Department of Physics and Astronomy

Heidelberg University

Master thesis

in Physics

submitted by

Michael Kolpin

born in Köln

October 2013

Measurement of the $B^0 - \overline{B}^0$ mixing frequency Δm_d in the decay channel $B^0 \to D^-(\to K^+\pi^-\pi^-)\mu^+X$ at the LHCb experiment

This Master thesis has been carried out by Michael Kolpin

at the

Physikalisches Institut Heidelberg

under the supervision of

Prof. Dr. Ulrich Uwer

Abstract

This thesis presents a measurement of the $B^0 - \overline{B}^0$ mixing frequency Δm_d, using the semileptonic decay channel $B^0 \to D^-(\to K^+\pi^-\pi^-)\mu^+ X$.
Semileptonic decays are not fully reconstructed due the non-detection of neutrinos. In this analysis a new approach is used to account for the missing reconstructed B^0 momentum caused by this. Applying this approach to the measurement of Δm_d will allow for the extraction of its value with high precision.
The dataset used in this analysis was collected by the LHCb experiment in 2011 and corresponds to an integrated luminosity of $\mathcal{L} = 1.0\,\text{fb}^{-1}$. The value of Δm_d is extracted from a three-dimensional maximum-likelihood fit. This fit uses information about the mixing state of the B^0 meson and the distributions of the reconstructed B^0 decay time and the invariant D^- mass.
The result of this analysis is a new measurement of the mixing frequency $\Delta m_d = (0.5103 \pm 0.0065(\text{stat.}) \pm 0.0046(\text{syst.}))\,\text{ps}^{-1}$. This is in good agreement with the existing world average. This measurement is the first step of the measurement of the mixing asymmetry a^d_{sl}. Investigating this asymmetry is a way to probe CP violation and search for New Physics.

Zusammenfassung

Diese Arbeit stellt eine Messung der $B^0 - \overline{B}^0$ Mischungsfrequenz Δm_d in semileptonischen Zerfällen $B^0 \to D^-(\to K^+\pi^-\pi^-)\mu^+ X$ vor.
Da Neutrinos vom LHCb Detektor nicht detektiert werden, werden semileptonische Zerfälle nur teilweise rekonstruiert. Diese Analyse verwendet eine neue Methode, um den aufgrund des fehlenden Neutrinos zu klein rekonstruierten B^0-Impuls zu behandeln. Die Anwendung dieser Methode bei der Messung von Δm_d ermöglicht die Bestimmung dieser Größe mit hoher Präzision.
Der in dieser Analyse verwendete Datensatz wurde im Jahr 2011 vom LHCb Experiment aufgezeichnet und entspricht einer integrierten Luminosität von $\mathcal{L} = 1.0\,\text{fb}^{-1}$. Der Wert von Δm_d wird durch einen dreidimensionalen Maximum-Likelihood-Fit bestimmt. Dieser verwendet Informationen über den Mischungszustand der B^0-Mesonen sowie die Verteilung der rekonstruierten B^0-Zerfallszeit und invarianten D^--Masse.
Das Ergebnis dieser Analyse ist ein Wert der Mischungsfrequenz von $\Delta m_d = (0.5103 \pm 0.0065(\text{stat.}) \pm 0.0046(\text{syst.}))\,\text{ps}^{-1}$, was gut mit dem existierenden Weltdurchschnittswert übereinstimmt. Diese Messung ist der erste Schritt in der Messung der Mischungsasymmetrie a^d_{sl}. Die Untersuchung solcher Asymmetrien stellt eine Möglichkeit dar, Verletzungen der CP-Symmetrie zu untersuchen und nach Neuer Physik zu suchen.

Contents

Chapter 1

Introduction

With the discovery of the Higgs boson at the Large Hadron Collider in 2012 all particles predicted by the Standard Model of Particle Physics have been discovered. However, as successful as the Standard Model might be in describing the electromagnetic, weak and strong interactions, it does not include a theory of gravity or give an explanation for dark matter or the matter-antimatter asymmetry in the universe. Therefore, the search for Physics beyond which is described by the Standard Model is one of the primary goals in modern Particle Physics experiments.
There exist many theoretical models trying to expand the Standard Model to be able to describe phenomena for which so far no explanation exists, e.g. the Minimal Supersymmetric Standard Model (MSSM) might present a viable dark matter candidate. Most of those theories also predict particles that have not been observed yet. Therefore, the search for Physics Beyond the Standard Model is directly connected to the search for new particles.
New particles can contribute also as "virtual" particles to quantum loops, thus their existence would change the amplitudes of those processes compared to Standard Model predictions. Decay rates of rare decays and CP violating asymmetries are especially sensitive to New Physics. The LHCb experiment uses the decays of B mesons to investigate these quantities.
When the Large Hadron Collider operates at the instantaneous luminosity of $2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ (the design luminosity of LHCb), about 10^{12} $b\bar{b}$ pairs per year are produced in pp collisions. This large amount of available b hadrons allows for measurements with a precision unobtainable by previous experiments.
In this thesis the measurement of the $B^0 - \overline{B}^0$ mixing frequency Δm_d using semileptonic B^0 decays is presented. As every 10th B^0 meson decays via semileptonic modes, high statistical power is available for analyses involving these modes. However, the LHCb detector cannot detect neutrinos, which leads to an incomplete reconstruction of the invariant B^0 mass and the B^0 momentum in semileptonic decays. Time-dependent analyses use the decay time of B^0 mesons, which is calculated from the flight distance and momentum of the particle. Therefore, the effect of the missing neutrino momentum has to be accounted for.
In this thesis a new approach to correct for the missing momentum is presented and it is applied to the measurement of Δm_d. For the measurement the decay $B^0 \to D^- \mu^+ X$ where the D^- decays further via $D^- \to K^+\pi^-\pi^-$ is used[1]. The dataset used in this analysis was collected by the LHCb experiment in 2011. It corresponds to an integrated luminosity of $1.0\,\mathrm{fb}^{-1}$. The measurement of Δm_d is an important part of the time-dependent measurement of the semileptonic asymmetry a^d_{sl}. In case of a successful implementation of the new method used to account for the missing momentum this method will then further be used in the analysis of this asymmetry.

[1] CP conjugated modes are included

This thesis is structured as follows: Chapter 2 gives an introduction to the theoretical models describing the investigated phenomena. The LHCb experiment is introduced in Chapter 3. In order to measure the mixing frequency Δm_d it is necessary to determine the flavour of a B^0 meson at the time of production and decay. While the latter is identified from the lepton charge, the first can be determined using the technique of flavour tagging, which is briefly introduced in the same chapter.
Chapter 4 gives a short overview over the analysis strategy and introduce the mixing asymmetry a_{sl}^d.
Chapter 5 describes the signal event selection and the background suppression.
The fitting algorithm is introduced in Chapter 6. The used probability density functions describing the distributions of reconstructed B^0 decay time and invariant D^- mass is discussed and the validity of the fit algorithm is tested.
Chapter 7 presents the result of the fit and the uncertainties will be discussed in Chapter 8. The result is discussed and compared to expectations and former results. Furthermore future prospects to improve the results of this analysis and its impact on the a_{sl}^d analysis are briefly discussed.

Chapter 2

Theory

This chapter will provide a brief overview over the Standard Model of Particle Physics and the mechanism of neutral meson mixing described by it. Furthermore, the theoretical mechanisms behind the three categories of CP violation allowed by the Standard Model will be discussed. Also possible experimental approaches to measure these categories will be presented.

