Measurements of $\sin 2\beta$ using charmonium and open charm decays at LHCb

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

vorgelegt von
Frank Meier
geboren in Kiel

Lehrstuhl für Experimentelle Physik V
Fakultät Physik
Technische Universität Dortmund
2016

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

1. Gutachter: Prof. Dr. B. Spaan
2. Gutachter: Prof. Dr. K. Kröninger

Datum des Einreichens der Arbeit: 04.11.2016
Datum der mündlichen Prüfung: 13.12.2016

Abstract

The *CP* violation observables S_f and C_f in the decays of B^0 and $\overline{B}^0$ mesons to the $J/\psi\, K^0_{\rm S}$ final state and to the D^+D^- final state are measured with a data sample corresponding to an integrated luminosity of 3 fb^{-1} collected with the LHCb experiment in proton-proton collisions at centre-of-mass energies of 7 and 8 TeV.

The analysis of the decay-time evolution of 41 560 $B^0 \to J/\psi\, K^0_{\rm S}$ decays yields

$$\begin{aligned} S_{J/\psi\, K^0_{\rm S}} &= 0.731 \pm 0.035\,(\text{stat}) \pm 0.020\,(\text{syst})\,, \\ C_{J/\psi\, K^0_{\rm S}} &= -0.038 \pm 0.032\,(\text{stat}) \pm 0.005\,(\text{syst})\,, \end{aligned}$$

which is consistent with the current world averages and with the Standard Model expectations.

In a flavour-tagged, decay-time-dependent analysis of 1410 $B^0 \to D^+D^-$ decays the following results are determined:

$$\begin{aligned} S_{D^+D^-} &= -0.54\,^{+0.17}_{-0.16}\,(\text{stat}) \pm 0.05\,(\text{syst})\,, \\ C_{D^+D^-} &= 0.26\,^{+0.18}_{-0.17}\,(\text{stat}) \pm 0.02\,(\text{syst})\,. \end{aligned}$$

With these results high-order Standard Model corrections, which could lead to differences between the absolute values of the obtained observables of the two decay modes, are constrained to be small.

Zusammenfassung

Für die Zerfallskanäle $B^0 \to J/\psi\, K^0_{\rm S}$ und $B^0 \to D^+D^-$ werden die Observablen S_f und C_f bestimmt, welche die *CP*-Verletzung in diesen Zerfallskanälen beschreiben. Dazu wird ein Datensatz verwendet, der einer integrierten Luminosität von 3 fb^{-1} entspricht. Aufgenommen wurde er am LHCb-Experiment in Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von 7 und 8 TeV.

Durch Analyse der Zeitentwicklung von 41 560 $B^0 \to J/\psi\, K^0_{\rm S}$ Zerfällen werden die *CP*-Observablen zu

$$\begin{aligned} S_{J/\psi\, K^0_{\rm S}} &= 0.731 \pm 0.035\,(\text{stat}) \pm 0.020\,(\text{syst})\,, \\ C_{J/\psi\, K^0_{\rm S}} &= -0.038 \pm 0.032\,(\text{stat}) \pm 0.005\,(\text{syst})\,, \end{aligned}$$

bestimmt. Dieses Ergebnis ist kompatibel mit den aktuellen Weltmittelwerten und mit den Erwartungen des Standardmodells.

In der zeitabhängigen Analyse von 1410 getaggten $B^0 \to D^+D^-$ Zerfällen werden folgende Werte gemessen:

$$\begin{aligned} S_{D^+D^-} &= -0.54\,^{+0.17}_{-0.16}\,(\text{stat}) \pm 0.05\,(\text{syst})\,, \\ C_{D^+D^-} &= 0.26\,^{+0.18}_{-0.17}\,(\text{stat}) \pm 0.02\,(\text{syst})\,. \end{aligned}$$

Aus dem Vergleich der Werte für S_f und C_f in den beiden Zerfallskanälen lässt sich ablesen, dass Korrekturen höherer Ordnung im Standardmodell klein sind.

Contents

1 Introduction

From latest measurements the age of our universe is calculated to be 13.8 billion years [1]. Theoretical models exist that describe the evolution of the universe during this incredibly long period of time, starting from directly after the Big Bang [2] until today. But although there are models describing phenomena ranging from the largest (cosmological models) down to the smallest scales (Standard Model of particle physics (SM)), a comprehensive model has not been found so far. One of the most striking observations is that we are surrounded by matter, while there are no antimatter clusters. Though, according to big bang theories at the origin of the universe the same amount of matter and antimatter must have been created. Sakharov has proposed three conditions [3] that need to be fulfilled to explain this so called baryogenesis: At some point in the evolution of the universe the thermal equilibrium must have been imbalanced, the baryon number conservation is required to be violated, and the C (and even CP) symmetry has to be violated as well. While baryon number violation has not been observed so far, *e.g.* the lifetime of the proton is determined to be greater than 5.9×10^{33} years [4], evidence for a violation of the thermal equilibrium has been found [5]. Moreover, already fifty years ago C-violating and shortly thereafter CP-violating processes have been discovered [6]. Still, there are reasons why even after so many years this topic is interesting. On the one hand, the size of CP violation in the SM is orders of magnitude below what is required to explain the matter-antimatter asymmetry [7]. On the other hand, CP-violating processes are an excellent test bed for the predictions in the quark-flavour sector of the Standard Model of particle physics. The unitarity of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which describes the probability of quark transitions, is a fundamental requisite of the SM. It can be tested by studying the unitarity triangle, which represents one of the unitarity conditions of the CKM matrix. The determination of the angle β of this triangle, or more precisely of the derived quantity $\sin 2\beta$, is the common theme of this thesis. Precision measurements of CP violation in charmonium decays, *i.e.* in decay modes involving a $c\bar{c}$ resonance, and in open charm decays, *i.e.* in decay modes with at least one hadron containing exactly one c quark, are performed. If the final state of a decay mode is accessible for B^0 and $\overline{B}^0$ mesons, CP violation can occur in the interference between the direct decay and the decay after mixing. This is the case in $B^0 \to J/\psi\, K^0_S$ and in $B^0 \to D^+ D^-$ decays, which are studied in tagged decay-time dependent analyses.

The data exploited in these analyses is collected at the Large Hadron Collider (LHC), where proton bunches are accelerated and collided at centre-of-mass energies of up to 13 TeV. In 2011 and 2012, referred to as Run I, a data sample corresponding to an integrated luminosity of 3 fb^{-1} has been recorded with the LHCb detector.

This data sample, collected at centre-of-mass energies of 7 and 8 TeV, is the world's largest sample of B^0 mesons.

From a theoretical point of view the decay mode $B^0 \to J/\psi K^0_S$ offers a very clean determination of $\sin 2\beta$, as the dominating contributions can be calculated perturbatively. Experimentally, the quite large branching fractions and the clear signature in the detector allow a measurement with very high precision. In contrast, the main purpose of the study of $B^0 \to D^+D^-$ decays is to constrain higher-order Standard Model corrections occurring in measurements of the *CP*-violating phase. These contributions need to be controlled to distinguish them from effects caused by physics beyond the Standard Model of particle physics, often referred to as "New Physics". There are several reasons to believe that the SM needs to be extended. The measurements of rotation curves of galaxies [8] have lead to the assumption of the presence of dark matter. This is not a small effect. Around 26 % of the energy density in the universe are assigned to dark matter compared to around 5 % for normal baryonic matter [1]. However, dark matter is not accounted for in the SM, and no dark matter candidate has been found so far. Furthermore, in the SM the neutrinos are set to be massless, which is disproved by the observation of neutrino oscillations [9, 10], awarded with the Nobel prize in 2015 [11].

The thesis is structured as follows: First, the basics of the Standard Model of particle physics are shortly introduced (see Ch. 2). In Ch. 3 a more detailed description of the origin and nature of *CP* violation is given as well as ways to measure it. The LHCb experiment, namely the detector and the associated software, is described in Ch. 4. Some relevant techniques applied in data analysis of high energy physics are presented in Ch. 5. After these prerequisites are introduced, the analysis strategies of the measurements of *CP* violation in $B^0 \to J/\psi K^0_S$ decays (see Ch. 6) and in $B^0 \to D^+D^-$ decays (see Ch. 7) are discussed. The results of the two analyses are compared with previous measurements and with each other in Ch. 8 and a summary of the outcome of this thesis is given in Ch. 9.

2 Standard Model of Particle Physics

The Standard Model of particle physics (SM) is a renormalisable gauge-invariant quantum field theory, which describes the fundamental constituents of matter (see Sec. 2.1) and three of the four fundamental interactions between them (see Sec. 2.2). It is structured by the conservation and (spontaneous) breaking of symmetries (see Sec. 2.3). Despite being very successful in explaining experimental data and even predicting their results, there is a number of issues showing that extensions of the SM are required (see Sec. 2.4). The information given in this chapter is mainly inspired by the description in Ref. [12] and based on explanations from Refs. [13] and [14].

2.1 Particles

In the SM 12 fermions, which are elementary particles with spin ½, and the same number of antifermions, which have the opposite charge-related quantum numbers, are described. The fermions are divided into six quarks and six leptons. The quarks are further subdivided into three generations, which each contain an up-type and a down-type quark. The common matter, protons and neutrons, is built up from the quarks of the first generation, the up quark (u) and the down quark (d). Their heavier partners are the charm (c) and the top quark (t) respectively the strange (s) and the bottom quark (b). Due to confinement [15] quarks are always part of bound states, so called hadrons (terminology introduced by L. B. Okun [16]). A quark and an antiquark form a meson, three quarks a baryon, and just recently evidence for the existence of four and five quark bound states (tetraquarks respectively pentaquarks) has been found [17–19]. A colour charge is associated to the quarks, which can take three different types. However, the colour charges add up in a way that the hadrons are colourless. The electric charge of the up-type quarks is $+⅔$ and of the down-type quarks $-⅓$ of the elementary charge. The leptons are also classified in three families, each consisting of a negatively charged particle, in increasing order of mass the electron (e), the muon (μ) and the tauon (τ), and a corresponding neutral neutrino, which is set to be massless in the SM. Additionally, 12 gauge bosons with integer spin, which mediate the forces (see Sec. 2.2), are described in the SM. The gauge bosons are the massless photon (γ), the three massive mediators of the weak force (W^+, W^-, Z^0), and the eight gluons, which carry different colour charge configurations. The SM has been completed by the

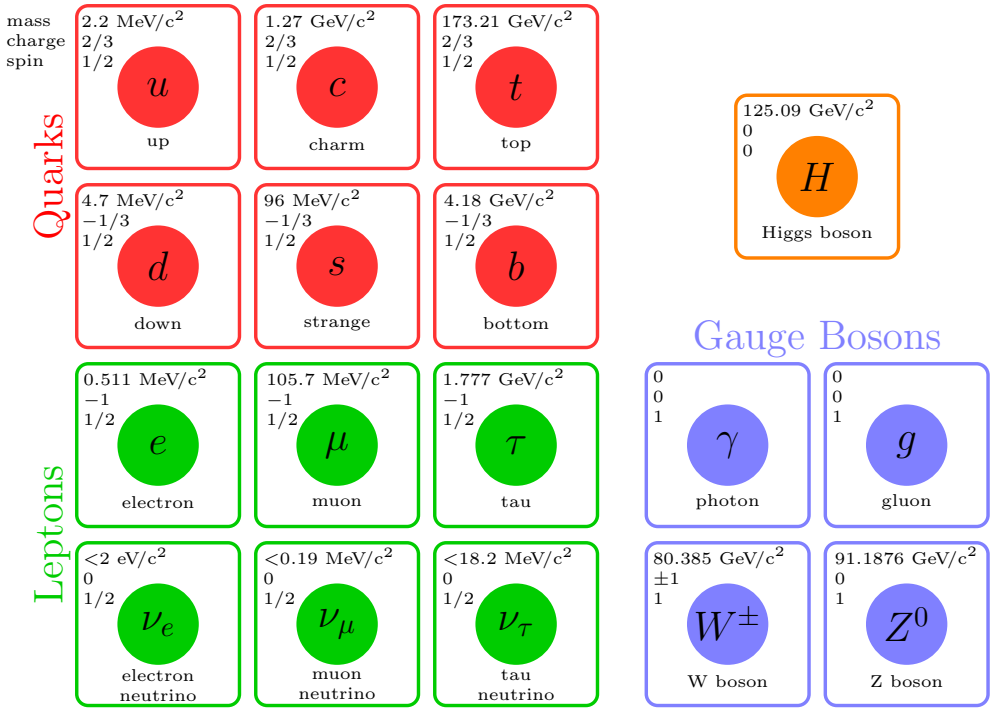

Figure 2.1: Summary of all SM particles. The values for the mass, the electromagnetic charge and the spin are taken from Ref. [22]. The corresponding antiparticles to the 12 fermions on the left have the same mass, but charges and spins of the opposite sign.

discovery of the Higgs boson [20, 21], a massive boson with spin 0. A summary of all SM particles, including their masses, charges and spins, is depicted in Fig. 2.1.

2.2 Forces and couplings

The three interactions that are described in the SM are the electromagnetic, the weak and the strong force. These differ dramatically in their strength and the time scales, in which the interactions proceed. Strong decays typically take 10^{-23} s, electromagnetic decays 10^{-16} s, while the decay time of weak interactions ranges from 10^{-13} s to a few minutes.

The classical approach of the electromagnetic interaction is given by Maxwell's equations. These are generalised into a relativistic quantum field theory by the Quantum Electrodynamics (QED) [23–25]. The QED can be derived from the Lagrangian of a free fermion field

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

by extending the global to a local $U(1)$ symmetry. This is done by replacing ∂_μ

with the corresponding covariant derivative D_μ

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

where the vector field A_μ can be identified as the photon, which mediates the electromagnetic force via the coupling to the electric charge. The dynamics is introduced by the kinetic term

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

with the field strength tensor $F_{\mu\nu}$, which is the compressed formulation of Maxwell's classical equations. However, in the SM the electromagnetic interaction is unified with the weak interaction in the electroweak $SU(2) \times U(1)$ symmetry group [26–28]. The weak part couples to the weak isospin and differs between left-handed and right-handed fermion fields, where handedness gives the orientation of the spin with respect to the momentum vector. Neutrinos have the peculiarity that they only exist as right-handed fermions. The electroweak gauge symmetry is broken, which becomes apparent, as the photon is massless, while the $W^\pm$ and Z^0 bosons are not. The Higgs mechanism [29] is responsible for this symmetry breaking, which comes along with the need for the massive Higgs boson. The masses of the quarks and leptons are generated through the Yukawa interactions between the Higgs and the fermion fields. However, the calculations do not contain any predictions for the coupling constants and thus for the numerical values of the masses. Moreover, the calculations show that the weak eigenstates of the down-type quarks (at least in the most common convention) are a superposition of the mass eigenstates, where the relation is given by the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix [30]. This topic is explained in more detail in Sec. 3.1.

The third fundamental force, called the strong force, is characterised by the Quantum Chromodynamics (QCD). It describes the binding between quarks and gluons, which are the mediators of this interaction, through the colour charge in a $SU(3)$ gauge symmetry group. The coupling heavily depends on the momentum scale, which in the renormalisation theory can be understood as a running of the coupling "constant" α_s. Gluon polarisation, which is possible as gluons carry colour charge by themselves and therefore can couple to each other, outperforms quark polarisation effects leading to asymptotic freedom of the quarks on very short distances [31, 32]. On the other hand, quarks can not separate too much from each other or — at least according to one possible scenario for confinement — a quark-antiquark pair is produced in between. Both effects can be summarised in the quark-antiquark potential

$$V_{\mathrm{QCD}} = -\frac{4}{3}\frac{\alpha_s}{r} + k\, r \,, \tag{2.4}$$

where r is the distance between the two fermions and $k \approx 1\,\mathrm{GeV\,fm^{-1}}$ [14].

2.3 Symmetries and conservation laws

Symmetries, global or local as well as continuous or discrete, structure the Standard Model. According to Noether's theorem every symmetry corresponds to a conservation law [33].

As a whole the SM is constructed to be invariant under gauge transformations, meaning that the physics is independent of the choice of the gauge. This symmetry is only broken by the vacuum expectation value of the Higgs field, which is the origin of the masses of the W and Z bosons. Gauge invariance is associated with the conservation of electric charge and colour charge. The invariance under space-time translation corresponds to the conservation of energy. However, Heisenberg's uncertainty principle [34] allows the violation of the conservation of energy for a very short period of time. This enables the existence of virtual, heavy particles in decay processes. The invariance of a system under translation in space and rotation leads to the conservation of momentum and angular momentum, respectively. Moreover, the baryon number (B) and the lepton number (L) are conserved respectively only broken by tiny non-pertubative effects. However, in the early universe these effects might have been larger [35], while (B − L) is an exact symmetry. The lepton family number, *i.e.* the individual lepton number for electrons, muons and tauons, also seems to be conserved, at least no significant asymmetry has been found yet. But there is no symmetry group evoking this conservation law and tests of the lepton universality by LHCb using $B^+ \to K^+\ell^+\ell^-$ decays show an asymmetry with a significance corresponding to 2.6 standard deviations [36]. While all the former symmetries are absolute, *i.e.* valid for all three interactions, there are also approximate symmetries that only apply for the electromagnetic and strong but not for the weak force. For example, flavour transitions are only possible in the weak interaction. Thus, flavour symmetry as well as the U-spin symmetry are approximate symmetries. The latter states that under the assumption that the masses of up, down and strange quarks are the same, processes are invariant under exchange of the two down-type quarks. This allows to transfer some findings from one decay mode to another, *e.g.* from decay modes of B^0 mesons to B_s^0 mesons [37, 38]. Another approximate symmetry is the parity operation (P), which performs a spatial inversion of all coordinates

$$P\Psi(r) = \Psi(-r)\,, \tag{2.5}$$

or in other words it transforms left-handed into right-handed fermions. In fact, the weak interaction even maximally violates parity, *e.g.* there are only left-handed neutrinos and right-handed antineutrinos. Charge conjugation (C) is another discrete symmetry. It changes the sign of all charges and the magnetic moment, and thus transforms particles into their antiparticles

$$C|p\rangle = |\bar{p}\rangle\,. \tag{2.6}$$

The combination of charge conjugation and parity (CP) is more stable, *e.g.* a left-handed neutrino becomes a right-handed antineutrino. Nevertheless, it is still

violated at the 10^{-4} level by the weak interaction. *CP* violation is explained in more detail in Ch. 3. Combining *CP* with time reversal (T), one of the most fundamental concepts of the SM is found, the *CPT* symmetry. The *CPT* theorem [39–41] states that particles and antiparticles have the same mass and the same lifetime.

2.4 Problems and possible extensions

Although the SM has proven to be a very successful and predictive theory, there are several issues that can not be explained in the SM and others that appeal very constructed. The latter leads to the idea of a more fundamental theory in which the SM is embedded. The concept of a unification has first been proposed by Georgi and Glashow [42]. A first step would be a generalisation of the electroweak with the strong interaction. Then, gravitation, the fourth fundamental force, could be included, whose effect is almost negligible at the energy scale, which is probed in today's high energy physic experiments, and therefore not part of the SM. Here, the difficulty is that even nowadays gravitation is still based on Einstein's general theory of relativity, so unlike the other theories of the SM not given in a quantum mechanical framework. The unification of the forces would probably emerge at energy scales of $10^{16}\,\mathrm{GeV}$. However, quantum corrections from those mass scales would heavily influence the Higgs mass, which is measured to be around $125\,\mathrm{GeV/c^2}$ [43]. In the SM this hierarchy problem is solved by a fine tuning of tree-level and loop contributions, which exactly cancel each other. Other explanations are given by extending the SM with new symmetries, like models including supersymmetry [44–46]. From a theoretical point of view it is also unsatisfactory that the SM includes so many free parameters, like the masses of the constituents or the number of generations. In addition, the SM only applies to the processes of ordinary matter, which makes only about 5 % of the total energy density in the universe [1], whereas it lacks an explanation for dark matter or dark energy. Furthermore, the amount of *CP* violation in the weak sector can not account for the baryon asymmetry in the universe, *i.e.* the dominance of matter without any large clusters of antimatter in the universe. The observation of neutrino oscillations [9, 10], which implicitly is an observation of mass differences between the neutrino generations, shows that neutrinos can not be massless as assumed in the SM. The special role of neutrinos in the SM to appear only left-handed could be corrected if they were Majorana particles, *i.e.* their own antiparticles [47].

3 *CP* Violation

In this chapter the concept of *CP* violation, its origin and manifestation in the SM, as well as the possibilities to measure *CP* violation, are described. The formalism closely follows Refs. [48] and [49].

3.1 The KM mechanism and the CKM matrix

Quarks get their mass through coupling to the Higgs field with vacuum expectation value v and via Yukawa interaction between the left-handed and the right-handed quark content. The Yukawa matrices $\mathbf{Y}_d$ and $\mathbf{Y}_u$ for down-type and up-type quarks involved in the corresponding Lagrangian

$$\mathcal{L}_{\text{Yukawa}} = -\frac{v}{\sqrt{2}}(\bar{d}_{\text{L}}\mathbf{Y}_d d_{\text{R}} + \bar{u}_{\text{L}}\mathbf{Y}_u u_{\text{R}}) + \text{h.c.} \tag{3.1}$$

are not necessarily diagonal. The mass eigenstates q' can be obtained by a unitary transformation

$$q'_{\text{A}} = \mathbf{V}_{\text{A},q} q_{\text{A}} \quad \text{for } q = u, d \text{ and } \text{A} = \text{L}, \text{R} \tag{3.2}$$

with $\mathbf{V}_{\text{A},q}\mathbf{V}^{\dagger}_{\text{A},q} = 1$. When applying this transformation in the Lagrangian that describes the charged-current interaction

$$\mathcal{L}_{\text{CC}} = -\frac{g_2}{\sqrt{2}}(\bar{u}_{\text{L}}\gamma^{\mu}W^{+}_{\mu}d_{\text{L}} + \bar{d}_{\text{L}}\gamma^{\mu}W^{-}_{\mu}u_{\text{L}}) \tag{3.3}$$

$$= -\frac{g_2}{\sqrt{2}}(\bar{u}'_{\text{L}}\gamma^{\mu}W^{+}_{\mu}\mathbf{V}_{\text{L},u}\mathbf{V}^{\dagger}_{\text{L},d}d'_{\text{L}} + \bar{d}'_{\text{L}}\gamma^{\mu}W^{-}_{\mu}\mathbf{V}_{\text{L},d}\mathbf{V}^{\dagger}_{\text{L},u}u'_{\text{L}}) \tag{3.4}$$

the Cabibbo-Kobayashi-Maskawa matrix $\mathbf{V}_{\text{CKM}} = \mathbf{V}_{\text{L},u}\mathbf{V}^{\dagger}_{\text{L},d}$ enters. As the Yukawa matrices are not diagonalised by the same unitary transformation, the CKM matrix is not the unit matrix and thus allows for flavour changes through the weak interaction. Thus, the CKM matrix can be understood as the connection between the mass eigenstates and the eigenstates to the weak interaction

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{3.5}$$

Being the product of two unitary matrices the CKM matrix itself is unitary as well. In general, a complex 3×3 matrix has 18 free parameters. However, the

unitarity removes nine degrees of freedom. Another five phases can be constrained by global rephasings between the six mass fields. Thus, four free parameters remain, of which three are real-valued angles and one is a complex phase. This single phase introduces *CP* violation to the SM. Kobayashi and Maskawa developed this concept, which explains the origin of *CP* violation and predicted the existence of the third quark generation [30]. The corresponding parametrisation of the CKM matrix is

$$\mathbf{V}_{\mathrm{CKM}} = \begin{pmatrix} c_1 & -s_1c_3 & -s_1s_3 \\ s_1c_2 & c_1c_2c_3 - s_2s_3e^{i\delta} & c_1c_2s_3 + s_2c_3e^{i\delta} \\ s_1s_2 & c_1s_2c_3 + c_2s_3e^{i\delta} & c_1s_2s_3 - c_2c_3e^{i\delta} \end{pmatrix} , \tag{3.6}$$

with c_i and s_i being shorthand for the cosine respectively sine of the three Euler angles, and δ being the irreducible phase. Tests of the SM concerning *CP* violation in the quark mixing sector are performed by examining the unitarity conditions of the CKM matrix. Six of the 12 equations are orthogonality relations, which can be interpreted as triangles in the complex plane. The area of all triangles is the same and given by half of the Jarlskog invariant

$$J_{CP} = \pm \mathcal{I}m(V_{ik}V_{jl}V_{il}^*V_{jk}^*) \quad (i \neq j, l \neq k) , \tag{3.7}$$

which expresses the amount of *CP* violation in the SM [50]. It is measured to be $J = (3.04^{+0.21}_{-0.20}) \times 10^{-5}$ [22]. However, the ratios of the side lengths of the unitarity triangles are very different. In two of them all sides are of comparable length, one of the conditions is given by

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 . \tag{3.8}$$

When depicting this triangle in the complex plane, it is convenient to scale the triangle by dividing all side lengths by $V_{cd}V_{cb}^*$. Then, the base matches the real axis with length one (see Fig. 3.1). Using the parametrisation of the CKM matrix by

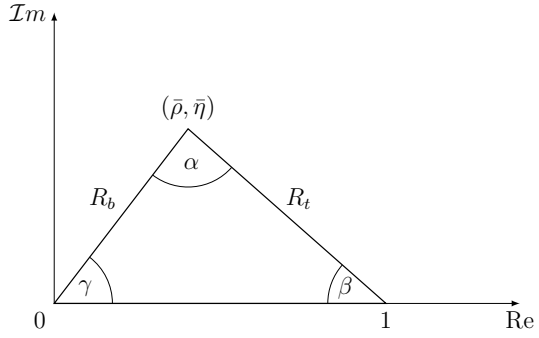

Figure 3.1: Schematic representation of the CKM unitarity triangle.

Wolfenstein [51], which is an expansion in powers of $\lambda \equiv |V_{us}| = 0.2248 \pm 0.0006$ [22],

$$\mathbf{V}_{\mathrm{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) , \tag{3.9}$$

the other two sides are given by

$$R_b = \left(1 - \frac{\lambda^2}{2}\right) \frac{1}{\lambda} \left|\frac{V_{ub}}{V_{cb}}\right| = \sqrt{\bar{\rho}^2 + \bar{\eta}^2}\,, \tag{3.10}$$

$$R_t = \frac{1}{\lambda} \left|\frac{V_{td}}{V_{cb}}\right| = \sqrt{(1-\bar{\rho})^2 + \bar{\eta}^2}\,, \tag{3.11}$$

where $\bar{\rho}$ and $\bar{\eta}$ define the position of the apex and are related to the Wolfenstein parameters through

$$\bar{\rho} = \rho(1 - \lambda^2/2) \quad \text{and} \quad \bar{\eta} = \eta(1 - \lambda^2/2)\,. \tag{3.12}$$

The three angles of the unitarity triangle are defined by

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{3.13}$$

The unitarity triangle is overconstrained, *i.e.* there are measurements of more independent parameters than necessary to fully characterise the shape of the triangle. The angle α can be studied with $B^0 \to \pi\pi$ decays [52–54], β is precisely measured using the decay-time-dependent *CP* asymmetry in $B^0 \to J/\psi\, K^0_S$ decays (see Sec. 3.4), and γ can be extracted from a combination of results in $B \to Dh$ decays [55]. Semileptonic b-hadron decays are used to determine the size of the triangle side R_b. Further information on $|V_{ub}|$ comes from studies of $B^+ \to \tau^+\nu_\tau$ decays [56–58]. The second non-trivial side length R_t is constrained by measurements of the mixing frequencies Δm_d and Δm_s in the system of neutral B^0 and B^0_s mesons [59]. Furthermore, information on the position of the apex can be gained from the measurement of *CP* violation in the neutral kaon system [22]. All these inputs are put into a global fit, which mainly checks how well the different constraints agree on the position of the apex. The latest result of the CKMfitter group in Fig. 3.2 shows a very good agreement of all present tests of *CP* violation in the SM, as the area for the position of the apex is relatively small.

3.2 The system of neutral B^0 mesons

In the system of neutral B^0 mesons four decay amplitudes occur. The decay amplitude A_f stands for the decay of a B^0 meson into a final state f, while $\bar{A}_f$ is the decay amplitude of an $\overline{B}^0$ meson into the same final state. Similarly, the decay amplitudes into the *CP*-conjugated final state $\bar{f}$ can be defined. By convention B^0 mesons consist of $\bar{b}$ and d quarks, while $\overline{B}^0$ mesons contain b and $\bar{d}$ quarks. These two mesons can mix, *i.e.* they can oscillate between the two flavour states. As flavour-changing neutral currents (FCNC) are forbidden in the SM, the B^0–$\overline{B}^0$ oscillation is in lowest order Standard Model described by quantum loops involving charged currents, as shown in the Feynman diagrams in Fig. 3.3. The corresponding

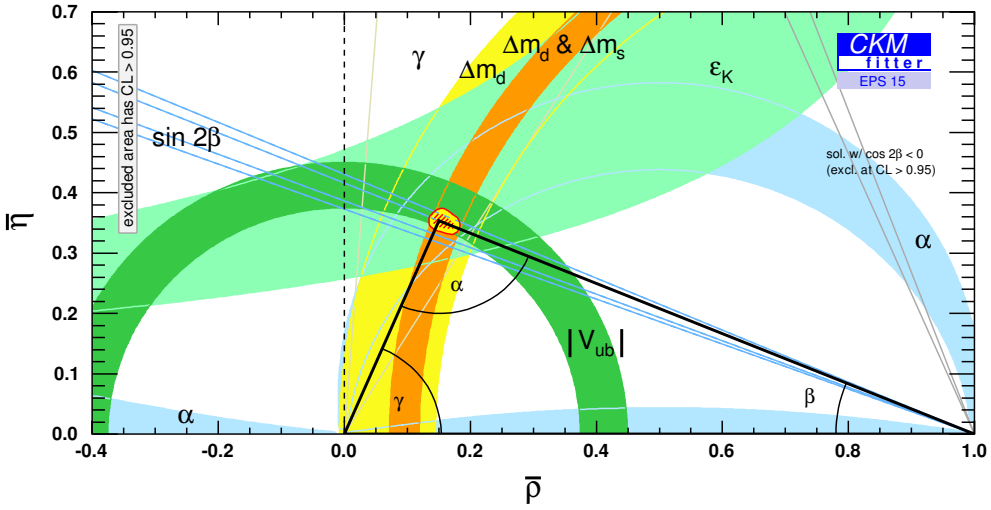

Figure 3.2: Unitarity triangle with constraints from measurements of various quantities [60].

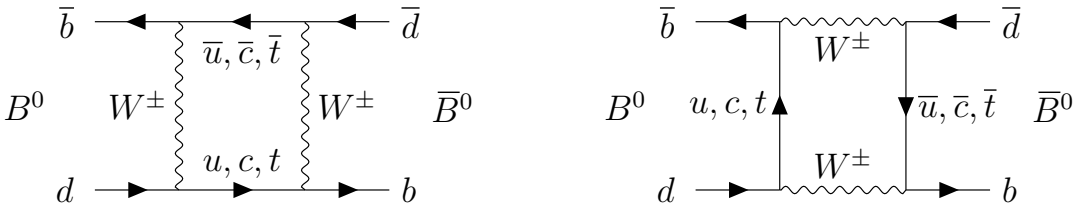

Figure 3.3: Box diagrams of B^0–$\overline{B}^0$ oscillation.

decay amplitude is given by

$$\langle B^0|H_{\text{eff}}|\overline{B}^0\rangle = \mathcal{A}(\overline{B}^0 \to B^0) = \frac{g^4}{6\,\pi^2\,M_W^2}(V_{tb}V_{td}^*)^2 B_{B^0} F_{B^0}^2 m_{B^0}^2 \eta_{2B} S(x_t)\,, \qquad (3.14)$$

with the coupling constant of the weak interaction g, the masses of the W boson M_W and the B^0 meson m_{B^0}, the bag parameter for the B^0 meson B_{B^0}, the weak decay constant F_{B^0}, the short-distance QCD correction factor η_{2B} [61], and the Inami-Lim function [62]

$$S(x_t) = x_t\left(\frac{1}{4} + \frac{9}{4}\frac{1}{1-x_t} - \frac{3}{2}\frac{1}{(1-x_t)^2}\right) + \frac{3}{2}\left(\frac{x_t}{x_t-1}\right)^3 \log x_t\,, \qquad (3.15)$$

where x_t is the squared fraction of the top quark mass with respect to the W boson mass. Here, it is accounted for the suppression of contributions with up and charm quarks in the loop due to $m_{u,c}^2 \ll m_t^2$.

