

Studies of $B \rightarrow D^{*+} D^{*-}$ decays at the LHCb experiment

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

vorgelegt von
Jan Langer
geboren am 31.01.1996 in Wickede (Ruhr)

Fakultät Physik
Technische Universität Dortmund

Dortmund, August 2024

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

Gutachter:

Prof. Dr. Johannes Albrecht

PD Dr. Dominik Elsässer

Vorsitzende der Prüfungskommission:

Prof. Dr. Mirko Cinchetti

Vertreter der wissenschaftlichen Mitarbeiter*innen:

Dr. Moritz Cygorek

Datum der mündlichen Prüfung:

10.10.2024

Abstract

In this thesis, two measurements probing the Standard Model of particle physics in $B \rightarrow D^{*+}D^{*-}$ decays are presented. Both measurements are performed using proton-proton collision data recorded by the LHCb experiment at the Large Hadron Collider, corresponding to an integrated luminosity of 9 fb^{-1} .

The first analysis is a branching fraction measurement of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay. The decay is observed with a significance of more than 10 standard deviations. The absolute time-integrated $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction is measured to be

$$\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-}) = (2.15 \pm 0.26 \pm 0.09 \pm 0.06 \pm 0.16) \times 10^{-4},$$

where the first uncertainty is statistical, the second systematic, the third the uncertainty of the fragmentation fraction ratio f_s/f_d and the fourth is due to the limited precision of the $B^0 \rightarrow D^{*+}D^{*-}$ branching fraction used in the normalisation.

The second analysis presented is the measurement of CP violation in $B^0 \rightarrow D^{*+}D^{*-}$ decays. As the final state is a mixture of CP -even and CP -odd states, the CP observable $S_{D^{*+}D^{*-}}$ is diluted. Therefore, an angular analysis is performed to determine the dilution factor. The current state of the analysis and its sensitivity are presented and discussed. With the current precision, the uncertainty of the measurement is expected to be statistically dominated.

Kurzfassung

In dieser Arbeit werden zwei Messungen vorgestellt, die das Standardmodell der Teilchenphysik in $B \rightarrow D^{*+}D^{*-}$ -Zerfällen untersuchen. Beide Messungen werden mit Daten aus Proton-Proton Kollisionen durchgeführt, die vom LHCb Experiment am Large Hadron Collider aufgezeichnet wurden und einer integrierten Luminosität von 9 fb^{-1} entsprechen.

Die erste Analyse ist eine Messung des Verzweigungsverhältnisses des $B_s^0 \rightarrow D^{*+}D^{*-}$ Zerfalls. Der Zerfall wird mit einer Signifikanz von mehr als 10 Standardabweichungen beobachtet. Das gemessene absolute zeitintegrierte $B_s^0 \rightarrow D^{*+}D^{*-}$ Verzweigungsverhältnis ist

$$\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-}) = (2.15 \pm 0.26 \pm 0.09 \pm 0.06 \pm 0.16) \times 10^{-4},$$

wobei die erste Unsicherheit statistisch ist, die zweite systematisch, die dritte die Unsicherheit des Verhältnisses der Fragmentierungswahrscheinlichkeiten f_s/f_d und die vierte die begrenzte Genauigkeit des bei der Normalisierung verwendeten $B^0 \rightarrow D^{*+}D^{*-}$ Verzweigungsverhältnisses ist.

Die zweite vorgestellte Analyse ist die Messung von CP -Verletzung in $B^0 \rightarrow D^{*+}D^{*-}$ Zerfällen. Da der Endzustand eine Mischung aus geraden und ungeraden CP -Zuständen ist, ist die CP -Observable $S_{D^{*+}D^{*-}}$ abgeschwächt. Daher wird eine Winkelanalyse durchgeführt, um den Abschwächungsfaktor zu bestimmen. Der aktuelle Stand der Analyse und ihre Sensitivität werden vorgestellt und diskutiert. Mit der derzeitigen Präzision wird erwartet, dass die Unsicherheit der Messung statistisch dominiert ist.

Table of contents

1	Introduction	1
2	Standard Model of particle physics	3
2.1	Particles and interactions	3
2.2	The CKM matrix	7
2.3	The system of neutral B mesons	10
2.4	Types of CP violation	13
2.5	The $B \rightarrow D^{*+} D^{*-}$ decays	16
2.6	Limitations of the Standard Model	22
3	The LHCb experiment at the LHC	25
3.1	The Large Hadron Collider	25
3.2	The LHCb detector	27
3.3	LHCb software	35
3.4	Flavour tagging at LHCb	37
4	Statistical methods	43
4.1	Reweighting approach	43
4.2	Maximum-likelihood fits	45
4.3	$sPlot$ method	46
4.4	Significance	47
5	Analysis strategies	49
6	Selection of signal decays	51
6.1	Preselection	52
6.2	Combinatorial background suppression	56
6.3	Multiple candidates removal	58
6.4	Physical background	60
6.5	Selection efficiencies	63
7	Observation of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay	69
7.1	Signal extraction	70
7.1.1	Fit model	70
7.1.2	Fit to simulation	72
7.1.3	Fit to data	73

7.2	Studies of systematic effects	76
7.2.1	Polarisation	77
7.2.2	Multiple candidates	79
7.2.3	Mass model	81
7.2.4	Mass fit range	85
7.3	Calculation of the ratio of branching fractions	86
7.4	Effect of the B_s^0 lifetime	88
7.5	Results	89
8	Measurement of CP violation in $B^0 \rightarrow D^{*+}D^{*-}$ decays	93
8.1	Extraction of B^0 candidates	93
8.1.1	Fit to simulation	94
8.1.2	Fit to data	94
8.2	Decay-time parametrisation	96
8.2.1	External asymmetries	97
8.2.2	Flavour tagging	97
8.2.3	Decay-time resolution	104
8.2.4	Decay-time-dependent efficiency	106
8.3	Extraction of the CP observables	107
8.3.1	Angular acceptance	109
8.3.2	Angular decay-time fit	110
8.4	Studies of systematic effects	114
8.4.1	Bootstrapping	117
8.4.2	Gaussian constraints	119
8.4.3	Fit bias	119
8.4.4	Mass model	122
8.4.5	Decay-time resolution model	123
8.4.6	Decay-time acceptance model	124
8.5	Results	124
9	Summary and outlook	127
	Bibliography	129

1 Introduction

The Standard Model of particle physics (SM) [1–3] is a well-established theory that describes the fundamental particles and their interactions. It has been tested with remarkable precision. At the microscopic level, it describes most of the observed phenomena and has allowed several predictions about particles that were later observed, such as the Z^0 boson [4], the W^\pm boson [5, 6] and the Higgs boson [7, 8]. Although the development of this model has significantly increased the understanding of nature over the last century, many questions remain unresolved. The observed neutrino oscillations [9, 10] and the resulting neutrino masses are not included in the SM. The nature of dark energy and dark matter is still unknown, although together, they account for about 95% of the energy density in the observed universe [11]. Furthermore, the origin of the matter-antimatter asymmetry in the universe is not yet understood. In 1967, Sakharov developed three conditions to be fulfilled for the occurrence of this asymmetry [12]. These include the violation of the baryon number, a deviation from thermal equilibrium and the violation of the CP symmetry. The last condition is particularly interesting because it is described in the SM, but only to an extent that is too small to explain the observed matter-antimatter asymmetry [13, 14]. In the SM, CP violation is incorporated through the complex phase in the Cabibbo-Kobayashi-Maskawa (CKM) matrix [15, 16], which describes the mixing between quark flavours. Precise measurements of the CKM parameters are performed to search for possible deviations from the SM, which could indicate New Physics (NP). These NP effects could potentially contribute to the understanding of the matter-antimatter asymmetry. Thus, the measurement of CP violation is an important part of research in particle physics.

Measurements of CP violation can be performed in B meson decays, which provide access to observables probing elements of the CKM matrix, such as the CP -violating weak phases β and ϕ_s . The most prominent analyses studying these parameters are performed in $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi \phi$ decays [17, 18]. However, the two parameters can also be measured in $B^0 \rightarrow D^{(*)+} D^{(*)-}$, $B_s^0 \rightarrow D^{(*)+} D^{(*)-}$ and $B_s^0 \rightarrow D_s^{(*)+} D_s^{(*)-}$ decays [19–26]. In the $B_s^0 \rightarrow D^{*+} D^{*-}$ and $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$ decays, CP violation has not been measured yet. Measurements in these decays are polluted by higher-order Standard Model effects [27–31]. It is important to determine these contributions to constrain them in the determination of ϕ_s [32–35], since they represent one of the largest uncertainties [36, 37]. This can be achieved by studying CP violation parameters or ratios of branching fractions in additional decays of the $B \rightarrow DD$ family [38, 39]. The measurements of the $B_s^0 \rightarrow D^{*+} D^{*-}$ branching fraction [40] and the CP violation in $B^0 \rightarrow D^{*+} D^{*-}$ decays are presented in this thesis. In contrast to the $B^0 \rightarrow D^{*+} D^{*-}$ decay, the $B_s^0 \rightarrow D^{*+} D^{*-}$

decay has not been observed yet. It has a low predicted branching fraction [41] but gives access to higher-order SM processes such as W -exchange and penguin-annihilation transitions. Besides constraining the higher order effects in CP violation measurements, the search for predicted decays and the validation of their predicted branching fractions also represents a test of the SM. The $B^0 \rightarrow D^{*+}D^{*-}$ decay allows to study the CKM parameter β . However, the final state is not a CP eigenstate and the CP parameters are diluted due to the interference of the two eigenstates. By determining the fraction of the CP eigenstates, the polarisation fraction can be calculated, which can be used to test SM predictions [30]. Additionally, these fractions are interesting because partially reconstructed $B^0 \rightarrow D^{*+}D^{*-}$ decays represent backgrounds in other more sensitive analyses [24, 42]. A more precise determination of the CP -odd fraction allows for a better description of this background.

The Large Hadron Collider (LHC) offers a well-suited environment for these measurements. It is located at the *European Organization for Nuclear Research* (CERN) near Geneva and is currently the largest and most powerful particle accelerator in the world. At the LHC, protons are collided at centre-of-mass energies of up to 14 TeV. One experiment at the LHC is the Large Hadron Collider beauty (LHCb) experiment, which provides the data analysed in this thesis. It focuses on precision measurements in decays of hadrons involving b and c quarks. The LHCb experiment also provides the largest data set of B_s^0 mesons and, therefore, enables the measurement of $B_s^0 \rightarrow D^{*+}D^{*-}$ decay. Overall, the data from this experiment allows to constrain the SM with measurements in the sector of B meson decays, such as the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays.

In this thesis, first, an overview of the SM and CP violation is provided in Chap. 2. Afterwards, the environment of the LHC and the LHCb experiment are described in Chap. 3. An outline of the statistical methods used in this work is given in Chap. 4. Both the measurement of the branching fraction of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay and the measurement of the CP violation in $B^0 \rightarrow D^{*+}D^{*-}$ decays are presented. The strategy for both analyses is formulated in Chap. 5. The selection of the two decays is then described in Chap. 6. The measurement of the branching fraction of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay is presented in Chap. 7. Afterwards, the measurement of CP violation in $B^0 \rightarrow D^{*+}D^{*-}$ decays is discussed in Chap. 8. While the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis has already been published in Ref. [40], the $B^0 \rightarrow D^{*+}D^{*-}$ analysis is still in internal review of the LHCb collaboration. Finally, this thesis concludes with a summary of all results presented and an outlook on future work in Chap. 9.

2 Standard Model of particle physics

The theoretical foundations related to the analyses presented in this thesis are discussed in this chapter. Furthermore, an overview of the current experimental status is given. If not further specified, the information outlined in this chapter is mainly based on Refs. [43–46].

The understanding of nature at the smallest possible level has changed fundamentally over the last century with the development of the SM. The SM is a well-tested theory and can predict almost all experimentally observed phenomena. It describes the fundamental particles and their interactions and is outlined in Sec. 2.1. The introduction of the SM is followed by a discussion of the CKM mechanism in Sec. 2.2. Sec. 2.3 elaborates CP violation in the SM in more detail, while Sec. 2.4 then focuses on the explanation of neutral meson mixing in the B^0 meson sector. In Sec. 2.5, the decays analysed in this thesis are introduced. Further, the theoretical foundations are applied to the framework of $B \rightarrow D^{*+}D^{*-}$ decays. First, the calculation of the branching fractions is discussed in Sec. 2.5.1, which addresses the first measurement of this thesis. Afterwards, Sec. 2.5.2 focusses on the CP violation in $B^0 \rightarrow D^{*+}D^{*-}$ decays, which gives the theoretical foundation for the second measurement performed in this thesis. The presence of a non- CP eigenstate as a final state introduces further complexity in the CP violation measurements. The consequences for the measurement in $B^0 \rightarrow D^{*+}D^{*-}$ decays are discussed in Sec. 2.5.3. Despite being the most successful theory, the SM cannot describe all observed phenomena. Therefore, Sec. 2.6 discusses its limitations.

2.1 Particles and interactions

The SM is a Lorentz invariant quantum field theory describing the elementary particles and their interactions. It has a total of 18 free parameters that must be determined experimentally. The elementary particles included in SM are shown in Fig. 2.1 with their properties such as mass, spin, and charge. The spin allows the particles to be categorised into two groups called fermions and bosons. The fermions carry half-integer spin while the bosons carry integer spin. The characteristics of the particles of each group and the interactions included in the SM are discussed in the following.

Fermions The fermions are the ingredients of all matter. They carry spin 1/2 and obey the Fermi-Dirac statistics. In total, 12 fermions exist. These can be further categorised as leptons l and quarks q , each including 6 particles. Their antiparticles are also considered



Figure 2.1 – Particles described by the SM. Each particle is shown with its mass, charge and spin. The quarks, leptons, gauge bosons and the Higgs boson are illustrated in red, green, blue and orange, respectively. The strong, electromagnetic (e.m.) and weak forces are indicated by lines showing the particles they interact with. The values are taken from Ref. [47].

as fermions and are included in the SM. They can be obtained by inverting the sign of each quantum number for each particle.

Leptons and quarks can both be divided into three families. Each lepton family consists of a charged and a neutral lepton. The charged leptons are the electron (e), the muon (μ), and the tau (τ), all carrying the charge $Q = -1e$, where e is the elementary charge. The neutral leptons with $Q = 0$ are the corresponding neutrinos, the electron neutrino (ν_e), the muon neutrino (ν_μ) and the tau neutrino (ν_τ). Differences between the individual leptons are the lepton family numbers and the masses of the particles. A strong mass hierarchy exists between the families, where the neutrinos are massless in the SM.

The quarks can also be divided into three families. Each family consists of an up-type and a down-type quark. The up-type quarks comprise the up quark (u), the charm quark (c) and the top quark (t). They carry the charge $Q = 2e/3$. In comparison, the down-type quarks carry the charge $Q = -1e/3$. This group comprises the down quark

(*d*), the strange quark (*s*) and the bottom quark (*b*). The individual quarks differ in their flavour quantum number and their masses, where a certain mass hierarchy is present between the families. In contrast to the leptons, no free quarks have been observed. As a consequence, the model of asymptotic freedom has been developed. According to this model, the quarks only behave freely at large energy scales, *i.e.* at small distances below 10^{-15} m. Above these distances, quarks only appear in bound states called hadrons. In addition to this phenomenon, the predecessor version of SM, also known as the early quark model [48, 49], introduced an issue with the postulation of the Δ^{++} particle. This particle is a bound state of three *u* quarks and has a spin of $3/2$, which, according to the model at that time, meant that all three quark spins are parallel. This is prohibited due to the Pauli principle. By introducing another quantum number for the quarks called the colour charge, the problem was solved. This colour charge can assume the values red, blue or green. So far, only colour-neutral hadrons have been observed. This phenomenon is called quark confinement [50].

The most common bound states of quarks are mesons and baryons. Mesons are formed by a $q\bar{q}$ pair, where both quarks carry the same colour charge. These particles are particularly interesting since both analyses performed in this thesis study decays of mesons either containing a *b* or *s* quark paired with a *d* quark. The respective particles are called B^0 and B_s^0 mesons. Baryons are formed by three quarks (qqq), where all quarks carry different colour charges. Further bound states that have been experimentally confirmed are tetraquarks and pentaquarks. Tetraquarks are states consisting of four quarks ($qq\bar{q}\bar{q}$) and were observed by the Belle experiment [51] and by the LHCb experiment [52–54]. The LHCb experiment also observed pentaquark states for the first time [55, 56], which consist of five quarks ($qqq\bar{q}\bar{q}$).

Bosons The bosons can be divided into a group of gauge bosons and the Higgs boson. The gauge bosons have spin $J = 1$ and mediate the fundamental forces described in the SM. The group comprises the photon (γ), the gluons (g), the W^\pm bosons and the Z^0 boson. The photon is the mediator of the electromagnetic force described by the theory of quantum electrodynamics (QED). It is massless and couples to the electric charge. The gluons mediate the strong force described by quantum chromodynamics (QCD). They are also massless but couple to the colour charge. In addition to coupling to quarks, gluons can couple to themselves since they carry a colour and an anticolour charge. Eight gluons exist, all with a different colour configuration but an electric charge of $Q = 0$. The remaining gauge bosons, the electrically charged W^\pm and neutral Z^0 bosons, mediate the weak force. The Z^0 bosons are responsible for the neutral currents and couple to all fermions. In contrast, the W^\pm bosons couple to charged weak currents, *i.e.* to left-handed fermions and right-handed antifermions. In comparison to the photon and the gluons, the W^\pm and Z^0 bosons have a finite lifetime due to their masses of $m_{W^\pm} = 80.377 \text{ GeV}/c^2$ and $m_{Z^0} = 91.1876 \text{ GeV}/c^2$ [47].

In contrast to the gauge bosons, the Higgs boson (H^0) is spinless, *i.e.* $J = 0$. It has a mass of $m_{H^0} = 125.25 \text{ GeV}/c^2$ [47], but does not carry an electric or colour charge. In general, the Higgs completes the SM. It was already postulated in 1964 [57, 58], but was only discovered in 2012 by ATLAS [7] and CMS [8]. The postulation of the Higgs boson originates from the introduction of a complex, scalar and massless field that should be invariant under a local gauge transformation and the study of the case of spontaneous symmetry breaking. This is also known as the Higgs mechanism. While the Higgs mechanism also gives the masses to the gauge bosons, the masses of the fermions do not directly originate from the mechanism. They arise from a Yukawa interaction [59] of the fermions with the Higgs field.

Symmetry group All fundamental interactions of the SM can be generated by introducing a global gauge invariance that should hold locally under certain transformations. This is also known as Yang-Mills theory, after C.-N. Yan and R. L. Mills, who were the first to apply it [60]. The electromagnetic interaction is generated by the requirement of a local gauge invariance under a $U(1)$ transformation of the free Dirac Lagrangian. For the strong interaction, the Lagrangian considers particles with three different colour charge states but the same mass. This Lagrangian must be invariant under a $SU(3)$ transformation. Accordingly, the symmetry group of the strong interaction is called $SU(3)_C$. The symmetry group of the weak interaction results from writing leptons and quarks as doublets in their respective families, which are characterised by the eigenvalues of the weak isospin. A Lagrangian considering these doublets must be invariant under a $SU(2)$ transformation. However, the electromagnetic and weak interactions can be unified into a combined electroweak symmetry group [1–3], called isospin-hypercharge symmetry $SU(2)_L \times U(1)_Y$. This was proposed in the 1960s and is also known as the Glashow-Weinberg-Salam model [61–63]. With the discovery of neutral currents by the Gargamelle experiment at Cern in 1973, this model has been experimentally confirmed [64]. Considering the unification of the electromagnetic and weak interaction, the total symmetry group of the SM can be written as

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \quad (2.1)$$

Symmetries The SM fulfils the *CPT* theorem resulting from quantum field theory, which requires invariance of any interaction under a *CPT* transformation [65, 66]. This transformation comprises several individual discrete transformations. One of them is the time reversion T , *i.e.* $T(t, \vec{x}) = (-t, \vec{x})$. The second transformation is the parity transformation P , which inverts the spatial coordinates in the form of $P(t, \vec{x}) = (t, -\vec{x})$. The last transformation inverts the sign of the charges and is the so-called charge conjugation C . Although every interaction in the SM is invariant under the *CPT* transformation, symmetries can be violated when applying one individual transformation or a combination of two transformations. For example, while the strong and electromagnetic interactions are invariant under C or P transformation, the weak interaction is maximal

parity violating [67, 68] and is not symmetric under C . Also, the combined CP symmetry is violated in the weak interaction, shown for the kaon sector in 1964 by J. W. Cronin and V. L. Fitch [69]. In the B^0 meson sector, CP violation was observed in 2001 by the Belle [70] and BaBar experiments [71].

2.2 The CKM matrix

The historical origin of the CKM matrix dates back to 1963. In this year, N. Cabibbo described the rates of $d \rightarrow uW^-$ and $s \rightarrow uW^-$ transitions by introducing the factors $\cos(\theta_C)$ and $\sin(\theta_C)$ with the Cabibbo angle $\theta_C \approx 13.15^\circ$ [15], respectively. At this time, the only known quarks were the up, down and strange quarks. The theory successfully predicted many decay rates consistent with the experimentally measured rates. However, it did not solve the puzzle observed in $K^0 \rightarrow \mu^+\mu^-$ decays. In this decay channel, the measured experimental limit was far smaller than the calculated rate. Glashow, Iliopoulos and Maiani proposed a solution for this issue in 1970 by introducing an additional fourth quark [72], *i.e.* the charm quark. In the case of the $K^0 \rightarrow \mu^+\mu^-$ decay, the charm quark contributes with destructive interference, causing a significant reduction of the rate. This is also known as the GIM mechanism. The charm quark was experimentally confirmed in 1974 with the observation of $J/\psi(c\bar{c})$ [73, 74]. Combining the GIM mechanism with Cabibbo's approach, a simple scheme was introduced to describe the mixing of the quarks, where the eigenstates of the weak interaction are rotated to the mass eigenstates. The scheme was later extended to three generations by Kobayashi and Maskawa in 1973, postulating two new quarks [16]. The resulting 3×3 rotation matrix is known as the CKM matrix, which allowed for the description of CP violation in the SM. With the discovery of the beauty quark in 1977 [75] and the top quark in 1995 [76, 77], the third generation has been confirmed experimentally.

The theoretical basis for the CKM matrix is the manual introduction of the fermion masses in the SM. These masses do not directly originate from the Higgs mechanism but are derived by a Yukawa coupling of the fermions with the Higgs field ϕ , which is introduced with the electroweak unification. Assuming massless neutrinos, the Lagrangian of the interaction is defined as

$$\mathcal{L}_{\text{Yukawa}} = \hat{Y}_{ij}^e \bar{L}_L^i \phi e_R^j + \hat{Y}_{ij}^d \bar{Q}_L^i \phi d_R^j + \hat{Y}_{ij}^u \bar{Q}_L^i \phi c u_R^j + \text{h.c.}, \quad (2.2)$$

where

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \quad \phi_c = i\sigma_2 \phi^*, \quad (2.3)$$

and where $h(x)$ is the physical Higgs field and $\hat{Y}_{ij}^{e,d,u}$ are the 3×3 complex Yukawa matrices. The observables L_L^i and Q_L^i are the left-handed lepton and quark doublets, while e_R^j , d_R^j and u_R^j are the right-handed singlets. By substituting v in Eq. (2.2) with

the vacuum expectation value $(\phi, \langle \phi \rangle) = (0, v/\sqrt{2})$ the mass terms of the quarks and leptons are derived. The non-Yukawa parts of the SM Lagrangian for the leptons are invariant under the transformations S and R of the $U(3)$ group so that $Y^e = S\hat{Y}^e R^\dagger$. Using a suitable choice of S and R , the Yukawa matrix Y^e can be selected as real, diagonal, and non-negative. In contrast, this is not straightforward for the quarks. The non-Yukawa part is also invariant under the $U(3)$ transformations S_u , R_d and R_u , where S_u and R_u can be chosen so that $Y^u = S_u\hat{Y}^u R_u^\dagger$ is real, diagonal and non-negative. However, the result of the transformation of $S_u\hat{Y}^d R_d^\dagger$ is not real and not diagonal, *i.e.*

$$S_u\hat{Y}^d R_d^\dagger = S_u S_d^\dagger S_d \hat{Y}^d R_d^\dagger. \quad (2.4)$$

Instead, Y_d is real, diagonal and non-negative with $Y^d = S_d\hat{Y}^d R_d^\dagger$. The matrix product $S_u S_d^\dagger$ in Eq. (2.4) is defined as the CKM matrix V_{CKM} . It follows that the matrix rotates the mass eigenstates of the quarks to the so-called flavour eigenstates, *i.e.*

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.5)$$

The CKM matrix can also be identified in the Lagrangian describing charged-current interactions, which are mediated by the W^\pm bosons. The weak currents do not couple to the quarks' mass eigenstates but to their flavour eigenstates. The Lagrangian is given by

$$\mathcal{L} \propto \frac{g}{\sqrt{2}} (u_L \quad c_L \quad t_L) \gamma^\mu W_\mu^+ V_{\text{CKM}} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.}, \quad (2.6)$$

where g is the weak coupling constant and γ^μ are the Dirac matrices. Since the transition probability of an up-type quark into a down-type quark is proportional to the squared matrix elements $|V_{ij}|^2$, the matrix describes the mixing of the quark generations. For $N = 3$ quark generations, the 3×3 CKM matrix can be expressed by $(N - 1)^2 = 4$ real parameters. For example, the three mixing angles θ_{12} , θ_{13} and θ_{23} and a complex phase δ can be used to parametrise the CKM matrix. This phase is the anchor of CP violation in the SM. Using these parameters and the abbreviations $s_{ij} = \sin(\theta_{ij})$ and $c_{ij} = \cos(\theta_{ij})$, the CKM matrix can be written as

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (2.7)$$

which is the so-called standard parametrisation of the CKM matrix. It has been developed by L.-L. Chau and W.-Y. Keung [78]. Another common parametrisation is the Wolfenstein parametrisation [79], which shows the strong hierarchy of the CKM matrix.

By introducing the relations

$$\lambda = s_{12}, \quad A = \frac{s_{23}}{\lambda}, \quad \rho + i\eta = \frac{s_{13}e^{i\delta}}{A\lambda^3}, \quad (2.8)$$

the CKM can be written as

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (2.9)$$

A value of $\lambda = 0.22500 \pm 0.00067$ [47] has been determined by combining several experimental measurements. By substituting λ in Eq. (2.9) with the measured value, the strong hierarchy of the CKM matrix is directly visible.

According to its construction, the CKM matrix is unitary. As a consequence, the matrix has to fulfil the following relations

$$\sum_{k=1}^3 V_{ki}V_{kj}^* = \delta_{ij}, \quad \text{with } i, j \in 1, 2, 3. \quad (2.10)$$

These equations are also known as the unitarity equations of the CKM matrix. In total, nine unitarity equations exist, but only six of them vanish ($i \neq j$). These conditions can be drawn in the complex plane, resulting in the so-called unitarity triangles. All of these triangles share the same area $\mathcal{J}/2$ with

$$\mathcal{J} \sum_{m,n} \epsilon_{ikm}\epsilon_{jln} = \Im(V_{ij}V_{kl}V_{il}^*V_{kj}^*), \quad (2.11)$$

where the observable \mathcal{J} is known as the Jarlskog invariant [80]. It is a phase convention independent observable and a measure for the CP violation described in the SM. The experimentally measured value is $\mathcal{J} = (3.08_{-0.13}^{+0.15}) \times 10^{-5}$ [47]. Most unitarity triangles are degenerated, and only two provide sides of similar lengths. One of them is based on the condition

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (2.12)$$

It is the most common triangle and presented in Fig. 2.2, where the sides are normalised to the most precisely known elements $V_{cd}V_{cb}^*$. The lengths of the triangle are defined as

$$R_a = 1, \quad R_b = \left| \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right|, \quad R_t = \left| \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} \right| \quad (2.13)$$

and the angles are given by

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \quad (2.14)$$

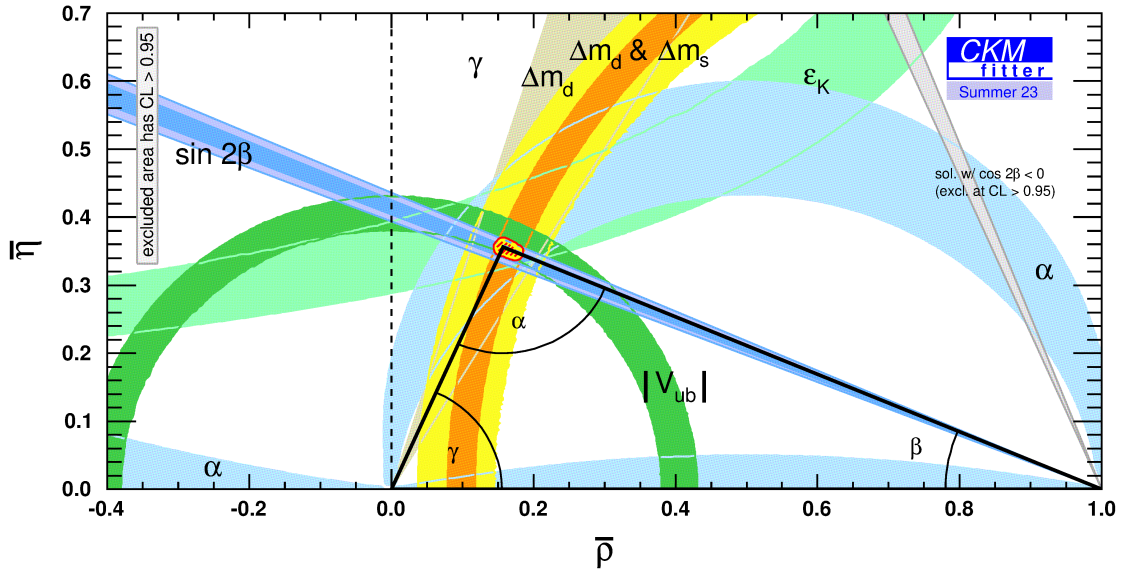


Figure 2.2 – Unitarity triangle of most common unitarity condition given in Eq. (2.12) drawn in the $\bar{\rho}$ - $\bar{\eta}$ -plane [81]. The experimental constraints are shown in coloured bands. No signs of New Physics have been observed since the current constraints agree with a triangle. The constraint on the apex of the triangle is shown in red.

A non-closing triangle would be a sign of New Physics (NP). Consequently, precision measurements of the properties of the triangles are a strong intrinsic test of the SM. Several measurements of the different properties of the unitarity triangle defined by Eq. (2.12) have been performed. Testing this condition is especially suitable since all parameters are accessible by studying B meson decays. Combinations of the results of the measurements are provided by the CKMfitter group [81], which are presented in Fig. 2.2. The current results are consistent, and no sign of NP has been observed so far.

2.3 The system of neutral B mesons

In the SM, quarks can change their flavour via charged currents. This mixing is described by the CKM matrix discussed in Sec. 2.2. Due to this mixing, it is also possible that neutral particles such as the K^0 , D^0 , B^0 , or B_s^0 mesons can transition into their own antiparticles and vice versa. Consequently, the particles can also oscillate between their flavour eigenstates. This is also known as neutral meson mixing. Such oscillations were first observed in the neutral kaon sector [69]. The B^0 system is the most prominent example of this mixing. In this system, the oscillation has been observed for the first time in 1987 [82, 83]. However, the same mathematical description can be applied to all systems. The Feynman diagrams of the B^0 - \bar{B}^0 oscillation are shown in Fig. 2.3.

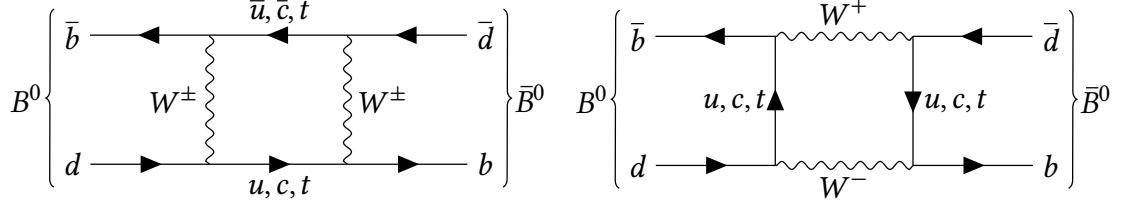


Figure 2.3 – Leading-order Feynman diagrams of the $B^0 - \bar{B}^0$ oscillation.

At the lowest order, the $B^0 - \bar{B}^0$ oscillation is mainly induced by the so-called box diagrams. Since this system is also particularly interesting for this thesis, it is discussed in more detail in the following. The B^0 and \bar{B}^0 mesons are bound states of a \bar{b} and a d or a b and a \bar{d} quark, respectively, *i.e.*

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

The Schrödinger equation describes the time evolution of the B^0 and \bar{B}^0 mesons and is given by

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

where \mathbf{M} and $\mathbf{\Gamma}$ are Hermitian 2×2 matrices. They are defined as the mass and decay matrices, respectively. As a consequence of the CPT invariance, the particles and their respective antiparticles are required to have the same mass and decay width, *i.e.*

$$m = \mathbf{M}_{11} = \mathbf{M}_{22}, \quad \Gamma = \mathbf{\Gamma}_{11} = \mathbf{\Gamma}_{22}. \quad (2.17)$$

The off-diagonal elements are induced by the box diagrams in Fig. 2.3 and are responsible for the $B^0 - \bar{B}^0$ oscillations. By diagonalising the Hamiltonian $\mathcal{H} = \mathbf{M} - \frac{i}{2} \mathbf{\Gamma}$ in Eq. (2.16), the eigenstates B_L and B_H are obtained. They are defined as

$$|B_L\rangle = p|B^0\rangle + q|\bar{B}^0\rangle, \quad |B_H\rangle = p|B^0\rangle - q|\bar{B}^0\rangle \quad (2.18)$$

with $|q|^2 + |p|^2 = 1$. These states are called the light (L) and heavy (H) eigenstates, and their time-evolution is given by

$$|B_{L,H}(t)\rangle = e^{-(im_{L,H} + \Gamma_{L,H}/2)t} |B_{L,H}\rangle, \quad (2.19)$$

where m_L and m_H are the respective masses and Γ_L and Γ_H are the respective decay widths. By defining the average mass and decay width

$$m = \frac{m_L + m_H}{2} = \mathbf{M}_{11}, \quad \Gamma = \frac{\Gamma_L + \Gamma_H}{2} = \mathbf{\Gamma}_{11} \quad (2.20)$$

and the differences

$$\Delta m = m_H - m_L, \quad \Delta \Gamma = \Gamma_H - \Gamma_L \quad (2.21)$$

the fraction of p over q is calculated to be

$$\frac{q}{p} = \frac{\Delta m + \frac{i}{2}\Delta\Gamma}{2(M_{12} - \frac{i}{2}\Gamma_{12})} = \frac{2(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}{\Delta m + \frac{i}{2}\Delta\Gamma}. \quad (2.22)$$

In the B^0 system, the mixing shown in the box diagrams in Fig. 2.3 is dominated by the top quark so that the q/p can be approximated by

$$\frac{q}{p} \approx \frac{V_{td}V_{tb}^*}{V_{tb}V_{td}^*}. \quad (2.23)$$

The time evolution of the flavour eigenstates B^0 and \bar{B}^0 is calculated by inverting the states in Eq. (2.18) and substituting B_L and B_H with Eq. (2.19), *i.e.*

$$\begin{aligned} |B^0(t)\rangle &= \frac{1}{2p} \left[e^{-(im_L + \Gamma_L/2)t} |B_L\rangle + e^{-(im_H + \Gamma_H/2)t} |B_H\rangle \right], \\ |\bar{B}^0(t)\rangle &= \frac{1}{2q} \left[e^{-(im_L + \Gamma_L/2)t} |B_L\rangle - e^{-(im_H + \Gamma_H/2)t} |B_H\rangle \right]. \end{aligned} \quad (2.24)$$

By introducing the time-dependent coefficients

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

the time evolution of the B^0 and \bar{B}^0 mesons can be expressed with the initial flavour, *i.e.*

$$|B^0(t)\rangle = g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\bar{B}^0\rangle, \quad |\bar{B}^0(t)\rangle = \frac{p}{q}g_-(t)|B^0\rangle + g_+(t)|\bar{B}^0\rangle. \quad (2.26)$$

The B^0 mesons are difficult to access experimentally. However, by measuring time-dependent decay rates of a B^0 or \bar{B}^0 meson into a final state f , the CP parameters can be quantified. The corresponding decay amplitudes are defined as

$$A_f = \langle f|B^0\rangle, \quad \bar{A}_f = \langle f|\bar{B}^0\rangle. \quad (2.27)$$

In addition to these amplitudes, a commonly defined quantity is

$$\lambda_f = \frac{1}{\lambda_f} = \frac{q \bar{A}_f}{p A_f}. \quad (2.28)$$

In general, the time-dependent decay rates are given by

$$\Gamma(B^0(t) \rightarrow f) = U_f |\langle f | B^0(t) \rangle|^2, \quad \Gamma(\bar{B}^0(t) \rightarrow f) = U_f |\langle f | \bar{B}^0(t) \rangle|^2, \quad (2.29)$$

where U_f is a time-independent normalisation factor. With the definitions given in Eqs. (2.26) to (2.28), the time-dependent decay rates are calculated to be

$$\begin{aligned} \Gamma_{B^0(t) \rightarrow f} &\propto e^{-\Gamma t} |A_f|^2 (1 + |\lambda_f|^2) \\ &\cdot \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + A_f^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma t}{2}\right) - S_f \sin(\Delta m t) + C_f \cos(\Delta m t) \right], \\ \Gamma_{\bar{B}^0(t) \rightarrow f} &\propto e^{-\Gamma t} |A_f|^2 (1 + |\lambda_f|^2) \left|\frac{p}{q}\right|^2 \\ &\cdot \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) + A_f^{\Delta\Gamma} \sinh\left(\frac{\Delta\Gamma t}{2}\right) + S_f \sin(\Delta m t) - C_f \cos(\Delta m t) \right], \end{aligned} \quad (2.30)$$

where $A_f^{\Delta\Gamma}$, S_f and C_f are the CP parameters defined as

$$A_f^{\Delta\Gamma} = -\frac{2\text{Re}(\lambda_f)}{1 + |\lambda_f|^2}, \quad S_f = \frac{2\text{Im}(\lambda_f)}{1 + |\lambda_f|^2}, \quad C_f = \frac{1 - |\lambda_f|^2}{1 + |\lambda_f|^2}. \quad (2.31)$$

These parameters fulfil the condition

$$(A_f^{\Delta\Gamma})^2 + (S_f)^2 + (C_f)^2 = 1. \quad (2.32)$$

They absorb all CP -violating effects in the time-dependent decay rates defined in Eq. (2.30). Consequently, measuring these rates allows to quantify the CP violation in the SM experimentally. The results of these measurements can be used to constrain the CKM matrix and to search for NP.

2.4 Types of CP violation

The measurement of CP violation is one of the essential tests of the SM. Measurements are performed exploiting various decays of B , K and D mesons. Depending on the decay channel, different manifestations of CP violation can be observed. The SM distinguishes three types of CP violation. These are discussed in more detail in this chapter. First, Sec. 2.4.1 discusses CP violation arising from different decay amplitudes of the decay and the CP conjugated decay. Second, indirect CP violation is discussed in Sec. 2.4.2. This type occurs when the mass eigenstates are not the CP eigenstates. Third, CP violation originating from the interference of direct decays and decays after mixing is discussed in Sec. 2.4.3. Since the second and third types require neutral meson mixing, they are only possible for neutral decays. However, the first type can be observed for charged and neutral mesons. In general, most decays only show one type of CP violation, and it can be concluded that $\lambda_f \neq 1$ is required to allow for CP violation.

2.4.1 Direct CP violation

Direct CP violation is a result of the different decay amplitudes of a decay $B^0 \rightarrow f$ with A_f and its CP conjugated decay $\bar{B}^0 \rightarrow \bar{f}$ with $\bar{A}_{\bar{f}}$. This type is commonly referred to as CP violation in the decay, and the asymmetry A_{CP} is defined as

$$A_{CP} = \frac{\Gamma(B^0 \rightarrow f) - \Gamma(\bar{B}^0 \rightarrow f)}{\Gamma(B^0 \rightarrow f) + \Gamma(\bar{B}^0 \rightarrow f)} = \frac{|\bar{A}_{\bar{f}}/A_f| - 1}{|\bar{A}_{\bar{f}}/A_f| + 1}. \quad (2.33)$$

Two complex phases enter the decay amplitudes, the strong phase δ and the weak phase ϕ . The strong phase arises from possible rescattering processes of intermediate states into the final state. Strong interactions usually dominate these rescattering processes. Consequently, the strong phase is invariant under CP transformation. The weak phase is induced by complex CKM matrix elements and changes its sign under CP transformation. With the weak and strong phase, the decay amplitudes can be written as

$$A_f = \sum_k^n A_k e^{i(\delta_k + \phi_k)}, \quad \bar{A}_{\bar{f}} = \sum_k^n A_k e^{i(\delta_k - \phi_k)}, \quad (2.34)$$

where n is the number of processes contributing to the decay amplitude. At least two processes are required to allow for CP violation, which differ in their weak and strong phases. From Eq. (2.33), it can be concluded that a non-vanishing asymmetry and consequently CP violation in the decay arises if

$$|A_f| \neq |\bar{A}_{\bar{f}}|. \quad (2.35)$$

This type of CP violation has already been observed and well studied in the sector of K and B mesons [84–86]. It has also been observed later in the D meson sector [87].

2.4.2 Indirect CP violation

Indirect CP violation is also called CP violation in the mixing. It occurs when the transition probabilities between the B^0 and the \bar{B}^0 mesons in the oscillation are different, *i.e.*

$$P(B^0 \rightarrow \bar{B}^0) \neq P(\bar{B}^0 \rightarrow B^0). \quad (2.36)$$

This phenomenon can be associated with the coefficients q and p . If the ratio of these coefficients equals 1, the transition probabilities are equal. Otherwise, the transition probabilities are different. With Eq. (2.22) the squared ratio of $|q/p|$ is calculated to be

$$\left| \frac{q}{p} \right|^2 = \left| \frac{2(M_{12}^* - \frac{i}{2}\Gamma_{12}^*)}{2(M_{12} - \frac{i}{2}\Gamma_{12})} \right|. \quad (2.37)$$

From this ratio, it can be derived that the ratio is not equal to 1 if a relative phase between M_{12} and Γ_{12} is present. This generally occurs when the mass eigenstates are not the CP eigenstates.

Indirect CP violation has already been observed in the kaon system [69]. However, in the $B_{(s)}^0$ system, indirect CP violation is expected to be small [88]. Measurements in this system are performed using flavour-specific semileptonic decays, where the asymmetry is defined as

$$A_{CP}^{\text{mix}} = \frac{\Gamma(B^0(t) \rightarrow \ell^- \nu_\ell X^+) - \Gamma(\bar{B}^0(t) \rightarrow \ell^+ \bar{\nu}_\ell X^-)}{\Gamma(B^0(t) \rightarrow \ell^- \nu_\ell X^+) + \Gamma(\bar{B}^0(t) \rightarrow \ell^+ \bar{\nu}_\ell X^-)} = \frac{1 - |q/p|^4}{1 + |q/p|^4}. \quad (2.38)$$

The results of these measurements are consistent with 0, although the precision is limited [89, 90]. This also confirms the approximation of the $|q/p|$ ratio in Eq. (2.23).

2.4.3 CP violation in interference

Another type of CP violation is CP violation in the interference. It can occur if the B^0 and the \bar{B}^0 decay into the same CP eigenstate f_{CP} . In this case, the decay amplitudes of the direct decay $B^0 \rightarrow f_{CP}$ and the decay after mixing $B^0 \rightarrow \bar{B}^0 \rightarrow f_{CP}$ interfere. To account for the decay into a CP eigenstate, the observable λ_f is redefined as

$$\lambda_{CP} = \eta_{CP} \frac{q \bar{A}_f}{p A_f}, \quad (2.39)$$

where η_{CP} is the CP eigenvalue of the final state f_{CP} . The CP eigenvalue is defined by the transformation of the final state under CP transformation, *i.e.*

$$CP|f_{CP}\rangle = |\bar{f}_{CP}\rangle = \eta_{CP}|f_{CP}\rangle = \pm 1|f_{CP}\rangle. \quad (2.40)$$

With the approximation $|q/p| \approx 1$ the decay-time dependent asymmetry is calculated to be

$$\begin{aligned} A_{CP}^{\text{int}} &= \frac{\Gamma(B^0(t) \rightarrow f) - \Gamma(\bar{B}^0(t) \rightarrow f)}{\Gamma(B^0(t) \rightarrow f) + \Gamma(\bar{B}^0(t) \rightarrow f)} \\ &= -\frac{(1 - |\lambda_{CP}|^2) \cos(\Delta mt) - 2\Im(\lambda_{CP}) \sin(\Delta mt)}{(1 + |\lambda_{CP}|^2) \cosh\left(\frac{\Delta\Gamma t}{2}\right) - 2\Re(\lambda_{CP}) \sinh\left(\frac{\Delta\Gamma t}{2}\right)}. \end{aligned} \quad (2.41)$$

The asymmetry is simplified by using the definitions of the CP observables given in Eq. (2.31) and by assuming $\Delta\Gamma_d = 0$. The resulting modified asymmetry is given by

$$A_{CP}^{\text{int}}(t) = \eta_{CP} S_f \sin(\Delta mt) - C_f \cos(\Delta mt). \quad (2.42)$$

Consequently, CP violation in the interference arises if $|\lambda_{CP}| \neq \pm 1$ or if the imaginary part of λ_{CP} is non-zero. Since this type of CP violation can also occur in the absence of

direct and indirect CP violation, it is associated with

$$\Im(\lambda_{CP}) \neq 0. \quad (2.43)$$

The most prominent examples of measuring CP violation originating from interference between direct decay and decay after mixing are the measurements in $B^0 \rightarrow J/\psi K_S^0$ and $B_s^0 \rightarrow J/\psi \phi$ decays [17, 18]. Both were performed by the LHCb experiment and contributed significantly to the world averages of the CP observables, allowing the CKM parameters to be further constrained.

2.5 The $B \rightarrow D^{*+} D^{*-}$ decays

The analyses performed in this thesis aim to measure several observables in the framework of $B \rightarrow D^{*+} D^{*-}$ decays. This section focuses on discussing these decays and the observables of interest. The Feynman diagrams of both decays are presented in Fig. 2.4.

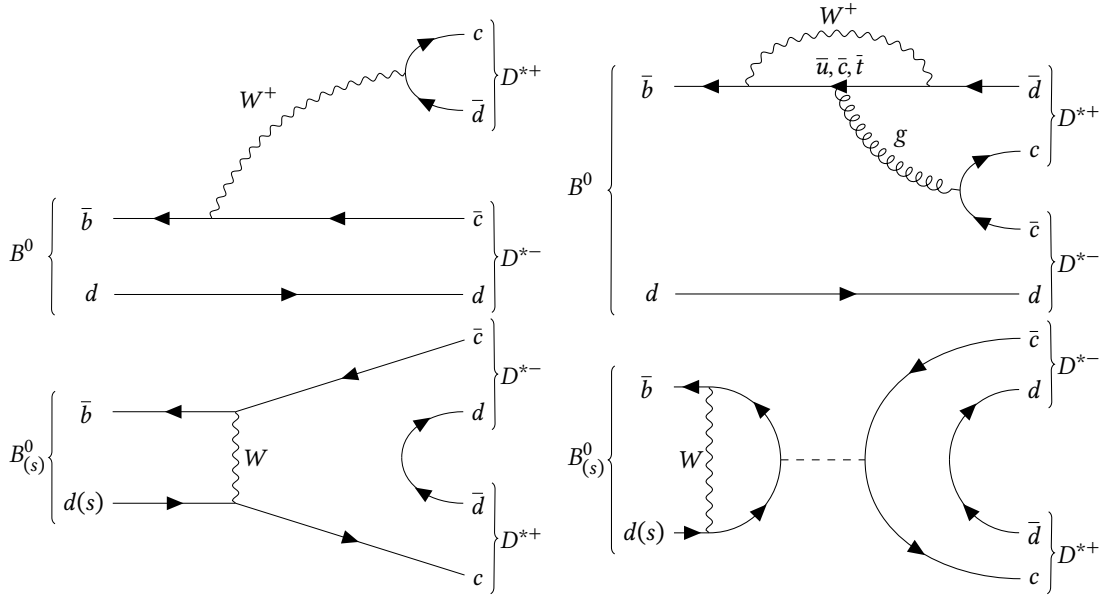


Figure 2.4 – Dominant Feynman diagrams of the $B_{(s)}^0 \rightarrow D^{*+} D^{*-}$ decays, *i.e.* the (top left) tree level, the (top right) penguin, the (bottom left) W -exchange and the (bottom right) penguin annihilation diagrams. The $B_s^0 \rightarrow D^{*+} D^{*-}$ is suppressed compared to the $B^0 \rightarrow D^{*+} D^{*-}$ decay since it is only possible via the W -exchange or the penguin annihilation diagrams.

The dominant contribution to the $B^0 \rightarrow D^{*+} D^{*-}$ decay is the tree level (T) diagram which is characterised by $b \rightarrow \bar{c} d$ transitions. Furthermore, penguin (P), W -exchange (E) and penguin annihilation (PA) diagrams contribute to the decay. It has been observed by the CLEO experiment [91] but was also measured by the Belle and BaBar

experiments [19, 92]. The current world average its branching fraction is given by

$$\mathcal{B}(B^0 \rightarrow D^{*+}D^{*-}) = (8.0 \pm 0.6) \cdot 10^{-4}, \quad (2.44)$$

which is taken from Ref. [47]. The $B_s^0 \rightarrow D^{*+}D^{*-}$ decay is suppressed compared to the $B^0 \rightarrow D^{*+}D^{*-}$ decay since it is only possible via the W -exchange or the penguin annihilation diagrams. This decay has not been observed so far. Its absolute branching fraction is predicted to be $(3.1_{-1.1}^{+1.3}) \times 10^{-4}$, where a perturbative QCD approach based on k_T factorisation is used in the calculation [41].

This thesis aims to measure the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay and its branching fraction relative to the $B^0 \rightarrow D^{*+}D^{*-}$ decay. The definition of this ratio is discussed in more detail in Sec. 2.5.1. Furthermore, this thesis aims to measure the CP observables in $B^0 \rightarrow D^{*+}D^{*-}$ decays. The general discussion of CP violation is given in Sec. 2.4. Sec. 2.5.2 combines this information and summarises the resulting properties of applying the theory on the $B^0 \rightarrow D^{*+}D^{*-}$ decays. The $B^0 \rightarrow D^{*+}D^{*-}$ decays introduce a further complication in the measurement of CP violation, as these are $B \rightarrow VV$ processes. This dilutes the measurement of the CP parameter $S_{D^{*+}D^{*-}}$ and is discussed in Sec. 2.5.3.

2.5.1 Ratio of branching fractions

The branching fraction \mathcal{B} indicates the probability that an initial state decays into the final state. It is defined as the ratio of the corresponding decay width relative to the total width. The time-integrated branching fraction of the decays $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ decays are defined as

$$\begin{aligned} \mathcal{B}(B^0 \rightarrow D^{*+}D^{*-}) &= \frac{\mathcal{N}_{B^0 \rightarrow D^{*+}D^{*-}}}{2 \cdot f_s \cdot \mathcal{L}_{\text{int}} \cdot \sigma_{b\bar{b}} \cdot \varepsilon_{B^0 \rightarrow D^{*+}D^{*-}}}, \\ \mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-}) &= \frac{\mathcal{N}_{B_s^0 \rightarrow D^{*+}D^{*-}}}{2 \cdot f_d \cdot \mathcal{L}_{\text{int}} \cdot \sigma_{b\bar{b}} \cdot \varepsilon_{B_s^0 \rightarrow D^{*+}D^{*-}}}, \end{aligned} \quad (2.45)$$

where f_i are the hadronisation probabilities, $\sigma_{b\bar{b}}$ is the $b\bar{b}$ production cross section, \mathcal{L}_{int} is the integrated luminosity, \mathcal{N}_i are the number of signal candidates and ε_i are the selection efficiencies of the respective decays. The factor 2 enters the definition because charge-conjugated decays are also possible. By measuring the $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction relative to the $B^0 \rightarrow D^{*+}D^{*-}$ fraction systematic uncertainties of the integrated luminosities \mathcal{L}_{int} and the cross sections $\sigma_{b\bar{b}}$ cancel. Additionally, possible systematic uncertainties induced from the $D^{*\pm}$ mesons or their decay products cancel as both decays are reconstructed using the same final states. Calculating the ratio of branching fractions with individual branching fractions defined in Eq. (2.45) results in

$$\frac{\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-})}{\mathcal{B}(B^0 \rightarrow D^{*+}D^{*-})} = \frac{\mathcal{N}_{B_s^0 \rightarrow D^{*+}D^{*-}}}{\mathcal{N}_{B^0 \rightarrow D^{*+}D^{*-}}} \cdot \frac{f_d}{f_s} \cdot \frac{\varepsilon_{B^0 \rightarrow D^{*+}D^{*-}}}{\varepsilon_{B_s^0 \rightarrow D^{*+}D^{*-}}}. \quad (2.46)$$

The dominant contribution to the ratio arises from the ratio of the candidates and the ratio of the hadronisation probabilities since the selection efficiencies of both decays are expected to be approximately equal.

