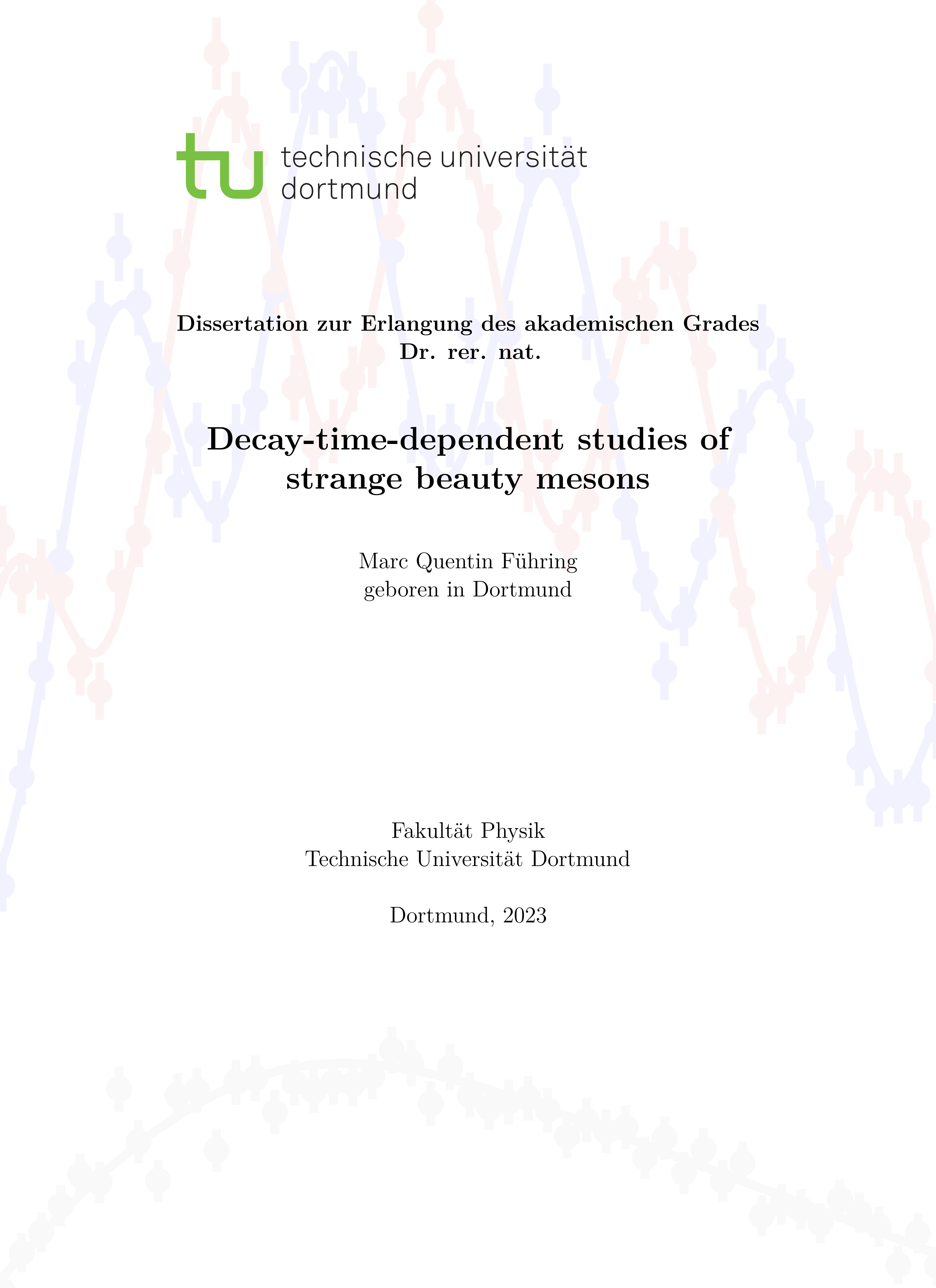


technische universität
dortmund

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

Decay-time-dependent studies of strange beauty mesons

Marc Quentin Führing
geboren in Dortmund

Fakultät Physik
Technische Universität Dortmund

Dortmund, 2023

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

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

Vorsitzender der Prüfungskommission:
Prof. Dr. Herre Jelger Risselada

Vertreter der wissenschaftlichen Mitarbeiter:
Dr. Ulf Berges

Datum der Einreichung der Dissertation:
16. September 2023

Datum der mündlichen Prüfung:
22. November 2023

Abstract

In this thesis, two decay-time-dependent analyses of B_s^0-meson decays are presented. Precise measurements are performed using proton-proton collision data recorded by the LHCb experiment between 2015 and 2018. Based on the reconstruction and selection of $378\,700 \pm 700$ $B_s^0 \to D_s^- \pi^+$ decays, the most precise measurement of the B_s^0-$\overline{B}_s^0$ oscillation frequency

$$\Delta m_s = (17.7683 \pm 0.0051 \pm 0.0032)\ \text{ps}^{-1}$$

in the world to date is achieved. This measurement alone is an essential test of the Standard Model of particle physics. In addition, it is a crucial ingredient to increase the precision of decay-time-dependent measurements of other B_s^0-decay modes. The presented analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays benefits from this updated measurement of Δm_s and from further improvements of algorithms and methods. The CKM angle

$$\gamma = (74 \pm 11)^{\circ}\,,$$

an important Standard Model parameter, is measured based on $20\,950 \pm 180$ selected decays. Compared to the analysis using data from 2011 and 2012, the precision is improved beyond what would be expected from the increase in data sample sizes.

Kurzfassung

In dieser Dissertation werden zwei zerfallszeitabhängige Messungen von B_s^0-Meson Zerfällen vorgestellt. Beide Messungen basieren auf der Analyse eines Datensatzes von Proton-Proton Kollisionen aufgezeichnet zwischen 2015 und 2018 am LHCb-Experiment. Die Rekonstruktion und Selektion von $378\,700 \pm 700$ $B_s^0 \to D_s^- \pi^+$ Zerfällen ermöglicht die zur Zeit weltweit präziseste Messung der B_s^0-$\overline{B}_s^0$ Oszillationsfrequenz

$$\Delta m_s = (17.7683 \pm 0.0051 \pm 0.0032)\ \text{ps}^{-1}\,.$$

Die Messung von Δm_s selbst ist ein wichtiger Test des Standardmodells der Teilchenphysik. Darüber hinaus liefert sie einen wesentlichen Beitrag um die Präzision weiterer zerfallszeitabhängiger Messungen von B_s^0-Meson Zerfällen zu steigern. Die vorgestellte Analyse von $B_s^0 \to D_s^{\mp} K^{\pm}$ Zerfällen profitiert von der präziseren Kenntnis von Δm_s, sowie von verbesserten Verfahren und Algorithmen. Der CKM Winkel

$$\gamma = (74 \pm 11)^{\circ}\ ,$$

ein wichtiger Standardmodell Parameter, wird anhand von $20\,950 \pm 180$ selektierten Zerfällen bestimmt. Im Vergleich zu der vorherigen Analyse der in 2011 und 2012 aufgezeichneten Daten fällt die Verbesserung damit höher aus, als die erhöhte Statistik alleine erwarten lässt.

Contents

1 Introduction

The decay-time-dependent analyses of strange beauty (B_s^0) mesons provide an excellent environment to prove the fascinating nature of quantum mechanics and to test the Standard Model of particle physics. The Standard Model (SM) [3–5], a relativistic quantum-field theory briefly introduced in Chapter 2, describes matter and its interactions in the microscopic world at remarkable precision. However, the Standard Model in its current form cannot describe all observations. The list of shortcomings is growing, ranging from neglecting gravitation, over not including descriptions of dark matter and energy [6], to a missing inclusion of neutrino oscillations [7]. This list motivates the development of new, more generalised theories, direct searches for New Physics phenomena and indirect searches for imperfections of the SM at high precision.

A well-suited laboratory for such precision measurements is the system of neutral B_s^0 mesons [8–11] discussed in Chapter 3. The large number of possible quark transitions via the weak interaction enables observable phenomena giving experimental access to various SM parameters. Two of these phenomena, studied in this thesis, are the B_s^0-$\overline{B}_s^0$ oscillation and *CP* violation. The oscillation between matter and antimatter state is an impressive example of the nature of quantum mechanics, while *CP* violation shows the contra-intuitively different behaviour of matter and antimatter.

The LHCb experiment [12] at the Large Hadron Collider (LHC) [13], presented in Chapter 4, is currently the best choice to perform decay-time-dependent studies in the B_s^0 system. Due to the high luminosity and collision energy at the LHC, the LHCb experiment provides the largest data set of B_s^0 mesons to date. Further, as it is designed to study B meson decays, it provides the necessary instrumentation of the forward region and an excellent vertex detector. The provided decay-time resolution of $\mathcal{O}(50\,\text{fs})$ is crucial to resolve the fast oscillations of the B_s^0 meson.

In this thesis, two analyses of B_s^0-meson tree-level decays into a $D_s^{\mp}$ meson and a charged, light hadron $h^{\pm}$ are presented utilising the LHCb data set [1, 2]. This data set, recorded between 2015 and 2018 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, corresponds to an integrated luminosity of $6\,\text{fb}^{-1}$. Both of these analyses have been performed by an international group of researchers within the LHCb collaboration with major contributions from the author of this thesis.

For the flavour-specific $B_s^0 \to D_s^- \pi^+$ decays, the presence of *CP* violating effects can be neglected even at very high precision. This allows the oscillation frequency Δm_s to be measured at remarkable precision in Part II of this thesis. The published result [1] is the most precise measurement of this quantity to date, dominating the updated average. Further, it is an important input to measurements of decay-time-dependent *CP* violation in B_s^0 mesons.

Such a measurement is presented in Part III of this thesis. Here, the oscillation frequency is constrained in the measurement of *CP* violation in $B_s^0 \to D_s^\mp K^\pm$ decays. The comparable magnitude of $B_s^0 \to D_s^- K^+$ and $B_s^0 \to D_s^+ K^-$ transitions enables sizeable interference effects in these decays. The observed *CP* violation is used to extract the CKM angle γ, which probes the consistency of the SM. The presented result is the most precise decay-time-dependent measurement of γ to date [2]. The systematic uncertainties have been significantly reduced compared to a previous analysis of $B_s^0 \to D_s^\mp K^\pm$ decays [14] by using new techniques and updated inputs such as the oscillation frequency Δm_s.

An essential technique in these analyses is the so-called flavour tagging. Here, machine learning algorithms are used to infer the not-directly measurable flavour of the B meson at its production by exploiting various mechanisms during the signal hadronisation. A prototype of a new inclusive flavour tagging algorithm, analysing the entire hadronisation process at once, is extensively studied in the context of this thesis. This alogorithm is briefly discussed in Chapter 14. The impact of this new algorithm and further developments on the statistical precision of the presented analyses is discussed. Finally, an outlook on the precision achievable in the future of these analyses is given.

Part I

Context And Foundations

2 The Standard Model of particle physics - Basic principles

The Standard Model of particle physics (SM) [3–5] is a relativistic renormalisable quantum field theory describing the observed matter and its interactions. Today the SM is well established, making remarkable precise predictions proven by experimental observations, such as the precise measurement of lepton flavour universality [15, 16]. With the discovery of the Higgs boson in 2012 [17, 18] all elementary particles postulated by the SM have been observed.

The large number of 19 free parameters, which have to be experimentally determined, gives reason to look for possible unifications within the SM. Additionally, there are observations of fundamental effects which cannot be described by the current SM. For example, the SM does not include gravitation, the dominant interaction in the macroscopic world. It does not make any predictions on the nature of dark matter and dark energy [6]. Also, the observed oscillations of neutrinos [7] require extensions of the established SM. These shortcomings make particle physics a vivid field of research. There are both direct and indirect searches for physics beyond the SM. Direct searches are focused on the observation of phenomena postulated by SM extensions, while indirect searches are precise analyses looking for small deviations between SM predictions and experimental measurements.

In the following, the SM's features relevant to the presented analyses are discussed. At first, the elementary particles and the fundamental interactions are introduced in Section 2.1. Then, discrete symmetries are discussed in Section 2.2. In Section 2.3 the CKM mechanism is explained, which is of special importance for the physics in neutral B_s^0 mesons reviewed in Chapter 3. This chapter is largely based on more general introductions to the SM [19, 20] and more detailed overviews on flavour physics [8–10].

2.1 Matter and interactions

In the SM, particles are represented by quantum fields $\psi(x)$, while the dynamics are described by a Lagrangian $\mathcal{L}(\psi,\, \partial_\mu\psi)$. All matter is composed of twelve fermions, spin-1/2 particles, or their bound states. Additionally, for every fermion, an antifermion with the same mass and inverted quantum numbers exists. The three fundamental interactions between particles in the SM (electromagnetic, weak and strong) are mediated by four gauge bosons with spin-1. A fifth boson with spin-0, the Higgs boson, represents the coupling to the Higgs field giving mass to all fundamental particles. The elementary particles of the SM and their relations are visualised in Figure 2.1.

The fermions can be divided into six quarks q and six leptons l. Both families consist of three generations each. A lepton generation is composed by a charged lepton l^- (electron e^-, muon μ^- or tau τ^-) and a corresponding neutral, massless[1] neutrino ν_l (electron neutrino ν_e, muon neutrino ν_μ or tau neutrino ν_τ). The lepton generations are characterised by their unique lepton flavour and differ by the lepton masses. The lepton flavour is a conserved quantity in all processes allowed within the SM.

Every quark generation is built by a pair of an up-type quark (up u, charm c or top t) and a down-type quark (down d, strange s or bottom, also known as beauty, b). The up-type quarks have an electric charge of $q = 2/3\,e$, the down-type quarks of $q = -1/3\,e$. Additionally, every quark carries one of three possible colour charges (red r, green g or blue b).

The strong interaction, mediated by the massless gluon g, couples to the colour charge. Every gluon itself carries a set of colour charge and anti-colour charge allowing the gluon to self-interact. Unlike the other interactions in the SM, the strength of the strong interaction increases with the distance between the interacting particles. Consequently, only colour-neutral particles can exist in a free state. As a consequence of this effect, known as colour confinement, all quarks form colour-neutral bound states, the so-called hadrons h. Hence, individual quarks cannot be observed directly. A hadron $q\bar{q}$ formed by a quark and a antiquark is called meson, a hadron qqq ($\overline{qqq}$) built by three (anti)quarks is a baryon. Also, states of more than three quarks are possible like the observed tetraquarks [21] and pentaquarks [22, 23]. Another consequence of the confinement is the finite range of the strong interaction.

There are two manifestations of the weak interaction. Neutral weak currents are mediated by the massive Z^0 boson, charged weak currents by the massive $W^\pm$ boson.

---

[1] Observations of neutrino oscillations [7] prove this assumption to be wrong and show that the current SM requires extensions. However, these effects are negligible in this thesis.

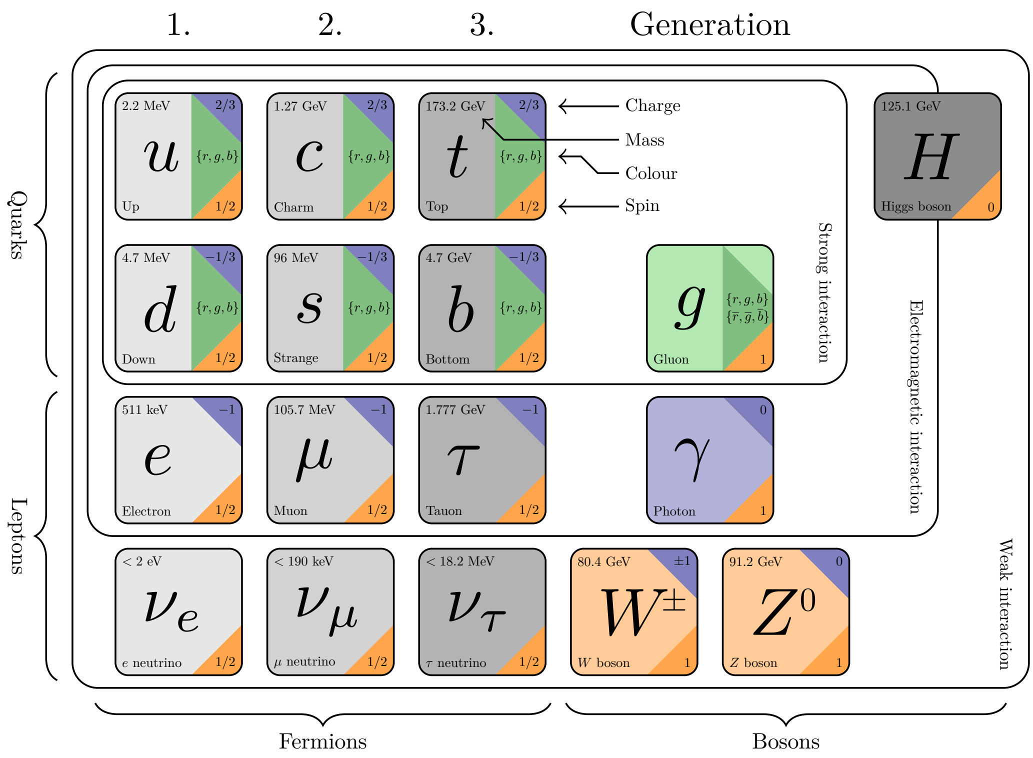


Figure 2.1: Overview of the elementary particles included in the SM and some of their properties.

While the neutral weak current is flavour conserving, the charged manifestation enables transitions between up- and down-type quarks even of different generations, as more detailed discussed in Section 2.3. Both Z^0 and $W^\pm$ bosons only couple to left(right)-handed (anti)particles leading to a maximal violation of the P symmetry, introduced in the following section.

The massless photon γ is the mediator of the electromagnetic interaction. Due to its infinite range and the stronger coupling compared to the weak interaction, the electromagnetic interaction is in the macroscopic world the dominant interaction described by the SM.

In the SM, the electromagnetic and the weak interactions are unified to the electroweak interaction by the Higgs mechanism [24–26]. The Higgs mechanism gives mass to the $W^\pm$ and Z^0 bosons while introducing the Higgs boson H discovered in 2012 [17, 18]. By a Yukawa interaction [27] with the Higgs boson all fermions enquire mass.

2.2 Symmetries of the Standard Model

Conserved quantities are helpful in the description of physical systems. Following Noether's theorem [28], there is a direct relation between conservation laws and symmetries. In the SM, the interactions are consequences of local gauge invariance and the mediators are the generators of the corresponding symmetry groups. For example, the photon field A^μ is the generator of the U(1) symmetry group. This generator is introduced with the covariant derivative

$$\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu \,. \tag{2.1}$$

This covariant derivative is used to make the Lagrangian of the electromagnetic interaction

$$\mathcal{L}_{\text{QED}} = \overline{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi \quad - \quad \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.2}$$

$$= \underbrace{\overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi}_{\text{free fermion field}} \quad - \quad \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{Maxwell dynamics}} \quad - \quad \underbrace{e\overline{\psi}\gamma^\mu \psi A_\mu}_{\text{photon-fermion interaction}}, \tag{2.3}$$

with the electromagnetic field tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, invariant under local gauge transformations using an arbitrary function $\theta(x)$

$$\psi' = \mathrm{e}^{-ie\theta(x)}\,\psi\,, \tag{2.4}$$

$$A'_\mu = A_\mu + \partial_\mu \theta(x) \ . \tag{2.5}$$

Similarly, the gluons are the generators of the SU(3) group of the strong interaction and the Z^0 and $W^\pm$ bosons the generators of the weak interaction's SU(2) symmetry.

In addition to these internal gauge symmetries, there are three discrete symmetries in the SM: The charge conjugation C inverts all internal quantum numbers, transforming a particle to an antiparticle $C\psi(\vec{x},\, t) = \overline{\psi}(\vec{x},\, t)$. The parity P mirrors the spatial coordinates $P\psi(\vec{x},\, t) = \psi(-\vec{x},\, t)$, while the time reversal T mirrors the time coordinate $T\psi(\vec{x},\, t) = \psi(\vec{x},\, -t)$. All of the three symmetries are multiplicative and can be combined.

In the electromagnetic interaction C, P, T and their combinations are strictly conserved. However, these symmetries can be violated by the weak and the strong interaction. While no violation of these symmetries has yet been observed in strong interactions, the individual C and the P symmetries are maximally violated by the weak interaction, which only couples to left(right)-handed (anti)particles. This symmetry violation in the weak interaction was first observed in β decays by Wu in 1956 [29]. As observed in kaon decays in 1964 [30], also the combination *CP* of these symmetries is violated by the weak interaction. In the SM, this *CP* violation is introduced by a complex phase within the CKM matrix, which is explained in Section 2.3.

In contrast to the combinations of two symmetries, the combination *CPT* of all three symmetries has to be conserved in the SM, following the *CPT* theorem [31]. Experimentally no violation of the *CPT* symmetry has been observed. However, this implies a violation of the *T* symmetry, which cannot be accessed experimentally.

2.3 The CKM mechanism

The quark fields in the SM provide mass eigenstates which are different to the flavour eigenstates. The Cabbibo-Kobayashi-Maskawa (CKM) mechanism [32] connects the two eigenstates introducing a basis transformation, based on the so-called CKM matrix V_{CKM}. This basis transformation is manifested in the Yukawa interaction

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}} &= \overline{Q}_L \phi^* Y_U \, U_R + \overline{Q}_L \phi \, Y_D \, D_R + \text{h.c.} \\ &= \overline{Q}_L M_U \, U_R + \overline{Q}_L M_D \, D_R + \text{h.c.}\,, \end{aligned} \tag{2.6}$$

introduced in the electroweak unification. Here, the Yukawa matrices $Y_{L(R)}$ and the Higgs field

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} v + h(x) \\ 0 \end{pmatrix} \tag{2.7}$$

couple to doublets of left-handed quarks

$$Q_L = \left\{ \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix} \right\} \tag{2.8}$$

to right-handed quark singlets

$$U_R = \left\{ u_R, \ \ c_R, \ \ t_R \right\} \quad \text{and} \quad D_R = \left\{ d_R, \ \ s_R, \ \ b_R \right\} . \tag{2.9}$$

The combination $M_{L(R)}$ of Higgs field and Yukawa matrices gives the left(right)-handed quark masses. By diagonalisation of these mass matrices

$$M_Q^{\text{diag}} = T_{Q,L} M_Q \, T_{Q,R}^{\dagger} \tag{2.10}$$

mass eigenstates of the quark fields

$$Q_{L(R)}^m = T_{Q,L(R)} \cdot Q_{L(R)} \tag{2.11}$$

can be constructed by the introduction of unitary matrices $T_{Q,L(R)}$. The eigenstates of the weak neutral current Lagrangian

$$\mathcal{L}_{\mathrm{NC}} \sim \overline{Q}_L \gamma_\mu Z^\mu Q_L = \overline{Q}^m_L \gamma^\mu Z_\mu \underbrace{T^\dagger_{Q,L} T_{Q,L}}_{\mathbb{1}} Q^m_L \tag{2.12}$$

can be identified as the mass eigenstates of the quark fields. However, this is not the case for the weak charged current

$$\begin{aligned} \mathcal{L}_{\mathrm{CC}} &\sim \overline{U}_L \gamma_\mu W^+_\mu D_L + \overline{D}_L \gamma_\mu W^-_\mu U_L \\ &\sim \overline{U}^m_L \gamma^\mu W^+_\mu \underbrace{T_{U,L} T^\dagger_{D,L}}_{V_{\mathrm{CKM}}} D^m_L + \overline{D}^m_L \gamma^\mu W^+_\mu \underbrace{T_{D,L} T^\dagger_{U,L}}_{V^\dagger_{\mathrm{CKM}}} U^m_L \,, \end{aligned} \tag{2.13}$$

where the CKM matrix is introduced by the basis transformation. The transition probability between an up-type quark $i = \{u,\, c,\, t\}$ and a down-type quark $j = \{d,\, s,\, b\}$ via a weak charged current is proportional to the square of the matrix element $|V_{ij}|^2$. Therefore, the finite off-diagonal elements of the CKM matrix enable transitions between different quark generations, which are not possible either in the strong, electromagnetic or neutral weak interaction. Due to its unitarity and the absorption of relative phases within the individual quark fields, the CKM matrix can be described by four independent parameters (three mixing angles and one complex phase)

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{2.14}$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\mathrm{e}^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\mathrm{e}^{i\delta} & c_{12}s_{23} - s_{12}s_{23}s_{13}\mathrm{e}^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\mathrm{e}^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}\mathrm{e}^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{2.15}$$

$$= \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \ . \tag{2.16}$$

The exact parametrisation [33] given in Equation 2.15, with $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$, is known as standard parametrisation and is based on the three Euler angles θ_{ij} and the complex phase δ. The strong hierarchy of the CKM matrix becomes visible in the Wolfenstein parametrisation [34] described by Equation 2.16, with

$$\lambda = s_{12}\,, \quad A = \frac{s_{23}}{s_{12}^2}\,, \quad \rho = \frac{s_{13}}{s_{12}s_{13}}\cos\delta \quad \text{and} \quad \eta = \frac{s_{13}}{s_{12}s_{13}}\sin\delta\,. \tag{2.17}$$

This expansion in terms of $\lambda = 0.226\,50 \pm 0.000\,48$ [35] shows the strong dominance of transitions within a generation compared to flavour-changing transitions.

Nine conditions arise from the unitarity of the CKM matrix:

$$\sum_{k=1}^{3} V_{ki} V_{kj}^* = \delta_{ij}\,, \quad \text{with } i,\, j \in \{1,\, 2,\, 3\}. \tag{2.18}$$

Six of these conditions ($i \neq j$) can be represented by triangles in the complex plane. The six triangles share the same area $J/2$, where J is the Jarlskog invariant

$$J = \left| \mathrm{Im}\left[V_{ij} V_{kl} V_{il}^* V_{kj}^* \right] \right|, \quad \text{with } i \neq k \wedge j \neq l\,. \tag{2.19}$$

which is a measure for the magnitude of *CP* violation allowed in the SM. The current experimentally measured world average is $J = \left(3.00^{+0.15}_{-0.09}\right) \cdot 10^{-5}$ [35]. Most commonly the triangle based on the condition

$$V_{ud} V_{ub}^* \;+\; V_{cd} V_{cb}^* \;+\; V_{td} V_{tb}^* \;=\; 0 \tag{2.20}$$

is studied as it provides sites of similar length in contrast to the remaining, very degenerated triangles. A visualisation of this triangle, including constraints from experimental measurements is given in Figure 2.2. In this representation, all sides are normalised by the most precise known element $V_{cd} V_{cb}^*$. Therefore, the apex is parameterised by $\overline{\rho} + i\overline{\eta} = -V_{ud} V_{ub}^* \,/\, V_{cd} V_{cb}^*$ with the generalised Wolfenstein parameters

$$\overline{\rho} \;=\; \rho \sum_n \frac{(-1)^n \lambda^n}{n!} \qquad \text{and} \qquad \overline{\eta} \;=\; \eta \sum_n \frac{(-1)^n \lambda^n}{n!}\,. \tag{2.21}$$

Substitution of the CKM matrix elements following Equation 2.16 can be used to define the triangle properties

$$\overline{\rho} + i\overline{\eta} \;=\; R_b \mathrm{e}^{i\gamma} \;=\; -\frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*}\,, \tag{2.22}$$

$$(1-\overline{\rho}) - i\overline{\eta} \;=\; R_c \mathrm{e}^{-i\beta} \;=\; -\frac{V_{td} V_{tb}^*}{V_{cd} V_{cb}^*} \tag{2.23}$$

and

$$\mathrm{e}^{i(\pi-\beta-\gamma)} \;=\; \mathrm{e}^{i\alpha} \;=\; -\frac{V_{td} V_{tb}^*}{V_{ud} V_{ub}^*}\,\frac{R_b}{R_c}\,. \tag{2.24}$$

With these relations, the angles

$$\alpha = \arg\left(-\frac{V_{td} V_{tb}^*}{V_{ud} V_{ub}^*} \right), \quad \beta = \arg\left(-\frac{V_{cd} V_{cb}^*}{V_{td} V_{tb}^*} \right), \quad \gamma = \arg\left(-\frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*} \right) \tag{2.25}$$

and the lengths of the triangle edges

$$R_a = \left| \frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*} \right| = 1\,, \;\; R_b = \left| \frac{V_{ud} V_{ub}^*}{V_{cd} V_{cb}^*} \right| = \sqrt{\overline{\rho}^2 + \overline{\eta}^2}\,, \;\; R_c = \left| \frac{V_{td} V_{tb}^*}{V_{cd} V_{cb}^*} \right| = \sqrt{(1-\overline{\rho})^2 + \overline{\eta}^2} \tag{2.26}$$

can be purely expressed in terms of the CKM matrix elements.

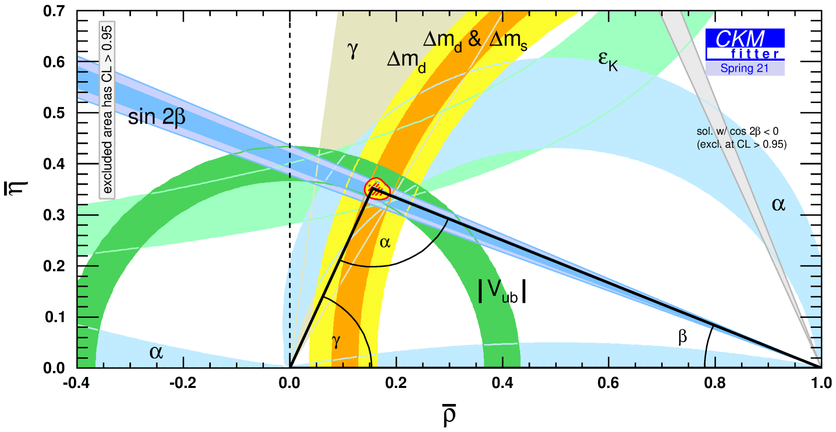


Figure 2.2: The unitarity triangle based on the condition defined by Equation 2.20 and its experimental constraints [36].

The SM can be tested by experimentally constraining the properties of the unitarity triangles. The unitarity of the CKM matrix, an intrinsic property of the SM, requires the exact closing of the triangle in the apex. This hypothesis is tested by combining precise measurements of physics parameters related to the CKM elements. These combinations, such as the ones presented in Figure 2.2 and Figure 2.3 are performed by the CKMfitter group [36]. At the current level of precision, the hypothesis of unitarity is well supported.

Since the unitarity triangle described in Equation 2.20 only depends on six of the nine CKM matrix elements directly, it is beneficial to also study a second triangle related to the missing three elements. The common choice is based on the condition

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0\,. \tag{2.27}$$

This triangle, shown in Figure 2.3, shows a manageable level of degeneration and introduces an additional angle

$$\beta_s = \arg\left(-\frac{V_{cs}V_{cb}^*}{V_{ts}V_{tb}^*}\right)\,. \tag{2.28}$$

The fascinating consequences of possible transitions between different quark generations introduced by the CKM mechanism are discussed in the following chapter.

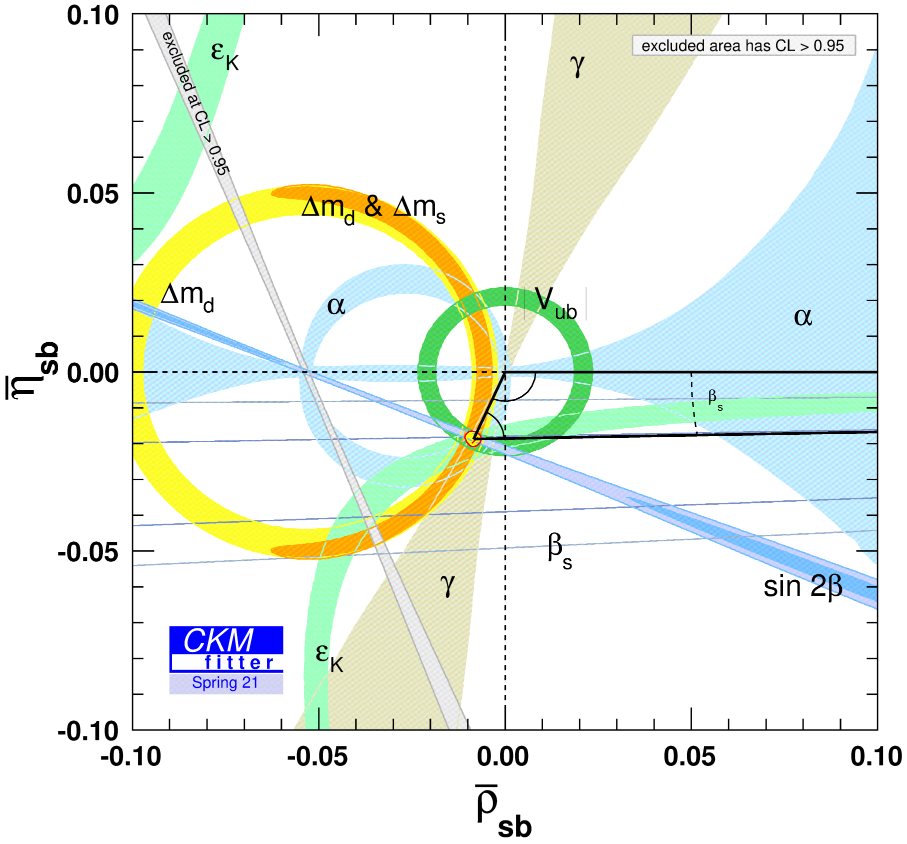


Figure 2.3: The unitarity triangle based on the condition defined by Equation 2.27 and its experimental constraints [36].

3 The B_s^0-meson system

The quark mixing introduced by the CKM mechanism discussed in Section 2.3 has consequences, which are of special interest in this thesis. A great laboratory to observe these phenomena, such as the oscillation of neutral mesons explained in Section 3.1 or *CP* violation discussed in Section 3.2, is the system of neutral B_s^0 mesons. Therefore, the tree-level $B_s^0 \to D_s^\mp h^\pm$ decays are introduced in Section 3.3, which can be used to extract properties of the SM, like the oscillation frequency Δm_s or the CKM angle γ, from experimental measurements of these phenomena. The following considerations are largely based on more detailed discussions on heavy flavour physics [8–11].

3.1 Oscillation of neutral mesons

The oscillation of neutral mesons, also referred to as neutral meson mixing, is the transition between a particle and its antiparticle state. This process is enabled by the quark mixing described in Section 2.3, which allows flavour-changing currents in the weak interaction. At the lowest order, the transition between meson and anti-meson can be described by two so-called box diagrams as displayed in Figure 3.1.

Historically, these transitions have first been observed in the system of neutral kaons [30]. However, this transition is possible (at different rates) for every neutral meson family (K^0 , D^0 , B^0 & B_s^0) independent of the exact quark content. In the following, the mechanism is explained in the B_s^0 $(\bar{b}, s)$ system, which is of special interest in this thesis.

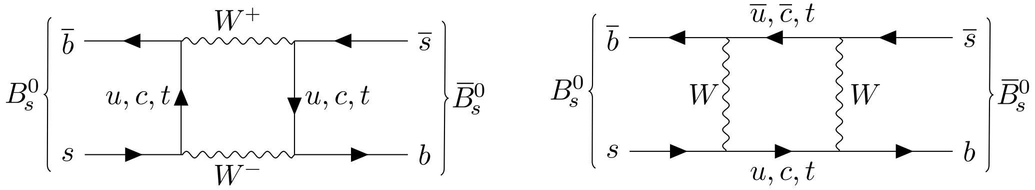


Figure 3.1: Feynman diagrams of the B_s^0-$\bar{B}_s^0$ oscillation process at the lowest-order possible in the SM.

The time evolution of this two-state system is described by an effective Schrödinger equation in the Wigner-Weisskopf formalism [37]

$$i\frac{d}{dt}\begin{pmatrix}|B_s^0\rangle\\|\bar{B}_s^0\rangle\end{pmatrix} = \boldsymbol{H}\begin{pmatrix}|B_s^0\rangle\\|\bar{B}_s^0\rangle\end{pmatrix} = \left(\boldsymbol{M}-\frac{i}{2}\boldsymbol{\Gamma}\right)\begin{pmatrix}|B_s^0\rangle\\|\bar{B}_s^0\rangle\end{pmatrix}, \tag{3.1}$$

with a Hamiltonian $\boldsymbol{H}$ composed of the hermitian mass $\boldsymbol{M}$ and decay-time width $\boldsymbol{\Gamma}$ matrices. The B_s^0-$\bar{B}_s^0$ oscillation is described by the off-diagonal elements of the Hamiltonian. The *CPT* invariance of the SM requires both, particle and antiparticle, to have the same masses $m_s = m_{11} = m_{22}$ and decay widths $\Gamma_s = \Gamma_{11} = \Gamma_{22}$ yielding an Hamiltonian

$$\boldsymbol{H} = \begin{pmatrix} m_s - \frac{i}{2}\Gamma_s & m_{12} - \frac{i}{2}\Gamma_{12} \\ m_{12}^* - \frac{i}{2}\Gamma_{12}^* & m_s - \frac{i}{2}\Gamma_s \end{pmatrix}. \tag{3.2}$$

The time evolution of the system

$$\left|B_{\mathrm{H/L}}(t)\right\rangle = \mathrm{e}^{-i\mu_{\mathrm{H/L}}t}\left|B_{\mathrm{H/L}}(0)\right\rangle \tag{3.3}$$

is trivial in the basis of a diagonalised Hamiltonian

$$i\frac{d}{dt}\begin{pmatrix}|B_{\mathrm{H}}\rangle\\|B_{\mathrm{L}}\rangle\end{pmatrix} = \begin{pmatrix}\mu_{\mathrm{H}} & 0\\ 0 & \mu_{\mathrm{L}}\end{pmatrix}\begin{pmatrix}|B_{\mathrm{H}}\rangle\\|B_{\mathrm{L}}\rangle\end{pmatrix} \tag{3.4}$$

with eigenvalues $\mu_{\mathrm{H/L}}$ and eigenvectors

$$|B_{\mathrm{H}}\rangle = p\left|B_s^0\right\rangle + q\left|\bar{B}_s^0\right\rangle, \quad |B_{\mathrm{L}}\rangle = p\left|B_s^0\right\rangle - q\left|\bar{B}_s^0\right\rangle. \tag{3.5}$$

While The eigenvectors, superpositions of the flavour eigenstates, can be identified as the heavier B_{H} and lighter B_{L} mass eigenstates of the system, the eigenvalues of the Hamiltonian

$$\mu_{\mathrm{H/L}} = m_{\mathrm{H/L}} - \frac{i}{2}\Gamma_{\mathrm{H/L}} \tag{3.6}$$

are composed of the corresponding masses

$$m_{\mathrm{H/L}} = m_s \pm \ \mathrm{Re}\left[\sqrt{\left(m_{12} - \frac{i}{2}\mathit{\Gamma}_{12}\right) \cdot \left(m_{12}^* - \frac{i}{2}\mathit{\Gamma}_{12}^*\right)}\right] \tag{3.7}$$

and decay widths

$$\mathit{\Gamma}_{\mathrm{H/L}} = \mathit{\Gamma}_s \mp 2\,\mathrm{Im}\left[\sqrt{\left(m_{12} - \frac{i}{2}\mathit{\Gamma}_{12}\right) \cdot \left(m_{12}^* - \frac{i}{2}\mathit{\Gamma}_{12}^*\right)}\right] . \tag{3.8}$$

Due to the normalisation the complex coefficients p and q are constrained to $|p|^2 + |q|^2 = 1$. Their ratio is given by

$$\frac{q}{p} = \sqrt{\frac{2m_{12}^* - i\mathit{\Gamma}_{12}^*}{2m_{12} - i\mathit{\Gamma}_{12}}} = \frac{\Delta m_s - \frac{i}{2}\Delta\mathit{\Gamma}_s}{2m_{12} - i\mathit{\Gamma}_{12}} , \tag{3.9}$$

with the mass and decay-width differences

$$\Delta m_s = m_{\mathrm{H}} - m_{\mathrm{L}} \qquad \text{and} \qquad \Delta\mathit{\Gamma}_s = \mathit{\Gamma}_{\mathrm{H}} - \mathit{\Gamma}_{\mathrm{L}} . \tag{3.10}$$

Analogue to the differences the mean values are defined as

$$m_s = \frac{m_{\mathrm{H}} + m_{\mathrm{L}}}{2} \qquad \text{and} \qquad \mathit{\Gamma}_s = \frac{\mathit{\Gamma}_{\mathrm{H}} + \mathit{\Gamma}_{\mathrm{L}}}{2} . \tag{3.11}$$

A basis transformation gives the time evolution of the flavour eigenstates

$$\begin{aligned} \left|B_s^0(t)\right\rangle &= g_+(t)\left|B_s^0(0)\right\rangle + \frac{q}{p}g_-(t)\left|\overline{B}_s^0(0)\right\rangle , \\ \left|\overline{B}_s^0(t)\right\rangle &= g_+(t)\left|\overline{B}_s^0(0)\right\rangle + \frac{p}{q}g_-(t)\left|B_s^0(0)\right\rangle , \end{aligned} \tag{3.12}$$

with the time-dependent coefficients

$$g_\pm(t) = \frac{1}{2}\left(\mathrm{e}^{-i\mu_{\mathrm{H}}t} \pm \mathrm{e}^{-i\mu_{\mathrm{L}}t}\right) . \tag{3.13}$$

The eigenstate of the B_s^0 meson is experimentally not directly accessible. However, decays into a final state f or the *CP* conjugated final state $\overline{f}$ give access to the time evolution of the B_s^0 meson. Four decay amplitudes

$$\begin{aligned} A_f &= \langle f|\,\boldsymbol{T}\,|B_s^0\rangle , \quad \overline{A}_f = \langle f|\,\boldsymbol{T}\,|\overline{B}_s^0\rangle , \\ A_{\overline{f}} &= \langle \overline{f}|\,\boldsymbol{T}\,|B_s^0\rangle , \quad \overline{A}_{\overline{f}} = \langle \overline{f}|\,\boldsymbol{T}\,|\overline{B}_s^0\rangle \end{aligned} \tag{3.14}$$

can be defined with a Hamiltonian matrix $\boldsymbol{T}$ describing the weak decay. The time-dependent decay rate of a B_s^0 meson decaying into the final state f

$$\begin{aligned}\Gamma(B_s^0(t) \to f) &= \left|\langle f|\,\boldsymbol{T}\,|B_s^0(t)\rangle\right|^2 \\ &= \left|\langle f|\,\boldsymbol{T}\,|B_s^0\rangle\,\langle B_s^0|B_s^0(t)\rangle \;+\; \langle f|\,\boldsymbol{T}\,|\overline{B}{}_s^0\rangle\,\langle \overline{B}{}_s^0|B_s^0(t)\rangle\right|^2 \\ &= \left|A_f\;\, g_+(t) \;+\; \overline{A}_f\,\frac{q}{p}\,g_-(t)\;\right|^2 \\ &= \left|A_f\left(g_+(t) \;+\; \lambda_f\, g_-(t)\right)\right|^2 \end{aligned} \tag{3.15}$$

composed of the two possible decay paths with oscillation $B_s^0(t) \to B_s^0 \to f$ and without $B_s^0(t) \to \overline{B}{}_s^0 \to f$ gives access to the time evolution of the $B_s^0(t)$. With the substitutions

$$\lambda_f \;=\; \frac{q}{p}\frac{\overline{A}_f}{A_f} \quad \text{and} \quad \lambda_{\bar f} \;=\; \frac{p}{q}\frac{\overline{A}_{\bar f}}{A_{\bar f}} \tag{3.16}$$

the four possible decay rates are defined as

$$\begin{aligned}\Gamma(B_s^0(t) \to f) \;&=\; |A_f|^2\Big[g_+(t)\,g_+^*(t) + |\lambda_f|^2 g_-(t)\,g_-^*(t) + \Big(\lambda_f^*\,g_-(t)\,g_+^*(t) + \lambda_f\,g_-^*(t)\,g_+(t)\Big)\Big] \\ &=\; \frac{1}{2}\mathrm{e}^{-\Gamma_s t}\,|A_f|^2\Big(1+|\lambda_f|^2\Big)\;\Bigg[\;\cosh\frac{\Delta\Gamma_s}{2}\,t \,+\, C_f\,\cos\Delta m_s t \\ &\qquad + D_f\,\sinh\frac{\Delta\Gamma_s}{2}\,t \,-\, S_f\,\sin\Delta m_s t\;\Bigg],\end{aligned} \tag{3.17}$$

$$\begin{aligned}\Gamma(\overline{B}{}_s^0(t) \to f) \;&=\; |A_f|^2\Big[|\lambda_f|^2 g_+(t)\,g_+^*(t) + g_-(t)\,g_-^*(t) + \Big(\lambda_f^*\,g_-^*(t)\,g_+(t) + \lambda_f\,g_-(t)\,g_+^*(t)\Big)\Big] \\ &=\; \frac{1}{2}\mathrm{e}^{-\Gamma_s t}\,|A_f|^2\Big(1+|\lambda_f|^2\Big)\left|\frac{p}{q}\right|^2\Bigg[\;\cosh\frac{\Delta\Gamma_s}{2}\,t \,-\, C_f\,\cos\Delta m_s t \\ &\qquad + D_f\,\sinh\frac{\Delta\Gamma_s}{2}\,t \,+\, S_f\,\sin\Delta m_s t\;\Bigg],\end{aligned} \tag{3.18}$$

$$\begin{aligned}\Gamma(B_s^0(t) \to \bar f) \;&=\; |A_{\bar f}|^2\Big[g_+(t)\,g_+^*(t) + |\lambda_{\bar f}|^2 g_-(t)\,g_-^*(t) + \Big(\lambda_{\bar f}^*\,g_-(t)\,g_+^*(t) + \lambda_{\bar f}\,g_-^*(t)\,g_+(t)\Big)\Big] \\ &=\; \frac{1}{2}\mathrm{e}^{-\Gamma_s t}\,|A_{\bar f}|^2\Big(1+|\lambda_{\bar f}|^2\Big)\;\Bigg[\;\cosh\frac{\Delta\Gamma_s}{2}\,t \,+\, C_{\bar f}\,\cos\Delta m_s t \\ &\qquad + D_{\bar f}\,\sinh\frac{\Delta\Gamma_s}{2}\,t \,-\, S_{\bar f}\,\sin\Delta m_s t\;\Bigg],\end{aligned} \tag{3.19}$$

$$\begin{aligned}\Gamma(\overline{B}{}_s^0(t) \to \bar f) \;&=\; |A_{\bar f}|^2\Big[|\lambda_{\bar f}|^2 g_+(t)\,g_+^*(t) + g_-(t)\,g_-^*(t) + \Big(\lambda_{\bar f}^*\,g_-^*(t)\,g_+(t) + \lambda_{\bar f}\,g_-(t)\,g_+^*(t)\Big)\Big] \\ &=\; \frac{1}{2}\mathrm{e}^{-\Gamma_s t}\,|A_{\bar f}|^2\Big(1+|\lambda_{\bar f}|^2\Big)\left|\frac{q}{p}\right|^2\Bigg[\;\cosh\frac{\Delta\Gamma_s}{2}\,t \,-\, C_{\bar f}\,\cos\Delta m_s t \\ &\qquad + D_{\bar f}\,\sinh\frac{\Delta\Gamma_s}{2}\,t \,+\, S_{\bar f}\,\sin\Delta m_s t\;\Bigg].\end{aligned} \tag{3.20}$$

The time dependence of the decay rates, an exponential decay enveloping hyperbolic and trigonometric contributions, is revealed by the substitutions

$$\begin{aligned} g_\pm(t)\, g_\mp^*(t) &= \frac{1}{2}\mathrm{e}^{-\Gamma_s t}\left(\cosh\frac{\Delta\Gamma_s}{2}t \pm \cos\Delta m_s t\right), \\ g_\pm^*(t)\, g_\mp(t) &= \frac{1}{2}\mathrm{e}^{-\Gamma_s t}\left(\sinh\frac{\Delta\Gamma_s}{2}t \pm i\sin\Delta m_s t\right). \end{aligned} \tag{3.21}$$

Together with these substitutions, six *CP* parameters are introduced

$$C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \qquad D_f = \frac{2\,\mathrm{Re}[\lambda_f]}{1+|\lambda_f|^2}, \qquad S_f = \frac{2\,\mathrm{Im}[\lambda_f]}{1+|\lambda_f|^2}, \tag{3.22}$$

$$C_{\bar f} = \frac{1-|\lambda_{\bar f}|^2}{1+|\lambda_{\bar f}|^2}, \qquad D_{\bar f} = \frac{2\,\mathrm{Re}[\lambda_{\bar f}]}{1+|\lambda_{\bar f}|^2}, \qquad S_{\bar f} = \frac{2\,\mathrm{Im}[\lambda_{\bar f}]}{1+|\lambda_{\bar f}|^2}. \tag{3.23}$$

By definition, these parameters are normalised by

$$C_f^2 + D_f^2 + S_f^2 = 1 \tag{3.24}$$

and

$$C_{\bar f}^2 + D_{\bar f}^2 + S_{\bar f}^2 = 1. \tag{3.25}$$

These parameters, which are a measure for *CP* violation in the B_s^0 system, can be experimentally accessed by measuring the time-dependent decay rates.

3.2 *CP* violation in neutral meson decays

As previously described the charged weak interaction allows a violation of the *CP* symmetry in the SM. The complex parameters λ_f and $\lambda_{\bar f}$ digest the CKM matrix elements present in the decay to a final state $f(\bar f)$ and with that also the complex phase enabling *CP* violation. Experimentally, these parameters can be constrained via the six *CP* parameters from Equation 3.22 by measuring the time-dependent decay rates. In general

$$\lambda_f = \frac{1}{\lambda_{\bar f}} \tag{3.26}$$

implies *CP* conservation. There are three manifestations of *CP* violation, which will be discussed in the following, cancelling this equality.

3.2.1 Direct *CP* violation in the decay

Differences in the rates of a $B_s^0 \to f$ decay and its *CP* conjugated decay $\overline{B}_s^0 \to \bar{f}$ are considered direct *CP* violation. Such a difference introduces an asymmetry

$$A_{CP} = \frac{\Gamma(B_s^0 \to f) - \Gamma(\overline{B}_s^0 \to \bar{f})}{\Gamma(B_s^0 \to f) + \Gamma(\overline{B}_s^0 \to \bar{f})} = \frac{\left|\overline{A}_{\bar{f}}/A_f\right|^2 - 1}{\left|\overline{A}_{\bar{f}}/A_f\right|^2 + 1} \,. \tag{3.27}$$

It can easily be seen, that the presence of direct *CP* violation is tied to the condition

$$\left|\overline{A}_{\bar{f}}\right|^2 \neq \left|A_f\right|^2 \,. \tag{3.28}$$

To fulfil this condition the two decays must be realised by $n \geq 2$ different processes with individual amplitudes A_i and phases

$$A_f = \sum_i^n A_i \mathrm{e}^{i(\delta_i + \phi_i)} \,, \qquad \overline{A}_{\bar{f}} = \sum_i^n A_i \mathrm{e}^{i(\delta_i - \phi_i)} \,. \tag{3.29}$$

Here, the strong phases δ originate from the possible on-shell scattering of final-state hadrons and are *CP* invariant. The weak phases ϕ originate from the complex CKM elements and change their sign under *CP* conjugation.

Direct *CP* violation is the only manifestation that is independent of the mixing process and can therefore also occur in baryons and charged mesons. A recent measurement of direct *CP* violation is performed by the LHCb experiment using decays of charged $B^\pm$ mesons into three light, charged mesons [38, 39]. As illustrated in Figure 3.2 asymmetries up to

$$A_{CP}^{\mathrm{local}}(B^\pm \to \pi^\pm \pi^+ \pi^-) = -0.745 \pm 0.027\,(\mathrm{stat.}) \pm 0.018\,(\mathrm{sys.}) \tag{3.30}$$

are observed in local regions of the phase space. While this is the largest amount of *CP* violation yet observed, it can not easily be interpreted. Strong final-state interactions and the large number of resonances in the rescattering of the hadronic final state introduce a large variation of amplitudes spread over the phase space [11, 40]. This makes it often very challenging to disentangle the strong and the weak phases.

3.2.2 Mixing induced *CP* violation

The manifestation of *CP* violation in the mixing corresponds to different transition probabilities for the mixing of a particle into its antiparticle and vice versa.

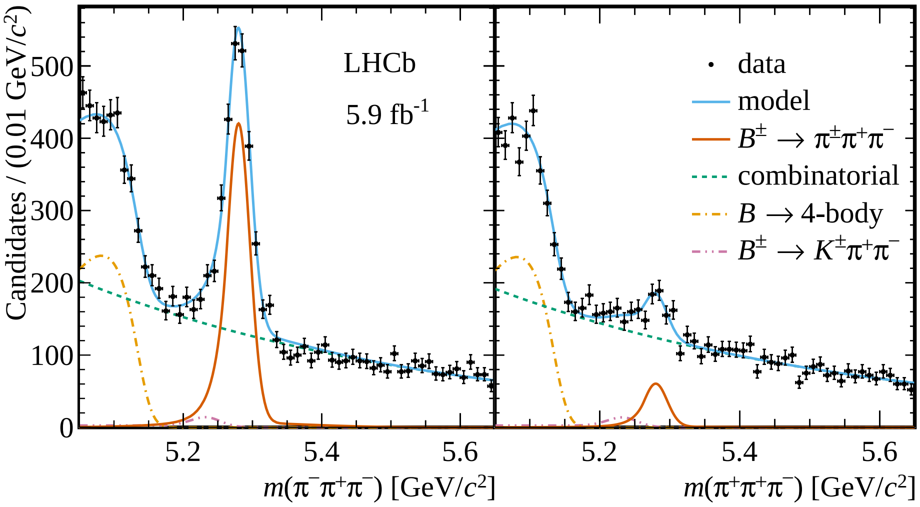


Figure 3.2: Fitted mass distribution of reconstructed B^- (left) and B^+ (right) candidates [38, 39]. The effect of direct *CP* violation manifests in the differently sized signal components (red line) for the two *CP* conjugated decays.

Therefore, it can only occur in neutral mesons. Following the decay rates defined in Equation 3.17 to 3.20, the transition probabilities can be expressed by

$$\left|\langle B_s^0|B_s^0(t)\rangle\right|^2 = \frac{1}{4}\mathrm{e}^{-\Gamma_s t}\left|\frac{q}{p}\right|^2\left(\mathrm{e}^{-\Gamma_\mathrm{H} t} + \mathrm{e}^{-\Gamma_\mathrm{L} t} + 2\mathrm{e}^{-\Gamma_s t}\cos\Delta m_s t\right)\,, \tag{3.31}$$

$$\left|\langle \overline{B}_s^0|B_s^0(t)\rangle\right|^2 = \frac{1}{4}\mathrm{e}^{-\Gamma_s t}\left|\frac{q}{p}\right|^2\left(\mathrm{e}^{-\Gamma_\mathrm{H} t} + \mathrm{e}^{-\Gamma_\mathrm{L} t} - 2\mathrm{e}^{-\Gamma_s t}\cos\Delta m_s t\right)\,, \tag{3.32}$$

$$\left|\langle B_s^0|\overline{B}_s^0(t)\rangle\right|^2 = \frac{1}{4}\mathrm{e}^{-\Gamma_s t}\left|\frac{p}{q}\right|^2\left(\mathrm{e}^{-\Gamma_\mathrm{H} t} + \mathrm{e}^{-\Gamma_\mathrm{L} t} - 2\mathrm{e}^{-\Gamma_s t}\cos\Delta m_s t\right)\,, \tag{3.33}$$

$$\left|\langle \overline{B}_s^0|\overline{B}_s^0(t)\rangle\right|^2 = \frac{1}{4}\mathrm{e}^{-\Gamma_s t}\left|\frac{p}{q}\right|^2\left(\mathrm{e}^{-\Gamma_\mathrm{H} t} + \mathrm{e}^{-\Gamma_\mathrm{L} t} + 2\mathrm{e}^{-\Gamma_s t}\cos\Delta m_s t\right)\,. \tag{3.34}$$

Here, *CP* violation in the mixing can easily be coupled to the condition

$$1 \neq \left|\frac{p}{q}\right|^2 \underbrace{=}_{3.9} \frac{1 - i\left|\frac{\Gamma_{12}}{2m_{12}}\right|\mathrm{e}^{-i\Delta\phi}}{1 - i\left|\frac{\Gamma_{12}}{2m_{12}}\right|\mathrm{e}^{+i\Delta\phi}}\,. \tag{3.35}$$

Following Equation 3.9, this condition can be transferred into the presence of a complex phase difference $\Delta\phi = \arg\left(m_{12}/\Gamma_{12}\right)$ in the off-diagonal Hamiltonian elements with a significant amplitude $|m_{12}/\Gamma_{12}|$. This manifestation of *CP* violation can be experimentally accessed by another asymmetry

$$A_{CP}^{\mathrm{mix}} = \frac{\mathit{\Gamma}(B_s^0(t) \to B_s^0) - \mathit{\Gamma}\big(\overline{B}{}_s^0(t) \to \overline{B}{}_s^0\big)}{\mathit{\Gamma}(B_s^0 \to B_s^0) + \mathit{\Gamma}\big(\overline{B}{}_s^0 \to \overline{B}{}_s^0\big)} = \frac{1 - |q/p|^4}{1 + |q/p|^4} = A_{\mathrm{sl}}^s \,. \tag{3.36}$$

This asymmetry is also referred to as semileptonic asymmetry, as it is usually measured in flavour-specific, semileptonic $B_s^0 \to l^- \nu X^+$ decays. While *CP* violation was first observed in the mixing of neutral kaons [30], the asymmetry in the B_s^0 system is predicted to be small $\mathcal{O}(10^{-5})$ [41]. The small size of this effect is reasoned by the ratio

$$\left|\frac{\mathit{\Gamma}_{12}}{m_{12}}\right| \approx \frac{m_b}{m_t} \approx \mathcal{O}(10^{-3}) \tag{3.37}$$

of the decay-matrix element, dominated by the mass available in the decay, and the mixing matrix element, with dominant contributions from the box diagrams involving t-quarks. At the moment the measured world average $A_{\mathrm{sl}}^s = -0.0006 \pm 0.0028$ [42] is in agreement with both the theoretical prediction and the absence of *CP* violation in the B_s^0 mixing. However, in New Physics scenarios sizeable *CP* violation in the B_s^0 mixing could be enabled by interactions or particles not yet included in the SM. The size of such effects can be constrained by measurements of the CKM angle β_s and the oscillation frequency Δm_s, which limits the magnitude of the corresponding *CP* violation to $\mathcal{O}\left(10^{-4}\right)$ [43]. Hence, such effects are currently (and will be for several years) well below the experimental sensitivity. This allows *CP* violation in the B_s^0 mixing to be reasonably neglected in the following considerations.