To derive the time evolution of initially produced B^0 and $\overline{B}^0$ mesons the Schrödinger equation [63] needs to be solved. Assuming the Wigner-Weisskopf

approximation [64], *i.e.* (excited) states do not know about their past, which is valid since the time scale of weak decays is significantly larger than the time scale of the production via the strong force, the effective Schrödinger equation for the wave function representing the system of B^0 and $\overline{B}^0$ mesons can be written as

$$i\frac{\mathrm{d}}{\mathrm{dt}}\begin{pmatrix}|B^0(t)\rangle\\|\overline{B}^0(t)\rangle\end{pmatrix} = \mathbf{H}\begin{pmatrix}|B^0(t)\rangle\\|\overline{B}^0(t)\rangle\end{pmatrix} = \left(\mathbf{M} - i\frac{\mathbf{\Gamma}}{2}\right)\begin{pmatrix}|B^0(t)\rangle\\|\overline{B}^0(t)\rangle\end{pmatrix}. \tag{3.16}$$

The Hamiltonian $\mathbf{H}$ is given by a non-Hermitian matrix, otherwise only oscillations but no decays would occur. It consists of Hermitian 2×2 mass $\mathbf{M}$ and decay matrices $\mathbf{\Gamma}$, which have contributions from virtual intermediate states respectively from physical final states accessible by both B^0 and $\overline{B}^0$ mesons. Due to the *CPT* theorem the diagonal elements of $\mathbf{M}$ and $\mathbf{\Gamma}$ are equal, *i.e.* $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$. The non-zero off-diagonal elements, else there would be no mixing, cause that the flavour eigenstates B^0 and $\overline{B}^0$ are not mass eigenstates. Instead, the light (L) and heavy (H) mass eigenstates are given by the linear combinations

$$\begin{aligned}|B_L\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle\,,\\ |B_H\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle\,,\end{aligned} \tag{3.17}$$

with the complex coefficients p and q, which fulfil the normalisation condition $|p|^2 + |q|^2 = 1$ and whose ratio can be expressed with the matrix elements as

$$\frac{q}{p} = \sqrt{\frac{2M_{12}^* - i\Gamma_{12}^*}{2M_{12} - i\Gamma_{12}}}\,. \tag{3.18}$$

Explicit calculations of Γ_{12}, as performed in Ref. [65], show that to a very good approximation Eq. (3.18) can be simplified to

$$\frac{q}{p} \approx \sqrt{\frac{M_{12}^*}{M_{12}}} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\,. \tag{3.19}$$

The well-defined masses and decay widths $m_{L,H}$ and $\Gamma_{L,H}$ of B_L and B_H lead to the eigenvalues

$$\begin{aligned}\mu_L &= m_L - {}^{i}\!/_{2}\Gamma_L\,,\\ \mu_H &= m_H - {}^{i}\!/_{2}\Gamma_H\,,\end{aligned} \tag{3.20}$$

which can be used to shortly express the time-evolution of the mass eigenstates

$$|B_{L,H}(t)\rangle = e^{-i\mu_{L,H}t}|B_{L,H}\rangle\,. \tag{3.21}$$

Using Eq. (3.17) the time evolution of B mesons purely produced as B^0 or $\overline{B}^0$ can be derived

$$\begin{aligned}|B^0(t)\rangle &= g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)|\overline{B}^0\rangle\,,\\ |B^0(t)\rangle &= g_+(t)|\overline{B}^0\rangle + \frac{p}{q}g_-(t)|B^0\rangle\,,\end{aligned} \tag{3.22}$$

with the time-dependent coefficients

$$\begin{aligned} g_+(t) &= e^{-i\frac{\mu_L+\mu_H}{2}t}\left[+\cosh\frac{\Delta\Gamma_d\,t}{4}\cos\frac{\Delta m_d\,t}{2} - i\sinh\frac{\Delta\Gamma_d\,t}{4}\sin\frac{\Delta m_d\,t}{2}\right], \\ g_-(t) &= e^{-i\frac{\mu_L+\mu_H}{2}t}\left[-\sinh\frac{\Delta\Gamma_d\,t}{4}\cos\frac{\Delta m_d\,t}{2} + i\cosh\frac{\Delta\Gamma_d\,t}{4}\sin\frac{\Delta m_d\,t}{2}\right]. \end{aligned} \tag{3.23}$$

The mass difference $\Delta m_d = m_H - m_L$ drives the oscillation and can be interpreted as mixing frequency, while $\Delta\Gamma_d = \Gamma_H - \Gamma_L$ is negligible in the B^0 meson sector [59].

The differential decay rates, *i.e.* the squared decay amplitudes, can be expressed by combining the terms in Eqs. (3.22) and (3.23) as

$$\begin{aligned} \Gamma(B^0(t)\to f) &= |A_f|^2\Big\{|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + 2\,Re[\lambda_f g_+^*(t)g_-(t)]\Big\}, \\ \Gamma(\overline{B}^0(t)\to f) &= |A_f|^2\left|\frac{p}{q}\right|^2\Big\{|g_-(t)|^2 + |\lambda_f|^2|g_+(t)|^2 + 2\,Re[\lambda_f g_+(t)g_-^*(t)]\Big\}, \end{aligned} \tag{3.24}$$

using

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}. \tag{3.25}$$

3.3 Types of *CP* violation

There are three different manifestations of *CP* violation. It can occur, when the decay amplitudes differ between *CP* conjugated processes (see Sec. 3.3.1), when the mass eigenstates are no *CP* eigenstates (see Sec. 3.3.2), and when there is interference between direct decays and decays to the same final state after mixing (see Sec. 3.3.3). While the first type can appear for charged and neutral hadrons, the latter are only possible for neutral decays.

All types of *CP* violation can be summarised with the condition $\lambda_f \neq 1$.

3.3.1 Direct *CP* violation

Two different type of phases can contribute to decay amplitudes, weak phases and strong phases. Weak phases can enter through the CKM matrix and take the opposite sign for A_f and $\bar{A}_{\bar{f}}$. Strong phases typically appear in scattering processes and originate from intermediate on-shell states. They occur with the same sign in A_f and $\bar{A}_{\bar{f}}$. However, only phase differences are physically meaningful, as the SM is a gauge-invariant theory and thus absolute phases could be removed by a rotation of the system. So, at least two terms with different weak and strong phases need to contribute to the decay amplitudes to have an effect. The superposition of several

contributions with individual magnitudes A_i, weak phases $e^{i\phi_i}$ and strong phases $e^{i\delta_i}$ leads to

$$\begin{aligned} A_f &= \sum_i A_i e^{i(\delta_i+\phi_i)}\,, \\ \bar{A}_{\bar{f}} &= e^{2i(\xi_f-\xi_B)} \sum_i A_i e^{i(\delta_i-\phi_i)}\,, \end{aligned} \tag{3.26}$$

where ξ_f and ξ_B are arbitrary phases coming from the *CP* transformation on the B^0 meson and the final state, respectively. If the final state f is a *CP* eigenstate, the term $e^{2i\xi_f} = \pm 1$ represents the *CP* eigenvalue. Direct *CP* violation is present for

$$\left|\frac{\bar{A}_{\bar{f}}}{A_f}\right| = \left|\frac{\sum A_i e^{i(\delta_i-\phi_i)}}{\sum A_i e^{i(\delta_i+\phi_i)}}\right| \neq 1\,. \tag{3.27}$$

This type of *CP* violation is observed in charmless two-body decays of neutral B mesons [66–68].

3.3.2 Indirect *CP* violation

Indirect *CP* violation occurs when the mass eigenstates are no *CP* eigenstates and instead a relative phase is present between M_{12} and Γ_{12}. Following Eq. (3.18) this means

$$\left|\frac{q}{p}\right| \neq 1\,. \tag{3.28}$$

So, it can be interpreted as difference of the mixing probabilities between B^0 and $\overline{B}^0$ mesons

$$\mathcal{P}(B^0 \to \overline{B}^0, t) \neq \mathcal{P}(\overline{B}^0 \to B^0, t)\,, \tag{3.29}$$

and thus is also called *CP* violation in mixing. While this type of *CP* violation has been observed in the system of neutral kaons, it is expected to be very small in the system of neutral B mesons. All measurements of the asymmetry of semileptonic decays

$$a_{\mathrm{sl}} = \frac{\Gamma(\overline{B}^0(t) \to \ell^+\nu X) - \Gamma(B^0(t) \to \ell^-\nu X)}{\Gamma(\overline{B}^0(t) \to \ell^+\nu X) + \Gamma(B^0(t) \to \ell^-\nu X)} = \frac{1-|q/p|^4}{1+|q/p|^4} \tag{3.30}$$

yield values consistent with zero [69, 70], though the precision is one to two magnitudes above the SM expectations. This lack in precision is not only due to statistical uncertainties. When extracting the *CP* asymmetry from the raw asymmetry further asymmetries, like detection and production asymmetries, need to be taken into account, and these are not precisely known. Nevertheless, these experimental findings confirm the validity of the approximation leading to Eq. (3.19).

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

Even if there is no direct or indirect *CP* violation, it is possible that *CP* violation occurs, in the interference between decay amplitudes with and without mixing. However, the final state has to be a *CP* eigenstate, *i.e.* accessible for decaying B^0 and $\overline{B}^0$ mesons. The definition from Eq. (3.25) slightly changes to

$$\lambda = \eta_{CP} \frac{q}{p} \frac{\bar{A}_{\bar{f}}}{A_f} \,. \tag{3.31}$$

Here, η_{CP} is the *CP* eigenvalue of the final state

$$CP|f_{CP}\rangle = |\bar{f}_{CP}\rangle = \eta_{CP}|f_{CP}\rangle = \pm 1|f_{CP}\rangle \,. \tag{3.32}$$

The condition that any deviation from unity for λ indicates *CP* violation holds. Indirect *CP* violation ($|q/p| \neq 1$) and direct *CP* violation ($|\bar{A}_{\bar{f}}/A_f| \neq 1$) affect the magnitude of λ, while *CP* violation in the interference is associated with

$$\mathcal{I}m\, \lambda \neq 0 \,. \tag{3.33}$$

The decay-time-dependent asymmetry

$$\mathcal{A}(t) \equiv \frac{\Gamma(\overline{B}^0(t) \to f_{CP}) - \Gamma(B^0(t) \to f_{CP})}{\Gamma(\overline{B}^0(t) \to f_{CP}) + \Gamma(B^0(t) \to f_{CP})} \tag{3.34}$$

can be used to measure *CP* violation in the interference of decay and decay after mixing. It compares the decay rates of initial ($t = 0$) $\overline{B}^0$ and B^0 mesons. Plugging in the expressions from Eq. (3.24) and using

$$\begin{aligned} |g_\pm(t)|^2 &= \frac{e^{-\Gamma t}}{2} \left[\cosh \frac{\Delta\Gamma_d\, t}{2} \pm \cos(\Delta m_d\, t) \right] , \\ g_+^*(t) g_-(t) &= \frac{e^{-\Gamma t}}{2} \left[-\sinh \frac{\Delta\Gamma_d\, t}{2} - i \sin(\Delta m_d\, t) \right] , \end{aligned} \tag{3.35}$$

$\mathcal{A}$ can be written as

$$\mathcal{A}(t) = \frac{2\,\mathcal{I}m\, \lambda \sin(\Delta m_d\, t) - (1 - |\lambda|^2)\cos(\Delta m_d\, t)}{(1 + |\lambda|^2)\cosh(\frac{\Delta\Gamma_d\, t}{2}) + 2\, Re\, \lambda \sinh(\frac{\Delta\Gamma_d\, t}{2})} \,. \tag{3.36}$$

It is apparent that this asymmetry only vanishes if either $|\lambda| \neq 1$ (direct or indirect *CP* violation) or if λ has an imaginary part, which is the condition for *CP* violation stated in Eq. (3.33). Defining

$$S_f = \frac{2\,\mathcal{I}m\, \lambda}{1 + |\lambda|^2} \quad \text{and} \quad C_f = \frac{1 - |\lambda|^2}{1 + |\lambda|^2} \,, \tag{3.37}$$

and neglecting $\Delta\Gamma_d$, the decay-time-dependent asymmetry simplifies to

$$\mathcal{A}(t) = S_f \sin \Delta m_d\, t - C_f \cos \Delta m_d\, t \,. \tag{3.38}$$

3.4 *CP* violation in $b \to c\bar{c}s$ decays

The gold-plated mode to measure *CP* violation in the system of neutral B mesons is $B^0 \to J/\psi K^0_S$. It proceeds via a Cabibbo-favored $b \to c\bar{c}s$ transition. Direct and indirect *CP* violation is strongly suppressed, which makes it a very clean mode to determine the weak mixing phase ϕ_d, and thus the CKM angle β, via *CP* violation in the interference of decay and decay after mixing. As the Feynman diagrams in Fig. 3.4 show, actually $B^0 \to J/\psi K^0$ and $\overline{B}^0 \to J/\psi \overline{K}^0$ decays take place. However,

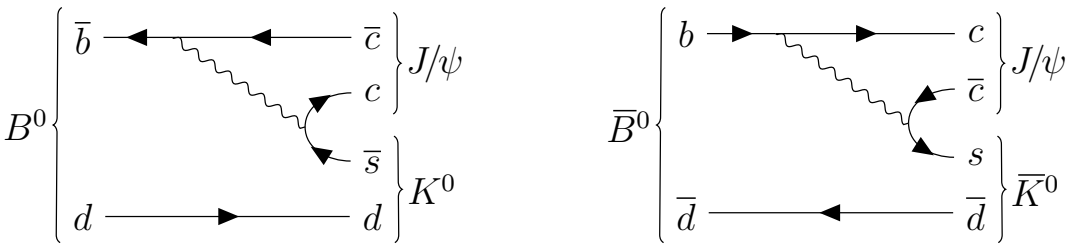

Figure 3.4: Tree Feynman diagrams of $B^0 \to J/\psi K^0_S$ for both flavours.

like for the B mesons the flavour eigenstates of the kaons are a superposition of the *CP* mass eigenstates, thus

$$|K^0_S\rangle = p_K|K^0\rangle - q_K|\overline{K}^0\rangle\,. \tag{3.39}$$

Therefore, the ratio of decay amplitudes is composed of two terms according to

$$\frac{\bar{A}_{J/\psi K^0_S}}{A_{J/\psi K^0_S}} = -\frac{p_K}{q_K}\frac{\bar{A}_{J/\psi \overline{K}^0}}{A_{J/\psi K^0}}\,. \tag{3.40}$$

The ratio of the mixing coefficients for the kaons can be calculated using Eq. (3.18). Different than for the B mesons the dominant contribution to the mixing diagrams arises from charm quarks in the loop:

$$\frac{p_K}{q_K} = -\frac{V_{cs}V^*_{cd}}{V^*_{cs}V_{cd}}\,. \tag{3.41}$$

Accounting only for the tree diagrams in Fig. 3.4, while neglecting loop processes, the ratio of the direct decay amplitudes can be expressed via the involved CKM matrix elements:

$$\frac{\bar{A}_{J/\psi \overline{K}^0}}{A_{J/\psi K^0}} = \frac{V_{cb}V^*_{cs}}{V^*_{cb}V_{cs}}\,. \tag{3.42}$$

Summarising these values and adding the ratio of CKM matrix elements for the mixing of the B^0 mesons (see Eq. (3.19)) the parameter describing *CP* violation from Eq. (3.25) becomes

$$\lambda_{J/\psi K^0_S} = -\frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}\frac{V_{cs}V^*_{cd}}{V^*_{cs}V_{cd}}\frac{V_{cb}V^*_{cs}}{V^*_{cb}V_{cs}} = -\frac{V^*_{tb}V_{td}}{V_{tb}V^*_{td}}\frac{V^*_{cd}V_{cb}}{V_{cd}V^*_{cb}} \tag{3.43}$$

The minus sign indicates that the final state $J/\psi K_S^0$ is CP-odd, as an angular momentum of $l = 1$ is necessary to compensate that the B^0 meson as initial state has spin zero, while the final state consists of a CP-even J/ψ meson with spin one and an almost CP-even[1] K_S^0 meson with spin zero. Due to only considering the dominant tree contribution the absolute value of $\lambda_{J/\psi K_S^0}$ is unity. Therefore, the CP observables from Eq. (3.37) become

$$C_{J/\psi K_S^0} = 0\,, \tag{3.44}$$

and

$$\begin{aligned} S_{J/\psi K_S^0} &= \mathcal{I}m\lambda_{J/\psi K_S^0} = \sin\left[\arg\left(-\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}\frac{V_{cd}^* V_{cb}}{V_{cd} V_{cb}^*}\right)\right] \\ &= \sin\left[\arg\left(-\left(\frac{V_{cd}^* V_{cb}}{V_{tb} V_{td}^*}\right)^2\right)\right] = \sin 2\beta\,. \end{aligned} \tag{3.45}$$

This shows that the measurement of CP violation in $B^0 \to J/\psi K_S^0$ offers a clean opportunity to determine the CKM triangle angle β. Studying other decay channels with $b \to c\bar{c}s$ transitions, like $B^0 \to J/\psi K_L^0$ or $B^0 \to \psi(2S) K_S^0$, further improves the sensitivity on β.

However, Eq. (3.45) is only valid in the SM, where $\phi_d = 2\beta$, while in the presence of New Physics effects it has to be extended to

$$S_{J/\psi K_S^0} = \sin(\phi_d + \phi_d^{\mathrm{NP}})\,. \tag{3.46}$$

The corresponding golden decay channel with an s quark as spectator quark is $B_s^0 \to J/\psi\,\phi$. In principle, this provides a clean measurement of the mixing phase ϕ_s of the B_s^0 meson system. However, to extract the CP observable an angular analysis is required to disentangle the different spin configurations of the two vector mesons in the final state.

3.5 CP violation in $b \to c\bar{c}d$ decays

The decay $B^0 \to D^+ D^-$ can be described with the Feynman diagrams in Fig. 3.5. The tree diagram (T) proceeds via a $b \to c\bar{c}d$ quark transition, which is CKM suppressed. The contributions from the other diagrams, especially the penguin diagrams ($P^{(q)}$ with $q = u$, c and t quarks in the loop), but also exchange (E) and penguin annihilation diagrams ($PA^{(q)}$), need to be taken into account as well, because they can carry different weak phases and are not Cabibbo-suppressed. Thus, the decay amplitude is given by [71–73]

$$\begin{aligned} A(B^0 \to D^+ D^-) = {} & V_{cb}^* V_{cd} T + V_{tb}^* V_{td} P^{(t)} + V_{cb}^* V_{cd} P^{(c)} + V_{ub}^* V_{ud} P^{(u)} \\ & + V_{cb}^* V_{cd} E + V_{cb}^* V_{td} PA^{(t)} + V_{cb}^* V_{cd} PA^{(c)} + V_{ub}^* V_{ud} PA^{(u)}\,. \end{aligned} \tag{3.47}$$

[1] When reconstructed in a pair of two pions it is fully CP-even.

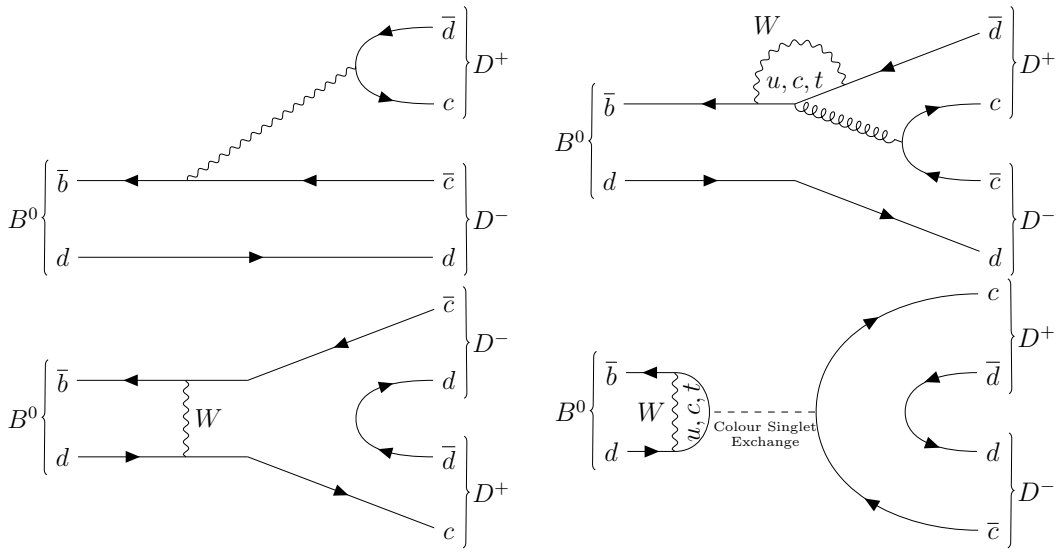

Figure 3.5: Main Feynman diagrams contributing to $B^0 \to D^+D^-$ decays. Apart from the tree diagram (top left), a penguin diagram (top right), an exchange diagram (bottom left) and a penguin annihilation diagram (bottom right) are shown.

Using the CKM unitarity condition

$$V_{tb}^* V_{td} + V_{cb}^* V_{cd} + V_{ub}^* V_{ud} = 0 \quad \Leftrightarrow \quad V_{tb}^* V_{td} = -V_{cb}^* V_{cd} - V_{ub}^* V_{ud} \tag{3.48}$$

the decay amplitude can be written as

$$\begin{aligned} A(B^0 \to D^+D^-) = V_{cb}^* V_{cd} \left(T + E + \left\{P^{(c)} + PA^{(c)}\right\} - \left\{P^{(t)} + PA^{(t)}\right\}\right) \\ + V_{ub}^* V_{ud} \left(\left\{P^{(u)} + PA^{(u)}\right\} - \left\{P^{(t)} + PA^{(t)}\right\}\right) . \end{aligned} \tag{3.49}$$

With the definition

$$\mathcal{A} \equiv \left[T + E + \left\{P^{(c)} + PA^{(c)}\right\} - \left\{P^{(t)} + PA^{(t)}\right\}\right] , \tag{3.50}$$

this can be further simplified to

$$A(B^0 \to D^+D^-) = V_{cb}^* V_{cd} \mathcal{A} \left[1 + \frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}} \frac{\left\{P^{(u)} + PA^{(u)}\right\} - \left\{P^{(t)} + PA^{(t)}\right\}}{\mathcal{A}}\right] . \tag{3.51}$$

Applying the Wolfenstein parametrisation and plugging in the definitions of the side length R_b (see Eq. (3.10)) and of the angle γ of the unitarity triangle (see Eq. (3.13))

$$\begin{aligned} \frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}} &= -\left(1 - \frac{\lambda^2}{2}\right) \frac{1}{\lambda} \left|\frac{V_{ub}}{V_{cb}}\right| \arg\left(\frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}}\right) \\ &= -R_b e^{-i\gamma} , \end{aligned} \tag{3.52}$$

the decay amplitude becomes

$$A(B^0 \to D^+D^-) = V_{cb}^* V_{cd} \mathcal{A}[1 - ae^{i\theta}e^{-i\gamma}]\,, \tag{3.53}$$

with

$$ae^{i\theta} \equiv R_b \left[\frac{\{P^{(u)} + PA^{(u)}\} - \{P^{(t)} + PA^{(t)}\}}{T + E + \{P^{(c)} + PA^{(c)}\} - \{P^{(t)} + PA^{(t)}\}}\right]. \tag{3.54}$$

While γ is a *CP*-violating weak phase, a and θ are hadronic *CP*-conserving parameters. Therefore, the corresponding $\overline{B}^0$ decay amplitude is

$$A(\overline{B}^0 \to D^+D^-) = V_{cb} V_{cd}^* \mathcal{A}[1 - ae^{i\theta}e^{i\gamma}]\,. \tag{3.55}$$

The parameter describing *CP* violation in the interference can be written as

$$\begin{aligned} \lambda_{D^+D^-} &= \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \frac{V_{cb} V_{cd}^*}{V_{cb}^* V_{cd}} \frac{1 - ae^{i\theta}e^{i\gamma}}{1 - ae^{i\theta}e^{-i\gamma}}\,, \\ &= e^{-i2\beta} \frac{1 - ae^{i\theta}e^{i\gamma}}{1 - ae^{i\theta}e^{-i\gamma}}\,, \end{aligned} \tag{3.56}$$

using the ratio of the mixing coefficients from Eq. (3.19) and the definition of the unitarity triangle β from Eq. (3.13). Different than for $B^0 \to J/\psi\, K_S^0$ (cf. Eq. (3.43)) the *CP* eigenvalue is $\eta_{CP} = +1$, since no angular momenta are involved in the decay $B^0 \to D^+D^-$. The hadronic parameters cannot be calculated reliably within QCD [73]. Thus, they must be determined through a measurement of the *CP* observables, which can be expressed via

$$C_{D^+D^-} = \frac{1 - |\lambda|^2}{1 + |\lambda|^2} = \frac{2a \sin\theta \sin\gamma}{1 - 2a\cos\theta\cos\gamma + a^2}\,, \tag{3.57}$$

$$S_{D^+D^-} = \frac{2\,\mathcal{I}m\,\lambda}{1 + |\lambda|^2} = -\frac{\sin\phi_d - 2a\cos\theta\sin(\phi_d + \gamma) + a^2\sin(\phi_d + 2\gamma)}{1 - 2a\cos\theta\cos\gamma + a^2}\,. \tag{3.58}$$

Due to interferences between tree and penguin contributions $C_{D^+D^-}$ might differ from zero. The term $S_{D^+D^-}$, which is caused by interference between the direct decay and the decay after mixing, gives access to the mixing phase ϕ_d. However, different than in the case of $B^0 \to J/\psi\, K_S^0$ decays only an effective phase

$$\phi_d^{\text{eff}} = \phi_d + \Delta\phi_d \tag{3.59}$$

with

$$\sin\phi_d^{\text{eff}} = -\frac{S_{D^+D^-}}{\sqrt{1 - C_{D^+D^-}^2}} \tag{3.60}$$

can be measured in $B^0 \to D^+D^-$. The phase shift $\Delta\phi_d$ is given by

$$\tan\Delta\phi_d = \frac{a^2\sin 2\gamma - 2a\cos\theta\sin\gamma}{1 - 2a\cos\theta\cos\gamma + a^2\cos 2\gamma}\,. \tag{3.61}$$

The decay channel $B_s^0 \to D_s^+ D_s^-$ is related to $B^0 \to D^+ D^-$ via U-spin symmetry. It is governed by a $b \to c\bar{c}s$ transition and gives access to ϕ_s. The measurement is as well polluted by hadronic penguin effects. Here, the phase shift is given by

$$\tan \Delta\phi_s = \frac{\epsilon^2 a'^2 \sin 2\gamma + 2\epsilon a' \cos\theta' \sin\gamma}{1 + 2\epsilon a' \cos\theta' \cos\gamma + \epsilon^2 a'^2 \cos 2\gamma} \,. \tag{3.62}$$

The factor

$$\epsilon \equiv \frac{\lambda^2}{1-\lambda^2} = 0.0536 \pm 0.0003 \tag{3.63}$$

reduces the impact of the phase shift. However, in the concept of U-spin symmetry

$$a e^{i\theta} = a' e^{i\theta'} \,. \tag{3.64}$$

Thus, with a measurement of *CP* violation in $B^0 \to D^+ D^-$ decays, where the phase shift is not suppressed by ϵ, the hadronic parameters a and θ can be determined precisely, and then transferred to the measurement of *CP* violation in $B_s^0 \to D_s^+ D_s^-$ decays.

Further decay modes from the family of $B \to D\overline{D}$ decays are $B^0 \to D^{*+} D^-$ and $B^0 \to D^{*+} D^{*-}$, which also enable a determination of ϕ_d^{eff}, but introduce further complications. For $B^0 \to D^{*+} D^-$ the final state is not a *CP* eigenstate as it can be distinguished by the charge of the $D^{*\pm}$ meson. Thus, four *CP* observables are needed to describe *CP* violation. Furthermore, from an experimentalist's point of view the final state is not symmetrical in terms of the charges of pions and kaons and thus a detection asymmetry has to be taken into account. In the measurement of *CP* violation using $B^0 \to D^{*+} D^{*-}$ decays, like for $B_s^0 \to J/\psi\,\phi$ decays, an angular-dependent analysis is required.

4 The LHCb Experiment at the LHC

The Large Hadron Collider beauty (LHCb) experiment is one of the four large experiments run at CERN, the European Organisation for Nuclear Research, in Geneva, Switzerland. The physics goal of the collaborators at LHCb is to test the standard model of particle physics (SM) by performing indirect searches for new physics with hadrons containing b or c quarks. In contrast to the direct searches conducted by the multipurpose experiments ATLAS (A Toroidal LHC Apparatus) [74] and CMS (Compact Muon Solenoid) [75], CP-violating processes and rare decay modes are precisely measured and the results are compared with the predictions of the SM. This allows to investigate effects caused by heavy unknown particles at energy scales far beyond what is accessible in direct searches. To do so, it is mandatory to have a very high vertex, momentum and decay time resolution. The detector, with which this demanding task is accomplished, is described in Sec. 4.2. Beforehand, a description of the accelerator complex, which provides the proton bunches to collide inside the detector, is given (see Sec. 4.1). Afterwards, the trigger system (see Sec. 4.3) and the specific software (see Sec. 4.4) are described.

4.1 The Large Hadron Collider

At CERN elementary particle physics is studied using the world's largest particle accelerator, the Large Hadron Collider (LHC). Located at the French-Swiss border area an accelerator complex including several linear and circular (pre-)accelerators is operated (see Fig. 4.1). Before particles enter the 27 km long ring of the LHC 50 to 175 m beneath ground, they have undergone accelerations to an energy of 450 GeV by the linear accelerator LINAC 2, the BOOSTER, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS). A total of 1232 superconducting dipole magnets keep the particle beams on the circular track, while quadrupole magnets focus them. The magnets have to be operated at a temperature of $-271.3\,^{\circ}\mathrm{C}$, which is achieved by a cooling system of liquid helium.

From 2010 on mainly proton bunches and for shorter periods of time also bunches of lead, on whose study ALICE (A Large Ion Collider Experiment) [77] is focused, are injected. The protons are collided at centre-of-mass energies of 7 and 8 TeV in 2011 and 2012 (Run I) and since 2015 (Run II) at 13 TeV.

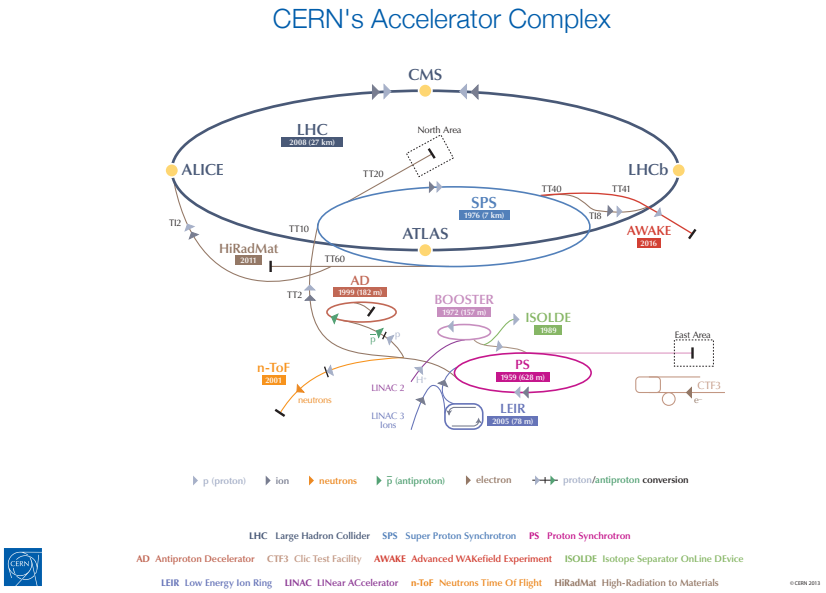

Figure 4.1: CERN's accelerator complex [76].

4.2 The LHCb detector

The LHCb detector, depicted in Fig. 4.2, is a single-arm forward spectrometer. This means that the individual detector components are sequentially arranged in the forward direction, starting from the interaction point. The angular acceptance in the horizontal plane is 10–300 mrad and in the vertical plane 10–250 mrad. Thereby, a pseudorapidity range $2 < \eta < 5$ is covered. Instrumenting only this part of the space has been found to be an optimal compromise between cost and output for LHCb's desired physics program, which is mainly to study particles containing b or c quarks. Simulations of the correlation between the angular distribution of $b\bar{b}$ quark pairs (see Fig. 4.3) show that a significant amount of the b and $\bar{b}$ quarks are produced in quite small cones around the beam axis. Of course, half of the $b\bar{b}$ quark pairs are going backwards but about 25 % of all $b\bar{b}$ quark pairs are inside the instrumented 4.5 % of the whole space.

More details on the structure of the LHCb detector can be found in Ref. [78] and an overview of the performance is given in Ref. [79].

