

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

1. Gutachter : Prof. Dr. Spaan

2. Gutachter :

Datum des Einreichens der Arbeit:

Abstract

The CP violation observables S_f and C_f in the decays of B^0 and \bar{B}^0 mesons to the $J/\psi K_s^0$ 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 time evolution of 41 560 $B^0 \rightarrow J/\psi K_s^0$ decays yields

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

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

In a flavour-tagged, decay-time-dependent analysis of 1410 $B^0 \rightarrow D^+D^-$ 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 obtained values of the two decay modes, are constrained to be small.

Zusammenfassung

Die CP -Verletzung beschreibenden Observablen S_f und C_f für die Zerfallskanäle $B^0 \rightarrow J/\psi K_s^0$ und $B^0 \rightarrow D^+D^-$ werden gemessen. 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 $B^0 \rightarrow J/\psi K_s^0$ Zerfällen werden die CP Observablen zu

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

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

In der zeitabhängigen Analyse von 1410 getaggten $B^0 \rightarrow D^+D^-$ 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	1
2	Standard Model of Particle Physics	3
2.1	Particles	3
2.2	Forces and couplings	4
2.3	Symmetries and conservation laws	5
2.4	Problems and possible extensions	6
3	CP Violation	9
3.1	The KM mechanism and the CKM matrix	9
3.2	The system of neutral B^0 mesons	11
3.3	Types of CP violation	14
3.3.1	Direct CP violation	14
3.3.2	Indirect CP violation	15
3.3.3	CP violation in the interference of decay and decay after mixing	16
3.4	CP violation in $b \rightarrow c\bar{c}s$ decays	17
3.5	CP violation in $b \rightarrow c\bar{c}d$ decays	18
4	The LHCb Experiment at the LHC	21
4.1	The Large Hadron Collider	21
4.2	The LHCb detector	22
4.3	The LHCb trigger system	26
4.4	The LHCb software	27
4.4.1	Reconstruction	27
4.4.2	Stripping	28
4.4.3	Monte Carlo simulation	28
4.5	Flavour tagging	29
4.5.1	Opposite-side flavour tagging	31
4.5.2	Same-side flavour tagging	31
5	Data Analysis Tools and Methods	33
5.1	Maximum likelihood method	33
5.2	Selection	33
5.2.1	Multivariate selection	34
5.2.2	Unfolding data distributions using sWeights	35
5.2.3	Figures of merit	36
5.3	Spline interpolation	38

Contents

5.4	Bootstrapping method	38
5.5	Blinding	38
5.6	Decay time resolution	39
5.7	Flavour-tagging calibration	39
5.7.1	Calibration using $B^0 \rightarrow D_s^+ D^-$	40
5.7.2	Calibration using $J/\psi X$ channels	42
6	Measurement of CP Violation in $B^0 \rightarrow J/\psi K_s^0$ Decays	45
6.1	Data Preparation	45
6.2	Decay time acceptance	46
6.3	Decay time resolution	47
6.4	Backgrounds	49
6.5	Nominal fit	51
6.6	Studies of systematic effects	52
7	Measurement of CP Violation in $B^0 \rightarrow D^+ D^-$ Decays	55
7.1	Selection	55
7.1.1	Preselection	55
7.1.2	Vetoos	56
7.1.3	Multivariate analysis	57
7.1.4	Final selection	60
7.2	Mass fit	61
7.3	Decay time fit	65
7.3.1	Decay time resolution	65
7.3.2	Decay time acceptance	67
7.3.3	External inputs	68
7.3.4	Results	68
7.4	Studies of systematic effects	71
7.4.1	Cross-checks	71
7.4.2	Decay Time Fit Bias	73
7.4.3	Fit Model	73
7.4.4	Further Studies	77
7.4.5	Total Systematic Uncertainty	78
8	Discussion	81
8.1	Comparison with previous measurements of $\sin 2\beta_{(\text{eff})}$	81
8.2	Comparison between CP violation in $B^0 \rightarrow J/\psi K_s^0$ and in $B^0 \rightarrow D^+ D^-$ decays	82
9	Conclusion	85
	Bibliography	88

¹ 1 Introduction

² From latest measurements the age of our universe is calculated to be 13.8 billion
³ years [1]. Theoretical models exist that describe the evolution of the universe during
⁴ this incredibly long period of time, starting from directly after the Big Bang [2]
⁵ until today. But although there are models describing phenomena ranging from
⁶ the largest (cosmological models) down to the smallest scales (Standard Model of
⁷ particle physics (SM)), a comprehensive model has not been found so far. One of
⁸ the most striking observations is that we are surrounded by matter, while there are
⁹ no antimatter clusters. Though, according to big bang theories at the origin of the
¹⁰ universe the same amount of matter and antimatter has been created. Sakharov
¹¹ has proposed three conditions [3] that need to be fulfilled to explain this so called
¹² baryogenesis: At some point in the evolution of the universe the thermal equilibrium
¹³ has to be imbalanced, the baryon number conservation is required to be violated,
¹⁴ and the C (and even CP) symmetry has to be violated as well. While baryon
¹⁵ number violation has not been observed so far, *e.g.* the lifetime of the proton
¹⁶ is determined to be greater than 5.9×10^{33} years [4], evidence for a violation of
¹⁷ the thermal equilibrium has been found [5]. Moreover, already fifty years ago
¹⁸ C -violating and shortly thereafter CP -violating processes have been discovered [6].
¹⁹ Still, there are reasons why even after so many years this topic is interesting. On
²⁰ the one hand, the size of CP violation in the SM is orders of magnitude below
²¹ what is required to explain the matter-antimatter asymmetry [7]. On the other
²² hand, CP -violating processes are an excellent test bed for the predictions in the
²³ quark-flavour sector of the Standard Model of particle physics. The unitarity of
²⁴ the Cabibbo-Kobayashi-Maskawa (CKM) matrix, which describes the probability
²⁵ of quark transitions, is a fundamental requisite of the SM. It can be tested by
²⁶ studying the unitarity triangle, which represents one of the unitarity conditions.
²⁷ The determination of the angle β of this triangle, or more precisely of the derived
²⁸ quantity $\sin 2\beta$, is the common theme of this thesis. Precision measurements of
²⁹ CP violation in charmonium decays, *i.e.* in decay modes involving a $c\bar{c}$ resonance,
³⁰ and in open-charm decays, *i.e.* in decay modes with at least one hadron containing
³¹ exactly one c quark, are performed. More specifically, CP violation is studied in
³² $B^0 \rightarrow J/\psi K_s^0$ and in $B^0 \rightarrow D^+ D^-$ decays. The two analyses exploit the full Run I
³³ data sample, which is the world's largest sample of B^0 mesons. It corresponds
³⁴ to an integrated luminosity of 3 fb^{-1} collected with the LHCb detector in proton-
³⁵ proton collisions at centre-of-mass energies of 7 and 8 TeV. While the decay mode
³⁶ $B^0 \rightarrow J/\psi K_s^0$ offers a very clean determination of $\sin 2\beta$ at a very high precision,
³⁷ the main purpose of the study of $B^0 \rightarrow D^+ D^-$ decays is to constrain higher-order
³⁸ Standard Model corrections occurring in measurements of the CP -violating phase.

1 Introduction

39 These contributions need to be controlled to distinguish them from effects caused
40 by physics beyond the Standard Model of particle physics, often referred to as “New
41 Physics”. There are several reasons to believe that the SM needs to be extended.
42 The measurements of rotation curves of galaxies [8] have lead to the assumption of
43 the presence of dark matter. This is not a small effect. Around 26 % of the energy
44 density in the universe are assigned to dark matter compared to around 5 % for
45 normal baryonic matter [1]. However, dark matter is not accounted for in the SM,
46 and no dark matter candidate has been found so far. Furthermore, in the SM the
47 neutrinos are set to be massless, which is disproved by the observation of neutrino
48 oscillations [9, 10], awarded with the Nobel prize in 2015 [11].

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

⁵⁹ 2 Standard Model of Particle Physics

⁶⁰ ⁶¹ 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).
⁶² 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, 14].

⁶⁹ 2.1 Particles

⁷⁰ In the SM 12 fermions, which are elementary particles with spin $\frac{1}{2}$, 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):

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \quad (2.1)$$

⁷⁸ 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 ⁷⁹ a meson, three quarks a baryon, and just recently evidence for the existence of ⁸⁰ four and five quark bound states (tetraquarks respectively pentaquarks) has been ⁸¹ found [17–19]. A colour charge is associated to the quarks, which can take three ⁸² different types. However, the colour charges add up in a way that the hadrons are ⁸³ colourless. The electric charge of the up-type quarks is $+\frac{2}{3}$ and of the down-type ⁸⁴ quarks $-\frac{1}{3}$ of the elementary charge. The leptons are also classified in three families, ⁸⁵ each consisting of a negatively charged particle, in increasing order of mass the ⁸⁶ electron (e), the muon (μ) and the tauon (τ), and a corresponding neutral neutrino,

2 Standard Model of Particle Physics

88 which is set to be massless in the SM:

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix} \quad (2.2)$$

89 Additionally, 12 gauge bosons with integer spin, which mediate the forces (see
90 Sec. 2.2), are described in the SM. The gauge bosons are the massless photon (γ),
91 the three massive mediators of the weak force (W^+ , W^- , Z^0), and the eight gluons,
92 which carry different colour charge configurations. The SM has been completed by
93 the discovery of the Higgs boson [20, 21], a massive boson with spin 0.

2.2 Forces and couplings

95 The three interactions that are described in the SM are the electromagnetic, the
96 weak and the strong force. These differ dramatically in their strength and the
97 time scales, in which the interactions proceed. Strong decays typically take 10^{-23} s,
98 electromagnetic decays 10^{-16} s, while the decay times of the weak interaction range
99 between 10^{-13} s and a few minutes.

100 The classical approach of the electromagnetic interaction is given by Maxwell's
101 equations. These are generalized into a relativistic quantum field theory by the
102 Quantum Electrodynamics (QED) [22–24]. The QED can be derived from the
103 Lagrangian of a free fermion field

$$\mathcal{L}_0 = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi \quad (2.3)$$

104 by extending the global to a local $U(1)$ symmetry. This is done by replacing ∂_μ
105 with the corresponding covariant derivative D_μ

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i e A_\mu, \quad (2.4)$$

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

$$\mathcal{L}_A = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad (2.5)$$

109 with the field strength tensor $F_{\mu\nu}$, which is the compressed formulation of Maxwell's
110 classical equations. However, in the SM the electromagnetic interaction is unified
111 with the weak interaction in the electroweak $SU(2) \times U(1)$ symmetry group [25–27].
112 The weak part couples to the weak isospin and differs between left-handed and
113 right-handed fermion fields, where handedness gives the orientation of the spin with
114 respect to the momentum vector. Neutrinos have the peculiarity that they only
115 exist as right-handed fermions. The electroweak gauge symmetry is broken, which
116 becomes apparent as the photon is massless while the W^\pm and Z^0 bosons are not.
117 The Higgs mechanism [28] is responsible for this symmetry breaking, which comes

along with the need of 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 exact values of the masses. Moreover, they 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 [29]. This topic is explained in more detail in Sec. 3.1.

The third fundamental, strong force is described 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 polarization, which is possible as gluons carry colour charge by themselves and therefore can couple to each other, outperforms quark polarization effects leading to asymptotic freedom of the quarks on very short distances [30,31]. 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 summarized in the quark-antiquark potential

$$V_{\text{QCD}} = -\frac{4}{3} \frac{\alpha_s}{r} + k r, \quad (2.6)$$

where r is the distance between the two fermions and $k \approx 1 \text{ GeV fm}^{-1}$ [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 [32].

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 [33] 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-perturbative effects. However, in the early universe these effects might have been larger [34], while $(B - L)$ is an exact symmetry. The lepton family number, *i.e.* the individual lepton number for electrons, muons and

2 Standard Model of Particle Physics

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 $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays show an asymmetry with a significance corresponding to 2.6 standard deviations [35]. 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 [36, 37]. Another approximate symmetry is the parity operation (P), which performs a spatial inversion of all coordinates

$$P\Psi(r) = \Psi(-r), \quad (2.7)$$

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

$$C|p\rangle = |\bar{p}\rangle. \quad (2.8)$$

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 [38–40] 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 [41]. A first step would be a generalization 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 physics 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 in a quantum mechanical

2.4 Problems and possible extensions

192 framework. The unification of the forces would probably emerge at energy scales
193 of 10^{16} GeV. However, quantum corrections from those mass scales would heavily
194 influence the Higgs mass, which is measured to be around $125 \text{ GeV}/c^2$ [42]. In the SM
195 this hierarchy problem is solved by a fine tuning of tree-level and loop contributions,
196 which exactly cancel each other. Other explanations are given by extending the
197 SM with new symmetries, like models including supersymmetry [43–45]. From a
198 theoretical point of view it is also unsatisfactory that the SM includes so many free
199 parameters, like the masses of the constituents or the number of generations. In
200 addition, the SM only applies to the processes of ordinary matter, which makes
201 only about 5 % of the total energy density in the universe [1], whereas it lacks
202 an explanation for dark matter or dark energy. Furthermore, the amount of CP
203 violation in the weak sector can not account for the baryon asymmetry in the
204 universe, *i.e.* the dominance of matter without any large clusters of antimatter in
205 the universe. The observation of neutrino oscillations [9, 10], which implicitly is
206 an observation of mass differences between the neutrino generations, shows that
207 neutrinos can not be massless as assumed in the SM. The special role of neutrinos
208 in the SM to appear only left-handed could be corrected if they were Majorana
209 particles, *i.e.* their own antiparticles [46].

210 3 CP Violation

211 In this chapter the concept of CP violation, its origin and manifestation in the SM,
212 as well as the possibilities to measure CP violation, are described. The formalism
213 closely follows Refs. [47] and [48].

214 3.1 The KM mechanism and the CKM matrix

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

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

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

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

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

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

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

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

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

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

3 CP Violation

removes nine of the degrees of freedom. Another five phases can be constrained by global rephasings between the six mass fields. So, 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 [29]. The corresponding parametrisation of the CKM matrix is

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

with c_i and s_i being shorthand for the cosine respectively sine of the three Euler angles, and δ is 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

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

which expresses the amount of *CP* violation in the SM [49]. It is measured to be $J = (3.04^{+0.21}_{-0.20}) \times 10^{-5}$ [50]. 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

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

When depicting the 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 as can be seen in Fig. 3.1. Using the parametrisation of the

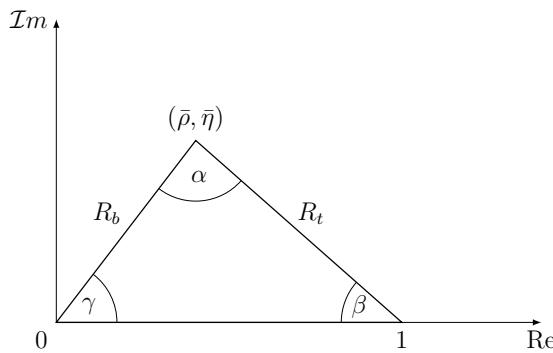


Figure 3.1: Schematic representation of the CKM unitarity triangle.

CKM matrix by Wolfenstein [51], which is an expansion in powers of $\lambda \equiv |V_{us}| =$

250 0.2248 ± 0.0006 [50],

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

251 the other two sides are given by

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

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

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

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

254 The three angles of the unitarity triangle are defined by

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

255 The unitarity triangle is overconstrained, *i.e.* there are measurements of more
256 independent parameters than necessary to fully characterise the shape of the triangle.
257 The angle α can be studied with $B^0 \rightarrow \pi\pi$ decays [52–54], β is precisely measured
258 using the time-dependent CP asymmetry in $B^0 \rightarrow J/\psi K_s^0$ decays (see Sec. 3.4), and γ
259 can be extracted from a combination of results in $B \rightarrow Dh$ decays [55]. Semileptonic
260 b -hadron decays are used to determine the size of the triangle side R_b . Further
261 information on $|V_{ub}|$ comes from studies of $B^+ \rightarrow \tau^+\nu_\tau$ decays [56–58]. The second
262 non-trivial side length R_t is constrained by measurements of the mixing frequencies
263 Δm_d and Δm_s in the system of neutral B^0 and B_s^0 mesons [59]. Furthermore,
264 information on the position of the apex can be gained from the measurement of
265 CP violation in the neutral kaon system [50]. All these inputs are put into a global
266 fit, which mainly checks how well the different constraints agree on the position of
267 the apex. The latest result of the CKMfitter group in Fig. 3.2 shows a very good
268 agreement of all present tests of CP violation in the SM as the area for the position
269 of the apex is relatively small.

270 3.2 The system of neutral B^0 mesons

271 In the system of neutral B^0 mesons four decay amplitudes occur. The decay
272 amplitude A_f stands for the decay of a B^0 meson into a final state f , while \bar{A}_f is
273 the decay amplitude of an \bar{B}^0 meson into the same final state. Similarly, the decay
274 amplitudes into the CP -conjugated final state \bar{f} can be defined. By convention

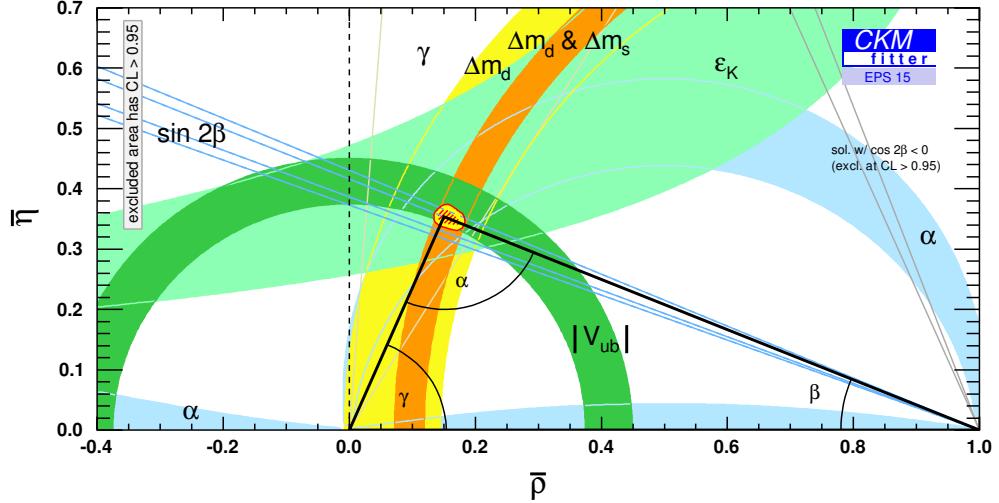


Figure 3.2: Unitarity triangle with constraints from measurements of various quantities [60].

²⁷⁵ B^0 mesons consist of \bar{b} and d quarks, while \bar{B}^0 mesons contain b and \bar{d} quarks.
²⁷⁶ These two mesons can mix, *i.e.* they can oscillate between the two flavour states.
²⁷⁷ As flavour-changing neutral currents (FCNC) are forbidden in the SM, the B^0 – \bar{B}^0
²⁷⁸ 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

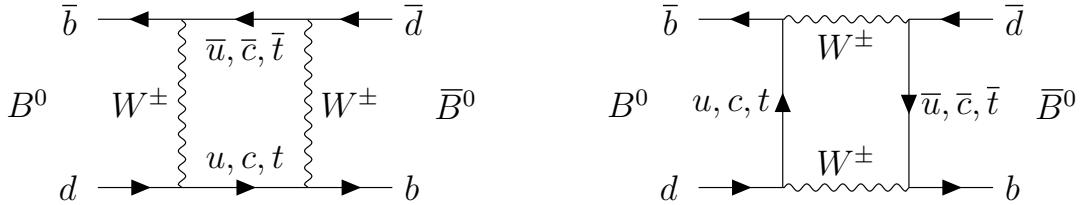


Figure 3.3: Box diagrams of B^0 – \bar{B}^0 oscillation.

²⁷⁹ decay amplitude is given by
²⁸⁰

$$\mathcal{A}(\bar{B}^0 \rightarrow B^0) \propto \frac{g_2^4}{16\pi^2} (V_{tb} V_{td}^*)^2 \frac{\langle B^0 | (\bar{d}_L \gamma_\mu b_L) (\bar{d}_L \gamma^\mu b_L) | \bar{B}^0 \rangle}{M_W^2} S(x_t), \quad (3.14)$$

²⁸¹ with the W boson mass M_W , the squared fraction of the top quark mass $x_t =$
²⁸² m_t^2/M_W^2 and the Inami-Lim function [61]

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

283 Here, it is accounted for the suppression of contributions with up and charm quarks
 284 in the loop due to $m_{u,c}^2 \ll m_t^2$. Formulae (3.14) and (3.15) are taken from Ref. [12].

285 To derive the time evolution of initially produced B^0 and \bar{B}^0 mesons the
 286 Schrödinger equation [62] needs to be solved. Assuming the Wigner-Weisskopf
 287 approximation [63], *i.e.* (excited) states do not know about their past, which is
 288 valid since the time scale of weak decays is significantly larger than the time scale
 289 of the production via the strong force, the effective Schrödinger equation for the
 290 wave function representing the system of B^0 and \bar{B}^0 mesons can be written as

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

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

$$\begin{aligned} |B_L\rangle &= p|B^0\rangle + q|\bar{B}^0\rangle, \\ |B_H\rangle &= p|B^0\rangle - q|\bar{B}^0\rangle, \end{aligned} \quad (3.17)$$

299 with the complex coefficients p and q , which fulfil the normalisation condition
 300 $|p|^2 + |q|^2 = 1$ and whose ratio can be expressed with the matrix elements as

$$\frac{q}{p} = \sqrt{\frac{2M_{12}^* - i\Gamma_{12}^*}{2M_{12} - i\Gamma_{12}}}. \quad (3.18)$$

301 Explicit calculations of Γ_{12} , as performed in Ref. [64], show that to a very good
 302 approximation Eq. (3.18) can be simplified to

$$\frac{q}{p} \approx \sqrt{\frac{M_{12}^*}{M_{12}}} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}. \quad (3.19)$$

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

$$\begin{aligned} \mu_L &= m_L - i/2\Gamma_L, \\ \mu_H &= m_H - i/2\Gamma_H, \end{aligned} \quad (3.20)$$

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

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

3 CP Violation

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

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

308 with the time-dependent coefficients

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

309 The mass difference $\Delta m_d = m_H - m_L$ drives the oscillation and can be interpreted
 310 as mixing frequency, while $\Delta\Gamma_d = \Gamma_H - \Gamma_L$ is negligible in the B^0 meson sector [59].
 311 The differential decay rates, *i.e.* the squared decay amplitudes, can be expressed
 312 by combining the terms in Eqs. (3.22) and (3.23) as

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

313 using

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f}. \quad (3.25)$$

3.3 Types of CP violation

314 There are three different manifestations of CP violation. It can occur, when the
 315 decay amplitudes differ between CP conjugated processes (see Sec. 3.3.1), when
 316 the mass eigenstates are no CP eigenstates (see Sec. 3.3.2), and when there is
 317 interference between direct decays and decays to the same final state after mixing
 318 (see Sec. 3.3.3). While the first type can appear for charged and neutral hadrons,
 319 the latter are only possible for neutral decays.
 320

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

3.3.1 Direct CP violation

322 Two different type of phases can contribute to decay amplitudes, weak phases and
 323 strong phases. Weak phases can enter through the CKM matrix and take the
 324 opposite sign for A_f and $\bar{A}_{\bar{f}}$. Strong phases typically appear in scattering processes
 325

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

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

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

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

This type of CP violation is observed in charmless two-body decays of neutral B mesons [65–67].