2.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) describes electromagnetic, weak and strong interactions at the level of elementary particles. It was developed from the middle of the 20th century onwards by multiple theoretical physicists [1–3], to describe observations made up to that point. From this model several particles were predicted, that had not been observed so far. Shortly after the SM had been developed, many of these predicted particles have been observed in experiments: first evidence for quarks was obtained 1969 at the Stanford Linear Accelerator Center (SLAC) [4]. There the existence of c quarks was shown by the discovery of the J/ψ in 1974 [5], which at the same time was also discovered at the Brookhaven National Laboratory [6]. In 1977 the b quark was discovered at Fermilab and the observation of the t quark followed in 1995 by CDF and D0 [7]. With the detection of the τ neutrino by the DONUT experiment in 2000 [8] and the discovery of the Higgs boson in 2012 at the LHC [9], all particle predictions of the SM have been verified. The Standard Model does not include a theory of gravity as General Relativity and neither contains a viable dark matter candidate. Therefore, it is not a complete theory describing all fundamental interactions. Due to these shortcomings the search for Physics Beyond the Standard Model (BSM), often also referred to as "New Physics", is conducted by current particle physics experiments.
The Standard Model is a quantum field theory, combining Quantum Chromo Dynamics (QCD) with the electroweak theory developed by Glashow, Salam and Weinberg. Local gauge transformations of the symmetry group $\mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ leave its Lagrangian invariant. The interactions described by this theory are a result of this gauge symmetry. More detailed reviews of the Standard Model can be found in various textbooks, e.g. on which this summary is based upon [10–13].

2.1.1 Fundamental particles and interactions

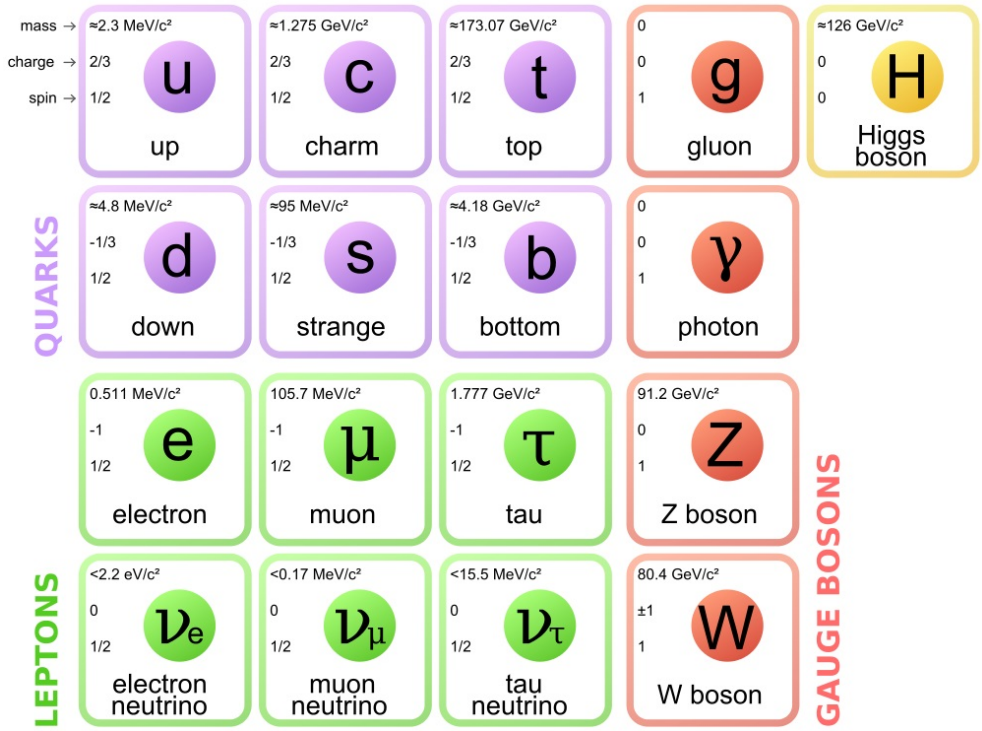


Figure 2.1: *The fundamental particles of the Standard Model [14].*

Figure 2.1 shows the fundamental particles of the Standard Model and their so far measured masses, charges and spins. These fundamental particles are of two basic types, fermions and bosons. Fermions are half-integer spin particles which all matter is composed of. For each of the fermions exists a corresponding antiparticle. The antiparticles have the same mass as their particle counterpart, while all quantum numbers (including charge) are reversed. In the following discussion, they will not be explicitly treated, as antiparticles mostly behave symmetric to their respective particle. Fermions are further divided into leptons and quarks. Leptons are split up into three lepton flavours, which each contain a charged lepton (e^-, μ^- and τ^-) and its corresponding neutrino (ν_e, ν_μ and ν_τ). The e^-, μ^- and τ^- each carry the charge -1, while the neutrinos are neutral. All leptons have the respective lepton flavour quantum number (L_e, L_μ and L_τ) of +1. Processes allowed in the Standard Model conserve these quantum numbers. Therefore, oscillations between neutrinos of different flavours are not included in the Standard Model.
In contrary to leptons, free quarks cannot be observed because of the properties of strong interaction. This phenomenon is called confinement. Quarks carry one of three possible colour charges. Due to the strong interaction, free particles must be colour neutral. Therefore, quarks form so-called hadrons. Hadrons are divided into baryons, consisting of three quarks or three antiquarks (one of each (anti-)colour charge), and mesons, consisting of one quark and one anti-quark (with one colour and the respective anticolour). The six quarks are divided into three generations, each containing a "up-type" quark of charge +2/3 and a "down-type" quark of charge -1/3. Due to that, all hadrons have integer charges.

The gauge bosons have integer spin and are the mediators of the fundamental interactions described by the Standard Model. The photon γ is the carrier of the electromagnetic interaction, which couples to all charged particles. As it is not charged itself, there is no fundamental photon-photon coupling.
The W^+, W^- and Z bosons are the mediators of the weak interaction. They couple to the weak isospin, which is dependent on the chirality of a particle. Due to that, they do not couple to right-handed neutrinos and left-handed antineutrinos. The weak interaction is the only known way for neutrinos to interact.
Gluons are the carriers of the strong interaction. They carry a colour charge and anti-charge and couple to every particle carrying a colour charge. Colour charge is carried by all quarks as well as the gluons themselves. This enables fundamental gluon-gluon coupling. There are eight gluons building a colour octet, which are superpositions of colour/anticolour combinations.
Through the Higgs-mechanism, masses are given to the gauge bosons of the weak interaction, while the photon remains massless. The Higgs boson is an excitation of the scalar field establishing this Higgs-mechanism. It does not mediate a fundamental interaction, as the gauge bosons do. The Higgs-field also couples to fermion fields via Yukawa-couplings, leading to the quark and charged lepton masses. The discovery of the Higgs boson in 2012 at the LHC [9] completes the validation of the Standard Model.