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

Even in the absence of both *CP* violation in the decay and in the mixing, which directly leads to $\left|\lambda_f\right| = 1/\left|\lambda_{\bar{f}}\right|$, *CP* violation can occur if

$$\arg \lambda_f + \lambda_{\bar{f}} \neq 0 \,. \tag{3.38}$$

Such an inequality can be introduced in the interference of the direct decay $B_s^0 \to f$ and the decay after mixing $B_s^0 \to \overline{B}{}_s^0 \to f$. This requires a common final state f shared among both B_s^0 and $\overline{B}{}_s^0$. In this case two time-dependent asymmetries

$$\begin{aligned} A_{CP}^{f(\bar{f})} &= \frac{\mathit{\Gamma}\big(\overline{B}{}_s^0(t) \to f\,(\bar{f})\big) - \mathit{\Gamma}\big(B_s^0(t) \to f\,(\bar{f})\big)}{\mathit{\Gamma}\big(\overline{B}{}_s^0(t) \to f\,(\bar{f})\big) + \mathit{\Gamma}\big(B_s^0(t) \to f\,(\bar{f})\big)} \\ &= \frac{-C_{f(\bar{f})} \cos\left(\Delta m_s\, t\right) + S_{f(\bar{f})} \sin\left(\Delta m_s\, t\right)}{\cosh\left(\Delta \mathit{\Gamma}_s\, t\,/\,2\right) + D_{f(\bar{f})} \sinh\left(\Delta \mathit{\Gamma}_s\, t\,/\,2\right)} \end{aligned} \tag{3.39}$$

can be defined. This manifestation of *CP* violation is of special interest in the later discussed $B_s^0 \to D_s^\mp K^\pm$ decays, which show negligible effects from both *CP* violation in the mixing and in the direct decay.

3.3 $B_s^0 \to D_s^- h^+$ decays

For this thesis, the $B_s^0 \to D_s^\mp h^\pm$ decays are of special interest, where the companion hadron $h^\pm \in \{\pi^+, K^\pm\}$ is either a pion or kaon. These decays are part of the family of weak non-leptonic colour-allowed tree-level decays. Weak tree-level decays give clean access to the CKM mechanism and allow the properties of the CKM matrix and the unitarity triangles to be measured. While *CP* violation in the mixing is negligible in the B_s^0 system in general, effects from direct *CP* violation are negligible in tree-level decays, too. In the flavour-specific case of $B_s^0 \to D_s^- \pi^+$, this gives clean access to the oscillation process as discussed in Section 3.3.1. In $B_s^0 \to D_s^\mp K^\pm$, where the final states are common to both the B_s^0 and $\overline{B}_s^0$ decays, significant effects from *CP* violation in the interference can be measured giving access to the CKM angle γ as discussed in Section 3.3.2. Besides, these decays benefit from the experimental perspective from a high abundance, charged final-state particles and the absence of neutrinos.

3.3.1 $B_s^0 \to D_s^- \pi^+$ and the oscillation frequency Δm_s

Due to the realisation via a tree-level process and due to the small[1] Cabibbo suppression factor $|V_{cb}^* V_{ud}| \approx \lambda^2$ the $B_s^0 \to D_s^- \pi^+$ decay mode is very abundant. Currently, the relative branching ratio [35] is measured to be

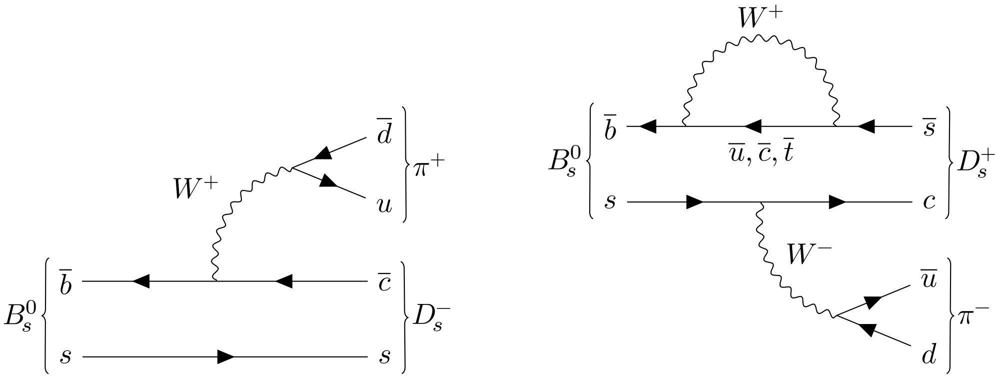
$$\mathcal{B}\left(B_s^0 \to D_s^- \pi^+\right) = 2.98 \pm 0.14 \cdot 10^{-3} \,. \tag{3.40}$$

At tree-level B_s^0 mesons can only decay to the $D_s^- \pi^+$ final state but not to its charge conjugated state. To reach the $D_s^+ \pi^-$ state without previous oscillation a weak loop has to be introduced. As a consequence of the massive vector boson and its couplings, the corresponding loop process is suppressed by an additional factor $G_F^2 \sim 0.1\ \text{MeV}^4$ compared to the tree process. Hence, contributions of $B_s^0 \to D_s^+ \pi^-$ can be neglected yielding a flavour-specific $B_s^0 \to D_s^- \pi^+$ decay. The Feynman diagrams for both decays are presented in Figure 3.3.

The absence of the loop-suppressed decays leads to negligible decay amplitudes $\left|A_{\bar{f}}\right| = \left|\overline{A}_f\right| \approx 0$. As a consequence the *CP* parameters can be simplified to

---

[1] Weak B-meson decays are at least suppressed by a factor λ^2 as there does not exist any lighter bound state of a third generation quark.



(a) $B_s^0 \rightarrow D_s^- \pi^+$ **(b)** $B_s^0 \rightarrow D_s^+ \pi^-$

Figure 3.3: Leading-order Feynman diagrams for the dominant, tree-level $B_s^0 \rightarrow D_s^- \pi^+$ decay (a) and the loop-suppressed $B_s^0 \rightarrow D_s^+ \pi^-$ decay (b).

$D_f = D_{\overline{f}} = S_f = S_{\overline{f}} = 0$ and $C_f = -C_{\overline{f}} = 1$. Together with the negligible effects from *CP* violation in the B_s^0 mixing $|p/q| \approx 1$ and from *CP* violation in the direct decay $\left|A_f\right| = \left|\overline{A}_{\overline{f}}\right| = A_{D_s\pi}$, this allows the time-dependent decay-rates in Equation 3.17 to 3.20 for $B_s^0 \rightarrow D_s^- \pi^+$ decays to be reduced to

$$\varGamma\left(B_s^0(t) \rightarrow D_s^- \pi^+\right) = \frac{1}{2} \mathrm{e}^{-\varGamma_s t} \left|A_{D_s\pi}\right|^2 \left(\cosh \frac{\Delta\varGamma_s}{2} t + \cos \Delta m_s t\right) , \tag{3.41}$$

$$\varGamma\left(\overline{B}_s^0(t) \rightarrow D_s^- \pi^+\right) = \frac{1}{2} \mathrm{e}^{-\varGamma_s t} \left|A_{D_s\pi}\right|^2 \left(\cosh \frac{\Delta\varGamma_s}{2} t - \cos \Delta m_s t\right) , \tag{3.42}$$

$$\varGamma\left(B_s^0(t) \rightarrow D_s^+ \pi^-\right) = \frac{1}{2} \mathrm{e}^{-\varGamma_s t} \left|A_{D_s\pi}\right|^2 \left(\cosh \frac{\Delta\varGamma_s}{2} t - \cos \Delta m_s t\right) , \tag{3.43}$$

$$\varGamma\left(\overline{B}_s^0(t) \rightarrow D_s^+ \pi^-\right) = \frac{1}{2} \mathrm{e}^{-\varGamma_s t} \left|A_{D_s\pi}\right|^2 \left(\cosh \frac{\Delta\varGamma_s}{2} t + \cos \Delta m_s t\right) . \tag{3.44}$$

These four decay rates can be summarised pairwise by the decay rates of unmixed

$$\varGamma_u(t) = \varGamma\left(B_s^0(t) \rightarrow D_s^- \pi^+\right) + \varGamma\left(\overline{B}_s^0(t) \rightarrow D_s^+ \pi^-\right) \tag{3.45}$$

and mixed

$$\varGamma_m(t) = \varGamma\left(B_s^0(t) \rightarrow D_s^+ \pi^-\right) + \varGamma\left(\overline{B}_s^0(t) \rightarrow D_s^- \pi^+\right) \tag{3.46}$$

decays allowing a single decay-time-dependent asymmetry to be defined by

$$A_{\text{mix}} = \frac{\varGamma_u(t) - \varGamma_m(t)}{\varGamma_u(t) + \varGamma_m(t)} = \frac{\cos \Delta m_s\, t}{\cosh \frac{\Delta\varGamma_s}{2} t} . \tag{3.47}$$

A measurement of this asymmetry or the decay-time-dependent rates gives clean access to the B_s^0-oscillation frequency Δm_s.

Implications of the oscillation frequency Δm_s

The oscillation frequency $\Delta m_s = 2|m_{12}|$ is directly related to the off-diagonal mixing-matrix element. In this way, the oscillation frequency can be expressed in terms of the SM parameters such as CKM matrix elements. The mixing matrix $\boldsymbol{M}$ summarises the off-shell transitions within the B_s^0 meson dominated by the box diagrams sketched in Figure 3.1. The effective Hamiltonian for these diagrams can be expressed by $\boldsymbol{H}_{\text{eff}}^{\text{box}}$ as for example derived in [10]. The matrix element is then given by

$$|m_{12}| = \left|\langle B_s^0 \right| \boldsymbol{H}_{\text{eff}}^{\text{box}} \left| \overline{B}_s^0 \rangle\right| \approx \frac{G_F^2 m_W^2}{12\pi^2} m_{B_s^0} f_{B_s^0}^2 B_{B_s^0} \eta_B S_0\left(\frac{m_t^2}{m_W^2}\right) \left(V_{tb}^* V_{ts}\right)^2 . \quad (3.48)$$

Here, perturbative QCD corrections are described by the parameter η_B, while the bag parameter $B_{B_s^0}$ and the decay constant $f_{B_s^0}$ are summarising non-perturbative QCD corrections. The Inami-Lim functions [44] are denoted by S_0, the Fermi constant by G_F. Further, the masses of the W boson m_W, of the B_s^0 meson $m_{B_s^0}$ and of the top quark m_t are involved. Due to the GIM mechanism [45] only the CKM matrix elements V_{tb} and V_{ts} are relevant, while other elements are suppressed by the small u-quark and c-quark masses. This expression allows the SM value of the oscillation frequency [41]

$$\Delta m_s^{\text{theo}} = 18.77 \pm 0.86\,\text{ps}^{-1} , \quad (3.49)$$

which is in agreement with previous observations as well as with the later in Part II presented measurement, to be predicted. Due to the cancellation of theoretical inputs in the ratio

$$\frac{\Delta m_d}{\Delta m_s} = \frac{m_{B^0}}{m_{B_s^0}} \frac{f_{B^0}^2}{f_{B_s^0}^2} \frac{B_{B^0}}{B_{B_s^0}} \left|\frac{V_{td}^2}{V_{ts}^2}\right| \quad (3.50)$$

a more precise prediction [46] is possible, which again is in agreement with the ratio of the experimental observations [42]

$$\left(\frac{\Delta m_d}{\Delta m_s}\right)_{\text{theo}} = 0.0298\,^{+\,0.0005}_{-\,0.0009} , \qquad \left(\frac{\Delta m_d}{\Delta m_s}\right)_{\text{exp}} = 0.0285 \pm 0.0001 . \quad (3.51)$$

Additionally, the relation between the oscillation frequencies and the CKM matrix elements can be used to constrain the unitarity triangles, as visualised as orange bands in Figure 2.2 and Figure 2.3. Hence, measuring the oscillation frequency Δm_s is a consistency check of the SM. Also, New Physics scenarios in $b \to sll$ transitions, which trigger interest due to a variety of tensions with SM prediction observed in this sector, can be probed by the oscillation frequency Δm_s [47, 48]. Possible scenarios to explain such anomalies influence the bs-coupling and consequently the B_s^0-$\overline{B}_s^0$ transitions, by introducing significant contributions, next to the box diagrams.

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

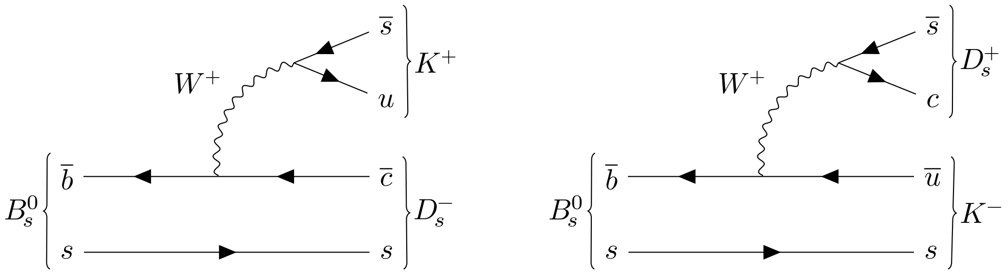
In the case of $B_s^0 \to D_s^{\mp} K^{\pm}$ the decay into both the $D_s^- K^+$ (f) and the charge conjugated $D_s^+ K^-$ ($\overline{f}$) final state is possible without the need of an oscillation. Both of these transitions are not only realised on tree level as illustrated in Figure 3.4, but also Cabibbo suppressed by a similar factor $|V_{cb}^* V_{us}| \approx \lambda^3 \approx |V_{ub}^* V_{cs}|$. Hence, all four decay amplitudes

$$\begin{aligned} A_f &= \langle D_s^- K^+ | \boldsymbol{T} | B_s^0 \rangle \;, & \overline{A}_f &= \langle D_s^- K^+ | \boldsymbol{T} | \overline{B}{}_s^0 \rangle \;, \\ A_{\overline{f}} &= \langle D_s^+ K^- | \boldsymbol{T} | B_s^0 \rangle \;, & \overline{A}_{\overline{f}} &= \langle D_s^+ K^- | \boldsymbol{T} | \overline{B}{}_s^0 \rangle \end{aligned} \tag{3.52}$$

are of the same order of magnitude causing significant interference effects. Additionally, the presence of the CKM matrix element V_{ub}, which is complex even at lowest-order approximations, predicts significant effects from *CP* violation. However, the increased Cabibbo suppression of λ^3 compared to the previously discussed $B_s^0 \to D_s^- \pi^+$ decays makes the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay mode less abundant. Nevertheless, with a relative branching ratio [35] of

$$\mathcal{B}\left(B_s^0 \to D_s^{\mp} K^{\pm}\right) \;=\; 2.25 \pm 0.12 \cdot 10^{-4} \tag{3.53}$$

this mode still contributes significantly to B_s^0-meson decays.



(a) $B_s^0 \to D_s^- K^+$ **(b)** $B_s^0 \to D_s^+ K^-$

Figure 3.4: Leading-order Feynman diagrams for the $B_s^0 \to D_s^+ K^-$ decay (a) and the $B_s^0 \to D_s^+ K^-$ decay (b).

As discussed in Section 3.2, the parameters λ_f and $\lambda_{\overline{f}}$ defined in Equation 3.16 summarise the *CP*-violating effects. The properties of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays can be used to simplify these parameters. The dominance of the tree-level transitions and the resulting absence of direct *CP* violation results in $\left|A_f\right| = \left|\overline{A}_{\overline{f}}\right| = A$ and $\left|A_{\overline{f}}\right| = \left|\overline{A}_f\right| = B$ leading to

$$\frac{\overline{A}_f}{A_f} = \frac{V_{ub} V_{cs}^*}{V_{cb}^* V_{us}} \left|\frac{B}{A}\right| \mathrm{e}^{i\delta} \qquad \text{and} \qquad \frac{\overline{A}_{\overline{f}}}{A_{\overline{f}}} = \frac{V_{cb} V_{us}^*}{V_{ub}^* V_{cs}} \left|\frac{A}{B}\right| \mathrm{e}^{-i\delta} \,. \tag{3.54}$$

Here, the strong phase difference δ is caused by final-state interactions, while the weak phases are absorbed in the CKM matrix elements. Further, following Equation 3.37 the ratio q/p from Equation 3.9 can be expressed by a single phase

$$\frac{q}{p} \approx \sqrt{\frac{m_{12}^*}{m_{12}}} = \mathrm{e}^{i\arg(m_{12})} . \tag{3.55}$$

This phase can be purely described by CKM matrix elements

$$\arg\frac{q}{p} \approx \arg\frac{V_{tb}^* V_{ts}}{V_{ts}^* V_{tb}} = \arg\frac{V_{cb}^* V_{cs}}{V_{cs}^* V_{cb}} - 2\beta_s \tag{3.56}$$

under the considerations from Section 3.3.1. By substitution of the unitarity triangle angles the *CP* parameters can be expressed by

$$\begin{aligned}
\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f} &= \frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*}\,\frac{V_{ub} V_{cs}^*}{V_{cb}^* V_{us}}\left|\frac{B}{A}\right| \mathrm{e}^{i\delta} \\
&= \frac{V_{cd}^* V_{cs}}{V_{ud}^* V_{us}}\,\frac{V_{ud}^* V_{ub}}{V_{cd}^* V_{cb}}\left|\frac{B}{A}\right| \mathrm{e}^{i(\delta-2\beta_s)} \\
&= \underbrace{\left(1-\frac{V_{td}^* V_{ts}}{V_{ud}^* V_{us}}\right)}_{1-\lambda^4\approx 1}\left|\frac{V_{ud}^* V_{ub}}{V_{cd}^* V_{cb}}\right|\left|\frac{B}{A}\right| \mathrm{e}^{i(\delta+\gamma-2\beta_s)} \\
&= \left|\lambda_f\right| \mathrm{e}^{i(\delta+\gamma-2\beta_s)}
\end{aligned} \tag{3.57}$$

and
$$\lambda_{\overline{f}} = \frac{p}{q}\frac{\overline{A}_{\overline{f}}}{A_{\overline{f}}} = \left|\lambda_{\overline{f}}\right| \mathrm{e}^{-i(\delta-\gamma+2\beta_s)} . \tag{3.58}$$

The combination of these two parameters can be used to either determine the strong phase difference

$$\delta = \frac{1}{2}\left(\arg\lambda_f - \arg\lambda_{\overline{f}}\right) \tag{3.59}$$

or to cancel it
$$\gamma - 2\beta_s = -\frac{1}{2}\left(\arg\lambda_f + \arg\lambda_{\overline{f}}\right) . \tag{3.60}$$

The previously discussed considerations leading to Equation 3.60 allow the CKM angle γ to be extracted from the measurable decay-time-dependent asymmetries given in Equation 3.39 with a minimum of assumptions using $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. In the literature, this procedure is sometimes referred to as a " theoretically clean determination of γ " [8]. However, even in these cases usually a few assumptions, like neglecting the B_s^0-mixing phase $2\beta_s := 0$, are made. In reality, a non-zero mixing phase, which for example can be measured in $B_s^0 \to J/\psi\phi$ decays $2\beta_s = \phi_s = -0.070 \pm 0.022$ [42], has to be constrained. However, this situation could be utilised to track anomalies, as the phases γ and $2\beta_s$ are introduced by different processes [49]. While the phase $2\beta_s$ is related to the mixing process, γ arises in the decay amplitudes.

Experimental status of the CKM angle γ

The CKM angle $\gamma^{\text{exp}} = \left(66.2^{+3.4}_{-3.6}\right)^{\circ}$ [42] is constrained by a wide variety of direct experimental measurements. The constraints can be probed against an indirect determination, where all measured properties (but γ) of the unitarity triangle are fitted to constrain the angle $\gamma^{\text{fit}} = (63.4 \pm 0.9)^{\circ}$ [50]. These two values are currently in good agreement with each other. However, small tensions can be observed between the individual direct measurements, which probe different processes.

The CKM angle γ is found in the phase introduced by interference effects between weak $b \to Wu$ and $b \to Wc$ transitions. It is the only CKM angle that can be measured from pure tree-level processes as it does not involve t-quark transitions (compare Equation 2.25). While the determination of γ from tree-level processes is of interest in this thesis, this parameter is also accessible via loop-level transitions. As these transitions could be significantly influenced by the presence of new non-SM particles, the comparison of tree and loop processes can be used to probe New Physics effects. Further, the measurements of γ can be divided into time-integrated and time-dependent analyses.

In tree-level decays, γ is accessible in the interference between $b \to Wu$ and $b \to Wc$ transitions. Hence, B-meson decays to D-mesons are analysed. In the time-integrated case usually $B^- \to D^{(*)0}K^{(*)-}$ decays are used. Here, the interference arises between the colour-favoured $B^- \to D^{(*)0}K^{(*)-}$ and the colour-suppressed $B^- \to \bar{D}^{(*)0}K^{(*)-}$ transitions. The corresponding analyses can be categorised by the final states of the D meson. The Gronau-London-Wyler (GLW) method [51, 52] makes use of *CP* eigenstates such as $D^0 \to K^+K^-$ or $D^0 \to \pi^+\pi^-$, while the Atwood-Dunietz-Soni (ADS) method [53, 54] relies on flavour-specific $D^0 \to K^+\pi(\pi^+\pi^-)$ decays. The Dalitz analysis of three-body final states like $D^0 \to K_{\mathrm{S}}^0\pi^+\pi^-$ and $D^0 \to K_{\mathrm{S}}^0K^+K^-$ is known as Giri-Grossman-Soffer-Zupan (GGSZ) method [55]. The sensitivity of all these measurements suffers under the imbalance of the decay amplitudes

$$r_B^{DK} = \left|\frac{A\left(B^- \to \bar{D}^0K^-\right)}{A\left(B^- \to D^0K^-\right)}\right| \approx 0.1 \quad \text{and} \quad r_B^{D\pi} = \left|\frac{A\left(B^0 \to D^+\pi^-\right)}{A\left(\bar{B}^0 \to D^+\pi^-\right)}\right| = \mathcal{O}(0.01) \tag{3.61}$$

caused by the colour suppression [42]. The same applies to the time-dependent analysis using $B^0 \to D^\mp \pi^\pm$ decays [56]. While the $B^0 \to D^\mp \pi^\pm$ analysis benefits from larger signal statistics $\mathcal{B}\left(B^0 \to D^\mp \pi^\pm\right) = 2.51 \pm 0.08 \cdot 10^{-3}$ [35], the previously discussed $B_s^0 \to D_s^\mp K^\pm$ decays are subject to larger interference effects [42]

$$r_B^{D_s K} = \left|\frac{A_{\bar{f}}}{A_f}\right| = \left|\frac{\overline{A}_f}{\overline{A}_{\bar{f}}}\right| \approx 0.37\,. \tag{3.62}$$

Besides, all time-dependent measurements require knowledge of the B meson flavour at production, which cannot be inferred trivially. To retrieve this information the flavour tagging technique introduced in Section 4.5 is needed, which again comes at the cost of statistical sensitivity. Until now, two time-dependent measurements of γ using tree-level B_s^0-meson decays have been performed. Both the measurement using $B_s^0 \to D_s^\mp K^\pm$ decays in the data set of the first $3\,\mathrm{fb}^{-1}$ recorded at LHCb [14] and the measurement based on $B_s^0 \to D_s^\mp K^\pm \pi^+ \pi^-$ decays in the full $9\,\mathrm{fb}^{-1}$ LHCb data set [57]

$$\gamma_{\mathrm{DsK}}^{3\,\mathrm{fb}^{-1}} = \left(128\,^{+\,17}_{-\,22}\right)^\circ\,, \qquad \gamma_{\mathrm{DsKpipi}}^{9\,\mathrm{fb}^{-1}} = (44 \pm 12)^\circ \tag{3.63}$$

trigger attention. The two results, which both rely on similar processes, are almost shifted by 90°. However, the tensions between the two measurements and with respect to the combination of γ are not significant. Nevertheless, next to the large measured value of γ, which would lead to a significant distortion of the unitarity triangle displayed in Figure 2.2, the measured branching ratios of $B_s^0 \to D_s^\mp K^\pm$ decays show small tensions to the corresponding predictions [58]. This situation triggers special interest in the analysis of $B_s^0 \to D_s^\mp K^\pm$ decays in the additional $6\,\mathrm{fb}^{-1}$ LHCb data set presented in Part III.

4 The LHCb experiment

The LHCb experiment is specialised to study b and c hadrons in pp collisions provided by the Large Hadron Collider (LHC). With an excellent decay-time resolution provided by the detector, this experiment is well suited for decay-time-dependent studies of B_s^0 mesons. Hence, the work presented in this thesis is based on data collected by the LHCb experiment during its second data-taking period (Run2) in the years 2015 to 2018. The following description of the experiment and associated facilities is referring to the state during this period. At first, the LHC is introduced in Section 4.1 based on Reference [13]. Then, the LHCb detector is presented in Section 4.2, followed by explanations of the trigger system in Section 4.3 and of the simulation in Section 4.4. Finally, the flavour tagging algorithms used at the LHCb experiment are introduced in Section 4.5.

The experiment is operated by an international collaboration, which currently includes 95 institutes in 22 different countries, with more than 1600 scientists and engineers.

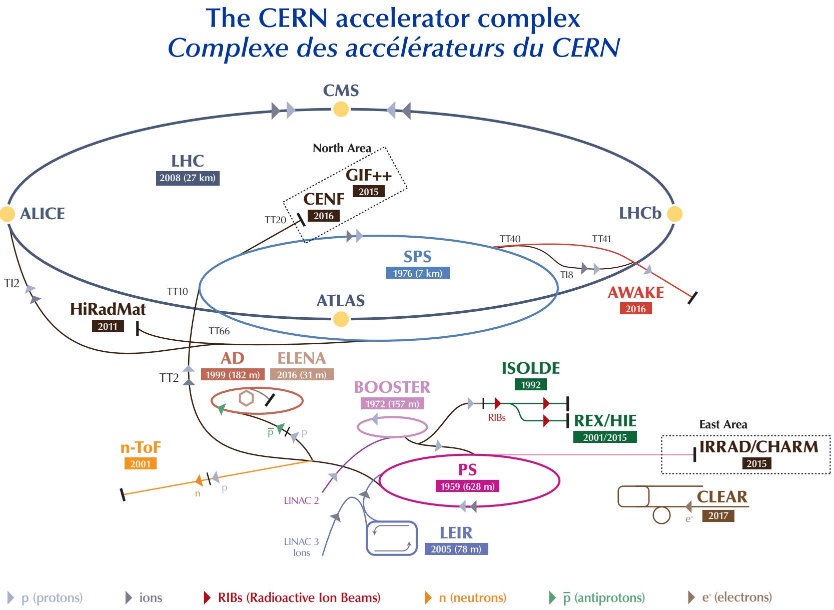


Figure 4.1: Schematic overview of the facilities used to accelerate proton and ions for the LHC at CERN [59]. The acceleration of protons colliding at LHCb is discussed in the text.

4.1 The Large Hadron Collider

The Large Hadron Collider is located in the Geneva area and is operated by the European Organisation for Nuclear Research (CERN). With a circumference of 26.7 km, it is the world's largest particle accelerator to date. The large circumference paired with the strong magnetic field of the superconducting bending magnets allows the LHC to accelerate and store high-energy proton beams. The LHC is designed to collide two counter-rotating beams with a centre-of-mass energy of $\sqrt{s} = 14$ TeV at a maximal luminosity of $2{\cdot}10^{34}\,\text{cm}^{-2}\text{s}^{-1}$.

To achieve the high luminosity a large number of protons has to be stored in the LHC. The beams consist of up to 2808 bunches containing $1.1{\cdot}10^{11}$ protons each. The nominal distance between the bunches is 25 ns allowing bunch-crossing rates up to 40 MHz. The high beam energy is reached by utilising many different accelerator facilities, including modernised predecessors of the LHC, to accelerate bunches of protons. While the illustration in Figure 4.1 shows the full variety of experimental facilities at the LHC complex, the following description focuses on the acceleration of protons colliding in the LHC.

The protons injected into the LHC are produced from ionised hydrogen. At first, the protons are accelerated by radio-frequency (RF) cavities using the LINAC2, a linear accelerator. Then a sequence of circular accelerators, starting with the BOOSTER, is used to increase the proton energy further. The Proton Synchrotron (PS), built in 1952, is re-utilised to accelerate the protons from 1.4 GeV to 25 GeV before they are transferred to the Super Proton Synchrotron (SPS). Here the protons are accelerated to energies of 450 GeV before they are finally fed to the LHC.

The LHC consists of eight bent and straight segments each. At the bent segments, in total 1232 dipole magnets are used to bend the proton beams. The super-conducting magnets are operated at temperatures below 2K and provide magnetic fields of more than 8T. At the straight segments, either experimental or utility facilities are located. One segment is instrumented with RF cavities to accelerate the protons to 6.5 TeV. Another straight hosts a mechanism to dump the proton beams. Two segments are equipped with systems of magnets to collimate the beams. The remaining four straight segments provide different interaction points (IP). where the two proton beams are crossed to allow interactions and collisions of proton bunches with centre-of-mass energies of up to $\sqrt{s} = 13$ TeV.

These points are equipped with large detectors to study the remnants of these interactions. At IP2 the ALICE experiment [60] is located, which is specialised for studies of quark-gluon plasma in heavy ion collisions. The ATLAS [61] and the CMS [62] experiments are located at IP1 and IP5, respectively. Both of these experiments are based on general-purpose detectors, which cover almost the full phase space and are designed to measure a wide variety of possible signals. Allowing similar measurements using mostly independent instruments, these two experiments can cross-check each other to provide profound results as proven by the discovery of the Higgs boson in 2012 [17, 18]. The LHCb experiment [12], which is located at IP8, will be discussed in more detail in the following.

4.2 The LHCb detector

The LHCb detector is a single-arm forward spectrometer covering an acceptance from 10 mrad to 300 mrad (250 mrad) in the horizontal (vertical) plane. The $b\bar{b}$ pairs produced in pp collisions at the high LHC energies are significantly boosted. Therefore, a large fraction of all $b\bar{b}$ pairs produced lay within the LHCb detector acceptance. This is illustrated in Figure 4.2 in the planes of the angle θ between the beam axis and momentum, and the pseudorapidity η of the quarks. The focus on the very forward direction allows for more extensive instrumentation of this area compared to the general purpose detectors covering the full spatial acceptance.

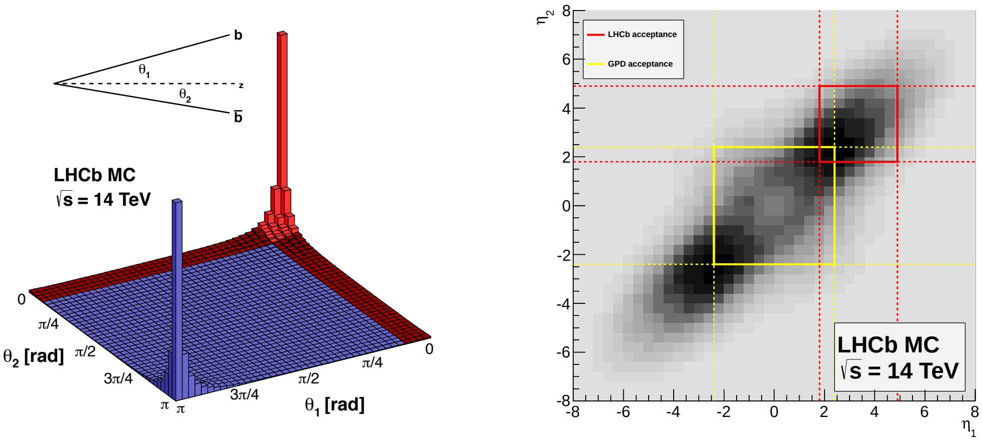


Figure 4.2: Production rates of $b\bar{b}$ pairs in simulated 14 TeV pp collisions for different kinematic regimes with the LHCb acceptance highlighted in red [64]. At the left, the rate is displayed for different angles between the $b(\bar{b})$ momentum and proton beam, demonstrating the boost in the forward region. The rate in bins of pseudorapidity is shown, together with a comparison to the acceptance of general-purpose detectors (yellow).

The LHCb experiment is designed to run at a reduced instantaneous luminosity of $2{\cdot}10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. At the cost of lower data statistics, the reduced luminosity increases the data quality, by decreasing the detector occupancy, and the radiation damage and by providing a constant luminosity level over a LHC fill. This modification is achieved by reducing the overlap of the proton beams at IP8 independently to the other interaction points, using a set of dipole magnets [63].

The detector system consists of different subsystems to reconstruct tracks and measure their momentum, and to identify the species of the underlying particle. A schematic overview of the full detector system is shown in Figure 4.3. The individual subsystems are described in the following.

4.2.1 Tracking detectors

The LHCb tracking system consists of a Vertex Locator (VELO) and a spectrometer based on tracking stations up and downstream from a magnet. The VELO [66] is a silicon strip detector which is located directly at the interaction point with a distance of only a few millimetres to the proton beams. This proximity between the detector and the interaction region is crucial for the excellent resolution for

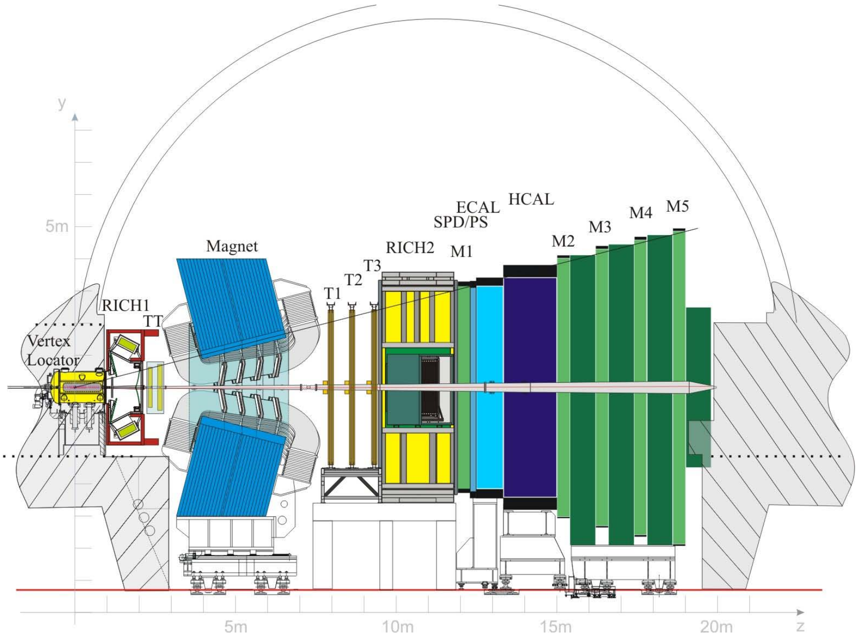


Figure 4.3: Schematic overview of the LHCb detector and its subsystems [65].

primary vertices (PV) of the *pp* interactions and secondary vertices (SV) from decays of long-lived particles. The achieved impact parameter resolution is below 35 µm for particles with large transverse momenta, $p_{\mathrm{T}} > 1\,\mathrm{GeV}/c$. This precise spatial resolution also translates into a resolution of less than 50 fs for the measurement of the decay time of long-lived particles decaying within the VELO, which is very important for decay-time-dependent measurements of B_s^0 mesons.

To protect the VELO and other sensitive components close to the beam pipe from being damaged by unstable proton beams, a beam condition monitor (BCM) system [67] is installed. The system consists of two circular stations, each instrumented with eight symmetrically-aligned sensors. The sensitive area of the sensors starts at a radial distance of 3.7 cm (5.05 cm) at the station located downstream (upstream) of the VELO. To provide the necessary radiation hardness, diamond sensors are used. As soon as the energy deposit measured by the BCM exceeds a defined threshold, a beam dump is requested to protect the LHCb components. Additionally, the VELO is opened during phases in which unstable beam conditions are expected, such as in the beam injection. This way, the distance of the detector to the beam is increased to avoid potential damage.

The momentum of a charged particle is inferred from the bending of its trajectory in a magnetic field. The magnetic field is provided by a conventional dipole magnet with a 1450 t iron yoke operated at room temperature. The integrated field strength along the beam axis is 4 Tm [68]. The polarity is regularly reversed to cancel effects from different detection efficiencies in different regions of the detector. The particle trajectory is reconstructed by information from three tracking stations (T1, T2 & T3) downstream to the magnet and a tracking station upstream (Tracker Turicenis, TT). Each of the three downstream tracking stations consists of an Inner Tracker (IT) and an Outer Tracker (OT).

The TT and the IT, also known as silicon trackers (ST) are silicon strip detectors [69] providing the necessary granularity to cope with the high detector occupancy close to the beam pipe. The hit efficiencies of the ST are above 99 % and the spatial hit resolutions are in the order of 50 µm [65]. The IT covers only the central 1.2 m × 0.4 m region. The OT then extends the instrumented area to the full LHCb acceptance. The OT is a straw tube detector filled with a gas mixture designed to provide low drift times of the same order of proton bunch gaps [70]. The spatial hit resolution of the OT is in the order of 200 µm. Each tracker station has four layers of modules arranged in a *xuvx* pattern. In the x layers the strips (straws) are vertically arranged. The u and v layers are rotated by a stereo angle of $-5°$ and $+5°$ respectively. This arrangement allows the 2-dimensional estimation of the hit position. The average track finding efficiency for the combined tracking system is above 96 %. The relative momentum resolution for charged tracks is in the range of 0.5 % for low momenta (below 20 GeV/c) and 0.8 % for high momenta (around 100 GeV/c).

4.2.2 Particle identification

Particle identification (PID) at LHCb is achieved by combining information from different detector systems, two Ring Imaging Cherenkov (RICH) detectors, a calorimeter system and a muon system. The RICH detectors measure the opening angle of Cherenkov light emitted by charged particles passing through a medium (fluorocarbon gas). Together with the momentum information from the tracking system, this angle gives an estimate of the particle's species as illustrated in Figure 4.4, since the momentum-speed relation is mass dependent and the angle depends on the emitting particle's speed. The two RICH detectors cover different momentum ranges [71]. In RICH1, which is located upstream to the tracking system, C_4F_{10} is used as radiator medium

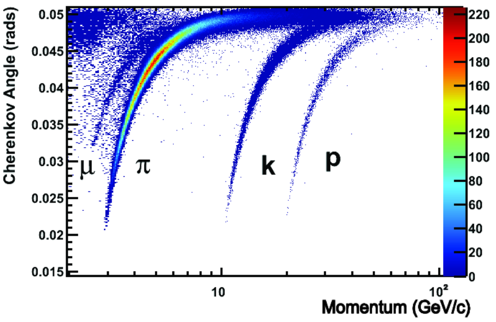


Figure 4.4: Reconstructed Cherenkov angles and momenta for isolated tracks of different species in the C_4F_{10} radiator [71].

to have a good resolution at low momenta (2-40 GeV/c) for particles, which possibly leave the detector acceptance, when they are deflected in the magnetic field. Located after the tracking system, RICH2 is based on CF_4 to cover the high momentum range (15-100 GeV/c) of particles, which are only slightly deflected passing through the magnetic field.

The calorimeter system consists of the Preshower (PS), the Scintillating Pad Detector (SPD), the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The PS and SPD are two very similar small detectors, based on scintillating pads separated by a 12 mm (2 radiation lengths) thick lead absorber [72]. Information from PS and SPD are mainly used to suppress background in the trigger described in Section 4.3. Both ECAL and HCAL are designed in the form of shashlik calorimeters, consisting of multiple alternating layers of scintillators and absorbers [73]. The ECAL uses lead absorbers and provides a combined thickness of 25 radiation lengths and 1.1 nuclear interaction lengths. The HCAL with a combined thickness of 5.6 nuclear interaction lengths utilises iron absorbers.

A system of multi-wire proportion chambers is used to track and identify muons with high precision. The first station (M1) is located upstream of the calorimeter system and is used to set a seed for muon tracks and for triggering. The inner part of this first station is based on GEM detectors [74]. Four additional stations (M2-M5) are placed downstream to the calorimeter system, which absorbs most of the particles produced in the collisions due to its combined thickness of 6.6 interaction lengths. The only particles passing through the calorimeter with high probability are muons, which are minimal ionising particles at the GeV/c momentum scale, and neutrinos. Since neutrinos do not interact with the detector at all, all tracks reconstructed in the full muon system are muons at a high probability. This enables a rejection rate of above 99 % for misidentified hadrons while maintaining a muon signal efficiency of above 98 % [75].

4.3 The LHCb trigger system

It is not possible to save the full event information at the high LHC bunch-crossing rate of 40 MHz due to technical limitations of the bandwidth. The rate is reduced by a trigger system. The trigger applies kinematic and topological criteria to events. This way the small fraction of events, in which b- or c-hadrons are considered to be produced, can be selected, while events probably not contributing to the LHCb physics program are discarded. The trigger system consists of a hardware trigger (L0) and a software trigger divided into two stages (HLT1 & HLT2). The following description of the trigger system is based on Reference [76].

4.3.1 Hardware trigger stage

The overall readout rate of the LHCb detector as a whole is limited to about 1 MHz. Therefore, field-programmable gate arrays with fixed latency are used as a hardware trigger to reduce the rate. Information from the calorimeter system and the muon system, which individually allow higher readout rates, are used for the L0 trigger decision. There are multiple L0-trigger lines, which are sets of requirements an event needs to fulfil to be accepted by the trigger. Most lines accept a limited number of hits in the SPD to discard extraordinarily complex events which would need large computation times in the software trigger.

Trigger lines to select muonic events require either a track (L0Muon) or a system of two tracks (L0DiMuon), which are reconstructible in the muon system and which transverse momentum exceeds a given threshold. These lines provide a signal efficiency of up to 90 % for the decay $B_s^0 \to \mu^+\mu^-$. The L0 signal efficiency for hadronic decays is lower, e.g. in the order of 50 % for $B_s^0 \to D_s^-\pi^+$ decays. Trigger lines designed for hadronic decays (L0Hadron) require clusters in the calorimeter system with a minimum energy E_T deposed transverse to the beam line. A similar, but looser criterion is required by trigger lines specialised to electromagnetic interacting particles (L0electron, L0photon). These particles are distinguished from hadrons by information on the shower from SPD and PS. Trigger lines dedicated to QCD measurements select events with a low particle multiplicity.

It needs to be taken into account that the selection thresholds for the hardware trigger have been varied and optimised over the data-taking period. This can lead to small differences between data sets collected in different years.

4.3.2 Software trigger stage

The high-level trigger HLT directly processes the output information from the hardware trigger L0 as a part of the LHCb data collection and storage. Therefore, the selection criteria applied by the HLT are also referred to as online selection. The HLT, purely implemented in software, runs on a computing farm, called the Event Filter Farm (EFF). The software trigger can be divided into two stages.

The HLT1 is running selections based on the first rudimentary reconstructions on the output of the L0 trigger. This way the rate of accepted events, which are written to a disk buffer of 10 PB size, is reduced to about 110 kHz. Due to timing constraints, events are only partially reconstructed in the first stage of the HLT. The HLT1 reconstruction starts by forming VELO hits into tracks. These tracks are combined to find primary vertices (PV). Additionally, VELO tracks are, if possible, extrapolated to the tracking stations, first to the TT then to the downstream trackers. The so-called long tracks passing through all three of the trackers are then fitted with a Kalman filter. The only particle identification available at this stage is the presence of the track in the muon system. Criteria applied to events in the HLT1 are based on the transverse momentum p_{T}, displacement relative to PVs or the fit quality of tracks.

The events stored at the buffer can be used for calibration and alignment of the detector in real time. Every task of the real-time calibration is based on the output of a dedicated HLT1 line and performed whenever a data sample of sufficient size is buffered. The calibrated constants of the aligned detector are then used for the HLT2.

In the second stage of the software trigger, a full event reconstruction is performed. The HLT2 reconstruction is divided into three steps. First tracks of charged particles are reconstructed. Then tracks of neutral particles are reconstructed before particle identification is applied. Due to the relaxed timing constraints compared to HLT1, more sophisticated algorithms are used for the track reconstruction, yielding precise momentum estimates. In addition, the full PID information from the RICH detectors and from the calorimeter system is available. Around 500 HLT2 lines are used to select events of interest for various physics cases based on this reconstruction. The combined output of all lines is stored at a rate around 12.5 kHz. The stored data is then available for offline analyses as presented in this thesis.

4.4 LHCb simulation

Simulated data can deliver valuable knowledge and is used on various occasions in data analysis. By providing full information on the true characteristics, simulation can be used to validate analysis procedures. Aside from this, it can be used to study reconstruction and selection effects, like efficiencies, or for the development of models and algorithms.

For the simulation of LHCb events the Gauss framework [77] is used. This framework assembles all software packages simulating different aspects of the event. The p p collision is simulated by Pythia [78]. Then the EvtGen package [79] handles the hadronisation and the decay of the particles produced in the collision, taking effects from mixing and *CP* violation into account. The radiation of photons from long-lived particles is simulated by Photos [80]. The Geant4 toolkit [81] simulates the interaction of particles with the detector material using a detailed model of the detector.

The Boole framework [82] is then used to map the generated particles to detector responses by performing the digitisation. This includes the simulation of electronics and the L0 hardware trigger. The provided detector responses are in the same format as the data collected by the real detector. So from this step on the simulation is handled in an equal manner to data, including the software trigger, reconstruction and offline processing.

Due to finite knowledge of the simulated processes and limited computational resources, the simulation does not always perfectly resemble the real data collected by the detector. Therefore, the simulation tools are constantly improving by incorporating the latest theoretical model, measurements and knowledge from the behaviour of collected data. To cope with these imperfections different methods are applied. Simulated events can be reweighted in a data-driven way to resemble observed data distributions. The PidCalib package [83] can be used to correct the complex particle identification estimates.

4.5 Flavour tagging at LHCb

An important input to studies of neutral mesons is the mesons' initial flavour state at the time of their production and their final flavour state at the time of the decay. This information is not directly accessible in data as the detector can only distinguish between different particle charges and not between the electroweak flavour states. However, in flavour-specific decays, such as $B_s^0 \to D_s^- \pi^+$, the flavour at decay can be reconstructed from the charges of the decay products. The initial flavour of neutral, mixing mesons cannot be determined from the measured decay products. Therefore, at LHCb multiple algorithms exploiting tracks of particles, produced in the course of the signal particle's hadronisation process, are used to obtain knowledge of the initial flavour of B mesons.

For every single event, these algorithms will give both a predicted production flavour, the tag decision d, and a predicted mistag η, which gives an estimate of the algorithm's probability of delivering the wrong tag decision. The tag decision is defined to be $d = 1$ ($d = -1$) if the algorithm predicts the initial flavour to be B ($\overline{B}$) and to be $d = 0$ in case the algorithm is not able to give information on the initial state. The predicted mistag, a probability to retrieve the wrong decision, is defined in the range $\eta \in [0, 0.5]$ A mistag prediction $\eta = 0.0$ corresponds to a perfect decision, $\eta = 0.5$ to a random guess. While the tag decision is usually based on charge information of a track expected to originate from the B meson's fragmentation, the predicted mistag is usually obtained from a multivariate classifier analysing the properties of the selected track and the signal.

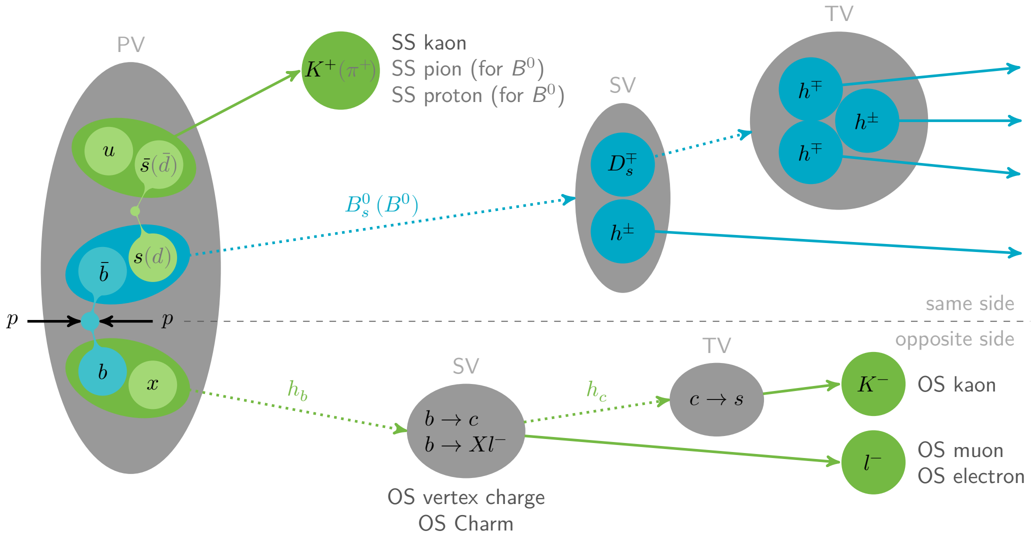


Figure 4.5: Schematic overview of the $B^0_{(s)}$-meson hadronisation process in a pp collision at LHCb adapted from [84]. The decay chain of the signal B meson is visualised in blue, and the fragmentation processes exploited by different flavour taggers in green.

At LHCb, this approach to obtain information on the initial B flavour is called Flavour Tagging (FT) and the algorithms used are called taggers. In the following the principles of these taggers are discussed. The algorithms can be divided into two classes, based on the different processes exploited: the opposite-side (OS) and the same-side (SS) taggers. A schematic overview of the different processes used by different taggers is shown in Figure 4.5.

Lately, a new inclusive FT approach has been developed at LHCb. In this case, a neural network is used to evaluate information on all non-signal tracks of a proton-bunch crossing at once to get information on the initial signal flavour. The network is expected to learn and exploit the same (and more) mechanisms as used by the individual SS and OS taggers. More details on the inclusive FT approach are given in Section 14.2.

Opposite-side taggers

The opposite-side taggers make use of the production mechanism of B mesons. The boosted b quarks within the LHCb detector acceptance are mostly produced in $b\bar{b}$ pairs by the strong interaction. So it is very likely that next to the signal B meson another b hadron is produced in the same event, whose production flavour is directly related to the initial signal flavour. The OS taggers try to identify parts of the decay chain of this accompanying b hadron. The baseline principle of the different taggers is sketched on the bottom half of Figure 4.5.

The OS muon and the OS electron taggers are using leptons from semi-leptonic decays of the accompanying b hadron. These decays are based on the weak decay $b \to W^+(\to l^+\nu_l)c$. So the lepton charge is directly related to the accompanying b hadron flavour at decay.

The OS kaon tagger's decision is based on the charge of a kaon originating from the decay chain $b \to c \to s$, where the kaon charge determines the b hadron's charge at the time of its decay.

The OS charm tagger is based on a similar, but more inclusive approach as it reconstructs c hadrons (D^0, D^+ or $\mathit{\Lambda}_c^+$) from the $b \to c$ decay by combining particles potentially originating from the OS decay. In the case of a flavour-specific c-hadron decay, the charges of the final-state particles can be used to get the accompanying b hadron's flavour at decay. However, in the case of a D^0 meson, there is a finite chance, that the meson oscillates before its decay, which is an additional source of wrong tag decisions based on this particle.

All of the previously mentioned OS taggers rely on an exclusive particle produced in the accompanying b-hadron decay chain. This can be problematic if multiple particles of the same species, but with different charge, are produced in the decay. A prominent example is the b-hadron decay into a charmonium resonance $(c\overline{c}) \to l^+l^-$ decaying into two leptons. In this case, selecting the correct lepton is impossible. A more inclusive strategy is used by the OS vertex charge tagger.

The OS vertex charge tagger reconstructs a possible secondary vertex from particle tracks related to the same primary interaction but separated from the signal candidate using a multivariate classifier. The average charge of the selected tracks weighted by the likelihood to originate from a second b hadron decay is used as an estimate of the b hadron's charge. Since all long-lived b hadrons share the same b flavour to charge relation (b: $q = +1$, $\overline{b}$: $q = -1$) the flavour of the accompanying hadron and therefore of the signal at production can be inferred from the charge. This approach fails if either the b hadron is neutral or short-lived. Additionally, selecting random tracks or discarding decay products of the accompanying hadron have negative impacts on the decision made by this tagger.

All OS taggers suffer from reduced performance due to dilution from accompanying b-quarks hadronising into neutral B mesons. In this case, the vertex charge tagger cannot infer any charge of a secondary vertex, while all other OS taggers provide the wrong tag decision at a rate matching the B meson's mixing probability. For

B^0 mesons the probability to mix before the decay is about 20 %, for the faster oscillating B_s^0 mesons approximately 50 %.

Same-side taggers

The same-side taggers exploit the fragmentation process of the signal B meson as sketched in the top half of Figure 4.5. For the hadronisation process of the $\overline{b}$ quark bound in the signal B^0 (B_s^0) meson an additional d (s) quark is necessary. This quark is created in a $d\overline{d}$ ($s\overline{s}$) pair in the course of the of the strong interaction. This way, the flavour state of the residual $\overline{d}$ ($\overline{s}$) quark at the time of the signal production is directly related to the production flavour of the signal B meson.

The SS taggers try to identify the hadron produced in the hadronisation of the left residual related to the signal flavour. Therefore, there are different taggers for B^0 mesons and B_s^0 mesons. The SS pion (proton) tagger selects a charged pion (proton) as a tagging particle, which is created close to the signal and moves in a similar direction, to tag B^0 mesons. The SS kaon tagger selects a charged kaon similarly. From the charge of the selected tagging particle, the initial flavour of the signal can then be inferred.

The tag information of the SS taggers is affected by the difficulty of selecting the correct hadron. It is possible in the fragmentation process, that the specific hadron species exploited by the tagger is either not or ambiguously produced. In this case, the necessary information is not available. Further, it may happen that a correct tagging particle does not fulfil the taggers' selection criteria or that a random hadron unrelated to the signal fragmentation by chance fulfils the criteria.

4.5.1 Flavour tagging characteristics

The FT performance is usually described by three characteristic measures: the tagging efficiency ε_{tag}, the true mistag ω and the effective tagging efficiency (also known as tagging power) ε_{eff}. The tagging efficiency gives the fraction of candidates within a sample, for which a tagger can give a decision, without considering the correctness of the decision

$$\varepsilon_{\text{tag}} = \frac{N_{\text{right}} + N_{\text{wrong}}}{N_{\text{right}} + N_{\text{wrong}} + N_{\text{untagged}}}. \tag{4.1}$$

The true mistag gives the probability that the tag decision for a candidate is wrong. For a sample of flavour-tagged candidates, it is described by the fraction

$$\omega = \frac{N_{\text{wrong}}}{N_{\text{right}} + N_{\text{wrong}}}. \tag{4.2}$$

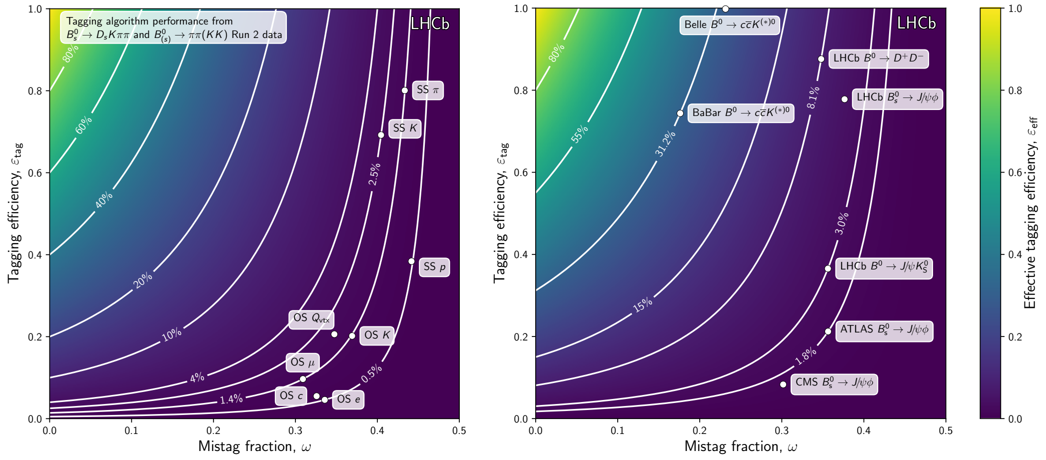
For an individual candidate, this probability can be derived from a calibration $\omega(\eta)$ of the predicted mistag derived on a larger sample as described in Section 4.5.2. Candidates with incorrectly estimated initial flavour dilute flavour tagged measurements. The size of the dilution $D = 1 - 2\omega$ is dependent on the mistag probability and can easily be derived from measured asymmetries:

$$\begin{aligned} \mathcal{A}_{\text{meas}} &= \frac{N(d=1) \, - \, N(d=-1)}{N(d=1) \, + \, N(d=-1)} \\ &= D \cdot \frac{N(B) \, - \, N(\overline{B})}{N(B) \, + \, N(\overline{B})} \; = \; D \cdot \mathcal{A}_{\text{true}} \, . \end{aligned} \tag{4.3}$$

Usually, there is a trade-off between efficiency and mistag probability while tuning a tagging algorithm. Therefore, the tagging power $\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} \cdot D^2$ takes both the tagging efficiency and dilution effects from mistagged candidates into account. This measure is best suited to compare the FT performance of different taggers or samples, as for every sample it can describe the decrease of statistical power originating from the imperfect knowledge of the the initial B flavour. While the statistical error of an asymmetry $\sigma_{\mathcal{A}_{\text{true}}} \sim \frac{1}{\sqrt{N}}$ is expected to decrease with the sample size N, an effective sample size $N_{\text{eff}} = \varepsilon_{\text{eff}} \cdot N$ can be defined to show the impact of the FT on the statistical sensitivity of a measurement

$$\sigma_{\mathcal{A}_{\text{meas}}} \sim \frac{1}{\sqrt{N_{\text{eff}}}} \, . \tag{4.4}$$

Hence, the statistical sensitivity is dependent on the effective sample size and can be improved both by higher statistics and by higher tagging power. An overview of the FT performances achieved in some representative analysis is presented in Figure 4.6 in terms of tagging efficiency, mistag probability and resulting tagging power. It must be stated, that the FT performance is related to the signal kinematics and the event topology and can be dependent on the selection applied to the data. Therefore, the tagging power of the three LHCb measurements presented in Figure 4.6b ranges between 3.0 % in the $B_s^0 \to J/\psi K_S^0$ analysis, where the selection benefits from the clear structure of the charmonium resonance, and 8.1 % in the $B^0 \to D^+ D^-$analysis, where the pure hadronic final state requires tighter selections. In comparison to LHCb, the B factories, Belle and BaBar, show much higher tagging powers $\mathcal{O}(30\,\%)$ due to their leptonic collision environment, which yields a much simpler event topology. However, this disadvantage is overcome by the very high luminosity at the LHC resulting in much larger data samples. In comparison to ATLAS and CMS, LHCb shows a much better FT performance benefiting from the experiment's focus on B physics and its superior instrumentation of the detector's forward region.