2.5.2 CP violation in $B^0 \rightarrow D^{*+}D^{*-}$ decays

Several processes contribute to the decay amplitude of the $B^0 \rightarrow D^{*+}D^{*-}$ decay. The dominant Feynman diagrams are shown in Fig. 2.4. Besides the main contribution from the tree-level diagram with a $b \rightarrow \bar{c}\bar{d}$ transition, the penguin, W -exchange and penguin annihilation diagrams have to be taken into account when calculating the decay amplitude. The amplitude is the same as for the $B^0 \rightarrow D^+D^-$ decay [32, 35, 39] since the $D^{*\pm}$ and D^\pm mesons share the same quark content and, therefore, the transition processes are the same. It is given by

$$A(B^0 \rightarrow D^{*+}D^{*-}) = V_{cd}V_{cb}^* (T + P_c + E + PA_c) + V_{ud}V_{ub}^* (P_u + PA_u) + V_{td}V_{tb}^* (P_t + PA_t), \quad (2.47)$$

where the index of the P and PA diagrams indicate the quark that is exchanged in the loop. Rearranging the unitarity equation given in Eq. (2.12) and introducing the observable \mathcal{A} defined as

$$\mathcal{A} := T + E + P_c + PA_c - P_t - PA_c \quad (2.48)$$

allows to simplify the decay amplitude to

$$A(B^0 \rightarrow D^{*+}D^{*-}) = V_{cd}V_{cb}^* \mathcal{A} \left[1 + \frac{V_{ud}V_{ub}^* (P_u + PA_u - P_t - PA_t)}{V_{cd}V_{cb}^* \mathcal{A}} \right]. \quad (2.49)$$

The Wolfenstein parametrisation of the CKM matrix given in Eq. (2.9) is exploited to simplify the decay amplitude further. Using the definitions of R_b and the CKM angle γ from Eq. (2.13) and Eq. (2.14), respectively, the ratio of the matrix elements in Eq. (2.49) is calculated to be

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} = -R_b e^{i\gamma}. \quad (2.50)$$

By defining

$$ae^{i\theta} := R_b \left(\frac{P_u + PA_u - P_t - PA_c}{\mathcal{A}} \right) \quad (2.51)$$

with the hadronic CP conserving parameters \mathcal{A} , a , and θ , the decay amplitude is simplified to

$$A(B^0 \rightarrow D^{*+}D^{*-}) = V_{cd}V_{cb}^* \mathcal{A} (1 - ae^{i\theta} e^{i\gamma}). \quad (2.52)$$

In contrast to the parameters a and θ , the CKM angle γ is a CP -violating weak phase. Taking into account the characteristics of the individual parameters under CP transformation, the amplitude of the \bar{B}^0 decay is determined to be

$$A(\bar{B}^0 \rightarrow D^{*+}D^{*-}) = V_{cd}^* V_{cb} \mathcal{A} (1 - ae^{i\theta} e^{-i\gamma}) . \quad (2.53)$$

Using the definition in Eq. (2.39), the parameter $\lambda_{D^{*+}D^{*-}}$ is calculated to be

$$\lambda_{D^{*+}D^{*-}} = \eta_{CP} e^{-i2\beta} \frac{1 - ae^{i\theta} e^{i\gamma}}{1 - ae^{i\theta} e^{-i\gamma}} . \quad (2.54)$$

The CP parameters are subsequently determined with the definitions given in Eq. (2.31). They are calculated to be

$$\begin{aligned} S_{D^{*+}D^{*-}} &= \frac{2\Im(\lambda)}{1 + |\lambda|^2} = -\eta_{CP} \frac{\sin(2\beta) - 2a \cos(\theta) \sin(\phi_d + \gamma) + a^2 \sin(\phi_d + 2\gamma)}{1 + a^2 - 2a \cos(\theta) \cos(\gamma)} , \\ C_{D^{*+}D^{*-}} &= \frac{1 - |\lambda|^2}{1 + |\lambda|^2} = \frac{2a \sin(\theta) \sin(\gamma)}{1 + a^2 - 2a \cos(\theta) \cos(\gamma)} . \end{aligned} \quad (2.55)$$

Only two of the three CP observables are calculated as the third observable $A_{D^{*+}D^{*-}}^{\Delta\Gamma}$ does not enter the decay rates in the B^0 system due to the assumption of $\Delta\Gamma_d = 0$. The observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ still include the hadronic parameters. The calculation of these hadronic parameters is not trivial due to QCD effects [39]. Accordingly, they need to be determined experimentally. In the absence of the hadronic contributions, the parameter $S_{D^{*+}D^{*-}}$ is equal to $-\sin 2\beta$. For the $B^0 \rightarrow D^{*+}D^{*-}$ decay, the SM predicts higher-order contributions from penguin diagrams in the order of a few percent [30, 31]. Consequently, direct CP violation, *i.e.* $C_{D^{*+}D^{*-}} \neq 0$, must be considered and only an effective phase $2\beta^{\text{eff}}$ defined as

$$2\beta^{\text{eff}} = \phi_d^{\text{eff}} = \phi_d + \Delta\phi_d \quad (2.56)$$

can be measured, where the phase shift $\Delta\phi_d$ is calculated to be

$$\tan(\Delta\phi_d) = \frac{a^2 \sin(2\gamma) - 2a \cos(\theta) \sin(\gamma)}{1 + a^2 \cos(2\gamma) - 2a \cos(\theta) \cos(\gamma)} . \quad (2.57)$$

The relation of this effective phase and the CP parameters $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ is given by

$$\sin(\phi_d^{\text{eff}}) = -\frac{S_{D^{*+}D^{*-}}}{\sqrt{1 - C_{D^{*+}D^{*-}}^2}} . \quad (2.58)$$

The current world averages and the individual measurements contributing to the averages of the $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ observables are shown in Fig. 2.5.

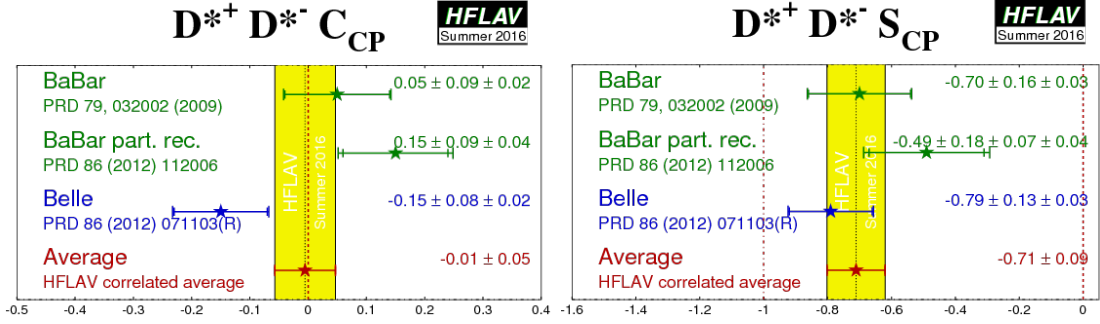


Figure 2.5 – Averages of the CP observables (left) $C_{D^{*+}D^{*-}}$ and (right) $S_{D^{*+}D^{*-}}$ in $B^0 \rightarrow D^{*+}D^{*-}$ decays taken from Ref. [93].

The CP observable $S_{D^{*+}D^{*-}}$ defined in Eq. (2.55) still depends on the CP eigenvalue η_{CP} of the final state. The $B^0 \rightarrow D^{*+}D^{*-}$ decays are $B \rightarrow VV$ processes, where the final state is a mixture of CP -odd and CP -even eigenstates. As a consequence, the CP eigenvalue of the final state has to be determined when measuring the weak mixing angle β^{eff} in $B^0 \rightarrow D^{*+}D^{*-}$ decays. This can be achieved by performing an angular analysis, which is discussed in more detail in Sec. 2.5.3.

2.5.3 Angular formalism in $B^0 \rightarrow D^{*+}D^{*-}$ decays

A non- CP eigenstate as a final state has to be considered when measuring CP violation. The B^0 mesons are pseudoscalar particles, *i.e.* they have a spin of $J^P = 0$. In contrast, $D^{*\pm}$ mesons are vector particles with a spin of $J^P = 1^-$. Consequently, three different relative orbital momenta between the D^{*+} and D^{*-} mesons are possible, *i.e.* $L = 0$, $L = 1$ and $L = 2$. The CP eigenvalues of each of the three states are given by

$$CP|D^*D^{*-}\rangle = (-1)^L|D^*D^{*-}\rangle. \quad (2.59)$$

The resulting possible CP eigenvalues of the final states are $\eta_{CP} = \pm 1$. They are called CP odd for $L = 1$ with $\eta_{CP} = -1$ and CP even for $L = 0$ and $L = 2$ with $\eta_{CP} = +1$. The final state thus consists of a mixture of CP -odd and CP -even eigenstates. Due to the different signs of the CP eigenvalues, the decay amplitudes of the CP -odd and CP -even states are different, and the overall CP asymmetry is diluted. With the amplitudes A_+ and A_- of the CP -even and CP -odd states, respectively, the dilution factor D is calculated to be

$$D = \frac{A_+^2 - A_-^2}{A_+^2 + A_-^2} = 1 - \frac{2A_-^2}{A_+^2 + A_-^2} = 1 - 2R_{\perp}, \quad (2.60)$$

where R_{\perp} is the fraction of the CP -odd component in the final state. As derived in Ref. [94], the CP parameter $S_{D^{*+}D^{*-}}$ is diluted by this factor D . In contrast, the parameter $C_{D^{*+}D^{*-}}$ as defined in Eq. (2.55) is independent of η_{CP} . By considering these dependencies, assuming $\Delta\Gamma_d = 0$ and neglecting contributions from indirect CP violation, *i.e.* $|q/p|^2 = 1$,

the time-dependent decay rates as given in Eq. (2.30) are determined to be

$$\begin{aligned}\Gamma_{B^0(t) \rightarrow D^{*+}D^{*-}} &\propto e^{-\Gamma t} (1 + A_{D^{*+}D^{*-}}) [1 - DS_{D^{*+}D^{*-}} \sin(\Delta m_d t) + C_{D^{*+}D^{*-}} \cos(\Delta m_d t)] \\ \Gamma_{\bar{B}^0(t) \rightarrow D^{*+}D^{*-}} &\propto e^{-\Gamma t} (1 + A_{D^{*+}D^{*-}}) [1 + DS_{D^{*+}D^{*-}} \sin(\Delta m_d t) - C_{D^{*+}D^{*-}} \cos(\Delta m_d t)]\end{aligned}\quad (2.61)$$

Both assumptions, *i.e.* $\Delta\Gamma_d = 0$ and $|q/p|^2 = 1$, are in agreement with the current experimental precision [47]. Since the dilution factor directly enters the decay rates, the CP eigenvalue of the final state has to be determined when measuring CP violation. This can be done by disentangling the CP -odd and CP -even components via an angular analysis, which allows to measure the CP asymmetry without losing statistical power [95]. The general approach in angular analyses is to study angular distributions. These are sensitive to spin correlations and reveal the properties of particles and their interactions. However, in the case of a spin-zero particle, such as the B^0 meson, the decay chain has to be analysed since there are no spin correlations in the production mechanism. A common angular basis for such an analysis is the helicity basis [47, 96]. The helicity of a particle is defined as the projection of the spin along the momentum [97]. For a B^0 meson decaying into the $D^{*+}D^{*-}$ final state, the spin projection of the final state on the decay axis in the B^0 rest frame has to be zero. Therefore, the $D^{*\pm}$ mesons are required to have the same helicity. Since the $D^{*\pm}$ meson has spin $J^P = 1^-$, the helicity can take three values, *i.e.* $h = -1$, $h = 0$ and $h = +1$. A complex amplitude H_h is associated with each helicity eigenstate, *i.e.* the longitudinal amplitude H_0 and the transverse amplitudes H_+ and H_- . Since this basis allows to determine the longitudinal rate directly, it is very popular for theoretical calculations. In contrast, the transversity basis [47, 95] is convenient for determining the CP -odd fraction. In the transversity basis, the amplitudes are defined as spin projections for one vector particle parallel and perpendicular to the plane of the decay of the other. In comparison with the helicity basis, the longitudinal amplitude A_0 remains unchanged. The other two transversity amplitudes are defined as the following linear combinations of the helicity amplitudes

$$A_0 = H_0, \quad A_\perp = \frac{1}{\sqrt{2}}(H_+ - H_-), \quad A_\parallel = \frac{1}{\sqrt{2}}(H_+ + H_-). \quad (2.62)$$

The decay products of the $D^{*\pm}$ mesons, *i.e.* the $D_{D^{*\pm}}^0$ mesons and the respective slow pions $\pi_{D^{*\pm}}^\pm$, are analysed to determine the angular momenta. The particles are used to define the angles in the transversity basis with which the decay rate of $B^0 \rightarrow D^{*+}D^{*-}$ decays can be described. The angles are shown in Fig. 2.6 and defined as follows. First, the polar angle $\theta_{D^{*-}}$ of the $\pi_{D^{*-}}^-$ in the D^{*-} rest frame, *i.e.* the angle between the momentum of the $\pi_{D^{*-}}^-$ and the momentum of the D^{*+} meson in the rest frame of the D^{*-} meson. Second, the polar angle θ_{tr} between the normal of the D^{*-} decay plane and the $\pi_{D^{*+}}^+$ line of flight. Third, the relative azimuthal angle ϕ_{tr} , *i.e.* the angle between the momentum of the D^{*+} meson in the B^0 rest frame and the momentum vector of the $\pi_{D^{*+}}^+$ projected onto the D^{*-} decay-plane in the D^{*+} rest frame.

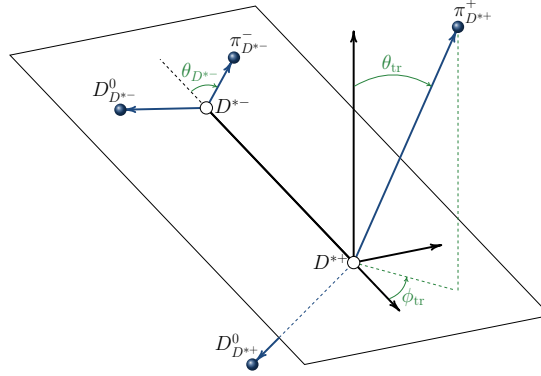


Figure 2.6 – Illustration of the definition of the angles in the transversity basis. The angles are defined in more detail in the text.

The $\theta_{D^{*-}}$ and θ_{tr} distributions of the CP -odd and CP -even contributions allow to separate the two components. By integrating three-dimensional angular-dependent decay rate [95] over the angle ϕ_{tr} , the two-dimensional decay rate is calculated to be

$$\frac{d^2\Gamma(R_0, R_\perp)}{d\cos(\theta_{D^{*-}})d\cos(\theta_{tr})} = \frac{9}{16} \sum_{i=0, \perp, \parallel} R_i H_i(\cos(\theta_{tr}), \cos(\theta_{D^{*-}})), \quad (2.63)$$

where

$$\begin{aligned} H_0 &= 2 \cos^2(\theta_{D^{*-}}) \sin^2(\theta_{tr}), \\ H_\perp &= 2 \sin^2(\theta_{tr}) \cos^2(\theta_{tr}), \\ H_\parallel &= \sin^2(\theta_{tr}) \sin^2(\theta_{tr}). \end{aligned} \quad (2.64)$$

The relation between the amplitudes A_i and the R_i is given by $\sqrt{R_i} = A_i$, and the real-valued parameters R_i are defined to have the sum of one. The observable R_\perp is the CP -odd fraction that directly enters the decay rates given in Eq. (2.61) via Eq. (2.60). The current world average of the CP -odd fraction is given by

$$R_\perp = 0.15 \pm 0.02 \quad (2.65)$$

as given in Ref. [93]. It is illustrated in Fig. 2.7, where also the individual measurements contributing to the average are shown. The theoretical prediction of the CP -odd fraction is $R_\perp^{\text{theor}} \approx 11\%$ [30].

2.6 Limitations of the Standard Model

The SM is in good agreement with experimental observations, and all theoretical predictions have been confirmed experimentally. Although the model describes the processes very accurately on a microscopic level, it fails to completely explain nature on a macro-

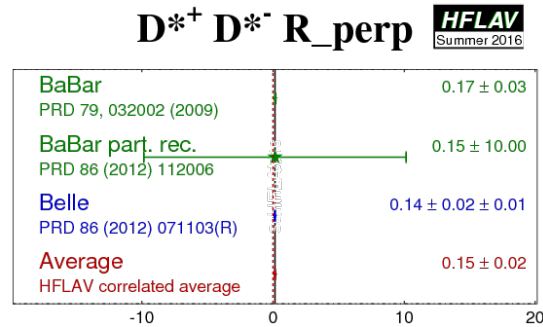


Figure 2.7 – Average of the CP -odd fraction in $B^0 \rightarrow D^{*+} D^{*-}$ decays taken from Ref. [93].

scopic level. The perfect example is that it does not describe the gravitational force, the dominant interaction in the macroscopic world. An extension of the current SM would include the gravitational force and describe it in a quantum-mechanical framework instead of the still-used basis of Einstein's general theory of relativity.

The SM also does not explain the existence of dark matter and energy. However, these two components of the universe comprise 95% of its energy density [11]. Accordingly, the model describes only 5% of the known matter. Furthermore, the known matter also leads to the question of why similarly large clusters of antimatter do not exist. This phenomenon is also referred to as the matter-antimatter asymmetry. In 1967, Sakharov established three conditions for this asymmetry to occur [12]. These are the violation of the baryon number, the deviation from thermal equilibrium in the early inflation of the universe and the violation of the CP symmetry. Although the latter condition has been observed experimentally, the amount of the measured CP violation is not sufficient to describe the matter-antimatter asymmetry [13, 14].

Another phenomenon that is not included in the SM are neutrino oscillations [9, 10]. These oscillations show that the neutrinos are not massless, as the SM implies. One solution to this problem would be to describe the neutrinos as Majorana particles, *i.e.* they would be their own antiparticles [98].

3 The LHCb experiment at the LHC

One of the most remarkable particle colliders that currently exists is the LHC. The LHC is operated by the CERN and is located near Geneva, Switzerland. A unique detector installed at the LHC is the LHCb experiment. The main focus of the LHCb experiment lies in testing the SM with precision measurements or searching for NP by studying hadrons containing b or c quarks.

In this chapter, the CERN facility with its accelerators is described in Sec. 3.1. This thesis uses data collected with the LHCb experiment. Therefore, a detailed overview of the LHCb detector is given in Sec. 3.2. An outline of the software used at the experiment is discussed in Sec. 3.3. Finally, Flavour Tagging (FT) at the LHCb experiment is explained in Sec. 3.4, which is used to tag the initial B meson flavour. If not specified in more detail, the information presented in this chapter is mainly based on Refs. [99–101].

3.1 The Large Hadron Collider

The LHC is a hadron accelerator and collider which has been installed in the old tunnel of the Large Electron-Positron Collider (LEP)[102]. The tunnel is located 45 m to 170 m beneath the surface and has a length of 26.7 km. Accordingly, it is currently the largest particle accelerator in the world. Various hadrons can be accelerated in the LHC, including protons or lead ions [103]. For this thesis, only the physics programme with protons is of interest. Therefore, the following description of the CERN facility will focus on the relevant information for this programme.

In the LHC ring, the protons are accelerated in opposite directions and are brought to collision at four interaction points. They are organised in bunches, which circulate in intervals of 25 ns, resulting in a bunch crossing frequency of 40 MHz. This leads to a total number of up to 2808 bunches for each beam, where one bunch contains approximately 1.1×10^{11} protons. A total of 1232 dipole magnets with a maximum field strength of 8.33 T keep the proton bunches in their orbit. These magnets are superconducting and cooled with liquid helium to reach a temperature below 2 K. The maximum field strength limits the maximum beam energy of the LHC to 7 TeV since the protons can not be kept on their trajectory for higher beam energies. To focus the beams, the LHC has about 450 quadrupoles installed. Both beam pipes are evacuated to an almost complete vacuum of 10^{-13} bar to 10^{-14} bar to avoid collisions with air molecules.

The LHC uses a series of pre-accelerators which have been built before the LHC was designed. All accelerators, as well as the large experiments, are shown in the schematic

layout of the whole CERN facility in ???. The first pre-accelerator is a linear accelerator called LINAC2, which is used to accelerate protons to an energy of 50 MeV. In the next step, the protons are injected into the Proton Synchrotron Booster (BOOSTER). After reaching an energy of 1.4 GeV in the BOOSTER, they are transferred into the Proton Synchrotron (PS) where they are accelerated up to an energy of 25 GeV. Finally, they are passed to the Super Proton Synchrotron (SPS), accelerated to an energy of 450 GeV and injected into the LHC storage ring. In the LHC, the protons are accelerated to their final energy of 7 TeV.

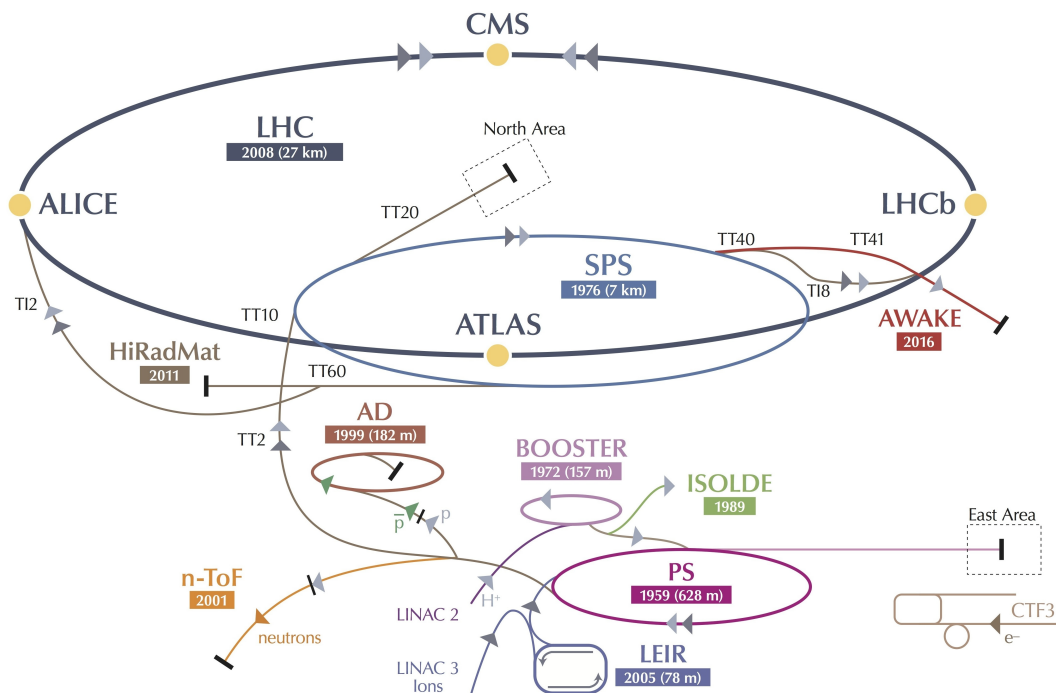


Figure 3.1 – CERN accelerator facility with all pre-accelerators and the LHC. The four large experiments ATLAS, CMS, ALICE and LHCb, as well as smaller experiments, are shown. All pre-accelerators and also characteristics of the LHC are described in the text. The scheme is taken from Ref. [104].

The design of the LHC allows a maximum centre-of-mass energy of up to 14 TeV and a peak luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ to be achieved at the intersection points of the LHC. However, during the first two data-taking periods of the LHC, the maximum centre-of-mass energy was not reached yet. The first period is defined as the years from 2011 to 2012. Due to safety reasons after a major fault in 2008, the LHC was operated with a reduced centre-of-mass energy of 7 TeV and 8 TeV in 2011 and 2012, respectively. In the second period, during the years from 2015 to 2018, the LHC was operated with a

centre-of-mass energy of 13 TeV. The data-taking periods can also be called a Run. In what follows, they are therefore further referenced as Run 1 and Run 2.

Four large experiments at CERN are installed at the interaction points of the LHC. These are the ATLAS [105], CMS [106], ALICE [107] and LHCb experiments. ATLAS and CMS are general-purpose experiments which study a broad particle physics programme. They are hermetic detectors covering the full angular range. Despite being also a hermetic detector, the ALICE experiment is focussed on heavy ion physics in lead collisions and on studies of quark-gluon plasma. The data analysed in this thesis was recorded with the LHCb experiment. This experiment has a unique geometry and is described in more detail in the following section.

3.2 The LHCb detector

The LHCb experiment is designed for precision measurements in the c and b quark sectors. It is also capable of performing fixed-target measurements by injecting different gases into the interaction point [108]. In contrast to the other large experiments at CERN, it is not centred around the collision point and does not cover the entire angular range but is instead built in one direction from the collision point. The unique angular range covered by the detector extends from ± 10 mrad to ± 250 mrad in the vertical and from about ± 10 mrad to ± 300 mrad in the horizontal plane which corresponds to a pseudorapidity range of $2 < \eta < 5$. Due to this special detector geometry, it is also called a single-arm forward spectrometer. This setup is particularly beneficial for the physics programme since b and c quarks are mainly produced in the forward direction. The reason is that the main production process of the quark pairs is gluon-gluon fusion. Since the energy threshold for b or c pairs is relatively small, there is a high probability that one of the gluons will carry a larger momentum, resulting in a boost of the produced quark pair. The production angle and the pseudorapidity of $b\bar{b}$ pairs at a centre-of-mass energy of 14 TeV are shown in Fig. 3.2. Overall, around 25% of the quark pairs are produced in a direction that is covered by the LHCb detector.

A scheme of the detector design is presented in Fig. 3.3, where all major subdetectors that are essential for taking quality data are shown. One other detector, which is not illustrated in Fig. 3.3, is the Beam Condition Monitor system (BCM) [110]. It is installed to monitor the stability of the beam around the interaction point at LHCb. Particularly during the injection of protons into the LHC, but also in general, it can happen that the particle beams are misaligned. Misaligned beams result in particle showers hitting sensitive material, which can cause crucial damage to the LHCb detector. To prevent such incidents, the BCM system measures the flux of charged particles around the beam pipe. If the beams are misaligned, the system can initiate a beam dump to protect the detector against radiation damage. Therefore, it is essential for the safety of the LHCb detector [111]. Two subdetectors are installed, both equipped with eight diamond sensors. One is placed downstream, and one is upstream of the interaction point.

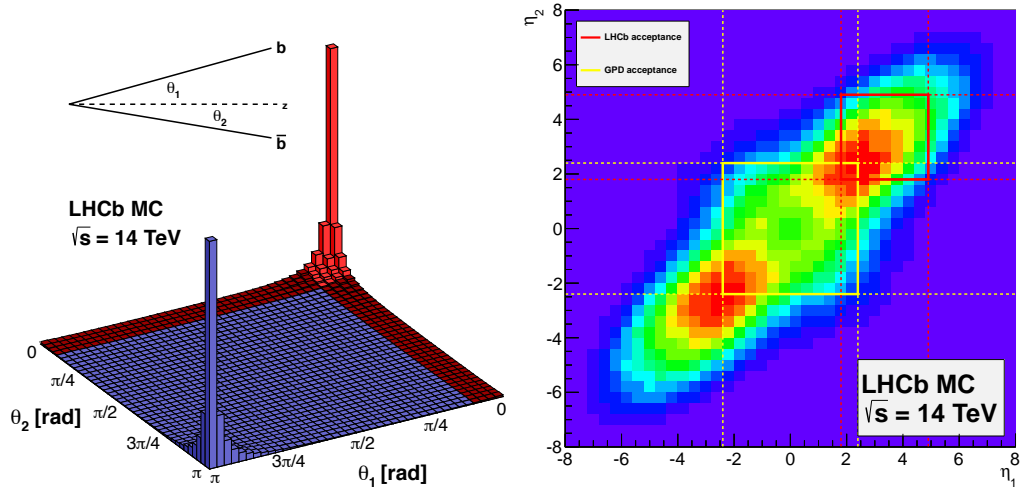


Figure 3.2 – The (left) production angle and (right) pseudorapidities of simulated $b\bar{b}$ pairs at a centre-of-mass energy of 14 TeV both taken from Ref. [109]. The LHCb acceptance region is coloured in red in the left plot and indicated by a red box in the right plot.

The LHCb experiment does not operate at the maximum instantaneous luminosity delivered by the LHC. Using a reduced luminosity keeps the amount of inelastic proton-proton scattering in the detector at a constant level. In detail, it allows to operate in a cleaner environment ensuring good data and reconstruction quality since otherwise, the track multiplicity in the detector would be too high to perform precision measurements. To achieve this, the two proton beams are repositioned so that they do not overlap completely. The procedure is called luminosity levelling and allows the LHCb experiment to operate at an instantaneous luminosity of $\mathcal{L} = 4 \cdot 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ [112]. Operating at this luminosity, the experiment recorded data sets corresponding to integrated luminosities of $\mathcal{L}_{\text{int}, 2011} = 1.11 \text{ fb}^{-1}$ and $\mathcal{L}_{\text{int}, 2012} = 2.08 \text{ fb}^{-1}$ during Run 1. The individual samples were recorded at centre-of-mass energies of $\sqrt{s}_{2011} = 7 \text{ TeV}$ and $\sqrt{s}_{2012} = 8 \text{ TeV}$, respectively. For Run 2 the centre-of-mass energy was increased to $\sqrt{s}_{\text{Run 2}} = 13 \text{ TeV}$. A data sample corresponding to an integrated luminosity of $\mathcal{L}_{\text{int}, \text{Run 2}} = 5.99 \text{ fb}^{-1}$ has been recorded during this period. Overall, an integrated luminosity of $\mathcal{L}_{\text{int}} = 9.18 \text{ fb}^{-1}$ has been recorded during Run 1 and Run 2.

In general, an accurate tracking and a reliable particle identification have to be ensured by the detector to record data of high quality. The individual subdetectors shown in Fig. 3.3 are discussed in more detail in the following.

3.2.1 Tracking system

The LHCb tracking system consists of several subdetectors to ensure accurate tracking of the particles travelling through the detector. It comprises a high-resolution vertex

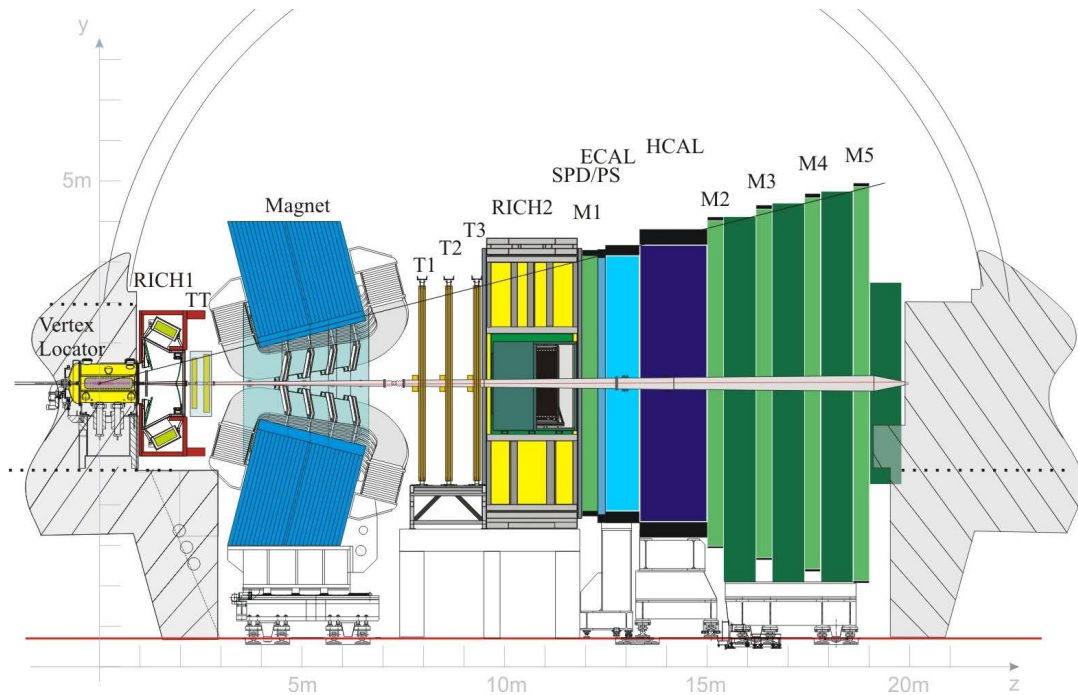


Figure 3.3 – Schematic overview of the LHCb detector [100]. All essential subdetectors are shown, *i.e.* the Vertex Locator (VELO), the two Ring Image Cherenkov detectors (RICH1, RICH2), the tracking stations (TT, T1 - T3), the Preshower Detector (PS), the Scintillating Pad Detectors (SPD), the electromagnetic calorimeter (ECAL), the hadronic calorimeter (HCAL), the muon chambers (M1 - M5) and the dipole magnet. They are explained in more detail in Sec. 3.2.1 and Sec. 3.2.2.

detector and several tracking stations. In addition, a dipole magnet is installed to bend the tracks of charged particles. All components of the tracking system are discussed in more detail below.

Vertex locator One of the most important subdetectors is the Silicon Vertex Locator (VELO). It is directly placed around the interaction point and allows to precisely measure the production and decay vertices of the particles, *e.g.* of produced B mesons. These vertices are further referenced as the primary vertex (PV) and the secondary vertex (SV). A high resolution on the measurement of the production and decay vertices is crucial for decay-time dependent studies as performed in this thesis. This is due to the short lifetime or rather short flight distances of the B mesons. The VELO consists of 21 semi-circular silicon modules that are installed in the beam pipe. The modules allow to measure the tracks of particles in cylindrical coordinates, *i.e.* the radial distance R and the azimuthal angle ϕ . The longitudinal component is measured using information about the position of the modules. For vertices with 25 tracks, a resolution of around $13\ \mu\text{m}$ and $71\ \mu\text{m}$

can be reached transversal and along to the beam axis [113], respectively. A decay-time resolution of approximately 50 fs is obtained [113]. This can be achieved by moving the VELO up to 5 mm close to the beam pipe when the beams are stable after the injection into the LHC. Since the VELO is a key subdetector for many analyses, it also has to be protected when there are no stable beams. For this reason, the VELO can be moved up to 3.5 cm away from the beam, preventing the system from radiation damage.

Tracking stations The tracking stations at the LHCb experiment gain information about charged particles. By combining the information obtained from the VELO and the tracking system, most of the tracks can be reconstructed. The tracking system itself can be divided into two subsystems. The first subsystem is the Tracker Turicensis (TT), which is directly located before the magnet. It covers the full detector acceptance and consists of four layers of silicon micro-strip sensors with vertically aligned strips. The inner two layers are rotated by an angle of $\pm 5^\circ$ to achieve a spatial resolution in the y direction. Accordingly, the layers are arranged in the $x-u-v-x$ scheme. The overall task of the TT is to gather information on long-lived neutral particles decaying outside of the VELO by measuring the tracks of the charged decay products.

The second subsystem is located directly behind the magnet and comprises the tracking stations T1 to T3. Each of these stations consists of an Inner Tracker (IT) and an Outer Tracker (OT). The IT is a high-resolution silicon-strip detector centred around the beam pipe. In contrast, the OT is a drift tube detector that covers an angular range further away from the beam pipe. The OT is filled with Ar/CO₂/O₂ (70:28.5:1.5% [114]) which is ionised by charged particles. By measuring the time of flight of the electron to the anode, the position of the particles can be determined. The reason for the different designs of the IT and OT is that the area closer to the beam pipe is significantly higher occupied with particles than the outer detector region. Consequently, a higher resolution is needed in the central region of the detector. In contrast, drift tube detectors usually have a lower resolution due to their higher delay time. However, this resolution is sufficient for the outer region of the tracking stations and therefore it is a cost-effective solution. The IT of each station consists of four layers arranged in the same scheme used for the TT, *i.e.* $x-u-v-x$. A single-hit resolution of about 50 μm [115] and 200 μm [114] is achieved for the IT and OT, respectively. The resolution for the TT and IT is the same.

Magnet A dipole magnet is installed at the LHCb experiment. It is placed between the TT and the remaining tracking stations T1 to T3. The tracking stations are exposed to the maximum field strength of the magnet, while the sub-detectors in front or behind them are only affected by a low field strength. This is necessary to keep the influence of the field on the subdetectors as low as possible. The field points in the vertical direction but is periodically inverted to study the charge-dependent asymmetries of the detector. In general, the main task of the magnet is to bend the tracks of charged particles. Information about the curvature from the other tracking subsystems allows

the determination of the momenta of these particles since the strength of the bending and the momentum are correlated.

Overall, a track finding efficiency of 96% for the reconstruction of charged particles is reached considering information from all tracking subsystems. Depending on the particle kinematics, a resolution of around 0.5% to 0.8% is achieved on the measurement of the momentum. The mass resolution is determined to be around 0.5%.

3.2.2 Particle identification system

The particle identification system at the LHCb experiment comprises several subdetectors such as the Ring Imaging Cherenkov Detector (RICH), a calorimeter system and a muon identification system. These subsystems are discussed in more detail in the following.

Ring imaging Cherenkov detectors The RICH [116, 117] detectors identify charged particles and are decisive in distinguishing between hadrons such as pions, kaons and protons. A strong differentiation and a small misidentification (misID) of these particles is also essential for this thesis since all final-state particles of the reconstructed decays are kaons and pions. Two RICH detectors are installed at LHCb, further referenced as RICH1 and RICH2. RICH1 is located directly behind the VELO and covers a momentum range from 2 GeV to 40 GeV. In comparison, RICH2 covers a range from 15 GeV to 100 GeV and is located directly behind the tracking stations T1 to T3.

Both subdetectors utilise the Cherenkov effect to identify particles. The effect occurs when charged particles traverse a medium with a velocity v that is faster than the speed of light in this medium c_M . In this case, the particles emit light in the form of a cone along their trajectory. The cosine of the opening angle of the cone θ_C is then given by $\cos(\theta_C) = 1/(n\beta)$, where n is the refractive index of the medium and β is defined as $\beta = v/c_M$. In the RICH detectors, the emitted light is projected over spherical and planar mirrors onto photodetectors. These photodetectors are placed outside of the detector acceptance to prevent them from radiation damage and to avoid possible influences on the particle tracks. By measuring the radius of the cone of the projected light, the opening angle θ_C can be determined, as these two quantities are proportional to each other. Since the refraction index of the gas inside of the RICH detectors is known, the velocity can be determined. By combining the velocity with information about the momentum from the tracking system discussed in Sec. 3.2.1, the particle mass can be calculated, and a particle hypothesis can be assigned. In Fig. 3.4, the angular distribution is plotted against the momentum for isolated events. It shows a good separation between the hadrons, such as pions, kaons and protons, as well as the muon.

The schematic structure of the RICH1 detector is also shown in Fig. 3.4. The RICH2 detector is not presented, but it uses a similar structure [118]. Behind the magnet, the momentum spectrum is shifted to higher values. Due to the dependence of the refraction

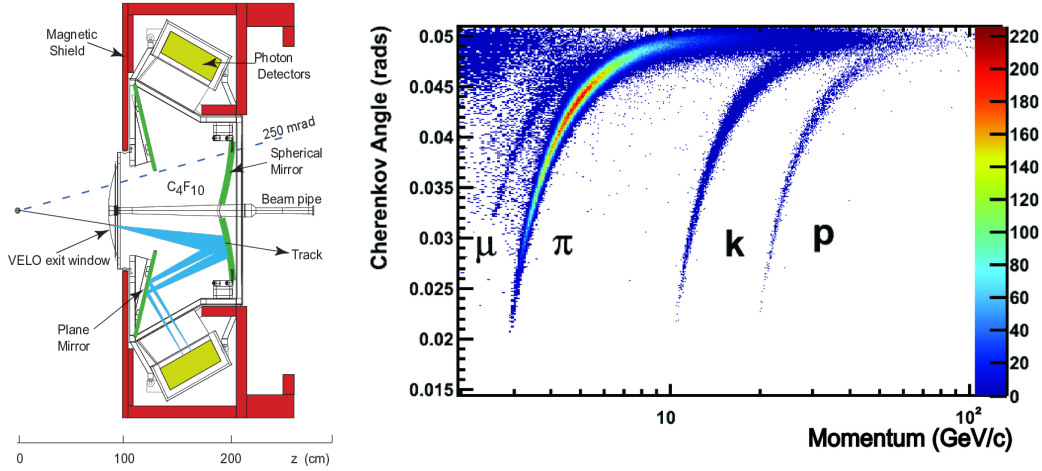


Figure 3.4 – The (left) schematic overview of the RICH1 detector [118] and the (right) reconstructed Cherenkov angle as a function of track momentum [117] in the C_4F_{10} radiator used in RICH1. The angular distribution is shown for isolated events. A good separation between the hadrons, such as pions, kaons and protons, as well as the muon, is achieved. The RICH2 detector uses the same principle and a similar structure, but a different radiator gas is used, *i.e.* CF_4 .

index, the momentum and the opening angle, different gases are used in RICH1 and RICH2 to ensure an excellent pion and kaon separation over the full momentum range. In detail, RICH1 is filled with C_4F_{10} while RICH2 is filled with CF_4 . Overall, a correct kaon identification of 95% and a pion to kaon misID of 5% is reached.

Calorimeter system The calorimeter system includes several subdetectors to measure the energy and position of particles. It also allows to identify hadrons, electrons and photons. The schematic structure with all subcomponents is shown in Fig. 3.5.

The first part of the calorimeter system is a double detector system consisting of the Scintillating Pad Detector (SPD) and the Preshower Detector (PS) with a 2.5 radiation length (15 mm) lead layer placed between both components. The radiation length is a characteristic quantity of a material that indicates the energy loss of high-energy particles interacting electromagnetically with the material. The SPD detects charged particles which deposit energy in the scintillator material. Radiation energy originating from electrons and photons interacting with the lead absorber is measured with the PS. Consequently, the SPD and PS are used to separate photons and electrons. The system also helps with the detection of neutral and charged pions. The charged pions also cause a signal in the SPD while the neutral pions do not. Electrons and charged pions can then be separated because the electrons shower already in the PS. Neutral pions with high transversal energy can be separated because they do not cause a signal in the SPD or PS.

Directly behind the double detector system, the electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL) are installed. The ECAL focuses on the energy meas-

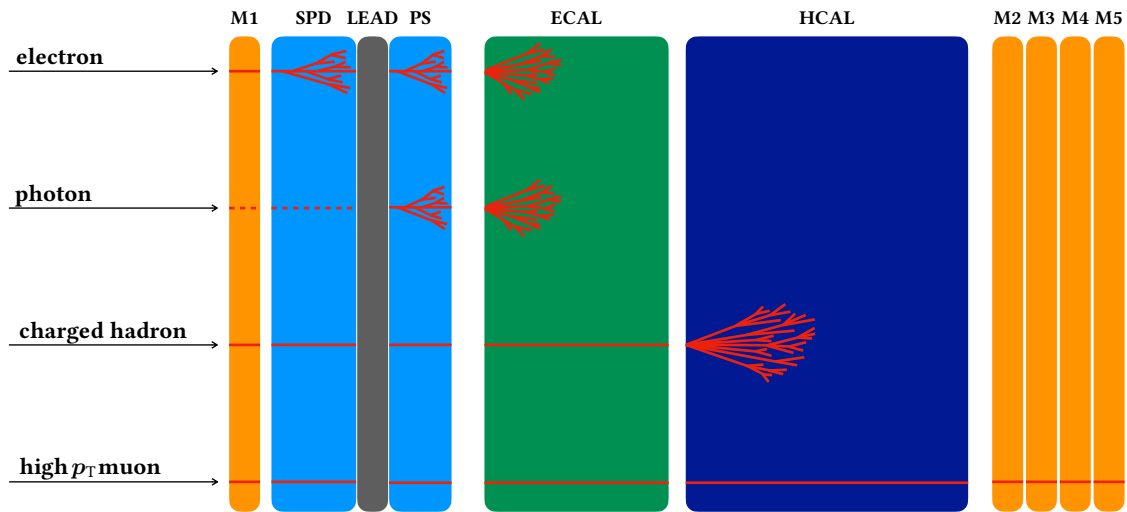


Figure 3.5 – The schematic overview of the calorimeter system installed at the LHCb experiment. The signal of electrons, photons, charged hadrons and muons with high transverse momentum in different components of the system is shown. The system consists of the (blue, SPD) Scintillating Pad Detector, (blue, PS) Preshower Detector, (green, ECAL) electromagnetic calorimeter and (dark blue, HCAL) hadronic calorimeter. A (grey) lead layer is installed in between the Scintillating Pad Detector and the Preshower Detector. In addition to the calorimeter system, the (orange, M1–M5) muon chambers are illustrated. The scheme has been adapted from Ref. [119].

urement of electrons and photons, while the HCAL measures the energy of hadrons. Both calorimeters use a similar structure of alternating layers of absorber and scintillating material. Particles penetrating the absorber layers in front of the scintillators cause electromagnetic and hadronic showers. The showers then excite the scintillating material. By radiation transition into the ground state, photons are emitted, which are then measured by photomultipliers. Multiple layers of absorber and scintillating material are used since interacting particles should deposit all their energy. The main difference between both calorimeters is that the ECAL uses lead layers, and the HCAL uses iron layers as absorber material. The reason is that the average radiation length of hadrons in lead is large and no significant signal will be caused in the ECAL by the hadrons. The resolution on the measurement of the energy deposition for both ECAL and HCAL is defined as $\sigma_E/E = a/\sqrt{E} + b$ where a is a stochastic term and b constant term. For the ECAL a and b are given by 8.5 % to 9.5 % and 0.8 %, respectively. The parameters for the HCAL are (69 ± 5) % and (9 ± 2) % for a and b , respectively.

Muon identification system The muon identification system is the last detector subsystem and consists of five individual muon chambers, where one (M1) is placed in front of the calorimeter system, and the remaining four (M2 to M5) are placed behind the system as shown in Fig. 3.5. Between the chambers M2 to M5, 80 cm thick iron absorbers

are installed. The muon system is an essential component of the LHCb experiment since it is elementary for reconstructing decays with muons. Such decays have a clean detector signature, and their reconstruction is of high importance since they play a huge role in the trigger system and in the FT, both explained later in this chapter. A resolution of around 20% on the measurement of the transverse momentum is achieved [120].

Muons are weakly interacting particles that traverse the whole detector without depositing their entire energy. Additionally, most of the other particles are stopped before the muon chambers, and therefore, only muons produce a signal in the chambers. Low energy muons are also stopped before and in the muon system due to the iron absorbers so that only muons with a momentum above 6 GeV/c traverse all five stations. Since the design of the muon chambers only allows the measurement of charged particles, neutrinos cannot influence a signal.

Two tasks can be assigned to the five muon chambers. First, M1 to M3 are responsible for the tracking direction and the measurement of the transverse momentum. Second, due to a smaller spatial resolution the task of M4 and M5 is to identify particles that penetrate the iron layers. All chambers are multi-wire proportional chambers filled with Ar/CO₂/CF₄ (40:55:5) except for the central region of M1. This region is exposed to a much higher particle flux since it is close to the beam pipe and is installed before the calorimeter system. Due to their radiation resistance, gas-electron multipliers are used for this region instead of the multi-wire proportional chambers. A total muon identification efficiency of 95% with a misidentification rate of below 1% is reached [121].

3.2.3 Trigger

The bunch crossing rate at the LHC is 40 MHz. Even with the approach of luminosity levelling used at LHCb, it is not possible to save all available data to disk. Additionally, the *b* and *c* quark production cross sections are relatively small. Therefore, the data has to be filtered, and only events with respect to the physics program are stored. A trigger system is installed to decide if events are interesting. Due to continuous development, the system has slightly changed between Run 1 and Run 2. The schemes of the individual trigger systems used in the respective runs are shown in Fig. 3.6.

In general, the system consists of one hardware and two software stages. The hardware stage is called the L0 trigger and operates directly at the bunch crossing rate. It uses information from the calorimeter and muon systems. Events with hadrons or electrons having high transverse energy deposition in the calorimeter system or one or two muons with a high transverse momentum are selected. To distinguish between events with one or more visible proton-proton interactions, information from the SPD is used to reject events with a high charged track multiplicity, and a pile-up system veto using information from the VELO is implemented [123]. Overall, the L0 trigger has an efficiency of roughly 60% [124–126] and allows to reduce the data rate to 1 MHz. After the hardware stage, all detector components can be read out and used in the software stage. The software stage is split into two separate trigger stages called HLT1 and HLT2. In the HLT1 stage, a

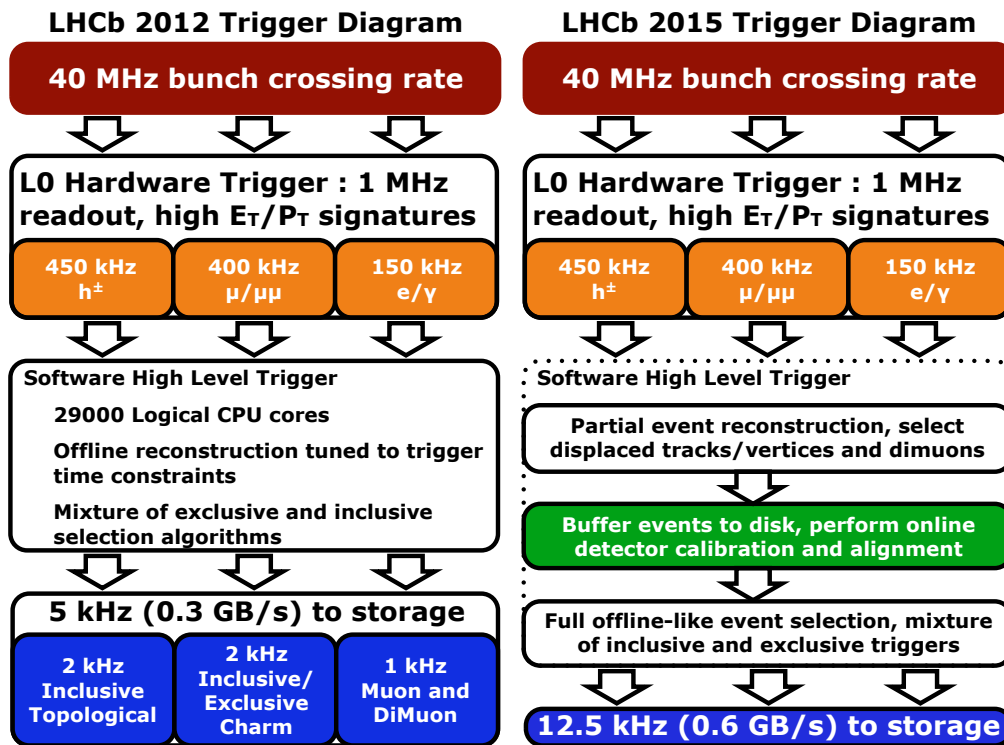


Figure 3.6 – Basic trigger schemes of the LHCb experiment [122]. The (left) layout of 2012 representative for Run 1 and the (right) layout for Run 2 are shown. The system is split into three stages, which are further discussed in the text.

partial event reconstruction is performed by either adding information from the VELO or the tracking stations. Only a part of the full detector information available is used to decrease the computing resources. Afterwards, several requirements on the track conditions are applied to confirm the decision of the L0 trigger. In the HLT2 stage the complete decay chain is then reconstructed and the conditions of the tracks are amplified in comparison to HLT1. Additionally, more detailed pre-defined selections for several physics cases are applied. Finally, after the trigger stages, the data rate is small enough to be completely written to disk. The total trigger efficiency for hadronic $B \rightarrow hh$ decays is roughly 40% [125]. The final rates that are written to disk are 3.5 kHz, 5 kHz and 12.5 kHz for 2011, 2012 and Run 2 [127], respectively.