2.1.2 CKM mechanism of the weak interaction

The weak interaction allows for transition between up-type and down-type quarks by transmission of a $W^\pm$ boson. The mass or flavour eigenstates of the quarks are not eigenstates of the weak interaction. The weak interaction allows for transition between different quark generations. The Cabbibo-Kobayashi-Maskawa mechanism [15] describes these transitions. The complex unitary CKM matrix is the matrix of change of basis between flavour (mass) eigenstates (d, s, b) and weak interaction eigenstates (d', s', b') via

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathrm{V_{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} \mathrm{V_{ud}} & \mathrm{V_{us}} & \mathrm{V_{ub}} \\ \mathrm{V_{cd}} & \mathrm{V_{cs}} & \mathrm{V_{cb}} \\ \mathrm{V_{td}} & \mathrm{V_{ts}} & \mathrm{V_{tb}} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.1}$$

The elements of the CKM matrix appear in the SM Lagrangian and therefore in the transition amplitudes for processes involving the weak interaction of quarks. Unitarity reduces the number of free parameters of this 3×3 matrix from 18 to nine. Five more parameters can be absorbed in physically unobservable quark phases. This leads to four physical parameters remaining that completely define $\mathrm{V_{CKM}}$. The standard parameterisation of the CKM matrix uses three Euler angles Θ_{12},Θ_{13} and Θ_{23} and one phase δ. $\mathrm{V_{CKM}}$ is then given by

$$\begin{aligned} \mathrm{V_{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \end{aligned} \tag{2.2}$$

where the abbreviations $c_{ij} = \cos\Theta_{ij}$ and $s_{ij} = \sin\Theta_{ij}$ were used. Another frequently used parameterisation of the CKM matrix is the so-called Wolfenstein parameterisation. It uses the definitions

$$\begin{aligned} \lambda &= s_{12} \\ A &= s_{23}/\lambda^2 \\ \rho + i\eta &= s_{13}e^{i\delta}/A\lambda^3, \end{aligned}$$

where the parameter $\lambda \approx 0.23$ can be used as an expansion parameter. That way $\mathrm{V_{CKM}}$ can be written as

$$\mathrm{V_{CKM}} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.3}$$

The Wolfenstein parameterisation shows that the diagonal elements are close to 1, while off-diagonal elements are small. Transitions within the same quark generation are the most likely, while cross-generation transitions are suppressed. Transitions between the first and second quark generations are only suppressed by a factor λ. This suppression is often also referred to as "Cabbibo" suppression, as these transitions were already allowed in the Cabbibo model, where only two quark generations were described. Other cross-generation transitions are even suppressed by a factor of λ^2 or λ^3, respectively. Processes involving these kind of transitions are often referred to as "CKM suppressed", when compared to processes with a lower order of suppression.
The non-invariance of the Standard Model Lagrangian under CP transformation was thought for a long time to be an explanation for the observed matter-antimatter imbalance in the universe. Today electroweak baryogenesis is disfavoured by the size of the Higgs mass and the size of the CP violation in the quark sector. As mentioned before, the elements of the CKM matrix show up in the transition amplitudes in weak interactions. CP transformation corresponds to hermitian conjugation of these amplitudes. Invariance under CP transformation would therefore reflect in all matrix elements being real, $(\mathrm{V^*_{CKM}})_{\mathrm{ij}} = (\mathrm{V_{CKM}})_{\mathrm{ij}}$. In the parameterisation used in formula 2.2 CP violation in the Standard Model is induced by a non-trivial phase δ. This is the only part of the Standard Model that allows for CP violation. The amount of CP violation predicted by the Standard Model though is not large enough to explain the observed matter-antimatter imbalance. Therefore, investigation of the magnitude of CP violation in the weak interaction is a well suited area in search for Physics beyond the Standard Model.

2.2 Neutral meson mixing

A phenomenon that is sensitive to CP violation is the mixing of neutral mesons. The theory describing it is well established and various literature exists concerning this topic. The summary that is be given in the following sections is based on [16], where it is reviewed more detailed.
Neutral mesons are flavour eigenstates consisting of a quark-antiquark pair, e.g. the B^0 meson

consists of a d and a $\bar{b}$ quark. These flavour eigenstates are not equal to the mass eigenstates, which are the eigenstates of the effective Hamiltonian $\mathcal{H}$. Thus they oscillate into their respective antiparticles via time development. In the following the $B^0 - \overline{B}^0$ mixing will be used to describe the neutral meson mixing, as the B^0 system was investigated in the present thesis. The formalism is as well valid for all other neutral mesons. The effective Hamiltonian is given by

$$\mathcal{H} = M - \frac{i}{2}\Gamma, \tag{2.4}$$

where M and Γ are 2×2 Hermitian matrices. The off-diagonal elements play an important role in CP violation, as will be discussed later in this chapter. Using this effective Hamiltonian, the time development of any linear combination of flavour eigenstates, $a|B^0\rangle + b|\overline{B}^0\rangle$, is described by the time-dependent Schrödinger equation

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \mathcal{H}\begin{pmatrix} a \\ b \end{pmatrix} = (M - \frac{i}{2}\Gamma)\begin{pmatrix} a \\ b \end{pmatrix}, \tag{2.5}$$

in the basis of the flavour eigenstates. $\mathcal{H}$ has two eigenstates, the **h**igh and **l**ow mass eigenstates $|B_H\rangle$ and $|B_L\rangle$, which compose of the flavour eigenstates as

$$\begin{aligned} |B_H\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle, \\ |B_L\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle. \end{aligned} \tag{2.6}$$

The complex coefficients q and p have to fulfill the normalisation condition $|p|^2 + |q|^2 = 1$. The time development of these mass eigenstates then takes the form

$$\begin{aligned} |B_H\rangle(t) &= e^{-iM_H t}e^{-\frac{\Gamma_H}{2}t}|B_H\rangle, \\ |B_L\rangle(t) &= e^{-iM_L t}e^{-\frac{\Gamma_L}{2}t}|B_L\rangle, \end{aligned} \tag{2.7}$$

with the eigenvalues $M_{H/L}$ and $\Gamma_{H/L}$ of the matrices M and Γ respectively for each of the two eigenstates. The mass difference Δm_d and decay width difference $\Delta\Gamma_d$ can be defined as

$$\Delta m_d = M_H - M_L, \tag{2.8}$$

$$\Delta\Gamma_d = \Gamma_L - \Gamma_H. \tag{2.9}$$

This way, Δm_d is constrained to be positive. To $\mathcal{O}(10^{-2})$ accuracy, the previously introduced quantities can be written as

$$\begin{aligned} \Delta m_d &\approx 2|M_{12}|, \\ \Delta\Gamma_d &\approx 2\Re(M_{12}\Gamma_{12}^*)/|M_{12}|, \\ q/p &\approx -|M_{12}|/M_{12}. \end{aligned} \tag{2.10}$$

Using equations 2.6 and 2.7 the time development of physical states, consisting at $t = 0$ of pure B^0 ($|B^0_{phys}\rangle$) or $\overline{B}^0$ ($|\overline{B}^0_{phys}\rangle$) respectively, can be expressed as

$$\begin{aligned} |B^0_{phys}\rangle(t) &= g_+(t)|B^0\rangle + (q/p)g_-(t)|\overline{B}^0\rangle, \\ |\overline{B}^0_{phys}\rangle(t) &= (p/q)g_-(t)|B^0\rangle + g_+(t)|\overline{B}^0\rangle. \end{aligned} \tag{2.11}$$