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

$$\left| \frac{q}{p} \right| \neq 1. \quad (3.28)$$

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

$$\mathcal{P}(B^0 \rightarrow \bar{B}^0, t) \neq \mathcal{P}(\bar{B}^0 \rightarrow B^0, t), \quad (3.29)$$

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

$$a_{\text{sl}} = \frac{\Gamma(\bar{B}^0(t) \rightarrow \ell^+ \nu X) - \Gamma(B^0(t) \rightarrow \ell^- \nu X)}{\Gamma(\bar{B}^0(t) \rightarrow \ell^+ \nu X) + \Gamma(B^0(t) \rightarrow \ell^- \nu X)} = \frac{1 - |q/p|^4}{1 + |q/p|^4} \quad (3.30)$$

yield values consistent with zero [68, 69], though the precision is one to two magnitudes above the SM expectations. This lack in precision is not only due to statistical

3 CP Violation

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 \bar{B}^0 mesons. The definition from Eq. (3.25) slightly changes to

$$\lambda = \eta_{CP} \frac{q}{p} \frac{\bar{A}_f}{A_f}. \quad (3.31)$$

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

$$CP|f_{CP}\rangle = |\bar{f}_{CP}\rangle = \eta_{CP}|f_{CP}\rangle = \pm 1|f_{CP}\rangle. \quad (3.32)$$

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

$$\mathcal{Im} \lambda \neq 0. \quad (3.33)$$

The time-dependent asymmetry

$$\mathcal{A}(t) \equiv \frac{\Gamma(\bar{B}^0(t) \rightarrow f_{CP}) - \Gamma(B^0(t) \rightarrow f_{CP})}{\Gamma(\bar{B}^0(t) \rightarrow f_{CP}) + \Gamma(B^0(t) \rightarrow f_{CP})} \quad (3.34)$$

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

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

\mathcal{A} can be written as

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

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

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

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

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

374 3.4 CP violation in $b \rightarrow c\bar{c}s$ decays

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

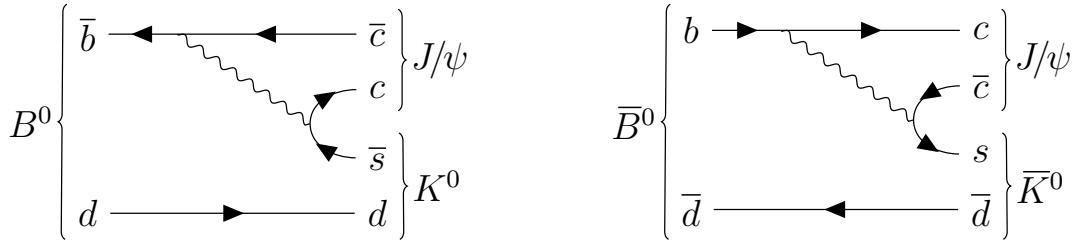


Figure 3.4: Tree Feynman diagrams of $B^0 \rightarrow J/\psi K_s^0$ for both flavours.

380 like for the B mesons the flavour eigenstates of the kaons are a superposition of the
 381 CP mass eigenstates:
 382

$$|K_s^0\rangle = p_K |K^0\rangle - q_K |\bar{K}^0\rangle \quad (3.39)$$

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

$$\frac{\bar{A}_{J/\psi K_s^0}}{A_{J/\psi K_s^0}} = -\frac{p_K}{q_K} \frac{\bar{A}_{J/\psi \bar{K}^0}}{A_{J/\psi K^0}}. \quad (3.40)$$

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

$$\frac{p_K}{q_K} = -\frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}}. \quad (3.41)$$

3 CP Violation

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

$$\frac{\bar{A}_{J/\psi \bar{K}^0}}{A_{J/\psi K^0}} = \frac{V_{cb} V_{cs}^*}{V_{cb}^* V_{cs}}. \quad (3.42)$$

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

$$\lambda_{J/\psi K_s^0} = -\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \frac{V_{cs} V_{cd}^*}{V_{cs}^* V_{cd}} \frac{V_{cb} V_{cs}^*}{V_{cb}^* V_{cs}} = -\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \frac{V_{cd}^* V_{cb}}{V_{cd} V_{cb}^*} \quad (3.43)$$

393 The minus sign indicates that the final state $J/\psi K_s^0$ is CP -odd, as an angular
 394 momentum of $l = 1$ is necessary to compensate that the B^0 meson as initial state
 395 has spin zero, while the final state consists of a CP -even J/ψ meson with spin one
 396 and an almost CP -even¹ K_s^0 meson with spin zero. Due to only considering the
 397 dominant tree contribution the absolute value of $\lambda_{J/\psi K_s^0}$ is unity. Therefore, the
 398 CP observables from Eq. (3.37) become

$$C_{J/\psi K_s^0} = 0, \quad (3.44)$$

399 and

$$S_{J/\psi K_s^0} = \text{Im} \lambda_{J/\psi K_s^0} = \sin \left[\arg \left(-\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \frac{V_{cd}^* V_{cb}}{V_{cd} V_{cb}^*} \right) \right] \quad (3.45)$$

$$= \sin \left[\arg \left(- \left(\frac{V_{cd}^* V_{cb}}{V_{tb} V_{td}^*} \right)^2 \right) \right] = \sin 2\beta. \quad (3.46)$$

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

404 The corresponding golden decay channel with an s quark as spectator quark is
 405 $B_s^0 \rightarrow J/\psi \phi$. In principle, this provides a clean measurement of the mixing phase
 406 ϕ_s of the B_s^0 meson system. However, to extract the CP observable an angular
 407 analysis is required to disentangle the different spin configurations of the two vector
 408 mesons in the final state.

409 3.5 CP violation in $b \rightarrow c\bar{c}d$ decays

410 The decay $B^0 \rightarrow D^+ D^-$ can be described with the Feynman diagrams in Fig. 3.5.
 411 The tree diagram (T) proceeds via a $b \rightarrow c\bar{c}d$ quark transition, which is CKM

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

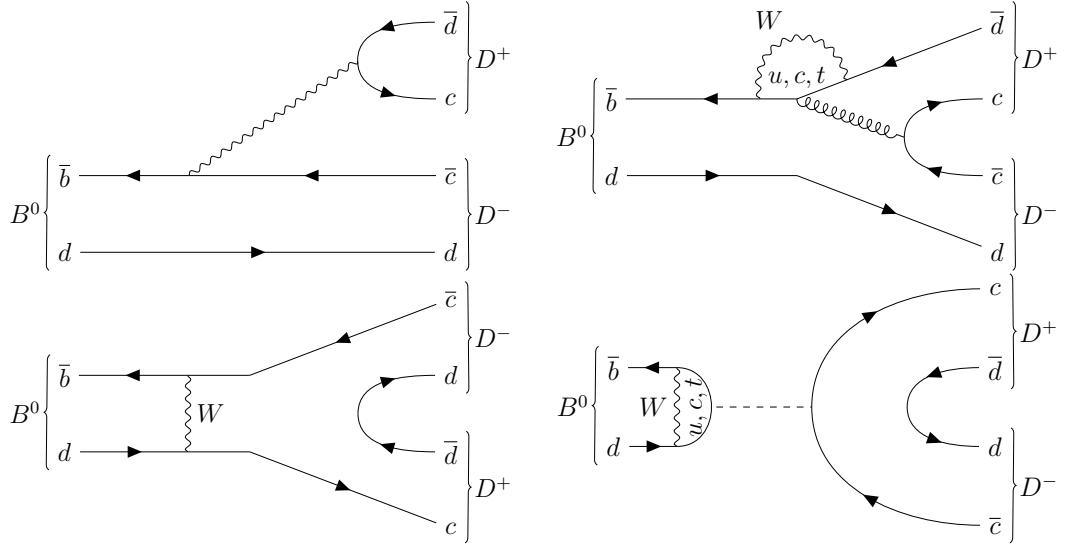


Figure 3.5: Main Feynman diagrams contributing to $B^0 \rightarrow D^+ D^-$ decays. Apart from the tree diagram (top left), a penguin diagram (top right), an exchange diagram (bottom left) and a penguin annihilation diagram (bottom right) are shown.

⁴¹² suppressed. The contributions from the other diagrams, especially the penguin
⁴¹³ diagrams ($P^{(q)}$ with $q = c$ and t quarks in the loop), but also exchange (E) and
⁴¹⁴ penguin annihilation diagrams ($PA^{(q)}$), need to be taken into account as well because
⁴¹⁵ they can carry different weak phases and are not Cabibbo-suppressed. Thus, the
⁴¹⁶ decay amplitude is given by [70–72]

$$A(B^0 \rightarrow D^+ D^-) = V_{cb} \mathcal{A}[1 - ae^{i\theta} e^{i\gamma}], \quad (3.47)$$

⁴¹⁷ with

$$\mathcal{A} \equiv V_{cd}^* [T + E + P^{(c)} + PA^{(c)} - P^{(t)} + PA^{(t)}], \quad (3.48)$$

⁴¹⁸ and

$$ae^{i\theta} \equiv R_b \left[\frac{P^{(u)} + PA^{(u)} - P^{(t)} + PA^{(t)}}{T + E + P^{(c)} + PA^{(c)} - P^{(t)} + PA^{(t)}} \right]. \quad (3.49)$$

⁴¹⁹ Here, R_b is a side length of the unitarity triangle defined in Eq. (3.10). The angle γ
⁴²⁰ of the unitarity triangle is a CP -violating weak phase, while a and θ are hadronic
⁴²¹ CP -conserving parameters. Therefore, the corresponding \bar{B}^0 decay amplitude is

$$A(\bar{B}^0 \rightarrow D^+ D^-) = V_{cb}^* \mathcal{A}[1 - ae^{i\theta} e^{-i\gamma}]. \quad (3.50)$$

⁴²² The parameter describing CP violation in the interference can be written as

$$\begin{aligned} \lambda_{D^+ D^-} &= \frac{V_{tb}^* V_{td} V_{cb} V_{cd}^*}{V_{tb} V_{td}^* V_{cb}^* V_{cd}} \frac{1 - ae^{i\theta} e^{i\gamma}}{1 - ae^{i\theta} e^{-i\gamma}}, \\ &= e^{-i2\beta} \frac{1 - ae^{i\theta} e^{i\gamma}}{1 - ae^{i\theta} e^{-i\gamma}}, \end{aligned} \quad (3.51)$$

3 CP Violation

423 using the ratio of the mixing coefficients from Eq. (3.19) and the definition of the
 424 unitarity triangle β from Eq. (3.13). Different than for $B^0 \rightarrow J/\psi K_s^0$ (cf. Eq. (3.43))
 425 the CP eigenvalue is $\eta_{CP} = +1$, since no spins are involved in the decay $B^0 \rightarrow D^+ D^-$.
 426 The hadronic parameters cannot be calculated reliably within QCD [72]. Thus,
 427 they must be determined through a measurement of the CP observables, which can
 428 be expressed via

$$S_{D^+ D^-} = -\frac{\sin \phi_d - 2a \cos \theta \sin(\phi_d + \gamma) + a^2 \sin(\phi_d + 2\gamma)}{1 - 2a \cos \theta \cos \gamma + a^2}, \quad (3.52)$$

$$C_{D^+ D^-} = \frac{2a \sin \theta \sin \gamma}{1 - 2a \cos \theta \cos \gamma + a^2}. \quad (3.53)$$

429 The term $S_{D^+ D^-}$ gives access to the mixing phase ϕ_d , which is related to β through

$$\phi_d = 2\beta + \phi_d^{\text{NP}}, \quad (3.54)$$

430 and thus considers new physics contributions as well. While $S_{D^+ D^-}$ is caused by
 431 interference between the direct decay and the decay after mixing, $C_{D^+ D^-}$ might differ
 432 from zero due to interferences between tree and penguin contributions. Different
 433 than in the case of $B^0 \rightarrow J/\psi K_s^0$ only an effective phase

$$\phi_d^{\text{eff}} = \phi_d + \Delta\phi_d \quad (3.55)$$

434 with

$$\sin \phi_d^{\text{eff}} = -\frac{S_{D^+ D^-}}{\sqrt{1 - C_{D^+ D^-}^2}} \quad (3.56)$$

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

$$\tan \Delta\phi = \frac{a^2 \sin 2\gamma - 2a \cos \theta \sin \gamma}{1 - 2a \cos \theta \cos \gamma + a^2 \cos 2\gamma}. \quad (3.57)$$

436 The decay channel $B_s^0 \rightarrow D_s^+ D_s^-$ is also governed by a $b \rightarrow c\bar{c}d$ transition. It gives
 437 access to ϕ_s . However, the measurement is as well polluted by hadronic penguin
 438 effects. Since $B_s^0 \rightarrow D_s^+ D_s^-$ is related to $B^0 \rightarrow D^+ D^-$ via U-spin symmetry, the
 439 phase shift $\Delta\phi$ can be transferred.

440 Further decay modes from the family of $B \rightarrow D\bar{D}$ decays are $B^0 \rightarrow D^{*+} D^-$ and
 441 $B^0 \rightarrow D^{*+} D^{*-}$, which also enable a determination of ϕ_d^{eff} , but introduce further
 442 complications. For $B^0 \rightarrow D^{*+} D^-$ the final state is no CP eigenstate as it can be
 443 distinguished by the charge of the $D^{*\pm}$ meson. Thus, four CP observables are
 444 needed to describe CP violation. Furthermore, from an experimentalist's point
 445 of view the final state is not symmetrical in terms of the charges of pions and
 446 kaons and thus a detection asymmetry has to be taken into account. In the
 447 measurement of CP violation using $B^0 \rightarrow D^{*+} D^{*-}$ decays, like for $B_s^0 \rightarrow J/\psi \phi$
 448 decays, an angular-dependent analysis is required.

⁴⁴⁹ 4 The LHCb Experiment at the LHC

⁴⁵⁰ The Large Hadron Collider beauty (LHCb) experiment is one of the four large
⁴⁵¹ experiments run at CERN, the European Organization for Nuclear Research, in
⁴⁵² Geneva, Switzerland. The physics goal of the collaborators at LHCb is to test
⁴⁵³ the standard model of particle physics (SM) by performing indirect searches for
⁴⁵⁴ new physics with hadrons containing b or c quarks. In contrast to the direct
⁴⁵⁵ searches conducted by the multi-purpose experiments ATLAS (A Toroidal LHC
⁴⁵⁶ Apparatus) [73] and CMS (Compact Muon Solenoid) [74], CP -violating processes
⁴⁵⁷ and rare decay modes are precisely measured and the results are compared with the
⁴⁵⁸ predictions of the SM. This allows to investigate effects caused by heavy unknown
⁴⁵⁹ particles at energy scales far beyond what is accessible in direct searches. To do so,
⁴⁶⁰ it is mandatory to have a very high vertex, momentum and decay time resolution.
⁴⁶¹ The detector, with which this demanding task is accomplished, is described in
⁴⁶² Sec. 4.2. Beforehand, a description of the accelerator complex, which provides the
⁴⁶³ proton bunches to collide inside the detector, is given (see Sec. 4.1). Afterwards, the
⁴⁶⁴ trigger system (see Sec. 4.3) and the specific software (see Sec. 4.4) are described.

⁴⁶⁵ 4.1 The Large Hadron Collider

⁴⁶⁶ At CERN elementary particle physics is studied using the world's largest particle
⁴⁶⁷ accelerator, the Large Hadron Collider (LHC). Located at the French-Swiss border
⁴⁶⁸ area an accelerator complex including several linear and circular (pre-)accelerators
⁴⁶⁹ is operated (see Fig. 4.1). Before particles enter the 27 km long ring of the LHC
⁴⁷⁰ 50 to 175 m beneath ground, they have undergone accelerations to an energy of
⁴⁷¹ 450 GeV by the linear accelerator LINAC 2, the BOOSTER, the Proton Synchrotron
⁴⁷² (PS) and the Super Proton Synchrotron (SPS). A total of 1232 superconducting
⁴⁷³ dipole magnets keep the particle beams on the circular track, while quadrupole
⁴⁷⁴ magnets focus them. The magnets have to be operated at a temperature of
⁴⁷⁵ -271.3°C , which is achieved by a cooling system of liquid helium.

⁴⁷⁶ From 2010 on, mainly proton bunches and for shorter periods of time also bunches
⁴⁷⁷ of lead, on whose study ALICE (A Large Ion Collider Experiment) [76] is focused,
⁴⁷⁸ are injected. The protons are collided at centre-of-mass energies of 7 and 8 TeV in
⁴⁷⁹ 2011 and 2012 (Run I) and since 2015 (Run II) at 13 TeV.

4 The LHCb Experiment at the LHC

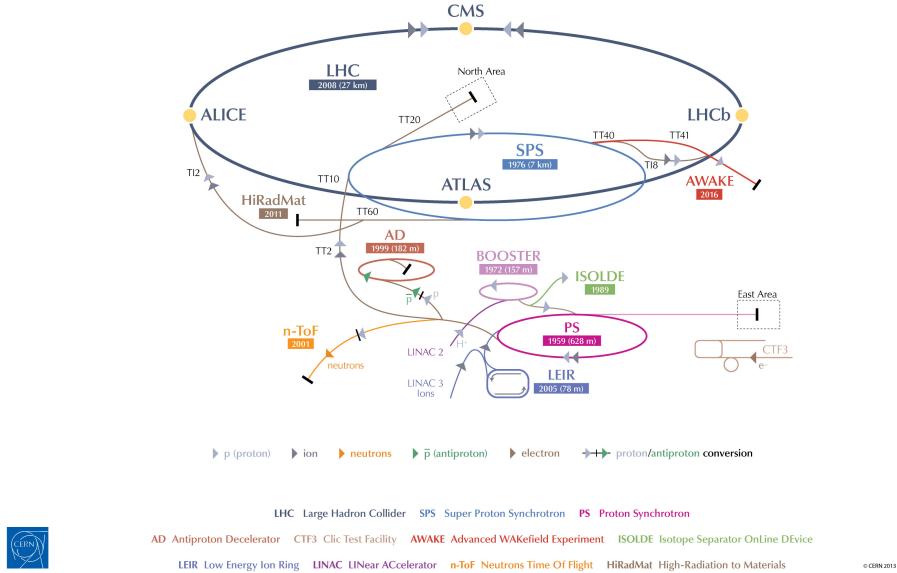


Figure 4.1: CERN’s accelerator complex [75].

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 the b and \bar{b} quarks are mainly 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. [77] and an overview of the performance is given in Ref. [78].

4.2.1 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 the curvature radius, information from tracking detector elements located upstream

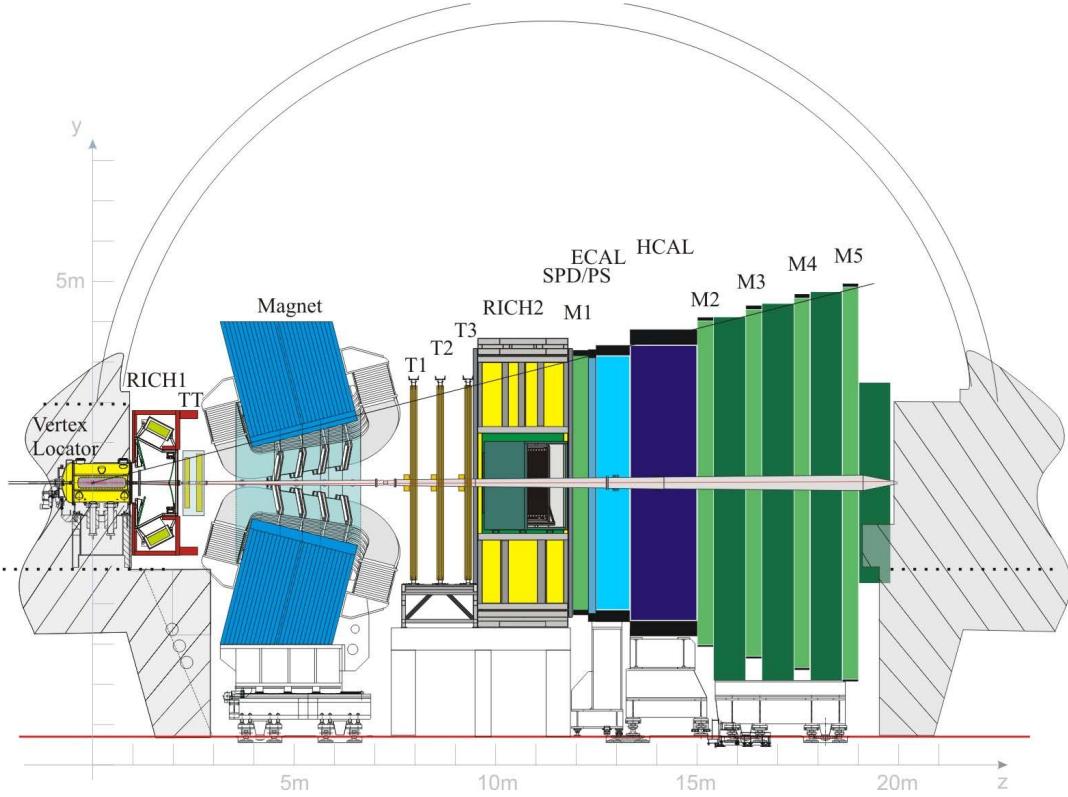


Figure 4.2: Schematic view on the LHCb detector [77].

and downstream of the magnet is needed. The pp interaction region is surrounded by a silicon-strip vertex locator (VELO) [79], 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 $4 \mu\text{m}$ 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 interaction called primary vertex (PV) and the displaced secondary decay vertices

4 The LHCb Experiment at the LHC

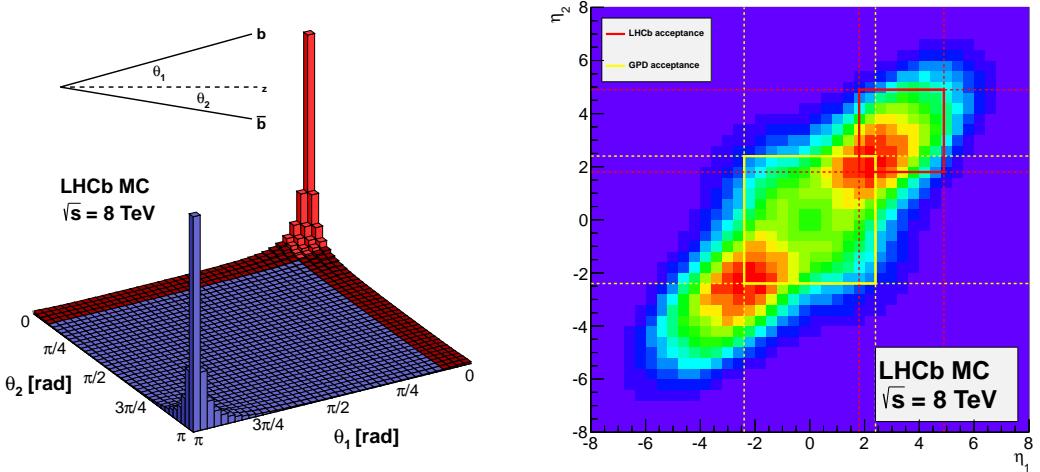


Figure 4.3: Correlation of angular acceptance (left) and pseudorapidity (right) of $b\bar{b}$ quark pairs. The frequency of produced $b\bar{b}$ quark pairs is indicated by the bin content in the plot showing the angular acceptance and by the colour code in the plot of the pseudorapidities, where purple 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.

