technische universität
dortmund

Measurement of the $B\overline{B}$ mixing frequency Δm_d and tagging calibration measurements with data from the LHCb Experiment

Messung der $B\overline{B}$ Mischungsfrequenz Δm_d und Kalibrierung des Flavour-Tagging mit Daten des LHCb Experiments

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

Vorgelegt von
Tobias Thomas Brambach

am
15. Mai 2013

Lehrstuhl für Experimentelle Physik V
Fakultät für Physik
Technische Universität Dortmund

Betreuer dieser Arbeit ist Prof. Dr. Bernhard Spaan.
Zweitgutachter ist PD Dr. Reiner Klingenberg.

Zusammenfassung

Am LHCb Experiment am CERN bei Genf wurden im Jahr 2011 Daten mit einer integrierten Luminosität von $1\,\mathrm{fb}^{-1}$ gesammelt. Auf Basis dieses Datensatzes wird die Mischungsfrequenz Δm_d des neutralen B^0 Mesons in den Zerfallskanälen $B^0 \to D^- \pi^+$ und $B^0 \to J/\psi\, K^{*0}$ gemessen.

$$\Delta m_d = 0{,}5156 \pm 0{,}0051\,(\mathrm{stat.}) \pm 0{,}0033\,(\mathrm{syst.})\,\mathrm{ps}^{-1},$$

dies entspricht der weltbesten Einzelmessung dieses Parameters. Die Messung wurde in *Physics Letters B* veröffentlicht [1].

Im Rahmen dieser Messungen ist außerdem eine Überprüfung des Flavour Tagging möglich. Um systematische Unsicherheiten in Messungen von *CP*-Verletzung zu untersuchen, werden verschiedene Tests durchgeführt.

Abstract

In 2011 the LHCb experiment at CERN (Geneva) collected data with an integrated luminosity of $1\,\mathrm{fb}^{-1}$. Using this dataset, the mixing frequency Δm_d of the neutral B^0 mesons is measured using the decay channels $B^0 \to D^- \pi^+$ and $B^0 \to J/\psi\, K^{*0}$.

$$\Delta m_d = 0.5156 \pm 0.0051\,(\mathrm{stat.}) \pm 0.0033\,(\mathrm{syst.})\,\mathrm{ps}^{-1},$$

corresponding to the world's best individual measurement. The measurement was published in *Physics Letters B* [1].

Additionally, the flavour tagging algorithms can be checked using a similar measurement. In order to examine systematic uncertainties on measurements of *CP* violation, different crosschecks have been performed.

Contents

1 Introduction 1

2 The Standard Model 3
- 2.1 Discrete symmetries . . . 4
- 2.2 Elementary particles and forces . . . 4
  - 2.2.1 Quarks and leptons . . . 5
  - 2.2.2 Forces and gauge bosons . . . 5
  - 2.2.3 Fermion masses . . . 6
  - 2.2.4 The weak interaction of quarks . . . 7
- 2.3 The CKM matrix . . . 8
  - 2.3.1 B^0 mixing . . . 9
  - 2.3.2 Constraining the CKM matrix . . . 13
  - 2.3.3 *CP* violation . . . 13
  - 2.3.4 New physics . . . 16
- 2.4 Particle production in *pp*-collisions . . . 17

3 The experimental setup 19
- 3.1 Tracking system . . . 21
  - 3.1.1 Vertex locator . . . 21
  - 3.1.2 Dipole magnet . . . 21
  - 3.1.3 Silicon trackers . . . 23
  - 3.1.4 Outer tracker . . . 23
- 3.2 Particle identification system . . . 24
- 3.3 Calorimeters . . . 25
- 3.4 Muon system . . . 26
- 3.5 Trigger system . . . 27
  - 3.5.1 Level 0 trigger . . . 28
  - 3.5.2 High level trigger . . . 28
- 3.6 Software . . . 29

4 An introduction to flavour tagging 31
- 4.1 Opposite side tagging . . . 32
- 4.2 Same side tagging . . . 33
- 4.3 Tagging efficiency . . . 34
- 4.4 Estimation of wrong tag probabilities . . . 34
- 4.5 Methods for tagging calibration . . . 36

5 Self tagging decay channels 37
5.1 The charmonium mode $B^0 \to J/\psi K^{*0}$ . . . 37
5.2 The hadronic mode $B^0 \to D^- \pi^+$ . . . 38

6 Datasets and selection 41
6.1 Observables . . . 41
6.2 Trigger and selection for $B^0 \to D^- \pi^+$ . . . 42
6.3 Trigger and selection for $B^0 \to J/\psi K^*$ . . . 43

7 Fitting model for tagging and oscillation measurements 47
7.1 Parametrisation for the $B^0 \to D^- \pi^+$ channel . . . 49
7.2 Parametrisation for the $B^0 \to J/\psi K^{*0}$-channel . . . 50

8 Physics background 53
8.1 Background in $B^0 \to D^- \pi^+$ . . . 53
8.2 Background in $B^0 \to J/\psi K^{*0}$ . . . 56

9 Effects on the reconstructed decay time 57
9.1 Decay time resolution . . . 57
9.2 Decay time acceptance . . . 58
9.2.1 Efficiency for candidates with small decay times . . . 58
9.2.2 Efficiency for candidates with large decay times . . . 59

10 Tagging calibration 61
10.1 Fitting strategy . . . 61
10.2 OST calibration . . . 63
10.2.1 Results for $B^0 \to D^- \pi^+$ decays . . . 64
10.2.2 Results for $B^0 \to J/\psi K^{*0}$ decays . . . 64
10.3 SSπ calibration . . . 65
10.3.1 Results for $B^0 \to D^- \pi^+$ decays . . . 65
10.3.2 Results for $B^0 \to J/\psi K^{*0}$ decays . . . 67
10.4 Combination of taggers . . . 67
10.5 Tagging related systematic uncertainties on measurements of *CPV* . 67
10.6 Summary . . . 71

11 Measurement of Δm_d 75
11.1 Fitting strategy . . . 75
11.2 Fit results . . . 76
11.3 Determination of systematic uncertainties . . . 84
11.3.1 Fit model . . . 84
11.3.2 Tagging . . . 85
11.3.3 Length scale of the VELO . . . 85
11.3.4 Momentum calibration . . . 86
11.3.5 Decay time acceptance . . . 86
11.3.6 Decay time resolution . . . 87

11.3.7 Combination of systematic uncertainties 88
11.4 Summary 88

12 Conclusion 93

Acknowledgements 95

Danksagung 97

Bibliography 99

1 Introduction

The Large Hadron Collider (LHC) project at the European Organisation for Nuclear Research (CERN) includes four major experiments that are dedicated to the search for the Higgs boson, as well as to searches for Dark Matter candidates and exotic particles, studies on the quark gluon plasma, and precision measurements in flavour physics. These precision measurements are supposed to test the *Standard Model* (SM) of particle physics and they are sensitive to contributions from particles and forces that have not been discovered, yet.

The Large Hadron Collider beauty (LHCb) experiment is designed to perform precision measurements on decays of mesons that contain b or c quarks. These are measurements of the SM parameters and the violation of *CP* invariance, and the search for contributions of physics beyond the SM.

The mixing frequency Δm_d of oscillations between B^0 and $\overline{B}^0$ mesons (also referred to as the mass difference between the mass eigenstates of the B^0 mesons) has been precisely measured at LEP [2], the Tevatron [3, 4], and the B factories [5, 6]. The current world average is $\Delta m_d = 0.507 \pm 0.004\,\mathrm{ps}^{-1}$ [7], whilst the best single measurement prior to this thesis is by the Belle experiment, $\Delta m_d = 0.511 \pm 0.005\,(\mathrm{stat.}) \pm 0.006\,(\mathrm{syst.})\,\mathrm{ps}^{-1}$ [6]. This thesis presents the results of the first competitive measurement of this parameter by LHCb [1] with decays of $B^0 \to D^- \pi^+$ and $B^0 \to J/\psi\, K^{*0}$. Throughout this thesis, the convention $\hbar = c = 1$ is used for all units.

The measurement is motivated by the possibility of the calculation of constraints on the apex of the CKM triangle [8, 9] (cf. Fig. 2.1) from lattice QCD calculations [10, 11]. This constraint can be calculated using the ratio $^{\Delta m_d}/_{\Delta m_s}$ and its uncertainty is thus dominated by the larger relative uncertaintiy of either of the two measurements. Together with the very precise measurement of Δm_s [12] a more precise knowledge of Δm_d serves as an important test of the SM [13, 14].

As for all time dependent analyses of neutral B mesons, the knowledge of the production state of the meson is of great importance. The determination of the flavour of the production state (whether it contains a b or a $\bar{b}$ quark) is called *flavour tagging*. The probability of a tag being wrong (referred to as wrong tag probability) has to be used as an input for this type of analyses. As it directly contributes to the measured asymmetry amplitudes, it is important to be able to predict the wrong tag probability accurately for each of the candidates that are used in the corresponding measurements.

The decay channels used in the measurement of Δm_d provide self tagging final states. That means, that from the charge of the final state particles the flavour of the decaying B meson can be identified. Thus, these decay channels can also be used to test the validity of the flavour tagging calibration that is performed using decays of charged B mesons [15, 16].

This thesis describes the measurement of the parameter Δm_d and the necessary theoretical and experimental prerequisites. Additionally, the calibration of the wrong tag probabilities is verified and tagging induced systematic uncertainties in measurements of *CPV* are investigated.

2 The Standard Model

About 50 years ago physicists started to describe the theory of subatomic particles. Since then the understanding of the elementary particles building up the known matter, and the fundamental forces that describe the interaction between those particles is summarised in the so called Standard Model of Particle Physics (SM).

For many years now, there has been an equilibrium between theory–describing experimental results and postulating new particles and forces–and experiments–testing theories and finding evidence for new facets of particle physics that might give input for new theories.

Recently, many experimental results gave proof for the SM, like the discovery of the neutral and charged currents of the weak force [17], the measurement of *CP*-violation in the Kaon sector [18], the CKM-matrix [19], and many more precision tests at the large collider experiments around the world. All these results and theoretical calculations lead to today's precise picture of particle physics. However, there are quite some aspects, that are not included in the SM and lead to puzzling questions. Galaxy movements for example, that cannot be explained by the visible matter in the galaxies, are leading to the postulation of dark matter, a type of matter that is interacting only gravitationally, and probably weakly, for which no candidates have been found, yet. Or the long sought Higgs particle [20], that was a still missing feature of the SM. It is necessary for all quarks, charged leptons, and the W and Z bosons to gather mass. For this particle a candidate has been discovered at the LHC [21, 22] in the summer of 2012. In newer measurements [23, 24, 25] its properties are found to be consistent with a SM Higgs boson.

Even though many aspects of the SM are well tested, it is the small number of mysteries that motivates more precise measurements. Machines like the Large Hadron Collider (LHC) at Geneva and the four major experiments ALICE, ATLAS, CMS, and LHCb are build to unravel the remaining secrets of the SM and to go beyond and find proof for new physics.

The following summary of SM physics relevant for this thesis is largely inspired by the lectures of G. Hiller[1] and the works of Branco et al. [26] and Bigi et al. [27]. The formalism for *CPV* in B^0 decays is kept close to a previous work on this topic by V. Gligorov [28].

[1] The script from the lecture was edited by D. van Dyk

2.1 Discrete symmetries

In particle physics there are three important discrete symmetry operations that are used in the description of the SM.

- The first is the reversal of time $\mathcal{T}$. This operation changes the direction of time, which would have no effect on movements in classical mechanics. Though, it is known from thermodynamics, that irreversible processes exist in physics.
- Secondly, the parity transformation $\mathcal{P}$ is a simple inversion in the spatial origin, meaning all spatial coordinates change their sign. Naively, it is expected, that physics should be invariant under such a transformation. However, there are interactions in particle physics that are not invariant under the $\mathcal{P}$ transformation.
- The charge conjugation $\mathcal{C}$ changes the charge of all particles. Electrodynamics are invariant under such a transformation.

The combination of the $\mathcal{C}$ and $\mathcal{P}$ operations is very important, especially for the physics program of LHCb. When a particle is transformed under $\mathcal{CP}$ it changes its state from matter to antimatter or vice versa. As described in the previous chapter, *CP* invariance is violated in particle physics. One of the most puzzling questions of modern physics is, if there is a source of *CP* violation (*CPV*) beyond the SM. This would lead to a possible explanation to the fact, that there is only one type of matter in the universe. This is curious, because matter and antimatter should emerge in equal parts from energy and they most probably did so in the big bang. But shortly after that one type of matter vanished nearly completely. So matter must be favoured over antimatter by physics.

Today it is estimated that in 10^9 parts of matter and antimatter only one part of matter survived the annihilation. It is expected that all visible matter consists of these remnants. *CPV* in the SM is not big enough to explain such a large asymmetry. Thus, experiments like LHCb search for contributions to *CPV* that come from physics beyond the SM.

2.2 Elementary particles and forces

The SM as we know it today consists of two different basic classes of particles: fermions, from which all the visible matter (and quite some more) is build up and bosons, carrying the forces between all particles and giving them the ability to interact with each other.

All elementary fermions have an intrinsic angular momentum–called spin–of $\pm 1/2$. The spin of bosons on the other hand is an integer.

2.2.1 Quarks and leptons

All elementary fermions may be divided in two groups of particles: quarks and leptons. The quarks consist of two subgroups, the so-called up-type quarks (up u, charm c, top t), which carry an electrical charge of $2/3$ and the down-type quarks (down d, strange s, bottom b) carrying an electrical charge of $-1/3$. All quarks have a corresponding anti-particle ($\bar{u}, \bar{d}, \bar{c}, \bar{s}, \bar{t}, \bar{b}$) with opposite charge and spin. The masses of the quarks are aligned with their families with the u and d quarks being the lightest and the t (and b) being the heaviest ones.

Leptons are slightly different from quarks. Firstly, because there are two completely different types: neutrinos and charged leptons. Neutrinos do not carry an electrical charge and therefore are only weakly interacting. The charged leptons have an electrical charge of -1 and do additionally interact through the electromagnetic force. As for the quarks there are also three families of leptons, where a charged lepton and one neutrino form a weak isospin doublet. These are the electron e and electron neutrino ν_e, the muon μ and its neutrino ν_μ, and the tau τ with the tau neutrino ν_τ.

Though it might seem arbitrary that there are three families of elementary fermions, measurements of the Z^0 boson mass show, that three is the only number of fermion families fitting the SM parameters [29].

2.2.2 Forces and gauge bosons

The forces in the SM describe the interaction of particles. These interactions can either be elastic or inelastic, where the latter include the production and decay of heavy and/or composite particles. Three different forces are known in the SM. They are described as local gauge symmetries. From invariances under gauge transformations the SM predicts gauge bosons, the mediators of the forces. The attributes of these bosons lead to very different properties of the corresponding forces.

The strong force is described by a $SU(3)$ symmetry group. From eight generators of the group follow eight massless bosons the so called gluons. The charge of the strong force is called colour and every free particle has to be neutral under that charge. As gluons carry colour themselves, they are able to change the colour of quarks as well as to interact with other gluons. This self-interaction is a unique property of the strong force and is a reason for its strength and thus for its name. Particles that decay governed by the strong force live shorter than approximately 10^{-24} s.

The electromagnetic (EM) force is known the longest and has a simple structure. From the $U(1)$ symmetry follows one massless gauge boson, the photon. The photon is uncharged and interacts only with particles that carry an electrical charge. The range of the electromagnetic force is infinite, but its strength is much smaller than

that of the strong force. Electromagnetic decaying particles have lifetimes of about 10^{-18} s.

The weak force has three massive gauge bosons the $W^{\pm}$ and the Z^0. Due to their mass the range of the weak interaction is very limited. The force is even weaker than the EM force and particles that decay weakly have long lifetimes of up to about 10^{-6} s. To properly include the weak force into the SM, it has to be joined with the electromagnetic force leading to the electroweak unification.

In the SM the EM and the weak force are combined in a $SU(2) \times U(1)$ symmetry that is broken by the Higgs field. The advantage of this compared to the two separated forces is that the three gauge bosons of the weak force gain mass from the Higgs field. Fermion masses can be described using Yukawa couplings to the Higgs field. Additionally, the theory describes that right handed neutrinos (and left handed anti neutrinos) do not couple to the weak interaction.

2.2.3 Fermion masses

The 12 elementary fermions are described as left handed doublets and right handed singulets. The doublets are chosen as follows:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L; \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L, \tag{2.1}$$

corresponding to the three uncharged neutrinos, the three charged leptons and the six quarks as introduced in Sec. 2.2.1. In these doublets, the third component of the weak isospin $T^3 = +1/2$ for the upper components, and $T^3 = -1/2$ for the lower ones. The singulets under $SU(2)$ are given by $e_r, \mu_R, \tau_R, u_R, d_R, c_R, s_R, t_R, b_R$, with $T^3 = 0$. As neutrinos do not carry an electrical charge, it follows that the neutrino singulets do not interact in the electroweak theory.

With the quark doublet $\bar{Q}_L$ containing u_L and d_L, the Yukawa couplings can be written as

$$\mathcal{L}_{\text{Yukawa}} = \bar{Q}_L \phi Y_d d_R + \bar{Q}_L \phi^c Y_u u_R + h.c., \tag{2.2}$$

with the Higgs field ϕ (and the charge conjugated field ϕ_c)

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} v + h^0(x) \\ 0 \end{pmatrix}, \; \phi^c(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h^0(x) \end{pmatrix}, \tag{2.3}$$

after spontaneous symmetry breaking. Replacing $1/\sqrt{2} v Y_{u,d}$ by the mass $M_{u,d}$, that yields mass terms like in QED and QCD:

$$\mathcal{L}_{\text{Yukawa}} = \bar{d}_L M_d d_R + \bar{u}_L M_u u_R + h.c. \tag{2.4}$$

This example for one generation of quarks can be easily expanded for all three known generations. Therefore, the Yukawa constants $Y_{u,d}$ and the masses $M_{u,d}$ are replaced by 3×3 matrices $\hat{M}_{ij}$ and $\hat{Y}_{ij}$, and Q_L is replaced by

$$Q_{L_i} = \begin{pmatrix} u_{L_i} \\ d_{L_i} \end{pmatrix} = \begin{pmatrix} u_L & c_L & t_L \\ d_L & s_L & b_L \end{pmatrix}. \tag{2.5}$$

The quark singulets are thus $u_{R_i} = (u_R, c_R, t_R)^T$ and $d_{R_i} = (d_R, s_R, b_R)^T$. The mass matrix $\hat{M}_{ij}$ is not necessarily diagonal, so Q_{L_i}, d_{R_i}, and u_{R_i} are gauge eigenstates, but no mass eigenstates that can be observed in experiments. The latter can be derived using four unitary matrices $V_u, U_u, V_d,$ and V_u to diagonalise the mass matrix $\hat{M}_{ij}$. The mass terms then write as

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= \bar{d}_{L_i} V_d^\dagger \cdot V_d \hat{M}_{ij} U_d^\dagger \cdot U_d d_R + \bar{u}_{L_i} V_d^\dagger \cdot V_d \hat{M}_{ij} U_d^\dagger \cdot U_d u_R + h.c. && (2.6) \\ &= d'_{L_i} \text{diag}(m_d, m_s, m_b) d'_{R_i} + u'_{L_i} \text{diag}(m_u, m_c, m_t) u'_{R_i} + h.c., && (2.7) \end{aligned}$$

with the mass eigenstates u'_{LR_i} and d'_{LR_i}.

2.2.4 The weak interaction of quarks

In the electroweak theory, so called neutral currents (NC) are mediated by the photon and the Z-boson, whereas charged currents (CC) occur with an exchange of W bosons. The part of the Lagrangian that describes the neutral currents with respect to the gauge eigenstates can be transformed to the mass eigenstates:

$$\begin{aligned} \mathcal{L}_{\text{NC}} &= \bar{u}_{L_i} \gamma_\mu Z^\mu u_{L_i} = \bar{u}_{L_i} V_u^\dagger V_u \gamma_\mu Z^\mu V_u^\dagger V_u u_{L_i} && (2.8) \\ &= \bar{u}'_{L_i} \gamma_\mu Z^\mu u'_{L_i}. && (2.9) \end{aligned}$$

thus, diagonalising the mass matrix has no effect to NCs. The same can be done for CCs, yielding

$$\begin{aligned} \mathcal{L}_{\text{CC}} &= W_\mu^+ \bar{u}_{L_i} \gamma^\mu d_{L_i} && (2.10) \\ &= W_\mu^+ \bar{u}_{L_i} V_u^\dagger V_u \gamma^\mu V_d^\dagger V_d d_{L_i} && (2.11) \\ &= W_\mu^+ \bar{u}'_{L_i} V_u V_d^\dagger \gamma^\mu d'_{L_i}. && (2.12) \end{aligned}$$

The term $V_u V_d^\dagger$ can be identified as the Cabibbo-Kobayashi-Maskawa (CKM) matrix that describes the transition between different generations of quarks in weak interactions.

2.3 The CKM matrix

The Cabibbo-Kobayashi-Maskawa (CKM) matrix [19] generalises the weak isospin symmetry of u- and d-quarks [30] that is broken by the quark masses in the second and third generation. It additionally allows translations between quarks of different generations. the CKM matrix is necessary to introduce CP-violation (CPV) to the SM. It is important to describe the branching ratios of weakly decaying mesons and measurements of its unitarity are a fundamental test of the consistency of the SM. Furthermore, through the CKM matrix, the mixing of electrically uncharged mesons like kaons, B mesons, and D mesons is possible.

The CKM matrix has four parameters. Three of them are rotation angles, where the fourth is a complex phase. The complex phase is the only source of CPV in the SM. The CKM-matrix can be parametrised in orders of the sine of the Cabibbo angle $\lambda_{\mathrm{C}} = \sin\theta_c \approx 0.23$. The Wolfenstein parametrisation [31] till the order $\mathcal{O}(\lambda_{\mathrm{C}}^3)$ is given by

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 1-\lambda_{\mathrm{C}}^2 & \lambda_{\mathrm{C}} & A\lambda_{\mathrm{C}}^3(\rho - \mathrm{i}\eta) \\ -\lambda_{\mathrm{C}} & 1-\lambda_{\mathrm{C}}^2 & A\lambda_{\mathrm{C}}^2 \\ A\lambda_{\mathrm{C}}^3(1-\rho-\mathrm{i}\eta) & -A\lambda_{\mathrm{C}}^2 & 1 \end{pmatrix}. \quad (2.13)$$

From the unitarity of the CKM-matrix follows for $i, j, k \in 1, 2, 3$

$$\sum_k V_{ki} V_{kj}^* = \delta_{ij} \text{ and } \sum_k V_{ik} V_{jk}^* = \delta_{ij}. \quad (2.14)$$

One of the most commonly used unitarity relations is

$$V_{ud} V_{ub}^* + V_{cd} V_{cb}^* + V_{td} V_{tb}^* = 0, \quad (2.15)$$

which describes the so called unitarity triangle (c.f. Fig. 2.1) on a complex plane. This triangle is one of two possible non-degenerate triangles, as only V_{td} and V_{ub} include large imaginary parts. It is used to visualise many different measurements of parameters of the CKM matrix. Any inconsistency in the apex of the triangle would be a hint for new physics.

The measurement of the parameters of the CKM-triangles, and thus of the CKM matrix, is one of the main physics cases of modern particle physics. Up to now, the SM predictions have been reproduced by many experiments to a high precision. Evidence for physics beyond the SM (BSM) is sought for in rare decays of heavy mesons and the size of CPV that is expected to exceed SM predictions under special circumstances.

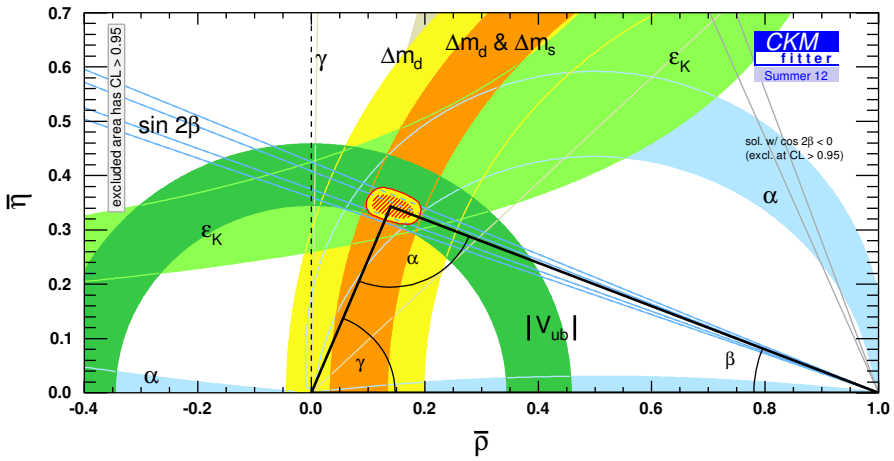


Fig. 2.1: Measured constraints on the CKM matrix including measurements of Δm_d and $\sin 2\beta$. [32]

2.3.1 B^0 mixing

Oscillations of neutral mesons have first been predicted in 1955 by Gell-Mann and Pais [33]. The observations of Ledermann et al. [34] then proved the existence of two weak eigenstates of the neutral kaon, and thus lead to the postulation of the CKM matrix that predicts similar oscillations also for B^0, B_s^0, and D^0 mesons. Oscillations of neutral mesons are allowed via the exchange of charged currents between the two quarks in a box diagramm (c.f. Fig. 2.2). This thesis focusses on oscillations of neutral mesons containing a b-quark.

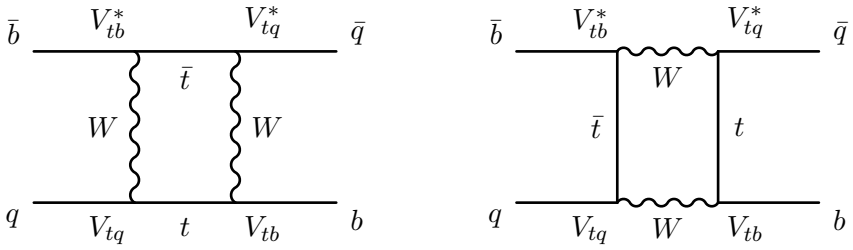


Fig. 2.2: Feynman diagrams for the mixing of B mesons.

In the Wigner-Weisskopf approximation [35], the time evolution of a beam of neutral mesons B^0 and $\overline{B}^0$ decaying at the time t can be described by

$$|\psi(t)\rangle = \begin{pmatrix} |B^0(t)\rangle \\ |\overline{B}^0(t)\rangle \end{pmatrix}. \tag{2.16}$$

The wave-function $|\psi\rangle$ then propagates according to

$$\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}|B^0(t)\rangle\\|\overline{B}^0(t)\rangle\end{pmatrix} = \hat{R}\begin{pmatrix}|B^0(t)\rangle\\|\overline{B}^0(t)\rangle\end{pmatrix}, \tag{2.17}$$

with the non hermitian matrix $\hat{R} = \hat{M} - {}^1\!/\!_2\hat{\Gamma}$, where $\hat{M}$ and $\hat{\Gamma}$ are hermitian. Eq. (2.17) can be solved using the similarity transformation

$$\hat{X}^{-1}\hat{R}\hat{X} = \mathrm{diag}(\mu_L, \mu_H) = \mathcal{H}, \tag{2.18}$$

with the eigenvalues μ_L and μ_H of $\hat{R}$. The corresponding eigenstates of $\hat{R}$ are called $(|B_L\rangle, 0)^T$ and $(0, |B_H\rangle)^T$, and thus

$$\begin{pmatrix}|B_L\rangle\\|B_H\rangle\end{pmatrix} = \hat{X}^{-1}\begin{pmatrix}|B^0\rangle\\|\overline{B}^0\rangle\end{pmatrix}. \tag{2.19}$$

The matrix $\hat{X}$ is assumed to take its most general form

$$\hat{X} = \begin{pmatrix}p_L & p_H\\ q_L & -q_H\end{pmatrix};\ \hat{X}^{-1} = \frac{1}{p_H q_L + p_L q_H}\begin{pmatrix}p_L & q_L\\ p_H & -q_H\end{pmatrix}. \tag{2.20}$$

This can be simplified, when demanding that $\hat{X}$ is unitary, meaning $p = p_H = p_L$ and $q = q_H = q_L$, with $|p|^2 + |q|^2 = 1$ and $|p|^2 = |q|^2$. This means, the matrix $\hat{R}$ has to be normal and therefore there is no *CPV* in the mixing (c.f. Sec. 2.3.3). Eqs. 2.20 then simplify to

$$\hat{X} = \begin{pmatrix}p & p\\ q & -q\end{pmatrix};\ \hat{X}^{-1} = \begin{pmatrix}p & q\\ p & -q\end{pmatrix}. \tag{2.21}$$

Thus, the states $|B_L\rangle$ and $|B_H\rangle$ can be expressed by the parameters p and q and the physical states $|B^0\rangle$ and $|\overline{B}^0\rangle$:

$$\begin{pmatrix}|B_L\rangle\\|B_H\rangle\end{pmatrix} = \begin{pmatrix}p & q\\ p & -q\end{pmatrix}\begin{pmatrix}|B^0\rangle\\|\overline{B}^0\rangle\end{pmatrix} \Leftrightarrow \begin{pmatrix}|B^0\rangle\\|\overline{B}^0\rangle\end{pmatrix} = \frac{1}{2}\begin{pmatrix}{}^1\!/\!_p & {}^1\!/\!_q\\ {}^1\!/\!_p & -{}^1\!/\!_q\end{pmatrix}\begin{pmatrix}|B_L\rangle\\|B_H\rangle\end{pmatrix}. \tag{2.22}$$

As the time evolution of the states $|B_{L,H}\rangle$ is given by

$$\begin{pmatrix}|B_L\rangle(t)\\|B_H\rangle(t)\end{pmatrix} = \mathrm{e}^{-\mathrm{i}\mathcal{H}t}\begin{pmatrix}|B_L\rangle\\|B_H\rangle\end{pmatrix}, \tag{2.23}$$

the time evolution of the flavour eigenstates $|B^0\rangle(t)$ and $|\overline{B}^0\rangle(t)$ can be expressed by

$$\begin{pmatrix}|B^0\rangle(t)\\|\overline{B}^0\rangle(t)\end{pmatrix} = \hat{X}^{-1}\begin{pmatrix}|B_L\rangle(t)\\|B_H\rangle(t)\end{pmatrix} = \mathrm{e}^{-\mathrm{i}\mathcal{H}t}\hat{X}^{-1}\begin{pmatrix}|B_L\rangle\\|B_H\rangle\end{pmatrix}. \tag{2.24}$$

Therefore, it follows for the state $|B^0\rangle$ (and equivalently for the state $|\overline{B}^0\rangle$)

$$\begin{aligned}
|B^0(t)\rangle &= \frac{1}{2}\left(\frac{1}{p}|B_L(t)\rangle + \frac{1}{p}|B_H(t)\rangle\right) \\
&= \frac{1}{2p}\left(\mathrm{e}^{-\mathrm{i}\mu_L t}|B_L\rangle + \mathrm{e}^{-\mathrm{i}\mu_H t}|B_H\rangle\right) \\
&= \frac{1}{2p}\left(\mathrm{e}^{-\mathrm{i}\mu_L t}(p|B^0\rangle + q|\overline{B}^0\rangle) + \mathrm{e}^{-\mathrm{i}\mu_H t}(p|B^0\rangle - q|\overline{B}^0\rangle)\right) \\
&= \frac{1}{2p}\left((\mathrm{e}^{-\mathrm{i}\mu_L t} + \mathrm{e}^{-\mathrm{i}\mu_H t})|B^0\rangle + \frac{p}{q}(\mathrm{e}^{-\mathrm{i}\mu_L t} - \mathrm{e}^{-\mathrm{i}\mu_H t})|\overline{B}^0\rangle\right) \qquad (2.25) \\
|\overline{B}^0(t)\rangle &= \frac{1}{2q}\left((\mathrm{e}^{-\mathrm{i}\mu_L t} + \mathrm{e}^{-\mathrm{i}\mu_H t})|\overline{B}^0\rangle + \frac{q}{p}(\mathrm{e}^{-\mathrm{i}\mu_L t} - \mathrm{e}^{-\mathrm{i}\mu_H t})|B^0\rangle\right). \qquad (2.26)
\end{aligned}$$

In order to allow a description of the time evolution by parameters that are measurable, the mass m and decay width Γ of the B^0-meson are defined as follows:

$$m \equiv \frac{m_L + m_H}{2} \text{ and} \qquad (2.27)$$

$$\Gamma \equiv \frac{\Gamma_L + \Gamma_H}{2}. \qquad (2.28)$$

Furthermore, it is useful to define the differences $\Delta\mu$, Δm, and $\Delta\Gamma$ of the corresponding parameters between the states $|B_L\rangle$ and $|B_H\rangle$:

$$\begin{aligned}
\Delta\mu &\equiv \mu_H - \mu_L \\
&= \Delta m - \mathrm{i}/2\Delta\Gamma \qquad (2.29) \\
\Delta m &\equiv m_H - m_L \qquad (2.30) \\
\Delta\Gamma &\equiv \Gamma_H - \Gamma_L. \qquad (2.31)
\end{aligned}$$

In one common convention Δm is defined to be positive, meaning $|B_L\rangle$ belongs to the lighter and $|B_H\rangle$ to the heavier mass eigenstate. The sign of $\Delta\Gamma$ can then only be determined experimentally. This convention will be used throughout this thesis.

