technische universität
dortmund

Measurements of CP violation in $B^0 \to D\bar{D}$ decays at the LHCb experiment

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

Philipp Ibis

11. Mai 2022

Lehrstuhl für Experimentelle Physik V
Fakultät für Physik
Technische Universität Dortmund

Der Fakultät Physik der Technischen Universität Dortmund zur Erlangung des akademischen Grades eines Dr. rer. nat. vorgelegte Dissertation.

Betreuer: Prof. Dr. Bernhard Spaan
1. Gutachter: Prof. Dr. Johannes Albrecht
2. Gutachter: Prof. Dr. Kevin Kröninger

Abstract

Precise measurements of CKM matrix parameters to probe the Standard Model of particle physics are a main focus of current flavour-physics experiments. Analyses of $B \to D\overline{D}$ decays allow to assess higher-order Standard Model corrections in these measurements. Two decay-time-dependent measurements of *CP* violation in different $B^0 \to D\overline{D}$ decay modes are presented using data collected by the LHCb experiment.

The analysis of $B^0 \to D^{*\pm}D^{\mp}$ decays utilises data corresponding to an integrated luminosity of $9\,\text{fb}^{-1}$ collected at centre-of-mass energies of 7, 8 and 13 TeV. The results of the *CP* parameters correspond to the most precise measurement in this channel. They are consistent with the previous world average.

The analysis of $B^0 \to D^+D^-$ decays uses data corresponding to $6\,\text{fb}^{-1}$ collected at 13 TeV. The analysis is ongoing and currently still blind. The evaluated sensitivities show that this will be the most precise measurement of *CP* violation in this channel.

Kurzfassung

Ein Schwerpunk von aktuellen Flavourphysik-Experimenten sind Präzisionsmessungen von Parametern der CKM-Matrix zur Untersuchung des Standardmodells der Teilchenphysik. Auswertungen von $B \to D\overline{D}$ Zerfällen erlauben es Standardmodell-Korrekturen höherer Ordnung in diesen Messungen zu untersuchen. Es werden zwei Messungen von *CP* Verletzung in unterschiedlichen $B^0 \to D\overline{D}$ Zerfallskanälen präsentiert, die Daten des LHCb Experimentes benutzen.

Die Analyse von $B^0 \to D^{*\pm}D^{\mp}$ Zerfällen wird mit einem Datensatz durchgeführt, der einer integierten Luminosität von $9\,\text{fb}^{-1}$ entspricht und bei Schwerpunktsenergien von 7, 8 and 13 TeV aufgenommen wurde. Die Ergebnisse der *CP*-Parameter entsprechen der präzisesten Messung in diesem Kanal. Sie sind konsistent mit den vorherigen Weltmittelwerten.

Die Analyse des $B^0 \to D^+D^-$ Zerfalls benutzt Daten, die $6\,\text{fb}^{-1}$ entsprechen und bei 13 TeV aufgenommen wurden. Die Analyse ist noch nicht abgeschlossen und noch blind. Die Unsicherheiten werden bewertet und die präziseste Messung von *CP*-Verletzung in diesem Kanal erwartet.

Table of contents

1 Introduction

The Standard Model of particle physics (SM) [1–3] aims to describe all elementary particles and their interactions. It has been tested to extraordinary precision and predicted several particles that were later observed, *e.g.* the Higgs boson [4, 5]. Still, there has to be New Physics, *i.e.* physics from beyond the SM, as it is incomplete: gravity is not included and the existence of dark matter and dark energy is not explained, which account for around 95% of the energy density of the universe [6]. Neutrinos are considered to be massless in the SM, which is refuted by the observation of neutrino oscillations [7, 8]. Furthermore, the size of the matter-antimatter asymmetry observed in the universe is not explained.

It is assumed that immediately after the Big Bang the amounts of matter and antimatter were equal. Today, only matter and no larger amounts of antimatter can be found in the universe. In 1967, three conditions were proposed to explain this phenomenon [9]: the baryon number is violated, the C and CP symmetries are violated, and there was a departure from the thermal equilibrium. C and CP violation have been observed already in the middle of the 20th century [10, 11] and are included in the SM. Still, the amount of CP violation in the SM is not large enough to explain the matter-antimatter asymmetry [12]. CP violation is anchored in the SM by the complex phase of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, the matrix describing the probability of quark-flavour transitions [13, 14]. The CKM parameter β can be measured with high theoretical [15] and experimental [16–18] precision using $B^0 \to J/\psi K^0_S$ decays. Measurements of β are also possible using $B^0 \to D^{(*)+}D^{(*)-}$ decays, but due to pollution by hadronic phases the results are not theoretically clean [19]. Instead, these decays can be used to distinguish this higher-order SM pollution from New Physics in other decays. An example for this is the $B^0_s \to D^+_s D^-_s$ decay, which can be used to measure the additional CKM parameter β_s. The most precise measurements of β_s are based on $B^0_s \to J/\psi\phi$, $B^0_s \to J/\psi K^+K^-$ and $B^0_s \to J/\psi\pi^+\pi^-$ decays [20–26]. Measurements using the $B^0_s \to D^+_s D^-_s$ decay are also performed [27], but suffer from lower selection efficiencies. With increasing experimental precision, hadronic effects will be the source of the leading uncertainty in the measurements of β_s [28, 29]. This makes the $B^0_s \to D^+_s D^-_s$ decay attractive, as the hadronic effects can be constrained by exploiting the wide array of $B \to D\overline{D}$ decays, which include B^0, B^+, B^0_s and B^+_c mesons, D^+, D^0 and D^+_s mesons, as well as their excited or baryonic counterparts. This can be done using not only CP-violation measurements [15, 30–32] but also branching-ratio measurements [19, 33]. Thus,

$B \to D\overline{D}$ decays are an important tool to investigate hadronic uncertainties in measurements of β and β_s.

The data analysed in this thesis was collected at the Large Hadron Collider (LHC). The LHC enables searches for New Physics at the energy frontier, *i.e.* its centre-of-mass energy of up to 13 TeV is the highest of any collider. With the launch of the LHC, the opportunities to find New Physics in direct searches were greater than ever: the high centre-of-mass energy allows to measure New Physics particles directly if they are lighter than the interaction energy. Until now, no New Physics particles were found, which makes the focus on indirect searches even more important. Indirect searches use precision measurements of SM processes that can be affected by New Physics particles through interference, including *CP*-violation and branching-ratio measurements. They allow to probe energy scales that are orders of magnitude higher than the centre-of-mass energy of the LHC. The LHCb experiment focusses on precision measurements involving b- and c-quark decays and provides the data used in this thesis.

The work of the author is mainly focussed on analysing $B \to D\overline{D}$ decays. In this thesis the analyses of the measurements of *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays [34] and in $B^0 \to D^+D^-$ decays are presented. The latter is still ongoing but close to being finished. Further analyses the author contributed to are not discussed in this work. They comprise the measurement of *CP* violation in $B_s^0 \to D_s^+D_s^-$ decays, which measures β_s and is performed in parallel to the $B^0 \to D^+D^-$ analysis; the search for the $B_s^0 \to D^{*\pm}D^{\mp}$ decay and the measurement of its branching ratio [35]; the search for and branching-ratio measurement of the $B_s^0 \to D^{*+}D^{*-}$ decay [36], which is close to being published; and the measurement of *CP* violation in the $B^0 \to D^{*+}D^{*-}$ decay, which has begun recently.

A summary of *CP* violation in the SM and $B \to D\overline{D}$ decays in particular is given in Chap. 2. In Chap. 3, a short description of the LHCb detector is provided. Chapter 4 gives an overview of utilised analysis methods. The strategy of the analyses is discussed in Chap. 5, before the analyses are presented in Chaps. 6 and 7. A discussion of the results and an outlook of future measurements is given in Chap. 8.

2 *CP* violation in the Standard Model of particle physics

The Standard Model of particle physics offers the theoretical description of fundamental particles and their interactions. The SM is a Lorentz invariant quantum field theory and fulfils the *CPT* theorem [37,38], *i.e.* it is invariant under the combination of the three discrete transformations C, P, and T. Charge conjugation C inverts all charges; parity transformation P inverts spatial coordinates so that $(t, x) \to (t, -x)$; and time reversal T inverts time, $(t, x) \to (-t, x)$. While *CPT* symmetry is always conserved in the SM, symmetries under any single transformation or any combination of two of these transformations, *e.g.* the *CP* symmetry, can be violated.

This chapter summarises the fundamentals for decay-time-dependent measurements of CP violation to assess CKM elements. An overview of the particles and forces in the SM is given based on Refs. [39,40] in Sec. 2.1. The CKM mechanism is outlined in Sec. 2.2. Decay-time-dependent measurements depend on the mixing of neutral mesons to extract *CP* parameters, which is summarised in Sec. 2.3. The different types of *CP* violation are summarised in Sec. 2.4. The descriptions of the CKM mechanism, neutral-meson mixing and *CP* violation are based on Refs. [41–45]. Finally, the *CP* parameters in $B \to D\overline{D}$ decays is discussed in Sec. 2.5 using Ref. 19.

2.1 Particles and interactions

The fundamental particles of the SM are shown in Fig. 2.1. They can be divided into fermions with spin $1/2$ and bosons with integer spin.

The fermions can be subdivided into six quarks and six leptons, which together form all observed matter. Each of the fermions has an antifermion carrying inverted charges. The three up-type quarks (u, c, t) carry electric charges of $+2/3\,e$, while the down-type quarks (d, s, b) carry electric charges of $-1/3\,e$. Here, e is the elementary charge. All quarks also carry a colour charge allowing them to interact strongly. Their strong interaction leads to confinement [46], *i.e.* quarks always occur in bound states, so-called hadrons, which are always colourless. The most common types of hadrons are mesons, consisting of one quark and one antiquark

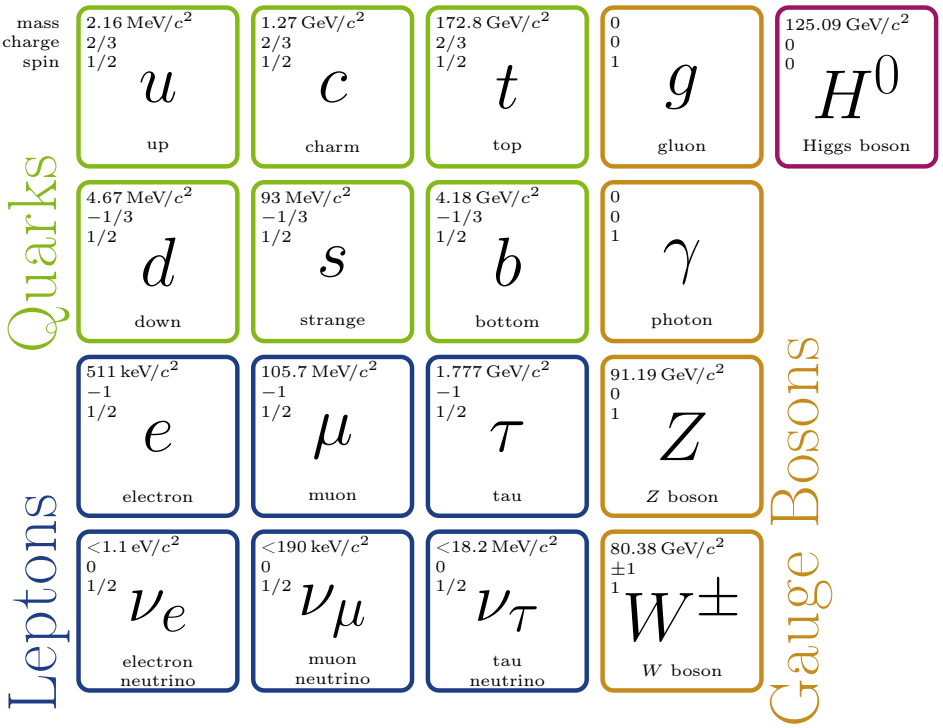

Figure 2.1: Schematic overview of the particles of the Standard Model. The properties are taken from Ref. 41.

($q\overline{q}$), and baryons, consisting of three quarks (qqq). Exotic tetraquarks ($qq\overline{qq}$) have been observed by the Belle experiment in 2003 [47]. Recently, the LHCb experiment found additional tetraquark states [48–50] and also observed pentaquarks ($qqq\overline{qq}$) for the first time [51, 52]. Leptons are subdivided into charged leptons with electric charges of $-1\,e$ (e, μ, τ), and the uncharged neutrinos (ν_e, ν_μ, ν_τ), which do not interact electromagnetically. Quarks and leptons can also be grouped in respectively three families, following a mass hierarchy. Each of the first three columns in Fig. 2.1 comprise one quark and one lepton family.

The bosons can be subdivided into two groups by their spin: the spin-1 gauge bosons, which mediate fundamental interactions, and the spin-0 Higgs boson. The mediators of the strong interaction are the eight neutral gluons (g), each carrying a colour and an anticolour charge with different configurations. They couple to the colour charge of quarks and other gluons. The neutral photon (γ) is the mediator of the electromagnetic interaction and couples to the electric charge. The neutral Z meson and the charged $W^\pm$ mesons are the mediators of the weak interaction. They couple to left-handed particles and right-handed antiparticles. The SM is completed by the Higgs boson (H^0). It was proposed as a consequence of the Higgs mechanism [53, 54], *i.e.* the mechanism that gives the gauge bosons their masses, and was discovered in 2012 [4, 5]. The Higgs mechanism mechanism is not directly responsible for the masses of the fermions. Instead, the masses of fermions

arise due to Yukawa interactions with the Higgs field, as explained in the following section.

2.2 The CKM mechanism

Yukawa interactions of the quark and Higgs fields are responsible for the mixing and masses of quarks,

$$\mathcal{L}_Y = -Y^d_{ij}\overline{Q^I_{Li}}\phi d^I_{Rj} - Y^u_{ij}\overline{Q^I_{Li}}\epsilon\phi^* u^I_{Rj} + \text{h.c.}, \tag{2.1}$$

where $Y^{u,d}_{ij}$ are 3×3 complex matrices, ϕ is the Higgs field, i and j are quark-family labels, ϵ is the 2×2 antisymmetric tensor, Q^I_{Li} are left-handed quark doublets, and u^I_{Rj} are right-handed up- and d^I_{Rj} right-handed down-type quark singlets. Due to the vacuum expectation value of ϕ, $\langle\phi\rangle = (0, v/\sqrt{2})$, Eq. (2.1) gives mass terms for the quarks. Using the four unitary matrices $V^{u,d}_{L,R}$ allows to diagonalize Y^f to

$$M^f_{\text{diag}} = V^f_L Y^f V^{f\dagger}_R \left(\frac{v}{\sqrt{2}}\right), \tag{2.2}$$

where $f = u, d$. The charged-current interactions, *i.e.* interactions mediated by $W^\pm$ mesons, couple to u_{Lj} and d_{Lj} quarks according to

$$\frac{-g}{\sqrt{2}}\left(u_L, c_L, t_L\right)\gamma^\mu {W^+}_\mu V_{\text{CKM}}\begin{pmatrix} d_L \\ s_L \\ b_L\end{pmatrix} + \text{h.c.} \tag{2.3}$$

Here g is the weak coupling constant, γ^μ are the Dirac matrices and V_{CKM} is the CKM matrix [13, 14] given by

$$V_{\text{CKM}} = V^u_L V^{d\dagger}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb}\end{pmatrix}. \tag{2.4}$$

The CKM matrix is a 3×3 unitary matrix. It can be parametrised by three mixing angles, θ_{12}, θ_{23} and θ_{13}, and the *CP*-violating phase, δ. All *CP*-violating processes in the SM involve flavour-changing processes and are caused by the phase δ. One possible parametrisation is the Chau-Keung parametrisation [55],

$$\begin{aligned} V_{\text{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23}\end{pmatrix}\begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13}\end{pmatrix}\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1\end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13}\end{pmatrix}, \end{aligned} \tag{2.5}$$

where $s_{ij} = \sin(\theta_{ij})$ and $c_{ij} = \cos(\theta_{ij})$. The Wolfenstein parametrisation [56], another parametrisation of the CKM matrix, uses the experimental finding that the CKM matrix shows the hierarchy $s_{13} \ll s_{23} \ll s_{12} \ll 1$. The exact Wolfenstein parametrisation [57] is obtained by imposing the relations

$$\begin{aligned} s_{12} &= \lambda\,, \\ s_{23} &= A\lambda^2\,, \\ s_{13}\mathrm{e}^{\mathrm{i}\delta} &= A\lambda^3(\rho+\mathrm{i}\eta) = \frac{A\lambda^3(\overline{\rho}+\mathrm{i}\overline{\eta})\sqrt{1-A^2\lambda^4}}{\sqrt{1-\lambda^2}\left[1-A^2\lambda^4(\overline{\rho}+\mathrm{i}\overline{\eta})\right]}\,. \end{aligned} \tag{2.6}$$

Since $(\overline{\rho}+\mathrm{i}\overline{\eta})$ is phase convention independent, the extension of the CKM matrix to any order of $\lambda \approx 0.22$ written in terms of λ, A, $\overline{\rho}$ and $\overline{\eta}$ is still unitary. The Wolfenstein parametrisation is commonly approximated to

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

Because of its unitarity, the CKM matrix fulfils the conditions

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk}\,, \quad \sum_j V_{ij}V_{kj}^* = \delta_{ik}\,. \tag{2.8}$$

Six of these combinations vanish and can be represented as triangles in the complex $\overline{\rho}$-$\overline{\eta}$-plane. All six triangles have the same area, *i.e.* half of the Jarlskog invariant J [58], defined by

$$\mathcal{I}m\left[V_{ij}V_{kl}V_{il}^*V_{kj}^*\right] = J\sum_{m,n}\epsilon_{ikm}\epsilon_{jln}\,. \tag{2.9}$$

This invariant is independent of phase convention and a measure of *CP* violation in the SM.

The most commonly used triangle is the *bd* triangle arising from the condition

$$\begin{aligned} & V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \\ \Leftrightarrow\ & \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + \quad 1 \quad + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0\,. \end{aligned} \tag{2.10}$$

While most of the CKM triangles have side lengths of different orders of λ, *i.e.* they are squashed, all side lengths of the *bd* triangle are of order $\mathcal{O}(\lambda^3)$. This along with the fact that its parameters are assessable using B-meson decays makes this triangle attractive for precision measurements of the SM. This is done by measuring the side lengths

$$R_b = \left|\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right|\,, \quad R_t = \left|\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right|\,, \tag{2.11}$$

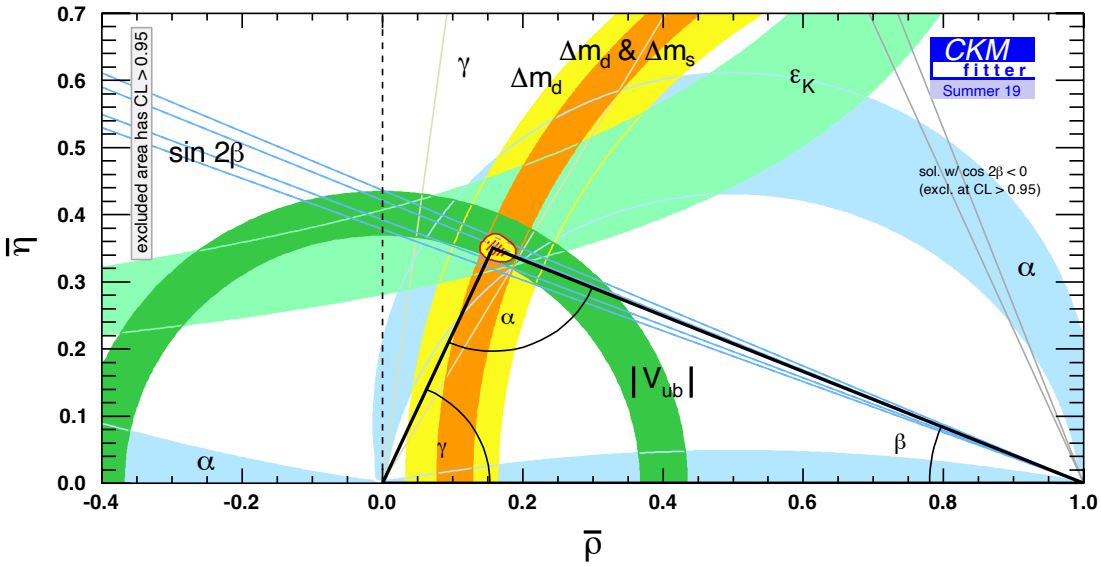

Figure 2.2: The *bd* unitary triangle in the $\overline{\rho}$-$\overline{\eta}$-plane [59]. Experimental constraints are shown as coloured bands. The red hatched and yellow areas around the apex correspond to the 68% and 95% confidence levels of the combination.

and the CKM angles,

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right), \tag{2.12}$$

and overconstraining the triangle. In the SM, the triangle has to close in its apex, due to the unitarity of the CKM matrix, and contradictory measurements would be proof of New Physics. The current status of measurements of its parameters can be seen in Fig. 2.2 and is in good agreement with the SM.

The unitarity condition

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0\,, \tag{2.13}$$

leads to the nearly degenerate *bs* triangle, as the side lengths are in the order of $\mathcal{O}(\lambda^4)$, $\mathcal{O}(\lambda^2)$ and $\mathcal{O}(\lambda^2)$. The parameters of this triangle are assessable using B_s^0-meson decays. Of these, the angle β_s is of articular interest, it is defined as

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

As β_s is the angle between the two longer sides, its value is expected to be small, of order one degree [42].

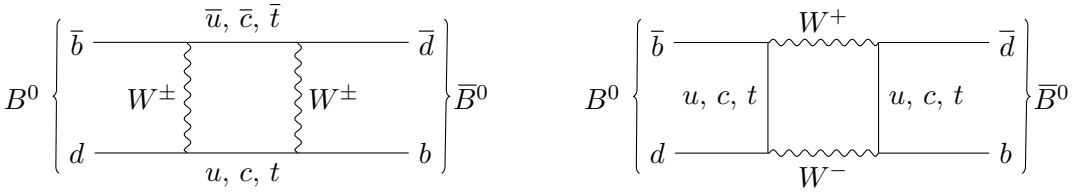

Figure 2.3: Dominant box diagrams of the B^0-$\overline{B}^0$ mixing.

2.3 The neutral meson system

Flavoured neutral mesons have the ability to oscillate between particle and antiparticle via flavour-changing charged currents. This process is called mixing and the lowest order Feynman diagrams of mixing in the B^0 system, so-called box diagrams, are shown in Fig. 2.3. Mixing and decays of flavoured neutral mesons involve the weak interaction and can thus be *CP*-violating. *CP* violation in these processes is particularly interesting, as many of the observables can be cleanly interpreted.

The following description is valid for K^0, D^0, B^0 and B_s^0 decays, but the outline follows the B^0 system with the initial flavour eigenstates

$$\left|B^0\right\rangle = \left|\bar{b}d\right\rangle = \left|B^0(t=0)\right\rangle \,, \quad \left|\overline{B}^0\right\rangle = \left|b\bar{d}\right\rangle = \left|\overline{B}^0(t=0)\right\rangle \,. \tag{2.15}$$

The sign convention from Ref. 43 is used. The time evolution of these states is described by a Schrödinger equation:

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\left|B^0(t)\right\rangle \\ \left|\overline{B}^0(t)\right\rangle\end{pmatrix} = \left(\mathbf{M} - \frac{\mathrm{i}}{2}\mathbf{\Gamma}\right)\begin{pmatrix}\left|B^0(t)\right\rangle \\ \left|\overline{B}^0(t)\right\rangle\end{pmatrix} \,. \tag{2.16}$$

Here, $\mathbf{M}$ is the mass matrix and $\mathbf{\Gamma}$ is the decay matrix, which are both time-independent Hermitian 2×2 matrices. Due to *CPT* invariance, the mass and lifetime of B^0 and $\overline{B}^0$ is identical and the respective diagonal entries must be the same:

$$M_{11} = M_{22} \,, \quad \Gamma_{11} = \Gamma_{22} \,. \tag{2.17}$$

The presence of non-zero off-diagonal elements allows transitions from B^0 to $\overline{B}^0$ and vice versa, *i.e.* mixing. This can be described by introducing the eigenstates of $\mathbf{M} - \frac{\mathrm{i}}{2}\mathbf{\Gamma}$:

$$\left|B_\mathrm{L}\right\rangle = p\left|B^0\right\rangle + q\left|\overline{B}^0\right\rangle \,, \quad \left|B_\mathrm{H}\right\rangle = p\left|B^0\right\rangle - q\left|\overline{B}^0\right\rangle \,. \tag{2.18}$$

Here, $|B_\mathrm{L}\rangle$ and $|B_\mathrm{H}\rangle$ are the light and heavy eigenstate with the masses m_L, m_H and decay widths Γ_L, Γ_H, respectively. Their time evolution is given by

$$\left|B_{\mathrm{L,H}}(t)\right\rangle = \mathrm{e}^{-\left(\mathrm{i}\,m_{\mathrm{L,H}} + \Gamma_{\mathrm{L,H}}/2\right)t}\left|B_{\mathrm{L,H}}\right\rangle \,. \tag{2.19}$$

The complex numbers p and q in Eq. (2.18) obey the normalisation $|q|^2+|p|^2=1$. Their fraction can be expressed in terms of M_{12} and Γ_{12},

$$\frac{q}{p}=\frac{\Delta m+\frac{\mathrm{i}}{2}\Delta\Gamma}{2(M_{12}-\frac{\mathrm{i}}{2}\Gamma_{12})}=\frac{2(M_{12}^*-\frac{\mathrm{i}}{2}\Gamma_{12}^*)}{\Delta m+\frac{\mathrm{i}}{2}\Delta\Gamma}\,, \tag{2.20}$$

using the definitions

$$\begin{aligned} m&=\frac{m_\mathrm{H}+m_\mathrm{L}}{2}=M_{11}\,, & \Gamma&=\frac{\Gamma_\mathrm{H}+\Gamma_\mathrm{L}}{2}=\Gamma_{11}\,,\\ \Delta m&=m_\mathrm{H}-m_\mathrm{L}\,, & \Delta\Gamma&=\Gamma_\mathrm{L}-\Gamma_\mathrm{H}\,, \end{aligned} \tag{2.21}$$

for the average mass and width, and the mass and width differences of the heavy and light eigenstate. Due to the fact that the contributions from top quarks are dominant in the diagrams in Fig. 2.3 [60], Eq. (2.20) can be approximated to

$$\frac{q}{p}\approx\frac{V_{td}V_{tb}^*}{V_{tb}V_{td}^*}\,. \tag{2.22}$$

The same approximation holds for B_s^0 mesons exchanging d and s quarks.

The time evolution of the flavour eigenstates can now be derived using Eqs. (2.18) and (2.19):

$$\begin{aligned} \left|B^0(t)\right\rangle&=\frac{1}{2p}\left[\mathrm{e}^{-(\mathrm{i}m_\mathrm{L}+\Gamma_\mathrm{L}/2)t}\left|B_\mathrm{L}\right\rangle+\mathrm{e}^{-(\mathrm{i}m_\mathrm{H}+\Gamma_\mathrm{H}/2)t}\left|B_\mathrm{H}\right\rangle\right]\,,\\ \left|\overline{B}^0(t)\right\rangle&=\frac{1}{2q}\left[\mathrm{e}^{-(\mathrm{i}m_\mathrm{L}+\Gamma_\mathrm{L}/2)t}\left|B_\mathrm{L}\right\rangle-\mathrm{e}^{-(\mathrm{i}m_\mathrm{H}+\Gamma_\mathrm{H}/2)t}\left|B_\mathrm{H}\right\rangle\right]\,. \end{aligned} \tag{2.23}$$

By using Eq. (2.18) again, the time evolution of the flavour eigenstates can be expressed with the initial flavour eigenstates:

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

where

$$\begin{aligned} g_+(t)&=\mathrm{e}^{-\mathrm{i}mt}\mathrm{e}^{-\Gamma t/2}\left[\;\;\cosh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta mt}{4}-\mathrm{i}\sinh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta mt}{4}\right]\,,\\ g_-(t)&=\mathrm{e}^{-\mathrm{i}mt}\mathrm{e}^{-\Gamma t/2}\left[-\sinh\frac{\Delta\Gamma t}{4}\cos\frac{\Delta mt}{4}+\mathrm{i}\cosh\frac{\Delta\Gamma t}{4}\sin\frac{\Delta mt}{4}\right]\,. \end{aligned} \tag{2.25}$$

The decay amplitudes of a pure B^0 or $\overline{B}^0$ state decaying into the final state f,

$$A_f=\left\langle f\,\middle|\,B^0\right\rangle\,,\quad \overline{A}_f=\left\langle f\,\middle|\,\overline{B}^0\right\rangle\,, \tag{2.26}$$

and the quantity

$$\lambda_f = \frac{1}{\overline{\lambda}_f} = \frac{q}{p}\frac{\overline{A}_f}{A_f}\,, \tag{2.27}$$

are useful to describe the time-dependent decay rates. These can be calculated with

$$\Gamma(B^0(t)\to f) = N_f \left|\left\langle f \,\middle|\, B^0(t)\right\rangle\right|^2\,, \quad \Gamma(\overline{B}^0(t)\to f) = N_f \left|\left\langle f \,\middle|\, \overline{B}^0(t)\right\rangle\right|^2\,, \tag{2.28}$$

where N_f is a time-independent normalisation factor stemming from the kinematics of the decay. The resulting formulas for the time-dependent decay rates are:

$$\begin{aligned}
\Gamma(B^0(t)\to f) &\propto \mathrm{e}^{-\Gamma t}|A_f|^2(1+|\lambda_f|^2) \\
&\qquad \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + A_f^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right) + C_f\cos\left(\Delta m t\right) - S_f\sin\left(\Delta m t\right)\right], \\
\Gamma(\overline{B}^0(t)\to f) &\propto \mathrm{e}^{-\Gamma t}|A_f|^2(1+|\lambda_f|^2)|\frac{p}{q}|^2 \\
&\qquad \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + A_f^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right) - C_f\cos\left(\Delta m t\right) + S_f\sin\left(\Delta m t\right)\right].
\end{aligned} \tag{2.29}$$

Here, the *CP* observables

$$A_f^{\Delta\Gamma} = -\frac{2\mathcal{R}e(\lambda_f)}{1+|\lambda_f|^2}\,, \quad C_f = \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}\,, \quad S_f = \frac{2\mathcal{I}m(\lambda_f)}{1+|\lambda_f|^2}\,, \tag{2.30}$$

fulfil the normalisation condition

$$(A_f^{\Delta\Gamma})^2 + (C_f)^2 + (S_f)^2 = 1\,. \tag{2.31}$$

They describe all *CP*-violating effects stemming from the complex phase of the CKM triangle.

Finally, the results for the *CP*-conjugate final state $\overline{f}$ can be obtained by replacing f with $\overline{f}$. In this case it is more useful to factor out $|\overline{A}_{\overline{f}}|$ and $|\overline{\lambda}_{\overline{f}}|$ instead of $|A_f|$ and $|\lambda_f|$:

$$\begin{aligned}
\Gamma(B^0(t)\to \bar{f}) &\propto \mathrm{e}^{-\Gamma t}|\overline{A}_{\overline{f}}|^2(1+|\overline{\lambda}_{\overline{f}}|^2)|\frac{q}{p}|^2 \\
&\qquad \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + A_{\overline{f}}^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right) + C_{\overline{f}}\cos\left(\Delta m t\right) - S_{\overline{f}}\sin\left(\Delta m t\right)\right], \\
\Gamma(\overline{B}^0(t)\to \bar{f}) &\propto \mathrm{e}^{-\Gamma t}|\overline{A}_{\overline{f}}|^2(1+|\overline{\lambda}_{\overline{f}}|^2) \\
&\qquad \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + A_{\overline{f}}^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right) - C_{\overline{f}}\cos\left(\Delta m t\right) + S_{\overline{f}}\sin\left(\Delta m t\right)\right].
\end{aligned} \tag{2.32}$$

The *CP* observables for the decays into f and $\overline{f}$ do not have to be identical, introducing the additional set of *CP* observables $A_{\overline{f}}^{\Delta\Gamma}$, $C_{\overline{f}}$ and $S_{\overline{f}}$.

Eqs. (2.29) and (2.32) are often called master equation of time evolution. They are commonly used to assess CKM matrix elements by measuring the *CP* observables. The relationship of the *CP* observables and the CKM matrix elements in $B \to D\overline{D}$ decays is shown in Sec. 2.5.

2.4 The three types of *CP* violation

The three observables

$$\left|\frac{q}{p}\right|, \quad \left|\frac{\overline{A}_{\overline{f}}}{A_f}\right|, \quad \lambda_f, \tag{2.33}$$

can be used to classify three different types of *CP* violation. In most decays only one of these types occurs.

2.4.1 Direct *CP* violation

Direct *CP* violation, often also called *CP* violation in decay, can be assessed using the quantity $|\overline{A}_{\overline{f}}/A_f|$. Two types of complex phases can enter the decay amplitudes: the so-called weak phases, which enter *CP*-conjugate amplitudes with opposite signs, and strong phases, whose sign do not change in the *CP*-conjugate case. Weak phases only occur in the CKM matrix while strong phases arise from rescattering from on-shell intermediate resonances into the final state. Hence, the two decay amplitudes can be expressed as

$$A_f = \sum_k A_k e^{i\delta_k + \phi_k}, \quad \overline{A}_{\overline{f}} = \sum_k A_k e^{i\delta_k - \phi_k}, \tag{2.34}$$

where k is the label of the different contributions, A_k is their magnitude, and δ_k and ϕ_k are their strong and weak phases. If

$$\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right| = \frac{\sum_k A_k e^{i\delta_k - \phi_k}}{\sum_k A_k e^{i\delta_k + \phi_k}} \neq 1, \tag{2.35}$$

CP is violated. This can occur as a result of interference between several terms in the decay amplitude, but requires at least two terms with different strong and weak phases. While direct *CP* violation has been known to occur in the kaon and B systems for a while [61–63], it was recently also observed in the D system [64].

2.4.2 Indirect *CP* violation

Indirect *CP* violation, or *CP* violation in mixing, occurs when the mass eigenstates differ from the *CP* eigenstates. In this case, the relative phase between M_{12} and Γ_{12} is non-zero so that

$$\left|\frac{q}{p}\right|^2 \neq 1\,, \tag{2.36}$$

using Eq. (2.20). While indirect *CP* violation has been observed in the kaon system [11], all measurements in the B^0 and B_s^0 system [43] lack the precision to observe the small value predicted by the SM [65].

2.4.3 Interference *CP* violation

Interference *CP* violation is also called *CP* violation in the interference of decay and decay after mixing as it arises from the interference between $B^0 \to f$ and $B^0 \to \overline{B}^0 \to f$ transitions. This can only occur if the final state is common to both B^0 and $\overline{B}^0$ mesons. It is defined by

$$\arg \lambda_f + \arg \lambda_{\overline{f}} \neq 0\,. \tag{2.37}$$

This type of *CP* violation is commonly measured in decays into a *CP*-eigenstate final state, f_{CP}. Equation (2.27) can then be rewritten to

$$\lambda_{f_{CP}} = \eta_{CP} \frac{q}{p} \frac{\overline{A}_{f_{CP}}}{A_{f_{CP}}}\,, \tag{2.38}$$

and in the case of *CP* violation

$$\mathcal{I}m\lambda_{f_{CP}} \neq 0\,, \tag{2.39}$$

holds. Interference *CP* violation can even occur in the absence of *CP* violation in decay and *CP* violation in mixing. In this case, the phases of q/p and $\overline{A}_{f_{CP}}/A_{f_{CP}}$ vanish, but the relative phase between the two terms does not. The most important of this type of *CP* violation are performed in $B^0 \to J/\psi\, K_S^0$ and $B_s^0 \to J/\psi\, \phi$ decays, where the most precise results come from the LHCb collaboration [18, 24]. The $B^0 \to D^{*\pm} D^{\mp}$ and $B^0 \to D^+ D^-$ decays both exhibit interference *CP* violation.