3.3 LHCb software

At LHCb, multiple software packages are used to translate the detector output into a high-level output format. Each of them is based on the GAUDI framework [128, 129], which is a common framework for many particle physics experiments.

The conversion into a data format that is accessible for physics analyses starts with the trigger selection discussed in Sec. 3.2.3 where the software trigger stages are implemented in the MOORE package [130]. Further software is provided for the reconstruction of the decays, the application of first analysis steps and the generation of simulation. All software packages used for these specific applications are described in the following.

3.3.1 Reconstruction

The reconstruction of decays consists of two parts. First, the tracking information from several subsystems has to be combined. Second, a particle hypothesis has to be assigned. Both steps are implemented in the BRUNEL package [131]. To combine the tracks, a forward tracking algorithm is used [132]. First, the algorithm combines the hits in the individual subdetectors. Afterwards, the tracks from the VELO are associated with tracks from the tracking stations T1 to T3. Finally, information from the TT is added. Duplicated tracks, which can occur when hits are used in the reconstruction of one or more tracks, are removed. A Kalman filter [133] is used for the reconstruction of the tracks where the trajectory of the particles is determined by fitting the individual extrapolated tracks of the various detector components. In detail, the filter uses a collection of hits, information from the magnetic field and multiple scattering and energy loss from interaction with detector material. For the assignment of the particle hypothesis information from the RICH, the calorimeter system and the muon stations are used. The RICH information is used to identify long-lived hadrons such as protons, charged pions and charged kaons. By matching the rings in the RICH detector with the reconstructed tracks, the particle hypothesis can be assigned. The calorimeter system identifies electrons, photons and neutral pions. Muons are identified by extrapolating tracks into the muon chambers. For each track likelihoods for various particle hypotheses are calculated.

3.3.2 Preselection

The reconstruction of the tracks and the particle identification (PID) assignment allows to fit the whole decay chain by combining the proto particles in the raw data. By fitting the whole decay chain, the final-state particles can be tracked to the mother particles, and the vertices and properties of intermediate particles can be calculated. Two approaches are available called DecayTreeFitter (DTF) [134] or LokiVertexFitter (LVF). The LVF uses a bottom-up approach where the decay chain is fitted from the final-state particles to the head of the decay chain. The DTF uses a Kalman filter [133]. It requires more computing resources but allows to correctly account for correlations and uncertainties of common vertex positions of the intermediate particles, particle momenta, flight distances, decay times and invariant masses. While fitting the decay chain with the DTF, constraints can be placed on various properties of the decay. This results in an improved resolution in the measurement of several observables, *e.g.* for the invariant mass or decay-time

variables. Observables extracted with a DTF are used in this thesis where constraints are placed on the PV of the B meson or on the masses of intermediate particles.

The reconstruction of the whole decay chain directly after trigger selection results in a lot of data which is only stored on magnetic tape in first place. An analysis-specific centralised preselection called Stripping [135] is applied after the reconstruction. It allows to filter out decay candidates that are not interesting for the physics analysis. The remaining data can be completely saved on disk. The Stripping is performed online, and loose requirements are applied for different decay modes. These requirements are defined in Stripping lines, which are described in the DAVINCI [136] software package. The computing time can be decreased by applying this selection online because selection steps used by multiple lines only have to be performed once. Furthermore, the lines are very inclusive or contain similar decays so that signal and control channels can use the same lines. This is also the case for the decays studied in this thesis, where the same Stripping line is used to select $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays.

3.3.3 Simulation

Simulation is crucial for most analyses performed at LHCb. It is used to develop or validate models that describe the data, *e.g.* models to describe invariant mass or decay-time distributions. Simulation is also essential to calculate the selection efficiencies in the measurement of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay. Another use case is the study of resolution effects on several observables. This is possible in simulation since the true generation values are known. Due to the broad spectrum of applications, accurate simulation is needed and discrepancies with data have to be avoided which is ensured by constantly updating the simulation conditions.

Multiple software packages are used in the generation of the simulation. First, proton-proton collisions are simulated with PYTHIA [137] where a specific LHCb configuration [138] is used. The decay of the generated unstable particles is then described with the EVTGEN package [139]. In the decay of the particles, the final-state radiation is generated using PHOTOS [140]. Interactions with the detector material and the detector response are simulated using the GEANT4 toolkit [141, 142] as described in Ref. [143]. All of the above mentioned software tools are part of the GAUSS [144] package. The BOOLE software package [145] digitalises the generated simulation so that it is equal to the real detector output. Therefore, the simulation can be processed in the same way as the recorded data after the BOOLE step.

3.4 Flavour tagging at LHCb

Time-dependent measurements of CP violation in the b hadron sector require knowledge of the production flavour of the B^0 meson since the decay rates defined in Eq. (2.30) are different for B^0 and \bar{B}^0 mesons. At LHCb, the production flavour is determined by several

independent FT algorithms known as taggers. Each tagger determines a tag decision d and uses a multivariate analysis (MVA) to assign a mistag estimate η for each candidate in the data set. The tag decision is the flavour at the production of the B^0 meson, and the mistag estimate gives the probability that the decision is wrong. In general, these two quantities are determined by selecting a tagging particle which is correlated with the production mechanism of the B^0 meson. The charge of the selected particle is used to assign the tag decision, and the kinematic and PID information are used to determine the mistag estimate with a MVA. A scheme of the FT used at LHCb, including all available algorithms, is shown in Fig. 3.7. The individual taggers can be classified as a same side (SS) or a opposite side (OS) tagger, where both utilise different topological information of the event and kinematic properties of the different tagging particles. The methods each tagger exploits are briefly described in this section. Additionally, the evaluation of the FT performance is discussed. In general, the information presented in this section is mainly based on Refs. [146–148].

OS tagging At the PV, the b quarks are produced in pairs by gluon-gluon fusion as explained in more detail in Sec. 3.2. The OS tagger exploits the second b quark that does not form the B^0 meson of the signal decay. Independent from the decay chain of the signal B^0 meson, the second b also hadronises. The decay products of the second b hadron are partially reconstructed so that the initial flavour of this hadron can be determined with the charges of the decay products. Such decay products can be electrons, muons, kaons or charm hadrons. Finally, the signal b quark has to have the opposite flavour of the b of the second hadron.

Overall, five OS taggers are available. By studying the charge of the electrons and muons from semileptonic $b \rightarrow Xl^-$ decays, the decision tag can be assigned directly. These taggers are called *OSMuon* and *OSElectron*. Another tagger, called *OSKaon*, analyses the charge of kaons produced in $b \rightarrow c \rightarrow s$ transitions. The *OSVtxCharge* tagger aims to find decays of charged B mesons. The tagger reconstructs their SVs and calculates the total effective charge by summing over the charges of the decay particles weighted with their transverse momentum. Finally, the *OSCharm* tagger reconstructs D mesons produced via $b \rightarrow c$ transitions to assign a tagging decision. If the D meson is not charged, the charge of the produced kaon is used to assign a tagging decision. The *OSCharm* tagger has been added after the end of Run 1 and is consequently only available for Run 2 [148].

SS tagging In comparison to the OS taggers, the SS taggers exploit the remnants of the hadronisation of the b quark forming the signal B^0 meson. The quark forming the B^0 meson together with the b quark, *i.e.* the d quark, is also created as a $d\bar{d}$ pair. The d not forming the B^0 meson also hadronises due to confinement as described in Sec. 2.1. Usually, this d quark forms protons or pions. The charges of these protons and pions are analysed by the *SSProton* and *SSPion* taggers to assign a tag decision. The *SSPion*

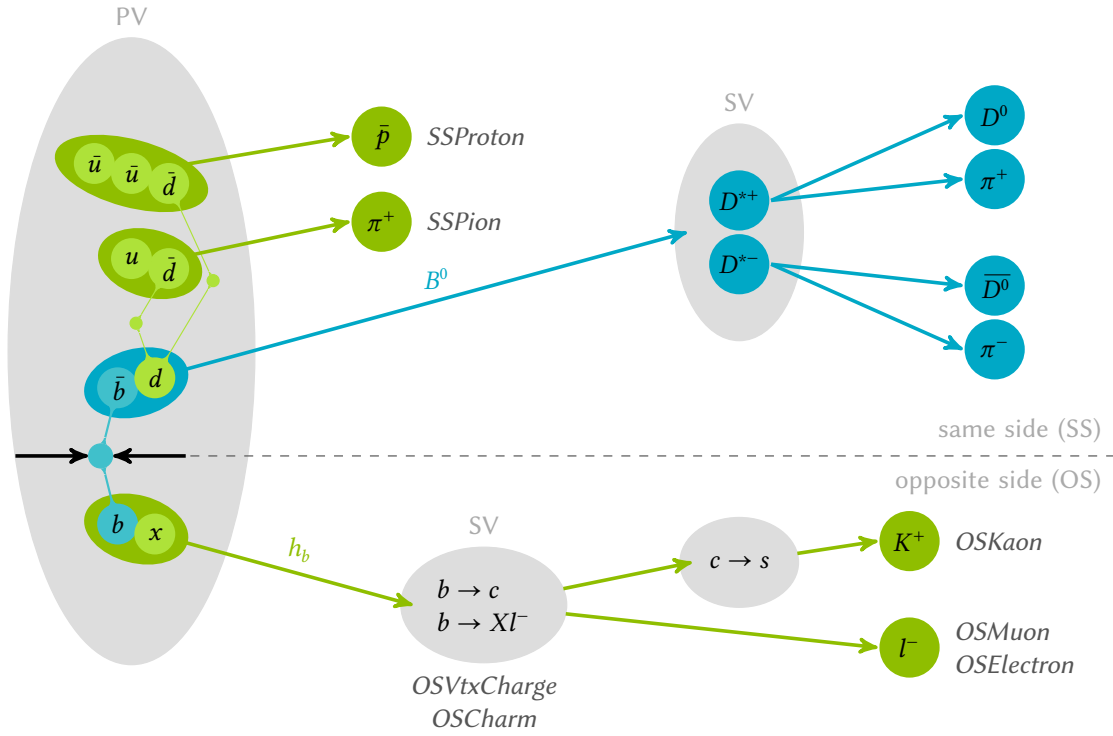


Figure 3.7 – Flavour tagging algorithms used at the LHCb experiment to tag the initial B^0 meson flavour. They can be divided into SS and OS taggers. In this analysis, the *SSProton* and *SSPion* taggers are available for the SS while for the OS the *OSElectron*, *OSMuon*, *OSKaon* and the *OSVtxCharge* are available. Additionally, the *OSCharm* tagger is used in the analysis of the Run 2 data. The taggers and their different approaches are discussed in more detail in the text. The signal b quark forming the signal B^0 meson, which decays into two charged D^* mesons, is indicated in blue.

tagger also covers another B^0 meson production mechanism in addition to production in the PV to tag the initial flavour. In detail, they also analyse events where excited B^0 mesons such as B^{*+} are produced and decay via $B^{*+} \rightarrow B^0 \pi^+$. In the case of analysing B_s^0 meson decays, another tagger called *SSKaon* is available [149]. The hadronisation of a B_s^0 leads to a s quark instead of a d quark. This s quark hadronises to a kaon. The tagging algorithm works similarly as the *SSPion* tagger for the B^0 meson, but it is not further discussed and used in this thesis since only B^0 meson decays are analysed.

Performance The selections to classify the particles are optimised centrally. Since the data-taking conditions between Run 1 and Run 2 have changed significantly, several taggers have been re-optimised for the Run 2 data-taking period [150]. A difficulty in classifying a particle arises when many tracks are present in an event. Accordingly, the performance of the FT depends on the event properties. Since the FT is not perfect, it

can happen that one or more taggers cannot assign a tag decision. The performance can be estimated using the tagging efficiency ε_{tag} , representing the fraction of candidates for which a tagger can provide a decision, and the average mistag $\langle\omega\rangle$.

All taggers aim to provide a tag decision. The possible values are +1, 0 or -1 for B^0 , no tag or \bar{B}^0 , respectively. The mistag probability lies between 0 and 0.5 where 0 is equivalent to a perfect tag decision and 0.5 is equivalent to no tag decision. The tagging efficiency and the mistag probability both affect the sensitivity of the CP observables. A measure that combines both is the effective tagging efficiency, defined as

$$\varepsilon_{\text{eff,tag}} = \frac{1}{N} \sum_i^N (1 - 2\omega_i)^2, \quad (3.1)$$

where N is the number of candidates in the data set. It is also known as tagging power since it represents a loss of statistical power compared to a perfectly tagged data set. Accordingly, this is a suitable quantity for the performance of the FT. The effect of the tagging power on the statistical uncertainty of the CP observables can be derived as $\sigma_{S,C} \sim 1/\sqrt{\varepsilon_{\text{eff,tag}} \cdot N}$.

3.4.1 Calibration

The selections for the training data of the taggers are optimised centrally as discussed in Sec. 3.4. However, the data analysed in this thesis can differ significantly from the data used in the training of the taggers. This is usually especially evident in the mistag distributions. Therefore, these differences have to be considered when performing the analysis. The corresponding procedure is known as FT calibration, which requires knowledge about the true mistag. The true mistag is defined as the ratio $\omega = N_{\text{false}}/(N_{\text{false}} + N_{\text{true}})$ where N_{true} are correctly and N_{false} falsely tagged decays. It can be determined with flavour-specific decays of neutral B mesons. These are chosen to be kinematically similar to the signal decays, such as $B^0 \rightarrow D_s^+ D^-$ decays. In these decays, the charge-conjugated decay is much more suppressed, and the charge of either the D_s^\pm or the D^\mp directly indicates the flavour of the B^0 meson. As the B^0 mesons can oscillate, only the decay flavour of the B^0 meson is known in these decays. The probability of the initial flavour being the decay flavour and the true mistag can be determined by extrapolating the decay flavour, constraining the oscillation frequency to its measured value. In general, the mistag estimate is used to obtain the true mistag probability $\omega(\eta)$ with a linear calibration of the form

$$\omega(\eta) = p_0 + p_1(\eta - \langle\eta\rangle). \quad (3.2)$$

The term $\eta - \langle\eta\rangle$ reduces correlations between the calibration parameters. In this thesis, the calibration is performed using the FTALIB software package [151]. By performing the calibration, two sets of parameters can be determined, one for B^0 and one for \bar{B}^0 . This also results in different true mistags for the two flavours. The average and the

difference of the true mistags are also calculated by FTCALIB, which are defined as

$$p_i = \frac{p_i(B^0) + p_i(\bar{B}^0)}{2} \quad (3.3)$$

and

$$\Delta p_i = p_i(B^0) - p_i(\bar{B}^0). \quad (3.4)$$

The parameters of a perfect tagger would be given by $p_0 = \langle \eta \rangle$, $p_1 = 1$, $\Delta p_0 = 0$ and $\Delta p_1 = 0$. In the calibration, potential asymmetries in the production of the B mesons, detection asymmetries or FT efficiency asymmetries are not considered.

3.4.2 Combination

The aim of any CP violation analysis is to measure the CP observables with the lowest possible statistical uncertainty. As already described, this uncertainty is directly correlated with the tagging power. Therefore, analyses aim to maximise the tagging power by using all individual FT algorithms simultaneously. This can be achieved by combining the outputs of the individual taggers. In this thesis, all described taggers are used, and two combinations are calculated, one for OS and SS. As for the calibration, the FTCALIB software package is used to calculate the combinations.

The procedure for combining several taggers is the following: First, the taggers that should be combined are calibrated individually since they have different mistag distributions. For every single tagger, the mistag probability is used to calculate probabilities for a b or \bar{b} quark. The probabilities are defined as

$$p_i(b) = \begin{cases} 1 - \omega & , \text{ if } d_i = 1 \\ \omega & , \text{ if } d_i = -1 \\ 1/2 & , \text{ if } d_i = 0 \end{cases} \quad (3.5)$$

and

$$p_i(\bar{b}) = \begin{cases} \omega & , \text{ if } d_i = 1 \\ 1 - \omega & , \text{ if } d_i = -1 \\ 1/2 & , \text{ if } d_i = 0 \end{cases} . \quad (3.6)$$

The combined probabilities are then calculated as

$$P(b) = \frac{\prod p_i(b)}{\prod p_i(b) + \prod p_i(\bar{b})} \quad (3.7)$$

and

$$P(\bar{b}) = 1 - P(b). \quad (3.8)$$

Depending on which probability is higher, the tag decision and mistag probability of the combination are assigned as

$$d_i, \omega = \begin{cases} 1, 1 - P(b) & , \text{if } P(b) > P(\bar{b}) \\ -1, P(b) & , \text{if } P(b) < P(\bar{b}) . \\ 0, \frac{1}{2} & , \text{if } P(b) = P(\bar{b}) \end{cases} \quad (3.9)$$

The combination is calibrated again to take into account the correlations between the individual taggers. This also allows to propagate the uncertainties of the individual taggers in the combination to achieve a more precise estimate for the tagging power of the combination.

4 Statistical methods

The measurements performed in this thesis rely on several statistical tools and techniques. The theoretical foundation of the most essential methods is summarised in this chapter. A common approach in high-energy physics is to reweight data sets to match the expected distributions. This reweighting is performed by applying weights to the data sets. These weights can be determined by exploiting machine learning techniques as outlined in Sec. 4.1. Maximum-likelihood fits are performed in this thesis to estimate several parameters of interest. This method is described in Sec. 4.2. Subsequently, the *sPlot* method is introduced in Sec. 4.3. This technique is used to statistically subtract individual contributions present in the data. Finally, the determination of the significance of a measurement is discussed in Sec. 4.4.

4.1 Reweighting approach

Reweighting kinematic or event-related distributions is a common task in high-energy physics analyses. This section summarises how machine learning techniques can be used to reweight data sets, where all presented information is taken from Ref. [152]. In this thesis, kinematic distributions of two different but similar decays are reweighted to match each other. Additionally, weights between generator-level simulation and full simulation are determined to account for acceptance effects.

In the following, the most common use case is considered, in which imperfect simulation has to be reweighted to match the recorded data. In general, the reweighting problem involves the estimation of the density ratio $f_{\text{data}}(\vec{x}_i)/f_{\text{sim}}(\vec{x}_i)$ as a function of the variables \vec{x}_i for an event i considered in the reweighting. The classical solution of this problem is to use histograms to estimate the densities. To reweight a distribution of the variables of interest, each bin is multiplied by the ratio of the total weights of recorded data $w_{\text{bin, data}}$ and simulation $w_{\text{bin, sim}}$ in the bin, *i.e.*

$$m_{\text{bin}} = \frac{w_{\text{bin, data}}}{w_{\text{bin, sim}}}. \quad (4.1)$$

However, despite being very intuitive, this approach introduces an issue. The number of variables that can be reweighted at the same time is limited because the amount of data that is needed to reliably estimate the density functions increases exponentially with the number of variables. This is known as the curse of dimensionality [153]. Since

this thesis requires the reweighting of multiple variables at the same time, a different approach is used.

As proposed in Ref. [154], machine learning techniques are used to estimate the density ratio. These methods are used to classify events providing the probabilities $p_{\text{data}}(\vec{x}_i)$ and $p_{\text{sim}}(\vec{x}_i) = 1 - p_{\text{data}}(\vec{x}_i)$ of a given event to be from data or simulation, respectively. These probabilities can be used to estimate the density ratio as

$$\frac{f_{\text{data}}(\vec{x}_i)}{f_{\text{sim}}(\vec{x}_i)} \sim \frac{p_{\text{data}}(\vec{x}_i)}{p_{\text{sim}}(\vec{x}_i)}. \quad (4.2)$$

Using this approach has the advantage of preventing the curse of dimensionality [153] but provides inaccurate predictions when the density ratio is high or low. The reason is that certain regions in the variable space are more significant for the reweighting, but not for the classification task. Predicting the class of an event in a region in the variable space where the density ratio is high or low is simple since the events belong to one class. However, since these regions provide smaller contributions to the loss function, a machine learning technique focuses on other regions. To solve this issue, a Gradient Boosted Decision Tree (GBDT) [155] is used to split the variable space into several regions. A GBDT consists of several decision trees, which naturally divide a variable space by checking simple decisions. The leaves of the tree are then associated with one region of the variable space. Suitable reweighting regions are found by a greedy optimisation of the symmetrised χ^2 defined as

$$\chi^2 = \sum_l \frac{(w_{l, \text{sim}} - w_{l, \text{data}})^2}{w_{l, \text{sim}} + w_{l, \text{data}}}, \quad (4.3)$$

where l indicates the leaves of the tree. By maximising the χ^2 important regions for the reweighting are found.

The GBDT used in the reweighting consists of many trees, which are trained one after another. In the training, first, a shallow tree is built to maximise the symmetrised χ^2 . Afterwards, the predictions

$$l_{\text{pred}} = \ln \frac{w_{l, \text{sim}}}{w_{l, \text{data}}} \quad (4.4)$$

are calculated in each leaf. Finally, the simulation distribution is reweighted with the weights

$$w = \begin{cases} w, & \text{if event from data} \\ w \cdot e^{p_e}, & \text{if event from simulation} \end{cases}, \quad (4.5)$$

where the per-event prediction p_e is equal to the prediction of a leaf containing the event. In general, the procedure is similar to the naive binning approach. The main difference arises from the difference in the selection of the bins. In this thesis, a Gradient Boosted

reweighter (GBreweighter) [152] included in the HEP-ML software package [156] is used to reweight the data sets.

4.2 Maximum-likelihood fits

Unbinned maximum-likelihood fits [157] are essential for this thesis to estimate the parameters of interest and their uncertainties from a data set. Assuming a data set containing n measured values x_1, x_2, \dots, x_n and a known probability density function (PDF) $P(\vec{x} | \vec{\rho})$ describing the measured values with a set of parameters $\vec{\rho}$, the likelihood function is defined as

$$\mathcal{L}(\vec{\rho}) = \prod_{i=1}^n P(\vec{x}_i | \vec{\rho}). \quad (4.6)$$

Accordingly, the higher the likelihood, the better the parameters $\vec{\rho}$ describe the measured values. However, in practice, it is common to maximise the logarithm of the likelihood function, the log-likelihood $L(\vec{\rho}) = \ln \mathcal{L}(\vec{\rho})$, instead. This approach reduces the complexity of the numerical calculations needed to determine the maximum. The reason is that it is computationally less expensive to calculate the sum of the logarithms of the probabilities than the product of the probabilities. Due to the monotonicity of the logarithm function, the maximum of the log-likelihood function is equal to the maximum of the likelihood function. Another adjustment is to minimise the negative log-likelihood $F(\vec{\rho})$. This is due to the fact that most algorithms are designed to minimise a function. The negative log-likelihood is defined as

$$F(\vec{\rho}) = -\ln \mathcal{L}(\vec{\rho}) = -\sum_{i=1}^n \ln P(\vec{x}_i | \vec{\rho}). \quad (4.7)$$

The computationally most expensive part of the minimisation is the normalisation of the PDF. For all parameters of $\vec{\rho}$ the PDF is normalised to unity, *i.e.*

$$\int P(\vec{x} | \vec{\rho}) d\vec{x} = 1, \quad (4.8)$$

which has to be maintained during all intermediate steps of the minimisation of $F(\vec{\rho})$.

The likelihood $\mathcal{L}(\vec{\rho})$ can be extended with a Poisson term, which indicates the probability of observing n candidates while expecting \mathcal{N} candidates. This allows the determination of the absolute number of candidates while also considering statistical fluctuations of the rate around the expectation value. The approach is called an extended maximum-likelihood fit [158]. The resulting likelihood function is defined as

$$\mathcal{L}_{\text{ext}}(\vec{\rho}) = e^{-\mathcal{N}} \frac{\mathcal{N}^n}{n!} \prod_{i=1}^n P(\vec{x}_i | \vec{\rho}). \quad (4.9)$$

The likelihood function is additionally modified by applying Gaussian constraints. This technique allows the inclusion of uncertainties of observables known from other measurements and not measured in this thesis. In detail, for each constraint, the likelihood is multiplied by a Gaussian distribution $\mathcal{G}(\rho_i; \mu, \sigma)$ defined as

$$\mathcal{G}(\rho_i; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\rho_i - \mu)^2}{2\sigma^2}}, \quad (4.10)$$

where the mean μ and the width σ of the constraint are the mean and uncertainty of the external input.

All fits in this thesis are performed using the MINUIT algorithm [159] implemented by the RooFIT package [160], which is part of the ROOT framework [161, 162]. The RooFIT software package is a commonly used tool for data modelling and fitting in high-energy physics. It provides a set of tools and methods to perform maximum-likelihood fits.

4.3 sPlot method

After the selection, the recorded data set still contains several other contributions besides the $B^0 \rightarrow D^{*+}D^{*-}$ signal decay. For example, one contribution is from the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay, which has similar selection efficiencies and kinematic properties. But also combinatorial background often remains in the data, since a fraction of this type of background cannot be removed even with very strict selection criteria. However, some maximum-likelihood fits performed in this thesis require a data set which only contains the signal decay. A solution to this problem is the application of the *sPlot* method [163, 164], which allows to unfold the signal decay from the data set. The technique requires a discriminating variable where the distributions of the contributions in this variable are known. In high-energy physics, the reconstructed invariant mass is often used as a discriminating variable.

Assuming that the individual contributions c in the discriminating variable m are described by the PDFs $\mathcal{P}_c(m)$, the individual shapes of the contributions can be determined by an extended maximum-likelihood fit as described in Sec. 4.2. With the determined shapes, the per-event weights $w_c^{\mathcal{S}}$ known as *sWeights* are defined as

$$w_c^{\mathcal{S}} = \frac{\sum_{j=1}^{N_c} V_{cj} \mathcal{P}_j(m)}{\sum_{k=1}^{N_c} \mathcal{N}_k \mathcal{P}_k(m)}, \quad (4.11)$$

where V is the covariance matrix of the expected candidates \mathcal{N} of the individual contributions and N_c is the number of contributions. The matrix is defined as

$$V_{cj}^{-1} = \sum_{i=1}^N \frac{\mathcal{P}_c(m_i) \mathcal{P}_j(m_i)}{\left(\sum_{k=1}^{N_c} \mathcal{N}_k \mathcal{P}_k(m_i) \right)^2}, \quad (4.12)$$

where N is the total number of events in the data set. By the construction of the *sWeights*, it is required that the sum of the weights $\sum_{i=1}^N w_c^{\delta_i}$ is equal to the number of events of the respective contribution.

The signal component can be unfolded from the data in a distribution of a certain variable of interest t , if the variable t factorises with the discriminating variable m in the form $\mathcal{P}_c(m, \vec{t}) = \mathcal{P}_c(m) \mathcal{P}_c(\vec{t})$. In this case, the weights w_c^{δ} can be used to statistically subtract the individual contributions from the data. In maximum-likelihood fits the weights are considered in the fit to a distribution t by implementing them into the likelihood function given in Eq. (4.7). The resulting likelihood function is defined as

$$F_{w_c^{\delta}}(\vec{\rho}) = - \sum_{i=1}^n w_c^{\delta}(\vec{t}_i) \ln P(\vec{t}_i | \vec{\rho}). \quad (4.13)$$

4.4 Significance

The significance is a measure of the probability that a deviation in the number of events relative to a null hypothesis is not a statistical background fluctuation. It provides information about whether the measurement equals an evidence or observation. Two arbitrary cut-off values are common in high-energy physics. If the significance is above 3σ , it equals an evidence for the measured decay, and if it is above 5σ , an observation of the measured decay has been made.

The significance of the measurement of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay given in standard deviations σ is calculated using Wilks theorem [165]. The theorem uses the likelihoods from two separate maximum-likelihood fits performed as described in Sec. 4.2 to determine the significance. The first fit is performed with the default PDF, and the second fit is a repetition of the first fit where the null hypothesis is applied to the parameter of interest. The null hypothesis is defined as the case where no $B_s^0 \rightarrow D^{*+} D^{*-}$ contribution is present. Consequently, the yield of the contribution is set to zero in the fit, *i.e.* $\mathcal{N}_{B_s^0} = 0$. These likelihoods are denoted as $\mathcal{L}_{\text{free}}$ and \mathcal{L}_0 for the default fit and the null hypothesis fit, respectively. The resulting significance is then given by

$$\text{significance} = \sqrt{-2 \ln (\mathcal{L}_0 / \mathcal{L}_{\text{free}})}. \quad (4.14)$$

5 Analysis strategies

Two separate analyses are performed, one focussing on the measurement of the branching fraction of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay¹, and the other on the measurement of CP observables in the $B^0 \rightarrow D^{*+}D^{*-}$ decay. Both measurements are performed using the Run 1 and Run 2 data sets of the LHCb experiment corresponding to integrated luminosities of 3 fb^{-1} and 6 fb^{-1} , respectively.

The $D^{*\pm}$ mesons from the B^0 or B_s^0 meson decays are reconstructed using $D^{*\pm} \rightarrow D^0\pi^\pm$ decays. Three final states for the subsequent D^0 decays are considered in this thesis, *i.e.* $D^0 \rightarrow K^-\pi^+$, $D^0 \rightarrow K^+K^-$, and $D^0 \rightarrow \pi^+\pi^-$. These are the most common two-body D^0 decays with charged final-state particles. The $D^0 \rightarrow K^+\pi^-$ decay is not considered since it is doubly Cabibbo suppressed. Neutral final-state particles are not taken into account as they are difficult to reconstruct within the LHCb environment. No three body D^0 decay modes are considered since the reconstruction efficiency decreases with more final-state particles. Only fully hadronic decay modes are considered in this thesis, since the reconstruction of semi-leptonic decays is more challenging due to the presence of neutrinos. The unreconstructed neutrinos dilute the mass resolution, which makes the extraction of signal decays and the unfolding of signal and background more challenging. Additionally, they worsen the decay-time resolution, which dilutes the CP parameters. The same decay chain is used in both analyses to reconstruct the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays. Further, the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis uses the $B^0 \rightarrow D^{*+}D^{*-}$ decay as a normalisation channel to reduce uncertainties. The identical reconstruction allows both decays to be selected simultaneously, provided that only loose requirements are applied to the B meson². In comparison with other B factories operating at electron-positron colliders, the hadronic environment at the LHC has a much higher track multiplicity. This requires an effective suppression of backgrounds. Typically, it is differentiated between two types of backgrounds, *i.e.* combinatorial background and physical background from other decays. Combinatorial background originates from random combinations of unrelated tracks, forming B meson candidates. Backgrounds from other decays mainly arise from the misidentification of final-state particles or from only partially reconstructed B decays. The selection of the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays and the suppression of backgrounds are described in Chap. 6. Both analyses use very similar selection criteria and are therefore described together to avoid redundancy.

¹Charge-conjugated processes are implicitly included in the following unless clearly stated.

²Throughout this thesis a B candidate or a $B \rightarrow D^{*+}D^{*-}$ decay is used to refer to a possible B_s^0 or B^0 candidate or the respective decays.

The measurement of the $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction is the first of its kind. It is measured relative to the $B^0 \rightarrow D^{*+}D^{*-}$ decay, which is also used to optimise the selection of this analysis. The calculation of the ratio of branching fractions as given in Eq. (2.46) requires knowledge about the selection efficiencies, the ratio of hadronisation fractions, and the ratio of signal decays. The selection efficiencies are determined in Chap. 6, and the ratio of hadronisation fractions is an external input. Consequently, the ratio of signal decays remains to be determined. The number of signal decays is extracted by performing extended maximum likelihood fits to the invariant $D^{*+}D^{*-}$ mass distribution. Systematic uncertainties on the ratio of branching fractions are also evaluated. Finally, the effects of the B_s^0 lifetime on the efficiency ratio are studied. The analysis has been published in Ref. [40] and is presented in Chap. 7 apart from the selection.

The second analysis presented in this thesis focuses on the measurement of CP violation in $B^0 \rightarrow D^{*+}D^{*-}$ decays. Here, maximum-likelihood fits to the invariant $D^{*+}D^{*-}$ mass are performed to determine $sWeights$. These are used to unfold the signal and background contributions in the subsequent analysis steps. Several experimental effects are considered, and the decay-time PDF is adjusted accordingly. These include the FT calibration, the decay-time acceptance and the decay-time resolution. Since the CP parameter $S_{D^{*+}D^{*-}}$ is diluted due to the interference of two CP eigenstates, these eigenstates are determined by performing a simultaneous fit to the angular distributions in the transversity basis as described in Sec. 2.5. Consequently, angular acceptance effects also have to be considered in the extraction of the CP observables. A fit to the angular and decay-time distributions is performed to extract the CP parameters and the CP -odd fraction. Finally, systematic uncertainties are evaluated. The complete analysis, apart from the selection, is presented in Chap. 8. The analysis is still in internal review within the LHCb collaboration. Consequently, it has not yet been published, and the analysis is still blinded to avoid any bias towards a naive expectation of the central values.

The author of this thesis is the main proponent of both analyses. He has implemented the frameworks and pushed forward all relevant parts of both analysis on his own. If something is adopted from other analyses, this is explicitly mentioned in the following.

6 Selection of signal decays

The LHCb experiment has collected a substantial amount of data during the Run 1 and Run 2 data-taking periods. However, only a small fraction of this data is relevant for the analyses performed in this thesis since the measured and expected branching fractions of the decays of interest are low. Therefore, the data has to be filtered, and only events with respect to the analyses are selected. To reduce contributions from other decays, several selection steps are applied to obtain clean data with a suitable signal to background ratio for further physics analyses.

The signal decays of the two analyses presented in this thesis are the $B_s^0 \rightarrow D^{*+} D^{*-}$ and $B^0 \rightarrow D^{*+} D^{*-}$ decays. The $B^0 \rightarrow D^{*+} D^{*-}$ decay is also used as a control channel in the $B_s^0 \rightarrow D^{*+} D^{*-}$ analysis because it is less suppressed and, therefore, offers higher statistics. The same selection strategy is applied in both analyses. To reduce textual redundancies, this chapter focuses on the discussion of the selection of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay. The differences to the $B^0 \rightarrow D^{*+} D^{*-}$ selection are explicitly mentioned in the following.

Both decays are reconstructed, exploiting the same topology and final-state particles. The B mesons are produced in proton-proton collisions at the interaction point. After flying approximately 1 cm through the detector, they decay into two oppositely charged $D^{*\pm}$ mesons at the SV. The $D^{*\pm}$ mesons decay almost instantaneously as $D^{*\pm} \rightarrow D^0 \pi^\pm$ decays, where the D^0 mesons decay after a measurable distance from the SV. This enables a separation of the decay vertices of the $D^{*\pm}$ and D^0 mesons. The D^0 mesons are then reconstructed in the most dominant two-hadron decay modes with charged final-state particles, *i.e.* the $D^0 \rightarrow K^- \pi^+$, $D^0 \rightarrow K^+ K^-$ and $D^0 \rightarrow \pi^+ \pi^-$ modes. No decays with neutral final-state particles are considered because their reconstruction efficiency is significantly lower. A similar statement holds for the three-body decay modes, which are not considered because more final-state particles would also significantly decrease the reconstruction efficiency. The branching fractions of the considered modes are given in Table 6.1.

Several selection steps are applied to reduce multiple types of backgrounds that can be present in the data. One such type is combinatorial background, which consists of falsely reconstructed candidates and occurs when a random combination of unrelated tracks is used to form a candidate. The cause of this background is the high track multiplicity in a proton-proton collision and the limited resolution of the detector. The combinatorial background cannot be completely removed, and its contribution must be considered in the later stages of the analyses. Not only combinatorial but also physical background can be present in the data. This background can arise from the misID of final-state

Table 6.1 – The measured branching fractions of the D^0 decay modes considered in this thesis. The most dominant two-body decays with charged final-state particles are studied. The values are taken from Ref. [47].

Mode	\mathcal{B} [10^{-3}]
$D^0 \rightarrow K^- \pi^+$	39.47 ± 0.30
$D^0 \rightarrow K^+ K^-$	4.08 ± 0.06
$D^0 \rightarrow \pi^+ \pi^-$	1.454 ± 0.024

particles, where tracks are wrongly reconstructed as $B \rightarrow D^{*\pm} D^{*-}$ decays. Physical background can also arise from charmless or single-charm decays. They can occur when no or only one $D^{*\pm}$ meson is present as an intermediate state. Partially reconstructed background is also considered a physical background. It can arise by an incomplete reconstruction of another multi-body decay, which mimics the signal decay and where at least one particle is not measured. Overall, the physical background cannot always be removed completely. Consequently, if it remains present after the application of the whole selection, it has to be taken into account later in the analyses.

The selection begins with a cut-based preselection discussed in Sec. 6.1. In this thesis, the B mesons are reconstructed with subsequent $D^{*\pm} \rightarrow D^0 \pi^\pm$ decays. In analyses with $D^{*\pm} \rightarrow D^0 \pi^\pm$ decays, it is common to apply requirements on the mass difference between the $D^{*\pm}$ and the D^0 mesons. Such requirements are further considered in the suppression of combinatorial background described in Sec. 6.2. After applying this selection step, multiple signal decays resulting from imperfections in the reconstruction can be present in one event. Given the predicted and the measured branching fractions of the signal decays, it is highly unlikely that the signal decays present in an event are truly independent. Therefore, they must be removed so that only one decay remains in each event. This procedure is described in Sec. 6.3. After applying the whole selection, multiple possible background sources are studied in Sec. 6.4. Finally, the selection efficiencies for both decay channels and the result of the selection are discussed in Sec. 6.5.

6.1 Preselection

A loose preselection is performed centrally by the LHCb experiment, which is referred to as Stripping. It is discussed in more detail in Sec. 3.3.2. Between the Run 1 and Run 2 data-taking periods, the Stripping conditions have changed due to a continuous experiment-wide development of the selection strategies. The requirements of the different Stripping versions are taken from Ref. [135] and listed in Table 6.2. The development and optimisation of these cuts is not an explicit part of this thesis. Therefore, this section only briefly explains the variables and various requirements.

Table 6.2 – Requirements of the loose centralised preselection known as Stripping. The selection step is discussed in more detail in Sec. 3.3.2. 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 compared to Run 1. The variables are explained in more detail in the text.

Candidate	Variable	Requirement	Unit
	#longtracks	< 500	
	Hlt2Topo or Hlt2IncPhi	passed	
B^0	$\chi_{\text{vtx}}^2/\text{ndf}$	< 10	
	DIRA _{best PV}	> 0.999	
	$\chi_{\text{IP, PV}}^2$	< 25	
	m	> 4750	MeV/c ²
	m	< 6000	MeV/c ²
	$\sum_{\text{daughters}} p_{\text{T}}$	> 5000	MeV/c
	t	> 0.2	ps
one daughter	p	> 10000	MeV/c
hadron	p_{T}	> 1700	MeV/c
	$\min(\text{IP}_{\text{any PV}})$	> 0.1	mm
	$\min(\chi_{\text{IP, PV}}^2)$	> 16	
$D^0 / D^{*\pm}$	$\chi_{\text{vtx}}^2/\text{ndf}$	< 10	
	DIRA _{best PV}	> 0	
	DOCA	< 0.5	mm
	$\chi_{\text{IP, PV}}^2$	> 36	
$D^{*\pm}$	$ m_{D^{*\pm}, \text{PDG}} - m_{D^{*\pm}} $	< 50 (< 600)	MeV/c ²
	$ m_{D^{*\pm}} - m_{D^0} $	– (< 200)	MeV/c ²
D^0	$\sum_{\text{daughters}} p_{\text{T}}$	> 1800	MeV/c
	$ m_{D^0, \text{PDG}} - m_{hh} $	< 100	MeV/c ²
one daughter	p_{T}	> 500	MeV/c
hadron	p	> 5000	MeV/c
	$\chi_{\text{track}}^2/\text{ndf}$	< 2.5 (< 4.0)	
π^\pm / K^\pm	$\chi_{\text{track}}^2/\text{ndf}$	< 3.0 (< 4.0)	
	p_{T}	> 100	MeV/c
	p	> 1000	MeV/c
	$\min(\chi_{\text{IP, PV}}^2)$	> 4.0	
	p_{ghost}	< 0.4	
π^\pm	DLL _{Kπ}	< 20	
K^\pm	DLL _{Kπ}	> –10	

Two general requirements related to event characteristics are applied. First, it is required that the number of long tracks is less than 500. Such tracks stem from charged particles that have their origin in the VELO and at least travel throughout all tracking stations. This requirement ensures that the event multiplicity is not too high to separate all tracks cleanly. No requirements are applied to the L0 hardware trigger or the HLT1 trigger, only to the HLT2 trigger. The events must have passed one of the topological or inclusive ϕ trigger lines. The topological lines require a two-, three- or four-track SV, which is significantly displaced from any PV. To identify secondary vertices that are consistent with a decay of a b hadron, a multivariate algorithm is used [166, 167]. Further, at least one charged particle must have a large transverse momentum and be inconsistent with originating from a PV. The inclusive ϕ trigger lines search for $\phi \rightarrow KK$ decays. However, the contribution of these lines compared to the topological lines is not significant. The fraction of events remaining after applying the whole selection triggered only by the inclusive ϕ lines is estimated using simulated $B_s^0 \rightarrow D^{*+} D^{*-}$ and $B^0 \rightarrow D^{*+} D^{*-}$ decays. The fractions are determined to be below 0.02% and 0.4% in Run 1 and Run 2, respectively. Conclusively, no further studies about the impact of the inclusive ϕ lines are performed. A more detailed description of the trigger lines can be found in Refs. [124–126].

The reconstructed particles have to fulfil several requirements. A high quality of reconstructed vertices and tracks is required by evaluating the reduced chi-squared statistic for the vertices $\chi_{\text{vtx}}^2/\text{ndf}$ and the tracks $\chi_{\text{track}}^2/\text{ndf}$, which is used in both cases for determining the goodness-of-fit of the reconstructed object. The quantity ndf represents the number of degrees of freedom. The $\chi_{\text{vtx}}^2/\text{ndf}$ requirement is applied to the B meson and the intermediate D^* mesons¹ and D^0 mesons, while the $\chi_{\text{track}}^2/\text{ndf}$ requirement is applied to all final-state particles. This also ensures that the D^* mesons form a common vertex. Additionally, the B , D^* , and D^0 mesons must have certain transverse momenta p_T , total momenta p and invariant masses m . For example, the daughters of the B and D^0 candidates must have a total transverse momentum of at least 5000 MeV/c and 1800 MeV/c, respectively. These requirements ensure the decay of a heavy meson. Additionally, the cuts on the masses reject wrongly reconstructed candidates because randomly combined candidates rarely fulfil these requirements. To suppress backgrounds arising from particles directly produced in the PV, the lifetime of the B meson has to be larger than 0.2 ps. The B candidates also must have a χ_{IP}^2 smaller than 25 with respect to the best reconstructed PV. This variable is defined as the difference in the vertex fit χ^2 of a given PV reconstructed with and without the candidate being considered. The final-state particles and the intermediate D mesons² are required to have a high χ_{IP}^2 with respect to any PV. The impact parameter $\text{IP}_{\text{any PV}}$ represents the shortest distance between any primary vertex and the reconstructed track. Applying requirements on the impact parameter for the D^* mesons ensures that they are not directly produced in the PV.

¹Throughout this thesis, the D^* meson always refers to the D^{*+} and the D^{*-} meson.

²In the following, D mesons refer to $D^{*\pm}$ and D^0 mesons, unless clearly stated.

Further, the cosine of the angle between the momentum and the flight direction, called DIRA, is checked for the B meson and the D mesons. The flight direction is calculated with the best PV and the decay vertex. In the case of the B mesons, the decay vertex is the SV, while for the D mesons, it is always the D^0 decay vertex. The best PV is chosen with respect to the lowest χ_{IP}^2 . The angle has to be small for the B candidates, and for the D mesons, the angle has to point into the detector plane. The distance of the closest approach (DOCA) between all possible combinations of tracks forming D vertices has to be smaller than 0.5 mm. This ensures a correct vertex reconstruction. Requirements are made not only on the masses themselves but also on mass differences. These are applied on the difference of the reconstructed D meson masses and their measured world averages taken from Ref. [47] and on the mass differences of the D^* and the corresponding D^0 mesons. By calculating the difference between the log-likelihoods of two-particle hypotheses, called delta log-likelihood DLL, requirements can be made on the identification of pions and kaons by using information from the PID system. Finally, a small ghost probability [168] is required for all final-state particles. This variable corresponds to the probability that the reconstructed track does not stem from a real particle but from a random combination of hits in the detector components.

The Stripping conditions have changed between Run 1 and Run 2, resulting in slightly different requirements for a few variables, which is indicated in Table 6.2. The requirements are aligned to allow a better comparison of the selection efficiency between the two runs. This is done by applying the tightest cut to both data sets. Despite only loose requirements being applied in the Stripping, signal decay can still be rejected. It can also happen that B candidates are not completely reconstructed and rejected because they do not decay within the detector acceptance. No separate simulation is available for both effects. The combined selection efficiencies for the Stripping and the reconstruction are determined to be

$$\epsilon_{\text{Str, Rec}}^{\text{Run 1}} = (0.3397 \pm 0.0033)\% \quad \text{and} \quad \epsilon_{\text{Str, Rec}}^{\text{Run 2}} = (0.4725 \pm 0.0034)\% \quad (6.1)$$

for the $B^0 \rightarrow D^{*+} D^{*-}$ decay and

$$\epsilon_{\text{Str, Rec}}^{\text{Run 1}} = (0.3529 \pm 0.0034)\% \quad \text{and} \quad \epsilon_{\text{Str, Rec}}^{\text{Run 2}} = (0.5087 \pm 0.0036)\% \quad (6.2)$$

for the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay. An improvement in the Run 2 efficiencies can be observed due to the increased centre-of-mass energy resulting in higher momenta and due to more efficient trigger and particle identification algorithms. The higher momentum improves the reconstruction efficiency as it implies a higher transverse momentum. A particle with a higher transverse momentum traverses the detector further away from the beam pipe. Thus, such particles can be better reconstructed due to a larger distance to the high particle stream near the beam pipe. The efficiency of the B_s^0 decay is slightly higher due to the higher mass of the B_s^0 meson compared to the B^0 meson, which results in higher transverse momenta of the daughter particles.

In both analyses, variables from the DTF described in Sec. 3.3.2 are used, where constraints are placed on the PV and on the masses of the $D^{*\pm}$ and D^0 mesons. An additional set of variables is used in the CP violation analysis. It is determined with a DTF where only the PV is constrained and is exploited in the combinatorial background suppression. In general, it is required that the fits of the DTF have to be converged. The selection efficiencies for the measurement of the ratio of branching fractions are determined to be

$$\epsilon_{\text{DTF}}^{\text{Run 1}} = (97.71 \pm 0.14)\% \quad \text{and} \quad \epsilon_{\text{DTF}}^{\text{Run 2}} = (97.75 \pm 0.11)\% \quad (6.3)$$

for the $B^0 \rightarrow D^{*+} D^{*-}$ decay and

$$\epsilon_{\text{DTF}}^{\text{Run 1}} = (97.25 \pm 0.16)\% \quad \text{and} \quad \epsilon_{\text{DTF}}^{\text{Run 2}} = (97.600.11)\% \quad (6.4)$$

for the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay.

6.2 Combinatorial background suppression

The suppression of combinatorial background is the main selection step, which is applied after the Stripping and the offline preselection. The suppression is achieved by exploiting the topology of the $B \rightarrow D^{*+} D^{*-}$ and the two subsequent $D^{*\pm} \rightarrow D^0 \pi^\pm$ decays. The available phase space is small in the $D^{*\pm} \rightarrow D^0 \pi^\pm$ decays since the $D^{*\pm}$ mass of $2010.26 \text{ MeV}/c^2$ [47] is only slightly higher than the D^0 mass of $1864.84 \text{ MeV}/c^2$ [47]. Consequently, the pion in this decay has a small kinematic energy. Because of this property, it is also called a slow pion. Following this insight, the mass difference between the $D^{*\pm}$ mesons and the respective D^0 mesons is calculated. A peak slightly higher than the pion mass of $139.57 \text{ MeV}/c^2$ [47] is then present in the distribution of the mass difference. All signal decays are found within this peak. At higher masses, only background events are present. This is validated by comparing recorded data with simulation, which is shown in Fig. 6.1. Overall, the cut is very efficient and has a high background suppression. It is applied for both $D^{*\pm}$ decays. The analysis of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decays uses the default reconstructed $D^{*\pm}$ and respective D^0 masses with $m_{D^{*\pm}} - m_{D^0} < 150 \text{ MeV}/c^2$, while the $B^0 \rightarrow D^{*+} D^{*-}$ analysis uses the masses reconstructed by a DTF constraining the PV with $143 \text{ MeV}/c^2 < m_{D^{*\pm}}^{\text{DTF}} - m_{D^0}^{\text{DTF}} < 148 \text{ MeV}/c^2$. The constraint on the PV further improves the resolution of the individual masses and the mass difference. This allows the use of a tighter requirement and enables a better background suppression. In the $B_s^0 \rightarrow D^{*+} D^{*-}$ analysis, one further requirement is applied to the quality of the DTF with constraints on the masses and the PV, where the candidates have to fulfil the requirement $\chi_{\text{DTF}}^2 < 50$. This variable also provides a good separation of signal and background, which is also shown in Fig. 6.1 by comparing simulation and data. For the $B^0 \rightarrow D^{*+} D^{*-}$ analysis, the χ_{DTF}^2 requirement has been removed since the background suppression of the mass difference calculated with the DTF variables is sufficient.

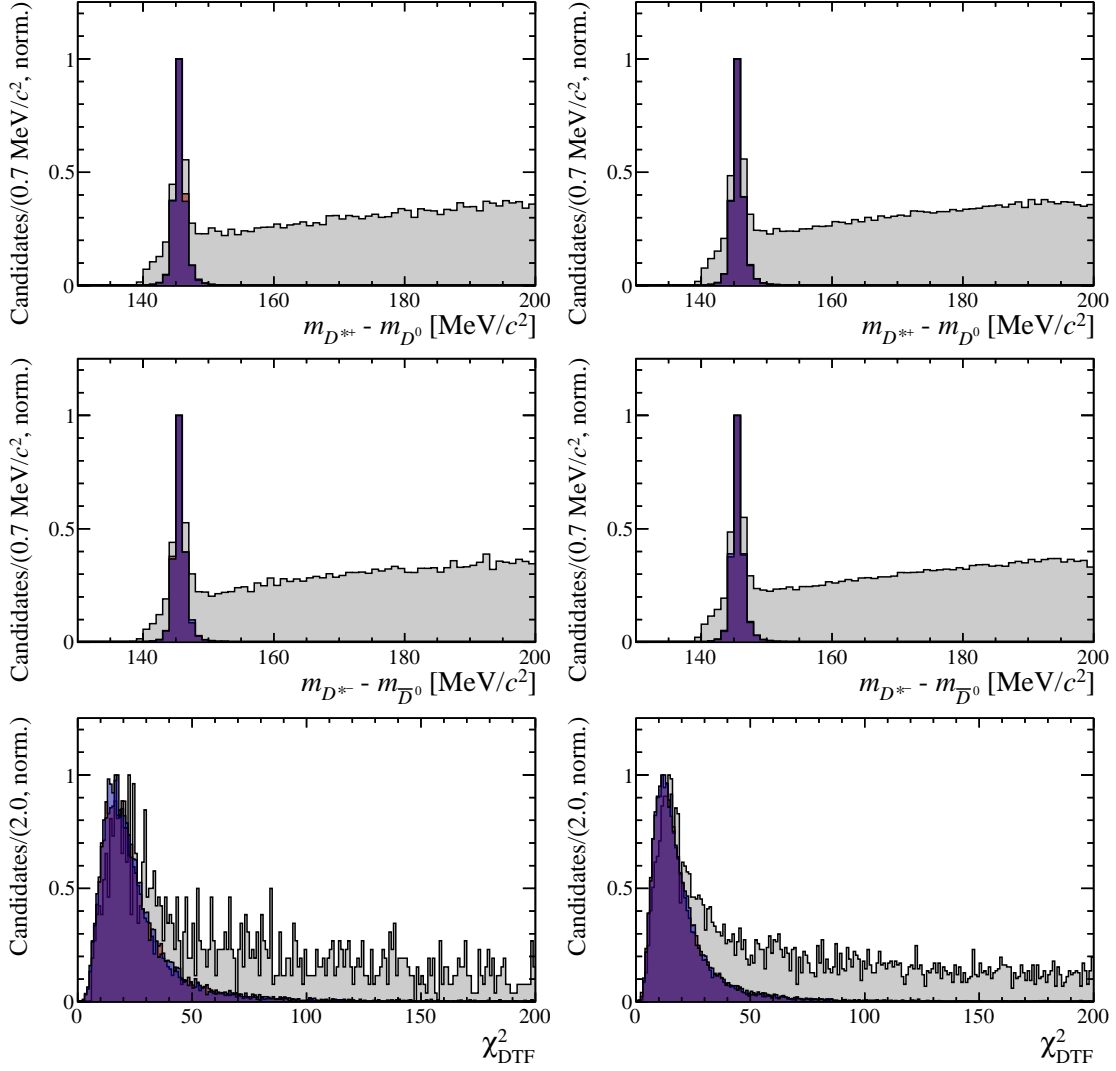


Figure 6.1 – Distributions of variables used in the suppression of combinatorial background in the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis, *i.e.* the differences of the invariant (top) D^{*+} and D^0 masses and of the (mid) D^{*-} and \bar{D}^0 masses and the (bottom) χ_{DTF}^2 distribution for (left) Run 1 and (right) Run 2, respectively. The distributions of $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ simulation are shown in blue and red, and the recorded data is shown in grey.

