technische universität
dortmund

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

Dissertation zur Erlangung des akademischen Grades

Dr. rer. nat.

vorgelegt von

Alex Birnkraut

Fakultät Physik
Technische Universität Dortmund

Dortmund, August 2018

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

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

Datum des Einreichens der Arbeit: 13.08.2018
Datum der mündlichen Prüfung: 16.10.2018

Abstract

A measurement of the *CP* asymmetries S_f and $S_{\bar{f}}$ in $B^0 \to D^{\mp}\pi^{\pm}$ decays is presented in this thesis. The analysed data set was recorded by the LHCb detector in proton-proton collision at centre-of-mass energies of 7 and 8 TeV in the years 2011 and 2012 and corresponds to an integrated luminosity of $3\,\text{fb}^{-1}$. The measured *CP* asymmetries are

$$S_f = 0.058 \pm 0.020\,(\text{stat.}) \pm 0.011\,(\text{syst.})\,,$$
$$S_{\bar{f}} = 0.038 \pm 0.020\,(\text{stat.}) \pm 0.007\,(\text{syst.})\,.$$

These results are more precise than and in agreement with previous determinations from the Belle and BaBar collaborations. Using these values, constraints are placed on the CKM angle γ and the CKM quantity $\sin(2\beta+\gamma)$. The obtained confidence intervals are consistent with the current world-average values.

Kurzfassung

In dieser Arbeit wird eine Messung der *CP*-Asymmetrien S_f und $S_{\bar{f}}$ in $B^0 \to D^{\mp}\pi^{\pm}$-Zerfällen präsentiert. Der analysierte Datensatz enthält dabei Ereignisse aus Proton-Proton-Kollisionen bei Schwerpunktsenergien von 7 und 8 TeV, die in den Jahren 2011 und 2012 vom LHCb-Detektor aufgenommen wurden und entspricht einer integrierten Luminosität von $3\,\text{fb}^{-1}$. Die *CP*-Asymmetrien sind gemessen als

$$S_f = 0.058 \pm 0.020\,(\text{stat.}) \pm 0.011\,(\text{syst.})\,,$$
$$S_{\bar{f}} = 0.038 \pm 0.020\,(\text{stat.}) \pm 0.007\,(\text{syst.})\,.$$

Diese Ergebnisse sind präziser und in Übereinstimmung mit bisherigen Messungen der Belle- und BaBar-Kollaborationen. Unter Ausnutzung dieser Ergebnisse werden Konfidenzintervalle für den CKM Winkel γ und die CKM Größe $\sin(2\beta+\gamma)$ bestimmt. Die ermittelten Intervalle stimmen mit den aktuellen Weltmittelwerten überein.

Contents

1 Introduction

The aim of particle physics is to understand the fundamental constituents of matter and their interactions. The theoretical model describing this is the so-called Standard Model (SM) of particle physics, established in the 1970s. During the last 40 years, all predicted particles such as the heavy t quark and the tau neutrino have been observed experimentally [1–3]. The SM was completed in 2012 when the Higgs Boson as the last missing particle was discovered [4, 5]. However, as successful as the SM is on the smallest observable scales, it fails to describe several macroscopic observations and phenomena: neither gravity is included, which is negligible in the interactions of elementary particles, nor the clear astronomical hints of dark matter and energy can be explained in the scope of the SM [6, 7]. Furthermore, the matter-antimatter asymmetry, which is observed in today's universe, is not accounted for in the SM. According to the Big-Bang theories, matter and antimatter were generated in equal parts. However, today only galaxies and clusters of matter can be observed.

In 1960, Andrei Sakharov formulated three necessary criteria for this asymmetry [8]: 1) violation of the baryon number conservation, 2) interactions out of the thermal equilibrium and 3) violation of the C and CP symmetries, *i.e.* particles and antiparticles behave differently even when inverting the spatial coordinates. So far, the baryon number is observed to be an extremly strong symmetry of nature, yielding in a lower bound for the proton lifetime of roughly 10^{34} years [9] - greater than the age of the universe. Departures from the thermal equilibrium are assumed to have occurred during the early development of the universe [10]. The violation of C and CP symmetry is allowed in the SM and was observed experimentally by the Wu and Fitch-Cronin experiments, respectively [11, 12]. However, the magnitude of this last effect is not large enough to explain the asymmetry in the universe [13] and therefore hints to physics beyond the SM, referred to as New Physics (NP).

In the SM, CP violation is possible in the strong and weak interactions, yet it is only observed in the latter one, stemming from the single complex phase of the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix [14]. This matrix is unitary by construction, what can be used for a strong self test of the SM. The unitarity can be represented graphically by a triangle in the complex plane. Determining the sides and angles of this triangle in independent measurements allows to overconstrain the position of the apex and to check if the triangle closes. One important set of measurements is the determination of the CKM angle γ. This angle is the least well known angle of the triangle and is the only angle that can be measured using both tree-level and loop processes. In this thesis constraints are

placed on the CKM quantity $\sin(2\beta + \gamma)$ and the CKM angle γ using a time-dependent *CP* violation measurement of $B^0 \to D^{\mp}\pi^{\pm}$ decays. Hereby, a focus is on the large number of signal candidates, resulting in the challenge to control small experimental effects, such as the kinematic differences between the flavour-tagging control modes and the $B^0 \to D^{\mp}\pi^{\pm}$ decay mode.

The analysed data set was recorded by the LHCb experiment located at the Large Hadron Collider (LHC) at CERN. The LHC, a circular proton-proton collider, is currently the largest and most powerful particle accelerator in the world. The LHCb experiment is designed to measure processes involving b and c hadrons in the forward direction. The two main fields are the determination of decay widths of rare B decays and the precise measurement of *CP* violation. One of the key challenges to measure *CP* violation time-dependently is to determine the production flavour of the B mesons in the harsh hadronic environment at the LHC.

The analysis described in this thesis was performed in collaboration between the LHCb groups from Dortmund and Lausanne. To present the complete analysis, also the contributions from Vincenzo Battista and Conor Fitzpatrick are described; the corresponding parts are indicated throughout the thesis. Apart from this direct contributions, also Julian Wishahi and Mirco Dorigo gave helpful input at many stages of the analysis.

The document is structured as follows: in Ch. 2 a short overview about the fundamental particles and interactions is given and the CKM matrix is introduced. The formalism of *CP* violation is explained and the manifestations of *CP* violation in the B meson sector are presented in Ch. 3, followed by a introduction and comparison of the experimental techniques to measure the CKM angle γ in Ch. 4. Subsequently, the LHCb detector is presented in Ch. 5, experimental techniques used in the analysis , including the flavour-tagging algorithms providing the initial B meson flavour at LHCb, are introduced in Ch. 6 and the analysed sample and the selection of signal candidates are described in Ch. 7. In Ch. 8 the mass fit to statistically separate signal and background candidates is detailed. Next, the flavour tagging strategy of the analysis with the training and calibration of the algorithms is described in Ch. 9. These ingredients are used in the decay-time fit, which is presented in Ch. 10, together with several performed cross-checks. As the last analysis steps, the estimation of systematic uncertainties is detailed in Ch. 11 followed by a summary of the measured *CP* asymmetries and their interpretation in terms of the CKM quantities γ and $\sin(2\beta + \gamma)$ in Ch. 12. Finally, a conclusion of the thesis is given in Ch. 13.

2 The standard model of particle physics

The following chapter gives an overview of fundamental particles and how they interact with each other. The elementary particles and forces are described following Refs. [15–17]. In the following, a short illustration how mediator particles emerge in the SM is given and discrete symmetries in the SM are introduced. Then, a more detailed discussion of the weak force is presented.

2.1 Fundamental particles and forces

The SM is a relativistic quantum field theory in which particles are produced and destroyed with fields $\phi(x)$ and the dynamics are described through Lagrangians $\mathscr{L}\left(\phi(x), \partial_\mu \phi(x)\right)$. In total twelve fundamental particles with half-integer spin exist: six quarks and six leptons. These twelve so-called fermions form all matter. Forces between the fermions are mediated by bosons which have integer spin. The masses of all these particles arise due to couplings to the Higgs field. These couplings are mediated by the so-called Higgs boson. A graphical representation of all fundamental particles is shown in Fig. 2.1.

Both quarks and leptons are classified in three families, where each family comprises a duplet of two particles. Further, the quarks are divided into up- and down-type quarks. The up-type quarks are the up (u), charm (c) and top quark (t) having an electric charge of $+\frac{2}{3}e$, the down-type quarks are the down (d), strange (s) and bottom quark (b) carrying a charge of $-\frac{1}{3}e$. The six leptons are classified by their electric charge. The electron (e^-), muon (μ^-) and tauon (τ^-) are negatively charged ($-1e$), whereas the corresponding neutrinos (ν_e, ν_μ, ν_τ) are uncharged. All twelve fermions have an antiparticle with opposite charge. The differentiation between particles and antiparticles is also denoted as flavour.

As previously mentioned, the fundamental forces in the SM are mediated by particles with integer spin. The so-called gauge bosons can be directly associated with these fundamental forces.

The force-carrier particles of the strong interaction are the eight massless gluons (g), which couple to the so-called colour charge. The only particles beside gluons carrying colour are the quarks. In contrast to the electrical charge, colour does not have only one but in total three pairs of distinct charges: red, green, blue and three corresponding anti-colours. Due to a phenomenon called confinement quarks and gluons cannot exist as isolated coloured particles but have to form bounded colourless states, *i.e.* quarks can

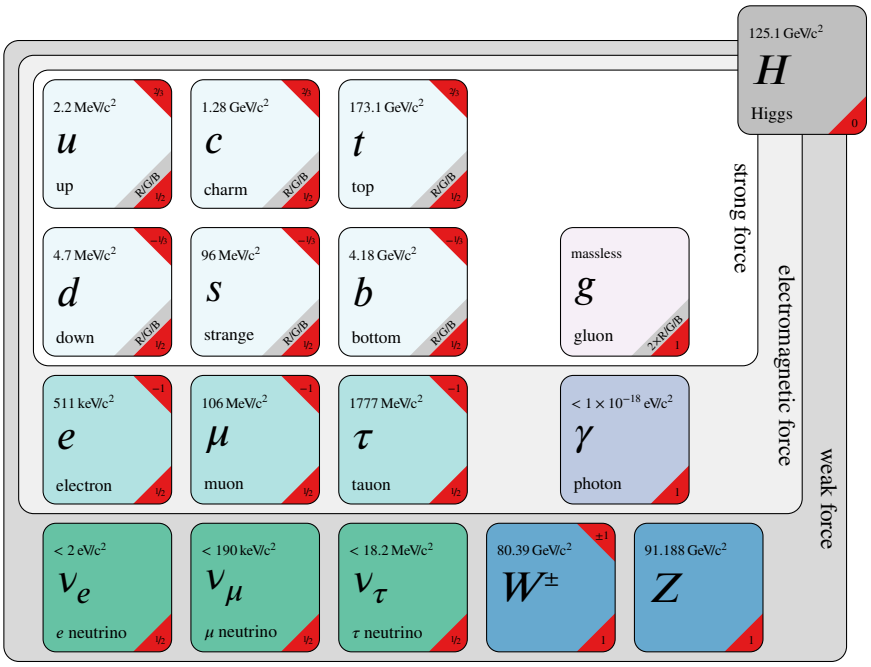

Fig. 2.1: Fundamental particles and forces of the SM. For each particle its spin and its possible colours are given in the bottom right corner. The top right corner shows the electric charge and the particles masses are shown in the top left part of the boxes. All numerical values are taken from [18].

only be observed in multi-quark states. Most commonly, three quarks (anti-quarks) form a baryon (anti-baryon), where each quark carries one of the three colours, or a quark and an anti-quark form a meson, where the anti-quark carries the anti-colour of the corresponding quark colour. As gluons carry both, a colour and an anticolour, gluonic self-couplings are allowed in the SM.

The electromagnetic force is mediated by the photon (γ) which couples to the electric charge. Accordingly, the only fermions not affected by the electromagnetic force are the uncharged neutrinos. Photons are also uncharged and thus do not couple to themselves.

The third interaction described in the SM is the weak interaction. The associated particles are the uncharged Z boson and the charged $W^{\pm}$ bosons. In contrast to the gluons and the photon, these are massive particles with masses of $M_W \approx 80$ GeV/c^2 [18] and $M_Z \approx 91$ GeV/c^2 [18]. They couple to all twelve fermions.

The last gauge boson is the Higgs boson (H), which is associated to the Higgs field and was discovered in 2012 [4, 5]. It has a mass of $M_H \approx 125$ GeV/c^2 [18] and interacts with all massive particles.

2.2 Symmetries in the standard model

Symmetries play an important role in the field of high energy physics, *e.g.* because of their relation to conservation laws (Noether's theorem). When discussing symmetries in the SM, one needs to distuingish between continous gauge symmetries and discrete symmetries. Requiring local invariance of the continous gauge symmetries leads to the interactions in the corresponding groups $U(1)$ (electromagnetic interaction), $SU(2)$ (weak interaction) and $SU(3)$ (strong interaction). In this instance the gauge bosons act as generators of the gauge transformation. This is exemplified for the $U(1)$ group, where the Lagrangian

$$\mathscr{L} = \overline{\psi}\left(i\partial\!\!\!/ - m\right)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \underbrace{e\overline{\psi}\gamma^{\mu}\psi}_{j^{\mu}}A_{\mu} \tag{2.1}$$

is invariant under the transformation

$$\psi \rightarrow \psi' = e^{-ie\theta(x)}\psi, \tag{2.2}$$

$$A_{\mu} \rightarrow A'_{\mu} = A_{\mu} + \partial_{\mu}\theta. \tag{2.3}$$

Interpreting the gauge field A_{μ} as the photon, the interaction term $j^{\mu}A_{\mu}$ can be identified. For the $SU(2)$ and $SU(3)$ groups equivalent transformations yield the $W^{\pm}$ and Z bosons and the gluons, respectively.

Furthermore there are also three discrete symmetries in the SM:

- The parity operator P is unitary ($P^{\dagger} = P^{-1}$) and reverses the momentum without flipping the spin of a particle. Hence, it describes spatial inversion $P\psi\left(t, \vec{x}\right) = \psi\left(t, -\vec{x}\right)$.

- The charge conjugation operator C is also unitary and transforms particles into their corresponding antiparticles. Its name is slightly misleading as the operator reverses not only the electric charge but also changes the sign of all internal quantum numbers.

- The third discrete symmetry is the time reversal T. It changes the sign of all temporal components $T\psi\left(t, \vec{x}\right) = \psi\left(-t, \vec{x}\right)$. Contrary to P and C, the time reversal operator is not unitary but antiunitary, *i.e.* $T^2 = 1$.

All discrete symmetries are conserved by the electromagnetic interaction while the weak and strong interactions can break them both individually and in combination with one other discrete symmetry (CP, PT, CT) in the SM. However, the breaking of the discrete symmetries has so far only been observed by the weak interaction. Moreover, based on the CPT theorem, the combination of all operations is an exact symmetry also for the weak and strong interactions. It assures that particles and antiparticles have the same invariant masses and lifetimes.

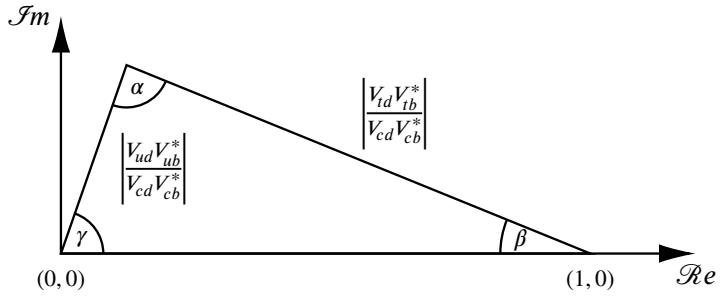

Fig. 2.2: CKM triangle in the complex plane.

2.3 The unitarity triangle

As explained previously, the weak interaction plays a special role in the SM by breaking the discrete symmetries. Consequently, the eigenstates to the weak interaction are not the same as the mass eigenstates. Under the assumption of massless neutrinos, this can be solved by a simple matrix rotation such that the eigenstates to the weak interaction and the mass are the same for the charged leptons. On the other hand, this is not possible for the up- and downtype quarks at the same time as none of them is massless. By convention the weak eigenstates of the downtype quarks *d*', *s*' and *b*' are chosen to be mixtures of their mass eigenstates *d*, *s* and *b*:

$$\begin{aligned}\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \\ &\approx \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}. \end{aligned} \tag{2.4}$$

This transformation matrix, denoted as CKM matrix [14, 19], has four degrees of freedom and is unitary by construction. As shown in Eq. (2.4), the matrix elements can be parametrised in the Wolfenstein parametrisation [20] with three real parameters $A \approx 0.81$, $\lambda \approx 0.22$, $\rho \approx 0.13$ [18] and one complex phase $\eta \approx 0.36$ [18]. It can be seen that the matrix elements become smaller with greater distance to the diagonal and therefore transitions between the quark families are suppressed.

As a consequence of the unitarity of the matrix, its elements are subject to the following constraints:

$$\sum_i V_{ij} V_{ik}^* = \delta_{jk} \quad \text{and} \quad \sum_j V_{ij} V_{kj}^* = \delta_{ik}. \tag{2.5}$$

The equations for which $j \neq k$ ($i \neq k$) can be represented as triangles in the complex plane, which are of great importance in modern particle physics. Their angles and sides can be

measured experimentally, so that the triangles can be overconstrained and the unitarity of the matrix can be tested. Experimental measurements of a non-closing triangle would be a clear sign of physics beyond the SM. Additionally, the complex phase of the CKM matrix is the only known source of *CP* violation in the SM. Therefore, the size of the triangles is a measure for the theoretically described size of *CP* violation as it is proportional to the Jarlskog Invariant J [21]. The most commonly used triangle is defined by the equation

$$V_{ud}V_{ub}^{*} + V_{cd}V_{cb}^{*} + V_{td}V_{tb}^{*} = 0\,, \tag{2.6}$$

as all terms are of similar size $\propto \lambda^3$ and can be measured from B meson decays. After normalising with $V_{cd}V_{cb}^{*}$, the corresponding triangle

$$\frac{V_{ud}V_{ub}^{*}}{V_{cd}V_{cb}^{*}} + 1 + \frac{V_{td}V_{tb}^{*}}{V_{cd}V_{cb}^{*}} = 0 \tag{2.7}$$

is obtained as presented in Fig. 2.2. Its angles can be parametrised using the CKM matrix elements in the following way:

$$\begin{aligned} \alpha &\equiv \arg\left(-V_{td}V_{ub}V_{tb}^{*}V_{ud}^{*}\right)\,, \\ \beta &\equiv \arg\left(-V_{cd}V_{tb}V_{cb}^{*}V_{td}^{*}\right)\,, \\ \gamma &\equiv \arg\left(-V_{ud}V_{cb}V_{ub}^{*}V_{cd}^{*}\right)\,. \end{aligned} \tag{2.8}$$

3 *CP* violation in the *B* meson sector

Since in the SM all physical interactions are conserved under the *CPT* transformation, the violation of *CP* is equivalent to a violation of the *T* symmetry. As described in Sec. 2.2, the *T* operator is antiunitary and it therefore transforms numbers into their complex conjugate. Hence, *CP* transformations only affect the complex phases of the wave functions describing initial and final states. However, the absolute values of phases describing transitions between different states are not physically meaningful as the initial and final states can be rephased convention dependent. The physical meaningful quantities are the relative phase differences between coherent contributions to a transition as these are invariant under global phase transformation. There are three types of phases arising in transition amplitudes: *weak* phases, changing sign under *CP* transformation (*CP* odd), *strong* phases, which do not change sign under *CP* transformation (*CP* even) and *spurious* phases, which usually arise due to conventional phase transformations. The denotations *weak* and *strong* do not mean that the phases originate in weak or strong interactions but only describe their behaviour under *CP* transformation. *Spurious* phases are global and, for simplification, will be ignored in the following discussion as they do not originate from any dynamics. Consequently, the *CP* transformations of the initial and final states are defined with *weak* phases ξ_i and ξ_f as

$$\begin{aligned} CP\left|B^0\right\rangle &= e^{i\xi_i}\left|\overline{B}^0\right\rangle \quad CP\left|\overline{B}^0\right\rangle = e^{-i\xi_i}\left|B^0\right\rangle , \\ CP\left|f\right\rangle &= e^{i\xi_f}\left|\overline{f}\right\rangle \quad CP\left|\overline{f}\right\rangle = e^{-i\xi_f}\left|f\right\rangle . \end{aligned} \tag{3.1}$$

Using this notation the time evolution of neutral mesons is described in this chapter and subsecently the formalism is applied to the B^0-$\overline{B}^0$ mixing. Then, the main equations describing *CP* violation are derived and the three types of *CP* violation are discussed. More details on these topics can be found in Refs. [22, 23].

3.1 Time evolution of neutral mesons

As previously described in Sec. 2.3 the mass eigenstates and the eigenstates of the weak interaction are not identical for quarks. The same applies for bound states of quarks like B mesons. Studying the system of a neutral particle P^0 and its antiparticle $\bar{P}^0$, the most general description to determine the time evolution is the Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix} P^0 \\ \bar{P}^0 \end{pmatrix} = H\begin{pmatrix} P^0 \\ \bar{P}^0 \end{pmatrix} = \left(m - \frac{i}{2}\Gamma\right)\begin{pmatrix} P^0 \\ \bar{P}^0 \end{pmatrix} \tag{3.2}$$

with m and Γ being hermitian 2x2 matrices. Hence, the matrix H is not hermitian and allows neutral particles P^0 to decay and not only to oscillate. In possible transitions, virtual intermediate states contribute to the matrix M, while real physical states to which P^0 and $\bar{P}^0$ decay contribute to the matrix Γ. Due to the CPT theorem, particles and antiparticles have the same masses and decay widths and the following constraints apply for the matrix elements:

$$\begin{aligned} m_{11} = m_{22} \equiv m\,, \quad m_{12} = m_{21}^*\,, \\ \Gamma_{11} = \Gamma_{22} \equiv \Gamma\,, \quad \Gamma_{12} = \Gamma_{21}^*\,. \end{aligned} \tag{3.3}$$

Interpreting P^0 and $\bar{P}^0$ as two states distinguished by an internal quantum number N_q the matrix elements can be classified by certain types of transitions: transitions with $\Delta N_q = 1$ are driven by the diagonal elements while the off-diagonal elements describe transitions with $\Delta N_q = 2$. These $\Delta N_q = 2$ processes include so-called particle-antiparticle oscillations.

To solve Eq. (3.2) and infer the time evolution of the initial states P^0 and $\bar{P}^0$, the matrix H needs to be diagonalised to obtain the mass eigenstates and the corresponding eigenvalues. These eigenstates can have different masses and lifetimes, however, the absolute sign of the mass difference Δm or decay-width difference $\Delta\Gamma$ has no physical meaning as interchanging the two eigenstates would lead to $\Delta m \to -\Delta m$ and $\Delta\Gamma \to -\Delta\Gamma$. Instead, only the relative sign between both quantities is of physical interest. With regard to the B^0 meson system, in which the eigenstates have quite different masses, in the following the mass eigenstates are denoted with P_H and P_L, referring to the heavier and lighter eigenstate, respectively. Using

$$F = \sqrt{\left(m_{12} - \frac{i}{2}\Gamma_{12}\right)\left(m_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)} \tag{3.4}$$

the eigenvalues can be expressed as

$$\begin{aligned} \mu_\mathrm{H} = m_\mathrm{H} - \frac{i}{2}\Gamma_\mathrm{H} = m + \mathscr{R}e(F) - \frac{i}{2}\left(\Gamma - 2\mathscr{I}m(F)\right)\,, \\ \mu_\mathrm{L} = m_\mathrm{L} - \frac{i}{2}\Gamma_\mathrm{L} = m - \mathscr{R}e(F) - \frac{i}{2}\left(\Gamma + 2\mathscr{I}m(F)\right) \end{aligned} \tag{3.5}$$

with the eigenstates

$$\begin{aligned} \left|P_\mathrm{H}\right\rangle &= p\left|P^0\right\rangle + q\left|\bar{P}^0\right\rangle\,, \\ \left|P_\mathrm{L}\right\rangle &= p\left|P^0\right\rangle - q\left|\bar{P}^0\right\rangle\,. \end{aligned} \tag{3.6}$$

The parameters p and q are constrained to fulfil $|p|^2 + |q|^2 = 1$ by construction and their ratio $\frac{q}{p}$ can be expressed in terms of the matrix elements:

$$\frac{q}{p} = \sqrt{\frac{m_{12}^* - \frac{i}{2}\Gamma_{12}^*}{m_{12} - \frac{i}{2}\Gamma_{12}}} = \frac{\Delta m - \frac{i}{2}\Delta\Gamma}{2\left(m_{12} - \frac{i}{2}\Gamma_{12}\right)} \tag{3.7}$$

where Δm and $\Delta\Gamma$ are the differences of the masses and decay widths of the eigenstates of the weak interaction defined as

$$\Delta m = m_{\mathrm{H}} - m_{\mathrm{L}} = 2\mathscr{R}e(F) \qquad \text{and} \qquad \Delta\Gamma = \Gamma_{\mathrm{L}} - \Gamma_{\mathrm{H}} = 4\mathscr{I}m(F)\,. \tag{3.8}$$

This definition is chosen such that it matches the convention used by the Heavy Flavour Averaging Group (HFLAV) [24].

After diagonalising the Schrödinger equation can be rewritten as

$$i\frac{d}{dt}\begin{pmatrix} P_{\mathrm{L}} \\ P_{\mathrm{H}} \end{pmatrix} = \begin{pmatrix} \mu_{\mathrm{L}} & 0 \\ 0 & \mu_{\mathrm{H}} \end{pmatrix}\begin{pmatrix} P_{\mathrm{L}} \\ P_{\mathrm{H}} \end{pmatrix}, \tag{3.9}$$

using the mass eigenvalues from Eq. (3.5) and mass eigenstates from Eq. (3.6), which can be easily solved and leads to the time evolution of the mass eigenstates with simple exponential functions $P_{\mathrm{L,H}}(t) = e^{-i\mu_{\mathrm{L,H}}t}P_{\mathrm{L,H}}$. Inverting Eq. (3.6) the time evolution for the flavour eigenstates follows:

$$\begin{aligned} \left|P^0(t)\right\rangle &= \left|P^0\right\rangle g_+ + \frac{q}{p}\left|\bar{P}^0\right\rangle g_-\,, \\ \left|\bar{P}^0(t)\right\rangle &= \left|\bar{P}^0\right\rangle g_+ + \frac{p}{q}\left|P^0\right\rangle g_- \end{aligned} \tag{3.10}$$

with $g_\pm = \frac{1}{2}\left(e^{-i\mu_{\mathrm{H}}t} \pm e^{-i\mu_{\mathrm{L}}t}\right)$. The associated masses and decay widths of the eigenstates of the weak interaction can be written as

$$m = \frac{m_{\mathrm{H}} + m_{\mathrm{L}}}{2} \qquad \text{and} \qquad \Gamma = \frac{\Gamma_{\mathrm{H}} + \Gamma_{\mathrm{L}}}{2}\,. \tag{3.11}$$

3.2 B^0-$\bar{B}^0$ mixing

As described above, the mixing of the flavour eigenstates B_q^0 and $\overline{B}_q^0$ is characterised by the mass difference Δm, the decay-width difference $\Delta\Gamma$ and the ratio q/p. All of these quantities are connected to the off-diagonal matrix elements $m_{12} - i/2\,\Gamma_{12}$ and hence m_{12} and Γ_{12} must be calculated to further probe mixing phenomena. In the SM, transitions from B_q^0 to $\overline{B}_q^0$ mesons can only happen through $\Delta N_q = 2$ dynamics, which can be further separated into short-distance transitions (transitions at quark-level) and long-distance transitions (transitions at hadron-level).

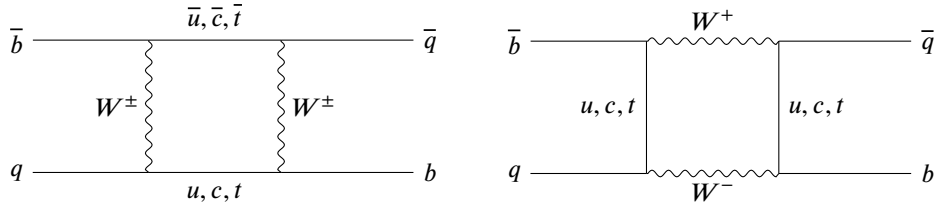

Fig. 3.1: Box diagrams of lowest order for B_q^0-$\overline{B}_q^0$-oscillations. Both diagrams are dominated by the t quark exchange[25].

Due to the large mass of the b quark, long-distance transitions are expected to be negligible for the B_q^0-$\overline{B}_q^0$ system in the SM.

Transitions at quark-level at lowest order can be represented by Feynman diagrams as shown in Fig. 3.1. The first corresponding matrix element m_{12} can be expressed as

$$m_{12} = -\frac{G_{\mathrm{F}}^2 M_W^2}{12\pi^2} f^2 m_{B_q^0} B \mathscr{F}^* \tag{3.12}$$

where G_{F} is the Fermi constant, M_W the W-boson mass, f the weak interaction constant and B the *bag* parameter, which describes strong interaction effects [22]. The quantity $\mathscr{F}$ sums over the different box diagrams containg a u-, c- or t quark, respectively. Using the short notation $\lambda_i = V_{ib}^* V_{iq}$, it can be written as

$$\mathscr{F} = \eta_1 \lambda_c^2 S_0(x_c) + \eta_2 \lambda_t^2 S_0(x_t) + 2\eta_3 \lambda_c \lambda_t S_0(x_c, x_t) \tag{3.13}$$

where λ_u is eliminated by λ_c and λ_t using Eq. (2.5). Here, η_i are QCD correction factors and S_0 are the Inami-Lin functions [26], which go with the up-type-quark masses through the ratio $x_q \equiv m_q^2/m_W^2$. In case of $t = d$ both λ_c and λ_t are of same magnitude λ^3, in case of $q = s$ both λ_c and λ_t are of magnitude λ^2. Hence, the summand containing $S_0\left(x_t\right)$ is dominant and with the replacement $\eta_2 = \eta_{B_q^0}$ one can approximate

$$\mathscr{F} \approx \eta_{B_q^0} \lambda_t^2 S_0(x_t) \approx m_t^2 \,. \tag{3.14}$$

The second matrix element Γ_{12} corresponding to the short-distance transitions is given by

$$\Gamma_{12} = \sum_f \left\langle f \middle| T \middle| B_q^0 \right\rangle^* \left\langle f \middle| T \middle| \overline{B}_q^0 \right\rangle \,. \tag{3.15}$$

Here f describes the possible physical states to which both B_q^0 and $\overline{B}_q^0$ decay. As the mass of the t quark is much larger than the mass of any B meson, B_q^0 and $\overline{B}_q^0$ cannot decay in any top-flavoured state. Therefore the contributing diagrams to Eq. (3.15) must be dominated by the available mass, *i.e.* by $m_{B_q^0}$.

Consequently, the off-diagonal matrix elements $m_{12} - i/2\Gamma_{12}$ are clearly dominated by m_{12} as

$$\left|\frac{\Gamma_{12}}{m_{12}}\right| \propto \frac{m_{B_q^0}^2}{m_t^2} \propto 10^{-3}\,. \tag{3.16}$$

This can be used to derive a prediction of the relative size of $\Delta\Gamma$ compared to Δm. The difference between the mass eigenvalues μ_{H} and μ_{L} can be expressed as

$$\Delta\mu = \mu_{\mathrm{H}} - \mu_{\mathrm{L}} = \Delta m - \frac{i}{2}\Delta\Gamma = 2F. \tag{3.17}$$

Squaring Eq. (3.17) and separating the real and imaginary parts leads to

$$\begin{aligned} \Delta m^2 - \frac{1}{4}\Delta\Gamma^{\,2} &= 4\left|m_{12}\right|^2 - \left|\Gamma_{12}\right|^2\,, \\ \Delta m\Delta\Gamma &= 4\mathscr{R}e\left(m_{12}^*\Gamma_{12}\right)\,. \end{aligned} \tag{3.18}$$

Taking into account the GIM enhancement [27] of m_{12} and the bound on Γ_{12} to be of order $m_{B_q^0}$ (Eq. (3.16)), this can be simplified to

$$\begin{aligned} \Delta m &\approx 2\left|m_{12}\right|\,, \\ \Delta\Gamma &\approx \frac{2\mathscr{R}e\left(m_{12}^*\Gamma_{12}\right)}{\left|m_{12}\right|}\,, \end{aligned} \tag{3.19}$$

which shows that for the B^0-system the decay width difference is expected to be much smaller than the mass difference.

Further, the ratio q/p from Eq. (3.7) can be expressed as

$$\frac{q}{p} \approx \frac{\left|m_{12}\right|}{m_{12}} = \frac{m_{12}^*}{\left|m_{12}\right|}\,, \tag{3.20}$$

by applying the reasoning from Eq. (3.16), *i.e.* the quantity q/p is a pure phase. Using Eq. (3.12), the ratio can be connected to the CKM matrix elements

$$\frac{q}{p} \approx -\frac{V_{tb}^* V_{tq}}{V_{tb} V_{tq}^*}\,. \tag{3.21}$$

As explained above, the CKM combination $V_{tb}^* V_{tq}$ appears here because the box diagrams for m_{12} shown in Fig. 3.1 are dominated by the top-quark contribution.

3.3 Master equations of *CP* violation

Using the time evolution presented in Sec. 3.1, one can also study the time evolution of decaying particles. In the following, the notation for the decay amplitudes

$$\begin{aligned} A_f &= \left\langle f \middle| T \middle| P^0 \right\rangle , \quad A_{\bar{f}} = \left\langle \bar{f} \middle| T \middle| P^0 \right\rangle , \\ \bar{A}_f &= \left\langle f \middle| T \middle| \bar{P}^0 \right\rangle , \quad \bar{A}_{\bar{f}} = \left\langle \bar{f} \middle| T \middle| \bar{P}^0 \right\rangle . \end{aligned} \tag{3.22}$$

is used. Denoting initially produced particles with $P^0(t)$ (Eq. (3.10)), the probability for the transition $\left|\left\langle f \middle| T \middle| P^0(t) \right\rangle\right|^2$ can be calculated as

$$\begin{aligned} \left|\left\langle f \middle| T \middle| P^0(t) \right\rangle\right|^2 &= \left| \left\langle f \middle| T \middle| P^0 \right\rangle g_+ + \frac{q}{p} \left\langle f \middle| T \middle| \bar{P}^0 \right\rangle g_- \right|^2 \\ &= \left|A_f\right|^2 \left| g_+ + \frac{q}{p}\frac{\bar{A}_f}{A_f} g_- \right|^2 = \left|A_f\right|^2 \left| g_+ + \lambda_f \, g_- \right|^2 \\ &= \left|A_f\right|^2 \left[g_+ g_+^* + \left|\lambda_f\right|^2 g_- g_-^* + \left(\lambda_f^* \, g_+ g_-^* + \lambda_f \, g_- g_+^* \right) \right] . \end{aligned} \tag{3.23}$$

In analogy the probabilites for an initially produced antiparticle $\bar{P}^0(t)$ and a second final state $\bar{f}$ are given by

$$\left|\left\langle f \middle| T \middle| \bar{P}^0(t) \right\rangle\right|^2 = \left|A_f\right|^2 \left|\frac{p}{q}\right|^2 \left[\left|\lambda_f\right|^2 g_+ g_+^* + g_- g_-^* + \left(\lambda_f^* g_- g_+^* + \lambda_f \, g_+ g_-^* \right) \right] , \tag{3.24}$$

$$\left|\left\langle \bar{f} \middle| T \middle| P^0(t) \right\rangle\right|^2 = \left|A_{\bar{f}}\right|^2 \left[g_+ g_+^* + \left|\lambda_{\bar{f}}\right|^2 g_- g_-^* + \left(\lambda_{\bar{f}}^* g_+ g_-^* + \lambda_{\bar{f}} \, g_- g_+^* \right) \right] , \tag{3.25}$$

$$\left|\left\langle \bar{f} \middle| T \middle| \bar{P}^0(t) \right\rangle\right|^2 = \left|A_{\bar{f}}\right|^2 \left|\frac{q}{p}\right|^2 \left[\left|\lambda_{\bar{f}}\right|^2 g_+ g_+^* + g_- g_-^* + \left(\lambda_{\bar{f}}^* g_- g_+^* + \lambda_{\bar{f}} \, g_+ g_-^* \right) \right] \tag{3.26}$$

where the quantities λ_f and $\lambda_{\bar{f}}$ are defined as

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} \quad \text{and} \quad \lambda_{\bar{f}} = \frac{q}{p}\frac{\bar{A}_{\bar{f}}}{A_{\bar{f}}} . \tag{3.27}$$

The transition probabilites can be expressed in terms of the physical quantities Δm, $\Delta\Gamma$ and Γ by using

$$g_\pm g_\pm^* = \frac{1}{2} e^{-\Gamma t} \left(\cosh\left(\frac{\Delta\Gamma}{2} t \right) \pm \cos(\Delta m t) \right) , \tag{3.28}$$

$$g_\pm^* g_\mp = \frac{1}{2} e^{-\Gamma t} \left(\sinh\left(\frac{\Delta\Gamma}{2} t \right) \pm i \sin(\Delta m t) \right) . \tag{3.29}$$

Substituting this in Eqs. (3.23) to (3.26), one obtains

$$\left|\left\langle f\left|T\right|P^0(t)\right\rangle\right|^2=\frac{1}{2}e^{-\Gamma t}\left|A_f\right|^2\left(1+\left|\lambda_f\right|^2\right)\quad\left[\cosh\left(\frac{\Delta\Gamma}{2}t\right)+A_f^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right)\right.$$
$$\left.-S_f\sin(\Delta mt)+C_f\cos(\Delta mt)\right], \tag{3.30}$$

$$\left|\left\langle f\left|T\right|\bar{P}^0(t)\right\rangle\right|^2=\frac{1}{2}e^{-\Gamma t}\left|A_f\right|^2\left(1+\left|\lambda_f\right|^2\right)\left|\frac{p}{q}\right|^2\left[\cosh\left(\frac{\Delta\Gamma}{2}t\right)+A_f^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right)\right.$$
$$\left.+S_f\sin(\Delta mt)-C_f\cos(\Delta mt)\right], \tag{3.31}$$

$$\left|\left\langle \bar{f}\left|T\right|P^0(t)\right\rangle\right|^2=\frac{1}{2}e^{-\Gamma t}\left|A_{\bar{f}}\right|^2\left(1+\left|\lambda_{\bar{f}}\right|^2\right)\quad\left[\cosh\left(\frac{\Delta\Gamma}{2}t\right)+A_{\bar{f}}^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right)\right.$$
$$\left.-S_{\bar{f}}\sin(\Delta mt)+C_{\bar{f}}\cos(\Delta mt)\right], \tag{3.32}$$

$$\left|\left\langle \bar{f}\left|T\right|\bar{P}^0(t)\right\rangle\right|^2=\frac{1}{2}e^{-\Gamma t}\left|A_{\bar{f}}\right|^2\left(1+\left|\lambda_{\bar{f}}\right|^2\right)\left|\frac{q}{p}\right|^2\left[\cosh\left(\frac{\Delta\Gamma}{2}t\right)+A_{\bar{f}}^{\Delta\Gamma}\sinh\left(\frac{\Delta\Gamma}{2}t\right)\right.$$
$$\left.+S_{\bar{f}}\sin(\Delta mt)-C_{\bar{f}}\cos(\Delta mt)\right] \tag{3.33}$$

where the coefficients in front of the trigonometric and hyperbolic functions are defined as

$$A_f^{\Delta\Gamma}=-\frac{2\mathscr{R}e\left(\lambda_f\right)}{1+\left|\lambda_f\right|^2},\qquad S_f=\frac{2\mathscr{I}m\left(\lambda_f\right)}{1+\left|\lambda_f\right|^2},\qquad C_f=\frac{1-\left|\lambda_f\right|^2}{1+\left|\lambda_f\right|^2}, \tag{3.34}$$

$$A_{\bar{f}}^{\Delta\Gamma}=-\frac{2\mathscr{R}e\left(\lambda_{\bar{f}}\right)}{1+\left|\lambda_{\bar{f}}\right|^2},\qquad S_{\bar{f}}=\frac{2\mathscr{I}m\left(\lambda_{\bar{f}}\right)}{1+\left|\lambda_{\bar{f}}\right|^2},\qquad C_{\bar{f}}=\frac{1-\left|\lambda_{\bar{f}}\right|^2}{1+\left|\lambda_{\bar{f}}\right|^2}. \tag{3.35}$$

These coefficients satisfy the conditions

$$S_f^2+C_f^2+A_f^{\Delta\Gamma 2}=1\quad\text{and}\quad S_{\bar{f}}^2+C_{\bar{f}}^2+A_{\bar{f}}^{\Delta\Gamma 2}=1\,. \tag{3.36}$$

Also, they are not necessarily constant over the whole phase space, *i.e.* for multibody decays the contributing phases originate from final state interactions as well (*i.e. strong* phases), which are not identical for different regions of phase space.

3.4 Classes of *CP* violation

Depending on the type of transition in which *CP* violation occurs, its manifestation is different. Transitions with $\Delta N_q = 1$ are affected by direct *CP* violation, transitions with $\Delta N_q = 2$ might be subject to *CP* violation in mixing and transitions affected by both $\Delta N_q = 1$ and $\Delta N_q = 2$ dynamics can be affected by interference *CP* violation in the interference of decay and decay after mixing. These three types of *CP* violation are described in more detail below.

3.4.1 Direct *CP* violation

Direct *CP* violation or *CP* violation in decay means that a specific decay amplitude differs between the decay of a particle into a specific final state and its corresponding antiparticle decay into the charge conjugated final state. It is the only type of *CP* violation which can occur for charged particles. In terms of the *CP* coefficients given in Eq. (3.34) and Eq. (3.35), this means that $C_f \neq C_{\bar{f}}$ or in case of neutral mesons, which decay into one common final state $C_f \neq 0$. Experimentally, direct *CP* violation can be measured with an asymmetry such as

$$A_{CP} = \frac{\left|\langle \bar{f} \left|T\right| \bar{P}\rangle\right|^2 - \left|\langle f \left|T\right| P\rangle\right|^2}{\left|\langle \bar{f} \left|T\right| \bar{P}\rangle\right|^2 + \left|\langle f \left|T\right| P\rangle\right|^2} = \frac{\left|\bar{A}_{\bar{f}}/A_f\right|^2 - 1}{\left|\bar{A}_{\bar{f}}/A_f\right|^2 + 1}\,. \tag{3.37}$$

Naively, one could expect that it is sufficient that one single amplitude contributes to a transition. For illustration one can consider a decay with just one amplitude

$$\begin{aligned} A_f &= A e^{i(\delta+\phi)}\,, \\ \bar{A}_{\bar{f}} &= A e^{i(\delta-\phi)} \end{aligned} \tag{3.38}$$

where A is a real positive number, ϕ is the *weak* phase and δ the *strong* phase. From Eq. (3.38) it immediately becomes obvious that the quantity $\left|\bar{A}_{\bar{f}}\right|^2 - \left|A_f\right|^2$ vanishes and therefore *CP* is conserved. When instead considering a decay with two contributing amplitudes with different *weak* and *strong* phases

$$\begin{aligned} A_f &= A_1 e^{i(\delta_1+\phi_1)} + A_2 e^{i(\delta_2+\phi_2)}\,, \\ \bar{A}_{\bar{f}} &= A_1 e^{i(\delta_1-\phi_1)} + A_2 e^{i(\delta_2-\phi_2)}\,, \end{aligned} \tag{3.39}$$

CP violation becomes possible if both the *weak* and the *strong* phases differ:

$$\left|A_f\right|^2 - \left|\bar{A}_{\bar{f}}\right|^2 = -4A_1A_2 \sin\left(\delta_1 - \delta_2\right) \sin\left(\phi_1 - \phi_2\right)\,. \tag{3.40}$$

For B mesons this has been measured by the LHCb experiment in the decay modes $B^0 \to K^+\pi^-$ and $B_s^0 \to K^-\pi^+$ [28] to be

$$\begin{aligned} A_{CP}\left(B^0 \to K^+\pi^-\right) &= -0.080 \pm 0.007\,(\text{stat.}) \pm 0.003\,(\text{syst.})\,, \\ A_{CP}\left(B_s^0 \to K^-\pi^+\right) &= 0.27 \pm 0.04\,(\text{stat.}) \pm 0.01\,(\text{syst.})\,, \end{aligned} \tag{3.41}$$

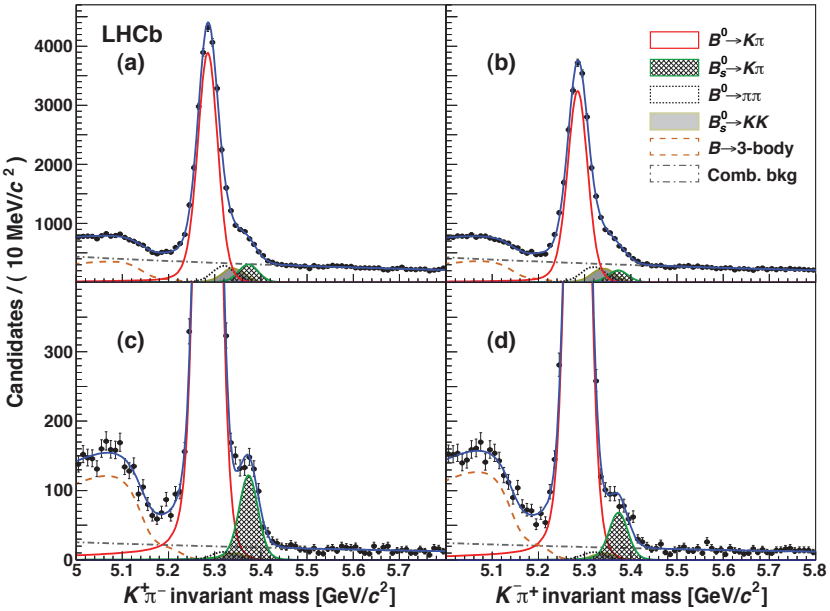

Fig. 3.2: Invariant mass spectra for B^0 (top) and B_s^0 (bottom) decays into $K^+\pi^-$ (left) and $K^-\pi^+$ (right). The result of the unbinned maximum likelihood fits are overlaid [28].

which corresponds to a statistical significance of 10.5σ and 6.5σ for the B^0 and the B_s^0 mode, respectively. Figure 3.2 shows the invariant mass spectra of B^0 and B_s^0 mesons decaying into $K^+\pi^-$.