Using Eqs. (2.25) and (2.26) and the substitutions

$$e^+ = \mathrm{e}^{-\mathrm{i}\mu_L t} + \mathrm{e}^{-\mathrm{i}\mu_H t} \text{ and} \qquad (2.32)$$

$$e^- = \mathrm{e}^{-\mathrm{i}\mu_L t} - \mathrm{e}^{-\mathrm{i}\mu_H t}, \qquad (2.33)$$

the probability that a B meson with the initial state $|B^0\rangle$ or $|\overline{B}^0\rangle$ decays to a self

tagging final state $\langle f|$ or $\langle \bar{f}|$ can then be expressed by the matrix elements

$$|\langle f|\hat{R}|B^0(t)\rangle|^2 = \frac{1}{4p^2}\left(A_f^2\, e^+ e^{+*}\right), \tag{2.34}$$

$$|\langle \bar{f}|\hat{R}|\overline{B}^0(t)\rangle|^2 = \frac{1}{4q^2}\left(\bar{A}_{\bar{f}}^2\, e^+ e^{+*}\right), \tag{2.35}$$

$$|\langle \bar{f}|\hat{R}|B^0(t)\rangle|^2 = \frac{1}{4p^2}\left(\frac{q}{p}\bar{A}_{\bar{f}}^2\, e^- e^{-*}\right), \tag{2.36}$$

$$|\langle f|\hat{R}|\overline{B}^0(t)\rangle|^2 = \frac{1}{4q^2}\left(\frac{q}{p}A_f^2\, e^- e^{-*}\right). \tag{2.37}$$

Herein, a self tagging final state is defined such, that $\langle f|\hat{R}|B^0\rangle = A_f$, $\langle \bar{f}|\hat{R}|\overline{B}^0\rangle = \bar{A}_{\bar{f}}$, and $\langle \bar{f}|\hat{R}|B^0\rangle = \langle f|\hat{R}|\overline{B}^0\rangle = 0$. Thus, only the one final state corresponding to the initial state is allowed in the decay.

Using the definitions from Eqs. (2.29)-(2.31), $e^{\pm}$ can be calculated depending on Δm and $\Delta\Gamma$. Replacing $\mu_L = \mu - {}^{\Delta m}/_2 + {}^{\mathrm{i}}/_4\Delta\Gamma$ and $\mu_H = \mu + {}^{\Delta m}/_2 - {}^{\mathrm{i}}/_4\Delta\Gamma$, the exponential functions can be written as

$$\begin{aligned} \mathrm{e}^{-\mathrm{i}\mu_L t} &= \mathrm{e}^{-\mathrm{i}(\mu - \frac{\Delta m}{2} + \frac{\mathrm{i}}{4}\Delta\Gamma)t} \\ &= \mathrm{e}^{-\mathrm{i}\mu t}\mathrm{e}^{-\mathrm{i}(\frac{\Delta m}{2} + \frac{\mathrm{i}}{4}\Delta\Gamma)t} \end{aligned} \tag{2.38}$$

$$\mathrm{e}^{-\mathrm{i}\mu_H t} = \mathrm{e}^{-\mathrm{i}\mu t}\mathrm{e}^{\mathrm{i}(\frac{\Delta m}{2} + \frac{\mathrm{i}}{4}\Delta\Gamma)t}. \tag{2.39}$$

Thus, the parameters $\Delta\Gamma$ and Δm get visible again in the terms of Eqs. (2.34) - (2.37)

$$\begin{aligned} e^+ e^{+*} &= \left(\mathrm{e}^{-\mathrm{i}\mu t}\mathrm{e}^{-\mathrm{i}(\frac{\Delta m}{2} + \frac{\mathrm{i}}{4}\Delta\Gamma)t} + \mathrm{e}^{\mathrm{i}(\frac{\Delta m}{2} + \frac{\mathrm{i}}{4}\Delta\Gamma)t}\right)^2 \\ &= \mathrm{e}^{-\mathrm{i}\mu t}\mathrm{e}^{\mathrm{i}\mu t}\left(\mathrm{e}^{-\mathrm{i}(\frac{\Delta m}{2} + \frac{\mathrm{i}}{4}\Delta\Gamma)t} + \mathrm{e}^{\mathrm{i}(\frac{\Delta m}{2} + \frac{\mathrm{i}}{4}\Delta\Gamma)t}\right)\left(\mathrm{e}^{\mathrm{i}(\frac{\Delta m}{2} - \frac{\mathrm{i}}{4}\Delta\Gamma)t} + \mathrm{e}^{-\mathrm{i}(\frac{\Delta m}{2} - \frac{\mathrm{i}}{4}\Delta\Gamma)t}\right) \\ &= \mathrm{e}^{-\Gamma t}\left(\mathrm{e}^{-\Delta m t} + \mathrm{e}^{\Delta m t} + \mathrm{e}^{-(\frac{\mathrm{i}}{2}\Delta\Gamma)t} + \mathrm{e}^{(\frac{\mathrm{i}}{2}\Delta\Gamma)t}\right) \\ &= 2\mathrm{e}^{-\Gamma t}\left(\cosh({}^{\Delta\Gamma}/_2\, t) + \cos(\Delta m t)\right) \end{aligned} \tag{2.40}$$

$$e^- e^{-*} = 2\mathrm{e}^{-\Gamma t}\left(\cosh({}^{\Delta\Gamma}/_2\, t) - \cos(\Delta m t)\right) \tag{2.41}$$

$$e^+ e^{-*} = 2\mathrm{e}^{-\Gamma t}\left(\sinh({}^{\Delta\Gamma}/_2\, t) + \mathrm{i}\sin(\Delta m t)\right) \tag{2.42}$$

$$e^- e^{+*} = 2\mathrm{e}^{-\Gamma t}\left(\sinh({}^{\Delta\Gamma}/_2\, t) - \mathrm{i}\sin(\Delta m t)\right). \tag{2.43}$$

In the B^0–$\overline{B}^0$ system, the difference in the decay width $\Delta\Gamma$ is measured to be compatible with zero [7]. Furthermore, ${}^p/_q = 1$, if there is no *CPV*. Keeping in mind Eqs. (2.40) to (2.43) for later chapters, the probability for the decay of a $\overline{B}^0$ to a self-tagging final state f can now be written as

$$|\langle f|\hat{R}|\overline{B}^0\rangle|^2 = \frac{1}{2}\mathrm{e}^{-\Gamma t}\left(1 - \cos(\Delta m_d t)\right). \tag{2.44}$$

2.3.2 Constraining the CKM matrix

Fig 2.1 shows the most recent constraints on the CKM triangle. The yellow and orange circles on the classic unitarity triangle show the constraints from measurements of Δm_d (and Δm_s). The radius of these circles is given by

$$R_t = \left| \frac{V_{td} V_{tb}^*}{V_{cd} V_{cb}^*} \right|. \tag{2.45}$$

The $B^0 \overline{B}^0$ mixing frequency Δm_d can be calculated with

$$\Delta m_d = \frac{G_F^2}{6\pi} \eta_B m_{B_d} (B_{B_d} F_{B_d}^2) M_W^2 S_0(x_t) |V_{td}|^2 \text{ [11]}, \tag{2.46}$$

where $S_0(x_t)$ is the Inami-Lim function, η_B is a QCD correction parameter, G_F is the Fermi coupling, M_W is the $W^\pm$ boson mass, and B_{B_d} and F_{B_d} are the Bag factor and the leptonic decay constant. Using this, the constraint R_t to the CKM triangle can be calculated as

$$R_t = 1.52 \frac{\Delta m_d}{\sqrt{S_0(x_t)} \eta_B \sqrt{B_{B_d}} F_{B_d} |V_{cb}|} \text{ [11]} \tag{2.47}$$

The accuracy of this expression is dominated by a large uncertainty on the factor

$$\sqrt{B_{B_d}} F_{B_d} \approx 200 \pm 40 \,\text{MeV} \text{ [11, 36, 37, 38]}.$$

The theoretical sensitivity can be increased, using a a simultaneous measurement of Δm_d and Δm_s, with

$$R_t = \frac{1}{\sqrt{R_{ds}}} \sqrt{\frac{\Delta m_d}{\Delta m_s}} \frac{1}{\lambda_C} \sqrt{1 - \lambda_C^2 (1 - 2\rho)} \text{ with } R_{ds} = \frac{m_{B_d}}{m_{B_s}} \left(\frac{F_{B_d}^2 B_{B_d}}{F_{B_s}^2 B_{B_s}} \right) \text{ [11]}.$$

Herein, λ_C and ρ are parameters from the Wolfenstein parametrisation of the CKM matrix and F_{B_s} and B_{B_s} are the leptonic decay constant and the Bagg factor for the B_s^0. It is claimed [11] that R_t can be constrained with a sensitivity of 10 % if $\Delta m_d / \Delta m_s$ is measured precisely.

2.3.3 *CP* violation

In the understanding of modern physics the difference between the laws of physics for matter and antimatter is a key to the validity of the SM and many cosmological phenomena. It has first been discovered by the experiments of Fitch and Cronin in the neutral kaon system [18] in 1964. In modern physics any *CP* violation that is not expected in the SM would be a hint for new physics. One of the main goals of

the LHCb experiment is to search for such contributions in the decays of B and D mesons.

The only source for *CPV* in the SM is the complex phase of the CKM matrix. Thus, any *CP* violating effect has to include CKM matrix elements in which this complex phase occurs (e.g. V_{td} or V_{ub}). In the SM, there are three different types of *CPV*.

- Direct *CPV*, or *CPV* in the decay,
- indirect *CPV*, or *CPV* in the mixing, and
- mixing induced *CPV*, or *CPV* in the interference of mixing and decay.

Direct *CPV* occurs if there are different decay amplitudes A_f and $\bar{A}_{\bar{f}}$ for the two *CP* conjugated decays $\langle f|i\rangle$ and $\langle \bar{f}|\bar{i}\rangle$, meaning that $|A_f \ / \ \bar{A}_{\bar{f}}| \neq 1$ in the case of *CPV*. The decay amplitudes are given by

$$A_f = \sum_i A_i \mathrm{e}^{\mathrm{i}(\delta_i+\phi_i)}. \tag{2.48}$$

Herein, $\sum_i A_i$ describes the fact, that several different terms might contribute to the total decay amplitude (e.g. CKM suppressed decays). The weak phase ϕ_i arrises from the CKM matrix elements involved in that amplitude, and the strong phase δ_i connected to absorbtive parts of the amplitude mostly due to the strong interaction. For the charge conjugated amplitude only the weak phase ϕ_i changes its sign. Thus, the amplitude for the charge conjugated mode can be written as

$$\bar{A}_{\bar{f}} = \sum_i A_i \mathrm{e}^{2i(\epsilon_f-\epsilon_B)} \mathrm{e}^{\mathrm{i}(\delta_i-\phi_i)}, \tag{2.49}$$

where the phases ϵ_B and ϵ_f can be chosen arbitrarily due to the flavour conservation of the strong interaction. That leads to

$$\left|\frac{A_f}{\bar{A}_{\bar{f}}}\right| = \left|\frac{\sum_i A_i \mathrm{e}^{\mathrm{i}(\delta_i+\phi_i)}}{\sum_i A_i \mathrm{e}^{\mathrm{i}(\delta_i-\phi_i)}}\right|. \tag{2.50}$$

Hence, it becomes clear that any direct *CPV* needs at least two different contributions with different weak and strong phases, as for $i = 1$ the result of Eq. (2.50) always equals one. Any *CPV* in decays of charged mesons has to be of this type.

Indirect *CPV* may be observed if the process $|i\rangle \to |\bar{i}\rangle$ is more probable than $|\bar{i}\rangle \to |i\rangle$ in the mixing of neutral mesons. The quantity describing indirect *CPV* is

$$\left|\frac{q}{p}\right| = \left|\frac{M_{12}^* - \mathrm{i}/2\Gamma_{12}^*}{M_{12} - \mathrm{i}/2\Gamma_{12}}\right|. \tag{2.51}$$

Therefore $q/p = 1$ if *CP* is conserved, implying that the mass eigenstates are *CP* eigenstates, so that the relative phase between the off-diagonal elements M_{12} and Γ_{12}

of $\hat{\Gamma}$ and $\hat{M}$ vanishes. Indirect *CPV* has been measured in the neutral kaon system [18] and is also accessible in the time dependent asymmetries in decays of neutral B mesons to semileptonic final states (e.g. $B^0 \to \mu\nu X$) [39, 40].

CP violation in the interference of decay and mixing can occur, when analysing decays of neutral B mesons to *CP* eigenstates that are accessible for both, B and $\overline{B}$ mesons. As this type of *CPV* combines the effects from decay and mixing, the relevant phase independent quantity here is

$$\lambda_{CP} = \eta_{f_{CP}} \frac{q}{p} \frac{\bar{A}_{\bar{f}}}{A_f}, \tag{2.52}$$

with $\eta_{f_{CP}}$ being the *CP* eigenvalue of the final state. Thus, $\lambda_{CP} \neq \pm 1$ in the case of direct and/or indirect *CPV*. If there is no direct, nor indirect *CPV*, it is still possible, that $|\lambda_{CP}| = \pm 1$, but $\Im(\lambda_{CP}) \neq 0$. This is referred to as *CP* violation in the interference of decays with and without mixing, as it exploits the fact, that the same final state $\langle f|$ has different time dependent amplitudes for the initial states $|i\rangle$ and $|\bar{i}\rangle$.

The search for interference *CPV* in the decay channels $B^0 \to J/\psi\, K^0_{\mathrm{S}}$, $B^0_s \to J/\psi\, f^0$, or $B^0_s \to J/\psi\, \phi$ gives a handle on the CKM angles $\beta = \arg\left(-V_{cd}V^*_{cb}/V_{td}V^*_{tb}\right)$ and $\phi_s = \arg\left(-V_{cs}V^*_{cb}/V_{ts}V^*_{tb}\right)$ and is one of the main fields of interest at the LHCb experiment.

Additionally, [28] shows, that the decay channel $B^0 \to D^-\pi^+$ gives an opportunity to measure *CPV* at the LHCb experiment. The final states of this decay are not completely self-tagging, but there is a largely CKM favoured and an unfavoured decay to the charge conjugated final state. That means, oppositely to the calculations leading to Eq. (2.44), the matrix elements $\langle \bar{f}|\hat{R}|B^0\rangle = A_{\bar{f}}$ and $\langle f|\hat{R}|\overline{B}^0\rangle = \bar{A}_f$ are allowed, but suppressed by the ratio of the corresponding CKM matrix elements.

Eqs. (2.34) to (2.37) then become slightly more complex, as the second possible final state has to be considered. The B^0 width difference is considered to be $\Delta\Gamma = 0$ from the beginning and the two *CP* violation parameters

$$\lambda = \frac{q}{p} \frac{\bar{A}_f}{A_f} \text{ and} \tag{2.53}$$

$$\bar{\lambda} = \frac{p}{q} \frac{A_{\bar{f}}}{\bar{A}_{\bar{f}}} \tag{2.54}$$

are introduced. Eq. (2.34) for the decay amplitudes of the decay $\langle f|\hat{R}|B^0(t)\rangle$ then becomes

$$
\begin{aligned}
|\langle f|\hat{R}|B^0(t)\rangle|^2 &= \frac{1}{4p^2}\left|A_f\, e^+ + \bar{A}_f \frac{p}{q} e^-\right|^2 \\
&= \frac{1}{4p^2} A_f^2 \left|e^+ + \frac{\bar{A}_f}{A_f}\frac{p}{q} e^-\right|^2 \\
&= \frac{1}{4p^2} A_f^2 \left|e^+ + \lambda e^-\right|^2 \\
&= \frac{1}{4p^2} A_f^2 \left(e^+ e^{+*} + |\lambda|^2 e^- e^{-*} + \lambda e^+ e^{-*} + \bar{\lambda} e^- e^{+*}\right) \\
&= \frac{1}{4p^2} A_f^2 \left((1+|\lambda|^2) + \cos(\Delta m t) - |\lambda|^2\cos(\Delta m t) + \lambda e^+ e^{-*} + \bar{\lambda} e^- e^{+*}\right) \\
&= \frac{1}{4p^2} A_f^2 \left((1+|\lambda|^2) + (1-|\lambda|^2)\cos(\Delta m t) + 2\mathcal{I}(\lambda)\sin(\Delta m t)\right).
\end{aligned}
\tag{2.55}
$$

Herein, it is used that all A_f are proportional to the CKM matrix elements that govern the decay. It is safe to say then, that $|A_f| = |\bar{A}_{\bar{f}}|$ and $|\bar{A}_f| = |A_{\bar{f}}|$ if there is no direct *CPV*. Additionally, $\bar{\lambda} = \lambda^*$ as no strong phase difference is expected between the favoured and the suppressed feynman diagrams for the decay. Analogously it follows for the other amplitudes (compare [28, 27])

$$|\langle \bar{f}|\hat{R}|\overline{B}^0(t)\rangle|^2 \propto \left((1+|\lambda|^2) + (1-|\lambda|^2)\cos(\Delta m t) - 2\mathcal{I}(\lambda)\sin(\Delta m t)\right), \tag{2.56}$$

$$|\langle \bar{f}|\hat{R}|B^0(t)\rangle|^2 \propto \left((1+|\lambda|^2) - (1-|\lambda|^2)\cos(\Delta m t) - 2\mathcal{I}(\lambda)\sin(\Delta m t)\right), \tag{2.57}$$

$$|\langle f|\hat{R}|\overline{B}^0(t)\rangle|^2 \propto \left((1+|\lambda|^2) - (1-|\lambda|^2)\cos(\Delta m t) + 2\mathcal{I}(\lambda)\sin(\Delta m t)\right). \tag{2.58}$$

Note, that in the case of a self-tagging final state $\lambda = \bar{\lambda} = 0$ and the equations then describe a decay with flavour oscillations and no visible *CPV*.

2.3.4 New physics

Since the discovery of *CPV* in the kaon system, similar effects where also predicted and found for heavier mesons. As the overall amount of *CPV* that has been found yet, cannot explain the matter/antimatter asymmetry in the universe, it is one main physics goal of the current experiments to do precision tests of the SM and the unitarity of the CKM matrix, in order to find possible tensions that might be explained by BSM physics.

Precision measurements of the angles of the CKM unitarity triangles (cf. Fig. 2.1) in measurements of time-dependent *CPV* and also of the lengths of the triangle sides (e.g. by the measurement of the oscillation frequencies of B mesons) are one possibility for such tests. Other efforts are taken in the measurements of rare decays (e.g. $B_s^0 \to \mu\mu$) that are expected to have larger decay amplitudes due to BSM physics than those predicted by the SM.

2.4 Particle production in pp-collisions

Recently, most precision measurements in hadron systems containing b quarks were performed at e^+e^- colliders at the BaBar and Belle experiments. In B factories B mesons are usually produced from the decays of the $\Upsilon(4S)$ resonance. These B events from electron-positron annihilation are very clean and usually do not contain tracks that do not belong to the decays of the B mesons. The B mesons from $\Upsilon(4S)$ decay back-to-back in the $\Upsilon(4S)$ center-of-mass (CMS) system and a boost of the $b\bar{b}$ system can only be achieved by boosting the e^+e^- CMS. When produced at a B factory, the $B\overline{B}$ system is quantum dynamically entangled. Thus, if the state (B or $\overline{B}$) of one B meson is known at a certain time, the entangled meson is known to be *CP* conjugated at the same time.

The situation at hadron colliders, especially at the LHC is much different. First of all, b quarks are not produced in a resonance but directly from gluon fusion. Therefore, the pairs of b hadrons are not entangled. Additionally, there is no defined center-of-mass system for the collision, as protons are complex particles which do not collide themselves. At the energy of the LHC, in most cases the gluons from the protons collide. Thus, the boost of the resulting b jets depends on the momenta of the colliding partons. As a result, the $b\bar{b}$ pairs are strongly boosted along the beam axis (cf. Fig. 2.3).

The layout of the LHCb detector is justified by this fact, as about one quarter of all produced $b\bar{b}$ pairs is expected to be within the LHCb acceptance of only about 300 mrad. As the boost of both, the b and the $\bar{b}$ quark, is strongly correlated in nearly all events with one b quark from a primary pp interaction, the corresponding $\bar{b}$ quark is likely to be found in the acceptance of the detector. That is essential, because if one b quark belongs to a signal decay, the second one can be used to tag the initial state of the signal. This so-called flavour tagging is neccessary for many analyses of *CPV* as well as for measurements of $B^0\overline{B}^0$ and $B_s^0\overline{B}_s^0$ mixing.

The $b\bar{b}$ production from partons also leads to a large number of additional charged tracks from the primary pp interaction. Thus, a strong trigger and powerful selection criteria become necessary to suppress the combinatorial background.

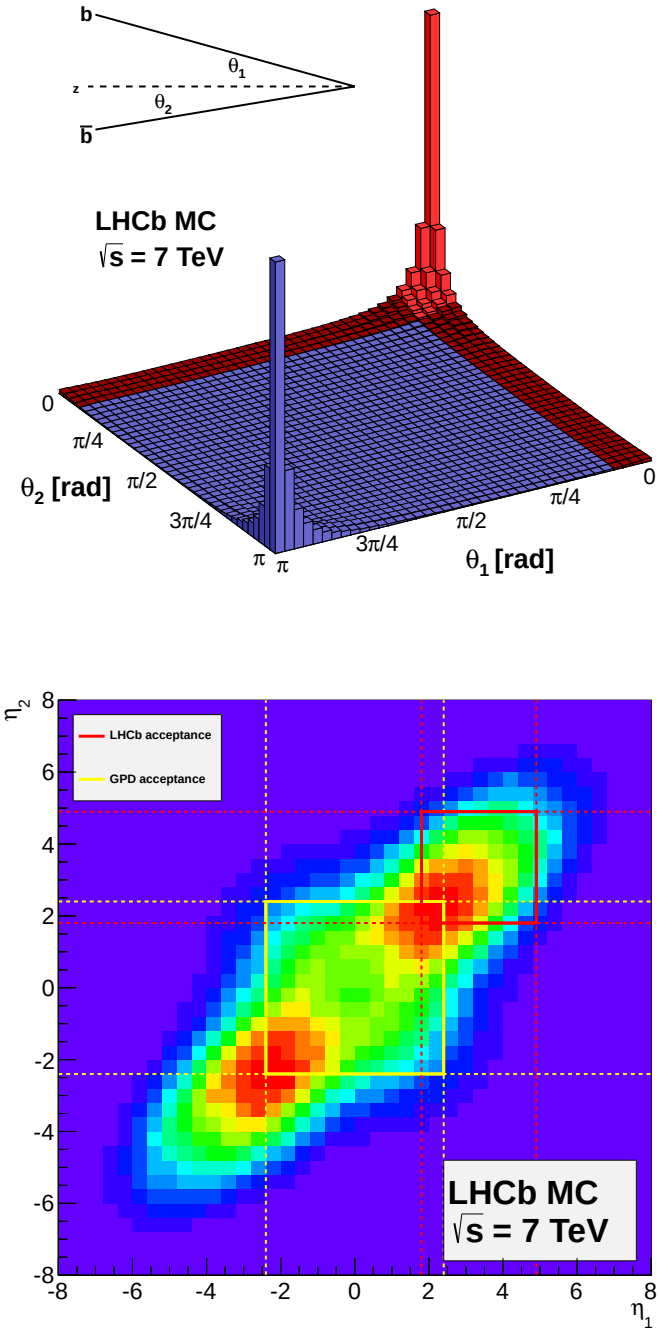


Fig. 2.3: (Top) angular correlation and (bottom) pseudorapidity distribution of the $b\bar{b}$ pairs produced in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. The red band and box indicate the acceptance of the LHCb detector of 300 mrad, the yellow box corresponds to the acceptance of a general purpose detector. [41].

3 The experimental setup

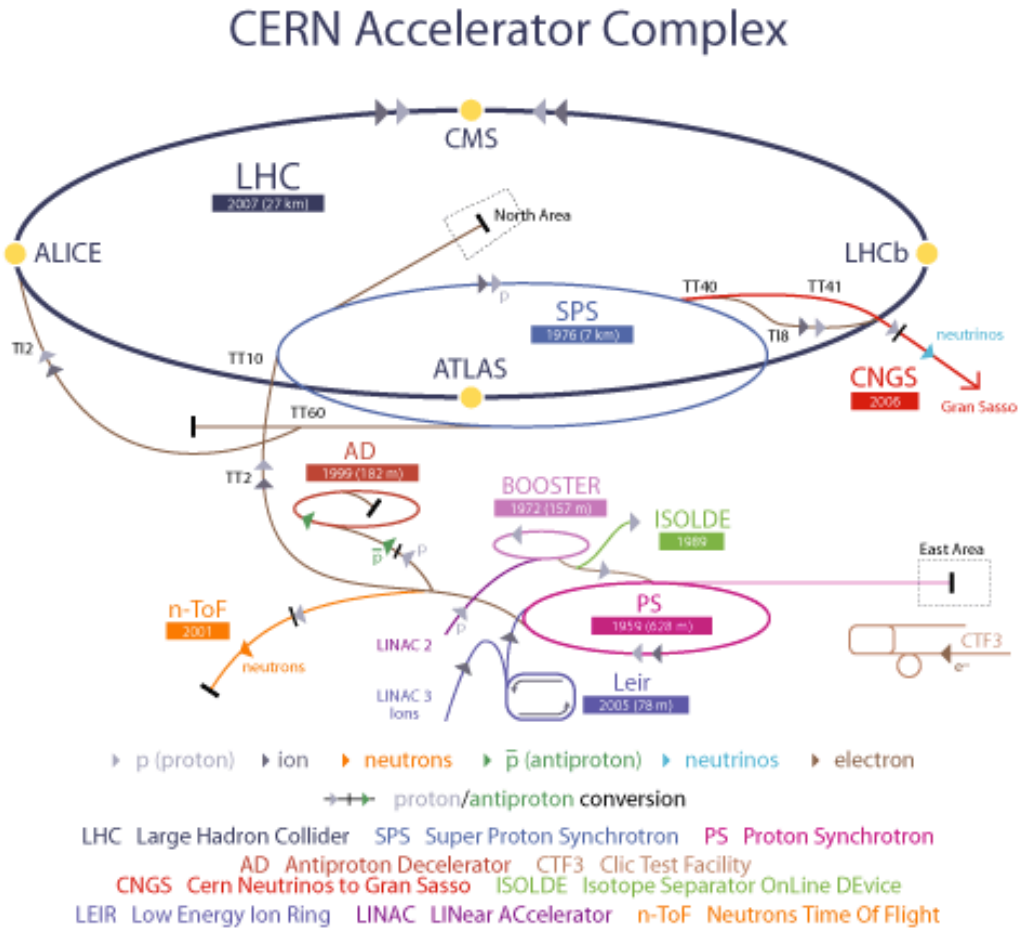


Fig. 3.1: Schematical view of the LHC accelerator complex. [42] The protons for the *pp* collisions run from the LINAC through the BOOSTER, the proton synchrotron (PS) and the super proton sychrotron (SPS) before they are injected into the LHC with an energy of 450 GeV.

The LHCb experiment is situated at the Large Hadron Collider (LHC) at the European Center for Nuclear Research (CERN) near Geneva. The LHC is a large proton synchrotron providing high intensity beams of protons that are collided in four different interaction points. It is designed to deliver proton-proton (*pp*) collisions at a

center of mass energy $\sqrt{s} = 14$ TeV. During the starting phase of the LHC, the energy has been ramped up slowly, so that the first significant amount of data was taken at $\sqrt{s} = 7$ TeV in 2010 and 2011, whereas the 2012 data taking period includes pp collisions at $\sqrt{s} = 8$ TeV. The protons are collided in bunches of about 10^{11} protons per bunch with an interaction rate of 20 MHz at each interaction point. This leads to an instantaneous luminosity of approximately $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

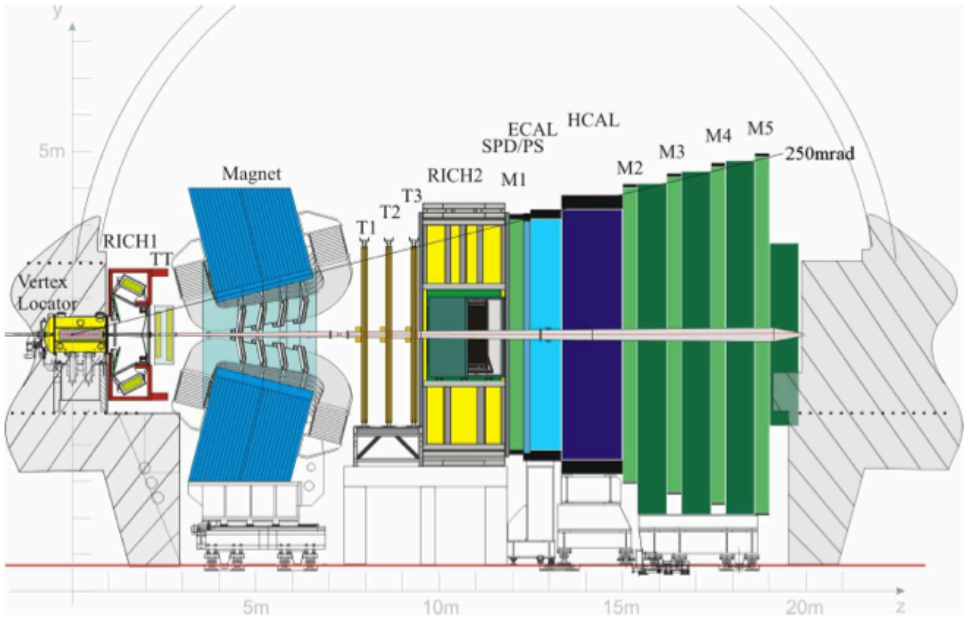


Fig. 3.2: Sideview of the LHCb detector layout. LHCb covers an approximately 300 mrad section of a standard 4π detector layout. Taken from [43].

As, contrary to the two general purpose experiments ATLAS and CMS, LHCb needs clean interactions with a limited number of pp interactions, the luminosity is reduced to $\mathcal{L} = 4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ at the LHCb interaction point by separating the beams from each other. Thus, it is possible to keep a constant instantaneous luminosity at the LHCb experiment by adjusting the beam separation during the run. Note, that the actual luminosity in the 2011 data-taking period of $\mathcal{L} = 4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is well beyond the design luminosity $\mathcal{L} = 2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ of the experiment. This is, because the number of primary interactions per event has been increased to 2 primary interactions on average per event (the design value was 0.7).

As the flight direction of the b quarks in $b\bar{b}$ pairs is highly correlated and they are produced predominantly in forward/backward (z) direction close to the beam axis (c.f. Fig. 3.1), the LHCb detector is designed to cover an angle of 300 mrad in the bending plane (x, z) of its dipole magnet (horizontally, perpendicular to the beam axis). In the vertical direction (y) an angle of 250 mrad is instrumented. The design of the LHCb experiment is illustrated in Fig. 3.2. The detector consists of a tracking system

with three different parts: The vertex locator (VELO), the Tracker Turicensis (TT) upstream of the dipole magnet, and the main tracking station with inner (IT) and outer tracker (OT) downstream of the magnet. Particle masses can be measured by a ring imaging cherenkov detector (RICH) system that covers three momentum ranges in two stations upstream and downstream of the magnet. Additionally, the detector has an electromagnetic (ECAL) and a hadronic calorimeter (HCAL) and a muon system for the identification and precision tracking of muons. Two stations of a Beam Conditions Monitor (BCM) are placed upstream and downstream of the primary interaction point to monitor beam losses and potential hazards for the detector. The trigger system consists of a hardware (L0) stage and a two stage software high level trigger (HLT) that reduce the initial pp interaction rate of 20 MHz to about 4 kHz of interactions stored on disc. All detector components are described in detail in the following sections.

3.1 Tracking system

3.1.1 Vertex locator

The VELO is a silicon strip detector located very close to the LHCb interaction point. It consists of 21 stations, each of which has an r- and a ϕ-sensor to measure the distance from and the angle around the beam axis, respectively. Additionally there are two so-called pile-up sensors on the upstream side of the VELO which only consist of the r-sensors.

The VELO is divided into two halves that can be moved along the x-axis of the detector, in order to allow to secure the detector in a safe distance from the beam at LHC injection or other unstable beam conditions. When closed, the minimum distance of the VELO sensors to the beam is about 8 mm. That allows a precise measurement of the positions of the pp interaction vertex and possible decay vertices of particles decaying inside the VELO (e.g. B and D mesons) with a resolution of 60 μm along the beam axis and 10 μm perpendicular to the beam axis.

3.1.2 Dipole magnet

For the measurement of particle momenta, the LHCb detector has a normal conducting dipole magnet with a horizontal bending plane. The integrated magnet field of 4 Tm for tracks with a length of 10 m is generated by a current of nearly 6 kA. The magnetic field is designed such that it has a strength of less than 2 mT in the region of the sensitive parts of the RICH detectors. The magnet polarity is changed several times during every year to be able to see detection asymmetries for oppositely charged particles due to the detector geometry.

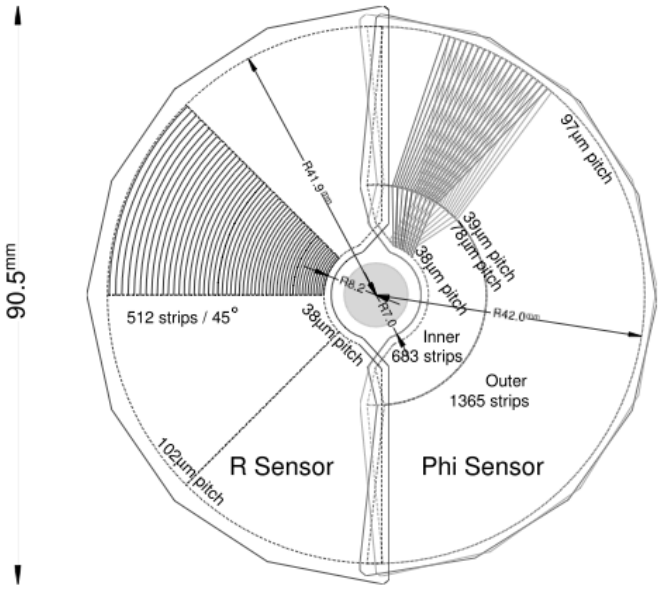


Fig. 3.3: Schematical view of one VELO station. The difference between the detectors for the radial distance (R Sensor) and for the opening angle around the detector axis (Phi Sensor) is visible on the left and right side of the image respectively. Taken from [43].

