Flavour-tagged Measurement of *CP* Observables in $B_s^0 \to D_s^{\mp} K^{\pm}$ Decays with the LHCb Experiment

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

vorgelegt von
Ulrich Paul Eitschberger
geboren in Dorsten

Fakultät Physik
Technische Universität Dortmund

Dortmund, im März 2018

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

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

Datum des Einreichens der Arbeit: 27.03.2018
Datum der mündlichen Prüfung: 28.05.2018

Abstract

In this thesis, the LHCb measurement of *CP* violation in $B_s^0 \to D_s^\mp K^\pm$ decays is discussed. In $B_s^0 \to D_s^\mp K^\pm$ decays, *CP* violation arises in the interference between decay and decay after B_s^0–$\bar{B}_s^0$ mixing. A decay-time-dependent measurement of the involved *CP* parameters is enabled by the flavour tagging, which infers the production flavour of the B_s^0 mesons.

The analysed data has been recorded by the LHCb experiment in proton-proton collisions at centre-of-mass energies of 7 and 8 TeV. The *CP* parameters in $B_s^0 \to D_s^\mp K^\pm$ are determined to be

$$
\begin{aligned}
C &= 0.73 \pm 0.14\,(\text{stat}) \pm 0.05\,(\text{syst})\,, \\
A_f^{\Delta\Gamma} &= 0.39 \pm 0.28\,(\text{stat}) \pm 0.15\,(\text{syst})\,, \\
A_{\bar{f}}^{\Delta\Gamma} &= 0.31 \pm 0.28\,(\text{stat}) \pm 0.15\,(\text{syst})\,, \\
S_f &= -0.52 \pm 0.20\,(\text{stat}) \pm 0.07\,(\text{syst})\,, \\
S_{\bar{f}} &= -0.49 \pm 0.20\,(\text{stat}) \pm 0.07\,(\text{syst})\,.
\end{aligned}
$$

Due to the interference between $b \to c\bar{s}u$ and $b \to u\bar{c}s$ transitions in $B_s^0 \to D_s^\mp K^\pm$ decays, the *CP* parameters are sensitive to the CKM angle γ. It is measured to be $\gamma = \left(128^{+17}_{-22}\right)^\circ$, representing the most precise γ measurement using B_s^0 decays to date.

Kurzfassung

In dieser Dissertation wird die Messung von *CP*-Verletzung in $B_s^0 \to D_s^\mp K^\pm$-Zerfällen mit Daten des LHCb-Experiments diskutiert. In Zerfällen von $B_s^0 \to D_s^\mp K^\pm$ tritt *CP*-Verletzung in der Interferenz zwischen Zerfall und Zerfall nach B_s^0–$\bar{B}_s^0$ Mischung auf. Eine zerfallszeitabhängige Messung der beteiligten *CP*-Parameter wird durch das Flavour Tagging ermöglicht, welches die Produktionszustände der B_s^0-Mesonen bestimmt.

Die analysierten Daten wurden vom LHCb-Experiment in Proton-Proton-Kollisionen bei Schwerpunktsenergien von 7 und 8 TeV aufgezeichnet und entsprechen einer integrierten Luminosität von 3 fb^{-1}. Die *CP*-Parameter im Zerfallskanal $B_s^0 \to D_s^\mp K^\pm$ werden gemessen zu

$$
\begin{aligned}
C &= 0{,}73 \pm 0{,}14\,(\text{stat}) \pm 0{,}05\,(\text{syst})\,, \\
A_f^{\Delta\Gamma} &= 0{,}39 \pm 0{,}28\,(\text{stat}) \pm 0{,}15\,(\text{syst})\,, \\
A_{\bar{f}}^{\Delta\Gamma} &= 0{,}31 \pm 0{,}28\,(\text{stat}) \pm 0{,}15\,(\text{syst})\,, \\
S_f &= -0{,}52 \pm 0{,}20\,(\text{stat}) \pm 0{,}07\,(\text{syst})\,, \\
S_{\bar{f}} &= -0{,}49 \pm 0{,}20\,(\text{stat}) \pm 0{,}07\,(\text{syst})\,.
\end{aligned}
$$

Die *CP*-Parameter sind sensitiv auf den CKM-Winkel γ, da im Zerfallskanal $B_s^0 \to D_s^\mp K^\pm$ Übergänge von $b \to c\bar{s}u$ und $b \to u\bar{c}s$ interferieren. Der CKM-Winkel wird bestimmt zu $\gamma = \left(128^{+17}_{-22}\right)^\circ$. Es handelt sich um die bisher präziseste γ-Messung mit B_s^0-Zerfällen.

Contents

1 Introduction

Particle physics aims to understand the nature of the smallest constituents of matter and the interactions among them. The most successful theory in this field is the Standard Model of particle physics (SM). Its current formulation was finalised during the mid-1970s after the existence of quarks had been experimentally confirmed [1,2]. All particles that were postulated by the SM, like the top quark or the tau neutrino, have been discovered by various experiments in the last decades. The last particle to be experimentally established was the Higgs boson [3–5]. While experiments have proven the reliability of the SM predictions up to high precision, there are still many motivations to keep on searching for physics beyond the SM. One of the most striking shortcomings of the SM is its inability to explain the matter-antimatter asymmetry that is observed in today's universe. According to big bang theories, matter and antimatter have been produced in equal amounts at the beginning of the universe. However, today we are surrounded by clusters of matter, while the antimatter seems to have almost completely vanished. This leads to the concept of baryogenesis, which implies that the matter excess was created after the big bang. In 1967, Andrei Sakharov proposed three conditions that are prerequisites for any baryogenesis [6]. Firstly, a process must exist that violates the conservation of the baryon number. Secondly, there need to be interactions outside of the thermal equilibrium. Lastly, both the C and the CP symmetries need to be violated, which expressed in simple terms states that matter and antimatter behave differently. Currently, there is no evidence for violation of the baryon number conservation as the lifetime of the proton is measured to be above 1.6×10^{34} years [7]. There are several scenarios for the emergence of thermal non-equilibrium in the immediate aftermath of the big bang [8]. Furthermore, processes that violate the C and CP symmetries were already found in 1964 [9]. The SM does include CP violation, but the allowed amount is magnitudes too small to explain the size of the observed matter-antimatter asymmetry [10]. This is a clear hint for the existence of physics beyond the SM and makes studying CP violation even fifty years after its discovery a promising field of research.

The only source of CP violation in the SM is a single complex phase in the Cabbibo-Kobayashi-Maskawa (CKM) quark-mixing matrix [11]. A fundamental requirement of the SM is that the CKM matrix is unitary. Confirming or falsifying the unitarity represents a strong test of the SM itself. This test is illustrated by means of the so-called CKM triangle, which represents the unitarity of the CKM matrix as a triangle in the complex plane. The sides and angles of the CKM triangle are measured independently in order to probe whether it closes at its apex. An important building block is γ, which is the least well-known of the CKM angles. The CP violation in the interference between decay and decay after mixing that occurs in tree-level decays

of $B^0_{(s)} \to D^\mp_{(s)} h^\pm$ ($h = \pi, K$) gives access to γ [12–14]. A consistency check of the SM picture of *CP* violation is given by comparing the values for γ obtained from tree-level processes with γ measurements that involve loop processes and the indirect determination of γ from the other unitarity parameters. The decay-time-dependent analysis of *CP* violation in $B^0_s \to D^\mp_s K^\pm$ decays [15] is the main topic of this thesis. Special focus is put on the flavour tagging, which is a technique that provides vital knowledge about the initial flavour of the B^0_s mesons. The flavour-tagging studies documented in this thesis exceed the scope of the $B^0_s \to D^\mp_s K^\pm$ analysis and include calibrations that were contributions by the author to other publications.

The analysed data was collected by the LHCb experiment, which is located at the world's most powerful particle accelerator, the Large Hadron Collider (LHC). The LHCb detector is dedicated to precision measurements in the sector of *b* and *c* hadrons with a focus on rare decays and *CP* violation. One of the main challenges in the analysis of $B^0_s \to D^\mp_s K^\pm$ decays is the hadronic environment caused by the *pp* collisions of the LHC. Hence, a signal selection needs to be developed that provides a strong separation between signal and background to allow for a precise measurement.

The presented analysis is an updated measurement of Ref. [16], which was performed on the 2011 data. Apart from exploiting the full Run 1 dataset, several improvements are implemented in regard to the analysis strategy.

The thesis is structured as follows. Chapter 2 gives a short introduction to the SM, while the subsequent Chap. 3 explains how *CP* violation is established in the SM. Afterwards, the LHCb experiment is discussed in Chap. 4, which includes details on the detector and on the software framework used to analyse the data. In Chap. 5, a few common analysis techniques are introduced, which are used throughout the thesis. Chapter 6 introduces the concept of the flavour tagging and documents flavour-tagging calibrations with $B^0 \to J/\psi K^{*0}$ decays. The preparation of the data, which involves the training of a boosted decision tree, is outlined in Chap. 7. In Chap. 8, the multidimensional fit is explained, which aims to provide weights that enable to perform the decay-time fit in a background subtracted way. The strategy and the results of the decay-time fit are the topic of Chap. 9, which includes the determination of the CKM angle γ from the *CP* parameters. Apart from a discussion of the results, Chap. 10 gives an outlook on future measurements in the same field. Finally, the thesis is concluded in Chap. 11.

2 The Standard Model of Particle Physics

The Standard Model of particle physics has been the established theoretical foundation of particle physics for more than forty years. It is a gauge invariant quantum field theory that describes the fundamental particles and the interactions between them (see Sec. 2.1). In its core the SM is based on the conservation and the breaking of continuous and discrete symmetries, as discussed in Sec. 2.2. Although the SM has been and still is a huge success, especially in terms of providing precise predictions, it is not able to explain all observed phenomena. A short summary of the Standard Model's shortcomings that demand extensions is given in Sec. 2.3. The information in this chapter is based on Refs. [17, 18].

2.1 Fundamental particles and interactions

The Standard Model describes twelve elementary particles with spin 1/2, so-called fermions, which are further grouped into six quarks and six leptons. Additionally, there is one antifermion for each fermion, which carries opposite charge-related quantum numbers. Both quarks and leptons are subdivided into three generations, ordered according to their increasing masses.

Each of the three quark generations, also referred to as families, consists of one up-type and one down-type quark. The electrical charge of up-type quarks is $+2/3e$, where e is the elementary charge, while down-type quarks carry $-1/3e$. The first quark family contains the quarks out of which protons and neutrons are composed, being the up quark (u) and the down quark (d). In the second quark family, the charm (c) and strange (s) quarks are included, while the third consists of the top (t) and the bottom (b) quarks. Quarks are the only fermions with a colour charge, which can be r, g, or b, or one of the three corresponding anticolours in case of an antiquark. Due to confinement [19] quarks appear only in bound states, so-called hadrons, of which different types exist. Firstly, there are baryons and antibaryons, built out of three quarks or three antiquarks, respectively. Secondly, bound states containing one quark and one antiquark are known as mesons. Finally, there were recent discoveries of hadrons consisting of four and five quark bound states, called tetra- and penta-quarks [20–23]. In all types of hadrons, the aggregated quark colours render the hadron colourless, due to the mixture of all three colours or equal numbers of colours and anticolours.

The lepton generations each contain one particle with an electric charge of $-1e$ and

one corresponding neutrino, an electrically neutral particle. Specifically, the first family consists of the electron (e) and the electron neutrino (ν_e), the second is comprised of the muon (μ) and the muon neutrino (ν_μ), and lastly, the third family includes the tau (τ) and the tau neutrino (ν_τ). All neutrinos are set to be massless in the SM.

Three of the four fundamental interactions are described within the SM, being the electromagnetic, the weak, and the strong force, whereas gravity is omitted. All embedded interactions are mediated by integer spin particles, called gauge bosons, which couple to force specific particle charges. The massless photon (γ) is the mediator of the electromagnetic force and couples to the electrical charge. Three massive bosons, the W^+, W^-, and Z^0 bosons, mediate the weak force whose corresponding charge is the weak isospin. Finally, eight massless gluons are the force carriers of the strong force, that couple to the colour charge. Only quarks are able to interact strongly, as they are the only fermions carrying a colour charge.

The scalar Higgs boson was the last particle of the Standard Model to be experimentally discovered [3–5]. It is an excitation of the Higgs field that generates particle masses. Figure 2.1 summarises the Standard Model's elementary particles. It includes properties of the particles and illustrates by which forces they are able to interact.
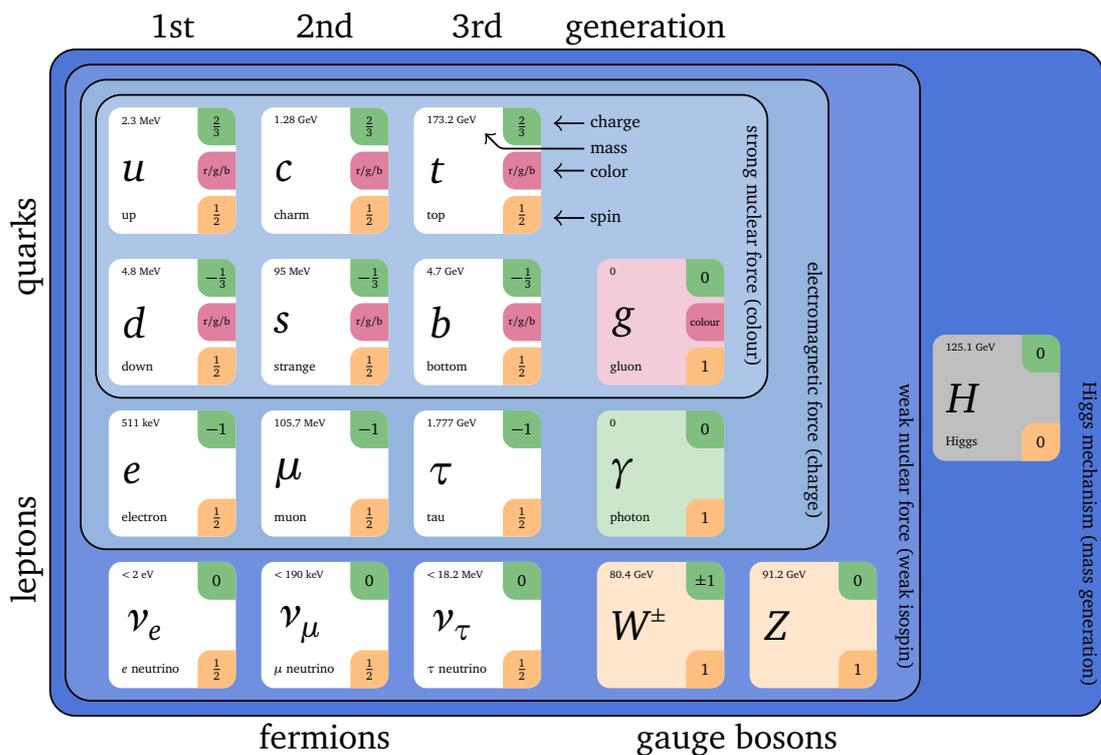

Figure 2.1: Overview of the fundamental fermions, gauge bosons, and forces in the Standard Model. Values for electric charge, mass, and spin are taken from Ref. [24].

2.2 Symmetries and symmetry violations

Symmetry considerations are fundamental in the Standard Model. A symmetry in terms of a quantum field theory is equivalent to a transformation of the fields that does not change the equations of motion and thus leaves the Lagrangian invariant. The Poincaré symmetry group is common to all relativistic quantum field theories. It includes invariance under translations, rotations, and boosts in the Minkowski spacetime. According to the fundamental principle of Noether's theorem [25], every symmetry of a physical system leads to a conserved quantity. For example, the conservation of energy is connected to the invariance of translations in spacetime. It can only be violated on very short timescales due to Heisenberg's uncertainty principle [26], which allows heavy particles to appear virtually in decay processes. Further examples are the conservations of momentum and angular momentum, which are consequences of invariance under translations and rotations in space, respectively.

For the definition of interactions in the SM, invariance under local gauge transitions is fundamental. The theory of the electromagnetic interaction, called Quantum Electrodynamics (QED), is a generalised quantum field theory of Maxwell's classical approach. It is established by requiring a local $U(1)$ gauge symmetry in the Lagrangian of a free fermion field Ψ with a corresponding mass m

$$\mathscr{L}_0 = \overline{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi\,. \tag{2.1}$$

This requirement is implemented by exchanging ∂_μ with the covariant derivative D_μ:

$$\partial_\mu \to D_\mu = \partial_\mu + i\,e\,A_\mu\,, \tag{2.2}$$

where the vector field A_μ can be interpreted as the photon field that mediates the interactions of electrically charged particles. The kinetic term

$$\mathscr{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.3}$$

is added to introduce the dynamics, where the electromagnetic field strength tensor $F_{\mu\nu}$ contains Maxwell's classical equations. This concludes the QED Lagrangian:

$$\mathscr{L}_{\mathrm{QED}} = \overline{\Psi}(i\gamma^\mu D_\mu - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\,. \tag{2.4}$$

One of the biggest achievements of the SM is the unification of the QED and the weak interaction via the Glashow-Salam-Weinberg (GSW) model [27, 28]. The combined electroweak interaction is represented by invariance under an $SU(2) \times U(1)$ gauge group, whose symmetry is spontaneously broken by the non-zero vacuum expectation value of the scalar Higgs field. In this way, the Higgs mechanism enables the photon to be massless, while the mediators of the weak interaction, the W^+, W^-, and Z^0 bosons, acquire masses. Further consequences of the Higgs mechanism are the different coupling strengths of the electromagnetic and the weak interactions at

low momentum scales as well as the existence of the Higgs boson. All fundamental particles are able to interact weakly, but the interaction is sensitive to the particle's helicity in the way that only left-handed particles and right-handed antiparticles are affected by it. The helicity of a particle is defined as the orientation of its spin in relation to its momentum vector. Neutrinos only exist left-handed and they are not attributed a mass through the Higgs mechanism. Finally, the quantum field theory of the strong interaction is called Quantum Chromodynamics (QCD) and it is invariant under the $SU(3)$ gauge group. The Lagrangian of the QCD

$$\mathscr{L}_{\text{QCD}} = \overline{\Psi}_i (i(\gamma^\mu D_\mu)_{ij} - m\delta_{ij})\Psi_j - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a. \tag{2.5}$$

contains the gluon field strength tensor $G^a_{\mu\nu}$ that characterises gluon interactions between quarks and between themselves. Two of the most peculiar properties of the QCD are the already mentioned confinement and the asymptotic freedom. The former causes the production of quark-antiquark pairs in between quarks that move too far away from each other. The latter is relevant at very short distances, where gluon polarisation outperforms quark polarisation effects, allowing quarks to move as free particles. Gluon self polarisation itself is enabled by the fact that gluons carry colour charge themselves. In conclusion, the Lagrangian of the Standard Model can be grouped as follows:

$$\mathscr{L} = \mathscr{L}_{SU(3)} + \mathscr{L}_{SU(2)\otimes U(1)}. \tag{2.6}$$

Besides continuous symmetries, like those of the global Poincaré group and the local gauge symmetries, several discrete symmetry operations are being examined with respect to their possible influences on physics processes. Especially the following three discrete transformations are of interest in the course of this thesis:

- the parity operation P, which reflects all spatial coordinates at the origin

$$P\Psi(r) = \Psi(-r), \tag{2.7}$$

- the charge conjugation C, which inverts the signs of the magnetic moment and all charges, and hence is equivalent to the transformation of a particle into its antiparticle

$$C|p\rangle = |\bar{p}\rangle, \tag{2.8}$$

- the time reversal T, which inverts the sign of the time coordinate and by that reverses the direction of motion as the signs of e.g. velocity, momentum, and angular momentum are flipped

$$T\Psi(t) = \Psi(-t). \tag{2.9}$$

When carried out individually, none of these transitions is a true symmetry in the SM. The parity operation, which can be understood as changing a left-handed fermion into a right-handed fermion, is maximally violated by the weak interaction, which differs between left-handed and right-handed fermion fields. Consequently, the weak interaction also violates the charge conjugation symmetry, due to the differentiation between particles and antiparticles. A symmetry under the combined *CP* operation could still be preserved, as it transforms the particle into the corresponding antiparticle state and its helicity at the same time. However, it is also only an approximate symmetry in the SM due to the established *CP* violation by the weak interaction, which is discussed in detail in Chap. 3. A truly conserved discrete symmetry in the SM is the combined transformation of C, P and T. This, in accordance with the *CPT* theorem [29, 30], can be summed up to particles and antiparticles having identical lifetimes and masses.

2.3 Limitations of the Standard Model

The Standard Model has proven to be a reliable theory, especially in the interplay between theoretical expectations and experimental verificatons. A well-known example of this is the Higgs boson, which was predicted in 1964 [31–33], about 50 years before it was discovered experimentally [3, 4]. Another example is the top quark, whose prediction [11] also included an accurate mass estimation, experimentally confirmed at Fermilab in 1995 [34, 35]. Furthermore, even though many experiments are searching for deviations from SM expectations, clear evidence of "New Physics" has yet to be found. It is nevertheless certain that physics beyond the SM exists, or that the current SM is only a part of a larger universal theory, because it lacks explanations for several observations. One example is the matter-antimatter asymmetry observed in today's universe that was already covered in the introduction. In the following, more examples of the SM's shortcomings are discussed.

The experimentally established neutrino oscillations [36–38], are an implicit discovery of mass differences between neutrino generations. Hence, the SM assumption of massless neutrinos is proven to be wrong. A minimal extension to the SM exists that allows for neutrino masses by assuming neutrinos are their own antiparticles, so-called Majorana particles [39–41].

The Standard Model only describes about 5 % of the energy content of today's universe, being common baryonic matter and leptons [42]. The remaining 95 %, which consist of so-called dark matter and dark energy, are not covered by the SM. Dark matter has a share of 26 % and is an unidentified type of matter that does not emit electromagnetic radiation, but interacts at least gravitationally. Its existence is postulated to explain the rotational speed of stars and gases in galaxies [43, 44]. Dark energy on the other hand is an unknown type of energy made responsible for the accelerating expansion of the universe [45, 46].

Besides delivering explanations for the mentioned observations, the most desirable prospect of an improved SM is providing a unified theoretical representation

of all four known fundamental interactions. This requires a quantum mechanical description of gravitation, which is not available, yet. In addition, there is a hierarchy problem of gravity being 10^{24} times weaker than the weak force. A unification of the forces would presumably emerge only at extremely high energy scales of 10^{25} eV/c^2 [47]. At these energies, the mass of the Higgs boson is strongly enlarged by diverging quantum corrections, which is in conflict to the measured Higgs mass of about 125 GeV/c^2 [5]. The SM solution to the hierarchy problem is fine-tuning tree-level and loop contributions to the same level, which is criticised to be too artificial. A different solution could be the existence of symmetrical partners to SM particles [48, 49], but until today experiments have not found any evidence for the existence of this so-called supersymmetry [50].

3 *CP* Violation

This chapter discusses *CP* violation in the context of the Standard Model and the time-dependent measurement of $B_s^0 \to D_s^\mp K^\pm$ decays. The sole origin of *CP* violation in the SM is a complex phase of the Cabbibo-Kobayashi-Maskawa (CKM) quark mixing matrix, which is introduced in Sec. 3.1. Derived from the unitary CKM matrix are so-called unitarity triangles, which are utilised in a stringent test of the description of *CP* violation in the SM (see Sec. 3.2). Many of the parameters in the CKM sector are accessible in processes involving the oscillation and decay of neutral *B* mesons, as explained in Sec. 3.3. The different types of *CP* violation are introduced in Sec. 3.4. In the following section 3.5 and 3.6, it is illustrated how *CP* violation appears in $B_s^0 \to D_s^\mp K^\pm$ decays and how the CKM angle γ can be determined from a measurement of *CP* violation in this channel. The information in this chapter is loosely following Refs. [24, 51, 52].

3.1 The CKM mechanism

The only source of *CP* violation in the Standard Model is in the weak sector, more precisely a single complex phase of the Cabbibo-Kobayashi-Maskawa (CKM) quark-mixing matrix [11]. This section explains how the matrix is introduced in the context of the quark mass generation, where the Yukawa interaction between right- and left-handed quarks is connected to the vacuum expectation value v of the Higgs field ϕ. The Yukawa Lagrangian

$$\mathscr{L}_{\text{Yukawa}} = \overline{Q}_L \phi Y_d d_R + \overline{Q}_L \phi^c Y_u u_R + h.c. \tag{3.1}$$

contains the Yukawa matrices Y_d and Y_u for the down-type and up-type quark fields Q_L, u_R and d_R. The notation of the quark fields takes into account that left-handed quarks (L) are described by doublets under an $SU(2)$ symmetry

$$Q_L \in \left\{ \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \right\},$$

while right-handed quarks (R) are singlets

$$u_R \in \left\{ u_R, c_R, t_R \right\}, \qquad d_R \in \left\{ d_R, s_R, b_R, \right\}.$$

By using the Higgs field ϕ and its charge conjugated field ϕ^c,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h^0(x) \end{pmatrix} \quad \text{and} \quad \phi^c(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} v + h^0(x) \\ 0 \end{pmatrix},$$

as well as replacing $\frac{v}{\sqrt{2}}Y_{u,d}$ by the mass matrices M_u and M_d, the Yukawa Lagrangian becomes

$$\mathcal{L}_{\text{Yukawa}} = \bar{d}_L M_d d_R + \bar{u}_L M_u u_R + h.c.,\tag{3.2}$$

now containing mass terms similar to those of the QCD and QED Lagrangians. Furthermore, $\bar{u}_L$ and $\bar{d}_L$ represent all left-handed up- and down-type quark fields, respectively. These fields are only gauge eigenstates and not mass eigenstates. In order to obtain the corresponding mass eigenstates q' and to diagonalise the mass matrices M_u and M_d, the unitary transformation $q'_A = V_{A,q} q_A$ is performed, where $A \in \{L, R\}$, $q \in \{u, d\}$ and $V_{A,q} V_{A,q}^\dagger = \mathbb{1}$.

Applying the same transformation in the Lagrangians that describe the neutral currents (NC) and the charged currents (CC) of the electroweak theory shows that for neutral currents, which are mediated by the photon and the Z boson, the gauge and the mass eigenstates are the same.

$$\begin{aligned}\mathcal{L}_{\text{NC}} &= \bar{u}_L \gamma_\mu Z^\mu u_L = \bar{u}_L V_{L,u}^\dagger V_{L,u} \gamma_\mu Z^\mu V_{L,u}^\dagger V_{L,u} u_L \\ &= \bar{u}'_L \gamma_\mu Z^\mu u'_L\end{aligned}$$

The charged currents on the other hand allow for transitions between up- and down-type quarks, therefore, the transformation matrices do not extinguish each other in this case.

$$\mathcal{L}_{\text{CC}} = -\frac{g_2}{\sqrt{2}}(\bar{u}_L \gamma^\mu W_\mu^+ d_L + \bar{d}_L \gamma^\mu W_\mu^- u_L)\tag{3.3}$$

$$= -\frac{g_2}{\sqrt{2}}(\bar{u}_L V_{L,u}^\dagger V_{L,u} \gamma^\mu W_\mu^+ V_{L,d}^\dagger V_{L,d} d_L + \bar{d}_L V_{L,d}^\dagger V_{L,d} \gamma^\mu W_\mu^- V_{L,u}^\dagger V_{L,u} u_L)\tag{3.4}$$

$$= -\frac{g_2}{\sqrt{2}}(\bar{u}'_L \gamma^\mu W_\mu^+ V_{L,u} V_{L,d}^\dagger d'_L + \bar{d}'_L \gamma^\mu W_\mu^- V_{L,d} V_{L,u}^\dagger u'_L)\tag{3.5}$$

The product $V_{L,u} V_{L,d}^\dagger$ is the CKM Matrix V_{CKM}. It connects the weak eigenstates of the down-type fermion fields with the corresponding mass eigenstates.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_L = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L\tag{3.6}$$

Because the Yukawa matrices are diagonalised by different unitary transformations, the CKM matrix is not just a unity matrix. As a consequence, mixing between different generations is enabled. The probability of the transition from an up-type quark $i \in \{u, c, t\}$ into a down-type quark $j \in \{d, s, b\}$ is proportional to the square of the CKM matrix element $|V_{ij}|^2$. A complex $n \times n$ matrix has generally $2n^2$ real parameters in total, which would lead to 18 free parameters for V_{CKM}. However, the CKM matrix is the product of two unitary matrices ($V_{L,u}$ and $V_{L,d}$) and is therefore a unitary matrix itself, which requires

$$V_{\text{CKM}}^\dagger V_{\text{CKM}} = \mathbb{1}.\tag{3.7}$$

By rewriting Eq. 3.7 into relations between the CKM matrix elements,

$$\sum_{k=1}^{3} V_{ki} V_{kj}^{*} = \delta_{ij}, \quad \text{with} \quad i, j \in \{1, 2, 3\}, \tag{3.8}$$

it becomes obvious that unitarity reduces the number of free parameters to nine. More-over, five parameters are physically irrelevant, because they can be absorbed in a quark field. Therefore, only four degrees of freedom remain, out of which three are real-valued rotation angles. Besides the three angles, there is one more parameter, the complex Kobayashi Maskawa phase δ_{KM}. This phase violates the *CP* symmetry of the weak charged currents (Eq. 3.5) and is the sole source of *CP* violation in the Standard Model. Chau and Keung suggested an exact parameterisation of the CKM matrix (see Ref. [53]), given by

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}s_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \tag{3.9}$$

In this parametrisation $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$ (with $i > j$, $i \in \{1, 2\}$, and $j \in \{2, 3\}$) contain the three Euler mixing angles $\theta_{ij} \in [0, \frac{\pi}{2}]$. The hierarchical structure of the CKM matrix elements, whose magnitudes decrease with increasing distance to the main diagonal, becomes apparent in the most common parametrisation of the CKM matrix

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{3.10}$$

This parametrisation, which was introduced by Wolfenstein [54], makes use of a series expansion in the parameter λ with

$$\lambda = s_{12}, \quad A = \frac{s_{23}}{s_{12}^2}, \quad \rho = \frac{s_{13}}{s_{12}s_{23}} \cos \delta, \quad \text{and} \quad \eta = \frac{s_{13}}{s_{12}s_{23}} \sin \delta. \tag{3.11}$$

The Wolfenstein parametrisation provides a representation of the CKM matrix that is unitary in all orders of λ and elucidates the hierarchical order of the matrix elements via the exponent of λ. Currently, the world average of λ is 0.22506 ± 0.00050 [24].

In summary, there are ten parameters in the flavour sector of the SM, which are the six quark masses and the four parameters of the CKM matrix. A precise experimental determination of the latter parameters represents a strong indirect test on the correctness of the SM, as explained in the next section.

3.2 Unitarity triangles

Six of the nine equations defined by Eq. 3.8 require the sum of three complex quantities to vanish and can therefore be geometrically interpreted as triangles in the complex plane. Four of these six unitarity triangles degenerate, due to the large differences between the magnitudes of the matrix elements. Nevertheless, all triangles have the same area $J_{CP}/2$, where the Jarlskog invariant

$$|J_{CP}| = |\mathrm{Im}(V_{ij}V_{kl}V_{il}^*V_{kj}^*)|, \quad \text{with} \quad i \neq k, j \neq l, \tag{3.12}$$

represents a phase-convention independent measure of *CP* violation allowed in the SM [55]. The current world average is $J_{CP} = 3.04^{+0.21}_{-0.20} \times 10^{-5}$ [24]. Of the two unitarity triangles with sides of comparable length,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{3.13}$$

is most commonly used. Prior to plotting the triangle in the complex plane, each side is divided by the best-known value $V_{cd}V_{cb}^*$ and the generalised Wolfenstein parameters

$$\overline{\rho} = \rho(1 - \lambda^2/2 + ...) \quad \text{and} \quad \overline{\eta} = \eta(1 - \lambda^2/2 + ...) \tag{3.14}$$

are introduced. This allows to define the vertex positions as $(0,0)$, $(1,0)$ and $(\overline{\rho}, \overline{\eta})$, where the apex

$$\overline{\rho} + i\overline{\eta} = -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \tag{3.15}$$

as well as the sides

$$R_c = \left|\frac{V_{cd}V_{cb}^*}{V_{cd}V_{cb}^*}\right| = 1, \quad R_t = \left|\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{(1-\overline{\rho})^2 + \overline{\eta}^2}, \quad R_u = \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right| = \sqrt{\overline{\rho}^2 + \overline{\eta}^2} \tag{3.16}$$

are phase-convention independent. The same holds for the angles of the triangle

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{3.17}$$

Angles and sides of the triangle are being measured separately in a wide range of processes, some of which incorporate *CP* violation. All different measurements are subsequently combined to achieve an over-constraint of the triangle, which enables a proficient inclusive test of the SM. If the measured triangle does not close at its apex, unitarity of the quark mixing matrix is not maintained in nature. Figure 3.1 shows the triangle from Eq. 3.13 in the complex plane including the current experimental constraints calculated by the CKMfitter group [56]. At the moment, the measurements are in good agreement with each other and with a unitary CKM matrix.
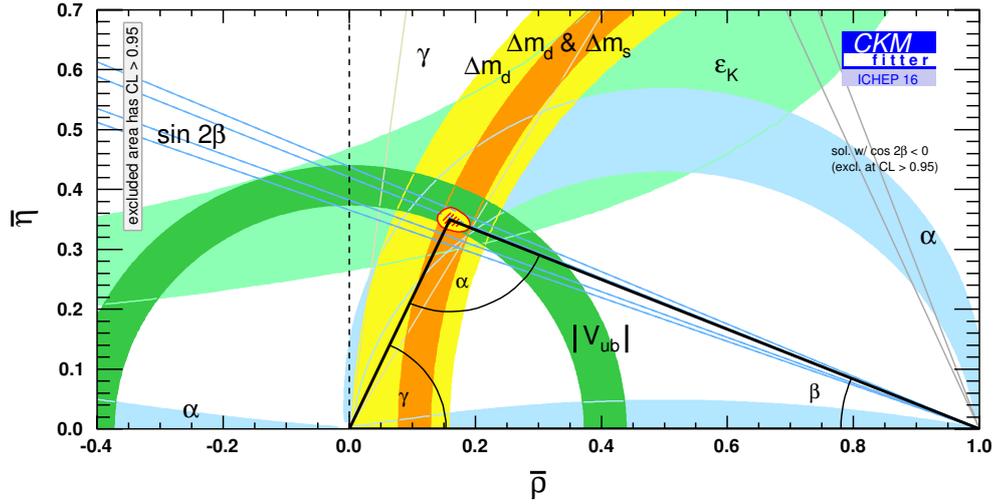
Figure 3.1: The unitarity triangle in the complex plane including the current experimental constraints [56]. The red hashed region is the 68 % confidence level area around the apex, which originates from a global fit that combines all existing measurements.

The second unitarity triangle that is not completely degenerated is given by

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \,. \tag{3.18}$$

One of this triangle's angles,

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right), \tag{3.19}$$

enters as phase in the B_s^0 meson mixing, as will be explained later in Sec. 3.5. It is not useful to draw the triangle, because the lengths of its sides are in the proportions $\mathcal{O}(\lambda^2) : \mathcal{O}(\lambda^2) : \mathcal{O}(\lambda^4)$. This is also the reason why the unitarity of the SM is primarily probed with the triangle shown in Fig. 3.1.

3.3 Mixing and decay of neutral *B* mesons

Mixing refers to the conversion of a particle into its antiparticle in systems of neutral self-conjugate pairs of mesons. It has been observed for $|B_s^0\rangle = |\bar{b}s\rangle$, $|B^0\rangle = |\bar{b}d\rangle$, $|K^0\rangle = |\bar{s}d\rangle$ and $|D^0\rangle = |c\bar{u}\rangle$ mesons. As the ground states of these mesons decay only via the weak interaction, they enable access to *CP* violation processes. The formalism that describes mixing is very similar for all mentioned mesons. Here, it will be outlined for the neutral *B* meson systems, using B_q^0 to denote either a B_s^0 or a

B^0 meson. The two flavour eigenstates

$$|B_q^0\rangle = |\bar{b}q\rangle \quad \text{and} \quad |\bar{B}_q^0\rangle = |b\bar{q}\rangle \tag{3.20}$$

have conjugated quark content and are eigenstates of the strong and electromagnetic interactions. A linear combination

$$|\Psi(t)\rangle = \Psi_1(t)|B_q^0\rangle + \Psi_2(t)|\bar{B}_q^0\rangle \tag{3.21}$$

of these eigenstates is governed by the effective Schrödinger equation for a two state system

$$i\frac{d}{dt}\begin{pmatrix}\Psi_1(t)\\\Psi_2(t)\end{pmatrix} = \mathbf{H}\begin{pmatrix}\Psi_1(t)\\\Psi_2(t)\end{pmatrix} = \left(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix}\Psi_1(t)\\\Psi_2(t)\end{pmatrix}, \tag{3.22}$$

with a non-hermitian Hamiltonian $\mathbf{H}$. The Hamiltonian consists of the hermitian matrices $\mathbf{M}$ and $\mathbf{\Gamma}$, which account for the masses and decay widths. Equality of the flavour-conserving transitions is guaranteed due to CPT invariance, which leads to $H_{11} = H_{22}$ and thus results in $|B_q^0\rangle$ and $|\bar{B}_q^0\rangle$ having the same mass $m_{11} = m_{22} = m$ and decay width $\Gamma_{11} = \Gamma_{22} = \Gamma$. The non-zero off-diagonal elements in

$$\mathbf{H} = \begin{pmatrix} m - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & m - \frac{i}{2}\Gamma \end{pmatrix}, \tag{3.23}$$

allow for the mixing of the flavour eigenstates, the transition of $B_q^0 \to \bar{B}_q^0$ and vice versa. More specifically, M_{12} enables virtual dispersive short-distance transitions, while Γ_{12} represents absorptive decays to physical final states common to both B_q^0 and $\bar{B}_q^0$. Flavour-changing neutral currents (FCNCs) are forbidden in the SM. Therefore, processes that involve quantum loops of charged currents are required to describe mixing. The corresponding Feynman diagrams are so-called box diagrams, of which the two diagrams of lowest-order are depicted in Fig. 3.2.
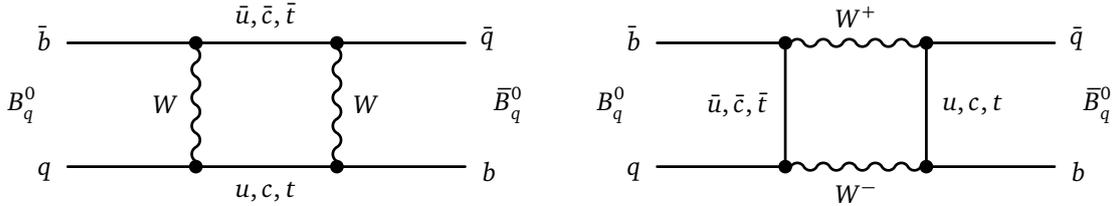
Figure 3.2: Short-distance box diagrams that depict the mixing of B_q^0 mesons in lowest order in the SM. The q denotes either a d or an s quark.

As the flavour-eigenstates are allowed to oscillate, they differ from physically observable mass eigenstates. However, mass eigenstates with well-defined masses and decay widths can be defined as linear combinations of the flavour eigenstates

$$|B_L\rangle = p|B_q^0\rangle + q|\bar{B}_q^0\rangle \quad \text{and} \quad |B_H\rangle = p|B_q^0\rangle - q|\bar{B}_q^0\rangle. \tag{3.24}$$

The complex coefficients p and q fulfil the normalisation condition $|p|^2 + |q|^2 = 1$ and have a ratio of

$$\frac{q}{p} = \sqrt{\frac{2M_{12}^* - i\Gamma_{12}^*}{2M_{12} - i\Gamma_{12}}} = -\frac{2\Delta m + i\Delta\Gamma}{2M_{12} - i\Gamma_{12}}. \tag{3.25}$$

The indices L and H denote the light and the heavy mass eigenstate, respectively, based on their mass difference

$$\Delta m = m_H - m_L \quad \text{with} \quad m_H + m_L = 2m. \tag{3.26}$$

Similarly, a decay width difference

$$\Delta\Gamma = \Gamma_L - \Gamma_H \quad \text{with} \quad \Gamma_H + \Gamma_L = 2\Gamma. \tag{3.27}$$

can be defined. The sign conventions of Δm and $\Delta\Gamma$ follow those suggested by the Heavy Flavour Averaging Group (HFLAV) [57]. Using the eigenvalues

$$\mu_L = m_L - \frac{i}{2}\Gamma_L \quad \text{and} \quad \mu_H = m_H - \frac{i}{2}\Gamma_H, \tag{3.28}$$

the time-propagation of the mass eigenstates can be written as

$$|B_{L,H}\rangle(t) = e^{-i\mu_{L,H}t}|B_{L,H}\rangle(0). \tag{3.29}$$

Combining Eqs. 3.24 to 3.29 leads to the time evolution of pure initial (*i.e.* at $t = 0$) flavour eigenstates

$$|B_q^0(t)\rangle = g_+(t)|B_q^0\rangle + \frac{q}{p}g_-(t)|\bar{B}_q^0\rangle, \tag{3.30}$$

$$|\bar{B}_q^0(t)\rangle = g_+(t)|\bar{B}_q^0\rangle + \frac{p}{q}g_-(t)|B_q^0\rangle. \tag{3.31}$$

The time-dependent coefficients $g_+(t)$ and $g_-(t)$ are given by

$$g_+(t) = e^{-i\frac{\mu_L+\mu_H}{2}t}\left[+\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} - i\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right], \tag{3.32}$$

$$g_-(t) = e^{-i\frac{\mu_L+\mu_H}{2}t}\left[-\sinh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta m t}{2} + i\cosh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta m t}{2}\right]. \tag{3.33}$$

Time evolutions of eigenstates are not directly observable. Instead, final states of their decays are being examined. Considering a final state f and its *CP* conjugated state $\bar{f}$ that are both common to B_q^0 and $\bar{B}_q^0$ gives four direct decay transitions

$$\begin{aligned} A_f &= \langle f|\mathbf{T}|B_q^0\rangle, \quad \bar{A}_f = \langle f|\mathbf{T}|\bar{B}_q^0\rangle, \\ A_{\bar{f}} &= \langle\bar{f}|\mathbf{T}|B_q^0\rangle, \quad \bar{A}_{\bar{f}} = \langle\bar{f}|\mathbf{T}|\bar{B}_q^0\rangle, \end{aligned} \tag{3.34}$$

governed by the Hamiltonian **T** of the weak interaction. Physically measurable are the time-dependent decay rates

$$\Gamma(B_q^0(t) \to f) = \left| \langle f | \mathbf{H} | B_q^0(t) \rangle \right|^2, \quad \Gamma(\bar{B}_q^0(t) \to f) = \left| \langle f | \mathbf{H} | \bar{B}_q^0(t) \rangle \right|^2, \quad (3.35)$$

$$\Gamma(B_q^0(t) \to \bar{f}) = \left| \langle \bar{f} | \mathbf{H} | B_q^0(t) \rangle \right|^2, \quad \Gamma(\bar{B}_q^0(t) \to \bar{f}) = \left| \langle \bar{f} | \mathbf{H} | \bar{B}_q^0(t) \rangle \right|^2, \quad (3.36)$$

of pure initial B_q^0 and $\bar{B}_q^0$ mesons that decay at the time t. The decay rates can be expressed in a phase-convention independent way as

$$\Gamma(B_q^0(t) \to f) = |A_f|^2 \qquad \left[|g_+(t)|^2 + |\lambda_f|^2 |g_-(t)|^2 - 2\mathrm{Re}\left(\lambda_f g_+^*(t) g_-(t) \right) \right],$$

$$\Gamma(\bar{B}_q^0(t) \to f) = |\bar{A}_f|^2 \left| \frac{p}{q} \right|^2 \left[|g_-(t)|^2 + |\lambda_f|^2 |g_+(t)|^2 - 2\mathrm{Re}\left(\lambda_f g_+(t) g_-^*(t) \right) \right],$$

$$\Gamma(B_q^0(t) \to \bar{f}) = |A_{\bar{f}}|^2 \left| \frac{q}{p} \right|^2 \left[|g_-(t)|^2 + |\bar{\lambda}_{\bar{f}}|^2 |g_+(t)|^2 - 2\mathrm{Re}\left(\bar{\lambda}_{\bar{f}} g_+(t) g_-^*(t) \right) \right],$$

$$\Gamma(\bar{B}_q^0(t) \to \bar{f}) = |\bar{A}_{\bar{f}}|^2 \qquad \left[|g_+(t)|^2 + |\bar{\lambda}_{\bar{f}}|^2 |g_-(t)|^2 - 2\mathrm{Re}\left(\bar{\lambda}_{\bar{f}} g_+^*(t) g_-(t) \right) \right].$$

$$(3.37)$$

Equations 3.37 are making use of the phase-convention independent quantities

$$\lambda_f = \frac{1}{\bar{\lambda}_f} = \frac{q}{p} \frac{\bar{A}_f}{A_f} \quad \text{and} \quad \bar{\lambda}_{\bar{f}} = \frac{1}{\lambda_{\bar{f}}} = \frac{p}{q} \frac{A_{\bar{f}}}{\bar{A}_{\bar{f}}}, \quad (3.38)$$

which are complex parameters with physical importance in terms of *CP* violation as explained in Sec. 3.4. By replacing

$$|g_\pm(t)|^2 = \frac{e^{-\Gamma t}}{2} \left[\cosh\left(\frac{\Delta\Gamma t}{2} \right) \pm \cos(\Delta m t) \right], \quad (3.39)$$

$$g_+^*(t) g_-(t) = \frac{e^{-\Gamma t}}{2} \left[\sinh\left(\frac{\Delta\Gamma t}{2} \right) + i \sin(\Delta m t) \right], \quad (3.40)$$

and introducing the six *CP* observables

$$A_f^{\Delta\Gamma} = \frac{-2\mathrm{Re}(\lambda_f)}{1 + |\lambda_f|^2}, \qquad C_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}, \qquad S_f = \frac{2\mathrm{Im}(\lambda_f)}{1 + |\lambda_f|^2},$$

$$A_{\bar{f}}^{\Delta\Gamma} = \frac{-2\mathrm{Re}(\lambda_{\bar{f}})}{1 + |\lambda_{\bar{f}}|^2}, \qquad C_{\bar{f}} = \frac{1 - |\lambda_{\bar{f}}|^2}{1 + |\lambda_{\bar{f}}|^2}, \qquad S_{\bar{f}} = \frac{2\mathrm{Im}(\lambda_{\bar{f}})}{1 + |\lambda_{\bar{f}}|^2}, \quad (3.41)$$

which fulfil the normalisation conditions

$$(A_f^{\Delta\Gamma})^2 + (C_f)^2 + (S_f)^2 = 1 \quad \text{and} \quad (A_{\bar{f}}^{\Delta\Gamma})^2 + (C_{\bar{f}})^2 + (S_{\bar{f}})^2 = 1, \quad (3.42)$$

the decay-rate equations from Eq. 3.37 can be rewritten as

$$
\frac{\Gamma(B_q^0(t) \to f)}{e^{-\Gamma t}} = \frac{1}{2}\left|A_f\right|^2 \quad (1 + |\lambda_f|^2)\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + A_f^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma t}{2}\right)\right.
$$
$$
\left. + C_f\cos(\Delta mt) - S_f\sin(\Delta mt)\right], \quad (3.43)
$$

$$
\frac{\Gamma(\bar{B}_q^0(t) \to f)}{e^{-\Gamma t}} = \frac{1}{2}\left|A_f\right|^2\left|\frac{p}{q}\right|^2 (1 + |\lambda_f|^2)\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + A_f^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma t}{2}\right)\right.
$$
$$
\left. - C_f\cos(\Delta mt) + S_f\sin(\Delta mt)\right], \quad (3.44)
$$

$$
\frac{\Gamma(B_q^0(t) \to \bar{f})}{e^{-\Gamma t}} = \frac{1}{2}\left|A_{\bar{f}}\right|^2 \quad (1 + |\lambda_{\bar{f}}|^2)\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + A_{\bar{f}}^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma t}{2}\right)\right.
$$
$$
\left. + C_{\bar{f}}\cos(\Delta mt) - S_{\bar{f}}\sin(\Delta mt)\right], \quad (3.45)
$$

$$
\frac{\Gamma(\bar{B}_q^0(t) \to \bar{f})}{e^{-\Gamma t}} = \frac{1}{2}\left|A_{\bar{f}}\right|^2\left|\frac{p}{q}\right|^2 (1 + |\lambda_{\bar{f}}|^2)\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + A_{\bar{f}}^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma t}{2}\right)\right.
$$
$$
\left. - C_{\bar{f}}\cos(\Delta mt) + S_{\bar{f}}\sin(\Delta mt)\right]. \quad (3.46)
$$

Terms with $\cos(\Delta mt)$ represent decays with a net oscillation, while those proportional to $\cosh(\frac{\Delta\Gamma}{2}t)$ can be identified as decays without a net oscillation of the initial state. Furthermore, terms with $\sin(\Delta mt)$ and $\sinh(\frac{\Delta\Gamma}{2}t)$ originate from the interference of direct decay and decay after mixing.

3.4 Manifestations of *CP* violation

CP violation can only be measured in quantities that are independent of phase conventions. Three different types of phases need to be considered in this context. The first type comprises spurious phases between a state and its *CP* conjugate. For example, applying a *CP* transformation to the initial and final states of the direct decay transitions given in Eq. 3.34 results in

$$
\begin{aligned}
CP|B_q^0\rangle &= e^{+i\xi_{B_q}}|\bar{B}_q^0\rangle, \quad CP|f\rangle = e^{+i\xi_f}|\bar{f}\rangle, \\
CP|\bar{B}_q^0\rangle &= e^{-i\xi_{B_q}}|B_q^0\rangle, \quad CP|\bar{f}\rangle = e^{-i\xi_f}|f\rangle.
\end{aligned} \quad (3.47)
$$

Here, the convention $(CP)^2 = 1$ is used. The spurious phases ξ_{B_q} and ξ_f are non-observable, as they do not carry any physical meaning. Next, there are strong phases, which are *CP* invariant, because they originate from final-state interaction scattering from on-shell states due to strong interactions. Finally, there are weak phases, which are *CP* odd and arise from complex couplings in the Lagrangian. In the SM, they manifest in phases of CKM matrix elements. Absolute values of strong and weak

phases are always convention-dependent, while phase differences are convention-independent and thus can be physically observed.