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 $(15 + 29/p_T)$ μm (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 [80], it is built of four layers. While the first and last layer are arranged vertically, the inner layers are tilted by -5° 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 [78]. The straw drift tubes of the outer tracker are filled with a gas mixture of 70 % Ar, 28.5 % CO₂ and 1.5 % O₂. The addition of the oxygen reduces the ageing rate [81]. 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 [78].

Different track types are distinguished based on which detector components provide information on the trajectory of the track. The category with the best mass, momentum and vertex resolution is the long category. Long tracks originate

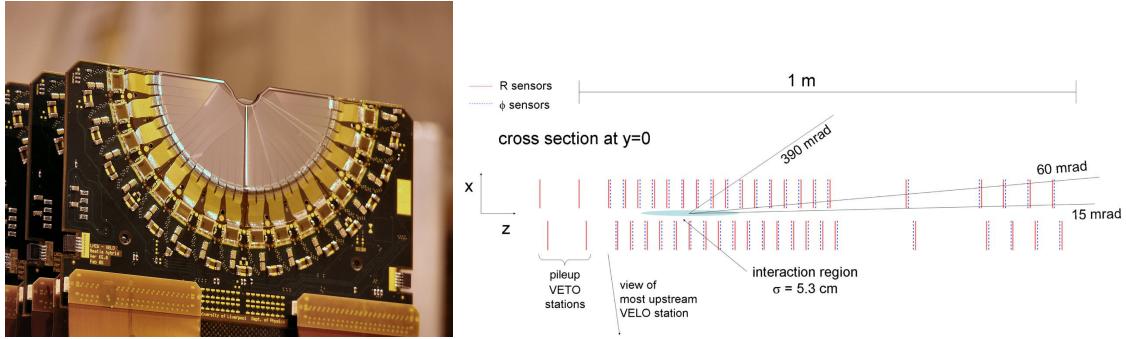


Figure 4.4: View on a single VELO half disk (left) and arrangement of all VELO modules [77].

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.

551 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 [82] 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 [83]. The calorimeters are built of alternating layers of metal and plastic. Polystyrene molecules in the plastic are excited by particle showers produced in the metal plates and produce ultraviolet light, whose amount is proportional

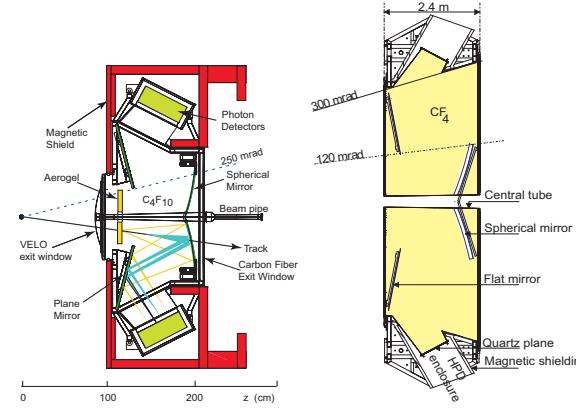


Figure 4.5: Side view schematic layout of RICH 1 (left) and top view schematic of RICH 2 (right) [77].

572 to the energy of the incident particle. The least interacting charged particles are
 573 muons, which are identified by five stations of multi-wire proportional chambers
 574 (MWPC) filled by a gas mixture of Ar, CO₂ and CF₄. Four of them are right at the
 575 end of the detector downstream of the calorimeters and one is located in between
 576 the second RICH and the calorimeter system. To stop the muons 80 cm thick layers
 577 of iron are put between the last four muon stations. Only muons with a momentum
 578 $p > 6 \text{ GeV}/c$ pass the whole detector. The detection of the muons is based on
 579 ionisation of the gas in the MWPCs. An electric field accelerates the ions and
 580 electrons. The emerging current is proportional to the energy of the muon.

581 4.3 The LHCb trigger system

582 Deliberately, the instantaneous luminosity at LHCb is reduced to $4 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$,
 583 which is significantly lower than at the other three experiments at the LHC. Though,
 584 the partial beam loss with time can be compensated by adjusting of the beam
 585 crossing so that a constant luminosity level can be kept throughout the whole
 586 fill [84]. Nevertheless, it is not possible to store the data of all visible proton-proton
 587 collisions. Instead, a two stage trigger system consisting of a hardware (L0) and a
 588 subsequent software level (HLT) is deployed. At the hardware trigger stage, which
 589 runs synchronously with the bunch crossing rate of 40 MHz, events are required to
 590 contain at least one muon with a high p_T (L0Muon), or two muons with a minimal
 591 product of their p_T (LODiMuon), or a hadron (LOHadron), a photon (LOPhoton) or
 592 an electron (LOElectron), which deposit high transverse energy in the calorimeters.
 593 Additionally, the number of allowed hits in the SPD is limited. These requirements
 594 reduce the data rate to 1 MHz, at which the full detector can be read out. The
 595 L0 signal efficiency varies a lot between muons and hadrons. While dimuon final
 596 states are triggered with more than 90 % efficiency, for fully hadronic final states
 597 like D^+D^- only around 60 % are reached [85, 86]. The high level trigger (HLT) is

598 a C++ application, which runs on an event filter farm of several thousand CPU
 599 nodes. It is again split into two stages. In the HLT1 basically the decisions of the
 600 L0 are checked. Due to the reduced data rate some more time is available. For
 601 all events the VELO tracks are reconstructed and a partial event reconstruction
 602 of all charged particles with $p_T > 500 \text{ MeV}/c$ in 2011 and $p_T > 300 \text{ MeV}/c$ in 2012
 603 is performed. This improves the momentum resolution and enables to calculate
 604 some invariant masses. The `Hlt1TrackMuon` trigger line requires a high p_T muon
 605 with a χ^2_{IP} with respect to any primary interaction greater than 16, where χ^2_{IP}
 606 is defined as the difference in χ^2 of a given PV reconstructed with and without
 607 the considered track. The `Hlt1DiMuonHighMass` trigger line accepts events if they
 608 contain two muons that form a good common vertex with an invariant mass above
 609 $2.7 \text{ GeV}/c^2$. In HLT2, a full reconstruction of the event is performed. Therefore,
 610 it is possible to further tighten the requirements applied in HLT1. Furthermore,
 611 for the `Hlt2DiMuonDetachedJPsi` trigger line a requirement on the flight distance
 612 is imposed. For hadrons, it is typically searched for two-, three- or four-track
 613 secondary vertices, which are identified via a multivariate algorithm [87].

614 The total output rate after all trigger stages has been increased from 3.5 kHz in
 615 2011 to 5 kHz in 2012 and 12.5 kHz in Run II.

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

620 A detailed description of the requirements imposed in the individual trigger lines
 621 is given in Ref. [85], while the latest performance numbers can be found in Ref. [86].

622 4.4 The LHCb software

623 4.4.1 Reconstruction

624 Two things need to be done in the reconstruction: tracks need to be find and
 625 combined and particle hypotheses have to be assigned to these tracks. The interface
 626 for the algorithms and tools is provided by the BRUNEL project [88], based on the
 627 GAUDI framework [89].

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

636 The particle identification (PID) of charged hadrons is performed via the allocation
 637 of rings in the RICH detectors to the tracks and calculating likelihoods for the

4 The LHCb Experiment at the LHC

638 different particle hypotheses. The calorimeters are used to identify electrons and
639 neutral pions, which decay into pairs of photons. One of the best signatures is given
640 by the muon system, which excludes respectively settles the muon hypothesis quite
641 reliable.

642 4.4.2 Stripping

643 Once tracking and PID are done the full decay chain can be fitted. However, the data
644 size after reconstruction is enormous. Right now, it is inevitable to use a centralized
645 selection called Stripping to handle it. In the DAVINCI framework [90] stripping
646 lines are defined, which basically are a set of requirements that describe certain
647 decay modes. Many selection steps can be shared between various stripping lines,
648 which saves a lot of computing time. For example, there are minimal requirements
649 for stable particles to start from. In the individual stripping lines these can then
650 be tightened. Only data selected by a stripping line can be analysed offline by the
651 users and data campaigns are usually only performed as often as once per year.
652 This makes the stripping so important. In the stripping the OfflineVertexFitter
653 (OVF) is used for the analysis of the $B^0 \rightarrow J/\psi K_s^0$ decays and the LoKiVertexFitter
654 (LVF) for the analysis of the $B^0 \rightarrow D^+ D^-$ decays. In order to correctly comprise
655 correlations and uncertainties on vertex positions, particle momenta, flight distances,
656 decay times, and invariant masses, the DecayTreeFitter (DTF) [91] can be used
657 in the reconstruction of decay chains. The decay time related observables in both
658 analyses covered in this thesis stem from a DTF fit where a constraint on the
659 production vertex of the B^0 mesons is applied using the knowledge about the
660 position of the primary vertex. The momenta and the invariant mass of the B^0
661 meson are determined with a DTF fit in which additionally the invariant masses of
662 the daughter hadrons are constrained to their nominal masses.

663 4.4.3 Monte Carlo simulation

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

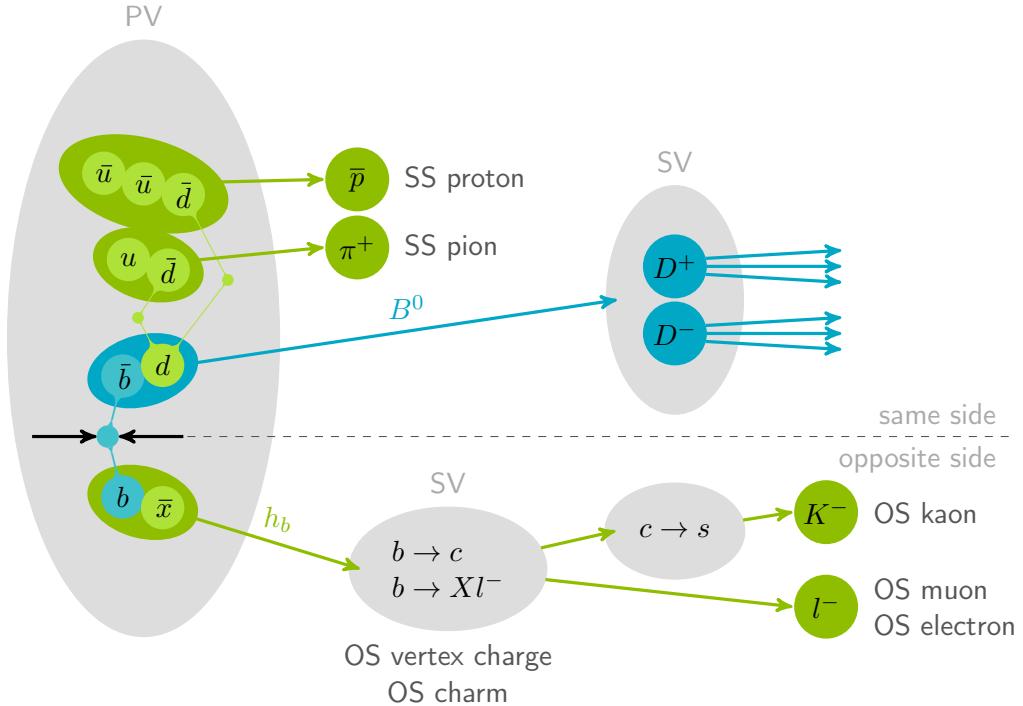
675 For LHCb the pp collisions are generated using PYTHIA [92] with a specific config-
676 uration [93]. The decays of the hadronic particles are simulated with EVTGEN [94].
677 To account for the radiation of photons in the final-state the package PHOTOS [95]

is used. The GEANT4 toolkit [96] is implemented as described in Ref. [97] to model the interaction of the generated particles with the detector material. The digitisation is realised using BOOLE [98]. The further processing is identical with the analysis of real collision data, starting with the trigger implemented in MOORE [99], the reconstruction done via BRUNEL [88] and the stripping using the DAVINCI package [90].

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 \bar{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 $B\bar{B}$ 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 \bar{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 predictions are based on the outcome of multivariate classifiers, which combine


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

719 kinematic and geometric properties of the tagging particle as well as information
 720 on the event.

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

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

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

734 While further information on the basic principles of LHCb's flavour-tagging
 735 algorithms can be found in Refs. [100, 101], a short description of the OS and SS
 736 taggers will be given in Secs. 4.5.1 and 4.5.2.

737 4.5.1 Opposite-side flavour tagging

738 Opposite-side flavour-tagging algorithms [102] infer the flavour of the signal B meson
 739 by studying the decay process of the second b hadron, which is produced from the
 740 same $b\bar{b}$ quark pair as the reconstructed signal B meson. Mistag probabilities are
 741 on the one hand introduced by selecting a wrong tagging particle and on the other
 742 hand intrinsically if the opposite-side b hadron is neutral and has already mixed at
 743 the time of decay.

744 In case of a semileptonic decay of the opposite-side b hadron the charges of the
 745 leptons are used by the OS electron (OSe) and OS muon ($\text{OS}\mu$) tagger to determine
 746 the flavour. These two taggers provide relatively good mistag estimates of around
 747 30 %, but have quite low tagging efficiencies of around 2 % (OSe) and 5 % ($\text{OS}\mu$)
 748 for charmonium respectively 3.5 % (OSe) and 8.5 % ($\text{OS}\mu$) for open charm modes.
 749 The efficiency for muons is a factor 2–3 higher than for electrons due to the better
 750 reconstruction and identification with the LHCb detector.

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

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

758 The recent OS charm tagger [103] reconstructs charm hadron candidates produced
 759 through $b \rightarrow c$ transitions of the opposite-side b hadron. The main contribution
 760 to the tagging power of the OS charm tagger comes from partially reconstructed
 761 charm hadrons in $K^-\pi^+X$ final states and from $D^0 \rightarrow K^-\pi^+$ decays. The overall
 762 tagging efficiency of the OS charm tagger is only 3–5 % with a mistag probability
 763 of around 35 %.

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

767 4.5.2 Same-side flavour tagging

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

4 The LHCb Experiment at the LHC

778 tagging pion. It has a tagging efficiency of around 15 % for $J/\psi X$ final states at
779 an average mistag probability of about 42 %. Recently, an improved SS π tagger
780 using a boosted decision tree (BDT) to select the tagging pion and based on the
781 very same principles also a SS p tagger have been developed [104, 105]. These two
782 flavour-tagging algorithms select completely disjoint tagging particles ensured by a
783 requirement on the distance log-likelihood (DLL) between the pion and the proton
784 hypothesis of the tagging particle. The tagging efficiency of the SS π tagger is very
785 high with 70–75 %. The SS p tagger provides non-zero tags for around 35 % of all
786 reconstructed signal candidates, of which about 80 % are also tagged by the SS π
787 (BDT). However, the high tagging efficiency comes along with rather large average
788 mistag probabilities of around 45 %. The response of the two BDT-based SS taggers
789 is combined into a common SS response.

790 5 Data Analysis Tools and Methods

791 5.1 Maximum likelihood method

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

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

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

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

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

808 5.2 Selection

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

5 Data Analysis Tools and Methods

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 [106] or scikit-learn [107], provide implementations of these algorithms.

A simple multivariate classifier is a decision tree [108], 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 [109], $p \cdot (1 - p)$ 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.

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 [110].

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) [111]. One possibility how the boosting can be realised is the AdaBoost method [112]. Events that are misidentified in the previous tree are weighted by

$$\alpha = \frac{1 - \varepsilon}{\varepsilon}, \quad (5.3)$$

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, $\alpha \rightarrow \alpha^\beta$. 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,

$$y = \frac{1}{N_{\text{trees}}} \sum_i^{N_{\text{trees}}} \ln(\alpha_i) h_i. \quad (5.4)$$

This means that an event is classified more signal-like the higher the BDT output classifier. Another boosting algorithm is the GradientBoost method [113]. 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) [114]. 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.

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 *sPlot* technique [115] the weights are calculated from yields of an extended maximum likelihood fit (see Sec. 5.1). Based on the fit results sWeights according to

$${}^s \mathcal{P}_n(x_i) = \frac{\sum_{j=1}^{N_s} \mathbf{V}_{nj} f_j(x_i)}{\sum_{k=1}^{N_s} N_k f_k(x_i)} \quad (5.5)$$

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 other floating parameters are fixed. The sWeights fulfil the condition that their sum over

5 Data Analysis Tools and Methods

890 one category returns the corresponding fitted yield. The sWeights can be applied to
 891 other observables if they are uncorrelated with the one used to obtain the sWeights.
 892 In sweighted histograms the uncertainty on the bin content of bin i is given by

$$\sigma(i) = \sqrt{\sum_{e \in i} ({}_s\mathcal{P}_n)^2}. \quad (5.6)$$

893 5.2.3 Figures of merit

894 One of the main questions when performing a selection is, how the requirements, be
 895 it a sequence of cuts or the classifier of a multivariate method, should be optimized.
 896 In an ideal world, one would find a selection that keeps all signal candidates and
 897 removes all background contributions. But this is unrealistic. Instead, suitable
 898 figures of merit have to be defined, whose optimum should correspond with the
 899 ideal selection. In a measurement of CP violation the goal is to obtain the optimal
 900 precision. Therefore, the sensitivity on the CP observables itself seems to be the
 901 best figure of merit. However, there are some caveats. On the one hand there is
 902 usually more than one observable describing the CP violation and a strategy needs
 903 to be found how the sensitivities of the different CP observables can be combined
 904 into a single figure of merit. On the other hand the absolute uncertainty might
 905 depend on the central value and thus small values for the CP observables could
 906 be preferred. Finally, the full time-dependent fit has to be performed, which is
 907 often very complex and takes a long time until it converges. For all these reasons,
 908 alternative figure of merits are developed. A very simple one is the pure signal
 909 efficiency ϵ_S . Under the assumptions that a high signal yield is more important
 910 than a low background contamination and that the selection requirements are
 911 effectively suppressing background contributions, a possible selection strategy is
 912 to judge requirements only by their signal efficiency. In the $B^0 \rightarrow J/\psi K_s^0$ analysis
 913 this approach is chosen using the requirement that the individual cuts have to
 914 be at least 99 % signal efficient. Including the background yield to the definition
 915 of the figure of merit should make it easier to find the optimal cut point, as the
 916 background contamination partly influences the achievable sensitivity. There are
 917 several possibilities how the signal yield S and the background yield B can be
 918 combined: The figure of merit

$$Q_1 = \frac{S}{S + B}, \quad (5.7)$$

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

$$Q_2 = \frac{S}{\sqrt{S + B}}, \quad (5.8)$$

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:

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

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 [116]

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

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 $t = \frac{\pi}{2\Delta m_d} \approx 3$ 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:

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

with $d_i = +1$ for B^0 and $d_i = -1$ for \bar{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 $f_i(t)$:

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

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 generalized figure of merit for decay-time-dependent measurements of CP violation is derived in Ref. [117].

950 5.3 Spline interpolation

951 In many cases phenomenological models are an efficient way of describing shapes,
 952 *e.g.* when parametrising acceptances, which are typically influenced by more effects
 953 than could realistically be analysed separately. Interpolating cubic splines, which
 954 are piecewise defined third order polynomials, are an useful implementation [118].
 955 They are parametrised by a set of knots and coefficients at these positions and
 956 can be written as the sum over base splines. The first and second derivatives are
 957 continuous throughout the domain. The choice of the number and positions of the
 958 knots determines how accurate the given shape can be described.

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

965 5.4 Bootstrapping method

966 The bootstrapping method (see *e.g.* Ref. [119]) is a frequentist model-independent
 967 approach to estimate confidence level intervals. Toy data samples are produced by
 968 drawing events from the nominal data sample, with replacement, until the statistics
 969 matches the number of candidates of the nominal data sample. This means that the
 970 same event can be drawn multiple times. The bootstrapping method can easily be
 971 used for multiple dimensions. Thus, it maintains correlations between the observ-
 972 ables without any assumptions or model dependencies. Therefore, bootstrapped
 973 samples serve as perfect representations of the original data. Calculations or fits
 974 that are originally done once on the nominal data sample can be repeated several
 975 times. So, reliable estimates for standard deviations or confidence level intervals
 976 can be derived.

977 5.5 Blinding

978 Especially when performing precision measurements or searching for rare decays
 979 it is advisable to blind the results throughout the analysis and even as analyst
 980 only look at them after some thorough (ideally external) review process. Blinding
 981 means that the central value or the final result is unknown to all involved people.
 982 This procedure avoids the experimenter's bias, *i.e.* the unintended biasing of a
 983 result towards a known or expected value or towards a (subconsciously) desired
 984 observation. The blinding transformation that has been applied in the measurements
 985 of CP violation in $B^0 \rightarrow J/\psi K_s^0$ and $B^0 \rightarrow D^+ D^-$ decays is adding a hidden offset
 986 to the fitted CP parameters S_f and C_f using the `RooUnblindUniform` method of
 987 ROOFIT's `RooBlindTools` [120]. 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. [121].

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

$$\mathcal{D} = e^{-\frac{\Delta m_d^2 \sigma^2}{2}}. \quad (5.13)$$

This formula is the special case for a Gaussian resolution model with width σ . The general formula is derived in Ref. [122]. 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 $\Delta m_d = (0.5064 \pm 0.0019) \text{ } \hbar \text{ ps}^{-1}$ [50] 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

$$\omega(\eta) = p_0 + p_1(\eta - \langle \eta \rangle) \quad (5.14)$$

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 be dependent 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 \bar{B}^0 are implemented:

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

5 Data Analysis Tools and Methods

1018 Equivalently, the calibration parameters for B^0 and \bar{B}^0 can be related through their
1019 mean and their difference:

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

1020 The asymmetry of the mistags can then be written as

$$\Delta\omega(\eta) = \Delta p_0 + \Delta p_1(\eta - \langle\eta\rangle). \quad (5.17)$$

1021 For the flavour-tagging calibration it is beneficial to use flavour-specific decay
1022 channels that are kinematically similar to the signal channel. Additionally, the
1023 selection between the two channels should be as close as possible. This allows
1024 to transfer the calibration results from the control to the signal channel without
1025 assigning large systematic uncertainties. On the other hand, the control channel
1026 should ideally be a mode with high statistics to reduce the statistical uncertainties on
1027 the flavour-tagging calibration parameters. A good compromise between these two
1028 requirements is found by choosing $B^0 \rightarrow D_s^+ D^-$ as calibration mode for $B^0 \rightarrow D^+ D^-$,
1029 while for $B^0 \rightarrow J/\psi K_s^0$ the OS tagging combination and the cut-based SS π tagging
1030 algorithm are calibrated with $B^+ \rightarrow J/\psi K^+$ and $B^0 \rightarrow J/\psi K^{*0}$ decays, respectively.