Vertexing and tracking

The tracking system consists of several detector components, one of which is a dipole magnet, which bends the tracks of charged particles with an integrated magnetic field of 4 Tm. To be able to study charge-dependent detection asymmetries the polarity of the dipole magnet is reversed periodically throughout data-taking. The momentum of the track can be derived from the curvature radius. To determine

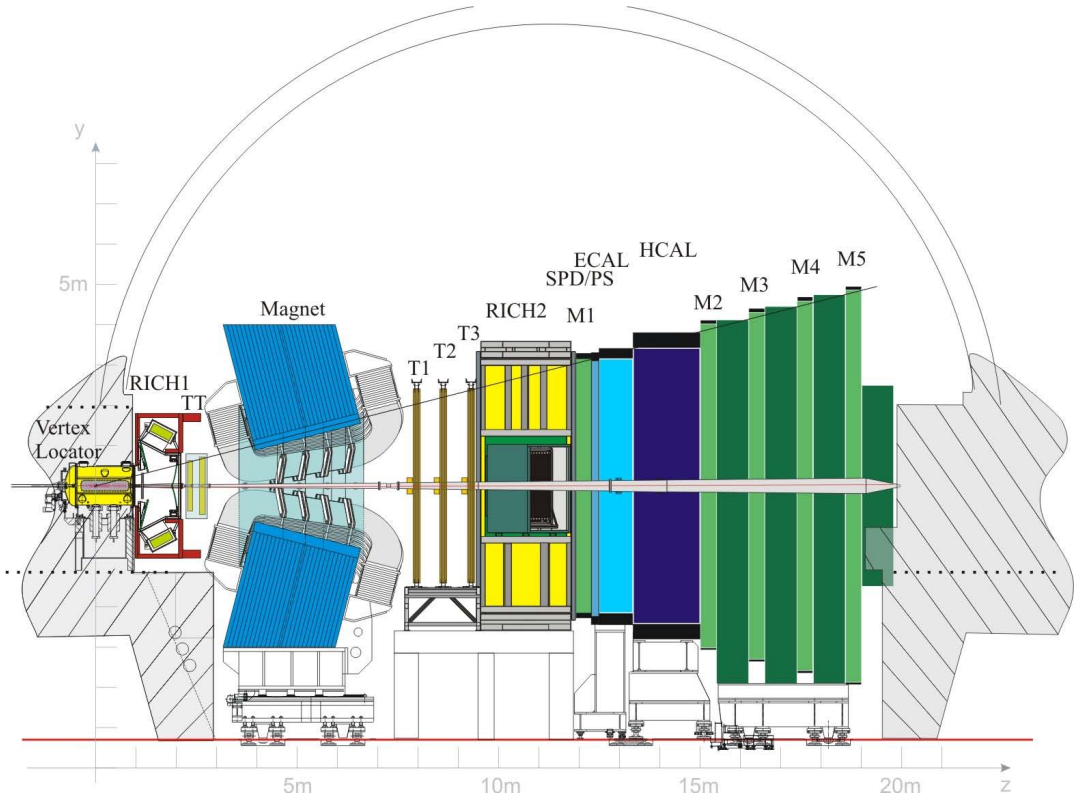

Figure 4.2: Schematic view on the LHCb detector [78].

the curvature radius, information from tracking detector elements located upstream and downstream of the magnet is needed. The *pp* interaction region is surrounded by a silicon-strip vertex locator (VELO) [80], which delivers the most precise information on the position of the tracks and vertices due to being installed very closely around the beam pipe. It is composed of 42 modules with R and ϕ sensors, which measure the positions of the tracks in cylindrical coordinates. Each module is a half disk (see Fig. 4.4), which can be pulled to a proximity of 5 mm to the beam axis. However, this is only done for stable beam conditions, otherwise the modules could be destroyed by the beam. To monitor the beam position a dedicated detector component called Beam Conditions Monitor (BCM) is installed at two locations in the vicinity of the beam. Via eight diamond sensors, which have been proven to be very radiation-hard, each station determines the particle flux and can trigger a beam dump in case of instabilities, which occur especially during the injection of proton bunches. The importance of this system is underlined by the fact that it has its own power supply and constantly reports its status. If no information from the BCM is received a beam dump is also initiated. The VELO achieves a single hit resolution of up to 4 µm at an efficiency of more than 99 %. The disks are arranged in a way that guarantees that even at the outermost acceptance of 300 mrad tracks hit at least three VELO stations (see Fig. 4.4). One of the main purposes of the VELO is to precisely determine the position of the proton-proton

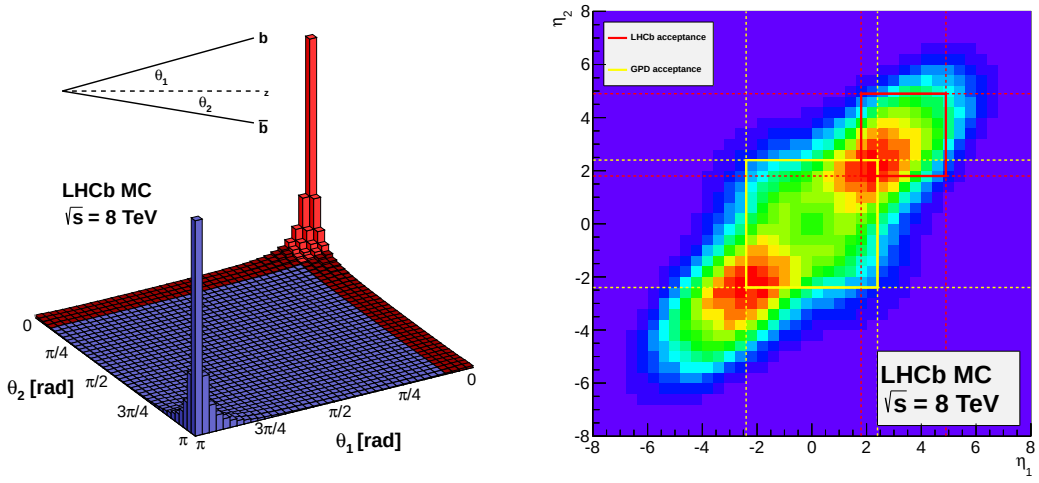

Figure 4.3: Correlation of angular acceptance (left) and pseudorapidity (right) of $b\bar{b}$ quark pairs. The frequency of produced $b\bar{b}$ quark pairs is indicated by the bin content in the plot showing the angular acceptance and by the colour code in the plot of the pseudorapidities, where purple means low and red corresponds to high. The region marked in red (left) respectively the region in the red square is instrumented by the LHCb detector.

interaction called primary vertex (PV) and the displaced secondary decay vertices of long-lived particles like B^0 mesons. The impact parameter (IP), which is the minimum distance of a track to a PV, is measured depending on the transversal momentum p_T with a spatial resolution of $(15 + 29/p_\mathrm{T})\,\mu\mathrm{m}$ (p_T given in units of GeV/c).

Between VELO and dipole magnet there is another silicon-strip detector, the tracker turicensis (TT). Like the three tracking stations located downstream of the magnet, which are subdivided into an inner silicon-strip tracker and an outer straw drift tube detector [81], it is built of four layers. While the first and last layer are arranged vertically, the inner layers are tilted by $-5\,^\circ$ and $+5\,^\circ$, respectively. Charged particles create electron-hole pairs in the silicon-strip detectors inducing a measurable current. A hit efficiency of at least 99.7 % and a hit resolution of 55 µm and better is achieved during data-taking in 2011 and 2012 [79]. The straw drift tubes of the outer tracker are filled with a gas mixture of 70 % Ar, 28.5 % CO_2 and 1.5 % O_2. The addition of the oxygen reduces the ageing rate [82]. Passing particles ionise the gas. Timing measurements on how long it takes for the electrons to reach the anode in the middle of the tube allow to reconstruct the position of the hit. In total, the tracking system provides a relative precision on the measurement of the momentum that varies from 0.5 % at low momentum to 0.8 % at 100 GeV/c [79].

Different track types are distinguished based on which detector components provide information on the trajectory of the track. The category with the best

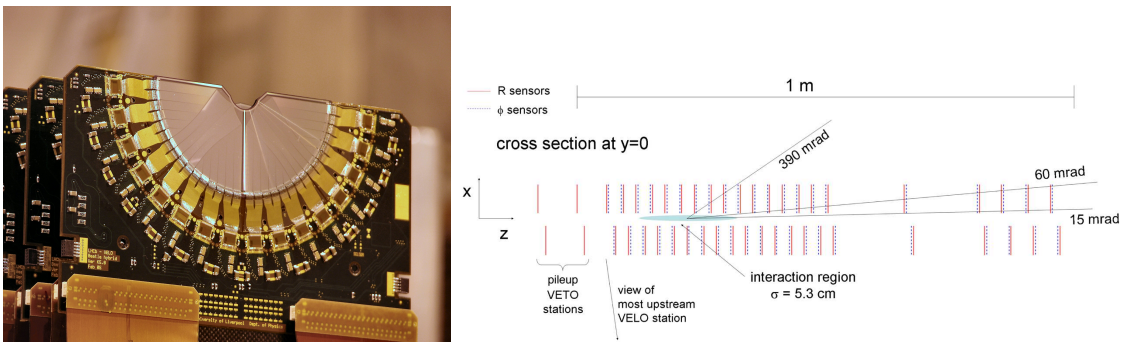

Figure 4.4: View on a single VELO half disk (left) and arrangement of all VELO modules [78].

mass, momentum and vertex resolution is the long category. Long tracks originate in the vertex detector and leave hits in all subsequent tracking stations. Long-lived particles like K_S^0 mesons might decay outside the VELO. If their tracks are detected in the TT and, after passing the magnet, in the tracking stations they are referred to as downstream. These two track categories are the only ones used in the analyses described in this thesis. Furthermore, tracks are classified as VELO tracks, if they have only left hits in the VELO, as upstream tracks, if in addition the TT delivers information, or as T tracks, if they are solely reconstructed in the tracking stations downstream the magnet.

Particle identification

Apart from detecting the tracks and reconstructing their trajectory it is important to estimate the identity of the particles. To distinguish pions from kaons and protons two ring-imaging Cherenkov (RICH) detectors [83] are used, which are installed between VELO and TT respectively downstream of the tracking stations. The RICH detector upstream of the magnet is filled with C_4F_{10} and during Run I additionally with Aerogel. It is designed for particles with momenta in the range of 1–60 GeV/c. Higher momentum particles are detected by the second RICH detector, which is filled with CF_4. When particles pass through these materials with a speed greater than the speed of light in the medium photons are emitted. The light is guided to hybrid photo detectors by a system of mirrors (see Fig. 4.5). From the radius of the light cones and the measurement of the momentum a particle hypothesis can be constructed. While photons and electrons are identified by an electromagnetic calorimeter (ECAL), the energy of protons, neutrons and other long-lived hadrons is measured in a hadronic calorimeter (HCAL). To suppress background from charged and neutral pions there is a preshower (PS) respectively a Scintillator Pad Detector (SPD) in front of the ECAL. The thickness of the lead in the PS is chosen as a compromise between energy resolution and trigger performance [84]. The calorimeters are built of alternating layers of metal and plastic. Polystyrene molecules in the plastic are excited by particle showers produced

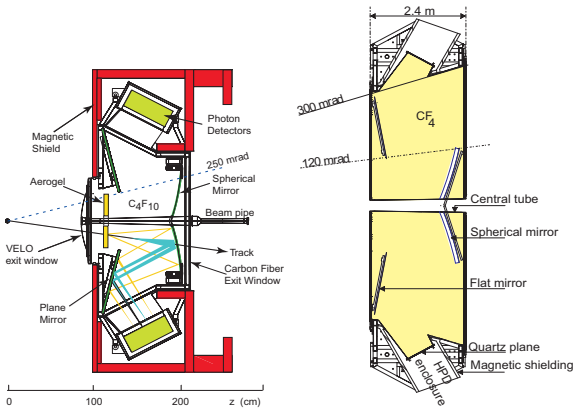

Figure 4.5: Side view schematic layout of RICH 1 (left) and top view schematic of RICH 2 (right) [78].

in the metal plates and produce ultraviolet light, whose amount is proportional to the energy of the incident particle. The least interacting charged particles are muons, which are identified by five stations of multi-wire proportional chambers (MWPC) filled by a gas mixture of Ar, CO_2 and CF_4. Four of them are right at the end of the detector downstream of the calorimeters and one is located in between the second RICH and the calorimeter system. To stop the muons 80 cm thick layers of iron are put between the last four muon stations. Only muons with a momentum $p > 6$ GeV/c pass the whole detector. The detection of the muons is based on ionisation of the gas in the MWPCs. An electric field accelerates the ions and electrons. The emerging current is proportional to the energy of the muon.

4.3 The LHCb trigger system

Deliberately, the instantaneous luminosity at LHCb is reduced to $4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is significantly lower than at the other three experiments at the LHC. Though, the partial beam loss with time can be compensated by adjusting of the beam crossing, so that a constant luminosity level can be kept throughout the whole fill [85]. Nevertheless, it is not possible to store the data of all visible proton-proton collisions. Instead, a two stage trigger system consisting of a hardware (L0) and a subsequent software level (HLT) is deployed. At the hardware trigger stage, which runs synchronously with the bunch crossing rate of 40 MHz, events are required to contain at least one muon with a high p_T (`L0Muon`), or two muons with a minimal product of their p_T (`L0DiMuon`), or a hadron (`L0Hadron`), a photon (`L0Photon`) or an electron (`L0Electron`), which deposit high transverse energy in the calorimeters. Additionally, the number of allowed hits in the SPD is limited. These requirements reduce the data rate to 1 MHz, at which the full detector can be read out. The L0 signal efficiency varies a lot between muons and hadrons. While dimuon final states are triggered with more than 90 % efficiency, for fully hadronic final states

like D^+D^- only around 60 % are reached [86, 87]. The high level trigger (HLT) is a C++ application, which runs on an event filter farm of several thousand CPU nodes. It is again split into two stages. In the HLT1 basically the decisions of the L0 are checked. Due to the reduced data rate some more time is available. For all events the VELO tracks are reconstructed and a partial event reconstruction of all charged particles with $p_T > 500$ MeV/c in 2011 and $p_T > 300$ MeV/c in 2012 is performed. This improves the momentum resolution and enables to calculate some invariant masses. The `Hlt1TrackMuon` trigger line requires a high p_T muon with a χ^2_{IP} with respect to any primary interaction greater than 16, where χ^2_{IP} is defined as the difference in χ^2 of a given PV reconstructed with and without the considered track. The `Hlt1DiMuonHighMass` trigger line accepts events if they contain two muons that form a good common vertex with an invariant mass above 2.7 GeV/c^2. In HLT2, a full reconstruction of the event is performed. Therefore, it is possible to further tighten the requirements applied in HLT1. Furthermore, for the `Hlt2DiMuonDetachedJPsi` trigger line a requirement on the flight distance is imposed. For hadrons, it is typically searched for two-, three- or four-track secondary vertices, which are identified via a multivariate algorithm [88].

The total output rate after all trigger stages has been increased from 3.5 kHz in 2011 to 5 kHz in 2012 and 12.5 kHz in Run II.

In the offline selection, trigger signals are associated with reconstructed particles. Selection requirements can therefore be made on the trigger selection itself and on whether the decision was due to the signal candidate (TOS), other particles produced in the pp collision (TIS), or a combination of both.

A detailed description of the requirements imposed in the individual trigger lines is given in Ref. [86], while the latest performance numbers can be found in Ref. [87].

4.4 The LHCb software

4.4.1 Reconstruction

Two things need to be done in the reconstruction: tracks need to be find and combined, and particle hypotheses have to be assigned to these tracks. The interface for the algorithms and tools is provided by the Brunel project [89], based on the Gaudi framework [90].

The forward tracking algorithm starts with straight VELO tracks, which are built from hits in the R and ϕ sensors of the VELO modules. These are extrapolated to match hits in the tracking stations taking into account the bending of the tracks by the magnet. Then, corresponding hits in the TT are added. A second tracking algorithm directly matches independent VELO and T tracks. It is possible that certain track segments are used for different tracks. In that case a Clone Killer algorithm selects one of the reconstructed tracks. Another difficulty are ghost tracks, which are randomly combined hits that do not stem from a real physics particle.

The particle identification (PID) of charged hadrons is performed via the allocation

of rings in the RICH detectors to the tracks and calculating likelihoods for the different particle hypotheses. The calorimeters are used to identify electrons and neutral pions, which decay into pairs of photons. One of the best signatures is given by the muon system, which excludes respectively settles the muon hypothesis quite reliable.

4.4.2 Stripping

Once tracking and PID are done the full decay chain can be fitted. However, the data size after reconstruction is enormous. Right now, it is inevitable to use a centralised selection called Stripping to handle it. In the DaVinci framework [91] stripping lines are defined, which basically are a set of requirements that describe certain decay modes. Many selection steps can be shared between various stripping lines, which saves a lot of computing time. For example, there are minimal requirements for stable particles to start from. In the individual stripping lines these can then be tightened. Only data selected by a stripping line can be analysed offline by the users and data campaigns are usually only performed as often as once per year. This makes the stripping so important. In the stripping the OfflineVertexFitter (OVF) is used for the analysis of the $B^0 \to J/\psi\, K^0_S$ decays and the LoKiVertexFitter (LVF) for the analysis of the $B^0 \to D^+ D^-$ decays. In order to correctly comprise correlations and uncertainties on vertex positions, particle momenta, flight distances, decay times, and invariant masses, the DecayTreeFitter (DTF) [92] can be used in the reconstruction of decay chains. The decay-time related observables in both analyses covered in this thesis stem from a DecayTreeFit, where a constraint on the production vertex of the B^0 mesons is applied using the knowledge about the position of the primary vertex. The momenta and the invariant mass of the B^0 meson are determined with a DecayTreeFit, in which additionally the invariant masses of the daughter hadrons are constrained to their nominal masses.

4.4.3 Monte Carlo simulation

In many ways data analyses benefit from the use of Monte Carlo simulations (MC). This reaches from the calculation of efficiencies to the development of selection strategies or to finding appropriate parametrisations to model data distributions. One of the advantages of MC is that, except for the need of enough computing and storage resources, the simulated samples can be very large, typically considerably larger than the real data sample. The main goal of the simulation is to be as close as possible to the conditions found on real data. Therefore, constant comparison, calibration and adjustments are needed. Whenever a deviation attracts attention, methods to compensate the effect are applied, *e.g.* the performance of the particle identification system is overestimated on MC, which can be corrected by applying a data-driven resampling.

For LHCb the pp collisions are generated using Pythia [93] with a specific configuration [94]. The decays of the hadronic particles are simulated with EvtGen [95].

To account for the radiation of photons in the final-state the package PHOTOS [96] is used. The GEANT4 toolkit [97] is implemented as described in Ref. [98] to model the interaction of the generated particles with the detector material. The digitisation is realised using BOOLE [99]. The further processing is identical with the analysis of real collision data, starting with the trigger implemented in MOORE [100], the reconstruction done via BRUNEL [89] and the stripping using the DAVINCI package [91].

Apart from the reconstructed properties of the particles, the true information is available as well. This allows to compare the two and study resolution and acceptance effects. Besides signal MC, where a specific decay mode is specified in all details in a configuration file, it is also possible to generate inclusive MC samples, which contain a whole family of similar decay modes, *e.g.* decays involving a J/ψ resonance and anything else, or even completely unbiased samples.

4.5 Flavour tagging

For measurements of *CP* violation in B decays it is essential to know the initial flavour of the decaying b hadron candidate, *i.e.* whether it contained a b or a $\overline{b}$ quark at production. When studying decays of charged B mesons the flavour at decay matches the production flavour. Therefore, the flavour can unambiguously be determined from the charges of the final state particles. Due to meson oscillations it is not as trivial for neutral mesons. Instead, dedicated methods called flavour-tagging algorithms are needed, which infer the initial flavour of a reconstructed candidate from other particles inside the event. The B-factories BaBar and Belle have been operated at the $\Upsilon(4S)$ resonance, which dominantly decays into a quantum-correlated pair of $B\overline{B}$ mesons. Therefore, by analysing the decay of the non-signal B meson, *e.g.* if it proceeds via a flavour-specific process, the flavour of the signal B meson at that time could be determined. Such correlations are not present at proton-proton colliders like the LHC, where b quarks are dominantly produced in $b\bar{b}$ quark pairs via gluon-gluon fusion. The LHCb collaboration has developed several flavour-tagging algorithms, which can be classified as *same-side* (SS) and *opposite-side* (OS) taggers. A schematic overview of all current taggers that can be exploited to tag B^0 mesons is given in Fig. 4.6. By convention each flavour-tagging algorithm provides a flavour tag of $d = +1$ if the tagger decides that it is more likely that the flavour of the initial B meson was a B^0 and of $d = -1$ if, based on the algorithm, a $\overline{B}^0$ flavour is more likely. However, when no appropriate tagging particle can be found for a reconstructed candidate, a tag decision of $d = 0$ is assigned. The tag decisions are either based on the charge of a single selected tagging particle or on the sign of the averaged charge of multiple tagging particles. Besides the tag decision each tagger also provides a prediction η on the probability that the tag decision is wrong. This mistag estimate η takes values between 0 and 0.5, where $\eta = 0$ means that there is no uncertainty on the tag decision and $\eta = 0.5$ basically corresponds to a random choice and is associated with $d = 0$. These

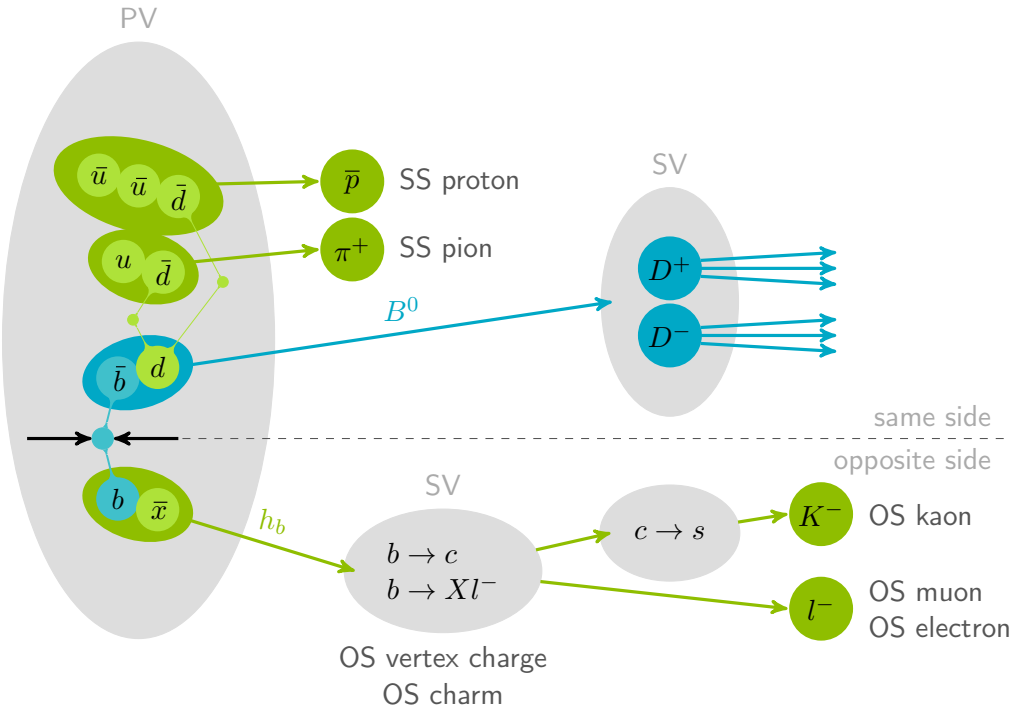

Figure 4.6: Available tagging algorithms to tag B^0 mesons at the LHCb experiment.

predictions are based on the outcome of multivariate classifiers, which combine kinematic and geometric properties of the tagging particle as well as information on the event.

The performance of flavour-tagging algorithms can be quantified by the tagging efficiency ε_{tag}, which specifies for how many reconstructed candidates a tag decision can be made, and by the true mistag probability ω. The relation between the predicted and the true mistag probability $\omega(\eta)$ is determined in calibration studies (see Sec. 5.7). From the mistag probability the tagging dilution $D = 1 - 2\omega$ can be derived, which indicates how much a measured amplitude is reduced with respect to the physical amplitude due to wrong tags. The product of the tagging efficiency and the squared tagging dilution $\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} D^2$ is called tagging power or effective tagging efficiency. It is widely used as figure of merit for tagging algorithms as it states the effective loss in statistics compared to a perfectly tagged sample. Ideally, the tagging power is calculated on a per-candidate basis by summing up the dilution of all N signal candidates according to

$$\varepsilon_{\text{eff}} = \frac{1}{N} \sum_{i=1}^{N} (1 - 2\omega(\eta_i))^2 \,, \tag{4.1}$$

with $\omega = 0.5$ ($D = 0$) for the untagged candidates.

While further information on the basic principles of LHCb's flavour-tagging algorithms can be found in Refs. [101, 102], a short description of the OS and SS taggers will be given in Secs. 4.5.1 and 4.5.2.

4.5.1 Opposite-side flavour tagging

Opposite-side flavour-tagging algorithms [103] infer the flavour of the signal B meson by studying the decay process of the second b hadron, which is produced from the same $b\bar{b}$ quark pair as the reconstructed signal B meson. Mistag probabilities are on the one hand introduced by selecting a wrong tagging particle and on the other hand intrinsically if the opposite-side b hadron is neutral and has already mixed at the time of decay.

In case of a semileptonic decay of the opposite-side b hadron the charges of the leptons are used by the OS electron (OSe) and OS muon (OSμ) tagger to determine the flavour. These two taggers provide relatively good mistag estimates of around 30 %, but have quite low tagging efficiencies of around 2 % (OSe) and 5 % (OSμ) for charmonium respectively 3.5 % (OSe) and 8.5 % (OSμ) for open charm modes. The efficiency for muons is a factor 2–3 higher than for electrons due to the better reconstruction and identification with the LHCb detector.

The OS kaon tagger selects kaons from a $b \to c \to s$ decay chain. Its tagging efficiency is around 17 % (21 %) for charmonium (open charm) modes at an average mistag probability of approximately 39 %.

The OS vertex charge tagger reconstructs the secondary vertex of the opposite-side b hadron and calculates the average charge of all tracks associated to this vertex. The tagging efficiency and mistag probability are comparable with the OS kaon tagger.

The recent OS charm tagger [104] reconstructs charm hadron candidates produced through $b \to c$ transitions of the opposite-side b hadron. The main contribution to the tagging power of the OS charm tagger comes from partially reconstructed charm hadrons in $K^-\pi^+X$ final states and from $D^0 \to K^-\pi^+$ decays. The overall tagging efficiency of the OS charm tagger is only 3–5 % with a mistag probability of around 35 %.

The four taggers described first are usually combined into an OS combination. Since the release of the OS charm tagger a new OS combination is defined, which is slightly better than the old one.

4.5.2 Same-side flavour tagging

Apart from b quarks being produced in $b\bar{b}$ quark pairs, also d quarks mainly stem from $d\bar{d}$ quark pairs. In the hadronisation process of the B signal candidate charged pions and protons can be produced, which contain the other quark from the $d\bar{d}$ quark pair, and whose charge is thereby correlated with the initial flavour of the reconstructed B signal candidate. Positive pions and antiprotons are associated with B^0 mesons, and negative pions and protons with $\overline{B}^0$ mesons. Additionally, B^0 mesons can originate from the decay $B^{*+} \to B^0\pi^+$ of excited charged B mesons. Then, the charge of the associated pion again determines the initial flavour. For quite some time, the only available same-side flavour-tagging algorithm for B^0 mesons was a SSπ tagger with a cut-based approach to select the appropriate

tagging pion. It has a tagging efficiency of around 15 % for $J/\psi X$ final states at an average mistag probability of about 42 %. Recently, an improved SSπ tagger using a boosted decision tree (BDT) to select the tagging pion and based on the very same principles also a SSp tagger have been developed [105, 106]. These two flavour-tagging algorithms select completely disjoint tagging particles ensured by a requirement on the distance log-likelihood (DLL) between the pion and the proton hypothesis of the tagging particle. The tagging efficiency of the SSπ tagger is very high with 70–75 %. The SSp tagger provides non-zero tags for around 35 % of all reconstructed signal candidates, of which about 80 % are also tagged by the SSπ. However, the high tagging efficiency comes along with rather large average mistag probabilities of around 45 %. The response of the two BDT-based SS taggers is combined into a common SS response.

5 Data Analysis Tools and Methods

5.1 Maximum likelihood method

For parameter estimations, especially in the case of multidimensional fits, the (extended) maximum likelihood method is often used. The extended likelihood function is defined as

$$\mathcal{L}(\vec{\lambda};\vec{x}) = \frac{e^{-N}N^n}{n!}\prod_s\prod_{i=1}^{N^s}\mathcal{P}^s(\vec{x}_i;\vec{\lambda}_s)\,. \tag{5.1}$$

Here, $\mathcal{P}^s$ is a properly normalised probability density function (PDF), which can differ between several simultaneously treated categories indexed via s. The vector $\vec{x}$ contains the values of all observables and the vector $\vec{\lambda}_s$ comprises all parameters, for which the optimal values, *i.e.* the ones that maximise the likelihood function, have to be found. Most numerical algorithms do not maximise the likelihood function but instead look for the minimum of the negative log-likelihood $-\ln\mathcal{L}$, which represents the same optimum. The prefactor takes the probability into account that $n = \sum_s N^s$ events are observed although N are expected. The value of this expectation is as well estimated in the likelihood fit. To include uncertainties on fixed parameters in the likelihood fit Gaussian PDFs

$$\mathcal{G}(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{5.2}$$

with mean μ and width σ are multiplied to the likelihood function. Herein, the parameter to be constrained is the variable and the Gaussian's mean and width are fixed to the parameter's value and uncertainty, respectively.

5.2 Selection

When reconstructing decays in a hadronic environment, like at LHCb, it is inevitable that some of the candidates do not stem from the signal decay chain that one wants to analyse. In fact, most of the reconstructed candidates are usually built from random combinations of tracks that have no common physical origin. But especially when searching for a very rare signal it is better to be careful and at first stage rather keep an event than throwing it away. However, this is limited by the available disk space and computational resources. Furthermore, the sensitivity of a measurement suffers from background contamination. Therefore, a selection needs

to be developed that separates signal from background candidates. The simplest selection is a requirement of the type $a < b$. Several of these cuts can be combined to a sequence but for each variable a maximum of two requirements can be applied, *i.e.* a minimal and a maximal value can be defined. This means that out of the whole phase space only a hyperrectangle is selected. But the simplicity is also a strength of the cut-based selection. It is very fast and the selection requirements can easily be understood and connected to event or particle properties. Additionally, the efficiency of the requirements can be determined individually. To account for dependencies between the variables a grid search can be performed, in which the optimal cut values are determined recursively. In principle, the cut-based selection can be extended by allowing case differentiation, *i.e.* connecting sequences with `OR` requirements, or by constructing combinations of variables, like products or ratios. However, a cut-based selection often leads to suboptimal selection performances.

5.2.1 Multivariate selection

More elaborate selection methods are based on machine learning algorithms, which more and more enter the field of particle physics. These multivariate techniques improve the possibility to separate signal from background contributions as they make use of correlations between input variables. Software packages, like TMVA [107] or scikit-learn [108], provide implementations of these algorithms.

A simple multivariate classifier is a decision tree [109], which splits the phase space according to repeated decisions on certain variables. Starting from a root node the variable and cut value is determined, where the best separation can be achieved according to a criterion like the Gini index $p \cdot (1-p)$ [110], with p being the signal purity, *i.e.* the fraction of signal in the total data sample. At each following node the cut value and even the variable used for the separation can be chosen dependent on the previous decision. The depth of the tree, *i.e.* the maximal number of consecutive decisions, is tunable. When a stop criterion is matched, *e.g.* the ratio of candidates reaching a node falls below a predefined threshold, no further decisions are applied. The tree is trained with a set of labelled data. Each final leaf is classified as signal or background depending on the class of the majority of training events ending in that leaf. Ideally, the decision tree should learn, which sequence of criteria achieves the best separation between signal and background. An improvement can be achieved by combining several decision trees to a forest, in which the classification of an event follows a majority vote of the individual trees, *e.g.* in a Random Forest [111].