(a) **(b)**

Figure 4.6: Comparison of the FT performances achieved by individual taggers used at LHCb (left) and achieved in representative analyses performed at different experiments (right). Due to different selection requirements, the tagging power of the three chosen analyses performed at LHCb varies in the range between 3.0 % and 8.1 %. [85]

4.5.2 Flavour tagging calibration

The multivariate classifiers used for the estimation of the predicted mistag are based on the kinematic properties of the tagging track and the signal candidate. The possibility of a dependence of the classifier from the specific decay channel or sample cannot be excluded. Possible reasons for such a dependency could be over-training in the process of the taggers' development, or simple kinematic differences due to different selections or differences in the reconstruction of different decay modes. Therefore, the mistag predicted by the algorithm cannot be taken as a perfectly predicted mistag probability.

Such effects usually can be corrected by applying a calibration $\omega(\eta)$ to the predicted mistag. The calibration can be derived from a sample with similar kinematics, where either the initial flavour is known, e.g. simulated samples or self-tagged charged B^+ modes, or can be statistically derived from the known decay flavour of a flavour-specific channel taking knowledge of the oscillation into account. In general, the calibration function follows the form

$$\omega(\eta) = g\left[g^{-1}(\eta) + \sum_k \theta_k \cdot P_k(g^{-1}(\eta))\right], \tag{4.5}$$

with calibration coefficients θ_k for a set of polynomial basis vectors P_k and a link function g. Often, a logistic link function is applied to maintain the probabilistic character of the mistag. However, in the scope of this thesis, a linear calibration function without the application of a link function turned out to be sufficient. Taking potential differences in the calibration of particle and antiparticle into account, this yields two calibration functions

$$\omega^{\bar{b}}(\eta) = \theta_0 - \frac{\Delta\theta_0}{2} + \left(\theta_1 - \frac{\Delta\theta_1}{2}\right) \cdot (\eta - \overline{\eta}) \quad , \text{ for } B \tag{4.6}$$

and

$$\omega^{b}(\eta) = \theta_0 + \frac{\Delta\theta_0}{2} + \left(\theta_1 + \frac{\Delta\theta_1}{2}\right) \cdot (\eta - \overline{\eta}) \quad , \text{ for } \overline{B} \tag{4.7}$$

At LHCb there are currently two software packages used to easily perform such calibration, the ROOT and c++ based EspressoPerformanceMonitor [86] and the newer python-based lhcb-ftcalib package [87].

4.5.3 Tagger combination

A large set of different taggers is used at the LHCb experiment. To get a single prediction of the initial flavour of a B meson the tag decisions and mistag estimates of multiple taggers can be combined. The combined tag decision and mistag estimate

$$d_{\text{comb}} = \text{sign}(P_{\bar{b}} - P_b)\,, \tag{4.8}$$

$$\eta_{\text{comb}} = 1 - \max(P_b\,,\, P_{\bar{b}}) \tag{4.9}$$

are based on the probabilities $P_b = \frac{p_b}{p_b + p_{\bar{b}}}$ and $P_{\bar{b}} = 1 - P_b$ for the signal candidate to contain a b or a $\bar{b}$ respectively. These are based on the per-candidate likelihoods

$$p_b\big(\vec{\omega},\, \vec{d}\big) = \prod_i \left(\frac{1 + d_i}{2} - d_i \cdot [1 - \omega_i(\eta_i)]\right)\,, \tag{4.10}$$

$$p_{\bar{b}}\big(\vec{\omega},\, \vec{d}\big) = \prod_i \left(\frac{1 - d_i}{2} + d_i \cdot [1 - \omega_i(\eta_i)]\right)\,, \tag{4.11}$$

with the index i running over the individual taggers. This combination is usually applied by the use of the EspressoPerformanceMonitor [86] or lhcb-ftcalib package [87]. To obtain precise estimates for the tagging power of a set of combined taggers, uncertainties originating from the calibration of individual taggers have to be propagated through the combination algorithm. This error propagation has been implemented within the lhcb-ftcalib package.

5 Statistical methods

A crucial part of high-energy physics analyses is the estimation of parameters from data and the separation of different data classes (e.g. background and signal). In this context, the presented analyses make use of different statistical methods. In the following, the most important techniques are presented. At first, two different multivariate classification algorithms are introduced in Section 5.1. Then, the maximum-likelihood method for parameter estimation from data is explained in Section 5.2 based on References [88, 89]. In Section 5.3 the sPlot technique [90, 91], used to unfold different data components based on a fitted model, is described.

5.1 Multivariate classifiers

A common task in data analyses is classification. In the field of high-energy physics, classification can be encountered in various exercises, like the selection of signal events from a background diluted sample or the discrimination of different particle states in the flavour tagging. Usually, a univariate classification based on a single feature x of the data set does not yield good discrimination. A multivariate analysis does not only improve the classification by using multiple features but also by exploiting potential correlations within the feature set $\vec{x}$.

Machine learning algorithms are commonly used for multivariate classifications. These algorithms can be divided into supervised learning and unsupervised learning algorithms. While supervised learning requires labelled training samples with a known true class y for every data set, this is not necessary for unsupervised learners, such as cluster algorithms. In the following, two supervised algorithms are introduced. In Section 5.1.1 boosted decision trees are presented based on References [92, 93], and in Section 5.1.2 recurrent neural networks are discussed.

5.1.1 Boosted decision trees

In the presented analyses, boosted decision trees (BDTs) are used to distinguish between signal and background candidates based on the candidate's kinematics and topology. A BDT is an ensemble of decision trees, which are trained using a boosting algorithm. A decision tree gives an estimate $\hat{y}$ for the true class y, by testing the given feature set $\vec{x}$ for different criteria c. These criteria are derived in the training process by splitting a data sample recursively. Starting at a root node, the sample is split by a rectangular cut on a single observable of the given feature set. Both the feature and the cut point are chosen to optimise the separation, by minimising a loss function $\mathcal{L}(\hat{y}, y)$ in both subsamples. This procedure is then repeated for each subsample until the separation of the sample cannot be further improved or a given depth is reached.

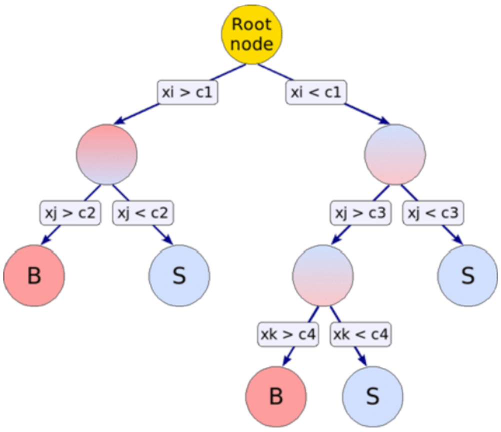


Figure 5.1: Scheme of a single decision tree distinguishing two classes S and B [94].

Usually, the discrimination using a single decision tree is rather low and yields large classification biases. This situation can be improved by using an ensemble of decision trees. One approach to create such an ensemble is boosting. In this approach, the classification of a candidate i is

$$\hat{y}_i = T(\vec{x}_i) = \frac{\sum_{k=1}^{N_\text{tree}} \alpha_k \, T_k(\vec{x}_i)}{\sum_{k=1}^{N_\text{tree}} \alpha_k} \tag{5.1}$$

the weighted average of the individual trees $T_k(\vec{x}_i)$. The weights α_k are commonly based on the performance of the corresponding tree. Usually, in boosting the k-th classifier is trained on a data sample where candidates misidentified by the previous classifier $T_{k-1}(\vec{x}_i)$ are weighted to be more important in the training. For a sufficiently large number of trees N_tree the classification becomes quasi-continuous.

The decision trees and the boosting algorithm used in the presented analyses are implemented in the XGBoost framework [93], which also provides different regularisation approaches to avoid overtraining. Additionally, a calibration of the classification is implemented in this framework, allowing a probabilistic interpretation of the classifier.

5.1.2 Recurrent neural networks

In Section 14.2 another multivariate classifier, a recurrent neural network (RNN) based on gated recurrent units (GRU), is used. Artificial neural networks (NN) are commonly used machine learning algorithms, which are discussed detailed for example in Reference [96]. The baseline architecture of a NN is a feed-forward network. Such a network is organised in layers. The input layer representing the input features $\vec{x}$ is mapped to the output layer $\hat{y}$ via a number of hidden layers $\vec{h}_n$ as sketched in Figure 5.2. The units of the $(n+1)$-th layer are linear transformations of the previous n-th layer yielding an output function

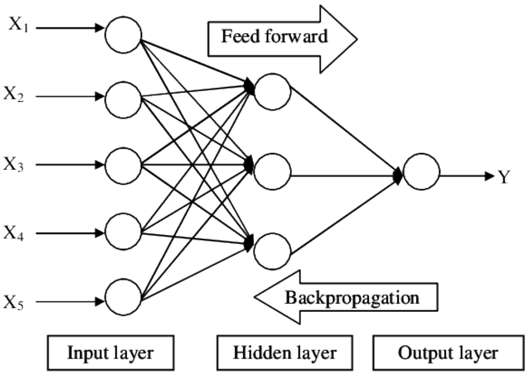


Figure 5.2: Sketch of typical feed-forward NN with one hidden layer [95].

$$\hat{y} = \vec{h}_{n+1}\big(\vec{h}_n\big) = \phi_n\big(W_n \cdot \vec{h}_n + \vec{b}_n\big)\,, \tag{5.2}$$

$$\vec{h}_0 = \vec{x}. \tag{5.3}$$

This linear transformation consists of a weight matrix W_n and a bias vector $\vec{b}_n$. Additionally, a non-linear activation function ϕ_n is applied to prevent the network from collapsing to a single linear transformation of the input vector. In the training process, the weight matrices and bias vectors are iteratively optimised to minimise a loss function $\mathcal{L}(\hat{y}, y)$ comparing the prediction to the true label. In every iteration the weight matrices

$$W_{nij} \to W_{nij} - \alpha \, \frac{\partial \mathcal{L}(\hat{y},\, y)}{\partial W_{nij}} \tag{5.4}$$

(and the bias vectors) are updated by back-propagating the gradient of the loss function. The hyperparameter α, the so-called learning rate, can be used to adjust the convergence of this procedure.

A large variety of extensions can be used to specialise neural networks for specific tasks. An example of such an extended architecture is a recurrent neural network. RNNs are designed to handle input in the form of sequential feature sets like time series commonly used in the field of speech recognition. A simple approach to adapt Equation 5.2 to handle a variable-length sequence of input features $\vec{x}_t$ is

$$\vec{h}_{n,\,t}\big(\vec{h}_{n-1,\,t},\, \vec{h}_{n,\,t-1}\big) = \phi_n\big(W_{n-1} \cdot \vec{h}_{n-1,\,t} + V_n \cdot \vec{h}_{n,\,t-1} + \vec{b}_{n-1}\big)\,, \tag{5.5}$$

$$\vec{h}_{0,\,t} = \vec{x}_t\,. \tag{5.6}$$

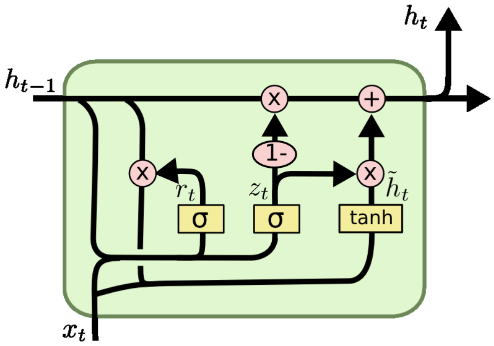


Figure 5.3: Sketch of a gated recurrent unit [97].

In this approach, each unit receives input from both, the previous layer of the same time step $\vec{h}_{n-1,\,t}$ and the same layer of the previous time step $\vec{h}_{n-1,\,t}$ transformed by a separate weight matrix V_n. However, this approach introduces the issues of exploding or vanishing gradients in the back propagation [98]. An RNN architecture commonly used to overcome these issues is the Long Short-Term Memory (LSTM) [99]. In Section 14.2 a similar, but simplified architecture based on gated recurrent units (GRU) [100, 101] is used. A GRU, sketched in Figure 5.3, utilises a reset gate r and an update gate z to control the information flow avoiding effects from exploding or vanishing gradients. The GRU activation

$$h_{n,\,t} \;=\; \left(1 - z_{n,\,t}\right) \cdot h_{n,\,t-1} \;+\; z_{n,\,t} \cdot \tilde{h}_{n,\,t} \tag{5.7}$$

is composed of the previous time step's activation $h_{n,\,t-1}$ and the candidate activation

$$\tilde{h}_{n,\,t} \;=\; \tanh\left[W_{n-1} \cdot h_{n-1,\,t} \;+\; V_n \cdot \left(r_{n,\,t} \odot h_{n,\,t-1}\right)\right], \tag{5.8}$$

where $\odot$ denotes an element-wise multiplication. Both the reset gate

$$r_{n,\,t} \;=\; \sigma\left(W_n^r \cdot h_{n-1,t} \;+\; V_n^r \cdot h_{n,t-1}\right) \tag{5.9}$$

and the update gate

$$z_{n,\,t} \;=\; \sigma\left(W_n^z \cdot h_{n-1,t} \;+\; V_n^z \cdot h_{n,t-1}\right) \tag{5.10}$$

use a logistic sigmoid activation σ and individual weight matrices $W^{r/z}$ and $V^{r/z}$. The GRU-based RNNs trained later are based on the Keras library [102] within the TensorFlow framework [103].

5.2 Maximum-likelihood method

The estimation of parameters from data is a crucial ingredient of the presented analyses. A well-established method for parameter estimation is the maximum-likelihood fit [88]. For n measurements of a set of observables $\vec{x}$ a likelihood

$$\mathcal{L}\left(\vec{\lambda}\right) = \prod_{i=1}^{n} \mathcal{P}\left(\vec{x}_i|\vec{\lambda}\right) \tag{5.11}$$

can be defined, based on a probability density function (PDF) $\mathcal{P}\left(\vec{x}_i|\vec{\lambda}\right)$ described by a set of parameters $\vec{\lambda}$. The optimal set of parameters

$$\hat{\vec{\lambda}} = \arg\max\, \mathcal{L}\left(\vec{\lambda}\right) \tag{5.12}$$

can be found by maximising the likelihood, the joint probability to measure the observables $\vec{x}$ for a given set of parameters. In practice, usually the negative logarithm of the likelihood is minimised

$$\hat{\vec{\lambda}} = \arg\min\left[-\ln\mathcal{L}\left(\vec{\lambda}\right)\right], \tag{5.13}$$

with

$$\ln\mathcal{L}\left(\vec{\lambda}\right) = \sum_{i=1}^{n} \ln\mathcal{P}\left(\vec{x}_i|\vec{\lambda}\right). \tag{5.14}$$

While the logarithm as a monotone function does not affect the result, it makes the computation numerically more stable by transferring the product to a sum. The negative value of the log-likelihood is optimised, because optimisers are usually designed for minimisation.

Multiple PDFs $\mathcal{P}_c(\vec{x}_i|\vec{\lambda}_c)$ can be combined to optimise parameter sets for a number N_s of different data categories c (e.g. backgrounds or data taking conditions) simultaneously. Different normalisations $\mathcal{N}_c$ can be taken into account for each PDF by extending the likelihood with a Poisson term as discussed in Reference [89]

$$\mathcal{L}\left(\vec{\lambda}\right) = \frac{\mathrm{e}^{\left(-\sum_c^{N_s}\mathcal{N}_c\right)}}{n!} \prod_{i=1}^{n} \left(\sum_{c}^{N_s} \left[\mathcal{N}_c \cdot \mathcal{P}_c\left(\vec{x}_i|\vec{\lambda}\right)\right] \right). \tag{5.15}$$

External inputs can be used in the maximum-likelihood method to constrain a parameter θ with a known expected value μ and known uncertainty σ by multiplying a Gaussian $\mathcal{G}(\theta|\mu, \sigma)$ to the likelihood.

The maximum-likelihood fits performed in these analyses are usually implemented in the ROOT framework [104, 105] using the Minuit package [106] for the log-likelihood minimisation.

5.3 The sPlot technique

In the analysed data, the signal, events of interest, is contaminated by background events. A fraction of these background events usually cannot be removed from the data set by a data selection. To statistically unfold the signal properties from the data, the sPlot technique [90, 91] can be used. This approach relies on discriminating observables $\vec{m}$ (usually masses of the signal candidates) to unfold the signal component of the observables of interest $\vec{t}$ (e.g. the candidate's decay time) similar to a side-band subtraction. An important condition for the use of this technique is the independence of the discriminant and the observable of interest, so that the total PDF factorises

$$\mathcal{P}\big(\vec{m},\,\vec{t}\big) \;=\; \mathcal{P}(\vec{m}) \cdot \mathcal{P}\big(\vec{t}\big)\,. \tag{5.16}$$

Based on the PDFs $\mathcal{P}_c(\vec{m})$ for different data categories of the discriminating observable, which usually are estimated by an extended maximum-likelihood fit as described in Section 5.2, per-event weights are calculated. These so-called sWeights are defined by

$$w_c^{\mathcal{S}}(\vec{m}_i) \;=\; \frac{\sum_j^{N_s}\big[V_{cj}\cdot\mathcal{P}_j(\vec{m}_i)\big]}{\sum_k^{N_s}\left[\mathcal{N}_k\cdot\mathcal{P}_k(\vec{m}_i)\right]}\,, \tag{5.17}$$

with the covariance matrix V of the different components' yields

$$V_{cj}^{-1} \;=\; \sum_i^n \frac{\mathcal{P}_n(\vec{m}_i)\cdot\mathcal{P}_j(\vec{m}_i)}{\left(\sum_k^{N_s}\left[\mathcal{N}_k\cdot\mathcal{P}_k(\vec{m}_i)\right]\right)^2}\,. \tag{5.18}$$

By construction, the sum over all sWeights of a given category gives the category's event yield $\sum_i^n w_c^{\mathcal{S}}(\vec{m}_i) \;=\; \mathcal{N}_c\,.$

Using these sWeights in Equation 5.14, the contribution of an individual category can be projected

$$\mathcal{L}_c\big(\vec{\lambda}\big) \;=\; \sum_{i=1}^{n}\Big[w_c^{\mathcal{S}}\big(\vec{t}_i\big)\cdot\ln\mathcal{P}\big(\vec{t}_i|\vec{\lambda}\big)\Big]\,. \tag{5.19}$$

However, this modification of the likelihood affects the error estimation of the fitted parameters $\vec{\lambda}$ as the covariance matrix is calculated from the Hessian matrix of the likelihood. Often a simplified approach is sufficient to achieve good coverage of the fitted parameters. In this approach, the likelihood is scaled by a factor

$$\alpha \;=\; \frac{\sum_i^n w_c^{\mathcal{S}}(\vec{m}_i)}{\sum_i^n\left[w_c^{\mathcal{S}}(\vec{m}_i)\right]^2}\,. \tag{5.20}$$

Another approach, which also provides asymptotic correctness of the estimated parameter uncertainties, but is also computationally more expensive, is discussed in Reference [91].

Part II

Measurement Of The Oscillation Frequency Δm_s Using $B_s^0 \to D_s^- \pi^+$ Decays

The measurement of the B_s^0-meson oscillation frequency Δm_s using data taken by the LHCb experiment is presented in the following. For this measurement, the B_s^0 meson is reconstructed in the $B_s^0 \to D_s^- \pi^+$ decay mode, where the D_s^- candidate is reconstructed from three charged hadrons. This flavour-specific decay mode allows the B_s^0-meson flavour to be reconstructed directly from its decay products at the time of the decay. Together with an estimate of the B_s^0-meson flavour at production provided by the LHCb flavour tagging algorithms, this gives access to the meson's mixing state. On a sufficiently large data sample this can be used to extract the decay-time-dependent mixing probability and so the oscillation frequency. The used decay mode is very abundant, yielding good statistical precision of the measurement. Additionally, its long-lived final state consists of charged hadrons only, allowing the B_s^0 candidate to be fully reconstructed in the LHCb data.

For this analysis, the LHCb data set recorded during LHC Run2 (2015-2018) is used providing an integrated luminosity of $6\,\text{fb}^{-1}$. The additional $3\,\text{fb}^{-1}$ data set recorded during LHC Run1 (2011 & 2012) is not used in this analysis, as it has already partially been the subject of a previous analysis of the 2011 data [107], measuring $\Delta m_s = 17.768 \pm 0.023 \pm 0.006$, where the first uncertainty is statistical and the second systematic. Besides, the data-taking conditions changed between the two data-taking periods. Most prominently the centre-of-mass energy has been increased from 7 TeV to 13 TeV. Therefore, the increase in statistical precision is expected to be overshadowed by increased systematic uncertainties in a combined analysis of the Run2 and the 2012 data sets.

The analysis is divided into three parts. At first, a time-independent analysis is performed aiming to reconstruct and select a sample of $B_s^0 \to D_s^- \pi^+$ decays. This part of the analysis is presented in Chapter 6 and consists of the online reconstruction, an offline selection for background suppression, and a statistical separation of signal and the remaining backgrounds using the candidates' reconstructed masses as the discriminator. Then, the $B_s^0 \to D_s^- \pi^+$ sample provided by the time-independent part is used for a decay-time-dependent analysis as presented in Chapter 7. Here, the oscillation frequency Δm_s is extracted by a maximum-likelihood fit of the sample's decay-time distribution after investigating and calibrating the necessary inputs. Lastly, potential sources of systematic uncertainties are discussed and studied in detail in Chapter 8.

The presented analysis has been performed by an international group of researchers, including the author of this thesis. Since none of the studies performed in the course of this analysis can be reasonably discussed without context, the full analysis is presented as a whole regardless of which contributor has performed a specific study. The parts with major contributions from the author are discussed in detail, while other parts are briefly summarised. If some studies investigated mainly by other contributors need to be discussed on a more detailed level due to their importance, the corresponding contributors are explicitly acknowledged.

This measurement of the B_s^0-$\overline{B}_s^0$ oscillation frequency Δm_s has been published together with an updated LHCb-wide combination in 2022 by Nature Physics [1].

6 Extraction of the $B_s^0 \to D_s^- \pi^+$ signal

The decay-time-dependent measurement of the oscillation frequency Δm_s relies on a precise analysis of the properties of $B_s^0 \to D_s^- \pi^+$ decays. Hence, the signal properties have to be extracted from data samples diluted by background processes. To do so, the sPlot approach introduced in Section 5.3 is used in Section 6.3. Here, the invariant masses of the B_s^0-meson and D_s^--meson candidates are used as discriminants to extract signal weights based on models fitted to the distributions. The fit and the modelling of the mass distributions require the level of background contamination present in the data sample to be reduced first. The background consists of two major components. Background decays, that are either partially reconstructed or based on wrong mass hypotheses, are summarised as physics backgrounds. These wrongly reconstructed decays are suppressed in Section 6.2 by the application of explicit vetoes or by tighter selection criteria compared to the baseline reconstruction requirements. Combinatorial background, a random combination of unrelated tracks yielding fake decays, is suppressed by a multivariate classifier as discussed in Section 6.1.3. The analysed data samples and the simulated samples used to study various properties of both signal and backgrounds are introduced in Section 6.1, together with the preselections applied during the reconstruction and the data processing.

6.1 Analysed data samples

In the course of this analysis, both data recorded by the LHCb detector and simulated events are used. The simulated samples compromise a larger number of decay modes used to study and model both signal and backgrounds. Next to the $B_s^0 \to D_s^- \pi^+$ signal, the random combination of D_s^- with a π^+ is simulated, which is used for the calibration of the decay-time resolution in Section 7.1 and of a decay-time bias in Section 7.2. Additionally, several simulated background decays are used to model the invariant mass distributions in Section 6.3, including $B_s^0 \to D_s^{*-} \pi^+$, $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$, $B_s^0 \to D_s^{\mp} K^{\pm}$ and $B^0 \to D^{\mp} \pi^{\pm}$ decays. These samples are also used in Section 8.3.2 in a bootstrapping approach to produce data-like simulated samples with background contamination.

In total, 41 M simulated events are available for this analysis, including 18 M $B_s^0 \to D_s^- \pi^+$ generated signal events. Since the neural-network-based PID information is known to behave differently on data and simulation, the information is corrected using the PIDCalib package [83, 108]. The full selection, which will be discussed in the following, is applied to the simulated samples to incorporate possible effects from the selection process. If this is not the case, this is stated explicitly.

6.1.1 Data sample composition

The full data set corresponds to an integrated luminosity of $\int \mathcal{L} \mathrm{d}t = 6\,\mathrm{fb}^{-1}$ of proton-proton collisions recorded at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ in the years 2015-2018. The sample is categorised by the year of data taking. Although all data is recorded during the same data-taking period, small changes in the data-taking conditions can occur between the different years. Most prominent is a difference in the detector alignment affecting the reconstruction. The integrated luminosity recorded in the different years split by the magnet polarity is given in Table 6.1. In the following, the subsamples recorded with different magnet polarities are combined. This way possible systematic effects from spatial differences in the detector efficiencies should cancel, as the composition of the up and down polarities is almost even. In addition, the 2015 and 2016 samples are merged as there are no significant changes in the data-taking conditions present between the two years. This yields a sample divided into three subsamples of similar size with homogeneous data conditions each.

In data, the $D_{(s)}^-$ candidates are reconstructed in two final states and four different decay modes. The $\pi^- \pi^+ \pi^-$ final state is used and the $K^- K^+ \pi^-$ final state is divided into the non-resonant $(K^- K^+ \pi^-)_{\mathrm{NR}}$ and the two resonant $\phi \pi^-$ and $K^{*0} K^-$ modes.

In the preselection, the two $B^0_{(s)}$ and the four $D^-_{(s)}$ modes are treated equally as no PID information is used. Different sources and levels of background contamination are expected for the different decay modes.

Table 6.1: Integrated luminosity in fb^{-1} split by magnet polarity and the year of data taking.

Sample	2015	2016	2017	2018	Combined
Magnet up	0.141	0.805	0.833	1.142	2.921
Magnet down	0.187	0.860	0.877	1.048	2.972
Combined	0.328	1.665	1.710	2.190	5.893

6.1.2 Preselection

The data analysed in this measurement is preselected at different stages. At first, selections are applied during data taking by the trigger system, discussed in Section 4.3. The trigger requirements applied in this analysis are loose. No specific requirements at the level of the L0 hardware trigger are applied. At the level of the HLT1, displaced tracks with high transverse momentum are required. In the HLT2, this requirement is tightened demanding the presence of a displaced multi-track vertex with sizeable transverse momentum.

To save computational resources, the data recorded has to pass a set of loose preselections common to a variety of similar decay modes. The set used in this analysis is designed to select $B^0_{(s)} \to D^-_{(s)} h^+$ decays, where the $D^-_{(s)}$-meson candidate decays into three charged hadrons. A displaced secondary vertex (SV) formed by the $D^-_{(s)}$-meson candidate and a charged accompanying hadron is required. The three charged hadrons from the $D^-_{(s)}$ decay have to form a displaced tertiary vertex (TV).

All four final-state tracks are required to be well-reconstructed with criteria on the track momenta and the quality of the fitted tracks. No criteria based on particle identification (PID) information is used at this stage. The invariant mass from the combination of the three hadrons' four-vector is expected to lay within a $100\,\text{MeV}/c^2$ window around the known mass of either the D^- or D^-_s meson.

The assumed vertices have to be well-reconstructed as well. Significant separation from any PV as well as from each other is required for both SV and TV. Additionally, requirements on the fit quality of the vertices and the distance of closest approach of the particles forming the vertex are applied. Finally, the combined $D^-_{(s)}$-meson candidate's momentum is required to point to the SV, and the combined $B_{(s)}^0$ momentum to the primary vertex (PV). For better reproducibility, more details on the preselection are given Appendix A.1.

6.1.3 BDT-based suppression of combinatorial background

The main source of background in the preselected samples is combinatoric. This background is suppressed by a BDT trained in a data-driven approach. The training follows the strategy used in the measurement of *CP* violation in $B_s^0 \to D_s^\mp K^\pm$ decays using the LHCb Run1 data set [14]. However, the classifier is retrained to account for the different data-taking conditions in Run2 and a new cut point is optimised. While the procedure is briefly summarised in the following, more details can be found in the corresponding thesis [109].

The BDT training is performed in a fully data-driven approach, using data collected by the LHCb experiment in the years 2015 and 2016, to avoid possible implications from a simulation-data mismatch. For the training, a labelled data set including both background and signal is needed. Therefore, the data sample is split into two regions. The signal region is defined by the $m(D_s^- \pi^+) \in [5310, 5430]\,\mathrm{MeV}/c^2$ mass window around the known B_s^0 mass. As a background proxy the upper mass side-band $m(D_s^- \pi^+) \in [5445, 5800]\,\mathrm{MeV}/c^2$ is used.

To reduce contributions from specific physics backgrounds in the training samples, a simplified version of the selection discussed in Section 6.2 is applied. Further, the sPlot method is used to extract a statistically pure signal sample from the signal-proxy region, which still is polluted from the combinatorial background. To do so, an unbinned maximum-likelihood fit to the $m(D_s^- \pi^+)$ mass is performed. The strategy here is similar to the nominal mass fit later presented in Section 6.3, but the fit is performed only in one dimension and using a simplified model. The combinatorial background is described by PDF consisting of an exponential and a constant, the signal component by a double Crystal Ball PDF.

To keep the full data sample available for the following analysis, the data sample is split randomly into halves. Two BDTs are trained. Each BDT is trained on one half of the sample and applied to the other. Next to keeping the full sample available, this brings a reduction of overtraining effects and the possibility to cross-validate the BDTs. The mass fit in the signal region to extract sWeights is performed on both sample halves individually.

The set of training features is chosen similarly to the Run1 $B_s^0 \to D_s^\mp K^\pm$ analysis [14]. The full list of training features, together with more details on the procedure is given in the thesis [109]. In addition to features used in the preselection, the radial flight distance of the B_s^0 and D_s candidates and the angle $\theta_{B_s^0}$ between the B_s^0 candidate's and the accompanying pion's direction is used. The vertex qualities are evaluated by decay-tree fits with or without constraining the B_s^0 candidate to point to the associated PV.

The optimal cut point for the combinatorial background suppression is found by optimising a figure-of-merit (FOM). As the BDT selection developed here is also designed to be used in the CP violation measurement in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays, the FOM taken from the previous $B_s^0 \to D_s^{\mp} K^{\pm}$ analysis [14] is used. It is constructed from the signal efficiency ε_{sig} and the signal significance P_{sig}

$$\text{FOM} = \varepsilon_{\text{sig}} \cdot P_{\text{sig}} = \varepsilon_{\text{sig}} \frac{N_{\text{sig}}}{\sqrt{N_{\text{sig}} + N_{\text{bkg}}}} , \tag{6.1}$$

with fitted signal yields N_{sig} and background yields N_{bkg}. These yields are determined for the different cut points by a multidimensional fit to the $m(D_s^- \pi^+)$ and $m(h^{\mp} h^+ h^-)$ masses as later described in Section 6.3. The signal efficiency is derived from the fitted signal yield with respect to the signal yield before a BDT selection. The FOM shows a plateau in the cut range between 0.4 and 0.6 with a maximum at 0.475, which is chosen as the cut applied to the data.

6.2 Suppression of physics backgrounds

After the preselection and the suppression of combinatorial backgrounds, there are still various backgrounds from heavy hadron decays left. These need to be further suppressed as they are challenging to be modelled in Section 6.3 and could possibly dilute the sample's properties. The physical backgrounds can be divided into two classes. Partially reconstructed decays yield candidates, which are reconstructed only from a subset of the true decay's final-state particles. Misidentified background is originating from candidates reconstructed assuming a wrong mass hypothesis of at least one final-state particle. Usually, hadron decays into a different final state but with a topology similar to the signal cause this kind of background, as these decays easily pass the preselection. A condensed summary of these selections is given in the appendix in Table A.2.

6.2.1 PID selection

The PID information is a powerful tool to suppress misidentified backgrounds as it gives an estimate of the final-state particles' species and the risk of applying a wrong mass hypothesis is minimised. Sophisticated observables are constructed to give an estimate of the particle species. The PROBNN_X variables assign a probability estimate to be a particle of the species X to every track. This probability is estimated by a neural network evaluating the PID information provided by the different detector components. The $\mathrm{PID}_X = \ln \mathcal{L}(X) - \ln \mathcal{L}(\pi)$ variables are defined by the difference of the log-likelihood to be a particle X and to be a pion. Here the likelihood $\mathcal{L}$ again is based on the PID information coming from the detector. In this analysis, the signal is selected by applying criteria, which can be specific to a D_s final state or general, to the PID of the final-state particles.

To provide reliable PID estimates every final-state track is required to have information from the RICH system available. Additionally, an absence of the track in the muon system is required to suppress backgrounds of misidentified muons. The desired $B_s^0 \to D_s^- \pi^+$ signal is then selected by applying PID criteria to the final-state tracks. The accompanying pion is required to fulfil a loose kaon suppression $\mathrm{PID}_K < 0$, removing $B_s^0 \to D_s^\mp K^\pm$ decays originating from a $\pi^+ \to K^+$ misidentification.

Backgrounds in the D_s sample do also propagate to the B_s^0 sample. These are dominated by decays of $\overline{\Lambda}_c^-$, D^-, D^0 and D_s^- into final states consisting of pions, kaons and protons reconstructed under the wrong mass hypothesis. The $\pi \leftrightarrow K$ and the $p \to K$ misidentifications can be suppressed by PID_K criteria, the rarer $p \to \pi$ misidentification by a loose PID_p criterion. The requirements applied to the D_s children have to be specific to the D_s-decay mode due to different sources of backgrounds and different levels of contamination.

In the $D_s^- \to \phi\pi^-$ mode, the very narrow $\phi(1020)$ resonance is selected in a small $20\,\mathrm{MeV}/c^2$ window around its known mass. The cleanness of the reconstructed ϕ sample propagates into a very clean D_s and B_s^0 spectrum. Hence, only a loose $\mathrm{PID}_K > -2$ criterion is applied to both final-state kaons.

The $K^{*0}(892)$ resonance is wider requiring to select a $50\,\mathrm{MeV}/c^2$ window around the resonance to select the $D_s^- \to K^{*0}K^-$ mode. Therefore, this sample provides a medium purity with some contributions from misidentified backgrounds. These are suppressed by a tighter $\mathrm{PID}_K > 5$ requirement on the kaon, that does not originate from the K^{*0} resonance.

The remaining $D_s^- \to K^-K^+\pi^-$ phase is summarised as the $(K^-K^+\pi^-)_{\mathrm{NR}}$ mode. This mode does not significantly benefit from the presence of resonances. Therefore, to suppress the significant background contamination the tighter $\mathrm{PID}_K > 5$ requirement is applied to both kaons and an additional $\mathrm{PID}_K < 10$ requirement to the pion from the D_s decay.

Similarly, the selection of the $D_s^- \to \pi^-\pi^+\pi^-$ mode does not exploit any resonance. Further, the pion as the lightest hadron is very abundant. Tight requirements $\text{PID}_K < 2$ and $\text{PID}_p < 5$ are applied to all D_s children.

6.2.2 Variation of the mass hypotheses and explicit vetoes

As the possible separation from the PID information is limited and as some of the backgrounds are introduced by very abundant decays, the previously presented PID selection alone is not sufficient to suppress all physics backgrounds. The dominant sources of remaining backgrounds originate from b-hadron decays without an intermediate D_s meson. These include decays to other intermediate charm hadrons (namely D^-, D^0 and $\overline{\Lambda}_c^-$), but also to charmonium resonances or charmless states. The latter ones are of special importance as these are necessarily not subject to a wrong mass hypothesis in the final state. Hence, these can not be distinguished by the sPlot method using the $m\left((h^-h^+h^-)_{D_s^-}\pi^+\right)$ mass spectrum.

The previous analysis of $B_s^0 \to D_s^\mp K^\pm$ decays in Run1 [14] has made efficient use of the finite lifetime of the D_s meson to reduce contributions from charmless backgrounds. The backgrounds including other charm hadrons than the D_s meson have been suppressed by specific vetoes. However, in this analysis decay-time biasing effects are observed from the application of requirements on the flight distance $\text{FD} > 0$ of the D_s meson, its significance $\chi^2_{\text{FD}} > 2$ (> 9 for $D_s \to \pi^-\pi^+\pi^-$), and the reconstructed D_s-meson decay time $t_{D_s^-} > 0$. While only a small bias of $0.27 \pm 0.10\,\text{fs}$ is observed on simulated samples after the application of the preselection and the BDT selection, this bias increases to $-1.42 \pm 0.11\,\text{fs}$ after applying the requirements on the D_s flight distance and decay time. This bias directly translates into a biased oscillation frequency Δm_s. Therefore, these requirements are dropped and a variety of explicit vetoes is studied. A potential $D_s^- \to K^-\pi^+\pi^-$ mode, which only provides small signal statistics, is completely dropped as this mode shows a level of background contamination, which is not manageable without the decay-time biasing selection.

The dominant misidentified background in the $D_s^- \to K^-K^+\pi^-$ modes is the $B^0 \to D^-(\to K^+\pi^-\pi^-)\pi^+$ where one of the pions is misidentified as a kaon. This background is suppressed by a tightened $\text{PID}_K > 10$ on the potentially misidentified kaon if the $m(K^-K^+\pi^-)_{K^-\to\pi^-}$ mass under application of the pion-mass hypothesis is within a $30\,\text{MeV}/c^2$ window around the known D^- mass. Similarly, $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^-\pi^+$ decays where the antiproton from a $\overline{\Lambda}_c^- \to \overline{p}K^+\pi^-$ is misidentified as a kaon is suppressed. Here, the tighter $\text{PID}_K - \text{PID}_p > 5$ criterion is applied to the possibly misidentified kaon for candidates with $m(K^-K^+\pi^-)_{K^-\to\overline{p}}$ masses reconstructed under a proton-mass hypothesis laying in a $30\,\text{MeV}/c^2$ window around the known $\overline{\Lambda}_c^-$ mass. The possible $\overline{p} \to \pi^-$ misidentification in $\overline{\Lambda}_c^- \to K^+K^-\overline{p}$ decays is vetoed by requiring $\mathrm{PID}_p < 10$ for the pion. Contributions from D^0 decays also need to be considered. A prominent source of this contribution are $B^0 \to D^{*-}(\to D^0\pi^-)\pi^+$ decays with $D^0 \to K^+K^-$, where no misidentification occurs. This process is removed by requiring the invariant mass of both kaons to be lower than $1800\,\mathrm{MeV}/c^2$. The $\overline{D}^0 \to K^-\pi^+$ decay is very abundant. This mode is suppressed by removing candidates with an invariant mass of the accompanying pion and the kaon within a $30\,\mathrm{MeV}/c^2$ window around the D^0 mass. The decay of a D^0 meson into a pion and kaon assigned to the D_s^- is not considered, as the invariant mass of this combination together with a residual kaon is significantly higher than the D_s^- mass.

In the $D_s^- \to \pi^-\pi^+\pi^-$ mode contributions from $\overline{\Lambda}_c^- \to \overline{p}K^+\pi^-$ decays occur due to simultaneous $\overline{p} \to \pi^-$ and $K^+ \to \pi^+$ misidentification. This background is suppressed by $\mathrm{PID}_K < 2$ and $\mathrm{PID}_p < 5$ requirements on the pions. Contributions from $D^0 \to \pi^-\pi^+$ decays are removed by a $m(\pi^+\pi^-) < 1700$ MeV/c^2 threshold. Assuming a $K^+ \to \pi^+$ misidentification for the accompanying hadron two backgrounds appear in the spectrum of the invariant mass in combination with two D_s^- children. The contribution of $D^+ \to K^+\pi^-\pi^+$ decays is vanished by the already applied PID criteria. The $D^{*-} \to D^0(\to K^+\pi^-)\pi^-$ background is handled by vetoing the intermediate D^0 state formed by a pion and the misidentified accompanying hadron. Through this veto candidates with an invariant $m(\pi^+\pi^-)_{\pi^+ \to K^+}$ mass within $30\,\mathrm{MeV}/c^2$ around the known D^0 mass are rejected.

In the previous analysis, the very short-lived charmonium resonances have been effectively rejected by the decay-time biasing requirements based on the finite D_s^- decay-time. After the removal of these criteria, significant contamination from $B^0_{(s)} \to J/\psi X$ decays are observed. This background originates from the $l^+l^- \to \pi^+\pi^-$ misidentification of the very abundant $J/\psi \to l^+l^-$ decays. A loose $\mathrm{PID}_e < 5$ requirement applied to the accompanying pion and to all oppositely charged pion is sufficient to reject all $J/\psi \to e^+e^-$ contributions. The $\mu^\pm \to \pi^\pm$ misidentification is suppressed by requiring the final-state tracks' absence in the muon system.

6.2.3 Investigation of charmless backgrounds

The presence of remaining charmless $B_s^0 \to h^-h^+h^-h^+$ backgrounds is extensively studied in the $m(h^\mp h^+h^-)$ mass sideband. The sideband is contaminated by residual misidentified $\overline{\Lambda}_c^-$ and D^- decays. As the lower-mass sideband additionally contains contributions from partially reconstructed and radiative D_s^- decays, as well as from genuine D^- decays, only the upper sideband $m(h^\mp h^+h^-) \in [2015, 2070]\,\mathrm{MeV}/c^2$ is used. As constraining the mass of the three hadrons to a non-genuine D_s^- state distorts the $m(D_s^-\pi^+)$ estimate, the unconstrained $m(h^-h^+h^-h^+)$ mass is also studied. The $m(D_s^-\pi^+)$ and the $m(h^-h^+h^-h^+)$ mass distributions taken from the upper-mass sidebands of the four D_s modes after application of the previously discussed selection is compared in Figure 6.1. The effects of the decay-time biasing selection are also shown. While no peaking structure can be observed in the $D_s \to \phi\pi^-$ mode, different structures are present in the other modes.

In the $m(K^-K^+\pi^-\pi^+)$ distributions of the $K^{*0}K^-$ and $(K^-K^+\pi^-)_{\mathrm{NR}}$ modes, a peak at the nominal B_s^0 mass is visible, which is shifted by the D_s^--mass constraint to lower masses in the $m(D_s^-\pi^+)$ spectra. This structure appears unaffected by the decay-time biasing constraint to a finite lifetime of the D_s state. Due to these characteristics, this structure is expected to originate from residual long-lived backgrounds. This hypothesis is supported by a study of simulated $B^0 \to D^\mp\pi^\pm$ decays. After the full selection is applied, the $B^0 \to D^\mp\pi^\pm$ mass distributions based on the upper $m(h^\mp h^+ h^-)$ sideband yield the observed structure. Further, the simulated $B^0 \to D^\mp\pi^\pm$ samples show the same behaviour in respect of a D_s^--mass constraint and the decay-time biasing selection. Hence, no charmless $B_s^0 \to h^-h^+h^-h^+$ backgrounds are found in the $D_s^- \to K^-K^+\pi^-$ modes.

The situation is different for the $D_s^- \to \pi^-\pi^+\pi^-$ mode. Here, a peaking structure is observed at the nominal B^0 mass in the $m(\pi^-\pi^+\pi^-\pi^+)$ distribution. This structure is significantly suppressed by the biasing selection and completely removed by the even tighter $\chi^2_{\mathrm{FD}} > 9$ requirement on the D_s candidate. Further, the structure is not only shifted to lower masses in the constrained $m(D_s^-\pi^+)$ mass but also blurred. Therefore, this structure is expected to originate from charmless $B_s^0 \to \pi^-\pi^+\pi^-\pi^+$ decays. However, the shift introduced by the mass constraint moves the structure below the $m(D_s^-\pi^+)$ threshold defined by the mass fit range in Section 6.3. Overall, contributions from charmless backgrounds are considered negligible after the full selection is applied.

Lastly, a restriction of the data sets to the ranges used in the fits performed in the course of this analysis is applied. The invariant $m(D_s^-\pi^+)$ and $m(h^\mp h^+ h^-)$ masses are constrained to the fit ranges defined in Section 6.3, the decay time $t_{B_s^0}$ and its uncertainty estimate $\sigma(t_{B_s^0})$ to the ranges defined in Section 7.4.

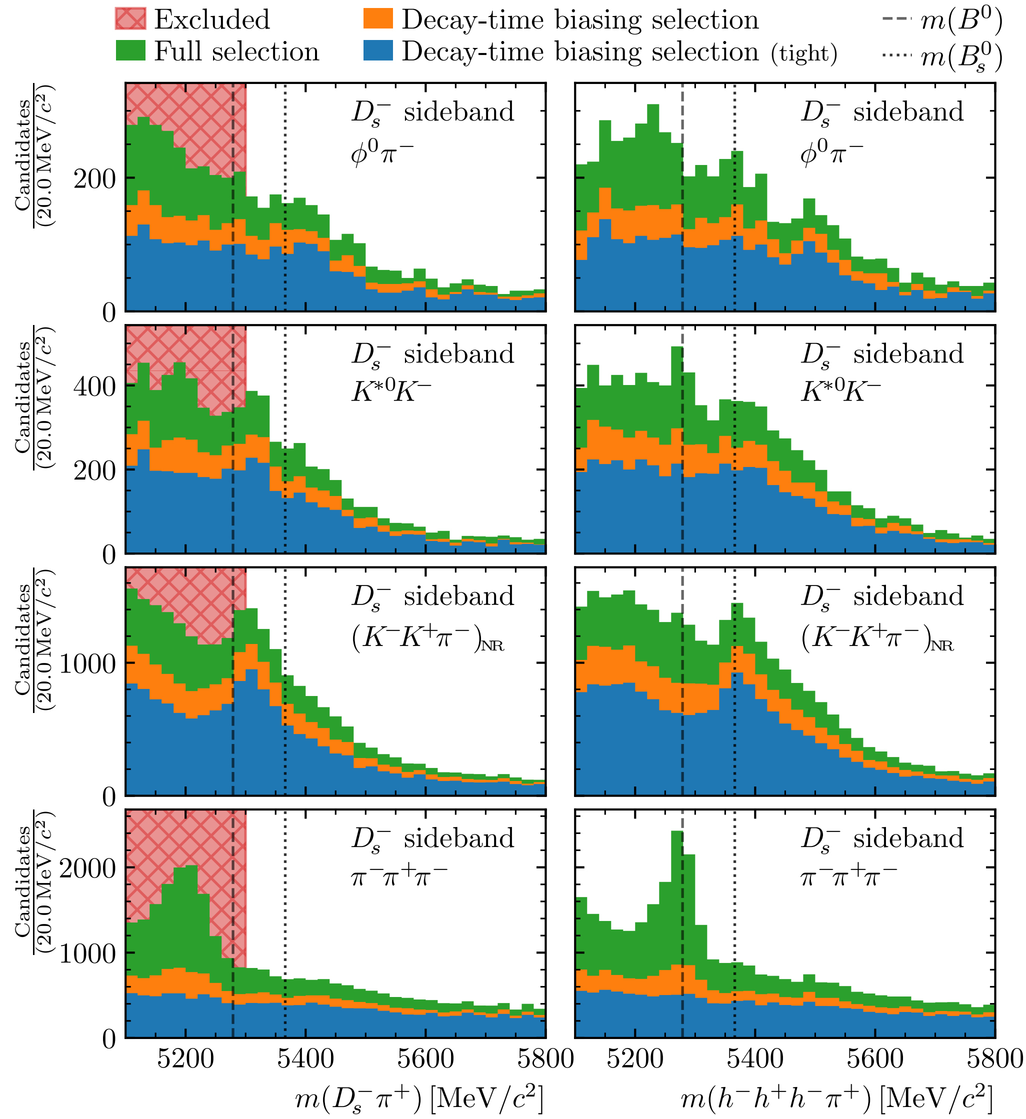


Figure 6.1: Distributions of the reconstructed B_s^0 mass taken from D_s^- sideband samples with different selections applied. The left side shows the $m(D_s^-\pi^+)$ distributions with a D_s^- mass constraint applied, the right side the unconstrained $m(h^-h^+h^-h^+)$ distribution, each split in the different decay modes $D_s^- \to \phi\pi^-$, $D_s^- \to K^{*0}K^-$, $D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$ and $D_s^- \to \pi^-\pi^+\pi^-$ (from top to bottom). The distributions after the full offline selection (green), with an additional, decay-time biasing selection (orange) and even tighter requirements (blue) are shown.

6.3 Multidimensional mass fit

To distinguish signal decays and residual background, the sPlot method introduced in Section 5.3 is utilised. The good mass resolution of the LHCb experiment allows the reconstructed mass to be used as a discriminant. Both the reconstructed B_s^0-candidates' and D_s^--candidates' masses are fitted simultaneously in a multidimensional, unbinned, extended maximum-likelihood fit, also referred to as MDFit. In this fit, the different D_s^--decay modes and data-taking periods are modelled individually with a few parameters shared among the subsamples, while the two magnet polarities are merged. The optimal parametrisation of the mass distributions has been extensively been studied [110]. The fitted model and the mass distributions are presented in Figure 6.2 for the full data set.

The signal component is described by the sum of a double-sided Hypatia $\mathcal{H}(m\,|\,\vec{p}^{\,H})$ and a Johnson SU $\mathcal{J}(m\,|\,\vec{p}^{\,J})$ function [111, 112] for both the B_s^0 and D_s^- masses

$$\begin{aligned}\mathcal{S}\Big(m_{D_s^-\pi^+}\,|\,\vec{p}^{\,H}_{B_s^0}, \vec{p}^{\,J}_{B_s^0}, f^S_{B_s^0}\Big) &= f^S_{B_s^0}\cdot\mathcal{H}\Big(m_{D_s^-\pi^+}\,|\,\vec{p}^{\,H}_{B_s^0}\Big) \\ &\quad+ \Big(1-f^S_{B_s^0}\Big)\cdot\mathcal{J}\Big(m_{D_s^-\pi^+}\,|\,\vec{p}^{\,J}_{B_s^0}\Big)\,,\end{aligned} \tag{6.2}$$

$$\begin{aligned}\mathcal{S}\Big(m_{h^-h^+h^-}\,|\,\vec{p}^{\,H}_{D_s^-}, \vec{p}^{\,J}_{D_s^-}, f^S_{D_s^-}\Big) &= f^S_{D_s^-}\cdot\mathcal{H}\Big(m_{h^-h^+h^-}\,|\,\vec{p}^{\,H}_{D_s^-}\Big) \\ &\quad+ \Big(1-f^S_{D_s^-}\Big)\cdot\mathcal{J}\Big(m_{h^-h^+h^-}\,|\,\vec{p}^{\,J}_{D_s^-}\Big)\,.\end{aligned} \tag{6.3}$$

Here, the parameter $f^S_{B_s^0/D_s^-}$ modulates the fraction of the two components. These fractions are fixed to values found in the simulation. The parameter sets $\vec{p}^{\,H/J}_{B_s^0/D_s^-}$ contain several shape parameters, which are also fixed for each subsample individually to values extracted from simulations. These include the Gaussian widths $\sigma^J_{B_s^0/D_s^-}$ of the Johnson SU for each mass and subsample. The means $\mu_{B_s^0/D_s^-}$ are shared among the two components and the different D_s^--decay modes, but are separated between the data-taking periods to account for effects from changed reconstruction. The overall signal PDF is defined by the product of both dimensions

$$\mathcal{S}(\vec{m}\,|\,\vec{p}^{\,S}) = \mathcal{S}\Big(m_{D_s^-\pi^+}\,|\,\vec{p}^{\,S}_{B_s^0}\Big)\cdot\mathcal{S}\Big(m_{h^-h^+h^-}\,|\,\vec{p}^{\,S}_{D_s^-}\Big)\,. \tag{6.4}$$

The shape of the combinatorial background is studied in the upper-mass sideband $m\,(D_s^-\pi^+) \in [5600, 6800]\,\mathrm{MeV}/c^2$. In the B_s^0 mass $m(D_s^-\pi^+)$ the combinatorial background is described by the sum of two exponential functions

$$\mathcal{C}\Big(m_{D_s^-\pi^+}\,|\,c^1_{B_s^0}, c^2_{B_s^0}, f^C_{B_s^0}\Big) = f^C_{B_s^0}\cdot\mathrm{e}^{c^1_{B_s^0}\cdot m_{D_s^-\pi^+}} + \Big(1-f^C_{B_s^0}\Big)\cdot\mathrm{e}^{c^2_{B_s^0}\cdot m_{D_s^-\pi^+}}\,. \tag{6.5}$$

While the smaller slope $c^2_{B_s^0}$ is fixed to the value found in the sideband, the larger slope $c^1_{B_s^0}$ and the fraction $f^C_{B_s^0}$ are floating in the fit. The combinatorial background can

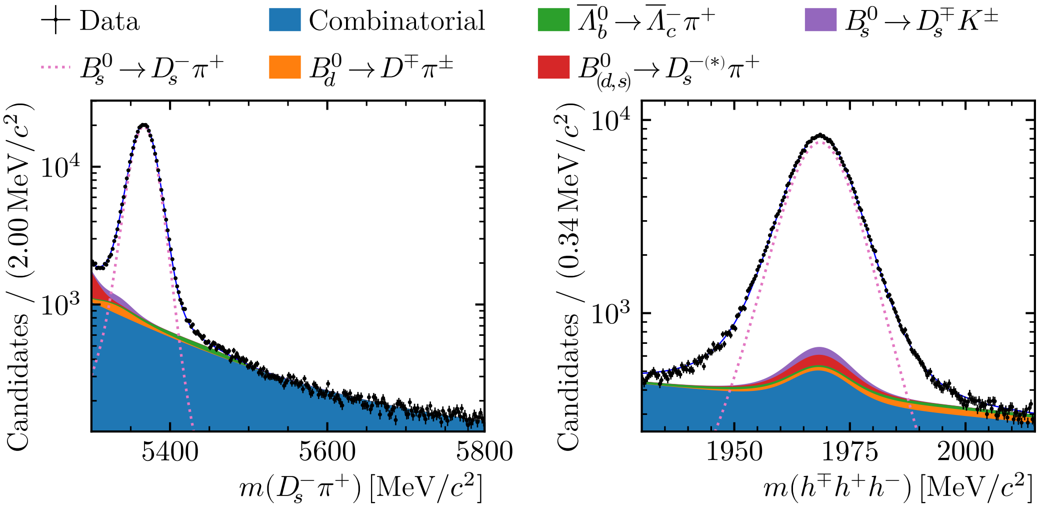


Figure 6.2: Fitted mass distributions of the $B_s^0 \to D_s^- \pi^+$ candidates, on the left the $m(D_s^- \pi^+)$, on the right the $m(h^\mp h^+ h^-)$ spectrum. The data distributions and the fitted PDFs shown here are merged among all subsamples.

originate from the random combination of four final-state tracks or the combination of a true D_s^- decay and a random pion. Hence, the combinatorial background is modelled by the sum of an exponential function and the signal component

$$\mathcal{C}\left(m_{h^-h^+h^-} \,|\, c_{D_s^-},\, f^C_{D_s^-},\, \vec{p}^{\,S}_{D_s^-}\right) = f^C_{D_s^-} \cdot \mathrm{e}^{c_{D_s^-} \cdot m_{h^-h^+h^-}} + \left(1 - f^C_{D_s^-}\right) \cdot \mathcal{S}\left(m_{h^-h^+h^-} \,|\, \vec{p}^{\,S}_{D_s^-}\right) . \tag{6.6}$$

in the D_s mass $m(h^\mp h^+ h^-)$, where the parameter set $\vec{p}^{\,S}_{D_s^-}$ is shared with the signal PDF. Again, the product of both dimensions $\mathcal{C}(\vec{m})$ defines the total PDF for the combinatorial background. A small contribution of physics backgrounds also passes the selection performed in Section 6.2. The $B^0 \to D_s^- \pi^+$ background shares a common final state with the signal and is described by the signal PDF shifted by the known mass difference $m_{B_s^0} - m_{B^0} = 86.8$ MeV/c^2 in the $m(D_s^- \pi^+)$ mass

$$\mathcal{B}_{B^0 \to D_s^- \pi^+}(\vec{m}) = \mathcal{S}\left(m_{D_s^- \pi^+} + 86.8 \text{ MeV/c}^2,\, m_{h^-h^+h^-}\right) . \tag{6.7}$$

The partially reconstructed $B_s^0 \to D_s^{*-} \pi^+$ and the misidentified $B_s^0 \to D_s^\mp K^\pm$ backgrounds contribute with true D_s^- decays to the $m(h^\mp h^+ h^-)$ spectrum and are hence described by the signal PDF in this dimension, too. However, both are modelled by simulation-based templates in the $m(D_s^- \pi^+)$ mass. The $B^0 \to D^\mp \pi^\pm$ and the $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ backgrounds are modelled by templates in both dimensions. The total PDF used in the fit is the sum of the Signal PDF and all aforementioned components extended with a set of yield parameters $\vec{N}$

$$\begin{aligned}\mathcal{P}(\vec{m}) \;=\; & N_{\mathrm{sig}} \cdot \mathcal{S}(\vec{m}) + N_{\mathrm{comb}} \cdot \mathcal{C}(\vec{m}) + N_{B_s^0 \to D_s^\mp K^\pm} \cdot \mathcal{B}_{B_s^0 \to D_s^\mp K^\pm}(\vec{m}) \\ & + N_{\mathrm{low}} \cdot \Big(f_{\mathrm{low}} \cdot \mathcal{B}_{B^0 \to D_s^- \pi^+}(\vec{m}) + (1 - f_{\mathrm{low}}) \cdot \mathcal{B}_{B_s^0 \to D_s^{*-} \pi^+}(\vec{m}) \Big) \\ & + N_{B^0 \to D^\mp \pi^\pm} \cdot \mathcal{B}_{B^0 \to D^\mp \pi^\pm}(\vec{m}) + N_{\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+} \cdot \mathcal{B}_{\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+}(\vec{m}) \;. \end{aligned} \tag{6.8}$$

In this parametrisation, the low-mass backgrounds $B^0 \to D^\mp \pi^\pm$ and $B_s^0 \to D_s^{*-} \pi^+$ share a yield, as the fit stability benefits from this modification in the pseudo-experiment studies discussed in Section 8.2.1. The fit shows only minor sensitivity to the fraction f_{low} of the two components. Hence, it is fixed to 0.5 and the effect is studied in the systematics Chapter 8. Only the signal, the combinatorial background and the low-mass background yields are floated in the fit. The fitted yields for the individual subsamples are listed in Table 6.2, while the fitted shape parameters of the signal and the combinatorial component can be found in the appendix in Table A.3. The yields of the remaining backgrounds are constrained based on the selection efficiencies extracted on simulation. The $B_s^0 \to D_s^\mp K^\pm$ yield is coupled to the signal yield by the ratio of the selection efficiencies and the known branching ratios. The $B^0 \to D^\mp \pi^\pm$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ yields are fixed to efficiency corrected values extracted from a fit to $B^0 \to D^\mp \pi^\pm$ data and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ simulation. The distributions of the two reconstructed masses are visualised in Figure 6.2 together with the fitted model. In total, 378663 ± 670 $B_s^0 \to D_s^- \pi^+$ signal decays are extracted.

