technische universität
dortmund

Measurement of *CP* violation in the decay $B^0 \to D^{*\pm} D^{\mp}$ with data from the LHCb experiment

Dissertation zur Erlangung des akademischen Grades

Dr. rer. nat.

vorgelegt von
Margarete Schellenberg
geboren in Dortmund

Lehrstuhl für Experimentelle Physik V
Fakultät Physik
Technische Universität Dortmund

Dortmund, 23. Oktober 2019

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

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

Datum des Einreichens der Arbeit: 23.10.2019
Datum der mündlichen Prüfung: 28.11.2019

Abstract

This thesis presents the first LHCb measurement of *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays. Due to the $b \to c\bar{c}d$ transition, $B^0 \to D^{*\pm}D^{\mp}$ decays give access to higher-order Standard Model contributions. The decay-time-dependent measurement is performed on a dataset corresponding to an integrated luminosity of $9\,\mathrm{fb}^{-1}$ recorded by the LHCb detector in proton-proton collisions at centre-of-mass energies of 7, 8 and 13 TeV. The *CP* observables are measured with an unbinned maximum likelihood fit as

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

To date, the analysis is the most precise single measurement of *CP* violation in this decay channel. All parameters are consistent with their current world average values.

Kurzfassung

In dieser Arbeit wird die erste LHCb-Messung von *CP*-Verletzung in Zerfällen von $B^0 \to D^{*\pm}D^{\mp}$ vorgestellt. Aufgrund des $b \to c\bar{c}d$-Übergangs, besteht Zugang zu Standardmodellbeiträgen höherer Ordnung. Die zerfallszeitabhängige Messung wird auf einem Datensatz durchgeführt, der einer integrierten Luminosität von $9\,\mathrm{fb}^{-1}$ entspricht, und der vom LHCb-Detektor bei Proton-Proton-Kollisionen mit Schwerpunktsenergien von 7, 8 und 13 TeV aufgenommen wurde. Die *CP*-Observablen werden in einem ungebinnten Maximum-Likelihood-Fit bestimmt zu

$$
\begin{aligned}
S_{D^*D} &= -0{,}861 \pm 0{,}077\,(\text{stat}) \pm 0{,}019\,(\text{syst})\,, \\
\Delta S_{D^*D} &= 0{,}019 \pm 0{,}075\,(\text{stat}) \pm 0{,}012\,(\text{syst})\,, \\
C_{D^*D} &= -0{,}059 \pm 0{,}092\,(\text{stat}) \pm 0{,}020\,(\text{syst})\,, \\
\Delta C_{D^*D} &= -0{,}031 \pm 0{,}092\,(\text{stat}) \pm 0{,}016\,(\text{syst})\,, \\
A_{D^*D} &= 0{,}008 \pm 0{,}014\,(\text{stat}) \pm 0{,}006\,(\text{syst})\,.
\end{aligned}
$$

Die Analyse stellt die bisher präziseste Einzelmessung von *CP*-Verletzung in diesem Zerfallskanal dar. Alle Parameter sind konsistent mit ihren aktuellen Weltmittelwerten.

Contents

1 Introduction

Over the last decades, the Standard Model of particle physics (SM) has been intensively evaluated through experiments and is established as an accurate description of the constituents of nature and their fundamental interactions. Although the SM is an extraordinary successful theory, especially in the interplay between theoretical predictions and experimental verifications, there are still weaknesses and unsolved questions: For example, the neutrinos are assumed to be massless in the SM, while experimental observations of neutrino oscillations revealed that mass differences between neutrino generations exist [1, 2]. Moreover, cosmological observations indicate that the SM describes only about 5% of the energy in today's universe [3]. Additional 26% consist of so-called dark matter, while the remaining 69% comprise so-called dark energy. Dark matter becomes evident through gravitational interactions, as it is needed to explain the rotational speed of stars, gases and dust clouds in galaxies [4, 5]. The dark energy on the other hand is postulated to explain the observed accelerated expansion of the universe. Both, the dark matter and the dark energy, are not described by the SM. In addition, gravity, best known from our daily lives, is not explained by the SM, since the integration of gravity into this theory has turned out to be a complicated challenge. Then, there is the biggest question that the SM cannot answer: What happened to the antimatter that was created after the Big Bang? The Big Bang took place about 13.8 billion years ago [3] and led to an equal production of matter and antimatter in the early universe. However, inspecting our universe today, from the smallest scales on Earth to the huge stellar formations, reveals an almost exclusive composition of matter. Hence, it is a major challenge in physics to discover what happened to the antimatter resulting in the baryonic asymmetry observed today.

The most common theory to explain the baryonic asymmetry is baryogenesis, which implies the fulfilment of three conditions that were proposed by Andrei Sakharov in 1967 [6]: Firstly, there must have been a deviation from the thermal equilibrium at the early time of the universe. Secondly, the baryon number and thirdly, the charge symmetry (C) as well as the combination of charge and parity symmetry (CP) must be violated. More than 50 years ago, C violation and shortly afterwards also CP violation were discovered in particle decays and successfully incorporated in the SM [7]. To date, measurements show that the size of the *CP* violation is compatible with the SM predictions, however, it is an order of magnitude too small to account for the observed matter-antimatter asymmetry [8]. In the SM, *CP* violation is manifested in the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which is unitary and describes the probabilities of quark-flavour transitions. The unitarity leads to conditions that can partly be illustrated as triangles in the complex plane, of which the most common one is called the CKM triangle. In order to probe the SM, high precision measurements of the side lengths and angles of the CKM triangle are performed. These measurements are then combined to reveal, if the CKM triangle is a closed triangle in nature and, hence, if the CKM matrix is indeed a unitary matrix.

One angle of the CKM triangle is represented by the parameter β. It can be extracted from the quantity $\sin 2\beta$, which can be measured with high precision in so-called golden plated modes like $B^0 \to J/\psi K^0_S$. Such decays allow for a very clean prediction from the theoretical and for an excellent signature in the detector from an experimental point of view. In contrast to $B^0 \to J/\psi K^0_S$ decays, the decay $B^0 \to D^{*\pm}D^{\mp}$ allows to measure the quantity $\sin 2\beta_{\mathrm{eff}}$, as further hadronic contributions and phases exist. Knowledge of these contributions is needed to separate them from effects of physics beyond the SM, also called New Physics.

In this thesis, a measurement of *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays is presented. The analysed data was taken in proton-proton collisions accelerated to a centre-of-mass energy of up to 13 TeV by the world's most powerful particle accelerator, the Large Hadron Collider (LHC). The data has been recorded by the LHCb detector and corresponds to an integrated luminosity of $9\,\mathrm{fb}^{-1}$. The $D^{\mp}$ meson is reconstructed as $K^{\pm}\pi^{\mp}\pi^{\mp}$, while the $D^{*\pm}$ meson is reconstructed as $D^{*\pm} \to \overset{(-)}{D}{}^0\pi^{\pm}$, where for the $\overset{(-)}{D}{}^0$ decays the final states $K^{\mp}\pi^{\pm}$ and $K^{\mp}\pi^{\mp}\pi^{\pm}\pi^{\pm}$ are considered. Due to the full hadronic final state, one big challenge in the presented analysis are physical background contributions. Hence, for a precise measurement it is necessary to develop a selection, which provides a strong separation between the signal and the background components.

The measurement of *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays was performed in collaboration between the LHCb physics groups from Dortmund, CERN, Ferrara and Milan. In order to present the complete measurement, contributions from proponents of the $B^0 \to D^{*\pm}D^{\mp}$ analysis group are shown. The corresponding results are indicated in this thesis.

The outline of the thesis is as follows: the fundamentals of flavour physics are presented in Ch. 2 and are followed by an introduction to the LHCb experiment in Ch. 3, including details on the sub-detectors and the software framework. In Ch. 4, the data analysis tools and techniques utilised in this thesis are explained. While Ch. 5 discusses the datasets and their preparations, Ch. 6 outlines the mass fit to the invariant $D^{*\pm}D^{\mp}$ mass distribution, which enables a statistical subtraction of the remaining background from the signal candidates. This is followed by the description of the decay-time-dependent fit performed to measure the *CP* parameters and the studies concerning systematic uncertainties in Ch. 7 and Ch. 8, respectively. Finally, the thesis concludes in Ch. 9 with a discussion of the results and an outlook on further measurements.

2 Fundamentals of Flavour Physics

In this chapter, the physical motivation of this thesis is presented. Firstly, the standard model of particle physics is introduced in Sec. 2.1, where the elementary particles and their interactions are discussed. Furthermore, *CP* violation and neutral-meson mixing are described in Sec. 2.2 to Sec. 2.4. Finally, in Sec. 2.5 the theoretical background of the signal channel investigated in this thesis, $B^0 \to D^{*\pm}D^{\mp}$, is explained. The information in this chapter are taken from Refs.[9, 10].

2.1 The Standard Model of particle physics

The Standard Model of particle physics (SM) is a relativistic quantum field theory that comprises the elementary particles and fundamental interactions occurring among them. In total, the SM includes twelve fermions and five bosons, where the latter represent the force carriers. The interactions are explained in two theories that are part of the SM. Firstly, there is the Glashow-Weinberg-Salam (GWS) model [11–13], which describes the electroweak interactions, and secondly Quantum Chromodynamics (QCD) [14, 15] defining the strong interaction.

A fundamental aspect of the SM is the validity of the *CPT* theorem [16, 17], which was developed by Wolfgang Pauli in 1955. It states that after the concurrent execution of the transformations charge conjugation C, space reflection $P\,(\vec{x} \to -\vec{x})$ and time reversal $T\,(t \to -t)$, the physical laws retain their validity. Until 1956 it was assumed that the individual operations are preserved as well. This assumption was experimentally refuted by Wu and Garwin [18]. A year later, Landau postulated that the weak interaction was invariant under *CP* transformation. The opposite was proven in 1964, when *CP* violation was discovered for the first time in decays of long-lived kaons into two pions [7]. Outside the kaon system, *CP* violation was firstly discovered in B meson decays in 2001 [19, 20]. Just recently *CP* violation was observed by the LHCb experiment additionally in the D meson system [21].

2.1.1 Elementary particles

The SM contains twelve elementary particles, six quarks (q) and six leptons (l). They are known as fermions, as they carry an intrinsic angular momentum of $1/2$ (spin). For each particle an anti-particle ($\overline{q}$, $\bar{l}$) exists, which has the same properties as its partner except for the charge like quantum numbers, which have an opposite sign. Ascending by mass, they are divided into three families, see Fig. 2.1.

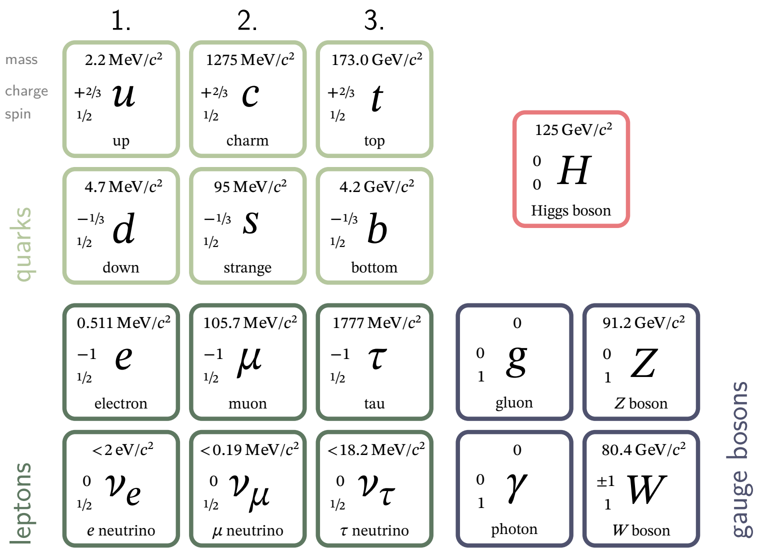

Figure 2.1 – Overview of the fermions and the gauge bosons in the SM. The values are taken from Ref. [22].

The up-type quarks (u, c, t) carry a fractional electric charge of $+2/3\,e$, while the electric charge of down-type quarks (d, s, b) is $-1/3\,e$, with e being the elementary charge. In addition to the electrical charge, (anti-)quarks carry a(n) (anti-)colour charge, whose possible states are called (anti-)red, (anti-)blue and (anti-)green. The combination of colour and anti-colour as well as the combination of red, green and blue results in neutral colour-charge. If a particle carries a colour charge it can only occur in bound states that are colour-neutral. This phenomenon is called confinement. The bound states of quarks are called hadrons and can be classified further as mesons ($q\overline{q}$), baryons (qqq, $\overline{qqq}$), tetraquarks ($qq\overline{qq}$) and pentaquarks ($qqqq\overline{q}$). Resonances compatible with the latter two types of bound states were observed by LHCb in 2015 and 2016 [23–25]. The six leptons are divided into three generations as well. Each generation consists of a charged lepton (electron e^-, muon μ^-, tau τ^-) and an associated neutral lepton, the neutrino (ν_e, ν_μ, ν_τ). The charged leptons carry a charge of $-1e$ and in case of an anti-particle $1e$.

2.1.2 Fundamental interactions

The construction of the SM results from symmetries under transformations of a gauge group. Local invariance of the gauge symmetries is considered to establish the interactions and the number of gauge bosons corresponding to the generators of the group. The bosons, which are particles of integer spin, are the mediators of interactions in the SM. The symmetry group of the SM is given as $SU(3) \otimes SU(2) \otimes U(1)$. As a part of it, the gauge group $SU(3)$, representing the QCD, describes the strong interactions and is associated with eight gluons (g). Gluons are massless and couple to the colour charge of the quarks. Since gluons carry colour and anti-colour, they can also interact with themselves.

The weak and electromagnetic interaction are unified in the electroweak theory and are associated with the gauge group $SU(2) \otimes U(1)$. The exchange particle of the electromagnetic interaction is the photon γ, which couples to the elementary charge. The photon couples to all fermions, except for the charge-neutral neutrinos. As the photon itself carries no charge, it cannot interact directly with itself like the gluons. The mediators of the weak interaction are the charged $W^{\pm}$ bosons and the neutral Z^0 boson. The range of the weak interaction is strongly limited by the masses of their exchange particles $m_W = 80.4\,\mathrm{GeV}/c^2$ and $m_Z = 91.2\,\mathrm{GeV}/c^2$, while the electromagnetic interaction has theoretically an infinite range, since the photon γ is massless. All fermions interact weakly. In addition to the $W^{\pm}$ and Z^0 bosons that carry a spin of 1, the Higgs boson, which is spinless, exists. It was predicted in 1964 [26–28] and has been discovered in 2012 by the ATLAS and CMS experiments [29, 30] and is a result of the Higgs mechanism [26, 27] that comprises a formalism of spontaneous symmetry breaking to allow for the $W^{\pm}$ and Z^0 bosons masses and enables photons to be massless.

2.2 Origin of *CP* violation in the Standard Model

The fermion masses and *CP* violation have a common origin in the SM, the Yukawa coupling to the Higgs field. It requires local gauge invariance between the left- and right-handed fermion fields and the Higgs field. Through spontaneous symmetry breaking, the fermions acquire their masses by their specific coupling constants to the Higgs field. These coupling constants are arbitrary matrices when allowing multiple generations. To obtain the physically observable mass eigenstates, the matrices are diagonalised. In the case of leptons, the matrices can be diagonalised for both e, μ, τ and ν_e, ν_μ, ν_τ in a way that matrices without mixing are created, due to the fact that neutrinos are considered to be massless in the SM. In contrast, this is not possible in the quark sector, where only one of the matrices can be diagonalised. By convention, it is performed for the up-type quarks, while for down-type quarks the mass eigenstates (d, s, b) are converted into weak interaction eigenstates (d', s', b') by the so-called CKM matrix V_{CKM} [31, 32]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{2.1}$$

The CKM matrix, established by Nicola Cabibbo, Makoto Kobayashi and Toshihide Maskawa, is a complex, unitary 3×3 matrix. A unitary matrix V fulfils the relation

$$V^{\dagger} V = \mathbb{1} . \tag{2.2}$$

Due to its non-zero off-diagonal elements, V_{CKM} enables quark transitions between up- and down-type quarks of different families. The transition probability of a quark q_i into a quark q_j is proportional to $|V_{q_i q_j}|^2$. The fact that the CKM-matrix contains a complex phase allows for the violation of the CP symmetry. The CKM matrix can be represented in many different forms. One parametrisation is referred to as the Chau-Keung parametrisation [33], which can be seen as a complex rotation matrix with three Euler mixing angles $\theta_{12}, \theta_{13}, \theta_{23} \in [0, \pi/2]$ and a phase $\delta \in (-\pi, \pi]$. With the abbreviations $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ it is given by

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}\mathrm{e}^{-\mathrm{i}\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}\mathrm{e}^{\mathrm{i}\delta} & c_{23}c_{13} \end{pmatrix} . \tag{2.3}$$

Another common way to define the CKM matrix V_{CKM} is the Wolfenstein parametrisation [34], which uses a series expansion in the parameter λ:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - \mathrm{i}\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - \mathrm{i}\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.4}$$

The Wolfenstein parametrisation illustrates the presence of the complex phase in the CKM matrix elements of V_{ub} and V_{td} in the first order of expansion, represented by the CP violating phase η. Hence, especially these transitions provide a good access to CP violation, which gives an essential motivation for studies of CP violation in B^0 $(\overline{b}d)$

decays. The relation between the parameters of Eq. (2.3) and Eq. (2.4) can be expressed as

$$\lambda = s_{12}\,, \quad A\lambda^2 = s_{23}\,, \quad A\lambda^3(\rho - \mathrm{i}\eta) = s_{13}\mathrm{e}^{-\mathrm{i}\delta}\,. \tag{2.5}$$

The four parameters of the Wolfenstein parameters are measured to be $A = 0.810^{+0.018}_{-0.024}$, $\lambda = 0.22548^{+0.00068}_{-0.00034}$, $\overline{\rho} = 0.145^{+0.013}_{-0.007}$ and $\overline{\eta} = 0.343^{+0.011}_{-0.012}$ [35], where the following relations hold

$$\overline{\rho} = \rho\left(1 - \lambda^2/2\right) \text{ and } \overline{\eta} = \eta\left(1 - \lambda^2/2\right)\,. \tag{2.6}$$

Moreover, in the Wolfenstein parametrisation, the hierarchical structure of the CKM matrix becomes visible: With increasing distance from the main diagonal, the magnitudes of the CKM matrix elements decrease. Due to the fact that the main diagonal elements are of $\mathcal{O}(1)$, quark transitions occur most likely within a family, as transitions are proportional to $|V_{ij}|^2$. A transition from a generation into a neighbouring generation (Cabibbo or CKM suppression) is more likely ($\mathcal{O}(\lambda^2)$) than skipping a generation ($\mathcal{O}(\lambda^3)$), which is referred to a doubly Cabibbo suppressed transition. Exploiting the unitarity requirement (see Eq. (2.2)), nine conditions can be constructed. Six of them are equal to zero and can be represented as triangles in the complex plane. The area of these so-called unitarity triangles is constant for all triangles and is corresponding to the half of the Jarlskog parameter J_{CP} [36], which is defined as

$$J_{CP} = \pm\mathrm{Im}(V_{ik}V_{jl}V_{il}^*V_{jk}^*) \text{ with } i \neq j,\ l \neq k\,. \tag{2.7}$$

The Jarlskog parameter can be interpreted as a measure of the amount of *CP* violation allowed in the SM, implying that *CP* symmetry is only violated if $J_{CP} \neq 0$. One of the triangles, which can be studied particularly effectively, as its side lengths have the same orders in λ, fulfils the following equation

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

This triangle is also known as the *bd*-triangle or the Unitarity Triangle, as all CKM-matrix elements include transitions of b or d quarks. In Fig. 2.2, the triangle from Eq. (2.8) is shown. Here Eq. (2.8) is divided by the experimentally best-known value $V_{cd}V_{cb}^*$ [22]. Hence, it has one side of unit length that is aligned along the horizontal axis. The corner points of the triangle have the coordinates $(0,0)$, $(1,0)$ and the apex positioned at $(\overline{\rho},\overline{\eta})$, defined by

$$\overline{\rho}+\mathrm{i}\overline{\eta}=-\frac{V_{ud}V_{ub}^{*}}{V_{cd}V_{cb}^{*}}\,. \tag{2.9}$$

The angles of the Unitarity Triangle are given by

$$\alpha=\arg\left(-\frac{V_{td}V_{tb}^{*}}{V_{ud}V_{ub}^{*}}\right),\ \ \beta=\arg\left(-\frac{V_{cd}V_{cb}^{*}}{V_{td}V_{tb}^{*}}\right)\ \text{ and }\ \gamma=\arg\left(-\frac{V_{ud}V_{ub}^{*}}{V_{cd}V_{cb}^{*}}\right). \tag{2.10}$$

One major goal of experimental flavour physics is to perform multiple measurements of the angles and the side lengths to overconstrain the Unitarity Triangle, which allow to verify, if it closes in its apex. In this way, an indirect search for New Physics is carried out. If experiments show that the sides do not converge at one point, it would be an indication of physics beyond the SM, as the unitarity is a fundamental building block of the SM. The current experimental results of the parameters and their uncertainties are combined by a fit of the CKM fitter group [37] and are in good agreement with each other and the SM. The CKM triangle and the experimental results are shown in Fig. 2.2.

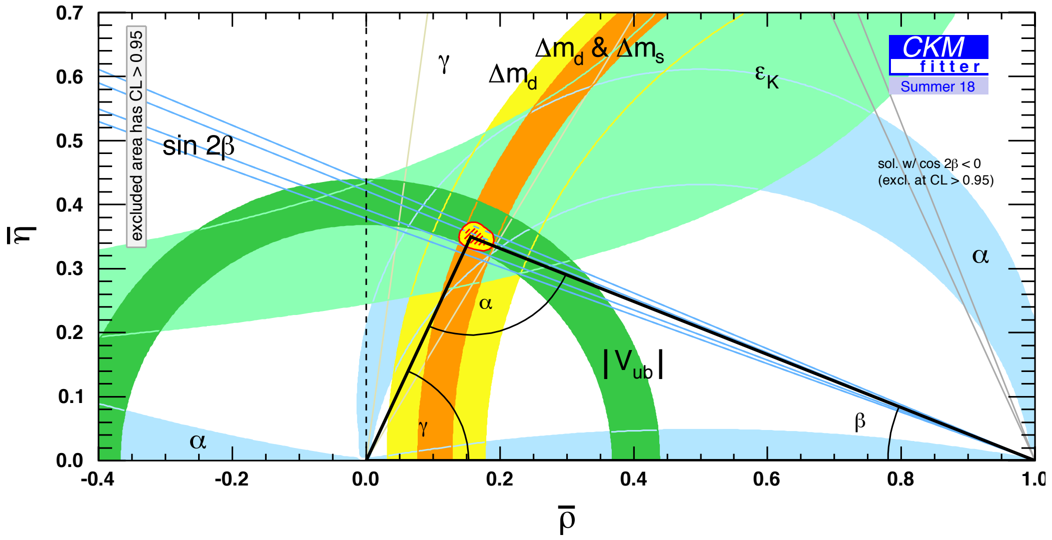

Figure 2.2 – The CKM triangle in the complex plane $(\bar{\rho},\bar{\eta})$ [37]. The coloured bands represent the experimental uncertainties of the respective parameters. The red shaded area shows the combination of all measurements.

2.3 Neutral meson system

One important effect that gives access to parameters in the flavour sector is neutral meson mixing. In this section, a short overview of the formalism for the description of neutral-meson mixing is given. A more detailed description of the formalism can be found in the Refs. [38, 39]. In the following, the neutral meson is referred to as P^0, which comprises K^0 ($\bar{s}d$), D^0 ($c\bar{u}$), B^0 ($\bar{b}d$) and B_s^0 ($\bar{b}s$) mesons.

Mesons that consist of a different quark and anti-quark type decay weakly allowing for *CP* violation and mixing. Mixing describes the conversion of a neutral meson into its anti-particle state and vice versa and is also called meson oscillation, as a time-dependent oscillation between the particle and anti-particle states appears. The decay amplitudes for a neutral meson P^0 or its *CP* conjugate $\bar{P}^0$ decaying into the final state f or its *CP* conjugate $\bar{f}$ can be expressed as

$$\begin{aligned} A_f &= A(P^0 \to f) &&= \langle f|H|P^0\rangle \,, & \bar{A}_f &= A(\bar{P}^0 \to f) &&= \langle f|H|\bar{P}^0\rangle \,, \\ A_{\bar{f}} &= A(P^0 \to \bar{f}) &&= \langle \bar{f}|H|P^0\rangle \,, & \bar{A}_{\bar{f}} &= A(\bar{P}^0 \to \bar{f}) &&= \langle \bar{f}|H|\bar{P}^0\rangle \,. \end{aligned} \tag{2.11}$$

The matrix H is the Hamiltonian of the weak interaction. Applying the *CP* operation on the initial and final states and taking $(CP)^2 = 1$ into account results in

$$\begin{aligned} CP\,|P^0\rangle &= \mathrm{e}^{+\mathrm{i}\xi_{P^0}}\,|\bar{P}^0\rangle \,, & CP\,|f\rangle &= \mathrm{e}^{+\mathrm{i}\xi_f}\,|\bar{f}\rangle \,, \\ CP\,|\bar{P}^0\rangle &= \mathrm{e}^{-\mathrm{i}\xi_{P^0}}\,|P^0\rangle \,, & CP\,|\bar{f}\rangle &= \mathrm{e}^{-\mathrm{i}\xi_f}\,|f\rangle \,. \end{aligned} \tag{2.12}$$

Due to convention, the phases ξ_{P^0} and ξ_f are arbitrary and not observable. The time-dependent wave function of a P^0 meson including the oscillation and decay is given by a superposition of P^0 and $\bar{P}^0$

$$|\psi(t)\rangle = \psi_1(t)\,|P^0\rangle + \psi_2(t)\,|\bar{P}^0\rangle \,. \tag{2.13}$$

The values of $\psi_1(t)$ and $\psi_2(t)$ can be expressed by the time-dependent Schrödinger-like equation

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = H\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = (M - {}^{\mathrm{i}}\!/\!_2\Gamma)\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} \,, \tag{2.14}$$

with the non-Hermitian Hamiltonian H constructed of the two Hermitian 2×2 matrices M and Γ. Due to the assumption of *CPT* invariance, the masses of P^0 and $\bar{P}^0$ and their decay widths Γ are identical. Subsequently, the main diagonal elements are equal

$(H_{11} = H_{22})$. The off-diagonal elements in

$$H = \begin{pmatrix} m - {}^{i}\!/\!{}_{2}\Gamma & M_{12} - {}^{i}\!/\!{}_{2}\Gamma_{12} \\ M_{12}^{*} - {}^{i}\!/\!{}_{2}\Gamma_{12}^{*} & m - {}^{i}\!/\!{}_{2}\Gamma \end{pmatrix} \tag{2.15}$$

are non-zero and therefore enable mixing of the flavour eigenstates, hence, the transition of P^0 into $\bar{P}^0$ and vice versa. The element M_{12} represents the short-distance contributions occurring via off-shell states and Γ_{12} describes the virtual intermediate decay to on-shell final states given by transitions of P^0 and $\bar{P}^0$ mesons.

Diagonalisation of the Hamilton operator results in the mass eigenstates P_H and P_L, which possess defined masses and decay widths. They can be represented as superpositions of the flavour eigenstates

$$\begin{aligned} |P_H\rangle &= p\,|P^0\rangle + q\,|\bar{P}^0\rangle \,, \\ |P_L\rangle &= p\,|P^0\rangle - q\,|\bar{P}^0\rangle \,. \end{aligned} \tag{2.16}$$

The complex coefficients p and q fulfil the normalisation condition $|p|^2 + |q|^2 = 1$. The decay width difference $\Delta\Gamma$ and the mass difference Δm of the mass eigenstates are defined in such a way that the mass difference is positive. The index L means light and H stands for heavy.

$$\Delta m \equiv m_H - m_L \qquad \text{and} \qquad \Delta\Gamma \equiv \Gamma_H - \Gamma_L \,. \tag{2.17}$$

The ratio of the complex coefficients p and q can be written with the matrix elements of M and Γ via

$$\left(\frac{q}{p}\right)^2 = \frac{M_{12}^{*} - {}^{i}\!/\!{}_{2}\Gamma_{12}^{*}}{M_{12} - {}^{i}\!/\!{}_{2}\Gamma_{12}}\,, \qquad \frac{q}{p} = -\frac{\Delta m + {}^{i}\!/\!{}_{2}\Delta\Gamma}{2(M_{12} - {}^{i}\!/\!{}_{2}\Gamma_{12})}\,. \tag{2.18}$$

The time evolutions of the mass eigenstates P_H and P_L are defined as

$$\begin{aligned} |P_H(t)\rangle &= \mathrm{e}^{-\mathrm{i}m_H t}\mathrm{e}^{-\Gamma_H/2\,t}\,|P_H\rangle \,, \\ |P_L(t)\rangle &= \mathrm{e}^{-\mathrm{i}m_L t}\mathrm{e}^{-\Gamma_L/2\,t}\,|P_L\rangle \,. \end{aligned} \tag{2.19}$$

Inserting Eqs. 2.16 into Eqs. 2.19 gives

$$\begin{aligned} |P^0(t)\rangle &= g_+(t)\,|P^0\rangle - \frac{q}{p}g_-(t)\,|\bar{P}^0\rangle \,, \\ |\bar{P}^0(t)\rangle &= g_+(t)\,|\bar{P}^0\rangle - \frac{p}{q}g_-(t)\,|P^0\rangle \,, \end{aligned} \tag{2.20}$$

with

$$g_\pm(t) = \frac{1}{2}\left(\mathrm{e}^{-\mathrm{i}m_H t}\mathrm{e}^{-\frac{\Gamma_H}{2}t} \pm \mathrm{e}^{-\mathrm{i}m_L t}\mathrm{e}^{-\frac{\Gamma_L}{2}t}\right). \tag{2.21}$$

The strong or electromagnetic interactions can produce the flavour eigenstates P^0 and $\bar{P}^0$, which enable an examination of their decays into a final state f or its charge conjugate $\bar{f}$ at a decay time t. Experimentally, the time-dependent decay rates $\Gamma(t)$ of the flavour eigenstates P^0 and $\bar{P}^0$ into the final state f or $\bar{f}$ are investigated:

$$\begin{aligned}
&\Gamma(P^0(t) \to f) = \left|\langle f|H|P^0(t)\rangle\right|^2, \qquad &&\Gamma(\bar{P}^0(t) \to f) = \left|\langle f|H|\bar{P}^0(t)\rangle\right|^2, \\
&\Gamma(P^0(t) \to \bar{f}) = \left|\langle \bar{f}|H|P^0(t)\rangle\right|^2, \qquad &&\Gamma(\bar{P}^0(t) \to \bar{f}) = \left|\langle \bar{f}|H|\bar{P}^0(t)\rangle\right|^2.
\end{aligned} \tag{2.22}$$

Using the decay amplitudes A_f, $A_{\bar{f}}$, $\bar{A}_f$ and $\bar{A}_{\bar{f}}$ from Eqs. 2.11 and the quotient of the coefficients from Eqs. 2.16, the following complex parameters λ_f and $\lambda_{\bar{f}}$ can be defined:

$$\lambda_f = \frac{1}{\bar{\lambda}_f} = \frac{q}{p}\frac{\bar{A}_f}{A_f}, \qquad \lambda_{\bar{f}} = \frac{1}{\bar{\lambda}_{\bar{f}}} = \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_{\bar{f}}}. \tag{2.23}$$

These complex quantities include all the factors related to the occurrence of *CP* violation, hence, they can be utilised to distinguish between the different types of *CP* violation that are explained in Sec. 2.4. If the *CP* symmetry would be preserved, then $\lambda_f = \pm 1$ would be valid.

The parameters λ_f and $\lambda_{\bar{f}}$ can be used to express the time-dependent rates as

$$\begin{aligned}
\Gamma(P^0(t) \to f) &= |A_f|^2 && \left(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 - 2\,\mathrm{Re}(\lambda_f g_+^*(t) g_-(t))\right), \\
\Gamma(\bar{P}^0(t) \to f) &= |A_f|^2 && \left|\frac{p}{q}\right|^2\left(|g_-(t)|^2 + |\lambda_f|^2|g_+(t)|^2 - 2\,\mathrm{Re}(\lambda_f g_+(t) g_-^*(t))\right), \\
\Gamma(P^0(t) \to \bar{f}) &= |A_{\bar{f}}|^2 && \left|\frac{q}{p}\right|^2\left(|g_-(t)|^2 + |\lambda_{\bar{f}}|^2|g_+(t)|^2 - 2\,\mathrm{Re}(\lambda_{\bar{f}} g_+(t) g_-^*(t))\right), \\
\Gamma(\bar{P}^0(t) \to \bar{f}) &= |A_{\bar{f}}|^2 && \left(|g_+(t)|^2 + |\lambda_{\bar{f}}|^2|g_-(t)|^2 - 2\,\mathrm{Re}(\lambda_{\bar{f}} g_+^*(t) g_-(t))\right).
\end{aligned} \tag{2.24}$$

The terms proportional to $\frac{p}{q}A_f^2$, $\frac{q}{p}A_{\bar{f}}^2$ and $g_-(t)^2$ are associated with decays after a net oscillation. Decays of P^0 ($\bar{P}^0$) that do not involve an oscillation, i.e. that were produced as P^0 ($\bar{P}^0$), are represented by the terms proportional to $g_+(t)^2$ and A^2 in Eq. (2.24). Terms proportional to $(g_\pm^{(*)}(t) g_\mp(t))$ result from interference of direct decay and decay after oscillation. A replacement of

$$\begin{aligned} g_+^*(t)g_-(t) &= -\frac{\mathrm{e}^{-\Gamma t}}{2}\left[\sinh\left(\frac{\Delta\Gamma t}{2}\right) + \mathrm{i}\sin(\Delta m t)\right] \\ \text{and}\quad |g_\pm(t)|^2 &= \frac{\mathrm{e}^{-\Gamma t}}{2}\left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) \pm \cos(\Delta m t)\right] , \end{aligned} \tag{2.25}$$

by a combination of trigonometric and hyperbolic functions, allows to write the decay rates in the simpler forms

$$\begin{aligned} \Gamma(P^0(t) \to f) =& \frac{1}{2}\left|A_f\right|^2\left(1+|\lambda_f|^2\right)\mathrm{e}^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D_f \sinh\left(\frac{\Delta\Gamma}{2}t\right)\right. \\ &\left. + C_f \cos(\Delta m t) - S_f \sin(\Delta m t)\right], \\ \Gamma(\bar{P}^0(t) \to f) =& \frac{1}{2}\left|A_f\right|^2\left(1+|\lambda_f|^2\right)\mathrm{e}^{-\Gamma t}\left|\frac{p}{q}\right|^2 \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D_f \sinh\left(\frac{\Delta\Gamma}{2}t\right)\right. \\ &\left. - C_f \cos(\Delta m t) + S_f \sin(\Delta m t)\right], \\ \Gamma(P^0(t) \to \bar{f}) =& \frac{1}{2}\left|A_{\bar{f}}\right|^2\left(1+|\lambda_{\bar{f}}|^2\right)\mathrm{e}^{-\Gamma t}\left|\frac{p}{q}\right|^2 \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma}{2}t\right)\right. \\ &\left. - C_{\bar{f}} \cos(\Delta m t) + S_{\bar{f}} \sin(\Delta m t)\right], \\ \Gamma(\bar{P}^0(t) \to \bar{f}) =& \frac{1}{2}\left|A_{\bar{f}}\right|^2\left(1+|\lambda_{\bar{f}}|^2\right)\mathrm{e}^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma}{2}t\right) + D_{\bar{f}} \sinh\left(\frac{\Delta\Gamma}{2}t\right)\right. \\ &\left. + C_{\bar{f}} \cos(\Delta m t) - S_{\bar{f}} \sin(\Delta m t)\right], \end{aligned} \tag{2.26}$$

exploiting the following *CP* observables

$$\begin{aligned} S_f &= \frac{2\mathrm{Im}(\lambda_f)}{1+|\lambda_f|^2}, & C_f &= \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, & D_f &= \frac{2\mathrm{Re}(\lambda_f)}{1+|\lambda_f|^2}, \\ S_{\bar{f}} &= \frac{2\mathrm{Im}(\lambda_{\bar{f}})}{1+|\lambda_{\bar{f}}|^2}, & C_{\bar{f}} &= \frac{1-|\lambda_{\bar{f}}|^2}{1+|\lambda_{\bar{f}}|^2}, & D_{\bar{f}} &= \frac{2\mathrm{Re}(\lambda_{\bar{f}})}{1+|\lambda_{\bar{f}}|^2}. \end{aligned} \tag{2.27}$$

The *CP* observables satisfy the relations $S_f^2 + C_f^2 + D_f^2 = 1$ and $S_{\bar{f}}^2 + C_{\bar{f}}^2 + D_{\bar{f}}^2 = 1$. The terms in Eq. (2.26), which are proportional to $\cos(\Delta m t)$ and $\cosh(\Delta\Gamma/2\, t)$, originate from decays with and without oscillation. Terms proportional to $\sin(\Delta m t)$ and $\sinh(\Delta\Gamma/2\, t)$ are associated to interference between direct decays and decays after mixing.

2.4 Classes of *CP* violation

CP violation is another very important aspect that allows to measure parameters in the flavour sector. As discussed in Sec. 2.2, the effect of *CP* violation occurs only in the electroweak interaction in the SM. It is a single irreducible complex phase in the CKM matrix that allows for *CP* violation. In weak decays of neutral and charged mesons, *CP* violation appears in three manifestations. Firstly, *CP* violation in the decay, often called direct *CP* violation, results in different amplitudes between *CP*-conjugated decays. Secondly, *CP* violation in mixing leads to different probabilities between a neutral meson oscillating in its antiparticle compared to the reversed process. Finally, *CP* violation in the interference of direct decay and decay after mixing exists. These three manifestations of *CP* violation are explained in the following.

2.4.1 Direct *CP* violation

If the magnitudes of the amplitudes of *CP*-conjugated processes differ according to

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| \neq 1\,, \tag{2.28}$$

CP violation in the decay exists. The required criteria for *CP* violation in the decay are at least two contributing decay amplitudes ($j \geq 2$) with not cancelling relative strong and weak phases. The total decay amplitudes are expressed as

$$A_f = \sum_j |a_j| \mathrm{e}^{\mathrm{i}(\delta_j + \phi_j)} \quad \text{and} \quad \bar{A}_{\bar{f}} = \sum_j |a_j| \mathrm{e}^{\mathrm{i}(\delta_j - \phi_j)}\,, \tag{2.29}$$

where δ_j and ϕ_j denote the strong and weak phases and a_j is associated with the single contribution amplitudes. While the strong phases arise due to hadronic effects, the weak phases stem from complex CKM matrix elements and appear with opposite signs in A_f and $\bar{A}_{\bar{f}}$. As the strong interaction is invariant under *CP* transformations, the strong phases of A_f and $\bar{A}_{\bar{f}}$ are equal. Expressing the difference of the squared decay amplitudes

$$|A_f|^2 - |\bar{A}_{\bar{f}}|^2 = 2\sum_{i,j} |a_i||a_j| \sin(\delta_i - \delta_j)\sin(\phi_i - \phi_j)\,, \tag{2.30}$$

the requirement for *CP* violation becomes visible. This difference is non-zero for the case of non-vanishing differences of the strong phases $\delta_i - \delta_j$ and weak phases $\phi_i - \phi_j$. Both conditions must be fulfilled for direct *CP* violation to be present. Due to the absence of mixing for charged mesons, this class of *CP* violation is the only possible type of *CP* violation in charged meson decays. In $B^0_{(d,s)}$ decays for example, direct *CP* violation has been significantly observed in decays of $B^0_{(d,s)} \rightarrow K^+\pi^-$ and on the charge conjugated channel at LHCb [40].

2.4.2 Indirect *CP* violation

Indirect *CP* violation, also named *CP* violation in the mixing, can occur, if

$$\left|\frac{q}{p}\right| \neq 1 \tag{2.31}$$

is fulfilled. This condition implies that the complex coefficients differ in more than just a phase. Considering flavour-specific decays, where a final state f ($\bar{f}$) can just be reached by the decay of P^0 ($\bar{P}^0$) and not by $\bar{P}^0$ (P^0), i.e. $|A_{\bar{f}}| = |\bar{A}_f| = 0$, and requiring no direct *CP* violation, i.e. $|A_f| = |\bar{A}_{\bar{f}}|$, then transitions like $P^0 \rightarrow \bar{f}$ or $\bar{P}^0 \rightarrow f$ can only appear after a net oscillation of a meson into its anti-particle state. Measurements of indirect *CP* violation, among other experiments, were performed at LHCb in semileptonic B meson decays [41, 42]. In these measurements, an asymmetry expressed by the time-dependent decay-rates of neutral $B^0_{(d,s)}$ and $\bar{B}^0_{(d,s)}$ mesons into the final states μ^+X and μ^-X according to

$$a_{\mathrm{sl}}^{(d,s)} = \frac{\Gamma(\bar{B}^0_{(d,s)}(t) \rightarrow \mu^+X) - \Gamma(B^0_{(d,s)}(t) \rightarrow \mu^+X)}{\Gamma(\bar{B}^0_{(d,s)}(t) \rightarrow \mu^+X) + \Gamma(B^0_{(d,s)}(t) \rightarrow \mu^+X)} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \tag{2.32}$$

is investigated, which has proven to be time-independent. In the neutral B meson system this class of *CP* violation is predicted to be small $a_{\mathrm{sl}}^{(d,s)} = \mathcal{O}\left(10^{-4}\right)$ [43, 44]. Currently, experimental results [41, 42] are in agreement with this prediction, and, due to the large uncertainties, compatible with zero. Hence, indirect *CP* violation in neutral B meson decays can be neglected.