2.5 *CP* violation in $B^0 \to D^+ D^-$ and $B^0 \to D^{*\pm} D^{\mp}$ decays

This section summarises the amplitudes and observables of the $B^0 \to D^+ D^-$ decay, which comprises an *CP*-eigenstate final state, before doing the same for the more

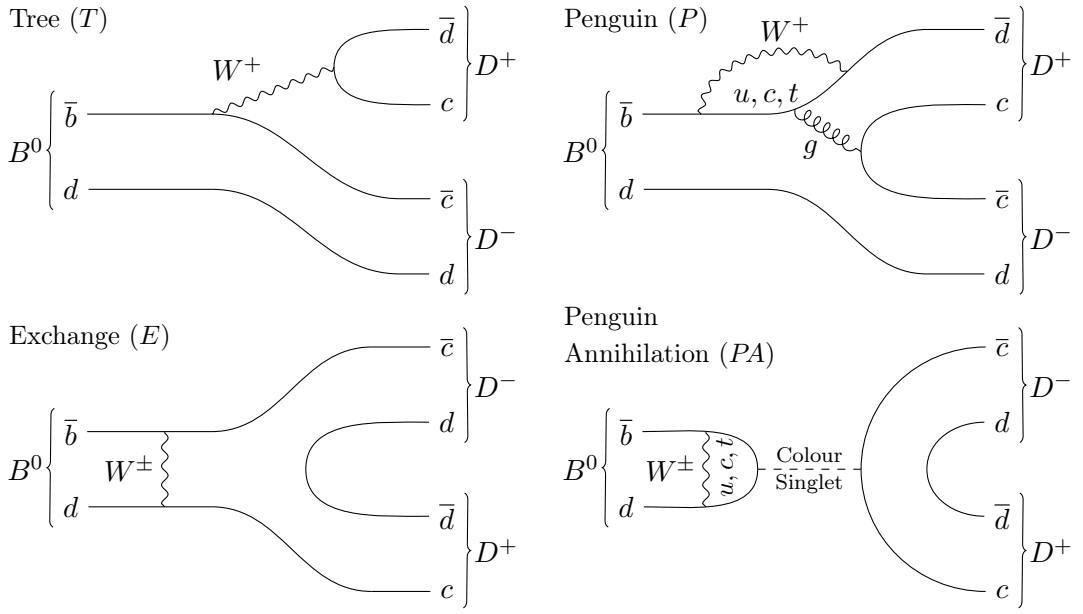

Figure 2.4: The (top left) tree, (top right) penguin, (bottom left) exchange and (bottom right) penguin annihilation diagrams of the $B^0 \to D^+D^-$ decay. Any of the two $D^{\pm}$ mesons can be replaced by an excited state to obtain the $B^0 \to D^{*\pm}D^{\mp}$ decay.

complicated $B^0 \to D^{*+}D^-$ and $B^0 \to D^{*-}D^+$ decays, which have two different final states allowing the measurement of additional *CP* observables. The dominant contributions to the $B^0 \to D^+D^-$ and $B^0 \to D^{*\pm}D^{\mp}$ decays are shown in Fig. 2.4. The amplitude of the $B^0 \to D^+D^-$ decay can be written as [15, 19, 32]

$$\begin{aligned} A(B^0 \to D^+D^-) = {} & V_{cd}V_{cb}^*(T + E + P_c + PA_c) \\ & + V_{ud}V_{ub}^*(P_u + PA_u) + V_{td}V_{tb}^*(P_t + PA_t)\,. \end{aligned} \tag{2.40}$$

The strong amplitudes are defined according to Fig. 2.4 and the index of P and PA denotes which up-type quark contributes in the respective loop. The CKM condition from Eq. (2.10) can be used to eliminate the $V_{td}V_{tb}^*$ term:

$$\begin{aligned} A(B^0 \to D^+D^-) = {} & V_{cd}V_{cb}^*(T + E + P_c + PA_c - P_t - PA_t) \\ & + V_{ud}V_{ub}^*(P_u + PA_u - P_t - PA_t). \end{aligned} \tag{2.41}$$

Using

$$\mathcal{A} = (T + E + P_c + PA_c - P_t - PA_t)\,, \tag{2.42}$$

the amplitude results in

$$A(B^0 \to D^+D^-) = V_{cd}V_{cb}^*\mathcal{A}\left(1 + \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\frac{P_u + PA_u - P_t - PA_t}{\mathcal{A}}\right)\,. \tag{2.43}$$

This can be further simplified with

$$-R_b \mathrm{e}^{\mathrm{i}\gamma} = \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\,, \tag{2.44}$$

from Eqs. (2.11) and (2.12) and

$$a\mathrm{e}^{\mathrm{i}\theta} = R_b \left(\frac{P_u + PA_u - P_t - PA_t}{\mathcal{A}} \right)\,, \tag{2.45}$$

resulting in

$$A(B^0 \to D^+D^-) = V_{cd}V_{cb}^*\mathcal{A}(1 - a\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{\mathrm{i}\gamma})\,. \tag{2.46}$$

Here, $\mathcal{A}$, a and θ are hadronic parameters with strong phases, while the CKM angle γ is a weak phase. Thus, the *CP*-conjugate amplitude is

$$A(\overline{B}^0 \to D^+D^-) = V_{cb}V_{cd}^*\mathcal{A}(1 - a\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\gamma})\,. \tag{2.47}$$

Inserting both amplitudes and Eq. (2.22) into the formula for λ_f in Eq. (2.27) results in

$$\lambda_{D^+D^-} = \frac{V_{td}V_{tb}^*}{V_{tb}V_{td}^*}\frac{V_{cb}V_{cd}^*}{V_{cd}V_{cb}^*}\frac{1 - a\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\gamma}}{1 - a\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{\mathrm{i}\gamma}}\,. \tag{2.48}$$

The remaining CKM matrix elements can be expressed by the CKM angle β with Eq. (2.12):

$$\lambda_{D^+D^-} = \mathrm{e}^{-\mathrm{i}2\beta}\frac{1 - a\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\gamma}}{1 - a\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{\mathrm{i}\gamma}}\,. \tag{2.49}$$

This can be used to find the *CP* observables using Eq. (2.30):

$$\begin{aligned} S_{D^+D^-} &= -\frac{\sin 2\beta - 2a\cos\theta\sin(2\beta+\gamma) + a^2\sin(2\beta+2\gamma)}{1 - 2a\cos\theta\cos\gamma + a^2}\,, \\ C_{D^+D^-} &= \frac{2a\sin\theta\sin\gamma}{1 - 2a\cos\theta\cos\gamma + a^2}\,. \end{aligned} \tag{2.50}$$

The third *CP* observable $A_f^{\Delta\Gamma}$ is not observable in the B^0 system, as $\Delta\Gamma_d \approx 0$, making the $A_f^{\Delta\Gamma}$ term in Eq. (2.29) vanish. Here, direct *CP* asymmetries, *i.e.* $C_{D^+D^-} \neq 0$ are caused by interference between tree and penguin contributions, while $S_{D^+D^-}$ measures mixing induced *CP* asymmetries, which originate in interference between mixing and decay, and depends on the mixing phase $\phi_d = 2\beta$. Due to the non-negligible contributions of the non-tree diagrams in Fig. 2.4, only the effective phase

$$\phi_{d,D^+D^-}^{\mathrm{eff}} = \phi_d + \Delta\phi_{d,D^+D^-}\,, \tag{2.51}$$

with

$$\sin \phi_{d,D^+D^-}^{\text{eff}} = -\frac{S_{D^+D^-}}{\sqrt{1 - {C_{D^+D^-}}^2}}\,, \tag{2.52}$$

can be measured. In contrast, the gold-plated $B^0 \to J/\psi\, K_S^0$ decay allows a clean measurement of ϕ_d with

$$\begin{aligned} S_{J/\psi\, K_S^0} &\approx \sin\phi_d\,, \\ C_{J/\psi\, K_S^0} &\approx 0\,. \end{aligned} \tag{2.53}$$

This direct measurement of the CKM angle β allows to probe the SM, while the measurement in the $B^0 \to D^+D^-$ decay is polluted by hadronic penguin effects. Still, the measurement of $\Delta\phi_{d,D^+D^-}$ can be an important input for the measurement of $\phi_s = -2\beta_s$ in the $B_s^0 \to D_s^+D_s^-$ decay [19]. This is done by exploiting the U-spin symmetry of the strong contributions in $B^0 \to D^+D^-$ and $B_s^0 \to D_s^+D_s^-$ decays, *i.e.* the decays are the same when exchanging all d quarks and s quarks. This leads to

$$a\mathrm{e}^{\mathrm{i}\theta} = a'\mathrm{e}^{\mathrm{i}\theta'}\,, \quad \mathcal{A} = \mathcal{A}'\,. \tag{2.54}$$

Here, $a'\mathrm{e}^{\mathrm{i}\theta'}$ and A' are the respective hadronic parameters of the $B_s^0 \to D_s^+D_s^-$ decay. As $a\mathrm{e}^{\mathrm{i}\theta}$ and $a'\mathrm{e}^{\mathrm{i}\theta'}$ are ratios of hadronic amplitudes, hadronic form factors and decay constants cancel within factorisation and factorisable U-spin corrections vanish [15, 32]. This allows to use a measurement of $\Delta\phi_{d,D^+D^-}$ via

$$\tan\Delta\phi_{d,D^+D^-} = \frac{a^2\sin 2\gamma - 2a\cos\theta\sin\gamma}{1 - 2a\cos\theta\cos\gamma + a^2\cos 2\gamma}\,, \tag{2.55}$$

to assess the shift

$$\tan\Delta\phi_{s,D_s^+D_s^-} = \frac{\epsilon^2 a'^2\sin 2\gamma + 2\epsilon a'\cos\theta'\sin\gamma}{1 + 2\epsilon a'\cos\theta'\cos\gamma + \epsilon^2 a'^2\cos 2\gamma}\,, \tag{2.56}$$

in the $B_s^0 \to D_s^+D_s^-$ decay. The factor

$$\epsilon = \frac{\lambda^2}{1-\lambda^2}\,, \tag{2.57}$$

leads to $\Delta\phi_{s,D_s^+D_s^-}$ being CKM suppressed.

Similar measurements can be done using the $B^0 \to D^{*\pm}D^\mp$ decay. Due to the fact that the final state is not an *CP* eigenstate, four amplitudes can be defined analogously to Eqs. (2.46) and (2.47):

$$\begin{aligned} A(B^0 \to D^{*+}D^-) &= V_{cd}V_{cb}^*\mathcal{A}(1 - a\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{\mathrm{i}\gamma}) \\ A(\overline{B}^0 \to D^{*+}D^-) &= V_{cb}V_{cd}^*\overline{\mathcal{A}}(1 - \overline{a}\mathrm{e}^{\mathrm{i}\overline{\theta}}\mathrm{e}^{-\mathrm{i}\gamma}) \\ A(B^0 \to D^{*-}D^+) &= V_{cd}V_{cb}^*\mathcal{A}^\dagger(1 - a^\dagger\mathrm{e}^{\mathrm{i}\theta^\dagger}\mathrm{e}^{\mathrm{i}\gamma}) \\ A(\overline{B}^0 \to D^{*-}D^+) &= V_{cb}V_{cd}^*\overline{\mathcal{A}^\dagger}(1 - \overline{a}^\dagger\mathrm{e}^{\mathrm{i}\overline{\theta^\dagger}}\mathrm{e}^{-\mathrm{i}\gamma})\,. \end{aligned} \tag{2.58}$$

These lead to four *CP* observables $S_{D^{*+}D^-}$, $S_{D^{*-}D^+}$, $C_{D^{*+}D^-}$, $C_{D^{*-}D^+}$ and the two effective mixing phases

$$\phi^{\text{eff}}_{d,D^{*\pm}D^{\mp}} = \phi_d + \Delta\phi_{d,D^*D} \pm \delta_{D^*D}\,, \tag{2.59}$$

which fulfil the relation in Eq. (2.52). Here, δ_{D^*D} is the relative strong phase between the $B^0 \to D^{*+}D^-$ and $B^0 \to D^{*-}D^+$ amplitudes. The four *CP* observables are oftentimes measured as

$$\begin{aligned} S_{D^*D} &= \frac{1}{2}(S_{D^{*+}D^-} + S_{D^{*-}D^+})\,, \qquad \Delta S_{D^*D} = \frac{1}{2}(S_{D^{*+}D^-} - S_{D^{*-}D^+})\,, \\ C_{D^*D} &= \frac{1}{2}(C_{D^{*+}D^-} + C_{D^{*-}D^+})\,, \qquad \Delta C_{D^*D} = \frac{1}{2}(C_{D^{*+}D^-} - C_{D^{*-}D^+})\,. \end{aligned} \tag{2.60}$$

An additional observable is the the time- and flavour-integrated charge asymmetry defined as

$$A_{D^*D} = \frac{\left(|A_{D^{*+}D^-}|^2 + |\overline{A}_{D^{*+}D^-}|^2\right) - \left(|A_{D^{*-}D^+}|^2 + |\overline{A}_{D^{*-}D^+}|^2\right)}{\left(|A_{D^{*+}D^-}|^2 + |\overline{A}_{D^{*+}D^-}|^2\right) + \left(|A_{D^{*-}D^+}|^2 + |\overline{A}_{D^{*-}D^+}|^2\right)}\,. \tag{2.61}$$

The observables C_{D^*D} and S_{D^*D} can be interpreted in the same way as for the $B^0 \to D^+D^-$ decay: C_{D^*D} is a measure of direct *CP* violation, which can only be caused by interference between tree and penguin effects, and S_{D^*D} measures interference *CP* violation. The parameter ΔC_{D^*D} states how flavour specific the decays are. For $\Delta C_{D^*D} = \pm 1$ each final state is only accessible by one flavour making interference *CP* violation impossible, while for $\Delta C_{D^*D} = \pm 0$ the sensitivity on interference *CP* violation is maximal. The relative strong phase δ_{D^*D} can be assessed with ΔS_{D^*D}. The last observable A_{D^*D} is another measure for direct *CP* violation.

Measurements of *CP* violation are also possible in $B \to D\overline{D}$ decays where both D mesons are excited, *i.e.* $B^0 \to D^{*+}D^{*-}$, allowing to measure $\phi^{\text{eff}}_{d,D^{*+}D^{*-}}$. This final state is a *CP* eigenstate so that there is no additional set of *CP* observables, but an angular analysis is required to measure the relative orbital momentum between both D^* mesons as the configuration can be *CP* even or *CP* odd. Finally, the SM can be tested by relating *CP* violation and branching ratios measurements of several additional $B \to D\overline{D}$ modes [19, 33].

3 The LHCb experiment

The LHCb experiment is one of the large experiments at the LHC. The focus lies on precise measurements of b and c decays. First, the LHC is introduced in Sec. 3.1 based on Ref. 66. The LHCb detector is introduced in Sec. 3.2 using Refs. 67 and 68. An overview of the LHCb software stack is given in Sec. 3.3. Finally, flavour tagging (FT) at the LHCb experiment is discussed in Sec. 3.4.

3.1 The LHC

The LHC is currently the most powerful particle accelerator and operated by the European Organization for Nuclear Research (CERN) near Geneva. The design centre-of-mass energy is 14 TeV with a design luminosity of $10^{34}\,\mathrm{cm}^{-2}s^{-1}$. Two proton beams are accelerated into opposite directions inside two rings with a circumference of roughly 27 km, which are located 50 to 175 m below ground. They are collided at interaction points inside the four large experiments ATLAS [69], ALICE [70], CMS [71] and LHCb [67]. The whole accelerator complex is shown in Fig. 3.1.

The protons pass through several accelerators: they start in a linear accelerator (LINAC2) before they traverse the BOOSTER, the proton synchrotron (PS) and the super-proton synchrotron (SPS). They are injected into the LHC ring with an energy of 450 GeV at two injection points left and right of the ATLAS experiment. Here, they are further accelerated until the final centre-of-mass energy is reached. The LHC operated at a centre-of-mass energy of 7 TeV in 2011, 8 TeV in 2012 and 13 TeV from 2015–2018. The period from 2011 to 2012 is commonly called Run 1 and the period from 2015 to 2018 is called Run 2. The protons are organised in over 2000 bunches per beam, with $\mathcal{O}(10^{11})$ protons each and a bunch crossing rate of 40 MHz. The LHC can also collide heavy ions or the beam pipe at the LHCb interaction point can be filled with various gases, emulating a fixed target experiment.

While the ATLAS and CMS experiments are general purpose experiments, the ALICE and LHCb experiments are more specialised: the ALICE experiment focusses on heavy ion collisions and the main purpose of the LHCb experiment is precision measurements, especially in decays of b and c hadrons.

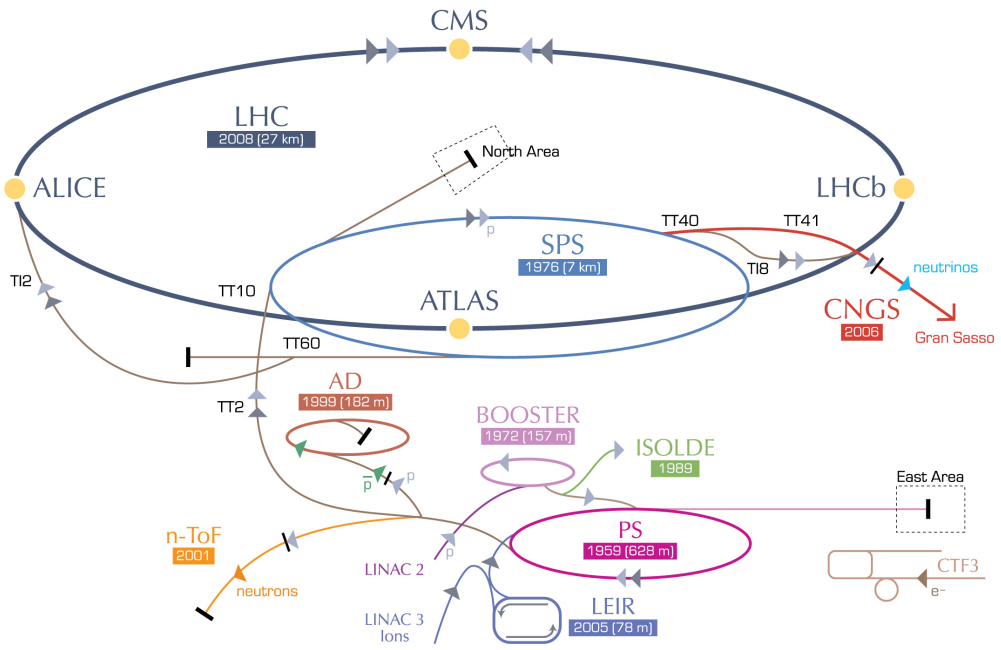

Figure 3.1: Overview of the CERN accelerator complex [72]. The protons are accelerated first in the LINAC2, then the BOOSTER, the PS and SPS to 450 GeV before they are injected into the LHC.

3.2 The LHCb detector

The name LHCb stands for LHC beauty experiment due to the focus on b-quark decays. Unlike the three other large experiments at the LHC, the LHCb does not cover the entire spatial angle. It is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$. The main production process of $b\bar{b}$ pairs in proton-proton collision is gluon-gluon fusion. This typically produces the b quarks with a large boost in forward or backward direction, as shown in Fig. 3.2. Roughly 25% of them can be reconstructed by the LHCb detector. The interaction point is shifted from the centre of the detector to use the maximum of the cavern space. To allow a optimal reconstruction and reduce resolution effects, the instantaneous luminosity at the LHCb interaction point is reduced to $4 \cdot 10^{32}\,\mathrm{cm}^{-2}s^{-1}$. This is achieved by not colliding the proton beams directly, but shifting them relative to each other. Another advantage is the possibility to level the luminosity, *i.e.* even though the maximal luminosity is decreasing during the period of a fill due to the loss of protons, the proton beams can be adjusted to deliver a functionally constant luminosity at the LHCb interaction point. The LHCb experiment collected data corresponding to an integrated luminosity of $1\,\mathrm{fb}^{-1}$ in the year 2011, $2\,\mathrm{fb}^{-1}$ in 2012 and $6\,\mathrm{fb}^{-1}$ in the years 2015–2018.

A cross-section of the LHCb detector and all subdetectors is shown in Fig. 3.3. During the proton injection and acceleration, the beams can become unstable and

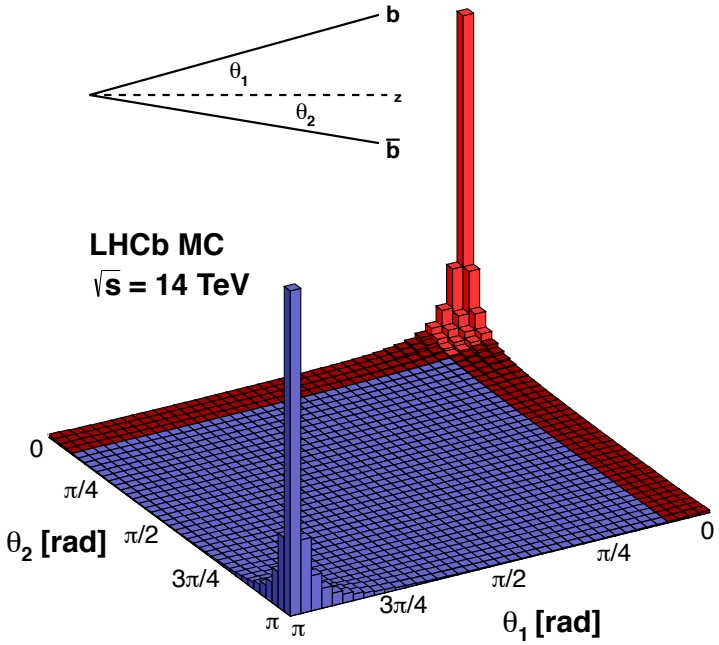

Figure 3.2: Production angles of simulated $b\bar{b}$ events at a centre-of-mass energy of 14 TeV [73]. The LHCb acceptance region is coloured in red. Roughly 25% of the $b\bar{b}$ pairs are inside the acceptance for centre-of-mass energies from 7 TeV to 14 TeV.

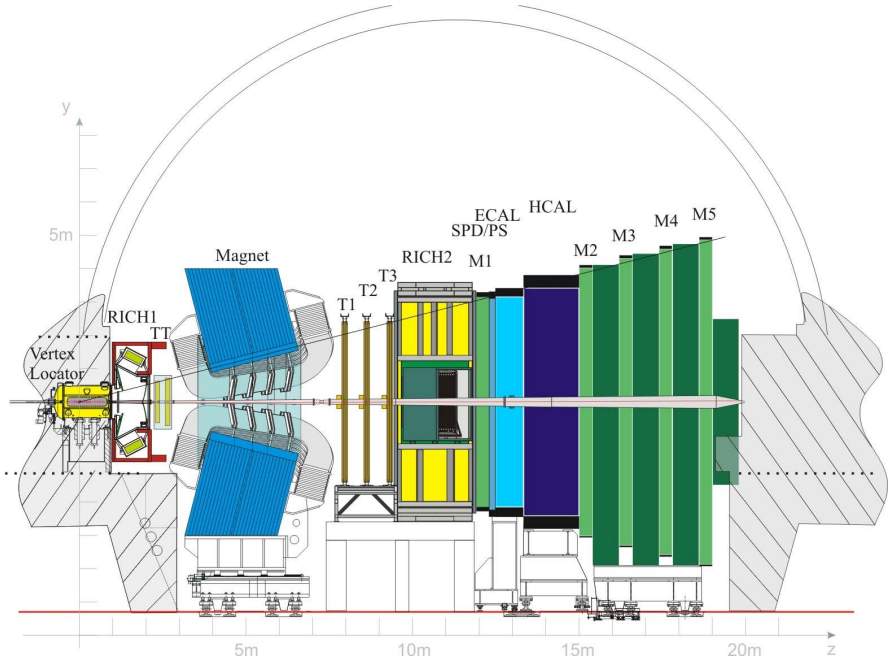

Figure 3.3: Cross-section of the LHCb detector and the subdetectors [67]. The interaction point lies inside of the VELO. The z-axis is parallel to the beam direction, the vertical y-axis is parallel to the magnetic field lines and the x-axis lies in the bending direction of the magnetic field.

damage the detector. The beam conditions monitor measures the flux of charged particles and initiates a beam dump if unstable beam conditions are detected [74]. It consists of two subdetectors, each equipped with eight diamond sensors, placed downstream and upstream of the interaction point.

3.2.1 Tracking

Tracks are reconstructed with information from the vertex locator (VELO) around the interaction point, the Tracker Turicensis (TT) upstream, and the tracking stations T1 to T3 downstream of the magnet.

Dipole magnet

The dipole magnet bends the trajectory of charged particles, allowing the determination of their momentum and charge. It is a conventional, non-superconducting magnet with saddle-shaped coils in a window-frame yoke with sloping poles. The integrated magnetic field is 4 Tm, pointing in vertical direction. The polarity is periodically inverted to allow studies of charge-dependent detections asymmetries.

VELO

The precise measurement of the decay times of B^0 decays is crucial for decay-time dependent measurements. The decay time is measured indirectly by determining the distance of the B^0 creation vertex, *i.e.* the primary vertex where the proton-proton collision happens (PV), and the secondary decay vertex (SV). The vertex locator (VELO) surrounds the interaction point and measures the exact positions of primary and secondary vertices. It consists of two halves with 21 silicon-half discs each, which are perpendicular to the beam axis. They use R and ϕ sensors to measure particle tracks. The VELO is located in a secondary vacuum outside of the beam pipe and its two halves can be opened. During standby, the halves are kept at a distance of roughly 4 cm and only once the proton beams are stable, the two halves are closed.

Tracking stations

The TT and the tracking stations T1 to T3 are positioned downstream of the VELO. The tracking stations T1 to T3 are divided into the Inner Tracker (IT), covering a cross-shape around the beam pipe and the Outer Tracker (OT), covering the outer part of the LHCb acceptance. The TT and the IT use silicon strip detectors. They consist of four layers, which are arranged in an x-u-v-x layout, where the second and third layer are rotated by $\pm 5^\circ$ compared to the first and last layer. This allows the measurement of both radial coordinates. The OT consists of

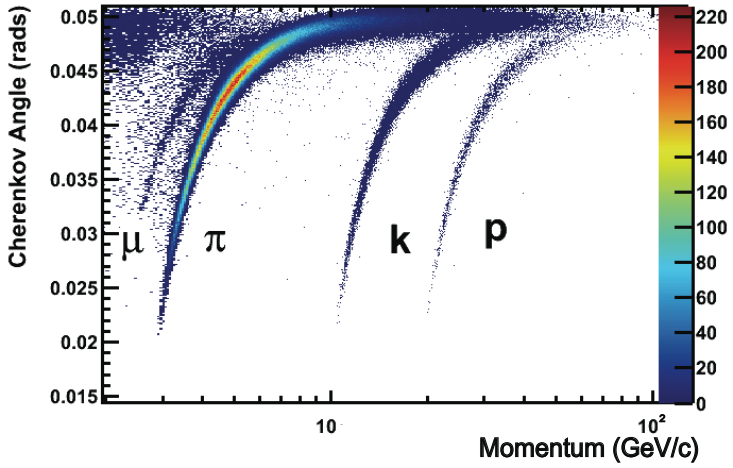

Figure 3.4: Reconstructed Cherenkov angle as a function of track momentum in the C_4F_{10} radiator of the RICH1 detector [76]. The pion, kaon and proton bands can be clearly separated. Even a distinct muon band can be observed, although the RICH detectors are primarily used to identify hadrons.

drift-time detectors, using the same x-u-v-x layout as the IT. The layers comprise straw tubes, which are filled with a mixture of 70% Ar, 28.5% CO_2 and 1.5% O_2 [75]. The TT covers the full detector acceptance. It is placed in front of the magnet to record low-momentum tracks that are curved outside of the acceptance by the magnetic field. The IT covers the inner part of the acceptance, where roughly 20% of the tracks pass through, with the OT covering the rest.

3.2.2 Particle identification

The LHCb detector has the capability to resolve separate tracks and identify the particle. The particle identification (PID) system comprises two ring-imaging Cherenkov detectors (RICH1 and RICH2) and the calorimeter system.

RICH detectors

The RICH detectors are essential for the particle identification at LHCb, especially to distinguish pions, kaons and protons. The Cherenkov effect can be used to identify particles, *i.e.* emission of photons by charged particles travelling through a medium faster than light. The photons are emitted in cones around the flight direction with the opening angle $\cos\theta = c/nv$, where v is the speed of the particle, c the speed of light in a vacuum, and n is the index of refraction in the medium. Combining the measured angles and momenta gives access to the particle mass, as shown in Fig. 3.4. The RICH1 detector is placed between the VELO and the TT.

It identifies particles with lower momenta in the range 2–40 GeV/c using C_4F_{10} as a radiator [76]. The RICH2 detector is located downstream of the three tracking stations, uses a CF_4 radiator and covers the momentum region 15–100 GeV/c. Both detectors use mirrors to reflect the incoming photons outside of the acceptance, focused on hybrid photo detectors.

Calorimeters

The calorimeter system identifies hadrons, electrons and photons. It also measures their energy and positions for the L0 trigger, see Sec. 3.2.3. The four subdetectors are the Scintillator Pad Detector (SPD), the preshower detectors (PS), the electromagnetic calorimeter and the hadronic calorimeter. SPD, PS and ECAL are segmented into three sections depending on the distance to the beam pipe. Similarly, the HCAL is segmented into two sections. The granularity of the active pad in each section grows with higher distances to the beam pipe. The SPD and PS are placed directly in front of the ECAL. They consist of two scintillator layers separated by a 15 mm thick lead plate. The radiation energy originating from electrons and photons in the plate is measured with the downstream scintillator. The calorimeters are sampling calorimeters built of alternating layers of scintillators and metal. Traversing particles shower in the metal layers and the energy deposition is measured with the scintillators. SPD, PS and ECAL are used to identify electron, photons and neutral pions. The energy of long-lived hadrons is measured in the HCAL.

Muon chambers

The muon-detecting system consists of the five stations M1 to M5, which increase in transverse size as the distance to the interaction point increases. They consist of multi-wire proportional chambers filled with a mixture of Ar, CO_2 and CF_4. The first station is placed in front of the calorimeters to increase the transverse-momentum resolution for the trigger. In front of each station M2 to M5, an 80 cm thick iron absorber is placed to slow the muons. Muons with momenta larger than 6 GeV/c pass through the whole detector.

3.2.3 Trigger

Even with the reduced luminosity at the LHCb interaction point, not all of the measured data can be saved due to limitations in storage space and bandwidth. Due to the small production cross-sections of b and c quarks [77, 78], most of the data are not interesting for the LHCb physics programme. The data rate is reduced by the LHCb trigger system from the initial bunch crossing rate of 40 MHz to 3 kHz, 5 kHz or 12.5 kHz for the year 2011, the year 2012 and Run 2,

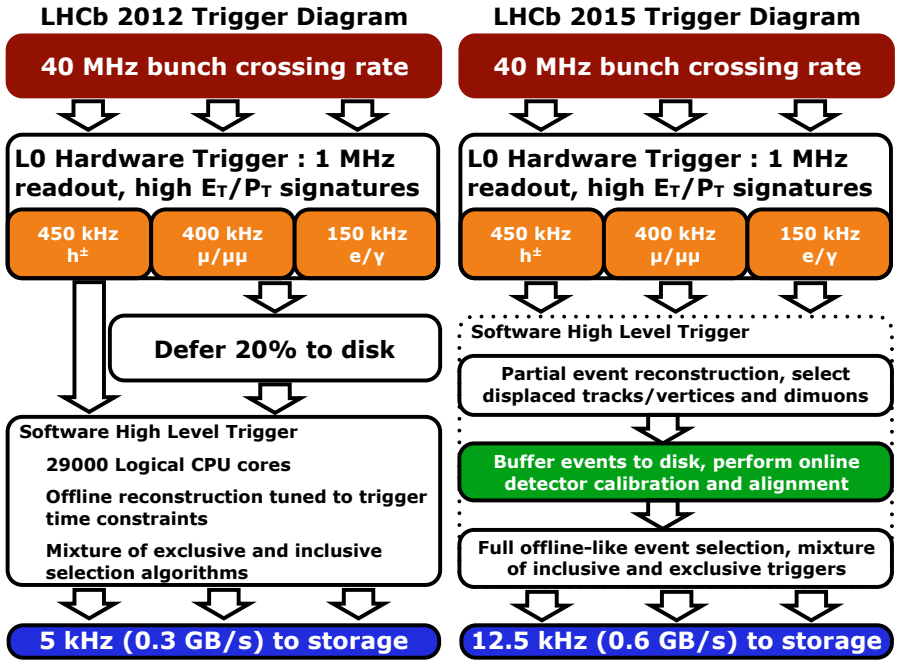

Figure 3.5: Trigger schemes for (left) 2012 and (right) Run 2 reducing the initial rate of 40 MHz to 5 kHz and 12.5 kHz, respectively [82].

respectively [79–81]. The trigger schemes are shown in Fig. 3.5. The LHCb trigger system in the first two Runs consists of two stages. A first-stage, hardware-based trigger (L0) and a software-level trigger (HLT).

L0

The L0 trigger operates at the bunch crossing rate of the LHC and reduces it to 1 MHz. It does not reconstruct the whole event, but uses specific information from the calorimeter and muon system. The trigger searches for hadrons, electrons and photons with high transverse-energy depositions in the calorimeters. Events with a muon or with muon pairs are retained if they have a high transverse momentum respectively high product of transverse momenta measured by the muon system. Additionally, events with too many tracks are rejected using information of the SPD. The L0 efficiency for hadronic decays, *e.g.* the decays investigated in this thesis, is roughly 60% [79–81].

HLT

The HLT trigger is a software-based trigger implemented in C++ consisting of the two stages HLT1 and HLT2. The HLT runs on the event filter farm, a computing cluster at CERN comprising several ten thousand CPU cores. The HLT1

performs a partial event reconstruction with information from the tracking system. For decays without leptons in the final state, the trigger decisions are taken based on the transverse momentum and displacement from the PV. The efficiency is around 90% for hadronic decays [79–81]. In the HLT2, a full event reconstruction can be performed. The trigger lines contributing in this analyses use inclusive selections based on multivariate algorithms. They search for two-, three-, or four-track secondary vertices using requirements on transverse momenta and displacements from the PV. Finally, all events that pass the HLT2 selection are saved on magnetic-tape. They are further selected with offline selections.

3.3 The LHCb software

The LHCb software is based on the GAUDI framework [83], a framework that is common to many particle physics experiments. It provides modules for the trigger (MOORE [84]), reconstruction, analysis and simulation.

3.3.1 Reconstruction

Tracking and particle identification is implemented in the BRUNEL [85] software package. Hits in the subdetectors of the tracking system are combined to form charged-particle tracks using pattern recognition. Matching hits are combined successively in the subdetectors. First, hits in a straight line in the VELO are matched with hits in the tracking stations T1 to T3 considering the magnetic field. Afterwards, matching hits in the TT are added to the track. Duplicate tracks happen when some detector hits are used in the reconstruction of two or more tracks. These duplicates are removed. A Kalman filter [86] calculates the track using the collection of hits and information from the magnetic field, multiple scattering and energy loss from interaction with the detector material. The resulting tracks are assigned to proto particles with an established charge and momentum.

Information from the PID system is used to calculate likelihoods for the tracks to originate from different particle types. The RICH detectors are used to identify charged long-lived hadrons, *i.e.* protons, charged pions and charged kaons. The tracks are extrapolated to the RICH detectors and the likelihoods for the particle masses are calculated using the known Cherenkov rings. The muon identification is very efficient and matches the established tracks to hits in the muon system. Finally, clusters in the calorimeter system are used to identify electrons and photons.

3.3.2 Preselection

Even after passing the trigger selection, the recorded LHCb data set is too large to be handled by individual analysts. The raw data is stored on magnetic tape and only a fraction, determined with a centralised preselection, is saved on a faster storage. This centralised preselection is called Stripping selection and is implemented in the DaVinci framework [87].

The proto particles in the raw data are combined to match the relevant decays including any intermediate states. This can be done with the LokiVertexFitter or the DecayTreeFitter (DTF) [88]. The former uses a bottom-up approach, combining final-state particles to the last intermediate state, then combining preceding intermediate states until the head of the decay chain is reached. The DTF performs a fit to the whole decay chain using a Kalman filter [86], which is more computationally intensive. The DTF exploits the knowledge of common vertices to obtain more accurate results. Additionally, the head of the decay chain can be forced to point to the PV and intermediate-resonance masses can be constrained. The decay-time variables used throughout this thesis use the PV constraint to improve the resolution. Similarly, the invariant B^0 masses use constraints on the masses of the intermediate resonances.

