurements comparable, the results first need to be rearranged into the form

$$P = \hat{P} \pm \sigma'_P \pm \delta'_P \pm \delta_P^{(\Delta m_d)'} \pm \delta_P^{(\Delta\Gamma_d)'} \pm \delta_P^{(\alpha)'} \tag{8.7}$$

instead, whereby the contributions of relevant systematic components are subtracted from the reported quantities σ_P and δ_P, leaving them as σ'_P and δ'_P to measure the uncorrelated statistical and systematic uncertainty contribution, which is unrelated to these external parameters. The primed values are not reported in the previous analyses but can be estimated based on the quantities $\delta_P^{(j)}$ and $\delta_E^{(j)}$ from reported toy studies in the respective internal documents [159, 160]. Using the fact that the systematic uncertainties approximate the derivative of P with respect to the external parameter,

the total systematic contribution can be scaled up, if needed, to account for the total parameter uncertainty of E:

$$\delta_P^{(j)\prime} = \underbrace{\frac{\delta_P^{(j)}}{\delta_E^{(j)}}}_{\approx \partial_E P} \sqrt{\left(\sigma_E^{(j)}\right)^2 + \left(\delta_E^{(j)}\right)^2}. \tag{8.8}$$

Similarly, the estimated influence of Gaussian constraints of E can be subtracted from σ_P using the approximate derivative $\partial_E P$ as a scaling factor, and the residual statistical and systematic uncertainties are then given by

$$\sigma_P' = \sqrt{\sigma_P^2 - \sum_j \left(\frac{\partial P}{\partial E}\sigma_E^{(j)}\right)^2} \qquad \delta_P' = \sqrt{\delta_P^2 - \sum_j \left(\delta_P^{(j)}\right)^2}. \tag{8.9}$$

These scalings are applied as appropriate for each analysis to achieve the needed separation of uncertainties.

The combination approach is to combine the components $\hat{P} \pm \sigma_P' \pm \delta_P'$ with GammaCombo, taking into account the correlation coefficients and sensitivities of S and C of each measurement. The systematic uncertainties determined in this measurement for the three parameters in question are then added in quadrature and assumed to be the same for all measurements. Consequently, the statistical and systematic uncertainties of the Run 1 measurements are improved. While a total uncertainty of 0.034 was reported in Ref. [14] for the LHCb Run 1 combination of S, we now find $\sigma(S)^{\text{Run 1}}_{\text{LHCb}} = 0.029$. In total, in all combinations involving Run 1 data, a systematic uncertainty of ± 0.0055 for the parameter S and ± 0.0028 for C is assigned for the influence of Δm_d, $\Delta\Gamma_d$ and α, which demonstrates again that the systematic uncertainties are dominated by the knowledge of external inputs, primarily influenced by the $\Delta\Gamma_d$ uncertainty. The LHCb Run 1 and Run 2 average is determined to be

$$\begin{aligned} S^{\text{LHCb}}_{\psi K^0_S} &= 0.724 \pm 0.014\,(\text{stat+syst})\,, \\ C^{\text{LHCb}}_{\psi K^0_S} &= 0.004 \pm 0.012\,(\text{stat+syst})\,, \end{aligned}$$

with a correlation coefficient of $\rho(S, C) = 0.40$. In Figure 8.4 (right), the new LHCb Run 1 combination is compared to the new Run 2 result, and the new LHCb global combination is shown as a black outline. Finally, a combination of all $B^0 \to J/\psi(\to \ell^+\ell^-)K^0_S$ decay channels is performed between Run 1 and Run 2 and results in

$$\begin{aligned} S^{\text{LHCb}}_{J/\psi K^0_S} &= 0.726 \pm 0.014\,(\text{stat+syst})\,, \\ C^{\text{LHCb}}_{J/\psi K^0_S} &= 0.010 \pm 0.012\,(\text{stat+syst})\,, \end{aligned}$$

with a correlation coefficient of 0.41.

9 Discussion

9.1 Updates of world averages and comparisons between experiments

The sensitivity achieved with the full Run 2 data set of LHCb exceeds the sensitivity of the world-average result for the parameters $S_{\psi K^0_S}$ and $C_{\psi K^0_S}$ [62]. The measurement presented in this thesis is therefore the most precise single measurement of *CP* violation in the B system and of the CKM angle β to date. It is thus worthwhile to explore some of the implications of this new result and to compare it to other measurements and theory predictions.

The *CP*-violating phase of the CKM matrix is a free parameter of the SM and cannot be derived from first principles. This fact does not mean that the consistency of results related to the CKM mechanism cannot be tested: the unitarity property of the CKM matrix is a strong condition that effectively correlates the values of CKM-matrix elements. Key observables like the side lengths and angles of unitarity triangles can thus be over-constrained and thereby predicted without inputs from direct measurements. Any measurement of CKM-matrix elements or their combinations always sharpens the global picture of the CKM mechanism. Global fits of the unitarity triangle apexes (CKM-fit), as they are performed by the CKMfitter group [52] or the UTfit collaboration [163, 164], are used to test the unitarity condition and to predict SM parameters for various scenarios. From the results of these global fits, the impact of new results can be quantified, and tensions can be uncovered. The selection of input parameters, the treatment of theory uncertainties, and the methods used for fitting differ between the two groups, but these choices do not significantly alter the implications of our new result.

A natural way of testing for New Physics is to verify whether the apex coordinates are consistent with a single point, as otherwise, the unitary assumption would not hold. It is furthermore possible to compare inclusive and exclusive predictions of a CKM parameter: inclusive predictions are obtained from fits in which the parameter value itself is an input to the fit and can therefore be viewed as a fit-average. Exclusive predictions of a CKM parameter include all relevant measurements except for direct measurements of the parameter itself. The latter result is referred to as a SM prediction in this context. A tension between an inclusive and an exclusive result does not necessarily point to New Physics in the CKM mechanism but can point to deficiencies in the theoretic assumptions made for certain measurements or the interpretation of their results.

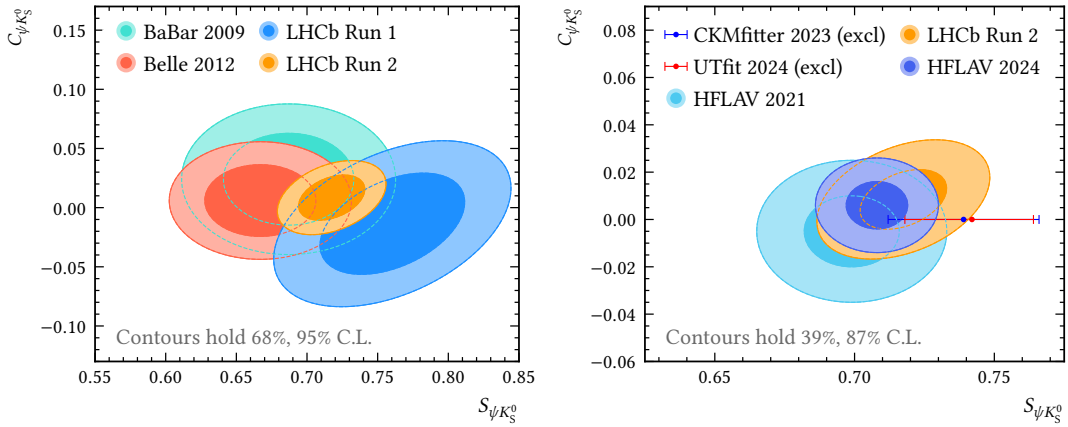

Figure 9.1 – Comparison of LHCb measurements and previous measurements by Belle and BaBar in the (S, C)-plane (left) and the change of HFLAV world-average values before and after this measurement in the (S, C)-plane (right). The exclusive predictions by the UTfit and CKMfitter groups for S [161, 162] are shown as data points, whereby $C = 0$ is set here.

Figure 9.1 (left) shows the results of previous measurements of the Belle, BaBar, and LHCb collaborations and the measurement presented here. There is excellent agreement between the results of different experiments so far. Both the CKMfitter and UTfit groups have published new unitarity triangle fits, which include our result [161, 162]. Furthermore, new HFLAV 2024 world averages[1] have been published [165]:

$$S^{\text{HFLAV}}_{J/\psi K^0_S} = 0.711 \pm 0.012 \qquad C^{\text{HFLAV}}_{J/\psi K^0_S} = 0.012 \pm 0.011$$
$$S^{\text{HFLAV}}_{\psi(2S) K^0_S} = 0.728 \pm 0.038 \qquad C^{\text{HFLAV}}_{\psi(2S) K^0_S} = -0.053 \pm 0.034$$
$$S^{\text{HFLAV}}_{[c\bar{c}]X} = 0.709 \pm 0.011 \qquad C^{\text{HFLAV}}_{[c\bar{c}]X} = 0.004 \pm 0.010.$$

So far, HFLAV world averages have been published without averaged correlation coefficients of the S and C parameters. The improvement between the current and previous world averages of S and C is shown in Figure 9.1 (right), where these experimental averages are compared to the new exclusive predictions published by CKMfitter and UTfit. The agreement between experimental averages and exclusive predictions for $\sin(2\beta)$ is below 1σ, as visualized in Figure 9.1. It should be noted that the significance of the difference between inclusive and exclusive fits by the CKMfitter group has increased from 0.8σ to 1.2σ between the 2021 and 2023 averages in light of the result presented in this thesis [52], which is only a small decrease in agreement. The significance of the difference reported by UTfit remains nearly the same between 2022 and 2023: it has increased from 1.4σ to 1.5σ [162]. In summary, the new inclusive SM predictions for

[1] These averages are published for the Moriond 2024 Winter conference and can be considered preliminary.

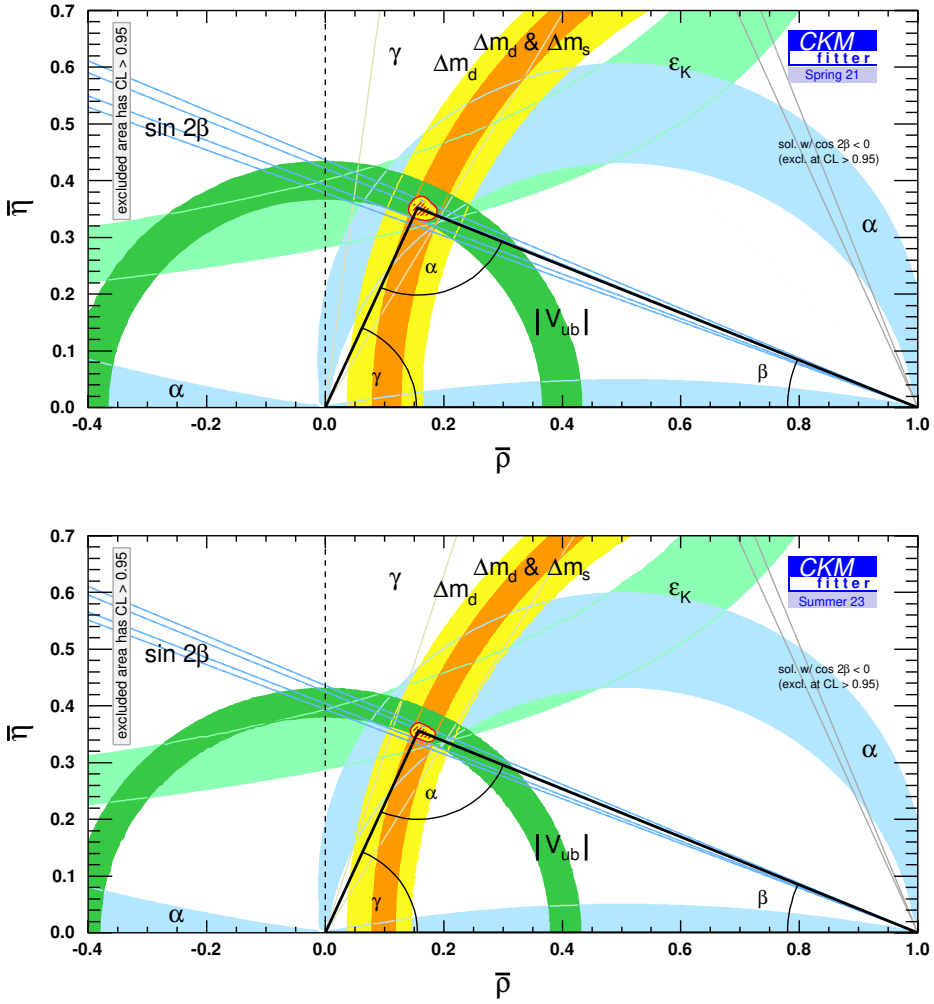

Figure 9.2 – Comparison of the global unitarity triangle fit by CKMfitter group showing the 2021 (top) and 2023 preliminary results (bottom) [52].