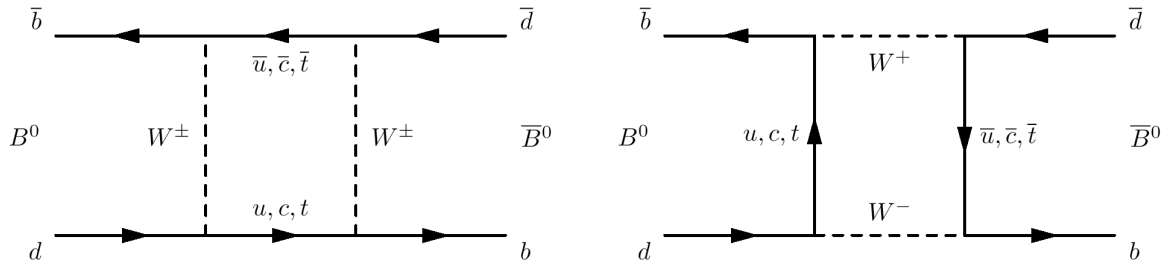


Figure 2.2: *Box diagrams contributing to $B^0 - \overline{B}^0$ mixing [17].*

CP conservation would correspond to $|\langle\overline{B}^0|B^0_{phys}\rangle(t)|^2 = |\langle B^0|\overline{B}^0_{phys}\rangle(t)|^2$, which is fulfilled only if $|q/p| = 1$. Neglecting the difference in Γ_H and Γ_L (as $\Delta\Gamma_d/\Gamma_d = \mathcal{O}(10^{-2})$), the coefficients $g_-(t)$ and $g_+(t)$ are given as

$$\begin{aligned} g_+(t) &= e^{-i\frac{M_H+M_L}{2}t}e^{-\Gamma_d t}\cos(\Delta m_d\, t/2), \\ g_-(t) &= e^{-i\frac{M_H+M_L}{2}t}e^{-\Gamma_d t} i\sin(\Delta m_d\, t/2). \end{aligned} \tag{2.12}$$

This corresponds to an oscillation of B^0 into $\overline{B}^0$ and back with the mixing frequency Δm_d overlayed by the decay of these states with the decay width $\Gamma_d \approx \Gamma_H \approx \Gamma_L$. In the Standard Model, these oscillations happen via box-diagrams, as shown for the B^0 meson in fig. 2.2. Intermediate u, c and t quarks contribute to these diagrams, but u and c contributions are heavily suppressed by the GIM mechanism [18]. Therefore, the t contribution dominates this process. The relevant CKM matrix elements involved in these diagrams are $|V_{tb}|$ and $|V_{td}|$ in $B^0 - \overline{B}^0$ mixing. As $|V_{td}|$ is an order of magnitude smaller than $|V_{ts}|$, which would be the corresponding element relevant for $B^0_s - \overline{B}^0_s$ mixing, the mixing in the B^0 sector is significantly slower than in the B^0_s sector. Expressed in terms of the mass differences, $\Delta m_d < \Delta m_s$. Precise knowledge of these values is important in the evaluation of mixing asymmetries and the investigation of *CP* violation derived from them, as will be explained in the next section.

2.3 Mixing asymmetry and CP violation

As formulated by Sakharov [19] one of the necessary conditions for the baryon asymmetry of the universe is CP violation, which is to a certain degree allowed in the Standard Model. The measurement of asymmetries between a decay and its CP conjugate is a way to investigate the magnitude of this CP violation and study if it exceeds the Standard Model prediction, giving hints towards New Physics. The advantage of measuring asymmetries is that due to their definition many systematical effects cancel. That way high precision can be achieved.
The flavour eigenstates of the B^0 system are related through CP transformation

$$\begin{aligned} CP|B^0\rangle &= -|\overline{B}^0\rangle, \\ CP|\overline{B}^0\rangle &= -|B^0\rangle, \end{aligned} \tag{2.13}$$

where a possible non-physical complex phase-factor has been neglected. Similarly, one can define a possible final state as $|f\rangle$ and its CP conjugate $CP|f\rangle = \pm|\overline{f}\rangle$, dropping again the phase-factor. From now on decay amplitudes will be written as

$$\begin{aligned} A_f &= \langle f|\mathcal{H}|B^0\rangle, \\ \overline{A}_f &= \langle f|\mathcal{H}|\overline{B}^0\rangle, \\ A_{\overline{f}} &= \langle \overline{f}|\mathcal{H}|B^0\rangle, \\ \overline{A}_{\overline{f}} &= \langle \overline{f}|\mathcal{H}|\overline{B}^0\rangle. \end{aligned} \tag{2.14}$$

CP violation is usually classified in three different categories:

1. CP violation in decay
2. CP violation in mixing
3. CP violation in interference between decays with and without mixing

For each of these categories different observables can be used to quantify the CP violation, which will further be discussed.

2.3.1 CP Violation in Decay

CP violation can be observed when comparing a decay $B \to f$ and its CP conjugate $\overline{B} \to \overline{f}$. This category of CP violation is often also referred to as "direct" CP violation. An observable in which CP violation in decay can be observed in is the absolute value of the ratio of the decay amplitudes:

$$\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right|. \tag{2.15}$$

From $\left|\overline{A}_{\overline{f}}/A_f\right| \neq 1$ one can immediately conclude CP violation, as the decay $B \to f$ and its CP conjugate differ. The quantity $\left|\overline{A}_{\overline{f}}/A_f\right|$ and thereby the degree of direct CP violation can be measured or example in charged B meson decays:

$$a_f = \frac{\Gamma(B^+ \to f) - \Gamma(B^- \to \overline{f})}{\Gamma(B^+ \to f) + \Gamma(B^- \to \overline{f})} = \frac{1 - \left|\overline{A}_{\overline{f}}/A_f\right|^2}{1 + \left|\overline{A}_{\overline{f}}/A_f\right|^2}, \tag{2.16}$$

where a_f is the asymmetry between the decay $B^+ \to f$ and its CP conjugate, $a_f \neq 0$ corresponds to CP direct violation.

2.3.2 CP Violation in Mixing

Another category of CP violation is its occurrence in mixing and is often also called "indirect" CP violation. It occurs if

$$\left|\frac{q}{p}\right|^2 = \left|\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}\right| \neq 1, \tag{2.17}$$

where q and p are the coefficients introduced in formula (2.6). M_{12} and Γ_{12} are the off-diagonal elements of the matrices appearing in the effective Hamiltonian from formula (2.4) in the basis of the flavour eigenstates. CP violation in mixing corresponds to the probability for a B^0 to oscillate into a $\overline{B}^0$ differing from the probability for a $\overline{B}^0$ to oscillate into a B^0. It can be observed in the mixing asymmetry of neutral mesons, for example in the case of semileptonic B^0 decays:

$$a_{sl} = \frac{\Gamma(\overline{B}^0_{phys}(t) \to l^+\nu_\ell X) - \Gamma(B^0_{phys}(t) \to l^-\overline{\nu}_\ell X)}{\Gamma(\overline{B}^0_{phys}(t) \to l^+\nu_\ell X) + \Gamma(B^0_{phys}(t) \to l^-\overline{\nu}_\ell X)}. \tag{2.18}$$

Again, the definition of this asymmetry leads to $a_{sl} = 0$, if CP is conserved. The decay amplitude $A = \langle l^+\nu_\ell X|\mathcal{H}|B^0\rangle$ is describing the semileptonic decay of the B^0 meson. Using

$$\langle l^-\overline{\nu}_\ell X|\mathcal{H}|B^0_{phys}(t)\rangle = (q/p)g_-(t)\overline{A},$$

and

$$\langle l^+\nu_\ell X|\mathcal{H}|\overline{B}^0_{phys}(t)\rangle = (p/q)g_-(t)A,$$

the asymmetry can be expressed as

$$a_{sl} = \frac{1 - |q/p|^4}{1 + |q/p|^4}. \tag{2.19}$$

The effects of CP violation from Standard Model calculations in these semileptonic B^0 decays are expected to be small ($\mathcal{O}(10^{-2})$). Therefore, to investigate effects that differ from these expectations (or otherwise to confirm them), experiments measuring them need to achieve a high accuracy.
The B^0 meson decays semileptonically with a high branching ratio ($\mathcal{BR}(B^0 \to l^+\nu_\ell X) = (10.33 \pm 0.28)\%$), thus a large number of B^0 mesons are available to probe very small asymmetries.