2.4.3 *CP* violation in the interference of decay and decay after mixing

The third type of *CP* violation, which is called *CP* violation in the interference of decay and decay after mixing, or shortly interference *CP* violation, can occur when a meson and the corresponding anti-meson have a common final state f. It arises in the interference between the direct decay $P^0 \rightarrow f$ and the decay after oscillation $P^0 \rightarrow \bar{P}^0 \rightarrow f$ and is caused by a phase difference between mixing and decay, which can be expressed with

$$\mathrm{Im}(\lambda_f) \neq 0\,. \tag{2.33}$$

This criterion contains the implication that this type of *CP* violation can be present in the absence of direct and indirect *CP* violation, which results in $|\lambda_f| = 1$. In order to observe interference *CP* violation in decays of initial P^0 and $\bar{P}^0$ into a common *CP* final state f_{CP}, a time-dependent *CP* asymmetry $A_{f_{CP}}(t)$ is defined as

$$A_{f_{CP}}(t) = \frac{\Gamma(\bar{P}^0(t) \to f_{CP}) - \Gamma(P^0(t) \to f_{CP})}{\Gamma(\bar{P}^0(t) \to f_{CP}) + \Gamma(P^0(t) \to f_{CP})}\,. \tag{2.34}$$

By exploiting the time-dependent decay-rates defined in Eqs. 2.26, as well as the *CP* observables from Eq. (2.27), the *CP* asymmetry can be expressed by

$$A_{f_{CP}}(t) = \frac{S_{f_{CP}} \sin(\Delta m t) - C_{f_{CP}} \cos(\Delta m t)}{\cosh\left(\frac{\Delta\Gamma t}{2}\right) + D_{f_{CP}} \sinh\left(\frac{\Delta\Gamma t}{2}\right)}\,. \tag{2.35}$$

In the system of B^0 mesons, the decay width difference $\Delta\Gamma$ is negligible, which simplifies the time-dependent *CP* asymmetry to

$$A_{f_{CP}}(t) = S_{f_{CP}} \sin(\Delta m t) - C_{f_{CP}} \cos(\Delta m t)\,. \tag{2.36}$$

The *CP* parameters $S_{f_{CP}}$ and $C_{f_{CP}}$ are the amplitudes of the sine and cosine oscillations and can be experimentally obtained by investigating decays with known initial state. Whenever the contributing amplitudes carry the same weak phase resulting in $|A_{f_{CP}}| = |\bar{A}_{f_{CP}}|$ and $|\lambda_{f_{CP}}| = 1$, no direct *CP* violation can be present and the *CP* parameter $S_{f_{CP}}$ can be expressed as

$$S_{f_{CP}} = \mathrm{Im}(\lambda_f)\,. \tag{2.37}$$

Thus, $S_{f_{CP}}$ is directly connected to the phases of the CKM matrix elements that are involved in the mixing. Decays with this property are referred to as golden modes, like $B^0 \to J/\psi K^0_S$ and $B^0_s \to J/\psi\phi$, where the most recent measurements were performed by the LHCb collaboration [45, 46].

2.5 *CP* violation in the decay $B^0 \to D^{*\pm}D^{\mp}$

Investigating decays that involve *CP* violation is an excellent way to verify the theoretical principles in the quark-flavour sector of the SM. In particular, such decays give access to the side lengths and angles of the unitary triangle (see Fig. 2.2), allowing to test the unitarity of the CKM matrix precisely. As an example, the angle β can be measured with high precision in the golden channel $B^0 \to J/\psi K^0_S$. This decay occurs via $b \to c\bar{c}s$ transitions. Moreover, direct and indirect *CP* violation are strongly suppressed, allowing to measure the weak mixing phase ϕ_d and, thus, the angle β without large theoretical uncertainties. From an experimental point of view, an analysis of $B^0 \to J/\psi K^0_S$ decays is supported by clean signatures in the detector, especially when reconstructing J/ψ in the two muon final state.

In contrast to $B^0 \to J/\psi K^0_S$, the final state of $B^0 \to D^{*\pm}D^{\mp}$ decays is not a *CP* eigenstate, as it is distinguishable by the charge configuration of the D mesons. This leads to additional measurable *CP* parameters , which will be introduced in the following sections. Moreover, the decay $B^0 \to D^{*\pm}D^{\mp}$, as part of the $B \to D\overline{D}$ family, involves $b \to c\bar{c}d$ transitions, which are CKM suppressed. Contributions of higher orders are not Cabibbo-suppressed, so that the main motivation of analyses of $B \to D\overline{D}$ decays is to constraint these higher order contributions in order to distinguish them from effects of New Physics. Consequently, instead of a pure ϕ_d, only an effective phase ϕ_d^{eff} can be measured in such decays.

2.5.1 B^0-$\bar{B}^0$ mixing phenomenology

A measurement of *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays can be carried out in terms of a decay-time-dependent analysis. Referring to Sec. 2.3, the complex quantity q/p, the mass difference Δm and the decay-width difference $\Delta\Gamma$ are important parameters that are necessary to define the time-dependent decay rates of neutral meson decays. All of these quantities are associated with the off-diagonal matrix elements of the Hamiltonian H in Eq. (2.15). They are expressed by M_{12} and Γ_{12} that need to be defined in a more detailed way to investigate the mixing phenomenology. In the SM, the $B^0 - \bar{B}^0$ mixing occurs at quark-level via so-called short-distance transitions, which contribute to M_{12} mostly. The matrix element M_{12} can be expressed by utilising the box diagrams illustrated in Fig. 2.3 as

$$M_{12} = -\frac{G_F^2 m_W^2}{12\pi^2} f_{B^0}^2 m_{B^0} B_{B^0} \mathcal{F}^* \,, \tag{2.38}$$

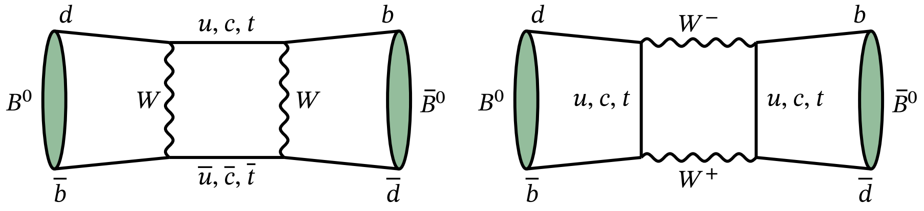

Figure 2.3 – Dominant box diagrams for the oscillation of the neutral B^0 meson. Of the up-type quarks, the t quark is dominant.

where G_f is the Fermi constant, m_W the mass of the W boson, m_{B^0} the B^0 meson mass and f_{B^0} the weak decay constant. Finally, B_{B^0} is called bag parameter and describes non-pertubative QCD corrections. By summing over all box diagrams, the function $\mathcal{F}$ is given by

$$\mathcal{F} = \eta_1 \lambda_c^2 S_0(x_c) + \eta_2 \lambda_t^2 S_0(x_t) + 2\eta_3 \lambda_c \lambda_t S_0(x_c, x_t)\,. \tag{2.39}$$

The term $\lambda_i = V_{ib}^* V_{id}$ denotes the CKM couplings, x_i is an abbreviation for m_i^2/m_W^2 and the factors η_i are the QCD corrections. The Inami-Lim function [47] S_0 fulfils

$$S_0(x_t) \gg S_0(x_c, x_t) > S_0(x_c)\,, \tag{2.40}$$

which allows the approximation

$$M_{12} \approx -\frac{G_F^2 m_W^2}{12\pi} f_{B^0} m_{B^0} B_{B^0} \eta_2 (V_{tb} V_{td}^*)^2 S_0(x_t)\,. \tag{2.41}$$

The matrix element Γ_{12} on the other hand is defined by the sum over all physical states f that are accessible by both B^0 and $\bar{B}^0$

$$\Gamma = \sum_f \langle f|T|B^0\rangle^* \langle f|T|\bar{B}^0\rangle\ . \tag{2.42}$$

Considering that, because of the high mass of the t quark, transitions to physical states are only allowed for hadrons containing c or u quarks. For these transitions two probability restrictions exist: either the decays of both the B^0 and $\bar{B}^0$ are Cabibbo-suppressed, or the decay is Cabibbo-allowed for B^0 and doubly Cabibbo-suppressed for $\bar{B}^0$ mesons or vice versa. Thus, the following estimations can be made

$$\left|\frac{\Gamma_{12}}{M_{12}}\right| \sim \frac{m_b^2}{m_t^2} \sim 10^{-3}\,. \tag{2.43}$$

Taking into account the following equations

$$\begin{aligned}(\Delta m)^2 - {}^1\!/\!_4(\Delta\Gamma)^2 &= 4|M_{12}|^2 - |\Gamma_{12}|^2\,, \\ (\Delta m)(\Delta\Gamma) &= 4\,\mathrm{Re}(M_{12}^*\Gamma_{12})\,,\end{aligned} \tag{2.44}$$

Δm and $\Delta\Gamma$ can be expressed as

$$\begin{aligned}\Delta m &= 2|M_{12}|\,, \\ \Delta\Gamma &= \frac{2\,\mathrm{Re}(M_{12}^*\Gamma_{12})}{|M_{12}|}\,.\end{aligned} \tag{2.45}$$

Hence, the decay width difference is expected to be small compared to the mass difference, which allows to express q/p from Eq. (2.18) in terms of CKM matrix elements as

$$\frac{q}{p} \approx \sqrt{\frac{M_{12}^*}{M_{12}}} = \frac{V_{td}V_{tb}^*}{V_{tb}V_{td}^*}\,. \tag{2.46}$$

This equation will be used in the next section, in order to express λ solely in terms of CKM matrix elements in case of the $B^0 \to D^{*\pm}D^{\mp}$ decay.

2.5.2 Decay amplitudes in $B^0 \to D^{*\pm}D^{\mp}$

Both, the $D^{*+}D^-$ and the $D^{*-}D^+$ final states, can be accessed by the B^0 and the $\bar{B}^0$ mesons with decay amplitudes of similar magnitudes. Thus, four amplitudes need to be considered: the ones of a B^0 meson into $D^{*+}D^-$ ($A(B^0 \to D^{*+}D^-)$) and into $D^{*-}D^+$ ($A(B^0 \to D^{*-}D^+)$) and equivalent for a $\bar{B}^0$ meson ($A(\bar{B}^0 \to D^{*+}D^-)$, $A(\bar{B}^0 \to D^{*-}D^+)$). In Fig. 2.4, the four Feynman diagrams contributing at comparable level to $B^0 \to D^{*\pm}D^{\mp}$ are shown. There are the tree diagram (T), which is CKM-suppressed, the penguin ($P^{(q)}$), exchange (E) and penguin annihilation ($PA^{(q)}$) diagrams. The q refers to the quark in the loop. The latter three diagrams need to be considered, as they are not Cabibbo-suppressed and carry a different weak phase, than the tree diagram. Hence, the amplitude of a B^0 into $D^{*+}D^-$ can be expressed as [38, 48–50]

$$\begin{aligned}A(B^0 \to D^{*+}D^-) = {}& V_{cd}V_{cb}^*T + V_{td}V_{tb}^*P^{(t)} + V_{cd}V_{cb}^*P^{(c)} + V_{ud}V_{ub}^*P^{(u)} \\ &+ V_{cd}V_{cb}^*E + V_{td}V_{tb}^*PA^{(t)} + V_{cd}V_{cb}^*PA^{(c)} + V_{ud}V_{ub}^*PA^{(u)}\,.\end{aligned} \tag{2.47}$$

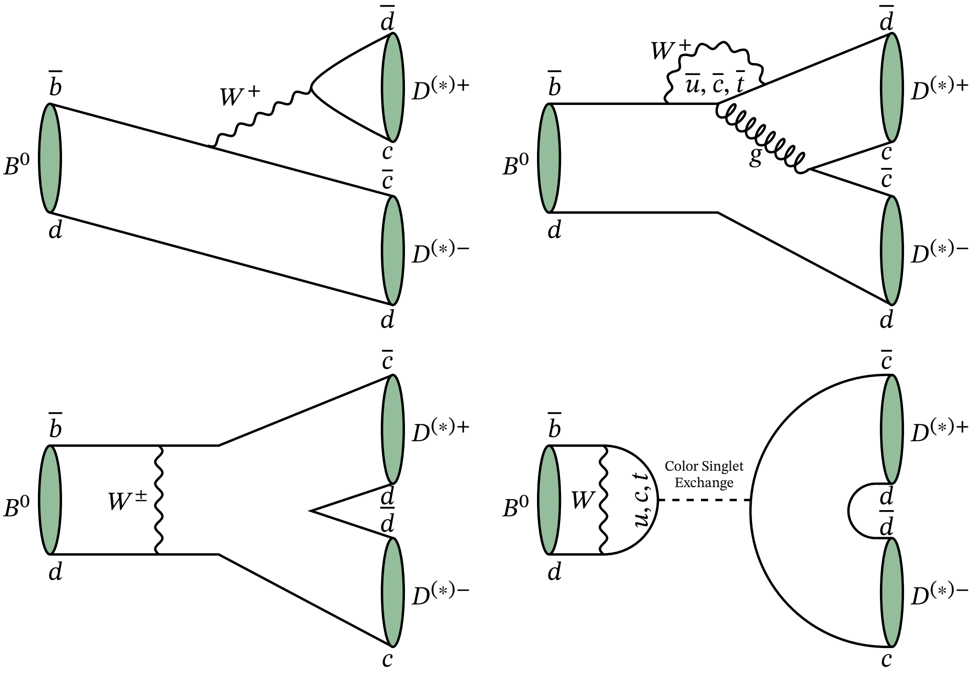

Figure 2.4 – Contributing Feynman diagrams to the decay $B^0 \to D^{*\pm}D^{\mp}$. While at the top the (left) tree and the (right) penguin diagrams are depicted, at the bottom the (left) exchange and penguin (right) annihilation diagrams are presented.

Exploiting the CKM unitary condition Eq. (2.8)

$$V_{td}V_{tb}^* + V_{cd}V_{cb}^* + V_{ud}V_{ub}^* = 0 \iff V_{td}V_{tb}^* = -V_{cd}V_{cb}^* - V_{ud}V_{ub}^* \tag{2.48}$$

the decay amplitude is given by

$$\begin{aligned} A(B^0 \to D^{*+}D^-) = {} & V_{ud}V_{ub}^*(P^{(u)} + PA^{(u)} - P^{(t)} - PA^{(t)}) \\ & + V_{cd}V_{cb}^*(\underbrace{T + E + P^{(c)} + PA^{(c)} - P^{(t)} - PA^{(t)}}_{\mathcal{M}}) . \end{aligned} \tag{2.49}$$

By factorising out $V_{cd}V_{cb}^*\mathcal{M}$, the amplitude can be simplified to

$$A(B^0 \to D^{*+}D^-) = V_{cd}V_{cb}^*\mathcal{M}\left(1 + \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\frac{P^{(u)} + PA^{(u)} - P^{(t)} - PA^{(t)}}{\mathcal{M}}\right) . \tag{2.50}$$

Utilising the definition of the CKM angle γ (Eq. (2.10))

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} = -R_b \mathrm{e}^{\mathrm{i}\gamma} , \tag{2.51}$$

where R_b is the side length of the CKM triangle enclosed by the angles α and γ, the amplitude can be written with

$$u\mathrm{e}^{\mathrm{i}\theta} \equiv R_b \left(\frac{P^{(u)} + PA^{(u)} - P^{(t)} - PA^{(t)}}{\mathcal{M}} \right) \tag{2.52}$$

as

$$A(B^0 \to D^{*+}D^-) = V_{cd}V_{cb}^*\mathcal{M}(1 - u\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{\mathrm{i}\gamma})\,. \tag{2.53}$$

Here, γ is the *CP*-violating weak phase, while θ and u are hadronic parameters that conserve *CP*.

The amplitude for the charge conjugated decay $\bar{B}^0 \to D^{*-}D^+$ is given by

$$A(\bar{B}^0 \to D^{*-}D^+) = V_{cb}V_{cd}^*\mathcal{M}(1 - u\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\gamma})\,. \tag{2.54}$$

The sign of the weak phase γ is flipped, while the sign of the strong phase θ stays the same, as a result of *CP* invariance. The two missing decay amplitudes are defined as

$$\begin{aligned} A(B^0 \to D^{*-}D^+) &= V_{cd}V_{cb}^*\overline{\mathcal{M}}(1 - \bar{u}\mathrm{e}^{\mathrm{i}\bar{\theta}}\mathrm{e}^{\mathrm{i}\gamma})\,, \\ A(\bar{B}^0 \to D^{*+}D^-) &= V_{cb}V_{cd}^*\overline{\mathcal{M}}(1 - \bar{u}\mathrm{e}^{\mathrm{i}\bar{\theta}}\mathrm{e}^{-\mathrm{i}\gamma})\,. \end{aligned} \tag{2.55}$$

For the amplitudes $A(B^0 \to D^{*-}D^+)$ and $A(\bar{B}^0 \to D^{*+}D^-)$ another strong phase $\bar{\theta}$ and hadronic parameters $\overline{\mathcal{M}}$ and $\bar{u}$ are introduced. Inserting all four amplitudes into Eq. (2.23) and considering the approximation of q/p from Eq. (2.46) gives access to the parameters describing *CP* violation of decay and decay after mixing:

$$\lambda_{D^{*+}D^-} = \frac{q}{p}\frac{A(\bar{B}^0 \to D^{*+}D^-)}{A(B^0 \to D^{*+}D^-)} = \frac{V_{td}V_{tb}^*}{V_{tb}V_{td}^*}\frac{V_{cb}V_{cd}^*}{V_{cd}V_{cb}^*}\frac{\overline{\mathcal{M}}}{\mathcal{M}}\frac{1 - \bar{u}\mathrm{e}^{\mathrm{i}\bar{\theta}}\mathrm{e}^{-\mathrm{i}\gamma}}{1 - u\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{\mathrm{i}\gamma}}\,, \tag{2.56}$$

$$\lambda_{D^{*-}D^+} = \frac{q}{p}\frac{A(\bar{B}^0 \to D^{*-}D^+)}{A(B^0 \to D^{*-}D^+)} = \frac{V_{td}V_{tb}^*}{V_{tb}V_{td}^*}\frac{V_{cb}V_{cd}^*}{V_{cd}V_{cb}^*}\frac{\mathcal{M}}{\overline{\mathcal{M}}}\frac{1 - u\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\gamma}}{1 - \bar{u}\mathrm{e}^{\mathrm{i}\bar{\theta}}\mathrm{e}^{\mathrm{i}\gamma}}\,. \tag{2.57}$$

By taking the definition of the angle β of the Unitarity Triangle (see Eq. (2.10)) into account, $\lambda_{D^{*+}D^-}$ and $\lambda_{D^{*-}D^+}$ can be written as

$$\begin{aligned} \lambda_{D^{*+}D^-} &= \mathrm{e}^{-\mathrm{i}2\beta}\frac{\overline{\mathcal{M}}}{\mathcal{M}}\frac{1 - \bar{u}\mathrm{e}^{\mathrm{i}\bar{\theta}}\mathrm{e}^{-\mathrm{i}\gamma}}{1 - u\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{\mathrm{i}\gamma}}\,, \\ \lambda_{D^{*-}D^+} &= \mathrm{e}^{-\mathrm{i}2\beta}\frac{\mathcal{M}}{\overline{\mathcal{M}}}\frac{1 - u\mathrm{e}^{\mathrm{i}\theta}\mathrm{e}^{-\mathrm{i}\gamma}}{1 - \bar{u}\mathrm{e}^{\mathrm{i}\bar{\theta}}\mathrm{e}^{\mathrm{i}\gamma}}\,. \end{aligned} \tag{2.58}$$

Since the hadronic parameters $u\mathrm{e}^{\mathrm{i}\theta}$ and $\bar{u}\mathrm{e}^{\mathrm{i}\bar{\theta}}$ cannot be calculated reliably within QCD [51], an experimental determination via the *CP* parameters

$$\begin{aligned} S_{D^{*+}D^-} &= \frac{-2\left[\sin 2\beta - u\mathcal{K}_+ - \bar{u}\overline{\mathcal{K}}_- + u\bar{u}\sin(2\beta + 2\gamma - \delta)\right]}{\frac{\overline{\mathcal{M}}}{\mathcal{M}}(1 - 2\bar{u}\cos(\bar{\theta} - \gamma) + \bar{u}^2) + \frac{\mathcal{M}}{\overline{\mathcal{M}}}(1 - 2u\cos(\bar{\theta} + \gamma) + u^2)} \\ S_{D^{*-}D^+} &= \frac{-2\left[\sin 2\beta - \bar{u}\overline{\mathcal{K}}_+ - u\mathcal{K}_- + \bar{u}u\sin(2\beta + 2\gamma + \delta)\right]}{\frac{\mathcal{M}}{\overline{\mathcal{M}}}(1 - 2u\cos(\theta - \gamma) + u^2) + \frac{\overline{\mathcal{M}}}{\mathcal{M}}(1 - 2\bar{u}\cos(\bar{\theta} + \gamma) + \bar{u}^2)} \end{aligned} \tag{2.59}$$

$$\begin{aligned} C_{D^{*+}D^-} &= \frac{\frac{\mathcal{M}}{\overline{\mathcal{M}}}(1 - 2u\cos(\theta + \gamma) + u) - \frac{\overline{\mathcal{M}}}{\mathcal{M}}(1 - 2\bar{u}\cos(\bar{\theta} - \gamma) + \bar{u})}{\frac{\mathcal{M}}{\overline{\mathcal{M}}}(1 - 2u\cos(\theta + \gamma) + u) + \frac{\overline{\mathcal{M}}}{\mathcal{M}}(1 - 2\bar{u}\cos(\bar{\theta} - \gamma) + \bar{u})} \\ C_{D^{*-}D^+} &= \frac{\frac{\overline{\mathcal{M}}}{\mathcal{M}}(1 - 2\bar{u}\cos(\bar{\theta} + \gamma) + \bar{u}) - \frac{\mathcal{M}}{\overline{\mathcal{M}}}(1 - 2u\cos(\theta - \gamma) + u)}{\frac{\overline{\mathcal{M}}}{M}(1 - 2\bar{u}\cos(\bar{\theta} + \gamma) + \bar{u}) + \frac{\mathcal{M}}{\overline{\mathcal{M}}}(1 - 2u\cos(\theta - \gamma) + u)} \end{aligned} \tag{2.60}$$

when exploiting Eq. (2.27) can be performed. The quantity $\delta = \theta - \bar{\theta}$ is the relative strong phase between the $B^0 \to D^{*+}D^-$ and $B^0 \to D^{*-}D^+$ decay amplitudes and

$$\overset{(-)}{\mathcal{K}}_{\pm} = \sin(2\beta + \gamma)\cos\overset{(-)}{\theta} \pm \cos(2\beta + \gamma)\sin\overset{(-)}{\theta} \tag{2.61}$$

is introduced for convenience. Interferences between tree and penguin contributions can cause the *CP* parameters $C_{D^{*+}D^-}$ and $C_{D^{*-}D^+}$ to be different from zero. The parameters $S_{D^{*+}D^-}$ and $S_{D^{*-}D^+}$ give access to $\beta = 2\phi_d^{\mathrm{eff}}$. As a golden mode, like $B^0 \to J/\psi K_S^0$, is not consicered, decays in $B^0 \to D^{*\pm}D^{\mp}$ give access to an effective phase ϕ_d^{eff}.

The *CP* parameters fulfil the relation

$$\sin(\phi_d^{\mathrm{eff}} \pm \delta) = -\frac{S_{D^{*\pm}D^{\mp}}}{\sqrt{1 - C^2_{D^{*\pm}D^{\mp}}}}\,, \tag{2.62}$$

where ϕ_d^{eff} is defined by

$$\phi_d^{\mathrm{eff}} = 2\beta + \Delta\phi_d\,. \tag{2.63}$$

Here, the phase shift $\Delta\phi_d$ includes the hadronic amplitudes and phases introduced by the loop, exchange and penguin annihilation decays. In case of *CP* invariance, $S_{D^{*+}D^-} = -S_{D^{*-}D^+}$ is fulfilled, which motivates the following combination of the parameters

$$\begin{aligned} S_{D^*D} &= \frac{1}{2}(S_{D^{*+}D^-} + S_{D^{*-}D^+})\,, \\ \Delta S_{D^*D} &= \frac{1}{2}(S_{D^{*+}D^-} - S_{D^{*-}D^+})\,, \\ C_{D^*D} &= \frac{1}{2}(C_{D^{*+}D^-} + C_{D^{*-}D^+})\,, \\ \Delta C_{D^*D} &= \frac{1}{2}(C_{D^{*+}D^-} - C_{D^{*-}D^+})\,. \end{aligned} \tag{2.64}$$

In case of *CP* invariance, S_{D^*D} and C_{D^*D} become zero. While ΔS_{D^*D} is related to the relative strong phase, the parameter ΔC_{D^*D} indicates how flavour specific a decay is: $\Delta C_{D^*D} = \pm 1$ denotes a complete flavour-specific decay, in which no *CP* violation in the interference between decay and decay after mixing is possible, while decays with $\Delta C_{D^*D} = 0$ have the highest sensitivity to mixing induced *CP* violation. With the relations from Eq. (2.64) the Eq. (2.26) can be written as

$$\begin{aligned} \Gamma(B^0 \to D^{*+}D^-)(t) = {}& \frac{e^{-t/\tau}}{8\tau}(1 + A_{D^*D})\times \\ & (1 - (S_{D^*D} + \Delta S_{D^*D})\sin(\Delta m_d t) + (C_{D^*D} + \Delta C_{D^*D})\cos(\Delta m_d t))\,, \\ \Gamma(B^0 \to D^{*-}D^+)(t) = {}& \frac{e^{-t/\tau}}{8\tau}(1 - A_{D^*D})\times \\ & (1 - (S_{D^*D} - \Delta S_{D^*D})\sin(\Delta m_d t) + (C_{D^*D} - \Delta C_{D^*D})\cos(\Delta m_d t))\,, \\ \Gamma(\bar{B}^0 \to D^{*+}D^-)(t) = {}& \frac{e^{-t/\tau}}{8\tau}(1 + A_{D^*D})\times \\ & (1 + (S_{D^*D} + \Delta S_{D^*D})\sin(\Delta m_d t) - (C_{D^*D} + \Delta C_{D^*D})\cos(\Delta m_d t))\,, \\ \Gamma(\bar{B}^0 \to D^{*-}D^+)(t) = {}& \frac{e^{-t/\tau}}{8\tau}(1 - A_{D^*D})\times \\ & (1 + (S_{D^*D} - \Delta S_{D^*D})\sin(\Delta m_d t) - (C_{D^*D} - \Delta C_{D^*D})\cos(\Delta m_d t))\,. \end{aligned} \tag{2.65}$$

Here, A_{D^*D} is the time- and flavour-integrated charge asymmetry that gives access to direct *CP* violation and is defined as follows

$$A_{D^*D} = \frac{\left(|A_{D^{*+}D^-}|^2 + |\bar{A}_{D^{*+}D^-}|^2\right) - \left(|A_{D^{*-}D^+}|^2 + |\bar{A}_{D^{*-}D^+}|^2\right)}{\left(|A_{D^{*+}D^-}|^2 + |\bar{A}_{D^{*+}D^-}|^2\right) + \left(|A_{D^{*-}D^+}|^2 + |\bar{A}_{D^{*-}D^+}|^2\right)}\,. \tag{2.66}$$

The main goal of the presented analysis is to measure the *CP* parameters S_{D^*D}, ΔS_{D^*D}, C_{D^*D}, and ΔC_{D^*D}. At a late stage in the analysis it was decided to measure A_{D^*D} additionally. Theoretical models predict possible direct *CP* violation contributions in $B^0 \to D^{*\pm}D^{\mp}$ to be a few percent [52]. The same level is expected for the deviation of the interference *CP* violation in $B^0 \to D^{*\pm}D^{\mp}$ decays compared to $\sin(2\beta)$ [53].

By studying modes from the family of $B \rightarrow D\overline{D}$ decays, a complementary measurement to that from charmonium decays can be carried out. In the group of the $B^0 \rightarrow D\overline{D}$ decays, apart from the $B^0 \rightarrow D^{*\pm}D^{\mp}$ decay, also the $B^0 \rightarrow D^+D^-$ and the $B^0 \rightarrow D^{*+}D^{*-}$ decays are sensitive to ϕ_d^{eff}. In a measurement of *CP* violation in $B^0 \rightarrow D^+D^-$, the four parameters S, ΔS, C, and ΔC are merged to one parameter for S and one for C, as the decay $B^0 \rightarrow D^+D^-$ is a *CP* eigenstate. In order to measure *CP* violation in decays of $B^0 \rightarrow D^{*+}D^{*-}$, the different spin configurations of the two vector mesons of the D^* mesons need to be separated.

3 The LHCb Experiment at the LHC

The Large Hadron Collider beauty (LHCb) experiment is one of four big experiments located at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN), near Geneva, Switzerland. CERN was established on the 29th September 1954 with the participation of 12 countries. Since then, the number of member states has risen to 23. In total, there are more than 2500 people employed at CERN and furthermore up to 12 200 scientists from 110 nationalities contribute to the experiments [54]. At one of the four interaction points of the LHC the LHCb experiment is located. The main focus of LHCb are indirect searches for physics beyond the SM in decays of b and c hadrons. The research program comprises precise measurements of rare decays and CP violation. To fulfil the needs of these analyses the LHCb detector is designed to provide an excellent vertex reconstruction for determining decay times and for an outstanding particle identification. One big challenge is the high track multiplicity occurring in the hadronic collisions of the LHC. Highly-developed systems are installed to reconstruct the physical processes in order to separate signal from background candidates. Nevertheless, data from simulated collisions is often additionally needed to achieve measurements of high precision. The following sections introduce the LHC (Sec. 3.1) and the individual components of the LHCb detector (Sec. 3.2). Furthermore, the LHCb software (Sec. 3.3) including details on the simulation are covered. Prospects of data processing at LHCb are discussed in Sec. 3.4. The information in this chapter are mainly based on Refs. [55, 56].

3.1 The LHC

The Large Hadron Collider (LHC) is a hadron-hadron collider, mostly colliding protons, with a circumference of 26.7 km located in the former Large Electron–Positron Collider (LEP) tunnel 50 to 170 m beneath the ground. Fig. 3.1 shows the CERN accelerator complex. The protons are pre-accelerated in the Linear Accelerator 2 (LINAC2), the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS) and in the Super Proton Synchrotron (SPS) before being injected into the LHC. In the LHC the beams are further accelerated to their nominal energies. Each beam consists of up to 2808 bunches and every bunch contains $1.15 \cdot 10^{11}$ protons. The original design of the LHC incorporates a luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ with a bunch spacing of 25 ns at a centre-of-mass energy of $\sqrt{s} = 14\,\mathrm{TeV}$. In separate beam pipes the protons circulate in opposite directions. These high-energetic beams are controlled by magnetic fields with a strength of up to 8.33 T. A number of 1232 NbTi superconducting dipole magnets, which are cooled to a temperature of $-271.9\,°\mathrm{C}$ using liquid Helium, keep the protons on their circular track. In addition, quadrupole magnets quench the beam.

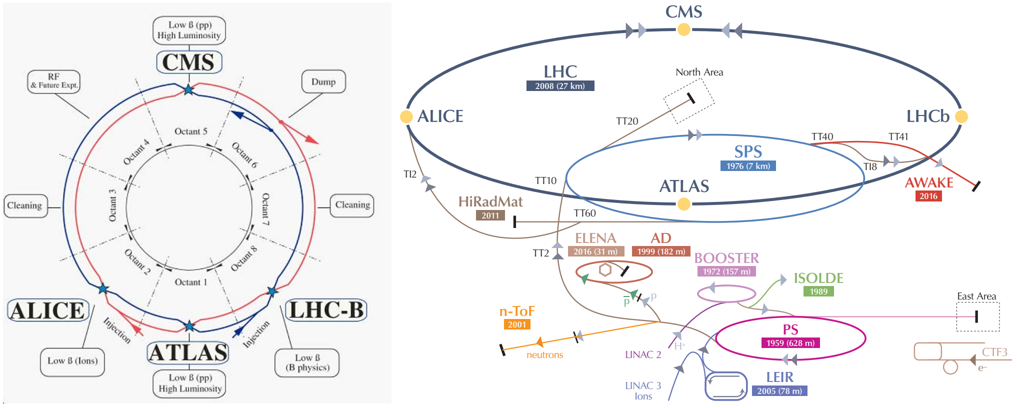

Figure 3.1 – Scheme of the LHC (left) [57]. The red beam circulates clockwise while the blue beam circulates counter-clockwise. The CERN accelerator complex is shown on the right [58]. After passing the LINAC2, BOOSTER, PS and SPS the beams are injected into the LHC. The beams collide at the collisions points, where the detectors of ATLAS, ALICE, CMS and LHCb are positioned.

Seven experiments are distributed at the four collision points. Besides LHCb, there are three further major experiments namely, ATLAS [59], CMS [60] and ALICE [61], and three small experiments, TOTEM [62], MoEDAL [63] and LHCf [64]. The two largest experiments, ATLAS (A Toroidal LHC Apparatus) and CMS (Compact Muon Solenoid), are general-purpose detectors (GPDs). Their field of research is large and ranges from the exploration of dark matter, over hints for supersymmetry to the search for exotic particles. Another area of interest are processes involving top quarks. One of the most notable measurements of ATLAS and CMS is the discovery of the Higgs boson [65] in 2012.

At the ALICE (A Large Ion Collider Experiment) experiment a quark gluon plasma is investigated, which is a state of high temperature and density simulating the universe shortly after the Big Bang. In order to create the quark gluon plasma, lead-ions are brought to collision in the LHC instead of protons.

TOTEM (Total Elastic and Diffractive Cross Section Measurement) is a smaller experiment that makes use of the collisions at the CMS interaction point. The research program aims at the precise measurement of the proton-proton interaction cross section as well as the detailed investigation of the proton structure. MoEDAL (Monopole and Exotics Detector at the LHC) is installed close to the collision point of the LHCb experiment and searches for highly ionising particles such as magnetic monopoles or electrically charged, massive particles. Such particles are predicted in various theoretical models describing physics beyond the standard model. The LHCf (Large Hadron Collider forward) experiment studies particles that are emitted in the forward direction. The cascades created in the collisions are similar to the ones appearing in the upper atmosphere caused by cosmic rays. The results can be used to calibrate large-scale cosmic-ray experiments. LHCf is located at the same interaction point as ALICE.

The first data taking period, the so-called Run 1, took place from 2010 to 2012 at centre-of-mass energies of $\sqrt{s} = 7\,\text{TeV}$ (2010 and 2011) and $\sqrt{s} = 8\,\text{TeV}$ with a bunch spacing of 50 ns. For the Run 2 data taking period from 2015 to 2018 the energy was increased to $\sqrt{s} = 13\,\text{TeV}$ and the bunch spacing has been reduced to 25 ns. The presented analysis is performed on data recorded in 2011 to 2012 and 2015 to 2018.

3.2 The LHCb detector

The details about the LHCb detector given in this sections are based on [56, 66], if not quoted differently. The LHCb detector, depicted in Fig. 3.2, is designed as a single-arm forward spectrometer. Hence, the individual detector components are ordered sequentially in the forward direction, starting from the point of interaction. In contrast to that, the experiments ATLAS and CMS have a 4π detector that covers almost the entire solid angle range. LHCb uses a right-handed coordinate system. The z-axis starts at the collision point and extends along the beam-axis into the detector. The y-axis is oriented vertically and the x-axis is oriented horizontally. At the LHC, b quarks are produced dominantly by gluon fusion and quark-antiquark annihilation known as pair creation. Comparing these two processes, the bigger contribution comes from gluon fusion, since at the energy scale of the LHC the parton distribution functions of the protons are dominated by gluons. Due to the low production threshold of a $b\bar{b}$ pair compared to the centre-of-mass energy of the LHC, it is very likely that gluons with different momenta create the $b\bar{b}$ pairs. This implies that the $b\bar{b}$ pairs receive a large boost along the beam axis. The boost causes a relative long flight distance, which leads to well-separated decay vertices. This is exploited by the LHCb experiment to reach an

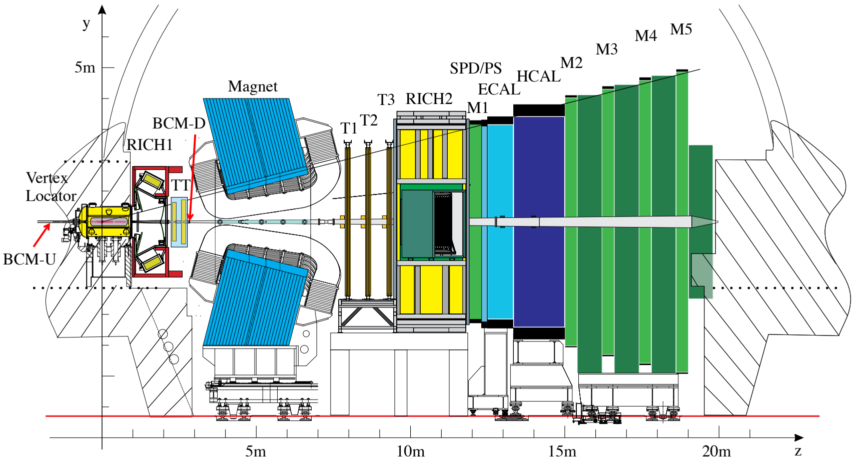

Figure 3.2 – Schematic side view of the LHCb detector. The pp interactions take place on the far left in the vertex locator. It is followed to the right by the ring-imaging Cherenkov detector (RICH1), the Tracker Turicensis (TT), the magnet, the tracking stations T1–T3, RICH2, one muon chamber, the scintillator pad and preshower detectors (SPD, PS), the electromagnetic (ECAL) and hadron (HCAL) calorimeters and four more muon chambers [55].

excellent decay-time resolution of approximately 50 fs. In Fig. 3.3 the simulated angular distribution of $b\bar{b}$ pairs at a centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$ is presented. These facts motivate the design of the LHCb detector that covers a pseudo-rapidity range of $2 < \eta < 5$. The instrumented angular range in the vertical plane is 10 mrad to 250 mrad and in the horizontal plane 10 mrad to 300 mrad. With this construction, 24 % of all produced $b\bar{b}$ pairs are within the detector acceptance, despite covering only 4.5% of the entire space.

In contrast to the pairwise $b\bar{b}$ production, the hadronisation processes to B^0 (B^+) and $\bar{B}^0$ (B^-) mesons are asymmetric. This is caused by the interaction of the b and $\bar{b}$ quarks with the beam remnants, comprising the proton valence quarks [68–70]. The production asymmetry between B^0 and $\bar{B}^0$ mesons

$$A_{\text{prod}} = \frac{R_{\bar{B}^0} - R_{B^0}}{R_{\bar{B}^0} + R_{B^0}} \tag{3.1}$$

needs to be taken into account in an analysis of *CP* asymmetry like the one presented in this thesis. It is defined by the relative variation between the production rate $R_{\bar{B}^0}$ for $\bar{B}^0$ mesons and the production rate R_{B^0} for B^0 mesons.

In order to perform precision measurements, it is necessary to reconstruct decay

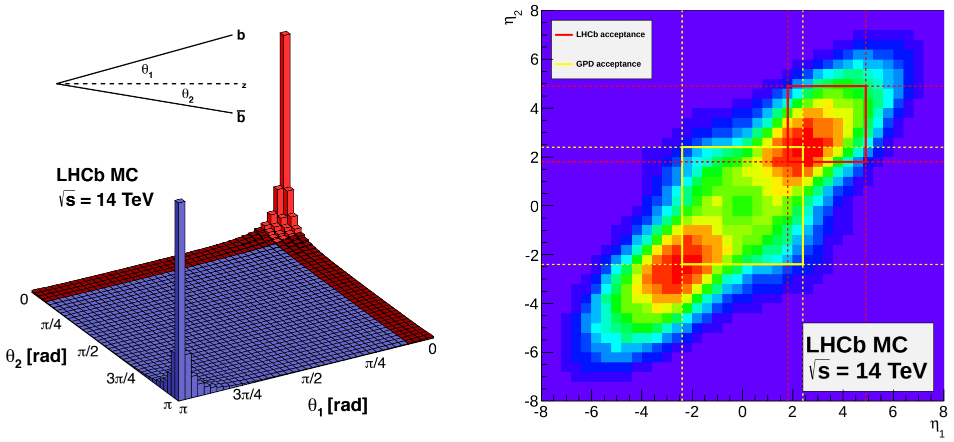

Figure 3.3 – Correlation between $b\bar{b}$ pair production and (left) the azimuthal angles $\theta_{1,2}$ and (right) the pseudorapidities $\eta_{1,2}$. The red area in the left plot represents the acceptance of the LHCb detector. In the right plot this is indicated by a red box, while the yellow box shows the area covered by a GPD like CMS or ATLAS. [67]

processes as accurately as possible. This is particularly challenging in the hadronic environment of the LHC. Therefore, the number of pp collisions per bunch crossing at LHCb is reduced, which leads to a lower number of tracks and a smaller constant instantaneous luminosity of $\mathcal{L} \approx 4 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This so-called luminosity levelling is achieved by displacing the two beams against each other. At beam interaction rates of 20 million proton bunches per second, 10 million visible processes per second are created. Since these events also contain a large number of events without B hadrons, an efficient selection is needed. The selection consists of online and offline steps, which are described in Secs. 3.2.3 and 3.3.2.

The detector components can be divided into two systems: The track reconstruction system and the particle identification system. The former is responsible for the detection of hits that are are used to locate the tracks of the passing particles, while the latter classifies the particles. In the following, the components of these systems are described in detail.

3.2.1 Track reconstruction system

The track reconstruction system comprises the vertex locator that is positioned at the interaction point, the Tracker Turicensis and the tracking stations T1–T3 (see Fig. 3.2). It is complemented by the magnet that bends the tracks of the charged particles.

Vertex Locator

The B hadrons are produced at the pp collision point, which is called the primary vertex (PV). After flight distances of about ~ 1 cm, they decay into daughter particles forming the secondary vertex. To enable a high decay time resolution, these two vertices must be reconstructed very precisely, which is the purpose of the VELO. The VELO is a silicon micro-strip detector and consists of 42 silicon half-disks with diameters of 42 mm, which are aligned cylindrically perpendicular to the beam axis and are positioned close to the pp interaction point. Fig. 3.4 shows the arrangement of the VELO disks and the disk structure. A silicon disk consists of R and ϕ sensors, which detect the distance (R) and angular coordinate (ϕ) of a passing charged particle. Thus, it can indicate the position of a single hit in cylindrical coordinates. If hits are registered with at least three different modules of the VELO, their track can be reconstructed in the VELO. To prevent collisions with air molecules, the VELO is under vacuum. The angular acceptance of the VELO is in the range of 15 to 390 mrad [71].