Table 6.2: The three fitted yields split among the individual subsamples. The yields of further background components are constrained as described in the text. The total signal yield combining all subsamples is $N_{\mathrm{sig}} = 378\,663 \pm 670$.

Parameter	Year	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$	$D_s^- \to \pi^-\pi^+\pi^-$
N_{sig}	2015 & 16	$49\,963 \pm 232$	$36\,142 \pm 197$	$22\,944 \pm 161$	$23\,378 \pm 165$
	2017	$43\,372 \pm 216$	$31\,357 \pm 183$	$19\,902 \pm 149$	$20\,142 \pm 152$
	2018	$49\,700 \pm 232$	$36\,384 \pm 197$	$22\,612 \pm 159$	$22\,765 \pm 162$
N_{comb}	2015 & 16	3444 ± 86	3301 ± 79	6885 ± 104	$11\,734 \pm 127$
	2017	2760 ± 77	2772 ± 71	5589 ± 94	9208 ± 112
	2018	3491 ± 88	3111 ± 75	6817 ± 103	$10\,897 \pm 122$
N_{low}	2015 & 16	515 ± 46	362 ± 39	195 ± 34	256 ± 39
	2017	397 ± 41	274 ± 33	139 ± 29	249 ± 36
	2018	360 ± 43	477 ± 38	202 ± 33	186 ± 37

Redefinition of the sWeighting PDF

The sPlot approach requires independence of the discriminant and the variable of interest. This assumption does not hold perfectly for the mass-fit and decay-time-fit observables. Therefore, the resulting effects are extensively studied in Section 8.3. In the course of these studies, it is observed, that here the effect from neglected correlations, in the discrimination of signal from a combined background component, shows a significant effect on the measured oscillation frequency Δm_s. The resulting bias is significantly reduced by the usage of a combined background component, that composes the individually modelled background components. Therefore, sWeights are extracted from a fit to a modified PDF, where all previously floating parameters are fixed to the fitted values. Only the signal yield N_{sig} and the combined background yield N_{bkg} are floating. The fractions of the background components are fixed based on the sum N'_{bkg} of the individual background yields in the first fit. The modified PDF is defined by

$$\begin{aligned}\mathcal{P}'(\vec{m}) \;=\; & N_{\text{sig}} \cdot \mathcal{S}(\vec{m}) + N_{\text{bkg}} \cdot \left(\frac{N_{\text{comb}}}{N'_{\text{bkg}}} \cdot \mathcal{C}(\vec{m}) + \frac{N_{B_s^0 \to D_s^\mp K^\pm}}{N'_{\text{bkg}}} \cdot \mathcal{B}_{B_s^0 \to D_s^\mp K^\pm}(\vec{m}) \right. \\ & \quad + \frac{N_{\text{low}}}{N'_{\text{bkg}}} \cdot \Big(f_{\text{low}} \cdot \mathcal{B}_{B^0 \to D_s^- \pi^+}(\vec{m}) + (1 - f_{\text{low}}) \cdot \mathcal{B}_{B_s^0 \to D_s^{*-} \pi^+}(\vec{m}) \Big) \\ & \quad \left. + \frac{N_{B^0 \to D^\mp \pi^\pm}}{N'_{\text{bkg}}} \cdot \mathcal{B}_{B^0 \to D^\mp \pi^\pm}(\vec{m}) + \frac{N_{\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+}}{N'_{\text{bkg}}} \cdot \mathcal{B}_{\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+}(\vec{m}) \right) . \end{aligned} \tag{6.9}$$

The residual effects, introduced by neglected correlations in the sWeighting approach, are investigated in Section 8.3, while the selected and sWeighted data can be used for the decay-time-dependent measurement in the following Chapter 7.

7 Decay-time-dependent analysis of $B_s^0 \to D_s^- \pi^+$ decays

The oscillation frequency Δm_s is measured by extracting the decay-time-dependent decay rates defined by Equation 3.41 from data. A statistically clean sample of $B_s^0 \to D_s^- \pi^+$ decays is provided by the previously in Chapter 6 discussed decay-time-independent analysis. The fitted PDF describing the decay rates and incorporating experimental effects is discussed in Section 7.4. While the flavour-specific decay mode $B_s^0 \to D_s^- \pi^+$ provides by itself information on the B_s^0 flavour at decay, the flavour at production is estimated by flavour tagging algorithms introduced in Section 4.5. These algorithms are set up for usage in this analysis in Section 7.3. The precision necessary for this measurement requires a careful treatment of the decay time and its uncertainty estimate, which are calibrated in Section 7.2 and Section 7.1, respectively.

7.1 Calibration of the decay-time uncertainty estimate

The LHCb detector provides an excellent decay-time resolution of about 50 fs. However, the fast B_s^0-$\overline{B}{}_s^0$ oscillation shows a length of approximately 350 fs. Thus, resolution effects play a significant role and need to be properly accounted for in Section 7.4 for the decay-time fit. The finite resolution is incorporated into the decay-time PDF by convolution with a Gaussian resolution function

$$\mathcal{R}(t) = \mathcal{G}(t \mid \mu = 0, \sigma_t) \tag{7.1}$$

with decay-time uncertainty σ_t. The track and vertex reconstruction provides a per-event estimate δt for the uncertainty of the reconstructed decay time t, which yields a better precision than the usage of an average resolution. For the usage in this analysis, this estimate is calibrated as discussed in detail in the thesis [113] and briefly summarised in the following. The decay-time uncertainty estimate is calibrated in a data-driven approach using so-called prompt $D_s^- \pi^+$ candidates. The portability of the calibration to the signal channel is then studied using simulation.

Prompt candidates are constructed by combining unrelated D_s^- decays and π^+ tracks from a common PV yielding fake decays with the same final state as the $B_s^0 \to D_s^- \pi^+$ signal. These candidates are expected to have a vanishing decay time as they do not originate from the decay of long-lived particles. Hence, the reconstructed decay time of the prompt candidates is solely caused by resolution effects.

The prompt candidates are processed similarly to the signal decay with small variations only. A similar preselection compared to Section 6.2 is applied, where only the requirement on the B_s^0-decay time is omitted. Further, the BDT selection, aiming to remove fake-B_s^0 decays, is dropped. The D_s^- decay is reconstructed in the $\phi\pi^-$ mode as it is both the cleanest and the most abundant mode. The calibration is proven to be consistent among the different D_s^--decay modes in simulation. Combinatorial background in the D_s^- sample is subtracted using the sPlot technique based on a fit to the $m(h^{\mp}h^{+}h^{-})$ distribution.

For the calibration, the prompt data sample is divided into ten bins of the decay-time estimate. In each bin, the true decay-time uncertainty σ_t is extracted by a fit to the reconstructed decay-time distribution. Here, the distribution of the decay-time uncertainty estimate is modelled by a triple Gaussian

$$\begin{aligned} \mathcal{P}_t(t) \,=\, & f_1 \cdot \mathcal{G}(t \,|\, \mu,\, \sigma_1) + (1 - f_1)\, f_2 \cdot \mathcal{G}(t \,|\, \mu,\, \sigma_2) \\ & + (1 - f_1)\,(1 - f_2) \cdot \mathcal{G}(t \,|\, \mu,\, \sigma_3) \end{aligned} \tag{7.2}$$

with a shared mean μ, individual width σ_i and relative fractions f_i. To avoid contamination from long-lived decays, the positive tail of the distribution is discarded in the $[-500\,\text{fs},\, \langle\delta t\rangle]$ fit range. The calibration maps the uncertainty estimate to its true uncertainty using a linear function

$$\sigma_t\,(\delta t) \,=\, p_0 + p_1 \cdot \delta t\,. \tag{7.3}$$

The offset p_0 and the scale factor p_1 are estimated by a fit of the calibration function to the set $(\sigma_{\text{eff}},\, \langle\delta t\rangle)_i$ of effective measured resolution and average uncertainty estimate in each bin. The effective resolution is defined by

$$\sigma_{\text{eff}} \,=\, \sqrt{-2/\Delta m_s^2 \cdot \ln D} \tag{7.4}$$

with a dilution factor

$$D \,=\, \sum_i f_i \cdot \mathrm{e}^{\sigma_i^2 \cdot \Delta m_s^2/2} \tag{7.5}$$

digesting the individual Gaussian widths. The calibration is estimated for each data-taking period individually. The fitted calibration parameters are given in Table 7.1. In Section 8.2, systematic effects from the choice of the resolution model are studied, using two alternative parametrisations considering either the narrowest or widest Gaussian only.

Table 7.1: Calibration parameters of the decay-time uncertainty estimate and average uncertainty estimate extracted from prompt candidates for the different data-taking periods.

Sample	p_0 [fs]	p_1	δt [fs]
2015 & 16	8.4 ± 2.6	1.002 ± 0.059	46.97
2017	6.1 ± 2.5	1.048 ± 0.059	45.90
2018	5.6 ± 2.5	1.052 ± 0.059	45.74

7.2 Correction of the decay-time bias

During the previously discussed calibration of the decay-time uncertainty estimate, a decay-time bias $\mathit{\Delta} t = \mu$ in the form of a non-zero mean of the reconstructed decay-time distribution of prompt $D_s^- \pi^+$ candidates is observed. This bias has been extensively studied as it directly propagates to a bias of the oscillation frequency Δm_s, which is magnitudes larger than the expected uncertainty. At first, the bias is significantly reduced by omitting the flight-distance criteria by using a more explicit selection in Section 6.2. However, still a significant bias in the order of 5 fs is observed.

The origin of this bias is found in the alignment of the VELO detector. The dominant misalignment is a shift ΔT_x between the two VELO halves in the x direction. The resulting shift Δx of reconstructed tracks translates into a bias $\Delta z = \Delta x / \tan \alpha$ of the SV position along the z axis parallel to the beam line. Here, α denotes the angle between a track and the beam axis. The tracks detected at LHCb are heavily boosted in the forward direction. Hence, the z position of a SV dominates the reconstructed decay length, which is proportional to the decay time.

The decay-time bias can be directly accessed in samples of prompt $D_s^- \pi^+$ candidates as used in Section 7.1. However, it neither can be observed in $B_s^0 \to D_s^- \pi^+$ data, nor directly be ported, as the bias is very dependent on the candidates' kinematics in terms of α. Hence, a calibration function is used to translate the bias observed in prompt candidates Δt_{pr} found in a sample to a bias present in the signal sample Δt_{sig}. Here, the calibration is determined on simulation as discussed in [113]. Ten simulated samples are generated with an artificial misalignment in the range $\Delta T_x \in [0,\, 9]$µm. In each sample, the decay-time bias of both prompt and signal samples is evaluated with the help of the known generated decay time. A linear fit of the signal bias as a function of the bias from prompt candidates yields the calibration function

$$\Delta t_{\mathrm{sig}}(\Delta t_{\mathrm{pr}}) \;=\; (0.031 \pm 0.197)\,\mathrm{fs} \;+\; (0.496 \pm 0.032) \cdot \Delta t_{\mathrm{pr}}\,. \tag{7.6}$$

Further, minor contributions from a rotation ΔR_y in y direction and from direct translation ΔT_z in z of the modules are identified in the misalignment. However, the resulting effects are small and found to be properly addressed by the aforementioned calibration.

The total bias observed in the prompt samples and the corresponding bias evaluated for the signal samples is given in Table 7.2. These effective bias estimates are used to correct the decay time in Section 7.4 by shifting the mean μ of the Gaussian resolution function $\mathcal{R}(t \,|\, \delta t,\, \Delta t)$ for each data-taking period. More details on this correction are given in the dedicated thesis [113].

Table 7.2: Decay-time bias found in the data samples of prompt $D_s^- \pi^+$ candidates and the calibrated bias expected for the signal candidates.

Sample	Δt_{pr} [fs]	Δt_{sig} [fs]
2015 & 16	-4.562 ± 0.165	-2.254 ± 0.109
2017	-6.157 ± 0.152	-3.047 ± 0.112
2018	-4.844 ± 0.141	-2.394 ± 0.107

7.3 Preparation of the flavour tagging

The initial B_s^0 flavour q_i needed to retrace the oscillation is provided by a variety of flavour tagging algorithms as discussed in Section 4.5, while the flavour at decay q_f is defined by the charge of the accompanying pion in the flavour-specific $B_s^0 \to D_s^- \pi^+$ mode. The predicted initial flavour d comes with an intrinsic mistag rate ω. As a consequence the mixing state

$$q \;=\; q_f \cdot q_i \;=\; q_f \cdot d\,(1 - 2\omega) \tag{7.7}$$

dividing between Equation 3.45 and 3.46 is diluted. The tagging algorithms provide a per-event mistag estimate η, which can be used to incorporate the diluting factor properly after a calibration. As the dilution is independent of the oscillation frequency and only affects the amplitude of the measured decay rates, the calibration

$\omega(\eta)$ can be optimised in the decay-time fit itself. This assumption has been tested using pseudo experiments in Section 7.4, which do not show any biasing effect. Nevertheless, preparations are still required to set up the flavour tagging correctly. These preparations are performed in the course of thesis [109].

In this analysis, a number of tagging algorithms, namely the SS kaon, OS kaon, OS muon, OS electron, OS charm and OS vertex charge, are used simultaneously to achieve the highest possible sensitivity. The OS taggers are combined into a single tag decision d_{OS} and mistag estimate η_{OS}. The combination described in Section 4.5.3 is performed here by using the EspressoPerformanceMonitor (EPM) [86]. The combination algorithm is based on probabilistic properties. Hence, the individual OS taggers are precalibrated beforehand to provide solid mistag probabilities.

For the precalibrations sWeighted $B^+ \to J/\psi K^+$ data provided by the analysis of $B^0_{(s)} \to (c\bar{c})K^0_{\mathrm{S}}$ decays [114] is used. Due to its self-tagging nature, B^+ data allows a tagging calibration without the need for accounting data-simulation differences or further assumptions. Besides, the independence of the OS taggers to the signal B species permits this strategy. Further, the $B^+ \to J/\psi K^+$ decay mode is abundant and clean mode due to its charmonium resonance. However, the kinematics of $B^+ \to J/\psi K^+$ differs from the $B^0_s \to D^-_s \pi^+$ signal kinematic, because of the different B species and the different decay topology. To ensure good portability between the two modes, the $B^+ \to J/\psi K^+$ calibration sample is reweighted to match the signal kinematics. The per-event weights are determined based on the number of tracks present in the event, the transversal momentum and the pseudorapidity of the B candidate by the GBReweighter algorithm implemented in the hep_ml package [115]. The precalibration of the individual OS taggers is then performed using the EPM again. The model describing the dependence of true and estimated mistag best is found to be a second-order polynomial with a logistic link function

$$g(\eta) = \ln \frac{\eta}{1-\eta} \,. \tag{7.8}$$

As the OS combination and the SS kaon can be assumed to be independent of each other, no further combination is required. Instead, the two taggers are individually calibrated in the decay-time fit and incorporated into the decay-time PDF.Although the calibrations of the OS combination and the SS kaon are floating in the decay-time fit, they are studied using the EPM beforehand to evaluate a suitable calibration model. A linear function with the identity link $g(\eta) = \eta$ as described in Equation 4.7 and 4.6 is found to be sufficient.

7.4 Decay-time fit of the $B_s^0 \to D_s^- \pi^+$ signal

In the decay-time fit of the $B_s^0 \to D_s^- \pi^+$ data not only the theoretical decay rates from Equation 3.45 and 3.46 have to be modelled, but also experimental effects. Next to the previously discussed flavour tagging calibration and resolution function $\mathcal{R}(t - \Delta t \,|\, \delta t)$ incorporating the decay-time uncertainty and bias, the decay-time-dependent selection and reconstruction efficiency $\varepsilon_t(t)$ and the production and

$$a_{\text{prod}} = \frac{\sigma(B_s^0) - \sigma(\overline{B}_s^0)}{\sigma(B_s^0) + \sigma(\overline{B}_s^0)} \tag{7.9}$$

and detection asymmetries

$$a_{\text{det}} = \frac{\varepsilon(D_s^+(\to h^+h^-\pi^+)\pi^-) - \varepsilon(D_s^-(\to h^-h^+\pi^-)\pi^+)}{\varepsilon(D_s^+(\to h^+h^-\pi^+)\pi^-) + \varepsilon(D_s^-(\to h^-h^+\pi^-)\pi^+)} \tag{7.10}$$

need to be accounted for. Here, the asymmetries account for possible differences in the production cross-section $\sigma(i)$ of the $B_s^0/\overline{B}_s^0$ states and in the detection efficiency $\varepsilon(f)$ of the two oppositely charged final states. The full decay-time PDF is described by

$$\mathcal{P}\Big(t \,|\, \delta t,\, \Delta t,\, q_f,\, \vec{d},\, \vec{\eta}\Big) \sim \varepsilon_t(t) \cdot \int_{-\infty}^{\infty} \mathcal{R}(t - t' \,|\, \delta t,\, \Delta t) \cdot \mathcal{P}_{\text{eff}}\Big(t' \,|\, q_f,\, \vec{d},\, \vec{\eta}\Big)\, \mathrm{d}t' \tag{7.11}$$

with an effective PDF

$$\begin{aligned} \mathcal{P}_{\text{eff}}\Big(t \,|\, q_f,\, \vec{d},\, \vec{\eta}\Big) \sim \mathrm{e}^{-\Gamma_s t}\Big[&(1 + q_f \cdot a_{\text{det}}) \cdot \Big(\Phi^-\Big(\vec{d},\, \vec{\eta}\Big) - \Phi^+\Big(\vec{d},\, \vec{\eta}\Big) \cdot a_{\text{prod}}\Big) \cosh\left(\frac{\Delta\Gamma_s t}{2}\right) \\ &+ q_f\,(1 - q_f \cdot a_{\text{det}}) \cdot \Big(\Phi^+\Big(\vec{d},\, \vec{\eta}\Big) - \Phi^-\Big(\vec{d},\, \vec{\eta}\Big) \cdot a_{\text{prod}}\Big) \cos(\Delta m_s t)\Big] \end{aligned} \tag{7.12}$$

adapting the theoretical PDF, given in Equation 3.41, for imperfect knowledge of the tagged initial flavour and the presence of asymmetries. Here, the tagging effects are digested by the parameters $\Phi^\pm$ based on the flavour-specific mistag calibrations, defined in Equation 4.7 and 4.6, and tagging efficiencies

$$\varepsilon^{\overline{b}} = \varepsilon_{\text{tag}} + \frac{1}{2}\Delta\varepsilon_{\text{tag}} \qquad \text{and} \qquad \varepsilon^{b} = \varepsilon_{\text{tag}} - \frac{1}{2}\Delta\varepsilon_{\text{tag}}\,. \tag{7.13}$$

In the case of untagged events $d_1 = d_2 = 0$, these parameters are described by

$$\Phi^\pm = \left(1 - \varepsilon_1^b - \varepsilon_2^b - \varepsilon_1^b\varepsilon_2^b\right) \pm \left(1 - \varepsilon_1^{\overline{b}} - \varepsilon_2^{\overline{b}} - \varepsilon_1^{\overline{b}}\varepsilon_2^{\overline{b}}\right)\,. \tag{7.14}$$

For tagged events, the parameters are defined by

$$\begin{aligned}\Phi^{\pm}(d_1,\, \eta_1) \;&=\; \frac{1}{2}\varepsilon_1^b\, \left[1-\varepsilon_2^b + d_1(1-\varepsilon_2^b - 2\omega_1^b(\eta_1) + 2\omega_1^b(\eta_1)\,\varepsilon_2^b)\right] \\ &\pm\; \frac{1}{2}\varepsilon_1^{\bar{b}}\, \left[1-\varepsilon_2^{\bar{b}} + d_1\left(1-\varepsilon_2^{\bar{b}} - 2\omega_1^{\bar{b}}(\eta_1) + 2\omega_1^{\bar{b}}(\eta_1)\,\varepsilon_2^{\bar{b}}\right)\right] \qquad (7.15)\end{aligned}$$

if only one tagger provides a decision $d_1 \neq d_2 = 0$ or by

$$\begin{aligned}\Phi^{\pm}\left(\vec{d},\, \vec{\eta}\right) \;&=\; \frac{1}{4}\varepsilon_1^b\varepsilon_2^b\left[1 + d_1(1-2\omega_1^b(\eta_1)) + d_2(1-2\omega_2^b(\eta_2))\right. \\ &\qquad\left. +\; d_1 d_2(1 - 2\omega_1^b(\eta_1) - 2\omega_2^b(\eta_2) - 4\omega_1^b(\eta_1)\,\omega_2^b(\eta_2))\right] \\ &\pm\; \frac{1}{4}\varepsilon_1^{\bar{b}}\varepsilon_2^{\bar{b}}\left[1 + d_1\left(1-2\omega_1^{\bar{b}}(\eta_1)\right) + d_2\left(1-2\omega_2^{\bar{b}}(\eta_2)\right)\right. \\ &\qquad\left. + d_1 d_2\left(1 - 2\omega_1^{\bar{b}}(\eta_1) - 2\omega_2^{\bar{b}}(\eta_2) - 4\omega_1^{\bar{b}}(\eta_1)\,\omega_2^{\bar{b}}(\eta_2)\right)\right] \qquad (7.16)\end{aligned}$$

if both provide a decision $d_{1/2} \neq 0$. The resolution function with the bias correction is defined by

$$\mathcal{R}(t\,|\,\delta t,\, \Delta t) \;=\; \frac{1}{\sqrt{2\pi\sigma_t(\delta t)}}\mathrm{e}^{-(t-\Delta t)/(\sigma_t^2(\delta t))} \qquad (7.17)$$

as discussed in Section 7.1. The decay-time-dependent efficiency, also referred to as acceptance, is modelled by a cubic spline function

$$\varepsilon_t(t) \;=\; \sum_{i=1}^{8} c_i \cdot B_{i,3}\big(t\,|\,\vec{t}\,\big)\;, \qquad (7.18)$$

where the eight individual B-splines are recursively defined by

$$B_{i,1}\big(t\,|\,\vec{t}\,\big) \;=\; \begin{cases} 1 & \text{, if } t_i \le t < t_{i+1} \\ 0 & \text{, else} \end{cases} \qquad (7.19)$$

and $$B_{i,k+1}\big(t\,|\,\vec{t}\,\big) \;=\; \frac{t-t_i}{t_{i+k}-t_i}B_{i,k}\big(t\,|\,\vec{t}\,\big) \;+\; \left(1-\frac{t-t_{i+1}}{t_{i+1+k}-t_{i+1}}\right)B_{i+1,k}\big(t\,|\,\vec{t}\,\big)\;. \qquad (7.20)$$

The knot positions $\vec{t} = [0.4,\, 0.4,\, 0.5,\, 1.0,\, 1.5,\, 2.0,\, 3.0,\, 12.0,\, 15.0,\, 15.0,\, 15.0]\,\mathrm{ps}$ covering the full fit range $t \in [0.4,\, 15.0]\,\mathrm{ps}$ are fixed. The shape of the acceptance model is estimated in the fit by floating the coefficients c_i. For better stability, the last coefficient is constrained as the linear extrapolation of the previous two

$$c_8 \;=\; c_7 + \frac{t_8 - t_7}{t_7 - t_6}\,(c_7 - c_6)\;. \qquad (7.21)$$

To reduce ambiguities, one knot is fixed to the value $c_7 = 1.0$. The fit stability of the model with a floating acceptance is justified by the negligible sensitivity of the splines to the cosine of the oscillation in Equation 7.12, while the values of Γ_s and $\Delta\Gamma_s$ are fixed to values taken from [116] in the exponential and the hyperbolic terms. Further, this approach is validated with pseudo experiments in Chapter 8.

The decay-time fit is performed simultaneously for all three data-taking periods. Hereby, the acceptance, resolution and flavour tagging calibration are modelled individually for each period, while the physics quantities Δm_s, Γ_s and $\Delta\Gamma_s$, as well as the two asymmetries, are shared among the subsamples. Within the subsamples, it is not distinguished between the individual D_s^- decay modes or magnet polarities. Both the fixed parameters and the parameters fitted are reported in Appendix A.3. The detection asymmetry can not be disentangled from a physical asymmetry between the two final states $a_{\mathrm{fs}} = A_{\mathrm{sl}}^s$, which gives access to the small effects of indirect *CP* violation in the mixing as described in Section 3.2.2. Hence, the result of this parameter is kept blinded for an unbiased analysis of indirect *CP* violation in $B_s^0 \to D_s^- \pi^+$ in the future. The extracted value of the oscillation frequency is

$$\Delta m_s = 17.7683 \pm 0.0051 \text{ ps}^{-1} .$$

The uncertainties of all fitted parameters are corrected for the usage of sWeights based on the asymptotic approach developed in reference [91], which shows good uncertainty coverage in pseudo-experiment-based studies. The decay-time distribution together with the fitted PDF is presented in Figure 7.1 split into contributions from mixed, unmixed and untagged candidates. The disentangled oscillation and the contribution of the floating acceptance model are also presented. The estimated tagging power of the data sample based on the flavour tagging calibration obtained from the fit is listed in Table 7.3.

Table 7.3: The tagging power of the two individual taggers and their combination, evaluated by the lhcb_ftcalib [87] package, based on the calibration parameters obtained from the decay-time fit.

Sample	OS combination [%]	SS kaon [%]	Combined OS + SS [%]
2015 & 16	3.65± 0.19	2.31± 0.15	5.80± 0.23
2017	3.84± 0.20	2.25± 0.15	5.94± 0.24
2018	4.03± 0.19	2.39± 0.15	6.22± 0.23
2015 - 18	3.84± 0.11	2.32± 0.09	5.99± 0.13

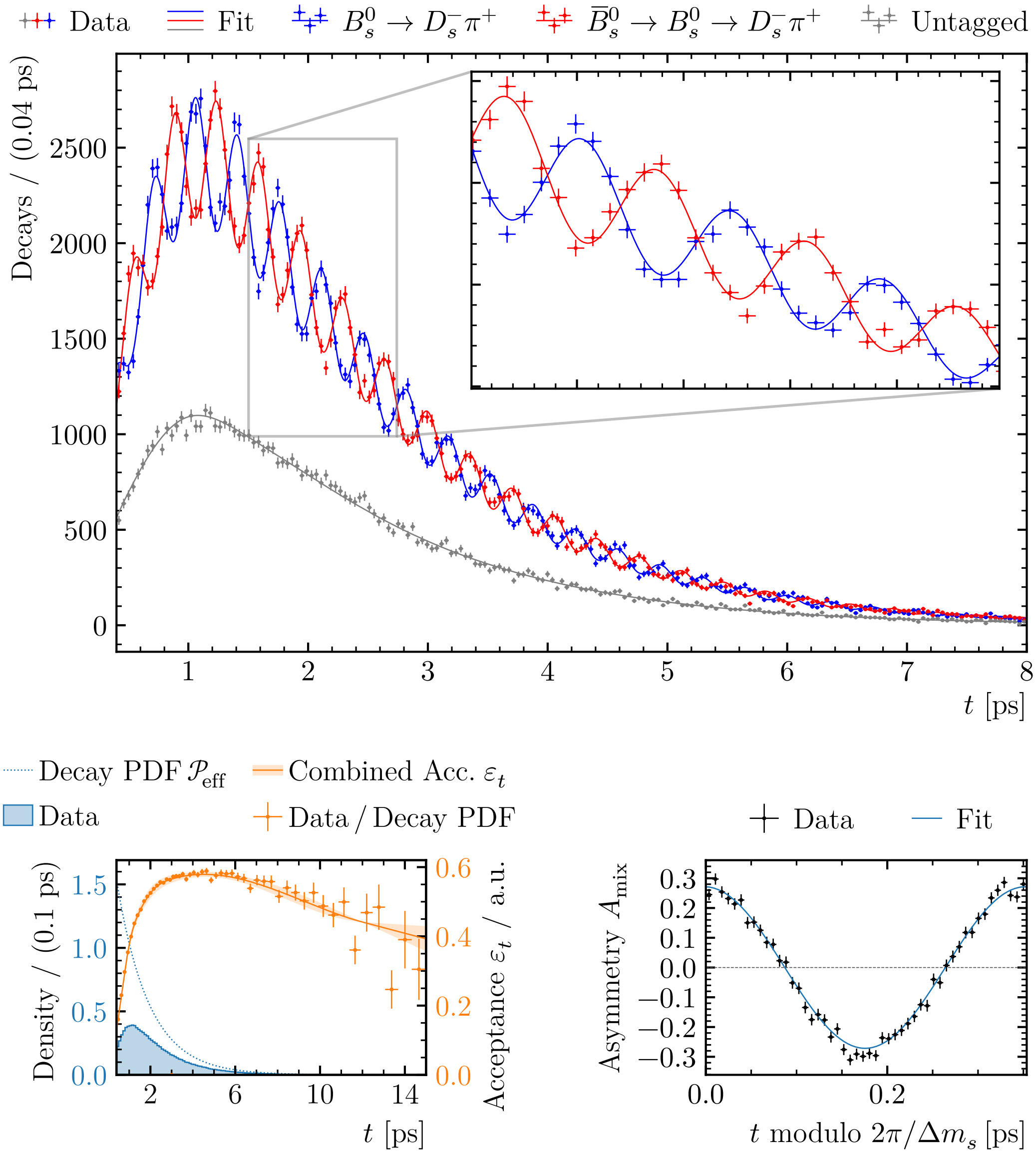


Figure 7.1: The fitted decay-time distribution in different representations: On top, the distribution split into untagged (grey), unmixed (blue) and mixed (red) decays; Below on the left, the effect from the decay-time acceptance ε_t (orange), disentangled from the decay PDF $\mathscr{P}_{\mathrm{eff}}$ (light blue); Below on the right, the projection of the experimentally-diluted, decay-time-dependent asymmetry A_{mix} (Equation 3.47), folded to a single oscillation period.

8 Systematic studies and cross-checks on the oscillation frequency Δm_s

A variety of studies is performed to prove the robustness of the fit strategy discussed in the previous chapters 6 and 7 and to study systematic effects on on the measured value of Δm_s. These studies include variations in the fit strategy and fits performed on simulation. Here, simulation corresponds to both full LHCb simulation, including the full reconstruction procedure, and simplified pseudo experiments, where the detector effects are explicitly parameterised. In Section 8.1 systematic effects introduced by the reconstruction due to a finite precision of the detector are discussed. Effects related to the chosen model are studied in Section 8.2. Lastly, the effect of an imperfect assumption of independent observables in the sWeighting approach is investigated in detail in Section 8.3. A summary of all studies, which show a non-vanishing systematic effect on the measured value of the oscillation frequency Δm_s, is given in Table 8.1. Overall, the measured value

$$\Delta m_s = 17.7683 \pm 0.0051 \pm 0.0032 \text{ ps}^{-1} ,$$

where the first uncertainty is the statistical precision from the decay-time fit and the second is accounting for all systematic effects, is still statistically limited.

Table 8.1: The full list of the systematic studies performed and the corresponding uncertainty on the oscillation frequency Δm_s.

Systematic	$\sigma_{\text{syst}}^{\Delta m_s}$ [ps^{-1}]
Reconstruction effects:	
Momentum scale uncertainty	0.0007
VELO z-scale	0.0018
VELO alignment correction	0.0020
Closure tests:	
Subsample splits	0.0003
Pseudo experiment based closure test	–
Multidimensional mass fit model:	
Signal parametrisation	–
Specific background parametrisation	0.0002
Combinatorial background parametrisation	–
Fixed yields	–
Fixed fraction f_{low}	0.0005
Decay-time fit model:	
Decay-time resolution model	0.0011
Decay-time acceptance model	–
Flavour tagging model	–
Validity of the sWeighting approach	0.0011
Fixed Γ_s, $\Delta\Gamma_s$	–
Total systematic uncertainty	0.0032
Statistical uncertainty	0.0051

8.1 Reconstruction effects

The high precision of this measurement requires an exact understanding of the experimental setup. In the following, detector effects affecting the reconstructed decay time are investigated. At LHCb the decay time is reconstructed from the particle mass m, the measured decay length l and the measured momentum p as

$$t = \frac{m \cdot l}{p} \,. \tag{8.1}$$

Both momentum and mass of the reconstructed signal candidate are reconstructed from the individual final-state particle momenta. Hence, the uncertainty on the reconstructed decay time is dominated by the momentum resolution of the tracking system and the precision of the VELO-based decay-length reconstruction.

Momentum scale uncertainty

The momentum reconstruction of charged particle tracks at LHCb is calibrated using the known masses of several narrow resonances. With this calibration a precision of 0.03 % [117] is achieved for the momentum reconstruction of LHCb data recorded between 2015 and 2018. The impact on the oscillation frequency Δm_s of this resolution effect is studied using full LHCb simulation. To study the effect of (over-)underestimation of the reconstructed momentum, all final-state momenta are scaled by a factor of 0.9997 (1.0003) in the simulation. The dependent quantities are recomputed and a decay-time fit of the modified sample is performed. The two fit results differ by 0.0014 ps^{-1} and a symmetric uncertainty of $\pm 0.0007\ \mathrm{ps}^{-1}$ is assigned as a systematic uncertainty for the imperfect knowledge of the momentum scale.

VELO z-scale uncertainty

The relative position of the VELO modules along the beam axis has been determined before the first data-taking and has been unchanged since then. The accuracy of the z position is $100\,\mu\mathrm{m/m} = 0.01\,\%$ [107]. This uncertainty directly propagates over reconstructed decay length and time to the measured oscillation frequency. Based on the value of Δm_s this corresponds to an absolute value of 0.0018 ps^{-1} which is assigned as systematic uncertainty due to the VELO z-scale uncertainty.

VELO alignment correction

In Section 7.2 the effects of misaligned VELO modules are discussed. A correction is applied to compensate effects of the misalignment on the reconstructed decay time. The correction is validated in simulated signal samples, which are reconstructed with an artificial misalignment introducing the need for a correction. A decay-time fit is performed for each of these samples. After the correction, the oscillation frequency measured in the misaligned simulation samples is found to agree with the correctly aligned sample within 0.002 ps^{-1}. This value is taken as a residual systematic uncertainty from the VELO misalignment after the decay-time bias correction.

Overall, the reconstruction effects are the dominant source of systematic uncertainties in this measurement. Nevertheless, they are not limiting the sensitivity of this measurement as they reach 40 % of the statistical precision at most. However, they could become a limiting factor for future measurements as they do not rely on the data or simulation statistics and are directly dependent on the detector resolution only.

8.2 Model-dependent systematics

During the analysis, several data properties and detector effects are modelled. The corresponding models are usually empirically chosen or rely on parameters only known at limited precision. Therefore, the effects of varying parameters and alternative parametrisations are tested in the following to study the effects of the chosen models.

8.2.1 Closure tests

The full analysis strategy as a whole is validated by two closure tests. A data-driven closure test is performed by performing all fits on various subsamples and averaging the resulting oscillation frequencies. Pseudo experiments with the possibility to simulate various properties by chosen parametrisations are generated and fitted to prove that the fit procedure leads to unbiased results.

Subsample fits

The decay-time fit is repeated on the data sample split into the two magnet polarities, the four D_s^- decay modes and the three data-taking periods. Further, both the MDFit and the decay-time fit are repeated on subsamples split into two bins of the B_s^0 momentum and the BDT classification. Additionally, fits of the data sample are performed only using the tagging information of either the OS taggers or of the SS kaon tagger. The results from the individual subsamples are then averaged per set and compared to the nominal fit result. The deviations to the nominal result are listed in Table 8.2. The largest deviation is used to assign an additional uncertainty of 0.0003 ps^{-1}.

Pseudo-experiment based closure test

A set of 1500 pseudo-experiment samples is generated in the course of the thesis [109] to validate the robustness of the fit strategy. The strategy of generating and evaluating the pseudo-experiment samples discussed here is also the baseline to the studies summarised in Section 12.5 Each sample consists of subsamples for all data-taking periods, D_s^--decay modes and for the signal and background components modelled in the mass fit in Section 6.3. The sample size is randomly drawn from a Poisson distribution corresponding to the fitted yields reported in Table 6.2. The $m(D_s^-\pi^+)$ and $m(h^\mp h^+ h^-)$ masses are generated from the fitted MDFit PDF. The

Table 8.2: The difference of the oscillation frequency Δm_s obtained from the weighted average of subsamples and the nominal result. The uncertainties are computed accounting for correlations among the subsamples and the combined sample.

Averaged subsamples:	Deviation to nominal [ns^{-1}]
Magnet polarity:	
Up, down	0.031 ± 0.020
D_s^--decay mode:	
$\phi\pi^-$, $K^{*0}K^-$, $(K^-K^+\pi^-)_{\mathrm{NR}}$, $\pi^-\pi^+\pi^-$	-0.152 ± 0.006
Data-taking period:	
2015 & 16, 2017, 2018	-0.022 ± 0.004
Flavour tagging:	
OS only, SS kaon only	0.24 ± 0.07
B_s^0 momentum:	
$p_{B_s^0} \leq 120$ GeV/c, $p_{B_s^0} > 120$ GeV/c	-0.324 ± 0.015
BDT:	
$\alpha_{\mathrm{BDT}} \in [0.475, 0.94)$, $\alpha_{\mathrm{BDT}} \in [0.94, 1.0]$	0.339 ± 0.033

reconstructed decay time t, the final-state flavour q_f and the tag decisions $d_{\mathrm{OS/SS}}$ are simultaneously drawn from the fitted decay-time PDF, while the decay-time error estimate δt and the two mistag estimates $\eta_{\mathrm{OS/SS}}$ are drawn from the observed data distributions.

The full fit strategy is mimicked for each combined pseudo-experiment sample. At first, the MDFit is performed to extract sWeights. Then the decay-time fit is used to extract the oscillation frequency Δm_s. To study potential fit biases and the uncertainty coverage, the pull distribution

$$\theta_i = \frac{x_i - \mu_x}{\sigma_i} \tag{8.2}$$

of the fitted variables x and uncertainties σ with respect to the generated value μ_x are studied. The pull distribution is expected to be Gaussian with a mean $\overline{\theta}$ indicating potential systematic biases and a width $\sigma_\theta = 1.0$ in case of good coverage.

The pull distributions of the MDFit variables are overall in good agreement with a normal distribution with biases of 10 % at most. These small biases in the MDFit can be neglected as they do not impact the oscillation frequency Δm_s, which shows an unbiased pull distribution with good coverage. Similar to the MDFit, the nuisance parameter in the decay-time fit show smaller biases, which are neglected to the unbiased value of Δm_s. Nevertheless, alternative decay-time acceptance parametrisations are tested in the pseudo-experiment fits since the coefficient c_2 shows a significant bias. Using an alternative parametrisation with modified knot positions and fixed coefficients $c_7 = c_8 = 1$ an unbiased acceptance can be extracted. However, this change does not affect the fitted value of Δm_s. Hence, no systematic effect is found in this closure test.

At an earlier stage of the analysis, the standard sWeight correction defined in Equation 5.20 was used for the decay-time fits to reduce the computational cost. This way the fitted toy uncertainties show an under coverage of approximately 5 % due to this choice of sWeight correction. However, the application of the asymptotically correct sWeight correction [91] in the decay-time fit yields good coverage in the pseudo experiments.

8.2.2 Mass fit model

The mass fit model is discussed in Section 6.3. The signal shape is studied on reweighted simulation and is described best by the sum of a Hypatia function and a Johnson SU function. No alternative model is found to yield a reasonable fit quality based on a χ^2 test of the fitted PDF and the mass distribution. The robustness of the exact analytical shape is tested against the simulation-to-data reweighting. While in the nominal strategy, the simulation is corrected based on the spatial uncertainties of the reconstructed B_s^0 and D_s^- vertices in z direction, an alternative reweighting based on the B_s^0 momentum and the number of tracks within an event is tested. The result of the decay-time fit using sWeights extracted with the signal parametrisation based on the alternative reweighting shows a negligible deviation of $0.069\,\mathrm{ns}^{-1}$ to the nominal fit. Hence, no systematic uncertainty is assigned.

Two modifications are tested for the parametrisation of the combinatorial background. In one setup the second exponential slope $c^2_{B_s^0}$ is allowed to float in the $m(D_s^-\pi^+)$ mass. In the other, a second exponential function is added to the nominal $m(h^\mp h^+ h^-)$ model, described by the sum of the signal shape and an exponential function. The maximal change in Δm_s observed for the different combinatorial parametrisations is $0.006\,\mathrm{ns}^{-1}$ and concluded to be negligible.

The physics backgrounds in the mass spectra are either modelled by templates, based on kernel estimators [118], or by analytical shapes each fitted on simulation. To vary these models, again the alternative simulation-data correction is applied. A deviation of $0.149\,\mathrm{ns}^{-1}$ is observed for the fitted oscillation frequency when the specific backgrounds in the MDFit are modelled with the alternative reweighting. For this effect a conservative uncertainty of 0.0002 ps^{-1} is assigned.The effect of some background yields being fixed is tested by varying the yields individually by ±50 % of the nominal value. The impact on the oscillation frequency Δm_s is $0.06\,\mathrm{ns}^{-1}$ at most. No additional systematic uncertainty is assigned, as this minor effect is considered to be accounted for by the conservative uncertainty assigned for the shape parametrisation.

The fraction between the $B^0 \to D_s^- \pi^+$ and the $B_s^0 \to D_s^{*-} \pi^+$ backgrounds is fixed to an arbitrary value of $f_{\mathrm{low}} = 0.5$ as the fit shows no sensitivity to this parameter. The parameter is varied to the most extreme scenarios $f_{\mathrm{low}} \in \{0.0,\, 1.0\}$. The largest effect of this variation observed on the oscillation frequency Δm_s is $0.474\,\mathrm{ns}^{-1}$, which is assigned as a systematic uncertainty.

8.2.3 Decay-time fit model

In Chapter 7, a number of different calibrations is applied to the data before or during the decay-time fit. The robustness of these calibrations is studied by varying the corresponding models.

Decay-time resolution model

The calibration of the decay-time error estimate, performed in Section 7.1, is based on an effective width of the resolution, observed in prompt $D_s^- \pi^+$ candidates and modelled by the sum of three Gaussian functions. As an alternative, resolution models based on the width of either the narrowest or the widest Gaussian core only are used. Applying the calibration overestimating the resolution using the wide Gaussian, a $0.0007\ \mathrm{ps}^{-1}$ shift of the oscillation frequency Δm_s is observed. The underestimation of the calibration, based on the narrow Gaussian, yields a shift of $-0.0015\ \mathrm{ps}^{-1}$. The corresponding systematic uncertainty is symmetrised by assigning the average of the absolute deviations $0.0011\ \mathrm{ps}^{-1}$.

Decay-time acceptance model

Two alternative parametrisations are tested for the decay-time acceptance model showing negligible impact to the measured value of Δm_s each. One model varies the knot positions of the eight B-splines while fixing the coefficient of the last spline to $c_8 = 1$, the other one makes use of an increased number of 14 B-splines. Further, the effect of the floating coefficients is studied by fixing the first three coefficients to ±10 % of their values extracted in the nominal fit. The resulting effect on the oscillation frequency Δm_s is also negligible. Hence, no systematic uncertainty is assigned for the decay-time acceptance model.

Flavour tagging model

The mistag calibrations of the SS kaon tagger and of the OS combination are modelled by linear functions. Investigations using the EspressoPerformanceMonitor do not show any improvement using higher-order polynomials. Instead, the addition of a logistic link function, improving the probabilistic interpretation of the mistag at its boundaries, is used in the decay-time fit as an alternative model. The effect on the fitted value of Δm_s is found to be negligible. Besides, the decay-time fit is performed using an alternative OS combination, with the individual OS taggers not being precalibrated before the combination. Again, no deviation of the oscillation frequency is observed. Hence, no systematic uncertainty is assigned for the flavour tagging model.

Effect of the fixed values of $\varGamma_s$ and $\Delta\varGamma_s$

Both $\varGamma_s = 0.6600 \pm 0.0020\ \mathrm{ps}^{-1}$ [116] and $\Delta\varGamma_s = 0.085 \pm 0.006\ \mathrm{ps}^{-1}$ [116] are fixed in the decay-time fit neglecting the corresponding uncertainties. Further, the correlation of the two quantities $\rho(\varGamma_s, \Delta\varGamma_s) = -0.124$ [116] is neglected. The effect of these simplifications is tested using fits to the generated pseudo-experiment samples. The fit of each sample is performed with a set of fixed values for $\varGamma_s$ and $\Delta\varGamma_s$, randomly drawn from a bivariate Gaussian distribution and resembling their correlated uncertainties. Both the mean and the width of the resulting deviations of Δm_s, with respect to the pseudo-experiment fits using the nominal fixed values, are negligible and no systematic uncertainty is assigned.

In summary, the choice of parametrisation and models only shows a minor impact on the measurement with the model-dependent uncertainties being small or negligible. The largest effect is found to be the choice of the decay-time resolution model, which again is closely related to the detector resolution. It is of a similar size as the residual systematic introduced by the sWeighting approach, which is studied in detail in the following due to its complexity and importance.

8.3 Validation of the sWeighting approach

The sPlot method introduced in Section 5.3 plays a crucial role in this analysis as it allows the properties of the $B_s^0 \to D_s^- \pi^+$ decays to be statistically separated from residual backgrounds. This approach requires the discriminant variables (the $m(D_s^- \pi^+)$ and $m(h^\mp h^+ h^-)$ masses) to be independent of the variables of interest used in the decay-time fit. In Chapter 7 the two sets of fit variables are assumed to fulfil this requirement. However, this assumption is not expected to hold perfectly. Therefore, the validity of this assumption and possible systematic effects on the oscillation frequency Δm_s need to be studied.

8.3.1 Correlations among fit observables

Small dependencies of mass and decay-time variables can be introduced by a variety of different mechanisms, especially in the reconstruction. Hence, these dependencies cannot be simulated trivially in the aforementioned pseudo experiments. Instead, the full LHCb simulation is used, which on the one hand gives access to such dependencies, while lacking control over data-taking related nuisance parameters on the other hand.

At first, correlations of the two reconstructed masses and the decay time t and its uncertainty estimate δt are studied in simulated signal and background samples, as the presence of correlations directly indicates dependence. The correlation observed in the simulation is given in Table 8.3 for every component modelled in the MDFit, together with the yield fitted on data and the available sample size after applying the full selection. While all D_s^--decay modes are available in the signal simulation, only the $(K^-K^+\pi^-)_{\mathrm{NR}}$ mode is available for the simulated background components and used as a representative for the other modes, too. Due to its nature, the combinatorial background cannot be easily simulated. Hence, data sideband $m(D_s^-\pi^+) \in [5600, 6800]\,\mathrm{MeV}/c^2$ is used instead of simulation as a proxy for the combinatorial background. In this case. the $m(D_s^-\pi^+)$ mass is moved to the signal region using a linear transformation.

The correlations between MDFit and decay-time observables are found to be negligible in the signal simulation. While this does not exclude dependency, it does not indicate one, either. However, in the background samples such correlations can be observed indicating dependence, especially between $m(D_s^-\pi^+)$ mass and decay-time uncertainty. Hence, the requirement of independent discriminant observables for the sPlot method is not fulfilled and the effect on the oscillation frequency Δm_s from this misassumption needs to be studied.

8.3.2 Effect of neglected correlations

A bootstrapping approach is used to study the effect of the present correlations making use of the large simulation samples, exceeding the statistics of the recorded data. As shown in Table 8.3, the available statistics are of the same size as the observed yield in data for the combinatorial background proxy, while it is even

Table 8.3: Pearson correlation coefficient ρ for different pairs of the reconstructed masses used in the MDFit and of the decay-time observables in the decay-time fit listed for different samples. Also, the available statistic of the sample after the full selection is compared to the yield observed in the data MDFit.

Component	Available stat.	Nominal yield	$\rho\big(m(h^{\pm}h^{+}h^{-}), t\big)$	$\rho\big(m(h^{\pm}h^{+}h^{-}), \delta t\big)$	$\rho\big(m(D_s^{-}\pi^{+}), t\big)$	$\rho\big(m(D_s^{-}\pi^{+}), \delta t\big)$
Signal						
$D_s^- \to K^{*0}K^-$	4 030 186	103 180	−0.008	0.006	0.0	−0.007
$D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$	2 543 872	64 920	−0.007	0.005	−0.001	−0.007
$D_s^- \to \phi\pi^-$	4 745 053	141 640	−0.005	0.005	−0.001	−0.007
$D_s^- \to \pi^-\pi^+\pi^-$	859 994	65 130	−0.007	0.004	0.0	−0.006
Combinatorial						
$D_s \to K^{*0}K^-$	5124	9511	−0.007	0.011	−0.124	−0.015
$D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$	10 409	19 540	0.003	0.002	−0.092	0.017
$D_s^- \to \phi\pi^-$	6553	10 474	0.002	0.013	−0.075	−0.043
$D_s^- \to \pi^-\pi^+\pi^-$	48 005	32 763	0.009	0.006	−0.036	−0.018
$B_s^0 \to D_s^{\mp}K^{\pm}$						
$D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$	494 663	499	−0.009	0.007	0.013	0.095
$B^0 \to D^{\mp}\pi^{\pm}$						
$D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$	2 586 650	3064	−0.005	0.038	0.036	0.236
$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$						
$D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$	122 760	2197	0.007	0.004	0.038	−0.147
$B_s^0 \to D_s^{*-}\pi^+$						
$D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$	4959	831	0.012	−0.009	−0.077	−0.041

orders of magnitudes larger in the case of the simulated samples. In the course of the bootstrapping, events for each component are randomly drawn from the proxy samples. The number of events drawn is itself randomly taken from a Poisson distribution with a mean defined by the yield observed in data for each component. Combining the bootstrapped components, a data-like sample is obtained, which requires the signal to be extracted by the usage of sWeights before it can be used in the decay-time fit.

The effect of the correlations can be studied by running the analysis pipeline both in the presence and absence of the correlations. To do so, each bootstrapped sample is duplicated. The original sample, referred to as the correlated sample, maintains the correlations. In the duplication, referred to as the uncorrelated sample, the correlations are artificially vanished. To vanish the correlations, the set of mass variables is randomly shuffled between all candidates of each component in the bootstrapped sample, while every candidate keeps its original set of decay-time observables.

A number of 300 samples is bootstrapped. The problematic correlations are vanished as expected in the uncorrelated samples, while the correlated samples reproduce the original correlations from Table 8.3 well. The average correlations observed in both sets of samples are reported in Appendix A.4 for completeness.

A MDFit is run on all bootstrapped samples to extract sWeights. The sWeighted samples are then used in a decay-time fit to extract the oscillation frequency Δm_s, which is generated with a value of $\Delta m_s^0 = 17.8\ \mathrm{ps}^{-1}$ in the simulated signal samples. In the uncorrelated samples, the generated value is well reproduced by the average of the fitted value $\langle \Delta m_s^u \rangle = 17.800\,16 \pm 0.000\,23\ \mathrm{ps}^{-1}$, while a small deviation is observed in the correlated samples $\langle \Delta m_s^c \rangle = 17.799\,04 \pm 0.000\,23\ \mathrm{ps}^{-1}$. The distributions of the fitted values for both the correlated and the uncorrelated samples, as well as their per-sample comparison are displayed in Figure 8.1. An average deviation of $\langle \Delta m_s^c - \Delta m_s^u \rangle = \langle \delta(\Delta m_s) \rangle = 1.12 \pm 0.08\,\mathrm{ns}^{-1}$ is observed and assigned as systematic uncertainty for neglecting correlations among the fit observables in the sWeighting approach.

Cross-checks on the bootstrapping strategy

Due to its complexity, the robustness of the bootstrapping study, used to assign a systematic uncertainty for the neglecting correlations among the fit observables, itself is tested. In the absence of backgrounds, the sWeights are expected to not affect the analysis, even if there are significant dependencies among the fit observables. Hence, a null test is performed using only the signal component in the bootstrapping. In 300 samples of pure bootstrapped signal the observed per-sample differences $\delta(\Delta m_s) = \mathcal{O}(10^{-11}\,\mathrm{ps}^{-1})$ are negligible and expected to originate purely from numerical effects. Consequently, both the averaged fit result for the correlated and uncorrelated samples match $\langle \Delta m_s^c \rangle = \langle \Delta m_s^u \rangle = 17.799\,79 \pm 0.000\,23\ \mathrm{ps}^{-1}$ and are in agreement with the generated value. The corresponding fit results are displayed against each other in Figure 8.2.

Alternative approaches to obtain the combinatorial background proxy are studied as this background is the dominant one. Further, the available statistic is rather small so that multiple counting of candidates could affect the bootstrapping study. To increase the available statistics, the BDT requirement aiming to reduce this background is loosened for this component. The looser BDT cut values of $\alpha_{\mathrm{BDT}} \in \{0.4, 0.3, 0.2, 0.1\}$ do not show a significant impact on the correlations present in the sample. Besides, the larger statistics allow the sideband range to be reduced. Two control regions are defined, a lower-mass region $[5800, 6300]\,\mathrm{MeV}/c^2$ and an upper-mass region $[6300, 6800]\,\mathrm{MeV}/c^2$. In the lower-mass region, similar correlations compared to the nominal sideband are observed, in the upper-mass region the correlations between $m(D_s^- \pi^+)$ mass and decay time are reduced to the order of $\rho(m(D_s^-\pi^+), t) \sim 1\,\%$. The full bootstrapping study is repeated with the alternative background proxies. The observed deviation between the correlated and the uncorrelated samples is compatible among most of the different combinatorial background proxies. Only the setup using the upper-mass control region, which shows decreased correlations, also provides smaller deviations. Overall, the study is concluded to be robust against the combinatorial background proxy.

The impact of specific physics backgrounds is tested by individually adding these components. The bootstrapping study is repeated in the absence of all physics backgrounds. Further, the signal and combinatorial background are tested together with each of the individual physics backgrounds, the misidentified $\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^- \pi^+$, the misidentified $B_s^0 \to D_s^\mp K^\pm$ and the low-mass background composed of $B^0 \to D^\mp \pi^\pm$ and $B_s^0 \to D_s^{*-} \pi^+$. Only in the presence of the $B_s^0 \to D_s^\mp K^\pm$ background a significant neglecting the correlations is observed. Still, this effect is a factor of two smaller than in the nominal scenario. It is concluded, that these effects are not caused by the presence of a single specific background but by their interplay in the sWeighting PDF. All of the aforementioned bootstrapping setups and the observed deviations are summarised in Table 8.4.

Table 8.4: Averaged per-sample deviations of the correlated and uncorrelated bootstrapped samples, with respect to each other and to the generated value $\Delta m_s^0 = 17.8\ \text{ps}^{-1}$, using different background contaminations and combinatorial background proxies.

Backgrounds	BDT cut	Sideband region	$\langle \Delta m_s^c - \Delta m_s^0 \rangle\ [\text{ns}^{-1}]$	$\langle \Delta m_s^u - \Delta m_s^0 \rangle\ [\text{ns}^{-1}]$	$\langle \Delta m_s^c - \Delta m_s^u \rangle\ [\text{ns}^{-1}]$
All	0.475	nominal	-0.96 ± 0.23	0.16 ± 0.23	-1.12 ± 0.08
–	0.475	–	0.21 ± 0.23	0.21 ± 0.23	$(0.012 \pm 0.013) \cdot 10^{-6}$
All	0.400	nominal	-1.46 ± 0.23	-0.63 ± 0.23	-0.83 ± 0.09
All	0.300	nominal	-1.38 ± 0.23	-0.31 ± 0.24	-1.07 ± 0.09
All	0.200	nominal	-2.00 ± 0.23	-0.75 ± 0.24	-1.25 ± 0.09
All	0.100	nominal	-1.37 ± 0.23	-0.25 ± 0.24	-1.11 ± 0.07
All	0.200	lower	-2.00 ± 0.23	-0.80 ± 0.24	-1.20 ± 0.08
All	0.200	upper	-1.27 ± 0.23	-0.67 ± 0.24	-0.60 ± 0.09
Comb	0.475	nominal	0.32 ± 0.23	0.22 ± 0.23	0.10 ± 0.07
Comb, $\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^- \pi^+$	0.475	nominal	0.38 ± 0.23	0.20 ± 0.23	0.18 ± 0.07
Comb, $B_s^0 \to D_s^\mp K^\pm$	0.475	nominal	-0.47 ± 0.23	0.20 ± 0.23	-0.67 ± 0.07
Comb, $B^0 \to D^\mp \pi^\pm$,$B_s^0 \to D_s^{*-} \pi^+$	0.475	nominal	0.01 ± 0.23	0.18 ± 0.23	-0.17 ± 0.07

8.3.3 Motivation to redefine the PDF used for the sWeighting

In Section 6.3 a redefined PDF is introduced to obtain the sWeights to subtract the residual backgrounds. The original PDF defined in Equation 6.8 and used for the MDFit distinguishes between all described background components. The redefined PDF only distinguishes between a signal and an effective background component summarising all individual background contamination. This modification is reasoned by a large bias found in the bootstrapping studies to evaluate the neglecting correlations among the fit observables at an earlier stage of the analysis. Biased fit results from the correlated samples cause a large deviation of $\langle \delta(\Delta m_s) \rangle = -0.010\,70 \pm 0.000\,22$ ps^{-1} from the uncorrelated samples, using the original PDF for the sWeight extraction.