2.3.3 CP Violation in Interference of Decays with and without Mixing

The last category of CP violation is its occurrence in the decay of neutral B mesons into final states that are CP eigenstates, denoted as f_{CP}. The quantity sensitive to this category is

$$\lambda = \eta_{f_{CP}} \frac{q}{p} \frac{\overline{A}_{f_{CP}}}{A_{f_{CP}}}, \tag{2.20}$$

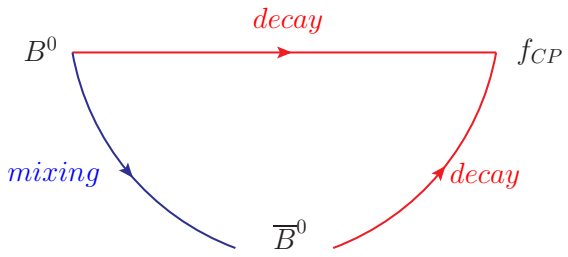


Figure 2.3: *Interference of decays with and without mixing.*

where $\eta_{f_{CP}} = \pm 1$ is the CP eigenvalue of f_{CP}. As discussed in the two former categories, CP conservation implies $|q/p| = 1$ and $|\overline{A}_{\overline{f}_{CP}}/A_{f_{CP}}| = 1$. In addition to that, the relative phase between q/p and $\overline{A}_{\overline{f}_{CP}}/A_{f_{CP}}$ has to vanish. This phase results from interference between the direct decay $B^0 \to f_{CP}$ with the decay $\overline{B}^0 \to f_{CP}$, where the $\overline{B}^0$ originated from oscillation of a B^0. This is illustrated in fig. 2.3. This leads to the presence of CP violation if $\lambda \neq \pm 1$. λ can be observed by measuring the asymmetry in the decay of neutral B mesons into CP eigenstates

$$a_{f_{CP}} = \frac{\Gamma(B^0_{phys}(t) \to f_{CP}) - \Gamma(\overline{B}^0_{phys}(t) \to f_{CP})}{\Gamma(B^0_{phys}(t) \to f_{CP}) + \Gamma(\overline{B}^0_{phys}(t) \to f_{CP})}. \tag{2.21}$$

The relation between $a_{f_{CP}}$ and λ is complicated:

$$a_{f_{CP}} = \frac{(1 - |\lambda_{f_{CP}}|^2)\cos(\Delta m_B t) - 2\Im(\lambda_{f_{CP}})\sin(\Delta m_B t)}{1 + |\lambda_{f_{CP}}|^2}. \tag{2.22}$$

So-called "clean" decay modes are those in which the first two categories of CP violation do not occur. This leads to $|\lambda_{f_{CP}}| = 1$ for these modes and thus $a_{f_{CP}}$ simplifies to

$$a_{f_{CP}} = -\Im(\lambda_{f_{CP}})\sin(\Delta m_B t). \tag{2.23}$$

Another interesting asymmetry type related to CP violation in interference of decays with and without mixing can be obtained by observing the decay of a B^0 into two different CP eigenstates $f_{CP,a}$ and $f_{CP,b}$. The quantity

$$\eta_a \lambda_a - \eta_b \lambda_b = \frac{q}{p}\left(\frac{\overline{A}_{\overline{a}}}{A_a} - \frac{\overline{A}_{\overline{b}}}{A_b}\right), \tag{2.24}$$

is $\neq 0$ if CP violation occurs in the clean modes. The advantage of investigating the latter asymmetry is that production and detection asymmetries, that contribute to both the decays to final states $f_{CP,a}$ and $f_{CP,b}$ cancel due to the subtraction.

Chapter 3

The LHCb experiment

This chapter will describe the Large Hadron Collider beauty (LHCb) experiment. The Large Hadron Collider (LHC), at which it takes places is also introduced. The detector components important for this analysis will be discussed and furthermore the data and simulation samples used in this analysis will be presented.

3.1 The Large Hadron Collider

The Large Hadron Collider is a proton-proton collider located at the European Laboratory for Particle Physics CERN in Geneva, Switzerland. It is the largest particle accelerator build so far and is designed to operate at center-of-mass energy of $2 \times 7\,\text{TeV}$. The four major particle physics experiments at the LHC are LHCb, ATLAS, CMS and ALICE. Figure 3.1 schematically shows the LHC and the locations of these four experiments. The LHC tunnel is 27 km long, where each of the two proton beams is accelerated and both are brought to collision at the four interaction points at which the experiments are placed. The design value

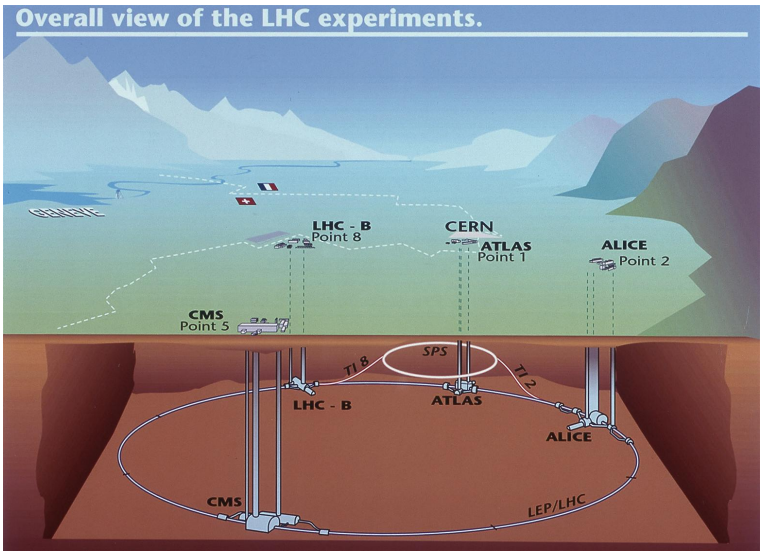


Figure 3.1: *A schematic overview of the Large Hadron Collider and the placement of the four major experiments [17].*

for the instantaneous luminosity is $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, at which each proton beam is separated in 2808 bunches. Each of these bunches contains $\sim 10^{11}$ protons. The time distance between two bunches is 25 ns, resulting in an interaction rate of 40 MHz. At start-up the LHC was operated for 10 days in 2008 after which technical defects in the accelerator's magnets forced it to be shut down. After finishing repairs in 2009 test runs were made at a center-of-mass energy of $\sqrt{s}$ =450 GeV. In 2010 the LHC operated at a center-of-mass energy of $\sqrt{s}$ =7 TeV, reaching a peak instantaneous luminosity of $\sim 2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The center-of-mass energy was increased to $\sqrt{s}$ =8 TeV in 2012. At the same time also a significant increase in the integrated luminosity to up to $\sim 8 \cdot 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ was achieved. Currently the LHC is shut down to prepare it for the operation at its design collision energy of $\sqrt{s}$ =14 TeV. It is planed to continue data collection in 2015.