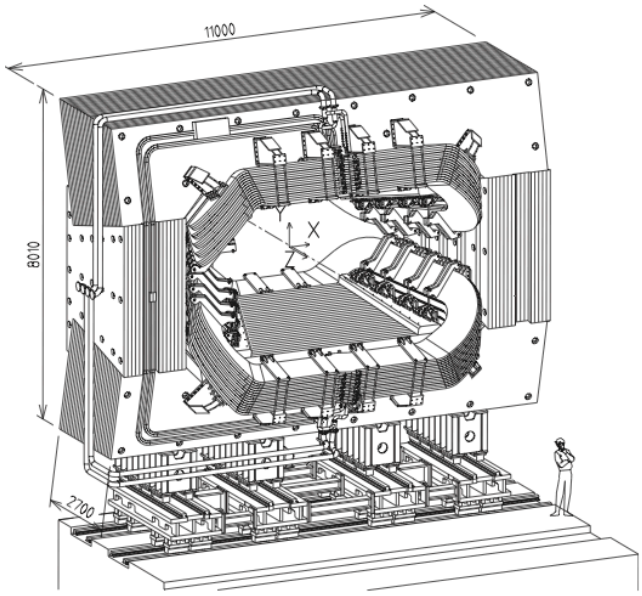


Fig. 3.4: Schematical view of the dipole magnet. Taken from [43].

3.1.3 Silicon trackers

The silicon trackers are situated in two different locations upstream and downstream of the magnet. The upstream detector is the tracker turicensis, a tracking detector that covers the full acceptance between the first RICH and the magnet. It consists of two stations and is particullary useful to find tracks of daughter particles of long living mesons like the K_S^0.

The inner tracker is the high occupancy part of the tracking system downstream of the magnet. The three IT stations are situated close to the beam pipe and are enclosed by the modules of the outer tracker. A schematical view of the silicon tracker (ST) stations is given in Fig. 3.5.

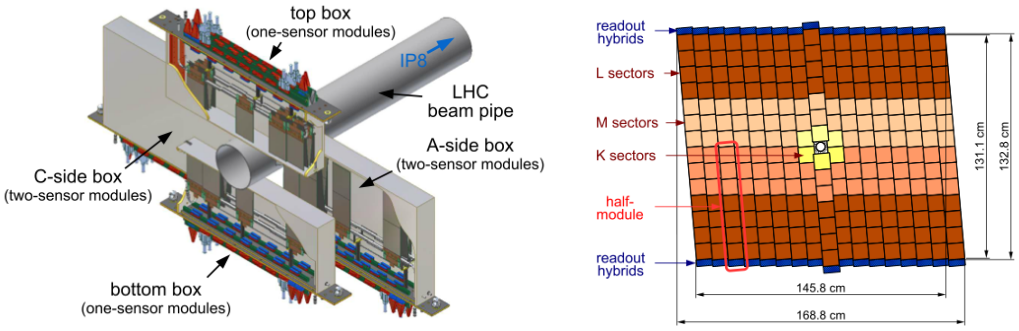


Fig. 3.5: Schematical view of the silicon trackers. (Left) the layout of the IT modules arround the beam pipe, and (right) the layout of the TT modules. Taken from [43].

The ST stations use silicon strips with a pitch of 200 µm giving a single hit resolution of about 50 µm. Each station consists of 4 layers $(x-u-v-x)$, where the x layers are aligned vertically and the u/v layers are rotated by $\pm 5\,^\circ$ against that axis, in order to be able to measured the y position of a hit. The ST stations have to deal with charged particle multiplicities of up to $5\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in the inner region of the TT.

3.1.4 Outer tracker

The outer tracker is a drift time detector that covers the low occupancy regions around the IT stations downstream of the magnet. Its drift tubes have an inner diameter of 4.9 mm and are filled with a mixture of 70 % Argon and 30 % CO_2 to guarantee a drift time below 50 μs. It is laid out in the same $(x-u-v-x)$ geometry as the ST stations. The inner boundary of the OT is designed such, that the occupancy does not exceed 10 % at the design luminosity of $2\times10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. A schematical view of the detector can be found in Fig. 3.6.

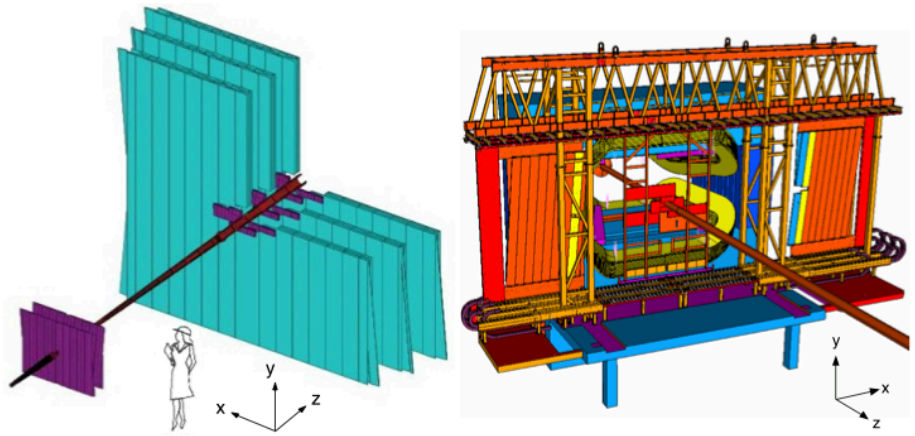


Fig. 3.6: (Left) schematical view of the OT (blue) and the ST stations (violet). (Right) the support structure of the OT and the dipole magnet. Taken from [43].

3.2 Particle identification system

The particle identification system consists of two ring imaging Cherenkov detectors. Both make use of the Cherenkov Effect for highly relativistic particles. When combined with the momentum measurements from the tracking system each track can be assigned a mass hypothesis due to the measurement of the Cherenkov angle in the RICH system.

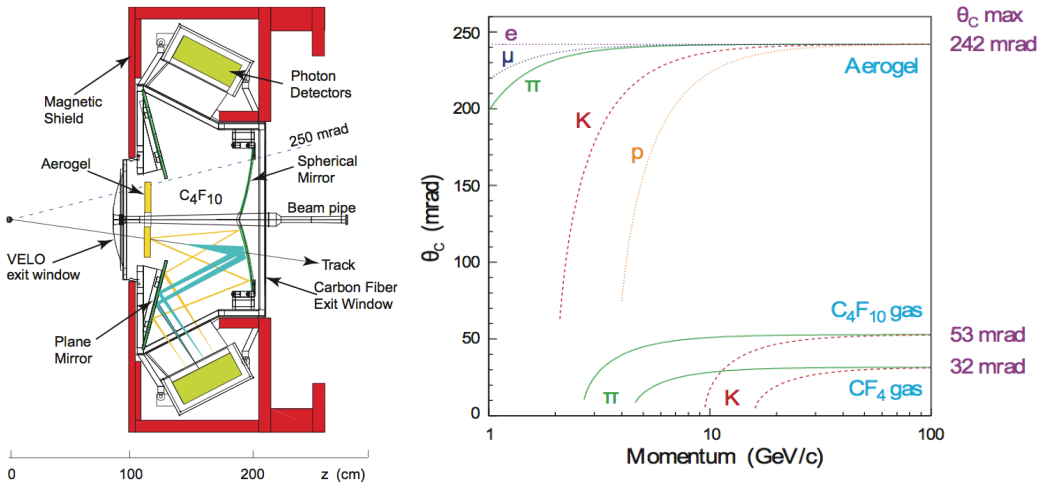


Fig. 3.7: Schematical view of RICH1 (left) and the opening angles of the Cherenkov cones for the different radiators (right). Taken from [43].

In order to deal with the wide momentum spectrum of charged particles in the

detector, two stations with different momentum acceptances are installed. RICH 1 is situated upstream of the magnet and covers the full acceptance. The radiators aerogel and C_4F_{10} enable it to cover a momentum range of about $1 - 60$ GeV. As most of the tracks with low momenta will be bent to the outer regions or even outside the detector acceptance by the dipole magnet, RICH 2 covers the inner region only with an angular acceptance of about ± 15 mrad to ± 120 mrad in the horizontal and ± 100 mrad in the vertical plain. With the radiator CF_4 it covers a momentum spectrum of about 15 GeV to more than 100 GeV. The difference between the Cherenkov angles of different particle with given momenta and a schematical view of the RICH 1 layout can be seen in Fig. 3.7.

3.3 Calorimeters

The calorimeter provides important information for the trigger system that will be described in Sec. 3.5. It measures the energy of charged and uncharged particles and is therefore crucial for measurements relying on the reconstructions of photons or uncharged pions.

In order to separate electrons from pions, a pre-shower (PS) and a scintillating pad detector (SPD) are installed upstream of the typical ECAL and HCAL setup. Charged particles produce hits in the SPD scintillators and can thus be distinguished from uncharged particles. This is essentially useful to discriminate uncharged pions from electrons. The PS helps to separate electrons from charged pions. Both PS and SPD allow the trigger to select electrons with high transverse energy with a low background from pions.

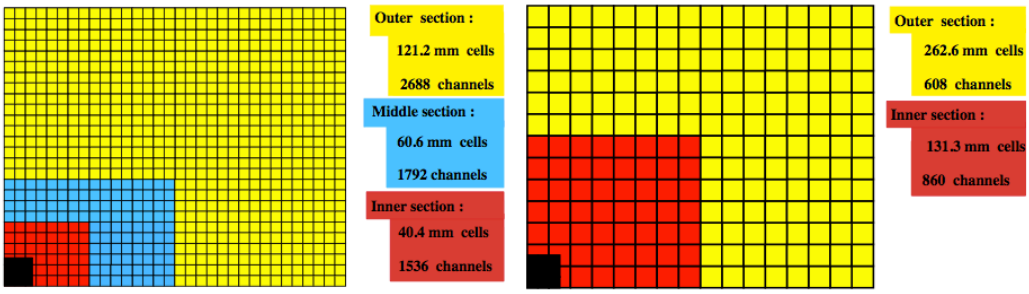


Fig. 3.8: Structure of the calorimeter layout which is similar for (left) EM and (right) hadronic calorimeters. The calorimeter cell sizes are 40, 60 and 120 mm for the inner (red), middle (blue), and outer (yellow) sections of the EM calorimeter, respectively. In the hadronic calorimeter the cell sizes are 130 mm in the inner (red) section and 260 mm in the outer (yellow) section. Taken from [43].

The thickness of the ECAL is chosen to be 25 radiation lengths of high energy photons to optimise the energy resolution. Due to space limitations the HCAL has a thickness

of 5.6 interaction lengths for hadronic showers. The calorimeters are constructed in a shashlik style with alternating layers of lead and scintillating material. The overall structure of the PS/SPD, ECAL and HCAL can be seen in Fig. 3.8. The cell size is chosen to be smaller for the high occupancy regions around the beam pipe, as there is a variation of two orders of magnitude in hit density along the calorimeter surface.

3.4 Muon system

The identification of muons is essential for many of the key analyses (e.g. *CPV* measurement in $B^0 \to J/\psi K^0_S$ and $B^0_s \to J/\psi\,\phi$ or the rare decay $B^0_s \to \mu\mu$). The muon system consists of five stations M1–M5 where M1 is placed upstream of the calorimeters to improve the p_T resolution for the trigger and all others are situated downstream of the calorimeters. Plates of 80 cm thick iron are placed between the downstream chambers to select penetrating muons. This is the case for muons with a momentum of at least 6 GeV.

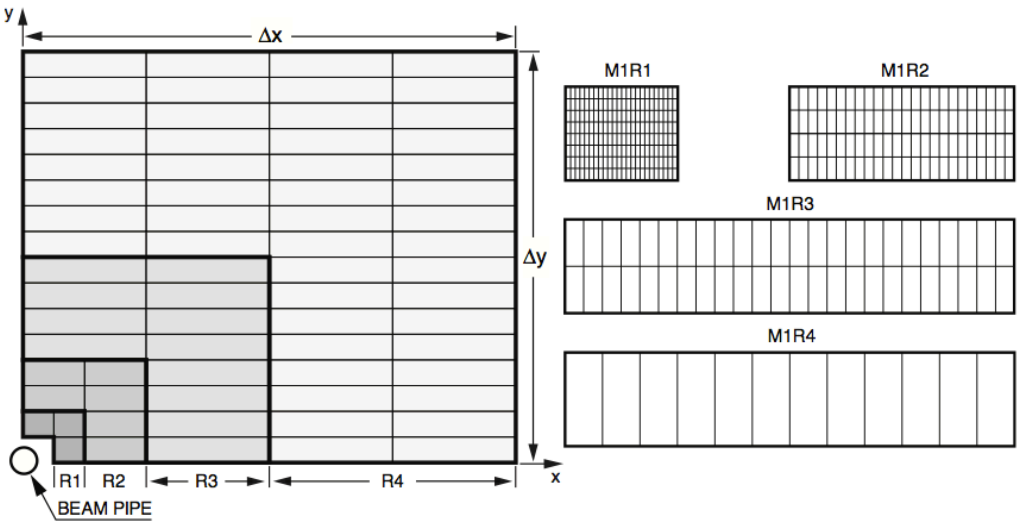


Fig. 3.9: Layout of the muon system. R1, R2, R3, and R4 are the different regions of the muon chambers corresponding to a different spacial resolution. The boxes on the right show the inner layout of the modules in each of the regions of the muon chamber M1. Taken from [43].

The muon chambers use multi-wire proportional chambers (and triple GEM chambers in the inner region of M1). The spatial resolution of M1–M3 is designed to measure the muon p_T with a resolution of about 20 %. The stations M4 and M5 are used to identify penetrating muons. The general layout is shown in Fig. 3.9.

3.5 Trigger system

At the design luminosity of $2 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and with a bunch crossing rate of 40 MHz a visible $b\bar{b}$ rate of 15 kHz is expected, where all decay products are within the detector acceptance. In fact, the LHC has been running with a reduced bunch crossing rate of 20 MHz during 2011 and 2012, the luminosity at the LHCb experiment was increased to $4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ leading to a higher number of pp interactions per bunch crossing. The three trigger stages L0, HLT1, and HLT2 are designed to reduce the event rate that is stored to about 4 kHz. The trigger criteria are optimised to preserve most of the decays of b hadrons and to suppress the combinatorial background.

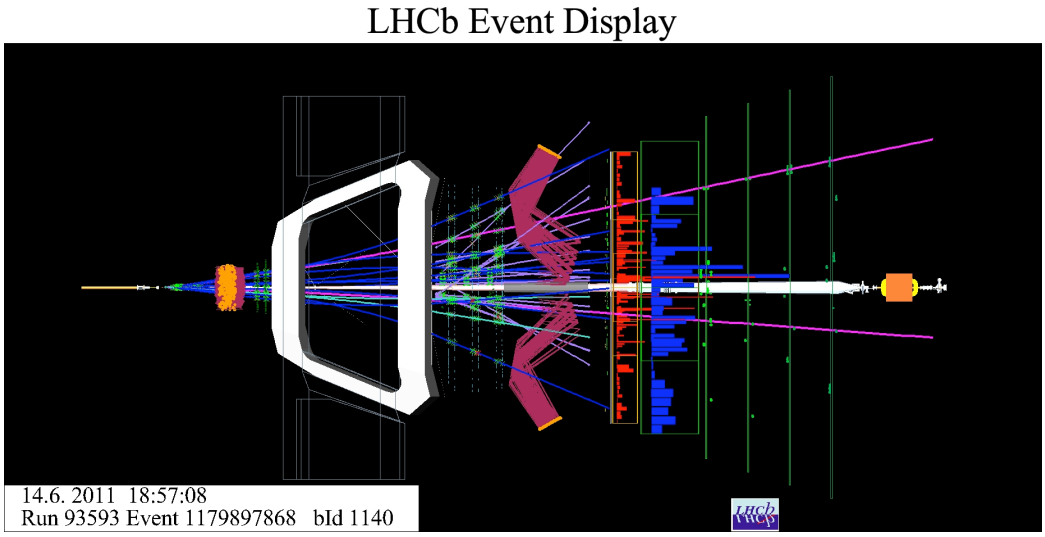


Fig. 3.10: Event display of a typical LHCb event. The straight purple tracks correspond to two particles identified as muons. The green and orange crosses mark hits in the tracking system, and the blue and red bars show energy deposition in the hadronic and EM calorimeter, respectively. The pink lines in the middle region show the photons in the Cherenkov detectors. [44]

When analysing an LHCb event (cf. Fig. 3.10) all software has to deal with a very high number of hits in the detector and thus potential particle tracks. In particular, this is a challenge for the trigger, as the large number of tracks leads to many possible combinations that could form a candidate for a decayed particle. Additionally, the processing time for each event is limited depending on the trigger stage. Furthermore, the large number of particles coming from the pp collision plays a role in flavour tagging (cf. Chapter 4), as it gets more difficult to identify the correct tagging tracks with a higher number of candidates.

3.5.1 Level 0 trigger

The L0 trigger operates synchronously with the nominal bunch crossing frequency of 40 MHz. It has to reduce the event rate to 1 MHz, as the complete detector can be read out at this frequency. In order to achieve this, information from the VELO, from the CALO, and from the muon system is used.

In the CALO, the L0 tries to reconstruct electron, photon or hadron clusters with very high transverse energy E_{T}. The muon system is used to select muons with a high transverse momentum p_{T} and the VELO and SPD can be used to suppress high multiplicity events in which the detector occupancy might be too high for good data quality. However, due to the changes in the LHC running conditions (lower interaction rate, higher luminosity) the average detector occupancy exceeds the design expectations. Nevertheless, the data quality is very good for the LHCb physics program for data taken during 2011 and 2012.

The selection of high E_{T} and p_{T} particles at the L0 stage leads to an enrichment of the data with $b\bar{b}$ events, since b hadrons favour decays to such particles due to their high mass. Additionally, it is possible to run the L0 in a random trigger mode, where it randomly reduces the interaction rate. This is necessary for studies of the global event shapes or of the production cross sections for certain particles.

3.5.2 High level trigger

The HLT is again divided into two separate parts. Both run on a computer farm and can be rerun on stored data. HLT1 has access to the full event. However, due to time limitations it is not possible to reconstruct the tracks in the event at this stage. Thus, HLT1 forms a decision based on information from the subdetectors. This information includes the impact parameters (IP) of VELO tracks, information about L0 candidates forming tracks in the tracking stations or the combination of information from the VELO and the TT. The HLT1 reduces the event rate to about 100 kHz.

The rate of accepted events coming from HLT1 is sufficiently low to allow an offline track reconstruction at the HLT2 stage. Thus, HLT2 can be used to do a full selection of exclusive candidates including the search for displaced vertices or two muons with a small distance of closest approach. The HLT2 Lines are specialised for certain or a small group of decay channels. All these lines are executed in parallel and the events are flagged by the lines with a positive decision for the offline analysis. The HLT lines relevant for this thesis will be discussed in more detail in Chapter 6.

3.6 Software

The general layout of the software packages used for this thesis can be seen in Fig. 3.11. Most of the software is integrated in the Gaudi framework [45, 46] that provides interfaces and conventions for the different programs. The software can be divided into three general parts. The simulation package Gauss, the software for the online analysis (e.g. trigger, tracking, reconstruction, pre-selection), and for the offline analysis (final selection and fitting).

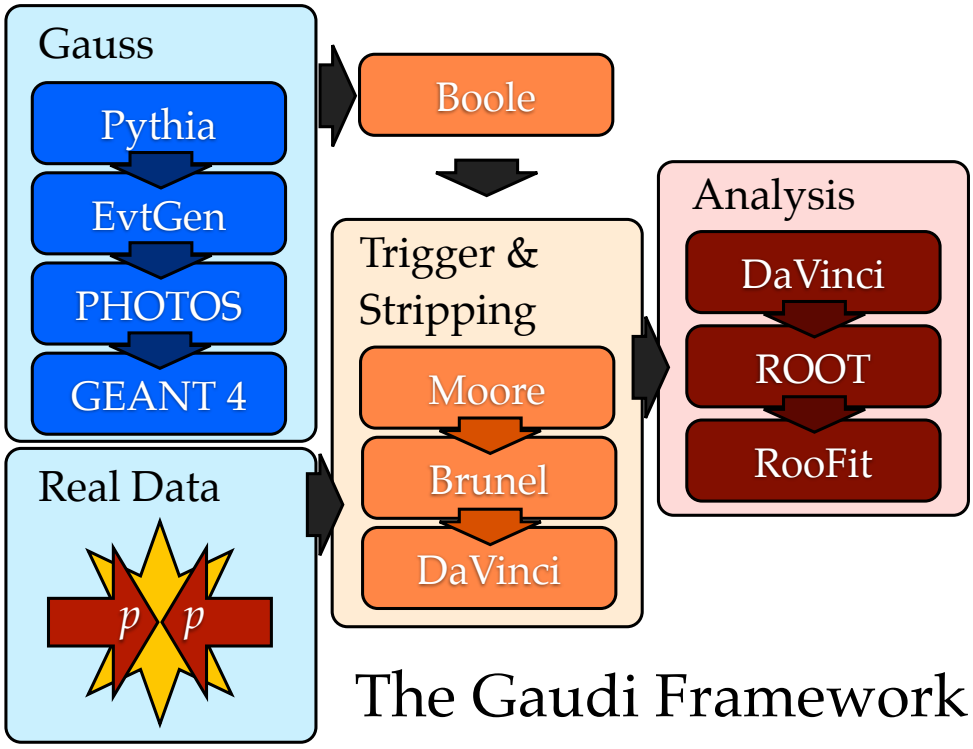


Fig. 3.11: Schematical view of the LHCb software packages for simulation (blue), as well as offline (orange) and online (red) analysis.

For the simulation studies, *pp* collisions are generated using PYTHIA 6.4 [47] with a specific LHCb configuration [48]. Decays of hadronic particles are described by EVTGEN [49] in which final state radiation is generated using PHOTOS [50]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [51, 52] as described in Ref. [53]. The software package Boole is used as an interface between the simulation software and the usual analysis chain that is also used for real data. It generates a simulated electrical response of the detector that is similar to the response to real data.

The online analysis tools include the software that runs during data taking. While

the trigger software Moore is used to decide which events are written to the data storages, the reconstruction using Brunel and the stripping selections with DaVinci can be redone on the saved datasets. As both is a major effort for the full datasets, these steps are redone only once or twice in a year when the LHC is not running. In the stripping procedure candidates for all decay channels that are to be studied are pre-selected, in order to reduce the CPU usage for the final (offline) selections.

For the final selection the DaVinci package is used again, as it provides methods to select the data with tighter selection criteria but following the same algorithms as the stripping. The selected n-tuples (n variables per candidate) are then processed using software based on the ROOT [54] and RooFit [55] packages. Minimizations for all likelihood fits are done by MINUIT [56]. The analysis code is based on a joint software package developed for different analyses with participation of the Dortmund working group.

4 An introduction to flavour tagging

Measurements of neutral B meson mixing or time dependent *CPV* in the decays of such mesons rely on the knowledge of the production flavour $d = +1$ ($d = -1$) of the B meson that contains a b ($\bar{b}$) quark. For neutral B mesons, d is not accessible in a trivial way, as only the flavour of the decay can be measured directly, if the final state is a self tagging final state, i.e. the final state f is only accessible from the B decay, not from the $\overline{B}$ and vice versa. The decay $B^0 \to J/\psi K^{*0}$ is such a decay. In the case of K^{*0} decaying to $K^+\pi^-$ the final state particles originate from a B^0 meson, while in the case of a decay $\overline{K}^{*0} \to K^-\pi^+$ the final state tags a decaying $\overline{B}^0$ meson. In $B^0 \to D^-\pi^+$ decays the kaon coming from the $D \to K\pi\pi$ decay reflects f.

The flavour tagging algorithms provide the information about the production flavour using similar techniques. As b quarks are produced in pairs where mostly both b and $\bar{b}$ quark are within the detector acceptance, the second quark can be used to determine the initial flavour of the signal particle (i.e. containing a $\bar{b}$ quark). The algorithms exploiting this are categorised as opposite side taggers (OST), because the tag is formed from information about the acompaniing hadron that contains the b quark.

Furthermore, it is possible to gain information about the signal $\bar{b}$ quark through particles that are produced in the hadronisation process. When a B^0 meson ($\bar{b}d$) is formed, an additional $\bar{d}$ quark becomes available, as it can also only be produced in $d\bar{d}$ pairs. This does now likely form a pion which has a 50 % chance of being a π^+. If now this pion can be associated to the B^0 production vertex, its charge tags the B^0 initial flavour. Note, that pions can also come from decays of excited states like B^*. In this case, the pions also give a tag on the B^0 originating from a B^* decay. This method is referred to as same side tagger (SST). Similarly, for B_s^0 mesons the SST uses a kaon instead of a pion, as from the hadronisation an additional $\bar{s}$ quark becomes available. In all cases the tag is inferred from the charge of the particle that is used to tag the event. A schematical overview of the tagging algorithms is given in Fig. 4.1.

The tagging algorithms provide additional information about the probability ω of a tag decision to be wrong. This may happen due to particles that are wrongly selected as tagging particles or also due to the intrinsic dilution of OS tags that use neutral B mesons and are thus biased by their oscillation. The knowledge of the wrong tag probability is essential, as the measurements of *CPV* rely on this information, because the tagging dilution $D = 1 - 2\omega$ enters the time dependent amplitudes in the same way as *CP* violation does. Thus, the tagger estimate η for the wrong tag probability

ω has to be calibrated carefully. The calibration procedure will be described later in this chapter. The following sections contain a summary of [15] and [16]. A detailed description of the tagging algorithms and the combination procedure can be found there.

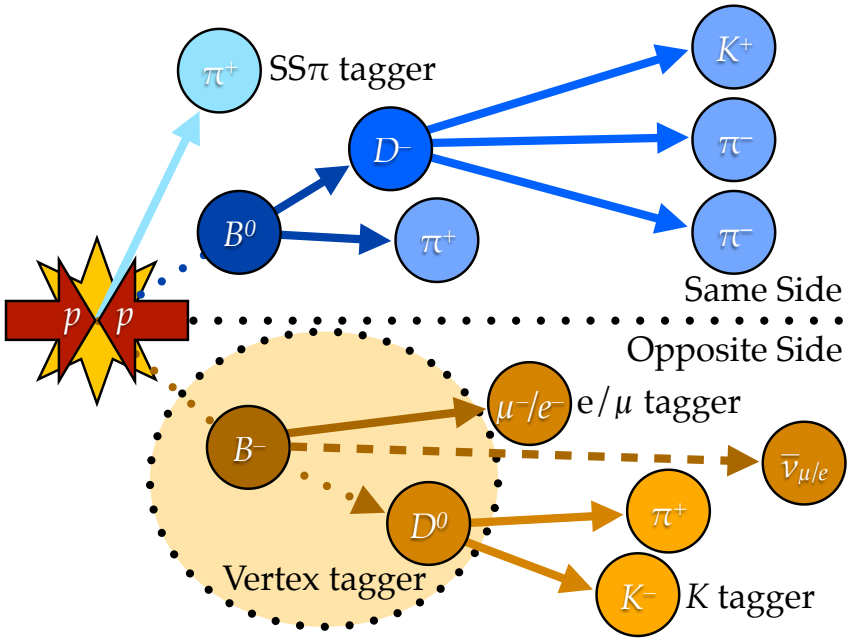


Fig. 4.1: Illustration of the tagging algorithms with a signal decay $B^0 \to D^- \pi^+$.

4.1 Opposite side tagging

To select candidates for the tagging algorithms, the particles are required to not belong to the signal decay and to be well separated from the signal tracks. Additional selection criteria are applied to ensure the tagging candidates do not originate from any additional primary interaction in the event. In the case of more than one candidate for a certain tagger passing the pre-selection and the individual selection for that tagger, the candidate with the highest p_{T} is chosen.

Nearly all b hadrons decay via the decay chain $b \to c \to s$, where a large fraction has a charged kaon in the final state. The charge of that kaon then tags the initial state. For the OS kaon tagger only candidates with a $p_{\mathrm{T}} > 0.5$ GeV and a large impact parameter with respect to the PV associated to the signal and to any PVs from pile-up are used. Furthermore, information from the PID system is used to suppress possible contamination from pions or protons.

About 10 % of the b hadrons decay via semileptonic decay channels with an electron or a muon in the final state. The electron tagger uses the electrons from decays

$b \to c\ e^-\ \bar{\nu}_e$. Tagging candidates for the electron tagger are required to have a $p_{\mathrm{T}} > 1.2$ GeV. Additional requirements include the PID and a hit on the ECAL as well as an $E/p > 0.8$.

Following the same principle as the electron tagger, the muon tagger exploits decays $b \to c\ \mu^-\ \bar{\nu}_\mu$. In order to reduce background from $b \to c \to s\ \mu^+\ \nu_\mu$ decays that include an oppositely charged muon, only muons with a $p_{\mathrm{T}} > 1.2$ GeV are used. Additionally, criteria on particle ID are applied. As for the semileptonic decays it is not clear if the leptons originate from a $b \to c\ \ell^-\ \bar{\nu}_\ell$ or a $b \to c \to s\ \ell^+\ \nu_\ell$ decay, the tags on the second decay chain are another source of wrong tags.

The vertex charge tagger forms a tag decision from an inclusive reconstruction of an additional secondary vertex (w.r.t. the signal vertex). In order to do so, a seed consisting of two tracks is formed from the combination of all tracks in the event that fulfil requirements on the track quality and geometrical cuts. From all possible combinations the one seed is selected that has the highest probability to come from a b hadron decay. This probability is evaluated using a likelihood method or multivariate methods trained with simulated events. Then additional tracks are added to the inclusive secondary vertex which have a small IP and a small distance of closest approach w.r.t. the seed. The vertex charge Q_{Vtx} is then calculated from a p_{T} weighted sum of the charges of the tracks coming from the inclusive secondary vertex

$$Q_{\mathrm{Vtx}} = \frac{\sum_i {p_{\mathrm{T}}}^k(i) Q_i}{\sum_i {p_{\mathrm{T}}}^k(i)}, \tag{4.1}$$

where the parameter k is optimised to maximise the tagging power. Finally, only vertices with a $|Q_{\mathrm{Vtx}}| > 0.25$ are considered as taggers and additional cuts on the mass and the momentum of the inclusive secondary vertex are applied. The sign of Q_{Vtx} then is the tag on the signal b hadron.

4.2 Same side tagging

When a $b\bar{b}$ pair is produced and it hadronises, the additional quarks are produced in pairs from the vacuum. Thus, when a B^0 (B_s^0) is produced there is an additional $\bar{d}$ ($\bar{s}$) available at the same place, that is likely to form a charged pion (kaon). These particles then provide a tag of the initial b quark flavour.

In the case of the SS pion tagger, the charged pion can also come from excited B mesons that decay to the state that is selected as a signal B candidate. For all tagging candidates the same selection criteria are applied as for the SS pion tagger, and they have to fulfil stronger cuts on the PID. Furthermore, the difference between the reconstructed invariant mass of the B meson and the invariant mass of the $B\pi$ combination is required to be smaller than 1.5 GeV.

For the tagging of B_s^0 signal decays, the SS kaon tagger is used. It works in a very similar manner as the SS pion tagger.

4.3 Tagging efficiency

The statistical significance of measurements of time dependent *CPV* depends on two factors connected to flavour tagging: the tagging efficiency ε_{tag}, and the wrong tag probability ω. The sensitivity on such measurements then scales with the square root of the effective tagging efficiency $\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} \cdot D^2$. The tagging efficiency can be calculated from the number U of untagged events and from W and R, the numbers of wrong and right tags, respectively

$$\varepsilon_{\text{tag}} = \frac{R+W}{U+R+W}. \tag{4.2}$$

The dilution is proportional to the asymmetry amplitude for mixing asymmetries (cf. Eq. (4.6)) or *CP* asymmetries that are constructed in a similar way. Thus, especially for measurements of time dependent *CPV* an exact knowledge of D is crucial.

4.4 Estimation of wrong tag probabilities

For every single tagger t an artificial neural net (NN) is trained to be able to reflect the wrong tag probability ω^t for a certain event. The neural nets are modelled by multi-layer perceptrons (MLP) as implemented in ROOT [54]. They consist of three layers with one hidden layer (cf. Fig 4.2). NNs are a powerful tool to discriminate events on a stochastical base. As in most cases only parts of the tagging B meson can be reconstructed, a NN allows to exploit a variety of attributes of the reconstructed particles as well as global event information to calculate the wrong tag probability.

The NNs are trained with MC simulated events, exploiting the true tag information to define wether the events are correctly or wrongly tagged. The NN input consists of several observables of the tagging candidates and the event itself. Since the distributions of some of these inputs show deviations between data and MC simulated events, the NNs have to be calibrated with data using control channels. For the calibration decays of charged and uncharged B mesons with self-tagging final states can be used. The neural net output nn^t for the tagger t is in the interval $[0, 1]$, where $nn^t = 0$ corresponds to a wrong tag and $nn^t = 1$ to a correct tag, respectively. This output is then mapped to values of $\eta^t \in [0, 0.5]$.