The deviations observed in 100 bootstrapped samples, using the original PDF for the sWeighting procedure, are reported for different setups in Table 8.5. Per definition, no deviations between correlated and uncorrelated samples are found in the bootstrapping of pure signal. Also by the addition of combinatorial background, no significant effect is observed, while the significant bias is in the presence of the physical backgrounds. Again, the addition of the $B_s^0 \to D_s^{\mp} K^{\pm}$ background has the largest impact, but no specific background responsible for this bias can be identified.

Next to the bias, it is observed that using the original PDF for the sWeighting yields a very unsteady distribution of the sWeights as a function of the $m(D_s^- \pi^+)$ mass in the case of a 1d-mass fit. Hence, the sWeighted phase space is expected to be distorted in this scenario. However, by introducing the PDF redefinition the behaviour of the sWeight is smoothed out and the effect of neglecting correlations is reduced to the reasonable size reported in Section 8.3.2.

The change of the sWeighting PDF from Equation 6.8 to 6.9 introduces a shift of $7.7 \pm 0.4\,\mathrm{ns}^{-1}$. This behaviour is well aligned with the expectations based on the bootstrapping result. The observed shift exceeds the statistical precision of the measurement by more than a factor of 1.5, proving the importance of this measure. Finally, it must be stated that this change of the strategy has been established before the unblinding of the measured value of the oscillation frequency Δm_s and hence is not affected by possible expectations of the result.

Table 8.5: Averaged per-sample deviations of the correlated and uncorrelated bootstrapped samples, with respect to each other and to the generated value $\Delta m_s^0 = 17.8$ ps^{-1}, *before* the redefinition of the sWeighting PDF.

Backgrounds	BDT cut	Sideband region	$\langle \Delta m_s^c - \Delta m_s^0 \rangle$ [ns^{-1}]	$\langle \Delta m_s^u - \Delta m_s^0 \rangle$ [ns^{-1}]	$\langle \Delta m_s^c - \Delta m_s^u \rangle$ [ns^{-1}]
–	0.475	–	0.3 ± 0.4	0.3 ± 0.4	$(0.00 \pm 0.05) \cdot 10^{-6}$
Comb	0.475	nominal	0.2 ± 0.4	0.3 ± 0.4	0.14 ± 0.14
All	0.475	nominal	-10.7 ± 0.4	0.0 ± 0.4	-10.70 ± 0.22

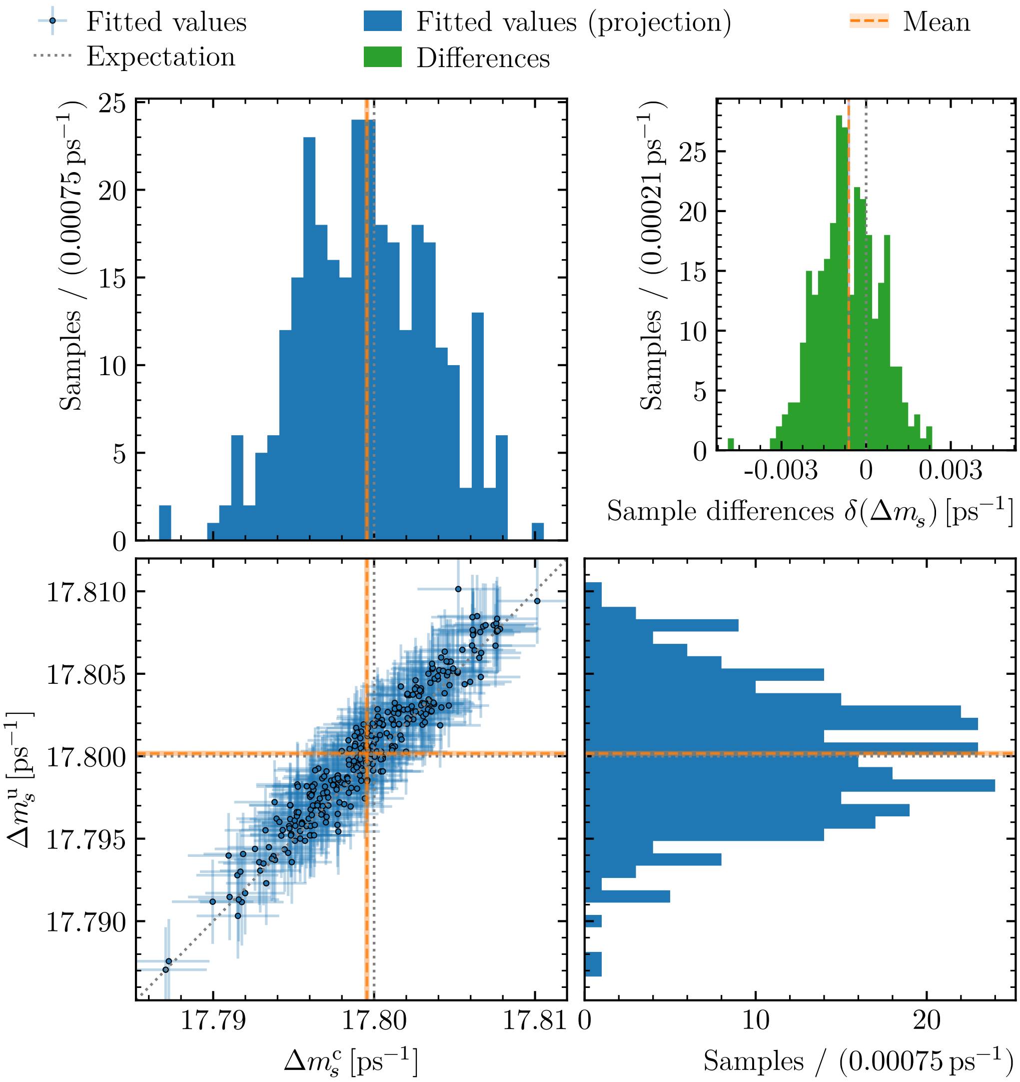


Figure 8.1: The impact of correlations in the sWeighting on the measured value of the oscillation frequency Δm_s. The fitted results of 300 bootstrapped samples in the presence and absence of correlations are compared in the bottom left figure. The corresponding projections are presented at the top left and the bottom right, respectively. While the mean (orange) reproduces the generated value well in the absence of correlations, a significant bias can be observed in its presence. Similarly, the per-sample differences δ given at the top right show a significant deviation from zero.

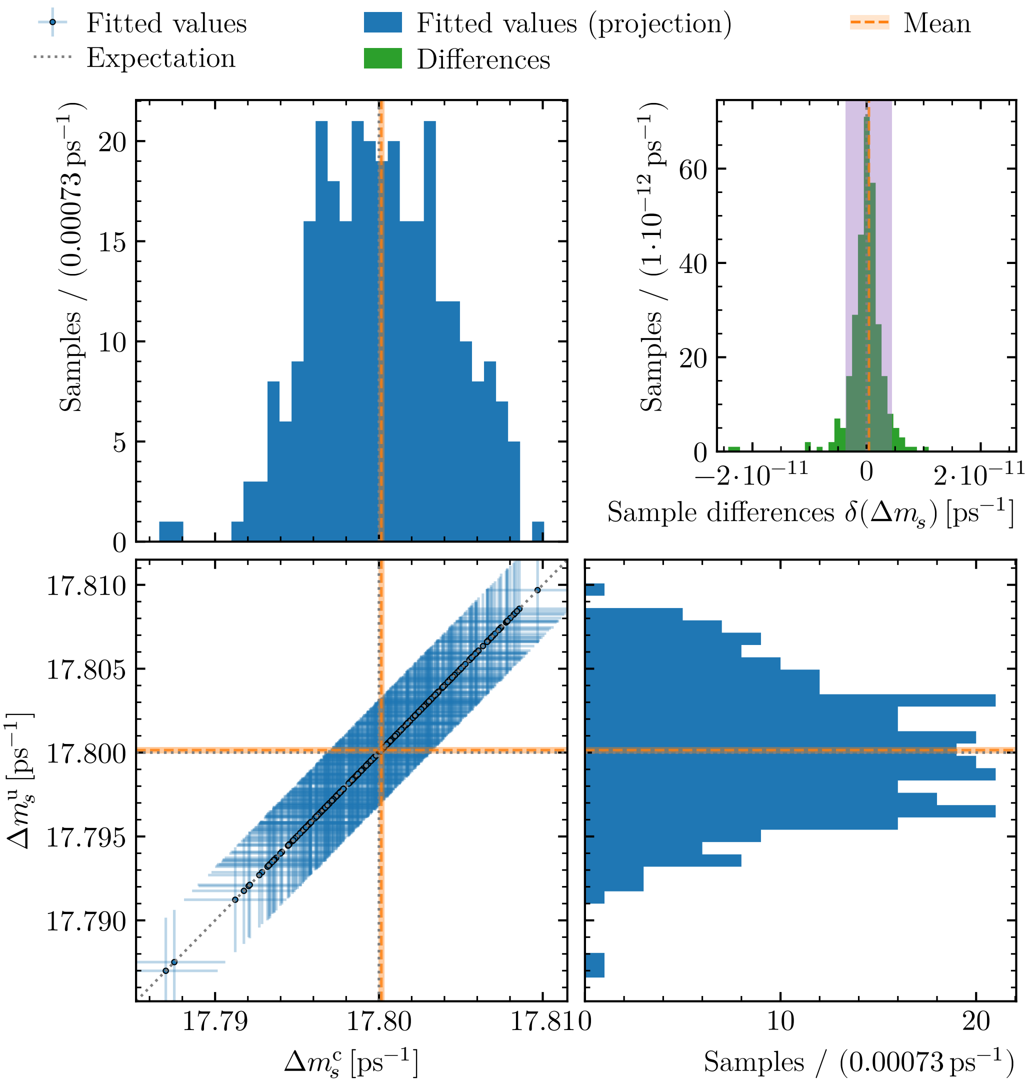


Figure 8.2: Impact of neglected correlations in absence of backgrounds as a null test. In contrast to Figure 8.1, the fit results of 300 bootstrapped samples in the presence (top left) and absence (bottom right) of correlations align well as presented at the bottom left. In both cases, the fitted mean agrees with the generated oscillation frequency. Also, the observed per-sample differences of $\mathcal{O}\left(1\cdot 10^{-12}\ \mathrm{ps}^{-1}\right)$ are negligible and the corresponding average agrees well with the null hypothesis.

9 Measurement of the oscillation frequency Δm_s

The oscillation frequency Δm_s can be directly extracted from the decay-time-dependent rates of mixed and unmixed $B_s^0 \to D_s^- \pi^+$ decays as performed by the decay-time fit in Section 7.4. However, the statistical uncertainty estimated by the decay-time fit alone does not provide a robust uncertainty estimate. To cover the uncertainty correctly systematic effects possibly impacting the measurement need to be studied as shown in Chapter 8.

In these studies, a significant bias introduced by the usage of sWeights is revealed in Section 8.3 and recovered by the implementation of an updated sWeighting strategy. After this correction, the measurement is found to be robust by performing both pseudo-experiment and data-based closure tests in Section 8.2.1. Also, variations of the applied models and parametrisations show only a minor impact on the analysis with model-dependent systematic uncertainties up to 20 % of the statistical precision.

The largest systematic effects are found in the detector resolution. These uncertainties up to 0.002 ps^{-1} are mainly introduced by the finite precision on the position and alignment of the VELO modules. While these effects are not a limiting factor to the precision of this measurement, their impact will be more significant in the future as they are introduced by mechanical limitations.

Considering all possible sources of systematic effects studied in Chapter 8, the analysis of $B_s^0 \to D_s^- \pi^+$ decays reconstructed in the LHCb Run2 data set corresponding to 6 fb^{-1} yields the most precise measurement of the oscillation frequency today

$$\Delta m_s = 17.7683 \pm 0.0051 \pm 0.0032\,\text{ps}^{-1},$$

with the first uncertainty being statistical, and the second systematic. The precision can be improved even further by combining this measurement with further results. For this combination, the previous result from the analysis of $B_s^0 \to D_s^- \pi^+$ decays in the 2011 LHCb data set [107], the analysis of $B_s^0 \to D_s^- \pi^+ \pi^- \pi^+$ decays in the full LHCb data set [57] and LHCb measurements of $B_s^0 \to J/\psi K^+ K^-$ decays [119, 120] are considered. Considering the correlations between the systematic uncertainties among the different measurements, especially from the detector-related uncertainties, the combination yields

$$\Delta m_s \;=\; 17.7656 \pm 0.0057\,\mathrm{ps}^{-1},$$

where the reported uncertainty resembles both statistical and systematic precision. This combination, which is dominated by the presented measurment, is published by Nature Physics [1], together with the presented measurement using $B_s^0 \to D_s^- \pi^+$ decays.

The oscillation frequency measured in $B_s^0 \to D_s^- \pi^+$ decays is used in the following determination of the CKM angle γ to constrain the mixing of $B_s^0 \to D_s^\mp K^\pm$ decays. In this process, the high precision in Δm_s reduces a systematic uncertainty significantly, which is a dominating effect in the previous analysis of $B_s^0 \to D_s^\mp K^\pm$ decays [14].

Review of the author's contributions

The analysis has already been progressed when I joined the analysis in 2019. Hence, the field I have contributed to the most in this measurement is the investigation of systematic effects. However, the observation of a decay-time biasing effect, introduced by the signal selection, reasoned a revision of the suppression of physics backgrounds, presented in Section 6.2. In the course of this revision, I have performed extensive studies of a variety of background decays and mass spectra. As a result of these studies, the vetoes explained in Section 6.2.2 are introduced, which significantly reduce the contamination from residual charmless backgrounds. The framework developed for these studies is later used by collaborators to perform the analogue studies for the analysis of $B_s^0 \to D_s^\mp K^\pm$ decays presented in Section 10.2.1.

My major responsibility in the analysis of $B_s^0 \to D_s^- \pi^+$ decays has been the validation of the sWeighting approach presented in Section 8.3. In the course of this study, I have adapted the full analysis to a large number of simulation-based, bootstrapped samples. This study revealed an effect which significantly biases the oscillation frequency Δm_s. This effect is corrected by the redefinition of the sWeighting PDF discussed in Section 6.3, preventing a significant bias in the measurement.

In addition, I have performed the full analysis pipeline split into subsamples as discussed in Section 8.2.1. In this context, I have gained the expert knowledge on this analysis, which allows me to study and reevaluate all details of the analysis and which is essential for the analysis of $B_s^0 \to D_s^\mp K^\pm$ decays.

Part III

Determination Of The CKM Angle γ Using $B_s^0 \to D_s^{\mp} K^{\pm}$ Decays

---

The determination of the CKM angle γ using data taken by the LHCb experiment is presented. For this measurement, the B_s^0 meson is reconstructed in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay mode, where the D_s^- candidate is reconstructed from three charged hadrons. This decay mode gives access to the CKM angle γ by the large interference between the $b \to c$ and the $b \to u$ tree-level transitions. Its long-lived final state only consists of charged hadrons allowing the B_s^0 candidate to be fully reconstructed in the LHCb data. Further, this decay mode is very similar to the previously analysed $B_s^0 \to D_s^- \pi^+$ decays both in topology and kinematics. This results in a good portability of models and calibrations, which can only be inferred from flavour-specific modes.

For this analysis, the LHCb data set recorded during LHC Run2 (2015-2018) is used, providing an integrated luminosity of $6\,\mathrm{fb}^{-1}$. The additional $3\,\mathrm{fb}^{-1}$ data set recorded during LHC Run1 (2011 & 2012) is not used in this analysis, as it has already been the subject of a previous analysis measuring $\gamma = \left(128^{+17}_{-22}\right)^\circ$ [14].

This analysis aims to extract the CKM angle γ in Chapter 13 from the five *CP* parameters. These parameters and their uncertainties are inferred from the data in three steps. At first, a time-independent analysis is performed aiming to reconstruct and select a sample of $B_s^0 \to D_s^\mp K^\pm$ decays. This part of the analysis is presented in Chapter 10 and consists of the online reconstruction, an offline selection for background suppression and a statistical separation of signal and the remaining background decays using the candidates' reconstructed masses as the discriminator. Then, the $B_s^0 \to D_s^\mp K^\pm$ sample provided by the time-independent part is used for a decay-time-dependent analysis as presented in Chapter 11. The *CP* parameters are extracted by a fit to the decay-time distribution of the sample. This fit relies on several inputs determined in the course of the analysis of $B_s^0 \to D_s^- \pi^+$ decays presented in Part II. Therefore, the portability is studied and if necessary corrections are applied. Lastly, potential sources of systematic uncertainties are discussed and studied in detail in Chapter 12.

This analysis has been performed to large extents by the author of this thesis as a member of the same group of researchers that has performed the previously presented measurement of the oscillation frequency Δm_s. The full analysis is presented, including briefly summarised studies performed by explicitly acknowledged collaborators, as the full picture of studies is essential for the understanding of the analysis.

The result of this measurement is made publicly available by the LHCb collaboration [2]. A combination with the previous result based on an independent data sample [14] is planned to be performed, before the result is submitted to a journal. The compatibility of the two results is good. However, careful handling of updated nuisance parameters in the previous result is required in the combination. This process is expected to be completed in a few months from now.

10 Extraction of the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal

A crucial part of the presented analysis is to extract the properties of the $B_s^0 \to D_s^{\mp} K^{\pm}$ decays from a background diluted data sample. Due to kinematic and topological similarities, this part of the analysis follows the strategy for signal extraction from the analysis of $B_s^0 \to D_s^- \pi^+$ decays presented in Chapter 6. In the following, the procedure is briefly summarised for completeness. The peculiarities of this analysis are discussed in detail. At first, the data samples and their preselection are introduced in Section 10.1. In the course of the preselection, also other selection steps having a large overlap with the $B_s^0 \to D_s^- \pi^+$ analysis are briefly mentioned. The suppression of physics backgrounds specific to the $B_s^0 \to D_s^{\mp} K^{\pm}$ mode is discussed in Section 10.2, before the signal extraction based on a modified model of the mass spectra for the sPlot method is presented in Section 10.3.

10.1 Processing of the event samples

In the presented analysis, both data recorded by the LHCb detector and simulated events are used. In the data, the signal mode $B_s^0 \to D_s^{\mp} K^{\pm}$ and two control modes with similar topology and kinematics ($B_s^0 \to D_s^- \pi^+$ and $B^0 \to D^{\mp} \pi^{\pm}$ decays) are reconstructed. The simulated samples cover a larger number of decay modes used to study and model both signal and backgrounds. Both data and simulation are treated equivalently, meaning both the same reconstruction algorithms and the same selections are applied.

The same data set as in the $B_s^0 \to D_s^- \pi^+$ analysis described in Section 6.1 is used, corresponding to an integrated luminosity of $\int \mathcal{L} \mathrm{d}t = 6\,\mathrm{fb}^{-1}$ of proton-proton collisions recorded at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$. The sample is divided into three subsamples, the combined sample recorded in 2015 and 2016, the sample recorded in 2017 and recorded in 2018, respectively. Each of these samples comprises data taken at both possible magnet polarities. The exact composition of the data sets in terms of magnet polarity and the year of data taking can be found in Table 6.1. In contrast to the $B_s^0 \to D_s^- \pi^+$ analysis, the $K^- \pi^+ \pi^-$ final state is also used to reconstruct $D_{(s)}^-$ candidates. This addition results in both increased signal statistics and additional backgrounds which need to be suppressed and modelled.

In total, 45 M simulated events are available for this analysis, including 13.5 M $B_s^0 \to D_s^{\mp} K^{\pm}$ generated signal events. In addition to the simulated samples used in Chapter 6, the background decays $B_s^0 \to D_s^- \rho^+$, $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- K^+$, $B^0 \to D^- K^+$, $\Lambda_b^0 \to D_s^- p$, $\Lambda_b^0 \to D_s^{*-} p$ are studied in simulation. Similar to the previous analysis, these simulated samples are primarily used to evaluate the mass shapes used in Section 10.3 and in the systematic studies of bootstrapped simulation in Section 12.4. Again, the neural-network-based PID information in the simulated samples is corrected using the PIDCalib package [83, 108]. Further, the full data selection, discussed in the following, is applied, if not explicitly stated differently.

The preselection to filter $B_{(s)}^0 \to D_{(s)}^- h^+$ decays discussed in Section 6.1.2 is applied. However, an additional requirement is introduced at the level of the L0-trigger selection, which is found to be beneficial in the course of this analysis. The L0 has either to be triggered by a track associated with the signal decay (TOS), which fulfils one of the L0 criteria for hadrons, or by a track independent of the later reconstructed signal (TIS) passing any L0 criterion. After the preselection, the multivariate classifier discussed in Section 6.1.3 is used to reduce the level of contamination from the combinatorial background.

10.2 Suppression of physics backgrounds

After the preselection and the BDT-based suppression of combinatorial background, a suppression of physics backgrounds is applied. The sources of backgrounds differ from the processes investigated in Section 6.2 due to the different types of accompanying hadron. Additionally, the overall level of contamination increased due to the smaller branching fraction of $B_s^0 \to D_s^{\mp} K^{\pm}$ with respect to the background processes. For better reproducibility, more details on the selection discussed in the following are given in Appendix B.1.

The PID requirement applied to the accompanying hadron is inverted to reject pions and to accept kaons. The requirements on the four final-state hadrons and on the D_s meson decay are not changed. The additional $D_s^- \to K^- \pi^+ \pi^-$ final state shows large contributions from misidentified $\overline{\Lambda}_c^- \to \overline{p} \pi^+ \pi^-$ and doubly misidentified

$D^- \to \pi^- K^+ \pi^-$ decays. Therefore tight PID requirements are applied to both pions suppressing contributions from misidentified kaon and protons and a high kaon likelihood is expected from the fina-tate kaon.

A revised version of the selection used in the Run1 analysis [14] is utilised to suppress specific backgrounds. This selection makes use of the finite lifetime of the D_s^- meson. It is observed that the cuts on the D_s^- flight distance χ^2_{FD} cause a bias of the sample's decay-time distribution. In comparison to the $B_s^0 \to D_s^- \pi^+$ analysis, the level of contamination is increased due to the additional Cabibbo suppression of the signal mode. Hence, this selection cannot be completely omitted. A loose $\chi^2_{\text{FD}} > 2$ requirement is applied to the $D_s^- \to K^- K^+ \pi^-$ modes, while a tighter $\chi^2_{\text{FD}} > 9$ cut is applied to the more contaminated $D_s^- \to K^- \pi^+ \pi^-$ and $D_s^- \to \pi^- \pi^+ \pi^-$ modes. The effects introduced by the resulting decay-time bias are studied in Section 11.2 and 12.5. In the following, the strategy to veto backgrounds under a variation of the mass hypotheses is revised and potentially residual charmless backgrounds are studied in detail.

10.2.1 Variation of the mass hypotheses and explicit vetoes

The dominant contribution of physics backgrounds comes from the presence of intermediate states with charmed hadrons other than D_s^- mesons. The exact process and its suppression are dependent on the D_s^- decay mode.

Contributions from $B^0_{(s)} \to J/\psi X$ decays are vetoed by a suppression of the charmonium resonance. This background originates from the double misidentification of leptons from $J/\psi \to l^+ l^-$ decays, where one lepton is identified as accompanying kaon and the other one as oppositely charged pion. The $J/\psi \to \mu^- \mu^+$ process is sufficiently suppressed by requiring the absence in the muon system of the corresponding tracks, the $J/\psi \to e^- e^+$ process by loose $\text{PID}_e < 5$ cuts on accompanying kaon and oppositely charged pion, as in Section 6.2.

In the $D_s^- \to K^- K^+ \pi^-$ mode backgrounds involving intermediate D^- mesons and $\overline{\Lambda}_c^-$ baryons can contribute via the misidentification of a pion from $D^- \to K^- \pi^+ \pi^-$ decays or of a proton from $\overline{\Lambda}_c^- \to \overline{p} K^+ \pi^-$ decays as a kaon. These backgrounds are suppressed by a tightened $\text{PID}_K > 10$ ($\text{PID}_K - \text{PID}_p > 5$) criterion for kaon carrying the same charge as the D_s^- candidate, if it is reconstructed in a $30\,\text{MeV}/c^2$ window around the known D^- ($\overline{\Lambda}_c^-$) mass under the corresponding mass hypothesis. A further source of background is the presence of intermediate D^0 states. These are rejected if the invariant mass of both kaons is above a $1800\,\text{MeV}/c^2$ threshold or if the invariant mass of the accompanying kaon (under a pion hypothesis) and the oppositely charged kaon is within a $30\,\text{MeV}/c^2$ window around the known D^0 mass.

The D^0 veto also suppresses $B_{(s)}^0 \to D^{*-}h^+$ backgrounds, as the D^{*-} dominantly decays via $D^{*-} \to D^0\pi^-$.

The $D_s^- \to K^-\pi^+\pi^-$ mode is subject to $\overline{\Lambda}_c^-$ and D^- backgrounds as well. Contributions from $\overline{\Lambda}_c^- \to \overline{p}\pi^+\pi^-$ decays, where the proton is misidentified as a kaon, are again suppressed by a tightened $\mathrm{PID}_K - \mathrm{PID}_p > 5$ cut if the reconstructed $\overline{\Lambda}_c^-$ mass under the corresponding mass hypothesis is within a $30\,\mathrm{MeV}/c^2$ window around its known value. Contamination from $D^- \to K^-K^+\pi^-$ originates from misidentifying the kaon as a pion. This background is vetoed by excluding a $30\,\mathrm{MeV}/c^2$ mass window around the known D^- mass after reconstructing the mass under exchange of the pion and kaon mass hypotheses. The D^0 backgrounds here are dominated by $D^0 \to \pi^+K^-$ decays. Hence, the invariant mass of kaon and pion is required to be below $1750\,\mathrm{MeV}/c^2$. Misidentification including the accompanying kaon is suppressed by explicitly rejecting the invariant mass of the accompanying kaon under the pion hypothesis and of the other kaon within a $30\,\mathrm{MeV}/c^2$ window around the D^0 mass.

In the $D_s^- \to \pi^-\pi^+\pi^-$ mode, all contributions from misidentified $\overline{\Lambda}_c^- \to \overline{p}K^+\pi^-$ decays are already rejected by the PID criteria. Two processes are suppressed to control the D^0 contamination of the D_s^- candidates. The $D^0 \to \pi^+\pi^-$ decays are rejected by requiring the invariant mass of two oppositely charged pion to be below $1700\,\mathrm{MeV}/c^2$. The $D^0 \to K^+\pi^-$ decays are vetoed in a $30\,\mathrm{MeV}/c^2$ range around the known D^0 mass in the invariant mass spectrum of the accompanying kaon and an oppositely charged pion.

10.2.2 Investigation of charmless backgrounds

Potential residual contamination from charmless $B_{(s)}^0 \to h^+h^-h^+h^-$ decays is studied in a similar way to what has been discussed in Section 6.2 using the upper D_s^- mass sideband $m(h^\mp h^+h^-) \in [2015,\, 2070]\,\mathrm{MeV}/c^2$. This study [121] has been performed using the framework developed in Section 6.2.3. Significant contributions from B^0 decays are found in the D_s^- sideband of the $D_s^- \to K^-K^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$ modes and a significant B_s^0 contribution is found in the $D_s^- \to K^-\pi^+\pi^-$ sideband.

A variety of resonances is observed in the invariant mass spectra of two-particle combinations involving the accompanying kaon. All modes are contaminated by residual D^0 and K^{*0} decays. In the $D_s^- \to K^-K^+\pi^-$ and $D_s^- \to K^-\pi^+\pi^-$ modes these are vetoed by rejecting candidates with an invariant mass of the accompanying kaon and an oppositely charged kaon (pion) within a $30\,\mathrm{MeV}/c^2$ window around the known D^0 (K^{*0}) mass.

In the $D_s^- \to \pi^-\pi^+\pi^-$ mode, the same K^{*0} veto is applied, while the D^0 contribution is excluded using a $30\,\mathrm{MeV}/c^2$ window around the known mass in the combination of the accompanying kaon and an oppositely charged pion. Additionally, D^- decays are rejected in a $30\,\mathrm{MeV}/c^2$ window around the D^- mass in the invariant mass of the accompanying kaon and both oppositely charged pions. Further, this mode is subject to a variety of backgrounds from light resonances. The contaminated phase space is excluded by requiring the invariant mass of accompanying kaon and oppositely charged pion to be above a $1\,\mathrm{GeV}/c^2$ threshold.

While in the $D_s^- \to K^-\pi^+\pi^-$ mode, a small contribution of charmless B^0 decays is effectively removed by these vetoes, in all D_s^- modes the dominant charmless contributions are only slightly reduced. In contrast to the $B_s^0 \to D_s^-\pi^+$ analysis, these contributions are not moved out of the fit range sufficiently by applying the D_s^- mass constraint. Therefore, the decay-time biasing requirements on the D_s^- flight distance have to be applied to control those contributions. The $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$ modes, which show the largest level of contamination, are required to fulfil $\chi^2_{\mathrm{FD}} > 9$. In the $D_s^- \to K^-K^+\pi^-$ modes with less contamination a looser $\chi^2_{\mathrm{FD}} > 2$ requirement is applied to reduce the induced bias.

10.3 Multidimensional mass fit

The $B_s^0 \to D_s^\mp K^\pm$ signal is statistically separated from residual backgrounds passing the previous selection using the sPlot approach introduced in Section 5.3. Here, the good mass resolution of the LHCb experiment is utilised to use the reconstructed $m(D_s^-\pi^+)$ and $m(h^\mp h^+ h^-)$ masses as a discriminant. Similar to Section 6.3, the two mass distributions are fitted simultaneously for all subsamples. Due to the different sources and level of background contamination compared to $B_s^0 \to D_s^-\pi^+$ the fit model has been revised [121, 122]. The fitted model is displayed in Figure 10.1 together with the corresponding mass distributions.

The signal model $\mathcal{S}(\vec{m})$, described by the sum of a Hypatia and a Johnson SU function [111, 112], is not changed. Also the combinatorial background $\mathcal{C}(\vec{m})$ is modelled by the sum of the signal PDF and an exponential function in the $m(h^\mp h^+ h^-)$ mass as previously discussed. To increase the stability with the decreased statistics, the combinatorial model is simplified in the $m(D_s^\mp K^\pm)$ spectrum. In the $D_s^- \to K^-K^+\pi^-$ modes, the smaller exponential slope is fixed to $c^2_{B_s^0} = 0.0$, yielding effectively the sum of an exponential and a constant function. The $D_s^- \to \pi^-\pi^+\pi^-$ and the additional $D_s^- \to K^-\pi^+\pi^-$ modes are sufficiently described by a single exponential function. The fixed parameters of the signal and combinatorial PDFs are reevaluated based on simulation and data sideband, respectively.

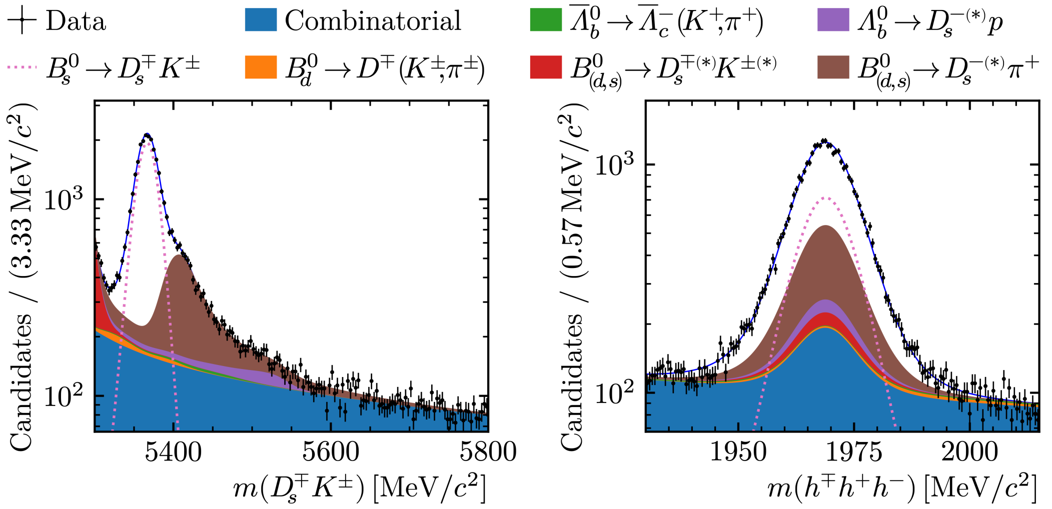


Figure 10.1: Fitted mass distributions of the $B_s^0 \to D_s^\mp K^\pm$ candidates, on the left the $m(D_s^\mp K^\pm)$ (left), on the right the $m(h^\mp h^+ h^-)$ spectrum. The data distributions and the fitted PDFs shown here are merged among all subsamples

Next to signal and combinatorial background, a variety of residual physics backgrounds $\mathcal{B}_i(\vec{m})$ are modelled. The $B^0 \to D_s^\mp K^\pm$ background decays to the exactly same final state as the signal. This background is modelled by a double-sided Hypatia function with the shape parameters extracted from simulation in the $m(D_s^\mp K^\pm)$ spectrum and by the signal PDF in $m(h^\mp h^+ h^-)$. The yield of this background is linked to the signal by a floating parameter shared among all subsamples

$$N_{B^0 \to D_s^\mp K^\pm} = f_1 \cdot N_{B_s^0 \to D_s^\mp K^\pm} \,. \tag{10.1}$$

Another source of backgrounds are decays into a $D_s^- \pi^+$ final state, where the pion is misidentified as kaon. In the $m(D_s^\mp K^\pm)$ mass the fully-reconstructed $B_s^0 \to D_s^- \pi^+$ background is modelled similarly to the signal by the sum of an Hypatia and a Johnson SU function. Further, the partially-reconstructed $B_s^0 \to D_s^{*-} \pi^+$ component is described by a double-sided Hypatia function, while the $B_s^0 \to D_s^- \rho^+$ component is sufficiently described by a double exponential function, as it only contributes with its upper-mass tail to the fit region. The corresponding shape parameters are fixed to values found in simulation. All these $D_s^- \pi^+$-like backgrounds contribute as additional sources of signal in the $m(h^\mp h^+ h^-)$ spectrum. The same is true for the $D_s^- p$-like backgrounds, where the proton is misidentified as a kaon. Here, the $m(D_s^\mp K^\pm)$ is described by a template from corrected simulation for both the fully-reconstructed $\Lambda_b^0 \to D_s^- p$ and the partially-reconstructed $\Lambda_b^0 \to D_s^{*-} p$ backgrounds. All $D_s^- \pi^+$-like and $D_s^- p$-like components share a common yield floating in the fit for better stability. The floating fraction $f_{\pi/p}$ modulates the contributions of the two groups and is shared among all subsamples. The fractions $f_{21} = 0.7664$ and $f_{22} = 0.8138$ of the individual $D_s^-\pi^+$-like components are fixed to values based on the efficiency-corrected yields observed in the branching ratio measurement of $B^0 \to D_s^-\pi^+$ decays [123]. The fraction between the two D_s^-p-like components is set to $f_{31} = 0.75$ following predictions.

Further misidentified backgrounds modelled in the PDF are $B^0 \to D^-h^+$ and $\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^- h^+$ decays, where the accompanying hadron is either a kaon or a misidentified pion. These backgrounds are each modelled by simulation-based templates in both the $m(D_s^\mp K^\pm)$ and the $m(h^\mp h^+ h^-)$ spectra. As in the $B_s^0 \to D_s^-\pi^+$-mass fit, described in Section 6.3, the corresponding yields are fixed. The fixed yield of the $B^0 \to D^\mp\pi^\pm$ component is based on fits to the in data reconstructed $m\left(D^-\pi^+\right)$ mass and corrected by simulation-based efficiencies. The $B^0 \to D^-K^+$ yield is then derived using the relative branching ratio of both modes and is again efficiency corrected. Similarly, the fixed $\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^-\pi^+$ and $\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^- K^+$ yields are estimated from a fit to the reconstructed $m\left(\overline{\Lambda}{}_c^-\pi^+\right)$ masses. The full extended PDF used for the fit is defined as

$$
\begin{aligned}
\mathcal{P}(\vec{m}) = N_{\text{sig}} \cdot \mathcal{S}(\vec{m}) + N_{\text{comb}} \cdot \mathcal{C}(\vec{m}) + \left(f_1 \cdot N_{\text{sig}}\right) \cdot \mathcal{B}_{B^0 \to D_s^\mp K^\pm}(\vec{m}) \\
+ N_{D_s^-\pi^+/p} \cdot \Big(f_{\pi/p} \cdot \big(f_{21} \cdot \mathcal{B}_{B_s^0 \to D_s^-\pi^+}(\vec{m}) \\
+ (1 - f_{21})\, f_{22} \cdot \mathcal{B}_{B_s^0 \to D_s^{*-}\pi^+}(\vec{m}) \\
+ (1 - f_{21})\,(1 - f_{22}) \cdot \mathcal{B}_{B_s^0 \to D_s^-\rho^+}(\vec{m}) \big) \\
+ \left(1 - f_{\pi/p}\right) \cdot \big((1 - f_{31}) \cdot \mathcal{B}_{\Lambda_b^0 \to D_s^- p}(\vec{m}) \\
+ (1 - f_{31}) \cdot \mathcal{B}_{\Lambda_b^0 \to D_s^{*-} p}(\vec{m}) \big) \Big) \\
+ N_{B^0 \to D^- K^+} \cdot \mathcal{B}_{B^0 \to D^- K^+}(\vec{m}) + N_{B^0 \to D^\mp \pi^\pm} \cdot \mathcal{B}_{B^0 \to D^\mp \pi^\pm}(\vec{m}) \\
+ N_{\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^- K^+} \cdot \mathcal{B}_{\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^- K^+}(\vec{m}) + N_{\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^- \pi^+} \cdot \mathcal{B}_{\overline{\Lambda}{}_b^0 \to \overline{\Lambda}{}_c^- \pi^+}(\vec{m}) \ . \qquad (10.2)
\end{aligned}
$$

The fitted yields and the floating yield fractions are listed in Table 10.1. The overall $B_s^0 \to D_s^\mp K^\pm$-signal yield of all subsamples combined is $N_{\text{sig}} = 20\,249 \pm 180$. For completeness, the fitted values for the shape parameters floating in the fit are given in Appendix B.2. These include the exponential slopes $c_{B_s^0/D_s}$ and the fractions $f_{B_s^0/D_s}$ of the two components modelling the combinatorial background and the widths $\sigma^H_{B_s^0/D_s}$ and means $\mu_{B_s^0/D_s}$ of the signal Hypatias.

In contrast to the $B_s^0 \to D_s^-\pi^+$ analysis (Part II), the fitted nominal PDF is used for the extraction of sWeights. The option to repeat the fit for the sWeighting with a signal component and a combined background component is also studied.

However, in this analysis the pseudo-experiment-based studies in Section 12.5 show a biasing effect on the *CP* parameters in the presence of this modification. Further, the effect of neglecting correlations among the observables of the two fits studied in Section 12.4 is larger in the presence of this modification.

Table 10.1: The three fitted yields split among the individual subsamples. The yields of further background components are constrained as described in the text. The total signal yield combining all subsamples is $N_{\text{sig}} = 20\,249 \pm 180$.

Parameter	Year	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to (K^-K^+\pi^-)_{\text{NR}}$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
N_{sig}	2015 & 16	2470 ± 62	1889 ± 53	1223 ± 44	458 ± 30	1067 ± 44
	2017	2232 ± 60	1655 ± 48	1132 ± 40	417 ± 27	881 ± 39
	2018	2681 ± 62	1946 ± 54	1242 ± 44	539 ± 31	1117 ± 42
N_{comb}	2015 & 16	1055 ± 74	789 ± 53	411 ± 41	214 ± 33	402 ± 48
	2017	1042 ± 65	825 ± 45	428 ± 40	182 ± 30	360 ± 42
	2018	1308 ± 55	828 ± 56	604 ± 46	184 ± 33	519 ± 49
$N_{D_s^-\pi^+/p}$	2015 & 16	879 ± 72	706 ± 57	1224 ± 51	1539 ± 49	2170 ± 66
	2017	555 ± 75	486 ± 44	920 ± 46	1218 ± 46	1835 ± 59
	2018	593 ± 46	710 ± 59	1116 ± 53	1446 ± 51	1987 ± 65
f_1	All			0.0385 ± 0.0026		
$f_{\pi/p}$	All			0.902 ± 0.015		

11 Decay-time-dependent analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays

---

In this analysis, the CKM angle γ is constrained by measuring six *CP* parameters introduced in Chapter 3. Here, these parameters are extracted from data by measuring the decay-time-dependent decay rates. The used data providing a statistically clean sample of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays is prepared by the decay-time-independent analysis discussed in Chapter 10.

The time-dependent analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ shares many similarities with the analysis of $B_s^0 \to D_s^- \pi^+$ decays presented in Chapter 7. However, some effects cannot be properly disentangled anymore, due to the increased number of interfering amplitudes. In this case, external inputs are needed. Most of these inputs are provided from the decay-time-dependent analysis of $B_s^0 \to D_s^- \pi^+$ decays presented in Chapter 7. However, the corresponding decay-time fit is repeated according to small updates in the analysis strategy. In all cases, the signal selection is updated with the additional L0 and X^2_{FD} requirements introduced in Chapter 10. The decay-time uncertainty estimate is calibrated using an updated model in Section 11.1. The strategy of evaluating the alignment-induced decay-time bias is not updated, but the portability between the two decay modes is investigated in Section 11.2. The parametrisation of the decay-time acceptance is extended with additional splines and corrected to account for kinematic differences between the modes as described in Section 11.3. The portability of the flavour-tagging calibration, which is Gaussian-constrained in the decay-time fit, is studied in Section 11.4.

11.1 Calibration of the decay-time uncertainty estimate

The *CP* parameters are directly attached to the fast B_s^0-flavour oscillation in the time-dependent decay rates of $B_s^0 \to D_s^\mp K^\pm$. Hence, it is necessary to resolve the oscillation properly to maintain sensitivity to the *CP* parameters. In the decay-time fit performed in Section 11.5, the diluting effects of the finite detector resolution are accounted by the convolution of the decay-time PDF with Gaussian resolution function $\mathcal{R}(t \,|\, \Delta t,\, \sigma_t)$. Here, the per-event resolution $\sigma_t(\delta t)$ is obtained from a calibration of the decay-time uncertainty estimate δt, while the decay-time bias Δt is studied in Section 11.2.

The calibration of the uncertainty estimate is obtained following Section 7.1 with an updated fit model increasing the stability. The observed decay-time resolution is fitted in ten bins of the per-event uncertainty estimate. Then, the linear relation between the true resolution and the uncertainty estimate is evaluated by performing a fit of Equation 7.3.

The true resolution is again measured using prompt $D_s^- \pi^+$ candidates from data. The prompt $D_s^- \pi^+$ sample is used as it provides larger statistics in comparison to $D_s^\mp K^\pm$. Further, studies [122] comparing samples of selected $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ simulation prove the calibration to be independent of the accompanying-hadron species. The $D_s^- \to \phi\pi^-$ mode is used solely as it is both the cleanest and the most abundant mode, while the compatibility among the decay modes is shown in the previous.

The resolution of the prompt candidates is estimated using an updated fit strategy. The decay-time distribution of the prompt candidates is fitted in a symmetric $t \in [-1000,\, 1000]$ fs interval. This improves the stability but requires to describe backgrounds from long-lived particle decays. Consequently, the fit model is updated. The resolution of the prompt candidates is described by the sum of two Gaussian functions with shared mean μ. Backgrounds from wrong PV associations of the prompt candidates are modelled by a symmetrised exponential function, while contributions from SVs of long-lived decays, dominantly affecting positive decay times, are accounted for by an exponential function. Overall, the updated PDF

$$\begin{aligned} \mathcal{P}_t'(t) = N_{\text{res}} \cdot [f_1 \cdot \mathcal{G}(t \,|\, \mu,\, \sigma_1) + (1 - f_1) \cdot \mathcal{G}(t \,|\, \mu,\, \sigma_2)] \\ + N_3 \cdot \mathrm{e}^{-|t|/\tau_3} + N_4 \cdot \mathrm{e}^{-t/\tau_4} \end{aligned} \tag{11.1}$$

is used to model the decay-time distribution of the prompt candidates. The normalisation N_i, the exponential slopes τ_i, the Gaussian widths σ_i and mean μ as well as the relative fractions f_i are floating in the fits.

The two fitted Gaussian cores are then combined into an effective resolution, following Equation 7.4 and 7.5, for each bin and data-taking period. The fit of the estimated width as a function of the uncertainty estimate yields the calibration parameters reported in Table 11.1.

Table 11.1: Calibration parameters of the decay-time uncertainty estimate and average uncertainty estimate extracted from prompt candidates for the different data-taking periods.

Sample	p_0 [fs]	p_1	δt [fs]
2015 & 16	12.5 ± 2.8	0.950 ± 0.064	47.21
2017	9.9 ± 2.7	0.955 ± 0.061	46.08
2018	9.8 ± 2.7	0.948 ± 0.060	45.99

11.2 Correction of the decay-time bias: Portability

The decay-time bias introduced by the misalignment of the VELO needs to be carefully corrected, as this bias has a significant impact on the fitted *CP* parameters. In this analysis, the correction estimated in Section 7.2 is used without modifications. To justify this approach, the portability between the $B_s^0 \to D_s^- \pi^+$ channel, used to estimate the correction, and the $B_s^0 \to D_s^\mp K^\pm$ signal has to be shown.

The propagation of the misalignment into the decay-time bias decay is known to be dependent on the decay kinematics. Because of this, a calibration is necessary to transfer the bias determined using prompt candidates to the signal sample. Hence, the kinematic properties, in form of the candidates' momentum $p(D_s^- h^+)$ and the opening angle of the accompanying hadron $p_x(h^+)/p_z(h^+) = \tan\alpha(h^+)$, are compared between, the prompt $D_s^- \pi^+$, the $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ samples. The distributions observed in the data show small differences in the momentum between the three samples. The opening angle shows only small differences between the two decay modes, while the prompt sample provides significantly smaller opening angles.

The kinematic differences of the prompt sample with respect to the resonant samples are expected and accounted for by the bias calibration. However, the differences between the two decay modes are undesired and require further investigation. The origin of this deviation is found in the different PID criteria applied to the accompanying hadron. While $\mathrm{PID}_K > 5$ is required for the accompanying kaon, the requirement $\mathrm{PID}_K < 0$ is applied to the accompanying pion. Removing these requirements yields matching kinematic properties for the two decay modes.

The kinematic requirements in the selection cannot be omitted. Hence, the criteria are symmetrised by a tighter $\mathrm{PID}_K < -5$ requirement on the accompanying pion. With this tighter pion selection, the kinematic properties of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^\mp K^\pm$ show good agreement with only minor differences. Even the momentum distribution of the prompt sample matches the resonant samples with this modification, while the prompt candidates still provide significantly smaller opening angles.

To study the effect of the kinematic differences introduced by the different PID requirements, the calibration is reevaluated under the tighter pion criterion and found as

$$\Delta t_{\mathrm{sig}}^{\mathrm{tight}}(\Delta t_{\mathrm{pr}}) = (0.119 \pm 0.216)\,\mathrm{fs} + (0.504 \pm 0.035) \cdot \Delta t_{\mathrm{pr}}\,. \tag{11.2}$$

This calibration is in good agreement with the nominal calibration reported in Equation 7.6. Hence, the application of the decay-time bias correction extracted in Section 7.2 is justified. The systematic effects introduced by this correction are studied in Chapter 12.

11.3 Adaption of the decay-time-dependent efficiency

A decay-time-dependent efficiency is introduced by the finite detector acceptance, reconstruction effects and selection. This effect needs to be modelled to properly analyse the time-dependent decay rates. Similar to the $B_s^0 \to D_s^- \pi^+$ decay-time fit described in Section 7.4, the acceptance is modelled by B-spline functions. It is possible to keep the coefficients parameterising the acceptance model floating in the fit of $B_s^0 \to D_s^- \pi^+$ decays, where the decay rates are reduced to the form of Equation 3.41. This is not possible in the analysis of $B_s^0 \to D_s^\mp K^\pm$ decays, where the *CP* parameters D_f and $D_{\bar{f}}$ of the hyperbolic terms present in the full decay rates are strongly correlated to the acceptance. The information on the acceptance effects has to be externally provided. Here, the parametrisation found in the $B_s^0 \to D_s^- \pi^+$ fit is used, while it is extended to accommodate eleven spline functions instead of eight as previously reported. This update improves the stability of the model.

The $B_s^0 \to D_s^- \pi^+$ signal sample provides the same decay topology, and similar kinematics and is selected with very similar requirements. The residual differences in the kinematics and the selection, especially from the PID requirement, are accounted for using a simulation-based correction. Every coefficient c_i is taken from the $B_s^0 \to D_s^- \pi^+$ fit and corrected by the ratio v_i of the coefficients found in the $B_s^0 \to D_s^\mp K^\pm$ and $B_s^0 \to D_s^- \pi^+$ simulation, respectively.

The simulated samples are reweighted to reproduce the kinematics observed in the data. The individual subsamples, namely the D_s^--decay mode, the magnet polarity and the data-taking period, are weighted to resemble the proportions present in the data. Hence, effects from the addition of the $D_s^- \to K^-\pi^+\pi^-$ mode, which is neglected in the $B_s^0 \to D_s^-\pi^+$ analysis, are also incorporated. Further, the simulated samples have to pass the full selection corresponding to the decay mode, including the different PID requirements on the accompanying hadron.

The knowledge of the generated physics parameters in the simulated samples, allows these parameters to be fixed in the decay-time fit and to extract the acceptance in both modes. As in data, the acceptance is fitted for each of the three data-taking periods. It is not distinguished between the different D_s^--decay modes, as the corresponding decay-time distributions do not show significant deviations. The ratios $v_i = c_i^{D_s^{\mp}K^{\mp}} / c_i^{D_s^-\pi^+}$ calculated from the acceptance coefficients fitted in the simulation. The deviation from unity is small for all ratios, proving the hypothesis of similar acceptance effects in both samples. However, a clear trend is observed, where the ratio decreases in the coefficients dominated by smaller decay times. This hints at a less efficient reconstruction or selection of $B_s^0 \to D_s^{\mp}K^{\pm}$ at low decay times with respect to $B_s^0 \to D_s^-\pi^+$.

A cross-check on the effect of the PID requirement is performed. The selection applied to the accompanying kaon is loosened to $\mathrm{PID}_K > 0$, symmetric to the $\mathrm{PID}_K < 0$ requirement on the accompanying pion. In this scenario, all ratios show only negligible deviations from unity, while the trend in the ratios is vanished. This supports the hypothesis of the difference in the decay-time-dependent efficiency being dominated by the PID requirements.

11.4 Constraining the flavour tagging calibration

In this analysis, it is necessary to distinguish between the four decay-time-dependent decay rates defined in Equation 3.17 to Equation 3.20. While the final state f is defined by the charge of the accompanying kaon, flavour tagging is needed to determine the initial B_s^0 state q_i. Hence, flavour tagging plays an important role.

The pre-calibration of the individual OS taggers shows only negligible effects on the OS combination in the $B_s^0 \to D_s^-\pi^+$ analysis. Therefore, it is assumed that these pre-calibrations, extracted from reweighted $B^+ \to J/\psi K^+$ data in Section 7.3, can also be used here. The pre-calibrations are applied using the EPM, which is also used to perform the combination of the OS taggers.

In Section 7.4 it is possible to retrieve the mistag calibration of the two taggers, the OS combination and the SS kaon, from the decay-time fit of $B_s^0 \to D_s^- \pi^+$ decays. This is not possible in the case of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. For the calibration information of the B_s^0-meson flavour at its decay is needed. While it can be derived from the final state in the flavour-specific case of $B_s^0 \to D_s^- \pi^+$, this information is not available in $B_s^0 \to D_s^{\mp} K^{\pm}$. Further, the flavour tagging dilutes the oscillation amplitudes. While the oscillation frequency Δm_s is independent of the amplitudes, the *CP* parameters studied here are directly dependent. Therefore, an external calibration is needed.

In Section 11.5, the flavour tagging is constrained using the calibration extracted in the decay-time fit of $B_s^0 \to D_s^- \pi^+$ data in Section 7.4. The topological and kinematic similarities between the two decay modes promise good portability of calibration, which is studied in Section 11.4.1. The limited knowledge of the calibration is incorporated by applying a Gaussian constraint instead of fixing the parameters. This way, the calibration uncertainties are propagated directly into the statistical uncertainties of the *CP* parameters extracted in the decay-time fit. To account for the correlations among the parameters, two multivariate Gaussian functions are used in the constraint based on the fitted covariance of the four parameters in each calibration. Since the model-dependent systematic effects discussed in Section 8.2 could also affect the flavour tagging calibration, these studies are reviewed in terms of the extracted calibration parameters in Section 11.4.2.

11.4.1 Portability studies

The tag decision and the mistag estimate provided by flavour tagging algorithms are independent of the B_s^0 final state. Further, the topology and kinematics of $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ decays are very similar. Therefore, good portability of the flavour tagging calibration between the two modes is expected. However, the two data samples are known to show small differences in their kinematics due to different selection requirements, especially in the perspective of the accompanying hadron's PID selection. Hence, the portability among the two modes is studied using simulated samples, which allow a flavour tagging calibration to be performed, based on the known simulated production flavour in both modes.

A few representative properties of the two simulated samples are compared in Figure 11.1. The number of tracks n_{tracks} present in the simulated events is a good measure for the detector occupancy in the selected events, which is known to have a significant impact on the flavour tagging performance. No significant difference can be observed here between the two decay modes. However, the kinematic distributions show small differences. Both the transversal momentum p_{T} and the pseudorapidity

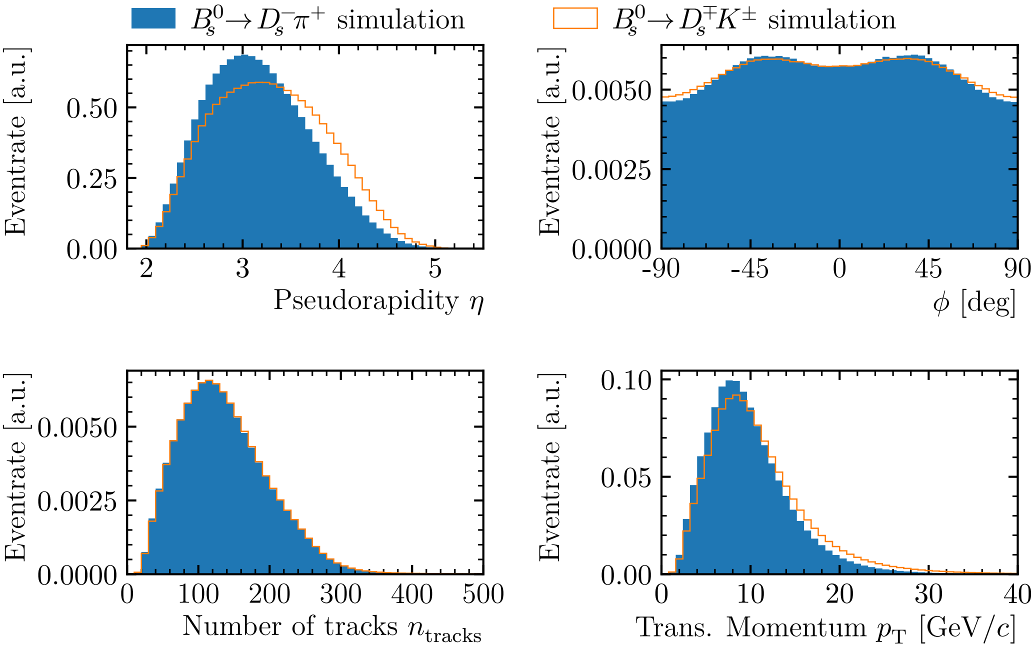


Figure 11.1: Comparison of flavour-tagging related properties of the simulated $B_s^0 \to D_s^- \pi^+$ (blue) and $B_s^0 \to D_s^\mp K^\pm$ (orange) samples. Due to the large statistics of the simulated samples, the uncertainties of the distributions are negligible.

of the B_s^0 candidates are shifted towards smaller values in $B_s^0 \to D_s^\mp K^\pm$ compared to $B_s^0 \to D_s^- \pi^+$. The angle $\phi = \arctan p_y/p_x$, defining the direction of flight in the xy plane, shows a slightly narrower distribution in the $B_s^0 \to D_s^\mp K^\pm$ sample, which hints at acceptance effects.

The mistag predictions $\eta^{\text{OS/SS}}$ presented in Figure 11.2 show similar distributions for both decay modes. However, the $B_s^0 \to D_s^- \pi^+$ provides a slightly larger fraction of events with small mistag predictions. This behaviour is expected as the flavour tagging performance is known to improve with increasing transversal momentum. The different fractions of high-quality tags lead to the expectation of slightly decreased tagging powers in $B_s^0 \to D_s^\mp K^\pm$ compared to the $B_s^0 \to D_s^- \pi^+$ analysis. However, this does not necessarily impact the calibration. To study the effect on the calibration, the lhcb_ftcalib package [87] is used to perform calibrations based on the known initial B_s^0 flavour in simulation. The extracted calibrations are displayed in Figure 11.3 and the corresponding parameters are given in Appendix B.3 for completeness. Only minor differences can be observed. Hence, the portability of the flavour tagging calibration from $B_s^0 \to D_s^- \pi^+$ to $B_s^0 \to D_s^\mp K^\pm$ is concluded to be validated. The residual differences are accounted as systematic effects on the calibration in Section 11.4.2.