An alternative is to alter the impact of the training events in a decision tree by applying weights. This procedure is called boosting and leads to Boosted Decision Trees (BDT) [112]. One possibility how the boosting can be realised is the AdaBoost method [113]. Events that are misidentified in the previous tree are weighted by

$$\alpha = \frac{1-\varepsilon}{\varepsilon}\,, \tag{5.3}$$

with $\varepsilon < 0.5$ being the misclassification rate of the previous tree. This criterion is fulfilled as long as the decision tree performs better than a random choice. The learning rate can be further modified by using an exponent β for the weight, $\alpha \to \alpha^{\beta}$. To have the same effective number of events as before, *i.e.* the same sum of weights, the events are renormalised before training the next tree. The BDT output classifier y is given by the weighted mean of the tree's output h, which is $+1$ for signal and -1 for background,

$$y = \frac{1}{N_{\text{trees}}} \sum_{i}^{N_{\text{trees}}} \ln(\alpha_i) h_i \,. \tag{5.4}$$

This means that an event is classified more signal-like the higher the BDT output classifier. Another boosting algorithm is the GradientBoost method [114]. It is based on minimising a defined loss function, which describes the deviation between the classification and the truth, by calculating the gradient of the loss function.

The importance of a feature in a (boosted) decision tree can be determined by counting how often it is used and by considering how important the cuts are in terms of events reaching the corresponding node and in terms of the separation power between signal and background candidates.

Another multivariate method is the Artificial Neural Network (ANN) [115]. It consists of several neurons, divided into input units, hidden units and output units. The way how these units are connected, the corresponding strength and the effect, either excitatory or inhibitory, defines the neural network.

Multivariate selection methods have the disadvantage that their success has to be limited or in other words that it is not necessarily ideal if a perfect separation is achieved. The problem is that the training data are only proxies for the real data and that the multivariate method eventually learns to distinguish the different categories based on unphysical properties that do not inhere in real data. These could for example be statistical fluctuations or differences between simulated and real data. This so called overtraining leads to an overestimation of the performance of the selection. Overtraining can be avoided if the training data samples are large and as close as possible to the real data, on which the selection should be applied in the end. However, this does not mean that the same data should be used for the training and for the analysis. In the selection of $B^0 \to D^+ D^-$ decays the high mass sideband is taken as proxy for the combinatorial background and excluded from the subsequent analysis. A typical check for overtraining is performed by comparing the output classifier distributions of the training data and of a test sample, which has been removed randomly from a common data set.

5.2.2 Unfolding data distributions using sWeights

Another approach to unfold data distributions is to statistically remove background contributions by applying weights. To do so, the shape of the signal and background contributions for one or several dimensions is needed. While the shapes of signal

decays are mostly defined by some theoretical considerations, which eventually need to be modified to account for experimental effects, like resolutions or acceptances, it is usually difficult to motivate the shape for background contributions from first principles. An observable that has proven to provide a good separation is the invariant mass. Reliable parametrisations of the invariant mass distribution of signal and background contributions can be found quite easily. In the ${}_s\mathcal{P}lot$ technique [116] the weights are calculated from yields of an extended maximum likelihood fit (see Sec. 5.1). Based on the fit results sWeights according to

$$ {}_s\mathcal{P}_\mathrm{n}(x_i) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{\mathrm{n}j} f_j(x_i)}{\sum_{k=1}^{N_s} N_k f_k(x_i)} \tag{5.5} $$

can be calculated for each candidate. Herein, the indices j and k sum over the N_s categories described in the PDF f. The matrix $\mathbf{V}$ contains the covariances between the yields N and needs to be determined from an individual fit, in which all shape parameters are fixed. The sWeights fulfil the condition that their sum over one category returns the corresponding fitted yield. The sWeights can be applied to other observables if they are uncorrelated with the observable used to obtain the sWeights. In sweighted histograms the uncertainty on the bin content of bin i is given by

$$ \sigma(i) = \sqrt{\sum_{e \subset i} ({}_s\mathcal{P}_\mathrm{n})^2}\,. \tag{5.6} $$

5.2.3 Figures of merit

One of the main questions when performing a selection is, how the requirements, be it a sequence of cuts or the classifier of a multivariate method, should be optimised. In an ideal world, one would find a selection that keeps all signal candidates and removes all background contributions. But this is unrealistic. Instead, suitable figures of merit have to be defined, whose optimum should correspond with the ideal selection. In a measurement of *CP* violation the goal is to obtain the optimal precision. Therefore, the sensitivity on the *CP* observables itself seems to be the best figure of merit. Indeed, the optimisation of the BDT output classifier in the selection of $B^0 \to D^+D^-$ decays is based on this very criterion. However, there are some caveats. On the one hand, there is usually more than one observable describing the *CP* violation and a strategy needs to be found how the sensitivities of the different *CP* observables can be combined into a single figure of merit. On the other hand, the absolute uncertainty might depend on the central value and thus small values for the *CP* observables could be preferred. Finally, the full decay-time-dependent fit has to be performed, which is often very complex and takes a long time until it converges. For all these reasons, alternative figures of merit are developed. A very simple one is the pure signal efficiency ϵ_S. Under the assumptions that a high signal yield is more important than a low background contamination and that

the selection requirements are effectively suppressing background contributions, a possible selection strategy is to judge requirements only by their signal efficiency. In the $B^0 \to J/\psi\, K^0_S$ analysis this approach is chosen using the requirement that the individual cuts have to be at least 99 % signal efficient. Including the background yield to the definition of the figure of merit should make it easier to find the optimal cut point, as the background contamination partly influences the achievable sensitivity. There are several possibilities how the signal yield S and the background yield B can be combined: The figure of merit

$$Q_1 = \frac{S}{S+B}\,, \tag{5.7}$$

called purity, describes the fraction of signal candidates. It is limited to unity, which is reached when no background candidates are left over. Using merely the purity does not necessarily lead to an optimal selection, *e.g.* when lots of signal candidates would be thrown away in order to remove one last remaining background candidate. Instead, the signal significance

$$Q_2 = \frac{S}{\sqrt{S+B}}\,, \tag{5.8}$$

which states by how many standard deviations the signal yield exceeds zero, is widely used. To enhance the impact of a high signal yield, the signal significance can be multiplied with the purity:

$$Q_3 = \frac{S^2}{\sqrt{(S+B)^3}}\,. \tag{5.9}$$

For decay-time-independent studies, like determinations of branching ratios, the figure of merit Q_3 is appropriate. In searches for very rare decay modes, where a certain significance expressed in number of standard deviations a is desired, Punzi's figure of merit [117]

$$Q_4 = \frac{\epsilon_S}{a/2 + \sqrt{B}}\,, \tag{5.10}$$

is often used.

However, for decay-time-dependent studies the figure of merit can be improved through an extension that takes into account that the contribution of a signal candidate to the sensitivity on *CP* observables depends on its decay time. For instance, the sine term in the decay-time-dependent asymmetry (see Eq. (3.38)) has its maximum at $t = \frac{\pi}{2\,\Delta m_d} \approx 3\,\text{ps}$. Thus, signal candidates with such decay times have a higher impact. This can be expressed by adding up the square of the differentiation of the log-likelihood with respect to the parameter of interest, here S_f, of all N_S signal candidates:

$$Q_5 = \sum_{i=1}^{N_S} \left[\frac{\sin(\Delta m_d\, t_i)}{1 + d_i\, S_f\, \sin(\Delta m_d\, t_i)} \right]^2 , \tag{5.11}$$

with $d_i = +1$ for B^0 and $d_i = -1$ for $\overline{B}^0$. Additionally, it has to be considered that the decay time distribution of background candidates usually follows an exponential function with an effective lifetime that is significantly smaller than the B^0 signal lifetime. This can be incorporated by using a decay-time-dependent purity $f_i(t)$:

$$Q_6 = \sum_{i=1}^{N} f_i^2 \cdot \left[\frac{\sin(\Delta m_d \, t_i)}{1 + d_i \, S_f \, \sin(\Delta m_d \, t_i)} \right]^2 . \tag{5.12}$$

Here, the sum is built over all N candidates, signal and background. The purity can be determined via a fit to the invariant mass distribution or more specifically via sWeights (see Sec. 5.2.2). Moreover, the influence of further experimentally induced dilutions can be added, *e.g.* from the flavour tagging (see Sec. 4.5) or the decay time resolution (see Sec. 5.6). Such a generalised figure of merit for decay-time-dependent measurements of *CP* violation is derived in Ref. [118]. Although it sounds as if the figure of merit Q_6 is most qualified, it is neither used in the selection of $B^0 \to J/\psi \, K^0_S$ decays nor in the selection of $B^0 \to D^+ D^-$ decays. The reason is that the systematic uncertainty caused by complex figures of merit is difficult to estimate. Therefore, although alternative figures of merit have been tested, the signal efficiency respectively the direct statistical uncertainty are chosen as figures of merit, as they have been trusted more in the internal review processes.

5.3 Spline interpolation

In many cases phenomenological models are an efficient way of describing shapes, *e.g.* when parametrising acceptances, which are typically influenced by more effects than could realistically be analysed separately. Interpolating cubic splines, which are piecewise defined third order polynomials, are an useful implementation [119]. They are parametrised by a set of knots and coefficients at these positions and can be written as the sum over base splines. The first and second derivatives are continuous throughout the domain. The choice of the number and positions of the knots determines how accurate the given shape can be described.

In the $B^0 \to J/\psi \, K^0_S$ analysis described in this thesis cubic splines are used to parametrise the shape of the mistag distributions and to transfer a histogram of the decay-time-dependent efficiency into an unbinned, analytically integrable representation. In the $B^0 \to D^+ D^-$ analysis cubic splines are used to model the deviation of the decay time distribution from a pure exponential distribution, which is not caused by *CP*-violating effects.

5.4 Bootstrapping method

The bootstrapping method (see *e.g.* Ref. [120]) is a frequentist model-independent approach to estimate confidence level intervals. Toy data samples are produced by drawing events from the nominal data sample, with replacement, until the statistics

matches the number of candidates of the nominal data sample. This means that the same event can be drawn multiple times. The bootstrapping method can easily be used for multiple dimensions. Thus, it maintains correlations between the observables without any assumptions or model dependencies. Therefore, bootstrapped samples serve as perfect representations of the original data. Calculations or fits that are originally done once on the nominal data sample can be repeated several times. So, reliable estimates for standard deviations or confidence level intervals can be derived.

5.5 Blinding

Especially when performing precision measurements or searching for rare decays it is advisable to blind the results throughout the analysis and even as analyst only look at them after some thorough (ideally external) review process. Blinding means that the central value or the final result is unknown to all involved people. This procedure avoids the experimenter's bias, *i.e.* the unintended biasing of a result towards a known or expected value or towards a (subconsciously) desired observation. The blinding transformation that has been applied in the measurements of *CP* violation in $B^0 \to J/\psi\, K^0_S$ and $B^0 \to D^+D^-$ decays is adding a hidden offset to the fitted *CP* parameters S_f and C_f using the `RooUnblindUniform` method of RooFit's `RooBlindTools` [121]. With this method the uncertainty on the extracted parameters does not change and can still be used for optimising the selection. Here, the offset is drawn from a uniform distribution between -2 and $+2$ using a random number generator whose seed is generated from a so-called *blinding string*. As the physical range of $\sin 2\beta$ is $[-1, 1]$, this ensures a good opacity. A review of blind analyses is given in Ref. [122].

5.6 Decay time resolution

Uncertainties in the determination of the position of vertices and in the measurement of momenta (although thanks to the VELO (see Sec. 4.2) pretty accurate at LHCb) lead to a finite decay time resolution σ, which dilutes the observed *CP* asymmetry by a factor

$$\mathcal{D} = e^{\frac{-\Delta m_d{}^2\,\sigma^2}{2}}\,. \tag{5.13}$$

This formula is the special case for a Gaussian resolution model with width σ. The general formula is derived in Ref. [123]. For B^0 mesons the dilution induced by the decay time resolution has only minor influence on the measurement of *CP* observables because the oscillation frequency of B^0 mesons $\Delta m_d = (0.5064 \pm 0.0019)\,\hbar\,\mathrm{ps}^{-1}$ [22] is quite low. Even for a decay time resolution of 100 fs, which would be almost two times larger than what is usually found in analyses performed by LHCb, the dilution factor is greater than 99 %.

5.7 Flavour-tagging calibration

As mentioned in Sec. 4.5 the output of the flavour-tagging algorithms η needs to be calibrated to ensure that it reflects the true mistag probability ω. Usually, a linear function

$$\omega(\eta) = p_0 + p_1(\eta - \langle\eta\rangle) \tag{5.14}$$

is chosen. Shifting the function by the average mistag estimate $\langle\eta\rangle$ reduces the correlation between the calibration parameters p_0 and p_1, which in case of a perfect calibration are $p_0 = \langle\eta\rangle$ and $p_1 = 1$.

Due to different interaction rates of the tagging particles with the detector material or detection asymmetries the performance of the flavour-tagging algorithms can depend on the initial flavour. This behaviour is quite unfortunate in the measurement of *CP* violation as it can dilute or enhance the observed asymmetry. To account for these tagging asymmetries separate parametrisations for the flavour-tagging calibrations of initial B^0 and $\overline{B}^0$ are implemented:

$$\begin{aligned} \omega^{B^0}(\eta) &= p_0^{B^0} + p_1^{B^0}\,(\eta - \langle\eta\rangle)\,, \\ \omega^{\overline{B}^0}(\eta) &= p_0^{\overline{B}^0} + p_1^{\overline{B}^0}\,(\eta - \langle\eta\rangle)\,. \end{aligned} \tag{5.15}$$

Equivalently, the calibration parameters for B^0 and $\overline{B}^0$ can be related through their mean and their difference:

$$p_i = \frac{p_i^{B^0} + p_i^{\overline{B}^0}}{2}\,, \quad \Delta p_i = p_i^{B^0} - p_i^{\overline{B}^0}\,, \quad \text{with } i = 0, 1\,. \tag{5.16}$$

The asymmetry of the mistags can then be written as

$$\Delta\omega(\eta) = \Delta p_0 + \Delta p_1(\eta - \langle\eta\rangle)\,. \tag{5.17}$$

For the flavour-tagging calibration it is beneficial to use flavour-specific decay channels that are kinematically similar to the signal channel. Additionally, the selection between the two channels should be as close as possible. This allows to transfer the calibration results from the control to the signal channel without assigning large systematic uncertainties. On the other hand, the control channel should ideally be a mode with high statistics to reduce the statistical uncertainties on the flavour-tagging calibration parameters. A good compromise between these two requirements is found by choosing $B^0 \to D_s^+ D^-$ as calibration mode for $B^0 \to D^+ D^-$, while for $B^0 \to J/\psi\, K_S^0$ the OS tagging combination and the cut-based SSπ tagging algorithm are calibrated with $B^+ \to J/\psi\, K^+$ and $B^0 \to J/\psi\, K^{*0}$ decays, respectively.

5.7.1 Calibration using $B^0 \to D_s^+ D^-$

The flavour-tagging calibration using $B^0 \to D_s^+ D^-$ decays, described in the following, has been provided by collaborators from Milano.

The decay of B^0 mesons via the decay mode $B^0 \to D_s^+ D^-$ proceeds flavour specific as the charge of the D_s^+ meson unambiguously determines the flavour of the decaying B^0 meson. When reconstructing the D_s^+ meson via $D_s^+ \to K^- K^+ \pi^+$ and the D^+ meson via $D^+ \to K^- \pi^+ \pi^+$, a very similar selection to the one for $B^0 \to D^+ D^-$ (described in Sec. 7.1) can be applied. The only differences are that the invariant $m_{KK\pi}$ mass is required to lie within $\pm 25\,\mathrm{MeV/c^2}$ of the known D_s^+ mass [124] and that the vetoes against misidentified backgrounds are not applied. A maximum likelihood fit to the invariant $m_{D_s^+ D^-}$ mass distribution is performed to statistically subtract the remaining background via the ${}_s\mathcal{P}lot$ technique [116]. Apart from the $B^0 \to D_s^+ D^-$ component, which is parametrised by the sum of two Crystal Ball functions (common mean but different widths, and tail parameters taken from MC), the fit model includes components for $B_s^0 \to D_s^- D^+$ decays and for combinatorial background. The total $B^0 \to D_s^+ D^-$ yield is found to be $16\,736 \pm 134$ at a quite low background level, as can be seen in Fig. 5.1.

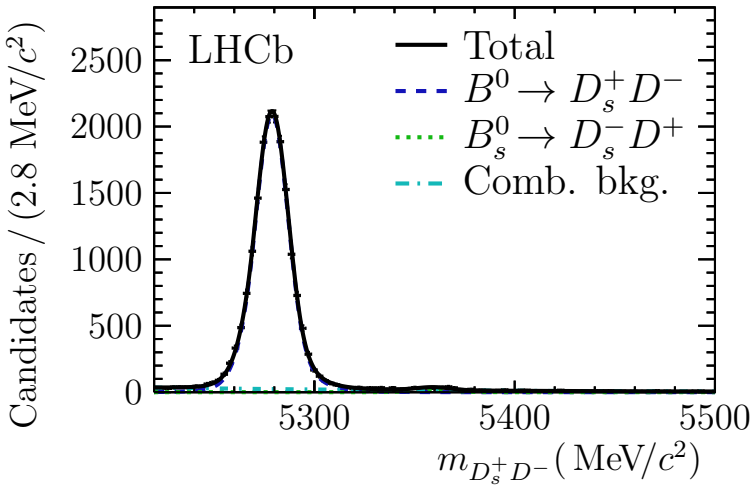

Figure 5.1: Masses of $B^0 \to D_s^+ D^-$ candidates and projected PDFs.

The B^0–$\overline{B}^0$ mixing prevents to infer the flavour of the B^0 meson at production. Therefore, a mixing analysis is performed to determine the true mistag probability ω from the amplitude of the mixing asymmetry

$$\mathcal{A}^{\mathrm{mix}}_{\mathrm{meas}}(t) \equiv \frac{N_{\mathrm{unmixed}}(t) - N_{\mathrm{mixed}}(t)}{N_{\mathrm{unmixed}}(t) + N_{\mathrm{mixed}}(t)} = (1 - 2\omega)\cos(\Delta m_d\, t)\,, \tag{5.18}$$

where N_{unmixed} is the number of $B^0 \to D_s^+ D^-$ decays with a final state that does correspond to the flavour tag, and N_{mixed} the number with a final state that does not. Other experimentally induced effects are corrected for, like the detection asymmetry $\mathcal{A}_{\mathrm{det}}$, the production asymmetry A_{P} and the flavour-specific asymmetry a^d_{sl}. With unbinned maximum likelihood fits to the decay time and tag distributions the results listed in Table 5.1 are determined, one fit for the sample with a non-zero

tag of the OS tagging combination and one for the sample with a non-zero tag of the SS tagging combination. This means that some candidates are used for both calibrations. The systematic uncertainties are dominated by the background subtraction and the calibration method.

Table 5.1: Flavour-tagging calibration parameters from $B^0 \to D_s^+ D^-$. The first uncertainty is statistical and the second accounts for systematic uncertainties.

Parameter	OS	SS
p_1	$1.07 \pm 0.07 \pm 0.01$	$0.84 \pm 0.09 \pm 0.01$
p_0	$0.369 \pm 0.008 \pm 0.010$	$0.430 \pm 0.006 \pm 0.009$
$\langle\eta\rangle$	0.3627	0.4282
Δp_1	$0.03 \pm 0.11 \pm 0.03$	$0.07 \pm 0.13 \pm 0.05$
Δp_0	$0.009 \pm 0.012 \pm 0.001$	$-0.007 \pm 0.009 \pm 0.001$

The raw mixing asymmetries in Fig. 5.2 represent graphically that the OS tagging combination on average provides better mistag estimates but has a lower tagging efficiency than the SS tagging combination. This can be derived from the larger amplitude and the larger error bars.

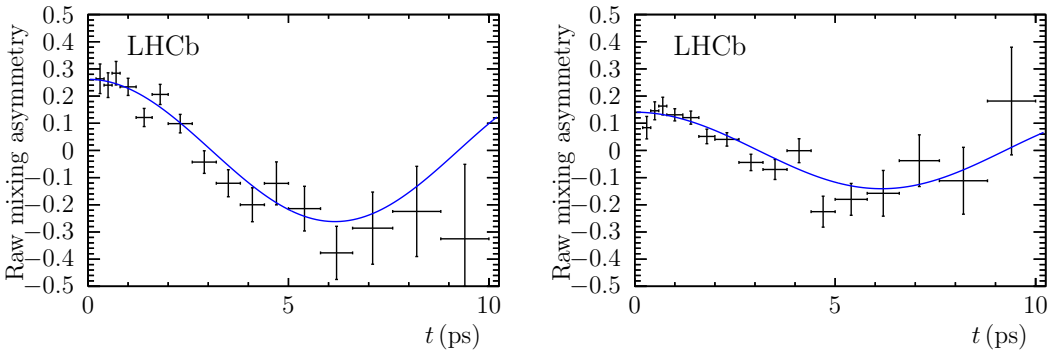

Figure 5.2: Raw mixing asymmetry as a function of the B^0 decay time for events tagged by (left) the OS and (right) the SS tagging combination. The solid line represents the PDF projection.

Thanks to the improved flavour-tagging algorithms and the kinematic properties of the selected $B^0 \to D^+ D^-$ decays, which for example have on average high p_T, an effective tagging efficiency of $\varepsilon_{\text{tag}} D^2 = (8.1 \pm 0.6)\,\%$ is achieved. This splits into a tagging power of $(1.02 \pm 0.09)\,\%$ from events that are tagged only by OS taggers, $(1.36 \pm 0.19)\,\%$ from events tagged only by SS taggers, and $(5.7 \pm 0.5)\,\%$ from events tagged by tagging algorithms of both sides. To date this is the highest effective tagging efficiency in tagged *CP* violation measurements at LHCb.

5.7.2 Calibration using $J/\psi X$ channels

The flavour-tagging calibration for $B^0 \to J/\psi K^0_S$ decays has been performed by collaborators from Lausanne and Dortmund.

For the charged $B^+ \to J/\psi K^+$ decay, which is used to determine the calibration of the OS tagging combination for the analysis of $B^0 \to J/\psi K^0_S$ decays, a comparison of the charge of the kaon with the tag decision directly tells if the tag decision is correct or not. Binning the sample in terms of the mistag estimate a χ^2 fit using Eq. (5.14) is performed to the (η, ω) pairs and reveals the following results for the calibration parameters:

$$\begin{aligned} p_0^{\mathrm{OS}} &= 0.3815 \pm 0.0011\,(\mathrm{stat.}) \pm 0.0016\,(\mathrm{syst.})\,, \\ p_1^{\mathrm{OS}} &= 0.978 \pm 0.012\,(\mathrm{stat.}) \pm 0.009\,(\mathrm{syst.})\,, \\ \langle \eta^{\mathrm{OS}} \rangle &= 0.3786\,. \end{aligned} \tag{5.19}$$

Repeating the same fit with a split of the sample into the initial flavours gives access to the asymmetry parameters Δp_0 and Δp_1, which are determined to be

$$\begin{aligned} \Delta p_0^{\mathrm{OS}} &= 0.0148 \pm 0.0016\,(\mathrm{stat.}) \pm 0.0008\,(\mathrm{syst.})\,, \\ \Delta p_1^{\mathrm{OS}} &= 0.070 \pm 0.018\,(\mathrm{stat.}) \pm 0.004\,(\mathrm{syst.})\,. \end{aligned} \tag{5.20}$$

A cross-check of the calibration in a control sample of $B^0 \to J/\psi K^{*0}$ decays confirms the validity of transferring the calibration from B^+ to B^0 decays.

Despite the advantages of $B^+ \to J/\psi K^+$ as control channel (charged decay mode, very high statistics), for the calibration of the cut-based SSπ tagging algorithm $B^0 \to J/\psi K^{*0}$ decays are used because differences in the composition of the fragmentation products in the B^+ and B^0 hadronisation are expected. Like for $B^0 \to D_s^+ D^-$ a decay-time-dependent mixing analysis is needed. Here, a twodimensional fit to both the reconstructed decay time and mass distributions is performed. From a simultaneous fit in five evenly filled bins of the mistag estimate η the calibration parameters for the SSπ tagging algorithm are determined to be

$$\begin{aligned} p_0^{\mathrm{SS}\pi} &= 0.4232 \pm 0.0029\,(\mathrm{stat.}) \pm 0.0028\,(\mathrm{syst.})\,, \\ p_1^{\mathrm{SS}\pi} &= 1.011 \pm 0.064\,(\mathrm{stat.}) \pm 0.031\,(\mathrm{syst.})\,, \\ \Delta p_0^{\mathrm{SS}\pi} &= -0.0026 \pm 0.0043\,(\mathrm{stat.}) \pm 0.0027\,(\mathrm{syst.})\,, \\ \Delta p_1^{\mathrm{SS}\pi} &= -0.171 \pm 0.096\,(\mathrm{stat.}) \pm 0.04\,(\mathrm{syst.})\,, \\ \langle \eta^{\mathrm{SS}\pi} \rangle &= 0.425\,. \end{aligned} \tag{5.21}$$

The systematic uncertainties for both calibrations cover two different types, one for intrinsic uncertainties and one for the kinematic differences between the control mode and the signal decay $B^0 \to J/\psi K^0_S$. The effect of these two sources is of the same order.

The combined effective tagging efficiency is $(3.02 \pm 0.05)\,\%$, which is composed of a tagging efficiency of $\varepsilon_{\mathrm{tag}} = (36.54 \pm 0.14)\,\%$ and an effective mistag probability

of $\omega_{\text{eff}} = (35.62 \pm 0.12)\,\%$. The major contribution comes from the OS tagging combination, which has an inclusive tagging power of $(2.63 \pm 0.04)\,\%$. The cut-based SSπ tagging algorithm adds $(0.376 \pm 0.024)\,\%$.

6 Measurement of CP Violation in $B^0 \to J/\psi K^0_S$ Decays

The measurement of CP violation in $B^0 \to J/\psi K^0_S$ decays [125] is performed on LHCb's full Run I data sample, which corresponds to an integrated luminosity of 3 fb^{-1}. The analysis is based on a collaborative work with Christophe Cauet [126]. Therefore, my main contributions, namely the decay time resolution study (see Sec. 6.3), the analysis of the backgrounds (see Sec. 6.4), the fitter development, and the studies of systematic uncertainties (see Sec. 6.6), are described in more detail, while the other parts are only briefly summarised to complete the picture. The initial flavour of the B^0 mesons is determined with the OS combination, excluding the OS charm flavour-tagging algorithm, and with the cut-based SSπ tagger. The calibration of these taggers is described in Sec. 5.7.2.

6.1 Data preparation

Candidate $B^0 \to J/\psi K^0_S$ decays are reconstructed through the subsequent decays $J/\psi \to \mu^+\mu^-$ and $K^0_S \to \pi^+\pi^-$. Only combinations of two long (LL) or two downstream tracks (DD) are considered for the pions. All events must have passed either the `L0Muon` or the `L0DiMuon` trigger line. This is implicitly given by the requirement that a positive J/ψ TOS decision of the `DiMuonHighMass` or the `TrackMuon` trigger line in the Hlt1 stage and of the `DiMuonDetachedJPsi` trigger line in the Hlt2 stage exists. These trigger requirements have a total signal efficiency of about 85 %.

In the stripping loose requirements on the quality of the J/ψ, the K^0_S and the B^0 vertex are applied. The invariant masses of the $\mu^+\mu^-$ and the $\pi^+\pi^-$ combination are required to be roughly consistent with the known J/ψ and K^0_S masses, respectively. Moreover, the K^0_S candidates must have a significant decay length and some further kinematic requirements need to be fulfilled. Despite these rather loose selection requirements, the signal to background ratio is already quite high. Therefore, in the offline selection almost all cuts are tuned to have a high signal efficiency. The probability of the muon and pion tracks to be ghost tracks (see Sec. 4.4.1) is reduced to 20 and 30 %, respectively. The mass window around the J/ψ meson is tightened to correspond to five standard deviations. For the K^0_S meson the mass window is adapted to the track type of the pions. It contains four and eight standard deviations for the long and the downstream candidates, respectively. A specific treatment against misidentified $\Lambda^0_b \to J/\psi \Lambda$ and $B^0 \to J/\psi K^{*0}$ decays is performed. The former are rejected by tighter PID requirements on the pions if the invariant

mass under a πp mass hypothesis is compatible with the Λ mass [124]. The latter are suppressed with a cut on the K^0_S decay time in units of its uncertainty.

The invariant B^0 mass, which is restricted to candidates inside 5230–5330 MeV/c^2, is obtained from a fit to the whole decay chain (DecayTreeFit [92]) with constraints on the J/ψ and K^0_S masses as well as the constraint that the momentum vector of the B^0 meson has to point back to the PV. The decay-time related observables stem from a DecayTreeFit, where only the PV constraint is applied. To facilitate the description of the decay time acceptance and to further suppress prompt combinatorial background, only candidates with $t > 0.3$ ps are kept. In the last step of the selection one candidate is chosen randomly for events where multiple candidates have survived the previously described selection.

6.2 Decay time acceptance

The trigger line requirements applied in the selection of the $B^0 \to J/\psi K^0_S$ candidates partially bias the decay time distribution. The sample is divided into an *almost unbiased* subset (AU), defined by

AU $\equiv$ `Hlt1DiMuonHighMass && Hlt2DiMuonDetachedJPsi`

and an *exclusively biased* subset (EB), defined by

EB $\equiv$ `(Hlt1TrackMuon && !Hlt1DiMuonHighMass) && Hlt2DiMuonDetachedJPsi`

Here, the TOS decisions with respect to the J/ψ meson of these trigger lines are used. For both samples the relative efficiency is calculated as the ratio between the number of signal candidates fulfilling these trigger requirements and those fulfilling the requirements of an unbiased collection of trigger lines (`Hlt1DiMuonHighMass && Hlt2DiMuonJPsi`). A simultaneous fit of the mass distribution in ten bins of the decay time is performed to obtain the signal yield ratios. The resulting histograms (see Fig. 6.1) are fitted with cubic splines using the bin centres as knots for the splines. In the fit of the spline function the bin contents are allowed to vary inside their uncertainties, which are estimated as binomial errors for the AU sample and via Gaussian error propagation for the EB sample. The acceptance shape at and beyond the decay time limits is unknown, so it is assumed that the efficiency shape is flat in front of the first and behind the last bin centre.

Another efficiency loss is observed at high decay times. This is mainly caused by a reconstruction inefficiency of the VELO. To account for this effect the lifetime τ is modified according to

$$\widetilde{\tau} = \frac{\tau}{1 + \beta_\tau \tau} \,. \tag{6.1}$$

The correction factor β_τ is determined in a fit to simulated data, where the generation value for the lifetime is known and thus can be fixed. To avoid further decay-time biasing effects an unbiased sample is used. The values, which are obtained individually for the two data-taking conditions and the two track types, are listed in

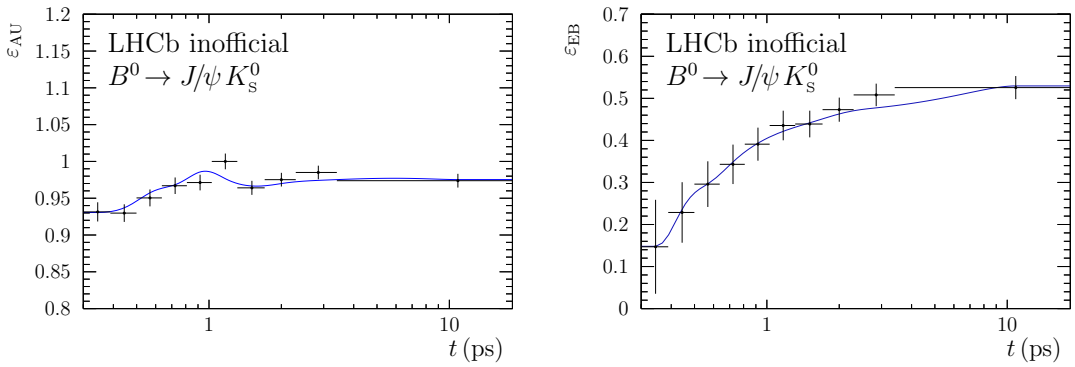

Figure 6.1: Histograms of the trigger acceptance for the almost unbiased (left) and the exclusively biased (right) sample. The blue curve shows the fitted acceptance using splines.

Table 6.1: Decay-time correction factor β_τ in ps^{-1}.

	2011	2012
downstream	0.0036 ± 0.0029	0.0084 ± 0.0032
long track	0.018 ± 0.004	0.035 ± 0.005

Table 6.1. It is apparent that the correction factors are roughly two times larger in 2012 than in 2011 and four to five times larger for long tracks than for downstream tracks. While the former can be explained with changes in the reconstruction software, the latter comparison clearly shows that the VELO is the main reason for this acceptance effect.