Differences between the MC simulated events and real data taken by the experiment lead to slightly incorrect estimates of the wrong tag probabilities η^t. For every tagger

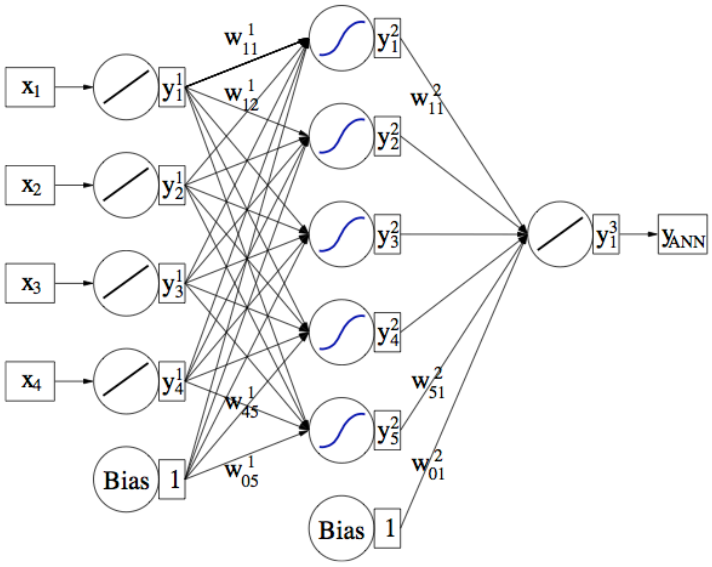


Fig. 4.2: Illustration of an MLP neural network with one hidden layer taken from [57].

a calibration function

$$\eta_c^t = p_0^t + p_1^t(\eta^t - \langle\eta^t\rangle) \tag{4.3}$$

with two free parameters p_0^t and p_1^t is implemented to reflect a wrong tag probability η_c^t calibrated on data. The value of $\langle\eta^t\rangle$ is fixed to the mean of the distribution of η^t to minimise the correlation between p_0 and p_1. Depending on the tagger, second and third order terms can be added to the calibration function. The same procedure is used to verify the calibration with data as described in Sec. 4.5.

If a signal B is tagged by more than one tagger, the decisions and wrong tag probabilities are combined. The combination can be calculated by

$$\mathcal{P}(b) = \frac{p(b)}{p(b) + p(\bar{b})}, \tag{4.4}$$

where $\mathcal{P}(b)$ is the probability for the B candidate to contain a b quark instead of a $\bar{b}$ and $p(b)$ and $p(\bar{b})$ are given by

$$p(b) = \prod_t \left(\frac{1+d^t}{2} - d^t(1-\eta_c^t)\right) \qquad p(\bar{b}) = \prod_t \left(\frac{1-d^t}{2} + d^t(1-\eta_c^t)\right), \tag{4.5}$$

calculated from the single tag decisions. Thus, for $\mathcal{P}(b) > 0.5$ the combined tagging decision is $d = -1$ and $\eta_c = 1 - \mathcal{P}(b)$ and also for $\mathcal{P}(b) < 0.5$ it follows $d = +1$ and $\eta_c = \mathcal{P}(b)$.

4.5 Methods for tagging calibration

The calculation of the combined tagging response from Eq. (4.4) does not include possible correlations between the different taggers. This assumption is a valid approximation for the SS tagger with respect to the OS taggers, since the criteria applied to select the tagging particle are complementary. Though, in the case of the combination of OS taggers a correlation is expected. As a consequence the combined wrong tag probability in Eq. (4.4) is a slightly overestimated and requires an additional calibration that was done using $B^+ \to J/\psi K^+$ events on data.

For measurements of *CPV*, the goodness of the calibration is additionally verified on data. In every case it can be decided, whether the tagging response should be corrected using the result of a calibration measurement or a systematic uncertainty has to be assigned. The following method is also used for the calibration of the single taggers as described in Eq. (4.3).

In order to be sufficiently sure, the tagging calibration is similar for different decay channels, the tagging response has to be tested in several kinematically different decay modes. Using data, there is clearly a difference in the analysis of charged and uncharged B meson decays. For decays of charged B mesons the final state is self tagging by definition, as the sum of the charges of all daughter particles will always be ± 1. Thus, the tag has to be compared to the final state flavour and it is possible to simply count the number of (in)correctly tagged candidates. These studies are performed for the training of the neural nets on data and for tagging calibration but they are not part of the work for this thesis.

A measurement of the tagging quality in decays of neutral B mesons can be performed by analysing the amplitude of the mixing asymmetry

$$\mathcal{A}_{\text{mix}} = \frac{N_{\text{unmixed}} - N_{\text{mixed}}}{N_{\text{unmixed}} + N_{\text{mixed}}} = D \cos(\Delta m t) \tag{4.6}$$

of self-tagging final states. Self-tagging means, that on tree level the assignment $B \to f$ and $\overline{B} \to \bar{f}$ is unambiguous, except for the $B\overline{B}$ mixing. The tagging dilution D can then be measured from the mixing asymmetry amplitude and the mixing frequency Δm is also accessible. The influence of the decay time resolution is discussed in Sec. 9.1 and negligible for B^0 decays.

The knowledge of the tagging dilution D is crucial for measurements of *CPV* as it is a linear factor to the decay rate asymmetry, as in the measurement of $\sin 2\beta$ [58]. Thus, for measurements of *CPV* it is necessary to find a kinematically similar decay mode to the signal decay that is used for those measurements. For a measurement of $\sin 2\beta$ in $B^0 \to J/\psi K^0_{\text{S}}$ decays the calibration channel is $B^0 \to J/\psi K^{*0}$.

In the work for this thesis tagging calibration measurements are performed in decays of $B^0 \to J/\psi K^{*0}$ and $B^0 \to D^- \pi^+$. The specialities of these decay modes are discussed in Chapter 5.

5 Self tagging decay channels

In the following, the two decay channels for tagging calibration with B^0 mesons and a measurement of Δm_d are introduced. These are the charmonium mode $B^0 \to J/\psi\, K^{*0}$, and the hadronic mode $B^0 \to D^-\pi^+$. Each of them offers different opportunities and brings in different challenges to the analysis strategies. This chapter will focus on the peculiarities of the different decays and their impact on the analysis. Additional tagging calibration studies have been performed with the semileptonic mode $B^0 \to D^{*-}\mu^+\nu_\mu$ [16]. These are not a part of this thesis.

5.1 The charmonium mode $B^0 \to J/\psi\, K^{*0}$

The decay of the neutral B^0 meson to a J/ψ and a K^{*0} has a lower yield compared to $B^0 \to D^-\pi^+$ decays. Though, due to the two muons in the final state and thus the clear signature, this channel yields very low systematic uncertainties in a measurement of Δm_d.

The two muons in the final state are responsible for a good reconstruction efficiency and a clear trigger signal. The combinatorial background is low compared to decays with only hadrons in the final state. Due to the CKM matrix elements V_{cb} and V_{cs} that are involved in the decay (cf. Fig. 5.1), the branching ratio is quite high but it is clearly suppressed by the probability to form a K^{*0} meson from the $d\bar{s}$ combination. The K^{*0} meson is crucial as it tags the flavour of the B^0 meson at the time of decay. A decay $K^{*0} \to K^+\pi^-$ indicates that the K^{*0} meson comes from a B^0, whereas a K^- meson in the final state belongs to a decaying $\overline{B}^0$.

In this decay channel the final state is self-tagging, meaning that on tree-level there is no possibility for a B^0 meson to decay into another final state than the one including a K^+ particle. Thus, the mixing asymmetry for this decay is given by

$$\mathcal{A}_{\text{mix}}(t) = \frac{N_{\text{unmixed}}(t) - N_{\text{mixed}}(t)}{N_{\text{unmixed}}(t) + N_{\text{mixed}}(t)} = \cos(\Delta m t), \tag{5.1}$$

where $N_{\text{mixed/unmixed}}(t)$ refers to the number of B^0 candidates that have/have not oscillated at the given time t. Eq. (5.1) follows directly from Eqs. (2.34) - (2.37).

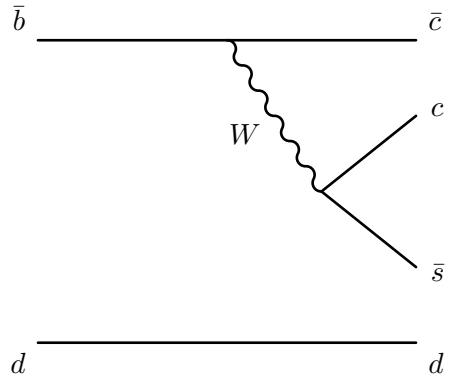


Fig. 5.1: Feynman diagram for tree-level decays of $B^0 \to J/\psi K^{*0}$.

5.2 The hadronic mode $B^0 \to D^-\pi^+$

Although, the decay $B^0 \to D^-\pi^+$ is quite similar to $B^0 \to J/\psi K^{*0}$ some more specifics have to be considered. First of all, the event yield is higher, due to the high probability for B mesons to decay into a D meson. $B^0 \to D^-\pi^+$ is the decay with the highest branching ratio of all hadronic B^0 decays and a final state that consists of charged particles only. A part of this advantage is lost by the worse reconstruction efficiency. This is due to the final state with three pions and only one kaon from the $D^\pm \to K^\mp\pi^\pm\pi^\pm$ decay. As the primary vertex (PV) background mostly consists of kaons and pions, the selection has to filter a large number of combinatorial background originating from the PV. This is done using the B^0 flight distance. The price for the background reduction is thus a bias on the exponential shape of the B^0 decay time distribution.

On the other hand, many $B^0 \to D^-\pi^+$ candidates pass the trigger without the signal tracks being used. That means that the tracks belonging to the second b hadron are potentially better reconstructed as they were already used in the trigger. This is a big advantage for the flavour tagging algorithms. Thus, the overall tagging quality in this channel is better than for $B^0 \to J/\psi K^{*0}$ candidates.

The decay channel $B^0 \to D^-\pi^+$ is not completely self-tagging (cf. Fig. 5.2). In addition to the CKM favoured decay in which a K^+ particle in the final state belongs to a B^0 meson as for $B^0 \to J/\psi K^{*0}$ candidates, there is a CKM suppressed mode which leads to a charge conjugated final state. The suppression of the CKM suppressed matrix element with respect to the favoured one is of the order λ^2. Two effects result from this fact. Firstly, the decay channel becomes sensitive to ínterference *CPV*, due to the final state being accessible from B^0 as well as $\overline{B}^0$ (cf. [28] and $B_s^0 \to D_s^+K^-$ analysis [59]). Secondly, the CKM suppressed mode introduces a dilution to the mixing asymmetry that potentially biases measurements of the tagging quality.

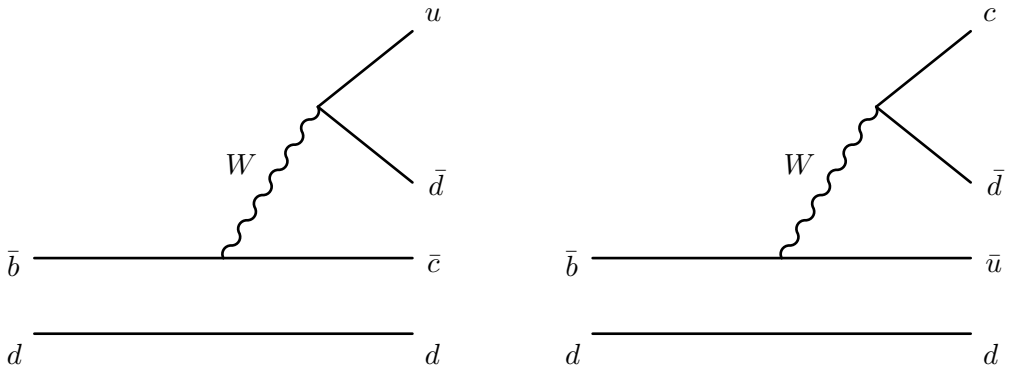


Fig. 5.2: Feynman diagrams for tree-level decays of $B^0 \to D^- \pi^+$.

When analysing the Eqs. (2.55) to (2.58), it is evident, that all sine terms cancel when summarising mixed ($\langle f|\overline{B}^0\rangle, \langle \bar{f}|B^0\rangle$) and unmixed ($\langle \bar{f}|\overline{B}^0\rangle, \langle f|B^0\rangle$) events respectively. The mixing asymmetry in this decay channel is then

$$\mathcal{A}_{\text{mix}}(t) = \frac{N_{\text{unmixed}}(t) - N_{\text{mixed}}(t)}{N_{\text{unmixed}}(t) + N_{\text{mixed}}(t)} = \frac{1 - |\lambda|^2}{1 + |\lambda|^2} \cos(\Delta m t). \tag{5.2}$$

Herein, the fraction including the $|\lambda|^2$ terms comes from the contribution of the CKM suppressed decays. As $|\lambda|^2 \approx \lambda_{\text{C}}^4$, the dilution on the asymmetry amplitude is less than three permille. It is negligible at the current level of statistics, but it will have to be included as a systematic uncertainty on the tagging calibration with $B^0 \to D^- \pi^+$ decays in future measurements.

6 Datasets and selection

The data used in this analysis are taken in the LHC runs with pp colisions at a center-of-mass energy of 7 TeV during the year 2011. That corresponds to an integrated luminosity of $1.0\,\mathrm{fb}^{-1}$. Fig. 6.1 shows an overview of the machine's and the detector's performance during the last years.

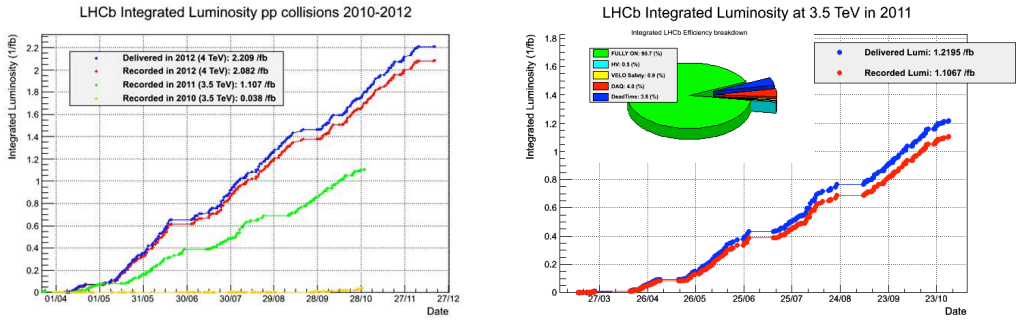


Fig. 6.1: Performance[1] of LHC and the LHCb detector for all data taking periods (left) and for the year 2011 (right). Only data taken in 2011 is used for the analyses in this thesis.

6.1 Observables

The decay time t of a B^0 candidate is evaluated from the measured momenta and from a vertex fit [60] that constrains the B^0 candidate to originate from the associated PV. No mass constraints on the intermediate resonances are applied. For the calculation of the invariant mass m, no mass constraints are used in the $B^0 \to D^-\pi^+$ channel, while the J/ψ mass is constrained to the world average[2] in the analysis of the decay $B^0 \to J/\psi\, K^{*0}$.

[1] The plots are taken from the LHCbOperations twiki on Feb. 04$^{\mathrm{th}}$, 2013
http://lbweb.cern.ch/groups/online/OperationsPlots/index_files/
IntegratedLumiLHCbTime_Yearly.png
2012IntegratedLumiLHCbTime.png

[2] Whenever a *world average* for a particle property is mentioned, it is taken from [7].

6.2 Trigger and selection for $B^0 \to D^-\pi^+$

Events including $B^0 \to D^-\pi^+$ decays are required to have tracks with high transverse momenta p_{T} to pass the hardware trigger. The software trigger requires a two-, three- or four-track secondary vertex with a large sum of the p_{T} of the tracks, significant displacement from the associated primary vertex (PV), at least one track with $p_{\mathrm{T}} > 1.7\,\mathrm{GeV}$ and a large impact parameter with respect to that PV, and a good track fit. A multivariate algorithm is used for the identification of the secondary vertices [61] on trigger level.

The kaon and pion candidates are separated by the particle ID system, exploiting the different Cherenkov angles in the RICH detectors. One kaon and two pions with a charge opposite of the kaon charge then form a D^- candidate. The $B^0 \to D^-\pi^+$ selection requires that the D^- reconstructed mass is in a range of $\pm 100\,\mathrm{MeV}$ around the world average. Furthermore, the D^- decay vertex is required to be downstream of the PV associated to the B^0 candidate that is formed from the D^- and a bachelor pion.

The sum of the D^- and π^+ transverse momenta must be larger than $5\,\mathrm{GeV}$. The B^0 candidate invariant mass must be in the interval $5000 \leq m(K^+\pi^-\pi^-\pi^+) < 5700\,\mathrm{MeV}$. Additionally, the cosine of the pointing angle between the B^0 momentum vector and the line segment between PV and secondary vertex is required to be larger than 0.999.

Candidates are classified by a boosted decision tree (BDT) [62, 63] with the so-called AdaBoost algorithm[64].[3] The BDT is trained with $B_s^0 \to D_s^-\pi^+$ candidates with no particle ID criteria applied to the daughter pions and kaons. The cut on the BDT classifier is optimised in order to maximise the significance of the $B^0 \to D^-\pi^+$ signal. Several input variables are used: the impact parameter (IP) significance, the flight distance perpendicular to the beam axis, the vertex quality of the B^0 and the D^- candidate, the angle between the B^0 momentum and the line segment between PV and B^0 decay vertex, the angle between the D^- momentum and the line segment between PV and the D^- decay vertex, the angle between the D^- momentum and the line segment between the B^0 decay vertex and D^- decay vertex, the IP and p_{T} of the π^+ track, and the angle between the π^+ momentum and the line segment between PV and B^0 decay vertex. Only B^0 candidates with a decay time $t > 0.3\,\mathrm{ps}$ are accepted.

To suppress potential background from misidentified kaons in $D_s^- \to K^-K^+\pi^-$ decays, all D^- candidates are removed if they have a daughter pion candidate that might pass a loose kaon selection and are within a $\pm 25\,\mathrm{MeV}$ mass window around the D_s^- mass (the D^- mass resolution is smaller than $10\,\mathrm{MeV}$) when that pion is reconstructed under the kaon mass hypothesis.

[3] The development and training of the BDT was performed by A. Dziurda.

An overview of the full set of the selection cuts and the BDT input variables can be found in Tables 6.1 and 6.2.

6.3 Trigger and selection for $B^0 \to J/\psi K^*$

Events in the decay $B^0 \to J/\psi K^{*0}$ are first required to pass a hardware trigger which selects a single muon with $p_T > 1.48\,\text{GeV}$. In the subsequent software trigger [61], at least one of the final state particles is required to have $p_T > 0.8\,\text{GeV}$ and a large IP with respect to all PVs in the event. Finally, the tracks of two or more of the final state particles are required to form a vertex which is significantly displaced from the PVs in the event.

A K^{*0} candidate is formed by the combination of a kaon and a pion that are separated by the PID system. It is required, that the K^{*0} candidate has a $p_T > 2\,\text{GeV}$ and $826 \leq m(K^+\pi^-) < 966\,\text{MeV}$.

The unconstrained $\mu^+\mu^-$ invariant mass must be within $\pm 80\,\text{MeV}$ of the world average for the J/ψ mass. B^0 candidates are required to have a large IP with respect to other PVs in the event and the B^0 decay vertex must be significantly separated from the PV. Additionally, B^0 candidates are required to have a reconstructed decay time $t > 0.3\,\text{ps}$ and an invariant mass in the range $5230 \leq m(J/\psi K^+\pi^-) < 5330\,\text{MeV}$. To suppress potential background from misidentified $B_s^0 \to J/\psi\,\phi$ decays, all candidates are removed for which the $K^+\pi^-$ mass is within a $\pm 10\,\text{MeV}$ window around the nominal $\phi(1020)$ mass when computed under the kaon mass hypothesis for the pion.

The resulting mass distributions for the two decay channels are shown in Fig. 6.2. A summary of the stripping and final selection criteria can be found in Table 6.3.

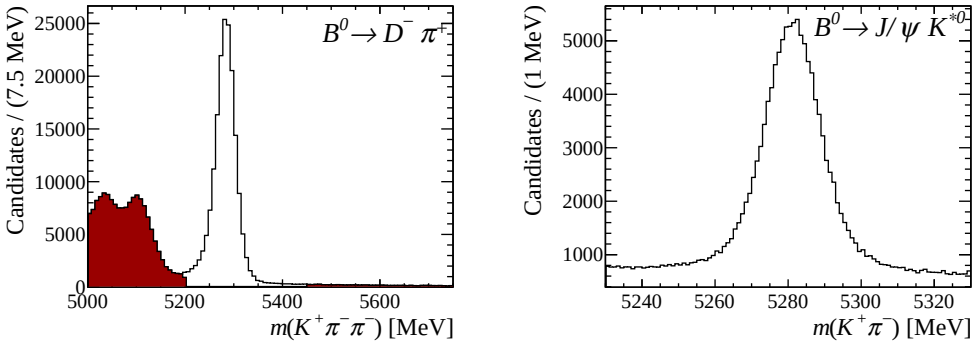


Fig. 6.2: Invariant mass distributions of the two decay channels (left) $B^0 \to D^-\pi^+$ and (right) $B^0 \to J/\psi K^{*0}$. The filled areas correspond to the invariant mass range excluded in the fits to Δm_d.

Tab. 6.1: Pre-selection for $B^0 \to D^- \pi^+$.

Cuts on B^0 candidate	
sum of input particles' transverse momenta	$> 5000\,\mathrm{MeV/c}$
χ^2 of primary vertex	< 25
primary vertex χ^2/ ndof	< 10
cos of $\sphericalangle[\lvert PV, B^0\mathrm{Vtx}\rvert, \vec{p}(B^0)]$	> 0.999
decay time	$> 0.2\,\mathrm{ps}$
$m(K\pi\pi\pi)_{\mathrm{min}}$	$4750\,\mathrm{MeV/c^2}$
$m(K\pi\pi\pi)_{\mathrm{max}}$	$5800\,\mathrm{MeV/c^2}$
Cuts on D^- candidate	
sum of input particles' transverse momenta	$> 1800\,\mathrm{MeV/c}$
vertex χ^2/ndof	< 10
best χ^2 of primary vertex	> 36
cos of $\sphericalangle[\lvert PV, D^+\mathrm{Vtx}\rvert, \vec{p}(D^+)]$	> 0
$m(K\pi\pi) - m(D^{\pm})_{\mathrm{PDG}}$	$< 100\,\mathrm{MeV/c}$
Cuts on K^+ and $\pi^{\pm}$ candidates	
DLL$(K-\pi)$ for pions	< 20
DLL$(K-\pi)$ for kaons	> -10
transverse momentum	$> 100\,\mathrm{MeV/c}$
momentum	$> 1000\,\mathrm{MeV/c}$
track χ^2 / ndof	< 4
minimum impact parameter χ^2	> 4

Tab. 6.2: Parameter list for the BDT selection of $B^0 \to D^- \pi^+$.

B^0-candidate	D^0-candidate
minimum IP χ^2	minimum IP χ^2
transversal distance to PV	transversal distanz to PV
vertex χ^2/ndof	vertex χ^2/ndof
lifetime fit vertex χ^2/ndof	$\sphericalangle[\lvert PV, B^0\mathrm{Vtx}\rvert, \vec{p}(D^-)]$
$\sphericalangle[\lvert PV, B^0\mathrm{Vtx}\rvert, \vec{p}(B^0)]$	$\sphericalangle[\lvert B^0\mathrm{Vtx}, D^-\mathrm{Vtx}\rvert, \vec{p}(D^-)]$
bachelor π	All π and $K^{\pm}$
minimum IP χ^2	ghost probability
transverse momentum	
cos of $\sphericalangle[\lvert PV, B^0\mathrm{Vtx}\rvert, \vec{p}(\pi)]$	

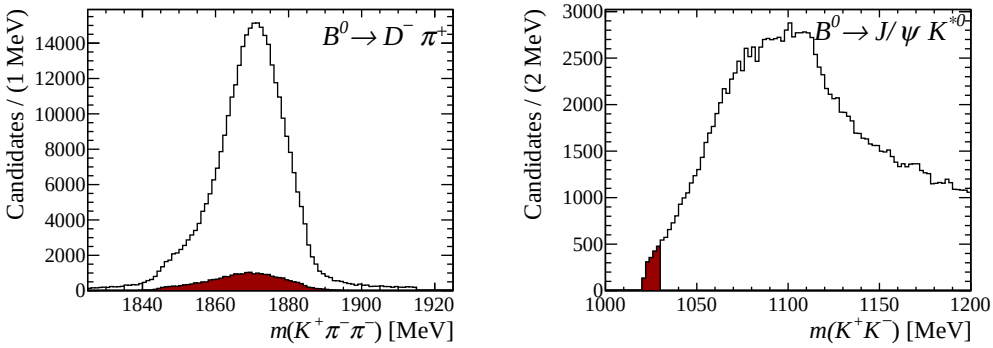


Fig. 6.3: (Left) Spectrum of $D^- \to K^+\pi^-\pi^-$ invariant mass with candidates not passing the D_s^+ veto shown in the filled histogram. (Right) Spectrum of pseudo ϕ (K^+K^-) mass when calculating the K^* invariant mass with a kaon mass hypothesis for the pion. The filled area shows candidates that do not pass the phi veto. Both plots show the data after applying the pre-selection before the final selection.

Tab. 6.3: Pre-selection and final selection for $B^0 \to J/\psi K^{*0}$. The next best PV (c.f. Cuts on B^0 candidate) is the one with the next best χ^2 w.r.t. the chosen PV.

	Stripping 17a	Final
Cuts on B^0 candidate		
vertex χ^2 / ndof	< 10	< 10
impact parameter χ^2 / ndof	–	< 5
impact parameter χ^2 w.r.t. nextbest PV	–	< 50
distance of flight χ^2 / ndof	–	< 5
decay time	$> 0.2\,\mathrm{ps}$	$> 0.3\,\mathrm{ps}$
$m(\mu\mu K\pi)_{\mathrm{min}}$	$5100\,\mathrm{MeV/c^2}$	$5230\,\mathrm{MeV/c^2}$
$m(\mu\mu K\pi)_{\mathrm{max}}$	$5400\,\mathrm{MeV/c^2}$	$5330\,\mathrm{MeV/c^2}$
Cuts on J/ψ candidate		
muon DLL μ, π	> 0	> 0
minimum transverse momentum (μ^+, μ^-)	–	$> 0.5\,\mathrm{GeV/c}$
vertex χ^2 / ndof (J/ψ)	< 16	< 16
$\lvert m(\mu\mu) - m(J/\psi)_{\mathrm{PDG}}\rvert$	$< 80\,\mathrm{MeV/c^2}$	$< 80\,\mathrm{MeV/c^2}$
Cuts on K^{*0} candidate		
kaon DLL K, π	> -2	> 0
kaon DLL K, p	–	> -2
pion DLL K,π	–	< 0
transverse momentum	$> 1\,\mathrm{GeV/c}$	$> 2\,\mathrm{GeV/c}$
$m(K\pi)_{\mathrm{min}}$	$826\,\mathrm{MeV/c^2}$	$826\,\mathrm{MeV/c^2}$
$m(K\pi)_{\mathrm{max}}$	$966\,\mathrm{MeV/c^2}$	$966\,\mathrm{MeV/c^2}$
vertex χ^2	< 16	< 16
all tracks		
track χ^2 / ndof	< 5	< 4
clone distance	–	> 5000

7 Fitting model for tagging and oscillation measurements

For the measurement of Δm_d and for the purpose of determining the tagging calibration parameters an unbinned extended maximum likelihood method is used. The minimisation of the negative log-likelihood is done by MINUIT [56] interfaced by the RooFit package [55]. As observables the reconstructed mass m, the decay time t and the mixing state q of the B^0 candidate are used in the fit. The extended likelihood function is defined as

$$\mathcal{L}(\vec{\lambda}) = \frac{\mathrm{e}^{-N} N^n}{n!} \prod_s \prod_{i=1}^{N_s} \mathcal{P}^s(\vec{x}_i; \vec{\lambda}_s). \tag{7.1}$$

Here, s is the number of simultaneous subsamples, $\mathcal{P}$ is the probability density function (pdf) with the observables x_i and the parameters $\vec{\lambda}_s$, and N, N_s, and n are the expected number of events, the number of events in one subsample and the sum of the events in all subsamples, respectively. $n = \sum_s N_s$ where N is estimated by the fit. The fit parameters for the two slightly different models in both decay channels are listed in sections 7.1 and 7.2.

To describe the signal and background components of the samples analysed, the following pdfs are used.

- Gaussian function with mean μ and width σ, used for peaks in the mass spectra

$$\mathcal{G}(m; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2} \tag{7.2}$$

- Crystal Ball function [65]: Gaussian function with an power law tail which needs two additional parameters α and ι. The tail can be used to describe candidates which have a lower invariant mass due to not reconstructed photons from final state radiation

$$\mathcal{CB}(m; \mu, \sigma, \iota, \alpha) \propto \begin{cases} \exp(-\frac{(m-\mu)^2}{2\sigma^2}), & \text{for } \frac{m-\mu}{\sigma} > -\alpha \\ A \cdot (B - \frac{m-\mu}{\sigma})^{-\iota}, & \text{for } \frac{m-\mu}{\sigma} \leqslant -\alpha \end{cases} \tag{7.3}$$

where

$$A = \left(\frac{\iota}{|\alpha|}\right)^{\iota} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \tag{7.4}$$

$$B = \frac{\iota}{|\alpha|} - |\alpha| \tag{7.5}$$

- Exponential function with parameter M to describe the combinatorial background in the mass spectra

$$\mathcal{X}(m; M) \propto \mathrm{e}^{\frac{m}{M}} \tag{7.6}$$

- Gaussian kernel pdf [66]: superposition of Gaussian functions to describe distributions extracted from MC or data more smoothly than with histograms.

- Decay pdf: exponential function with negative exponent cut off for negative decay times; may be convolved analytically with different decay time resolution models

$$\mathcal{D}(t; \tau) \propto \mathrm{e}^{-\frac{t}{\tau}} \mid t \geq 0 \tag{7.7}$$

- B^0 mixing pdf: same as Decay pdf but with two components for the two different mixing states q; a mean mistag rate ω and the mixing frequency Δm_d can be fit.

$$\mathcal{M}(t, q; \tau, \Delta m, \omega) \propto \mathrm{e}^{-\frac{t}{\tau}} \left(1 + q(1-2\omega)\cos\left(\Delta m_d t\right)\right) \mid t \geq 0. \tag{7.8}$$

  Note, that ω may be a parameter when using tagging categories or be replaced by (10.1) (cf. Sec. 10) in the case the per event wrong tag probabilities are used.

- Gaussian resolution: Gaussian resolution model to be convolved with the lifetime pdfs, using a mean decay time resolution s

$$\mathcal{R}(t; s) \propto \frac{1}{s\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{t}{s}\right)^2} \tag{7.9}$$

For describing the mass distributions, combinations of multiple Gaussian pdfs or one Crystal Ball function added to Gaussian pdfs are used. Combinatorial backgrounds are described using exponential functions.

Hadronic backgrounds in the $B^0 \to D^-\pi^+$ channel's mass spectra, are described by pdf shapes extracted from MC as described in Sec. 8.1.

7.1 Parametrisation for the $B^0 \to D^-\pi^+$ channel

In $B^0 \to D^-\pi^+$ a one dimensional fit to the mass distribution over the full range of $5000\,\mathrm{MeV} < m(B^0) < 5700\,\mathrm{MeV}$ is performed using Gaussian kernel pdfs from MC to ensure there is no peaking background under the signal peak. The nominal fit to Δm_d is then performed in the reduced mass range $5200\,\mathrm{MeV} < m(B^0) < 5450\,\mathrm{MeV}$ to avoid the hadronic background components.

The signal mass pdf is a Crystal Ball function added to a Gaussian pdf.

$$\begin{aligned}\mathcal{P}_{\mathrm{sig}}(m;\mu,\sigma_1,\sigma_2,\alpha,\iota,f_{m;\mathrm{sig}}) =& f_{m,\mathrm{sig}}\mathcal{CB}(m;\mu,\sigma_1,\iota,\alpha)+ \\ & (1-f_{m;\mathrm{sig}})\mathcal{G}(m,\mu,\sigma_2)\end{aligned} \tag{7.10}$$

The longlived combinatorial background (lbg) is described by an exponential pdf

$$\mathcal{P}_{\mathrm{lbg}}(m;\lambda_{\mathrm{lbg}}) = \mathcal{X}(m;\lambda_{\mathrm{lbg}}). \tag{7.11}$$

Two additional hadronic modes (cf. Figure 8.1) which are reconstructed in the given mass range are parametrised by Gaussian kernel pdfs from MC $\mathcal{P}_{\mathrm{bgi}}(m)$, where $\mathrm{i} \in 1,2$ denotes the number of the component.