Due to different origins and interplays of strong and weak phases, *CP* violation can be categorised into three manifestations. Firstly, *CP* violation in the decay leads to different decay amplitudes between *CP* conjugated processes. Secondly, indirect *CP* violation has the consequence that the mixing of a neutral meson into its antiparticle has a different probability than the reversed process. Lastly, *CP* violation can occur in the interference of direct decay and decay after mixing. In the following sections, the three manifestation of *CP* violation will be discussed in detail.

3.4.1 Direct *CP* violation

For the occurrence of direct *CP* violation it is necessary that at least two amplitudes A_i with different strong δ_i and weak phases ϕ_i are contributing to a decay, so that

$$A_f = \sum_i A_i e^{i(\delta_i + \phi_i)} \quad \text{and} \quad \overline{A}_{\bar{f}} = e^{2i(\xi_f - \xi_{B_q})} \sum_i A_i e^{i(\delta_i - \phi_i)}. \tag{3.48}$$

Here, ξ_f and ξ_{B_q} are spurious phases, which are, as discussed in the previous section, arbitrary. If f is a *CP* eigenstate, its *CP* eigenvalue is given by $e^{2i\xi_f} = \pm 1$. The convention-independent quantity that measures direct *CP* violation is

$$\left| \frac{A_f}{\overline{A}_{\bar{f}}} \right| = \left| \frac{\sum_i A_i e^{i(\delta_i + \phi_i)}}{\sum_i A_i e^{i(\delta_i - \phi_i)}} \right| \neq 1. \tag{3.49}$$

Direct *CP* violation is the only manifestation of *CP* violation that can not only arise for neutral hadrons, but also for electrically charged ones. First evidence for direct *CP* violation was found 1988 in the kaon system by the NA31 experiment at CERN [58]. The measurement was confirmed by the successive NA48 experiment as well as the KTeV experiment at Fermilab in 1999 [59, 60]. For B^0 mesons, direct *CP* violation was first observed in 2001 by the *B* factories [61, 62]. It is now well established for both B^0 and B_s^0 mesons, for example in charmless two-body decays [63–65].

3.4.2 Indirect *CP* violation

Indirect *CP* violation requires the mass eigenstates to be non *CP* eigenstates. This is the case if a non vanishing phase difference

$$\phi = \phi_q - \phi_\Gamma \neq 0 \tag{3.50}$$

exists between the matrix elements

$$M_{12} = |M_{12}|e^{i\phi_q} \quad \text{and} \quad \Gamma_{12} = |\Gamma_{12}|e^{i\phi_\Gamma}. \tag{3.51}$$

Here, M_{12} and Γ_{12} are the elements of the transition matrix defined in Eq. 3.23. Following Eq. 3.25, this relative phase leads to

$$\left|\frac{p}{q}\right| \neq 1 , \qquad (3.52)$$

which subsequently gives different mixing rates for B_q^0 and $\bar{B}_q^0$,

$$\Gamma(B_q^0 \to \bar{B}_q^0) \neq \Gamma(\bar{B}_q^0 \to B_q^0) . \qquad (3.53)$$

Therefore, indirect *CP* violation is referred to as *CP* violation in mixing. It was the first type of *CP* violation to be observed, in 1964 for neutral kaons by Christenson et al. [9]. For neutral *B* mesons, the SM predicts this type of *CP* violation to be very small as $|\Gamma_{12}| \ll |M_{12}|$ for both B_s^0 and B^0. This prediction is reasoned for B_s^0 mesons as outlined in the following. In general, the matrix element Γ_{12} is the sum over all states f that are common final states of both B_s^0 and $\bar{B}_s^0$

$$\Gamma_{12} = \sum_f \langle f | T | B_q^0 \rangle \cdot \langle f | T | \bar{B}_q^0 \rangle . \qquad (3.54)$$

Tree-level processes are the dominant contribution to the absorptive transitions in the SM. These are either Cabbibo-suppressed decays for both B_s^0 and $\bar{B}_s^0$, or a Cabbibo-favoured decay for one meson and a doubly Cabbibo-suppressed decay for the other. Therefore, it can be assumed that $\Gamma_{12} \ll \Gamma$. Due to $M_{12} \approx \Delta m_s/2$ and $x_s = \Delta m_s/\Gamma_s \approx 27$, as given in Ref. [24], it is reasoned that the dispersive mixing transition is dominating. Thus, the relation $|\Gamma_{12}| \ll |M_{12}|$ is justified in terms of the SM, leading to negligible indirect *CP* violation for neutral *B* mesons. This prediction is probed by searching for a time-dependent decay rate asymmetry

$$a_{sl}^q = \frac{\Gamma(\bar{B}_q^0(t) \to l^+ \nu X) - \Gamma(B_q^0(t) \to l^- \nu X)}{\Gamma(\bar{B}_q^0(t) \to l^+ \nu X) - \Gamma(B_q^0(t) \to l^- \nu X)} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \qquad (3.55)$$

in semi-leptonic B_q^0 decays. To date, all measurements in the B_s^0 and B^0 systems, are in agreement with the SM prediction. The latest results come from the LHCb collaboration [66, 67].

3.4.3 *CP* violation in interference

The third type of *CP* violation does not require the presence of either direct or indirect *CP* violation. Direct *CP* violation cannot exist, if only a single weak phase ϕ_f is contributing to a decay $A_f = |A_f| e^{i(\phi_f + \delta_f)}$, while indirect *CP* violation can be excluded if $q/p = e^{i\phi_s}$ as a result of $|\Gamma_{12}/M_{12}| \approx 0$. Absence of both direct and indirect *CP* violation leads to $|\lambda_f| = 1$, with λ_f being the parameter defined in 3.38. However, *CP* violation can still occur in decays, in which the direct decay $B_q^0 \to f$ and the

decay after mixing $B_q^0 \rightarrow \overline{B}_q^0 \rightarrow f$ are able to interfere with each other. This type of *CP* violation is enabled by a phase difference between mixing and decay, which results in

$$\mathrm{Im}(\lambda_f) \neq 0 \,. \tag{3.56}$$

In the case of one *CP* final state f_{CP} that is common to both B_q^0 and $\overline{B}_q^0$, *CP* violation manifests in a single non-vanishing time-dependent asymmetry

$$\mathscr{A}_{f_{CP}}(t) = \frac{\Gamma(\overline{B}_q^0(t) \rightarrow f_{CP}) - \Gamma(B_q^0(t) \rightarrow f_{CP})}{\Gamma(\overline{B}_q^0(t) \rightarrow f_{CP}) + \Gamma(B_q^0(t) \rightarrow f_{CP})} \,. \tag{3.57}$$

By making use of the decay rate equations 3.43 and 3.44 as well as of the *CP* observables defined in 3.41, the *CP* asymmetry can be written as

$$\mathscr{A}_{f_{CP}}(t) = \frac{-C_f \cos(\Delta m t) + S_f \sin(\Delta m t)}{\cosh(\Delta \Gamma t/2) - A_f^{\Delta \Gamma} \sinh(\Delta \Gamma t/2)} \,. \tag{3.58}$$

As *CP* violation in the interference between decay and decay after mixing gives access to several parameters of the CKM sector, it is studied in several decay modes. Two of the most important ones are $B^0 \rightarrow J/\psi K_s^0$ and $B_s^0 \rightarrow J/\psi \phi$, where the most recent measurements were performed by the LHCb collaboration [68, 69]. The third type of *CP* violation is also the one that is relevant for the measurement presented in this thesis. It will be discussed in the context of $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ decays in the next section.

3.5 *CP* violation in $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ decays

In $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ decays, the conditions for *CP* violation in the interference between decay and decay after mixing are fulfilled, because both B_s^0 and $\overline{B}_s^0$ can decay into the same final state $D_s^- K^+$ (denoted f) without the need for a net oscillation. Both eigenstates are also able to directly decay into the charge conjugated final state $D_s^+ K^-$, labelled $\bar{f}$. Thus, all four direct decay amplitudes from 3.34 can be assigned to specific states as

$$\begin{aligned} A_f &= \langle D_s^- K^+ | \mathbf{T} | B_s^0 \rangle \,, \quad \overline{A}_f = \langle D_s^- K^+ | \mathbf{T} | \overline{B}_s^0 \rangle \,, \\ A_{\bar{f}} &= \langle D_s^+ K^- | \mathbf{T} | B_s^0 \rangle \,, \quad \overline{A}_{\bar{f}} = \langle D_s^+ K^- | \mathbf{T} | \overline{B}_s^0 \rangle \,. \end{aligned} \tag{3.59}$$

Moreover, this leads to the existence of all four decay rates defined in Eq. 3.43- 3.46, two time-dependent *CP* asymmetries $\mathscr{A}_f$ and $\mathscr{A}_{\bar{f}}$ (compare Eq. 3.58), and six *CP* observables as given in Eq. 3.41. A special feature of $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ decays is that all transitions in Eq. 3.59 are at the lowest order solely described by tree-level processes. Therefore, direct *CP* violation can be neglected, which leads to

$$|A_f| = |\overline{A}_{\bar{f}}| \quad \text{and} \quad |A_{\bar{f}}| = |\overline{A}_f| \,. \tag{3.60}$$
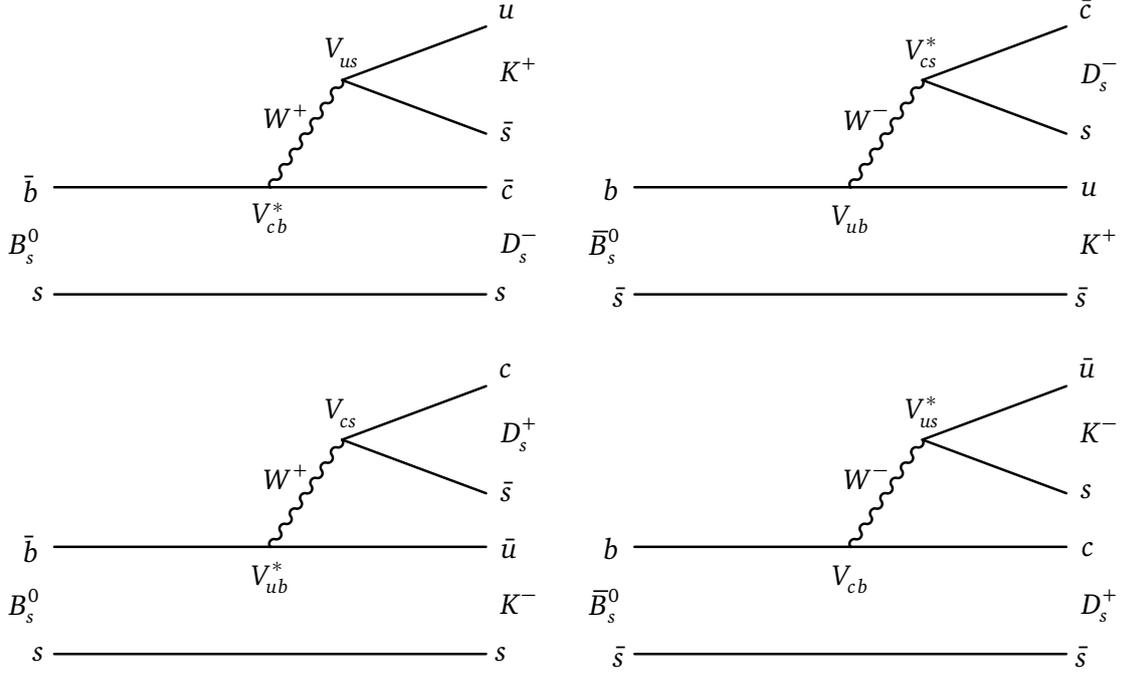
Figure 3.3: Tree-level diagrams of the four transitions from Eq. 3.59 lowest order, with A_f on the top left, $\overline{A}_f$ on the top right, $A_{\bar{f}}$ on the bottom left, and $\overline{A}_{\bar{f}}$ on the bottom right.

The four tree-level diagrams of the direct decays are shown in Fig. 3.10. They are all of the same order λ^3 with respect to the parameter of the Wolfenstein parametrisation (see Eq. 3.10), which favours interference. The branching ratio is currently measured to be [24]

$$\Gamma(B_s^0 \to D_s^\mp K^\pm) = (2.27 \pm 0.19) \times 10^{-4}. \tag{3.61}$$

In order to understand which parameters are accessible through *CP* violation in decays of $B_s^0 \to D_s^\mp K^\pm$, it is useful to define λ_f from Eq. 3.38 in terms of CKM matrix elements. The first step is to define the amplitude ratios, which can be calculated as

$$\frac{\overline{A}_f}{A_f} = \frac{V_{ub}V_{cs}^*}{V_{cb}^*V_{us}}\left|\frac{A_2}{A_1}\right| e^{i\delta} \quad \text{and} \quad \frac{\overline{A}_{\bar{f}}}{A_{\bar{f}}} = \frac{V_{cb}V_{us}^*}{V_{ub}^*V_{cs}}\left|\frac{A_1}{A_2}\right| e^{-i\delta}. \tag{3.62}$$

Here, δ is the strong phase difference in $B_s^0 \to D_s^\mp K^\pm$ decays, while A_1 and A_2 represent the hadronic amplitudes, *i.e.* what is left of the amplitudes once the CKM factors have been taken out. The second step is the ratio of the coefficients q/p for the B_s^0–$\overline{B}_s^0$ system. As mentioned in Sec. 3.4.2, *CP* violation in mixing can be neglected for neutral B mesons. Thus, q/p becomes a pure phase with $|q/p| \approx 1$. Furthermore,

Eq. 3.25 can be simplified to

$$\frac{q}{p} \approx \sqrt{\frac{M_{12}^*}{M_{12}}} = \sqrt{e^{2i\,\arg(M_{12}^*)}}, \tag{3.63}$$

which is sensible due to the assumption $|\Gamma_{12}| \ll |M_{12}|$ that was reasoned in Sec. 3.4.2. Neglecting spurious phases, the matrix element M_{12} for B_s^0 mesons is given by

$$M_{12} = -\frac{G_f^2 m_W^2}{12\pi} m_{B_s^0} f_{B_s^0}^2 B_{B_s^0} \left[\eta_1 (V_{cs}^* V_{cb})^2 S_0\left(\frac{m_c^2}{m_W^2}\right) + \eta_2 (V_{ts}^* V_{tb})^2 S_0\left(\frac{m_t^2}{m_W^2}\right) \right.$$
$$\left. + 2\eta_3 (V_{cs}^* V_{cb})(V_{ts}^* V_{tb}) S_0\left(\frac{m_c^2}{m_W^2}, \frac{m_t^2}{m_W^2}\right) \right]. \tag{3.64}$$

Here, m_W and $m_{B_s^0}$ are the masses of the W boson and the B_s^0 meson, respectively. Moreover, m_c represents the mass of the c quark and m_t that of the t quark. The Fermi constant is denoted by G_f, while $B_{B_s^0}$ is the bag parameter that describes, together with the weak decay constant $f_{B_s^0}$, corrections from non-perturbative QCD corrections that are related to the transition from bound to free quarks. Moreover, the η_i parameters originate from perturbative QCD corrections, while S_0 are the Inami-Lim functions [70], which depend on the mass ratio between the quark and the W boson. Due to the hierarchy of

$$S_0\left(\frac{m_t^2}{m_W^2}\right) \gg S_0\left(\frac{m_c^2}{m_W^2}, \frac{m_t^2}{m_W^2}\right) \gg S_0\left(\frac{m_c^2}{m_W^2}\right), \tag{3.65}$$

Eq. 3.64 can be simplified to

$$M_{12} = -\frac{G_F^2 m_W^2}{12\pi} m_{B_s^0} f_{B_s^0}^2 B_{B_s^0} \eta_2 (V_{ts}^* V_{tb})^2 S_0\left(\frac{m_t^2}{m_W^2}\right). \tag{3.66}$$

The hadronic parameters cancel, which leaves the following connection of q/p to the CKM matrix elements

$$\frac{q}{p} = \frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*} \quad \text{and} \quad \arg\left(\frac{q}{p}\right) = 2\beta_s. \tag{3.67}$$

Thus, the B_s^0 mixing phase is equivalent to the angle β_s introduced in Eq. 3.19. Combining Eq. 3.67 with Eq. 3.62 results in

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} = \frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*} \frac{V_{ub} V_{cs}^*}{V_{cb}^* V_{us}} \left|\frac{A_2}{A_1}\right| e^{i\delta} = |\lambda_f| e^{i(\delta-(\gamma-2\beta_s))},$$
$$\lambda_{\bar{f}} = \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_{\bar{f}}} = \frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*} \frac{V_{cb} V_{us}^*}{V_{ub}^* V_{cs}} \left|\frac{A_1}{A_2}\right| e^{-i\delta} = |\lambda_{\bar{f}}| e^{-i(\delta+(\gamma-2\beta_s))}. \tag{3.68}$$

Here, γ is one of the CKM angles defined in 3.17. Measuring both λ_f and $\lambda_{\bar{f}}$ in $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ enables an elimination of the strong phase δ, as

$$-\frac{1}{2}\left(\arg(\lambda_f) + \arg(\lambda_{\bar{f}})\right) = \gamma - 2\beta_s. \tag{3.69}$$

3.6 Determination of the CKM angle γ from $B_s^0 \to D_s^{\mp} K^{\pm}$ decays

Figure 3.1 illustrates that γ is the least well-known CKM angle as it has the broadest error band across the angles. The current world average of the Particle Data Group [24], which is calculated from a combination of all direct measurements, is

$$\gamma = (73.2^{+6.3}_{-7.0})^\circ. \tag{3.70}$$

Mostly contributing to the current precision are the LHCb experiment as well as the *B* factories Belle and BaBar.

Sensitivity to γ arises from the interference between favoured $b \to cW$ and suppressed $b \to uW$ transitions. The CKM angle γ does not depend on CKM elements that involve the top quark (see Eq. 3.17). Hence, it is the only CKM angle that can be measured solely with tree-level *B* decays. Examples for such tree-level decays are $B_s^0 \to D_s^{\mp} K^{\pm}$, $B^+ \to Dh^+$, $B^0 \to DK^{*0}$, $B^0 \to DK^+\pi^-$ and $B^+ \to Dh^+\pi^-\pi^+$. Here, *h* denotes a pion or a kaon and *D* refers to a neutral charm meson, indicating a mixture of the D^0 and $\bar{D}^0$ flavour eigenstates. The comparison of tree-level γ measurements with γ measurements that use decays such as $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$, which involve loop processes, allows for an interesting probe of New Physics contributions. Here, the focus are tree-level γ measurements and $B_s^0 \to D_s^{\mp} K^{\pm}$ in particular.

The tree-level measurements of γ can be categorised into time-dependent and time-integrated analyses. Time-dependent measurements of γ have been performed by the Belle and BaBar collaborations using $B^0 \to D^{(*)\mp}\pi^{\pm}$ decays [71–74] and by the LHCb collaboration using decays of $B_s^0 \to D_s^{\mp} K^{\pm}$ [75]. The latter analysis used ⅓ of the Run 1 data and is superseded by Ref. [15], which is performed in the course of this thesis. The time-integrated analyses can further be subdivided into four categories depending on the considered *D* final states. Firstly, the Gronau-London-Wyler (GLW) method uses *D* decays into *CP* eigenstates, like K^+K^- or $\pi^+\pi^-$ [76, 77]. Secondly, there is the Atwood-Dunietz-Soni (ADS) strategy [78, 79], which adds doubly Cabbibo-suppressed non-*CP* eigenstates such as π^-K^+. Very similar are the Grossman-Ligeti-Soffer (GLS) analyses [80], which exploit singly Cabbibo-suppressed *D* decays like $D \to K_S^0 K^-\pi^+$. Finally, there is the Giri-Grossman-Soffer-Zupan (GGSZ) method makes use of self-conjugate multibody decays as for example $D \to K_S^0 \pi^+\pi^-$ [81]. The GGSZ method is also often called the Dalitz method.

Both, time-integrated and time-dependent measurements, unites that their sensitivity to γ for a given $B \to X$ decay is approximately inversely proportional to the ratio r_B^X between the V_{cb} and V_{ub} transition amplitudes. Examples are

$$r_B^{DK} = \left| \frac{A(B^- \to \bar{D}^0 K^-)}{A(B^- \to D^0 K^-)} \right| \approx 0.1 \quad \text{and} \quad r_B^{D\pi} = \left| \frac{A(B^0 \to D^+\pi^-)}{A(\bar{B}^0 \to D^+\pi^-)} \right| = \mathcal{O}(0.01), \tag{3.71}$$

where the numerical values are taken from Ref. [24]. The value of $r_B^{D\pi}$ is about a magnitude smaller than r_B^{DK}, which can be explained by the enhancement of the fa-

voured amplitude through $|V_{ud}|/|V_{us}|$, while the suppressed transition is even further reduced by $|V_{cd}|/|V_{cs}|$.

For $B_s^0 \to D_s^{\mp} K^{\pm}$ decays, the amplitudes are proportional to the same order λ^3 of the Wolfenstein parameter. Due to incorporating V_{ub}, the processes $\overline{A}_f$ and $A_{\bar{f}}$, are slightly suppressed compared to A_f and $\overline{A}_{\bar{f}}$, leading to

$$r_B^{D_s K} = r_{D_s K} = \left| \frac{A_{\bar{f}}}{A_f} \right| = \left| \frac{\overline{A}_f}{\overline{A}_{\bar{f}}} \right| \approx |(\rho - i\eta)| \approx 0.37 \,. \tag{3.72}$$

In terms of sensitivity to γ, the large value of $r_{D_s K}$ compensates for the small doubly Cabbibo-suppressed branching ratios (see Eq. 3.61) as it favours interference between the amplitudes. Furthermore, negligible contributions from higher order loop processes enable small theoretical uncertainties of about $\delta\gamma/\gamma = 10^{-7}$ [82], which allows for a model-independent measurement.

CP violation in the interference between decay and decay after mixing leads to a non vanishing imaginary part of λ_f. Sensitivity to the CKM angle γ enters in the phase of λ_f. Subsequently, the *CP* observables from Eq. 3.41 can be redefined for $B_s^0 \to D_s^{\mp} K^{\pm}$ by using Eq. 3.68

$$
\begin{aligned}
A_f^{\Delta\Gamma} &= \frac{-2|\lambda_f|\cos(\delta - (\gamma - 2\beta_s))}{1 + |\lambda_f|^2} \,, & A_{\bar{f}}^{\Delta\Gamma} &= \frac{-2|\lambda_{\bar{f}}|\cos(\delta + (\gamma - 2\beta_s))}{1 + |\lambda_{\bar{f}}|^2} \,, \\
S_f &= \frac{2|\lambda_f|\sin(\delta - (\gamma - 2\beta_s))}{1 + |\lambda_f|^2} \,, & S_{\bar{f}} &= \frac{-2|\lambda_{\bar{f}}|\sin(\delta + (\gamma - 2\beta_s))}{1 + |\lambda_{\bar{f}}|^2} \,, \\
C_f &= \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2} \,, & C_{\bar{f}} &= -C_f \,.
\end{aligned}
\tag{3.73}
$$

The relation $C_{\bar{f}} = -C_f = C$ originates from equality between the absolute values of

$$|\lambda_f| = |\overline{\lambda}_{\bar{f}}| = \frac{1}{|\overline{\lambda}_f|} = \frac{1}{|\lambda_{\bar{f}}|} \,, \tag{3.74}$$

which itself is a consequence of the absence of direct *CP* violation.

The first goal of the analysis is to extract the *CP* observables from the measured decay rates. Afterwards, the weak phase $(\gamma - 2\beta_s)$ can be determined. The uncertainty on this difference is expected to be large compared to the experimental uncertainty on β_s, which, neglecting penguin pollution and assuming no contributions beyond the SM, is connected to the weak phase ϕ_s from $B_s^0 \to J/\psi\,\phi$. The current world average of this phase is [57]

$$\phi_s = -\arg\left(\frac{q}{p}\right) = -2\beta_s = -(0.030 \pm 0.033)\,\text{rad} \,. \tag{3.75}$$

Using ϕ_s as an external input allows to determine γ from the time-dependent measurement of *CP* violation in $B_s^0 \to D_s^{\mp} K^{\pm}$.

4 The LHC and the LHCb Experiment

The Large Hadron Collider (LHC) at the European Organisation for Nuclear Research (CERN) near Geneva, Switzerland, is the world's most powerful particle accelerator. One of the four large detectors operated at the LHC is the LHCb experiment that recorded the data analysed in context of this thesis. The LHC accelerator complex, as well as the remaining three experiments ALICE, ATLAS, and CMS, are briefly discussed in Sec. 4.1. In the following Sec. 4.2, the LHCb detector is described in detail, focussing on the parts responsible for track reconstruction and particle identification. Very important for the success of the LHCb experiment is its trigger system, introduced in Sec. 4.3, which is responsible for reducing the data rate to a handleable amount while keeping most physical processes of interest. The analysis and processing of the recorded data makes use of a specific software framework, which is described in Sec. 4.4. Furthermore, the comparison with simulated data is crucial for the understanding of the recorded data. Types of simulated data used in the context of the presented analysis are explained in Sec. 4.5. For the precision analyses performed by the LHCb collaboration, it is important to take data taking conditions into account, as these can change between different run periods. The run conditions that are relevant for the data analysed in this thesis are outlined in Sec. 4.6.

4.1 The LHC and the CERN accelerator complex

The LHC is a proton-proton (pp) collider with a circumference of 26.7 km, which makes it the largest synchrotron in the world [83]. It is built in the tunnel of the former Large Electron-Positron Collider (LEP) at a depth of about 50 to 175 m. Two counter-rotating proton beams are accelerated inside the LHC. In 2011 and 2012, called the Run 1 data taking period, the centre-of-mass energy $\sqrt{s}$ reached 7 and 8 TeV, respectively. For Run 2, which started in 2015, $\sqrt{s}$ was increased to 13 TeV. Before protons are injected into the LHC, they have already been boosted to an energy of 450 GeV by a couple of pre-accelerating stages. The complete CERN accelerator complex is depicted in Fig. 4.1. The protons, which are produced from hydrogen gas by stripping of the electrons, enter the accelerator complex through the Linear Accelerator 2 (LINAC2). Subsequently, the particles are accelerated by the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), which is the last stage prior to the LHC. The eight arc sections of the LHC make use of 1232 superconducting dipole magnets to bend the beams on
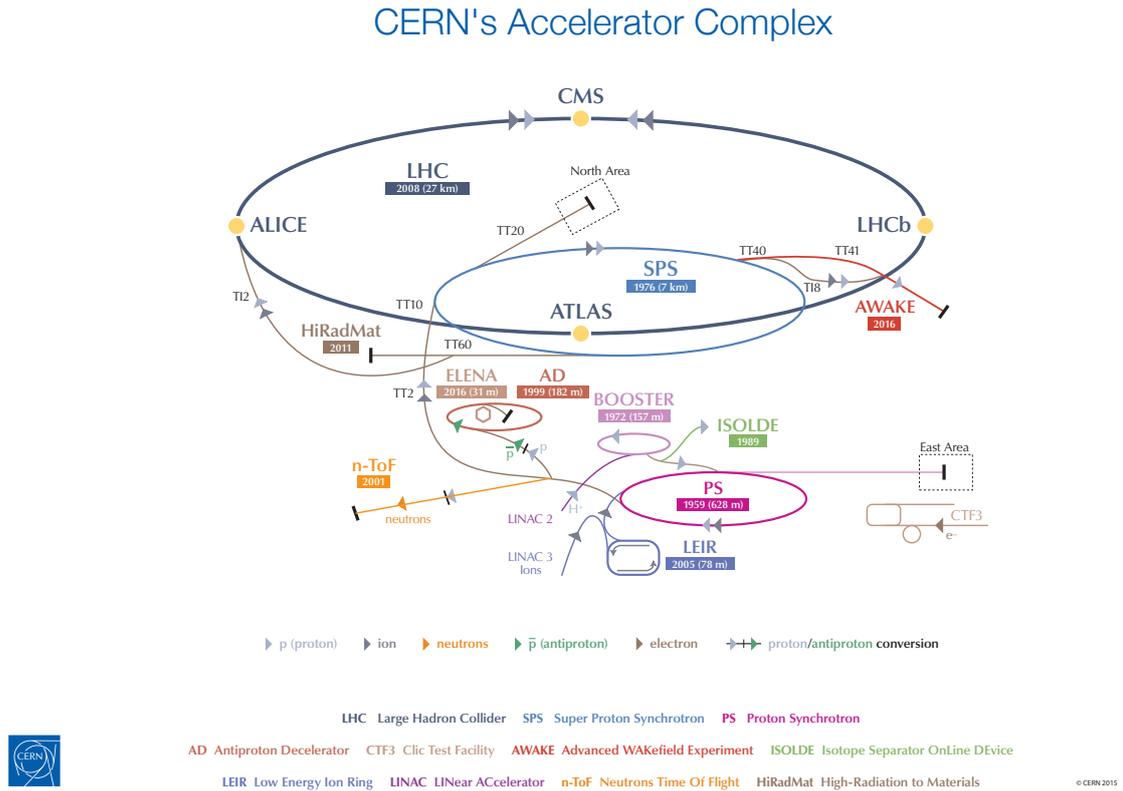
Figure 4.1: Overview of the accelerator complex operated at CERN [84].

the quasi circular track. Liquid ^{4}He cooles the NbTi dipole magnets to their operating temperature of $-271.3\,°$C. The eight straight sections of the LHC accelerate the beams and take care of beam cleaning and beam dumping. Focussing of the beams is done with quadrupole magnets. Each beam consists of up to 2808 bunches, which themselves comprise about 10^{11} protons. Inside the four detectors LHCb, ATLAS, CMS, and ALICE the beams are brought to collision at a rate of 20 MHz (40 MHz) in Run 1 (Run 2). The LHCb detector will be covered in detail in the next sections, while the other three experiments will be briefly introduced in the following paragraph.

Both ATLAS (A Toroidal LHC Apparatus, [85]) and CMS (Compact Muon Solenoid, [86]) are general purpose detectors (GPD). Their broad physics programmes vastly overlap in order to allow them to cross-check each other's results. The GPDs are designed to search for New Physics in a direct way. With the discovery of the Higgs boson [3–5] in 2012, ATLAS and CMS reached one of their major goals. The focus of the two experiments is now to find signatures of unknown particles, *e.g.* to find a candidate for dark matter. ALICE (A Large Ion Collider Experiment, [87]) focuses on studying a quark-gluon plasma, a state of matter produced in high energy ion-ion collisions. For this purpose, lead ions are injected into the LHC to perform lead-lead and lead-proton collisions. However, these lead runs only take place during small parts of the LHC data taking periods.

4.2 The LHCb detector

The Large Hadron Collider beauty (LHCb) experiment [88] performs precision measurements of hadrons containing beauty and charm quarks, focussing on *CP* violation and rare decays. It puts the Standard Model of particle physics to the test by comparing its experimental results with corresponding theoretical predictions. LHCb's indirect searches for New Physics are able to probe far higher energy scales than the direct searches performed by ATLAS and CMS, as unknown heavy particles may appear virtually in the observed processes. The LHCb detector, which is depicted in Fig. 4.2, is built as a single-arm forward spectrometer. It covers only the high pseudorapidity range of $2 < \eta < 5$. The pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where θ denotes the angle between a particle's momentum and the beam axis. LHCb's unique structure among the LHC experiments originates in the production process of *b* quarks at the LHC. In the *pp* collisions, *b* quarks are predominantly produced as $b\bar{b}$ pairs via gluon-gluon fusion [89]. Due to the large momentum asymmetry between the gluons present at LHC energies, the $b\bar{b}$ pairs are produced with a boost along the beam axis. Figure 4.3 shows a simulation of the kinematic correlation of produced $b\bar{b}$ quark pairs. By instrumenting an angular space of 10 mrad to 300 mrad, which translates to about 4.5 % of the whole space, the LHCb detector covers about 25 % of all produced $b\bar{b}$ quark pairs. Thus, a forward spectrometer is a cost effective solution for the physics program of the LHCb collaboration.
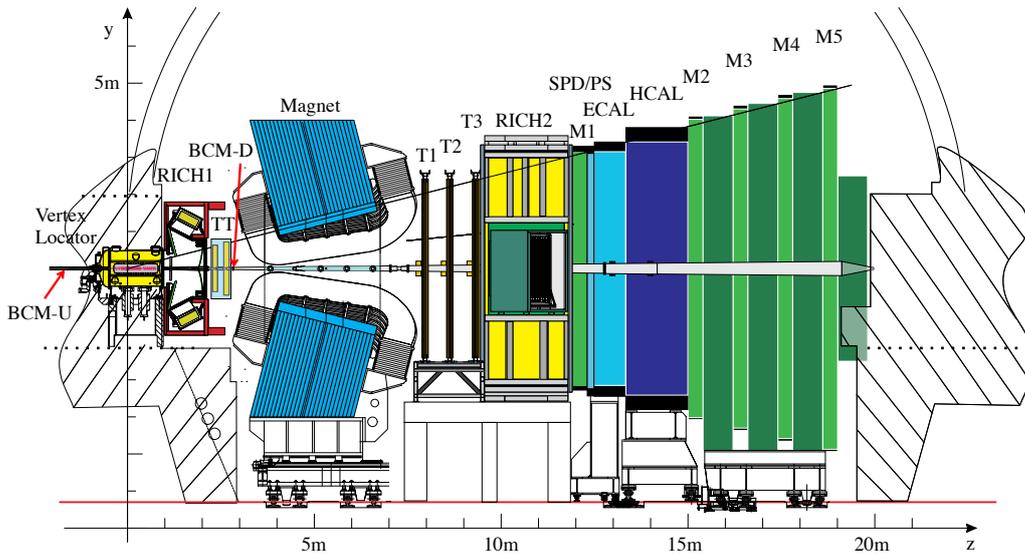

Figure 4.2: Schematic cross-section of the LHCb detector among the bending plane (y, z) at $x = 0$ [88]. The main *pp* interaction region is surrounded by the VELO located at the coordinate system's origin $(x = y = z = 0)$.
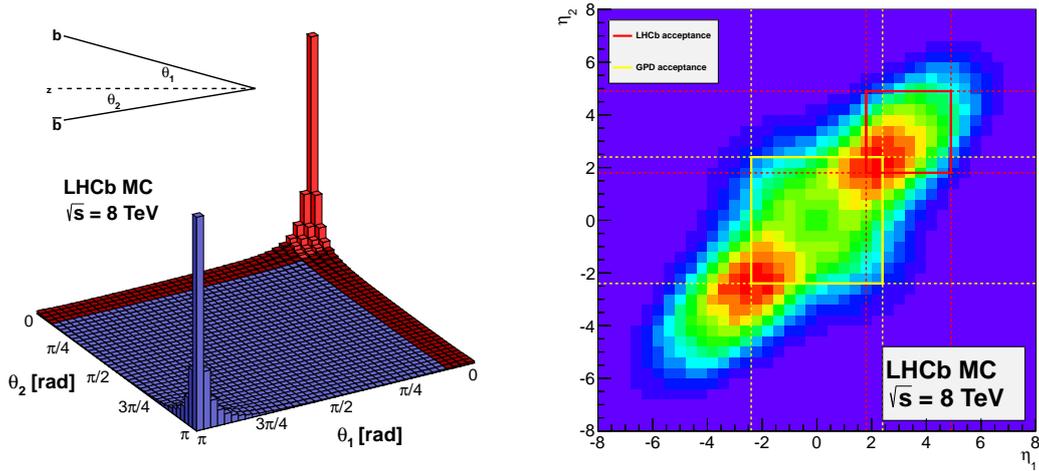
Figure 4.3: Correlations of $b\bar{b}$ quark pair production at $\sqrt{s} = 8\,\text{TeV}$ in terms of azimuthal angles $\theta_{1,2}$ (left) and pseudorapidities $\eta_{1,2}$ (right). The red region in the left plot indicates the acceptance of the LHCb detector. In the right plot the red rectangle illustrates the instrumented area of the LHCb experiment, while the yellow rectangle depicts the same for a general purpose detector like ATLAS or CMS.

For the physics analyses of the LHCb experiment, the precise determination of vertices, momenta, and decay times is vital, which is performed by the tracking system that is introduced in Sec. 4.2.1. Another crucial point is the identification of single particles and final states (see Sec. 4.2.2). A summary of the detector's performance can be found in Ref. [90].

4.2.1 Tracking system

The tracking system of the LHCb detector is responsible for determining vertices and tracks as well as momenta of traversing particles. It is comprised of several components out of which the first part is the Vertex Locator (VELO), a silicon strip detector surrounding the *pp* interaction point. Further parts are the large-area silicon strip detector Tracker Turicensis (TT) and the three tracking stations (T1–T3). The inner parts of the latter make use of silicon microstrips and form the so-called Inner Tracker (IT), while the outer parts are instrumented as drift-time detectors, shaping the Outer Tracker (OT). The final part of the tracking system is the warm dipole magnet, which is mounted between the TT and the T1 station. It provides a magnetic field with an integrated field strength of 4 Tm that bends tracks of charged particles and hence enables measurements of momenta. The magnet's polarity is switched regularly allowing studies on possible charge-dependent detection asymmetries. The different parts of the tracking will be discussed in more detail in the following.

The VELO is comprised of 42 half-disk shaped modules [91]. In order to measure

track hit positions in cylindrical coordinates related to the beam pipe, all modules have a ϕ sensor on one side and an R sensor on the other. The sensors use silicon microstrips with pitches varying between 40 to 100 µm. Traversing charged particles create electron-hole pairs in the silicon strips, which results in measurable currents. Figure 4.4 contains a photograph of the modules and illustrates the structure of the VELO. The arrangement of the modules in the VELO aims to cover the forward region in a way that all tracks inside LHCb's nominal acceptance of 10–300 mrad hit at least three of the modules. The VELO reaches a hit efficiency of 99 % and a single hit resolution of up to 4 µm. For data taking the modules are pulled to a proximity as close as 5 mm to the beam axis, with the half-disk slightly overlapping. However, during injection the beam aperture is larger than the innermost 8 mm radius of the sensitive module area. In order to prevent damage to the VELO, the halves are opened to a safe position of about 30 mm away from the beam until stable beam conditions are present. A crucial task of the VELO is to find and exactly localise the primary pp interaction, called primary vertex (PV), and the points where long-lived particles originating from the PV decayed. For B mesons, these so-called secondary vertices (SV) are typically detached by a few millimetres from the PV. The distance between the PV and SV is crucial for the measurement of decay times. In case of B mesons the VELO achieves a decay-time resolution of about 50 fs [91]. Furthermore, for tracks with $p_\mathrm{T} > 1\,\mathrm{GeV}/c$, where p_T denotes the transverse momentum, an impact parameter (IP) resolution of < 35 µm is accomplished [91]. The IP denotes the shortest distance of a track to the PV and is determined dependent on a particle's transverse momentum. An important protection system of the VELO is the Beam Conditions Monitor (BCM). The two parts of the BCM are installed upstream (BCM-U) and downstream (BCM-D) of the VELO. By using eight radiation-hard diamond sensors, the BCM measures the particle flux and initiates a dump of the beams when instabilities are detected. Moreover, the BCM has an independent power supply and reports its status constantly. Whenever the BCM is not responding, a beam dump is triggered immediately.

The IT, which is the inner part of the tracking stations T1–T3, and the TT are both silicon microstrip detectors with a strip pitch of about 200 µm [90]. The outer parts of the tracking stations make use of straw drift tubes with a diameter of about 5 mm, which contain a gas that is ionised when a particle surpasses the tube. By measuring the time it takes for the electrons to reach the anode in the middle of the tube, it is possible to conclude back to the hit position. The TT is installed in front of the magnet while the tracking stations are installed after the magnet. Having tracking information from both sides of the magnet is crucial to determine curvature radii of tracks in order to calculate their momenta. Common to the TT and the three tracking stations is that they are composed out of four layers. These layers are arranged in a so-called *x-u-v-x* configuration, which means the first and last layers (x) are arranged vertically, while the second and third layers are tilted by $-5°$ (u) and $+5°$ (v), respectively. This arrangement allows to gain sensitivity in the y-axis. During Run 1, the hit efficiencies of TT and IT were greater than 99.7 %, while hit resolutions were between 50 µm and 55 µm [90]. The OT reached a hit efficiency of 99.2 % for
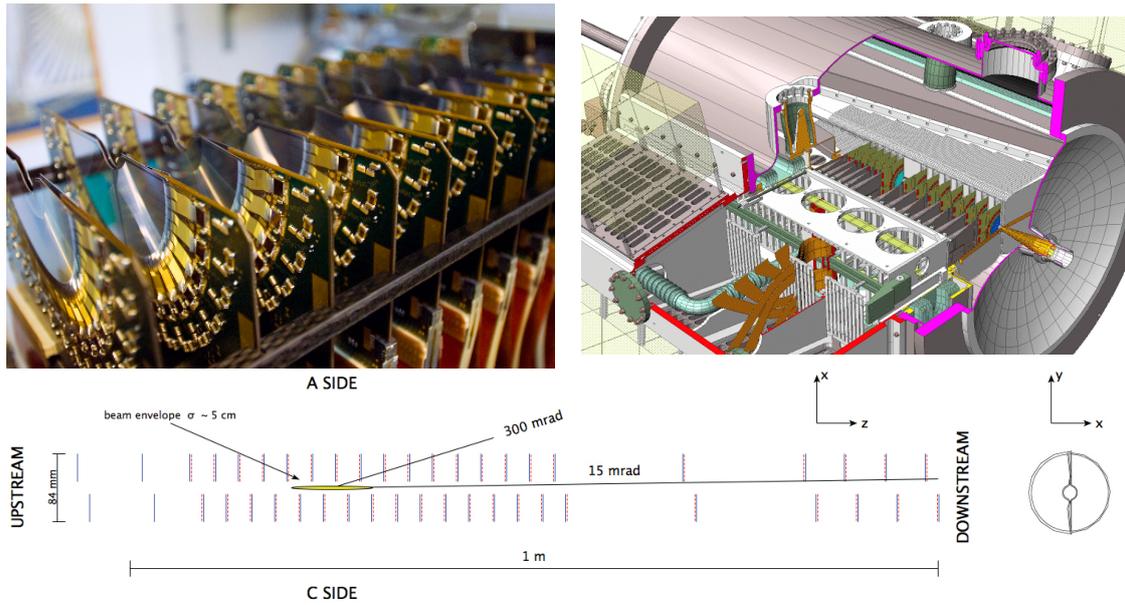
Figure 4.4: The top left figure shows a photograph of the VELO sensors during assembly. On the top right the VELO vacuum tank is illustrated, where the cut away enables to see the modules on the left side of the beam pipe. The bottom diagram depicts the arrangement of the modules in the VELO by a cross-section in the xz plane at $y = 0$. All figures are taken from Ref. [91].

the central half of the straw with a spatial resolution of 200 μm [92].

Tracks are divided into five categories depending on the parts of the tracking system they left hits in. VELO tracks have only hits inside the VELO, while tracks with additional hits in the TT are classified as upstream tracks. A track is a T track, if only hits in the tracking stations downstream of the magnet are present. Long-lived particles are able to decay outside of the VELO. Tracks of their daughter particles might only have hits in the TT and the tracking stations, which defines them as downstream tracks. The last track type are long tracks, corresponding to particles that left hits in all parts of the tracking system. The first three described track types are mostly feasible to improve the understanding of the detector, while downstream tracks are for example helpful when analysing decay modes with K_s^0 mesons in the final state. In the course of this thesis only long tracks are used.

4.2.2 Particle identification system

A very good particle identification (PID) is a prerequisite for the precision measurements the LHCb detector was built for. The PID system of the LHCb experiment consists of two ring-imaging Cherenkov (RICH) detectors, an electromagnetic and a hadronic calorimeter (ECAL, HCAL), as well as five muon chambers (M1-M5).

RICH1 is located between VELO and TT, while RICH2 is installed downstream of the tracking stations. The RICH detectors exploit the Cherenkov effect in order to

discriminate pions, kaons and protons. The Cherenkov effect describes the emission of a light cone by a charged particle in a dielectric when it is moving faster than the speed of light $c' = c/n$ present in the medium it is traversing [93, 94]. Considering a material with a refractive index n, the relation

$$\cos \theta_C = \frac{1}{n\beta} = \frac{c'}{v} \quad \text{with} \quad \beta = \frac{v}{c} \tag{4.1}$$

holds between the angle θ_C of the photon emission and the particle's speed v. In both RICH detectors, the Cherenkov light is reflected by a system of mirrors onto hybrid photo detectors (HPDs), as can be seen from the example of RICH1 in Fig. 4.5. The HPDs measure the radius of the light cone which allows to infer the emission angle. Combining measurements of momentum and Cherenkov angle enables to derive a hypothesis on a particle's identity, as can be seen in the right part of Fig. 4.5 for isolated tracks in RICH1. An isolated track is a track whose Cherenkov ring is not intersected by another ring from the same radiator. These tracks, which are about 2 % of all tracks, are useful to test the performance of a RICH detector. RICH1 is filled with C_4F_{10} (during Run 1 silica aerogel was used additionally) and covers particles in the momentum range from 1 GeV/c to 60 GeV/c. RICH2 is filled with CF_4 and provides discrimination tracks with higher momenta between 15 GeV/c and 100 GeV/c. The RICH systems achieved kaon identification efficiencies of about 95 % at pion misidentification rates of 10 % in Run 1 [95]. By applying stricter requirements, it is possible to lessen the misidentification rate to 3 %, while maintaining a kaon identification efficiency of 85 %.
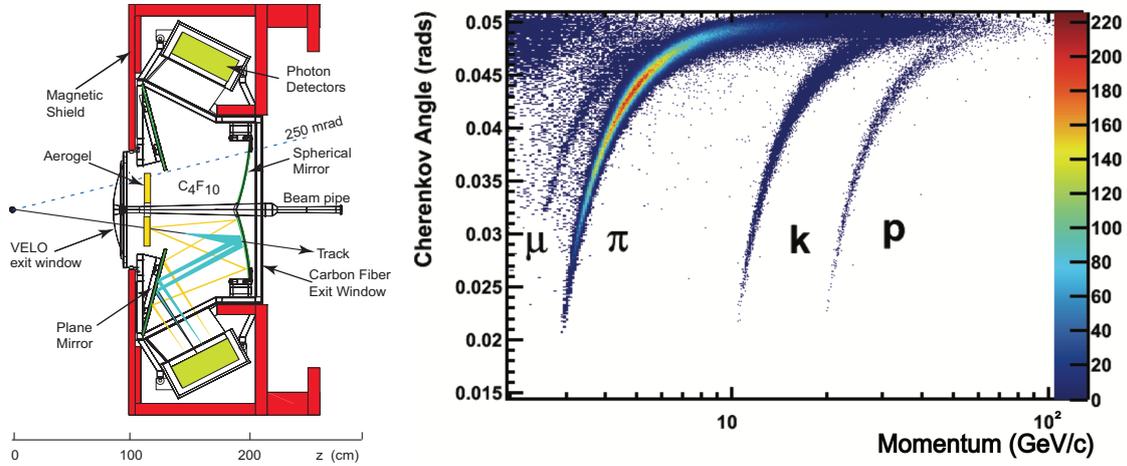

Figure 4.5: Side view of the RICH1 detector (left) [88]. The reconstructed Cherenkov angle as function of the track momentum for isolated tracks in the C_4F_{10} radiator of RICH1 [95]. Clear separations of pions, kaons, and protons are visible. A muon band exists as well.

The calorimeters are built out of alternating layers of metal and plastic. A surpassing particle creates showers in the metal layers that excite polystyrene molecules in the plastic layers, which themselves produce ultraviolet light in an amount proportional to the particle's energy. The ECAL is responsible for the identification of electrons and photons, while the HCAL measures energies of protons, neutrons and long-lived hadrons. A lead preshower (PS) and a Scintillator Pad Detector (SPD) are installed in front of the ECAL to reduce background from pions.

The PID system is completed by the five muon stations. Four of them are installed as the last components of the detector, while M1 is situated between RICH2 and the calorimeter system. Muons are the least interacting charged particles and are typically not stopped before M2. The muon chambers make use of multi-wire proportional chambers that are filled with a mixture of CO_2, Ar, and CF_4. A passing muon ionises the gas and an electrical field accelerates the electrons and ions. This creates an electrical current proportional to the muon energy. Layers of $80\,cm$ thick iron are installed between the last muon chambers, which are able to stop all muons with momenta $p < 6\,GeV/c$.

Information from all parts of the particle identification system are combined in order to calculate global PID likelihoods for particular particle hypotheses, as explained in Sec. 4.4.1.

4.3 The LHCb trigger system

During the Run 1 data taking period (see Sec. 4.6), the LHC provided a peak instantaneous luminosity of $6.8 \times 10^{33}\,cm^{-2}\,s^{-1}$, which is fully used by the general purpose detectors ATLAS and CMS. However, to enable the proposed precision analyses the LHCb detector runs at a significantly lower instantaneous luminosity of $4 \times 10^{32}\,cm^{-2}s^{-1}$. It is adjusted by reducing the overlap between the proton beams, which effectively reduces the amount of pp collisions per bunch crossing. However, even with the reduced rate the LHCb detector is running at, it is not possible to keep information from all collisions and to save everything at the full bunch crossing rate. Moreover, it would also be not efficient, as a $b\bar{b}$ quark pair is on average only produced every 200 to 350th pp collision. Therefore, the LHCb detector features a trigger system that brings the event rate down to about $5\,kHz$ ($3.5\,kHz$) in 2012 (2011), which is then low enough to be fully saved. Since 2015, *i.e.* the start of Run 2, it was possible to further increase the total trigger output rate to $12.5\,kHz$. The following subsections will describe the hardware and the software trigger, which are the two trigger stages of the LHCb experiment.

4.3.1 Hardware trigger

The hardware stage of the trigger (L0) runs synchronously to the full bunch crossing frequency and delivers a reduced event rate of around $1\,MHz$. It deduces information from the calorimeters and the muon chambers in order to decide whether an

event is rejected or maintained. The selection criteria from the L0 trigger aim to favour events that contain particles with high transverse energy, which are likely to be daughter particles of *b* hadrons. Thus, the L0 searches for clusters of high transverse energies in the calorimeters, which originate from electron, photon or hadron hits (L0Electron, L0Photon, L0Hadron). Only a limited cluster multiplicity in the SPD is allowed to avoid storing events in which the detector occupancy is too high. Furthermore, an event is kept if either the transverse momentum p_T (L0Muon) of at least one muon reaches a threshold, or if the p_T product of the two muons with the highest momenta is large enough (L0DiMuon). The specific limits are stored for every run period in trigger configuration keys (TCKs) as they are changed from time to time due to optimisations. Efficiencies of the L0 differ between muon and hadrons. The former are triggered more than 90 % efficient, while decays with fully hadronic final states (like $B_s^0 \to D_s^\mp K^\pm$) are triggered at lower efficiencies of about 60 % [96, 97].