No additional MVA is performed after the selection step presented in this section, as it is common in high-energy physics analyses. The reason is that a MVA is not efficient since only a few training candidates are left for the background proxy. The number of background candidates that could be used in the training, *i.e.* the candidates that have an invariant $D^{*+}D^{*-}$ mass greater than $5400 \text{ MeV}/c^2$, amounts to 17 for Run 1 and 104 for Run 2. Another approach could be completely replacing the selections presented in this section with a MVA. This has been studied in the early stages of the analysis,

where a GBDT [155] from the XGBOOST framework [169] is used to suppress the combinatorial background [170, 171]. Although the suppression of the background was slightly higher with the GBDT than with the $B_s^0 \rightarrow D^{*+}D^{*-}$ selection presented in this section [170], the GBDT is mainly pushed by the χ^2 variable of the DTF. Besides that, not enough background is left in the data after applying the GBDT to adequately describe its contributions in discriminative variables as the invariant mass. Consequently, only cut-based selections are used in both analyses.

A cross-check is performed to show that the simulation describes the recorded data well in all variables that are used in the selection step applied in this section. This cross-check is given in Fig. 6.2 for $B^0 \rightarrow D^{*+}D^{*-}$ decays, where the distributions of simulation are shown overlaid with signal-weighted data distributions. The comparison for $B_s^0 \rightarrow D^{*+}D^{*-}$ weighted data is not shown since it shows the same trend but is not as conclusive as the $B^0 \rightarrow D^{*+}D^{*-}$ distributions due to the significantly lower statistics. The weights are determined using the *sPlot* method described in Sec. 4.3, where the discriminative variable is the invariant $D^{*+}D^{*-}$ mass. Overall, no significant deviation between the simulation and the signal-weighted data is observed.

6.3 Multiple candidates removal

The selection steps are applied to each candidate present in the data sets. However, more than one signal decay can be present in one event. These are called multiple candidates and are rejected in this section so that only one signal decay remains in each event. Two categories of multiple candidates are considered in the selection. The first category contains multiple candidates that involve unrelated final-state particles in the event. They occur in 0.2% and 1.31% of all events for Run 1 and Run 2, respectively, and are treated by randomly removing all but one candidate in each event. The second category contains multiple candidates that originate from misID of a final-state particle. They are identified by having identical tracks for all final-state particles, but where one kaon has been misidentified as a pion or vice versa. For example, in addition to a correctly reconstructed $D^0 \rightarrow K^+K^-$ decay, the decay $D^0 \rightarrow K^-\pi^+$ is reconstructed due to misidentification. Both reconstructed candidates can pass the selection and be interpreted as signal decay. This type of multiple candidates occurs in 1.43% and 2.48% of all events for Run 1 and Run 2, respectively. The number of multiple candidates for each category in recorded data is listed in Table 6.3. In the $B^0 \rightarrow D^{*+}D^{*-}$ analysis, the fractions of multiple candidates differ due to the changes in the selection requirements. The combined fractions for both categories are 1.9% and 2.04% for Run 1 and Run 2, respectively. In addition, the modified selection suppresses multiple candidates from misID as the fraction of events with such candidates is reduced to 0% and 0.09% for Run 1 and Run 2, respectively. The fraction of multiple candidates in simulation is found to be zero since a requirement that a reconstructed candidate is fully associated with the signal decay suppresses multiple candidates.

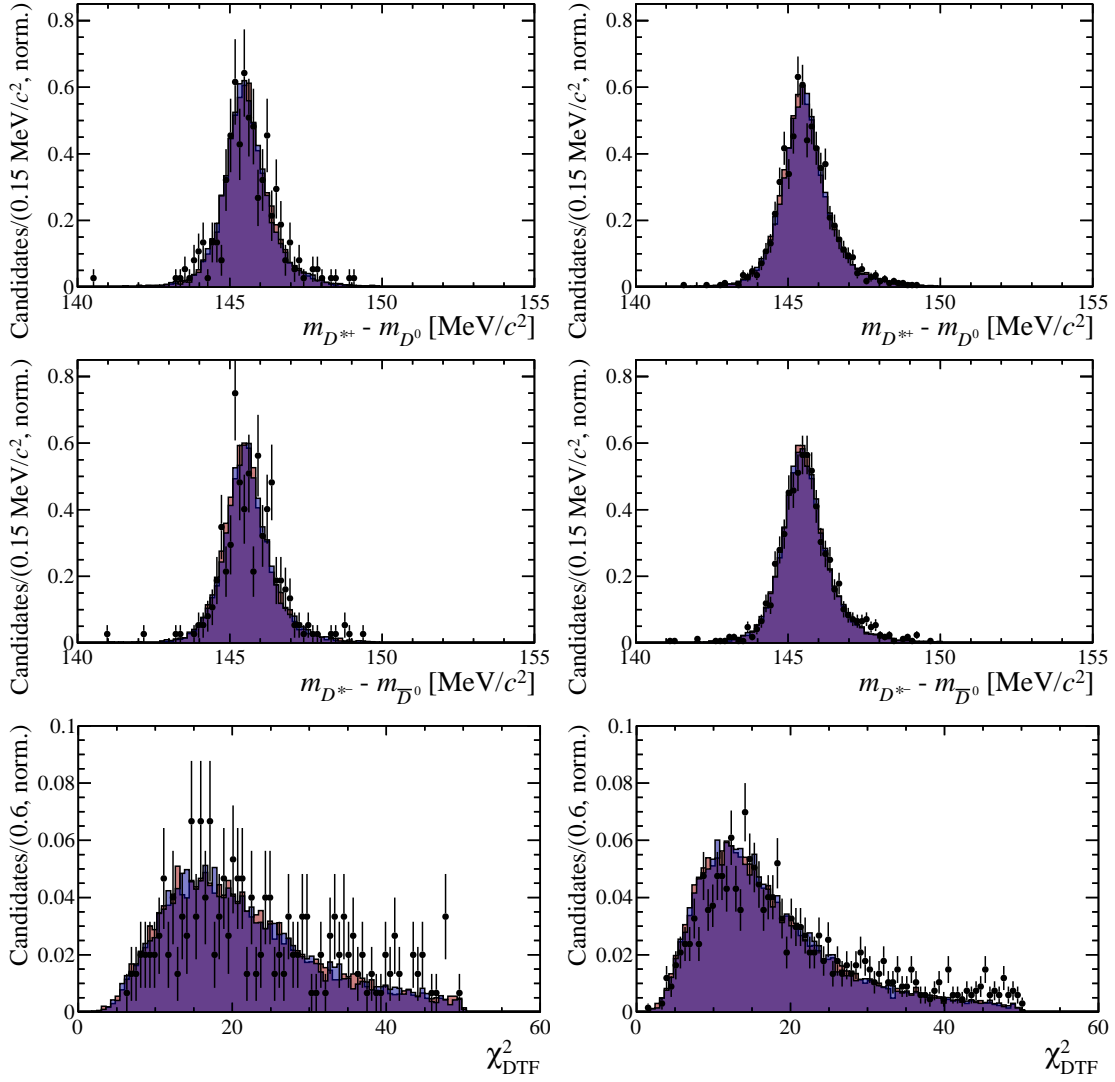


Figure 6.2 – Distributions of the differences of the invariant (top) D^{*+} and D^0 masses and of the (mid) D^{*-} and \bar{D}^0 masses and the (bottom) χ^2_{DTF} distribution for (left) Run 1 and (right) Run 2, respectively. The blue and red distributions show simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays, while the data points show $B^0 \rightarrow D^{*+}D^{*-}$ weighted data. The weights are obtained with the *sPlot* method [163].

Table 6.3 – Fraction of events without multiple candidates, with multiple candidates originating and not originating from misidentification for Run 1 and Run 2 in recorded data. The average number of candidates in these events is also listed.

Category	Fraction of events		Average number of candidates	
	Run 1 [%]	Run 2 [%]	Run 1	Run 2
Without mult. candidates	98.36	96.21	1.0	1.0
Mult. cand. without misID	0.2	1.31	2.0	2.0
Mult. cand. with misID	1.43	2.48	2.0	2.0

A different approach than the random rejection is used for events with multiple candidates originating from misID due to the high fraction of these events. In more detail, the candidates will be kept depending on the final state and the measured branching fractions of the D^0 decays. The fraction of the branching ratio of each multiple candidate in an event to the sum of the branching ratios of all multiple candidates in an event is calculated, *i.e.*

$$p_i = \frac{\mathcal{B}(i)}{\sum_{j=1}^N \mathcal{B}(j)}, \quad (6.5)$$

for the respective candidates $i \in [1, \dots, N]$, where $\mathcal{B}(i)$ corresponds to the measured branching fraction of the respective final state taken given in Table 6.1. The value of p_i is then interpreted as a probability, and the candidate to be kept is chosen according to that probability. This method is used since it is expected that most of the real decays are present in the $D^0 \rightarrow K^- \pi^+$ mode and the other D^0 modes are not as likely as this channel. If not the most likely candidate is selected, this would lead to a bias in the wrong direction. However, even if the not most likely candidate is selected, still a real signal decay is kept since all candidates in this category are real signal decays.

6.4 Physical background

After applying the whole selection, backgrounds can still be present in the data. Possible remaining backgrounds are discussed and studied in this section. For example, partially reconstructed backgrounds are not explicitly suppressed in this analysis but are further investigated. They appear at lower invariant masses with respect to the nominal B^0 mass. Contributing backgrounds could be $B^{0(+)} \rightarrow D^{*+} D^{*-} \pi^{0(+)}$ decays, including intermediate $D^* \pi^\pm$ resonances. Such resonances are present in $B^{0(+)} \rightarrow D_1(2420)^{+(0)} D^{*-(+)}$ or $B^{0(+)} \rightarrow D_1(2430)^{+(0)} D^{*-(+)}$ decays, where the $D_1(2420)$ and $D_1(2430)$ mesons decay into $D^{*+(-)} \pi^{0(+)}$. These backgrounds could mimic the signal, when the $\pi^{0(+)}$ from the

$D_1(2420)^{+(0)}$ and $D_1(2430)^{+(0)}$ decays or even directly from $B^{0(+)}$ decay is not reconstructed.

The decays are studied using the RapidSim software [172], which is a fast simulation tool for the study of phase-space decays of beauty and charm quark hadrons. Data sets with 100000 candidates are generated at a centre-of-mass energy of 13 TeV, which is equivalent to the centre-of-mass energy reached during Run 2. The efficiencies of the backgrounds to exceed a threshold of $5100 \text{ MeV}/c^2$ in the invariant $D^{*+}D^{*-}$ mass are given in Table 6.4. The invariant mass distributions of the individual generated data sets are shown in Fig. 6.3. None of the backgrounds has significant contributions above $5100 \text{ MeV}/c^2$, which is the lower mass threshold used later in the analysis to extract signal decays. Due to the smaller centre-of-mass energy in Run 1, it is assumed that the same holds for Run 1 data. Additionally, the number of candidates below $5100 \text{ MeV}/c^2$ in the recorded data sets is low, with only 67 and 374 candidates present in Run 1 and Run 2, respectively. Since some of these candidates are combinatorial background, the potential amount of partially reconstructed background is small. Extrapolating the exponential function resulting from the mass fit performed to extract the signal decays in Chap. 7 leads to 37 and 206 combinatorial background candidates in this region for Run 1 and Run 2, respectively. Because of the resulting small number of potential partially reconstructed background candidates in the region from $4800 \text{ MeV}/c^2$ to $5100 \text{ MeV}/c^2$ and the small efficiencies of the investigated backgrounds to exceed $5100 \text{ MeV}/c^2$, these partially reconstructed backgrounds are neglected in the further analysis.

Table 6.4 – Studied partially reconstructed backgrounds using data sets comprising 100000 events, which are generated with RapidSim [172]. The determined efficiencies of the generated decays to exceed $5100 \text{ MeV}/c^2$ in the invariant $D^{*+}D^{*-}$ mass are given. The backgrounds are studied for a centre-of-mass energy of 13 TeV. The reported efficiencies do not include any selection steps and only refer to the requirement exceeding the mass threshold.

Background	Efficiency [%]
$B^0 \rightarrow D^{*+}D^{*-}\pi^0$	0.825 ± 0.029
$B^+ \rightarrow D^{*+}D^{*-}\pi^+$	0.675 ± 0.026
$B^+ \rightarrow D_1(2420)^0(\rightarrow D^{*-}\pi^+)D^{*+}$	0.033 ± 0.006
$B^0 \rightarrow D_1(2420)^+(\rightarrow D^{*+}\pi^0)D^{*-}$	0.044 ± 0.007
$B^+ \rightarrow D_1(2430)^0(\rightarrow D^{*-}\pi^+)D^{*+}$	0.447 ± 0.021
$B^0 \rightarrow D_1(2430)^+(\rightarrow D^{*+}\pi^0)D^{*-}$	0.578 ± 0.024

Additionally, checks on backgrounds from $\Lambda_c^+ \rightarrow K^- p \pi^+$ and $D_s^\pm \rightarrow K^\mp K^\pm \pi^\pm$ are performed, which can occur due to the misidentification of one final-state particle. The checks are given in Fig. 6.4, where the combined invariant mass of all final-state particles originating from the D^{*+} meson is shown and where either the kaon or proton mass hypothesis has been applied to one of the pions in the $D^{*+} \rightarrow D^0(\rightarrow K^- \pi^+) \pi^+$ decays. No peaking structures around the D_s^+ mass of $1968.35 \text{ MeV}/c^2$ [47] or the Λ_c^+

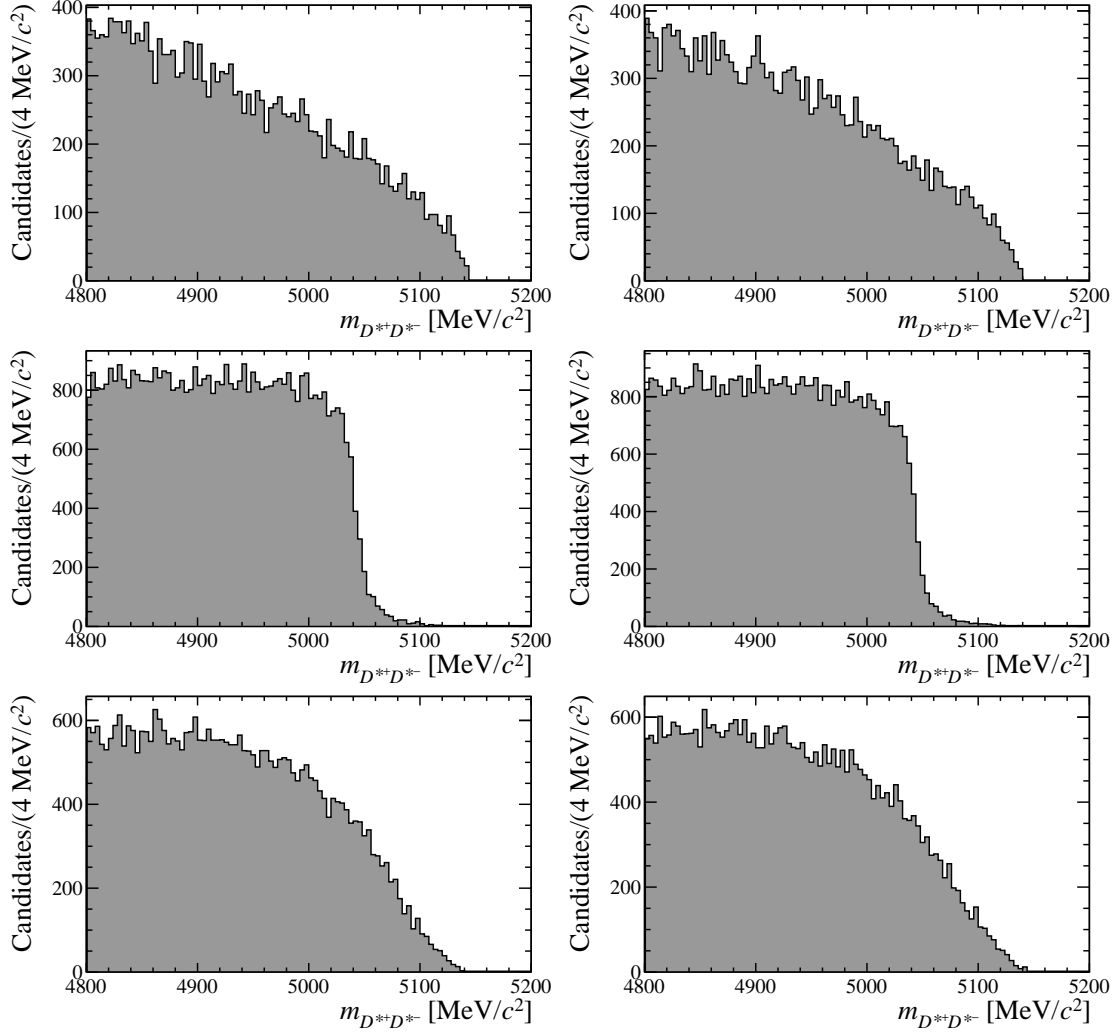


Figure 6.3 – Invariant $D^{*+}D^{*-}$ mass distribution of possible partial reconstructed background contributions. Simulated candidates are shown, which are generated with the RapidSim [172] package at a centre-of-mass energy of 13 TeV. The study includes backgrounds from (top left) $B^0 \rightarrow D^{*+}D^{*-}\pi^0$, (top right) $B^+ \rightarrow D^{*+}D^{*-}\pi^+$, (mid left) $B^+ \rightarrow D_1(2420)^0(\rightarrow D^{*-}\pi^+)D^{*+}$, (mid right) $B^0 \rightarrow D_1(2420)^+(\rightarrow D^{*+}\pi^0)D^{*-}$, (bottom left) $B^+ \rightarrow D_1(2430)^0(\rightarrow D^{*-}\pi^+)D^{*+}$ and (bottom right) $B^0 \rightarrow D_1(2430)^+(\rightarrow D^{*+}\pi^0)D^{*-}$ decays. Only the mass range from 4800 MeV/c^2 to 5200 MeV/c^2 is shown.

mass of $2286.46 \text{ MeV}/c^2$ [47] mass. Therefore, no signs of significant misidentification backgrounds are observed. No signs of misidentification backgrounds are present in the distributions for the D^{*-} meson either, which are not presented in this thesis since they show the same trends.

Further studies are performed to check for backgrounds from $B^0 \rightarrow D^{**}D$ decays, where the invariant $D^{*\pm}\pi^\mp$ mass is studied. These studies are given in Fig. 6.5 and Fig. 6.6. The distribution is analysed for the recorded data and polarised simulation, *i.e.* transversely and longitudinally polarised $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays. For simulation, only the Run 2 $m_{D^{*+}\pi^-}$ distributions are shown since the Run 1 and the $m_{D^{*-}\pi^+}$ distributions look similar. Assuming a larger contribution for longitudinally polarised decays as it was measured by the Belle experiment [92], the double peak structure of the data distribution can be explained with distributions in simulation and no clear sign of $B^0 \rightarrow D^{**}D$ decays is present.

No significant contribution of single-charm background, *i.e.* background from decays with no intermediate $D^{*\pm}$ mesons, is observed. To check if such background remains after the suppression of the combinatorial background, Fig. 6.7 shows the invariant $D^{*+}D^{*-}$ mass with the additional conditions $|m_{D^{*\pm}}^{\text{PDG}} - m_{D^{*\pm}}| > 30 \text{ MeV}/c^2$. The efficiency of these requirements is estimated with simulation. Using this efficiency, the number of expected $B^0 \rightarrow D^{*+}D^{*-}$ candidates and $B_s^0 \rightarrow D^{*+}D^{*-}$ candidates is calculated to be 3 (22) and 0.2 (2) for each requirement for Run 1 (Run 2), respectively. Combining these requirements with a decay length requirement of $l_{D^{*\pm}} < 0.4 \text{ mm}$ for the respective $D^{*\pm}$ meson, the number of expected candidates is determined to be 0.3 (2) and 0.02 (0.1) for each requirement for Run 1 (Run 2), respectively. This is in agreement with the number of candidates seen around the B_s^0 and B^0 masses in Fig. 6.7. Additionally, the invariant mass of each final-state particle combined with the slow pion is checked since no flight distance cuts are used in this analysis. No peaking distributions are found in these distributions.

All studied backgrounds do not contribute or only contribute marginally and are not considered in the further analysis. The cross-checks are performed for the $B_s^0 \rightarrow D^{*+}D^{*-}$ selection only. Although the χ_{DTF}^2 requirement is discarded in the $B^0 \rightarrow D^{*+}D^{*-}$ analysis, it is assumed that the same applies for the selection of this analysis. It is concluded that the only non-negligible background in the data is the combinatorial background.

6.5 Selection efficiencies

After the development of the selection, the discriminative variable to extract the signal can be compared for the individual selection steps. This variable is the invariant $D^{*+}D^{*-}$ mass. The distributions after the preselection and after the whole selection are presented in Fig. 6.8 for the Run 1 and Run 2 data. As the data contains both the $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ decays, the comparison proves that the requirements used in this thesis allow the selection of both decays. The distributions of only the selected candidates are

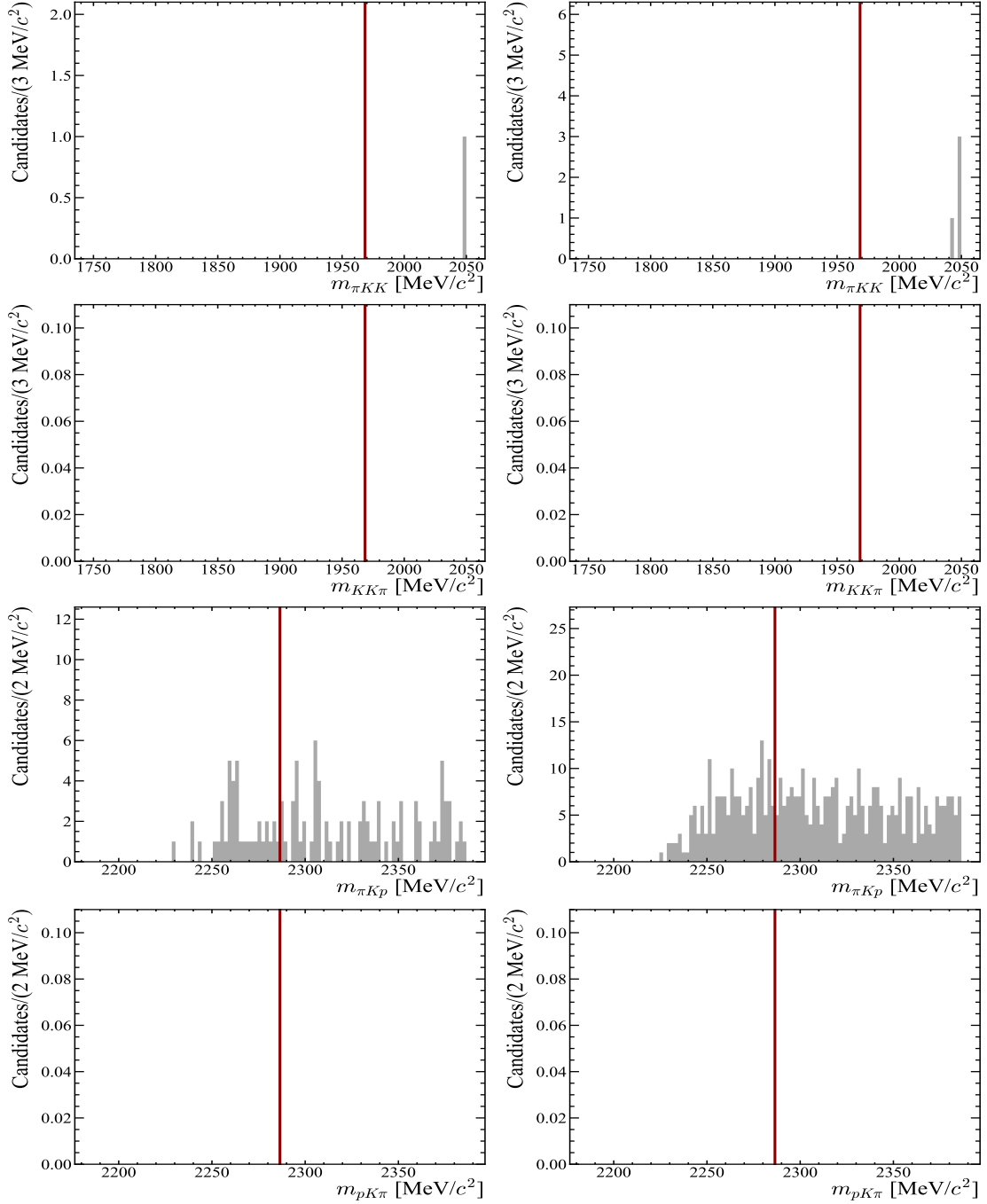


Figure 6.4 – Invariant $m_{h_0 h_1 h_2}$ mass distribution of the three final-state particles of the D^{*+} meson for the $D^0 \rightarrow K^- \pi^+$ final state for (left) Run 1 and (right) Run 2, where h_0 is always the slow pion. The (1st and 2nd row) kaon and (3rd and 4th row) proton mass hypothesis is applied to one of the pions to check for backgrounds from $D_s^\pm \rightarrow K^\mp K^\pm \pi^\pm$ and $\Lambda_c^+ \rightarrow K^- p \pi^+$ decays, respectively. The vertical red line indicates the respective D_s^+ and Λ_c^+ masses taken from Ref. [173].

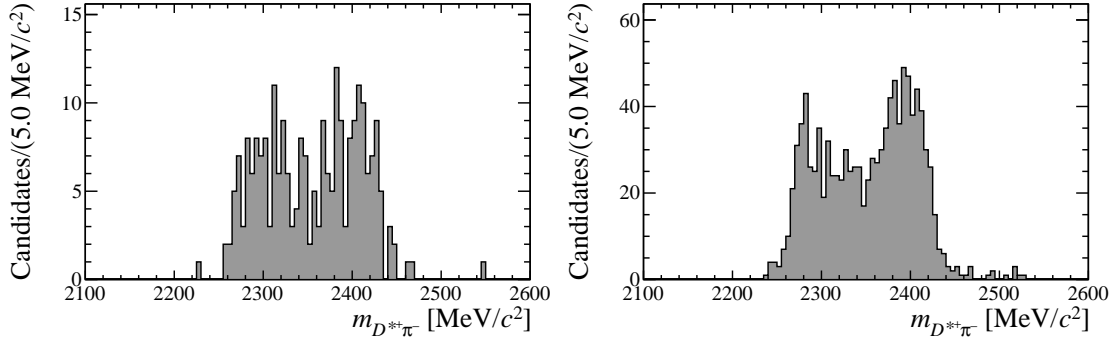


Figure 6.5 – Invariant $D^{*+}\pi^-$ mass distributions for the recorded data after the rejection of multiple candidates weighted according to the $B^0 \rightarrow D^{*+}D^{*-}$ signal for (left) Run 1 and (right) Run 2. The weights are obtained with the *sPlot* method [163]. Only events in the $D^{*+}D^{*-}$ mass range from $5100 \text{ MeV}/c^2$ to $6000 \text{ MeV}/c^2$ are shown. The invariant $D^{*+}\pi^-$ mass distributions are not explicitly presented, but they show the same structure for Run 1 and Run 2 weighted data.

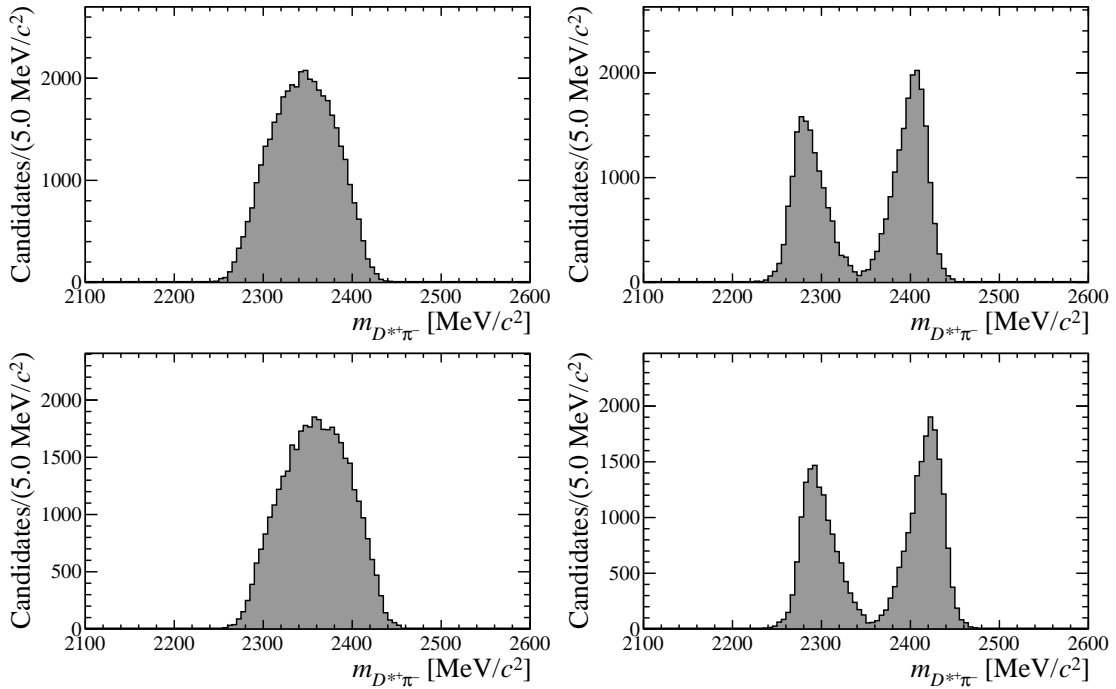


Figure 6.6 – Invariant $D^{*+}\pi^-$ mass distributions for the (left) transversely and (right) longitudinally polarised (top) $B^0 \rightarrow D^{*+}D^{*-}$ and (bottom) $B_s^0 \rightarrow D^{*+}D^{*-}$ simulation for Run 2, respectively. The Run 1 and invariant $D^{*+}\pi^-$ distributions show the same structures and are not shown explicitly.

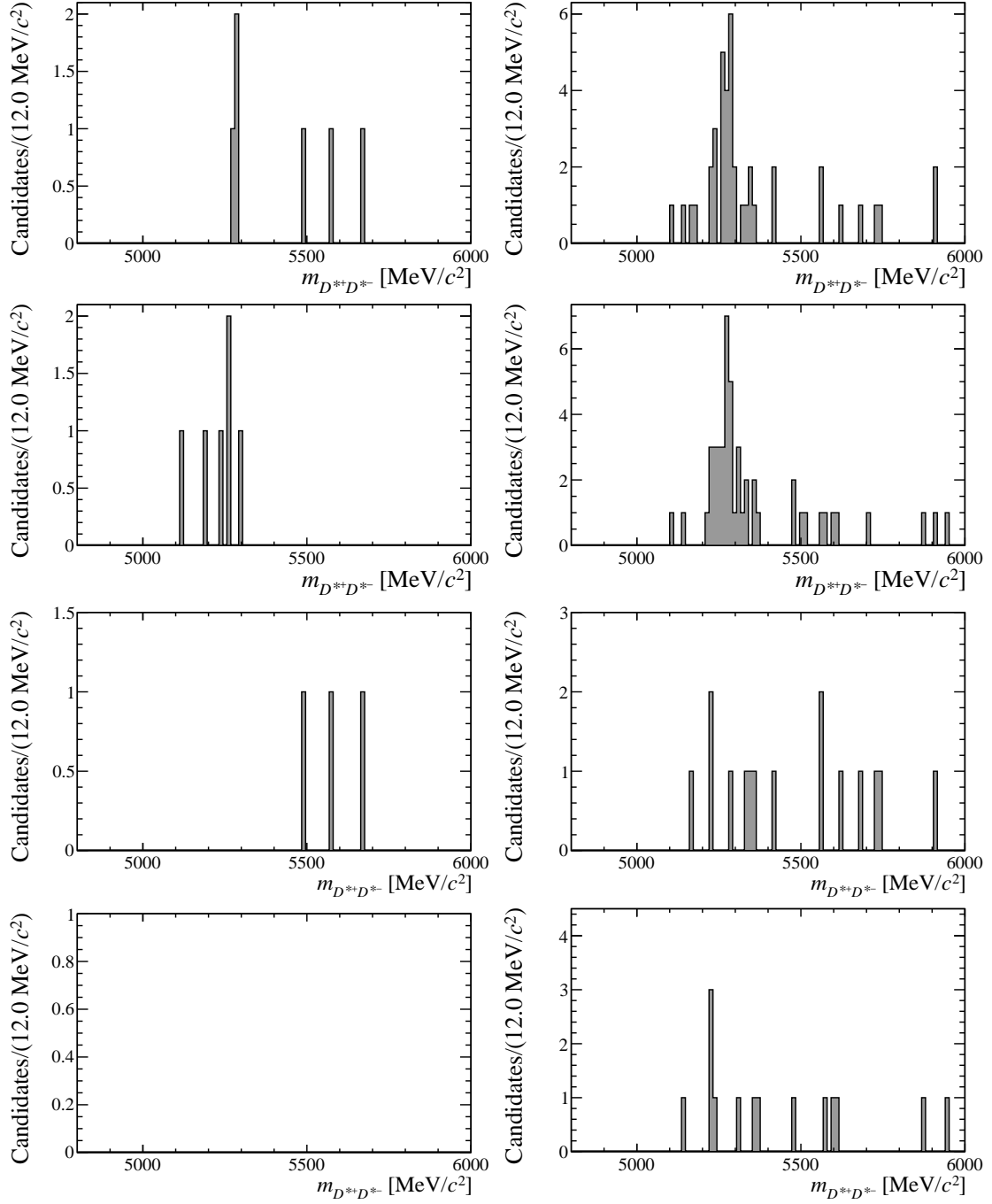


Figure 6.7 – Invariant $D^{*+}D^{*-}$ mass distributions after the application of the selection for (left) Run 1 and (right) Run 2. The additional requirements (1st row) $|m_{D^{*+}}^{\text{PDG}} - m_{D^{*+}}| > 30 \text{ MeV}/c^2$, (2nd row) $|m_{D^{*-}}^{\text{PDG}} - m_{D^{*-}}| > 30 \text{ MeV}/c^2$, (3rd row) $|m_{D^{*+}}^{\text{PDG}} - m_{D^{*+}}| > 30 \text{ MeV}/c^2$ and $l_{D^{*+}} < 0.4 \text{ mm}$ and (4th row) $|m_{D^{*-}}^{\text{PDG}} - m_{D^{*-}}| > 30 \text{ MeV}/c^2$ and $l_{D^{*-}} < 0.4 \text{ mm}$ are applied to the data sets, where l describes the decay length of the respective particles.

also shown separately. Fig. 6.8 shows a clear narrow peak around the B^0 mass after the whole selection and a clear increase in the signal-to-background ratio for the selection step of suppressing combinatorial background. Accordingly, a strong suppression of this background is achieved corresponding to a rejection of 99.86% and 99.73% for Run 1 and Run 2, respectively. The upper mass sideband with $m_{D^{*+}D^{*-}} > 5400 \text{ MeV}/c^2$ of data is used to calculate these efficiencies. It is also visible in Fig. 6.8 that no other peaking backgrounds are present in the shown invariant mass range.

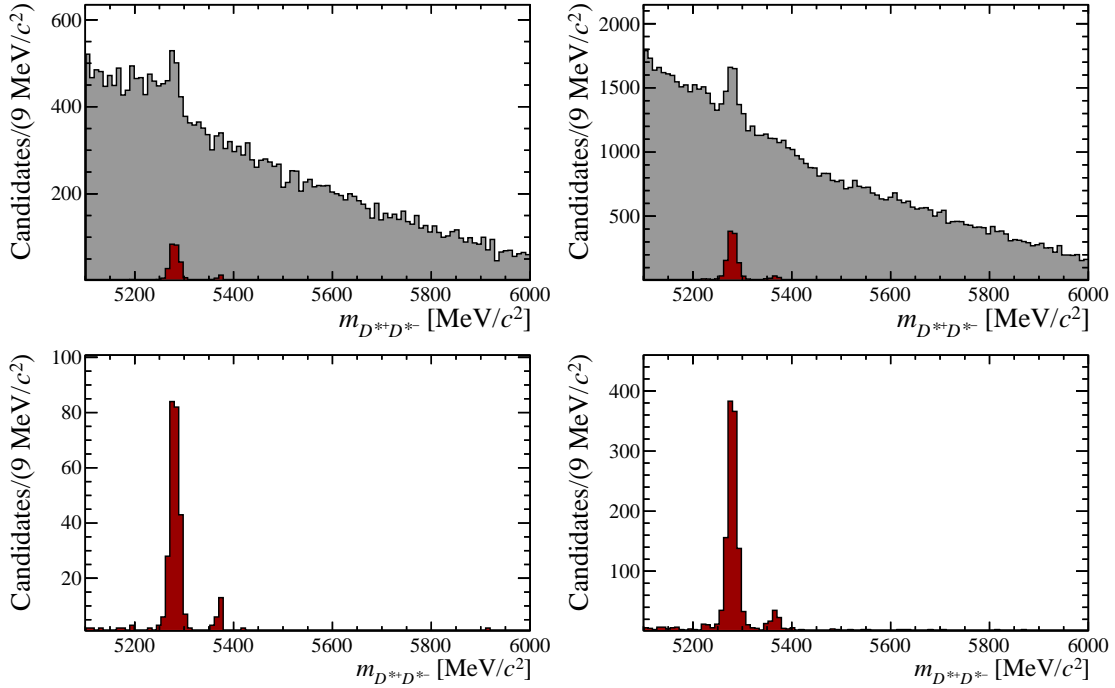


Figure 6.8 – Invariant $D^{*+}D^{*-}$ mass distribution for (top) the comparison before and after the suppression of combinatorial background and for (bottom) the remaining candidates after the whole selection. The comparison and the distribution of the selected candidates are both shown for (left) Run 1 and (right) Run 2 in the mass range from 5100 MeV/c^2 to 6000 MeV/c^2 . The selected candidates are presented in red, and the candidates before the suppression of the combinatorial background are shown in grey.

The selection efficiencies are a fundamental part of the calculation of the ratio of branching fractions. They are evaluated with simulated $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ decays separately for Run 1 and Run 2, where only true signal decays are considered. To take into account the different amounts of data between the years, the mean values for Run 1 and Run 2 are calculated by weighting the efficiencies of the individual years with the respective integrated luminosity. Each selection step is applied to the simulation in the same way as for the recorded data. The efficiencies of each step and the total efficiencies are summarised in Table 6.5 for the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis. The latter are calculated as the product of the individual efficiencies. The efficiencies for the

$B^0 \rightarrow D^{*+}D^{*-}$ analysis are similar and are not given since the CP violation measurement does not directly depend on the efficiencies. Besides the efficiencies for the preselection and further selection steps, acceptance effects of the detector are taken into account. This is because only decays that decay completely within the detector acceptance can be reconstructed and considered in the analyses. This efficiency is called geometric acceptance and is determined as the ratio of generated decays by the number of particles that decay within the acceptance. In more detail, it is required that all decay products of the decay chain must lie within the acceptance of the detector. The acceptance of the LHCb detector in forward direction is roughly given by 10 mrad to 300 mrad. However, in simulation, all events with decay products with a polar angle smaller than 400 mrad are accepted. This allows the particles to be recovered by the bending of the magnet. The Stripping efficiencies also include the step of aligning the Stripping selection between the two runs. No efficiency is listed for the rejection of multiple candidates since the fraction of multiple candidates in simulation is zero and therefore the selection efficiency is 100%. The $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays show very similar efficiencies. The main difference occurs for the preselection efficiencies, where the $B_s^0 \rightarrow D^{*+}D^{*-}$ decays have a slightly higher efficiency due to higher transverse momenta of the daughter particles.

Table 6.5 – Selection efficiencies of the $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ decays for Run 1 and Run 2 for selection of the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis. The efficiencies are given for the geometric acceptance, the preselection, the DTF requirement and the combinatorial background suppression (Comb. bkg.). They are calculated using simulated decays, where only the $D^0 \rightarrow K^- \pi^+$ decay mode is considered.

Step	$B^0 \rightarrow D^{*+}D^{*-}$		$B_s^0 \rightarrow D^{*+}D^{*-}$	
	$\epsilon^{\text{Run 1}}[\%]$	$\epsilon^{\text{Run 2}}[\%]$	$\epsilon^{\text{Run 1}}[\%]$	$\epsilon^{\text{Run 2}}[\%]$
Geometric	15.289 ± 0.020	16.285 ± 0.020	15.243 ± 0.020	16.233 ± 0.020
Preselection	0.3397 ± 0.0033	0.4725 ± 0.0034	0.3529 ± 0.0034	0.5087 ± 0.0036
DTF	97.71 ± 0.14	97.75 ± 0.11	97.25 ± 0.16	97.60 ± 0.11
Comb. bkg.	84.78 ± 0.35	87.61 ± 0.24	85.29 ± 0.34	88.42 ± 0.23
Total	0.0430 ± 0.0005	0.0659 ± 0.0005	0.0446 ± 0.0005	0.0713 ± 0.0005

In the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis, the simulation only comprises the $D^0 \rightarrow K^- \pi^+$ mode. The only requirements applied on the final-state particles are loose requirements in the Stripping. Therefore, the efficiencies are assumed to be similar for the other D^0 decay modes. Additionally, small differences because of different simulation versions and configurations used between the individual years and the effects of different detector performances are neglected. Both decays use the same configuration for every year, so differences between the years are cancelled out by calculating the ratio of the efficiencies.

7 Observation of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay

The measurement of the ratio of branching fractions is presented in this chapter. It requires knowledge about the selection efficiencies ϵ , the ratio of hadronisation fractions f_s/f_d and the number of selected $B_s^0 \rightarrow D^{*+} D^{*-}$ and $B^0 \rightarrow D^{*+} D^{*-}$ candidates \mathcal{N} . As derived in Sec. 2.5.1, the ratio is defined as

$$\frac{\mathcal{B}(B_s^0 \rightarrow D^{*+} D^{*-})}{\mathcal{B}(B^0 \rightarrow D^{*+} D^{*-})} = \frac{\mathcal{N}_{B_s^0 \rightarrow D^{*+} D^{*-}}}{\mathcal{N}_{B^0 \rightarrow D^{*+} D^{*-}}} \cdot \frac{f_d}{f_s} \cdot \frac{\epsilon_{B^0 \rightarrow D^{*+} D^{*-}}}{\epsilon_{B_s^0 \rightarrow D^{*+} D^{*-}}}. \quad (7.1)$$

The ratio of hadronisation fractions depends on the centre-of-mass energy and b -hadron momentum and is an external input in this analysis. Several analyses have been performed to measure the ratio at the various collision energies of Run 1 and Run 2. For example, it has been measured in semileptonic $B \rightarrow D\mu X$ decays at 7 and 13 TeV [174, 175]. Additionally, measurements have been performed at centre-of-mass energies of 7, 8 and 13 TeV in hadronic $B \rightarrow Dh$ [176, 177] and $B \rightarrow J/\psi X$ decays [178]. However, this analysis uses the latest results obtained by a phase-space integrated combination [179], which allowed to reduce the uncertainties compared to the individual measurements. The values obtained with the combination are

$$\begin{aligned} \frac{f_s}{f_d}(7 \text{ TeV}) &= 0.2390 \pm 0.0076, \\ \frac{f_s}{f_d}(8 \text{ TeV}) &= 0.2385 \pm 0.0075, \\ \frac{f_s}{f_d}(13 \text{ TeV}) &= 0.2539 \pm 0.0079. \end{aligned} \quad (7.2)$$

The combination in Ref. [179] also provides a measurement of the 13 TeV Run 2 value relative to a combined Run 1 result. This allows correlations between these values to be considered in the Run 1 and Run 2 combination of the ratio of branching fractions. The relative value is given by

$$\frac{f_s/f_d(\text{Run 2})}{f_s/f_d(\text{Run 1})} = 1.064 \pm 0.007. \quad (7.3)$$

While the ratio of hadronisation fractions is known and the selection efficiencies are calculated with simulation in Sec. 6.5, the candidate's ratio remains to be determined.

The extraction of the numbers of signal decays is described in Sec. 7.1. Multiple sources of systematic uncertainties on the ratio of branching fractions are studied in Sec. 7.2. The ratio itself is calculated in Sec. 7.3. The measurement is performed separately for Run 1 and Run 2, and the individual results are combined in the end. A possible effect of the short lifetime of the B_s^0 meson on the ratio of branching fractions is discussed in Sec. 7.4. Lastly, the results of this analysis are summarised in Sec. 7.5.

7.1 Signal extraction

The invariant $D^{*+}D^{*-}$ mass distribution is used to discriminate between signal and background. Extended maximum-likelihood fits are performed to extract the number of $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays, which is described in this section. The mass estimate from a DTF with constraints placed on the PV and the masses of the intermediate D mesons is used. This allows the B^0 mass resolution to be increased by around 50%. The model used to describe the invariant mass is discussed in Sec. 7.1.1. Fits to the distribution of simulated decays to validate the model are described in Sec. 7.1.2. Finally, the number of signal decays in data is determined in Sec. 7.1.3. All fits are performed to data sets where the whole selection described in Chap. 6 is applied.

7.1.1 Fit model

The model used to describe the invariant $D^{*+}D^{*-}$ mass distribution of the signal decay is a double-sided Hypatia function [180]. This function consists of the sum of two Crystal Ball (CB) functions [181] with opposite tails, *i.e.* one tail for lower and one for higher masses. Additionally, the central Gaussian kernel of the CB functions is replaced by a generalised hyperbolic distribution $G(m - \mu, \sigma, \lambda, \zeta, \beta)$ defined as

$$G(m - \mu, \sigma, \lambda, \zeta, \beta) = ((m - \mu)^2 + \sigma^2 A_\lambda^2(\zeta))^{\frac{\lambda}{2} - \frac{1}{4}} e^{\beta(m - \mu)} K_{\lambda - \frac{1}{2}} \left(\zeta \sqrt{1 + \frac{(m - \mu)^2}{\sigma^2 A_\lambda^2(\zeta)}} \right), \quad (7.4)$$

where the parameter $A_\lambda(\zeta)$ is given by

$$A_\lambda(\zeta) = \sqrt{\zeta \frac{K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}}. \quad (7.5)$$

The functions $K_\lambda(\zeta)$ are cylindrical Bessel functions of the second kind. The complete Hypatia function is defined as

$$H(m, \mu, \sigma, \lambda, \zeta, \beta, \alpha_1, \alpha_2, n_1, n_2) = \begin{cases} A(B - (m - \mu))^{-n_1} & , \text{if } \frac{m - \mu}{\sigma} < -\alpha_1 \\ C(D + (m - \mu))^{-n_2} & , \text{if } \frac{m - \mu}{\sigma} > \alpha_2 \\ G(m - \mu, \sigma, \lambda, \zeta, \beta) & , \text{otherwise} \end{cases}, \quad (7.6)$$

where the parameters A and B describe the power-law part of the CB functions. They are given by

$$A = \left(\frac{n_1}{|\alpha_1|} \right) \cdot \exp\left(-\frac{|\alpha_1|^2}{2}\right), \quad B = \frac{n_1}{|\alpha_1|} - |\alpha_1|. \quad (7.7)$$

The parameters C and D are defined identically but with the parameters α_2 and n_2 , which can differ from α_1 and n_1 . The parameter α describes the position at which the generalised hyperbolic distribution transitions into the power-law part of the CB function. Since this model has opposite tails, it is perfectly suited to describe asymmetric distributions, such as the invariant mass distribution of the B candidates. These distributions can become asymmetric by bremsstrahlung effects or by effects of the detector resolution.

The selection described in Chap. 6 allows the selection of the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays from the same data set. As discussed in Sec. 6.5, no peaking backgrounds or other physical backgrounds that would have to be further modelled are present in the invariant $D^{*+}D^{*-}$ mass distribution after the selection. The only remaining background is the combinatorial background, which cannot be removed completely. Besides the $B_s^0 \rightarrow D^{*+}D^{*-}$ signal, both additional contributions, *i.e.* the $B^0 \rightarrow D^{*+}D^{*-}$ signal and the combinatorial background, have to be described in the fit to the invariant $D^{*+}D^{*-}$ mass. Consequently, the PDF \mathcal{P} consists of three components \mathcal{P}_j : the $B_s^0 \rightarrow D^{*+}D^{*-}$ signal with $j = B_s^0$, the $B^0 \rightarrow D^{*+}D^{*-}$ signal with $j = B^0$ and the combinatorial background with $j = \text{Comb}$. The total PDF is defined as the sum of these three components and given by

$$\mathcal{N}\mathcal{P} = \mathcal{N}_{B^0}\mathcal{P}_{B^0} + \mathcal{N}_{B_s^0}\mathcal{P}_{B_s^0} + \mathcal{N}_{\text{Comb}}\mathcal{P}_{\text{Comb}}, \quad (7.8)$$

where \mathcal{N}_j is the yield of the respective component. The individual PDFs are described by the following models:

$B^0 \rightarrow D^{*+}D^{*-}$ signal: The distribution of the reconstructed invariant mass for the $B^0 \rightarrow D^{*+}D^{*-}$ signal component is parametrised by a double-sided Hypatia function.

$B_s^0 \rightarrow D^{*+}D^{*-}$ signal: The $B_s^0 \rightarrow D^{*+}D^{*-}$ signal component is parametrised by a double-sided Hypatia function, where the mean value μ of the $B_s^0 \rightarrow D^{*+}D^{*-}$ signal PDF is shifted relative to the mean value of the $B^0 \rightarrow D^{*+}D^{*-}$ signal. The shift is fixed to the known mass difference between the B_s^0 and B^0 mesons, which is given by $\Delta\mu_{B^0} = 87.26 \text{ MeV}/c^2$, calculated with the measured masses taken from Ref. [182].

Combinatorial background: The reconstructed invariant mass distribution for the combinatorial background component is modelled by an exponential function, where the slope is described by the parameter λ_{exp} .

7.1.2 Fit to simulation

Maximum-likelihood fits to the invariant $D^{*+}D^{*-}$ mass distribution of $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ simulation are performed separately. They are performed to validate the model used to describe the invariant mass and to determine the shape parameters of the respective distributions. The model used in the fits is discussed in more detail in Sec. 7.1.1. The fits and their results are described in this section.