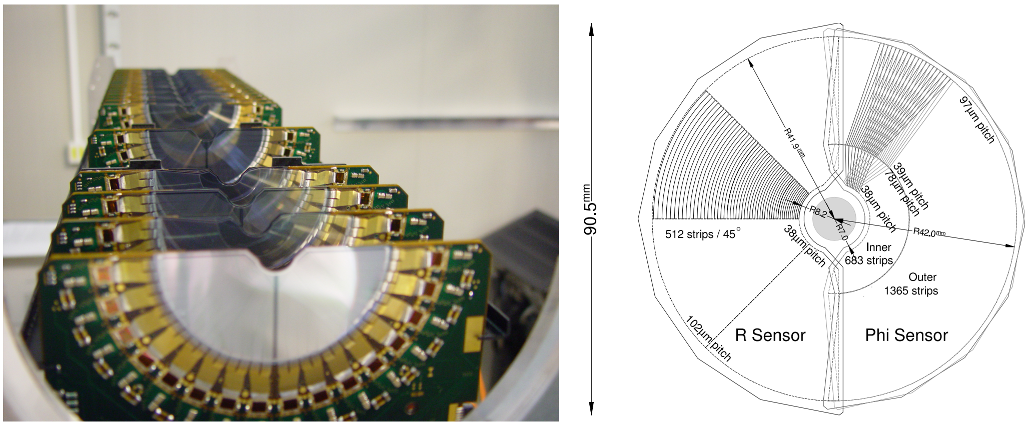

Figure 3.4 – Photograph showing the arrangement of the VELO sensors (left) [72]. Illustration of the $R\phi$ geometry of one single sensor (right) [55].

Before the actual data taking, the VELO modules are kept at a distance of 3.5 cm from the beam axis. As soon as stable beam conditions are ensured, the modules are moved closer to the beam to a distance of about 7 mm. The stability of the beam is continuously monitored by the Beam Conditions Monitor that consists of two modules positioned directly before and after the interaction point around the beam pipe. These two modules measure the particle rate at these positions. Each module consists of eight diamond sensors to provide a high radiation tolerance. If certain thresholds for the particle flux are exceeded, the beam is automatically directed into a graphite block. Uncontrolled beam conditions would otherwise damage important detector components close to the beam, such as the VELO.

Tracking stations and Tracker Turicensis

The remaining part of the track reconstruction system consists of the TT and the tracking stations T1–T3. Each tracking station can be divided into an outer part, the so-called Outer Tracker (OT), and an inner part, which is called Inner Tracker (IT). This structure is utilised to resolve the much higher track multiplicities close to the beam axis. The TT and IT are Silicon Tracker (ST), while the OT is a straw tube drift-gas detector. The TT is located between the VELO and the magnet, while the tracking stations are placed downstream of the magnet. With this arrangement, particles with low momentum that are bent out of the detector by the magnet can also be reconstructed. The alignment of the silicon strips of the ST and the straw-tubes of the OT follow a x-u-v-x alignment. The first and the last layer are aligned vertically, while the layers in between are tilted by $\pm 5^\circ$ around the beam axis. These so-called stereo layers enable a sensitivity in the direction of the y-axis. The gas mixture in the OT consists of 70% Argon, 28.5% CO_2 and 1.5% O_2.

LHCb's magnet is located between the TT and the tracking stations. It is a warm dipole magnet that bends the tracks of charged particles with an integrated magnetic field of 4 T m. Depending on the deflection, the momentum of the particles can be measured. The momentum resolution results in $\Delta p/p = 0.5\%$ for a low momentum and rises to $\Delta p/p = 1.0\%$ for momenta at about 200 GeV/c. A good momentum resolution enables a better mass resolution and thus reduces the combinatorial background. The effective time resolution for decays with six final state particles, as in the decay analysed in this thesis, is about 57 ps. This leads to a mass resolution of approximately 10 MeV/c^2.

3.2.2 Particle identification system

In addition to the reconstruction of tracks, it is necessary to identify the particles that pass the detector. That is the purpose of the particle identification system. It comprises two ring-imaging Cherenkov detectors (RICHs), the calorimeters and the muon chambers. Combining the information of all components leads to a probability for particle type hypotheses. The components are described in the following.

Ring-imaging Cherenkov detectors

At the LHCb experiment the RICH detectors are used to separate kaons, protons and pions from each other. In order to achieve this distinction, the Cherenkov effect is exploited. This effect describes that charged particles emit photons at a defined angle ϑ, if they are travelling through a medium with refractive index n at a higher speed v than the specific speed of light in that medium. For the detection, spherical and flat mirrors (see Fig. 3.5) are installed, so that the Cherenkov light propagates through the RICH to the hybrid photon detector. Considering the photon emission angle ϑ and the relation $\cos(\vartheta) = c(nv)^{-1}$, the velocity of the particles can be measured. Here, c denotes the speed of light in vacuum. In combination with the momentum measurement from the tracking system, different particle hypotheses can be analysed. The efficiency to identify a kaon is 95% at a pion mis-identification rate of 10%. When applying harder requirements, the mis-identification rate can be reduced to 3%, while maintaining an efficiency of 85%. RICH1 was filled with a mixture of aerogel and C_4F_{10} during Run 1 and is optimised for particles of the low- and intermediate-momentum region of 2 GeV/c to 40 GeV/c. In Run 2, only C_4F_{10} was used as a radiator. RICH2 is filled with CF_4 and covers a high-momentum range of 15 GeV/c to 100 GeV/c [73].

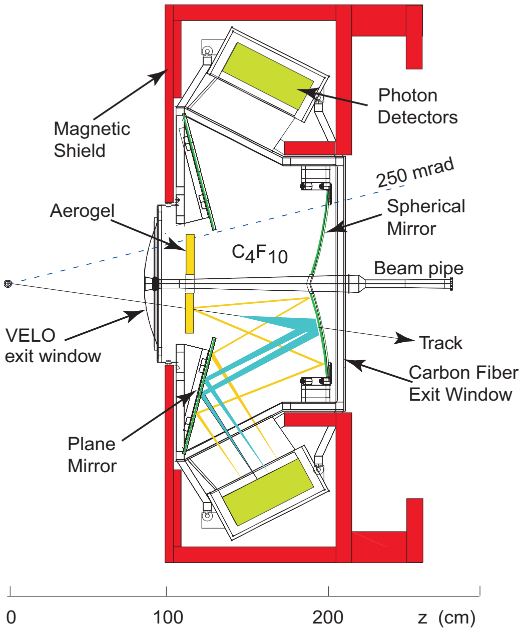

Figure 3.5 – Schematic side view showing RICH1 and its optical systems [55].

Calorimeters

The calorimeter system provides information that allows to distinguish between photons, electrons and hadrons. In addition, the deposited energies and flight positions of the particles are measured. These measurements are essential to decide 4 μs after the collision whether an event is stored or rejected. The calorimeter system comprises the scintillator pad detector (SPD), the preshower detector (PS), the electromagnetic calorimeter (ECAL) and the hadron calorimeter (HCAL). They are positioned in the same order as mentioned and downstream to RICH2. The SPD is installed to distinguish between electrons and photons. Electrons shower both in the SPD and in the ECAL, whereby the photons create showers in the ECAL only. With the PS it is possible to separate hadronic and electromagnetic showers. In contrast to electrons, hadrons deposit very little energy in the PS. All calorimeters are constructed as an alternating order of metal layers and scintillators. In this structure, the metal serves as an absorber material, in which passing particles can produce showers of secondary particles. These particles emit light in the scintillators. Lead is used as an absorber in the ECAL, while the HCAL utilises iron. The ECAL provides the measurement of the energies of photons and electrons. The HCAL is located directly behind the ECAL and measures the energy of hadrons.

Muon Chambers

Five muon detectors (M1–M5) are deployed for the detection of muons. Muons have a very small cross section of interaction with matter and, therefore, can pass trough all previously described detector components. Due to this fact, the muon chambers are located at the end of the LHCb detector. To cross all five chambers, a muon must have a minimum momentum of about 6 GeV/c. The first muon chamber M1 is positioned between RICH2 and the calorimeters, while the remaining M2–M5 are downstream of the calorimeters. They are separated by iron absorbers with a thickness of 80 cm. The chambers are Multi Wire Proportional Chambers (MWPCs), except for the inner region of M1. As there is a high particle rate present, it is constructed as a gas-electron multiplier (TripleGEM).

3.2.3 Trigger

The possible recording speed and the available storage space are not sufficient to store the data of all collisions. However, it is known that just a small fraction of all collisions contain decays that are interesting for physics analyses. The purpose of the trigger is to select the relevant events. After the trigger selection, the event rate is reduced to 5 kHz (12.5 kHz) in Run 1 (Run 2). The trigger consists of two stages: the hardware (L0 Trigger) and the software trigger (high level trigger), as illustrated in Fig. 3.6.

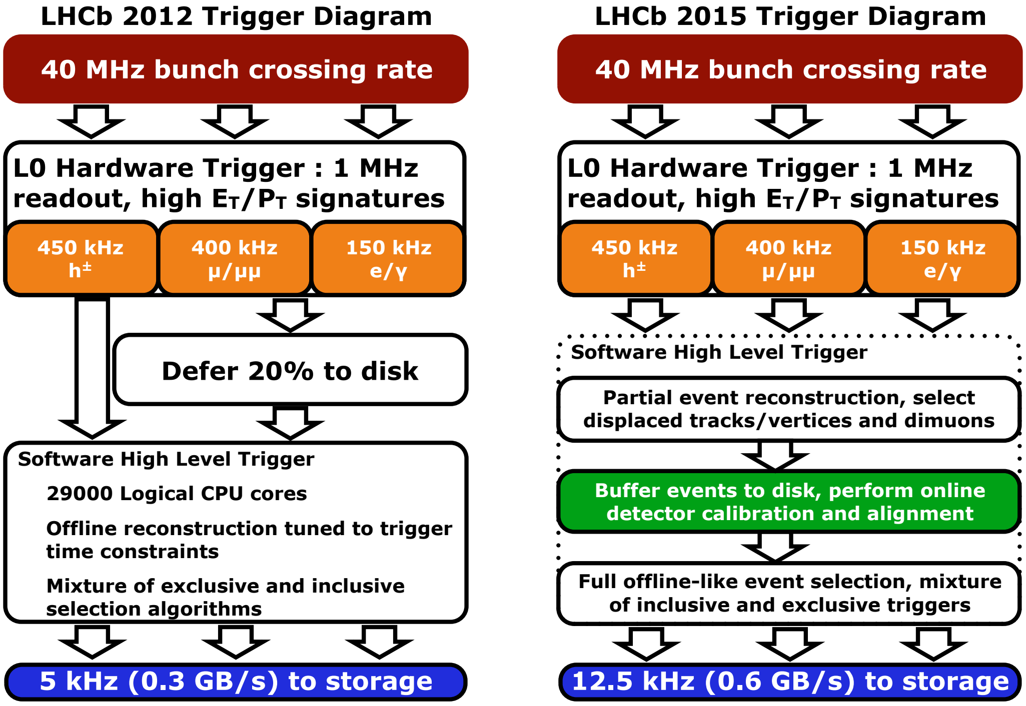

Figure 3.6 – Trigger scheme for (left) Run 1 and (right) Run 2. The collision rate of 40 MHz is reduced to 5 kHz in Run 1 and to 12.5 kHz in Run 2. [74]

L0 Triggers The first stage of the trigger is a hardware trigger, the so-called level-zero trigger. It operates at the full bunch crossing rate of 40 MHz. Necessary for decisions taking at this high frequency are extremely quick detector responses, which are provided by the muon an the calorimeter systems. The trigger uses information from the calorimeters to search for hadrons, electrons and photons with the highest transverse energy to identify hadron candidates (`L0Hadron`), electron candidates (`L0Electron`) and photon candidates (`L0Photon`). Additional events are triggered, if they include a muon with a minimum threshold for the transverse momentum (`L0Muon`) or if two muons in combination exceed a minimum of transverse momentum (`L0DiMuon`). In addition to these L0 triggers, the L0 pile-up system searches for events with a lower PV multiplicity. Higher multiplicities lead to an increased mis-identification probability, which subsequently complicates physics analyses and requires even more time in the reconstruction. With the much lower rate of 1 MHz after the L0 trigger, it is possible to perform a more detailed analysis covering more sophisticated reconstruction algorithms in the next stage.

High Level Trigger The high level trigger consists of C++ algorithms, which run on the big computing farm, the so-called event filter farm, at CERN. The HLT consists of two stages, the HLT1 and HLT2. HLT1 performs a partial event reconstruction with information from the VELO and the tracking stations. It searches for high transverse momentum (p_{T}) candidates to confirm the events selected at L0 stage. HLT1 reduces the rate further, which allows more time-consuming algorithms at the HLT2 stage. HLT2 has access to the full event information. By reconstructing the full event it searches for inclusive and exclusive decays. While specific decays are triggered in the exclusive case, a more general approach is used in the inclusive one: Events that contain at least two charged particles with a displaced decay vertex are selected.

As presented in Fig. 3.6 the trigger configuration changed between Run 1 and Run 2. The biggest change is the real-time alignment and calibration introduced in Run 2. During data taking the recorded data is used to perform new calibrations and alignments in the detector components. This real-time corrections improve the physics performance [75]. More details on the performance and configurations of the trigger are given in Refs. [76–78].

3.3 LHCb Software

The LHCb software is based on the GAUDI framework [79] that contains different packages. One package is MOORE [80] that comprises the HLT software already described in Sec. 3.2.3. The following section will concentrate on the packages that follow in the data processing chain after MOORE, which are BRUNEL [81], that is responsible for the full event reconstruction, and DAVINCI [82] that is used to perform a centralised preselection, the so-called stripping.

3.3.1 Event reconstruction

This section summarises how the electrical responses from the LHCb detector are processed to receive fully reconstructed events. To enable physical analyses of decays, the trajectories of the particles involved, the position of the vertices and information about the particle types are fundamental.

Tracking

The trajectories of charged particles, so-called tracks, are created by combining electrical signals, so-called hits, in each subdetector of the tracking system (VELO, TT, IT, OT). Then, pattern recognition and tracking algorithms are used to identify specific signatures in the subdetectors. Different track types are introduced depending on the subdetectors that contribute to the track. In the presented analysis only so-called long tracks are relevant. Tracks of this type cover all detector components. Two kinds of algorithms exists to search for long tracks. They are called forward tracking algorithm and track matching algorithm. Both start with a track in the VELO, by combining hits that lie on a straight line. In the forward tracking the VELO track is extended to match hits in the stations T1–T3, considering the deflection by the magnet. The last step is the addition of TT hits to the track. Furthermore, the track matching algorithm combines the VELO track with an independently formed track in the tracking stations. For a long track further compatible hits from the TT are added.

Finally, a global Kalman fit is performed [83]. Taking into account multiple scattering effects and energy losses. It determines the tracks and momenta with high precision.

Particle Identification

After the tracks have been established, particle hypotheses need to be assigned to them. Therefore information from all components of the particle identification system, RICH1, RICH2, the calorimeter and muon system, is utilised. To identify charged hadrons, the tracks are extrapolated into the RICH detectors. Combining the tracks and pixel hits a global likelihood is calculated by variation of the particle hypotheses. The identification of muons is performed very efficiently by searching for hits in the muon system near extrapolated traces. In the end each track has the likelihoods for each particle hypothesis. The likelihood of hypothesis h is generally calculated as the difference to the pion hypothesis

$$\mathrm{DLL}_{h\pi} = \Delta \log \mathcal{L}_{h\pi} = \log \mathcal{L}_h - \log \mathcal{L}_\pi \,, \tag{3.2}$$

as pions are the most commonly produced and detected particle type at LHCb. The greater the value of $\mathrm{DLL}_{h\pi}$, the higher the probability that the particle is an h-type particle. Smaller values $\mathrm{DLL}_{h\pi}$ indicate a pion.

3.3.2 Stripping

The last step that needs to be accomplished before the analysts access the data is the global event filtering, the so-called stripping. Within the LHCb software framework it is implemented in the DaVinci project. After the event reconstruction proto-particles are saved. In the DaVinci project these proto-particles and their particle hypotheses are combined to intermediate state particles until the full decay chain is reconstructed. In this combination, DaVinci calculates various observables that can be used for the physics analyses. It gives access e.g. to many kinematic observables of all particles involved, to the trigger responses and various track and vertex fit qualities. This output is used to develop selections to identify the specific signal decays and to discard background as much as possible.

The stripping is performed globally at LHCb, in order to avoid that all analysts have to process the full set of data. As many studies are performed in very similar decays like e.g. $B \to DX$, it would be very inefficient with regard to computing resources, if the complete dataset would be reconstructed in each analysis. In the stripping a set of requirements that are defined by the analysts are applied. These requirements are looser than applied for the final selection to enable more studies, but implying more background candidates at this stage. The analysts then use the data after the stripping to create ROOT ntuples, that comprise all relevant observables needed for the analyses.

3.3.3 Simulation

In many physical analyses it is useful to check the correctness of the analysis strategy. For this purpose simulated data, often called Monte Carlo (MC) can be used. Furthermore, different experimental effects influence the data and need to be understood. Simulated data is produced to be as similar as possible to real data, taking into account the run conditions and configuration of the detector. The Gauss package [84, 85] is used to generate the physics processes. In this package, external generators, such as Pythia [86, 87] to generate pp collisions are utilised. EvtGen [88] simulates decays of B mesons, while Geant4 [89, 90] considers the behaviour of the particles in the magnetic field and the interaction with the detector material. Subsequently, the Boole package [91] processes the output of Gauss and simulates the response of the detector in each subdetector. From this point on, the rest of the process is analogous to the data processing of real data and starting the application Moore. The advantage of MC datasets is that the true information about the processes and their properties is stored and can be retrieved at the end of processing. Hence, underlying physical quantities, such as the lifetime τ_{B^0} or CP parameters, are known. The correctness of the reconstruction can be proven on MC using the so-called truth matching, where reconstructed particles are compared to the originally generated ones. The true decay time can be used, for example, to determine resolution effects. In addition, the knowledge gained from MC datasets can be transferred to a certain extent to real data. To check an analysis, it can be run on MC at first and then the results can be compared with the original values used during generation.

3.4 Prospects of data processing at LHCb

In the context of the collaborative research centre CRC 876 at the TU Dortmund University there is a joint effort of the physics and computer science departments to investigate the feasibility of using new industry standards for data processing in high energy physics. Currently, the LHCb software is a custom development of primarily physicists. Some of the core elements had to be implemented nearly 15 years ago. In particular, the components for data processing were developed from scratch due to the lack of available solutions capable of handling the amount of produced data. Since then, the demand for Big Data applications has strongly increased, especially from the industry. This demand has led to the creation of industry standards built to store and process huge amounts of data.

The author contributed to the sub-projects C5 of the CRC 876, called real-time analysis and storage for high-volume data from particle physics. The C5 project has two main objectives. Firstly, it is evaluated, whether applications of Field Programmable Gate Arrays (FPGAs) and General Purpose Computation on Graphics Processing Units (GPGPUs) can enable future performance increases of the HLT. Secondly, the distributed storage and analysis of data with modern data processing technologies is investigated. As the author was involved in this second topic, the following paragraph summarises some of the work being done.

Using modern data processing technologies in particle physics pursues the vision of accessing and analysing data by sending and receiving queries from a big database. However, there is a gap between complex analyses, which are performed in particle physics, and query capabilities. Hence, together with computer scientists ”DeLorean” [92] was developed, an intelligent storage back-end meant to accelerate data analyses. It is built on top of Apache Drill [93], which is the open source counterpart to Google’s Dremel system [94]. Apache Drill follows the Hadoop approach for storing and processing huge amounts of data. In its core, this approach utilises Google’s MapReduce [95] algorithm to divide large processing tasks into smaller ones. As the tasks can be processed in parallel, large computing cluster can be fully utilised to speed up the processing massively. The second main part of the Hadoop approach is HDFS [96], which enables the storage, access and management of data in a distributed way on clusters.

In DeLorean, the analysis tasks are separated into a data-intensive part and a compute intensive part. The former is the equivalent to the filtering, currently called the stripping, while the latter covers the actual analysis. In order to reduce the data volume, DeLorean makes use of column-store technology. Additionally, the scanning of the data is optimised by creating a synopsis of the original dataset in the column-store. The performance of Delorean is evaluated in terms of a reference analysis. In this analysis, simulated data is used to create two muon tracks, combine them and compute their invariant mass. Furthermore, a cut on the mass is applied in order to select the decay $B_s^0 \to \mu^+\mu^-$.

The DeLorean software shows performance increases in the compute intensive part of up to 4.6 compared to DaVinci. More tests are necessary, especially using complex analyses. However, the first steps are made on the way to implement modern database technologies in particle physics.

4 Data Analysis Techniques

This chapter introduces data analysis methods and techniques used in the course of this thesis. It starts in Sec. 4.1 with an explanation of the maximum likelihood formalism, which is the foundation of the performed fits in this thesis. Afterwards, in Sec. 4.2, the sPlot technique is discussed, which is used to subtract background from signal, while in Sec. 4.3 boosted decision trees are introduced, which are implemented as a multivariate algorithm that is utilised by means of machine learning for classification. In addition, the bootstrapping method (Sec. 4.4), a statistical method for resampling, and the decay-chain fitting (Sec. 4.5) are described. Finally, the flavour tagging is explained in Sec. 4.6, which is utilised to determine the initial flavour of the B^0 meson.

4.1 Maximum likelihood method

The maximum likelihood method is a statistical technique for the determination of parameters and their uncertainties involving an adjustment of a model in form of a probability density function (PDF) to the distribution of measured data [97–99]. Thus, a small amount of relevant information can be obtained from a large amount of data. Given are N independent measurements $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N$ of a random variable $\vec{x}$, where $\vec{x}$ can be a single variable or a vector of variables. The goal is to determine the best estimate $\hat{\theta}$ for the (one or more) parameters $\vec{\theta}$ of the corresponding PDF $\mathcal{P}(\vec{x}|\vec{\theta})$. For this purpose, the maximum likelihood technique can be used, which derives from a one- or multi-dimensional probability density function $\mathcal{P}(\vec{x}|\vec{\theta})$ of the measurements to construct the likelihood function $\mathcal{L}(\vec{\theta})$. It is given by the product of individual probability densities of the measured values as a function of the parameter set $\vec{\theta}$

$$\mathcal{L}(\vec{\theta}) = \mathcal{P}(\vec{x}_1|\vec{\theta}) \cdot \mathcal{P}(\vec{x}_2|\vec{\theta}) \cdot \dots \cdot \mathcal{P}(\vec{x}_N|\vec{\theta}) = \prod_{i=1}^{N} \mathcal{P}(\vec{x}_i|\vec{\theta})\,. \tag{4.1}$$

The likelihood function indicates an estimate of the probability to measure $\vec{x}$ for a given set of parameters $\vec{\theta}$. The best estimate $\hat{\theta}$ is given by the maximum of the likelihood function. In practice, it is easier to maximise a sum than a product, hence, a logarithm is applied to the likelihood function. In addition, most algorithms solve minimisation rather than maximisation problems, which can be achieved by applying a negative sign to the log likelihood function. This results in the negative log likelihood function

$$-\ln \mathcal{L}(\vec{\theta}) = -\sum_{i=1}^{N} \ln \mathcal{P}(\vec{x}_i|\vec{\theta})\,. \tag{4.2}$$

Since the number of measured events follows a Poisson distribution, a Poisson term $\mathrm{e}^{-N}N^n/n!$ is multiplied to the log likelihood function

$$\mathcal{L}_{\text{ext}}(\vec{\theta}) = \frac{\mathrm{e}^{-N}N^n}{n!} \prod_{i=1}^{n} \mathcal{P}(\vec{x}_i|\vec{\theta})\,. \tag{4.3}$$

This term defines the probability to observe n events, when N are expected. The modified log likelihood function $\mathcal{L}_{\text{ext}}$ is called the extended maximum likelihood function. Furthermore, by multiplying a Gaussian PDF with mean μ and width σ

$$\mathcal{G}(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.4}$$

to the likelihood function, creating a so-called Gaussian constraint, external input parameters are allowed to float within their uncertainties in the fit.

4.2 sPlot technique

Background candidates can remain in the dataset, after the full signal selection chain is applied. To subtract these background candidates statistically, the sPlot method [100] can be used. It requires an observable that is distributed differently for signal and background candidates. Moreover, the PDF of this observable need to be known and it needs to be uncorrelated to the actual observables of interest. A suitable discriminating observable is the reconstructed B^0 mass. By applying an extended maximum likelihood fit (see Sec. 4.1) to the mass distribution, so-called sWeights can be determined for all candidates. In this extended maximum likelihood fit all parameters are fixed values from a previous extended maximum likelihood fit, leaving just the yields of the different components floating. The sWeights are defined as

$${}_s\mathcal{P}_{\mathrm{n}}(y_e) = \frac{\sum_{j=1}^{N_s} V_{\mathrm{n}j} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}\,. \tag{4.5}$$

The indices j and k sum over the N_s components comprised in the PDF f. The covariance matrix $V_{\mathrm{n}j}$ is given by the second derivative of the negative log likelihood function and contains the covariances between the yields. The sum over all sWeights in one component is equivalent to the fitted yield value. In the decay-time-dependent *CP* fit, these sWeights are applied as weights to statistically subtract the background. If sWeights are applied to the candidates in a histogram the uncertainty on each bin content i is defined as

$$\sigma(\mathrm{bin}) = \sqrt{\sum_{i \subset \mathrm{bin}} ({}_s\mathcal{P}_{\mathrm{n}})^2}\,. \tag{4.6}$$

4.3 Boosted decision tree

A boosted decision tree (BDT) is a multivariate method to solve decision problems and to classify data. Multivariate classifications have the advantage that they can take correlations of input features into account. In the scope of the presented analysis BDTs are mostly used to separate signal from background candidates. BDTs are an advancement of simple decision trees. A decision tree [101] applies ordered requirements to a dataset to sort the individual data points into a set of classes. As presented in Fig. 4.1, a decision tree can be illustrated as a flowchart. In a binary tree – as used in this thesis – that distinguishes between two classes (signal and background), the decision is made by starting at a root node and, depending on the result of the rule applied, leading to one of the two possible branches. Then, it continues downwards until a leaf is reached. At each node a decision is made with respect to an attribute, and a rectangular cut is applied to separate the two classes. These rules are determined in the training phase of the classifier. During the training, labelled data, here the labels are signal and background, is used to build the tree. In order to identify the variable with the highest separating power, the Gini index $p \cdot (1 - p)$, with p given by the signal purity, can be utilised. Once the leaf is reached, each data object receives its classification. The depth of the tree corresponds to the number of cuts and is configurable. The accuracy of the classification can be improved if not only one tree, but multiple trees that are created by a boosting method are used. A weighted average of all trees is calculated, in order to improve the performance. For this method, decision trees are created consecutively. Whenever a tree has been trained, the data points that have been classified incorrectly are given a higher weight (boosting) making them more relevant for the training of the next tree. The boosting method that was used in the scope of this analysis is the adaptive boost (AdaBoost) method [102]. The weights of data points that are mis-classified in the

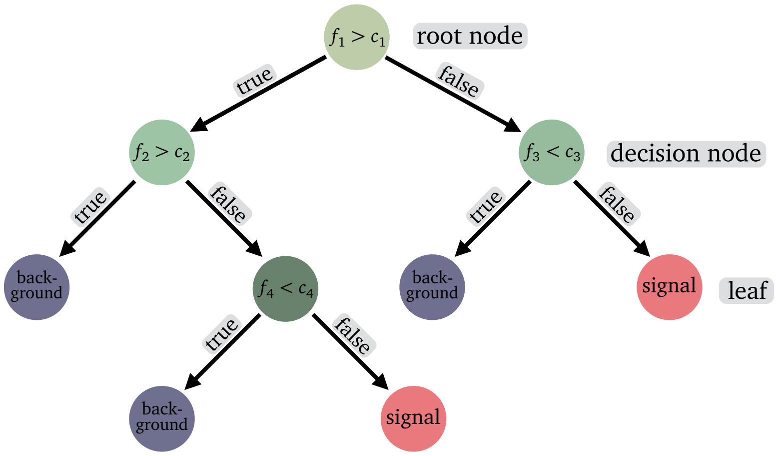

Figure 4.1 – Schematic illustration of a binary decision tree. Cuts in the features f_i are applied at cut values c_i to separate the signal from the background.

previous tree are multiplied with the boost weight

$$\alpha^{\beta} = \left(\frac{1-\epsilon_{\mathrm{mis}}}{\epsilon_{\mathrm{mis}}}\right)^{\beta}, \tag{4.7}$$

where ϵ_{mis} is the mis-classification rate. The learning rate, which controls how much the weights are adjusted, can be controlled by the parameter β. The weights are normalised so that the sum of all weights is always constant. In order to determine the total classification for one data point y_{Boost}, the weighted decisions of all classifiers N_{trees} are cumulated via

$$y_{\mathrm{Boost}} = \frac{1}{N_{\mathrm{trees}}} \cdot \sum_{i=1}^{N_{\mathrm{trees}}} \ln(w_i) h_i \,. \tag{4.8}$$

Here, w_i refers to the weight and h_i denotes the classifier's decision, which is -1 for background and $+1$ for signal. The configuration of a BDT includes the selection of variables with the highest possible separation power and the hyperparameters, such as number of trees for training, tree depth, boost factor (β) and the minimal amount of candidates that has to remain in a leaf. When selecting training variables, it should be noted that the differences between the signal and background datasets that contribute to the separation power have a physical nature and are not based on non-physical properties, which are not present in real data. There is a potential difficulty that the multivariate classifier will train on exactly these differences to distinguish signal from background. Generally it should be prevented that fluctuations are considered by the classifier. This phenomenon is referred to as overtraining or overfitting, whereas it can be controlled by tuning the hyperparameters. A less complex BDT is more robust against overtraining.

4.4 Bootstrapping method

The bootstrapping method [103] is a frequentist statistical procedure allowing to determine precise uncertainty estimates. For this technique, datasets of the same predefined size are sampled randomly from the original dataset. This random sampling is executed with replacement. As a consequence, some of the original data points occur more than once in the resampled dataset, others do not appear at all. One advantage is that the correlations to other observables are preserved, enabling measurements to be carried out in multiple dimensions. Usually, uncertainties are estimated in the fit to the one dataset that has been measured. However, more reliable statements about uncertainties can be made by applying the bootstrapping method.

4.5 Decay-chain fitting

In contrast to a vertex-to-vertex reconstruction, starting with the fitting at the final state particles most downstream and getting step by step upstream until the mother particle is reached, the decay-tree (or decay-chain) fitter (DTF) incorporates an algorithm that covers the entire decay chain from the final state particle to the mother at once [104]. The chain parameterisation includes vertex positions, momentum parameters, and decay lengths. All parameters are adjusted simultaneously in the fit. In this way, correlations and uncertainties are correctly comprised. While 4-momentum conservation at each vertex are considered, further boundary conditions as mass constraints or forcing the reconstructed momentum direction towards the PV (PV constraint) can be added.

4.6 Flavour tagging

For measurements of decay-time-dependent *CP* violation, knowledge of the flavour state (B^0 or $\bar{B}^0$) at the time of production ($t = 0$) is essential. At LHCb this information is provided by the so-called flavour tagging that can be divided into two classes of algorithms. The algorithms that determine the production flavour by exploiting the charged particles, that are most likely remnants of the hadronisation process of the signal B candidate, are categorised as same-side (SS) taggers. The opposite-side (OS) taggers make use of the predominant production process of B mesons. A schematic representation of the various taggers is shown in Fig. 4.2. Each tagging algorithm delivers a tag $d \in \{-1, 0, +1\}$. While the value -1 corresponds to the production flavour of a $\bar{B}^0$, $+1$ indicates to that of a B^0 and 0 means that the candidate is untagged, i.e. no statement can be made about its flavour state at the time of production. Each tag is provided with a mistag probability $\eta \in [0, 0.5]$, which is an estimate of the probability of a tag being wrong. If the mistag probability is 0, a candidate has been perfectly tagged, while 0.5 is equivalent to an untagged candidate. These mistag probabilities are the output of multivariate classifiers that are trained on datasets of flavour-specific decays. A decay is flavour-specific when only the transitions $B^0 \to f$ and $\bar{B}^0 \to \overline{f}$, but not $B^0 \to \overline{f}$ and $\bar{B}^0 \to f$ are allowed. Since the training datasets differ from the signal datasets, e.g. in terms of trigger and selection criteria, the output of the algorithms must be calibrated. The calibration is performed on control samples that are very similar to the signal in terms of kinematic topology. On these control channels a calibration function is extracted to convert the mistag probability estimate η of the algorithms into a mistag probability $\omega(\eta)$ measured in the control sample. Based on the mistag probability $\omega(\eta)$, a tagging dilution $D = (1 - 2\omega)$ can be determined, which indicates how much the measured amplitude is reduced by the imperfect flavour tagging. In addition, a limited tagging efficiency ε_{tag} exists, which characterises the ratio of tagged candidates over all candidates. The product of the squared tagging dilution with the tagging efficiency indicates the tagging power that represents the effective statistical loss compared to a perfectly tagged sample. By summing over N signal candidates the tagging power can be determined on a per-candidate basis reading

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} D^2 = \frac{1}{N} \sum_{i=1}^{N} (1 - 2\omega(\eta_i))^2 \,. \tag{4.9}$$

The tagging power ε_{eff} is a good quantity to compare the flavour-tagging performance between different channels and experiments. In case of a perfect flavour tagging a tagging response exists for all candidates and a mistag probability of zero is assigned to all of them. As a tagging efficiency of $\varepsilon_{\text{eff}} = 100\%$ is not achievable in practise, $\varepsilon_{\text{eff}} N$ indicates the ratio of perfectly tagged candidates compared to all candidates. Then, the precision of the CP parameters is limited by

$$\sigma_{CP} \propto \frac{1}{\sqrt{\varepsilon_{\text{eff}} N}} \,. \tag{4.10}$$

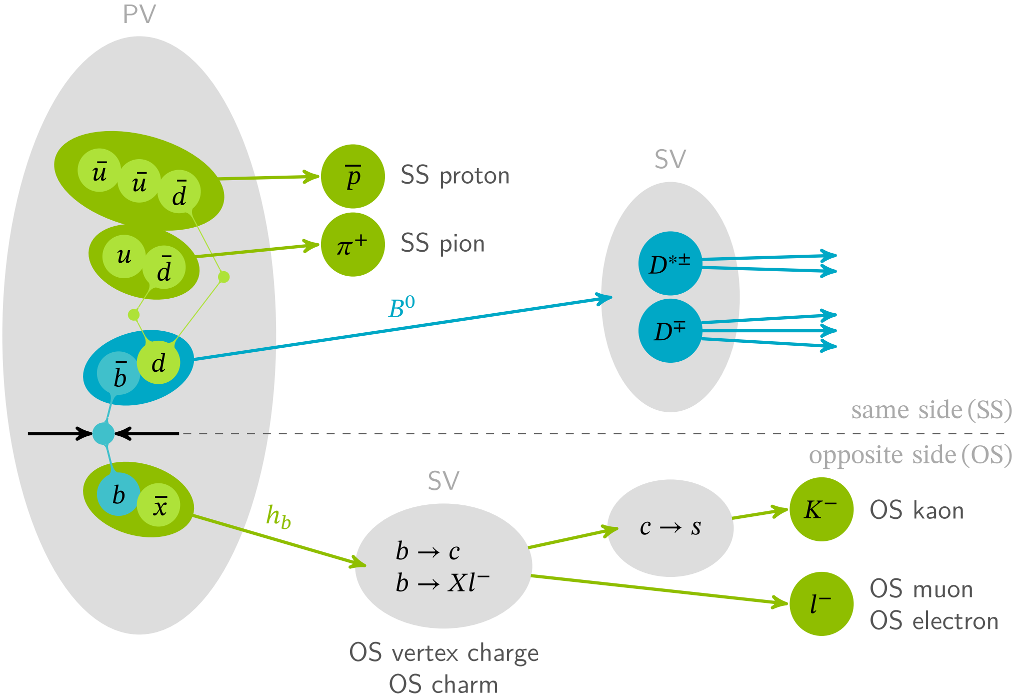

Figure 4.2 – Schematic overview of the OS and SS tagging algorithms used in the analysis of $B^0 \to D^{*\pm}D^{\mp}$ decays (illustration based on [105]).

4.6.1 Same-side flavour tagging

Same-side taggers make use of the hadronisation remnants of the signal B^0 meson. Similar to the $b\bar{b}$ quark production, d quarks are mostly created in $d\bar{d}$ pairs. When one of the d quarks becomes a constituent of the signal B^0 meson the second d quark usually combines with u quarks to form a pion or a proton. The charges of the pion and proton give access to the initial flavour of the signal B candidate. While protons and negative pions refer to $\bar{B}^0$ mesons, anti-protons and positively charged pions are associated with B^0 mesons. The two tagging algorithms that are applied in the presented analysis are the SS proton (SSp) and SS pion (SSπ) taggers. They are developed using BDTs for selecting the tagging particles [106]. The tagging efficiency of the SSp tagger is $\sim$39% and for the SSπ tagger $\sim$72%. In total $\sim$80% of all candidates are tagged by at least one of the SS taggers. While this efficiency is very high, it comes along with a high average mistag probability of around $\sim$45%. The values are taken from Ref. [106].

On the same-side, wrong tags mainly occur from falsely selecting particles from the underlying event as tagging particles. Such mis-identifications are especially challenging for the SSπ tagger, as pions are the most commonly produced type of particles in LHCb events.

4.6.2 Opposite-side flavour tagging

The OS tagging algorithms [107] are based on the pair production of b and $\overline{b}$ quarks and derive the flavour of the signal B meson by investigating the flavour of the second produced b hadron (tagging B). If the tagging B decays semileptonic, the charged leptons indicate the flavour of the tagging B. This charge is exploited by the OS muon (OSμ) and the OS electron (OSe) taggers. Due to the better reconstruction of muons compared to electrons at the LHCb, the tagging efficiency of the OSμ ($\sim 5\%$) is higher than of the OSe ($\sim 2\%$). While the tagging efficiency is relatively low, the mistag probability is low at around 30–34%. The efficiencies quoted here were determined on charmonium channels.

The OS kaon (OSK) tagger considers kaons produced in the $b \to c \to s$ decay chain. With this tagger a tagging efficiency of $\sim 11\%$ at a mistag probability of $\sim 40\%$ is achieved.

The OS vertex charge (OSvtx) tagger analyses the secondary vertex of the tagging B and determines its flavour by calculating the average charge corresponding to all tracks connected to this secondary vertex. At a tagging efficiency of $\sim 15\%$ this tagger attains a mistag probability similar to the OSK tagger.

The OS charm (OSCharm) tagger [108] reconstructs decays of charm hadrons that are created trough $b \to c$ transitions of the tagging B. Due to the challenge of the full reconstruction, only a rather low tagging efficiency of $\sim 3\%$ is achieved. However, the OS charm tagger reaches a low mistag probability of $\sim 35\%$. The values are taken from Ref. [107].

Wrong tags can occur when wrong particles that do not originate from the partner hadron, are inspected. In addition, it may happen that the partner b hadron has oscillated into the opposite flavour before decaying.

5 Datasets and Selection

The measurement of *CP* violation in the decay $B^0 \to D^{*\pm}D^{\mp}$ is performed on the full available dataset of the LHCb experiment. This dataset can be divided into Run 1 and Run 2. The dataset from the first data taking period (Run 1) comprises $3\,\mathrm{fb}^{-1}$ of pp collisions and was recorded at centre-of-mass energies of 7 TeV and 8 TeV. In Run 2, an integrated luminosity of $6\,\mathrm{fb}^{-1}$ was collected, taken at 13 TeV.

In the presented analysis, $B^0 \to D^{*\pm}D^{\mp}$ decays are reconstructed in a completely hadronic final state. Hence, the background level is significantly higher than for example the level in $B^0 \to J/\psi K^0_S$, where the characteristic two muon final state of the J/ψ allows to reduce the background severely. In order to improve the signal to background ratio in the $B^0 \to D^{*\pm}D^{\mp}$ analysis and to enable a precise measurement of the *CP* parameters, a selection chain containing multiple steps is developed. After explaining the reconstruction of $B^0 \to D^{*\pm}D^{\mp}$ decays in Sec. 5.1, the description of the selection chain starts with the stripping selection in Sec. 5.2. This global preselection is followed by a customised preselection explained in Sec. 5.3. After introducing physical backgrounds in Sec. 5.4, their suppression is explained in Sec. 5.5. Finally, a multivariate method is utilised in the selection in order to reduce combinatorial background (Sec. 5.6).

5.1 Reconstruction

The B^0 mesons are produced in the PV at the interaction point. After a mean flight distance of approximately 1 cm, they decay at the SV with a probability of $6 \cdot 10^{-2}\%$ into a $D^{*\pm}$ and $D^{\mp}$ [22]. The reconstruction of the $D^{\mp}$ meson is performed via its most dominant decay into a kaon and two pions. With respect to the $D^{\mp}$ meson, the kaon has the opposite charge ($K^{\pm}$), while the pions carry the same charge ($\pi^{\mp}\pi^{\mp}$). The branching ratio of the $D^{\mp} \to K^{\pm}\pi^{\mp}\pi^{\mp}$ decay is about 9%. The $D^{*\pm}$ is reconstructed as $D^{*\pm} \to \overset{(-)}{D}{}^0\pi^{\pm}$, which has a branching ratio of about 68%. Here, and in the latter of the thesis, charge conjugation is applied. In this analysis, two final states are considered for the D^0 decay: $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^-\pi^+\pi^+$. The branching ratio of the former decay (about 4%) is half as big as of the latter (about 8%). Nevertheless, the latter decay

$D^0 \to K^-\pi^-\pi^+\pi^+$ is harder to reconstruct due to tracking and selection inefficiencies. All quoted branching ratios are taken from [22].