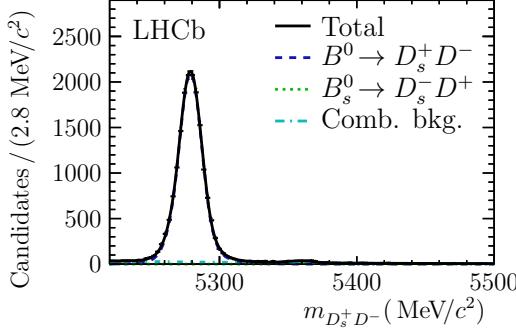
1031 5.7.1 Calibration using $B^0 \rightarrow D_s^+ D^-$

1032 The decay of B^0 mesons via the decay mode $B^0 \rightarrow D_s^+ D^-$ proceeds flavour specific
1033 as the charge of the D_s^+ meson unambiguously determines the flavour of the
1034 decaying B^0 meson. When reconstructing the D_s^+ meson via $D_s^+ \rightarrow K^- K^+ \pi^+$
1035 and the D^+ meson via $D^+ \rightarrow K^- \pi^+ \pi^+$, a very similar selection to the one for
1036 $B^0 \rightarrow D^+ D^-$ (described in Sec. 7.1) can be applied. The only differences are that
1037 the invariant $m_{KK\pi}$ mass is required to lie within ± 25 MeV/c² of the known D_s^+
1038 mass [123] and that the vetoes against misidentified backgrounds are not applied. A
1039 maximum likelihood fit to the invariant $m_{D_s^+ D^-}$ mass distribution is performed to
1040 statistically subtract the remaining background via the *sPlot* technique. Apart from
1041 the $B^0 \rightarrow D_s^+ D^-$ component, which is parametrised by the sum of two Crystal Ball
1042 functions (common mean but different widths, and tail parameters taken from MC),
1043 the fit model includes components for $B_s^0 \rightarrow D_s^- D^+$ decays and for combinatorial
1044 background. The total $B^0 \rightarrow D_s^+ D^-$ yield is found to be $16\,736 \pm 134$ at a quite
1045 low background level as can be seen in Fig. 5.1.

1046 The $B^0 - \bar{B}^0$ mixing prevents to infer the flavour of the B^0 meson at production.
1047 Therefore, a mixing analysis is performed to determine the true mistag probability
1048 ω from the amplitude of the mixing asymmetry

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

1049 where N_{unmixed} is the number of $B^0 \rightarrow D_s^+ D^-$ decays with a final state that does
1050 correspond to the flavour tag, and N_{mixed} the number with a final state that does

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

not. Other experimentally induced effects are corrected for, like the detection asymmetry \mathcal{A}_{det} , the production asymmetry A_P and the flavour-specific asymmetry a_{sl}^d . 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 $B^0 \rightarrow D_s^+ D^-$. 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.

Thanks to the improved flavour-tagging algorithms and the kinematic properties of the selected $B^0 \rightarrow D^+ D^-$ decays, which for example have on average a higher 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 it's the highest effective tagging efficiency in tagged CP violation measurements at LHCb.

5 Data Analysis Tools and Methods

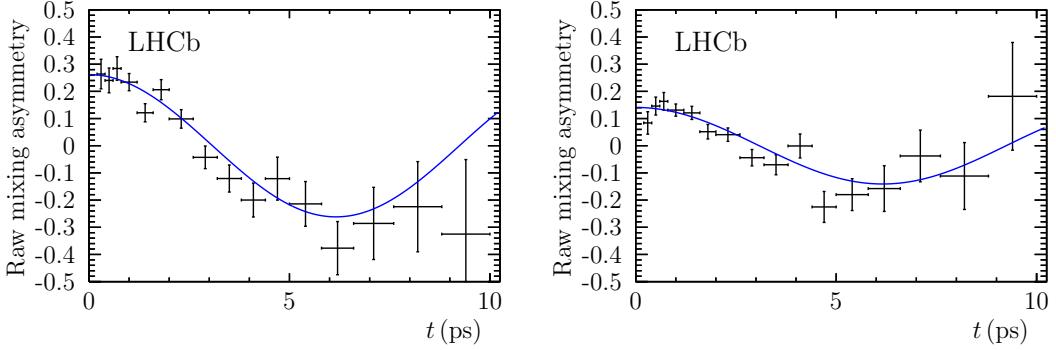


Figure 5.2: Raw mixing asymmetry as a function of the B^0 decay time for events tagged by (left) the OS and (right) the SS tagging combination. The solid line represents the PDF projection.

1070 The flavour-tagging calibration using $B^0 \rightarrow D_s^+ D^-$ decays has been provided by
1071 collaborators from Milano.

1072 5.7.2 Calibration using $J/\psi X$ channels

1073 For the charged $B^+ \rightarrow J/\psi K^+$ decay, which is used to determine the calibration of
1074 the OS tagging combination for the analysis of $B^0 \rightarrow J/\psi K_s^0$ decays, a comparison
1075 of the charge of the kaon with the tag decision directly tells if the tag decision is
1076 correct or not. Binning the sample in terms of the mistag estimate a χ^2 fit using
1077 Eq. (5.14) is performed to the (η, ω) pairs and reveals the following results for the
1078 calibration parameters:

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

1079 Repeating the same fit with a split of the sample into the initial flavours gives
1080 access to the asymmetry parameters Δp_0 and Δp_1 , which are determined to be

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

1081 A cross-check of the calibration in a control sample of $B^0 \rightarrow J/\psi K^{*0}$ decays confirms
1082 the validity of transferring the calibration from B^+ to B^0 decays.

1083 Despite the advantages of $B^+ \rightarrow J/\psi K^+$ as control channel (charged decay mode,
1084 very high statistics), for the calibration of the cut-based SS π tagging algorithm
1085 $B^0 \rightarrow J/\psi K^{*0}$ decays are used because differences in the composition of the fragmen-
1086 tation products in the B^+ and B^0 hadronisation are expected. Like for $B^0 \rightarrow D_s^+ D^-$
1087 a time-dependent mixing analysis is needed. Here, a two dimensional fit to both the

1088 reconstructed decay time and mass distributions is performed. From a simultaneous
1089 fit in five evenly filled bins of the mistag estimate η the calibration parameters for
1090 the SS π tagging algorithm are determined to be

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

1091 The systematic uncertainties for both calibrations cover two different types, one
1092 for intrinsic uncertainties and one for the kinematic differences between the control
1093 mode and the signal decay $B^0 \rightarrow J/\psi K_s^0$. The effect of these two sources is of the
1094 same order.

1095 The combined effective tagging efficiency is $(3.02 \pm 0.05)\%$, which is composed
1096 of a tagging efficiency of $\varepsilon_{\text{tag}} = (36.54 \pm 0.14)\%$ and an effective mistag probability
1097 of $\omega_{\text{eff}} = (35.62 \pm 0.12)\%$. The major contribution comes from the OS tagging
1098 combination, which has an inclusive tagging power of $(2.63 \pm 0.04)\%$. The cut-based
1099 SS π tagging algorithm adds $(0.376 \pm 0.024)\%$.

1100 **6 Measurement of CP Violation in 1101 $B^0 \rightarrow J/\psi K_s^0$ Decays**

1102 The measurement of CP violation in $B^0 \rightarrow J/\psi K_s^0$ decays [124] is performed on
1103 LHCb's full Run I data sample, which corresponds to an integrated luminosity of
1104 3 fb^{-1} . The analysis is based on a collaborative work of Christophe Cauet [125] and
1105 myself. Therefore, my main contributions, namely the decay time resolution study
1106 (see Sec. 6.3), the analysis of the backgrounds (see Sec. 6.4), the fitter development,
1107 and the studies of systematic uncertainties (see Sec. 6.6), are described in more detail,
1108 while the other parts are only briefly summarised to complete the picture. The initial
1109 flavour of the B^0 mesons is determined with the OS combination excluding the OS
1110 charm flavour-tagging algorithm and the cut-based SS π tagger. The calibration of
1111 these taggers is described in Sec. 5.7.2.

1112 **6.1 Data Preparation**

1113 Candidate $B^0 \rightarrow J/\psi K_s^0$ decays are reconstructed through the subsequent decays
1114 $J/\psi \rightarrow \mu^+ \mu^-$ and $K_s^0 \rightarrow \pi^+ \pi^-$. Only combinations of two long (LL) or two down-
1115 stream tracks (DD) are considered for the pions. All events must have passed either
1116 the L0Muon or the L0DiMuon trigger line. This is implicitly given by the requirement
1117 that a positive J/ψ TOS decision of the DiMuonHighMass or the TrackMuon trigger
1118 line in the Hlt1 stage and of the DiMuonDetachedJPsi trigger line in the Hlt2 stage
1119 exists. These trigger requirements have a total signal efficiency of about 85 %.

1120 In the stripping loose requirements on the quality of the J/ψ , the K_s^0 and the B^0
1121 vertex are applied. The invariant masses of the $\mu^+ \mu^-$ and the $\pi^+ \pi^-$ combination are
1122 required to be roughly consistent with the known J/ψ and K_s^0 masses, respectively.
1123 Moreover, the K_s^0 candidates must have a significant decay length and some further
1124 kinematic requirements need to be fulfilled. Despite these rather loose selection
1125 requirements, the signal to background ratio is already quite high. Therefore, in
1126 the offline selection almost all cuts are tuned to have a high signal efficiency. The
1127 probability of the muon and pion tracks to be ghost tracks (see Sec. 4.4.1) is reduced
1128 to 20 and 30 %, respectively. The mass window around the J/ψ meson is tightened
1129 to correspond to five standard deviations. For the K_s^0 meson the mass window
1130 is adapted to the track type of the pions. It contains four and eight standard
1131 deviations for the long and the downstream candidates, respectively. A specific
1132 treatment against misidentified $\Lambda_b^0 \rightarrow J/\psi \Lambda$ and $B^0 \rightarrow J/\psi K^{*0}$ decays is performed.
1133 The former are rejected by tighter PID requirements on the pions if the invariant

6 Measurement of CP Violation in $B^0 \rightarrow J/\psi K_s^0$ Decays

mass under a πp mass hypothesis is compatible with the Λ mass [123]. The latter are suppressed with a cut on the K_s^0 decay time in units of its uncertainty.

The invariant B^0 mass, which is restricted to candidates inside 5230–5330 MeV/c², is obtained from a fit to the whole decay chain (DTF) with constraints on the J/ψ and K_s^0 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 DTF 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 $B^0 \rightarrow J/\psi K_s^0$ candidates partially bias the decay time distribution. The sample is divided into an *almost unbiased* subset (AU), defined by

```
AU ≡ Hlt1DiMuonHighMass && Hlt2DiMuonDetachedJPsi
```