6.3 Decay time resolution

The most obvious effect of the decay time resolution are candidates that are reconstructed with negative decay times. However, these are ideal candidates to determine the decay time resolution. An unbinned likelihood fit to the decay time distribution of an unbiased $B^0 \to J/\psi K^0_S$ sample with true J/ψ candidates (selected by a fit of the invariant $m_{\mu^+\mu^-}$ mass distribution) is performed. The fit model consists of a component for the prompt peak around 0 ps, *i.e.* the decay time resolution model, a component to describe events where a wrong PV has been associated and therefore a large difference between true and reconstructed decay time occurs, and long-lived components. The long-lived components are parametrised with two exponentials, which have different pseudo lifetimes and which are themselves convolved with the decay time resolution model. The decay time resolution depends on characteristics of the event. The DecayTreeFit provides predictions for the per-event decay time resolution σ_t, which can be used as the

width of a Gaussian resolution model. However, these predictions are not perfect and a calibration needs to be applied. Additionally, to account for different sources causing the decay time resolution an effective model with two Gaussian functions, which share a common mean but have different calibrations, is used.

The first step is to find a reasonable calibration model. A linear (f_1) and a quadratic (f_2) calibration model are tested:

$$\begin{aligned} f_1: \quad \sigma_{\text{true}}(\sigma_t) &= & b_i\,\sigma_t + c_i\ , \\ f_2: \quad \sigma_{\text{true}}(\sigma_t) &= \alpha_i\,\sigma_t^2 + & \beta_i\,\sigma_t + \gamma_i\ . \end{aligned} \tag{6.2}$$

The data sample is divided into 20 equally filled bins of the decay time resolution predictions σ_t. This is done separately for the downstream and the long track sample as the decay time resolution of long track candidates is expected to be significantly better. Under the assumption that σ_t is constant inside the bins average widths can be set for the two Gaussian functions. An unbinned likelihood fit to the decay time distribution, simultaneous in all bins, sharing all fit parameters except the widths of the Gaussian resolution functions is performed. The widths are plotted in the corresponding bins and a χ^2-fit with the two calibration functions is executed. For the downstream sample the results are depicted in Fig. 6.2. Both functions fit

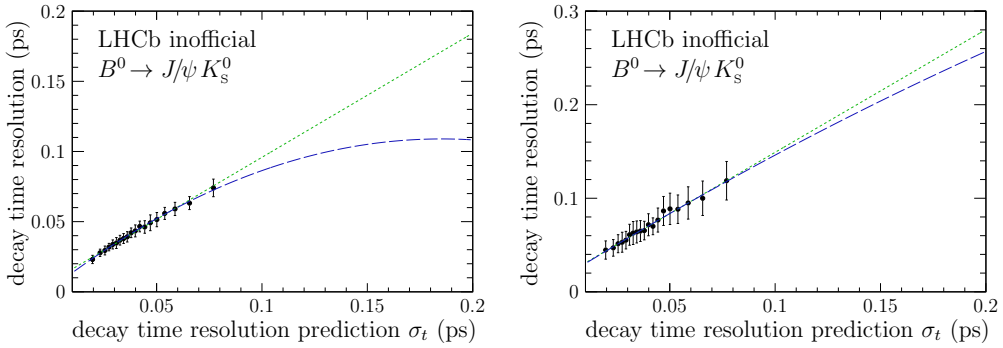

Figure 6.2: Fit of a linear (green short-dashed) and a quadratic calibration function (blue long-dashed) to the narrower (left) and to the wider width (right) of the downstream sample.

equally well, so the simpler linear model with less degrees of freedom is preferred. The linear model is also chosen for the long track sample.

With the calibration functions at hand an unbinned likelihood fit in decay times and decay time resolution predictions is performed and the nominal values of the calibration parameters are determined. The dilution factor induced by the decay time resolution (see Eq. (5.13)) is calculated to be 0.986 for downstream and 0.989 for long track candidates.

6.4 Backgrounds

Although $B^0 \to J/\psi\, K^0_S$ is an experimentally very clean decay channel, care has to be taken to properly identify, suppress or even reject backgrounds. While the two muons can be identified quite effectively, the pions of the K^0_S decay might actually be kaons or protons that have been misidentified. This would lead to background contributions from $B^0 \to J/\psi\, K^{*0}$ and $\Lambda^0_b \to J/\psi\, \Lambda$. To analyse the $p \to \pi$ misidentification the proton mass hypothesis is assigned to one of the pions and the invariant mass of the proton-pion pair $m_{p\pi}$ is recalculated. An excess of candidates at the Λ mass $M_\Lambda = 1115.683\,\mathrm{MeV/c^2}$ [124] can be seen, which is reduced by applying a tighter requirement on the difference of the proton-pion log-likelihood for candidates close to M_Λ. With $\Lambda^0_b \to J/\psi\, \Lambda$ signal MC it is checked that after reconstruction, stripping and all offline selection requirements, including the veto described above, the expected yield is a sub-percent effect. For $K \to \pi$ misidentification the broad width of the K^{*0} does not allow an analogous approach. But studies on $B^0 \to J/\psi\, K^{*0}$ MC show that the expected contribution is even lower than for $\Lambda^0_b \to J/\psi\, \Lambda$. The main reason is the short lifetime of the K^{*0}, which is exploited by the lifetime significance cut on the K^0_S. So, it can basically be assumed that besides the signal candidates almost only combinatorial background is present in the data sample. Nevertheless, it has to be studied whether the background shows any tagging-dependent asymmetry, which would dilute the measured *CP* asymmetry.

By performing a fit to the invariant mass distribution the ${}_s\mathcal{P}lot$ technique [116] provides a possibility to study the tagging-dependent distributions of the combinatorial background. First of all, the time-integrated asymmetry

$$\mathcal{A}^{\mathrm{int}}_{\mathrm{bkg}} = \frac{N^{\bar{B}^0}_{\mathrm{bkg}} - N^{B^0}_{\mathrm{bkg}}}{N^{\bar{B}^0}_{\mathrm{bkg}} + N^{B^0}_{\mathrm{bkg}}} \tag{6.3}$$

is calculated for both track type categories and separately for the OS tagging combination and the SSπ tagging algorithm. Out of the four values listed in Table 6.2 only the one for the downstream OS tagged sample, which has the highest statistics, stands out, as it disfavours *CP* symmetry at more than 3 standard deviations.

Table 6.2: Decay-time-integrated asymmetry of sweighted background distributions and p-value for compatibility of decay-time-dependent asymmetry with flat line at zero for downstream (DD) and long track (LL) OS and SSπ tagged events.

category	$\mathcal{A}^{\mathrm{int}}_{\mathrm{bkg}}$	p-value
DD OS	0.017 ± 0.005	0.100
DD SSπ	-0.016 ± 0.011	0.437
LL OS	-0.005 ± 0.012	0.617
LL SSπ	0.044 ± 0.034	0.969

But even decay-time-integrated asymmetries compatible with zero do not exclude decay-time-dependent asymmetries. The latter are investigated by plotting the raw background asymmetry in ten bins of the decay time with bin sizes chosen to increase exponentially (see Fig. 6.3). However, it is difficult to judge by eye if

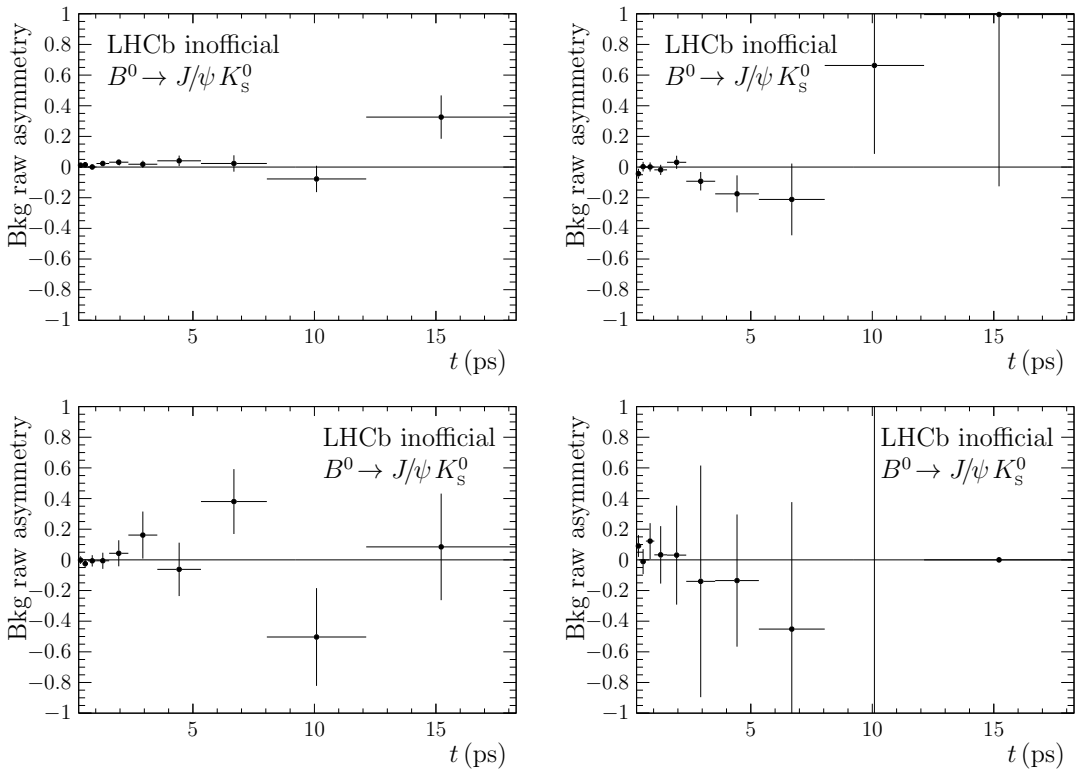

Figure 6.3: Raw background asymmetry in sweighted data with logarithmic binning. DD in top, LL in bottom plots. OS tagged candidates on the left, SSπ tagged candidates on the right side.

there is a significant oscillation. Thus, χ^2-tests against the null-hypothesis, *i.e.* a flat line at zero, are performed and the corresponding p-values are calculated (see Table 6.2). None of the p-values is very low, so the deviations from zero can be explained with statistical fluctuations. The same procedure (mass fit $\to$ sWeights $\to$ histograms of $\mathcal{A}_{\text{bkg}}$) is performed with cocktail MC consisting of signal MC and background Toy MC. Here, no asymmetry is generated for the background and the resulting p-values are very similar to the ones for the nominal data sample. Finally, an unbinned likelihood fit is performed to the sweighted background decay time distribution using two/three exponentials with different pseudo-lifetimes and allowing for tagging-dependent asymmetries in the PDF. All asymmetry parameters are compatible with zero at a significance of two standard deviations.

6.5 Nominal fit

A multidimensional fit to the distributions of the reconstructed mass m, the decay time t and its error prediction σ_t, the OS and SSπ tags and mistag probabilities is performed, mainly to extract the *CP* observables $S_{J/\psi K^0_S}$ and $C_{J/\psi K^0_S}$. Thanks to the selection, which cleans up the sample from other background contributions, besides the signal component only the combinatorial background component needs to be parametrised.

The mass distribution, which enables the best discrimination of the two components, is modelled by a modified Hypatia PDF [127] and an exponential function. All shape parameters are allowed to differ between the two track type categories. The four tail parameters of the signal parametrisation are taken from fits to simulated events.

The PDF describing the signal decay time distribution is basically given by

$$\frac{\mathrm{d}\Gamma(t,d)}{\mathrm{d}t} \propto \mathrm{e}^{-t/\tau}\Big(1 - d\,S_f \sin\left(\Delta m_d\, t\right) + d\,C_f \cos\left(\Delta m_d\, t\right)\Big)\,, \tag{6.4}$$

with $d = +1$ for B^0 and $d = -1$ for $\overline{B}^0$. The decay width difference $\Delta\Gamma_d$ is set to zero. The theoretical parametrisation is extended by taking into account the production asymmetry and the experimental effects of mistagging. Furthermore, it is convolved with the decay time resolution model presented in Sec. 6.3 and multiplied by the decay-time-dependent efficiency correction for low and high decay times developed in Sec. 6.2. The background decay time distribution is parametrised with the sum of exponential functions. In extensive studies using sweighted background distributions (see Sec. 6.4) it is determined how many exponential functions give the best description for the categories, into which the whole data sample is divided.

Lognormal distributions have been found to fit best when describing the decay time error distributions. A combination of double and single lognormals with some parameters being shared among the categories is used. Especially the downstream and the long track distributions differ as can be seen in Fig. 6.4.

The mistag distributions are described with cubic splines. The seven knots for the SSπ distribution, and similarly the 12 knots for the OS distribution, are positioned where the shape visibly changes. It is checked that the two mistag estimates are uncorrelated with each other and with the decay time, at least within the available statistical precision. This allows to simply multiply the corresponding PDFs.

In the fit 11 external input parameters are constrained within their statistical uncertainties. These are the production asymmetries for 7 and 8 TeV, for which the procedure is explained in more detail in Sec. 7.3.3, the oscillation frequency Δm_d [124] and the flavour-tagging calibration parameters (Sec. 5.7.2).

To stabilise the fit the decay time resolution parameters, the spline coefficients of the mistag parametrisation, the decay time error parameters and all parameters included in the decay time acceptance model are fixed. Thus, 72 floating parameters remain for the fit, of which 48 are yields.

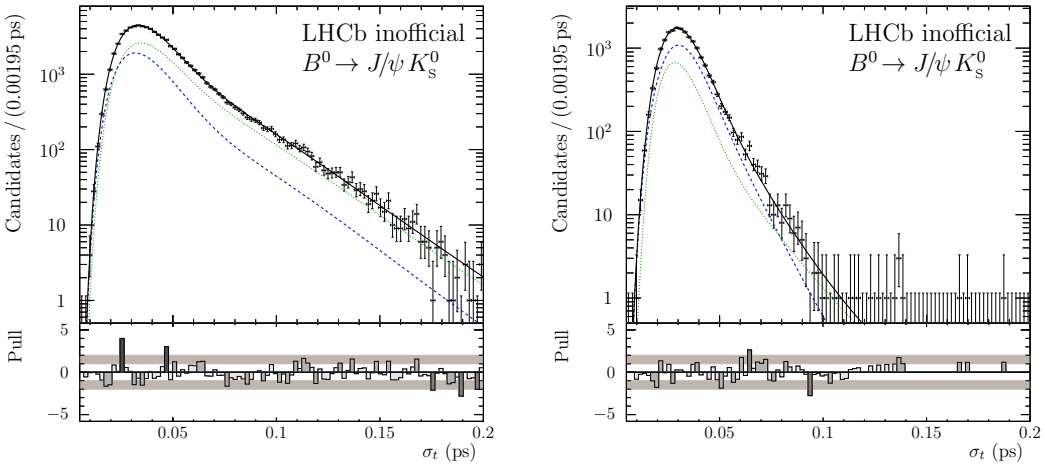

Figure 6.4: Decay-time error distributions of downstream (left) and long track (right). The solid black line shows the fit projection, while the blue dashed (green dotted) line shows the signal (background) component.

From 41 560 flavour-tagged $B^0 \to J/\psi K^0_S$ decays the *CP* observables are determined to be

$$S_{J/\psi K^0_S} = \quad 0.731 \pm 0.035\,(\text{stat}) \pm 0.020\,(\text{syst})\,,$$
$$C_{J/\psi K^0_S} = -0.038 \pm 0.032\,(\text{stat}) \pm 0.005\,(\text{syst})\,,$$

with a statistical correlation of 0.483. In these results corrections of $+0.002$ for $S_{J/\psi K^0_S}$ and -0.005 for $C_{J/\psi K^0_S}$ are included, which account for *CP* violation in K^0–$\overline{K}^0$ mixing and different nuclear cross-sections in material between K^0 and $\overline{K}^0$ [128].

The distributions of the invariant mass and the decay time are depicted in Fig. 6.5.

6.6 Studies of systematic effects

To check if and how various effects systematically influence the measurement of the *CP* violation parameters, the likelihood fit is performed

- with a second independent fitter [126],
- without a parametrisation of the background using sWeights extracted with the ${}_s\mathcal{P}lot$ technique,
- on subsamples split by the K^0_S track type, the trigger requirements, the tagging algorithms, the magnet polarity, and the year of data-taking.

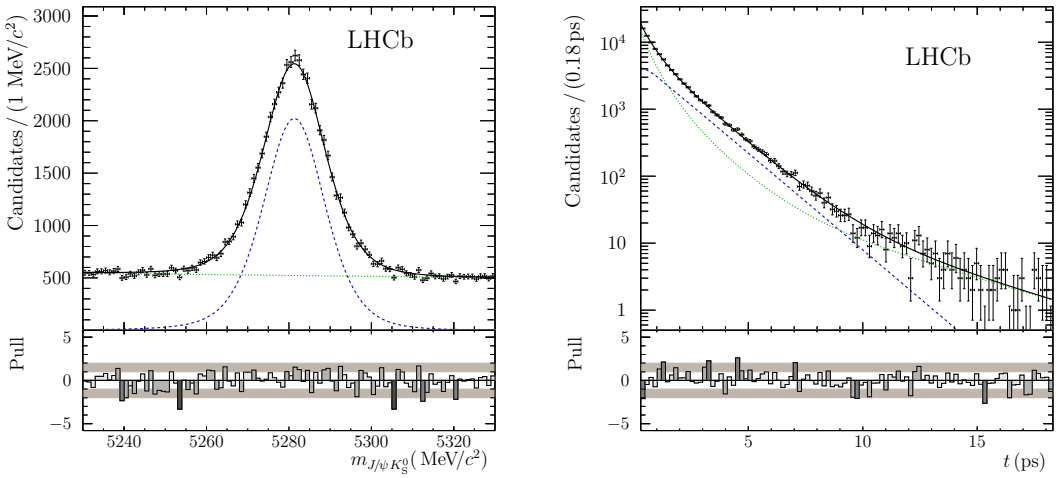

Figure 6.5: Invariant mass distribution (left) and decay time distribution (right). The solid black line shows the fit projection, while the blue dashed (green dotted) line shows the signal (background) component.

All results show good agreement with the nominal results. Additionally, the results from a pure time-dependent and from a pure time-integrated fit are compatible with each other and with the nominal fit, which comprises both effects.

Systematic uncertainties from several effects, especially from possible mismodelling of PDFs and from systematic uncertainties on external input parameters, are considered. Pseudoexperiments are generated using PDFs that contain a slight modification compared to the nominal PDF, which is used for the subsequent fit of the samples. Whenever the mean of the pull distribution exceeds zero by more than 0.032, a systematic uncertainty is assigned. Pull distributions show the difference between the individual fit result and the generation value for the *CP* observable normalised by the fit uncertainty. The value of the criterion is defined by one standard deviation according to the statistics of 1000 pseudoexperiments performed for each study. The size of the systematic uncertainty is taken from the residual distributions, which show the same as the pull distributions apart from not being normalised.

The possible tagging asymmetry of the background contribution is treated as a source for a systematic uncertainty on the *CP* observables using the asymmetry parameters determined from an sweighted fit to the background decay time distribution (see last part of Sec. 6.4). It is found to account for 83 % of the systematic uncertainty on $S_{J/\psi K^0_S}$ and 8 % for $C_{J/\psi K^0_S}$. The systematic uncertainties on the flavour-tagging calibration parameters (see Sec. 5.7.2) are transferred to systematic uncertainties on $S_{J/\psi K^0_S}$ and $C_{J/\psi K^0_S}$ by shifting in the generation of the pseudoexperiments all calibration parameters related to p_0 upwards by one systematic uncertainty and those related to p_1 downwards, while setting them to

their nominal values in the fit. This yields 9 % of the systematic uncertainty on $S_{J/\psi K^0_S}$, and 21 % for $C_{J/\psi K^0_S}$. The assumption $\Delta\Gamma_d = 0$ is responsible for 6 % of the systematic uncertainty on $S_{J/\psi K^0_S}$. It is determined generating the pseudoexperiments setting $\Delta\Gamma_d = 0.007$ ps^{-1}, the current experimental uncertainty [124]. The largest contribution (42 %) to the systematic uncertainty on $C_{J/\psi K^0_S}$ arises from the systematic uncertainty on the world average of Δm_d. Further effects, whose possible systematic effect are analysed, are the decay time resolution model, the uncertainty on the length scale of the vertex detector, the decay time acceptance model, the correlation between the invariant mass and the decay time, and the production asymmetry. They are all small or even negligible compared to the previously described effects.

The values of all individual systematic uncertainties and of the total systematic uncertainty, which is calculated as the sum of all contributions in quadrature, are listed in Table 6.3. The systematic uncertainty on $S_{J/\psi K^0_S}$ of ± 0.020 is more than 40 % smaller than the statistical uncertainty. For $C_{J/\psi K^0_S}$ the total uncertainty even increases by only 1 % through the systematic uncertainty.

Table 6.3: Systematic uncertainties on $S_{J/\psi K^0_S}$ and $C_{J/\psi K^0_S}$. Entries marked with a dash represent studies where no significant effect is observed.

Origin	$\sigma_{S_{J/\psi K^0_S}}$	$\sigma_{C_{J/\psi K^0_S}}$
Background tagging asymmetry	0.018	0.0015
Tagging calibration	0.006	0.0024
$\Delta\Gamma_d$	0.005	—
Fraction of wrong PV component	0.0021	0.0011
z-scale	0.0012	0.0023
Δm_d	—	0.0034
Upper decay time acceptance	—	0.0012
Correlation between mass and decay time	—	—
Decay time resolution calibration	—	—
Decay time resolution offset	—	—
Low decay time acceptance	—	—
Production asymmetry	—	—
Sum	0.020	0.005

7 Measurement of CP Violation in $B^0 \to D^+D^-$ Decays

In this chapter the analysis of $B^0 \to D^+D^-$ decays with the goal to determine the observables $S_{D^+D^-}$ and $C_{D^+D^-}$, which describe the CP violation in this decay mode, is presented [129]. After a description of the selection (see Sec. 7.1) the fit of the invariant mass distribution, performed to extract signal sWeights, is described (see Sec. 7.2). This is followed by a summary of the decay time fit (see Sec. 7.3). The chapter is concluded with a presentation of the studies on systematic uncertainties (see Sec. 7.4). The new OS combination including the OS charm tagger and the SS combination of SSπ and SSp tagger are used to determine the flavour tag of the B^0 mesons. The calibration of these flavour-tagging algorithms using $B^0 \to D_s^+D^-$ decays (see Sec. 5.7.1) is the only part not performed by myself but by collaborators from Milano.

7.1 Selection

The amount of background in $B^0 \to D^+D^-$ is too high to perform a significant measurement of CP violation without any selection. The selection is divided into three parts: a preselection with many high signal efficiency requirements, a dedicated treatment of misidentified backgrounds and a multivariate analysis to further reduce combinatorial background.

7.1.1 Preselection

Only events that have been triggered by a topological trigger line or by the inclusive ϕ line and that in total contain less than 500 long tracks are considered. All candidate kaon and pion tracks have to be long tracks and have to satisfy quality criteria. Lower limits on the momentum ($p > 1\,\mathrm{GeV/c}$) and on the transverse momentum ($p_\mathrm{T} > 100\,\mathrm{MeV/c}$) are required. The candidates should be inconsistent with originating from the PV and the particle identification (PID) system needs to classify them as pions or kaons with only a small probability to be a ghost.

Of all the possible combinations of three charged hadron tracks forming a D^+ meson candidate only the two possibilities $D^\pm \to K^\mp\pi^\pm\pi^\pm$ and $D^\pm \to K^\mp K^\pm\pi^\pm$ are selected. The vertex needs to be significantly displaced from all PVs in the event and the distance of the closest approach between all pairs of particles forming the vertex has to be below 0.5 mm. The scalar sum of the p_T has to exceed 1800 MeV/c

and the combined invariant mass has to be in the range $\pm 25\,\mathrm{MeV/c^2}$ around the nominal D^+ mass [124]. The tightened mass window as well as requiring that the χ^2 of the flight distance of each $D^\pm$ meson with respect to the B^0 decay vertex has to be larger than 2 reduces the amount of (partially) charmless contributions. On top of that, a cut on the decay time significance of the $D^\pm$ mesons, defined as their decay time with respect to the B^0 decay vertex divided by the corresponding uncertainty, is supposed to further suppress the (partially) charmless contributions. The optimal cut value is estimated under the assumption that a very tight cut leaves only candidates with resonant $D^\pm$ mesons. Gradually loosening the cut the value can be found where the product of the B^0 signal yield, extracted from a fit on data, and the signal efficiency, determined on MC, exceeds the estimation from the initial tight cut scenario. If both $D^\pm$ mesons are reconstructed via $D^\pm \to K^\mp\pi^\pm\pi^\pm$ decays, the decay time significance has to be greater than 0. It needs to be greater than 3 if one of the $D^\pm$ mesons is reconstructed in the $KK\pi$ and the other in the $K\pi\pi$ final state. Although in this case the final states of the $D^\pm$ mesons differ, the same cut is applied to both $D^\pm$ mesons, as on signal MC the comparison of the distributions of the decay time significance shows a good agreement between $D^\pm \to K^\mp\pi^\pm\pi^\pm$ and $D^\pm \to K^\mp K^\pm\pi^\pm$ decays.

The vertex formed by a pair of oppositely charged $D^\pm$ candidates needs to be of good quality. The scalar sum of the p_T of the $D^\pm$ mesons must exceed $5\,\mathrm{GeV/c}$. In the stripping a BDT to select B^0 candidates is applied. The BDT is based on the p_T and the flight distance χ^2 of the B^0 as well as on the sum of the B^0 and both D vertex χ^2 divided by the sum of the degrees of freedom of these vertex fits. Moreover, the B^0 candidates are required to have $p > 10\,\mathrm{GeV/c}$ and to have $\chi^2_\mathrm{IP} < 25$, where χ^2_IP is defined as the difference in the vertex fit χ^2 of the associated PV with and without the B^0 candidate.

The reconstructed decay time t of the B^0 candidate is determined from a DecayTreeFit [92], in which the B^0 production vertex is constrained to the position of the associated PV. The invariant mass $m_{D^+D^-}$ of the B^0 candidate is calculated from a DecayTreeFit, in which the invariant masses of $K\pi\pi$ and $KK\pi$ are additionally constrained to the known D^+ mass. It is required that these fits have converged. Further outliers are removed by requiring that the uncertainties on the invariant mass and on the decay time have to be below $30\,\mathrm{MeV/c^2}$ and 0.2 ps, respectively, and that the absolute value of the z coordinate of the PV is smaller than 250 mm.

The signal efficiency of the preselection is 82 % for the final state, where both $D^\pm$ mesons are reconstructed as $K\pi\pi$, at a background rejection of 94 %. For the second final state, which in total contains three kaons, the signal efficiency is 67.5 % at a background rejection of 98 %. The total number of candidates after the preselection is about 110 000. The emerging mass distribution in Fig. 7.1 shows that most of these candidates are due to background contributions. In fact, the highest peak in the mass spectrum is not arising from the signal decay but from misidentified background. The vetoes, which are applied to suppress this contribution, are described in the next section.

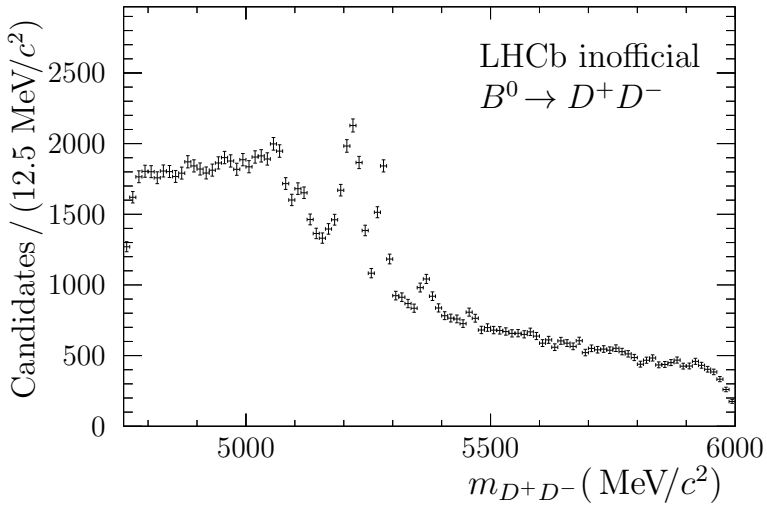

Figure 7.1: Invariant mass distribution of $B^0 \to D^+D^-$ decays after all preselection requirements are applied.

7.1.2 Vetoes

Misidentification of kaons as pions can lead to background contributions from $D_s^+ \to K^-K^+\pi^+$, which predominantly proceeds through $D_s^+ \to \phi\pi^+$. To reduce these D_s^+ contributions the kaon mass hypothesis is assigned to the pion with the higher transverse momentum of $D^+ \to K^-\pi^+\pi^+$ candidates. The candidate is rejected if the invariant mass of the hypothetical kaon pair is compatible with the ϕ mass of $M_\phi = 1019.461\,\mathrm{MeV/c^2}$ [124] within $\pm 10\,\mathrm{MeV/c^2}$. It is also rejected if the invariant mass of the $KK\pi$ combination is compatible with the D_s^+ mass of $M_{D_s^+} = 1968.30\,\mathrm{MeV/c^2}$ [124] within $\pm 25\,\mathrm{MeV/c^2}$ and at the same time the pion with the higher p_T (the one that the kaon mass hypothesis is assigned to) is more kaon- than pion-like. When the kaon mass hypothesis is assigned to the pion with the lower p_T, no vetoes are applied as no resonant structures at the ϕ or the D_s^+ mass are found.

To reduce $p \to \pi$ misidentification the proton mass hypothesis is assigned to the pion with the higher p_T of $D^+ \to K^-\pi^+\pi^+$ candidates. The candidate is rejected if the invariant mass of the $Kp\pi$ combination is compatible with the Λ_c^+ mass of $M_{\Lambda_c^+} = 2286.46\,\mathrm{MeV/c^2}$ [124] within $\pm 25\,\mathrm{MeV/c^2}$ and the proton probability of the pion with the higher p_T is larger than the pion probability.

7.1.3 Multivariate analysis

BDT training

To further suppress combinatorial background a Boosted Decision Tree (BDT) [109, 112] based on the implementation in TMVA [107] is trained using a signal MC

sample and the upper mass sideband with $m_{D^+D^-} > 5500\,\mathrm{MeV/c^2}$. The training is performed on half of these samples, while the other half is used to test the BDT performance. The selection steps described above, are applied before the training.

Two BDTs separated by the number of kaons in the B^0 final state are trained. The importance of the 21 features of the training differs, which is considered by their order in Table 7.1. One of the features is the ratio of the kaon over the sum

Table 7.1: List of features used in the training of the BDTs.

BDT for $K^-\pi^+\pi^+K^+\pi^-\pi^-$	BDT for $K^-K^+\pi^+K^+\pi^-\pi^-$
min($D^\pm$ τ significance)	PID ratio of $K^\pm$
B direction angle	B direction angle
log(DTF χ^2/ndof)	PID ratio of K^+
PID ratio of K^-	log(DTF χ^2/ndof)
PID ratio of K^+	PID ratio of K^-
min p_T of $K^\pm$	min($D^\pm$ τ significance)
log(B impact parameter χ^2)	log(min(h Velo χ^2/ndof))
log(min($\pi^\pm$ Velo χ^2/ndof))	p_T of $K^\pm$
p_T of π^- with lower p_T	log(min($K^\pm$ T-track χ^2/ndof))
log(min($K^\pm$ T-track χ^2/ndof))	log(B impact parameter χ^2)
log(min($\pi^\pm$ T-track χ^2/ndof))	PID ratio of $\pi^\pm$ with lower p_T
PID ratio of π^- with higher p_T	log(min(h VELO-T-Match χ^2))
p_T of π^+ with lower p_T	log(min($K^\pm$ Velo χ^2/ndof))
PID ratio of π^- with lower p_T	PID ratio of single $\pi^\pm$
PID ratio of π^+ with higher p_T	p_T of $\pi^\pm$ with higher p_T
p_T of π^+ with higher p_T	log(min(h T-track χ^2/ndof))
PID ratio of π^+ with lower p_T	p_T of $\pi^\pm$ with lower p_T
log(min($K^\pm$ Velo χ^2/ndof))	min p_T of K^+ and K^-
log(min($\pi^\pm$ VELO-T-Match χ^2))	p_T of single $\pi^\pm$
log(min($K^\pm$ VELO-T-Match χ^2))	log(min($K^\pm$ VELO-T-Match χ^2))
p_T of π^- with higher p_T	PID ratio of $\pi^\pm$ with higher p_T

of the kaon and the pion probabilities:

$$\text{PID ratio} = \frac{\texttt{ProbNN}K}{\texttt{ProbNN}K + \texttt{ProbNN}\pi}\,. \tag{7.1}$$

It turns out that this ratio performs a little bit better than just using the simple ProbNN variables. Among the other features are observables related to the kinematics of the decay like transverse momenta, decay time significances and direction angles, qualities of the track segments in the VELO and the T-stations, and vertex qualities.