In the fits to simulation, the invariant $D^{*+}D^{*-}$ mass is restricted to be in the range of $5200 \text{ MeV}/c^2$ to $5500 \text{ MeV}/c^2$. The fits are performed for the Run 1 and Run 2 data sets individually. The parameters ζ and β are set to zero in the fits to achieve a more stable fit since not enough statistics are available to let them float. All remaining parameters are floating in the fit. The results for the fits to the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ simulation are listed in Table 7.1. The invariant mass distributions overlaid with the projection of the results are given in Fig. 7.1. It is concluded that the double-sided Hypatia function describes the simulation accurately for both runs and the invariant mass distributions of both decays. As a cross-check, the fit quality is validated by calculating the respective χ^2/ndf values, where ndf is the number of degrees of freedom, *i.e.* $\text{ndf} = 7$. The values are determined within the RooFIT package [160] and listed in Table 7.1. The values show a good agreement between the simulation and the fit. Therefore, the model and the results of the simulation fits are used for the individual fits to Run 1 and Run 2 recorded data.

Table 7.1 – Results of the fits to the invariant $D^{*+}D^{*-}$ mass distribution of simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays for Run 1 and Run 2. The parameters ζ and β are fixed to zero in the fits to simulation to achieve a more stable fit. The resulting χ^2/ndf values are also listed.

Parameter	$B_s^0 \rightarrow D^{*+}D^{*-}$		$B^0 \rightarrow D^{*+}D^{*-}$	
	Run 1	Run 2	Run 1	Run 2
μ [MeV/ c^2]	5367.17 ± 0.11	5367.20 ± 0.08	5279.90 ± 0.11	5279.96 ± 0.09
σ [MeV/ c^2]	11.31 ± 0.23	11.8 ± 0.4	9.9 ± 0.4	10.85 ± 0.31
λ_{Hyp}	-2.58 ± 0.23	-2.47 ± 0.27	-3.9 ± 1.1	-2.78 ± 0.34
α_1	3.0 ± 0.5	1.90 ± 0.16	1.54 ± 0.13	1.48 ± 0.13
α_2	3.5 ± 0.4	2.32 ± 0.33	2.13 ± 0.26	2.21 ± 0.22
n_1	1.7 ± 0.6	2.69 ± 0.29	6.6 ± 1.5	5.1 ± 0.9
n_2	3.6 ± 1.3	3.8 ± 0.7	3.6 ± 0.6	4.1 ± 0.6
β	0	0	0	0
ζ	0	0	0	0
χ^2/ndf	0.73	0.69	0.7	0.81

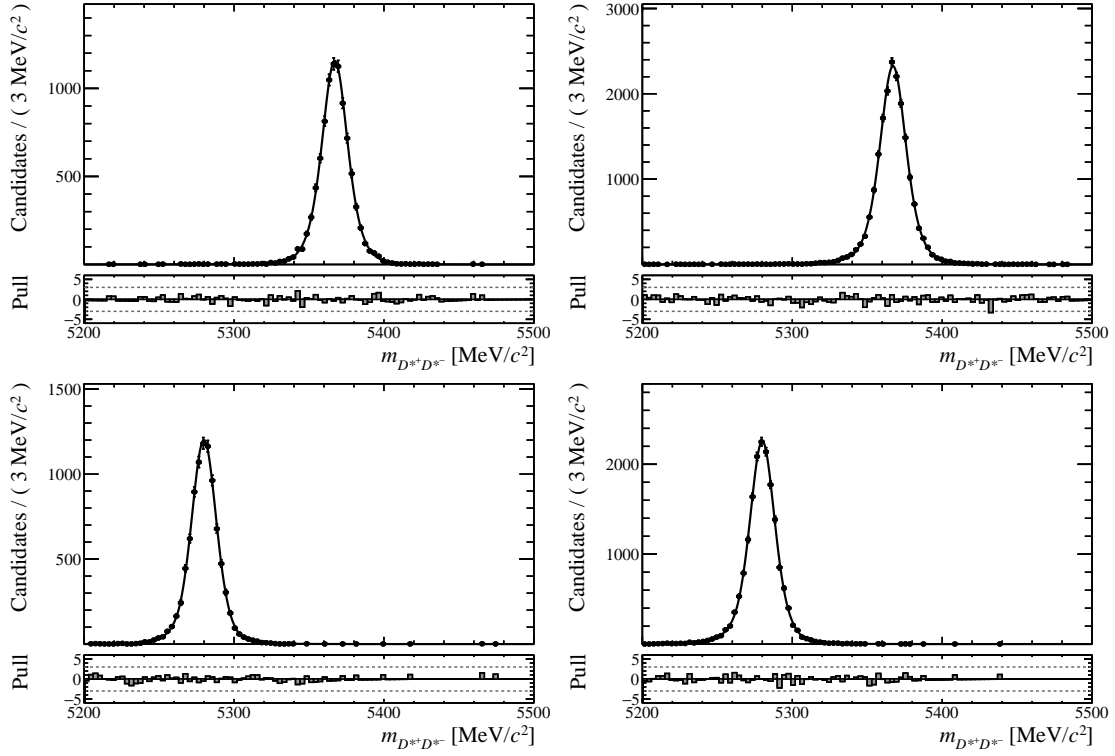


Figure 7.1 – Reconstructed invariant $D^{*+}D^{*-}$ mass of simulated (top) $B_s^0 \rightarrow D^{*+}D^{*-}$ and (bottom) $B^0 \rightarrow D^{*+}D^{*-}$ decays shown together with the fitted PDF and pull distributions for (left) Run 1 and (right) Run 2, respectively.

7.1.3 Fit to data

The extended maximum-likelihood fits to the invariant $D^{*+}D^{*-}$ mass distribution of recorded data are described in this section. As in Sec. 7.1.2 for simulation, separate fits are performed for Run 1 and Run 2 data. The invariant $D^{*+}D^{*-}$ mass is restricted to be in the range of $5100 \text{ MeV}/c^2$ to $6000 \text{ MeV}/c^2$. The choice of this mass range reduces partially reconstructed backgrounds to a negligible level. Such backgrounds appear at lower invariant masses with respect to the known B^0 mass. Potential sources of partially reconstructed backgrounds are $B^{0(+)} \rightarrow D^{*+}D^{*-}\pi^{0(+)}$ decays, which may involve intermediate $D_1(2420)^{0(+)} \rightarrow D^{*+}\pi^{-(0)}$ or $D_1(2430)^{0(+)} \rightarrow D^{*+}\pi^{-(0)}$ decays. These and further backgrounds due to misidentification are discussed in Sec. 6.4, but no significant contributions are found in data.

The fit to data uses the parametrisation of the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ signal contributions determined with the fits to simulation. However, the recorded data set comprises all considered D^0 decay modes, while only the $D^0 \rightarrow K^-\pi^+$ decay mode is available for the simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays. Accordingly, fixing the shape parameters assumes that the resolution is similar for all considered D^0 decay modes. As a cross-check, the invariant $D^{*+}D^{*-}$ mass distribution for the various

final states in data is compared and found to be compatible. All parameters listed in Table 7.1 are used and fixed in the fit except for the mean and the widths. The remaining parameters of the model, *i.e.* the yields of the three contributions and the slope of the exponential function, are floating. The results of the fits to Run 1 and Run 2 data are listed in Table 7.2, where also the fixed parameters are given. These can be identified since they are listed without uncertainties. Fig. 7.2 shows the invariant $D^{*+}D^{*-}$ mass distribution and the fit projection for Run 1 and Run 2. The individual contributions of the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ signal and the combinatorial background are also shown. The noticeably smaller width of the $B_s^0 \rightarrow D^{*+}D^{*-}$ distribution compared to the $B^0 \rightarrow D^{*+}D^{*-}$ distribution in the Run 1 fit is considered a statistical fluctuation. A further study is performed to cross-check this assumption, where the B_s^0 mean is allowed to float in the fit to data. The study shows that a floating B_s^0 mean does not resolve the discrepancy or change the number of decays. To ensure that the B^0 contribution does not affect the B_s^0 contribution, another study is performed in which the right tail of the B^0 distribution is removed. In detail, the $B^0 \rightarrow D^{*+}D^{*-}$ contribution is described with a CB function with a tail towards lower masses instead of a double-sided Hypatia function. This study also shows no significant change in the number of signal decays.

Figure 7.2 – Reconstructed mass of the $B \rightarrow D^{*+}D^{*-}$ data sets shown together with the fitted PDF for (left) Run 1 and (right) Run 2 with logarithmic scale [40].

Overall, the fits to recorded data after the application of the whole selection allow

$$\mathcal{N}_{B_s^0, \text{Run 1}} = 20 \pm 5 \quad \text{and} \quad \mathcal{N}_{B_s^0, \text{Run 2}} = 79 \pm 10 \quad (7.9)$$

$B_s^0 \rightarrow D^{*+}D^{*-}$ candidates and

$$\mathcal{N}_{B^0, \text{Run 1}} = 251 \pm 16 \quad \text{and} \quad \mathcal{N}_{B^0, \text{Run 2}} = 1123 \pm 34 \quad (7.10)$$

$B^0 \rightarrow D^{*+}D^{*-}$ candidates to be extracted for Run 1 and Run 2, respectively. Only a small fraction of combinatorial background is found. The number of combinatorial background candidates is given by

$$\mathcal{N}_{\text{Comb, Run 1}} = 44 \pm 7 \quad \text{and} \quad \mathcal{N}_{\text{Comb, Run 2}} = 228 \pm 18 \quad (7.11)$$

Table 7.2 – Parameters resulting from the fit to recorded data. The results are separated into Run 1 and Run 2. The parameters fixed to values from the fit to simulation are listed without uncertainty. In addition, the constant shift $\Delta\mu_{B^0}$ of the B_s^0 mean relative to the B^0 mean is given.

Parameter	Run 1	Run 2
$\mathcal{N}_{B_s^0}$	20 ± 5	79 ± 10
\mathcal{N}_{B^0}	251 ± 16	1123 ± 34
$\mathcal{N}_{\text{Comb}}$	44 ± 7	228 ± 18
μ_{B^0} [MeV/c ²]	5281.1 ± 0.6	5279.55 ± 0.29
λ_{exp} [c ² /MeV]	-0.0027 ± 0.0007	-0.00232 ± 0.00029
σ^{B^0} [MeV/c ²]	9.8 ± 0.6	10.32 ± 0.31
$\sigma^{B_s^0}$ [MeV/c ²]	7.5 ± 1.5	10.4 ± 1.7
$\Delta\mu_{B^0}$ [MeV/c ²]	87.26	87.26
$\lambda_{\text{Hyp}}^{B_s^0}$	-3.9	-2.78
$\lambda_{\text{Hyp}}^{B^0}$	-2.58	-2.47
$\alpha_1^{B^0}$	1.54	1.48
$\alpha_1^{B_s^0}$	3.0	1.90
$\alpha_2^{B^0}$	2.13	2.21
$\alpha_2^{B_s^0}$	3.5	2.32
$n_1^{B^0}$	6.6	5.1
$n_1^{B_s^0}$	1.7	2.69
$n_2^{B^0}$	3.6	4.1
$n_2^{B_s^0}$	3.6	3.8
β	0	0
ζ	0	0

for Run 1 and Run 2, respectively. The fits show a clear sign of a $B_s^0 \rightarrow D^{*+}D^{*-}$ contribution. With a simultaneous extended maximum-likelihood fit to the invariant $D^{*+}D^{*-}$ mass distributions of Run 1 and Run 2 data, the significance of the $B_s^0 \rightarrow D^{*+}D^{*-}$ contribution is calculated to be more than 10σ . This corresponds to an observation of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay. In the simultaneous fit, shape parameters for the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ contributions are used, which are determined by a simultaneous fit of the corresponding Run 1 and Run 2 simulations. All parameters except for the yields and the slope of the exponential function are shared between both runs. The exact procedure for the determination of the significance is described in Sec. 4.4.

7.2 Studies of systematic effects

Different sources of systematic uncertainties on the ratio of branching fractions are considered in this analysis. These are studied separately for Run 1 and Run 2. All considered uncertainties and their impact on the Run 1, Run 2 and combined results are summarised in Table 7.3. A significant contribution to the systematic uncertainty is the ratio of hadronisation fractions f_s/f_d , which is calculated with the values taken from Ref. [179]. In addition, uncertainties on the ratios of the selection efficiencies are considered systematic uncertainties since they depend on the number of decays available in the simulation. The uncertainties are taken from Table 6.5. Since only unpolarised simulation is used in the analysis, a possible effect of the polarisations of the $B_s^0 \rightarrow D^{*+} D^{*-}$ and $B^0 \rightarrow D^{*+} D^{*-}$ decay products on the efficiency ratio is also considered as an uncertainty. This is discussed in more detail in Sec. 7.2.1. Systematic uncertainties arising from the method used to reject multiple candidates are also included since the fraction of multiple candidates due to the misID of final-state particles is high. This is described in Sec. 7.2.2. Another source of a systematic uncertainty arises from the choice of the model used to describe the $B_s^0 \rightarrow D^{*+} D^{*-}$ and $B^0 \rightarrow D^{*+} D^{*-}$ signal distributions in the fit to the invariant mass. A systematic uncertainty is also assigned for the choice of the range in which the invariant mass is fitted. Pseudo-experiments are performed to investigate the influence of the assumptions made during the mass fit procedure. These studies are described in Sec. 7.2.3 and Sec. 7.2.4 for the model and the fit range, respectively. The general setup of the pseudo-experiments is described in the following. In total, 100 000 pseudo-experiment data sets with $B_s^0 \rightarrow D^{*+} D^{*-}$, $B^0 \rightarrow D^{*+} D^{*-}$ and combinatorial background candidates are generated according to the PDF used to describe invariant $D^{*+} D^{*-}$ mass. The number of candidates used in the generation is drawn from a Poisson distribution with the mean value of the respective component from the fit to data. Subsequently, an extended maximum-likelihood fit is performed to the invariant mass of the generated data set. Afterwards, the resulting ratio of signal decays and its deviation from the nominal ratio, which is determined with the results from Table 7.2, is calculated for each fit. Finally, the mean value of these 100 000 deviations is calculated. This mean value is considered the systematic uncertainty, where the whole procedure is performed individually for Run 1 and Run 2. To study the effects of the assumptions mentioned above, the model and the fit range used to generate the pseudo-experiment data sets are adjusted to test the two hypotheses. Additionally, a study is performed to check for any potential fit biases from the model. In this study, pseudo-experiments are performed, where both the generation of the data and the following fit are performed with the default mass model and in the default fit range. The results do not show any biases.

The systematic uncertainties are summarised in Table 7.3. In the combination of the ratio of branching fractions, correlations between the Run 1 and Run 2 uncertainties have to be considered. The uncertainties are partially correlated between Run 1 and Run 2 for the fragmentation fraction f_s/f_d . The corresponding correlation is taken

from Ref. [179] and is given in Eq. (7.3). In more detail, the Run 1 and Run 2 values share the same systematic uncertainties, but the Run 1 measurement has an additional independent uncertainty of about 0.7%. The systematic uncertainties occurring from the size of the simulation and from the method of rejecting multiple candidates are assumed to be uncorrelated. In contrast, the systematic uncertainties for the mass model, for the range of the mass fit and for the effect of the polarisation on the efficiency ratio are assumed to be fully correlated.

Table 7.3 – Systematic uncertainties considered for the measurement of the ratio of branching fractions. All uncertainties are specified for Run 1 and Run 2. The combinations of Run 1 and Run 2 are also given, where the respective correlations are considered. The values of f_s/f_d for Run 2 and the combination seem to be the same because of the rounding.

Source	Run 1 [%]	Run 2 [%]	Combined [%]
f_s/f_d	3.18	3.11	3.11
Size of simulation	1.50	1.09	0.93
Polarisation	4.47	3.66	3.82
Multiple candidates	0.12	1.05	0.85
Mass model	0.08	0.09	0.09
Mass fit range	1.49	0.51	0.70
Total	5.88	5.06	5.13

7.2.1 Polarisation

The $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays are $B \rightarrow VV$ processes, *i.e.* processes where a pseudoscalar decays into two vector mesons. Accordingly, the final states of both decays are not pure CP eigenstates but a mixture of CP -even and CP -odd final states. This is discussed in more detail in Sec. 2.5.3

The simulation used in this analysis is not generated considering any polarisation of the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays. Instead, only the phase space is taken into account in the generation. Accordingly, only unpolarised decays are used to calculate the selection efficiencies in Sec. 6.5. Fully longitudinally and transversely polarised simulation is analysed to consider the possible effect of polarised $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays on the efficiency ratio. This additional simulation was not available at the beginning of the analysis but was generated later during the review process. Additionally, the transversely polarised simulation only considers the H_+ helicity amplitude but not the H_- amplitude. This is because the efficiencies are expected to be similar for the H_+ and H_- helicity amplitudes. The entire selection is applied to the polarised simulations, and the efficiencies are calculated like for the default simulation.

The resulting total selection efficiencies are listed in Table 7.4, where also the efficiencies for the unpolarised simulation are given.

Table 7.4 – Total selection efficiencies of the polarised simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays. The selection efficiencies calculated with unpolarised decays are also listed, which are taken from Sec. 6.5.

Polarisation	$B^0 \rightarrow D^{*+}D^{*-}$		$B_s^0 \rightarrow D^{*+}D^{*-}$	
	$\epsilon^{\text{Run 1}}[\%]$	$\epsilon^{\text{Run 2}}[\%]$	$\epsilon^{\text{Run 1}}[\%]$	$\epsilon^{\text{Run 2}}[\%]$
Unpolarised	0.0430 ± 0.0005	0.0659 ± 0.0005	0.0446 ± 0.0005	0.0713 ± 0.0005
Longitudinal	0.0424 ± 0.0004	0.0632 ± 0.0004	0.0454 ± 0.0004	0.0680 ± 0.0004
Transverse	0.0449 ± 0.0004	0.0673 ± 0.0004	0.0470 ± 0.0004	0.0726 ± 0.0004

Table 7.4 shows that the selection efficiencies of polarised decays partly differ significantly from the efficiencies calculated in Sec. 6.5. Therefore, effects on the efficiency ratio and, accordingly, on the ratio of branching fractions are studied. In more detail, the fractions of the CP eigenstates that have been measured by the Belle experiment [92] are used, which are given in Sec. 2.5.3. They are exploited to calculate the fractions of longitudinally and transversely polarised decays. These are further used to weight the total $B^0 \rightarrow D^{*+}D^{*-}$ selection efficiencies between the polarised simulated decays.

The first measurement of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay is performed in this analysis. Therefore, the polarisation of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay has not been measured yet. Different assumptions for the polarisation configuration of the B_s^0 decay products are analysed to determine a systematic uncertainty. In more detail, the selection efficiencies for the two polarised simulations are weighted in multiple combinations. The efficiency ratio is calculated for each configuration, where the weighted efficiency of the $B^0 \rightarrow D^{*+}D^{*-}$ decay is used. After calculating the ratio, the respective deviation from the efficiency ratio of the selection of unpolarised decays is calculated. The study is performed for both runs individually. The deviations are shown in Fig. 7.3 for Run 2. The results of the Run 1 study are not presented explicitly, but they show the same trend. For both runs, the largest deviation occurs if the total $B_s^0 \rightarrow D^{*+}D^{*-}$ selection efficiency is either weighted to a maximum longitudinal or a transverse polarisation. The deviations for both cases are summarised in Table 7.5 for Run 1 and Run 2. An additional case is studied, where the $B_s^0 \rightarrow D^{*+}D^{*-}$ selection efficiencies are weighted in the same way between the individual polarised simulated decays as the $B^0 \rightarrow D^{*+}D^{*-}$ efficiencies. The effect on the combined Run 1 and Run 2 selection efficiency ratio is studied for all cases. In these studies, also correlations between Run 1 and Run 2 are taken into account, *i.e.* Run 1 and Run 2 are treated as fully correlated.

The configuration leading to the highest uncertainty on the combined Run 1 and Run 2 measurement of the ratio of branching fractions is considered the systematic uncertainty. This assumes a maximum transverse polarisation of the $B_s^0 \rightarrow D^{*+}D^{*-}$

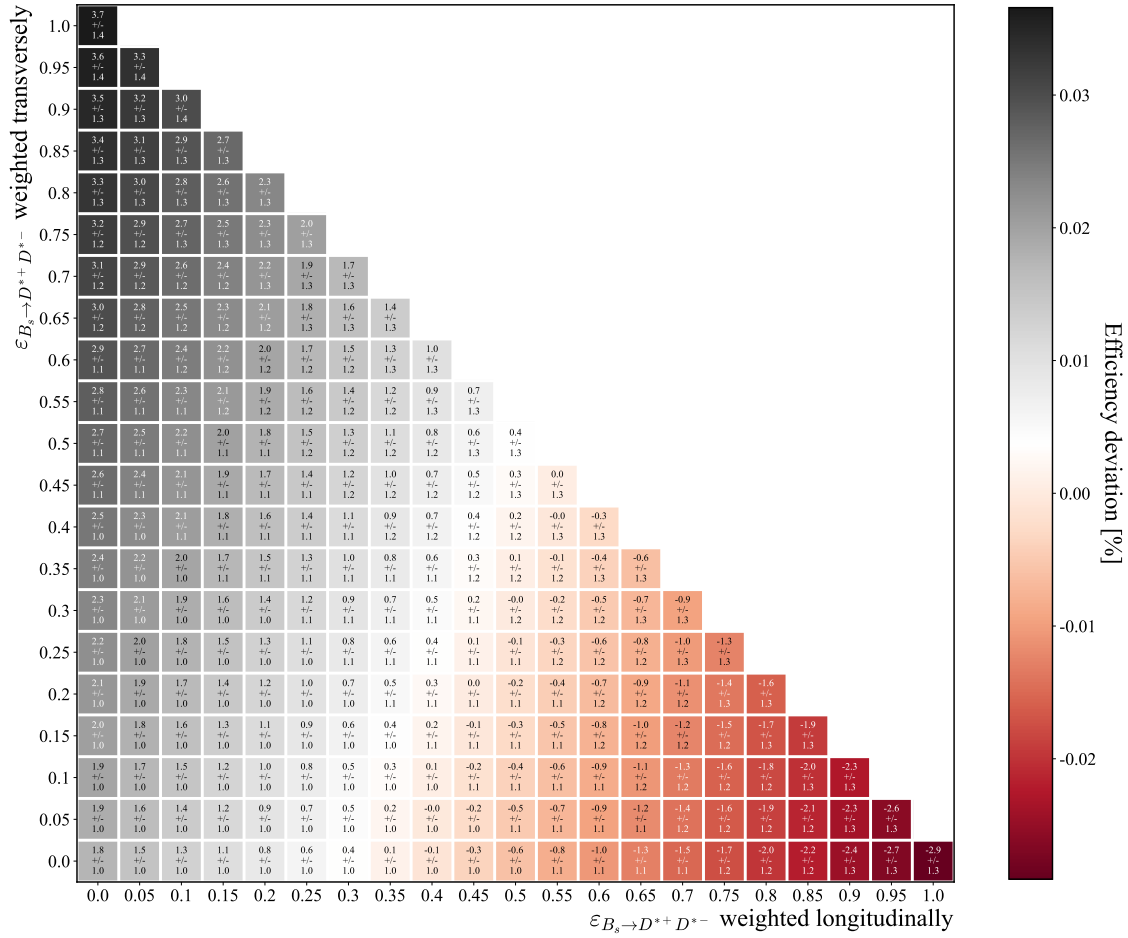


Figure 7.3 – Deviation from the nominal selection-efficiency ratio for Run 2 simulation. The deviation is shown for different weighting configurations of the selection efficiency of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay between the transversely and longitudinally polarised simulation. The total selection efficiency for the $B^0 \rightarrow D^{*+} D^{*-}$ decay is kept constant in this study. It is weighted to the fractions measured by the Belle experiment [92].

decay for Run 1 and Run 2. A further study where the longitudinal fraction is varied within its uncertainties shows no significant change in the uncertainties. The resulting systematic uncertainty on the overall combined ratio of branching fractions is determined to be 3.82%, and the individual uncertainties for Run 1 and Run 2 are 4.47% and 3.66%, respectively.

7.2.2 Multiple candidates

The fraction of multiple candidates originating from misID is high. Therefore, the method for rejecting this type of multiple candidates is considered a source of systematic

Table 7.5 – Resulting uncertainties on the efficiency ratio for Run 1, Run 2 and the combined result for different weighting configurations of the selection efficiency of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay. Correlations are also considered in the combination of Run 1 and Run 2.

Run 1		Run 2		Combination	
Polarisation	Unc. [%]	Polarisation	Unc. [%]	Correlation	Unc. [%]
longitudinal	1.04	longitudinal	2.91	fully correlated	2.54
transverse	4.47	transverse	3.66	fully correlated	3.82
like B^0	2.33	like B^0	0.44	fully correlated	0.81

uncertainty. The default method for rejecting these candidates is described in Sec. 6.3. Two additional methods are considered in this section. In the first method, the multiple candidates originating from misID are rejected randomly. The second method is to use a probability of how correctly the candidates are identified to decide which ones will be kept. The ProbNN variables of the final-state particles are used to calculate this quantity. These indicate the probability for a particle to be a pion or kaon and are provided by neural networks. The ProbNN variables of the four final-state particles of the D^0 decays, *i.e.* the kaons and pions, are used to calculate the ratios

$$p_K = \frac{\text{ProbNN}_K}{\text{ProbNN}_K + \text{ProbNN}_\pi} \quad p_\pi = \frac{\text{ProbNN}_\pi}{\text{ProbNN}_K + \text{ProbNN}_\pi}. \quad (7.12)$$

For each candidate, the ratios of all four particles are then added. The resulting sum is an estimate of how well all final-state particles have been correctly identified. For each event with multiple candidates, the candidate with the highest value of this sum is kept.

After applying one of the additional methods of rejecting multiple candidates instead of the default method, a separate fit to the invariant $D^{*+} D^{*-}$ mass is performed. With these fits, the numbers of $B_s^0 \rightarrow D^{*+} D^{*-}$ and $B^0 \rightarrow D^{*+} D^{*-}$ decays are extracted for Run 1 and Run 2 following the strategy discussed in Sec. 7.1.3. The resulting yields for the signal decays for the two additional methods and the default method are given in Table 7.6. For all methods, the ratio of signal decays is calculated, and the results are also listed in Table 7.6.

The systematic uncertainty is assumed to be the largest deviation between the ratio of signal decays of the default method and the additional methods. The largest deviation occurs in the methods of randomly rejecting the candidates and rejecting the candidates with respect to the correct identification for Run 1 and for Run 2, respectively. Consequently, the uncertainties on the method for rejecting multiple candidates that originate from misidentification are given by

$$\left. \frac{\mathcal{B}_{\text{default}} - \mathcal{B}_{\text{random}}}{\mathcal{B}_{\text{default}}} \right|_{\text{Run 1}} = 0.12\%, \quad \left. \frac{\mathcal{B}_{\text{default}} - \mathcal{B}_{\text{PID}}}{\mathcal{B}_{\text{default}}} \right|_{\text{Run 2}} = 1.05\%. \quad (7.13)$$

Table 7.6 – Number of signal decays resulting from the fit to the invariant $D^{*+}D^{*-}$ mass for the individual methods of rejecting multiple candidates originating from misID, *i.e.* the default method, random method or the correct PID method. These values are separated into Run 1 and Run 2. Furthermore, each ratio of signal decays is given. Some values seem to be the same because they are rounded according to their uncertainty.

Method	Run 1			Run 2		
	\mathcal{N}_{B^0}	$\mathcal{N}_{B_s^0}$	$\mathcal{N}_{B_s^0}/\mathcal{N}_{B^0}$	\mathcal{N}_{B^0}	$\mathcal{N}_{B_s^0}$	$\mathcal{N}_{B_s^0}/\mathcal{N}_{B^0}$
default	251 ± 16	20 ± 5	0.081 ± 0.020	1123 ± 34	79 ± 10	0.071 ± 0.010
correct PID	250 ± 16	20 ± 5	0.081 ± 0.020	1122 ± 34	78 ± 10	0.070 ± 0.010
random	249 ± 16	20 ± 5	0.081 ± 0.020	1122 ± 34	79 ± 10	0.071 ± 0.010

7.2.3 Mass model

A double-sided Hypatia function is the model used to describe the invariant $D^{*+}D^{*-}$ mass distributions of the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays. The effect of the mass-model choice on the ratio of signal candidates is studied in this section. Pseudo-experiments are performed to estimate a systematic uncertainty, where an alternative mass model is used to generate the data sets. As an alternative model, the sum of two CB functions [181] is considered. This model also accurately describes the signal distributions in the invariant $D^{*+}D^{*-}$ mass of simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays. One CB function consists of a Gaussian function for the central region and a power-law tail for the region of either low or high values. It is defined as

$$f_{\text{CB},i}(x, \alpha_i, n_i, \mu_i, \sigma_i) \propto \begin{cases} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right) & , \text{if } \frac{x-\mu_i}{\sigma_i} > -\alpha_i \\ A \cdot \left(B - \frac{x-\mu_i}{\sigma_i}\right)^{-n_i} & , \text{if } \frac{x-\mu_i}{\sigma_i} \leq -\alpha_i \end{cases} . \quad (7.14)$$

The parameters A and B are the same as in the definition of the Hypatia function and are given in Eq. (7.7). The overall PDF $\mathcal{P}_{\text{CB}}^j$ for each signal contribution j consists of two CB functions. While a shared mean value μ is used for both CB functions, the parameters α_i , n_i and σ_i are different for each of the CB functions. Additionally, opposite directions of the power-law tails are required by $\alpha_1 < 0$ and $\alpha_2 > 0$. The PDF used to describe the invariant mass distribution of each signal contribution is then given by

$$\mathcal{P}_{\text{CB}}^j = g \cdot \mathcal{P}_{\text{CB},1}^j + (1-g) \cdot \mathcal{P}_{\text{CB},2}^j, \quad (7.15)$$

where g is the fraction between the CB functions.

The parameters used in the generation of the pseudo-experiments are obtained by fits following the same procedure as described in Sec. 7.1. First, the parameters for the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ signal shapes are determined with fits to the respective

simulation, where the alternative mass model is used. The detailed results are given in Table 7.7, where also the χ^2/ndf values are listed. The χ^2/ndf values show that the model works similarly well compared to the default model for both decays and runs.

Table 7.7 – Results of the fits to the invariant $D^{*+}D^{*-}$ mass distribution of simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays with the alternative mass model. The parameters are separated into Run 1 and Run 2. The resulting χ^2/ndf values are also listed.

Parameter	$B_s^0 \rightarrow D^{*+}D^{*-}$		$B^0 \rightarrow D^{*+}D^{*-}$	
	Run 1	Run 2	Run 1	Run 2
μ [MeV/c ²]	5367.19 ± 0.29	5367.04 ± 0.19	5279.54 ± 0.23	5280.03 ± 0.24
σ_1 [MeV/c ²]	12.0 ± 1.4	7.4 ± 0.5	8.0 ± 0.5	8.1 ± 0.4
σ_2 [MeV/c ²]	7.6 ± 0.8	13.4 ± 1.2	17.0 ± 2.0	10.6 ± 1.6
α_1	-1.59 ± 0.22	-2.2 ± 0.4	-2.64 ± 0.33	-1.44 ± 0.17
α_2	1.5 ± 0.4	1.66 ± 0.17	1.55 ± 0.31	0.70 ± 0.21
g	0.56 ± 0.13	0.42 ± 0.07	0.67 ± 0.07	0.65 ± 0.09
n_1	2.6 ± 0.7	0.73 ± 0.33	0.64 ± 0.31	1.99 ± 0.28
n_2	1.32 ± 0.35	1.7 ± 0.4	1.9 ± 1.5	5 ± 4
χ^2/ndf	0.73	0.69	0.7	0.81

After the fits to simulation, the recorded data is fitted with the alternative mass model. A scaling factor s for the widths of the B_s^0 and B^0 signal shapes is introduced to consider resolution differences between simulation and data, *i.e.* $\sigma_i^{\text{data}} = s\sigma_i^{\text{sim}}$. Only a single scaling factor for both signal shapes is used, as the resolution effects are expected to be the same. In Sec. 7.1.3, it was noted that the width of the B_s^0 distribution for Run 1 is noticeably smaller compared to the B^0 distribution and also compared to the fit to the simulation. By using a single scaling factor for the widths of the CB functions in the alternative model, the difference in the widths between the B^0 and B_s^0 signal shapes is further covered by the systematic uncertainty. In the fit to recorded data, the parameters of the B_s^0 and B^0 signal shapes are fixed to the values obtained with the fits to simulation. Additionally, the B_s^0 mean is assumed to be shifted by $87.26 \text{ MeV}/c^2$ [182] compared to the B^0 mean, *i.e.* $\mu_{B_s^0} = 87.26 \text{ MeV}/c^2 + \mu_{B^0}$. As for the fit described in Sec. 7.1.3, the remaining background is described by an exponential function. The free parameters in the fit to recorded data are the slope of the exponential function, the B^0 mean, the scaling factor s and the yield of each contribution. The result from this fit is used in the generation of the data sets of the pseudo-experiments except for the yields and the exponential slope. These parameters are taken from the fit to recorded data performed with the default model given in Table 7.2. The final parameters used in the generation with the alternative mass model are summarised in Table 7.8. After the generation of the data sets, fits to the invariant $D^{*+}D^{*-}$ mass are performed using the default mass model

described in Sec. 7.1.1. In these fits, all parameters are fixed, except for the B^0 mean, the widths of the B_s^0 and B^0 signal PDFs, the slope of the exponential background and the yields. The yields of the B^0 and B_s^0 signal contributions are extracted for each fit. With these yields, the ratio of signal decays and its deviation from the nominal ratio, which is calculated with the results from Table 7.2, is determined. The resulting deviations are assigned as the systematic uncertainties on the mass model. The procedure is performed for Run 1 and Run 2 separately.

Table 7.8 – Parameters used in the generation of the pseudo-experiments with the alternative model. The yields and the slope of the exponential function are taken from Sec. 7.1.3. The parameters used from simulation fit with the alternative mass model are taken from Table 7.7. The parameters resulting from the fit to simulation with the alternative mass model are taken from Table 7.7.

Origin	Parameter	Run 1	Run 2	
From default	$\mathcal{N}_{B_s^0}$	21 ± 5	80 ± 11	
fit to data	\mathcal{N}_{B^0}	263 ± 17	1171 ± 35	
	$\mathcal{N}_{\text{Comb}}$	31 ± 7	178 ± 17	
	λ_{exp} [c ² /MeV]	-0.0034 ± 0.0012	-0.00198 ± 0.00035	
From fit	μ_{B^0} [MeV/c ²]	5280.9 ± 0.6	5279.58 ± 0.28	
to data	s	0.86 ± 0.05	0.878 ± 0.028	
From fit	$\sigma_1^{B^0}$ [MeV/c ²]	8.0 ± 0.5	8.1 ± 0.4	
to simulation	$\sigma_2^{B^0}$ [MeV/c ²]	17.0 ± 2.0	10.6 ± 1.6	
	$\sigma_1^{B_s^0}$ [MeV/c ²]	12.0 ± 1.4	7.4 ± 0.5	
	$\sigma_2^{B_s^0}$ [MeV/c ²]	7.6 ± 0.8	13.4 ± 1.2	
	$\alpha_1^{B^0}$	-2.64 ± 0.33	-1.44 ± 0.17	
	$\alpha_2^{B^0}$	1.55 ± 0.31	0.70 ± 0.21	
	$\alpha_1^{B_s^0}$	-1.59 ± 0.22	-2.2 ± 0.4	
	$\alpha_2^{B_s^0}$	1.5 ± 0.4	1.66 ± 0.17	
	g^{B^0}	0.67 ± 0.07	0.65 ± 0.09	
	$g^{B_s^0}$	0.56 ± 0.13	0.42 ± 0.07	
	$n_1^{B^0}$	0.64 ± 0.31	1.99 ± 0.28	
	$n_1^{B_s^0}$	2.6 ± 0.7	0.73 ± 0.33	
	$n_2^{B^0}$	1.9 ± 1.5	5 ± 4	
	$n_2^{B_s^0}$	1.32 ± 0.35	1.7 ± 0.4	
	Always fixed	$\Delta\mu_{B^0}$ [MeV/c ²]	87.26	87.26

The residuals and pulls of the B^0 and B_s^0 yields are shown in Fig. 7.4 and Fig. 7.5, but not directly used in the calculation of the uncertainty. No bias is expected due to the mean of the pull distribution of the number of B_s^0 decays being slightly off from zero as the bias corresponds to the mean of the residual, which is close to zero. The exact values of the means of the residual distributions are 0.379 and 0.469 for Run 1 and Run 2, respectively. Dividing these by the number of generated events (20 for Run 1; 79 for Run 2) would result in a relative bias of 1.80% and 0.6% for Run 1 and Run 2, respectively. This potential bias of the B_s^0 yield is neglected in the following since it is smaller compared to the assigned systematic uncertainty. The pull can be off due to different effects, *e.g.* the uncertainty estimation by MINUIT could be inaccurate. Such higher-order effects, as seen in the pull distribution, are neglected in this analysis. Overall, the systematic uncertainties resulting from this study are 0.08% and 0.09% for Run 1 and Run 2, respectively.

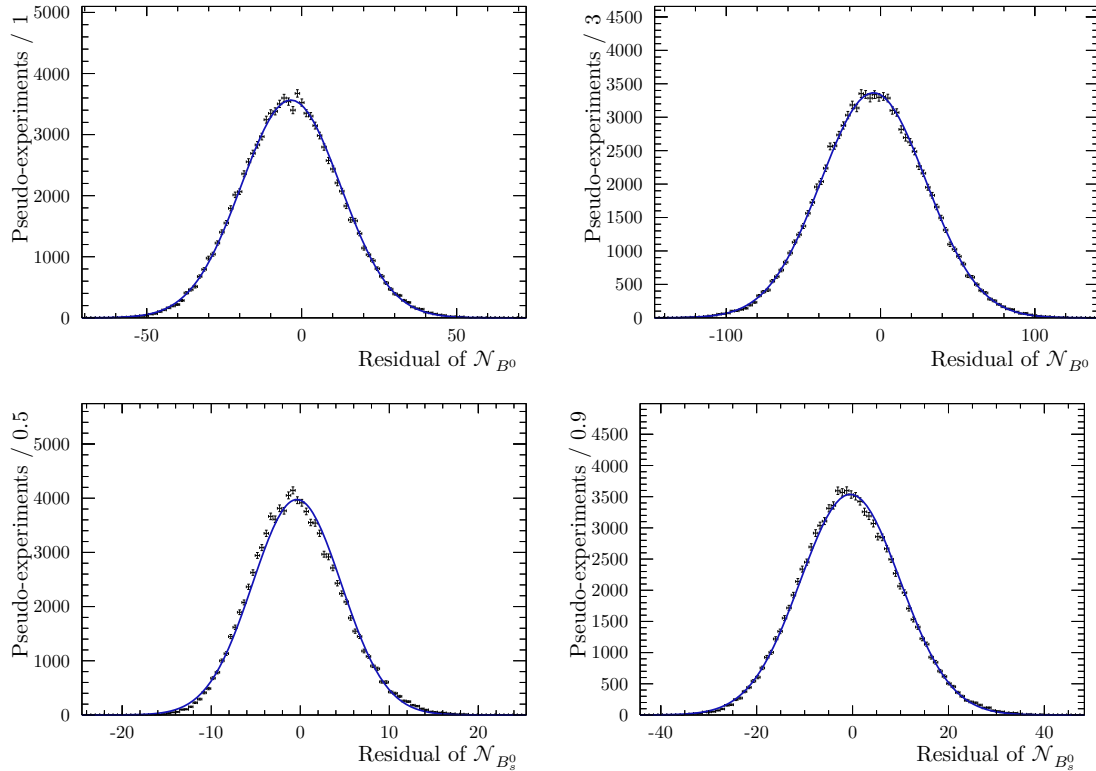


Figure 7.4 – Residuals of the yields resulting from the fits for the mass model study for the (top) $B^0 \rightarrow D^{*+}D^{*-}$ and (bottom) $B_s^0 \rightarrow D^{*+}D^{*-}$ decays for (left) Run 1 and (right) Run 2, respectively. The blue line shows the result of a maximum-likelihood fit to the distributions using a Gaussian function. The mean values resulting from these fits are close to zero, *i.e.* the systematic uncertainty is expected to be small.

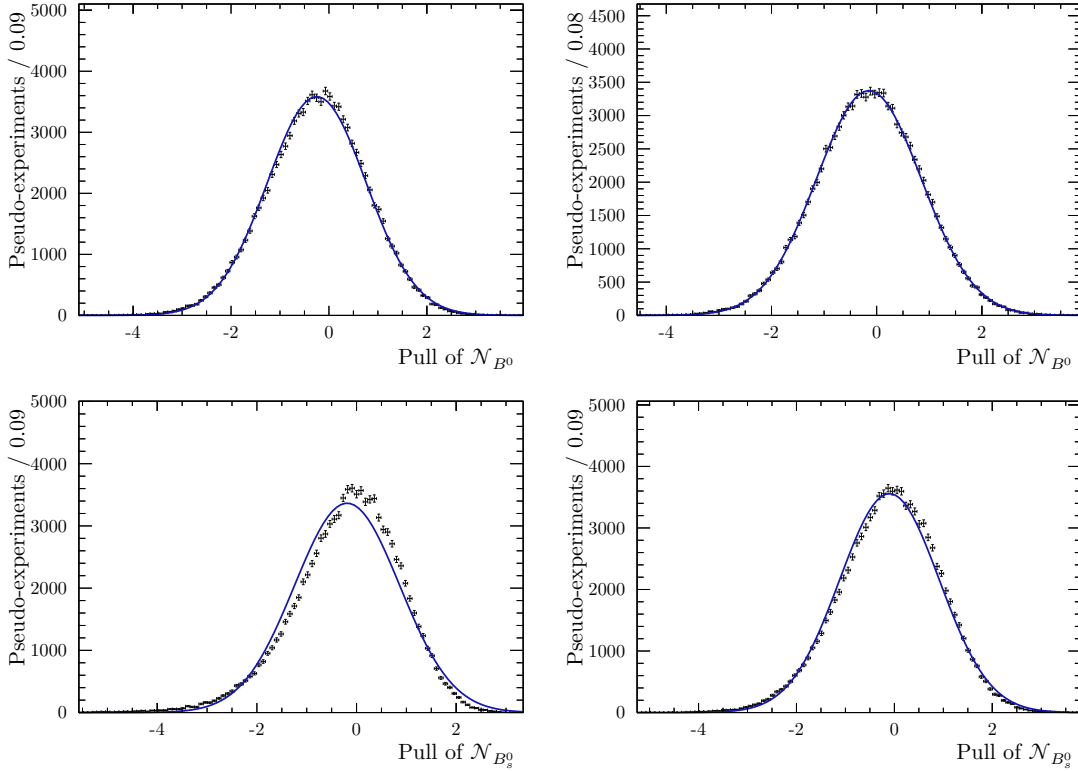


Figure 7.5 – Pulls of the yields resulting from the fits for the mass model study for the (top) $B^0 \rightarrow D^{*+}D^{*-}$ and (bottom) $B_s^0 \rightarrow D^{*+}D^{*-}$ decays for (left) Run 1 and (right) Run 2, respectively. The blue line shows the result of a maximum-likelihood fit to the distributions using a Gaussian function. Although the B_s^0 mean resulting from these fits deviates slightly from zero, no bias is expected. This is explained in more detail in the text.

7.2.4 Mass fit range

The fits to the invariant $D^{*+}D^{*-}$ mass to extract the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ candidates are performed in the range of $5100 \text{ MeV}/c^2$ to $6000 \text{ MeV}/c^2$ as described in Sec. 7.1.3. Potential sources of partially reconstructed background are found to be negligible in the considered mass region. This is studied in more detail in Sec. 6.4. Still, contributions from the physical background that are not described in the mass fit could change the shape of the combinatorial background. Therefore, this section studies the effect of the choice of the mass range on the ratio of branching fractions.

The size of the effect is estimated by performing pseudo-experiments, where data is generated according to the default model but using parameters for the background component that are determined on alternative mass ranges. Alternative ranges that are considered in this study are $5050 \text{ MeV}/c^2$ to $6000 \text{ MeV}/c^2$ and $5150 \text{ MeV}/c^2$ to $6000 \text{ MeV}/c^2$. The parameters for the description of the background are then obtained by fitting the

invariant $D^{*+}D^{*-}$ mass in the respective ranges. The parameters for the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ signal contributions are taken from the results of the nominal mass fit given in Table 7.2. The data sets of the pseudo-experiments are generated within the alternative ranges and a subsequent maximum-likelihood fit in the same range is performed. The starting values for the parameters of these fits are those used to generate the data sets, and the floating and fixed parameters are the same as for the nominal mass fit. This procedure is performed separately for Run 1 and Run 2.

A systematic uncertainty is assigned by calculating the ratio of signal decays resulting from the fits and its deviation from the nominal ratio. For this purpose, the maximum deviation occurring for the two ranges is used for each run. This gives uncertainties of 1.49% for Run 1 and 0.51% for Run 2. The residual and pull distributions of the B^0 and B_s^0 yields for both fit ranges are not explicitly shown in this thesis. However, the residuals are close to zero, and the pulls show the same trend as for the study of the mass model in Sec. 7.2.3. Since the mean values of the residuals are very close to zero, no bias is expected in the study of the fit ranges either.

7.3 Calculation of the ratio of branching fractions

The ratio of branching fractions of the $B_s^0 \rightarrow D^{*+}D^{*-}$ and the $B^0 \rightarrow D^{*+}D^{*-}$ decays can be calculated with the selection efficiencies, the number of signal decays and the hadronisation probabilities of the two decays. It is defined in Eq. (2.46). According to Eq. (7.2) and Eq. (7.3) the ratios of hadronisation fractions are given by

$$\left. \frac{f_s}{f_d} \right|_{\text{Run 1}} = 0.239 \pm 0.008, \quad \left. \frac{f_s}{f_d} \right|_{\text{Run 2}} = 0.254 \pm 0.008. \quad (7.16)$$

The efficiencies of the entire selection are calculated in Sec. 6.5 for both decays. The resulting efficiency ratios for Run 1 and Run 2 are given by

$$\left. \frac{\varepsilon_{B_s^0}}{\varepsilon_{B^0}} \right|_{\text{Run 1}} = 1.037 \pm 0.016, \quad \left. \frac{\varepsilon_{B_s^0}}{\varepsilon_{B^0}} \right|_{\text{Run 2}} = 1.082 \pm 0.012, \quad (7.17)$$

where it is assumed that the ratio is the same for all final states. The efficiencies of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay are higher due to the better signal efficiency of the central preselection and the geometric acceptance. The ratios of the numbers of signal decays are calculated with the results from the fit to the invariant $D^{*+}D^{*-}$ mass of the Run 1 and Run 2 data sets. The fits are discussed in Sec. 7.1.3. The resulting ratios are

$$\left. \frac{\mathcal{N}_{B_s^0}}{\mathcal{N}_{B^0}} \right|_{\text{Run 1}} = 0.081 \pm 0.020, \quad \left. \frac{\mathcal{N}_{B_s^0}}{\mathcal{N}_{B^0}} \right|_{\text{Run 2}} = 0.071 \pm 0.010. \quad (7.18)$$

With the individual values for the efficiencies and the numbers of signal decays, the ratios of branching fractions are calculated to be

$$\left. \frac{\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-})}{\mathcal{B}(B^0 \rightarrow D^{*+}D^{*-})} \right|_{\text{Run 1}} = 0.33 \pm 0.08 \pm 0.01 \pm 0.02, \quad (7.19)$$

$$\left. \frac{\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-})}{\mathcal{B}(B^0 \rightarrow D^{*+}D^{*-})} \right|_{\text{Run 2}} = 0.257 \pm 0.035 \pm 0.008 \pm 0.010, \quad (7.20)$$

where the first uncertainty is statistical, the second systematic, and the third is due to the uncertainty of the fragmentation-fraction ratio f_s/f_d . The systematic uncertainties are evaluated in Sec. 7.2. A combination that is weighted with the quadratic sum of the uncertainties of the individual ratios of Run 1 and Run 2 results in

$$\frac{\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-})}{\mathcal{B}(B^0 \rightarrow D^{*+}D^{*-})} = 0.269 \pm 0.032 \pm 0.008 \pm 0.011, \quad (7.21)$$

where also the correlations of the individual uncertainties between Run 1 and Run 2 are considered. The individual and the combined results are illustrated in Fig. 7.6.

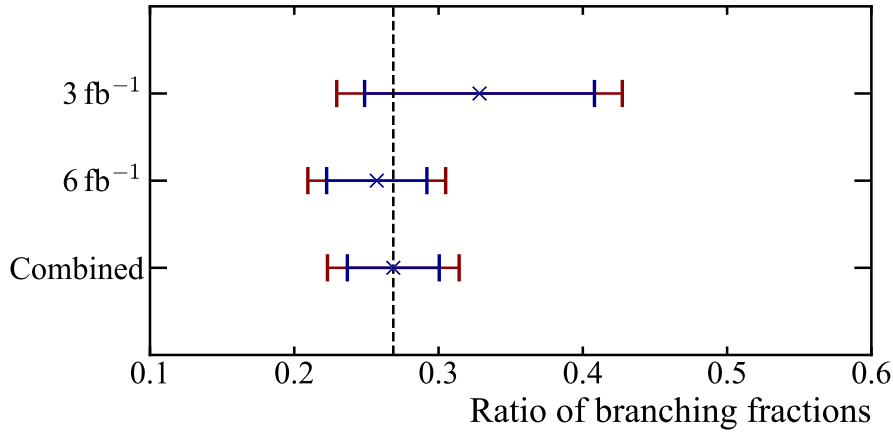


Figure 7.6 – Ratio of branching fractions for Run 1 with 3 fb^{-1} , for Run 2 with 6 fb^{-1} and for the combination of the individual results. The statistical and overall uncertainties are indicated in blue and red for each ratio, respectively.

As discussed in Sec. 2.5, the absolute branching fraction of the $B^0 \rightarrow D^{*+}D^{*-}$ decay has already been measured. By using the current world average of the branching fraction of this decay, which is given in Eq. (2.44) and taken from Ref. [173], the $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction is calculated to be

$$\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-}) = (2.15 \pm 0.26 \pm 0.09 \pm 0.06 \pm 0.16) \times 10^{-4},$$

where the first three sources of uncertainty are the same as in Eq. (7.20), and the fourth uncertainty arises from the limited precision of the $B^0 \rightarrow D^{*+}D^{*-}$ branching fraction.

7.4 Effect of the B_s^0 lifetime

This analysis exploits a time-integrated approach to extract the number of signal decays needed to calculate the central value of the ratio of branching fractions. Afterwards, the absolute $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction is calculated by using the world average of the $B^0 \rightarrow D^{*+}D^{*-}$ branching fraction. Consequently, $B_s^0 - \bar{B}_s^0$ mixing effects are not taken into account in the determination of the $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction. According to Ref. [183], the experimental branching fractions \mathcal{B}_{exp} can be converted into a ratio that is independent of $B_s^0 - \bar{B}_s^0$ mixing. The theoretical branching fraction \mathcal{B}_{th} of the B_s^0 meson decaying into any final state f is given by

$$\mathcal{B}_{\text{th}}(B_s^0 \rightarrow f) = \left(\frac{1 - y_s^2}{1 + \mathcal{A}_{\Delta\Gamma}^f y_s} \right) \mathcal{B}_{\text{exp}}(B_s^0 \rightarrow f), \quad (7.22)$$

where the parameters y_s and $\mathcal{A}_{\Delta\Gamma}^f$ are defined as

$$y_s = \frac{\Delta\Gamma_s}{2\Gamma_s} \quad (7.23)$$

and

$$\mathcal{A}_{\Delta\Gamma}^f = \frac{R_{\text{H}}^f - R_{\text{L}}^f}{R_{\text{H}}^f + R_{\text{L}}^f} \quad (7.24)$$

with the decay amplitudes of the heavy (H) and light (L) eigenstates of the B_s^0 meson R_{H}^f and R_{L}^f , respectively.