$\sin(2\beta)$ have been reported as [161, 162]

$$S_{2024}^{\text{UTfit}} = 0.705 \pm 0.014 \tag{9.1}$$

$$S_{2023}^{\text{CKMFitter}} = 0.7155^{+0.0079}_{-0.0071} \tag{9.2}$$

Penguin contributions have so far not been considered in the world averages due to a lack of empirical data, *i.e.* $S = \sin(2\beta)$ has been assumed for simplicity. This simplification is further justified by the small value of $\Delta\phi_d$ [68–73]. The relative uncertainty of $\Delta\phi_d$ is currently large and future measurements will be able to constrain its value more precisely [68]. Using a recent $\Delta\phi_d$ estimate from Ref. [69], $\Delta\phi_d(J/\psi K^0) = (-0.8^{+0.7}_{-1.8})^\circ$, one can trivially estimate the true size of the theoretical parameter $\sin(2\beta)$

$$\sin(\underbrace{\arcsin(0.709) - \Delta\phi_d}_{2\beta}) = 0.719^{+0.031}_{-0.001}, \tag{9.3}$$

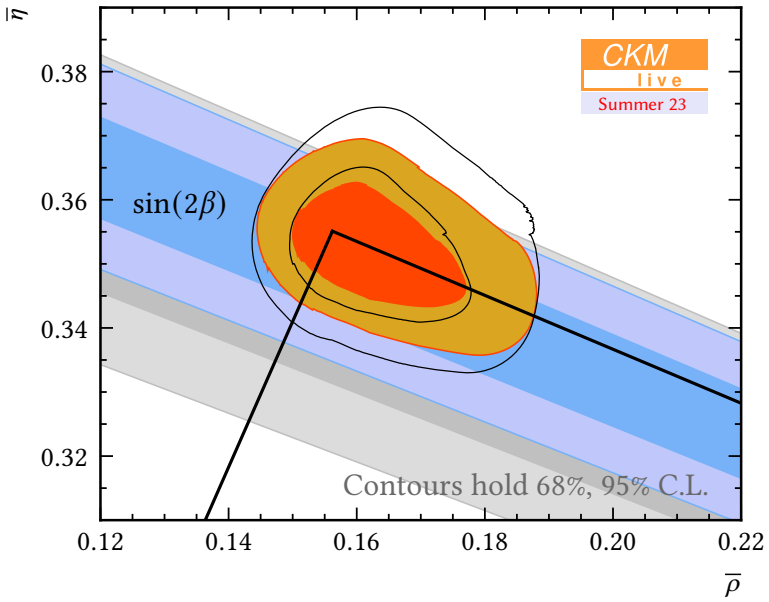

Figure 9.3 – Result of two unitarity triangle apex fits in the $\overline{\rho}$-$\overline{\eta}$ plane for the published 2023 CKMfitter inputs [161], illustrating the isolated impact of the measurement presented here. The black outline and gray confidence interval represent the contours of the 2023 preliminary fit where the old $\sin(2\beta)$ average is used as input, and the colored confidence intervals are the fit result for the new, world-average value of $\sin(2\beta)$.

where the given uncertainty stems from the $\Delta\phi_d$ uncertainty only. A negative value of $\Delta\phi_d$ corrects the value of $\sin(2\beta)$ upwards. The uncertainties of $\Delta\phi_d$ and $S_{\psi K^0_S}$ are therefore presently too large to interpret differences between inclusive and exclusive measurements as anything beyond the SM. In the future, measurements of $\Delta\phi_d$ will be of high relevance once more precise measurements of $\sin(2\beta)$ by LHCb and Belle II become available.

Figure 9.2 shows the result of the global CKM triangle fit as published by the CKMfitter group for the years 2021 and 2023 [52], where the latter fit includes our new result, amongst others. Improved measurements of γ, Δm_s, $|V_{ub}|$, and $\sin(2\beta)$ visibly increase the precision of the global CKM picture and the resolution of the unitarity triangle apex.

Using the CKMlive tool [166], which is kindly being provided by the CKMfitter group to the public, the impact of our result on the unitarity triangle apex contours can be explored. To that end, two global fits of the unitarity triangle apex are performed using the latest input parameters as of Summer 2023 listed in [161]. The first fit is performed using all the latest inputs, including the new CKMfitter-average for S, given in Equation (9.1) in a region close to the unitarity triangle apex. In the second fit, the same setup is used, where the previous $\sin(2\beta)$ world-average value is assumed instead. Figure 9.3 shows the results of the two unitarity triangle apex fits and illustrates the isolated impact of this measurement of $\sin(2\beta)$. Only a minimal movement of the

apex coordinates between 2021 and 2023 can be attributed to the new central value of $\sin(2\beta)$. The previous and the current predictions for the unitarity angle β by the CKMfitter group [52, 161] are shown as grey and blue regions. The improvement of the uncertainty of $\sin(2\beta)$ that is contributed by our measurement greatly decreases the uncertainty of the apex coordinates of the unitarity triangle.

9.2 Conclusion

Charge-parity violation is an integral part of the Standard Model of particle physics and is characteristic of processes mediated by the weak force. While most weak decays are *CP*-conserving, a relatively small subset violates *CP* symmetry. The small but fundamental differences between matter and antimatter interactions in nature become apparent in these decays. Time-dependent measurements of neutral B-meson oscillation provide a unique opportunity for precision measurements of *CP* violation and are thus useful for constraining the free parameter space of the Standard Model. More specifically, such measurements probe the Cabibbo-Kobayashi-Maskawa mechanism, which predicts *CP* violation in some particle decays where all three quark generations are involved in the decay amplitude. The source of *CP* violation in the CKM mechanism is a difference between the weak and mass-eigenstates of quarks which is parametrized by three mixing angles and a single, *CP*-violating phase. These parameters are encoded in a single, unitary CKM matrix and are, to our current knowledge, fundamental and cannot be precisely predicted from first principles. Only through a global analysis of all relevant measurements, the CKM parameters can be constrained to provide possible hints for New Physics.

The *CP*-violating decay $B^0 \to \psi K^0_S$ is exceptionally well suited for precision measurements of *CP*-violation for multiple reasons. The tree-level transition $b \to c\bar{c}s$ is strongly dominant and from a ratio of tree-level amplitudes parametrizing $B^0 \to \psi K^0_S$ and $\bar{B}^0 \to \psi K^0_S$ transitions, the CKM-angle β is accessible in a theoretically nearly clean way. The leading-order contribution from penguin-type decays contributes with the same weak phase as the tree-level process. The remaining higher-order contributions are strongly CKM suppressed (see Section 2.4). Experimentally, the decay signature of the $B^0 \to \psi K^0_S$ decay has characteristics that differentiate it decisively from combinatorial background: the B decay vertex is separated from the primary interaction point due to the Lorentz boost of B mesons produced in a hadronic collision environment and because of the comparably long lifetime of B^0 mesons. In addition, the identification of two high-momentum muons can be accomplished with high precision. Finally, other physical backgrounds can be well separated from the signal due to the excellent particle identification capabilities of the LHCb detector. From a measurement of the time-dependent asymmetry of the decay rates $\Gamma(B^0(t) \to \psi K^0_S)$ and $\Gamma(\bar{B}^0(t) \to \psi K^0_S)$, the strikingly large *CP*-violating effect arising from an interference of decay amplitudes with and without mixing is measured. The result from a combined fit to the full Run 2 data set, which was recorded by the LHCb detector in the time span from 2015 to 2018, is reported in Ref. [15] as

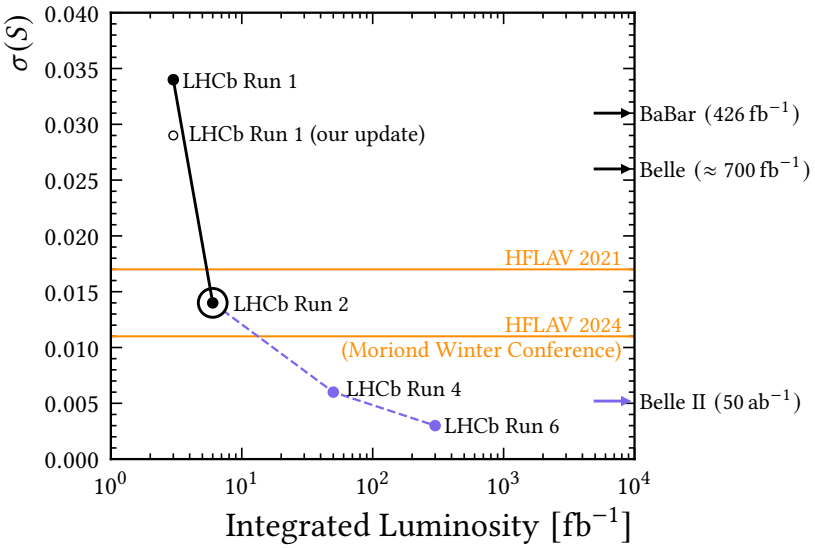

Figure 9.4 – Past, present (black) and expected, future sensitivities (blue) for the *CP*-violation parameter S shown for LHCb, and the experiments at the B-factories [10, 11] and recent world-averages (orange). The LHCb Run 1 average has been effectively improved by an improvement of systematic uncertainties of our result and is marked as "our update".

$$S_{\psi K^0_S} = 0.717 \pm 0.013\,(\text{stat}) \pm 0.008\,(\text{syst}),$$
$$C_{\psi K^0_S} = 0.008 \pm 0.012\,(\text{stat}) \pm 0.003\,(\text{syst}),$$

and surpasses the most recent world-average value of these parameters as of 2021 [62] in precision. The result is well compatible with previous measurements by the B-factories and by LHCb, and no significant increase in tensions pointing to New Physics is observed in light of this result. The result's sensitivity is still limited by the statistical uncertainty, and improved future measurements can therefore be performed with a larger future data set and currently available experimental methods at LHCb.