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

```
EB ≡ (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

$$\tilde{\tau} = \frac{\tau}{1 + \beta_\tau \tau}. \quad (6.1)$$

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

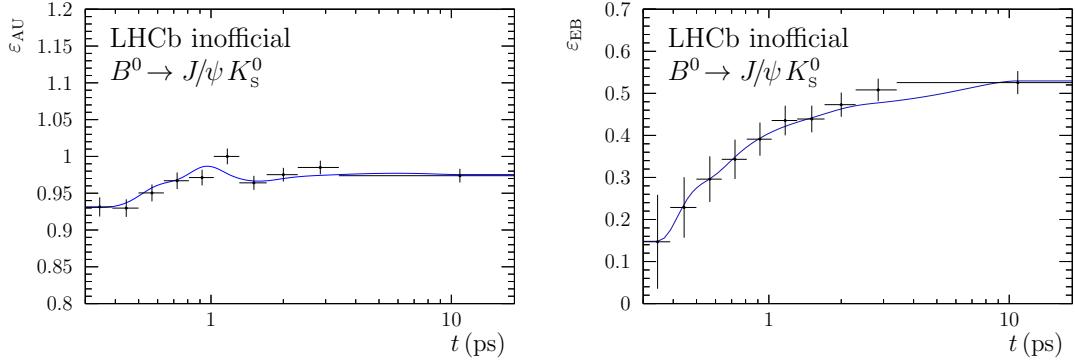


Figure 6.1: Histograms of the trigger acceptance for the almost unbiased (left) and the exclusively biased (right) sample. The blue curve shows the fitted acceptance using splines.

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

1171 Table 6.1. It is apparent that the correction factors are roughly two times larger in
 1172 2012 than in 2011 and four to five times larger for long tracks than for downstream
 1173 tracks. While the former can be explained with changes in the reconstruction
 1174 software, the latter comparison clearly shows that the Velo is the main reason for
 1175 this acceptance effect.

1176 6.3 Decay time resolution

1177 The most obvious effect of the decay time resolution are candidates that are
 1178 reconstructed with negative decay times. However, these are ideal candidates to
 1179 determine the decay time resolution. An unbinned likelihood fit to the decay
 1180 time distribution of an unbiased $B^0 \rightarrow J/\psi K_s^0$ sample with true J/ψ candidates
 1181 (selected by a fit of the invariant $m_{\mu^+\mu^-}$ mass distribution) is performed. The fit
 1182 model consists of a component for the prompt peak around 0 ps, *i.e.* the decay
 1183 time resolution model, a component to describe events where a wrong PV has
 1184 been associated and therefore a large difference between true and reconstructed
 1185 decay time occurs, and long-lived components. The long-lived components are
 1186 parametrised with two exponentials, which have different pseudo lifetimes and which
 1187 are themselves convolved with the decay time resolution model. The decay time
 1188 resolution depends on characteristics of the event. The DTF provides predictions for
 1189 the per-event decay time resolution σ_t , which can be used as the width of a Gaussian

6 Measurement of CP Violation in $B^0 \rightarrow J/\psi K_s^0$ Decays

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:

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

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

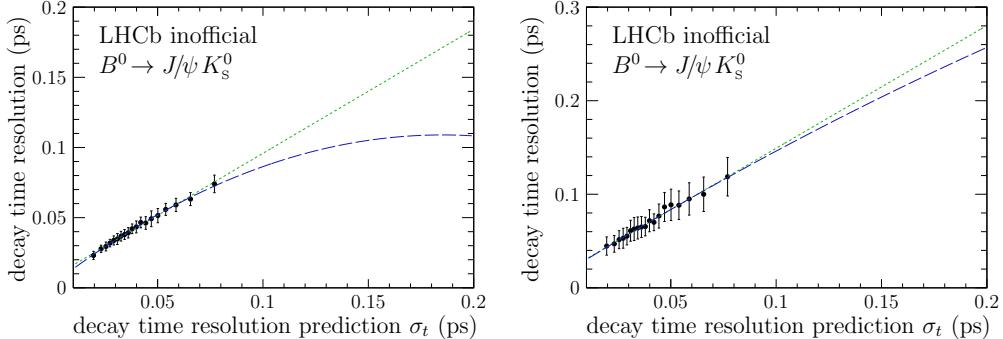


Figure 6.2: Fit of a linear (green short-dashed) and a quadratic calibration function (blue long-dashed) to the narrower (left) and to the wider width (right) of the downstream sample.

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 $B^0 \rightarrow J/\psi K_s^0$ 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_s^0 decay might actually be kaons or protons that have been mis-identified. This would lead to background contributions from $B^0 \rightarrow J/\psi K^{*0}$ and $\Lambda_b^0 \rightarrow J/\psi \Lambda$. To analyse the $p \rightarrow \pi$ mis-ID 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 $M_\Lambda = 1115.683 \text{ MeV}/c^2$ [123] 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 $\Lambda_b^0 \rightarrow J/\psi \Lambda$ 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 $K \rightarrow \pi$ mis-ID the broad width of the K^{*0} does not allow an analogous approach. But studies on $B^0 \rightarrow J/\psi K^{*0}$ MC show that the expected contribution is even lower than for $\Lambda_b^0 \rightarrow J/\psi \Lambda$. The main reason is the short lifetime of the K^{*0} , which is exploited by the lifetime significance cut on the K_s^0 . 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 *sPlot* technique provides a possibility to study the tagging-dependent distributions of the combinatorial background. First of all, the time-integrated asymmetry

$$\mathcal{A}_{\text{bkg}}^{\text{int}} = \frac{N_{\text{bkg}}^{\bar{B}^0} - N_{\text{bkg}}^{B^0}}{N_{\text{bkg}}^{\bar{B}^0} + N_{\text{bkg}}^{B^0}} \quad (6.3)$$

is calculated for both track type categories and separately for the OS tagging combination and the SS π 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 weighted 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 SS π tagged events.

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

1239

6 Measurement of CP Violation in $B^0 \rightarrow J/\psi K_s^0$ Decays

1240 But even decay-time-integrated asymmetries compatible with zero do not exclude
 1241 decay-time-dependent asymmetries. The latter are investigated by plotting the
 1242 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

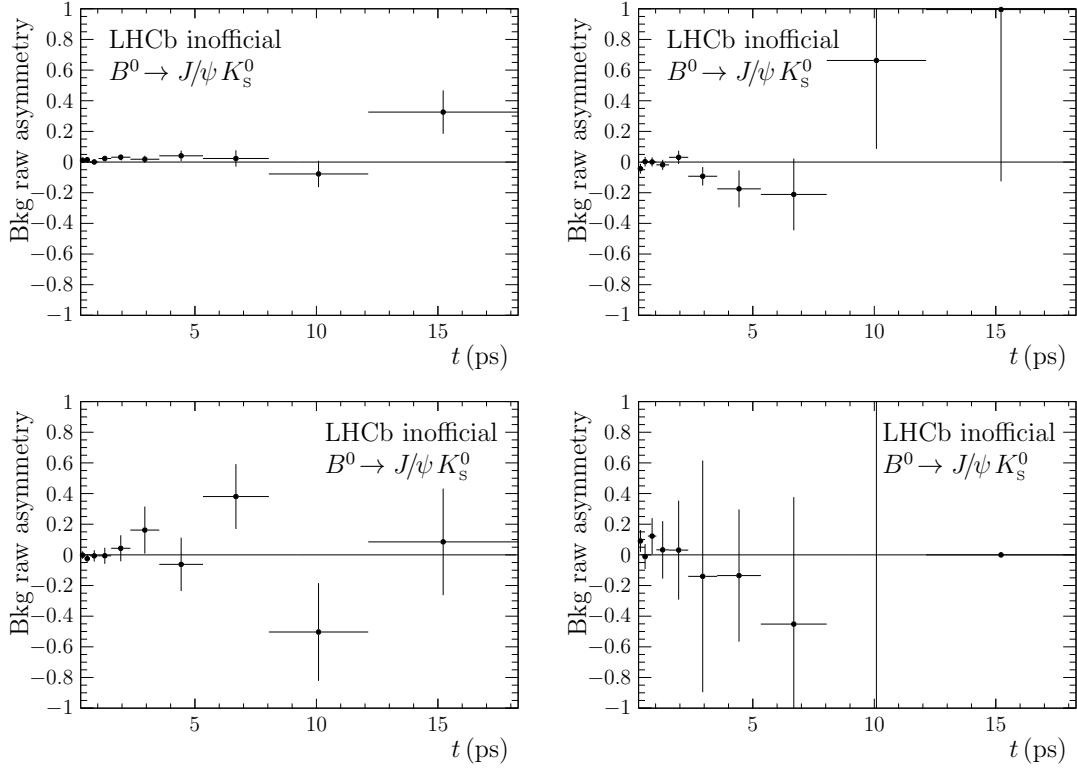


Figure 6.3: Raw background asymmetry in sweighted data with logarithmic binning. DD in top, LL in bottom plots. OS tagged candidates on the left, SS π tagged candidates on the right side.

1243 there is a significant oscillation. Thus, χ^2 -tests against the null-hypothesis, *i.e.* a
 1244 flat line at zero, are performed and the corresponding p -values are calculated (see
 1245 Table 6.2). None of the p -values is very low, so the deviations from zero can be
 1246 explained with statistical fluctuations. The same procedure (mass fit \rightarrow sWeights
 1247 \rightarrow histograms of \mathcal{A}_{bkg}) is performed with cocktail MC consisting of signal MC and
 1248 background Toy MC. Here, no asymmetry is generated for the background and
 1249 the resulting p -values are very similar to the ones for the nominal data sample.
 1250 Finally, an unbinned likelihood fit is performed to the sweighted background decay
 1251 time distribution using two/three exponentials with different pseudo-lifetimes and
 1252 allowing for tagging-dependent asymmetries in the PDF. All asymmetry parameters
 1253 are compatible with zero at a significance of two standard deviations.

6.5 Nominal fit

A multi-dimensional fit to the distributions of the reconstructed mass m , the decay time t and its error prediction σ_t , the OS and SS π tags and mistag probabilities is performed, mainly to extract the CP observables $S_{J/\psi K_S^0}$ and $C_{J/\psi K_S^0}$. 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 [126] 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

$$\frac{d\Gamma(t, d)}{dt} \propto e^{-t/\tau} \left(1 - d S_f \sin(\Delta m t) + d C_f \cos(\Delta m t) \right), \quad (6.4)$$

with $d = +1$ for B^0 and $d = -1$ for \bar{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 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 weighted 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 SS π 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 [123] and the flavour-tagging calibration parameters (Sec. 5.7.2).

To stabilize 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.

6 Measurement of CP Violation in $B^0 \rightarrow J/\psi K_s^0$ Decays

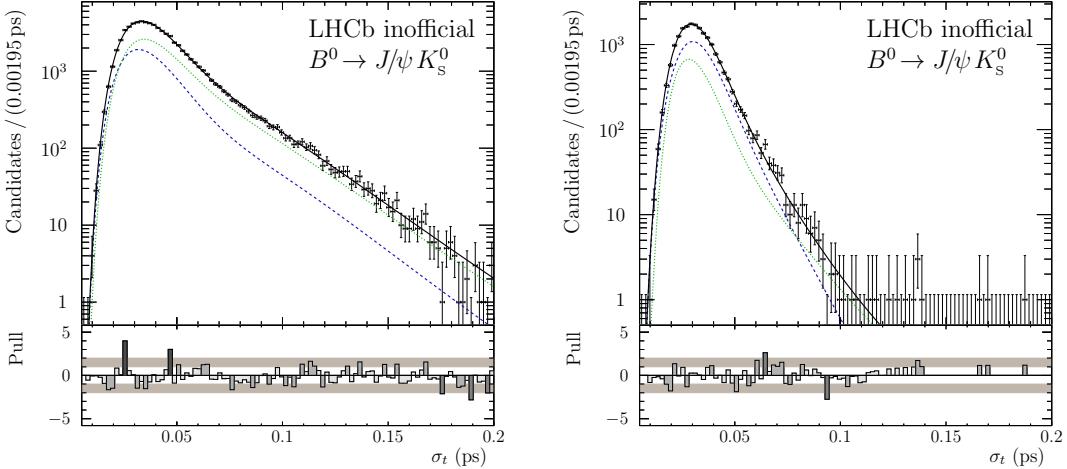


Figure 6.4: Decay-time error distributions of downstream (left) and long track (right). The solid black line shows the fit projection, while the blue dashed (green dotted) line shows the signal (background) component.

1294 From the 41 560 flavour-tagged $B^0 \rightarrow J/\psi K_s^0$ decays the CP observables are
 1295 determined to be

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

1296 with a statistical correlation of 0.483. In these results corrections of +0.002 for
 1297 $S_{J/\psi K_s^0}$ and -0.005 for $C_{J/\psi K_s^0}$ are included, which account for CP violation in
 1298 $K^0 - \bar{K}^0$ mixing and different nuclear cross-sections in material between K^0 and
 1299 \bar{K}^0 [127].

1300 The distributions of the invariant mass and the decay time are depicted in Fig. 6.5.

1301 6.6 Studies of systematic effects

1302 To check if and how various effects systematically influence the measurement of the
 1303 CP violation parameters, the likelihood fit is performed

- 1304 • with a second independent fitter [125],
- 1305 • without a parametrisation of the background using sWeights extracted from
 1306 the *sPlot* technique,
- 1307 • on subsamples split by the K_s^0 track type, the trigger requirements, the tagging
 1308 algorithms, the magnet polarity, and the year of data-taking.

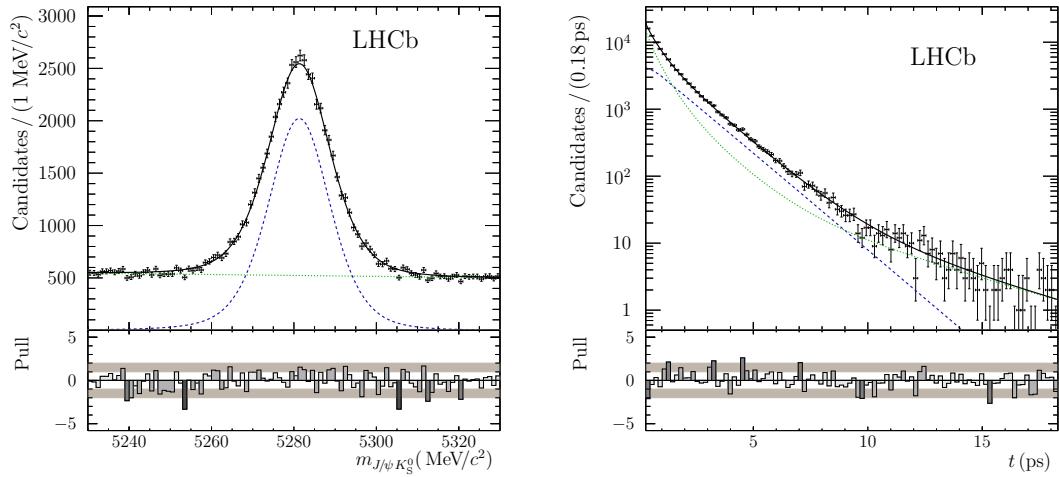


Figure 6.5: Invariant mass distribution (left) and decay time distribution (right). The solid black line shows the fit projection, while the blue dashed (green dotted) line shows the signal (background) component.

1309 All results show good agreement with the nominal results. Additionally, the results
 1310 from a pure time-dependent and from a pure time-integrated fit are compatible
 1311 with each other and with the nominal fit, which comprises both effects.

1312 Systematic uncertainties from several effects, especially from possible mismodelling of PDFs and from systematic uncertainties on external input parameters,
 1313 are considered. Pseudoexperiments are generated using PDFs that contain a slight
 1314 modification compared to the nominal PDF, which is used for the subsequent fit of
 1315 the samples. Whenever the mean of the pull distribution exceeds zero by more than
 1316 0.032, a systematic uncertainty is assigned. Pull distributions show the difference
 1317 between the individual fit result and the generation value for the CP observable
 1318 normalised by the fit uncertainty. The value of the criterion is defined by one
 1319 standard deviation according to the statistics of 1000 pseudoexperiments performed
 1320 for each study. The size of the systematic uncertainty is taken from the residual
 1321 distributions, which show the same as the pull distributions apart from not being
 1322 normalised.

1324 The possible tagging asymmetry of the background contribution is treated as a
 1325 source for a systematic uncertainty on the CP observables using the asymmetry
 1326 parameters determined from a weighted fit to the background decay time distribution
 1327 (see last part of Sec. 6.4). It is found to account for 83 % of the systematic
 1328 uncertainty on $S_{J/\psi K_S^0}$ and 8 % for $C_{J/\psi K_S^0}$. The systematic uncertainties on the
 1329 flavour-tagging calibration parameters (see Sec. 5.7.2) are transferred to systematic
 1330 uncertainties on $S_{J/\psi K_S^0}$ and $C_{J/\psi K_S^0}$ by shifting all calibration parameters related
 1331 to p_0 upwards by one systematic uncertainty and those related to p_1 downwards
 1332 in the generation of the pseudoexperiments while setting them to their nominal

6 Measurement of CP Violation in $B^0 \rightarrow J/\psi K_S^0$ Decays

values in the fit. This yields 9 % of the systematic uncertainty on $S_{J/\psi K_S^0}$, and 21 % for $C_{J/\psi K_S^0}$. The assumption $\Delta\Gamma_d = 0$ is responsible for 6 % of the systematic uncertainty on $S_{J/\psi K_S^0}$. It is determined generating the pseudoexperiments setting $\Delta\Gamma_d = 0.007 \text{ ps}^{-1}$, the current experimental uncertainty [123]. The largest contribution (42 %) to the systematic uncertainty on $C_{J/\psi K_S^0}$ 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 $S_{J/\psi K_S^0}$ of ± 0.020 is more than 40 % smaller than the statistical uncertainty. For $C_{J/\psi K_S^0}$ the total uncertainty even increases by only 1 % through the systematic uncertainty.

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

Origin	$\sigma_{S_{J/\psi K_S^0}}$	$\sigma_{C_{J/\psi K_S^0}}$
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

¹³⁴⁸ 7 Measurement of CP Violation in ¹³⁴⁹ $B^0 \rightarrow D^+ D^-$ Decays

¹³⁵⁰ In this chapter the analysis of $B^0 \rightarrow D^+ D^-$ decays with the goal to determine the
¹³⁵¹ observables $S_{D^+ D^-}$ and $C_{D^+ D^-}$, which describe the CP violation in this decay mode,
¹³⁵² is presented [128]. After a description of the selection (see Sec. 7.1) the fit of the
¹³⁵³ invariant mass distribution, performed to extract signal sWeights, is described (see
¹³⁵⁴ Sec. 7.2). This is followed by a summary of the decay time fit (see Sec. 7.3). The
¹³⁵⁵ chapter is concluded with a presentation of the studies on systematic uncertainties
¹³⁵⁶ (see Sec. 7.4). The new OS combination including the OS charm tagger and the SS
¹³⁵⁷ combination of SS π and SS p tagger are used to determine the flavour tag of the
¹³⁵⁸ B^0 mesons. The calibration of these flavour-tagging algorithms using $B^0 \rightarrow D_s^+ D^-$
¹³⁵⁹ decays (see Sec. 5.7.1) is the only part not performed by myself but by collaborators
¹³⁶⁰ from Milano.

¹³⁶¹ 7.1 Selection

¹³⁶² The amount of background in $B^0 \rightarrow D^+ D^-$ is too high to perform a significant
¹³⁶³ measurement of CP violation without any selection. The selection is divided
¹³⁶⁴ into three parts: a preselection with many high signal efficiency requirements, a
¹³⁶⁵ dedicated treatment of mis-identified backgrounds and a multivariate analysis to
¹³⁶⁶ further reduce combinatorial background.

¹³⁶⁷ 7.1.1 Preselection

¹³⁶⁸ Only events that have been triggered by a topological trigger line or by the inclusive
¹³⁶⁹ ϕ line and that in total contain less than 500 long tracks are considered. All
¹³⁷⁰ candidate kaon and pion tracks have to be long tracks and have to satisfy quality
¹³⁷¹ criteria. Lower limits on the momentum ($p > 1 \text{ GeV}/c$) and on the transverse
¹³⁷² momentum ($p_T > 100 \text{ MeV}/c$) are required. The candidates should be inconsistent
¹³⁷³ with originating from the PV and the particle identification (PID) system needs to
¹³⁷⁴ classify them as pions or kaons with only a small probability to be a ghost.

¹³⁷⁵ Of all the possible combinations of three charged hadron tracks forming a D^+
¹³⁷⁶ meson candidate only the two possibilities $D^\pm \rightarrow K^\mp \pi^\pm \pi^\pm$ and $D^\pm \rightarrow K^\mp K^\pm \pi^\pm$
¹³⁷⁷ are selected. The vertex needs to be significantly displaced from all PVs in the event
¹³⁷⁸ and the distance of the closest approach between all pairs of particles forming the
¹³⁷⁹ vertex has to be below 0.5 mm. The scalar sum of the p_T has to exceed 1800 MeV/c

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

and the combined invariant mass has to be in the range $\pm 25 \text{ MeV}/c^2$ around the nominal D^+ mass [123]. 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 $D^\pm \rightarrow K^\mp \pi^\pm \pi^\pm$ 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 $D^\pm \rightarrow K^\mp \pi^\pm \pi^\pm$ and $D^\pm \rightarrow K^\mp K^\pm \pi^\pm$ 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 $5 \text{ GeV}/c$. 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 's divided by the sum of the degrees of freedom of these vertex fits. Moreover, the B^0 candidates are required to have $p > 10 \text{ GeV}/c$ and to have $\chi_{\text{IP}}^2 < 25$, where χ_{IP}^2 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 [91], in which the B^0 production vertex is constrained to the position of the associated PV. Only candidates with decay times in the range 0.25 – 10.25 ps are kept. The invariant mass $m_{D^+ D^-}$ of the B^0 candidate has to be in the range 5150 – $5500 \text{ MeV}/c^2$. It is calculated from a DTF, 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 uncertainty on the invariant mass and on the decay time has to be below $30 \text{ MeV}/c^2$ 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 for the final state with two kaons is 82 % at a background rejection of 94 %. For the final state with three kaons the signal efficiency is 67.5 % at a background rejection of 98 %.

7.1.2 Veto

A $K \rightarrow \pi$ mis-ID can lead to background contributions from $D_s^+ \rightarrow K^- K^+ \pi^+$, which predominantly proceeds through $D_s^+ \rightarrow \phi \pi^+$. To reduce these D_s^+ contributions the

1422 kaon mass hypothesis is assigned to the pion with the higher transverse momentum
 1423 of $D^+ \rightarrow K^-\pi^+\pi^+$ candidates. The candidate is rejected if the invariant mass of the
 1424 hypothetical kaon pair is compatible with the ϕ mass of $M_\phi = 1019.461 \text{ MeV}/c^2$ [123]
 1425 within $\pm 10 \text{ MeV}/c^2$ or if the invariant mass of the $KK\pi$ combination is compatible
 1426 with the D_s^+ mass of $M_{D_s^+} = 1968.30 \text{ MeV}/c^2$ [123] within $\pm 25 \text{ MeV}/c^2$ and the
 1427 pion with the higher p_T (the one that the kaon mass hypothesis is assigned to)
 1428 has a larger $\text{ProbNN}K$ than $\text{ProbNN}\pi$ probability. When assigning the kaon mass
 1429 hypothesis to the pion with the lower p_T no vetoes are applied as no resonant
 1430 structures at the ϕ or the D_s^+ mass are found.

1431 To reduce $p \rightarrow \pi$ mis-ID the proton mass hypothesis is assigned to the pion
 1432 with the higher p_T of $D^+ \rightarrow K^-\pi^+\pi^+$ candidates. The candidate is rejected
 1433 if the invariant mass of the $Kp\pi$ combination is compatible with the Λ_c^+ mass
 1434 of $M_{\Lambda_c^+} = 2286.46 \text{ MeV}/c^2$ [123] within $\pm 25 \text{ MeV}/c^2$ and the proton probability
 1435 $\text{ProbNN}p$ of the pion with the higher p_T is larger than $\text{ProbNN}\pi$.

1436 7.1.3 Multivariate analysis

1437 BDT training

1438 To further suppress combinatorial background a Boosted Decision Tree (BDT) [108,
 1439 111] based on the implementation in TMVA [106] is trained using a signal MC
 1440 sample and the upper mass sideband with $m_{D^+D^-} > 5500 \text{ MeV}/c^2$. The training is
 1441 performed on half of these samples while the other half is used to test the BDT
 1442 performance. The selection steps described above, are applied before the training.

1443 Two BDTs separated by the number of kaons in the B^0 final state are trained.
 1444 The importance of the 21 features of the training differs, which is considered by
 1445 their order in Table 7.1. One of the features is the ratio of the kaon over the sum
 1446 of the kaon and the pion probabilities:

$$\text{PID ratio} = \frac{\text{ProbNN}K}{\text{ProbNN}K + \text{ProbNN}\pi}. \quad (7.1)$$

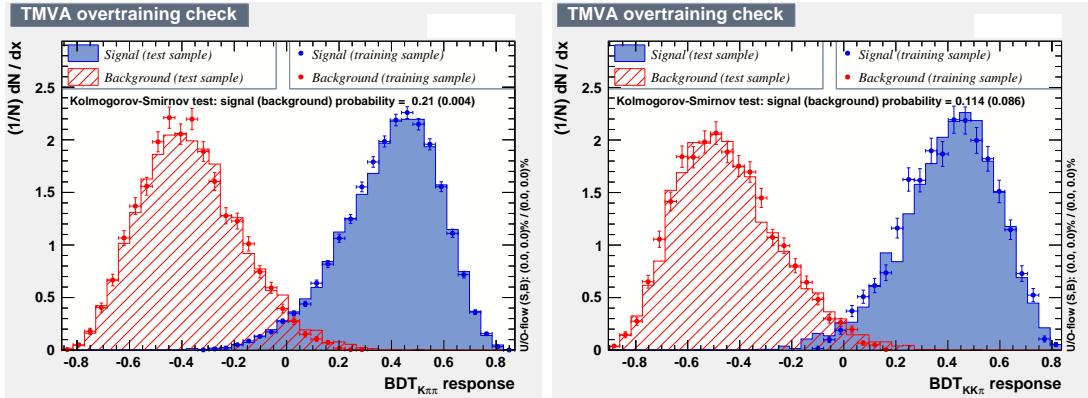
1447 It turns out that this ratio performs a little bit better than just using the simple
 1448 ProbNN variables. Among the other features are observables related to the kine-
 1449 matics of the decay like transverse momenta, decay time significances and direction
 1450 angles, qualities of the track segments in the VELO and the T-stations, and vertex
 1451 qualities.

1452 Before the training the features are transformed to decorrelate and decompose
 1453 them into the principal components, which improves the performance of the BDT.
 1454 The BDTs are each built out of 700 trees. The depth of the trees is limited to
 1455 three. At each node at least 3 % of the training events have to be present. The
 1456 variables are scanned at 40 points to find the optimal cut value. For the boosting
 1457 the AdaBoost method [112] with a boost factor of $\beta = 0.1$ is deployed.

1458 Overtraining is checked by applying the BDT on both the training and the testing
 1459 sample (see Fig. 7.1). Using simulations in the selection contains the possibility

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

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


Figure 7.1: 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).

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 the signal MC and

¹⁴⁶³ background-subtracted data show a quite large disagreement for both final states
¹⁴⁶⁴ as can be seen in Fig. 7.2. The performance is clearly overestimated in the training.
This would be a problem if the selection efficiencies had to be calculated using the

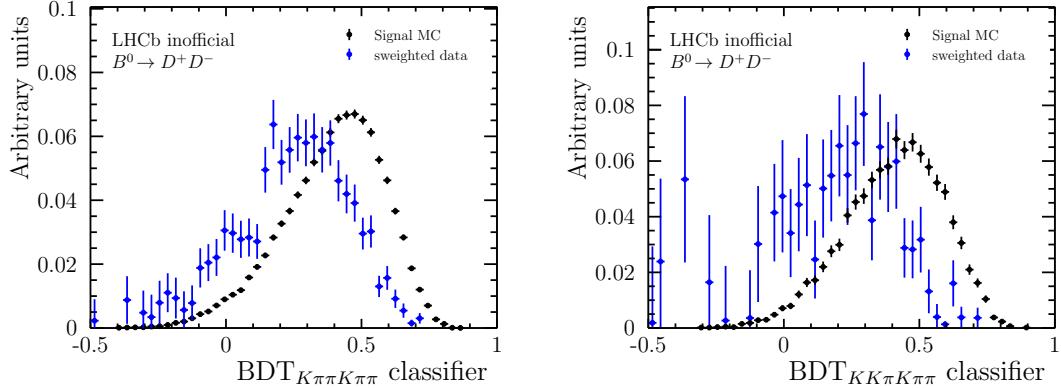


Figure 7.2: Distribution of the BDT output for background-subtracted data (blue) and signal MC (black) for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state (left) and the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state (right).

¹⁴⁶⁵
¹⁴⁶⁶ MC sample. But for a measurement of CP violation it is mainly important that the
¹⁴⁶⁷ amount of background can somehow be reduced while most of the signal is kept.