For the decay time pdf only two components have to be used in the nominal fit to Δm_d, as the fit uses a reduced mass range. These are the signal and the longlived combinatorial background. For the signal the decay function including the mixing model (7.8) is used

$$\mathcal{P}_{\mathrm{sig}}(t,q;\tau,\Delta m_d,\omega,s_{\mathrm{sig}}) = \mathcal{M}(t,q;\tau,\Delta m_d,\omega)\otimes\mathcal{R}(t;s_{\mathrm{sig}}). \tag{7.12}$$

The exponential model for the combinatorial background component is then

$$\mathcal{P}_{\mathrm{lbg}}(t,q;\tau_{\mathrm{lbg}},s_{\mathrm{lbg}},\omega_{\mathrm{lbg}}) = \mathcal{M}(t,q;\tau_{\mathrm{lbg}},0,\omega_{\mathrm{lbg}})\otimes\mathcal{R}(t;s_{\mathrm{lbg}}). \tag{7.13}$$

This pdf is able to describe a non vanishing constant mixing asymmetry ω_{bkg} in the background (cf. Fig 11.1), to account for a possible asymmetry in the number of positive and negative tags included in the background events. Note, that this does not have a physical meaning. The full pdf for the two dimensional fit for Δm_d and tagging calibration is given by

$$\begin{aligned}\mathcal{P}_{B^0\to D^-\pi^+}(t,q,m;\vec{\lambda}) =& N_{\mathrm{sig}}\mathcal{P}_{\mathrm{sig}}(t,q;\tau,\Delta m_d,\omega,s_{\mathrm{sig}})\cdot\epsilon(t)\cdot \\ & \mathcal{P}_{\mathrm{sig}}(m;\mu,\sigma_1,\sigma_2,f_{m,\mathrm{sig}},\alpha,\iota)+ \\ & N_{\mathrm{lbg}}\mathcal{P}_{\mathrm{lbg}}(t;\tau_{\mathrm{lbg}},s_{\mathrm{lbg}},\omega_{\mathrm{lbg}})\cdot\mathcal{P}_{\mathrm{lbg}}(m;\lambda_{\mathrm{lbg}}),\end{aligned} \tag{7.14}$$

with the decay time acceptance function $\epsilon(t)$. This efficiency function is introduced by trigger and selection criteria (like cuts on the B^0 flight distance) and changes the exponential decay time distribution of the signal candidates.

In the nominal fit for Δm_d all parameters are floated. In the crosscheck fits for different tagging settings those parameters are fixed to the fitted values.

7.2 Parametrisation for the $B^0 \to J/\psi\, K^{*0}$ channel

For the $B^0 \to J/\psi\, K^{*0}$-channel the mass pdf does not contain contributions from peaking background in the mass spectrum. The background is treated as two longlived combinatorial components with different lifetimes.

The signal mass pdf is described as a triple gaussian

$$\begin{aligned}\mathcal{P}_{\mathrm{sig}}(m;\mu,\sigma_1,\sigma_2,\sigma_3,f^{12}_{m;\mathrm{sig}},f^{23}_{m;\mathrm{sig}}) =& f^{12}_{m;\mathrm{sig}}\mathcal{G}(m;\mu,\sigma_1)+ \\ & (1-f^{12}_{m;\mathrm{sig}})f^{23}_{m;\mathrm{sig}}\mathcal{G}(m;\mu,\sigma_2)+ \\ & (1-(1-f^{12}_{m;\mathrm{sig}})f^{23}_{m;\mathrm{sig}})\mathcal{G}(m;\mu,\sigma_3)\end{aligned} \tag{7.15}$$

and the combinatorial background is described by an exponential mass distribution

$$\mathcal{P}_{\mathrm{bg}}(m;\lambda_{\mathrm{bg}}) = \mathcal{X}(m;\lambda_{\mathrm{bg}}). \tag{7.16}$$

For the signal decay time the same parametrisation as in $B^0 \to D^-\pi^+$ can be used

$$\mathcal{P}_{\mathrm{sig}}(t,q;\tau,\Delta m_d,\omega,s_{\mathrm{sig}}) = \mathcal{M}(t,q;\tau,\Delta m_d,\omega)\otimes\mathcal{R}(t;s_{\mathrm{sig}}). \tag{7.17}$$

In the combinatorial background two lifetime components are observed. Two decay pdfs are therefore defined by

$$\mathcal{P}_{\mathrm{sbg}}(t;\tau_1,s_{\mathrm{sbg}}) = \mathcal{E}(t;\tau_1)\otimes\mathcal{R}(t;s_{\mathrm{sbg}}) \tag{7.18}$$

$$\mathcal{P}_{\mathrm{lbg}}(t,q;\tau_2,s_{\mathrm{lbg}},\omega_{\mathrm{lbg}}) = \mathcal{M}(t,q;\tau_2,0,\omega_{\mathrm{lbg}})\otimes\mathcal{R}(t;s_{\mathrm{lbg}}) \tag{7.19}$$

As it is seen on data in Fig 11.1 (cf. Sec. 11.1), a non vanishing tagging asymmetry is visible for one of the combinatorial background components. This can be described by a mixing pdf with $\Delta m_d = 0$. The drop of this asymmetry for low decay times leads to the assumption, that this effect is not existent for the background component with the shorter lifetime in $B^0 \to J/\psi\, K^{*0}$. That leads to the full two dimensional, two component pdf

$$\begin{aligned}\mathcal{P}_{B^0\to J/\psi\, K^{*0}}(t,q,m;\vec{\lambda}) =& N_{\mathrm{sig}}\mathcal{P}_{\mathrm{sig}}(t,q;\tau,\Delta m_d,\omega,s_{\mathrm{sig}})\cdot\epsilon(t)\cdot \\ & \mathcal{P}_{\mathrm{sig}}(m;\mu,\sigma_1,\sigma_2,\sigma_3,f^{12}_{m,\mathrm{sig}},f^{23}_{m,\mathrm{sig}})+ \\ & N_{\mathrm{sbg}}\mathcal{P}_{\mathrm{sbg}}(t;\tau_1,s_{\mathrm{sbg}})\cdot\mathcal{P}_{\mathrm{sbg}}(m;\lambda_{\mathrm{sbg}})+ \\ & N_{\mathrm{lbg}}\mathcal{P}_{\mathrm{lbg}}(t,q;\tau_2,s_{\mathrm{lbg}},\omega_{\mathrm{lbg}})\cdot\mathcal{P}_{\mathrm{lbg}}(m;\lambda_{\mathrm{lbg}}),\end{aligned} \tag{7.20}$$

where $\epsilon(t)$ denotes a decay time acceptance function introduced by trigger criteria. This effect is discussed in detail in Sec. 9.2.

The fitting model is tested using fully simulated signal MC with 500k (10M) generated events in $B^0 \to D^- \pi^+$ ($B^0 \to J/\psi K^{*0}$). The fitted values of Δm_d on MC are found to be

$$\Delta m_d^{B^0 \to J/\psi K^{*0}} = 0.506 \pm 0.003\,\mathrm{ps}^{-1} \tag{7.21}$$

$$\Delta m_d^{B^0 \to D^- \pi^+} = 0.523 \pm 0.014\,\mathrm{ps}^{-1} \tag{7.22}$$

at a generation value of $\Delta m_d = 0.502\,\mathrm{ps}^{-1}$.

8 Physics background

The reconstructed final states in both decay channels rely on a proper particle identification, especially the separation between kaons and pions. The vetos that suppress background from misidentified particles are introduced in Sec. 6. The procedure implementing the veto will be described in this chapter.

Additionally, possible physics background coming from missing particles (e.g. not reconstructed uncharged pions) is investigated. Such background contributions lead to peaks in the invariant mass distributions at masses lower than that of the initial particle. In the following, the studies concerning those hadronic background contributions are described separately for the two decay channels.[1]

8.1 Background in $B^0 \to D^-\pi^+$

In a study based on MC simulated events, the main contributions to peaking background near the $B^0 \to D^-\pi^+$ mass peak are found to come from decays including ρ and $D^{*-}(2012)$ mesons. In the MC samples, simulated events containing decays of B^0, $B^\pm$, B_s^0, and Λ_b^0 to at least one D meson are regarded. The sample sizes are 40M for the B^0 and $B^\pm$ samples and 10M each for the B_s^0 and Λ_b^0 samples, reflecting the measured hadronisation fractions for these particles at LEP [67] and LHCb [68]. Thus, the ratios of the measured number of candidates from the different channels are expected to represent those on data.

The only significant contributions of background come from the B^0 and Λ_b^0 samples. Three different channels are found to yield significant background after the preselection:

- $B^0 \to \rho^+(\pi^0\pi^+)D^-(K^+\pi^-\pi^-)$,
- $B^0 \to D^{*-}(2010)(\pi^0 D^-(K^+\pi^-\pi^-))\pi^+$, and
- $\Lambda_b^0 \to \Lambda_c^+(p\,\overline{K}^{*0}(K^-\pi^+))\pi^-$,

[1] The work presented in this chapter is a major part of the master thesis by U. Eitschberger *"Studies on systematic uncertainties in the measurement of the $B\overline{B}$ oscillation frequency Δm_d with the decay channel $B^0 \to D^-\pi^+$ at LHCb."* – September 2012

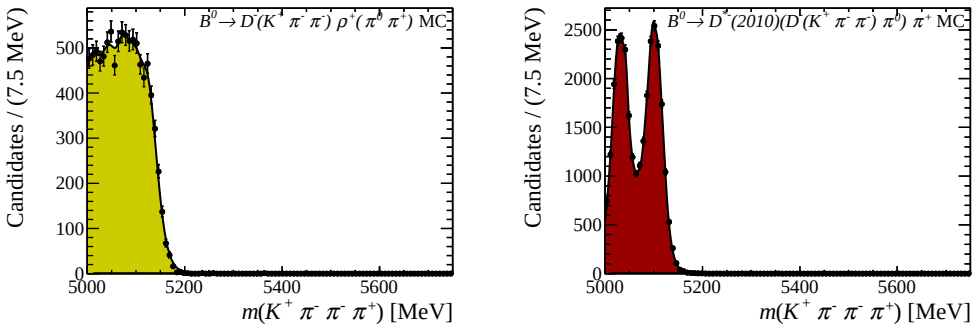


Fig. 8.1: The two figures show the background mass distributions extracted from MC simulated events without trigger requirements. Both distributions are used as templates for the fit to data. The double peak structure in the right plot is due to the spin ($s = 1$) of the D^* meson and the low mass difference $m(D^*) \approx m(D^-) + m(\pi^0)$ between the initial and final state. The $\cos^2\theta$ angular distribution of the $D^* \to D^-\pi^0$ decay then leads to a split-up in the partially reconstructed (with missing π^0) invariant mass of the B^0 candidate for the D^- emitted at $\theta \approx 0$ and $\theta \approx \pi$.

where the background from Λ_b^0 decays can be easily suppressed by harder PID cuts, to avoid reconstructing protons as pions. The partially reconstructed B^0 background persists, as it is not due to mis-id but missing π^0 particles. In order to show that this type of background does not peak below the signal, the shapes of the background components are taken from MC simulated samples that include large numbers of these decays. The shapes of the invariant mass distributions are then parametrised by Gaussian kernel pdfs [66]. The distributions and the Gaussian kernel pdfs can be seen in Fig. 8.1.

The Gaussian kernel pdfs are then used to fit the invariant mass spectrum of the $B^0 \to D^-\pi^+$ data sample in a wide mass range of $5000 \leq m(B^0) < 5750$ MeV. In addition to the two peaking background components, the pdfs for the signal mass peak and the longlived combinatorial background (c.f. Sec. 7.1) are used in a one dimensional, four component, extended maximum likelihood fit. As can be seen in Fig. 8.2, the model agrees very well with the data. No indication for missing background components is found. Additionally, a reduced mass range of $5200 \leq m(B^0) < 5450$ MeV for the fit to Δm_d will reduce the amount of the physics background components to a negligible number.

Although, no indication for contributions of background including D_s^+ mesons is found, a veto on $D_s^+ \to K^+K^-\pi^+$ decays is implemented. Such decays could enter the data sample due to misidentified kaons. If one of the pions cannot be properly separated from a kaon PID hypothesis, the invariant mass for the reconstructed $K\pi\pi$ combination is calculated, using a kaon mass hypothesis for that pion. If the result is within 25 MeV of the world average for the D_s^+ mass, the event is rejected.

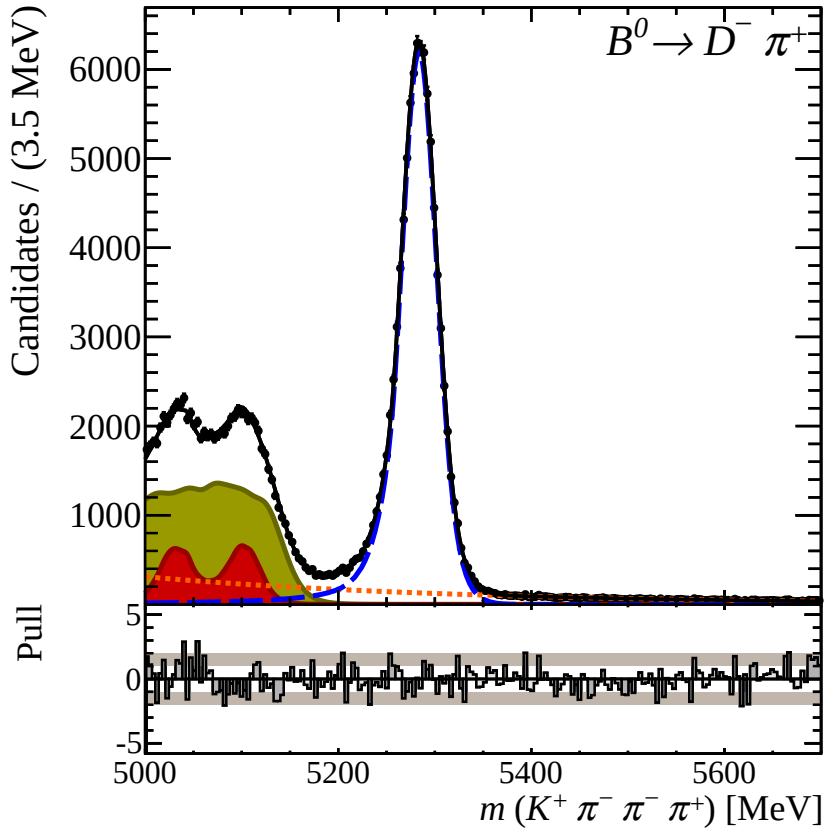


Fig. 8.2: Projection of the fit result (black line) of the one dimensional fit to the B^0 invariant mass spectrum using the background mass shapes extracted from MC simulated events. The shape parameters of the signal and the yields for signal and background components are free in the fit. Signal (blue dashed), longlived background (orange dotted), $B^0 \to D^{*-}(2010)(\pi^0 D^-(K^+\pi^-\pi^-))\pi^+$ (red filled area), $B^0 \to \rho^+(\pi^0\pi^+)D^-(K^+\pi^-\pi^-)$ (green filled area).

8.2 Background in $B^0 \to J/\psi\, K^{*0}$

In the $B^0 \to J/\psi\, K^{*0}$ channel there is no indication for any partially reconstructed background like in $B^0 \to D^-\pi^+$. Though, it might be possible to see tails of the mass peak of $B_s^0 \to J/\psi\,\phi$ decays in the case of incorrectly identified kaons. Thus, a veto on ϕ resonances is implemented similarly to the D_s^+ veto in $B^0 \to D^-\pi^+$. For each event the invariant mass of the $K\pi$ combination is calculated with a kaon mass hypothesis for the pion (cf. Fig. 6.3). All candidates with an invariant mass smaller than 1030 MeV are rejected.

9 Effects on the reconstructed decay time

9.1 Decay time resolution

For decays of B mesons the LHCb experiment has a decay time resolution of roughly 0.050 ps. A plain resolution effect on data without a bias means that candidates are reconstructed with a decay time that statistically fluctuates around the physical value, where the fluctuation is given by a resolution function (e.g. one or more Gaussians). For a simple decay this has the effect of a Gaussian tail at decay times smaller than zero. This effect becomes invisible due to the efficiency effects that are described below.

For the measurement of $B\overline{B}$ oscillations the resolution can additionally dilute the oscillation amplitude. Especially in measurements of interference CPV in B_s^0 mesons this can become an important factor as it is inline with the tagging dilution and the CP asymmetry amplitude. For $B^0\overline{B}^0$ oscillations however, the resolution is very small compared to the oscillation period of 12 ps. Thus, the influence of the decay time resolution on the measurement of the B^0 oscillation frequency and amplitude is insignificant (cf. Fig. 9.1). A fixed resolution of 0.050 ps (estimated from MC simulated events and background from the primary interaction) is used in all fits.

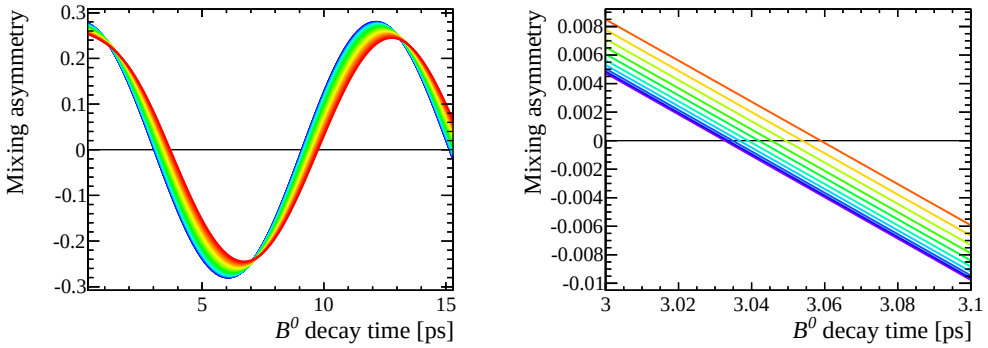


Fig. 9.1: Plot of a convolution of the mixing pdf with $\Delta m_d = 0.518\,\text{ps}^{-1}$ and a Gaussian resolution function for different resolution values. (Left) violet corresponds to a resolution of $\sigma_t = 0.02$ ps and red to $\sigma_t = 1.00$ ps. (Right) zoomed in to the first zero-crossing of the oscillation with violet corresponding to $\sigma_t = 0.02$ ps and red to $\sigma_t = 0.10$ ps. For the actual detector resolution of around $\sigma_t = 0.05$ ps the bias on the zero-crossing and thus on a measurement of Δm_d is at permille level.

9.2 Decay time acceptance

Studies of MC simulated events and real data show that the decay time distributions of the data do not follow a simple exponential function as could be expected due to the very good decay time resolution. For small decay times the distribution rises steeply from zero, before it saturates in an exponential decay function after about 1 ps depending on the decay channel. Another effect becomes visible when fitting the lifetime on MC simulated events using only candidates with large decay times. The systematically low measured lifetimes lead to an efficiency effect increasing with larger decay times. Both effects are described in the following section.

9.2.1 Efficiency for candidates with small decay times

If a B meson has a small decay time, its daughter particles are likely to have tracks with a very small distance to the PV. Depending on the final state and on the multiplicity of the particle types that belong to this final state, the selection requirements have to reduce the combinatorial background from particles that have been produced directly in the primary interaction. A most efficient criterion is a requirement on a large time of flight or a large flight distance in the lab system. Requirements like this directly impose an efficiency effect for candidates with a small decay time. They are used for selecting $B^0 \to D^-\pi^+$ candidates.

A J/ψ resonance in the final state reduces the possibility for background from particles that originate from the PV, as it is possible to use requirements on the J/ψ invariant mass. Though, selections of $B^0 \to J/\psi K^{*0}$ candidates often use cuts on the IP or IP significance. These have an indirect effect on the reconstruction efficiency, as for candidates with large decay times the probability for the daughter tracks to have a large IP (significance) increases (c.f. Fig. 9.2). These effects modify the visible distribution of the decay time and have to be described in the fitting pdfs to properly separate signal and background contributions.

In the $B^0 \to D^-\pi^+$ channel the effect is large and there is no suitable control channel to describe the exact shape of the decay time efficiency. Though, studies with MC simulated events lead to an analytic parametrisation that can be used in the fits to data[1](cf. Fig 9.3). The function

$$\epsilon(t) = \arctan(t\mathrm{e}^{(a_1 t + a_2)}) \tag{9.1}$$

[1]The parametrisation has been developed and its use for the description of the decay time efficiency was tested in the work of the master theses of R. Niet *"Studies on the Propertime Resolution in the Decay $B_s^0/\overline{B}_s^0 \to D_s^{\mp}K^{\pm}$ for the measurement of the CKM angle γ at LHCb"* – September 2012 – and F. Meier *"Untersuchung der Zerfallszeitbeschreibung im Zerfall $B^0 \to J/\psi K_S^0$ bei LHCb"* – September 2012.

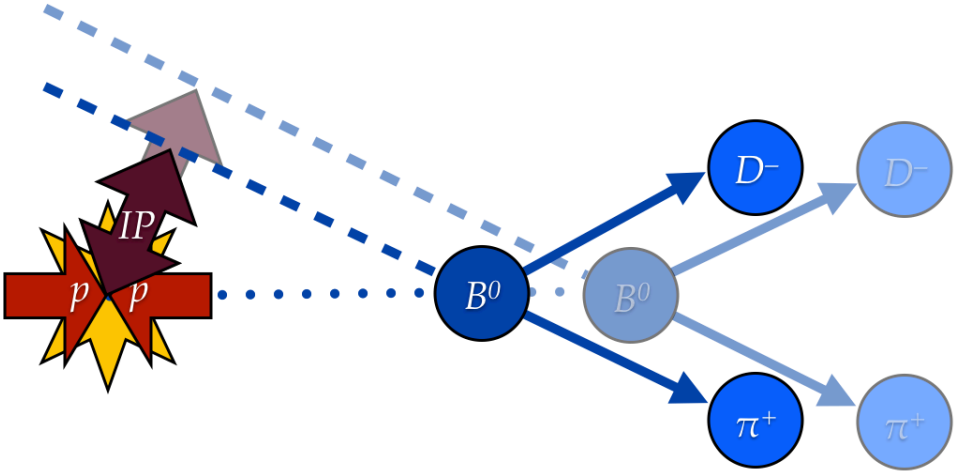


Fig. 9.2: Illustration of possible biases on the reconstructed decay time due to acceptance effects. Candidates with a higher decay time are selected more likely, if requirements on the IP or the IP significance are used.

with its two free parameters a_1 and a_2 can be multiplied with the exponential decay time pdf to describe the decay time distribution on data.

The decay time distribution of the $B^0 \to J/\psi\, K^{*0}$ channel is influenced by requirements imposed by the trigger algorithms. In addition to the so called lifetime unbiased triggers HLT1X and HLT2X there is an additional trigger HLT2L that is used for the selection of the data. If now S_{HLTxx} represents the set of events passing a certain trigger and N(S) is the number of events in that set the fraction

$$\epsilon(t) = \frac{N(S_{\mathrm{HLT1X}} \cap S_{\mathrm{HLT2X}} \cap S_{\mathrm{HLT2L}})(t)}{N(S_{\mathrm{HLT1X}} \cap S_{\mathrm{HLT2X}})(t)} \tag{9.2}$$

can be used to calculate a histogram of the decay time efficiency introduced by HLT2L in bins of the decay time of the candidates. A pdf based on this histogram is then used to describe the effect on data.

Fig. 9.3 shows the measured decay time acceptance on data and the corresponding models that are used in the fit to Δm_d and for tagging calibration. It is not possible to plot the data without a fit of the effective lifetime since the acceptance effect for high decay times (see following section) is neglected in the fit.

9.2.2 Efficiency for candidates with large decay times

From studies with MC simulated events it is known that the measured lifetime for b hadron decays with the LHCb detector is systematically biased to lower lifetimes.

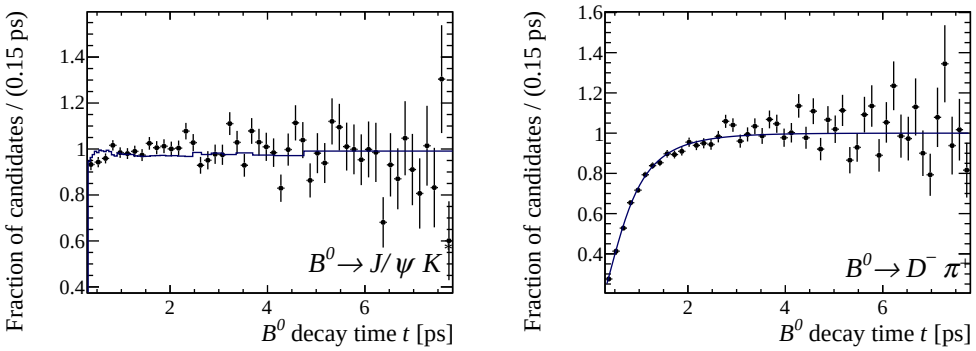


Fig. 9.3: Projection of the signal components (extracted using *sWeights* [69] from a fit to the invariant mass spectrum) of the decay time t weighted by $\exp(t/\tau)$, where τ is the fit result for the signal lifetime (black points). The decay time acceptance functions for both channels are visualised by the solid line. For these projections the nominal fit results from Tab. 11.1 and Fig. 11.5 are used.

This is due to a reconstruction inefficiency in the VELO[2]. The pattern recognition algorithms prefer tracks with a small IP. As tracks coming from b hadron decays with larger decay times have larger IPs (c.f. Fig. 9.2), these decays suffer from the reconstruction efficiency. Therefore, the decay time distribution that normally follows an exponential distribution gets a steeper slope. This is modelled by an additional linear component $1 - \beta t$, leading to a decay time distribution

$$\mathcal{P}(t) = \mathrm{e}^{-t/\tau}(1 - \beta t). \tag{9.3}$$

However, the shape of the decay time distribution can be described sufficiently accurate by a simple exponential. This is done in this thesis to the cost of the lifetime parameter τ, which then only describes an effective lifetime. The efficiency effect for large decay times is much higher in the $B^0 \to D^-\pi^+$ channel compared to decay channels including a J/ψ resonance.

[2]These results were presented in a talk on the measurement of $b \to J/\psi X$ lifetimes by Y. Amis during the LHCb Week in Davos (September 5th, 2012). There is no publication for reference up to now.

10 Tagging calibration

In preparation of several LHCb measurements that use flavour tagging, the optimised tagging calibration has to be verified in different decay channels. Additionally, potential systematic uncertainties due to the polarisation of the magnet field, production and reconstruction asymmetries, kinematic differences and several other reasons are estimated. The calibration measurements were performed before the actual measurement of Δm_d. The methods and results presented in this thesis reflect the calibration as described in [15, 16, 70]. The details of the selection and fitting strategy are slightly different to the measurement of Δm_d, as the analysis strategy has been refined for the measurement of Δm_d. The improvements will be used for future calibration measurements.

The tagging algorithms are optimised to maximise the effective tagging efficiency $\varepsilon_{\text{eff}} = \varepsilon D^2$. This value is proportional to the sensitivity of the sample with respect to a measurement of the mixing or *CP* amplitude as it is necessary for a measurement of time-dependent *CPV*. That means a sample with 100 $B^0 \to J/\psi\, K^0_{\text{S}}$ signal candidates with perfect tagging ($\varepsilon_{\text{eff}} = 1$) has the same sensitivity to $\sin 2\beta$ as a sample containing 10000 candidates with a dilution of $D = 0.2$ and a tagging efficiency of $\varepsilon = 0.25$.

10.1 Fitting strategy

The tagging calibration measurements use a fitting model similar to that of the measurement of Δm_d as it is described in Sec. 7. However, the value of Δm_d is fixed to $0.5\,\text{ps}^{-1}$ in all fits and in the PDF for the combinatorial background Eq. (7.8) the parameter ω is fixed to zero, allowing no asymmetry in the background. Additionally, there is negligible background in the $B^0 \to D^-\pi^+$ channel due to the different selection. Therefore, the fit model for $B^0 \to D^-\pi^+$ decays also neglects any background contribution, which is exact to a very good precision.

The fits are performed in several disjoint categories (cf. Fig 10.1) of the predicted wrong tag probability η. They run simultaneously in these categories and share all parameters but the component yields and the mean wrong tag probability of the categories. There are two different aims of the calibration measurements:

- Do a measurement on a subset of the data (e.g. only OS/SS taggers) and fit the mean wrong tag probability in categories of the neural net prediction η_c.

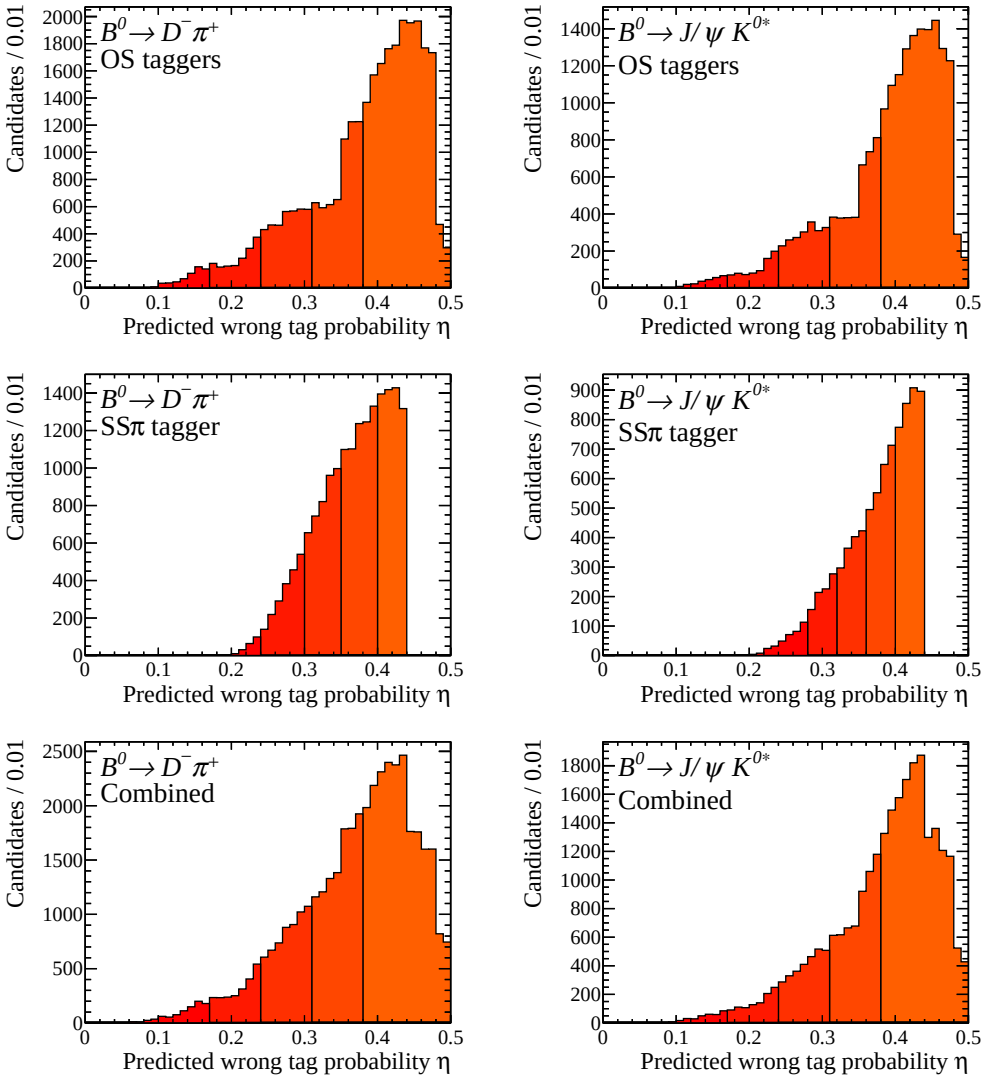


Fig. 10.1: Distribution of the predicted mistag probability η_c in (left) $B^0 \to D^-\pi^+$ and (right) $B^0 \to J/\psi K^{*0}$ events. The top row shows the combination of OS taggers, the middle row shows the SS pion tagger and the bottom row shows the full combination of all taggers. The different colours and vertical separations indicate the limits of the tagging categories. The events for these plots are selected in a ± 25 MeV mass range around the B^0 mass world average and with a requirement on the B^0 decay time $t > 1$ ps to suppress remaining background.

This can then be used to plot the neural net prediction against the fit result. In a perfect world the parameters of the calibration function

$$\omega(\eta_c) = p_0 + p_1(\eta_c - \langle\eta_c\rangle) + p_2(\eta_c - \langle\eta_c\rangle)^2 \tag{10.1}$$

would then show a slope $p_1 = 1$ and an offset $p_0 - \langle\eta_c\rangle = 0$. The parameter of the quadratic term is expected to be $p_2 = 0$. Any significant differences in one of the parameters would give a hint to problems in the calibration.

- Measure the average wrong tag probability (or a calibration plot as described above) in a subset of data binned in variables that are expected to be correlated to the tagging performance. The results can be used to estimate systematic uncertainties on measurements of *CPV*, especially in decay modes that show different distributions in these variables.

The mean values of η_c in each category are calculated using the *sweighted*[69] distributions of η_c for the signal. The *sWeights* are calculated from two component fits (signal and background) to the mass distributions of the data samples.

The following sections show studies performed on the decay modes $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi K^{*0}$. The benchmark value of ε_{eff} will be given for every tested combination of taggers. Due to different trigger strategies for both decay channels a different performance of the taggers is expected. $B^0 \to J/\psi K^{*0}$ candidates are triggered by muon triggers that give only small information about the quality of the other tracks. $B^0 \to D^-\pi^+$ candidates on the other hand have hadronic triggers that include the event topology and thus have a higher probability of additional good tracks that do not belong to the signal but can be used for tagging.

10.2 OST calibration

For the measurement of the OST calibration a tagging decision is formed combining the information of all active OS taggers like described in Sec. 4.4. The tagged B candidates are then divided into five bins of the predicted wrong tag probability η_c. The category with the worst tags (highest values of η_c) is referred to as category 1, where the best tags belong to category 5.

The dataset is fit simultaneously and a value for the mixing asymmetry amplitude corresponding to $1 - 2\omega$ is measured in each category. The measured value for ω is then compared to the mean value of the wrong tag probability predictions in the corresponding tagging category. The function (10.1) is then fitted to the resulting (η_c, ω) data points. The OS taggers have been calibrated on data using $B^+ \to J/\psi K^+$ events.