The preselection is organized in Stripping lines, which are sets of requirements to filter certain decay modes. Selection steps common to multiple lines only have to be performed once, decreasing the computing time. The requirements are typically very inclusive or contain at least multiple similar decays, *i.e.* the same set of requirements is used for the signal and control channels throughout this thesis.

3.3.3 Simulation

Data sets generated with Monte Carlo simulations are often used to verify the analysis procedure. They can be generated in much larger number than recorded data, minimising statistical uncertainties. The presented analyses use simulated data to calculate efficiencies, parametrise distributions, and train multivariate classifiers. Also, the knowledge of true generation values is used, *e.g.* to study resolution effects or validate fitting processes. The simulation conditions are constantly calibrated to ensure an accurate description and avoid discrepancies to recorded data. The generation of simulated data at LHCb uses Pythia [89] with a specific LHCb configuration [90] to generate proton-proton collisions. The decays are described by EvtGen [91] and final-state radiation is generated using PHOTOS [92]. The interaction of particles and the detector is simulated with the Geant4 framework [93] as described in Ref. [94]. The digitisation of the data is handled by the Boole [95] framework making the output equivalent to raw detector data. The further process is similar to the process for recorded data. The $B^0 \to D^+D^-$ analysis uses

an approach where the simulated underlying event is reused multiple times [96]. Only the signal decay and decays of other heavy particles are simulated every time, decreasing the computing time by a factor of 10 to 20.

3.4 Flavour tagging at LHCb

To measure *CP* violation in decays of neutral B^0 mesons, their production flavour has to be known. This is done using several different flavour-tagging algorithms, called taggers. Each algorithm provides a tag d, which defines the production flavour of the B^0 candidate, and a probability estimate of the tag to be wrong, the so-called mistag, η. The tagging is not perfectly efficient and it can happen that no tag can be assigned. A scheme of the taggers used for B^0 decays is given in Fig. 3.6. They can be divided in two types: opposite-side (OS) and same-side (SS) taggers. These are multivariate classifiers exploiting the topological information of the decay and kinematic properties of the tagging particle.

Opposite-side flavour tagging

The fact that b quarks are commonly produced in pairs by gluon-gluon fusion is exploited in the OS tagging. The second b quark is called the OS b quark. Several possible decay modes of the the OS b quark are the focus of the single OS taggers [97]. They are partially reconstructed by the OS algorithm, and through

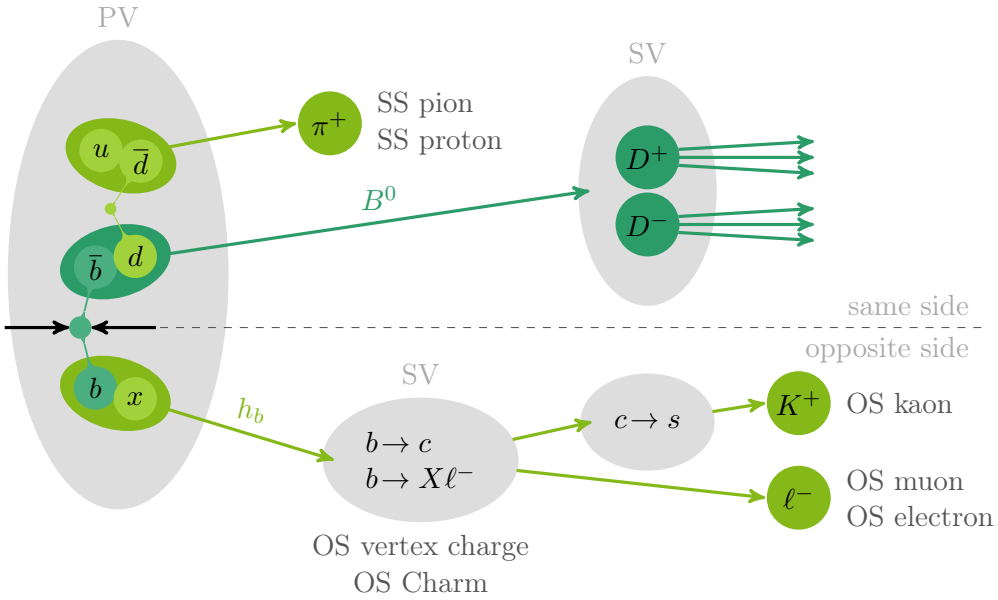

Figure 3.6: Flavour tagging algorithms used to tag B^0 mesons at the LHCb experiment. They are explained in the text.

the charges of the final-state particles the production flavour of the OS b quark is inferred. The signal b quark has to have the opposite production flavour. In case of semi-leptonic decays to muons or electrons, the charges of the leptons can directly be used to assign the tag. The OS kaon tagger exploits the charges of kaons from $b \to c \to s$ transitions. The OS vertex charge tagger reconstructs a secondary vertex and calculates the total charge by summing over the charges of the decay particles weighted with their transverse momentum. The OS charm tagger [98] was added after the Run 1 data-taking period ended and is not available for the respective data sets. It reconstructs $H_c \to K^-\pi^+X$ decays to infer the tag, where H_c is a charm hadron.

The performance of taggers can be estimated with their efficiency ε and their average mistag $\langle\omega\rangle$, via $\varepsilon_{\text{eff}} \approx \varepsilon(1-2\langle\omega\rangle)$. A more accurate calculation is introduced in Sec. 3.4.1. The most important OS tagger is the OS μ, with a tagging efficiency of about 11% and a very good average mistag of about 40%. The vertex charge is the next most performant tagger despite its high average mistag of almost 43%. The efficiency is very high at 23%. Next, is the OS kaon tagger with a similar mistag to the vertex charge tagger, but a tagging efficiency of 17%. The new OS charm has an average mistag of 41%, but an efficiency of only 6%. The OS e has the best average mistag at slightly less than 40%, but its efficiency is only about 2%. These values were determined with a data set of $B^0 \to D_s^+D^-$ decay recorded in Run 2 of the LHCb experiment. The values for Run 1 differ slightly.

Same-side flavour tagging

Like the $b\bar{b}$ pair, which is exploited by the OS taggers, the d in the B^0 meson is typically created as a $d\bar{d}$ pair, which is exploited by the SS taggers. Together with other quarks, the d quark forms hadrons, commonly protons or pions. Alternatively, the B^0 mesons can be produced as excited state and pions can originate from $B^{*+} \to B^0\pi^+$ decays. Using the charged pions and protons, the d quark flavour and thus the signal b quark flavour can directly be determined. The performance of both taggers is increased by upgrading them using multivariate approaches [99]. The tagging efficiencies in the presented in Run 2 are about 85% for the SS pion respectively 35% for the SS proton. Their average mistag is about 46%.

3.4.1 Flavour-tagging calibration and combination

The taggers return a tag and the mistag as the output. These parameters are only correct for the training data of the taggers. They have to be calibrated with data from flavour-specific decays to obtain accurate values, which are used in *CP*-violation analyses. Also, a combination of these values have to be performed, which is used in the analyses.

Performance

The tag decision d can be 1,-1 or 0 for a B^0 tag, a $\overline{B}^0$ tag, or no tag. The mistag probability estimate η lies between 0 and 0.5. A mistag of 0 means a perfect tag and a mistag of 0.5 is equivalent to no tag. The effect on the uncertainty of CP observables is the deciding performance parameter for the FT. This depends on the tagging efficiency ε as well as the mistag. The effective tagging efficiency ε_{eff} indirectly combines both:

$$\varepsilon_{\text{eff}} = \frac{1}{N}\sum_{i}^{N}(1-2\omega)^2\,. \tag{3.1}$$

This parameter is also called tagging power. It represents the loss of statistical power compared to a perfectly tagged data sample. The statistical uncertainty of CP observables is directly dependent on the tagging power via

$$\sigma_{S,C} = \frac{1}{\sqrt{\varepsilon_{\text{eff}}}}\,. \tag{3.2}$$

Calibration

There are often substantial differences between the data samples that are used to train the taggers and the data that is used to extract the CP observables. This leads to different mistag distributions, which have to be adjusted. The real mistag distribution of the signal channel can typically not be determined on the signal channel itself. This calibration needs a flavour-specific channel, as for these the B^0 flavour at the decay can be inferred from the final-state charge configuration. Using charged B^+ decays, which always decay flavour-specifically, the real mistag can be calculated by counting the correctly and falsely tagged decays. For neutral but flavour-specific B^0 decays, the oscillation of the B^0 meson has to be taken into account leading to a decay-time dependent probability of the B^0 flavour. The mistag estimate is used to obtain the true mistag probability $\omega(\eta)$ with a linear calibration according to

$$\omega(\eta) = p_0 + p_1(\eta - \langle\eta\rangle)\,, \tag{3.3}$$

where the average mistag $\langle\eta\rangle$ is subtracted from η to reduce the correlation of the two calibration parameters p_1 and p_2. The calibration is performed using the EPM tool [100]. Two sets of parameters can be defined, one for B^0 mesons and one for $\overline{B}^0$ mesons, leading to different true mistags for both flavours. The EPM tool returns their average and difference according to

$$p_i = \frac{p_i(B^0) + p_i(\overline{B}^0)}{2}\,, \qquad \Delta p_i = p_i(B^0) - p_i(\overline{B}^0)\,. \tag{3.4}$$

Example outputs of the EPM tool are shown in Fig. 3.7. The tool does not account for potential asymmetries in the production of the B^0 mesons, their detection or FT efficiency asymmetries.

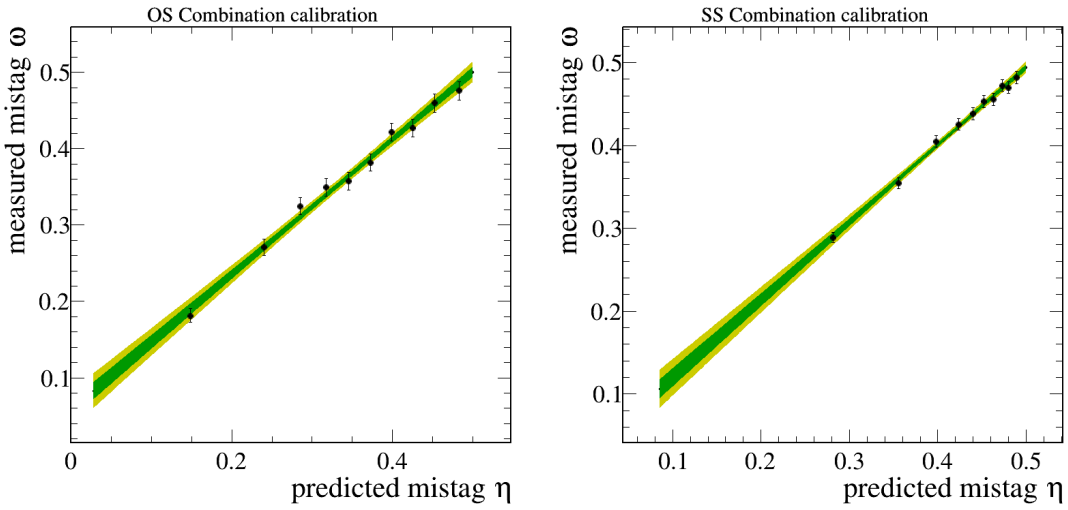

Figure 3.7: Examples of the (left) OS- and (right) SS-combination calibration using $B^0 \to D_s^+ D^-$ decays recorded in Run 2 by the LHCb detector. The measured mistag is plotted as a function of the predicted mistag. The calibration function is shown with (green) 1σ and (yellow) 2σ bands.

Combination

To profit from all of the several tagging algorithms, their outputs are combined. Two combinations are obtained: one combination of the SS taggers and one of the OS taggers. This is also done with the EPM tool. The mistag is used to calculate probabilities for a b or $\bar{b}$ quark according to one tagger:

$$p^i(b) = \begin{cases} 1-\omega & \text{if } d^i = b \\ \omega & \text{if } d^i = -b \\ 1/2 & \text{if } d^i = 0 \end{cases}, \qquad p^i(\bar{b}) = \begin{cases} \omega & \text{if } d^i = b \\ 1-\omega & \text{if } d^i = -b \\ 1/2 & \text{if } d^i = 0 \end{cases}. \tag{3.5}$$

These probabilities are used to obtain the combined probability of all taggers,

$$P(b) = \frac{\prod p^i(b)}{\prod p^i(b) + \prod p^i(\bar{b})}, \qquad P(\bar{b}) = 1 - P(b). \tag{3.6}$$

The tag decision and mistag are then assigned depending on whether the probability for a b- or $\bar{b}$-quark is higher:

$$(d, \omega) = \begin{cases} (\ 1\ , 1-P(b)) & \text{if } P(b) > P(\bar{b}) \\ (-1, \quad P(b)\ ) & \text{if } P(b) < P(\bar{b}) \\ (\ 0\ , \quad 1/2\ ) & \text{if } P(b) = P(\bar{b}) \end{cases}. \tag{3.7}$$

The resulting combination has to be calibrated again to describe the signal accurately. This is necessary even though the single taggers entering the combination are already calibrated because of different correlations between the taggers for different signal decays.

4 Experimental methods

A summary is given of the methods and tools used throughout this thesis. The maximum likelihood technique, a technique used to estimate parameters through fits, is described in Sec. 4.1. The sPlot technique is used to statistically separate different contributions in data samples and outlined in Sec. 4.2. Section 4.3 introduces boosted decisions trees (BDTs), multivariate tools to classify data.

4.1 The maximum likelihood technique

Parameters of a given function can be determined from a data set with the maximum likelihood technique [101]. The likelihood function of a probability density function (PDF) $f(x;a)$, with the underlying set of observables a, is defined as

$$L(a) = \prod_{i=1}^{n} f(x_i;a)\,, \tag{4.1}$$

where $x_1, x_2, \ldots, x_n$ are n measured values of the observable x. Thus, the higher the value of the likelihood function, the better the set of observables describes the measured values. Maximising $L(a)$ gives the best estimate of the underlying set of observables. Instead of maximising $L(a)$, the negative logarithm $-\ln L(a)$ is commonly minimised. Calculating the sum instead of a product decreases the computational effort. Also, most algorithms are focused on minimisation instead of maximisation. The likelihood function can be extended by multiplication with a Poisson term $\mathcal{P}(n;N)$. This accounts for the possibility to observe n events when N are expected. Uncertainties from observables that are known from other measurements can be included in the likelihood via Gaussian constraints. This is implemented by multiplying Gaussian PDFs $\mathcal{G}(a_i;\mu,\sigma)$ to the likelihood function. The mean μ and width σ of the function are the measured value and the uncertainty of the observable a_i. The fits presented in this thesis are maximum likelihood fits using the minimiser Minuit [102] and the ROOT [103] framework.

4.2 The sPlot technique

In particle-physics analyses it is common that there are still background candidates in a data sample, even after a thorough selection. A statistical subtraction of

remaining background candidates can be performed with the sPlot technique [104]. This is done by finding the parametrisation of all contributions in a discriminating variable. The discriminating variable is commonly the invariant mass of the reconstructed decay. The sWeights are defined as

$$ {}_s\mathcal{P}_n(x_i) = \frac{\sum_j V_{nj} f_j(x_i)}{\sum_j N_j f_j(x_i)} , \tag{4.2} $$

where the index j sums over all components PDFs f_j, and V_{nj} is the covariance of the yields N_n and N_j. The covariances are typically determined in a separate fit, where the shape parameters of the PDFs are fixed and only the yields are free. The sWeights are normalised so that their sum is equal to the respective yield. This approach is only valid if the investigated variables are uncorrelated to the discriminating variable.

Maximum likelihood fits to sWeighted data set potentially underestimate uncertainties of the fit results. This can be avoided by modifying the likelihood as described in Ref. 105. This modification of the likelihood is computationally expensive. In this thesis, a more naive approach is performed. The sWeights are multiplied with the correction factor

$$ c = \frac{\sum_i {}_s\mathcal{P}_n(x_i)}{\sum_i ({}_s\mathcal{P}_n(x_i))^2} , \tag{4.3} $$

i.e. the sum of the sWeights divided by the sum of their squares.

The resulting uncertainties are validated with bootstrapping studies [106]. The bootstrap method is a frequentist model-independent procedure, where new data sets are created from the original data set. Events are drawn from the original data set until the original number of events is reached. The same event can be drawn multiple times. This maintains correlations of the observables and allows to repeat the fits several times. The deviation in the fit results is used to estimate the statistical uncertainty.

4.3 Boosted decision trees

Multivariate methods are commonly used to process and classify data sets. They achieve higher separation powers than cut-based methods by exploiting the correlations of the input features.

Binary decision trees are relatively simple multivariate classifiers, which classify data as two categories. Figure 4.1 shows a schematic illustration of a binary decision tree classifying data as signal (S) or background (B). The tree consists of several nodes, where cut-based requirements are used to separate the data set.

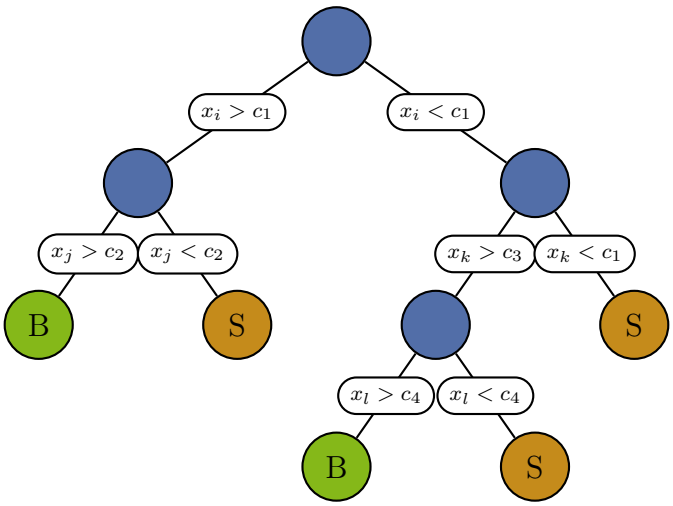

Figure 4.1: Schematic illustration of a decision tree with variables x_i to x_l and values c_1 to c_4.

The final nodes of the tree are reached once a stopping condition is fulfilled, *e.g.* the number of candidates is too small or too many subsequent requirements were performed. The data is classified as signal or background at the final nodes. The tree can be trained, *i.e.* the variables and requirements are changed to find an optimal configuration. This is done by giving access to data with labels for signal and background, defining a measure of the performance and finding the maximum of the measure by trying different variables and requirements.

The quality of the classification can be increased by training multiple trees, where the trees following the first are weighted to increase the separation power [107, 108]. This weighting is called boosting leading to the name BDT. The optimal parameters and requirements in the BDT are found with the help of a loss function $L(y, F(x))$. The BDT can be described as the function $F(x)$ that maps the variables x to the label y. The approximation of $F(x)$ that minimises the expected value of the loss function for all (y, x) values,

$$F = \arg\min_{F} E_{y,x} L(y, F(x)) \,, \tag{4.4}$$

gives the optimal BDT configuration [109]. Here, $E_{y,x}$ is the expectation value with respect to y and x. The loss functions used in this thesis are exponential loss

$$L(y, F(x)) = \mathrm{e}^{-yF(x)} \,, \tag{4.5}$$

which leads to so-called adaptive boosting [110], and binomial log-likelihood loss

$$L(y, F(x)) = \ln\left(1 + \mathrm{e}^{-2yF(x)}\right) , \tag{4.6}$$

called gradient boosting [109]. Solving Eq. (4.4) typically happens iteratively, starting with the first binary decision tree, but depends on the implementation. The presented analyses use the TMVA [111] and XGBOOST [112] frameworks.

5 Analysis strategy

This chapter discusses the general strategy of the analyses of the $B^0 \to D^{*\pm}D^{\mp}$ and $B^0 \to D^+D^-$ decays with the ultimate goal to measure the *CP* observables outlined in Sec. 2.5. Section 5.1 covers the selection of signal decays. Section 5.2 provides an overview of the extraction of the *CP* observables. The effects of FT calibration, detector resolution and efficiency, as well as other decay-time-dependent efficiencies on the measurement are discussed in detail.

5.1 Selection and mass fit

The first analysis step is selecting the recorded data by suppressing backgrounds to obtain a signal sample that is as clean as possible. The $B^0 \to D^{*\pm}D^{\mp}$ analysis uses the full data set that was recorded by LHCb corresponding to an integrated luminosity of $9\,\text{fb}^{-1}$. The $B^0 \to D^+D^-$ analysis uses data corresponding to $6\,\text{fb}^{-1}$ complementing the existing measurement in this channel [113]. Both decay modes are reconstructed using fully hadronic decay modes. Leptonic or semi-leptonic modes allow a more efficient selection, but typically include neutrinos, which worsen the decay-time resolution and complicate the unfolding of signal and background which relies on the reconstructed mass. The hadronic environment at the LHC has a much higher track multiplicity compared to electron-positron colliders and necessitates an effective suppression of backgrounds. Backgrounds are divided into two classes: combinatorial backgrounds, *i.e.* a combination of random tracks that do not originate from a common decay, and backgrounds from other decays, mainly misidentified and partially-reconstructed $B \to D\overline{D}$ decays or single-charm $B \to Dhhh$ decays.

Both analyses use very similar centralised preselections, where the main discrepancies lie in the different requirements for the $D^{*\pm}$ and $D^{\pm}$ meson. An additional preselection is applied that mainly constrains the reconstructed mass of the D mesons to further suppress combinatorial backgrounds. Single-charm backgrounds are rejected using the flight distance of the D^+ meson. Misidentified backgrounds are rejected with combinations of reconstructed mass and PID requirements. Partially-reconstructed backgrounds are not further suppressed. At this stage, the data samples are still dominated by combinatorial background. BDTs are used to bring their proportion to manageable levels.

Finally, the mass distribution is modelled to unfold the signal and the remaining backgrounds using the sPlot method, as described in Sec. 4.2. The shape of the signal decay and of remaining backgrounds from other decays are modelled using simulated data sets. The combinatorial background is known to have an exponential shape. Extended maximum likelihood fits to the mass distribution are performed, as discussed in Sec. 4.1, and sWeights are determined.

5.2 Measurement of the *CP* observables

The *CP* observables are extracted using a maximum likelihood fit to the decay-time distribution of the B^0 candidates. The construction and validation of the fitting framework for this fit is a time-intensive task. Here, the similarities of both analyses are exploited and the same framework is used for both measurements.

The master equations

The master equations of time evolution in Eqs. (2.29) and (2.32) are not sufficient to describe the decay-time distribution, because the effects from production and detection asymmetries, the FT, the decay-time resolution, and the decay-time-dependent efficiency have to be considered in the decay-time PDF. The master equations are simplified by assuming $\Delta\Gamma_d \approx 0$ and no indirect *CP* violation, *i.e.* $|q/p|^2 \approx 1$, which are both good assumptions within the current experimental precision [43]. The averages and differences of the *CP* observables according to Eq. (2.60) and the charge asymmetry from Eq. (2.61) are used to obtain

$$\begin{aligned}
\Gamma_{\text{theo}}(B^0(t) \to f) &\propto \mathrm{e}^{-\Gamma t}(1 + A_{\hat{f}}) \\
&\quad \left[1 + (C_{\hat{f}} + \Delta C_{\hat{f}}) \cos(\Delta m_d t) - (S_{\hat{f}} + \Delta S_{\hat{f}}) \sin(\Delta m_d t)\right], \\
\Gamma_{\text{theo}}(\overline{B}^0(t) \to f) &\propto \mathrm{e}^{-\Gamma t}(1 + A_{\hat{f}}) \\
&\quad \left[1 - (C_{\hat{f}} + \Delta C_{\hat{f}}) \cos(\Delta m_d t) + (S_{\hat{f}} + \Delta S_{\hat{f}}) \sin(\Delta m_d t)\right], \\
\Gamma_{\text{theo}}(B^0(t) \to \bar{f}) &\propto \mathrm{e}^{-\Gamma t}(1 - A_{\hat{f}}) \\
&\quad \left[1 + (C_{\hat{f}} - \Delta C_{\hat{f}}) \cos(\Delta m_d t) - (S_{\hat{f}} - \Delta S_{\hat{f}}) \sin(\Delta m_d t)\right], \\
\Gamma_{\text{theo}}(\overline{B}^0(t) \to \bar{f}) &\propto \mathrm{e}^{-\Gamma t}(1 - A_{\hat{f}}) \\
&\quad \left[1 - (C_{\hat{f}} - \Delta C_{\hat{f}}) \cos(\Delta m_d t) + (S_{\hat{f}} - \Delta S_{\hat{f}}) \sin(\Delta m_d t)\right].
\end{aligned} \tag{5.1}$$

Here, $\hat{f}$ refers to the final state and its *CP* conjugate: $\hat{f} = D^+D^-$ for the $B^0 \to D^+D^-$ decay and $\hat{f} = D^*D$ for the $B^0 \to D^{*\pm}D^{\mp}$ decay. These four

formulas can be reduced to the single formula

$$\mathcal{P}_{\text{theo}}(t,q,b) \propto \mathrm{e}^{-\Gamma t}(1+qA_{\hat{f}}) \\ \left[1+b(C_{\hat{f}}+q\Delta C_{\hat{f}})\cos(\Delta m_d t)-b(S_{\hat{f}}+q\Delta S_{\hat{f}})\sin(\Delta m_d t)\right], \tag{5.2}$$

where $b = \pm 1$ for B^0 and $\overline{B}^0$ decays, and $q = \pm 1$ for the final states f and $\overline{f}$, respectively. Identical constant factors are neglected. For *CP*-eigenstate final states, *e.g.* in the $B^0 \to D^+D^-$ decay, $q = 0$ applies.

External Asymmetries

The detection and production asymmetries are defined as

$$A_{\text{det}}^{\hat{f}} = \frac{\varepsilon(f)-\varepsilon(\overline{f})}{\varepsilon(f)+\varepsilon(\overline{f})}, \qquad A_{\text{prod}} = \frac{\sigma(\overline{B}^0)-\sigma(B^0)}{\sigma(\overline{B}^0)+\sigma(B^0)}, \tag{5.3}$$

where ε is the efficiency and σ is the production cross-section. Including the asymmetries yields the PDF

$$\mathcal{P}_{\text{asym}}(t,q,b) \propto e^{-\Gamma t}(1-bA_{\text{prod}})(1+qA_{\text{raw}}^{\hat{f}}) \\ \left[1+b(C_{\hat{f}}+q\Delta C_{\hat{f}})\cos(\Delta m_d t)-b(S_{\hat{f}}+q\Delta S_{\hat{f}})\sin(\Delta m_d t)\right]. \tag{5.4}$$

The detection asymmetry cannot be disentangled from the intrinsic charge asymmetry of the decay and only the parameter

$$A_{\text{raw}}^{\hat{f}} = A_{\text{det}}^{\hat{f}} + A_{\hat{f}}, \tag{5.5}$$

can be measured. The detection asymmetry includes asymmetries in the detection of particles, but also asymmetries due to reconstruction and particle identification inefficiencies. It is not relevant for the $B^0 \to D^+D^-$ decay due to the symmetric final state. The final state of the $B^0 \to D^{*\pm}D^{\mp}$ decay is also charge symmetric. Any detection asymmetry can only arise due to the different kinematic distributions of the final-state particles, which lead to different efficiencies. Thus, $A_{\text{det}}^{D^*D}$ is expected to be small. The approximation

$$A_{\text{det}}^{D^*D} \approx A_{\text{det}}^{D^*} - A_{\text{det}}^{D} = \frac{\varepsilon(D^{*-})-\varepsilon(D^{*+})}{\varepsilon(D^{*-})+\varepsilon(D^{*+})} - \frac{\varepsilon(D^-)-\varepsilon(D^+)}{\varepsilon(D^-)+\varepsilon(D^+)}, \tag{5.6}$$

shows that $A_{\text{det}}^{D^*D}$ can be determined by measuring the detection asymmetries of $D^{*\pm}$ and $D^{\pm}$ decays. The asymmetries are measured using $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ decays, where the $D^{\pm}$ mesons are created in the PV. As the detection asymmetry depends on the kinematic distribution, the momenta and transverse momenta of the final state particles are adjusted to match the distributions in $B^0 \to D^{*\pm}D^{\mp}$ decays.

The production asymmetry can be taken from external measurements [114] or be determined with decay-time-dependent fits to flavour-specific control channels. This is done using the nominal PDF shown below in Eq. (5.13) and fixing the values of the *CP* observables of the control channels to their known values. As there is no current measurement of A_{prod} for Run 2 when the $B^0 \to D^{*\pm}D^{\mp}$ analysis is carried out, the value is determined using the flavour-specific $B^0 \to D_s^+ D^-$ and $B^0 \to D_s^+ D^{*-}$ decays. The $B^0 \to D^+ D^-$ analysis reuses the determined value.

Flavour tagging

The FT parameters have to be parametrised and integrated into the decay-time PDF. The FT calibration and combination uses $B^0 \to D_s^+ D^-$ and $B^0 \to D_s^+ D^{*-}$ decays, or only $B^0 \to D_s^+ D^-$ decays, as calibration channels for the $B^0 \to D^{*\pm}D^{\mp}$ and $B^0 \to D^+ D^-$ analyses, respectively. The single taggers are calibrated and combined to an SS and an OS combination. The combinations are calibrated once more to account for correlation between the taggers. The SS and OS combinations are not combined to an overall combination, but are combined in the PDF. This allows to constrain the calibration parameters of the SS and OS combination in the fit instead of fixing them, which is done to incorporate the uncertainties on the calibration parameters in the statistical uncertainty of the *CP* observables.

The probability to obtain the tag d^i of the tagger i for the true tag b is given as

$$P^i(d^i|b) = \begin{cases} \varepsilon_b^i(1-\omega_b^i) & \text{if } d^i = b \\ \varepsilon_b^i \omega_b^i & \text{if } d^i = -b \\ (1-\varepsilon_b^i) & \text{if } d^i = 0 \end{cases}, \tag{5.7}$$

where ε_b^i is the tagging efficiency and ω_b^i is the calibrated mistag and depends implicitly on the the estimated mistag η^i. In the following, $\varepsilon_b^i = \varepsilon^i$ for B^0 mesons, $\varepsilon_b^i = \overline{\varepsilon}^i$ for $\overline{B}^0$ mesons, and identical definitions for ω_b^i are used. The combined probability of both taggers is then defined as

$$P(d|b) = P^{\text{OS}}(d^{\text{OS}}|b) P^{\text{SS}}(d^{\text{SS}}|b)\,. \tag{5.8}$$

The PDF that combines the FT response with Eq. (5.4) is given by

$$\begin{aligned} \mathcal{P}(t, q, \vec{d}|\vec{\eta}) &= P(d|B^0)\mathcal{P}_{\text{asym}}(t, q, B^0) + P(d|\overline{B}^0)\mathcal{P}_{\text{asym}}(t, q, \overline{B}^0) \\ &= \mathrm{e}^{-\Gamma t}(1 + q A_{\text{raw}}^{\hat{f}}) \Big[(\Delta^+ - A_{\text{prod}}\Delta^-) \\ &\quad + (\Delta^- - A_{\text{prod}}\Delta^+)(C_{\hat{f}} + q\Delta C_{\hat{f}}) \cos(\Delta m_d t) \\ &\quad - (\Delta^- - A_{\text{prod}}\Delta^+)(S_{\hat{f}} + q\Delta S_{\hat{f}}) \sin(\Delta m_d t) \Big]\,. \end{aligned} \tag{5.9}$$

The expressions $\Delta^{\pm}$ absorb all FT parameters. They are defined as

$$\begin{aligned}\Delta^{\pm} = &\frac{1}{4}\varepsilon^{\mathrm{OS}}[1+d^{\mathrm{OS}}(1-2\omega^{\mathrm{OS}})]\varepsilon^{\mathrm{SS}}[1+d^{\mathrm{SS}}(1-2\omega^{\mathrm{SS}})] \\ &\pm\frac{1}{4}\overline{\varepsilon}^{\mathrm{OS}}[1+d^{\mathrm{OS}}(1-2\overline{\omega}^{\mathrm{OS}})]\overline{\varepsilon}^{\mathrm{SS}}[1+d^{\mathrm{SS}}(1-2\overline{\omega}^{\mathrm{SS}})]\,,\end{aligned} \tag{5.10}$$

if both tagger combinations provide a tag,

$$\begin{aligned}\Delta^{\pm} = &\frac{1}{2}\varepsilon^{\mathrm{OS}}[1+d^{\mathrm{OS}}(1-2\omega^{\mathrm{OS}})](1-\varepsilon^{\mathrm{SS}}) \\ &\pm\frac{1}{2}\overline{\varepsilon}^{\mathrm{OS}}[1+d^{\mathrm{OS}}(1-2\overline{\omega}^{\mathrm{OS}})](1-\overline{\varepsilon}^{\mathrm{SS}})\,,\end{aligned} \tag{5.11}$$

if only the OS combination provides a tag, and

$$\Delta^{\pm} = (1-\varepsilon^{\mathrm{OS}})(1-\varepsilon^{\mathrm{SS}}) \pm (1-\overline{\varepsilon}^{\mathrm{OS}})(1-\overline{\varepsilon}^{\mathrm{SS}})\,, \tag{5.12}$$

if no combination provides a tag. They fourth case, in which only the SS combination provides a tag, is described by switching SS and OS in Eq. (5.11).

Decay-time resolution and decay-time-dependent efficiency

The effects of the decay-time resolution and the decay-time-dependent efficiency are included in the final PDF

$$P(t,q,\vec{d}|\vec{\eta}) = \varepsilon(t)\int \mathcal{P}(t',q,\vec{d}|\vec{\eta})R(t-t')\mathrm{d}t'\,. \tag{5.13}$$

Here, $\varepsilon(t)$ is the decay-time-dependent efficiency, $R(t')$ is the decay-time resolution function and t' is the true decay time. The momentum and vertex resolution of the detector is accounted for by convolving Eq. (5.9) with the resolution function. This resolution comes from the finite precision of the momentum and vertex measurements of the detector. It smears out the measured decay times, reducing the amplitudes of the oscillation and affecting the precision of the measurement of the *CP* observables. For B^0 decays, the oscillation frequency $\Delta m_d = 0.5065 \pm 0.0019\,\mathrm{ps}^{-1}$ [43] is much larger than the detector resolution, which is roughly 60 fs in the investigated decays. Thus, the effect of the resolution is limited. The resolution function is modelled using the sum of three Gaussian functions $\mathcal{G}$ with the same mean μ, but different widths σ_i, and fractions g_i:

$$R(t-t') = \sum_i g_i \mathcal{G}(t-t';\mu,\sigma_i)\,. \tag{5.14}$$

The parameters are determined using simulated data samples where the true decay time is known. For B_s^0 decays or decays with much more statistical power, a more precise description of the resolution is needed. This is commonly done using recorded candidates where the final state is produced in the PV, estimates of

the decay-time uncertainty, which are provided by the vertex algorithms, or a combination of both.

The decay-time-dependent efficiency is caused by the reconstruction and selection. Most of the background suppressed in the selection stems directly from the PV, leading to large selection inefficiencies especially at low decay times. The efficiency is described using the sum of n cubic B-splines b_i with the coefficients v_i:

$$\varepsilon(t) = \sum_{i=1}^{n} v_i b_i(t)\,. \tag{5.15}$$

B-splines are defined by their $n+2$ interval boundaries, called knot positions. Fits to the decay-time distributions of simulated data are used to find a set of knot positions. These are fixed during the extraction of the *CP* observables, while the coefficients are left free as their correlation to the *CP* observables is negligible.

Finally, Eq. (5.13) is used to extract the *CP* observables with a weighted maximum likelihood fit to the decay-time distribution. The fit is carried out simultaneously for all years and final states. Due to larger differences between the final states, and between Run 1 and Run 2 in the $B^0 \to D^{*\pm}D^{\mp}$ analysis, some parameters are allowed to take different values in the respective category. This is not necessary for the $B^0 \to D^+D^-$ analysis, as the final states are more similar and only Run 2 data is used.