¹⁴⁶⁸ This can be achieved with the current setting.

¹⁴⁶⁹ BDT cut optimisation

¹⁴⁷⁰ As explained in Sec. 5.2.3 the best figure of merit for a measurement of CP violation
¹⁴⁷¹ is the sensitivity on the CP observables themselves. So the requirement on the
¹⁴⁷² BDT classifier output is scanned performing a fit to the invariant D^+D^- mass
¹⁴⁷³ spectrum followed by a decay time fit of the background-subtracted sample for
¹⁴⁷⁴ each scan point. Initially, only the subsample with two kaons in the B^0 final
¹⁴⁷⁵ state is analysed. In Fig. 7.3 the statistical uncertainties of $S_{D^+D^-}$ and $C_{D^+D^-}$
¹⁴⁷⁶ are plotted as a function of the requirement on the BDT classifier output. The
¹⁴⁷⁷ uncertainty on $C_{D^+D^-}$ improves with tighter requirements on the BDT classifier
¹⁴⁷⁸ until it reaches an optimum shortly after zero. This can be explained with the
¹⁴⁷⁹ fact that the sensitivity on $C_{D^+D^-}$ mainly comes from candidates at low decay
¹⁴⁸⁰ times because the cosine function is maximal there. The suppression of the rather
¹⁴⁸¹ short-lived combinatorial background compensates the loss in signal efficiency for a
¹⁴⁸² quite long range. In contrast, the uncertainty on $S_{D^+D^-}$ is mainly driven by the
¹⁴⁸³ amount of signal candidates. So it is more or less flat for loose requirements on the
¹⁴⁸⁴ BDT classifier where only few signal candidates are lost and this is compensated
¹⁴⁸⁵ by the higher purity and reaches its optimum around -0.25 before it starts to get
¹⁴⁸⁶ worse. Now that both observables are of interest and the optima are not at the
¹⁴⁸⁷ same cut value it is decided to require the BDT classifier to be greater than -0.10 .
¹⁴⁸⁸ This is a good compromise between both observables as the uncertainties of $S_{D^+D^-}$

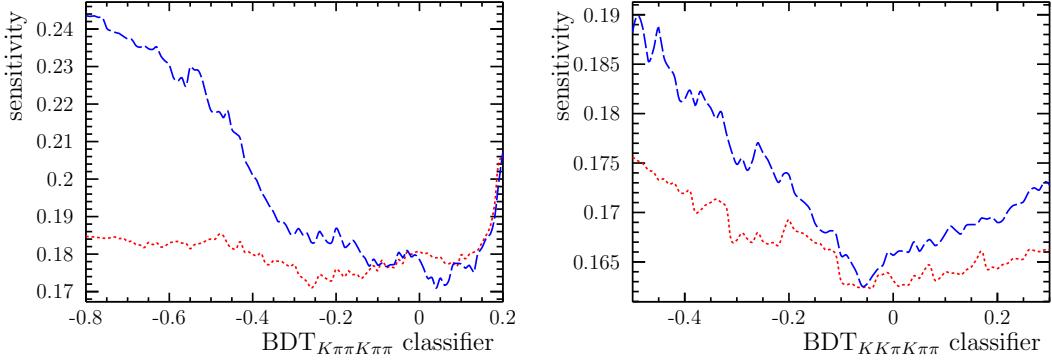


Figure 7.3: Sensitivity of S_{D+D^-} (red short-dashed) and C_{D+D^-} (blue long-dashed) as a function of the BDT classifier output for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state (left) and the $K^-K^+\pi^+K^+\pi^-\pi^-$ final state (right).

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 dataset 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. 7.3. 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 $5150\text{--}5500\,\text{MeV}/c^2$, which eliminates some backgrounds, like misreconstructed $B^0 \rightarrow D^{*+}D^-$, 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 $0.25\text{--}10.25\,\text{ps}$ to avoid edge effects. 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.

1511 7.2 Mass fit

1512 In this section, the fit of the invariant $m_{D^+D^-}$ mass distribution is described, which
 1513 is used to calculate signal candidates weights via the $_s\mathcal{P}lot$ method, and thereby
 1514 discriminates between signal and background candidates. As the linear Pearson
 1515 correlation coefficient between the invariant mass and the decay time is determined
 1516 to be $\rho = 0.007$ it is valid to apply the sWeights in the decay time fit later on (see
 1517 Sec. 7.3) to obtain the CP observables.

1518 The mass distribution is parametrised with a PDF \mathcal{P} consisting of five components,
 1519 $B^0 \rightarrow D^+D^-$ signal, $B_s^0 \rightarrow D^+D^-$ background, background from $B^0 \rightarrow D_s^+D^-$,
 1520 background from $B_s^0 \rightarrow D_s^-D^+$, and combinatorial background:

$$N^s \mathcal{P}^s = N_{B^0}^s \mathcal{P}_{B^0}^s + N_{B_s^0}^s \mathcal{P}_{B_s^0}^s + N_{B^0 \rightarrow D_s^+D^-}^s \mathcal{P}_{B^0 \rightarrow D_s^+D^-}^s + N_{B_s^0 \rightarrow D_s^-D^+}^s \mathcal{P}_{B_s^0 \rightarrow D_s^-D^+}^s + N_{\text{Bkg}}^s \mathcal{P}_{\text{Bkg}}^s . \quad (7.2)$$

1521 In the extended maximum likelihood fit four disjoint categories are simultaneously
 1522 fitted. It is distinguished between the two years of data-taking 2011 and 2012
 1523 and between the two final states $K^-\pi^+\pi^+K^+\pi^-\pi^-$ and $K^-K^+\pi^+K^+\pi^-\pi^-$ ($s =$
 1524 $\{2011, K\pi\pi\}, \{2011, KK\pi\}, \{2012, K\pi\pi\}, \{2012, KK\pi\}$). The tagging output is
 1525 not split.

1526 **$B^0 \rightarrow D^+D^-$ signal:** The $B^0 \rightarrow D^+D^-$ signal mass component is modelled by
 1527 the sum of three Crystal Ball functions [129], which share a common peak position
 1528 μ_{B^0} but have different width parameters σ_i . Two of the Crystal Ball functions
 1529 have a tail towards lower masses and one has a tail towards higher masses. The
 1530 parameters α_1 to α_3 of the power law functions, the ratio between the widths, and
 1531 the fractions f_1 and f_2 between the Crystal Ball functions are determined from a fit
 1532 to the invariant D^+D^- mass distribution of $B^0 \rightarrow D^+D^-$ signal MC in the range
 1533 4800–5400 MeV/c². This MC sample consists of both final states generated in the
 1534 ratio of the current world averages [50]. Apart from the mass range the full selection
 1535 is applied. The background categories 0, *i.e.* Signal, and 50, *i.e.* LowMassBkg with
 1536 missing photons, are considered. Events of the latter category create a very long tail
 1537 towards lower masses, which require the third Crystal Ball function. The exponent
 1538 n of all power law parts is fixed to 10. The widths σ_1^{MC} to σ_3^{MC} are multiplied by
 1539 a common scale factor R in the fit to data to account for differences in the mass
 1540 resolution between simulation and data. The fit results are listed in Table 7.2 and a
 1541 plot of the distribution overlaid with the projection of the PDF is given in Fig. 7.4.

1542 **$B_s^0 \rightarrow D^+D^-$ background:** Apart from the B^0 also the heavier B_s^0 can decay
 1543 to the D^+D^- final state. Almost the same parametrisation as for the B^0 signal
 1544 component is used, *i.e.* same width and tail parameters, while the peak position is
 1545 shifted by the world average $\Delta m_{B_s^0-B^0} = \mu_{B_s^0} - \mu_{B^0} = 87.35$ MeV/c² [50].

1546 **$B^0 \rightarrow D_s^+D^-$ background:** The vetoes applied in the selection suppress the
 1547 contribution from misidentified kaons. Nevertheless, a significant amount of

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

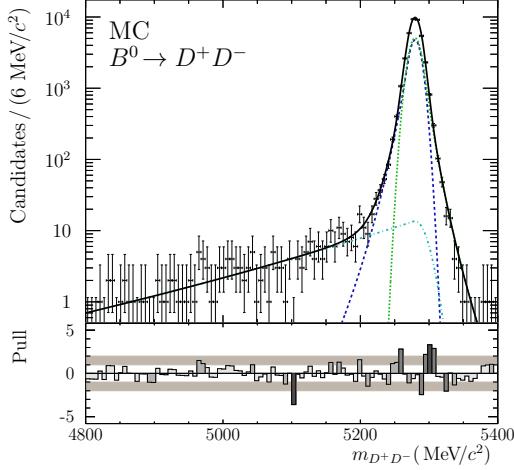


Figure 7.4: Mass distribution of the $B^0 \rightarrow D^+D^-$ signal MC sample overlaid with the projection of the fitted PDF. The blue dotted, green dashed and turquoise short-dash-dotted lines represent the three Crystal Ball components.

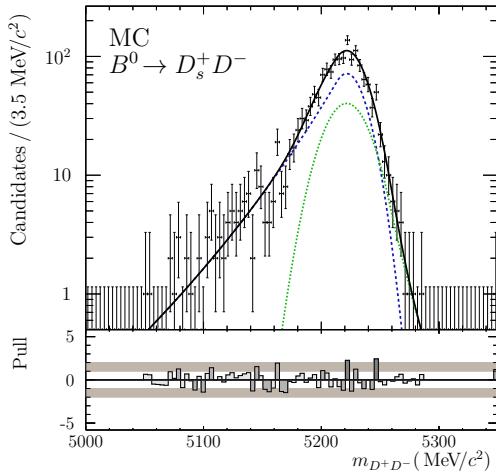
1548 $B^0 \rightarrow D_s^+D^-$ decays remains in the data sample. A fit to the invariant mass
1549 distribution of simulated $B^0 \rightarrow D_s^+D^-$ events reconstructed as $B^0 \rightarrow D^+D^-$ is
1550 performed. The full selection is applied to the simulated sample as this can change
1551 the shape of the $B^0 \rightarrow D_s^+D^-$ background contribution. The sum of two Crystal
1552 Ball PDFs with both power law exponents n fixed to 10 is used to parametrise the
1553 invariant mass distribution. The fit results are listed in Table 7.3 and the corre-
1554 sponding plot is shown in Fig. 7.5. The fraction parameter and the tail parameters
1555 are taken from this fit. The width parameters as well as the peak position are
1556 floating parameters in the fit to data.

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

Parameter	Value
$\mu_{B^0}^{\text{MC}}$ (MeV/c ²)	5279.70 ± 0.09
σ_1^{MC} (MeV/c ²)	8.5 ± 0.4
σ_2^{MC} (MeV/c ²)	16 ± 5
σ_3^{MC} (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

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

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

Figure 7.5: Mass distribution of the $B^0 \rightarrow D_s^+ D^-$ MC sample reconstructed as $B^0 \rightarrow D^+ D^-$ overlaid with the projection of the two Crystal Ball PDFs shown in blue dashed and green dotted.

¹⁵⁵⁷ **$B_s^0 \rightarrow D_s^- D^+$ background:** Although only few candidates of $B_s^0 \rightarrow D_s^- D^+$ decays
¹⁵⁵⁸ are expected this contribution is included in the nominal fit. Like the B^0 component
¹⁵⁵⁹ it is parametrised with the sum of two Crystal Ball PDFs. The peak position is
¹⁵⁶⁰ constrained to the sum of the peak position of the B^0 component and the mass
¹⁵⁶¹ difference $\Delta m_{B_s^0 - B^0}$. All other shape parameters are shared with the ones for the
¹⁵⁶² B^0 component.

¹⁵⁶³ **Combinatorial background:** The reconstructed mass PDF of the combinatorial
¹⁵⁶⁴ background is modelled by an exponential function with individual slopes $\beta^{K\pi\pi}$
¹⁵⁶⁵ and $\beta^{KK\pi}$ based on the number of kaons in the final state.

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

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

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

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

average of $\mu_{B^0}^{\text{WA}} = (5279.62 \pm 0.15) \text{ MeV}/c^2$ [50]. 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. In Fig. 7.6 the complete data-sample is plotted overlaid with the PDF projections and its components.

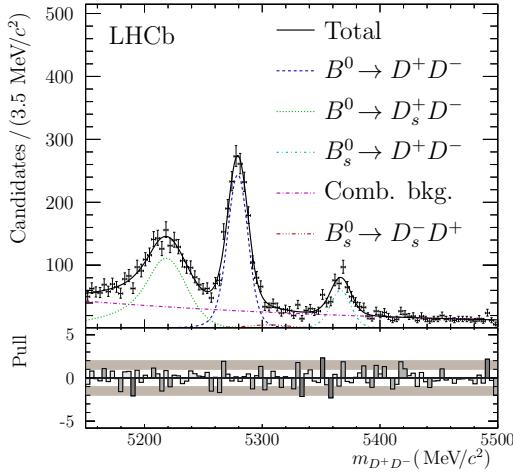


Figure 7.6: Plot of the reconstructed mass of the $B^0 \rightarrow D^+ D^-$ data sample with the projected PDF and pull distribution.

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

1578 and this can indeed be observed for the fitted yields. Additionally, there are around
1579 five times more signal candidates in the final state with two kaons than with three
1580 kaons, which also meets the expectations from the branching ratios.

1581 7.3 Decay time fit

1582 The conditional PDF describing the reconstructed decay time t' and tag decisions
1583 $\vec{d}' = (d_{\text{OS}}, d_{\text{SS}})$, given a per-event decay time resolution $\sigma_{t'}$ and per-event mistag
1584 probability estimates $\vec{\eta} = (\eta_{\text{OS}}, \eta_{\text{SS}})$, is

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

1585 where

$$\mathcal{P}(t, \vec{d}' | \vec{\eta}) \propto \sum_d \mathcal{P}(\vec{d}' | d, \vec{\eta}) [1 - d A_{\text{P}}] e^{-t/\tau} \{1 - d S \sin(\Delta m t) + d C \cos(\Delta m t)\}, \quad (7.4)$$

1586 and where t is the true decay time, d is the true production flavour, A_{P} is the
1587 production asymmetry, and $\mathcal{P}(\vec{d}' | d, \vec{\eta})$ is a two-dimensional binomial PDF describing
1588 the distribution of tagging decisions given $\vec{\eta}$ and d . Normalisation factors are omitted
1589 for brevity.

1590 7.3.1 Decay time resolution

1591 The prediction of the DTF on the decay time error is used to determine the decay
1592 time resolution. Like in the analysis of $B^0 \rightarrow J/\psi K_s^0$ (see Sec. 6.3) these predictions
1593 are calibrated using linear functions with parameters b and c . To account for
1594 different sources introducing the decay time resolution an effective model consisting
1595 of two Gaussians with per-event widths is used. Besides this common resolution
1596 effect the decay time resolution model is also supposed to describe the effect of
1597 events matched to the wrong PV, which can cause a large deviation between the true
1598 and the reconstructed decay time. The wrong PV component is parametrised with
1599 a broad Gaussian distribution using the same mean μ_t as the other two Gaussians
1600 and one width parameter σ_{PV} . The complete parametrisation of the resolution
1601 model is given by

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

1602 The first two Gaussian components have different calibration parameters b_i and
1603 c_i and thus different widths. Together with the fraction f_{PV} of the wrong PV
1604 component the fractions of the two Gaussian components g_1 and g_2 sum up to

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

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. 7.7). 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 $B^0 \rightarrow D^+D^-$ 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

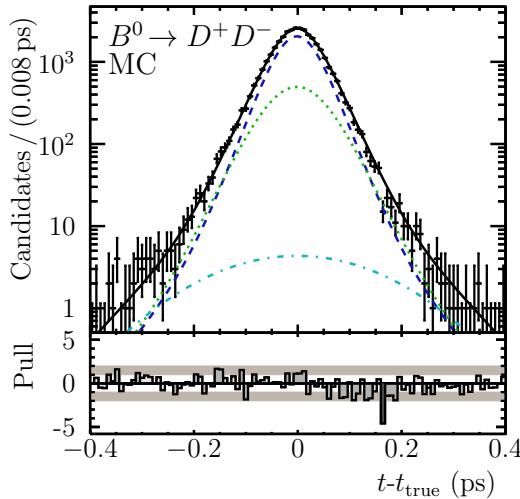


Figure 7.7: Fit of per event resolution model to the difference of true and reconstructed decay time in signal MC. The black solid line is the projection of the full PDF. The blue dashed and the green dotted line represent the two per-event components and the turquoise dashed-dotted line shows the wrong PV component.

differ between signal MC and data. In the analysis of $B_s^0 \rightarrow D_s^- \pi^+$ decays [130] it is found to be 1.15 times higher in data than in MC. If the same applied for $B^0 \rightarrow D^+D^-$ 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

1614 previously described per-event decay time resolution model without any corrections
1615 can be used in the nominal decay time fit.

1616 7.3.2 Decay time acceptance

1617 The trigger requirements as well as some input variables to the BDT result in a
1618 decay-time-dependent efficiency. Additionally, the VELO reconstruction (*i.e.* the
1619 FastVelo algorithm [131]) causes a drop in decay time acceptance for events with
1620 large decay times. In order to correctly describe these effects the B^0 lifetime is
1621 constrained to $\tau = (1.519 \pm 0.005)$ ps [123] in the nominal fit and any deviation of
1622 the decay time distribution (summed over the tags) from a pure exponential shape
1623 is supposed to be described by cubic splines (see Sec. 5.3). Knots are positioned
1624 on the rising edge at 0.8 ps, approximately at the turning point at 2 ps, and at the
1625 boundaries of the decay time range (0.25 and 10.25 ps). The normalisation of the
1626 splines is arbitrary and it has been decided to fix the second to last spline coefficient
1627 to 1.0.

1628 On signal MC the truth information is available so the shape of the decay time
1629 acceptance can be separated from the exponential decay. This shape is compared
1630 with the spline method described above. As the BDTs are trained and applied
1631 separately for the two final states and might have different effects on the shape of
1632 the decay time acceptance these two categories are studied individually.

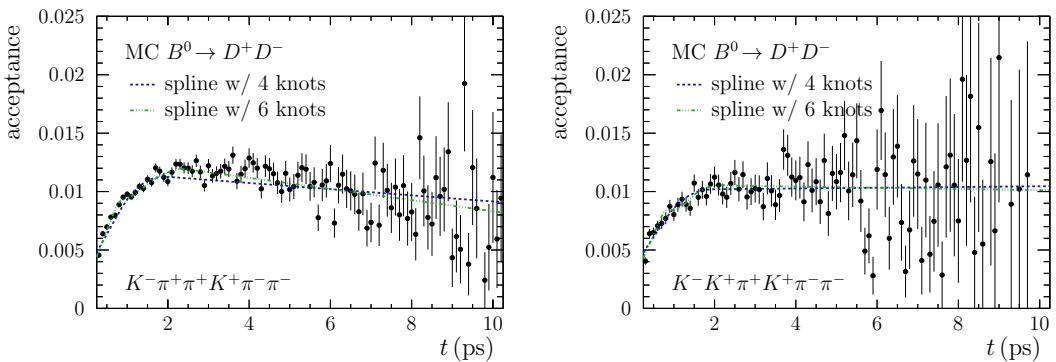


Figure 7.8: Decay time acceptance of truth-matched signal MC for the $K^- \pi^+ \pi^+ K^+ \pi^- \pi^-$ final state (left) and the $K^- K^+ \pi^+ K^+ \pi^- \pi^-$ final state (right). The black data points show the true decay time acceptance determined by dividing the reconstructed by the true decay time distribution. The blue dashed line is the spline acceptance function with four knots and the green dashed-dotted line with six knots.

1633 Looking at the plots in Fig. 7.8 it is apparent that compared to $B^0 \rightarrow J/\psi K_s^0$ there
1634 is a quite large efficiency loss at high decay times. This might be related to the fact
1635 that both B^0 daughter particles (D^+ and D^-) are relatively long-lived. The true MC
1636 decay time acceptance is overlaid with the shape of two spline functions. Besides the

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

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 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 [132]. Taking those distributions from $B^0 \rightarrow D^+D^-$ individual weighted averages for the 2011 and 2012 subsamples are calculated yielding

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

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 $A_P^{12} = A_P^{11} + \Delta A_P$ with $\Delta A_P = -0.0024 \pm 0.0018$ (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 [68] and is used as the width of a Gaussian constraint. The B^0 oscillation frequency and the B^0 lifetime are constrained to $\Delta m = (0.510 \pm 0.003) \hbar \text{ps}^{-1}$ [123] and $\tau = (1.519 \pm 0.005) \text{ ps}$ [123], respectively. The flavour-tagging calibration parameters (Table 5.1) are constrained within their combined statistical and systematic uncertainties, determined in the calibration using $B^0 \rightarrow D_s^+D^-$ decays. The decay time resolution parameters (Table 7.5) and the B^0 lifetime difference $\Delta\Gamma = 0 \text{ ps}^{-1}$ are fixed in the likelihood fit.

7.3.4 Results

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

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

Only after rescaling the sWeights via

$$w_i = w_i \frac{\sum w_i}{\sum w_i^2}, \quad (7.8)$$

correct asymmetric uncertainty estimates are delivered by MINOS, which is ROOT's standard method to analyse the likelihood shape. A plot of the decay time distribution and the projection of the acceptance model are shown in Fig. 7.9. Good agreement between the latter and the shape on signal MC (cf. Fig. 7.8) can be observed but the low statistics leading to rather large uncertainties indicated by the error band diminishes the significance of the comparison.

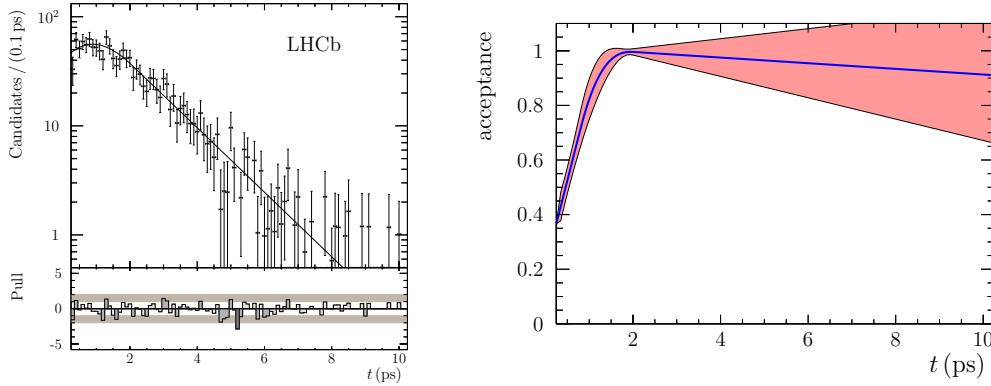


Figure 7.9: Plot of the decay time distribution of the background-subtracted $B^0 \rightarrow D^+ D^-$ data sample with the projection of the PDF and the pull distribution on the left. The y-axis is plotted in logarithmic scale. Plot of the nominal decay time acceptance model on the right. The red area indicates the 1σ 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

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

1691 statistical uncertainties for $S_{D^+ D^-}$ and $C_{D^+ D^-}$ from the bootstrapping including the
1692 impact of the uncertainty of the flavour-tagging calibration parameters are given
1693 by scaling the bootstrapping uncertainties by the width of the pull distributions:

$$\sigma_{S_{D^+ D^-}}(\text{bootstrapping}) = {}^{+0.17}_{-0.16}, \quad (7.9)$$

$$\sigma_{C_{D^+ D^-}}(\text{bootstrapping}) = {}^{+0.18}_{-0.17}. \quad (7.10)$$

1694 These uncertainties match the nominal ones from MINOS quite well. The 1D likelihood scans in Fig. 7.10 show a nice parabolic shape with a clear minimum.

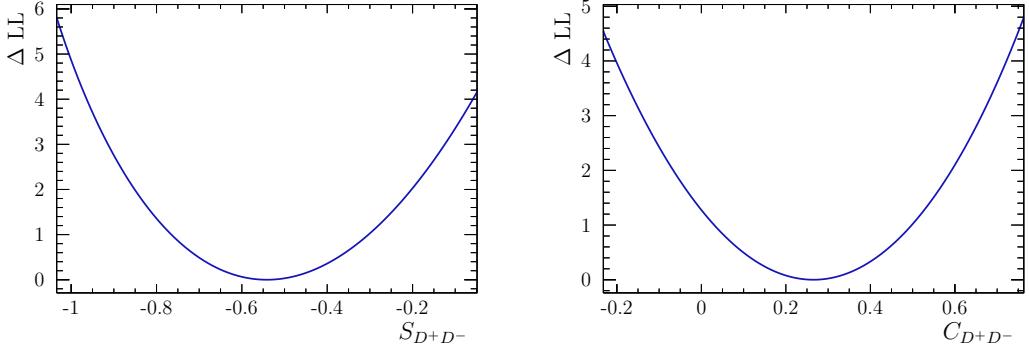


Figure 7.10: One-dimensional likelihood profile scans for $S_{D^+ D^-}$ and $C_{D^+ D^-}$.

1695
1696 Apart from a quite high positive correlation between the parameters of the
1697 acceptance spline function and the already quoted correlation of about 0.5 between
1698 $S_{D^+ D^-}$ and $C_{D^+ D^-}$, which is expected from first principles (see Ref. [133]), no large
1699 correlation between fitted parameters is present as can be seen from the correlation matrix visualised in Fig. 7.11. A possible correlation between Δm and $C_{D^+ D^-}$ is

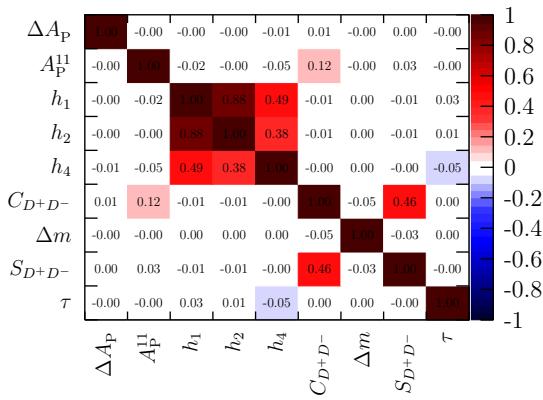


Figure 7.11: Visualised correlation matrix of the fit parameters in the decay time fit to data. Positive correlations are represented by the red palette on the z axis, while negative correlations are represented by the blue palette of the z axis.

1700

significantly reduced by the constraint applied on Δm , 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 is maintained in the nominal setup.

In Fig. 7.12 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

$$\mathcal{A}^{\text{meas}}(t) = \frac{\Delta\omega + A_P^{11}(1 - 2\omega) + (1 - 2\omega + A_P^{11}\Delta\omega)\mathcal{A}^{\text{theo}}(t)}{1 + A_P^{11}(S_{D^+D^-} \sin(\Delta m t) - C_{D^+D^-} \cos(\Delta m t))}, \quad (7.11)$$

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.

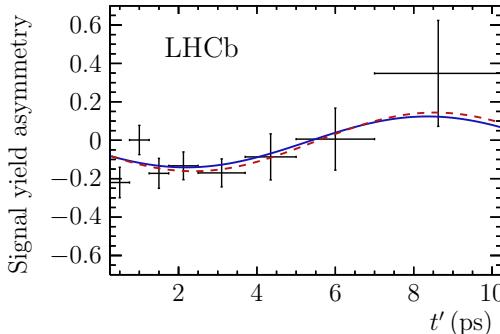


Figure 7.12: Decay-time-dependent signal yield asymmetry. The solid blue curve is the projection of the signal PDF given in Eq. (7.3) and the dashed red curve is the pure time-dependent fit function from Eq. (7.11)

1712

1713 7.4 Studies of systematic effects

1714 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

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

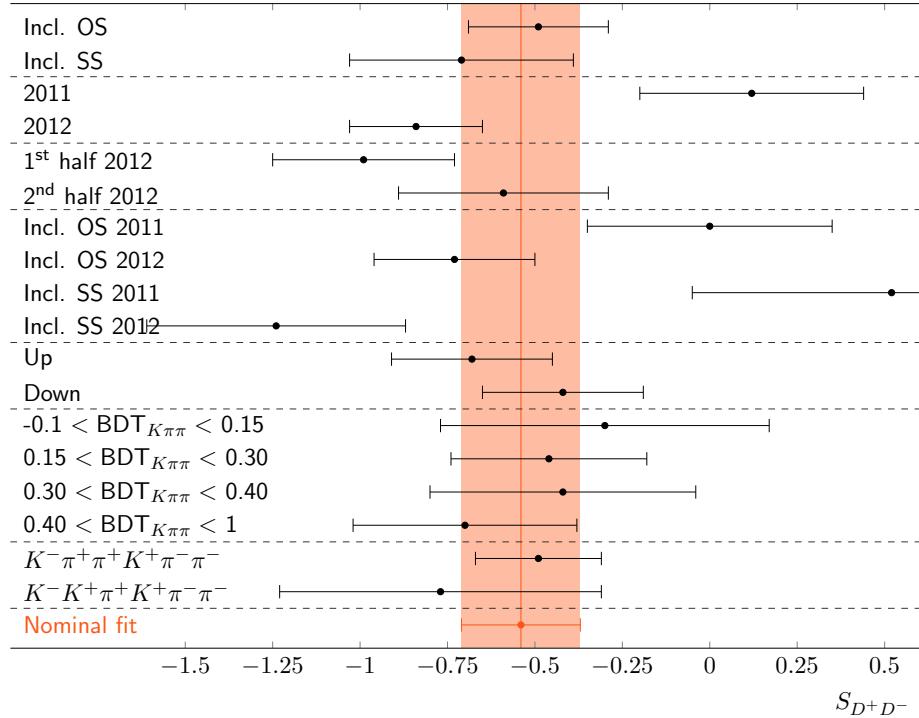


Figure 7.13: Comparison of fit results of $S_{D^+D^-}$ for fits on various subsamples.

1719 final states ($K^-\pi^+\pi^+K^+\pi^-\pi^-$ vs. $K^-\bar{K}^+\pi^+\pi^-\pi^-$) and for four different slices
1720 of the BDT classifier for the $K^-\pi^+\pi^+K^+\pi^-\pi^-$ final state.