3.4.2 *CP* violation in mixing

Indirect *CP* violation, also denoted as *CP* violation in mixing, implies that the transition probabilities for a meson P^0 to oscillate into its antiparticle state $\bar{P}^0$ and vice versa are different. Due to charge conservation, mixing is only possible for uncharged mesons and hence this type of *CP* violation cannot occur for charged particles. Using the time evolution from Eq. (3.10), the probabilities of *e.g.* initially produced P^0 ($\bar{P}^0$) mesons to have oscillated into a $\bar{P}^0$ (P^0) at proper-time t are

$$\left|\left\langle P^0 \middle| \bar{P}^0(t) \right\rangle\right|^2 = \frac{1}{4}\left|\frac{p}{q}\right|^2 \left(e^{-\Gamma_H t} + e^{-\Gamma_L t} - 2e^{-\Gamma t}\cos\left(\Delta m t\right)\right) , \tag{3.42}$$

$$\left|\left\langle \bar{P}^0 \middle| P^0(t) \right\rangle\right|^2 = \frac{1}{4}\left|\frac{q}{p}\right|^2 \left(e^{-\Gamma_H t} + e^{-\Gamma_L t} - 2e^{-\Gamma t}\cos\left(\Delta m t\right)\right) . \tag{3.43}$$

To obtain the same probabilities for both processes, the condition

$$\left|\frac{q}{p}\right| = \left|\frac{p}{q}\right| \Rightarrow \left|\frac{q}{p}\right| = 1 \tag{3.44}$$

has to be satisfied. According to Eq. (3.7), this means that indirect *CP* violation occurs if the matrix elements m_{12} and Γ_{12} have different complex phases. Using neutral B mesons as an example, the *CP* asymmetry in case of indirect *CP* violation is accordingly defined as

$$A_{CP} = \frac{\Gamma\left(B^0(t) \to \overline{B}^0\right) - \Gamma\left(\overline{B}^0(t) \to B^0\right)}{\Gamma\left(B^0(t) \to \overline{B}^0\right) + \Gamma\left(\overline{B}^0(t) \to B^0\right)} = \frac{1 - |p/q|^4}{1 + |p/q|^4} \,. \tag{3.45}$$

However, as neutral B mesons do not just oscillate but also decay this asymmetry cannot be used directly to measure *CP* violation in mixing. Instead, the B mesons need to be reconstructed in flavour specific decays, *i.e.* only the transitions $B^0 \to f$ and $\overline{B}^0 \to \bar{f}$, but not $B^0 \to \bar{f}$ and $\overline{B}^0 \to f$ are allowed. Thus, the flavour of the meson at decay can be inferred from the final state and compared to the initial production flavour. For the B^0 and B_s^0 meson system, *CP* violation in mixing has been measured to be negligible [24], which is in good agreement with the SM predictions (see Sec. 3.2).

3.4.3 Interference *CP* violation

So far *CP* violation arising from a clash between the phases of two interfering decay amplitudes or a clash between the phases of m_{12} and Γ_{12} has been discussed. The third possibility is a clash between the phase of q/p and the phase of the decay amplitude what results in the so-called *CP* violation in the interference of decay and decay after mixing. For this type of *CP* violation the initial particle P^0 and antiparticle $\overline{P}^0$ must decay into both the final state f and its *CP*-conjugate $\bar{f}$.

Inverting the requirement for *CP* violation in mixing shows that *CP* is conserved in mixing if there is a phase ξ' such that

$$\begin{aligned} m_{12}^* &= e^{2i\xi'} m_{12} \,, \\ \Gamma_{12}^* &= e^{2i\xi'} \Gamma_{12} \,, \end{aligned} \tag{3.46}$$

which leads directly to $q^2/p^2 = e^{2i\xi'}$ (Eq. (3.20)). Using Eq. (3.1) the *CP* conjugated amplitudes $\bar{A}_{\bar{f}}$ and $A_{\bar{f}}$ can be expressed as

$$\bar{A}_{\bar{f}} = e^{i(\xi_f - \xi_i)} A_f \,, \tag{3.47}$$

$$A_{\bar{f}} = e^{i(\xi_f + \xi_i)} \bar{A}_f \,. \tag{3.48}$$

yielding $\left| A_f \right| = \left| \bar{A}_{\bar{f}} \right|$ and $\left| \bar{A}_f \right| = \left| A_{\bar{f}} \right|$ after eliminating the phases and shows that these amplitudes are not subject to direct *CP* violation. Multiplying Eq. (3.47) and Eq. (3.48) gives the relation

$$A_f \, A_{\bar{f}} = e^{2i\xi_i} \, \bar{A}_{\bar{f}} \, \bar{A}_f \,. \tag{3.49}$$

Under the assumption that the phase ξ' of q/p and the *weak* phase ξ_i are the same, one finds *CP* conservation and

$$\arg\left(\frac{p^2}{q^2} A_f \bar{A}_f^* A_{\bar{f}} \bar{A}_{\bar{f}}^* \right) = 0 \tag{3.50}$$

applies. However, $\xi_i = \xi'$ is a very specific case and thus even without the presence of *CP* violation in decay or mixing, *CP* is not necessarily conserved. This also can be expressed using the parameters λ_f and $\lambda_{\bar{f}}$. In the case of *CP* conservation in decay or mixing $|\lambda_f| = |\lambda_{\bar{f}}| = \pm 1$ holds while *CP* is not conserved in case of

$$\arg(\lambda_f) + \arg(\lambda_{\bar{f}}) \neq 0\,. \tag{3.51}$$

This means the *CP* coefficients C_f and $C_{\bar{f}}$ are not affected by this type of *CP* violation, while for the coefficients $(S_f, S_{\bar{f}})$ and $(A_f^{\Delta\Gamma}, A_{\bar{f}}^{\Delta\Gamma})$ this condition can be reformulated to

$$S_f \neq -S_{\bar{f}} \quad \text{and} \quad A_f^{\Delta\Gamma} \neq A_{\bar{f}}^{\Delta\Gamma}\,. \tag{3.52}$$

In case that both, particle and antiparticle, decay into only one common final state these conditions simplifies to $\arg(\lambda_f) \neq 0$ and $S_f \neq 0$, $A_f^{\Delta\Gamma} \neq 0$.

CP violation in the interference of decay and decay after mixing was first observed by the *B*-factories BaBar [29] and Belle [30] in the so-called golden mode $B^0 \to J/\psi K_S^0$. This measurement is the most prominent one for this type of *CP* violation allowing to determine $\sin(2\beta)$. For $B^0 \to J/\psi K_S^0$ no *CP* violation in decay and mixing is expected and with the current experimental precision $\Delta\Gamma = 0$ can be assumed. Therefore, the *CP* asymmetry in this case can be expressed as

$$A_{CP}(t) = \frac{\Gamma(\bar{B}^0 \to J/\psi K_S^0) - \Gamma(B^0 \to J/\psi K_S^0)}{\Gamma(\bar{B}^0 \to J/\psi K_S^0) + \Gamma(B^0 \to J/\psi K_S^0)} \approx S_f \sin(\Delta m_d t) \tag{3.53}$$

where the parameter S_f can be identified with $\sin(2\beta)$. The most recent measurement of S_f was performed by LHCb [31] yielding a result of

$$S_f = 0.731 \pm 0.035\,(\text{stat.}) \pm 0.005\,(\text{syst.})\,, \tag{3.54}$$

and is consistent with the SM expectation. The resulting time-dependent signal yield asymmetry is shown in Fig. 3.3

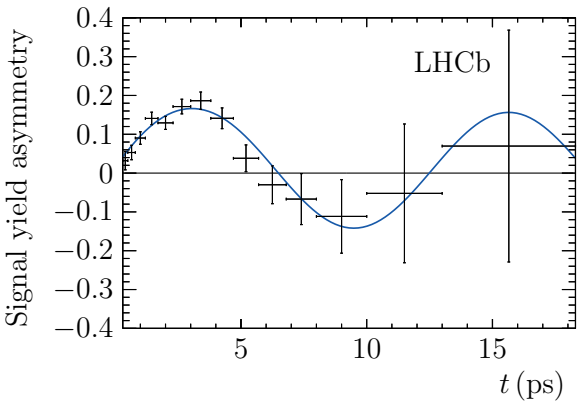

Fig. 3.3: Time-dependent signal yield asymmetry $\left(N_{\bar{B}^0} - N_{B^0}\right) / \left(N_{\bar{B}^0} + N_{B^0}\right)$ for the decay $B^0 \to J/\psi K^0_S$. The black points represent the data sample, the blue solid curve is the projection of the signal PDF [31].

4 The CKM angle γ

As mentioned before, overconstraining the CKM triangle relations following from the unitarity of the CKM matrix is a null test of the SM. The CKM angle γ is one of five observables (two sides, three angles) parametrising the CKM triangle described in Eq. (2.7). The current experimental constraints on this triangle are shown in Fig. 4.1. One can see that γ is currently the least well known parameter. Hence, a more accurate determination of γ is one of the main tasks of current research in the field of flavour physics.

This chapter is organised as follows: firstly, a description on how γ can be accessed in general is given in section Sec. 4.1, especially the determination using tree-level decays (Sec. 4.1.1) and loop-processes (Sec. 4.1.2) is emphasized, followed by the explanation how the decay mode $B^0 \to D^{\mp}\pi^{\pm}$ can be used to derive constraints on γ in Sec. 4.2.

4.1 Accessing the angle γ

From the form in Eq. (2.8) it can be seen that the angle γ is the sole angle which is independent of CKM elements involving the top quark. This makes γ the only angle, which can be determined from tree-level decays in a theoretically clean way. On the other hand, the experimental challenges are large as transitions sensitive to γ need to be proportional to V_{ub}, *i.e.* such transitions are highly suppressed by order $\mathcal{O}\left(\lambda^3\right)$. So either the precision of single measurements is limited due to small interference effects and branching fractions or in addition to the tree-level transition also penguin contributions of similar size contribute. These gluonic penguins usually carry a *weak* phase different from the one in the tree-level diagram.

The latter possibility is briefly discussed using the example of the $B_s^0 \to \rho K_S^0$ decay. Similar to the case of the golden mode $B^0 \to J/\psi K_S^0$, the final state ρK_S^0 is a *CP*-eigenstate and only the parameter λ_f needs to be calculated. Using the amplitudes

$$\begin{aligned} A_f &= \left\langle \rho K_S^0 \middle| T \middle| B_s^0 \right\rangle = -\frac{1}{2q_K}\left\langle \rho \overline{K}^0 \middle| T \middle| B_s^0 \right\rangle \propto -\frac{1}{2q_K} V_{ub}^* V_{ud}\,, \\ \bar{A}_f &= \left\langle \rho K_S^0 \middle| T \middle| \overline{B}_s^0 \right\rangle = \frac{1}{2p_K}\left\langle \rho K^0 \middle| T \middle| \overline{B}_s^0 \right\rangle \propto \frac{1}{2p_K} V_{ub} V_{ud}^* \end{aligned} \tag{4.1}$$

where q_K and p_K are the mixing parameters for the neutral kaon system, similar to what is shown in Eq. (3.7) for the B^0 meson system. With $q_K/p_K = V_{cs}^* V_{cd}/V_{cs} V_{cd}^*$ [22] the parameter

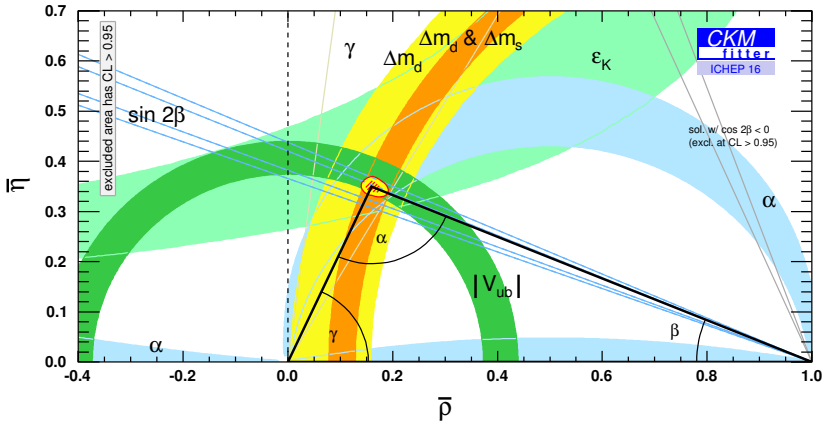

Fig. 4.1: CKM triangle in the complex plane. The coloured bands show the experimental constraints. The red hashed and the yellow area around the apex represent the currrent uncertainties at 68 % and 95 % confidence level [32].

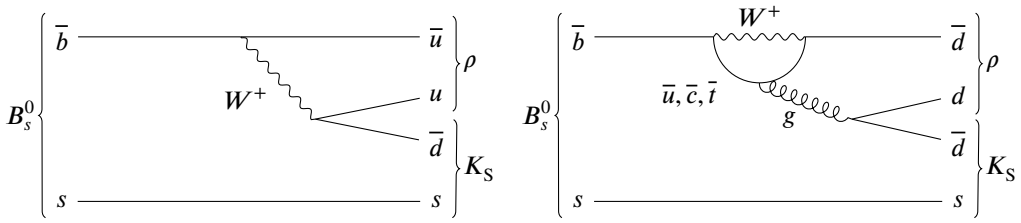

Fig. 4.2: Tree-level diagram of $B_s^0 \to \rho K_S^0$ (left) and the dominantly contributing gluonic penguin (right) [25].

λ_f can be expressed in terms of the CKM matrix elements

$$\lambda_f = -\frac{q_K}{p_K}\frac{q}{p}\frac{\langle \rho K^0 | T | \bar{B}_s^0 \rangle}{\langle \rho \bar{K}^0 | T | B_s^0 \rangle} = -\frac{V_{cs}^* V_{cd}}{V_{cs} V_{cd}^*}\frac{V_{tb}^* V_{ts}}{V_{tb} V_{ts}^*}\frac{V_{ub} V_{ud}^*}{V_{ub}^* V_{ud}} . \tag{4.2}$$

what simplifies λ_f being a pure phase when using the Wolfenstein parametrisation shown in Eq. (2.4)

$$\lambda_f = e^{-2i\gamma} . \tag{4.3}$$

However, the final state ρ in $B_s^0 \to \rho K_S^0$ has a $u\bar{u}$ and a $d\bar{d}$ component. Therefore, not only the tree-level transition $\bar{b} \to \bar{u}\,[u\bar{d}]$ but also the transition $\bar{b} \to \bar{d}\,[d\bar{d}]$ via the gluonic-penguin is possible beside the spectator s-quark (see Fig. 4.2). For both diagrams the CKM factor is $\propto A\lambda^3$ but the *weak* phases are γ and β for the tree-level diagram and the penguin, respectively. The effect of an additional contributing amplitude with a *weak* phase differing from the one of the tree-level process can be quantified by considering two *weak* phases

contributing to the transitions A_f ($B_s^0 \to \rho K_S^0$) and $\bar{A}_f$ ($\bar{B}_s^0 \to \rho K_S^0$)

$$\begin{aligned} A_f &= A_1 e^{i\left(\Phi_{A_1}+\delta_1\right)} + A_2 e^{i\left(\Phi_{A_2}+\delta_2\right)} , \\ \bar{A}_f &= \eta_f \left[A_1 e^{i\left(-\Phi_{A_1}+\delta_1\right)} + A_2 e^{i\left(-\Phi_{A_2}+\delta_2\right)} \right] . \end{aligned} \tag{4.4}$$

As shown in Eq. (3.20), the quantity q/p is a pure phase and hence one can derive $q/p = -e^{2i\Phi_M}$, resulting in

$$\lambda_f = -\eta_f e^{2i\Phi_M} \frac{A_1 e^{i\left(-\Phi_{A_1}+\delta_1\right)} + A_2 e^{i\left(-\Phi_{A_2}+\delta_2\right)}}{A_1 e^{i\left(\Phi_{A_1}+\delta_1\right)} + A_2 e^{i\left(\Phi_{A_2}+\delta_2\right)}} \tag{4.5}$$

with the *weak* phases Φ_{A_i} and the *strong* phases δ_i with $i = 1, 2$. The phases Φ_{A_1}, Φ_{A_2} and Φ_M are not invariant under global phase transformations, but the relative phases $\Phi_1 \equiv \Phi_{A_1} - \Phi_M$, $\Phi_2 \equiv \Phi_{A_2} - \Phi_M$ and $\Delta = \delta_2 - \delta_1$ can be measured. The form of λ_f in Eq. (4.5) already shows that a penguin contribution makes it impossible to measure the single phases Φ_1 and Φ_2. However, even with the approximation that $r = A_2/A_1$ is small, *i.e.* both contributing amplitudes are not of same magnitude, one finds

$$\begin{aligned} \lambda_f &= -\eta_f e^{-2i\Phi_1} \frac{1 + r e^{i(\Delta - \Phi_2 + \Phi_1)}}{1 + r e^{i(\Delta + \Phi_2 - \Phi_1)}} \\ &\approx -\eta_f e^{-2i\Phi_1} \left[1 + 2r \sin\Delta \sin\left(\Phi_2 - \Phi_1\right) - 2ir \cos\Delta \sin\left(\Phi_2 - \Phi_1\right) \right] . \end{aligned} \tag{4.6}$$

Obviously if a gluonic-penguin contributes with a different *weak* phase from the one of the tree-level diagram, *i.e.* $r \neq 0$ and $\Phi_1 \neq \Phi_2$, it is not possible to measure a single *weak* phase. Even in the case of vanishing final state interactions, *i.e.* $\Delta = 0$, λ_f can just be written as

$$\lambda_f = -\eta_f e^{-2i\left(\Phi_1 - \delta_{\Phi_1}\right)} , \tag{4.7}$$

where δ_{Φ_1} is defined by

$$\tan\left(\delta_{\Phi_1}\right) = \frac{r \sin\left(\Phi_1 - \Phi_2\right)}{1 + r \cos\left(\Phi_1 - \Phi_2\right)} . \tag{4.8}$$

For the decay $B_s^0 \to \rho K_S^0$, the amplitudes A_1 and A_2 can be identified as the tree-level and gluonic-penguin amplitudes, respectively, where the phases Φ_1 and Φ_2 would represent the CKM angles γ and β, such that γ cannot not be determined.

Additionally, it is important to note that hadronic matrix elements cannot be calculated reliably [22]. Consequently, in case they do not cancel out as they do if only one amplitude contributes to a specific decay, the resulting *CP* asymmetries cannot be interpreted without large uncertainties that need to be propagated into the determination of the sides and angles of the unitarity triangle.

Therefore, the current strategy to decrease the uncertainty on γ is to measure it in many different decay modes and combine the results into one value for the angle γ [33]. These decay modes can be divided into the two mentioned classes: the first class are tree-level processes where either decays of charged B mesons or neutral B mesons are exploited. The second class are processes involving penguin contributions, similar to the contribution to $B_s^0 \to \rho K_S^0$ described above.

4.1.1 Determination of γ in tree-level decays

As γ is propotional to the phase of the matrix element V_{ub}, a natural way to measure its value is to exploit interference effects between the Cabibbo-favoured $b \to c$ transitions and the Cabibbo-suppressed $b \to u$ transitions. This interference appears in decay channels such as $B^+ \to DK^+$ and $B_s^0 \to D_s^{\mp} K^{\pm}$. In the first case exploring different D decay chaines requires slightly different experimental methods, whereas in the latter case a time-dependent analysis is needed.

The basic principle when exploiting decays of charged B mesons is always the same. The initial $B^{\pm}$ meson decays into $D^0 K^{\pm}$ or $\overline{D}^0 K^{\pm}$ and subsequently the D meson is reconstructed in a final state common to both D^0 and $\overline{D}^0$. Therefore the amplitudes for the decay of the B meson can be defined as

$$\begin{aligned} A\big(B^+ \to D^0 K^+\big) &= V_{ub}^* V_{cs} = A e^{i(\delta+\gamma)}\,, \\ A\big(B^+ \to \overline{D}^0 K^+\big) &= V_{cb}^* V_{us} = \overline{A}\, e^{i\overline{\delta}}\,, \end{aligned} \tag{4.9}$$

where the *weak* phase is directly idenitified as γ, the parameters δ and $\overline{\delta}$ are the *strong* phases and A and $\overline{A}$ are the moduli of the amplitudes. For the D-meson decay the amplitudes can be defined accordingly as

$$\begin{aligned} A\big(D^0 \to f\big) &= A_D e^{i\delta_D}\,, \\ A\big(\overline{D}^0 \to f\big) &= \overline{A}_D\, e^{i\overline{\delta}_D}\,. \end{aligned} \tag{4.10}$$

Hence, the transition of $B^+ \to f K^+$ has two contributing amplitudes, which produce an interference term containing γ in the total decay rate

$$\left|A\big(B^+ \to f K^+\big)\right|^2 = A^2 \overline{A}_D^2 \left(1 + r_B^2 r_D^2 + 2\, r_B r_D \cos\big(\Delta + \Delta_D + \gamma\big)\right)\,, \tag{4.11}$$

where the notations $r_B = \overline{A}/A$, $r_D = A_D/\bar{A}_D$, $\Delta = \overline{\delta} - \delta$ and $\Delta_D = \delta_D - \overline{\delta}_D$ are used. To obtain the total decay rate for the B^- decay only the sign of γ must be reversed.

The first method exploiting this interference is the so-called GLW method [34, 35]. The idea is simply to reconstruct the intermediate D meson in CP eigenstates such as K^+K^- (CP-even) or $K_S^0\pi^0$ (CP-odd). As a consequences of such choice the unknowns from the D decay can be reduced because $r_D = 1$ and $\Delta_D = 0$, so the total decay rates for *e.g.* a CP-even D decay can be written as

$$\left|A\big(B^{\pm} \to f_{CP} K^{\pm}\big)\right|^2 = A^2 A_D^2 \left(1 + r_B^2 + 2 r_B \cos(\Delta \pm \gamma)\right)\,. \tag{4.12}$$

Experimentally, for the *CP*-even final state of the D^0 ($\overline{D}^0$) meson a *CP* asymmetry

$$A_{CP}^{\mathrm{GLW}} = \frac{\left|A\left(B^- \to f_{CP}K^-\right)\right|^2 - \left|A\left(B^+ \to f_{CP}K^+\right)\right|^2}{\left|A\left(B^- \to f_{CP}K^-\right)\right|^2 + \left|A\left(B^+ \to f_{CP}K^+\right)\right|^2} = \frac{2r_B \sin(\Delta)\sin(\gamma)}{1 + r_B^2 + 2r_B\cos(\Delta \pm \gamma)} \quad (4.13)$$

and a *CP* ratio

$$R_{CP}^{\mathrm{GLW}} = \frac{\left|A\left(B^- \to f_{CP}K^-\right)\right|^2 + \left|A\left(B^+ \to f_{CP}K^+\right)\right|^2}{2\left|A\left(B^- \to D^0K^-\right)\right|^2} = 1 + r_B^2 + 2r_B\cos(\Delta)\cos(\gamma) \quad (4.14)$$

can be measured. If the D meson in the denominator of R_{CP}^{GLW} is reconstructed in a hadronic final state such as $K\pi$, decays proceeding via a D^0 or a $\overline{D}^0$ cannot be distinguished as both decay into $K^+\pi^-$ and $K^-\pi^+$. However, this can be avoided by instead measuring the double ratio

$$R_{+}^{\mathrm{GLW}} = \frac{\left|A\left(B^- \to f_{CP}K^-\right)\right|^2 + \left|A\left(B^+ \to f_{CP}K^+\right)\right|^2}{\left|A\left(B^- \to f_{CP}\pi^-\right)\right|^2 + \left|A\left(B^+ \to f_{CP}\pi^+\right)\right|^2} \times \frac{1}{R_{K/\pi}} \quad (4.15)$$

with

$$R_{K/\pi} = \frac{\left|A\left(B^- \to D\left(\to K^-\pi^+\right)K^-\right)\right|^2 + \left|A\left(B^+ \to D\left(\to K^+\pi^-\right)K^+\right)\right|^2}{\left|A\left(B^- \to D\left(\to K^-\pi^+\right)\pi^-\right)\right|^2 + \left|A\left(B^+ \to D\left(\to K^+\pi^-\right)\pi^+\right)\right|^2}, \quad (4.16)$$

where f_{CP} denotes always a *CP*-even final state of the D meson from Eq. (4.9). This double ratio R_{+}^{GLW} is identical to R_{CP}^{GLW} when neglecting all terms proportional to r_D^2, $r_{B_\pi}^2 = \left(A(B^+ \to D^0\pi^+)/A(B^- \to D^0\pi^-)\right)^2$ or $r_D r_{B_\pi}$ which are of $\mathcal{O}(1\,\%)$. However, the GLW method has two disadvantages: firstly, the interference term is proportional to r_B (see Eq. (4.12)) which is of order 10 % and therefore decreases the sensitivity on γ. Secondly, there is not a single solution for γ, but an eight-fold discrete ambiguity due to the fact that γ and Δ always appear together in the product of two sine or cosine functions as in Eq. (4.13) and Eq. (4.14).

The second, so-called ADS method [36] directly counteracts the first disadvantage of the GLW method. Instead of using decays into *CP*-eigenstates of the D meson, the D is reconstructed in a final state for which $D^0 \to f$ is suppressed relative to $\overline{D}^0 \to f$ (*e.g.* f could be $K^-\pi^+$). This enhances the interference term to be of the same order as the other ones, but on the other hand r_D and Δ_D do not cancel as before. Hence, to determine γ from similar asymmetries and ratios as for the GLW method, external input for both quantities r_D and Δ_D is needed.

The last method described in this section using charged B decays probably is the most sophisticated one. The GGSZ method [37] makes use of multibody D decays like $K_S^0\pi^+\pi^-$. On the one hand, this requires an analysis of the Dalitz plane as the strong phase originating from the D decay varies over the Dalitz plane, but on the other hand γ can be extracted without the disadvantages of the GLW or the ADS method. The basic idea however of

exploiting the interference between B^+ decays via D^0 or $\overline{D}^0$ mesons is the same as for the GLW and ADS methods. In case of multibody D decays the interference term can be expressed in the form $\cos(\phi)\, c_i \pm \sin(\phi)\, s_i$, where ϕ is the difference (sum) of the *strong* phase δ_B originating from the B^- (B^+) decay and γ. The coefficients c_i and s_i contain the varying *strong* phases from the D decay. Using k pairs of bins which are symmetrically around the $m_{K^0_S\pi^-} \leftrightarrow m_{K^0_S\pi^+}$ axis in the Dalitz plot, and exploiting the symmetry of the c_i and s_i around this axis leads to $4k$ equations. These are related to the $2k + 3$ unknowns c_i, s_i, r_B, δ_B and γ, so that a partition of the D meson Dalitz plot into four or more bins allows to determine all unknowns.

Finally it is important to mention, that the described methods can be applied to any decay of type $B^\pm \to DX^\pm$ where X is any state with the flavour of a kaon.

When exploiting decays of uncharged B mesons like $B^0_s \to D^\mp_s K^\pm$ or $B^0 \to D^\mp \pi^\pm$ to determine the angle γ, the procedure is different as this class of decays is only affected by ***CP*** violation in the interference of decay and decay after mixing and therefore a time-dependent analysis is required in order to achieve sensitivity on γ. The basic formalism has already been described in Sec. 3.3. Due to the ***CP*** conservation in the decay, $C_{\bar{f}} = -C_f$ and the number of independent coefficients as given in Eq. (3.34) and Eq. (3.35) is reduced to five. Furthermore the ***CP*** coefficient C_f contains the ratio r_B, while via the coefficients S_f, $S_{\bar{f}}$, $A_f^{\Delta\Gamma}$, $A_{\bar{f}}^{\Delta\Gamma}$ the phase information are measurable. This experimental technique to measure ***CP*** violation can also be applied to $B^0 \to D^\mp \pi^\pm$ and therefore it will be described more thoroughly in Sec. 4.2.

4.1.2 Determination of γ in loop processes

Instead of exploiting decays where γ can be accessed via the interference in tree-level diagrams also processes containing loop transitions can be used to infer information about γ. The first and probably most obvious possibility is to derive a result for γ from experimental measurements of other CKM triangle quantities as the remaining two angles of the CKM triangle, α and β, which are only accesible in loop processes. As an example, the angle β can be measured quite precisely in the decay $B^0 \to J/\psi K^0_S$ while the angle α is accessible via an isospin analyses of the decays $B^0 \to \pi^+\pi^-$, $B^0 \to \pi^0\pi^0$ and $B^+ \to \pi^+\pi^0$ [38] or measurements of longitudinally polarized $B^0 \to \rho\rho$ decays [39, 40]. Though, the best precision is achieved when performing a *global* fit using all available inputs as performed by the CKMfitter and UTfit collaborations [32, 41].

However, such tests of a closing CKM triangle do not directly indicate a specific process which is affected by effects of New Physics (NP). Therefore, a more direct determination of γ is also important. A possibility for such determination is to extract γ from ***CP*** violation measurements in $B^0 \to \pi\pi$ and $B^0_s \to K^+K^-$ [42, 43]. For this, the U-spin flavour symmetry of strong interactions between the two decay channels, *i.e.* the decays are related by interchanging all d-quarks with s-quarks, is exploited.

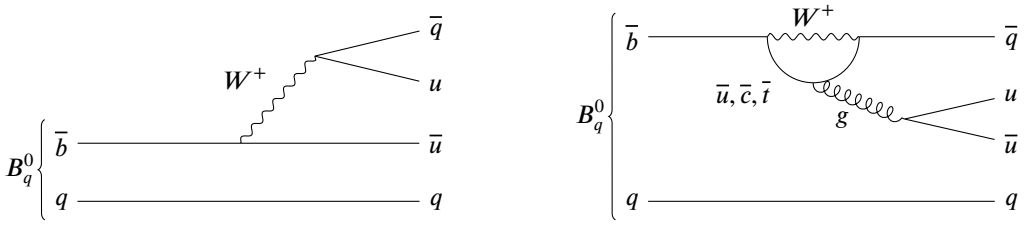

Fig. 4.3: Feynman diagrams of the Tree-level (left) and penguin (right) contributions to $B^0 \to \pi^+\pi^-$ ($q = d$) and $B_s^0 \to K^+K^-$($q = s$) [25].

In order to do so, the transition amplitudes and *CP* observables for the decay modes $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$ have to be defined. In Fig. 4.3, the dominant $\overline{b} \to \overline{u}\,[u\overline{d}]$ processes are shown, which leads to the transition amplitude

$$A\left(B^0 \to \pi^+\pi^-\right) = \lambda_u \left(A_{\mathrm{cc}}^u + A_{\mathrm{pen}}^u\right) + \lambda_c A_{\mathrm{pen}}^c + \lambda_t A_{\mathrm{pen}}^t \tag{4.17}$$

where the amplitudes A_{pen}^q and A_{c}^q describe the penguin and charged-current contributions, respectively. Furthermore, the short notation $\lambda_q = V_{qd}V_{qb}^*$ with $q = d, s$ was applied. Making use of the unitarity of the CKM matrix and applying the Wolfenstein parametrisation [20] allows to determine the parameter λ_f and the corresponding *CP* coefficients

$$\begin{aligned} C_f^{B^0\to\pi^+\pi^-} &= -\left[\frac{2d\sin(\theta)\sin(\gamma)}{1 - 2d\cos(\theta)\cos(\gamma) + d^2}\right], \\ S_f^{B^0\to\pi^+\pi^-} &= \left[\frac{\sin(2\beta+2\gamma) - 2d\cos(\theta)\sin(2\beta+\gamma) + d^2\sin(2\beta)}{1 - 2d\cos(\theta)\cos(\gamma) + d^2}\right]. \end{aligned} \tag{4.18}$$

The quantities d and θ are defined as

$$d\,e^{i\theta} = \frac{1}{\left(1 - \lambda^2/2\right) R_b}\left(\frac{A_{\mathrm{pen}}^{ct}}{A_{\mathrm{cc}}^u + A_{\mathrm{pen}}^{ut}}\right) \tag{4.19}$$

with $A_{\mathrm{pen}}^{qt} = A_{\mathrm{pen}}^q - A_{\mathrm{pen}}^t$ and the CKM factor $R_b = 1/\lambda \left|V_{ub}/V_{cb}\right|$, where λ is the Wolfenstein parameter introduced in Eq. (2.4). The *CP* coefficient $A_f^{\Delta\Gamma}$ is neglected as the decay width difference for B^0 mesons is expected to be negligibly small.

Equivalently to this, the *CP* coefficients for $B_s^0 \to K^+K^-$ decays can be calculated as

$$\begin{aligned} C_f^{B_s^0\to K^+K^-} &= \left[\frac{2\tilde{d}'\sin\left(\theta'\right)\sin(\gamma)}{1 - 2\tilde{d}'\cos(\theta')\cos(\gamma) + \tilde{d}'^2}\right], \\ S_f^{B_s^0\to K^+K^-} &= \left[\frac{\sin\left(\phi_s+2\gamma\right) - 2\tilde{d}'\cos(\theta)\sin\left(\phi_s+\gamma\right) + \tilde{d}'^2\sin\left(\phi_s\right)}{1 - 2\tilde{d}'\cos(\theta')\cos(\gamma) + \tilde{d}'^2}\right], \\ A_f^{\Delta\Gamma, B_s^0\to K^+K^-} &= -\left[\frac{\cos\left(\phi_s+2\gamma\right) + 2\tilde{d}'\cos\left(\theta'\right)\cos\left(\phi_s+\gamma\right) + \tilde{d}'^2\cos\left(\phi_s\right)}{1 - 2\tilde{d}'\cos(\theta')\cos(\gamma) + \tilde{d}'^2}\right] \end{aligned} \tag{4.20}$$

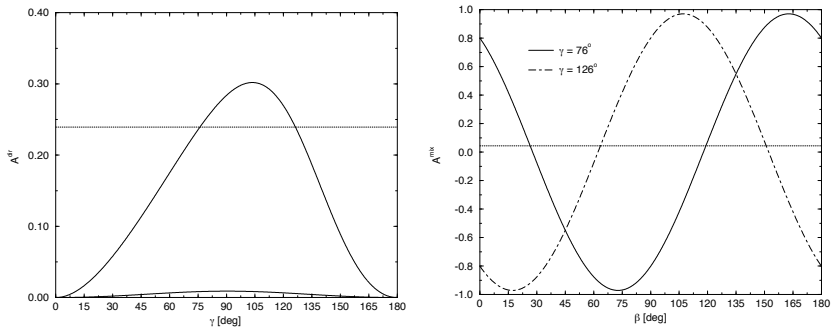

Fig. 4.4: Dependence of $C_f^{B^0\to\pi^+\pi^-}$ (A^{dir}) on γ (left) and $S_f^{B^0\to\pi^+\pi^-}$ (A^{mix}) on β (right) for a specific example as given in Ref. [42].

where the short notation $\tilde{d}' = \left(1-\lambda^2/\lambda^2\right) d'$ was applied. Assuming the U-spin flavour symmetry to be exact, implies $d' = d$ and $\theta' = \theta$.

This theoretical setup can now be used in several strategies to determine γ among other observables as β, d and θ. In the following, the basic idea of most strategies will be briefly outlined. The first approach [42] uses an external input for the phase ϕ_s, so that the four unknows d, θ, β and γ can be determined using only the four observables $C_f^{B\to hh}$, and $S_f^{B\to hh}$. As shown in Ref. [42], the parameters d and θ can be expressed as functions of γ and the *CP* coefficients $C_f^{B_s^0\to K^+K^-}$ and $S_f^{B_s^0\to K^+K^-}$. Inserting this into $C_f^{B^0\to\pi^+\pi^-}$ then allows to determine γ up to a two-fold ambiguity as shown in Fig. 4.4. Using these two solutions for γ and inserting them together with the obtained expressions for d and θ into $S_f^{B^0\to\pi^+\pi^-}$ gives another twofold ambiguity for β (see Fig. 4.4). Hereby, for both ambiguities the assumption is made that $\gamma \in [0, 180]^\circ$ and $\beta \in [0, 180]^\circ$. If additionally the phase β is used, this ambiguities can be resolved by expressing d and θ as functions of γ and the *CP* coefficients $C_f^{B^0\to\pi^+\pi^-}$ and $S_f^{B^0\to\pi^+\pi^-}$ and consequently fixing the contours in the $\gamma - d$ plane. However, assuming a 15 % and 20 % U-spin-breaking effect on the parameters d and θ as proposed in Ref. [44], respectively, an additional uncertainty on the extracted value of γ of roughly 20 % could arise [45]. Adding also the modes $B^0 \to \pi^0\pi^0$ and $B^+ \to \pi^0\pi^+$ also allows to either constrain NP effects in the $b\to s$ penguin contributions or in mixing [43].

4.1.3 Comparison of tree-level and loop determinations of γ

Both strategies described in the previous sections were and are used to determine the angle γ. As shown before, the determination from tree-level decays provides theoretically clean measurements, which are only affected by low sensitivity or discrete ambiguities. Comparing these measurements with measurements of loop-processes could provide aa constraint on potentially contributing NP effects. However, to properly compare the different approaches, the experimental uncertainties need to be reduced to a sensible

amount. Therefore, the current experimental status and required improvements in the experimental precision are discussed below.

At LHCb, the measurements exploiting tree-level processes are studied in many different decay modes and combined into one single γ-combination. The most recent combination was presented in Ref. [33] yielding a result of

$$\gamma = \left(74.0^{+5.0}_{-5.8}\right)^{\circ}. \tag{4.21}$$

This result includes 15 different decay modes, of which ten were analysed using data corresponding to 3 fb^{-1} collected at centre-of-mass energies of 7 and 8 TeV. The remaining five measurements use additional data corresponding to 2 fb^{-1} collected at centre-of-mass energies of 13 TeV. Deriving constraints on γ by performing a *global* fit using all available inputs gives

$$\gamma = \left(65.4^{+0.97}_{-1.16}\right)^{\circ} \quad \text{and} \quad \gamma = (65.8 \pm 1.9)^{\circ}\,, \tag{4.22}$$

determined by the CKMfitter [32] and UTfit [41] collaborations, respectively. Determining γ from loop processes with a similar strategy exploiting U-spin-symmetry as outlined in Sec. 4.1.2 leads to

$$\gamma = \left(63.5^{+7.2}_{-6.7}\right)^{\circ} \tag{4.23}$$

where inputs from the BaBar, Belle, CDF and LHCb collaborations were used [46]. In contrast to the combination of tree-level measurements, the inputs from LHCb are based on analyses using only 1 fb^{-1} collected at a centre-of-mass energy of 7 TeV.

The determinations from the *global* fits are by far the most precise. Furthermore the results using loop processes seem to agree quite well, while there seems to be the potential for a tension between the determinations using loop-processes and tree-level transitions. However, the discrepancy between the result from Eq. (4.21) and Eq. (4.22) is only $\approx 1.5\sigma$ due to the large uncertainties, *i.e.* the precision with which γ is determined in tree-level decays must be increased to draw a reliable conclusion.

4.2 Measuring γ in $B^0 \to D^{\mp}\pi^{\pm}$ decays

The decay channel $B^0 \to D^{\mp}\pi^{\pm}$ allows to measure *CP* violation in the interference between decay and decay after mixing as both B^0 and $\overline{B}^0$ mesons can decay in the final states $D^-\pi^+$ and $D^+\pi^-$ as shown in Fig. 4.5. More precisely, the amplitudes of the $b \to c\,[\overline{u}d]$ and $b \to u\,[\overline{c}d]$ transitions interfer. Consequently, measuring the resulting *CP* asymmetries gives access to the *weak* and *strong* phases originating from these transitions. Neglecting constants such as the Fermi constant, which cancels in the amplitude ratio anyway, the amplitudes of the four contributing transitions can be expressed in terms of the CKM matrix elements V_{ij}, hadronic matrix elements M and $\overline{M}$ and potential *strong* phases Δ and

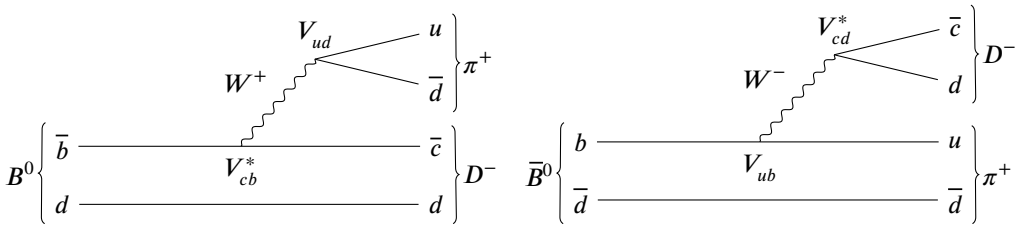

Fig. 4.5: Feynman diagrams of the Cabibbo-favoured $B^0 \to D^-\pi^+$ (left) and Cabibbo-suppressed $\overline{B}^0 \to D^-\pi^+$ (right) decays [25].

$\overline{\Delta}$:

$$A\left(B^0 \to D^-\pi^+\right) = V_{ud}V_{cb}^* \times M e^{i\Delta}, \tag{4.24}$$

$$A\left(\overline{B}^0 \to D^-\pi^+\right) = V_{ub}V_{cd}^* \times \overline{M}\, e^{i\bar{\Delta}}, \tag{4.25}$$

$$A\left(B^0 \to D^+\pi^-\right) = V_{ub}^*V_{cd} \times \overline{M}\, e^{i\bar{\Delta}}, \tag{4.26}$$

$$A\left(\overline{B}^0 \to D^+\pi^-\right) = V_{ud}^*V_{cb} \times M e^{i\Delta}. \tag{4.27}$$

Denoting the final state $D^-\pi^+$ and $D^+\pi^-$ as f and $\bar{f}$, respectively, and using the expression for q/p from Eq. (3.21) the parameters λ_f and $\lambda_{\bar{f}}$ can be defined as

$$\lambda_f = \frac{q}{p}\frac{A\left(\overline{B}^0 \to D^-\pi^+\right)}{A\left(B^0 \to D^-\pi^+\right)} = -\frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}\frac{V_{ub}V_{cd}^*}{V_{ud}V_{cb}^*}\frac{\overline{M}}{M}e^{i\delta}, \tag{4.28}$$

$$\lambda_{\bar{f}} = \frac{q}{p}\frac{A\left(\overline{B}^0 \to D^+\pi^-\right)}{A\left(B^0 \to D^+\pi^-\right)} = -\frac{V_{tb}^*V_{td}}{V_{tb}V_{td}^*}\frac{V_{ud}^*V_{cb}}{V_{ub}^*V_{cd}}\frac{M}{\overline{M}}e^{-i\delta} \tag{4.29}$$

where the abbreviation $\delta = \bar{\Delta} - \Delta$ was applied. Using the definition of the angles of the unitarity triangle from Eq. (2.8) the fraction of CKM matrix elements can be further simplified to

$$\lambda_f = -\left|\frac{V_{ub}V_{cd}}{V_{ud}V_{cb}}\right|\frac{\overline{M}}{M}e^{-i(2\beta+\gamma-\delta)} = -re^{-i(2\beta+\gamma-\delta)}, \tag{4.30}$$

$$\lambda_{\bar{f}} = -\left|\frac{V_{ud}^*V_{cb}^*}{V_{ub}^*V_{cd}^*}\right|\frac{M}{\overline{M}}e^{-i(2\beta+\gamma+\delta)} = -\frac{1}{r}e^{-i(2\beta+\gamma+\delta)}. \tag{4.31}$$

Here, the ratio of the CKM matrix elements and the ratio of hadronic matrix elements were combined into the parameter r. Thus, the decay allows to probe $2\beta + \gamma$ together with the *strong* phase difference δ.

Looking at the four contributing decay rates, the expressions from Eqs. (3.30) to (3.33)

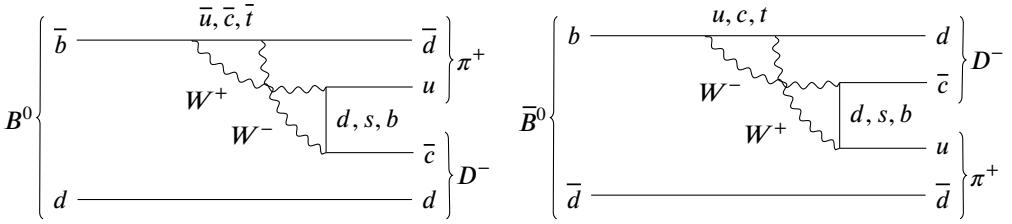

Fig. 4.6: First higher-order corrections to the Cabibbo-favoured $B^0 \to D^-\pi^+$ (left) and Cabibbo-suppressed $\overline{B}^0 \to D^-\pi^+$ (right) decays [25].

reduce to

$$\Gamma\big(B^0(t) \to D^-\pi^+\big) = \frac{A}{2} e^{\Gamma t}\big[1 - S_f \sin(\Delta m t) + C_f \cos(\Delta m t)\big]\,, \tag{4.32}$$

$$\Gamma\big(\overline{B}^0(t) \to D^-\pi^+\big) = \frac{A}{2} e^{\Gamma t}\big[1 + S_f \sin(\Delta m t) - C_f \cos(\Delta m t)\big]\,, \tag{4.33}$$

$$\Gamma\big(B^0(t) \to D^+\pi^-\big) = \frac{\overline{A}}{2} e^{\Gamma t}\big[1 - S_{\bar{f}} \sin(\Delta m t) + C_{\bar{f}} \cos(\Delta m t)\big]\,, \tag{4.34}$$

$$\Gamma\big(\overline{B}^0(t) \to D^+\pi^-\big) = \frac{\overline{A}}{2} e^{\Gamma t}\big[1 + S_{\bar{f}} \sin(\Delta m t) - C_{\bar{f}} \cos(\Delta m t)\big]\,, \tag{4.35}$$

when assuming that the decay width difference $\Delta\Gamma$ is zero and q/p is a pure phase. Additionally $\overline{(A)} = A\big(B^0 \to D^{\mp}\pi^{\pm}\big)\left(1 + \left|\lambda_{\bar{f}}\right|^2\right)$ was used. From $\Delta\Gamma = 0$ follows, that the parameters $A_f^{\Delta\Gamma}$ and $A_{\bar{f}}^{\Delta\Gamma}$ do not contribute to the decay widths while the remaining *CP* parameters

$$S_f = \frac{2r \sin(2\beta + \gamma - \delta)}{1 + r^2}\,, \tag{4.36}$$

$$S_{\bar{f}} = \frac{2r \sin(2\beta + \gamma + \delta)}{1 + r^2}\,, \tag{4.37}$$

$$C_f = -C_{\bar{f}} = \frac{1 - r^2}{1 + r^2} \tag{4.38}$$

give access to the *weak* phase $2\beta + \gamma$ and the *strong* phase difference δ up to a two-fold ambiguity in the range $[0, 180]^\circ$. Alternatively using inputs from other measurements of β, the CKM angle γ can be determined. Estimating the ratio of interfering amplitudes r by expressing the CKM matrix elements in the Wolfenstein parametrisation and assuming $\overline{M}/M \approx \mathcal{O}(1)$ leads to

$$r \approx \frac{V_{ub}V_{cd}}{V_{ud}V_{cb}} = \frac{\lambda^2\sqrt{\rho^2 + \eta^2}}{1 - \lambda^2/2} \approx 2\,\% \tag{4.39}$$

what is in good agreement with the experimental determinations [47, 48]. As the parameters S_f and $S_{\bar{f}}$ are proportional to this ratio the interference and thus the sensitivity on γ is expected to be quite small.