4.3.2 Software trigger

The high level trigger (HLT) is the second trigger stage and is fully implemented in software. It is further divided into the two stages HLT1 and HLT2, which are running on the event filter farm (EFF). The EFF is a computer cluster with about 30 000 processor cores. Both HLT1 and HLT2 are organised in so-called trigger lines. An event survives, if it is accepted by at least one of the included lines in each of the stages.

At the 1 MHz output rate of the L0 trigger it is possible to fully read out all detector components. However, the available per-event time at the HLT1 stage is only sufficient to perform a partial event reconstruction in order to reconsider the L0 decisions. Hence, the HLT1 reconstructs the VELO tracks for all events and identifies vertices using at least five tracks. Primary vertices are considered to be vertices closer than 300 μm to the primary *pp* interaction, which itself is established at the beginning of a data taking run. Lines that do not use information from the muon system look for VELO tracks based on their smallest impact parameter with respect to any PV. For example, the Hlt1TrackAllL0 line selects tracks with a good quality by applying the requirement of a transverse momentum p_T larger than 1.6 GeV/c and a displacement from the primary vertex. This line takes the dominant part of the HLT1 bandwidth, with about 58 kHz in 2012 [97]. It is especially important for hadronic modes. Important HLT1 lines for decays with muons are the Hlt1TrackMuon and the Hlt1DiMuonHighMass. The former selects good muon candidates by requiring $p_T > 1$ GeV/c and a displacement from the primary vertex. The latter selects dimuon candidates by requiring their invariant mass to be larger than 2.5 GeV/c and demanding their tracks to originate from a common vertex. Altogether, the HLT1 lines had a rate of about 80 kHz in 2012 [97].

The HLT2 is able to fully reconstruct all incoming events with a minimum transverse momentum of 300 MeV/c. It contains more than 100 lines that implement inclusive and exclusive selections, further tightening the requirements of HLT1. For hadronic decays there are lines that make use of multivariate algorithms to find two-,

three- or four-track secondary vertices with large transverse momenta and significant displacements from the PV. All events accepted by HLT2 are stored on disk.

A novel feature of the LHCb experiment is its capability to exploit the times in which the LHC is not delivering collisions to run a so-called deferred trigger [98]. For this purpose, about 20 % of L0 triggered events are saved on the EFF's disks and evaluated by the HLT in between data taking runs. This strategy guarantees an optimal usage of the available processing power and exploits the storage opportunities of the EFF nodes.

4.4 Offline data processing with the LHCb software

The LHCb software contains software packages for the online and offline stages of the data processing. It is based on the GAUDI framework [99], which itself makes use of the ROOT framework [100, 101]. One of the packages in the LHCb software is the MOORE package [102]. It runs online and has implicitly already been introduced in the last section, as it is the package that holds the code of the HLT. This section focusses on the BRUNEL [103] and DAVINCI [104] packages, which run offline on triggered data, both from the detector and simulations, in order to perform preparations necessary for end-user analyses. The first step is a full event reconstruction done by the BRUNEL package. It is described in Sec. 4.4.1. The second step is a centralised preselection, the so-called stripping, which is taken care of by the DAVINCI package. It is explained in Sec. 4.4.2.

4.4.1 Event reconstruction

The event reconstruction comprises two tasks. Firstly, the tracking forms and combines tracks from hits of charged particles in the tracking system. Secondly, particle hypotheses are assigned to the tracks using information from the particle identification system.

Different tracking methods are utilised corresponding to the different track types (see Sec. 4.2.1). Two algorithms exist to find long tracks, which is the only track type relevant in the presented analysis. Both start with the forming of VELO tracks from hits in the ϕ and R sensors of the VELO that lie on a straight line. In case of the forward tracking algorithm, these VELO tracks are subsequently combined with matching hits in the tracking stations T1-T3, taking into account the bending from the magnetic field, and finally assigned to hits in the TT. The second algorithm is called the track matching algorithm, and it directly matches the VELO tracks to independent T tracks. In a second step compatible TT hits are added. After all track finding algorithms have scanned the event in order to find their designated track types, a Clone Killer algorithm removes all duplicate tracks. The principle of the clone killing is to delete tracks that share more hits than a specific limit with another longer one. A global Kalman fit [105] is performed as final step of the offline tracking in order to precisely determine the trajectories and momenta of particles. It considers both

multiple scattering effects as well as energy losses and takes the largest portion of the tracking's computing time. The outcome of the tracking is a state vector $\vec{s} = (x, y, t_x, t_y, q/p)^{\mathrm{T}}$ for each reconstructed track. Here, x and y are coordinates, t_x and t_y the slopes in the xz and xy projections, and q/p is the charge divided by the momentum.

After tracks have been established, particle hypotheses need to be assigned. Information from the RICH detectors identifies charged hadrons. Hence, all tracks are extrapolated into RICH1 and RICH2 and information from the HPD hits is added. Finally, global likelihood values are calculated for each track corresponding to different applied particle hypotheses. Muon hypotheses can be settled or excluded very efficiently, by searching for hits in the muon chambers that lie in the vicinity of extrapolated tracks. Clusters in the electromagnetic calorimeter are used to identify electrons and photons, with pairs of photons often indicating decays of neutral pions. Finally, all available information is used to conclude global PID likelihoods for specific particle hypotheses. Differences of the logarithmic likelihoods (Delta log-likelihood, DLL) with respect to the pion hypothesis are calculated in order to express particle discriminations as decisions between two types of particles. For example, the DLL of a kaon versus a pion hypothesis is given by

$$\mathrm{DLL}(K - \pi) = \ln(\mathscr{L}_K) - \ln(\mathscr{L}_\pi), \tag{4.2}$$

while for a muon-kaon separation it is calculated as

$$\mathrm{DLL}(\mu - K) = \mathrm{DLL}(\mu - \pi) - \mathrm{DLL}(K - \pi). \tag{4.3}$$

Hence, a $\mathrm{DLL}(A - B) > 0$ implies a higher probability of the particle to be A, whereas $\mathrm{DLL}(A - B) < 0$ favours particle type B. In the following, the most important DLLs are abbreviated as

$$\mathrm{PID}K = \mathrm{DLL}(K - \pi), \ \mathrm{PID}p = \mathrm{DLL}(p - \pi), \ \text{and} \ \mathrm{PID}\mu = \mathrm{DLL}(\mu - \pi). \tag{4.4}$$

4.4.2 Global preselection

The full LHCb dataset is far too large to be handled efficiently in offline physics analyses such as the presented *CP* violation measurement in $B_s^0 \to D_s^\mp K^\pm$. Thus, a centralised preselection, known as stripping, is performed. It consists of a large number of stripping lines, which define selection criteria dedicated to the broad range of decay modes and analyses examined by the LHCb collaboration. The software that is used to fill these lines is called the DAVINCI framework.

The DAVINCI framework provides algorithms that further process the so-called protoparticles, which are the output of the preceding event reconstruction. These algorithms transform the protoparticles into intermediate states and particle objects, calculate kinematic properties, and build full decay chains. The different stripping lines often share several selection steps in order to save computing time. As an example, there are minimal requirements applied to stable particles, which can subsequently be tightened in individual lines. In general, large overlaps exist between selection requirements of lines meant for decay modes with similar or identical particles

in the final state. However, the computing effort of the stripping remains enormous and new revisions are usually not released more often than once a year. Data that is rejected by the stripping is lost for any further analyses, as analysts can only access stripped data. Thus, stripping selections are much looser than the so-called offline selections applied in specific analyses such as the presented one. In this way the stripping maintains very high signal efficiencies, while at the same time enough candidates for studies on background properties are kept.

The stripping lines used in the course of the presented analysis do not inherit any PID selection criteria, but apply requirements on the particle kinematics and the displacement from the primary interaction. Details on the stripping requirements applied in the presented analysis are discussed in Sec. 7.1.

In the offline decay chain reconstruction, the DecayTreeFitter algorithm [106] is used to compute invariant masses and lifetime observables. The DecayTreeFitter, which is also provided in the DaVinci framework, takes correlations and uncertainties of its inputs, mostly being vertex positions, particle momenta, and flight distances into account. Furthermore, when the B invariant mass is computed, the invariant D mass is constrained to its nominal value [24], separately for the D and the D_s hypotheses. Additionally, in the calculation of the B decay times the B candidates are constrained to point to the associated PVs.

4.5 Simulated data

Simulated data, often referred to as Monte Carlo simulation (MC), is an important tool to understand and improve various parts of physics analyses at LHCb. A few examples are the estimations of efficiencies and systematic uncertainties, the modelling of observable distributions, or validity tests of analysis strategies like fitting procedures. The amount of MC is only limited by processing and storing capabilities. An essential input for the presented analysis is fully simulated MC, which has the aim to resemble real data as closely as possible. The generation process of this type of simulated data, which is explained in the context of LHCb in Sec. 4.5.1, is very computing intensive. Therefore, especially for the estimation of systematic uncertainties, less sophisticated and faster to generate types of MC are also used. These so-called pseudo-experiments are introduced in the Sec. 4.5.2.

4.5.1 Fully simulated Monte Carlo

The generation of the fully simulated MC at LHCb is performed by the GAUSS framework [107, 108], which makes use of different software packages in the process. It starts with the simulation of *pp* collisions using the PYTHIA event generator [109, 110]. The decays of the produced intermediate particles are generated with EVTGEN [111], which is well suited for decays of *b*-hadrons, as it includes effects like mixing and *CP* violation. Final state radiation of photons is taken into account by the PHOTOS package [112]. After the particles have been generated, their interaction with the

detector material is simulated with the GEANT4 toolkit [113, 114]. The final step in the generation of the MC is performed by the BOOLE software [115]. It provides the digitisation, which is the transformation of particle interactions with the detector into responses of the particular detector components. As an example, particle showers in the calorimeters are converted into corresponding calorimeter signals. Moreover, it takes the readout electronics as well as the L0 trigger into account, and it is able to generate spillover events originating from previous or following beam crossings.

Simulated data can be and is treated like real detector responses from this point on and undergoes the same subsequent processing chain. This implies that the MC is processed subsequently by MOORE, BRUNEL, and DAVINCI, in order to calculate the high level trigger responses, do the event reconstruction, and perform the stripping selection. However, an important difference to data is that in MC the information about the true nature of the generated particles is maintained. Thus, DAVINCI provides the possibility to match true and reconstructed particles, with a technique called truth matching. With the truth matching it is *e.g.* possible to verify selections, measure efficiencies, or to examine remaining physics background contributions.

4.5.2 Pseudo-experiments and toy studies

Pseudo-experiments, often called toy MCs, are the type of simulated data that consumes the least amount of processing time during its generation. It is produced with a hit-and-miss procedure using a probability density function (PDF). The main purpose of toy MC in the course of this analysis is to verify the reliability of the fitting procedure. This is done by generating toy data and fitting it with the same or differing PDFs. Possible biases can be quantified by a commonly used technique called toy study. In such studies, the process of generating and fitting toy samples is repeated thousands of times, while saving the fit results of each step. Finally, a pull distribution is plotted and analysed. This is the cumulated distribution of the differences between the generation value and the fitted values during the study, normalised to the fit uncertainty. In case of an unbiased fit with correct error estimates the pull distribution is a standard normal distribution, which can easily be tested with a fit. Toy studies are especially useful to estimate systematic effects by introducing differences between the PDFs used for the generation and the fitting.

4.6 Run conditions

The data analysed in the course of this thesis was recorded in the years 2011 and 2012, which represent LHC's Run 1 data taking period. In 2011, an integrated luminosity of about $1\,\text{fb}^{-1}$ at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ was taken, while in 2012 the centre-of-mass energy was $\sqrt{s} = 8\,\text{TeV}$ and the recorded luminosity reached about $2\,\text{fb}^{-1}$. Figure 4.6 illustrates the development of the recorded integrated luminosity during Run 1 and the first three years of Run 2. In order to improve the beam stability, the LHC ran at a bunch crossing rate of $20\,\text{MHz}$ throughout Run 1, which
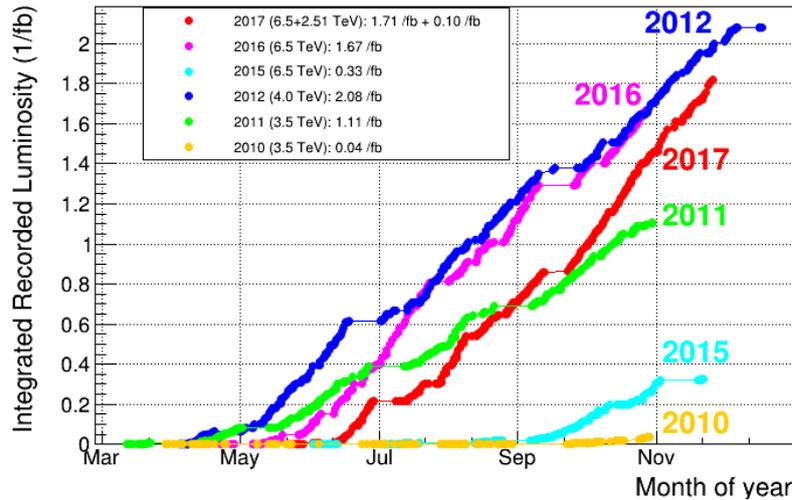
Figure 4.6: Evolution of the integrated Luminosity of the LHCb experiment during Run 1 as well as 2015–2017, the first three years of Run 2 [116].

was only half of its design rate. It was reduced as a consequence of the accident that happened shortly after the first proton beams were circulated in the LHC in 2008. A faulty electrical connection between two dipole magnets led to massive mechanical damages as helium was released from the magnet cold mass into the tunnel. The accident delayed the LHC program until late 2009. Despite the reduced bunch crossing rate, the LHC was still able to provide a peak instantaneous luminosity of about $6.8 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ in 2012, close to the target value of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This was achieved by enlarging the bunch intensities by 30 % to 1.5×10^{11}, which consequently increased the number of *pp* interactions per bunch crossing. As mentioned in Sec. 4.3, the LHCb detector runs at a significantly lower instantaneous luminosity. In 2011 it reached $3.5 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, while it peaked at $4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in 2012 [90]. However, LHCb maintains a constant instantaneous luminosity throughout an LHC fill, while those of ATLAS and CMS decline corresponding to the decreasing amount of protons left in the ring. A steady instantaneous luminosity is accomplished by a procedure called luminosity levelling [117], which constantly adjusts the crossings of the beams. The design pile-up of LHCb, which denotes the visible number of *pp* collisions per bunch crossing, is about 0.4, while the actual number during Run 1 reached up to 2.4. This pile-up value beyond the detector's specifications did not hinder LHCb's capability of disentangling and identifying most of the interesting tracks originating from the *pp* collisions, but more than compensated the lower bunch crossing frequency of the LHC. The data taking efficiency of LHCb, being the ratio of the recorded and the delivered luminosity, reached about 93 % in 2011 and 95 % in 2012.

5 Analysis Tools

This chapter introduces analysis tools and techniques that are used in different parts of this thesis. In Sec. 5.1, the maximum likelihood method is explained, which is fundamental for all parameter estimations performed in the course of the presented analysis. The *sPlot* technique is discussed in Sec. 5.2. This method allows to statistically separate different components, like signal and background, from each other. Finally, in Sec. 5.3 an introduction to boosted decision trees (BDTs) is given, which are an example for a multivariate selection technique. A BDT is utilised both in the selections of $B_s^0 \rightarrow D_s^\mp K^\pm$ decays and $B^0 \rightarrow J/\psi K^{*0}$ decays, where the latter are used for flavour tagging calibrations discussed later in Sec. 6.6.

5.1 Maximum likelihood method

The (unbinned) maximum likelihood method is a common method to estimate parameters in multidimensional fits and it is used throughout the analysis. Given an underlying probability density function (PDF), $\mathscr{P}(\vec{x}|\vec{\lambda})$, the principle of the maximum likelihood method is to find the set of parameters $\vec{\lambda}$ that maximises the probability to observe the measured sample of n observable vectors $\vec{x}$. The parameters are found by maximising the likelihood function

$$\mathscr{L}(\vec{\lambda}|\vec{x}) = \mathscr{P}(\vec{x}_1|\vec{\lambda}) \cdot \mathscr{P}(\vec{x}_1|\vec{\lambda}) \cdots \mathscr{P}(\vec{x}_n|\vec{\lambda}) = \prod_{i=1}^{n} \mathscr{P}(\vec{x}_i|\vec{\lambda}). \tag{5.1}$$

Multiplying a Poisson term leads to the extended maximum likelihood function

$$\mathscr{L}(\vec{\lambda}|\vec{x}) = \frac{e^{-N}N^n}{n!} \prod_{i=1}^{n} \mathscr{P}(\vec{x}_i|\vec{\lambda}), \tag{5.2}$$

which includes the possibility that n events are observed although N events are expected. In practice, the negative log likelihood

$$-\ln \mathscr{L}(\vec{\lambda}|\vec{x}) = -\sum_{i=1}^{n} \ln\left(\mathscr{P}(\vec{x}_i|\vec{\lambda})\right), \tag{5.3}$$

is minimised, which leads to the same parameters, but is a numerically more stable evaluation. Different data categories c, which originate from differing final states and backgrounds, can be fit simultaneously by using specific PDF representations, *i.e.*

$$\mathscr{P}(\vec{x}_i|\vec{\lambda}) = \prod_{c} \mathscr{P}^c(\vec{x}_i|\vec{\lambda}_c). \tag{5.4}$$

Furthermore, it is straightforward to parallelise the minimisation, both in terms of data and independent parts of the PDF. In the course of this analysis the ROOT implementation of the maximum likelihood method is used [100, 101], which utilises the MINUIT package [118]. It not only determines the central values of the parameters, but also gives an estimation of their uncertainties. Moreover, for fixed parameters it is possible to apply a Gaussian constraint. This implies multiplying the fit PDF with a Gaussian PDF

$$\mathcal{G}(\lambda|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{1}{2}\left(\frac{\lambda-\mu}{\sigma}\right)^2}, \tag{5.5}$$

which uses the parameter λ as variable and the parameter's central value and uncertainty as its mean μ and width σ, respectively.

5.2 The *sPlot* technique

The *sPlot* technique [119] utilises an extended maximum likelihood fit, the *sPlot* fit, in one or more dimensions to compute so-called *sWeights*. These per-event weights allow to perform fits to the data sample in a statistically background subtracted way. In all dimensions of the *sPlot* fit, the shape of signal and background contributions needs to be well described. The *sWeights* are computed based on the resulting yields N of the *sPlot* fit as

$$_sW_n(y_e) = \frac{\sum_{j=1}^{N_s} V_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}. \tag{5.6}$$

Here, f is the PDF of the fit, y_e denotes the set of discriminating variables for an event e, while j and k sum over the N_s categories. Moreover, V is the covariance matrix of the yields that is evaluated in a separate fit, in which all shape related parameters are fixed. A sum over all *sWeights* of one category returns the yield of this category. Furthermore, the uncertainty of the content of bin i in an *sWeighted* histogram is calculated by

$$\sigma(i) = \sqrt{\sum_{e \subset i} (_sW_n)^2}. \tag{5.7}$$

The *sWeights* can only be used in fits to dimensions which are not correlated to the dimensions of the *sPlot* fit. In the context of this thesis, *sWeights* are used to separate signal and background for the training of multivariate selections and to perform fits to the decay time distribution. A proper corresponding choice for the *sPlot* dimensions are invariant masses. Apart from being independent from the decay time, invariant masses have the advantage that there is often a reliable parameterisation for both background and signal components. In case of the flavour tagging calibration with $B^0 \to J/\psi K^{*0}$ (see Sec. 6.6), the reconstructed B^0 mass is used. For the measurement

of *CP* violation in $B_s^0 \rightarrow D_s^\mp K^\pm$ the reconstructed masses of the B_s^0 and the D_s meson are utilised as *sPlot* dimensions (see Chap. 8). Since the hadronic $B_s^0 \rightarrow D_s^\mp K^\pm$ decay has a comparable small branching fraction, adding the PID distribution as a third dimension of the *sPlot* fit improves the disentanglement of background decays with different bachelor particles such as pions or protons.

5.3 Decision trees and boosting

A decision tree [120] without boosting is a simple multivariate classifier. It partitions data consisting of candidates with a feature set of variables $(V_1, V_2, ...)$, by repeatedly performing decisions on these variables, *i.e.* checking if they are below or above certain thresholds $(a, b, ..)$. Figure 5.1 illustrates a decision tree with a depth of two, *i.e.* that concludes two consecutive decisions as a flowchart. The tree starts in its
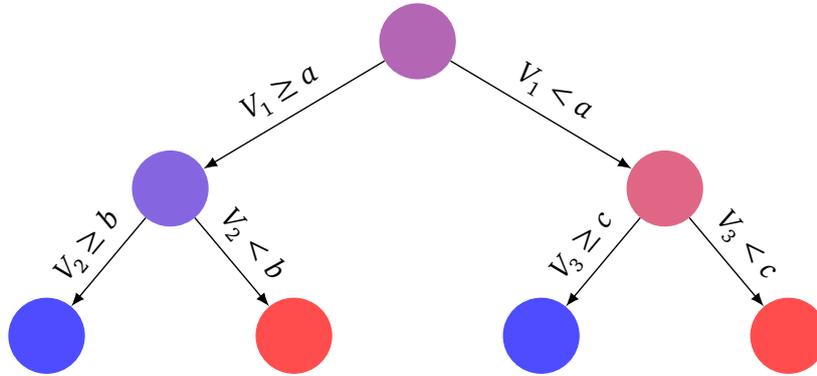
Figure 5.1: Float chart of a simple decision tree that classifies data into two classes, represented by the colours blue and red. The growing purity is illustrated by the increase in colour saturation of the nodes.

root node with the first decision, ideally using the variable with the best separation power (here V_1). Subsequently, more decisions are made using different variables and thresholds. An estimate of the separation power is given by the Gini index [121], which is defined as $p \cdot (1-p)$. Here, p is the signal fraction in the full data sample, the so-called signal purity. Correlations between variables are exploited as the selection of variables and thresholds is based on the previous decisions. The decision making comes to an end as soon as certain criteria are fulfilled, for example if the ratio of candidates that are sorted into a node reaches a certain lower limit. Per definition, all candidates in such a final leaf are classified as belonging to one class of data, *e.g.* as being signal or background candidates. Before they are applied to the real data, decision trees are trained with labelled data, *i.e.* data in which the class of each candidate is known, in order to achieve a useful classification. Final leafs are labelled based on the class that the majority of their associated training candidates belongs to.

An improvement of the classification power can be achieved by training and combining many decision trees with a boosting method resulting in a BDT. The boosting applies weights to wrongly classified candidates, which gives them a higher impact in the next iteration of the training. Several boosting methods exist, like the AdaBoost method [122] or the gradient boosting technique [123]. The latter is the boosting method used in the BDT that is developed in the course of this thesis. Given an input vector of features x, gradient boosting minimises a loss function $L(y, F(x))$, which describes the discrepancy between the classification model $F(x)$ and the true classification y. The algorithm starts with a constant function $F_0(x)$ and expands it incrementally. At each training step m, the model looks like

$$F_m(x) = F_{m-1}(x) + y_m h_m(x), \tag{5.8}$$

where

$$y_m = \arg\min_y \sum_{i=1}^n L\left(y_i, F_{m-1}(x_i) + y h_m(x_i)\right). \tag{5.9}$$

Here, n is the number of training candidates and $h_m(x)$ denotes the output of a single decision tree, which is either 1 (signal) or -1 (background) for each candidate. In the end, a BDT classifier output $F(x)$ is computed as a real number between 1 and -1. The closer it gets to 1 or -1, the safer it is to assume that the candidate is a signal or background candidate, respectively. Gradient boosting of decision trees is further improved by Friedman's TreeBoost algorithm, which is described in Ref. [123].

The input features are ranked by their relevance during the training. These importances are estimated by counting how often a feature is used to split decision tree nodes, where each split occurrence is weighted by the separation gain-squared it has accomplished and by the number of events in the node.

Training data sets are always only templates for the real data that a multivariate algorithm is developed for. For example, differences between simulated and real data can cause bad performance. Additionally, there is the risk that an MVA classifies data due to unphysical statistical fluctuations in the training data. This effect, which is known as overtraining, leads to an overestimation of the performance of a MVA. A way to reduce the risk of overtraining is to use large training data sets and to make sure that simulated data used in the training looks as much as real data as possible. Apart from increasing the quality of the classification, boosting is another way to increase the stability of the MVA. However, it is important to check if a MVA suffers from overtraining. A common test for overtraining is conducted as follows. First, the training data is split randomly before the training, which leads to a training data set and a testing data set. After using the former in the training, the latter can be used to compare the classifier response on an independent sample. It can be assumed that there is no overtraining, if the two classifier distributions agree with each other.

There are a couple of software frameworks that provide implementations of MVAs. Here, the TMVA framework is used [124], while scikit-learn [125] is another example. Moreover, the XGBoost tree boosting system [126] exists, which provides an alternative implementation of gradient boosted decision trees.

6 Flavour Tagging

The aim of the presented analysis is the determination of the CKM angle γ from the *CP* parameters in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays, which themselves are extracted from a decay-time-dependent measurement of the decay rates introduced in Eqs. 3.43 to 3.46. Necessary for such a measurement is knowledge of the initial B meson flavour of the reconstructed signal decay, *i.e.* whether the signal B meson contained a b or a $\bar{b}$ quark. For charged B mesons this can be reconstructed from the final state of their decays. However, this is not possible for decays of neutral B_s^0 and B^0 mesons, as they are able to oscillate (see Sec. 3.3). Even in flavour-specific decays, *i.e.* where $B_{(s)}^0$ decays only into f and $\bar{B}_{(s)}^0$ into $\bar{f}$, it is not possible to reconstruct the initial flavour, but only the flavour at the time of the decay. Thus, the initial flavour is deduced by a process called flavour tagging. The principle of this technique differs fundamentally between LHCb and the B factories [51].

At a B factory, electrons and positrons are collided, mostly at centre-of-mass energies identical to the mass of the $\Upsilon(4S)$ resonance. This resonance decays dominantly into a quantum-correlated system of one particle and one antiparticle B meson state. Thus, the production flavour of the signal meson can then be determined from reconstructing the final state of the second meson, called the tagging meson, in case it decayed flavour-specifically. Furthermore, the decay time of the signal meson is determined relatively to that of the tagging meson. Hence, it is also possible to measure negative decay times. However, the LHCb detector is operated at a *pp* collider, where B mesons are not produced in a quantum-correlated pair. Therefore, a different strategy for the flavour tagging is required.

At LHCb, the initial flavour is deduced by a variety of algorithms that are divided into two groups [127, 128]. These groups are the same-side (SS) and the opposite-side (OS) tagging algorithms. In principle, all algorithms search for charged particles associated with the signal b hadron to infer its production flavour. The two classes are motivated by the origin of the so-called tagging particles, which is illustrated in Fig. 6.1. Same-side taggers, introduced in Sec. 6.1, make use of charged hadronisation remnants of the signal B meson. OS taggers, discussed in Sec. 6.2, exploit the paired b quark production at the LHC and reconstruct particles from the decay chain of the non-signal b quark, *i.e.* the one that does not formed the signal B meson.

Section 6.3 introduces the two per-event output observables that are common for all tagging algorithms, which are the tagging decision (d) and the estimated probability of the decision to be wrong (η). Additionally, the section discusses how the performance of tagging algorithms is evaluated. The overall tagging performance of an analysis is increased by using information from more than one tagger. Common
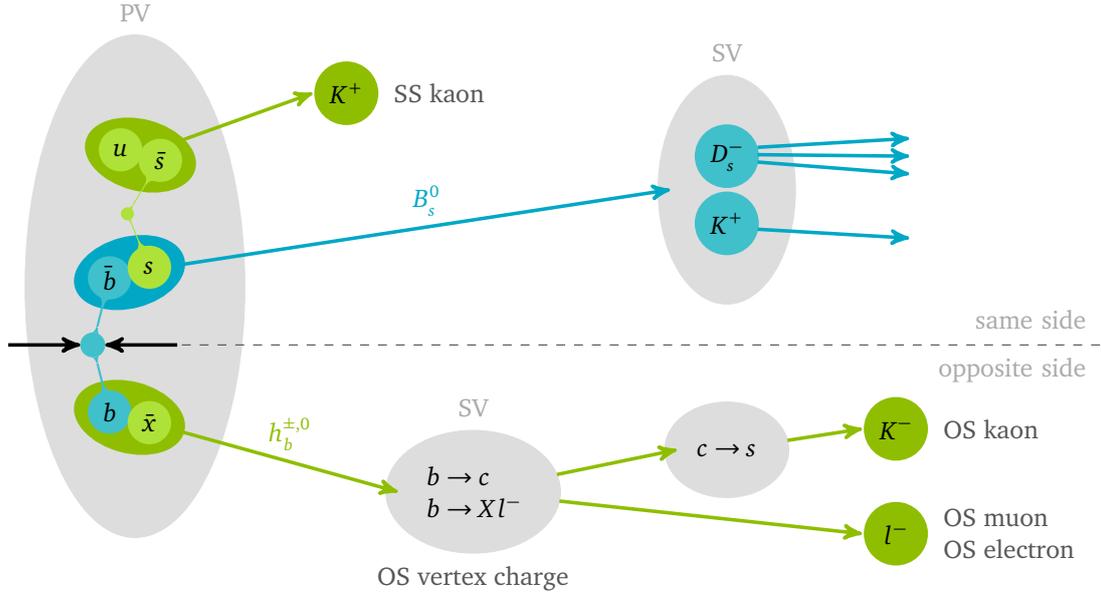
Figure 6.1: Illustration of the flavour tagging principle at the LHCb experiment. The scheme shows only the flavour tagging algorithms used in the presented analysis.

ways to combine taggers are discussed in Sec. 6.4. Afterwards, Sec. 6.5 explains the data-driven calibration of η in general, while Sec. 6.6 discusses flavour-tagging calibrations with $B^0 \to J/\psi K^{*0}$ decays in particular. The latter section also contains two specific examples of calibrations, which are contributions to the *CP* violation measurement in $B^0 \to J/\psi K^0_S$ [68] (see Sec. 6.6.2) and the development of the OS charm tagger [129] (see Sec. 6.6.3). For the presented analysis, the calibration using $B^0_s \to D^-_s \pi^+$ decays is summarised later in Sec. 9.5.

6.1 Same-side tagging

Same-side taggers exploit the fact that not only b quarks, but all quarks are produced in pairs of a quark and an antiquark. Thus, in the case of a B^0_s meson, there is a second s quark that can join with an u quark to build a charged kaon. Such kaons are the tagging particles for the SS kaon tagger, because their charges are directly correlated to the initial flavour of the B^0_s mesons. As illustrated in the top of Fig. 6.1, a positively charged kaon indicates a produced B^0_s meson. The SS kaon tagger cannot be used for B^0 mesons, as they consist of a $\overline{b}$ and a d quark. Instead, two different same-side taggers exist to tag B^0 mesons. These taggers are the SS pion (sometimes abbreviated SSπ) and the SS proton taggers, which follow the same principle as the SS kaon tagger used for B^0_s mesons by searching for pions and respectively protons that contain the second d quark.

Two SS kaon taggers are implemented in the LHCb software, which pursue different strategies for the selection of the tagging kaon. The initially developed version uses

rectangular cuts to select the tagging particle, while the newer one utilises a neural net selection. This change resulted in a significant improvement of the SS kaon tagging efficiency [130]. Hence, the presented analysis uses only the newer SS kaon algorithm. The training of the multivariate algorithm was performed on $B_s^0 \rightarrow D_s^- \pi^+$ decays using input variables such as track fit qualities, polar angles, momenta, transverse momenta and PID requirements. Wrong tags mostly arise if a kaon from the opposite side or from underlying events is mistakenly selected. In order to reduce this source of mistags, the angle ϕ between the B_s^0 meson and the potential tagging kaon is also used in the neural net. More details on the development, optimisation and implementation of the SS kaon tagger can be found in Ref. [130].

6.2 Opposite-side tagging

The opposite-side taggers that are used in the presented analysis are illustrated in the bottom half of Fig. 6.1. Unlike SS taggers, all available OS taggers can be used to tag B_s^0 and B^0 mesons, as they exploit the second b quark produced in the initial pp collision, *i.e.* the one that does not form the signal b hadron. The tag decision is based on the charged decay remnants of the b hadron formed by the second b quark. There are a range of opposite-side tagging algorithms, which search for different charged particles to infer the tagging information. In the presented analysis, the standard opposite-side tagger combination is used, which is comprised out of the four most established OS taggers. These taggers, which are the OS kaon, the OS muon, the OS electron, and the OS vertex charge, are briefly introduced in the next paragraphs. Details on the combination process are given in Sec. 6.4. The fifth available OS tagging algorithm is the OS charm tagger, which was recently developed [129]. Due to the time scale it is not used in the presented analysis, but it is discussed in Sec. 6.6.3 in the context of its calibration with $B^0 \rightarrow J/\psi K^{*0}$ decays.

The tagging particle of the OS kaon tagger is a charged kaon originating from a $b \rightarrow c \rightarrow s$ transition. A negatively (positively) charged kaon that survives the selection, which includes requirements on PID, transverse momentum and impact parameter, indicates a signal $\bar{b}$ (b) quark. If more than one kaon candidate passes the selection, the candidate with the highest p_{T} is chosen. Main backgrounds are prompt kaons and those from primary $\bar{b} \rightarrow \bar{c}W^+(\rightarrow c\bar{s})$ transitions, which lead to reconstructing the wrong charge.

The OS muon and OS electron taggers exploit semi-leptonic decays on the opposite-side. If the tagging b hadron decays via the transition $b \rightarrow cW^-(\rightarrow l^-\bar{\nu})$, the negative (positive) charge of the lepton indicates an initial signal $\bar{b}$ (b) quark. The leptons are identified with cuts on PID and transverse momentum, effectively reducing leptons originating from secondary charm meson decays. Furthermore, clone tracks are suppressed and a sufficient track fit quality for muons is required. Electron tracks have to lay in the HCAL acceptance and need to deposit a substantial amount of energy in the ECAL with respect to their momentum. Moreover, a maximal deposition of ionisation charge in the VELO is defined. This requirement suppresses the conversion

of photons close to the primary interaction region. Both the OS muon and the OS electron tagger base their tag decision on the charge of the respective candidate with the highest p_{T}.

The OS vertex charge tagger does not use a single particle to deduce its tagging decision, but tries to fully reconstruct the decay vertex of the opposite-side b hadron. Then, the charges of all particles belonging to the vertex are summed up. In case of charged opposite-side b hadrons, the effective charge gives the correct tagging decision. A negative (positive) charge indicates an initial $\bar{b}$ (b) in the signal b hadron. Wrong tagging decisions can originate from mixing up the opposite-side b hadron vertex with the one of a secondary charm meson, *i.e.* a charm meson from a B decay. In order to reduce this type of background, a good separation from the PV is required. Furthermore, decays of long-lived K_s^0 mesons are suppressed by starting the reconstruction with a seed algorithm that ignores all two particle combinations compatible with the mass of a K_s^0 meson. Additionally, all particles have to fulfil requirements on their momentum p and transverse momentum p_{T}. The summed charge of the vertex Q_{vtx} is weighted by $p_{\mathrm{T}}^{0.4}$ and a tagging decision is only taken, if $|Q_{\mathrm{vtx}}| \geq 0.275$.

An intrinsic source of incorrect tags for all OS taggers occurs if the opposite-side b hadron is a neutral B meson that oscillates into its antiparticle state before it decays. In this case, OS taggers deduce the wrong tagging information, even if the decay remnants are correctly reconstructed.

6.3 Output and performance of tagging algorithms

Common to all tagging algorithms is the output of two per-event observables, which are the tag decision d and an estimated mistag probability η. The former has a response of 1 if an initial $\bar{b}$ quark is identified, and -1 in case of an initial b quark. Furthermore, the tag decision is 0 if no tag decision could be derived, because *e.g.* no tagging particle meets the tagger's selection criteria. The second per-event output η gives a probability estimate of the tag to be incorrectly assigned and is defined in the range between 0 and 0.5. A response of 0 represents no uncertainty on the tag decision, while 0.5 implies a candidate remains untagged, *i.e.* $\eta = 0.5$ is connected to a tag decision of $d = 0$. A few individual reasons why incorrect tag decisions occur were discussed in the previous sections about same- and opposite-side tagging. Details on how the mistag probability estimate η is calibrated to obtain the true mistag probability $\omega(\eta)$ are given in the following section.

In order to estimate the performance of a tagging algorithm, two quantities are important: the tagging efficiency ϵ_{tag} and the effective tagging efficiency ϵ_{eff}. The former is the ratio of tagged candidates over all reconstructed candidates, given by

$$\epsilon_{\mathrm{tag}} = \frac{N_{\mathrm{r}} + N_{\mathrm{w}}}{N_{\mathrm{r}} + N_{\mathrm{w}} + N_{\mathrm{u}}} \,. \tag{6.1}$$

Here, N_{r} and N_{w} are the numbers of correctly and incorrectly tagged candidates, respectively. Furthermore, N_{u} denotes the amount of events without a tag. In order to

derive ϵ_{eff}, it is useful to define the true mistag probability as the fraction of candidates with wrong tags over all tagged candidates

$$\omega = \frac{N_{\text{w}}}{N_{\text{r}} + N_{\text{w}}}. \tag{6.2}$$

By using ω the tagged decay rates Γ^{tagged} can be related to the true decay rates Γ as

$$\Gamma^{\text{tagged}}(\bar{B}_q^0(t) \to f)(t) = (1-\omega)\,\Gamma(\bar{B}_q^0(t) \to f)(t) + \omega\,\Gamma(B_q^0(t) \to f)(t) \tag{6.3}$$

$$\Gamma^{\text{tagged}}(B_q^0(t) \to f)(t) = (1-\omega)\,\Gamma(B_q^0(t) \to f)(t) + \omega\,\Gamma(\bar{B}_q^0(t) \to f)(t). \tag{6.4}$$

This relation is the origin of a tagging induced dilution D_ω of all measured tagged time-dependent asymmetries like the *CP* asymmetry introduced in Eq. 3.58

$$\mathscr{A}_f^{\text{meas}}(t) = (1-2\omega)\mathscr{A}_f(t) = D_\omega \mathscr{A}_f(t) \tag{6.5}$$

The effective tagging efficiency can be defined as

$$\epsilon_{\text{eff}} = \epsilon_{\text{tag}}(1-2\omega)^2. \tag{6.6}$$

For any tagging algorithm ϵ_{eff} is computed on a per-event basis i, as

$$\epsilon_{\text{eff}} = \frac{1}{N} \sum_i^N (1-2\omega)^2. \tag{6.7}$$

A perfectly working tagger would always yield a mistag probability of 0 and assign a tag to all candidates, hence $\epsilon_{\text{eff}} = 100\%$. This cannot be achieved in reality, thus, $\epsilon_{\text{eff}}N$ represents the proportion of perfectly tagged candidates over all reconstructed candidates. This proportion of candidates would give the same statistical sensitivity σ_{CP} on the *CP* parameters as the full data set with imperfect tagging, therefore

$$\sigma_{CP} \propto \frac{1}{\sqrt{\epsilon_{\text{eff}}N}}. \tag{6.8}$$

6.4 Combination of tagging algorithms

The presented analysis uses the only available SS tagger for B_s^0 mesons, the SS kaon tagger, as well as the most established OS taggers. However, the latter are not utilised as individual inputs, but in terms of combined OS tagging decisions d and mistag probability estimates η. These outputs are determined by calculating

$$P(b) = \frac{p(b)}{p(b) + p(\bar{b})} \quad \text{and} \quad P(\bar{b}) = 1 - P(b), \tag{6.9}$$

which are the probabilities that the initial B meson contained a b or a $\bar{b}$ quark, respectively. The individual outputs of all OS taggers i that provided a non-zero tag decision enter in

$$p(b) = \prod_i \left(\frac{1 + d_i}{2} - d_i(1 - \eta_i) \right) \quad \text{and} \quad p(\bar{b}) = \prod_i \left(\frac{1 - d_i}{2} + d_i(1 - \eta_i) \right). \quad (6.10)$$

The combined outputs are obtained by comparing the probabilities from Eq. 6.9. If $P(b)$ is larger than $P(\bar{b})$, it is implied that the probability for an initial b is higher, resulting in a combined tag of $d = -1$ and a mistag probability of $\eta = 1 - P(b)$. Otherwise, the combined outputs are $d = 1$ and $\eta = 1 - P(\bar{b})$. The contribution of tags with poor tagging power is reduced, by requiring the mistag probabilities of the OS kaon and the OS vertex charge tagger to be below 0.46. An overview on the performances of the OS tagging algorithms is given in Tab. 6.1.

Table 6.1: Performances of the OS tagging algorithms in Run 1. The values, which include only statistical uncertainties, are determined on the control channel $B^+ \to J/\psi K^+$ [131].

Tagger	ϵ_{tag} [%]	ω [%]	ϵ_{eff} [%]
OS muon	4.8 ± 0.1	29.9 ± 0.7	0.77 ± 0.07
OS electron	2.2 ± 0.1	33.2 ± 1.1	0.25 ± 0.04
OS kaon	11.6 ± 0.1	38.3 ± 0.5	0.63 ± 0.05
OS vertex charge	15.1 ± 0.1	40.0 ± 0.4	0.60 ± 0.06
OS combination	27.3 ± 0.2	36.2 ± 0.5	2.07 ± 0.11

6.5 Calibration of flavour-tagging algorithms

The per-event mistag probabilities η computed by the neural nets of the tagging algorithms only represent predictions. Additionally, in case of the OS tagger combination, correlations between the OS taggers can lead to an overestimation of the combined probability. Therefore, mistag correction functions are derived, allowing to transform the estimated mistag probability η into the true mistag probability $\omega(\eta)$. This relation is usually sufficiently described by the linear function

$$\omega(\eta) = p_0 + p_1(\eta - \langle \eta \rangle). \quad (6.11)$$

The correlation between the calibration parameters p_0 and p_1 is reduced by shifting the calibration function by the value of the average mistag estimate $\langle \eta \rangle$ that is observed for the signal candidates. In case of a perfect calibration $p_0 = \langle \eta \rangle$ and $p_1 = 1$. Different interaction rates of the tagging particles with the detector material and detection asymmetries can lead to a dependence between the tagging parameters and

the initial flavour of the B meson. Thus, ω is defined as the mean of the mistag probabilities for initial B and $\bar{B}$ mesons

$$\omega = \frac{\omega^B + \omega^{\bar{B}}}{2} \tag{6.12}$$

and separate calibration functions for initial B and $\bar{B}$ mesons are introduced as

$$\omega^B(\eta) = p_0^B + p_1^B(\eta - \langle\eta\rangle) \quad \text{and} \quad \omega^{\bar{B}}(\eta) = p_0^{\bar{B}} + p_1^{\bar{B}}(\eta - \langle\eta\rangle). \tag{6.13}$$

The individual calibration parameters of initial B and $\bar{B}$ mesons are connected in terms of their means p_i and differences Δp_i as

$$p_i = \frac{p_i^B + p_i^{\bar{B}}}{2}, \quad \Delta p_i = p_i^B - p_i^{\bar{B}}, \quad \text{where } i = 0, 1. \tag{6.14}$$

Based on the calibration parameter differences, the mistag asymmetry is given by

$$\Delta\omega(\eta) = \omega^B(\eta) - \omega^{\bar{B}}(\eta) = \Delta p_0 + \Delta p_1(\eta - \langle\eta\rangle). \tag{6.15}$$

An asymmetry of the tagging efficiency is taken into account equivalently by introducing individual efficiencies for B (ϵ_{tag}^B) and $\bar{B}$ ($\epsilon_{\text{tag}}^{\bar{B}}$), as well as a corresponding mean tagging efficiency ϵ_{tag} and difference $\Delta\epsilon_{\text{tag}}$.

The calibration parameters cannot be determined from simulation, as the flavour tagging depends on kinematic and particle identification observables that are not modelled perfectly in simulations. Instead, calibrations are derived from data by either exploiting the self-tagging nature of charged B mesons decays, or utilising flavour-specific decays of neutral B mesons. Charged B decays, like $B^+ \to J/\psi K^+$ or $B^+ \to D^0 \pi^+$, allow to simply count the number of correctly and incorrectly tagged decays by comparing the charge of the final state with the tagged initial state. Tagging calibrations that are performed with neutral B decays require to take the mixing of B_s^0 and B^0 mesons into account (see Sec. 3.3), for example by performing decay-time-dependent analyses. The following section introduces how the flavour-tagging is calibrated using $B^0 \to J/\psi K^{*0}$ decays and gives examples. For the presented analysis, the calibration is performed with $B_s^0 \to D_s^- \pi^+$ decays (see Sec. 9.5). The choice of the calibration channel depends on the physics channel. Preferably, kinematically similar channels are used, in order to reduce systematic uncertainties from the flavour tagging. Additionally, these channels should feature large branching fractions and low background levels.

6.6 Flavour-tagging calibrations with $B^0 \to J/\psi K^{*0}$ decays

The flavour-specific charmonium mode $B^0 \to J/\psi K^{*0}$ is a commonly used channel for flavour-tagging studies, such as calibrations. This section introduces characteristics of the decay mode as well as the selection that was developed in order to separate signal candidates of the decay from background. Afterwards, the general strategy of flavour-tagging calibrations with this decay channel is outlined. As an example, the calibration of the OS tagger combination for the full Run 1 dataset is discussed in Sec. 6.6.1. Subsequently, the calibrations of the SS pion and the OS charm tagger are discussed in the Secs. 6.6.2 and 6.6.3, respectively. The latter two calibrations are both direct inputs to publications of the LHCb collaboration [68, 129]. Throughout this section, the $B^0 \to J/\psi K^{*0}$ data samples are reconstructed with stripping v20r1 (2011) and v20r0 (2012) conditions.

Candidates of the $B^0 \to J/\psi K^{*0}$ decay are reconstructed in the $J/\psi \to \mu^+\mu^-$ and $K^{*0} \to K^+\pi^-$ final states. The charge configuration of the latter decay allows to identify the B^0 meson flavour at the time of decay, as B^0 ($\overline{B}{}^0$) mesons decay only into the $J/\psi K^{*0}$ ($J/\psi \overline{K}{}^{*0}$) final state and because the K^{*0} resonance decays flavour specifically as well. Hence, a $K^{*0} \to K^+\pi^-$ decay indicates a B^0 meson decay, while $\overline{K}{}^{*0} \to K^-\pi^+$ is related to the decay of a $\overline{B}{}^0$ meson. The $J/\psi \to \mu^+\mu^-$ final state is chosen instead of $J/\psi \to e^+e^-$ or $J/\psi \to K^+K^-\pi^+\pi^-$, as it can be triggered very well by the LHCb experiment, which allows much lower background levels. All $B^0 \to J/\psi K^{*0}$ candidates need to be triggered by the HLT1 line `Hlt1DiMuonHighMass` and the HLT2 line `Hlt2DiMuonDetachedJPsi`. Details on the requirements of these trigger lines can be found in Ref. [132].

The utilised stripping line (called `StrippingBetaSBd2JpsiKstarDetachedLine`) applies loose requirements on the quality of the J/ψ, the K^{*0}, and the B^0_s vertices. Additionally, the line demands the invariant masses of the two muons and the $K^+\pi^-$ combination to be roughly consistent with the known J/ψ and K^{*0} masses [133], respectively.

To further reduce background contributions, an offline selection is applied after the trigger and stripping selections. The focus of the offline selection is to maintain a high signal efficiency of $> 90\,\%$. Firstly, the invariant $J/\psi K^{*0}$ mass is reduced to 5230–5330 MeV/c^2. Secondly, the B^0_s decay time is restricted to 0.3–15.3 ps, which reduces contributions from prompt background. The final step of the offline selection suppresses combinatorial background by making use of a multivariate selection based on a boosted decision tree (BDT). An introduction to boosted decision trees can be found in Sec. 5.3. The BDT training is performed separately for the 2011 and 2012 subsamples of the Run 1 data. Both trainings follow the same general strategy of using weighted data from the respective year of data taking as proxies for the signal and background components. The weights are computed by utilising the *sPlot* technique (see Sec 5.2) in fits to the invariant $J/\psi K^{*0}$ mass distributions. For both, 2011 and 2012 data, the fit to the invariant mass is performed with a PDF that con-

sists of two components, the signal and combinatorial background. While the former is described by a Hypatia PDF [134], an exponential function is used for the latter. A Hypatia function is composed of a generalised hyperbolic core and a power-law tail to each side. The slope of the exponential and the mean of the Hypatia are the only floating parameters in the fit. All further parameters of the Hypatia are fixed to values determined on simulated data. Figure 6.2 shows the *sPlot* fits to the reconstructed $J/\psi K^{*0}$ mass. The pull distributions, which are shown underneath the invariant mass distributions, confirm that the data is well described by the fit. Here, a pull is defined as the difference between a data point and the functional value of the PDF at the same spot, divided by the uncertainty on the data point. Thus, the pull gives the numbers of sigma, that the PDF and the data deviate. In case of a good fit, the pulls follow a standard normal distribution. Prior to the training, the weighed data is split randomly into two halves. One half is used for the training and the other for the testing of the trained classifier. The BDT training input variables, which are listed in Tab. 6.2, are selected to deliver large separation power between signal and background. While the muon PIDs are used as inputs of the BDT training for the 2011 data, they are removed in the 2012 training, as a worse performance is observed when they are included. Figure 6.3 shows the signal and background distributions of the 2011 and 2012 classifiers, separately for the training and testing samples. No signs of overtraining are observed, as the distribution of the training and testing samples are compatible. In Fig. 6.4 the receiver operating curves of the two BDTs are illustrated. The latter figures roughly illustrate that both BDTs enable
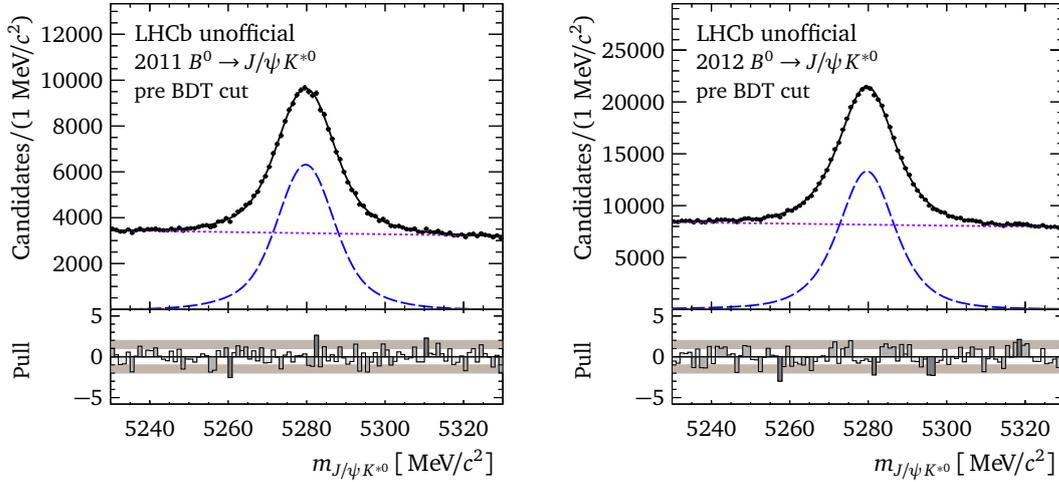
Figure 6.2: Invariant $J/\psi K^{*0}$ mass distributions of the 2011 (left) and 2012 (right) data after applying all selection steps prior to the BDT selection. Both plots contain fit projections of the fits that are used to calculate the *sWeights* for the BDT trainings. The solid black curve is the total fit PDF, while the $B^0 \to J/\psi K^{*0}$ signal and the combinatorial background are plotted as dashed blue and shortly-dashed purple lines, respectively.