Before the training the features are transformed to decorrelate and decompose them into the principal components, which improves the performance of the BDT.

The BDTs are each built out of 700 trees. The depth of the trees is limited to three. At each node at least 3 % of the training events have to be present. The variables are scanned at 40 points to find the optimal cut value. For the boosting the AdaBoost method [113] with a boost factor of $\beta = 0.1$ is deployed.

Overtraining is checked by applying the BDT on both the training and the testing sample. As can be seen in Fig. 7.2, the classifier output distributions are compatible, which means that no inherent overtraining is present for the BDT. However, using

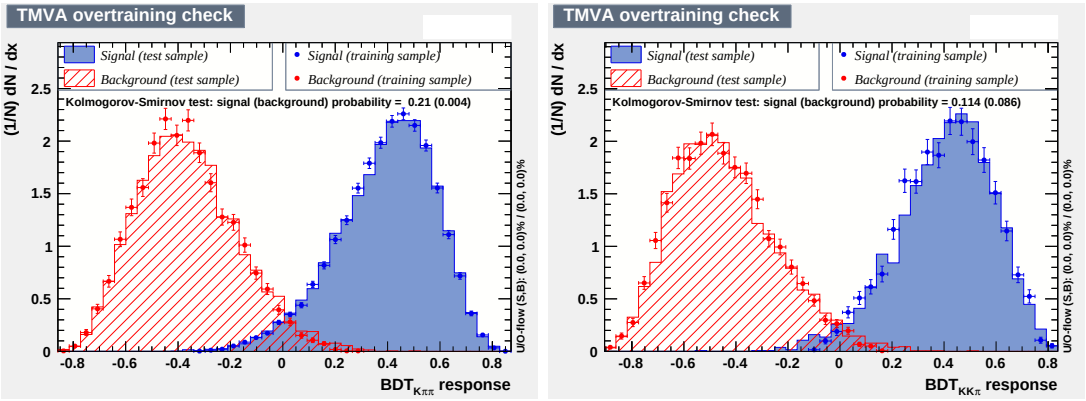

Figure 7.2: Comparison of the BDT response on training and test sample for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state (left) and the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state (right).

simulated data in the selection involves the possibility that certain distributions are not modelled properly and differences between the simulation and real data are exploited instead of differences between signal and background. Indeed, the classifier output distributions of signal MC and of background-subtracted data show a quite large disagreement for both final states, as can be seen in Fig. 7.3. The distributions of signal MC are shifted towards higher, thus more signal-like,

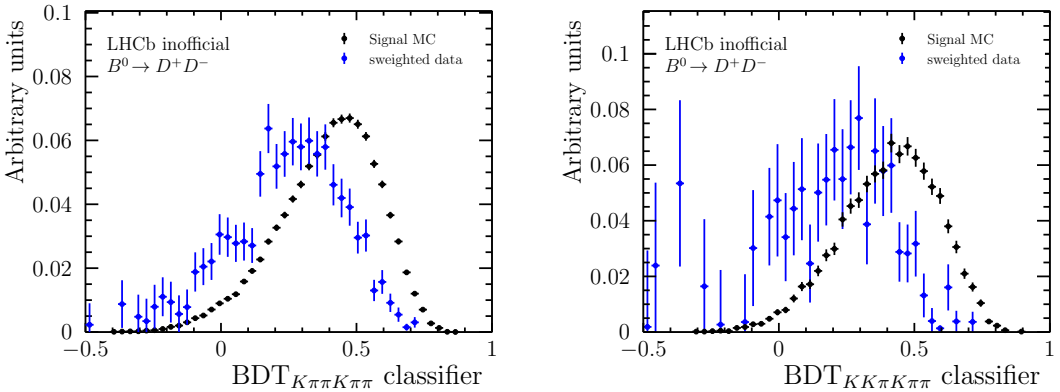

Figure 7.3: Distribution of the BDT output for background-subtracted data (blue) and signal MC (black) for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state (left) and the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state (right).

values. This means that the performance of the BDT classifier is overestimated in the training. If selection efficiencies had to be calculated using the MC sample, this would be a problem. But for a measurement of *CP* violation it is mainly important that the amount of background can somehow be reduced, while most of the signal is kept. This can be achieved with the given setting.

BDT cut optimisation

As explained in Sec. 5.2.3 the best figure of merit for a measurement of *CP* violation is the sensitivity on the *CP* observables themselves. Thus, the requirement on the BDT classifier output is scanned performing a fit to the invariant D^+D^- mass spectrum followed by a decay time fit of the background-subtracted sample for each scan point. Initially, only the subsample with two kaons in the B^0 final state is analysed. In Fig. 7.4 the statistical uncertainties of $S_{D^+D^-}$ and $C_{D^+D^-}$ are plotted as a function of the requirement on the BDT classifier output. The

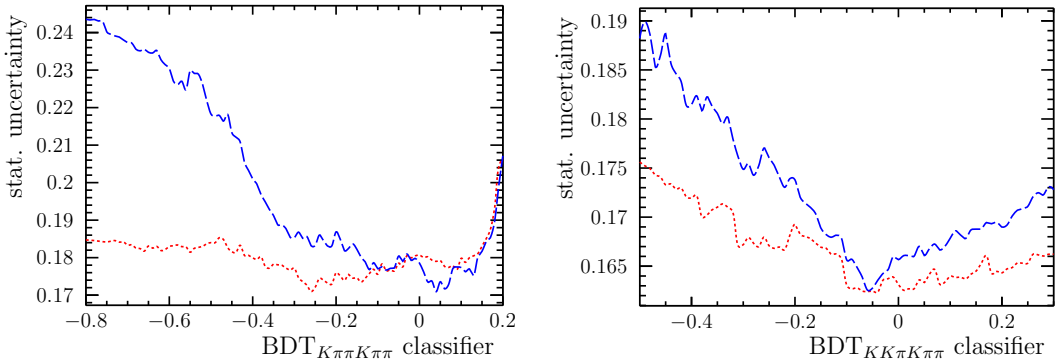

Figure 7.4: Statistical uncertainty on $S_{D^+D^-}$ (red short-dashed) and $C_{D^+D^-}$ (blue long-dashed) as a function of the BDT classifier output for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state (left) and the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state (right).

uncertainty on $C_{D^+D^-}$ improves with tighter requirements on the BDT classifier, until it reaches an optimum just above zero. This can be explained with the fact that the sensitivity on $C_{D^+D^-}$ mainly comes from candidates at low decay times because the cosine function is maximal there. The suppression of the rather short-lived combinatorial background compensates the loss in signal efficiency for a quite long range. In contrast, the uncertainty on $S_{D^+D^-}$ is mainly driven by the amount of signal candidates. Thus, it is more or less flat for loose requirements on the BDT classifier, where only few signal candidates are lost and this loss is compensated by the higher purity, and reaches its optimum around -0.25, before it starts to get worse. Now that both observables are of interest and the optima are not at the same cut value, it is decided to require the BDT classifier to be greater than -0.10. This is a good compromise between both observables as the uncertainties of $S_{D^+D^-}$ and $C_{D^+D^-}$ are almost the same and close to their optima. The requirement has a

signal efficiency of $(96.5 \pm 0.5)\,\%$ and rejects $(84.18 \pm 0.34)\,\%$ of the combinatorial background.

In a second step the requirement on the BDT classifier for the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state is optimised. The $K^-K^+\pi^+K^+\pi^-\pi^-$ subsample is quite small, which makes individual fits on this subsample rather unstable. This can be solved by performing a simultaneous fit to the whole data sample with the previously determined BDT cut applied to the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ subsample. Scanning the BDT classifier output for the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state results in the sensitivities on $S_{D^+D^-}$ and $C_{D^+D^-}$ plotted in Fig. 7.4. Both uncertainties show a minimum at around -0.05, which is chosen as cut value. This cut removes $(90.75 \pm 0.33)\,\%$ of the combinatorial background at a signal efficiency of $(87.2 \pm 1.9)\,\%$.

7.1.4 Final selection

Finally, the fit range of the invariant $m_{D^+D^-}$ mass is reduced to 5150–5500 $\mathrm{MeV/c^2}$, which eliminates some backgrounds, like misreconstructed $B^0 \to D^{*+}D^-$, at low masses, prevents overtraining effects on the high-mass sideband used in the training of the BDT, and leaves enough candidates in the upper mass sideband to determine the shape of the combinatorial background. Additionally, the decay time is restricted to be in the range 0.25–10.25 ps. Already in the stripping the candidates are required to have decay times greater than 0.2 ps. But as in the stripping the LokiVertexFitter and in the analysis the DecayTreeFitter is used, the lower decay time boundary is not well defined. That is the reason why the cut is tightened a little bit. Very few signal candidates are lost by this requirement, while the determination of a smooth decay time acceptance (see Sec. 7.3.2) is simplified. In 0.8 % of the selected events more than one candidate remains, which is very unlikely given the low branching fraction. Therefore, only one of the multiple candidates is kept, which is chosen randomly.

After all selection steps 6209 candidates remain in the data sample, of which about 25 % are signal candidates. The exact number of signal candidates is determined in a mass fit described in the next section.

7.2 Mass fit

In this section the fit of the invariant $m_{D^+D^-}$ mass distribution is described, which is used to calculate signal weights via the ${}_s\mathcal{P}lot$ method [116], and thereby discriminates between signal and background candidates. As the linear Pearson correlation coefficient between the invariant mass and the decay time is determined to be $\rho = 0.007$, it is valid to apply the sWeights in the decay time fit to obtain the CP observables (see Sec. 7.3).

The mass distribution is parametrised with a PDF $\mathcal{P}$ consisting of five components, $B^0 \to D^+D^-$ signal, $B_s^0 \to D^+D^-$ background, background from $B^0 \to D_s^+D^-$,

background from $B_s^0 \to D_s^- D^+$, and combinatorial background:

$$N^s \mathcal{P}^s = N^s_{B^0} \mathcal{P}^s_{B^0} + N^s_{B_s^0} \mathcal{P}^s_{B_s^0} + N^s_{B^0 \to D_s^+ D^-} \mathcal{P}^s_{B^0 \to D_s^+ D^-} + N^s_{B_s^0 \to D_s^- D^+} \mathcal{P}^s_{B_s^0 \to D_s^- D^+} + N^s_{\text{Bkg}} \mathcal{P}^s_{\text{Bkg}} \quad . \tag{7.2}$$

In the extended maximum likelihood fit four disjoint categories are simultaneously fitted. It is distinguished between the two years of data-taking 2011 and 2012 and between the two final states $K^-\pi^+\pi^+K^+\pi^-\pi^-$ and $K^-K^+\pi^+K^+\pi^-\pi^-$ ($s = \{2011, K\pi\pi\}, \{2011, KK\pi\}, \{2012, K\pi\pi\}, \{2012, KK\pi\}$). The tagging output is not split. In Fig. 7.5 the complete data sample is plotted overlaid with the PDF projections and its components. The individual shapes are explained in the following.

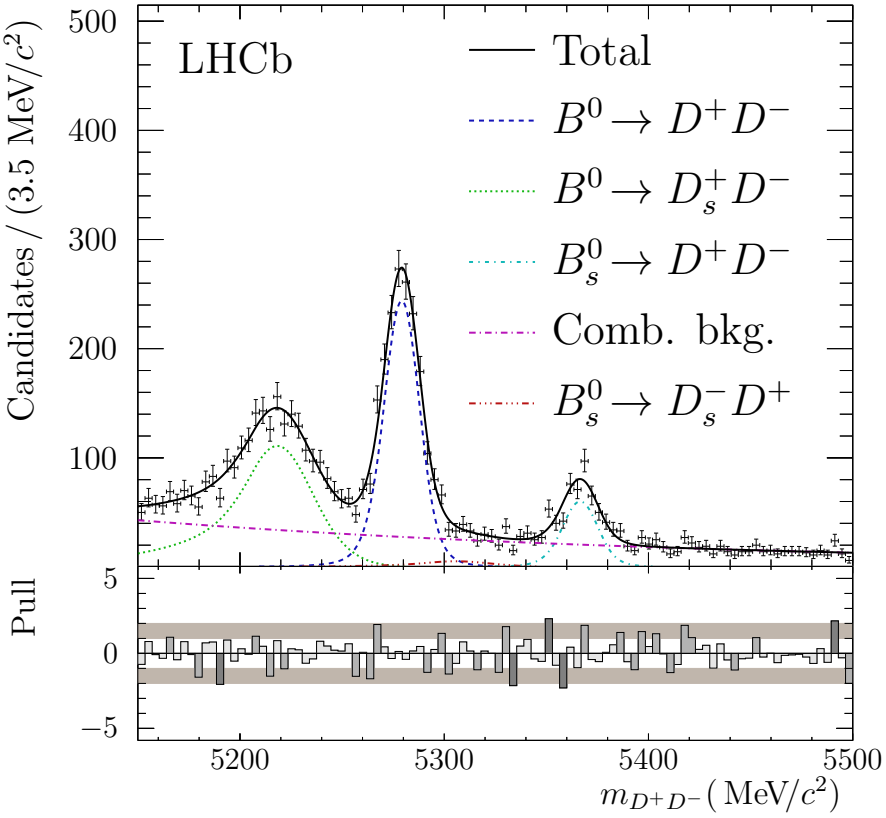

Figure 7.5: Plot of the reconstructed mass of the $B^0 \to D^+D^-$ data sample with the projected PDF and pull distribution.

$B^0 \to D^+D^-$ signal: The $B^0 \to D^+D^-$ signal mass component is modelled by the sum of three Crystal Ball functions [130], which share a common peak position μ_{B^0} but have different width parameters σ_i. Two of the Crystal Ball functions have a tail towards lower masses and one has a tail towards higher masses. The parameters α_1 to α_3 of the power law functions, the ratio between the widths, and

the fractions f_1 and f_2 between the Crystal Ball functions are determined from a fit to the invariant D^+D^- mass distribution of $B^0 \to D^+D^-$ signal MC in the range 4800–5400 MeV/c^2. This MC sample consists of both final states generated in the ratio of the current world averages [22]. Apart from the mass range the full selection is applied. Events, where photons are missed in the reconstruction, create the very long tail towards lower masses, which requires the third Crystal Ball function. The exponent of all power law parts is fixed to 10. The widths σ_1^{MC} to σ_3^{MC} are multiplied by a common scale factor R in the fit to data to account for differences in the mass resolution between simulation and data. The fit results are listed in Table 7.2 and a plot of the distribution overlaid with the projection of the PDF is given in Fig. 7.6.

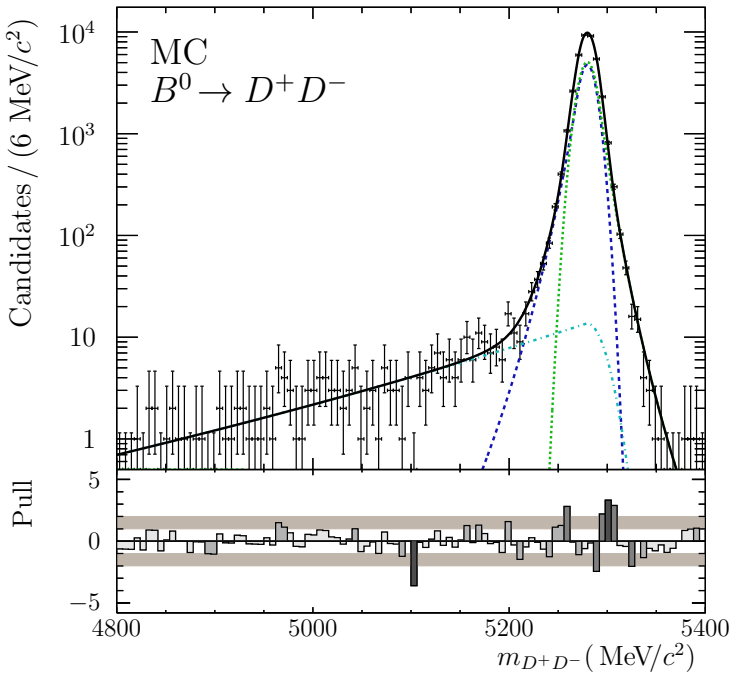

Figure 7.6: Mass distribution of the $B^0 \to D^+D^-$ signal MC sample overlaid with the projection of the fitted PDF. The blue dotted, green dashed and turquoise short-dash-dotted lines represent the three Crystal Ball components.

$B_s^0 \to D^+D^-$ background: Apart from the B^0 also the heavier B_s^0 can decay to the D^+D^- final state. Almost the same parametrisation as for the B^0 signal component is used, *i.e.* same width and tail parameters, while the peak position is shifted by the world average $\Delta m_{B_s^0-B^0} = \mu_{B_s^0} - \mu_{B^0} = 87.35\,\mathrm{MeV/c^2}$ [22].

Table 7.2: Fit results of the mass fit to $B^0 \to D^+D^-$ signal MC.

Parameter		Value
$\mu_{B^0}^{\mathrm{MC}}$	$(\mathrm{MeV/c^2})$	5279.70 ± 0.09
σ_1^{MC}	$(\mathrm{MeV/c^2})$	8.5 ± 0.4
σ_2^{MC}	$(\mathrm{MeV/c^2})$	16 ± 5
σ_3^{MC}	$(\mathrm{MeV/c^2})$	9.0 ± 0.4
f_1^{MC}		0.48 ± 0.06
f_2^{MC}		0.0098 ± 0.0011
α_1^{MC}		1.18 ± 0.08
α_2^{MC}		0.12 ± 0.04
α_3^{MC}		-1.46 ± 0.08

$B^0 \to D_s^+D^-$ background: The vetoes applied in the selection suppress the contribution from misidentified kaons. Nevertheless, a significant amount of $B^0 \to D_s^+D^-$ decays remains in the data sample. A fit to the invariant mass distribution of simulated $B^0 \to D_s^+D^-$ events reconstructed as $B^0 \to D^+D^-$ is performed. The full selection is applied to the simulated sample as this can change the shape of the $B^0 \to D_s^+D^-$ background contribution. The sum of two Crystal Ball PDFs with both power law exponents fixed to 10 is used to parametrise the invariant mass distribution. The fit results are listed in Table 7.3 and the corresponding plot is shown in Fig. 7.7. The fraction parameter and the tail parameters are taken from this fit. The width parameters as well as the peak position are floating parameters in the fit to data.

Table 7.3: Fit results of the mass fit to $B^0 \to D_s^+D^-$ MC.

Parameter		Value
$\mu_{B^0 \to D_s^+D^-}^{\mathrm{MC}}$	$(\mathrm{MeV/c^2})$	5222.2 ± 0.9
$\sigma_{1,D_s^+D^-}^{\mathrm{MC}}$	$(\mathrm{MeV/c^2})$	15.0 ± 1.5
$\sigma_{2,D_s^+D^-}^{\mathrm{MC}}$	$(\mathrm{MeV/c^2})$	20.7 ± 2.1
$f_{1,D_s^+D^-}^{\mathrm{MC}}$		0.78 ± 0.13
$\alpha_{1,D_s^+D^-}^{\mathrm{MC}}$		0.60 ± 0.09
$\alpha_{2,D_s^+D^-}^{\mathrm{MC}}$		-1.8 ± 0.4

$B_s^0 \to D_s^-D^+$ background: Although only few candidates of $B_s^0 \to D_s^-D^+$ decays are expected in the data sample, a component for this contribution is included in the nominal fit. It is parametrised with the sum of two Crystal Ball PDFs. All shape parameters are shared with the B^0 component, apart from the peak position, which is constrained to be $\Delta m_{B_s^0 - B^0}$ above the peak position of the B^0 component.

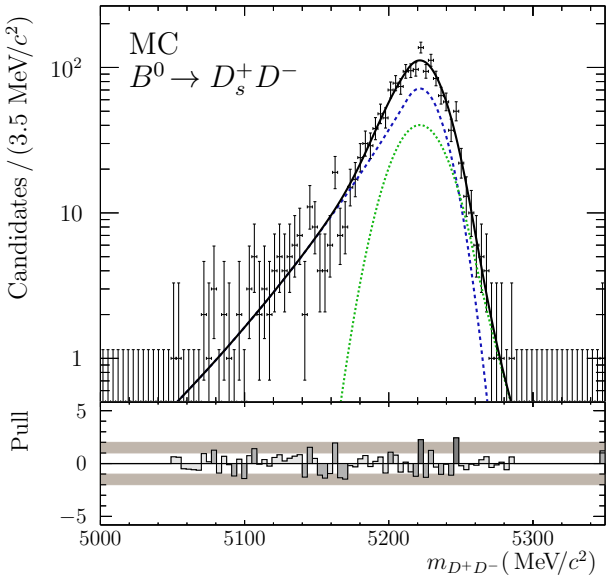

Figure 7.7: Mass distribution of the $B^0 \to D_s^+ D^-$ MC sample reconstructed as $B^0 \to D^+ D^-$ overlaid with the projection of the two Crystal Ball PDFs shown in blue dashed and green dotted.

Combinatorial background: The reconstructed mass PDF of the combinatorial background is modelled by an exponential function with individual slopes $\beta_{K^-\pi^+\pi^+K^+\pi^-\pi^-}$ and $\beta_{K^-K^+\pi^+K^+\pi^-\pi^-}$ based on the number of kaons in the final state.

Total fit: In Table 7.4 the results of the floating shape parameters of the mass fit to data are shown. The B^0 peak position is in good agreement with the current world average of $\mu_{B^0}^{\text{WA}} = (5279.62 \pm 0.15)\,\text{MeV/c}^2$ [22]. The scale factor R is compatible with unity, which means that the mass resolution of the signal component is well simulated. The slopes of the combinatorial background differ significantly between the two subsamples showing the benefit of splitting them to achieve an improved mass description.

The total number of B^0 signal candidates is $N_{B^0} = 1610 \pm 50$. Due to the two times higher integrated luminosity and the increased production cross-section, which in first order scales with the centre-of-mass energies (8 TeV/7 TeV), one expects 2.3 times more signal candidates in the 2012 subsample than in the 2011 subsample and this can indeed be observed for the fitted yields. Additionally, there are around five times more signal candidates in the final state with two kaons than with three kaons, which also meets the expectations from the branching ratios.

Table 7.4: Results of the floating shape parameters in the mass fit to data.

Parameter		Value
μ_{B^0}	$(\mathrm{MeV}/\mathrm{c}^2)$	5279.26 ± 0.29
R_{B^0}		0.995 ± 0.032
$\mu_{D_s^+D^-}$	$(\mathrm{MeV}/\mathrm{c}^2)$	5218.2 ± 1.1
$\sigma_{1,D_s^+D^-}$	$(\mathrm{MeV}/\mathrm{c}^2)$	19.2 ± 2.7
$\sigma_{2,D_s^+D^-}$	$(\mathrm{MeV}/\mathrm{c}^2)$	14.3 ± 3.1
$\beta_{K^-\pi^+\pi^+K^+\pi^-\pi^-}$	$(1/(\mathrm{MeV}/\mathrm{c}^2))$	-0.0031 ± 0.0005
$\beta_{K^-K^+\pi^+K^+\pi^-\pi^-}$	$(1/(\mathrm{MeV}/\mathrm{c}^2))$	-0.0041 ± 0.0006

7.3 Decay time fit

The conditional PDF describing the reconstructed decay time t' and tag decisions $\vec{d'} = (d_{\mathrm{OS}}, d_{\mathrm{SS}})$, given a per-event decay time resolution $\sigma_{t'}$ and per-event mistag probability estimates $\vec{\eta} = (\eta_{\mathrm{OS}}, \eta_{\mathrm{SS}})$, is

$$P\left(t', \vec{d'} \mid \sigma_{t'}, \vec{\eta}\right) \propto \epsilon(t') \left(\mathcal{P}(t, \vec{d'} \mid \vec{\eta}) \otimes \mathcal{R}(t' - t \mid \sigma_{t'})\right) , \tag{7.3}$$

where

$$\mathcal{P}(t, \vec{d'} \mid \vec{\eta}) \propto \sum_d \mathcal{P}(\vec{d'} \mid d, \vec{\eta})[1 - d\, A_{\mathrm{P}}]\, e^{-t/\tau} \left\{1 - d\, S \sin(\Delta m_d t) + d\, C \cos(\Delta m_d t)\right\} , \tag{7.4}$$

and where t is the true decay time, d is the true production flavour, A_{P} is the production asymmetry, and $\mathcal{P}(\vec{d'}|d, \vec{\eta})$ is a two-dimensional binomial PDF describing the distribution of tagging decisions given $\vec{\eta}$ and d. Normalisation factors are omitted for brevity.

7.3.1 Decay time resolution

The prediction of the DecayTreeFit on the decay time error is used to determine the decay time resolution. Like in the analysis of $B^0 \to J/\psi\, K^0_{\mathrm{S}}$ (see Sec. 6.3) these predictions are calibrated using linear functions with parameters b and c. To account for different sources introducing the decay time resolution an effective model consisting of two Gaussians with per-event widths is used. Besides this common resolution effect the decay time resolution model is also supposed to describe the effect of events matched to the wrong PV, which can cause a large deviation between the true and the reconstructed decay time. The wrong PV component is parametrised with a broad Gaussian distribution using the same mean μ_t as the other two Gaussians and one width parameter σ_{PV}. The complete

parametrisation of the resolution model is given by

$$\begin{aligned}\mathcal{R}(t - t_{\text{true}}|\sigma_t) = \sum_{i=1}^{2} g_i \cdot \frac{1}{\sqrt{2\pi}(c_i + b_i \cdot \sigma_t)} \exp\left(-\frac{(t - t_{\text{true}} - \mu_t)^2}{2(c_i + b_i \cdot \sigma_t)^2}\right) \\ + f_{\text{PV}} \frac{1}{\sqrt{2\pi}\sigma_{\text{PV}}} \exp\left(-\frac{(t - t_{\text{true}} - \mu_t)^2}{2\sigma_{\text{PV}}^2}\right) .\end{aligned} \tag{7.5}$$

The first two Gaussian components have different calibration parameters b_i and c_i and thus different widths. Together with the fraction f_{PV} of the wrong PV component the fractions of the two Gaussian components g_1 and g_2 sum up to unity. The shift of the Gaussian mean μ_t is shared between all components. To extract the parameter values an unbinned maximum likelihood fit to the simulated events where the difference between true and reconstructed decay time is below 0.4 ps is performed (see Fig. 7.8). The results listed in Table 7.5 correspond to a decay-time-resolution related dilution of 0.9996. The decay time resolution might

Table 7.5: Fit parameters of the decay time resolution function determined on $B^0 \to D^+D^-$ signal MC.

Parameter		Value
μ_t	(ps)	-0.00156 ± 0.00023
b_1		1.022 ± 0.031
c_1	(ps)	0.0036 ± 0.0012
b_2		1.24 ± 0.08
c_2	(ps)	0.0127 ± 0.0035
g_2		0.23 ± 0.12
σ_{PV}	(ps)	0.16 ± 0.04
f_{PV}		0.0024 ± 0.0014

differ between signal MC and data. In the analysis of $B_s^0 \to D_s^-\pi^+$ decays [131] it is found to be 1.15 times higher in data than in MC. If the same applies for $B^0 \to D^+D^-$ decays the dilution would be 0.9995. This marginal difference is not expected to influence the determination of the *CP* observables at all. Thus, the previously described per-event decay time resolution model without any corrections is used in the nominal decay time fit.

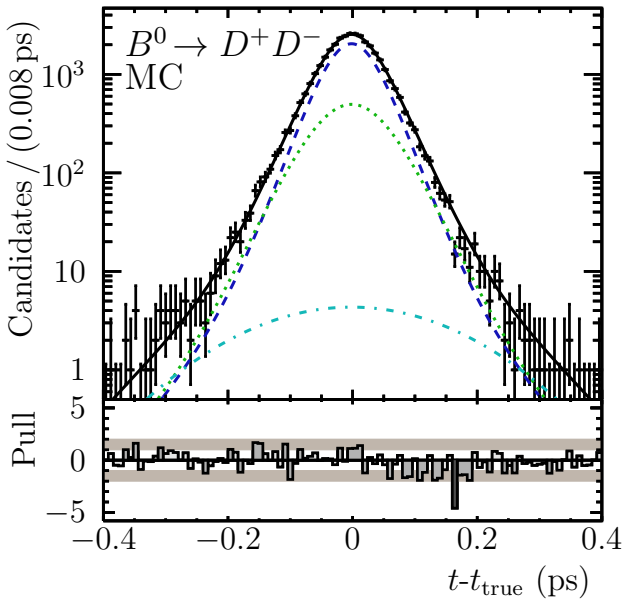

Figure 7.8: Fit of per-event resolution model to the difference of true and reconstructed decay time in signal MC. The black solid line is the projection of the full PDF. The blue dashed and the green dotted line represent the two per-event components and the turquoise dashed-dotted line shows the wrong PV component.

7.3.2 Decay time acceptance

The trigger requirements as well as some input variables to the BDT result in a decay-time-dependent efficiency. Additionally, the VELO reconstruction (*i.e.* the FastVelo algorithm [132]) causes a drop in the decay time acceptance for events with large decay times. In order to correctly describe these effects the B^0 lifetime is constrained to $\tau = (1.519 \pm 0.005)$ ps [124] in the nominal fit and any deviation of the decay time distribution (summed over the tags) from a pure exponential shape is supposed to be described by cubic splines (see Sec. 5.3). Knots are positioned on the rising edge at 0.8 ps, approximately at the turning point at 2 ps, and at the boundaries of the decay time range (0.25 and 10.25 ps). The normalisation of the splines is arbitrary and it has been decided to fix the second to last spline coefficient to 1.0.

On signal MC the truth information is available, so the shape of the decay time acceptance can be separated from the exponential decay. This shape is compared with the spline method described above. As the BDTs are trained and applied separately for the two final states and might have different effects on the shape of the decay time acceptance these two categories are studied individually.

Looking at the plots in Fig. 7.9 it is apparent that there is a quite large efficiency loss at high decay times. This might be related to the fact that both B^0 daughter particles (D^+ and D^-) are relatively long-lived. The true MC decay time acceptance

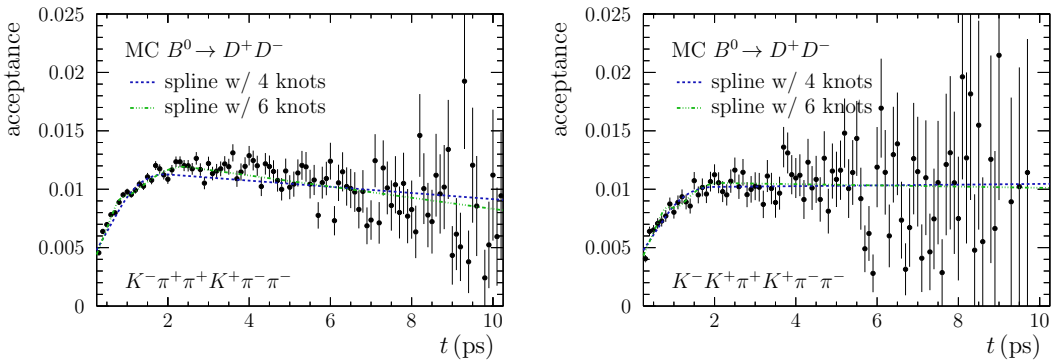

Figure 7.9: Decay time acceptance of truth-matched signal MC for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state (left) and the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state (right). The black data points show the true decay time acceptance determined by dividing the reconstructed by the true decay time distribution. The blue dashed line is the spline acceptance function with four knots and the green dashed-dotted line with six knots.

is overlaid with the shape of two spline functions. Besides the spline function with the nominal number of four knots an additional spline function with two more knots and slightly changed positions $(0.25, 0.7, 1.0, 1.5, 2.5, 10.25\,\text{ps})$ is plotted, which gives a better description. But it has to be considered that the statistics of the MC sample is 25 times larger than the real data. Therefore, the spline function with four knots is chosen, otherwise statistical fluctuations and not acceptance effects would be described. The low statistics of the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state on real data does also not allow to use separate spline coefficients for the two final states, although with the increased MC statistics some differences become visible.