1721 The fit results in the various scenarios are illustrated in Figs. 7.13 and 7.14. While
1722 almost all splits show compatible results a rather large difference can be observed
1723 between the 2011 and the 2012 subsample for $S_{D^+D^-}$. This is even more pronounced
1724 when using only SS tagging. When determining the flavour-tagging calibration
1725 parameters separately for 2011 and 2012 data small non-significant differences are
1726 observed but these can not explain the different results of the CP observables. So
1727 the best explanation is that the difference is due to a statistical fluctuation.

1728 For the nominal fit the decay times and the decay time errors from the DTF
1729 are used. The central values of the CP observables slightly change when using the
1730 decay time (error) from the LVF, $S_{D^+D^-} = -0.539$ (LVF) vs. -0.541 (DTF) and
1731 $C_{D^+D^-} = 0.266$ (LVF) vs. 0.263 (DTF). But this difference is clearly below the
1732 statistical significance.

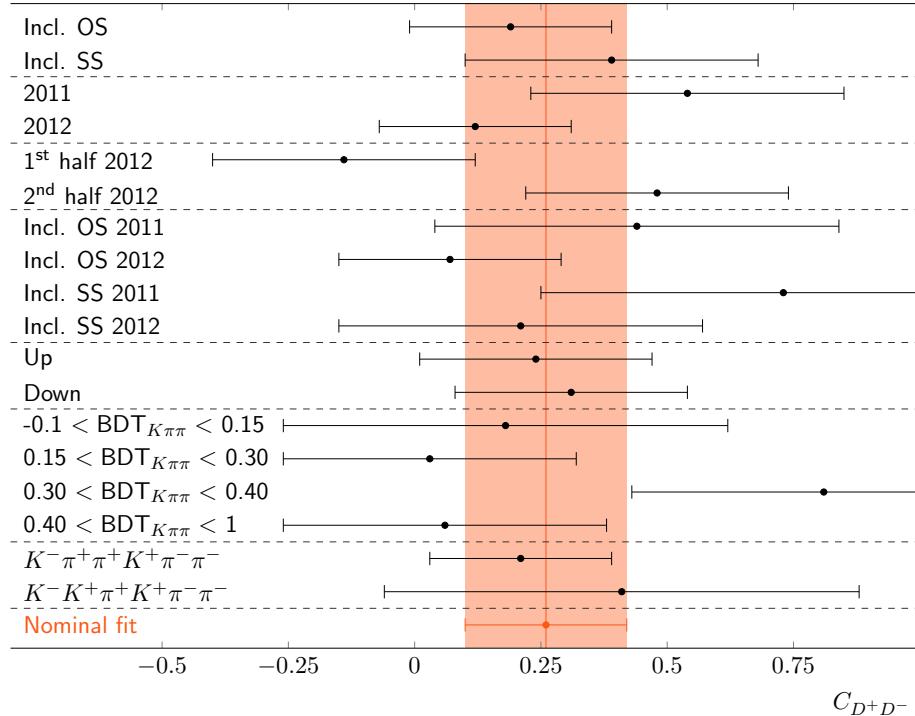


Figure 7.14: Comparison of fit results of C_{D+D-} for fits on various subsamples.

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. 7.15 show a very small deviation of the mean value from zero. Multiplying it with the statistical uncertainty the systematic uncertainty is calculated to be

$$s_{S_{D+D-}}^{\text{fit}} = 0.004, \quad s_{C_{D+D-}}^{\text{fit}} = 0.0025. \quad (7.12)$$

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 $B \rightarrow D^{*+}D^-$, with $D^{*+} \rightarrow D^+\pi^0$ or $D^{*+} \rightarrow D^+\gamma$, can mimic the $B \rightarrow D^+D^-$ decay. In the rest frame of the D^{*+} resonance, the missing momentum

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

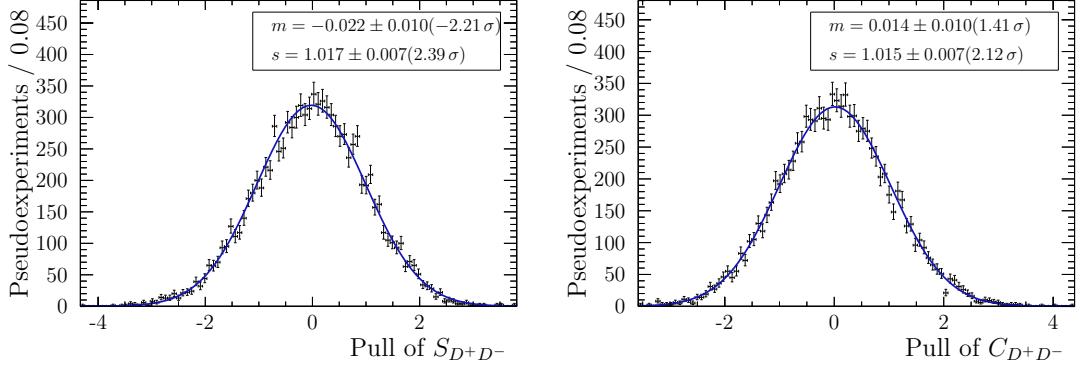


Figure 7.15: Pull distributions of $S_{D^+D^-}$ and $C_{D^+D^-}$ for a study on the systematic uncertainty due to the likelihood fitter.

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 θ of the missing π^0 . This leads to a double-horned structure approximately $140\text{ MeV}/c^2$ below the nominal B mass (see Ref. [134] for more details on the shape of this background). As the lower boundary on the invariant $m_{D^+D^-}$ mass is set to $5150\text{ MeV}/c^2$ the $B^0 \rightarrow D^{*+}D^-$ contribution lies completely outside the mass range used for the fit. However, the $B_s^0 \rightarrow D^{*+}D^-$ contribution enters the fit region. But since the expected number of $B_s^0 \rightarrow D^{*+}D^-$ 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 $B_s^0 \rightarrow D^{*+}D^-$ and for (partially) charmless background are included in the generation but excluded from the fit procedure.

The shape of $B_s^0 \rightarrow D^{*+}D^-$ is parametrised with two single Gaussian functions centred around $5150\text{ MeV}/c^2$ and $5200\text{ MeV}/c^2$. 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 $10\text{ MeV}/c^2$ is chosen. The mean is set to the same position as the B^0 signal. The $B_s^0 \rightarrow D^{*+}D^-$ 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 $B^0 \rightarrow D^{*+}D^-$ decays [135] a significant contribution of $B_s^0 \rightarrow D^{*+}D^-$ candidates is observed. The ratio between the two yields is determined to be 1:20. Under the assumption that the efficiencies for $B \rightarrow D^+D^-$ and $B \rightarrow D^{*+}D^-$ are

1775 the same the expected number of $B_s^0 \rightarrow D^{*+} D^-$ candidates can be calculated via

$$N_{B_s^0 \rightarrow D^{*+} D^-} = \frac{1}{20} N_{B^0 \rightarrow D^+ D^-} \frac{\mathcal{B}(B^0 \rightarrow D^{*+} D^-) \mathcal{B}(D^{*+} \rightarrow D^+(\pi^0 || \gamma))}{\mathcal{B}(B^0 \rightarrow D^+ D^-)} . \quad (7.13)$$

1776 Using the world averages for the branching ratios [50] the number of candidates
 1777 to be generated in the pseudoexperiments is estimated to be $N(B_s^0 \rightarrow D^{*+} D^-) =$
 1778 66 ± 9 .

1779 To determine how many (partially) charmless background candidates need to
 1780 be generated the D mass window is widened to $\pm 40 \text{ MeV}/c^2$ and the nominal D
 1781 mass window of $\pm 25 \text{ MeV}/c^2$ is vetoed for one or for both D candidates. Fits to
 1782 the invariant B mass without the D mass constraint are performed in the various
 1783 scenarios. The fitted yields, which are constrained to positive values, are scaled to
 1784 account for the applied D mass window. The total amount of residual contamination
 1785 ($B^0 \rightarrow Dhhh$ or $B^0 \rightarrow hhhhhh$ decays) surviving the $B^0 \rightarrow D^+ D^-$ selection is found
 1786 to be 28.7 ± 19.5 candidates for the $K^- K^+ \pi^+ K^+ \pi^- \pi^-$ final state and 0.0 ± 27.8
 1787 candidates for the $K^- \pi^+ \pi^+ K^+ \pi^- \pi^-$ final state. For the pseudoexperiments the
 1788 number of (partially) charmless background is drawn from Gaussian distributions
 1789 using these values for mean and width. When the outcome is negative the procedure
 1790 is repeated until a positive yield is drawn.

1791 The systematic uncertainties on $S_{D^+ D^-}$ and $C_{D^+ D^-}$ are calculated as the product
 1792 of the bias on the mean parameter of the pull distributions and the statistical
 1793 uncertainty:

$$s_{S_{D^+ D^-}}^{\text{mass},1} = 0.05 , \quad s_{C_{D^+ D^-}}^{\text{mass},1} = 0.013 .$$

1794 **Mismodelling of mass components** The BDT is trained with MC samples that
 1795 are known to not perfectly model the PID information. As a result the BDT
 1796 classifier distributions of background-subtracted data and MC show a quite big
 1797 discrepancy. Some shape parameters are estimated on MC samples and might
 1798 be distorted by the data/MC differences. Therefore, different alternative mass
 1799 parametrisations are tested against the nominal model: the component of the
 1800 $B^0 \rightarrow D^+ D^-$ signal (and of the $B_s^0 \rightarrow D^+ D^-$ 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 $B \rightarrow D_s D$ 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

$$s_{S_{D^+ D^-}}^{\text{mass},2} = 0.004 , \quad s_{C_{D^+ D^-}}^{\text{mass},2} = 0.006 .$$

1808 **7.4.3.2 Correlation between decay time and mistags**

1809 The correlation between the decay time distribution and the per-event mistags
1810 is studied by calculating the linear Pearson correlation coefficient $\rho(\eta, t)$. The
1811 significance of the correlation value, *i.e.* 95 % confidence level interval, is determined
1812 using the bootstrapping method (Sec. 5.4) with 10 000 repetitions. The correlation
1813 coefficients are found to be small. The profile histogram of the OS tagging com-
1814 bination, which shows the average η_{OS} value as a function of the decay time, is
1815 flat within statistics. For the SS tagging combination the profile histogram slowly
1816 increases with decay time. This can be confirmed by analysing the larger signal
MC sample (see Fig. 7.16). Performing a χ^2 fit in the decay time range 0.25–8.25 ps

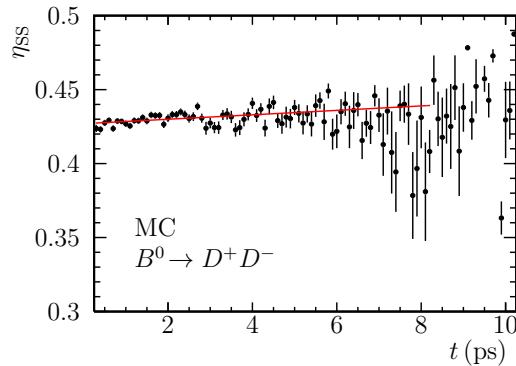


Figure 7.16: Profile histogram for the decay time dependence on η_{SS} for signal MC. The black data points represent the mean value of η_{SS} and its uncertainty for each bin in t . The red curve is the fitted linear function.

1817
1818 with the linear function

$$\eta_{\text{SS}} = a_{\eta_{\text{SS}},t} t + b_{\eta_{\text{SS}},t} \quad (7.14)$$

1819 yields a slope of $a_{\eta_{\text{SS}},t} = (1.50 \pm 0.27) \text{ ns}^{-1}$. Although this is a significant deviation
1820 from zero the correlation is not taken into account in the nominal fit model. Instead a
1821 study on the systematic uncertainty from neglecting this effect is performed. In 1000
1822 pseudoexperiments the SS mistag is generated using a Gaussian distribution whose
1823 mean is drawn from the linear function defined in Eq. (7.14) thereby introducing
1824 the correlation with the decay time. In the subsequent fit the correlation is again
1825 ignored. This leads to systematic uncertainties of

$$s_{S_{D^+D^-}}^{\text{corr}} = 0.0007, \quad s_{C_{D^+D^-}}^{\text{corr}} = 0.007.$$

1826 **7.4.3.3 Decay Time Resolution Model**

1827 As calculated in Sec. 7.3.1 even an underestimation of the decay time resolution
1828 by 15 % has only a minor effect on the resolution related dilution. Nevertheless,
1829 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

$$s_{S_{D^+D^-}}^{\text{res}} = 0.0020, \quad s_{C_{D^+D^-}}^{\text{res}} = 0.0023.$$

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. 7.8). Small differences are observed. As the low statistics in the $K^-K^+\pi^+\pi^-\pi^-$ 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:

$$s_{S_{D^+D^-}}^{\text{acc}} = 0.007, \quad s_{C_{D^+D^-}}^{\text{acc}} = 0.0027.$$

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 $\sigma_{z\text{-scale}} = 0.022\%$ [136]. 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:

$$s_{S_{D^+D^-}}^{z\text{-scale}} = 0.0031, \quad s_{C_{D^+D^-}}^{z\text{-scale}} = 0.0028.$$

7.4.4.2 Production Asymmetry

The systematic uncertainty on the production asymmetry A_P^{11} is studied using 1000 pseudoexperiments. The nominal value is used in the generation and the procedure described in Ref. [137] is applied in the fit: Before fitting the data sample the mean of the Gaussian constraint for A_P^{11} 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_P^{11} in the fit. Both shifts, upwards and downwards, are tested and the larger deviation is taken as systematic uncertainty:

$$s_{S_{D^+D^-}}^{A_P} = 0.0015, \quad s_{C_{D^+D^-}}^{A_P} = 0.004.$$

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 $\sigma(\Delta\Gamma_d) = \pm 0.007 \text{ ps}^{-1}$ [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

$$s_{S_{D^+D^-}}^{\Delta\Gamma_d} = 0.014, \quad s_{C_{D^+D^-}}^{\Delta\Gamma_d} = 0.0021.$$

7.4.4.4 B^0 Mass Difference Δm_d

The systematic uncertainty on the world average of Δm_d ($\pm 0.002 \hbar \text{ps}^{-1}$ [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

$$s_{S_{D^+D^-}}^{\Delta m_d} = 0.0025, \quad s_{C_{D^+D^-}}^{\Delta m_d} = 0.006,$$

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

¹⁸⁹⁵ 8 Discussion

¹⁸⁹⁶ 8.1 Comparison with previous measurements of ¹⁸⁹⁷ $\sin 2\beta_{(\text{eff})}$

¹⁸⁹⁸ Measurements of CP violation in $B^0 \rightarrow J/\psi K_s^0$ decays have been performed since the
¹⁸⁹⁹ end of the nineties. The first results by OPAL [138], ALEPH [139] and CDF [140]
¹⁹⁰⁰ had an uncertainty on $\sin 2\beta$ no better than ± 0.4 . One of the main purposes
¹⁹⁰¹ of the B -factories was to improve the precision, which succeeded with results of
¹⁹⁰² $S_{J/\psi K_s^0} = 0.657 \pm 0.036 \pm 0.012$ by BaBar [141] and $S_{J/\psi K_s^0} = 0.670 \pm 0.029 \pm 0.013$
¹⁹⁰³ by Belle [142]. Averaging the results of the B -factories and combining them
¹⁹⁰⁴ with measurements in various other charmonium modes leads to an average of
¹⁹⁰⁵ $\sin 2\beta = 0.679 \pm 0.020$ [59]. The measurement presented in this thesis, using
¹⁹⁰⁶ data corresponding to an integrated luminosity of 3 fb^{-1} collected with the LHCb
¹⁹⁰⁷ experiment, further improves on this yielding an updated world average of $\sin 2\beta$
¹⁹⁰⁸ $= 0.691 \pm 0.017$ [59]. There is also the possibility to constrain β from a global fit
¹⁹⁰⁹ to the CKM triangle. When not using the inputs from the direct measurements
¹⁹¹⁰ described above ($\beta = (21.85 \pm 0.70)^\circ$), a value of $\beta = (24.3^{+1.3}_{-1.4})^\circ$ [60] is found.
¹⁹¹¹ This shows that the value found by LHCb improves the compatibility between the
¹⁹¹² direct and the indirect measurements by shifting $\sin 2\beta$ slightly upwards.

¹⁹¹³ While for $B^0 \rightarrow J/\psi K_s^0$ decays mainly the value of $S_{J/\psi K_s^0}$ is interesting, as it
¹⁹¹⁴ is expected to correspond to $\sin 2\beta$ with only small corrections, the same does
¹⁹¹⁵ not apply for the measurement of CP violation using $B^0 \rightarrow D^+ D^-$ decays. Both
¹⁹¹⁶ observables, $S_{D^+ D^-}$ and $C_{D^+ D^-}$, have to be considered simultaneously for a proper
¹⁹¹⁷ interpretation of the results. In Fig. 8.1 the latest results from BaBar, Belle, and
¹⁹¹⁸ LHCb, as well as the average of these three are plotted in the two-dimensional plane
¹⁹¹⁹ of $C_{D^+ D^-}$ versus $S_{D^+ D^-}$. Looking at the uncertainty ellipses it is apparent that the
¹⁹²⁰ precision of LHCb matches the one of Belle, while being significantly better than
¹⁹²¹ the one of BaBar. The orientation of the ellipses shows that in the measurements
¹⁹²² of the B -factories the two CP observables are determined almost uncorrelated. For
¹⁹²³ a comparison of the central values it is useful to take into account the arc defined
¹⁹²⁴ by the condition $S_{D^+ D^-}^2 + C_{D^+ D^-}^2 = 1$. It represents the extreme case of $\lambda_{D^+ D^-}$
¹⁹²⁵ being purely imaginary and delimits the physically allowed region. The result by
¹⁹²⁶ Belle [143] of $S_{D^+ D^-} = -1.06$ and $C_{D^+ D^-} = -0.43$ lies outside, while the results of
¹⁹²⁷ BaBar [144], which are $S_{D^+ D^-} = -0.63$ and $C_{D^+ D^-} = -0.07$, as well as the result
¹⁹²⁸ of LHCb (see Eq. (7.7)) are inside.

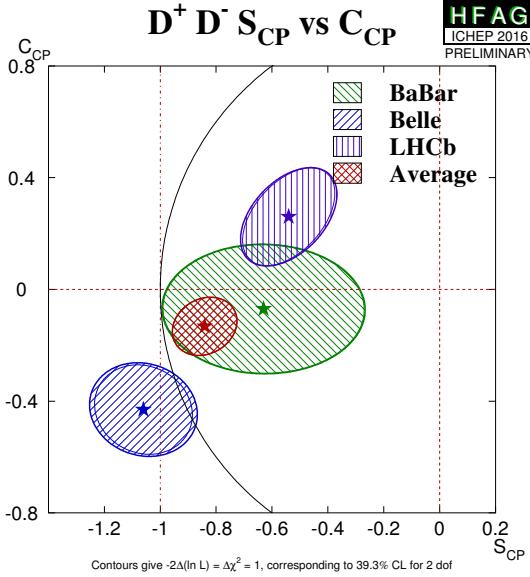


Figure 8.1: Comparison of CP observables from $B^0 \rightarrow D^+D^-$ decays in $(S_{D^+D^-}, C_{D^+D^-})$ plane [59]. The contours give $-2\Delta(\ln L) = \Delta\chi^2 = 1$, thereby corresponding to a coverage of 39.3 % for 2 degrees of freedom.

1929 8.2 Comparison between CP violation in 1930 $B^0 \rightarrow J/\psi K_s^0$ and in $B^0 \rightarrow D^+D^-$ decays

1931 Measurements in decays with $b \rightarrow c\bar{s}$ transitions have established the presence of
1932 CP violation in the system of neutral B mesons with very high precision. Comparing
1933 the plots of the decay-time-dependent signal yield asymmetry for $B^0 \rightarrow J/\psi K_s^0$ and
1934 $B^0 \rightarrow D^+D^-$ decays (see Fig. 8.2), the difference in sensitivity becomes apparent.
1935 However, the main motivation for the measurement of CP violation in $B^0 \rightarrow D^+D^-$
1936 decays, presented in this thesis, is not to achieve the single best measurement of
1937 CP violation. In fact, given the uncertainties the conservation of CP symmetry can
1938 be excluded by only 4.0 standard deviations with this measurement. Instead, the
1939 phase shift $\Delta\phi$ can be obtained, which is caused by higher-order Standard Model
1940 contributions and complicates the determination of β and β_s similarly (see Sec. 3.5).
1941 The fit results of $S_{D^+D^-}$ and $C_{D^+D^-}$ (see Eq. (7.7)) correspond to

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

1942 where the statistical uncertainty is estimated by generating three million sets of
1943 $S_{D^+D^-}$ and $C_{D^+D^-}$ using a two-dimensional Gaussian distribution including their
1944 correlation, calculating $\sin(\phi_d + \Delta\phi)$ for each of them, and then taking the two-sided
1945 68 % confidence intervals. This result can be combined with the corresponding

8.2 Comparison between CP violation in $B^0 \rightarrow J/\psi K_s^0$ and in $B^0 \rightarrow D^+ D^-$ decays

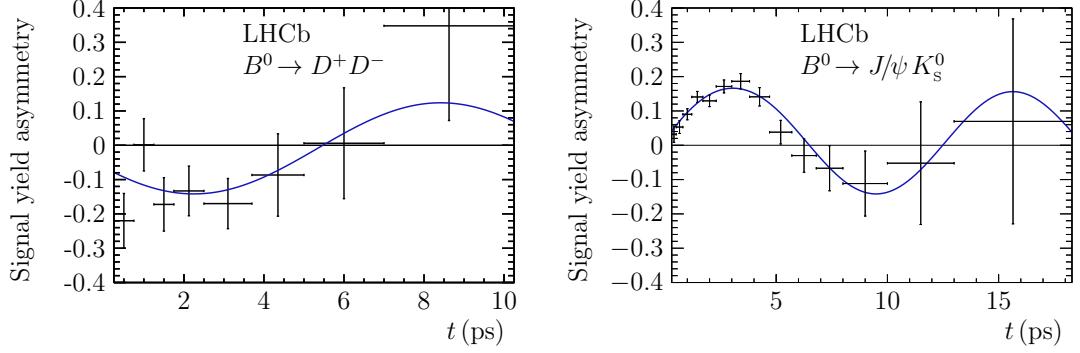


Figure 8.2: Signal yield asymmetry from $B^0 \rightarrow D^+ D^-$ decays [128] (left) and $B^0 \rightarrow J/\psi K_s^0$ decays [124] (right).

value from $B^0 \rightarrow J/\psi K_s^0$ to extract the phase shift $\Delta\phi$. The calculation yields $\Delta\phi = -0.16^{+0.19}_{-0.21}$ rad, which is significantly more precise than the previous measurement using the results by the B -factories of $\Delta\phi = (30^{+23})^\circ = 0.52^{+0.40}_{-0.56}$ [72], 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.

¹⁹⁵² 9 Conclusion

¹⁹⁵³ Since the discovery of CP violation in 1964 by Christenson, Cronin, Fitch and
¹⁹⁵⁴ Turlay [6] there have been many experiments searching for CP violation, first in
¹⁹⁵⁵ the sector of neutral K mesons but later on also for neutral B mesons. The most
¹⁹⁵⁶ significant indication of CP violation for B^0 mesons is found by the determination
¹⁹⁵⁷ of $\sin 2\beta$ using $B^0 \rightarrow J/\psi K_s^0$ decays. But although LHCb is not the first experiment,
¹⁹⁵⁸ even not at a hadron collider, to measure CP violation it plays an important role
¹⁹⁵⁹ in the further exploration of the quark-mixing sector.

¹⁹⁶⁰ The measurement of CP violation in $B^0 \rightarrow J/\psi K_s^0$ decays [124] using proton-
¹⁹⁶¹ proton collision data corresponding to an integrated luminosity of 3 fb^{-1} , which is
¹⁹⁶² presented in this thesis, yields

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

¹⁹⁶³ which is the most precise determination of these CP observables at a hadron collider
¹⁹⁶⁴ to date and is almost as precise as the previous measurements by BaBar [141] and
¹⁹⁶⁵ Belle [142]. The central values are compatible with the world averages and with
¹⁹⁶⁶ the Standard Model expectations. Thus, it serves as a benchmark measurement
¹⁹⁶⁷ showing the capability of LHCb to perform flavour-tagged precision measurements
¹⁹⁶⁸ of CP violation. The experimental difficulties, *e.g.* decay time resolution, production
¹⁹⁶⁹ asymmetries and asymmetries induced by the flavour tagging, are well under control
¹⁹⁷⁰ as the result is statistically limited. The largest systematic uncertainty on $S_{J/\psi K_s^0}$
¹⁹⁷¹ is introduced by a possible tagging asymmetry of the background, which is not
¹⁹⁷² accounted for in the likelihood fit (see Sec. 6.6). With an increased statistics this
¹⁹⁷³ effect can probably be analysed, controlled and suppressed better. Furthermore, in
¹⁹⁷⁴ the meantime new same-side flavour-tagging algorithms have been developed [105],
¹⁹⁷⁵ which are used for the first time in the measurement of CP violation in $B^0 \rightarrow D^+ D^-$
¹⁹⁷⁶ decays [128] yielding

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

¹⁹⁷⁷ If the flavour-tagging performance was the same as in the $B^0 \rightarrow J/\psi K_s^0$ analysis, the
¹⁹⁷⁸ 70 times lower number of available untagged signal candidates ($114\,000 B^0 \rightarrow J/\psi K_s^0$
¹⁹⁷⁹ decays vs. $1610 B^0 \rightarrow D^+ D^-$ decays) would only allow a sensitivity of ± 0.29 and
¹⁹⁸⁰ ± 0.27 for $S_{D^+ D^-}$ and $C_{D^+ D^-}$, respectively. However, the kinematic properties of the
¹⁹⁸¹ selected $B^0 \rightarrow D^+ D^-$ candidates lead to a significantly higher tagging power and

9 Conclusion

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 $B^0 \rightarrow J/\psi K_s^0$ decays. The value of $\varepsilon_{\text{eff}} = 8.1\%$ for the $B^0 \rightarrow D^+ D^-$ sample is the highest tagging power to date in a tagged CP violation measurement at LHCb.

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

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

can be transferred to the measurement of CP violation in $B_s^0 \rightarrow D_s^+ D_s^-$ decays [145], where ϕ_s , the mixing phase of the B_s^0 meson sector, can be determined, but only in a sum with this phase shift.

The analysis of $B^0 \rightarrow D^+ D^-$ decays is only the starting point for similar measurements in other $B^0 \rightarrow D_{(s)}^{(*)+} D_{(s)}^{(*)-}$ decay modes. First studies using $B^0 \rightarrow D^{*+} D^-$ decays have already been performed [135]. Recent calculations taking into account the flavour-tagging performance seen in $B^0 \rightarrow D^+ D^-$ and the increase in statistics when adding data from Run II indicate that the sensitivity of BaBar [144] and Belle [143] 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. Therefore, more and more data is collected at the LHC during Run II. Due to the higher centre-of-mass energy of currently 13 TeV the $b\bar{b}$ cross section, which roughly scales linearly with the centre-of-mass energy, is even increased with respect to the proton-proton collisions recorded during Run I. The same applies for direct searches at ATLAS and CMS, which can extend their direct searches to higher masses. In principle, for LHCb the instantaneous luminosity is another adjusting screw to further increase the data samples, though it is already higher than the design value originally planned in the proposal for the detector [146]. 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 [147]. 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