So far, only the tree-level contributions shown in Fig. 4.5 have been taken into account. Therefore, also possible contributions from penguin contributions are briefly discussed in the following. The first irreducible error on the determination of γ arises from the box diagrams as shown in Fig. 4.6. These contributions carry potentially a different *weak* phase and thus induce a shift δ_γ as shown in Sec. 4.1. However only the contribution to the $b \to c\,[\bar{u}d]$ transition needs to be taken into account as effects of $b \to u\,[\bar{c}d]$-transitions are suppressed by λ^2 and can be neglected. In good approximation δ_γ can be calculated by investigating the effect on effective couplings leading $\delta_\gamma \approx 10^{-6}$ to 10^{-4} [49]. This shift is much smaller than any expected experimental uncertainty expected in the next years and can therefore safely be ignored.

5 The LHCb experiment

The Large Hadron Collider (LHC) is a circular proton-proton collider at the European Organisation for Nuclear Research (CERN). Along with ATLAS, CMS and ALICE, the LHCb experiment is one of the four major experiments at the LHC. The experiment is specialized on precision measurements of physics processes involving b- and c-quarks. Below, first the LHC is briefly introduced, followed by a more detailed description of the LHCb detector and its components, based on Refs. [50] and [51]. At the end of this chapter the LHCb software stack will be outlined shortly.

5.1 The Large Hadron Collider

The LHC is a ring accelerator at CERN near Geneva, Switzerland, with a circumference of about 27 km. It is designed to collide protons at a centre-of-mass energy of up to $\sqrt{s} = 14\,\text{TeV}$ at a luminosity of $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ [53]. Two proton beams are accelerated in opposite directions and brought to collision at four points, where the ATLAS, CMS, ALICE and the LHCb experiments are located. The first protons in the LHC were accelerated in 2008.

The LHC operates in running periods, the first from 2010 to 2012 (Run I), followed by the long shutdown 1 in 2013 and 2014. During Run I, the proton beams consisted of about 1380 proton bunches, colliding at centre-of-mass energies of $\sqrt{s} = 7\,\text{TeV}$ (2010 and 2011) and 8 TeV (2012) [54]. For the currently ongoing second running period (Run II) the energy and the number of proton bunches per beam were increased to $\sqrt{s} = 13\,\text{TeV}$ and 2220, respectively [54]. The latter was achieved by reducing the bunch spacing from 50 ns to 25 ns.

Before the protons are injected into the LHC accelerator ring, they have to be pre-accelerated. This is first done in a linear accelerator, the LINAC 2, followed by the booster, the proton synchrotron (PS) and the super-proton synchrotron (SPS), all of which are circular accelerators. From the SPS, the protons are then injected into the LHC ring with an energy of 450 GeV [53] (see Fig. 5.1). In order to keep the high-energetic protons on their circular path, 1232 superconducting dipole magnets with a length of 14.3 m each and a field strength of up to 8.33 T are in operation.

As already mentioned, four major experiments are located at the LHC. ATLAS and CMS are general-purpose experiments collecting data at maximum luminosity, ALICE mainly studies quark-gluon plasmas and the fourth experiment, LHCb, performs primarily

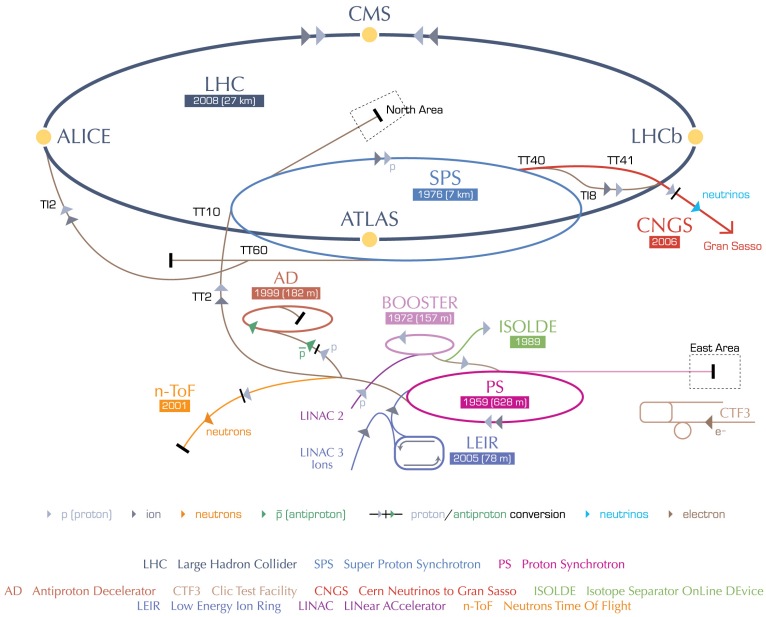

Fig. 5.1: Schematic view of the CERN accelerator complex. The protons for the LHC are first pre-accelerated to an energy of 450 GeV in the LINAC 2, the booster, the proton synchrotron (PS) and the super proton synchrotron (SPS). Their path is indicated in the illustration by the light grey arrows [52].

precision measurements in the field of flavour physics, especially with b- and c-hadrons.

5.2 The LHCb detector

In contrast to the three other experiments, the LHCb detector does not cover the entire spatial angle, but is designed as a single-arm forward spectrometer with an angular coverage of about 10 to 300 mrad in the bending plane. The detector geometry is based on the fact that the mainly investigated $b\bar{b}$-quark pairs have a high probability of being produced in forward or backward direction (see Fig. 5.2). Despite the limited angular coverage, about 25 % of produced $b\bar{b}$ quark pairs are within the detector acceptance. Furthermore, it is important at LHCb to resolve individual processes as detailed as possible. This is easier for events with less particles traversing the detector, *i.e.* for events with ideally only one proton-proton collisions occuring simultaneously. Therefore, LHCb does not use the maximum luminosity provided by the LHC, but a constant luminosity of about $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [54]. This is adjusted by reducing the overlap of the colliding proton bunches in the LHCb detector compared to the other experiments. However, since the collision rate during Run I with 20 MHz was still too high to store all events directly, powerful trigger systems are in place, which already make an initial selection of the data

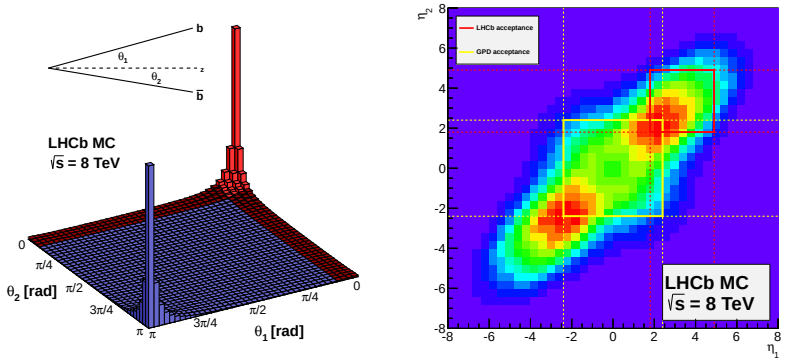

Fig. 5.2: Angular distribution of the $b\bar{b}$-quark pairs with respect to the beam axis in a proton-proton collision with a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ as expected from simulations; the LHCb detector acceptance is shown in red (left). Simulated pseudo-rapidity distribution for two b-quarks produced in a proton-proton collision with a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ where the yellow box represents the acceptance of a general-purpose detector as CMS or ATLAS and the red box shows the LHCb detector acceptance (right) [55].

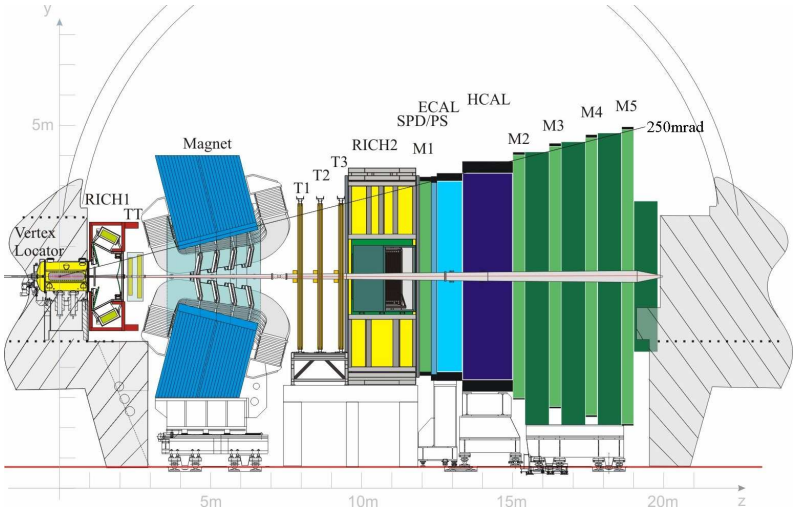

Fig. 5.3: Schematic structure of the LHCb detector: The VELO, the two RICH detectors, as well as the tracking system with magnet, calorimeter and the muon chambers. The collision point is located on the left side, enclosed by the VELO [50]. The z-coordinate is defined along the beam pipe, the x- and y-coordinates are defined vertically to the beam pipe.

and preserving as interesting physics processes as possible.

During the injection and accelerating phases of the LHC beam instabilities may arise, which could cause a significant damage in detector components close to the beam pipe. To protect the detector from such incidents, it is equipped with several safety systems. The main protection system is the beam conditions monitor (BCM), which measures the particle flux in close vicinity of the beam pipe at two locations downstream and upstream of the VELO [56]. In case of any unexpected beam displacement the BCM is able to remove the beam-permit signal and therefore to dump the LHC beam in order to protect the detector equipment. Accordingly, the BCM is powered by the LHC-mains through an uninterruptible power supply.

In the following, the individual components of the LHCb detector (see Fig. 5.3) based on Ref. [50] are explained, separately for the components of the tracking system, the particle identification system and the LHCb trigger. Though it is important to note that neither the components described in Sec. 5.2.1 nor the components described in Sec. 5.2.2 are exclusively used for tracking or particle identification purposes, but always a combination of all components is needed for the final reconstruction.

5.2.1 The tracking system

The tracking system consists of the VELO, which encloses the collision point, the TT and the tracking stations T1-T3. Furthermore, a dipole magnet bending the tracks of charged particles is part of the tracking system.

The VELO

The vertex locator (VELO) is the detector component closest to the collision point. It is used to resolve primary and secondary vertices with high precision. The VELO is composed of 21 semicircular silicon modules which can be moved up to 8 mm to the beam. These measure the r and ϕ coordinates of the hits left by a traversing charged particle and are mounted along the beam axis as shown in Fig. 5.4. Furthermore, the inner part close to the collision point of the protons is distinguished from an outer part downstream of the nominal interaction point.

To reconstruct a track for a tranversing particle in the VELO, it is required that the particle generates hits in at least three stations. To achieve an angular acceptance of 300 mrad of the VELO under this condition, the distance between the inner stations is smaller than 5 cm with a sensor radius of 42 mm. This small distance leads to a quite short extrapolation distance from the first measured hit to the interaction point. The resolution of the interaction point, also denoted as primary vertex (PV), depends on the number of tracks forming the vertex. For a PV with 25 tracks, the resolution was 13 µm in the x and y coordinates and 71 µm in z in Run I [51]. For particles, which are created at $z = 10.6$ cm downstream of the nominal interaction point, the lower limit of the angular accceptance is

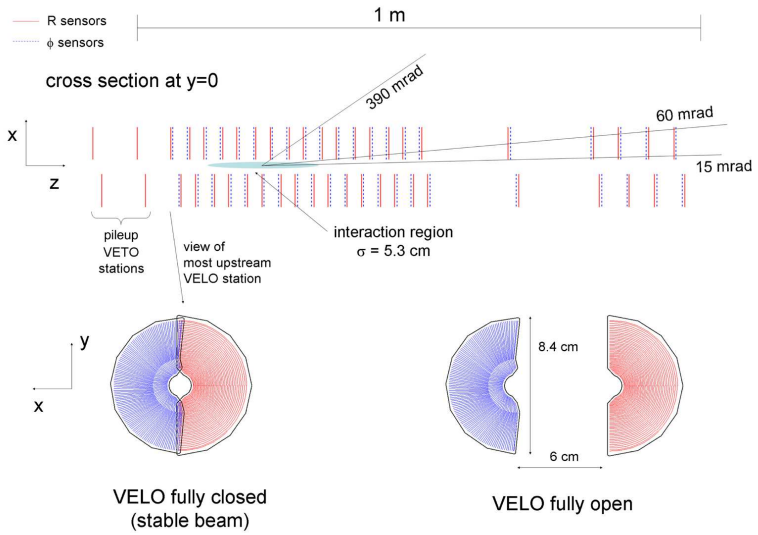

Fig. 5.4: View through the (x, z)-plane of the VELO at $y = 0$ with closed modules (top). View from the beam direction onto a module in closed and open state (bottom). The two halves for ϕ- (blue) and r-measurement (red) are shown [50].

15 mrad.

To cover the full azimuthal acceptance, the two detector halves of each module overlap. This is possible because the z-positions of the respective halves are shifted by 1.5 cm with respect to each other (see Fig. 5.4).

Tracking stations

The tracking stations comprise the Tracker Turicensis (TT) and the three tracking stations T1 to T3 . The stations T1 to T3 are further divided into an inner area close to the beam pipe, the Inner Tracker (IT), and more distant areas from the beam pipe, denoted as Outer Tracker (OT). A dipole magnet is located between the TT and the tracking station T1, which is described in Sec. 5.2.1. Technologically, both the TT and the IT are silicon trackers (ST).

The STs are made of silicon strips with a width of 200 µm, what leads to a spatial resolution of 50 µm [51] in the IT and the TT. The TT is directly located downstream of the RICH1 while the IT is installed in the central part around the beam pipe of the three stations T1 to T3. It covers a region 120 cm wide and 40 cm high as shown in Fig. 5.5.

The OT is a detector made of drift tubes filled with gas. The drift time of ionized gas atoms and their electrons is measured in the drift tubes and from this drift time the ionization spot is determined. The gas mixture in the drift tubes consists of 70 % argon, 28.5 % CO_2 and 1.5 % O_2. This composition guarantees fast drift times of about 35 ns

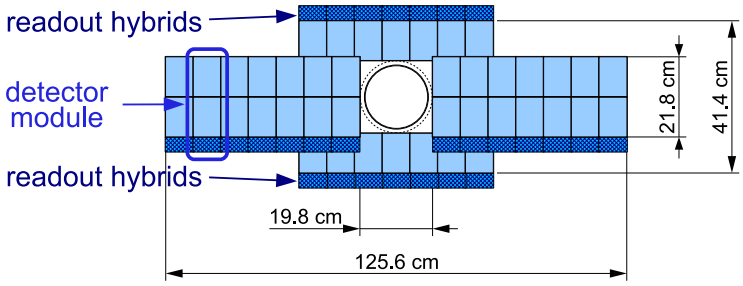

Fig. 5.5: Schematic view of a layer of IT with readout electronics. One station consists of four such layers, whereby the middle layers are rotated around the z-axis to obtain additional angular information [50].

and a high spatial resolution of 205 µm [51]. The momentum resolution of the OT is approximately 0.4 %, with an overall reconstruction efficiency of 80 % [50].

All tracking stations, *i.e.* both, the TT and the stations T1 to T3, are constructed out of four layers each with the middle layers rotated by ±5° around the beam axis so that the y-coordinate of a hit left by a traversing particle can be determined.

The dipole magnet

As mentioned previously, the dipole magnet is located between the TT and the tracking station T1. It covers an acceptance of ±250 mrad vertically and ±300 mrad horizontally. The magnet is designed as a conventional magnet with saddle-shaped coils. The integrated magnetic field strength is 4 T m for tracks with a length of 10 m. Its strength along the z-axis is shown in Fig. 5.6. Tracks of charged particles are bent within the accelerator plane (x-plane) due to the vertical magnetic field. The relative precision of the magnetic field required to achieve the desired momentum resolution is of $\mathcal{O}(10^{-4})$. During data taking, the polarity is inverted regularly to prevent systematic effects due to *e.g.* different performances detector areas which the particles traverse.

5.2.2 The particle identification system

The particle identification system consists of two ring-imaging cherenkov (RICH) detectors, the first upstream of the TT, the second downstream of the tracking station T3. Further, downstream of the second RICH detector, the electromagnetic and hadronic calorimeters are located, followed by the muon chambers.

The RICH detectors

An important factor at LHCb is the particle identification, especially the distinction between pions and kaons. To realize this, LHCb has two ring imaging Cherenkov detectors, RICH1

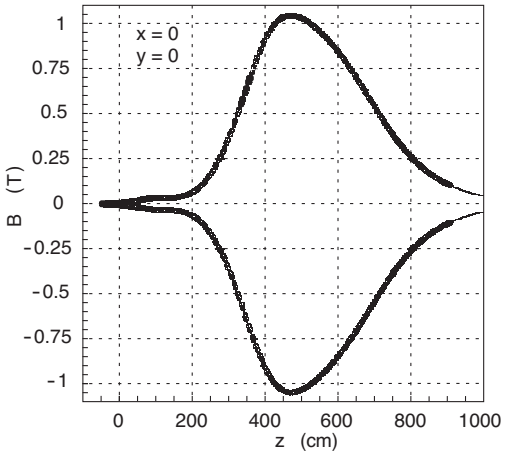

Fig. 5.6: Magnetic field along the z-axis [50].

and RICH2. These detectors exploit the Cherenkov effect, where electromagnetic radiation is emitted when a charged particle passes through a medium faster than the speed of light in that medium. This electromagnetic radiation is emitted with an opening angle θ, where θ depends on the speed of the particles and the refractive index n of the medium:

$$\cos(\theta) = \frac{1}{n\beta}\,. \tag{5.1}$$

Together with the momentum information from other detector components, different particels can be distuingished. Since the momentum spectrum changes with the polar angle, the RICH system consists of the two components RICH1 and RICH2. RICH1 identifies particles with small momenta of about 1 GeV to 60 GeV [50] with a mixture of aerogel and C_4F_{10}, while RICH2 uses a CF_4 gas and distinguishes particles with larger momenta in the range of 15 GeV to 100 GeV [50].

Averaging over the momentum, the efficiency to identify a kaon as such is about 95 % with a pion misidentification rate of about 10 %. This latter number can be reduced by slightly stricter requirements yielding a pion misidentification rate of about 3 % at a loss in kaon identification efficiency of about 10 % [51].

The calorimeters

The calorimeters have various functions. On the one hand they support the e-, γ- and hadron identifications, on the other hand they measure particle energies and positions. They also select candidates for the first trigger stage, the L0 trigger, which makes first decisions already 4 µs after a proton-proton interaction. Overall, the calorimeter setup at LHCb follows the classical arrangement: an electromagnetic calorimeter (ECAL) is followed by a hadronic calorimeter (HCAL), where the ECAL is *e.g.* responsible for the e identification.

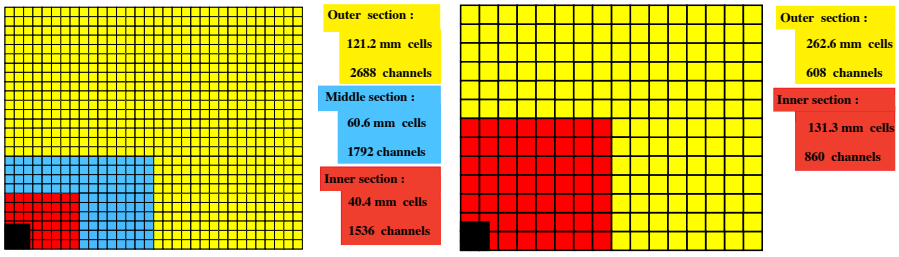

Fig. 5.7: Lateral division of the PS, SPD and ECAL (left) and HCAL (right). For all components one quarter of the front view of the detector can be seen. The area cut out for the beam pipe is shown in black. [50].

To suppress backgrounds from charged pions, the Preshower (PS) detector is installed upstream of the ECAL. For the trigger, backgrounds from π^0 with high transverse energy E_T are suppressed by the scintillating pad detector (SPD). Since the hit density varies by two orders of magnitude over the calorimeter surface, the lateral division of the calorimeters increases closer to the beam pipe (see Fig. 5.7).

The muon chambers

The detection of muons is of fundamental importance at LHCb. They are used in the reconstruction of many *CP*-sensitive decays, such as $B^0 \to J/\psi K^0_S$ in which the J/ψ is reconstructed in the decay into two muons, as well as in the reconstruction of rare B decays with flavour-changing neutral currents, such as $B^0_s \to \mu^+\mu^-$. There are five muon chambers at LHCb. The chamber M1 is located upstream of the calorimeters, while M2 to M5 are positioned downstream of the HCAL. The individual stations consist of 80 cm thick iron absorbers, so that the minimum momentum of a muon to pass all five stations must be 6 GeV. Stations M1 to M3 have a relatively high spatial resolution along the x-coordinate. They are mainly used to identify the track directions and to measure the transverse momentum p_T of the muon candidates with a resolution of 20 %. The stations M4 and M5 are used for particle identification of traversing muons.

5.2.3 Trigger

In contrast to the general-purpose experiments ATLAS and CMS, the LHCb experiment does not operate at the maximum luminosity of $L = 7 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ provided by the LHC in Run I but at a constant instantaneous luminosity of $L = 4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ in order to reduce the pile-up per bunch crossing. Nevertheless, to be stored, the data rate must still be reduced from about 20 MHz to about 4 kHz. Two trigger stages are available for this purpose: the first stage (L0) works synchronously to the interaction rate of 40 MHz and reduces it to 1 MHz, whereupon the second trigger stage, the high-level trigger (HLT), processes the data independently of the interaction rate. At a luminosity

of $L = 4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, events containing a B meson are generated at a rate of about 15 kHz. Furthermore, many partial decay widths of the B mesons are smaller than 10^{-3}, so that the trigger system is optimized to select these interesting decays for subsequent analyses with maximum efficiency, while the backgrounds created in the hadronic environment at the LHC are suppressed to the minimum.

The Level 0 trigger is a pure hardware trigger. It identifies hadrons, electrons and photons which have maximum transverse energies E_T, as well as the two muons with the highest transverse momenta p_T. The L0 consists of three components: A L0 pile-up system, the L0 calorimeter trigger and the L0 muon trigger. The aim of the pile-up system is to distinguish between events with one or more visible proton-proton interactions. The calorimeter and muon components search for the maximum E_T or p_T for the corresponding particles.

The HLT is a C++ application and runs on the Event Filter Farm (EFF), a large-capacity computer at CERN. Every application has full access the full information of an event. However, since the data rate coming from the L0 trigger is very high, the HLT consists of two stages. The HLT1 reconstructs the partial candidates in the VELO and in the tracking stations, which the L0 transfers. In addition for photons and neutral pions the absence of charged particles that could be associated with these candidates is confirmed. Overall, the HLT1 has the task of reducing the data rate to such an extent that a full full pattern recognition is possible in the next step, the HLT2. At a sufficiently low data rate, the HLT2 then even reconstructs specific B decays .

5.3 The LHCb software stack

The LHCb software is based on the Gaudi framework [57], in which different software packages are running. The order in which the various packages are executed is shown in Fig. 5.8. The first action on the data recorded by the detector happens in the HLT through the software package Moore [58, 59] as described in Sec. 5.2.3. Afterwards the raw data is reconstructed in Brunel [60–62] and the particles are combined in so-called protoparticles. These protoparticles contain track information and particle identification (PID) information of the concerning particle. The data is then available as so-called Data Summary Tape Files (DSTs) and is further processed for analysis in DaVinci. In the DaVinci software project, a first preselection (stripping) takes place and the final reconstruction of the different decays is done. After the stripping, the data has the form of so-called nTuples, which can be analysed by individual analysts.

The generation and simulation of events within the LHCb detector is implemented in the Gauss package. The proton-proton-collisions and the hadronisation process are generated using Pythia [63, 64], which runs with a special LHCb configuration [65]. Decays and the interaction with the detector are then simulated using the EvtGen [66] and Geant4 [67, 68] packages, respectively. This is followed by the Boole project, in which the data is digitized so it can then be further processed in the same way as the raw data collected by the detecor.

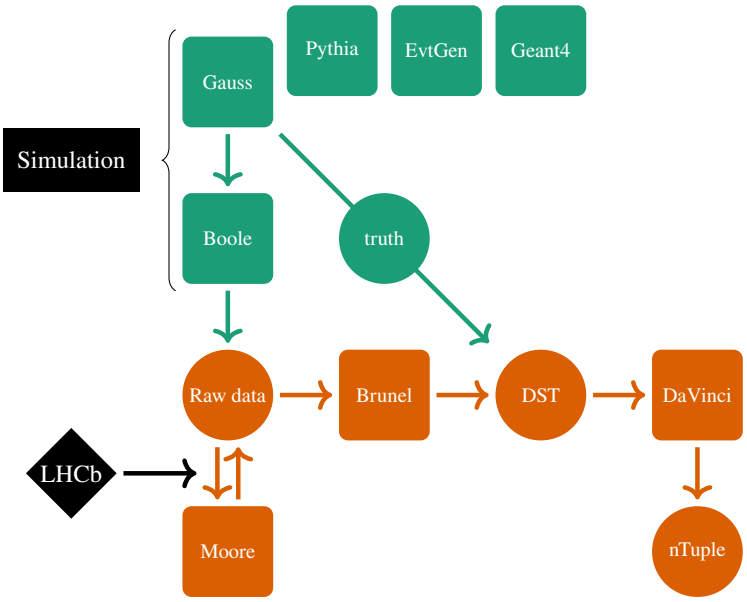

Fig. 5.8: Sequence of data processing within the LHCb software. In red the steps to simulate data, in blue the processing of real data is shown. The software packages are displayed in rectangles, the transferred data formats in circles.

Furthermore, the generated information (truth information) is stored and can be retrieved after the simulated events have been processed. Thus, in addition to the detector response, the initially generated states are also known for simulated events, which is important to check the reliability of reconstruction steps or experimental techniques such as the flavour tagging described in Ch. 9.

6 Experimental techniques

In this chapter several experimental techniques used in the analysis are introduced. First, three statistical tools are described: a multivariate classifier called boosted decision tree (BDT) is described in Sec. 6.1 based on Ref.[69]. In the analysis, BDTs are used in the selection of $B^0 \to D^{\mp}\pi^{\pm}$ candidates and when determining the production flavour of the B^0 mesons. The maximum-likelihood method which is used to fit the invariant B^0 mass and to estimate the *CP* asymmetries is detailed in Sec. 6.2 [69]. The last statistical tool, the *sPlot* technique [70], is described in Sec. 6.3, which is used to statistically separate signal from background candidates. Finally, the flavour tagging is introduced in Sec. 6.4, which provides algorithms to infer the initial flavour of B^0 and B_s^0 mesons at LHCb.

6.1 Boosted decision trees

Decision trees are multivariate classifiers, *i.e.* they analyse multiple randomly distributed variables simultaneously in contrast to simple univariate analyses that examine each variable individually. A decision tree classifies a data set of different types of events by applying hierarchically ordered logical rules to different properties of these events. The tree always consists of a root node and an arbitrary number of sub nodes, as well as at least two leaves. For binary decision trees, as used in the scope of this thesis, the rule applied at each node has only two possible outcomes. Figure 6.1 shows a decision tree with a depth of two, *i.e.* two consecutive binary decisions are made before the events are assigned to the classes A and B, depending on their final leave. In principle, each decision tree starts applying a rule to the variable with the best separation power (here v_1) in the root node. Subsequently, further rules are applied in the lower sub nodes until certain limits are reached, *e.g.* a minimum number of events in each leave.

Before applying a decision tree to a data sample, first it needs to be trained with data samples labelled according to their class. Furthermore, an estimate of the separating power of the different variables is needed. An easy way for example is to maximise the difference $\Delta N = N_r - N_w$ between right (N_r) and wrong (N_w) assignments in a node. However, the cut point of the binary rule remains random to some extent, since the value ΔN only changes when the cut-point is shifted by so much that it hits the nearest variable value. A slightly more complicated, but very popular measure is the following: the purity P of events from class A

$$P_A = \frac{N_A}{N_A + N_B}\,, \tag{6.1}$$

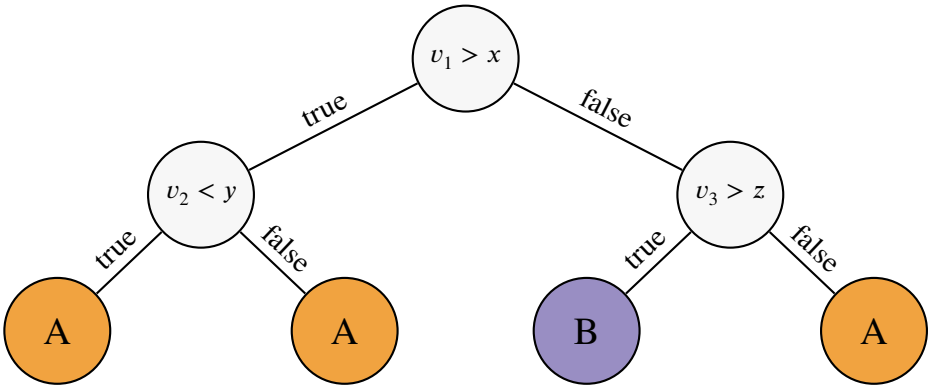

Fig. 6.1: Binary decision tree with a depth of two, dividing events into the classes A and B based on the variables v_i.

where N_i are the number of events from class A and B, becomes 0 or 1 if the classification is perfect. Therefore, minimising the Gini Index [69]

$$G = P_A \left(1 - P_A\right) + P_B \left(1 - P_B\right) \tag{6.2}$$

provides presumably the variable with the largest separation power. As the selection of variables depend on the chosen variables and applied cuts in higher nodes of the decision tree, correlations are automatically taken into account.

The overall classification performance can be improved by not only using one moderately effective classifier but instead calculating the weighted average of many different decision trees. In order to do so, many decision trees are grown. Each time a new tree is built, events, which were assigned to the wrong class by the previous tree, are weighted (boosted) in order to reduce the probability that they are wrongly classified again. Such a combination of decision trees is denoted as boosted decision tree (BDT). Well known boosting algorithms are the gradient boosting technique [71] or the adaptive boosting (AdaBoost) method [72]. Since the latter is used in the selection of $B^0 \to D^{\mp}\pi^{\pm}$ candidates in Ch. 7 and the development of flavour-tagging algorithms in Ch. 9 it is shortly described here.

Considering a boosted decision tree consisting of T trees, where the trees are distinguished by the subscript t. The first decision tree is trained with N events equally weighted providing a hypothesis $h_1(x^i)$ being 1 (-1) for signal (background) for the i-th event. This hypothesis is compared with the label y_i and the total error rate is calculated as

$$\varepsilon_1 = \sum_{i=1}^{N} p_1^i \left| h_1(x^i) - y^i \right| , \tag{6.3}$$

where $p_1^i = w_1^i / \sum_i w_1^i$ is the normalised weight for each event. This can be generalised by replacing the subscript 1 with t. Using this error rate the weights for the next tree are

computed using the boost weight

$$\alpha^i = \left(\frac{\varepsilon_t}{1-\varepsilon_t} \right)^{\beta\left(1-\left|h_t(x^i)-y^i\right|\right)} \tag{6.4}$$

as $w^i_{t+1} = \alpha w^i_t$, where β is the boosting factor, which, in simple terms, modifies the learning rate during the training. The final BDT output is then given by

$$h_f(x^i) = \frac{1}{T}\sum_{i=1}^{T} \ln\left(\alpha^i\right) \times h_t(x^i)\,, \tag{6.5}$$

which is distributed between -1 and 1, implying that an event is more likely signal (background) when the value is close to 1 (-1).

When training a BDT, some caveats must be taken into account. On the one hand, the labelled data samples used in the training are usually only proxies for the *real* data. For example, differences between simulated events and *real* data can lead to a bad performance of the BDT. Furthermore, the BDT can learn to distinguish the training data samples by statistical fluctuations. This phenomenon, denoted as over-fitting or overtraining, can be avoided by splitting the labelled training data sample before developing the BDT. The first half is then used as training data set and following the BDT is applied to the second data set denoted as test sample. If the BDT output distributions on both data sets are the same, it can be assumed that there is no overtraining.

To train BDTs, various implementations exist. In the course of this thesis always the implementation from TMVA [73] is used.

6.2 The maximum-likelihood method

The maximum-likelihood method is a common tool for parameter estimation from a data sample. In simple terms, the parameter values are selected as an estimate according to which the shape of the observed data appears most probable. This estimation is possible in either one or multiple dimensions. Assuming n measurements of a set of observables $\vec{x}$ the maximum-likelihood function is given as

$$\mathscr{L}(\vec{a}) = \mathscr{P}(\vec{x}_1|\vec{a}) \times \mathscr{P}(\vec{x}_2|\vec{a}) \times ... \times \mathscr{P}(\vec{x}_n|\vec{a}) = \prod_{i=1}^{n} \mathscr{P}(\vec{x}_i|\vec{a})\,, \tag{6.6}$$

where $\mathscr{P}(\vec{x}_i|\vec{a})$ are the properly normalised probability densitiy function (PDF) with a set of parameters $\vec{a}$ to estimate. The function $\mathscr{L}(\vec{a})$ gives the probability of obtaining the measured values for a set of parameters $\vec{a}$ in a sample $\vec{x}_i$. However, even if the maximum-likelihood function becomes maximal for the maximum probability of obtaining the data set $\vec{x}_i$, it is not a probability density in the parameters $\vec{a}$. Extending the function in Eq. (6.6) with a Poisson term

$$\mathscr{L}(\vec{a}) = \frac{e^{-n}n^N}{N!} \prod_{i=1}^{n} \mathscr{P}(\vec{x}_i|\vec{a})\,, \tag{6.7}$$

where n is the number of expected events, although the sample containts N measurements allows to distinguish different categories of events by summing up several likelihood functions. Usually, the negative logarithmic likelihood function is minimised because this is numerically more stable and it leads to the same results, since the logarithm is a monotone function.

Besides, it is possible to use an external input to constrain a parameter μ to be $\mu_0 \pm \Delta\mu$ by means of a Gaussian function. This implies that the likelihood is multiplied by a Gaussian with the mean and width set to μ_0 and $\Delta\mu$, respectively. In this analysis, the maximum-likelihood fits were implemented using the Root framework [74], which makes use of the Minuit package [75] for the minimisation of the likelihood function.

6.3 The sPlot technique

The *sPlot* technique [70] uses a maximum-likelihood fit to calculate the so-called *sWeights* by performing a *sPlot* fit to one or multiple discriminating observables. Considering a data sample, containing a mixture of N_c different categories of events, the *sWeights* are per-event weights ${}_s w$, which allow to reconstruct the distributions of variables separately for each category present in the initial sample. However, one important assumption is that the *sWeights* are applied to observables, which are independent of the discriminating observables. To perform a *sPlot* fit the PDFs for all categories of events are needed, so that the *sWeights* can be calculated as

$$ {}_s w = \frac{\sum_{j=1}^{N_c} V_{nj}\, f_j(\vec{y}_e)}{\sum_{k=1}^{N_c} N_k\, f_k(\vec{y}_e)}, \tag{6.8} $$

where the sums iterate over all categories of events. Moreover, the functions f_i are the corresponding PDFs of the discriminating set of observables $\vec{y}$ for an event e, N_k is the yield in the corresponding category and V the covariance matrix of the yields. In practice, a first fit to the observables $\vec{y}$ is performed to determine the parameters of the PDFs $f_i(\vec{y})$, before all parameters except for the yields are fixed and the *sWeights* are calculated in a second fit. The normalisation of the *sWeights* is such that the sum over the weights for one category provides the number of events N of this category in the sample. The statistical uncertainty on this number of events is defined for each bin δx by

$$ \sigma_N = \sqrt{\sum_{e \subset \delta x} ({}_s w)^2}\,. \tag{6.9} $$

In the scope of this analysis, signal and background candidates for the signal decay $B^0 \to D^{\mp}\pi^{\pm}$ and for the flavour tagging control modes $B^0 \to J/\psi K^{*0}$ and $B^+ \to D^0\pi^+$ are separated using this technique by performing fits to the invariant mass distributions (more details in Ch. 8 and Ch. 9). The choice of the invariant mass as discriminating observable has two advantages: on the one hand, it is independent of the decay-time, for which the

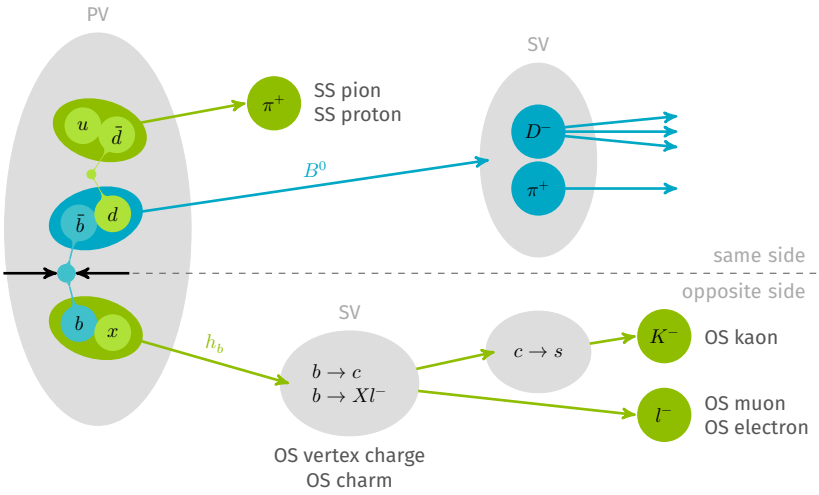

Fig. 6.2: Schematic overview of all available B^0 tagging algorithms.

distribution of signal candidates is needed to *e.g.* extract the *CP* asymmetries. On the other hand, the distributions of the different contributions in the distribution are well known and allow a reliable parametrisation.

6.4 Flavour tagging

To measure interference *CP*-violation the production flavour of *B*-mesons under study must be known. At LHCb this is inferred using the so-called flavour tagging. A decision (tag) d whether a B candidate was initially produced as a B^0-meson or a $\overline{B}^0$-meson and a probability-estimate (mistag) η of being wrong with this decision is provided by the flavour tagging algorithms (taggers). They can be divided into two classes: opposite side (OS) and same side (SS) algorithms. In the following, a general description of the different algorithms available at LHCb, their performance characteristics and their calibration is given.

6.4.1 Tagging algorithms

At LHCb several tagging algorithms exist to infer the initial B flavour of which some differ for B^0 and B_s^0 mesons. In Fig. 6.2 a schematic representation of the tagging algorithms for B^0 mesons is shown. They can be separated into so-called opposite side (OS) and same side (SS) algorithms.

The OS algorithms exploit the production and decay of the second b-quark which is produced in the proton-proton collision. By partially reconstructing single decay products as electrons, muons, kaons and D-mesons associated with the decay of the opposite side b-hadron the initial flavour is inferred. Furthermore, charged tracks which originate from

a secondary vertex, which is displaced from the PV, are used to make a decision on the production flavour of the signal B-meson. As the hadronisation and the decay of the OS b-hadron is independent of the signal B-meson, these algorithms can be used for both B^0 and B_s^0 mesons. Based on [76, 77] the OS algorithms are briefly described below:

- The OS muon and OS electron tagger use the charge of muons and electrons from semileptonic $b \to X l^-$ decays to take a decision on the initial B-flavour. The charged leptons are selected using a simple cut-based selection. To suppress contributions from $b \to c \to l^+$ decays, which would give the wrong tag decision, for example the transverse momentum of the muon (electron) is required to be larger than 1.2 GeV/c (1.0 GeV/c). Electrons have additionally to satisfy criteria on electron identification variables such as the ratio $E/p > 0.8$. Here E denotes the energy deposited in the ECAL and p the electron momentum. If more than one muon or electron per event survives the selection, the lepton with the highest transverse momentum is chosen to define the flavour of the signal B. The mistag is estimated with an artificial neural network, which takes as inputs event properties as the number of PVs and tracks in the event, B-properties as the transverse momentum and various geometrical and kinematic properties of the tagging lepton.

- The OS kaon tagger explores the charge of kaons produced in the decay chain $b \to c \to s$. Very similar to the lepton taggers the tagging kaon is selected using rectangular cuts based on kinematic and PID observables. In case multiple kaons per event pass this selection, the kaon with the highest transverse momentum is fed into an artificial neural network with similar inputs as for the lepton taggers to calculate the mistag estimate η.

- The OS charm tagger selects D-mesons produced via $b \to c$ decays. In case of a charged D-meson the charge of the meson directly hints at the initial flavour, in case of an uncharged D-meson the charge of the produced kaon is used to infer the flavour of the signal B-meson. In contrast to the other single track taggers, a BDT is used to select the D-meson and estimate the mistag. As the OS charm is the newest development on the OS it was developed to have a small overlap concerning the used tagging particles with the other taggers.

- The OS vertex charge tagger is the only algorithm which does not reconstruct single particles, but uses the weighted charge of a secondary vertex (SV) associated with the opposite side b-hadron instead. In order to do this, the track pair with the highest probability of originating from the opposite side b-hadron is used to build a vertex. Following, particles which are compatible with coming from this two-track vertex but not from the PV are added to it. Finally all tracks of the final SV are weighted

with their transverse momentum, p_{T}, and used to calculate a charge

$$Q_{\mathrm{vtx}} = \frac{\sum_i p_{\mathrm{T}}^k(i) Q_i}{\sum_i p_{\mathrm{T}}^k(i)}\,, \tag{6.10}$$

where the parameter k is optimised to maximise the performance of the tagging algorithm. Based on this charge the initial flavour of the signal B-meson is then determined.

The SS algorithms use remnants of the hadronisation of the signal B-meson to infer the initial flavour. As the companion quark of the b-quark is different for B^0 and B_s^0-mesons, different algorithms must be used to deduce the initial flavour of the signal B.

In case of a B^0 ($\overline{b}d$) a free $\overline{d}$-quark is produced which can hadronise to a pion or proton. Additionally, the production mechanisms, *e.g.* via the strong decay of an excited $B^{**+} \to B^{(*)0}\pi^+$ can be exploited [78]. The SS pion and SS proton taggers use the charge of these companion particles to infer a tag decision. They were developed on $B^0 \to D^{\mp}\pi^{\pm}$ decays assuming $S_f = S_{\bar{f}} = 0$, *i.e.* the decay $B^0 \to D^{\mp}\pi^{\pm}$ is CP-conserving. Therefore a potential bias of the analysis cannot be excluded when using these algorithms and they are retrained on $B^0 \to J/\psi K^{*0}$. The basic strategy is similar to the one in Ref. [78]: First, tagging particles from the same region of phase space as the signal B are selected using requirements on PID, kinematic and geometrical observables. Then, these particles are all used to train a BDT, which further selects the final tagging particle and estimates the mistag. This means in case multiple particles per event pass the selection, the SS pion and SS proton taggers do not select the particle with highest transverse momentum but for all particles the BDT response is calculated and the tagging candidate with the largest BDT response is chosen to infer the initial flavour of the signal B.

On the other hand the hadronisation of a B_s^0 ($\overline{b}s$) leads to a $\overline{s}$-quark which can hadronise to a kaon [79]. The SS kaon tagger was developed to identify such kaons and works similar for kaons as the SS pion tagger for pions produced in B^0 events. Since this analysis covers decays of neutral B^0 mesons and the SS kaon tagger is not used, this algorithm is not discussed any further.