7.3.3 External inputs

LHCb has performed a measurement of the production asymmetry as a function of transverse momentum and pseudorapidity using 7 TeV data [133]. Taking those distributions from $B^0 \to D^+D^-$ individual weighted averages for the 2011 and 2012 subsamples are calculated yielding

$$\begin{aligned} A_{\text{P}}^{11} &= -0.0047 \pm 0.0106\,(\text{stat}) \pm 0.0014\,(\text{syst})\,, \\ A_{\text{P}}^{12} &= -0.0071 \pm 0.0107\,(\text{stat}) \pm 0.0014\,(\text{syst})\,. \end{aligned} \tag{7.6}$$

As the measurement of the production asymmetry has been performed on 2011 data only, the numbers for A_{P}^{11} and A_{P}^{12} are highly correlated. So, the latter is modelled as $A_{\text{P}}^{12} = A_{\text{P}}^{11} + \Delta A_{\text{P}}$ with $\Delta A_{\text{P}} = -0.0024 \pm 0.0018\,(\text{syst})$. The systematic uncertainty accounts for the difference of the production asymmetries observed for the two data-taking conditions in the measurement of the semileptonic CP asymmetry [69] and is used as the width of a Gaussian constraint. The B^0 oscillation frequency

and the B^0 lifetime are constrained to $\Delta m_d = (0.510 \pm 0.003)\,\hbar\,\mathrm{ps}^{-1}$ [124] and $\tau = (1.519 \pm 0.005)\,\mathrm{ps}$ [124], respectively. The flavour-tagging calibration parameters (Table 5.1) are constrained within their combined statistical and systematic uncertainties, determined in the calibration using $B^0 \to D_s^+ D^-$ decays. The decay time resolution parameters (Table 7.5) and the B^0 lifetime difference $\Delta\Gamma = 0\ \mathrm{ps}^{-1}$ are fixed in the likelihood fit.

7.3.4 Results

The fit results of the *CP* observables from the decay time fit are

$$\begin{aligned} S_{D^+D^-} &= -0.54 \pm {}^{0.17}_{0.16}\,, \\ C_{D^+D^-} &= 0.26 \pm 0.17\,, \\ \rho(S_{D^+D^-}, C_{D^+D^-}) &= 0.48\,. \end{aligned} \tag{7.7}$$

Only after rescaling the sWeights via

$${}_s\mathcal{P}_\mathrm{n} = {}_s\mathcal{P}_\mathrm{n} \frac{\sum {}_s\mathcal{P}_\mathrm{n}}{\sum {}_s\mathcal{P}_\mathrm{n}{}^2}\,, \tag{7.8}$$

the correct asymmetric uncertainty estimates are determined from an analysis of the likelihood shape. A plot of the decay time distribution and the projection of the acceptance model are shown in Fig. 7.10. Good agreement between the latter and the shape on signal MC (cf. Fig. 7.9) can be observed but the low statistics leading to rather large uncertainties indicated by the error band diminishes the significance of the comparison.

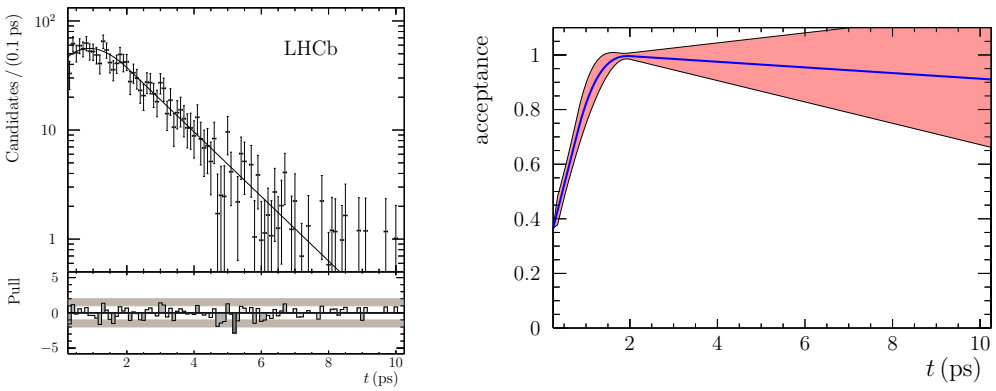

Figure 7.10: Plot of the decay time distribution of the background-subtracted $B^0 \to D^+D^-$ data sample with the projection of the PDF and the pull distribution on the left. The y-axis is plotted in logarithmic scale. Plot of the nominal decay time acceptance model on the right. The red area indicates the $1\,\sigma$ error band taking into account the statistical uncertainties.

To check if the coverage is guaranteed the bootstrapping method is applied. The nominal fit procedure, *i.e.* performing the mass fit, calculating the sWeights and fitting the weighted tagged decay time distribution, is executed and the fit results are stored. The drawing and fitting is done 10 000 times. It turns out that half of the fits fail if the flavour-tagging calibration parameters are constrained within their statistical uncertainties. When fixing them to their central values the fit failure rate drops to a per-mille effect. From the distribution of fit results the two-side 68 % confidence intervals are extracted. To account for the uncertainties on the flavour-tagging calibration parameters 10 000 pseudoexperiments are performed, in which the nominal model is used to generate the signal decay time distribution and the fit results of the nominal fit are chosen for the *CP* observables $S_{D^+D^-}$ and $C_{D^+D^-}$. Before generating the flavour-tagging calibration parameters are drawn from Gaussian distributions around their central values using the combined statistical + systematic uncertainties. In the subsequent fit the flavour-tagging calibration parameters are fixed to their central values, like in the fits to the bootstrapped samples. The resulting pull distributions are broader than standard normal distributions. The deviation of the width from unity shows how much the statistical uncertainties are underestimated in the likelihood fit due to not accounting for the variation of the flavour-tagging calibration parameters. So, the statistical uncertainties for $S_{D^+D^-}$ and $C_{D^+D^-}$ from the bootstrapping including the impact of the uncertainty of the flavour-tagging calibration parameters are given by scaling the bootstrapping uncertainties by the width of the pull distributions:

$$\begin{aligned} \sigma_{S_{D^+D^-}}(\text{bootstrapping}) &= {}^{+0.17}_{-0.16}\,, \\ \sigma_{C_{D^+D^-}}(\text{bootstrapping}) &= {}^{+0.18}_{-0.17}\,. \end{aligned} \tag{7.9}$$

These uncertainties match the nominal ones from the likelihood shape quite well. The one-dimensional likelihood scans in Fig. 7.11 support the validity as they show nice parabolic shapes with clear minima.

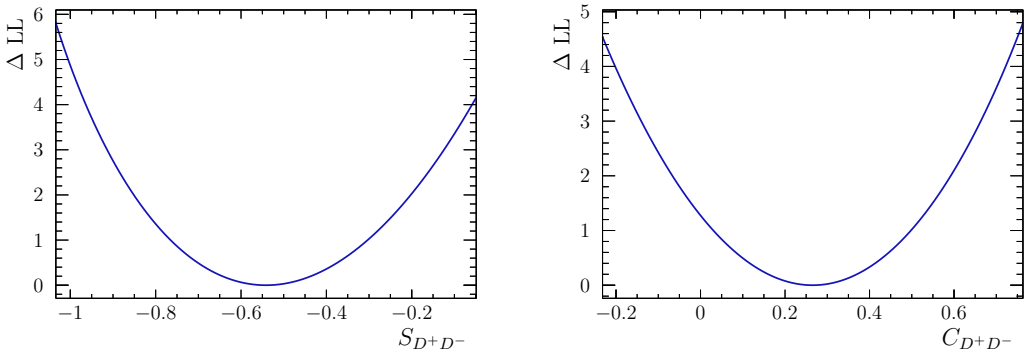

Figure 7.11: One-dimensional likelihood profile scans for $S_{D^+D^-}$ and $C_{D^+D^-}$.

Apart from a quite high positive correlation between the parameters of the acceptance spline function and the already quoted correlation of about 0.5 between

$S_{D^+D^-}$ and $C_{D^+D^-}$, which is expected from first principles (see Ref. [134]), no large correlation between fitted parameters is present, as can be seen from the correlation matrix visualised in Fig. 7.12. A possible correlation between Δm_d and $C_{D^+D^-}$

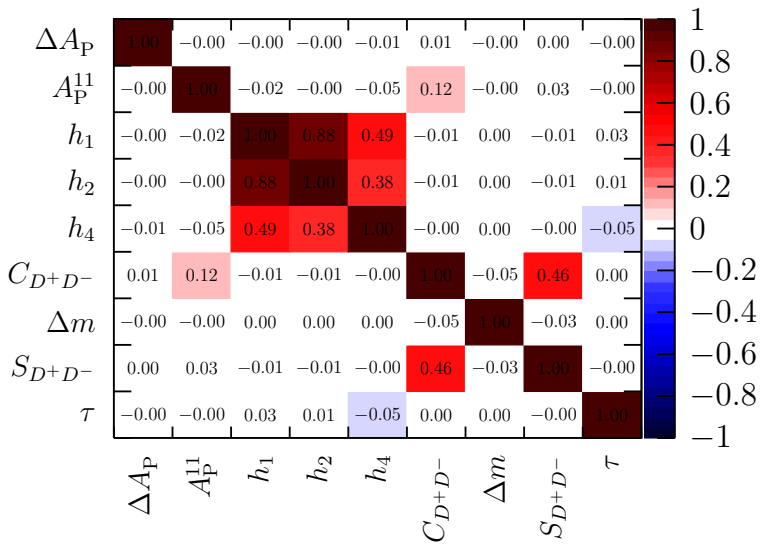

Figure 7.12: Visualised correlation matrix of the fit parameters in the decay time fit to data. Positive correlations are represented by the red palette on the z axis, while negative correlations are represented by the blue palette of the z axis.

is significantly reduced by the constraint applied on Δm_d, which is a lot tighter than the sensitivity accessible from the data sample. When releasing this constraint the correlation coefficient becomes -0.8. But the sensitivity on $C_{D^+D^-}$ would significantly decrease in this scenario, so the constraint on Δm_d is maintained in the nominal setup.

In Fig. 7.13 the signal yield asymmetry is plotted in eight bins of the decay time. A binned χ^2-fit to this signal asymmetry is performed using

$$\mathcal{A}^{\text{meas}}(t) = \frac{\Delta\omega + A_P^{11}(1-2\omega) + (1-2\omega+A_P^{11}\Delta\omega)\mathcal{A}^{\text{theo}}(t)}{1 + A_P^{11}(S_{D^+D^-}\sin(\Delta m_d\, t) - C_{D^+D^-}\cos(\Delta m_d\, t))}\,, \tag{7.10}$$

which is a modified version of the theoretical signal asymmetry in Eq. (3.38) and accounts for the mistag probability ω and the asymmetries induced by flavour tagging ($\Delta\omega$) and production asymmetry (A_P^{11}). The fit results

$$\begin{aligned} S_{D^+D^-} &= -0.65 \pm 0.25\,, \\ C_{D^+D^-} &= 0.24 \pm 0.26\,, \end{aligned}$$

are compatible with those from the unbinned fit presented in Eq. (7.7) but not as sensitive.

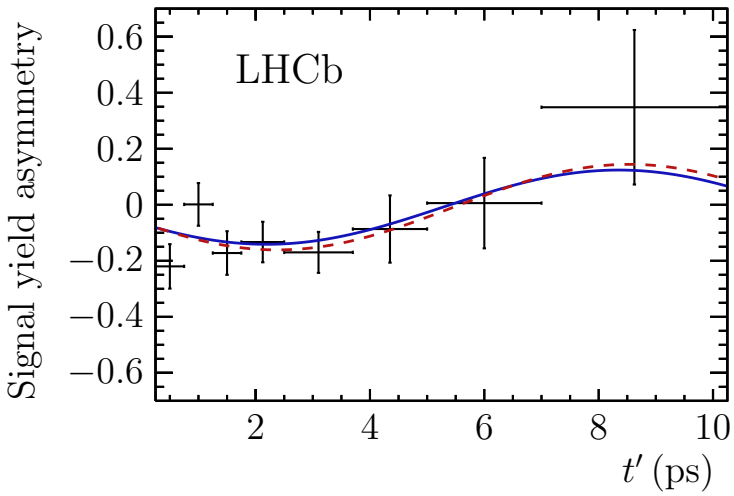

Figure 7.13: Decay-time-dependent signal yield asymmetry. The solid blue curve is the projection of the signal PDF given in Eq. (7.3) and the dashed red curve is the pure decay-time-dependent fit function from Eq. (7.10).

7.4 Studies of systematic effects

7.4.1 Cross-checks

To check for possible systematic effects, fits in different subsamples of the nominal data set are performed. The cross-checks are performed for the two tagging algorithms (OS vs. SS (not exclusive samples)), the two years of data-taking (2011 vs. 2012) , combinations of those, the magnet polarities (Up vs. Down), the two final states ($K^-\pi^+\pi^+K^+\pi^-\pi^-$ vs. $K^-K^+\pi^+K^+\pi^-\pi^-$) and for four different slices of the BDT classifier for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state.

The fit results in the various scenarios are illustrated in Figs. 7.14 and 7.15. While almost all splits show compatible results, a rather large difference can be observed between the 2011 and the 2012 subsample for $S_{D^+D^-}$. This is even more pronounced when using only SS tagging. However, when calculating the difference of the log likelihoods it turns out that the two results are compatible within 2.6 standard deviations. Nevertheless, the flavour-tagging calibration parameters are also determined separately for 2011 and 2012 data. Only small non-significant differences are observed, which can not explain the different results of the *CP* observables. Therefore, the best explanation is that the difference is due to a statistical fluctuation.

For the nominal fit the decay times and the decay time errors from the DecayTreeFitter (DTF) are used. The central values of the *CP* observables slightly change when using the decay time (error) from the LoKiVertexFitter (LVF), $S_{D^+D^-} = -0.539$ (LVF) vs. -0.541 (DTF) and $C_{D^+D^-} = 0.266$ (LVF) vs. 0.263 (DTF). But this difference is clearly below the statistical significance.

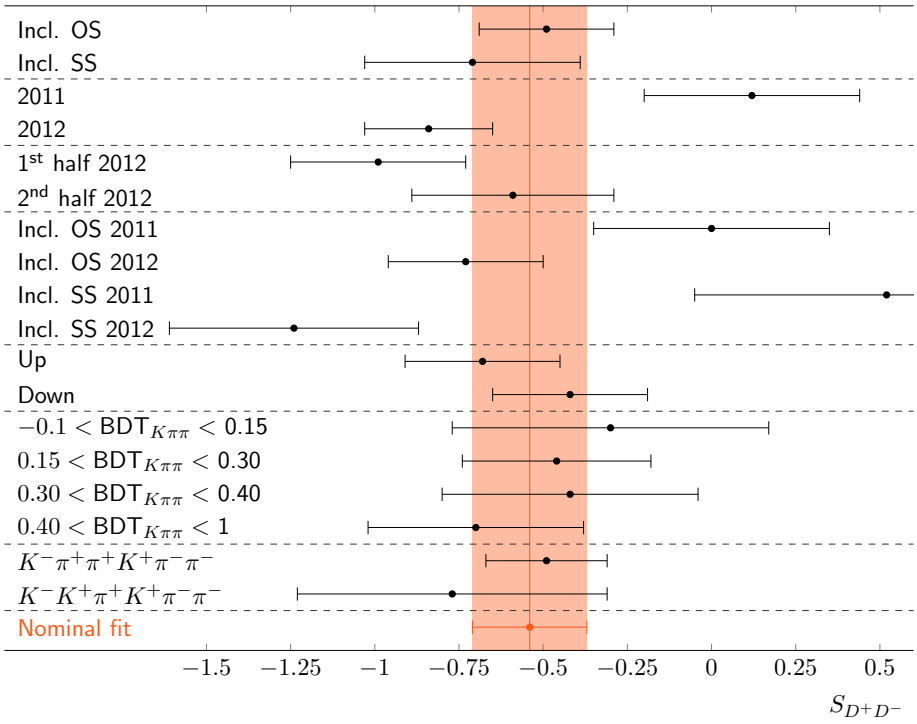

Figure 7.14: Comparison of fit results of $S_{D^+D^-}$ for fits on various subsamples.

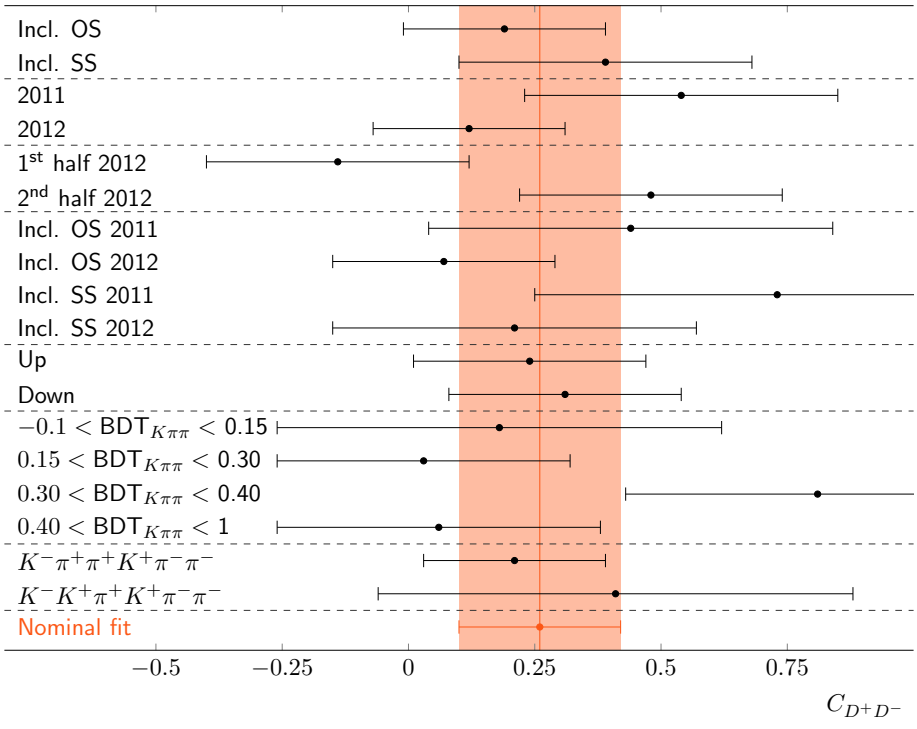

Figure 7.15: Comparison of fit results of $C_{D^+D^-}$ for fits on various subsamples.

7.4.2 Decay time fit bias

The likelihood fit itself might be biased. The nominal fit results for the *CP* observables are used to generate 10 000 pseudoexperiments. The pull distributions in Fig. 7.16 show a very small deviation of the mean value from zero. Multiplying

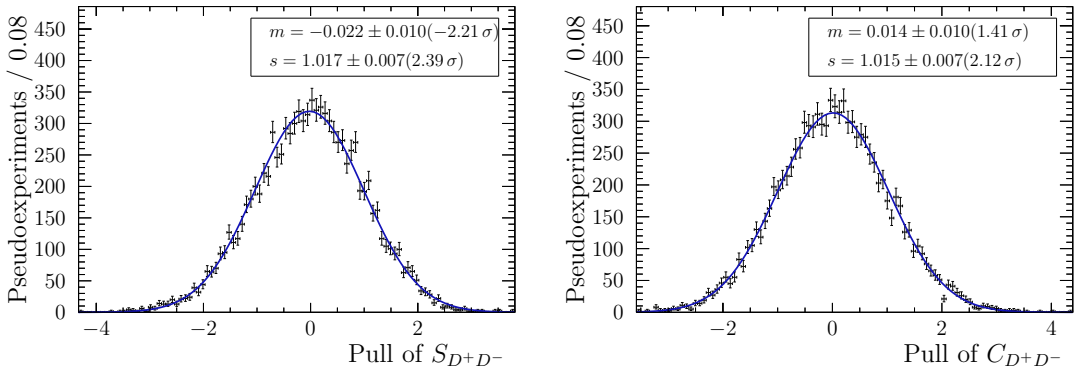

Figure 7.16: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainty due to the likelihood fitter.

it with the statistical uncertainty the systematic uncertainty is calculated to be

$$s^{\text{fit}}_{S_{D^+D^-}} = 0.004\ , \qquad s^{\text{fit}}_{C_{D^+D^-}} = 0.0025\ . \tag{7.11}$$

For all following studies of systematic uncertainties the residuals are corrected for the decay time fit bias. Otherwise, even effects that are actually not biasing would be misinterpreted due to the fit bias.

7.4.3 Fit model

7.4.3.1 Mass Model

Two different aspects of the mass model are studied regarding systematic uncertainties: the impact of neglecting contributions and of mismodelling components.

Neglected contributions If a neutral π^0 or a photon is missed in the reconstruction the decay $B \to D^{*+}D^-$, with $D^{*+} \to D^+\pi^0$ or $D^{*+} \to D^+\gamma$, can mimic the $B \to D^+D^-$ decay. In the rest frame of the D^{*+} resonance, the missing momentum of the π^0 is fixed, but it needs to be boosted when transferred into the rest frame of the B meson. So, the reconstructed mass depends on the helicity angle of the missing π^0. This leads to a double-horned structure approximately 140 MeV/c^2 below the nominal B mass (see Ref. [135] for more details on the shape of this background). As the lower boundary on the invariant $m_{D^+D^-}$ mass is set to 5150 MeV/c^2 the $B^0 \to D^{*+}D^-$ contribution lies outside the mass range used for the fit. However, the $B^0_s \to D^{*+}D^-$ contribution enters the fit region. But since the expected number

of $B_s^0 \to D^{*+}D^-$ candidates is low, it is not included in the nominal mass model. Another contribution that is neglected in the nominal mass fit model is (partially) charmless background where at least one of the hadron triplets is not originating from a D decay. The systematic uncertainty on the determination of the CP observables arising from neglecting these two contributions is estimated using 1000 pseudoexperiments. Components for $B_s^0 \to D^{*+}D^-$ and for (partially) charmless background are included in the generation but excluded from the fit procedure.

The shape of $B_s^0 \to D^{*+}D^-$ is parametrised with two single Gaussian functions centred around $5150\,\mathrm{MeV/c^2}$ and $5200\,\mathrm{MeV/c^2}$. The (partially) charmless background is modelled with a single Gaussian function. When optimising the decay time significance cut it has been observed that the width of the (partially) charmless background is approximately 10 % wider than the signal component. Therefore, a width of $10\,\mathrm{MeV/c^2}$ is chosen. The mean is set to the same position as the B^0 signal. The $B_s^0 \to D^{*+}D^-$ component is generated without any tagging asymmetry, while for the (partially) charmless background the worst case scenario of maximal CP violation with the opposite CP eigenvalue ($S_f = +1.0$) is tested.

In studies of $B^0 \to D^{*+}D^-$ decays [136] a significant contribution of $B_s^0 \to D^{*+}D^-$ candidates is observed. The ratio between the two yields is determined to be 1:20. Under the assumption that the efficiencies for $B \to D^+D^-$ and $B \to D^{*+}D^-$ are the same the expected number of $B_s^0 \to D^{*+}D^-$ candidates can be calculated via

$$\mathrm{N}_{B_s^0 \to D^{*+}D^-} = \frac{1}{20}\mathrm{N}_{B^0 \to D^+D^-} \frac{\mathcal{B}(B^0 \to D^{*+}D^-)\mathcal{B}(D^{*+} \to D^+(\pi^0||\gamma))}{\mathcal{B}(B^0 \to D^+D^-)}\,. \tag{7.12}$$

Using the world averages for the branching ratios [22] the number of candidates to be generated in the pseudoexperiments is estimated to be $\mathrm{N}(B_s^0 \to D^{*+}D^-) = 66 \pm 9$.

To determine how many (partially) charmless background candidates need to be generated the D mass window is widened to $\pm 40\,\mathrm{MeV/c^2}$ and the nominal D mass window of $\pm 25\,\mathrm{MeV/c^2}$ is vetoed for one or for both D candidates. Fits to the invariant B mass without the D mass constraint are performed in the various scenarios. The fitted yields, which are constrained to positive values, are scaled to account for the applied D mass window. The total amount of residual contamination ($B^0 \to Dhhh$ or $B^0 \to hhhhhh$ decays) surviving the $B^0 \to D^+D^-$ selection is found to be 28.7 ± 19.5 candidates for the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state and 0.0 ± 27.8 candidates for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state. For the pseudoexperiments the number of (partially) charmless background is drawn from Gaussian distributions using these values for mean and width. When the outcome is negative the procedure is repeated, until a positive yield is drawn.

The systematic uncertainties on $S_{D^+D^-}$ and $C_{D^+D^-}$ are calculated as the product of the bias on the mean parameter of the pull distributions and the statistical uncertainty:

$$s^{\mathrm{mass,1}}_{S_{D^+D^-}} = 0.05\,, \qquad s^{\mathrm{mass,1}}_{C_{D^+D^-}} = 0.013\,.$$

Mismodelling of mass components The BDT is trained with MC samples that are known to not perfectly model the PID information. As a result the BDT classifier distributions of background-subtracted data and MC show a quite big discrepancy. Some shape parameters are estimated on MC samples and might be distorted by the data/MC differences. Therefore, different alternative mass parametrisations are tested against the nominal model: the component of the $B^0 \to D^+D^-$ signal (and of the $B_s^0 \to D^+D^-$ background) is parametrised with a single Gaussian function; the combinatorial background is described with a second order Chebyshev polynomial of first kind; the tail parameters of $B \to D_s D$ are once extracted from the MC sample without applying the BDT and once applying a tight cut on the BDT classifier. The mass fit is performed with these new models, sWeights are calculated for each approach, and the decay time fit is performed. The results of the *CP* observables are then compared with the nominal central values. The largest deviations for $S_{D^+D^-}$ and $C_{D^+D^-}$ are

$$s^{\text{mass,2}}_{S_{D^+D^-}} = 0.004\,, \qquad s^{\text{mass,2}}_{C_{D^+D^-}} = 0.006\,.$$

7.4.3.2 Correlation between decay time and mistags

The correlation between the decay time distribution and the per-event mistags is studied by calculating the linear Pearson correlation coefficient $\rho(\eta, t)$. The significance of the correlation value, *i.e.* 95 % confidence level interval, is determined using the bootstrapping method (Sec. 5.4) with 10 000 repetitions. The correlation coefficients are found to be small. The profile histogram of the OS tagging combination, which shows the average η_{OS} value as a function of the decay time, is flat within statistics. For the SS tagging combination the profile histogram slowly increases with decay time. This can be confirmed by analysing the larger signal MC sample (see Fig. 7.17). Performing a χ^2 fit in the decay time range 0.25–8.25 ps with the linear function

$$\eta_{\text{SS}} = a_{\eta_{\text{SS}},t}\, t + b_{\eta_{\text{SS}},t} \tag{7.13}$$

yields a slope of $a_{\eta_{\text{SS}},t} = (1.50 \pm 0.27)$ ns^{-1}. Although this is a significant deviation from zero, the correlation is not taken into account in the nominal fit model. Instead, a study on the systematic uncertainty from neglecting this effect is performed. In 1000 pseudoexperiments the SS mistag is generated using a Gaussian distribution whose mean is drawn from the linear function defined in Eq. (7.13) thereby introducing the correlation with the decay time. In the subsequent fit the correlation is again ignored. This leads to systematic uncertainties of

$$s^{\text{corr}}_{S_{D^+D^-}} = 0.0007\,, \qquad s^{\text{corr}}_{C_{D^+D^-}} = 0.007\,.$$

7.4.3.3 Decay Time Resolution Model

As calculated in Sec. 7.3.1 even an underestimation of the decay time resolution by 15 % has only a minor effect on the resolution related dilution. Nevertheless,

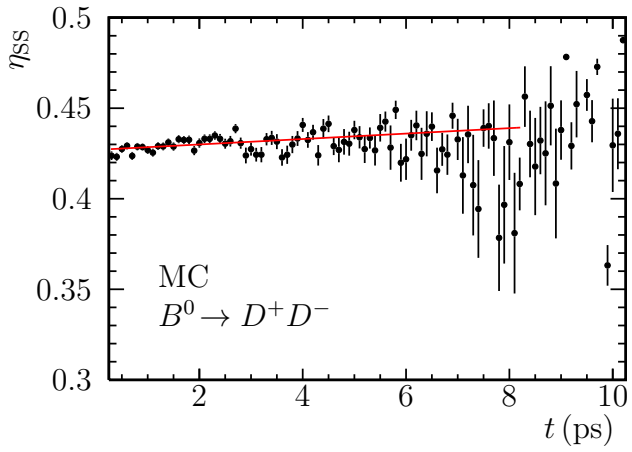

Figure 7.17: Profile histogram for the decay time dependence on η_{SS} for signal MC. The black data points represent the mean value of η_{SS} and its uncertainty for each bin in t. The red curve is the fitted linear function.

1000 pseudoexperiments are performed, in which the scale factors and the offset parameters (b_i and c_i from Table 7.5) are enlarged by 15 % in the generation and fixed to their nominal values in the fit. Additionally, the mean parameter of the Gaussians is set to the value obtained in the MC study for the generation and, like in the nominal setup, fixed to zero in the fit. The systematic uncertainties are calculated as the product of the biases on the mean parameter and the statistical uncertainty to be

$$s^{\text{res}}_{S_{D^+D^-}} = 0.0020\,, \qquad s^{\text{res}}_{C_{D^+D^-}} = 0.0023\,.$$

7.4.3.4 Decay Time Acceptance Model

On signal MC the decay time acceptance is determined separately for the two final states (see Fig. 7.9). Small differences are observed. As the low statistics in the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state on data does not allow for an individual spline model, a study is performed to estimate a possible systematic uncertainty from neglecting this difference. In 1000 pseudoexperiments the decay time distribution is generated using the histograms of the true decay time acceptance from signal MC, split by final state, and fitted with the spline acceptance as done in the nominal fit. The use of the histograms with 100 bins should also cover uncertainties from the choice of the number and position of the knots. The pull between the fit results and the generation values is calculated. The systematic uncertainty due to the decay time acceptance model is calculated as the product of the shift in the pull distribution and the statistical uncertainty of the nominal fit:

$$s^{\text{acc}}_{S_{D^+D^-}} = 0.007\,, \qquad s^{\text{acc}}_{C_{D^+D^-}} = 0.0027\,.$$

7.4.4 Further studies

7.4.4.1 z-scale

The decay times are determined by measuring the distance between PV and decay vertex. So, any uncertainty on the positioning of detector elements (especially the VELO modules) leads to biased decay times. Due to the high boosting the main contribution to the flight distance is in z direction. The scale uncertainty in z direction has been estimated to be $\sigma_{z\text{-scale}} = 0.022\,\%$ [137]. The influence on the measurement of the *CP* observables is studied by performing 1000 pseudoexperiments. For each pseudoexperiment a value for the uncertainty on the z-scale is drawn from a Gaussian distribution around zero of width $\sigma_{z\text{-scale}}$. The sum of 50 fs and the product of this value with the decay time is used as width of the Gaussian function modelling the decay time resolution in the generation. In the fit the width is set to 50 fs. The product of the bias from the pull distributions of the pseudoexperiments and the nominal statistical uncertainty is taken as systematic uncertainty:

$$s^{z\text{-scale}}_{S_{D^+D^-}} = 0.0031\,, \qquad s^{z\text{-scale}}_{C_{D^+D^-}} = 0.0028\,.$$

7.4.4.2 Production Asymmetry

The systematic uncertainty on the production asymmetry A^{11}_{P} is studied using 1000 pseudoexperiments. The nominal value is used in the generation and the procedure described in Ref. [138] is applied in the fit: Before fitting the data sample the mean of the Gaussian constraint for A^{11}_{P} is shifted by one systematic uncertainty. The resulting Gaussian distribution is used to draw a new value for the mean. Then, the new Gaussian distribution is used to constrain A^{11}_{P} in the fit. Both shifts, upwards and downwards, are tested and the larger deviation is taken as systematic uncertainty:

$$s^{A_{\mathrm{P}}}_{S_{D^+D^-}} = 0.0015\,, \qquad s^{A_{\mathrm{P}}}_{C_{D^+D^-}} = 0.004\,.$$

For the production asymmetry difference ΔA_{P} the systematic uncertainty is already included in the Gaussian constraint of the nominal fit.

7.4.4.3 Decay Width Difference $\Delta\Gamma_d$

The decay width difference $\Delta\Gamma_d$ is expected to be very small and therefore fixed to zero in the nominal fit. But experimentally it has a relatively large uncertainty. This is taken into account by performing 1000 pseudoexperiments where the current statistical precision $\sigma(\Delta\Gamma_d) = \pm 0.007\ \mathrm{ps}^{-1}$ [59] is used in the generation of the data samples while it is, like in the nominal model, neglected in the fit. The mean parameters of the pull distributions are converted into systematic uncertainties of

$$s^{\Delta\Gamma_d}_{S_{D^+D^-}} = 0.014\,, \qquad s^{\Delta\Gamma_d}_{C_{D^+D^-}} = 0.0021\,.$$

7.4.4.4 B^0 Mass Difference Δm_d

The systematic uncertainty on the world average of Δm_d ($\pm 0.002\,\hbar\,\mathrm{ps}^{-1}$ [59]) is not covered by the Gaussian constraint that is used in the nominal fit. Instead, it is analysed using 1000 pseudoexperiments. In the generation the nominal model is used. Before performing the fit the mean of the Gaussian distribution (its width is the statistical precision of the world average) is shifted by one systematic uncertainty (once up and once down) and a new value is drawn from the distribution. This new constraint is then used in the minimisation. Looking at the resulting pull distributions systematic uncertainties of

$$s^{\Delta m_d}_{S_{D^+D^-}} = 0.0025\,, \qquad s^{\Delta m_d}_{C_{D^+D^-}} = 0.006\,,$$

are assigned.