Table 6.2: Input variables of the 2011 and 2012 $B^0 \to J/\psi K^{*0}$ BDT training.

Variable	Variable
B^0:	K^{*0}:
$\quad p_{\mathrm{T}}$	$\quad p_{\mathrm{T}}$
$\quad$ endvertex χ^2	$\quad$ endvertex χ^2
	$\quad$ invariant $K^+\pi^-$ mass
J/ψ:	K^+:
$\quad$ invariant $\mu^+\mu^-$ mass	$\quad$ PIDK
$\quad$ error of invariant $\mu^+\mu^-$ mass	$\quad$ PIDp
	$\quad$ track χ^2/ndof
both muons:	$\quad$ track ghost prob
$\quad p_{\mathrm{T}}$	π^-:
$\quad$ track χ^2/ndof	$\quad$ track χ^2/ndof
$\quad$ track ghost prob	$\quad$ track ghost prob
$\quad$ PIDμ (only 2011)	

a background suppression of 90 % while maintaining about 90 % of the signal candidates. The cut points of the BDT classifiers are chosen to maintain a high signal yield by optimising with respect to the signal significance $S/(\sqrt{S+B})$, where S (B) is the number of signal (background) candidates. In this way, the cut values for the BDT classifiers are chosen to be about 0.1 for both the 2011 and 2012 classifiers. Figure 6.5 shows the invariant $J/\psi K^{*0}$ mass distributions for the 2011 and 2012 data after all selection steps, including the BDT classifier cuts, have been applied. From this point on the 2011 and 2012 datasets are combined to perform calibrations on the full Run 1 data.

After the selection of the signal candidates is finished, the calibration of the flavour tagging parameters is performed. In contrast to calibrations with decays of the
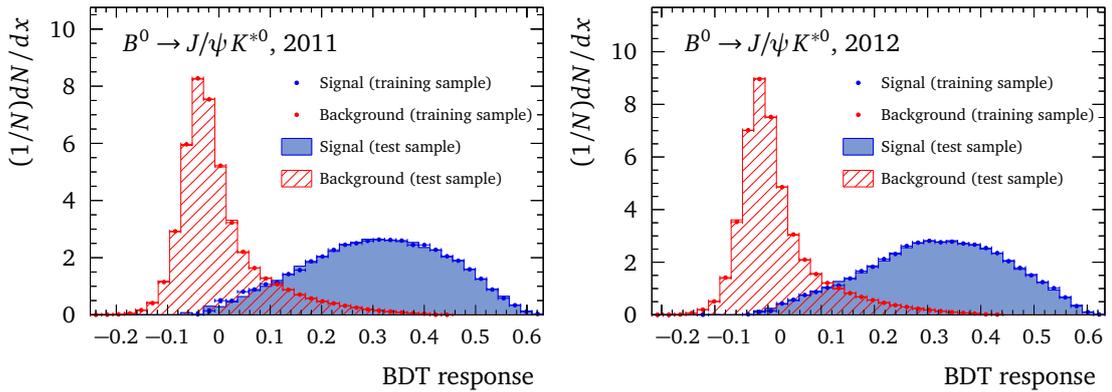
Figure 6.3: The signal and background BDT responses of the $B^0 \to J/\psi K^{*0}$ BDT, superimposed are the training and the testing data.
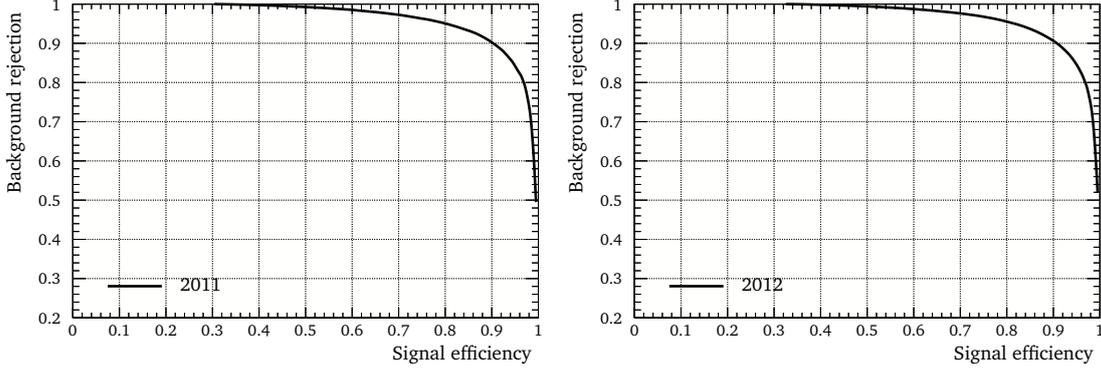
**Figure 6.4:** Receiver operating curves of the BDTs trained for $B^0 \to J/\psi K^{*0}$.
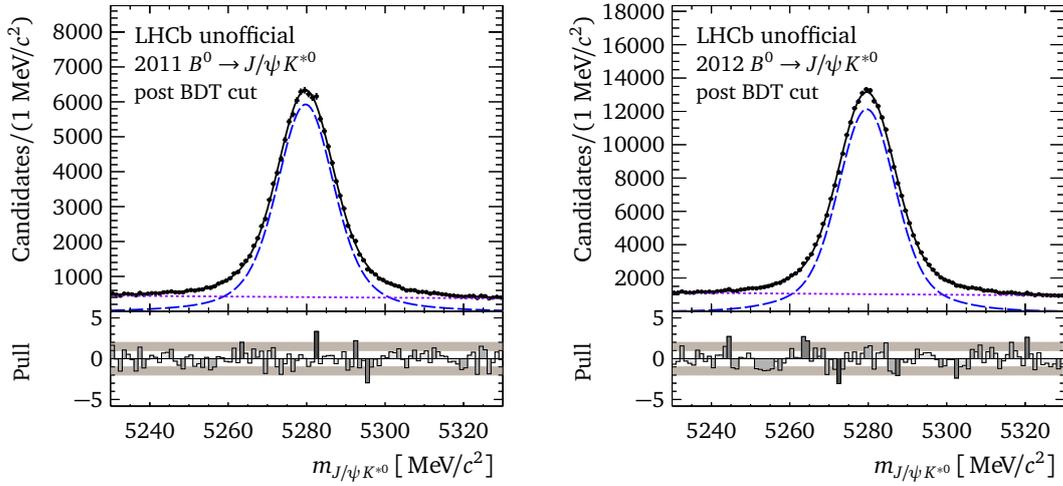
Figure 6.5: Invariant $J/\psi K^{*0}$ mass distribution of the 2011 (left) and 2012 (right) data after applying all selection steps including the BDT selection. The solid black curve is the total fit PDF, while the $B^0 \to J/\psi K^{*0}$ signal and the combinatorial background are plotted as dashed blue and shortly-dashed purple lines, respectively.

charged B^+ meson, it is not possible to count wrongly tagged candidates by comparing the final state charge with the tag decision as the neutral B^0 mesons are able to mix before their decay. Thus, a decay-time-dependent analysis is performed. Due to the flavour-specific nature of the $B^0 \to J/\psi K^{*0}$ decay, there is no time-dependent *CP* asymmetry present (compare Eq. 3.57). Instead, it is possible to measure a raw time-dependent mixing asymmetry $\mathscr{A}_{\mathrm{mix}}$, which is defined as

$$\mathscr{A}_{\mathrm{mix}}(t) = \frac{N_{\mathrm{unmixed}}(t) - N_{\mathrm{mixed}}(t)}{N_{\mathrm{unmixed}}(t) + N_{\mathrm{mixed}}(t)} = \cos(\Delta m_d t). \tag{6.16}$$

Here, N_{mixed} and N_{unmixed} are the decay-time-dependent numbers of candidates that have and have not oscillated before their decay, respectively. The B^0 mixing frequency is denoted by Δm_d. The theoretical mixing amplitude is reduced by the same tagging

induced dilution factor that was introduced in Eq. 6.5, hence

$$\mathscr{A}_{\text{mix}}^{\text{meas}}(t) = (1 - 2\omega)\cos(\Delta m_d t). \tag{6.17}$$

Therefore, the tagging algorithm output can be calibrated by performing a decay-time-dependent mixing analysis and interpreting the amplitude as $(1 - 2\omega(\eta))$. As there is still combinatorial background present in the data (compare Fig. 6.5), a two dimensional fit of the decay time t and the invariant $J/\psi K^{*0}$ mass is performed, where the latter allows to disentangle signal and remaining background candidates.

The mass dimension is again parameterised by a Hypatia PDF for the signal and an exponential function for the combinatorial background. For the description of the decay time t, the signal PDF is given by

$$P_{\text{sig}}(t, d, q) = \epsilon(t) \cdot \left(M_{\text{tag}}(t, d, q) \otimes R(t - t_{\text{true}} | \sigma_t) \right). \tag{6.18}$$

Its variables and components are introduced in the following. The main part is the B mixing PDF

$$M_{\text{tag}}(t, d, q) \propto e^{-t/\tau} \left(1 - d\Delta\omega + dq(1 - 2\omega)\cos(\Delta m_d t) \right). \tag{6.19}$$

Here, d is the tag decision and q denotes the reconstructed flavour of the final state, *i.e.* it contains the information if a B^0 ($q = 1$) or a $\bar{B}^0$ ($q = -1$) decayed at the time t. Furthermore, ω denotes the arithmetic mean of the mistag probability and $\Delta\omega$ is the difference of $\omega^B(\eta)$ and $\omega^{\bar{B}}(\eta)$ as introduced in Eq. 6.15. The B mixing PDF is derived from the decay-rate coefficients defined in the Eqs. 3.43 to 3.46, taking into account that $B^0 \to J/\psi K^{*0}$ is a flavour-specific decay and the consequential absence of CP violation. Additionally, a decay-width difference $\Delta\Gamma_d = 0$ is implied, which is justified in the B^0 system [57]. The mixing PDF is convolved with a single Gaussian resolution model

$$R(t - t_{\text{true}} | \sigma_t) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{(t - t_{\text{true}})^2}{2\sigma_t^2} \right). \tag{6.20}$$

with a fixed width of $\sigma_t = 50\,\text{fs}$ that is obtained from simulations. Here, t denotes the reconstructed decay time, while t_{true} represents the true decay time. A simple model with a fixed mean resolution instead of a per-event resolution is sufficient as the dilution induced by the decay-time resolution is negligible in the B^0 system (see Sec. 9.3). The final component of the signal PDF is a decay-time-dependent efficiency function

$$\epsilon(t) = \arctan\left(t e^{\alpha t + \beta} \right). \tag{6.21}$$

It describes a decay-time bias of the selection for candidates with low decay times. The parameters α and β are fixed in the fit to values that were determined with simulations. In contrast to the signal component, there is no PDF motivated by physics

that parameterises the decay-time distribution of combinatorial background candidates. Thus, the PDF is chosen phenomenologically by means of simulated data. A sufficient description of the simulations is achieved by the sum of two exponentials, one with a lower and one with a larger lifetime.

The two dimensional fit is performed simultaneously in roughly evenly filled bins of the mistag prediction η. The choice of the number of bins as well as their positions depend on the tagging algorithm. For the SS pion and the OS charm taggers, a stable fit is only possible with as many as 5 bins, while the larger number of tagged candidates allows to use 8 bins in case of the OS tagger combination. The η distributions of the three taggers are shown for $B^0 \to J/\psi K^{*0}$ signal candidates in Fig. 6.6, highlighting the η intervals in different colours. The aim of the calibration is to measure the tagging calibration parameters introduced in Eq. 6.13. Therefore, ω and $\Delta\omega$ are directly parameterised by the calibration functions

$$\omega(\eta) = p_0 + p_1\left(\eta - \langle\eta\rangle_i\right) \quad \text{and} \quad \Delta\omega(\eta) = \Delta p_0 + \Delta p_1\left(\eta - \langle\eta\rangle_i\right), \qquad (6.22)$$

in the B mixing PDF from Eq. 6.19. The mean values of the mistag prediction $\langle\eta\rangle_i$ are fixed in the calibration fit using their respective values of the different bins i. These $\langle\eta\rangle_i$ values are obtained prior to the calibration fit utilising the *sPlot* technique. Apart from the flavour tagging calibration parameters, floating parameters in the fit are the B lifetime, the yield parameters of the η bins, the B oscillation frequency Δm_d, the slopes of the exponential functions describing the background, and the mean parameter of the Hypatia signal mass PDF.
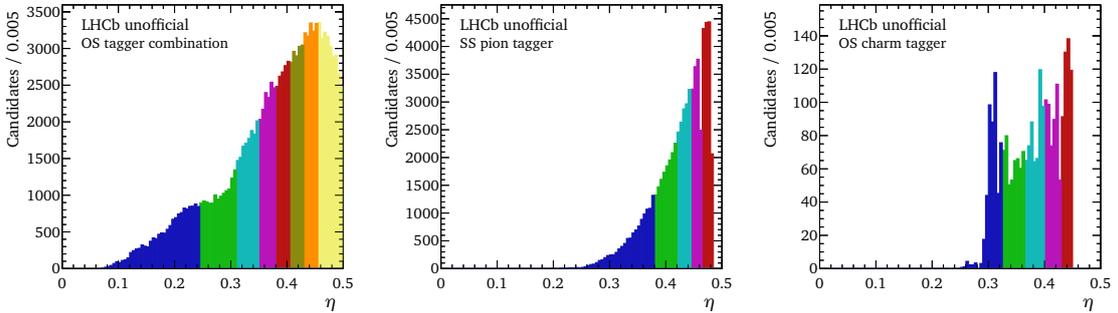
Figure 6.6: Mistag prediction distributions of the selected $B^0 \to J/\psi K^{*0}$ signal candidates for the OS tagger combination (left), the SS pion tagger (center), and the OS charm tagger (right). The colours illustrate the η intervals in which the fit is performed simultaneously.

In general, two types of systematic uncertainties are considered for the flavour tagging calibration parameters. The so-called type I systematic uncertainties are the intrinsic uncertainties of the calibration method, such as parametrisation of signal and background, uncertainties on fixed input parameters, or neglecting asymmetries. In case of the $B^0 \to J/\psi K^{*0}$ calibration, these uncertainties are estimated by means of toy studies, a technique introduced in Sec. 4.5.2. Type II systematic uncertainties

cover kinematic differences between the control mode (here $B^0 \to J/\psi K^{*0}$) and the physics mode which the calibration is applied to. Hence, they are only calculated if a calibration is applied in a specific analysis, like in case of the SS pion tagger calibration used in the *CP* violation measurement with $B^0 \to J/\psi K_S^0$ decays. Before this SS pion calibration is summarised in Sec. 6.6.2, the calibration of the OS tagger combination is introduced in Sec. 6.6.1. Finally, a calibration cross-check of the newly developed OS charm tagger is discussed (see Sec. 6.6.3).

6.6.1 Calibration of the OS tagger combination

This section summarises the results of the OS tagger calibration on the full Run 1 dataset as an example of a calibration with $B^0 \to J/\psi K^{*0}$ decays. About 120 000 selected $B^0 \to J/\psi K^{*0}$ signal candidates in the combined 2011 and 2012 dataset have a tag decision from the OS tagger combination. This corresponds to a tagging efficiency of $\epsilon_{\text{tag}} = (33.13 \pm 0.10)\,\%$. The effective tagging efficiency is measured to be $\epsilon_{\text{eff}} = (2.63 \pm 0.08)\,\%$. The fit is performed in simultaneously eight bins of the mistag estimate. Figure 6.7 contains the invariant mass and the decay-time distributions. Additionally, the figure contains projections of the fitted PDFs. Both, the data distributions and the PDFs, are summed over all η bins in the figures. A projection of the raw mixing asymmetry is shown in the left part of Fig. 6.8. In the right plot of the same figure the measured pairs of ω_i and $\langle \eta \rangle_i$ are plotted together with the linear calibration function. The plot illustrates that a linear calibration model is sufficient to describe the relation.
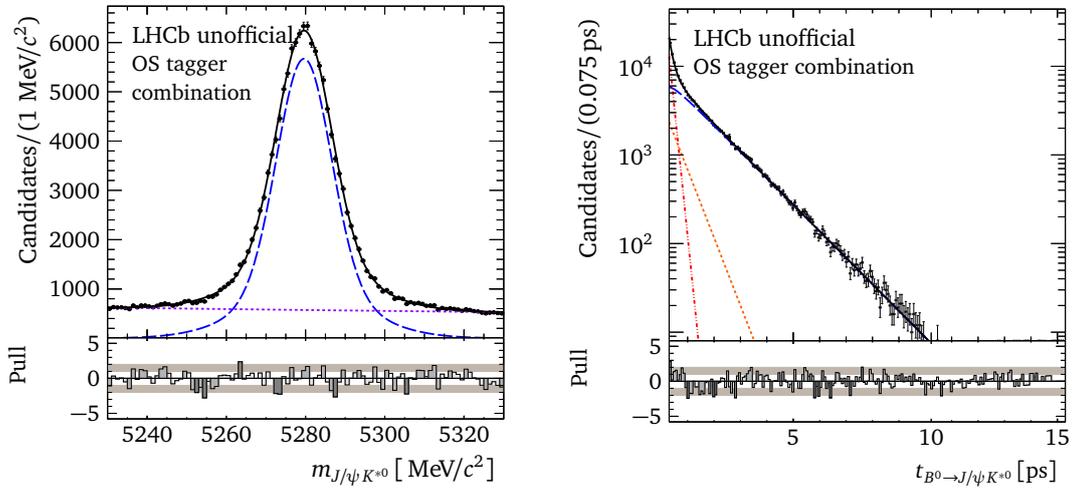
Figure 6.7: The invariant $J/\psi K^{*0}$ mass (left) and the decay-time (right) distributions of the $B^0 \to J/\psi K^{*0}$ candidates tagged by the OS tagger combination. The total PDF is plotted as a black solid line and the signal is shown as dashed blue line. The combinatorial background is denoted by the shortly-dashed purple line in the mass distribution, while it is illustrated in the decay-time distribution by the shortly-dashed orange and dashed-dotted red lines.
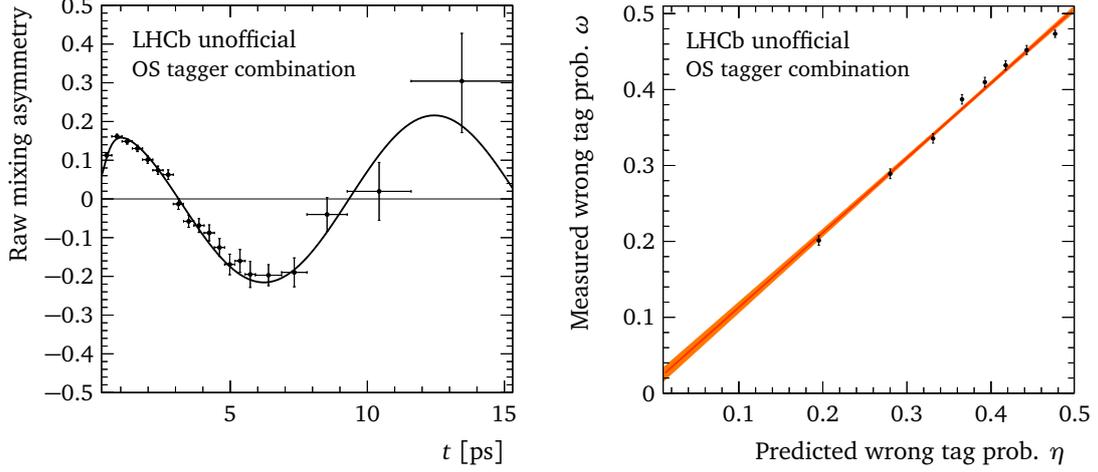
Figure 6.8: The raw mixing asymmetry for all $B^0 \rightarrow J/\psi K^{*0}$ candidates tagged by the combination of OS taggers (left). The calibration of the wrong tag estimate for the OS tagger combination based on the decay-time fit of the $B^0 \rightarrow J/\psi K^{*0}$ candidates (right).

The fit determines the tagging calibration parameters to be

$$
\begin{aligned}
p_0^{OS} &= 0.3872 \pm 0.0021 \text{ (stat.)} \quad \pm 0.0011 \text{ (syst. type I)} , \\
p_1^{OS} &= 0.985 \;\pm 0.024 \;\; \text{(stat.)} \quad \pm 0.012 \;\; \text{(syst. type I)} , \\
\Delta p_0^{OS} &= 0.0210 \pm 0.0031 \text{ (stat.)} \quad \pm 0.0014 \text{ (syst. type I)} , \quad\quad (6.23) \\
\Delta p_1^{OS} &= 0.059 \;\pm 0.035 \;\; \text{(stat.)} \quad \pm 0.016 \;\; \text{(syst. type I)} , \\
\langle \eta^{OS} \rangle &= 0.378 .
\end{aligned}
$$

The p_1 parameter is compatible with unity within $1\,\sigma$, while there is a deviation of about $2.5\,\sigma$ between p_0 and the global mean $\langle \eta \rangle = 0.378$. The correlations between the tagging calibration parameters are small (see Tab. 6.3). Different sources of systematic uncertainties are studied, like the decay-time acceptance and resolution, asymmetries in production and detection, and the choices of the fit models. All systematic uncertainties are estimated by means of toy studies, in which the generated pseudo-experiments and the corresponding fits differ with respect to the potential source of interest. For example, a per-event decay-time resolution is used in the generation, but the nominal average resolution model is used in the fit. This specific test results no significant uncertainty. The largest source of systematic uncertainty is the choice of the acceptance description. It is conservatively estimated in a toy study in which samples generated with the default acceptance model are fit with a PDF without any acceptance model. Overall, the statistical uncertainty remains larger for all parameters, which will allow to improve the precision of flavour-tagging calibrations with $B^0 \rightarrow J/\psi K^{*0}$ using the larger Run 2 data.

The presented results of the OS tagger calibration, especially the value of p_0, show the necessity of applying calibration parameters to the raw tagging output when

Table 6.3: Correlations between the calibration parameters of the OS tagger combination.

	p_0^{OS}	p_1^{OS}	Δp_0^{OS}	Δp_1^{OS}
p_0^{OS}	1	0.07	−0.019	0.004
p_1^{OS}		1	−0.008	−0.004
Δp_0^{OS}			1	−0.0025
Δp_1^{OS}				1

used in physics analyses like *CP* measurements. In case of the OS tagger calibration, the standard calibration is obtained with $B^+ \to J/\psi K^+$ decays. The latter decay features an even higher statistic than $B^0 \to J/\psi K^{*0}$ and allows for an easier calibration method, as it is a decay of a charged *B* meson. Hence, the statistical and the type I systematic uncertainties are about a factor 2 smaller compared to $B^0 \to J/\psi K^{*0}$. Considering the uncertainties, both channels deliver compatible results for the OS tagger combination. This gives a good legitimation for the described calibration method to be applied on different taggers like the SS pion tagger, for which $B^+ \to J/\psi K^+$ decays are not suitable.

6.6.2 Calibration of the SS pion tagger for the LHCb Run 1 sin(2β) measurement

This section discusses the calibration of the SS pion tagger that the author contributed to the *CP* violation measurement with decays of $B^0 \to J/\psi K_S^0$ [68]. Similar to $B_s^0 \to D_s^\mp K^\pm$ decays, it is possible to measure *CP* violation in the interference between decay and decay after mixing (compare 3.4.3) with $B^0 \to J/\psi K_S^0$ decays. However, in contrast to $D_s^\mp K^\pm$ final state, there is only one final state, $J/\psi K_S^0$, which is a *CP* eigenstate. Another difference is that the decay-width difference $\Delta\Gamma_d$ in the B^0 system is compatible with zero [57]. Consequently, only a single decay-time-dependent *CP* asymmetry (see Eq. 3.58) exists for $B^0 \to J/\psi K_S^0$ decays

$$\mathscr{A}_{J/\psi K_S^0}(t) \approx S_{J/\psi K_S^0} \sin(\Delta m_d t), \tag{6.24}$$

with only one non-zero *CP* parameter

$$S_{J/\psi K_S^0} = \sin(2\beta). \tag{6.25}$$

The angle β is one of the CKM angles introduced in Eq. 3.17. Due to the clean relation between the *CP* asymmetry and the CKM angle, the decay $B^0 \to J/\psi K_S^0$ is often referred to as the golden channel for the measurement of *CP* violation in the B^0 meson system [135].

The calibration that is presented in the following is an input for LHCb's Run 1 sin(2β) measurement performed with $B^0 \to J/\psi K_S^0$ decays [68]. Apart from the same side pion tagger, the *CP* violation analysis uses the standard combination of OS taggers as well. However, for the OS tagger combination, the calibration is performed

with decays of $B^+ \to J/\psi K^+$. The $B^0 \to J/\psi K^{*0}$ decay is used for the calibration of the SS pion tagger, as the composition of fragmentation products is different between the B^0 and B^+ hadronisation, which consequently leads to differing tagging responses. The output of the SS pion tagger is pre-calibrated since DAVINCI v35r1 with the parameters

$$p_0 = 0.425 \quad \text{and} \quad p_1 = 0.939, \tag{6.26}$$

which were determined in a previous calibration with $B^0 \to J/\psi K^{*0}$ decays. This earlier calibration was also performed in the scope of this thesis utilising the same strategy and selection as presented above, but based on an earlier version of the reconstruction software. Due to the pre-calibration, it is expected that the tagger is already well calibrated. However, a set of calibration parameters and their corresponding uncertainties are determined and applied on top of the pre-calibration in the fitter of the *CP* violation analysis. This tagging calibration strategy allows to propagate the uncertainties and correlations of the calibration parameters to the *CP* parameters. This propagation is implemented in the $B^0 \to J/\psi K_S^0$ *CP* fit by applying Gaussian constraints on the tagging parameters corresponding to the central values, uncertainties, and correlations obtained in the tagging calibration presented here.

The calibration with $B^0 \to J/\psi K^{*0}$ decays is performed by utilising the strategy discussed in Sec. 6.6. After the selection there are about 62 000 $B^0 \to J/\psi K^{*0}$ signal candidates with a non-zero tag decision of the SS pion tagger. The tagging efficiency is measured to be $\epsilon_{\text{tag}} = (16.87 \pm 0.07)\%$, while the effective tagging efficiency is $\epsilon_{\text{eff}} = (0.532 \pm 0.034)\%$. Five bins of the mistag prediction η are fit simultaneously in order to deduce the tagging calibration. The calibration parameters are determined as

$$
\begin{aligned}
p_0^{\text{SS}\pi} &= 0.4232 \pm 0.0029 \,(\text{stat.}) \pm 0.0020 \,(\text{syst. type I}) \pm 0.0019 \,(\text{syst. type II}), \\
p_1^{\text{SS}\pi} &= 1.011 \pm 0.064 \,(\text{stat.}) \pm 0.009 \,(\text{syst. type I}) \pm 0.030 \,(\text{syst. type II}), \\
\Delta p_0^{\text{SS}\pi} &= -0.0026 \pm 0.0043 \,(\text{stat.}) \pm 0.0024 \,(\text{syst. type I}) \pm 0.0013 \,(\text{syst. type II}), \\
\Delta p_1^{\text{SS}\pi} &= -0.171 \pm 0.096 \,(\text{stat.}) \pm 0.029 \,(\text{syst. type I}) \pm 0.027 \,(\text{syst. type II}), \\
\langle \eta^{\text{SS}\pi} \rangle &= 0.425.
\end{aligned}
\tag{6.27}
$$

The results confirm that the tagging output is already very well calibrated, as p_0 and p_1 are compatible with $\langle \eta \rangle$ and unity, respectively. Table 6.4 contains the correlations between the calibration parameters, which are very small. Figure 6.9 contains the invariant $J/\psi K^{*0}$ mass and decay-time distributions together with projections of the fitted PDF. The raw mixing asymmetry is plotted in the left part of Fig. 6.10, while the right part of the figure shows the cross check of the linearity of the calibration.

The type I systematic uncertainties are estimated based on toy-studies in which 1000 samples are generated and fitted. Various possible sources are tested, including the decay-time resolution, the decay-time acceptance, asymmetries in production and detection, and the choice of the signal and background mass models. The largest

Table 6.4: Correlations between the calibration parameters of the SS pion tagger.

|	$p_0^{\text{SS}\pi}$	$p_1^{\text{SS}\pi}$	$\Delta p_0^{\text{SS}\pi}$	$\Delta p_1^{\text{SS}\pi}$
$p_0^{\text{SS}\pi}$	1	0.04	-0.007	0.0004
$p_1^{\text{SS}\pi}$		1	0.0016	-0.006
$\Delta p_0^{\text{SS}\pi}$			1	0.03
$\Delta p_1^{\text{SS}\pi}$				1
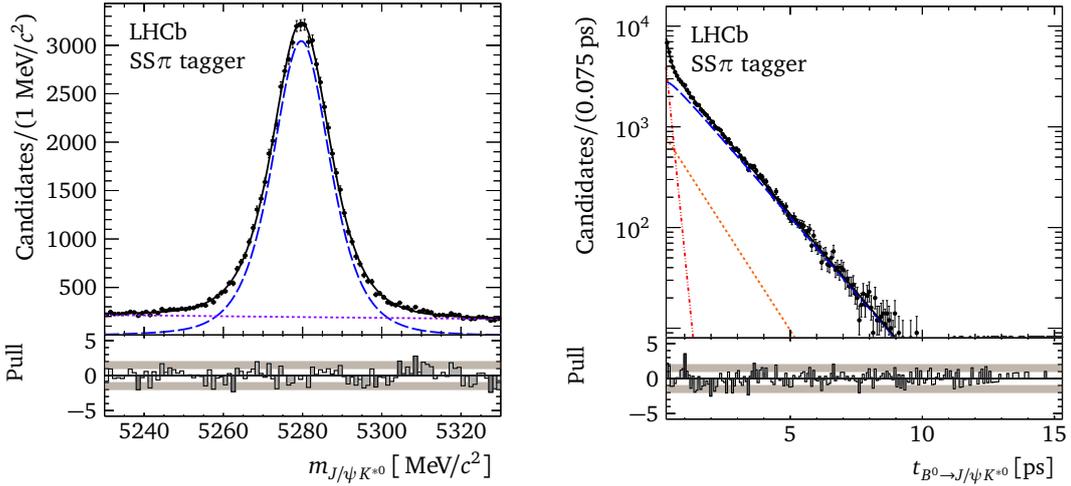
Figure 6.9: The invariant $J/\psi K^{*0}$ mass (left) and the decay-time (right) distributions of the $B^0 \to J/\psi K^{*0}$ candidates tagged by the SS pion tagger. The total PDF is plotted as a black solid line and the signal is shown as dashed blue line. The combinatorial background is depicted by the shortly-dashed purple line in the mass distribution, while it is illustrated in the decay-time distribution by the shortly-dashed orange and dashed-dotted red lines.

type I systematic uncertainties, which are still smaller than the statistical uncertainties, originate from the comparison of the nominal calibration method using a two dimensional fit and a fit utilising the *sPlot* technique. This particular source of uncertainty is investigated as the estimation of the type II uncertainties relies on weighted calibrations. Type II uncertainties cover kinematic differences between the calibration channel $B^0 \to J/\psi K^{*0}$ and the physics channel $B^0 \to J/\psi K_S^0$. They are estimated by weighting different kinematic observables, which are correlated to the flavour tagging output, in the calibration channel by the respective distributions in the signal decay. After weighting the observables one-by-one, the flavour-tagging calibration is repeated, but only performing a fit of the signal-decay-time PDF utilising *sWeights*. The differences between the weighted calibrations and the nominal one are taken as systematic uncertainties. The weighting is performed for the number of tracks (nTracks), the number of PVs (nPVs), the azimuthal angle (ϕ), the transverse momentum of the B (p_{T}), and the pseudo rapidity (η). Table 6.5 contains the observed
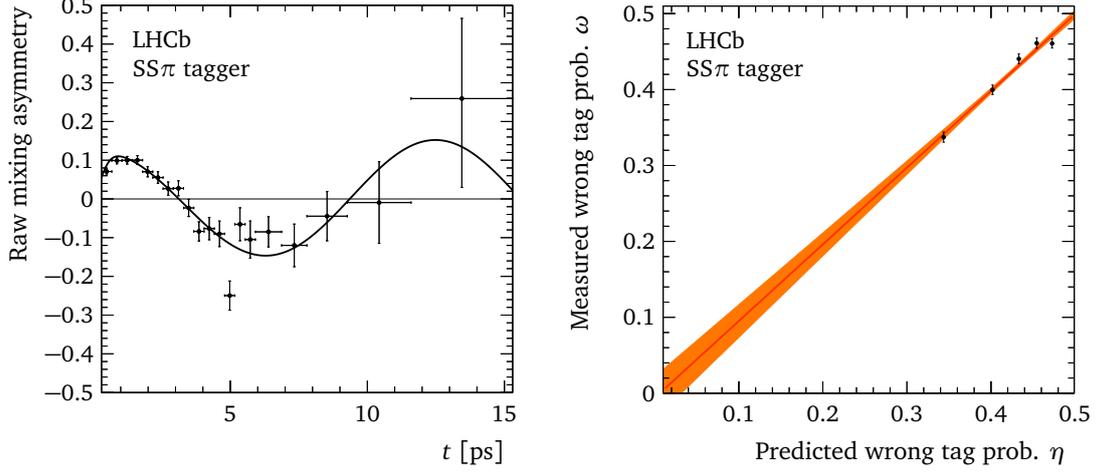
Figure 6.10: he raw mixing asymmetry for all $B^0 \to J/\psi K^{*0}$ candidates tagged by the SS pion tagger (left). The calibration of the wrong tag estimate for the SS pion tagger based on the decay-time fit of the $B^0 \to J/\psi K^{*0}$ candidates (right).

differences for the calibration parameters. The largest changes are observed when weighting in in p_T and nTracks. However, all type II systematics are well below the statistical uncertainties.

Table 6.5: Type II systematic uncertainties of the SS pion calibration.

	$\delta p_0 (10^{-3})$	$\delta p_1 (10^{-2})$	$\delta \Delta p_0 (10^{-3})$	$\delta \Delta p_1 (10^{-2})$
$p_T(B)$	1.3	2.1	0.6	2.7
$\eta(B)$	0.23	0.020	0.7	0.5
$\phi(B)$	0	0.6	0.14	0.27
nTracks	1.5	2.0	1.1	0.8
nPVs	0.6	0.5	0	0.05
Total	1.9	3.0	1.3	2.7
Percentage of stat. uncert.	65.5 %	46.9 %	30.2 %	28.1 %

The Run 1 analysis of CP violation with $B^0 \to J/\psi K_S^0$ decays results in a sensitivity similar to that of the B factories' measurements in the same channel [68]. One major change of the analysis with the full Run 1 data compared to the preceding LHCb measurement with 1 fb^{-1} [136] is the use of the same side pion tagger. This addition is only made possible by the presented calibration.

6.6.3 Calibration of the new OS charm tagger

In order to further improve the overall tagging performance of the experiment, the LHCb collaboration is constantly working on the implementation of new tagging algorithms. An example of such a new algorithm is the OS charm tagger. This algorithm exploits the correlation between the initial *B* meson flavour and the charge of a reconstructed secondary charm hadron, which originates from the decay of the other *b* hadron (see Fig. 6.1). The selection of the different charm hadrons and decay channels is done using boosted decision trees trained on simulated data. The output of the tagger is calibrated with the self-tagging decay $B^+ \to J/\psi K^+$ and cross-checked with $B^0 \to J/\psi K^{*0}$ decays. More details on the OS charm tagger can be found in the corresponding LHCb publication [129]. The calibration cross-check that is presented in the following is a part of this publication.

The tagging efficiency of the OS charm tagger is small due to the challenges in the reconstruction. In the selected $B^0 \to J/\psi K^{*0}$ Run 1 data a tagging efficiency of $\epsilon_{\text{tag}} = (3.32 \pm 0.04)\,\%$ is measured. However, the low mistag rate allows for a decent tagging efficiency of $\epsilon_{\text{eff}} = (0.30 \pm 0.03)\,\%$. The calibration is performed in five bins of the mistag estimate and determines the calibration parameters to be

$$
\begin{aligned}
p_0^{\text{OSC}} &= 0.361 \pm 0.008\,\text{(stat.)} \quad \pm 0.003\,\text{(syst. type I)}\,, \\
p_1^{\text{OSC}} &= 1.16 \ \pm 0.17 \ \text{(stat.)} \quad \pm 0.02 \ \text{(syst. type I)}\,, \\
\Delta p_0^{\text{OSC}} &= 0.023 \pm 0.011\,\text{(stat.)} \quad \pm 0.004\,\text{(syst. type I)}\,, \\
\Delta p_1^{\text{OSC}} &= 0.21 \ \pm 0.25 \ \text{(stat.)} \quad \pm 0.04 \ \text{(syst. type I)}\,, \\
\langle \eta^{\text{OSC}} \rangle &= 0.379\,.
\end{aligned}
\tag{6.28}
$$

In Tab. 6.6, the correlations between the calibration parameters are listed. The projection of the decay-time and invariant mass distributions are shown in Fig. 6.11. In the left part of Fig. 6.12 the projection of the mixing asymmetry is plotted. Notably is the large amplitude compared to the OS tagger combination (Fig. 6.8) and the SS pion tagger (Fig. 6.10). The relation between ω and η is well modelled by a linear function, as illustrated in the right part of Fig. 6.12. The calibration is well compatible with that in the $B^+ \to J/\psi K^+$ channel [129], which confirms the usability of the new tagging algorithm. The first analysis to use the OS charm tagger is the *CP* violation analysis with $B^0 \to D^+ D^-$ decays [137]. This analysis incorporates the highest effective tagging efficiency to date in a tagged *CP* violation measurement at LHCb.

Table 6.6: Correlations between the calibration parameters of the OS charm tagger.

	p_0^{OSC}	p_1^{OSC}	Δp_0^{OSC}	Δp_1^{OSC}
p_0^{OSC}	1	0.04	-0.004	0.02
p_1^{OSC}		1	0.02	-0.02
Δp_0^{OSC}			1	0.03
Δp_1^{OSC}				1
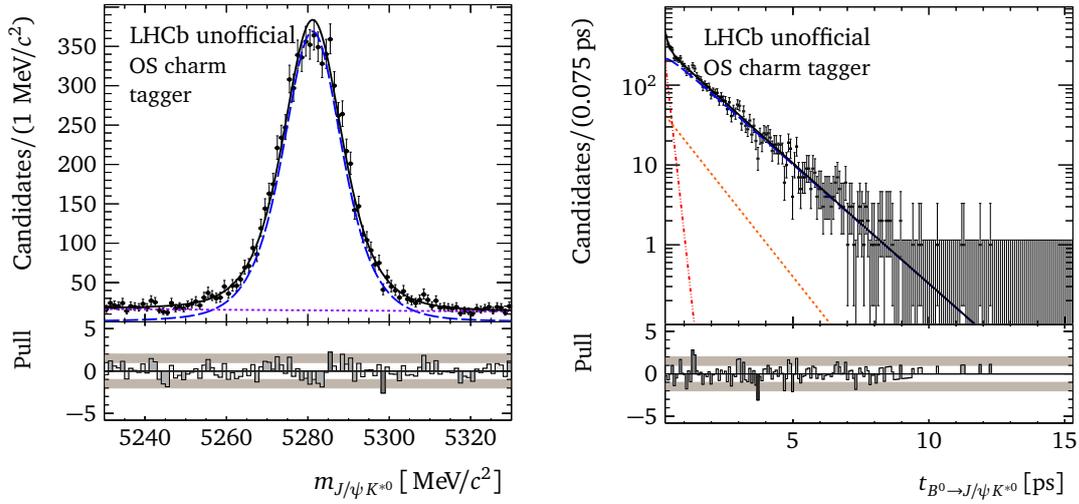
Figure 6.11: The invariant $J/\psi K^{*0}$ mass (left) and the decay-time (right) distributions of the $B^0 \to J/\psi K^{*0}$ candidates tagged by the OS charm tagger. The total PDF is plotted as a black solid line and the signal is shown as dashed blue line. The combinatorial background is depicted by the shortly-dashed purple line in the mass distribution, while it is illustrated in the decay-time distribution by the shortly-dashed orange and dashed-dotted red lines.
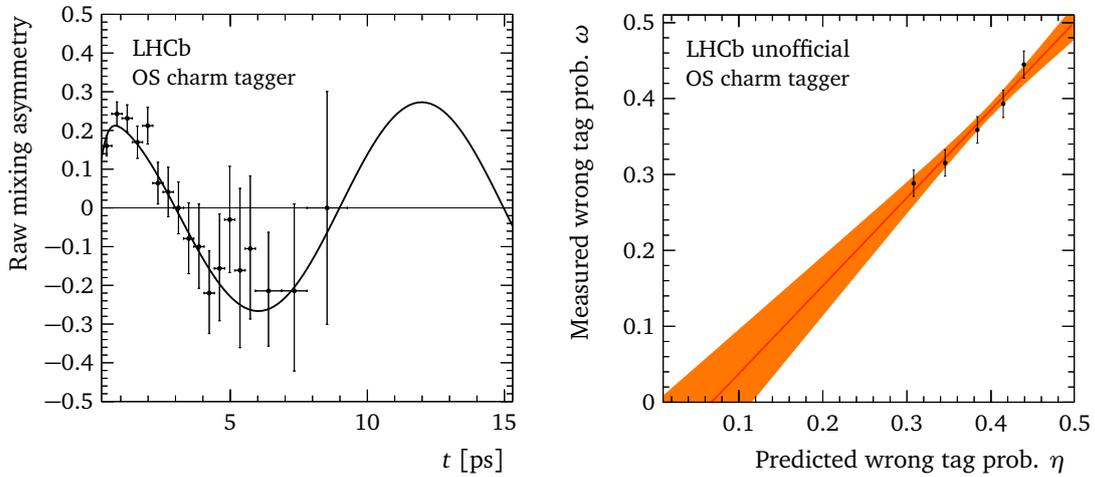
Figure 6.12: The left plot shows the raw mixing asymmetry for all $B^0 \to J/\psi K^{*0}$ candidates tagged by OS charm tagger. The right plot shows the calibration of the wrong tag estimate for the OS charm tagger based on the decay-time fit of the $B^0 \to J/\psi K^{*0}$ candidates.

7 Data Preparation

The measurement of *CP* violation in $B_s^0 \to D_s^\mp K^\pm$ decays is performed on a dataset from the Run 1 data taking period of the LHC. The dataset has an integrated luminosity of $0.98\,\mathrm{fb}^{-1}$ $(1.99\,\mathrm{fb}^{-1})$ of *pp* collisions recorded with the LHCb detector at centre-of-mass energies of $7\,\mathrm{TeV}$ $(8\,\mathrm{TeV})$ in 2011 (2012). The magnet polarity was switched several times during the data taking, resulting in $0.42\,\mathrm{fb}^{-1}$ and $1.00\,\mathrm{fb}^{-1}$ taken with the up direction in 2011 and 2012, respectively, while for the rest of the dataset the down polarity was present. Performing the analysis on subsets separated by the year of data taking or the magnet polarity allows to cross-check the full analysis chain.

Decay modes with purely hadronic final states, like $B_s^0 \to D_s^- \pi^+$ or $B_s^0 \to D_s^\mp K^\pm$, have higher background levels compared to decays such as $B^0 \to J/\psi K^{*0}$, which can be triggered by muons from the J/ψ decay (compare Sec. 6.6). Hence, the background has to be suppressed in order to allow for a significant measurement of *CP* violation. Therefore, a multistep selection is developed, which selects both the signal decay $B_s^0 \to D_s^\mp K^\pm$, as well as the two control modes $B^0 \to D^- \pi^+$ and $B_s^0 \to D_s^- \pi^+$. The latter two decays are kinematically similar to $B_s^0 \to D_s^\mp K^\pm$, but feature higher branching ratios. They allow modelling of remaining background components in the $B_s^0 \to D_s^\mp K^\pm$ fits as well as the study of detector effects like the decay time acceptance. The selection starts with a trigger and stripping based preselection, as explained in Sec. 7.1. Subsequently, cuts on kinematic properties are tightened and from then on PID information are used, in order to separate the decay modes and veto physics backgrounds. The offline selection requirements for $B_s^0 \to D_s^\mp K^\pm$ and $B_s^0 \to D_s^- \pi^+$ are discussed in Sec. 7.2, while those for $B^0 \to D^- \pi^+$ can be found in Sec. 7.3. A multivariate algorithm is trained and applied as the final stage of the selection (see Sec. 7.4), which leads to a significant reduction of combinatorial background.

Fully simulated data is a vital input to different parts of the analysis. An overview of the used samples is given in Sec. 7.5.

7.1 Reconstruction and preselection

The reconstruction strategy is similar for the signal decay $B_s^0 \to D_s^\mp K^\pm$ and the two control modes $B^0 \to D^- \pi^+$ and $B_s^0 \to D_s^- \pi^+$. It is outlined in this section exemplarily for a negatively charged charm meson. However, charge conjugation is implied. The reconstruction starts by building a $D_{(s)}^-$ candidate from three charged hadrons. In order to form the *B* candidate, a single kaon or pion is added. This particle is called the bachelor particle. The D^- mesons of the $B^0 \to D^- \pi^+$ decays are reconstructed

Table 7.1: Branching fractions of $B_s^0 \to D_s^{\mp} K^{\pm}$ [138], $B_s^0 \to D_s^- \pi^+$, and the reconstructed D_s^- decays [24]. Additionally, the fragmentation fraction f_s/f_d is given, which is taken from the LHCb hadronic/semi-leptonic combination [139].

Channel	Branching fraction
$B_s^0 \to D_s^{\mp} K^{\pm}$	$(2.29 \pm 0.05\,(\text{stat}) \pm 0.06\,(\text{syst}) \pm 0.17(B_{B_s^0})) \times 10^{-4}$
$B_s^0 \to D_s^- \pi^+$	$(3.04 \pm 0.23)) \times 10^{-3}$
$D_s^- \to K^- K^+ \pi^-$	$(5.45 \pm 0.17) \times 10^{-2}$
$D_s^- \to K^- \pi^+ \pi^-$	$(6.6 \pm 0.4) \times 10^{-3}$
$D_s^- \to \pi^- \pi^+ \pi^-$	$(1.09 \pm 0.05) \times 10^{-2}$
f_s/f_d	$0.259 \pm 0.015\,(\text{stat\&syst})$

in the $D^- \to K^- \pi^+ \pi^-$ mode, while three different final states are used to build D_s^- mesons, namely $D_s^- \to K^- K^+ \pi^-$, $D_s^- \to K^- \pi^+ \pi^-$, and $D_s^- \to \pi^- \pi^+ \pi^-$. Furthermore, the two resonant decays $D_s^- \to \phi \pi^-$ and the $D_s^- \to K^{*0} K^-$ are being considered for the $D_s^- \to K^- K^+ \pi^-$ mode. Non-resonant decays of $D_s^- \to K^- K^+ \pi^-$ are called $D_s^- \to (KK\pi)_{\text{nonres}}$ in the following. Table 7.1 contains the branching fractions of the signal decay and $B_s^0 \to D_s^- \pi^+$, as well as those of the further D_s^- decays. The different D_s^- final states are distinguished by kinematic and PID requirements that are discussed in Sec. 7.2.

Requirements on trigger and stripping are chosen to perform a loose preselection that reduces the data to a manageable amount, while maintaining a very high signal efficiency. No explicit requirements are made on the L0 triggers. In terms of the high level trigger (see Sec. 4.3.2), all events are required to be triggered by the `HLT1TrackAllL0` line in HLT1 and by either the inclusive ϕ line (`HLT2IncPhi`), the two-, or the three-body topological trigger line (`HLT2TopoBBDT`) in HLT2. Details about the trigger lines can be found in dedicated public notes (see Refs. [140, 141]). During the data taking in 2011 and 2012, the trigger conditions were stable.

The used stripping versions (see Sec. 4.4.2) are `Stripping21r1` and `Stripping21` for the data taken in 2011 and 2012, respectively. Candidates for the analysis originate from the output of two stripping lines. which only differ concerning the applied mass hypothesis of the bachelor particle. The kaon mass is used for $B_s^0 \to D_s^{\mp} K^{\pm}$ decays (`StrippingB02DKD2HHHBeauty2CharmLine`), while for the control modes $B_s^0 \to D_s^- \pi^+$ and $B^0 \to D^- \pi^+$ the pion mass is used (`StrippingB02DPiD2HHHBeauty2CharmLine`). Apart from accepting only events with less than 500 long tracks, the stripping lines include requirements on the kinematics of the reconstructed particles and their displacement from the primary interaction. Both stripping lines do not apply any PID criteria, but use a bagged boosted decision tree (BDT) [142, 143] to further reduce the background rate. A short introduction to BDTs in general can be found in Sec. 5.3. The specific stripping cuts are summarised in the following paragraphs.

All kaon and pion candidates are required to be long tracks and have a track $\chi^2/\text{ndf} < 3$, where ndf denotes number degrees of freedom. Furthermore, a ghost

probability below 0.4, a transverse momentum $p_T > 100\,\mathrm{MeV}/c$, a momentum $p > 1\,\mathrm{GeV}/c$, and a minimum impact parameter $\chi^2_{IP} > 4$ to any primary vertex are required. Stronger requirements are applied on the properties of the bachelor particle, as it needs to have a track $\chi^2/\mathrm{ndf} < 2.5$, $p_T > 500\,\mathrm{MeV}/c$, and $p > 5\,\mathrm{GeV}/c$.