6.4.2 Performance characteristics

The predictions of the flavour tagging algorithms are not perfect. Of N reconstructed candidates only N' candidates get a tag d and mistag η assigned, while N_{U} are untagged. The N' candidates can be further divided into N_{W} candidates which are wrongly tagged and N_{R} correctly tagged candidates. These imperfections can be reflected by a tagging efficiency

$$\varepsilon_{\mathrm{tag}} = \frac{N_{\mathrm{R}} + N_{\mathrm{W}}}{N_{\mathrm{R}} + N_{\mathrm{W}} + N_{\mathrm{R}} + N_{\mathrm{U}}} \tag{6.11}$$

and a mistag probability

$$\omega = \frac{N_{\mathrm{W}}}{N_{\mathrm{R}} + N_{\mathrm{W}}} \,. \tag{6.12}$$

Therefore, in an analysis using flavour tagging to infer the initial B flavour, the number $N_{B^0}(t)$ ($N_{\bar{B}^0}(t)$) of measured initial B^0 ($\overline{B}^0$) candidates are

$$\begin{aligned} N_{B^0}(t) &= (1-\omega) N^{\mathrm{true}}_{B^0}(t) + \omega N^{\mathrm{true}}_{\bar{B}^0}(t) \,, \\ N_{\bar{B}^0}(t) &= \omega N^{\mathrm{true}}_{B^0}(t) + (1-\omega) N^{\mathrm{true}}_{\bar{B}^0}(t) \end{aligned} \tag{6.13}$$

where $N^{\mathrm{true}}_{B^0}$ and $N^{\mathrm{true}}_{\bar{B}^0}$ denotes the true number of initial B^0 and $\overline{B}^0$ candidates, respectively. These quantities need to be transferred further into measurements of CP-asymmetries such as

$$A_{CP}(t) = \frac{N^{\mathrm{true}}_{B^0}(t) - N^{\mathrm{true}}_{\bar{B}^0}(t)}{N^{\mathrm{true}}_{B^0}(t) + N^{\mathrm{true}}_{\bar{B}^0}(t)} \,. \tag{6.14}$$

For a measured asymmetry the true numbers of initial B^0- and $\overline{B}^0$-mesons need to be replaced with the observed yields which leads to

$$A^{\mathrm{meas}}_{CP}(t) = \frac{N_{B^0}(t) - N_{\bar{B}^0}(t)}{N_{B^0}(t) + N_{\bar{B}^0}(t)} = (1-2\omega) A^{\mathrm{theo}}_{CP}(t) \tag{6.15}$$

with the dilution $D = 1 - 2\omega$. However, experimentally not only the dilution affects the measured asymmetry but also intrinsic asymmetries I like an unequal production of B^0- and $\overline{B}^0$-mesons might influence a measurement so that the measured asymmetry can be expressed as

$$A^{\mathrm{meas}}_{CP}(t) = D A_{CP}(t) + I \,. \tag{6.16}$$

As the mistag probability is defined in the range $[0, 0.5]$ the dilution can take values between 0 and 1. A large dilution factor is equivalent to a vanishing mistag and hence leads to a smaller experimental sensitivity as will be shown below. To simplify the following discussion the intrinsic asymmetries and dilution are assumed to be time-independent, even if that is not generally valid. The theoretical asymmetry can then be expressed as

$$A_{CP}(t) = \frac{1}{D} \left(A^{\mathrm{meas}}_{CP}(t) - I \right) \,. \tag{6.17}$$

Assuming that all quantities are uncorrelated and gaussian distributed the uncertainty on the theoretical asymmetry is given by

$$\begin{aligned} \sigma^2_{A_{CP}} &= \left(\frac{\partial A_{CP}}{\partial N_{B^0}} \right)^2 \sigma^2_{N_{B^0}} + \left(\frac{\partial A_{CP}}{\partial N_{\bar{B}^0}} \right)^2 \sigma^2_{N_{\bar{B}^0}} + \left(\frac{\partial A_{CP}}{\partial I} \right)^2 \sigma^2_I + \left(\frac{\partial A_{CP}}{\partial D} \right)^2 \sigma^2_D \\ &= \frac{1}{D^2} \frac{1}{N_{B^0}(t) + N_{\bar{B}^0}(t)} \left(1 - A^{\mathrm{meas}}_{CP}(t) \right) + \frac{\sigma^2_I}{D^2} + \frac{A^2_{CP}(t)}{D^2} \sigma^2_D \,. \end{aligned} \tag{6.18}$$

Neglecting the uncertainties on the intrinsic asymmetries and dilution factor and further assuming that the measured asymmetries are small, this expression can be reduced to

$$\sigma^2_{A_{CP}} = \frac{1}{D^2}\frac{1}{N_{B^0}(t) + N_{\overline{B}^0}(t)}\,. \tag{6.19}$$

Here it is useful to identify the number of measured B^0- and $\overline{B}^0$-candidates as

$$N_{B^0}(t) + N_{\overline{B}^0}(t) = \varepsilon(t)N \tag{6.20}$$

where ε includes all efficiencies either in the trigger, reconstruction, selection or the flavour tagging. Regarding the flavour tagging this means that the uncertainty on a *CP*-asymmetry is given by

$$\sigma_{A_{CP}} = \frac{1}{\sqrt{\varepsilon_{\text{tag}} D^2 N}} = \frac{1}{\sqrt{\varepsilon_{\text{eff}} N}} \tag{6.21}$$

where the effective tagging efficiency $\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} D^2$ is introduced. As one can see, this efficiency defines the experimental sensitivity of the measurement as it effectively reduces the number of candidates. It also becomes obvious that the tagging efficiency introduced in Eq. (6.11) and the mistag probability defined in Eq. (6.12) are not individually suitable for determining the performance of different tagging algorithms. Instead the tagging efficiency, which is also denoted as tagging power, must be used.

Rather than using an average mistag ω the mistag estimate η of the various tagging algorithms can be used. To do this, the estimated mistag has to be calibrated with a calibration function $\omega(\eta)$ (more details on the calibration are given in Sec. 6.4.3), what gives a per-event tagging power, defined as

$$\varepsilon_{\text{eff}} = \frac{1}{N}\sum_{i=1}^{N} D_i^2 = \frac{1}{N}\sum_{i=1}^{N}\left(1 - 2\omega(\eta_i)\right)^2\,. \tag{6.22}$$

Here the sum is iterating over all candidates and for untagged candidates the mistag probability is defined to be 0.5 ($D_i = 0$).

When considering only tagged candidates, the effective tagging efficiency reduces to a pseudo tagging power, which is effectively the same as the average of the dilution squared in the respective sample:

$$\left\langle D^2 \right\rangle = \frac{1}{N_{\text{R}} + N_{\text{W}}}\sum_{i=1}^{N_{\text{R}}+N_{\text{W}}}\left(1 - 2\omega(\eta_i)\right)^2\,. \tag{6.23}$$

6.4.3 Combination and calibration of flavour tagging algorithms

To improve the overall performance of the flavour tagging, the individual algorithms are combined to form one single tag decision and mistag for the OS (d_{OS} and η_{OS}) and a tag

decision and mistag for the SS (d_{SS} and η_{SS}). This is done by calculating the combined probability, that a B-candidate contains a b-quark

$$P(b) = \frac{p(b)}{p(b) + p(\bar{b})} \quad \text{and} \quad P(\bar{b}) = 1 - P(b)\,. \tag{6.24}$$

The probabilities $p(b)$ and $p(\bar{b})$ are defined as

$$p(b) = \prod_i \left(\frac{1 + d_i}{2} - d_i \left(1 - \eta_i\right) \right) \tag{6.25}$$

and

$$p(\bar{b}) = \prod_i \left(\frac{1 - d_i}{2} + d_i \left(1 - \eta_i\right) \right) \tag{6.26}$$

where d_i (η_i) are the tag decisions (mistag estimates) of the individual tagging algorithms. The combined tag decision and mistag are now defined as $d = -1$ and $\eta = 1 - P(b)$ if $P(b) > P(\bar{b})$ and as $d = +1$ and $\eta = P(b)$ if $P(b) < P(\bar{b})$.

As mentioned before, the output of the flavour tagging algorithms is mostly the result of multivariate classifiers, which are trained on flavour specific B decays. This output is then transformed into a mistag estimate η and crosschecked on another flavour specific validation sample. However, the training and validation samples are usually different from the signal decay used in a CP-violation measurement. This differences are caused by different trigger and selection criteria and can influence the distributions which are used by the multivariate classifier to estimate the mistag. Consequently, the mistag must be calibrated on a dedicated flavour specific decay, which shows at best kinematically similar distributions compared to the signal decay. For the OS taggers this is most often done using charged decay modes, as the charge of the final state particles allows to directly infer the production flavour. Instead, to calibrate the SS taggers a decay mode with the same initial B flavour is needed, as these taggers highly depend on the hadronisation process. In the course of this analysis, the OS taggers are calibrated using $B^+ \to D^0\pi^+$, where the bachelor pion allows to infer the initial flavour, while the SS taggers are calibrated using $B^0 \to J/\psi K^{*0}$.

So far, for all analyses at LHCb a linear calibration function of the form

$$\omega(\eta) = \tilde{p}_0 + \tilde{p}_1 \left(\eta - \langle\eta\rangle\right)\,, \tag{6.27}$$

where the arithmetic mean $\langle\eta\rangle$ of the estimated mistag is used to decorrelate the calibration parameters p_0 and p_1, was sufficient. Using this calibration function, a perfect calibration, *i.e.* $\omega = \eta$, would result in $\tilde{p}_0 = \langle\eta\rangle$ and $\tilde{p}_1 = 1$. Yet, the performance of the tagger can depend on the initial B flavour: If the interaction rates of charged decay products (*e.g.* the kaons used by the OS kaon tagger) with the detector material depend on the charge, the mistags will also depend on the flavour of the initial B meson. Such asymmetry yields in an additional dilution factor, which needs to be understood and properly described. This

can be achieved by using different calibration functions for B^0- and $\overline{B}^0$-mesons. In the simple linear case these calibration functions are

$$\begin{aligned} \omega &= p_0 + p_1 + (\eta - \langle\eta\rangle) \ , \\ \overline{\omega} &= \overline{p}_0 + \overline{p}_1 + (\eta - \langle\eta\rangle) \ . \end{aligned} \tag{6.28}$$

These different calibration parameters are furthermore linked to each other via the average calibration parameters $\tilde{p}_i$ used in Eq. (6.27) and corresponding differences Δp_i defined as

$$\tilde{p}_i = \frac{p_i + \overline{p}_i}{2} \quad \text{and} \quad \Delta p_i = p_i - \overline{p}_i \ . \tag{6.29}$$

However, due to the large number of $B^0 \to D^{\mp}\pi^{\pm}$ signal candidates small effects that were hidden in the statistical uncertainties before become significant. Therefore more sophisticated calibration functions are needed to calibrate the flavour tagging algorithms. Possible models are the so-called generalised linear models (GLM) [80] haing the following form:

$$\overset{(-)}{\omega}(\eta) = g\left(h(\eta)\right) = g\left(g^{-1}(\eta) + \sum_{i=1}^{N}\left(\tilde{p}_i \overset{(-)}{+} \frac{\Delta p_i}{2}\right) f_i(\eta)\right) . \tag{6.30}$$

The functions f_i are denoted as *basis functions*, which for example can be simple poynomials or natural spline functions [81]. To minimise the correlation between the $\tilde{p}_i$ and Δp_i parameters the *basis functions* are orthogonalised using the Gram-Schmidt method [82]. The function g is referred to as *link function*, which is usually defined as the inverse cumulative distribution function, to map all input values to the range $[0, 1]$, *i.e.* an interval which can be interpreted as a probability. As the mistag is only defined in the range $[0, 0.5]$ it is possible, that after applying a calibration function with this *link function* the mistag is larger than 0.5. In this case an arbitrary decision has to be taken how such candidates are further treated. Possible options are to either flip the corresponding tag decision $d \to -d$ and adjust the mistag $\omega \to 1 - \omega$ or to mark the candidate as untagged with a mistag of 0.5. The latter possibility leads to fit instabilities in the decay-time fit when the calibration parameters are not fixed, but allowed to float or constrained by means of a Gaussian function. This is due to the fact that for a floating or constrained calibration the ratio of tagged and untagged candidates varies during the minimisation and the changes to the likelihood are not continous. On the other hand, a flip of the tagdesicion may yield in a bias of the *CP*-parameters as shown in Sec. 10.2.1. Therefore, the modified logistic function

$$g(h) = \frac{1}{2\left(1 + e^{h}\right)} \tag{6.31}$$

is used in the course of this thesis as *link function*, which maps all input values into the range $[0, 0.5]$ and hence assures that all calibrated mistags are well defined.

To not rely on possible binning effects on the estimated mistag η or the mistag probability ω, the calibration functions are determined using an unbinned maximum-likelihood method, the so-called binomial regression [83].

7 Data sample and selection

This analysis is done on the data sample recorded by the LHCb experiment in 2011 and 2012 at centre-of-mass energies of 7 and 8 TeV, respectively. In 2011, the detector collected 1 fb^{-1}, while in 2012 2 fb^{-1} were recorded.

Before doing a *sPlot* fit physical backgrounds with different *CP* characteristics turning up in the same invariant mass region as the signal $B^0 \to D^{\mp}\pi^{\pm}$ candidates need to be removed as they cannot be distinguished in this dimension. Furthermore, as much combinatorial background as possible is rejected for two main reasons. On the one hand a cleaner sample simplifies the parametrisation of the invariant mass as the signal shape becomes more significant and on the other hand the background contamination, which dilutes the *sWeights* for the final decay-time fit in Ch. 10 is reduced. In this chapter, the used data samples as well as the simulated samples are described in Sec. 7.1. The selection procedure is reported in Sec. 7.2, divided into preselection and trigger requirements (Sec. 7.2.1), vetoes to suppress *e.g.* misidentified background candidates (Sec. 7.2.2) and a multivariate classifier to reduce combinatorial background (Sec. 7.2.3 and Sec. 7.2.4). Last, the handling of multiple B candidates in one event is presented in Sec. 7.2.5 and the selection performance is given in Sec. 7.2.6.

7.1 Data and simulation samples

Candidates from the decay $B^0 \to D^{\mp}\pi^{\pm}$[1] are reconstructed in the hadronic $D^{\mp} \to K^{\pm}\pi^{\mp}\pi^{\mp}$ decay with one additional pion, which will be denoted as bachelor pion in the following. The D decay is chosen despite the purely hadronic final state, as it is the one with the largest decay width. It is reconstructed inclusively, *i.e.* no resonances such as the decay via a K^{*0} meson into a $K\pi$ final state are excluded.

To distinguish between the charged hadrons in the final state, a likelihood function assuming the respective particle to be a pion or kaon (proton) is computed for every particle using information from the PID system. Then, the difference between the two logarithmic likelihoods is calculated, which in the following is referred to as $\text{DLL}_{K\pi}$ ($\text{DLL}_{p\pi}$) [51].

As the PID observables are not described well in the simulation, selection requirements on such observables can result in different distributions in other correlated observables. Therefore the $\text{DLL}_{K\pi}$ and $\text{DLL}_{p\pi}$ variables are corrected using calibration samples of kinematically-clean $D^{*+} \to D^0\left(\to K^-\pi^+\right)\pi^+$ decays. The correction is done in bins of

[1] Charge conjugation is implied throughout the whole document if not stated otherwise

Tab. 7.1: Simulated samples used in this analysis with a short note in which analysis step the samples are used and the number of available candidates before applying any analysis specific selection step. Charged D mesons are always generated with the decay $D^- \to K^+\pi^-\pi^-$, uncharged D mesons with the decay $\overline{D}^0 \to K^+\pi^-$.

Sample	Analysis step	Candidates [10^6]
$B^0 \to D^\pm \pi^\mp$	all steps	3.2
$B_s^0 \to D_s^- \left(\to K^+K^+\pi^-\right) \pi^+$	selection	1.2
$\Lambda_b^0 \to \Lambda_c^- \left(\to K^+\overline{p}\pi^-\right) \pi^+$	selection	0.46
$B^0 \to D^- K^+$	mass fit	0.26
$B^0 \to D^- \rho^+ \left(\to \pi^+\pi^0(\to \gamma\gamma)\right)$	mass fit	0.62
$B^0 \to D^{*-} \left(\to D^-\pi^0\right) \pi^+$	mass fit	0.16
$B^0 \to D^- K^{*+} \left(\to K^+\pi^0\right)$	mass fit	0.03
$B^0 \to J/\psi \left(\to \mu^+\mu^-\right) K^{*0} \left(\to K^+\pi^-\right)$	flavour tagging	3.5
$B^+ \to \overline{D}^0 \pi^+$	flavour tagging	5.1
$B^+ \to \overline{D}^0 K^+$	flavour tagging	0.05
$B^+ \to \overline{D}^* \left(\to D^0\gamma\right) \pi^+$	flavour tagging	0.06
$B^+ \to \overline{D}^0 K^{*+} \left(\to K^+\pi^0\right)$	flavour tagging	0.04
$B^0 \to \overline{D}^0 \pi^+\pi^-$	flavour tagging	0.05

transverse momentum p_T and pseudorapidity η. For every candidate in the simulation, the corresponding PID distribution of the calibration sample in the (p_T, η)-bin is built and used to randomly sample a PID value [84]. Possible effects due to the chosen (p_T, η)-binning are evaluated in the systematic uncertainties described in Ch. 11. The simulated samples used in this analysis are listed in Tab. 7.1, together with a short reference, in which analysis step they are needed.

Last, to obtain the correct correlations and uncertainties between vertex positions, particle momenta, decay times and invariant masses, kinematic fits to the whole decay chain are performed on data and simulated events [85]. These fits allow to determine several parameters such as decay times, particle momenta, track positions and the corresponding uncertainties and correlations. In total, three of these fits are performed: To determine observables correlated with the decay time, the PV is constrained to the known position of the proton-proton collision. Observables correlated with the invariant mass stem from a fit, where the mass of the D^- meson is constrained to its known mass of $m_{D^-}^\mathrm{PDG} = 1869.61\,\mathrm{MeV/c^2}$ [18]. A third fit is performed without any constraint as this would lead to wrong results for selection steps like optimising the vetoes described in Sec. 7.2.2.

7.2 Selection

As a first step, the so-called stripping is applied to build $B^0 \to D^{\mp}\pi^{\pm}$ candidates with $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$. The stripping is a first loose preselection common to a set of kinematically similar decays. Events with more than 500 tracks, which are constructed from hits in the VELO and the tracking stations T1 to T3, are rejected. The criteria on the charged tracks depend on whether the charged track is considered as the bachelor particle, or as a D^- daughter. Three of these charged tracks are then used to form a D^- meson, where the (transverse) momentum of one of the three tracks has to exceed (500 MeV/c) 5 GeV/c and its track χ^2/ndof has to be less than 2.5. This D^- meson is then combined with a charged bachelor track to form a B^0 meson. Finally a boosted decision tree trained on simulation is applied, and its response is required to be larger than 0.05. These requirements together with all cuts on single particles given in Tab. 7.2, consisting of cuts to the vertex χ^2_{IP}, where the χ^2_{IP} is defined as the difference in the vertex-fit of a given PV reconstructed with and without the originating particles, momenta, track and vertex fit qualities and flight direction quantities, aim to select B^0 decays into a charged $D^{\mp}$ meson including $D^{*\mp}$ mesons and pion, where the $D^{\mp}$ mesons decays into three hadrons from combinatorial backgrounds candidates. Thereby, the term combinatorial background means candidates originating from random combinations of tracks in an event. Following the stripping, a decay-specific selection described in the following is applied.

7.2.1 Preselection and trigger requirements

Before applying further selections to reduce the various background components, requirements on the trigger are made. In principle there are two different classes of trigger decisions at LHCb: a trigger can fire due to a particle or event property directly connected to the signal decay - denoted as trigger on signal (TOS) - or it can fire due to some property separate to the signal decay what is denoted as a decision independent of the signal (TIS). Furthermore, each trigger stage has various lines, triggering on different event properties and thus being differently effective depending on the specific decay. Hence, a requirement on which trigger line has fired and whether this decision is TOS or TIS results in characteristic distributions of observables such that a decision which requirement is made needs to be done analysis specific.

For this analysis, no specifc requirements at the L0 level are applied, *i.e.* events from all available L0 trigger lines and also both TOS and TIS triggered events are accepted. On the HLT1 level, the B^0 candidates are required to be TOS on the `Hlt1TrackAllL0Decision` line. At the HLT2 trigger stage, the $B^0 \to D^{\mp}\pi^{\pm}$ candidates are required to form a SV out of two, three or four tracks with a significant separation from the PV, *i.e.* they need to be TOS on one of the `Hlt2Topo` lines. More details on the trigger lines can be found in Ref. [86]. In Fig. 7.1, the invariant mass distributions of the B^0 and D^- candidates are shown. The B^0 peak is clearly visible together with structures from the partially

Tab. 7.2: Stripping cuts for the decay $B^0 \to D^{\mp}\pi^{\pm}$ with $D^{\pm} \to K^{\mp}\pi^{\pm}\pi^{\pm}$. For the charged tracks, the more stringent requirements on the bachelor pion are given in brackets. The decay vertex of the B^0 meson is denoted as SV, for the impact parameter the shortcut IP is used and the distance of closest appraoch of the D^- daughter particles w.r.t. each other is denoted as DOCA.

charged tracks requirements	
track χ^2/ndof	$< 3.0(2.5)$
momentum p	$> 1(5)$ GeV/c
transverse momentum p_{T}	$> 100(500)$ MeV/c
χ^2_{IP} w.r.t. any PV	> 4.0
track ghost probability	< 0.4
D^- meson requirements	
$\sum p_{\mathrm{T}}\,(hhh)$	> 1800 MeV/c
DOCA	< 0.5 mm
m_{D^-}	1769.92 to 2068.49 MeV/c^2
SV χ^2/ndof	< 10.0
vertex separation χ^2 to any PV	> 36.0
cos of $\sphericalangle\left[\lvert \mathrm{PV}, D^-\text{-Vtx}\rvert, \vec{p}(D^-)\right]$	> 0.0
B^0 meson requirements	
SV χ^2/ndof	< 10.0
reconstructed decay time t	> 0.2 ps
χ^2_{IP} w.r.t. the associated PV	< 25.0
cos of $\sphericalangle\left[\lvert \mathrm{PV}, \mathrm{SV}\rvert, \vec{p}\left(B^0\right)\right]$	> 0.999

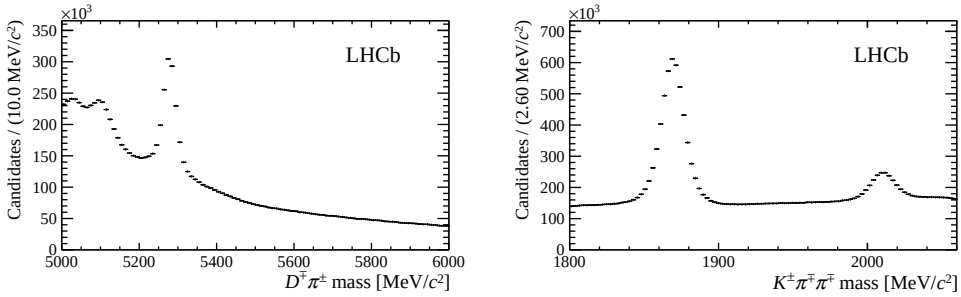

Fig. 7.1: Invariant mass distributions of the $D^-\pi^+$ combination using the kinematic decay chain fit with a D^- mass constraint (right) and of the $K^+\pi^-\pi^-$ combination without any constraint (right).

Tab. 7.3: Preselection cuts applied after the trigger requirements.

D^- daughter requirements	
$\mathrm{DLL}_{K\pi}$ for pions	< 8.0
$\mathrm{DLL}_{K\pi}$ for kaons	> -2.0
D^- and B^0 meson requirements	
$\left\lvert m_{K^+\pi^-\pi^-} - m_{D^-}^{\mathrm{PDG}} \right\rvert$	$< 35\,\mathrm{MeV/c^2}$
B^0 decay time	$> 0.2\,\mathrm{ps}$

reconstructed decays $B^0 \to D^-\rho^+$ and $B^0 \to D^{*-}\pi^+$ in the lower mass region. The distribution of the invariant mass of the D meson shows the D^- peak around $1870\,\mathrm{MeV/c^2}$ and a D^{*-} peak at $2010\,\mathrm{MeV/c^2}$. After the trigger requirements, some loose sanity cuts to remove clear combinatorial background candidates are applied. These cuts are listed in Tab. 7.3.

7.2.2 Background vetoes

When a multivariate classifier is trained on and applied to a data sample, this sample should at best only consist of the candidates which are supposed be separated by the classifier. In this case the classifier described in Sec. 7.2.3 is supposed to separate $B^0 \to D^{\mp}\pi^{\pm}$ signal candidates and combinatorial background candidates. Therefore, backgrounds due to kinematic failures as misidentifications of particles in the reconstruction and wrong associations between a B^0 candidate and a PV are vetoed.

Mass vetoes

First, the investigated sources of backgrounds which arise due to failures in the reconstrucion, and if necessary, the applied vetoes are described. Such failures can be missed neutral particles or misidentified particles in the reconstruction.

The first type of backgrounds arises from the decay $B^0 \to D^-\mu^+\nu_\mu$. When missing the neutrino and misidentifying the muon as a pion this decay would falsely be reconstructed as the signal decay $B^0 \to D^{\mp}\pi^{\pm}$. This is vetoed with a binary requirement on the bachelor particle not to be a muon. This binary requirement is based on the momentum of the track of the corresponding particle and the number and the region of the muon stations where hits are found [87].

The following kinematic backgrounds are all due to misidentification of particles in the final state. Misidentification between protons and pions can lead to backgrounds from $\Lambda_b^0 \to \Lambda_c^-\pi^+$ decays where the Λ_c^- decays into a kaon, a pion and an antiproton. To identify such background candidates a proton-mass hypothesis is applied to both daughter pions of the D^- meson. After combining the three D daughters again a peak around the

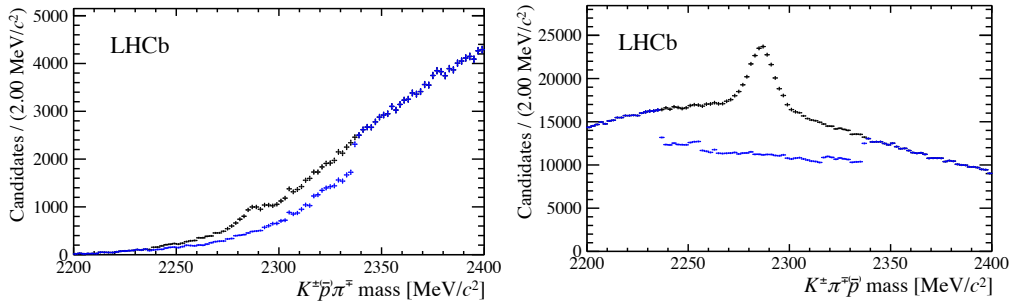

Fig. 7.2: Invariant mass distributions of the $K\pi p$ combinations for both daughter pions of the D^- meson. The distributions are shown without the veto (black) and with the veto applied (blue). The left (right) plot shows the proton-mass hypothesis applied to the pion with lower (higher) transverse momentum.

Λ_c^+ mass becomes visible in the distributions shown in Fig. 7.2. The distributions look different, because the pions are originally sorted by transverse momentum. To remove this background a two-stage veto is applied: In the first stage candidates are rejected if the invariant mass of the three hadrons is inside a $\pm 30\,\mathrm{MeV/c^2}$ window around the nominal Λ_c^- mass $m^{\mathrm{PDG}}_{\Lambda_c^-} = 2286.46\,\mathrm{MeV/c^2}$ and the $\mathrm{DLL}_{p\pi}$ is larger than -8.0. At the second stage, the mass window is enlarged to $\pm 50\,\mathrm{MeV/c^2}$ around the nominal Λ_c^- mass, but the $\mathrm{DLL}_{p\pi}$ requirement is loosened, only requiring $\mathrm{DLL}_{p\pi} > -5.0$. This strategy is adopted in order to remove all Λ_c^- candidates, where the width of the mass windows was optimised on simulated Λ_c^- candidates showing a resololution about $20\,\mathrm{MeV/c^2}$, after applying the reconstruction for $B^0 \to D^{\mp}\pi^{\pm}$ candidates and subsequently again the proton mass hypothesis. After the preselection, $(99.720 \pm 0.004)\,\%$ of the $\Lambda_b^0 \to \Lambda_c^-\pi^+$ candidates are rejected. This veto rejects another $(76.6 \pm 0.6)\,\%$ at a signal efficiency of $(93.48 \pm 0.06)\,\%$.

In the same way as pions and protons can be misidentified, kaons can be falsely identified as one of the D^- daughter pions. Such misidentification gives rise to a potential background contamination from $B_s^0 \to D_s^-\pi^+$ candidates. As previously, the kaon mass hypothesis is applied to both pions and the three D^- daughters are combined. However, these distributions do not show a clear mass peak. Therefore, they are compared for different kinematic regions of the invariant $\left[K^+\pi^-\pi^-\right]\pi^+$ mass: after applying the kaon mass hypothesis to the pions, $B_s^0 \to D_s^-\pi^+$ candidates should end up in a B_s^0 signal region, whereas no $B_s^0 \to D_s^-\pi^+$ candidates are expected in the upper-mass sideband $m_{[K^+\pi^-\pi^-]\pi^+} > 5500\,\mathrm{MeV/c^2}$. Therefore, the invariant mass distributions of the $KK\pi$ system are compared for candidates from the B_s^0 signal range $5330\,\mathrm{MeV/c^2}$ to $5400\,\mathrm{MeV/c^2}$ and a background range $5500\,\mathrm{MeV/c^2}$ to $5700\,\mathrm{MeV/c^2}$. The visible difference in Fig. 7.3 stems from the fact that those two distributions arise from different kinematic regions. To further ensure that no significant contamination from $B_s^0 \to D_s^-\pi^+$ candidates is present in the data, resonances like K^{*0}- or ϕ mesons, which arise in possible D_s^- decays are studied. Those resonances would become visible in the KK (ϕ) and $K\pi$ (K^{*0}) invariant mass distributions. Figure 7.4

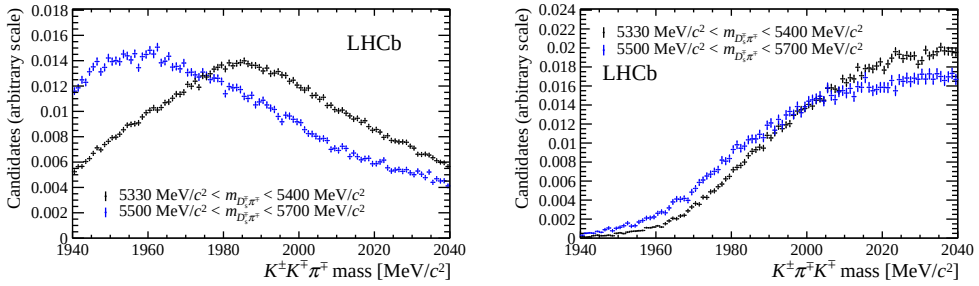

Fig. 7.3: Invariant mass distributions of the $KK\pi$ combinations for both daughter pions of the D^- meson. The distributions are shown in the B_s^0 signal region from 5330 MeV/c^2 to 5400 MeV/c^2 (black) and in a background region from 5500 MeV/c^2 to 5700 MeV/c^2 (blue) In the left (right)plot the kaon-mass hypothesis is applied to the pion with lower (higher) transverse momentum.

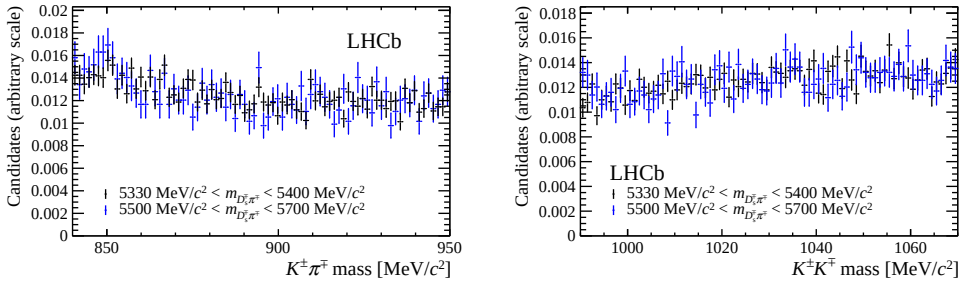

Fig. 7.4: Invariant mass distributions of the $K\pi$ (left) and KK (right) combinations for the daughter pion of the D^- meson with larger transverse momentum with the daughter kaon or remaining daughter pion, respectively. The distributions are shown in the in the B_s^0 signal region from 5330 MeV/c^2 to 5400 MeV/c^2 (black) and in a background region from 5500 MeV/c^2 to 5700 MeV/c^2 (blue). Only candidates with 1940 MeV/c^2 < $m_{KK\pi}$ < 2040 MeV/c^2 are considered.

shows representatively the invariant mass distributions of the KK ($K\pi$) combinations of the the daughter kaon with the pion with larger transverse momentum under the kaon mass hypothesis (under the initial pion mass hypothesis). In addition to the beforehand defined signal and background regions, only candidates within a range from 1940 to 2040 MeV/c^2 from Fig. 7.3 are considered for these plots as the invariant mass of true D_s^+ candidates would be in this range. Consequently, one would expect a peaking structure in these plots in case of a significant contamination with $B_s^0 \to D_s^- \pi^+$ decays. As no distribution in Fig. 7.3 and Fig. 7.4 shows such a structure, background candidates from $B_s^0 \to D_s^- \pi^+$ decays are assumed to be negligible.

The last considered background comes from $B^0 \to \overline{(}D\overline{)}K\pi$ decays, arising due to a kaon-pion misidentification of the bachelor particle or a D^- daughter particle followed by a combination of the bachelor particle with a D^- daughter particle. Performing the

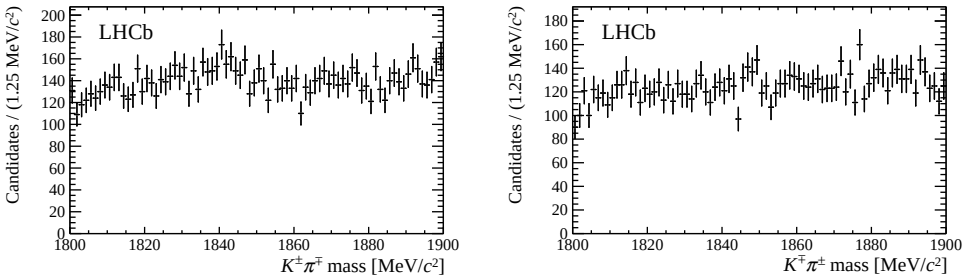

Fig. 7.5: Invariant mass distributions for the combinations of the bachelor particle with the D^- daughter pion with higher transverse momentum. The distributions shows the kaon hypothesis is applied to the bachelor pion (left) and the kaon hypothesis applied to the D^- daughter pion(right).

four possible combinations of the bachelor particle with the kaon mass hypothesis applied with the two D^- daughter pions shows that all distributions have a flat shape. Hence, backgrounds from $B^0 \to \overline{(D)}^0 K\pi$ decays are assumed to be negligible as well. In Fig. 7.5, the combinations of the bachelor track with the D^- daughter pion with higher transverse momentum are shown for illustration.

Wrongly associated PVs

The average number of pp-collisions per bunch crossing at LHCb is $\nu = 2.5$. Therefore, a considerable amount of events has more than one PV and in these events a B^0 candidate can be associated with each of them. Besides that, an event can also contain more than one B^0 candidate; in this case the B^0 candidate is chosen randomly (more details in Sec. 7.2.5). However, in case of multiple PVs per event, the B^0 candidate can be associated with the wrong PV leading to an incorrect decay time for this candidate. Usually a decay-time dependent selection efficiency is expected at LHCb, which strongly increases at small decay times up to ≈ 2 ps and shows a flat or slightly dropping distribution for high decay times. This efficiency, further denoted as decay-time acceptance, is caused by the track reconstruction in the VELO and certain trigger requirements (more details are given in Sec. 10.1.2). Yet, due to the wrong PV association, a large unexpected tail at high decay times arises. This can be checked on simulation, where the true decay time is known. Weighting each (B^0,PV)-pair with an exponential using the true lifetime of the B^0 candidates, shows an excess of (B^0,PV)-pairs at high decay times (see Fig. 7.6). In the simulation, the true z-position of the PV is known and can be compared with the reconstructed z-position. Requiring that the distance between the true z-position of the PV and the reconstructed z-position does not exceed five times the uncertainty on the reconstructed z-position removes the wrong associated (B^0,PV)-candidates from the sample and the expected decay time acceptance becomes visible (Fig. 7.6). As this approach cannot be applied to data, a criterion is developed requiring for each (B^0,PV)-candidate

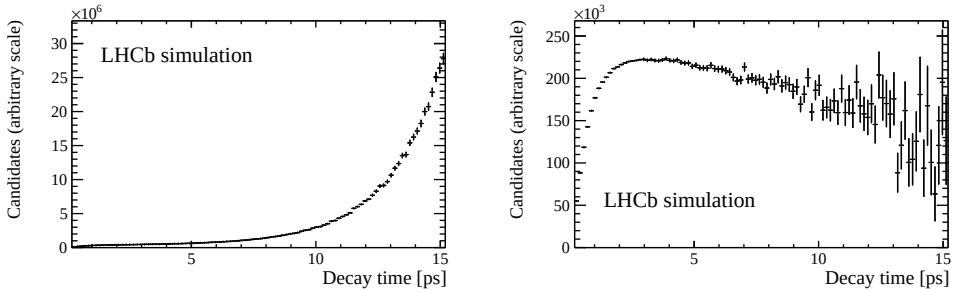

Fig. 7.6: Decay time distribution of simulated $B^0 \to D^{\mp}\pi^{\pm}$ candidates, weighted with an exponential using the true lifetime of the B^0 candidates. At high decay times an excess of (B^0,PV)-candidates can be seen (left). After applying a cut on the distance between the true and the reconstructed z position of the PV the expected acceptance distribution at LHCb is visible (right).

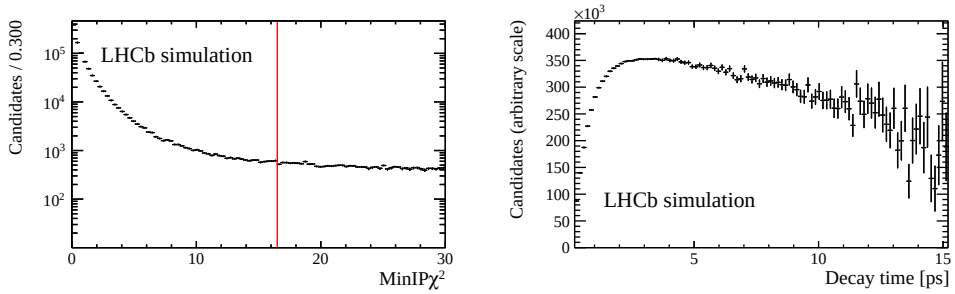

Fig. 7.7: Distributions of the MinIPχ^2 variable with the chosen cut point at 16.5 in a narrow range (left) and the decay time of simulated $B^0 \to D^{\mp}\pi^{\pm}$ candidates, wheighted with an exponential using the simulated B^0 lifetime after applying the cut on MinIPχ^2 (right). The MinIPχ^2 the distribution also remains flat for larger values.

the impact parameter χ^2 (χ^2_{IP}) with any other PV in the event (denoted as MinIPχ^2) to be larger than a certain value. This means that in case this MinIPχ^2 is too small, the two corresponding PVs cannot be distinguished sufficiently well from each other, and the (B^0,PV)-candidate is rejected. Events which contain just one PV are always kept. The cut on the MinIPχ^2 variable is optimised such that 98 % of the events in the simulated $B^0 \to D^{\mp}\pi^{\pm}$ sample are retained. In Fig. 7.7, the distribution of the MinIPχ^2 variable together with the cut point at 16.5 and the resulting decay-time-acceptance distribution of simulated candidates after rejecting candidates with MinIP$\chi^2 \leq 16.5$ is shown.

In contrast the PV could also be chosen using some “best” PV criterion, *e.g.* choosing the PV with the smallest χ^2_{IP} with respect to the B^0 candidate. Such a criterion also removes most of the wrongly associated PVs, but potentially biases the decay time distribution, whereas the strategy described above treats all PVs equally and thus the decay time distribution remains unbiased.

Tab. 7.4: List of input variables used in the training of the BDT

B^0 candidate	cos of $\sphericalangle\left[\left\|\mathrm{PV},\mathrm{SV}\right\|,\vec{p}(B^0)\right]$
	SV χ^2
D^- candidate	χ^2_{IP} w.r.t. the SV
	χ^2_{IP} w.r.t. the associated PV
	radial flight distance
	flight distance χ^2 w.r.t. the SV
	D^- vertex χ^2/ndof
	transverse momentum p_{T}
	cos of $\sphericalangle\left[\left\|\mathrm{SV},D^-\text{-}\mathrm{Vtx}\right\|,\vec{p}(D)\right]$
bachelor π	χ^2_{IP} w.r.t. the associated PV
	transverse momentum p_{T}
	track χ^2/ndof
D^- daughters	χ^2_{IP} of the associated PV
decay chain fit	χ^2 of the kinematic fit with PV constraint

7.2.3 Development of a MVA classifier

Combinatorial background is suppressed using a multivariate analysis (MVA), more precisely a BDT is used. In this case, $B^0 \to D^{\mp}\pi^{\pm}$ candidates should be separated from combinatorial background candidates and therefore the BDT needs to be provided with proxies for these classes of candidates. As proxy for the $B^0 \to D^{\mp}\pi^{\pm}$ candidates, simulation based on the conditions of the 2012 data taking is used. The upper mass sideband of the 2012 data with $m_{[K^-\pi^+\pi^+]\pi^-} > 5500\,\mathrm{MeV/c^2}$ mimics the background. Both proxy samples are divided into a training and a test sample of same size. Since the BDT should be applied to the dataset after all specific backgrounds like $\Lambda_b^0 \to \Lambda_c^-\pi^+$ or the wrong PV associations are removed, all previous selection steps are applied to the signal- and background-proxy samples for the training. In total, the BDT uses 16 input variables which are listed in Tab. 7.4 and shown in Figs. 7.8 and 7.9. To reduce the number of input variables, from pairs of variables which have a correlation of $> 97\,\%$ with each other the variable with the smaller separation power is removed. The final variables cover mostly various χ^2_{IP} variables, flight distances, momenta and flight directions in order to obtain a classifier output which is independent of the invariant mass of the B^0 meson as this is used in Ch. 8 to further identify signal candidates. The BDT consists of 1700 decision trees with a maximum depth of four. The variables are scanned at 20 points to find the optimal cut value and each node has to contain at least 2.5 % of the training candidates. The boosting algorithm (AdaBoost [72]) is chosen with a boost factor $\beta = 0.5$. The approach to find this configurations is iteratively, *i.e.* the complexity of the BDT is increased as long as no

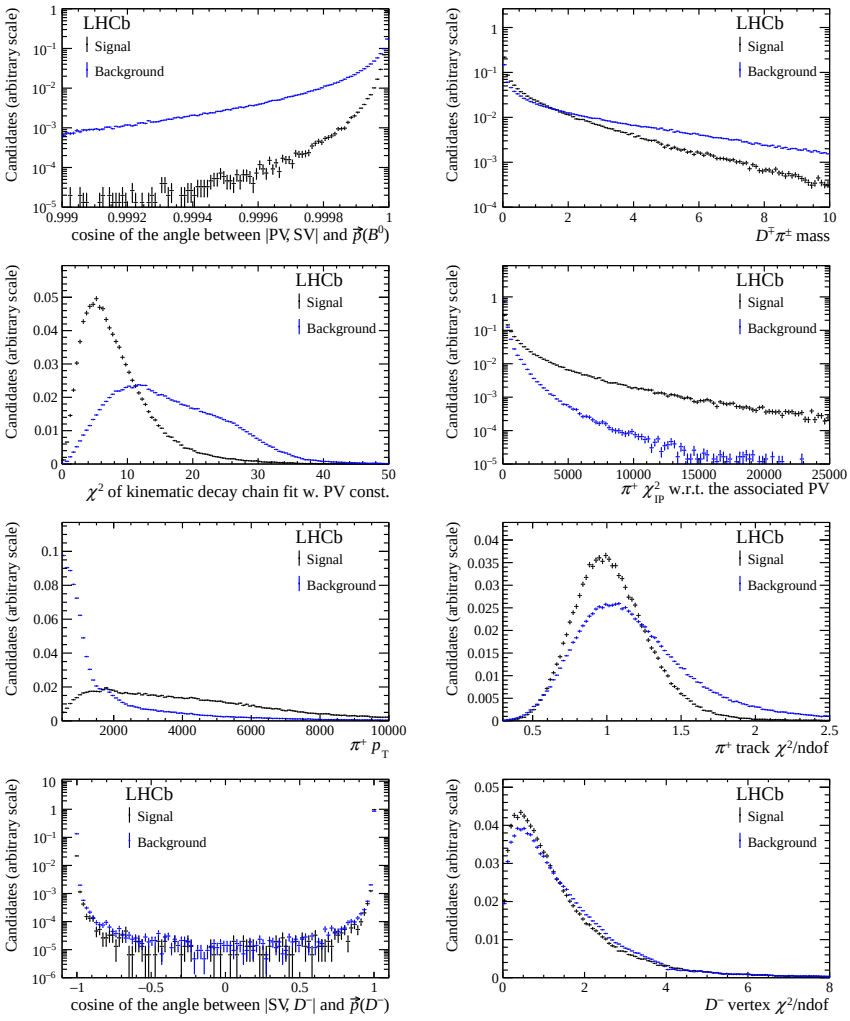

Fig. 7.8: Distributions of the input variables used in the BDT training. The black (blue) points represent the signal-proxy (background-proxy) samples.

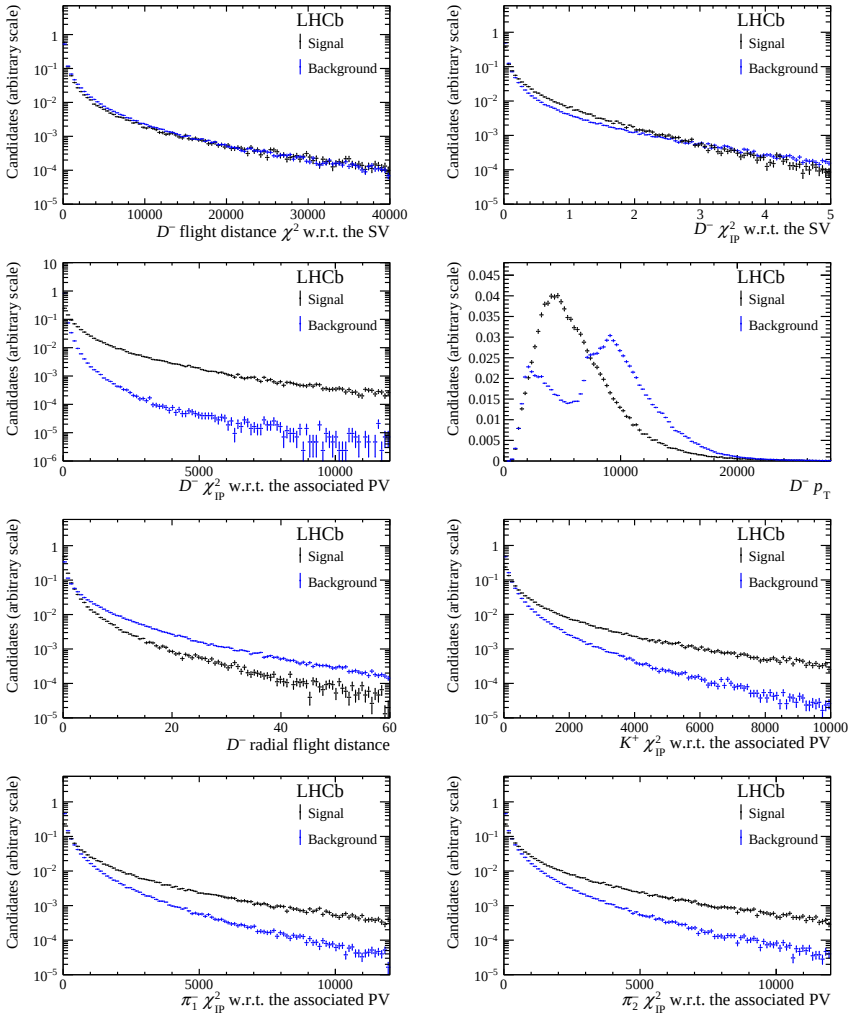

Fig. 7.9: Distributions of the input variables used in the BDT training. The black (blue) points represent the signal-proxy (background-proxy) samples.