Due to the fact, that the mass difference of the D^{*+} and D^0 mesons is only slightly larger than the mass of a pion, a low kinetic energy is provided to the pion. Hence, a so-called slow pion is produced. The slow pion is an important characteristic of D^{*+} decays. Another characteristic is that the D^0 meson does not decay promptly. Before decaying into a kaon and one or three pions at the tertiary vertex (TV), it flies a short distance downstream, which leads to a separation of the SV and TV. The slow pion as well as the separated vertices are features that can be utilised for an efficient selection.

This thesis covers the approach and strategy for the $D^0 \to K^-\pi^+$ final state. The second final state $D^0 \to K^-\pi^-\pi^+\pi^+$ was studied by other participants of the analysis group. Therefore, mainly the analysis of $B^0 \to D^{*\pm}D^{\mp}$ with the final state $D^0 \to K^-\pi^+$ is discussed in the following. However, results, such as the mass fit on data, are presented for both final states, as the candidates of the two final states are jointly used to measure CP violation in $B^0 \to D^{*\pm}D^{\mp}$.

5.2 Stripping selection

Events that pass the criteria of the `StrippingB02DstDBeauty2CharmLine` stripping line are used for the $D^0 \to K^-\pi^+$ final state. The data for the year 2012 (2011) is processed under Stripping S21(r1) conditions. For 2015, 2016, 2017 and 2018 the respective Stripping versions S24r1, S28r1, S29r2 and S34 are used. All requirements of the stripping are listed in Tab. 5.1.

Only events that are triggered by a topological trigger line or by the inclusive ϕ line are considered. The inclusive ϕ line is triggered by a $\phi \to K^+K^-$ decay. Thus, it is especially useful for the final state $D^{\mp} \to K^{\pm}K^{\mp}\pi^{\mp}$ that can occur via $D^{\mp} \to \phi\pi^{\mp}$, which is considered as background in this analysis. With a requirement on the number of long tracks in an event, the multiplicity is restricted. In addition, all pion and kaon candidates are required to be long tracks with a good track quality. Requiring high transverse momenta ensures that decays of heavy mesons are present. The quantity χ^2 is a measure for the quality of a performed fit, e.g. how accurate the position of the PV could be determined. Often, the χ^2/ndof is used as a measure of quality, where ndof denotes the number of degrees of freedom. The requirements on the vertex χ^2 and the shortest distance between two particles, which is called DOCA (distance of closest approach), ensure a correct vertex reconstruction. The parameter DIRA (direction angle) is the cosine of the angle between the momentum direction of the respective particle

Table 5.1 – Requirements of the stripping line `StrippingB02DstDBeauty2CharmLine`. The requirements are the same in all Stripping versions except for a few cases when values in parenthesis refer to the changes in Run 2. The variables are explained in the text.

Candidate	Variable	Requirement	Unit
	#longtracks	< 500	
	`Hlt2Topo` or `Hlt2IncPhi`	passed	
$\pi^\pm/K^\pm$	$\chi^2_{\text{track}}/\text{ndf}$	$< 3\,(< 4)$	
	p_{T}	> 100	MeV/c
	p	> 1000	MeV/c
	$\min(\chi^2_{\text{IP to any PV}})$	> 4	
	p_{ghost}	< 0.4	
$\pi^\pm$	$\text{DLL}_{K\pi}$	< 20	
$K^\pm$	$\text{DLL}_{K\pi}$	> -10	
$D^{*\pm}/D^\mp/D^0$	$\Sigma_{\text{daughters}}\, p_{\text{T}}$	> 1800	MeV/c
	$\lvert m_{D^*} - m_{D^*,\text{PDG}} \rvert$	$< 50\,(< 600)$	MeV/c^2
	$\lvert m_{D^0} - m_{D^0,\text{PDG}} \rvert$	< 100	MeV/c^2
	$m_{D^\pm}$	$1769.62 - 2068.49$	MeV/c^2
	$\lvert m_{D^*} - m_{D^0} \rvert$	$-\,(< 200)$	MeV/c^2
	DOCA	< 0.5	mm
	$\chi^2_{\text{vtx}}/\text{ndf}$	< 10	
	$\chi^2_{\text{distance to best PV}}$	> 36	
	$\text{DIRA}_{\text{best PV}}$	> 0	
one daughter hadron	p_{T}	> 500	MeV/c
	p	> 5000	MeV/c
	$\chi^2_{\text{track}}/\text{ndf}$	$< 2.5\,(< 4.0)$	
B^0	$\Sigma_{\text{daughters}}\, p_{\text{T}}$	> 5000	MeV/c
	m	$4750 - 6000$	MeV/c^2
	$\chi^2_{\text{vtx}}/\text{ndf}$	< 10	
	t	> 0.2	ps
	$\chi^2_{\text{IP to any PV}}$	< 25	
	$\text{DIRA}_{\text{best PV}}$	> 0.999	
one hadron in decay chain	p_{T}	> 1700	MeV/c
	p	$> 10\,000$	MeV/c
	$\chi^2_{\text{track}}/\text{ndf}$	$< 2.5\,(< 4.0)$	
	$\min(\chi^2_{\text{IP to any PV}})$	> 16	
	$\min(\text{IP}_{\text{any PV}})$	> 0.1	mm

and the vector between origin vertex and the vertex of the decay of the particle. To ensure, that the pions and kaons do not originate from the PV, requirements are applied on the so-called impact parameter (IP), which is the shortest distance between the reconstructed track and the PV. Furthermore, cuts are introduced to the IP significance $\chi^2_{\text{IP to any PV}}$ and the χ^2 of the distance to the best PV $\chi^2_{\text{distance to best PV}}$. In Fig. 5.1, some parameters are schematically illustrated.

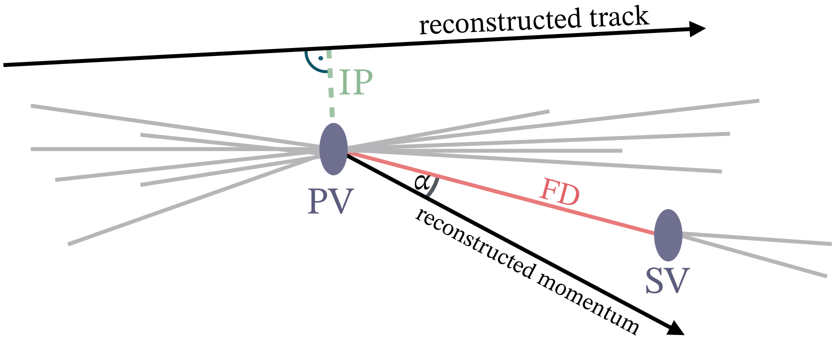

Figure 5.1 – Schematic illustration of parameters that are relevant for the selection. The PV and SV are the primary and secondary vertices, IP denotes the impact parameter, FD is the flight distance and DIRA refers to the direction angle α.

The parameter ω_{ghost} indicates the probability that a reconstructed track does not describe the trajectory of a real particle, but is instead just a so-called ghost, a random combination of detector hits. As explained in Sec. 3.3.1, DLL stands for Delta-Log-Likelihood and gives a measure for the probability that the information from the particle identification system favours a certain particle hypothesis over another. A cut is applied to the decay time t of the B^0 meson to suppress prompt background, where the reconstructed particles were produced directly in a pp collision and do not originate from the decay of a long-lived particle. The invariant mass of the B^0 ($D^{\pm}$) meson is required to lie within a mass window, which covers both the B^0 ($D^{\pm}$) and B^0_s ($D^{\pm}_s$) signal candidates as well as background regions that are needed for e.g. multivariate analyses.

Small differences in the requirements exist between the different Stripping versions. Some cuts are looser in Run 2 than in Run 1 and one new quantity has been introduced in Run 2: the mass difference between the $D^{*\pm}$ and D^0 mesons.

Though the stripping requirements for $B^0 \to D^{*\pm}D^{\mp}$ are not very strict, the selection efficiency $\epsilon_{\text{stripping}}$ is still limited. In addition, there is a reconstruction inefficiency due to the fact that the algorithms sometimes fail in reconstructing the B^0 candidates. The efficiency $\epsilon_{\text{stripping}}$ considers also inefficiencies in the detection of tracks. The combined efficiency for simulated signal candidates in the decay channel $B^0 \to D^{*\pm}D^{\mp}$ is given by $\epsilon_{\text{stripping,reco}} = (0.727 \pm 0.002)\%$ in Run 1 and $\epsilon_{\text{stripping,reco}} = (1.030 \pm 0.005)\%$ in Run 2. The higher centre-of-mass energy in Run 2 causes changes in different distributions, e.g. in p_{T}, so that significantly less candidates are rejected. The uncertainty in Run 1 is smaller due to a larger sample of simulated signal decays under Run 1 conditions.

5.3 Preselection

In this analysis, the observables for the reconstructed mass and reconstructed decay time of the B^0 meson are determined by a DecayTreeFit (see Sec. 4.5). All D meson masses are constrained to their nominal masses taken from the PDG [22]. Additionally, a PV constraint is introduced and it is required that these DecayTreeFits have converged. Background is reduced by introducing requirements on the D^0 and the $D^\pm$ masses as well as on the mass difference between $D^{*\pm}$ and D^0. The cut values are chosen to fulfil that at least 99% of the signal simulation is kept in the Run 1 sample of simulated $B^0 \to D^{*\pm}D^{\mp}$ decays. All requirements, which are also applied to the Run 2 data samples, are listed in Tab. 5.2. While a mass window of 50 MeV/c^2 around the $D^\pm$ mass is chosen, a mass window of 40 MeV/c^2 around the D^0 mass is selected. The third cut is applied on the mass difference between the $D^{*\pm}$ and D^0 mesons, requiring it to be less than 150 MeV/c^2. This cut has a high separating power between signal and background, which results from the fact that mass results effects are reduced by subtracting the $D^{*\pm}$ and D^0 masses. Additionally, a large number of false reconstructed candidates are discarded, since the mass difference between the $D^{*\pm}$ and the D^0 peaks at ~145 MeV/c^2. Candidates with higher mass differences, have a larger probability to be mis-reconstructed. This cut suppresses about 75% of the combinatorial background. While all preselection cuts have a background suppression of more than 90%, a signal efficiency on simulated $B^0 \to D^{*\pm}D^{\mp}$ decays of more than 95% is retained. The background rejection is determined on the upper mass side band of the data sample. The

Table 5.2 – Cuts applied in the preselection. The nominal masses used for these cuts are taken from Ref. [39].

Variable	Requirement	Unit
Difference between $D^{*\pm}$ and D^0 masses	$m_{D^{*\pm}} - m_{D^0} < 150$	MeV/c^2
D^0 mass	$\lvert m_{D^0} - m_{D^0}^{\text{PDG}} \rvert < 40$	MeV/c^2
$D^\pm$ mass	$\lvert m_{D^\pm} - m_{D^\pm}^{\text{PDG}} \rvert < 50$	MeV/c^2

distributions of the variables that are used in the preselection are presented in Fig. 5.2. In Fig. 5.3, the reconstructed $D^{*\pm}D^{\mp}$ mass distributions after the preselection for Run 1 and Run 2 are shown. Although most candidates are background contributions, a clear peak at the nominal B^0 mass is visible. The background contributions include combinatorial and physical backgrounds. The different types of background and corresponding suppressions are described in the next sections.

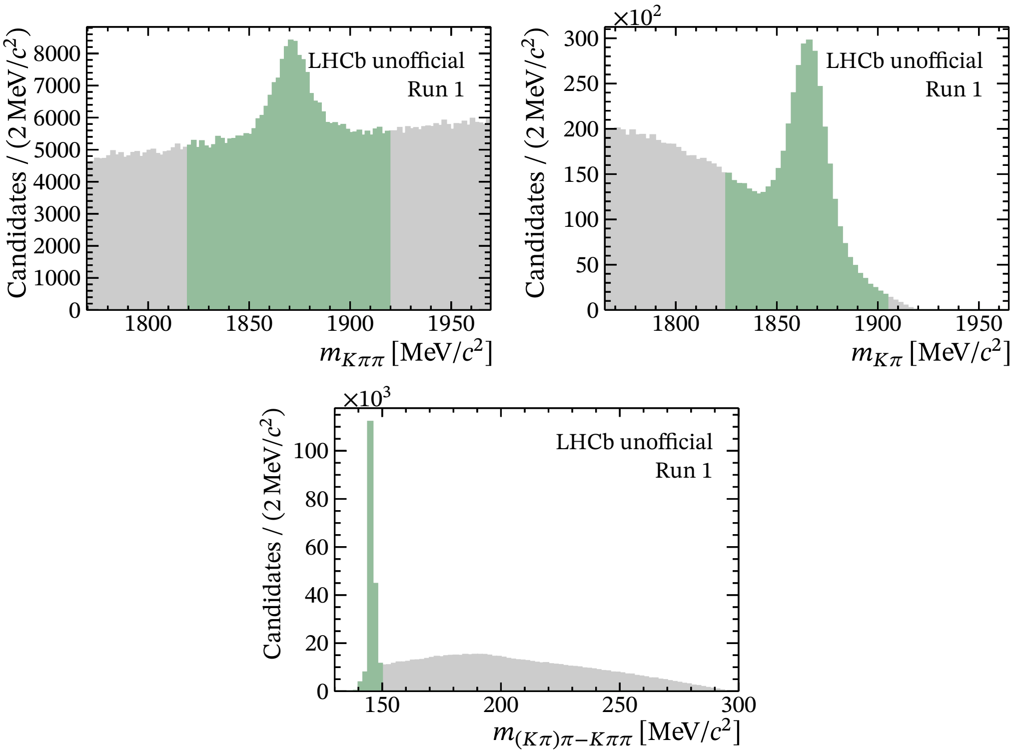

Figure 5.2 – Distributions of variables used in the preselection. Requirements are applied on the invariant D^0 and $D^\pm$ masses as well as on the mass difference between the invariant masses of $D^{*\pm}$ and D^0. The green area indicates the candidates that are selected, the candidates located in the grey areas are discarded. The plots show the distributions of the Run 1 dataset.

5.4 Types of backgrounds

In contrast to leptonic colliders, the track multiplicity in hadronic colliders is much higher, which results in a greater background level. At the LHCb experiment, the background is dominated by combinatorial background, where random particle tracks that do not originate from a common physical decay are combined. Combinatorial background appears as a flat, exponentially decreasing distribution in the invariant

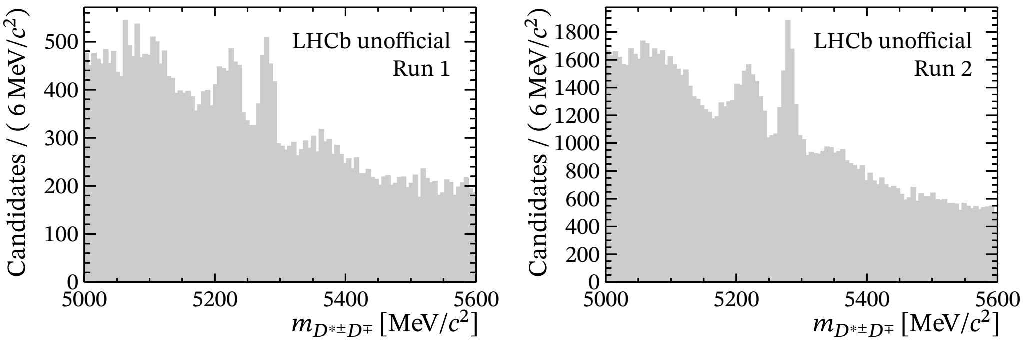

Figure 5.3 – Distributions of the invariant $D^{*\pm}D^{\mp}$ mass after applying the preselection for (left) Run 1 and (right) Run 2. A clear signal peak is visible at the nominal B^0 mass of 5280 MeV/c^2. Physical background is present towards smaller masses. Combinatorial background represents the dominant background contribution.

$D^{*\pm}D^{\mp}$ mass distribution. In addition, there are physical backgrounds, which appear as peaking structures in the $D^{*\pm}D^{\mp}$ mass distribution.

Different kinds of physical background exist: exclusive background decays mimic the signal due to mis-identification or mis-reconstruction. Additionally, there are partially reconstructed backgrounds. Here, single particles of the physical background process are not reconstructed, so that the remaining final state particles can be formed to a signal candidate wrongly. Finally, there are non-resonant background decays, which do not occur via a resonance like a $D^{*\pm}$ or $D^{\mp}$ meson as in the signal decay $B^0 \to D^{*\pm}D^{\mp}$, but whose final state particles match the ones from signal decays. These peaking backgrounds should be suppressed as efficiently as possible. If such contributions are still partially left in the invariant $D^{*\pm}D^{\mp}$ mass distribution, they must be described by PDFs, especially if they lie within the B^0 signal-mass region.

5.5 Vetoes of physical background

The presence of physical background contributions due to mis-identification can be studied by calculating the invariant mass of the respective contribution after applying different mass hypotheses to the presumable mis-identified particles. The background gets visible as a peak at the nominal mass of the background mother particle. These peaks can be reduced by cuts on variables that give information about the particle type. Here, the ProbNN variables are used, which provide an a posteriori probability of the particle to be a proton, muon, electron, kaon, or pion. These quantities are provided by neural networks [109]. A neural network is a multivariate analysis method, which is used to find patterns in datasets to allow for the identification of particles in this analysis. Due to the fact that ProbNN variables show deviations between simulated and recorded data, a variable transformation is performed to achieve a better MC-data agreement. This procedure uses a kernel density estimation technique in four dimensions [110], which are the ProbNN, the transverse momentum, the pseudorapidity and the number of tracks in the event, keeping the correlations between the ProbNN variables to other variables. These observables are chosen, as they make it possible to categorise the ProbNN observables in the dimensions that describe the kinematic properties of the particles. The MC-data agreement can be even more improved by using the ratio of the probability describing that a particle is a pion over the sum of the probabilities that it is a pion and one other particle type x (e.g. kaon, proton)

$$\mathrm{PIDratio}_x = \frac{\mathrm{ProbNN}\pi}{\mathrm{ProbNN}\pi + \mathrm{ProbNN}x}\,. \tag{5.1}$$

A good MC-data agreement is especially conducive in multivariate analyses as discussed in Sec. 5.6. In case of the vetoes selection, the corrected MC enables a more accurate signal efficiency estimation.

As already mentioned in the previous section, after the cut on the mass difference of the D^{*+} and D^0 mesons is applied, a lot of background caused by mis-identification of D^{*+} meson decays is reduced. Hence, no physical background occurs in the invariant mass distribution of the D^{*+} meson. In contrast to that physical background contributions can be found in the D^- mass distribution. These arise due to decays of Λ_c^+ baryons as well as of D_s^+ and ϕ mesons.

$D_s^\pm$ veto A background contribution from $D_s^+ \to K^-K^+\pi^+$ decays can occur if a kaon is mis-identified as a pion. In reconstructed $B^0 \to D^{*\pm}D^\mp$ decays, this background component is identified by assigning the kaon mass hypothesis to the pion of $D^+ \to K^-\pi^+\pi^+$ candidates with the higher transverse momentum. The resulting invariant mass distribution of the $K^-K^+\pi^+$ combination is shown in the left plot of Fig. 5.4. A clear peak is visible at the nominal D_s^+ mass of $m_{D_s^+} = 1968.34\,\mathrm{MeV}/c^2$ [22]. A candidate is rejected if the invariant $K^-K^+\pi^+$ mass is compatible with $m_{D_s^+}$ within $\pm 25\,\mathrm{MeV}/c^2$ and if the probability of the mis-identified pion to be a kaon (ProbNNK) is higher than the probability to be a pion (ProbNNπ). As can be seen in the left plot of Fig. 5.4, this veto is quite loose and a resonant structure is still present after applying this veto. However, it is not possible to suppress this background completely without a large signal loss. Thus, this loose veto is used to allow for an adequate level of the $B^0 \to D^{*-}D_s^+$ component to describe it in the invariant $D^{*\pm}D^{\mp}$ mass distribution. The invariant $D^{*\pm}D^{\mp}$ mass distribution before and after applying the veto is shown in the right plot of Fig. 5.4, the rejected candidates are overlaid additionally.

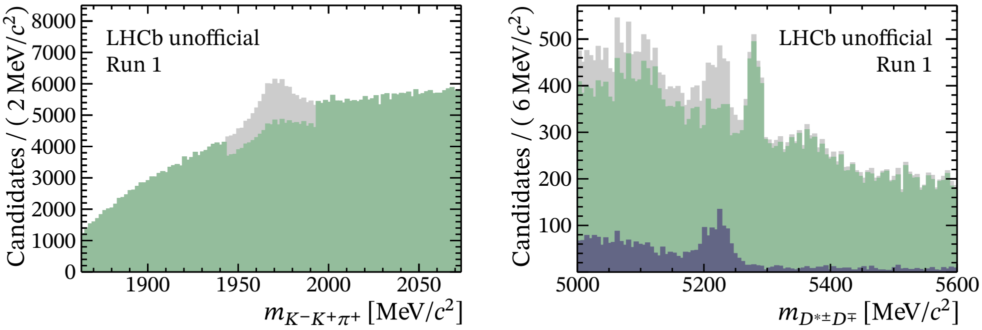

Figure 5.4 – Invariant (left) $K^-K^+\pi^+$ and (right) $D^{*\pm}D^{\mp}$ mass distributions. In both figures the distributions are shown (grey) before and (green) after applying the D_s^+ veto. In the right plot, the dark blue distribution illustrates the rejected candidates.

ϕ veto The background decay $D_s^+ \to K^-K^+\pi^+$ predominantly proceeds via the ϕ resonance through $D_s^+ \to \phi\pi^+$. By again assigning the kaon mass to the pion with higher p_T, a clear resonance at the ϕ mass of $m_\phi = 1019.461\,\mathrm{MeV}/c^2$ [22] becomes visible in the invariant K^+K^- mass, as can be seen in the left plot of Fig. 5.5.

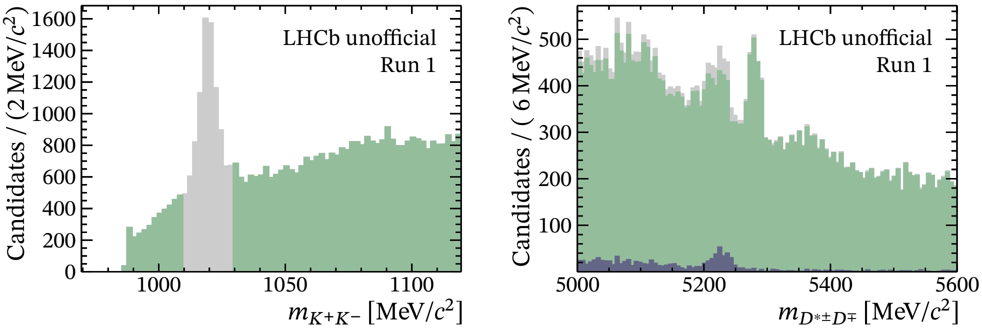

Figure 5.5 – Invariant (left) K^+K^- and (right) $D^{*\pm}D^{\mp}$ mass distributions. In both figures the distributions are shown (grey) before and (green) after applying the ϕ veto. In the right plot, the dark blue distribution illustrates the rejected candidates.

This background contribution is suppressed by rejecting all candidates for which the invariant mass of the kaon pair lies within a $10\,\mathrm{MeV}/c^2$ window around m_ϕ. No resonant structures are visible at the nominal D_s^+ or ϕ masses, when the kaon mass hypothesis is given to the pion with the lower p_T. Hence, no further $D_s^\pm$ or ϕ vetoes are applied.

Furthermore, checks on simulated data show that the probability that the pion with the lower p_{T} is mis-identified is lower than 5%. After applying the $D_s^{\pm}$ and ϕ vetoes, 75% of these physical backgrounds are rejected in Run 1 and 82% in Run 2, while 96% and 94% of the signal is preserved in Run 1 and Run 2, respectively. The background rejections are calculated on datasets of simulated $B^0 \to D^{*-}D_s^+$ decays that are reconstructed as $B^0 \to D^{*\pm}D^{\mp}$, while the signal efficiency is determined on simulated signal decays. The invariant $D^{*\pm}D^{\mp}$ mass distribution before and after applying the veto is shown in the right plot of Fig. 5.5, the rejected candidates are overlaid additionally.

Λ_c^+ veto The mis-identification of a proton as a pion leads to the background contribution $\Lambda_c^+ \to K^- p\pi^+$. If a proton is mis-identified as a pion, the background contribution from decays can be present. By assigning the proton mass hypothesis to the pion of $D^+ \to K^-\pi^+\pi^+$ candidates with the larger transverse momentum, this background component is identified as a resonant structure in the invariant mass distribution of the $K^-p\pi^+$ combination at the nominal Λ_c^+ mass of $m_{\Lambda_c^+} = 2286.46\,\mathrm{MeV}/c^2$ [22] (see left plot in Fig. 5.6). A candidate is rejected, if the invariant $K^-p\pi^+$ mass is compatible with $m_{\Lambda_c^+}$ within $\pm 25\,\mathrm{MeV}/c^2$ and if the $\mathrm{PIDratio}_p$ is lower than 0.3. When proceeding with the pion with lower p_{T}, no peak at $m_{\Lambda_c^+}$ in the invariant $K^-p\pi^+$ mass distribution is visible. After applying the Λ_c^+ veto, 95% of this physical background is rejected in Run 1 and 96% in Run 2, while 97% of the signal is preserved in both runs. The background rejections are calculated on datasets of simulated $\Lambda_b^0 \to \Lambda_c^+ D^{*-}$ decays, while the signal efficiency is determined on simulated signal decays. The invariant $D^{*\pm}D^{\mp}$ mass distribution before and after applying the veto is shown in the right plot of Fig. 5.6, the rejected candidates are overlaid additionally.

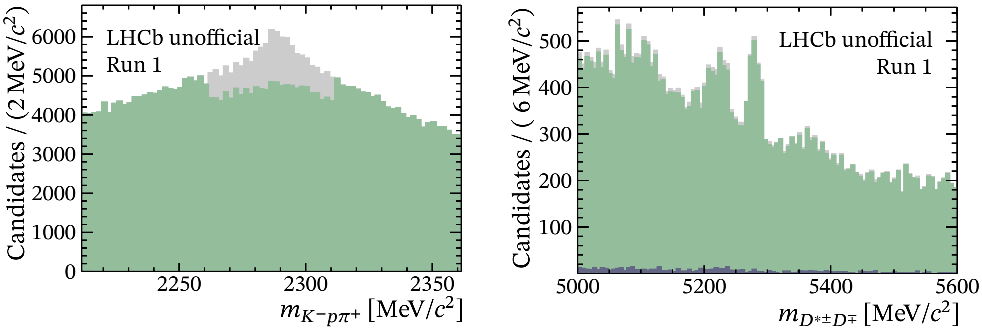

Figure 5.6 – Invariant (left) $K^-p\pi^+$ and (right) $D^{*\pm}D^{\mp}$ mass distributions. In both figures the distributions are shown (grey) before and (green) after applying the Λ_c^+ veto. In the right plot, the dark blue distribution illustrates the rejected candidates.

Single-charmless veto A contamination from $B^0 \to D^{*\pm}\pi^{\pm}\pi^{\mp}\pi^{\mp}$ decays, where the three pions are not produced via a $D^{\pm}$ resonance, but stem directly from the B^0 decay and where a kaon–pion mis-identification appears, is considered and suppressed by the following veto: Candidates are rejected if the invariant mass of the $D^{*+}\pi^+\pi^-\pi^-$ combination is compatible with the B^0 mass within $\pm 40\,\text{MeV}/c^2$ and if the PIDratio_K is lower than 0.3 or if the χ^2 of the flight distance of the $D^{\pm}$ with respect to the B^0 decay vertex is lower than four. With the latter requirement a separation between the B^0 and $D^{\pm}$ decay vertices is ensured and, thus, that the three pions are not created in the SV. The invariant $D^{*+}\pi^+\pi^-\pi^-$ and $D^{*\pm}D^{\mp}$ mass distributions before and after applying this veto are shown in Fig. 5.7.

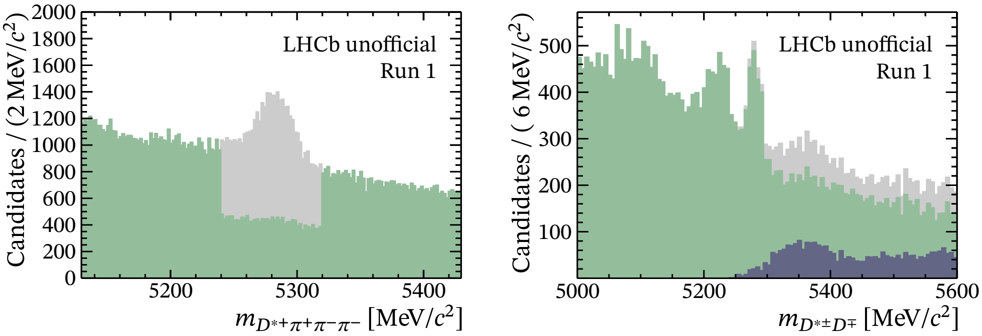

Figure 5.7 – Invariant (left) $D^{*+}\pi^+\pi^-\pi^-$ and (right) $D^{*\pm}D^{\mp}$ mass distributions. In both figures the distributions are shown (grey) before and (green) after applying the single charmless veto. In the right plot, the dark blue distribution illustrates the rejected candidates.

5.6 Multivariate selection

After suppressing physical backgrounds with vetoes, a multivariate selection with a boosted decision tree is utilised to reduce the level of combinatorial background. At first, the training of the BDT is described in Sec. 5.6.1. Afterwards, the optimisation of the BDT selection is presented in Sec. 5.6.2.

5.6.1 Training of the boosted decision tree

In this analysis, separated BDTs are trained for the two run periods, but with the same set of configurations and under the same conditions. The BDT training is based on the framework of TMVA [111] and utilises the AdaBoost algorithm. For the signal proxy simulated $B^0 \to D^{*\pm}D^{\mp}$ decays are used, while the background sample is given by reconstructed B^0 candidates with invariant masses outside the signal region. Since physical backgrounds are still present towards lower masses of the nominal B^0 mass, only the upper mass sideband ($m_{B^0} > 5400\,\mathrm{MeV}/c^2$) is used for the BDT training. In order to apply the technique of a 2-fold cross-validation, the dataset is randomly divided into two equally sized parts. Two separate BDTs are trained, one on each half, but with exactly the same training options and then applied on the dataset that was not used for the training. This method is employed, to avoid a possible overtraining in the upper mass sideband, since the region $5400\,\mathrm{MeV}/c^2 < m_{B^0} < 5600\,\mathrm{MeV}/c^2$ is also used for the mass fit (see Sec. 6.2).

In order to train the BDT, a set of input features is needed. Therefore, many variables are selected that show differences in their signal and background distributions. Using this large set of features, a first BDT is trained. After the BDT training, a list of the variables is provided, sorted according to their importance. For the final selection of features, all observables showing a low importance are discarded. In Tab. 5.3, the 26 features used for the final training are listed. Since four BDTs are trained in total, two for each run, and fluctuations always occur in the ranking of the variables, the order in the table represents an approximate ranking.

Table 5.3 – List of input variables used in the BDT training. The order of the input variables, left then right column, reflects their importance in the training.

BDT Variables	
$D^\pm$ decay-time significance	$D^{*\pm}$ p_T (PV constraint)
$\mathrm{PIDratio}_K$ from $K\,(D^\pm)$	$\cos(\angle(D^0,\pi))$
$\log(B^0$ IP χ^2 wrt own PV)	B^0 p_T (PV constraint)
$\log(D^\pm$ FD χ^2 wrt own PV)	$\cos(\angle(B^0,D^{*\pm}))$
π_2 from $D^\pm$ p_T	$\log(D^\pm$ IP χ^2 wrt own PV)
$\log(D^0$ FD χ^2 wrt own PV)	$D^\pm$ p_T (PV constraint)
$m(D^{*\pm}) - m(D^0)$	$\cos(\angle(D^\pm,\pi_2))$
K from $D^\pm$ p_T (PV constraint)	$\mathrm{PIDratio}_K$ from $\pi_2(D^\pm)$
$\mathrm{PIDratio}_K$ from $K\,(D^0)$	$\mathrm{PIDratio}_K$ from $\pi_1(D^\pm)$
$\log(\mathrm{DTF}\ \chi^2)$	$\mathrm{PIDratio}_K$ from π_slow
D^0 decay-time significance	$\log(D^0$ IP χ^2 wrt own PV)
$\cos(\angle(D^\pm,\pi_2))$	π_1 from $D^\pm$ p_T
$\cos(\angle(D^\pm,K))$	$\mathrm{PIDratio}_K$ from $\pi(D^0)$

The set of features comprises kinematic variables like the $D^{\pm}$ and D^0 decay-time significance t/σ_t and transverse momenta p_{T} as well as vertex qualities like the flight distance quality χ^2_{FD} and the impact parameter significance χ^2_{IP}. Furthermore, the $\mathrm{PIDratio}_K$ of corrected ProbNN variables (see Sec. 5.5) for all final state particles, the decay-tree-fit godness, the mass difference of the $D^{*\pm}$ and the D^0, and angles between reconstructed tracks are included. In order to optimise the hyperparameters, the so-called ROC (receiver operating characteristic) curve is utilised. The ROC curve represents the background rejection as a function of the signal selection efficiency. The larger the area under the curve, the better the BDT performance. The BDT is optimised by changing the hyperparameters iteratively until the area under the curve does not increase in size. The hyperparameters described in the following are used in the final BDT training. The BDTs consist of 800 trees, with the depth of the trees being limited to three. At each node of the tree 3% of the training dataset must be remaining. In order to find the best cut point for each variable, a scan is performed at 40 points. For the boosting a factor of $\beta = 0.1$ is used (see Eq. (4.7)). To ensure a prediction that is more robust, the bagging method is employed. Bagging is used to reduce the variance of BDTs. For every tree this procedure creates subsets with a size of, here, 80% compared to the provided datasets. In this so-called resampling it is allowed that events are picked several times. In addition, for every tree 10 random features are used. In Fig. 5.8, the BDT classifier distributions of the training and test samples are shown for one fold of each Run 1 and Run 2 training. The training and test samples are in a good agreement, thus, it can be assumed that no overtraining is present.

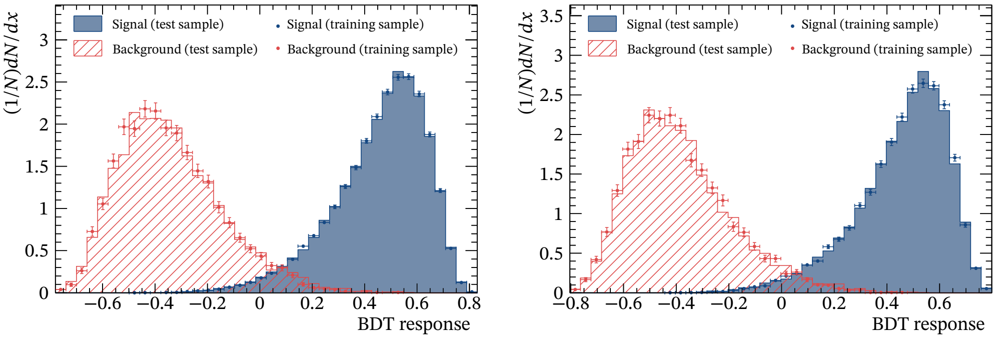

Figure 5.8 – Comparison between the BDT responses on the (data points) training and the (filled/shaded area) test samples of the (blue) signal and (red) background components for (left) Run 1 and (right) Run 2.

Fig. 5.9 shows the comparison between the classifier distributions of the BDTs outputs from the two folds. In both runs, the two distributions are compatible with each other, hence, the BDT classifiers from the two folds can be combined and are treated as one in the following.

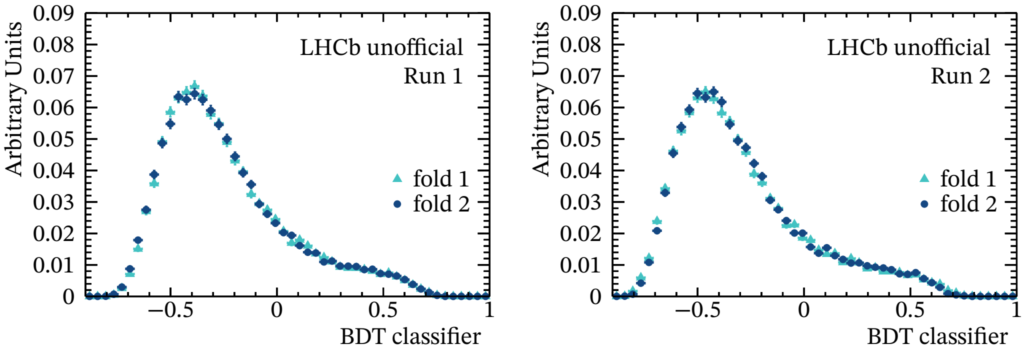

Figure 5.9 – Classifier distributions on data of the two trained BDTs, when applied to the candidates they were not trained on, for (left) Run 1 and (right) Run 2.

5.6.2 Classifier cut optimisation

After training the BDT, the BDT response is added to the data samples. Then, the selection itself is performed by applying a cut on the BDT response. The goal is to find the optimal cut point in order to maximise the sensitivity on the *CP* parameters in the decay-time-dependent and flavour-tagged measurement of $B^0 \to D^{*\pm}D^{\mp}$ decays. Hence, a measure called figure of merit (FOM) is needed, which reflects the sensitivity of the analysis. Then, this FOM is maximised with respect to the requirement on the BDT response. In this analysis, a FOM is used that takes several properties of the dataset into account that have an impact on the statistical power of the sample in a *CP* measurement.

In addition to a high effective signal size, a good performance of the flavour tagging is important, as especially small mistag probabilities ω_i lead to a more precise measurement. Moreover, the decay time resolution is considered in the FOM. It is expected to not have a great impact on *CP* measurements of B^0 decays, different to *CP* measurements in decays of the more rapidly oscillating B_s^0 mesons. However, it is taken into account in favour of a very generally usable FOM. Furthermore, the sensitivity of the decay-time-dependent asymmetry itself is also influenced by the reconstructed decay time t' of the B^0 candidates. The distributions of these properties may differ depending on the applied BDT cut point, which needs to be taken into account in order to find the optimal selection point that enables the best sensitivity on the *CP* parameters.

The FOM represents the inverse variance of the respective parameter to be measured and is obtained by creating the square of the differentiation of the log-likelihood with respect to this parameter summing up over all candidates. This FOM has been used firstly in Ref. [112] and is extended in Ref. [113] to use sWeights (see Sec. 4.2). The latter extension needs to be applied in the presented analysis, as the *CP* measurement is performed in a background subtracted way. The determination is done in terms of a maximum likelihood fit to the $D^{*\pm}D^{\mp}$ invariant mass distribution. Details on the procedure are given in Ch. 6. In this analysis, $\sin(\phi_d^{\mathrm{eff}})$ (compare Eq. (2.62)) is chosen as a single reference parameter, as it combines the parameters S_{D^*D} and C_{D^*D}. The FOM results in

$$\mathrm{FOM} \equiv \frac{\left(\sum_i s_{w_i}\right)^2}{\sum_i s_{w_i}^2}\mathcal{D}\,, \tag{5.2}$$

where s_{w_i} denotes the sWeights and $\mathcal{D}$ incorporates a dilution term that is defined as

$$\mathcal{D} = \frac{1}{\sum_i s_{w_i}}\sum_i (1-2\omega_i)^2 \mathrm{e}^{-(\Delta m_d \sigma(t_i'))^2} \cdot \mathcal{C}_i \cdot s_{w_i}\,. \tag{5.3}$$

It is averaged over all candidates taking the decay-time error estimates $\sigma(t_i')$ and the mistag probabilities ω_i into account. The parameter Δm_d represents the mass difference in the $B^0-\bar{B}^0$ system (compare Eq. (2.17)). The term $\mathcal{C}_i$ incorporates the parameter of interest $\sin(\phi_d^{\mathrm{eff}})$, and is given by

$$\mathcal{C}_i = \left[\frac{2d_i|\lambda|s}{1+|\lambda|^2 + d_i\mathcal{D}_{\mathrm{FT}}\mathrm{e}^{-(\Delta m_d\sigma(t_i'))^2/2}\left(-2|\lambda|s\sin(\phi_d^{\mathrm{eff}}) - (1-|\lambda|^2)c\right)}\right]^2, \tag{5.4}$$

where d_i denotes the decision for the production flavour of the B^0 mesons. The abbreviations $s = \sin(\Delta m_d t_i')$, $c = \cos(\Delta m_d t_i')$ and $\mathcal{D}_{\mathrm{FT}} = (1-2\omega_i)$ are introduced, which lead to a more concise definition of the term. In the BDT cut optimisation, the value of $\sin(\phi_d^{\mathrm{eff}}) = 0.54$ is used from the analysis of *CP* violation in $B^0 \to D^+D^-$ decays [114], neglecting the strong phase difference δ. The parameter λ (see Eq. (2.23)) is set to 0.75. As $|q/p = 1|$ is a good approximation for B^0 mesons and the decay amplitudes into the two final states $D^{*+}D^-$ and $D^{*-}D^+$ are of comparable size, $\lambda = 1$ would be a more intuitive choice. However, cross checks show that both choices of $\lambda = 1$ and using the world average of $\sin(\phi_d^{\mathrm{eff}}) = 0.54$ [50] result in very similar maxima in the BDT cut scan. The BDT cut optimisation aims to find the optimal cut point with regard to the FOM and, hence, the best sensitivity on the *CP* Parameters possible with the data. It is Important to note that the central values of the *CP* parameters can not biased by the cut scan, as only the sensitivities are reflected in the FOM. In order to investigate the effects of the individual terms, the FOM is divided into the following components. The effective flavour-tagging efficiency is given by

$$\mathrm{FOM}_{\epsilon D^2} \equiv \frac{1}{\sum_i s_{w_i}} \sum_i (1 - 2\omega_i)^2 \cdot s_{w_i}\,. \tag{5.5}$$

Considering the effect of decay-time-resolution results in the expression

$$\mathrm{FOM}_{\sigma_t} \equiv \frac{1}{\sum_i s_{w_i}} \sum_i \mathrm{e}^{-(\Delta m_d \sigma(t_i'))^2} \cdot s_{w_i}\,. \tag{5.6}$$

The effective signal size is given by

$$\mathrm{FOM}_{\mathrm{sig}_{\mathrm{eff}}} \equiv \frac{(\sum_i s_{w_i})^2}{\sum_i s_{w_i}^2}\,. \tag{5.7}$$

Assuming a perfect flavour tagging, i.e. no mistags, no limited tagging efficiency, and a vanishing decay-time resolution, the FOM from Eq. (5.2) gives

$$\mathrm{FOM}_{B^0} \equiv \frac{1}{\sum_i s_{w_i}} \sum_i \mathcal{C}_i \cdot s_{w_i}\,. \tag{5.8}$$

In order to find the maximal FOM value with respect to the cut on the BDT classifier, the BDT response is scanned on recorded data, separately for each run. At each cut point, sWeights are calculated by a maximum likelihood fit to the $D^{*\pm}D^{\mp}$ invariant mass distribution. The flavour-tagging parameters are calibrated as explained in Sec. 7.2. Then, the FOM is calculated including its partial contributions (see Eq. (5.5) to Eq. (5.8)). Figure 5.10 shows the total FOM along with its components determined in the BDT cut scans for Run 1 and Run 2. The optimal cut point is found at a BDT classifier value of 0.18 for both runs. Compared to the maximum of the effective signal size component, the cut point of the total FOM is shifted towards harder cut values. This is caused by the flavour-tagging component and also by the term FOM_{B^0}, which is depending on $\sin(\phi_d^{\mathrm{eff}})$. As the oscillation period given by $2\pi/\Delta m_d$ is large compared to the decay time of B^0 mesons, the term containing the decay-time resolution is constant on the full scan range. Thus, it has no impact on the optimal cut value. A signal efficiency of around 92% at a background rejection of more than 98% is achieved with this optimised BDT requirement in Run 1 and Run 2.