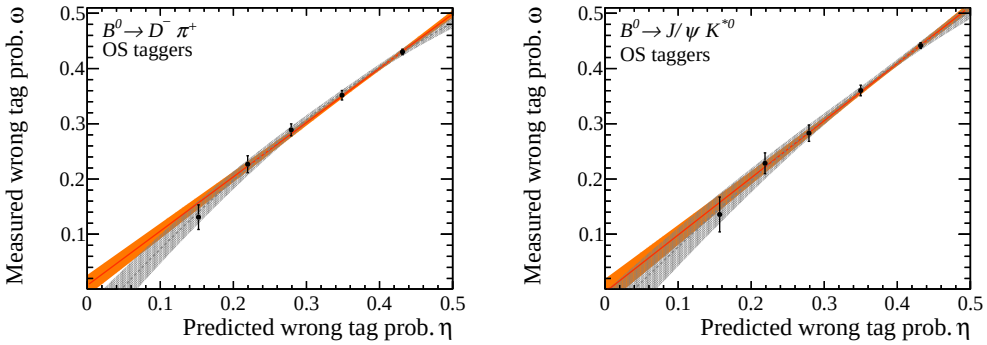


Fig. 10.2: Tagging calibration function for the combination of OS Taggers with (left) $B^0 \to D^-\pi^+$ and (right) $B^0 \to J/\psi K^{*0}$ events. The red curve and orange band show the linear fit with 1σ error band. The grey curve and error band correspond to the parabolic calibration function. The black points show the results from the fits to the average wrong tag probabilities in the five tagging categories. The numerical results of the calibration fits can be found in Tab. 10.3.

10.2.1 Results for $B^0 \to D^-\pi^+$ decays

The calibration for the OS taggers is measured to be excellent for $B^0 \to D^-\pi^+$ decays. The fits to the mixing asymmetry amplitudes are shown in Fig. 10.5, and the results for the tagging quality are presented in Tab. 10.1. $B^0 \to D^-\pi^+$ decays yield an effective tagging power $\varepsilon_{\text{eff}} = 2.98 \pm 0.16$ from OST only.

The calibration parameters (cf. Fig. 10.2) are compatible with the expected values, with $p_0 - \langle\eta_c\rangle = 0.00 \pm 0.004$, and $p_1 = 0.982 \pm 0.049$. The test with a parabolic parametrisation of the calibration function yields $p_2 = -0.737 \pm 0.640$, which is compatible with zero.

10.2.2 Results for $B^0 \to J/\psi K^{*0}$ decays

As $B^0 \to J/\psi K^{*0}$ decays are triggered by the powerful muon triggers, the situation is different to $B^0 \to D^-\pi^+$ decays triggered in the hadronic lines. Due to the different trigger configurations $B^0 \to J/\psi K^{*0}$ decays are nearly exclusively due to tracks that belong to the signal (e.g. one or two muons from the J/ψ), where many candidates from the hadronic lines are triggered by information from the opposite side b hadron. Therefore, $B^0 \to D^-\pi^+$ decays have a higher probability of a well reconstructed opposite side and the tagging power in $B^0 \to J/\psi K^{*0}$ decays is measured to be $\varepsilon_{\text{eff}} = 2.04 \pm 0.14$ and thus smaller than in the hadronic channel.

Still, the results from the calibration measurements show well calibrated OS taggers for $B^0 \to J/\psi K^{*0}$ decays (cf. Fig 10.6). The results from the calibration measurement

are $p_0 - \langle\eta_c\rangle = 0.009 \pm 0.004$, and $p_1 = 1.035 \pm 0.060$ and yield very good agreement with the expectations. As in the $B^0 \to D^-\pi^+$ channel, the parabolic terms from the reference fit are insignificant.

Tab. 10.1: Performance of the OST in $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi K^{*0}$ decays.

$B^0 \to D^-\pi^+$	ε (%)	ω	D	εD^2 (%)
$0.38 \leq \eta_c < 0.50$	23.478 ± 0.135	0.430 ± 0.005	0.140 ± 0.010	0.459 ± 0.066
$0.31 \leq \eta_c < 0.38$	7.691 ± 0.085	0.352 ± 0.009	0.297 ± 0.017	0.677 ± 0.078
$0.24 \leq \eta_c < 0.31$	4.618 ± 0.067	0.289 ± 0.011	0.422 ± 0.022	0.822 ± 0.084
$0.17 \leq \eta_c < 0.24$	2.010 ± 0.045	0.227 ± 0.015	0.546 ± 0.031	0.600 ± 0.068
$0.00 \leq \eta_c < 0.17$	0.770 ± 0.028	0.131 ± 0.022	0.738 ± 0.045	0.419 ± 0.051
Total	38.567 ± 0.180	0.376 ± 0.004	0.249 ± 0.008	2.977 ± 0.157

$B^0 \to J/\psi K^{*0}$	ε (%)	ω	D	εD^2 (%)
$0.38 \leq \eta_c < 0.50$	21.831 ± 0.137	0.442 ± 0.005	0.117 ± 0.010	0.298 ± 0.052
$0.31 \leq \eta_c < 0.38$	6.150 ± 0.079	0.360 ± 0.010	0.279 ± 0.019	0.480 ± 0.066
$0.24 \leq \eta_c < 0.31$	3.488 ± 0.061	0.283 ± 0.015	0.434 ± 0.030	0.657 ± 0.090
$0.17 \leq \eta_c < 0.24$	1.271 ± 0.037	0.229 ± 0.019	0.543 ± 0.038	0.375 ± 0.052
$0.00 \leq \eta_c < 0.17$	0.439 ± 0.022	0.136 ± 0.032	0.728 ± 0.063	0.233 ± 0.040
Total	33.179 ± 0.175	0.398 ± 0.004	0.205 ± 0.008	2.042 ± 0.140

10.3 SSπ calibration

In decays of B^0 mesons, the additional d quark leads to the production of an additional pion. If the pion is charged it can be used to form a same side tag. A priori it is not clear if the process is similar between $B^\pm$ and B^0 mesons. Thus, it is essential here to do the calibration in decays of neutral B mesons.

The calibration strategy is similar to that for the OS taggers, using 5 tagging categories. However, the category limits are adapted to reflect the wrong tag probability distribution of the SS pion tagger and to avoid empty categories. They are chosen to be different for both decay channels. The SSπ tagger is not calibrated on data. The decay channels $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi K^{*0}$ are used to verify the calibration of the tagger and to justify the calibration model.

10.3.1 Results for $B^0 \to D^-\pi^+$ decays

The measurement of the SS pion calibration using $B^0 \to D^-\pi^+$ decays shows a very high effective efficiency $\varepsilon_{\text{eff}} = 1.32 \pm 0.11\,\%$ (cf. Tab. 10.2). Though, the calibration function is evidently not correctly described by a linear model. The data points in

Tab. 10.2: Performance of the SS pion tagger in $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi K^{*0}$ decays.

$B^0 \to D^-\pi^+$	ε (%)	ω	D	εD^2 (%)
$0.40 \leq \eta_c < 0.50$	6.965 ± 0.114	0.450 ± 0.013	0.101 ± 0.026	0.071 ± 0.037
$0.35 \leq \eta_c < 0.40$	7.763 ± 0.120	0.400 ± 0.012	0.200 ± 0.024	0.312 ± 0.076
$0.30 \leq \eta_c < 0.35$	5.489 ± 0.102	0.378 ± 0.014	0.244 ± 0.028	0.326 ± 0.076
$0.24 \leq \eta_c < 0.30$	2.708 ± 0.073	0.302 ± 0.020	0.395 ± 0.040	0.423 ± 0.086
$0.00 \leq \eta_c < 0.24$	0.285 ± 0.024	0.203 ± 0.056	0.594 ± 0.112	0.101 ± 0.038
Total	23.211 ± 0.210	0.394 ± 0.007	0.213 ± 0.014	1.232 ± 0.148

$B^0 \to J/\psi K^{*0}$	ε (%)	ω	D	εD^2 (%)
$0.40 \leq \eta_c < 0.50$	6.043 ± 0.081	0.456 ± 0.010	0.087 ± 0.019	0.046 ± 0.020
$0.36 \leq \eta_c < 0.40$	4.346 ± 0.069	0.419 ± 0.011	0.163 ± 0.022	0.115 ± 0.031
$0.32 \leq \eta_c < 0.36$	2.667 ± 0.055	0.390 ± 0.014	0.220 ± 0.028	0.130 ± 0.033
$0.28 \leq \eta_c < 0.32$	1.576 ± 0.042	0.339 ± 0.017	0.322 ± 0.035	0.164 ± 0.035
$0.00 \leq \eta_c < 0.28$	0.688 ± 0.028	0.293 ± 0.027	0.414 ± 0.055	0.118 ± 0.031
Total	15.320 ± 0.130	0.412 ± 0.006	0.175 ± 0.012	0.573 ± 0.069

Fig. 10.3 show a drop in $\omega(\eta_c)$ for low values of η_c, where the linear model seems to hold for high η_c values. The fit with the parabolic function results in a significant contribution of the terms that are of second order in η_c. This observation is confirmed in the semileptonic decay channel $B^0 \to D^*\mu\nu_\mu$.[1]

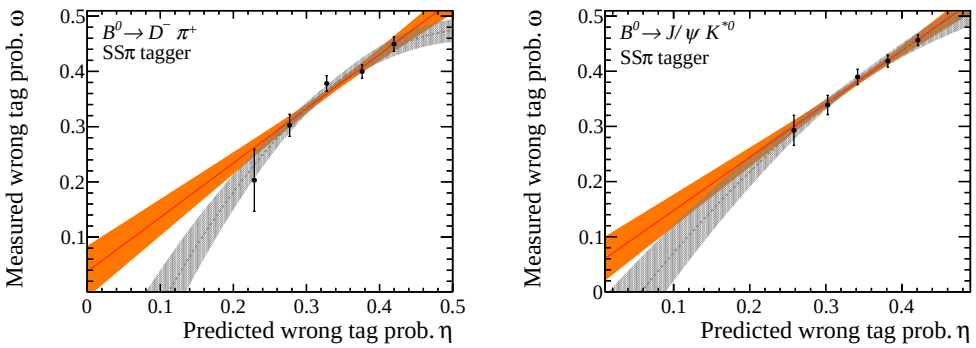


Fig. 10.3: Tagging calibration function for the combination the SSπ Tagger with (left) $B^0 \to D^-\pi^+$ and (right) $B^0 \to J/\psi K^{*0}$ events. The red curve and orange band show the linear fit with 1σ error band. The grey curve and error band correspond to the parabolic calibration function. The black points show the results from the fits to the average wrong tag probabilities in the five tagging categories. The numerical results of the calibration fits can be found in Tab. 10.3.

[1] This result was presented in a talk on the 63rd LHCb week (2012-03-29) by Stefania Vecchi

10.3.2 Results for $B^0 \to J/\psi K^{*0}$ decays

The effective efficiency of the SS pion tags in $B^0 \to J/\psi K^{*0}$ decays is measured to be $\varepsilon_{\text{eff}} = 0.627 \pm 0.074\,\%$ (cf. Tab. 10.2) and thus is significantly smaller than that in $B^0 \to D^-\pi^+$ decays. The calibration function however does not show the same features. Therefore, there is no evidence in these decays for contributions of a non-linear component to the calibration function. The linear calibration yields $p_0 - \langle\eta_c\rangle = 0.004 \pm 0.007$, and $p_1 = 1.236 \pm 0.151$ and is thus compatible with the expectations.

The statistical sensitivity of the high tagging categories is very low in $B^0 \to J/\psi K^{*0}$ decays. Thus, the significant observations of a non-linear calibration function in two other decay channels still justify the use of the parabolic model for the SS tagging calibration.

10.4 Combination of taggers

For a measurement of Δm_d the statistical sensitivity can be increased by combining the responses of OS and SS taggers. However, due to possible correlations between them, the calibration is not expected to be perfect a priori. The combined response can be tested in the same way as the single OS and SS taggers in the previous sections. The results in Fig. 10.4 show a small difference in the parameter p_0. As the parabolic model shows no conclusive advantage over the linear one, the latter is used for the physics measurement. Thus, as the decay channels to measure Δm_d have self-tagging final states, the calibration parameters can be varied in a fit to Δm_d to ensure the asymmetry amplitude due to the calibration of the wrong tag probability is described correctly. This is crosschecked by different tagging calibration settings.

10.5 Tagging related systematic uncertainties on measurements of *CPV*

Measurements of *CPV* require an excellent knowledge of the wrong tag probabilities in the corresponding decay channels, as the dilution $D = 1 - 2\omega$ is a direct factor to many *CP* violating amplitudes. The knowledge of D is limited by differences in the kinematic and other variables between the data used for calibration and that used for the physics measurements. In the following sections sources of systematic uncertainties to measurements of *CPV* are presented.

The systematic uncertainties are estimated by measurements of the effective tagging efficiency with the data binned in different event categories or kinematic variables.

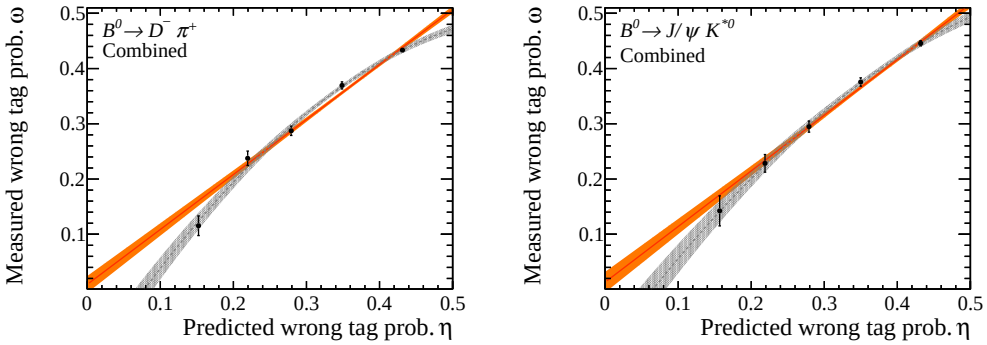


Fig. 10.4: Tagging calibration function for the combination of OS and SS Taggers with (left) $B^0 \to D^-\pi^+$ and (right) $B^0 \to J/\psi K^{*0}$ events. The red curve and orange band show the linear fit with 1σ error band. The grey curve and error band correspond to the parabolic calibration function. The black points show the results from the fits to the average wrong tag probabilities in the five tagging categories. The influence of the uncalibrated SSπ tagger is clearly visible in the calibration. Still, none of the models shows a significantly better fit performance than the other. The numerical results of the calibration fits can be found in Tab. 10.3.

The differences in the results of several decay channels are then used to calculate the uncertainties to be used in analyses that measure *CPV*.

Magnet polarity

The LHCb detector covers only an angle of about 300 mrad and differently charged particles are bent in opposite direction by the dipole magnet. This can lead to different reconstruction efficiencies for oppositely charged tracks and therefore influence the tagging efficiency. These differences can arise from asymmetries in the detector and corresponding systematic uncertainties can be estimated by analysing the tagging quality separately for different directions of the dipole magnet field. The data is divided in two parts, each one belonging to the runs with magnet-up or magnet-down fields. The difference in the effective tagging efficiencies is taken as the systematic uncertainty due to the dipole magnet polarisation.

Number of primary vertices

The LHCb detector is designed to handle events with one primary proton-proton collision and thus one PV. Due to recent developments in the LHC running scheme, the average number of PVs in an LHCb event is between two and three leading to the regular appearance of events with up to seven PVs. The number of primary

Tab. 10.3: Tagging calibration results for the OS combination, the SSπ tagger and the combination of all taggers in the $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi K^{*0}$ channel.

Channel	Function	$p_0 - \langle\eta_c\rangle$	p_1	p_2
$B^0 \to D^-\pi^+$ OST	linear	0.000 ± 0.004	0.982 ± 0.049	
	parabolic	0.004 ± 0.005	0.904 ± 0.083	-0.737 ± 0.640
$B^0 \to D^-\pi^+$ SSπ	linear	0.031 ± 0.007	0.983 ± 0.141	
	parabolic	0.038 ± 0.010	0.894 ± 0.166	-2.716 ± 2.666
$B^0 \to D^-\pi^+$ All	linear	0.007 ± 0.003	0.995 ± 0.041	
	parabolic	0.016 ± 0.004	0.859 ± 0.060	-1.588 ± 0.506
$B^0 \to J/\psi K^{*0}$ OST	linear	0.009 ± 0.004	1.035 ± 0.060	
	parabolic	0.011 ± 0.006	0.984 ± 0.109	-0.460 ± 0.822
$B^0 \to J/\psi K^{*0}$ SSπ	linear	0.002 ± 0.006	0.966 ± 0.125	
	parabolic	0.002 ± 0.008	0.915 ± 0.165	-1.226 ± 2.546
$B^0 \to J/\psi K^{*0}$ All	linear	0.016 ± 0.004	1.007 ± 0.049	
	parabolic	0.021 ± 0.005	0.897 ± 0.080	-1.183 ± 0.674

interactions in an event has a strong influence on the number of tracks und thus the number of tagging candidates in that event. Additionally, there is a non-zero probability of assigning tracks to the wrong PV and thus bias the tagging algorithms. For statistical reasons the dataset is divided into subsets with one, two, or three PVs. The measured differences give rise to a systematic uncertainty due to variations in the number of PVs.

Number of tracks

The number of tracks in an event influences the tagging algorithms in a similar way as the number of PVs. With only one PV there is no chance to assign a tagging candidate to an incorrect PV. Still, the systematic uncertainties due to the number of tracks and PVs are expected to be correlated to each other.

Initial/final state asymmetries

For measurements of *CPV* it is crucial to know, whether there are any differences in the tagging algorithms depending on the initial flavour of the B meson. The dataset can be divided depending on the tagging decision and on the final state (in the case of a self-tagging calibration channel). Differences in the tagging performances between these are referred to as systematic uncertainties due to the initial or final state asymmetries.

Tab. 10.4: Comparison of the tagging efficiency of the OS taggers under different running conditions in $B^0 \to D^-\pi^+$ decays.

	ε_{tag} [%]	ω	$(1-2\omega)$	$\varepsilon_{\text{tag}}(1-2\omega)^2$ [%]
Mag. down	38.910 ± 0.175	0.389 ± 0.004	0.223 ± 0.008	1.930 ± 0.140
Mag.up	38.210 ± 0.206	0.382 ± 0.005	0.235 ± 0.010	2.114 ± 0.174
K^+	38.655 ± 0.190	0.386 ± 0.004	0.229 ± 0.009	2.024 ± 0.156
K^-	38.580 ± 0.187	0.387 ± 0.004	0.227 ± 0.009	1.986 ± 0.153
One PV	39.993 ± 0.221	0.378 ± 0.005	0.243 ± 0.010	2.369 ± 0.196
Two PV	38.486 ± 0.213	0.388 ± 0.005	0.225 ± 0.010	1.940 ± 0.172
More PV	36.718 ± 0.270	0.396 ± 0.007	0.208 ± 0.013	1.583 ± 0.199
B^0	—	0.393 ± 0.004	0.214 ± 0.009	—
$\overline{B}^0$	—	0.379 ± 0.004	0.242 ± 0.009	—

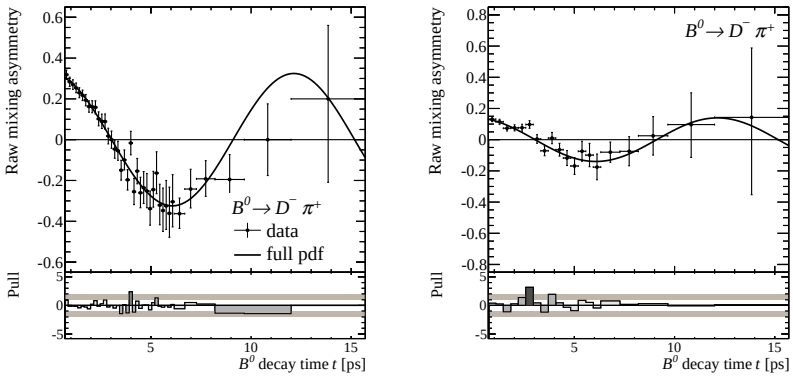
Tab. 10.5: Comparison of the tagging efficiency of the OS taggers under different running conditions in $B^0 \to J/\psi K^{*0}$ decays.

	ε_{tag} [%]	ω	$(1-2\omega)$	$\varepsilon_{\text{tag}}(1-2\omega)^2$ [%]
Mag. down	32.593 ± 0.222	0.391 ± 0.006	0.218 ± 0.012	1.554 ± 0.176
Mag. Up	32.680 ± 0.188	0.399 ± 0.005	0.202 ± 0.010	1.333 ± 0.138
K^+	32.231 ± 0.205	0.390 ± 0.006	0.219 ± 0.011	1.552 ± 0.161
K^-	33.045 ± 0.202	0.401 ± 0.006	0.199 ± 0.011	1.304 ± 0.146
One PV	32.581 ± 0.241	0.389 ± 0.007	0.222 ± 0.013	1.603 ± 0.192
Two PV	32.924 ± 0.234	0.399 ± 0.006	0.201 ± 0.013	1.336 ± 0.171
More PV	32.320 ± 0.279	0.399 ± 0.008	0.203 ± 0.016	1.327 ± 0.206
B^0	—	0.414 ± 0.006	0.171 ± 0.011	—
$\overline{B}^0$	—	0.376 ± 0.005	0.247 ± 0.011	—

10.6 Summary

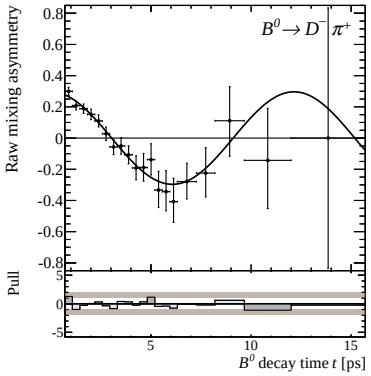
The calibration of the OS taggers is ready for physics measurements. A new parametrisation for the SS pion tagger calibration is found and validated in collaboration with the analysis of semileptonic decays $B^0 \to D^* \mu \nu_\mu$. The input of this analysis is thus used for an improved calibration of the SS pion tagger and the combination of OS and SS taggers. From the results presented in Tabs. 10.4 and 10.5 the main contribution of the flavour tagging to systematic uncertainties in time dependent *CPV* measurements comes from a wrong tag probability asymmetry depending on the tag of the initial flavour of the B meson. The difference of the measured wrong tag probability for candidates tagged as B^0 or $\overline{B}^0$ from the average of the sample is 0.007 ± 0.003 in $B^0 \to D^- \pi^+$ events and 0.019 ± 0.004 in $B^0 \to J/\psi K^{*0}$ decays. However, this uncertainty is diluted by the oscillations of the B^0 mesons. To get rid of this dilution and estimate a systematic uncertainty to measurements of *CPV*, the difference depending on the initial flavour can be measured using charged B mesons. The other tests show no significant deviations.

For the measurement of Δm_d in $B^0 \to D^- \pi^+$ and $B^0 \to J/\psi K^{*0}$ decays the taggers do not have to be perfectly calibrated. As the linear calibration function has proved to describe the combined wrong tag probability accurately, the parameters p_0 and p_1 can be varied in the fits to Δm_d.

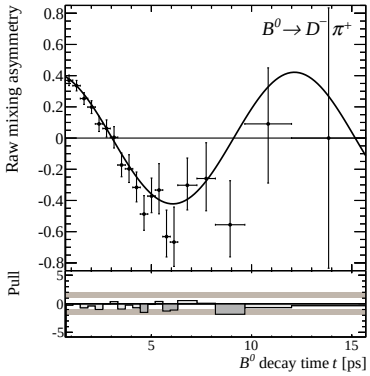


(a) Average of categories

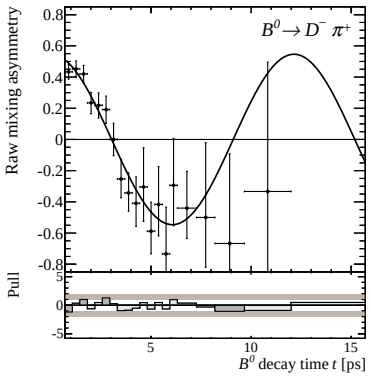
(b) Cat. 1: $0.48 \geq \omega > 0.40$



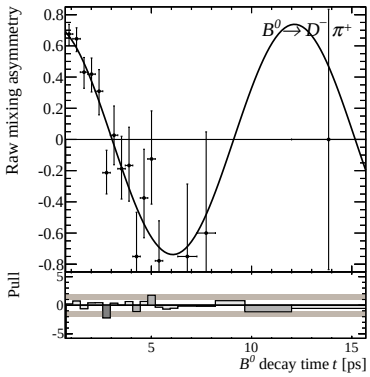
(c) Cat. 2: $0.40 \geq \omega > 0.32$



(d) Cat. 3: $0.32 \geq \omega > 0.24$

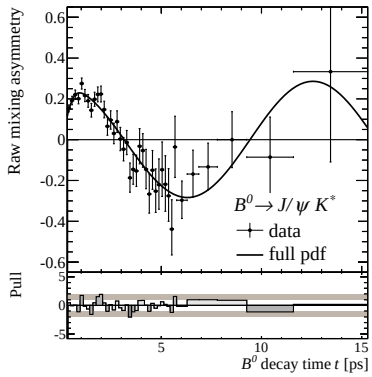


(e) Cat. 4: $0.24 \geq \omega > 0.17$

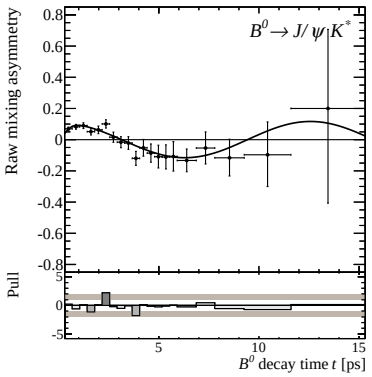


(f) Cat. 5: $0.17 \geq \omega > 0.00$

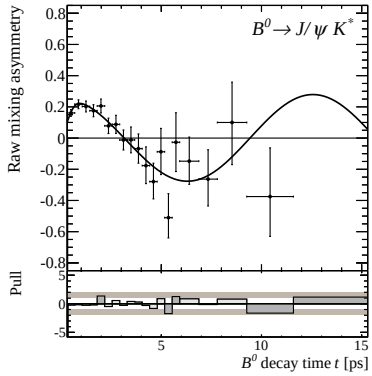
Fig. 10.5: Tagging performance in the mixing asymmetry (black curve) of $B^0 \to D^-\pi^+$ decays using the combined OS tagging decision on average for the full data sample and in five different categories of the predicted wrong tag probability.



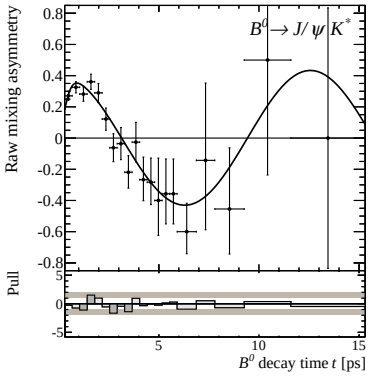
(a) Average of categories



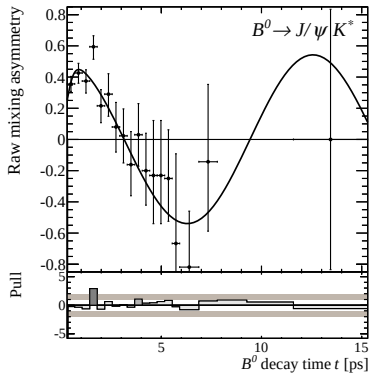
(b) Cat. 1: $0.48 \geq \omega > 0.40$



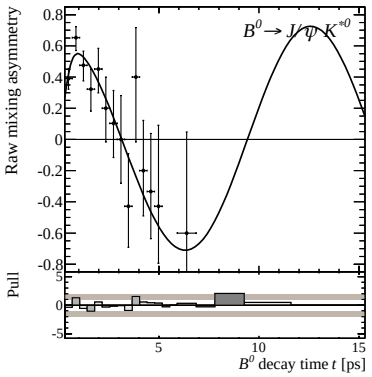
(c) Cat. 2: $0.40 \geq \omega > 0.32$



(d) Cat. 3: $0.32 \geq \omega > 0.24$



(e) Cat. 4: $0.24 \geq \omega > 0.17$



(f) Cat. 5: $0.17 \geq \omega > 0.00$

Fig. 10.6: Tagging performance in the mixing asymmetry (black curve) of $B^0 \to J/\psi\, K^{*0}$ decays using the combined OS tagging decision on average for the full data sample and in five different categories of the predicted wrong tag probability.

11 Measurement of Δm_d

The fit for Δm_d is handled slightly different for the two decay channels $B^0 \to D^- \pi^+$ and $B^0 \to J/\psi K^{*0}$ due to different background components and mass ranges. The fits yield two statistically independent results for Δm_d. Additionally, three different methods for the handling of the wrong tag probability are used to crosscheck the measurement. The use of per event wrong tag predictions η_c together with the tagging calibration function Eq. (10.1) is referred to as *nominal* method. Both results from the fits to each in the two decay channels are combined to form one result for the parameter Δm_d.

11.1 Fitting strategy

For $B^0 \to D^- \pi^+$ decays, a one dimensional fit to the mass distribution in the full mass range $5000 < m(B^0) < 5750$ MeV is performed, to ensure the background of the low mass region does not interfere with the signal (c.f. Fig. 8.2 and Sec. 8.1).

For the fit to Δm_d the mass range is then reduced to $5200 < m(B^0) < 5450$ MeV, where the results from the fit to the full mass range (cf. Fig. 8.2) are taken as starting values for the corresponding parameters of the multidimensional fit to Δm_d. In this mass range only the signal and one background component are relevant. The corresponding PDF for the extended maximum likelihood fit is described in Sec. 7.1.

In the $B^0 \to J/\psi K^{*0}$ channel, no hadronic background components near the signal region have to be taken into account. Though, there is an additional combinatorial background component from prompt J/ψ candidates. This component has a short effective lifetime compared to the longlived combinatorial background. Thus, the fit of Δm_d to $B^0 \to J/\psi K^{*0}$ candidates is a multidimensional three component extended maximum likelihood fit.

The background pdf that can describe a non-zero mixing asymmetry in the background decay time distribution is motivated by the analysis of the sideband data as it is shown in Fig. 11.1. All accessible mass sidebands show a positive mixing asymmetry that has to be accounted for in the measurement of Δm_d. The asymmetry is compatible with a constant asymmetry in the longlived combinatorial background component. It shows the typical suppression at low decay times in the $B^0 \to J/\psi K^{*0}$ channel, where the shortlived component of the background is dominating.

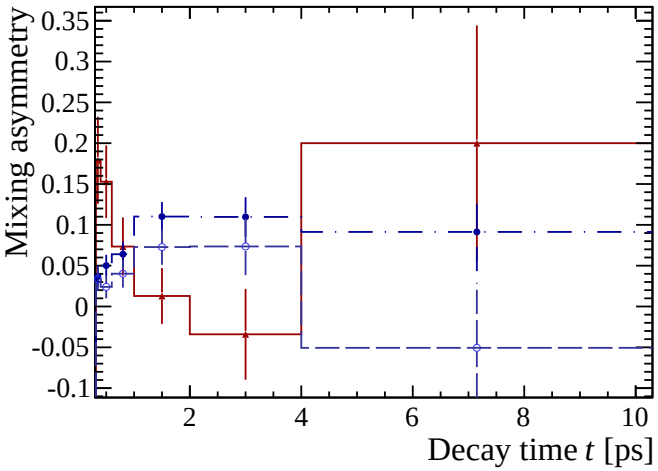


Fig. 11.1: Mixing asymmetry of the invariant mass sidebands. Red triangles and solid line for the high mass sideband of $B^0 \to D^-\pi^+$ candidates. Dark blue filled circles with dashed dotted line corresponds to the $B^0 \to J/\psi K^{*0}$ high mass sideband and the empty light blue circles with the dashed line to the $B^0 \to J/\psi K^{*0}$ low mass sideband.

In the fits to both decay channels, all parameters of the pdfs from Eqs. (7.15) and (7.20) are allowed to vary within ranges that are large against the expected uncertainty. The fit result of Δm_d has been blinded[1] during the development of the analysis.

In both fits, the pdfs $\mathcal{P}(m, t, \eta_c)$ are constructed neglecting that in fact they are conditional pdfs $\mathcal{P}(m, t|\eta_c)$. This is still correct, as the distributions of η_c are similar for all different components in the fit (cf. Fig. 11.2). Remaining differences are analysed in Sec. 11.3.2.

11.2 Fit results

The numerical results for the most relevant parameters are presented in Tab. 11.1 and a compilation of the results for all parameters can be found in Tabs. 11.2 and 11.3. The results of both decay channels are compatible with each other within one standard deviation. The tagging calibration parameters p_0 and p_1 are close to a perfect tagging calibration. Still, p_0 does not exactly equal $\langle\eta_c\rangle$ as the combination of the SSπ and the OS taggers is not calibrated for the data used in the analysis. The effective lifetimes are below the expectation for B^0 decays, as the decay time efficiency for large decay times is not described in the fit.

[1] The *RooUnblindUniform* algorithm is implemented in the RooFit package [55] and described in the online documentation.