Systematic uncertainties

Several cross-checks are carried out and systematic uncertainties are quantified, mainly using pseudo-experiments. In these pseudo-experiments, specific distributions are generated according to predefined PDFs, which takes much less computing power than generating fully-simulated samples. The mass and decay-time distributions are generated using PDFs, which use results from the fits to recorded data. They may be altered to investigate the systematic uncertainties that result from the changes. The fit also depends on the mistag distributions, which are generated according to histograms, and inputs for the parameters that are Gaussian-constrained, which are are generated according to the constraint applied [115]. The nominal procedure is then performed with the data samples: a mass fit is carried out to extract the sWeights and a fit to the decay-time distribution is used to extract the *CP* observables. This is repeated several times. Analysing the fit results allows the effects on the *CP* observables to be quantified. The difference between the mean of the fit results and the input parameters is assigned as the systematic uncertainty. The distribution of the pull, *i.e.* the difference between the fit result and input value divided by the statistical uncertainty from the fit, can show biases or an over- or underestimation of the statistical uncertainty. In both analyses the systematic uncertainty is expected to be much smaller than the statistical uncertainty.

6 Measurement of *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays

The measurement of the five *CP* observables S_{D^*D}, C_{D^*D}, ΔS_{D^*D}, ΔC_{D^*D} and A_{D^*D} is presented in this chapter. The analysis has been published in Ref. [34]. Previous measurements were conducted by the BaBar and Belle collaborations [116, 117].

The analysis uses the full LHCb data set. The decay is reconstructed in two decay modes of the D^0 meson that is produced in the D^{*+} decay.

The analysis was performed in collaboration between the LHCb groups at CERN and from Dortmund, Ferrara and Milan. The author worked on the selection and mass fit of the $D^0 \to K^-\pi^+$ mode with the exception of finding the best requirement on the BDT response, the decay-time fit with the exception of the determination of the instrumental asymmetries, and various cross-checks and systematic effects. The focus lies on the parts the author worked on while the remaining parts are summarised briefly. The contributions from Nicoletta Belloli [118], Marta Calvi, Basem Khanji, Margarete Schellenberg [119], Giulia Tellarini and Stefania Vecchi are described to give a coherent overview of the analysis.

The signal selection is described in Sec. 6.1 and the following mass fit is described in Sec. 6.2. The decay time description and the fit to extract the *CP* observables is described in Sec. 6.3. Section 6.4 gives an overview of the systematic uncertainties and Sec. 6.5 presents the results.

6.1 Selection

The LHCb data sets recorded in Run 1 and Run 2 are used in this analysis. The Run 1 data corresponds to an integrated luminosity of $1\,\text{fb}^{-1}$ recorded at a centre-of-mass energy of 7 TeV in the year 2011, and $2\,\text{fb}^{-1}$ recorded at 8 TeV in 2012. The Run 2 data was taken in 2015–2018 at 13 TeV and corresponds to an integrated luminosity of $6\,\text{fb}^{-1}$. The $B^0 \to D^{*+}D^-$ decay is reconstructed via the $D^- \to K^+\pi^-\pi^-$ and the $D^{*+} \to D^0\pi^+$ decays. The D^0 meson is reconstructed in two different modes: the $D^0 \to K^-\pi^+$ decay and the $D^0 \to K^-\pi^+\pi^-\pi^+$ decay. These decays are chosen as they are the most common modes that are easy to reconstruct by the LHCb detector, *i.e.* they do not contain neutral particles in the

final state. While the $D^0 \to K^-\pi^+\pi^-\pi^+$ decay is roughly twice as common as the $D^0 \to K^-\pi^+$ decay [41], the latter gives the higher yield due to better tracking and selection efficiencies. The selection of the data set containing $D^0 \to K^-\pi^+$ decays is discussed in detail. The selection of the complimentary data set containing $D^0 \to K^-\pi^+\pi^-\pi^+$ is discussed in more detail in Ref. 118.

6.1.1 Preselection

In the central preselection, the data is filtered to bring it to a manageable size and suppress obvious backgrounds. The number of tracks is restricted to allow a separation of tracks. Events are retained if they are triggered by a topological trigger. The triggers search for decay topologies with vertices with a significant displacement from the PV with two, three or four tracks. One charged particle is required to have a large transverse momentum p_{T} and cannot be consistent with originating from a PV.

The track quality of the final-state particles is required to be good. The probability of the track to be a ghost track [120] is used to avoid such tracks, *i.e.* a random combination of unrelated detector hits or detector hits originating from multiple real tracks. Observables provided by the PID system are used to ensure high likelihoods of correct particle identifications of the final-state particles. The smallest distance of all tracks that form a vertex is required to be small.

The final-state particles and the D mesons are required to have high momenta to ensure the decay of a heavy particle. The invariant mass of the D^{*+} (D^0) meson is required to lie inside a mass window $\pm 50\,\mathrm{MeV}$ ($\pm 100\,\mathrm{MeV}$) around the known mass [41]. The reconstructed D^+ mass is required to lie in a mass window that starts $100\,\mathrm{MeV}$ below the known D^+ mass and ends $100\,\mathrm{MeV}$ above the known D_s^+ mass. This allows to reconstruct the $B^0 \to D_s^+ D^{*-}$ decay at the same time. The invariant mass of the B^0 meson is required to lie in the mass window 4750–6000 MeV, which allows to use the sidebands for background studies.

The qualities of the D and B^0 decay vertices are required to be good. The distance of closest approach is the smallest distance between all tracks that form a vertex. It is required to be small to guarantee that the tracks originate from the same vertex. The impact parameter, IP, is the smallest distance between a track and a vertex. The χ^2_{IP} is defined as the difference of the vertex-fit χ^2 of a given vertex reconstructed with and without a given track. The IP and the χ^2_{IP} are used to suppress tracks that emerge directly in the PV, as B^0 mesons typically fly a significant distance before decaying. The B^0 candidate has is required to have an χ^2_{IP} that is consistent with originating from a PV. If multiple PVs are reconstructed in the same event, the PV is associated to the decay for which the B^0 candidate has the lowest χ^2_{IP}. The cosine of the angle between the momentum of a particle and the distance from its production and decay vertex is also called direction angle

Table 6.1: Requirements used in the preselection.

Variable	Requirement
difference of invariant D^{*+} and D^0 masses	$\lvert m(D^{*+}) - m(D^0)\rvert < 50\,\mathrm{MeV}/c^2$
invariant D^0 mass	$\lvert m(D^0) - m_{D^0,\mathrm{PDG}}\rvert < 40\,\mathrm{MeV}/c^2$
invariant D^+ mass	$\lvert m(D^+) - m_{D^+,\mathrm{PDG}}\rvert < 50\,\mathrm{MeV}/c^2$

or DIRA. It is used to guarantee that the D mesons travel in forward direction and to guarantee that the B^0 candidate momentum direction and vertex distance have the same direction.

Due to the use of trigger requirements in the centralised preselection, the preselection efficiency cannot be easily separated from the trigger and reconstruction efficiencies. The combined efficiency, calculated with simulated data, is determined to be around 0.7% in Run 1 and 1.0% in Run 2. The higher efficiency in Run 2 mainly comes from better trigger efficiencies and from the higher centre-of-mass energy, which leads to higher momenta in Run 2.

An additional selection on top of the centralised preselections is applied using requirements on the invariant D masses. The requirements are listed in Table 6.1. The mass windows of the D^0 and the D^+ mesons are tightened to reduce combinatorial background. A requirement on the mass difference of the D^{*+} and the D^0 meson reduces the background significantly. The known mass difference of the two mesons is only slightly larger than the pion mass. The pion only gets minimal momentum in the D^{*+} decay and is thus called slow pion. The mass difference cancels resolution effects of the D^0 final-state particles, leading to a sharp peak around the nominal mass difference for signal decays. This requirement already reduces the combinatorial background by around 75%, which is calculated on the upper-mass sideband, *i.e.* recorded data with $m_{B^0} > 5400\,\mathrm{MeV}$.

The combination of the additional preselection requirements suppresses more than 90% of the combinatorial background, while keeping more than 95% of the signal candidates. The distributions of the used masses and the requirements are shown in Fig. 6.1.

6.1.2 Vetoes

Vetoes are requirements that are used to suppress backgrounds from specific decays mimicking the signal. Here all vetoes suppress backgrounds from decays that include misidentified particles. By recalculating the invariant masses under the assumption of alternative mass hypotheses for the final state particles, background contributions become visible.

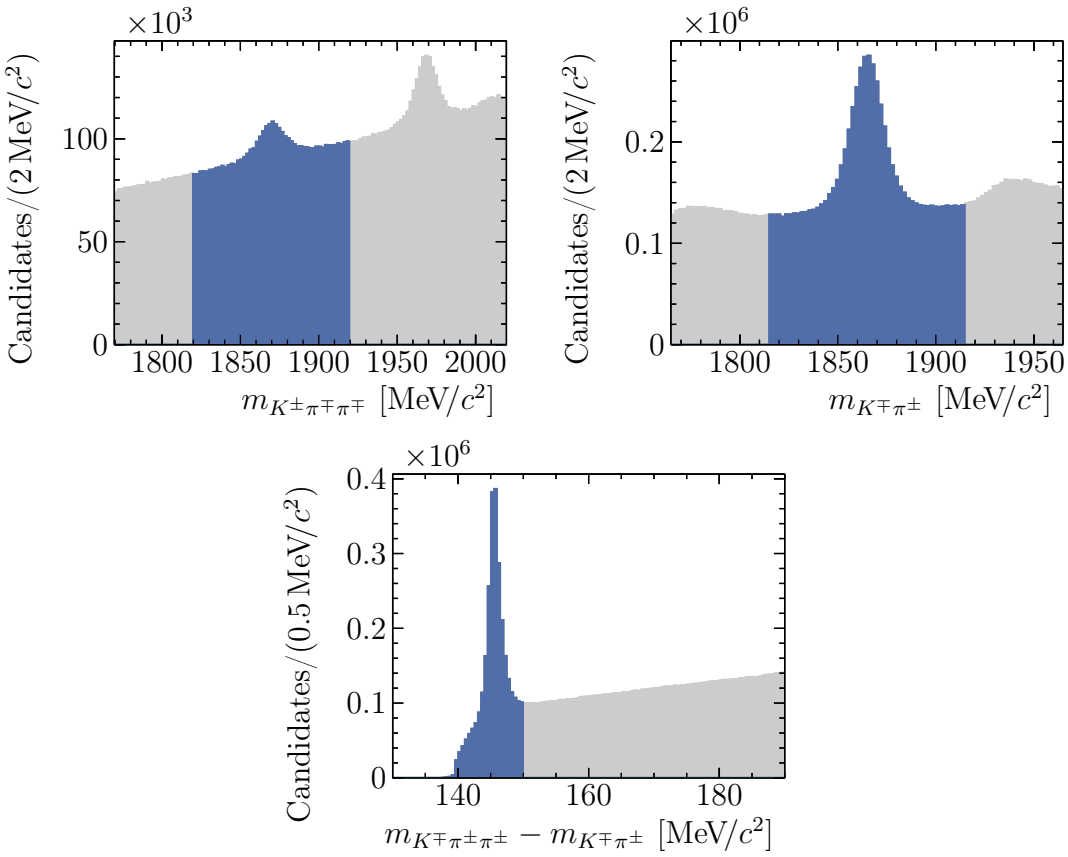

Figure 6.1: Effect of the preselection requirements and the resulting distributions: the invariant (top left) $D^{\mp}$, (top right) D^0 masses, and (bottom) the difference of the $D^{*\pm}$ and D^0 masses. The grey data is rejected by the selection requirement and the blue data is retained. Only data recorded in Run 2 is shown. The distributions in Run 1 look similar.

The backgrounds can be reduced by requirements on PID variables. In this case, ProbNN variables are used, which give the probability for the particle to be a pion, kaon or proton. They are provided by neural networks and show differences between simulated and recorded data. The differences are adjusted with variable transformations based on kernel density estimations in the four dimensions ProbNN, p_{T}, η and track multiplicity [121]. This transformation keeps the correlation of the adjusted ProbNN values with the rest of the event. This is not necessarily given for other adjustment techniques and allows multivariate analyses to exploit this correlation. The ratio

$$\mathrm{ProbNN}_{p_1 p_2} = \frac{\mathrm{ProbNN}_{p_1}}{\mathrm{ProbNN}_{p_1} + \mathrm{ProbNN}_{p_2}}\,, \tag{6.1}$$

increases the separation power between the mass hypothesis p_1 and p_2 and further decreases discrepancies between simulated and recorded data.

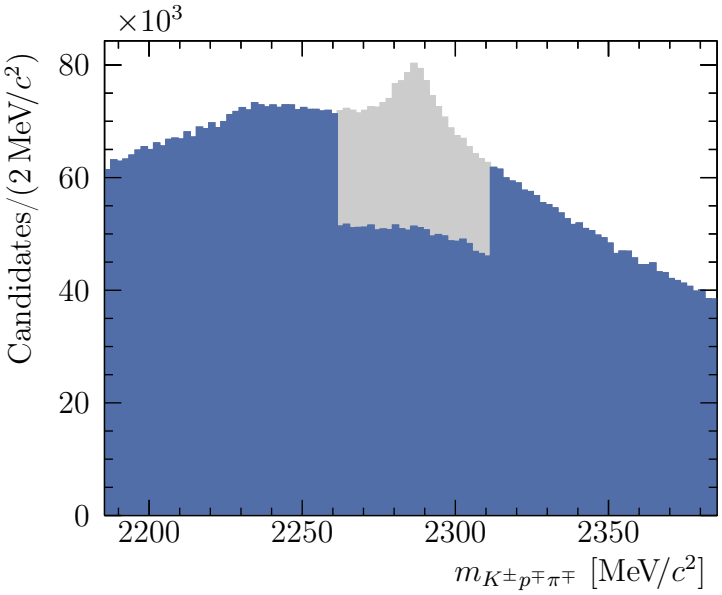

Figure 6.2: Invariant $K^{\pm}p^{\mp}\pi^{\mp}$ mass distribution. The grey data is rejected by the veto requirement and the blue data is retained. The contribution at the known Λ_c^+ mass is suppressed. Only data recorded in Run 2 is shown. The distributions in Run 1 look similar. The preselection is not applied for this plot to enhance the Λ_c^+ contribution.

All backgrounds from misidentification involve the D^+ meson, *i.e.* other decays are identified as the $D^+ \to K^-\pi^+\pi^+$ decay. The $D^{*+} \to D^0\pi^+$ candidates are very clean due to the unique topology of the decay and the requirement on the difference of the D^0 and D^{*+} masses.

Λ_c^+ veto

The $\Lambda_c^+ \to K^-p\pi^+$ decay can be reconstructed as a D^+ candidate if the proton is misidentified as a pion. The main source of this background is the $\Lambda_b^0 \to \Lambda_c^+ D^{*-}$ decay, which has a similar topology to the signal decay. Assigning the proton mass to the pion from the D^+ decay with the higher p_{T} shows a background contribution in the invariant $K^-p\pi^+$ mass around the known Λ_c^+ mass, as can be seen in Fig. 6.2. Performing the same procedure with the pion with the lower p_{T} shows no visible background contribution. The background is suppressed by rejecting candidates if the ratio $\mathrm{ProbNN}_{\pi p}$ is lower than 0.7 and the $K^-p\pi^+$ mass lies in the mass window within $\pm 25\,\mathrm{MeV}$ around the known Λ_c^+ mass. This veto reduces more than 95% of the background while keeping more than 96% of the signal candidates. The efficiencies are calculated with simulated data.

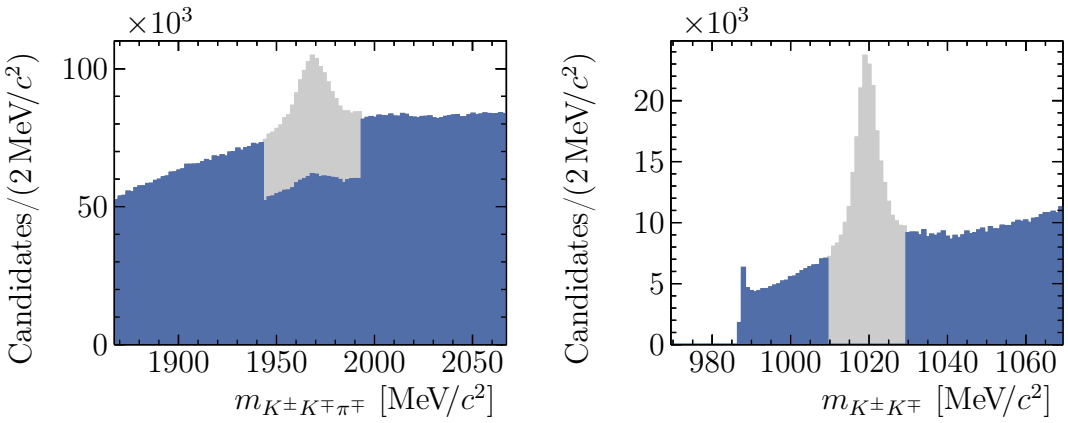

Figure 6.3: Invariant (left) $K^{\pm}K^{\mp}\pi^{\mp}$ and (right) $K^{\pm}K^{\mp}$ mass distributions. The grey data is rejected by the veto requirements and the blue data is retained. The misidentified backgrounds at the known D_s^+ and ϕ masses are reduced, but not completely suppressed. Only data recorded in Run 2 is shown. The distributions in Run 1 look similar. The preselection is not applied for this plot to enhance the background contributions.

D_s^+ veto

Similar backgrounds occur in the $D_s^+ \to K^-K^+\pi^+$ decay when the positive kaon is misidentified as a pion and the candidate is reconstructed as a D^+ decay. The main contributor to these is the $B^0 \to D_s^+ D^{*-}$ decay. A background contribution becomes visible when applying the kaon mass hypothesis to the pion from the D^+ decay with the higher p_{T} and calculating the invariant mass of the final state particles. Candidates are rejected if $\mathrm{ProbNN}_k > \mathrm{ProbNN}_\pi$ and the recalculated invariant mass is compatible with the known D_s^+ mass within $\pm 25\,\mathrm{MeV}$.

One possible intermediate resonance of the D_s^+ decay is the $\phi \to K^-K^+$ decay. Applying the kaon mass to the pion with the higher p_{T} and calculating the K^-K^+ mass shows a narrow background contribution around the known ϕ mass. This allows to suppress additional background by rejecting all candidates that are compatible with the known ϕ mass within $\pm 10\,\mathrm{MeV}$.

The plots of the mass distributions and the vetoes are shown in Fig. 6.3. The vetoes suppress 75% and 82% of the background while retaining 96% and 94% of the signal in Run 1 and Run 2, respectively. The background contribution is still significant and has to be modelled in the mass fit. Applying stronger PID requirements leads to drastically reduced signal efficiencies.

$B^0 \to D^{*-}\pi^-\pi^+\pi^+$ veto

The single-charm $B^0 \to D^{*-}\pi^-\pi^+\pi^+$ decay can be reconstructed as the signal decay, if the negative pion is misidentified as a kaon and the three hadrons are

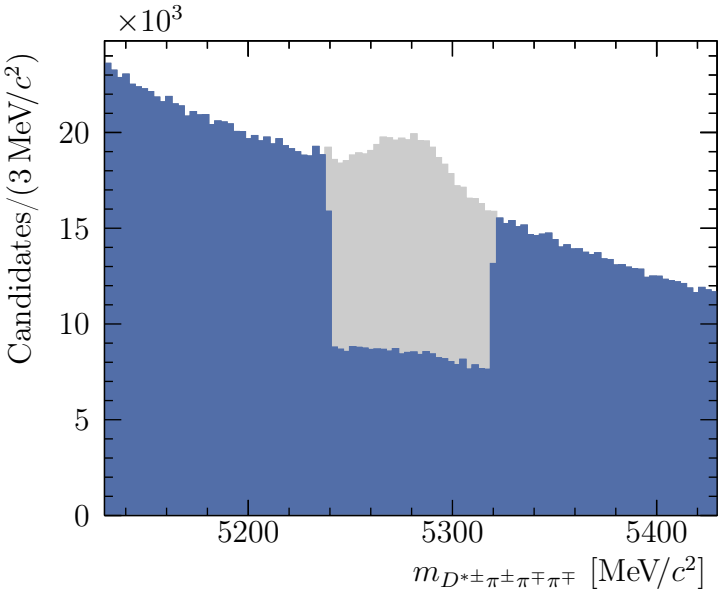

Figure 6.4: Invariant $D^{*\pm}\pi^{\pm}\pi^{\mp}\pi^{\mp}$ mass distribution. The grey data is rejected by the veto requirement and the blue data is retained. The misidentified background at the known B^0 mass is suppressed. Only data recorded in Run 2 is shown. The distributions in Run 1 look similar. The preselection is not applied for this plot to enhance the single-charm contribution.

falsely assigned to a D^+ decay. This decay is suppressed by rejecting candidates where the invariant $D^{*-}\pi^-\pi^+\pi^+$ lies in the mass window within $\pm 40\,\text{MeV}$ around the known B^0 mass, if one of two conditions is fulfilled: the ratio $\text{ProbNN}_{\pi K}$ is lower than 0.3 or the χ^2 of the flight distance (FD) of the D^+ meson is lower than 4. The first condition ensures a correct particle identification and the second ensures that the D^+ flight distance is significant, *i.e.* its three daughter hadrons are not produced directly in the B^0 decay vertex. The invariant $D^{*-}\pi^-\pi^+\pi^+$ mass and the effect of the veto are shown in Fig. 6.4. The veto keeps more than 99% of the signal candidates.

6.1.3 Multivariate analysis

At this stage, the data samples are still dominated by combinatorial background. This is reduced with a BDT based on the TMVA framework [111] using adaptive boosting [110]. The BDT training is performed separately for Run 1 and Run 2, but uses the same hyperparameters, *i.e.* parameters that change the general BDT architecture. Simulated data is used as the signal proxy and the upper-mass sideband of recorded data is used as the background proxy. The previous selection steps are applied on both data sets. A two-folding method [122] is applied to

Table 6.2: List of input features used in the BDT training. They are sorted by the approximate importance, top to bottom, then left to right, and explained in the text. The mothers of the final state particles are given in the subscript. The pion with the higher and lower p_{T} from the D^- decay are given as π_1^- and π_2^-, respectively.

Input features		
$t/\sigma_t(D^-)$	$\log \chi^2_{\mathrm{DTF}}(B^0)$	$p_{\mathrm{T}}(D^-)$
$\mathrm{ProbNN}_{\pi K}(K^+_{D^-})$	$t/\sigma_t(D^0)$	$\cos\angle(D^-, \pi_2^-)$
$\log \chi^2_{\mathrm{IP}}(B^0)$	$\cos\angle(D^-, \pi_1^-)$	$\mathrm{ProbNN}_{\pi K}(\pi_2^-)$
$\log \chi^2_{\mathrm{FD}}(D^-)$	$\cos\angle(D^-, K^+)$	$\mathrm{ProbNN}_{\pi K}(\pi_1^-)$
$p_{\mathrm{T}}(\pi_2^-)$	$p_{\mathrm{T}}(D^{*+})$	$\mathrm{ProbNN}_{\pi K}(\pi^+_{D^{*+}})$
$\log \chi^2_{\mathrm{FD}}(D^0)$	$\cos\angle(D^0, \pi^+)$	$\log \chi^2_{\mathrm{IP}}(D^0)$
$m(D^{*+}) - m(D^0)$	$p_{\mathrm{T}}(B^0)$	$p_{\mathrm{T}}(\pi_1^-)$
$p_{\mathrm{T}}(K^+_{D^-})$	$\cos\angle(B^0, D^{*+})$	$\mathrm{ProbNN}_{\pi K}(\pi^+_{D^0})$
$\mathrm{ProbNN}_{\pi K}(K^-_{D^0})$	$\log \chi^2_{\mathrm{IP}}(D^-)$	

avoid overfitting, *i.e.* the training on statistical fluctuation in data samples. The data is split randomly in two evenly sized parts, two separate BDTs are trained and applied to the other half.

The input features are found iteratively. In the first step, a large number of input features is added to the BDT, which are commonly used in decays with long-lived intermediate states. The BDT training provides a list of importances of the features. The least important features are removed from the BDT until the receiver operating characteristic (ROC) score starts to decrease. The ROC score is the area under the ROC curve, where the true positive rate is plotted as a function of the false positive rate. The remaining features are listed by their approximate importance in Table 6.2. The input features include the decay-time significance t/σ_t of the long-lived D mesons, *i.e.* the decay time divided by the decay-time uncertainty, the adjusted ProbNN ratios of the final-state particles, IP and flight-distance variables χ^2_{IP} and χ^2_{FD}, transversal momenta, and the cosine of the angles between mother particles M and their daughters D, $\cos\angle(M, D)$. The difference of the D^{*+} and D^0 masses, $m(D^{*+}) - m(D^0)$, is included and improves the BDT even though there is a requirement on this variable in the preselection. The goodness of the fit to the whole decay chain χ^2_{DTF} is the final input feature. Some variables are transformed, *i.e.* their logarithm is used. This is due to the fact that the TMVA algorithm scans the input features in bins. By taking the logarithm, low-density bins are avoided. The agreement between the distributions of the BDT features in simulated and background-subtracted data is verified. The background-subtracted data is obtained with the procedure described in Sec. 6.2. The agreement is crucial for the correct calculation of efficiencies on simulated data. The calculation of efficiencies is not necessary for decay-time-dependent analyses of *CP* violation.

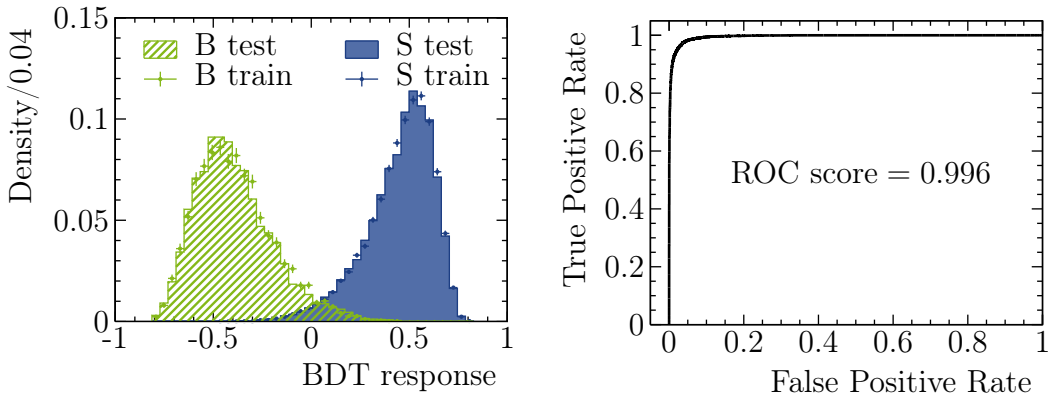

Figure 6.5: Distribution of the (green) background and (blue) signal distributions of the BDT response is shown on the left. The (shaded areas) test and (points) training distributions are compatible, which is an indicator that there is no overfitting. The ROC curve is shown on the right.

Still, disagreements between simulated and recorded data lead to a worse BDT performance on recorded data.

The hyperparameters of the BDT are also determined iteratively. They are changed until the maximum of the ROC score is found. Each BDT comprises 800 trees, the depth is limited to a maximum of three and at each node 3% of the data have to be retained. Each feature is scanned in 40 bins, to find the optimal value. The requirement with the highest GiniIndex [123] is chosen. The GiniIndex is defined by $p(1-p)$, where p is the purity, *i.e.* the ratio of signal events over all events. The learning rate is set to 0.1, which defines the boost weight of subsequent decision trees. The upper-mass sideband has a limited amount of entries, which makes the BDT prone for overfitting. To mitigate this, each tree of the BDT gets a random subset which amounts to 80% of the training data. Each tree is then trained with ten random features of the feature set. The output distributions of the BDT and the ROC curve are shown in Fig. 6.5 for one fold.

Next the optimal requirement on the BDT response must be found. The procedure is only briefly summarised here and more details can be found in Ref. 119. A figure of merit is developed that yields the BDT requirement with the best sensitivity to the effective phase ϕ_d^{eff}. The procedure is based on Ref. 124, with modifications to include sWeights ${}_s\mathcal{P}$. The sensitivity is not only given by the signal yield and purity, but also by additional characteristics of the data sample. The additional properties only affect the optimal BDT requirement due to correlations to the BDT response. The sensitivity is directly dependent on the effective tagging efficiency, which is given by the tagging decisions d and the calibrated mistags ω. Small decay-time-error estimates σ_t should increase the sensitivity, though this effect is almost negligible for B^0 decays due to the low oscillation frequency. The decay-

time t influences the sensitivity to ϕ_d^{eff}. This can be explained by the fact that the CP observables are amplitudes of oscillating terms which depend on t. Thus, the sensitivity is highest at the extrema of the oscillations. The figure of merit is given by

$$\text{FOM} = \frac{\sum_i {}_s\mathcal{P}_i}{\sum_i {}_s\mathcal{P}_i^2} \sum_i {}_s\mathcal{P}_i (1-2\omega_i)^2 \mathrm{e}^{-(\Delta m_d \sigma_{ti})^2} X_i \,, \tag{6.2}$$

with

$$X_i = \left[\frac{2 d_i |\lambda| s}{1 + |\lambda|^2 + d_i (1-2\omega_i) \mathrm{e}^{-(\Delta m_d \sigma_{ti})^2/2} \left(-2|\lambda| s \sin \phi_d^{\text{eff}} - (1-|\lambda|^2) c\right)} \right]^2 , \tag{6.3}$$

and $s = \sin(\Delta m_d t_i)$ and $c = \cos(\Delta m_d t_i)$. The CP parameter λ and ϕ_d^{eff} are fixed to the world averages [43]. The effective signal size $(\sum_i w_i)^2 / \sum_i w_i^2$ is the dominant term when finding the optimal requirement. The tagging dilution is included via the factor $(1-2\omega_i)^2$, the decay-time dilution is included via $\mathrm{e}^{-(\Delta m_d \sigma_{ti})^2}$ and X_i includes the dependence on the decay time. The best requirement is found by scanning different BDT requirements and performing a fit to the mass distribution for each requirement to extract the sWeights and calculate the figure of merit. The maximum of the figure of merit corresponds gives the best requirement on the BDT response. Using only the effective signal size leads only to slightly changed requirements and reduces the expected sensitivity by less than 5%. The BDT requirement has a signal efficiency of about 92% and suppressed more than 98% of the upper-mass sideband.

6.1.4 Multiple candidates

In about 2% for Run 1 and 4% for Run 2 of all events, more than one signal candidate is found. The higher track multiplicity in Run 2 causes the higher number of multiple candidates in Run 2. The multiple candidates typically share several final state tracks so that only one of them can be a real signal decay. To avoid any bias, the multiple candidates are rejected randomly so that only one candidate per event remains.

6.2 Mass fit

The sPlot method is used to subtract background contributions using the reconstructed mass of the B^0 candidate as the discriminating variable. For this, a fit to the mass distribution has to be performed. The fit range is chosen as 5000–5600 MeV. This gives a good handle on contributions from other decays, which can be found at lower masses, and the combinatorial background, which

is the only contribution at higher masses. At this stage, the data sample contains signal $B^0 \to D^{*\pm}D^{\mp}$ decays, similar $B_s^0 \to D^{*\pm}D^{\mp}$ decays of the B_s^0 meson, misidentified $B^0 \to D_s^+ D^{*-}$ decays, partially-reconstructed $B^0 \to D^{*+}D^{*-}$ and $B_s^0 \to D^{*+}D^{*-}$ decays, and combinatorial background. No contribution from misidentified $B_s^0 \to D_s^+ D^{*-}$ decays is observed. This decay is highly suppressed compared to the similar B^0 decay. The extended PDF describing the mass distribution is given as

$$\begin{aligned} N\mathcal{P} = N_{B^0\to D^{*\pm}D^{\mp}}\mathcal{P}_{B^0\to D^{*\pm}D^{\mp}} + N_{B_s^0\to D^{*\pm}D^{\mp}}\mathcal{P}_{B_s^0\to D^{*\pm}D^{\mp}} + N_{B^0\to D_s^+D^{*-}}\mathcal{P}_{B^0\to D_s^+D^{*-}} \\ + N_{B^0\to D^{*+}D^{*-}}\mathcal{P}_{B^0\to D^{*+}D^{*-}} + N_{B_s^0\to D^{*+}D^{*-}}\mathcal{P}_{B_s^0\to D^{*+}D^{*-}} + N_{\text{Comb}}\mathcal{P}_{\text{Comb}}\,, \end{aligned} \tag{6.4}$$

where N is the number of candidates and $\mathcal{P}$ the respective PDF. The fits are performed separately for Run 1 and Run 2. The parametrisation of the separate contributions is first developed on simulated data, as described in Sec. 6.2.1, and then the fit to the recorded data is applied, as described in Sec. 6.2.2. Only the procedure for the $D^0 \to K^-\pi^+$ sample is discussed in detail. An overview of the very similar procedure for the $D^0 \to K^-\pi^+\pi^-\pi^+$ sample can be found in Ref. 118.

6.2.1 Determination of the mass shapes

Maximum likelihood fits to simulated data are used to find parametrisations of the contributions from signal and other decays. All samples are reconstructed as the signal $B^0 \to D^{*\pm}D^{\mp}$ decay and the selection is applied. The combinatorial background is known to have an exponentially decreasing distribution, which is free in the fit to data.

$B^0 \to D^{*\pm}D^{\mp}$ and $B_s^0 \to D^{*\pm}D^{\mp}$ decays

The sum of three Crystal Ball ($\mathcal{CB}$) functions [125] is used to parametrise the signal $B^0 \to D^{*\pm}D^{\mp}$ decay:

$$\mathcal{P}_{B^0\to D^{*\pm}D^{\mp}}(m) = \sum_i^3 g_i \mathcal{CB}_i(m;\mu,\sigma_i,\alpha_i,n_i)\,. \tag{6.5}$$

The Crystal Ball function has a Gaussian-like core with the mean μ and the width σ. Below the threshold $(m-\mu)\sigma < -\alpha$, it follows a power law function with the exponent $-n$. The sign of α establishes the direction of the power-law tail: the tail goes to lower masses if the tail parameter α is positive or to higher masses if α is

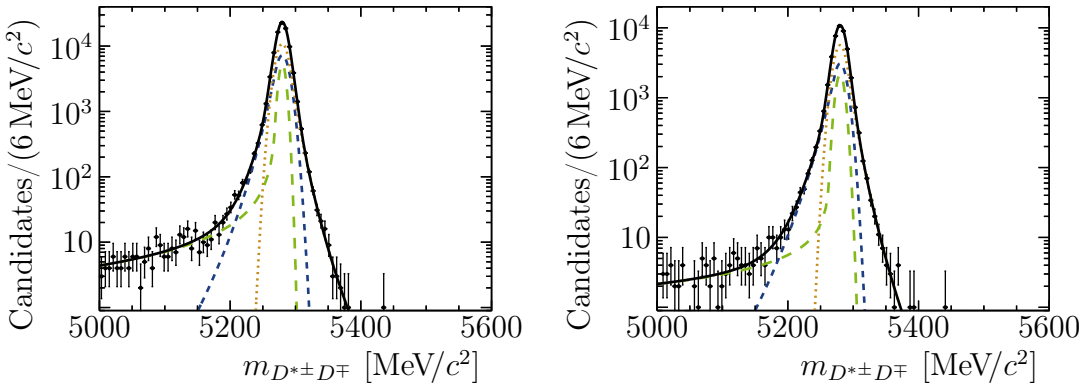

Figure 6.6: Mass distributions and PDF projections of simulated signal decays for (left) Run 1 and (right) Run 2. The projection of the total PDF is shown in solid black while the projections of the three Crystal-Ball components are shown in dashed blue, long-dashed green, and dotted yellow.

negative. The three Crystal Ball functions share the same mean. The fractions g_i are defined recursively, *i.e.*

$$g_i = f_i \cdot \prod_{j=1}^{i-1}(1 - f_j)\,, \tag{6.6}$$

with $f_3 = 1$ and $g_3 = (1-g_1)(1-g_2)$. Two of the tail parameters are negative and one is positive. The larger tail to lower masses is caused by final-state radiation which is not reconstructed. The resulting PDFs from a fit to simulated data and the mass distributions are shown in Fig. 6.6. The resulting parameters are fixed in the fit to recorded data, except for the mean, which is free. The shape of the $B_s^0 \to D^{*\pm}D^{\mp}$ is found to be compatible and the parameters are fixed to the same values. The mean is fixed to the known mass difference $\mu_{B_s^0} - \mu_{B^0} = (87.23 \pm 0.24)\,\text{MeV}$ [41].