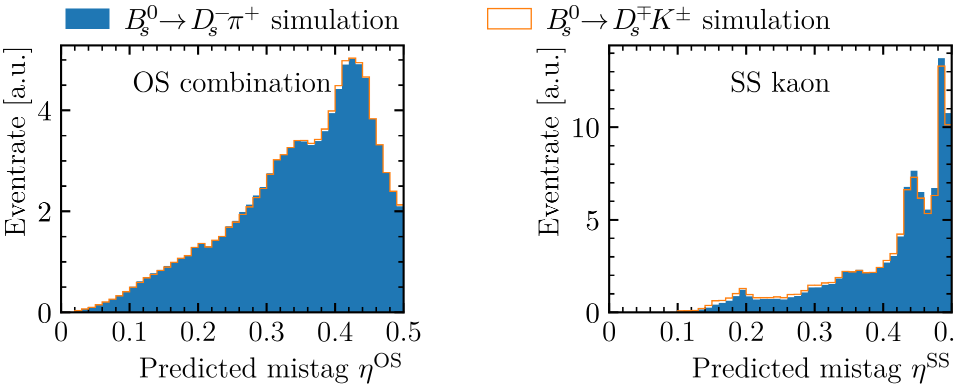


Figure 11.2: Comparison of the mistag probability predicted by the OS combination (left) and the SS kaon tagger (right) for simulated $B_s^0 \to D_s^- \pi^+$ (blue) and $B_s^0 \to D_s^{\mp} K^{\pm}$ (orange) samples.

11.4.2 Systematic effects to the calibration

As the flavour tagging calibration parameters in the decay-time fit are Gaussian constrained, the uncertainties of the corresponding parameters are of special interest. Thus, systematic effects on these parameters are studied. Different systematic effects are considered and tested by modifications to the decay-time fit presented in Section 7.4.

In a number of checks negligible effects on the calibration are found. The effect from the mass fit model is checked by varying the low-mass background fraction $f_{\mathrm{low}} \in \{0.0, 0.5, 1.0\}$, while the acceptance coefficients c_1, c_2 and c_3 are fixed to values $\pm 10\,\%$ of their nominal values to study the impact of the acceptance. Both studies show no effect on the flavour tagging calibration. Besides, negligible effects are found, when the floating detection asymmetry a_{det} is fixed to 0.0 or when the fixed tagging efficiencies $\varepsilon_{\mathrm{tag}}^{\mathrm{OS/SS}}$ are allowed to float in the fit.

Sizeable effects are found by varying the decay-time resolution model and in the analyses of subsamples. The resolution model is varied by using the calibration based on the wide or narrow Gaussian core as explained in Section 8.2. The largest deviation with respect to the nominal calibration parameter found in the alternative fits is taken as systematic. The effects introduced by different subsamples are also evaluated following Section 8.2. Here, the decay-time fit is performed split in the different D_s^--decay modes, the two magnet polarities and in two bins of the B_s^0 momentum and the BDT classification each. The fitted calibration parameters are averaged over each set of subsamples and the average is then compared to the nominal value. The deviation found in each split is considered a systematic effect.

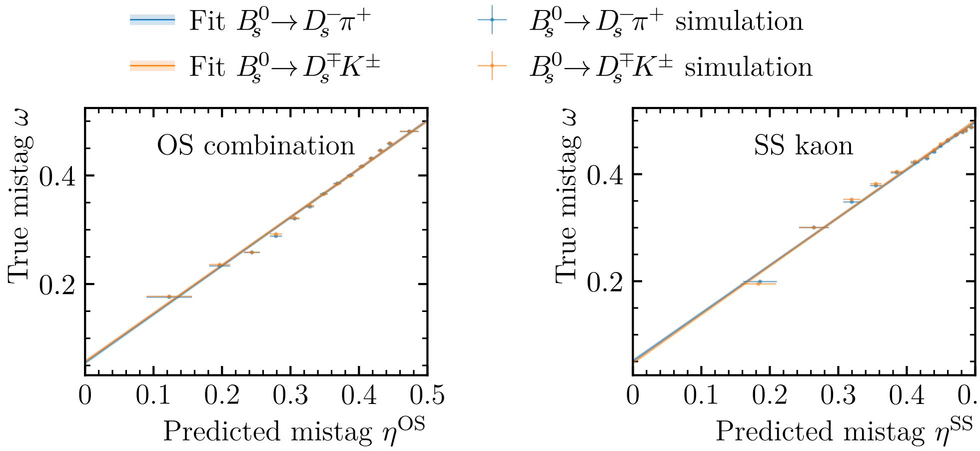


Figure 11.3: Comparison of the mistag calibration for the OS combination (left) and the SS kaon tagger (right). The calibrations are based on the known initial flavour in the simulated $B_s^0 \to D_s^- \pi^+$ (blue) and $B_s^0 \to D_s^\mp K^\pm$ (orange) samples.

The small differences found in the calibration of simulated samples in Section 11.4.1 are also assumed as systematic. In Table 11.2 the full list of systematics for the calibration of the SS kaon in the 2018 sample is shown as a representative for the calibrations of both taggers and each data-taking period. The full systematic uncertainty on each parameter is obtained by summing the individual effects in quadrature. A summary of the calibration parameters extracted in the $B_s^0 \to D_s^- \pi^+$ decay-time fit together with their fitted and systematic uncertainties is given in Table 11.3.

Table 11.2: Difference of SS 2018 calibration parameters of different fit strategies to the nominal strategy.

Systematic	p_0	p_1	Δp_0	Δp_1
Polarity split	0.0004 ± 0.0010	-0.008 ± 0.011	0.0000 ± 0.0009	-0.003 ± 0.010
D_s^--mode split	0.0000 ± 0.0026	0.036 ± 0.030	0.0012 ± 0.0023	-0.012 ± 0.025
Momentum split	0.0043 ± 0.0018	0.016 ± 0.012	0.0008 ± 0.0015	0.001 ± 0.010
BDT split	0.0006 ± 0.0010	-0.003 ± 0.011	0.0002 ± 0.0009	-0.003 ± 0.011
Resolution	-0.0018 ± 0.0001	0.015 ± 0.001	0.0031 ± 0.0001	-0.007 ± 0.001
Portability	0.0077 ± 0.0004	-0.010 ± 0.004	-0.0004 ± 0.0008	0.012 ± 0.009
Sum in quadrature	0.0090 ± 0.0009	0.044 ± 0.025	0.0034 ± 0.0009	0.019 ± 0.017
Stat. uncertainty	0.0033	0.038	0.003	0.033
Tot. uncertainty	0.0097	0.063	0.0046	0.042

Table 11.3: Flavour tagging calibration parameters extracted from the decay-time fit to $B_s^0 \to D_s^- \pi^+$ data. The first uncertainty is statistical, the second is systematic.

Parameter	2015 & 16	2017	2018
OS combination			
p_0^{OS}	$0.386 \pm 0.004 \pm 0.008$	$0.376 \pm 0.005 \pm 0.006$	$0.374 \pm 0.004 \pm 0.013$
p_1^{OS}	$0.985 \pm 0.044 \pm 0.110$	$0.853 \pm 0.043 \pm 0.069$	$0.870 \pm 0.040 \pm 0.064$
Δp_0^{OS}	$0.009 \pm 0.004 \pm 0.008$	$0.004 \pm 0.004 \pm 0.008$	$0.010 \pm 0.004 \pm 0.002$
Δp_1^{OS}	$0.016 \pm 0.041 \pm 0.035$	$0.070 \pm 0.039 \pm 0.102$	$0.032 \pm 0.035 \pm 0.053$
SS kaon			
p_0^{SS}	$0.435 \pm 0.003 \pm 0.012$	$0.437 \pm 0.004 \pm 0.009$	$0.437 \pm 0.003 \pm 0.009$
p_1^{SS}	$0.748 \pm 0.038 \pm 0.068$	$0.732 \pm 0.041 \pm 0.020$	$0.793 \pm 0.038 \pm 0.044$
Δp_0^{SS}	$-0.016 \pm 0.003 \pm 0.003$	$-0.024 \pm 0.003 \pm 0.002$	$-0.015 \pm 0.003 \pm 0.003$
Δp_1^{SS}	$-0.002 \pm 0.035 \pm 0.043$	$0.060 \pm 0.036 \pm 0.038$	$0.055 \pm 0.033 \pm 0.019$

11.5 Decay-time fit of the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal

Next to the theoretical decay rates, the different experimental effects described in the previous have to be modelled in the decay-time fit. Similar to Equation 7.11, the fitted PDF is defined by the product of the acceptance model $\varepsilon_t(t)$ and the convolution of the resolution model $\mathcal{R}(t \,|\, \Delta t,\, \delta t)$ and an effective PDF. This effective PDF

$$\begin{aligned}\frac{\mathcal{P}_{\text{eff}}\big(t \,|\, q_f,\, \vec{d},\, \vec{\eta}\big)}{\mathrm{e}^{-\mathit{\Gamma}_s t}} \sim & \left(1 + q_f \cdot a_{\text{det}}\right) \left(\Phi^-\big(\vec{d},\, \vec{\eta}\big) - \Phi^+\big(\vec{d},\, \vec{\eta}\big) \cdot a_{\text{prod}}\right) \cosh\left(\frac{\Delta\mathit{\Gamma}_s t}{2}\right) \\ & + q_f \left(1 - q_f \cdot a_{\text{det}}\right) \left(\Phi^+\big(\vec{d},\, \vec{\eta}\big) - \Phi^-\big(\vec{d},\, \vec{\eta}\big) \cdot a_{\text{prod}}\right) \ \cos\left(\Delta m_s t\right) \cdot C \\ & + \quad \left(1 + q_f \cdot a_{\text{det}}\right) \left(\Phi^-\big(\vec{d},\, \vec{\eta}\big) - \Phi^+\big(\vec{d},\, \vec{\eta}\big) \cdot a_{\text{prod}}\right) \sinh\left(\frac{\Delta\mathit{\Gamma}_s t}{2}\right) \cdot D(q_f) \\ & + q_f \left(1 - q_f \cdot a_{\text{det}}\right) \left(\Phi^+\big(\vec{d},\, \vec{\eta}\big) - \Phi^-\big(\vec{d},\, \vec{\eta}\big) \cdot a_{\text{prod}}\right) \ \sin\left(\Delta m_s t\right) \cdot S(q_f)\end{aligned} \tag{11.3}$$

incorporates the detection and production asymmetries as well as the flavour tagging to the theoretical decay rates from Equation 3.17 to 3.20. The calibration and combination of the two flavour taggers is implemented by the coefficients $\Phi^{\pm}(d_1,\, \eta_1)$ defined in Equation 7.14 to 7.16. The contribution of the *CP* parameters

$$D(q_f) = \begin{cases} D_f, & \text{for } q_f = 1, \\ D_{\bar{f}}, & \text{for } q_f = -1, \end{cases} \qquad S(q_f) = \begin{cases} S_f, & \text{for } q_f = 1, \\ S_{\bar{f}}, & \text{for } q_f = -1 \end{cases} \tag{11.4}$$

is modulated in the PDF by the final-state flavour q_f defined by the kaon charge. In the absence of *CP* violation in the B_s^0 mixing and of direct *CP* violation in the $B_s^0 \to D_s^{\mp} K^{\pm}$ decay, the parameter $C = C_f = C_{\bar{f}}$ is invariant to the final state as discussed in Section 3.3.

Similar to $B_s^0 \to D_s^- \pi^+$ fit in Section 7.4, the decay-time bias and the calibration of the uncertainty estimate are fixed to the values found in Section 11.1 and 11.2, respectively. In contrast to $B_s^0 \to D_s^- \pi^+$, the detection and production asymmetries, the acceptance parametrisation and the flavour tagging calibration are fixed in the fit. For the production asymmetry the value estimated in Section 7.4 is used. The detection asymmetry is taken from [124], where for the 2017 and 2018 samples the same value as observed in 2016 is assumed. In the fit, the detection asymmetry is fixed to the weighted average of all years. The acceptance is fixed to the parametrisation found in the updated $B_s^0 \to D_s^- \pi^+$ fit with a simulation-based correction, which is introduced in Section 11.3. The decay width $\mathit{\Gamma}_s$ and the corresponding difference $\Delta\mathit{\Gamma}_s$ are fixed to values taken from latest LHCb combination [116]. The oscillation frequency is set to the value measured in Part II. As discussed in Section 11.4, the flavour tagging calibrations are Gaussian constrained in the fit based on calibration parameters and their covariances fitted in $B_s^0 \to D_s^- \pi^+$. An overview of all fixed parameters in the fit is given in Appendix B.4.

The fit is performed simultaneously for all subsamples, where the parametrisation of the decay-time resolution and bias, the acceptance and the tagging calibration are split between the three data-taking periods. The fitted decay-time distribution of the combined sample is presented in Figure 11.4. The oscillating character of the decay-time-dependent mixing asymmetry introduced in Equation 3.39 is clearly visible for both final states. Besides, each of the four decay rates is individually resolved.

The fitted *CP* parameters and the corresponding uncertainties are given in Table 11.4a. The fitted covariance of these five parameters is essential for the determination of the CKM angle γ in Chapter 13. The information can be derived from the given uncertainties and the correlation matrix provided in Table 11.4b. The highest correlation of 50 % between the parameters D_f and $D_{\bar{f}}$ is found to be the largest. The correlations of other parameter combinations are 13 % at most. Following Equation 3.38, *CP* violation in the interference of mixing and decay is observed at a significance of 8.6 by evaluating the inequality $S_f + S_{\bar{f}} \neq 0$ taking the corresponding correlations into account.

Table 11.4: *CP* parameters and the corresponding correlations extracted in the decay-time fit to $B_s^0 \to D_s^\mp K^\pm$ data. The fitted central values are still blinded, as the analysis is currently under review.

Parameter	Fitted value
C	0.791 ± 0.061
D_f	-0.051 ± 0.134
$D_{\bar{f}}$	-0.303 ± 0.125
S_f	-0.571 ± 0.084
$S_{\bar{f}}$	-0.503 ± 0.084

(a) Fitted *CP* parameters.

	C	D_f	$D_{\bar{f}}$	S_f	$S_{\bar{f}}$
C	1	0.134	0.130	0.039	0.022
D_f	0.134	1	0.501	−0.108	−0.036
$D_{\bar{f}}$	0.130	0.501	1	−0.056	−0.067
S_f	0.039	−0.108	−0.056	1	0.006
$S_{\bar{f}}$	0.022	−0.036	−0.067	0.006	1

(b) Correlation matrix.

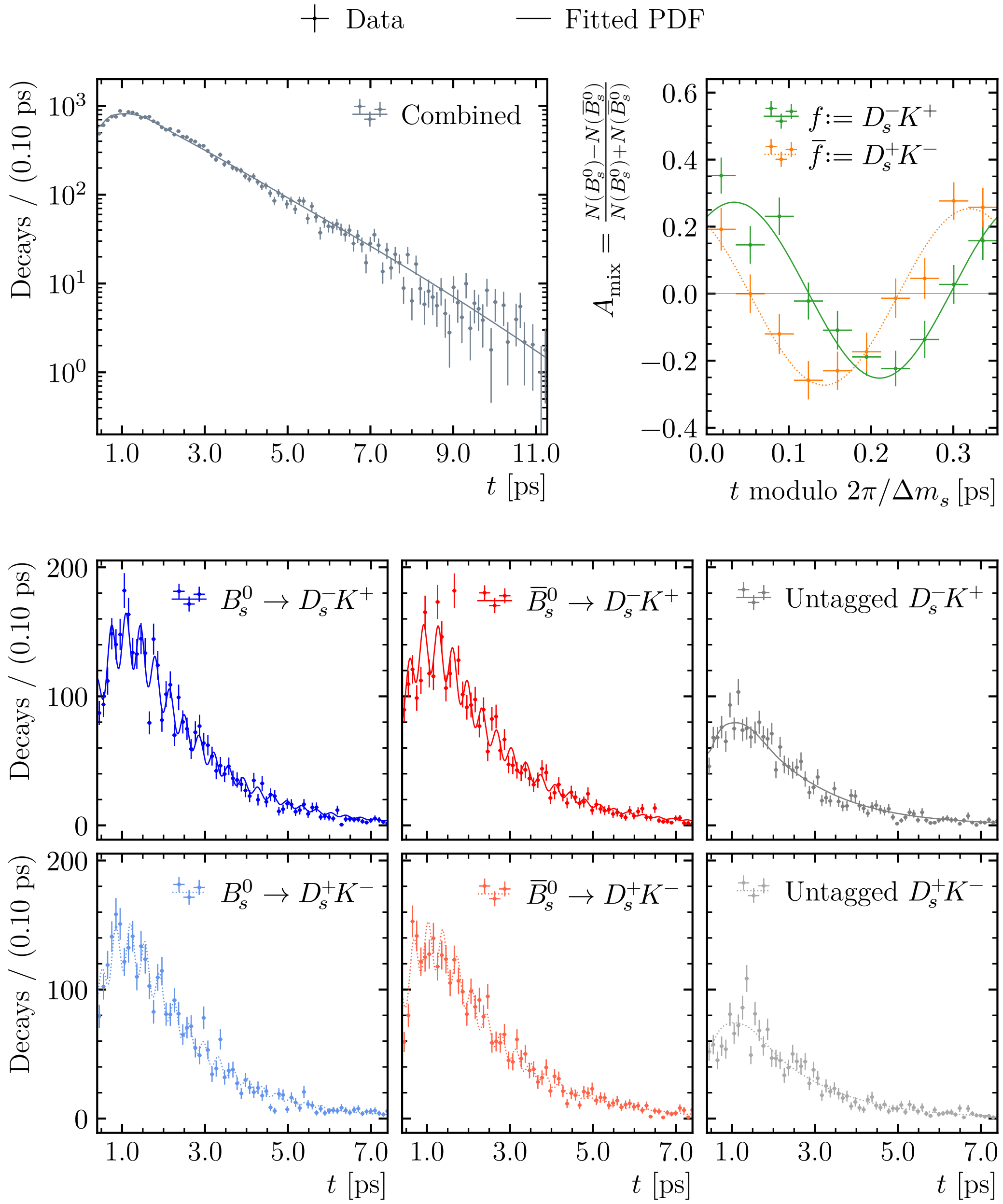


Figure 11.4: Decay-time distribution of the $B_s^0 \to D_s^\mp K^\pm$ signal sample and the corresponding fitted PDF. The upper left plot shows the combined decay rate on a logarithmic scale, while in the bottom six plots the time-dependent decay rates, defined in Equation 3.17 to 3.20, are split by the tagged initial flavour (columns) and by the final state (rows). In the upper right plot, the time-dependent asymmetries A_{mix} from Equation 3.39 are projected to a single oscillation period for the two final states.

12 Systematic studies and cross-checks on the *CP* parameters

The robustness of the fit strategy and the stability of the fitted parameters against model changes need to be proven. Hence, a variety of cross-checks, ranging from the variation of fixed parameters over alternative fit models up to full closure tests, is performed and systematic effects on the fitted *CP* parameters are studied.

In Section 12.6, the systematic effects introduced by the mass-fit model and the constrained flavour-tagging calibration are evaluated in a data-driven approach by rerunning the decay-time fit with an alternative setup. While in Chapter 8 such a data-driven approach is often sufficient to evaluate systematic effects on the oscillation frequency Δm_s, the most model-dependent systematic effects here are studied using pseudo-experiments, as explained in Section 12.5. This way, the correlation among the shifts of the five *CP* parameters introduced by systematic effects is accessible, which provides a more precise evaluation of the effects on the angle γ in Chapter 13.

The bootstrapping approach developed in Section 8.3 is also used for the analysis validation in Section 12.4. Next to the bootstrapping, the large simulated signal samples are used in closure tests fitting the full signal sample in Section 12.3. Similar to the $B_s^0 \to D_s^- \pi^+$ analysis, another cross-check is performed by analysing different subsamples of the data set individually in Section 12.2. Further, the compatibility of the fitted *CP* parameters with respect to previous results is tested in Section 12.1.

An overview of the systematic effects is given in Table 12.1, and the corresponding covariance is shown in Table 12.2. The detailed evaluation of these effects is discussed in the following.

Table 12.1: Overview of the different systematic effects found to affect the *CP* parameters normalised to the statistical uncertainty.

Parameter	C_f	D_f	$D_{\bar{f}}$	S_f	$S_{\bar{f}}$
Neglecting correlations	0.137	0.081	0.054	0.135	0.043
Δm_s	0.007	0.004	0.004	0.108	0.103
Production asymmetry	—	—	—	—	—
Detection asymmetry	—	0.079	0.083	0.006	0.007
Decay-time bias	0.062	0.027	0.046	0.188	0.167
Decay-time resolution model	0.195	0.002	0.003	0.058	0.167
Decay-time acceptance, $\varGamma_s$, $\Delta\varGamma_s$	0.006	0.225	0.231	0.003	0.003
Decay-time acceptance ratios	0.001	0.018	0.018	—	—
Flavour tagging portability	0.256	0.026	0.028	0.012	0.070
Multidimensional fit	0.045	0.095	0.121	0.088	0.112
Total	0.358	0.273	0.285	0.278	0.294

Table 12.2: Correlation matrix obtained from the combined covariance of all systematic uncertainties.

	C	D_f	$D_{\bar{f}}$	S_f	$S_{\bar{f}}$
C	1	0.004	0.010	−0.092	−0.260
D_f	0.004	1	0.778	0.012	−0.002
$D_{\bar{f}}$	0.010	0.778	1	0.000	−0.005
S_f	−0.092	0.012	0.000	1	0.099
$S_{\bar{f}}$	−0.260	−0.002	−0.005	0.099	1

12.1 Compatibility to previous results

A significant incompatibility of the extracted fit result with respect to previous measurements of the *CP* parameters would hint at a bias in either of the measurements. Hence, a study of the compatibility with previous results is a reasonable cross-check. The measure chosen to evaluate the compatibility between two sets of fit results is the Mahalanobis distance [125]

$$d_{\mathrm{M}} = \sqrt{(\vec{x}-\vec{y})^{\mathrm{T}} \cdot \boldsymbol{\Sigma}_{\vec{x}}^{-1} \cdot (\vec{x}-\vec{y})} = \sqrt{\Delta\vec{x}^{\mathrm{T}} \cdot \boldsymbol{\Sigma}_{\Delta\vec{x}}^{-1} \cdot \Delta\vec{x}}\,. \tag{12.1}$$

The Mahalanobis distance is designed to study the distance between a point $\vec{y}$ and a distribution defined by a vector $\vec{x}$ and a covariance $\boldsymbol{\Sigma}_{\vec{x}}$. However, the distance between the two distributions is here of interest. This misconception is overcome by studying the distance between the point-like origin of the parameter space $\vec{0}$ and the distributed difference of the two results $\vec{x}_1$ and $\vec{x}_2$, which is parameterised by the difference vector

$$\Delta\vec{x}(\vec{x}_1,\, \vec{x}_2) \;=\; \begin{pmatrix}1 & -1\end{pmatrix} \cdot \begin{pmatrix}\vec{x}_1 \\ \vec{x}_2\end{pmatrix} \tag{12.2}$$

and the corresponding covariance

$$\boldsymbol{\Sigma}_{\Delta\vec{x}} \;=\; \boldsymbol{J}^{\mathrm{T}}_{\Delta\vec{x}} \cdot \boldsymbol{\Sigma}_{\vec{x}_1,\vec{x}_2} \cdot \boldsymbol{J}_{\Delta\vec{x}}\,. \tag{12.3}$$

Here, the Jacobian matrix of the mapping $\Delta\vec{x}(\vec{x}_1,\, \vec{x}_2)$ is denoted as $\boldsymbol{J}_{\Delta\vec{x}}$. In case of two independent fit results, the combined covariance $\boldsymbol{\Sigma}_{\vec{x}_1,\vec{x}_2} = \mathrm{diag}(\boldsymbol{\Sigma}_{\vec{x}_1},\, \boldsymbol{\Sigma}_{\vec{x}_2})$ is a diagonal block matrix of the two fitted covariances. The Mahalanobis distance is χ distributed. Hence, the cumulative distribution function (CDF) of the χ distribution can be used to transform the distance into a p-value

$$p \;=\; 1 - \chi_{\mathrm{CDF}}(d_{\mathrm{M}},\, n_{\mathrm{dof}}) \tag{12.4}$$

measuring the probability to find a result with the given distance or larger. In the case of the fitted *CP* parameters, the degree of freedom is $n_{\mathrm{dof}} = 5$.

Two previous measurements of γ using $B_s^0 \to D_s^{\mp}K^{\pm}$ decays are considered, the analysis of $B_s^0 \to D_s^{\mp}K^{\pm}$ decays using the LHCb Run1 data set [14] and the analysis of $B_s^0 \to D_s^{\mp}K^{\pm}\pi^{+}\pi^{-}$ decays using the full LHCb data set [57]. The corresponding fitted *CP* parameters and the difference between the two sets are given in Table 12.3. Although the absolute deviation of the two determined values $\gamma_{\mathrm{DsK}}^{3\,\mathrm{fb}^{-1}} = \left(128^{+17}_{-22}\right)^{\circ}$ and $\gamma_{\mathrm{DsKpipi}}^{9\,\mathrm{fb}^{-1}} = (44 \pm 12)^{\circ}$ is large, the distance in the *CP* parameter space $d_{\mathrm{M}} = 2.958$, $p = 12.0\,\%$ not too significant.

The distance of the Run1 result [14] with respect to the nominal fit described in Section 11.5 is found to be $d_{\mathrm{M}} = 2.222$ corresponding to $p = 42.4\,\%$. This compatibility estimate is considered to be conservative as only the statistical covariances are considered and systematic effects are neglected. Besides, it must be considered, that the fixed values of the detection asymmetry a_{det}, production asymmetry a_{prod} and the oscillation frequency Δm_s are fixed to significantly different values in the Run1 fit. These parameters are known to significantly impact the measurement, as shown by the systematic studies summarised in Table 12.1 and by the previous measurement. Hence, the decay-time fit is rerun using the fixed input parameters corresponding to the Run1 measurement. This alternative fit result and its deviation from the nominal result is reported in Table 12.4. The distance $d_{\mathrm{M}} = 1.887$ to the Run1 result is decreased as expected and the compatibility $p = 61.4\,\%$ is improved, still neglecting further systematic effects. The compatibility check with respect to the $B_s^0 \to D_s^{\mp}K^{\pm}\pi^{+}\pi^{-}$ analysis [57] yields a distance $d_{\mathrm{M}} = 2.144$ corresponding to

Table 12.3: *CP* parameters extracted in the analysis of $B_s^0 \to D_s^\mp K^\pm$ decays using the LHCb Run1 data set [14] and in the analysis of $B_s^0 \to D_s^\mp K^\pm \pi^+ \pi^-$ decays using the full LHCb data set [57], as well as the corresponding difference.

Parameter	$B_s^0 \to D_s^\mp K^\pm$ Run1	$B_s^0 \to D_s^\mp K^\pm \pi^+ \pi^-$	Difference
C	0.73 ± 0.14	0.63 ± 0.10	0.10 ± 0.17
D_f	0.39 ± 0.28	-0.33 ± 0.23	0.7 ± 0.4
$D_{\bar{f}}$	0.31 ± 0.28	-0.70 ± 0.22	1.00 ± 0.35
S_f	-0.52 ± 0.20	-0.42 ± 0.14	-0.10 ± 0.24
$S_{\bar{f}}$	-0.49 ± 0.20	-0.46 ± 0.13	-0.03 ± 0.24

Table 12.4: *CP* parameters extracted using fixed inputs aligned to the Run1 analysis of $B_s^0 \to D_s^\mp K^\pm$ decays [14] and the difference to the nominal result.

Parameter	Fitted value	Difference to nominal
C	0.780 ± 0.061	-0.0104 ± 0.0009
D_f	0.120 ± 0.136	-0.171 ± 0.009
$D_{\bar{f}}$	-0.146 ± 0.128	-0.157 ± 0.009
S_f	-0.572 ± 0.085	0.0014 ± 0.0017
$S_{\bar{f}}$	-0.530 ± 0.084	0.0276 ± 0.0009

$p = 46.7\,\%$. Here, systematic effects are again neglected and the fixed inputs are not corrected. The compatibility with the previous two results is concluded to be good.

12.2 Analyses of subsamples

An important cross-check is the compatibility between the nominal result and *CP* parameters extracted on specific subsamples of the data. Additionally, the combination of subsample results yields an important closure test as it is expected to be equivalent to the nominal result. The subsamples analysed individually are the two magnet polarities, the three data-taking periods, the five D_s^--decay modes, two bins of the B_s^0 momentum and a split into three exclusive flavour tagging categories, events tagged exclusively by one specific tagger or tagged by both.

The compatibility of the subsample results is evaluated using the Mahalanobis distance and the corresponding p-value as described in Section 12.1. However, the two results are not independent here. Hence, the combined covariance

$$\boldsymbol{\Sigma}_{\vec{x},\vec{x}_i} = \begin{pmatrix} \boldsymbol{\Sigma}_{\vec{x}} & \boldsymbol{Z}_{\vec{x},\vec{x}_i} \\ \boldsymbol{Z}^T_{\vec{x},\vec{x}_i} & \boldsymbol{\Sigma}_{\vec{x}_i} \end{pmatrix} \tag{12.5}$$

for the nominal result $\vec{x}$ and a subsample result $\vec{x}_i$ is not diagonal anymore and knowledge of the cross-correlation between the two measurements described by the matrix $\boldsymbol{Z}_{\vec{x},\vec{x}_i}$ is required. This information is not trivially accessible. Here, an estimate for this cross-correlation is obtained by making use of the residual subsamples. It is assumed, that the nominal result resembles the average of all subsamples

$$\vec{x} := \operatorname{avg}\left(\vec{x}_i,\, i \in \{1\,..\,n\}\right) = \sum_i \boldsymbol{\alpha}_i \cdot \vec{x}_i\,, \tag{12.6}$$

with a weight tensor

$$\boldsymbol{\alpha}_i = \left[\boldsymbol{\Sigma}_{\vec{x}_i} \cdot \sum_j \boldsymbol{\Sigma}^{-1}_{\vec{x}_j}\right]^{-1} \tag{12.7}$$

incorporating the uncertainties of the individual measurements. Hence, the difference between a specific subsample and the nominal result

$$\vec{x} - \boldsymbol{\alpha}_i \cdot \vec{x}_i = \operatorname{avg}\left(\vec{x}_j,\, j \neq i\right) \tag{12.8}$$

is defined by the average of the remaining subsamples. Since the individual subsample results are independent, the covariance of the average can easily be constructed. Using proper uncertainty propagation, similar to Equation 12.3,

$$\begin{pmatrix} 1 \\ -\boldsymbol{\alpha}_i \end{pmatrix}^{\mathrm{T}} \cdot \begin{pmatrix} \boldsymbol{\Sigma}_{\vec{x}} & \boldsymbol{Z}_{\vec{x},\vec{x}_i} \\ \boldsymbol{Z}^T_{\vec{x},\vec{x}_i} & \boldsymbol{\Sigma}_{\vec{x}_i} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -\boldsymbol{\alpha}_i \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}_j, \\ j \in \{1..n\}, \\ j \neq i \end{pmatrix}^{\mathrm{T}} \cdot \operatorname{diag}\begin{pmatrix} \boldsymbol{\Sigma}_{\vec{x}_j}, \\ j \in \{1..n\}, \\ j \neq i \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\alpha}_j, \\ j \in \{1..n\}, \\ j \neq i \end{pmatrix} \tag{12.9}$$

the estimate of the cross-correlation

$$\boldsymbol{Z}_{\vec{x},\vec{x}_i} = \frac{1}{2}\boldsymbol{\alpha}_i^{-1}\left[\boldsymbol{\Sigma}_{\vec{x}} + \boldsymbol{\alpha}_i^2 \cdot \boldsymbol{\Sigma}_{\vec{x}_i} - \sum_{j \neq i}\left(\boldsymbol{\alpha}_j^2 \cdot \boldsymbol{\Sigma}_{\vec{x}_j}\right)\right] \tag{12.10}$$

can be constructed from the fitted covariances. The application of the derived covariance to the proper uncertainty propagation from Equation 12.3 yields the simple relation

$$\boldsymbol{\Sigma}_{\vec{x}-\vec{x}_i} = \boldsymbol{\Sigma}_{\vec{x}} - \boldsymbol{\Sigma}_{\vec{x}_i}\,. \tag{12.11}$$

Following this approach, the compatibility of the different subsamples with respect to the nominal result is calculated and reported in Table 12.5 in terms of the Mahalanobis distance and the corresponding p-value. The details on the corresponding deviations of the individual *CP* parameters found in the different subsamples are presented in Appendix B.5.

Table 12.5: Compatibility of the *CP* parameters extracted on subsample splits with respect to the nominal result.

Split	Subsample	Mahalanobis distance	p-value
Magnet polarity	Up	2.52	0.272
	Down	2.20	0.434
Data-taking period	2015 & 16	2.52	0.275
	2017	2.10	0.495
	2018	1.66	0.736
D_s^--decay mode	$(K^-K^+\pi^-)_{\mathrm{NR}}$	2.81	0.162
	$K^{*0}K^-$	2.68	0.209
	$\phi\pi^-$	3.65	0.021
	$K^-\pi^+\pi^-$	1.59	0.770
	$\pi^-\pi^+\pi^-$	2.05	0.520
B_s^0 momentum	$p_{B_s^0} > 120$ GeV/c	0.86	0.980
	$p_{B_s^0} < 120$ GeV/c	1.08	0.947
FT	SS	2.35	0.358
	OS	1.23	0.913
	Both	1.71	0.713

Overall, the observed compatibility is good. Further, it must be considered that the decay-time acceptance and flavour-tagging calibration are reevaluated for each subsample fit following the procedures from Section 11.3 and 11.4. Hence, each result is subject to the related systematic uncertainties, which are neglected in this study, making this compatibility estimate very conservative.

12.3 Simulation-based closure tests

The closure test of fitting simulated samples provides an important fit validation. The use of pure signal simulations allows the sWeight approach to be omitted for background subtraction and the related MDFit. However, the full selection is applied to account for effects from the PID requirements and from kinematic cuts. Like in data, neither the acceptance parametrisation nor the flavour tagging calibration can be left free in the fit, since the *CP* parameters are required to float in the closure test. Hence, both need to be taken from a fit to simulated $B_s^0 \to D_s^-\pi^+$ decays as in data. The decay-time uncertainty calibration and the decay-time bias are extracted based on the true decay time known in the simulation with respect to the reconstructed decay time.

At first, a fit of simulated $B_s^0 \to D_s^- \pi^+$ decays is performed to extract the tagging calibration and the decay-time acceptance. In the $B_s^0 \to D_s^\mp K^\pm$ fit, the flavour tagging calibration is then constrained by the fitted covariance. The acceptance is fixed to values found in the $B_s^0 \to D_s^- \pi^+$ fit and corrected by the ratio of the acceptance found in both modes. However, since this study was performed before the acceptance parametrisation was extended from a set of eight to a set of eleven splines, in this study the original parametrisation from Section 7.4 is used. As this update is found to have a negligible effect on the *CP* parameters in data, the closure test with the original parametrisation is concluded to be a reasonable crosscheck and is not updated. To extract the correction, independent simulated samples are required. There are three different simulated samples available for each magnet polarity, data-taking period and D_s^--decay mode:

- Samples of $B_s^0 \to D_s^\mp K^\pm$ decays simulated in the presence of *CP* violation. These samples are considered the "data-like" $B_s^0 \to D_s^\mp K^\pm$ samples as they are used in the decay-time fit to extract the five *CP* parameters.
- Simulated samples of $B_s^0 \to D_s^\mp K^\pm$ decays neglecting all effects of *CP* violation. Still, these so-called "simulation-like" samples are used for the extraction of the acceptance.
- Simulation of $B_s^0 \to D_s^- \pi^+$ decays. These samples are needed both for the decay-time fit and the acceptance evaluation. Hence, they are randomly split into halves, a "data-like" half used for the decay-time fit and a "simulation-like" half for the acceptance.

The simulation-like samples are reweighted to match the same proportions in terms of polarity, data-taking period and D_s^- mode as the corresponding data-like samples. In the fits of the $B_s^0 \to D_s^- \pi^+$ samples, the $D_s^- \to K^- \pi^+ \pi^-$ mode is discarded to recreate the same situation as in the data. The simulation-like samples are used to extract acceptance ratios reported in Appendix B.6.

The extracted *CP* parameters in comparison to their generated values are reported in Table 12.6. The Mahalanobis distance of the fitted result with respect to the generated parameters is $d_\mathrm{M} = 3.10$ with a corresponding p-value of $p = 8.7\,\%$. Although the p-value is rather low, it is still well above a 2σ threshold corresponding to $p < 4.6\,\%$. Further, the simulated samples provide much more statistics than the data samples. As a cross-check the Mahalanobis distance $d'_\mathrm{M} = 0.45$ corresponding to $p' = 99.9\,\%$ is recalculated with a covariance scaled to reproduce the uncertainties seen in data. In combination with the later in Section 12.5 discussed pseudo experiments, which show perfectly unbiased fit results, the small deviations observed are considered to originate from the imperfect knowledge of the inputs in simulation. No systematic uncertainty is assigned as these effects are small and already accounted for by the specific systematics. The closure test is also performed using the different data-taking periods and D_s^- modes individually. The individual results align well with the nominal closure test and no unexpected behaviour is found.

Effect of the PID requirements

The PID requirement on the accompanying hadron is known to affect the signal kinematics and the decay-time acceptance as seen in Section 11.3. Hence, the effect is also studied in the simulation-based closure test. The fits of the simulated $B_s^0 \to D_s^\mp K^\pm$ samples are repeated, while the PID requirement $\mathrm{PID}_K > 5$ of the accompanying kaon is tightened to $\mathrm{PID}_K > 10$ or loosened to $\mathrm{PID}_K > 0$. The latter case yields a symmetric splitting with respect to the $\mathrm{PID}_K < 0$ requirement on the accompanying pion, which is not modified.

As expected, the deviation from 1.0 of acceptance ratios decreases with the looser requirement and increases with the tighter. Nevertheless, the compatibility is not significantly affected for the CP parameters extracted from the simulated samples with the modified PID requirement reported in Table 12.7. In the presence of the tight requirement, the Mahalanobis distance is $d_\mathrm{M} = 3.05$ corresponding to $p = 9.8\,\%$ and $d_\mathrm{M} = 3.22$ corresponding to $p = 6.6\,\%$ under the looser requirement.

Table 12.6: The CP parameters extracted in the closure test using full LHCb simulation in comparison to the generated values. The difference found is also presented normalised by the statistical uncertainty of the corresponding parameter found in data $\sigma_\mathrm{stat}^\mathrm{data}$.

Parameter	Generation value	Fit result	Difference	Difference / $\sigma_\mathrm{stat}^\mathrm{data}$
C	0.759	0.776 ± 0.009	-0.017 ± 0.009	-0.28 ± 0.15
D_f	−0.314	-0.310 ± 0.030	-0.004 ± 0.030	-0.03 ± 0.24
$D_{\bar{f}}$	−0.101	-0.088 ± 0.029	-0.012 ± 0.029	-0.10 ± 0.25
S_f	−0.570	-0.580 ± 0.013	0.010 ± 0.013	0.12 ± 0.16
$S_{\bar{f}}$	−0.643	-0.669 ± 0.013	0.026 ± 0.013	0.31 ± 0.15

Distorted simulation-like samples

In the nominal data fit, the acceptance is extracted from $B_s^0 \to D_s^- \pi^+$ data, while the correction is based on simulation. Hence, this part of the analysis is prone to be subject to effects from simulation-data differences. In Section 11.3 a reweighting is

Table 12.7: The *CP* parameters extracted in the closure test using alternative PID requirements in comparison to the generated values. The difference found is also presented normalised by the statistical uncertainty of the corresponding parameter found in data $\sigma_{\text{stat}}^{\text{data}}$.

Parameter	Fit result	Difference to generation	Difference / $\sigma_{\text{stat}}^{\text{data}}$
	Tight $\text{PID}_K > 10$ requirement		
C	0.772 ± 0.010	-0.012 ± 0.010	-0.21 ± 0.16
D_f	-0.322 ± 0.032	0.008 ± 0.032	0.06 ± 0.25
$D_{\bar{f}}$	-0.076 ± 0.031	-0.025 ± 0.031	-0.21 ± 0.26
S_f	-0.583 ± 0.014	0.013 ± 0.014	0.16 ± 0.17
$S_{\bar{f}}$	-0.673 ± 0.014	0.030 ± 0.014	0.36 ± 0.16
	Tight $\text{PID}_K > 10$ requirement		
C	0.767 ± 0.008	-0.008 ± 0.008	-0.13 ± 0.13
D_f	-0.276 ± 0.027	-0.038 ± 0.027	-0.30 ± 0.21
$D_{\bar{f}}$	-0.060 ± 0.026	-0.040 ± 0.026	-0.34 ± 0.22
S_f	-0.577 ± 0.011	0.007 ± 0.011	0.09 ± 0.13
$S_{\bar{f}}$	-0.662 ± 0.011	0.019 ± 0.011	0.23 ± 0.13

used to correct these differences before the acceptance ratio is extracted. Nevertheless, the acceptance extracted from simulated $B_s^0 \to D_s^- \pi^+$ decays in Section 11.3 does not perfectly match the acceptance found in data in Section 7.4, hinting for residual differences.

In the closure test, these differences are not present, as both the simulation-like and data-like samples are simulated. Here, such an effect is artificially introduced by distorting the simulation-like samples. The BDT selection aiming to reduce combinatorial background is known to show large effects on the decay-time acceptance, especially at low decay times. The simulation-like samples are randomly split into halves. One half is required to pass a tighter BDT selection $\alpha_{\text{BDT}} > 0.7$, the other a looser $\alpha_{\text{BDT}} > 0.4$. This modification significantly distorts the acceptance of the simulation-like samples and avoids a potential cancellation of the $B_s^0 \to D_s^- \pi^+$ acceptance, when the acceptance of data-like samples is multiplied by the ratio from the simulation-like samples.

The *CP* parameters extracted in the closure test using ratios based on these distorted simulation-like samples as corrections are presented in Table 12.8. The change in the observed ratios is minor, although the individual acceptances are changed significantly. This shows the robustness of the acceptance correction against simulation-data differences. Nevertheless, the effect on the parameters D_f and $D_{\bar{f}}$ is sizeable and the compatibility $d_{\mathrm{M}} = 3.64$ corresponding to $p = 2.1\,\%$ with the generated values is reduced. This shows the sensitivity of these two parameters on the acceptance. The change in C, S_f and $S_{\bar{f}}$ however is of negligible size. No systematic is assigned here, as a dedicated study evaluates the impact of the acceptance in Section 12.5.

Table 12.8: The *CP* parameters extracted in the closure test based on acceptance ratios extracted from artificially distorted simulation-like samples.

Parameter	Fit result	Difference to generation	Difference / $\sigma_{\mathrm{stat}}^{\mathrm{data}}$
C	0.774 ± 0.009	-0.014 ± 0.009	-0.24 ± 0.15
D_f	-0.273 ± 0.030	-0.041 ± 0.030	-0.32 ± 0.24
$D_{\bar{f}}$	-0.052 ± 0.029	-0.049 ± 0.029	-0.41 ± 0.24
S_f	-0.578 ± 0.013	0.008 ± 0.013	0.10 ± 0.15
$S_{\bar{f}}$	-0.667 ± 0.013	0.024 ± 0.013	0.29 ± 0.15

12.4 Bootstrapping studies

The simulated samples are also used to perform bootstrapping studies similar to Section 8.3. While the primary use case is to study the effect of neglecting correlations among the fit observable in the sWeighting approach, the simulation-based bootstrapping studies can also be considered as an alternative closure test.

As in the $B_s^0 \to D_s^- \pi^+$ analysis, events are randomly drawn from simulated signal and background samples and from data sideband, as a proxy for combinatorial background, to resemble the contributions found in the selected data sample in Section 10.3. The bootstrapped samples are then duplicated with one sample keeping the initial correlations among the fit observables and the other having them artificially removed. The two versions are analysed using the full fit strategy consisting of the MDFit to extract sWeights and the decay-time fit to fit the *CP* parameters. In Section 8.3 the approach is discussed in more detail.

For the decay-time fit of the bootstrapped samples again external information on the asymmetries, the acceptance and the tagging calibration is needed. This information is taken from the fit to the full data-like samples of simulated $B_s^0 \to D_s^- \pi^+$ decays performed in Section 12.3. The acceptance correction is taken from the corresponding simulation-like samples discussed in Section 12.3. The choice of using the full simulated samples instead of bootstrapped $B_s^0 \to D_s^- \pi^+$ samples is made to reduce unnecessary disturbance from fluctuation in the input parameters. To reduce the computational costs, the decay-time fit strategy is modified by using a fixed flavour tagging calibration, instead of Gaussian constraining it.

The correlations found in the initial samples used for the bootstrapping align well with the findings of Section 8.3 and are not explicitly listed here. The mean deviation of the fitted *CP* parameters from the generated values is reported for both 300 correlated and uncorrelated bootstrapped samples in Table 12.9 together with the mean of the difference between the two samples, which is accounted as a systematic for neglecting the correlations. The corresponding results of the bootstrapped samples neglecting all background contributions are summarised in Table 12.10. In this scenario, effects introduced by the sWeighting are significantly reduced as expected. Further, it perfectly reproduces the results of the full sample closure test reported in Table 12.6 within the small uncertainties.

Table 12.9: The mean of the five *CP* parameters X extracted from the uncorrelated (u) and correlated (c) bootstrapped samples with respect to the generated (g) values and each other.

Parameter	$\langle X^{(c)} - X^{(g)} \rangle$	$\langle X^{(u)} - X^{(g)} \rangle$	$\langle X^{(c)} - X^{(u)} \rangle$
C	0.033 ± 0.002	0.025 ± 0.002	0.008 ± 0.002
D_f	-0.023 ± 0.008	-0.033 ± 0.008	0.009 ± 0.005
$D_{\bar{f}}$	-0.015 ± 0.007	-0.019 ± 0.007	0.004 ± 0.005
S_f	0.002 ± 0.004	-0.009 ± 0.003	0.011 ± 0.002
$S_{\bar{f}}$	-0.028 ± 0.003	-0.025 ± 0.003	-0.003 ± 0.002

Table 12.10: The mean of the five *CP* parameters X extracted from the uncorrelated (u) and correlated (c) bootstrapped samples neglecting the presence of backgrounds with respect to the generated (g) values and each other.

Parameter	$\langle X^{(c)} - X^{(g)} \rangle$	$\langle X^{(u)} - X^{(g)} \rangle$	$\langle X^{(c)} - X^{(u)} \rangle$
C	0.017 ± 0.002	0.016 ± 0.002	0.0009 ± 0.0002
D_f	0.000 ± 0.007	0.000 ± 0.007	-0.0001 ± 0.0007
$D_{\bar{f}}$	0.015 ± 0.007	0.015 ± 0.007	0.0002 ± 0.0006
S_f	-0.001 ± 0.003	-0.001 ± 0.003	-0.0003 ± 0.0003
$S_{\bar{f}}$	-0.025 ± 0.003	-0.024 ± 0.003	-0.0017 ± 0.0003

As in Section 8.3, Cross-checks on the stability of the bootstrapping studies are performed by loosening the BDT requirement, by obtaining the combinatorial background from two alternative sideband regions and by using the acceptance ratios extracted from the distorted simulation-like samples. Again, no significant impact on the extracted systematic is observed in these cross-checks. As discussed in Section 12.3, the usage of the acceptance ratios taken from the distorted samples has a small impact on the fitted *CP* parameters. However, this effect is found to be identical in the correlated and the uncorrelated samples and hence completely cancels in the systematic from neglecting correlations.

12.5 Pseudo-experiment based studies

Pseudo-experiments are used to study a number of systematic effects. The nominal setup is used for a closure test to exclude biases in the baseline fit procedure. Modified setups are used to study systematic effects, like the finite knowledge of parameters fixed in the decay-time fit. The pseudo experiments are generated as discussed in Section 8.2 allowing all sample properties, including the true acceptance and tagging calibration, to be arbitrarily generated.

In the generation of the pseudo experiments, the parametrisation of the mass distributions of all components is also exactly parameterised. This enables the validation of the MDFit model extracted in Section 10.3. The mean and the width of the pull distributions of the MDFit parameters fitted in 1500 pseudo-experiment samples are studied.While most of the parameters are unbiased and show good coverage, a few problematic parameters can be identified. The shared fractions f_1 and $f_{\pi/p}$ modulating the relative yields of backgrounds, the signal mean $\mu_{B_s^0}$, the floating signal width $\sigma^H_{B_s^0}$ and the slope $c_{B_s^0}$ of the combinatorial background each in the $m(D_s^{\mp}K^{\pm})$ mass show small biases up to 20 % of the statistical uncertainty. The combinatorial background slope and fraction $f_{B_s^0}$ also reveals an undercoverage in some subsamples. These findings hint at a potential unidentified mismatch between the parametrisation used in the generation and in the MDFit. However, the absolute effect is rather small as the fitted signal yield in all subsamples combined shows an effect of only 20 candidates. Further and even more important, no effect is observed on the fitted *CP* parameters, as the values agree well with the generation. The deviations found are not statistically significant and orders of magnitude smaller than the statistical uncertainty. Hence, the fit strategy is considered unbiased and no systematic is assigned for a fit bias. Instead, this result is taken as evidence of the robustness of the analysis against the mass model.

In the following, a number of systematic effects is evaluated by performing modified decay-time fits on the pseudo-experiment samples. The systematic uncertainty of

$$\sigma^{\text{syst}} = \sqrt{\langle\mu^{\text{mod}}\rangle^2 + (\sigma^{\text{mod}})^2} \tag{12.12}$$

each modification is derived from the mean deviation $\langle\mu^{\text{mod}}\rangle = \langle x^0 - x^{\text{mod}}\rangle$ of the modified fit results x^{mod} with respect to the nominal fit results x^0 and the corresponding standard deviation σ^{mod}. The systematic covariance is calculated from the correlations among the observed deviations μ^{mod} scaled by the systematic uncertainties

$$\boldsymbol{\Sigma}_{ij}^{\mathrm{syst}} = \frac{\boldsymbol{\Sigma}_{ij}^{\mathrm{mod}}}{\sqrt{\boldsymbol{\Sigma}_{ii}^{\mathrm{mod}} \cdot \boldsymbol{\Sigma}_{jj}^{\mathrm{mod}}}} \cdot \sigma_i^{\mathrm{syst}} \cdot \sigma_j^{\mathrm{syst}} \,, \tag{12.13}$$

where the indices i and j run over the five *CP* parameters. The covariance of the deviations in N pseudo experiments is given by

$$\boldsymbol{\Sigma}_{ij}^{\mathrm{mod}} = \frac{1}{N-1} \sum_{k=1}^{N} \left(\mu_{ik}^{\mathrm{mod}} - \langle\mu_i^{\mathrm{mod}}\rangle\right) \cdot \left(\mu_{jk}^{\mathrm{mod}} - \langle\mu_j^{\mathrm{mod}}\rangle\right) . \tag{12.14}$$

Fixed asymmetries and oscillation frequency

The decay-time fit is performed with the oscillation frequency, as well as the production and detection asymmetries, fixed to their known values. To evaluate systematic effects originating from the finite knowledge of these fixed parameters, the decay-time fit performed on the pseudo-experiment samples is modified by shifting the parameters according to their known uncertainties. For the oscillation frequency the value measured in Part II is used, hence the value is shifted by $\pm 0.006\ \mathrm{ps}^{-1}$. Similarly, the production asymmetry is shifted by the fitted uncertainty $\pm 0.32\,\%$. In the data decay-time fit, the detection asymmetry is fixed to the weighted average of the data-taking years. Here, the value is shifted up and down by the corresponding uncertainty $\pm 0.15\,\%$. The average of the systematic deviations of the two modifications is assigned as systematic uncertainty in Table 12.1. The effect of the oscillation frequency is found to be sizeable for the parameters C, S_f and $S_{\bar{f}}$ related to the oscillation terms, while being negligible for the hyperbolic parameters D_f and $D_{\bar{f}}$. The situation is the opposite for the effects introduced by the detection asymmetry, while all effects from the production asymmetry are negligible.

Decay-time resolution model

The effect of the decay-time resolution model is studied by using alternative resolution models in the fit of the pseudo-experiment samples, which are generated with the nominal model. Two alternative models are considered, the calibration of the decay-time uncertainty based only on the width of the narrow and only of the wide Gaussian, respectively. The decay-time resolution is known to be correlated to the flavour tagging calibration. Therefore, the tagging calibration is altered together with the decay-time uncertainty calibration. The altered tagging calibration is obtained from decay-time fits to pseudo-experiment samples of $B_s^0 \to D_s^- \pi^+$ decays using the corresponding resolution model. The average of the systematic deviations with respect to the fits using the nominal resolution model is assigned as a systematic uncertainty. The impact on the hyperbolic parameters is small, while significant effects are observed for the oscillating terms.

Decay-time acceptance parametrisation

The coefficients of the acceptance parametrisation are correlated with each other and with the decay width Γ_s and its difference $\Delta\Gamma_s$. Hence, these parameters are not shifted individually. Instead, the fixed acceptance coefficients and the decay width (difference) are randomly drawn for each pseudo-experiment sample from a multi-variate Gaussian, with the mean being defined by the nominal values and the width by the covariance of the parameters. The found deviations are assigned as systematic. The systematic effects introduced by the acceptance are found to be negligible for the oscillation coefficients, while being sizeable for the hyperbolic terms, which are heavily correlated with the acceptance model. With a magnitude of above 20 % relative to the statistical uncertainty, the acceptance is found to be the dominant effect on these terms.

Similarly, the effect of the acceptance ratios is studied. For each sample, the ratios are randomly drawn from a multi-variate Gaussian defined by the nominal ratios and the according covariance. The deviations are found to be negligible in the oscillation terms and small but sizeable in the hyperbolic terms.

12.6 Data-driven studies of systematic effects

Systematic effects introduced by the mass-fit model in Section 10.3 and by the Gaussian-constrained flavour-tagging calibration discussed in Section 11.4 are studied in a data-driven approach. Since no information on correlations between the *CP* parameters is available in this approach, conservatively no correlations are assumed, yielding only diagonal contributions to the correlation matrix given in Table 12.2.

The mass model used to extract the signal component using sWeights in Section 10.3 includes a large number of fixed parameters. These include yields and fractions of background components calculated from known branching ratios and simulation-based selection efficiencies, as well as shape parameters derived from simulation. As the number of parameters is large, the effects of possible variations are not studied using a full pseudo-experiment study. Instead, the mass fit is repeated [126] varying each fixed parameter up and down within its uncertainties. Based on the resulting sWeights the decay-time fit is performed to reevaluate the *CP* parameters for each alternative mass model. The largest deviations observed are in the order of 5 −12 % and assigned as a systematic uncertainty.

The flavour-tagging calibration is Gaussian-constrained in the decay-time fit. this way, the uncertainties of the calibration parameters are propagated and already included in the statistical uncertainties of the fitted parameters. Hence, no systematic uncertainty for the finite knowledge of the calibration needs to be assigned. However, the strategy itself might have a biasing impact on the decay-time fit. These effects are studied by repeating the fit with alternative strategies for the flavour-tagging calibration. In one approach, the fit is performed using a fixed calibration. In another, the possible parameter space is expanded either by neglecting the correlations among the parameters or by constraining not only the statistical uncertainties from the $B_s^0 \to D_s^- \pi^+$ decay-time fit but also the systematic uncertainties on the calibration derived in Section 11.4.2. The deviations with respect to the nominal result found using the alternative calibration strategies are reported in Table 12.11. To assign a systematic uncertainty, the observed deviations of the different strategies are summed in quadrature. The corresponding uncertainty is found to be the dominant systematic effect on the parameter C with a relative magnitude of 26 %.

In the following, all systematic uncertainties found in the previous will be accounted for, when the CKM angle γ is extracted from the fitted *CP* parameters and the corresponding uncertainties.

Table 12.11: Deviation of the fitted *CP* parameters using alternative flavour-tagging strategies with respect to the nominal result.

Calibration	C	D_f	$D_{\bar{f}}$	S_f	$S_{\bar{f}}$
Fixed	-0.0020 ± 0.0003	-0.0009 ± 0.0002	0.0005 ± 0.0001	-0.0007 ± 0.0001	0.0009 ± 0.0002
Neglected correlations	-0.0001 ± 0.0001	-0.0014 ± 0.0002	-0.0004 ± 0.0001	-0.0001 ± 0.0001	0.0000 ± 0.0001
Constrained systematics	0.0155 ± 0.0033	0.0031 ± 0.0006	0.0034 ± 0.0007	-0.0008 ± 0.0018	-0.0058 ± 0.0019
Systematic	0.0156	0.0035	0.0035	0.0010	0.0059

13 Determination of the CKM angle γ

The five *CP* parameters extracted in Chapter 11 can be used to constrain the CKM angle γ as explained in Section 3.3.2. The combination is performed using the GammaCombo framework [127, 128]. A likelihood-ratio test is performed to estimate the parameter set of the amplitude ratio $r_B^{D_sK}$, the strong phase difference δ and the CKM angle γ fitting the *CP* parameters measured in Section 11.5 best. The relation of the five *CP* parameters and the parameters of interest are described by Equation 3.22 and 3.60. External input is used to constrain the weak mixing phase $\phi_s = -0.031 \pm 0.018$ [129]. The likelihood scan yields the CKM angle

$$\gamma = (74 \pm 11)^\circ ,$$

as well as the strong phase and the amplitude ratio

$$\delta = (346.9 \pm 6.6)^\circ , \qquad r_B^{D_sK} = 0.327 \pm 0.038 .$$

The reported uncertainties of the three parameters describe the 68 % confidence interval. In the fit, the combined covariance matrix of both the statistical uncertainties from Chapter 11 and the systematic effects found in Chapter 12 is constrained. The constrained parameters are visualised in Chapter 13. The confidence interval of the angle γ is shown in Chapter 13.

The compatibility of the determined parameters with respect to the previous result $\gamma_{\mathrm{DsK}}^{3\,\mathrm{fb}^{-1}} = \left(128^{+17}_{-22}\right)^\circ$ [14] is evaluated. The agreement is found to be $1.8\,\sigma$ corresponding to a p-value of 7 %. This small deviation is driven by the compatibility of $2.5\,\sigma$ for the CKM angle γ. The strong phase ($0.7\,\sigma$) and the amplitude ratio ($0.4\,\sigma$) individually show good agreement.