The $B_s^0 \rightarrow D^{*+}D^{*-}$ simulation used in this analysis is generated with the average B_s^0 lifetime since no measurement or prediction is available for the B_s^0 lifetime in this decay. However, the heavy and light eigenstates have significantly different lifetimes with $\tau_{B_s^0, \text{L}} = 1.423$ ps and $\tau_{B_s^0, \text{H}} = 1.620$ ps, both taken from Ref. [173]. Since the B_s^0 selection efficiency depends on the lifetime, correction factors for the efficiency are calculated for the scenarios where either a purely heavy or a purely light B_s^0 eigenstate is considered. Assuming an average B_s^0 lifetime corresponds to $\mathcal{A}_{\Delta\Gamma}^f = 0$, the lifetime of a heavy or light B_s^0 eigenstate corresponds to $\mathcal{A}_{\Delta\Gamma}^f = \pm 1$. The effect of a change in the B_s^0 lifetime on the selection efficiency is studied with simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ decays, where the complete selection is applied. The lifetime of each simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ decay is

weighted with

$$w_{L,H} = \frac{\exp(-t/\tau_{L,H})}{\tau_{L,H}} \bigg/ \frac{\exp(-t/\tau)}{\tau}, \quad (7.25)$$

i.e. the normalised lifetime PDF of the heavy or light eigenstate divided by the normalised lifetime PDF of the average. The correction factor is then calculated as the sum of the weights divided by the total number of decays. In Fig. 7.7, the weighted B_s^0 lifetime is shown for Run 1 and Run 2 simulation.

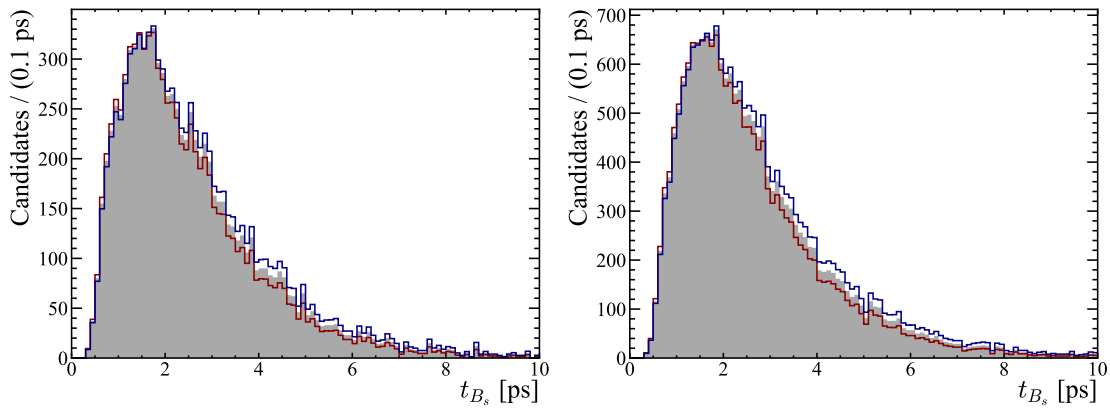


Figure 7.7 – Distribution of the B_s^0 lifetime of simulated $B_s^0 \rightarrow D^{*+}D^{*-}$ decays for (left) Run 1 and (right) Run 2. The distribution generated with the average B_s^0 lifetime is shown in grey, while the distributions weighted by the theoretical lifetime of the heavy and light B_s^0 mesons are shown in blue and red, respectively.

The correction factors for a change to the lifetime of the heavy or light B_s^0 eigenstate are determined to be 1.041 and 0.951 for Run 1 and 1.045 and 0.946 for Run 2, respectively. The time-integrated correction factors for the efficiency are 0.947 and 1.045 for light and heavy B_s^0 eigenstates, respectively. If the lifetime of the B_s^0 meson in the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay is found in further studies to differ significantly from the average B_s^0 lifetime, the central value of the $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction presented in this thesis can be corrected. The equivalent effect also occurs for the $B^0 \rightarrow D^{*+}D^{*-}$ decay but is neglected due to the small value of $\Delta\Gamma_d$ [47].

7.5 Results

In this chapter, the first measurement of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay is presented. The measurement is performed with the full Run 1 and Run 2 data sets of the LHCb experiment. The selection of these samples is described in Chap. 6.

First, extended maximum-likelihood fits to the invariant $D^{*+}D^{*-}$ mass are performed to extract the numbers of $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays. This is described in Sec. 7.1. Overall, the numbers of $B_s^0 \rightarrow D^{*+}D^{*-}$ decays are 20 ± 5 and 79 ± 10 for Run 1 and Run 2, respectively. The numbers of $B^0 \rightarrow D^{*+}D^{*-}$ decays are 251 ± 16 and 1123 ± 34 for Run 1 and Run 2, respectively. This corresponds to an observation of the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay with a significance of more than 10σ , which is determined with the procedure described in Sec. 4.4. The individual Run 1 and Run 2 fit results are shown in Fig. 7.2, and a combination is presented in Fig. 7.8.

Figure 7.8 – Reconstructed invariant $D^{*+}D^{*-}$ mass shown together with the PDF of a Run 1 and Run 2 combination with a logarithmic scale [184]. The combination is calculated as the sum of the individual PDFs.

Systematic uncertainties considered in the measurement are discussed in Sec. 7.2. Uncertainties regarding the mass model to describe the invariant $D^{*+}D^{*-}$ mass and fit range in which the fit is performed are considered. Systematic uncertainties are also assigned for the method of rejecting the high fraction of multiple candidates originating from misID and for the size of the simulation. The dominant systematic uncertainties originate from the external input of the ratio of hadronisation fractions f_s/f_d and an effect of polarised $B \rightarrow D^{*+}D^{*-}$ decays on the selection efficiencies. The simulation used in this thesis is generated without considering a polarisation present in the $B \rightarrow D^{*+}D^{*-}$ decay topology. A detailed and conservative study of possible effects results in systematic uncertainties of 4.47% and 3.66% for Run 1 and Run 2, respectively.

By using the total selection efficiencies reported in Sec. 6.5, the ratio of hadronisation fractions and the extracted number of signal decays, the ratio of branching fractions is calculated. This is described in more detail in Sec. 7.3. The resulting ratio of branching fractions of a Run 1 and Run 2 combination is calculated to be

$$\frac{\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-})}{\mathcal{B}(B^0 \rightarrow D^{*+}D^{*-})} = 0.33 \pm 0.08 \pm 0.01 \pm 0.02, \quad (7.26)$$

where the first uncertainty is statistical, the second systematic, and the third is due to the uncertainty of the fragmentation-fraction ratio f_s/f_d . In the combination, the individual

ratios of Run 1 and Run 2 are weighted by the quadratic sum of their uncertainties. Correlations of the individual uncertainties between Run 1 and Run 2 are also considered. By using the current world average of the branching fraction of the $B^0 \rightarrow D^{*+} D^{*-}$ decay taken from Ref. [173], the absolute time-integrated $B_s^0 \rightarrow D^{*+} D^{*-}$ branching fraction is determined to be

$$\mathcal{B}(B_s^0 \rightarrow D^{*+} D^{*-}) = (2.15 \pm 0.26 \pm 0.09 \pm 0.06 \pm 0.16) \times 10^{-4},$$

where the first three sources of uncertainty are the same as in Eq. (7.20), and the fourth uncertainty arises from the limited precision of the $B^0 \rightarrow D^{*+} D^{*-}$ branching fraction. An effect of the B_s^0 lifetime on the selection efficiency is discussed in Sec. 7.4. Correction factors are calculated for the scenarios where either a purely heavy or a purely light B_s^0 eigenstate is considered. The time-integrated correction factors for the efficiency are determined to be 0.947 and 1.045 for light and heavy B_s^0 eigenstates, respectively. These can be used to correct the central value of the $B_s^0 \rightarrow D^{*+} D^{*-}$ branching fraction if the lifetime of the B_s^0 meson in the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay is found in further studies to differ significantly from the average B_s^0 lifetime.

8 Measurement of CP violation in $B^0 \rightarrow D^{*+} D^{*-}$ decays

The measurement of CP violation in $B^0 \rightarrow D^{*+} D^{*-}$ decays is presented in this chapter. The selection of the decays is described in Chap. 6, which is similar but slightly different compared to the $B_s^0 \rightarrow D^{*+} D^{*-}$ analysis. Consequently, the extraction of signal decays by performing a maximum-likelihood fit is repeated in this analysis. This is discussed in Sec. 8.1, where also *sWeights* are determined, which allow to unfold the signal decay contribution from the data set. The CP observables are measured by a fit to the decay-time distribution of statistically subtracted $B^0 \rightarrow D^{*+} D^{*-}$ candidates. Consequently, knowledge about the time-dependent decay rate of the $B^0 \rightarrow D^{*+} D^{*-}$ decays is required. The time evolution equations describing the oscillation of a particle and an antiparticle in the B^0 system are given in Eq. (2.61). They can be simplified to a single formula defined as

$$\mathcal{P}_{\text{th}}(t, q) \propto e^{-\Gamma t} (1 + A \hat{f}) \cdot [1 - q(1 - R_{\perp}) S_{D^{*+} D^{*-}} \sin(\Delta m_d t) + q C_{D^{*+} D^{*-}} \cos(\Delta m_d t)] , \quad (8.1)$$

where identical constant factors are neglected and where the parameter q can take the values $+1$ and -1 for B^0 and \bar{B}^0 mesons, respectively. However, this simplified formula is insufficient to describe the decay-time distribution measured by the detector since it neglects several experimental effects. Such effects arise from the FT, production and detection asymmetries, decay-time-dependent selection efficiency and the finite resolution in the decay-time measurement. Consequently, the decay-time PDF has to be adjusted to account for these effects. In general, this analysis uses the decay-time estimate from a DTF with constraints placed on the PV, which allows to increase the resolution. A detailed description of the experimentally measured decay-time parametrisation is given in Sec. 8.2. Afterwards, the extraction of the CP observables is discussed in Sec. 8.3. Studies of systematic effects on the measurement of the CP observables are presented in Sec. 8.4. Finally, the results of the measurement are summarised in Sec. 8.5.

8.1 Extraction of B^0 candidates

An extended maximum-likelihood fit to the reconstructed invariant $D^{*+} D^{*-}$ masses of the selected $B^0 \rightarrow D^{*+} D^{*-}$ decays is performed to discriminate between signal and

background. After the fit, the *sPlot* method is used to extract signal weights to later statistically subtract the background in the angular and decay-time distributions. The *sPlot* technique is described in more detail in Sec. 4.3.

As for the $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction measurement, the mass estimate from a DTF with constraints placed on the PV and the masses of the intermediate charm mesons is used. This allows to increase the B^0 mass resolution by about 50%. The $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis shows that a double-sided Hypatia function describes the invariant $D^{*+}D^{*-}$ mass distribution well. Consequently, the model is also used in the $B^0 \rightarrow D^{*+}D^{*-}$ analysis to parametrise the $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ contributions. It is discussed in more detail in Sec. 7.1.1. The fit strategy is also adopted from the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis. First, the shape of the Hypatia functions for the $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ contributions is determined using simulation. The shape of the signal contributions in the $B^0 \rightarrow D^{*+}D^{*-}$ analysis can differ due to the different selections. Consequently, the determination of the Hypatia parameters is repeated. This step is described in Sec. 8.1.1. Afterwards, a fit to data is performed in Sec. 8.1.2 to extract the *sWeights*.

8.1.1 Fit to simulation

The fits to the invariant $D^{*+}D^{*-}$ mass distribution of simulated $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ decays are performed in the range of $5200 \text{ MeV}/c^2$ to $5500 \text{ MeV}/c^2$. The data sets for each decay and for Run 1 and Run 2 are fitted separately. The parameters ζ and β are set to zero to achieve a more stable fit. All remaining parameters are floating.

The results of the fits to $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ simulation are listed in Table 8.1. Despite minor differences are present due to the different selections, the results consistent with the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis. To avoid redundancy, the invariant $D^{*+}D^{*-}$ mass distributions overlaid with the projection of the results are not shown. However, the double-sided Hypatia function describes the simulation accurately for all data sets. The results are used for the fits to Run 1 and Run 2 recorded data.

8.1.2 Fit to data

Extended maximum-likelihood fits to the invariant $D^{*+}D^{*-}$ mass distribution of recorded data are presented in this section. The invariant $D^{*+}D^{*-}$ mass is restricted to be in the range of $5100 \text{ MeV}/c^2$ to $5500 \text{ MeV}/c^2$. This range reduces partially reconstructed backgrounds to a negligible level, which appear at lower invariant masses with respect to the nominal B^0 mass. Potential sources of partially reconstructed backgrounds are discussed in Sec. 6.4. These could arise from $B^{0(+)} \rightarrow D^{*+}D^{*-}\pi^{0(+)}$ decays, which may involve intermediate $D_1(2420)^{0(+)} \rightarrow D^{*+}\pi^{-(0)}$ or $D_1(2430)^{0(+)} \rightarrow D^{*+}\pi^{-(0)}$ decays. Additionally, cross-checks for backgrounds arising from the misID of final-state particles are considered in Sec. 6.4. However, no significant background contributions are present in the data. All checks are performed for the selection of the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis, but it is assumed that the same applies to the $B^0 \rightarrow D^{*+}D^{*-}$ analysis. The upper mass limit

Table 8.1 – Results of the fits to the invariant $D^{*+}D^{*-}$ mass distribution of simulated $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ decays. They are separated into Run 1 and Run 2. The parameters ζ and β are fixed to zero to simulation to achieve a more stable fit. All remaining parameters are floating.

Parameter	$B_s^0 \rightarrow D^{*+}D^{*-}$		$B^0 \rightarrow D^{*+}D^{*-}$	
	Run 1	Run 2	Run 1	Run 2
μ_{B^0} [MeV/c ²]	5367.05 ± 0.04	5367.089 ± 0.035	5279.870 ± 0.032	5280.016 ± 0.035
σ [MeV/c ²]	10.57 ± 0.33	11.15 ± 0.21	10.56 ± 0.14	10.40 ± 0.30
$\lambda_{\text{Hyp}}^{B_s^0}$	-3.1 ± 0.4	-2.67 ± 0.20	-2.74 ± 0.15	-3.0 ± 0.4
α_1	1.56 ± 0.10	1.56 ± 0.07	1.83 ± 0.10	1.46 ± 0.06
α_2	1.72 ± 0.16	1.83 ± 0.19	2.1 ± 0.5	1.79 ± 0.21
n_1	5.8 ± 0.7	4.9 ± 0.4	4.2 ± 0.4	5.8 ± 0.4
n_2	10.0 ± 1.4	8.7 ± 1.2	12.8 ± 2.8	8.5 ± 1.5
β	0.0	0.0	0.0	0.0
ζ	0.0	0.0	0.0	0.0

is chosen to achieve a more accurate *sWeights* extraction. In detail, the shrinking of the range has the advantage that the correction of the *sWeights* applied in the following is smaller.

The fit to simulation performed in Sec. 8.1.1 only considers the $D^0 \rightarrow K^- \pi^+$ decay mode. However, the resolution is assumed to be similar for all considered decay modes. This assumption is validated in the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis. The recorded data set comprises all considered D^0 decay modes, which are fitted simultaneously. Three contributions have to be considered in the fit to the invariant $D^{*+}D^{*-}$ mass of recorded data, *i.e.* the $B^0 \rightarrow D^{*+}D^{*-}$ decays, the $B_s^0 \rightarrow D^{*+}D^{*-}$ decays and the combinatorial background. The parametrisations of the $B^0 \rightarrow D^{*+}D^{*-}$ and $B_s^0 \rightarrow D^{*+}D^{*-}$ signal contributions determined with the fits to simulation are used. The combinatorial background is described by an exponential function. In more detail, all parameters listed in Table 8.1 are used and fixed in the fit to data except for the B^0 mean and the widths. The B_s^0 mean is shifted relative to the B^0 mean by a fixed value calculated with the measured masses of the B^0 and B_s^0 mesons taken from Ref. [47]. The remaining parameters of the model, *i.e.* the yields of the three contributions and the slope of the exponential function, are floating. Separate fits are performed for Run 1 and Run 2 data, and the results are listed in Table 8.2. Fig. 8.1 shows the invariant $D^{*+}D^{*-}$ mass distribution and the projection of the fit result for Run 1 and Run 2. The individual contributions of the $B_s^0 \rightarrow D^{*+}D^{*-}$ and $B^0 \rightarrow D^{*+}D^{*-}$ decays and the combinatorial background are also shown.

The PDFs resulting from the fits and the *sPlot* method are used to determine the *sWeights*. More details about the *sPlot* method are given in Sec. 4.3. These *sWeights* are used in the following to unfold the signal and background distributions in the angular decay-time fit. After determining the *sWeights*, a likelihood correction factor of

Table 8.2 – Parameters resulting from the fit to the reconstructed invariant $D^{*+}D^{*-}$ mass distribution of recorded data for Run 1 and Run 2. Only the floating parameters and the constant shift $\Delta\mu_{B^0}$ of the B_s^0 mean relative to the B^0 mean are listed. The remaining parameters of the mass model are fixed to the values obtained with the fits to simulation given in Table 8.1.

Parameter	Run 1	Run 2
$\mathcal{N}_{B_s^0}$	25 ± 6	92 ± 15
\mathcal{N}_{B^0}	282 ± 18	1220 ± 40
\mathcal{N}_{BKG}	163 ± 15	540 ± 30
μ_{B^0} [MeV/c ²]	5281.1 ± 0.6	5279.79 ± 0.31
λ_{exp} [c ² /MeV]	-0.0027 ± 0.0007	-0.0022 ± 0.0004
σ_{B^0} [MeV/c ²]	11.4 ± 0.8	10.85 ± 0.35
$\sigma_{B_s^0}$ [MeV/c ²]	7.1 ± 1.7	12.4 ± 2.4
$\Delta\mu_{B^0}$ [MeV/c ²]	87.26	87.26

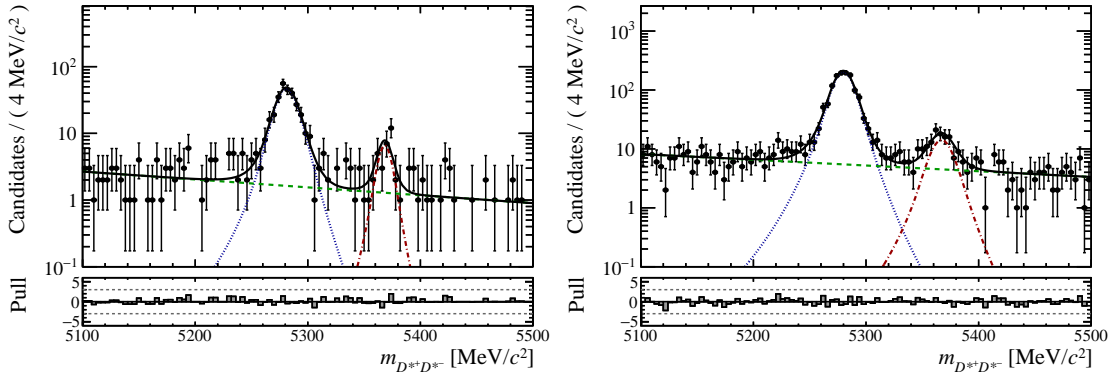


Figure 8.1 – Reconstructed invariant $D^{*+}D^{*-}$ mass shown together with the fitted PDF and pull distribution for (left) Run 1 and (right) Run 2 data with logarithmic scale. Besides the data points overlaid with the full PDF (solid black), the individual components are shown: the (dotted blue) B^0 decays, the (dot-dashed red) B_s^0 decays and the (dashed green) combinatorial background.

$\sum_i w_i / \sum_i w_i^2$ is applied to each weight. It is taken from Ref. [164] and ensures that the sensitivity of the fit is not underestimated. The approach shows a good coverage of the statistical uncertainty, which is further validated with bootstrapping studies [185] in the following.

8.2 Decay-time parametrisation

This section describes the modifications of the theoretical decay-time PDF given in Eq. (8.1) that are required to account for several experimental effects. The modified PDF

of the reconstructed decay-time t is defined as

$$P(t, \vec{d} | \vec{\eta}) = \varepsilon_{\text{sel}}(t) \cdot \left(\mathcal{P}(t', \vec{d} | \vec{\eta}) \otimes R(t - t') \right), \quad (8.2)$$

where the PDF $\mathcal{P}(t', \vec{d} | \vec{\eta})$ describes the distribution of the true decay time t' . Eq. (8.2) includes effects from external asymmetries like the production and detection asymmetries. Another effect arises from the FT, which experimentally enters the decay-time distribution with the tag decisions $\vec{d} = (d^{\text{OS}}, d^{\text{SS}})$ and the per-event mistag probability estimates $\vec{\eta} = (\eta^{\text{OS}}, \eta^{\text{SS}})$. A finite decay-time resolution $R(t - t')$ and a decay-time-dependent efficiency $\varepsilon_{\text{sel}}(t)$ are also considered.

8.2.1 External asymmetries

The detection and production asymmetries can be defined with the detection efficiencies ε_{det} and production cross sections σ_{prod} , respectively, *i.e.*

$$A_{\text{det}}^{\hat{f}} = \frac{\varepsilon_{\text{det}}(f) - \varepsilon_{\text{det}}(\bar{f})}{\varepsilon_{\text{det}}(f) + \varepsilon_{\text{det}}(\bar{f})}, \quad A_{\text{prod}} = \frac{\sigma_{\text{prod}}(\bar{B}^0) - \sigma_{\text{prod}}(B^0)}{\sigma_{\text{prod}}(\bar{B}^0) + \sigma_{\text{prod}}(B^0)}. \quad (8.3)$$

The detection asymmetry refers to asymmetries in the detection of particles but also includes asymmetries in the reconstruction of the particles and PID inefficiencies. It is impossible to measure the detection asymmetry directly since it is directly coupled to the intrinsic charge asymmetry. As a consequence, only the raw asymmetry

$$A_{\text{raw}}^{\hat{f}} = A_{\text{det}}^{\hat{f}} + A^{\hat{f}} \quad (8.4)$$

can be measured. However, the final state in the $B^0 \rightarrow D^{*+}D^{*-}$ analysis is symmetric. Accordingly, an asymmetry can only arise from different kinematic distributions of the final-state particles. These effects are expected to be small, and the detection asymmetry A_{det} is neglected in the decay-time fit. Including the asymmetries, the PDF given in Eq. (8.1) is extended to

$$\begin{aligned} \mathcal{P}_{\text{asym}}(t, q) \propto e^{-\Gamma t} (1 - qA_{\text{prod}}) (1 + A_{\text{raw}}^{\hat{f}}) \\ \cdot [1 - q(1 - R_{\perp})S_{D^{*+}D^{*-}} \sin(\Delta m_d t) + qC_{D^{*+}D^{*-}} \cos(\Delta m_d t)], \end{aligned} \quad (8.5)$$

The production asymmetry is not measured directly in this analysis. Instead, it is an external input. Due to the different centre-of-mass energy, different values need to be constrained for the two data-taking periods. Both values are taken from Ref. [24].

8.2.2 Flavour tagging

The measurement of the CP observables requires knowledge about the initial flavour of the B^0 meson. This information is provided by the FT, which is discussed in more

detail in Sec. 3.4. The FT algorithms have to be calibrated using a control channel since the centrally optimised training data for the FT can differ significantly from the data analysed in this thesis. This applies in particular to the mistag distribution, which can be distorted significantly. Consequently, the mistag of each tagger has to be calibrated using flavour-specific control modes that are topologically similar to the signal. The resulting calibration parameters have to be integrated into the decay-time PDF since the uncalibrated taggers are later used in the decay-time fit. In this analysis, the OS and SS combinations are used in the decay-time fit and no overall combination is determined. This allows to account for correlations among them. The inclusion of the FT parameters, the selection of the control channel and the calibration of the taggers are discussed in more detail in this section. All FT related results presented in the following are obtained with the FTALIB software package [151].

Using the tagging efficiencies ε_q and the respective calibrated mistag probabilities ω_q , which directly depend on the estimated mistag η , the probability to obtain the tag d for the true tag q of one single tagger can be defined as

$$\mathcal{P}_{\text{tag}}(d|q) = \begin{cases} \varepsilon_q(1 - \omega_q) & , \text{ if } d_i = q \\ \varepsilon_q\omega_q & , \text{ if } d_i = -q \\ 1 - \varepsilon_q & , \text{ if } d_i = 0 \end{cases} . \quad (8.6)$$

The tagging efficiencies for the B^0 (\bar{b}) and \bar{B}^0 (b) are defined as

$$\varepsilon_{\bar{b}} := \varepsilon = \varepsilon_{\text{tag}} + \frac{1}{2}\Delta\varepsilon_{\text{tag}}, \quad \varepsilon_b := \bar{\varepsilon} = \varepsilon_{\text{tag}} - \frac{1}{2}\Delta\varepsilon_{\text{tag}}, \quad (8.7)$$

respectively. The corresponding mistag probabilities are ω and $\bar{\omega}$. Multiple tag responses can be combined based on likelihoods derived from the mistag probabilities. The respective combinations of the OS and SS tagging information can both be included together in the description of the decay time given in Eq. (8.5) by defining the quantity Φ^\pm . The resulting modified PDF is given by

$$\begin{aligned} \mathcal{P}(t, d) &= \mathcal{P}_{\text{tag}}(d|B^0)\mathcal{P}_{\text{asym}}(t, B^0) + \mathcal{P}_{\text{tag}}(d|\bar{B}^0)\mathcal{P}_{\text{asym}}(t, \bar{B}^0) \\ &= e^{-\Gamma t}(1 + A_{\text{raw}}^{\hat{f}})[(\Phi^+ - A_{\text{prod}}\Phi^-) \\ &\quad - (\Phi^+ - A_{\text{prod}}\Phi^-)(1 - R_\perp)S_{D^{*+}D^{*-}} \sin(\Delta m_d t) \\ &\quad + (\Phi^+ - A_{\text{prod}}\Phi^-)C_{D^{*+}D^{*-}} \cos(\Delta m_d t)] \end{aligned} \quad (8.8)$$

In the case that both taggers provide a tag decision, ϕ^\pm is given by

$$\begin{aligned} \Phi^\pm &= \frac{1}{4}\varepsilon^{\text{OS}}[1 + d^{\text{OS}}(1 - 2\omega^{\text{OS}})]\varepsilon^{\text{SS}}[1 + d^{\text{SS}}(1 - 2\omega^{\text{SS}})] \\ &\quad \pm \frac{1}{4}\bar{\varepsilon}^{\text{OS}}[1 + d^{\text{OS}}(1 - 2\bar{\omega}^{\text{OS}})]\bar{\varepsilon}^{\text{SS}}[1 + d^{\text{SS}}(1 - 2\bar{\omega}^{\text{SS}})]. \end{aligned} \quad (8.9)$$

If only one tag decision is available, e.g. the OS tag, the observable is given by

$$\begin{aligned} \Phi^\pm &= \frac{1}{2}\varepsilon^{\text{OS}}[1 + d^{\text{OS}}(1 - 2\omega^{\text{OS}})](1 - \varepsilon^{\text{SS}}) \\ &\pm \frac{1}{2}\bar{\varepsilon}^{\text{OS}}[1 + d^{\text{OS}}(1 - 2\bar{\omega}^{\text{OS}})](1 - \bar{\varepsilon}^{\text{SS}}). \end{aligned} \quad (8.10)$$

By changing OS to SS, the same definition can be used to calculate the observable for the case in which only the SS tag is available. In the last case, where no tag is available, the observable is given by

$$\Phi^\pm = (1 - \varepsilon^{\text{OS}})(1 - \varepsilon^{\text{SS}}) \pm (1 - \bar{\varepsilon}^{\text{OS}})(1 - \bar{\varepsilon}^{\text{SS}}). \quad (8.11)$$

Since the FT directly enters the decay-time PDF, a precise calibration of the taggers is essential for accurately measuring the CP observables. This typically requires high statistics decay modes as control channels. For $B^0 \rightarrow D^{*+}D^{*-}$ decays, the obvious choice would be the utilisation of the $B^0 \rightarrow D^{*-}D_s^{*+}$ decay whose branching fraction is measured to be $\mathcal{B}(B^0 \rightarrow D^{*-}D_s^{*+}) \approx (1.77 \pm 0.14)\%$ [47]. The D_s^{*+} mesons strongly decay via $D_s^{*+} \rightarrow D_s^+\gamma$ with about $(93.5 \pm 0.7)\%$ [47]. Since the LHCb detector is not well suited to reconstruct photons, this decay is not used for the FT calibration in this thesis. Instead, the $B^0 \rightarrow D_s^+D^-$ decay channel is used to calibrate the mistag probabilities. The selection of this channel is mainly adopted from a CP violation analysis in $B^0 \rightarrow D^+D^-$ and $B_s^0 \rightarrow D_s^+D_s^-$ decays [26]. However, adjustments are required due to the differences of the analyses, which are explicitly mentioned in the short explanation of the selection discussed in the following.

$B^0 \rightarrow D_s^+D^-$ decays

The parts of the selection, which are directly taken from the $B_{(s)}^0 \rightarrow D_{(s)}^+D_{(s)}^-$ analysis [26], are summarised concisely. In contrast, small adjustments applied to the selection are discussed in more detail. These adjustments are mainly due to the different data sets used in the two analyses. The $B_{(s)}^0 \rightarrow D_{(s)}^+D_{(s)}^-$ analysis is performed using Run 2 data only, while this analysis performs a simultaneous Run 1 and Run 2 measurement. Additionally, an inclusive measurement of CP violation in $B^0 \rightarrow D^+D^-$ and $B_s^0 \rightarrow D_s^+D_s^-$ decays is performed in the $B_{(s)}^0 \rightarrow D_{(s)}^+D_{(s)}^-$ analysis, and a multivariate analysis is trained using both signal and control channels simultaneously. This is not necessary in this analysis, and only the $B^0 \rightarrow D_s^+D^-$ channel is used. To account for these differences, the whole selection of the $B^0 \rightarrow D_s^+D^-$ decays is implemented in the framework of the $B^0 \rightarrow D^{*+}D^{*-}$ analysis to select $B^0 \rightarrow D_s^+D^-$ Run 1 and Run 2 data.

The final states considered in the selection of the $B^0 \rightarrow D_s^+D^-$ decays are summarised in Table 8.3. As described in Sec. 3.3.2, a centralised preselection is applied, where a Stripping line is used that contains the $B^0 \rightarrow D_s^+D^-$ decays. Afterwards, a loose preselection is applied with requirements on the mass window of the $D_{(s)}^+$ mesons, the χ_{FD}^2 of both $D_{(s)}^+$ mesons and the status of the DTF fit, where the DTF fit uses constraints

on the PV or the masses of the intermediate $D_{(s)}^+$ mesons. The χ^2 requirement on the flight distance (FD) ensures that the decay vertex of the $D_{(s)}^+$ mesons is displaced from the SV. The misID of final-state particles leads to backgrounds in the data possibly leaking into the signal region. These can be found by assigning different mass hypotheses to the final-state particles and recalculating the invariant mass. Most of the backgrounds are suppressed by applying requirements on the recalculated invariant masses and variables indicating the probability of a particle being a kaon or a pion. These requirements are referred to as vetoes. The detailed requirements are not listed in this thesis but can be found in Ref. [26].

Table 8.3 – Final states considered for the $B^0 \rightarrow D_s^+ D^-$ channel. The final states are ordered according to their branching fraction. The first parenthesis indicates the final state of the D_s^+ meson, and the second is the final state of the D^- meson.

Decay	Final state
$B^0 \rightarrow D_s^+ D^-$	$(K^- K^+ \pi^+)(K^+ \pi^- \pi^-)$
	$(\pi^- K^+ \pi^+)(K^+ \pi^- \pi^-)$
	$(\pi^- \pi^+ \pi^+)(K^+ \pi^- \pi^-)$
	$(K^- K^+ \pi^+)(K^+ K^- \pi^-)$

One background which arises from the misID between the $D^+ \rightarrow K^- \pi^+ \pi^+$ and $D_s^+ \rightarrow K^- K^+ \pi^+$ decays cannot be rejected by the cut-based vetoes. This background is more challenging because of the similar mass shift between π^+ and K^+ compared to D^+ and D_s^+ mesons. For both decays, the signal component largely overlaps with the misID background in the invariant mass distribution of the respective $D_{(s)}^+$ meson. This makes a cut-based veto requirement very inefficient, and a multivariate analysis is used to separate the two decays. The `xgboost` [169] module from the `scikit-learn` package [186] is used to train a GBDT [155, 187] that can distinguish between the two decays. Simulated $D_s^+ \rightarrow K^- K^+ \pi^+$ and $D^+ \rightarrow K^- \pi^+ \pi^+$ meson decays from the $B^0 \rightarrow D_s^+ D^-$ control channel are used for the training. All variables and hyperparameters used in the training and the cuts applied to the output of the GBDT are taken from the $B_{(s)}^0 \rightarrow D_{(s)}^+ D_{(s)}^-$ analysis. The same applies to the training on Run 1 data, where the variables, hyperparameters, and cuts are not further optimised but are adopted from the Run 2 training. Both trainings are performed separately. Additionally, the procedure is applied for each D meson separately. More details about the procedure of the training and requirements on the GBDT output can be found in Ref. [26].

After the vetoes, the data still contains many combinatorial background candidates. These are also suppressed by a GBDT following the strategy from the $B_{(s)}^0 \rightarrow D_{(s)}^+ D_{(s)}^-$ analysis, where multiple signal and control channels are simultaneously used as a signal proxy. In contrast, only simulated $B^0 \rightarrow D_s^+ D^-$ decays are used in this analysis. The upper-mass sideband of the data corresponding to $m_{D_s^+ D^-} > 5600 \text{ MeV}/c^2$ is used as a background proxy. Due to the different signal proxies, the GBDT is retrained in this

analysis using the same variables, hyperparameters and procedure as in the $B_{(s)}^0 \rightarrow D_{(s)}^+ D_{(s)}^-$ analysis. All selection steps except the vetoes are applied before the training of the GBDT. Since a different signal proxy is used in this analysis, new optimal cuts on the GBDT output are determined. A Figure of Merit (FoM) defined as

$$\text{FoM} = \frac{S}{\sqrt{S+B}} \quad (8.12)$$

is used to maximise the signal significance in the data samples, where S and B are the numbers of $B^0 \rightarrow D_s^+ D^-$ and background candidates, respectively. The FoM is calculated for different requirements of the BDT output. The number of signal and background candidates for each requirement is determined and calculated by performing a fit to the invariant $D_s^+ D^-$ mass. These fits are performed using the same procedure and fit model described later in this section. The final cut value is the requirement that maximises the FoM. The procedure is performed separately for all individual $B^0 \rightarrow D_s^+ D^-$ final states and for Run 1 and Run 2 data, respectively. In Table 8.4, the resulting cut values are listed. The FoM for the most dominant final state $(K^- K^+ \pi^+)(K^+ \pi^- \pi^-)$ and for Run 2 data is shown in Fig. 8.2. The distributions of the other final states and for Run 1 data are not explicitly presented due to redundancy. Despite the maximum occurring for a different requirement on the GBDT output, they generally show the same trend. As a last selection step, the multiple candidates are rejected randomly.

Table 8.4 – Requirements on the GBDT output that maximise the signal significance and the corresponding number of candidates extracted with the final extended maximum-likelihood fit to the invariant $D_s^+ D^-$ mass for the different $B^0 \rightarrow D_s^+ D^-$ final states and for Run 1 and Run 2 data.

Final state	BDT requirement		Signal candidates	
	Run 1	Run 2	Run 2	Run 1
$(K^- K^+ \pi^+)(K^+ \pi^- \pi^-)$	> 0.6	> 0.6	22080 ± 170	93090 ± 320
$(\pi^- K^+ \pi^+)(K^+ \pi^- \pi^-)$	> 0.9	> 0.93	2130 ± 60	10400 ± 110
$(\pi^- \pi^+ \pi^+)(K^+ \pi^- \pi^-)$	> 0.92	> 0.9	4920 ± 80	23100 ± 170
$(K^- K^+ \pi^+)(K^+ K^- \pi^-)$	> 0.96	> 0.95	1990 ± 50	9570 ± 110

Finally, the mass fits to extract the $B^0 \rightarrow D_s^+ D^-$ candidates are performed in the same procedure as for the $B^0 \rightarrow D^{*+} D^{*-}$ channel described in Sec. 8.1. The model used to describe the invariant $D_s^+ D^-$ mass distribution consists of two Hypatia functions to describe the $B^0 \rightarrow D_s^+ D^-$ and $B_s^0 \rightarrow D_s^- D^+$ contributions and an exponential function to describe the combinatorial background. Simulation is fitted to determine the shape of the $B^0 \rightarrow D_s^+ D^-$ component. Since no $B_s^0 \rightarrow D_s^- D^+$ simulation is available in this analysis, it is impossible to determine the respective shape from the simulation. However, the $B_s^0 \rightarrow D_s^- D^+$ contribution can be described with the same shape as the $B^0 \rightarrow D_s^+ D^-$ contribution since both decays share the same final state. In the fit to data,

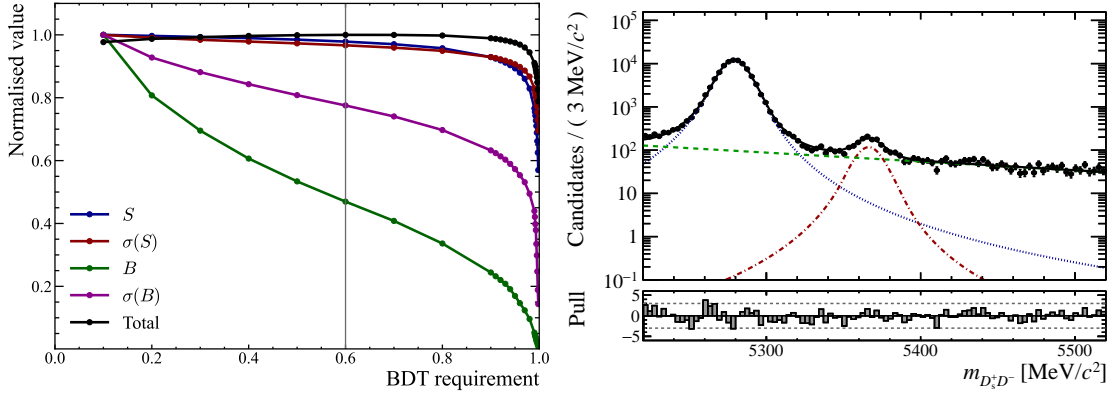


Figure 8.2 – The (left) FoM as a function of the BDT output and (right) the result of the final extended maximum-likelihood fit to the invariant $D_s^+ D^-$ mass for the most dominant $B^0 \rightarrow D_s^+ D^-$ final state, *i.e.* the $(K^- K^+ \pi^+)(K^+ \pi^- \pi^-)$ final state. The result of the mass fit shows the invariant $D_s^+ D^-$ mass distribution together with the (solid black) fitted PDF, pull distribution and the individual contributions, *i.e.* the (dotted blue) B^0 signal, the (dot-dashed red) B_s^0 signal and the (dashed green) combinatorial background.

the $B^0 \rightarrow D_s^+ D^-$ contribution is fixed to the shape determined with simulation. The fit is performed separately for all $B^0 \rightarrow D_s^+ D^-$ final states and for Run 1 and Run 2 data. The number of signal decays resulting from the fit to the invariant $D_s^+ D^-$ mass are given in Table 8.4. The result for the most dominant final state and Run 2 data is shown in Fig. 8.2. Finally, the *sWeights* are determined using the *sPlot* method following the same procedure described for the signal channel in Sec. 8.1.

Calibration results

The $B^0 \rightarrow D_s^+ D^-$ decay offers high statistics but is topologically different compared to the $B^0 \rightarrow D^{*+} D^{*-}$ decay since no intermediate $D^{*\pm}$ mesons are present. To account for that, the $B^0 \rightarrow D_s^+ D^-$ data is reweighted to match the $B^0 \rightarrow D^{*+} D^{*-}$ data in the distributions of a subset of variables. These variables are the transverse momentum of the B^0 meson, the pseudorapidity, the number of tracks and the number of PVs. They are chosen since it is known that they are related to the tagging characteristics. As discussed in Sec. 4.1, a GBreweighter is trained to reweight the data.

The calibration is performed for Run 1 and Run 2 data individually but simultaneously for all considered $B^0 \rightarrow D_s^+ D^-$ final states. No implications are expected from this approach as the event kinematics are similar. Additionally, only minor differences from the different final-state particles are expected. The calibration is applied simultaneously to all considered $B^0 \rightarrow D^{*+} D^{*-}$ final states. All available taggers are used in the calibration. However, as discussed in Sec. 3.4, the *OSCharm* tagger is not available in Run 1 and is therefore only considered in the OS combination of Run 2. The single taggers are calibrated first and then combined to obtain the OS and SS combinations.

Afterwards, the combinations are calibrated again to correctly account for correlations. They are combined in the decay-time PDF given in Eq. (8.8). This allows for constraining the calibration parameters of the combinations and for correctly accounting for the uncertainties and correlations among the two combinations in the fit. Consequently, the uncertainties propagate into the statistical uncertainty on the extracted CP observables. In each calibration step, $sWeights$ are applied to obtain the calibration function. The results of the calibrations are given in Table 8.5 and Table 8.6. The calibrations of the OS combination and the SS combination are presented in Fig. 8.3 for Run 2 data. The Run 1 results offer fewer statistics but look similar and are not shown explicitly.

Table 8.5 – Results of the FT calibration for the $B^0 \rightarrow D_s^+ D^-$ decay for Run 1.

Tagger	$p_0 - \langle \eta \rangle$	p_1	Δp_0	Δp_1	$\langle \eta \rangle$
<i>SSProton</i>	0.030 ± 0.015	1.18 ± 0.28	0.031 ± 0.021	-0.0 ± 0.4	0.445
<i>SSPion</i>	-0.021 ± 0.011	1.14 ± 0.18	0.000 ± 0.016	0.45 ± 0.26	0.431
<i>OSMuon</i>	0.049 ± 0.025	1.19 ± 0.26	0.01 ± 0.04	0.3 ± 0.4	0.303
<i>OSElectron</i>	0.10 ± 0.04	1.2 ± 0.7	0.00 ± 0.06	1.0 ± 1.0	0.289
<i>OSVtxCh</i>	0.022 ± 0.016	0.99 ± 0.25	-0.028 ± 0.023	-0.38 ± 0.36	0.381
<i>OSKaon</i>	-0.031 ± 0.018	1.23 ± 0.30	0.029 ± 0.025	-0.6 ± 0.4	0.393
SS combination	-0.001 ± 0.010	0.98 ± 0.14	0.005 ± 0.015	0.24 ± 0.20	0.41
OS combination	0.044 ± 0.013	1.01 ± 0.12	0.003 ± 0.018	-0.19 ± 0.17	0.346

Table 8.6 – Results of the FT calibration for the $B^0 \rightarrow D_s^+ D^-$ decay for Run 2.

Tagger	$p_0 - \langle \eta \rangle$	p_1	Δp_0	Δp_1	$\langle \eta \rangle$
<i>SSProton</i>	-0.005 ± 0.006	1.00 ± 0.11	0.027 ± 0.008	0.05 ± 0.16	0.451
<i>SSPion</i>	0.000 ± 0.004	1.09 ± 0.06	-0.003 ± 0.005	-0.05 ± 0.09	0.432
<i>OSMuon</i>	-0.001 ± 0.010	1.55 ± 0.15	-0.011 ± 0.014	-0.28 ± 0.22	0.329
<i>OSElectron</i>	0.003 ± 0.015	1.21 ± 0.21	0.024 ± 0.021	0.05 ± 0.29	0.355
<i>OSVtxCh</i>	-0.010 ± 0.007	1.20 ± 0.11	0.024 ± 0.010	0.04 ± 0.16	0.384
<i>OSKaon</i>	-0.015 ± 0.007	1.49 ± 0.16	0.016 ± 0.010	0.16 ± 0.23	0.4
<i>OSCharm</i>	-0.001 ± 0.014	0.87 ± 0.27	0.001 ± 0.020	0.2 ± 0.4	0.365
SS combination	-0.001 ± 0.004	0.95 ± 0.05	0.003 ± 0.005	-0.06 ± 0.07	0.425
OS combination	0.035 ± 0.005	0.80 ± 0.04	0.015 ± 0.007	0.04 ± 0.06	0.334

The uncalibrated SS and OS combinations are later used in the decay-time fit, and the calibration parameters are Gaussian constrained to the values obtained in the calibration procedure. However, to estimate the tagging performance, a full SS and OS combination is used, which results in an expected tagging power of $\epsilon_{\text{eff, Run 1}} = (6.3 \pm 1.0)\%$ and $\epsilon_{\text{eff, Run 2}} = (7.0 \pm 0.5)\%$ for Run 1 and Run 2, respectively.

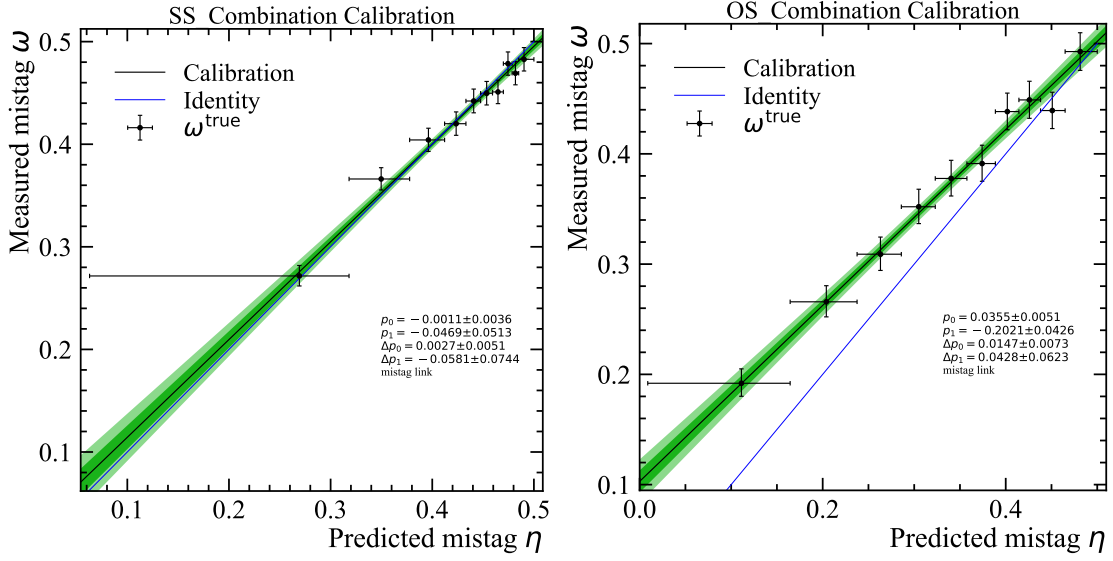


Figure 8.3 – Calibration of the (left) SS and (right) OS combination for the $B^0 \rightarrow D_s^+ D^-$ in Run 2 data. The Run 1 calibration results look similar but offer fewer statistics. Hence, they are not explicitly shown.

8.2.3 Decay-time resolution

The decay-time PDF has to account for the resolution effects of the detector. This is because the reconstructed decay times t are smeared out due to the finite precision of the measurements of the momenta and vertex positions. This leads to a reduction of the amplitudes of the $B^0 - \bar{B}^0$ oscillation and, therefore, to a reduction of the sensitivity to the CP observables. These effects are considered in the decay-time PDF by convolving the CP violation physics PDF $\mathcal{P}(t', \vec{d} | \vec{\eta})$ with a decay-time resolution function $R(t - t')$, where t' is the true decay time. The effective detector resolution of the decays analysed in this thesis is roughly 60 fs. This resolution is much smaller than the oscillation frequency Δm_d of the B^0 meson, which is $(0.5065 \pm 0.0019) \text{ ps}^{-1}$ [47]. Consequently, the effects of the decay-time resolution are limited in this analysis.

The model used to describe the resolution effects consists of the sum of three Gaussian functions with a shared mean μ and individual widths σ_i . One of the fractions f_i between the Gaussians can be eliminated because their sum is required to be equal to one, *i.e.* $\sum_i f_i = 1$. The overall resolution model R is defined as

$$R(t - t') = \sum_{i=1}^3 f_i \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(t-t'-\mu)^2}{2\sigma_i^2}}. \quad (8.13)$$

Since both the reconstructed decay time t and true decay time t' are known in simulation, it is exploited to determine the parameters of the model. The simulation only contains the most dominant D^0 decay mode considered in this thesis, *i.e.* the $D^0 \rightarrow K^- \pi^+$ decay.

However, the resolution is assumed to be similar for all considered D^0 decay modes. Separate maximum-likelihood fits to the $t-t'$ distributions of Run 1 and Run 2 simulation are performed. The resulting parameters are given in Table 8.7 and the resulting PDF is shown in Fig. 8.4.

Table 8.7 – Results of the fit to the $t-t'$ distributions of simulated $B^0 \rightarrow D^{*+}D^{*-}$ decays to determine the decay-time resolution model. The fits are performed for Run 1 and Run 2 separately.

Parameter	Run 1	Run 2
μ [ps]	0.00166 ± 0.00017	0.00124 ± 0.00017
σ_1 [ps]	0.106 ± 0.004	0.114 ± 0.006
σ_2 [ps]	0.0562 ± 0.0019	0.0599 ± 0.0026
σ_3 [ps]	0.0319 ± 0.0016	0.0358 ± 0.0012
f_1	0.105 ± 0.017	0.079 ± 0.017
f_2	0.62 ± 0.05	0.49 ± 0.04

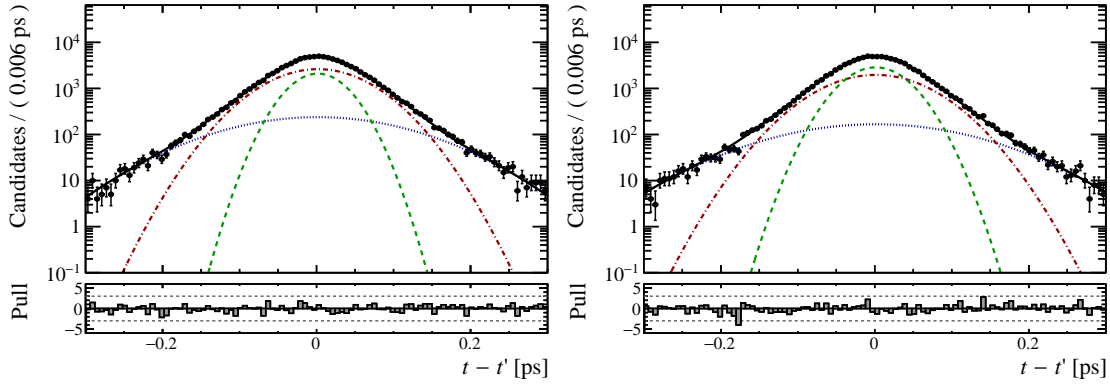


Figure 8.4 – The $t-t'$ distributions of simulated $B^0 \rightarrow D^{*+}D^{*-}$ decays overlaid with the PDF resulting from the fit to determine the decay-time resolution model with a logarithmic y-axis scale for (left) Run 1 and (right) Run 2. The pull distributions are also shown.

The effective resolution of the model can be quantified as

$$\sigma_{\text{eff}} = \sqrt{(-2/\Delta m_d^2) \ln(Q)}, \quad (8.14)$$

where Q is defined as the decay-time dilution and given by

$$Q = \sum_{i=1}^3 f_i e^{-\frac{\sigma_i^2 \Delta m_d^2}{2}}. \quad (8.15)$$

Using the the fit results given in Table 8.7, the effective resolutions for Run 1 and Run 2 are calculated to be (58.5 ± 2.0) fs and (57.9 ± 2.5) fs, respectively. These values are compatible within their uncertainties. The parameters of the decay-time resolution model are fixed in the fits to the decay-time distribution to extract the CP observables.

8.2.4 Decay-time-dependent efficiency

The reconstruction and selection lead to a decay-time-dependent efficiency. This is introduced in the PDF given in Eq. (8.2) by the quantity $\varepsilon_{\text{sel}}(t)$. The most significant effect is the rejection of low decay times in the selection, which allows the suppression of most of the backgrounds produced directly in the PV. For example, a dedicated requirement on the lifetime of the B^0 meson is applied in the Stripping with $t > 0.2$ ps. The Stripping is discussed in more detail in Sec. 3.3.2, and the requirements are given in Table 6.2.