The requirements on the $D_{(s)}^-$ are applied before vertexing, in order to speed up the processing time of the stripping selection. These requirements include that the scalar p_T sum of its daughters is larger than $1.8\,\mathrm{GeV}/c$, the maximal distance of closest approach (DOCA) between all pairs of particles forming the vertex is below 0.5 mm, and that the reconstructed mass is within $100\,\mathrm{MeV}/c^2$ of the nominal D^- or $D_s^{\mp}$ mass. Moreover, at least one of the D_s^- daughter tracks must fulfil the same tighter requirements regarding track χ^2/ndf, p_T, and p that are applied to the bachelor particle. After the vertexing, additionally a vertex $\chi^2/\mathrm{ndf} < 10$ and a minimal vertex separation χ^2 larger than 16 to any other primary vertex is required.

When the B candidates are built by combining D_s^- candidates with bachelor particles, it is required that good secondary vertices are formed ($\chi^2/\mathrm{ndf} < 10$). Furthermore, a displacement of the SV and the PV is ensured, by demanding a B decay time of more than 0.2 ps. Another requirement on the B candidates is a χ^2_{IP} below 25, where χ^2_{IP} is the difference in the vertex fit χ^2 of the associated PV with and without the B candidate. The latter cut makes sure that the B candidates' trajectories are consistent with originating from the PV. Moreover, the cosine of the angle between the B momentum and its direction of flight has to be larger than 0.999, which leads to B momentum vectors pointing along the path between the PV and the SV. In addition to the previous cuts, a BDT is used to select B candidates. This BDT is based on the p_T of the B candidates, their vertex separation χ^2 from the nearest PV, and the sum of the B- and the D-vertex χ^2 divided by the sum of their numbers of degrees of freedom. The cut on the BDT response (> 0.05) is about 100 % efficient for the corresponding signal B decays.

The normalised distributions of the invariant $D_s^{\mp} h^{\pm}$ mass and the B decay time after the stripping are shown in Fig. 7.1. The figure shows the distributions of the data taken in 2011 and 2012 superimposed, which show a good agreement with each other. In the invariant $D_s^{\mp} h^{\pm}$ mass distribution on data (left plot of Fig. 7.1) a clear peak around the B_s^0 mass of $5366.82\,\mathrm{MeV}/c^2$ [24] is already visible. However, it is dominated by combinatorial background.

7.2 Cut-based selection of $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^{-} \pi^{+}$

This section discusses the offline selection requirements, which are applied to the $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^{-} \pi^{+}$ candidates as rectangular cuts prior to the subsequent multivariate selection (see Sec. 7.4). All requirements are listed in Tabs. 7.2 and 7.3. Momenta, invariant masses, and decay-time related observables used in the offline selection originate from the DecayTreeFitter algorithm [106].
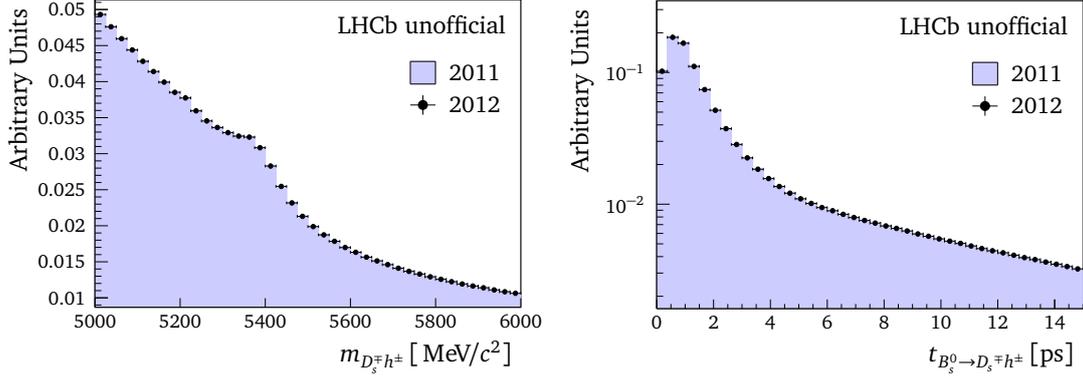
**Figure 7.1:** Distributions of the $m(D_s^{\mp}h^{\pm})$ invariant mass (left) and decay time (right) in $B_s^0 \to D_s^{\mp}h^{\pm}$ data after the stripping. The blue area denotes 2011 data, while black points indicate 2012 data.

The candidates need to have a measured B_s^0 decay time larger than 0.4 ps, which reduces the amount of prompt background and allows for a better description of the decay-time inefficiency (see Sec. 9.4). The decay time of the D_s^- with respect to the B_s^0 has to be greater than 0 ps. Moreover, an invariant mass $m(D_s^{\mp}h^{\pm})$ between 5300 and 5800 MeV/c^2 is required, while the combined invariant mass of the D_s^- daughter hadrons needs to be within 1930 and 2015 MeV/c^2. As mentioned in Sec. 7.1, the D_s^- final state $K^-K^+\pi^-$ can emerge from the resonant decays $D_s^- \to \phi\pi^-$ and $D_s^- \to K^{*0}K^-$, as well as without an intermediate resonance. The three modes are separated as outlined in the following. Candidates are identified as $D_s^- \to \phi\pi^-$ decays if the invariant mass of their K^+K^- pair is within 20 MeV/c^2 of the $\phi(1020)$ mass. The remaining candidates are categorised as $D_s^- \to K^{*0}K^-$ if the invariant mass of their $K^+\pi^-$ pair lies within a 50 MeV/c^2 window around the $K^{*0}(892)$ mass, otherwise they are classified as non-resonant decays.

Table 7.2: Requirements on invariant masses and decay times of the B_s^0 and D_s^- candidates.

Applied to	Description	Requirement
All modes	$m(D_s^+h^{\pm})$	$\in [5300, 5800]$ MeV/c^2
	$m(h^-h^+h^-)$	$\in [1930, 2015]$ MeV/c^2
	B_s^0 decay time (from DTF)	> 0.4 ps
	D_s decay time (wrt. B_s^0)	> 0 ps
$D_s^- \to K^-K^+\pi^-$		
$\quad D_s^- \to \phi\pi^-$	$m(K^+K^-)$	$\in [1000, 1040]$ MeV/c^2
$\quad D_s^- \to K^{*0}K^-$	$m(K^+K^-)$	$\notin [1000, 1040]$ MeV/c^2
	$m(\pi^-K^+)$	$\in [842, 942]$ MeV/c^2
$\quad D_s^- \to (KK\pi)_{\mathrm{nonres}}$	$m(K^+K^-)$	$\notin [1000, 1040]$ MeV/c^2
	$m(\pi^-K^+)$	$\notin [842, 942]$ MeV/c^2

Cuts on the particle identification are necessary to reject physical backgrounds, *i.e.* decays, which are (partly) misreconstructed or in which one or more particles are not reconstructed at all. An important background in the reconstruction of $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ decays are decays of $B_s^0 \rightarrow D_s^- \pi^+$ and vice versa. This type of bachelor particle misidentification is strongly reduced by requiring a bachelor PID$K > 5$ for $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ decays and a PID$K < 0$ for $B_s^0 \rightarrow D_s^- \pi^+$ decays. Moreover, semi-leptonic background decays, *e.g.* $B^0 \rightarrow D^- \mu \bar{\nu}_\mu$, are suppressed by applying a PID$\mu < 2$ cut for the bachelor particle. Besides the bachelor particle, PID requirements are used for the daughter particles of the D_s^-. However, the choice of the PID cuts depends on the respective reconstructed D_s^- final state as different background contributions are possible. The optimisation strategy of the PID cuts chosen to veto these backgrounds is always the same. The PID requirements are being iteratively tightened until no peaking background structure is left in the D_s^- invariant mass spectrum. Two peaking background contributions exist for the $D_s^- \rightarrow K^- \pi^+ \pi^-$ mode. Firstly, there are $\overline{\Lambda}_c^- \rightarrow \overline{p} \pi^+ \pi^-$ decays in which the antiproton is misidentified as a kaon. It is vetoed by requiring a cut of $(\mathrm{PID}K - \mathrm{PID}p) > 5$ for the kaon and excluding candidates with an invariant $h^- h^+ h^-$ mass between 2255 and 2315 MeV/c^2. For the latter, the invariant mass is computed under the $\overline{\Lambda}_c^-$ hypothesis, *i.e.* the potentially misreconstructed kaon is constrained to the proton mass. The second background decay is $D^- \rightarrow \pi^- K^+ \pi^-$, where the kaon and one of the pions are misidentified. This background is vetoed through tight cuts on the PID of the oppositely charged pion and kaon. In addition, the invariant $h^- h^+ h^-$ mass is calculated with switched hypotheses of the kaon and one of the pions and candidates with a resulting invariant mass around the nominal D^- mass are excluded. The same two background decays are found in the spectrum of the $D_s^- \rightarrow K^- K^+ \pi^-$ final state. Here, the $\overline{\Lambda}_c^-$ background occurs in the case of a double misidentification, while for the D^- background decay only a single misidentification needs to takes place. The vetoes that are applied are similar to those for the $D_s^- \rightarrow K^- \pi^+ \pi^-$ final state. In case of the $D_s^- \rightarrow \pi^- \pi^+ \pi^-$ final state, no peaking background falls into the relevant mass range. However, loose requirements are applied for the pion PIDs, which reduce combinatorial background. All vetoes and further PID requirements can be found in Tab. 7.3. This table also contains the cuts applied on the D_s vertex separation χ^2 from the B_s^0 vertex, which suppress charmless backgrounds. These backgrounds are challenging, as they peak directly in the signal B_s^0 mass region. Due to CKM factor suppression relative to $B_s^0 \rightarrow K K \pi \pi$, it is possible to apply a softer cut (> 2) for $D_s^- \rightarrow K^- K^+ \pi^-$ with respect to $D_s^- \rightarrow \pi^- \pi^+ \pi^-$ and $D_s^- \rightarrow K^- \pi^+ \pi^-$ (> 9). In order to check, if there is remaining charmless background, the sideband $m(h^- h^+ h^-) > 2030$ MeV/c^2 is inspected after applying the full offline selection. No significant contributions persist.

Table 7.3: PID requirements and vetoes for $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ decays.

Applied to	Description	Requirement
All modes	Semi-leptonic backgrounds veto:	
	PIDμ of bachelor	< 2
$B_s^0 \to D_s^- \pi^+$	PIDK of bachelor	< 0
$B_s^0 \to D_s^{\mp} K^{\pm}$	PIDK of bachelor	> 5
$D_s^- \to K^- K^+ \pi^-$	D_s^- vertex separation χ^2 (wrt. B_s^0)	> 2
	D^0 veto:	
	$m(K^+ K^-)$	$< 1840 \,\mathrm{MeV}/c^2$
	D^- veto:	
	PIDK of same charge K	> 10, or
	D_s^- under D^- hypothesis	$\notin [1840, 1900]\,\mathrm{MeV}/c^2$
	$\overline{\Lambda}_c^-$ veto:	
	p veto, same charge K	PID$K - PIDp > 5$, or
	D_s^- under $\overline{\Lambda}_c^-$ hypothesis	$\notin [2255, 2315]\,\mathrm{MeV}/c^2$
$D_s^- \to \phi \pi^-$	PIDK of both K	> -2
$D_s^- \to K^{*0} K^-$	PIDK of same charge K	> 5
	PIDK of opposite charge K	> -2
$D_s^- \to (KK\pi)_{\mathrm{nonres}}$	PIDK of both K	> 5
	PIDK of π	< 10
$D_s^- \to K^- \pi^+ \pi^-$	D_s^- vertex separation χ^2 (wrt. B_s^0)	> 9
	PIDK of K	> 10
	PIDK of both pions	< 5
	PIDp of both pions	< 10
	D^0 veto:	
	$m(K^+ \pi^-)$	$< 1750 \,\mathrm{MeV}/c^2$
	D^- veto:	
	PIDK of opposite charge π	< -10, or
	PIDK of opposite charge K	> 20, or
	D_s^- under D^- hypothesis	$\notin [1839, 1899]\,\mathrm{MeV}/c^2$
	$\overline{\Lambda}_c^-$ veto:	
	p veto, same charge K	PID$K - PIDp > 5$, or
	D_s^- under $\overline{\Lambda}_c^-$ hypothesis	$\notin [2255, 2315]\,\mathrm{MeV}/c^2$
$D_s^- \to \pi^- \pi^+ \pi^-$	D_s^- vertex separation χ^2 (wrt. B_s^0)	> 9
	PIDK of all pions	< 10
	PIDp of all pions	< 10
	D^0 veto:	
	Both $m(\pi^+ \pi^-)$	$< 1700 \,\mathrm{MeV}/c^2$

7.3 Cut-based selection of $B^0 \rightarrow D^- \pi^+$

The main purpose of the $B^0 \rightarrow D^- \pi^+$ sample is the correction of differences between simulated and detector data (see Sec. 8.1). As these corrections are needed for the $B_s^0 \rightarrow D_s^\mp K^\pm$ and $B_s^0 \rightarrow D_s^- \pi^+$ decays, it is important that the $B^0 \rightarrow D^- \pi^+$ selection does not introduce substantial differences in kinematic observables between the three decays. Due to the topological and geometrical properties of the decays, the selection requirements for $B^0 \rightarrow D^- \pi^+$ can be chosen very similar to the other decays.

All selection criteria for $B^0 \rightarrow D^- \pi^+$ decays are listed in Tab. 7.4. Compared to $B_s^0 \rightarrow D_s^\mp K^\pm$ and $B_s^0 \rightarrow D_s^- \pi^+$ decays, the same cut on the PIDμ of the bachelor is applied. Furthermore, the D^- has to fulfil the same requirements on its decay time and vertex separation χ^2 with respect to the B as the D_s^- in the other two decays. The invariant B^0 mass has to be within the larger range of 5000 to 6000 MeV/c^2. This allows to model low mass backgrounds in the subsequent fit of the invariant mass. Decays with $\overline{\Lambda}_c^-$ and D_s^- are suppressed by vetoes. Except for the vetoes, no specific requirements on PID information for either the bachelor particle or the D daughters have to be fulfilled, which prevents a distortion of observables such as momenta or the number of tracks.

Table 7.4: Offline selection requirements for $B^0 \rightarrow D^- \pi^+$ candidates.

Description	Requirement
$m(D^- \pi^+)$	$[5000, 6000]$ MeV/c^2
$m(K^+ \pi^- \pi^-)$	$[1830, 1920]$ MeV/c^2
D^- decay time (wrt. to B^0, from DTF)	> 0 ps
D^- vertex separation χ^2 (wrt. B^0)	> 9
Semi-leptonic backgrounds veto:	
$\quad$ PIDμ of bachelor	< 2
$\overline{\Lambda}_c^-$ veto:	
$\quad p$ veto for pions	PID$p < 0$, or
$\quad D^-$ under $\overline{\Lambda}_c^-$ hypothesis	$\notin [2255, 2315]$ MeV/c^2
D_s^- veto:	
$\quad$ kaon veto for pions	PID$K > 0$, or
$\quad D^-$ under D_s^- hypothesis	$\notin [1950, 2030]$ MeV/c^2

7.4 Multivariate selection

The final step of the selection is the training and application of a multivariate algorithm (MVA), in order to suppress the still dominating combinatorial background. MVAs comprise a family of statistical methods that are able to perform classification tasks by making use of more than one variable at a time. In contrast to simple univariate analyses, they profit from correlations among their input variables. Well-known examples for MVAs are artificial neural networks (ANNs) [142], boosted decision trees (BDTs, see Sec. 5.3) [144], and random forests [145]. In particle physics it is nowadays quite common to train MVAs by exploiting machine learning techniques in order to utilise them for the separation of signal and background contributions. In the course of this thesis, a BDT is trained to reduce the amount of combinatorial background, while maintaining a high signal efficiency. The training is described in Sec. 7.4.1. Afterwards, the optimisation of the cut on the classifier response is discussed in Sec. 7.4.2.

7.4.1 Training of the BDT

Before the author became responsible for the training of the BDT it had already been tested that new BDTs trained with simulated data perform worse than the BDT of the previous $B_s^0 \to D_s^\mp K^\pm$ analysis [75]. Thus, a data-driven BDT training strategy is developed, which is similar to the one of the former analysis. Similarities include the selection of input variables and choosing $B_s^0 \to D_s^- \pi^+$ data for the training. However, apart from the use of the full Run 1 dataset, a couple of improvements are implemented. One of them is the use of a 2-fold cross-validation. In order to apply this technique, the training data is split into candidates with even and odd event numbers. This is equivalent to a random selection, as the event number has no physical meaning and is assigned continuously to the candidates. Then, two BDTs are trained with the same options, one on each of the two equally sized sets of training data. This allows to test and apply the two classifiers to the training sample of the other BDT. Hence, a strong check on the presence of overtraining is enabled and it is prevented that the BDT is applied onto its own training candidates. The latter would be the case without the 2-folding approach, as the BDT is not only used in $B_s^0 \to D_s^\mp K^\pm$, but also in the control channels $B_s^0 \to D_s^- \pi^+$ and $B^0 \to D^- \pi^+$. As expected, and shown later on, the two BDTs give very similar responses. Therefore, they can be optimised and treated as one BDT.

The BDT training uses $B_s^0 \to D_s^- \pi^+$ decays, where the D_s is reconstructed only in the $K^- K^+ \pi^-$ final state, which is the one with the highest statistics (see Tab. 7.1). The training candidates are selected by the trigger and stripping requirements described in Sec. 7.1. Additionally, a preselection consisting of the cuts listed in Tab. 7.5 is applied. This preselection, which is only used for the $B_s^0 \to D_s^- \pi^+$ data sample utilised in the BDT training, includes slightly tighter requirements compared to the standard offline selection cuts described in the previous section. The invariant B_s^0 mass distribution of the $B_s^0 \to D_s^- \pi^+$ Run 1 sample used for the training is shown in the left plot of

Table 7.5: Requirements of the $B_s^0 \to D_s^- \pi^+$ sample used in the BDT training.

Description	Requirement
Bachelor	PID$K < 0$
Both kaons	PID$K > 0$
$m(D_s^{\mp} h^{\pm})$	$\in [5300, 5800]\,\mathrm{MeV}/c^2$
$m(h^- h^+ h^-)$	$\in [1940, 1990]\,\mathrm{MeV}/c^2$
D^- veto:	
PIDK of same charge K, or	> 10, or
$m(h^- h^+ h^-)$ under $m(K^- \pi^+ \pi^-)$ hypothesis	$< 1850\,\mathrm{MeV}/c^2$
$\overline{\Lambda}_c^-$ veto:	
p veto, same charge K, or	PID$K -$ PID$p > 5$, or
$m(h^- h^+ h^-)$ under $m(\bar{p} K^+ \pi^-)$ hypothesis	$\notin [2250, 2320]\,\mathrm{MeV}/c^2$

Fig. 7.2 together with an unbinned maximum likelihood fit (see Sec. 5.1). Due to the very clean distribution after the preselection, it is possible to describe the data with a simple model consisting of a Gaussian for the B_s^0 signal peak and an exponential function for the flat combinatorial background. Candidates with an invariant $D_s^- \pi^+$ mass above $5445\,\mathrm{MeV}/c^2$ are used as templates for the combinatorial background in the BDT training. For the signal, the *sPlot* technique [119] is exploited. This technique uses event weights, called *sWeights*, to statistically subtract background. It has been introduced in more detail in Sec. 5.2. The *sWeights* for the signal candidates are obtained from *sPlot* fits to the invariant B_s^0 mass range between 5310 and $5430\,\mathrm{MeV}/c^2$.
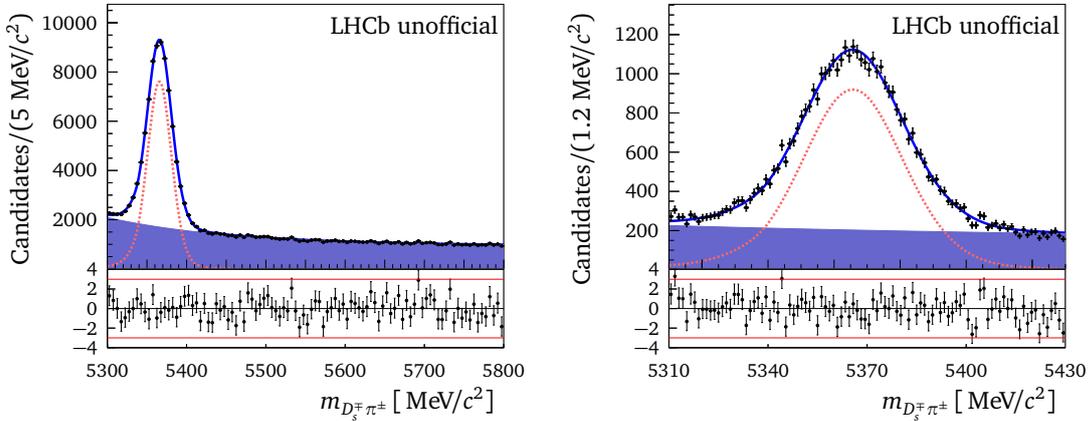
Figure 7.2: Invariant mass distributions of the $B_s^0 \to D_s^- \pi^+$ sample used for the BDT training. The left plot contains the full training sample, while the right plot shows the signal window of one half of the data. Both plots include as solid blue line the projection of a fit PDF (see Sec. 5.1), which consists of a Gaussian distribution for the signal (orange dashed line) and an exponential for the combinatorial background (filled blue area).

The *sPlot* fits are performed separately for the two halves of the training data, but in both cases the same fit model and fixed parameter values from the fit to the full range are used. For one half of the training data the *sPlot* fit is shown in the right part of Fig. 7.2. The *sWeights* are extracted in a narrower signal mass window, as it omits lots of background candidates from the upper mass sideband that would be assigned negative *sWeights* and, hence, be excluded from the training in any case.

Table 7.6 lists the input variables of the BDT training and contains their corresponding relevance in the training measured by TMVA. Overall, all variables have similar importance in the training and no input stands out. Compared to the previous analysis, the direction angle (DIRA) of the D with respect to its associated PV is dropped, as it does not add significant separation power. Furthermore, the logarithm is applied to some of the input variables. This is done to spread their distributions, which leads to an improved classification in TMVA [124]. In order to ensure that the BDTs can be used for all different D_s^- final states as well as the control decays, no PID information is used as input.

Table 7.6: Input variables of the BDT and importances in the training.

Variable	Importance (10^{-2})
B_s^0:	
log(radial flight distance)	7.663
minimum impact parameter χ^2	7.373
log(direction angle own PV)	7.105
log(lifetime vertex χ^2/ndf)	6.062
log(vertex χ^2/ndf)	4.040
Bachelor:	
log(p_{T})	10.18
log(minimum impact parameter χ^2)	6.861
cos(θ)	5.822
D_s:	
log(radial flight distance)	7.763
log(minimum impact parameter χ^2)	7.512
vertex χ^2/ndf	4.428
direction angle origin vertex	3.852
D_s children:	
log(minimum impact parameter χ^2)	7.369
log(minimum p_{T})	6.810
Bachelor and D_s daughers:	
maximum track ghost prob	7.160

The configuration of the BDTs is developed iteratively by training and comparing the performance of several BDTs that utilised different hyper parameters. In the end, the best performance is seen with gradient BDTs in the following configuration. Each of the two BDTs consists of 300 trees that feature a maximum depth of two. At least 4 % of the training candidates have to be present at each node. The cut values on the input variables are found by scanning at 40 points. Candidates with negative *sWeights* are excluded from the training, which is the default treatment of negative weights for gradient boosted decision trees in TMVA. Two more options are provided by TMVA, called "Pray" and "PairNegWeightsGlobal". The option "Pray" uses events with negative weights in the training without any transformation, while "PairNegWeightsGlobal" pairs candidates with negative and positive event weights and annihilates them. Both methods are marked as experimental. Nevertheless, it is checked whether they improve the performance. This is done by training three BDTs on the same part of the training data with exactly the same hyper parameters, except for the treatment of negative weights. Then, the performances are compared on the part of the data that was not used in the training. The resulting receiver operating curves (ROC), which represent a plot of the background rejection against the signal efficiency, are shown in Fig. 7.3. The better a BDT classifies, the closer its ROC gets to the upper right corner. All three BDTs perform very similar, but a little advantage exists for the BDT that ignores negative weights in the training. Therefore, this option is chosen for the final BDTs.

In the left plot of Fig. 7.4 the classifier output distributions of the two BDTs are shown. Both are applied to their testing samples, *i.e.* to the subset of candidates not used in their trainings. Moreover, Kolmogorov-Smirnov tests [146, 147] for both, signal and background, are performed to check the compatibility of the two BDT out-
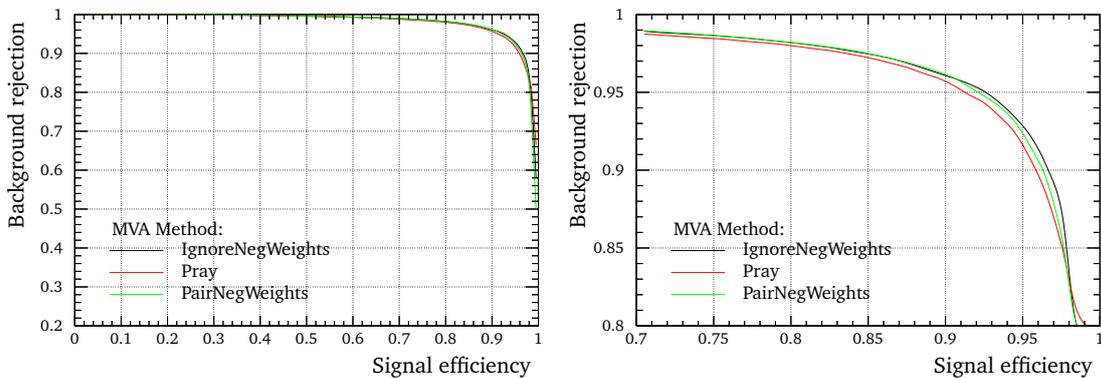
Figure 7.3: Receiver operating curves of three BDTs that differ only in their treatment of negative event weights. The green line corresponds to the BDT utilising "PairNeg-WeightsGlobal", the red line to the one that uses "Pray", and the black line represents the BDT that ignores negative event weights in the training. The left plot contains the full range, while the right one highlights performance differences by magnifying the upper right corner of the plot.
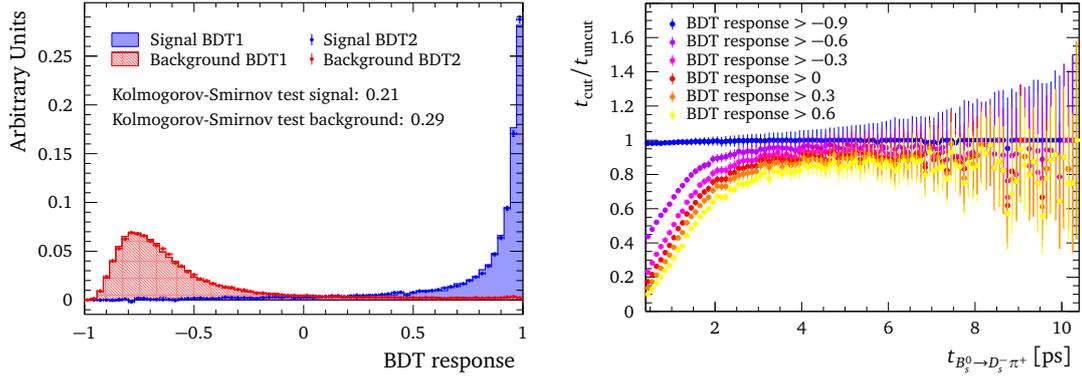
Figure 7.4: The left plot contains the classifier output distributions of signal (blue) and background (red), superimposing the distributions of the two BDTs when applied to the candidates they were not trained on. The right plot shows the quotient of the B_s^0 decay time after applying different cuts on the BDT response and the uncut decay time distribution.

puts. Their values confirm compatibility between the responses, as they are not close to 0 or 1. Thus, while ensuring an application only on their respective candidates, the two BDTs are treated as a single BDT throughout the analysis.

A crucial ingredient of the decay time fit is a proper description of the decay time inefficiency (see Sec. 9.4), which is mostly introduced by the detector and the selection. The BDTs use decay time biasing input variables like the impact parameter χ^2 of the B_s^0 meson. Thus, it is investigated if the potential decay-time inefficiency of the BDT leads to steps in the decay-time distribution, which would make a description challenging. The test is done on the full preselected $B_s^0 \to D_s^- \pi^+$ sample that is used in the training, with each BDT being applied to their respective part of the data they were not trained on. Then, several cuts on the BDT responses are applied and the resulting B_s^0 decay-time distribution is divided by the distribution without a cut. The right part of Fig. 7.4 shows the resulting ratios. As expected, a decay-time-dependent inefficiency is introduced, but the acceptance changes smoothly as a function of decay time. This allows to model the decay-time acceptance of the MVA alongside other acceptance effects using the spline method, as explained in Sec. 9.4.

7.4.2 BDT classifier cut optimisation

The cut on the BDT classifier response has to serve the goal of the analysis, which is to measure the *CP* parameters with the optimal precision in the given data sample. In terms of the sensitivity of the *CP* parameters, the most important factor is the amount of signal candidates. Therefore, the BDT classifier cut is optimised with respect to $B_s^0 \to D_s^\mp K^\pm$ data as follows. First, the selection requirements described in Sec. 7.1 and 7.2 are applied. Then, the data is repeatedly fit with the fit procedure that is explained in Chap. 8, while different cuts on the BDT classifier are applied. For each

fit and respective BDT classifier cut value the signal significance is calculated, which is given by

$$S_{\text{Sig}} = \frac{N_{\text{Sig}}}{\sqrt{N_{\text{Sig}} + N_{\text{Bkg}}}} \,. \tag{7.1}$$

It connects the number of signal candidates N_{Sig} and the sum of all background yields N_{Bkg}. Both values originate from the fit, but are taken from the reduced signal mass region of $[5320, 5420]\,\text{MeV}/c^2$. The signal significance can be understood as the number of standard deviations that the signal yield is larger than zero and is an often used figure of merit in the optimisation of selections. In order to favour large signal yields even more, the signal significance can be multiplied with the signal efficiency. Figure 7.5 shows the latter figure of merit together with its two components in the BDT response range $\in [-1, 0.6]$. All D_s^- final states are taken into account. The optimisation curve shows a plateau starting at a classifier value of 0.1, which is why it is chosen as the nominal cut value. Figure 7.6 depicts the distributions of the reconstructed B_s^0 mass and decay time for $B_s^0 \to D_s^{\mp} K^{\pm}$ decays after applying the full offline selection. In the invariant B_s^0 mass, the larger peak are $B_s^0 \to D_s^{\mp} K^{\pm}$ signal decays, while the smaller peak is due to $B^0 \to D_s^- K^+$ decays. In Fig. 7.7, the B_s^0 mass and time distributions are shown for $B_s^0 \to D_s^- \pi^+$ decays. In comparison to Fig. 7.1, which shows the data after the stripping, the strong combinatorial background suppression of the BDT is evident in Figs. 7.6 and 7.7. The same BDT cut value is applied for all selected decays, being $B_s^0 \to D_s^{\mp} K^{\pm}$, $B_s^0 \to D_s^- \pi^+$, and $B^0 \to D^- \pi^+$.
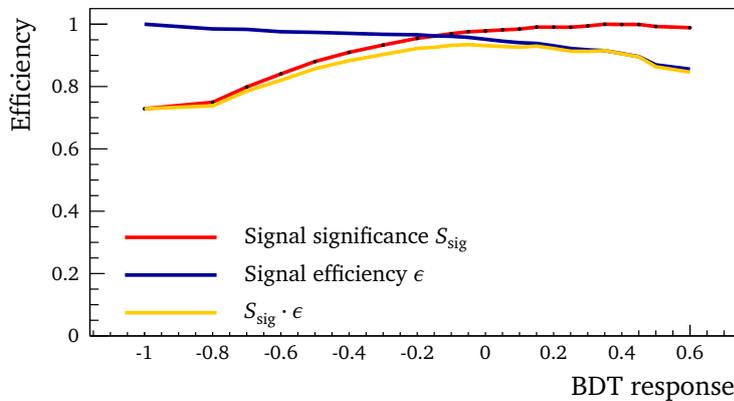

**Figure 7.5:** Optimisation curve and its components in the range of the BDTG response.
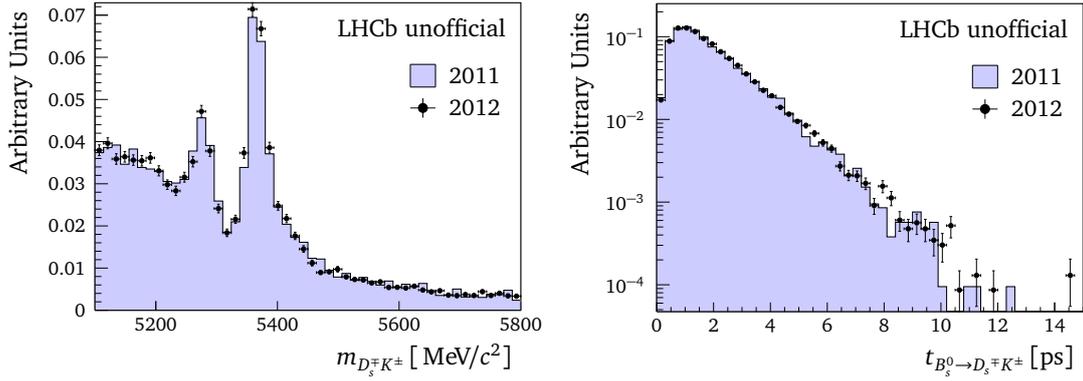
Figure 7.6: Normalised distributions of the invariant $D_s^{\mp}K^{\pm}$ mass (left) and decay time (right) of $B_s^0 \to D_s^{\mp}K^{\pm}$ data after applying the full offline selection. The data from the two years of data taking are superimposed.
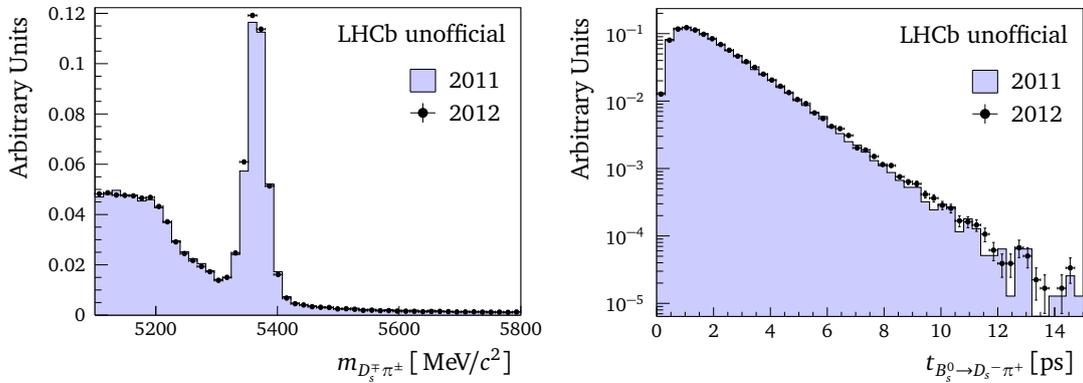
Figure 7.7: Normalised distributions of the invariant $D_s^{\mp}\pi^{\pm}$ mass (left) and decay time (right) of $B_s^0 \to D_s^{-}\pi^{+}$ data after applying the full offline selection. The data from the two years of data taking are superimposed.

7.5 Simulated data samples

Different signal and background decays are simulated and used in the course of the presented analysis. Table 7.7 lists the generated decay modes and the sample sizes. Moreover, for filtered MC the retention rate is given, which represents the fraction of events that pass the stripping requirements. Three signal MC samples include *CP* violation in their generation and are labelled as "CPV MC" in the table. These samples are used to perform a closure test of the *CP* fit (see Sec. 9.7.2). Their productions are based on MC configuration files that the author created for the presented analysis. Most of the other simulated samples are used to extract proper descriptions of signal and background distributions, labelled "Def. PDF" in the table. Furthermore, the $B_s^0 \to D_s^{-}\pi^{+}$ samples are needed for studies on the flavour tagging (see Sec. 9.5), while $B^0 \to D^{-}\pi^{+}$ MC helps to correct differences between data and MC differences (see Sec. 8.1).

Table 7.7: Simulated data samples used in the analysis.

Sample	Daughter decay	Sample size	Prod. info	Usage
$B_s^0 \to D_s^- \pi^+$	$D_s^- \to K^- K^+ \pi^-$	1 252 327	Filt. (1.9 %)	Def. PDF, FT
$B_s^0 \to D_s^- \pi^+$	$D_s^- \to K^- \pi^+ \pi^-$	238 978	Filt. (1.9 %)	Def. PDF, FT
$B_s^0 \to D_s^- \pi^+$	$D_s^- \to \pi^- \pi^+ \pi^-$	243 679	Filt. (1.9 %)	Def. PDF, FT
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to K^- K^+ \pi^-$	1 132 917	Filt. (3.5 %)	Def. PDF
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to K^- \pi^+ \pi^-$	217 758	Filt. (3.5 %)	Def. PDF
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to \pi^- \pi^+ \pi^-$	309 932	Filt. (3.5 %)	Def. PDF
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to K^- K^+ \pi^-$	5 121 168	CPV MC	Closure test
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to K^- \pi^+ \pi^-$	562 885	CPV MC	Closure test
$B_s^0 \to D_s^{\mp} K^{\pm}$	$D_s^- \to \pi^- \pi^+ \pi^-$	1 018 581	CPV MC	Closure test
$B_s^0 \to D_s^{*-} \pi^+$	$D_s^- \to K^- K^+ \pi^-$	2 101 381		Def. PDF
$B_s^0 \to D_s^- \rho^+$	$D_s^- \to K^- K^+ \pi^-$	2 067 161		Def. PDF
$\Lambda_b^0 \to D_s^- p$	$D_s^- \to K^- K^+ \pi^-$	506 641		Def. PDF
$\Lambda_b^0 \to D_s^{*-} p$	$D_s^- \to K^- K^+ \pi^-$	572 773		Def. PDF
$\Lambda_b^0 \to \Lambda_c^+ \pi^-$	$\Lambda_c^+ \to p K^- \pi^+$	490 014	Filt. (2.7 %)	Def. PDF
$\Lambda_b^0 \to \Lambda_c^+ K^-$	$\Lambda_c^+ \to p K^- \pi^+$	542 701		Def. PDF
$B^0 \to D^- \pi^+$	$D_s^- \to K^+ \pi^- \pi^-$	3 338 807	Filt. (3.5 %)	Def. PDF, Data/MC corr.
$B^0 \to D^- K^+$	$D_s^- \to K^+ \pi^- \pi^-$	277 392	Filt. (3.5 %)	Def. PDF

8 Fit to invariant mass and PID distributions

This chapter discusses the so-called multidimensional (MD) fit, which is an unbinned maximum likelihood fit to the B_s^0 mass, the D_s^- mass, and the PID distribution of the bachelor particle. It is performed simultaneously for the different D_s^- final states and independently for $B_s^0 \to D_s^\mp K^\pm$ and $B_s^0 \to D_s^- \pi^+$ decays. The MD fit is an *sPlot* fit (see Sec. 5.2). It computes the *sWeights* that allow to perform the decay-time fits, which are covered in the next chapter, in a statistically background subtracted way. Many of the PDF components in the MD fit rely on simulated data, thus, they might be affected by differences between data and MC. In order to minimise such disagreements, they are reweighted based on $B^0 \to D^- \pi^+$ data (see Sec. 8.1). Section 8.2 discusses the fit components and the definition of their PDFs. Afterwards, the results of the MD fits to $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ decays are summarised in the Sec. 8.3. Finally, the validation of the MD fit is outlined in Sec. 8.4.

8.1 Control mode $B^0 \to D^- \pi^+$

The $B^0 \to D^- \pi^+$ control mode serves two purposes in the presented analysis. Firstly, it is used to constrain two background yields in the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ MD fits. These two backgrounds are decays of $B^0 \to D^- K^+$ and those of the $B^0 \to D^- \pi^+$ channel itself. Secondly, weights are extracted from $B^0 \to D^- \pi^+$ data, which correct differences in kinematic observables between simulated and detector data. These corrections are necessary as the description of many MD fit components relies on templates from simulated data. Especially the PDFs used to describe the PID distributions cannot be established from first principles. Hence, the MC used to extract these PDFs needs to be carefully corrected for potential differences to the real data. The $B^0 \to D^- \pi^+$ decay is a good choice to extract these corrections. It is kinematically similar to $B_s^0 \to D_s^\mp K^\pm$ and $B_s^0 \to D_s^- \pi^+$ decays, which allows for a nearly identical selection (see Sec. 7.3). Furthermore, $f_s/f_d \approx 0.26$ [139] and the larger branching ratio of $B^0 \to D^- \pi^+$ result in a much larger sample of candidates. This section discusses the *sPlot* fit performed to the $m(D^- \pi^+)$ invariant mass distribution and the principle of the data/MC reweighting. The description is brief as the work has mainly been done by collaborators in the analysis group.

  The reweighting is performed independently for the two years of data taking and the two magnet polarities, which takes possible differences into account. Furthermore, this strategy allows for a consistency cross check, in which the four samples are

processed individually by the complete analysis chain (see Sec. 8.4). Figure 8.1 shows the invariant mass distributions and the respective fit projections of the samples. The signal is parametrised with the sum of two Crystal Ball (CB) functions [148] that share a common mean. A CB function corresponds to a Gaussian PDF that has a power law tail to one side. Here, tails in either direction of the invariant mass are used to take reconstruction and final state radiation effects into account. Except for the mean and a scale factor of the mass resolutions, all shape parameters of the CB functions are fixed to values obtained in $B^0 \to D^- \pi^+$ signal MC. The description of the background contributions $B^0 \to D^- K^+$, $B_s^0 \to D_s^- \pi^+$, $\Lambda_b^0 \to \Lambda_c^+ \pi^-$, $B^0 \to D^- \rho^+$, and $B^0 \to D^{*-} \pi^+$ is based on template PDFs, which consist of Gaussian kernel PDFs that are able to describe any arbitrary shape [100]. These PDFs, so-called `RooKeysPdf`, are extracted from dedicated MC samples. For the backgrounds described by `RooKeysPdf` tem-
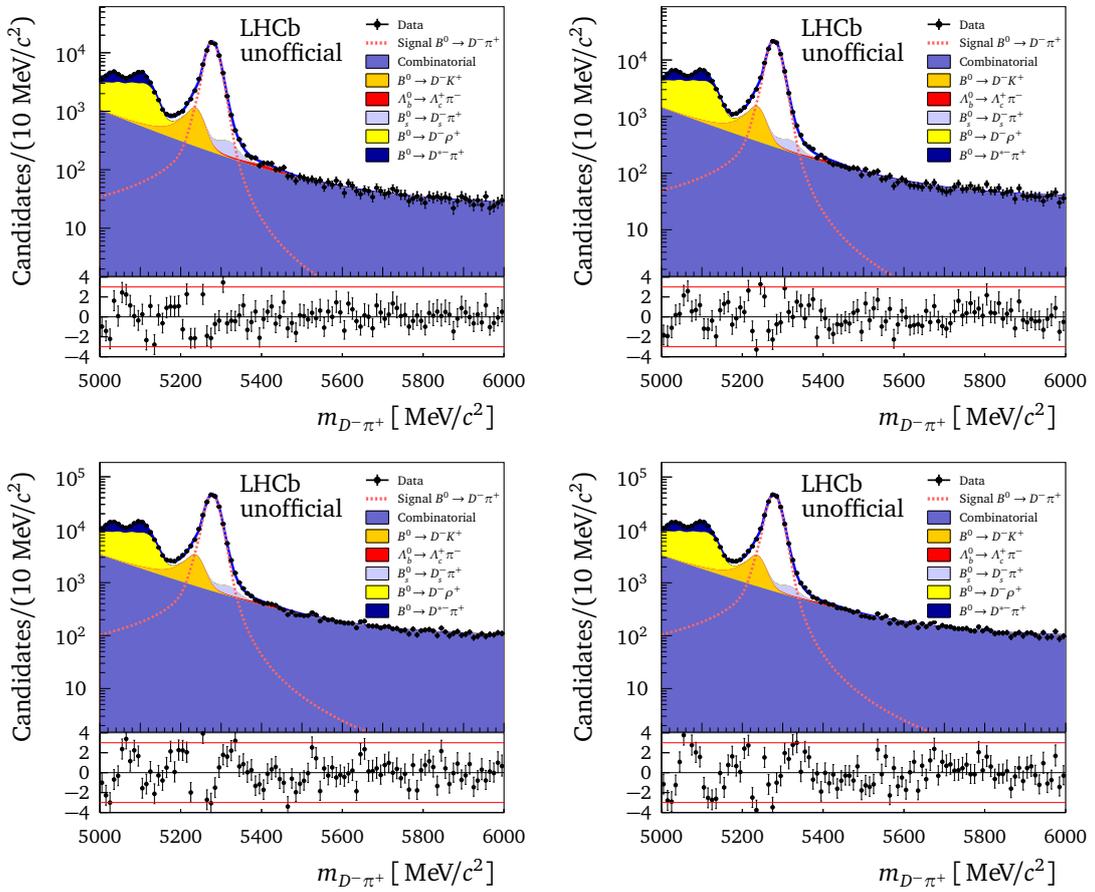
Figure 8.1: Distributions of the invariant $D^- \pi^+$ mass obtained in the fit to $B^0 \to D^- \pi^+$ candidates in data (data points). The plots are split by the year of data taking, 2011 (top row) and 2012 (bottom row), as well as by the magnet polarities, up (left) and down (right). The solid, blue curve corresponds to the fit described in the text. Different contributions to the fit are shown as coloured areas (for backgrounds) or a dashed line (for signal), as described in the legend.

plates, the only floating parameters are their respective yield parameters. The last component in the fit is the combinatorial background, which is parametrised by the sum of two exponential PDFs.

The *sWeights* allow to statistically subtract the background and thus, to use only $B^0 \to D^- \pi^+$ signal candidates for the correction of differences between simulated and detector data. Differences are known to exist especially for the PID information and the observables it depends on. These differences need to be corrected, as some of the components in the MD fit are described by templates taken directly from simulated data. In the presented analysis, a reweighting cannot be performed directly in the PID distribution of the bachelor particle, as it is one of the MD fit's dimensions. As the bachelor PID depends mostly on the bachelor momentum (p) and the track multiplicity (nTracks), a two-dimensional reweighting in these two observables is performed. Therefore, two-dimensional histograms in log(nTracks) and log(p) are created for the truth-matched $B^0 \to D^- \pi^+$ MC and the sWeighed data sample. Then, the correction matrix is calculated as the ratio of the data and the MC histograms. The logarithm is applied to both observables, as it leads to widened distributions and more equally filled bins. In each of the two observables Figure 8.2 shows the correction matrix for the two 2012 samples. Some bins get large corrections, which proves the necessity of the procedure. Furthermore, Fig. 8.3 shows a one-dimensional projection of log(nTracks) and log(p), superimposing the data, the unweighed MC, and the weighted MC. The distributions of the weighted MC matches the data much better than the unweighted MC.
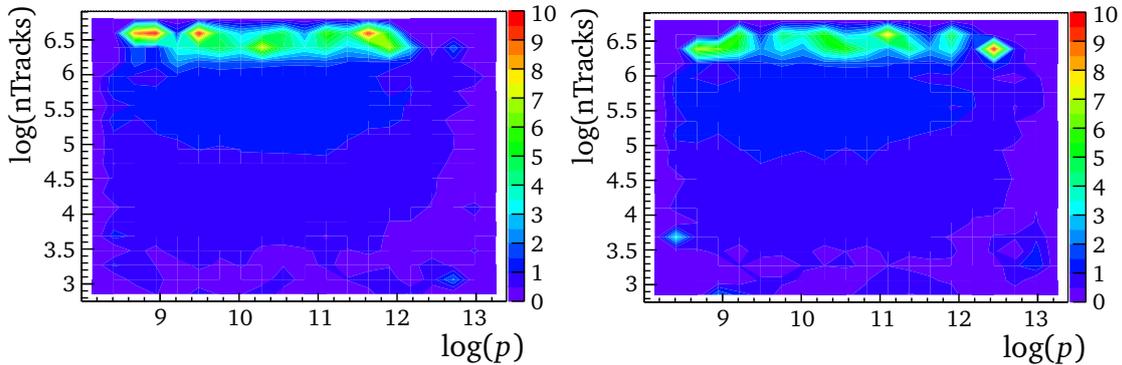

Figure 8.2: The two-dimensional data/MC ratios of the 2012 magnet up (left) and down (right) samples, obtained from $B^0 \to D^- \pi^+$.
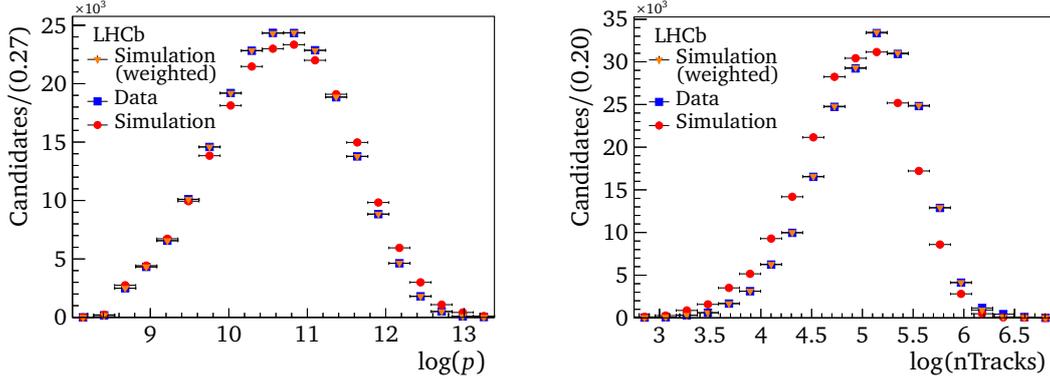
Figure 8.3: Projection of log(p) and log(nTracks) illustrating the two-dimensional data/MC weighting (only 2012, magnet polarity up).