In the medium-term future, major improvements in measurements of $\sin(2\beta)$ can be expected. The LHCb detector is currently taking data in the Run 3 data taking phase. It will continue to take data in the high luminosity phase of the LHC, which is scheduled to start with Run 5 and to last at least until 2042 (up to Run 6). The LHCb detector is planned to record a total of $50\,\text{fb}^{-1}$ in Run 3 and Run 4 and is going to record $300\,\text{fb}^{-1}$ during Run 5 and Run 6 [167]. In addition, Belle II is expected to contribute with a competitive measurement once it has recorded $50\,\text{ab}^{-1}$ of data [168]. With an LHCb data set size of $300\,\text{fb}^{-1}$, an uncertainty as low as 0.003 is expected for S [167], which is somewhat lower than the currently projected Belle II uncertainty for this parameter and corresponds to an angular resolution of $\sigma(\beta) \approx 0.1°$. The decrease of the statistical and systematic uncertainty of S as a function of integrated luminosity at LHCb is shown in Figure 9.4. Integrated luminosities of the B factories are not projected onto the same axis due to the drastically different collision environment in e^+e^- collisions compared to proton-proton collisions. The LHCb Run 1 average for

S has been effectively improved with the Run 2 systematic uncertainties determined in this measurement, and its effective value, which is included in the updated LHCb-average (see Section 8.2), is shown in the plot ("our update"). Future measurements of $\sin(2\beta)$ in $B^0 \to \psi K^0_S$ decays by the ATLAS and CMS collaborations cannot be entirely ruled out, but are not likely, given that no such measurement has been published so far by these two collaborations [62].

Systematic uncertainties can be expected to reduce in size at LHCb once improved measurements of $\Delta\Gamma_d$ and Δm_d become available. Significant advances in our understanding of the proton-proton collision will be needed to reduce the sub-leading systematic uncertainties, which we attribute to flavor tagging transferability and determine in simulation. In particular, an accurate simulation of the charged track multiplicity created in proton collisions is needed, which influences particle identification performance and, in turn, flavor tagging. One possible approach to partially improve flavor tagging realism would be to train flavor tagging algorithms with already available, calibrated particle identification measures instead of the output of multivariate classifiers, which operate on raw input data from either data or simulation.

With a future angular precision for β at the level of $0.1°$, the measurement of the parameter $\Delta\phi_d(\psi K^0_S)$ potentially comes into reach. To differentiate this parameter from the phase 2β, which will be essential in the future, measurements of CP-violation in $B^0_s \to J/\psi K^0_S$, and $B^0 \to J/\psi\{\pi^0, \phi, \rho\}$ decays, amongst others, are needed to constrain the value of $\Delta\phi_d$ [68, 69] further. Such measurements would provide sensitivity for New Physics effects contributing to the loop transitions of these decays in addition [68].

Measurements of the oscillation parameter Δm_d, which is closely related to m_{12} in the B^0 system, constrain the unitarity triangle apex to a circle of radius proportional to $|V_{td}V^*_{tb}|$ in the $\overline{\rho}$-$\overline{\eta}$-plane (compare with Equation (2.13)) for the unitarity condition given in Equation (2.8). This constraint is represented in Figure 9.2 as a yellow circle around the $(1, 0)$ point. With the data taken so far, LHCb has the potential for competitive, if not highest precision measurements of the oscillation parameter Δm_d in the decay modes $B^0 \to \psi K^0_S$ and $B^0 \to \psi K^*$ alone, and other decay modes previously used for this measurement [169]. Precise knowledge of the parameters Δm_d and $\Delta\Gamma_d$ deepens our understanding of B-mixing and offers an interesting opportunity for New Physics searches. Specifically, it allows us to probe for the influence of new particles at energy scales of up to 20 TeV [170]. Finally, future measurements of Δm_d and β at LHCb will allow us to decisively constrain New Physics contributions to the mixing process [171].

Leading up to and after publication of our result, the Belle II collaboration published two measurements of $S_{J/\psi K^0_S}$ in $B^0 \to J/\psi K^0_S$ decays [172, 173], demonstrating their readiness for this measurement once the Belle II data has been recorded. These results are in agreement with our result and do not supersede it in precision.

In addition to the planned major increase of recorded data, which will result in significant improvements in all measurements of CP-violation, the optimization of flavor tagging algorithms is of central importance for LHCb. Currently, the tagging power is measurement-specific and ranges from ca. 4 to 6 %, depending on the B-meson type, signal kinematics and applied selection requirements. An increase in flavor

tagging performance, specifically tagging power, enables time-dependent *CP*-violation measurements to be performed on a larger subset of the available data. The current flavor tagging algorithms are specialized for selecting the most frequent physical processes occurring in the hadronizations and decays of the initial $b\bar{b}$ production. These algorithms analyze small selections of tracks, which likely originate from a selection of the most frequent processes that can be expected. In the past few years, the development of an *inclusive* flavor tagging algorithm was enforced by the collaboration, which performs a simultaneous evaluation of all reconstructed tracks in an event. Through this novel approach, the analysis of the rich physics of proton-proton collisions, which results in a multitude of correlations of produced tracks, will be possible. This approach utilizes machine-learning techniques, and the suitability of various models, ranging from BDT-adjacent models [174], to recurrent neural networks and most recently, Deep-Sets [175] has been tested and promising improvements of overall tagging power have been observed on simulation and data. The optimization and validation of this model are currently the subject of research and will be documented in a future publication.

Appendix

E Equations

E.1 Hypatia parameters

$$\begin{aligned} c_2 &= \frac{\sigma(\alpha_1^2(1-2\lambda)+n_1(2+2\lambda-\alpha_1^2))}{\alpha_1(2\lambda-1)} \quad c_1 = \left(1-\frac{\alpha_1^2}{2(\lambda+1)}\right)^{\left(\lambda-\frac{1}{2}\right)}(\alpha_1\sigma+c_2)^{n_1} \\ c_4 &= \frac{\sigma(\alpha_2^2(1-2\lambda)+n_2(2+2\lambda-\alpha_2^2))}{\alpha_2(2\lambda-1)} \quad c_3 = \left(1-\frac{\alpha_2^2}{2(\lambda+1)}\right)^{\left(\lambda-\frac{1}{2}\right)}(\alpha_2\sigma+c_4)^{n_2} \end{aligned} \tag{E.1}$$

E.2 Decay rate coefficients for two taggers

In the case of two taggers, labeled "A" and "B", there are four categories of data: candidates can be tagged by both taggers (AB), they can be tagged only by A or B, ($A\overline{B}$ or $\overline{A}B$), and lastly, they can be untagged by both ($\overline{AB}$). For a single tagger, the CP pdf that includes untagged events is

$$\begin{aligned} \mathcal{P}(q,\eta) \propto &\left[\epsilon^{+}\frac{|q|}{2}(1+q(1-\omega^{+}(\eta)))+(1-|q|)(1-\epsilon^{+})\right]P_{B^0}(t) \\ &+\left[\epsilon^{-}\frac{|q|}{2}(1+q(1-\omega^{-}(\eta)))+(1-|q|)(1-\epsilon^{-})\right]P_{\bar{B}^0}(t) \\ &\equiv \pi^{+}(\eta)P_{B^0}(t)+\pi^{-}(\eta)P_{\bar{B}^0}(t) \end{aligned} \tag{E.2}$$

whereby calibration parameters and the η-dependency of ω are implied from here on and $P_{B^0}(t)$ and $P_{\bar{B}^0}(t)$ are given in Equation (5.4). In the case of two taggers, the product probabilities of measuring initial B^0 (+) or $\bar{B}^0$ (−) mesons become

$$\pi^{\pm}(\boldsymbol{\eta}) = \prod_{n=\{A,B\}} |q_n|\epsilon_n^{\pm}\frac{1\pm q_n(1-2\omega_n^{\pm}(\eta_n))}{2}+(1-|q_n|)(1-\epsilon_n^{\pm}). \tag{E.3}$$

Expanding this expression and reordering it into the four tagging categories gives

$$\pi^{\pm}(\boldsymbol{\eta}) = \pi^{\pm}_{AB}(\boldsymbol{\eta})+\pi^{\pm}_{A\overline{B}}(\boldsymbol{\eta})+\pi^{\pm}_{\overline{A}B}(\boldsymbol{\eta})+\pi^{\pm}_{\overline{AB}}(\boldsymbol{\eta}), \tag{E.4}$$

with

$$\pi^{\pm}_{AB} = |q_A q_B| \times \frac{\epsilon^{\pm}_A \epsilon^{\pm}_B}{4}[1 \pm q_A D^{\pm}_A][1 \pm q_B D^{\pm}_B], \tag{E.5}$$

$$\pi^{\pm}_{A\overline{B}} = |q_A|(1 - |q_B|) \times \frac{\epsilon^{\pm}_A}{2}(1 - \epsilon^{\pm}_B)[1 \pm q_A D^{\pm}_A], \tag{E.6}$$

$$\pi^{\pm}_{\overline{A}B} = |q_B|(1 - |q_A|) \times \frac{\epsilon^{\pm}_B}{2}(1 - \epsilon^{\pm}_A)[1 \pm q_B D^{\pm}_B], \tag{E.7}$$

$$\pi^{\pm}_{\overline{AB}} = (1 - |q_A|)(1 - |q_B|) \times (1 - \epsilon^{\pm}_A)(1 - \epsilon^{\pm}_B), \tag{E.8}$$

whereby $D^{\pm}_{A,B}(\boldsymbol{\eta}) = (1 - 2\omega^{\pm}_{A,B}(\boldsymbol{\eta}))$. Equations (E.5) to (E.8) are then inserted into Equation (E.2) and regrouped into common coefficients for S, C and H:

$$\begin{aligned}\mathcal{P}(\boldsymbol{q},\boldsymbol{\eta})e^{\Gamma_d t/\hbar} \propto\ & \pi^{+}(\boldsymbol{\eta}) \cdot (1-\alpha)[H - S\sin(\Delta m_d t) + C\cos(\Delta m_d t)] \\ & + \pi^{-}(\boldsymbol{\eta}) \cdot (1+\alpha)[H + S\sin(\Delta m_d t) - C\cos(\Delta m_d t)]\end{aligned} \tag{E.9}$$

$$= S \times \sum_{X=\{AB,A\overline{B},\overline{A}B,\overline{AB}\}} (\pi^{-}_X - \pi^{+}_X) + \alpha(\pi^{+}_X + \pi^{-}_X) \tag{E.10}$$

$$+ C \times \sum_{X=\{AB,A\overline{B},\overline{A}B,\overline{AB}\}} (\pi^{+}_X - \pi^{-}_X) - \alpha(\pi^{+}_X + \pi^{-}_X) \tag{E.11}$$

$$+ H \times \sum_{X=\{AB,A\overline{B},\overline{A}B,\overline{AB}\}} (\pi^{+}_X + \pi^{-}_X) + \alpha(\pi^{-}_X - \pi^{+}_X). \tag{E.12}$$

Terms (E.10) to (E.12) represent the coefficients $\tilde{S}(\boldsymbol{q},\boldsymbol{\eta})$, $\tilde{C}(\boldsymbol{q},\boldsymbol{\eta})$ and $\tilde{H}(\boldsymbol{q},\boldsymbol{\eta})$ in Equation (5.18), respectively. As demonstrated in Section 5.1, the substitutions $\epsilon^{\pm} \to (\epsilon_{\text{tag}} \pm \frac{\Delta\epsilon}{2})$, $\Delta\omega_{A,B} = \omega^{+}_{A,B} - \omega^{-}_{A,B}$ can be made, yet expanding the obtained expression further would not aid our understanding. Despite being more complex, these expressions closely resemble the single-tagger case qualitatively. It should be noted that since the data category $\overline{AB}$ is not included in our fit, the tagging efficiencies of each tagger are computed at the time of fitting based on the tagged statistics of the input data.