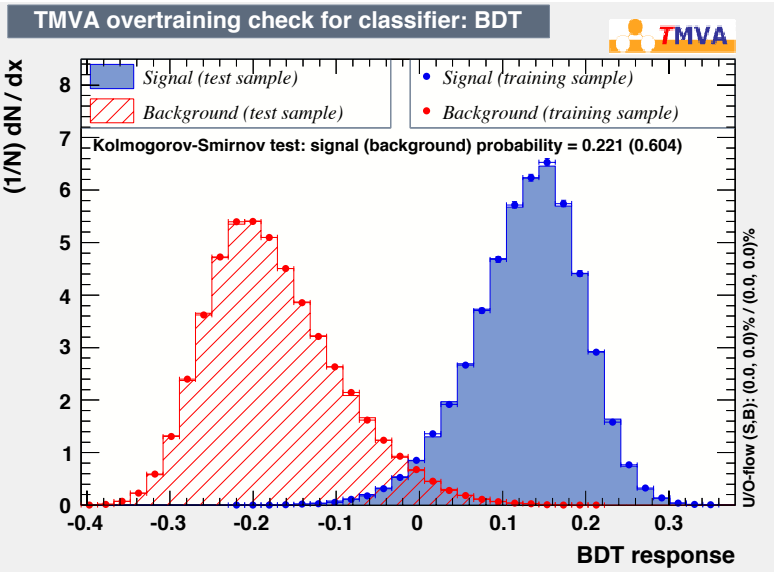

Fig. 7.10: Comparison of BDT responses on the training and test sample.

overtraining is visible. The final overtraining check is shown in Fig. 7.10 where the good agreement between the BDT output distributions on the training and test samples is also indicated at least for the background by a Kolmogorov-Smirnov test [69]. For this test the empirical distribution functions are compared and the supremum of all deviations between the two studied distributions is calculated. The given values represent the confidence values that the distributions are the same.

7.2.4 BDT selection optimisation

To optimise the cut on the BDT response, the uncertainties on the *CP* asymmetries S_f and $S_{\bar{f}}$ are used. For the optimisation, all preselection cuts, the vetoes for background from semileptonic decays and $\Lambda_b^0 \to \Lambda_c^- \pi^+$ decays and the veto for wrongly associated PVs are applied to the full Run I data set. To determine the uncertainty on S_f and $S_{\bar{f}}$ depending on the BDT response the following strategy is adopted: the BDT output is scanned with a step size of 0.01 within a narrow range from -0.15 to 0.1 and with a step size of 0.05 in the outer regions. For each cut on the BDT response a fit to the invariant B^0 mass in the range 5200 MeV/c^2 to 5500 MeV/c^2 using the *pion*-sample is performed. This fit is used to determine *sWeights* [70] and to obtain the mass shape which corresponds to the respective cut on the BDT output. Applying the *sWeights* to other observables such as the decay time makes these distributions appear like signal-only [88]. The tagging efficiencies, the shapes of the mistag distributions of the OS and SS tagging algorithms and the shape of the decay-time acceptance are obtained for these *sWeighted* distributions. The shape of the latter is obtained in the same way as described in Sec. 10.1.2. At each cut point

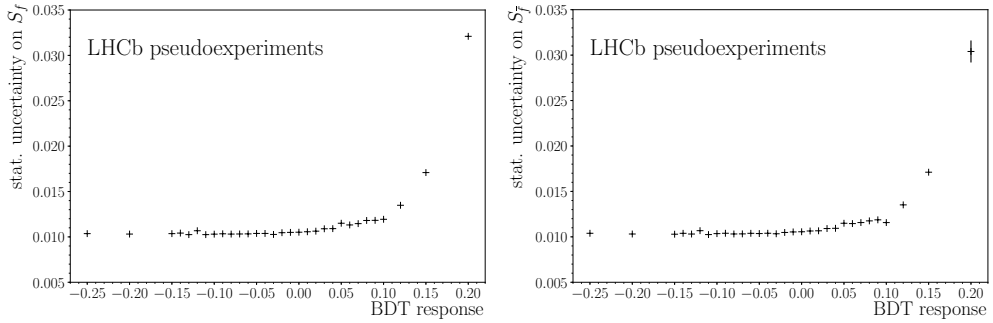

Fig. 7.11: Uncertainty on S_f (left) and $S_{\bar{f}}$ (right) as a function of the BDT response.

the invariant mass distribution is generated using the obtained signal and background yields and shapes from the mass fit. For the mistag and decaytime distribution, signal and background are both generated using the signal shapes. The generated values for the *CP* parameters are taken from simulation. For this pseudoexperiment sample the statistical uncertainty on the *CP* parameters is then determined: after determining *sWeights* from a fit to the invariant B^0 mass distribution, a *CP* fit to the *sWeighted* decay time distribution is performed, where the flavour tagging calibration is assumed to be perfect. In Fig. 7.11, the uncertainties on S_f and $S_{\bar{f}}$ are shown as a function of the BDT response. The final cut point is chosen to be 0.0, to minimise the background as much as possible while achieving the best possible sensitivity.

7.2.5 Multiple *B* candidates

In the stripping and trigger, all used variables rely on the association of the B^0 candidate with the PV to which the candidate has the smallest χ^2_{IP}, also denoted as best PV. To be consistent with this strategy for each event also the best PV in terms of χ^2_{IP} is chosen, after the wrong associations are rejected in Sec. 7.2.2. Events where the formerly best PV is no longer present after the selection are removed. Besides, events can contain more than one B^0 candidate. After the stripping this is the case for 9 % of the events and 18 to 20 % of the B^0 candidates share the same event. This is reduced after applying the described selection steps, so that after selection only 0.4 % of the events contain multiple candidates and 0.8 % of the B^0 candidates share one event. Following the proposal in Ref. [89]. the remaining candidates are assumed to be equally likely signal candidates and one candidate per event is chosen at random.

7.2.6 Selection performance and cross checks

The final selection performance is determined on the upper mass sideband of the data with $m_{[K^-\pi^+\pi^+]\pi^-} > 5500\,\text{MeV/c}^2$ and simulated $B^0 \to D^{\mp}\pi^{\pm}$ candidates. The corresponding signal efficiencies ε_{sig} and background rejections $1 - \varepsilon_{\text{bkg}}$ are given in Tab. 7.5. The low

Tab. 7.5: Performances of all selection steps. As proxy for the $B^0 \to D^{\mp}\pi^{\pm}$ candidates, simulation is used, while the rejection of the combinatorial background is calculated on data for candidates with $m_{D^{\mp}\pi^{\pm}} > 5500\,\mathrm{MeV/c^2}$. All efficiencies are calculated with respect to the previous selection step indicated by the vertical lines. The overall performance is given in the last row.

Selection step	$\varepsilon_{\mathrm{sig}}$	$1-\varepsilon_{\mathrm{bkg}}$
preselection	$(93.61 \pm 0.06)\,\%$	$(85.20 \pm 0.02)\,\%$
$\Lambda_c^{\pm}$-veto	$(93.48 \pm 0.06)\,\%$	$(9.85 \pm 0.03)\,\%$
semileptonic veto	$(98.96 \pm 0.03)\,\%$	$(7.66 \pm 0.03)\,\%$
mass vetoes combined	$(92.51 \pm 0.07)\,\%$	$(16.77 \pm 0.04)\,\%$
wrongly associated PVs veto	$(97.75 \pm 0.04)\,\%$	$(15.81 \pm 0.04)\,\%$
BDT selection	$(83.63 \pm 0.10)\,\%$	$(97.18 \pm 0.01)\,\%$
overall	$(70.7 \pm 0.1)\,\%$	$(99.911 \pm 0.002)\,\%$

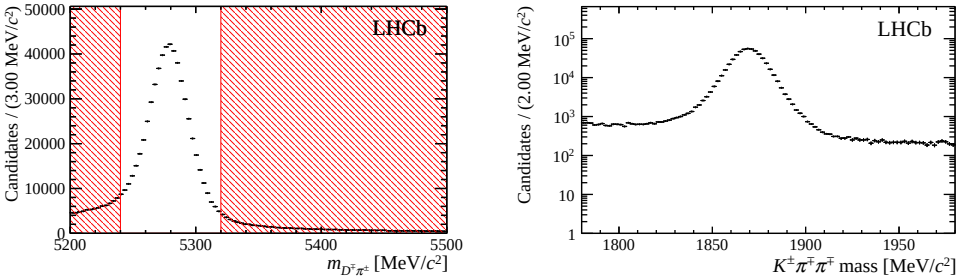

Fig. 7.12: Distribution of the invariant B^0 mass distribution with the selected signal region between the red shaded areas (right) and the resulting invariant D^- mass distribution (right).

background suppression of the vetoes is expected as the vetoes aim to suppress specific backgrounds but not the combinatorial background, which is used to determine the given efficiencies.

Finally, the contamination with non-resonant $B^0 \to K^+\pi^-\pi^-\pi^+$ decays is determined, as such background would show the same structure as $B^0 \to D^{\mp}\pi^{\pm}$ candidates in the invariant $m_{K^+\pi^-\pi^-\pi^+}$ distribution but the *weak* phase would be different. To probe this, two strategies are implemented, both using the data sample after applying all selection steps, except for the cut on the invariant D^- mass given in Tab. 7.3: for the first strategy the invariant D^- mass is plotted for candidates in a tight B^0 signal window from $5240\,\mathrm{MeV/c^2}$ to $5320\,\mathrm{MeV/c^2}$ as shown in Fig. 7.12. Combinatorial background candidates in the resulting D^- mass distribtution could stem from non-resonant $K^{\pm}\pi^{\mp}\pi^{\mp}\pi^{\pm}$ candidates. from this D^- mass distribution, the non-resonant contamination is estimated to a maximum at the percent level . In the second strategy a cut on the invariant D^- mass distribution

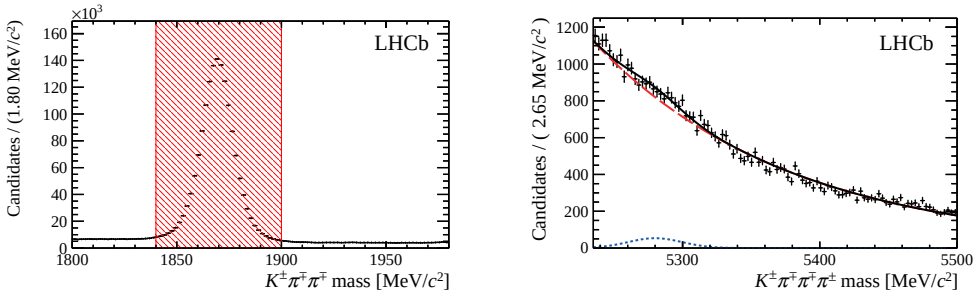

Fig. 7.13: Distributions of the invariant D^- mass with the excluded signal region in the red shaded areas (left) and the resulting invariant B^0 mass distribution with the fit overlaid (right). The red dashed line describes the combinatorial background, the blue dotted line the signal component.

removing the D^- peak is applied. Then the invariant B^0 mass without constraint on the D^- mass is inspected and a fit using an exponential to model the background and a Gaussian with fixed shape from simulation to model the signal is performed to estimate the number of non-resonant $B^0 \to K^+\pi^-\pi^-\pi^+$ candidates (see Fig. 7.13). This fit yields 645 ± 242 signal candidates which is negligible compared to the expected number of $B^0 \to D^\mp\pi^\pm$ signal candidates. Therefore the contamination of the sample with non-resonant background candidates is assumed to be negligible.

8 Mass fit

After the full selection, the data sample is split into two samples using the $\mathrm{DLL}_{K\pi}$ of the bachelor pion, referred to as *pion* and *kaon* sample: the *pion* sample has $\mathrm{DLL}_{K\pi} \leq 5.0$ and the *kaon* sample $\mathrm{DLL}_{K\pi} > 5.0$. This distinction is useful when separating $B^0 \to D^{\mp}\pi^{\pm}$ from $B^0 \to D^- K^+$ candidates as no dedicated selection cut is needed. Instead this separation is done statistically in the fit to the invariant mass distribution.

The invariant B^0 mass distributions of candidates with known initial flavour based on one of the flavour tagging algorithms are fitted simultaneously in these samples in order to calculate *sWeights* [70], which are used in the following analysis steps to statistically separate signal from background candidates. It is important to note that the work described in this chapter was done by a collaborator. However, since this is an essential part of the analysis required to follow the subsequent steps, it is not omitted completely, but the extent to which *e.g.* experimental techniques are described is less comprehensive compared to the other parts.

When parametrising the invariant B^0 mass, all contributing components need to be described. Backgrounds from semileptonic decays, $\Lambda_b^0 \to \Lambda_c^- \pi^+$ decays and $B_s^0 \to D_s^- \pi^+$ decays are either removed in the selection or found to be at a negligible level. However, besides the signal component and the combinatorial background both, the *pion*- and *kaon* sample show additional backgrounds which arise due to missing neutral particles in the reconstruction or pion-kaon-misidentifications. In the *pion* sample contributions from $B^0 \to D^- \rho^+ \left(\to \pi^+ \pi^0\right)$ and $B^0 \to D^{*-} \left(\to D^- \pi^0/\gamma\right) \pi^+$ decays arise, the *kaon* sample shows a component from $B^0 \to D^- K^{*+} \left(\to K^+ \pi^0\right)$ and also $B^0 \to D^- \rho^+ \left(\to \pi^+ \pi^0\right)$ decays which need to be described. Furthermore, both samples show a cross-feed component from each other. The number of cross-feed $B^0 \to DK$ candidates in the *pion* sample is expressed from the yield in the *kaon* sample and vice versa as

$$
\begin{aligned}
N^K_{B^0 \to D\pi} &= \frac{1 - \varepsilon_{\mathrm{PID}}\left(B^0 \to D\pi\right)_\pi}{\varepsilon_{\mathrm{PID}}\left(B^0 \to D\pi\right)_\pi} \times N^\pi_{B^0 \to D\pi} \,, \\
N^\pi_{B^0 \to DK} &= \frac{1 - \varepsilon_{\mathrm{PID}}\left(B^0 \to DK\right)_K}{\varepsilon_{\mathrm{PID}}\left(B^0 \to DK\right)_K} \times N^K_{B^0 \to DK} \,.
\end{aligned}
\tag{8.1}
$$

Here, the quantities $N^Y_{B^0 \to DX}$ $(X, Y = \pi, K)$ are the numbers of $B^0 \to DX$ candidates in the Y sample, while $\varepsilon_{\mathrm{PID}}\left(B^0 \to DX\right)_Y$ are the corresponding efficiencies of the $\mathrm{DLL}_{K\pi}$ requirements on simulation.

Before describing the mass fit to data Sec. 8.2, the probability density functions (PDFs) used for the different components are introduced in the following.

8.1 Probability densitiy functions

The various peaking components in the invariant mass distributions are described by a phenomenological approach, where the description is first estimated on simulated decays. In contrast, the combinatorial background is determined directly in the fit to data. For the *pion* sample, the following PDFs are used to parameterise the peaking components:

- $B^0 \to D^{\pm}\pi^{\mp}$: The signal component is described by a double-sided Hypatia and a Johnson SU function. Assuming a symmetric peak of the distribution, the Hypatia function [90] is defined as

$$\mathcal{J}(m;\mu,\sigma,\lambda,\zeta,a_1,n_1,a_2,n_2) \propto \begin{cases} G(m,\mu,\sigma,\lambda,\zeta), & -a_1 < \frac{m-\mu}{\sigma} < a_2 \\ \frac{G(\mu-a_1\sigma,\mu,\sigma,\lambda,\zeta)}{\left(1-mR_1^{-1}\right)^{n_1}}, & -a_1 > \frac{m-\mu}{\sigma} \\ \frac{G(\mu-a_2\sigma,\mu,\sigma,\lambda,\zeta)}{\left(1-mR_2^{-1}\right)^{n_2}}, & a_2 < \frac{m-\mu}{\sigma} \end{cases} \tag{8.2}$$

with

$$R_i = n_i \frac{G(\mu - a_i\sigma,\mu,\sigma,\lambda,\zeta)}{G'(\mu - a_i\sigma,\mu,\sigma,\lambda,\zeta)} - a_i\sigma\,, \tag{8.3}$$

where $G(m,\mu,\sigma,\lambda,\zeta,a,n)$ is a limit case of the generalised hyperbolic function

$$G(x,\mu,\sigma,\lambda,\zeta) = \left((x-\mu)^2 + A_\lambda^2(\zeta)\sigma^2\right)^{\frac{1}{2}\lambda-\frac{1}{4}} K_{\lambda-\frac{1}{2}}\left(\zeta\sqrt{1+\left(\frac{m-\mu}{A_\lambda(\zeta)\sigma}\right)^2}\right). \tag{8.4}$$

The quantities $K_\lambda(\zeta)$ are the cylindrical harmonics also needed to parametrise

$$A_\lambda^2(\zeta) = \frac{\zeta K_\lambda(\zeta)}{K_{\lambda+1}(\zeta)}\,. \tag{8.5}$$

This parametrisation describes a distribution with a Gaussian core with mean μ and width σ. The parameters λ and ζ are further degrees of freedom to model the shape of the peak while the parameters a_i and n_i describe power-law tails to both sides of the distribution. The Johnson SU function [91] is defined as

$$\mathcal{J}(m;\mu,\sigma,\delta,\gamma) \propto \frac{1}{\sqrt{2\pi\left((m-\mu)^2+\sigma^2\right)}} \times e^{-\frac{1}{2}\left(\gamma+\delta\sinh^{-1}\left(\frac{m-\mu}{\sigma}\right)\right)^2\delta}. \tag{8.6}$$

The mean μ' of this distribution, which features a single peak though its overall shape is completely determined by its parameter values, is given by

$$\mu' = \mu - e^{\frac{1}{2\delta^2}} \sigma \sinh\left(\frac{\gamma}{\delta}\right) . \tag{8.7}$$

Thereby, the parameter μ describes the location of the peak, σ mainly contributes to the width of the distribution and δ and γ parametrise the shape.

- $B^0 \to D^- K^+$: The cross-feed component is parametrised by a double-sided Hypatia function as described in Eq. (8.2).

- $B^0 \to D^- \rho^+ \left(\to \pi^+ \pi^0\right)$: The first partially-reconstructed background is described by a single-sided Crystal Ball function and a Gaussian function. The single-sided Crystal Ball function is defined as

$$\mathscr{CB}(m; \mu, \sigma, \alpha, n) \propto \begin{cases} e^{-\frac{(m-\mu)^2}{2\sigma^2}}, & \frac{m-\mu}{\sigma} > -\alpha \\ A \left(B - \frac{m-\mu}{\sigma}\right)^{-n}, & \frac{m-\mu}{\sigma} \leq -\alpha \end{cases} \tag{8.8}$$

  with

$$A = \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{|\alpha|^2}{2}} \qquad \text{and} \qquad B = \frac{n}{|\alpha|} - |\alpha| . \tag{8.9}$$

  The parameters μ and σ are the mean and width of a Gaussian core, while the parameters a and n describe a power-law tail.

- $B^0 \to D^{*-} \left(\to D^- \pi^0\right) \pi^+$: The second partially-reconstructed component is modelled by the sum of a single-sided Crystal Ball function defined in Eq. (8.8) and a Gaussian function.

For the *kaon* sample the peaking components are modelled as following:

- $B^0 \to D^- K^+$: The signal component ist described by a single-sided Hypatia function. The single-sided Hypatia can be derived from the double-sided Hypatia function as described in Eq. (8.2) by setting the parameters $n_2 = 0$ and $a_2 \to +\infty$, *i.e.* fixing a_2 to a large value.

- $B^0 \to D^\pm \pi^\mp$: The cross-feed component is parametrised by a double-sided Hypatia function as described in Eq. (8.2).

- $B^0 \to D^- \rho^+ \left(\to \pi^+ \pi^0\right)$: The partially reconstructed and further misidentified background is parametrised by a sum of two Gaussian functions. The sum is normalised using fractions f and $1 - f$ for the two Gaussian functions.

- $B^0 \to D^- K^{*+} \left(\to K^+ \pi^0\right)$: The partially reconstructed background is modelled with a Gaussian function.

The combinatorial background is described with a sum of two exponentials in the *pion* sample, while for the *kaon* sample a single exponential function is sufficient.

Tab. 8.1: Fitted yields of the signal $B^0 \to D^{\mp}\pi^{\pm}$ component and the combination of all backgrounds in Fit B.

Parameter	Yield
$N^{\pi}_{B^0 \to D\pi}$	$479\,045 \pm 732$
N^{π}_{bkg}	$34\,381 \pm 300$

8.2 Fit to data

To determine *sWeights* [70], the invariant B^0 mass is fitted in two stages: first, a simultaneous binned extended maximum-likelihood fit in the range $[5090, 6000]\,\text{MeV/c}^2$ of the invariant-mass is performed to the *pion* and *kaon* samples (Fit A). This allows to control the contamination with $B^0 \to D^- K^+$ candidates in the *pion* sample. A binned fit is performed due to the very large number of candidates. In this fit, the mean and width parameters of the signal components in the *pion* and *kaon* samples are floated, while for the tail parameters the values are taken from simulation, multiplied by a scale factor, which is then floated to allow for differences between the simulation and data. All yields in the fit are floating, except for the cross-feed yields, which are constrained using Eq. (8.1), *i.e.* the efficiencies $\varepsilon_{\text{PID}}\left(B^0 \to D\pi\right)_{\pi}$ and $\varepsilon_{\text{PID}}\left(B^0 \to DK\right)_{K}$ are constrained by means of a Gaussian function to their values determined from simulated $B^0 \to D^{\pm}\pi^{\mp}$ and $B^0 \to D^- K^+$ candidates. Furthermore, the mean parameters of the $B^0 \to D^- K^{*+}$ in the *kaon* sample and of the cross-feed $B^0 \to D^- K^+$ and $B^0 \to D^{\mp}\pi^{\pm}$ components are floated. All other parameters are fixed to values determined from simulated events. In Fig. 8.1, the invariant $D\pi$ and DK mass distributions with the fit projections overlaid are shown.

After Fit A, an unbinned maximum-likelihood fit to the *pion* sample is performed for which all background components are combined into one single background component and the shapes are fixed to the values found in Fit A. Furthermore, the fit range is reduced to $[5220, 5600]\,\text{MeV/c}^2$ in order to prevent a dilution of the *sWeights* [70] in later steps of the analysis as all background components, which are neither combinatorial nor close to the signal are removed. In this second fit (Fit B) only two parameters are floating: the yield $N^{\pi}_{B^0 \to D\pi}$ of the signal $B^0 \to D^{\mp}\pi^{\pm}$ component and a yield N^{π}_{bkg} for the combination of all backgrounds. These fitted yields are given in Tab. 8.1.

As a crosscheck, the full sample is also split by year, polarity and final state and the whole fit procedure is repeated. The sums of the yields of the corresponding sub samples are then compared to the result given in Tab. 8.1. All comparisons show satisfactory good agreement as shown in Tab. 8.2.

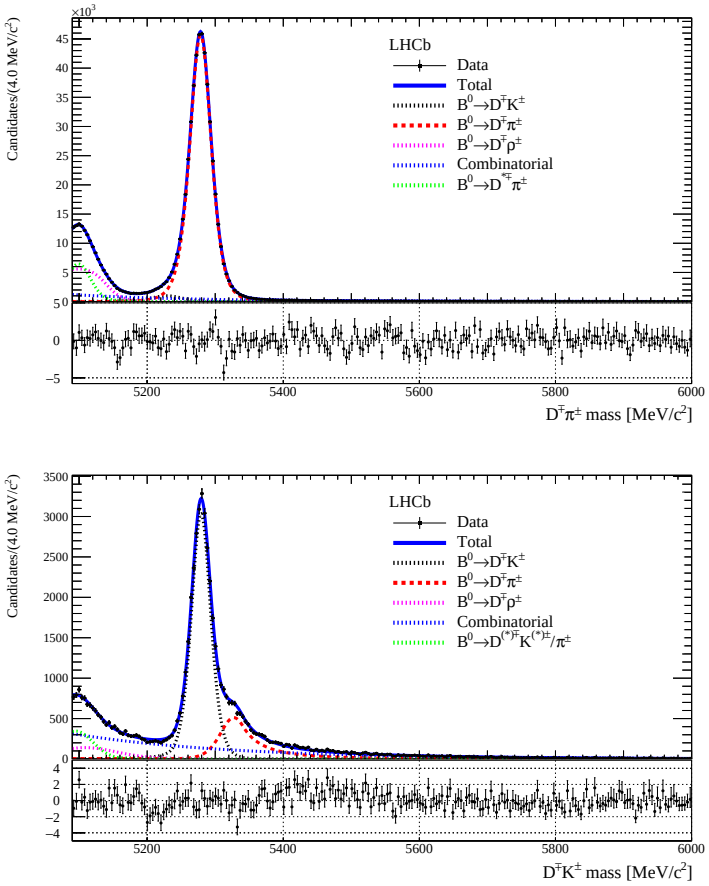

Fig. 8.1: Invariant mass distributions of the DX mass in the *pion* sample (top) and *kaon* sample (bottom). The fit projections of Fit A are overlaid.

Tab. 8.2: Fitted signal yields in fit B to the pion sample split by year of data taking, magnet polarity and finalstate. The last column shows the sum for each split and can be compared with the fitted signal yield in the nominal fit B (Tab. 8.1).

2011	2012	Sum
138 300 ± 400	342 400 ± 600	480 700 ± 700
Magnet up	Magnet down	Sum
226 300 ± 500	2523 ± 500	478 600 ± 700
$D^-\pi^+$	$D^+\pi^-$	Sum
242 100 ± 500	237 300 ± 500	479 400 ± 700

9 Flavour tagging calibration

The aim of the presented analysis is a determination of the CKM angle γ in a time-dependent measurement of *CP* violation in the decay $B^0 \to D^{\mp}\pi^{\pm}$. As can be seen from the time-dependent decay rates in Eqs. (4.32) to (4.35) and the corresponding coefficients in Eqs. (4.36) to (4.38) the angle γ is accessible via the *CP* asymmetries S_f and $S_{\bar{f}}$. To distinguish between the different decay rates and measure these asymmetries, the production flavour of B mesons under study must be known. In this chapter, the tagging strategy for this analysis is outlined in Sec. 9.1 and the required retraining and calibration of the SS and OS tagging algorithms is presented in Sec. 9.2 and Sec. 9.3, respectively. The work presented in the last section was done by a collaborator and is added to delineate the whole analysis procedure, but the explanations are less extensive than for the same side tagging algorithms.

9.1 Flavour tagging strategy

In this analysis the combination of all available OS algorithms is used, while for the SS both taggers designed to infer the initial flavour of B^0 mesons are combined into one single tag decision and mistag. The OS taggers were all developed and trained on a data sample of $B^+ \to J/\psi K^+$ decays, except for the OS charm tagger, whose BDT was trained on a mixed sample of simulated $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to J/\psi \phi$ decays. As the OS taggers are independent of the initial signal B they can be used in this default version. As control channel the decay $B^+ \to D^0\pi^+$ is used. As mentioned before, the SS pion and the SS proton taggers were developed and trained on a data sample of $B^0 \to D^{\mp}\pi^{\pm}$ decays, assuming no *CP* violation. The use of these algorithms could therefore bias the measurement of S_f and $S_{\bar{f}}$ and hence both taggers are retrained (more details given in Sec. 9.2.2 and Sec. 9.2.3).

To include the flavour tagging in the decay rates given in Eqs. (4.32) to (4.35) the expressions for the *CP*-parameters S_f and C_f need to be extended to

$$\begin{aligned} S_f &\to \left(\Delta^- - \Delta^+\right) S_f \,, \\ C_f &\to \left(\Delta^- - \Delta^+\right) C_f \,. \end{aligned} \tag{9.1}$$

Similar equations hold also for $S_{\bar{f}}$ and $C_{\bar{f}}$. The coefficients $\Delta^{\pm}$ contain the calibration

functions and tagging efficiencies $\varepsilon_{\text{tag}}^{\text{OS}}$ and $\varepsilon_{\text{tag}}^{\text{SS}}$. They are defined as

$$\begin{aligned} \Delta^{\pm} = &\frac{1}{2}\varepsilon_{\text{tag}}^{\text{OS}}\left[1-\varepsilon_{\text{tag}}^{\text{SS}}+d^{\text{OS}}\left(1-\varepsilon_{\text{tag}}^{\text{SS}}-2\omega(\eta^{\text{OS}})\left(1+\varepsilon_{\text{tag}}^{\text{SS}}\right)\right)\right] \\ &\pm\frac{1}{2}\varepsilon_{\text{tag}}^{\text{OS}}\left[1-\varepsilon_{\text{tag}}^{\text{SS}}+d^{\text{OS}}\left(1-\varepsilon_{\text{tag}}^{\text{SS}}-2\overline{\omega}(\eta^{\text{OS}})\left(1+\varepsilon_{\text{tag}}^{\text{SS}}\right)\right)\right] \end{aligned} \tag{9.2}$$

for candidates which are only tagged by the OS tagger combination. To obtain the expression for candidates only tagged by the SS taggers, all superscripts need to be exchanged with OS $\leftrightarrow$ SS. For candidates tagged by both, the OS tagger combination and the SS taggers the coefficients can be written as

$$\begin{aligned} \Delta^{\pm} = &\frac{1}{4}\varepsilon_{\text{tag}}^{\text{OS}}\varepsilon_{\text{tag}}^{\text{SS}}\left[1+\sum_{j=\text{OS, SS}} d_j\left(1-2\omega(\eta_j)\right)+d^{\text{OS}}d^{\text{SS}}\left(1-2\omega(\eta_j)+2\omega(\eta^{\text{OS}})\omega(\eta^{\text{SS}})\right)\right] \\ &\pm\frac{1}{4}\varepsilon_{\text{tag}}^{\text{OS}}\varepsilon_{\text{tag}}^{\text{SS}}\left[1+\sum_{j=\text{OS, SS}} d_j\left(1-2\overline{\omega}(\eta_j)\right)+d^{\text{OS}}d^{\text{SS}}\left(1-2\overline{\omega}(\eta_j)+2\overline{\omega}(\eta^{\text{OS}})\overline{\omega}(\eta^{\text{SS}})\right)\right]. \end{aligned} \tag{9.3}$$

The flavour specific control samples ($B^+ \to D^0\pi^+$ for the OS, $B^0 \to J/\psi K^{*0}$ for the SS) are used to determine the functional form of the calibration function $\omega(\eta)$. Unlike other flavour tagged analyses at LHCb the calibration parameters can be determined directly in the decay-time fit as floating nuisance parameters of the likelihood function in $B^0 \to D^{\mp}\pi^{\pm}$. This is possible because the parameters C_f and $C_{\bar{f}}$ are fixed to 1 and -1, respectively, as the ratio r (see Eq. (4.39)) is expected to be too small to allow a significant measurement of these parameters. As a result, the cosine term allows to determine the calibration parameters independently of the sine term. A qualitative explanation of this is given in Fig. 9.1.

When determining all calibration parameters directly in the signal sample no assumptions for the portability of the calibration from a control channel is needed. This portability in principle can be limited due to kinematic differences between the control sample and the signal sample leading to different flavour tagging calibrations. It is probed using simulated events of both, the control sample and the signal sample. When checking the portability of the calibrations from $B^+ \to D^0\pi^+$ and $B^0 \to J/\psi K^{*0}$ to $B^0 \to D^{\mp}\pi^{\pm}$ a bias on S_f and $S_{\bar{f}}$ of roughly the size of their statistical uncertainty arises. This bias can be reduced when floating the calibration parameters (more details on this are given in Sec. 10.2). Furthermore, the precision of the calibration parameters for the SS tagger combination obtained in the decay-time fit to $B^0 \to D^{\mp}\pi^{\pm}$ is much smaller compared to the precision on the control channel, while for the OS tagger combination the uncertainty on the signal sample increases only slightly compared to the uncertainty obtained on the control sample. However, determining the calibration parameters in the final decay-time fit leads to an increased statistical uncertainty on the *CP* asymmetries S_f and $S_{\bar{f}}$ due to the additional degrees of freedom in the fit. This is compensated by the fact that the systematic uncertainty for the portability of the calibration, which is often the largest systematic uncertainty for flavour tagged analyses at LHCb, is not needed.

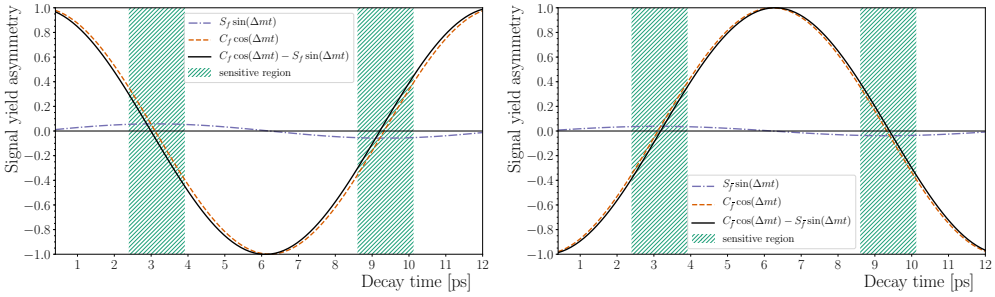

Fig. 9.1: Time-dependent asymmetries for $\overline{B}^0$ versus B^0 for the $D^-\pi^+$ (left) and $D^+\pi^-$ (right) final states. The values for the *CP* parameters are taken from simulation and no experimental dilutions are included. Since the deviation of the sum of both trigonometric functions from a single cosine function is very small due to the tiny values of S_f and $S_{\bar{f}}$, the maximal sensitivity is achieved in the regions of the zero-points, *i.e.* where the amplitude of $\cos(\Delta mt)$ and $\sin(\Delta mt)$ are of similar size. These regions are denoted as "sensitive regions" in both plots. In the "non-sensitive regions", where the cosine-function is completely dominating, additional diluting effects due to the tagging calibration can be determined.

A quantitative validation of this strategy was done by a collaborator on pseudoexperiments: Generating pseudoexperiments and floating the parameters C_f and $C_{\bar{f}}$ in the fit showed a non-negligible bias for the parameters S_f and $S_{\bar{f}}$, while fixing the coefficients C_f and $C_{\bar{f}}$ in the fit showed unbiased results on all fitted parameters. Therefore this strategy is adopted for the extraction of the calibration parameters. Finally it should be noted here that possible deviations from the assumption of $C_f = -C_{\bar{f}} = 1$ are taken into account in the systematic uncertainties in Ch. 11.

For these reasons, the studies presented in the following sections do not aim to obtain the calibration parameters. Instead, on the one hand, they are used to determine the functional form of the calibration function, which is used in the decay-time fit described in Sec. 10.1.3. On the other hand, the obtained parameter values of the calibration function can be used as reference values for the decay-time fit, as even if they are not expected to agree perfectly, the parameter sets obtained on the respective control sample and on the signal decay should be in similar.

9.2 Same side tagging calibration

The SS tagger combination is retrained and calibrated on the data sample of $B^0 \to J/\psi K^{*0}$ candidates corresponding to $3\,\text{fb}^{-1}$, recorded in 2011 and 2012 at centre-of-mass energies of 7 TeV and 8 TeV, respectively. It is used because $B^0 \to J/\psi K^{*0}$ is the only flavour-specific decay mode of B^0 mesons with a sufficient large statistic beside the signal mode of $B^0 \to D^{\mp}\pi^{\pm}$ at LHCb. The strategy follows mainly the one which was used to train the original taggers on $B^0 \to D^{\mp}\pi^{\pm}$, described in Ref. [78]. Only one additional step is

implemented: The distributions of $B^0 \to J/\psi K^{*0}$ candidates are kinematically weighted to match the distributions of $B^0 \to D^{\mp}\pi^{\pm}$ candidates. The general procedure is similar for both tagging algorithms and follows these steps:

- The $B^0 \to J/\psi K^{*0}$ sample is selected and a fit to the invariant mass is performed in order to separate the signal and background components statistically. Using the *sPlot* [70] method weights are calculated, which allow to obtain the signal-component distributions of other variables. This is reported in Sec. 9.2.1

- The distributions of transverse momentum p_{T}, pseudo-rapidity η' and azimuthal angle ϕ of the B^0 candidate, the number of tracks and the number of PVs in the event and a distribution of HLT2 trigger decisions for $B^0 \to J/\psi K^{*0}$ candidates are weighted to match those of the $B^0 \to D^{\mp}\pi^{\pm}$ candidates. This has two reasons: first, the performance of the BDT is potentially improved when applying it to $B^0 \to D^{\mp}\pi^{\pm}$, second, the portability of the functional form of the calibration is assured, and the obtained calibration parameters can be better compared and used as reference values.

- The tagging particles, whose charge is correlated with the initial B flavour are selected

- Finally, the data sample is divided into three sub samples:
  - The first third of the sample is used to train a BDT to separate correctly and wrongly charged tagging particles. Hereby correctly tagged means that the charge of the tagging particle indicates the correct tag, while wrongly tagged denotes exactly the opposite situation: The tagging particles' charge is indicating the wrong tag decision. To reduce the number of oscillated B^0 mesons, only B^0 mesons with a decay-time smaller than the first maximum of the oscillation are used. Therefore, following the procedure in Ref. [78], B^0 mesons used in the training are required to have a a decay-time smaller than 2.2 ps.
  - On the second sub sample the BDT is tested in order to avoid effects as over-training. Furthermore, both, the first training sample and this second testing sample are used to determine the mistag as a function of the BDT response. For this latter step the requirement on the B^0 decay-time is removed, as a full time-dependent analysis of the B^0 oscillation is performed.
  - On the last third of the data sample the SS taggers are combined and the functional form of the calibration is determined together with reference values for the calibration parameters.

Tab. 9.1: Stripping cuts for the decay $B^0 \to J/\psi K^{*0}$ with $J/\psi \to \mu^+\mu^-$ and $K^{*0} \to K^+\pi^-$. DOCA denotes the distance of closest approach of the respective K^{*0} or J/ψ mesons and $m_{p^+p^-}$ denotes the invariant mass of the $\mu^+\mu^-$ or $K^+\pi^-$ combination.

μ requirements		
transverse momentum p_T	$> 500\,\text{MeV/c}$	
$\text{DLL}_{\mu\pi}$	> 0.0	
resonances-requirements	J/ψ	K^{*0}
transverse momentum p_T	-	$> 500\,\text{MeV/c}$
DOCA χ^2	< 20.0	< 30.0
$\left\lvert m_{p^+p^-} - m^{\text{PDG}}_{J/\psi}\right\rvert$	$150\,\text{MeV/c}^2$	$300\,\text{MeV/c}^2$
decay vertex χ^2	< 16.0	< 25.0
B^0 meson requirements		
reconstructed decay-time t	$> 0.2\,\text{ps}$	

Tab. 9.2: Cuts to select $B^0 \to J/\psi K^{*0}$ candidates. The transverse momentum is denoted as p_T, for the impact parameter the shortcut IP is used and h denotes a kaon or pion.

$J/\psi/\mu$ requirements	$K^{*0}/(K, \pi)$ requirements
$\text{DLL}_{\mu\pi} > 0$	$\text{DLL}_{K\pi} > 0$ for kaons
track $\chi^2/\text{ndof} < 4$	track $\chi^2/\text{ndof} < 4$
$p_T(\mu^+) > 500\,\text{MeV/c} \,\vert\vert\, p_T(\mu^-) > 500\,\text{MeV/c}$	$p_T(K^{*0}) > 1000\,\text{MeV/c}$
vertex χ^2/ndof< 16 of J/ψ	vertex χ^2/ndof< 16 of K^{*0}
$\left\lvert m_{J/\psi} - m_{\text{PDG}}\right\rvert < 80\,\text{MeV/c}^2$	$\left\lvert m_{K^{*0}} - m^{\text{PDG}}\right\rvert < 70\,\text{MeV/c}^2$

9.2.1 Preparation of the $B^0 \to J/\psi K^{*0}$ samples

The decay $B^0 \to J/\psi K^{*0}$ is reconstructed using the muonic decay $J/\psi \to \mu^+\mu^-$ and the domninant $K^{*0} \to K^+\pi^-$ decay. The cuts applied in the stripping are listed in Tab. 9.1. Again, to achieve better compatibility between $B^0 \to J/\psi K^{*0}$ and the signal decay $B^0 \to D^{\mp}\pi^{\pm}$, the same trigger requirements as for the signal mode are applied, although the signal efficiency is only about 66 %. After the trigger requirements, the K^{*0} and the J/ψ meson are requried to fulfill the cuts shown in Tab. 9.2. Additionally only B^0 candidates with an invariant mass within the range $[5100, 5450]\,\text{MeV/c}^2$, a transverse momentum larger than 1 GeV/c, the minimal χ^2 of the impact parameter w.r.t. the PV smaller than 25 and χ^2/ndof of the decay vertex smaller than ten are selected. Last the χ^2/ndof of the kinematic decay chain fit is required to be smaller than five and the number of PVs in the event has to be one or in case an event has more than one PV the impact parameter χ^2 has to be

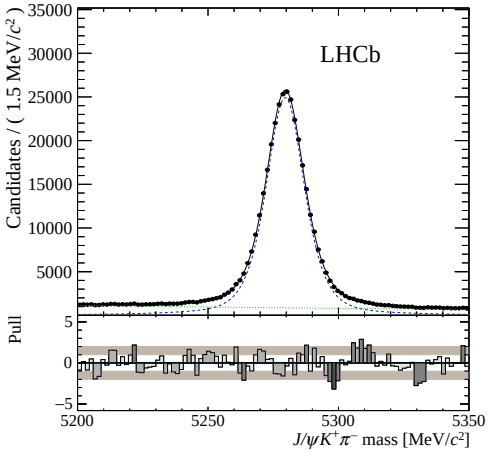

Fig. 9.2: Fit to the invariant mass of the $B^0 \to J/\psi K^{*0}$ candidates with a constraint on the known J/ψ mass. The data is represented by the black points, the black solid line shows the full PDF, the blue dashed line shows the signal componente and the background component is represented by the dotted yellow line.

Tab. 9.3: Fitted Yields of the $B^0 \to J/\psi K^{*0}$ component and the combinatorial background.

Parameter	Yield
$N^{\pi}_{B^0 \to D\pi}$	$351\,507 \pm 662$
N^{π}_{bkg}	$88\,630 \pm 419$

larger than 50. In case of multiple B^0 candidates the one with the smallest χ^2/ndof of the kinematic decay chain fit is chosen.

To extract *sWeights* [70] a fit to the invariant B^0 mass, stemming from the kinematic decay chain fit with a constraint on the known J/ψ mass is performed. The signal model consists of a sum of two double-sided Hypatia functions with a shared mean value, the background is parametrised using an exponential function. The first Hypatia function describes mainly $B^0 \to J/\psi K^{*0}$ candidates without photon radiation in the final state, while the second Hypatia shows a large tail towards smaller invariant masses. The tail parameters a_i and n_i of both Hypatia functions are determined on simulated events. Moreover the fraction between both signal components and the parameter λ of the Hypatia describing the B^0 candidates with harder photon radiation are taken from fits to simulated samples. All remaining parameters are floating. The fit that is performed in the range $[5200, 5350]\,\text{MeV/c}^2$ is shown in Fig. 9.2 and the resulting yields are given in Tab. 9.3.

The last preparation step for the data sample is to weight certain *sWeighted* distributions of $B^0 \to J/\psi K^{*0}$ candidates to match the corresponding distributions of the $B^0 \to D^{\mp}\pi^{\pm}$

candidates. The observables which are weighted are chosen because they are known to influence the tagging responses most: These observables are the transverse momentum p_{T}, the pseudo-rapidity η' and the azimuthal angle ϕ of the B^0 candidate as well as the number of tracks and the number of PVs in the event. Additionally, an observable containing the composition of the used trigger lines is weighted. For this observables the categories 1, 2 and 3 contain candidates which are exclusively triggered TOS by one of the `Hlt2Topo` lines. Category 4 contains candidates triggered TOS exclusively by the overlap of two- and three-body `Hlt2Topo` lines, candidates triggered exclusively TOS by the overlap of the two- and four-body `Hlt2Topo` lines are classified in category 5 and candidates triggered exclusively TOS by the overlap of the three- and four-body `Hlt2Topo` lines are in category 6. The last category (7) contains candidates which are triggered by all three used HLT2 lines.