7.4.5 Total systematic uncertainty

The systematic uncertainties are summarised in Table 7.6. The full systematic uncertainty is calculated by summing the individual uncertainties in quadrature.

Table 7.6: Systematic uncertainties on the *CP* observables $S_{D^+D^-}$ and $C_{D^+D^-}$.

Origin	$\sigma_{S_{D^+D^-}}$	$\sigma_{C_{D^+D^-}}$
Neglecting components in mass model	0.05	0.013
$\Delta\Gamma_d$	0.014	0.0021
Decay time acceptance	0.007	0.0027
Correlation between mass and decay time	0.0007	0.007
Parametrisation of PDFs in mass model	0.004	0.006
Δm_d	0.0025	0.006
Fit bias	0.004	0.0025
Production asymmetry	0.0015	0.004
z-scale	0.0031	0.0028
Decay time resolution	0.0020	0.0023
Sum	0.05	0.018

8 Discussion

8.1 Comparison with previous measurements of $\sin 2\beta_{(\mathrm{eff})}$

Measurements of CP violation in $B^0 \to J/\psi K^0_S$ decays have been performed since the end of the nineties. The first results by OPAL [139], ALEPH [140] and CDF [141] had an uncertainty on $\sin 2\beta$ no better than ± 0.4. One of the main purposes of the B-factories was to improve the precision, which succeeded with results of $S_{J/\psi K^0_S} = 0.657 \pm 0.036 \pm 0.012$ by BaBar [142] and $S_{J/\psi K^0_S} = 0.670 \pm 0.029 \pm 0.013$ by Belle [143]. Averaging the results of the B-factories and combining them with measurements in various other charmonium modes leads to an average of $\sin 2\beta = 0.679 \pm 0.020$ [59], corresponding to $\beta = (21.4 \pm 0.8)^\circ$. The measurement presented in this thesis, using data corresponding to an integrated luminosity of 3 fb^{-1} collected with the LHCb experiment, further improves on this, yielding an updated world average of $\sin 2\beta = 0.691 \pm 0.017$ ($\beta = (21.9 \pm 0.7)^\circ$) [59]. The angle β can also be constrained from a global fit to the CKM triangle. When not using the inputs from the direct measurements described above, a value of $\beta = (24.3^{+1.3}_{-1.4})^\circ$ [60] is found. This shows that the value found by LHCb improves the compatibility between the direct and the indirect measurements by shifting $\sin 2\beta$ slightly upwards.

While for $B^0 \to J/\psi K^0_S$ decays mainly the value of $S_{J/\psi K^0_S}$ is interesting, as it is expected to correspond to $\sin 2\beta$ with only small corrections, the same does not apply for the measurement of CP violation using $B^0 \to D^+D^-$ decays. Both observables, $S_{D^+D^-}$ and $C_{D^+D^-}$, have to be considered simultaneously for a proper interpretation of the results. In Fig. 8.1 the latest results from BaBar, Belle, and LHCb, as well as the average of these three are plotted in the two-dimensional plane of $C_{D^+D^-}$ versus $S_{D^+D^-}$. Looking at the uncertainty ellipses it is apparent that the precision of LHCb matches the one of Belle, while being significantly better than the one of BaBar. The orientation of the ellipses shows that in the measurements of the B-factories the two CP observables are determined almost uncorrelated. For a comparison of the central values it is useful to take into account the arc defined by the condition $S^2_{D^+D^-} + C^2_{D^+D^-} = 1$. It represents the extreme case of $\lambda_{D^+D^-}$ being purely imaginary and delimits the physically allowed region. The result by Belle [144] of $S_{D^+D^-} = -1.06$ and $C_{D^+D^-} = -0.43$ lies outside, while the results of BaBar [145], which are $S_{D^+D^-} = -0.63$ and $C_{D^+D^-} = -0.07$, as well as the results of this thesis of $S_{D^+D^-} = -0.54$ and $C_{D^+D^-} = 0.26$ are inside.

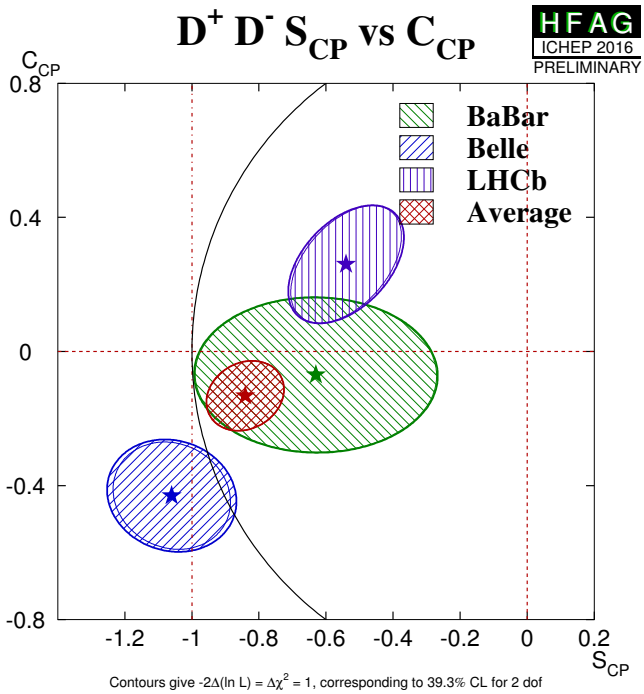

Figure 8.1: Comparison of *CP* observables from $B^0 \to D^+D^-$ decays in the ($S_{D^+D^-}$, $C_{D^+D^-}$) plane [59]. The contours give $-2\Delta(\ln L) = \Delta\chi^2 = 1$, thereby corresponding to a coverage of 39.3 % for 2 degrees of freedom. The marked arc is defined by the condition $S^2_{D^+D^-} + C^2_{D^+D^-} = 1$.

8.2 Comparison between *CP* violation in $B^0 \to J/\psi\, K^0_S$ and in $B^0 \to D^+D^-$ decays

Measurements in decays with $b \to c\bar{c}s$ transitions have established the presence of *CP* violation in the system of neutral B mesons with very high precision. Comparing the plots of the decay-time-dependent signal yield asymmetry for $B^0 \to J/\psi\, K^0_S$ and $B^0 \to D^+D^-$ decays (see Fig. 8.2), the difference in sensitivity becomes apparent. However, the main motivation for the measurement of *CP* violation in $B^0 \to D^+D^-$ decays, presented in this thesis, is not to achieve the single best measurement of *CP* violation. In fact, given the uncertainties the conservation of *CP* symmetry can be excluded by only 4.0 standard deviations with this measurement. Instead, the phase shift $\Delta\phi_d$ can be obtained, which is caused by higher-order Standard Model contributions and complicates the determination of β and β_s similarly (see Sec. 3.5). The fit results of $S_{D^+D^-}$ and $C_{D^+D^-}$ (see Eq. (7.7)) correspond to

$$\sin(\phi_d + \Delta\phi_d) = \frac{-S_{D^+D^-}}{\sqrt{1 - C^2_{D^+D^-}}} = 0.56 \pm {}^{0.16}_{0.17}\,,$$

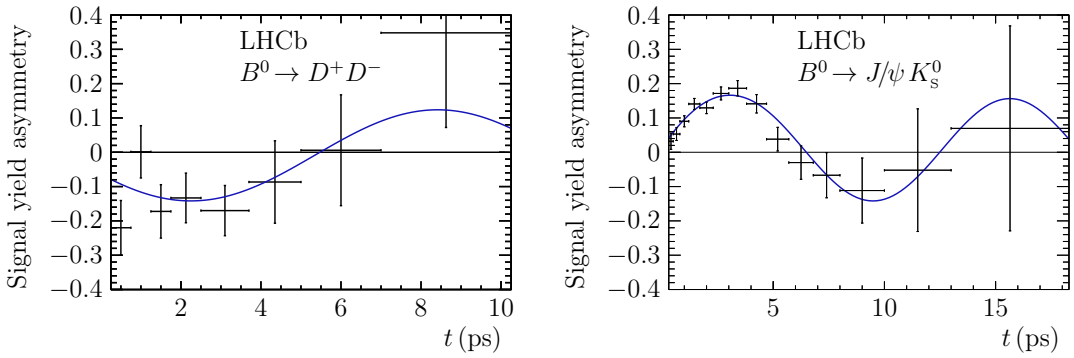

Figure 8.2: Signal yield asymmetry from $B^0 \to D^+D^-$ decays [129] (left) and $B^0 \to J/\psi K^0_S$ decays [125] (right).

where the statistical uncertainty is estimated by generating three million sets of $S_{D^+D^-}$ and $C_{D^+D^-}$ using a two-dimensional Gaussian distribution including their correlation, calculating $\sin(\phi_d + \Delta\phi_d)$ for each of them, and then taking the two-sided 68 % confidence intervals. This result can be combined with the corresponding value from $B^0 \to J/\psi K^0_S$ to extract the phase shift $\Delta\phi_d$. The calculation yields $\Delta\phi_d = -0.16\,^{+0.19}_{-0.21}$ rad, which is significantly more precise than the previous measurement using the results of the B-factories of $\Delta\phi_d = (30\,^{+23}_{-30})^\circ = 0.52\,^{+0.40}_{-0.56}$ rad [73], and thus represents the world's most precise determination of this quantity. Moreover, the size of the higher-order corrections is reduced by more than a factor of three and is compatible with zero.

9 Conclusion

Since the discovery of *CP* violation in 1964 by Christenson, Cronin, Fitch and Turlay [6] there have been many experiments searching for *CP* violation, first in the sector of neutral K mesons but later on also in neutral B mesons. The most significant indication of *CP* violation for B^0 mesons is found by the determination of $\sin 2\beta$ using $B^0 \to J/\psi\, K^0_S$ decays. But although LHCb is not the first experiment, even not at a hadron collider, to measure *CP* violation, it plays an important role in the further exploration of the quark-mixing sector.

The measurement of *CP* violation in $B^0 \to J/\psi\, K^0_S$ decays [125] using proton-proton collision data corresponding to an integrated luminosity of 3 fb^{-1}, which is presented in this thesis, yields

$$\begin{aligned} S_{J/\psi\, K^0_S} &= 0.731 \pm 0.035\,(\text{stat}) \pm 0.020\,(\text{syst})\,, \\ C_{J/\psi\, K^0_S} &= -0.038 \pm 0.032\,(\text{stat}) \pm 0.005\,(\text{syst})\,, \end{aligned}$$

which is the most precise determination of these *CP* observables at a hadron collider to date and is almost as precise as the previous measurements by BaBar [142] and Belle [143]. The central values are compatible with the world averages and with the Standard Model expectations. Thus, it serves as a benchmark measurement showing the capability of LHCb to perform flavour-tagged precision measurements of *CP* violation. The experimental difficulties, *e.g.* decay time resolution, production asymmetries and asymmetries induced by the flavour tagging, are well under control as the result is statistically limited. The largest systematic uncertainty on $S_{J/\psi\, K^0_S}$ is introduced by a possible tagging asymmetry of the background, which is not accounted for in the likelihood fit. With an increased statistics this effect can probably be analysed, controlled and suppressed better. Furthermore, in the meantime new same-side flavour-tagging algorithms have been developed [106], which are used for the first time in the measurement of *CP* violation in $B^0 \to D^+D^-$ decays [129] yielding

$$\begin{aligned} S_{D^+D^-} &= -0.54\,^{+0.17}_{-0.16}\,(\text{stat}) \pm 0.05\,(\text{syst})\,, \\ C_{D^+D^-} &= 0.26\,^{+0.18}_{-0.17}\,(\text{stat}) \pm 0.02\,(\text{syst})\,. \end{aligned}$$

If the flavour-tagging performance was the same as in the $B^0 \to J/\psi\, K^0_S$ analysis, the 70 times lower number of available signal candidates (114 000 $B^0 \to J/\psi\, K^0_S$ decays vs. 1610 $B^0 \to D^+D^-$ decays) would only allow a sensitivity of ± 0.29 and ± 0.27 for $S_{D^+D^-}$ and $C_{D^+D^-}$, respectively. However, the kinematic properties of the selected $B^0 \to D^+D^-$ candidates lead to a significantly higher tagging power and

on top of that the usage of the improved flavour-tagging algorithms increases the tagging power by another 20 %. The latter improvement can probably be exploited in future measurements of $\sin 2\beta$ with $B^0 \to J/\psi K^0_S$ decays. The value of $\varepsilon_{\text{eff}} = 8.1\,\%$ for the $B^0 \to D^+D^-$ sample is the highest tagging power to date in a tagged *CP* violation measurement at LHCb.

The main achievement of the measurement of *CP* violation in $B^0 \to D^+D^-$ decays is to constrain the contribution of higher-order Standard Model corrections to be small. The result of

$$\Delta\phi_d = -0.16\,^{+0.19}_{-0.21}\,\text{rad}$$

can be transferred to the measurement of *CP* violation in $B^0_s \to D^+_s D^-_s$ decays [146], where ϕ_s, the mixing phase of the B^0_s meson sector, can be determined, but only in a sum with a phase shift $\Delta\phi_s$, which shows a very similar structure as $\Delta\phi_d$ (see Sec. 3.5).

The analysis of $B^0 \to D^+D^-$ decays is only the starting point for similar measurements in other $B^0 \to D^{(*)+}_{(s)} D^{(*)-}$ decay modes. First studies using $B^0 \to D^{*+}D^-$ decays have already been performed [136]. Recent calculations taking into account the flavour-tagging performance seen in $B^0 \to D^+D^-$ and the increase in statistics when adding data from Run II indicate that the sensitivity of BaBar [145] and Belle [144] could be reached and even be topped.

Up to now no significant deviations from the Standard Model expectations have been observed, neither in direct nor in indirect searches. It is obvious that only a tiny effect corresponding to a quite high energy scale is possible. But the question is not really if new physics is needed but rather how it looks like. No dark matter candidate has been found so far. The origin of the baryonic universe with the absence of antimatter can not be explained by the amount of *CP* violation incorporated in the SM. Explanations can only be found with more data. Currently, Run II of the LHC is ongoing. Herein, data is collected at an increased centre-of-mass energy of 13 TeV, which leads to a higher $b\bar{b}$ cross section with respect to the proton-proton collisions recorded during Run I, as the cross section roughly scales linearly with the centre-of-mass energy. The same applies for direct searches at ATLAS and CMS, which can be extended to higher masses. In principle, for LHCb the instantaneous luminosity could also be raised to further increase the data samples, though it is already higher than the design value originally planned in the proposal for the detector [147]. However, a key to significant improvements, especially for decays with hadronic final states but also for measurements of charmonium, is the performance of the trigger system. After the upgrade in 2018–2020 it is planned to read out the full detector at 30 MHz [148]. Right now, the signal efficiency of the hardware trigger is not higher than 50 %. Therefore, a large potential for improvements exists. Another important aspect of tagged *CP* violation measurements is the performance of the flavour-tagging algorithms. The higher the occupancy in the detector the more difficult it is to find the appropriate tagging particle. There are ideas, at least for Run II, how to accommodate for this and the higher centre-of-mass energy helps in

regaining the flavour-tagging performance of Run I. Additionally, in the sector of B^0 mesons the restart of Belle II, planned for 2017, with an improved detector should increase the sensitivity of many measurements significantly. Only the combination of all these efforts to increase the available amount of data to be analysed might result in the observation of deviations from the Standard Model expectations and thus a hint for new physics.

Bibliography

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

[2] G. Gamow, *Expanding universe and the origin of elements*, Phys. Rev. **70** (1946) 572.

[3] A. D. Sakharov, *Violation of CP invariance, c asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32, [Usp. Fiz. Nauk161,61(1991)].

[4] Super-Kamiokande collaboration, K. Abe *et al.*, *Search for proton decay via* $p \to \nu K^+$ *using* 260 kiloton $\cdot$ year *data of Super-Kamiokande*, Phys. Rev. **D90** (2014) 072005, arXiv:1408.1195.

[5] R. A. McKee and J. P. Stark, *Violation of thermal equilibrium in a thermomigration experiment*, Phys. Rev. B **11** (1975) 1374.

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

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

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

[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.*, *Measurement of the rate of* $\nu_e + d \to p + p + e^-$ *interactions produced by* ^8B *solar neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001) 071301, arXiv:nucl-ex/0106015; *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] Nobel Media, *The nobel prize in physics 2015*, Nobelprize.org, Oct, 2016.

[12] I. C. Brock and T. Schorner-Sadenius, eds., *Physics at the Terascale*, 2011.

[13] D. Griffiths, *Introduction to elementary particles*, 2008.

[14] D. H. Perkins, *Introduction to high energy physics*, 1982.

[15] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D **10** (1974) 2445.

[16] L. B. Okun, *The theory of weak interaction*, in *High-energy physics. Proceedings, 11th International Conference, ICHEP'62, Geneva, Switzerland, Jul 4-11, 1962*, p. 845, 1962.

[17] LHCb collaboration, R. Aaij *et al.*, *Observation of exotic $J/\psi\phi$ structures from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, arXiv:1606.07895, submitted to Phys. Rev. Lett.; LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, arXiv:1606.07898, submitted to PRD.

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

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

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

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

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

[23] S. Tomonaga, *On a relativistically invariant formulation of the quantum theory of wave fields*, Progress of Theoretical Physics **1** (1946) 27.

[24] J. Schwinger, *On quantum-electrodynamics and the magnetic moment of the electron*, Phys. Rev. **73** (1948) 416; J. Schwinger, *Quantum electrodynamics. I. A covariant formulation*, Phys. Rev. **74** (1948) 1439.

[25] R. P. Feynman, *Space-time approach to quantum electrodynamics*, Phys. Rev. **76** (1949) 769; R. P. Feynman, *The theory of positrons*, Phys. Rev. **76** (1949) 749; R. P. Feynman, *Mathematical formulation of the quantum theory of electromagnetic interaction*, Phys. Rev. **80** (1950) 440.

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

[27] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.

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

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

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

[31] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343.

[32] H. D. Politzer, *Reliable perturbative results for strong interactions?*, Phys. Rev. Lett. **30** (1973) 1346.

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

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

[35] V. A. Rubakov and M. E. Shaposhnikov, *Electroweak baryon number nonconservation in the early universe and in high-energy collisions*, Usp. Fiz. Nauk **166** (1996) 493, arXiv:hep-ph/9603208, [Phys. Usp.39,461(1996)].

[36] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using* $B^+ \to K^+\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.

[37] H. J. Lipkin, *Penguins, trees and final state interactions in B decays in broken SU(3)*, Phys. Lett. **B415** (1997) 186, arXiv:hep-ph/9710342.

[38] M. Gronau, *U spin symmetry in charmless B decays*, Phys. Lett. **B492** (2000) 297, arXiv:hep-ph/0008292.

[39] J. S. Schwinger, *The theory of quantized fields. 1.*, Phys. Rev. **82** (1951) 914.

[40] G. Lüders, *On the equivalence of invariance under time reversal and under particle-antiparticle conjugation for relativistic field theories*, Kong. Dan. Vid. Sel. Mat. Fys. Med. **28N5** (1954) 1.

[41] W. Pauli, *Niels Bohr and the Development of Physics: Essays Dedicated to Niels Bohr on the Occasion of His Seventieth Birthday*, McGraw-Hill, 1955.

[42] H. Georgi and S. L. Glashow, *Unity of all elementary particle forces*, Phys. Rev. Lett. **32** (1974) 438.

[43] ATLAS collaboration, CMS collaboration, G. Aad *et al.*, *Combined measurement of the Higgs boson mass in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments*, Phys. Rev. Lett. **114** (2015) 191803, arXiv:1503.07589.

[44] J.-L. Gervais and B. Sakita, *Field theory interpretation of supergauges in dual models*, Nucl. Phys. **B34** (1971) 632.

[45] Yu. A. Golfand and E. P. Likhtman, *Extension of the algebra of poincare group generators and violation of p invariance*, JETP Lett. **13** (1971) 323, [Pisma Zh. Eksp. Teor. Fiz.13,452(1971)].

[46] D. V. Volkov and V. P. Akulov, *Possible universal neutrino interaction*, JETP Lett. **16** (1972) 438, [Pisma Zh. Eksp. Teor. Fiz.16,621(1972)]; *Is the Neutrino a Goldstone Particle?*, Phys. Lett. **B46** (1973) 109.

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

[48] BaBar, D. Boutigny *et al.*, *The BABAR physics book: Physics at an asymmetric B factory*, in *Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting) Pasadena, California, September 22-24, 1997*, 1998.

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

[50] C. Jarlskog, *Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP violation*, Phys. Rev. Lett. **55** (1985) 1039, Erratum ibid. **58** (1987) 1698; *A basis independent formulation of the connection between quark mass matrices, CP violation and experiment*, Z. Phys. **C29** (1985) 491.

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

[52] BaBar collaboration, J. P. Lees *et al.*, *Measurement of CP asymmetries and branching fractions in charmless two-body B-meson decays to pions and kaons*, Phys. Rev. **D87** (2013) 052009, arXiv:1206.3525.

[53] Belle collaboration, I. Adachi *et al.*, *Measurement of the CP violation parameters in $B^0 \to \pi^+\pi^-$ decays*, Phys. Rev. **D88** (2013) 092003, arXiv:1302.0551.

[54] LHCb collaboration, R. Aaij *et al.*, *First measurement of time-dependent CP violation in $B_s^0 \to K^+K^-$ decays*, JHEP **10** (2013) 183, arXiv:1308.1428.

[55] LHCb collaboration, *Measurement of the CKM angle γ from a combination of $B \to DK$ analyses*, LHCb-CONF-2016-001.

[56] BaBar collaboration, J. P. Lees *et al.*, *Evidence of $B^+ \to \tau^+\nu$ decays with hadronic B tags*, Phys. Rev. **D88** (2013) 031102, arXiv:1207.0698.

[57] Belle collaboration, I. Adachi *et al.*, *Evidence for $B^- \to \tau^-\bar{\nu}_\tau$ with a hadronic tagging method using the full data sample of Belle*, Phys. Rev. Lett. **110** (2013) 131801, arXiv:1208.4678.

[58] Belle collaboration, B. Kronenbitter *et al.*, *Measurement of the branching fraction of $B^+ \to \tau^+\nu_\tau$ decays with the semileptonic tagging method*, Phys. Rev. **D92** (2015) 051102, arXiv:1503.05613.

[59] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, arXiv:1412.7515, updated results and plots available at http://www.slac.stanford.edu/xorg/hfag/.

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

[61] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125, arXiv:hep-ph/9512380.

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

[63] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Annalen der Physik **384** (1926) 361.

[64] V. Weisskopf and E. P. Wigner, *Berechnung der natürlichen Linienbreite auf Grund der Diracschen Lichttheorie*, Z. Phys. **63** (1930) 54; *Über die natürliche Linienbreite in der Strahlung des harmonischen Oszillators*, Z. Phys. **65** (1930) 18.

[65] A. J. Buras, W. Slominski, and H. Steger, *$B^0\overline{B}^0$ mixing, CP violation and the B-meson decay*, Nucl. Phys. **B245** (1984) 369.

[66] BaBar collaboration, J. P. Lees *et al.*, *Measurement of CP asymmetries and branching fractions in charmless two-body B-meson decays to pions and kaons*, Phys. Rev. **D87** (2013) 052009, arXiv:1206.3525.

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

[68] 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.

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

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

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

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

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

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

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

[76] F. Marcastel, *CERN's Accelerator Complex. La chaîne des accélérateurs du CERN*, Oct, 2013. General Photo.

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

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

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

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

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

[82] S. Bachmann *et al.*, *Ageing in the LHCb outer tracker: Phenomenon, culprit and effect of oxygen*, Nucl. Instrum. Meth. **A617** (2010) 202.

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

[84] E. Guschin and S. V. Laptev, *Monte-carlo study of LHCb preshower*, Tech. Rep. LHCb-2000-030, CERN, Geneva, Oct, 2000.

[85] R. Alemany-Fernandez, F. Follin, and R. Jacobsson, *The LHCb online luminosity control and monitoring*, in *Proceedings, 4th International Particle Accelerator Conference (IPAC 2013)*, 2013. CERN-ACC-2013-0028.

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

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

[88] 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.

[89] R. Aaij *et al.*, *The Brunel project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/, Oct, 2016.

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

[91] R. Aaij *et al.*, *The DaVinci project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/, Oct, 2016.

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

[93] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175; T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

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

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

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

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

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

[99] R. Aaij *et al.*, *The Boole project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/, Oct, 2016.

[100] R. Aaij *et al.*, *The Moore project*, http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/, Oct, 2016.

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

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

[103] 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.

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

[105] D. Fazzini, M. Calvi, and B. Khanji, *Development of "same side" flavour tagging algorithms for measurements of flavour oscillations and CP violation in the B^0 mesons system*, CERN-THESIS-2015-040.

[106] LHCb collaboration, R. Aaij *et al.*, *New algorithms for identifying the flavour of B^0 mesons using pions and protons*, arXiv:1610.06019, submitted to Eur. Phys. J.

[107] A. Hoecker *et al.*, *TMVA: Toolkit for multivariate data analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[108] F. Pedregosa *et al.*, *Scikit-learn: Machine learning in Python*, J. Mach. Learn. Res. **12** (2011) 2825.

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

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

[111] L. Breiman, *Random forests*, Machine Learning **45** (2001) 5.

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

[113] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning and an application to boosting*, Jour. Comp. and Syst. Sc. **55** (1997) 119.

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

[115] B. H. Denby, *Neural networks and cellular automata in experimental high-energy physics*, Comput. Phys. Commun. **49** (1988) 429.

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

[117] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. arXiv:physics/0308063.

[118] Y. Xie, *Formalism for simulation-based optimization of measurement errors in high energy physics*, arXiv:0901.3305.

[119] C. de Boor, *A practical guide to splines*, vol. 27 of *Applied Mathematical Sciences*, Springer-Verlag New York, 1978.

[120] O. Behnke, K. Kröninger, T. Schörner-Sadenius, and G. Schott, eds., *Data analysis in high energy physics*, Wiley-VCH, Weinheim, Germany, 2013.

[121] W. Verkerke and D. Kirkby, *RooFit*. http://roofit.sourceforge.net.

[122] J. R. Klein and A. Roodman, *Blind analysis in nuclear and particle physics*, Ann. Rev. Nucl. Part. Sci. **55** (2005) 141.

[123] H. G. Moser and A. Roussarie, *Mathematical methods for $B^0\overline{B}^0$ oscillation analyses*, Nucl. Instrum. Meth. **A384** (1997) 491.

[124] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.

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

[126] C. Cauet, *Precision measurement of the CKM parameter* $\sin(2\beta)$ *with the LHCb experiment*, PhD thesis, TU Dortmund University, Aug, 2015, CERN-THESIS-2015-291.

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

[128] W. Fetscher *et al.*, *Regeneration of arbitrary coherent neutral kaon states: A new method for measuring the $K^0 - \overline{K}^0$ forward scattering amplitude*, Z. Phys. **C72** (1996) 543; B. R. Ko, E. Won, B. Golob, and P. Pakhlov, *Effect of nuclear interactions of neutral kaons on CP asymmetry measurements*, Phys. Rev. **D84** (2011) 111501, arXiv:1006.1938.

[129] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B \to D^+D^-$ decays*, arXiv:1608.06620, submitted to Phys. Rev. Lett.

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

[131] S. Ali *et al.*, *Measurement of time-dependent CP-violation observables in $B_s^0 \to D_s^{\mp}K^{\pm}$*, LHCb-ANA-2012-068.

[132] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, LHCb-PUB-2011-001.

[133] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\overline{B}^0$–B^0 and $\overline{B}_s^0$–B_s^0 production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739** (2014) 218, arXiv:1408.0275.

[134] T. Brambach *et al.*, *Measurement of CP violation in the time-dependent analysis of $B^0 \to J/\psi K_S^0$ decays with the 2010 data*, LHCb-ANA-2011-004. LHCb-CONF-2011-004.

[135] P. Gandini, *Measurement of the B_s^0 meson lifetime in the flavour specific hadronic decay $B_s^0 \to D_s^-\pi^+$*, LHCb-ANA-2014-015.

[136] M. Schellenberg, *Messung der zerfallszeitabhängigen CP-Asymmetrie im Zerfall $B^0 \to D^{*+}D^-$ mit dem LHCb Experiment*, Master's thesis, TU Dortmund University, Sep, 2015.

[137] Y. Amhis *et al.*, *b-hadrons absolute lifetime measurements*, LHCb-ANA-2011-055. Linked to LHCb-PAPER-2013-065.

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

[139] OPAL collaboration, K. Ackerstaff *et al.*, *Investigation of CP violation in $B^0 \to J/\psi K_S^0$ decays at LEP*, Eur. Phys. J. **C5** (1998) 379, arXiv:hep-ex/9801022.

[140] ALEPH collaboration, R. Barate *et al.*, *Study of the CP asymmetry of $B^0 \to J/\psi K_S^0$ decays in ALEPH*, Phys. Lett. **B492** (2000) 259, arXiv:hep-ex/0009058.

[141] CDF collaboration, T. Affolder *et al.*, *A measurement of $\sin(2\beta)$ from $B \to J/\psi K_S^0$ with the CDF detector*, Phys. Rev. **D61** (2000) 072005, arXiv:hep-ex/9909003.

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

[143] Belle collaboration, I. Adachi *et al.*, *Precise measurement of the CP violation parameter* $\sin 2\phi_1$ *in* $B^0 \to (c\bar{c})K^0$ *decays*, Phys. Rev. Lett. **108** (2012) 171802, arXiv:1201.4643.

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

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

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

[147] *LHCb: Technical Proposal*, Tech. Proposal, CERN, Geneva, 1998.

[148] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*, CERN-LHCC-2014-016. LHCb-TDR-016.

Acknowledgements

Zunächst möchte ich mich bei Herrn Spaan bedanken. Seitdem ich im Oktober 2011 an Ihren Lehrstuhl gekommen bin, um erst meine Masterarbeit und nun diese Doktorarbeit zu schreiben, waren Sie immer mit Rat und Tat für mich da. Zusätzlich haben Sie uns immer wieder mit interessanten Anekdoten unterhalten.

Mein Dank gilt auch Professor Kröninger, der sich bereit erklärt hat, als Zweitgutachter für meine Dissertation zu fungieren.

In den ersten zwei Jahren meiner Dissertation habe ich gemeinsam mit Christophe an der Messung von CP Verletzung in $B^0 \to J/\psi\, K^0_{\mathrm{S}}$ Zerfällen gearbeitet. Wir haben uns gegenseitig unterstützt, und so nicht nur eine gelungene Analyse abgeliefert, sondern auch jeweils Dissertationsschriften verfassen können. Danke schön.

High energy physics is often a complex field of study, where progress is impossible without collaborative work. Thanks to Paul, Nicoletta and Marta for working with me on the $B^0 \to D^+ D^-$ analysis, especially for performing and providing the flavour-tagging calibration.

Ich möchte mich auch ganz herzlich bei meinen Bürokollegen Alex, Margarete und insbesondere Uli bedanken. Es war immer eine wunderbare Arbeitsatmosphäre, in der ich mich wohl gefühlt habe. Probleme wurden untereinander diskutiert und auch für Gespräche außerhalb des Unialltags war immer Zeit.

Weiterer Dank gebührt Julian für gute Ratschläge, für das Teilen von Erfahrungen, produktive Diskussionen und allgemein Hilfe bei vielen Passagen, die Eingang in diese Arbeit gefunden haben.

Wenn man zu lange an einem Dokument wie dieser Dissertation schreibt, bemerkt man oftmals gar nicht, dass Bezüge fehlen oder Sachverhalte ungenügend erklärt worden sind. Deshalb bin ich Timon, Moritz, Vanessa und Alex sehr dankbar dafür, sich die Zeit genommen und Teile meiner Arbeit Korrektur gelesen zu haben.

Plots problemlos ansprechend aussehen zu lassen, komplexe Systematikstudien mit wenigen Zeilen Code erstellen und auswerten zu können, sowie etliche nützliche Funktionen zur Verfügung zu haben, ist ein Luxus, an den man sich all zu schnell gewöhnt, obwohl eine Menge Arbeit dahinter gesteckt hat. Vielen Dank, Florian, dass du das DooSoftware-Framework erstellt hast.

Meine Begeisterung für die Physik wurde in der Oberstufe durch den leider viel zu früh verstorbenen Herrn Bär, meinen Physiklehrer im Leistungskurs, geweckt. Auch wenn Sie das nicht lesen können, vielen Dank, ohne Sie wäre ich wohl nicht hier gelandet.