The small tension to the previous result in γ is not considered to be an issue. The statistical precision is increased by a factor of two, while a good agreement to the current world average $\gamma^{\mathrm{avg}} = \left(66.2^{+3.4}_{-3.6}\right)^\circ$ [42] is observed.

As a further cross-check, the unbiasedness and the coverage are tested using the pseudo experiments discussed in Section 12.5. For a subsample of 500 pseudo experiments, the CKM angle γ is extracted from the fitted *CP* parameters. The pull distribution of the extracted angles shows good coverage with a width of $\sigma^\gamma = 1.12 \pm 0.04$ and no hint at a bias is found in the mean $\mu^\gamma = 0.04 \pm 0.05$.

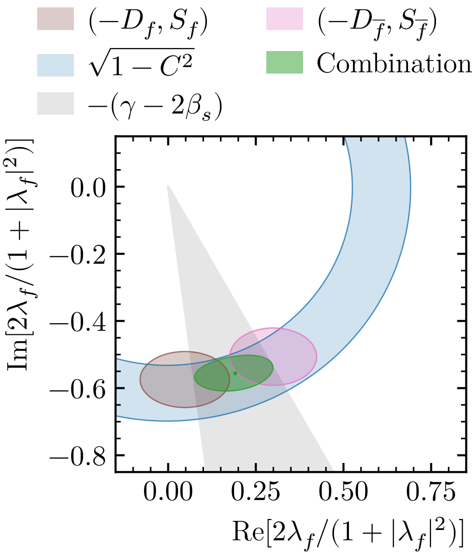


Figure 13.1: Combination of the constrained *CP* parameters. The visualised areas represent the 68 % (39 % in case of the ellipses) confidence level and are based on the combined covariance of both statistical and systematic uncertainties.

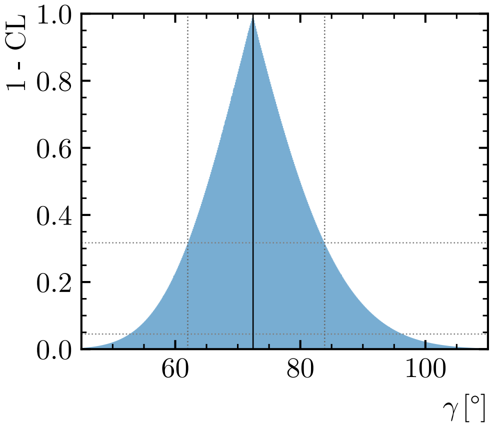


Figure 13.2: Confidence intervals of the CKM angle γ, accounting both for statistical and systematic uncertainties.

The precision of the CKM angle γ from $B_s^0 \to D_s^{\mp} K^{\pm}$ can be increased further by combining the presented result with the previous result, which is extracted from an independent data set. This combination is currently in preparation as it requires careful handling of the updated fixed inputs in both the fit to the Run1 data set and the evaluation of the corresponding systematics. The updated Run1 result and the combination will be submitted to a journal together with the presented measurment using the Run2 data set in the future. Hence, this measurment is preliminarily published by the LHCb collaboration as a stand-alone result [2].

Review of the author's contributions

After gaining expert knowledge in the analysis of $B_s^0 \to D_s^- \pi^+$, I have been involved in this analysis from the start to the publication of the result [2]. Further, I have been responsible for large extents of the analysis, especially in the decay-time-dependent part and in the evaluation of systematic effects presented in Chapter 11 and Chapter 12, respectively.

In the decay-time-dependent analysis, I am responsible for performing the decay-time fit discussed in Section 11.5. This covers the extraction of the decay-time-acceptance and its correction as discussed in Section 11.3, as well as the implementation of the constrained flavour tagging to the existing B2DXFitter framework, as discussed in Section 11.4. In addition, I rerun the full analysis of $B_s^0 \to D_s^- \pi^+$ decays to accommodate the constrained input values to match the strategies that have been updated with respect to Part II.

Except for effects from the MDFit model, I have evaluated all systematic uncertainties on the *CP* parameters listed in Table 12.1. For these evaluations, I have performed the full set of studies and cross-checks presented in Chapter 12, which include the analysis of fully-simulated samples, bootstrapped samples and pseudo experiments, as well as data-based compatibility studies. In the course of these studies, I have adapted the fit strategy to accommodate the different characteristics of the various samples used. This includes rerunning the full analysis procedure, including the decay-time-independent part, of both the $B_s^0 \to D_s^\mp K^\pm$ and $B_s^0 \to D_s^- \pi^+$ analyses.

As an expert and as the major contributor to this analysis I have become contact author of the publication, responsible for the correspondence and defence of the analysis in and outside of the LHCb collaboration. Accordingly, I will give the first presentation on the analysis to the public at the CKM conference [130].

Part IV

Prospects And results

14 Future improvements of the experimental setup

While the analyses presented in Part II and Part III already provide impressive precision, future developments will enable further improvements. In the following, the potential future sensitivity to the parameters will be discussed in Section 14.3. Two major ingredients to increase the sensitivities are the LHCb Upgrade I, reviewed in Section 14.1, and an inclusive flavour tagger, which is introduced in Section 14.2.

14.1 The LHCb Upgrade I

The analyses presented in this thesis are based on data taken in the period between the years 2015 and 2018, the so-called LHC Run2. However, a new data-taking period, the LHC Run3 started in 2022. In this data-taking period, which is scheduled to continue until the end of 2025, additional data will be recorded at the LHCb experiment. While the centre-of-mass energy of the *pp*-collisions is slightly increased to $\sqrt{s}$ = 13.6 TeV, the LHCb collaboration aims to increase the instantaneous luminosity of the collisions recorded by a factor of five to $\mathcal{L} = 2{\cdot}10^{33}\,\text{cm}^{-2}\text{s}^{-1}$. This way, the recorded LHCb data set is expected to be doubled by the end of LHC Run3 and to reach a total of 50 fb^{-1} by the year 2030.

The increased luminosity brings a proportionally increased detector occupancy. To accommodate these changes while maintaining the previous detector performances, most of the detector systems described in Section 4.2 have been replaced in the course of the LHCb Upgrade I [131, 132]. Also, the trigger system introduced in Section 4.3 has been upgraded. In the following, a brief overview of some important aspects of the upgrades is given.

The efficiency of the L0 hardware trigger, which was already close to saturation in the LHC Run2, cannot easily be improved. Instead, the corresponding selection processes are completely integrated into the HLT. The removal of the L0 trigger requires a detector readout at the LHC bunch-crossing rate of upto 40 MHz. All front-end and readout electronics of the detector are upgraded to accommodate these challenging needs.

Next to the increase of the hadronic trigger efficiencies by a factor of 2 - 4, the upgraded trigger system brings the possibility to perform the detector alignment calibration already during data taking, improving the selection quality. However, this comes at the cost of expensive computations. Hence, the event filter farm is upgraded and a large part of the computations will be performed using GPUs. Additionally, the output of the HLT1 can be buffered on disk allowing to perform the HLT2 computations asynchronously to the data taking.

In the course of the upgrade, a large fraction of the detector submodules have been replaced. The full tracking system (except for the dipole magnet) is replaced. Additionally, the first muon station M1 is removed together with the pre-shower and the scintillating pad detector, as these three components were mainly used for the L0 trigger. The only components kept are the electromagnetic and hadronic calorimeters, as well as the following muon chambers. However, even here the front-end electronics are replaced to accommodate the updated data-taking and trigger strategy.

The VELO is upgraded by a pixel-based silicon detector [133]. This design yields a better hit resolution and is able to handle the increased occupancy. Additionally, the upgraded VELO comes with a decreased material budget and will be placed closer to the interaction region, at a distance of 5.1 mm. Overall, these improvements are not only expected to maintain the current performance in the new data-taking conditions, but also to improve the impact-parameter and decay-time resolution as well as the tracking efficiency significantly.

The upstream tracker is replaced by an improved silicon strip detector [134]. While the working principle of this component is not changed, the design is updated. Improved silicon material is used to achieve the radiation hardness needed for higher luminosity. Further, the segmentation and geometric acceptance are increased. Additionally, the downstream tracking stations are replaced by a new system based on scintillating fibres. This system comes with better segmentation than the previous gas detector used in the OT, as well as better radiation hardness and a lower material budget compared to the silicon-based IT. Overall, the full tracking system after the upgrade is expected to achieve a slightly reduced tracking efficiency, while increasing the momentum resolution.

The two RICH detectors have received an upgrade of their mechanics and readout electronics, which have been upgraded for all detector components to accommodate. The improved mechanics increase the focal length of the mirrors significantly, which effectively reduces the detector occupancy by a factor of 2. This is essential to provide reliable PID information in the upgraded data-taking conditions.

14.2 Prospects on inclusive flavour tagging

The statistical precision of the presented analyses can be improved not only by increasing the data statistics but also by increasing the statistical power of the data. The effective tagging efficiency $\varepsilon_{\text{eff}} \approx 6\,\%$ observed in the presented analyses is a large achievement as it is a relative improvement of about 20 % compared to the previous analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays [14], and more than 50 % concerning the previous $B_s^0 \to D_s^- \pi^+$ analysis [107]. Though, this leaves large space for improvements with 94 % of all selected B_s^0 candidates being effectively unused.

However, it must be stated that a perfect tagging power $\varepsilon_{\text{eff}} = 100\,\%$ can conceptually not be achieved, even with a hypothetical perfect detector, as the $B_{(s)}^0$ hadronisation process itself can still be ambiguous. Nevertheless, there is still room for improvements expected as the B factories BaBar and Belle(II) reach tagging powers in the order of 30 %. This performance is based on the clean leptonic collision environment, which allows for a good reconstruction and association of the associated B meson produced on the OS. However, the hadronic environment at the LHC enables fragmentation processes, from which the LHCb experiment benefits in the form of the SS taggers.

Currently, a new approach to maximise the flavour tagging performance is being investigated at the LHCb experiment, an inclusive flavour tagger (InclFT). Here, all tracks reconstructed in the event are evaluated simultaneously instead of analysing individual processes. This way, all possible fragmentation processes, both on the SS and the OS, and all possible decay chains of the OS B meson can potentially be exploited independently on the specific particle species, instead of using only the processes explicitly implemented in a specific algorithm. Besides, correlations among the different processes can be exploited in this approach.

The inclusive tagger prototype

In the course of this thesis, a prototype of this tagger has been developed and made available within the LHCb software stack. This prototype is based on a recurrent neural network (RNN) for the classification of the initial B flavour, which is designed to handle sequential inputs as discussed in Section 5.1.2. Each non-signal track of an event is considered a sequence with individual features. In total 18 track features, based on the properties commonly used in the existing flavour taggers, are used for the classification. These include PID information and kinematic properties of the track and the system of the signal candidate and the track. Additionally, the track's charge is considered as this property is crucial to determine the initial flavour of the track and the possibly associated signal candidate.

The implementation of this prototype is based on the Keras package of the TensorFlow library [102, 103]. This implementation of the gated recurrent units does not allow for a variable sequence length (track number). Hence, only a fixed number of 40 tracks per event are evaluated. In the case of fewer tracks, the missing information is padded. If more tracks are reconstructed in the event, only the 40 tracks with the highest transverse momentum are considered. However, the impact of this cut-off is studied by an alternative model evaluating up to 100 tracks per event. No significant change in the classification performance of the algorithm is observed.

The training of the classifier requires a labelled data set. Hence, simulated decays with known initial B flavour are used. As the hadronisation process of the signal depends on the specific B-meson species, three separated classifiers are trained. A classifier for B^0 is trained on simulated $B^0 \to J/\psi K^{*0}$ decays and a classifier for B_s^0 on $B_s^0 \to D_s^- \pi^+$ simulation. Additionally, $B^+ \to J/\psi K^+$ simulation is used to train a tagger for the non-oscillating B^+ meson as a cross-check.

The prototype InclFT achieves significant improvements compared to the combination of the established taggers described in Section 4.5 in simulation. For simulated $B^0 \to J/\psi K^{*0}$ decays the tagging power after calibration is increased from $5.15 \pm 0.09\,\%$ to $6.62 \pm 0.08\,\%$, for simulated $B_s^0 \to D_s^- \pi^+$ decays from $8.31 \pm 0.05\,\%$ to $9.47 \pm 0.04\,\%$, and for $B^+ \to J/\psi K^+$ simulation from $6.43 \pm 0.03\,\%$ to $10.13 \pm 0.04\,\%$. The large improvement in the latter, charged mode has to be taken with caution as the established SS taggers are not optimised for this mode.

The relative improvement of the flavour tagging performance observed in recorded data is smaller, but still significant. Application of the InclFT to the sWeighted $B_s^0 \to D_s^- \pi^+$ data set obtained in Chapter 6 yields a calibrated tagging power of $6.43 \pm 0.16\,\%$. This is equivalent to a relative improvement of $7\,\%$ with respect to the performance reported in Table 7.3. This improvement in statistical power matches the improvement in statistical sensitivity well when the decay-time fit from Section 7.4 is performed using the InclFT, yielding a measured oscillation frequency $\Delta m_s^{\text{InclFT}} = 17.7646 \pm 0.0048\ \text{ps}^{-1}$. However, it must be stated here, that the systematic implications of using this tagger have not yet been studied and that the change of the central value is significant, assuming strong correlations among the two fits.

Ongoing developments

The algorithm discussed in the previous is the first prototype for inclusive flavour tagging at the LHCb experiment and no publications upon its development or application have been made yet. Instead, various improvements and extensions to this approach are currently in development.

The DeepSet architecture [135] is a promising alternative to the RNN. In this approach, all event tracks are evaluated independently instead of sequentially. Further, the current PyTorch-based DeepSet implementation [136] allows for a variable number of per-event tracks. Besides, the DeepSet model can be trained one order of magnitude faster compared to the prototype algorithm, boosting further developments and studies.

While the established tagging algorithms are carefully designed to exploit specific hadronisation processes in the context of the B-meson production, the classifier of the InclFT is expected to reconstruct these processes itself. One approach to increase the performance of the classifier is to constrain the knowledge of the physical processes in the training process. A pre-classification of all tracks shows the potential for a relative improvement of up to 10 % in a master thesis. Four track categories are considered in this study, tracks originating from the signal fragmentation, from the decay or from the fragmentation of the accompanying b hadron on the OS, and tracks unrelated to the signal production.

The InclFT training is simulation-based. Hence, the simulation-data mismatch is a possible reason for the reduced improvement of the tagging performance observed in data with respect to the performance on simulation. An approach to train the InclFT with data shows further improvements. Here, the InclFT is split into OS and SS. While the SS classifier still needs to be trained using simulation specific to the B-meson species, the OS classifier can commonly be used for both neutral $B^0_{(s)}$ and charged B^+ mesons. This allows training the OS classifier on data of a self-tagging charged mode like $B^+ \to J/\psi K^+$. Another strategy to possibly overcome a simulation-data mismatch is domain adaption [137, 138], which will be studied in the context of inclusive flavour tagging in the future.

14.3 Outlook on decay-time-dependent analyses of B_s^0-meson decays

Competition and combined effort in the experimental landscape can be very beneficial, as the discovery of the Higgs boson in 2012 [17, 18] based on data from both the ATLAS and CMS experiments demonstrates. However, the sector of hadronic B_s^0-meson decays is expected to be still dominated by LHCb measurements in the upcoming years. Currently, no new experiment capable to collect a significant B_s^0-meson data set is under construction or planned within the current decade. Consequently, competition is only expected from the already data-taking experiments Belle2, ATLAS and CMS.

The Belle2 experiment, which provides many competitive measurements in the B meson sector, records only a small subset of its data at the energy of the $\Upsilon(5S)$ resonance. Further, this resonance only decays at a rate of about 20 % [35] into B_s^0 mesons. Hence, the B_s^0-meson sample recorded at Belle2 will not provide sufficient statistics to contribute to the field before the Belle2 upgrade scheduled for the year 2027 and presumably further. The ATLAS and CMS collaborations have published competitive measurements [139, 140] of the mixing phase ϕ_s in $B_s^0 \to J/\psi\,\phi$ decays. However, these modes profit from the two narrow resonances, of which the charmonium resonance decays into two muons. The capability to perform competitive measurements in pure hadronic decay modes with challenging PID requirements has still to be proven.

Nevertheless, the statistical precision in B_s^0-meson analyses is expected to increase significantly in the upcoming years. The LHCb collaboration aims to collect about 23 fb^{-1} of data by the end of Run3. Together with the increased trigger efficiencies after the upgrade introduced in Section 14.1, this will increase the B_s^0 meson data set by a factor of 10. In the long term, a 50 fb^{-1} data set is expected at the end of the decade and 300 fb^{-1} until the end of the LHCb operations. Assuming similar selection efficiencies as in the presented analysis and a trigger and reconstruction efficiency improved by a factor of two, a data set of approximately $1.8{\cdot}10^6$ ($9.9{\cdot}10^4$) $B_s^0 \to D_s^- \pi^+$ ($B_s^0 \to D_s^\mp K^\pm$) candidates is expected during Run3 and additional $3.2{\cdot}10^7$ ($1.8{\cdot}10^6$) candidates until the end of the LHCb operations. In Figure 14.1, the sensitivity to the oscillation frequency Δm_s and the *CP* parameters is extrapolated for these data sets.

Although the tagging performance is expected to decrease in Run3 due to the increased luminosity, many possibilities arise to recover the performance and even improve it beyond. A key here is the quality of the simulation, which is expected to improve in the future. Also, the understanding and treatment of residual simulation-data mismatches will be crucial. Further, it is to mention, that improvements in the flavour tagging will not only increase the statistical power of future data sets, but could also be used to improve the data already recorded. The effective tagging efficiency of about 6 % observed in the presented analysis for example leaves large room to increase the statistical power of the data set. A potential inclusive flavour tagger will increase the effective sample size by 10 %. Even 20 % or more seem to be achievable at the moment for future evolutions of the algorithm.

The statistical precision is however not the only factor which must be considered. Although the two presented measurements are limited by the statistical sensitivity at the moment, a raising importance of the systematic effects can be foreseen as displayed in Figure 14.1. Assuming a data set increased by a factor of 10 by the end of Run3, the statistical uncertainty of the oscillation frequency Δm_s would possibly be decreased to 0.0017 ps^{-1}. Thus, the projected statistical sensitivity comes close to the dominant systematic uncertainties observed in Chapter 8. As these effects arise from the detector alignment, they cannot be reduced by more sophisticated analysis strategies and improvements in the detector development will be necessary to avoid systematical limitations. The upgraded VELO might help here with improved resolution and reduced material budget. However, the practical precision of the VELO alignment still needs to be evaluated.

The picture is slightly different in the measurement of the CKM angle γ with $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. Due to its Cabibbo suppression, the branching ratio and the observed signal yield are reduced by more than an order of magnitude compared to $B_s^0 \to D_s^- \pi^+$. Further, the majority of the systematic uncertainties observed in Chapter 12 arise from external inputs. The majority of systematic effects are not directly dependent on the detector performance. Hence, these effects are expected to either scale with the statistics or to be reducible by increased simulation statistics and improved simulation quality. However, from the extrapolation in Figure 14.1 a systematic limitation of the *CP* parameters C and $S_{f(\overline{f})}$ is expected before the end of the LHCb operations. This limitation, which is driven by the effects introduced by the finite decay-time resolution, will presumably not play a role during the LHC Run3 and is in general not expected to affect the parameters $D_{f(\overline{f})}$. The systematic contribution to the uncertainty budget of the CKM angle γ cannot be evaluated as easily. Nevertheless, as the *CP* parameters are not expected to be systematically limited in Run3, the same expectation applies to the CKM angle. Further, the current tensions among the time-dependent measurements of γ are driven by the parameters $D_{f(\overline{f})}$, which only shows minor sensitivity to resolution effects. Hence, the sensitivity from $B_s^0 \to D_s^{\mp} K^{\pm}$ decays to the CKM angle γ is expected to be statistically limited at the level of $\sigma(\gamma) \approx 3.5^{\circ}$ at the end of Run3, assuming a data set increased by a factor 10 again. This results in a significant constraint to the tensions observed in this sector, the Standard Model and possible New Physics effects.

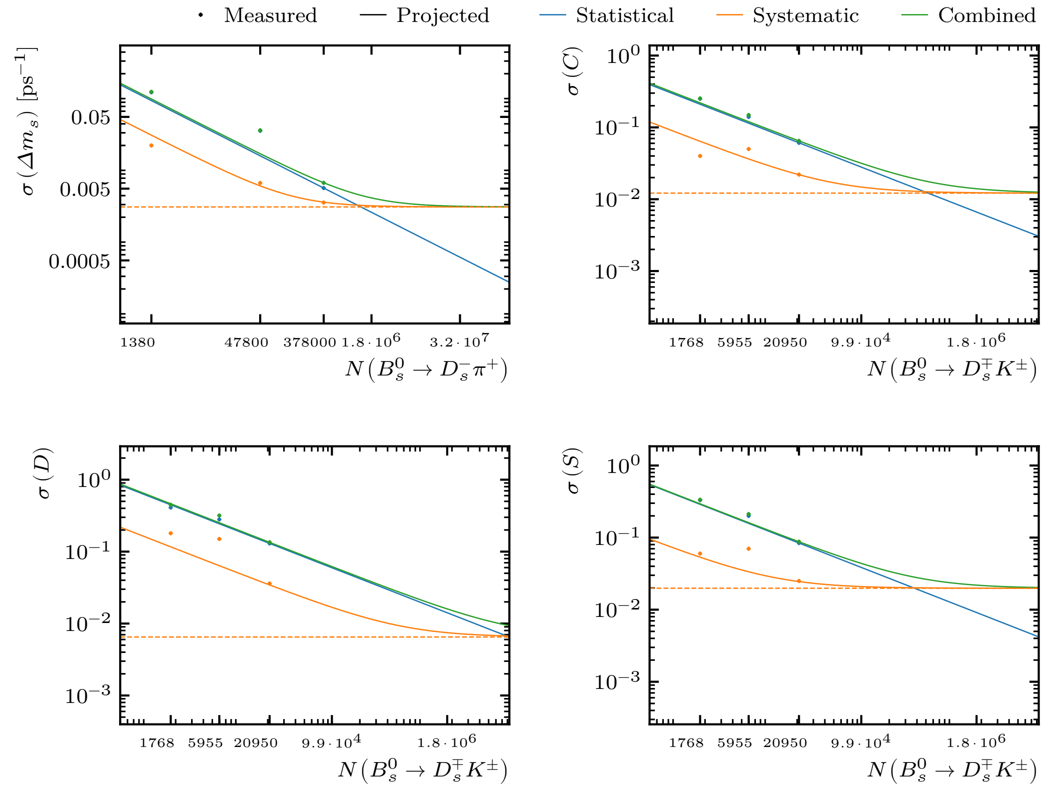


Figure 14.1: Projections of the sensitivities of the oscillation frequency Δm_s from $B_s^0\to D_s^-\pi^+$ decays (top left), and the *CP* parameters C (top right), $S_{f(\bar{f})}$ (bottom left) and $D_{f(\bar{f})}$ (bottom right) from $B_s^0\to D_s^\mp K^\pm$ decays each. The projections are based on uncertainties evaluated in the Run2 measurements performed in Part II and III, respectively. The statistical uncertainties (blue) assume a constant statistical power of the data samples, which is primarily affected by the dilution from the decay-time resolution and the flavour tagging. The projection of the systematic effects (orange) assumes the uncertainties from the VELO alignment and the decay-time resolution to be constant and all other uncertainties to scale with the statistics. The combined sensitivity (green) incorporates both statistical and systematic effects. The data points corresponding to previous measurements of the quantities[14, 107, 141, 142] show a trend similar to the projections.

15 Summary

The oscillation of neutral B_s^0 mesons provides experimental access to important properties of the Standard Model of particle physics. Due to its well-suited instrumentation, the LHCb experiment provides high-quality data of the large amount of B_s^0 mesons produced in the high energy proton-proton collisions at the LHC. In the data set recorded between 2015 and 2018 at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of $\mathcal{L}_{\text{int}} = 6\ \text{fb}^{-1}$, a sample of $378\,700 \pm 700$ $B_s^0 \to D_s^- \pi^+$ decays is reconstructed. These large statistics, in combination with the high decay-time resolution of the LHCb experiment, enable the world's most precise measurement of the oscillation frequency [1]

$$\Delta m_s = (17.7683 \pm 0.0051 \pm 0.0032)\ \text{ps}^{-1}$$

to date. The result is in agreement with previous measurements, as well as with theory predictions [41, 46]. Because of the high precision, this measurement dominates the updated LHCb-wide average of the oscillation frequency [1]

$$\Delta m_s = (17.7656 \pm 0.0057)\ \text{ps}^{-1},$$

calculated in the context of this measurement. The measured oscillation frequency can be used to constrain the unitarity triangle, an important closure test of the Standard Model. Together with measurements of the decay-width differences $\Delta\Gamma_s$ it can be used to constrain possible New Physics effects, as many of the theory uncertainties cancel in the ratio of these two quantities. Further, the oscillation frequency is a crucial ingredient to all decay-time-dependent studies of B_s^0-meson decays. Such a measurement is the presented analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. Significant *CP* violating effects are present in this decay mode due to large interference effects of equally suppressed decay paths. In total, $20\,950 \pm 180$ $B_s^0 \to D_s^{\mp} K^{\pm}$ decays are reconstructed and selected. The decay-time-dependent analysis of this sample allows a precise determination of the five *CP* parameters

$$\begin{aligned} C &= 0.791 \pm 0.061 \pm 0.022\,, \\ D_f &= -0.051 \pm 0.134 \pm 0.037\,, & D_{\bar{f}} &= -0.303 \pm 0.125 \pm 0.036\,, \\ S_f &= -0.571 \pm 0.084 \pm 0.023 \quad \text{and} & S_{\bar{f}} &= -0.503 \pm 0.084 \pm 0.025\,. \end{aligned}$$

A significant improvement of the sensitivity in comparison to the analysis of the 2011 and 2012 data set [14] is observed. This increase of precision is achieved not only by the larger data set but also by the improved experimental techniques and methods and by improved knowledge of input parameters such as the oscillation frequency Δm_s. The CKM angle

$$\gamma = (74 \pm 11)^\circ$$

can be extracted from the obtained *CP* parameters, as well as the strong phase $\delta = (346.9 \pm 6.6)^\circ$ and the amplitude ratio $r_B^{D_sK} = 0.327 \pm 0.038$ [2]. This result is the most precise decay-time-dependent measurement of the angle γ to date. The measured CKM angle can be probed against time-integrated measurements to constrain the size of New Physics effects in the B_s^0-meson mixing and in the amplitudes of $B_s^0 \to D_s^\mp K^\pm$ decays. In the future, this result will be included in the upcoming combination of γ measurements. This combination is used to constrain the CKM sector of the Standard Model, which today is not able to accommodate the observed baryon asymmetry in the universe.

While the presented analyses provide impressive precision already, future improvements are expected, as the upgraded LHCb experiment continues data taking and improved experimental techniques, such as inclusive flavour tagging, are under development to maximise the statistical power of the data. These improvements will increase the statistical sensitivity significantly in the coming years. On the other hand, the detector performance and alignment will play a more important role in the future. Without improvements in the detector alignment, the sensitivity to the oscillation frequency is expected to be systematically limited at the level of $\sigma(\Delta m_s) \approx 0.0028\ \text{ps}^{-1}$ already during the ongoing data-taking period of LHC Run3. At the same time, the time-dependent measurement of the CKM angle γ is expected to reach a possible sensitivity of $\sigma(\gamma) \approx 3.5^\circ$, which will still be statistically limited.

Part V

Supplements

A Technical details on the analyses of $B_s^0 \to D_s^- \pi^+$ decays

In the following, details on the analyses of $B_s^0 \to D_s^- \pi^+$ decays are given, which are not essential to understand the measurement, but are helpful for experts reproducing the measurement.

A.1 Signal selection

Several preselection steps are performed globally by the LHCb collaboration to handle the large amount of data collected by the detector. Next to the trigger discussed in Section 4.3, which reduces the data rate already during data taking, further preselections are applied in the processing of the data, before it is provided to the community. Within the LHCb collaboration, this processing is also referred to as stripping. In contrast to the trigger requirements, these selections can be revised as the data passing the trigger is preserved on tape. However, due to their computational intensity, these selections are not performed regularly but in collaboration-wide campaigns. In these campaigns, different sets of requirements, so-called stripping lines, are applied, with the accepted data having to pass at least one of these sets. Each set represents a preselection common to a group of analyses. In this selection process, the reconstructed decay trees are refitted, including the individual tracks and vertices. In contrast to the trigger selection, this part of the data processing is performed offline avoiding constraints on the computation time. Therefore, a variety of constraints can be used in the fits used for the decay reconstruction, allowing for more sophisticated requirements. Additionally, updated detector alignments can be considered.

Only loose trigger requirements are applied in this analysis. No specific requirements at the level of the L0 hardware trigger are applied. At the level of the HLT1, it is required that either the TrackMVA or the TwoTrackMVA line is triggered on a signal track to filter events including displaced tracks with high transverse momentum. In the HLT2, the requirements of the 2-/3-/4-body topo line or the inclusive-ϕ line need to be fulfilled. These lines are triggered by the presence of a displaced multi-track vertex with sizeable transverse momentum.

In the further preselection, the requirements of the StrippingB02DPiD2HHHBeauty2CharmLine stripping line are used. This line is designed to select candidates for $B^0_{(s)}$ mesons based on displaced secondary vertices (SV) formed by a charged $D^-_{(s)}$-meson candidate and a charged accompanying hadron. The $D^-_{(s)}$-meson candidate itself is built by three charged hadrons forming a displaced tertiary vertex (TV). The combined $D^-_{(s)}$-meson candidate's momentum is required to point to the SV, and the combined $B^0_{(s)}$ momentum to the primary vertex (PV). No criteria based on particle identification (PID) information is used at this stage. Hence, both all $B^0_{(s)} \to D^-_{(s)} h^+$ decays are covered by this line. Different stripping versions are applied to the different years of data taking. Version s24r1 is applied to the 2015 data, s28r1 to 2016, s28r2 to 2017 and s34 to 2018. However, the selection defined by the line does not depend on the applied version. The full selection of this stripping line is given in Table A.1 and discussed in the following.

All four final-state tracks are required to have a well-reconstructed track fitted with a $\chi^2_{\text{track}}/\text{ndf} < 4$ and a neural-network-based probability $P(\text{ghost})$ to be a ghost built from random detector hits below 40 %. Additionally, the impact parameter with respect to any reconstructed PV is required to show a significant displacement of the tracks $\text{IP}^{\text{PVs}}_{\chi^2} > 4$. All signal tracks have to provide sizeable transverse $p_{\text{T}} > 100\,\text{MeV}/c$ and absolute $p > 1\,\text{GeV}/c$ momenta. The track assigned to the accompanying hadron has tighter momentum requirements $p_{\text{T}} > 500\,\text{MeV}/c$ and $p > 5\,\text{GeV}/c$.

The transverse momenta of the three tracks forming the $D^-_{(s)}$ candidate are required to sum a value larger than $1.8\,\text{GeV}/c$, with at least one track having $p_{\text{T}} > 500\,\text{MeV}/c$ and $p > 5\,\text{GeV}/c$. Further, the invariant mass from the combination of the three hadrons' four-vector is expected to lay within a $100\,\text{MeV}/c^2$ window around the known mass of either the D^- or D^-_s meson. To form a reasonable vertex, the three tracks are required to provide a pairwise distance of closest approach (DOCA) of 0.5 mm at maximum. To save computation time, the aforementioned requirements are applied before the vertex fit of the $D^-_{(s)}$ candidate. The fit is then expected to provide a high-quality vertex with $\chi^2_{\text{vtx}}/\text{ndf} < 10$ and a significant displacement $\min \text{IP}^{\text{PVs}}_{\chi^2} > 16$ with respect to any reconstructed PV.

The fit for the $B^0_{(s)}$ candidate has to yield a high-quality vertex $\chi^2_{\text{vtx}}/\text{ndf} < 10$ combining the accompanying track and the $D^-_{(s)}$ candidate. A displacement from the PV is required in terms of the reconstructed $B^0_{(s)}$ lifetime being longer than 0.2 ps. But still, the trajectory of the reconstructed $B^0_{(s)}$ candidate has to be compatible with originating from the PV. Therefore, the impact parameter significance with respect to the associated PV has to be $\text{IP}^{\text{PV}}_{\chi^2} < 25$ and the cosine of the angle between the reconstructed momentum and the line connecting PV and reconstructed SV, known as direction angle (DIRA), is required to be above 0.999.

Additionally, the stripping line expects the event to provide less than 500 tracks and to have the topological multi-body trigger lines or the inclusive-ϕ line to be triggered either by a signal track (TOS) or by an independent track (TIS). However, these requirements are already fulfilled by the analysis' trigger requirements.

The offline selection discussed in Section 6.2 is summarised in Table A.2 in a more handy format.

Table A.1: Summarised preselection of $B^0_{(s)} \to D^{\mp}_{(s)}(\to h^{\mp}h^{+}h^{-})\, h^{\pm}$ candidates applied in the stripping process.

Candidate	Observable		Cut	Unit
All	$\chi^2_{\text{track}}/\text{ndf}$	$<$	4	
	$P(\text{ghost})$	$<$	0.4	
	$\text{IP}^{\text{PVs}}_{\chi^2}$	$>$	4	
	p_{T}	$>$	100	MeV/c
	p	$>$	1000	MeV/c
Accompanying $h^{\pm}$	p_{T}	$>$	500	MeV/c
	p	$>$	5000	MeV/c
$D^{\mp}_{(s)}$ children	$\chi^2_{\text{vtx}}/\text{ndf}$	$<$	10	
	$\min \text{IP}^{\text{PVs}}_{\chi^2}$	$>$	16	
(pairwise)	$\text{DOCA}(h^{\pm}_i,\, h^{\pm}_j)$	$<$	0.5	mm
	$\lvert m(h^{\mp}h^{+}h^{-}) - m^{\text{PDG}}_{D^{-}_{(s)}}\rvert$	$<$	100	MeV/c^2
	$\sum p_{\text{T}}$	$>$	1800	MeV/c
(at least one child)	p_{T}	$>$	500	MeV/c
(at least one child)	p	$>$	5000	MeV/c
$B^0_{(s)}$ candidate	t	$>$	0.2	ps
	$\text{IP}^{\text{PV}}_{\chi^2}$	$<$	25	
	$\chi^2_{\text{vtx}}/\text{ndf}$	$<$	10	
	DIRA: $\cos\measuredangle(\vec{p},\, \overrightarrow{\text{SV}} - \overrightarrow{\text{PV}})$	$>$	0.999	

Table A.2: Full offline selection used to suppress backgrounds in the $B_s^0 \to D_s^- \pi^+$ signal mode. The origin associated with the final-state hadrons is denoted by a subscript to prevent ambiguities. Additionally, all final-state tracks are required to have associated information from the RICH system and are expected to not appear in the muon system. Further, the BDT selection discussed in Section 6.1.3 is applied.

Decay mode		Requirement			Unit
All modes		$m(D_s^- \pi^+_{B_s^0})$	$\in$	[5300, 5800]	MeV/c^2
		$m(h^-_{D_s^-} h^+_{D_s^-} h^-_{D_s^-})$	$\in$	[1920, 2015]	MeV/c^2
		$t_{B_s^0}$	$\in$	[0.4, 15]	ps
		$\sigma(t_{B_s^0})$	$\in$	[0.01, 0.1]	ps
		$\mathrm{PID}_K(\pi^+_{B_s^0})$	$<$	0	
		$\mathrm{PID}_e(\pi^+_{B_s^0})$	$<$	5	
		$\mathrm{PID}_e(\pi^-_{D_s^-})$	$<$	5	
$D_s^- \to K^- K^+ \pi^-$	or	$\mathrm{PID}_K(K^-_{D_s^-})$	$>$	10	
		$m(K^-_{D_s^-} K^+_{D_s^-} \pi^-_{D_s^-})_{K^- \to \pi^-}$	$\notin$	[1839, 1899]	MeV/c^2
	or	$\mathrm{PID}_K(K^-_{D_s^-}) - \mathrm{PID}_p(K^-_{D_s^-})$	$>$	5	
		$m(K^-_{D_s^-} K^+_{D_s^-} \pi^-_{D_s^-})_{K^- \to \overline{p}}$	$\notin$	[2255, 2315]	MeV/c^2
		$\mathrm{PID}_p(\pi^-_{D_s^-})$	$<$	10	
		$m(K^+_{D_s^-} K^-_{D_s^-})$	$<$	1800	MeV/c^2
		$m(\pi^+_{B_s^0} K^-_{D_s^-})$	$\notin$	[1834, 1894]	MeV/c^2
$\to \phi \pi^-$		$\mathrm{PID}_K(K^\mp_\phi)$	$>$	-2	
		$m(K^+_\phi K^-_\phi)$	$\in$	[1000, 1040]	MeV/c^2
$\to K^{*0} K^-$		$\mathrm{PID}_K(K^-_{D_s^-})$	$>$	5	
		$\mathrm{PID}_K(K^+_{K^{*0}})$	$>$	-2	
		$m(K^+_{K^{*0}} K^-_{D_s^-})$	$\notin$	[1000, 1040]	MeV/c^2
		$m(K^+_{K^{*0}} \pi^-_{D_s^-})$	$\in$	[842, 942]	MeV/c^2
$\to (K^- K^+ \pi^-)_{\mathrm{NR}}$		$\mathrm{PID}_K(K^\mp_{D_s^-})$	$>$	5	
		$\mathrm{PID}_K(\pi^-_{D_s^-})$	$<$	10	
		$m(K^+_{D_s^-} K^-_{D_s^-})$	$\notin$	[1000, 1040]	MeV/c^2
		$m(K^+_{D_s^-} \pi^-_{D_s^-})$	$\notin$	[842, 942]	MeV/c^2
$D_s^- \to \pi^- \pi^+ \pi^-$		$\mathrm{PID}_K(\pi^\mp_{D_s^-})$	$<$	2	
		$\mathrm{PID}_p(\pi^\mp_{D_s^-})$	$<$	5	
		$m(\pi^+_{D_s^-} \pi^-_{D_s^-})$	$<$	1700	MeV/c^2
		$m(\pi^+_{B_s^0} \pi^-_{D_s^-})_{\pi^+ \to K^+}$	$\notin$	[1834, 1894]	MeV/c^2

A.2 Mass fit

In Table A.3, the floated shape parameters of mass models introduced in Section 6.3 are listed for the $B_s^0 \to D_s^- \pi^+$ signal and the combinatorial background component.

Table A.3: Fitted values of the shape parameters floating in the multi-dimensional mass fit. The mean μ of the signal shapes is shared among the four D_s^- final states in both mass spectra.

Parameter	Year	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to (K^-K^+\pi^-)_{\rm NR}$	$D_s^- \to \pi^-\pi^+\pi^-$
$f^C_{B_s^0}$	2015 & 16	0.69 ± 0.03	0.70 ± 0.03	0.83 ± 0.05	0.40 ± 0.23
	2017	0.74 ± 0.03	0.81 ± 0.05	0.80 ± 0.04	0.39 ± 0.25
	2018	0.73 ± 0.02	0.69 ± 0.04	0.85 ± 0.04	0.42 ± 0.21
$f^C_{D_s}$	2015 & 16	0.609 ± 0.021	0.783 ± 0.022	0.956 ± 0.012	0.976 ± 0.011
	2017	0.614 ± 0.024	0.828 ± 0.023	0.958 ± 0.013	0.978 ± 0.012
	2018	0.607 ± 0.022	0.864 ± 0.022	0.953 ± 0.012	0.958 ± 0.011
$c^1_{B_s^0}$ [c^2/GeV]	2015 & 16	-10.06 ± 0.69	-9.89 ± 0.66	-6.52 ± 0.48	-8.68 ± 0.64
	2017	-9.26 ± 0.74	-7.76 ± 0.69	-6.35 ± 0.45	-8.98 ± 0.75
	2018	-11.08 ± 0.65	-8.94 ± 0.74	-5.99 ± 0.41	-9.38 ± 0.61
c_{D_s} [c^2/GeV]	2015 & 16	-7.49 ± 1.07	-6.46 ± 0.94	-4.01 ± 0.58	-5.94 ± 0.41
	2017	-6.29 ± 1.18	-8.46 ± 0.99	-5.34 ± 0.64	-6.25 ± 0.46
	2018	-7.91 ± 1.07	-7.58 ± 0.92	-5.47 ± 0.58	-5.89 ± 0.43
$\sigma^J_{B_s^0}$ [MeV/c^2]	2015 & 16	16.13 ± 0.08	15.78 ± 0.09	15.57 ± 0.12	15.97 ± 0.14
	2017	15.59 ± 0.09	15.34 ± 0.09	14.99 ± 0.12	15.54 ± 0.14
	2018	15.66 ± 0.08	15.46 ± 0.09	14.89 ± 0.12	15.36 ± 0.12
$\sigma^J_{D_s}$ [MeV/c^2]	2015 & 16	6.92 ± 0.03	6.98 ± 0.04	6.86 ± 0.05	9.76 ± 0.09
	2017	6.77 ± 0.03	6.84 ± 0.04	6.69 ± 0.05	9.55 ± 0.09
	2018	6.83 ± 0.03	6.89 ± 0.04	6.64 ± 0.05	9.37 ± 0.09
$\mu_{B_s^0}$ [MeV/c^2]	2015 & 16	5366.20 ± 0.05			
	2017	5366.30 ± 0.05			
	2018	5366.10 ± 0.05			
μ_{D_s} [MeV/c^2]	2015 & 16	1968.50 ± 0.02			
	2017	1968.50 ± 0.02			
	2018	1968.60 ± 0.02			

A.3 Decay-time fit

The parameters fixed in the decay-time fit performed in Section 7.4 are listed in Table A.4. The fitted values of the floating parameters are given in Table A.5.

Table A.4: Parameters set to constant values in the decay-time fit.

Param.	2015 & 16	2017	2018	Unit
c_7	1.0	1.0	1.0	
Δt	-2.254	-3.047	-2.394	fs
$p_0^{\delta t}$	8.377	6.092	5.5521	fs
$p_1^{\delta t}$	1.002	1.048	1.052	
$\overline{\eta}^{\mathrm{OS}}$	0.3562	0.3463	0.3464	
$\overline{\eta}^{\mathrm{SS}}$	0.4162	0.4164	0.4156	
$\varepsilon_{\mathrm{tag}}^{\mathrm{OS}}$	0.4126	0.4084	0.4123	
$\varepsilon_{\mathrm{tag}}^{\mathrm{SS}}$	0.6918	0.6992	0.6973	
$\mathit{\Gamma}_s$	0.6600	0.6600	0.6600	ps^{-1}
$\Delta\mathit{\Gamma}_s$	0.085	0.085	0.085	ps^{-1}

Table A.5: Fitted results of the parameters floating in the decay-time fit. The fitted value of the detection asymmetry is kept blinded for future analyses.

Param.	2015 & 16	2017	2018	Unit
c_1	0.325 ± 0.010	0.345 ± 0.011	0.382 ± 0.011	
c_2	0.468 ± 0.015	0.477 ± 0.017	0.533 ± 0.018	
c_3	0.779 ± 0.024	0.830 ± 0.026	0.860 ± 0.027	
c_4	0.958 ± 0.029	0.906 ± 0.029	1.024 ± 0.031	
c_5	1.107 ± 0.029	1.078 ± 0.030	1.132 ± 0.030	
c_6	1.33 ± 0.05	1.18 ± 0.05	1.31 ± 0.05	
p_0^{OS}	0.385 ± 0.004	0.376 ± 0.004	0.374 ± 0.004	
p_0^{SS}	0.4345 ± 0.0032	0.4373 ± 0.0032	0.4373 ± 0.0030	
p_1^{OS}	0.99 ± 0.04	0.88 ± 0.04	0.882 ± 0.035	
p_1^{SS}	0.747 ± 0.035	0.71 ± 0.04	0.783 ± 0.033	
Δp_0^{OS}	0.008 ± 0.004	0.004 ± 0.005	0.012 ± 0.004	
Δp_0^{SS}	-0.0163 ± 0.0033	-0.0220 ± 0.0035	-0.0123 ± 0.0033	
Δp_1^{OS}	0.00 ± 0.04	0.07 ± 0.04	0.02 ± 0.04	
Δp_1^{SS}	0.01 ± 0.04	0.06 ± 0.04	0.05 ± 0.04	
$\Delta\varepsilon_{\mathrm{tag}}^{\mathrm{OS}}$	0.007 ± 0.005	0.003 ± 0.005	-0.003 ± 0.005	
$\Delta\varepsilon_{\mathrm{tag}}^{\mathrm{SS}}$	-0.001 ± 0.005	0.002 ± 0.005	-0.007 ± 0.005	
a_{det}		[redacted] ± 0.0016		
a_{prod}		-0.0031 ± 0.0032		
Δm_s		17.7683 ± 0.0051		ps^{-1}

A.4 Bootstrapping studies

In Table 8.3 the neglected correlations among the fit observables observed in the simulated samples and the upper mass sideband are reported. In the course of the bootstrapping studies, samples are constructed to either vanish or reproduce these correlations. Table A.6 and Table A.7 respectively show the average correlations found in the two sets of bootstrapped samples. Both samples provide the yields observed in the mass fit (Table 6.2), while the correlations are vanished in Table A.6 and maintained in Table A.7.

Table A.6: Mean correlations among the fit observables in 300 uncorrelated bootstrapped samples. The average number of bootstrapped candidates is in good agreement with the yield observed in the data fit, while all problematic correlations vanish.

Component	Bootstrapped candidates	$\rho\big(m(h^{\pm}h^{+}h^{-}),\, t\big)$	$\rho\big(m(h^{\pm}h^{+}h^{-}),\, \delta t\big)$	$\rho\big(m(D_s^{-}\pi^{+}),\, t\big)$	$\rho\big(m(D_s^{-}\pi^{+}),\, \delta t\big)$
Signal					
$D_s^- \to K^{*0}K^-$	$103\,090 \pm 853$	0.000 ± 0.003	0.000 ± 0.003	0.0004 ± 0.0026	-0.0003 ± 0.0025
$D_s^- \to (K^-K^+\pi^-)_{\text{NR}}$	$64\,864 \pm 689$	0.000 ± 0.004	0.000 ± 0.004	0.000 ± 0.004	-0.001 ± 0.004
$D_s^- \to \phi\pi^-$	$141\,535 \pm 1000$	0.0000 ± 0.0026	-0.0001 ± 0.0025	0.0002 ± 0.0028	0.0001 ± 0.0026
$D_s^- \to \pi^-\pi^+\pi^-$	$65\,078 \pm 699$	0.000 ± 0.004	0.000 ± 0.004	0.000 ± 0.004	0.000 ± 0.004
Combinatorial					
$D_s^- \to K^{*0}K^-$	9456 ± 271	0.001 ± 0.009	0.001 ± 0.011	0.001 ± 0.009	0.001 ± 0.011
$D_s^- \to (K^-K^+\pi^-)_{\text{NR}}$	$19\,463 \pm 387$	0.000 ± 0.007	0.001 ± 0.007	0.000 ± 0.008	0.002 ± 0.007
$D_s^- \to \phi\pi^-$	$10\,416 \pm 284$	-0.003 ± 0.010	0.000 ± 0.010	-0.001 ± 0.009	-0.001 ± 0.009
$D_s^- \to \pi^-\pi^+\pi^-$	$32\,662 \pm 503$	0.001 ± 0.006	0.000 ± 0.005	0.000 ± 0.006	0.000 ± 0.006
$B_s^0 \to D_s^{\mp}K^{\pm}$					
$D_s^- \to (K^-K^+\pi^-)_{\text{NR}}$	496 ± 55	0.00 ± 0.05	0.00 ± 0.04	0.00 ± 0.04	0.00 ± 0.04
$B^0 \to D^{\mp}\pi^{\pm}$					
$D_s^- \to (K^-K^+\pi^-)_{\text{NR}}$	3055 ± 141	0.000 ± 0.017	0.005 ± 0.016	0.000 ± 0.018	-0.002 ± 0.019
$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$					
$D_s^- \to (K^-K^+\pi^-)_{\text{NR}}$	2196 ± 122	-0.004 ± 0.023	0.001 ± 0.021	-0.001 ± 0.021	-0.001 ± 0.024
$B_s^0 \to D_s^{*-}\pi^+$					
$D_s^- \to (K^-K^+\pi^-)_{\text{NR}}$	828 ± 75	0.01 ± 0.04	0.00 ± 0.03	0.00 ± 0.03	0.00 ± 0.03

Table A.7: The mean correlation observed among the fit observables in 300 correlated bootstrapping samples. Both, the number of candidates and the correlations present in the bootstrapped samples, are in agreement with the observed yields and with the baseline correlations listed in Table 8.3.

Component	Bootstrapped candidates	$\rho\big(m(h^\pm h^+ h^-), t\big)$	$\rho\big(m(h^\pm h^+ h^-), \delta t\big)$	$\rho\big(m(D_s^- \pi^+), t\big)$	$\rho\big(m(D_s^- \pi^+), \delta t\big)$
Signal					
$D_s^- \to K^{*0} K^-$	$103\,090 \pm 853$	-0.008 ± 0.003	0.006 ± 0.004	0.001 ± 0.003	-0.0070 ± 0.0029
$D_s^- \to (K^- K^+ \pi^-)_{\rm NR}$	$64\,864 \pm 689$	-0.007 ± 0.004	0.005 ± 0.004	-0.001 ± 0.003	-0.007 ± 0.004
$D_s^- \to \phi \pi^-$	$141\,535 \pm 1000$	-0.0034 ± 0.0026	0.0047 ± 0.0026	-0.0011 ± 0.0027	-0.0070 ± 0.0028
$D_s^- \to \pi^- \pi^+ \pi^-$	$65\,078 \pm 699$	-0.006 ± 0.004	0.005 ± 0.004	-0.001 ± 0.004	-0.008 ± 0.004
Combinatorial					
$D_s^- \to K^{*0} K^-$	9456 ± 271	0.017 ± 0.011	0.012 ± 0.011	-0.125 ± 0.010	-0.016 ± 0.009
$D_s^- \to (K^- K^+ \pi^-)_{\rm NR}$	$19\,463 \pm 387$	0.002 ± 0.006	0.000 ± 0.007	-0.092 ± 0.007	0.018 ± 0.007
$D_s^- \to \phi \pi^-$	$10\,416 \pm 284$	0.003 ± 0.012	0.014 ± 0.009	-0.076 ± 0.010	-0.041 ± 0.010
$D_s^- \to \pi^- \pi^+ \pi^-$	$32\,662 \pm 503$	0.010 ± 0.006	0.007 ± 0.006	-0.036 ± 0.005	-0.019 ± 0.005
$B_s^0 \to D_s^\mp K^\pm$					
$D_s^- \to (K^- K^+ \pi^-)_{\rm NR}$	496 ± 55	0.00 ± 0.05	0.00 ± 0.04	0.02 ± 0.05	0.10 ± 0.05
$B^0 \to D^\mp \pi^\pm$					
$D_s^- \to (K^- K^+ \pi^-)_{\rm NR}$	3055 ± 141	-0.007 ± 0.019	0.039 ± 0.019	0.035 ± 0.019	0.234 ± 0.021
$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$					
$D_s^- \to (K^- K^+ \pi^-)_{\rm NR}$	2196 ± 122	0.002 ± 0.024	0.000 ± 0.027	0.043 ± 0.021	-0.140 ± 0.026
$B_s^0 \to D_s^{*-} \pi^+$					
$D_s^- \to (K^- K^+ \pi^-)_{\rm NR}$	828 ± 75	0.02 ± 0.04	-0.04 ± 0.04	-0.08 ± 0.03	-0.03 ± 0.04

B Technical details on the analyses of $B_s^0 \to D_s^- K^+$ decays

In the following, details on the analyses of $B_s^0 \to D_s^\mp K^\pm$ decays are given, which are not essential to understand the measurement, but are helpful for experts reproducing the measurement.

B.1 Signal selection

The preselection of $B_s^0 \to D_s^\mp K^\pm$ decays is common with the preselection of $B_s^0 \to D_s^- \pi^+$ reviewed in Appendix A.1 and summarised in Table A.1. However, additional requirements at the stage of the L0 trigger are applied. The L0 is required to be either triggered by the L0hadron lines on a signal track (L0hadron TOS) or by any line independent of the signal (L0global TIS).

In contrast to the preselection, the suppression of physics backgrounds is specialised to the decay mode. The selection reviewed in Section 10.2 is summarised in Table B.1.

Table B.1: Full offline selection used to suppress backgrounds in the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal mode. The origin associated with the final-state hadrons is denoted by a subscript to prevent ambiguities. Additionally, all final-state tracks are required to have associated information from the RICH system and are expected to not appear in the muon system. Further, the BDT selection discussed in Section 6.1.3 is applied.

Decay mode	Requirement			Unit
All modes	$m(D_s^{\mp} K^{\pm}_{B_s^0})$	$\in$	[5300, 5800]	MeV/c^2
	$m(h^{\mp}_{D_s^{\mp}} h^{+}_{D_s^{\mp}} h^{-}_{D_s^{\mp}})$	$\in$	[1920, 2015]	MeV/c^2
	$t_{B_s^0}$	$\in$	[0.4, 15]	ps
	$\sigma(t_{B_s^0})$	$\in$	[0.01, 0.1]	ps
	$\text{PID}_K(K^{\pm}_{B_s^0})$	$>$	5	
	$\text{PID}_e(K^{\pm}_{B_s^0})$	$<$	5	
	$\text{PID}_e(\pi^{\mp}_{D_s^{\mp}})$	$<$	5	
$D_s^{\mp} \to K^- K^+ \pi^{\mp}$	or $\begin{cases} \text{PID}_K(K^{\mp}_{D_s^{\mp}}) \\ m(K^{-}_{D_s^{\mp}} K^{+}_{D_s^{\mp}} \pi^{\mp}_{D_s^{\mp}})_{K^{\mp} \to \pi^{\mp}} \end{cases}$	$>$ $\notin$	10 [1839, 1899]	 MeV/c^2
	or $\begin{cases} \text{PID}_K(K^{\mp}_{D_s^{\mp}}) - \text{PID}_p(K^{\mp}_{D_s^{\mp}}) \\ m(K^{-}_{D_s^{\mp}} K^{+}_{D_s^{\mp}} \pi^{\mp}_{D_s^{\mp}})_{K^{\mp} \to p} \end{cases}$	$>$ $\notin$	5 [2255, 2315]	 MeV/c^2
	$\text{PID}_p(\pi^{\mp}_{D_s^{\mp}})$	$<$	10	
	$m(K^{\pm}_{\text{all}} K^{\mp}_{D_s^{\mp}})$	$<$	1800	MeV/c^2
	$m(K^{\pm}_{B_s^0} \pi^{\mp}_{D_s^{\mp}})$	$\notin$	[1834, 1894]	MeV/c^2
	$m(K^{\pm}_{B_s^0} \pi^{\mp}_{D_s^{\mp}})$	$\notin$	[862, 922]	MeV/c^2
	$\chi^2_{\text{FD}}(D_s^{\mp})$	$>$	0	
$\to \phi \pi^{\mp}$	$\text{PID}_K(K^{\mp}_{\phi})$	$>$	-2	
	$m(K^{+}_{\phi} K^{-}_{\phi})$	$\in$	[1000, 1040]	MeV/c^2
$\to K^{*0}\ K^{\mp}$	$\text{PID}_K(K^{\mp}_{D_s^{\mp}})$	$>$	5	
	$\text{PID}_K(K^{\pm}_{K^{*0}})$	$>$	-2	
	$m(K^{\pm}_{K^{*0}} K^{\mp}_{D_s^{\mp}})$	$\notin$	[1000, 1040]	MeV/c^2
	$m(K^{\pm}_{K^{*0}} \pi^{\mp}_{D_s^{\mp}})$	$\in$	[842, 942]	MeV/c^2
$\to (K^- K^+ \pi^-)_{\text{NR}}$	$\text{PID}_K(K^{\mp}_{D_s^{\mp}})$	$>$	5	
	$\text{PID}_K(\pi^{\mp}_{D_s^{\mp}})$	$<$	10	
	$m(K^{+}_{D_s^{\mp}} K^{-}_{D_s^{\mp}})$	$\notin$	[1000, 1040]	MeV/c^2
	$m(K^{\pm}_{D_s^{\mp}} \pi^{\mp}_{D_s^{\mp}})$	$\notin$	[842, 942]	MeV/c^2

Continued on the next page

Table B.1 (continued): Full offline selection used to suppress backgrounds in the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal mode. The origin associated with the final-state hadrons is denoted by a subscript to prevent ambiguities. Additionally, all final-state tracks are required to have associated information from the RICH system and are expected to not appear in the muon system. Further, the BDT selection discussed in Section 6.1.3 is applied.