Instead of weighting the six-dimensional parameter space using a six-dimensional histogram a BDT based approach is chosen. In this approach, the parameter space is split into several large regions, which are separated using binary decision trees. These trees maximise a symmetrised χ^2

$$\chi^2 = \sum_i \frac{\left(N_{i,\mathrm{source}} - N_{i,\mathrm{target}}\right)^2}{N_{i,\mathrm{source}} + N_{i,\mathrm{target}}} \tag{9.4}$$

where $N_{i,\mathrm{source}}$ are the number of candidates in the ith bin of the distribution which should be weighted and $N_{i,\mathrm{target}}$ are the number of candidates in the i-th bin of the target distribution. For all leaves of a decision tree a prediction p is calculated following

$$p = \log \frac{N_{i,\mathrm{target}}}{N_{i,\mathrm{source}}}\,, \tag{9.5}$$

what is used to calculate the weights as

$$w = \begin{cases} w, & \text{if event from target distribution} \\ we^p, & \text{if event from source distribution}\,. \end{cases} \tag{9.6}$$

These steps are then repeated for each decision tree. In this case the BDT is built out of 40 trees with a maximum depth of three. Each leave has to contain a minimum number of 200 candidates and the boosting method is Gradient boosting [92]. The distributions of the original and weighted $B^0 \to J/\psi K^{*0}$ and $B^0 \to D^{\mp}\pi^{\pm}$ candidates is shown in Fig. 9.3.

9.2.2 Retraining of the SS pion tagger

For the retraining of the SS pion tagger the tagging particles first have to fulfill a set of selection cuts, which are partly applied only to the tagging track and partly to the system of the tagging track and the B candidate. These cuts have two aims: First separating pions form other charged tracks in the event and second selecting particles which are in the same

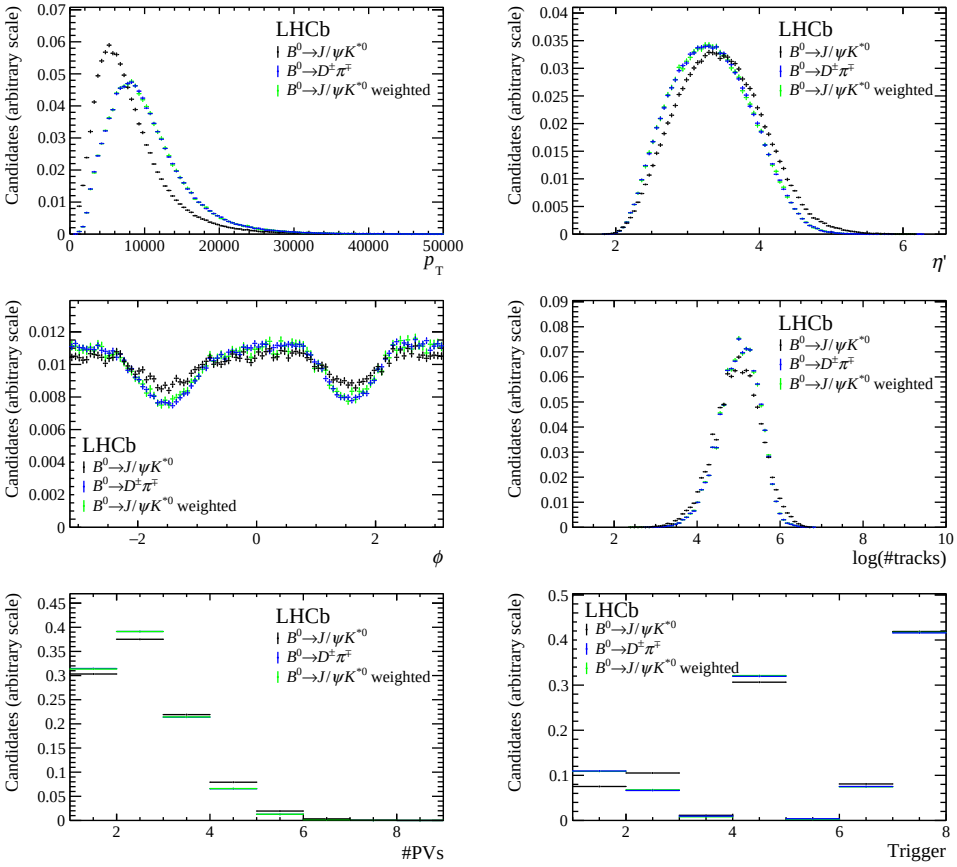

Fig. 9.3: Left, top to bottom: Normalised *sWeighted* distributions of the transverse momentum p_T, azimuthal angle ϕ and number of tracks per event. Right, top to bottom: Normalised *sWeighted* distributions of the pseudo-rapidit η', logarithm of number of tracks per event and distribution of trigger decisions.

Tab. 9.4: Selection cuts for the SS pion algorithm. The first set of cuts is applied only top the tagging-particle track, while the second set is applied to the "tagging-particle track+B" system.

Variable	Cut
transverse momentum p_{T}	$> 0.4\,\mathrm{GeV/c}$
$\mathrm{IP}/\sigma_{\mathrm{IP}}$	< 4
track ghost probability g_{prob}	< 0.5
$\mathrm{DLL}_{p\pi}$	< 5
$\mathrm{DLL}_{K\pi}$	< 5
track χ^2/ndof	< 5
$\mathrm{IP}_{\mathrm{PU}}/\sigma_{\mathrm{IP}_{\mathrm{PU}}}$	> 3
transverse momentum p_{T}	$> 3\,\mathrm{GeV/c}$
$\Delta Q = m_{B+\pi} - m_B - m_\pi$	$< 1.2\,\mathrm{GeV/c^2}$
$\Delta\eta'$	< 1.2
$\Delta\phi$	< 1.1
vertex χ^2	< 100

region of phase space as the B candidate. For each tagging-particle candidate passing the requirements shown in Tab. 9.4, the B meson gets a preliminary tag decision assigned, corresponding to the charge of the tagging particle ($\pi^+ \to d = +1$, $\pi^- \to d = -1$). The quantities $\Delta\eta'$ and $\Delta\phi$ describe the difference of pseudo-rapidity and azimuthal angle between the signal and tagging-particle track. To further assure that the tagging-particle is also produced at the PV with which the B candidate has the smallest χ^2_{IP}, the impact parameter significance with the next "best" PV in the event (denoted as pile-up vertex) is used. The corresponding quantities are $\mathrm{IP}/\sigma_{\mathrm{IP}}$ and $\mathrm{IP}_{\mathrm{PU}}/\sigma_{\mathrm{IP}_{\mathrm{PU}}}$, respectively.

Subsequently, a BDT is trained to determine the mistag and to select the best tagging-particle candidate in events with more than one tagging-particle candidate passing the selection in Tab. 9.4. In the training, all candidates with the correct correlation between the preliminary tag and the B flavour are considered as signal, all with the wrong correlation are considered to be background. A correct or wrong correlation for the SS pion algorithm means

$$\text{Correct tag: } B^0\pi^+ \text{ or } \overline{B}^0\pi^-\,,$$
$$\text{Wrong tag: } B^0\pi^- \text{ or } \overline{B}^0\pi^+\,,$$

where only B candidates with a decay-time smaller than 2.2 ps are used in the BDT training to reduce the number of oscillated B candidates. The BDT further consists of 3000 trees with a maximum depth of two. Each tree uses up to five of the input variables scanned at 30 points to find the optimal cut point. As boost algorithm AdaBoost [72] is chosen

Tab. 9.5: List of input variables used in the training of the BDT for the SS pion tagger. The quantity ProbNNK is the probability for a particle to be a kaon, calculated by an artificial neural network using information from the LHCb PID system. The momentum of a particle is denoted as p.

tagging-particle inputs	$\log\left(p_{\mathrm{T}}\right)$
	$\log\left(p\right)$
	$\log\left(\mathrm{IP}/\sigma_{\mathrm{IP}}\right)$
	$\log\left(g_{\mathrm{prob}}\right)$
	$\log\left(\mathrm{track}\chi^2/\mathrm{ndof}\right)$
	$\log\left(\mathrm{ProbNNK}\right)$
tagging-particle + B^0 inputs	ΔQ
	ΔR
	$\log\left(\Delta\eta'\right)$
	$\log\left(\Delta\phi\right)$
	$\log\left(p_{\mathrm{T}}\right)$
event properties	PVndof

with a boost factor of $\beta = 0.005$. A list of the input variables is given in Tab. 9.5. Most of these variables were already used in the selection, PVndof is the number of degrees of freedom in the fit to determine the position of the PV and $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta'^2}$. Moreover a logarithmic transformation is applied to some of the input variables to improve the performance of the BDT. As one can see in Fig. 9.4 the distributions of the signal and background samples are quite similar. This small, but significant difference is propagated to the distribution of the BDT response as shown in Fig. 9.5. For events with more than one tagging particle, the best one is identified as the particle with the largest BDT output.

Due to the flavour oscillations, a time-dependent analysis is necessary to transform the BDT response into a mistag. For this, an unbinded maximum-likelhihood fit to the decay-time t, the tag decisions d and the mixing state ξ is performed. The mixing state takes values of $+1$ when the B-meson flavour at decay is the same as the tag decision and -1 otherwise. As the fit is performed to the *sWeighted* decay-time distribution the describing PDF consists only of a signal-component

$$\mathscr{M}(t; d, \xi) = \mathscr{N} a(t) e^{-t'/\tau} \left((1 - d\Delta\omega) + \xi(1 - 2\omega)\cos\left(\Delta m t'\right)\right) \otimes \mathscr{R}(t - t'), \qquad (9.7)$$

where $\mathscr{N}$ is a normalisation factor and ω and $\Delta\omega$ are the mistag and mistag difference for initial B^0- and $\overline{B}^0$ mesons, respectively. $\mathscr{R}(t-t')$ describes an average decay-time resolution of 50 fs. The lifetime is constrained by means of a Gaussian function to be $\tau = 1.518 \pm 0.004$ what allows to float the parameters of the cubic-splines describing the decay-time acceptance $a(t)$. These cubic-splines have five knots placed at $[0.2, 0.25, 2.0, 13.8, 14.5]$ ps (more details in Sec. 10.1.2). At first, the fit is performed in bins of the BDT response to

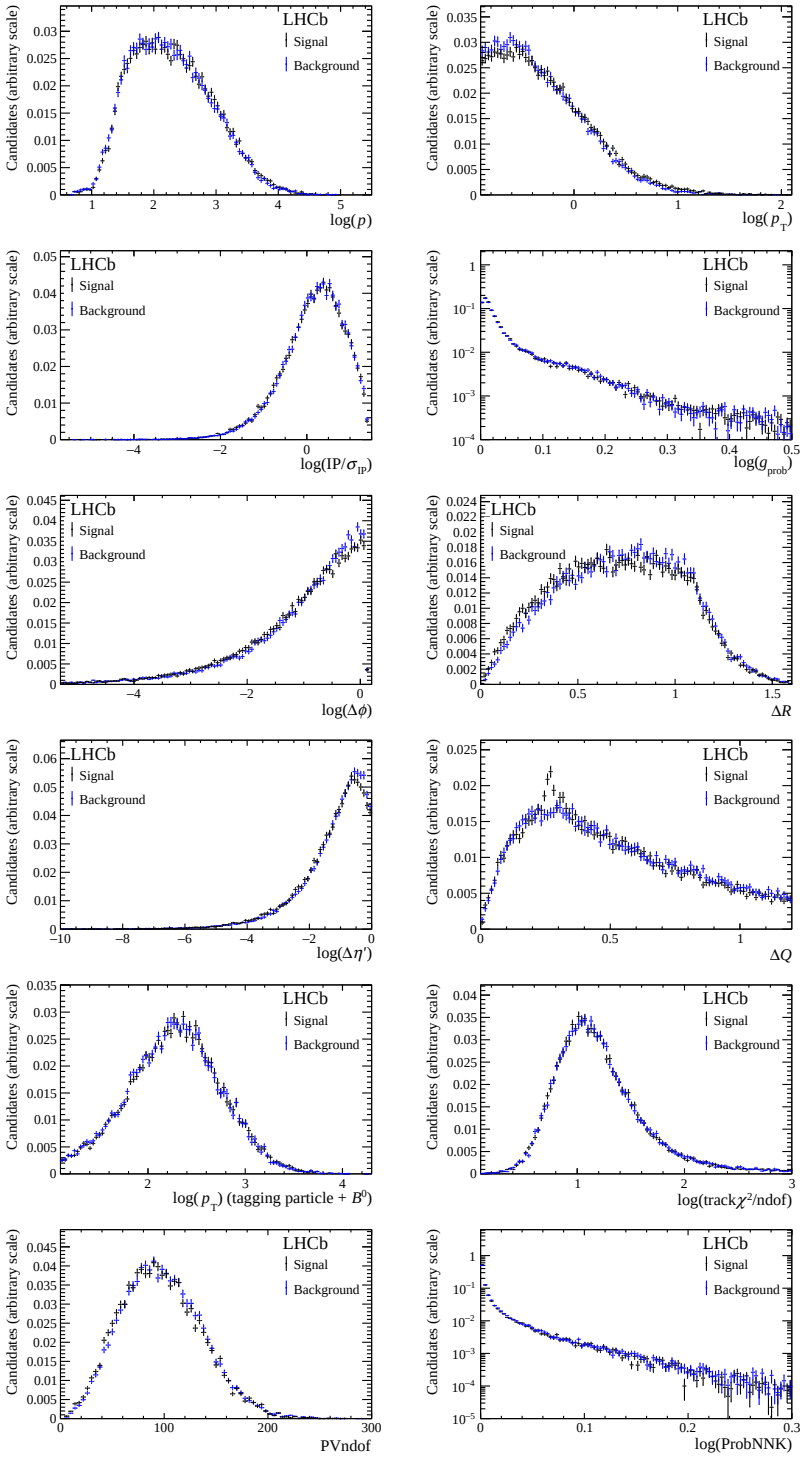

Fig. 9.4: Distributions of the input variables used in the BDT training. The black points represent the right charge correlation while the blue points correspond to the wrong ones.

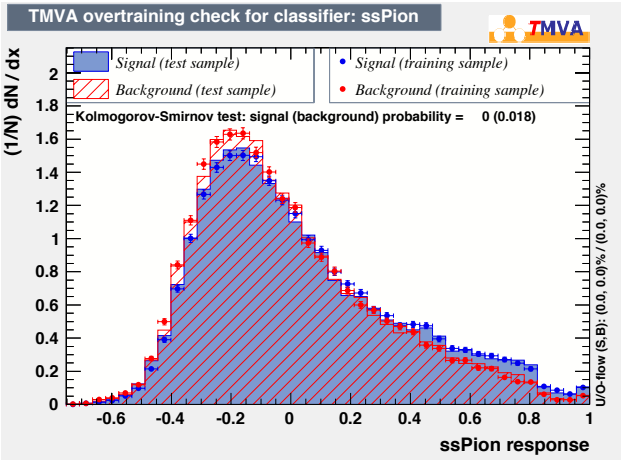

Fig. 9.5: Output of the SS pion BDT ($\alpha^{\mathrm{SS}\pi}_{\mathrm{BDT}}$). The blue distribution represents the right charge correlated tracks while the red distribution corresponds to the wrong charge correlated tracks.

Tab. 9.6: Final transformation parameters for the BDT responst of the SS pion tagger.

b_0	b_1	b_2	b_3
0.469 ± 0.003	-0.066 ± 0.012	-0.022 ± 0.048	-0.026 ± 0.056

determine pairs of the average BDT response and mistag. These pairs are then used to determine the functional form of the transformation. A 3$^{\mathrm{rd}}$ order polynomial of the form

$$\omega\left(\alpha^{\mathrm{SS}\pi}_{\mathrm{BDT}}\right) = b_0 + b_1 \times \alpha^{\mathrm{SS}\pi}_{\mathrm{BDT}} + b_2 \times \alpha^{\mathrm{SS}\pi\,2}_{\mathrm{BDT}} + b_3 \times \alpha^{\mathrm{SS}\pi\,3}_{\mathrm{BDT}} \tag{9.8}$$

is found to describe the relation well. However, to determine the final transformation parameters, an unbinned fit, where the mistag is parametrised directly in the PDF, is performed. Figure (9.6) shows the projected mixing asymmetry from the decay-time fit, defined as

$$A(t) = \frac{N_{\mathrm{unmix}}(t) - N_{\mathrm{mix}}(t)}{N_{\mathrm{unmix}}(t) + N_{\mathrm{mix}}(t)} \propto (1 - 2\omega)\cos(\Delta m t) \tag{9.9}$$

where $N_{\mathrm{unmix}}(t)$ and $N_{\mathrm{mix}}(t)$ are the number of unmixed and mixed B^0 candidates for a decay-time t, respectively. The resulting transformation parameters are given in Tab. 9.6, while Fig. 9.7 shows the corresponding 3$^{\mathrm{rd}}$ order polynomial. Using this transformation, for each B^0 candidate the final mistag η is calculated. In case the computed mistag η' is larger than 0.5, the tag decision is inverted and the final mistag is calculated as $\eta = 1 - \eta'$.

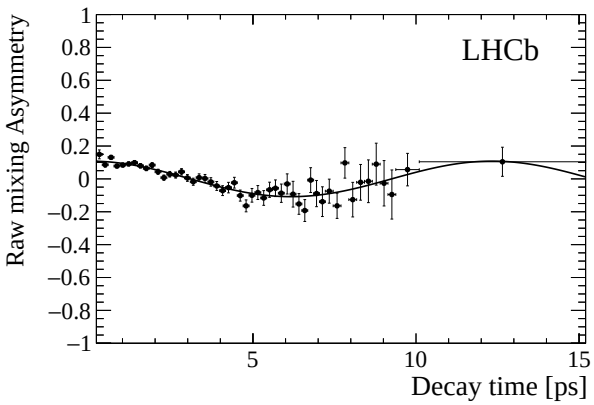

Fig. 9.6: Mixing asymmetry for $B^0 \to J/\psi K^{*0}$ candidates tagged by the SS pion tagger. The curve overlaid is the fit projection of the decay-time fit.

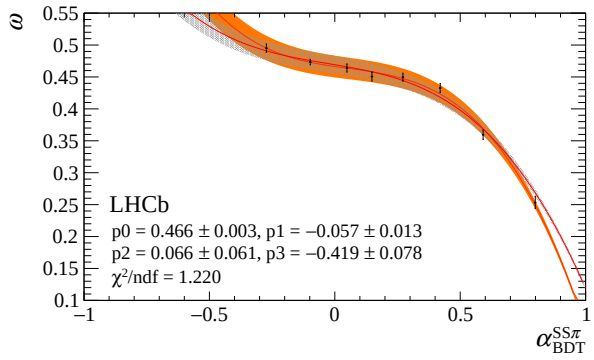

Fig. 9.7: Polynomial curve of ω versus $\alpha^{SS\pi}_{BDT}$ on the first two thirds of the $B^0 \to J/\psi K^{*0}$ data sample. The fit from the binned approach is shown in orange, the curve of the nominal unbinned method is shown in grey . The parameter values in the plot correspond to the orange curve/binned approach.

9.2.3 Retraining of the SS proton tagger

The procedure for tuning the SS proton tagger BDT and transforming the BDT response is very similar to the one described in Sec. 9.2.2 for the SS pion tagger.

Besides pions, protons, which have a correlation with the initial flavour of the B meson, may also be formed in the hadronisation process. These tagging-proton candidates must first meet a set of requirements for the proton track itself and the system of the tagging-proton candidate and the B meson. These cuts are listed in Tab. 9.7. The correlation between the charge of the proton and the B meson flavour is as follows for the SS proton tagger:

$$\text{Correct tag: } B^0\overline{p} \text{ or } \overline{B}^0 p\,,$$
$$\text{Wrong tag: } B^0 p \text{ or } \overline{B}^0\overline{p}\,.$$

Based on the charge of the selected particles, a preliminary tag is assigned to the B candidates ($p \to d = -1$, $\overline{p} \to d = +1$).

Here it is important to note that the correlation is the exact opposite to the one of the SS pion tagger. As a consequence, a misidentification of pions and protons would lead to a degraded performance of both taggers and the requirement on $\mathrm{DLL}_{p\pi}$ is chosen exactly the opposite for both algorithms. Furthermore, this requirement reduces the correlation between both taggers as the sets of selected tagging particles are completely disjoint.

In the same way as for the SS pion tagger, proton candidates with the correct charge correlation with the flavor of the initial B meson are assumed to be signal while proton candidates with the incorrect charge correlation are considered as background for the BDT training. The input variables for the BDT are listed in Tab. 9.8, most of them again already used for the selection. As in Sec. 9.2.2, the distributions of the input variables are very similar for the data sets considered as signal and background. Again, only B candidates with a decay-time below 2.2 ps are used for the BDT training and some of the input variables are logarithmically transformed to improve the performance. All further configurations of the BDT are same as described in Sec. 9.2.2, in Fig. 9.8 the BDT response is shown. The separation power of the BDT is again small but significant. In addition, the test and training samples show some small differences, which probably are a result of a

Tab. 9.7: Selection cuts for the SS proton algorithm. The first set of cuts is applied only on the tagging-particle track, while the second set is applied to the "tagging-particle track+B" system. The quantity ΔQ is defined as $\Delta Q = m_{B+p} - m_B - m_p$.

Variable	p_T	$\mathrm{IP}/\sigma_\mathrm{IP}$	g_prob	$\mathrm{DLL}_{p\pi}$	$\mathrm{IP}_\mathrm{PU}/\sigma_{\mathrm{IP}_\mathrm{PU}}$
Cut	$> 0.4\,\mathrm{GeV/c}$	< 4	< 0.5	> 5	> 3
Variable	p_T	ΔQ	$\Delta\eta'$	$\Delta\phi$	vertex χ^2
Cut	$> 3\,\mathrm{GeV/c}$	$< 1.2\,\mathrm{GeV/c^2}$	< 1.2	< 1.1	< 100

Tab. 9.8: List of input variables used in the training of the BDT for the SS proton tagger. The momentum of a particle is denoted as p.

tagging particle inputs	$\log\left(p_{\mathrm{T}}\right)$ $\log\left(p\right)$ $\log\left(\mathrm{IP}/\sigma_{\mathrm{IP}}\right)$ $\log\left(\mathrm{DLL}_{p\pi}\right)$
tagging particle + B^0 inputs	ΔQ ΔR $\log\left(\Delta\eta'\right)$
event properties	PVndof

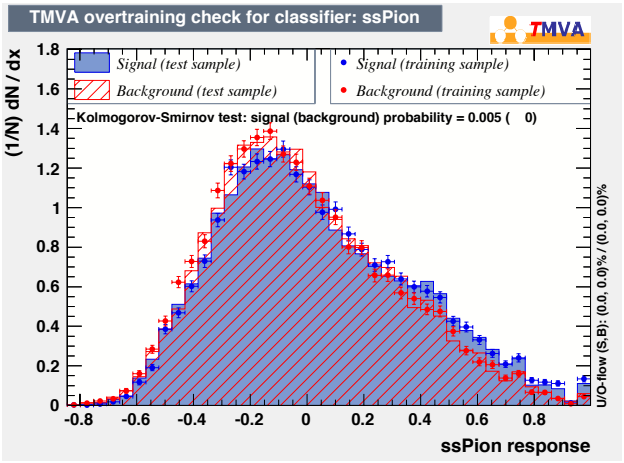

Fig. 9.8: Output of the SS proton BDT ($\alpha^{\mathrm{SS}p}_{\mathrm{BDT}}$). The blue distribution represents the right charge correlated tracks while the red distribution corresponds to the wrong charge correlated tracks.

Tab. 9.9: Final transformation parameters for the BDT response of the SS proton tagger.

b_0	b_1	b_2	b_3
0.479 ± 0.004	-0.097 ± 0.016	0.009 ± 0.043	-0.215 ± 0.064

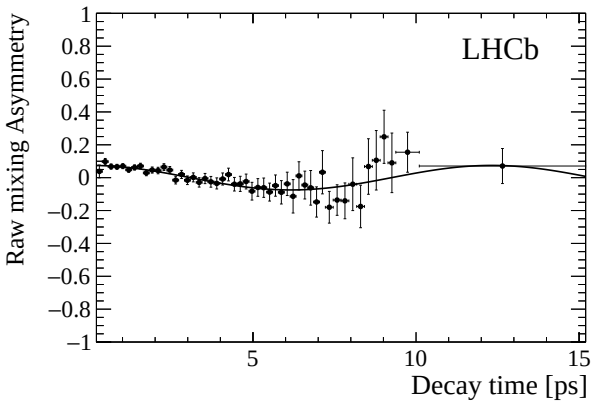

Fig. 9.9: Mixing asymmetry for $B^0 \to J/\psi K^{*0}$ candidates tagged by the SS proton tagger. The curve overlaid is the fit projection of the decay-time fit.

slight overtraining. However, possible effects arising due to this effect are corrected when transforming and calibrating the BDT response. As for the SS pion algorithm, in case of multiple tagging-proton candidates per event the tag decision and the mistag are defined by the candidate with the largest BDT output.

To determine the relation between the mistag and the BDT output, a time-dependent mixing analysis is performed. This means, an unbinned maximum-likelihood fit to the distribution of the decay-time t, the tag decision d and the mixing state ξ is done. Since the *sWeigthed* distributions are used for the fit again, only the signal component must be described with the PDF from Eq. (9.7). Firstly the functional relation between the mistag and $\alpha^{\mathrm{SS}p}_{\mathrm{BDT}}$ is determined with a binned fit in the BDT output $\alpha^{\mathrm{SS}p}_{\mathrm{BDT}}$. A 3$^{\mathrm{rd}}$ order polynomial as described in Eq. (9.8) describes the data well. An unbinned fit, in which the mistag is parametrised directly in the PDF, finally provides the final transformation parameters, which can be found in Tab. 9.9. Figure 9.9 shows the time-dependent mixing asymmetry for the SS proton tagger, the transformation functions of the binned and unbinned approach are presented in Fig. 9.10.

Subsequently, the mistag η is determined using the parameterisation from Tab. 9.9 If the calculated mistag η' is greater than 0.5, the tag decision is inverted and the mistag is calculated as $\eta = 1 - \eta'$.

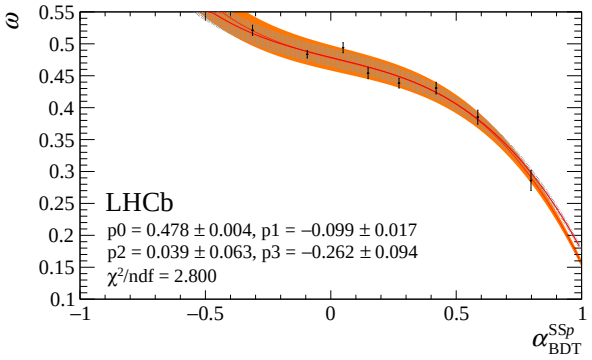

Fig. 9.10: Polynomial curve of ω versus $\alpha_{\mathrm{BDT}}^{\mathrm{SS}p}$ on the first two thirds of the $B^0 \to J/\psi K^{*0}$ data sample. The fit from the binned approach is shown in orange, the curved of the nominal unbinned method is shown in grey. The parameter values in the plot correspond to the orange curve/binned approach.

Tab. 9.10: GOF metrics for two different calibration models for the SS taggers.

GOF metric	Score (4 parameters)	Score (6 parameters)
χ^2	3.3214	−2.0864
G^2	2.5524	−2.6624
CR	3.1819	−2.7763
S	1.5426	−2.3908

9.2.4 Calibration of the SS tagger combination

After training the taggers of the same side in Sec. 9.2.2 and Sec. 9.2.3 they need to be combined. This is achieved using the formalism from Sec. 6.4.3, yielding one global SS tagger with a tag decision d^{SS} and a mistag η^{SS}.

This global SS combination then must be calibrated. In order to do this, the adopted model is a GLM with a 1$^{\mathrm{st}}$ order polynomial as *basis function* and the modified logistic function from Eq. (6.31) as *link function*. The number of free parameters in the model is 4 and is selected to achieve several satisfactory goodness-of-fit (GOF) metrics. The comparisons of GOF metrics are shown in Table Tab. 9.10. All presented metrix are expected to be normally distributed, *i.e.* to follow Gaussian distributions with a mean of zero and a width of one. This means, that the given values, can be interpreted directly as the deviation in terms of standard deviations from the perfectly calibrated case. Since the deviance G^2 and the le Cessie-van Houwelingen-Copas-Hosmer metric (S) seem to prefer the more simple model, while the Pearson χ^2 and the Cressie-Read (CR) metric prefer the more complicated model, both are assumed to fit the data equally well and the more simple one is chosen (all metrics are described in Ref. [93]). The Calibration is then performed

Tab. 9.11: SS calibration parameters obtained on the last third of the $B^0 \to J/\psi K^{*0}$ data sample.

p_0	p_1	Δp_0	Δp_1
-0.091 ± 0.059	-0.027 ± 0.065	0.034 ± 0.084	0.032 ± 0.094

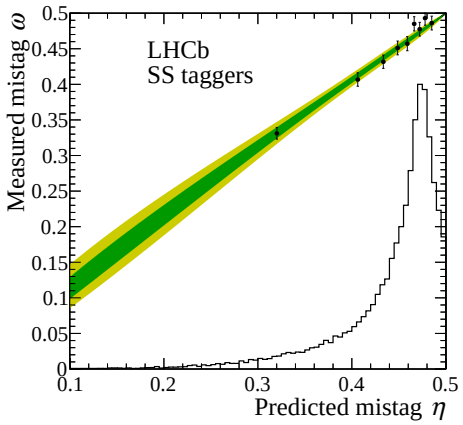

Fig. 9.11: Calibration retreived for the SS tagger combination. The black histogram is the distributions of the mistag probabilities in arbitrary units. The green areas correspond to the 68 % and 95 % confidence level regions of the calibration functions.

using the *sWeights* [70] together with the weights from Sec. 9.2.1, giving the calibration parameters shown in Tab. 9.11. A graphical representation can be found in Fig. 9.11.

The portability of the SS tagger combination is tested on simulated $B^0 \to D^{\mp}\pi^{\pm}$ and $B^0 \to J/\psi K^{*0}$ candidates. This is done using the simulated truth information such that the calibration can be performed in the same way as on a charged decay mode, after equalising the number of initial B^0- and $\overline{B}^0$ mesons to separate tagging asymmetries from asymmetries in the prodcution of B mesons or to CP violation. Both, the individual calibration parameters and a "full" comparison from a χ^2 test including the correlations between the parameters, show good agreement (*e.g.* 0.09σ for the full test). Despite this good agreement, the parameters in the CP-fit on $B^0 \to D^{\mp}\pi^{\pm}$ are left free. On the one hand, this is motivated by the higher accuracy of the calibration parameters on the $B^0 \to D^{\mp}\pi^{\pm}$ data sample. On the other hand this strategy matches the approach for the OS tagger combination, where the portability is not given (Sec. 9.3). In addition, this approach does not require any systematic uncertainty accounting for the portability of the calibration parameters.

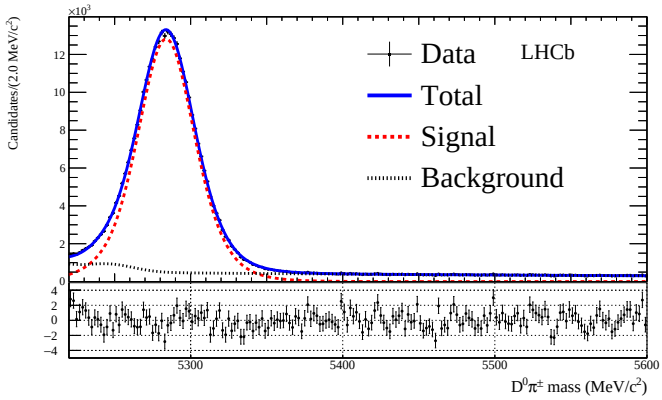

Fig. 9.12: Invariant mass distributions of the $D^0\pi^+$ mass. The fit projection of Fit B is overlaid.

Tab. 9.12: Fitted yields of the $B^+ \to D^0\pi^+$ control channel from Fit B.

Parameter	Yield
$N_{B^+\to D^0\pi^+}$	$319\,974 \pm 612$
N_{bkg}	$85\,687 \pm 377$

9.3 Opposite side tagging calibration

As mentioned before, the work in this section was done by a collaborator and therefore the explanations will be less detailed than in most other parts of this thesis.

As control channel for the OS tagger combination (namely the OS muon, OS electron, OS kaon, OS vertex charge and OS charm) the charged decay mode $B^+ \to D^0\pi^+$ is used. As this decay mode is kinematically similar to the signal decay $B^0 \to D^\mp\pi^\pm$, the same trigger requirements can be applied with a high signal efficiency. The $B^+ \to D^0\pi^+$ candidates are further selected by simple rectangular cuts on track quality variables, PID variables, the decay-time of the B candidate and the invariant mass of the D^0. Then a fit to the invariant B^+ mass is performed to extract *sWeights* using the same strategy as in Ch. 8. The invariant mass distribution overlaid with the fit projection from Fit B is shown in Fig. 9.12 while Tab. 9.12 shows the resulting signal and background yields. Furthermore, the $B^+ \to D^0\pi^+$ candidates are weighted in the distributions of transverse momentum, pseudo-rapidity and decay-time of the B candidate, the number of tracks and PVs in the event and a distribution of trigger decisions to match the distributions of the $B^0 \to D^\mp\pi^\pm$ candidates.

The calibration model is a GLM with a natural spline function as *basis function* and the modified logistic function from Eq. (6.31) as *link function*. The number of free parameters in the calibration model is 10 and is again chosen such that the goodness-of-fit metrics

Tab. 9.13: GOF metrics for two different calibration models for the OS taggers.

GOF metric	Score (8 parameters)	Score (10 parameters)
χ^2	4.1312	−2.197
G^2	−3.8699	0.7199
CR	2.9273	−1.6896
S	−4.2701	1.8470

Tab. 9.14: OS calibration parameters obtained on the $B^+ \to D^0\pi^+$ data sample.

p_0	p_1	p_2	p_3	p_4
-0.136 ± 0.019	-0.006 ± 0.022	-0.0107 ± 0.0083	-0.45 ± 0.10	-0.85 ± 0.46
Δp_0	Δp_1	Δp_2	Δp_3	Δp_4
-0.129 ± 0.038	0.042 ± 0.045	-0.020 ± 0.017	0.42 ± 0.21	1.91 ± 0.92

used in Sec. 9.2.4 provide sufficiently satisfactory results. The comparison with the next simpler model with 8 parameters shown in Tab. 9.13 indicates that a simpler model is not describing the data well. The Calibration is performed using the *sWeights* together with the weights calculated before such that the $B^+ \to D^0\pi^+$ distributions match the $B^0 \to D^\mp\pi^\pm$ ones, giving the calibration parameters shown in Tab. 9.14. A graphical representation can be found in Fig. 9.13.

The portability of the calibration is checked on simulated candidates. The calibration is determined using the simulated true initial flavours on both decay channels after equalising the number of initial B^0 (B^+) and $\overline{B}^0$ (B^-) mesons. The parameters p_0 and p_1 show deviations of more than 2.5σ, while the full comparison taking into account the correlations between the parameters shows a discrepancy of 2.0σ. Although one could presume this overall discrepancy is sufficiently small, a study on simulated events presented in Sec. 10.2 shows that applying the calibration determined on $B^+ \to D^0\pi^+$ causes a bias for the *CP*-parameters S_f and $S_{\bar{f}}$. Presumably, some residual discrepancies remain despite the weighting of the $B^+ \to D^0\pi^+$ candidates. For this reason, the portability is not considered to be given and the calibration is determined directly in the *CP* fit on $B^0 \to D^\mp\pi^\pm$.

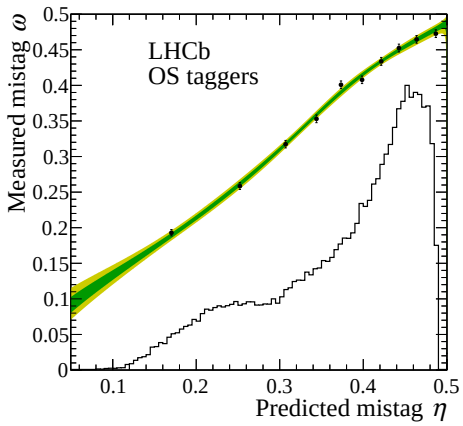

Fig. 9.13: Calibration retreived by the for the OS tagger combination. The black histogram is the distributions of the mistag probabilities in arbitrary units. The shaded areas correspond to the 68 % and 95 % confidence level regions of the calibration functions.

10 Decay-time fit

In this chapter the decay-time fit on $B^0 \to D^{\mp}\pi^{\pm}$ to extract the *CP* observables S_f and $S_{\bar{f}}$ is presented. Section 10.1.1 and 10.1.2 describe the parameterisation of the decay time resolution and acceptance before the extraction of the *CP* parameters is presented in Sec. 10.1.3. The validation of the fit is detailed in Sec. 10.2. After comparing the resulting values of nuisance parameters with reference values, the *link function* used for the calibration function of the OS and SS taggers is validated (Sec. 10.2.1). At last, the fit to extract the *CP* parameters is repeated on different sub samples of the data set (Sec. 10.2.2) and the entire strategy is also tested on simulated events (Sec. 10.2.3).

10.1 Fit to data

In the following, the decay-time fit to extract S_f and $S_{\bar{f}}$ and its components are presented. The fit is performed on the decay-time range from 0.4 to 12 ps. The lower limit was chosen in order to obtain a good description of the decay-time distribution at low decay-times without losing sensitivity to the parameters S_f and $S_{\bar{f}}$. The upper limit was set to a value such that the statistics is already too small so that an enlarged range would no longer add sensitivity to the *CP* parameters.

10.1.1 Decay time resolution

Since the work in this section was done by a collaborator, the contents are described only briefly.

The decay-time resolution is determined on a sample of *fake* B^0 candidates, formed from a prompt $D^{\pm}$ candidate and another track originating from the PV. These *fake* B^0 candidates are expected to have a decay time of zero and therefore the sample also is referred to as prompt sample. The candidates are selected with the same selection as presented in Sec. 7.2, except for the cut on the BDT output. Additionally, the *fake* B^0 candidates are required to have an χ^2_{IP} with the PV less than nine and the number of PVs in the event must be one to exclude wrong PV associations. Subsequently, *sWeights* [70] are determined by a fit to the invariant mass of the $D^{\pm}$ meson in order to examine only signal distributions in the following. Since the time resolution depends on the transverse momentum of the bachelor particle, this needs to be corrected in the sample of *fake* B^0 candidates. Therefore, the prompt sample is weighted by the ratio of the distributions of the logarithmic transverse momenta of the bachelor candidate in the signal $B^0 \to D^{\mp}\pi^{\pm}$ and the prompt sample.

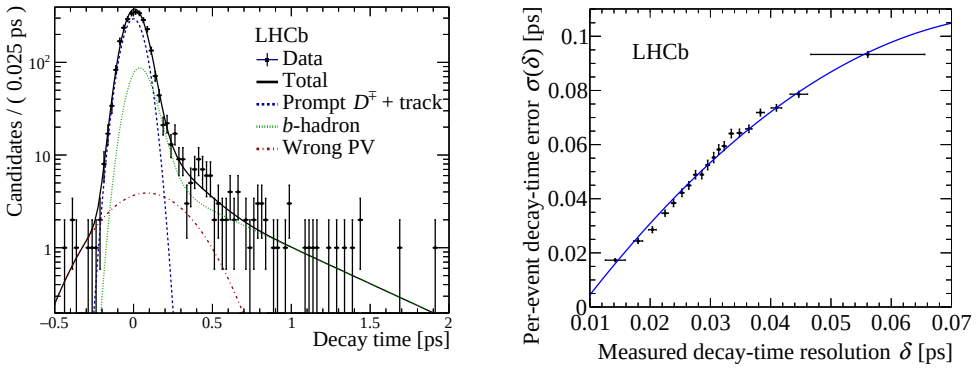

Fig. 10.1: Distribution of the decay-time resolution for one representative bin in per-candidate decay-time error for *fake* B^0 candidates (left) and measured resolution versus average per-candidate decay-time error, determined from fits to the decay time in bins of decay-time error (right).

To resolve the decay-time resolution, fits are then performed to the decay-time distribution of the prompt sample in 20 bins of the decay-time error. Since this sample does not contain real B^0 candidates, the decay-time resolution can be derived from the width of the decay-time distribution. The binning is chosen such that the sum of *sWeights* in each bin is equal. The fit model consists of three components: a delta function convolved with a Gaussian function to describe true prompt $D^{\pm}$ +track candidates, a pair of exponential functions convolved with the same Gaussian function to describe candidates from b hadrons and a wide Gaussian function to describe backgrounds due to wrongly associated PVs. The fit is shown for one representative bin in Fig. 10.1. A measured resolution $\langle\sigma\rangle_i$ per bin is obtained from this fit, which can be related to the corresponding average decay-time error $\langle\delta\rangle_i$. Following, a χ^2 fit to the $(\langle\delta\rangle_i, \langle\sigma\rangle_i)$ pairs of the form

$$\langle\sigma\rangle_i = \langle\sigma\rangle + p_1 \times \left(\langle\delta\rangle_i - \langle\delta\rangle\right) + p_2 \times \left(\langle\delta\rangle_i - \langle\delta\rangle\right)^2 \tag{10.1}$$

is performed, where $\langle\delta\rangle$ is the average per-event decay-time error of the whole unbinned sample. This χ^2 fit is shown in Fig. 10.1. It provides an average decay time resolution $\langle\sigma\rangle$ and a trend, from which a global average resolution of $\sigma(\langle\delta\rangle) = (54.91 \pm 0.38)$ fs is determined.

10.1.2 Decay-time dependent efficiency

Due to *e.g.* some selection criteria and trigger requirements, as well as inefficiencies in the VELO reconstruction, the detector efficiency is not constant over the B^0 decay-time. This efficiency, referred to as acceptance $a(t)$, decreases very quickly towards zero for low decay times, reaches a plateau for intermediate decay times, and slightly drops again at high decay times.

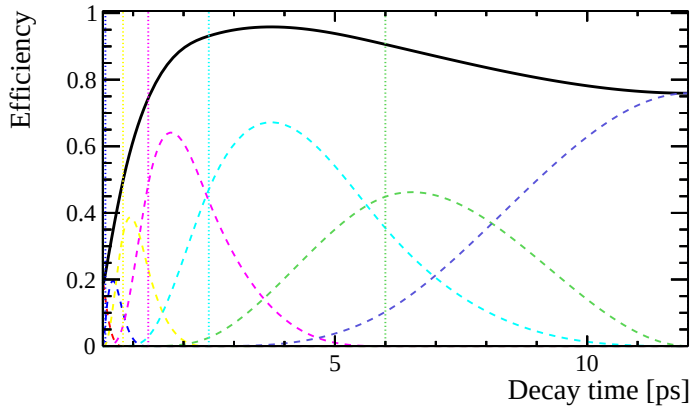

Fig. 10.2: Graphical representation of the acceptance for $B^0 \to D^{\mp}\pi^{\pm}$ decays. The dotted vertical lines represent the knot positions, the dashed lines show the underlying cubic polynomials, where the same colour is chosen for the associated knot and polynomial.

For this analysis, two models were developed in parallel, which give almost identical results for the *CP* parameters S_f and $S_{\bar{f}}$. The model used in the final decay-time fit was developed by a collaborator and has an additional degree of freedom, while the model described below is used as a crosscheck and for estimating systematic uncertainties. In both models, the acceptance is parametrised by splines, which are implemented analytically in the decay-time fit as described in Ref. [94]. These splines consist of cubic polynomials defined piecewise in decay-time.

The final acceptance parameterisation is characterised by the limits of the ranges on which the cubic polynomials are defined (also denoted as knots) and associated coefficients. It is optimised in order to find the ideal knot positions giving a good description of the decay-time while minimising the number of knots. This is done on simulated $B^0 \to D^{\mp}\pi^{\pm}$ events by performing a maximum-likelihood fit to the decay-time with the PDF defined as

$$\mathscr{A}(t) \propto a(t) \int dt' \mathscr{R}\left(t - t'\right) e^{t'/\tau} \tag{10.2}$$

where the resolution $\mathscr{R}\left(t - t'\right)$ with the true and reconstructed decay times t and t' is taken from Sec. 10.1.1 and the lifetime τ is fixed to the value used in the generation. It is further checked if the obtained model also describes the *sWeighted* decay-time distribution in the $B^0 \to D^{\mp}\pi^{\pm}$ sample. Instead of fixing the lifetime on data, it is constraint by means of a Gaussian function to the world average $\tau = (1.518 \pm 0.004)\,\text{ps}$ [18]. A good description was found using seven knots at $[0.4, 0.45, 0.8, 1.3, 2.5, 6.0, 12.0]\,\text{ps}$, where the coefficient at 2.5 ps is set to one to fix the overall normalisation. Figure 10.2 shows a graphical representation of the used parameterisation with the coefficients obtained on $B^0 \to D^{\mp}\pi^{\pm}$ data, the numerical values of the coefficients are given in Tab. 10.1.

Tab. 10.1: Spline coefficients v_i as obtained for the decay-time distribution on $B^0 \to D^{\mp}\pi^{\pm}$. The coefficient v_5 is set to one to fix the overall normalisation.

Parameter	Value
v_1	0.187 ± 0.004
v_2	0.306 ± 0.005
v_3	0.557 ± 0.005
v_4	0.870 ± 0.010
v_5	1.0
v_6	0.880 ± 0.023
v_7	0.759 ± 0.023

10.1.3 Extraction of *CP* observables

The *CP* parameters S_f and $S_{\bar{f}}$ are determined through a multi-dimensional unbinned maximum-likelihood fit to the *sWeighted* (background-subtracted) distributions of the $B^0 \to D^{\mp}\pi^{\pm}$ candidates. The PDF to describe the decay time t, the tags $\vec{d} = (d^{\text{OS}}, d^{\text{SS}})$ and the final state F taking the values f and $\bar{f}$ given the mistags $\vec{\eta} = (\eta^{\text{OS}}, \eta^{\text{SS}})$ is defined by

$$\mathscr{P}(t, F, \vec{d}|\vec{\eta}) \propto a(t) \left(P(t', F, \vec{d}|\vec{\eta}) \otimes R(t' - t) \right) \tag{10.3}$$

where $P(t', F, \vec{d}|\vec{\eta})$ describes the true decay time, $R(t'-t)$ is the resolution from Sec. 10.1.1 and $a(t)$ parametrises the acceptance described in Sec. 10.1.2. Furthermore, the function $P(t', F, \vec{d}|\vec{\eta})$ corresponds to the decay rates from Eqs. (4.32) to (4.35) taking into account the corrections from Eq. (9.1). Besides, production and detection asymmetry must be described. These are defined as

$$A_{\text{P}} = \frac{\sigma(\overline{B}^0) - \sigma(B^0)}{\sigma(\overline{B}^0) + \sigma(B^0)} \quad \text{and} \quad A_{\text{D}} = \frac{\varepsilon(f) - \varepsilon(\bar{f})}{\varepsilon(f) + \varepsilon(\bar{f})} \tag{10.4}$$

where ε is the decay-time integrated reconstruction and selection efficiency for the final states f and $\bar{f}$ and σ is the production cross-section for B^0 and $\overline{B}^0$ mesons. Both asymmetries were determined to be at the percent level in independent measurements at the LHC [95]. As both are further known to be decay-time independent they can be described by modifying the expressions for the *CP* coefficients from Eq. (9.1) further to

$$\begin{aligned} \left(\Delta^- - \Delta^+\right) S_f &\to \left(\Delta^- - A_{\text{P}}\,\Delta^+\right)(1 + A_{\text{D}}) S_f\,, \\ \left(\Delta^- - \Delta^+\right) C_f &\to \left(\Delta^- - A_{\text{P}}\,\Delta^+\right)(1 + A_{\text{D}}) C_f\,. \end{aligned} \tag{10.5}$$

The same expressions also apply to $S_{\bar{f}}$ and $C_{\bar{f}}$ with the substitution $A_{\text{D}} \to -A_{\text{D}}$.