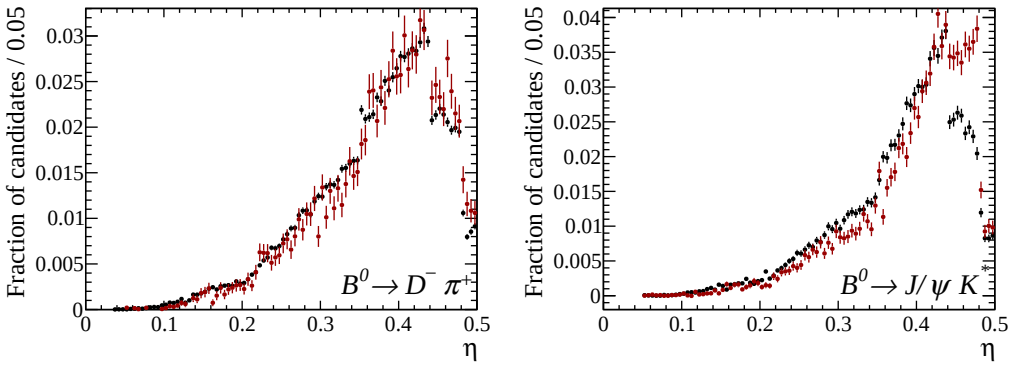


Fig. 11.2: *sWeighted* (black) signal and (red) background distributions of the per event estimations for the wrong tag probability for (left) $B^0 \to D^-\pi^+$ and (right) $B^0 \to J/\psi K^{*0}$ decays on data.

The deviations of the measured means of the B^0 mass distributions are due to the momentum scale calibration of the detector. The observed shifts are compatible with the results found for the systematic uncertainty on the reconstructed decay time in Sec. 11.3.4.

Tab. 11.1: Summary of fit results for the nominal fits. Find the full list and the results for the crosschecks in Tabs. 11.2 and 11.3.

Parameter & Unit		$B^0 \to D^-\pi^+$			$B^0 \to J/\psi K^{*0}$		
Δm_d	ps^{-1}	0.518	$\pm$	0.006	0.510	$\pm$	0.011
τ	ps	1.431	$\pm$	0.009	1.489	$\pm$	0.008
$m(B^0)$	MeV	5283.60	$\pm$	0.10	5281.20	$\pm$	0.05
p_0		0.3907	$\pm$	0.0025	0.3779	$\pm$	0.0039
p_1		1.022	$\pm$	0.0029	1.041	$\pm$	0.047
N_{sig}		88200	$\pm$	500	39150	$\pm$	$^{310}_{300}$
N_{sbg}					11800	$\pm$	$^{260}_{270}$
N_{lbg}		7170	$\pm$	$^{350}_{390}$	6290	$\pm$	300

Figs. 11.3 and 11.4 show the projections of the fit results to the mass and decay time distributions of the datasets. The pull distributions below the plots show the differences between data and model normalised to the uncertainties on the data points. From the pull distributions it becomes visible, that the data is well described by the fit model. The small differences for low lifetimes in the $B^0 \to D^-\pi^+$ channel are due to a statistical fluctuation in the data which only occurs for data taken with the LHCb magnet in *up* polarisation. The small fluctuations in the $B^0 \to J/\psi K^{*0}$ channel are due to decay time efficiency effects for the shortlived combinatorial background. Both effects are taken care of in the studies of systematic uncertainties due to the fit model in Sec. 11.3.1.

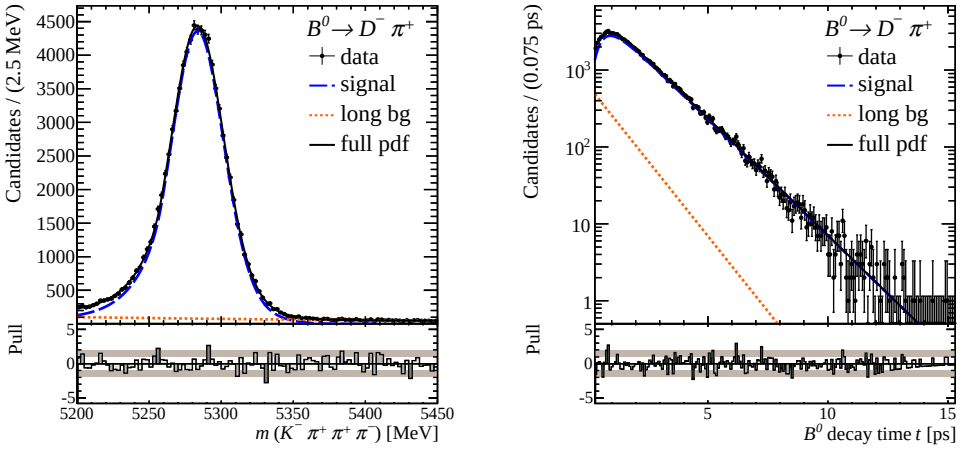


Fig. 11.3: Projections of the mass (left) and decay time (right) distributions from the nominal fit in $B^0 \to D^-\pi^+$. The blue long-dashed line shows the signal pdf projection, where the orange short-dashed line corresponds to combinatorial background. The solid black line is the full pdf projection.

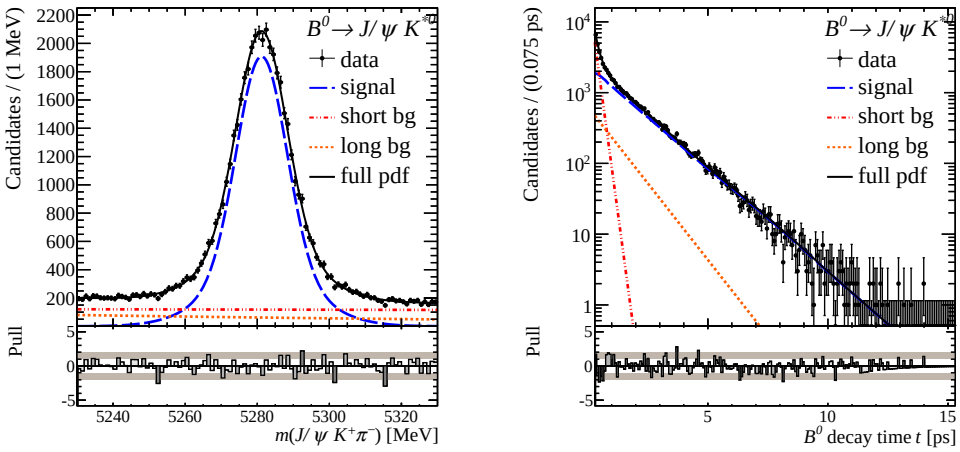


Fig. 11.4: Projections of the mass (left) and decay time (right) distributions from the nominal fit in $B^0 \to J/\psi\, K^{*0}$. The signal correspond to the blue long-dashed line. The two components of the longlived background are shown by the red dash-dotted and orange short-dashed lines. The solid black line is the full pdf projection.

Fig. 11.5 shows the projection of the asymmetry function from Eq. (10.1) to the data. Note, that the background also contributes to the asymmetry, resulting in a time dependent dilution of the asymmetry amplitude which is larger for small decay times.

The results of the log-likelihood scans for the parameter Δm_d are presented in Fig. 11.6. The log-likelihood functions are parabolic in both decay channels. Finally, Figs. 11.7 and 11.8 show the correlation matrices for the fits. The main correlations of Δm_d are with the tagging parameters p_0 and p_1 as well as the asymmetry ω_{lbg} of the background. No significant correlations with other parameters are found.

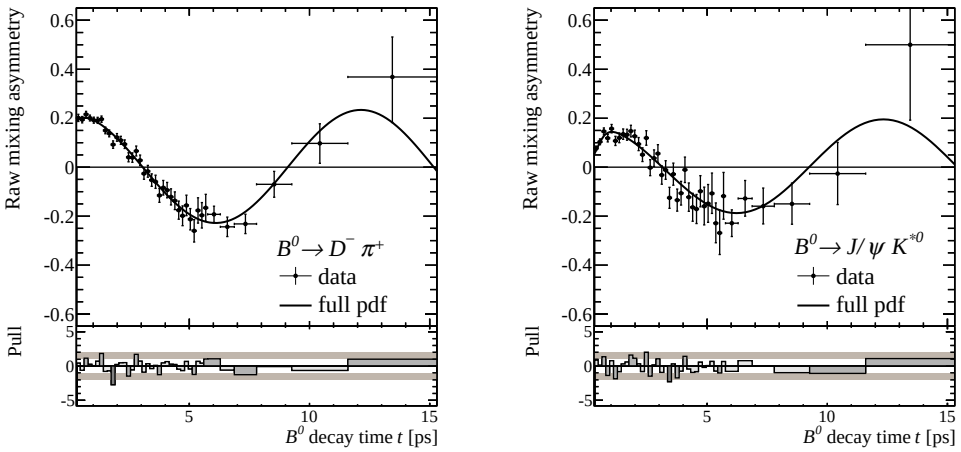


Fig. 11.5: Projections of the asymmetry functions from the nominal fits using per event mistags and floating calibration parameters.

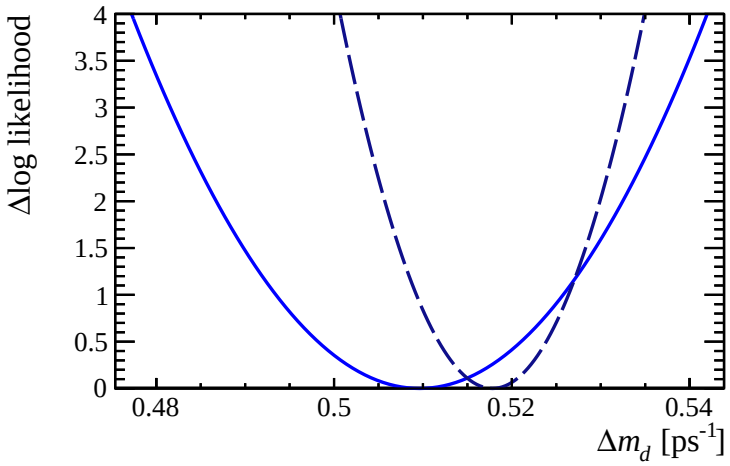


Fig. 11.6: Minima of the log-likelihood functions for the nominal fits in dashed dark blue for $B^0 \to D^- \pi^+$ and solid light blue for $B^0 \to J/\psi K^{*0}$. A Δlog-likelihood of 0.5 (2.0) corresponds to a 1 (2) σ deviation from the best fit value.

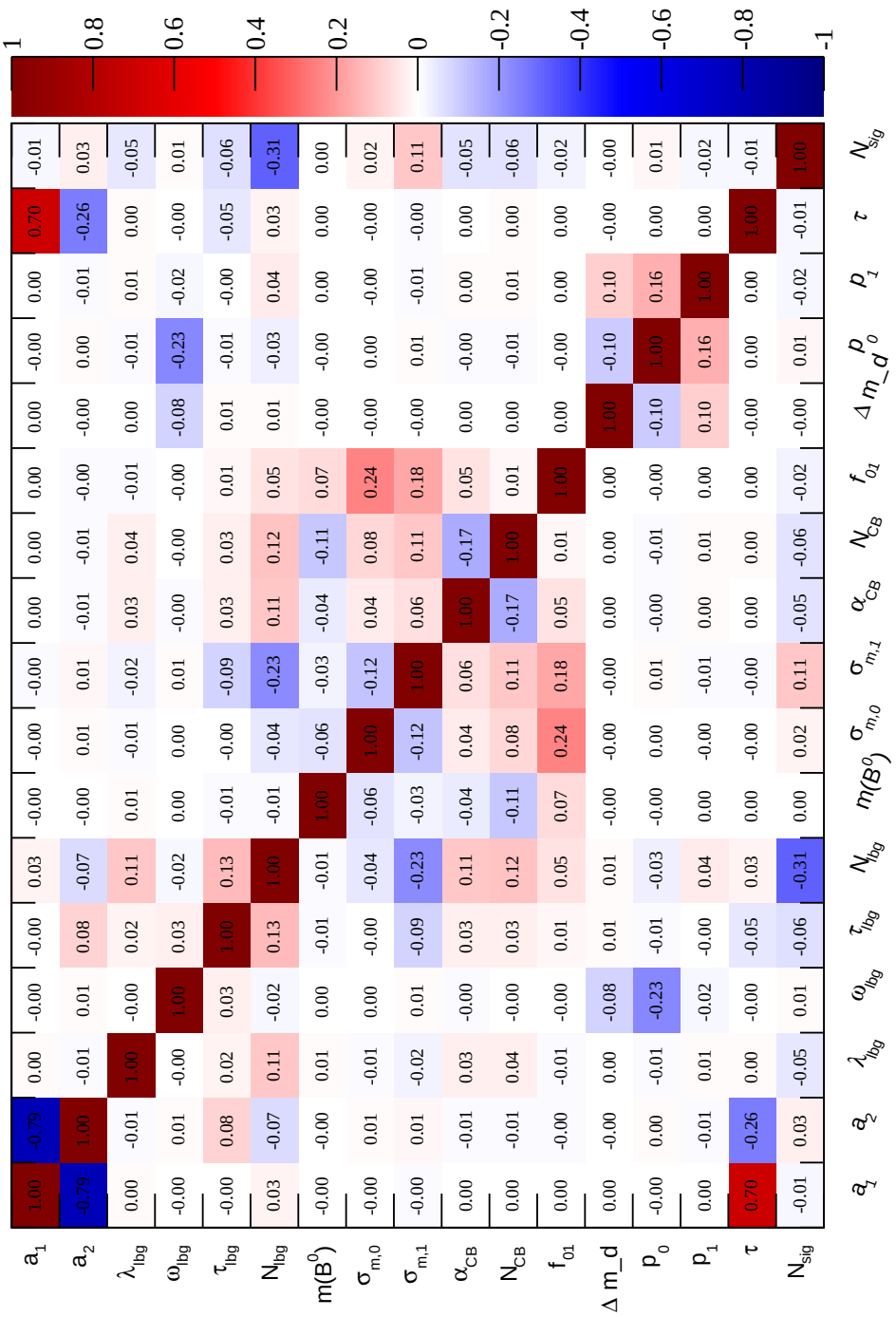


Fig. 11.7: Correlation coefficients of all fit parameters for $B^0 \to D^-\pi^+$ from the nominal fit result.

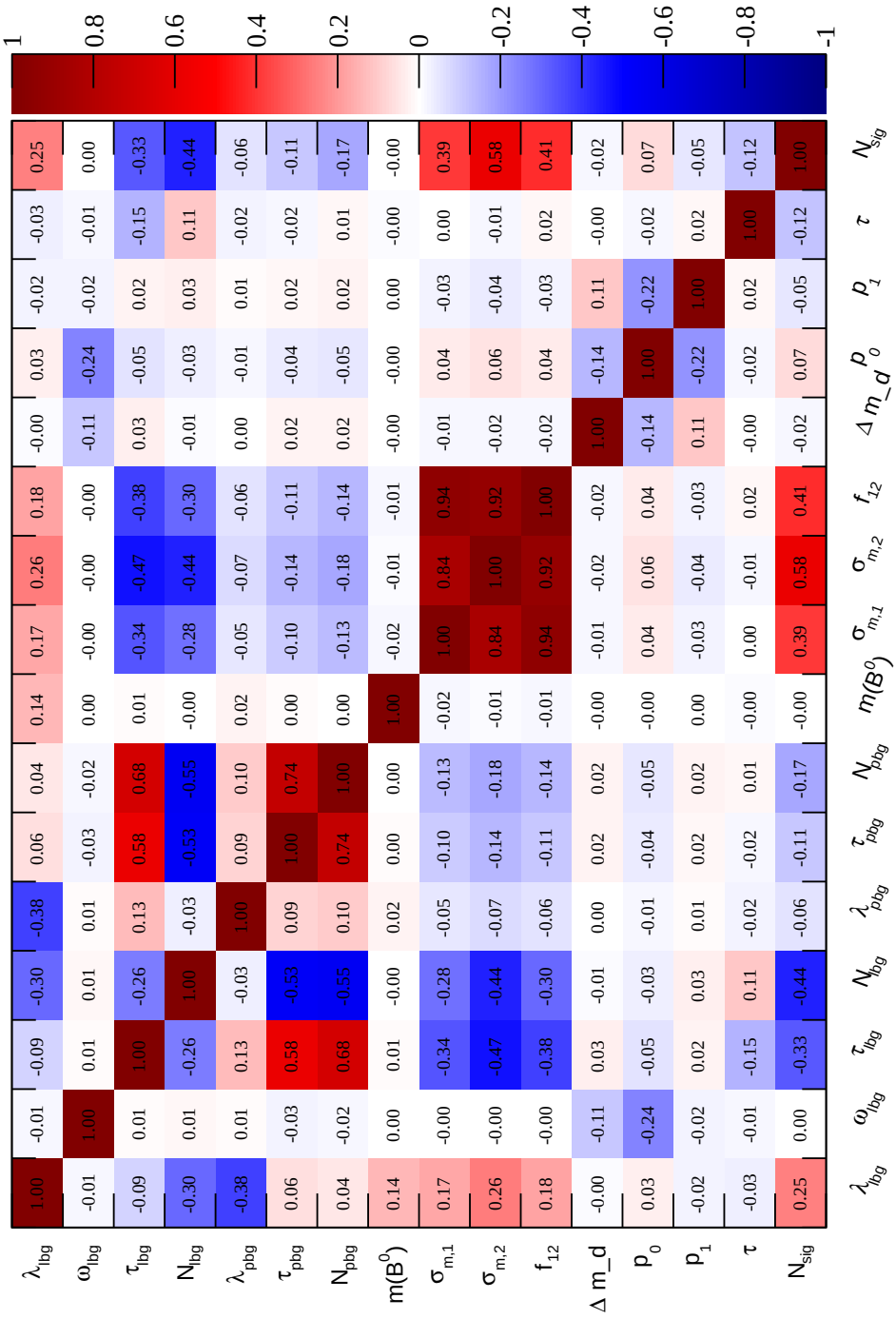


Fig. 11.8: Correlation coefficients of all fit parameters for $B^0 \to J/\psi K^{*0}$ from the nominal fit result.

Tab. 11.2: Summary of fit results and crosschecks for all fit parameters in $B^0 \to D^- \pi^+$ from the nominal fit result.

Parameter & Unit		nominal	mean η	5 cats
Δm_d	ps^{-1}	0.518 ± 0.006	0.517 ± 0.008	0.518 ± 0.006
τ	ps	1.431 ± 0.009	1.432 ± 0.009	1.433 ± 0.009
$m(B^0)$	MeV	5283.6 ± 0.1	5283.6 ± 0.1	5283.7(const)
p_0		0.391 ± 0.0025		
p_1		1.02 ± 0.029		
N_{sig}		88200 ± 500	88080^{+450}_{-440}	
$N_{\mathrm{sig},1}$				46750 ± 230
$N_{\mathrm{sig},2}$				22040 ± 160
$N_{\mathrm{sig},3}$				12250 ± 120
$N_{\mathrm{sig},4}$				4680 ± 70
$N_{\mathrm{sig},5}$				1890 ± 50
N_{lbg}		7170^{+350}_{-390}	7340^{+340}_{-360}	
$N_{\mathrm{lbg},1}$				4310 ± 100
$N_{\mathrm{lbg},2}$				1990 ± 70
$N_{\mathrm{lbg},3}$				1020 ± 50
$N_{\mathrm{lbg},4}$				374^{+32}_{-30}
$N_{\mathrm{lbg},5}$				107^{+17}_{-15}
a_1		0.93 ± 0.08	0.93 ± 0.08	0.94 ± 0.08
a_2		$-0.100^{+0.032}_{-0.033}$	$-0.103^{+0.032}_{-0.033}$	$-0.109^{+0.017}_{-0.032}$
ω			0.3819 ± 0.0025	
ω_1				0.443 ± 0.004
ω_2				0.360 ± 0.005
ω_3				0.285 ± 0.006
ω_4				0.231 ± 0.010
ω_5				$0.117^{+0.006}_{-0.013}$
ι		$1.8^{+0.7}_{-0.5}$	$1.7^{+0.6}_{-0.4}$	1.77(const)
α		1.44 ± 0.06	1.45 ± 0.06	1.421(const)
f_{12}		0.75 ± 0.09	$0.72^{+0.09}_{-0.08}$	0.69(const)
σ_0	MeV	17.3 ± 0.6	$17.1^{+0.6}_{-0.5}$	16.76(const)
σ_1	MeV	$27.8^{+3.7}_{-2.0}$	$26.9^{+2.7}_{-1.5}$	25.66(const)
τ_{lbg}	ps	$1.107^{+0.027}_{-0.029}$	$1.112^{+0.025}_{-0.026}$	$1.134^{+0.022}_{-0.021}$
λ_{lbg}	MeV^{-1}	$0.0030^{+0.0004}_{-0.0005}$	$0.0031^{+0.0004}_{-0.0005}$	

Tab. 11.3: Summary of fit results and crosschecks for all fit parameters in $B^0 \to J/\psi K^{*0}$ from the nominal fit result.

Parameter & Unit		nominal	mean η	5 cats
Δm_d	ps^{-1}	0.510 ± 0.011	$0.512^{+0.014}_{-0.015}$	0.507 ± 0.012
τ	ps	1.489 ± 0.008	1.489 ± 0.008	1.489 ± 0.008
$m(B^0)$	MeV	5281.2 ± 0.05	5281.2 ± 0.05	5281.2 ± 0.05
p_0		0.378 ± 0.004		
p_1		1.04 ± 0.05		
N_{sig}		39150^{+310}_{-300}	39090^{+310}_{-300}	
$N_{\mathrm{sig},1}$				23840^{+220}_{-210}
$N_{\mathrm{sig},2}$				8740 ± 110
$N_{\mathrm{sig},3}$				4430 ± 80
$N_{\mathrm{sig},4}$				1580 ± 40
$N_{\mathrm{sig},5}$				550^{+27}_{-26}
N_{sbg}		11800^{+260}_{-270}	11830 ± 260	
$N_{\mathrm{pbg},1}$				8320 ± 190
$N_{\mathrm{pbg},2}$				2110 ± 80
$N_{\mathrm{pbg},3}$				970 ± 50
$N_{\mathrm{pbg},4}$				353^{+27}_{-26}
$N_{\mathrm{pbg},5}$				94^{+15}_{-14}
N_{lbg}		6290 ± 300	6350 ± 300	
$N_{\mathrm{lbg},1}$				4190 ± 210
$N_{\mathrm{lbg},2}$				1210 ± 90
$N_{\mathrm{lbg},3}$				680 ± 60
$N_{\mathrm{lbg},4}$				150^{+28}_{-26}
$N_{\mathrm{lbg},5}$				62^{+18}_{-16}
ω			0.399 ± 0.004	
ω_1				0.4463 ± 0.005
ω_2				0.3748 ± 0.008
ω_3				$0.289^{+0.014}_{-0.010}$
ω_4				0.222 ± 0.016
ω_5				$0.136^{+0.027}_{-0.026}$
f_{12}		$0.680^{+0.034}_{-0.037}$	0.670 ± 0.04	0.670 ± 0.04
σ_1	MeV	$6.88^{+0.15}_{-0.16}$	$6.85^{+0.15}_{-0.16}$	$6.85^{+0.15}_{-0.16}$
σ_2	MeV	$13.8^{+0.8}_{-0.7}$	$13.6^{+0.8}_{-0.7}$	$13.7\pm^{+0.8}_{-0.7}$
τ_{lbg}	ps	1.02 ± 0.04	1.01 ± 0.04	1.01 ± 0.04
λ_{lbg}	MeV^{-1}	0.0050 ± 0.0007	0.0049 ± 0.0007	0.0049 ± 0.0007
τ_{sbg}	ps	0.173 ± 0.004	0.173 ± 0.004	0.173 ± 0.004
λ_{sbg}	MeV^{-1}	0.0004 ± 0.0004	0.0004 ± 0.0004	0.0005 ± 0.0004

11.3 Determination of systematic uncertainties

The measurement of Δm_d depends largely on a correct description of the asymmetry pdf. For the estimation of systematic uncertainties to the measurement, effects that influence the amplitude or the zero-crossing point of the asymmetry are considered. These are the fit model and the background description, that dilute the asymmetry amplitude for low decay times, as well as the decay time acceptance and the decay time resolution that have a potential to bias the amplitude and the zero-crossing point. Additionally, the influence of the tagging and other detector effects is studied.

11.3.1 Fit model

The main systematic uncertainty to the measurement of Δm_d is the one due to the fit model (e.g. the parametrisation of the signal and background decay time pdfs). The parametrisation is non-trivial, as for the signal component, the effects due to the decay time efficiencies are different between data and MC simulated events. For the background there is an insufficient number of combinatorial background candidates in the simulated events and thus parametrisations might depend on statistical fluctuations. The effects of the decay time efficiency are analysed exclusively in Sec. 11.3.5. The main uncertainty is assigned to the fit model, as any incorrect description of the time dependent signal $S(t)$ and background $B(t)$ yields dilutes the asymmetry amplitude, as

$$\mathcal{A}_{\text{mixing}}(t) \propto \frac{S(t)}{S(t)+B(t)} \cos \Delta m_d t. \tag{11.1}$$

From the correlation matrices in Figs. 11.7 and 11.8 it can be seen, that the amplitudes of the mixing asymmetry (represented by the tagging calibration parameters p_0 and p_1) are correlated with Δm_d and thus any inaccuracies in the model can bias the measurement.

In this analysis a novel approach is used to get a realistic estimate of the uncertainty due to the fit model. From a fit to the mass distribution including the signal peak and an exponential combinatorial background *sWeighted* [69] datasets are generated. Using these, Gaussian kernel pdfs are fit to the *sWeighted* distributions of the decay time distributions for signal and background. These then replace the exponential $\mathrm{e}^{-t/\tau}$ terms of the decay time pdfs (resolution effects are neglected). Thus, a model independent description of the decay time distributions including acceptance effects is found. The fits to Δm_d are then redone on the same datasets as the nominal fits. The differences to the nominal result of both decay channels are taken as a systematic uncertainty. Fig. 11.9 shows the corresponding projections of the mass pdfs and the Gaussian kernel pdfs. The systematic uncertainties on Δm_d due to the fit model inaccuracies are found to be $0.0037\,\mathrm{ps}^{-1}$ in the $B^0 \to D^-\pi^+$ channel and $0.0022\,\mathrm{ps}^{-1}$ in the $B^0 \to J/\psi K^{*0}$ channel.

11.3.2 Tagging

In the development of the analysis different strategies concerning the handling of the wrong tag probabilities have been used. In addition to the nominal method that uses the neural net wrong tag predictions of every single event, the events can also be divided into subsamples with similar values of the wrong tag probability. In every subsample the mean wrong tag probability ω_i is then a parameter of the fit (instead of the calibration parameters p_0 and p_1). The per event information for η_c is not used in this case.

These crosschecks are performed using only one parameter for the mean value of the wrong tag probability (1 Category) and also with the data divided into five categories of η_c. The results of the fits compared to the nominal results in both decay channels are summarised in Tab. 11.4. It is clearly visible, that the uncertainties for the fits with only one category are slightly higher than when using five categories or the per event information. That is due to the dependence of the sensitivity of Δm_d to the tagging dilution $D = 1 - 2\omega$. The more accurate the tagging is, the better the amplitude and the frequency of the oscillation can be described.

Tab. 11.4: Results for Δm_d in ps^{-1} for both channels in the different fitting settings and weighted averages.

	$B^0 \to J/\psi\, K^{*0}$	$B^0 \to D^-\pi^+$	average
per event	0.5096 ± 0.0114	0.5178 ± 0.0061	0.5156 ± 0.0051
1 Category	0.5121 ± 0.0145	0.5169 ± 0.0075	0.5154 ± 0.0065
5 Categories	0.5074 ± 0.0117	0.5181 ± 0.0061	0.5157 ± 0.0052

Additionally, the results from the nominal fit and the fit using five tagging categories are almost identical. That leads to the conclusion, that no significant bias to the analysis is expected due to the handling of the wrong tag probabilities. Thus, no systematic uncertainty is assigned to to the handling of the tagging information.

11.3.3 Length scale of the VELO

The measurement of Δm_d crucially depends on the length scale of the detector, because any inaccuracy directly translates to a systematic deviation of the measured value of the oscillation frequency. The z-axis of the detector is calibrated to a precision of 0.1 % (with reference to the measurement of ϕ_s in $B_s^0 \to J/\psi\,\phi$ [71]). As $t \approx {}^{m}/_{p}\Delta z$, the relative systematic uncertainty on Δm_d is of the same size. Thus, an absolute systematic uncertainty of $0.0005\,\mathrm{ps}^{-1}$ is assigned to the value of Δm_d due to the length scale calibration of the detector. This uncertainty is correlated across both decay channels.

11.3.4 Momentum calibration

As the decay time of each candidate can be calculated from $t = {}^{m}/_{p} \cdot L$, where L is the flight distance of the candidate, the momentum calibration has an influence on Δm_d similar to the influence of the length scale calibration. The uncertainty on the reconstructed momenta is measured to be smaller than $r_p = 0.15\,\%$ in decays of charmonia. Both the candidate mass m and the absolute value p of its three momentum depend on the measured momenta of the daughter particles. Therefore, parts of the uncertainty on the momentum calibration cancel in the fraction ${}^{m}/_{p}$.

To calculate the uncertainty on Δm_d taking into account the correlations in ${}^{m}/_{p}$ a study on MC simulated events is performed. The study uses only signal candidates that are considered to originate from true simulated signal events. The invariant mass and the momentum of the B^0 candidates are calculated from the four momenta of the daughter particles using the mass hypotheses for the corresponding particles. This is done twice using the true momenta in one case and scaling all momenta with $1 + r_p$ in the other case. Thus, two different decay times t_{true} (correct momenta) and t_{test} (scaled momenta) can be evaluated. The difference $t_{\text{true}} - t_{\text{test}}$ is then divided by t_{true} for every candidate. The mean of the resulting distribution (c.f. Fig. 11.10) is taken as the systematic uncertainty due to the momentum calibration.

This results in a relative uncertainty to Δm_d of $0.009\,\%$ in the $B^0 \to D^- \pi^+$ channel and $0.011\,\%$ in the $B^0 \to J/\psi\, K^{*0}$ channel. These uncertainties are one order of magnitude smaller than those on the length scale of the detector and thus negligible.

11.3.5 Decay time acceptance

An inaccurate description of the decay time acceptance can lead to a bias on Δm_d due to the same reason as for a fit model that does not describe the data well. If the time dependent numbers of signal $S(t)$ and background $B(t)$ candidates is not known exactly, the amplitudes can be biased and this affects the measurement of the oscillation frequency (c.f. Eq. (11.1)). Due to the different shapes of the decay time acceptance in both decay channels (c.f. Fig 9.3) a different influence on Δm_d is expected.

In order to estimate the systematic uncertainties coming from the description of the decay time acceptance, studies on candidates simulated from the fitting pdfs are performed. In these so called toy studies 10000 toy datasets are generated. Each dataset has a number of signal and background candidates similar to the corresponding real dataset. One fit is performed to every dataset and the systematic uncertainty on a certain fit parameter is evaluated by analysing the distributions of the residuals (difference between fit value and nominal value) and pulls (fraction of residual and parameter error) of that parameter. These distributions are fit by Gaussian functions to evaluate potential biases. In the case of significant shifts from zero, the mean of

the residual distribution is taken as the systematic uncertainty. In all toy studies no per event wrong tag information is used. The toy datasets reproduce the average wrong tag probability of the real data. In both decay channels the toy datasets are generated from the nominal pdf that is used for the measurement of Δm_d with the parameter values from the fit to data, including the measured acceptance. The value of $\Delta m_d = 0.5\,\mathrm{ps}^{-1}$ is used for toy generation.

In the fit to the toy datasets the decay time acceptance function is constant one and thus neglecting any acceptance effect in the data. All parameters with significant correlations to Δm_d are free in the fits to the toy datasets. These include the signal and background yields and their effective lifetimes, the average wrong tag probability, and Δm_d. The remaining parameters are fixed due to CPU time limitations.

In the $B^0 \to D^-\pi^+$ channel, the background contributions are additionally neglected, because the background is very low and flat under the signal peak. The fits to the toy datasets are performed in a reduced mass range of $5250 \leq m < 5310$ MeV. The fit results are presented in Fig. 11.11. A shift on Δm_d of about $0.0005\,\mathrm{ps}^{-1}$ is observed and assigned as the systematic uncertainty on Δm_d due to the decay time acceptance in this channel.

For $B^0 \to J/\psi\, K^{*0}$ decays, the shifts on Δm_d is very small compared to the statistical uncertainties. The results of this study are presented in Fig. 11.12. The uncertainty on Δm_d due to the decay time acceptance on this channel is $0.00001\,\mathrm{ps}^{-1}$.

11.3.6 Decay time resolution

For the evaluation of possible biases due to the decay time resolution a toy study similar to that in Sec. 11.3.5 is performed. The only difference is that instead of neglecting acceptance effects in the fits, the resolution is not included in the fit model. The results from 10000 toy experiments yield a marginal bias to Δm_d of $0.0002\,\mathrm{ps}^{-1}$ in both decay channels. The results of the study are presented in Figs. 11.13 and 11.14.

Additionally, the influence of not taking into account the per event information on the decay time resolution is tested. Therefore, toy data samples similar to the previous ones but with different decay time resolution for every single candidate are used. The decay time resolution values are generated from *sWeighted* distributions (c.f. Fig. 11.15) of the per event prediction on the decay time resolution. These are acquired from fits to the invariant mass distributions of the corresponding samples. No significant biases are found in this study.

11.3.7 Combination of systematic uncertainties

In the following, the systematic uncertainties are summarised in Table 11.5. The uncertainties on Δm_d are uncorrelated across the different channels and will be therefore added in quadrature in the final result. Only the uncertainty coming from the z-scale calibration has the same source in both cases. It will contribute to the final systematic uncertainty of the combined result as a relative error of one permille.

Tab. 11.5: Summary of systematic uncertainties on Δm_d refering to the nominal fit.