$B^0 \to D_s^+ D^{*-}$ decays

Due to the misidentification in $B^0 \to D_s^+ D^{*-}$ decays, the mass distribution is shifted to lower masses and the shape is visibly different from the signal shape. The mass is parametrised with the sum of two Crystal Ball functions. The shapes are compatible between Run 1 and Run 2 within their uncertainties. A combined set of parameters is determined to reduce the statistical uncertainty. The mean and the exponent are the same for both Crystal Ball functions. The PDF resulting from the fit and the simulated data are shown in Fig. 6.7. All parameters are fixed to the resulting values the fit to recorded data, with the exception of the mean.

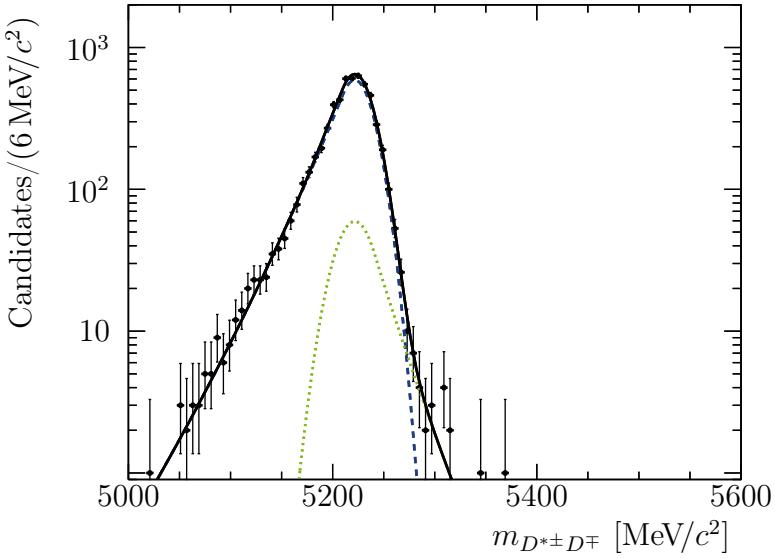

Figure 6.7: Mass distributions and PDF projections of simulated misidentified $B^0 \to D_s^+ D^{*-}$ decays. The projection of the total PDF is shown in solid black while the projections of the two Crystal-Ball components are shown in dashed blue, and dotted green.

$B^0 \to D^{*+} D^{*-}$ and $B_s^0 \to D^{*+} D^{*-}$ decays

The $B^0 \to D^{*+} D^{*-}$ and $B_s^0 \to D^{*+} D^{*-}$ decays can be reconstructed as signal if one of the $D^{*\pm}$ mesons performs the $D^{*+} \to D^+ \pi^0$ or the $D^{*+} \to D^+ \gamma$ decay. The $D^{*+} \to D^+ \pi^0$ decay has a branching fraction of $(30.7 \pm 0.5)\%$ while the branching fraction of the $D^{*+} \to D^+ \gamma$ decay is $(1.6 \pm 0.4)\%$ [41]. If the neutral pion or the photon are not reconstructed, the remaining final state is identical to the signal final state. The $B^0 \to D^{*+} D^{*-}$, $D^{*+} \to D^+ \pi^0$ and $D^{*+} \to D^+ \gamma$ modes are two-body decays, so that the absolute momenta are unambiguously defined. Thus, the reconstructed mass only depends on the helicity angle of the unreconstructed particle with respect to the boost of the D^{*+} meson in the B^0 rest frame. The distribution of the helicity angles depends on the polarisation in the $B^0 \to D^{*+} D^{*-}$ decay, the spin-parity of the unreconstructed π^0 or γ particle and the fact that the D^{*+} hadron is a vector meson.

If the pion is missing, decays with longitudinal polarisation show two peaks and decays with transverse polarisation show a single broad peak. The smaller contribution from missing photons leads to even broader distributions. In 55% of all cases the $D^{*\pm}$ mesons in the $B^0 \to D^{*+} D^{*-}$ decay are longitudinally polarised assuming factorisation [126]. The PDF is the sum of the two different contribu-

tions:

$$\mathcal{P}_{B^0 \to D^{*+}D^{*-}}(m) = g_{D^{*+}D^{*-}}\mathcal{P}_{\text{longitudinal}}(m) + (1 - g_{D^{*+}D^{*-}})\mathcal{P}_{\text{transverse}}(m)\,. \quad (6.7)$$

The longitudinal contribution is modelled as the sum of four Gaussian functions

$$\mathcal{P}_{\text{longitudinal}}(m) = \sum_i^4 g_i \mathcal{G}(m; \mu_i, \sigma_i)\,, \quad (6.8)$$

where the first two describe the broader shape from the missing photon and have a shared mean, and the last two describe the double-peak structure and have a shared width. The fractions g_i are defined iteratively as described before. The transverse shape is modelled as

$$\begin{aligned}\mathcal{P}_{\text{transverse}}(m) = {} & g_1 \mathcal{U}(m) \circledast \mathcal{G}_1(m; \mu = 0, \sigma_1) \\ & + (1 - g_1)[g_3 \mathcal{G}_3(m; \mu_3, \sigma_3) + (1 - g_3)\mathcal{G}_4(m; \mu_4, \sigma_4)]\,.\end{aligned} \quad (6.9)$$

The contribution from the missing photon is modelled as the convolution of a Gaussian function and a uniform distribution $\mathcal{U}$, the range of which is determined from the kinematics of the decay. The broader peak is modelled by two Gaussian functions sharing the same width. The shapes for Run 1 and Run 2 are found to be compatible. Only a combined set of parameters is determined. Fits to the mass distributions of simulated $B^0 \to D^{*+}D^{*-}$ and $B_s^0 \to D^{*+}D^{*-}$ decays are performed to extract the parameters. The resulting PDFs and the mass distributions are shown in Fig. 6.8. All results of the fit parameters are fixed in fit to recorded data. The ratio of the longitudinal and transverse components $g_{D^{*+}D^{*-}}$ is assumed to be the same for B^0 and B_s^0 mesons and is free in the fit to recorded data.

6.2.2 Fit to recorded data

The complete PDF from Eq. (6.4) is fitted to the recorded data using extended maximum likelihood fits. The fits are performed separately for Run 1 and Run 2. The only free parameters are the means of the signal and $B^0 \to D_s^+ D^{*-}$ contributions, the slope of the combinatorial background β_{Comb}, the ratio of the longitudinal and transverse contributions of the partially-reconstructed backgrounds, and the yields of all components. The PDFs and mass distributions are shown in Fig. 6.9 together with the respective plots for the $D^0 \to K^-\pi^+\pi^-\pi^+$ sample.

The fit results are listed in Table 6.3. The mean of the signal contribution is compatible with the known B^0 mass [41]. The ratio of the longitudinal- and transverse-polarisation components are compatible the prediction from Ref. 126. The fit yields 856 ± 32 (469 ± 28) and 3265 ± 61 (1570 ± 48) signal candidates for the $D^0 \to K^-\pi^+$ ($D^0 \to K^-\pi^+\pi^-\pi^+$) sample for Run 1 and Run 2, respectively. The results are used to extract sWeights, which are used in the fit to extract the CP

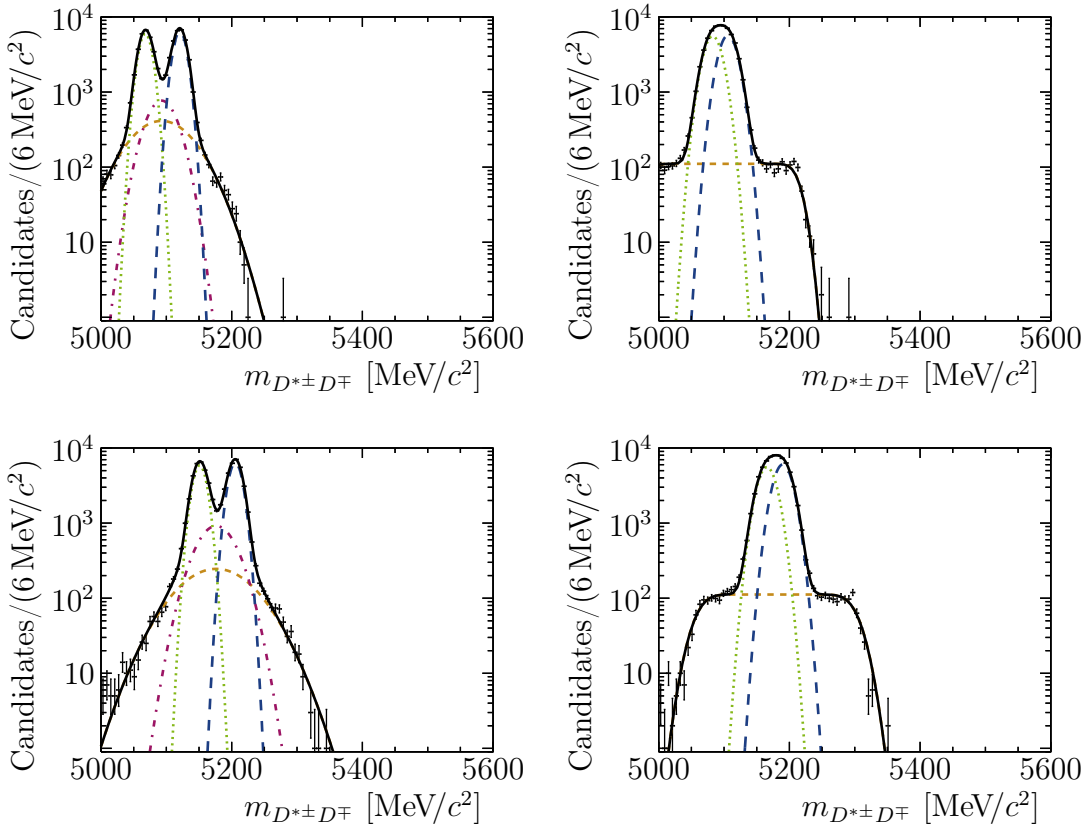

Figure 6.8: Mass distributions and PDF projections of simulated partially-reconstructed (top) $B^0 \to D^{*+}D^{*-}$ and (bottom) $B_s^0 \to D^{*+}D^{*-}$ decays with (left) longitudinal and right (transverse) polarisation. The projection of the total PDF is shown in solid black. The two contributions from decays, in which a neutral pion is not reconstructed, are modelled with two Gaussian PDFs and shown in long-dashed blue and dotted green. The contribution from decays, in which a photon is not reconstructed, uses different models depending on the polarisation. In the longitudinal case, two Gaussian PDFs are used, which are shown in short-dashed yellow and dash-dotted magenta. In the transverse case, the convolution of a uniform distribution and a Gaussian function is used, which is shown in short-dashed yellow.

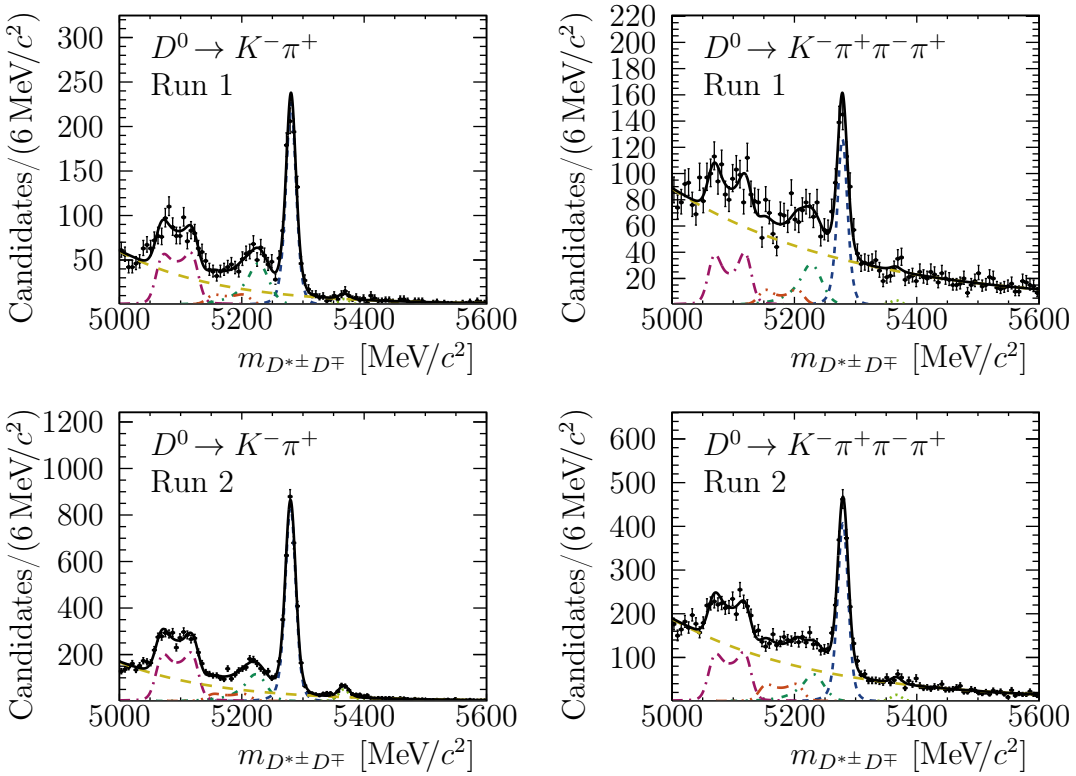

Figure 6.9: Mass distributions and PDF projections of the (top) Run 1 and (bottom) Run 2 data samples with (left) $D^0 \to K^-\pi^+$ and (right) $D^0 \to K^-\pi^+\pi^-\pi^+$ decays. The projection of the total PDF is shown in solid black. The (dashed blue) $B^0 \to D^{*\pm}D^{\mp}$ signal contribution is visible at the known B^0 mass. Also shown are the projections of the (dotted green) $B^0_s \to D^{*\pm}D^{\mp}$ background, (dash-dotted dark green) $B^0 \to D^+_s D^{*-}$ background, (long-dash-dotted magenta) $B^0 \to D^{*+}D^{*-}$ background, (dash-three-dotted red) $B^0_s \to D^{*+}D^{*-}$ background and (long-dashed yellow) combinatorial background.

Table 6.3: Results of the mass fit to recorded data to extract sWeights.

Parameter	Run 1	Run 2
$\mu_{B^0\to D^{*\pm}D^{\mp}}$ (MeV/c^2)	5280.4 ± 0.4	5279.6 ± 0.2
$\mu_{B^0\to D^+_s D^{*-}}$ (MeV/c^2)	5226.9 ± 2.1	5223.0 ± 1.4
β_{Comb} (c^2/ MeV $\cdot 10^{-3}$)	-6.16 ± 0.22	-6.24 ± 0.14
$g_{D^{*+}D^{*-}}$	0.49 ± 0.06	0.52 ± 0.03
$N_{B^0\to D^{*\pm}D^{\mp}}$	856 ± 32	3265 ± 61
$N_{B^0_s\to D^{*\pm}D^{\mp}}$	24 ± 8	197 ± 19
$N_{B^0\to D^+_s D^{*-}}$	404 ± 33	1007 ± 57
$N_{B^0\to D^{*+}D^{*-}}$	693 ± 23	2451 ± 75
$N_{B^0_s\to D^{*+}D^{*-}}$	141 ± 36	395 ± 64
N_{Comb}	1579 ± 66	4359 ± 111

observables. The fit also yields roughly 280 ± 30 candidates for the $B_s^0 \to D^{*\pm} D^{\mp}$ decay summed for the Run periods and D^0 decay modes. This decay was not yet experimentally observed when this analysis was carried out. An independent analysis was launched that resulted in the first observation of the decay and the measurement of the branching ratio [35].

6.3 Decay-time fit

To measure the *CP* observables, a fit to the decay-time in the window 0.3–10.3 ps is performed using the background-subtracted data set containing both final states. The central preselection suppresses data at very low decay times. The upper boundary lies is more than six times the B^0 lifetime, so that almost no B^0 candidates contribute beyond the boundary. The necessary preparations are performed as described in Chap. 5.

6.3.1 Instrumental asymmetries

The procedure to determine the instrumental asymmetries is only summarised briefly, more information can be found in Ref. 119. Using $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ decays, where the $D^{\pm}$ mesons are created promptly in the PV, the detection asymmetries of the D mesons $A_{\text{det}}^{D^*}$ and A_{det}^{D} are determined. The data was recorded in Run 2, samples with Run 1 data are not available, as this study was only added in a late phase of the analysis. The data is selected using the PID requirements from the nominal centralised preselection and a requirement on χ^2_{IP} to suppress secondary $D^{\pm}$ decays. The sample is then split in four categories: one per D meson for both samples containing the different D^0 final states.

The samples are adjusted to the signal samples using the momenta and transversal momenta of the respective final-state particles. The additional $\pi^+\pi^-$ pair in the $D^0 \to K^-\pi^+\pi^-\pi^+$ has no representatives in the prompt $D^{\pm}$ sample. As the momentum and transversal momentum distributions of the negative and positive pion is found to be compatible, they contribute in the same way in D^{*+} and D^{*-} decays. Thus, they have no effect on the detection asymmetry. Fits to the invariant D mass on halves of the data samples are performed to subtract the background. The sWeighted samples are adjusted using a two-folded multivariate approach [127] to match the distribution of the signal decays. The adjusted samples are used to perform additional fits to extract the yields per D meson charge. These yields are used to calculate the asymmetry for the four categories. The results are listed in Table 6.4. The total asymmetry $A_{\text{det}}^{D^*D}$ is compatible with zero. The asymmetry relies on kinematic asymmetries, which are even smaller in the simulated signal samples for Run 1. Thus, the asymmetry for Run 1 is assumed

Table 6.4: Detection asymmetries for the D mesons and their combination for Run 2.

D^0 decay	$A_{\text{det}}^{D^*}$ [%]	A_{det}^{D} [%]	$A_{\text{det}}^{D^*D}$ [%]
$D^0 \to K^-\pi^+$	1.69 ± 0.36	1.58 ± 0.18	0.11 ± 0.40
$D^0 \to K^-\pi^+\pi^-\pi^+$	1.46 ± 0.22	1.38 ± 0.15	0.09 ± 0.26

to be zero. The uncertainty is conservatively assumed to be twice the uncertainty of the Run 2 sample.

6.3.2 Flavour-tagging calibration

The $B^0 \to D_s^+ D^{*-}$ and $B^0 \to D_s^+ D^-$ decays are chosen as control channels. These channels are kinematically very similar to the signal channels. Usually, an adjustment of the calibration samples using topological variables is necessary to obtain valid calibration parameters. These variables typically include a subset of the transversal momentum and pseudorapidity of the B^0 candidate, and the number of tracks and PVs in the event. Studies using simulated data show that the $B^0 \to D_s^+ D^{*-}$ and $B^0 \to D_s^+ D^-$ decays are similar enough and an adjustment is not necessary. A comparison using recorded data is not feasible, to the relatively low sample size of the signal decay. The statistical power of the calibration channels is also limited and not suitable to calibrate the single taggers, especially in Run 1. Thus, the single taggers are calibrated using adjusted $B^0 \to D^-\pi^+$ samples. The former samples are used to calibrate the combination of the SS and OS taggers, respectively. To increase the statistical power, the resulting calibration parameters are combined for both channels. The calibrations are performed simultaneously for the two samples containing the different D^0 final states, but separately for Run 1 and Run 2. The FT algorithms defined in Sec. 3.4 are used, with the exception of the OSCharm tagger, which is only available for the Run 2 calibration and is not used for Run 1.

The LHCb FT group provides selected and background-subtracted $B^0 \to D^-\pi^+$ samples. The sample is adjusted to match the transversal-momentum distribution of the recorded $B^0 \to D_s^+ D^{*-}$ and $B^0 \to D_s^+ D^-$ data. The FT algorithms are calibrated using the EPM tool [100]. This procedure neglects *CP* asymmetries in the $B^0 \to D^-\pi^+$ decay, which can influence the FT calibration but are measured to be negligible [128].

The calibrated single taggers are combined to obtain an SS and an OS combination. The combinations are then calibrated using $B^0 \to D_s^+ D^{*-}$ and $B^0 \to D_s^+ D^-$ samples. The decays are reconstructed using the same intermediate decays as for the signal decays. Additionally, the D_s^+ mesons are reconstructed using the

$D_s^+ \to K^-K^+\pi^+$ decay. The samples are selected using the invariant masses of the D mesons and requirements on the final-state-particle momentum, transverse momentum and PID. The selection is similar to the signal selection, but uses no multivariate analysis. Fits to the reconstructed B^0 masses are used to be able to subtract the remaining backgrounds using the sPlot method. The EPM tool is not used to extract the calibration parameters, as it relies on the assumptions of no asymmetries due to detection, production or FT. Instead, a fit to the decay-time distribution using Eq. (5.13) is performed. *CP* violation in these decays is assumed to be negligible and the decays are assumed to be completely flavour-specific. The lifetime τ_d and mass difference Δm_d of B^0 mesons are fixed to their known value and the lifetime difference $\Delta\Gamma_d$ is fixed to 0, which is a good approximation [43]. This allows a different set of calibration parameters to be obtained for the two different tags, and allows A_{prod} to be measured.

Systematic uncertainties on the calibration parameters need to be evaluated. The quadratic sum of the systematic and statistical uncertainties is used to apply Gaussian constraints in the fit to extract the *CP* observables. Effects from the choices of the mass model, of the decay-time-dependent efficiency and the decay-time resolution, as well as from external parameters are evaluated. The uncertainties are determined by changing different parts of the FT calibration procedure. The variation of the calibration parameters with the respect to the nominal fit results is assigned as the systematic uncertainty. The systematic uncertainty due to the mass model is determined by performing alternative fits instead of fits to the reconstructed B^0 mass. Bi-dimensional fits to either the D_s^+ and D^- mass distributions, or the distribution of the D_s^+ mass and the difference of the D^{*-} and $\overline{D}^0$ masses, are performed. The uncertainty due to the model of the decay-time-dependent efficiency is evaluated with a model using a rational function instead of the B-spline-based model. The uncertainty due to the choice of the decay-time resolution model is evaluated by increasing or decreasing the effective resolution by 50%. The effect is negligible. The systematic uncertainty due to the external inputs is calculated by performing four additional fits, where τ_d and Δm_d are varied by adding or subtracting their current uncertainty. The effect of $\Delta\Gamma_d$ on the FT calibration was previously measured to be negligible [129]. The systematic uncertainty due to the mass model is dominant, while the other sources are mostly negligible. An additional cross-check using the EPM tool yields compatible results.

Finally, the weighted average for both control channels is calculated. This assumes no correlation of the statistical uncertainties and full correlation of the systematic uncertainties. The results are used as inputs when extracting the *CP* observables. The results for the production asymmetries are

$$\begin{aligned} A_{\text{prod}}^{\text{Run 1}} &= -\,0.0107 \pm 0.0081\,(\text{stat}) \pm 0.0028\,(\text{syst})\,, \\ A_{\text{prod}}^{\text{Run 2}} &= 0.0037 \pm 0.0046\,(\text{stat}) \pm 0.0017\,(\text{syst})\,, \end{aligned}$$

Table 6.5: Tagging performance parameters from the $B^0 \to D^{*\pm} D^{\mp}$ sample with $D^0 \to K^-\pi^+$ decays: the average of the predicted mistag probability $\langle\eta\rangle$, the tagging efficiency ε_{tag} and the effective tagging efficiency ε_{eff}.

Sample	Tagger	$\langle\eta\rangle$	ε_{tag} [%]	ε_{eff} [%]
Run 1	OS	0.3507	39.9 ± 1.8	$3.47 \pm 0.30(\text{stat}) \pm 0.47(\text{syst})$
	SS	0.4257	68.1 ± 1.7	$2.34 \pm 0.20(\text{stat}) \pm 0.37(\text{syst})$
	total		80.2 ± 1.4	$5.61 \pm 0.36(\text{stat})$
Run 2	OS	0.3444	40.1 ± 0.9	$3.94 \pm 0.16(\text{stat}) \pm 0.25(\text{syst})$
	SS	0.4289	89.3 ± 0.6	$2.92 \pm 0.09(\text{stat}) \pm 0.18(\text{syst})$
	total		93.5 ± 0.5	$6.61 \pm 0.19(\text{stat})$

where the first uncertainty is statistical and the second one is systematic. Both values are compatible with zero. To determine the FT performance, the calibration parameters are applied to the $B^0 \to D^{*\pm} D^{\mp}$ sample and the results are given in Table 6.5. The tagging power is $\varepsilon_{\text{eff}} = 5.61 \pm 0.36$ and $\varepsilon_{\text{eff}} = 6.61 \pm 0.19$ for Run 1 and Run 2, respectively.

6.3.3 Decay-time resolution

The resolution function from Eq. (5.14) is used to describe the distribution of the difference of the true and the recorded decay time of simulated decays. The distributions are compatible between Run 1 and Run 2. As the overall effect of the resolution is almost negligible, only the results determined using the simulated Run 2 data is used. The resulting PDF of the maximum likelihood fit and the time-difference distribution for the sample containing $D^0 \to K^-\pi^+$ decays is shown in Fig. 6.10. The resulting parameters are fixed in the fit to recorded data. The resolution can be quantified with the effective decay-time resolution

$$\sigma_{\text{eff}} = \sqrt{(-2/\Delta m_d^2) \ln D}\,, \tag{6.10}$$

where D is the decay-time Dilution

$$D = \sum_i g_i \mathrm{e}^{-\sigma_i^2 \Delta m_d^2/2}\,, \tag{6.11}$$

where g_i and σ_i are the resolution parameters of the separate Gaussian functions from Eq. (5.14). The result of the effective decay-time resolution is $\sigma_{\text{eff}} \approx 55\,\text{fs}$.

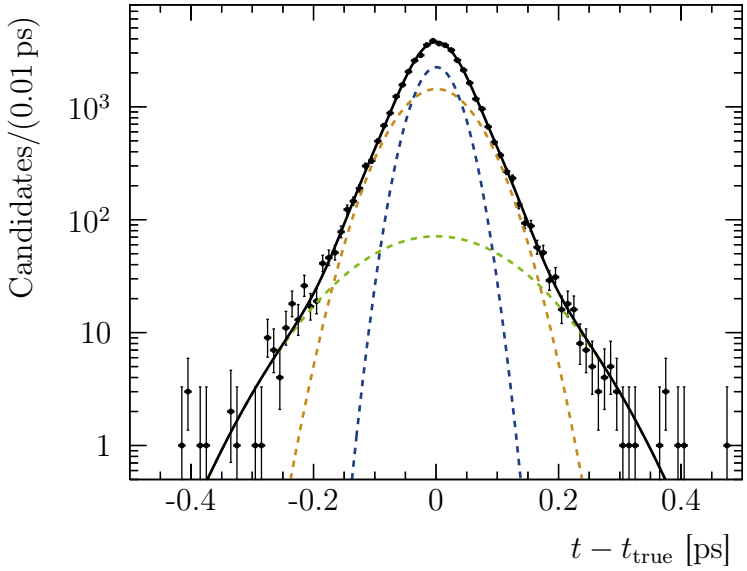

Figure 6.10: Distribution of the difference of the true and reconstructed decay time and the projection of the PDF describing the decay-time resolution. The sample contains $D^0 \to K^-\pi^+$ decays and is produced with Run 2 conditions. The resolution PDF comprises three Gaussian functions. The total PDF projection is shown in solid black and the three components are shown in dashed blue, dashed orange and dashed green.

6.3.4 Decay-time-dependent efficiency

The decay-time-dependent efficiency leads to a deviation of the decay-time distribution from the expected shape: a pure exponential function that is convolved with the resolution function. To parametrise the decay-time-dependent efficiency, the part that depends on the *CP* observables in the PDF from Eq. (5.13) is replaced with a function that is exponentially decreasing according to the B^0 lifetime. The decay-time-dependent efficiency is parametrised with the B-spline model described in Chap. 5. The number of knots and the knot positions are varied until an adequate description of the distribution is achieved. This procedure is performed using simulated decays before checking that the parametrisation also describes the recorded data. This assumes the same knots, but potentially different coefficients for simulated and recorded data. The minimal number of knots is used that allows an accurate parametrisation. The distributions of the two D^0 final states are found to be compatible and are treated the same. The distributions of Run 1 and Run 2 are not compatible and are treated separately in the fit to recorded data. The PDFs resulting from fits to simulated and recorded data and the distributions are shown in Fig. 6.11.

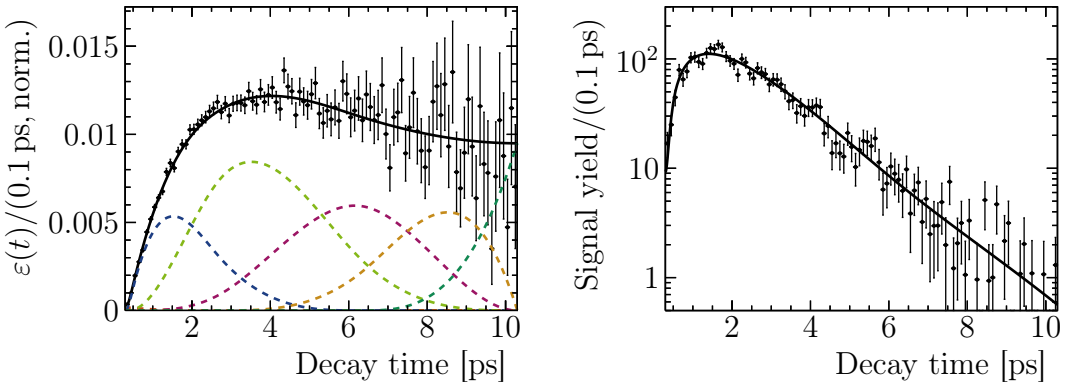

Figure 6.11: Projection of the PDF that describes the decay-time-dependent efficiency on the left with the distribution of simulated data, where the effect of the exponential decay is eliminated. The total PDF projection is shown in solid black. The contributions from the different B-splines are shown in different colours. The projection of the PDF of the decay-time-dependent efficiency multiplied with an exponential decay with the B^0 lifetime, and the background-subtracted data are shown on the right. Both plots show the results for the Run 2 and $D^0 \to K^-\pi^+$ samples.

The resulting knots are positioned at 0.3, 0.5, 2.7, 6.3 and 10.3 ps. The first and last positions correspond to the boundaries of the decay-time-description. More knots are placed at low decay-time because of the higher density of the data in this region. The coefficients are free in the fit to recorded data. This is reasonable due to the very low correlation of the coefficients and the other parameters.

6.3.5 Extraction of CP observables

The final inputs for the extraction of the CP parameters are the average mistag estimates and external parameters. The average mistag estimates are measured in the sWeighted data samples. The values for the $D^0 \to K^-\pi^+$ ($D^0 \to K^-\pi^+\pi^-\pi^+$) sample are $\langle\eta^{\text{OS}}\rangle = 0.3507$ (0.3527) and 0.3444 (0.3452), and $\langle\eta^{\text{SS}}\rangle = 0.4257$ (0.4323) and 0.4289 (0.4266), for Run 1 and Run 2, respectively. The external parameters are τ_d and Δm_d, which are taken from Ref. [43]. The lifetime difference $\Delta\Gamma_d$ is fixed to zero. The production asymmetry and the FT-calibration parameters are implemented with Gaussian constraints using the values and uncertainties determined in Sec. 6.3.2. The B-spline coefficients are free in the fit.

The PDF resulting from the maximum likelihood fit to extract the CP parameters and the decay-time distribution in shown in Fig. 6.12. The results of the CP

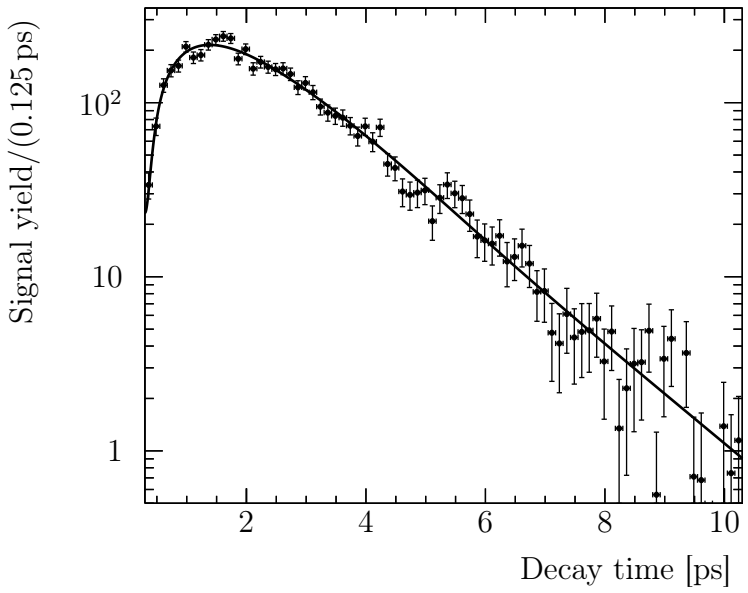

Figure 6.12: Decay time distribution of signal candidates and the projection of the PDF. The data set comprises the recorded Run 1 and Run 2, and $K\pi$ and $K\pi\pi\pi$ subsamples.

observables and $A_{\text{raw}}^{\hat{f}}$ values are

$$\begin{aligned} S_{D^*D} &= -0.861 \pm 0.077\,, & \Delta S_{D^*D} &= 0.019 \pm 0.075\,, \\ C_{D^*D} &= -0.059 \pm 0.092\,, & \Delta C_{D^*D} &= -0.031 \pm 0.092\,, \\ A_{\text{raw}}^{K\pi,\text{Run1}} &= 0.018 \pm 0.037\,, & A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}} &= 0.079 \pm 0.059\,, \\ A_{\text{raw}}^{K\pi,\text{Run2}} &= -0.029 \pm 0.019\,, & A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}} &= 0.009 \pm 0.030\,. \end{aligned}$$

The correlations of the *CP* parameters are

$$\rho_{ij} = \left(\begin{array}{c|cccc} & S_{D^*D} & \Delta S_{D^*D} & C_{D^*D} & \Delta C_{D^*D} \\ \hline S_{D^*D} & 1 & 0.07 & 0.44 & 0.05 \\ \Delta S_{D^*D} & & 1 & 0.04 & 0.46 \\ C_{D^*D} & & & 1 & 0.04 \\ \Delta C_{D^*D} & & & & 1 \end{array}\right).$$

The correlation of the various parameters of A_{raw} to the other parameters is well below 1%. The correlation of the B-spline coefficients to the *CP* observables is also well below 1%.

6.4 Systematic effects

Systematic uncertainties are evaluated mainly using pseudo-experiments. The procedure follows the description in Sec. 5.2. The PDFs describing the mass and decay-time distributions are assumed to be uncorrelated. The contributions of the data sample, *i.e.* signal, backgrounds from other decays and combinatorial background, are generated separately. This allows to use different decay-time PDFs for the different components. The combinatorial background assumes no *CP* violation or oscillation. Their lifetime is set to $\tau = 0.5\,\text{ps}$, which is determined with a fit on the upper-mass sideband. The PDFs for the contributions from B^0 and B_s^0 decays use the respective values for the lifetime and oscillation parameters. Only the signal contribution and the contribution from $B^0 \to D^{*+} D^{*-}$ decays assume *CP* violation. The parameters are taken from the fit to recorded data respectively Ref. 43 for the $B^0 \to D^{*+} D^{*-}$ decay. The PDF of the flavour-specific $B^0 \to D_s^+ D^{*-}$ decay uses $\Delta C = 1$. The mistag distributions are generated according to histograms of the background-subtracted data sample. The SS combination is assumed to provide random tags for B_s^0 decays and the tags of both combinations are assumed to be random for combinatorial background. Parameters that are Gaussian-constrained in the fit to recorded data are generated using Gaussian functions that correspond to the constraints [115].

The total systematic uncertainties are reported in Table 6.6. The dominant systematic uncertainties come from the mass model and the FT. The different sources are discussed in the following sections.

Table 6.6: Total systematic uncertainties and the quadratic sum.