8.2 Components

Due to the topological and geometrical similarities of the decays, the $B_s^0 \to D_s^\mp K^\pm$ and $B_s^0 \to D_s^- \pi^+$ MD fits comprise largely the same components and parameterisations. Thus, this section jointly describes the modelling of the components in the fits of the two decays. Besides the signal decays, the MD fit needs to take physical backgrounds, which occur due to mis-reconstructions or partial reconstructions, and the combinatorial background into account. The descriptions of the components in the invariant mass dimensions can mostly be taken directly from the simulated data samples, which have been reweighted by the procedure introduced in Sec. 8.1. However, parameterisations of the PID distribution need to be taken from detector data, as simulations do not reproduce them well enough. Thus, the PID PDFs are templates that originate from high statistics PID calibration samples, which themselves are taken from $D^{*\pm} \to D^0(K^\mp \pi^\pm)\pi^\pm$ decays (for pure kaon and pion samples) and $\Lambda_c^+ \to pK^-\pi^+$ decays (for the proton sample) [149]. Their backgrounds are subtracted using the *sPlot* technique. Furthermore, the offline PIDK requirements are applied to the calibration samples, which are PID$K < 0$ for $B_s^0 \to D_s^- \pi^+$ and PID$K > 5$ for $B_s^0 \to D_s^\mp K^\pm$ decays. In order to have compatible kinematics and event track multiplicity as in the signal decays, the PID control samples are reweighted by the same strategy as the MC using the *sWeighted* $B^0 \to D^- \pi^+$ data sample. This reweighting is done separately for each fit component as well as the two magnet polarities. In the data fits, no free parameter other then the normalisation are used for the PID PDFs.

8.2.1 Signal candidates

In the invariant mass dimensions, signal candidates are parameterised by the sum of two Crystal Ball functions [148]. The two summed CB functions, hereafter called double CB, feature a common mean and separate tails to lower and higher masses, which model non-Gaussian reconstruction effects and final state photon radiation. In

order to obtain values for the CBs' shape parameters, fits on simulated signal decays are performed separately for all D_s^- final states. These parameters are fixed in the data fits, with two exceptions. Firstly, the shared means of the CBs are floating in the fit. Secondly, as the resolutions of simulated and detector data do not match perfectly, for both, the D_s^- and the B_s^0 invariant mass, scale parameters R_{B_s} and R_{D_s} are floated in the fit to the $B_s^0 \rightarrow D_s^- \pi^+$ data to correct the widths of the Gaussians. In the subsequent fit to the $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ candidates, these scale parameters are fixed to their respective values obtained from the $B_s^0 \rightarrow D_s^- \pi^+$ data fit. In the bachelor PID dimension the signal component is modelled from pure calibration samples, originating from $D^{*\pm}$ decays, reweighted by the nominal procedure.

8.2.2 Fully reconstructed backgrounds

Fully reconstructed background decays, which are the dominant source of physical backgrounds, are the result of misidentification in the final state, or originate from B^0 decays. All background decays of this type that are considered in the $B_s^0 \rightarrow D_s^- \pi^+$ and $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ MD fits are listed in Tab. 8.1. The table also gives the respective misidentifications that lead to the occurrence of the backgrounds. The description of the fully reconstructed backgrounds in the $m(D_s^{\mp} h^{\pm})$ mass depends on the final state. If the final state is the same as the respective signal decay, the signal double Crystal Ball function is used. However, the mean is shifted according to the known mass difference $m(B_s^0) - m(B) = 86.9 \, \text{MeV}/c^2$ [24]. Moreover, the widths of the CBs are corrected by the width ratio $B_s^0 \rightarrow D_s^{\mp} h^{\pm} / B^0 \rightarrow D_s^{\mp} h^{\pm}$, which is measured on MC samples. All other fully reconstructed backgrounds are described by `RooKeysPdf` templates, which are taken from reweighted simulated data. In the $m(h^- h^+ h^-)$ mass, decays with true D_s^- mesons are parameterised by the same PDF as the signal component. Backgrounds with D^- or $\overline{\Lambda}_c^-$ mesons are described by templates from simulated data. Finally, in the PID distribution the fully reconstructed backgrounds are described by predefined shapes without free parameters, which were taken from the reweighted PID calibration samples. If the final state is identical to the one of the signal, the signal PID PDF is utilised.

For each fully reconstructed background an expected number of candidates in the $B_s^0 \rightarrow D_s^- \pi^+$ fit is calculated. These expectations are calculated similarly to the previous analysis [75, 150]. The $B^0 \rightarrow D^- \pi^+$ yield is extracted from the control sample (see Sec. 8.1), taking the misidentification rate obtained from simulated $B^0 \rightarrow D^- \pi^+$ samples and the different selections into account. The number of $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ background candidates is estimated from a fit to a special $B_s^0 \rightarrow D_s^- \pi^+$ data sample. This sample is reconstructed under the $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$ hypothesis and has an adjusted selection, which mainly reverses the $\overline{\Lambda}_c^-$ veto. The $B_s^0 \rightarrow D_s^{\mp} K^{\pm}$ background yield is computed through a factor that accounts for the relative branching fractions, the misidentification rate, and the efficiency of the PID cut. Further background yields are based on known branching fractions and relative efficiencies measured on simulated data. All expected background yields that are below 2 % of the signal yield are fixed in

Table 8.1: Fully reconstructed backgrounds described in the MD fits of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates.

Decay	Background to	Reason of occurrence
$B^0 \to D_s^- \pi^+$	$B_s^0 \to D_s^- \pi^+$	Same final state
$B_s^0 \to D_s^- K^+$	$B_s^0 \to D_s^- \pi^+$	$K^+ \to \pi^+$ (bachelor)
$B^0 \to D^- \pi^+ \to (K^+\pi^-\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (K^+K^-\pi^-)\pi^+$	$\pi^- \to K^-$ (D^-)
$B^0 \to D^- \pi^+ \to (K^+\pi^-\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (\pi^+K^-\pi^-)\pi^+$	$K^+ \to \pi^+$ and
		$\pi^- \to K^-$ (D^-)
$\overline{\Lambda_b^0} \to \overline{\Lambda_c^-} \pi^+ \to (\overline{p}K^+\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (K^+K^-\pi^-)\pi^+$	$\overline{p} \to K^-$ $(\overline{\Lambda_c^-})$
$\overline{\Lambda_b^0} \to \overline{\Lambda_c^-} \pi^+ \to (\overline{p}K^+\pi^-)\pi^+$	$B_s^0 \to D_s^- \pi^+ \to (\pi^+K^-\pi^-)\pi^+$	$\overline{p} \to K^-$ $(\overline{\Lambda_c^-})$ and
		$K^+ \to \pi^+$ $(\overline{\Lambda_c^-})$
$B^0 \to D_s^- K^+$	$B_s^0 \to D_s^- K^+$	Same final state
$B_s^0 \to D_s^- \pi^+$	$B_s^0 \to D_s^- K^+$	$\pi^+ \to K^+$ (bachelor)
$\Lambda_b^0 \to D_s^- p$	$B_s^0 \to D_s^- K^+$	$p \to K^+$ (bachelor)
$B^0 \to D^- K^+ \to (K^+\pi^-\pi^-)K^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\pi^- \to K^-$ (D^-)
$B^0 \to D^- \pi^+ \to (K^+\pi^-\pi^-)\pi^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\pi^- \to K^-$ (D^-) and
		$\pi^+ \to K^+$ (bachelor)
$\overline{\Lambda_b^0} \to \overline{\Lambda_c^-} K^+ \to (\overline{p}K^+\pi^-)K^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\overline{p} \to K^-$ $(\overline{\Lambda_c^-})$
$\overline{\Lambda_b^0} \to \overline{\Lambda_c^-} \pi^+ \to (\overline{p}K^+\pi^-)\pi^+$	$B_s^0 \to D_s^- K^+ \to (K^+K^-\pi^-)K^+$	$\overline{p} \to K^-$ $(\overline{\Lambda_c^-})$ and
		$\pi^+ \to K^+$ (bachelor)

the fit to the $B_s^0 \to D_s^- \pi^+$ data, while the others are fixed within gaussian constraints. Fully floating background yields would cause the fit to become unstable, due to the similarities to the signal in the different dimensions. Furthermore, in the $B_s^0 \to D_s^{\mp} K^{\pm}$ fit all fully reconstructed background yields are fixed. The values are based on those calculated for or obtained in the $B_s^0 \to D_s^- \pi^+$ fit. A factor $1/15$ is used to correct the relative branching fraction of the modes $B^0 \to D^- K^+$ and $B^0 \to D^- \pi^+$ [24]. Moreover, the PID efficiencies are corrected according to the different bachelor PID cuts in the two decays.

8.2.3 Partially reconstructed backgrounds

Partially reconstructed decays typically peak in the lower $m(D_s^{\mp}h^{\pm})$ mass sideband as the energy of one or more particles are missing. However, these backgrounds may extend into the region around the B_s^0 meson mass, which is why they need to be parameterised accurately. In order to do so, it is important to define the lower $m(D_s^{\mp}h^{\pm})$ mass boundary in a way that a description is still possible, *i.e.* to not cut too close to the nominal B_s^0 mass. Table 8.2 lists all partially reconstructed backgrounds included in the MD fits. Decays of $B_s^0 \to D_s^{*-} \pi^+$, in which a photon is not reconstructed, are considered for both $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ decays. Additionally, decays of $B_s^0 \to D_s^- \rho^+$ and of $\Lambda_b^0 \to D_s^{*-} p$ are taken into account in the $B_s^0 \to D_s^{\mp} K^{\pm}$ MD

Table 8.2: Partially reconstructed backgrounds described in the MD fits of $B_s^0 \to D_s^\mp K^\pm$ and $B_s^0 \to D_s^- \pi^+$ candidates.

Decay	Background to	Not reconstructed particle and misidentification
$B_s^0 \to D_s^{*-} \pi^+$	$B_s^0 \to D_s^- \pi^+$	photon from $D_s^{*-} \to D_s^- \gamma$, no misidentification
$B_s^0 \to D_s^- \rho^+$	$B_s^0 \to D_s^- K^+$	π^0 from $\rho^+ \to \pi^+ \pi^0$ and $\pi^+ \to K^+$ (bachelor)
$B_s^0 \to D_s^{*-} \pi^+$	$B_s^0 \to D_s^- K^+$	photon from $D_s^{*-} \to D_s^- \gamma$ and $\pi^+ \to K^+$ (bachelor)
$\Lambda_b^0 \to D_s^{*-} p$	$B_s^0 \to D_s^- K^+$	photon from $D_s^{*-} \to D_s^- \gamma$ and $p \to K^+$ (bachelor)

fit. The former occurs, when the π^0 is not reconstructed from the $\rho^+ \to \pi^+ \pi^0$ decay and the charged pion is misidentified as a kaon. In the latter decay, the photon from $D_s^{*-} \to D_s^- \gamma$ is missed and the proton is reconstructed as a kaon. In the invariant $m(D_s^\mp h^\pm)$ mass distribution all mentioned partially reconstructed backgrounds are described by templates extracted from simulations, while for the $m(h^- h^+ h^-)$ mass the respective signal PDFs are used. Depending on the bachelor particle type, the PID PDFs are based on the proton or the pion calibration sample.

8.2.4 Combinatorial background

Combinatorial background is the flat, non-peaking component in the $m(D_s^\mp h^\pm)$ distribution. It does not contain true B_s^0 meson decays, but consists of randomly selected particle tracks or true D_s^- meson decays to which an arbitrary bachelor particle is assigned. Combinatorial background is parameterised in the $m(D_s^\mp h^\pm)$ dimension by the sum of an exponential function and a constant. This parametrisation is developed and tested on pure combinatorial candidates from the upper mass sideband $m(D_s^\mp h^\pm) > 5800 \, \text{MeV}/c^2$. The slope of the exponential function and a fraction between the two components are free parameters in the MD fits. In the $m(h^- h^+ h^-)$ mass, combinatorial background is parametrised by the sum of an exponential function and a double Crystal Ball. The latter is added to take decays of true D_s^- mesons into account. A fraction between the two components of the PDF is left floating in the fit, while the shape parameters of the double CB are fixed to the same values used for signal candidates. In case of the $D_s^- \to K^- K^+ \pi^-$ final state, the double CB is the dominant proportion of the combinatorial background. The PID distribution of the combinatorial background can not easily be described, as it contains kaons, pions, and protons, which all have distinctive shapes. Thus, the individual PDFs of the three bachelor particle types are summed and the fractions are left floating in the fit. The PDFs are taken from the $D^{*\pm}$ and $\overline{\Lambda}_c^-$ calibration samples, which have been weighted to match the combinatorial background in terms of the kinematic properties and the track multiplicity. This weighting is done as follows. A pure combinatorial background sample, which originates from the upper mass sideband $m(D_s^\mp h^\pm) \in [5600, 6800] \, \text{MeV}/c^2$, is partitioned in 10 bins. The p_T and nTracks distributions are fitted in each bin using a Landau function [151]. From the out-

put of these fits the p_T and nTracks distributions are extrapolated into the fit region $m(D_s^\mp h^\pm) \in [5300, 5800]$ MeV/c^2 and then used to reweight the calibration samples. This reweighting is done individually for the D_s^- final states, as the different PID selections lead to different kinematic distributions.

8.3 Results

The MD fits are performed on the full $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ samples, *i.e.* the magnet polarities and years of data taking are combined. Moreover, the D_s^- final state are fit simultaneously. The ranges of the invariant mass dimensions are in both fits set to the boundaries given in Tab. 7.3. However, the fit range of the PID distribution differs between the two decays. In the $B_s^0 \to D_s^- \pi^+$ fit, it is limited to the range $\log(|\mathrm{PID}K|) \in [-7, 5]$, while in case of $B_s^0 \to D_s^\mp K^\pm$ it is $\log(\mathrm{PID}K) \in [\ln(5), \ln(150)]$. Figure 8.4 shows the projections of the three MD fit dimensions from the fit of the $B_s^0 \to D_s^- \pi^+$ candidates, while Fig. 8.5 contains the same for the $B_s^0 \to D_s^\mp K^\pm$ sample; the data is well described by the fitted PDFs. The Tabs. 8.3 and 8.4 contain the fit results of the floating parameters obtained in the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ MD fits, respectively. The fit results prove that the selection, in particular the newly trained BDT, performs very well. Compared to the previous 1 fb^{-1} analysis [75], the number of $B_s^0 \to D_s^- \pi^+$ signal candidates increases by more than a factor of three (96942 compared to 28260), while there are less than three times as many combinatorial background candidates (25789 compared to 9030). More importantly, the same holds for $B_s^0 \to D_s^\mp K^\pm$ decays, where the presented analysis observes 5955 signal candidates and 6100 combinatorial background candidates, while the previous analysis found 1768 signal candidates and 2430 combinatorial background candidates. These values give rise to the expectation that the presented analysis gives a significant improvement in terms of statistical uncertainties on the *CP* parameters with respect to the previous analysis.
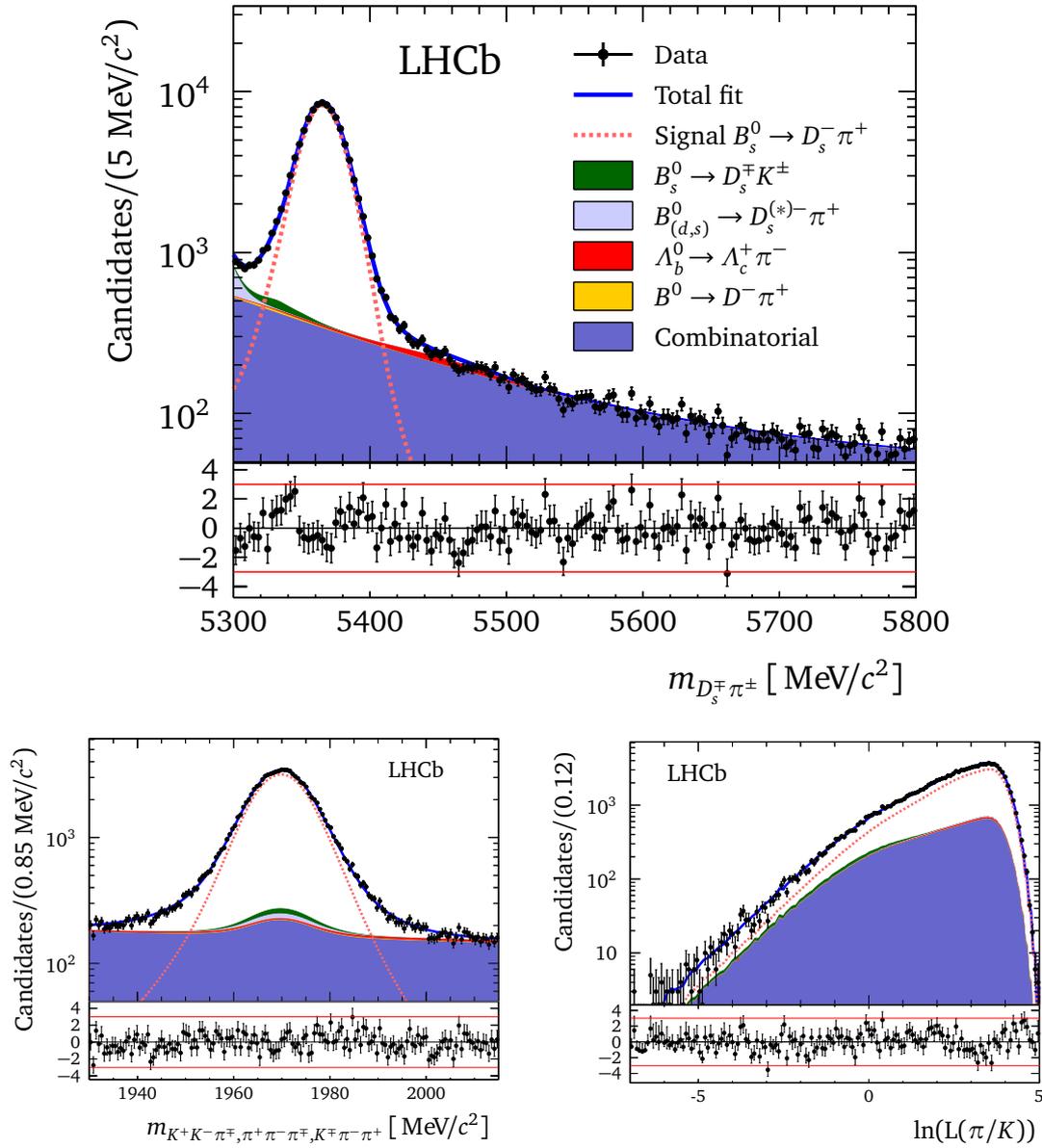
Figure 8.4: Projections of the three MD fit dimensions from the fit to the $B_s^0 \to D_s^- \pi^+$ candidates. From top to bottom right: $m(D_s^\mp h^\pm)$ invariant mass, $m(h^- h^+ h^-)$ invariant mass, $\log(|\text{PID}K|)$ variable. Both magnet polarities, both years of data taking, and all D_s^- final states are combined. Different components of the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend of the plot at the top.
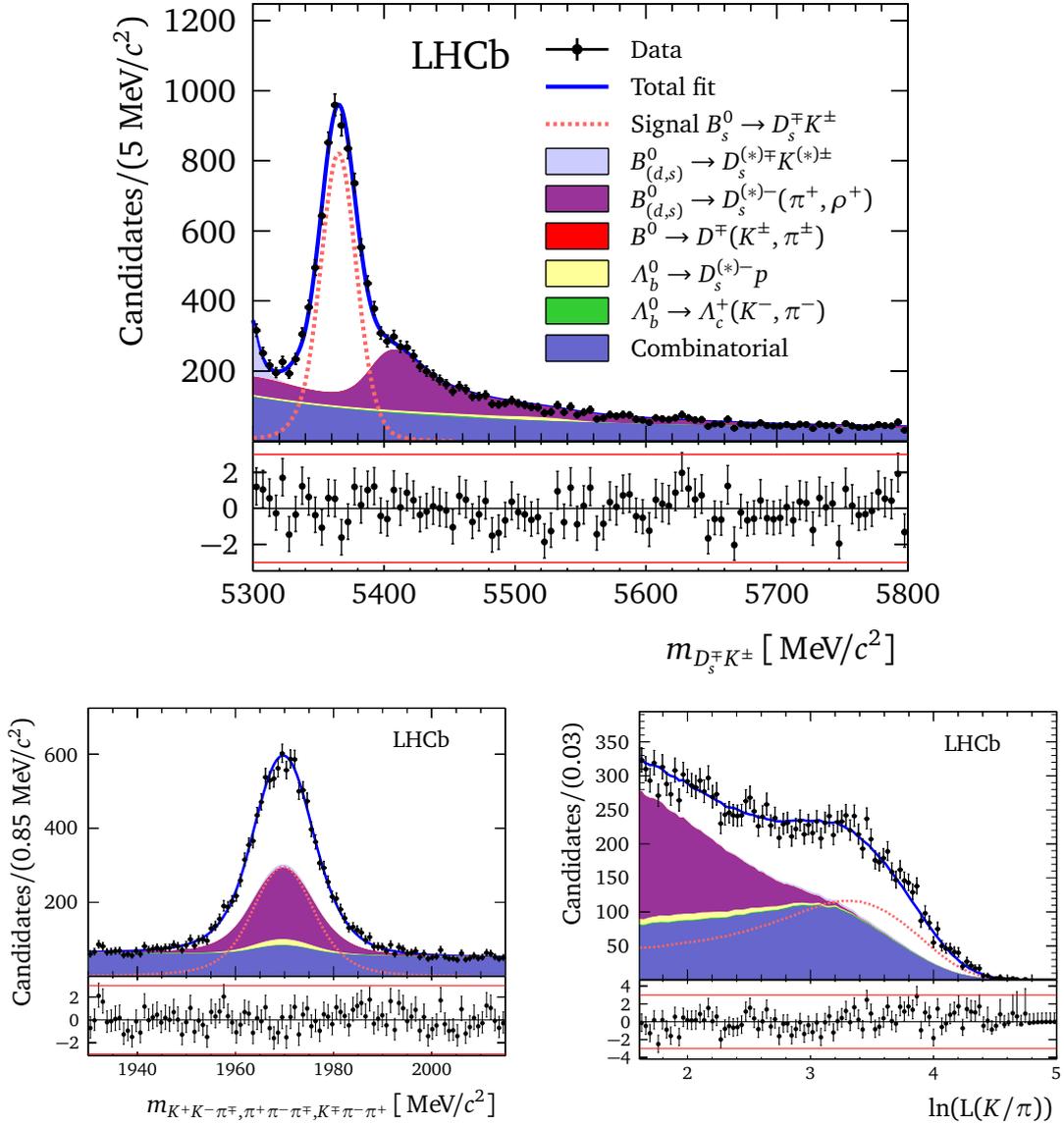
Figure 8.5: Projections of the three MD fit dimensions from the fit to the $B_s^0 \to D_s^\mp K^\pm$ candidates. From top to bottom right: $m(D_s^\mp h^\pm)$ invariant mass, $m(h^- h^+ h^-)$ invariant mass, $\log(|\mathrm{PID}K|)$ variable. Both magnet polarities, both years of data taking, and all D_s^- final states are combined. Different components of the fit are shown as coloured areas (for backgrounds) or dashed line (for signal) as described in the legend of the plot at the top.

Table 8.3: Fit results of the $B_s^0 \to D_s^- \pi^+$ MD fit. The parameter μ_{B_s} and μ_{D_s} are the means of the double Crystal Balls used to describe the signal in the invariant B_s^0 and D_s mass distributions, respectively. The parameters β_{B_s} and β_{D_s} are the slopes of the exponential function that parameterises the combinatorial background in the invariant mass dimensions, given in $(\mathrm{MeV}/c^2)^{-1} \cdot 10^{-3}$ units. The fraction f_{Comb,B_s} is the fraction between the exponential and the constant function in the B_s^0 mass, whereas f_{Comb,D_s} is the fraction between the signal shape and the exponential in the invariant D_s mass. The $f_{\mathrm{Comb},\mathrm{PID}K}$ is the fraction between pion and kaon components in the PIDK combinatorial shape.

| Parameter	$D_s^- \to K^-K^+\pi^-$ non resonant	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$N_{B_s \to D_s\pi}$	$16\,056 \pm 145$	$34\,355 \pm 201$	$25\,596 \pm 173$	5728 ± 86	$15\,206 \pm 145$
N_{partReco}	87 ± 25	215 ± 31	168 ± 31	38 ± 14	94 ± 24
N_{Comb}	9185 ± 123	3116 ± 103	3769 ± 93	2765 ± 65	6952 ± 109
Tot. $N_{B_s \to D_s\pi}$			$96\,942 \pm 345$		
Tot. N_{partReco}			643 ± 59		
Tot. N_{Comb}			$25\,789 \pm 233$		
μ_{B_s} [MeV/c^2]			5365.10 ± 0.06		
R_{B_s}	1.082 ± 0.010	1.082 ± 0.006	1.082 ± 0.007	1.077 ± 0.016	1.070 ± 0.010
μ_{D_s} [MeV/c^2]			1969.80 ± 0.02		
R_{D_s}	1.040 ± 0.009	1.056 ± 0.006	1.053 ± 0.006	1.047 ± 0.015	1.049 ± 0.010
β_{B_s}	-7.35 ± 0.28	-9.74 ± 0.60	-9.22 ± 0.53	-6.28 ± 0.99	-4.70 ± 0.53
f_{Comb,B_s}	0.86 ± 0.02	0.75 ± 0.03	0.75 ± 0.03	0.52 ± 0.06	0.70 ± 0.06
β_{D_s}	-0.29 ± 0.46	-4.66 ± 0.99	-4.14 ± 0.78	-1.23 ± 0.84	-4.13 ± 0.53
f_{Comb,D_s}	0.98 ± 0.01	0.73 ± 0.02	0.89 ± 0.02	0.95 ± 0.02	0.97 ± 0.02
$f_{\mathrm{Comb},\mathrm{PID}K}$	0.86 ± 0.01	0.76 ± 0.02	0.81 ± 0.02	1.00 ± 0.01	1.00 ± 0.01

Table 8.4: Fit results of the $B_s^0 \to D_s^{\mp} K^{\pm}$ MD fit. Group23 denotes $B_{(d,s)}^0 \to D_s^{(*)-}(\pi^+, \rho^+)$ and $\Lambda_b^0 \to D_s^{(*)-}p$ background decays. The parameter μ_{B_s} and μ_{D_s} are the means of the double Crystal Balls used to describe the signal in the invariant B_s^0 and D_s mass distributions, respectively. The parameters β_{B_s} and β_{D_s} are the slopes of the exponential function that parameterises the combinatorial background in the invariant mass dimensions, given in $(\mathrm{MeV}/c^2)^{-1} \cdot 10^{-3}$ units. The fraction f_{Comb,B_s} is the fraction between the exponential and the constant function in the B_s^0 mass, whereas f_{Comb,D_s} is the fraction between the signal shape and the exponential in the invariant D_s mass. The parameters $f_{\mathrm{Comb,PID}K1}$ and $f_{\mathrm{Comb,PID}K2}$ are the fractions between pion, kaon and proton components in the PIDK combinatorial shape.

Parameter	$D_s^- \to K^-K^+\pi^-$ non resonant	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$N_{B_s \to D_s K}$	1055 ± 38	1957 ± 51	1616 ± 46	391 ± 24	936 ± 37
$N_{B \to D_s K}$	32 ± 10	49 ± 13	50 ± 12	24 ± 7	41 ± 211
N_{Group23}	676 ± 36	1513 ± 57	1062 ± 50	208 ± 23	763 ± 43
N_{Comb}	1370 ± 47	736 ± 55	707 ± 53	1113 ± 40	2173 ± 59
Tot. $N_{B_s \to D_s K}$			5955 ± 90		
Tot. $N_{B \to D_s K}$			196 ± 42		
Tot. N_{Group23}			4223 ± 97		
Tot. N_{Comb}			6100 ± 117		
μ_{B_s} [MeV/c^2]			5365.2 ± 0.2		
μ_{D_s} [MeV/c^2]			1969.7 ± 0.1		
β_{B_s}	-6.8 ± 1.0	-12.2 ± 0.9	-7.9 ± 1.5	-1.40 ± 0.24	-1.63 ± 0.18
f_{Comb,B_s}	0.64 ± 0.07	0.37 ± 0.05	0.63 ± 0.08	not fitted	not fitted
β_{D_s}	-0.4 ± 0.1	-7.6 ± 2.1	-1.6 ± 1.9	-1.0 ± 1.2	-3.8 ± 0.9
f_{Comb,D_s}	0.99 ± 0.09	0.61 ± 0.05	0.72 ± 0.05	1.00 ± 0.07	1.00 ± 0.02
$f_{\mathrm{Comb,PID}K1}$	0.28 ± 0.04	0.12 ± 0.07	0.31 ± 0.07	0.01 ± 0.04	0.15 ± 0.03
$f_{\mathrm{Comb,PID}K2}$	0.77 ± 0.08	0.55 ± 0.09	0.93 ± 0.16	0.66 ± 0.07	0.60 ± 0.05

8.4 Validation

The robustness of the MD fit is tested by splitting the full data set into independent subsamples and repeating the complete procedure for both decays. The subsamples are created by splitting the full sample corresponding to

1. the two magnet polarities (up and down),

2. the two years of data taking (2011 and 2012),

3. low $(0.10 - 0.95)$ and high $(0.95 - 1.00)$ values of the BDT classifier,

4. low $(p_{B_s^0} < 120 \, \text{GeV/c})$ and high $(p_{B_s^0} \geq 120 \, \text{GeV/c})$ B_s^0 momenta.

The latter two splits are chosen in a way that they result in roughly the same number of signal candidates in the two samples. Moreover, the split by BDT classifier leads to one sample containing almost all combinatorial background candidates, while the other contains mostly signal candidates. This split enables to test if the signal and background shapes are obtained in a robust way. In all of the subsamples, a converging MD fit is attained. The signal yields that are obtained in the different subsamples are listed for $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates in Tabs. 8.5 and 8.6, respectively. Summing the yields from the fits to the different subsamples gives compatible results with each other and those on the full sample, which confirms the robustness of the MD fit procedure. The subsample containing only the 2011 data allows for a comparison with the previous analysis [75]. The new selection improves the signal to background ratio from 3.13 to 3.99 in case of $B_s^0 \to D_s^- \pi^+$ candidates, while for $B_s^0 \to D_s^{\mp} K^{\pm}$ it increases from 0.73 to 1.00. This is again proving the good performance of the newly trained BDT, as the signal yield remains very similar, while maintaining a better rejection of combinatorial background.

Further studies regarding the robustness of the MD fit can be found later in Sec. 9.7, where the systematic uncertainties on the measured *CP* parameters are discussed. Specifically, large amounts of pseudo-experiments are generated and used to repeat the whole fitting procedure including the MD fit (see Sec. 9.7.1). Furthermore, the influence of ignoring correlations among the MD fit observables is investigated, as summarised in Sec. 9.7.5.

Table 8.5: Signal yields in the different $B_s^0 \to D_s^- \pi^+$ subsamples.

Subsamples	$D_s^- \to K^- K^+ \pi^-$ non resonant	$D_s^- \to \phi \pi^-$	$D_s^- \to K^{*0} K^-$	$D_s^- \to K^- \pi^+ \pi^-$	$D_s^- \to \pi^- \pi^+ \pi^-$
Up	7569 ± 93	$16\,051 \pm 132$	$12\,077 \pm 114$	2646 ± 55	7250 ± 91
Down	8494 ± 99	$18\,322 \pm 140$	$13\,529 \pm 120$	3084 ± 59	7962 ± 96
2011	4741 ± 74	$10\,485 \pm 105$	7497 ± 90	1619 ± 43	4337 ± 71
2012	$11\,325 \pm 114$	$23\,893 \pm 161$	$18\,117 \pm 140$	4113 ± 69	$10\,882 \pm 112$
low BDT	8517 ± 102	$17\,347 \pm 138$	$14\,470 \pm 126$	3074 ± 60	8472 ± 101
high BDT	7365 ± 90	$16\,920 \pm 132$	$10\,967 \pm 107$	2605 ± 52	6595 ± 84
low $p_{B_s^0}$	$10\,015 \pm 107$	$19\,781 \pm 144$	$15\,652 \pm 129$	3399 ± 62	8029 ± 94
high $p_{B_s^0}$	6098 ± 85	$14\,672 \pm 127$	9976 ± 104	2302 ± 52	7244 ± 93
Up+Down			$96\,985 \pm 345$		
2011+2012			$97\,010 \pm 345$		
low+high BDT			$96\,334 \pm 342$		
low+high $p_{B_s^0}$			$97\,170 \pm 346$		

Table 8.6: Signal yields in the different $B_s^0 \to D_s^\mp K^\pm$ subsamples.

Subsamples	$D_s^- \to K^- K^+ \pi^-$ non resonant	$D_s^- \to \phi \pi^-$	$D_s^- \to K^{*0} K^-$	$D_s^- \to K^- \pi^+ \pi^-$	$D_s^- \to \pi^- \pi^+ \pi^-$
Up	516 ± 27	932 ± 36	772 ± 32	187 ± 16	429 ± 25
Down	534 ± 26	1035 ± 37	845 ± 32	205 ± 17	507 ± 28
2011	294 ± 20	577 ± 27	478 ± 25	116 ± 13	322 ± 21
2012	759 ± 32	1379 ± 43	1138 ± 38	275 ± 20	615 ± 31
low BDT	566 ± 29	1004 ± 38	906 ± 35	223 ± 19	542 ± 31
high BDT	468 ± 24	940 ± 34	691 ± 30	159 ± 14	376 ± 21
low $p_{B_s^0}$	795 ± 35	1389 ± 44	1112 ± 39	286 ± 23	598 ± 31
high $p_{B_s^0}$	261 ± 20	541 ± 28	485 ± 27	124 ± 14	320 ± 26
Up+Down			5962 ± 91		
2011+2012			5953 ± 90		
low+high BDT			5873 ± 90		
low+high $p_{B_s^0}$			5911 ± 95		

9 Measurement of *CP* Violation in $B_s^0 \to D_s^{\mp} K^{\pm}$ Decays

The *CP* observables introduced in Eq. 3.41 are obtained by fitting the decay-time-dependent decay rates $\Gamma(t)$ of the B_s^0 and $\bar{B}_s^0$ mesons (see Eqs. 3.43–3.46). This unbinned maximum likelihood fit is performed in a statistically background subtracted way, as the *sWeights* extracted from the multidimensional fit (see Chap. 8) are applied to the candidates. Thus, only the signal component has to be taken into account in the decay-time PDF. Section 9.1 describes this signal PDF before any experimental imperfections are included. The following Sec. 9.2 explains how production and detection asymmetries are introduced into the decay-time PDF. Another experimental effect that needs to be considered is the limited decay-time resolution. It is determined with prompt decays, as described in Sec. 9.3. The final nuisances of the fit, which are the decay-time-dependent efficiency and the flavour-tagging calibration, are obtained from decay-time fits to $B_s^0 \to D_s^- \pi^+$ signal decays. The former requires additional input from simulations (see Sec. 9.4), while the flavour tagging is calibrated solely with $B_s^0 \to D_s^- \pi^+$ data, as explained in Sec. 9.5. The result of the *CP* fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal candidates is given in Sec. 9.6, while the evaluation of systematic uncertainties is summarised in Sec. 9.7. From the measured *CP* parameters the CKM angle γ is determined. This is summarised in Sec. 9.8.

9.1 Decay-time PDF without detector effects

The PDF used to describe the decay-time distribution of the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal candidates can be derived directly from the decay-rate equations given in the equations 3.43 to 3.46. Omitting all experimental effects and normalisations, the PDF is given by

$$
\begin{aligned}
P_{D_s K}^{\text{theo}}(t,d,q) \propto (1+|\lambda_f|^2)\, e^{-\Gamma_s t} \Bigg(& \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) - d\, C_f \cos(\Delta m_s t) \\
& + \delta_{q,1}\,\left(A_f^{\Delta\Gamma_s} \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + d\, S_f \sin(\Delta m_s t) \right) \\
& + \delta_{q,-1}\,\left(A_{\bar{f}}^{\Delta\Gamma_s} \sinh\left(\frac{\Delta\Gamma_s t}{2}\right) + d\, S_{\bar{f}} \sin(\Delta m_s t) \right) \Bigg).
\end{aligned}
\tag{9.1}
$$

Here, t is the true decay time and d indicates the true production state, *i.e.* $d = 1$ in case of an initial B_s^0 and $d = -1$ for initial $\bar{B}_s^0$ mesons. The charge of the bachelor kaon in the final state is denoted by q, *i.e.* $q = 1$ is used for the final state f ($D_s^- K^+$)

and $q = -1$ for the conjugated final state $\bar{f}$ ($D_s^+ K^-$). The Kronecker deltas $\delta_{q,1}$ and $\delta_{q,-1}$ allow to unite the PDFs of the two final states.

As $B_s^0 \to D_s^- \pi^+$ is a flavour-specific decay, *i.e.* B_s^0 decays only to $D_s^- \pi^+$ and $\bar{B}_s^0$ to $D_s^+ \pi^-$, there is no *CP* violation in the interference between decay and decay after mixing. Hence, a simplified PDF is used to describe the decay time of the $B_s^0 \to D_s^- \pi^+$ candidates

$$P_{D_s \pi}^{\text{theo}}(t, d) \propto e^{-\Gamma_s t} \left(\cosh\left(\frac{\Delta\Gamma_s t}{2} \right) - d \cos(\Delta m_s t) \right). \tag{9.2}$$

The decay width Γ_s and the decay-width difference $\Delta\Gamma_s$ are fixed in all decay-time fits to the central values of the HFLAV averages [57]

$$\Gamma_s = (0.6643 \pm 0.0020)\ \text{ps}^{-1},$$
$$\Delta\Gamma_s = (0.083 \pm 0.006)\ \text{ps}^{-1}, \tag{9.3}$$
$$\rho(\Gamma_s, \Delta\Gamma_s) = -0.239.$$

Here, $\rho(\Gamma_s, \Delta\Gamma_s)$ is the correlation between Γ_s and $\Delta\Gamma_s$. The given uncertainties are used to determine systematic uncertainties with pseudo-experiments (see Sec. 9.7).

The theoretically motivated decay-time PDFs given in the Eqs. 9.1 and 9.2 need to be extended to take the experimental conditions into account. The different effects that need to be considered as well as their modelling are introduced in Secs. 9.2 to 9.5.

9.2 Asymmetries in production and detection

An intrinsic asymmetry I induced by the experimental setup shifts an observed time-dependent asymmetry compared to the true underlying asymmetry, *i.e.*

$$\mathscr{A}^{\text{meas}}(t) = \mathscr{A}(t) + I. \tag{9.4}$$

The intrinsic asymmetry can be understood as the sum of asymmetries from different sources. In the presented analysis, two sources of such asymmetries need to be considered, namely the production and the detection asymmetries.

In the *pp* collisions at the LHC, it is expected that B^0 and B^+ mesons are produced with an excess at the percent level compared to $\bar{B}^0$ and B^- mesons. The explanation are possible coalescences of perturbatively produced b or $\bar{b}$ quarks with u and d valence quarks present in the beam remnant. Consequently, this asymmetry is expected to be compensated by opposite production asymmetries of other b hadrons. For B_s^0 mesons, the production asymmetry is defined as

$$A_{\text{P}}(B_s^0) \equiv \frac{\sigma(\bar{B}_s^0) - \sigma(B_s^0)}{\sigma(\bar{B}_s^0) + \sigma(B_s^0)}. \tag{9.5}$$

Here, $\sigma(B_s^0)$ and $\sigma(\bar{B}_s^0)$ are the production cross-sections of B_s^0 and $\bar{B}_s^0$ mesons, respectively. In the fits, the production asymmetry is fixed corresponding to the LHCb measurement [152]

$$A_{\mathrm{p}}(B_s^0) = (1.1 \pm 2.7)\,\%. \tag{9.6}$$

A detection asymmetry describes different efficiencies ϵ in the reconstruction of charge-conjugated states f and $\bar{f}$. Thus, it is defined as

$$A_{\mathrm{D}} \equiv \frac{\epsilon_{\bar{f}} - \epsilon_f}{\epsilon_{\bar{f}} + \epsilon_f}. \tag{9.7}$$

Sources of the detection asymmetry are misalignments, inefficiencies in subdetector modules, or differing interaction cross-sections of charge-conjugated final state particles with the detector material. In the presented analysis an asymmetry of

$$A_{\mathrm{D}}(K^-\pi^+) = (1 \pm 1)\,\% \tag{9.8}$$

is assumed. This value is based on a measurement of A_{D} that was performed alongside the analysis of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays [153].

Compared to the theoretically motivated decay-time PDFs introduced in the Eqs. 9.1 and 9.2, the PDFs need to be modified to take A_{p} and A_{D} into account. This is done by multiplying corresponding terms with the PDF, *i.e.*

$$P(t,d,q) = (1 - d\,A_{\mathrm{p}})(1 - q\,A_{\mathrm{D}})\,P^{\mathrm{theo}}(t,d,q). \tag{9.9}$$

9.3 Decay-time resolution

The precision of measured decay times is limited due to the uncertainties present in the determination of momenta and in the localisation of vertices. Thus, a finite decay-time resolution exists, which leads to a dilution D_t of measured time-dependent oscillations

$$\mathscr{A}_f^{\mathrm{meas}}(t) = e^{-\frac{\Delta m_s^2 \sigma_t^2}{2}} \mathscr{A}_f(t) = D_t \mathscr{A}_f(t). \tag{9.10}$$

Here, a Gaussian resolution model is assumed with a width σ_t, while a deduction of a more general formula can be found in Ref. [154]. The dilution of the decay time resolution reduces the sensitivity of the measured CP parameters, independent from, but similar to the dilution D_ω implied by the imperfection of the flavour tagging (see Sec. 6.3). Thanks to the very good spatial and momentum resolutions of the VELO (see Sec. 4.2.1), the dilution factor D_t is greater than 99 % in case of B^0 mesons. Since the oscillation frequency Δm_s is about 35 times larger than Δm_d, the dilution is expected to be approximately 65 % for B_s^0 mesons. In order to avoid biasing the CP parameters, an accurate modelling of the decay-time resolution is required in the CP

fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ candidates. Therefore, the decay-time PDF is convolved with a Gaussian resolution model that features a different width for each candidate

$$P(t', d, q | \sigma_t) = (1 - dA_P)(1 - qA_D) P^{\text{theo}}(t, d, q) \otimes R(t' - t | \sigma_t). \qquad (9.11)$$

Here, t is the true decay time, t' is the measured decay time, σ_t is the decay-time error, and the resolution model is defined as

$$R(t' - t | \sigma_t) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{(t' - t)^2}{2\sigma_t^2}\right) \qquad (9.12)$$

The presented analysis makes use of the DecayTreeFitter algorithm [106], which determines per-event predictions $\sigma_{t'}$ of the decay-time errors alongside the decay times. Ideally, these predictions would be identical to the true decay-time errors σ_t. However, they are typically underestimating the true decay-time resolution and need to be corrected. This calibration process is outlined in the following. The description is held shortly, as the largest part of the work has been done by collaborators.

In order to find the relation between the estimated decay-time errors $\sigma_{t'}$ and the true decay-time resolution σ_t, candidates with known true decay times are needed. Thus, promptly produced D_s candidates are reconstructed in the final state $KK\pi$ using the lifetime-unbiased `StrippingB02DKLTUBD2HHH` stripping line. Subsequently, they are combined with a random track that originates directly from the PV as well. In this way a sample of fake B_s^0 mesons is obtained with true decay times of $t = 0\,\text{ps}$. However, due to the decay-time resolution the reconstructed decay times are spread around $0\,\text{ps}$, which allows for a calibration of the decay-time errors. Before the calibration is performed, combinatorial background is statistically subtracted. This is done by applying *sWeights* obtained in an *sPlot* fit [119] to the D_s mass distribution (see Fig. 9.1). Afterwards, the sample is split into 20 equally filled bins of the estimated per-event errors $\sigma_{t'}$. Figure 9.2 shows the decay-time distributions of the fake B_s^0 candidates for two of these bins, which cover the ranges $\sigma_{t'} \in [20, 25]\,\text{fs}$ and $\sigma_{t'} \in [50, 55]\,\text{fs}$. The plots of Fig. 9.2 also contain the PDF, which is fit separately in each of the 20 bins. In order to cover different sources of the decay-time resolution, the sum of two Gaussians is used as fit PDF. The fits are performed in the range $[-300\,\text{fs}, \overline{\sigma}_{t'}]$, where $\overline{\sigma}_{t'}$ is the numerical value of the centre of the respective per-event-error bin. A limited fit range is used, because the decay-time distributions show a bias towards positive decay times, as the D_s^- mesons sometimes originate from decays of real long-lived particles. For each bin, the obtained widths σ_1 and σ_2 of the double Gaussians are combined into an effective width

$$\sigma_{t,\text{eff}} = \sqrt{(-2/\Delta m_s^2) \ln D_t}, \qquad (9.13)$$

of the per-event error, using the dilution

$$D_t = f e^{-\sigma_1^2 \Delta m_s^2 / 2} + (1 - f) e^{-\sigma_2^2 \Delta m_s^2 / 2}, \qquad (9.14)$$
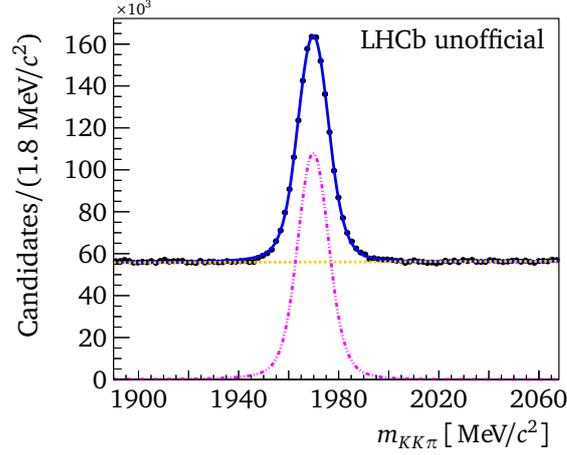
Figure 9.1: The invariant mass distribution of the promptly produced D_s candidates. The overall fit PDF is shown as solid blue line, signal D_s candidates are depicted as pink dashed-dotted line, and combinatorial background is illustrated as yellow dashed line.

where f is the fraction between the two Gaussians. Hence, $\sigma_{t,\text{eff}}$ can be interpreted as the width of a single Gaussian resolution that has the same diluting effect on the time-dependent *CP* asymmetry in a specific bin as the measured one that is based on two Gaussians. The next step in the calibration is to fit the relation between the determined $\sigma_{t,\text{eff}}$ and the estimated $\sigma_{t'}$, where the latter are the centres of the 20 bins. A linear dependence is found to be sufficient, given by

$$\sigma_{t,\text{eff}}(\sigma_{t'}) = s_0 + s_1 \cdot \sigma_{t'} = (10.262 \pm 1.523)\,\text{fs} + (1.280 \pm 0.042)\,\sigma_{t'}. \qquad (9.15)$$
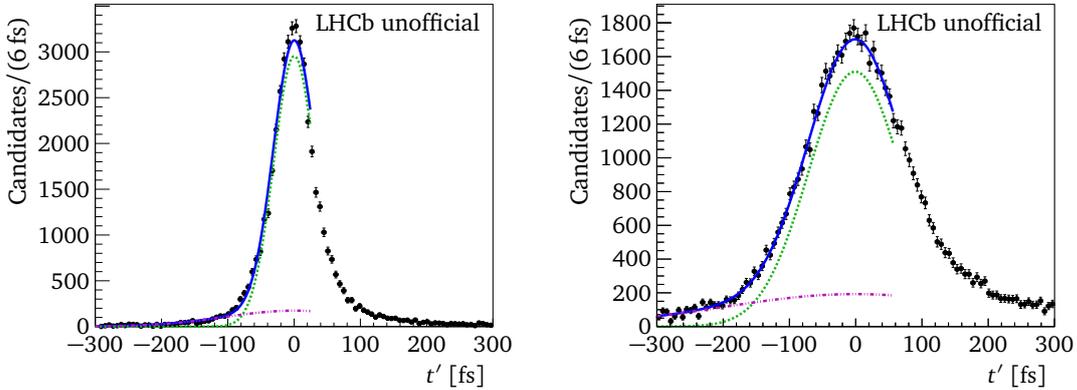

Figure 9.2: Decay-time distribution of fake B_s^0 candidates, which are built by combining promptly produced D_s candidates and a random bachelor track, for two bins in the per-event error, (left) 3$^{\text{rd}}$ bin $\sigma_{t'} \in [20, 25]\,\text{fs}$ and (right) 18$^{\text{th}}$ bin, $\sigma_{t'} \in [50, 55]\,\text{fs}$. The total fit is shown as solid blue line, while the narrow and the broad Gaussians are shown as green dashed and purple dashed-dotted lines, respectively.
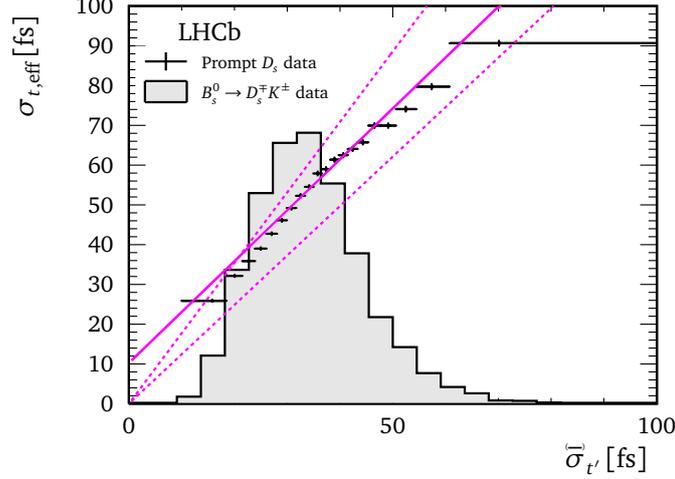
Figure 9.3: Data points show the measured resolution $\sigma_{t,\text{eff}}$ as a function of the bin centres $\overline{\sigma}_{t'}$ of the per-candidate errors for prompt D_s candidates combined with a random track. The solid line shows the linear fit to the presented data set as discussed in the text. Overlaid as a grey bar chart in arbitrary units is the distribution of the per-candidate decay-time errors $\sigma_{t'}$ for $B_s^0 \to D_s^\mp K^\pm$ candidates. The dashed lines indicate alternative approaches discussed later in Sec. 9.7.4.