T Supplementary tables

T.1 BDT features

The BDT training features are defined as follows

$\mathrm{IP}(x)$	Impact parameter of reconstructed track or trajectory of particle x
$\mathrm{DOCA}(x, y)$	Distance of closest approach between two tracks or trajectories of particles x and y
$\chi^2_{\mathrm{DOCA}}(x, y)$	χ^2 value of the computed distance of closest approach between two tracks or trajectories of particles x and y
$p_{\mathrm{T}}(x)$	Transverse momentum of particle x
$p_z(x)$	z-component of momentum of particle x
χ^2_{Fit}	Global fit χ^2 value of the decay tree fitter [119] of an event, which is configured to substitute mass hypotheses for ψ and K^0_{S} candidates
$\chi^2_{\mathrm{Vtx}}(x)$	χ^2 value of the decay-vertex fit of a given particle x
$\eta(x)$	Pseudorapidity of particle x
$\mathrm{FD}(x)$	Flight distance of particle x, *i.e.* distance between reconstructed origin and decay vertices of particle x
$\mathrm{MinIP}(x)$	Minimum impact parameter of particle x with respect to any PV
$\tau(x)$	Reconstructed decay time of particle x

The following BDT features are used per channel and track type:

Dimuon channels

All track types $\mathrm{IP}(B^0)$, $\mathrm{IP}(J/\psi)$, $\mathrm{IP}(K^0_{\mathrm{S}})$, $p_{\mathrm{T}}(\pi^+)$, $p_{\mathrm{T}}(\pi^-)$, $p_{\mathrm{T}}(K^0_{\mathrm{S}})$, χ^2_{Fit}, $\eta(B^0)$, $\chi^2_{\mathrm{Vtx}}(B^0)$, $\mathrm{FD}(K^0_{\mathrm{S}})$, $\mathrm{DOCA}(\pi^+, \pi^-)$, $\eta(K^0_{\mathrm{S}})$, $p_z(K^0_{\mathrm{S}})$, $\mathrm{MinIP}(\mu^\pm)$, $\min(p_{\mathrm{T}}(\pi^+), p_{\mathrm{T}}(\pi^-))$

Exclusive to LL $\mathrm{DOCA}(J/\psi, K^0_{\mathrm{S}})$, $\chi^2_{\mathrm{DOCA}}(J/\psi, K^0_{\mathrm{S}})$, $\mathrm{MinIP}(K^0_{\mathrm{S}})$

Exclusive to DD $\mathrm{IP}(\pi^+)$, $\mathrm{IP}(\pi^-)$, $\tau(K^0_{\mathrm{S}})$, $\mathrm{MinIP}(\pi^\pm)$, $\mathrm{MinIP}(\pi^-)$, $\mathrm{MinIP}(\pi^+)$

Exclusive to LD or UL $\mathrm{IP}(\pi^+)$, $\mathrm{IP}(\pi^-)$, $\mathrm{DOCA}(J/\psi, K^0_{\mathrm{S}})$, $\chi^2_{\mathrm{DOCA}}(J/\psi, K^0_{\mathrm{S}})$, $\mathrm{MinIP}(K^0_{\mathrm{S}})$, $\tau(K^0_{\mathrm{S}})$, $\mathrm{MinIP}(\pi^\pm)$, $\mathrm{MinIP}(\pi^-)$, $\mathrm{MinIP}(\pi^+)$

Dielectron channel

All track types $\mathrm{IP}(B^0)$, $\mathrm{IP}(J/\psi)$, $\mathrm{IP}(K^0_{\mathrm{S}})$, $\mathrm{IP}(\pi^+)$, $p_{\mathrm{T}}(\pi^+)$, $\mathrm{IP}(\pi^-)$, $p_{\mathrm{T}}(\pi^-)$, $p_{\mathrm{T}}(K^0_{\mathrm{S}})$, χ^2_{Fit}, $\eta(B^0)$, $\chi^2_{\mathrm{Vtx}}(B^0)$, $\mathrm{MinIP}(e^\pm)$, $\mathrm{FD}(K^0_{\mathrm{S}})$, $\mathrm{DOCA}(\pi^+, \pi^-)$, $\eta(K^0_{\mathrm{S}})$, $p_z(K^0_{\mathrm{S}})$, $\tau(K^0_{\mathrm{S}})$, $\mathrm{MinIP}(\pi^\pm)$, $\min(p_{\mathrm{T}}(\pi^+), p_{\mathrm{T}}(\pi^-))$

Exclusive to LL $\mathrm{DOCA}(J/\psi, K^0_{\mathrm{S}})$, $\chi^2_{\mathrm{DOCA}}(J/\psi, K^0_{\mathrm{S}})$

T.2 Other tables

Table T.2.1 – Values of relevant branching ratios and constants.

$\Gamma_i/\Gamma_{\text{tot}}$	Value	Source
$\mathcal{B}(B^0 \to J/\psi K^0)$	$(8.91 \pm 0.21) \times 10^{-4}$	[116]
$\mathcal{B}(B^0 \to \psi(2S) K^0)$	$(5.8 \pm 0.5) \times 10^{-4}$	[116]
$\mathcal{B}(B_s^0 \to J/\psi K_S^0)$	$(1.92 \pm 0.14) \times 10^{-5}$	[116]
$\mathcal{B}(B_s^0 \to \psi(2S) K^0)$	$(1.9 \pm 0.5) \times 10^{-5}$	[128]
f_s/f_d (13 TeV)	0.2539 ± 0.0079	[176]

Table T.2.2 – Summary of selection and reconstruction efficiencies for $B^0 \to J/\psi(\to \mu\mu)K_S^0$ candidates determined in simulation.

	Efficiency [%]			
Selection	LL	DD	LD	UL
LHCb geometry		19.897 ± 0.083		
Reconstruction	1.592 ± 0.003	3.792 ± 0.004	0.255 ± 0.001	0.562 ± 0.002
Λ-Veto	99.051 ± 0.015	98.209 ± 0.014	99.254 ± 0.034	98.585 ± 0.032
BDT selection	87.96 ± 0.05	79.37 ± 0.04	79.96 ± 0.16	79.52 ± 0.11
K^*-Veto	97.961 ± 0.024	100.0 ± 0	99.972 ± 0.008	99.205 ± 0.027
B^+-Veto	99.536 ± 0.012	100.0 ± 0	99.709 ± 0.024	99.818 ± 0.013
Decay-time range	98.814 ± 0.018	98.507 ± 0.014	99.25 ± 0.04	99.605 ± 0.019
Multiple candidates	99.678 ± 0.010	99.689 ± 0.007	99.25 ± 0.04	98.839 ± 0.033
Tagging efficiency	84.20 ± 0.06	87.04 ± 0.04	84.56 ± 0.16	82.20 ± 0.12
Total	0.2304 ± 0.0011	0.5026 ± 0.0022	0.033 47 ± 0.000 22	0.0703 ± 0.0004

Table T.2.3 – Summary of selection and reconstruction efficiencies for $B^0 \to \psi(2S)K_S^0$ candidates determined in simulation.

	Efficiency [%]	
Selection	LL	DD
LHCb geometry	19.57 ± 0.05	
Reconstruction	2.104 ± 0.004	4.274 ± 0.006
Λ-Veto	99.161 ± 0.019	98.152 ± 0.019
BDT selection	85.12 ± 0.07	82.84 ± 0.05
K^*-Veto	98.246 ± 0.029	100.0 ± 0
B^+-Veto	99.494 ± 0.016	100.0 ± 0
Decay-time range	98.418 ± 0.028	97.563 ± 0.025
Multiple candidates	99.978 ± 0.003	99.757 ± 0.008
Tagging efficiency	84.19 ± 0.08	87.25 ± 0.05
Total	0.2814 ± 0.0019	0.5775 ± 0.0018

Table T.2.4 – Summary of selection and reconstruction efficiencies for $B^0 \to J/\psi(\to ee)K^0_S$ candidates determined in simulation. The main difference in the preselection efficiency depend on the flight distance requirements of the LL and DD kaons.

Selection	Efficiency [%] LL	DD
LHCb geometry	19.75 ± 0.05	
Reconstruction	0.372 ± 0.002	1.037 ± 0.003
Λ-Veto	99.384 ± 0.035	98.321 ± 0.034
BDT selection	94.55 ± 0.10	76.91 ± 0.11
K^*-Veto	99.887 ± 0.015	100.0 ± 0
B^+-Veto	99.517 ± 0.032	100.0 ± 0
Decay-time range	100.0 ± 0	99.994 ± 0.002
Multiple candidates	99.937 ± 0.012	99.763 ± 0.015
Tagging efficiency	89.23 ± 0.14	90.37 ± 0.09
Total	0.061 26 ± 0.000 34	0.1396 ± 0.0006

Correlation matrix applied in the fit to $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$:

	p_0^{SS}	p_1^{SS}	Δp_0^{SS}	Δp_1^{SS}	p_0^{OS}	p_1^{OS}	Δp_0^{OS}	Δp_1^{OS}	α
p_0^{SS}	1	0.038	−0.011	−0.004	0	0.001	0	0	0
p_1^{SS}		1	−0.005	0.004	0.001	0	0	0	0.001
Δp_0^{SS}			1	0.016	0	0	0.008	−0.005	−0.071
Δp_1^{SS}				1	0	0	−0.007	0.005	0.058
p_0^{OS}					1	0.112	−0.008	−0.007	0.001
p_1^{OS}						1	−0.007	−0.001	0.001
Δp_0^{OS}							1	0.089	−0.060
Δp_1^{OS}								1	0.037
α									1

(T.13)

Correlation matrix applied in the fit to $B^0 \to J/\psi(\to e^+e^-)K^0_S$:

	p_0^{SS}	p_1^{SS}	Δp_0^{SS}	Δp_1^{SS}	p_0^{OS}	p_1^{OS}	Δp_0^{OS}	Δp_1^{OS}	α
p_0^{SS}	1	0.057	0.004	−0.007	0	0.004	0.003	−0.004	−0.001
p_1^{SS}		1	−0.007	0.023	0.006	0.003	0	0.001	0
Δp_0^{SS}			1	0.023	0	0	0.007	−0.004	−0.027
Δp_1^{SS}				1	0.001	0	−0.006	0.003	0.022
p_0^{OS}					1	0.078	0.001	−0.017	0
p_1^{OS}						1	−0.018	−0.035	0.001
Δp_0^{OS}							1	0.044	−0.024
Δp_1^{OS}								1	0.014
α									1

(T.14)