Source	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
Fit bias	0.002	0.002	0.002	0.002
$\Delta m_d, \tau_d, \Delta\Gamma_d$	0.001	0.001	0.003	0.001
Mass model	0.011	0.003	0.014	0.006
Flavour tagging	0.015	0.012	0.014	0.015
Sum	0.019	0.012	0.020	0.016

Source	$A_{\text{raw}}^{K\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi,\text{Run2}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$
Fit bias	0.0008	0.0004	0.0013	0.0007
$\Delta m_d, \tau_d, \Delta\Gamma_d$	0.0002	0.0001	0.0003	0.0002
Mass model	0.0021	0.0016	0.0025	0.0024
Decay-time resolution	0.0001	0.0001	0.0002	0.0001
Decay-time acceptance	0.0002	0.0001	0.0003	0.0001
Flavour tagging	0.0001	0.0001	0.0001	0.0001
Sum	0.0023	0.0016	0.0028	0.0025

6.4.1 Cross-checks and fit bias

Decay-time fits to fully-simulated data are used to confirm the validity of the fit model. The resulting *CP* observables are consistent with the results used in the generation of the data.

The inherent fit bias is examined by performing 2000 pseudo-experiments using the nominal models in the generation and in the fits. The results are listed in Table 6.7. This study is repeated using alternative *CP* observables in the generation. None of these studies provide statistically different biases. The uncertainty of the means of the fit results are taken as systematic uncertainty. They are of order $\mathcal{O}(10^{-3})$ or lower for all parameters. The statistical uncertainties seem to be overestimated, as the widths of the pull distribution is smaller than one. The problem seems to lie in the generation of the FT calibration parameters, as the overestimation becomes smaller when fixing them in the generation. Another possible source of wrongly estimated uncertainties is the use of sWeights in the fit [105]. The statistical uncertainty is examined with a bootstrapping study [106], as explained in Sec. 4.2, performing the nominal mass and decay-time fits to 1000 bootstrapped data samples. The widths of the residual distributions, *i.e.* the difference of the fit result and the generated value, the means of the fit uncertainties, and the widths of the pull distributions are listed in Table 6.8. The widths of the residual distributions and the means of the fit uncertainties are compatible. The widths of the pull distributions are compatible with one. This indicates that the fit uncertainties are correctly estimated.

Additional cross-checks are performed by splitting the data set in different categories and comparing the fit results. The splits include the D^0 mode, SS-tagged

Table 6.7: Means and widths of distributions of the fit results and the pull for the fit validation. There is no significant bias. The statistical uncertainties seem to be overestimated.

	Pull		Fit results	
Parameter	Mean	Width	Mean	Width
S_{D^*D}	-0.003 ± 0.019	0.924 ± 0.017	-0.862 ± 0.002	0.069 ± 0.001
ΔS_{D^*D}	0.008 ± 0.023	0.954 ± 0.018	0.017 ± 0.002	0.068 ± 0.001
C_{D^*D}	0.025 ± 0.024	0.970 ± 0.015	-0.058 ± 0.002	0.091 ± 0.001
ΔC_{D^*D}	0.010 ± 0.023	0.981 ± 0.018	-0.031 ± 0.002	0.089 ± 0.001
$A_{\text{raw}}^{K\pi,\text{Run1}}$	-0.010 ± 0.023	0.972 ± 0.018	-0.001 ± 0.001	0.036 ± 0.001
$A_{\text{raw}}^{K\pi,\text{Run2}}$	0.006 ± 0.022	0.942 ± 0.018	0.001 ± 0.001	0.018 ± 0.001
$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	-0.003 ± 0.023	0.977 ± 0.017	-0.002 ± 0.001	0.058 ± 0.001
$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$	0.021 ± 0.023	0.980 ± 0.018	0.004 ± 0.001	0.029 ± 0.001

Table 6.8: Results of the fits to bootstrapped data: the widths of the residuals, the means of the fit uncertainties and the widths of the pull distributions.

Parameter	Residual Width	Fit uncertainty Mean	Pull Width
S_{D^*D}	0.0757 ± 0.0018	0.07759 ± 0.00010	0.994 ± 0.024
ΔS_{D^*D}	0.0776 ± 0.0018	0.07505 ± 0.00010	1.048 ± 0.025
C_{D^*D}	0.0912 ± 0.0022	0.09198 ± 0.00009	0.997 ± 0.024
ΔC_{D^*D}	0.0897 ± 0.0021	0.09174 ± 0.00009	0.979 ± 0.023
$A_{\text{raw}}^{K\pi,\text{Run1}}$	0.0369 ± 0.0009	0.036907 ± 0.000023	1.000 ± 0.024
$A_{\text{raw}}^{K\pi,\text{Run2}}$	0.0190 ± 0.0005	0.018632 ± 0.000006	1.023 ± 0.024
$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	0.0615 ± 0.0015	0.05894 ± 0.00006	1.055 ± 0.025
$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$	0.0298 ± 0.0007	0.029767 ± 0.000015	1.000 ± 0.024

events vs OS-tagged events, magnet polarity, and Run 1 vs Run 2. The Run 1 sample does not have the statistical power to guarantee a stable decay-time fit. Instead, a comparison of the complete data sample and the Run 2 sample is performed. The splits show no statistically significant deviation. Finally, the extraction of the *CP* observables is repeated without the suppression of multiple candidates. The deviations to the nominal results are negligible.

A study using 3000 pseudo-experiments is performed, where a sizeable asymmetry $A_{\text{raw}}^{D_s D^*}$ is used in the generation of the $B^0 \to D_s^+ D^{*-}$ contribution. No biases are observed and no systematic uncertainty is assigned.

The potential contribution from misidentified $B_s^0 \to D_s^+ D^{*-}$ decays is neglected in the mass fit. The mass fit is repeated by adding this contribution, using the same parametrisation as for the $B^0 \to D_s^+ D^{*-}$ contribution and fixing the relative mean of the contribution to the known mass difference $\mu_{B_s^0} - \mu_{B^0} = (87.23 \pm 0.24)\,\text{MeV}$ [41]. The number of $B_s^0 \to D_s^+ D^{*-}$ candidates is compatible with zero and the number of signal candidates is changed by less than 1%. No systematic uncertainty is assigned.

The production asymmetry is included with Gaussian constraints in the fit. The effect from its uncertainty is thus included in the statistical uncertainty of the fit. The decay-time fit is repeated with fixed values of the production asymmetries which leads to reduced statistical uncertainties. The values of the *CP* observables are unchanged. The contribution to the uncertainties is calculated with the quadratic difference of the nominal and reduced uncertainty. The uncertainties are $\sigma_{S_{D^*D}} = 0.002$ for S_{D^*D} and negligible for the other parameters. Increasing the uncertainty on A_{prod} to ± 10, which is 10 to 20 times larger than the nominal value, increases the uncertainties on the *CP* parameters to $\sigma_{S_{D^*D}} = 0.029$ and $\sigma_{C_{D^*D}} = 0.023$. The other uncertainties are negligible.

6.4.2 External inputs

The parameters τ_d, Δm_d and $\Delta\Gamma_d$ are fixed in the decay-time fit. A study using 3000 pseudo-experiments is performed, where the external inputs are generated according to Gaussian distributions. The functions use the known value and uncertainty from Ref. [43] as mean and width. The systematic uncertainties are assigned as described before and listed in Table 6.6. They are of order $\mathcal{O}(10^{-3})$ for the *CP* observables and order $\mathcal{O}(10^{-4})$ for the A_{raw} parameters.

6.4.3 Mass model

Several variations of the parametrisation of the mass distribution are investigated. The systematic uncertainties are calculated as the quadratic sum of the mean and width of the differences of the fit results and the generated *CP* observables.

While the mass is parametrised using three Crystal Ball functions for the $D^0 \to K^-\pi^+$ sample, only two are used to parametrise the $D^0 \to K^-\pi^+\pi^-\pi^+$ sample. Two alternative models which use two or three Crystal Ball functions in the fits for both final states are used to perform a study using 2000 pseudo-experiments. A third study is performed, where the generation uses three Crystal Ball functions, but only two are used in the fits.

An alternative model of the combinatorial background is used in the generation step of pseudo-experiments. A second order polynomial is chosen that describes the recorded data accurately.

A third systematic uncertainty due to the mass model is assigned without using pseudo-experiments. To account for the difference of the mass resolution between simulated and recorded data, a scaling factor is introduced. This factor is multiplied to the widths of the Crystal Ball functions describing the signal decays. The difference of the results of the following decay-time fit and the nominal values are assigned as systematic uncertainty.

The different systematic uncertainties due to the mass model and their quadratic sum is are listed in Table 6.9. They are the dominant uncertainties for the A_{raw} parameters with values of order $\mathcal{O}(10^{-3})$. They are also significant for the averages of the *CP* observables with values of order $\mathcal{O}(10^{-2})$, but an order of magnitude smaller for the differences of the *CP* observables.

6.4.4 Decay-time resolution

The uncertainty due to the parametrisation of the decay-time resolution is evaluated by changing the resolution by +10% during the fit. The difference of the

Table 6.9: Systematic uncertainties due to the mass model and the quadratic sum. Negligible values are not listed.

Source	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
3CB in fit	0.0013	0.0009	0.0017	0.0017
2CB in fit	0.0011	0.0008	0.0017	0.0013
3CB in generation	0.0072	–	0.0051	–
Combinatorial background	0.008	0.002	0.007	0.002
Mass resolution	0.002	0.002	0.011	0.005
Sum	0.011	0.003	0.014	0.006

Source	$A_{\text{raw}}^{K\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi,\text{Run2}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$
3CB in fit	0.0002	0.0013	0.0011	0.0009
2CB in fit	0.0007	0.0003	0.0001	0.0001
3CB in generation	–	–	–	0.0019
Combinatorial background	0.001	–	0.002	0.001
Mass resolution	0.0017	0.0008	0.0010	0.0007
Sum	0.0021	0.0016	0.0025	0.0024

results using the nominal and the alternative resolution is calculated for 3000 pseudo-experiments. The uncertainties is assigned as the quadratic sum of the mean and width of the distribution. They are only significant for the A_{raw} parameters, with values of order $\mathcal{O}(10^{-4})$, and listed in Table 6.6.

6.4.5 Decay-time-dependent efficiency

Two sources are evaluated relating to the decay-time-dependent efficiency. For the first source, an alternative set of knots are found that provides stable results on simulated data. Fits with the nominal and alternative model are performed for 3000 pseudo data sets that are generated with the nominal model. The uncertainties is assigned as for the decay-time resolution, *i.e.* the quadratic sum of the mean and width of the distributions of the difference of both fit results. The second source is the assumption that the same parametrisation of the decay-time-dependent efficiency can be used for both D^0 modes. A bootstrapping study [106] is performed with two fits to each bootstrapped data set, with a shared and a separate parametrisation for both D^0 modes. The systematic uncertainties are assigned as before. The quadratic sum of both studies is given in Table 6.6, only the effect on the A_{raw} parameters is non-negligible and of order $\mathcal{O}(10^{-4})$.

6.4.6 Flavour tagging calibration

The statistical and systematic uncertainties on the FT calibration parameters are included in the statistical uncertainty of the *CP* parameters due to the Gaussian constraints. The contribution to S_{D^*D}, $\sigma_{S_{D^*D}} = 0.002$, is the only one that is non-negligible calculated with an additional fit where the FT parameters are fixed instead of Gaussian constrained. A study is performed using the correlations between the calibration parameters reported by the decay-time fit to the control channels. The correlation is included in the Gaussian constraints. The difference of the results of the *CP* parameters is two orders of magnitude smaller than their statistical uncertainty.

On simulated data, the calibration function of the SS combination of the two control channel is not identical to the signal channel. In 3000 pseudo-experiments, data sets are generated that contain two sets of calibration parameters: the nominal set and one set that reproduces the differences observed in the simulated data. Two fits to each data set are then performed, each using a different set of calibration parameters. The systematic uncertainties are then calculated as described before. The effect on the A_{raw} parameters is of order $\mathcal{O}(10^{-5})$. The effect on the *CP* observables is of order $\mathcal{O}(10^{-2})$ making this the dominant source of systematic uncertainty.

The OS tagging efficiency shows a dependency on the B^0 decay time in simulated data. A similar study as for the SS calibration is performed: two different OS tagging efficiencies are generated for 3000 pseudo-experiments, the nominal one and one that includes the dependency on the decay time. Two fits are then respectively performed using both efficiencies and the systematic uncertainties are calculated as described above. The systematic uncertainties are of order $\mathcal{O}(10^{-5})$ for all parameters. The quadratic sum of the systematic uncertainties due to the SS calibration and OS tagging efficiency is reported in Table 6.6.

6.4.7 Instrumental asymmetries

The systematic uncertainties on the detection asymmetries are listed in Table 6.10. A detailed summary of the determination of the detection asymmetries can be found in Ref. 119.

The production asymmetry of the prompt D mesons is neglected during the calculation of the detection asymmetry. The production asymmetry is estimated using a measurement at 7 TeV [130]. As the data-taking conditions do not match the conditions of the prompt D samples, the production asymmetry is not taken into account when calculating the detection asymmetries. Instead, the difference of the production asymmetries for $D^{*\pm}$ and $D^{\pm}$ mesons is assigned as systematic uncertainties.

Table 6.10: Systematic uncertainties of the detection asymmetry.

Origin	$\sigma_{A_{\text{det}}^{K\pi}}$	$\sigma_{A_{\text{det}}^{K\pi\pi\pi}}$
$D^{\pm}$ production asymmetry	0.0002	0.0010
`L0Hadron` trigger asymmetry	0.0022	0.0015
PID asymmetry due to vetoes	0.0020	0.0020
PID asymmetry due to BDT	0.0003	0.0036
Secondary $D^{\pm}$ mesons	0.0001	0.0023
Mass model	0.0001	0.0010
Sum	0.0030	0.0051

A large fraction of the prompt D sample is triggered by decisions from the `L0Hadron` trigger due to particles in the signal decay chain. The asymmetry of this trigger as a function of the transversal momentum is defined as

$$A_{\texttt{L0HadTOS}}(p_{\text{T}}) = \frac{N^{+}_{\texttt{L0HadTOS\&L0TIS}}(p_{\text{T}})}{N^{+}_{\texttt{L0TIS}}(p_{\text{T}})} - \frac{N^{-}_{\texttt{L0HadTOS\&L0TIS}}(p_{\text{T}})}{N^{-}_{\texttt{L0TIS}}(p_{\text{T}})}\,, \tag{6.12}$$

where $N^{\pm}$ is the number of positive or negative D mesons that are triggered by either both the `L0Hadron` trigger due to a particle in signal decay chain or any `L0` trigger due to a particle not in the decay chain. The asymmetry is calculated per D meson as the integral of the asymmetries over the p_{T} distribution. The total asymmetry is the difference of the asymmetries for $D^{*\pm}$ and $D^{\pm}$ mesons and assigned as systematic uncertainty.

Uncertainties due to the usage of PID variables for the veto and BDT requirements are estimated using simulated decays. Asymmetries are calculated by counting the number of events of both $D^{*\pm}D^{\mp}$ charge configurations before and after the veto or BDT requirement. They are assigned as systematic uncertainties. The PID requirements in the centralised preselection are already taken into account in the measured detection asymmetry.

Another source of systematic uncertainties are secondary $D^{\pm}$ mesons, *i.e.* $D^{\pm}$ mesons that are produced in the decay of a long-lived particle. The change of the detection asymmetry by applying a tighter requirement on the χ^2_{IP} is assigned as the systematic uncertainty.

The $D^{\pm}$ mass model is the last source of systematic uncertainty that is considered. The variation of the detection asymmetry by choosing different models is assigned as the systematic uncertainty.

6.5 Results

The four values of A_{raw} from the decay-time fit, listed in Eq. (6.12), and the detection asymmetries evaluated in Sec. 6.3.1 are used to calculate A_{D^*D}. The average of the time- and flavour-integrated charge asymmetry is

$$A_{D^*D} = 0.008 \pm 0.014\,,$$

which is compatible with zero. The uncertainty is purely statistical. Altogether, the final results are

$$\begin{aligned}
S_{D^*D} &= -0.861 \pm 0.077\,(\text{stat}) \pm 0.019\,(\text{syst})\,,\\
\Delta S_{D^*D} &= 0.019 \pm 0.075\,(\text{stat}) \pm 0.012\,(\text{syst})\,,\\
C_{D^*D} &= -0.059 \pm 0.092\,(\text{stat}) \pm 0.020\,(\text{syst})\,,\\
\Delta C_{D^*D} &= -0.031 \pm 0.092\,(\text{stat}) \pm 0.016\,(\text{syst})\,,\\
A_{D^*D} &= 0.008 \pm 0.014\,(\text{stat}) \pm 0.006\,(\text{syst})\,,
\end{aligned}$$

where the only significant correlations are between the S_{D^*D} and C_{D^*D} parameters and the ΔS_{D^*D} and ΔC_{D^*D} parameters, with $\rho(S_{D^*D}, C_{D^*D}) = 0.44$ and $\rho(\Delta S_{D^*D}, \Delta C_{D^*D}) = 0.46$. This is compatible with and more precise than all previous measurements in this channel as can be seen in Fig. 6.13. This measurement has better sensitivity to S_{D^*D} and ΔS_{D^*D} than to C_{D^*D} and ΔC_{D^*D}, while the opposite is true for the measurements of the B-factories BaBar and Belle. This discrepancy comes from the decay-time-dependent efficiency at LHCb and the fact that the sensitivity to the CP parameters is highest at the extrema of the respective trigonometric function. The decay-time-dependent efficiency is lowest around decay times of zero, where the sensitivity to C_{D^*D} and ΔC_{D^*D} is best, and higher at longer decay times, where the sensitivity to S_{D^*D} and ΔS_{D^*D} is best. Still, the new results is the most precise for all of these parameters due to the much higher signal yield of more than 6000 $B^0 \to D^{*\pm}D^{\mp}$ candidates. The measurement of A_{D^*D} is even more sensitive: it is five times more precise than the measurements of the B-factories. This comes from the fact that the determination of this parameter does not rely on the flavour tagging. The B-factories reach tagging powers of around 30% [116, 131], while the measured tagging power is around 6% in the presented analysis, which is still relatively high compared to other LHCb measurements.

The result of the S_{D^*D} parameter excludes the hypothesis of CP conservation in the interference of decay and mixing at more than 10 standard deviations, obtained using Wilk's theorem [132]. The amount of direct CP violation is measured to be small as the result of C_{D^*D} is compatible with zero. The parameters ΔC_{D^*D} measures the flavour-specificity of the decay, and is compatible with zero. The relative strong phase between the $B^0 \to D^{*+}D^-$ and $B^0 \to D^{*-}D^+$ decays is

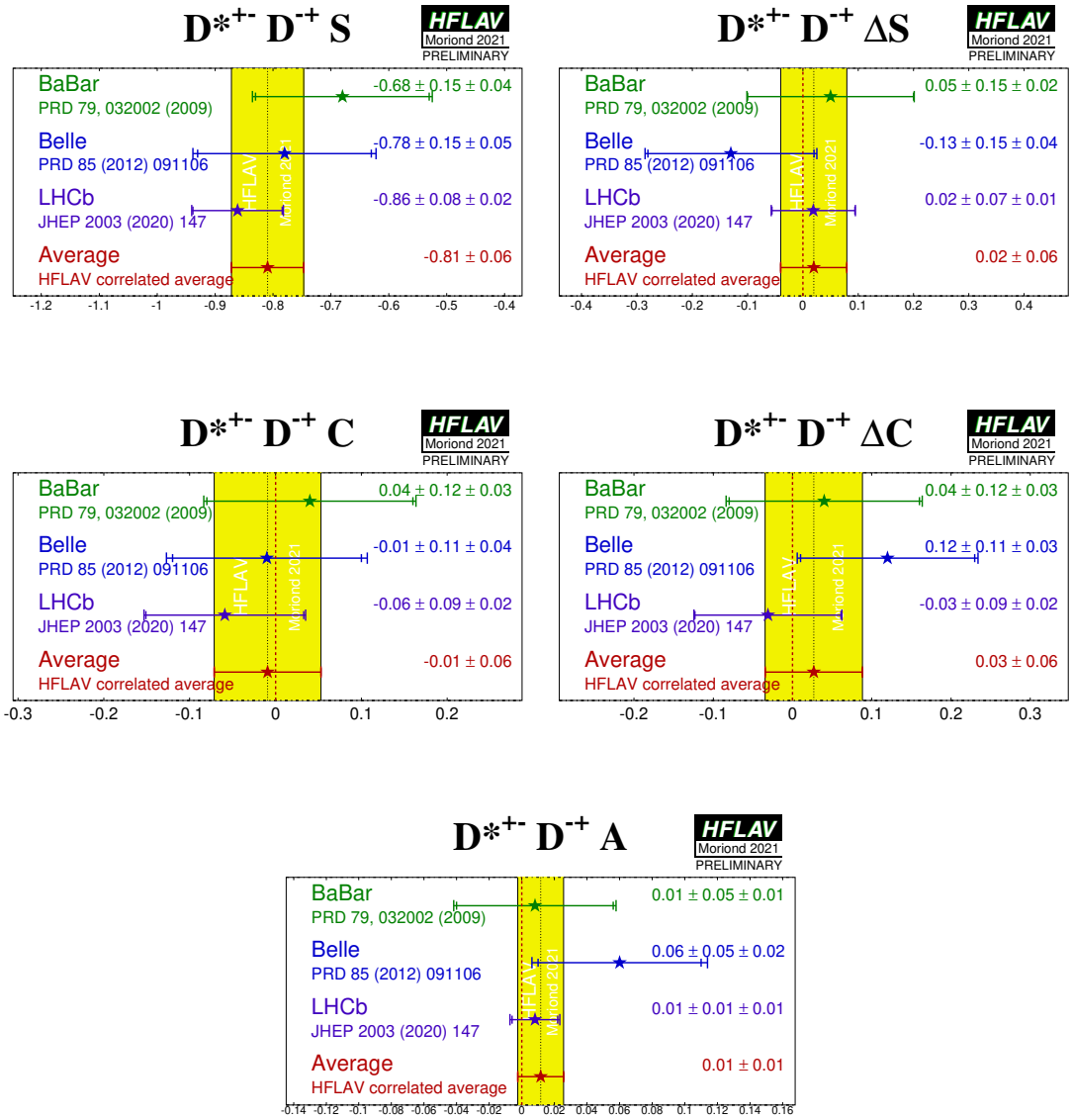

Figure 6.13: Results of the measurements of the CP observables (top left) S_{D^*D}, (top right) ΔS_{D^*D}, (middle left) C_{D^*D}, (middle right) ΔC_{D^*D} and (bottom) A_{D^*D} [43]. The previous measurements from the BaBar [116] and Belle [117] collaborations, and the result from the presented analysis [34] are listed. They are compared in the text.

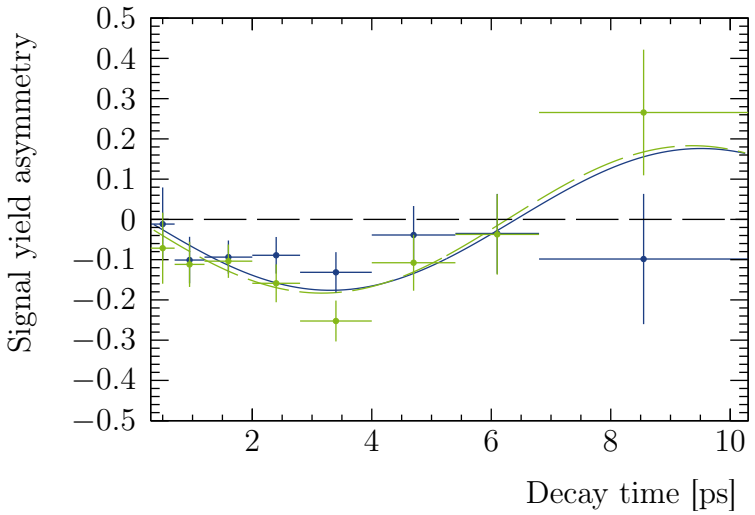

Figure 6.14: Asymmetry between the B^0 and $\overline{B}^0$ signal yields as a function of the decay time for the (long-dashed green) $D^{*+}D^-$ and (solid blue) $D^{*-}D^+$ data samples and (similar colours) the projections of the PDFs.

measured to be small, as ΔS_{D^*D} is also compatible with zero. Finally, the time- and flavour-integrated asymmetry is compatible with zero.

The decay-time-dependent signal yield asymmetry is defined as

$$A_{\text{yield}}(t) = \frac{N_{\overline{B}^0}(t) - N_{B^0}(t)}{N_{\overline{B}^0}(t) + N_{B^0}(t)}\,, \tag{6.13}$$

where N is the signal yield. It is shown for both $D^{*\pm}D^{\mp}$ charge configurations in Fig. 6.14. The almost pure sine wave can be seen with almost no contribution from the cosine.

The results can be used to assess the contributions from higher-order Standard Model effects in $B \to D\overline{D}$ decays and calculate the shift $\Delta\phi_{d,D^*D}$ of the mixing phase ϕ_d due to these effects. Neglecting the strong phase difference of the $B^0 \to D^{*+}D^-$ and $B^0 \to D^{*-}D^+$ decays and using Eq. (2.52), the effective phase results in

$$\sin\phi^{\text{eff}}_{d,D^*D} = 0.86 \pm 0.08.$$

Using the world average from $B^0 \to \psi(nS)K^0_{\text{S}}$ decays, $\sin\phi_d = 0.699 \pm 0.017$ [43], the phase shift is calculated to be

$$\Delta\phi_{d,D^*D} = (0.27 \pm 0.16)\,\text{rad}\,.$$

Thus, contributions from higher-order Standard Model effects are constrained to be small, which is an important input for the measurement of ϕ_s in $B^0_s \to D^+_s D^-_s$ decays.

7 Measurement of *CP* violation in $B^0 \to D^+D^-$ decays

The analysis of the $B^0 \to D^+D^-$ decay aims to measure the *CP* observables $S_{D^+D^-}$ and $C_{D^+D^-}$. The Run 2 data set collected by the LHCb detector is used while the Run 1 data set was analysed in Ref. [113]. Additional analyses were conducted by the BaBar and Belle collaborations [116, 117]. The analysis is still ongoing and blind. Currently, systematic uncertainties are evaluated while the rest of the analysis is already finished.

The analysis is performed in collaboration with Louis Gerken and Antje Mödden from the LHCb group from Dortmund. In parallel to the analysis of the signal decay, an analysis of the $B_s^0 \to D_s^+D_s^-$ decay is conducted by the same authors. This parallel procedure profits from many similarities of the two decays, but only the analysis of the $B^0 \to D^+D^-$ decay is presented here. Also, the flavour-specific $B^0 \to D_s^+D^-$ control channel is selected in parallel, which will be used to calibrate the FT response. The author contributed to all parts of the analysis with only smaller inputs to the figure-of-merit scan of the BDT response and the FT calibration. The FT calibration is based on the bachelor's thesis of Aron Kordt [133], which was supervised by the authors of the presented analysis. All parts of the analysis are described to give a coherent overview of the analysis.

The selection is discussed in Sec. 7.1 and the mass fit is described in Sec. 7.2. These chapters describe the procedures for the $B^0 \to D^+D^-$ and $B^0 \to D_s^+D^-$ decays simultaneously. The fit to the decay time is described in Sec. 7.3. Finally, a summary of significant expected systematic uncertainties is given in Sec. 7.4.

7.1 Selection

The data set collected by the LHCb detector in Run 2 is used. It corresponds to an integrated luminosity of $6\,\text{fb}^{-1}$ collected at a centre-of-mass energy of 13 TeV over the years 2015–2018. The decays are reconstructed using two decay modes of the D^+ meson: the $D^+ \to K^-\pi^+\pi^+$ decay and the $D^+ \to K^-K^+\pi^+$ decay. The former is roughly ten times more likely than the latter [41]. Still, the statistical power of the data set is increased by 20%, as either of the two $D^\pm$ mesons is allowed to decay via the $D^+ \to K^-K^+\pi^+$ mode. The contribution from decays where both

D^+ mesons decay into the less likely decay is very small and not retained. Thus, the two contributing final state combinations are

- $D^+ \to K^-\pi^+\pi^+$ and $D^- \to K^+\pi^-\pi^-$,

and

- $D^\pm \to K^\mp K^\pm \pi^\pm$ and $D^\mp \to K^\pm \pi^\mp \pi^\mp$.

The D_s^+ meson is reconstructed in three channels: the $D_s^+ \to K^-K^+\pi^+$ decay is the most common one and roughly five times more likely than the $D_s^+ \to \pi^-\pi^+\pi^+$ decay and eight times more likely than the $D_s^+ \to \pi^-K^+\pi^+$ decay. The additional D_s^+ decays increase the statistical power of the control channel moderately and are mainly considered due to their relevance in the $B_s^0 \to D_s^+ D_s^-$ analysis. Not all combinations of the mentioned D_s^+ and D^- decays are retained in the reconstruction of the $B^0 \to D_s^+ D^-$ decay. Either the D_s^+ or the D^- meson has to decay via its most common decay, leaving the four combinations

- $D_s^+ \to K^-K^+\pi^+$ and $D^- \to K^+\pi^-\pi^-$,
- $D_s^+ \to \pi^-\pi^+\pi^+$ and $D^- \to K^+\pi^-\pi^-$,
- $D_s^+ \to \pi^-K^+\pi^+$ and $D^- \to K^+\pi^-\pi^-$,

and

- $D_s^+ \to K^-K^+\pi^+$ and $D^- \to K^+K^-\pi^-$,

ordered from most to least likely.

7.1.1 Preselection

The centralised requirements are almost identical as in the $B^0 \to D^{*\pm} D^\mp$ analysis. An explanation of the variables can be found in Sec. 6.1.1. The only additional requirement is on the difference of the z coordinates of the $D^\pm$ and B^0 vertices $z(D^\pm) - (B^0)$. This ensures that the $D^\pm$ candidate decays downstream of the B^0 decay. The central-preselection efficiency of both decays is roughly 1.6 to 1.7%.

The mass windows of the D mesons are tightened to reduce combinatorial background. To suppress single-charm background, the source of the dominant systematic uncertainty in the Run 1 analysis, requirements on the flight distance of the D mesons are applied. The flight distance requirements lead to a relatively low efficiency in the preselection.

The efficiency of the additional requirements on simulated decays is about 76% and 69% for $B^0 \to D^+ D^-$ and $B^0 \to D_s^+ D^-$ decays, respectively. The lower efficiency for the control channel comes from the shorter average flight distance of D_s^+ mesons compared to D^+ mesons. About 88–90% of the combinatorial background

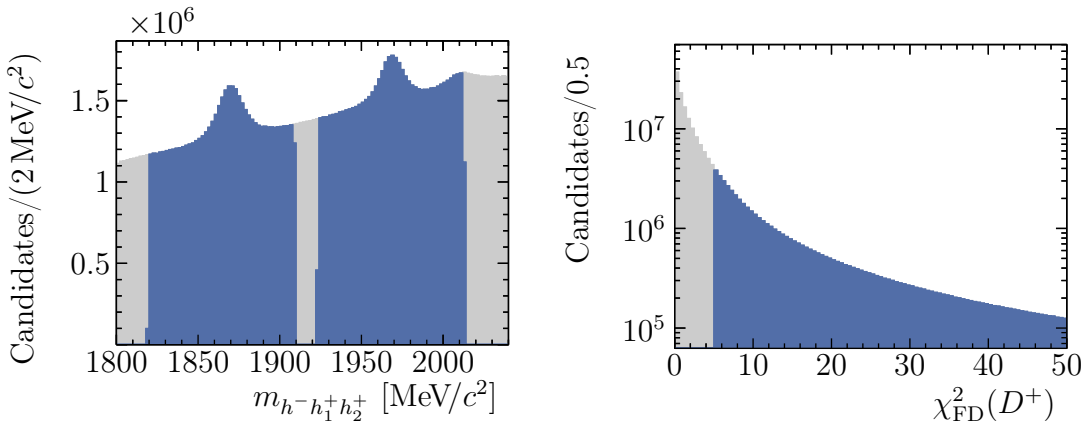

Figure 7.1: Effect of the preselection requirements and the resulting distributions: the invariant (left) $D^+_{(s)}$ mass and (right) the flight distance quality of the $D^+_{(s)}$ candidate. The grey data is rejected by the selection requirement and the blue data is retained. The distributions for the negative $D^-_{(s)}$ candidates look similar. The mass distribution shows the requirements for D^+ and D^+_s candidates. A smaller third contribution at the known D^{*+} mass is visible, which interferes with the D^+_s mass region. These candidates will be rejected with veto requirements.

is suppressed, evaluated using the upper-mass sideband with $m_{B^0} > 5600\,\text{MeV}/c^2$. The distributions of the variables used in the preselection are shown in Fig. 7.1.

7.1.2 Vetoes

Veto requirements are used to suppress backgrounds from other decays. These include mainly misidentified decays. Most single-charm backgrounds are suppressed by the requirement on the flight distance in the preselection. The remaining single-charm backgrounds also have misidentified final state particles, which is used to reduce them further with PID requirements. The veto requirements are similar to the ones presented in Sec. 6.1.2. They use requirements on the ProbNN_{hh}, defined in Eq. (6.1), in mass windows around the known mass of the background hadron. The ProbNN values are adjusted using the same kernel-density-based approached as described in Sec. 6.1.2. The cut-based veto requirements for the $D^\pm$ mesons are listed in Table 7.1. The effects of the requirements for the $B^0 \to D^+D^-$ decay are shown in Figs. 7.2 and 7.3. Several additional requirements are performed on the D^+_s mesons of the $B^0 \to D^+_s D^-$ decay, suppressing mainly misidentified backgrounds from D^0, D^+ and Λ^+_c decays along with D^+_s decays into other final states. The cut-based veto requirements have signal efficiencies of 92–98% for

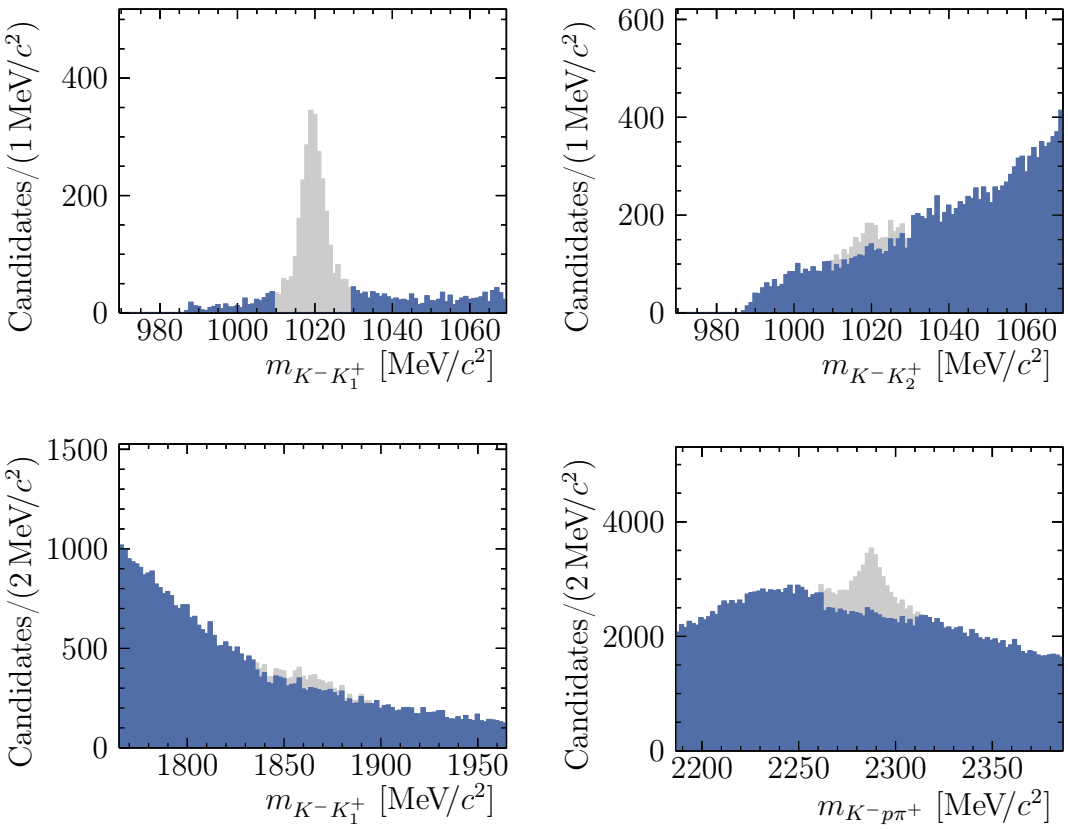

Figure 7.2: Invariant mass distributions of $D^+ \to K^-\pi^+\pi^+$ decays under alternative mass hypotheses for the final state particles. The h_1^+ hadron is the pion with the higher p_T compared to the h_2^+ hadron. The grey data is rejected by the veto requirement and the blue data is retained. The misidentified backgrounds at the known ϕ, D^0 and Λ_c^+ masses are suppressed by the veto requirements. The requirements are listed in the first part of Table 7.1, the order here and in the table is the same. Only data for the $D^\pm$ meson from the $B^0 \to D^+ D^-$ decay with the higher p_T is shown. Most misidentified backgrounds are less prominent but still present in the data for the $D^\mp$ meson with the lower p_T. The combinatorial BDT, presented in Sec. 7.1.3, is applied to enhance the misidentified backgrounds.