3.2 The LHCb detector

The purpose of the LHCb experiment is the investigation of CP violation in the decay of B and D mesons and the search for rare B decays. In the LHCb detector b quark pairs are mainly created by gluon fusion or the annihilation of a $q\bar{q}$ pair. The momentum fraction carried by the two colliding partons can differ by a lot. This leads to a massive boost of the rest frame of the created $b\bar{b}$ pair along the beam axis. Figure 3.2 shows the distribution of the polar angles of the b and $\bar{b}$ quarks in the detector frame derived from simulation. Thus b hadron

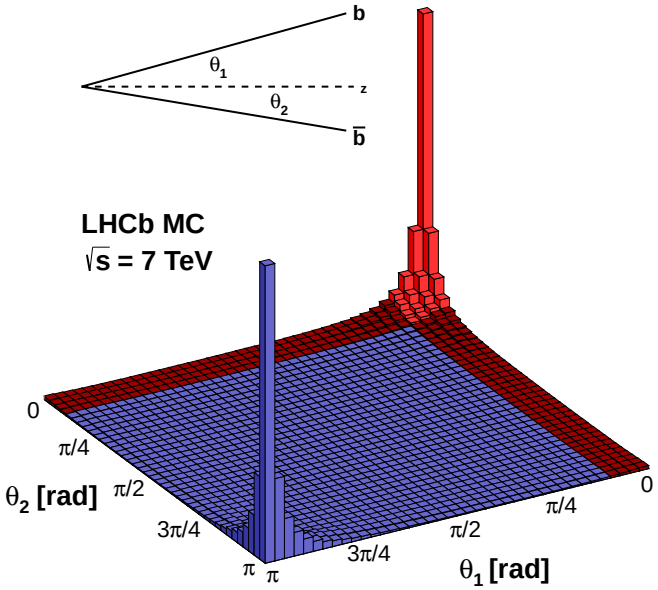


Figure 3.2: *Simulated distribution of $b\bar{b}$ polar angles in the detector frame [17].*

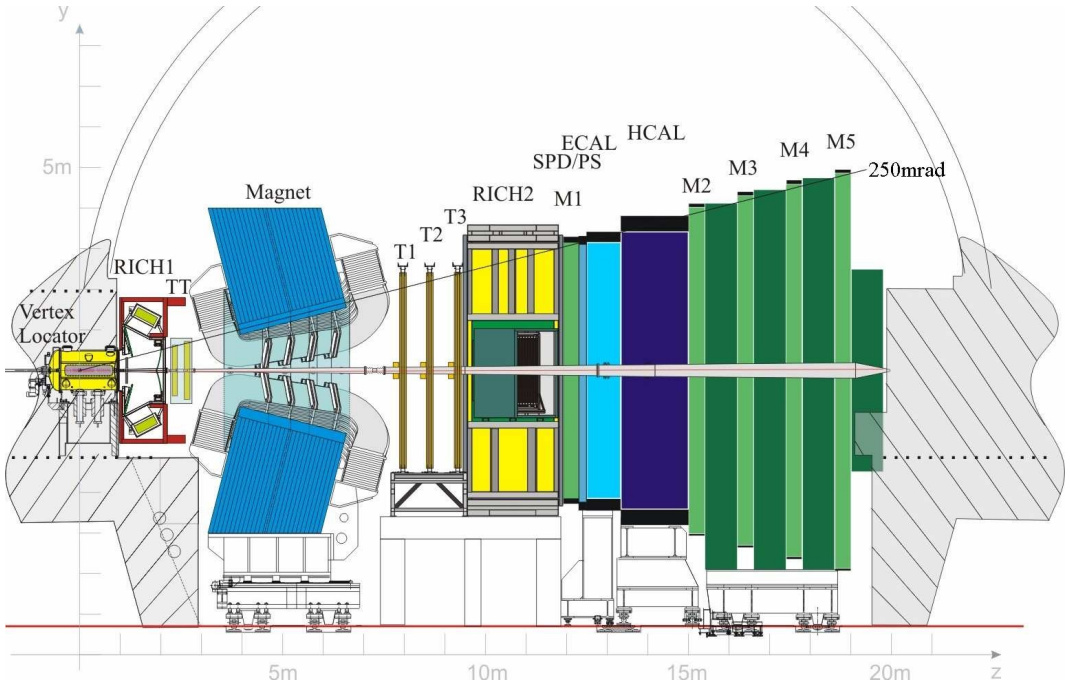


Figure 3.3: *Schematic overview of the LHCb detector and its components [20].*

pairs are mostly produced in a common forward or backward cone. As the LHCb experiment is especially interested in B meson decays, it is therefore not designed to cover the whole solid angle, in contrast to the 4π general purpose detectors ATLAS and CMS. The LHCb detector is shown schematically in fig. 3.3. It covers a forward angle from 10 mrad to 250 mrad in the non-bending plane (y-direction) and from 10 mrad to 300 mrad in the bending plane (x-direction). The detector covers only the forward direction, due to the limited space available in the detector site inherited from the DELPHI experiment. However, about 25% of the B mesons produced in proton-proton collisions can be detected within the angular acceptance of the LHCb detector. The detector is designed to operate at an instantaneous luminosity of $\mathcal{L} = 2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, significantly lower than the LHC design luminosity, to maximise the number of events with only a single interaction. Figure 3.4 shows the integrated luminosities recorded by the LHCb experiment in the years 2010 to 2012. While the data collection run of 2010 delivered about 40 nb^{-1}, it increased significantly in 2011, where data corresponding to an integrated luminosity of 1.1 fb^{-1} was recorded. This increased further in 2012 to $2.1\,\mathrm{fb}^{-1}$. The total data corresponds to an integrated luminosity of $3.2\,\mathrm{fb}^{-1}$.

The LHCb detector consists of various components for tracking, momentum measurement, calorimetry and particle identification.