Figure 9.3 illustrates the fit as well as the underlying distribution of the estimated $\sigma_{t'}$ from $B_s^0 \to D_s^\mp K^\pm$ signal candidates. The obtained calibration indicates that candidates with an estimated decay-time error of about 30 fs, which is in the range of the most frequent decay-time errors (see Fig. 9.3), a true resolution of about 50 fs needs to be applied. This corresponds to a scaling of the estimated decay-time errors of about 1.5. In the previous $B_s^0 \to D_s^\mp K^\pm CP$ analysis, the estimated decay-time errors were corrected by multiplying a constant scale factor of 1.37 [75] to all candidates.

Different studies are performed to verify the calibration method. The main difference between the fake B_s^0 candidates and the signal B_s^0 decays is the p_T spectrum of the bachelor particle. Therefore, the calibration process is repeated in different bins of the bachelor p_T. The study does not observe a strong dependence of the resolution on the bachelor p_T and finds parameters that are compatible with the nominal result given in Eq. 9.15. Another check is performed regarding a possible dependence between the decay-time resolution and the measured decay times. This check is useful, as the presented calibration method is per construction valid for $t = 0$. Thus, a sample of simulated $B_s^0 \to D_s^\mp K^\pm$ decays is used to measure the resolution as a function of the reconstructed B_s^0 decay time. The study finds a scale factor of about 1.2 that is stable for all decay times. Additionally, comparisons between simulated prompt D_s mesons and samples of simulated $B_s^0 \to D_s^\mp K^\pm$ and $B_s^0 \to D_s^- \pi^+$ decays are performed. These comparisons indicate a good agreement, which gives confidence that prompt D_s mesons are a good proxy to quantify the decay-time resolution for B_s^0 decays. The discussion of the systematic uncertainties arising from the decay-time resolution and its calibration method can be found in Sec. 9.7.4.

9.4 Decay-time acceptance

The term decay-time acceptance relates to a decay-time-dependent efficiency that is introduced by the signal selection (see Sec. 7). In case of the presented analysis, this efficiency is a consequence of the trigger lines, the requirements on the vertex separation χ^2 (see Sec. 7.2), and the BDT classifier (see Sec. 7.4). The latter biases the decay time, as it uses decay-time biasing observables like the IPχ^2 of the B_s^0 (see Fig. 7.4). In the decay-time fit the decay-time acceptance is taken into account by multiplying the decay-time PDF (see Eq. 9.11) by an efficiency function $a(t')$ after convolving with the resolution model, *i.e.*

$$P(t', d, q|\sigma_t) = a(t') \left((1 - d A_{\rm P})(1 - q A_{\rm D}) P^{\rm theo}(t, d, q) \otimes R(t' - t|\sigma_t) \right). \quad (9.16)$$

Correlations prevent a measurement of $a(t')$ in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay-time fit alongside the *CP* parameters. Thus, $a(t')$ is evaluated independently and used as a fixed template in the *CP* fit.

The shape of the decay-time acceptance is influenced by many different effects that cannot be analysed separately. Therefore, $a(t')$ is interpolated by a phenomenological model using cubic spline functions [155]. Spline functions are piecewise polynomial functions parameterised by a vector of knots and coefficients. The splines are implemented in an analytical way following the method discussed in Ref. [156]. The number of knots and the choice of their positions determine how precise a given shape can be described. Here, eight coefficients v_i (with $i \in [1, 8]$) and six internal knots at the positions $t = \{0.5, 1.0, 1.5, 2.0, 3.0, 12.0\}$ ps are used. The first and the last internal knot are repeated to receive the end points of the splines. Thus, v_1 and v_2 belong to the first knot position, while v_7 and v_8 are related to the position of the last knot. In order to cope with the rapid change in efficiency introduced by the BDT (see 7.4), the density of knots is higher at lower decay times. Less knots are needed for higher decay times, as the VELO reconstruction of the FastVelo algorithm [157] introduces only a flat drop in efficiency for candidates with higher lifetimes. A normalisation is established by fixing $v_7 = 1$. Additionally, the efficiency is stabilised by fixing v_8 to

$$v_8 = v_7 + \frac{v_6 - v_7}{t_6 - t_7} \times (t_{\rm max} - t_7). \quad (9.17)$$

Here, t_i corresponds to the position of knot i and $t_{\rm max} = 15$ ps is the upper boundary of the decay-time distribution. The spline parametrisation is expanded outside of the knot boundaries by extrapolating the first derivative at this point. Figure 9.4 illustrates the decay-time acceptance and the basic quadratic splines (B-splines) it consists of.

Three sets of spline coefficients are determined to calculate the coefficients v_i used in the $B_s^0 \to D_s^{\mp} K^{\pm}$ *CP* fit. The first set is obtained in a decay-time fit to the $B_s^0 \to D_s^- \pi^+$ signal candidates, while the other two sets originate from fits to simulated $B_s^0 \to D_s^- \pi^+$
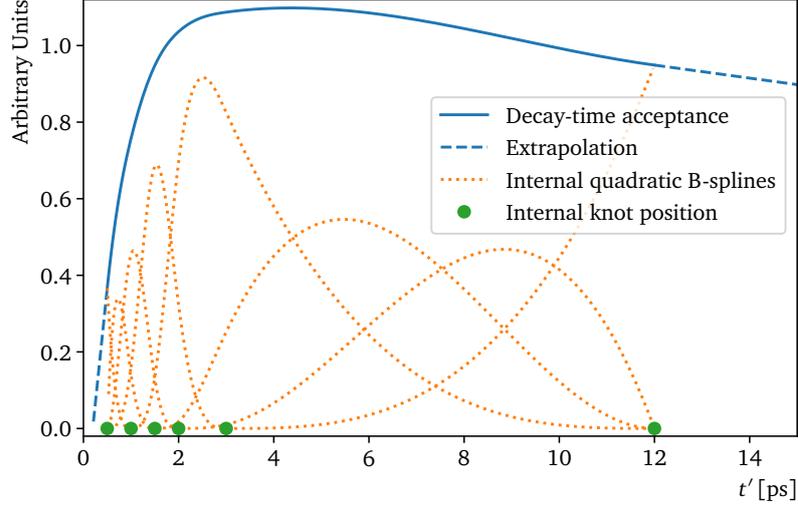
Figure 9.4: Decay-time acceptance and its lower-order basic components (B-splines).

and $B_s^0 \to D_s^\mp K^\pm$ decays. The nominal $B_s^0 \to D_s^\mp K^\pm$ coefficients are obtained by multiplying the $B_s^0 \to D_s^- \pi^+$ data coefficients with the ratio of coefficients from the simulated samples.

In the $B_s^0 \to D_s^- \pi^+$ data fit the *sWeights* calculated in the multidimensional fit (see Sec. 8) are utilised. The signal candidates are fitted with a PDF based on Eq. 9.16, where P^{theo} is only the exponential component of Eq. 9.2. The decay-time resolution is taken into account as introduced in Eqs. 9.11 and 9.12. As mentioned in Sec. 9.1, the decay width and the decay-width difference are fixed to the HFLAV averages [57]. This is particularly important for the determination of $a(t')$ as it is strongly correlated with Γ_s. Thus, the only remaining floating parameters in the fit are the spline coefficients v_i (with $i \in [1,6]$) and the oscillation frequency Δm_s. Figure 9.5 shows the fit of the $B_s^0 \to D_s^- \pi^+$ decay-time distribution, while Tab. 9.1 contains the obtained parameters. The central value of Δm_s is blinded as it is likely that Δm_s will be measured in a dedicated analysis on the same data set in the future. The most recent published LHCb measurement of Δm_s, which was also performed in the $B_s^0 \to D_s^- \pi^+$ channel, uses only the 2011 part of the Run 1 data [158].

The $B_s^0 \to D_s^- \pi^+$ and the $B_s^0 \to D_s^\mp K^\pm$ simulated samples have to fulfil the signal selection and are weighted for data/simulation differences as well as the effect of the PID requirement. The spline coefficients are determined in similar fits as that applied to the $B_s^0 \to D_s^- \pi^+$ data described above. As the $B_s^0 \to D_s^\mp K^\pm$ sample was simulated without *CP* violation, the same PDF can be used as in $B_s^0 \to D_s^- \pi^+$ decays. All parameters except for those of the acceptance are fixed to their generated values. A potential dependence of the acceptance on the D_s final state is tested and differences in the range of 5 to 10 % are found, depending on the coefficient and final state. In order to compensate these differences, events of the different final states are weighted according to their yields observed in data. The calculated ratio of the $B_s^0 \to D_s^\mp K^\pm$ and
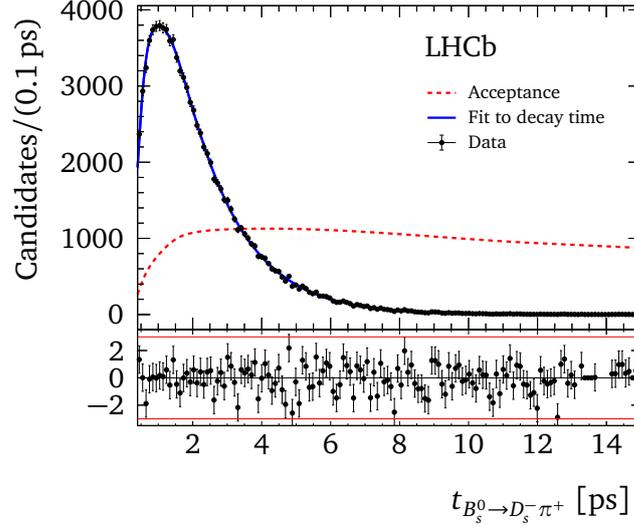
Figure 9.5: Decay-time distribution of $B_s^0 \to D_s^- \pi^+$ candidates obtained by the *sPlot* technique. The solid blue curve is the result of the *sFit* procedure and the dashed red curve shows the measured decay-time acceptance in arbitrary units. Normalised residuals are shown underneath.

$B_s^0 \to D_s^- \pi^+$ acceptances is compatible within statistical uncertainties throughout all final states and parameters. Figure 9.6 shows the two decay-time fits to the simulated samples, while Tab. 9.2 contains the two sets of coefficients as well as their ratio. The final $B_s^0 \to D_s^\mp K^\pm$ acceptance coefficients, which are calculated as the product of the $B_s^0 \to D_s^- \pi^+$ data coefficients and the last column of Tab. 9.2, are listed in Tab. 9.3.

Besides the strategy depicted in the preceding paragraphs, is is in principle possible to model the decay-time acceptance solely data-driven. This strategy, known as swimming, was analysed and prepared by the author for the presented analysis, but it did not become the nominal strategy. Section 10.4 discusses details about the swimming approach, which might be considered in future analyses.

Table 9.1: Result of the *sFit* performed to the $B_s^0 \to D_s^- \pi^+$ data sample. The central value of Δm_s is blinded.

| Parameter	Fit result
Δm_s	XXX $\pm$ 0.0129 ps^{-1}
v_1	0.390 $\pm$ 0.014
v_2	0.596 $\pm$ 0.024
v_3	0.790 $\pm$ 0.031
v_4	1.014 $\pm$ 0.038
v_5	1.099 $\pm$ 0.036
v_6	1.189 $\pm$ 0.062
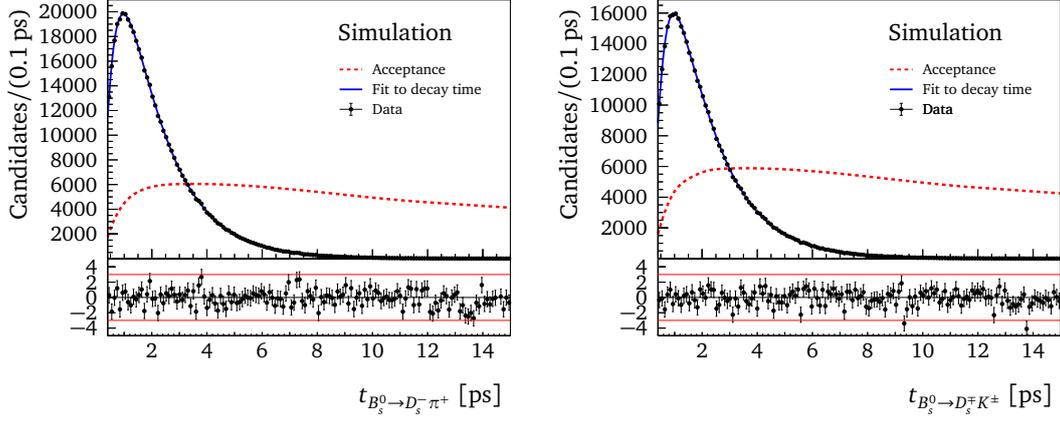
Figure 9.6: Decay-time acceptances determined on $B_s^0 \to D_s^- \pi^+$ (left) and $B_s^0 \to D_s^{\mp} K^{\pm}$ (right) simulation samples. The acceptance function is shown in red, while the total fit curve is blue.

Table 9.2: The fit parameters for the acceptance fit to the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ simulation samples, and their ratio.

Parameters	Fit to $B_s^0 \to D_s^{\mp} K^{\pm}$	Fit to $B_s^0 \to D_s^- \pi^+$	$B_s^0 \to D_s^{\mp} K^{\pm}/B_s^0 \to D_s^- \pi^+$
v_1	0.447 ± 0.005	0.475 ± 0.005	0.939 ± 0.015
v_2	0.646 ± 0.009	0.679 ± 0.008	0.952 ± 0.017
v_3	0.923 ± 0.012	0.935 ± 0.011	0.987 ± 0.017
v_4	1.043 ± 0.013	1.095 ± 0.013	0.952 ± 0.016
v_5	1.162 ± 0.013	1.195 ± 0.012	0.973 ± 0.015
v_6	1.225 ± 0.022	1.263 ± 0.020	0.970 ± 0.023

Table 9.3: The parameters of the $B_s^0 \to D_s^{\mp} K^{\pm}$ acceptance.

Parameters	Fit Values	correlation among parameters					
		v_1	v_2	v_3	v_4	v_5	v_6
v_1	0.366 ± 0.015	1	0.84	0.74	0.88	0.88	0.83
v_2	0.567 ± 0.025		1	0.54	0.85	0.79	0.78
v_3	0.779 ± 0.033			1	0.65	0.83	0.70
v_4	0.966 ± 0.040				1	0.82	0.88
v_5	1.070 ± 0.039					1	0.74
v_6	1.153 ± 0.066						1

9.5 Flavour tagging

The analysis uses two tagging responses, the same side kaon tagger and the standard opposite side tagger combination, where the latter includes the OS muon, the OS electron, the OS kaon, and the OS vertex charge tagger. This selection is equivalent to that of the preceding analysis of *CP* violation with $B_s^0 \to D_s^{\mp} K^{\pm}$ decays [75]. However, OS and SS were combined at the level of the data set in the previous analysis, while in the presented analysis the combination is performed directly inside the PDF. This approach has the advantage of allowing an easy propagation of the OS and SS calibration parameters' uncertainties to the *CP* parameters. The combination inside the PDF is performed analogously to the way introduced in Sec. 6.4. As the flavour tagging is only an imperfect determination of the true initial B_s^0 flavour d, Eq. 9.16 is changed to a conditional PDF of the tagging decision $\vec{d}' = (d'_{\mathrm{OS}}, d'_{\mathrm{SS}})$ given the corresponding wrong tag estimates $\vec{\eta} = (\eta_{\mathrm{OS}}, \eta_{\mathrm{SS}})$. Hence, the measured signal decay time becomes

$$P(t', \vec{d}', q | \sigma_t, \vec{\eta}) = a(t') \left(P(t, \vec{d}', q | \vec{\eta}) \otimes R(t' - t | \sigma_t) \right), \qquad (9.18)$$

with

$$P(t, \vec{d}', q | \vec{\eta}) = \sum_d P(\vec{d}' | d, \vec{\eta}) \left((1 - d A_{\mathrm{P}})(1 - q A_{\mathrm{D}}) P^{\mathrm{theo}}(t, d, q) \right). \qquad (9.19)$$

The two-dimensional binomial PDF $P(\vec{d}' | d, \vec{\eta})$ describes the distribution of tagging decisions for given d and $\vec{\eta}$.

The calibration of the flavour tagging is performed on the flavour-specific decay $B_s^0 \to D_s^- \pi^+$. A good portability of the flavour-tagging calibration from $B_s^0 \to D_s^- \pi^+$ to the signal decay $B_s^0 \to D_s^{\mp} K^{\pm}$ is expected as the decays are kinematically very similar. Using the same selection for both decay modes (see Chap. 7) avoids that differences in terms of trigger and reconstruction lead to a bad portability. Additionally, distributions of observables that are correlated with the tagging, such as the azimuthal angle, the detector occupancy, the pseudorapidity and the transverse B momentum, as well as the distribution of the wrong tag estimates are compared between simulated truth-matched $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ events. All distributions are compatible for the two decay modes. Additionally, calibrations are performed on the simulated $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B_s^0 \to D_s^- \pi^+$ events (see Tab. 9.4). The tagging parameters measured in the simulations are compatible between the two decays within less than 1 and 2 σ, respectively.

The calibration parameters utilised in the $B_s^0 \to D_s^{\mp} K^{\pm}$ *CP* fit are determined with $B_s^0 \to D_s^- \pi^+$ decays from data using a fit of the decay time, similar to the calibrations performed with $B^0 \to J/\psi K^{*0}$ decays, which are described in Sec. 6.6. However, instead of binning the data in the mistag estimates η, the calibration is performed in an unbinned way, using the OS and SS mistag estimates as per-event observables. The *sWeights* from the multidimensional fit (see Chap. 8) are applied. The decay-time PDF of the signal is given by Eq. 9.18, where the theoretical $B_s^0 \to D_s^- \pi^+$ PDF from

Table 9.4: Calibration parameters of the OS tagger combination and the SS kaon tagger extracted from simulated $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ signal events.

	Tagger	p_0	p_1	$\langle \eta \rangle$
$B_s^0 \to D_s^- \pi^+$	OS comb.	0.3607 ± 0.0013	0.936 ± 0.012	0.3597
$B_s^0 \to D_s^\mp K^\pm$	OS comb.	0.3604 ± 0.0013	0.914 ± 0.011	0.3597
$B_s^0 \to D_s^- \pi^+$	SS kaon	0.4267 ± 0.0011	1.186 ± 0.011	0.4314
$B_s^0 \to D_s^\mp K^\pm$	SS kaon	0.4276 ± 0.0011	1.184 ± 0.010	0.4311

Eq. 9.2 is plugged in as P^{theo}. The acceptance parameters are fixed to the values from Tab. 9.6, while the B_s^0 oscillation frequency Δm_s is fixed to its world average [57]

$$\Delta m_s = (17.757 \pm 0.021)\,\text{ps}^{-1}. \qquad (9.20)$$

Table 9.5 contains the results of the OS and SS calibrations including only the statistical uncertainties of the calibration parameters. The estimation of the systematic uncertainties is discussed in the next paragraph. Correlations between the tagging parameters for the OS and the SS calibrations can be found in Tabs. 9.6 and 9.7, respectively. In Tab. 9.8, the performance of the flavour tagging is listed for the $B_s^0 \to D_s^- \pi^+$ and the $B_s^0 \to D_s^\mp K^\pm$ candidates. The numbers are given separately for the candidates that are tagged by either the OS or the SS taggers as well as for the sample of candidates with tags from both SS and the OS algorithms. Compatible performances are observed in the two decay modes. A higher effective tagging efficiency is observed compared to the charmonium decay $B^0 \to J/\psi K^{*0}$ (compare Sec. 6.6). This is reasoned by the very good performance of the SS kaon tagger compared to the SS pion tagger for B^0 mesons as well as the kinematic properties of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ decays. For example, the triggering of the latter decays leads to an on average higher p_T of the B meson compared to $B^0 \to J/\psi K^{*0}$ decays, which are mostly triggered by the muons from the J/ψ decay.

Table 9.5: Calibration parameters, tagging asymmetries and performance of the OS and the SS tagger measured on $B_s^0 \to D_s^- \pi^+$ decays including their statistical uncertainties.

Parameter	OS	SS
p_0	0.374 ± 0.006	0.441 ± 0.005
p_1	1.09 ± 0.06	1.08 ± 0.07
Δp_0	0.014 ± 0.006	-0.018 ± 0.004
Δp_1	0.13 ± 0.06	0.13 ± 0.07
$\langle \eta \rangle$	0.370	0.437
$\varepsilon_{\text{tag}}\,[\%]$	37.15 ± 0.17	63.90 ± 0.17
$\Delta \varepsilon_{\text{tag}}\,[\%]$	-1.1 ± 0.7	0.8 ± 0.7
$\varepsilon_{\text{eff}}\,[\%]$	3.89 ± 0.29	2.08 ± 0.21

Table 9.6: Correlations among the flavour tagging parameters and asymmetries for the OS tagger from the $B_s^0 \to D_s^- \pi^+$ *sFit* on data.

Parameter	$\Delta\varepsilon_{\text{tag}}$	Δp_0	Δp_1	p_0	p_1	ε_{tag}
$\Delta\epsilon_{\text{tag}}$	1	−0.322	0.269	−0.005	−0.009	−0.019
Δp_0		1	−0.061	−0.017	0.008	0.004
Δp_1			1	0.003	−0.017	−0.003
p_0				1	0.018	0.00
p_1					1	0.00
ε_{tag}						1

Table 9.7: Correlations among the flavour tagging parameters and asymmetries for the SS tagger from the $B_s^0 \to D_s^- \pi^+$ *sFit* on data.

Parameter	$\Delta\varepsilon_{\text{tag}}$	Δp_0	Δp_1	p_0	p_1	ε_{tag}
$\Delta\epsilon_{\text{tag}}$	1	−0.119	0.138	−0.003	−0.005	−0.016
Δp_0		1	−0.008	0.026	0.017	0.002
Δp_1			1	0.013	0.003	−0.003
p_0				1	0.054	0.00
p_1					1	0.00
ε_{tag}						1

The systematic uncertainties of the tagging parameters are obtained by collaborators, which is why only the results are summarised in the following. The strategy of obtaining the uncertainties remains mostly the same compared to the previous analysis of *CP* violation in $B_s^0 \to D_s^\mp K^\pm$ decays [75]. It is documented in detail for the preceding analysis in Ref. [159]. The largest source of systematic uncertainty is the decay-time resolution. As it also directly affects the *CP* parameters measured in the $B_s^0 \to D_s^\mp K^\pm$ fit, it is treated separately from the other systematic uncertainties to avoid double counting (see Sec. 9.7). All other significant sources of systematic uncertainty on the flavour-tagging parameters are listed in Tab. 9.9. The uncertainty on the calibration method is obtained by comparing the nominal result, which uses a per-event mistag probability, to a fit that is binned in η. The other two relevant systematic uncertainties are both related to the description of background components in the multidimensional fit. As any residual background candidate can potentially alter the *sweighted* flavour tagging calibration, the MD fit is repeated with an alternative modelling of the background. In particular, fits are performed with a modified fixed fraction of $B_s^0 \to D_s^\mp K^\pm$ candidates, testing a halved and a doubled value. The flavour-tagging calibration is repeated using the *sWeights* from the modified MD fits and the difference in the calibration parameters estimates the systematic uncertainty. Analogously, the MD fit is repeated implementing two alternative models for the combinatorial background component. Instead of the sum of an exponential and a constant,

Table 9.8: The flavour-tagging performances for $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ signal candidates on the selected Run 1 data. Apart from values for the full selected samples, the performances are given for the candidates that are only tagged by the OS or the SS, as well as for the candidates with tags from both OS and SS taggers.

| Candidates tagged by | $B_s^0 \to D_s^- \pi^+$ | | $B_s^0 \to D_s^{\mp} K^{\pm}$ | |
	ε_{tag} [%]	ε_{eff} [%]	ε_{tag} [%]	ε_{eff} [%]
OS only	12.94 ± 0.11	1.41 ± 0.11	13.58 ± 0.44	1.44 ± 0.12
SS only	39.70 ± 0.16	1.29 ± 0.13	38.65 ± 0.63	1.18 ± 0.12
Both OS-SS	24.21 ± 0.14	3.10 ± 0.18	23.37 ± 0.55	3.05 ± 0.20
Total	76.85 ± 0.24	5.80 ± 0.25	75.60 ± 1.30	5.67 ± 0.26

a single exponential and a double exponential are tested. All three significant sources of systematic uncertainty are added in quadrature with the statistical uncertainties. In this way total uncertainties are obtained, which are used in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay-time fit to define Gaussian constraints on the tagging calibration parameters. Further possible sources of systematic uncertainty are considered, but found to have negligible impact. These include a comparison between calibrations of the 2011 and 2012 subsamples of the Run 1 data, a variation of the decay-time acceptance parameters, and letting Δm_s and $\Delta\Gamma_s$ float in the fit. Overall, the statistical uncertainties strongly dominate the systematic uncertainties (see last line of Tab. 9.9.), which leaves room for future improvements when performing the calibration on larger datasets.

Table 9.9: Systematic uncertainties of the calibration parameters for the OS and SS taggers evaluated using $B_s^0 \to D_s^- \pi^+$ data. The total systematic uncertainty corresponds to the sum in quadrature of all quoted uncertainties. As a comparison the statistical uncertainties are reported in the last line of the table.

| Systematic uncertainties | OS [10^{-3}] | | | | SS [10^{-3}] | | | |
	σ_{p_0}	σ_{p_1}	$\sigma_{\Delta p_0}$	$\sigma_{\Delta p_1}$	σ_{p_0}	σ_{p_1}	$\sigma_{\Delta p_0}$	$\sigma_{\Delta p_1}$
Calibration method	0.2	12	–	–	0.2	3.5	–	–
Fraction of $B_s^0 \to D_s^{\mp} K^{\pm}$	0.1	0.7	0.06	0.9	0.1	1.5	0.09	1.24
Combinatorial PDF($m(B_s)$)	0.4	1.8	0.01	1.5	0.11	4.1	0.15	0.89
Sum in quadrature	0.4	12	0.06	1.7	0.2	6	0.17	1.5
Stat. uncertainties	6.1	63	6	62	4.7	68	4	67

9.6 $B_s^0 \to D_s^{\mp} K^{\pm}$ decay-time fit

The PDF of the unbinned log-likelihood fit to the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay time is defined analogously to the $B_s^0 \to D_s^- \pi^+$ decay-time PDF, which is given in Eqs. 9.19 and 9.18. In order to describe *CP* violation in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay, P^{theo} is taken from Eq. 9.1. The full list of parameters of the PDF is given by

$$C, A_f^{\Delta\Gamma}, A_{\bar{f}}^{\Delta\Gamma}, S_f, S_{\bar{f}},$$

$$\Gamma_s, \Delta\Gamma_s, \Delta m_s, A_{\text{P}}, A_{\text{D}}, s_0, s_1, v_i,$$

$$\vec{p}_0, \vec{p}_1, \Delta\vec{p}_0, \Delta\vec{p}_1, \langle\vec{\eta}\rangle, \vec{\epsilon}_{\text{tag}}, \text{ and } \Delta\vec{\epsilon}_{\text{tag}}.$$

The five *CP* parameters are the only free parameters in the fit, while the decay width Γ_s and the decay-width difference $\Delta\Gamma_s$ are fixed to the values listed in Eq. 9.3. Furthermore, the oscillation frequency Δm_s is fixed to 17.757 ps^{-1} (from Ref. [57]). The asymmetries in production A_{P} and detection A_{D} are fixed to the values given in Eqs. 9.6 and 9.8, respectively. For the calibration parameters s_0 and s_1 of the decay-time resolution the values quoted in Eq. 9.15 are used, while the acceptance parameters v_i are chosen as explained in Sec. 9.4 and listed in Tab. 9.1. Finally, the flavour-tagging calibration and asymmetry parameters are set to their central values from Tab. 9.5, but are not fixed. Instead, they are allowed to float within Gaussian constraints defined as the sum in quadrature of the statistical and systematic uncertainties, where the latter are the values from Tab. 9.9. Hence, no systematic uncertainty from the flavour-tagging calibration parameters needs to be evaluated, as the uncertainties are already contained in the statistical uncertainty of the *CP* parameters reported by the fit. The observables in the fit are the reconstructed decay time t', the corresponding estimated decay-time uncertainty $\sigma_{t'}$, the charge of the bachelor particle q, as well as the tagging decisions $\vec{d}$ and the corresponding wrong tag estimates $\vec{\eta}$ of the SS and OS tagging responses. In the fit the background is statistically subtracted by applying the *sWeights* from the multidimensional fit (see Sec. 8).

The fit is performed to the full Run 1 data set, *i.e.* combining all $D_s^{\mp}$ final states. In Fig. 9.7 a plot of the decay-time distribution including a projection of the fitted PDF is shown. Table 9.10 contains the results for the *CP* parameters obtained by the fit including their estimated statistical uncertainties. For comparison, the table also includes the results of the previous $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis [75]. With respect to the increased number of candidates (see Sec. 8.3), the observed improvement in precision of about 30–40% lies within the expected level. As the sensitivity depends on external nuisance parameters, in particular $\Delta\Gamma_s$ for $A_f^{\Delta\Gamma}$ and $A_{\bar{f}}^{\Delta\Gamma}$, which have changed since the previous measurement, it is not expected that the uncertainties of the *CP* parameters scale in the same way. Taking the statistical uncertainties into account, all parameters are compatible within 1σ between the current and the previous measurement. The correlations between the *CP* parameters are listed in Tab. 9.11. As can be seen, the only large correlation (about 50%) exists between $A_f^{\Delta\Gamma}$ and $A_{\bar{f}}^{\Delta\Gamma}$. This is expected as these parameters are both sensitive to the slowly varying hyperbolic terms induced by
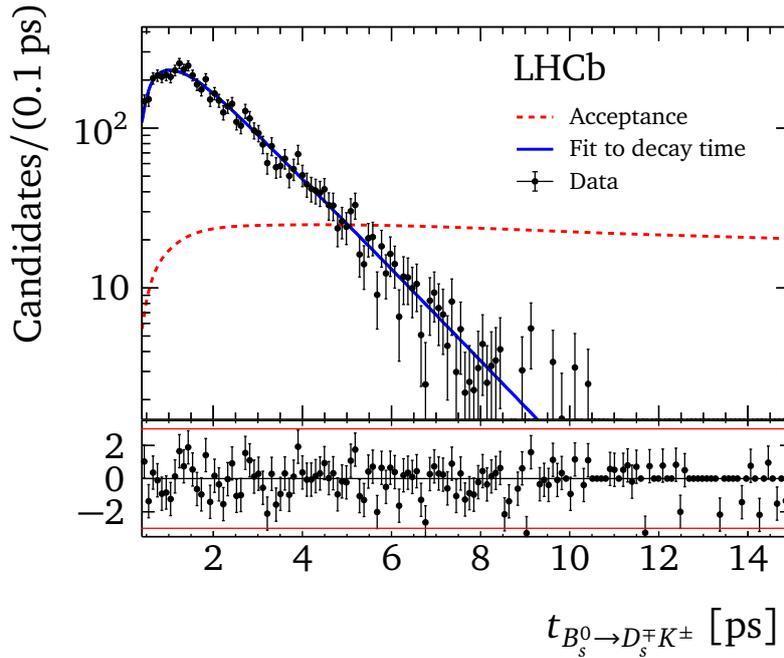
Figure 9.7: Decay-time distribution of the *sWeighted* $B_s^0 \to D_s^\mp K^\pm$ candidates with logarithmic scale on the vertical axis. The dashed red line corresponds to the decay-time acceptance and is plotted with an arbitrary normalisation.

$\Delta\Gamma_s$ (compare the decay-rate equations in Eqs. 3.43 to 3.46). The projections of the decay-time-dependent *CP* asymmetries for the two final states are shown in Fig. 9.8. The next section discusses the validation of the decay-time fit and the estimation of the systematic uncertainties on the *CP* parameters.

Table 9.10: Measured *CP* parameters from the *sWeighted* decay-time fit performed to the $B_s^0 \to D_s^\mp K^\pm$ data sample. Additionally, the results of the previous 1 fb^{-1} analysis [75] are given.

| Parameter	Fitted value	Result 1 fb^{-1} analysis
C	0.73 ± 0.14	0.52 ± 0.25
$A_f^{\Delta\Gamma}$	0.39 ± 0.28	0.29 ± 0.42
$A_{\bar{f}}^{\Delta\Gamma}$	0.31 ± 0.28	0.14 ± 0.41
S_f	-0.52 ± 0.20	-0.90 ± 0.31
$S_{\bar{f}}$	-0.49 ± 0.20	-0.36 ± 0.34
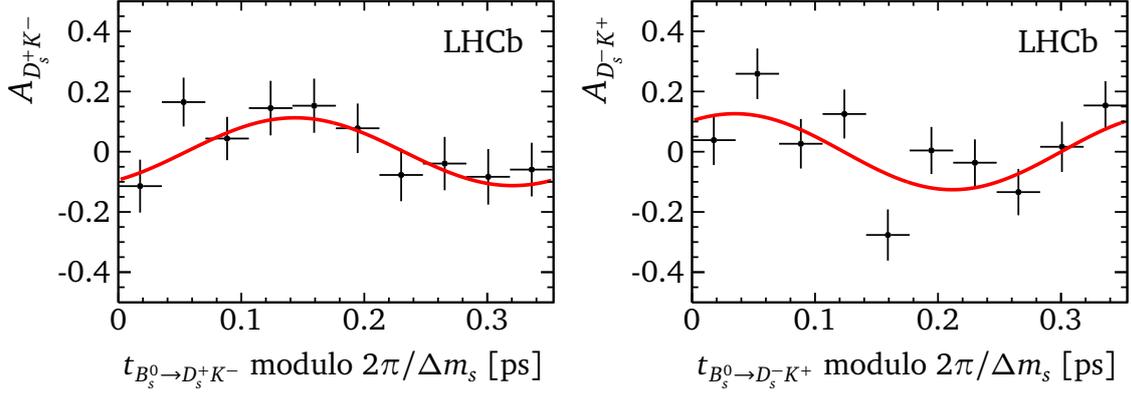
Figure 9.8: The *CP* asymmetry plots for (left) the $D_s^+K^-$ final state and (right) the $D_s^-K^+$ final state, folded into one mixing period $2\pi/\Delta m_s$.

Table 9.11: The correlation matrix of the $B_s^0 \to D_s^\mp K^\pm$ decay-time fit.

Parameter	C	$A_f^{\Delta\Gamma}$	$A_{\bar{f}}^{\Delta\Gamma}$	S_f	$S_{\bar{f}}$
C	1.000	0.092	0.078	0.008	-0.057
$A_f^{\Delta\Gamma}$	0.092	1.000	0.513	-0.083	-0.004
$A_{\bar{f}}^{\Delta\Gamma}$	0.078	0.513	1.000	-0.042	-0.003
S_f	0.008	-0.083	-0.042	1.000	0.001
$S_{\bar{f}}$	-0.057	-0.004	-0.003	0.001	1.000

9.7 Systematic uncertainties

Different potential sources of systematic effects are probed to estimate the systematic uncertainties on the *CP* parameters. The most elementary systematic error would be a bias in the fitting procedure itself. For the presented analysis, the fit procedure is validated to be unbiased by fitting large sets of pseudo-experiments as well as disjoint parts of the data. These studies are summarised in Sec. 9.7.1. In order to further prove the validity of the fitting procedure, a fully simulated MC sample, which includes the modelling of *CP* violation, is generated and fitted. These fits, which are called closure tests, are outlined in Sec. 9.7.2. Some parameters of the PDF are fixed in the fit. This procedure ignores that the knowledge about their true values is limited and represents another possible source of systematic uncertainty (see Sec 9.7.3). Furthermore, different decay-time resolution models are tested as explained in Sec. 9.7.4. Using a factorised PDF in the MD fit ignores possible correlations among the observables, causing a potential source of systematic uncertainty (see Sec. 9.7.5). The systematic effects arising from the modelling of the decay-time acceptance are estimated in Sec. 9.7.6. Finally, Sec. 9.7.7 concludes the total system-

atic uncertainties of the *CP* parameters after mentioning further potential sources of systematic uncertainties that were investigated, but found to be negligible. The work on the systematic effects is done in collaboration with the other members of the analysis group. Contributions from myself are described in Secs. 9.7.1, 9.7.2, and 9.7.6.

9.7.1 Validation of the fit procedure

In order to validate the complete fit procedure, 1000 toy data samples are generated (see Sec. 4.5.2 for an introduction to pseudo-experiments) taking into account all dimensions of the MD fit and the decay-time fit. In each toy dataset, the number of generated signal and background candidates is chosen corresponding to the yields that are observed in data (see Tab. 8.4). All shape parameters are set according to their values in the data fits. The toy data undergoes the same fit procedure as the recorded data, *i.e.* both the MD fit and the decay-time fit are performed with the corresponding PDFs. Pull distributions are obtained for the *CP* parameters and are fitted with Gaussian PDFs. Table 9.12 contains the measured means μ and widths σ of the Gaussians. All means are compatible with 0 within less than three σ of their statistical uncertainty, which confirms an unbiased fit procedure. Furthermore, the estimated fit uncertainties are reasonable, as the widths of the pull distributions show compatibility with 1 within less than three σ of their statistical uncertainty. Hence, no systematic uncertainty on the *CP* parameters from the fit procedure is applied.

An additional cross-check of the fitting procedure is performed by fitting disjoint subsamples of the full dataset. These decay-time fits are performed to the same subsamples used in the context of the MD fit validation (see Sec. 8.4). This cross-check implies redoing the determination of the acceptance parameters as well as the flavour-tagging calibration. Table 9.13 compares the fit results of the different subsamples with each other and the nominal fit result from Tab. 9.10. The results are in agreement within about 1.5 σ regarding their statistical uncertainty. For $A_f^{\Delta\Gamma}$ and $A_{\bar{f}}^{\Delta\Gamma}$ their correlation and the systematic uncertainty from the acceptance need to be considered. Among all of the decay-time fits on the subsamples, that on the 2011 subsample did not converge, which is the reason for the missing comparison between the years of

Table 9.12: Means and widths of the pull distributions of the nominal $B_s^0 \to D_s^{\mp} K^{\pm}$ pseudo-experiments. Additionally, the fit uncertainties observed in the studies are given.

Parameter	μ	σ	fit uncertainty
C	-0.004 ± 0.033	0.937 ± 0.028	0.139 ± 0.002
$A_f^{\Delta\Gamma}$	0.101 ± 0.036	1.022 ± 0.027	0.259 ± 0.004
$A_{\bar{f}}^{\Delta\Gamma}$	0.074 ± 0.038	1.080 ± 0.030	0.261 ± 0.004
S_f	0.062 ± 0.034	0.957 ± 0.026	0.195 ± 0.004
$S_{\bar{f}}$	0.007 ± 0.035	0.986 ± 0.028	0.196 ± 0.004

data taking. Several tests are performed in order to solve the issue, including changes in the dimension ranges, using different initial values for the *CP* parameters, searching for pathological events and including only OS or SS tagged candidates. None of the attempts results in a converging fit. Finally, toy datasets are generated according to the number of candidates that is observed in the part of the data that was taken in 2011. When fitting these toy datasets a failing rate of more than 30 % is observed. This leads to the assumption that the failing fit of the 2011 part of the data is an unfortunate consequence of the small number of candidates.

Table 9.13: Comparisons of the fits in the subsamples of the $B_s^0 \to D_s^{\mp} K^{\pm}$ data. The data is split by magnet polarity, as well as low and high values of the B_s^0 momentum ($p_{B_s^0}$) and the BDT classifier (see Sec. 8.4). The top set of numbers gives the differences between the corresponding splits of the data, while the bottom set contains the differences between the average of the subsamples and the nominal fit.

	Magnet polarity	$p_{B_s^0}$ split	BDT split
C	0.21 ± 0.29	-0.39 ± 0.30	0.32 ± 0.29
$A_f^{\Delta\Gamma}$	-0.92 ± 0.55	0.64 ± 0.66	0.49 ± 0.63
$A_{\bar{f}}^{\Delta\Gamma}$	-0.60 ± 0.55	-0.26 ± 0.63	1.19 ± 0.63
S_f	0.00 ± 0.41	-0.61 ± 0.50	0.25 ± 0.42
$S_{\bar{f}}$	-0.54 ± 0.39	-0.08 ± 0.39	0.21 ± 0.39
C	-0.01 ± 0.02	0.03 ± 0.01	0.03 ± 0.01
$A_f^{\Delta\Gamma}$	-0.03 ± 0.04	-0.11 ± 0.05	0.18 ± 0.14
$A_{\bar{f}}^{\Delta\Gamma}$	-0.03 ± 0.04	-0.11 ± 0.05	0.14 ± 0.15
S_f	-0.01 ± 0.02	-0.01 ± 0.06	0.01 ± 0.03
$S_{\bar{f}}$	-0.03 ± 0.04	-0.05 ± 0.03	0.03 ± 0.03

9.7.2 Closure tests with fully simulated events

The fit procedure is further validated by fitting a special fully simulated signal MC sample. This sample incorporates *CP* violation in its generation by using the so-called SSD_CP generator from EVTGEN [111]. The generated MC takes all three D_s final states into account (see Tab. 7.7) and is exclusively produced for the three closure tests that are described in the following. The first test is a decay-time fit to the MC sample, *i.e.* it represents a fit of the signal events only. In the second test, combinatorial background from data is added to the MC sample. The combinatorial background candidates are taken from the upper mass sideband $m(D_s^{\mp} h^{\pm}) \in [5600, 6100]$ MeV/c^2 shifted by 300 MeV/c^2 to match the signal fit region, and from the interval $m(D_s^{\mp} h^{\pm}) \in [6100, 6600]$ MeV/c^2 shifted by 800 MeV/c^2. The addition of background candidates allows to perform an MD fit, which subsequently enables an *sWeighted* decay-time

fit equivalent to the strategy applied to data. The last closure test is performed similarly to the second, but prior to the MD fit simulated $B_s^0 \to D_s^- \pi^+$ signal events are added to the sample maintaining the same proportions found in data. In all tests, the same selection criteria are applied to the simulated data as described in Chap. 7. The decay-time fits are performed analogously to that on data, which implies leaving the same parameters free in the fit, using per-event decay-time errors and utilising tagging responses from the OS and SS. The values for the decay-time acceptance are obtained from simulated $B_s^0 \to D_s^- \pi^+$ samples.

Table 9.14 lists the input values that are chosen to be used in the generation by the SSD_CP generator. Due to a bug in the documentation of the generator, the value for $\Delta\Gamma_s/\Gamma_s$ is sign flipped compared to the HFLAV definition. Following these inputs, Eqs. 3.72 to 3.74 allow to calculate the corresponding expected values for the CP parameters. These values are listed in Tab. 9.15 together with the results of the first closure test, which consists of just a decay-time fit to the signal events and hence, is expected to give the most accurate fit result. As can be seen in the last column of the table, which contains the differences between the generated values and the obtained fit results, all CP parameters are reproduced within less than $1\,\sigma$ corresponding to their statistical uncertainties. The relative sensitivities of the different CP parameters do not match what is observed on data, which is a consequence of differences in the flavour-tagging performance as well as in the nuisance parameters. The results of the second and third closure tests are compared to that of the first one in Tab. 9.16 by calculating the differences of the fit results. Overall, the results agree within less than $2\,\sigma$. The statistical uncertainties observed in the first closure test, which are reported in Tab 9.15, are added as a systematic uncertainty and listed in Tab. 9.17 as 'closure test'. This uncertainty covers the intrinsic uncertainty related to the fit procedure.

Table 9.14: Input parameters used in the generation of the $B_s^0 \to D_s^\mp K^\pm$ signal MC that incorporates CP violation.

Parameter	Input value	Reference
Δm_s	$17.757\,\mathrm{ps}^{-1}$	HFLAV [57]
$\Delta\Gamma_s/\Gamma_s$	-0.122	$(-1)\cdot$ HFLAV [57]
$|q/p|$	1	assuming no CP violation in mixing
$\arg(q/p) = -2\beta_s$	$-0.0363\,\mathrm{rad}$	CKM fitter group [56]
$|A_f|$	1	assuming $A_f = e^{i\gamma}$
$\arg(A_f) = \gamma$	$1.2776\,\mathrm{rad}$	CKM fitter group [56]
$|\overline{A}_f| = |\lambda_f|$	0.37	see Eq. 3.72
$\arg(\overline{A}_f) = \delta$	$0.1745\,\mathrm{rad}$	assuming a strong phase $\delta = 10°$
$|A_{\bar{f}}| = |\overline{\lambda}_{\bar{f}}|$	0.37	see Eq. 3.72
$\arg(A_{\bar{f}}) = \delta$	$0.1745\,\mathrm{rad}$	assuming a strong phase $\delta = 10°$
$|\overline{A}_{\bar{f}}|$	1	assuming $\overline{A}_{\bar{f}} = e^{-i\gamma}$
$\arg(\overline{A}_{\bar{f}}) = -\gamma$	$-1.2776\,\mathrm{rad}$	CKM fitter group [56]

Table 9.15: Comparison of generated and fitted $B_s^0 \to D_s^\mp K^\pm$ CP parameters for the first closure test that uses only the generated signal events.

Parameter	Generated value	Fit result	Difference
C	0.75917	0.769 ± 0.018	-0.010 ± 0.018
$A_f^{\Delta\Gamma}$	0.31436	0.301 ± 0.054	0.013 ± 0.054
$A_{\bar{f}}^{\Delta\Gamma}$	0.10046	0.059 ± 0.052	0.041 ± 0.052
S_f	-0.56995	-0.579 ± 0.025	-0.009 ± 0.025
$S_{\bar{f}}$	-0.64309	-0.653 ± 0.024	-0.010 ± 0.024

Table 9.16: Differences between the CP parameters obtained in the first closure test (signal-only, see Tab. 9.15) and the second (signal + combinatorial) and third (signal + combinatorial + $B_s^0 \to D_s^- \pi^+$) closure test fits.

Parameter	Signal-only vs. signal + combinatorial	Signal-only vs. signal + combinatorial + $B_s^0 \to D_s^- \pi^+$
C	-0.044 ± 0.027	-0.039 ± 0.031
$A_f^{\Delta\Gamma}$	0.009 ± 0.048	0.098 ± 0.070
$A_{\bar{f}}^{\Delta\Gamma}$	-0.013 ± 0.059	-0.086 ± 0.088
S_f	-0.027 ± 0.030	0.016 ± 0.044
$S_{\bar{f}}$	0.013 ± 0.040	-0.021 ± 0.046

9.7.3 Influence of fixed parameters

Some parameters are fixed in the fit, neglecting their uncertainties. This section discusses, how the systematic uncertainties arising from fixing parameters that are not correlated to other fixed parameters are determined. This includes Δm_s, the detection asymmetry and the tagging asymmetry parameters. Section 9.7.6 explains, how the correlated fixed parameters (parameters of the decay-time acceptance, Γ_s, and $\Delta\Gamma_s$) are treated. For the uncorrelated parameters the resulting systematic effects on the CP parameters are estimated by means of the 1000 nominal toy data sets, which are mentioned in Sec. 9.7.1. For each fixed parameter in question, the toy data sets are fitted varying the parameter up and down by 1σ corresponding to its uncertainties. Given a fictional parameter μ, the systematic uncertainty is calculated as

$$\sigma_{\text{syst}} = \sqrt{\langle\mu\rangle^2 + \sigma_{\text{shifts}}^2},\qquad(9.21)$$

where $\langle\mu\rangle$ is the average shift observed in the test fits computed as $\mu = (\Delta_{\text{up}} - \Delta_{\text{down}})/2$, and σ_{shifts} is the width of the μ distribution. The systematic correlations of the CP parameters are calculated from the covariances of the dataset of the test fit

results,

$$V_{ij} = \frac{1}{N-1} \sum_{k=1}^{N} (\mu_{ik} - \langle \mu_i \rangle)(\mu_{jk} - \langle \mu_j \rangle), \tag{9.22}$$

where (i, j) run over the *CP* parameters $(C, S_f, S_{\bar{f}}, A_f^{\Delta\Gamma}, A_{\bar{f}}^{\Delta\Gamma})$, and k runs over the toy experiments. Since a contribution to the error from a non-zero shift is also considered (notice the $\langle \mu \rangle$ term in Eq. (9.21)), an according correction is applied to the covariance as

$$V_{ij}' = \frac{V_{ij}}{\sqrt{V_{ii}V_{jj}}} \sigma_{\text{syst},i} \sigma_{\text{syst},j}. \tag{9.23}$$

The procedure is performed for the detection asymmetry A_{D} (see Eq. 9.8), the oscillation frequency Δm_s (see Eq. 9.20), and the asymmetries of the tagging efficiency (see Tab. 9.5). The resulting uncertainties are reported in Tab. 9.17 in Sec. 9.7.7. In case of the S_f and $S_{\bar{f}}$ parameters, Δm_s is among largest systematic uncertainties, but its size remains only about one fifth of the statistical uncertainties.

9.7.4 Decay-time resolution

The fast B_s^0 oscillations driven by Δm_s require an accurate description of the decay-time resolution. Systematic uncertainties originating from the choice of the decay-time resolution model are investigated by testing two alternative approaches for the calibration. Firstly, it is assumed that only the core of the decay-time distribution of the fake B_s^0 candidates is representative for the decay-time resolution. Hence, just the width of the narrow Gaussian is used to calculate the effective dilution following Eq. 9.10. By fitting the widths of this single Gaussian in bins of the per-event error the calibration

$$\sigma_{t,\text{eff}}(\sigma_{t'}) = s_0 + s_1 \cdot \sigma_{t'} = (-0.568 \pm 1.570)\,\text{fs} + (1.243 \pm 0.044)\,\sigma_{t'}, \tag{9.24}$$

is obtained. In the second alternative calibration approach a single Gaussian is fitted to the wider decay-time range of $[-3.25\,\overline{\sigma}_t, 1.3\overline{\sigma}_t]$, where $\overline{\sigma}_t$ is the numeric value of the centre of the respective per-event-error bin. The resulting calibration function of the decay-time resolution is given by

$$\sigma_{t,\text{eff}}(\sigma_{t'}) = 0 + s_1 \cdot \sigma_{t'} = (1.772 \pm 0.012)\,\sigma_{t'}. \tag{9.25}$$

The systematic uncertainties on the *CP* parameters are estimated by fitting the nominal 1000 toy datasets incorporating the two alternative resolution models. As the flavour-tagging calibration is correlated to the decay-time resolution, specific tagging calibrations are derived on $B_s^0 \to D_s^- \pi^+$ decays for the corresponding decay-time resolution models. The strategy to use corresponding sets of flavour tagging and decay-time resolution calibrations avoids double counting the systematic effects. Systematic uncertainties are estimated equivalent to Eq. 9.21, *i.e.* calculating the toy-by-toy difference between the nominal set of results and those using the alternative resolution. Table 9.17 contains the obtained uncertainties, which are well below the statistical uncertainties.