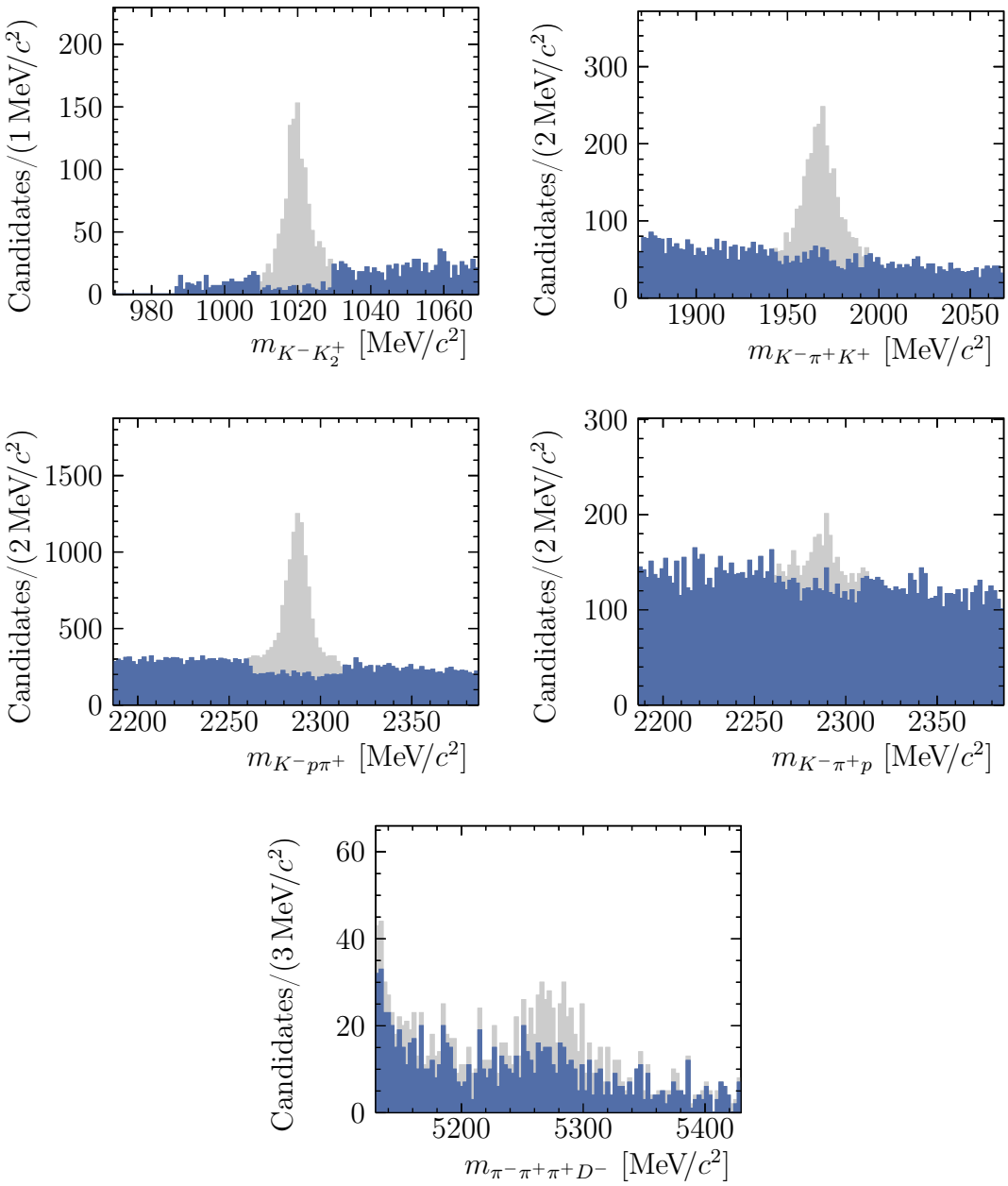

Figure 7.3: Invariant mass distributions of $D^+ \to K^-K^+\pi^+$ decays under alternative mass hypotheses for the final state particles. The h_1^+ hadron is the kaon and the h_2^+ hadron is the pion with the same charge. The grey data is rejected by the veto requirement and the blue data is retained. The misidentified backgrounds at the known ϕ, D_s^+, Λ_c^+ and B^0 masses are suppressed by the veto requirements. The requirements are listed in the second part of Table 7.1, the order here and in the table is the same. The combinatorial BDT, presented in Sec. 7.1.3, is applied to enhance the misidentified backgrounds.

Table 7.1: Cut-based veto requirements per D^+ final state. The final state is given as $h^-h_1^+h_2^+$, where h^- has the opposite charge of the other two hadrons, h_1^+ is either the kaon with the same charge or the pion with the higher p_T, and h_2^+ is the remaining pion. Candidates are rejected if they lie inside the given mass window around the known masses [41] of the background particle and if they do not fulfil the ProbNN$_{hh}$ requirement. If two ProbNN$_{hh}$ requirements are listed, the candidate has to fulfil both requirements. The requirements are applied to all $D^\pm$ mesons from the $B^0 \to D^+D^-$ and $B^0 \to D_s^+D^-$ decays. The requirements in parentheses are only applied on the $D^\pm$ candidate with the higher p_T.

$h^-h_1^+h_2^+$	Background	Window	ProbNN requirement
$K^-\pi^+\pi^+$	$\phi \to K^-K_1^+$	10 MeV/c^2	–
	$\phi \to K^-K_2^+$	10 MeV/c^2	$\pi K(h_2^+) > 0.95$
	$D^0 \to K^-K_1^+$	30 MeV/c^2	$\pi K(h_1^+) > 0.8$
	$\Lambda_c^+ \to K^-p\pi^+$	25 MeV/c^2	$\pi p(h_1^+) > 0.8$ (0.3)
$K^-K^+\pi^+$	$\phi \to K^-K_2^+$	10 MeV/c^2	$\pi K(h_2^+) > 0.98$
	$D_s^+ \to K^-\pi^+K^+$	25 MeV/c^2	$\pi K(h_1^+) < 0.1$
			$\pi K(h_2^+) > 0.9$
	$\Lambda_c^+ \to K^-p\pi^+$	25 MeV/c^2	$Kp(h_1^+) > 0.7$
	$\Lambda_c^+ \to K^-\pi^+p$	25 MeV/c^2	$\pi K(h_1^+) < 0.1$
			$\pi p(h_2^+) > 0.9$
	$B^0 \to \pi^-\pi^+\pi^+D^-$	–	$(\pi K(h^-) < 0.5)$
			$(\pi K(h_1^+) < 0.5)$

$B^0 \to D^+D^-$ and $B^0 \to D_s^+D^-$ decays, respectively. Only the $D_s^+ \to \pi^-K^+\pi^+$ category has a reduced efficiency of 83%.

The $D^+ \to K^-\pi^+\pi^+$ final state

Misidentified backgrounds in the $D^+ \to K^-\pi^+\pi^+$ final state come from similar sources as in the similar final state in the $B^0 \to D^{*\pm}D^\mp$ analysis. Contributions from $\phi \to K^-K^+$ decays arise if the kaon is identified as a pion. If the kaon mass is assigned to the pion with the higher p_T, all candidates in the mass window are rejected. If the kaon mass is assigned to the pion with the lower p_T, rejecting all candidates in the mass windows leads to a reduced signal efficiency. Instead, these candidates have to fulfil requirements on the ProbNN$_{\pi K}$ ratio. Contributions from $D^0 \to K^-K^+$ and $\Lambda_c^+ \to K^-p\pi^+$ decays arise also if the kaon respectively proton is identified as a pion. They are suppressed by rejecting candidates in the respective mass window if they do not fulfil ProbNN$_{\pi h}$ requirements after assigning the kaon respectively proton mass to the high-p_T pion. Applying the mass hypotheses to the low-p_T pion shows no misidentified background contributions.

The $D_s^+ \to K^-K^+\pi^+$ decay can be identified as the signal decay, when the positive kaon is identified as a pion. This decay is suppressed with a BDT, as described below.

The $D^+ \to K^-K^+\pi^+$ final state

Contributions from doubly-misidentified $D^+ \to K^-\pi^+K^+$ decays arise if the positive kaon is identified as a pion and the positive pion is identified as a kaon. They are suppressed by calculating the related invariant mass and performing $\text{ProbNN}_{\pi K}$ requirements on both positive hadron in a mass window around the known D_s^+ mass. This decay can also involve $\phi \to K^-K^+$ decays, which are used to apply stricter requirements in the respective mass window around the known ϕ mass. The $\Lambda_c^+ \to K^-p\pi^+$ can be identified as the signal decay if the proton is identified as a kaon, which is a single misidentification, or the proton is misidentified as a pion and the pions is misidentified as a kaon. These candidates are suppressed by rejecting candidates in the respective mass windows using one respectively two ProbNN_{hh} requirements. Finally, there is a contribution from misidentified single-charm $B^0 \to \pi^-\pi^+\pi^+D^-$ decays. This contribution only occurs for the $D^\pm$ meson with the higher p_T and is not visible for the other $D^\mp$ meson. It is rejected by applying $\text{ProbNN}_{\pi K}$ requirements for both kaons for the whole mass range.

Multivariate separation of $D^+ \to K^-\pi^+\pi^+$ and $D_s^+ \to K^-K^+\pi^+$ decays

Using similar strategies to reduce misidentification between $D^+ \to K^-\pi^+\pi^+$ and $D_s^+ \to K^-K^+\pi^+$ decays is either not efficient or cannot suppress the background completely, as is the case for the $B^0 \to D^{*\pm}D^\mp$ analysis. This is due to the similar mass differences between D_s^+ and D^+ mesons, and K^+ and π^+ mesons. When calculating the invariant mass of the D^+ decay and assigning the kaon mass to one of the pions, the distribution overlaps to a large extent with the distribution of correctly identified D_s^+ candidates. Thus, the use of a mass window has no effect and only the PID requirement can be used to suppress the decay. Instead, a multivariate approach is employed. Gradient boosting [109] with the XGBOOST framework [112] is utilised. Simulated $B^0 \to D^+D^-$, $B^0 \to D_s^+D^-$ and $B_s^0 \to D_s^+D_s^-$ decays are merged to a single sample and used in the training. The previous selection steps are performed on the data set. The sample is then divided in two categories: in one category the respective D meson is generated as a D^+ meson that decays via the $D^+ \to K^-\pi^+\pi^+$ decay, and in the other category the D meson is generated as a D_s^+ meson that decays via the $D_s^+ \to K^-K^+\pi^+$ decay. A five-folding approach is used, where four fifths of the data are used to train a BDT, which is then applied on the remaining fifth of the data [122]. This is repeated until the BDT is applied on the whole data set.

Table 7.2: List of input features used in the veto BDT training. The features are roughly sorted by their importance and explained in the text. The final state is defined as $D^+ \to K^- h_1^+ h_2^+$. One of the $h_{1,2}$ hadrons is a kaon while the other is a pion in case of a $D_s^+ \to K^- K^+ \pi^+$ decay and both are pions for $D^+ \to K^- \pi^+ \pi^+$ decays.

Input features
$m(K^- \pi_1^+ \pi_2^+)$
$\text{ProbNN}_K(h^+_{(1,2)})$
$m(K^- \pi_1^+ K_2^+)$
$m(K^- K_1^+ \pi_2^+)$
$\text{ProbNN}_\pi(h^+_{(1,2)})$
$m(K^- \pi^+_{(1,2)})$
$m(K^- K^+_{(1,2)})$
$\text{FD}(D^+_{(s)})$

The input features use the known differences of the decays. The D^+ meson has a significantly longer flight distance FD than the D_s^+ mesons on average. The ProbNN variables of the two final state particles with the same charge are additional input features. These are the particles that are potentially misidentified. Finally, the D decays can involve different intermediate resonances or at least have different probabilities to involve these resonances, *e.g.* ϕ and K^{*0} resonances [41]. These resonances are present in the simulated data and are exploited in the BDT. The invariant mass m of various final-state combinations is added to the BDT. This is done using pion and kaon mass hypotheses for the potentially misidentified particles. All input features are listed in Table 7.2. The agreement between the distributions of the BDT features in simulated and background-subtracted data is examined. The distributions are compatible. As discussed in Sec. 6.1.3, the agreement between simulated and recorded data is primarily important for a good BDT performance, as the calculation of efficiencies is not needed in decay-time-dependent analyses.

The hyperparameters are taken from the BDT that suppresses combinatorial background, presented in the following section. Changing the hyperparameters has no significant effect on the ROC score. Each BDT has at most 1000 decision trees and a learning rate of 0.5. Each decision tree has a depth of three. To avoid overfitting, no new decision trees are added if the ROC score is not increased by adding ten more decision trees. The distribution of the classifiers and the ROC curve are shown in Fig. 7.4. The BDT requirements are chosen so that almost all of the misidentified decays are rejected. They reject significantly more than 99% of the misidentified backgrounds in simulated data samples. The signal efficiency is roughly 88–89% for decays where the requirement has to be applied to both D mesons. For $B^0 \to D^+ D^-$ decays, this happens when both D mesons decay via the

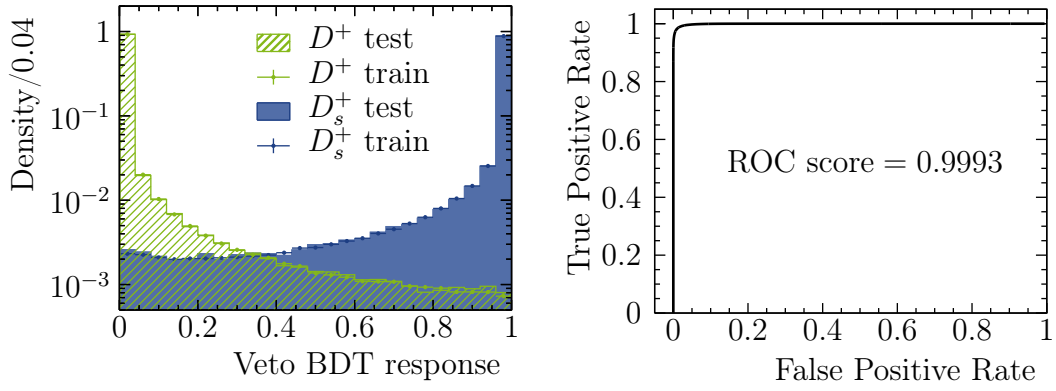

Figure 7.4: The distribution of the (green) D^+ and (blue) D_s^+ BDT response is shown on the left. The (shaded areas) test and (points) training distributions are compatible, which is an indicator that there is no overfitting. The ROC curve is shown on the right. The results for D^- candidates look similar.

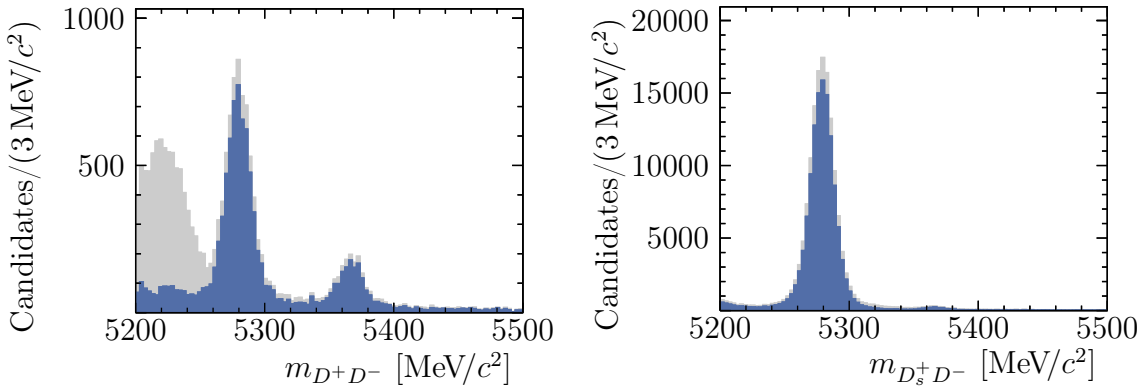

Figure 7.5: Invariant (left) D^+D^- and (right) $D_s^+D^-$ mass distributions before and after application of the veto BDT requirement. The grey data is rejected by the selection requirement and the blue data is retained. The combinatorial BDT, presented in Sec. 7.1.3, is applied to enhance the misidentified backgrounds. In the left plot, a large contribution from misidentified $B^0 \to D_s^+D^-$ background is suppressed that lies at masses below 5250 MeV. The reduction of the signal contribution at the known B^0 mass, a potential $B_s^0 \to D^+D^-$ contribution at the known B_s^0 mass, and the combinatorial background is much smaller. In the right plot, the contribution of the $B^0 \to D_s^+D^-$ control channel dominates the data set. Contributions from misidentified $B^0 \to D^+D^-$ or $B_s^0 \to D_s^+D_s^-$ decays are not apparent.

Table 7.3: List of input features used in the BDT training. They are roughly sorted by their importance and explained in the text.

Input features
$\chi^2_{\text{IP}}(B^0)$
$\lvert m(D^\pm) - m_{D^{*+},\text{PDG}}\rvert$
$\chi^2_{\text{IP}}(D^\pm)$
$\chi^2_{\text{DTF}}(B^0)$
$p_{\text{T}}(\pi^\pm/K^\pm)$
$p_{\text{T}}(D^\pm)$
$\chi^2_{\text{FD}}(D^\pm)$
$\cos\angle(D^\pm, \pi^\pm/K^\pm)$

$D^\pm \to K^\mp\pi^\pm\pi^\pm$ channel, and for $B^0 \to D_s^+D^-$ decays when the D mesons decay via the $D_s^+ \to K^-K^+\pi^+$ and $D^- \to K^+\pi^-\pi^-$ channels. In the other final-state configurations, the BDT requirement is only applied on one D meson and the efficiency is roughly 94–95%. The effect on the invariant B^0 mass distributions is shown in Fig. 7.5.

7.1.3 Multivariate analysis

Remaining combinatorial background is reduced with a multivariate analysis. Gradient boosting [109] via the XGBOOST framework is used [112]. The same BDT is trained simultaneously for $B^0 \to D^+D^-$, $B^0 \to D_s^+D^-$ and $B_s^0 \to D_s^+D_s^-$ decays. This allows to use the same BDT requirements on the signal and control channel for both analyses and retains the similarities of the kinematic distributions, which is important *e.g.* in the FT calibration. Simulated data of the three decays are used as the signal proxy. The upper-mass sideband is used as the background proxy. The previous selection steps are performed on all data sets. The BDT is trained five-folded [122].

First, the input features and then the hyperparameters are found iteratively as described in Sec. 6.1.3. The input features are listed in Table 7.3. They include the χ^2_{IP} variables of the B^0 and $D^\pm$ mesons to check whether they originate in the PV, the distances of the invariant $D^\pm$ masses to the known mass $|m(D^\pm) - m_{D^{*+},\text{PDG}}|$, the goodness of the fit to the whole decay chain χ^2_{DTF}, the transverse momenta of the final-state particles and the $D^\pm$ mesons, the χ^2_{FD} variables of the long-lived $D^\pm$ mesons and the angles between the $D^\pm$ mesons and their decay products. The accurate simulation of the variable distributions is checked using background-subtracted data. At most 1000 decision trees are added to a BDT. The addition of new trees is stopped if the ROC score is not increased after ten additional trees.

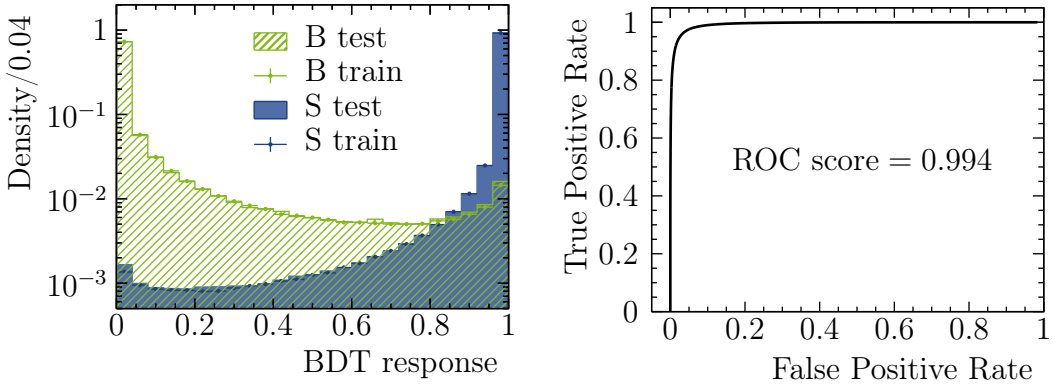

Figure 7.6: The distribution of the (green) signal and (blue) background BDT response is shown on the left. The (shaded areas) test and (points) training distributions are compatible, which is an indicator that there is no overfitting. The ROC curve is shown on the right.

The learning rate is 0.5. The distribution of the BDT response and the ROC curve are shown in Fig. 7.6.

The optimal requirement on the BDT response is found with the same procedure and figure of merit as described in Sec. 6.1.3. The same requirement is applied to the $B^0 \to D^+D^-$ and $B^0 \to D_s^+D^-$ samples, to ensure similar kinematic distributions. The efficiencies of the BDT requirement are 86–87% and 84–88% on simulated $B^0 \to D^+D^-$ and $B^0 \to D_s^+D^-$ decays, respectively, depending on the final state. About 99% of the upper-mass sideband is suppressed for all final states.

7.1.4 Multiple candidates

After the selection about 1.5% of the events contain more than one signal candidate. These are rejected randomly to avoid any bias.

7.2 Mass fit

The sPlot method is used to subtract background candidates. A fit to the mass distribution of the selected data samples is performed in the ranges 5200–5500 MeV/c^2 and 5220–5520 MeV/c^2 for the $B^0 \to D^+D^-$ and $B^0 \to D_s^+D^-$ samples, respectively. These ranges exclude partially reconstructed background. This was not possible in the $B^0 \to D^{*\pm}D^{\mp}$ analysis, as the model of the misidentified background becomes unstable with such a narrow region. The use of the veto BDTs

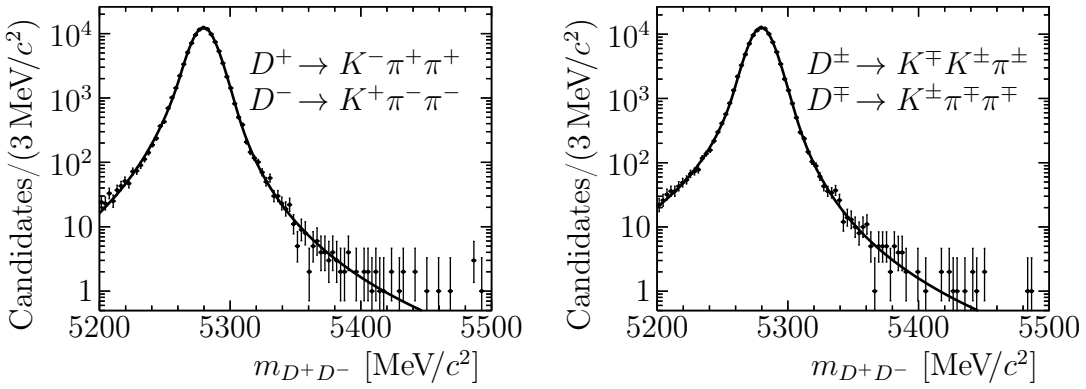

Figure 7.7: Mass distributions and PDF projections of simulated $B^0 \to D^+D^-$ decays for both final-state categories. The total PDF consists only of the Hypatia contribution, shown in solid black.

allows to suppress the respective misidentified background in this analysis completely. The only remaining backgrounds are the $B_s^0 \to D^+D^-$ and $B_s^0 \to D_s^+D^-$ decays for the signal and calibration channel, respectively. A separate fit is performed for each final-state configuration. The extended PDF describing the mass distributions is

$$N\mathcal{P} = N_{B^0}\mathcal{P}_{B^0} + N_{B_s^0}\mathcal{P}_{B_s^0} + N_{\text{Comb}}\mathcal{P}_{\text{Comb}}\,, \tag{7.1}$$

where N is the number of candidates and $\mathcal{P}$ the respective PDF.

7.2.1 Determination of the mass shapes

The B^0 and B_s^0 contributions are modelled with Hypatia functions [134]:

$$\mathcal{P}_{B^0}(m) = \mathcal{I}(m; \mu, \sigma, \lambda, \zeta, \beta, \alpha_1, n_1, \alpha_2, n_2)\,. \tag{7.2}$$

The Hypatia function is the generalisation of the Crystal Ball function with a hyperbolic core and one tail to each side. The parameters ζ and β describe asymmetries in the core and are fixed to 0. The parameter λ models the sharpness of the core. The further parameters have the same effect as for the Crystal Ball function. The parameters μ and σ are the mean and width of the model. The exponents of the power-law tails are $n_{(1,2)}$ and the transitions of the core to the power-law tails are defined by $\alpha_{(1,2)}$.

The projections of the PDFs resulting from the maximum likelihood fit to simulated data and the mass distributions are shown in Figs. 7.7 and 7.8, for the $B^0 \to D^+D^-$ and $B^0 \to D_s^+D^-$ samples, respectively. The fit results are used to fix the parameters in the fit to recorded data, where only the mean and the width are

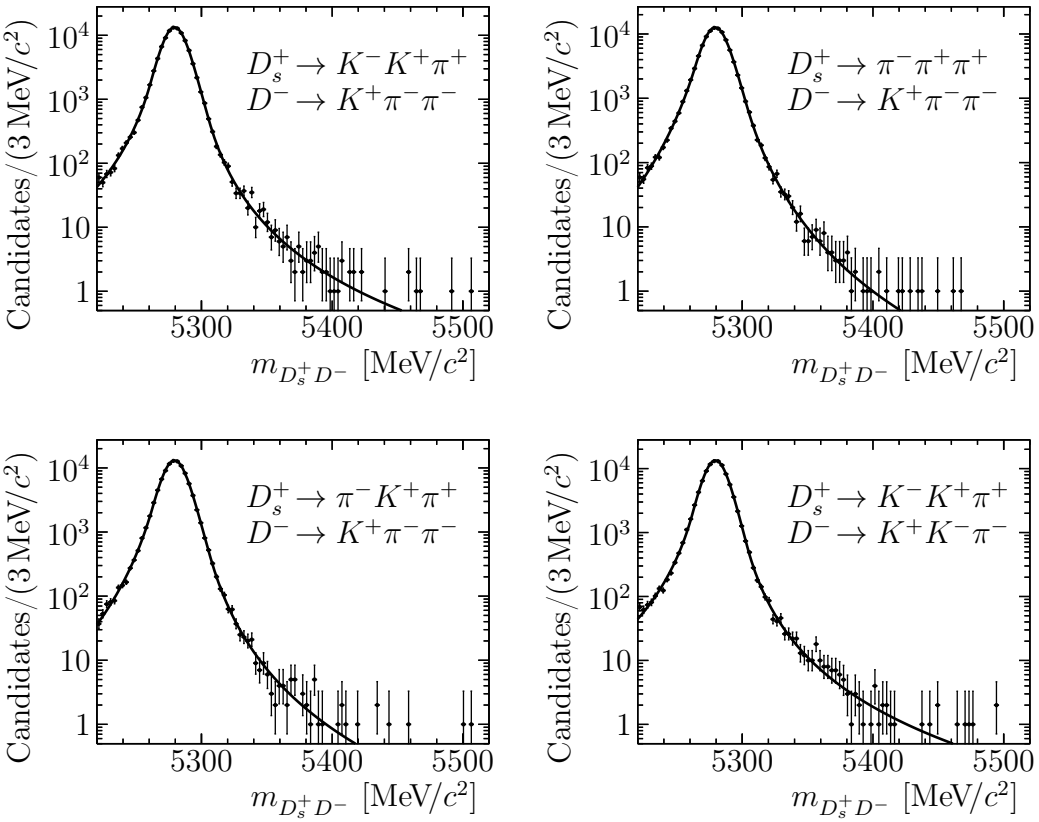

Figure 7.8: Mass distributions and PDF projections of simulated $B^0 \to D_s^+ D^-$ decays for all final-state categories. The total PDF consists only of the Hypatia contribution, shown in solid black.

free. The shape of the B_s^0 backgrounds is set to the same value as for the B^0 signal. The mean is fixed to the known mass difference $\mu_{B_s^0} - \mu_{B^0} = (87.23 \pm 0.24)\,\text{MeV}$ [41].

7.2.2 Fit to recorded data

An extended maximum likelihood fit is used to fit the PDF describing Eq. (7.1) to the mass distribution of recorded data. The free parameters are the mean of the B^0 contribution, the shared width of the B^0 and B_s^0 contributions, the slope of the combinatorial background, and the yields of all components. Figures 7.9 and 7.10 show the resulting PDF projections and mass distributions for the signal and control channel. The fit results of the fits to the $B^0 \to D^+D^-$ distributions is listed in Table 7.4. The mean values are compatible with the known B^0 mass [41] and the widths are compatible with the results on simulated data. The fit results in 5525 ± 85 $B^0 \to D^+D^-$ candidates and $120\,138 \pm 364$ $B^0 \to D_s^+D^-$

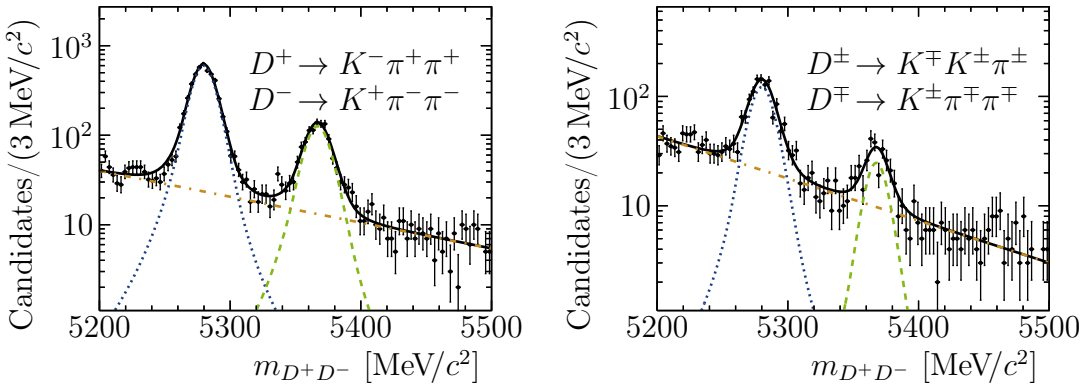

Figure 7.9: Mass distributions and PDF projections of the data samples with the two $B^0 \to D^+D^-$ final-state categories. Also shown are the projections of the (dotted blue) $B^0 \to D^+D^-$ decay, (dashed green) $B^0_s \to D^+D^-$ background and (dash-dotted orange) combinatorial background.

Table 7.4: Results of the mass fit to recorded data to extract sWeights for both decays.

Parameter	$D^+ \to K^-\pi^+\pi^+$ $D^- \to K^+\pi^-\pi^-$	$D^\pm \to K^\mp\pi^\pm\pi^\pm$ $D^\mp \to K^\pm K^\mp\pi^\mp$
μ_{B^0} (MeV/c^2)	5279.61 ± 0.16	5280.13 ± 0.42
σ (MeV/c^2)	10.35 ± 0.17	10.98 ± 0.50
β_{Comb}	-6.6 ± 0.3	-8.9 ± 0.4
N_{B^0}	4541 ± 74	984 ± 42
$N_{B^0_s}$	1004 ± 37	185 ± 19
N_{Comb}	1748 ± 59	1507 ± 52

candidates. To perform the fit to the decay-time distribution, discussed in the next chapter, sWeights are determined. Also, 1189 ± 42 $B^0_s \to D^+D^-$ candidates and 1224 ± 48 $B^0_s \to D^+_s D^-$ candidates are measured. These can be used to update the measurements of the branching ratio of these decays, where the most precise measurements were performed using a subset of the LHCb data set [135, 136].

7.3 Decay-time fit

A fit to the decay-time distribution in the range 0.3–10.3 ps is performed to extract the *CP* observables. The data at lower decay times is almost completely suppressed by the central preselection. Almost all B^0 candidates decay before the upper boundary, so that choosing a higher boundary has no effect. The data sets

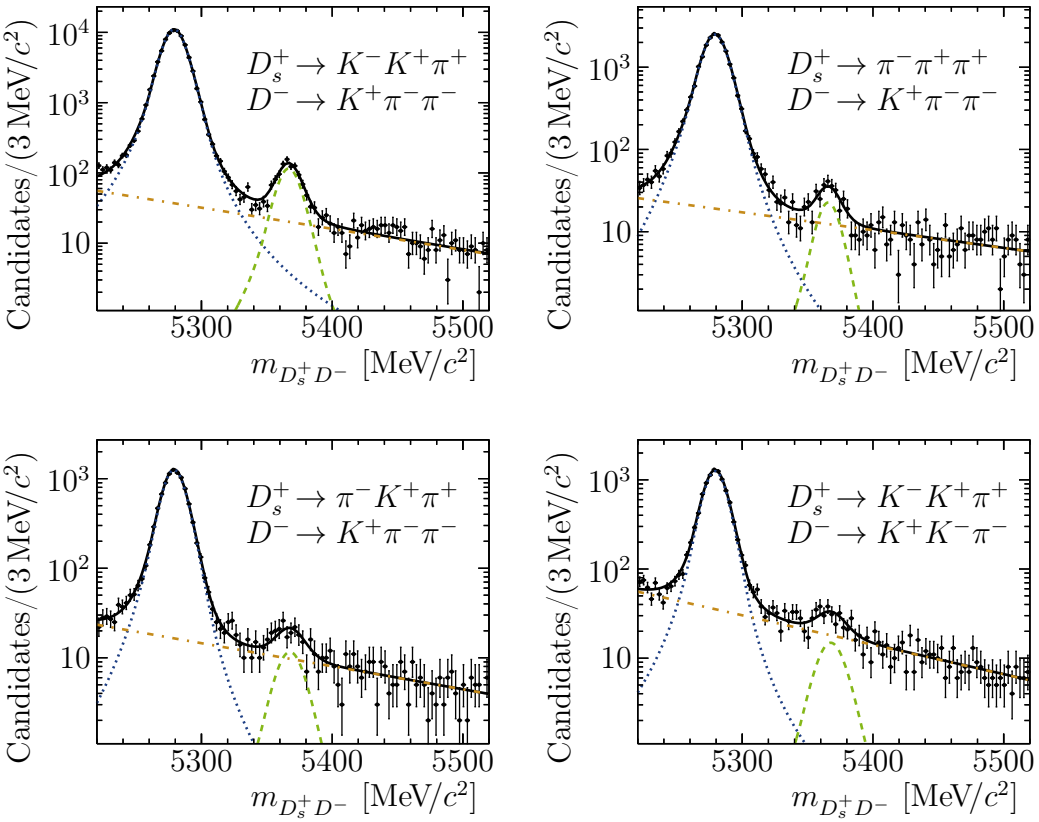

Figure 7.10: Mass distributions and PDF projections of the data samples with the four $B^0 \to D_s^+ D^-$ final-state categories. Also shown are the projections of the (dotted blue) $B^0 \to D_s^+ D^-$ decay, (dashed green) $B_s^0 \to D_s^+ D^-$ background and (dash-dotted orange) combinatorial background.

of all final states are merged and used simultaneously in the following analysis steps. In contrast to the $B^0 \to D^{*\pm} D^{\mp}$ analysis, instrumental asymmetries do not play a role, as the final state is symmetric.

7.3.1 Flavour-tagging calibration

The $B^0 \to D_s^+ D^-$ channel is used to calibrate the FT response. The statistical power is large enough to also calibrate the single taggers and no additional control channel is needed. The kinematic distributions of the signal and control channels match, and no adjustment is needed. The FT calibration is performed simultaneously for all $B^0 \to D^+ D^-$ final states and using all $B^0 \to D_s^+ D^-$ final states. All available FT algorithms, as defined in Sec. 3.4, are used.