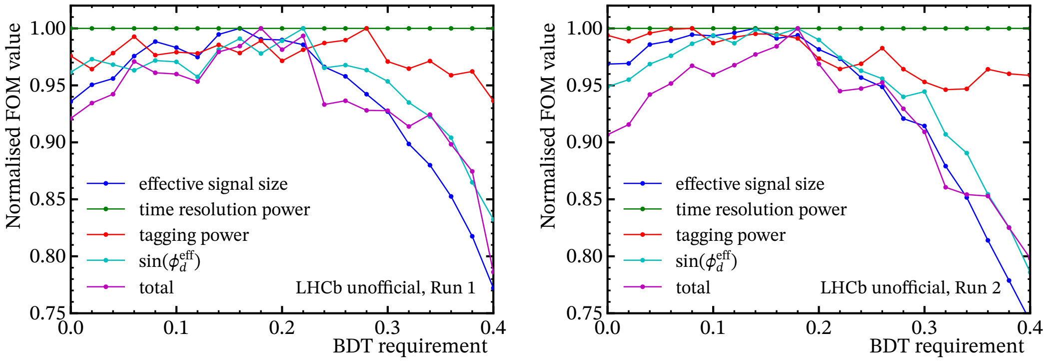

Figure 5.10 – The FOM on recorded data for different cuts on the BDT response. The full FOM is presented in violet. The other graphs show the single components that are defined in the text. All graphs are normalised to the highest value. The left plot shows Run 1, while the right plot illustrates Run 2.

5.7 Final selection

In the last selection step, all B^0 candidates are retained that have invariant masses and decay times within the ranges of $5000\,\text{MeV}/c^2$–$5600\,\text{MeV}/c^2$ and 0.3 ps–10.3 ps, respectively. These cuts show a signal efficiency of almost 100%. The lower boundary of the decay time is set to 0.3 ps. In the Stripping decay times are required to be greater than 0.2 ps, but it uses a decay-time observable created by a different algorithm (see Sec. 4.5) than the one used in the decay-time fit to measure the *CP* parameters. Hence, to ensure a well defined lower boundary, this requirement on the decay time is tightened. Consequently, in order to have a covered range of exactly 10 ps, the upper boundary is put to 10.3 ps. While the B^0 decay time is determined using the DTF with only the PV constraint, the invariant B^0 mass is calculated using the DTF with constraints on the PV and the three charmed mesons ($D^{*\pm}$, D^0 and $D^{\mp}$).

There is a chance that in an event more than one reconstructed signal candidate fulfils all selection criteria. The fraction of events with such multiple candidates is about 2% and about 4% for Run 1 and Run 2, respectively. Due to a higher track multiplicity in Run 2, the fraction of multiple candidates is higher in Run 2. Since the branching fraction of the $B^0 \to D^{*\pm}D^{\mp}$ decay is very small, only one signal candidate per event is expected. All multiple candidates are excluded randomly, as all candidates should be equally valid after surviving all selection steps. At the end of the selection, 3696 and 11 672 candidates remain in the datasets of Run 1 and Run 2, respectively. The next step is to fit the invariant $D^{*\pm}D^{\mp}$ mass distribution in order to describe and differentiate the remaining background components from the signal.

6 Modelling of the invariant Mass Distribution in the Signal Region

The goal of this thesis is to determine the *CP* parameters of the decay $B^0 \to D^{*\pm}D^{\mp}$. This measurement is performed in terms of a fit to the decay-time distribution of $B^0 \to D^{*\pm}D^{\mp}$ signal candidates. However, background contributions are still present in the data after the full selection. This becomes apparent, when looking at the invariant mass distribution. As can be seen in Fig. 6.1, the potential signal peak is already very significant, but the datasets still contain combinatorial and physical background candidates. It is essential that the remaining background is taken into account in the analysis. The signal distribution of the decay time can be extracted using the *sPlot* method described in section 4.2 by performing a maximum likelihood fit to the invariant $D^{*\pm}D^{\mp}$ mass, since the linear Pearson correlation between the B^0 decay time and the invariant $D^{*\pm}D^{\mp}$ mass is small, $\rho = 0.004$ in Run 1 and 0.007 in Run 2, respectively. In order to apply this method, a good description of the invariant $D^{*\pm}D^{\mp}$ mass distribution must be provided. The total PDF $\mathcal{P}$ comprises five components, the $B^0 \to D^{*\pm}D^{\mp}$ signal, the peaking backgrounds $B_s^0 \to D^{*\pm}D^{\mp}$, $B_{(s)}^0 \to D^{*+}D^{*-}$, $B^0 \to D^{*-}D_s^+$ and the combinatorial background

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

Here, N_i denotes the number of candidates of component i. The mass parameterisations for the components are described in the next section.

6.1 Extraction of parameterisations from simulated events

The mass parameterisations for the individual components, with the exception of the combinatorial background, are developed on simulated events. The parameters of the individual PDFs are determined in a maximum likelihood fit to the invariant $D^{*\pm}D^{\mp}$ mass distribution of simulated decays reconstructed as $B^0 \to D^{*\pm}D^{\mp}$. These fits are

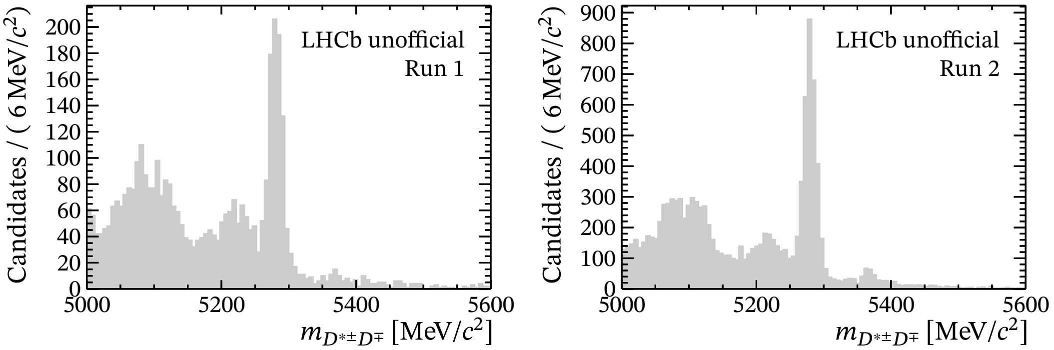

Figure 6.1 – Invariant $D^{*\pm}D^{\mp}$ mass distribution after all selection steps for (left) Run 1 and (right) Run 2. At the nominal B^0 mass at 5279 MeV/c^2 a clear signal peak can be seen. Physical backgrounds are still visible, especially for smaller invariant masses.

performed on datasets considering Run 1 and Run 2 separately after the full selection is applied.

6.1.1 Signal candidates

For the signal component, the reconstructed B^0 mass is modelled with the sum of three Crystal Ball functions [115–117]. A Crystal Ball function consists of a Gaussian core that converts into a power law with exponent n below the threshold $(m-\mu)/\sigma \leqslant -\alpha$, where σ and μ are the width and the mean of the Gaussian function. The sign of α defines whether the power law describes the distribution towards lower ($\alpha > 0$) or higher ($\alpha < 0$) masses.

In order to describe the B^0 signal, three Crystal Ball functions are combined, by setting the means μ_i equal to a common parameter μ. Two α parameters are negative, while one is positive. One of the Crystal Ball functions with a negative α parameter is responsible for the long tail towards lower masses, which is caused by candidates where photons are lost in the reconstruction. The other two Crystal Ball functions describe the main core of the distribution with a tail to each side. In Fig. 6.2, the plot of the invariant $D^{*\pm}D^{\mp}$ mass distribution in the range 5000 to 5600 MeV/c^2 is shown, overlaid with the projection of the fitted PDF. Below the actual mass distribution a so-called pull distribution is displayed. The entries are calculated by the difference between the data points and the PDF is normalised to the uncertainty on the data point.

The mass range is adapted to the range used in the fit to detector data, in which the shared mean varies, while all further parameters are fixed to the values obtained in the fit to simulated data (see Tab. 6.1). For the $B_s^0 \to D^{*\pm}D^{\mp}$ component, the same model as for the signal is used, except that the mean is shifted by the central value of the world average $\mu_{B_s^0} - \mu_{B^0} = (87.23 \pm 0.24)\,\mathrm{MeV}/c^2$ [22].

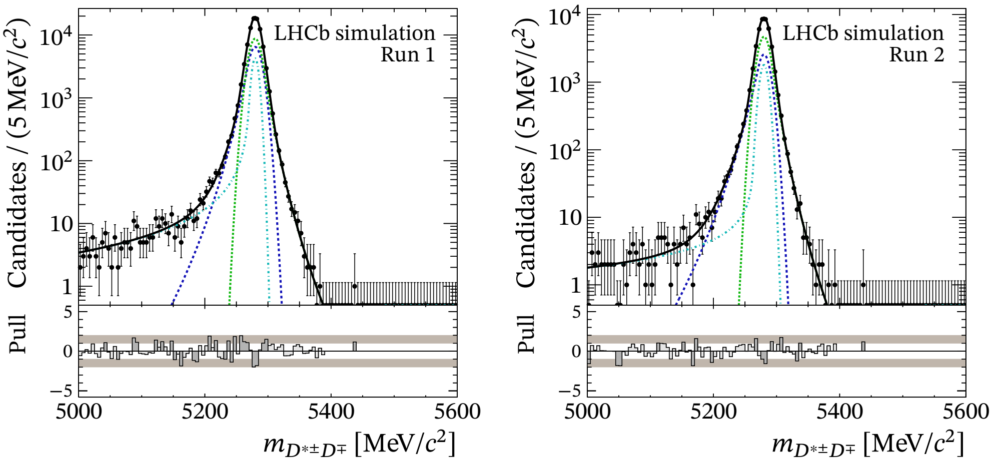

Figure 6.2 – Mass distributions of the invariant $D^{*\pm}D^{\mp}$ mass of simulated $B^0 \to D^{*\pm}D^{\mp}$ decays for (left) Run 1 and (right) Run 2. The fitted total PDF is overlaid in black, the three Crystal Ball components are presented by the lines in green dotted, blue dashed and turquoise dash-dotted.

Table 6.1 – Fit results of the mass fit to simulated $B^0 \to D^{*\pm}D^{\mp}$ decays for Run 1 and Run 2. The indices of the parameters indicate the correspondence to the three Crystal Ball shapes used in the fit. The parameter $n_{1,2}$ is shared by two Crystal Ball functions and the fractions g_i^{MC}, which combine the components, are defined recursively.

Parameter	Fitted Value Run 1	Fitted Value Run 2	Unit
μ_{B^0}	5279.80 ± 0.04	5280.00 ± 0.08	MeV/c^2
σ_1	9.71 ± 0.28	9.35 ± 0.72	MeV/c^2
σ_2	9.29 ± 0.24	9.31 ± 0.49	MeV/c^2
σ_3	5.32 ± 0.40	6.52 ± 0.61	MeV/c^2
α_1	1.17 ± 0.09	0.92 ± 0.09	
α_2	-1.43 ± 0.08	-1.39 ± 0.08	
α_3	1.91 ± 0.22	2.64 ± 0.34	
g_1^{MC}	0.38 ± 0.05	0.32 ± 0.05	
g_2^{MC}	0.77 ± 0.06	0.79 ± 0.15	
$n_{1,2}$	9.27 ± 1.29	9.41 ± 1.78	
n_3	1.20 ± 0.22	0.68 ± 0.31	

6.1.2 Mis-identified background

As already explained in Sec. 5.5, the contribution from $B^0 \to D^{*-}D_s^+$ decays due to mis-identified kaons is not fully suppressed and needs to be modelled in the invariant $D^{*\pm}D^{\mp}$ mass distribution. This background is parameterised by the sum of two Crystal Ball PDFs, which have a shared mean and tails to opposite sides. The maximum likelihood fit is performed on simulated $B^0 \to D^{*-}D_s^+$ decays that are reconstructed as the signal decay. As the distributions of the simulated samples corresponding to Run 1 and Run 2 are compatible within statistical uncertainties, the samples are combined in the fit. In Fig. 6.3, the invariant $D^{*\pm}D^{\mp}$ mass distribution in the range of 5000 to 5600 MeV/c^2 is shown, overlaid with the projection of the fitted PDF. In the fit to detector data the mean is floating, while all other parameters are fixed to the values obtained in the fit to simulated data that are given in Tab. 6.2.

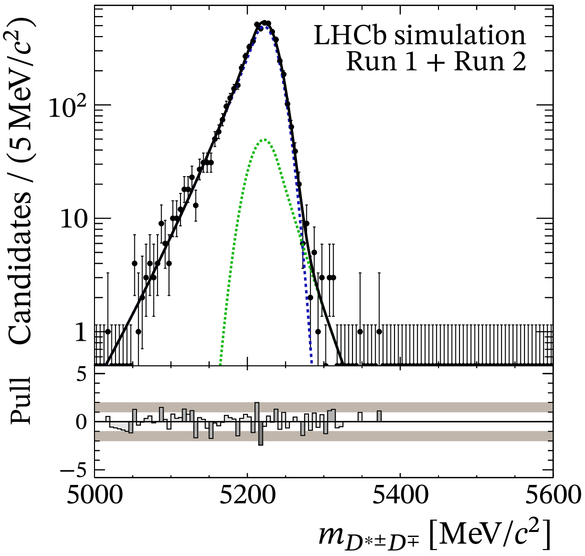

Figure 6.3 – Mass distributions of the invariant $D^{*\pm}D^{\mp}$ mass of simulated $B^0 \to D^{*-}D_s^+$ decays reconstructed as signal for the merged data sample of Run 1 and Run 2. The fitted PDF is overlaid in black, the two Crystal Ball components are presented by the lines in green dotted an blue dashed.

6.1.3 Partially reconstructed background

Besides $\overset{(-)}{D}{}^0\pi^{\pm}$, the $D^{*\pm}$ meson is able to decay in $D^{\pm}\pi^0$ and $D^{\pm}\gamma$. The final state particles of the decay $B^0_{(s)} \to D^{*+}D^{*-}$ can be falsely reconstructed as signal, as a π^0 or a photon are not considered by the reconstruction. The $D^{*\pm} \to D^{\pm}\pi^0$ decay has a probability of

Table 6.2 – Fit results of the mass fit to simulated $B^0 \to D^{*-}D_s^+$ decays reconstructed as signal for Run 1 and Run 2 merged. The indices of the parameters indicate the correspondence to the two Crystal Ball functions used in the fit. The parameter $n_{1,2}$ is shared by the two Crystal Ball functions and the fractions g_i^{MC}, which combine the components, are defined recursively.

Parameter	Fitted Value Run 1 + Run 2	Unit
μ_{B^0}	5221.7 ± 0.5	MeV/c^2
σ_1	16.9 ± 0.5	MeV/c^2
σ_2	18.88 ± 3.9	MeV/c^2
α_1	0.69 ± 0.06	
α_2	-1.01 ± 0.32	
g^{MC}	0.91 ± 0.06	
$n_{1,2}$	18.7 ± 9.4	

30%, while the $D^{*\pm} \to D^{\pm}\gamma$ is far less probable (2%) and, hence, will be neglected in the following. However, the reconstructed invariant mass will be smaller than the B^0 mass by at least the mass of the missed pion. Depending on the polarisation of the $D^{*\pm}$, the reconstructed invariant B^0 mass has two peaks (longitudinal) or a single peak (transverse). The characteristic shape of a double peak structure is a consequence of the spin of the D^{*+} meson ($s = 1$) and the small mass difference between the D^{*+} mass and the combined mass of the D^+ and π^0, which gives the latter daughter particles only a small boost. The angular distribution of the $D^{*+} \to D^+\pi^0$ decay causes a separation in the invariant $D^{*\pm}D^{\mp}$ mass distribution if the uncharged pion is not reconstructed. Therefore, the two peaks are related to the π^0 flight direction with respect to the D^{*+} boost direction. Either it is emitted in approximately the same or in the opposite direction compared to the D^{*+} boost. Simulated $B^0 \to D^{*+}D^{*-}$ and $B_s^0 \to D^{*+}D^{*-}$ decays are used to investigate the mass model for the two cases of pure longitudinal and pure transverse polarisations. As the resulting distributions are compatible within the statistical uncertainties between the runs, they are merged for simplicity. The model for the longitudinal case consists of four Gaussian functions. Two of them describe the two peaks, thus, they have different means, while their widths are shared. For the wider component under the double peak, two Gaussian functions are modelled with a common mean. In the distribution of the transverse case, the peak is described by two Gaussians with shared widths and separate means. The broad component is parameterised by a function that is defined as a uniform distribution between two kinematically constrained values and that is convolved with a Gaussian function with floating width. Since the distribution theoretically ends abruptly at the kinematic boundaries, experimental smearings are present at the edges, which are described by the Gaussian component.

In Fig. 6.4, the plot of the invariant $D^{*\pm}D^{\mp}$ mass distribution in the range 5000 to 5600 MeV/c^2 is shown, overlaid with the projection of the fitted PDF. In the fit to detector data, the relative fraction of the transverse and longitudinal contributions is floating and all parameters are fixed to the values obtained in the fit to simulated events given in Tab. 6.3.

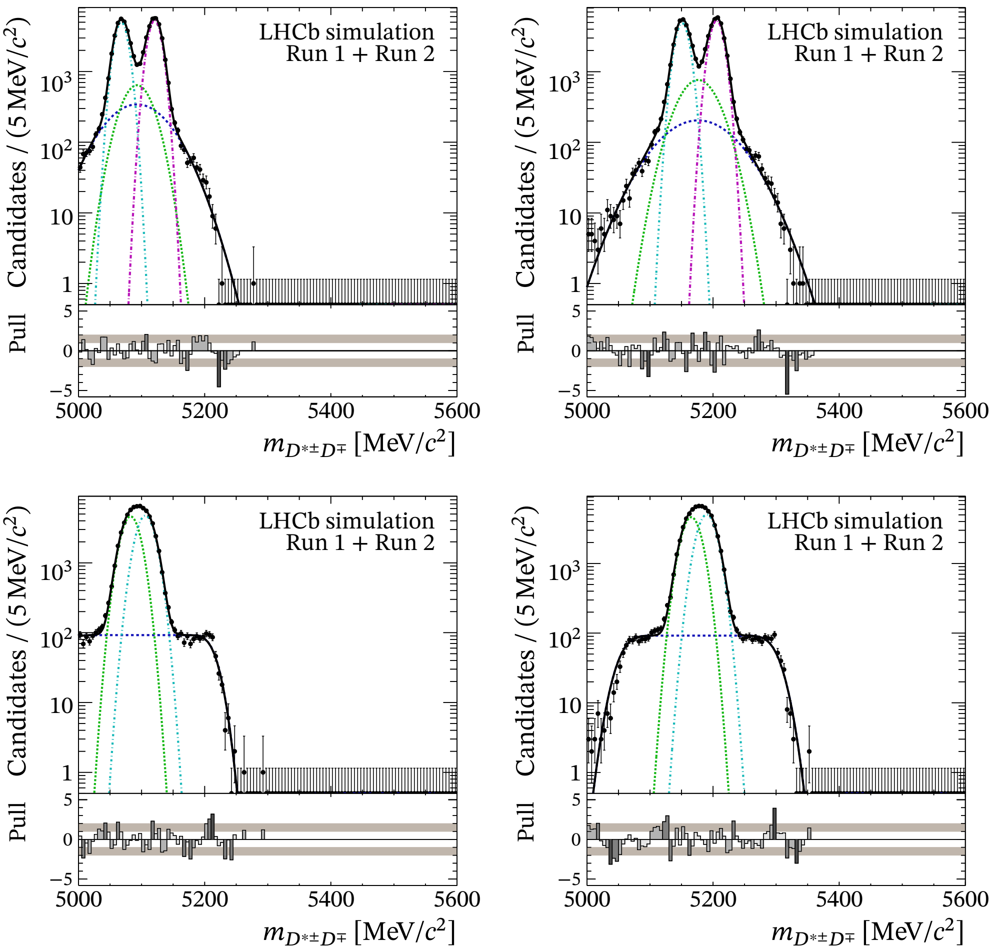

Figure 6.4 – Mass distributions of the invariant $D^{*\pm}D^{\mp}$ mass of simulated (left) $B^0 \to D^{*+}D^{*-}$ and (right) $B^0_s \to D^{*+}D^{*-}$ decays reconstructed as signal for the merged dataset of Run 1 and Run 2. The upper plots refer to the longitudinal polarisation, bottom plots correspond to the transverse polarisation. The fitted PDF is overlaid in black.

Table 6.3 – Fit results of the mass fit to simulated $B^0 \to D^{*+}D^{*-}$ and $B_s^0 \to D^{*+}D^{*-}$ decays, which are reconstructed as signal and generated with pure longitudinal or transverse polarisations. The indices of the parameters indicate the correspondence to the Gaussian component used in the fit. The fractions g_i^{MC}, which combine the components, are defined recursively. Events for Run 1 and Run 2 are merged.

Sample	Parameter	Fitted Value		Unit
		longitudinal	transverse	
$B^0 \to D^{*+}D^{*-}$	μ_1^{MC}	5092.8 ± 0.6	–	MeV/c^2
	μ_2^{MC}	$\equiv \mu_1^{\mathrm{MC}}$	–	MeV/c^2
	μ_3^{MC}	5067.8 ± 0.1	5082.4 ± 0.2	MeV/c^2
	μ_4^{MC}	5120.8 ± 0.9	5106.0 ± 0.2	MeV/c^2
	σ_1^{MC}	44.8 ± 1.1	13.8 ± 0.8	MeV/c^2
	σ_2^{MC}	21.5 ± 1.9	–	MeV/c^2
	σ_3^{MC}	9.68 ± 0.08	13.45 ± 0.10	MeV/c^2
	σ_4^{MC}	$\equiv \sigma_3^{\mathrm{MC}}$	$\equiv \sigma_3^{\mathrm{MC}}$	MeV/c^2
	g_1^{MC}	0.117 ± 0.010	0.060 ± 0.001	
	g_2^{MC}	0.122 ± 0.018	–	
	g_3^{MC}	0.483 ± 0.003	0.494 ± 0.007	
$B_s^0 \to D^{*+}D^{*-}$	μ_1^{MC}	5176.3 ± 0.5	–	MeV/c^2
	μ_2^{MC}	$\equiv \mu_1^{\mathrm{MC}}$	–	MeV/c^2
	μ_3^{MC}	5150.9 ± 0.1	5165.4 ± 0.2	MeV/c^2
	μ_4^{MC}	5205.6 ± 0.1	5189.9 ± 0.2	MeV/c^2
	σ_1^{MC}	53.5 ± 1.4	16.7 ± 0.6	MeV/c^2
	σ_2^{MC}	27.3 ± 0.9	–	MeV/c^2
	σ_3^{MC}	10.14 ± 0.08	13.93 ± 0.10	MeV/c^2
	σ_4^{MC}	$\equiv \sigma_3^{\mathrm{MC}}$	$\equiv \sigma_3^{\mathrm{MC}}$	MeV/c^2
	g_1^{MC}	0.081 ± 0.008	0.064 ± 0.001	
	g_2^{MC}	0.168 ± 0.013	–	
	g_3^{MC}	0.479 ± 0.003	0.479 ± 0.006	

6.2 Extended maximum likelihood fit to LHCb data

An extended maximum likelihood fit (see Sec. 4.1) is performed to detector data, using the shape parameters that are fixed to the values given in Tab. 6.1 to 6.3. The combinatorial background is modelled by an exponential function with the slope β that is floating in the fit to detector data.

Separate fits are performed for Run 1 and Run 2 in each final state $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^-\pi^+\pi^+$. For $D^0 \to K^-\pi^-\pi^+\pi^+$, the fits are performed by a proponent of the analysis group. The configurations of the parameterisation between the final states are set up as similar as possible. Still, one difference exists for the signal component. In

$D^0 \to K^-\pi^-\pi^+\pi^+$ instead of three only two Crystal Ball functions are used, since the tail towards smaller masses is less prominent. The background components have the same model as in $D^0 \to K^-\pi^+$, except that the values are determined on simulation adapted for the $D^0 \to K^-\pi^-\pi^+\pi^+$ final state.

In Fig. 6.5, the plots of the reconstructed mass distributions in the range of 5000 to 5600 MeV/c^2 are shown. The projections of the fitted PDFs are superimposed in the figures. In Tab. 6.4, the fit results for the yields and floating parameters are given. The peak position of the B^0 signal is compatible with the current world average for the nominal B^0 mass of (5279.63 ± 0.15) MeV/c^2 [22]. It is tested if the mass resolution on data and simulation is compatible by introducing a scaling factor for the widths of the mass components. The values for this scaling factor are in very good agreement with unity. Thus, it can be concluded that the mass resolution in MC is already well simulated and it is continued without scaling. The impact of neglecting the scaling will be investigated later in terms of a systematic uncertainty. Considering the different cross sections for $pp \to b\overline{b} + X$ decays taken from [118–120] and the different luminosities for the years of data taking, in Run 2 compared to Run 1, 3.5 times more signal candidates are expected. In the final state $D^0 \to K^-\pi^+$, a ratio of 3.81 ± 0.16 and in $D^0 \to K^-\pi^-\pi^+\pi^+$ a ratio of 3.35 ± 0.22 is determined. Both values are in good agreement with the expectation.

The fit to the invariant $D^{*\pm}D^{\mp}$ mass distribution is utilised to extract sWeights, which are needed to subtract remaining background from the signal to enable a fit to the B^0 decay time of the signal candidates. The decay-time fit to extract the *CP* parameters is described in the next chapter.

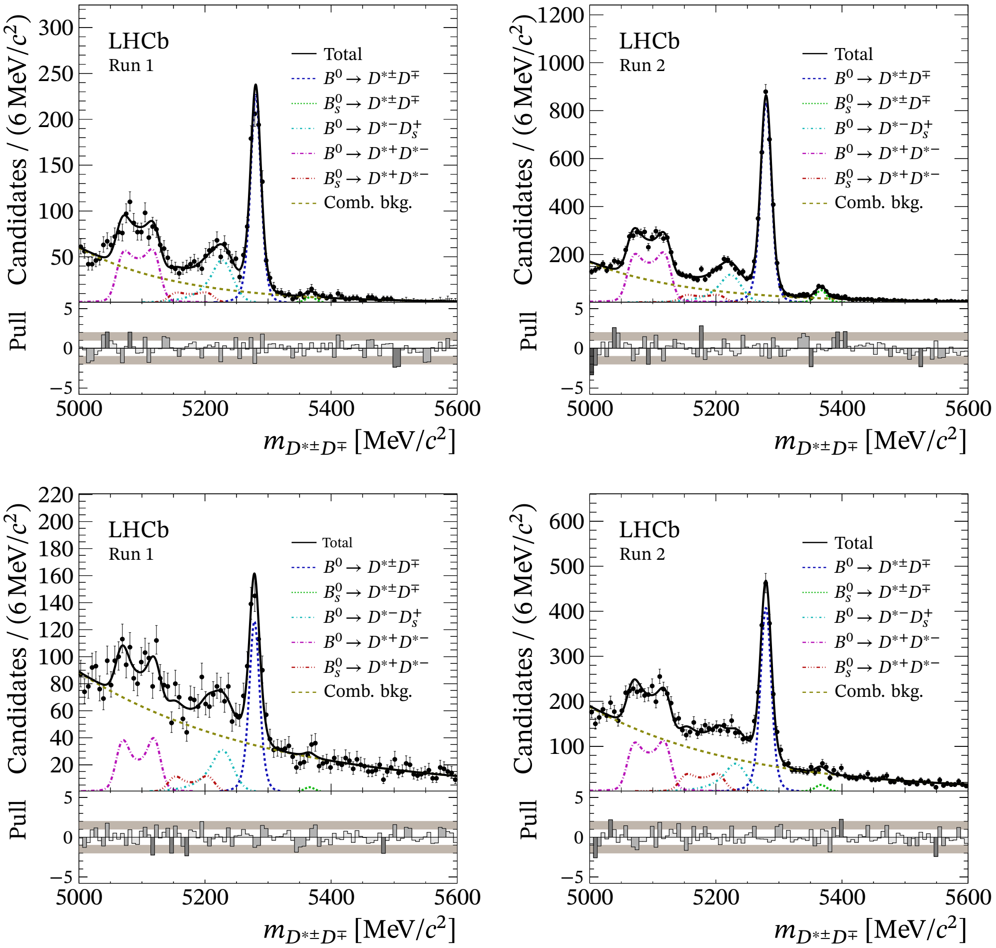

Figure 6.5 – Mass distribution of the reconstructed $B^0 \to D^{*\pm}D^{\mp}$ dataset for (left) Run 1 and (right) Run 2 and for the two final states $D^0 \to K^-\pi^+$ (top) and $D^0 \to K^-\pi^-\pi^+\pi^+$ (bottom). The fitted PDF is overlaid in black. In addition, projections of the components of the PDF are displayed, namely the B^0 signal (dashed blue), the $B_s^0 \to D^{*\pm}D^{\mp}$ background (dotted green), the $B^0 \to D^{*-}D_s^+$ background (long-dash-dotted turquoise), the $B^0 \to D^{*+}D^{*-}$ background (dash-dotted magenta), the $B_s^0 \to D^{*+}D^{*-}$ background (dash-dotted brown) and the combinatorial background (dotted dark green).

Table 6.4 – Results of the floating shape parameters in the extended maximum likelihood fit to Run 1 and Run 2 data for the final states $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^-\pi^+\pi^+$.

Final state	Parameter	Fitted Value		Unit
		Run 1	Run 2	
$D^0 \to K^-\pi^+$	μ_{B^0}	5280.4 ± 0.4	5279.6 ± 0.2	MeV/c^2
	$\mu_{B^0\to D^{*-}D_s^+}$	5226.9 ± 2.1	5223.0 ± 1.4	MeV/c^2
	$\beta \times 10^3$	-6.16 ± 0.22	-6.24 ± 0.14	c^2/MeV
	$g^{\text{HEL}}_{B^0\to D^{*+}D^{*-}}$	0.49 ± 0.06	0.52 ± 0.03	
	$N_{B^0\to D^{*\pm}D^{\mp}}$	856 ± 32	3265 ± 61	
	$N_{B_s^0\to D^{*\pm}D^{\mp}}$	24 ± 8	197 ± 19	
	$N_{B^0\to D^{*-}D_s^+}$	404 ± 33	1007 ± 57	
	$N_{B^0\to D^{*+}D^{*-}}$	693 ± 23	2451 ± 75	
	$N_{B_s^0\to D^{*+}D^{*-}}$	141 ± 36	395 ± 64	
	N_{Comb}	1579 ± 66	4359 ± 111	
$D^0 \to K^-\pi^-\pi^+\pi^+$	μ_{B^0}	5278.5 ± 0.7	5279.3 ± 0.3	MeV/c^2
	$\mu_{B^0\to D^{*-}D_s^+}$	5227.2 ± 3.7	5231.2 ± 2.8	MeV/c^2
	$\beta \times 10^3$	-3.37 ± 0.12	-4.17 ± 0.09	c^2/MeV
	$g^{\text{HEL}}_{B^0\to D^{*+}D^{*-}}$	0.65 ± 0.09	0.53 ± 0.05	
	$N_{B^0\to D^{*\pm}D^{\mp}}$	469 ± 28	1570 ± 48	
	$N_{B_s^0\to D^{*\pm}D^{\mp}}$	13 ± 13	54 ± 18	
	$N_{B^0\to D^{*-}D_s^+}$	288 ± 40	573 ± 52	
	$N_{B^0\to D^{*+}D^{*-}}$	433 ± 49	1300 ± 72	
	$N_{B_s^0\to D^{*+}D^{*-}}$	123 ± 48	461 ± 63	
	N_{Comb}	3798 ± 99	6911 ± 135	

7 Measurement of *CP* Violation in $B^0 \to D^{*\pm} D^{\mp}$ Decays

This chapter describes the measurement of the *CP* parameters S_{D^*D}, ΔS_{D^*D}, C_{D^*D}, ΔC_{D^*D} and A_{D^*D} that are obtained in a multi-dimensional unbinned maximum-likelihood fit to the decay-time distribution of $B^0 \to D^{*\pm} D^{\mp}$ signal candidates. The background in the sample is subtracted by applying the sWeights from the fit to the B^0 invariant-mass distribution. The results regarding the decay-time fit are obtained in collaboration with the proponents of the analysis group.

The decay-time parameterisation is described in Sec. 7.1, while experimental influences and their consideration in the measurement are introduced in Sec. 7.2 to Sec. 7.5. The extraction of the *CP* parameter is discussed in Sec. 7.6.

7.1 Decay-time parameterisation

The PDF of the reconstructed decay time t, the tag decisions $\vec{d} = (d^{\mathrm{OS}}, d^{\mathrm{SS}})$ and final state f given by the per-event mistag probability estimates $\vec{\eta} = (\eta^{\mathrm{OS}}, \eta^{\mathrm{SS}})$, is defined as

$$\mathcal{P}\left(t, f, \vec{d} \,\middle|\, \vec{\eta}\right) \propto \varepsilon(t) \cdot \left(P(t', f, \vec{d}|\vec{\eta}) \otimes \mathcal{R}(t - t')\right). \tag{7.1}$$

Here, $\varepsilon(t)$ is the decay-time-dependent efficiency discussed in Sec. 7.4. The PDF $P(t', f, \vec{d}|\vec{\eta})$ describes the true decay times t' of the B^0 mesons and $\mathcal{R}(t - t')$ incorporates the decay-time resolution described in Sec. 7.3. Depending on the B^0 flavour and final state, the PDF describing the B^0 decay time can be expressed as follows

$$\begin{aligned} P^{\mathrm{theo}}_{D^{*\pm}D^{\mp}}(t') \propto{} & \mathrm{e}^{-t/\tau}(1 \pm A_{D^*D}) \times \\ & (1 - d_{\mathrm{true}}[(S_{D^*D} \pm \Delta S_{D^*D}) \sin(\Delta m_d t) + (C_{D^*D} \pm \Delta C_{D^*D}) \cos(\Delta m_d t)]), \end{aligned} \tag{7.2}$$

where all experimental effects are omitted for now. The true production flavour of the B^0 meson is given by d_{true} and $\Delta\Gamma$ is assumed to be zero. As already mentioned in Sec. 3.2, the production asymmetry of B^0 and $\bar{B}^0$ mesons needs to be taken into account via Eq. (3.1). Furthermore, a detection asymmetry is incorporated, which considers different efficiencies ϵ in the detection and reconstruction of the final states f and $\bar{f}$. The detection asymmetry, which is defined as

$$A_{\mathrm{det}} = \frac{\epsilon(f) - \epsilon(\bar{f})}{\epsilon(f) + \epsilon(\bar{f})}\,, \tag{7.3}$$

can arise, if interaction cross-sections with the detector material depend on the charge of final state particles and from mis-alignments or inefficiencies of detector components.

The OS and SS tag decisions are combined in the PDF of the B^0 decay-time fit. To express the combination of the two tags, it is useful to define the quantity $\Delta^{\pm}$. Considering two taggers i and j, if there is only a tag decision from tagger i, $\Delta^{\pm}$ results in

$$\begin{aligned}\Delta^{\pm} =& \frac{1}{2}\varepsilon^{i}_{\mathrm{tag}}\Big[1 - \varepsilon^{j}_{\mathrm{tag}} + d^{i}\Big(1 - \varepsilon^{j}_{\mathrm{tag}} - 2\omega^{i}_{B^0}(1 - \varepsilon^{j}_{\mathrm{tag}})\Big)\Big] \\ &\pm \frac{1}{2}\varepsilon^{i}_{\mathrm{tag}}\Big[1 - \varepsilon^{j}_{\mathrm{tag}} - d^{i}\Big(1 - \varepsilon^{j}_{\mathrm{tag}} - 2\omega^{i}_{\bar{B}^0}(1 - \varepsilon^{j}_{\mathrm{tag}})\Big)\Big]\,,\end{aligned} \tag{7.4}$$

where ω_{B^0} and $\omega_{\bar{B}^0}$ are the calibrated mistag probabilities for a B^0 and $\bar{B}^0$ tagged by the tagger i and $\varepsilon^{\mathrm{OS,SS}}_{\mathrm{tag}}$ are the tagging efficiencies for the tags of OS and SS. These probabilities are expressed in terms of the observable η, the predicted mistag probability, by means of the calibration function of Eq. (7.9). If both OS and SS taggers provide tag decisions, $\Delta^{\pm}$ can be written as

$$\begin{aligned}\Delta^{\pm} =& \frac{1}{4}\varepsilon^{\mathrm{SS}}_{\mathrm{tag}}\varepsilon^{\mathrm{OS}}_{\mathrm{tag}}\Big[1 + \sum_{j=\mathrm{OS,SS}} d_j(1 - 2\omega^{j}_{B^0}) + d^{\mathrm{OS}}d^{\mathrm{SS}}\Big(1 - 2\omega^{j}_{B^0} + 2\omega^{\mathrm{OS}}_{B^0}\omega^{\mathrm{SS}}_{B^0}\Big)\Big] \\ &\pm \frac{1}{4}\varepsilon^{\mathrm{SS}}_{\mathrm{tag}}\varepsilon^{\mathrm{OS}}_{\mathrm{tag}}\Big[1 + \sum_{j=\mathrm{OS,SS}} d_j(1 - 2\omega^{j}_{\bar{B}^0}) + d^{\mathrm{OS}}d^{\mathrm{SS}}\Big(1 - 2\omega^{j}_{\bar{B}^0} + 2\omega^{\mathrm{OS}}_{\bar{B}^0}\omega^{\mathrm{SS}}_{\bar{B}^0}\Big)\Big]\,.\end{aligned} \tag{7.5}$$

Considering the experimental effects of the flavour tagging as well as the production and detection asymmetries, the PDF for the decay-time of the B^0 mesons is given by

$$P(t', f, \vec{d}) = (1 + qA_{\mathrm{raw}})(\Delta^{-} - A_{\mathrm{prod}}\Delta^{+})P^{\mathrm{theo}}(t', f, \vec{d})\,, \tag{7.6}$$

where $q = 1$ denotes the $D^{*+}D^{-}$ and $q = -1$ indicates the $D^{*-}D^{+}$ final state. The parameter A_{raw} combines the detection asymmetry A_{det} and the CP parameter A_{D^*D} via

$$A_{\mathrm{raw}} = A_{\mathrm{det}} + A_{D^*D}\,. \tag{7.7}$$

In order to extract A_{D^*D}, A_{det} needs to be measured separately (see Sec. 7.5).

7.2 Flavour tagging calibration

The predicted mistag probability η needs to be calibrated in order to measure the true flavour tagging error rate that is used in the decay-time PDF. This quantity is defined as the output of multivariate classifiers, which are trained with datasets of flavour-specific B decays. Hence, it is necessary to ensure that a reliable mistag probability is used in the fit. The linear function

$$\omega(\eta) = p_0 + p_1\,(\eta - \langle\eta\rangle) \tag{7.8}$$

is an approximation at first order that converts the predicted mistag probability estimate η into the corrected mistag probability ω. Using the average mistag probability estimate $\langle\eta\rangle$ leads to a decorrelation of the parameters p_0 and p_1. A perfect calibration would result in $p_0 = \langle\eta\rangle$ and $p_1 = 1$. The calibration allows calibrated mistag probabilities $\omega(\eta)$ to be larger than 0.5. In such cases the tag is reversed, i.e. $d' = -d$, and the mistag probability is adjusted to $\omega'(\eta) = 1 - \omega(\eta)$.

Tagging asymmetries can occur due to asymmetries in the reconstruction of the tagging particles. It is important to account for these asymmetries, as they can bias the measured *CP* asymmetry. Therefore, two separate parameterisations for the calibrations of initial B^0 and $\bar{B}^0$ mesons are introduced,

$$\begin{aligned} \omega_{B^0}(\eta) &= p_0^{B^0} + p_1^{B^0}\,(\eta - \langle\eta\rangle)\;, \\ \omega_{\bar{B}^0}(\eta) &= p_0^{\bar{B}^0} + p_1^{\bar{B}^0}\,(\eta - \langle\eta\rangle)\;. \end{aligned} \tag{7.9}$$

The calibration parameters $p_i^{B^0}$ and $p_i^{\bar{B}^0}$ with $i = 0, 1$ are combined to

$$p_i = \frac{1}{2}(p_i^{B^0} + p_i^{\bar{B}^0}),\;\; \Delta p_i = p_i^{B^0} - p_i^{\bar{B}^0}\;. \tag{7.10}$$

The mistag probability asymmetry is defined as the difference between the mistag probabilities of true initial B^0 and $\bar{B}^0$ mesons

$$\begin{aligned} \Delta\omega(\eta) &= \omega_{B^0}(\eta) - \omega_{\bar{B}^0}(\eta) \\ &= \Delta p_0 + \Delta p_1\,(\eta - \langle\eta\rangle)\;. \end{aligned} \tag{7.11}$$

7.2.1 Strategy

The calibration used in the *CP* violation measurement presented in this thesis has been provided by collaborators of the analysis group. It is beneficial to the flavour tagging calibration if the kinematic properties between calibration and signal channel are as similar as possible. This allows systematic uncertainties from the flavour tagging to be reduced in size. In addition, the control channel should have a larger number of candidates in order to minimise statistical uncertainties on the calibration parameters.