²⁰²³ mesons the restart of Belle II, planned for 2017, with an improved detector should
²⁰²⁴ increase the sensitivity of many measurements significantly. Only the combination
²⁰²⁵ of all these efforts to increase the available amount of data to be analysed might
²⁰²⁶ result in the observation of deviations from the Standard Model expectations and
²⁰²⁷ thus a hint for new physics.

2028

Bibliography

- 2029 [1] Planck collaboration, P. A. R. Ade *et al.*, *Planck 2015 results. XIII. Cosmo-*
- 2030 *logical parameters*, Astron. Astrophys. **594** (2016) A13, [arXiv:1502.01589](#).
- 2031 [2] G. Gamow, *Expanding universe and the origin of elements*, Phys. Rev. **70**
- 2032 (1946) 572.
- 2033 [3] A. D. Sakharov, *Violation of CP invariance, c asymmetry, and baryon asym-*
- 2034 *metry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32, [*Usp. Fiz.*
- 2035 *Nauk* **161**, 61 (1991)].
- 2036 [4] Super-Kamiokande collaboration, K. Abe *et al.*, *Search for proton decay via*
- 2037 *$p \rightarrow \nu K^+$ using 260 kiloton · year data of Super-Kamiokande*, Phys. Rev. **D90**
- 2038 (2014) 072005, [arXiv:1408.1195](#).
- 2039 [5] R. A. McKee and J. P. Stark, *Violation of thermal equilibrium in a thermomi-*
- 2040 *gration experiment*, Phys. Rev. B **11** (1975) 1374.
- 2041 [6] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the*
- 2042 *2π decay of the K_2^0 meson*, Phys. Rev. Lett. **13** (1964) 138.
- 2043 [7] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP*
- 2044 *violation*, Phys. Rev. **D51** (1995) 379, [arXiv:hep-ph/9404302](#).
- 2045 [8] H. W. Babcock, *The rotation of the Andromeda Nebula*, Lick Observatory
- 2046 *Bulletin* **19** (1939) 41.
- 2047 [9] Super-Kamiokande collaboration, Y. Fukuda *et al.*, *Evidence for os-*
- 2048 *cillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562,
- 2049 [arXiv:hep-ex/9807003](#).
- 2050 [10] SNO collaboration, Q. R. Ahmad *et al.*, *Measurement of the rate of $\nu_e + d \rightarrow$*
- 2051 *$p + p + e^-$ interactions produced by 8B solar neutrinos at the Sudbury Neutrino*
- 2052 *Observatory*, Phys. Rev. Lett. **87** (2001) 071301, [arXiv:nucl-ex/0106015](#);
- 2053 *Direct evidence for neutrino flavor transformation from neutral current in-*
- 2054 *teractions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002)
- 2055 011301, [arXiv:nucl-ex/0204008](#).
- 2056 [11] Nobel Media, *The nobel prize in physics 2015*, Nobelprize.org, Oct, 2016.
- 2057 [12] I. C. Brock and T. Schorner-Sadenius, eds., *Physics at the Terascale*, 2011.

Bibliography

- 2058 [13] D. Griffiths, *Introduction to elementary particles*, 2008.
- 2059 [14] D. H. Perkins, *Introduction to high energy physics*, 1982.
- 2060 [15] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D **10** (1974) 2445.
- 2061 [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.
- 2064 [17] LHCb collaboration, R. Aaij *et al.*, *Observation of exotic $J/\psi\phi$ structures from amplitude analysis of $B^+ \rightarrow J/\psi\phi K^+$ decays*, arXiv:1606.07895, submitted to Phys. Rev. Lett.; LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of $B^+ \rightarrow J/\psi\phi K^+$ decays*, arXiv:1606.07898, submitted to PRD.
- 2065 [18] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.
- 2066 [19] LHCb collaboration, R. Aaij *et al.*, *Evidence for exotic hadron contributions to $\Lambda_b^0 \rightarrow J/\psi p \pi^-$ decays*, Phys. Rev. Lett. **117** (2016) 082003, Erratum ibid. **117** (2016) 109902, arXiv:1606.06999.
- 2067 [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.
- 2068 [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.
- 2069 [22] S. Tomonaga, *On a relativistically invariant formulation of the quantum theory of wave fields*, Progress of Theoretical Physics **1** (1946) 27.
- 2070 [23] 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.
- 2071 [24] 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.
- 2072 [25] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.
- 2073 [26] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.

- 2093 [27] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.
- 2094 [28] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.
- 2095
- 2096 [29] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.
- 2097
- 2098 [30] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343.
- 2099
- 2100 [31] H. D. Politzer, *Reliable perturbative results for strong interactions?*, Phys. Rev. Lett. **30** (1973) 1346.
- 2101
- 2102 [32] E. Noether, *Invariante Variationsprobleme*, Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse (1918) 235.
- 2103
- 2104 [33] W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinetik und Mechanik*, Z. Phys. **43** (1927) 172.
- 2105
- 2106 [34] 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)].
- 2107
- 2108
- 2109 [35] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv:1406.6482.
- 2110
- 2111 [36] 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.
- 2112
- 2113 [37] M. Gronau, *U spin symmetry in charmless B decays*, Phys. Lett. **B492** (2000) 297, arXiv:hep-ph/0008292.
- 2114
- 2115 [38] J. S. Schwinger, *The theory of quantized fields. 1.*, Phys. Rev. **82** (1951) 914.
- 2116
- 2117 [39] 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. Medd. **28N5** (1954) 1.
- 2118
- 2119 [40] W. Pauli, *Niels Bohr and the Development of Physics: Essays Dedicated to Niels Bohr on the Occasion of His Seventieth Birthday*, McGraw-Hill, 1955.
- 2120
- 2121 [41] H. Georgi and S. L. Glashow, *Unity of all elementary particle forces*, Phys. Rev. Lett. **32** (1974) 438.
- 2122
- 2123 [42] 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.
- 2124
- 2125
- 2126

Bibliography

- [43] J.-L. Gervais and B. Sakita, *Field theory interpretation of supergauges in dual models*, Nucl. Phys. **B34** (1971) 632.
- [44] 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)].
- [45] 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.
- [46] E. Majorana, *Teoria simmetrica dell'elettrone e del positrone*, Nuovo Cim. **14** (1937) 171.
- [47] 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.
- [48] G. C. Branco, L. Lavoura, and J. P. Silva, *CP Violation*, Int. Ser. Monogr. Phys. **103** (1999) 1.
- [49] 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.
- [50] C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001.
- [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](https://arxiv.org/abs/1206.3525).
- [53] Belle collaboration, I. Adachi *et al.*, *Measurement of the CP violation parameters in $B^0 \rightarrow \pi^+ \pi^-$ decays*, Phys. Rev. **D88** (2013) 092003, [arXiv:1302.0551](https://arxiv.org/abs/1302.0551).
- [54] LHCb collaboration, R. Aaij *et al.*, *First measurement of time-dependent CP violation in $B_s^0 \rightarrow K^+ K^-$ decays*, JHEP **10** (2013) 183, [arXiv:1308.1428](https://arxiv.org/abs/1308.1428).
- [55] LHCb collaboration, *Measurement of the CKM angle γ from a combination of $B \rightarrow DK$ analyses*, LHCb-CONF-2016-001.
- [56] BaBar collaboration, J. P. Lees *et al.*, *Evidence of $B^+ \rightarrow \tau^+ \nu$ decays with hadronic B tags*, Phys. Rev. **D88** (2013) 031102, [arXiv:1207.0698](https://arxiv.org/abs/1207.0698).

- [57] Belle collaboration, I. Adachi *et al.*, *Evidence for $B^- \rightarrow \tau^- \bar{\nu}_\tau$ with a hadronic tagging method using the full data sample of Belle*, Phys. Rev. Lett. **110** (2013) 131801, [arXiv:1208.4678](https://arxiv.org/abs/1208.4678).
- [58] Belle collaboration, B. Kronenbitter *et al.*, *Measurement of the branching fraction of $B^+ \rightarrow \tau^+ \nu_\tau$ decays with the semileptonic tagging method*, Phys. Rev. **D92** (2015) 051102, [arXiv:1503.05613](https://arxiv.org/abs/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](https://arxiv.org/abs/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](https://arxiv.org/abs/hep-ph/0406184), updated results and plots available at <http://ckmfitter.in2p3.fr>.
- [61] T. Inami and C. S. Lim, *Effects of superheavy quarks and leptons in low-energy weak processes $K_L \rightarrow \mu \bar{\mu}$, $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ and $K^0 \leftrightarrow \bar{K}^0$* , Prog. Theor. Phys. **65** (1981) 297, Erratum *ibid.* **65** (1981) 1772.
- [62] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Annalen der Physik **384** (1926) 361.
- [63] 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.
- [64] A. J. Buras, W. Slominski, and H. Steger, *$B^0 \bar{B}^0$ mixing, CP violation and the B-meson decay*, Nucl. Phys. **B245** (1984) 369.
- [65] 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](https://arxiv.org/abs/1206.3525).
- [66] Belle collaboration, Y.-T. Duh *et al.*, *Measurements of branching fractions and direct CP asymmetries for $B \rightarrow K\pi$, $B \rightarrow \pi\pi$ and $B \rightarrow KK$ decays*, Phys. Rev. **D87** (2013) 031103, [arXiv:1210.1348](https://arxiv.org/abs/1210.1348).
- [67] 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](https://arxiv.org/abs/1304.6173).
- [68] LHCb collaboration, R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in B^0 - \bar{B}^0 mixing*, Phys. Rev. Lett. **114** (2015) 041601, [arXiv:1409.8586](https://arxiv.org/abs/1409.8586).

Bibliography

- [69] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP asymmetry in $B_s^0 - \bar{B}_s^0$ mixing*, Phys. Rev. Lett. **117** (2016) 061803, [arXiv:1605.09768](#).
- [70] R. Fleischer, *Extracting γ from $B_{s(d)} \rightarrow J/\psi K_S$ and $B_{d(s)} \rightarrow D_{d(s)}^+ D_{d(s)}^-$* , Eur. Phys. J. **C10** (1999) 299, [arXiv:hep-ph/9903455](#).
- [71] R. Fleischer, *Exploring CP violation and penguin effects through $B_d^0 \rightarrow D^+ D^-$ and $B_s^0 \rightarrow D_s^+ D_s^-$* , Eur. Phys. J. **C51** (2007) 849, [arXiv:0705.4421](#).
- [72] L. Bel *et al.*, *Anatomy of $B \rightarrow D\bar{D}$ decays*, JHEP **07** (2015) 108, [arXiv:1505.01361](#).
- [73] ATLAS collaboration, G. Aad *et al.*, *The ATLAS experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [74] CMS collaboration, S. Chatrchyan *et al.*, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [75] F. Marcastel, *CERN's Accelerator Complex. La chaîne des accélérateurs du CERN*, Oct, 2013. General Photo.
- [76] ALICE collaboration, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.
- [77] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.
- [78] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, [arXiv:1412.6352](#).
- [79] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, [arXiv:1405.7808](#).
- [80] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, [arXiv:1311.3893](#).
- [81] S. Bachmann *et al.*, *Ageing in the LHCb outer tracker: Phenomenon, culprit and effect of oxygen*, Nucl. Instrum. Meth. **A617** (2010) 202.
- [82] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, [arXiv:1211.6759](#).
- [83] E. Guschin and S. V. Laptev, *Monte-carlo study of LHCb preshower*, Tech. Rep. LHCb-2000-030, CERN, Geneva, Oct, 2000.
- [84] 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.

- [85] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.
- [86] 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.
- [87] 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.
- [88] R. Aaij *et al.*, *The Brunel project*, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/>, Oct, 2016.
- [89] G. Barrand *et al.*, *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001) 45.
- [90] R. Aaij *et al.*, *The DaVinci project*, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/>, Oct, 2016.
- [91] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.
- [92] 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.
- [93] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.
- [94] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.
- [95] 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.
- [96] 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.
- [97] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.
- [98] R. Aaij *et al.*, *The Boole project*, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/>, Oct, 2016.

Bibliography

- 2264 [99] R. Aaij *et al.*, *The Moore project*, <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/>, Oct, 2016.
- 2265
- 2266 [100] LHCb collaboration, *Optimization and calibration of the LHCb flavour tagging performance using 2010 data*, LHCb-CONF-2011-003.
- 2267
- 2268 [101] M. Grabalosa and M. Musy, *Flavour Tagging developments within the LHCb experiment*, PhD thesis, Barcelona U., 2012, CERN-THESIS-2012-075.
- 2269
- 2270 [102] 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](https://arxiv.org/abs/1202.4979).
- 2271
- 2272 [103] LHCb collaboration, R. Aaij *et al.*, *B flavour tagging using charm decays at the LHCb experiment*, JINST **10** (2015) P10005, [arXiv:1507.07892](https://arxiv.org/abs/1507.07892).
- 2273
- 2274 [104] 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.
- 2275
- 2276
- 2277 [105] LHCb collaboration, R. Aaij *et al.*, *New algorithms for identifying the flavour of B^0 mesons using pions and protons*, [arXiv:1610.06019](https://arxiv.org/abs/1610.06019), submitted to Eur. Phys. J.
- 2278
- 2279
- 2280 [106] A. Hoecker *et al.*, *TMVA: Toolkit for multivariate data analysis*, PoS **ACAT** (2007) 040, [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039).
- 2281
- 2282 [107] F. Pedregosa *et al.*, *Scikit-learn: Machine learning in Python*, J. Mach. Learn. Res. **12** (2011) 2825.
- 2283
- 2284 [108] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.
- 2285
- 2286
- 2287 [109] C. Gini, *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche.*, Tipogr. di P. Cuppini, 1912.
- 2288
- 2289 [110] L. Breiman, *Random forests*, Machine Learning **45** (2001) 5.
- 2290
- 2291 [111] 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](https://arxiv.org/abs/physics/0408124).
- 2292
- 2293 [112] 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.
- 2294
- 2295
- 2296 [113] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*, Ann. Statist. **29** (2001) 1189.
- 2297

- [114] B. H. Denby, *Neural networks and cellular automata in experimental high-energy physics*, Comput. Phys. Commun. **49** (1988) 429.
- [115] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, [arXiv:physics/0402083](https://arxiv.org/abs/physics/0402083).
- [116] 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](https://arxiv.org/abs/physics/0308063).
- [117] Y. Xie, *Formalism for simulation-based optimization of measurement errors in high energy physics*, [arXiv:0901.3305](https://arxiv.org/abs/0901.3305).
- [118] C. de Boor, *A practical guide to splines*, vol. 27 of *Applied Mathematical Sciences*, Springer-Verlag New York, 1978.
- [119] O. Behnke, K. Kröniger, T. Schörner-Sadenius, and G. Schott, eds., *Data analysis in high energy physics*, Wiley-VCH, Weinheim, Germany, 2013.
- [120] W. Verkerke and D. Kirkby, *Roofit*. <http://roofit.sourceforge.net>.
- [121] J. R. Klein and A. Roodman, *Blind analysis in nuclear and particle physics*, Ann. Rev. Nucl. Part. Sci. **55** (2005) 141.
- [122] H. G. Moser and A. Roussarie, *Mathematical methods for $B^0\bar{B}^0$ oscillation analyses*, Nucl. Instrum. Meth. **A384** (1997) 491.
- [123] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001.
- [124] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B^0 \rightarrow J/\psi K_S^0$ decays*, Phys. Rev. Lett. **115** (2015) 031601, [arXiv:1503.07089](https://arxiv.org/abs/1503.07089).
- [125] 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.
- [126] D. M. Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, [arXiv:1312.5000](https://arxiv.org/abs/1312.5000).
- [127] W. Fetscher *et al.*, *Regeneration of arbitrary coherent neutral kaon states: A new method for measuring the $K^0 - \bar{K}^0$ 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](https://arxiv.org/abs/1006.1938).
- [128] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation in $B \rightarrow D^+ D^-$ decays*, [arXiv:1608.06620](https://arxiv.org/abs/1608.06620), submitted to Phys. Rev. Lett.

Bibliography

- 2332 [129] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics,
2333 Krakow, 1986, DESY-F31-86-02.
- 2335 [130] S. Ali *et al.*, *Measurement of time-dependent CP-violation observables in*
2336 $B_s^0 \rightarrow D_s^\mp K^\pm$, LHCb-ANA-2012-068.
- 2337 [131] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the*
2338 *Velo*, LHCb-PUB-2011-001.
- 2339 [132] LHCb collaboration, R. Aaij *et al.*, *Measurement of the \bar{B}^0 - B^0 and \bar{B}_s^0 - B_s^0*
2340 *production asymmetries in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B739**
2341 (2014) 218, [arXiv:1408.0275](#).
- 2342 [133] T. Brambach *et al.*, *Measurement of CP violation in the time-dependent*
2343 *analysis of $B^0 \rightarrow J/\psi K_S^0$ decays with the 2010 data*, LHCb-ANA-2011-004.
2344 LHCb-CONF-2011-004.
- 2345 [134] P. Gandini, *Measurement of the B_s^0 meson lifetime in the flavour specific*
2346 *hadronic decay $B_s^0 \rightarrow D_s^- \pi^+$* , LHCb-ANA-2014-015.
- 2347 [135] M. Schellenberg, *Messung der zerfallszeitabhängigen CP-Asymmetrie im Zer-*
2348 *fall $B^0 \rightarrow D^{*+} D^-$ mit dem LHCb Experiment*, Master's thesis, TU Dortmund
2349 University, Sep, 2015.
- 2350 [136] Y. Amhis *et al.*, *b-hadrons absolute lifetime measurements*, LHCb-ANA-2011-
2351 055. Linked to LHCb-PAPER-2013-065.
- 2352 [137] T. M. Karbach and M. Schlupp, *Constraints on yield parameters in extended*
2353 *maximum likelihood fits*, [arXiv:1210.7141](#).
- 2354 [138] OPAL collaboration, K. Ackerstaff *et al.*, *Investigation of CP violation in $B^0 \rightarrow J/\psi K_S^0$ decays at LEP*, Eur. Phys. J. **C5** (1998) 379,
2355 [arXiv:hep-ex/9801022](#).
- 2357 [139] ALEPH collaboration, R. Barate *et al.*, *Study of the CP asymmetry of $B^0 \rightarrow J/\psi K_S^0$ decays in ALEPH*, Phys. Lett. **B492** (2000) 259,
2358 [arXiv:hep-ex/0009058](#).
- 2360 [140] CDF collaboration, T. Affolder *et al.*, *A measurement of $\sin(2\beta)$ from*
2361 *$B \rightarrow J/\psi K_S^0$ with the CDF detector*, Phys. Rev. **D61** (2000) 072005,
2362 [arXiv:hep-ex/9909003](#).
- 2363 [141] BaBar collaboration, B. Aubert *et al.*, *Measurement of time-dependent*
2364 *CP asymmetry in $B^0 \rightarrow c\bar{c}K^{(*)0}$ decays*, Phys. Rev. **D79** (2009) 072009,
2365 [arXiv:0902.1708](#).

- 2366 [142] Belle collaboration, I. Adachi *et al.*, *Precise measurement of the CP violation*
 2367 *parameter $\sin 2\phi_1$ in $B^0 \rightarrow (c\bar{c})K^0$ decays*, Phys. Rev. Lett. **108** (2012) 171802,
 2368 [arXiv:1201.4643](https://arxiv.org/abs/1201.4643).
- 2369 [143] Belle collaboration, M. Rohrken *et al.*, *Measurements of branching fractions*
 2370 *and time-dependent CP violating asymmetries in $B^0 \rightarrow D^{(*)\pm}D^\mp$ decays*, Phys.
 2371 Rev. **D85** (2012) 091106, [arXiv:1203.6647](https://arxiv.org/abs/1203.6647).
- 2372 [144] BaBar collaboration, B. Aubert *et al.*, *Measurements of time-dependent CP*
 2373 *asymmetries in $B^0 \rightarrow D^{(*)+}D^{(*)-}$ decays*, Phys. Rev. **D79** (2009) 032002,
 2374 [arXiv:0808.1866](https://arxiv.org/abs/0808.1866).
- 2375 [145] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CP-violating*
 2376 *phase ϕ_s in $\bar{B}_s^0 \rightarrow D_s^+ D_s^-$ decays*, Phys. Rev. Lett. **113** (2014) 211801,
 2377 [arXiv:1409.4619](https://arxiv.org/abs/1409.4619).
- 2378 [146] *LHCb: Technical Proposal*, Tech. Proposal, CERN, Geneva, 1998.
- 2379 [147] LHCb collaboration, *LHCb Trigger and Online Technical Design Report*,
 2380 CERN-LHCC-2014-016. LHCb-TDR-016.