As explained in Sec. 9.1, the parameters C_f and $C_{\bar{f}}$ are fixed to 1 and -1, due to the expected small value of r (see Sec. 4.2). Moreover, since possible tagging efficiency

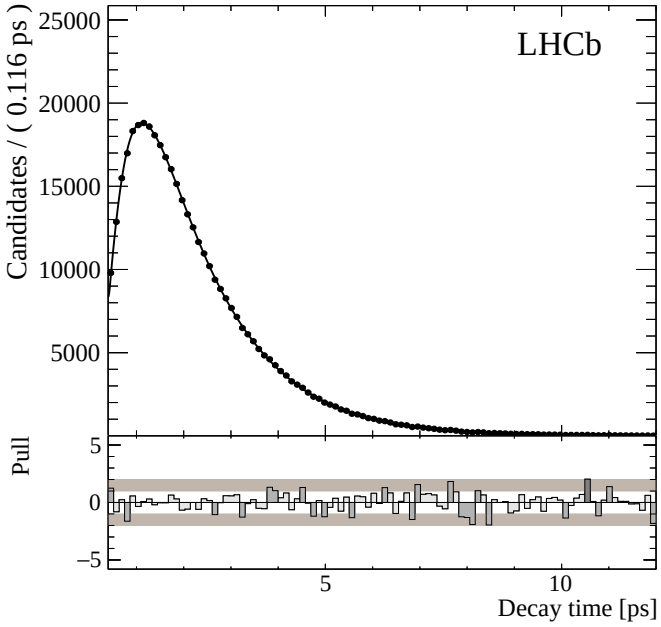

Fig. 10.3: Background-subtracted decay-time distribution of $B^0 \to D^{\mp}\pi^{\pm}$ candidates. The solid curve is the projection of the PDF, the black points represent the data. The lower histogram shows the distributions of pulls, *i.e.* the difference of the binned data and the fitted PDF divided by the data uncertainty in each bin.

asymmetries are measured in simulation to be compatible with zero, they are fixed to this value for the OS and SS taggers. Possible systematic effects due to one of both assumptions are taken into account in Ch. 11. Furthermore, the the B^0 lifetime and the oscillation frequency are constrained by means of a Gaussian function to $\tau = (1.518 \pm 0.004)$ ps [18] and $\Delta m = (0.5050 \pm 0.0023)\,\text{ps}^{-1}$ [96]. Hence, the completely floating parameters in the fit are the *CP* parameters S_f and $S_{\bar{f}}$, the production and detection asymmetry, the calibration parameters of the OS and SS taggers and the acceptance parameters. The fitted values for the parameters S_f, $S_{\bar{f}}$, Δm, $\Delta\Gamma$, A_{P} and A_{D} are shown in Tab. 10.2. Figure 10.3 shows the projection of the PDF onto the decay-time distribution. For the *CP* parameters S_f and $S_{\bar{f}}$, it is important to note that the given uncertainties are not purely statistical, but also include the systematic contributions from Δm and τ via the applied constraints. Repeating the fit with Δm and τ fixed to the central values of the constraints, the central values for S_f and $S_{\bar{f}}$ stay unchanged, but the uncertainties decrease to 0.020.

Tab. 10.2: Fit results for S_f, $S_{\bar{f}}$, Δm, $\Delta\Gamma$, A_{P} and A_{D} from the nominal decay-time fit in $B^0 \to D^{\mp}\pi^{\pm}$. The uncertainties on S_f and $S_{\bar{f}}$ are not purely statistical, but contain the systematic contributions from the constraints on Δm and τ.

Parameter	Value
S_f	0.058 ± 0.021
$S_{\bar{f}}$	0.038 ± 0.021
Δm	$(0.5054 \pm 0.0022)\,\mathrm{ps}^{-1}$
τ	$(1.5180 \pm 0.0040)\,\mathrm{ps}$
A_{P}	-0.0064 ± 0.0028
A_{D}	0.0086 ± 0.0019

In Fig. 10.4, the signal yield asymmetries given by

$$\begin{aligned} A_{CP}^{f}(t) &= \frac{\Gamma\left(B^0 \to D^-\pi^+\right) - \Gamma\left(\overline{B}^0 \to D^-\pi^+\right)}{\Gamma\left(B^0 \to D^-\pi^+\right) + \Gamma\left(\overline{B}^0 \to D^-\pi^+\right)} \propto -S_f \sin(\Delta m t) + C_f \cos(\Delta m t)\ , \\ A_{CP}^{\bar{f}}(t) &= \frac{\Gamma\left(B^0 \to D^+\pi^-\right) - \Gamma\left(\overline{B}^0 \to D^+\pi^-\right)}{\Gamma\left(B^0 \to D^+\pi^-\right) + \Gamma\left(\overline{B}^0 \to D^+\pi^-\right)} \propto -S_{\bar{f}} \sin(\Delta m t) + C_{\bar{f}} \cos(\Delta m t)\ , \end{aligned} \tag{10.6}$$

when neglecting the asymmetries from the flavour tagging and in the detection and production are shown. Obviously, for these asymmetries values for S_f and $S_{\bar{f}}$ different from zero would show a shifted oscillation with respect to the cosine oscillation, which stems from the B^0-$\overline{B}^0$ mixing. However, such shift would only be a necessary but not sufficient condition for *CP* violation as pointed out in Eq. (3.52): in the case that *CP* violation occurs in the decay $B^0 \to D^{\mp}\pi^{\pm}$ the shift must be different for the asymmetries $A_{CP}^{f}(t)$ and $A_{CP}^{\bar{f}}(t)$. Unfortunately, the values for S_f and $S_{\bar{f}}$ are small, so that these asymmetries are mainly dominated by the cosine term and therefore even a shift is barely visible. A better alternative, not suffering from this discrepancies in magnitude between $S_f/S_{\bar{f}}$ and $C_f/C_{\bar{f}}$, are the signal yield asymmetries between candidates tagged as B^0 and $\overline{B}^0$ split according to the favoured (F) $\overline{b} \to \overline{c}\,[u\overline{d}]$ and the suppressed (S) $\overline{b} \to \overline{u}\,[c\overline{d}]$ transitions. Neglecting again asymmetries in the flavour tagging and in the detection and production these can be expressed as

$$\begin{aligned} A_{\mathrm{F}}(t) &= \frac{\Gamma\left(B^0 \to D^-\pi^+\right) - \Gamma\left(\overline{B}^0 \to D^+\pi^-\right)}{\Gamma\left(B^0 \to D^-\pi^+\right) + \Gamma\left(\overline{B}^0 \to D^+\pi^-\right)} \propto -\left(S_f + S_{\bar{f}}\right) \sin(\Delta m t)\ , \\ A_{\mathrm{S}}(t) &= \frac{\Gamma\left(\overline{B}^0 \to D^-\pi^+\right) - \Gamma\left(B^0 \to D^+\pi^-\right)}{\Gamma\left(\overline{B}^0 \to D^-\pi^+\right) + \Gamma\left(B^0 \to D^+\pi^-\right)} \propto \left(S_f + S_{\bar{f}}\right) \sin(\Delta m t)\ . \end{aligned} \tag{10.7}$$

Apparently, the cosine terms in the numerator vanishes, so that in case of no *CP* violation, *i.e.* $S_f = S_{\bar{f}}$, the asymmetries vanish. However, as shown in Fig. 10.4 an oscillation indicating

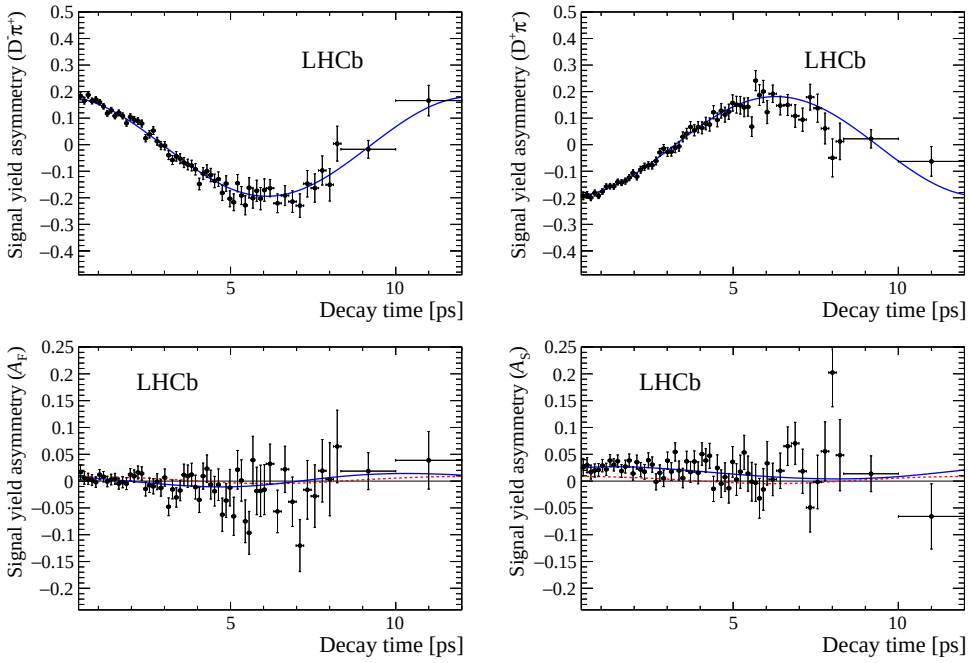

Fig. 10.4: Distributions of the decay-time dependent signal yield asymmetry for the $D^-\pi^+$ (top left) and the $D^+\pi^-$ (top right) finalstate and of the decay-time dependent signal yield asymmetry for the favoured (bottom left) and the suppressed (bottom right) transitions as defined in Eq. (10.7). The blue solid curve is the projection of the fitted PDF from the nominal fit, the red dotted curve in the lower plots shows the projection of a second fit under the assumption of no *CP* violation.

CP violation can be seen. Yet, neglecting the contributing experimental asymmetries in Eq. (10.7) oversimplified the expressions, as indicated by an alternative fit under the asumption of no *CP* violation: the resulting fit projection is not flat, but also shows an oscillating structure from contributions of the asymmetries in the numerator on the one hand and from sine and cosine contributions arising in the denominator of $A_{\mathrm{F}}(t)$ and $A_{\mathrm{S}}(t)$ on the other hand. The difference between these two fits yields a significance of 2.7σ for *CP* violation according to Wilk's theorem [97].

Using the results of the fitted calibration parameters (the results themselves are discussed in Sec. 10.2) and the fitted tagging efficiencies of $\varepsilon_{\mathrm{tag}}^{\mathrm{OS}} = (43.24 \pm 0.07)\,\%$ and $\varepsilon_{\mathrm{tag}}^{\mathrm{SS}} = (93.05 \pm 0.04)\,\%$, the tagging performances in the sample can be computed as shown in Eq. (6.22). The average dilution as defined in Eq. (6.23) is $(9.53 \pm 0.03)\,\%$ and $(2.789 \pm 0.009)\,\%$ for the OS and SS taggers, respectively. This leads to an overall average dilution of $(6.55 \pm 0.02)\,\%$. Including the untagged candidates, which were removed in Ch. 8, the total effective tagging efficiency is calculated to be $(5.59 \pm 0.01)\,\%$.

Tab. 10.3: Calibration parameters obtained in the decay-time fit in $B^0 \to D^{\mp}\pi^{\pm}$. The deviations are calculated with respect to the calibration parameters derived from the control modes $B^+ \to D^0\pi^+$ (Tab. 9.14) and $B^0 \to J/\psi K^{*0}$ (Tab. 9.11).

	OS		SS	
Parameter	Value	Deviation	Value	Deviation
p_0	-0.152 ± 0.021	-0.56	-0.041 ± 0.021	0.80
p_1	-0.035 ± 0.024	-0.89	-0.012 ± 0.022	0.22
p_2	-0.007 ± 0.009	-0.33	-	-
p_3	-0.32 ± 0.11	0.90	-	-
p_4	-0.47 ± 0.49	0.57	-	-
Δp_0	-0.079 ± 0.049	0.81	-0.085 ± 0.044	-1.25
Δp_1	0.140 ± 0.036	1.72	0.042 ± 0.033	0.11
Δp_2	-0.024 ± 0.013	-0.19	-	-
Δp_3	-0.26 ± 0.16	-2.66	-	-
Δp_4	-0.52 ± 0.71	-2.11	-	-

10.2 Fit validation

To validate the decay-time fit, first the fitted nuisance parameters like the production and detection asymmetry and the flavour tagging calibration parameters are compared to reference values. The results of $A_{\mathrm{P}} = (-0.64 \pm 0.28)\,\%$ and $A_{\mathrm{D}} = (0.86 \pm 0.19)\,\%$ are well in agreement with the values from an independent LHCb measurement [95], which *e.g.* range from $(-1.43 \pm 0.86)\,\%$ to $(-0.56 \pm 0.30)\,\%$ for measurements of the production asymmetry at centre-of-mass energies of 7 and 8 TeV in the decay channels $B^0 \to J/\psi K^{*0}$ and $B_s^0 \to D_s^- \pi^+$. The obtained calibration parameters are compared to those computed on the control channels $B^+ \to D^0\pi^+$ and $B^0 \to J/\psi K^{*0}$ for the OS and SS, respectively. In Tab. 10.3, the parameters from the decay-time fit in $B^0 \to D^{\mp}\pi^{\pm}$ are listed and the deviation of each parameter to the calibrations given in Tab. 9.11 and Tab. 9.14 are calculated. The largest deviation can be found for Δp_3^{OS} and Δp_4^{OS} being larger than two standard deviations. Additionally, taking into account the correlations between the parameters an overall discrepancy is calculated yielding 0.91σ for the OS and 0.29σ for the SS taggers demonstrating that the results of flavour tagging calibrations are quite similar despite the not given portability.

In a second step, the two-dimensional contour plots for the *CP* parameters S_f and $S_{\bar{f}}$ and for the detection and production asymmetries are checked. As shown in Fig. 10.5 both do not show any unexpected behaviour, indicating that the corresponding uncertainties are well understood.

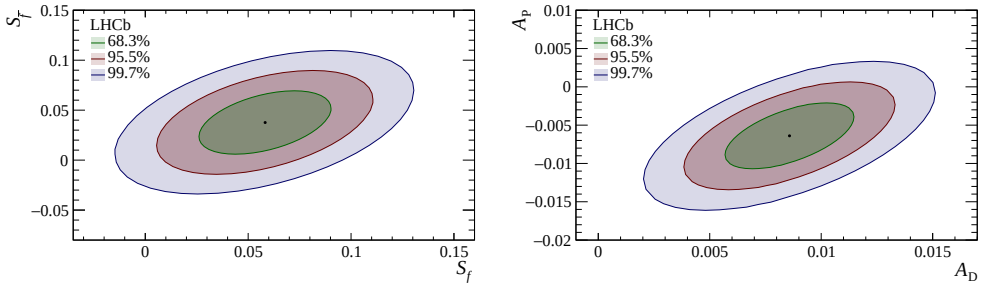

Fig. 10.5: Contour plot for (S_f, $S_{\bar{f}}$) (left) and (A_P, A_D) (right) showing the one, two and three sigma contours. The uncertainties include the full statistical uncertaintiy and the systematic uncertainty due to the constraints on Δm and τ.

10.2.1 Validation of link function for mistags

As mentioned before, the handling of candidates with a mistag close to 0.5 is important, both in case of a calibration with parameters constrained by means of a Gaussian function and completely floating calibration parameters, to guarantee a stable and unbiased decay-time fit. Therefore, the two scenarios presented additionally to the nominal scenario in Sec. 6.4.3 are tested using pseudoexperiments.

In each study presented below, 1000 pseudoexperiments are generated according to the PDF from Eq. (10.3). To simplify the used model and reduce the number of parameters, the flavour tagging calibration functions are reduced to linear models as *basis functions* (see Eq. (6.27)) and the identity as *link function*. The calibration parameters used for the generation are obtained from a linear calibration with the identity as *link function* on the control channels $B^+ \to D^0\pi^+$ and $B^0 \to J/\psi K^{*0}$. It should be noted that the calibration for the OS taggers is shifting the estimated mistags to higher values, *i.e.* $p_1^{\mathrm{OS}} > 1$, while the calibration for the SS taggers shows the opposite, *i.e.* $p_1^{\mathrm{SS}} < 1$, behaviour. Furthermore, the calibration function is implemented such that the mistag probability ω is not defined outside the range $[0, 0.5]$, *i.e.* if the mistag probability exceeds 0.5, the tag-decision is set to $d = 0$ and the corresponding mistag to $\omega = 0.5$. For all studies presented in the following, the pull distributions of the floating parameters are checked, where the pull is defined as the fitted value minus the value used in the generation of the simulated sample divided by the uncertainty on the fitted value. The pull distributions obtained from each set of pseudoexperiments are then fitted with a Gaussian function in order to determine the mean and width. A deviation of more than one standard deviation of the mean value from zero indicates a possible bias, while a deviation of more than three standard deviations is interpreted as a clear bias. These generated samples are then fitted with different approaches:

- In the first approach,the tag decision is flipped in case the mistag probability ω' exceeds 0.5 and the mistag probability is calculated as $\omega = 1 - \omega'$. In the fit, the

calibration parameters are constrained by means of a Gaussian function. This leads to biased calibration parameters for the OS algorithm. To understand if a possible bias on S_f and $S_{\bar{f}}$ is just "absorbed" by the calibration parameters, the same samples are also fitted with the calibration parameters fixed. In this case, the *CP* parameters show a small deviation of 2σ and 1.3σ for S_f and $S_{\bar{f}}$, respectively.

- To further understand if this small deviation is just a fluctuation or a real bias, two possible sources of the bias are investigated: in a first study the tagging asymmetry parameters Δp_i^{OS} are artificially increased *i.e.* the parameters are increased in both steps during generation and fitting. In a second study the tagging asymmetry parameters Δp_i^{OS} are reduced to their nominal values but instead the parameter p_1^{OS} is increased. In the fit, the tag decision is flipped in both studies if the mistag probability ω' exceeds 0.5 and the mistag probability is calculated as $\omega = 1 - \omega'$. To ensure that a possible bias is not "absorbed" by the calibration parameters, the calibration parameters are fixed in the fit . The resulting pull distributions for S_f and $S_{\bar{f}}$ are shown in Fig. 10.6. One can see that in case of the increased flavour tagging asymmetry parameters, the *CP* parameters are clearly biased by more than ten standard deviations, while the result is unbiased in case of the enlarged p_1^{OS} parameter.

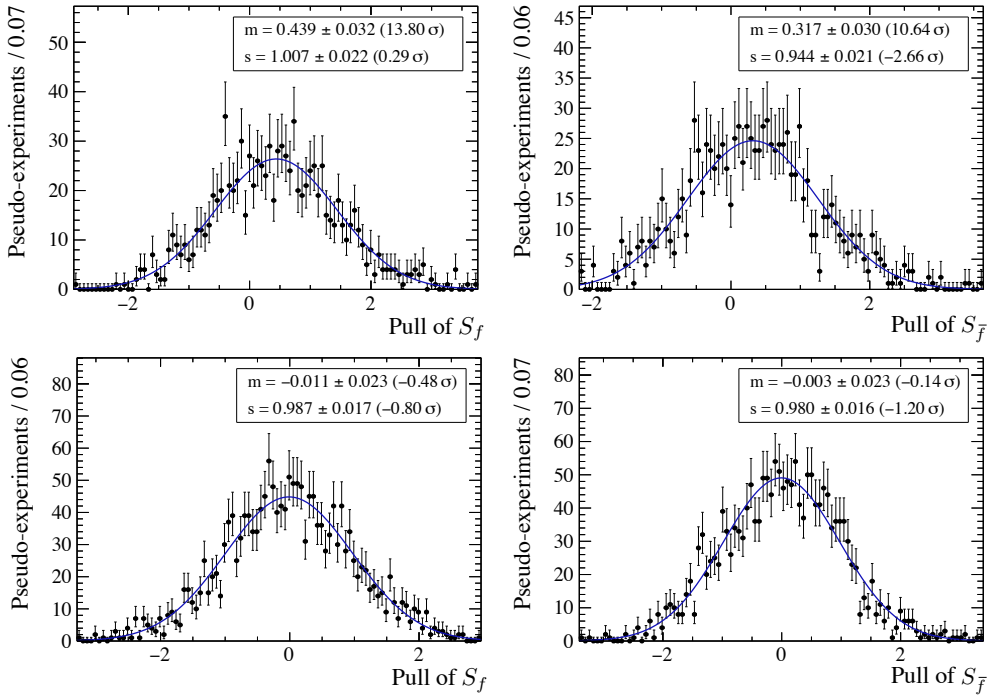

Fig. 10.6: Pull distributions of S_f (left) and $S_{\bar{f}}$ (right) when generating pseudoexperiments with artificially enlarged mistag asymmetry calibration parameters and a flip of the tag decision if $\omega' > 0.5$ (top) and with the artificially enlarged parameter p_1^{OS} and a flip of the tag decision if $\omega' > 0.5$ (bottom).

- In a last study, the pseudoexeriments are generated with artificially increased tagging asymmetry parameters. But instead of flipping the tag decision, it is set to $d = 0$ when the mistag probability exceeds 0.5. In order to achieve a stable fit the distribution of estimated mistags η for the OS and SS is reduced beforehand to

$$\eta < \frac{0.5 - (p_0 + \delta p_0) + (p_1 + \delta p_1 \langle\eta\rangle}{p_1 + \delta p_1}, \tag{10.8}$$

  where δp_i are the uncertainties of the calibration parameters. This assures that the mistag probabilities do not exceed 0.5. This strategy yields unbiased results for S_f and $S_{\bar{f}}$.

From this studies, it can be concluded that the flip of the tag decision can bias the measurement of *CP* parameters. However, the size of the bias depends on the specific values of the calibration parameters and this needs to be studied for each specific set of values. On the other hand, reducing the allowed range of estimated mistags prevents a bias on the measurement of *CP* parameters, but depending on the cut that needs to be applied, this could reduce the statistical sensitivity of the analysis. Therefore, the modified *link function* as used in the nominal approach currently provides the best unbiased appraoch as will be shown in Sec. 10.2.3.

10.2.2 Cross checks on sub samples

The stability of the fit is checked by also performing the fit in different sub samples of the full data set. The data set is split in several ways, namely by data taking conditions, used tagging algorithms or kinematic properties of the B^0 meson and properties of the event.

When splitting according to data taking conditions, the $B^0 \to D^{\mp}\pi^{\pm}$ sample is divided by the year of data taking and magnetic polarity. For each sub sample, the *sWeights* are determined with a dedicated mass fit according to the procedure from Ch. 8. A comparison of the fitted values for S_f and $S_{\bar{f}}$ between the four sub samples is shown in Fig. 10.7. The obtained results for S_f and $S_{\bar{f}}$ show good agreement and the average result from the fits in the sub samples is well compatible with the result of the nominal fit.

When using two classes of tagging algorithms, the full sample is divided into three independent samples. The first sub sample contains candidates tagged exclusively by the OS algorithms while the second sample consists of candidates which are only tagged by the SS algorithms. The third class contains candidates which are tagged by both, OS taggers and SS taggers. Again, the results for all sub samples show good agreement (see Fig. 10.8). Furthermore, this agreement gives additional confidence that the strategy of floating the calibration parameters in the decay-time fit provides a stable result for the *CP* parameters S_f and $S_{\bar{f}}$.

Finally, the data set is split in four bins in the transverse momentum of the B^0 mesons, three bins in the number of reconstructed PVs and tracks in the event and in four bins in the

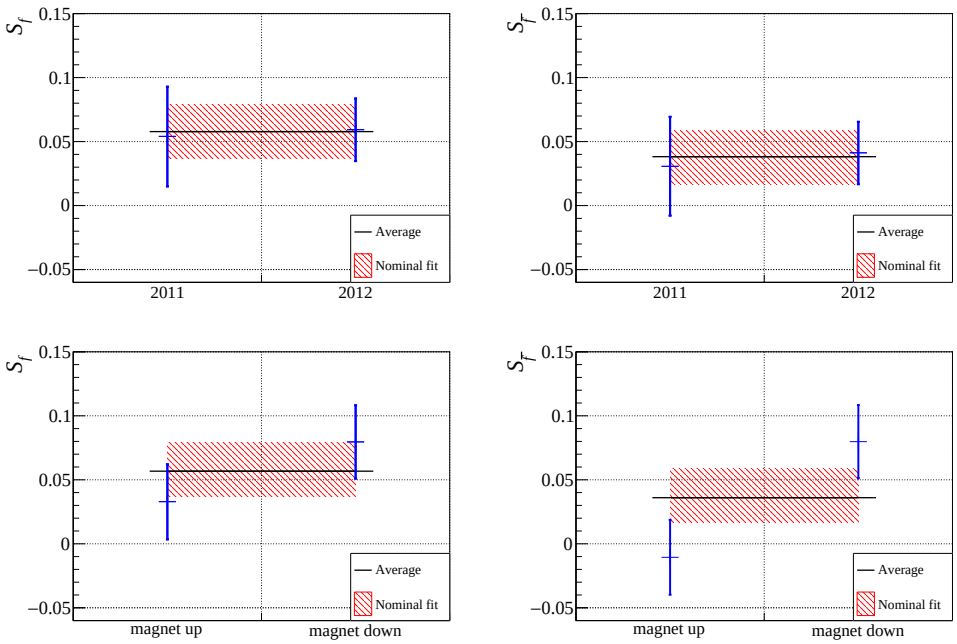

Fig. 10.7: Comparison between the fitted values of S_f (left) and $S_{\bar{f}}$ (right) in sub samples split by year of data taking (top) and magnet polarity (bottom). The blue points are the results of the fits in the sub samples, the red dashed area represents the result of the nominal fit and the black line is the average of the results obtained in the sub samples.

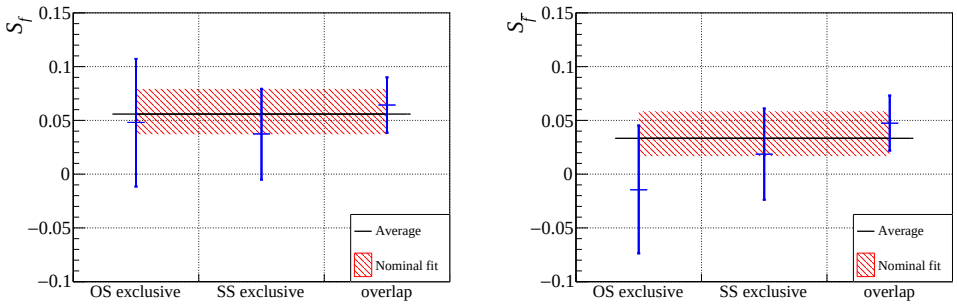

Fig. 10.8: Comparison between the fitted values of S_f (left) and $S_{\bar{f}}$ (right) when considering candidates exclusively tagged by the OS, SS or both classes of tagging algorithms. The blue points are the results of the fits in the sub samples, the red dashed area represents the result of the nominal fit and the black line is the average of the results obtained in the sub samples.

difference in pseudo-rapidity between the $D^{\pm}$ meson and the bachelor particle. The reason for these splits is that the flavour tagging calibrations partly depend on these observables, and therefore could cause a bias in the corresponding splits. Moreover, the difference in pseudo-rapidity is also sensitive to possible misalignments in the detector, which could influence the measurement of S_f and $S_{\bar{f}}$. However, all results show compatible results and no trends are observed.

10.2.3 Decay-time fits to simulated events

To validate the fit using simulated events, these are bootstrapped, *i.e.* the simulated data sample is resampled n times, whereby it is allowed that single events can be taken more than once, *e.g.* a bootstrapped sample can contain the same event multiple times. This is statistically valid because individual events are not correlated with each other. Each generated sample then contains as many candidates as signal candidates in the full $B^0 \to D^{\mp}\pi^{\pm}$ data sample used in Sec. 10.1.3 in order to obtain the same statistical uncertainties.

After generation, the samples are fitted with the same strategy as the nominal fit to extract the *CP* observables. The constrained parameters τ and Δm are treated as follows: for each fit, a value is generated randomly from the respective Gaussian function with which τ and Δm are constrained. This new value is then used in the fit as the mean value of the constraints. This allows the correct fluctuation for both parameters and prevents an underestimation of the fitted uncertainties. For the nominal constraint, the generation values of the simulated sample are used as the mean value, while for the width, the same value as on data is used. This means that the lifetime is constrained to $\tau = (1.519 \pm 0.004)\,\mathrm{ps}$ and the oscillation frequency to $\Delta m = (0.5100 \pm 0.0023)\,\mathrm{ps}^{-1}$.

For all settings described below, the distributions of residuals are studied for S_f and $S_{\bar{f}}$, whereby the residual is defined as the fitted value minus the value used in the generation of the simulated sample. This residual distributions are fitted with a Gaussian function in order to determine the mean and width. A mean value deviating from zero hints to a biased result, while the width of the distribution allows to determine the expected uncertainty of the parameter. Performing such a study with 1000 bootstrapped samples with the nominal strategy yields a mean of 0.0064 ± 0.0007 for S_f and -0.0024 ± 0.0007 for $S_{\bar{f}}$. This corresponds to a deviation of roughly one third of the statistical uncertainty for S_f and about 10 % of the statistical uncertainty for $S_{\bar{f}}$ as shown in Fig. 10.9. Furthermore, the following configurations are also investigated with 1000 bootstrapped samples each:

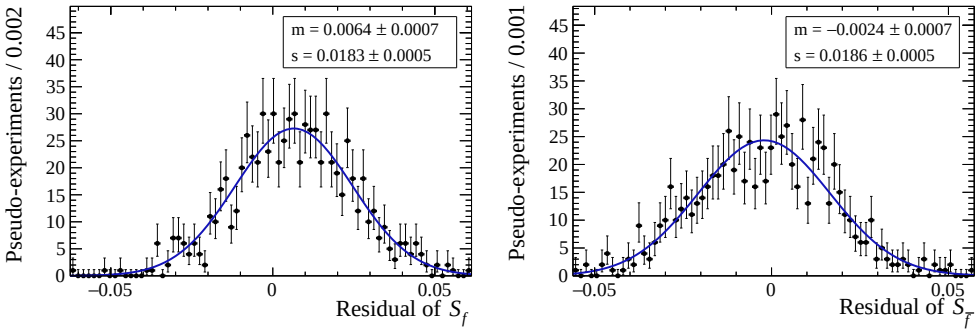

Fig. 10.9: Distribution of residuals for S_f (left) and $S_{\bar{f}}$ (right) using the nominal fit strategy with floating calibration.

- Using the true generated flavour of the B candidate instead of the tag decision and mistag estimate provided by the real tagging algorithms leads to an unbiased distribution of residuals for S_f and $S_{\bar{f}}$.
- A *cheated* tagger is implemented for simulated data. Instead of using the perfect tagging as in the first appraoch, the truth information for each candidate is resampled depending on the mistag probability. This way, the mistag is used as a conditional observable as is done in the nominal fit, but still the truth information from the simulation is exploited. This appraoch also gives unbiased results for S_f and $S_{\bar{f}}$.
- The retraining of the SS tagging algorithms is performed on simulated samples in the same way as described in Sec. 9.2. Afterwards, the calibration for the OS and SS algorithms is obtained from $B^0 \to D^{\mp}\pi^{\pm}$ using the true generated flavour as done for the portability checks in Sec. 9.2 and Sec. 9.3. Performing the fits to the bootstrapped simulation samples of $B^0 \to D^{\mp}\pi^{\pm}$, no bias on S_f and $S_{\bar{f}}$ is observed.
- The retraining and calibration of the tagging algorithms is performed on simulated samples in the same way as described in Sec. 9.2 and Sec. 9.3. This calibration is applied in the fits to the bootstrapped simulation samples of $B^0 \to D^{\mp}\pi^{\pm}$, what leads to a bias on S_f and $S_{\bar{f}}$ of the size of the statistical uncertainty of both parameters.
- Instead of fixing the calibration parameters obtained on simulated samples, they are constrained by means of Gaussian functions in the decay-time fits to the bootstrapped simulation samples of $B^0 \to D^{\mp}\pi^{\pm}$. These constraints are implented by multidimensional Gaussian functions taking into account the correlations on the simulated control samples. This approach reduces the bias on S_f and $S_{\bar{f}}$ to a value of the order of half the statistical uncertainty of both parameters.

This confirms that leaving the flavour tagging calibration parameters free in the *CP*-fit is the best choice. However, since the source of this potential, but anyway small bias cannot

be narrowed down further than coming from the flavour tagging calibration, it is included as systematic uncertainty. To confirm the size of the bias, a second study was performed by a collaborator yielding as mean values of the distribution of residuals 0.0071 ± 0.0006 and -0.0013 ± 0.0006 for S_f and $S_{\bar{f}}$, respectively. Finally, the average value from both studies, *i.e.* 0.0068 ± 0.0005 for S_f and -0.0018 ± 0.0005 for $S_{\bar{f}}$ is assumed as systematic uncertainty (see Ch. 11).

11 Systematic uncertainties

The maximum-likelihood fits to the invariant mass and decay-time of the $B^0 \to D^{\mp}\pi^{\pm}$ candidates are designed to describe the data and take uncertainties on all parameters correctly into account. However, not all effects possibly influencing the result can be accounted for directly in the fit and therefore systematic uncertainties need to be calculated: Parameters are constrained by means of Gaussian functions in various steps of the analysis and the influence of these constraints on the resulting uncertainties is estimated in Sec. 11.1. Further ensembles of pseudoexperiments are used to test the systematic effects of certain assumptions such as $\Delta\Gamma = 0$ as described in Sec. 11.2. Systematic effects related to the mass model and the associated determination of the *sWeights* are discussed in Sec. 11.3. Furthermore, the potential fit biases observed in the fits to simulated events in Sec. 10.2.3 are assigned as systematic uncertainties.

It is further important to mention that all work in this chapter related to the mass fit was done by a collaborator. This includes parts of the estimated uncertainties in Sec. 11.1 and the estimations presented in Sec. 11.3.

A summary of all systematic uncertainties is given in Tab. 11.1; all are assumed to be symmetric for both *CP* parameters. The total uncertainty of 0.0111 and 0.0073 on S_f and $S_{\bar{f}}$, respectively, is calculated from the sum of the squared individual contributions, *i.e.* the individual contributions are assumed to be uncorrelated. The correlation between the systematic uncertainties on the *CP* parameters is $-41\,\%$.

11.1 Systematic uncertainties from Gaussian constraints

Gaussian constraints have been used in two different steps of the analysis. On the one hand, such onstraints on the B^0 oscillation frequency Δm and the B^0 lifetime τ are used in the decay-time fit to account for the limited knowledge of those parameters. On the other hand, the uncertainties on the efficiencies of the $\mathrm{DLL}_{K\pi}$ cuts in Ch. 8 are propagated through constraints by means of Gaussian functions in the mass fit. This second uncertainty was calculated by a collaborator and hence the procedure is described briefly.

In order to investigate the composition of the uncertainties of S_f and $S_{\bar{f}}$ in the decay-time fit, the constrained parameters are set to the central value of the Gaussian constraint to obtain the purely statistical uncertainty. These purely statistical uncertainties are 0.0198 and 0.0199 for S_f and $S_{\bar{f}}$, respectively, with a correlation of $60\,\%$. From the difference of

Tab. 11.1: Systematic uncertainties on the *CP* parameters S_f and $S_{\bar{f}}$ listed by decreasing order for S_f. The "fit biases" are the residuals of the fits to bootstrapped simulated candidates described in Sec. 10.2.3. The total uncertainty is calculated from the sum of the squared individual contributions. The correlation between the uncertainties on S_f and $S_{\bar{f}}$ is −41 %.

Source	S_f	$S_{\bar{f}}$
uncertainty on Δm	0.0073	0.0061
fit biases	0.0068	0.0018
background subtraction	0.0042	0.0023
flavour-tagging models	0.0011	0.0015
flavour-tagging efficiency asymmetries	0.0012	0.0015
decay-time resolution	0.0012	0.0008
$\mathrm{DLL}_{K\pi}$ efficiencies	0.0008	0.0008
acceptance model	0.0007	0.0007
assumption on $\Delta\Gamma$	0.0007	0.0007
assumption on C_f and $C_{\bar{f}}$	0.0006	0.0006
total	0.0111	0.0073
statistical uncertainties	0.0198	0.0199

the squared uncertainties with and without constraint, the systematic uncertainty on the corresponding parameter is then obtained. For Δm this results in a fully anti-correlated systematic uncertainty of 0.0073 and 0.0061 for S_f and $S_{\bar{f}}$, respectively. The systematic uncertainty for the B^0 lifetime τ is found to be negligible.

The efficiency of the requirement on $\mathrm{DLL}_{K\pi}$ depends on the binning of several observables used in the determinaton of $\mathrm{DLL}_{K\pi}$. This effect is refleceted in the uncertainties of the efficiency which are further propagated to the mass fit by Gaussian constraints. The mass fit is repeated with the $\mathrm{DLL}_{K\pi}$ requirement efficiencies fixed to the mean value of the Gaussian constraint. The resulting *sWeights* are then used in an alternative decay-time fit to extract the *CP* parameters. The difference in quadrature between the result for S_f and $S_{\bar{f}}$ of the nominal fit and this alternative fit yields 0.0008 for both S_f and $S_{\bar{f}}$, which is used as systematic uncertainty.

11.2 Estimations with pseudoexperiments

Systematic uncertainties are determined using ensembles of pseudoexperiments by generating data samples of the same size as the $B^0 \to D^{\mp}\pi^{\pm}$ signal yield. In this generation, all parameters are set to the values found in the nominal decay-time fit with exception of S_f and $S_{\bar{f}}$. In order to prevent observer bias, the analysis was performed blind for these

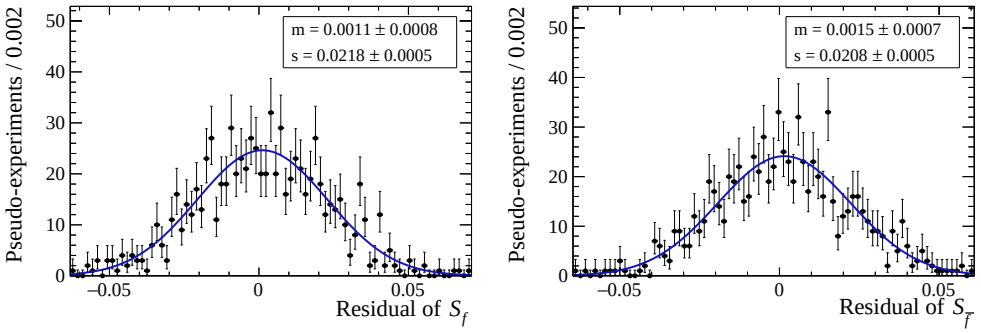

Fig. 11.1: Distribution of residuals for S_f (left) and $S_{\bar{f}}$ (right) to determine the systematic uncertainty due to the flavour tagging calibration model.

parameters and therefore the values used in the generation of the simulated events are adopted for S_f and $S_{\bar{f}}$. For each pseudoexperiment, the PDF from Eq. (10.3) is modified with alternate models in the generation, corresponding to the various assumptions, which are made in the analysis and are then fitted with the nominal model. Each study consists of 1000 pseudoexperiments for which the distribution of residuals of S_f and $S_{\bar{f}}$ are studied. Like in Sec. 10.2.3, this residual distributions are fitted with Gaussian functions, where the deviation from zero of the mean value of this function is taken as systematic uncertainty; if the mean value is compatible within one standard deviation with zero, the error on the mean is taken as systematic uncertainty. In this way, uncertainties are determined for the flavour-tagging calibration model, the assumption on the flavour-tagging efficiency asymmetries, the acceptance model, the decay-time resolution and the assumptions on $\Delta\Gamma$ and C_f.

Flavour tagging calibration model

For the SS taggers, the nominal model with a first-order polynomial is used in the generation while the model for the OS taggers is reduced by one degree compared to the nominal one. In the fit, the polynomials of the calibration models are then increased by one degree of freedom for both types of tagging algorithms, compared to what is used during generation. Figure 11.1 shows the distribution of residuals. The systematic uncertainty is 0.0011 and 0.0015 for S_f and $S_{\bar{f}}$, respectively.

Fixed flavour tagging efficiency asymmetries

Pseudoexperiments are generated with the flavour tagging efficiency asymmetries set to the values obtained on simulated events decreased by their uncertainty, namely $-0.14\,\%$ and $-0.13\,\%$ for the OS and SS taggers, respectively. The samples are then fitted with the efficiency asymmetries fixed to zero. The distributions of residuals for the *CP* parameters

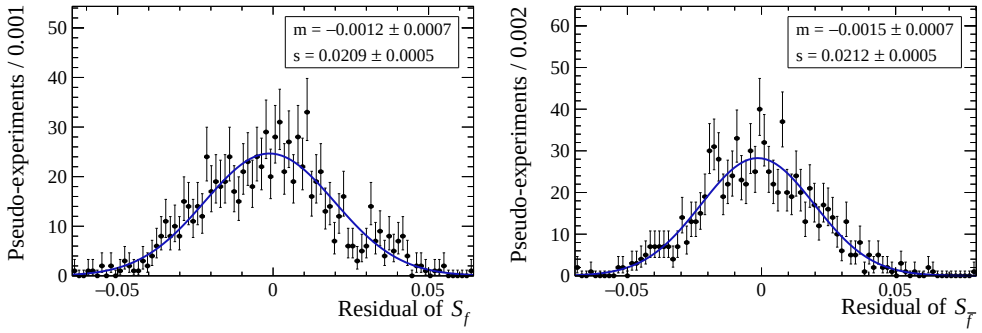

Fig. 11.2: Distribution of residuals for S_f (left) and $S_{\bar{f}}$ (right) to determine the systematic uncertainty due to the flavour tagging efficiency asymmetry.

shown in Fig. 11.2 results in systematic uncertainties of 0.0012 and 0.0015 for S_f and $S_{\bar{f}}$, respectively.

Decay-time resolution

Two different sets of samples are generated to determine the systematic uncertainty due to the decay-time resolution: the first with an average resolution of 20 fs higher than the nominal resolution of 54.91 fs, the second with an average resolution of 20 fs lower than the nominal one. In both cases the nominal resolution is used in the fit. The residual distributions of both studies are shown in Fig. 11.3. The larger uncertainty resulting from the two studies for S_f and $S_{\bar{f}}$ is then chosen as systematic uncertainty. The result is 0.0012 and 0.0008 for S_f and $S_{\bar{f}}$, respectively.

Acceptance model

The samples of the pseudoexperiments are generated with the acceptance model presented in Sec. 10.1.2. The model used in the fit has knots which are located at the following decay times: $[0.4, 0.5, 1.0, 1.5, 2.0, 2.3, 2.6, 3.0, 4.0, 10.0, 12.0]$ ps. The tenth coefficient v_{10} at 10 ps is set to one to fix the overall normalisation, and the eleventh coefficient is determined by a linear extrapolation of the two preceding coefficients (as stated in Sec. 10.1.2, this second model was developed by a collaborator). The distributions of residuals for S_f and $S_{\bar{f}}$ are shown in Fig. 11.4. Since the mean values of the fitted distributions are compatible with zero, a systematic uncertainty of the uncertainty of the mean values of 0.0007 follows for both *CP* parameters.

Assumption on $\Delta\Gamma$

The value of $\Delta\Gamma$ is set to the world average increased by its uncertainty in the generation, namely $0.0079\,\text{ps}^{-1}$ [24]. Since the hyperbolic sine from Eqs. (3.30) to (3.33) does not

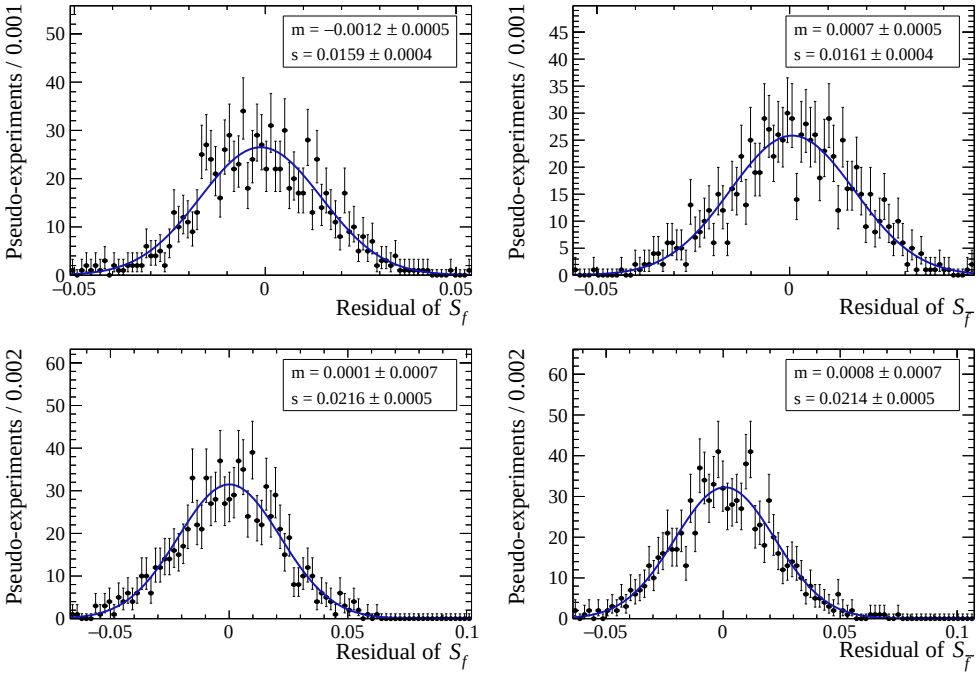

Fig. 11.3: Distribution of residuals for S_f (left) and $S_{\bar{f}}$ (right) to determine the systematic uncertainty due to the decay-time resolution. In the top (bottom) row, the resolution being 20 fs higher (lower) than the nominal model is used.