F Supplementary figures

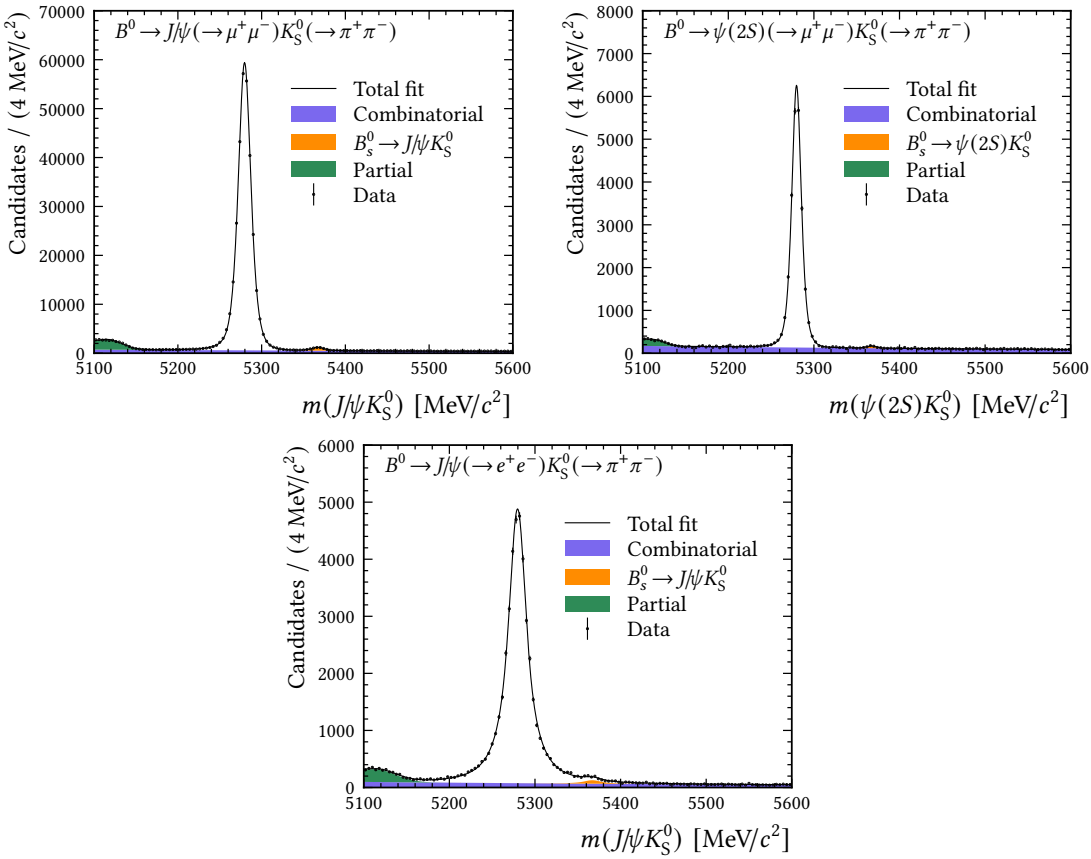

Figure F.1 – Mass fits of the analyzed decay modes of the full Run 2 data set in linear y-scale.

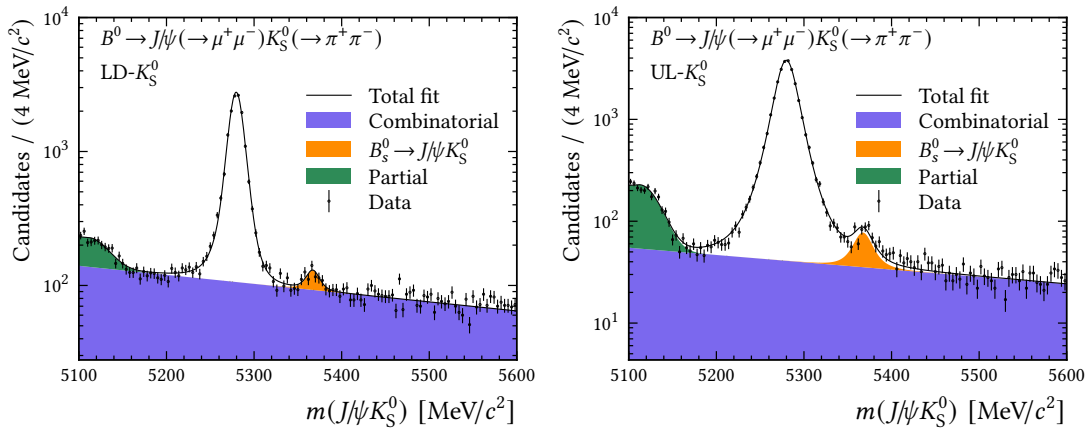

Figure F.2 – Mass fits of the decay mode $B^0 \to J/\psi(\to \mu^+\mu^-)K^0_S$ for the two new kaon reconstruction classes LD (left) and UL (right).

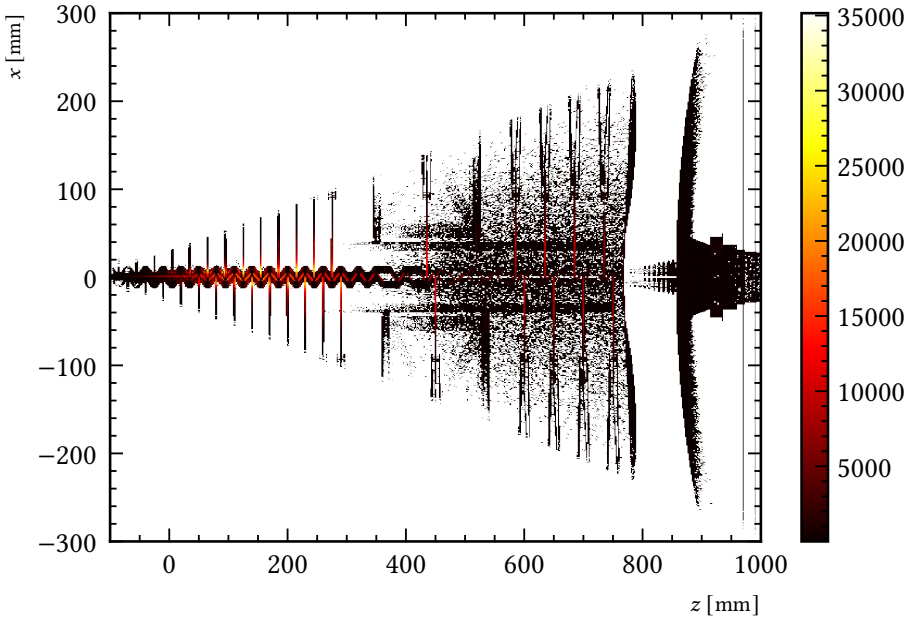

Figure F.3 – Control plot mapping the intersections of kaon tracks with the LHCb detector volumes in the x-z-plane. Most interactions occur in the VELO RF-foil.

References

[1] J. H. Christenson et al. *Evidence for the 2π decay of the K_2^0 meson*. Phys. Rev. Lett. **13** (1964) 138.
DOI: 10.1103/PhysRevLett.13.138.

[2] A. B. Carter and A. I. Sanda. *CP violation in B-meson decays*. Phys. Rev. D **23** (1981) 1567.
DOI: 10.1103/PhysRevD.23.1567.

[3] I. I. Bigi and A. I. Sanda. *Notes on the observability of CP violations in B decays*. Nucl. Phys. B **193** (1981) 85.
DOI: 10.1016/0550-3213(81)90519-8.

[4] I. I. Bigi and A. I. Sanda. *CP violation in heavy flavor decays: Predictions and search strategies*. Nucl. Phys. B **281** (1987) 41.
DOI: 10.1016/0550-3213(87)90246-X.

[5] R. Barate et al. (ALEPH collaboration). *Study of the CP asymmetry of $B^0 \to J/\psi K_S^0$ decays in ALEPH*. Phys. Lett. B **492** (2000) 259.
DOI: 10.1016/S0370-2693(00)01091-1.

[6] K. Ackerstaff et al. (OPAL collaboration). *Investigation of CP violation in $B^0 \to J/\psi K_S^0$ decays at LEP*. Eur. Phys. J. C **5** (1998) 379.
DOI: 10.1007/s100520050284.

[7] D. Bortoletto. *The measurement of* $\sin(2\beta)$. Proceedings of Science **hf8** (1999) 038.
DOI: 10.22323/1.003.0038.

[8] K. Abe et al. (Belle collaboration). *Observation of large CP violation in the neutral B meson system*. Phys. Rev. Lett. **87** (2001) 091802.
DOI: 10.1103/PhysRevLett.87.091802.

[9] B. Aubert et al. (BaBar collaboration). *Observation of CP violation in the B^0 meson system*. Phys. Rev. Lett. **87** (2001) 091801.
DOI: 10.1103/PhysRevLett.87.091801.

[10] I. Adachi et al. (Belle collaboration). *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.

[11] B. Aubert et al. (BaBar collaboration). *Measurement of time-dependent CP asymmetry in $B^0 \to c\bar{c}K^{(*)0}$ decays*. Phys. Rev. D **79** (2009) 072009.
DOI: 10.1103/PhysRevD.79.072009.

[12] R. Aaij et al. (LHCb collaboration). *Measurement of the time-dependent CP asymmetry in $B^0 \to J/\psi K^0_S$ decays*. Phys. Lett. B **721** (2013) 24.
DOI: 10.1016/j.physletb.2013.02.054.

[13] R. Aaij et al. (LHCb collaboration). *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.

[14] R. Aaij et al. (LHCb collaboration). *Measurement of CP violation in $B^0 \to J/\psi K^0_S$ and $B^0 \to \psi(2S) K^0_S$ decays*. J. High Energy Phys. **11** (2017) 170.
DOI: 10.1007/JHEP11(2017)170.

[15] R. Aaij et al. (LHCb collaboration). *Measurement of CP violation in $B^0 \to \psi(\to \ell^+\ell^-) K^0_S(\to \pi^+\pi^-)$ decays*. Phys. Rev. Lett. **132** (2024) 021801.
DOI: 10.1103/PhysRevLett.132.021801.

[16] P. Li and V. Jevtić. *Measurements of $\sin(2\beta)$ and ϕ_s with the full LHCb data sample*. LHC Seminar, CERN. 2023.
URL: https://cds.cern.ch/record/2862950.

[17] M. Gell-Mann. *The eightfold way: A theory of strong interaction symmetry* (1961).
DOI: 10.2172/4008239.

[18] A. Salam. *Gauge unification of fundamental forces*. Rev. Mod. Phys. **52** (1980) 525.
DOI: 10.1103/RevModPhys.52.525.

[19] S. Weinberg. *A model of leptons*. Phys. Rev. Lett. **19** (1967) 1264.
DOI: 10.1103/PhysRevLett.19.1264.

[20] A. J. Buras. *Gauge theory of weak decays: The standard model and the expedition to new physics summits*. Cambridge University Press, 2020.
ISBN: 978-1-107-03403-7.

[21] A. D. Sakharov. *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*. Sov. Phys. Usp. **34** (1991) 392.
DOI: 10.1070/PU1991v034n05ABEH002497.

[22] G. C. Branco, L. Lavoura, and J. P. Silva. *CP violation*. International series of monographs on physics: 103. Oxford University Press, 2014.
ISBN: 9780198716754.

[23] H. Georgi and S. L. Glashow. *Unity of all elementary-particle forces*. Phys. Rev. Lett. **32** (1974) 438.
DOI: 10.1103/PhysRevLett.32.438.

[24] S. Navas et al. (Particle Data Group). *Review of particle physics (PDGLive)*. To be published in Phys. Rev. D. 110, 030001 (2024). Accessed: 17.06.24.
URL: https://pdglive.lbl.gov.

[25] Q. R. Ahmad et al. (SNO collaboration). *Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory*. Phys. Rev. Lett. **89** (2002) 011301.
DOI: 10.1103/PhysRevLett.89.011301.