For this analysis, $B^0 \to D^{*-}D_s^+$ and $B^0 \to D_s^+D^-$ are chosen as calibration channels for the OS and SS tagger combinations due to their kinematic similarities with the signal channel. However, the number of candidates of these two control channels are not sufficient to calibrate individual taggers. Hence, single taggers are pre-calibrated with the channel $B^0 \to D^-\pi^+$, which has nine times more candidates compared to $B^0 \to D^{*-}D_s^+$ and $B^0 \to D_s^+D^-$ combined.

The calibration parameters for the single taggers are estimated by the Espresso Performance Monitor (EPM) tool [121] separately for Run 1 and Run 2. The EPM is developed by the LHCb collaboration to perform unbinned, numerical calibrations of flavour tagging algorithms with binomial regression. It comprises useful features, e.g. it allows to calibrate tagging algorithms with charged self-tagging decays and neutral flavour-specific decays. Additionally, it supports background subtraction by applying sWeights and accounts for the decay-time resolution for B^0 decays. Due to the fact that it does not consider production, detection or tagging-efficiency asymmetries, the EPM tool is only used for the single tagger calibration and for cross-checks of the tagger combination calibration. The final flavour tagging parameters are obtained from a fit to the decay-time distributions of $B^0 \to D_s^+D^-$ and $B^0 \to D^{*-}D_s^+$ decays separately for Run 1 and Run 2. The averages of the results from the calibration channels $B^0 \to D_s^+D^-$ and $B^0 \to D^{*-}D_s^+$, weighted by their statistical uncertainties, are used in the *CP* measurement.

All flavour tagging algorithms described in Sec. 4.6 are used, except for the OSCharm tagger, which is only available for Run 2 datasets.

7.2.2 Single tagger calibration

The single tagger calibration is performed on $B^0 \to D^-\pi^+$ decays. This channel is very commonly used to calibrate taggers, as it has a large branching fraction and a controllable background level [122]. In this channel the effects of *CP* violation are very small and are neglected for the purpose of this analysis. The datasets are provided by LHCb's flavour tagging group and already include a $B^0 \to D^-\pi^+$ signal selection and sWeights. The number of signal $B^0 \to D^-\pi^+$ candidates in the Run 1 and Run 2 samples are about 417 000 and 710 000, respectively. In order to make the kinematic of $B^0 \to D^-\pi^+$ more similar to that of the signal channel, the transverse momentum distribution of the B^0 candidate is re-weighted in such a way that it matches the p_{T} distribution of the signal decay. This is important, as the tagging performance, especially of the SS taggers, is correlated to the transverse momentum of the B^0 meson. The tagging calibration is performed with the EPM tool considering the sWeights, which were obtained from a fit to the reconstructed B^0 mass distribution beforehand, and the weights for the p_{T} correction. The calibration parameters for Run 1 and Run 2 are listed in Tab. 7.1 and Tab. 7.2. A selection of corresponding calibration plots is presented in Fig. 7.1.

Table 7.1 – Tagging calibration parameters of the OS and SS single taggers for Run 1. The calibration is performed with the EPM tool using $B^0 \to D^-\pi^+$ decays from the 2012 LHCb dataset re-weighted such to reproduce the p_{T} distribution of the signal B^0.

Tagger	$\langle\eta\rangle$	$p_0 - \langle\eta\rangle$	p_1
OSμ	0.3099	0.0139 ± 0.0047	1.063 ± 0.051
OSe	0.2986	0.0317 ± 0.0075	1.061 ± 0.117
OSK	0.3957	-0.0136 ± 0.0031	0.942 ± 0.052
OSVtx	0.3900	0.0078 ± 0.0029	0.984 ± 0.046
SSπ	0.4338	-0.0006 ± 0.0016	1.036 ± 0.029
SSp	0.4542	-0.0011 ± 0.0022	0.999 ± 0.046

Table 7.2 – Tagging calibration parameters of the OS and SS single taggers for Run 2. The calibration is performed with the EPM tool using $B^0 \to D^-\pi^+$ decays from the 2016 LHCb dataset re-weighted such to reproduce the p_{T} distribution of the signal B^0.

Tagger	$\langle\eta\rangle$	$p_0 - \langle\eta\rangle$	p_1
OSμ	0.3341	0.0011 ± 0.0039	1.336 ± 0.057
OSe	0.3594	0.0133 ± 0.0058	1.206 ± 0.099
OSK	0.4026	-0.0194 ± 0.0057	1.273 ± 0.115
OSVtx	0.3899	-0.0044 ± 0.0026	1.035 ± 0.041
OSCharm	0.3655	-0.0173 ± 0.0047	0.977 ± 0.095
SSπ	0.4336	0.0019 ± 0.0016	1.007 ± 0.031
SSp	0.4517	-0.0038 ± 0.0018	1.079 ± 0.038

7.2.3 Tagger combination calibration

For the calibration of the OS and SS tagger combinations the neutral control channels $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*-} D_s^+$ are used. These decays are flavour specific, hence, the charge of the $D_s^\pm$ meson allows to reconstruct the B^0 meson flavour at the time of decay.

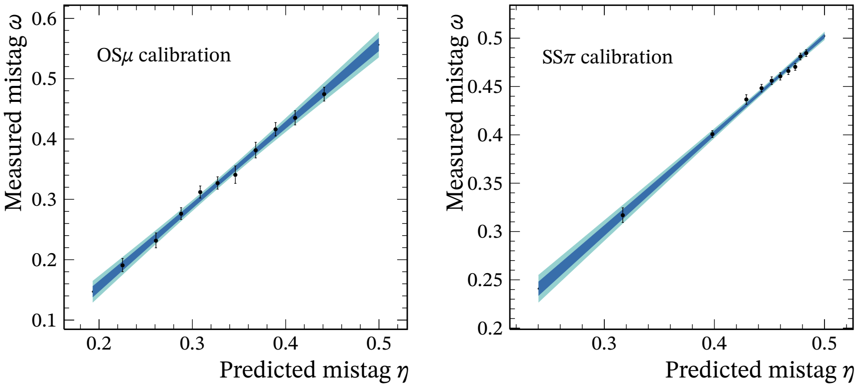

Figure 7.1 – The figures show exemplary results from the EPM tagging calibrations for (left) the OSμ and (right) SSπ taggers, obtained using p_{T} corrected $B^0 \to D^-\pi^+$ decays from 2016 data. The blue and light blue areas show the one and two sigma confidence intervals.

In both control channels, the D_s^+ mesons are reconstructed as $D_s^+ \to K^-K^+\pi^+$, while the D^- and D^{*-} are reconstructed as in the signal channel.

The selection applied to the control channels is very similar to the cut-based selection of the signal selection. Requirements on all D meson masses are introduced. All reconstructed D meson masses are required to lie within a defined mass window around the nominal D meson mass. Furthermore, DLL requirements are applied to reduce $K \leftrightarrow \pi$ mis-identification in the D_s^+ and D^- reconstructions. After the preselection in $B^0 \to D_s^+D^-$, vetoes against contaminations from $B^0 \to D^+D^-$ or $B^0 \to D_s^+D_s^-$ decays are applied. Additionally, candidates with a reconstructed invariant B^0 mass below $5220\,\mathrm{MeV}/c^2$ are neglected, in order to exclude physical background contributions.

Extended maximum-likelihood fits to the invariant $m_{D_s^+D^-}$ and $m_{D^{*-}D_s^+}$ masses are performed to statistically subtract background, that passes the selection using the sPlot method (see Sec. 4.2). In $B^0 \to D^{*-}D_s^+$ the mass parameterisation comprises the signal and a combinatorial background component. The B^0 signal is modelled by two Gaussian functions with a shared mean and different widths. The combinatorial background is parameterised as in $B^0 \to D^{*\pm}D^{\mp}$ with an exponential function. All shape parameters and yields are optimised by the fit.

Besides the signal and combinatorial background contributions, there is a contamination of $B_s^0 \to D_s^+D^-$ decays present in the $B^0 \to D_s^+D^-$ channel. It is modelled by the same PDF as $B^0 \to D_s^+D^-$ candidates, with the mean position shifted by the central value of the world average $m_{B_s^0} - m_{B^0} = (87.23 \pm 0.24)\,\mathrm{MeV}/c^2$ [22]. In Fig. 7.2, the invariant B^0 mass distributions of the data for $B^0 \to D_s^+D^-$ and $B^0 \to D^{*-}D_s^+$ control channels are shown with the overlaid fit projections.

From the charge configuration of the final state of the calibration channels, it is possible to reconstruct the meson flavour at the time of decay. In order to relate the production flavour and, thus, to calibrate the flavour tagging it is necessary to take the $B^0-\bar{B}^0$ mixing into account. Hence, a decay-time fit similar to the *CP* fit to $B^0 \to D^{*\pm}D^{\mp}$ decays is performed, in which the presumably true mistag probability ω is determined. In this fit, no *CP* violation and full flavour-specificity ($S=0, C=1$) is assumed. The flavour tagging parameters float in the fit.

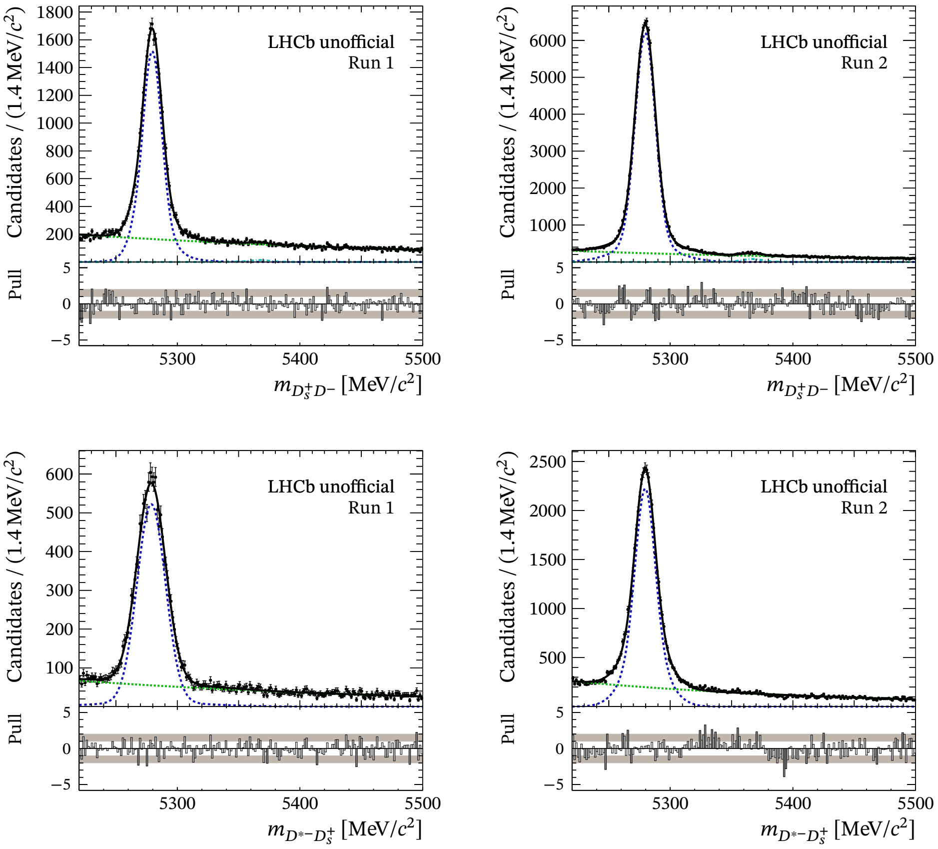

Figure 7.2 – Invariant mass distributions of (top) $B^0 \to D_s^+D^-$ candidates and (bottom) $B^0 \to D^{*-}D_s^+$ for the (left) Run 1 and (right) Run 2 data samples. The fitted PDF is overlaid in black. In addition, projections of the components of the PDF are displayed, namely the signal (dashed blue), combinatorial background (dotted green) and, only in the top plots, $B_s^0 \to D_s^+D^-$ (long-dash-dotted turquoise).

In this analysis, other experimental impacts can be taken into account: The production asymmetry of B^0 mesons and the detection asymmetry of the final state particles. The flavour tagging is performed using a linear calibration as introduced in Eq. (7.8) and tagging asymmetries are taken into account (see Eq. (7.11)). The calibration result from the unbinned maximum likelihood fits to the decay time, tag distributions and decay-time-error estimates in the two control channels for Run 1 and Run 2 are listed in Tab. 7.3 and Tab. 7.4.

Table 7.3 – Flavour tagging parameters of the fit to the control channels $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*-} D_s^+$ for the flavour tagging calibration for Run 1. For all calibration parameters a weighted average of the results from the two channels is given, using the statistical uncertainties as weights. The given uncertainty for each weighted average is the combination of the statistical and systematic uncertainties. They are used in the fit to the decay-time distribution of the $B^0 \to D^{*\pm}D^{\mp}$ signal candidates.

Parameter	$B^0 \to D_s^+ D^-$	$B^0 \to D^{*-} D_s^+$	Weighted Average
A_{det}	-0.0119 ± 0.0071	0.0229 ± 0.0096	—
A_{prod}	$-0.0164 \pm 0.0098 \pm 0.0020$	$0.0018 \pm 0.0145 \pm 0.0082$	-0.0107 ± 0.0086
$\Delta\varepsilon_{\text{tag}}^{\text{OS}}$	$0.0101 \pm 0.0090 \pm 0.0010$	$0.0026 \pm 0.0147 \pm 0.0120$	0.0081 ± 0.0094
Δp_0^{OS}	$0.0195 \pm 0.0082 \pm 0.0030$	$0.0270 \pm 0.0127 \pm 0.0380$	0.0217 ± 0.0150
Δp_1^{OS}	$0.2160 \pm 0.0760 \pm 0.0840$	$0.1530 \pm 0.1140 \pm 0.1560$	0.1970 ± 0.1200
$\varepsilon_{\text{tag}}^{\text{OS}}$	0.4051 ± 0.0027	0.4056 ± 0.0040	—
p_0^{OS}	0.3794 ± 0.0054	0.3831 ± 0.0090	—
$p_0^{\text{OS}} - \langle\eta^{\text{OS}}\rangle$	$0.0369 \pm 0.0054 \pm 0.0058$	$0.0362 \pm 0.0090 \pm 0.0070$	0.0367 ± 0.0075
p_1^{OS}	$0.9490 \pm 0.0500 \pm 0.0420$	$0.9590 \pm 0.0860 \pm 0.2900$	0.9510 ± 0.1120
$\langle\eta^{\text{OS}}\rangle$	0.3425	0.3470	—
$\Delta\varepsilon_{\text{tag}}^{\text{SS}}$	$-0.0046 \pm 0.0053 \pm 0.0030$	$-0.0086 \pm 0.0081 \pm 0.0080$	-0.0058 ± 0.0065
Δp_0^{SS}	$-0.0039 \pm 0.0065 \pm 0.0060$	$-0.0034 \pm 0.0096 \pm 0.0160$	-0.0037 ± 0.0110
Δp_1^{SS}	$-0.0500 \pm 0.0970 \pm 0.0390$	$0.0100 \pm 0.1400 \pm 0.0740$	-0.0300 ± 0.1000
$\varepsilon_{\text{tag}}^{\text{SS}}$	0.6762 ± 0.0026	0.700 ± 0.0040	—
p_0^{SS}	0.433 ± 0.0044	0.434 ± 0.0070	—
$p_0^{\text{SS}} - \langle\eta^{\text{SS}}\rangle$	$0.0062 \pm 0.0044 \pm 0.0080$	$0.0077 \pm 0.0070 \pm 0.0020$	0.0066 ± 0.0072
p_1^{SS}	$0.8770 \pm 0.0670 \pm 0.0720$	$1.0810 \pm 0.1020 \pm 0.0740$	0.9380 ± 0.0091
$\langle\eta^{\text{SS}}\rangle$	0.4265	0.4263	—

In addition, the tagging efficiencies $\varepsilon_{\text{tag}}^{\text{OS,SS}}$ are determined, but not used in the *CP* fit, as they can be calculated directly in the measurement. For the OS tagging combination all events tagged by OS taggers and for the SS tagging combination all events tagged by SS taggers are considered. This implies that some events that have an OS and SS tag are used in both combinations. The results of the calibration parameters are given with a statistical and systematic uncertainty. The central value as well as the systematic uncertainties of the weighted averages account for the statistical uncertainties of the OS and SS calibration parameters. The systematic uncertainties are assumed to be fully correlated between the two channels. The dominant systematic uncertainty is related to the sWeight determination. The impact on the calibration is tested by using a different procedure in the sWeight procedure. Instead of fitting the invariant B^0 mass distribution, a bi-dimensional fit to the invariant D_s^+ and D^- or $D^{*-} - D^0$ and D^- masses is performed. The resulting systematic uncertainties are of the same order of magnitude as the statistical uncertainties. Further possible sources of systematic uncertainties are investigated, e.g. the parameterisation of the decay-time-dependent efficiency, or fixing parameters such as the lifetime of the B^0 meson and Δm_d. These other possible sources for systematic uncertainties are found to be negligible.

In Tab. 7.5, the effective tagging efficiency ε_{eff} (see Eq. (4.9)) is listed for both runs. Both the effective tagging efficiency of the OS and the SS combination increase from Run 1 to Run 2. The SS tagger algorithms deliver better efficiencies for higher centre-of-mass energies, i.e. higher transverse momenta, therefore both the tagging efficiency and the effective tagging efficiency are increased in Run 2. In addition, Run 1 and Run 2 are optimised differently.

Table 7.4 – Flavour tagging parameters of the fit to the control channels $B^0 \to D_s^+ D^-$ and $B^0 \to D^{*-} D_s^+$ for the flavour tagging calibration for Run 2. For all calibration parameters considering the statistical uncertainties a weighted average of the results from the two channels is given. The given uncertainty for each weighted average is the combination of the statistical and systematic uncertainties. They are used in the fit to the decay-time distribution of the signal $B^0 \to D^{*\pm} D^{\mp}$ candidates.

Parameter	$B^0 \to D_s^+ D^-$	$B^0 \to D^{*-} D_s^+$	Weighted Average
A_{det}	0.0169 ± 0.0037	0.0169 ± 0.0065	—
A_{prod}	$0.0033 \pm 0.0051 \pm 0.0040$	$0.0054 \pm 0.0102 \pm 0.0071$	0.0037 ± 0.0049
$\Delta\varepsilon_{\text{tag}}^{\text{OS}}$	$-0.0044 \pm 0.0046 \pm 0.0060$	$0.0261 \pm 0.0103 \pm 0.0241$	0.0006 ± 0.0097
Δp_0^{OS}	$0.0107 \pm 0.0043 \pm 0.0040$	$0.0026 \pm 0.0091 \pm 0.0020$	0.0092 ± 0.0050
Δp_1^{OS}	$0.0210 \pm 0.0370 \pm 0.0060$	$0.0210 \pm 0.0820 \pm 0.0076$	0.0210 ± 0.0034
$\varepsilon_{\text{tag}}^{\text{OS}}$	0.4271 ± 0.0014	0.4257 ± 0.0028	—
p_0^{OS}	0.3688 ± 0.0029	0.3669 ± 0.0068	—
$p_0^{\text{OS}} - \langle\eta^{\text{OS}}\rangle$	$0.0320 \pm 0.0029 \pm 0.0030$	$0.0285 \pm 0.0068 \pm 0.0130$	0.0315 ± 0.0053
p_1^{OS}	$0.8570 \pm 0.0250 \pm 0.0150$	$0.9700 \pm 0.0590 \pm 0.0300$	0.8740 ± 0.0290
$\langle\eta^{\text{OS}}\rangle$	0.3368	0.3383	—
$\Delta\varepsilon_{\text{tag}}^{\text{SS}}$	$-0.0004 \pm 0.0015 \pm 0.0018$	$-0.0020 \pm 0.0033 \pm 0.0008$	-0.0007 ± 0.0021
Δp_0^{SS}	$0.0000 \pm 0.0030 \pm 0.0029$	$0.0048 \pm 0.0062 \pm 0.0009$	0.0009 ± 0.0037
Δp_1^{SS}	$-0.0870 \pm 0.0470 \pm 0.0160$	$-0.2070 \pm 0.0930 \pm 0.0317$	-0.1110 ± 0.0460
$\varepsilon_{\text{tag}}^{\text{SS}}$	0.8875 ± 0.0009	0.8738 ± 0.0019	—
p_0^{SS}	0.4291 ± 0.0021	0.4224 ± 0.0047	—
$p_0^{\text{SS}} - \langle\eta^{\text{SS}}\rangle$	$-0.0008 \pm 0.0021 \pm 0.0004$	$-0.0053 \pm 0.0047 \pm 0.0090$	-0.0015 ± 0.0026
p_1^{SS}	$0.8769 \pm 0.0320 \pm 0.0230$	$0.8751 \pm 0.0680 \pm 0.0652$	0.8770 ± 0.0420
$\langle\eta^{\text{SS}}\rangle$	0.4298	0.4277	—

Table 7.5 – Average mistag probability estimate $\langle\eta\rangle$, tagging efficiency ε_{tag} and effective tagging efficiency ε_{eff} for the signal samples.

Sample	Tagger	$\langle\eta\rangle$	ε_{tag} [%]	ε_{eff} [%]
Run 1	OS tagged	0.3507	36.8 ± 1.8	3.46 ± 0.30(stat) ± 0.47(syst)
	SS tagged	0.4257	67.7 ± 1.7	2.34 ± 0.20(stat) ± 0.37(syst)
	total		78.8 ± 1.5	5.61 ± 0.36(stat)
Run 2	OS tagged	0.3444	40.2 ± 0.9	3.94 ± 0.16(stat) ± 0.25(syst)
	SS tagged	0.4289	89.3 ± 0.6	2.92 ± 0.09(stat) ± 0.18(syst)
	total		93.4 ± 0.5	6.61 ± 0.19(stat)

7.3 Decay-time resolution

The theoretical decay-time distribution needs to be modified concerning effects caused by the detector. One effect is the decay-time resolution. The detector measures momenta and positions of vertices with finite precision. As the uncertainties on these quantities cause a deviation of the reconstructed decay time t' from the true decay time t, this effect must be considered in the parameterisation of the decay time. Therefore, the theoretical decay-time PDF is convolved with the resolution model $\mathcal{R}$, parameterising the probability density of the deviation $t - t'$. A possibility to investigate the decay-time resolution with detector data is given by reconstructing pseudo-signal candidates with tracks produced in the primary vertex. In these combinations negative reconstructed decay times are created, as a result of the limited detector resolution, which allows a determination of the decay-time resolution. This method enables a very accurate modelling of the decay-time resolution, which is especially essential in decay-time-dependent measurements of B_s^0 decays, as the oscillation frequency Δm_s of the B_s^0 system is about 35 times larger than Δm_d. However, the decay-time resolution has only a small effect to decay-time-dependent CP measurements in the B^0 system, so that the modelling of the resolution can also be derived from simulated signal decays. In Sec. 8.3.3, systematic uncertainty studies are performed in order to measure the influence of possible mis-modelling.

The decay-time resolution effects is modelled using three Gaussian functions with a common mean μ_t

$$\begin{aligned}\mathcal{R}(t-t') = {} & f_1 \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(t-t'-\mu_t)^2}{2\sigma_1^2}\right) + f_2 \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(t-t'-\mu_t)^2}{2\sigma_2^2}\right) \\ & + (1-f_1-f_2) \frac{1}{\sqrt{2\pi}\sigma_3} \exp\left(-\frac{(t-t'-\mu_t)^2}{2\sigma_3^2}\right),\end{aligned} \tag{7.12}$$

that are combined with the fractions f_1 and f_2. The widths of the Gaussian functions are σ_1 to σ_3. The Gaussian function with the width σ_1, which is the broadest of the three, corresponds to candidates that are associated to a wrong PV. An unbinned maximum likelihood fit to the difference of reconstructed and true decay times on simulated signal decays is performed separately for both runs. The results of the parameters are listed in Tab. 7.6 and the corresponding plots are shown in Fig. 7.3. The results from Run 2 are used for the measurement of the *CP* parameters, as the results between the runs are compatible, and because the larger number of candidates of the Run 2 dataset dominates the *CP* measurement.

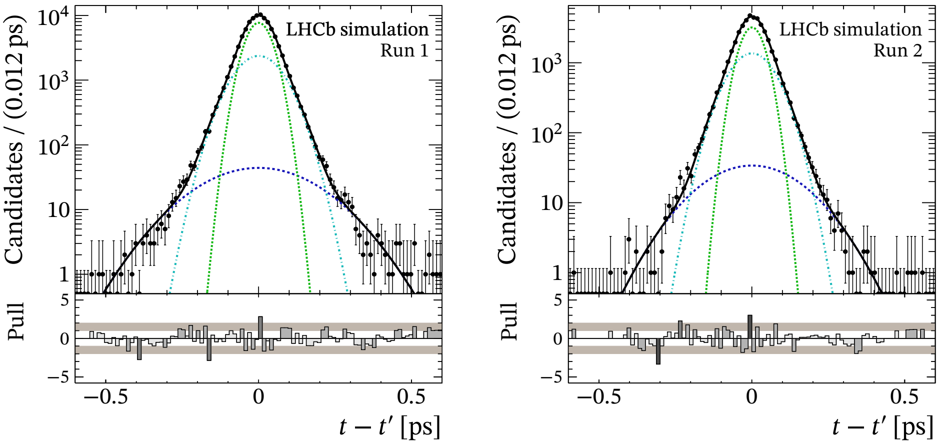

Figure 7.3 – Unbinned maximum likelihood fit to the difference of reconstructed t and true decay times t' for simulated signal decays. The left plot shows results for events simulated under Run 1 conditions, while the right plots shows the result for Run 2 simulations. The black solid line represents the projection of the full PDF. The green dotted line and the turquoise dashed-dotted line show two Gaussian functions, while the third Gaussian function, which is plotted in blue dashed, represents the wrong PV component.

Table 7.6 – Fit parameters of the decay-time resolution function for Run 1 and for Run 2 simulations.

Parameter	Fitted Value		Unit
	Run 1	Run 2	
μ_t	0.00018 ± 0.00017	0.00048 ± 0.00024	ps
f_1	0.017 ± 0.0033	0.023 ± 0.007	
f_2	0.348 ± 0.024	0.431 ± 0.044	
σ_1	0.161 ± 0.008	0.146 ± 0.011	ps
σ_2	0.070 ± 0.0018	0.0657 ± 0.0028	ps
σ_3	0.0387 ± 0.0060	0.0358 ± 0.0012	ps

7.4 Decay-time-dependent efficiency

Different experimental effects lead to a decay-time-dependent efficiency in the reconstruction of B^0 candidates. It is very significant at low decay times, as in the selection, e.g. in the trigger, requirements on the parameters that enable a separation of the decay vertex of the B^0 candidates from the PV, like IP, flight distances and decay time, are applied. At high decay times, reconstruction inefficiencies occur in the VELO [123, 124].

All deviations from the pure exponential behaviour are covered by a decay-time-dependent efficiency function. As these deviations are only experimental effects, the function can comprise any mathematical function as long as it fulfils smoothness in terms of an accurate interpolation of the efficiency and an efficient evaluation of integrals.

Cubic splines fulfil these requirements [125]. They are parameterised by k knots and their coefficients c_i. The model for the decay-time-dependent efficiency is defined by the sum over all base splines v_i

$$\epsilon(t) = \sum_{i=1}^{k} c_i v_i(t) \qquad \text{with} \qquad \sum_{i=1}^{k} v_i(t) = 1\,. \tag{7.13}$$

Here, B-splines [126] are used as base splines. The knots are the boundaries that define the valid region for a base spline. Thus, the density of knots defines the accuracy of interpolation. As only the relative values of the coefficients are important, the second to last coefficient is set to 1, which incorporates the total PDF normalisation and allows more comparability between fit results.

The set of knots are studied on simulated signal decays, where the B^0 lifetime is fixed to the generation value. The knots are placed at 0.3, 0.5, 2.7, 6.3 and 10.3 ps. A higher knot density is chosen towards lower decay times to enable a better description in the range where the efficiency changes the most. In the *CP* fit to detector data, the same knot positions are chosen and the coefficients float, while the B^0 lifetime is fixed to the recent world average [22]. In the top plots of Fig. 7.4, the decay-time distributions and the projection of the decay-time fits to simulated data are shown for both runs. In addition, the determined decay-time-dependent efficiency function is presented in Fig. 7.5. The coefficients of the cubic spline functions are shared among the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^-\pi^+\pi^+$ final states in the *CP*-parameter measurement, as their decay-time distributions are very similar (see bottom left plot in Fig. 7.4). Large differences are visible between the Run 1 and Run 2 distributions, especially at lower decay times (see bottom right plot in Fig. 7.4). Thus, the coefficients for the cubic spline functions are split for Run 1 and Run 2. In order to account for possible mis-modelling of the decay-time-dependent efficiency, systematic uncertainty studies are presented in Sec. 8.3.2, where a different set of knot positions is tested. In addition, it has been tested whether using different coefficients among all categories has an impact on the measurement of *CP* parameters.

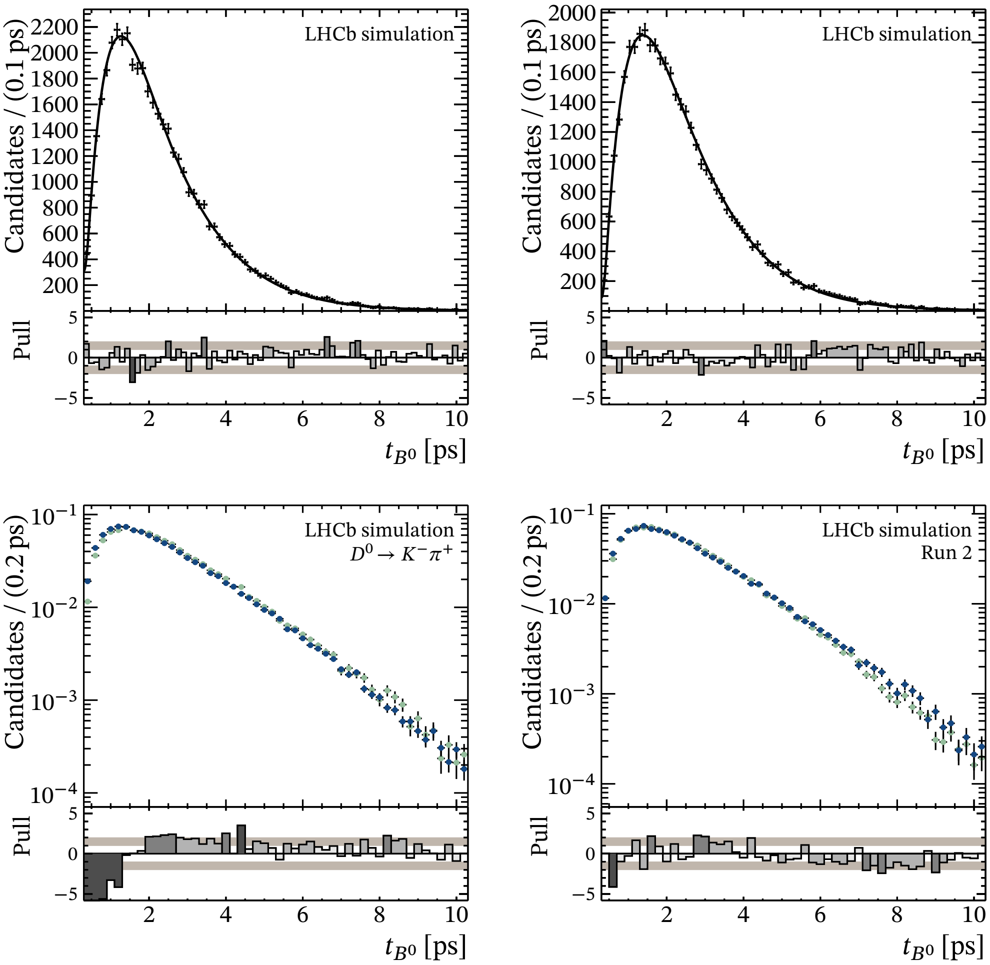

Figure 7.4 – At the top the unbinned maximum likelihood fits to the reconstructed decay-time distributions of simulated B^0 signal decays for (left) Run 1 and (right) Run 2 are presented. The decay-time-dependent efficiency is modelled using cubic splines. At the bottom the decay-time distributions of simulated signal decays (left) comparing (blue) Run 1 and (green) Run 2 for the $D^0 \rightarrow K^-\pi^+$ final state and (right) comparing the distributions between the final states (blue) $D^0 \rightarrow K^-\pi^+$ and (green) $D^0 \rightarrow K^-\pi^-\pi^+\pi^+$ for Run 2 are shown.

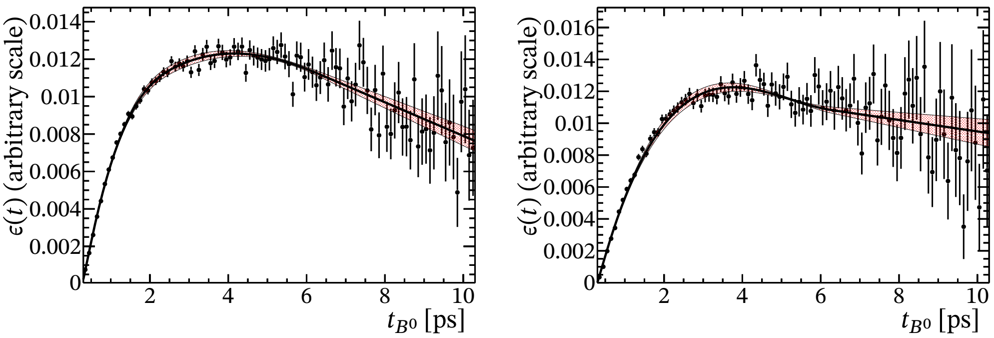

Figure 7.5 – The decay-time-dependent efficiency function is shown for (left) Run 1 and (right) Run 2. The red area represents the 1σ error band.

7.5 Instrumental asymmetries

As already mentioned in Sec. 7.1, the overall asymmetry A_{raw} (see Eq. (7.7)) is measured in the *CP* fit. In order to determine the *CP* parameter A_{D^*D}, A_{raw} needs to be corrected for the detection asymmetry (see Eq. (7.3)). This asymmetry arises due to detection, reconstruction and particle identification inefficiencies that are induced by the different interaction cross-sections with matter for positive and negative pions and kaons.

The $B^0 \to D^{*\pm}D^{\mp}$ final state is charge-symmetric, however, since the p and p_{T} distributions of kaons and pions in the $D^{*\pm}$ and $D^{\mp}$ decays are slightly different, and all instrumental efficiencies are momentum-dependent, so the cancellation is not expected to be complete. A comparison between the momentum distributions between the $D^{*\pm}$ and $D^{\mp}$ final state particles is shown in Fig. 7.8. It is notable that only small differences between the p distributions of the two kaons exist. However, larger differences are present between the slow pion and the pion with the lower p_{T} of the $D^{\mp}$ meson decay.

7.5.1 Strategy

The detection asymmetry for the $B^0 \to D^{*\pm}D^{\mp}$ decay can be approximated by the difference of the separated detection efficiencies for the $D^{*\pm}$ and $D^{\mp}$ mesons

$$\begin{aligned} A_{\text{det}}^{D^*D} &= \frac{\epsilon(D^{*-}D^{+}) - \epsilon(D^{*+}D^{-})}{\epsilon(D^{*-}D^{+}) + \epsilon(D^{*+}D^{-})} = \frac{\epsilon(D^{*-})\epsilon(D^{+}) - \epsilon(D^{*+})\epsilon(D^{-})}{\epsilon(D^{*-})\epsilon(D^{+}) + \epsilon(D^{*+})\epsilon(D^{-})} \\ &\simeq \frac{\epsilon(D^{*-}) - \epsilon(D^{*+})}{\epsilon(D^{*-}) + \epsilon(D^{*+})} - \frac{\epsilon(D^{-}) - \epsilon(D^{+})}{\epsilon(D^{-}) + \epsilon(D^{+})} = A_{\text{det}}^{D^*} - A_{\text{det}}^{D} \,. \end{aligned} \tag{7.14}$$

The approximation involves a non trivial conversion, where the quadratic term in the difference of efficiencies in the denominator is neglected. The detection asymmetry

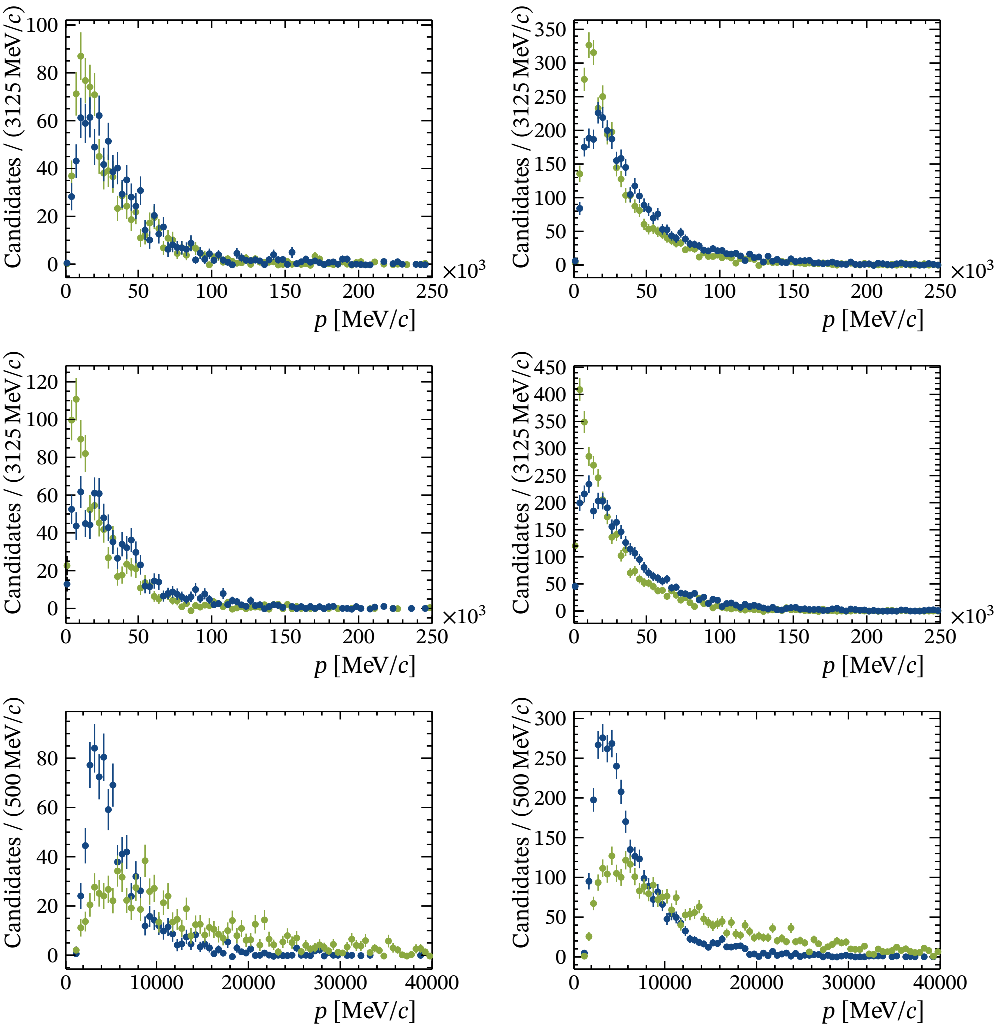

Figure 7.6 – Comparison of momentum distributions between corresponding $D^{\mp}$ and $D^{*\pm}$ daughter particles. The left plots show Run 1 data, while the right ones show Run 2 data. At the top, the distributions of the kaons are shown, while in the middle the pion coming from the D^0 decay and the pion with the higher p_{T} of the $D^{\mp}$ decay are compared. At the bottom, the distributions of the slow pion is compared to the pion with the lower p_{T} of the $D^{\mp}$ decay. The blue data points correspond to the final state particles of the $D^{\mp}$ meson and green to the particles originating from the $D^{*\pm}$ meson. In all plots, the background is subtracted using sWeights.

is measured separately for the $D^{*\pm}$ and $D^{\mp}$ mesons utilising a recorded data sample of $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ decays, where the $D^{\mp}$ mesons are prompt, hence, directly created in the PV. As the detection asymmetry is (transverse) momentum dependent, the kinematic properties of the final state particles of the prompt $D^{\pm}$ mesons are adjusted to match the kinematic properties of the signal channel. After this so-called re-weighting, the detection asymmetry is determined from the measured number of prompt D^{+} and D^{-} when being re-weighted according to the different kinematic distributions. Asymmetries induced by particle identification variables in the BDT training, as well as production asymmetries of the prompt $D^{\pm}$ mesons, are considered in the systematic uncertainty on A_{det} discussed in Sec. 8.3.6.

7.5.2 Measurement of the detection asymmetry

In order to measure A_{D^*D}, it is necessary to determine a possible detection asymmetry A_{det} between $D^{*-}D^{+}$ and $D^{*+}D^{-}$ independently, as A_{det} and A_{D^*D} are fully correlated. This determination of A_{det} is performed with prompt decays of $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$. The processing of a sample of this prompt decay is very time consuming due to the large amount of candidates. As the decision to measure A_{D^*D} was taken at a late stage in the analysis process, A_{det} is measured only on the already available $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$ sample from 2016. The 2016 samples comprises approximately six million events in total. To this data sample the same requirements on particle identification variables are applied that are used in the stripping line for the signal decay (see Tab. 5.1). In order to reject secondary $D^{\pm}$ mesons, a cut on the IPχ^2 of smaller than ten is applied. The available sample, which comprises both magnet polarities, is divided into four subsamples, two subsamples are used for the final state $D^0 \to K^-\pi^+$ and two for the $D^0 \to K^-\pi^-\pi^+\pi^+$ final state.