The single taggers are calibrated using the EPM tool and then combined to obtain

Table 7.5: FT calibration parameters obtained with the EPM tool using all final states of the $B^0 \to D_s^+ D^-$ decay.

Parameter	Opposite side	Same side
ε_{tag} [%]	32.14 ± 0.13	88.37 ± 0.09
$p_0 - \langle\eta\rangle$	0.0180 ± 0.0036	-0.0013 ± 0.0022
p_1	0.884 ± 0.035	0.936 ± 0.033
Δp_0	0.0041 ± 0.0051	0.0004 ± 0.0031
Δp_1	0.012 ± 0.050	-0.052 ± 0.048

Table 7.6: Tagging performance parameters from the $B^0 \to D^+ D^-$ sample: the average of the predicted mistag probability $\langle\eta\rangle$, the tagging efficiency ε_{tag} and the effective tagging efficiency ε_{eff}. The uncertainty is statistical only.

Tagger	$\langle\eta\rangle$	ε_{tag} [%]	ε_{eff} [%]
OS	0.3462	32.76 ± 0.47	3.35 ± 0.08
SS	0.4258	88.10 ± 0.32	3.18 ± 0.05
total		91.25 ± 0.28	6.29 ± 0.10

the SS and OS combinations. The combinations are calibrated again using the EPM tool to obtain the calibration parameters that are Gaussian-constrained in the fit to the decay-time distribution. A repeat of the procedure used in the $B^0 \to D^{*\pm} D^{\mp}$ analysis is not necessary, *i.e.* a fit to the decay-time distribution of the $B^0 \to D_s^+ D^-$ sample. This was done to also extract the production asymmetry. The effect of the production asymmetry on the results of the *CP* parameters is shown to be small and the value of the production asymmetry determined in the $B^0 \to D^{*\pm} D^{\mp}$ analysis is reused in this analysis.

The evaluation of systematic uncertainties on the calibration parameters is still ongoing. The $B^0 \to D^{*\pm} D^{\mp}$ analysis has shown that the uncertainties on the calibration parameters only have a small effect on the uncertainties of the *CP* observables. The results of the calibration parameters are listed in Table 7.5. The calibration is applied to the signal sample to evaluate the performance. The performance parameters are listed in Table 7.6. The tagging power is measured to be $\varepsilon_{\text{eff}} = 6.29 \pm 0.10$.

7.3.2 Decay-time resolution

The difference of the true and reconstructed decay times is described by a PDF following the resolution function from Eq. (5.14). The PDF is fitted to simulated decays. The PDF projection and data distribution are shown in Fig. 7.11. The

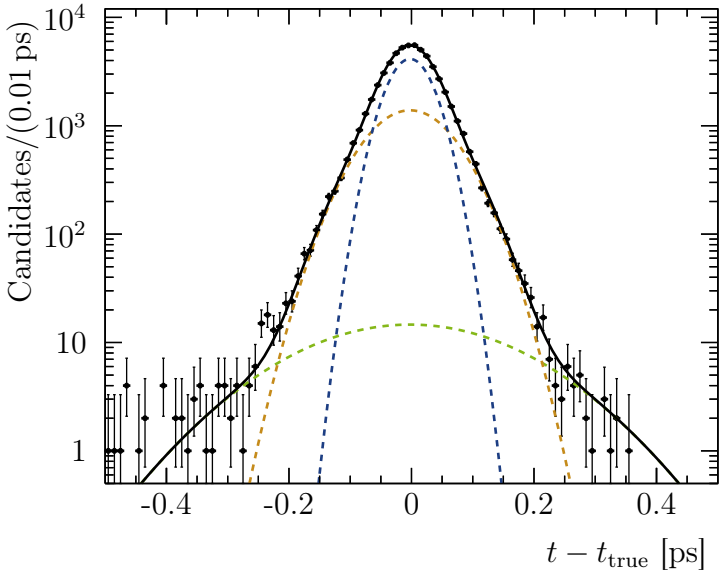

Figure 7.11: Distribution of the difference of the true and reconstructed decay time and the projection of the PDF describing the decay-time resolution. The sample contains all $D^{\pm}$ final states. The resolution PDF comprises three Gaussian functions. The total PDF projection is shown in solid black and the three components are shown in dashed blue, dashed orange and dashed green.

fit results are fixed in the maximum likelihood fit to recorded data. Using the results to calculate the effective decay-time resolution according to Eq. (6.10), yields $\sigma_{\text{eff}} \approx 53\,\text{fs}$.

7.3.3 Decay-time-dependent efficiency

The decay-time acceptance is modelled using the knots as in the $B^0 \to D^{*\pm}D^{\mp}$ analysis. The knot positions are 0.3, 0.5, 2.7, 6.3 and 10.3 ps. The model is validated by performing a maximum likelihood fit to the decay-time distributions of simulated and recorded data. The PDF corresponds to Eq. (5.13), but the part that depends on the *CP* parameters is exchanged for an exponential function decaying with the B^0 lifetime. The PDF projections and distributions are shown in Fig. 7.12. The coefficients are not fixed to the shown results, but are free in the fit to extract the *CP* observables.

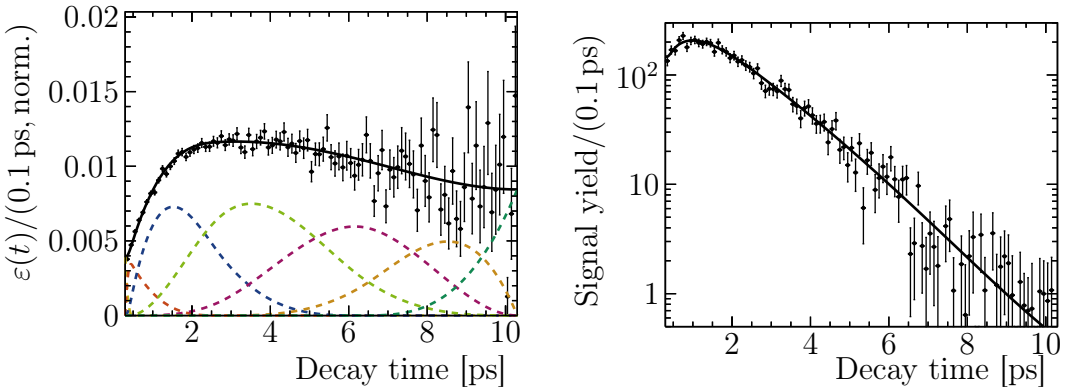

Figure 7.12: Projection of the PDF that describes the decay-time-dependent efficiency on the left with the distribution of simulated data, where the effect of the exponential decay is eliminated. The contributions from the different B-splines are shown in different colours. The projection of the PDF of the decay-time-dependent efficiency multiplied with an exponential decay with the B^0 lifetime, and the background-subtracted data is shown on the right. Both plots show samples containing all $D^{\pm}$ final states.

7.3.4 Extraction of CP observables

The missing inputs are the average mistag estimates and external parameters. The average mistag estimates are measured using the sWeighted signal samples. The results are $\langle \eta^{\text{OS}} \rangle = 0.3462$ and $\langle \eta^{\text{SS}} \rangle = 0.4258$. The external parameters τ_d and Δm_d are fixed to their known values [43] and $\Delta\Gamma_d$ is fixed to zero. Gaussian constraints are used for the production asymmetry and FT-calibration parameters. The coefficients of the decay-time-dependent efficiency are free in the fit.

The uncertainties from the maximum likelihood fit to the sWeighted data sample are

$$\sigma_{S_{D^+D^-}} = 0.10,$$
$$\sigma_{C_{D^+D^-}} = 0.10.$$

The statistical correlation is $\rho = 0.48$. The correlation of the other parameters to the CP observables is well below 1%. The projection of the PDF and the distribution of the data is shown in Fig. 7.13.

7.4 Systematic effects

The evaluation of systematic effects is still ongoing. The dominant source of systematic uncertainty in the Run 1 analysis come from neglecting single-charm back-

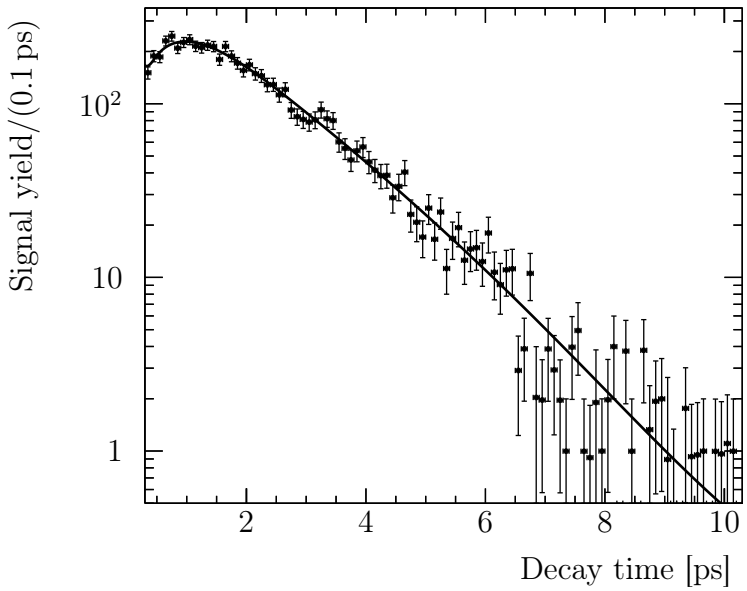

Figure 7.13: Decay time distribution of signal candidates and the projection of the PDF.

grounds in the fit to the B^0 mass. It is assumed that a significant number of $B \to D^- K^+ K^- \pi^-$ decays can be present, although the decay has still not been experimentally observed.. The invariant mass distribution of this decay is very similar to the signal mass shape. The contribution is estimated to be 2% of the signal yield. This is achieved by widening the D^+ mass window, but rejecting the inner part. This suppresses real D^+ meson almost completely, so that remaining candidates have to be single-charm backgrounds. Fits to the resulting invariant B^0 mass are performed. This yields 29 ± 20 candidates, which is roughly 2% of the total signal yield.

The presented, updated analysis introduced requirements on the flight distance of the $D^\pm$ mesons in the preselection, reducing any additional contribution from single-charm decays to a negligible level. The contribution of other sources of systematic uncertainty was evaluated to be almost an order of magnitude smaller in the Run 1 analysis, but no uncertainties due to the flavour tagging were considered.

The dominant sources of systematic uncertainty in the similar $B^0 \to D^{*\pm} D^\mp$ analysis come from the flavour tagging: the first source is the difference of the SS calibration functions in simulated signal and control samples. The second source is the dependence of the OS tagging efficiency on the decay time. The FT is expected to also be the dominant source of systematic uncertainty in the updated $B^0 \to D^+ D^-$ analysis. Following the $B^0 \to D^{*\pm} D^\mp$ analysis, the total systematic uncertainties are expected to be in the low percent range, so that the result will

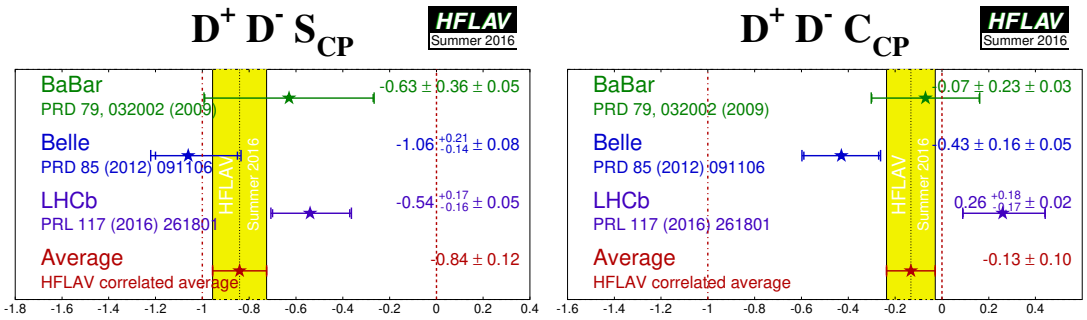

Figure 7.14: Results of the measurements of the CP observables (left) $S_{D^+D^-}$ and(right) $C_{D^+D^-}$ [43]. The previous measurements from the BaBar [116] and Belle [117] collaborations, and the result from the LHCb analysis of a smaller data set [113] are listed. They are compared in the text.

still be statistically limited.

7.5 Expected results

A measurement of CP violation in the $B^0 \to D^+D^-$ decay is presented. The analysis is ongoing and still blind. The statistical sensitivities are measured to be

$$\sigma_{S_{D^+D^-}} = 0.10,$$
$$\sigma_{C_{D^+D^-}} = 0.10.$$

The statistical correlation between the parameters is $\rho = 0.48$. Currently, systematic uncertainties are evaluated. They are expected to be in the low percent range by considering the systematic uncertainties of the previous LHCb analysis [113] and the similar $B^0 \to D^{*\pm}D^{\mp}$ analysis [34]. Once the evaluation of the systematic uncertainties is finished and the internal review process is finished, the results can be published. The sensitivity to the CP parameters will be the most precise measurement of CP violation in this channel and be competitive with the current world average, as can be seen in Fig. 7.14.

The presented analysis uses a similar procedure to the previous LHCb analysis, but introduces additional requirements to suppress single-charm backgrounds, the source of the largest systematic uncertainty, and a multivariate approach to suppress misidentified $B^0 \to D_s^+D^-$ background. The results will be used to assess the shift of the mixing phases in $B \to D\overline{D}$ like decays, especially the shift of ϕ_s in $B_s^0 \to D_s^+D_s^-$ decays.

8 Discussion and outlook

Two measurements of *CP* violation in different $B \to D\overline{D}$ decay channels are presented. The measurement of *CP* violation in the $B^0 \to D^{*\pm}D^{\mp}$ decay results in

$$\begin{aligned} S_{D^*D} &= -0.861 \pm 0.077\,(\text{stat}) \pm 0.019\,(\text{syst})\,, \\ \Delta S_{D^*D} &= 0.019 \pm 0.075\,(\text{stat}) \pm 0.012\,(\text{syst})\,, \\ C_{D^*D} &= -0.059 \pm 0.092\,(\text{stat}) \pm 0.020\,(\text{syst})\,, \\ \Delta C_{D^*D} &= -0.031 \pm 0.092\,(\text{stat}) \pm 0.016\,(\text{syst})\,, \\ A_{D^*D} &= 0.008 \pm 0.014\,(\text{stat}) \pm 0.006\,(\text{syst})\,. \end{aligned}$$

The largest statistical correlation are between the S_{D^*D} and C_{D^*D} parameters and the ΔS_{D^*D} and ΔC_{D^*D} parameters, with $\rho(S_{D^*D}, C_{D^*D}) = 0.44$ and $\rho(\Delta S_{D^*D}, \Delta C_{D^*D}) = 0.46$. This is the most precise measurement in this decay channel. It is compatible with the previous measurements by the BaBar [116] and Belle [117] collaborations. The uncertainties are close to the previous world average for C_{D^*D} and ΔC_{D^*D}, slightly better than the previous world average for S_{D^*D} and ΔS_{D^*D}, and much better for A_{D^*D} [43]. The results are used to calculate the shift of the mixing phase ϕ_d in this decay

$$\Delta\phi_{d,D^*D} = (0.27 \pm 0.16)\,\text{rad}\,.$$

This result is compatible with zero, which implies only a small contribution from higher-order corrections and is an important input for the measurement of ϕ_s in $B_s^0 \to D_s^+ D_s^-$ decays.

The analysis of the $B^0 \to D^+D^-$ decay is still ongoing and blind. The statistical sensitivities to the *CP* parameters are measured to be

$$\begin{aligned} \sigma_{S_{D^+D^-}} &= 0.10, \\ \sigma_{C_{D^+D^-}} &= 0.10. \end{aligned}$$

The correlation between the parameters is $\rho = 0.48$. The measurements will be the most precise measurement of *CP* violation in this decay. The systematic uncertainties are expected to be significantly smaller, which can be derived from the Run 1 analysis [113] or the similar measurement of *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays.

These measurements can be extended to perform updated branching-ratio measurements of the $B_s^0 \to D^+D^-$ and $B_s^0 \to D_s^+D^-$ decays. The current values were determined on only a small part of the LHCb data set and are statistically limited [135, 136]. Furthermore, the $B^0 \to D_s^+D_s^-$ decay has yet to be observed, which can be possible by adapting the presented selection. These measurements together with a series of additional measurements of *CP* violation and the branching ratio in $B \to D\overline{D}$ decays can be used to assess the phase shift in the measurement of β_s in $B_s^0 \to D_s^{(*)+}D_s^{(*)-}$ decays. Future measurements in these decays will come from the Belle II experiment [137] and from the upgraded LHCb experiment [138]. The LHCb experiment is scheduled to record data corresponding to an integrated luminosity of $50\,\text{fb}^{-1}$ in Run 3. The impact of the LHCb upgrade will be particularly high for $B \to D\overline{D}$ decays. Currently, these suffer from low efficiencies in the hardware trigger. After the upgrade, the trigger system will be exchanged to a GPU-based implementation, removing the hardware trigger completely [139]. Additionally, new developments in the flavour tagging will increase the sensitivity of decay-time-dependent analyses [140]. They are based on deep neural networks and show promising first results, increasing the effective tagging efficiency significantly, especially for B_s^0 decays. The main competitor of the LHCb experiment will be the Belle II experiment. The Belle II experiment will collect data corresponding to $50\,\text{ab}^{-1}$, which is 50 times larger than the data set from the Belle collaboration [137]. The combination of the results of both experiments will be crucial in the search for New Physics.

Bibliography

[1] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.

[2] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[3] A. Salam, *Weak and electromagnetic interactions*, Conf. Proc. **C680519** (1968) 367.

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

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

[6] Planck collaboration, N. Aghanim *et al.*, *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020) A6, Erratum ibid. **652** (2021) C4, arXiv:1807.06209.

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

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

[9] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, Sov. Phys. Usp. **34** (1991) 392.

[10] C. S. Wu *et al.*, *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

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

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

[13] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[14] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[15] R. Fleischer, *Extracting γ from $B_{s(d)} \to J/\psi K_S$ and $B_{d(s)} \to D^+_{d(s)} D^-_{d(s)}$*, Eur. Phys. J. **C10** (1999) 299, arXiv:hep-ph/9903455.

[16] BaBar collaboration, B. Aubert *et al.*, *Measurement of time-dependent CP asymmetry in $B^0 \to c\bar{c}K^{(*)0}$ decays*, Phys. Rev. **D79** (2009) 072009, arXiv:0902.1708.

[17] Belle collaboration, I. Adachi *et al.*, *Precise measurement of the CP violation parameter $sin2\phi_1$ in $B^0 \to (c\bar{c})K^0$ decays*, Phys. Rev. Lett. **108** (2012) 171802, arXiv:1201.4643.

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

[19] L. Bel *et al.*, *Anatomy of $B \to D\overline{D}$ decays*, JHEP **07** (2015) 108, arXiv:1505.01361.

[20] S. Stone and L. Zhang, *S-waves and the measurement of CP violating Phases in B_s decays*, Phys. Rev. **D79** (2009) 074024, arXiv:0812.2832.

[21] D0 collaboration, V. M. Abazov *et al.*, *Measurement of the CP-violating phase $\phi_s^{J/\psi\phi}$ using the flavor-tagged decay $B^0_s \to J/\psi\phi$ in 8 fb^{-1} of $p\bar{p}$ collisions*, Phys. Rev. **D85** (2012) 032006, arXiv:1109.3166.

[22] CDF collaboration, T. Aaltonen *et al.*, *Measurement of the bottom-strange meson mixing phase in the full CDF data set*, Phys. Rev. Lett. **109** (2012) 171802, arXiv:1208.2967.

[23] CMS collaboration, V. Khachatryan *et al.*, *Measurement of the CP-violating weak phase ϕ_s and the decay width difference $\Delta\Gamma_s$ using the $B^0_s \to J/\psi\phi(1020)$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. **B757** (2016) 97, arXiv:1507.07527.

[24] LHCb collaboration, R. Aaij *et al.*, *Updated measurement of time-dependent CP-violating observables in $B^0_s \to J/\psi K^+K^-$ decays*, Eur. Phys. J. **C79** (2019) 706, Erratum ibid. **C80** (2020) 601, arXiv:1906.08356.

[25] ATLAS collaboration, G. Aad *et al.*, *Measurement of the CP-violating phase ϕ_s in $B^0_s \to J/\psi\phi$ decays in ATLAS at 13* TeV, Eur. Phys. J. **C81** (2021) 342, arXiv:2001.07115.

[26] LHCb collaboration, R. Aaij *et al.*, *First measurement of the CP-violating phase in $B^0_s \to J/\psi(e^+e^-)\phi$ decays*, Eur. Phys. J. **C81** (2021) 1026, arXiv:2105.14738.

[27] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating phase ϕ_s in $\overline{B}^0_s \to D^+_s D^-_s$ decays*, Phys. Rev. Lett. **113** (2014) 211801, arXiv:1409.4619.

[28] S. Faller, R. Fleischer, and T. Mannel, *Precision physics with $B^0_s \to J/\psi\,\phi$ at the LHC: The quest for new physics*, Phys. Rev. **D79** (2009) 014005, arXiv:0810.4248.

[29] B. Bhattacharya, A. Datta, and D. London, *Reducing penguin pollution*, Int. J. Mod. Phys. **A28** (2013) 1350063, arXiv:1209.1413.

[30] A. Datta and D. London, *Extracting γ from $B^0_d(t) \to D^{(*)+}D^{(*)-}$ and $B^0_d \to D^{(*)+}D^{(*)-}$ decays*, Phys. Lett. **B584** (2004) 81, arXiv:hep-ph/0310252.

[31] J. Albert, A. Datta, and D. London, *A measurement of γ from the decays $B^0_d(t) \to D^{(*)+}D^{(*)-}$ and $B^0_d \to D^{(*)+}D^{(*)-}$*, Phys. Lett. **B605** (2005) 335, arXiv:hep-ph/0410015.

[32] R. Fleischer, *Exploring CP violation and penguin effects through $B^0 \to D^+D^-$ and $B^0_s \to D^+_s D^-_s$*, Eur. Phys. J. **C51** (2007) 849, arXiv:0705.4421.

[33] M. Jung and S. Schacht, *Standard model predictions and new physics sensitivity in $B \to DD$ decays*, Phys. Rev. **D91** (2015) 034027, arXiv:1410.8396.

[34] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B^0 \to D^{*\pm}D^{\mp}$ decays*, JHEP **03** (2020) 147, arXiv:1912.03723.

[35] LHCb collaboration, R. Aaij *et al.*, *Observation of the $B^0_s \to D^{*\pm}D^{\mp}$ decay*, JHEP **03** (2021) 099, arXiv:2012.11341.

[36] J. Langer, P. Ibis, and A. Mödden, *Search for the decay $B^0 \to D^{*+}D^{*-}$*, LHCb-PUB-2019-010, 2019.

[37] G. Lüders, *On the equivalence of invariance under time reversal and under particle-antiparticle conjugation for relativistic field theories*, Dan. Mat. Fys. Medd. **28** (1954) 1.

[38] W. Pauli and L. F. Curtiss, *Niels Bohr and the development of physics*, Am. J. Phys. **24** (1956) 292.

[39] D. J. Griffiths, *Introduction to elementary particles*, Wiley, 2nd ed., 2008.

[40] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*, Westview, 1995.

[41] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01.

[42] K. Anikeev *et al.*, *B physics at the Tevatron: Run II and beyond*, in *Workshop on B physics at the Tevatron: Run II and Beyond, Batavia, Illinois, September 23-25, 1999*, 2001, arXiv:hep-ph/0201071.

[43] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, Eur. Phys. J. **C81** (2021) 226, arXiv:1909.12524, updated results and plots available at http://hflav.web.cern.ch.

[44] G. C. Branco, L. Lavoura, and J. P. Silva, *CP violation*, Clarendon Press, 1999.

[45] I. I. Bigi and A. I. Sanda, *CP violation*, Cambridge University Press, 2nd ed., 2009.

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

[47] Belle collaboration, S. K. Choi *et al.*, *Observation of a narrow charmonium-like state in exclusive $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}J/\psi$ decays*, Phys. Rev. Lett. **91** (2003) 262001, arXiv:hep-ex/0309032.

[48] LHCb collaboration, R. Aaij *et al.*, *Observation of exotic $J/\psi\phi$ structures from amplitude analysis of $B^{+} \to J/\psi\phi K^{+}$ decays*, Phys. Rev. Lett. **118** (2017) 022003, arXiv:1606.07895.

[49] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^{+} \to J/\psi\phi K^{+}$ decays*, Phys. Rev. **D95** (2017) 012002, arXiv:1606.07898.

[50] LHCb collaboration, R. Aaij *et al.*, *Observation of structure in the J/ψ-pair mass spectrum*, Science Bulletin **65** (2020) 1983, arXiv:2006.16957.

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

[52] LHCb collaboration, R. Aaij *et al.*, *Observation of a narrow pentaquark state, $P_c(4312)^{+}$, and of two-peak structure of the $P_c(4450)^{+}$*, Phys. Rev. Lett. **122** (2019) 222001, arXiv:1904.03947.

[53] F. Englert and R. Brout, *Broken symmetry and the masses of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[54] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

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

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

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

[58] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP violation*, Phys. Rev. Lett. **55** (1985) 1039.

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

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

[61] NA31 collaboration, H. Burkhardt *et al.*, *First evidence for direct CP violation*, Phys. Lett. **B206** (1988) 169.

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

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

[64] LHCb collaboration, R. Aaij *et al.*, *Observation of CP violation in charm decays*, Phys. Rev. Lett. **122** (2019) 211803, arXiv:1903.08726.

[65] T. Jubb, M. Kirk, A. Lenz, and G. Tetlalmatzi-Xolocotzi, *On the ultimate precision of meson mixing observables*, Nucl. Phys. **B915** (2017) 431, arXiv:1603.07770.

[66] L. Evans and P. Bryant, *LHC machine*, Journal of Instrumentation **3** (2008) S08001.

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

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

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

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

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

[72] C. Lefèvre, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, http://cds.cern.ch/record/1260465, 2008.

[73] LHCb collaboration, C. Elsässer, *$b\bar{b}$ production angle plots*, http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[74] C. Ilgner *et al.*, *The Beam Conditions Monitor of the LHCb Experiment*, arXiv:1001.2487.

[75] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, arXiv:1708.00819.

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

[77] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13* TeV *pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140.

[78] LHCb collaboration, R. Aaij *et al.*, *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s}$ =13* TeV, JHEP **03** (2016) 159, Erratum ibid. **09** (2016) 013, Erratum ibid. **05** (2017) 074, arXiv:1510.01707.

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

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

[81] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, arXiv:1812.10790.

[82] LHCb collaboration, *Trigger schemes*, http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html.

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

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

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

[86] R. E. Kalman, *A new approach to linear filtering and prediction problems*, Journal of Basic Engineering **82** (1960) 35.

[87] R. Aaij *et al.*, *The* DaVinci *project*, http://lhcbdoc.web.cern.ch/lhcbdoc/davinci/.

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

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

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

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

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

[93] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

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

[95] R. Aaij *et al.*, *The* Boole *project*, http://lhcbdoc.web.cern.ch/lhcbdoc/boole/.

[96] D. Müller, M. Clemencic, G. Corti, and M. Gersabeck, *ReDecay: A novel approach to speed up the simulation at LHCb*, Eur. Phys. J. **C78** (2018) 1009, arXiv:1810.10362.

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

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

[99] LHCb collaboration, R. Aaij *et al.*, *New algorithms for identifying the flavour of B^0 mesons using pions and protons*, Eur. Phys. J. **C77** (2017) 238, arXiv:1610.06019.

[100] J. Wimberly *et al.*, *EspressoPerformanceMonitor*, http://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor.

[101] V. Blobel and E. Lohrmann, *Statistische und numerische Methoden der Datenanalyse*, Teubner, 1998.

[102] F. James and M. Roos, *Minuit: A system for function minimization and analysis of the parameter errors and correlations*, Comput. Phys. Commun. **10** (1975) 343.

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

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

[105] C. Langenbruch, *Parameter uncertainties in weighted unbinned maximum likelihood fits*, arXiv:1911.01303.

[106] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979) 1.

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

[108] B. P. Roe *et al.*, *Boosted decision trees as an alternative to artificial neural networks for particle identification*, Nucl. Instrum. Meth. **A543** (2005) 577, arXiv:physics/0408124.

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

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

[111] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, *TMVA - Toolkit for Multivariate Data Analysis with ROOT*, PoS **ACAT** (2007) 040.

[112] 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*, **KDD16** 785, 2016, arXiv:1603.02754.

[113] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B \to D^+ D^-$ decays*, Phys. Rev. Lett. **117** (2016) 261801, arXiv:1608.06620.

[114] LHCb collaboration, R. Aaij *et al.*, *Measurement of B^0, B_s^0, B^+ and Λ_b^0 production asymmetries in 7 and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139, arXiv:1703.08464.

[115] T. M. Karbach and M. Schlupp, *Constraints on yield parameters in extended maximum likelihood fits*, arXiv:1210.7141.

[116] BaBar collaboration, B. Aubert *et al.*, *Measurements of time-dependent CP asymmetries in $B^0 \to D^{(*)+} D^{(*)-}$ decays*, Phys. Rev. **D79** (2009) 032002, arXiv:0808.1866.

[117] Belle collaboration, M. Rohrken *et al.*, *Measurements of branching fractions and time-dependent CP violating asymmetries in $B^0 \to D^{(*)\pm}D^{\mp}$ decays*, Phys. Rev. **D85** (2012) 091106, arXiv:1203.6647.

[118] N. Belloli, *Measurement of CP violation using $B^0 \to D^{(*)\pm}D^{\mp}_{(s)}$ decays at LHCb*, PhD thesis, Milan Bicocca University, 2018, http://cds.cern.ch/record/2309847.

[119] M. Schellenberg, *Measurement of CP violation in the decay $B^0 \to D^{*\pm}D^{\mp}$ with data from the LHCb experiment*, PhD thesis, TU Dortmund University, 2019, doi: 10.17877/DE290R-20349.

[120] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, LHCb-PUB-2017-011, 2017.

[121] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, JINST **10** (2015) P02011, arXiv:1411.5528.

[122] A. Blum, A. Kalai, and J. Langford, *Beating the hold-out: bounds for k-fold and progressive cross-validation*, in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, **COLT99** 203–208, 1999.

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

[124] Y. Xie, *Principles to optimize event selections for measurements of CP asymmetries*, LHCb-INT-2009-012, CERN, 2009, http://cds.cern.ch/record/1187570.

[125] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[126] J. L. Rosner, *Determination of psendoscalar-charmed-meson decay constants from B-meson decays*, Phys. Rev. **D42** (1990) 3732.

[127] A. Rogozhnikov, *Reweighting with boosted decision trees*, J. Phys. Conf. Ser. **762** (2016) 012036, arXiv:1608.05806.

[128] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B^0 \to D^{\pm}\pi^{\mp}$ decays*, JHEP **06** (2018) 084, arXiv:1805.03448.

[129] LHCb collaboration, R. Aaij *et al.*, *New algorithms for identifying the flavour of B^0 mesons using pions and protons*, Eur. Phys. J. **C77** (2017) 238, arXiv:1610.06019.

[130] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $D^{\pm}$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B718** (2013) 902, arXiv:1210.4112.

[131] Belle collaboration, H. Kakuno *et al.*, *Neutral B flavor tagging for the measurement of mixing induced CP violation at Belle*, Nucl. Instrum. Meth. **A533** (2004) 516, arXiv:hep-ex/0403022.

[132] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math. Statist. **9** (1938) 60.

[133] A. Kordt, *Flavour Tagging Studien für eine Messung von CP-Verletzung im Zerfallskanal $B_s^0 \to D_s^+ D_s^-$ mit Daten des LHCb-Experiments*, Bachelor's thesis, TU Dortmund University, 2020.

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

[135] LHCb collaboration, R. Aaij *et al.*, *First observations of $\overline{B}_s^0 \to D^+ D^-, D_s^+ D^-$ and $D^0 \overline{D}^0$ decays*, Phys. Rev. **D87** (2013) 092007, arXiv:1302.5854.

[136] LHCb collaboration, R. Aaij *et al.*, *Study of beauty hadron decays into pairs of charm hadrons*, Phys. Rev. Lett. **112** (2014) 202001, arXiv:1403.3606.

[137] Belle II collaboration, T. Abe *et al.*, *Belle II technical design report*, arXiv:1011.0352.

[138] LHCb collaboration, *Framework TDR for the LHCb upgrade: Technical design report*, CERN-LHCC-2012-007, 2012.

[139] R. Aaij *et al.*, *Allen: A high level trigger on GPUs for LHCb*, Comput. Softw. Big Sci. **4** (2020) 7, arXiv:1912.09161.

[140] T. Likhomanenko, D. Derkach, and A. Rogozhnikov, *Inclusive flavour tagging algorithm*, J. Phys. Conf. Ser. **762** (2016) 012045, arXiv:1705.08707.

Danksagung

Während dieser Arbeit habe ich von der Hilfe vieler Menschen profitiert. Viele haben mir indirekt geholfen, indem sie zum Beispiel die LHCb Kollaboration unterstützten, Open Source-Software geschrieben oder sich um die lokale E5 Infrastruktur gekümmert haben. Aber einigen Menschen will ich namentlich danken.

Zuallererst möchte ich Johannes Albrecht danken, der die letzten Monate meiner Promotion betreut hat und sich bereiterklärt hat als Erstgutachter einzupringen. Außerdem möchte ich Herrn Kröninger dafür danken, dass er sich die Zeit genommen hat meine Arbeit zu bewerten.

Die Person, die mich eigentlich zum Lehrstuhl gebracht hat und mir die Promotion überhaupt erst ermöglicht hat, ist Herr Spaan. Er hat bereits meine Bachelor- und meine Masterarbeit betreut, genauso wie den größten Teil meiner Promotion. Er hat mir die einmalige Chance gegeben in einer internationalen Kollaboration mitzuarbeiten. Herr Spaan war eine Konstante am Lehrstuhl, dass er nun nicht mehr da ist, ist unvorstellbar. Seine Fachkenntnis wird genauso vermisst wie seine Anekdoten und Geschichten. Danke für alles!

I want to thank the $B^0 \to D^{*\pm} D^{\mp}$ team, Basem, Giulia, Margarete, Marta, Nicoletta, and Stefania, for the successful and fun time working on the analysis. I really enjoyed working with you and the visits to Milan were some of the highlight during my time as a PhD student. Ebenso danke ich der Dortmunder $B \to D\overline{D}$ Gruppe, Antje, Jan, Louis und Sophie, für die gemeinsame Arbeit. Es war eine einzigartige Erfahrung erst euer Betreuer zu sein und dann zusammen an Analysen zu arbeiten. Ich danke den Personen, die Teile meiner Arbeit korrektur gelesen haben. Danke Antje, Sophie, Gerwin, Jan, Louis, Quentin und Vukan dafür, dass ihr mir geholfen habt, die Arbeit in den jetzigen Zustand zu bringen.

Weiterer Dank geht an die Mitglieder des Lehrstuhls E5. Die Atmosphäre war stets angenehm und jeder hat sich immer Zeit genommen, um über Physik oder andere Themen zu reden. Besonders möchte ich meinen ehemaligen Bürokollegen danken, die Atmosphäre im Büro war stets angenehm und entspannt, so dass man immer gerne zur Uni gekommen ist (wenn wir nicht gerade eine Pandemie hatten). Danke Kevin, Patrick, Titus, Alex, Janine, Margarete, Gerwin, Martin und Jan. Außerdem möchte ich Matthias und Britta danken. Ohne euch würde der Lehrstuhl nicht funktionieren und ihr konntet mir mit den unterschiedlichsten Sachen weiterhelfen.

Meiner Familie möchte ich dafür danken, dass sie mich während und auch vor dem Studium immer unterstützt haben. Ohne euch hätte ich meine Promotion nicht abschließen können. Zuletz danke ich Antje. Du warst immer für mich da, egal ob es um Physik oder andere Themen ging. Du hast mir Arbeit abgenommen, wenn ich zu viel zu tun hatte und mir geholfen Stress abzubauen. Ich danke dir für die schöne Zeit und für deine Liebe.