[26] R. Aaij et al. (LHCb collaboration). *Observation of an exotic narrow doubly charmed tetraquark*. Nat. Phys. **18** (2022) 751.
DOI: 10.1038/s41567-022-01614-y.

[27] R. Aaij et al. (LHCb collaboration). *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi K^- p$ decays*. Phys. Rev. Lett. **115** (2015) 072001.
DOI: 10.1103/PhysRevLett.115.072001.

[28] R. Aaij et al. (LHCb collaboration). *Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$*. Phys. Rev. Lett. **122** (2019) 222001.
DOI: 10.1103/PhysRevLett.122.222001.

[29] E. Fermi. *Versuch einer Theorie der β-Strahlen. I.* Z. Phys. **88** (1934) 161.
DOI: 10.1007/BF01351864.

[30] P. W. Higgs. *Broken symmetries and the masses of gauge bosons*. Phys. Rev. Lett. **13** (1964) 508.
DOI: 10.1103/PhysRevLett.13.508.

[31] F. Englert and R. Brout. *Broken symmetry and the mass of gauge vector mesons*. Phys. Rev. Lett. **13** (1964) 321.
DOI: 10.1103/PhysRevLett.13.321.

[32] Y. Nambu. *Quasi-particles and gauge invariance in the theory of superconductivity*. Phys. Rev. **117** (1960) 648.
DOI: 10.1103/PhysRev.117.648.

[33] J. Goldstone. *Field theories with «Superconductor» solutions*. Nuovo Cimento **19** (1961) 154.
DOI: 10.1007/BF02812722.

[34] J. Goldstone, A. Salam, and S. Weinberg. *Broken symmetries*. Phys. Rev. **127** (1962) 965.
DOI: 10.1103/PhysRev.127.965.

[35] S. L. Glashow. *Partial-symmetries of weak interactions*. Nucl. Phys. **22** (1961) 579.
DOI: 10.1016/0029-5582(61)90469-2.

[36] M. Banner et al. (UA2 collaboration). *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider*. Phys. Lett. B **122** (1983) 476.
DOI: 10.1016/0370-2693(83)91605-2.

[37] F. J. Hasert et al. (Gargamelle collaboration). *Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment*. Phys. Lett. B **46** (1973) 138.
DOI: 10.1016/0370-2693(73)90499-1.

[38] G. Aad et al. (ATLAS collaboration). *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** (2012) 1.
DOI: 10.1016/j.physletb.2012.08.020.

[39] S. Chatrchyan et al. (CMS collaboration). *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. Phys. Lett. B **716** (2012) 30.
DOI: 10.1016/j.physletb.2012.08.021.

[40] T. D. Lee and C. N. Yang. *Question of parity conservation in weak interactions*. Phys. Rev. **104** (1956) 254.
DOI: 10.1103/PhysRev.104.254.

[41] T. D. Lee and C. N. Yang. *Parity nonconservation and a two-component theory of the neutrino*. Phys. Rev. **105** (1957) 1671.
DOI: 10.1103/PhysRev.105.1671.

[42] C. S. Wu et al. *Experimental test of parity conservation in beta decay*. Phys. Rev. **105** (1957) 1413.
DOI: 10.1103/PhysRev.105.1413.

[43] R. L. Garwin, L. M. Lederman, and M. Weinrich. *Observations of the failure of conservation of parity and charge conjugation in meson decays: The magnetic moment of the free muon*. Phys. Rev. **105** (1957) 1415.
DOI: 10.1103/PhysRev.105.1415.

[44] R. P. Feynman and M. Gell-Mann. *Theory of the Fermi interaction*. Phys. Rev. **109** (1958) 193.
DOI: 10.1103/PhysRev.109.193.

[45] N. Cabibbo. *Unitary symmetry and leptonic decays*. Phys. Rev. Lett. **10** (1963) 531.
DOI: 10.1103/PhysRevLett.10.531.

[46] M. Kobayashi and T. Maskawa. *CP-violation in the renormalizable theory of weak interaction*. Prog. of Theor. Phys. **49** (1973) 652.
DOI: 10.1143/PTP.49.652.

[47] N. Tuning. *Lecture notes on CP violation*. 2020.
URL: https://www.nikhef.nl/~h71/Lectures/2015/ppII-cpviolation-29012015.pdf.

[48] L.-L. Chau and W.-Y. Keung. *Comments on the parametrization of the Kobayashi-Maskawa matrix*. Phys. Rev. Lett. **53** (1984) 1802.
DOI: 10.1103/PhysRevLett.53.1802.

[49] C. Jarlskog. *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation.* Phys. Rev. Lett. **55** (1985) 1039.
DOI: 10.1103/PhysRevLett.55.1039.

[50] L. Wolfenstein. *Parametrization of the Kobayashi-Maskawa matrix.* Phys. Rev. Lett. **51** (1983) 1945.
DOI: 10.1103/PhysRevLett.51.1945.

[51] A. Ceccucci, Z. Ligeti, and Y. Sakai. *CKM quark-mixing matrix* (2020). Review in Ref. [116].

[52] J. Charles et al. (CKMfitter Group). *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories.* Eur. Phys. J. C **41** (2005) 1. Updated results and plots available at: http://ckmfitter.in2p3.fr.
DOI: 10.1140/epjc/s2005-02169-1.

[53] M. Gell-Mann and A. Pais. *Behavior of neutral particles under charge conjugation.* Phys. Rev. **97** (1955) 1387.
DOI: 10.1103/PhysRev.97.1387.

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

[55] R. H. Good et al. *Regeneration of neutral K mesons and their mass difference.* Phys. Rev. **124** (1961) 1223.
DOI: 10.1103/PhysRev.124.1223.

[56] T. Gershon and Y. Nir. *CP violation in the quark sector* (2022). Review in Ref. [116].

[57] O. Schneider. *B^0-$\bar{B}^0$ mixing* (2022). Review in Ref. [116].

[58] H. Cheng. *CP-violating effects in heavy-meson systems.* Phys. Rev. D **26** (1982) 143.
DOI: 10.1103/PhysRevD.26.143.

[59] A. J. Buras, W. Słominski, and H. Steger. *B^0–$\bar{B}^0$ mixing, CP violation and the B-meson decay.* Nucl. Phys. B **245** (1984) 369.
DOI: 10.1016/0550-3213(84)90437-1.

[60] T. Inami and C. S. Lim. *Effects of superheavy quarks and leptons in low-energy weak processes $K_L^0 \to \mu\bar{\mu}$, $K^+ \to \pi^+\nu\bar{\nu}$ and $K^0 \leftrightarrow \bar{K}^0$.* Prog. of Theor. Phys. **65** (1981) 297.
DOI: 10.1143/PTP.65.297.

[61] C. Gay. *B mixing.* Ann. Rev. Nucl. Part. Sci. **50** (2000) 577.
DOI: 10.1146/annurev.nucl.50.1.577.

[62] Y. Amhis et al. *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021*. Phys. Rev. D **107** (2023) 052008. Updated results and plots available at `https://hflav.web.cern.ch`.
DOI: 10.1103/PhysRevD.107.052008.

[63] H. Wahl. *First observation and precision measurement of direct CP violation: The experiments NA31 and NA48*. Physics Reports **403-404** (2004) 19.
DOI: 10.1016/j.physrep.2004.08.007.

[64] A. Alavi-Harati et al. (KTeV collaboration). *Observation of direct CP violation in $K_{S,L} \to \pi\pi$ decays*. Phys. Rev. Lett. **83** (1999) 22.
DOI: 10.1103/PhysRevLett.83.22.

[65] L. Wolfenstein, C.-J. Lin, and T. G. Trippe. *CP violation in K^0_L decays* (2016). Review in Ref. [116].

[66] 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. C **10** (1999) 299.
DOI: 10.1007/s100529900099.

[67] Y. H. Ahn, H.-Y. Cheng, and S. Oh. *Wolfenstein parametrization at higher order: Seeming discrepancies and their resolution*. Phys. Lett. B **703** (2011) 571.
DOI: 10.1016/j.physletb.2011.08.047.

[68] K. De Bruyn and R. Fleischer. *A roadmap to control penguin effects in $B^0 \to J/\psi K^0_S$ and $B^0_s \to J/\psi\phi$*. J. High Energy Phys. **03** (2015) 145.
DOI: 10.1007/JHEP03(2015)145.

[69] M. Z. Barel et al. *In pursuit of new physics with $B^0_d \to J/\psi K^0$ and $B^0_s \to J/\psi\phi$ decays at the high-precision frontier*. J. Phys. G **48** (2021) 065002.
DOI: 10.1088/1361-6471/abf2a2.

[70] M. Ciuchini, M. Pierini, and L. Silvestrini. *Effect of penguin operators in the $B^0 \to J/\psi K^0$ CP asymmetry*. Phys. Rev. Lett. **95** (2005) 221804.
DOI: 10.1103/PhysRevLett.95.221804.

[71] S. Faller et al. *The golden modes $B^0 \to J/\psi K_{S,L}$ in the era of precision flavour physics*. Phys. Rev. D **79** (2009) 014030.
DOI: 10.1103/PhysRevD.79.014030.

[72] M. Jung. *Determining weak phases from $B \to J/\psi P$ decays*. Phys. Rev. D **86** (2012) 053008.
DOI: 10.1103/PhysRevD.86.053008.

[73] 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.

[74] D. Boutigny et al. (BaBar collaboration). *The BABAR physics book: Physics at an asymmetric B factory*. 1998.
DOI: 10.2172/979931.

[75] I. Adachi et al. (BaBar collaboration and Belle collaboration). *Measurement of* $\cos 2\beta$ *in* $B^0 \to D^{(*)}h^0$ *with* $D \to K^0_S\pi^+\pi^-$ *decays by a combined time-dependent Dalitz plot analysis of BaBar and Belle data*. Phys. Rev. D **98** (2018) 112012.
DOI: 10.1103/PhysRevD.98.112012.

[76] Y. Grossman and H. R. Quinn. *Removing discrete ambiguities in CP asymmetry measurements*. Phys. Rev. D **56** (1997) 7259.
DOI: 10.1103/PhysRevD.56.7259.

[77] B. Kayser and D. London. *B-decay CP asymmetries, discrete ambiguities and new physics*. Phys. Rev. D **61** (2000) 116012.
DOI: 10.1103/PhysRevD.61.116012.

[78] 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. C **72** (1996) 543.
DOI: 10.1007/s002880050277.

[79] B. R. Ko et al. *Effect of nuclear interactions of neutral kaons on CP asymmetry measurements*. Phys. Rev. D **84** (2011) 111501.
DOI: 10.1103/PhysRevD.84.111501.

[80] O. S. Brüning et al. *LHC design report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004.
DOI: 10.5170/CERN-2004-003-V-1.

[81] E. Mobs. *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN- Août 2018* (2018).
URL: https://cds.cern.ch/record/2636343.

[82] D. J. Warner. *Project study for a new 50 MeV linear accelerator for the C.P.S.* CERN-MPS-LINP-73-1 (1973).
URL: https://cds.cern.ch/record/414071.

[83] J. Wenninger. *Operation and configuration of the LHC in Run 2* (2019). CERN-ACC-NOTE-2019-0007.
URL: https://cds.cern.ch/record/2668326.

[84] G. Aad et al. (ATLAS collaboration). *The ATLAS experiment at the CERN Large Hadron Collider*. J. Instrum. **3** (2008) S08003.
DOI: 10.1088/1748-0221/3/08/S08003.