The Vertex Locator (VELO) surrounds the interaction point and used to locate primary and secondary vertices with high precision. Besides the VELO, the tracking system of the LHCb detector consists of four tracking stations: the Tracker Turicensis (TT) directly downstream of the VELO and the three T-stations (T1, T2 and T3) downstream of the dipole magnet. While the VELO and TT use silicon microstrip detectors, the T-stations use silicon microstrip sensors only close to the beam pipe. This region is also referred to as Inner Tracker (IT). The outer area of the T-stations, the Outer Tracker (OT) uses a gas detector based on straw-tube

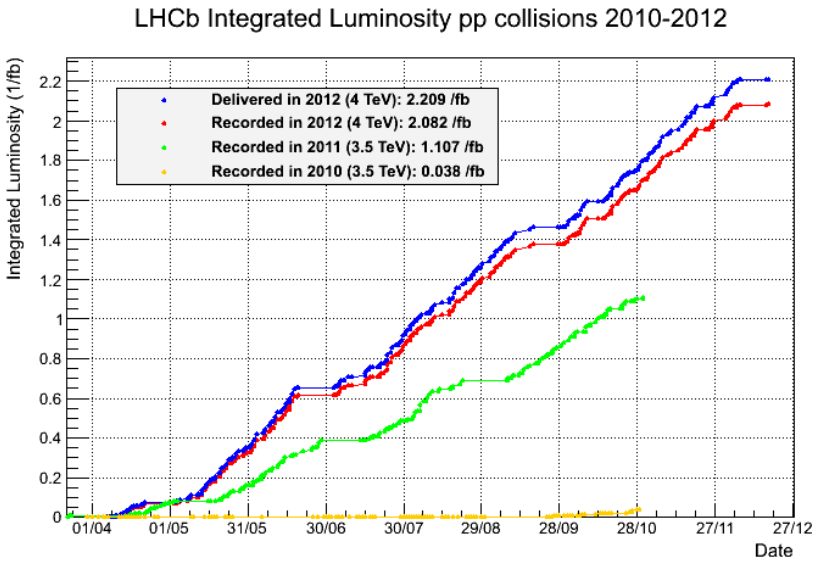


Figure 3.4: *Integrated luminosities recorded by the LHCb detector in the years 2010 to 2012 [17].*

technology to track charged particle movement.
The dipole magnet is used to bend a charged particle's path depending on its charge. It affects charged particles over a length of 10 m with an integrated magnetic field of 4 Tm. An important feature of this magnet is the ability to invert its polarity. That way asymmetries induced by the detector geometry can be canceled.
The two Ring Imaging Cherenkov Detectors (RICH) are used to identify particles, especially to distinguish pions from kaons. They each cover different momentum and angular acceptance ranges and will be described more detailed later on.
The LHCb detector uses four muon chambers, one upstream of the calorimeters and three downstream of them. They are used for particle identification as well as for the trigger of LHCb.
The electromagnetic (ECAL) and hadronic calorimeters (HCAL) serve multiple purposes. They are used in the first stage of the trigger system, the level-0 trigger, to select events with high E_{T} and p_{T}. Furthermore, they measure the energies of the respective particles and identify electrons, photons and hadrons.

3.2.1 Vertex Locator

The Vertex Locator is an important part of the tracking system of the LHCb detector. It is placed close to the interaction region, to provide high sensitivity in the measurement of the track coordinates. These are used to precisely identify displaced secondary vertices. Figure 3.5 schematically shows the placement of the VELO components. The VELO consists of multiple silicon strip detectors arranged in 21 stations, which are placed alongside the beam direction. Each VELO station consists of two subsensors, one measuring the distance from the beam axis R, the other the polar angle ϕ. An important goal of the VELO is to achieve high sensitivity

for the location of even those secondary vertices that are only slightly displaced from the primary vertex. Therefore the VELO stations are placed at 7 mm distance from the beam axis. However, when a new beam is being injected into the LHC it takes some time until it is stable. To prevent high radiation damage during this procedure, the subsensors can be opened.

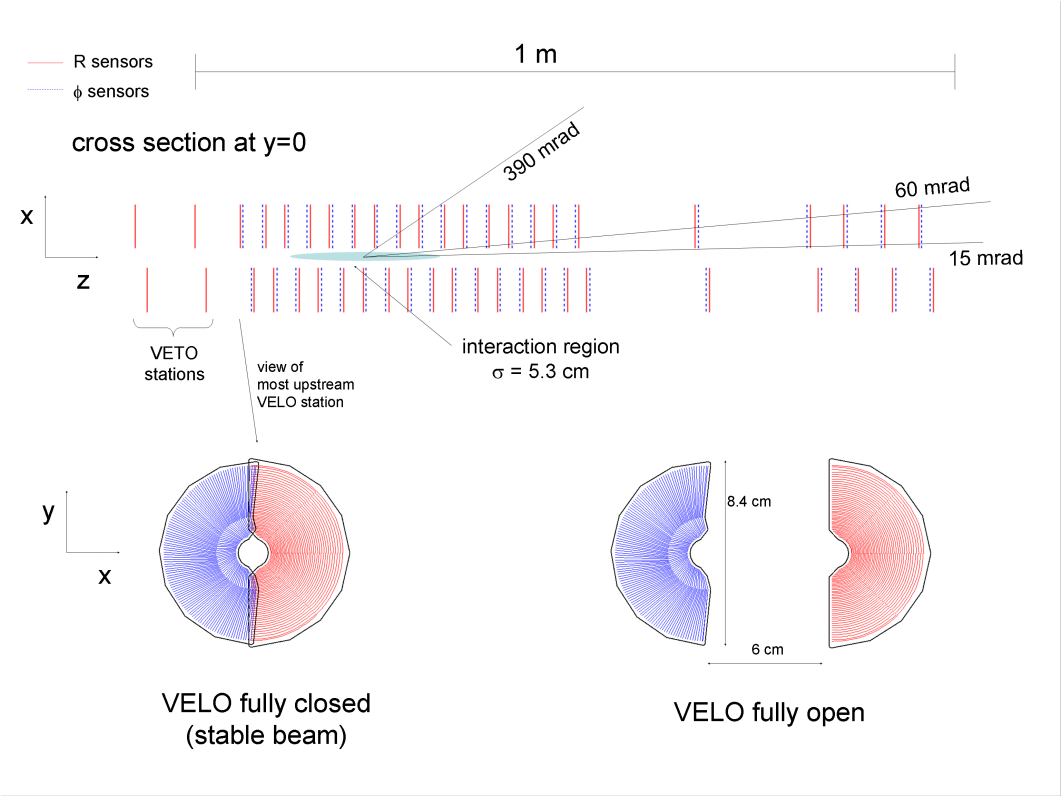


Figure 3.5: *Schematical overview of the placement of the Vertex Locator along the beam axis. The subsensors of each station can be opened during beam injection [20].*

3.2.2 Ring Imaging Cherenkov Detectors

The Ring Imaging Cherenkov Detectors RICH 1 and RICH 2 are used in the particle identification process. The RICH 1 detector is placed upstream of the dipole magnet. It provides particle identification for charged particles in the momentum range from about 1 GeV/c to 60 GeV/c and covers the angular range from 25 mrad to 300 mrad (x-direction) and 250 mrad (y-direction), respectively. The RICH 2 detector is placed upstream of the dipole magnet and covers a smaller angular range than RICH 1, from 15 mrad to 120 mrad and 100 mrad, respectively. However, this range covers the region at which high momentum particles can be detected (from ∼15 GeV/c to 100 GeV/c). Figure 3.6 shows the layout of the RICH 1 detector. The Cherenkov detectors exploit the Cherenkov effect: charged particles which pass a medium with a velocity that is higher than the speed of light in that medium emit photons.