The theoretical decay-time distribution should follow an exponential behaviour. However, the decay-time-dependent selection efficiency causes a deviation from the expected behaviour by suppressing particularly low decay times. In this analysis, the model to parametrise these effects is generally given by a sum of n cubic B-splines, *i.e.*

$$\varepsilon_{\text{sel}}(t) = \sum_{i=1}^n c_i b_{i,3}(t, \vec{t}), \quad (8.16)$$

with the spline coefficients c_i and a set of knot positions \vec{t} , which is fixed in the following. The B-splines are recursively defined as

$$b_{i,k+1}(t, \vec{t}) = \frac{t - t_i}{t_{i+k} - t_i} b_{i,k}(t, \vec{t}) + \left(1 - \frac{t - t_{i+1}}{t_{i+k+1} - t_{i+1}}\right) b_{i+1,k}(t, \vec{t}) \quad (8.17)$$

and

$$b_{i,1}(t, \vec{t}) = \begin{cases} 1 & , \text{if } t_i \leq t < t_{i+1} \\ 0 & , \text{otherwise} \end{cases} . \quad (8.18)$$

The model to describe the decay-time-dependent efficiency is determined by replacing the CP violation physics PDF $\mathcal{P}(t', \vec{d} | \vec{\eta})$ in Eq. (8.2) with a pure exponential function that is convolved with the decay-time resolution function. The pure exponential function decreases according to the known B^0 lifetime [47]. The number of knots and their positions are varied until an adequate description of the acceptance is achieved. By performing a maximum-likelihood fit to the decay-time distribution of simulation comprising the $D^0 \rightarrow K^- \pi^+$ decay mode, the coefficients of the model are determined. In this fit, the knots are fixed. The knots that describe the decay-time distribution sufficiently are placed at 0.3, 0.5, 0.9, 1.6, 2.7, 4.5, 6.8 and 10.3 ps, where the first and the last knot are placed at the lower and upper boundary of the fit range of the decay-time fit, respectively. More knots are placed at lower decay times since the data density is higher, and the change in the efficiency is more significant in this region. The resulting coefficients are

given in Table 8.8 and the resulting PDF is shown in Fig. 8.5. The model is determined using simulation only comprising the $D^0 \rightarrow K^- \pi^+$ decay mode, but the decay-time-dependent efficiency is assumed to be similar for all D^0 decay modes considered in the analysis. Separate coefficients are determined for Run 1 and Run 2, but the same knots

Table 8.8 – Results of the fit to the decay-time distribution of simulated $B^0 \rightarrow D^{*+} D^{*-}$ decays to determine the decay-time acceptance model. The fits are performed for Run 1 and Run 2 separately. The values listed without uncertainties are fixed in the fit.

Parameter	Run 1	Run 2
c_0^S	0.0056 ± 0.0016	0.00010 ± 0.00018
c_1^S	0.149 ± 0.022	0.0678 ± 0.0032
c_2^S	0.77 ± 0.11	0.628 ± 0.018
c_3^S	1.43 ± 0.21	1.331 ± 0.034
c_4^S	1.87 ± 0.26	1.85 ± 0.05
c_5^S	2.0 ± 0.4	2.00 ± 0.17
c_6^S	1.74 ± 0.26	1.90 ± 0.07
c_7^S	1.40 ± 0.21	1.58 ± 0.06
μ [ps]	0.0	0.0
σ_{eff} [ps]	0.059	0.058
τ [ps]	1.519	1.519

are used. In the final decay-time fit, the coefficients and the knots of the splines are fixed as the correlation with the CP observables is negligible. Additionally, not enough statistics are available in the $B^0 \rightarrow D^{*+} D^{*-}$ decay channel to let them float.

8.3 Extraction of the CP observables

This section describes the extraction of the CP observables from the decay-time distribution of the $B^0 \rightarrow D^{*+} D^{*-}$ candidates. The construction of the decay-time PDF $P(t, \vec{d} | \vec{\eta})$ is discussed in more detail in Sec. 8.2. It is defined in Eq. (8.2). However, the $B^0 \rightarrow D^{*+} D^{*-}$ decay is a $P \rightarrow VV$ decay, and the decay-time distribution depends on the fraction of the CP-odd eigenstates R_{\perp} , as the CP parameter $S_{D^{*+} D^{*-}}$ is diluted by the factor $(1 - 2R_{\perp})$. This is described in more detail in Sec. 2.5.3. Instead of using the CP-odd fraction as an external input, this analysis performs a fit to the angular distributions in the transversity basis to determine the CP-odd fraction simultaneously with the CP observables. The transversity basis is explained in more detail in Sec. 2.5.3. Since the measurement of the fraction is expected to be competitive with the current world average, this reduces the systematic uncertainties on the CP observables. Additionally, correlations between

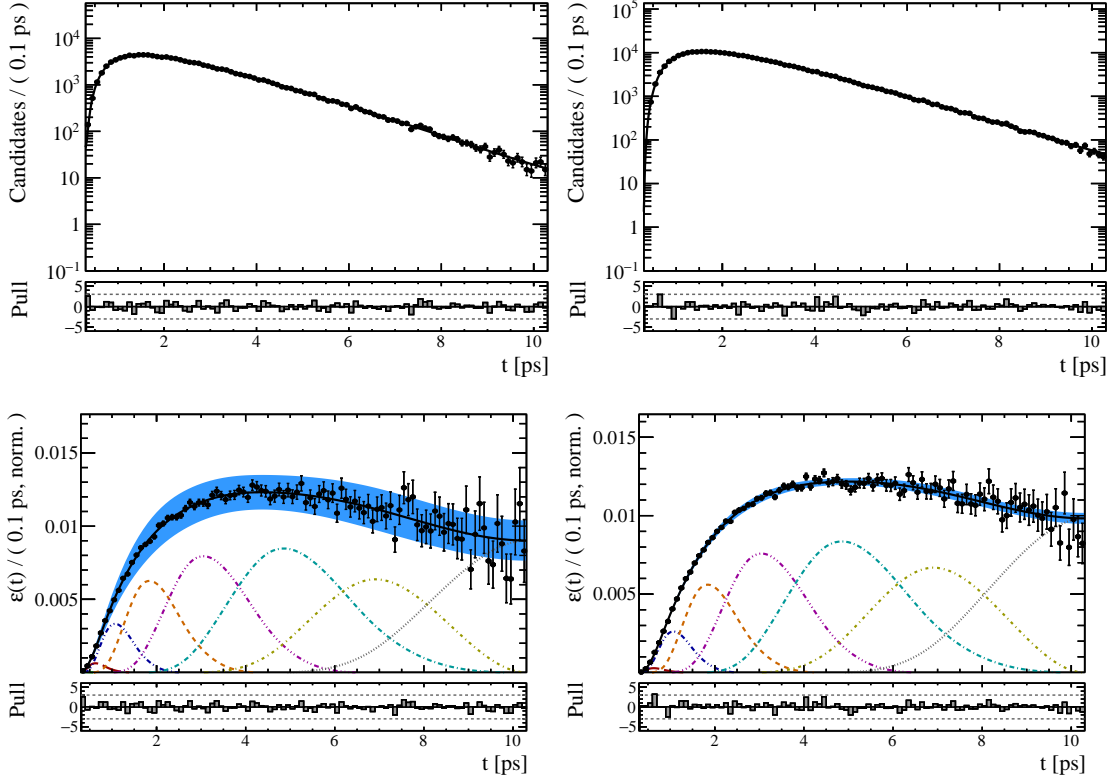


Figure 8.5 – Lifetime distribution of simulated $B^0 \rightarrow D^{*+}D^{*-}$ decays shown together with the fit result of the acceptance PDF and the pull distribution for (left) Run 1 and (right) Run 2. The (top) result is shown with a logarithmic scale, where the PDF is multiplied with an exponential function. Additionally, the (bottom) resulting pure acceptance PDF is presented, where the lifetime distribution is divided by an exponential function decreasing according to the known B^0 lifetime [47]. The 1σ error band of the acceptance PDF is indicated by the blue band, and the individual splines are shown in several colours.

the CP -odd fraction and the CP observables are directly considered in the fit. This is achieved by multiplying the decay-time PDF given in Eq. (8.2) with the theoretical two-dimensional angular decay rate given in Eq. (2.63). The resulting fit model is given by

$$P_{\text{fit}}(t, \cos(\theta_{D^{*-}}), \cos(\theta_{\text{tr}}), \vec{d} | \vec{\eta}) = \frac{d^2 \Gamma(R_0, R_{\perp})}{d\cos(\theta_{D^{*-}})d\cos(\theta_{\text{tr}})} \cdot P(t, \vec{d} | \vec{\eta}). \quad (8.19)$$

However, adding the angular dependency in the CP fit introduces another experimental effect that must be considered in the PDF. This effect is called the angular acceptance and is discussed in more detail in Sec. 8.3.1. The fit framework is extensively tested on simulation before it is applied to recorded data. In both cases, weighted data is fitted to account for the effects of the angular acceptance. A simultaneous maximum-likelihood

fit of the Run 1 and Run 2 data sets is performed to determine the angular and CP observables. In this fit, the full angular phase space is considered while the decay time is restricted to the range of 0.3 ps to 10.3 ps. The lower boundary is chosen because of the low number of B^0 mesons contributing below the threshold of 0.3 ps, which is due to the inefficient selection at lower decay times. The upper boundary is about 6 times the average B^0 lifetime to ensure that the decay-time distribution is fully described. Also above the threshold of 10.3 ps, the number of B^0 candidates is very low. The extraction of the CP observables is performed in Sec. 8.3.2, where also the external inputs and inputs determined in previous analysis steps are summarised. Additionally, the complete fit framework is validated. Lastly, the angular decay-time fit to background-subtracted data is performed.

8.3.1 Angular acceptance

The angle-dependent decay-time PDF in Eq. (8.19) only considers the theoretical distribution in the true angular phase space. However, due to several experimental effects, the reconstructed angular distributions in data are distorted compared to the true distributions. These effects have to be considered to measure the correct CP -odd fraction and, consequently, the correct CP observables. They can arise from the detector acceptance, the detector resolution and the selection, all explained in more detail in the following. First, the detector acceptance, which is defined by the geometry of the detector, leads to certain regions of the angular phase space being excluded. Second, the momenta and, consequently, the angles are measured with finite precision due to the detector resolution and are smeared out. This causes events being migrated from their true location in phase space to another point. Third, the application of selection requirements can affect the kinematic distributions of the selected decays. This causes certain regions of the angular phase to be removed preferentially. To account for these effects, the angular distributions of fully selected simulation and generator-level simulation are compared. The latter is generated applying only the EVTGEN step without any detector simulation or further steps described in Sec. 3.3.3. Both simulations are generated using the same EVTGEN model and parameters. Consequently, the differences in the distributions are purely due to the acceptance effects.

The angular distributions in the transversity basis, *i.e.* the $\cos(\theta_{D^{*-}})$ and $\cos(\theta_{\text{tr}})$ distributions, of full and generator-level simulation are used to train a GBreweighter [152] included in the HEP-ML software package [156]. The approach is discussed in more detail in Sec. 4.1. A comparison between the generator-level and full simulation distributions and the reweighted full simulation distributions is presented in Fig. 8.6. After training the reweighter with generator-level and full simulation, it is applied to data. The reconstructed and reweighted angular distributions for recorded data are also shown in Fig. 8.6. Only the Run 2 distributions are shown for both comparisons, but the Run 1 distributions look similar and show the same trends. The determined angular weights are normalised and by multiplication combined with the $sWeights$, which are determined

with the *sPlot* method as described in Sec. 8.1.2. Lastly, the combination is normalised to the sum of the *sWeights*. The procedure is performed for Run 1 and Run 2 separately.

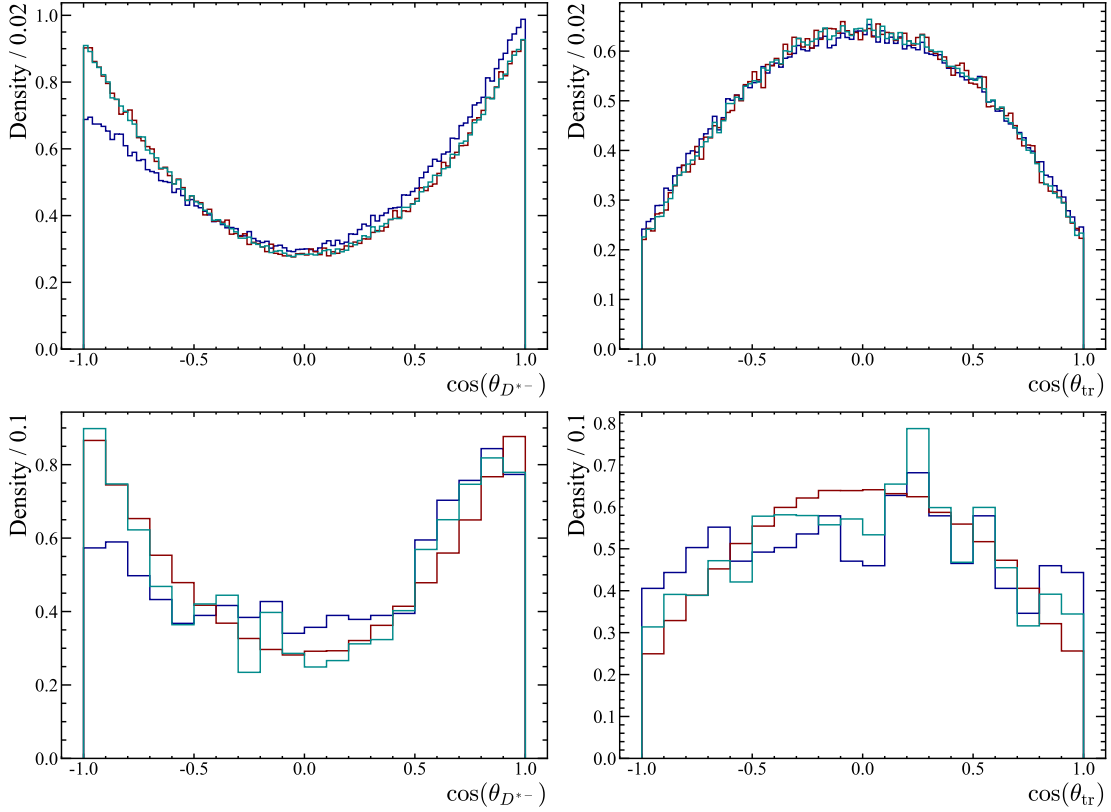


Figure 8.6 – The (left) $\cos(\theta_{D^{*-}})$ and (right) $\cos(\theta_{tr})$ distributions used to reweight the (top) simulation and (bottom) recorded data for Run 2, respectively. For simulation, the (red) generator-level, (blue) full simulation and (cyan) reweighted full simulation distributions are shown. For recorded data, the (red) generator-level simulation, the (blue) data and the (cyan) reweighted data distributions are presented. The data distributions are reweighted using the reweighter determined on simulation.

8.3.2 Angular decay-time fit

In the angular decay-time fit, not only external inputs are used, but also inputs determined in previous analysis steps. The parameters of the decay-time resolution and acceptance models are fixed to the values determined in Sec. 8.2.3 and Sec. 8.2.4, respectively. The reason to fix the coefficients of the acceptance model is that the statistics in the $B^0 \rightarrow D^{*+}D^{*-}$ decay channel are not sufficient to let them float. Effects from this choice on the CP parameters and CP -odd fraction are evaluated in Sec. 8.4. The measured value of $\Delta\Gamma_d$ is compatible with 0 [47]. Consequently, the parameter is fixed to this value

in the fit. As discussed in Sec. 8.2, the detection asymmetry A_{det} is neglected in this analysis and also fixed to zero.

Gaussian constraints are used in the fit to account for the uncertainties of other external input parameters. The technique is discussed in more detail in Sec. 4.2. The world averages and their uncertainties are used for the external parameters as the B^0 lifetime τ_{B^0} and the B^0 oscillation frequency Δm_d . These values are taken from Ref. [47]. The production asymmetry is taken from a CP violation measurement in $B^0 \rightarrow D^{*\pm} D^\mp$ decays [24]. The kinematics are very similar between these decays, and no correction is required. Also, the FT calibration parameters are Gaussian constrained in the fit. As discussed in Sec. 8.2, the uncalibrated mistag distributions of the combined SS and OS taggers are used in the fit, and each calibration parameter is constrained in the fit using the results from Table 8.5 and Table 8.6. Constraining the FT parameters and using the uncalibrated mistag distributions in the fit allows to consider their correlations. All input parameters of the fit that are not given in Tables 8.5 to 8.8 are summarised in Table 8.9.

Table 8.9 – Input parameters used in the angular decay-time fit. The parameters listed without uncertainty are fully fixed, while the parameters given with uncertainty are Gaussian constrained in the fit. The values and uncertainties are taken from Ref. [47] for $\Delta\Gamma_d$ and Δm_d and from Ref. [24] for A_{prod} .

Parameter	Run 1	Run 2
Δm_d [ps ⁻¹]	0.5065 ± 0.0019	0.5065 ± 0.0019
τ_{B^0} [ps]	1.519 ± 0.004	1.519 ± 0.004
A_{prod}	-0.011 ± 0.009	0.004 ± 0.005
A_{det}	0	0
$\Delta\Gamma_d$ [ps ⁻¹]	0	0

Fit to simulation

The PDF presented in Eq. (8.19) is used to extract the CP observables with a fit to the $\cos(\theta_{D^{*-}})$, $\cos(\theta_{\text{tr}})$ and decay-time distributions. It is a multi-dimensional function and depends on many input parameters and models. Due to its complexity, the model is extensively tested before applying the fit procedure to the data. These tests are performed in this section using maximum-likelihood fits to the identical distributions of simulation, where the true generated values of the angular and CP observables are known. The values used in the generation are summarised in Table 8.10. The model used in the generation requires the helicity amplitudes and the weak phase β . The helicity amplitudes needed in the generation are calculated using the latest Belle measurement [92], while the weak phase is taken from Ref. [188] corresponding to $S_{D^{*+}D^{*-}} = -0.71$ and $C_{D^{*+}D^{*-}} = 0$.

Table 8.10 – Input parameters used in the generation of the simulation. The parameter Δm_d is taken from Ref. [47] while the parameter β is taken from Ref. [188] and R_\perp and R_0 are calculated with the results presented in Ref. [92].

Parameter	Combined
Δm_d [ps ⁻¹]	0.5065
β	0.387
R_\perp	0.138
R_0	0.624

In general, the simulation is processed the same way as the recorded data, *i.e.* the same selection criteria are applied. Additionally, the same FT strategy is applied using the same taggers, the same calibration functions and $B^0 \rightarrow D_s^+ D^-$ simulation. Since the simulation only contains true signal decays, *sWeights* are neglected, and only angular weights are considered in the fits. These weights are determined by applying the reweighter trained in Sec. 8.3.1 to the simulation itself. The fits also use the same input parameters as the final fit on recorded data.

Multiple fits are performed to validate the angular and time-dependent fit model defined in Eq. (8.19). First, a fit to the weighted angular distributions is performed only using the theoretical angular decay rate given in Eq. (2.63). Afterwards, an individual fit to the decay-time distribution is performed using the model given in Eq. (8.2). Finally, a combined fit to the weighted angular and decay-time distributions is performed. This fit uses the final model defined in Eq. (8.19). In each study, a simultaneous fit to the Run 1 and Run 2 distributions of simulation comprising only the $D^0 \rightarrow K^- \pi^+$ decay mode is performed. However, for the validation of the model, it is sufficient to use simulation, which only comprises the most dominant D^0 decay mode. In the simultaneous fits, the shared parameter between the Run 1 and Run 2 depend on the study and are given by the angular observables R_0 and R_\perp , the CP observables S_f and C_f or both the angular and CP observables. The fit results of the main parameters of interest are given in Table 8.11 for each study. The Gaussian-constrained parameters are not listed but are compatible with their initial values. Overall, within the uncertainties, the fit results are compatible with the true generated values given in Table 8.10. The fit result of the combined angular and decay-time fit is shown in Fig. 8.7.

Fit to data

The angular decay-time fit is performed to a combined data set comprising all considered D^0 decay modes. A simultaneous fit to weighted Run 1 and Run 2 data sets is performed, where the weights are a combination of the *sWeights* extracted in Sec. 8.1.2 and the angular weights determined in Sec. 8.3.1. The angular weights account for the acceptance effects in the angular phase space. The combination of both weights is corrected

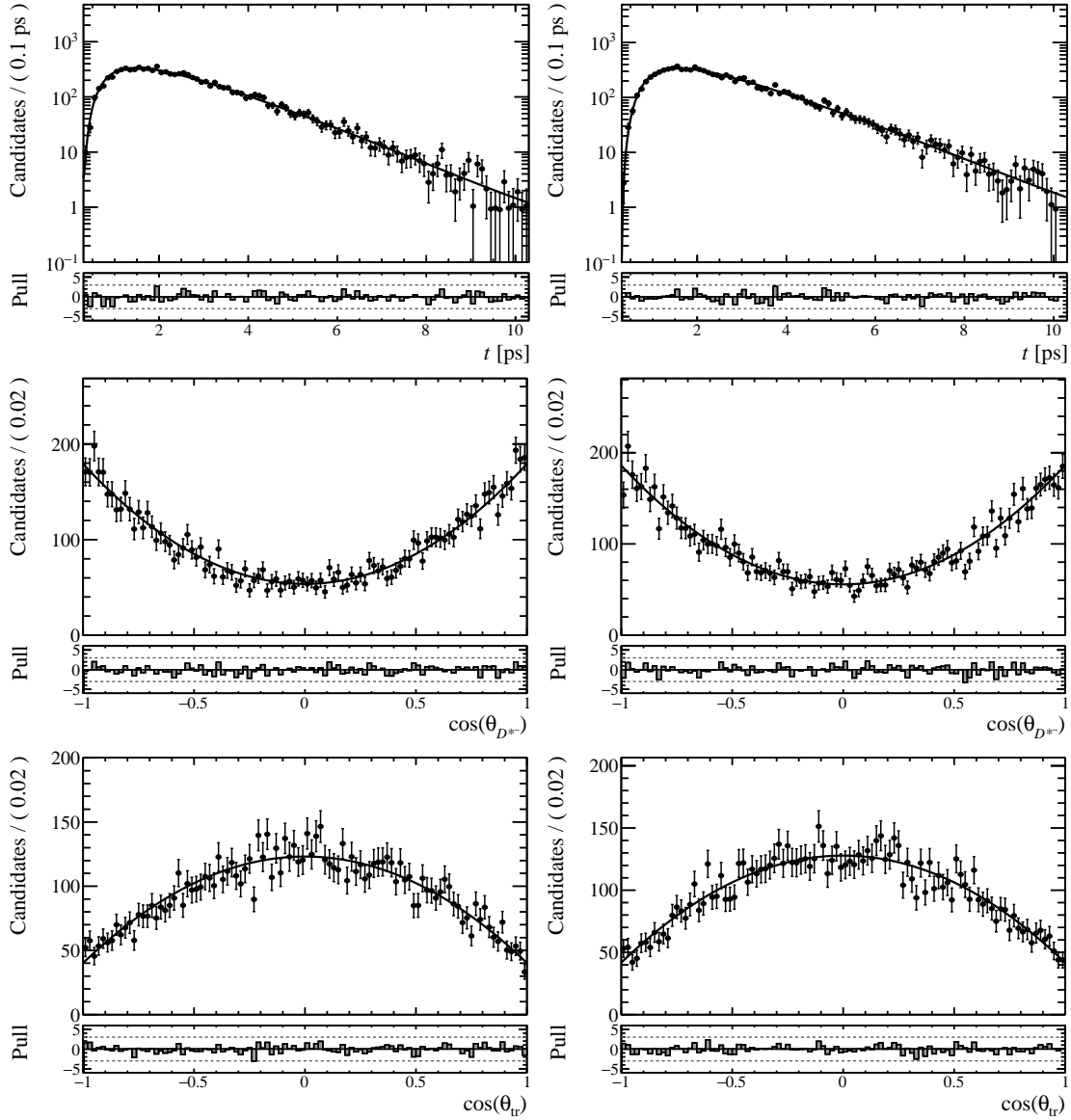


Figure 8.7 – Angular and lifetime distributions of simulated $B^0 \rightarrow D^{*+}D^{*-}$ decays shown together with the fit result of the combined study and the pull distributions for (left) Run 1 and (right) Run 2. The (top) decay-time distribution t and the angular distributions (middle) $\cos(\theta_{D^{*-}})$ and (bottom) $\cos(\theta_{tr})$ are presented. A simultaneous fit to the Run 1 and Run 2 simulations is performed.

Table 8.11 – Results of the individual studies to validate the final fit model given in Eq. (8.19). The studies are the individual angular fit, the individual decay-time fit and the combined fit. All fits are performed using $B^0 \rightarrow D^{*+}D^{*-}$ simulation, which only comprises the $D^0 \rightarrow K^- \pi^+$ decay mode. The fits are performed simultaneously for Run 1 and Run 2.

Study	Parameter	Value
Angular	R_0	0.6216 ± 0.0012
	R_{\perp}	0.1390 ± 0.0009
Time	$C_{D^{*+}D^{*-}}$	-0.05 ± 0.05
	$S_{D^{*+}D^{*-}}$	-0.74 ± 0.07
Combined	R_0	0.625 ± 0.005
	R_{\perp}	0.140 ± 0.004
	$C_{D^{*+}D^{*-}}$	-0.05 ± 0.05
	$S_{D^{*+}D^{*-}}$	-0.74 ± 0.05

according to Sec. 8.1.2 to avoid underestimating the statistical uncertainties in the fit. This correction is applied to the Run 1 and Run 2 samples individually. The shared parameters between both data sets are the transversity amplitudes R_0 and R_{\perp} and the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$. The decay-time resolution and acceptance models are fixed to the parameters determined during the evaluation of the models. External inputs and the FT calibration parameters are Gaussian constrained in the fit. At this point, the analysis is still blinded so that only the uncertainties of the parameters are presented. They are given with the complete fit result in Table 8.12. The statistical correlations of the physics parameters are given in Table 8.13. Additionally, the angular and decay-time distributions overlaid with the fit result are presented in Fig. 8.8.

8.4 Studies of systematic effects

This section summarises different sources of systematic uncertainties on the measurement of the CP observables and the CP -odd fraction considered in this thesis. The statistical uncertainties resulting from the final decay-time fit are cross-checked in Sec. 8.4.1 by exploiting the bootstrapping method. Additionally, the impact of the Gaussian constraints on the statistical uncertainties is checked in Sec. 8.4.2. Further systematic effects are studied using pseudo-experiments, where the influence of several assumptions on the fit inputs or the fit model itself is investigated. The general setup of the pseudo-experiments is explained in Sec. 8.4.3 where also a possible fit bias by the model given in Eq. (8.19) is studied. Subsequently, the model used to generate the data sets of the pseudo-experiments is adjusted to test the various assumptions. For each study, the nominal result is subtracted from the results of the pseudo-experiments to obtain the residual distribution. The mean of the distribution is assigned as a system-

Table 8.12 – Results of the fit to data using the combined angular and decay-time fit model given in Eq. (8.19). The fit is performed to weighted data samples of Run 1 and Run 2, where the $sWeights$ are combined with angular weights to account for acceptance effects in the angular phase space. The fit is performed simultaneously for Run 1 and Run 2.

Parameter	Combined
$\sigma(R_0)$	0.019
$\sigma(R_\perp)$	0.016
$\sigma(C_{D^*+D^{*-}})$	0.16
$\sigma(S_{D^*+D^{*-}})$	0.23
Δm_d [ps ⁻¹]	0.5065 ± 0.0019
τ [ps]	1.520 ± 0.004
$A_{\text{prod}}^{\text{Run 1}}$	-0.011 ± 0.009
$A_{\text{prod}}^{\text{Run 2}}$	0.004 ± 0.005
$\epsilon_{\text{tag, Run 1}}^{\text{OS}}$	0.425 ± 0.030
$\epsilon_{\text{tag, Run 2}}^{\text{OS}}$	0.440 ± 0.015
$\epsilon_{\text{tag, Run 1}}^{\text{SS}}$	0.611 ± 0.030
$\epsilon_{\text{tag, Run 2}}^{\text{SS}}$	0.889 ± 0.009
$p_0^{\text{OS, Run 1}}$	0.389 ± 0.012
$p_0^{\text{OS, Run 2}}$	0.370 ± 0.005
$p_0^{\text{SS, Run 1}}$	0.409 ± 0.010
$p_0^{\text{SS, Run 2}}$	0.4236 ± 0.0035
$p_1^{\text{OS, Run 1}}$	1.02 ± 0.12
$p_1^{\text{OS, Run 2}}$	0.79 ± 0.04
$p_1^{\text{SS, Run 1}}$	1.01 ± 0.14
$p_1^{\text{SS, Run 2}}$	0.94 ± 0.05

Table 8.13 – Statistical correlation of the parameters resulting from the angular decay-time fit.

	R_0	R_\perp	$S_{D^*+D^{*-}}$	$C_{D^*+D^{*-}}$
R_0	1.0	-0.34	0.09	0.00
R_\perp		1.0	-0.26	0.00
$S_{D^*+D^{*-}}$			1.0	0.35
$C_{D^*+D^{*-}}$				1.0

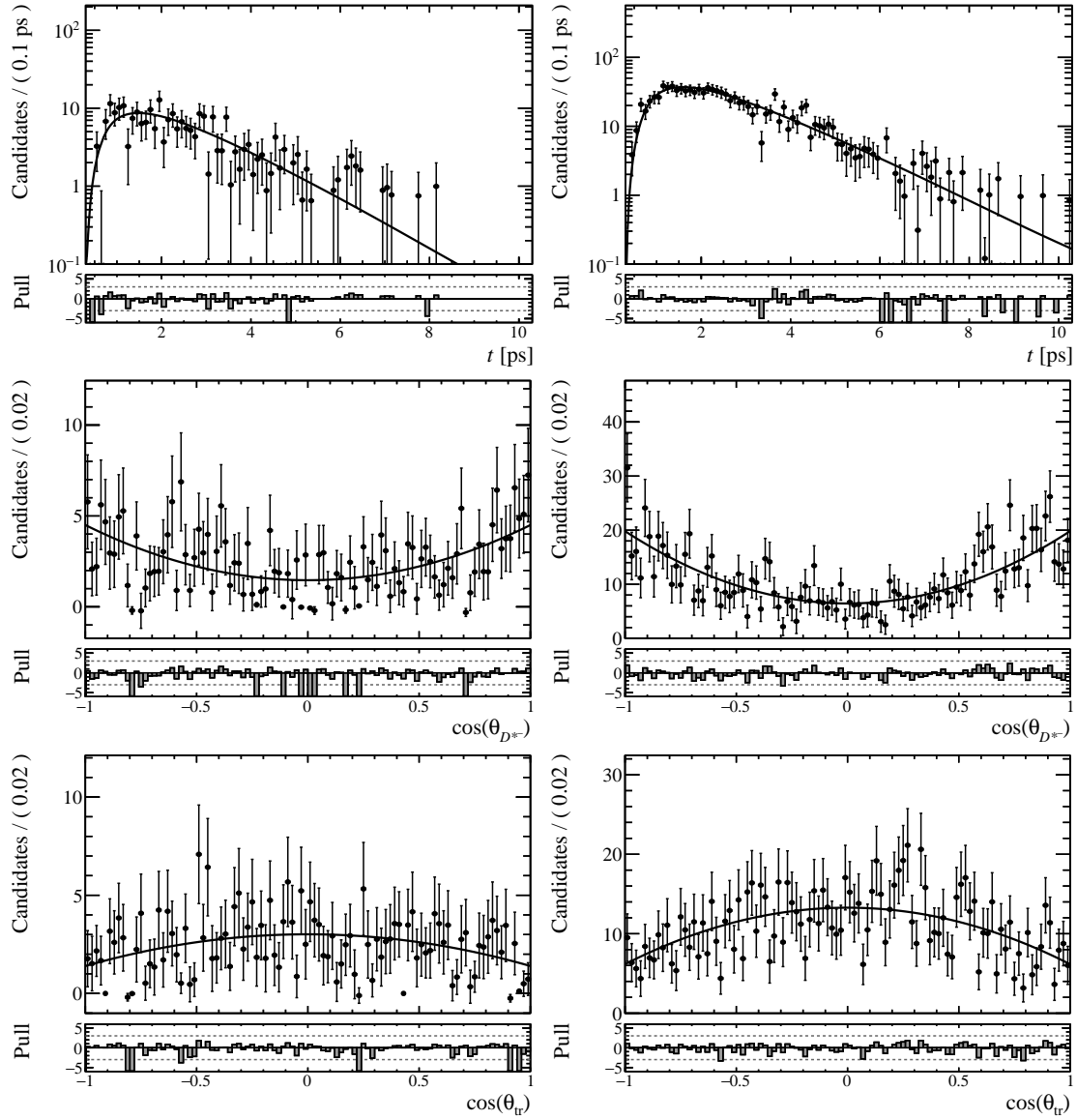


Figure 8.8 – Weighted angular and lifetime distributions of recorded data shown together with the result of the angular decay-time fit and the pull distributions for (left) Run 1 and (right) Run 2. The (top) decay-time distribution t and the angular distributions (middle) $\cos(\theta_{D^{*-}})$ and (bottom) $\cos(\theta_{tr})$ are presented. The weights are obtained by combining the $sWeights$ and the angular weights, which is discussed in more detail in the text. A simultaneous fit to the Run 1 and Run 2 simulations is performed.

atic uncertainty. The impact of the models to describe the invariant $D^{*+}D^{*-}$ mass, the decay-time resolution and the decay-time acceptance are studied in Secs. 8.4.4 to 8.4.6 by performing such pseudo-experiments. The total systematic uncertainty is obtained by adding all individual uncertainties in quadrature. All considered systematic uncertainties and the total uncertainty are summarised in Table 8.14.

Table 8.14 – Summary of the systematic uncertainties. For each observable, the individual uncertainties are added in quadrature to calculate the total systematic uncertainty.

Study	$S_{D^{*+}D^{*-}}$	$C_{D^{*+}D^{*-}}$	R_0	R_{\perp}
Mass model	0.0078 ± 0.0025	0.0016 ± 0.0023	0.00057 ± 0.00026	0.00046 ± 0.0002
Acceptance	0.0091 ± 0.0033	0.0019 ± 0.0032	0.0006 ± 0.0004	0.00042 ± 0.00029
Resolution	0.01 ± 0.0035	0.0033 ± 0.0033	0.0009 ± 0.0004	0.00019 ± 0.00029
Total	0.0156 ± 0.0032	0.0041 ± 0.0031	0.0012 ± 0.0004	0.00065 ± 0.00025

8.4.1 Bootstrapping

The statistical uncertainties resulting from the angular decay-time fit can be underestimated due to the usage of *sWeights* [164]. Bootstrapping studies are performed to estimate the correct statistical uncertainties of the fit results [185]. This technique is a frequentist, model-independent method as it is characterised by creating new data sets based on the original data set. To create a new set, candidates are drawn from the original data set until the original number of candidates is reached. After drawing a candidate, it is kept in the original data set so that single candidates can be drawn multiple times. This approach maintains the correlations of the observables, and by creating multiple data sets, it allows for repeating the angular decay-time fit several times. The variation in the fit results is used to estimate the statistical uncertainty.

A total of 6000 bootstrapped data sets is created. The default mass fit is repeated, and the angular-reweighting procedure is applied to each data set. These steps are performed separately for Run 1 and Run 2. Finally, the angular decay-time fit is performed simultaneously to the Run 1 and Run 2 data sets of each study following the same strategy as described in Sec. 8.3. The residual distributions are obtained by subtracting the fit results from the nominal results. The residuals of the angular observables R_0 and R_{\perp} and the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ are shown in Fig. 8.9. The width of the distribution is a measure of the statistical uncertainty of the fit results. It is determined by a Gaussian fit to the residual distribution. The pull distributions, *i.e.* the difference between the fit result and nominal value divided by the statistical uncertainty from the fit of each observable, are compatible with a Gaussian distribution with a mean of 0 and a width of 1, ensuring a good coverage of the statistical uncertainties. They are not explicitly shown in this thesis.

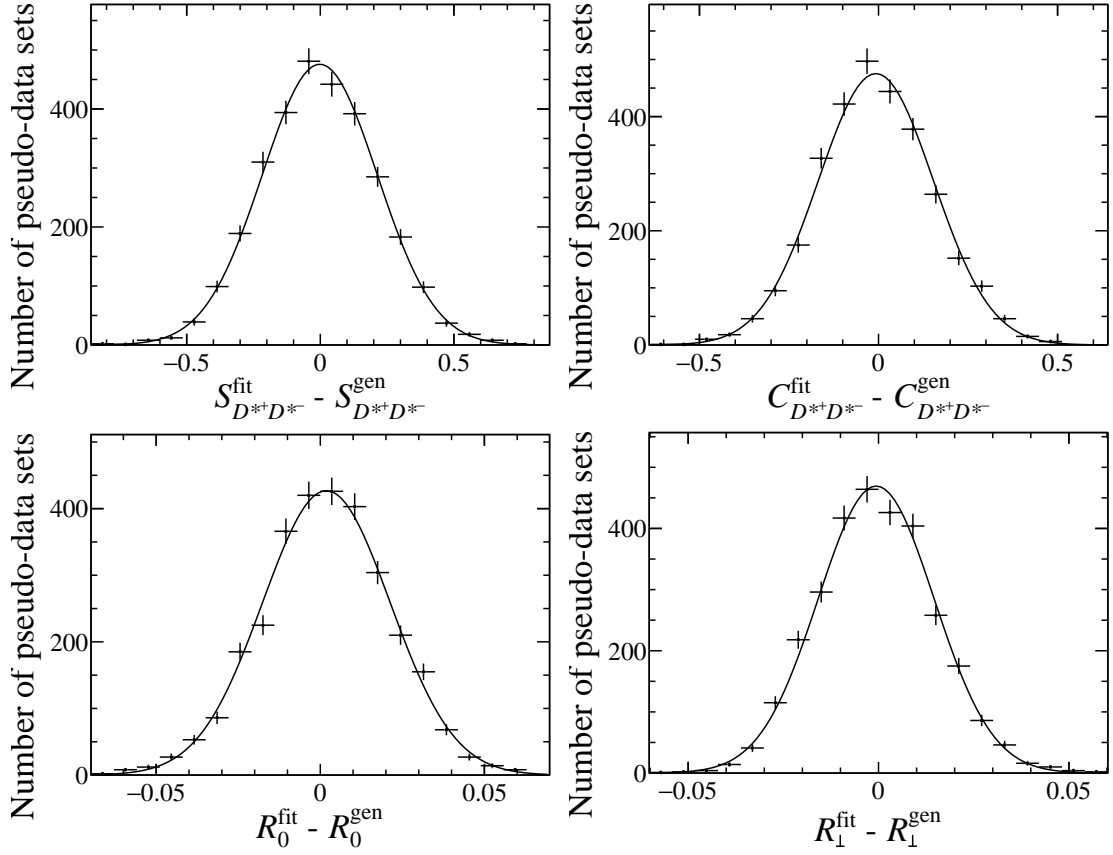


Figure 8.9 – Residual distribution of the CP observables (top left) $S_{D^{*+}D^{*-}}$ and (top right) $C_{D^{*+}D^{*-}}$ and of the angular observables (bottom left) R_0 and (bottom right) R_\perp obtained from the bootstrapping procedure.

The statistical uncertainties resulting from the bootstrapping procedure are summarised in Table 8.15. They are compatible with the statistical uncertainties provided by the default fit given in Table 8.12. This validates the uncertainties provided by the default fit, which are therefore used in the following.

Table 8.15 – Statistical uncertainties of the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ and the angular observables R_0 and R_\perp resulting from the bootstrapping procedure.

Observable	Statistical uncertainty
$S_{D^{*+}D^{*-}}$	0.2151 ± 0.0028
$C_{D^{*+}D^{*-}}$	0.1605 ± 0.0022
R_0	0.01945 ± 0.00028
R_\perp	0.01525 ± 0.00020

8.4.2 Gaussian constraints

The parameters and the respective uncertainties of several external and internal inputs in the angular decay-time fit are Gaussian constrained. These constraints can impact the statistical uncertainties of the fit result. The constrained parameters in the fit are $\mathcal{A}_{\text{prod}}$, Δm_d , τ_{B^0} and the tagging parameters. To estimate the impact of each constraint, the angular decay-time fit is repeated, where one of these parameters, or in the case of the tagging parameters, a set of parameters, is fixed to their input values, neglecting the uncertainty. The procedure is performed for each of the input parameters. The resulting statistical uncertainty is subtracted in quadrature from the nominal uncertainty for each observable. The impact on the statistical uncertainty of each constraint is given in Table 8.16. The largest impact is observed for the tagging parameters, while other parameters do not change the statistical uncertainties significantly. With the number provided in Table 8.16, the pure statistical uncertainty can be calculated if needed.

Table 8.16 – Impact of each Gaussian constraint, or in case of the tagging parameters, a set of Gaussian constraints, on the statistical uncertainties of the fit result.

Constraint	$S_{D^{*+}D^{*-}}$	$C_{D^{*+}D^{*-}}$	R_0	R_{\perp}
Tagging	0.032	0.015	0.0004	0.0005
A_{prod}	0.007	0.005	0.00013	0.00026
$\tau(B^0)$	0.0025	0.0019	0.00009	0.00011
Δm	0.0018	0.0022	0.00009	0.00010

8.4.3 Fit bias

The angle-dependent decay-time PDF defined in Eq. (8.19) used to describe the angular and decay-time distributions is validated in Sec. 8.3.2. The parameters resulting from a fit to simulation are consistent with the true generated values for each observable. However, the model is complex, and the fit might be biased. To determine the inherent bias of the model, pseudo-experiments are performed. These pseudo-experiments are beneficial since much less computing power is required compared to full simulation. This is because the data in these pseudo-experiments is generated according to predefined PDFs or directly according to histograms present in the recorded data. Consequently, they offer a convenient approach to study the impact of certain assumptions on the fit model.

The data is generated separately for each run in the pseudo-experiments, but a simultaneous fit to both samples is performed to determine the angular and CP observables. The detailed procedure of the generation of one pseudo-experiment for one run is described in the following: First, the invariant $D^{*+}D^{*-}$ distribution is generated using the default mass model and the fit result given in Sec. 8.1.2. Each contribution described

by the mass model is generated separately, *i.e.* the $B^0 \rightarrow D^{*+}D^{*-}$, the $B_s^0 \rightarrow D^{*+}D^{*-}$ and the combinatorial background contribution. The number of candidates used in the generation is drawn from a Poisson distribution where the mean value is the yield of the respective component taken from the mass fit to recorded data. As a next step, the angular and decay-time distributions are generated. The PDFs of these distributions are assumed to be uncorrelated with the PDF describing the mass distribution. This allows the use of different PDFs for the various contributions present in the invariant $D^{*+}D^{*-}$ distribution. The angular PDF does not depend on any inputs and is directly generated according to the theoretical PDF, which is the two-dimensional decay rate. The decay-time PDF depends on the mistag distributions. These are generated according to the histograms of the recorded data sets. The inputs for the $B^0 \rightarrow D^{*+}D^{*-}$ signal contribution are fixed to the parameters resulting from the default angular decay-time fit performed in Sec. 8.3.2. The Gaussian constraints are considered in the generation by setting the constrained parameters to values randomly drawn from a Gaussian distribution, with the mean and width corresponding to the value of the parameter and its uncertainty. The blinded parameters are replaced with the values measured by the Belle experiment [92] for the angular observables and the world averages taken from Ref. [188] for the CP observables, *i.e.*

$$R_0 = 0.624, \quad R_{\perp} = 0.138, \quad S_{D^{*+}D^{*-}} = -0.71, \quad C_{D^{*+}D^{*-}} = -0.01.$$

Different inputs are used in the generation for the background components, *e.g.* the combinatorial background assumes no CP violation or $B^0-\bar{B}^0$ oscillation. The tagging parameters for the combinatorial background component are set to provide a random decision, *i.e.* $p_0 = 0.5$ and $p_1 = 0$. The same applies to the SS taggers for the $B_s^0 \rightarrow D^{*+}D^{*-}$ contribution because the *SSProton* and *SSPion* taggers are developed for B^0 decays. The parameters of the OS taggers and the R_0 and R_{\perp} amplitudes of the $B_s^0 \rightarrow D^{*+}D^{*-}$ contribution are set to the values of the $B^0 \rightarrow D^{*+}D^{*-}$ contribution. The transversity amplitudes of the two decays are assumed to be similar because the $H_0 > H_+ > H_-$ hierarchy in the helicity amplitudes increases with a higher momentum of the B meson [47]. For the combinatorial background, the transversity amplitudes are randomly chosen to be between 0 and 1. The inputs that differ in the generation of the combinatorial background and $B_s^0 \rightarrow D^{*+}D^{*-}$ contributions compared to the $B^0 \rightarrow D^{*+}D^{*-}$ signal contribution are listed in Table 8.17.

Table 8.17 – Input parameters for the $B_s^0 \rightarrow D^{*+}D^{*-}$ and combinatorial background contributions used in the generation of the pseudo-experiments.

Background	τ [ps]	Δm [ps ⁻¹]	S	C	R_0	R_{\perp}	tagging SS	tagging OS
Combinatorial	0.5	0	0	0	0.353	0.375	random	random
$B_s^0 \rightarrow D^{*+}D^{*-}$	1.527	17.765	0	0	0.624	0.138	random	as signal

The default analysis strategy is applied to the generated Run 1 and Run 2 data sets of each pseudo-experiment. An extended maximum-likelihood fit is performed to the invariant $D^{*+}D^{*-}$ distributions to extract the $sWeights$. Since the data sets are directly generated with the theoretical PDF of the angular distribution, the angular reweighting procedure is not applied. Finally, a simultaneous fit to the angular and decay-time distributions between Run 1 and Run 2 is performed to extract the angular and CP observables. The whole procedure is repeated for each pseudo-experiment. A total of 6000 pseudo-experiments is analysed. The residual distribution is obtained by subtracting the nominal fit result from the fit results of the pseudo-experiments. The distributions are shown in Fig. 8.10 for the angular observables R_{\perp} and R_0 and the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$.

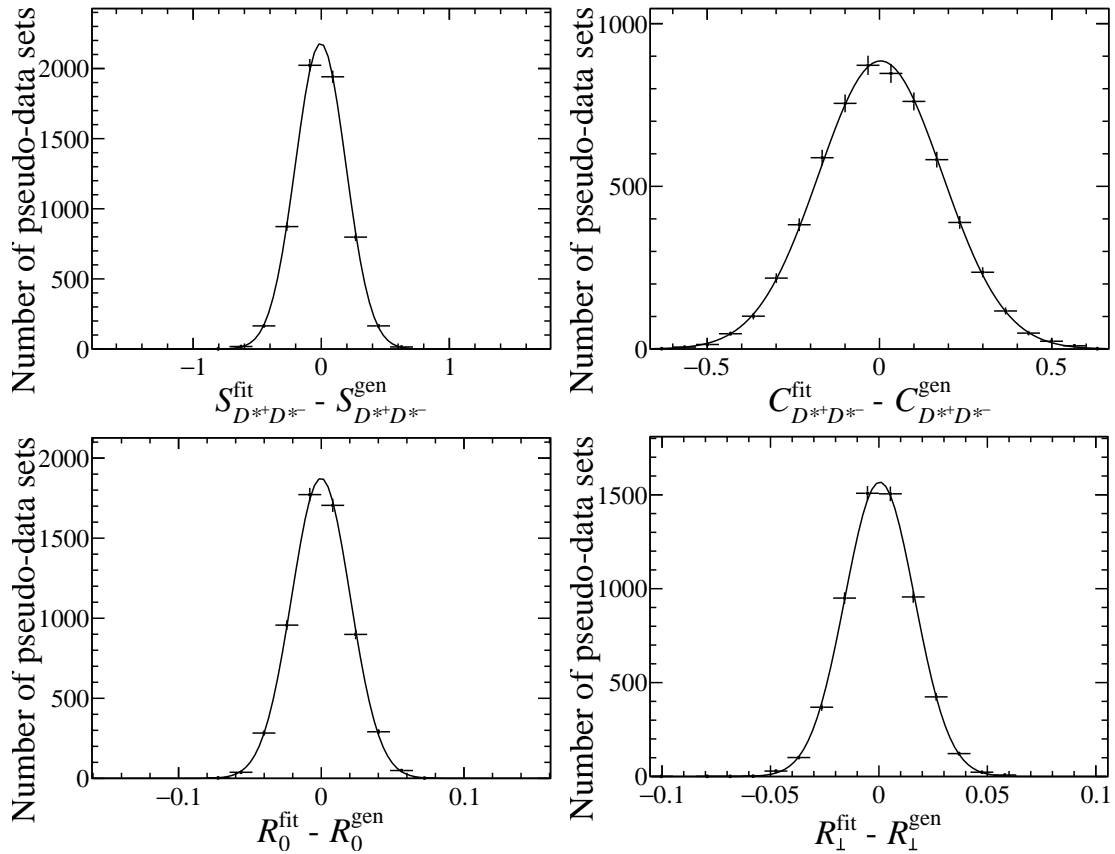


Figure 8.10 – Residual distributions of the CP observables (top left) $S_{D^{*+}D^{*-}}$ and (top right) $C_{D^{*+}D^{*-}}$ and of the angular observables (bottom left) R_0 and (bottom right) R_{\perp} obtained from the pseudo-experiments with the default setup in the generation to extract the fit bias.

Maximum-likelihood fits to the residual distribution are performed using a Gaussian. The resulting mean of the Gaussian is considered as the fit bias. The results of the fits to

the residual distributions are given in Table 8.18, where the fit bias is compatible with 0 for each observable. Consequently, no systematic uncertainty is assigned, and possible biases are neglected in the following systematic studies. The pull distributions of this study are compatible with a Gaussian distribution with a mean of 0 and a width of 1. This ensures a good coverage of the statistical uncertainties. The distributions are not explicitly shown in this thesis. To test assumptions made on the model used in the angular decay-time fit, further pseudo-experiments are performed in the following, where the model used in the generation may be altered to investigate systematic uncertainties from these assumptions.

Table 8.18 – Fit bias of the angular observables R_0 and R_\perp and of the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ obtained from the pseudo-experiments.

Study	$S_{D^{*+}D^{*-}}$	$C_{D^{*+}D^{*-}}$	R_0	R_\perp
Fit bias	-0.0050 ± 0.0025	0.0032 ± 0.0023	-0.00020 ± 0.00027	0.00039 ± 0.00021

8.4.4 Mass model

The choice of the mass model is considered a possible source of systematic uncertainties since it is essential for extracting the $sWeights$. To check the influence of the choice on the angular and CP observables, the default mass model is replaced with an alternative model that consists of two CB functions [181]. This model is also used to cross-check the choice of the mass model in the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis, and its definition is given in Sec. 7.2.3. An extended maximum-likelihood fit to the invariant $D^{*+}D^{*-}$ mass with the alternative model is performed to obtain the values used to generate the data sets of the pseudo-experiments. Despite the different mass model in the generation, the general procedure of the evaluation of a systematic uncertainty discussed in Sec. 8.4.3 remains unchanged. This implies that the default model is used to determine the $sWeights$, and the default decay-time fit is performed to extract the angular and CP observables. The residual distributions for these observables are shown in Fig. 8.11. The pull distributions are not explicitly shown, but the means and widths are compatible with 0 and 1, respectively, ensuring a good coverage of the statistical uncertainties. The resulting systematic uncertainties are given in Table 8.19 for each observable. They are found to be small compared to the statistical uncertainty.

Table 8.19 – Systematic uncertainties on the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ and on the angular observables R_0 and R_\perp due to the change of the mass model.