[85] S. Chatrchyan et al. (CMS collaboration). *The CMS experiment at the CERN LHC*. J. Instrum. **3** (2008) S08004.
DOI: 10.1088/1748-0221/3/08/S08004.

[86] K. Aamodt et al. (ALICE collaboration). *The ALICE experiment at the CERN LHC*. J. Instrum. **3** (2008) S08002.
DOI: 10.1088/1748-0221/3/08/S08002.

[87] A. A. Alves Jr. et al. (LHCb collaboration). *The LHCb detector at the LHC*. J. Instrum. **3** (2008) S08005.
DOI: 10.1088/1748-0221/3/08/S08005.

[88] R. Aaij et al. (LHCb collaboration). *LHCb detector performance*. Int. J. Mod. Phys. A **30** (2015) 1530022.
DOI: 10.1142/S0217751X15300227.

[89] R. Aaij et al. (LHCb collaboration). *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*. J. Instrum. **14** (2019) P04013.
DOI: 10.1088/1748-0221/14/04/P04013.

[90] LHC commissioning. *Longer term LHC schedule*. Accessed: July 5, 2024. 2024.
URL: https://lhc-commissioning.web.cern.ch/schedule/LHC-long-term.htm.

[91] P. Fernandez Declara et al. *A parallel-computing algorithm for high-energy physics particle tracking and decoding using GPU architectures*. IEEE Access **7** (2019) 91612.
DOI: 10.1109/ACCESS.2019.2927261.

[92] P. R. Barbosa-Marinho et al. (LHCb collaboration). *LHCb VELO (VErtex LOcator): Technical design report*. CERN-LHCC-2001-011. Geneva, 2001.
URL: https://cds.cern.ch/record/504321.

[93] R. Aaij et al. *Performance of the LHCb vertex locator*. J. Instrum. **9** (2014) P09007.
DOI: 10.1088/1748-0221/9/09/P09007.

[94] Ch. Ilgner et al. *The Beam Conditions Monitor of the LHCb experiment* (2010).
DOI: 10.48550/arXiv.1001.2487.

[95] P. R. Barbosa-Marinho et al. (LHCb collaboration). *LHCb inner tracker: Technical design report*. CERN-LHCC-2002-029. Geneva, 2002.
URL: https://cds.cern.ch/record/582793.

[96] P. R. Barbosa-Marinho et al. (LHCb collaboration). *LHCb outer tracker: Technical design report*. CERN-LHCC-2001-024. Geneva, 2001.
URL: https://cds.cern.ch/record/519146.

[97] P. d'Argent et al. *Improved performance of the LHCb outer tracker in LHC Run 2*. J. Instrum. **12** (2017) P11016.
DOI: 10.1088/1748-0221/12/11/P11016.

[98] S. Amato et al. (LHCb collaboration). *LHCb magnet: Technical design report*. CERN-LHCC-2000-007. Geneva, 2000.
URL: https://cds.cern.ch/record/424338.

[99] P. R. Barbosa-Marinho et al. (LHCb collaboration). *LHCb muon system: Technical design report*. CERN-LHCC-2001-010. Geneva, 2001.
URL: https://cds.cern.ch/record/504326.

[100] S. Amato et al. (LHCb collaboration). *LHCb RICH: Technical design report*. CERN-LHCC-2000-037. Geneva, 2000.
URL: https://cds.cern.ch/record/494263.

[101] M. Adinolfi et al. *Performance of the LHCb RICH detector at the LHC*. Eur. Phys. J. C **73** (2013) 2431.
DOI: 10.1140/epjc/s10052-013-2431-9.

[102] P. A. Čerenkov. *Visible radiation produced by electrons moving in a medium with velocities exceeding that of light*. Phys. Rev. **52** (1937) 378.
DOI: 10.1103/PhysRev.52.378.

[103] S. Amato et al. (LHCb collaboration). *LHCb calorimeters: Technical design report*. CERN-LHCC-2000-036. Geneva, 2000.
URL: https://cds.cern.ch/record/494264.

[104] S. Borghi. *Novel real-time alignment and calibration of the LHCb detector and its performance*. Nucl. Instrum. Methods Phys. Res. A **845** (2017) 560. In: Proceedings of the Vienna Conference on Instrumentation 2016.
DOI: 10.1016/j.nima.2016.06.050.

[105] M. Clemencic et al. *The LHCb simulation application, Gauss: Design, evolution and experience*. J. Phys. Conf. Ser. **331** (2011) 032023.
DOI: 10.1088/1742-6596/331/3/032023.

[106] T. Sjöstrand, S. Mrenna, and P. Skands. *PYTHIA 6.4 physics and manual*. J. High Energy Phys. **05** (2006) 026.
DOI: 10.1088/1126-6708/2006/05/026.

[107] T. Sjöstrand, S. Mrenna, and P. Skands. *A brief introduction to PYTHIA 8.1*. Comput. Phys. Commun. **178** (2008) 852.
DOI: 10.1016/j.cpc.2008.01.036.

[108] I. Belyaev et al. *Handling of the generation of primary events in Gauss, the LHCb simulation framework*. J. Phys. Conf. Ser. **331** (2011) 032047.
DOI: 10.1088/1742-6596/331/3/032047.

[109] A. Buckley et al. *LHAPDF6: parton density access in the LHC precision era*. Eur. Phys. J. C **75** (2015) 132.
DOI: 10.1140/epjc/s10052-015-3318-8.

[110] B. Andersson et al. *Parton fragmentation and string dynamics*. Physics Reports **97** (1983) 31.
DOI: 10.1016/0370-1573(83)90080-7.

[111] D. J. Lange. *The EvtGen particle decay simulation package*. Nucl. Instrum. Meth. A **462** (2001) 152.
DOI: 10.1016/S0168-9002(01)00089-4.

[112] N. Davidson, T. Przedzinski, and Z. Was. *PHOTOS interface in C++: Technical and physics documentation*. Comp. Phys. Comm. **199** (2016) 86.
DOI: 10.1016/j.cpc.2015.09.013.

[113] S. Agostinelli et al. (GEANT4 collaboration). *Geant4: A simulation toolkit*. Nucl. Instrum. Meth. A **506** (2003) 250.
DOI: 10.1016/S0168-9002(03)01368-8.

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

[115] LHCb collaboration. *The Boole project* (2024).
URL: http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[116] R. L. Workman et al. (Particle Data Group). *Review of particle physics*. Prog. Theor. Exp. Phys. **2022** (2022) 083C01 and 2023 update.
DOI: 10.1093/ptep/ptac097.

[117] G. Meier. *Precision measurements of* $\sin(2\beta)$ *with the LHCb experiment*. Presented 06 Dec 2023. TU Dortmund University, 2023.
DOI: 10.17877/DE290R-24244.

[118] G. A. Cowan, D. C. Craik, and M. D. Needham. *RapidSim: An application for the fast simulation of heavy-quark hadron decays*. Comput. Phys. Commun. **214** (2017) 239.
DOI: 10.1016/j.cpc.2017.01.029.

[119] W. D. Hulsbergen. *Decay chain fitting with a Kalman filter*. Nucl. Instrum. Meth. A **552** (2005) 566.
DOI: 10.1016/j.nima.2005.06.078.

[120] T. Chen and C. Guestrin. *XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. New York, NY, USA: ACM, 2016 785.
DOI: 10.1145/2939672.2939785.

[121] S. Murthy. *Automatic construction of decision trees from data: A multi-disciplinary survey*. Data Min. Knowl. Disc. **2** (1998) 345.
DOI: 10.1023/A:1009744630224.

[122] A. Blum, A. Kalai, and J. Langford. *Beating the hold-out: Bounds for k-fold and progressive cross-validation. Proceedings of the twelfth annual conference on computational learning theory*. COLT '99. New York, NY, USA: ACM, 1999 203.
DOI: 10.1145/307400.307439.

[123] M. Pivk and F. R. Le Diberder. *sPlot: A statistical tool to unfold data distributions*. Nucl. Instrum. Meth. A **555** (2005) 356.
DOI: 10.1016/j.nima.2005.08.106.

[124] H. Dembinski et al. *Custom Orthogonal Weight functions (COWs) for event classification*. Nucl. Instrum. Methods Phys. Res. A **1040** (2022) 167270.
DOI: 10.1016/j.nima.2022.167270.

[125] E. Prilepin and S. Husheer. *CSAPS: Cubic spline approximation*. 2022.
URL: https://github.com/espdev/csaps.

[126] C. de Boor. *A practical guide to splines*. **27**. Springer New York, NY, 2001.
ISBN: 9780387953663.

[127] D. Martínez Santos and F. Dupertuis. *Mass distributions marginalized over per-event errors*. Nucl. Instrum. Meth. A **764** (2014) 150.
DOI: 10.1016/j.nima.2014.06.081.

[128] A. Tumasyan et al. (CMS collaboration). *Observation of $B^0 \to \psi(2S)K^0_S\pi^+\pi^-$ and $B^0_s \to \psi(2S)K^0_S$ decays*. Eur. Phys. J. C **82** (2022) 499.
DOI: 10.1140/epjc/s10052-022-10315-y.

[129] J. Wishahi. *Flavour tagging conference plots: Schematic overview of the underlying principles of LHCb's flavour tagging algorithms. (Internal webpage)*. (modified source material from https://gitlab.cern.ch/lhcb-ft/TikZPics). 2016.
URL: https://twiki.cern.ch/twiki/bin/view/LHCb/FlavourTaggingConferencePlots.

[130] R. Aaij et al. (LHCb collaboration). *Opposite-side flavour tagging of B mesons at the LHCb experiment*. Eur. Phys. J. C **72** (2012) 2022.
DOI: 10.1140/epjc/s10052-012-2022-1.

[131] D. Fazzini. *Flavour Tagging in the LHCb experiment*. Proceedings of Science **LHCP2018** (2018) 230.
DOI: 10.22323/1.321.0230.

[132] R. Aaij et al. (LHCb collaboration). *Optimization and calibration of the LHCb flavour tagging performance using 2010 data*. LHCb-CONF-2011-003 (2011).
URL: https://cds.cern.ch/record/1328957.

[133] R. Aaij et al. (LHCb collaboration). *B flavour tagging using charm decays at the LHCb experiment*. J. Instrum. **10** (2015) P10005.
DOI: 10.1088/1748-0221/10/10/P10005.

[134] K. Heinicke. *Optimization of flavour tagging algorithms for the LHCb experiment*. Presented 30 Sep 2016. TU Dortmund University, 2016.
URL: https://cds.cern.ch/record/2229990.

[135] E. Norrbin and T. Sjöstrand. *Production and hadronization of heavy quarks*. Eur. Phys. J. C **17** (2000) 137.
DOI: 10.1007/s100520000460.

[136] T. Skwarnicki. *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*. Institute of Nuclear Physics, Krakow. DESY-F31-86-02 (1986).
URL: http://inspirehep.net/record/230779/.

[137] A. Rogozhnikov. *Reweighting with boosted decision trees*. J. Phys. Conf. Ser. **762** (2016) 012036.
DOI: 10.1088/1742-6596/762/1/012036.

[138] J. Wimberley. *EspressoPerformanceMonitor*. 2022.
URL: https://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor.

[139] G. Van Rossum and Fred L. Drake. *Python 3 reference manual*. Scotts Valley, CA: CreateSpace, 2009.
ISBN: 1441412697.

[140] Q. Führing and V. Jevtić. *lhbc-ftcalib: A software package for the calibration of flavour-tagged LHCb data* (2024).
DOI: 10.5281/zenodo.12156328.