As a reminder the final state particles in case D^0 is reconstructed as $K^-\pi^+$ are $(K^{\pm}\pi^{\mp}\pi^{\mp})_{D^{*\pm}}(K^{\mp}\pi^{\pm}\pi^{+})_{D^{\mp}}$. For the measurement of the detection asymmetry in this case one of the prompt $D^{\pm}$ subsample is re-weighted to match the kinematic properties of the final state particles in the signal $D^{*\pm}$ meson and one further subsample is re-weighted to match the kinematic properties of the final state particles of the signal $D^{\mp}$ meson. The same procedure is applied for the final state in which D^0 is reconstructed as $K^-\pi^-\pi^+\pi^+$, but the re-weighting is only performed considering the slow pion, the kaon and one pion from the D^0 decay. The remaining $\pi^+\pi^-$ pair does not contribute to the detection asymmetry as it appears in the same way for both D^{*+} and D^{*-}. Moreover, the p, p_{T} and pseudorapidity distributions of the two pions are in very good agreement.

As the data samples of the prompt $D^{\pm}$ decays contain signal and background candidates, an extended maximum likelihood fit to the invariant $K^{\mp}\pi^{\pm}\pi^{+}$ mass is performed on each subsample to determine sWeights, which are then used in the re-weighting. The prompt $D^{\pm}$ signal model is parametrised by two Crystal Ball functions with a common mean and tails to both sides. The parameters $n_{1,2}$ are fixed to 100, while $\alpha_{1,2}$ and σ_1 are fixed to values obtained by the least squares method χ^2 [127], where a fit to binned data is performed. The background is modelled by an exponential function. One of the four samples, including the PDF, is shown in Fig. 7.7 and the corresponding fit results are listed in Tab. 7.7. The other three results are in agreement with the presented one.

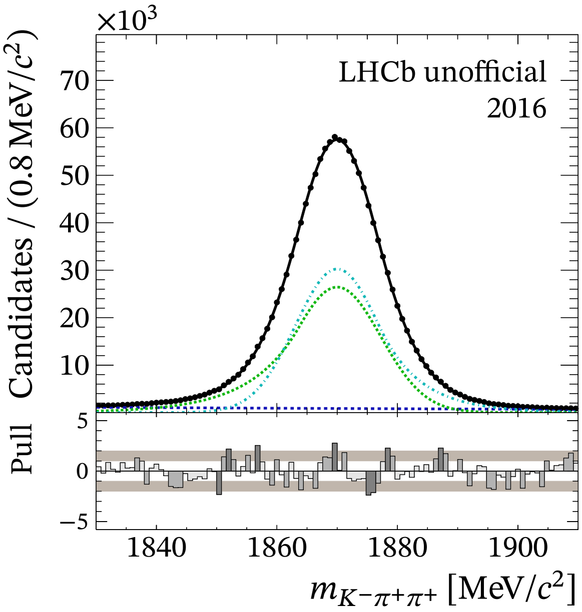

Figure 7.7 – Results of one of the four fits to the invariant $K^{-}\pi^{+}\pi^{+}$ mass distribution of prompt $D^{\pm}$ data recorded in 2016. The signal is modelled by three Crystal Ball functions, while for the background an exponential function is used.

The `GBreweighter` tool [128] is used to transform the momentum and transverse momentum distributions of the final state particles from prompt $D^{\pm}$ decays, where sWeights are applied, into those from the $D^{*\pm}$ (or $D^{\mp}$) meson decays of the $B^0 \to D^{*\pm}D^{\mp}$ signal decays. The distributions for the latter are taken from simulated signal events that passed the full selection. This is feasible, as the p and p_{T} observables are very well simulated and a larger amount of events is available compared to recorded data.

The individual asymmetries for the $D^{*\pm}$ and $D^{\mp}$ mesons, $A_{\mathrm{det}}^{D^*}$ and A_{det}^{D}, respectively, are calculated from the yields determined in the fits to the mass distributions of the prompt $D^{\pm}$ samples, where the samples are split by the charge of the $D^{\pm}$ meson. In

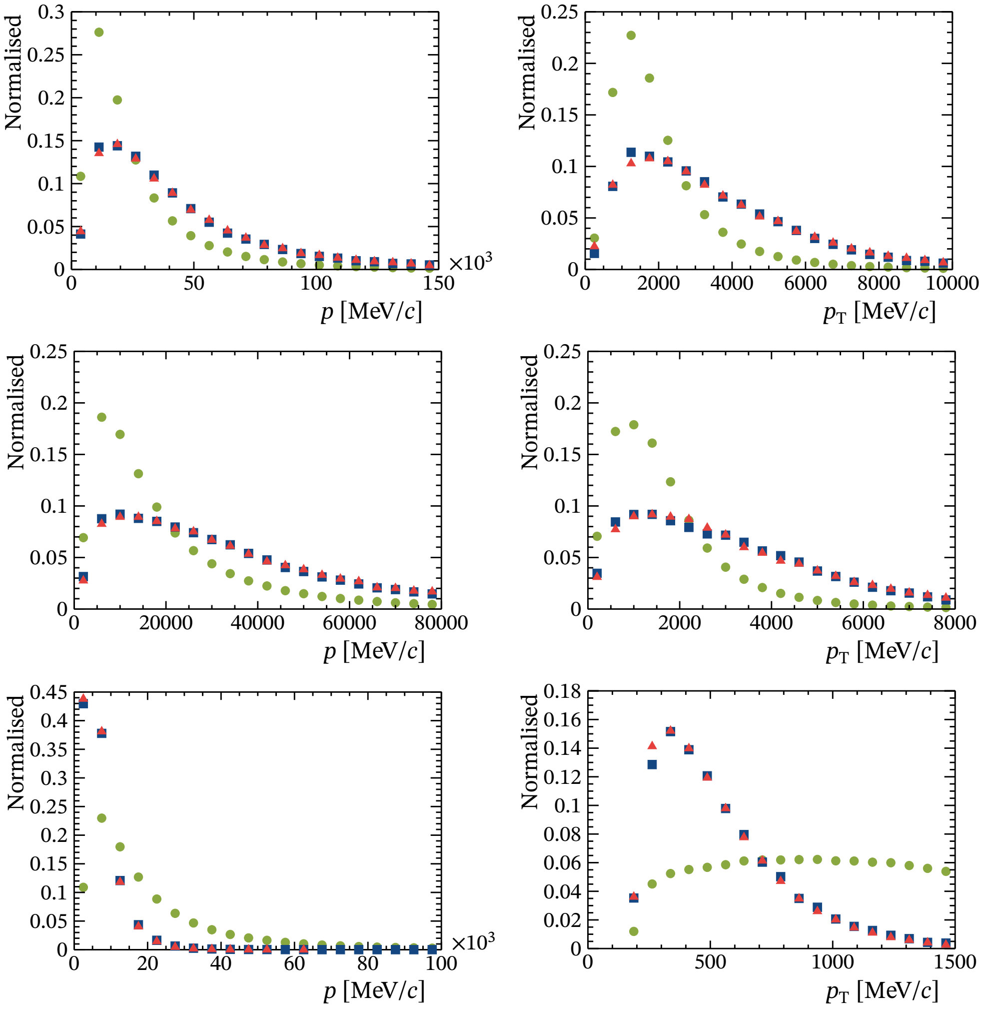

Figure 7.8 – Distributions of the (left) momentum and (right) transverse momentum for the final state particles of the (green circle) prompt $D^{\pm}$, (red triangle) signal simulation and (blue square) re-weighted prompt $D^{\pm}$ mesons. The top plots show the kaon distributions, the middle plots the ones for the pion and the bottom plots contain distributions for the slow pion.

Table 7.7 – Result of one of the four sPlot fits to the invariant $K^-\pi^+\pi^+$ mass distribution of prompt $D^\pm$ data of the year 2016 to determine sWeights. The parameters without uncertainties are fixed in the fit.

Parameter	Values
μ_D (MeV/c^2)	1870.000 ± 0.009
σ_1 (MeV/c^2)	7.24
σ_2 (MeV/c^2)	6.932 ± 0.013
α_1	0.895
α_2	-1.041
$n_{1,2}$	100
g	0.486 ± 0.003
$\beta \times 10^3$ (c^2/MeV)	-6.34 ± 0.27
N_D	1371140 ± 1322
N_{BKG}	86149 ± 680

these mass fits, the weights from the `GBreweighter` are applied. Two exemplary results, the mass fits for D^+ and D^- in the $D^0 \to K^-\pi^+$ are given in Tab. 7.8. The corresponding plots are shown in Fig. 7.9. The fit model is the same as in the previous cases using a shared floating width of the Crystal Ball functions among the two charges. The resulting yields for all $D^{(*)\pm}$ mesons and the calculated detection asymmetries are summarised in Tab. 7.9 and Tab. 7.10. Given the result listed in Tab. 7.10, it can be concluded that the difference between the detection asymmetries $A_{\text{det}}^{D^*}$ and A_{det}^{D}, is consistent with zero with a statistical uncertainty of 0.4%, or less. For Run 1, different values for the detection asymmetries are expected that are mainly related to the differences in the (transverse) momentum distributions of the final state particles. However, the differences between the kinematic distributions of $D^\mp$ and $D^{*\pm}$ are smaller in Run 1 compared to Run 2 (see Fig. 7.8). As no suitable samples for Run 1 are available to perform the same studies for A_{det}, it is assumed that the detection asymmetry is zero in Run 1. As a conservative assumption two times the statistical uncertainty of the Run 2 result is taken as uncertainty for the Run 2 asymmetry.

7.6 Determination of *CP* parameters

In order to determine the *CP* parameters, the ingredients discussed in Sec. 7.2 to Sec. 7.5 are utilised. In the decay-time fit both the description of the decay-time resolution (see Tab. 7.6) and the average mistag probability estimates $\langle\eta^{\text{OS}}\rangle$ and $\langle\eta^{\text{SS}}\rangle$ are fixed. For the latter, the results calculated in each signal sample are used. The val-

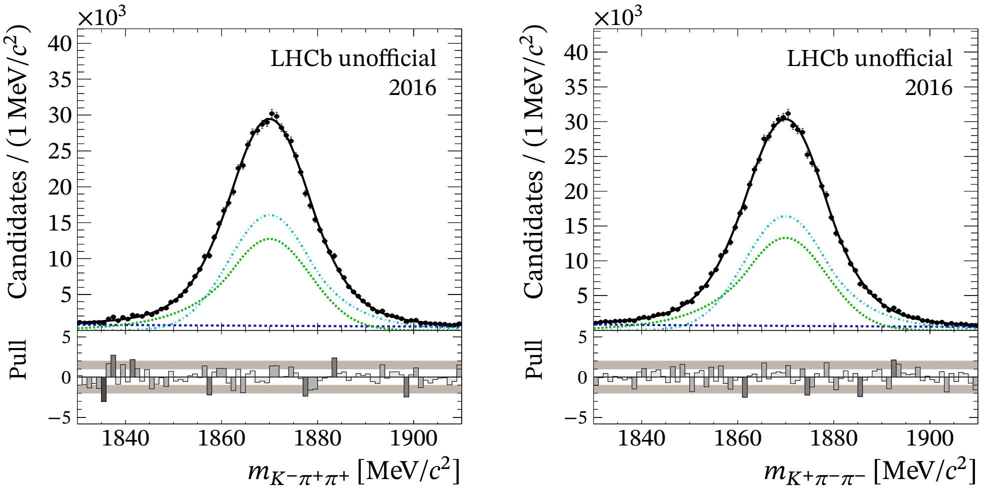

Figure 7.9 – Example for the mass fits to the invariant $K^{\mp}\pi^{\pm}\pi^{\pm}$ distribution split by the charge of the prompt $D^{\pm}$ meson to extract the yields for (left) D^- and (right) D^+. The signal is modelled by three Crystal Ball functions and the background by an exponential.

Table 7.8 – Fit results of two exemplary mass fits to the invariant $K^-\pi^+\pi^+$ mass of prompt $D^{\pm}$ data split by the charge of the $D^{\pm}$ meson. The values without uncertainties are fixed in the fits.

Parameter	Values	
	D^- sample	D^+ sample
μ_D (MeV/c^2)	1870.00 ± 0.05	1870.00 ± 0.05
σ (MeV/c^2)	8.507 ± 0.056	8.484 ± 0.056
α_1	0.895	0.895
α_2	−1.041	−1.041
$n_{1,2}$	100	100
g	0.455 ± 0.021	0.450 ± 0.021
$\beta \times 10^3$ (c^2/MeV)	-6.26 ± 1.45	-5.50 ± 1.46
N_D	688343 ± 3499	665409 ± 3417
N_{BKG}	52203 ± 2258	51285 ± 2189

Table 7.9 – Yields of prompt D^+ and D^- mesons re-weighted to match the kinematics of $D^{\pm}$ and $D^{*\pm}$ from signal $B^0 \to D^{*\pm}D^{\mp}$ decays in the $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^-\pi^+\pi^+$ decay modes.

D^0 final state	Yield	$D^{\pm}$ kinematic	$D^{*\pm}$ kinematic
$K^-\pi^+$	D^-	$688\,375 \pm 1833$	$688\,343 \pm 3499$
	D^+	$666\,972 \pm 1594$	$665\,409 \pm 3417$
$K^-\pi^-\pi^+\pi^+$	D^-	$687\,431 \pm 1689$	$689\,045 \pm 2105$
	D^+	$668\,749 \pm 1132$	$669\,160 \pm 2067$

Table 7.10 – Detection asymmetries of $D^{\pm}$, $D^{*\pm}$ and $B^0 \to D^{*\pm}D^{\mp}$ decays obtained from prompt $D^{\pm}$ decays.

D^0 final state	$A_{\text{det}}^{D^*}$	A_{det}^{D}	$A_{\text{det}}^{D^*D}$
$K^-\pi^+$	0.0169 ± 0.0036	0.0158 ± 0.0018	0.0011 ± 0.0040
$K^-\pi^-\pi^+\pi^+$	0.0146 ± 0.0022	0.0138 ± 0.0015	0.0009 ± 0.0026

ues for the final state $D^0 \to K^-\pi^+$ ($D^0 \to K^-\pi^-\pi^+\pi^+$) are $\langle\eta^{\text{OS}}\rangle = 0.3527\,(0.3507)$ and $\langle\eta^{\text{SS}}\rangle = 0.4323\,(0.4257)$ for Run 1, while for Run 2 $\langle\eta^{\text{OS}}\rangle = 0.3452\,(0.3444)$ and $\langle\eta^{\text{SS}}\rangle = 0.4266\,(0.4289)$ are used. The lifetime difference $\Delta\Gamma_d$ is fixed to zero, while the B^0 lifetime τ_{B^0} and the mass difference Δm_d are fixed to the world averages of $\tau_{B^0} = (1.520 \pm 0.004)\,\text{ps}$ and $\Delta m_d = (0.050\,65 \pm 0.001\,90)\,\text{ps}^{-1}$ [129]. All further parameters are introduced in the *CP* fit via Gaussian constraints (see Eq. (4.4)), like the flavour tagging calibration parameters obtained from the analysis of the control channels $B^0 \to D_s^+D^-$ and $B^0 \to D^{*-}D_s^+$ given in the right column in Tab. 7.3 and Tab. 7.4.

Another necessary input to the *CP* fit is the production asymmetry A_{prod}. An LHCb measurement for A_{prod} exists for Run 1 [130], but it needs to be adjusted in order to be used in $B^0 \to D^{*\pm}D^{\mp}$. Hence, a weighting procedure is performed in bins of transverse momentum and pseudorapidity of the B^0 meson, in order to match the kinematics of the signal dataset. This procedure results in $A_{\text{prod}} = -0.0014 \pm 0.0068$. The value is in agreement with the Run 1 value for A_{prod} obtained in the analysis of the control channels (see Sec. 7.2). For Run 2, no LHCb measurement exists that could be used in the same way. For consistency, the values of the production asymmetries obtained in the fit to the control channels for both Run 1 and Run 2 are used in the *CP* measurement of $B^0 \to D^{*\pm}D^{\mp}$, introduced with Gaussian constraints.

The parameter A_{raw} (see Eq. (7.6) and Eq. (7.7)) is expected to be different between the final states and the data taking periods. Therefore, four independent A_{raw} parameters are introduced that all can be adjusted in the fit, as only small correlations to other parameters are present.

Results

The *CP* observables resulting from the fit to the decay time are

$$\begin{aligned} S_{D^*D} &= -0.861 \pm 0.077\,, & \Delta S_{D^*D} &= 0.019 \pm 0.075\,, \\ C_{D^*D} &= -0.059 \pm 0.092\,, & \Delta C_{D^*D} &= -0.031 \pm 0.092\,, \\ A_{D^*D} &= 0.008 \pm 0.014\,, \end{aligned} \tag{7.15}$$

where the given uncertainties for the parameters S_{D^*D}, ΔS_{D^*D}, C_{D^*D} and ΔC_{D^*D} are estimated by the maximum likelihood fit representing the statistical uncertainties based on the number of signal candidates and the external parameters that are deployed as Gaussian constraints. The correlations among S_{D^*D}, ΔS_{D^*D}, C_{D^*D} and ΔC_{D^*D} are reported in Eq. (7.16).

$$\rho_{CP} = \left(\begin{array}{c|cccc} & S_{D^*D} & \Delta S_{D^*D} & C_{D^*D} & \Delta C_{D^*D} \\ \hline S_{D^*D} & 1 & 0.07 & 0.44 & 0.05 \\ \Delta S_{D^*D} & & 1 & 0.04 & 0.46 \\ C_{D^*D} & & & 1 & 0.04 \\ \Delta C_{D^*D} & & & & 1 \end{array}\right). \tag{7.16}$$

The fit results for A_{raw}, which represents the sum of CP and detection asymmetries (see Eq. (7.7)), for the two D^0 final states are listed in Tab. 7.11. The detection asymmetries from Tab. 7.10 are subtracted to each A_{raw} value. Then, the weighted mean of the four values of A_{D^*D} is calculated, using as weights the statistic uncertainties of A_{raw}. The statistical and systematic uncertainties of $A_{\text{det}}^{D^*D}$ will be considered in the systematic uncertainty of A_{D^*D}, described in Sec. 8.3.6. The four A_{raw} parameters have correlations that are less than one per mille with all other fit parameters, the only exception is the correlation to ΔS_{D^*D}, which reaches from a few per mille up to 1%. The fit results for the spline coefficients are listed in Tab. 7.12. While in Fig. 7.10 the decay-time distribution of the full selected dataset is shown with the fit PDF overlaid, Fig. 7.11 displays the decay-time-dependent asymmetry for both the $D^{*+}D^-$ and $D^{*-}D^+$ charge configurations.

Table 7.11 – Results of the CP fit of A_{raw} for Run 1 and Run 2 and the final states $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^-\pi^+\pi^+$.

Parameter	Value	
	Run 1	Run 2
$A_{\text{raw}}^{K\pi\pi\pi}$	0.0791 ± 0.0591	0.0093 ± 0.0298
$A_{\text{raw}}^{K\pi}$	0.0177 ± 0.0370	-0.0286 ± 0.0186

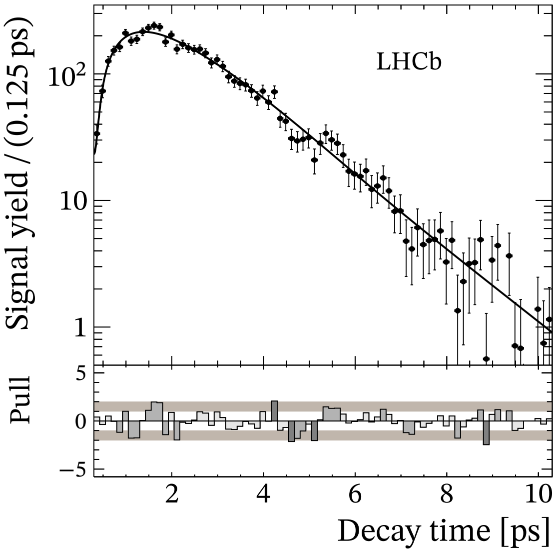

Figure 7.10 – Reconstructed B^0 decay time for the total dataset. The fitted PDF is overlaid in black.

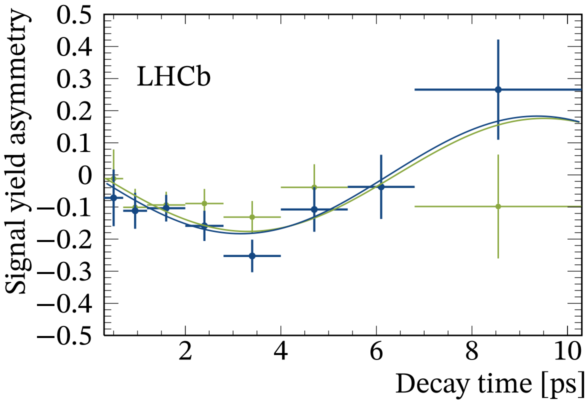

Figure 7.11 – Signal yield asymmetry of reconstructed B^0 or $\bar{B}^0$ mesons decaying into the $D^{*+}D^-$ (green) or into the $D^{*-}D^+$ (blue) charge configuration. This decay-time-dependent asymmetry is defined as $(N_{\bar{B}^0}(t) - N_{B^0}(t))/(N_{\bar{B}^0}(t) + N_{B^0}(t))$, where $N_{B^0}(t)$ $(N_{\bar{B}^0}(t))$ corresponds to the number of reconstructed signal decays at the decay time t for a tagged B^0 $(\bar{B}^0)$.

Table 7.12 – Results of the *CP* fit for the coefficients of the cubic spline parameterisation of the decay-time-dependent efficiency for Run 1 and Run 2.

Parameter	Fitted Value	
	Run 1	Run 2
c_1	0.032 ± 0.019	0.039 ± 0.013
c_2	0.62 ± 0.13	0.74 ± 0.12
c_3	0.78 ± 0.27	1.51 ± 0.33
c_5	0.43 ± 0.31	1.24 ± 0.40

8 Fit Validation and Studies of systematic Effects

Apart from statistical uncertainties that are estimated by the maximum likelihood fit, it is important to take possible uncertainties arising from methods and strategies into account. These so-called systematic uncertainties and the validity of the decay-time fit are investigated using pseudo-experiments, often called toy MCs. Pseudo-experiments are produced using a hit-and-miss procedure following a PDF, which reduces computing resources compared to full physics simulation. This facilitates producing a large number of toy MCs in different scenarios easily to probe the fitting procedure and the impact of different fit configurations. Generating and fitting toy MCs, in order to estimate systematic uncertainties, is called a toy study. In this analysis, toy studies are performed by generating thousands of toy MCs with different parameterisations and assumptions, which are described in Sec. 8.1. Based on these configurations, cross-checks of the fitter are performed (see Sec. 8.2) and the estimation of systematic uncertainties is carried out, as described in Sec. 8.3. In Sec. 8.4 the total systematic uncertainties are given.

8.1 Model for fit validation and estimation of systematic uncertainties

In the toy studies, the same fit strategy as in the *CP* analysis is utilised, namely the background is subtracted with sWeights in the fit to the B^0 decay time. This strategy is mimicked by generating toy MC for each component and every single observable that is needed in the *CP* PDF, namely the reconstructed B^0 mass, the reconstructed decay time, the tag decision and the corresponding mistag probability estimates. Since these observations are almost uncorrelated, they can be generated independently and the individual PDFs can be multiplied with each other. The distributions are generated separately for the runs and D^0 final states $K^-\pi^+$ and $K^-\pi^-\pi^+\pi^+$. The models and assumptions for the reconstructed mass and decay time are described in the following.

Reconstructed B^0 mass model

For the modelling of the invariant B^0 mass the same PDF as described in Ch. 6 is used. The parameters are fixed to the same values as in the fit to data and the same number of events for the individual components is generated as observed in data (see Tab. 6.4).

Decay-time model

The signal component is modelled by the nominal parameterisation used in the *CP* fit as described in Sec. 7.1. The coefficients for the time-dependent efficiency are fixed to the values obtained in the fit to recorded data (see Tab. 7.12). All parameters that are employed with Gaussian constraints (see Sec. 4.1) in the *CP* fit are set to values randomly generated according to the constraint [131]. For the *CP* parameter S, the world average of $S_{D^*D} = -0.73$ is used, while the other *CP* parameters (C_{D^*D}, ΔS_{D^*D} and ΔC_{D^*D}) are set to zero, as their measurements do not differ significantly from zero.

The individual physical background components are generated with the same model as the signal in the nominal *CP* fit. For the components $B_s^0 \to D^{*\pm}D^{\mp}$ and $B_s^0 \to D^{*+}D^{*-}$, the *CP* parameters are set to zero, as no *CP* violation has been measured in these decays. In the decay of $B^0 \to D^{*-}D_s^+$, the parameter ΔC is set to 1, as this decay is flavour specific. For the component $B^0 \to D^{*+}D^{*-}$, the values are set to $S = -0.71$ and $C = -0.01$, which state the current world averages [132–134]. Besides the *CP* parameters, the lifetime and oscillation frequency are set according to $\tau_{B_s^0}$ and Δm_s for the physical background components arising from B_s^0 decays.

Mistag probability estimates

In order to generate per-event OS and SS mistag probability estimates, histogram PDFs are utilised. The histogram PDFs are obtained from sWeighted η^{OS} and η^{SS} distributions for the signal component. The background components are treated in the same way, except for the background from B_s^0 decays, for which the SS tagging is assumed to provide random decisions. For combinatorial background, both OS and SS tagging is assumed to provide random decisions.

8.2 Cross-checks

Various cross-checks are performed to verify the correctness of the *CP* fitter and to confirm consistency of the measurement. In particular, a toy study (see Sec. 8.2.1) and a fit to simulated data are utilised (see Sec. 8.2.2), subsets of the data are fit (see Sec. 8.2.3) and the statistical uncertainties are validated (see Sec. 8.2.4).

8.2.1 Fit validation using pseudo-experiments

It is investigated, whether the *CP* fit itself creates a bias by performing a toy study using the nominal fit model as described in Sec. 8.1 for generating and fitting. For each fit parameter and each of the N generation and fit steps the difference of the observed X_{toy} and expected value X_{exp} is calculated, relative to the uncertainty $\sigma_{X_{\text{toy}}}$ of the parameter, resulting in the so-called pull distribution

$$\text{Pull}(X) = \frac{X_{\text{toy}} - X_{\text{exp}}}{\sigma_{X_{\text{toy}}}} \,. \tag{8.1}$$

In case of no bias in a measurement of a parameter X, the pull distribution has a mean value $\mu(\text{Pull})$ that is in agreement with zero. In addition, if the fit uncertainty of the parameter is correctly estimated, the width of the pull distribution σ_{Pull} is compatible with one. Hence, if deviations of the mean from zero are present, the fit is biased, while deviations of the widths from one reflect under- ($\sigma_{\text{Pull}} > 1$) or over-estimation ($\sigma_{\text{Pull}} < 1$) of the uncertainty.

In Tab. 8.1, the means and widths from the pull distributions of the toy study with 2400 pseudo-datasets are listed. No significant deviations from zero are observed for the means of the pull distributions. Thus, no systematic uncertainties need to be assigned. However, the toy study results that the uncertainties of S_{D^*D} and ΔS_{D^*D} are slightly overestimated by the fit, as the widths of the pull distributions are significantly below one. In order to verify the correctness of the fit uncertainty estimates, a bootstrapping study is performed, which is presented in Sec. 8.2.4.

8.2.2 Fit validation using simulated signal events

An additional validation of the fitting procedure and the fitter implementation is carried out by fitting simulated signal events. The same selection chain and the same fit model as in data are applied. The total number of simulated events corresponds to about 35 times the statistics of signal candidates in recorded data. The fit is performed as in the nominal fit to data, i.e. simultaneously in the four samples according to Run 1 and Run 2 for the final states $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^-\pi^+\pi^+$. The results of the *CP* parameters are consistent with the values used in the generation of the simulation within less than 2σ, as can be seen in Tab. 8.2.

Table 8.1 – Results of the toy study for fit validation. The mean and widths of the pull distributions from the *CP* and asymmetry parameters are given.

Parameter	Mean	Sigma
S_{D^*D}	-0.036 ± 0.021	0.955 ± 0.015
ΔS_{D^*D}	0.022 ± 0.020	0.929 ± 0.015
C_{D^*D}	0.000 ± 0.021	0.982 ± 0.016
ΔC_{D^*D}	0.010 ± 0.022	0.997 ± 0.016
$A_{\mathrm{raw}}^{K\pi,\mathrm{Run1}}$	0.024 ± 0.021	0.997 ± 0.016
$A_{\mathrm{raw}}^{K\pi,\mathrm{Run2}}$	0.016 ± 0.021	0.992 ± 0.016
$A_{\mathrm{raw}}^{K\pi\pi\pi,\mathrm{Run1}}$	0.026 ± 0.022	0.992 ± 0.016
$A_{\mathrm{raw}}^{K\pi\pi\pi,\mathrm{Run2}}$	-0.024 ± 0.021	0.977 ± 0.016

Table 8.2 – Results of the *CP* fit to the reconstructed decay time of simulated signal events.

Parameter	Value in Generation	Fitted Value
S_{D^*D}	0.7298	0.747 ± 0.011
$\Delta\, S_{D^*D}$	0.0	-0.009 ± 0.011
C_{D^*D}	0.0	0.012 ± 0.013
$\Delta\, C_{D^*D}$	0.0	0.012 ± 0.013

8.2.3 Cross-checks on subsets of the data

It is cross-checked whether results obtained on subsets of the data are compatible. Hence, the dataset is split into several sub-samples and the *CP* fit to the decay time is performed on each subset. The data is split by the following criteria:

- By D^0 final states ($D^0 \to K^-\pi^+$, $D^0 \to K^-\pi^-\pi^+\pi^+$)
- By tagging algorithms (OS, SS)
- By magnet polarities (magnet up, magnet down)
- By data taking periods (Run 1, Run 2)

The parameters determined when splitting by final state and magnet polarity are compatible within less than 1σ. When splitting by tagging algorithms the compatibility is less than 2σ. Issues with non-converging fits have been found in Run 1 data only, as the number of candidates is very low. However, the Run 2 results show a very good agreement with the results of the full dataset.

8.2.4 Estimation of statistical uncertainties

In order to validate the statistical uncertainties on the *CP* parameters obtained from the *CP* fit to recorded data, a bootstrapping procedure is utilised (see Sec. 4.4). Since the widths of the pull distributions (see Sec. 8.2.1) are compatible with one within 2σ and because the widths of the residual distributions are in agreement with the means of the obtained statistical uncertainty estimations within 2σ, it is valid to assume that the reported statistical uncertainty estimations are accurate.

8.3 Estimation of systematic uncertainties

In the toy studies, discussed in this section, the toy MC is generated with the PDF as described in Sec. 8.1 with the same number of candidates as observed in the fit to data and then fitted twice. In the first fit the PDF is unchanged. In the second fit, the PDF is modified, depending on the investigated source of the systematic uncertainty. By subtracting these two fit results, e.g. for the parameter p, $p_{\mathrm{nom}} - p_{\mathrm{mod}}$, the impact of the systematic change can be extracted, by cancelling fit biases that can appear in the fit procedure itself, as reported in Sec. 8.2.1.

The systematic uncertainty s for such a toy study is then obtained by calculating $s = \sqrt{\mu^2 + \mathrm{RMS}^2}$, whereby μ is the mean of the distribution given by $p_{\mathrm{nom}} - p_{\mathrm{mod}}$ and RMS is the root mean square of the distribution.

8.3.1 Mass model

Signal models The nominal model of the signal component comprises three Crystal Ball functions for the $D^0 \to K^-\pi^+$ final state and two Crystal Ball functions for the $D^0 \to K^-\pi^-\pi^+\pi^+$ final state. Toy studies are performed using two alternative models. In the first, for both decay modes two and in the second three Crystal Ball functions are used. An additional study is utilised by generating the toys using three Crystal Ball functions and fitting with two Crystal Ball functions for both final states. The results are given in in Tab. 8.3.

Mass resolution In the nominal fit to the invariant $D^{*\pm}D^{\mp}$ mass, the widths of the Crystal Ball and Gaussian components are fixed to the values obtained from simulated candidates. To account for possible mass resolution differences between recorded data and simulated candidates, the mass fit to the invariant $D^{*\pm}D^{\mp}$ mass distribution is

Table 8.3 – Systematic uncertainties on the *CP* parameters due to alternative models for the signal shapes.

Parameter	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
Signal mass model	0.011	0.003	0.014	0.006

Parameter	$A_{\text{raw}}^{K\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi,\text{Run2}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$
Signal mass model	0.0021	0.0016	0.0025	0.0024

repeated with a free scaling factor. This factor is multiplied to all parameters corresponding to the widths of the models of the signal as well as the $B_s^0 \to D^{*\pm}D^{\mp}$ and $B^0 \to D^{*-}D_s^+$ components. In Tab. 8.4, the results of the modified mass fit, which are used as inputs for the toy study, are shown. The scaling factor is consistent with unity within less than 2σ and the number of signal and background candidates is consistent with the nominal values given in Tab. 6.4 within less than 1σ. In the *CP* fit to recorded data, where the sWeights obtained with the modified mass model are used, a variation of less than 10% compared to the statistical uncertainties of the *CP* parameters (see Tab. 8.5) is observed.

Table 8.4 – Relevant results of the mass fit to data in case the mass models of signal, $B_s^0 \to D^{*\pm}D^{\mp}$ and $B^0 \to D^{*-}D_s^+$ backgrounds include a common, free scaling factor for their width parameters.

	$D^0 \to K^-\pi^+$		$D^0 \to K^-\pi^-\pi^+\pi^+$	
Parameter	Run 1	Run 2	Run 1	Run 2
Scaling Factor	1.03 ± 0.04	0.987 ± 0.019	1.136 ± 0.076	1.080 ± 0.038
$N_{B^0 \to D^{*\pm}D^{\mp}}$	862 ± 33	3257 ± 62	490 ± 31	1608 ± 52

Table 8.5 – Systematic uncertainty on *CP* parameters due to an alternative model for the signal, $B_s^0 \to D^{*\pm}D^{\mp}$ and $B^0 \to D^{*-}D_s^+$ shapes, using a scaling factor for the width parameters.

Parameter	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
Syst. unc.	0.008	−0.002	0.007	−0.002

Parameter	$A_{\text{raw}}^{K\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi,\text{Run2}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$
Syst. unc.	0.001	$6 \cdot 10^{-5}$	−0.002	−0.001

Model of combinatorial background In the nominal fit to the invariant $D^{*\pm}D^{\mp}$ mass distribution, the combinatorial background is modelled with an exponential function. As an alternative parameterisation, a second order polynomial is used. In Tab. 8.6, the resulting systematic uncertainties are given.

Table 8.6 – Systematic uncertainties on *CP* parameters due to alternative mass model for the combinatorial background.

Parameter	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
Syst. unc.	0.002	0.002	0.011	0.005

Parameter	$A_{\text{raw}}^{K\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi,\text{Run2}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$
Syst. unc.	0.0017	0.0008	0.001	0.0007

Neglected background component In the nominal fit to the invariant $D^{*\pm}D^{\mp}$ mass distribution, the potential background component $B_s^0 \to D^{*-}D_s^+$ arising from misidentification of a kaon with a pion is neglected. The mass fit to recorded data is repeated after adding a component for the $B_s^0 \to D^{*-}D_s^+$ background, mirroring the same model and parameters used for the $B^0 \to D^{*-}D_s^+$ component, but shifting the mean by the world average of the B_s^0-B^0 mass difference of 87.23 MeV/c^2 [22]. As the number of signal candidates changes by less than 1% and the number of $B_s^0 \to D^{*-}D_s^+$ candidates is compatible with zero within less than 1σ, no systematic uncertainty is assigned.

8.3.2 Decay-time-dependent efficiency

In the nominal fit to the decay time of signal candidates, the decay-time dependent efficiency is modelled with cubic splines including fixed knot positions at 0.3, 0.5, 2.7, 6.3 and 10.3 ps. In a toy study, an alternative set of knot positions (0.3, 1.3, 2.2, 6.3, 10.3 ps), which also shows a reasonable parameterisation on simulated data, is used. To assess the systematic uncertainty arising from the decay-time-dependent efficiency, a toy MC is generated using both the nominal and the alternative models and then is fit by the nominal model.

In addition, it is studied whether separate cubic spline coefficients for the final states cause any biases. The determined systematic uncertainties, which are obtained from the differences of the fit results, are shown in Tab. 8.7.

Table 8.7 – Systematic uncertainties on *CP* parameters due to an alternative set of knot positions of the cubic spline parameterisation of the decay-time-dependent efficiency and due to separated parameterisations for Run 1 and Run 2.

Parameter	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
Syst. unc. $[10^{-4}]$ (knots)	1.0	1.0	0.9	0.9
Syst. unc. $[10^{-4}]$ (split)	3.3	3.1	3.9	3.4

Parameter	$A_{\mathrm{raw}}^{K\pi,\mathrm{Run1}}$	$A_{\mathrm{raw}}^{K\pi,\mathrm{Run2}}$	$A_{\mathrm{raw}}^{K\pi\pi\pi,\mathrm{Run1}}$	$A_{\mathrm{raw}}^{K\pi\pi\pi,\mathrm{Run2}}$
Syst. unc. $[10^{-4}]$ (knots)	2.0	1.0	2.0	1.0
Syst. unc. $[10^{-4}]$ (split)	0.5	0.4	1.7	0.4

8.3.3 Decay-time resolution

Since the decay time resolution is obtained from simulated signal decays, possible differences between simulated and recorded data are considered. An alternative resolution model is created with an increased resolution of 10% by enlarging the width of the three Gaussian functions. The determined systematic uncertainties, which are obtained from the differences of the fit results are shown in Tab. 8.8.

Table 8.8 – Systematic uncertainties on *CP* parameters due to enlarging the decay time resolution by 10%.

Parameter	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
Systematic $[10^{-4}]$	0.7	0.3	1.7	0.2

Parameter	$A_{\mathrm{raw}}^{K\pi,\mathrm{Run1}}$	$A_{\mathrm{raw}}^{K\pi,\mathrm{Run2}}$	$A_{\mathrm{raw}}^{K\pi\pi\pi,\mathrm{Run1}}$	$A_{\mathrm{raw}}^{K\pi\pi\pi,\mathrm{Run2}}$
Systematic $[10^{-4}]$	1.0	1.0	2.0	1.0

8.3.4 External inputs to the decay-time fit

In the *CP* fit, the parameters τ_d, Δm_d and $\Delta\Gamma_d$ are fixed to their world averages. In order to account for systematic uncertainties induced by uncertainties of these external inputs, a toy study is performed with a generation of τ_d, Δm_d and $\Delta\Gamma_d$ according to a Gaussian centred around the central value, with a width of the uncertainty of the parameter. Then two fits are performed to the produced toy MC. The first uses fixed values, while in the second a Gaussian constraint is applied to the parameters. The arising systematic uncertainties are shown in Tab. 8.9.

Table 8.9 – Systematic uncertainties on *CP* parameters due to fixing the parameters τ_d, Δm_d and $\Delta\Gamma_d$.

Parameter	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
Syst. unc.	0.001	0.0006	0.003	0.0005

Parameter	$A_{\text{raw}}^{K\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi,\text{Run2}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$
Syst. unc.	0.00015	0.00013	0.00027	0.00023

8.3.5 Flavour tagging calibration

Studies on simulated data show that differences between the calibration obtained from the two control channels and a calibration on the signal channel exist, e.g. in the linearity of the SS tagger combination in $\omega(\eta^{\text{SS}})$. A toy study is performed in which data is generated that reproduces the mis-calibration effect observed in simulations. Two kinds of SS tagging information are generated for each dataset. One SS tagging response is generated with a linear calibration and the second one is a modification of the first one, which behaves like the non-linearity effect seen in simulations. Then, the *CP* fit is performed twice, taking the two SS tagging information into account, but with a calibration function assumed to be linear in both fits. The results of this study are reported in Tab. 8.10.

A similar toy study is performed to investigate the decay-time dependency of the OS tagging efficiency in the B^0 decay time, which is neglected in the nominal *CP* fit. For each dataset, two kinds of OS tagging information are generated. One OS tagging response is generated corresponding to a tagging efficiency information independent to the decay time, while the second one is a modification of the first one introducing the decay-time dependency. Then, the *CP* fit is performed twice considering one of the two OS tagging responses in each fit. In both fits, a decay-time independent tagging efficiency is assumed. The results of this study are listed in Tab. 8.10.

8.3.6 Systematic uncertainty on the determination of the detection asymmetry

Four possible different sources of systematic uncertainty are investigated for the determination of A_{det}. The strategy of the estimation of these sources and the results are discussed in the following paragraphs.

Table 8.10 – Systematic uncertainties due to possible limitations in the SS tagger calibration portability from the control channels to signal channel and a possible decay-time dependency of the OS tagger efficiency that are neglected in the nominal *CP* fit.

Variable	S_{D^*D}	ΔS_{D^*D}	C_{D^*D}	ΔC_{D^*D}
SS Tagging Calibration	0.015	0.012	0.014	0.015
Time dependency of OS Tagging	$6 \cdot 10^{-5}$	$4 \cdot 10^{-4}$	$6 \cdot 10^{-4}$	0.001

Variable	$A_{\text{raw}}^{K\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi,\text{Run2}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run1}}$	$A_{\text{raw}}^{K\pi\pi\pi,\text{Run2}}$
SS Tagging Calibration	$6 \cdot 10^{-5}$	$3 \cdot 10^{-5}$	$8 \cdot 10^{-5}$	$5 \cdot 10^{-5}$
Time dependency of OS Tagging	$9 \cdot 10^{-5}$	$4 \cdot 10^{-5}$	$1 \cdot 10^{-4}$	$7 \cdot 10^{-5}$

Table 8.11 – Asymmetry on the number of signal candidates obtained before and after the BDT selection and the absolute difference.