Study	$S_{D^{*+}D^{*-}}$	$C_{D^{*+}D^{*-}}$	R_0	R_\perp
Mass model	0.0078 ± 0.0025	0.0016 ± 0.0023	0.00057 ± 0.00026	0.00046 ± 0.0002

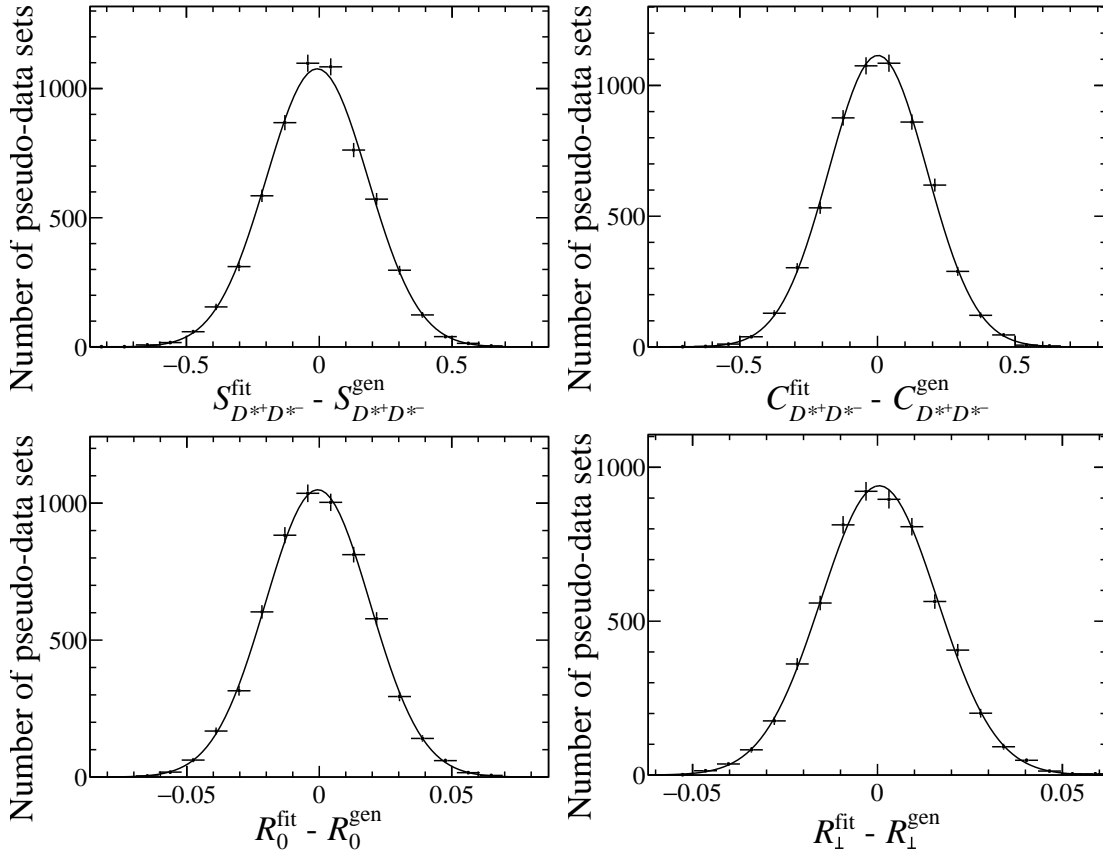


Figure 8.11 – Residual distributions of the CP observables (top left) $S_{D^{*+}D^{*-}}$ and (top right) $C_{D^{*+}D^{*-}}$ and of the angular observables (bottom left) R_0 and (bottom right) R_{\perp} obtained from the pseudo-experiments with a different mass model in the generation. The alternative mass model consists of two CB functions.

8.4.5 Decay-time resolution model

The decay-time resolution model is obtained from simulation. All parameters of the model are fixed in the decay-time fit. However, the resolution between simulation and data might differ. To account for such differences, two separate pseudo-experiment studies are performed. One uses a resolution model in the generation that is 10% wider than the default model, and the other uses a resolution model that is 10% narrower. Besides the broader or narrower resolution in the generation, the pseudo-experiments follow the same strategy as described in Sec. 8.4.3. This implies that the default resolution model is used in the decay-time fit to extract the angular and CP observables. A total of 3000 pseudo-experiments is performed in each study. The resulting uncertainties from both studies are given in Table 8.20. They are obtained by performing a maximum-likelihood fit of a Gaussian to the residual distributions. The systematic uncertainties are found

to be small compared to the statistical uncertainty. The larger of the two uncertainties is assigned for each observable as a final systematic uncertainty. In both studies, the pull distributions of each observable are compatible with a Gaussian distribution with a mean of 0 and a width of 1, ensuring a good coverage of the statistical uncertainties.

Table 8.20 – Systematic uncertainties on the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ and on the angular observables R_0 and R_{\perp} due to the change of the decay-time resolution model.

Study	$S_{D^{*+}D^{*-}}$	$C_{D^{*+}D^{*-}}$	R_0	R_{\perp}
Lower	0.01 ± 0.0035	0.0033 ± 0.0033	0.0009 ± 0.0004	0.00019 ± 0.00029
Higher	0.0073 ± 0.0034	0.0008 ± 0.0033	0.0007 ± 0.0004	$4e - 05 \pm 0.00029$

8.4.6 Decay-time acceptance model

The influence of the choice of the decay-time acceptance model is evaluated in this section. Two pseudo-experiment studies are performed, where different acceptance models are used to generate the data sets. One model uses fewer knots than the default model, while the other uses more knots to describe the decay-time-dependent efficiency. In both cases, an alternative set of knots is found, providing stable results on simulation. They are placed at 0.3, 0.5, 0.9, 1.6, 2.9, 5.5 and 10.3 ps and at 0.3, 0.5, 0.8, 1.3, 2.4, 4.1, 5.5, 7.3 and 10.3 ps for the models with less and more knots, respectively. For each model, the decay-time acceptance fit described in Sec. 8.2.4 is repeated to determine the coefficients of the respective models. Separate pseudo-experiments are performed for each study following the strategy described in Sec. 8.4.3 to determine a systematic uncertainty.

A total of 3000 pseudo-experiments is generated in each study using the alternative decay-time acceptance models. The default acceptance model is used to determine the angular and CP observables. The resulting systematic uncertainties from both studies are summarised in Table 8.21. They are obtained by performing a maximum-likelihood fit of a Gaussian to the residual distributions. The systematic uncertainties are found to be small compared to the statistical uncertainty. The larger of the two uncertainties is assigned for each observable as a final systematic uncertainty. The pull distributions are compatible with a Gaussian distribution with a mean of 0 and a width of 1 for each observable in both studies, ensuring a good coverage of the statistical uncertainties.

8.5 Results

A measurement of the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ and the CP-odd fraction R_{\perp} in $B^0 \rightarrow D^{*+}D^{*-}$ decays is presented. The full Run 1 and Run 2 data sets of the LHCb experiment are used to perform this measurement, which corresponds to a total integrated luminosity of 9 fb^{-1} . The *sPlot* technique is used to calculate *sWeights*

Table 8.21 – Systematic uncertainties on the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ and on the angular observables R_0 and R_{\perp} due to the change of the decay-time acceptance model.

Study	$S_{D^{*+}D^{*-}}$	$C_{D^{*+}D^{*-}}$	R_0	R_{\perp}
Less	0.0062 ± 0.0034	0.0005 ± 0.0033	0.0003 ± 0.0004	0.00034 ± 0.00029
More	0.0091 ± 0.0033	0.0019 ± 0.0032	0.0006 ± 0.0004	0.00042 ± 0.00029

based on an extended maximum-likelihood fit to the invariant $D^{*+}D^{*-}$ mass distribution of the $B^0 \rightarrow D^{*+}D^{*-}$ decays. These signal weights are used in the subsequent angular decay-time fit to unfold the $B^0 \rightarrow D^{*+}D^{*-}$ contribution from the data sets. The total (untagged) $B^0 \rightarrow D^{*+}D^{*-}$ yields are determined to be 282 and 1220 for Run 1 and Run 2, respectively. The flavour-specific decay channel $B^0 \rightarrow D_s^+D^-$ is used to calibrate the FT algorithms. Using the OS and the SS taggers, an effective tagging power of up to 7.0% is achieved in the $B^0 \rightarrow D^{*+}D^{*-}$ channel.

By performing a simultaneous angular decay-time fit of the Run 1 and Run 2 data sets, the CP observables and the CP -odd fraction are measured to be

$$\begin{aligned} S_{D^{*+}D^{*-}} &= X \pm 0.226 \pm 0.016, \\ C_{D^{*+}D^{*-}} &= Y \pm 0.162 \pm 0.004, \\ R_{\perp} &= Z \pm 0.0158 \pm 0.0007, \end{aligned}$$

where the first uncertainty is statistical and the second systematic. The letters X , Y and Z denote the central values of the respective parameters, which are not given since the analysis is still blinded. The systematic uncertainties are taken from Sec. 8.4. A high sensitivity on the CP -odd fraction is achieved with a precision that is of the same order as the world average. No previous measurement of the CP observables in $B^0 \rightarrow D^{*+}D^{*-}$ decays is available from the LHCb experiment, and the world average only comprises measurement from B -factories as the Belle and BaBar experiments. The most precise and latest of these measurements was performed by the Belle experiment [92] and the results of the CP observables are $S_{D^{*+}D^{*-}} = -0.79 \pm 0.13 \pm 0.03$ and $C_{D^{*+}D^{*-}} = -0.15 \pm 0.08 \pm 0.02$, where the first uncertainty is statistical and the second systematic.

In general, it is expected that the LHCb experiment achieves a higher sensitivity on the measurements of $S_{D^{*+}D^{*-}}$ compared to $C_{D^{*+}D^{*-}}$. The opposite assumption should hold for the measurements of the B -factories. The sensitivity of each CP parameter is the highest at the extrema of the respective trigonometric function, *i.e.* for $C_{D^{*+}D^{*-}}$ it is the highest at decay times of zero, while for $S_{D^{*+}D^{*-}}$ it is the highest at longer decay times. Accordingly, the expectation can be explained by the decay-time-dependent selection efficiency at the LHCb experiment, which particularly suppresses low decay times. However, the uncertainties presented in this thesis do not reflect this expectation. A reason could be that the CP observable $S_{D^{*+}D^{*-}}$ is diluted by the factor $(1 - 2R_{\perp})$ and that the statistical correlation of $S_{D^{*+}D^{*-}}$ and R_{\perp} is about -0.26 . Both observables are

determined by a simultaneous fit of the angular and decay time distributions. This inclusive approach is chosen because the expected uncertainty on the CP -odd fraction is smaller than the uncertainty of the world average. The sensitivity on the CP observables $S_{D^{*+}D^{*-}}$ and $C_{D^{*+}D^{*-}}$ and consequently on the measurement of $\sin(2\beta^{\text{eff}})$ in this analysis is limited in general compared to the world average and the previous Belle measurement. This is due to the lower tagging power of the FT algorithms used in this analysis. The B -factories reach tagging powers of around 30% [20, 189], while the measured tagging power in this analysis is around 7% which is still relatively high compared to other LHCb measurements. In contrast, the high sensitivity on the CP -odd fraction R_{\perp} is achieved because the measurement of this observable does not rely on the FT.

In the absence of the hadronic contributions, it follows that the parameter $S_{D^{*+}D^{*-}}$ is equal to $-\sin 2\beta$. However, Eq. (2.58) takes these contributions into account by calculating the effective weak phase β^{eff} . By using this definition, the weak phase β^{eff} is determined to be

$$\beta^{\text{eff}} = W \pm 0.246 \pm 0.016,$$

where W denotes the blinded central value and the uncertainties are statistical and systematic. When the analysis is unblinded, the value of β^{eff} can be compared to the world average of β to calculate the phase shift $\Delta\Phi_{d,D^{*+}D^{*-}}$.

9 Summary and outlook

In this thesis, extensive studies of $B \rightarrow D^{*+}D^{*-}$ decays are presented. Two analyses are performed using data from the LHCb experiment corresponding to an integrated luminosity of 9 fb^{-1} . This data is recorded at centre-of-mass energies of 7, 8 and 13 TeV during the Run 1 and Run 2 data-taking periods of the LHCb experiment.

The measurement of $B_s^0 \rightarrow D^{*+}D^{*-}$ decays is presented, which has been published in Ref. [40]. The decay is observed with a significance of more than 10 standard deviations, and its branching fraction is measured relative to the $B^0 \rightarrow D^{*+}D^{*-}$ branching fraction. The time-integrated ratio is measured to be

$$\frac{\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-})}{\mathcal{B}(B^0 \rightarrow D^{*+}D^{*-})} = 0.269 \pm 0.032 \pm 0.008 \pm 0.011,$$

where the first uncertainty is statistical, the second systematic and the third due to the uncertainty of the fragmentation fraction ratio f_s/f_d . The latter is given separately as it is an external input in this analysis. The absolute time-integrated $B_s^0 \rightarrow D^{*+}D^{*-}$ branching fraction is determined to be

$$\mathcal{B}(B_s^0 \rightarrow D^{*+}D^{*-}) = (2.15 \pm 0.26 \pm 0.09 \pm 0.06 \pm 0.16) \times 10^{-4},$$

where the fourth uncertainty is due to the limited knowledge of the $B^0 \rightarrow D^{*+}D^{*-}$ branching fraction. This result is in agreement with the theoretical prediction using a perturbative QCD approach [41]. The measurement is statistically limited, and the dominant systematic uncertainties are due to the effect of polarised decays on the selection efficiencies and from the external input of the fragmentation fraction ratio.

A second analysis focussing on the measurement of CP violation in $B^0 \rightarrow D^{*+}D^{*-}$ decays is also presented. The analysis is still blinded, as it is currently in internal review within the LHCb collaboration. A simultaneous fit to the angular and decay-time distributions of the $B^0 \rightarrow D^{*+}D^{*-}$ decays of Run 1 and Run 2 yields

$$S_{D^{*+}D^{*-}} = X \pm 0.226 \pm 0.016, \quad C_{D^{*+}D^{*-}} = Y \pm 0.162 \pm 0.004, \quad R_{\perp} = Z \pm 0.0158 \pm 0.0007,$$

where X , Y , and Z denote the blinded central values, the first uncertainty is statistical and the second systematic. The weak phase β^{eff} is consequently calculated to be

$$\beta^{\text{eff}} = W \pm 0.246 \pm 0.016,$$

where W denotes the blinded central value. With this statistical precision, the result of the CP -odd fraction will be the most precise measurement in this decay. Overall, the result is expected to be statistically dominated. However, an upgraded LHCb detector [190] has already started taking additional data, and more precise results are expected in this decay channel. With the upgraded detector, the LHCb experiment is scheduled to record data corresponding to an integrated luminosity of 50 fb^{-1} until the end of 2032 [191]. The upgraded detector includes a new GPU-based trigger system [192], which will be beneficial for the $B^0 \rightarrow D^{*+}D^{*-}$ analysis as channels in the $B \rightarrow DD$ family currently suffer from low efficiencies in the hardware trigger. Additionally, a newly developed tagger based on deep neural networks will improve the effective tagging efficiency, showing promising first results [193]. As the measurement of the CP observables is especially limited by the FT compared to the B factories, this will further help to increase the ability of the LHCb experiment to compete in this decay channel. The new FT algorithm is especially beneficial for the $B_s^0 \rightarrow D^{*+}D^{*-}$ analysis, where a first measurement of CP violation is expected with the increased statistics. Scaling the current data to an integrated luminosity of 50 fb^{-1} and assuming similar detector efficiencies, the number of expected $B_s^0 \rightarrow D^{*+}D^{*-}$ candidates is estimated to be around 550. Accordingly, a first measurement of the helicity amplitudes in the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay is also expected, which would allow to discard the dominant systematic uncertainty in the measurement of the branching fraction.

Both measurements presented in this thesis can contribute to constrain higher-order SM effects in the measurement of the CP -violating parameter ϕ_s in $B_s^0 \rightarrow D_s^{(*)+}D_s^{(*)-}$ decays. Consequently, improved measurements would further allow to place stronger constraints, reducing one of the most significant uncertainties [36, 37]. Measuring the helicity amplitudes and ϕ_s in the $B_s^0 \rightarrow D^{*+}D^{*-}$ decay is especially interesting as these observables have yet to be measured. The framework for such an analysis has already been tested in this thesis, but the statistics are too limited to measure ϕ_s or the helicity amplitudes.

Measurements in both decay channels are also expected from the Belle II experiment, which is scheduled to record data corresponding to an integrated luminosity of 50 ab^{-1} [194]. This is roughly 50 times larger than the data set from the Belle experiment. Consequently, this experiment will be the main competitor of the LHCb experiment in the next years. After running at the $Y(4S)$ resonance, the Belle II experiment might increase their centre-of-mass energy to take data at the $Y(5S)$ resonance [195]. In this case, the Belle II experiment can also contribute to the measurement of CP violation in the $B_s^0 \rightarrow D^{*+}D^{*-}$ decays. Otherwise, the LHCb experiment will be the only experiment to perform this measurement. Overall, the future of the $B \rightarrow D^{*+}D^{*-}$ analyses at LHCb is promising. Based on this thesis, the measurements of the properties of the SM can be further extended in this sector. This is particularly interesting as these measurements can provide more insights and a deeper understanding of the SM, which will be an important part of solving the open questions of this model.

Bibliography

- [1] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- [2] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367.
- [3] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579, ISSN: 0029-5582.
- [4] UA1 collaboration, G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV² at the CERN SPS Collider*, Phys. Lett. B **126** (1983) 398.
- [5] UA1 collaboration, G. Arnison et al., *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. B **122** (1983) 103.
- [6] UA2 collaboration, M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider*, Phys. Lett. B **122** (1983) 476.
- [7] ATLAS collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv:1207.7214.
- [8] CMS collaboration, S. Chatrchyan et al., *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv:1207.7235.
- [9] Super-Kamiokande collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003.
- [10] SNO collaboration, Q. R. Ahmad et al., *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, arXiv:nucl-ex/0204008.
- [11] Planck collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020) A6, arXiv:1807.06209, [Erratum: Astron.Astrophys. 652, C4 (2021)].
- [12] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

- [13] M. B. Gavela et al., *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. A **9** (1994) 795, arXiv:hep-ph/9312215.
- [14] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. D **51** (1995) 379, arXiv:hep-ph/9404302.
- [15] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.
- [16] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.
- [17] LHCb collaboration, R. Aaij et al., *Measurement of CP violation in $B^0 \rightarrow \psi(\rightarrow \ell^+ \ell^-) K_S^0(\rightarrow \pi^+ \pi^-)$ decays*, Phys. Rev. Lett. **132** (2024) 021801, arXiv:2309.09728.
- [18] LHCb collaboration, R. Aaij et al., *Improved measurement of CP violation parameters in $B_s^0 \rightarrow J/\psi K^+ K^-$ decays in the vicinity of the $\phi(1020)$ resonance*, Phys. Rev. Lett. **132** (2024) 051802, arXiv:2308.01468.
- [19] BaBar collaboration, B. Aubert et al., *Measurement of branching fractions and CP-violating charge asymmetries for B meson decays to $D^* \bar{D}^*$, and implications for the CKM angle γ* , Phys. Rev. D **73** (2006) 112004, arXiv:hep-ex/0604037.
- [20] BaBar collaboration, B. Aubert et al., *Measurements of time-dependent CP asymmetries in $B^0 \rightarrow D^{(*)+} D^{(*)-}$ decays*, Phys. Rev. D **79** (2009) 032002, arXiv:0808.1866.
- [21] Belle collaboration, S. Fratina et al., *Evidence for CP violation in $B^0 \rightarrow D^+ D^-$ decays*, Phys. Rev. Lett. **98** (2007) 221802, arXiv:hep-ex/0702031.
- [22] Belle collaboration, M. Röhrken et al., *Measurements of branching fractions and time-dependent CP violating asymmetries in $B^0 \rightarrow D^{(*)\pm} D^\mp$ decays*, Phys. Rev. **D85** (2012) 091106, arXiv:1203.6647.
- [23] LHCb collaboration, R. Aaij et al., *Measurement of CP violation in $B^0 \rightarrow D^+ D^-$ decays*, Phys. Rev. Lett. **117** (2016) 261801, arXiv:1608.06620.
- [24] LHCb collaboration, R. Aaij et al., *Measurement of CP violation in $B^0 \rightarrow D^{*\pm} D^\mp$ decays*, JHEP **03** (2020) 147, arXiv:1912.03723.
- [25] LHCb collaboration, R. Aaij et al., *Measurement of the CP-violating phase ϕ_s in $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$ decays*, Phys. Rev. Lett. **113** (2014) 211801, arXiv:1409.4619.
- [26] LHCb collaboration, R. Aaij et al., *Measurement of CP violation in $B^0 \rightarrow D^+ D^-$ and $B_s^0 \rightarrow D_s^+ D_s^-$ decays*, LHCb-PAPER-2024-027 (2024), In internal review.
- [27] R. Aleksan et al., *The decay $B \rightarrow D \bar{D}^* + D^* \bar{D}$ in the heavy quark limit and tests of CP violation*, Phys. Lett. **B317** (1993) 173, ISSN: 0370-2693.
- [28] A. Sanda and Z.-z. Xing, *Towards determining ϕ_1 with $B \rightarrow D^{(*)} \bar{D}^{(*)}$* , Phys. Rev. **D56** (1997) 341, arXiv:hep-ph/9702297.

- [29] Z.-z. Xing, *Measuring CP violation and testing factorization in $B_d \rightarrow D^{*\pm} D^\mp$ and $B_s \rightarrow D_s^{*\pm} D_s^\mp$ decays*, Phys. Lett. **B443** (1998) 365, arXiv:hep-ph/9809496.
- [30] Z.-z. Xing, *CP violation in $B^0 \rightarrow D^+ D^-$, $D^{*+} D^-$, $D^+ D^{*-}$ and $D^{*+} D^{*-}$ decays*, Phys. Rev. D **61** (2000) 014010, arXiv:hep-ph/9907455.
- [31] X.-Y. Pham and Z.-z. Xing, *CP asymmetries in $B^0 \rightarrow D^{*+} D^{*-}$ and $B_s^0 \rightarrow D_s^{*+} D_s^{*-}$ decays: P wave dilution, penguin and rescattering effects*, Phys. Lett. B **458** (1999) 375, arXiv:hep-ph/9904360.
- [32] R. Fleischer, *Extracting γ from $B_{s(d)} \rightarrow J/\psi K_S^0$ and $B_{d(s)} \rightarrow D_{d(s)}^+ D_{d(s)}^-$* , Eur. Phys. J. C **10** (1999) 299, arXiv:hep-ph/9903455.
- [33] A. Datta and D. London, *Extracting γ from $B^0(t) \rightarrow D^{(*)+} D^{(*)-}$ and $B^0 \rightarrow D_s^{(*)+} D^{(*)-}$ decays*, Phys. Lett. B **584** (2004) 81, arXiv:hep-ph/0310252.
- [34] J. Albert, A. Datta and D. London, *A Measurement of γ from the decays $B^0(t) \rightarrow D^{(*)+} D^{(*)-}$ and $B^0 \rightarrow D_s^{(*)+} D^{(*)-}$* , Phys. Lett. B **605** (2005) 335, arXiv:hep-ph/0410015.
- [35] R. Fleischer, *Exploring CP violation and penguin effects through $B_d^0 \rightarrow D^+ D^-$ and $B_s^0 \rightarrow D_s^+ D_s^-$* , Eur. Phys. J. C **51** (2007) 849, arXiv:0705.4421.
- [36] S. Faller, R. Fleischer and T. Mannel, *Precision Physics with $B_s^0 \rightarrow J/\psi \phi$ at the LHC: The Quest for New Physics*, Phys. Rev. D **79** (2009) 014005, arXiv:0810.4248.
- [37] B. Bhattacharya, A. Datta and D. London, *Reducing Penguin Pollution*, Int. J. Mod. Phys. A **28** (2013) 1350063, arXiv:1209.1413.
- [38] M. Jung and S. Schacht, *Standard model predictions and new physics sensitivity in $B \rightarrow DD$ decays*, Phys. Rev. D **91** (2015) 034027, arXiv:1410.8396.
- [39] L. Bel et al., *Anatomy of $B \rightarrow D\bar{D}$ decays*, JHEP **07** (2015) 108, arXiv:1505.01361.
- [40] LHCb collaboration, R. Aaij et al., *Observation of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay*, JHEP **07** (2023) 119, arXiv:2210.14945.
- [41] R.-H. Li et al., *Decays of B meson to two charmed mesons*, Phys. Rev. **D81** (2010) 034006, arXiv:0910.1424.
- [42] LHCb collaboration, R. Aaij et al., *Observation of the $B_s^0 \rightarrow D^{*\pm} D^\mp$ decay*, JHEP **03** (2021) 099, arXiv:2012.11341.
- [43] D. Griffiths, *Introduction to elementary particles* (2008), ISBN: 978-3-527-40601-2.
- [44] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory* (1995), Westview Press Inc, ISBN: 978-0-201-50397-5.
- [45] C. Berger, *Elementarteilchenphysik, Von den Grundlagen zu den modernen Experimenten* (2014), Springer Spektrum, ISBN: 978-3-642-41753-5.
- [46] K. Anikeev et al., *B physics at the Tevatron: Run II and beyond* (2001), arXiv:hep-ph/0201071.

- [47] Particle Data Group, R. L. Workman et al., *Review of particle physics*, Prog. Theor. Exp. Phys. **2022** (2022) 083C01.
- [48] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 2*, CERN-TH-412 (1964).
- [49] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.
- [50] K. G. Wilson, *Confinement of Quarks*, Phys. Rev. D **10** (1974) 2445.
- [51] Belle collaboration, S. K. Choi et al., *Observation of a narrow charmonium-like state in exclusive $B^\pm \rightarrow K^\pm \pi^+ \pi^- J/\psi$ decays*, Phys. Rev. Lett. **91** (2003) 262001, arXiv:hep-ex/0309032.
- [52] LHCb collaboration, R. Aaij et al., *Observation of exotic $J/\psi\phi$ structures from amplitude analysis of $B^+ \rightarrow J/\psi\phi K^+$ decays*, Phys. Rev. Lett. **118** (2017) 022003, arXiv:1606.07895.
- [53] LHCb collaboration, R. Aaij et al., *Amplitude analysis of $B^+ \rightarrow J/\psi\phi K^+$ decays*, Phys. Rev. **D95** (2017) 012002, arXiv:1606.07898.
- [54] LHCb collaboration, R. Aaij et al., *Observation of structure in the J/ψ -pair mass spectrum*, Science Bulletin **65** (2020) 1983, arXiv:2006.16957.
- [55] LHCb collaboration, R. Aaij et al., *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.
- [56] LHCb collaboration, R. Aaij et al., *Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$* , Phys. Rev. Lett. **122** (2019) 222001, arXiv:1904.03947.
- [57] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.
- [58] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.
- [59] H. Yukawa, *On the Interaction of Elementary Particles I*, Proc. Phys. Math. Soc. Jap. **17** (1935) 48.
- [60] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191.
- [61] S. L. Glashow, *Towards a unified theory: Threads in a tapestry*, Rev. Mod. Phys. **52** (1980) 539.
- [62] S. Weinberg, *Conceptual foundations of the unified theory of weak and electromagnetic interactions*, Rev. Mod. Phys. **52** (1980) 515.
- [63] A. Salam, *Gauge unification of fundamental forces*, Rev. Mod. Phys. **52** (1980) 525.

-
- [64] Gargamelle Neutrino, F. J. Hasert et al., *Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment*, Phys. Lett. B **46** (1973) 138.
- [65] G. Lüders, *On the equivalence of invariance under time reversal and under particle-antiparticle conjugation for relativistic field theories*, Dan. Mat. Fys. Medd. **28** (1954) 1.
- [66] W. Pauli and L. F. Curtiss, *Niels Bohr and the Development of Physics*, American Journal of Physics **24** (1956) 292, ISSN: 0002-9505.
- [67] T. D. Lee and C.-N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1956) 254.
- [68] C. S. Wu et al., *Experimental Test of Parity Conservation in β Decay*, Phys. Rev. **105** (1957) 1413.
- [69] J. H. Christenson et al., *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.
- [70] Belle collaboration, K. Abe et al., *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, arXiv:hep-ex/0107061.
- [71] BaBar collaboration, B. Aubert et al., *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, arXiv:hep-ex/0107013.
- [72] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) 1285.
- [73] SLAC-SP-017, J. E. Augustin et al., *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406.
- [74] J. J. Aubert et al., *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.
- [75] E288 collaboration, S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.
- [76] D0 collaboration, S. Abachi et al., *Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **74** (1995) 2422, arXiv:hep-ex/9411001.
- [77] CDF collaboration, F. Abe et al., *Observation of top quark production in $p\bar{p}$ collisions*, Phys. Rev. Lett. **74** (1995) 2626, arXiv:hep-ex/9503002.
- [78] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) 1802.
- [79] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.
- [80] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation*, Phys. Rev. Lett. **55** (1985) 1039.

- [81] CKMfitter Group, J. C. et al., *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories (updated results and plots available at: <http://ckmfitter.in2p3.fr>)*, Eur. Phys. J. C **41** (2005) 1, arXiv:hep-ph/0406184.
- [82] ARGUS collaboration, H. Albrecht et al., *Observation of $B^0-\bar{B}^0$ Mixing*, Phys. Lett. B **192** (1987) 245.
- [83] UA1 collaboration, C. Albajar et al., *Search for $B^0-\bar{B}^0$ oscillations at the CERN proton-antiproton collider*, Phys. Lett. B **186** (1987) 247, [Erratum: Phys.Lett.B 197, 565 (1987)].
- [84] NA31 collaboration, H. Burkhardt et al., *First Evidence for Direct CP Violation*, Phys. Lett. B **206** (1988) 169.
- [85] BaBar collaboration, B. Aubert et al., *Measurement of CP violating asymmetries in B^0 decays to CP eigenstates*, Phys. Rev. Lett. **86** (2001) 2515, arXiv:hep-ex/0102030.
- [86] LHCb collaboration, R. Aaij et al., *First observation of CP violation in the decays of B_s^0 mesons*, Phys. Rev. Lett. **110** (2013) 221601, arXiv:1304.6173.
- [87] LHCb collaboration, R. Aaij et al., *Observation of CP Violation in Charm Decays*, Phys. Rev. Lett. **122** (2019) 211803, arXiv:1903.08726.
- [88] T. Jubb et al., *On the ultimate precision of meson mixing observables*, Nucl. Phys. B **915** (2017) 431, arXiv:1603.07770.
- [89] LHCb collaboration, R. Aaij et al., *Measurement of the semileptonic CP asymmetry in $B^0-\bar{B}^0$ mixing*, Phys. Rev. Lett. **114** (2015) 041601, arXiv:1409.8586.
- [90] LHCb collaboration, R. Aaij et al., *Measurement of the CP asymmetry in $B_s^0-\bar{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016) 061803, arXiv:1605.09768.
- [91] CLEO collaboration, M. Artuso et al., *First observation of the decay $B^0 \rightarrow D^{*+} D^{*-}$* , Phys. Rev. Lett. **82** (1999) 3020, arXiv:hep-ex/9811027.
- [92] Belle collaboration, B. Kronenbitter et al., *First observation of CP violation and improved measurement of the branching fraction and polarization of $B^0 \rightarrow D^{*+} D^{*-}$ decays*, Phys. Rev. D **86** (2012) 071103, arXiv:1207.5611.
- [93] Heavy Flavor Averaging Group, Y. Amhis et al., *Averages of b-hadron, c-hadron, and τ -lepton properties as of 2018*, Eur. Phys. J. **C81** (2021) 226, arXiv:1909.12524, updated results and plots available at <https://hflav.web.cern.ch>.
- [94] B. Kayser et al., *B Hadronic Asymmetries and Spin Correlations*, Phys. Lett. B **237** (1990) 508.
- [95] I. Dunietz et al., *How to extract CP violating asymmetries from angular correlations*, Phys. Rev. D **43** (1991) 2193.
- [96] G. Kramer and W. F. Palmer, *Branching ratios and CP asymmetries in the decay $B \rightarrow VV$* , Phys. Rev. D **45** (1992) 193.

-
- [97] M. Jacob and G. C. Wick, *On the General Theory of Collisions for Particles with Spin*, *Annals Phys.* **7** (1959) 404.
- [98] E. Majorana, *Teoria simmetrica dell'elettrone e del positrone*, *Nuovo Cim.* **14** (1937) 171.
- [99] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.
- [100] LHCb collaboration, A. A. Alves Jr. et al., *The LHCb detector at the LHC*, *JINST* **3** (2008) S08005.
- [101] LHCb collaboration, R. Aaij et al., *LHCb detector performance*, *Int. J. Mod. Phys. A* **30** (2015) 1530022, arXiv:1412.6352.
- [102] CERN, *LEP Design Report: Vol.2. The LEP Main Ring*, CERN-LEP-84-01 (1984) .
- [103] B. Goddard and F. Zimmermann, *RLIUP: Review of LHC and Injector Upgrade Plans: Archamps, France, October 29-31, 2013*, CERN, Geneva (2014) .
- [104] J. Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, OPEN-PHO-ACCEL-2013-056 (2013) .
- [105] ATLAS collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.
- [106] CMS collaboration, S. Chatrchyan et al., *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.
- [107] ALICE collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.
- [108] S. J. Brodsky et al., *Physics Opportunities of a Fixed-Target Experiment using the LHC Beams*, *Phys. Rept.* **522** (2013) 239, arXiv:1202.6585.
- [109] LHCb collaboration, C. Elsässer, *$\bar{b}b$ production angle plots* (visited on 22/08/2024).
- [110] C. Ilgner et al., *The Beam Conditions Monitor of the LHCb Experiment* (2010), arXiv:1001.2487.
- [111] M. Bieker, *MIBAB - An FPGA-based readout system for the LHCb Beam Conditions Monitor* (2024), TU Dortmund, PhD thesis.
- [112] LHCb collaboration, R. Aaij et al., *LHCb Detector Performance*, *Int. J. Mod. Phys. A* **30** (2015) 1530022, arXiv:1412.6352.
- [113] R. Aaij et al., *Performance of the LHCb Vertex Locator*, *JINST* **9** (2014) P09007, arXiv:1405.7808.
- [114] R. Arink et al., *Performance of the LHCb Outer Tracker*, *JINST* **9** (2014) P01002, arXiv:1311.3893.
- [115] LHCb Silicon Tracker Group, M. Tobin, *The LHCb Silicon Tracker*, *Nucl. Instrum. Meth. A* **831** (2016) 174.

- [116] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037 (2000).
- [117] M. Adinolfi et al., *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.
- [118] R. Calabrese et al., *Performance of the LHCb RICH detectors during LHC Run 2*, JINST **17** (2022) P07013, arXiv:2205.13400.
- [119] LHCb collaboration, E. Picatoste Olloqui, *LHCb preshower (PS) and scintillating pad detector (SPD): Commissioning, calibration, and monitoring*, J. Phys. Conf. Ser. **160** (2009) 012046.
- [120] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010 (2001).
- [121] F. Archilli et al., *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.
- [122] LHCb collaboration, *Trigger Schemes* (visited on 22/07/2024).
- [123] LHCb collaboration, J. Albrecht, *The LHCb trigger system*, Nucl. Phys. B Proc. Suppl. **187** (2009) 237.
- [124] R. Aaij et al., *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.
- [125] LHCb HLT project, J. Albrecht et al., *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, arXiv:1310.8544.
- [126] 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, arXiv:1812.10790.
- [127] LHCb collaboration, B. Sciascia, *LHCb Run 2 trigger performance*, PoS **BEAUTY2016** (2016) 029.
- [128] G. Barrand et al., *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.
- [129] CERN, *The GAUDI Project* (visited on 07/03/2024).
- [130] LHCb collaboration, *The MOORE Project* (visited on 07/03/2024).
- [131] LHCb collaboration, *The BRUNEL Project* (visited on 07/03/2024).
- [132] M. Benayoun and O. Callot, *The forward tracking, an optical model method*, CERN-LHCb-2002-008 (2002).
- [133] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Journal of Basic Engineering **82** (1960) 35, ISSN: 0021-9223.
- [134] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

-
- [135] LHCb collaboration, *The Stripping Project* (visited on 07/03/2024).
- [136] LHCb collaboration, *The DAVINCI Project* (visited on 07/03/2024).
- [137] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.
- [138] I. Belyaev et al., *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.
- [139] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [140] N. Davidson, T. Przedzinski and Z. Was, *PHOTOS interface in C++: Technical and physics documentation*, Comp. Phys. Comm. **199** (2016) 86, arXiv:1011.0937.
- [141] Geant4 collaboration, J. Allison et al., *Geant4 developments and applications*, IEEE Trans.Nucl.Sci. **53** (2006) 270.
- [142] Geant4 collaboration, S. Agostinelli et al., *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.
- [143] M. Clemencic et al., *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [144] LHCb collaboration, *The GAUSS Project* (visited on 07/03/2024).
- [145] LHCb collaboration, *The BOOLE Project* (visited on 07/03/2024).
- [146] LHCb collaboration, R. Aaij et al., *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. **C72** (2012) 2022, arXiv:1202.4979.
- [147] LHCb collaboration, R. Aaij et al., *New algorithms for identifying the flavour of B^0 mesons using pions and protons*, Eur. Phys. J. **C77** (2017) 238, arXiv:1610.06019.
- [148] LHCb collaboration, R. Aaij et al., *B flavour tagging using charm decays at the LHCb experiment*, JINST **10** (2015) P10005, arXiv:1507.07892.
- [149] LHCb collaboration, R. Aaij et al., *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016) P05010, arXiv:1602.07252.
- [150] K. Heinicke, *Optimization of Flavour Tagging Algorithms for the LHCb Experiment* (2016), TU Dortmund, Presented 30 Sep 2016.
- [151] Q. Führung and V. Jevtić, *lhcb-ftcalib: A software package for the calibration of flavour-tagged LHCb data*, Zenodo (2024).
- [152] A. Rogozhnikov, *Reweighting with boosted decision trees*, J. Phys. Conf. Ser. **762** (2016) 012036, arXiv:1608.05806.
- [153] R. E. Bellman, *Dynamic Programming* (1957), Princeton University Press, Sixth Printing, 1972, ISBN: 0-691-07951-X.
- [154] D. Martschei et al., *Advanced event reweighting using multivariate analysis*, J. Phys. Conf. Ser. **368** (2012) 012028.

- [155] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*. The Annals of Statistics **29** (2001) 1189.
- [156] A. Rogozhnikov, *HEP-ML: HEP-ML provides specific machine learning tools for purposes of high energy physics* (visited on 07/03/2024).
- [157] Blobel, V. and Lohrmann, E., *Statistische und numerische Methoden der Datenanalyse* (1998), Vieweg+Teubner Verlag, ISBN: 978-3-519-03243-4.
- [158] R. J. Barlow, *Extended maximum likelihood*, Nucl. Instrum. Meth. A **297** (1990) 496.
- [159] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343.
- [160] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116.
- [161] I. Antcheva et al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, Comput. Phys. Commun. **182** (2011) 1384.
- [162] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81.
- [163] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.
- [164] C. Langenbruch, *Parameter uncertainties in weighted unbinned maximum likelihood fits*, Eur. Phys. J. C **82** (2022) 393, arXiv:1911.01303.
- [165] S. S. Wilks, *The large-sample distribution of the likelihood ratio for testing composite hypotheses*, Ann. Math.Stat. **9** (1938) 60.
- [166] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, arXiv:1210.6861.
- [167] T. Likhomanenko et al., *LHCb topological trigger reoptimization*, J. Phys. Conf. Ser. **664** (2015) 082025.
- [168] M. De Cian et al., *Fast neural-net based fake track rejection in the LHCb reconstruction*, LHCb-PUB-2017-011 (2017) .
- [169] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, ACM (2016) .
- [170] LHCb collaboration, S. Hollitt et al., *Search for the decay $B_s^0 \rightarrow D^{*+} D^{*-}$* , CERN-LHCb-ANA-2019-058 (2019) .
- [171] J. Langer, *Search for the decay $B_s^0 \rightarrow D^{*+} D^{*-}$ and studies of its control channel $B^0 \rightarrow D^{*+} D^{*-}$ with the LHCb experiment* (2019), TU Dortmund, Master thesis.
- [172] G. A. Cowan, D. C. Craik and M. D. Needham, *RapidSim: an application for the fast simulation of heavy-quark hadron decays*, Comput. Phys. Commun. **214** (2017) 239, arXiv:1612.07489.

-
- [173] Particle Data Group, P. A. Zyla et al., *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01.
- [174] LHCb collaboration, R. Aaij et al., *Measurement of b hadron production fractions in 7 TeV pp collisions*, Phys. Rev. **D85** (2012) 032008, arXiv:1111.2357.
- [175] LHCb collaboration, R. Aaij et al., *Measurement of b -hadron fractions in 13 TeV pp collisions*, Phys. Rev. **D100** (2019) 031102(R), arXiv:1902.06794.
- [176] LHCb collaboration, R. Aaij et al., *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286.
- [177] LHCb collaboration, R. Aaij et al., *Measurement of the branching fraction of the $B^0 \rightarrow D_s^+ \pi^-$ decay*, Eur. Phys. J. **C81** (2021) 314, arXiv:2010.11986.
- [178] LHCb collaboration, R. Aaij et al., *Measurement of f_s/f_u variation with proton-proton collision energy and $-$ meson kinematics*, Phys. Rev. Lett. **124** (2020) 122002, arXiv:1910.09934.
- [179] LHCb collaboration, R. Aaij et al., *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B_s^0 decay branching fractions*, Phys. Rev. **D104** (2021) 032005, arXiv:2103.06810.
- [180] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.
- [181] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances* (1986), Institute of Nuclear Physics, Krakow, DESY-F31-86-02, PhD thesis.
- [182] Particle Data Group, M. Tanabashi et al., *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.
- [183] K. De Bruyn et al., *Branching Ratio Measurements of B_s Decays*, Phys. Rev. D **86** (2012) 014027, arXiv:1204.1735.
- [184] LHCb collaboration, R. Aaij et al., *Observation of the $B_s^0 \rightarrow D^{*+} D^{*-}$ decay*, LHCb-PAPER-2022-023 (2023), Supplementary material.
- [185] B. Efron, *Bootstrap Methods: Another Look at the Jackknife*, The Annals of Statistics **7** (1979) 1.
- [186] F. Pedregosa et al., *Scikit-learn: Machine learning in Python*, J. Machine Learning Res. **12** (2011) 2825, arXiv:1201.0490.
- [187] L. Breiman et al., *Classification and regression trees* (1984), Wadsworth international group, Belmont, California, USA.
- [188] Heavy Flavor Averaging Group, Y. Amhis et al., *Averages of b -hadron, c -hadron, and τ -lepton properties as of 2021*, Phys. Rev. **D107** (2023) 052008, arXiv:2206.07501, updated results and plots available at <https://hflav.web.cern.ch>.

- [189] Belle collaboration, H. Kakuno et al., *Neutral B flavor tagging for the measurement of mixing induced CP violation at Belle*, Nucl. Instrum. Meth. A **533** (2004) 516, arXiv:hep-ex/0403022.
- [190] LHCb collaboration, *LHCb Framework TDR for the LHCb Upgrade II Opportunities in flavour physics, and beyond, in the HL-LHC era*, CERN-LHCC-2021-012 (2022) .
- [191] LHCb collaboration, *Physics case for an LHCb Upgrade II – Opportunities in flavour physics, and beyond, in the HL-LHC era* (2018), arXiv:1808.08865.
- [192] LHCb collaboration, *LHCb Upgrade GPU High Level Trigger Technical Design Report*, CERN-LHCC-2020-006 (2020) .
- [193] T. Likhomanenko, D. Derkach and A. Rogozhnikov, *Inclusive Flavour Tagging Algorithm*, J. Phys. Conf. Ser. **762** (2016) 012045, arXiv:1705.08707.
- [194] Belle II collaboration, T. Abe et al., *Belle II Technical Design Report*, KEK-REPORT-2010-1 (2010), arXiv:1011.0352.
- [195] A. Gaz, *Physics Prospects at SuperKEKB/Belle II*, PoS **KMI2017** (2017) 005.

Acknowledgement

First of all, I would like to thank Prof. Dr. Bernhard Spaan, who unfortunately passed away far too early and whom I would like to remember at this point. He was present for most of my time at university and supervised both my Bachelor's and Master's theses. He also gave me the opportunity to further deepen my studies in $B \rightarrow D^{*+} D^{*-}$ decays by measuring CP violation, which resulted in this thesis. His heartwarming humour and his always positive nature as well as his strong commitment to research will be remembered forever.

I would also like to thank Prof. Dr. Johannes Albrecht. He took over my supervision during a challenging phase for all of us and gave me the opportunity to complete this thesis. I would like to thank him for his outstanding supervision, in which he always found time and advice for me despite the large number of other PhD students. With the truly outstanding KET excursion and the inspiring KET lecture, he also played a significant role in my decision to choose particle physics as a field of research. With a very informative and interesting discussion after the excursion about possible topics for a Bachelor thesis, he brought me to the working group in which I have now written all my theses.

I would also like to thank PD Dr. Dominik Elsässer, who agreed to review this thesis, for taking the time required to do so. I would especially like to thank him for the very encouraging and pleasant discussions before the submission and the exam. His enthusiasm for reading the thesis gave me the confidence I needed for the final stages of the thesis.

I would also like to thank Prof. Dr. Mirko Cinchetti and Dr. Moritz Cygorek for taking the time to complete the examination commission.

As usual in experimental particle physics, the results presented in this thesis were obtained in collaboration with many colleagues. The $B_s^0 \rightarrow D^{*+} D^{*-}$ analysis was started during my master thesis, which was supervised by Antje and Philipp. They later remained part of the analysis team and helped with the publication of the results. During this time, Sophie also joined the team. All three also accompanied the beginnings of the $B^0 \rightarrow D^{*+} D^{*-}$ analysis, to which Quentin has recently joined for the publication process. Overall, I would like to thank everyone for the good cooperation and the many helpful discussions and constructive feedback.

During my time at the chair, I shared the office with many colleagues who I would like to mention here: initially with Martin and Philipp; then this office became the admin office with Fabio, Jannis and Louis; finally, I moved to the so-called 'Schreib' office with Maik, Vukan and Louis. I really enjoyed the time and atmosphere in each constellation.

I would also like to thank the chair AG Albrecht in general for the always pleasant atmosphere and the administration of various infrastructural services that I used. It would go beyond the scope of this acknowledgement to name all the people who are responsible for these services. I would also like to take this opportunity to thank you for your many constructive suggestions for improving my work.

Of course, I would also like to thank my friends, both from my time at university and from home. They have always provided me with the necessary balance and distraction, both during my studies and in my free time. I would especially like to mention my former fellow students Alexander and Louis, with whom I was able to spend the entire time at university. Together, we overcame a wide variety of challenges, always having a lot of fun and entertainment and always being able to rely on each other.

Of course, I would also like to thank my family, who have always supported me unconditionally. First and foremost my parents Jutta and Andreas, without your trust and patience this degree would not have been possible. But I would also like to thank my brother Tim, who always had an open ear for me. You always stood by my side in difficult moments and always encouraged me. Thanks also go to my 'future' parents-in-law, for the distraction and extensive support for the preparations to make the exam day the special day it turned out to be.

Finally, I would like to thank my girlfriend Mara, who has been with me throughout my entire PhD, always standing by my side and supporting me in the most stressful moments. Especially in the last few months, she had my back and gave me the necessary calm. This allowed me to concentrate purely on completing the thesis. This time was certainly not the easiest, but I am grateful that you are by my side.

Without the support of all these people, this work would not have been possible. Many thanks to everyone who has supported me along the way!

Danksagung

Als erstes möchte ich mich bei Prof. Dr. Bernhard Spaan bedanken, der leider viel zu früh verstorben ist und dem ich an dieser Stelle gedenken möchte. Er hat mich den Großteil meiner Studienzeit begleitet und dabei sowohl meine Bachelor- als auch meine Masterarbeit betreut. Er gab mir auch die Möglichkeit, die Studien in den $B \rightarrow D^{*+} D^{*-}$ Zerfällen durch Messung der CP -Verletzung weiter zu vertiefen, woraus diese Arbeit entstanden ist. Sein herzwärmender Humor und seine stets positive Art sowie sein starkes Engagement für die Forschung werden für immer in Erinnerung bleiben.

Nicht weniger Dank meinerseits gilt Herrn Prof. Dr. Johannes Albrecht. Er hat mich in einer für uns alle herausfordernden Phase als Doktorand übernommen und mir die Möglichkeit gegeben diese Arbeit fertig zu stellen. Ich danke Ihm für seine hervorragende Betreuung, in der er trotz der großen Anzahl an weiteren Doktoranden immer Zeit und Rat für mich gefunden hat. Mit der damals wirklich herausragenden KET-Exkursion und der inspirierenden KET-Vorlesung hat er auch einen signifikanten Anteil daran, dass ich mich für die Teilchenphysik als Forschungsgebiet entschieden habe. Durch ein sehr informatives und interessantes Gespräch nach der Exkursion über mögliche Bachelorarbeitsthemen hat er mich zur Arbeitsgruppe gebracht, in der ich nun alle meine Arbeiten durchgeführt habe.

Bedanken möchte ich mich auch bei Herrn PD. Dr. Dominik Elsässer, der sich bereit erklärt und die Zeit genommen hat diese Arbeit zu begutachten. Speziell danke ich Ihm für die sehr ermutigenden und angenehmen Gespräche vor der Abgabe und Prüfung. Sein Enthusiasmus die Arbeit zu lesen hat mir auf dem letzten Stück die nötige Zuversicht gegeben.

Mein Dank gilt auch Herrn Prof. Dr. Mirko Cinchetti und Herrn Dr. Moritz Cygorek, dafür dass sie sich die Zeit genommen haben die Prüfungskommission zu vervollständigen.

Wie in der experimentellen Teilchenphysik üblich wurden die in dieser Arbeit vorgestellten Ergebnisse in Zusammenarbeit mit vielen Kollegen erzielt. Die durchgeführte $B_s^0 \rightarrow D^{*+} D^{*-}$ Analyse wurde während meiner Masterarbeit angefangen, welche von Antje and Philipp betreut wurden. Sie blieben später ein Teil des Analyse Teams und halfen bei der Veröffentlichung der Ergebnisse. Während dieser Zeit ist auch Sophie dazu gestoßen. Alle drei begleiteten auch die Anfänge der $B^0 \rightarrow D^{*+} D^{*-}$ Analyse, zu der seit kurzem auch Quentin für den Veröffentlichungsprozess dazu gestoßen ist. Insgesamt möchte ich mich bei allen für die gute Zusammenarbeit und die vielen hilfreichen Diskussionen sowie konstruktiven Verbesserungsvorschläge bedanken.

Während meiner Zeit am Lehrstuhl habe ich das Büro mit vielen Arbeitskollegen geteilt, die ich an dieser Stelle nicht unerwähnt lassen möchte: anfangs mit Martin und Philipp; anschließend wurde aus diesem Büro das Admin-Büro mit Fabio, Jannis und Louis; schlussendlich bin ich ins sogenannte Schreib-Büro mit Maik, Vukan und Louis umgezogen. In jeder Konstellation habe ich die Zeit und Atmosphäre sehr genossen.

Auch dem Lehrstuhl AG Albrecht insgesamt bedanke ich mich für die immer angenehme Atmosphäre und die Administration diverser von mir genutzter infrastruktureller Services. Alle Personen persönlich zu nennen, die diese Services betreuen, würde den Rahmen sprengen. Auch bedanken möchte ich mich an dieser Stelle für die vielen konstruktiven Verbesserungsvorschläge für die Arbeit.

Ein großer Dank gilt natürlich auch meinen Freunden, aus dem Studium als auch aus meiner Heimat. Sowohl im Studium als auch in der Freizeit haben sie mir immer den nötigen Ausgleich gegeben und die nötige Ablenkung verschafft. Besonders erwähnen möchte ich meine ehemaligen Kommilitonen Alexander und Louis, mit denen ich das ganze Studium verbringen durfte. Zusammen haben wir die unterschiedlichsten Herausforderungen gemeistert, wobei wir immer viel Unterhaltung und Spaß hatten und uns immer auf einander verlassen konnten.

Dank gebührt natürlich auch meiner Familie, die mich immer bedingungslos unterstützt hat. An erster Stelle meine Eltern Jutta und Andreas, ohne euer Vertrauen und eure Geduld wäre dieser Abschluss nicht möglich gewesen. Aber auch meinem Bruder Tim möchte ich danken, der immer ein offenes Ohr für mich hatte. Ihr standet in schwierigen Momenten stets an meiner Seite und habt mich stets ermutigt. Dank gilt natürlich auch meinen Schwiegereltern in spe, für die Ablenkung und die ausgiebige Unterstützung bei den Vorbereitungen, um den Prüfungstag zu dem besonderen Tag zu machen der er geworden ist.

Als letztes möchte ich meiner Freundin Mara danken, die mich während der gesamten Promotion begleitet hat, mir immer zur Seite stand und mich in den stressigsten Momenten unterstützt hat. Besonders in den letzten Monaten hat sie mir den Rücken freigehalten und die nötige Ruhe gegeben. Dadurch konnte ich mich rein auf die Fertigstellung der Arbeit konzentrieren. Diese Zeit war sicherlich nicht die einfachste, aber ich bin dankbar, dass du an meiner Seite bist.

Ohne die Unterstützung all dieser Personen nicht wäre diese Arbeit nicht möglich gewesen. Vielen Dank an alle, die mich auf diesem Weg begleitet haben!