[141] C. R. Harris et al. *Array programming with NumPy*. Nature **585** (2020) 357.
DOI: 10.1038/s41586-020-2649-2.

[142] The pandas development team. *pandas-dev/pandas: Pandas*. 2020.
DOI: 10.5281/zenodo.3509134.

[143] P. Virtanen et al. *SciPy 1.0: Fundamental algorithms for scientific computing in Python*. Nature Methods **17** (2020) 261.
DOI: 10.1038/s41592-019-0686-2.

[144] H. Dembinski, P. Ongmongkolkul, et al. *scikit-hep/iminuit* (2020).
DOI: 10.5281/zenodo.3949207.

[145] 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.
DOI: 10.1016/0010-4655(75)90039-9.

[146] J. Wimberley. *Semitauonic B_c^+ decays and quark flavor identification methods*. Presented 13 Dec 2017. University of Maryland, 2017.
URL: https://cds.cern.ch/record/2313446.

[147] P. Mackowiak. *Measurements of beauty hadron to charmonium decays – Branching fraction of $\Lambda_b^0 \to \psi(2S)\Lambda$ and CP violation in $B^0 \to \psi K_S^0$*. Presented 22 Mar 2021. TU Dortmund University, 2023.
DOI: 10.17877/DE290R-22118.

[148] I. J. Schoenberg. *Cardinal spline interpolation*. Society for Industrial and Applied Mathematics, 1973.
DOI: 10.1137/1.9781611970555.

[149] R. A. Briere and B. Winstein. *Determining the phase of a strong scattering amplitude from its momentum dependence to better than 1°: The example of kaon regeneration*. Phys. Rev. Lett. **75** (1995) 402.
DOI: 10.1103/PhysRevLett.75.402.

[150] A. Gsponer et al. *Precise coherent K_S^0 regeneration amplitudes for C, Al, Cu, Sn, and Pb nuclei from 20 to 140 GeV/c and their interpretation*. Phys. Rev. Lett. **42** (1979) 13.
DOI: 10.1103/PhysRevLett.42.13.

[151] J. Roehrig et al. *Coherent regeneration of K_S's by carbon as a test of Regge-Pole-Exchange theory*. Phys. Rev. Lett. **38** (1977) 1116. Erratum given in Ref. [177].
DOI: 10.1103/PhysRevLett.38.1116.

[152] R. Brun et al. *root-project/root: v6.18/02*. Version v6-18-02. 2020.
DOI: 10.5281/zenodo.3895860.

[153] W. Verkerke and D. P. Kirkby. *The RooFit toolkit for data modeling*. eConf **C0303241** (2003) MOLT007.
DOI: 10.48550/arXiv.physics/0306116.

[154] P. C. Mahalanobis. *On the generalized distance in statistics*. Sankhya A **80** (2018) 1. Reprint of 1936 publication *Proc. Natl. Acad. Sci. India A 2 (1936) 49*.
DOI: 10.1007/s13171-019-00164-5.

[155] F. Mölder et al. *Sustainable data analysis with Snakemake*. F1000Research **10** (2021) 33.
DOI: 10.12688/f1000research.29032.1.

[156] W. Hulsbergen (Nikhef). personal communication.

[157] T. Gershon (University of Warwick). personal communication.

[158] M. Kenzie et al. *GammaCombo framework for combinations of measurements and computation of confidence intervals*. 2020.
DOI: 10.5281/zenodo.3371421.

[159] LHCb collaboration. *Measurement of CP violation in $B^0 \to J/\psi K^0_S$*. Internal analysis documentation of Ref. [13]. Version 2.2 (unpublished). 2014. LHCb-ANA-2013-046.

[160] LHCb collaboration. *Measurement of CP violation in $B^0 \to J/\psi(\to e^+e^-)K^0_S$ and $B^0 \to \psi(2S)(\to \mu^+\mu^-)K^0_S$ decays*. Internal analysis documentation of Ref. [14]. Version 3.1 (unpublished). 2017. LHCb-ANA-2014-085.

[161] J. Charles et al. (CKMfitter Group). *Updated results on the CKM Matrix – Including results presented up to Summer 23 – Preliminary*. http://ckmfitter.in2p3.fr/www/results/plots_summer23/num/ckmEval_results_summer23.pdf. Accessed: 2024-03-21, Related reference: [52].

[162] M. Bona et al. (UTfit collaboration). *Overview and theoretical prospects for CKM matrix and CP violation from the UTfit collaboration*. Proceedings of Science **WIFAI2023** (2024) 007.
DOI: 10.22323/1.457.0007.

[163] M. Bona et al. (UTfit collaboration). *The unitarity triangle fit in the Standard Model and hadronic parameters from lattice QCD: A reappraisal after the measurements of Δm_s and $BR(B \to \tau\nu_\tau)$*. J. High Energy Phys. **10** (2006) 081.
DOI: 10.1088/1126-6708/2006/10/081.

[164] M. Bona et al. (UTfit collaboration). *New UTfit analysis of the unitarity triangle in the Cabibbo-Kobayashi-Maskawa scheme*. Rend. Lincei Sci. Fis. Nat. **34** (2023) 37.
DOI: 10.1007/s12210-023-01137-5.

[165] Y. Amhis et al. (Heavy Flavor Averaging group [62]). *Results on time-dependent CP violation and measurements of the angles of the unitarity triangle: Winter conferences (Moriond, etc)* (2024).
URL: https://hflav-eos.web.cern.ch/hflav-eos/triangle/moriond2024/.

[166] J. Charles et al. (CKMfitter Group). *CKMlive Web Project*. http://ckmlive.in2p3.fr/. Accessed: 2024-03-21, Related reference: [52].

[167] R. Aaij et al. (LHCb collaboration). *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era* (2016).
DOI: 10.17181/CERN.QZRZ.R4S6.

[168] A. Gaz. *Physics prospects at SuperKEKB/Belle II*. Proceedings of Science **KMI2017** (2017) 005.
DOI: 10.22323/1.294.0005.

[169] Roel Aaij et al. (LHCb collaboration). *A precise measurement of the B^0 meson oscillation frequency*. Eur. Phys. J. C **76** (2016) 412.
DOI: 10.1140/epjc/s10052-016-4250-2.

[170] J. Charles et al. *Future sensitivity to new physics in B_d, B_s, and K mixings*. Phys. Rev. D **89** (2014) 033016.
DOI: 10.1103/PhysRevD.89.033016.

[171] J. Charles et al. *New physics in B meson mixing: Future sensitivity and limitations*. Phys. Rev. D **102** (2020) 056023.
DOI: 10.1103/PhysRevD.102.056023.

[172] I. Adachi et al. (Belle II collaboration). *Measurement of decay-time-dependent CP violation in $B^0 \to J/\psi K^0_S$ decays using 2019-2021 Belle II data* (2023).
DOI: 10.48550/arXiv.2302.12898.

[173] I. Adachi et al. (Belle II collaboration). *New graph-neural-network flavor tagger for Belle II and measurement of sin $2\phi_1$ in $B^0 \to J/\psi K^0_S$ decays*. Phys. Rev. D **110** (2024) 012001.
DOI: 10.1103/PhysRevD.110.012001.

[174] T. Likhomanenko, D. Derkach, and A. Rogozhnikov. *Inclusive flavour tagging algorithm*. J. Phys. Conf. Ser. **762** (2016) 012045.
DOI: 10.1088/1742-6596/762/1/012045.

[175] M. Zaheer et al. *Deep Sets*. CoRR **abs/1703.06114** (2017).
DOI: 10.48550/arXiv.1703.06114.

[176] R. Aaij et al. (LHCb collaboration). *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B^0_s decay branching fractions*. Phys. Rev. D **104** (2021) 032005.
DOI: 10.1103/PhysRevD.104.032005.

[177] J. Roehrig et al. *Erratum: Coherent regeneration of K_S's by carbon as a test of Regge-Pole-Exchange theory*. Phys. Rev. Lett. **39** (1977) 674.
DOI: 10.1103/PhysRevLett.39.674.

[178] O. Klein. *On the theory of charged fields* (1939). In: New theories in physics: Conference organized in collaboration with the international union of physics and the Polish intellectual co-operation committee, Warsaw, May 30th - June 3rd, 1938. LCCN: 40007679 77; O. Klein. *Mesons and nucleons*. Nature **161** (1948) 897.
DOI: 10.1038/161897a0.

Danksagung

Mein besonderer Dank gilt Herrn Prof. Dr. Bernhard Spaan, der es mir ermöglicht hat, mein Masterstudium und mein Promotionsstudium an seinem Lehrstuhl anzutreten. Bernhard Spaan vertraute mir diese so wichtige und interessante Messung an und unterstützte mich über einen großen Zeitraum meines Studiums. Bernhard Spaan ist im Dezember 2021 viel zu früh verstorben. Die unerschöpfliche Fachkenntnis, die Leidenschaft und der Humor, mit denen Bernhard Spaan mit uns Studierenden oft auch weit über die Physik hinaus diskutiert hat, werden bis heute oft mit Bewunderung in Erinnerung gerufen.

Ich bin Prof. Dr. Johannes Albrecht zu großem Dank verpflichtet, da er es mir ermöglicht hat, mein Promotionsstudium an seinem Lehrstuhl fortzusetzen und zu einem erfolgreichen Abschluss zu bringen. In der Endphase der Arbeit war Johannes' Rat besonders wertvoll und trug wesentlich zum Erfolg dieser Messung und des gesamten Projekts bei. Außerdem danke ich Herrn Prof. Dr. Kevin Kröninger, der meine Arbeit als Zweitgutachter gelesen hat, und ich danke der gesamten Prüfungskommission.

Ich danke dem BMBF für die langjährige Finanzierung meiner Forschung.

Ich danke meinen Kommilitonen Dr. Patrick Mackowiak und Dr. Gerwin Meier, ohne deren intensive Arbeit die Durchführung dieser Messung kaum vorstellbar gewesen wäre. Es war eine große Freude, mit Patrick und mit Gerwin zu arbeiten. Obwohl der Arbeitsaufwand angesichts des Projektumfangs für ein so kleines Forscherteam enorm war, konnten wir uns immer voll und ganz aufeinander verlassen.

In this context, I would particularly like to thank Dr. Sophie Hollitt, who was on hand to advise us during the final phase of the measurement and did much to ensure that communication between the collaboration and ourselves was maintained at all times, despite the amount of work involved, so that no question remained unanswered in the end.

I thank the LHCb collaboration for their many years of support. Many individual members of the collaboration have given us important advice over the period of this project's development, and I thank every one of them. A lot of work is done by members of the collaboration tasked with reviewing measurements, and I am glad to have profited from their inexhaustible expertise.

Ich danke meinen Eltern und meiner Familie, die mich seit jeher und in jeder erforderlichen Weise unterstützen und mir so meine akademische Laufbahn erst ermöglicht haben.

Meinen Bürokollegen, Maik Becker, Gerwin Meier, Patrick Mackowiak, Quentin Führing, Louis Gerken und Jan Langer danke ich für eine tolle gemeinsame Zeit und für die Unterstützung, die ich stets durch sie erfahren habe.

Zuletzt möchte ich sagen, dass ich mich glücklich schätze Teil dieses Lehrstuhls zu sein und die Gelegenheit habe mit so vielen motivierten und hilfsbereiten Wissenschaftler*innen zusammenzuarbeiten.