	$B^0 \to J/\psi\, K^{*0}$	$B^0 \to D^-\pi^+$
acceptance	0.0001	0.0004
decay time resolution	0.0002	0.0002
background model	0.0022	0.0037
combined	0.0022	0.0037
momentum scale	0.0000	0.0001
z-scale	0.0005	0.0005

11.4 Summary

The B^0 $\overline{B}^0$ oscillation frequency Δm_d has been measured in two independent datasets with $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi\, K^{*0}$ decays and is found to be

$$\Delta m_d(B^0 \to D^-\pi^+) = 0.5178 \pm 0.0061\,(\text{stat.}) \pm 0.0037\,(\text{syst.})\,\text{ps}^{-1} \text{ and}$$
$$\Delta m_d(B^0 \to J/\psi\, K^{*0}) = 0.5096 \pm 0.0114\,(\text{stat.}) \pm 0.0022\,(\text{syst.})\,\text{ps}^{-1}.$$

The combined value for Δm_d is calculated as the weighted average of the individual results taking correlated systematic uncertainties (cf. Tab. 11.5) into account. This leads to

$$\Delta m_d = 0.5156 \pm 0.0051\,(\text{stat.}) \pm 0.0033\,(\text{syst.})\,\text{ps}^{-1}.$$

The measurement is in good agreement with the PDG average of $\Delta m_d = 0.507 \pm 0.004\,\text{ps}^{-1}$ [7]. It is the current most precise measurement of this parameter.

With data taken in 2012 at the LHCb detector, the statistical sensitivity can be improved to below $0.004\,\text{ps}^{-1}$. Depending on the tagging performance a combined measurement with a statistical sensitivity of about $0.003\,\text{ps}^{-1}$ is possible.

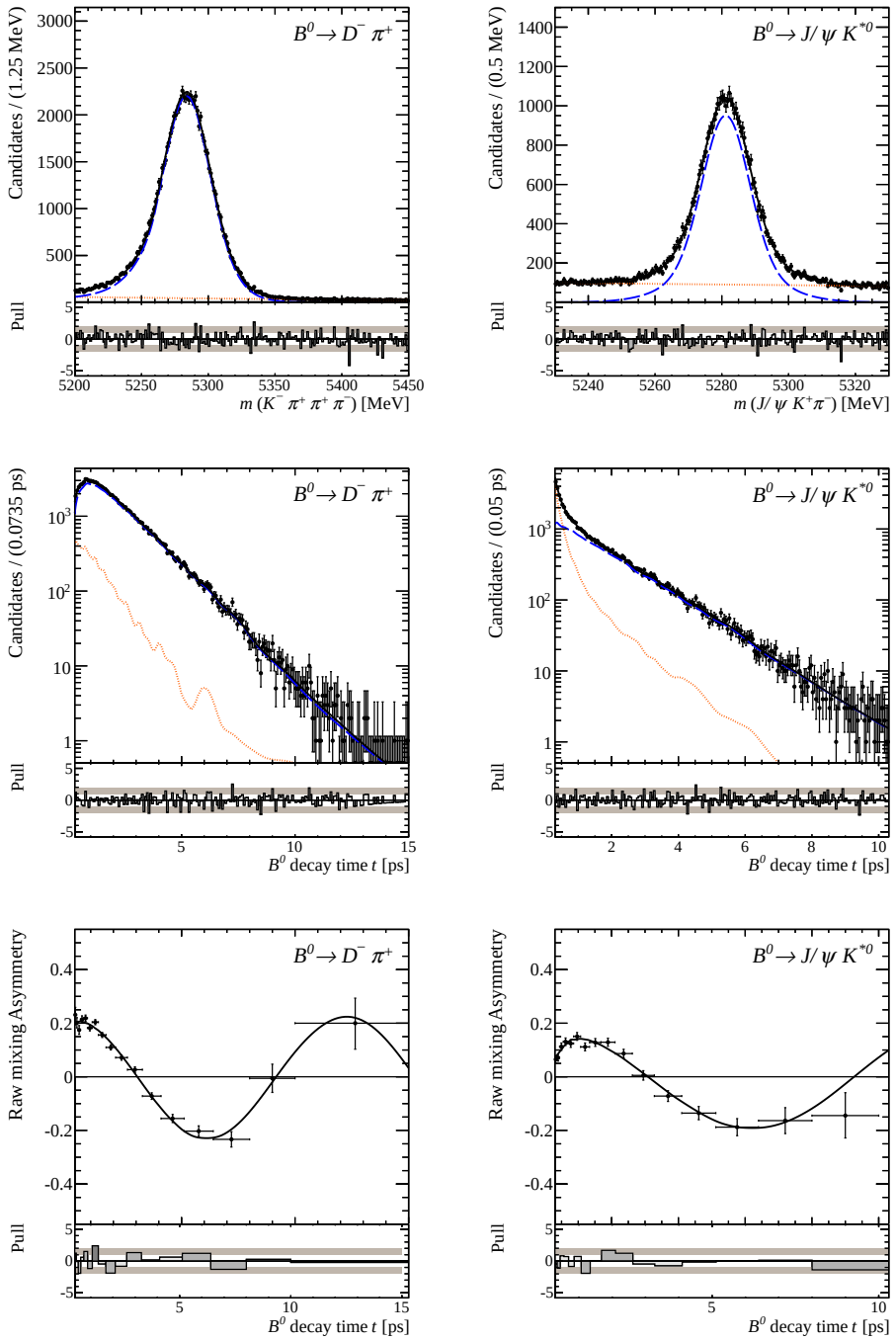


Fig. 11.9: Projections of the cross check in (left) $B^0 \to D^-\pi^+$ and (right) $B^0 \to J/\psi K^{*0}$. The blue line shows the signal, the orange line the combinatorial background and the black line the sum of both. The two component mass fits can be found on the top. The center plots are showing the projections of the s-weighted decay time distributions parametrized by Gaussian kernel pdfs and on the bottom, the resulting mixing asymmetry functions are plotted.

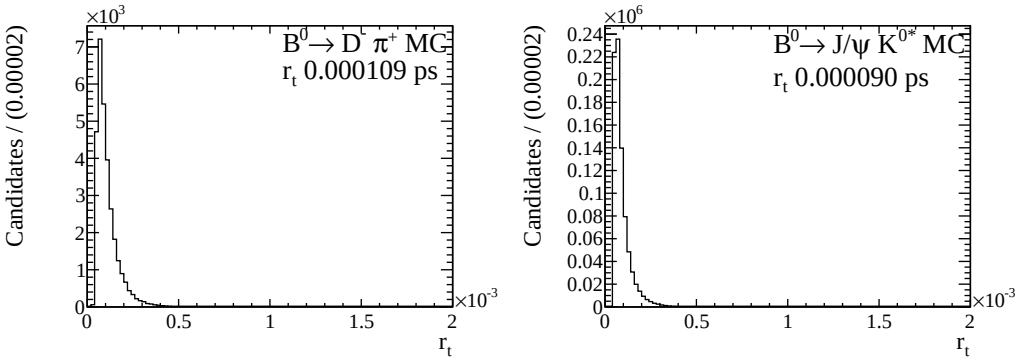


Fig. 11.10: Distributions of the relative differences of the decay times reconstructed with correct momenta and those with momenta modified by the momentum scale uncertainty r_p from MC11.

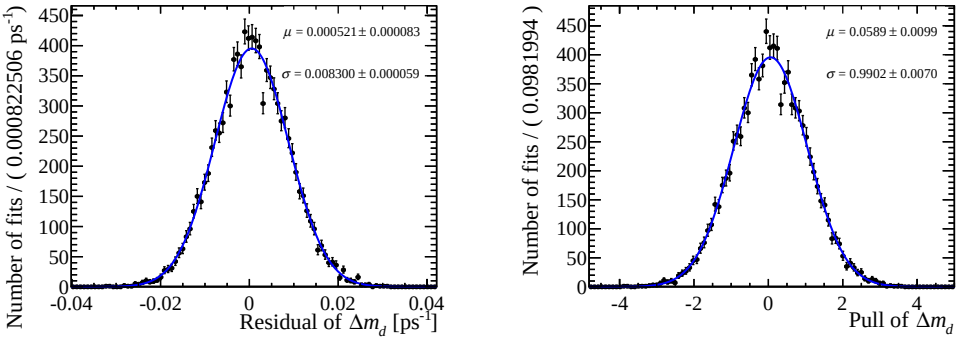


Fig. 11.11: (Left) residuals and (right) pulls on Δm_d in the decay time acceptance toy study for $B^0 \to D^- \pi^+$ with 10000 generated datasets. A Gaussian is fitted to each pull or residual distribution.

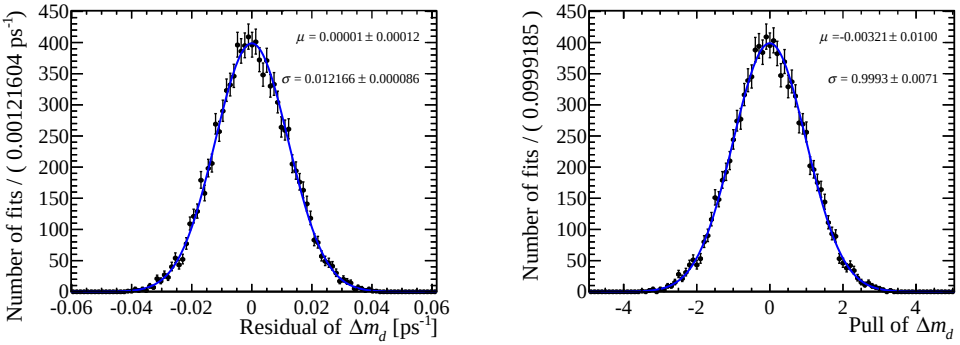


Fig. 11.12: (Left) residuals and (right) pulls on Δm_d in the decay time acceptance toy study for $B^0 \to J/\psi K^{*0}$ with 10000 generated datasets. A Gaussian is fitted to each pull or residual distribution.

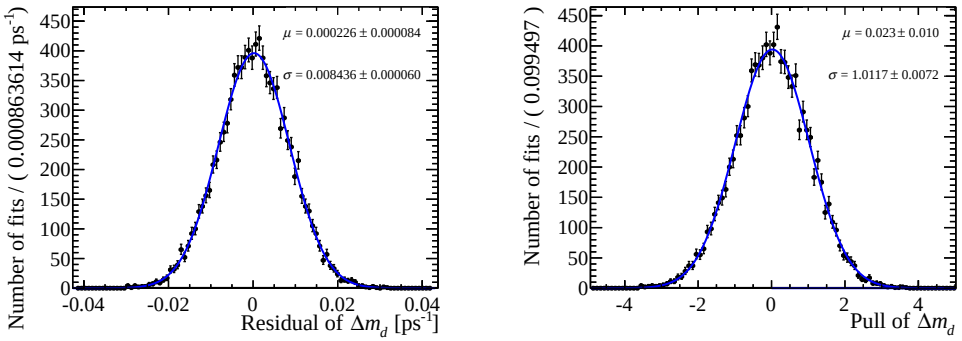


Fig. 11.13: (Left) residuals and (right) pulls on Δm_d in the decay time resolution toy study for $B^0 \to D^- \pi^+$ with 10000 generated datasets. A Gaussian is fitted to each pull or residual distribution.

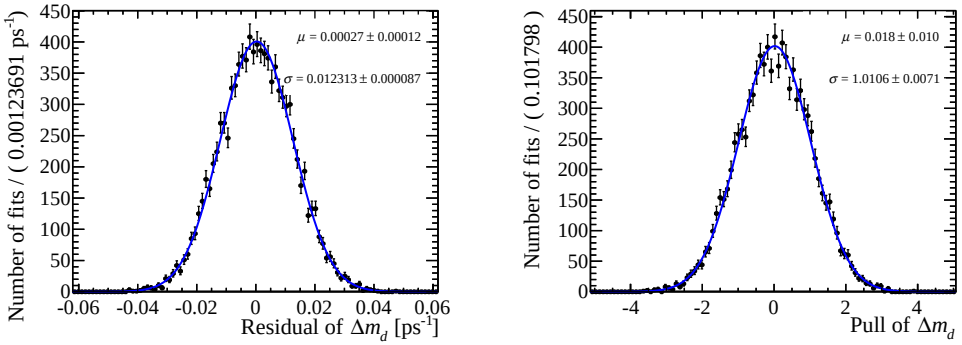


Fig. 11.14: (Left) residuals and (right) pulls on Δm_d in the decay time resolution toy study for $B^0 \to J/\psi K^{*0}$ with 10000 generated datasets. A Gaussian is fitted to each pull or residual distribution.

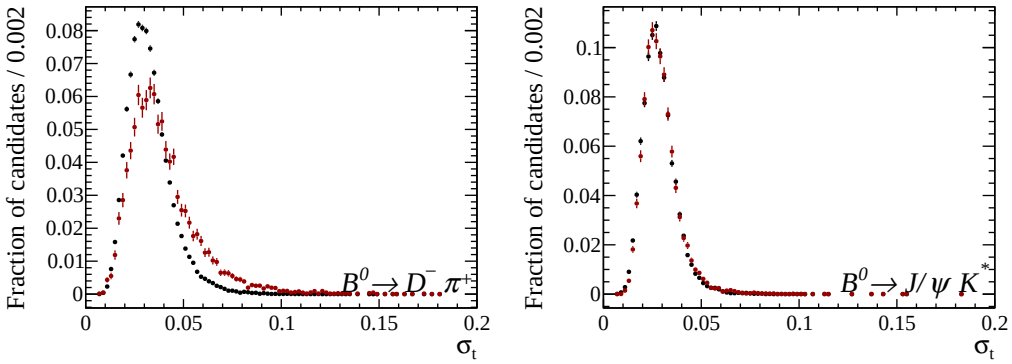


Fig. 11.15: Signal (black) and background (red) Distributions of the per event estimations for the decay time resolution for (left) $B^0 \to D^- \pi^+$ and (right) $B^0 \to J/\psi K^{*0}$ decays.

12 Conclusion

In the work for this thesis a measurement of the $B\overline{B}$ mixing frequency Δm_d has been performed using samples of $B^0 \to D^- \pi^+$ and $B^0 \to J/\psi K^{*0}$ events collected in 1.0 fb^{-1} of pp collisions at $\sqrt{s} = 7\,\text{TeV}$. The measurement yields a value of

$$\Delta m_d = 0.5156 \pm 0.0051\,(\text{stat.}) \pm 0.0033\,(\text{syst.})\,\text{ps}^{-1}$$

and it is currently the most precise measurement of this parameter. The relative uncertainty on Δm_d is 1.2 %, where it is around 0.6 % for Δm_s [12]. Thus, the uncertainty on the ratio $\Delta m_d / \Delta m_s$ is dominated by Δm_d. As the systematic uncertainties in the Δm_d and Δm_s measurements are small, the error on the ratio can be further improved with more data.

Additionally, the measurements have been used to measure the flavour tagging calibration and improve the calibration of the same side pion tagger. The calibration of the opposite side taggers and their combination is found to be suitable for measurements of *CPV*. The studies on the same side pion tagger are useful for the calibration on 2012 data. An inclusion of SSπ tags in the analysis of $\sin 2\beta$ in $B^0 \to J/\psi K^0_S$ decays can lead to an improvement in statistical sensitivity of about 10 %.

The tagging calibration measurements yield differences in the wrong tag probabilities depending on the initial state of the B mesons of up to 2 %. This asymmetry has to be taken into account in time dependent measurements of *CPV*, as they have an influence on the asymmetry distributions and thus the measured amplitudes.

The efforts taken for the measurement of Δm_d will be used for the verification of the tagging calibration on data taken in 2012.

Acknowledgements

I want to express my gratitude to all the people that helped making this work possible, especially to those who did so and did not find their way into this list.

I would like to thank Christophe Cauet and Ulrich Eitschberger for co-authoring the publication with me and for their work on many aspects of the analysis. Many thanks to Stefania Vecchi, Marc Grabalosa, and Marta Calvi for their assistance concerning all questions on flavour tagging and for the lot of work they put into the tagging software. To Stephanie Hansmann-Menzemer who gave me a very intense introduction into the topic and who participated in the analysis as convener of the working group together with Greig Cowan.

The documentation and strategy of the analysis benefitted strongly from the fruitful discussions with Olaf Steinkamp, Manuel Schiller, and Niels Tuning, the referees for the Δm_d measurement and the corresponding publication. I have to thank the members of the editorial board and all other people who commented on the analysis and the publication and helped to improve the paper. Furthermore, I owe a debt of gratitude to Ailsa Sparks, Agnieszka Dziurda, and Vladimir Gligorov for their assistance with the selections for both decay channels that were developed for other analyses and adapted for the Δm_d measurement. For a good start into the depths of the LHCb software and lots of help concerning the simulation of data I thank Gloria Corti and Patrick Robbe.

The measurement had not been possible without the excellent performance of the LHC and the LHCb detector and the people who spent day and night keeping the machines running and collecting high quality data.

The work in a large collaboration requires communication, patience, and a lot of travelling. I am grateful for the support of the German Ministry of Education and Research that gave me the opportunity to meet my international colleagues in person. Only this way a close collaboration was possible.

The development of large parts of the software used for this analysis was done in close collaboration with my colleagues at TU Dortmund University. Thanks a lot for the help with debugging the analysis code, producing n-tuples, optimising cuts, answering statistics questions, proof-reading, ...

Furthermore, I am glad I did not have to nurse our IT infrastructure. Especially the PhiDo-Cluster and our two high performance nodes for fast fitting. Thanks to Matthias Domke, Klaus Wacker, Carsten Raas, Fabian Clevermann, and Danny van

Dyk, who kept the systems running through our analysis. Additional thanks to Julian Wishahi and Florian Kruse who felt responsible for most of our shared software.

I would like to thank my reviewer Reiner Klingenberg for reading this thesis and finally, for answering my questions, supporting my work, and leading the strong company of our working group I just can say *Thank You!* to my supervisor Bernhard Spaan.

Danksagung

Nach dem offiziellen Teil möchte ich mich bei den Menschen bedanken, die mir nicht nur auf direktem Weg bei meiner Arbeit geholfen, sondern mich in den letzten Jahren begleitet und unterstützt haben.

Allen voran möchte ich mich bei meiner Frau Birgit bedanken, die mich in den letzten Jahren unterstützt, meine häufige Abwesenheit (und auch Anwesenheit) ertragen und mich in den anstrengenden Phasen wieder aufgebaut hat. Auch wenn die Bereitschaft mit mir über Physik zu sprechen bei dem Einen oder Anderen in den letzten Jahren aus mir unerfindlichen Gründen merklich abgenommen hat, haben mir alle meine Eltern und meine Familie in allen anderen Bereichen geholfen und mir ermöglicht mich auf die wesentlichen Dinge zu konzentrieren. Danke.

Zwischen den Phasen der Anspannung zu Konferenzen und Veröffentlichungen war ich froh, dass auch Freunde, mit denen ich seit Monaten nicht mehr gesprochen habe, noch wussten wer ich bin und ein offenes Ohr für mich hatten. Vielen Dank für die Ablenkung beim Basketball, Skifahren auf Geburtstagsfeiern, in den Bergen, im Westfalenpark und im Schwimmbad oder einfach nur bei einer Tasse Tee.

Es ist mir ein besonderes Anliegen, all die hervor zu heben, die mir in den letzten Jahren nicht aus dem Weg gehen konnten, und das glücklicherweise auch nicht versucht haben. Meine Kollegen Sebastian, Julian, Florian und Christophe waren immer die ersten Ansprechpartner für viele Probleme. Klaus, Max, Michael Ulrich, Frank, Ramon und Johannes haben im Rahmen ihrer Diplom- und Masterabeiten oder in Gruppen- und Analysebesprechungen an vielen Details der Analyse mitgewirkt. Unsere Ehemaligen Mirco, Magnus, Jesko, Heiko und Moritz waren immer mit Rat und Tat zur Stelle, wenn es um die Untiefen bei der Arbeit in der Hochenergiephysik und in einer Kollaboration ging. Dabei möchte ich nicht die Kollegen Mirco, Stefan, Robert und Moritz aus unserer Hardwareabteilung vergessen die an vielen Punkten wichtige Details beigetragen haben und mich in unzähligen Probevorträgen und Reviews unterstützt haben. Das Wichtigste daran ist jedoch, dass alle Mitglieder der Arbeitsgruppe an vorderster Stelle auch für die Stimmung in der Arbeitsgruppe und für deren Zusammenhalt gearbeitet haben, so dass der Weg ins Büro immer auch ein Weg zu Freunden ist.

Ein ganz besonderer Dank geht an Frau Lorenz. Ohne ihre Hilfe wäre ich auf dem Weg wahrscheinlich irgendwo in den Wirren der Bürokratie verloren gegeangen. Danke für die Erinnerungen, für die Hilfe beim Papierkrieg und für die Unterstützung bei Anträgen und Abrechnungen.

Bibliography

[1] **LHCb Collaboration**, R. Aaij *et al.*, "Measurement of the B^0–$\bar{B}^0$ oscillation frequency Δm_d with the decays $B^0 \to D^- \pi^+$ and $B^0 \to J\,\psi K^{*0}$," *Phys.Lett.* **B719** (2013) 318–325, arXiv:1210.6750 [hep-ex].

[2] **ALEPH, CDF, DELPHI, L3, OPAL and SLD Collaborations**, D. Abbaneo *et al.*, "Combined results on B hadron production rates, lifetimes, oscillations and semileptonic decays," arXiv:hep-ex/0009052 [hep-ex].

[3] **D0 Collaboration**, V. Abazov *et al.*, "Measurement of B^0 mixing using opposite-side flavor tagging," *Phys.Rev.* **D74** (2006) 112002, arXiv:hep-ex/0609034 [hep-ex].

[4] **CDF Collaboration**, T. Affolder *et al.*, "Measurement of the B^0–$\overline{B}^0$ oscillation frequency using $\ell^- D^{*+}$ pairs and lepton flavor tags," *Phys.Rev.* **D60** (1999) 112004, arXiv:hep-ex/9907053 [hep-ex].

[5] **BaBaR Collaboration**, B. Aubert *et al.*, "Measurement of the $\overline{B}^0$ lifetime and the $B^0\overline{B}^0$ oscillation frequency using partially reconstructed $\overline{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ decays," *Phys.Rev.* **D73** (2006) 012004, arXiv:hep-ex/0507054 [hep-ex].

[6] **Belle Collaboration**, K. Abe *et al.*, "Improved measurement of CP-violation parameters $\sin 2\phi_1$ and $|\lambda|$, B-meson lifetimes, and B^0 $\overline{B}^0$ mixing parameter Δm_d," *Phys.Rev.* **D71** (2005) 072003, arXiv:hep-ex/0408111 [hep-ex].

[7] **Particle Data Group**, J. Beringer *et al.*, "Review of Particle Physics (RPP)," *Phys.Rev.* **D86** (2012) 010001.

[8] **CKMfitter Group**, J. Charles *et al.*, "CP violation and the CKM matrix: assessing the impact of the asymmetric B factories," *Eur.Phys.J.* **C41** (2005) 1–131, arXiv:hep-ph/0406184 [hep-ph].

[9] **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 $\mathcal{B}(B \to \tau\nu_\tau)$," *JHEP* **0610** (2006) 081, arXiv:hep-ph/0606167 [hep-ph].

[10] A. Hocker, H. Lacker, S. Laplace, and F. Le Diberder, "A new approach to a global fit of the CKM matrix," *Eur.Phys.J.* **C21** (2001) 225–259, arXiv:hep-ph/0104062 [hep-ph].

[11] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, "Weak decays beyond leading logarithms," *Rev.Mod.Phys.* **68** (1996) 1125–1144, arXiv:hep-ph/9512380 [hep-ph].

[12] **LHCb Collaboration**, R. Aaij *et al.*, "Measurement of the $B_s^0 - \overline{B}_s^0$ oscillation frequency Δm_s in $B_s^0 \to D_s^-(3)\pi$ decays," *Phys.Lett.* **B709** (2012) 177–184, arXiv:1112.4311 [hep-ex].

[13] J. Urban, F. Krauss, C. Hofmann, and G. Soff, "B^0 $\overline{B}^0$ mixing involving supersymmetry," *Mod.Phys.Lett.* **A12** (1997) 419–426.

[14] J. Urban, F. Krauss, U. Jentschura, and G. Soff, "Next-to-leading order QCD corrections for the B^0 $\overline{B}^0$ mixing with an extended Higgs sector," *Nucl.Phys.* **B523** (1998) 40–58, arXiv:hep-ph/9710245 [hep-ph].

[15] M. Grabalosa, *Flavour tagging developments within the LHCb experiment.* PhD thesis, Barcelona U., 2012. CERN-THESIS-2012-075.

[16] **LHCb Collaboration**, R. Aaij and others", "Opposite-side flavour tagging of B mesons at the LHCb experiment," *Eur.Phys.J.* **C72** (2012) 2022, arXiv:1202.4979 [hep-ex].

[17] **UA2 Collaboration**, P. Bagnaia *et al.*, "Evidence for $Z0 \to e + e-$ at the CERN anti-p p Collider," *Phys.Lett.* **B129** (1983) 130–140.

[18] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, "Evidence for the 2π decay of the k_2^0 meson,"*Phys. Rev. Lett.* **13** (Jul, 1964) 138–140. http://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[19] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction,"*Progress of Theoretical Physics* **49** (Feb., 1973) 652–657.

[20] P. W. Higgs, "Broken symmetries and the masses of gauge bosons,"*Phys. Rev. Lett.* **13** (Oct, 1964) 508–509. http://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[21] **ATLAS Collaboration**, G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys.Lett.* **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[22] **CMS Collaboration**, S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys.Lett.* **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[23] **CMS Collaboration**, S. Chatrchyan *et al.*, "On the mass and spin-parity of the Higgs boson candidate via its decays to Z boson pairs," *Phys. Rev. Lett.* **110** (2013) 081803, arXiv:1212.6639 [hep-ex].

[24] **ATLAS Collaboration**, G. Aad *et al.*, "Study of the spin of the new boson with up to 25 fb^{-1} of atlas data," Tech. Rep. ATLAS-CONF-2013-040, CERN, Geneva, Apr, 2013.

[25] **ATLAS Collaboration**, G. Aad *et al.*, "Combined coupling measurements of the higgs-like boson with the atlas detector using up to 25 fb^{-1} of proton-proton collision data," Tech. Rep. ATLAS-CONF-2013-034, CERN, Geneva, Mar, 2013.

[26] G. Branco, L. Lavoura, and J. Silva, *CP Violation*. Oxford Science Publications. Clarendon Press, 1999.

[27] I. Bigi and A. Sanda, *CP violation*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology Series. Cambridge University Press, 2000.

[28] V. V. Gligorov, G. Wilkinson, and J. Rademacker, *Measurement of the CKM angle gamma and B meson lifetimes at the LHCb detector*. PhD thesis, Oxford U., Oxford, UK, 2008. Presented on 28 May 2008.

[29] **ALEPH Collaboration**, D. Decamp *et al.*, "Determination of the Number of Light Neutrino Species," *Phys.Lett.* **B231** (1989) 519.

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

[31] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix,"*Phys. Rev. Lett.* **51** (Nov, 1983) 1945–1947. http://link.aps.org/doi/10.1103/PhysRevLett.51.1945.

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

[33] M. Gell-Mann and A. Pais, "Behavior of neutral particles under charge conjugation,"*Phys. Rev.* **97** (Mar, 1955) 1387–1389. http://link.aps.org/doi/10.1103/PhysRev.97.1387.

[34] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman, and W. Chinowsky, "Observation of long-lived neutral v particles,"*Phys. Rev.* **103** (Sep, 1956) 1901–1904. http://link.aps.org/doi/10.1103/PhysRev.103.1901.

[35] V. Weisskopf and E. Wigner, "Berechnung der natürlichen Linienbreite auf Grund der Diracschen Lichttheorie," *Zeitschrift für Physik* **63** (1930) 54–73. http://dx.doi.org/10.1007/BF01336768.

[36] A. Duncan, E. Eichten, J. Flynn, B. R. Hill, G. Hockney, *et al.*, "Properties of B mesons in lattice QCD," *Phys.Rev.* **D51** (1995) 5101–5129, arXiv:hep-lat/9407025 [hep-lat].

[37] C. W. Bernard, J. N. Labrenz, and A. Soni, "Lattice computation of the decay constants of b and d mesons,"*Phys. Rev. D* **49** (Mar, 1994) 2536–2566. http://link.aps.org/doi/10.1103/PhysRevD.49.2536.

[38] J. E. Mandula and M. C. Ogilvie, "A Lattice calculation of the heavy quark universal form-factor," arXiv:hep-lat/9408006 [hep-lat].

[39] **D0 Collaboration**, V. M. Abazov *et al.*, "Evidence for an anomalous like-sign dimuon charge asymmetry," *Phys.Rev.* **D82** (2010) 032001, arXiv:1005.2757 [hep-ex].

[40] **LHCb Collaboration**, Z. Xing, "Measurement of the flavour-specific CP violating asymmetry $a_s l^s$ in B_s decays," arXiv:1212.1175 [hep-ex].

[41] LHCb Speaker's Bureau, "$b\bar{b}$ production angle plots," march, 2012. http://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[42] CERN public web page, "CERN Accelerator Complex," march, 2012. http://public.web.cern.ch/public/Objects/Research/AccComplex0700829.gif.

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

[44] LHCb public web page, "Hunting for New Physics continues," july, 2011. http://lhcb-public.web.cern.ch/lhcb-public/Images_2011/BsMuMUPt_b.jpg.

[45] G. Barrand, I. Belyaev, P. Binko, M. Cattaneo, R. Chytracek, *et al.*, "GAUDI - A software architecture and framework for building HEP data processing applications," *Comput.Phys.Commun.* **140** (2001) 45–55.

[46] M. Clemencic, H. Degaudenzi, P. Mato, S. Binet, W. Lavrijsen, *et al.*, "Recent developments in the LHCb software framework Gaudi," *J.Phys.Conf.Ser.* **219** (2010) 042006.

[47] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual," *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175 [hep-ph].

[48] I. Belyaev *et al.*, "Handling of the generation of primary events in Gauss, the LHCb simulation framework," *Nuclear Science Symposium Conference Record (NSS/MIC)* **IEEE** (2010) 1155.

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

[50] P. Golonka and Z. Was, "PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays," *Eur.Phys.J.* **C45** (2006) 97–107, arXiv:hep-ph/0506026 [hep-ph].

[51] **GEANT4 Collaboration**, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Dubois, *et al.*, "Geant4 developments and applications," *IEEE Trans.Nucl.Sci.* **53** (2006) 270.

[52] **GEANT4 Collaboration**, S. Agostinelli *et al.*, "GEANT4: a simulation toolkit," *Nucl. Instrum. Meth.* **A506** (2003) 250.

[53] M. Clemencic *et al.*, "The LHCb simulation application, GAUSS: design, evolution and experience," *J. of Phys.: Conf. Ser.* **331** (2011) 032023.

[54] ROOT, "ROOT Version v5-30-00," jan, 2012. `http://root.cern.ch`.

[55] W. Verkerke and D. Kirkby, "Roofit." `http://roofit.sourceforge.net`.

[56] F. James and M. Roos, "MINUIT-a system for function minimization and analysis of the parameter errors and correlations," *Comput. Phys. Commun.* **10** (1975) no. 6, 343–67.

[57] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, *et al.*, "TMVA - Toolkit for Multivariate Data Analysis," *PoS* **ACAT** (2007) 040, `arXiv:physics/0703039 [PHYSICS]`.

[58] **LHCb Collaboration**, R. Aaij *et al.*, "Measurement of the time-dependent *CP* asymmetry in $B^0 \to J/\psi K^0_{\rm S}$ decays," `arXiv:1211.6093 [hep-ex]`.

[59] **LHCb Collaboration**, R. Aaij and others", "Measurement of time-dependent CP-violation observables in $B^0_s \to D^{\mp}_s K^{\pm}$,". LHCb-CONF-2012-029.

[60] W. D. Hulsbergen, "Decay chain fitting with a Kalman filter," *Nucl. Instrum. Meth.* **A552** (2005) 566–575, `physics.comp-ph/0503191v1`. `http://www.sciencedirect.com/science/article/pii/S0168900205013963`.

[61] V. Gligorov, C. Thomas, and M. Williams, "The HLT inclusive B triggers,". LHCb-PUB-2011-016.

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

[63] B. P. Roe, H.-J. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor, "Boosted decision trees as an alternative to artificial neural networks for particle identification," *Nucl.Instrum.Meth.* **A543** (2005) 577–584, `arXiv:physics/0408124 [physics]`.

[64] 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.

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

[66] K. S. Cranmer, "Kernel estimation in high-energy physics," *Comput.Phys.Commun.* **136** (2001) 198–207, arXiv:hep-ex/0011057 [hep-ex].

[67] **ALEPH Collaboration, CDF Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration**, D. Abbaneo *et al.*, "Combined results on b hadron production rates and decay properties," arXiv:hep-ex/0112028 [hep-ex].

[68] **LHCb Collaboration**, R. Aaij *et al.*, "Measurement of b-hadron production fractions in 7 TeVpp collisions," *Phys.Rev.* **D85** (2012) 032008, arXiv:1111.2357 [hep-ex].

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

[70] **LHCb Collaboration**, R. Aaij *et al.*, "Performance of flavour tagging algorithms optimised for the analysis of $B_s^0 \to J/\psi\phi$,". LHCb-CONF-2012-026.

[71] **LHCb Collaboration**, R. Aaij *et al.*, "Measurement of the CP-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\,\phi$," *Phys.Rev.Lett.* **108** (2012) 101803, arXiv:1112.3183 [hep-ex].