Decay mode	Requirement			Unit
$D_s^{\mp} \to K^{\mp}\pi^{+}\pi^{-}$	or $\big\{$ $\mathrm{PID}_K(K^{\mp}_{D_s^{\mp}}) - \mathrm{PID}_p(K^{\mp}_{D_s^{\mp}})$	$>$	5	
	$m(K^{\mp}_{D_s^{\mp}}\pi^{+}_{D_s^{\mp}}\pi^{-}_{D_s^{\mp}})_{K^{\mp}\to p,\, \pi^{\pm}\to K^{\pm}}$	$\notin$	[2255, 2315]	MeV/c^2
	$\mathrm{PID}_K(\pi^{\mp}_{D_s^{\mp}})$	$<$	5	
	$\mathrm{PID}_p(\pi^{\mp}_{D_s^{\mp}})$	$<$	10	
	$\mathrm{PID}_K(K^{\mp}_{D_s^{\mp}})$	$>$	10	
	$m(K^{\mp}_{D_s^{\mp}}\pi^{+}_{D_s^{\mp}}\pi^{-}_{D_s^{\mp}})_{\pi^{\pm}\to K^{\pm}}$	$\notin$	[1839, 1899]	MeV/c^2
	$m(\pi^{\pm}_{D_s^{\mp}}K^{\mp}_{D_s^{\mp}})$	$<$	1750	MeV/c^2
	$m(K^{\pm}_{B_s^0}\pi^{\mp}_{D_s^{\mp}})$	$>$	1000	MeV/c^2
	$m(K^{\pm}_{B_s^0}K^{\mp}_{D_s^{\mp}})_{K^{\pm}\to\pi^{\pm}}$	$\notin$	[1834, 1894]	MeV/c^2
	$m(\pi^{\mp}_{D_s^{\mp}}K^{\pm}_{B_s^0}\pi^{\mp}_{D_s^{\mp}})$	$\notin$	[1839, 1899]	MeV/c^2
	$\chi^2_{\mathrm{FD}}(D_s^{\mp})$	$>$	9	
$D_s^{\mp} \to \pi^{\mp}\pi^{+}\pi^{-}$	$\mathrm{PID}_K(\pi^{\mp}_{D_s^{\mp}})$	$<$	2	
	$\mathrm{PID}_p(\pi^{\mp}_{D_s^{\mp}})$	$<$	5	
	$m(\pi^{+}_{D_s^{\mp}}\pi^{-}_{D_s^{\mp}})$	$<$	1700	MeV/c^2
	$m(K^{\pm}_{B_s^0}\pi^{\mp}_{D_s^{\mp}})$	$>$	1000	MeV/c^2
	$m(K^{\pm}_{B_s^0}\pi^{\mp}_{D_s^{\mp}})$	$\notin$	[1834, 1894]	MeV/c^2
	$m(\pi^{\mp}_{D_s^{\mp}}K^{\pm}_{B_s^0}\pi^{\mp}_{D_s^{\mp}})$	$\notin$	[1839, 1899]	MeV/c^2
	$\chi^2_{\mathrm{FD}}(D_s^{\mp})$	$>$	9	

B.2 Mass fit

In Table B.2, the floated shape parameters of mass models introduced in Section 10.3 are listed for the $B_s^0 \to D_s^{\mp} K^{\pm}$ signal and the combinatorial background component.

Table B.2: Fitted values of the shape parameters floating in the multi-dimensional mass fit. The mean μ of the signal shapes is shared among the five D_s^- final states in both mass spectra.

Parameter	Year	$D_s^- \to \phi\pi^-$	$D_s^- \to K^{*0}K^-$	$D_s^- \to (K^-K^+\pi^-)_{\mathrm{NR}}$	$D_s^- \to K^-\pi^+\pi^-$	$D_s^- \to \pi^-\pi^+\pi^-$
$f_{B_s^0}$	2015 & 16	0.54 ± 0.08	0.72 ± 0.06	0.56 ± 0.09	-	-
	2017	0.37 ± 0.06	0.41 ± 0.08	0.58 ± 0.11	-	-
	2018	0.30 ± 0.05	0.64 ± 0.08	0.52 ± 0.06	-	-
f_{D_s}	2015 & 16	0.50 ± 0.04	0.75 ± 0.05	0.93 ± 0.03	0.96 ± 0.02	0.92 ± 0.03
	2017	0.53 ± 0.05	0.77 ± 0.06	0.89 ± 0.03	0.94 ± 0.03	0.93 ± 0.03
	2018	0.60 ± 0.05	0.68 ± 0.05	0.99 ± 0.03	0.95 ± 0.03	0.95 ± 0.03
$c_{B_s^0}$ [c^2/GeV]	2015 & 16	-9 ± 2	-9.0 ± 1.2	-6.7 ± 1.4	-0.7 ± 0.2	-0.9 ± 0.2
	2017	-22 ± 7	-17 ± 6	-6.2 ± 1.5	-1.4 ± 0.2	-1.1 ± 0.2
	2018	-26 ± 7	-8.4 ± 1.5	-9.1 ± 1.5	-1.4 ± 0.2	-1.0 ± 0.2
c_{D_s} [c^2/GeV]	2015 & 16	-5.2 ± 2.0	-10.3 ± 1.7	-5.5 ± 1.1	-1.4 ± 1.0	-2.0 ± 1.0
	2017	-9.1 ± 2.5	-9.5 ± 2.0	-7.6 ± 1.3	-1.2 ± 1.2	-4.2 ± 1.0
	2018	-12.9 ± 2.4	-8.4 ± 1.8	-4.3 ± 1.1	-1.2 ± 1.1	-2.3 ± 1.0
$\sigma_{B_s^0}^H$ [MeV/c^2]	2015 & 16	14.4 ± 0.5	14.2 ± 0.5	16.4 ± 0.8	16.5 ± 1.6	16.5 ± 1.0
	2017	14.6 ± 0.5	14.6 ± 0.5	14.9 ± 0.7	14.3 ± 1.2	14.9 ± 1.0
	2018	15.0 ± 0.4	14.3 ± 0.5	14.9 ± 0.7	15.4 ± 1.2	15.2 ± 0.8
$\sigma_{D_s}^H$ [MeV/c^2]	2015 & 16	6.85 ± 0.11	6.79 ± 0.13	6.07 ± 0.17	8.4 ± 0.5	10.4 ± 0.4
	2017	6.63 ± 0.12	6.86 ± 0.14	6.51 ± 0.18	7.9 ± 0.4	9.0 ± 0.4
	2018	6.82 ± 0.11	7.00 ± 0.13	6.08 ± 0.17	8.2 ± 0.4	10.0 ± 0.3
$\mu_{B_s^0}$ [MeV/c^2]	2015 & 16			5366.82 ± 0.22		
	2017			5366.75 ± 0.23		
	2018			5366.08 ± 0.21		
μ_{D_s} [MeV/c^2]	2015 & 16			1968.62 ± 0.07		
	2017			1968.48 ± 0.07		
	2018			1968.70 ± 0.07		

B.3 Flavour-tagging portability

The calibration parameters extracted from $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ simulation using the lhcb-ftcalib backage [87] are given in Table B.3. The calibrations of both modes are in good agreement.

Table B.3: Tagging parameters extracted by a calibration based on the known initial flavour on simulated $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$ samples.

Parameter	2015 & 16	2017	2018
		$B_s^0 \to D_s^- \pi^+$	
$\overline{\eta}^{\mathrm{OS}}$	0.349 13 ± 0.000 06	0.349 40 ± 0.000 07	0.347 70 ± 0.000 06
p_0^{OS}	0.366 54 ± 0.000 31	0.367 99 ± 0.000 33	0.366 34 ± 0.000 30
p_1^{OS}	−0.1110 ± 0.0027	−0.1141 ± 0.0029	−0.1129 ± 0.0026
Δp_0^{OS}	0.0039 ± 0.0006	0.0042 ± 0.0007	0.0055 ± 0.0006
Δp_1^{OS}	0.011 ± 0.005	0.019 ± 0.006	0.000 ± 0.005
$\overline{\eta}^{\mathrm{SS}}$	0.406 83 ± 0.000 05	0.407 66 ± 0.000 05	0.405 56 ± 0.000 04
p_0^{SS}	0.415 91 ± 0.000 24	0.416 85 ± 0.000 26	0.415 50 ± 0.000 24
p_1^{SS}	−0.0918 ± 0.0025	−0.0969 ± 0.0027	−0.0922 ± 0.0024
Δp_0^{SS}	−0.0133 ± 0.0005	−0.0128 ± 0.0005	−0.0136 ± 0.0005
Δp_1^{SS}	0.014 ± 0.005	0.007 ± 0.005	0.009 ± 0.005
		$B_s^0 \to D_s^{\mp} K^{\pm}$	
$\overline{\eta}^{\mathrm{OS}}$	0.348 18 ± 0.000 09	0.348 57 ± 0.000 09	0.346 56 ± 0.000 09
p_0^{OS}	0.3655 ± 0.0004	0.3659 ± 0.0004	0.3647 ± 0.0004
p_1^{OS}	−0.109 ± 0.004	−0.109 ± 0.004	−0.104 ± 0.004
Δp_0^{OS}	0.0035 ± 0.0009	0.0057 ± 0.0008	0.0060 ± 0.0009
Δp_1^{OS}	0.020 ± 0.008	0.013 ± 0.007	0.019 ± 0.007
$\overline{\eta}^{\mathrm{SS}}$	0.412 25 ± 0.000 06	0.413 00 ± 0.000 06	0.411 29 ± 0.000 06
p_0^{SS}	0.4189 ± 0.0004	0.419 76 ± 0.000 34	0.418 71 ± 0.000 34
p_1^{SS}	−0.111 ± 0.004	−0.108 ± 0.004	−0.102 ± 0.004
Δp_0^{SS}	−0.0141 ± 0.0007	−0.0138 ± 0.0007	−0.0140 ± 0.0007
Δp_1^{SS}	0.020 ± 0.008	0.023 ± 0.007	0.021 ± 0.007

B.4 Decay-time fit

The parameters fixed in the decay-time fit performed in Section 11.5 are listed in Appendix B.4. The fitted values of the Gaussian-constrained parameters are given in Table B.5.

Table B.4: Parameters fixed to constant values in the decay-time fit of $B_s^0 \to D_s^{\mp} K^{\pm}$ data. The origin of the values is explained in the text.

Parameter	2015 & 16	2017	2018	Unit
c_1	0.291	0.316	0.349	
c_2	0.327	0.333	0.387	
c_3	0.526	0.545	0.592	
c_4	0.804	0.839	0.881	
c_5	1.002	0.942	1.046	
c_6	1.145	1.088	1.171	
c_7	1.225	1.146	1.193	
c_8	1.224	1.150	1.128	
c_9	1.203	1.107	1.076	
c_{10}	1.000	1.000	1.000	
Δt	−3.0	−3.8	−3.1	fs
$p_0^{\delta t}$	0.0125	0.0099	0.0098	fs
$p_1^{\delta t}$	0.950	0.955	0.948	
$\overline{\eta}^{\mathrm{OS}}$	0.3562	0.3463	0.3464	
$\overline{\eta}^{\mathrm{SS}}$	0.4162	0.4164	0.4156	
$\varepsilon_{\mathrm{tag}}^{\mathrm{OS}}$	0.4126	0.4084	0.4123	
$\varepsilon_{\mathrm{tag}}^{\mathrm{SS}}$	0.6918	0.6992	0.6973	
$\Delta\varepsilon_{\mathrm{tag}}^{\mathrm{OS}}$	0.009	0.003	−0.002	
$\Delta\varepsilon_{\mathrm{tag}}^{\mathrm{SS}}$	0.002	0.002	−0.006	
a_{det}		0.0096		
a_{prod}		−0.0033		
Γ_s		0.6563		ps^{-1}
$\Delta\Gamma_s$		0.085		ps^{-1}
Δm_s		17.7683		ps^{-1}

Table B.5: Fitted values of the Gaussian constrained FT calibration parameters. The nominal parameters can be found in Table A.5 for comparison.

Parameter	2015 & 16	2017	2018
p_0^{OS}	0.3812 ± 0.0010	0.3763 ± 0.0012	0.3732 ± 0.0011
p_0^{SS}	0.4339 ± 0.0008	0.4382 ± 0.0009	0.4369 ± 0.0008
p_1^{OS}	1.0142 ± 0.0104	0.8910 ± 0.0108	0.8699 ± 0.0109
p_1^{SS}	0.7624 ± 0.0099	0.7291 ± 0.0107	0.7900 ± 0.0106
Δp_0^{OS}	0.0086 ± 0.0011	0.0027 ± 0.0011	0.0117 ± 0.0010
Δp_0^{SS}	-0.0151 ± 0.0008	-0.0217 ± 0.0008	-0.0115 ± 0.0007
Δp_1^{OS}	0.0190 ± 0.0108	0.0725 ± 0.0093	-0.0017 ± 0.0090
Δp_1^{SS}	0.0189 ± 0.0091	0.0471 ± 0.0083	0.0532 ± 0.0078

B.5 Data splits

The deviations of the *CP* parameters fitted on the different subsamples as discussed in Section 12.2 with respect to the nominal result are listed in Table B.6. The corresponding Mahalanobis distances and the p-values are given in Table 12.5.

Table B.6: Deviations of the *CP* parameters extracted on subsample splits with respect to the nominal result.

Subsample	C	D_f	$D_{\bar{f}}$	S_f	$S_{\bar{f}}$
Magnet polarity split:					
Up	0.075 ± 0.056	-0.157 ± 0.120	-0.088 ± 0.116	-0.027 ± 0.073	-0.096 ± 0.080
Down	-0.085 ± 0.066	0.150 ± 0.143	0.073 ± 0.130	0.022 ± 0.096	0.104 ± 0.088
Data-taking period split:					
2015 & 16	0.149 ± 0.081	0.02 ± 0.14	0.19 ± 0.14	0.12 ± 0.10	0.06 ± 0.11
2017	-0.066 ± 0.072	0.10 ± 0.16	-0.10 ± 0.16	-0.01 ± 0.10	-0.07 ± 0.11
2018	-0.061 ± 0.067	-0.10 ± 0.15	-0.09 ± 0.13	-0.09 ± 0.10	0.010 ± 0.090
D_s^--decay mode split:					
$(K^- K^+ \pi^-)_{\mathrm{NR}}$	-0.10 ± 0.11	-0.30 ± 0.28	0.03 ± 0.23	-0.38 ± 0.17	-0.11 ± 0.18
$K^{*0} K^-$	0.040 ± 0.091	0.16 ± 0.17	-0.28 ± 0.19	-0.09 ± 0.17	-0.10 ± 0.15
$\phi \pi^-$	0.075 ± 0.066	0.22 ± 0.13	0.19 ± 0.12	0.188 ± 0.082	0.232 ± 0.078
$K^- \pi^+ \pi^-$	0.16 ± 0.24	-0.55 ± 0.57	-0.30 ± 0.51	0.20 ± 0.40	-0.04 ± 0.33
$\pi^- \pi^+ \pi^-$	-0.21 ± 0.16	-0.41 ± 0.32	0.02 ± 0.29	0.07 ± 0.27	-0.01 ± 0.18
B_s^0 momentum split:					
$p_{B_s^0} > 120$ GeV/c	0.020 ± 0.073	0.06 ± 0.17	0.02 ± 0.16	-0.006 ± 0.086	0.043 ± 0.084
$p_{B_s^0} < 120$ GeV/c	0.011 ± 0.042	-0.031 ± 0.059	-0.002 ± 0.056	-0.012 ± 0.054	-0.030 ± 0.063
Flavour tagging split:					
SS	0.145 ± 0.104	0.017 ± 0.146	-0.160 ± 0.143	0.154 ± 0.140	0.097 ± 0.151
OS	-0.098 ± 0.117	-0.212 ± 0.368	-0.198 ± 0.345	0.071 ± 0.195	-0.132 ± 0.172
Both	-0.039 ± 0.050	-0.062 ± 0.202	0.119 ± 0.186	-0.070 ± 0.064	0.011 ± 0.064

B.6 Closure test

In the simulation-based closure test discussed in Section 12.3, the acceptance corrections need to be reevaluated on the simulation-like samples. In Table B.7, the acceptance ratios for the nominal closure test are given. In Table B.8, these ratios are reported applying different PID requirements to the $B_s^0 \to D_s^\mp K^\pm$ samples. Effects are introduced at low decay times. Lastly, the ratios are given in Table B.9 assuming an artificial distortion of the simulation-like samples. The differences with respect to Table B.7 are small as expected since both modes are distorted similarly.

Table B.7: Acceptance ratios extracted from the simulation-like samples for the closure test.

Parameter	2016	2017	2018
v_1	0.875± 0.009	0.904± 0.010	0.902± 0.010
v_2	0.895± 0.011	0.920± 0.011	0.925± 0.012
v_3	0.932± 0.011	0.965± 0.011	0.949± 0.012
v_4	0.948± 0.011	0.978± 0.011	0.967± 0.012
v_5	0.952± 0.009	0.979± 0.010	0.973± 0.010
v_6	0.973± 0.015	1.011± 0.015	0.982± 0.016

Table B.8: Acceptance ratios extracted from the simulation-like samples for the closure test using a tighter $\text{PID}_K > 10$ or a looser $\text{PID}_K > 0$ requirement on the accompanying kaon.

Parameter	2016	2017	2018
	Tight $\text{PID}_K > 10$ requirement		
v_1	0.835± 0.009	0.864± 0.010	0.864± 0.010
v_2	0.865± 0.011	0.887± 0.011	0.895± 0.012
v_3	0.914± 0.011	0.950± 0.011	0.936± 0.012
v_4	0.933± 0.011	0.965± 0.011	0.960± 0.012
v_5	0.943± 0.010	0.972± 0.010	0.967± 0.010
v_6	0.965± 0.015	1.008± 0.016	0.984± 0.016
	Lose $\text{PID}_K > 0$ requirement		
v_1	0.961± 0.010	0.992± 0.010	0.989± 0.011
v_2	0.963± 0.011	0.987± 0.011	0.982± 0.012
v_3	0.964± 0.011	0.997± 0.011	0.981± 0.011
v_4	0.976± 0.011	1.007± 0.011	0.990± 0.011
v_5	0.974± 0.009	1.001± 0.009	0.992± 0.010
v_6	0.984± 0.014	1.022± 0.015	0.985± 0.015

Table B.9: Acceptance ratios for the closure test extracted from the artificially distorted simulation-like samples.

Parameter	2016	2017	2018
v_1	$0.869\pm$ 0.009	$0.898\pm$ 0.010	$0.894\pm$ 0.010
v_2	$0.887\pm$ 0.011	$0.911\pm$ 0.011	$0.917\pm$ 0.012
v_3	$0.927\pm$ 0.011	$0.959\pm$ 0.011	$0.941\pm$ 0.012
v_4	$0.945\pm$ 0.011	$0.973\pm$ 0.011	$0.966\pm$ 0.012
v_5	$0.951\pm$ 0.009	$0.979\pm$ 0.010	$0.973\pm$ 0.010
v_6	$0.974\pm$ 0.015	$1.010\pm$ 0.015	$0.984\pm$ 0.016

List of Figures

List of Tables

Bibliography

[1] LHCb Collaboration. "Precise determination of the B_s^0-$\overline{B}_s^0$ oscillation frequency." In: *Nature Physics* 18.1 (Jan. 2022), pp. 1–5. DOI: `10.1038/s41567-021-01394-x`.

[2] LHCb Collaboration. *Measurement of CP asymmetry in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays.* Tech. rep. 2023. URL: `http://cds.cern.ch/record/2873713`.

[3] S. Weinberg. "A Model of Leptons." In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: `10.1103/PhysRevLett.19.1264`.

[4] S. L. Glashow. "Partial-symmetries of weak interactions." In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: `https://doi.org/10.1016/0029-5582(61)90469-2`.

[5] A. Salam. "Weak and Electromagnetic Interactions." In: *Conf. Proc. C* 680519 (1968), pp. 367–377. DOI: `10.1142/9789812795915_0034`.

[6] V. Trimble. "Existence and Nature of Dark Matter in the Universe." In: *Annual Review of Astronomy and Astrophysics* 25.1 (1987), pp. 425–472. DOI: `10.1146/annurev.aa.25.090187.002233`.

[7] Y. Fukuda, T. Hayakawa, E. Ichihara, et al. "Evidence for Oscillation of Atmospheric Neutrinos." In: *Phys. Rev. Lett.* 81 (8 Aug. 1998), pp. 1562–1567. DOI: `10.1103/PhysRevLett.81.1562`.

[8] A. J. Buras and M. Lindner, eds. *Heavy flavours II.* Wsp, 1998. DOI: `10.1142/2706`.

[9] I. I. Bigi and A. I. Sanda. *CP violation.* Vol. 9. Cambridge University Press, Sept. 2009. ISBN: 978-0-511-57728-4, 978-0-521-84794-0, 978-1-107-42430-2. DOI: `10.1017/cbo9780511581014`.

[10] G. C. Branco, L. Lavoura, and J. P. Silva. *CP Violation.* Vol. 103. 1999.

[11] S. Hansmann-Menzemer and U. Nierste. "Quark-Flavour Physics." In: *The Large Hadron Collider: Harvest of Run 1.* Ed. by T. Schörner-Sadenius. Springer International Publishing, 2015, pp. 301–353. ISBN: 978-3-319-15001-7. DOI: `10.1007/978-3-319-15001-7_8`.

[12] LHCb Collaboration. "The LHCb Detector at the LHC." In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08005. DOI: `10.1088/1748-0221/3/08/s08005`.

[13] L. Evans and P. Bryant. "LHC Machine." In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001. DOI: `10.1088/1748-0221/3/08/s08001`.

[14] LHCb Collaboration. "Measurement of CP asymmetry in $B_s^0 \to D_s^{\mp} K^{\pm}$ decays." In: *Jhep* 11 (2014), p. 060. DOI: `10.1007/jhep11(2014)060`.

[15] LHCb Collaboration. *Measurement of lepton universality parameters in* $B^+ \to K^+\ell^+\ell^-$ *and* $B^0 \to K^{*0}\ell^+\ell^-$ *decays.* Aug. 2023. DOI: `10.1103/PhysRevD.108.032002`.

[16] LHCb Collaboration. *Test of Lepton Universality in* $b \to s\ell^+\ell^-$ *Decays.* Aug. 2023. DOI: `10.1103/PhysRevLett.131.051803`.

[17] ATLAS Collaboration et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC." In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/j.physletb.2012.08.020`.

[18] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC." In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/j.physletb.2012.08.021`.

[19] D. Griffiths. *Introduction to elementary particles.* 2008. ISBN: 978-3-527-40601-2.

[20] D. H. Perkins. *Introduction to high energy physics.* 1982. ISBN: 978-0-521-62196-0.

[21] LHCb Collaboration. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in ${}^0_b \to J/\psi K^- p$ Decays." In: *Phys. Rev. Lett.* 115 (7 Aug. 2015), p. 072001. DOI: `10.1103/PhysRevLett.115.072001`.

[22] Belle Collaboration. "Study of $e^+e^- \to \pi^+\pi^- J/\psi$ and Observation of a Charged Charmoniumlike State at Belle." In: *Phys. Rev. Lett.* 110 (25 June 2013), p. 252002. DOI: `10.1103/PhysRevLett.110.252002`.

[23] BESIII Collaboration. "Observation of a Charged Charmoniumlike Structure in $e^+e^- \to \pi^+\pi^- J/\psi$ at $\sqrt{s}$=4.26 GeV." In: *Phys. Rev. Lett.* 110 (25 June 2013), p. 252001. DOI: `10.1103/PhysRevLett.110.252001`.

[24] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons." In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: `10.1103/PhysRevLett.13.508`.

[25] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons." In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: `10.1103/PhysRevLett.13.321`.

[26] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles." In: *Phys. Rev. Lett.* 13 (20 Nov. 1964), pp. 585–587. DOI: `10.1103/PhysRevLett.13.585`.

[27] H. Yukawa. "On the Interaction of Elementary Particles I." In: *Proc. Phys. Math. Soc. Jap.* 17 (1935), pp. 48–57. DOI: `10.1143/ptps.1.1`.

[28] E. Noether. "Invariante Variationsprobleme." In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257. URL: `http://eudml.org/doc/59024`.

[29] C. S. Wu, E. Ambler, R. W. Hayward, et al. "Experimental Test of Parity Conservation in Beta Decay." In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: `10.1103/PhysRev.105.1413`.

[30] J. H. Christenson, J. W. Cronin, V. L. Fitch, et al. “Evidence for the 2π Decay of the K_2^0 Meson.” In: *Phys. Rev. Lett.* 13 (4 July 1964), pp. 138–140. DOI: `10.1103/PhysRevLett.13.138`.

[31] G. Lüders. “Proof of the TCP theorem.” In: *Annals of Physics* 2.1 (1957), pp. 1–15. ISSN: 0003-4916. DOI: `https://doi.org/10.1016/0003-4916(57)90032-5`.

[32] M. Kobayashi and T. Maskawa. “*CP*-Violation in the Renormalizable Theory of Weak Interaction.” In: *Progress of Theoretical Physics* 49.2 (1973), pp. 652–657. DOI: `10.1143/ptp.49.652`.

[33] L.-L. Chau and W.-Y. Keung. “Comments on the Parametrization of the Kobayashi-Maskawa Matrix.” In: *Phys. Rev. Lett.* 53 (19 Nov. 1984), pp. 1802–1805. DOI: `10.1103/PhysRevLett.53.1802`.

[34] L. Wolfenstein. “Parametrization of the Kobayashi-Maskawa Matrix.” In: *Phys. Rev. Lett.* 51 (21 Nov. 1983), pp. 1945–1947. DOI: `10.1103/PhysRevLett.51.1945`.

[35] C. Patrignani. “Review of Particle Physics.” In: *Chinese Physics C* 40.10 (Oct. 2016), p. 100001. DOI: `10.1088/1674-1137/40/10/100001`.

[36] CKM Fitter Group. *CKM Fitter Group.* 2005. URL: `http://ckmfitter.in2p3.fr/`.

[37] V. Weisskopf and E. Wigner. “Berechnung der natürlichen Linienbreite auf Grund der Diracschen Lichttheorie.” In: *Zeitschrift fur Physik* 63.1-2 (Jan. 1930), pp. 54–73. DOI: `10.1007/bf01336768`.

[38] LHCb Collaboration. “Direct *CP* violation in charmless three-body decays of $B^{\pm}$ mesons.” In: *Phys. Rev. D* 108 (1 July 2023), p. 012008. DOI: `10.1103/PhysRevD.108.012008`.

[39] LHCb Collaboration. “Search for direct *CP* violation in charged charmless $B \to PV$ decays.” In: *Phys. Rev. D* 108 (1 July 2023), p. 012013. DOI: `10.1103/PhysRevD.108.012013`.

[40] M. Suzuki. “Inelastic final-state interaction.” In: *Physical Review D* 77.5 (Mar. 2008). DOI: `10.1103/physrevd.77.054021`.

[41] A. Lenz and G. Tetlalmatzi-Xolocotzi. “Model-independent bounds on new physics effects in non-leptonic tree-level decays of *B*-mesons.” In: *Journal of High Energy Physics* 2020.7 (July 2020). DOI: `10.1007/jhep07(2020)177`.

[42] HFLAV Group. *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021.* Mar. 2023. DOI: `10.1103/PhysRevD.107.052008`.

[43] T. Gershon, A. Lenz, A. V. Rusov, et al. “Testing the Standard Model with *CP* asymmetries in flavor-specific nonleptonic decays.” In: *Phys. Rev. D* 105.11 (2022), p. 115023. DOI: `10.1103/PhysRevD.105.115023`.

[44] T. Inami and C. S. Lim. “Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L^0 \to \nu\nu$, $K^+ \to \pi^+\nu\nu$ and $K^0 \leftrightarrow K^0$.” In: *Progress of Theoretical Physics* 65.1 (Jan. 1981), pp. 297–314. ISSN: 0033-068x. DOI: `10.1143/ptp.65.297`.

[45] S. L. Glashow, J. Iliopoulos, and L. Maiani. “Weak Interactions with Lepton-Hadron Symmetry.” In: *Phys. Rev. D* 2 (7 Oct. 1970), pp. 1285–1292. DOI: `10.1103/PhysRevD.2.1285`.

[46] L. D. Luzio, M. Kirk, A. Lenz, et al. " Ms theory precision confronts flavour anomalies." In: *Journal of High Energy Physics* 2019.12 (Dec. 2019). DOI: `10.1007/jhep12(2019)009`.

[47] L. Di Luzio, M. Kirk, and A. Lenz. "Updated B_s-mixing constraints on new physics models for $b \to s\ell^+\ell^-$ anomalies." In: *Phys. Rev. D* 97 (9 May 2018), p. 095035. DOI: `10.1103/PhysRevD.97.095035`.

[48] K. De Bruyn, R. Fleischer, E. Malami, et al. "Studies of New Physics in $B_q^0 - \bar{B}_q^0$ Mixing and Implications for Leptonic Decays." In: *8th Symposium on Prospects in the Physics of Discrete Symmetries*. Jan. 2023. eprint: `2301.13649`.

[49] E. Malami. "*CP* Violation and mixing in beauty sector: A theoretical overview." In: *PoS* Lhcp2022 (2023), p. 056. DOI: `10.22323/1.422.0056`.

[50] D. King, M. Kirk, A. Lenz, et al. "$|V_{cb}|$ and γ from B-mixing - Addendum to "B_s mixing observables and $|V_{td}/V_{ts}|$ from sum rules"." In: (Nov. 2019). [Addendum: JHEP 03, 112 (2020)]. DOI: `10.1007/jhep03(2020)112`.

[51] M. Gronau and D. Wyler. "On determining a weak phase from charged B decay asymmetries." In: *Physics Letters B* 265.1 (1991), pp. 172–176. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/0370-2693(91)90034-N`.

[52] 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." In: *Physical Review D* 50.5 (Sept. 1994), pp. 3433–3446. DOI: `10.1103/physrevd.50.3433`.

[53] D. Atwood, I. Dunietz, and A. Soni. "Enhanced *CP* Violation with $B \to KD$ Modes and Extraction of the CKM Angle gamma." In: *Physical Review Letters* 78.17 (Apr. 1997), pp. 3257–3260. DOI: `10.1103/physrevlett.78.3257`.

[54] D. Atwood, I. Dunietz, and A. Soni. "Improved methods for observing *CP* violation in $B^\pm \to$ KD and measuring the CKM phase γ." In: *Physical Review D* 63.3 (Jan. 2001). DOI: `10.1103/physrevd.63.036005`.

[55] Y. Grossman, Z. Ligeti, and A. Soffer. "Measuring γ in $B^\pm \to K^\pm(KK^*)_D$ decays." In: *Physical Review D* 67.7 (Apr. 2003). DOI: `10.1103/physrevd.67.071301`.

[56] LHCb Collaboration. "Measurement of *CP* violation in $B^0 \to D^\mp\pi^\pm$ decays." In: *Journal of High Energy Physics* 2018.6 (June 2018). DOI: `10.1007/jhep06(2018)084`.

[57] LHCb Collaboration. "Measurement of the CKM angle γ and B_s^0-$\bar{B}_s^0$ mixing frequency with $B_s^0 \to D_s^\mp K^\pm\pi^+\pi^-$ decays." In: *Journal of High Energy Physics* 2021.3 (Mar. 2021). DOI: `10.1007/jhep03(2021)137`.

[58] E. Malami and R. Fleischer. "Puzzles in the $B_s^0 \to D_s^\pm K^\mp$ System." In: *PoS* Ckm2021 (2023), p. 077. DOI: `10.22323/1.411.0077`.

[59] E. Mobs. *The CERN accelerator complex in 2019.* General Photo. 2019. URL: `https://cds.cern.ch/record/2684277` (visited on 02/2023).

[60] ALICE Collaboration. "The ALICE experiment at the CERN LHC." In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08002. DOI: `10.1088/1748-0221/3/08/s08002`.

[61] ATLAS Collaboration et al. “The ATLAS Experiment at the CERN Large Hadron Collider.” In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003. DOI: `10.1088/1748-0221/3/08/s08003`.

[62] CMS Collaboration. “The CMS experiment at the CERN LHC.” In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004. DOI: `10.1088/1748-0221/3/08/s08004`.

[63] R. Alemany-Fernandez, F. Follin, and R. Jacobsson. “The LHCb Online Luminosity Control and Monitoring.” In: *4th International Particle Accelerator Conference*. 2013.

[64] C. Elsässer. *$\bar{b}b$ production angle plots*. URL: `https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb%5C_ProductionAngles.html` (visited on 02/2023).

[65] LHCb Collaboration. “LHCb Detector Performance.” In: *Int. J. Mod. Phys. A* 30.07 (2015), p. 1530022. DOI: `10.1142/s0217751x15300227`.

[66] LHCb Collaboration. “Performance of the LHCb Vertex Locator.” In: *Journal of Instrumentation* 9.09 (Sept. 2014), P09007. DOI: `10.1088/1748-0221/9/09/p09007`.

[67] C. Ilgner, M. D. M. Lieng, M. Nedos, et al. *The Beam Conditions Monitor of the LHCb Experiment*. 2010. eprint: `1001.2487`.

[68] LHCb Collaboration. *LHCb magnet: Technical Design Report*. Technical design report. LHCb. Cern, 2000. URL: `https://cds.cern.ch/record/424338`.

[69] C. Elsasser, B. Adeva, A. Gallas Torreira, et al. *The LHCb Silicon Tracker*. Tech. rep. On behalf of the LHCb Silicon Tracker Group. Cern, 2013. URL: `https://cds.cern.ch/record/1606842`.

[70] LHCb OuterTracker Group. “Performance of the LHCb Outer Tracker.” In: *Jinst* 9 (2014). Comments: 30 pages, 20 figures, P01002. DOI: `10.1088/1748-0221/9/01/p01002`.

[71] LHCb RICH Group. “Performance of the LHCb RICH detector at the LHC.” In: *Eur. Phys. J. C* 73 (2013), p. 2431. DOI: `10.1140/epjc/s10052-013-2431-9`.

[72] LHCb Collaboration. *LHCb calorimeters: Technical Design Report*. Technical design report. LHCb. Cern, 2000. URL: `https://cds.cern.ch/record/494264`.

[73] P. Perret and X. Vilasis-Cardona. “Performance of the LHCb calorimeters during the period 2010-2012.” In: *J. Phys.: Conf. Ser.* 587.1 (2015), p. 012012. DOI: `10.1088/1742-6596/587/1/012012`.

[74] P. R. Barbosa-Marinho, I. Bediaga, G. Cernicchiaro, et al. *LHCb muon system: Technical Design Report*. Technical design report. LHCb. Cern, 2001. URL: `https://cds.cern.ch/record/504326`.

[75] F. Archilli, W. Baldini, G. Bencivenni, et al. “Performance of the Muon Identification at LHCb.” In: *Journal of Instrumentation* 8.10 (Oct. 2013), P10020–p10020. DOI: `10.1088/1748-0221/8/10/p10020`.

[76] LHCb Collaboration. “Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC.” In: *Journal of Instrumentation* 14.04 (Apr. 2019), P04013. DOI: `10.1088/1748-0221/14/04/p04013`.

[77] I. Belyaev, T. Brambach, N. H. Brook, et al. "Handling of the generation of primary events in Gauss, the LHCb simulation framework." In: *Journal of Physics: Conference Series* 331.3 (Dec. 2011), p. 032047. DOI: `10.1088/1742-6596/331/3/032047`.

[78] T. Sjöstrand, S. Ask, J. R. Christiansen, et al. "An introduction to PYTHIA 8.2." In: *Computer Physics Communications* 191 (2015), pp. 159–177. ISSN: 0010-4655. DOI: `https://doi.org/10.1016/j.cpc.2015.01.024`.

[79] D. J. Lange. "The EvtGen particle decay simulation package." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001), pp. 152–155. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/S0168-9002(01)00089-4`.

[80] P. Golonka and Z. Was. "PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays." In: *The European Physical Journal C* 45.1 (Jan. 2006), pp. 97–107. DOI: `10.1140/epjc/s2005-02396-4`.

[81] D. J. Lange. "The EvtGen particle decay simulation package." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001), pp. 152–155. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/S0168-9002(01)00089-4`.

[82] G. Corti, M. Cattaneo, P. Charpentier, et al. "Software for the LHCb experiment." In: *IEEE Symposium Conference Record Nuclear Science 2004*. Vol. 4. 2004, 2048–2052 Vol. 4. DOI: `10.1109/nssmic.2004.1462666`.

[83] L. Anderlini, A. Contu, C. R. Jones, et al. *The PIDCalib package*. Tech. rep. Cern, 2016. URL: `https://cds.cern.ch/record/2202412`.

[84] J. T. Wishahi. "Measurement of *CP* violation in $B^0 \to J/\Psi K^0_S$ decays with the LHCb experiment." Presented 20 November 2013. PhD thesis. TU Dortmund, Nov. 2013.

[85] LHCb Collaboration. "Comparison of Flavour Tagging performances displayed in the ω-ε-plane." In: (2020). URL: `https://cds.cern.ch/record/2707121`.

[86] J. Wimberley. *Espresso Performance Monitor Documentation*. URL: `https://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor/wikis/home/` (visited on 02/2023).

[87] V. Jevtic and Q. Führing. *lhcb-ftcalib: Library for calibrating flavour tagging algorithms at LHCb*. URL: `https://pypi.org/project/lhcb-ftcalib/` (visited on 02/2023).

[88] R. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences (Manchester Physics Series)*. WileyBlackwell, 1989. ISBN: 978-0-471-92295-7.

[89] R. Barlow. "Extended maximum likelihood." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 297.3 (1990), pp. 496–506. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/0168-9002(90)91334-8`.

[90] M. Pivk and F. L. Diberder. ": A statistical tool to unfold data distributions." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 555.1-2 (Dec. 2005), pp. 356–369. DOI: `10.1016/j.nima.2005.08.106`.

[91] C. Langenbruch. "Parameter uncertainties in weighted unbinned maximum likelihood fits." In: *The European Physical Journal C* 82.5 (May 2022). DOI: `10.1140/epjc/s10052-022-10254-8`.

[92] Y. Coadou. "Boosted Decision Trees." In: *Artificial Intelligence for High Energy Physics*. World Scientific, Feb. 2022, pp. 9–58. DOI: `10.1142/9789811234033_0002`.

[93] 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*. Kdd '16. San Francisco, California, USA: Association for Computing Machinery, 2016, pp. 785–794. ISBN: 9781450342322. DOI: `10.1145/2939672.2939785`.

[94] A. Hocker, P. Speckmayer, J. Stelzer, et al. *TMVA - Toolkit for Multivariate Data Analysis*. Tech. rep. Cern, 2007. URL: `https://cds.cern.ch/record/1019880`.

[95] S. Jamalaldin, S. J. Hakim, J. Noorzaei, et al. "Application of artificial neural networks to predict compressive strength of high strength concrete." In: *International Journal of Physical Sciences* 6(5) (Mar. 2011).

[96] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. URL: `http://www.deeplearningbook.org`.

[97] S. Abdulwahab, M. Jabreel, and D. Moreno. "Deep Learning Models for Paraphrases Identification." PhD thesis. Sept. 2017. DOI: `10.13140/rg.2.2.15743.46240`.

[98] F. Informatik, Y. Bengio, P. Frasconi, et al. "Gradient Flow in Recurrent Nets: the Difficulty of Learning Long-Term Dependencies." In: *A Field Guide to Dynamical Recurrent Neural Networks* (Mar. 2003).

[99] S. Hochreiter and J. Schmidhuber. "Long Short-Term Memory." In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. ISSN: 0899-7667. DOI: `10.1162/neco.1997.9.8.1735`.

[100] K. Cho, B. van Merrienboer, C. Gulcehre, et al. *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation*. 2014. eprint: `1406.1078`.

[101] J. Chung, C. Gulcehre, K. Cho, et al. "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling." In: (2014). eprint: `1412.3555`.

[102] F. Chollet et al. *Keras*. 2015. URL: `https://keras.io`.

[103] Martín Abadi, Ashish Agarwal, Paul Barham, et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. 2015. URL: `https://www.tensorflow.org/`.

[104] I. Antcheva, M. Ballintijn, B. Bellenot, et al. "ROOT – A C++ framework for petabyte data storage, statistical analysis and visualization." In: *Computer Physics Communications* 182.6 (2011), pp. 1384–1385. ISSN: 0010-4655. DOI: `https://doi.org/10.1016/j.cpc.2011.02.008`.

[105] R. Brun and F. Rademakers. "ROOT – An object oriented data analysis framework." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997), pp. 81–86. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/S0168-9002(97)00048-X`.

[106] F. James and M. Roos. "Minuit - a system for function minimization and analysis of the parameter errors and correlations." In: *Computer Physics Communications* 10.6 (1975), pp. 343–367. ISSN: 0010-4655. DOI: `https://doi.org/10.1016/0010-4655(75)90039-9`.

[107] LHCb Collaboration. "Precision measurement of the B_s^0–oscillation frequency with the decay $B_s^0 \to D_s^- \pi^+$." In: *New Journal of Physics* 15.5 (May 2013). DOI: `10.1088/1367-2630/15/5/053021`.

[108] A. Bertolin. "Precise determination of the B_s^0-$\overline{B}_s^0$ oscillation frequency." Individual work performed in the context of the published measurement.

[109] K. Heinicke. "Strange oscillating beauty-mesons." PhD thesis. TU Dortmund, 2021. DOI: `10.17877/de290r-22378`.

[110] A. Dziurda. "Precise determination of the B_s^0-$\overline{B}_s^0$ oscillation frequency." Individual work performed in the context of the published measurement.

[111] D. Martínez Santos and F. Dupertuis. "Mass distributions marginalized over per-event errors." In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 764 (2014), pp. 150–155. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2014.06.081`.

[112] N. L. Johnson. "Systems Of Frequency Curves Generated By Methods Of Translation." In: *Biometrika* 36.1-2 (June 1949), pp. 149–176. ISSN: 0006-3444. DOI: `10.1093/biomet/36.1-2.149`.

[113] M. Veronesi. "Quantum oscillations of neutral particles with the beauty quark." Presented 15 Nov 2021. 2021. URL: `https://cds.cern.ch/record/2790773`.

[114] LHCb Collaboration. "Measurement of CP violation in $B^0 \to \psi(\to \ell^+\ell^-)K_S^0(\to \pi^+\pi^-)$ decays." In: (2023). eprint: `2309.09728`.

[115] A. Rogozhnikov. "Reweighting with Boosted Decision Trees." In: *Journal of Physics: Conference Series* 762 (Oct. 2016), p. 012036. DOI: `10.1088/1742-6596/762/1/012036`.

[116] HFLAV Group. "Averages of b-hadron, c-hadron, and tau-lepton properties as of 2018." In: *The European Physical Journal C* 81.3 (Mar. 2021). DOI: `10.1140/epjc/s10052-020-8156-7`.

[117] M. D. Needham. *Momentum scale calibration of the LHCb spectrometer*. Tech. rep. Cern, 2017. URL: `https://cds.cern.ch/record/2260848`.

[118] K. Cranmer. "Kernel estimation in high-energy physics." In: *Computer Physics Communications* 136.3 (May 2001), pp. 198–207. DOI: `10.1016/s0010-4655(00)00243-5`.

[119] LHCb Collaboration. "Precision Measurement of $B_s^0 \to J/\psi KK$ Decays." In: *Physical Review Letters* 114.4 (Jan. 2015). DOI: `10.1103/physrevlett.114.041801`.

[120] LHCb Collaboration. "Updated measurement of time-dependent *CP*-violating observables in $B_s^0 \to J/\psi KK$ decays." In: *The European Physical Journal C* 79.8 (Aug. 2019). DOI: `10.1140/epjc/s10052-019-7159-8`.

[121] A. Lupato. "Measurement of *CP* asymmetry in $B_s^0 \to D_s^\mp K^\pm$ decays." Individual work performed in the context of the published measurement.

[122] J. Butter. "b→u: Exploring the magnitude and phase of b→u quark transitions through open-charmed beauty decays." PhD thesis. Vrije Universiteit Amsterdam, June 2023. ISBN: 9789464198089. DOI: `10.5463/thesis.256`.

[123] LHCb Collaboration. "Measurement of the branching fraction of the $B^0 \to D_s^+\pi^-$ decay." In: *The European Physical Journal C* 81.4 (Apr. 2021). DOI: `10.1140/epjc/s10052-020-08790-2`.

[124] LHCb Collaboration. "Model-Independent Study of Structure in $B^+ \to D^+D^-K^+$ Decays." In: *Phys. Rev. Lett.* 125 (24 Dec. 2020), p. 242001. DOI: `10.1103/PhysRevLett.125.242001`.

[125] "Reprint of: Mahalanobis, P.C. (1936) "On the Generalised Distance in Statistics."" In: *Sankhya A* 80.1 (2018), pp. 1–7. DOI: `10.1007/s13171-019-00164-5`.

[126] A. Dziurda. "Measurement of *CP* asymmetry in $B_s^0 \to D_s^\mp K^\pm$ decays." Individual work performed in the context of the published measurement.

[127] LHCb Collaboration. "Measurement of the CKM angle gamma from a combination of LHCb results." In: *Journal of High Energy Physics* 2016.12 (Dec. 2016). DOI: `10.1007/jhep12(2016)087`.

[128] M. Kenzie. "Measurement of *CP* asymmetry in $B_s^0 \to D_s^\mp K^\pm$ decays." Individual work performed in the context of the published measurement.

[129] LHCb Collaboration. *Improved measurement of CP violation parameters in $B_s^0 \to J/\psi K^+K^-$ decays in the vicinity of the $\phi(1020)$ resonance.* Tech. rep. Cern, 2023. eprint: `2308.01468`.

[130] *12th International Workshop on the CKM Unitarity Triangle.* Santiago de Compostela, Spain, Sept. 2023.

[131] I. Bediaga, J. M. De Miranda, F. Ferreira Rodrigues, et al. *Framework TDR for the LHCb Upgrade: Technical Design Report.* Tech. rep. 2012. URL: `https://cds.cern.ch/record/1443882`.

[132] A. Piucci. "The LHCb Upgrade." In: *Journal of Physics: Conference Series* 878.1 (July 2017), p. 012012. DOI: `10.1088/1742-6596/878/1/012012`.

[133] T. L. Collaboration. *LHCb VELO Upgrade Technical Design Report.* Tech. rep. 2013. URL: `https://cds.cern.ch/record/1624070`.

[134] LHCb Collaboration. *LHCb Tracker Upgrade Technical Design Report.* Tech. rep. 2014. URL: `https://cds.cern.ch/record/1647400`.

[135] M. Zaheer, S. Kottur, S. Ravanbakhsh, et al. *Deep Sets.* 2018. eprint: `1703.06114`.

[136] A. Paszke, S. Gross, F. Massa, et al. *PyTorch: An Imperative Style, High-Performance Deep Learning Library.* 2019. eprint: `1912.01703`.

[137] Y. Ganin, E. Ustinova, H. Ajakan, et al. *Domain-Adversarial Training of Neural Networks.* 2016. eprint: `1505.07818`.

[138] S. J. Pan, I. W. Tsang, J. T. Kwok, et al. “Domain Adaptation via Transfer Component Analysis.” In: *IEEE Transactions on Neural Networks* 22.2 (2011), pp. 199–210. DOI: `10.1109/tnn.2010.2091281`.

[139] ATLAS Collaboration, G. Aad, et al. “Measurement of the *CP*-violating phase ϕ_s in $B_s^0 \to J/\psi\phi$ decays in ATLAS at 13 TeV.” In: *Eur. Phys. J. C* 81.4 (2021), p. 342. DOI: `10.1140/epjc/s10052-021-09011-0`.

[140] CMS Collaboration. “Measurement of the *CP*-violating phase ϕ_{s} in the $B_{\mathrm{s}}^0 \to J/\psi\,\phi(1020) \to \mu^+\mu^-K^+K^-$ channel in proton-proton collisions at $\sqrt{s} = 13$ TeV.” In: *Phys. Lett. B* 816 (2021), p. 136188. DOI: `10.1016/j.physletb.2021.136188`.

[141] LHCb Collaboration. “Measurement of the B_s^0–$\overline{B}_s^0$ oscillation frequency $B_s^0 \to D_s^-\pi^+\pi^-\pi^+$ decays.” In: *Physics Letters B* 709.3 (2012), pp. 177–184. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/j.physletb.2012.02.031`.

[142] LHCb Collaboration. “Measurement of *CP* asymmetry in $B_s^0 \to D_s^\mp K^\pm$ decays.” In: *Journal of High Energy Physics* 2014.11 (Nov. 2014). DOI: `10.1007/jhep11(2014)060`.

Acknowledgement

The first person I owe gratitude to in the context of this thesis and of my academic career is Bernhard Spaan, who left us in 2021 way too early. In 2016, Prof. Spaan gave me a home in his working group, providing the subject to both my bachelor's and master's thesis. Later, he gathered all the necessities to create my PhD position. During this period, he always encouraged me to strengthen my involvement in the field of time-dependent CP-violation measurements, the closely related flavour tagging, and in the realisation of the SciFi tracker. Unfortunately, he neither can examine this thesis himself, nor enjoy further fruits of his commitment to teaching and research, such as the LHCb upgrade. So, it is up to us to continue his work.

My thanks go no less to Johannes. He was there for the entire group in this time which was tough for himself personally. Not only to me but also to the large bunch of residual students, he offered himself as a supervisor and made it possible to continue our research projects. But even before this, Johannes had a significant impact on my career. From the day I joined the group, he played a role similar to a mentor already. Besides, at an early stage of my studies, he turned my attention to the field of particle physics with his sparking passion while teaching the ”Particle Physics I” lectures.

The latter also applies to Kevin Kröninger, who shared his enthusiasm with us students in the KET lectures at the same time. Hence, I am glad Prof. Kröninger agreed to be the second assessor of this thesis. This way, he will not only be part of my first oral examination as a physics student in the context of the ”Anfängerpraktikum” but also of my defence, my last examination in the role of a student.

I have enjoyed the cooperation in the D_sK/π analysis team, namely Alessandro B., Jordy B., Agnieszka D., Sevda E., Conor F., Emmy G., Kevin H., Matt K., Anna L., Michele V. and Niels T., within the LHCb collaboration. Every of these persons contributed at different degree to these two impressive measurements. Together we have been able to overcome every challenge in the analyses.

Here, a special thanks goes to Agnieszka. Although we have not yet met each other in person (hopefully this will change soon at the CKM conference), she has always been a mentor to me in the past years and contributed significantly to my education as a scientist. In the last months, she has been relentlessly involved in the day-to-day

business of the analysis work, although she originally intended to only contribute with her enormous experience to the analysis of $B_s^0 \to D_s^{\mp} K^{\pm}$ decays. This way, the two of us have managed as driving forces to push the analysis to its completion. Without her strong dedication to the measurement, this analysis would still require a large amount of time for its completion.

The large contributions to the analysis made by Jordy, who has always been available for lively discussions, need to be honoured, too. I am looking forward with joy to our future collaboration and face-to-face discussions in Cambridge in Matt's group.

No less my gratitude goes to my colleagues in Dortmund. Especially to Kevin. First, he has supervised my master's project. Then, he introduced me to the $B_s^0 \to D_s^- \pi^+$ analysis, where he was a driving force and opened some doors for me. At the same time, he provided many examples to me of how, both well-written and very improvised, code can look like.

Next to Kevin, I have shared the office for a long time with Antje, and later Louis. I have really enjoyed the atmosphere and spirit in our office. The same applies to the "Schreib-Büro" I have shared with Gerwin, Maik and Vukan in the last months and to the apartment in Geneva, which I have shared with Patrick during our time at the SciFi commissioning. Since I have enjoyed the atmosphere in the Dortmund working group in general as well, I am glad to continue being part of it. This way, I will be able to return the favour of proofreading to all the readers of my thesis, especially Jan and Louis.

In this context, the Lamarr Institute has to be mentioned as well, which already funds my position over the last months. Not only it will allow me to continue my research projects while providing profound expertise in the field of machine learning, but more importantly, it has given me the freedom to complete the analyses and my thesis over the last months.

I would also like to thank my friends and family who have always provided welcome distractions from the daily work. An important role here has been taken by my basketball team at TVG Kaiserau, with whom I have encountered both joyful wins and frustrating defeats in recent years. I am glad the team has shown so much understanding to my personal time constraints in the last months. Although I am aware of all the loving mockery on my abstract research subject of "not well-behaving protons", which I will continue to hear in the future, I am looking forward to our upcoming season.

From an early age, my parents encouraged my curiosity and interest in science. This way, they provided an essential foundation for this thesis. Further, they always have been my biggest supporters. Together with my girlfriend Jula, they have constantly provided emotional and logistical support to me during the last months. This gave

me the freedom to focus myself on the completion of the analyses and this thesis. At the same time, they have handled all my moods well and have always had my back.

Thank you.

Danksagung

Die erste Person, der ich im Kontext dieser Arbeit und meinem wissenschaftlichen Werdegang zu Dank verpflichtet bin, ist Bernhard Spaan, der uns im Jahr 2021 viel zu früh verlassen musste. Prof. Spaan nahm mich 2016 in seiner Arbeitsgruppe auf und gab bereits sowohl mein Bachelorarbeits- als auch Masterarbeitsthema aus, bevor er den nötigen Rahmen für meine Promotionsstelle schaffte. In dieser Zeit bestärkte er mich stets mich sowohl in zerfallszeitabhängigen CP Verletzungsmessungen und dem eng verbundenen Flavour Tagging zu involvieren als auch in der Verwirklichung des Scfi Trackers. Leider kann er die Ergebnisse dieser Arbeit nicht mehr selbst begutachten und sich auch nicht an weiteren Früchten seines Engagements in Lehre und Forschung, wie dem LHCb Upgrade, erfreuen. So ist es nun an us seine Arbeit fortzuführen.

Nicht weniger gilt mein Dank Johannes. Er war für die gesamte Gruppe auch in der für ihn persönlich schweren Zeit da, in der er nicht nur mir sondern auch einer großen Zahl anderer Doktoranden ermöglichte ihre Projekte fortzuführen und sich als Betreuer zur Verfügung stellte. Aber nicht erst in dieser Rolle hat Johannes meinen akademischen Werdegang mitgestaltet. Bereits seit meinem ersten Tag in der Gruppe hat er mich in einer einem Mentor ähnlichen Rolle unterstützt. Außerdem hat er schon in einer frühen Phase meines Studiums mit seiner überspringenden Leidenschaft für dieses Feld in der "Teilchenphysik I" Vorlesung mein Augenmerk auf die Teilchenphysik gelenkt.

Letzteres gilt auch für Kevin Kröninger, der zur selben Zeit in der KET Vorlesung seinen Enthusiasmus mit uns Studierenden teilte. Umso mehr freut es mich, dass sich Prof. Kröninger bereit erklärt hat als Zweitgutachter meiner Arbeit zu fungieren. Auf diese Weise wird er nicht nur der ersten mündlichen Prüfung meines Studiums im Rahmen des "Anfängerpraktikums" sondern wohl auch meiner Disputation, meiner letzten Prüfung als Student beigewohnt haben.

Mit Freude blicke ich auch auf die Zusammenarbeit im $D_s K/\pi$ Analyse Team innerhalb der LHCb Kollaboration, namentlich Alessandro B., Jordy B., Agnieszka D., Sevda E., Conor F., Emmy G., Kevin H., Matt K., Anna L., Michele V. und Niels T., zurück. Wenn auch in unterschiedlichem Maße, hat jede dieser Personen ihren

Beitrag zu den beiden beeindruckenden Messungen beigetragen. Dabei haben wir gemeinsam jede Hürde erfolgreich nehmen können.

Ein besonderer Dank geht dabei an Agnieszka. Auch wenn wir uns bis heute noch nicht persönlich begegnet sind (hoffentlich ändert sich dies in Kürze auf der CKM Konferenz), hat sie in den letzten Jahren stets als mein Mentor agiert und maßgeblich zu meiner Ausbildung zum Wissenschaftler beigetragen. Obwohl sie lange Zeit beabsichtigte lediglich ihre umfangreiche Erfahrung zur $B_s^0 \to D_s^{\mp} K^{\pm}$ Analyse beizusteuern, hat sie sich in den letzten Monaten unermüdlich auch im Tagesgeschäft der Analyse Arbeit eingebracht. Auf diese Weise ist es uns beiden gemeinsam als treibende Kräfte gelungen die Messung unter hohem persönlichen Einsatz zu vervollständigen. Ohne Agnieszkas Aufopferungsbereitschaft würde sich die Analyse noch über einen großen Zeitraum weiter ziehen.

Auch Jordy hat große Beiträge zu dieser Messung geliefert und stand jeder Zeit für rege Diskussionen zur Verfügung. Daher freue ich mich bereits auf unsere zukünftige Zusammenarbeit und den direkten Austausch in Cambridge unter der Obhut von Matt.

Nicht zu vergessen sind natürlich auch meine Dortmunder Kollegen. Allen voran Kevin. Zuerst hat er meine Masterarbeit mitbetreut. Danach hat er mich in das Analyse Team eingeführt, in dem er große Anteile zur $B_s^0 \to D_s^- \pi^+$ Analyse beigetragen hat und mir vielmals ein Wegbereiter war. Gleichzeitig konnte ich von Kevin sowohl lernen wie qualitativ hochwertiger, aber auch sehr improvisierter Code aussehen kann.

Neben Kevin habe ich mir lange Zeit das Büro mit Antje und später auch Louis geteilt. Für die angenehme Arbeitsatmosphäre bin euch sehr dankbar, genauso wie meinem zwischenzeitlichen Mitbewohner in Genf Patrick und meinen späteren "Schreib-Büro"-Kollegen Gerwin, Maik und Vukan, aber auch dem Rest des Lehrstuhls / der Arbeitsgruppe. Daher bin ich froh, dass ich auch weiterhin Teil dieser Arbeitsgruppe sein werde und mich in Zukunft auch bei den zahlreichen Korrekturlesern meiner Arbeit, insbesondere bei Louis und Jan, aber auch allen anderen, revanchieren werden kann.

In diesem Kontext muss auch das Lamarr Institut gennant werden, an dem ich bereits in den vergangenen Monaten angestellt bin. Nicht nur ermöglicht es mir die Fortsetzung meiner Forschung und unterstützt mich dabei mit einer breiten Expertise aus dem Machine Learning, auch hat es mir den nötigen Freiraum in den letzten Monaten gegeben um die Analysen und diese Arbeit abzuschließen.

Auch möchte ich meinen Freunden und meiner Familie danken, die stets Ablenkung vom Arbeitsalltag geliefert haben. Eine wichtige Rolle hat dabei meine Basketballmanschaft beim TVG Kaiserau eingenommen mit der ich in den letzten Jahren sowohl freudige Aufstiege als auch ernüchternde Niederlagen und Abstiege bewältigt habe. Ich bin froh, dass die Mannschaft in den vergangenen Monaten so viel Verständnis für meine zeitliche Einschränkung hatte und ich freue mich auf die kommende Saison, auch wenn ich mir in dem Rahmen viel liebevollen Spott zu meiner realitätsfernen Forschung an ”Protonen, die nicht machen was sie sollen” anhören werde.

Bereits von kleinauf haben meine Eltern meine Neugier und mein Interesse an der Naturwissenschaft gefördert und damit einen wesentlichen Grundstein für die Entstehung dieser Arbeit geliefert. Zusammen mit meiner Freundin Jula sind sie stets meine größten Unterstützer gewesen. Während der letzten Monate sind sie stets bedacht gewesen mich sowohl emotional als auch logistisch zu unterstützen. Dadurch haben sie den Freiraum geschaffen mich völlig auf den Abschluss der Analysen und der Promotion zu fokussieren. Dabei haben sie jede meiner Launen hingenommen und standen stets hinter mir.

Danke.

> “An expert is a person who has made all the mistakes that can be made in a very narrow field”
>
> — Niels Bohr