Sample	D^0 final state	A_{det} before BDT cut	A_{det} after BDT cut	$\lvert A_{\text{det,noBDT}}^{D^{*-}D^{+}} - A_{\text{det,BDT}}^{D^{*-}D^{+}}\rvert$
Run 1	$K^{\mp}\pi^{\pm}$	-0.0030 ± 0.0052	-0.0025 ± 0.0053	0.0005 ± 0.0010
	$K^{\mp}\pi^{\mp}\pi^{\pm}\pi^{\pm}$	-0.0062 ± 0.0058	-0.0086 ± 0.0059	0.0024 ± 0.0010
Run 2	$K^{\mp}\pi^{\pm}$	$+0.0046 \pm 0.0053$	$+0.0043 \pm 0.0055$	0.0003 ± 0.0015
	$K^{\mp}\pi^{\mp}\pi^{\pm}\pi^{\pm}$	-0.0053 ± 0.0053	-0.0017 ± 0.0055	0.0036 ± 0.0015

Particle identification asymmetries A possible asymmetry induced by the requirements on particle identification variables applied in the stripping is already included in the measured asymmetry. However, an additional asymmetry can be caused by the presence of particle identification variables in the BDT used in the $B^0 \to D^{*\pm}D^{\mp}$ selection. This possible asymmetry is estimated using a simulated signal sample generated without *CP* violation. The signal candidates of both charge configurations $D^{*-}D^{+}$ and $D^{*+}D^{-}$ are counted before and after the BDT selection. From the result, the asymmetry before and after the BDT selection is calculated (see Tab. 8.11). The additional asymmetry introduced by the BDT is used as a systematic uncertainty. A possible asymmetry arising from the requirements that are applied to the particle identification variables in the vetoes is neglected in the determination of A_{det}. A systematic uncertainty is assigned assuming a PID asymmetry of $A_{\text{PID}} = 5\%$ and considering the combined signal efficiencies of the vetoes $\varepsilon_{\text{veto}}$, which are 96.06% for the final state $D^0 \to K^-\pi^+$ and 96.57% for the final state $D^0 \to K^-\pi^-\pi^+\pi^+$. The resulting uncertainties, which are obtained as $(1 - \varepsilon_{\text{veto}}) \cdot A_{\text{PID}}/\varepsilon_{\text{veto}}$, are rounded up to 0.002 in both cases.

Table 8.12 – Production asymmetries of prompt $D^{\pm}$ mesons determined on the four $D^{\pm}$ samples, which are reweighted to match the kinematic properties of the $D^{*\pm}$ and $D^{\mp}$ mesons from the $B^0 \to D^{*\pm}D^{\mp}$ decay.

D^0 final state	$A_\mathrm{P}^{D^*}$ [%]	A_P^{D}[%]	$(A_\mathrm{P}^{D^*} - A_\mathrm{P}^{D})$[%]
$K^{\mp}\pi^{\pm}$	-0.5974 ± 0.00604	-0.6186 ± 0.0020	0.02115 ± 0.0063
$K^{\mp}\pi^{\mp}\pi^{\pm}\pi^{\pm}$	-0.7051 ± 0.00180	-0.6012 ± 0.0023	-0.1039 ± 0.0029

$D^{\pm}$ production asymmetry The strategy of the determination of A_det is based on the use of prompt $D^{\pm}$ mesons, assuming an equal amount of D^+ and D^- mesons is produced. In order to estimate a possible production asymmetry A_P of prompt $D^{\pm}$ mesons a measurement by the LHCb experiment performed at 7 TeV [135] is utilised. In this analysis, the $D^{\pm}$ production asymmetry is determined in bins of the transverse momentum and pseudorapidity. In order to determine A_P for the prompt $D^{\pm}$ samples used in the A_det determination, it has to be taken into account that the samples are reweighted to match the kinematic properties of the $D^{*\pm}$ and $D^{\mp}$ daughter particles of the $B^0 \to D^{*\pm}D^{\mp}$ decay. Hence, the $D^{\pm}$ production asymmetry A_P is estimated separately for the $D^{\mp}$ and $D^{*\pm}$ mesons of both final states by summing over all events

$$A_\mathrm{P} = \frac{\sum_{i=1}^{N} w_\mathrm{GB} v_i^\mathrm{bin}}{\sum_{i=1}^{N} w_\mathrm{GB}}, \qquad \sigma^2_{A_\mathrm{P}} = \frac{\sum_{i=1}^{N} w^2_\mathrm{GB} \sigma^2_{v_i^\mathrm{bin}}}{(\sum_{i=1}^{N} w_\mathrm{GB})^2}. \tag{8.2}$$

using the weights w_GB obtained by the `GBreweighter` tool and the LHCb measurement $v$bin provided in bins of η and p_T. The results are given in Tab. 8.12. As these values are determined considering the measurement performed on data samples from the year 2011, while the prompt $D^{\pm}$ samples are recorded in 2016, it is not corrected for the measured production asymmetry A_P, but the differences between the asymmetries $A_\mathrm{P}^{D^*}$ and A_P^{D} are considered as systematic uncertainties.

Secondary $D^{\pm}$ mesons In the selection, the requirement of IP$\chi^2 < 10$ is applied in order to suppress secondary $D^{\pm}$ mesons. It is tested, whether possibly remaining secondary $D^{\pm}$ mesons have an impact on the detection asymmetry by applying a stricter requirement of IP$\chi^2 < 5$ and repeating the determination of A_det. The results for A_det are given in Tab. 8.13. The absolute difference of the asymmetry obtained with the harder requirement compared to the nominal result for A_det is considered as a systematic uncertainty.

Table 8.13 – Detection asymmetry of $B^0 \to D^{*\pm}D^{\mp}$ decays obtained from prompt $D^{\pm}$ decays considering a harder requirement on the IPχ^2 and absolute difference to the nominal result for A_{det}.

D^0 final state	$A^{D^*D}_{\text{det,IP}}$	$\lvert A^{D^*D}_{\text{det,nominal}} - A^{D^*D}_{\text{det,IP}}\rvert$
$K^{\mp}\pi^{\pm}$	0.0010 ± 0.0045	0.0001 ± 0.0020
$K^{\mp}\pi^{\mp}\pi^{\pm}\pi^{\pm}$	0.0032 ± 0.0027	0.0023 ± 0.0004

Table 8.14 – Detection asymmetry of $B^0 \to D^{*\pm}D^{\mp}$ decays obtained from prompt $D^{\pm}$ decays considering a different mass parameterisation and the difference to the nominal result for A_{det}.

D^0 final state	$A^{D^*D}_{\text{det,3CB}}$	$\lvert A^{D^*D}_{\text{det,nominal}} - A^{D^*D}_{\text{det,3CB}}\rvert$
$K^{\mp}\pi^{\pm}$	0.0010 ± 0.0044	0.0001 ± 0.0018
$K^{\mp}\pi^{\mp}\pi^{\pm}\pi^{\pm}$	0.0019 ± 0.0029	0.0010 ± 0.0011

Mass model It has been checked, if the parameterisation of the invariant $K^{\pm}\pi^{\mp}\pi^{\mp}$ mass in the fits to the samples split by charge configurations has an impact on the A_{det} measurement. The parameterisation is changed from two to three Crystal Ball functions. The parameters $n_{1,2,3}$ are fixed to 100, while $\alpha_{1,2,3}$ and $\sigma_{1,2}$ are set to values obtained by the least squares method. The background is still modelled by an exponential function. The results for A_{det} are given in Tab. 8.14. The absolute difference between the result obtained with the modified mass model compared to the nominal result of A_{det} is considered as a systematic uncertainty.

8.4 Total systematic uncertainties

The total systematic uncertainty is obtained by the sum over all systematic uncertainties squared $s_{\text{tot}} = \sqrt{\sum_i s_i^2}$. The systematic uncertainties of the *CP* parameters, S_{D^*D}, ΔS_{D^*D}, C_{D^*D} and ΔC_{D^*D}, and of the overall charge asymmetries A_{raw} are listed in Tab. 8.15. The systematic uncertainties for the detection asymmetries are given in Tab. 8.16.

The total systematic uncertainty on A_{D^*D} is calculated by the square root of the sum of the total systematic uncertainties on A_{raw}, the statistical and the total systematic uncertainties on A_{det}, for each of the four data samples, in quadrature. Since parts of the systematic sources are correlated, as a conservative estimate the largest of the four resulting values is taken as systematic uncertainty on A_{D^*D}, which is ± 0.0056. This is 40% of the statistical uncertainty on A_{D^*D}.

Table 8.15 – Systematic uncertainties on the *CP* parameters and the asymmetry A_{raw} of the different final states for Run 1 and Run 2.

Origin	$\sigma_{S_{D^*D}}$	$\sigma_{\Delta S_{D^*D}}$	$\sigma_{C_{D^*D}}$	$\sigma_{\Delta C_{D^*D}}$
Fit bias	0.002	0.002	0.002	0.002
Mass model	0.011	0.003	0.014	0.006
$\Delta m_d, \tau_d, \Delta\Gamma_d$	0.001	0.001	0.003	0.001
Decay time resolution	–	–	–	–
Decay time efficiency	–	–	–	–
Flavour Tagging	0.015	0.012	0.014	0.015
Sum	0.019	0.012	0.020	0.016

Origin	$\sigma_{A_{\mathrm{raw}}^{K\pi,\mathrm{Run1}}}$	$\sigma_{A_{\mathrm{raw}}^{K\pi,\mathrm{Run2}}}$	$\sigma_{A_{\mathrm{raw}}^{K\pi\pi\pi,\mathrm{Run1}}}$	$\sigma_{A_{\mathrm{raw}}^{K\pi\pi\pi,\mathrm{Run2}}}$
Fit bias	0.0008	0.0004	0.0013	0.0007
Mass model	0.0021	0.0016	0.0025	0.0024
$\Delta m_d, \tau_d, \Delta\Gamma_d$	0.0002	0.0001	0.0003	0.0002
Decay time resolution	0.0001	0.0001	0.0002	0.0001
Decay time efficiency	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

Table 8.16 – Systematic uncertainties of the detection asymmetry.

Origin	$\sigma_{A_{\mathrm{det}}^{K\pi}}$	$\sigma_{A_{\mathrm{det}}^{K\pi\pi\pi}}$
$D^{\pm}$ production asymmetry	0.0002	0.0010
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	0.0020	0.0049

9 Conclusion and Outlook

The LHC, which is the largest machine ever built by mankind, performed excellently during the first two run periods. One of the biggest achievements of the LHC was the discovery of the Higgs Boson by the experiments ATLAS and CMS [29, 30], as the SM is completed now. However, the SM itself is not enough, as it fails to answer striking open questions. Direct searches for appropriate candidates of dark matter or super-symmetric particles at the current TeV scale have been performed without any success so far. While direct searches are limited by the given collision energies, indirect searches are very promising, since effects of New Physics can appear in higher order contributions, implicitly exploring much higher energy scales. The LHCb detector was developed to perform such indirect searches through high-precision measurements of parameters of the quark-flavour sector of the SM with decays of b- and c-hadrons. In this way, possible inconsistencies with SM predictions are aimed to be revealed. The LHCb experiment performed outstanding during the first run period of the LHC from 2010 to 2012, when $3\,\text{fb}^{-1}$ were recorded at centre-of-mass energies of 7 and 8 TeV. The performance was even improved in the second run period from 2015 to 2018, when additional $6\,\text{fb}^{-1}$ were taken at an energy of 13 TeV. The increasing amount of data makes it necessary to take higher order SM corrections into account. One way to do this is to investigate decays involving $b \to c\bar{c}d$ transitions, like $B^0 \to D^{*\pm}D^{\mp}$. In this thesis, a decay-time-dependent measurement of CP violation in $B^0 \to D^{*\pm}D^{\mp}$ decays is presented. The current full $9\,\text{fb}^{-1}$ dataset of the LHCb experiment, comprising 6160 reconstructed $B^0 \to D^{*\pm}D^{\mp}$ candidates, is used. In order to infer the production flavour, flavour tagging algorithms of the opposite- and same-side are utilised. The CP parameters are determined in an unbinned extended maximum likelihood fit, resulting in

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

with a statistical correlation coefficient between the parameters S_{D^*D} and C_{D^*D} of $\rho = 0.44$. Following Wilk's theorem [136], the conservation of *CP* symmetry can be excluded by 11.5 standard deviations with this measurement. The analysis results in the most precise single measurement of the *CP* parameters in $B^0 \to D^{*\pm}D^{\mp}$ decays. The results are consistent within 1σ with the previous measurements by Belle [137] and BaBar [132]. The precision is compatible with previous measurements of ΔC_{D^*D} and C_{D^*D}. For S_{D^*D}, ΔS_{D^*D} and A_{D^*D}, the precision is even better than the current world average [129]. Assuming the strong phase difference can be neglected in Eq. (2.62), which is reasoned by the compatibility of the parameters ΔS_{D^*D} and ΔC_{D^*D} with zero, the fit results can be used to determine

$$\sin(\phi_d^{\text{eff}}) = -\frac{S_{D^*D}}{\sqrt{1 - C^2_{D^*D}}} = 0.86\,^{+0.09}_{-0.08}\,.$$

Here, the estimation of the statistical uncertainties is performed by the generation of three million pairs of S_{D^*D} and C_{D^*D} following a two-dimensional Gaussian distribution considering their correlation. Then, $\sin(\phi_d^{\text{eff}})$ is calculated for each pair and the 68% confidence intervals are chosen to result the asymmetric uncertainties. In combination with the world average for $\sin 2\beta = (0.699 \pm 0.017)$ of all charmonium decays [129], where the quantity can be measured very clean, the phase shift $\Delta\phi_d$, describing higher-order corrections in the SM, can be extracted. The result

$$\Delta\phi_d = (0.27 \pm 0.16)\,\text{rad}$$

is compatible with zero within less than two standard deviations. The equivalent quantity has already been calculated in $B^0 \to D^+D^-$ decays from results of Belle [137] and BaBar [132] to $\Delta\phi = (30^{+23}_{-30})\,^\circ = (0.52^{+0.40}_{-0.56})\,\text{rad}$ [51]. While the most recent LHCb measurement only covers $3\,\text{fb}^{-1}$ of the available dataset, it results in the currently most precise single measurement of the quantity $\Delta\phi = (-0.16^{+0.19}_{-0.21})\,\text{rad}$ in $B^0 \to D^+D^-$ decays [114]. Additional comprehensive measurements are possible with $B^0 \to D^{(*)+}_{(s)}D^{(*)-}$ decays.

The measurement of $\sin(\phi_d^{\text{eff}})$ can be further improved by considering additional final states like $D^0 \to K^-K^+$ and $D^+ \to K^-K^+\pi^+$, where each of them would contribute 10% to the sample size. Furthermore, studies on additional tagging algorithm are ongoing, where one of this experimental flavour tagging algorithms promises improvements of up to 200%. This so-called inclusive flavour tagger is based on a Recursive Neural Net (RNN) [138], exploiting all available tracks in an event.

LHCb's biggest competitor will be the Belle II experiment, where the data taking already started in 2018. Since March 2019, the full detector installation was completed, thus, Belle II is ready to take physics data. It is planned to record an instantaneous luminosity of 50 times larger compared to the dataset available of the predecessor experiment Belle, resulting in $50\,\mathrm{ab}^{-1}$ [139]. If these expectations are fulfilled, the Belle II collaboration will improve the sensitivity of many *CP* violation measurements.

Currently, the LHC is in its second long shutdown. During this time, the LHCb detector will be upgraded in order to handle more demanding circumstances of the Run 3, namely the higher instantaneous luminosity of $\mathcal{L} = 2 \cdot 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and the increased $b\overline{b}$-pair production rate of 10^6 per second [140]. Part of the upgrade are a new hybrid pixel-sensor VELO [141], improvements to the calorimeter, RICH and muon systems [142] as well as enhanced tracking stations made of silicon micro-strips and scintillating-fibres [143]. Furthermore, the trigger is changed to a full software trigger that is able to read the detector out at the full bunch crossing rate of 40 MHz [144]. It is planned to restart the data taking in 2021 and to collect an integrated luminosity of $50\,\mathrm{fb}^{-1}$ in total during the Run 3. Assuming the conditions remain unchanged compared to the presented analysis, specifically the selection and flavour-tagging efficiencies, the sensitivities on $(\Delta)S_{D^*D}$ and $(\Delta)S_{D^*D}$ can be improved to 0.03 and 0.04, respectively.

While some competition among experiments is beneficial, it is important to state that only the joint efforts of the whole particle physics community will allow to increase the precision of the measurements to levels that will finally reveal physics beyond the Standard Model.

Bibliography

[1] Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, DOI: 10.1103/PhysRevLett.81.1562, arXiv:hep-ex/9807003 [hep-ex].

[2] 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, DOI: 10.1103/PhysRevLett.89.011301, arXiv:nucl-ex/0204008 [nucl-ex].

[3] P. A. R. Ade *et al.*, *Planck 2015 results. XIII. Cosmological parameters*, Astron. Astrophys. **594** (2016) A13, DOI: 10.1051/0004-6361/201525830, arXiv:1502.01589 [astro-ph.CO].

[4] V. Trimble, *Existence and Nature of Dark Matter in the Universe*, Ann. Rev. Astron. Astrophys. **25** (1987) 425–472, DOI: 10.1146/annurev.aa.25.090187.002233.

[5] H. W. Babcock, *The rotation of the Andromeda Nebula*, Lick Observatory Bulletin **19** (1939) 41–51, DOI: 10.5479/ADS/bib/1939LicOB.19.41B.

[6] A. Sakharov, *Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32–35, DOI: 10.1070/PU1991v034n05ABEH002497.

[7] J. Christenson *et al.*, *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, Phys. Rev. Lett. **13** (1964) 138–140, DOI: 10.1103/PhysRevLett.13.138.

[8] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. **D51** (1995) 379–394, DOI: 10.1103/PhysRevD.51.379, arXiv:hep-ph/9404302 [hep-ph].

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

[10] D. J. Griffiths, *Introduction to elementary particles*, Wiley, 2008, https://cds.cern.ch/record/111880.

[11] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588, DOI: `10.1016/0029-5582(61)90469-2`.

[12] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266, DOI: `10.1103/PhysRevLett.19.1264`.

[13] A. Salam, *Weak and Electromagnetic Interactions*, 8th Nobel Symposium, *Elementary Particle Theory*, ed. by N. Svaratholm, **C68-05-19**, Lerum, Sweden, 1968.

[14] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214–215, DOI: `10.1016/S0031-9163(64)92001-3`.

[15] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343–1346, DOI: `10.1103/PhysRevLett.30.1343`.

[16] G. Lüders, *On the Equivalence of Invariance under Time Reversal and under Particle-Antiparticle Conjugation for Relativistic Field Theories*, Kong. Dan. Vid. Sel. Mat. Fys. Med. **28N5** (1954) 1–17.

[17] W. Pauli and L. F. Curtiss, *Niels Bohr and the Development of Physics*, Am. J. Phys. **24** (1956) 292, DOI: `10.1119/1.1934212`.

[18] C. Wu *et al.*, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413–1414, DOI: `10.1103/PhysRev.105.1413`.

[19] B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, DOI: `10.1103/PhysRevLett.87.091801`, `arXiv:hep-ex/0107013 [hep-ex]`.

[20] K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, DOI: `10.1103/PhysRevLett.87.091802`, `arXiv:hep-ex/0107061 [hep-ex]`.

[21] R. Aaij *et al.*, *Observation of CP Violation in Charm Decays*, Phys. Rev. Lett. **122** (2019) 211803, DOI: `10.1103/PhysRevLett.122.211803`, `arXiv:1903.08726 [hep-ex]`.

[22] M. Tanabashi *et al.*, *Review of Particle Physics*, Phys. Rev. **D98** (2018) 030001, DOI: `10.1103/PhysRevD.98.030001`.

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

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

[25] R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi K^- p$ decays*, Phys. Rev. Lett. **115** (2015) 072001, DOI: `10.1103/PhysRevLett.115.072001`, `arXiv:1507.03414 [hep-ex]`.

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

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

[28] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587, DOI: `10.1103/PhysRevLett.13.585`.

[29] 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) 1–29, DOI: `10.1016/j.physletb.2012.08.020`, `arXiv:1207.7214 [hep-ex]`.

[30] 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) 30–61, DOI: `10.1016/j.physletb.2012.08.021`, `arXiv:1207.7235 [hep-ex]`.

[31] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533, DOI: `10.1103/PhysRevLett.10.531`.

[32] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657, DOI: `10.1143/PTP.49.652`.

[33] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802, DOI: `10.1103/PhysRevLett.53.1802`.

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

[35] J. Charles *et al.*, *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics*, Phys. Rev. **D91** (2015) 073007, DOI: `10.1103/PhysRevD.91.073007`, `arXiv:1501.05013 [hep-ph]`.

[36] 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, DOI: `10.1103/PhysRevLett.55.1039`.

[37] J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1–131, DOI: `10.1140/epjc/s2005-02169-1`, `arXiv:hep-ph/0406184 [hep-ph]`, updated results and plots available at: `http://ckmfitter.in2p3.fr`.

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

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

[40] R. Aaij *et al.*, *First observation of CP violation in the decays of B_s^0 mesons*, Phys. Rev. Lett. **110** (2013) 221601, DOI: `10.1103/PhysRevLett.110.221601`, `arXiv:1304.6173 [hep-ex]`.

[41] R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in $B^0 - \bar{B}^0$ mixing*, Phys. Rev. Lett. **114** (2015) 041601, DOI: `10.1103/PhysRevLett.114.041601`, `arXiv:1409.8586 [hep-ex]`.

[42] R. Aaij *et al.*, *Measurement of the CP asymmetry in $B_s^0 - \bar{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016), [Addendum: Phys. Rev. Lett.118,no.12,129903(2017)] 061803, DOI: `10.1103/PhysRevLett.118.129903,10.1103/PhysRevLett.117.061803`, `arXiv:1605.09768 [hep-ex]`.

[43] A. Lenz and U. Nierste, *Numerical Updates of Lifetimes and Mixing Parameters of B Mesons*, *CKM unitarity triangle. Proceedings, 6th International Workshop, CKM 2010, Warwick, UK, September 6-10, 2010*, 2011, `arXiv:1102.4274 [hep-ph]`.

[44] M. Artuso, G. Borissov, and A. Lenz, *CP violation in the B_s^0 system*, Rev. Mod. Phys. **88** (2016) 045002, DOI: `10.1103/RevModPhys.88.045002`, `arXiv:1511.09466 [hep-ph]`.

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

[46] R. Aaij *et al.*, *Precision measurement of CP violation in $B_s^0 \to J/\psi K^+K^-$ decays*, Phys. Rev. Lett. **114** (2015) 041801, DOI: `10.1103/PhysRevLett.114.041801`, `arXiv:1411.3104 [hep-ex]`.

[47] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L^0 \to \mu\bar{\mu}$, $K^+ \to \pi^+\nu\bar{\nu}$ and $K^0 \leftrightarrow \bar{K}^0$*, Prog. Theor. Phys. **65** (1981), [Addendum: Prog. Theor. Phys.65,1772(1981)] 297, DOI: `10.1143/PTP.65.297`.

[48] J. Albert, A. Datta, and D. London, *A Measurement of gamma from the decays $B_d^0(t) \to D^{(*)+}D^{(*)-}$ and $B_d^0(t) \to D_s^{(*)+}D^{(*)-}$*, Phys. Lett. **B605** (2005) 335–343, DOI: `10.1016/j.physletb.2004.11.046`, `arXiv:hep-ph/0410015 [hep-ph]`.

[49] A. Datta and D. London, *Extracting gamma from $B_d^0(t) \to D^{(*)+}D^{(*)-}$ and $B_d^0(t) \to D_s^{(*)+}D^{(*)-}$*, Phys. Lett. **B584** (2004) 81–88, DOI: `10.1016/j.physletb.2003.12.048`, `arXiv:hep-ph/0310252 [hep-ph]`.

[50] Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. **C77** (2017), updated results and plots available at `https://hflav.web.cern.ch` 895, DOI: `10.1140/epjc/s10052-017-5058-4`, `arXiv:1612.07233 [hep-ex]`.

[51] L. Bel *et al.*, *Anatomy of $B \to D\overline{D}$ decays*, JHEP **07** (2015) 108, DOI: `10.1007/JHEP07(2015)108`, `arXiv:1505.01361 [hep-ph]`.

[52] Z.-z. Xing, *CP violation in $B_d^0 \to D^+D^-, D^{*+}D^-, D^+D^{*-}$ and $D^{*+}D^{*-}$ decays*, Phys. Rev. D **61** (1 1999) 014010, DOI: `10.1103/PhysRevD.61.014010`.

[53] Z.-z. Xing, *Measuring CP violation and testing factorization in $B_d^0 \to D^{*\pm}D^{\mp}$ and $B_s^0 \to D_s^{*\pm}D_s^{\mp}$ decays*, Phys. Lett. B **443** (1998) 365, DOI: `10.1016/S0370-2693(98)01285-4`.

[54] *CERN Annual Personnel Statistics 2017* (2017), `https://cds.cern.ch/record/2317058`.

[55] J. Alves A. Augusto *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005, DOI: `10.1088/1748-0221/3/08/S08005`.

[56] R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, DOI: `10.1142/S0217751X15300227`, `arXiv:1412.6352 [hep-ex]`.

[57] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008), ed. by L. Evans S08001, DOI: `10.1088/1748-0221/3/08/S08001`.

[58] C. D. Melis, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, 2016, `https://cds.cern.ch/record/2119882/files/CERN%20Accelerator%20Complex.jpg` (visited on 2019-05-15).

[59] G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003, DOI: `10.1088/1748-0221/3/08/S08003`.

[60] S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004, DOI: `10.1088/1748-0221/3/08/S08004`.

[61] K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002, DOI: `10.1088/1748-0221/3/08/S08002`.

[62] G. Anelli *et al.*, *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007, DOI: `10.1088/1748-0221/3/08/S08007`.

[63] B. Acharya *et al.*, *The Physics Programme Of The MoEDAL Experiment At The LHC*, Int. J. Mod. Phys. **A29** (2014) 1430050, DOI: `10.1142/S0217751X14300506`, `arXiv:1405.7662 [hep-ph]`.

[64] O. Adriani *et al.*, *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006, DOI: `10.1088/1748-0221/3/08/S08006`.

[65] G. Aad *et al.*, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803, DOI: `10.1103/PhysRevLett.114.191803`, `arXiv:1503.07589 [hep-ex]`.

[66] M. De Cian, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC* (2019), `https://cds.cern.ch/record/2682906`.

[67] *$b\bar{b}$ production angle plots*, `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html` (visited on 2019-01-18).

[68] M. Chaichian and A. Fridman, *On a possibility for measuring effects of CP violation at pp colliders*, Phys. Lett. **B298** (1993) 218–223, DOI: `10.1016/0370-2693(93)91733-4`.

[69] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000) 137–161, DOI: `10.1007/s100520000460`, `arXiv:hep-ph/0005110 [hep-ph]`.

[70] E. Norrbin and R. Vogt, *Bottom production asymmetries at the LHC* (2000), `arXiv:hep-ph/0003056 [hep-ph]`.

[71] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, DOI: `10.1088/1748-0221/9/09/P09007`, `arXiv:1405.7808 [physics.ins-det]`.

[72] P. Collins, *Velo constraint system installation*, 2007, `http://cds.cern.ch/record/1017398/files/velo-2007-003_01.jpg` (visited on 2019-03-15).

[73] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, DOI: `10.1140/epjc/s10052-013-2431-9`, `arXiv:1211.6759 [physics.ins-det]`.

[74] *LHCb trigger schemes*, `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html` (visited on 2019-01-28).

[75] R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, DOI: `10.1088/1748-0221/14/04/P04013`, `arXiv:1812.10790 [hep-ex]`.

[76] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, DOI: `10.1088/1748-0221/8/04/P04022`, `arXiv:1211.3055 [hep-ex]`.

[77] J. Albrecht *et al.*, *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, DOI: `10.1088/1742-6596/513/1/012001`, `arXiv:1310.8544 [hep-ex]`.

[78] M. Frank *et al.*, *Deferred High Level Trigger in LHCb: A Boost to CPU Resource Utilization*, J. Phys. Conf. Ser. **513** (2014) 012006, DOI: `10.1088/1742-6596/513/1/012006`.

[79] G. Barrand *et al.*, *GAUDI - The software architecture and framework for building LHCb data processing applications*, Comput. Phys. Commun. **140** (2001) 92–95.

[80] R. Aaij *et al.*, *The Moore Project*, ed. by G. Raven, 2015, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/` (visited on 2019-03-15).

[81] R. Aaij *et al.*, *The Brunel Project*, ed. by M. Cattaneo and R. Lambert, 2015, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/` (visited on 2019-03-15).

[82] R. Aaij *et al.*, *The DaVinci Project*, ed. by E. Rodrigues, 2015, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/` (visited on 2019-03-15).

[83] R. E. Kalman, *A New Approach to Linear Filtering And Prediction Problems*, J. Basic Eng. **82** (1960) 35, DOI: `10.1115/1.3662552`.

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

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

[86] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, DOI: `10.1088/1126-6708/2006/05/026`, `arXiv:hep-ph/0603175 [hep-ph]`.

[87] T. Sjöstrand *et al.*, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, DOI: `10.1016/j.cpc.2015.01.024`, `arXiv:1410.3012 [hep-ph]`.

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

[89] S. Agostinelli *et al.*, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303, DOI: `10.1016/S0168-9002(03)01368-8`.

[90] J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270–278, DOI: `10.1109/TNS.2006.869826`.

[91] R. Aaij *et al.*, *The Boole Project*, ed. by M. Cattaneo, 2015, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/`.

[92] M. Kussmann *et al.*, *DeLorean: A Storage Layer to Analyze Physical Data at Scale*, *Datenbanksysteme für Business, Technologie und Web (BTW 2017)*, ed. by B. M. et al. (Hrsg.), **P-265**, LNI, 2017 413–422.

[93] M. Hausenblas and J. Nadeau, *Apache Drill: Interactive Ad-Hoc Analysis at Scale*, Big Data **1** (2013) 100–104, DOI: `10.1089/big.2013.0011`.

[94] S. Melnik *et al.*, *Dremel: Interactive Analysis of Web-Scale Datasets*, Commun. ACM **54** (2011) 114–123, `http://cacm.acm.org/magazines/2011/6/108648-dremel-interactive-analysis-of-web-scale-datasets/fulltext`.

[95] J. Dean and S. Ghemawat, *MapReduce: Simplified Data Processing on Large Clusters*, *OSDI'04: Sixth Symposium on Operating System Design and Implementation*, San Francisco, CA, 2004 137–150.

[96] K. Shvachko *et al.*, *The Hadoop Distributed File System*, *Proceedings of the 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*, MSST '10, Washington, DC, USA: IEEE Computer Society, 2010 1–10, DOI: `10.1109/MSST.2010.5496972`, `http://dx.doi.org/10.1109/MSST.2010.5496972`.

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

[98] G. D. Cowan, *Statistical data analysis*, Oxford: Oxford Univ. Press, 1998, `https://cds.cern.ch/record/358560`.

[99] F. James, *Statistical methods in experimental physics*, Hackensack, USA: World Scientific, 2006.

[100] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356–369, DOI: `10.1016/j.nima.2005.08.106`, `arXiv:physics/0402083 [physics.data-an]`.

[101] L. Breiman *et al.*, *Classification and regression trees*, Belmont, USA: Wadsworth international group, 1984.

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

[103] B. Efron, *Bootstrap Methods: Another Look at the Jackknife*, Ann. Stat. **7** (1979) 1–26.

[104] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566–575, DOI: `10.1016/j.nima.2005.06.078`, `arXiv:physics/0503191 [physics]`.

[105] J. T. Wishahi, *Measurement of CP Violation in $B^0 \to J/\psi K^0_S$ Decays with the LHCb Experiment*, Ph.D. thesis, TU Dortmund University, 2013, DOI: `10.17877/DE290R-13417`.

[106] R. Aaij *et al.*, *New algorithms for identifying the flavour of B^0 mesons using pions and protons*, Eur. Phys. J. **C77** (2017) 238, DOI: `10.1140/epjc/s10052-017-4731-y`, `arXiv:1610.06019 [hep-ex]`.

[107] R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, DOI: `10.1140/epjc/s10052-012-2022-1`, `arXiv:1202.4979 [hep-ex]`.

[108] R. Aaij *et al.*, *B flavour tagging using charm decays at the LHCb experiment*, JINST **10** (2015) P10005, DOI: `10.1088/1748-0221/10/10/P10005`, `arXiv:1507.07892 [hep-ex]`.

[109] D. Derkach *et al.*, *Machine-Learning-based global particle-identification algorithms at the LHCb experiment*, J. Phys. Conf. Ser. **1085** (2018) 042038, DOI: `10.1088/1742-6596/1085/4/042038`.

[110] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, JINST **10** (2015) P02011, DOI: `10.1088/1748-0221/10/02/P02011`, `arXiv:1411.5528 [physics.data-an]`.

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

[112] Y. Xie, *Principles to optimize event selections for measurements of CP asymmetries*, tech. rep. LHCb-INT-2009-012. CERN-LHCb-INT-2009-012, Geneva: CERN, 2009, `https://cds.cern.ch/record/1187570`.

[113] R. Aaij *et al.*, *First study of the CP-violating phase and decay-width difference in* $B_s^0 \rightarrow \psi(2S)\phi$ *decays*, Phys. Lett. **B762** (2016) 253–262, DOI: `10.1016/j.physletb.2016.09.028`, `arXiv:1608.04855 [hep-ex]`.

[114] R. Aaij *et al.*, *Measurement of CP violation in* $B^0 \rightarrow D^+D^-$ *decays*, Phys. Rev. Lett. **117** (2016) 261801, DOI: `10.1103/PhysRevLett.117.261801`, `arXiv:1608.06620 [hep-ex]`.

[115] M. Oreglia, *A Study of the Reactions* $\psi' \rightarrow \gamma\gamma\psi$, Ph.D. thesis, Stanford University, 1980.

[116] J. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the* J/ψ *and* ψ', Ph.D. thesis, Stanford University, 1982.

[117] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-Prime and Upsilon resonances*, Ph.D. thesis, Cracow, INP, 1986.

[118] R. Aaij *et al.*, *Measurement of* $\sigma(pp \rightarrow b\bar{b}X)$ *at* $\sqrt{s} = 7\,\mathrm{TeV}$ *in the forward region*, Phys. Lett. **B694** (2010) 209–216, DOI: `10.1016/j.physletb.2010.10.010`, `arXiv:1009.2731 [hep-ex]`.

[119] R. Aaij *et al.*, *Production of* J/ψ *and* Υ *mesons in* pp *collisions at* $\sqrt{s} = 8$ *TeV*, JHEP **06** (2013) 064, DOI: `10.1007/JHEP06(2013)064`, `arXiv:1304.6977 [hep-ex]`.

[120] R. Aaij *et al.*, *Measurement of forward* J/ψ *production cross-sections in* pp *collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **10** (2015), [Erratum: JHEP05,063(2017)] 172, DOI: `10.1007/JHEP05(2017)063,10.1007/JHEP10(2015)172`, `arXiv:1509.00771 [hep-ex]`.

[121] J. Wimberley *et al.*, *Espresso Performance Monitor*, `https://gitlab.cern.ch/lhcb-ft/EspressoPerformanceMonitor`, 2017, (visited on 2019-09-10).

[122] R. Aaij *et al.*, *Measurement of CP violation in* $B^0 \rightarrow D^{\mp}\pi^{\pm}$ *decays*, JHEP **06** (2018) 084, DOI: `10.1007/JHEP06(2018)084`, `arXiv:1805.03448 [hep-ex]`.

[123] R. Aaij *et al.*, *Measurements of the* B^+, B^0, B_s^0 *meson and* Λ_b^0 *baryon lifetimes*, JHEP **04** (2014) 114, DOI: `10.1007/JHEP04(2014)114`, `arXiv:1402.2554 [hep-ex]`.

[124] M. Schellenberg, *Analyse der VELO-Rekonstruktionseffizienz im Zerfall $B^0 \to J/\psi K^0_S$ am LHCb-Experiment*, B.Sc. thesis, TU Dortmund University, 2013.

[125] T. M. Karbach, G. Raven, and M. Schiller, *Decay time integrals in neutral meson mixing and their efficient evaluation* (2014), `arXiv:1407.0748 [physics.data-an]`.

[126] C. de Boor, *A practical guide to splines*, **27**, Applied Mathematical Sciences, Springer-Verlag New York, 1978, ISBN: 978-0-387-95366-3.

[127] J. Berkson, *Minimum Chi-Square, not Maximum Likelihood!*, Ann. Statist. **8** (1980) 457–487, DOI: `10.1214/aos/1176345003`.

[128] A. Rogozhnikov, *Brilliantly wrong, thoughts on science and programming: Reweighting with boosted decision trees*, `http://arogozhnikov.github.io/2015/10/09/gradient-boosted-reweighter.html`, 2019.

[129] Y. Amhis *et al.*, *Heavy Flavor Averaging Group*, `http://www.slac.stanford.edu/xorg/hfag/`, (visited on 2019-09-10).

[130] R. Aaij *et al.*, *Measurement of B^0, B^0_s, B^+ and Λ^0_b production asymmetries in 7 and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017) 139–158, DOI: `10.1016/j.physletb.2017.09.023`, `arXiv:1703.08464 [hep-ex]`.

[131] T. M. Karbach and M. Schlupp, *Constraints on Yield Parameters in Extended Maximum Likelihood Fits* (2012), `arXiv:1210.7141 [physics.data-an]`.

[132] B. Aubert *et al.*, *Measurements of time-dependent CP asymmetries in $B^0 \to D^{(*)+}D^{(*)-}$ decays*, Phys. Rev. **D79** (2009) 032002, DOI: `10.1103/PhysRevD.79.032002`, `arXiv:0808.1866 [hep-ex]`.

[133] J. P. Lees *et al.*, *Measurement of the Time-Dependent CP Asymmetry of Partially Reconstructed $B^0 \to D^{*+}D^{*-}$ Decays*, Phys. Rev. **D86** (2012) 112006, DOI: `10.1103/PhysRevD.86.112006`, `arXiv:1208.1282 [hep-ex]`.

[134] B. Kronenbitter *et al.*, *First observation of CP violation and improved measurement of the branching fraction and polarization of $B^0 \to D^{*+}D^{*-}$ decays*, Phys. Rev. **D86** (2012) 071103, DOI: `10.1103/PhysRevD.86.071103`, `arXiv:1207.5611 [hep-ex]`.

[135] R. Aaij *et al.*, *Measurement of the $D^\pm$ production asymmetry in 7 TeV pp collisions*, Phys. Lett. **B718** (2013) 902–909, DOI: `10.1016/j.physletb.2012.11.038`, `arXiv:1210.4112 [hep-ex]`.

[136] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statist. **9** (1938) 60–62, DOI: `10.1214/aoms/1177732360`, `https://doi.org/10.1214/aoms/1177732360`.

[137] 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, DOI: `10.1103/PhysRevD.85.091106`, `arXiv:1203.6647 [hep-ex]`.

[138] T. Mikolov, *Statistical Language Models Based on Neural Networks*, Ph.D. thesis, Brno University of Technology, 2012.

[139] T. Abe *et al.*, *Belle II Technical Design Report* (2010), `arXiv:1011.0352 [physics.ins-det]`.

[140] R. Aaij *et al.*, *Framework TDR for the LHCb Upgrade: Technical Design Report* (2012), `http://cds.cern.ch/record/1443882`.

[141] R. Aaij *et al.*, *LHCb VELO Upgrade Technical Design Report* (2013), `http://cds.cern.ch/record/1624070`.

[142] R. Aaij *et al.*, *LHCb PID Upgrade Technical Design Report* (2013), `http://cds.cern.ch/record/1624074`.

[143] R. Aaij *et al.*, *LHCb Tracker Upgrade Technical Design Report* (2014), `http://cds.cern.ch/record/1647400`.

[144] *LHCb Trigger and Online Upgrade Technical Design Report* (2014), `http://cds.cern.ch/record/1701361`.

Acknowledgements

Der Abschluss einer solchen Arbeit ist ohne die Unterstützung vieler Menschen nicht möglich. Diese letzte Seiten möchte ich denjenigen widmen, die mir auf unterschiedlichste Weise geholfen haben.

Zu allererst gilt mein Dank meinem Doktorvater Herrn Prof. Dr. Bernhard Spaan. Ich danke Ihnen nicht nur für die anhaltende Unterstützung und Beratung bei der Durchführung meiner Forschung, sondern auch dafür, dass Sie mir die Möglichkeit gegeben haben in einer internationalen Kollaboration mitzuwirken.

Ein besonderer Dank gebührt auch Herrn Prof. Dr. Kevin Kröninger, da er sich die Zeit genommen hat meine Arbeit als Zweitgutachter zu bewerten. Auch den weiteren Mitgliedern der Prüfungskommission, Herrn Prof. Dr. Roland Böhmer und Herrn Dr. Jörg Bünemann, möchte ich hier noch einmal explizit danken.

Darüber hinaus bedanke ich mich für die Unterstützung durch den Sonderforschungsbereich 876. Die interdisziplinäre Forschung ermöglichte mir eine zusätzliche Sichtweise auf die Datenanalyse.

A big thank you goes to the $B^0 \to D^{*\pm}D^{\mp}$ analysis group. It was a pleasure to work with you, Bassem, Giulia, Marta, Nicoletta, Philipp and Stefania.

Obwohl er den Lehrstuhl bereits seit einigen Jahren verlassen hat, merke ich immer noch den Einfluss, den Julian auf unsere Arbeitsgruppe und auch auf mich persönlich hinterlassen hat. Du warst, und bist immer noch, ein wichtiger Betreuer und Berater für mich.

Ein großer Dank geht an den gesamten Lehrstuhl E5. Die harmonische Atmosphäre und das große Gemeinschaftsgefühl ermöglichten einen angenehmen und produktiven Arbeitsalltag.

Vielen Dank an alle Personen, die meine Arbeit Korrektur gelesen haben. Danke für die zahlreichen sprachlichen und inhaltlichen Verbesserungsvorschläge. An der Stelle möchte ich insbesondere Leanna Müller danken, die sich die Mühe gemacht hat meine vollständige Arbeit zu lesen.

Meiner Familie und insbesondere meinen Eltern möchte ich ein ganz besonders großes Dankeschön aussprechen. Danke für die Ermöglichung meines Studiums und

Unterstützung in jeder Lebenslage.

Abschließend möchte ich meinem Partner Uli für den bedingungslosen Rückhalt und Ermutigung in den letzten Jahren danken.