9.7.5 Correlations among the MD fit observables

The MD fit uses a PDF that is factorised in the different dimensions, *i.e.* correlations among the fit observables are ignored. In principle, such correlations are possible and might differ for signal and background decays. Therefore, extensive studies were done by collaborators to estimate if this strategy introduces systematic errors [159]. These studies make use of the TGenPhaseSpace generator, which is provided by the ROOT package [100]. This generator is able to produce any n-body decay corresponding to phase space distributions, given the kinematic properties of the particle and the masses of its decay products. In this way a type of MC is produced that is more complex than a toy MC, but still less computing intensive as full simulations (see sections 4.5.1 and 4.5.2). The amount of correlations in the TGenPhaseSpace MC matches that observed in fully simulated MC. Systematic uncertainties are obtained by including correlations in the generation but neglecting them in the fit. These correlations are in the range of 10 %. The systematic uncertainty that is estimated for the *CP* parameters is small compared to the statistical uncertainties (see Tab. 9.17), but it is the leading systematic uncertainty for the $A_f^{\Delta\Gamma}$ and $A_{\bar{f}}^{\Delta\Gamma}$ *CP* parameters. In future measurements of *CP* violation in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays, when the larger Run 2 data set is incorporated, the decision of ignoring the correlations in the fit needs to be reconsidered.

9.7.6 Decay-time acceptance

The spline parameters of the decay-time acceptance are fixed in the fit, which neglects their uncertainties. These uncertainties have two sources: the fit to the $B_s^0 \to D_s^- \pi^+$ data and the parameter ratio from simulations. Due to the former fit, the acceptance is correlated to the decay width Γ_s and the decay-width difference $\Delta\Gamma_s$. The latter parameters are also fixed parameters in the $B_s^0 \to D_s^{\mp} K^{\pm}$ *CP* fit and thus, potential sources of systematic uncertainty. For this reason a systematic uncertainty originating from the decay-time acceptance and the fixed values of Γ_s and $\Delta\Gamma_s$ is estimated simultaneously.

The nominal pseudo-experiments are fitted using values of the spline coefficients, Γ_s, and $\Delta\Gamma_s$, which are fixed to randomly generated values related to multidimensional correlated Gaussian distributions centred at the nominal values. The considered correlations among the acceptance parameters, Γ_s and $\Delta\Gamma_s$ are chosen according to the $B_s^0 \to D_s^- \pi^+$ data fit and Eq. 9.3. In Tab. 9.17, the combined correlated systematic uncertainty is listed as 'acceptance data fit, Γ_s, $\Delta\Gamma_s$'. A similar approach is used to account for the uncertainty of the acceptance parameters arising from the coefficient ratio obtained on the simulated $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ samples. In this case the values, which the parameters are fixed to in the fits to the toy data, are randomly generated taking only correlations among the spline parameters and their corresponding statistical uncertainties from the ratio (see Tab. 9.2) into account. The estimated systematic uncertainty is listed as 'acceptance, simulation ratio' in the table with the total systematic uncertainties.

9.7.7 Negligible effects and total systematic uncertainties

On top of the sources of systematic uncertainty discussed in the previous sections additional effects are investigated, but are found to have only negligible impact. These studies include varying the production asymmetry, checking missing or imperfectly modelled backgrounds and probing fixed signal-shape parameters in the MD fit. Potential systematic effects from fixing background yields are evaluated by means of pseudo-experiments. In these studies, the nominal value for the yields is used in the generation, while in the fits the yields are fixed to twice or half their nominal values. No systematic uncertainty is assigned, as there is no significant bias observed. Furthermore, the decay-time fit is repeated with one or two additional knots in the spline parameterisation of the decay-time acceptance. No significant change in the fit result is observed. For the imperfect knowledge of the momentum and the longitudinal dimensions of the detector no systematic uncertainties are applied, as they are already taken into account by the systematic uncertainty on Δm_s, since the world average is dominated by the LHCb measurement [158]. Table 9.17 contains the total systematic uncertainties on the *CP* parameters relative to their statistical uncertainties. The correlations between the total systematic uncertainties, which are listed in Tab. 9.18, are obtained by adding the covariance matrices corresponding to each source of systematic uncertainty. Although the systematic uncertainties are conservatively estimated, they are significantly smaller than the statistical ones. This leaves room for future improvements with the larger Run 2 data from LHCb. In the next chapter, the determination of γ, $r_{D_s K}$ and the strong phase δ is discussed.

Table 9.17: Systematic uncertainties on the *CP* parameters, relative to the statistical uncertainties.

Source	C	$A_f^{\Delta\Gamma}$	$A_{\bar{f}}^{\Delta\Gamma}$	S_f	$S_{\bar{f}}$
Closure test	0.13	0.19	0.19	0.12	0.12
Detection asymmetry	0.02	0.28	0.29	0.02	0.02
Δm_s	0.11	0.02	0.02	0.20	0.20
Tagging efficiency asymmetry	0.02	0.00	0.00	0.02	0.02
Decay-time resolution	0.18	0.02	0.02	0.16	0.18
Correlation among observables	0.20	0.38	0.38	0.20	0.18
Acceptance data fit, Γ_s, $\Delta\Gamma_s$	0.01	0.18	0.17	0.00	0.00
Acceptance, simulation ratio	0.01	0.10	0.10	0.01	0.01
Total	0.32	0.55	0.55	0.35	0.35

Table 9.18: Correlation matrix of the total systematic uncertainties of the *CP* parameters.

Parameter	C	$A_f^{\Delta\Gamma}$	$A_{\bar{f}}^{\Delta\Gamma}$	S_f	$S_{\bar{f}}$
C	1	0.05	0.03	0.03	-0.01
$A_f^{\Delta\Gamma}$		1	0.42	0.02	0.02
$A_{\bar{f}}^{\Delta\Gamma}$			1	0.03	0.03
S_f				1	0.01
$S_{\bar{f}}$					1

9.8 Determination of the CKM angle γ

In order to determine $\gamma - 2\beta_s$, and subsequently γ, from the *CP* observables obtained in the decay-time fit, the χ^2 function $\chi^2(\vec{\alpha}) = -2\ln\mathscr{L}(\vec{\alpha})$ with

$$\mathscr{L}(\vec{\alpha}) = \exp\left(-\frac{1}{2}\left(\vec{A}(\vec{\alpha}) - \vec{A}_{\mathrm{obs}}\right)^T V^{-1}\left(\vec{A}(\vec{\alpha}) - \vec{A}_{\mathrm{obs}}\right)\right), \qquad (9.26)$$

is minimised following the strategy described in Ref. [160,161]. In Eq. 9.26 the vector of the physics parameters is denoted by $\vec{\alpha} = (\gamma, \beta_s, r_{D_sK}, \delta)$, while $\vec{A}(\vec{\alpha})$ is the vector of parameters defined in Eq. 3.73. Furthermore, $\vec{A}_{\mathrm{obs}}$ represents the vector of the measured *CP* parameters, and V is the (statistical and systematic) uncertainty covariance matrix from the measurement. Confidence levels (CL) are calculated by evaluating the test statistic $\Delta\chi^2 \equiv \chi^2(\vec{\alpha}'_{\mathrm{min}}) - \chi^2(\vec{\alpha}_{\mathrm{min}})$ (again following Ref. [160, 161]). Here, $\vec{\alpha}_{\mathrm{min}}$ is the global minimum of Eq. 9.26, while $\vec{\alpha}'_{\mathrm{min}}$ denotes the conditional minimum when the parameter of interest is fixed to the tested value. As discussed in the context of Eq. 3.75, the world average of $\phi_s = -2\beta_s$ is used as an external input [57].

The following values are obtained for the physics parameters

$$68.3\,\% \text{ CL:} \quad \gamma = \left(128^{+17}_{-22}\right)^\circ, \quad \delta = \left(358^{+13}_{-14}\right)^\circ, \quad r_{D_sK} = 0.37^{+0.10}_{-0.09}, \qquad (9.27)$$

$$95.4\,\% \text{ CL:} \quad \gamma = \left(128^{+32}_{-50}\right)^\circ, \quad \delta = \left(358^{+27}_{-29}\right)^\circ, \quad r_{D_sK} = 0.37^{+0.20}_{-0.19}, \qquad (9.28)$$

$$99.7\,\% \text{ CL:} \quad \gamma = \left(128^{+47}_{-81}\right)^\circ, \quad \delta = \left(358^{+43}_{-50}\right)^\circ, \quad r_{D_sK} = 0.37^{+0.31}_{-0.31}, \qquad (9.29)$$

where the phases are expressed modulo 360°. Figure 9.9 shows the $1 - \text{CL}$ curve for γ. Figure 9.10 contains the two-dimensional contours of the profile likelihood $\mathscr{L}(\vec{\alpha}'_{\mathrm{min}})$. The result of the γ measurement is visualised in Fig. 9.11 by inspecting the measured amplitude coefficients in the complex plane. While the points determined by $(-A_f^{\Delta\Gamma}, S_f)$ and $(-A_{\bar{f}}^{\Delta\Gamma}, S_{\bar{f}})$ are proportional to $r_{D_sK}e^{i(\pm\delta - (\gamma - 2\beta_s))}$, an additional constraint arises on r_{D_sK} from C. An excellent agreement is observed for $(-A_f^{\Delta\Gamma}, S_f)$ and $(-A_{\bar{f}}^{\Delta\Gamma}, S_{\bar{f}})$.
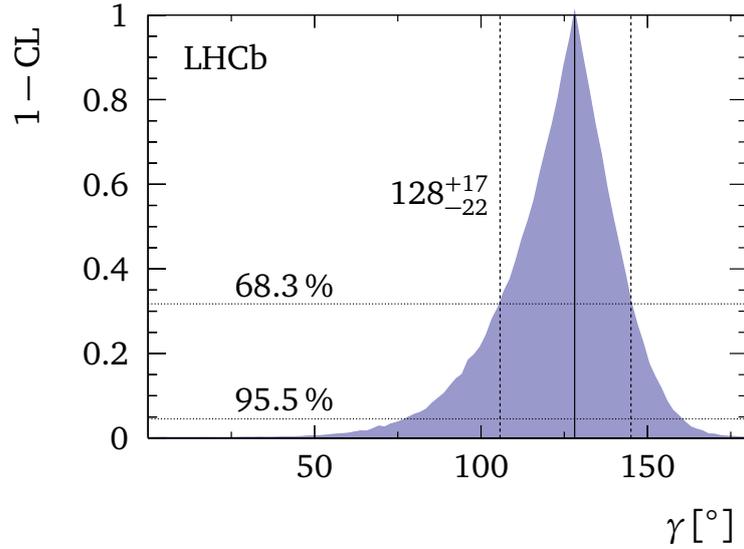
Figure 9.9: The graph shows $1 - \mathrm{CL}$ for the angle γ, together with the central value and the 68.3 % (95.5 %) CL interval as obtained from the frequentist method.

The significance of the measurement with respect to the presence of *CP* violation in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays is determined as follows. Firstly, the observed change in the fit log-likelihood between the combined best fit point and the origin in the complex plane is calculated. It indicates $3.8\,\sigma$ evidence for *CP* violation in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. Secondly, as a cross-check, the fit to $B_s^0 \to D_s^{\mp} K^{\pm}$ data is performed by applying the condition of no *CP* violation ($C = -C_{\bar{f}}$, $S_f = -S_{\bar{f}}$ and $A_f^{\Delta\Gamma} = A_{\bar{f}}^{\Delta\Gamma}$). Comparing the latter with the nominal fit leads to a significance of *CP* violation of $4\,\sigma$. Considering that the decay-time-dependent fit accounts for the statistical uncertainties only and that the difference in the number of degrees of freedom with the nominal fit has been neglected, the result is compatible to the $3.8\,\sigma$ obtained from the fit for γ.
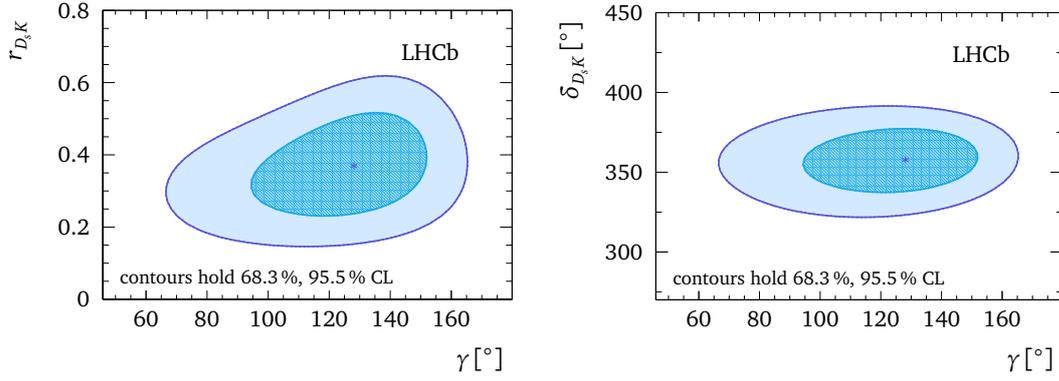
Figure 9.10: Profile likelihood contours of (left) r_{D_sK} vs. γ, and (right) δ vs. γ. The markers denote the best-fit values. The contours correspond to 68.3 % and 95.5 % CL, respectively.
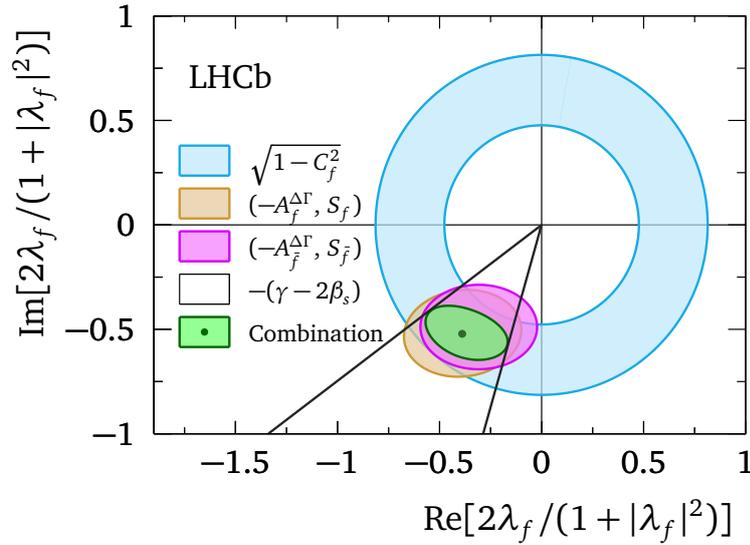
Figure 9.11: Based on the definitions of the *CP* parameters given in Eq. 3.73, the plot shows a visualisation of how each of the amplitude coefficients contributes towards the overall constraint on the weak phase, $\gamma - 2\beta_s$. The difference between the phase of $(-A_f^{\Delta\Gamma}, S_f)$ and $(-A_{\bar{f}}^{\Delta\Gamma}, S_{\bar{f}})$ is proportional to the strong phase δ, which is close to 360° and thus not indicated in the figure.

10 Discussion and Outlook

In this chapter the results of the measurement are discussed in the context of other γ measurements and the γ combination of the LHCb experiment (see Sec. 10.1). Furthermore, an outlook on future measurements of CP violation in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays is given in Sec. 10.2. The selection presented in this thesis will enable an updated measurement of the B_s^0 oscillation frequency Δm_s using the control channel $B_s^0 \to D_s^- \pi^+$. This upcoming measurement, which is briefly discussed in Sec. 10.3, will presumably lead to the world's best measurement of the parameter. The last section of this chapter introduces a new approach to describe the decay-time acceptance, which represents a potential improvement of the analysis strategy.

10.1 Compatibility of the γ measurement

Table 10.1 contains the results of the presented analysis and the previous $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis using the data taken in 2011 with 1 fb^{-1}. The results of the presented and the preceding measurements are compatible within less than $1\,\sigma$ regarding the uncertainties. Furthermore, a compatibility with the world average of $\gamma = \left(73.2^{+6.3}_{-7.0}\right)^{\circ}$ [57] is observed at the level of less than $2\,\sigma$. Another comparison can be performed with LHCb's γ combination. The LHCb collaboration computes a combination of all of its tree-level γ measurements and publishes an updated result from time-to-time, whenever added measurements lead to a significantly improved precision. The most recent γ combination, Ref. [162], represents the world's most precise determination

Table 10.1: The results for the CKM angle γ, the strong phase δ and the amplitudes ratio $r_{D_s K}$ (left column) in comparison to the previous results [75]. Both statistical and systematic uncertainties are considered. The intervals for the angles are expressed modulo 360°.

Parameter	This analysis (3 fb^{-1})	Previous analysis (1 fb^{-1})
γ	$\left(128^{+17}_{-22}\right)^{\circ}$	$\left(115^{+28}_{-43}\right)^{\circ}$
δ	$\left(358^{+13}_{-14}\right)^{\circ}$	$\left(3^{+19}_{-20}\right)^{\circ}$
$r_{D_s K}$	$0.37^{+0.10}_{-0.09}$	$0.53^{+0.17}_{-0.16}$

of γ from a single experiment to date, yielding

$$\gamma = \left(76.8^{+5.1}_{-5.7}\right)^{\circ} . \tag{10.1}$$

When removing the information from the measurement of γ with $B^0_s \to D^{\mp}_s K^{\pm}$ discussed in this thesis, a compatibility between LHCb's γ combination and the presented measurement is observed at the level of 2σ. The measurement of γ from $B^0_s \to D^{\mp}_s K^{\pm}$ decays has only a small impact on the γ combination at the moment, as the uncertainties are significantly larger than for example in the GLW measurement with $B^{\pm} \to D^{(*)} K^{\pm}$ and $B^{\pm} \to D^{(*)} \pi^{\pm}$ decays [163]. However, different to the latter analysis, there is no ambiguity in the measurement of γ from $B^0_s \to D^{\mp}_s K^{\pm}$ decays and hence, it remains a very useful input to the combination.

10.2 Perspective of decay-time-dependent γ measurements with $B^0_s \to D^{\mp}_s K^{\pm}$ decays

The LHC and hence the LHCb experiment are currently in the fourth year of the second period of data taking, the so-called Run 2. In the Run 2 the centre-of-mass energy is 13 TeV compared to the lower 7 and 8 TeV in 2011 and 2012, respectively. Since the start of the Run 2 in 2015, the LHCb detector took an additional 3.7 fb^{-1} of pp collisions. A first look into the data taken in the years 2015 and 2016, which corresponds to a sample size of about 2 fb^{-1}, shows about 9000 $B^0_s \to D^{\mp}_s K^{\pm}$ signal candidates. This number, which is only a very preliminary result without updating the selection, corresponds to an increase by 50 % compared to the full Run 1 sample. Hence, in Run 2 there are more candidates per recorded fb^{-1} of pp collisions than in Run 1. The reason is the larger $b\bar{b}$ cross section in Run 2, which scales approximately linearly with the centre-of-mass energy.

Apart from the increase in the amount of data the most promising way to improve the precision of the measurement originates in the flavour tagging. A straight forward enhancement of a future measurement is the addition of the OS charm tagger, a new tagging algorithm discussed in Sec. 6.6.3. In $B^0 \to J/\psi K^{*0}$, an effective tagging efficiency of $\varepsilon_{\mathrm{eff}} = 0.3\,\%$ is observed, which would correspond to about 5 % additional effective tagging efficiency in case of the presented analysis. Furthermore, the flavour tagging algorithms need to be retrained and optimised with respect to the Run 2 conditions in order to not loose tagging power. At the moment it is unclear whether the overall performance of the flavour tagging will in- or decrease in Run 2. On one hand, the increased energy will lead to higher p_{T} spectra of tagging particles, which will improve the purity of their selection. On the other hand, the larger amount of tracks will make it more challenging to maintain the excellent track reconstruction and tagging particle selection.

The current measurement shows a significance of *CP* violation of about $3.8\,\sigma$. In the next iteration of the measurement, *i.e.* using the full Run 2 data of approximately 5 fb^{-1}, it should be possible to exceed the level of $5\,\sigma$. Furthermore, in a $B^0_s \to D^{\mp}_s K^{\pm}$

analysis of the full Run 1 and Run 2 data set an uncertainty on γ in the range of 11°
can be expected.

Currently, the precision on γ is clearly worse than the precision of ϕ_s. Thus, it
is useful to utilise ϕ_s as an external input to measure γ with $B_s^0 \rightarrow D_s^\mp K^\pm$ decays.
However, the rapidly decreasing uncertainty of the world average of γ might allow
to provide an additional constrain on ϕ_s with $B_s^0 \rightarrow D_s^\mp K^\pm$ decays in the future while
using γ as an input.

10.3 Measurement of Δm_s with $B_s^0 \rightarrow D_s^\mp \pi^\pm$ decays

The most recent LHCb measurement of the B_s^0 oscillation frequency Δm_s was per-
formed using only the 1 fb^{-1} of data collected in the year 2011 [158]. The result of
the measurement, $\Delta m_s = 17.768 \pm 0.023\,(\text{stat}) \pm 0.006\,(\text{syst})\,\text{ps}^{-1}$, is the most pre-
cise measurement of the parameter to date and dominates the world average. With
the full Run 1 dataset and the improved selection presented in this thesis, a stat-
istical precision of about 0.013 ps can be expected (see Tab. 9.1). The decrease in
uncertainty is in line with the larger number of candidates. The published analysis
found about 34 000 candidates, while more than 95 000 candidates are observed in
the multidimensional fit to the $B_s^0 \rightarrow D_s^- \pi^+$ candidates here (see Tab. 8.5). The only
missing piece on the path to a publication of a measurement of Δm_s with the full
Run 1 dataset are studies on systematic uncertainties. These studies can built upon
the toy data samples and tools that have been created in the context of the presented
analysis.

10.4 Per-event decay-time acceptance

The modelling of the decay-time acceptance in the presented analysis depends on
data from the control channel $B_s^0 \rightarrow D_s^- \pi^+$ as well as on simulated $B_s^0 \rightarrow D_s^- \pi^+$ and
$B_s^0 \rightarrow D_s^\mp K^\pm$ events (see Sec. 9.4). Hence, multiple potential sources of systematic
uncertainties are involved in the current strategy. Still, in case of the presented Run 1
analysis, the overall estimated systematic uncertainty from the decay-time acceptance
is only about one fifth of the statistical uncertainty (see Tab. 9.17). However, with
the increased amount of data from the ongoing and the next data taking periods the
uncertainty from the current approach might contradict further improvements in the
measurement of *CP* violation with $B_s^0 \rightarrow D_s^\mp K^\pm$ decays.

A completely different approach to the modelling of the decay-time efficiency is
the so-called swimming. This method allows to extract the acceptance from the data,
here the reconstructed $B_s^0 \rightarrow D_s^\mp K^\pm$ candidates, itself. The swimming artificially moves
the position of the PV along the momentum of the B_s^0 meson and hence, gives the B_s^0
meson different decay times. For each position of the PV, the decay-time biasing
selection is rerun. In this way a decay-time acceptance is created that consists of a
series of step functions, as at each decay time a candidate either fails or passes the

selection. The swimming approach was first used at the NA11 spectrometer [164] and further developed by DELPHI [165, 166] and CDF [167]. The software trigger of the LHCb experiment can be rerun offline using the exact same configuration as applied when it runs online, which makes it very well suited for the the swimming strategy. Details on the realisation of the swimming algorithm for the LHCb experiment are given in Ref. [168]. The swimming has already successfully been used by several LHCb publications, *e.g.* Ref. [169] and Ref. [170].

The decay time biasing parts of the selection in the $B_s^0 \to D_s^\mp K^\pm$ analysis are the trigger, the stripping and the BDT. All parts need to be taken into account when running the swimming algorithm. Studies on simulated samples have been performed in order to test the swimming in context of the *CP* violation measurement with $B_s^0 \to D_s^\mp K^\pm$ decays. The studies revealed that it is technically challenging to include the BDT in the swimming. However, it is crucial that all BDT input variables are available and calculated correctly when the swimming algorithm checks whether a candidate passes or fails the BDT classifier cut value. During the testing and setting up of the swimming and the corresponding tools for the $B_s^0 \to D_s^\mp K^\pm$ analysis, the following strategy is applied. The swimming algorithm is run on simulated events to measure the intervals of the decay-time acceptance. Then, all steps of the selection are applied to the sample and swimming is used as a summed step function in a fit to the decay-time of the events. Finally, the measured lifetime of the fit is compared to the generated lifetime. After several iterations it was possible to reproduce the generated lifetime. However, the swimming approach was not used in the presented Run 1 analysis, mostly due to time constraints in the analysis process. It is very time and processing intensive to run the stripping on the full data set. At the point the bugs had been resolved with the help of simulated data, it was not feasible to wait for the swimming to finish on the detector data. The main advantage of the swimming is that per construction it delivers the correct decay-time acceptance for a given sample and hence, has no systematic uncertainties. However, as can be seen in Tab. 9.17, for the given size of the Run 1 dataset it was possible with the spline based approach to establish reasonable uncertainties that are well below the statistical uncertainties. In future measurements of *CP* violation with $B_s^0 \to D_s^\mp K^\pm$ decays, the amount of MC that is necessary to guarantee small enough uncertainties will rise. Furthermore, the swimming would allow to perform a flavour specific analysis of the $B_s^0 \to D_s^- \pi^+$ decay. Such an analysis represents a test of the SM as no *CP* violation is expected to exist in $B_s^0 \to D_s^- \pi^+$ due to its flavour-specific nature.

11 Conclusion

Phenomena like dark energy, dark matter, or the matter-antimatter asymmetry in the universe prove that our current understanding of the smallest constituents of matter is not complete. This is the main motivation for the LHC and its experiments. In the last couple of years the LHCb experiment established a leading role in the field of flavour physics. One of the main goals of the LHCb collaboration is to improve our understanding of the *CP* violation and the CKM sector of the Standard Model. With precision measurements of the involved parameters, a powerful indirect search for New Physics is carried out. This is the context of the measurement of *CP* violation in $B_s^0 \to D_s^\mp K^\pm$ decays [15] that is presented in this thesis. The analysis uses the full Run 1 dataset of the LHCb experiment and determines the *CP* parameters to be

$$
\begin{aligned}
C &= 0.73 \pm 0.14 \pm 0.05\,, \\
A_f^{\Delta\Gamma} &= 0.39 \pm 0.28 \pm 0.15\,, \\
A_{\bar{f}}^{\Delta\Gamma} &= 0.31 \pm 0.28 \pm 0.15\,, \\
S_f &= -0.52 \pm 0.20 \pm 0.07\,, \\
S_{\bar{f}} &= -0.49 \pm 0.20 \pm 0.07\,,
\end{aligned}
$$

where the first uncertainties are statistical and the second are systematic. The results are compatible with the preceding measurement [75] at the level of less than $2\,\sigma$ taking correlations into account. From the measured *CP* parameters, the least well-known CKM angle is determined to be $\gamma = \left(128^{+17}_{-22}\right)^\circ$, which agrees at the level of less than $1\,\sigma$ with the previous results obtained from $B_s^0 \to D_s^\mp K^\pm$ decays. The presented measurement represents the most precise determination of γ from B_s^0 meson decays to date. Evidence for *CP* violation in $B_s^0 \to D_s^\mp K^\pm$ decays is observed at the level of $3.8\,\sigma$.

While the analysis is one of the most challenging ones performed by the LHCb collaboration, its precision is not limited by systematic uncertainties. Hence, superseding analyses can profit from the larger Run 2 dataset and the increased $b\bar{b}$ production cross section, which itself is a consequence of the higher centre-of-mass energy of 13 TeV. However, with the increased amount of data it has to be checked whether the smaller statistical uncertainties and the systematic uncertainty arising from neglecting correlations among the MD fit observables level up. In this case, the strategy of the MD fit needs to be changed by taking the correlations between the observables into account. Another major challenge of the next $B_s^0 \to D_s^\mp K^\pm$ analysis is the flavour tagging. Currently, the algorithms are being retrained for the conditions of Run 2, but it is not certain if the performance will de- or increase. Further improvements of the

analysis can be reached by adding the decays $B_s^0 \to D_s^{*\mp} K^{\pm}$ and $B_s^0 \to D_s^{\mp} K^{*\pm}$. These decays allow measurements of γ analogously to the strategy applied to $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. While their reconstruction is more challenging due to the additional decays of the excited mesons in the final state, they can still add to the picture. Currently, analyses of these decays are ongoing following the strategy developed for $B_s^0 \to D_s^{\mp} K^{\pm}$ decays.

To date, there is no clear evidence for physics beyond the Standard Model at the level of elementary particle interactions. This holds for both direct and indirect searches. The current constraints on the CKM triangle show that the wide range of different measurements are very well in agreement. As an example, the current LHCb combination of γ measurements from tree-level processes [162] is compatible with a recent LHCb γ measurement that involves loop processes [171] within about 2σ of the involved uncertainties. However, the precision will be drastically improved in the next years. The LHCb collaboration is going to update most of its measurements based on the Run 2 data. In parallel, a substantial upgrade of the LHCb detector is being prepared and going to be implemented in the next long shutdown from 2019–2020. Afterwards, the detector will be read out at the full inelastic collision rate of 30 MHz, representing a strong challenge for the then solely software-based trigger system [172]. Further improvements, especially in the field of B^0 mesons, can be expected from the Belle II experiment. It is the successor of the Belle B factory and will start its data taking this year. Belle II will deliver a complementary approach and challenge the uncertainties of LHCb's measurements. It is certain that enormous efforts will be necessary, which especially involves the combination of results from many different measurements and experiments, to finally reveal hints for New Physics and to pave the way for a new SM.

Bibliography

[1] E598 collaboration, J. J. Aubert *et al.*, *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.

[2] G. S. Abrams *et al.*, *The Discovery of a Second Narrow Resonance in $e^+ e^-$ Annihilation*, Phys. Rev. Lett. **33** (1974) 1453.

[3] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, `arXiv:1207.7214`.

[4] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, `arXiv:1207.7235`.

[5] ATLAS collaboration, CMS collaboration, G. Aad *et al.*, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803, `arXiv:1503.07589`.

[6] A. D. Sakharov, *Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[7] Super-Kamiokande collaboration, K. Abe *et al.*, *Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in 0.31 megaton $\cdot$ years exposure of the super-kamiokande water cherenkov detector*, Phys. Rev. D **95** (2017) 012004, `arXiv:1610.03597`.

[8] W. Bernreuther, *CP violation and baryogenesis*, Lect. Notes Phys. **591** (2002) 237, `arXiv:hep-ph/0205279`.

[9] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[10] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. **D51** (1995) 379, `arXiv:hep-ph/9404302`.

[11] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[12] I. Dunietz and R. G. Sachs, *Asymmetry Between Inclusive Charmed and Anti-charmed Modes in B0, Anti-b0 Decay as a Measure of CP Violation*, Phys. Rev. **D37** (1988) 3186, Erratum ibid. **D39** (1989) 3515.

[13] R. Aleksan, I. Dunietz, and B. Kayser, *Determining the CP violating phase gamma*, Z. Phys. **C54** (1992) 653.

[14] R. Fleischer, *New strategies to obtain insights into CP violation through $B_s^0 \to D_s^{\pm} K^{\mp}$, $D_s^{*\pm} K^{\mp}$, ... and $B^0 \to D^{\pm} \pi^{\mp}, D^{*\pm} \pi^{\mp}$, ... decays*, Nucl. Phys. **B671** (2003) 459, `arXiv:hep-ph/0304027`.

[15] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays*, JHEP **03** (2018) 059, `arXiv:1712.07428`.

[16] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and the B_s^0 meson decay width difference with $B_s^0 \to J/\psi K^+ K^-$ and $B_s^0 \to J/\psi \pi^+ \pi^-$ decays*, Phys. Rev. **D87** (2013) 112010, `arXiv:1304.2600`.

[17] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley Publishing, 1995.

[18] D. J. Griffiths, *Introduction to elementary particles*, Wiley, 2008.

[19] K. G. Wilson, *Confinement of Quarks*, Phys. Rev. **D10** (1974) 2445.

[20] LHCb collaboration, R. Aaij *et al.*, *Observation of exotic $J/\psi \phi$ structures from amplitude analysis of $B^+ \to J/\psi \phi K^+$ decays*, Phys. Rev. Lett. **118** (2016) 022003, `arXiv:1606.07895`.

[21] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^+ \to J/\psi \phi K^+$ decays*, Phys. Rev. **D95** (2016) 012002, `arXiv:1606.07898`.

[22] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, `arXiv:1507.03414`.

[23] LHCb collaboration, R. Aaij *et al.*, *Evidence for exotic hadron contributions to $\Lambda_b^0 \to J/\psi p \pi^-$ decays*, Phys. Rev. Lett. **117** (2016) 082003, Erratum ibid. **117** (2016) 109902, **118** (2017) 119901, `arXiv:1606.06999`.

[24] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001.

[25] E. Noether, *Invariante Variationsprobleme*, Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse (1918) 235.

[26] W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, Z. Phys. **43** (1927) 172.

[27] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[28] H. Georgi and S. L. Glashow, *Unified weak and electromagnetic interactions without neutral currents*, Phys. Rev. Lett. **28** (1972) 1494.

[29] G. Lüders, *On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories*, Kong. Dan. Vid. Sel. Mat. Fys. Med. **28N5** (1954) 1.

[30] G. Lüders, *Proof of the TCP theorem*, Ann. Phys. **2** (1957) 1.

[31] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.

[32] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[33] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[34] D0 collaboration, S. Abachi *et al.*, *Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV*, Phys. Rev. Lett. **74** (1995) 2422, `arXiv:hep-ex/9411001`.

[35] CDF collaboration, F. Abe *et al.*, *Observation of top quark production in $\bar{p}p$ collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626, `arXiv:hep-ex/9503002`.

[36] B. T. Cleveland *et al.*, *Measurement of the solar electron neutrino flux with the Homestake chlorine detector*, Astrophys. J. **496** (1998) 505.

[37] Super-Kamiokande collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, `arXiv:hep-ex/9807003`.

[38] SNO collaboration, Q. R. Ahmad *et al.*, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, `arXiv:nucl-ex/0204008`.

[39] T. Asaka and M. Shaposhnikov, *The νMSM, dark matter and baryon asymmetry of the universe*, Phys. Lett. **B620** (2005) 17, `arXiv:hep-ph/0505013`.

[40] T. Asaka, S. Blanchet, and M. Shaposhnikov, *The νMSM, dark matter and neutrino masses*, Phys. Lett. **B631** (2005) 151, `arXiv:hep-ph/0503065`.

[41] E. Majorana, *Teoria simmetrica dell'elettrone e del positrone*, Nuovo Cim. **14** (1937) 171.

[42] Planck collaboration, P. A. R. Ade *et al.*, *Planck 2015 results. XIII. Cosmological parameters*, Astron. Astrophys. **594** (2016) A13, `arXiv:1502.01589`.

[43] H. W. Babcock, *The rotation of the Andromeda Nebula*, Lick Observatory Bulletin **19** (1939) 41.

[44] V. Trimble, *Existence and Nature of Dark Matter in the Universe*, Ann. Rev. Astron. Astrophys. **25** (1987) 425.

[45] G. Paal, I. Horvath, and B. Lukacs, *Inflation and compactification from galaxy redshifts?*, Astrophys Space Sci **191** (1992) 107.

[46] P. J. E. Peebles and B. Ratra, *The Cosmological constant and dark energy*, Rev. Mod. Phys. **75** (2003) 559, `arXiv:astro-ph/0207347`.

[47] G. G. Ross, *Grand Unified Theories*, Westview Press, 1984.

[48] J. Wess and B. Zumino, *Supergauge Transformations in Four-Dimensions*, Nucl. Phys. **B70** (1974) 39.

[49] P. Fayet, *Spontaneously Broken Supersymmetric Theories of Weak, Electromagnetic and Strong Interactions*, Phys. Lett. **69B** (1977) 489.

[50] ATLAS collaboration, CMS collaboration, C. Autermann, *Search for Supersymmetry at the LHC*, EPJ Web Conf. **164** (2017) 01028.

[51] BaBar collaboration, D. Boutigny *et al.*, *The BABAR physics book: Physics at an asymmetric B factory*, 1998.

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

[53] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802.

[54] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[55] 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.

[56] CKMfitter Group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, `arXiv:hep-ph/0406184`, updated results and plots available at: http://ckm-fitter.in2p3.fr.

[57] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, `arXiv:1612.07233`, updated results and plots available at: http://www.slac.stanford.edu/xorg/hfag/.

[58] NA31 collaboration, H. Burkhardt *et al.*, *First Evidence for Direct CP Violation*, Phys. Lett. **B206** (1988) 169.

[59] NA48 collaboration, V. Fanti *et al.*, *A New measurement of direct CP violation in two pion decays of the neutral kaon*, Phys. Lett. **B465** (1999) 335, `arXiv:hep-ex/9909022`.

[60] KTeV collaboration, A. Alavi-Harati *et al.*, *Observation of direct CP violation in $K_{S,L} \to \pi\pi$ decays*, Phys. Rev. Lett. **83** (1999) 22, `arXiv:hep-ex/9905060`.

[61] Belle collaboration, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, `arXiv:hep-ex/0107061`.

[62] BaBar collaboration, B. Aubert *et al.*, *Measurement of CP violating asymmetries in B^0 decays to CP eigenstates*, Phys. Rev. Lett. **86** (2001) 2515, `arXiv:hep-ex/0102030`.

[63] Belle collaboration, Y.-T. Duh *et al.*, *Measurements of branching fractions and direct CP asymmetries for $B \to K\pi$, $B \to \pi\pi$ and $B \to KK$ decays*, Phys. Rev. **D87** (2013) 031103, `arXiv:1210.1348`.

[64] BaBar collaboration, J. P. Lees *et al.*, *Measurement of CP Asymmetries and Branching Fractions in Charmless Two-Body B-Meson Decays to Pions and Kaons*, Phys. Rev. **D87** (2013) 052009, `arXiv:1206.3525`.

[65] LHCb collaboration, R. Aaij *et al.*, *First observation of CP violation in the decays of B_s^0 mesons*, Phys. Rev. Lett. **110** (2013) 221601, `arXiv:1304.6173`.

[66] LHCb collaboration, R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in B^0–$\bar{B}^0$ mixing*, Phys. Rev. Lett. **114** (2015) 041601, `arXiv:1409.8586`.

[67] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in B_s^0–$\bar{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016) 061803, `arXiv:1605.09768`.

[68] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B^0 \to J/\psi K_S^0$ decays*, Phys. Rev. Lett. **115** (2015) 031601, `arXiv:1503.07089`.

[69] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of CP violation in $B_s^0 \to J/\psi K^+ K^-$ decays*, Phys. Rev. Lett. **114** (2015) 041801, `arXiv:1411.3104`.

[70] T. Inami and C. S. Lim, *Effects of superheavy quarks and leptons in low-energy weak processes $K_L \to \mu\bar{\mu}$, $K^+ \to \pi^+ \nu\bar{\nu}$ and $K^0 \leftrightarrow \overline{K^0}$*, Prog. Theor. Phys. **65** (1981) 297, Erratum ibid. **65** (1981) 1772.

[71] BaBar collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP-violating asymmetries and constraints on $\sin(2\beta + \gamma)$ with partial reconstruction of $B \to D^{*\mp}\pi^{\pm}$ decays*, Phys. Rev. **D71** (2005) 112003, `arXiv:hep-ex/0504035`.

[72] BaBar collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP asymmetries in $B^0 \to D^{(*)\pm} \pi^{\mp}$ and $B^0 \to D^{\pm}\rho^{\mp}$ decays*, Phys. Rev. **D73** (2006) 111101, `arXiv:hep-ex/0602049`.

[73] Belle collaboration, F. J. Ronga *et al.*, *Measurements of CP violation in $B^0 \to D^{*-}\pi^+$ and $B^0 \to D^-\pi^+$ decays*, Phys. Rev. **D73** (2006) 092003, `arXiv:hep-ex/0604013`.

[74] Belle collaboration, S. Bahinipati *et al.*, *Measurements of time-dependent CP asymmetries in $B \to D^{*\mp}\pi^{\pm}$ decays using a partial reconstruction technique*, Phys. Rev. **D84** (2011) 021101, `arXiv:1102.0888`.

[75] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \to D_s^{\mp}K^{\pm}$ decays*, JHEP **11** (2014) 060, `arXiv:1407.6127`.

[76] M. Gronau and D. Wyler, *On determining a weak phase from CP asymmetries in charged B decays*, Phys. Lett. **B265** (1991) 172.

[77] A. J. Buras, M. E. Lautenbacher, and G. Ostermaier, *Waiting for the top quark mass, $K^+ \to \pi^+ \nu\bar{\nu}$, B_s^0–$\bar{B}_s^0$ mixing and CP asymmetries in B decays*, Phys. Rev. **D50** (1994) 3433, `arXiv:hep-ph/9403384`.

[78] D. Atwood, I. Dunietz, and A. Soni, *Enhanced CP violation with $B \to KD^0(\bar{D}^0)$ modes and extraction of the CKM angle γ*, Phys. Rev. Lett. **78** (1997) 3257, `arXiv:hep-ph/9612433`.

[79] D. Atwood, I. Dunietz, and A. Soni, *Improved methods for observing CP violation in $B^{\pm} \to KD$ and measuring the CKM phase γ*, Phys. Rev. **D63** (2001) 036005, `arXiv:hep-ph/0008090`.

[80] Y. Grossman, Z. Ligeti, and A. Soffer, *Measuring γ in $B^{\pm} \to K^{\pm}(KK^*)_D$ decays*, Phys. Rev. **D67** (2003) 071301, `arXiv:hep-ph/0210433`.

[81] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, *Determining γ using $B^{\pm} \to DK^{\pm}$ with multibody D decays*, Phys. Rev. **D68** (2003) 054018, `arXiv:hep-ph/0303187`.

[82] J. Brod and J. Zupan, *The ultimate theoretical error on γ from $B \to DK$ decays*, JHEP **01** (2014) 051, `arXiv:1308.5663`.

[83] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[84] C. De Melis, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, Jul, 2016. General Photo.

[85] ATLAS collaboration, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[86] CMS collaboration, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[87] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[88] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[89] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, `arXiv:hep-ph/0611148`.

[90] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.

[91] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[92] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, `arXiv:1311.3893`.

[93] P. A. Cerenkov, *Visible luminescence of pure fluids iduced by gamma rays*, Dokl. Akad. Nauk Ser. Fiz. **2** (1934) 451.

[94] P. A. Cerenkov, *Visible radiation produced by electrons moving in a medium with velocities exceeding that of light*, Phys. Rev. **52** (1937) 378.

[95] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[96] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[97] LHCb HLT project, J. Albrecht, V. V. Gligorov, G. Raven, and S. Tolk, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, `arXiv:1310.8544`.

[98] M. Frank *et al.*, *Deferred High Level Trigger in LHCb: A Boost to CPU Resource Utilization*, J. Phys. Conf. Ser. **513** (2014) 012006.

[99] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.

[100] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[101] I. Antcheva *et al.*, *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, Comput. Phys. Commun. **182** (2011) 1384.

[102] R. Aaij *et al.*, *The Moore project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/, Apr., 2017.

[103] R. Aaij *et al.*, *The Brunel project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/, Apr., 2017.

[104] R. Aaij *et al.*, *The DaVinci project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/, Apr., 2017.

[105] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, J. Basic Eng. **82** (1960) 35.

[106] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[107] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[108] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[109] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[110] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[111] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[112] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, arXiv:hep-ph/0506026.

[113] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[114] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[115] R. Aaij *et al.*, *The Boole project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/, Apr., 2017.

[116] *LHCb Integrated Recorded Luminosity in pp, 2010-2016*, http://lhcb-operationsplots.web.cern.ch/lhcb-operationsplots/index.htm (visited on 2018-02-24).

[117] R. Alemany-Fernandez, F. Follin, and R. Jacobsson, *The LHCb Online Luminosity Control and Monitoring*, in *Proceedings, 4th International Particle Accelerator Conference (IPAC 2013)*, 2013.

[118] 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.

[119] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[120] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[121] C. Gini, *Variabilità e mutuabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche.*, Tipogr. di P. Cuppini, 1912.

[122] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, J. Comput. Syst. Sci. **55** (1997) 119.

[123] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*, Ann. Statist. **29** (2001) 1189.

[124] A. Hocker *et al.*, *TMVA - toolkit for multivariate data analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[125] F. Pedregosa *et al.*, *Scikit-learn: Machine Learning in Python*, J. Machine Learning Res. **12** (2011) 2825, `arXiv:1201.0490`.

[126] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, in *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016.

[127] LHCb collaboration, *Optimization and calibration of the LHCb flavour tagging performance using 2010 data*, LHCb-CONF-2011-003.

[128] M. Grabalosa, *Flavour Tagging developments within the LHCb experiment*, PhD thesis, Barcelona U., 2012, CERN-THESIS-2012-075.

[129] LHCb collaboration, R. Aaij *et al.*, *B flavour tagging using charm decays at the LHCb experiment*, JINST **10** (2015) P10005, `arXiv:1507.07892`.

[130] LHCb collaboration, R. Aaij *et al.*, *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016) P05010, `arXiv:1602.07252`.

[131] LHCb collaboration, R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, `arXiv:1202.4979`.

[132] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

Bibliography

[133] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001, and 2017 update.

[134] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.

[135] A. J. Bevan *et al.*, *The Physics of the B Factories*, Eur. Phys. J. **C74** (2014) 3026, arXiv:1406.6311.

[136] LHCb collaboration, R. Aaij *et al.*, *Measurement of the time-dependent CP asymmetry in $B^0 \to J/\psi K_s^0$ decays*, Phys. Lett. **B721** (2013) 24, arXiv:1211.6093.

[137] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B \to D^+ D^-$ decays*, Phys. Rev. Lett. **117** (2016) 261801, arXiv:1608.06620.

[138] LHCb collaboration, R. Aaij *et al.*, *Determination of the branching fractions of $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B^0 \to D_s^- K^+$*, JHEP **02** (2015) 029, arXiv:1412.7654.

[139] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286, f_s/f_d value updated in LHCb-CONF-2013-011.

[140] V. V. Gligorov, *A single track HLT1 trigger*, LHCb public note LHCb-PUB-2011-003, CERN, 2011.

[141] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, LHCb public note LHCb-PUB-2011-016, CERN, 2011.

[142] B. H. Denby, *Neural Networks and Cellular Automata in Experimental High-energy Physics*, Comput. Phys. Commun. **49** (1988) 429.

[143] L. Breiman, *Bagging Predictors*, Machine Learning **24** (1996) 123.

[144] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577, arXiv:physics/0408124.

[145] L. Breiman, *Random Forests*, Machine Learning **45** (2001) 5.

[146] A. Kolmogorov, *Sulla determinazione empirica di una legge di distribuzione*, G. Ist. Ital. Attuari **4** (1933) 83.

[147] N. Smirnov, *Table for estimating the goodness of fit of empirical distributions*, Ann. Math. Statist. **19** (1948) 279.

[148] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[149] A. Powell *et al.*, *Particle identification at LHCb*, PoS **ICHEP2010** (2010) 020, LHCb-PROC-2011-008.

[150] A. Dziurda, *Studies of time-dependent CP violation in charm decays of B_s^0 mesons*, PhD thesis, Kracow, INP, 2015-04-01.

[151] L. Landau, *On the energy loss of fast particles by ionization*, J. Phys. (USSR) **8** (1944) 201.

[152] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\bar{B}^0$–B^0 and $\bar{B}_s^0$–B_s^0 production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218, arXiv:1408.0275.

[153] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, JHEP **07** (2014) 041, arXiv:1405.2797.

[154] H.-G. Moser and A. Roussarie, *Mathematical Methods for $B^0\bar{B}^0$ Oscillation Analyses*, Nucl. Instrum. Meth. **A384** (1997) 491 .

[155] C. de Boor, *A practical guide to splines*, vol. 27 of *Applied Mathematical Sciences*, Springer-Verlag New York, 1978.

[156] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation*, arXiv:1407.0748.

[157] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, LHCb-PUB-2011-001.

[158] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the B_s^0–$\bar{B}_s^0$ oscillation frequency in the decay $B_s^0 \to D_s^-\pi^+$*, New J. Phys. **15** (2013) 053021, arXiv:1304.4741.

[159] G. Tellarini, *B-flavour tagging calibration for CP violation measurement in $B_s^0 \to D_s^\mp K^\pm$ decays at LHCb*, PhD thesis, Università degli Studi di Ferrara, 2015.

[160] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of $B^\pm \to Dh^\pm$ analyses*, Phys. Lett. **B726** (2013) 151, arXiv:1305.2050.

[161] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of LHCb results*, JHEP **12** (2016) 087, arXiv:1611.03076.

[162] LHCb collaboration, *Measurement of the CKM angle γ from a combination of $B \to DK$ analyses*, LHCb-CONF-2017-004.

[163] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP observables in $B^\pm \to D^{(*)}K^\pm$ and $B^\pm \to D^{(*)}\pi^\pm$ decays*, Phys. Lett. **B777** (2017) 16, arXiv:1708.06370.

[164] R. Bailey *et al.*, *Measurement of the lifetime of charged and neutral D mesons with high resolution silicon strip detectors*, Z. Phys. **C28** (1985) 357.

[165] DELPHI collaboration, P. Abreu *et al.*, *A Measurement of the mean lifetimes of charged and neutral B hadrons*, Phys. Lett. **B312** (1993) 253.

[166] J. Rademacker, *Reduction of statistical power per event due to upper lifetime cuts in lifetime measurements*, Nucl. Instrum. Meth. **A570** (2007) 525, arXiv:hep-ex/0502042.

[167] CDF collaboration, T. Aaltonen *et al.*, *Measurement of the B^- lifetime using a simulation free approach for trigger bias correction*, Phys. Rev. **D83** (2011) 032008, arXiv:1004.4855.

[168] V. V. Gligorov *et al.*, *Swimming: A data driven acceptance correction algorithm*, J. Phys. Conf. Ser. **396** (2012) 022016.

[169] LHCb collaboration, R. Aaij *et al.*, *Measurement of the effective $B_s^0 \to K^+K^-$ lifetime*, Phys. Lett. **B707** (2012) 349, arXiv:1111.0521.

[170] LHCb collaboration, R. Aaij *et al.*, *Measurement of mixing and CP violation parameters in two-body charm decays*, JHEP **04** (2012) 129, arXiv:1112.4698.

[171] LHCb collaboration, R. Aaij *et al.*, *Determination of γ and $-2\beta_s$ from charmless two-body decays of beauty mesons*, Phys. Lett. **B739** (2015) 1, arXiv:1408.4368.

[172] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016. LHCb-TDR-016.

Acknowledgements