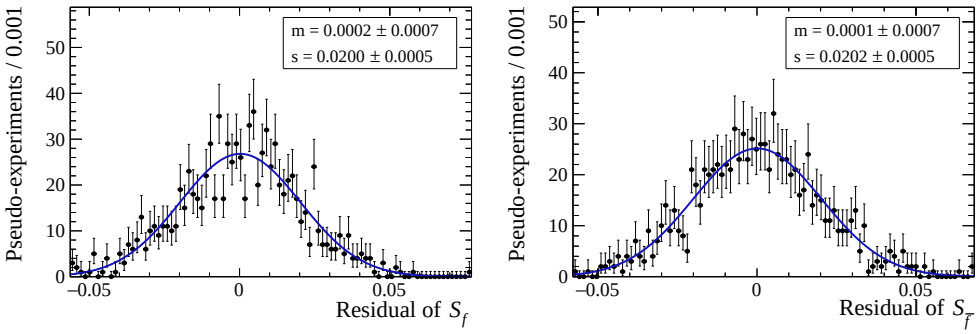

Fig. 11.4: Distribution of residuals for S_f (left) and $S_{\bar{f}}$ (right) to determine the systematic uncertainty due to the acceptance model.

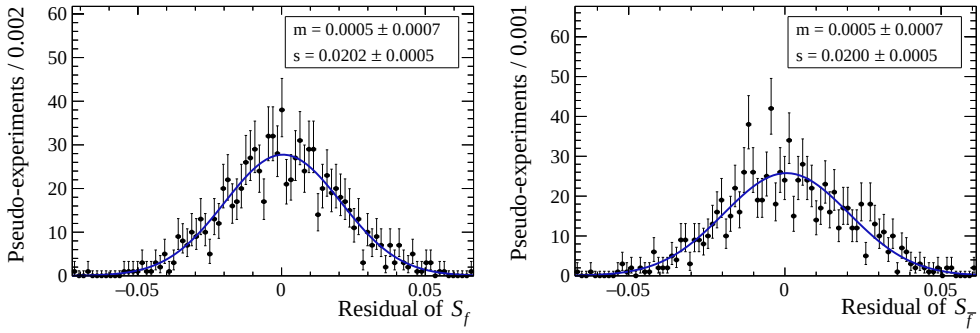

Fig. 11.5: Distribution of residuals for S_f (left) and $S_{\bar{f}}$ (right) to determine the systematic uncertainty due to the assumption on $\Delta\Gamma$.

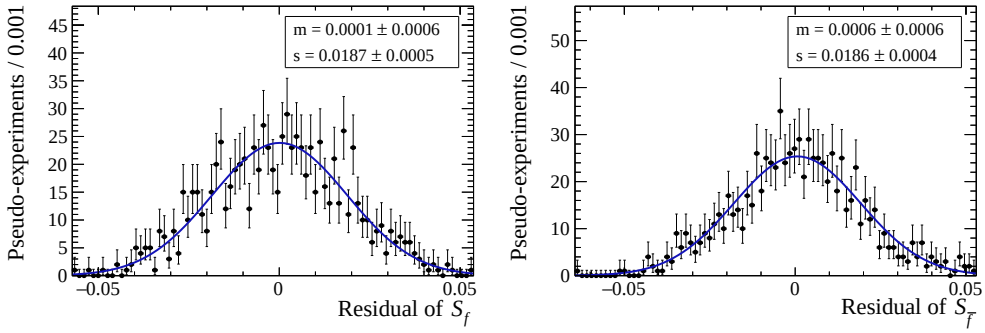

Fig. 11.6: Distribution of residuals for S_f (left) and $S_{\bar{f}}$ (right) to determine the systematic uncertainty due to the assumption on C_f and $C_{\bar{f}}$.

vanish with $\Delta\Gamma \neq 0$, the values for $A_f^{\Delta\Gamma}$ and $A_{\bar{f}}^{\Delta\Gamma}$ need to be defined. The same values as used in the generation of simulated events are used, namely -0.0103 and -0.0155. Then, the samples are fitted with the nominal strategy providing the residuals shown in Fig. 11.5. As systematic uncertainty follows 0.0007 for both, S_f and $S_{\bar{f}}$.

Assumption on C_f

Due to the small value of r (see Sec. 4.2) the values of C_f and $C_{\bar{f}}$ were fixed to 1 and -1 in the nominal fit as described in Sec. 10.1.3. To estimate the systematic uncertainty, in the generation, the values for C_f and $C_{\bar{f}}$ are calculated from the average measurements by Belle and BaBar for the parameter r increased by one statistical uncertainty, namely $C_f = 0.993$ [47, 48]. In the fit, C_f and $C_{\bar{f}}$ are then set to 1 and -1 as in the nominal strategy. The distribution of residuals for S_f and $S_{\bar{f}}$ is shown in Fig. 11.6 and yields 0.0006 as a systematic uncertainty on both parameters.

11.3 Mass model

Since the model to describe the invariant mass is the essential ingredient for the calculation of the *sWeights*, which are used in all subsequent steps of the analysis to statistically subtract background candidates, systematic effects from the parameterisation of the invariant mass can also influence the measurement of S_f and $S_{\bar{f}}$. In order to do this, the fit of the invariant mass as a tool to subtract the background is simply "replaced" by a narrow mass range of $[5250, 5330]\,\mathrm{MeV/c^2}$, *i.e.* no *sWeights* are calculated. This is possible due to the high purity in the signal range. The decay-time fit is then performed on a data sample containing both, signal and backgrounds which are distributed under the signal peak in the invariant mass distribution. The agreement between the nominal result and the result obtained without *sWeights* is 0.2σ and 1.3σ for S_f and $S_{\bar{f}}$, respectively. Due to this good agreement, despite the extreme test without any background suppression in the signal range, no further systematic uncertainty is assigned.

Further, a systematic uncertainty due to the fit strategy, *i.e.* the restriction of the invariant-mass range in Fit B, is estimated. For this purpose, Fit B is also performed in the wide range of the invariant mass, which leads to a larger background contamination in the subsequently used data sample. With the *sWeights* extracted from this fit, the decay-time fit is performed again. The result shows a deviation of 2.3σ and 1.8σ for S_f and $S_{\bar{f}}$, respectively. The difference between these newly obtained values and the nominal results for S_f and $S_{\bar{f}}$ parameters is taken as systematic uncertainty, namely 0.0042 and 0.0023.

Last, the strategy of splitting the data sample in order to control the $B^+ \to D^- K^+$ component is verified, by tightening the cut on the $\mathrm{DLL}_{K\pi}$, which defines the *pion* sample and repeating Fit A and B only for this sample in the narrow signal region $[5220, 5600]\,\mathrm{MeV/c^2}$. This test yields a good agreement of 0.4σ and 1.6σ for S_f and $S_{\bar{f}}$, respectively, and therefore no additional systematic uncertainty is assigned. This is further supported by the fact that a systematic uncertainty due to the requirement on the $\mathrm{DLL}_{K\pi}$ is already calculated in Sec. 11.1 and hence two separate systematic uncertainties would be taken into account for the same experimental effect.

12 Results

The *CP* asymmetries S_f and $S_{\bar{f}}$ are determined in the $B^0 \to D^{\mp}\pi^{\pm}$ decay on the full LHCb Run I data set at centre-of-mass energies of 7 and 8 TeV and measured to be

$$\begin{aligned} S_f &= 0.058 \pm 0.020\,(\text{stat.}) \pm 0.011\,(\text{syst.})\,, \\ S_{\bar{f}} &= 0.038 \pm 0.020\,(\text{stat.}) \pm 0.007\,(\text{syst.})\,, \end{aligned} \tag{12.1}$$

where the statistical and systematic correlations are 60 % and −41 %, respectively. These values are in agreement with, and more precise than, previous measurements from the Belle and BaBar collaborations [98, 99]. According to Wilk's theorem [97], they result in a significance of 2.7σ for *CP* violation. This result, even if it is not yet an evidence for *CP* violation, yields a larger significance for *CP* violation than the previous measurement from the Belle collaboration [98].

Furthermore, the *CP* asymmetries can be expressed using a parametrisation introduced by the BaBar collaboration [99] and adopted by HFLAV [24] with

$$\begin{aligned} a &= -\frac{2r}{1+r^2}\sin(2\beta+\gamma)\cos(\delta)\,, \\ c &= -\frac{2r}{1+r^2}\cos(2\beta+\gamma)\sin(\delta)\,. \end{aligned} \tag{12.2}$$

In this parametrisation only the parameter c is affected by the so-called tag-side interference, an experimental effect due to the coherent $B^0\overline{B}^0$ production at Belle and BaBar [100]. From a comparison with Eqs. (4.36) and (4.37) the transformation rules

$$a = -\frac{1}{2}\left(S_f + S_{\bar{f}}\right) \quad \text{and} \quad c = \frac{1}{2}\left(S_f - S_{\bar{f}}\right) \tag{12.3}$$

follow. Hence, the *CP* asymmetries can be expressed as

$$\begin{aligned} a &= -0.048 \pm 0.018\,(\text{stat.}) \pm 0.005\,(\text{syst.})\,, \\ c &= 0.010 \pm 0.009\,(\text{stat.}) \pm 0.008\,(\text{syst.})\,, \end{aligned} \tag{12.4}$$

where the statistical correlation is zero and the systematic correlation is −46 %.

The values for S_f and $S_{\bar{f}}$ are further interpreted in terms of the angles β and γ, as well as the amplitude ratio r and the *strong* phase δ (see Eq. (4.36) and (4.37)). This is done using a frequentistic approach as described in Ref. [101], where the PDFs f_i containing the experimental observables $\vec{A}_i$ are combined into one likelihood function

$$\mathscr{L}(\vec{\alpha}) = \prod_i f_i\left(\vec{A}_i^{\text{obs}} \middle| \vec{\alpha}\right)\,, \tag{12.5}$$

where $\vec{A}_i^{\text{obs}}$ are the experimentally measured parameters and $\vec{\alpha}$ is the set of parameters to be extracted. For all inputs, Gaussian distributions are assumed according to

$$f_i\left(\vec{A}_i^{\text{obs}}|\vec{\alpha}\right) \propto \exp\left(-\frac{1}{2}\left(\vec{A}_i(\vec{\alpha}) - \vec{A}_i^{\text{obs}}\right)^T V_i^{-1}\left(\vec{A}_i(\vec{\alpha}) - \vec{A}_i^{\text{obs}}\right)\right), \tag{12.6}$$

where V_i is the experimentally determined covariance matrix with the statistical and systematic uncertainties and the corresponding correlations. The best fit point is given as the minimum of a χ^2-function, defined as $\chi^2(\vec{\alpha}) = -2\ln\mathscr{L}(\vec{\alpha})$. The confidence level (CL) for a given parameter value, hereinafter γ_0, is calculated using a test statistic defined as $\Delta\chi^2 = \chi^2(\vec{\alpha}'_{\text{min}}(\gamma_0)) - \chi^2(\vec{\alpha}_{\text{min}})$, where $\chi^2(\vec{\alpha}_{\text{min}})$ is the global minimum and $\chi^2(\vec{\alpha}'_{\text{min}}(\gamma_0))$ is the new minumum with the parameter value γ_0.

The p-value or $1-\text{CL}$ is calculated by a procedure using pseudoexperiments: for each value γ_0, the test statistic $\Delta\chi^2$ is calculated and a set of pseudoexperiments $\vec{A}_j$ is generated according to Eq. (12.6). In this generation, the parameters $\vec{\alpha}$ are set to the values of the new minimum $\vec{\alpha}'$. For each pseudoexperiment, a new test statistic is then calculated by replacing $\vec{A}_{\text{obs}}$ with $\vec{A}_j$, which is again minimised with respect to $\vec{\alpha}$; once with the parameter γ free, once with γ set to γ_0. The $1-\text{CL}$ value is defined as the fraction of pseudoexperiments in which $\Delta\chi^2 < \Delta\chi^{2'}$. More details about this method can also be found in Ref. [102].

By adding external measurements for r, confidence intervals for the quantity $\sin(2\beta+\gamma)$ and the *strong* phase δ can be derived. The ratio r is determined from the branching fraction of $B^0 \to D_s^+\pi^-$ under the assumption of SU(3) symmetry with the same equations one finds in Refs [47, 48]

$$r = \tan(\theta_c)\frac{f_{D^+}}{f_{D_s^+}}\sqrt{\frac{\mathscr{B}\left(B^0 \to D_s^+\pi^-\right)}{\mathscr{B}\left(B^0 \to D^-\pi^+\right)}} \tag{12.7}$$

where $\tan(\theta_c) = 0.231\,01 \pm 0.000\,32$ is the tangent of the Cabibbo angle from Ref. [32]. Furthermore, for the branching fractions $\mathscr{B}\left(B^0 \to D_s^+\pi^-\right) = (2.16 \pm 0.26) \times 10^{-5}$ and $\mathscr{B}\left(B^0 \to D^-\pi^+\right) = (2.52 \pm 0.13) \times 10^{-3}$ the values reported in Ref. [18] are used. The ratio of decay constants $f_{D^+}/f_{D_s^+} = 1.173 \pm 0.003$ is taken from Refs. [103–105]. This results in $r = 0.0182 \pm 0.0012 \pm 0.0036$, where the second uncertainty is due to possible nonfactorisable SU(3)-breaking effects, which are assumed to be 20 % of the value of r. By also adding the known value of $\beta = (22.2 \pm 0.7)^\circ$ taken from Ref. [24], additional confidence intervals for γ can be determined.

The resulting confidence intervals are shown in Fig. 12.1 and 12.2. The numerical values are

$$\begin{aligned} |\sin(2\beta+\gamma)| &\in [0.77, 1.0]\,, \\ \gamma &\in [5, 86]^\circ \cup [185, 266]^\circ\,, \\ \delta &\in [-41, 41]^\circ \cup [140, 220]^\circ\,, \end{aligned}$$

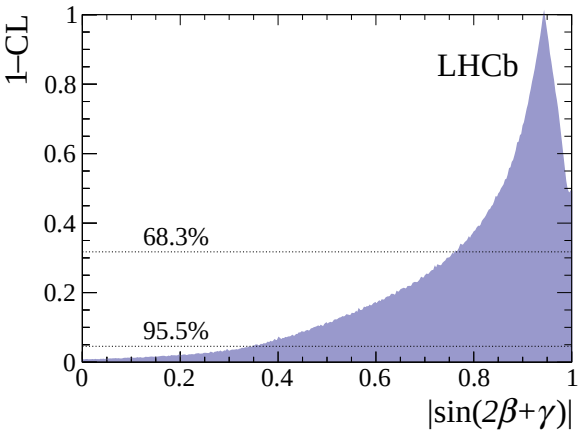

Fig. 12.1: Distribution of $1 - \mathrm{CL}$ for $|\sin(2\beta + \gamma)|$.

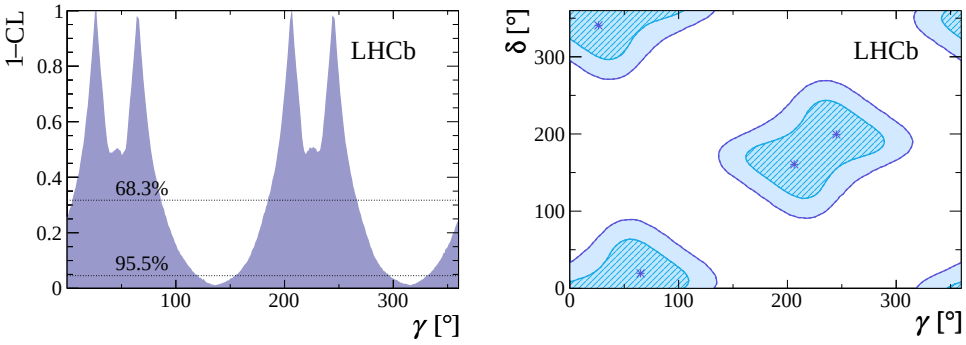

Fig. 12.2: Distribution of $1 - \mathrm{CL}$ for γ (left) and confidence regions for γ vs. δ (right). The regions hold the 39 % and 87 % CL, the points denote the obtained maxima.

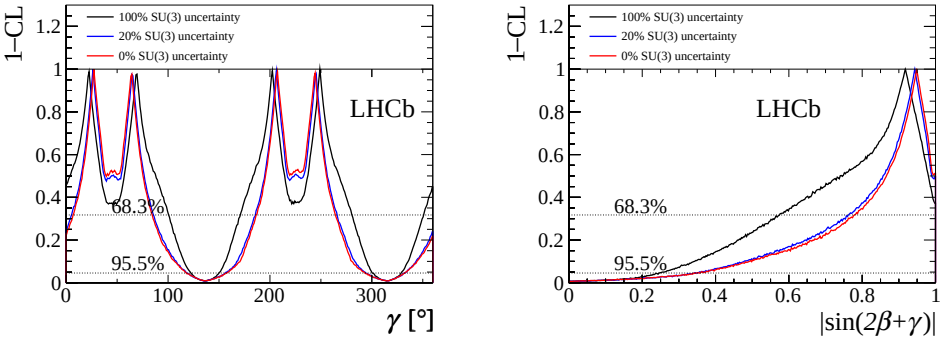

Fig. 12.3: Distribution of $1-\mathrm{CL}$ for γ (left) and $|\sin(2\beta+\gamma)|$ (right) for assumptions of 0 %, 20 % and 100 % for the SU(3) breaking uncertainty of r.

all at the 68 % CL. The $1-\mathrm{CL}$ interval for $|\sin(2\beta+\gamma)|$ is in agreement with, and more precise than, the previous measurements from the Belle and BaBar collaborations [98, 99], also the distribution of the $1-\mathrm{CL}$ shows the expected shape. The result for γ is compatible with all direct and indirect determinations presented in Sec. 4.1.3. However, the uncertainty is still quite large, so that on the one hand the precision of the measurement needs to be improved to make a conclusive statement about the agreement and on the other hand the contribution of this measurement to a combination of tree-level determinations of the angle γ is very small [33]. The obtained value for δ is compatible with the measured value in the time-dependent measurement of *CP* violation using $B_s^0 \to D_s^{\mp} K^{\pm}$ decays [106], which are related to $B^0 \to D^{\mp}\pi^{\pm}$ via the SU(3) symmetry [107]. The distributions of $1-\mathrm{CL}$ for γ and δ show the four-fold (two-fold) ambiguity for the range $[0, 360]^{\circ}$ ($[0, 180]^{\circ}$), which was discussed in Sec. 4.2. The uncertainties on r and β have a negligible impact on the confidence intervals.

As especially the assumption on SU(3) symmetry is highly unknown, the intervals are also determined for assumptions of 0 %, 20 % and 100 % for the SU(3) breaking uncertainty of r. These are representatively presented for γ and $|\sin(2\beta+\gamma)|$ in Fig. 12.3, which shows that also assuming a larger uncertainty on the SU(3) symmetry yields reasonable confidence intervals.

13 Conclusion and outlook

With the discovery of the Higgs boson in 2012 [4, 5], the SM was finally completed. However, it still fails to explain phenomena such as dark matter and dark energy or the observable matter-antimatter asymmetry in the universe, what clearly shows that there must be physics beyond the SM. While direct searches are limited by the available collision energies at accelerators, indirect searches are sensitive to NP effects, which exceed this energy threshold through higher-order contributions.

The measurement of the CKM angle γ is therefore interesting for different reasons. On the one hand, measurements in tree-level processes, which are not affected by higher-order contributions, can be compared with determinations using loop processes. On the other hand, the measurements of γ are an important part to probe the unitarity of the CKM matrix. The LHCb experiment, which is designed to measure processes containing b and c hadrons, showed an outstanding performance, recording high-quality data during the first LHC run period from 2010 to 2012. This is reflected in similar measurements of *CP* violation like in the golden mode $B^0 \to J/\psi K^0_S$ [31], where after two years of data taking already a similar precision was achieved compared to the previous measurements performed by the Belle and BaBar collaborations [108, 109].

One possibility to determine γ in a tree-level process is the time-dependent *CP* violation measurement in the decay $B^0 \to D^{\mp}\pi^{\pm}$. The analysis presented in this thesis was performed on a data set of proton-proton collisions recorded by the LHCb detector at centre-of-mass energies of 7 and 8 TeV, corresponding to an integrated luminosity of $3\,\text{fb}^{-1}$. The data sample contains 479 000 $B^0 \to D^{\mp}\pi^{\pm}$ candidates tagged by the combination of the OS or SS flavour-tagging algorithms. An unbinned maximum-likelihood fit to the decay-time, tags and finalstates yields the *CP* asymmetries

$$S_f = 0.058 \pm 0.020\,(\text{stat.}) \pm 0.011\,(\text{syst.})\,,$$
$$S_{\bar{f}} = 0.038 \pm 0.020\,(\text{stat.}) \pm 0.007\,(\text{syst.})\,,$$

with a correlation of the statistical and systematic uncertainties of 60 % and −41 %, respectively. This result is more precise and in agreement with previous determinations by the Belle and BaBar collaborations [98, 99]. Furthermore, even it is not yet a statistical evidence, the obtained values for S_f and $S_{\bar{f}}$ yield a significance of 2.7σ for *CP* violation according to Wilk's theorem [97].

To better compare this result with future measurements of the Belle II collaboration and the averaged values from the HFLAV collaboration, it is also transformed into a notation

that is less affected by the tag-side interference, an experimental effect arising due to the coherent $B^0\overline{B}^0$ production at the B factories Belle (II) and BaBar [100]. After such transformation the measured CP asymmetries are

$$a = -0.048 \pm 0.018\,(\text{stat.}) \pm 0.005\,(\text{syst.})\,,$$
$$c = 0.010 \pm 0.009\,(\text{stat.}) \pm 0.008\,(\text{syst.})\,,$$

where the statistical correlation is zero and the systematic correlation is $-46\,\%$.

Using the values for S_f and $S_{\bar{f}}$, confidence intervals are extracted for the CKM angle γ, the CKM quantity $\sin(2\beta+\gamma)$ and the *strong* phase difference δ. This is done by adding external input for the CKM angle β [24] and for the ratio r [32, 103–105], which is determined from the branching fraction of $B^0 \to D_s^+\pi^-$ assuming SU(3) symmetry. The obtained confidence intervals are

$$\gamma \in [5, 86]^\circ \cup [185, 266]^\circ\,,$$
$$|\sin(2\beta+\gamma)| \in [0.77, 1.0]\,,$$
$$\delta \in [-41, 41]^\circ \cup [140, 220]^\circ\,.$$

The result for γ is in agreement with all previous direct and indirect determinations, though the large uncertainties do not allow a conclusive statement yet. The confidence interval for $|\sin(2\beta+\gamma)|$ is more precise and in agreement with the previous determinations by the B factories [98, 99]. The determined value for δ can be compared to the result from the similar measurement in $B_s^0 \to D_s^\mp K^\pm$ decays yielding $\left(358^{+13}_{-14}\right)^\circ$ [106], also showing good agreement.

Furthermore, the analysed number of tagged signal candidates exceeds the respective number of signal candidates in the statistically largest time-dependent CP analysis at LHCb using $B^0 \to J/\psi K_S^0$ decays so far by about one order of magnitude. This successful measurement therefore shows that the recorded data is well understood and intrinsic asymmetries caused by *e.g.* the experimental setup are under control, so that this kind of analyses can be performed with the large number of signal candidates in the further run periods of the LHC. However, the systematic uncertainties of the CP violation measurement in the decay mode $B^0 \to D^\mp\pi^\pm$ are currently almost half as large as the statistical uncertainties. The leading systematic uncertainties are due to the uncertainty on B^0-oscillation frequency, potential fit biases on simulated events and the background subtraction. The systematic uncertainty due to uncertainty on Δm could be reduced in two ways: either the precision of the determination in the decay mode $B^0 \to D^{(*)-}\mu^+\nu_\mu$ [96] is improved using the data set recorded during Run II of the LHC, or the large number of $B^0 \to D^\mp\pi^\pm$ candidates is used directly to determine Δm and the CP asymmetries simultaneously. The two other uncertainties will need to be revisited: while the potential fit biases on simulated events will need to be investigated in greater depth to understand the exact source of this effect,

the uncertainty due to the background subtraction should be reduced with more data being available by examining the shapes of the contributing background components.

The largest competitor for LHCb in the sector of B mesons will probably be the Belle II experiment, aiming to start data taking in 2019. With an improved detector and a higher instantaneous luminosity compared to the previous Belle experiment, in total of $50\,\mathrm{ab}^{-1}$ [110] should be recorded, corresponding to 50 times the amount collected by the predecessor. To achieve the best sensitivity on CKM parameters like the angle γ, a joint effort of both collaborations will result in the best possible precision.

Currently, the confidence intervals for the CKM parameters γ and $|\sin(2\beta+\gamma)|$ are dominated by the uncertainties on S_f and $S_{\bar{f}}$ and the uncertainties on the external inputs β and r are negligible. However, assuming the same detector performance as achieved in Run I and only scaling the $B^0 \to D^{\mp}\pi^{\pm}$ yield, the expected statistical sensitivity for the CP asymmetries will drop to values of *e.g.* 0.005 for $50\,\mathrm{fb}^{-1}$. Yet, the estimation of the precision on γ and $|\sin(2\beta+\gamma)|$ is more challenging, since the precision of the external value of r will become the dominant source of systematic uncertainty. As the dominant uncertainty on r already comes from the calculations of nonfactorisable SU(3)-breaking effects, theoretical advancements are needed there.

Bibliography

[1] D0 collaboration, S. Abachi et al., *Search for high mass top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys. Rev. Lett. **74** (1995), pp. 2422–2426, arXiv: hep-ex/9411001 [hep-ex].

[2] CDF collaboration, F. Abe et al., *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995), pp. 2626–2631, arXiv: hep-ex/9503002 [hep-ex].

[3] DONUT collaboration, K. Kodama et al., *Observation of tau neutrino interactions*, Phys. Lett. **B504** (2001), pp. 218–224, arXiv: hep-ex/0012035 [hep-ex].

[4] 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), pp. 30–61, arXiv: 1207.7235 [hep-ex].

[5] 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), pp. 1–29, arXiv: 1207.7214 [hep-ex].

[6] E. Corbelli and P. Salucci, *The Extended Rotation Curve and the Dark Matter Halo of M33*, Mon. Not. Roy. Astron. Soc. **311** (2000), pp. 441–447, arXiv: astro-ph/9909252 [astro-ph].

[7] Supernova Cosmology Project, M. Kowalski et al., *Improved Cosmological Constraints from New, Old and Combined Supernova Datasets*, Astrophys. J. **686** (2008), pp. 749–778, arXiv: 0804.4142 [astro-ph].

[8] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967), pp. 32–35.

[9] Super-Kamiokande collaboration, H. Nishino et al., *Search for Proton Decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in a Large Water Cherenkov Detector*, Phys. Rev. Lett. **102** (2009) 141801, arXiv: 0903.0676 [hep-ex].

[10] E. W. Kolb and M. S. Turner, *The Early Universe*, Frontiers in physics, Avalon Publishing, 1994.

[11] C. S. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957), pp. 1413–1414.

[12] J. H. Christenson et al., *Evidence for the 2π decay of the K_2^0 meson*, Phys. Rev. Lett. **13** (1964), pp. 138–140.

[13] M. B. Gavela et al., *Standard model CP violation and baryon asymmetry*, Mod. Phys. Lett. **A9** (1994), pp. 795–810, arXiv: `hep-ph/9312215 [hep-ph]`.

[14] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973), pp. 652–657.

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

[16] D. H. Perkins, *Introduction to high-energy physics*, *4th ed.*, Cambridge Univ. Press, 2000.

[17] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Advanced book programm, Westview, 1995.

[18] Particle Data Group, M. Tanabashi et al., *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[19] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963), pp. 531–533.

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

[21] C. Jarlskog, *Commutator of the Quark Mass Matrices in the Standard Electroweak Model and a Measure of Maximal CP Nonconservation*, Phys. Rev. Lett. **55** (1985), pp. 1039–1042.

[22] G. C. Branco, L. Lavoura, and J. P. Silva, *CP violation*, Int. Ser. Monogr. Phys. Clarendon Press, 1999.

[23] I. I. Bigi and A. I. Sanda, *CP violation*, *2nd ed.*, Cambridge monographs on particle physics, nuclear physics and cosmology, Cambridge Univ. Press, 2009.

[24] Heavy Flavour Averaging Group, Y. Amhis et al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. **C77**.12 (2017) 895, arXiv: `1612.07233 [hep-ex]`, updated results and plots available at `https://hflav.web.cern.ch`.

[25] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, Comput. Phys. Commun. **210** (2017), pp. 103–123, arXiv: `1601.05437 [hep-ph]`.

[26] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes $K_L^0 \to \mu\overline{\mu}$, $K^+ \to \pi^+ \nu\overline{\nu}$ and $K^0 \leftrightarrow \overline{K}^0$*, Prog. Theor. Phys. **65** (1981) 297, [Erratum: Prog. Theor. Phys. **65** (1981) 1772].

[27] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970), pp. 1285–1292.

[28] 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 [hep-ex]`.

[29] BaBar collaboration, B. Aubert et al., *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, arXiv: hep-ex/0107013 [hep-ex].

[30] Belle collaboration, K. Abe et al., *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, arXiv: hep-ex/0107061 [hep-ex].

[31] LHCb collaboration, R. Aaij et al., *Measurement of CP violation in $B^0 \to J/\psi K^0_S$ decays*, Phys. Rev. Lett. **115** (2015) 3, p. 031601, arXiv: 1503.07089 [hep-ex].

[32] CKMfitter group, J. Charles et al., *Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ new physics*, Phys. Rev. **D91** (2015) 073007, arXiv: 1501.05013 [hep-ph], updated results and plots available at http://ckmfitter.in2p3.fr/.

[33] LHCb Collaboration, *Update of the LHCb combination of the CKM angle γ*, LHCb-CONF-2018-002 (2018).

[34] M. Gronau and D. Wyler, *On determining a weak phase from charged decay asymmetries*, Phys. Lett. **B265** (1991), pp. 172–176.

[35] M. Gronau and D. London, *How to determine all the angles of the unitarity triangle from $B^0 \to DK^0_S$ and $B^0_s \to D\phi$*, Phys. Lett. **B253** (1991), pp. 483–488.

[36] D. Atwood, I. Dunietz, and A. Soni, *Improved methods for observing CP violation in $B^\pm \to KD$ and measuring the CKM phase γ*, Phys. Rev. **D63** (2001) 036005, arXiv: hep-ph/0008090 [hep-ph].

[37] A. Giri et al., *Determining γ using $B^\pm \to DK^\pm$ with multibody D decays*, Phys. Rev. **D68** (2003) 054018, arXiv: hep-ph/0303187 [hep-ph].

[38] M. Gronau and D. London, *Isospin analysis of CP asymmetries in B decays*, Phys. Rev. Lett. **65** (1990), pp. 3381–3384.

[39] BaBar collaboration, B. Aubert et al., *Improved Measurement of $B^+ \to \rho^+\rho^0$ and Determination of the Quark-Mixing Phase Angle α*, Phys. Rev. Lett. **102** (2009) 141802, arXiv: 0901.3522 [hep-ex].

[40] Belle collaboration, P. Vanhoefer et al., *Study of $B^0 \to \rho^+\rho^-$ decays and implications for the CKM angle ϕ_2*, Phys. Rev. **D93** (2016) 032010, arXiv: 1510.01245 [hep-ex], [Erratum: Phys. Rev. **D94** (2016) 099903].

[41] UTfit collaboration, M. Bona et al., *The unitarity triangle fit in the standard model and hadronic parameters from lattice QCD: A reappraisal after the measurements of Δm_s and $BR(B \to \tau\nu_\tau)$*, JHEP **10** (2006) 081, arXiv: hep-ph/0606167 [hep-ph], updated results and plots available at http://www.utfit.org/.

[42] R. Fleischer, *New strategies to extract β and γ from $B_d \to \pi^+\pi^-$ and $B_s \to K^+K^-$*, Phys. Lett. **B459** (1999), pp. 306–320, arXiv: hep-ph/9903456 [hep-ph].

[43] M. Ciuchini et al., *Testing the Standard Model and searching for New Physics with $B^0 \to \pi\pi$ and $B_s^0 \to KK$ decays*, JHEP **10** (2012) 029, arXiv: 1205.4948 [hep-ph].

[44] R. Fleischer and R. Knegjens, *In Pursuit of New Physics With $B_s^0 \to K^+K^-$*, Eur. Phys. J. **C71** (2011) 1532, arXiv: 1011.1096 [hep-ph].

[45] R. Fleischer, *$B_{s,d} \to \pi\pi, \pi K, KK$: Status and Prospects*, Eur. Phys. J. **C52** (2007), pp. 267–281, arXiv: 0705.1121 [hep-ph].

[46] LHCb collaboration, R. Aaij et al., *Determination of γ and $-2\beta_s$ from charmless two-body decays of beauty mesons*, Phys. Lett. **B741** (2015), pp. 1–11, arXiv: 1408.4368 [hep-ex].

[47] Belle collaboration, A. Das et al., *Measurements of Branching Fractions for $B^0 \to D_s^+\pi^-$ and $\bar{B}^0 \to D_s^+K^-$*, Phys. Rev. **D82** (2010) 051103, arXiv: 1007.4619 [hep-ex].

[48] BaBar collaboration, B. Aubert et al., *Measurement of the Branching Fractions of the Rare Decays $B^0 \to D_s^{(*)+}\pi^-$, $B^0 \to D_s^{(*)+}\rho^-$, and $B^0 \to D_s^{(*)-}K^+$*, Phys. Rev. **D78** (2008) 032005, arXiv: 0803.4296 [hep-ex].

[49] J. Brod, *Electroweak effects in the extraction of the CKM angle γ from $B \to D\pi$ decays*, Phys. Lett. **B743** (2015), pp. 56–60, arXiv: 1412.3173 [hep-ph].

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

[51] LHCb collaboration, R. Aaij et al., *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv: 1412.6352 [hep-ex].

[52] C. Lefèvre, "The CERN accelerator complex.", Dec. 2008, url: http://cds.cern.ch/record/1260465.

[53] O. S. Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs, Geneva: CERN, 2004, url: https://cds.cern.ch/record/782076.

[54] CERN, *Accelerator Performance and Statistics*, June 2018, url: http://acc-stats.web.cern.ch/acc-stats/#lhc/super-table.

[55] LHCb collaboration, *$b\bar{b}$-production angle plots*, June 2018, url: https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[56] C. Ilgner et al., *The Beam Conditions Monitor of the LHCb Experiment*, 2010, arXiv: 1001.2487 [physics.ins-det].

[57] G. Barrand et al., *GAUDI - A software architecture and framework for building HEP data processing applications*, Comput. Phys. Commun. **140** (2001), pp. 45–55.

[58] R. Aaij et al., *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, arXiv: 1211.3055 [hep-ex].

[59] LHCb HLT project, J. Albrecht et al., *Performance of the LHCb High Level Trigger in 2012*, J. Phys. Conf. Ser. **513** (2014) 012001, arXiv: 1310.8544 [hep-ex].

[60] C. Parkes, T. Ruf, and T. Szumlak, *Reconstruction of Cluster Positions in the LHCb Velo*, tech. rep. LHCb-2007-151, Geneva: CERN, 2007.

[61] R. Van der Eijk, *Performance of the combined LHCb track reconstruction algorithms*, tech. rep. LHCb-2001-113, Geneva: CERN, 2001.

[62] M. Kucharczyk, P. Morawski, and M. Witek, *Primary Vertex Reconstruction at LHCb*, tech. rep. LHCb-PUB-2014-044, Geneva: CERN, 2014.

[63] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv: hep-ph/0603175 [hep-ph].

[64] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput.Phys.Commun. **178** (2008), pp. 852–867.

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

[66] D. Lange, *The EvtGen particle decay simulation package*, Nucl.Instrum.Meth. **A462** (2001), pp. 152–155.

[67] M. Asai, *Geant4-a simulation toolkit*, Trans.Amer.Nucl.Soc. **95** (2006) 757.

[68] J. Allison et al., *Geant4 developments and applications*, IEEE Trans.Nucl.Sci. **53** (2006) 270.

[69] G. Bohm and G. Zech, *Introduction to Statistics and Data Analysis for Physicists*, *3rd ed.*, Verlag Deutsches Elektronen-Synchrotron, 2017.

[70] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005), pp. 356–369, arXiv: physics/0402083 [physics.data-an].

[71] J. H. Friedman, *Greedy Function Approximation: A Gradient Boosting Machine*, Ann. of Statist. **29** (2001), pp. 1189–1232.

[72] 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.

[73] A. Hoecker et al., *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv: physics/0703039.

[74] I. Antcheva et al., *ROOT - A C++ framework for petabyte data storage, statistical analysis and visualization*, Comput. Phys. Commun. **180** (2009), pp. 2499–2512, arXiv: 1508.07749 [physics.data-an].

[75] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975), pp. 343–367.

[76] 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 [hep-ex].

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

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

[79] LHCb collaboration, R. Aaij et al., *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb*, JINST **11** (2016) P05010, arXiv: 1602.07252 [hep-ex].

[80] A. Agresti, *Categorical data analysis*, Wiley Series in Probability and Statistics, 2013.

[81] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning*, Springer Series in Statistics, 2001.

[82] W. Cheney and D. R. Kincaid, *Linear Algebra: Theory and Applications*, *2nd ed.*, Jones & Bartlett Learning Int., 2012.

[83] S. Weisberg, *Applied Linear Regression*, *4th ed.*, Wiley Series in Probability and Statistics, 2014.

[84] L. Anderlini et al., *The PIDCalib package*, LHCb-PUB-2016-021 (2016).

[85] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (Nov. 2005), pp. 566–575, arXiv: physics/0503191.

[86] 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 [physics.ins-det].

[87] F. Archilli et al., *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, arXiv: 1306.0249 [physics.ins-det].

[88] Y. Xie, *sFit: a method for background subtraction in maximum likelihood fit*, 2009, arXiv: 0905.0724 [physics.data-an].

[89] P. Koppenburg, *Statistical biases in measurements with multiple candidates*, 2017, arXiv: 1703.01128 [hep-ex].

[90] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014), pp. 150–155, arXiv: 1312.5000 [hep-ex].

[91] N. L. Johnson, *Systems of frequency curves generated by methods of translation*, Biometrika **36** (1949), pp. 149–176.

[92] J. H. Friedman, *Greedy Function Approximation: A Gradient Boosting Machine*, Annals of Statistics **29** (2000), pp. 1189–1232.

[93] D. W. Hosmer et al., *A comparison of goodness-of-fit tests for the logistic regression*, Statistics in Medicine **16** (1997), pp. 965–980.

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

[95] LHCb collaboration, R. Aaij et al., *Measurement of B^0, B_s^0, B^+ and Λ_b^0 production asymmetries in 7 and 8 TeV proton-proton collisions*, Phys. Lett. **B774** (2017), pp. 139–158, arXiv: 1703.08464 [hep-ex].

[96] LHCb collaboration, R. Aaij et al., *A precise measurement of the B^0 meson oscillation frequency*, Eur. Phys. J. **C76**.7 (2016) 412, arXiv: 1604.03475 [hep-ex].

[97] S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Ann. Math. Statst. **9** (1938), pp. 60–62.

[98] Belle collaboration, F. J. Ronga et al., *Measurements of CP violation in $B^0 \to D^{*-}\pi^+$ and $B^0 \to D^-\pi^+$ decays*, Phys. Rev. **D73** (2006) 092003, arXiv: hep-ex/0604013 [hep-ex].

[99] BaBar collaboration, B. Aubert et al., *Measurement of time-dependent CP asymmetries in $B^0 \to D^{*\pm}\pi^\mp$ and $B^0 \to D^\pm\rho^\mp$ decays*, Phys. Rev. **D73** (2006) 111101, arXiv: hep-ex/0602049 [hep-ex].

[100] O. Long et al., *Impact of tag-side interference on time-dependent CP asymmetry measurements using coherent $B^0\overline{B}^0$ pairs*, Phys. Rev. **D68** (2003) 034010, arXiv: hep-ex/0303030 [hep-ex].

[101] LHCb collaboration, R. Aaij et al., *Measurement of the CKM angle γ from a combination of LHCb results*, JHEP **12** (2016) 087, arXiv: 1611.03076 [hep-ex].

[102] S. Bodhisattva, M. Walker, and M. Woodroofe, *On the Unified Method with Nuisance Parameters*, Statist. Sinica **19** (2009), pp. 301–314.

[103] S. Aoki et al., *Review of lattice results concerning low-energy particle physics*, Eur. Phys. J. **C77**.2 (2017) 112, arXiv: 1607.00299 [hep-lat].

[104] Fermilab Lattice, MILC, A. Bazavov et al., *Charmed and light pseudoscalar meson decay constants from four-flavor lattice QCD with physical light quarks*, Phys. Rev. **D90**.7 (2014) 074509, arXiv: 1407.3772 [hep-lat].

[105] N. Carrasco et al., *Leptonic decay constants f_K, f_D and f_{D_s} with $N_f = 2+1+1$ twisted-mass lattice QCD*, Phys. Rev. **D91**.5 (2015) 054507, arXiv: 1411.7908 [hep-lat].

[106] LHCb collaboration, R. Aaij et al., *Measurement of CP asymmetry in $B_s^0 \to D_s^\mp K^\pm$ decays*, JHEP **03** (2018) 059, arXiv: 1712.07428 [hep-ex].

[107] R. Fleischer, *New Strategies to Obtain Insights into CP violation through* $B_s^0 \to D_s^{\pm} K^{\mp}$, $D_s^{*\pm} K^{\mp}$, ... *and* $d \to D^{\pm}\pi^{\mp}$, $D^{*\pm}\pi^{\mp}$, ... *Decays*, Nucl. Phys. **B671** (2003), pp. 459–482, arXiv: hep-ph/0304027 [hep-ph].

[108] BaBar collaboration, B. Aubert et al., *Measurement of Time-Dependent CP Asymmetry in* $B^0 \to c\bar{c}K^{(*)0}$ *Decays*, Phys. Rev. **D79** (2009) 072009, arXiv: 0902.1708 [hep-ex].

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

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

Acknowledgements

Zuallererst möchte ich mich bei meinem Doktorvater Herrn Professor Spaan bedanken. Nachdem er mich 2012 für die Bachelorarbeit am Lehrstuhl willkommen hieß, bin ich 2013 zur Masterarbeit zurückgekehrt und schließe nun meine Promotion 6 Jahre später ab. Während dieser Zeit, haben Sie mich in allen Aufgaben unterstützt und mir einige Konferenzbesuche, sowie diverse Aufenthalte am CERN ermöglicht, die ich sicher nicht vergessen werde.

Weiterhin möchte ich Herrn Professor Kröninger danken, dass er trotz einiger anderer Dissertationen, zugestimmt hat, auch Zeit als Zweitgutachter für diese Arbeit zu finden.

A huge thank-you goes to my analysis colleagues from Dortmund and Lausanne. Together we made this analysis possible, after some quite painful times, *e.g.* before rushing for CKM. A special thank goes to Julian, who supported me very strongly made it very easy for me to join the various LHCb working group and was always there to answer questions and give advice, although I surely was not the only PhD student regularly asking him. Also I want to say thank you to Vincenzo, who fighted with me though quite some barriers for this analysis. Finally, not to forget Conor and Mirco: It was a pleasure to work with you, and I learned many things in fruitful discussions!

Ein Dank geht weiterhin an alle Büro- und Arbeitsgruppenkollegen, sei es aus der lokalen "B to open charm" Arbeitsgruppe mit Frank, Philipp (vielen Dank vor allem auch an das teilweise prompte und intensive Korrekturlesen dieser Arbeit), Margarete und Ulrich (der ebenfalls mit mir am Flavour Tagging gearbeitet hat), oder aus der lokalen Flavour Tagging Arbeitsgruppe mit Kevin. Die vielen Diskussionen haben mir immer wieder hilfreiche Anstöße gegeben.

Zur etwa gleichen Zeit haben wir zumindest die Promotion, teilweise sogar die Masterarbeit am Lehrstuhl E5 begonnen und dabei auch des öfteren mal nicht über Physik geredet und so ab und an den Kopf freibekommen: Vielen Dank an Vanessa, Timon, Moritz und Janine.

Außerdem gilt mein Dank unserer Sekretärin Frau Stickel, die mir in so manch einer bürokratischen Angelegenheit weitergeholfen hat und ohne die manch eine Abrechnung oder Dienstreise nicht so einfach abgelaufen wäre.

Schlußendlich möchte ich noch allen weiteren Lehrstuhlmitgliedern bei E5 danken für die gute, gemeinschaftliche Atmosphäre!

Abseits der Universität möchte ich außerdem meiner Familie und dort zuallererst meinen Eltern danken, die mir das Physikstudium, und ebenfalls die anschließende Promotion

durch ihre unentwegte Unterstützung ermöglicht haben. Ebenfalls danke ich meinen beiden Brüdern, die mir immer beigestanden haben.

Zu guter Letzt geht mein Dank außerdem an Gina: Seit wir uns kennengelernt haben, hast du mich besonders immer wenn es eng und stressig wurde unterstützt, mir Aufgaben abgenommen, mich motiviert und mein Blickfeld erweitert. Nicht zu vergessen, als erste diese Arbeit auf sprachliche Fehler zu durchsuchen, und somit allen weiteren Korrekturlesern meine gröbsten Fehltritte vorweggenommen zu haben.