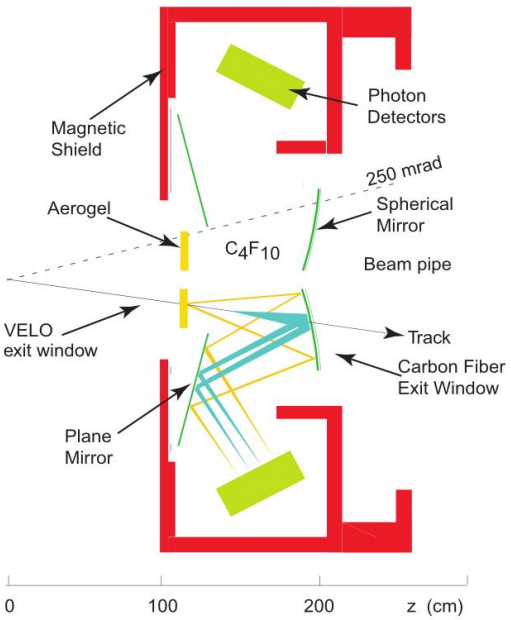


Figure 3.6: *Schematical overview of the layout of the RICH 1 detector [20].*

The angle Θ under which these photons are emitted is directly related to the particle's velocity $\beta = v/c$ via

$$\cos\Theta = \frac{1}{\beta\, n}, \tag{3.1}$$

where n is the refractive index of the medium. The medium used in RICH 1 is aerogel and C_4F_{10} gas, while RICH 2 uses CF_4 gas as radiator. In the RICH detectors the light emitted by charged particles is guided to the Hybrid Photo Detectors (HPD) via a system of mirrors. The photons emitted by the particles form a cone which is detected as a ring. The radius of these rings is depending on the angle Θ. Combined with the measurement of the momentum of the corresponding particle, a hypothesis can be made concerning its identity. The particle identification process also uses information from the calorimeters and the muon chambers. Its output used in the offline analysis is not a distinct identity, but a likelihood value, the DLL. This is the difference in the logarithm of the likelihoods of two particle hypothesis. For example $\mathrm{DLL}_{K\pi} = \log\mathcal{L}_K - \log\mathcal{L}_\pi$ denotes the difference of the hypotheses of the respective particle being a kaon or a pion. The kaon identification efficiency of the RICH detectors is about 95% in the momentum range from 2 - 100 GeV/c [21].

3.2.3 Muon system

Muons are present in the final states of many of the decays which are sensitive to CP violation and are investigated by the LHCb experiment. Therefore muon identification and triggering are essential for the LHCb detector. The muon system consists of five muon stations, one (M1) placed in front of the calorimeter system and four (M2-5) behind it. Figure 3.7 shows

the arrangement of these stations. All of these stations consist of multiple muon chambers, of which 1380 are used in the whole system. M1 is placed in front of the calorimeter to deliver an accurate p_{T} measurement for the trigger system. Stations M1-M3 have a high spatial resolution in the bending plane (x-direction), to define the direction of the tracks. Assuming the particle originates from the vertex they measure the particle's transverse momentum p_{T} with a resolution of 20%. The spatial resolution in M4 and M5 is lower, as their main purpose is to identify particles that penetrate all of the absorption layers. The muon chambers consists of multi-wire proportional chambers (MWPC) which are used to track the trajectory of the muons. An exception is the inner region of M1, where a high particle flux is expected and a triple-GEM (gas electron multiplier) detector is used.

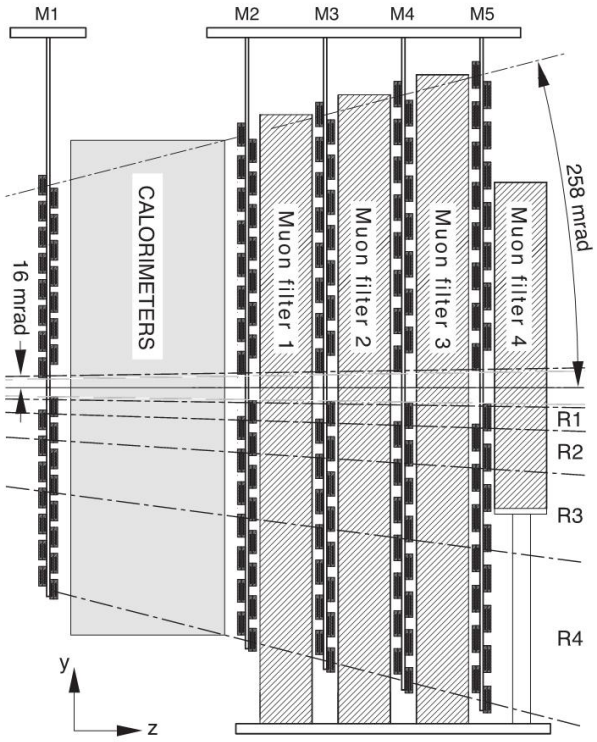


Figure 3.7: *Schematical overview of the LHCb muon system [20].*

3.3 The LHCb trigger system

The maximum possible beam collision rate at the LHC is about 40 MHz. However, to be able to store data for further offline analysis, this has to be reduced to a rate of about 5 kHz. To achieve this, two trigger levels are used, the Level-0 (L0) and High Level Trigger (HLT). Figure 3.8 schematically shows the trigger process. The L0 trigger is designed to reduce the bunch crossing rate of 40 MHz to a rate of 1.1 MHz, at which the readout electronics of the

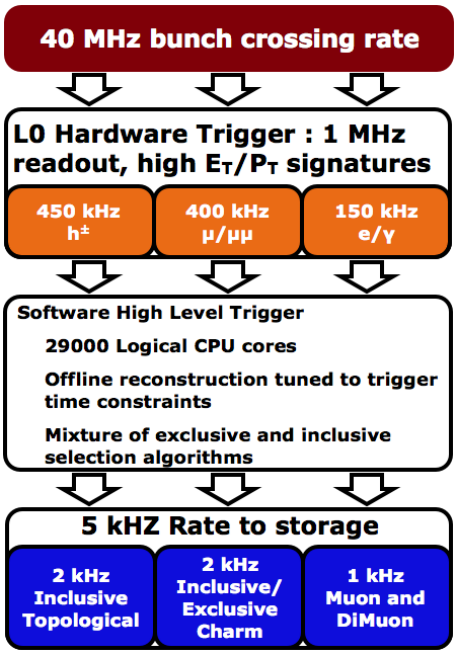


Figure 3.8: *Schematic overview of the LHCb trigger structure [17].*

LHCb detector can be read out. The L0 trigger exploits the fact that due to their high mass B mesons mostly decay into final state particles with high transverse momentum p_{T} and energy E_{T}. Therefore it searches for events with high E_{T} hadron, electron or photon clusters in the calorimeters and high p_{T} muons in the muon chambers. The L0 trigger decision unit allows for an overlap of different trigger conditions, for example it can be triggered by a hadron and a muon at the same time.
The HLT is divided into two stages, HLT1 and HLT2. It uses the full data of the event to reduce the data collection rate further to about 5 kHz. HLT1 reconstructs particles corresponding to the objects that triggered the L0 decision, using the VELO and T-stations. The purpose of this stage is to confirm the decisions made in the L0 trigger. The output of the HLT1 is sufficiently low to allow for a more advanced reconstruction, which is performed in the HLT2 stage. In the HLT2 tracks are reconstructed to meet requirements for certain composite particles, for example two muons that allow for the reconstruction of a $J/\psi \rightarrow \mu^-\mu^+$ decay. Furthermore, many inclusive decay patterns can be covered, like the decay of a B meson into a certain number of particles. Depending on which requirements an event fulfills, it is assigned to a specific trigger line, which can be accessed when analysing the data.

3.4 Flavour tagging at the LHCb experiment

To observe CP violation and mixing phenomena it is often necessary to know the flavour of a meson at the time of its creation. The method to reconstruct the flavour of a meson on

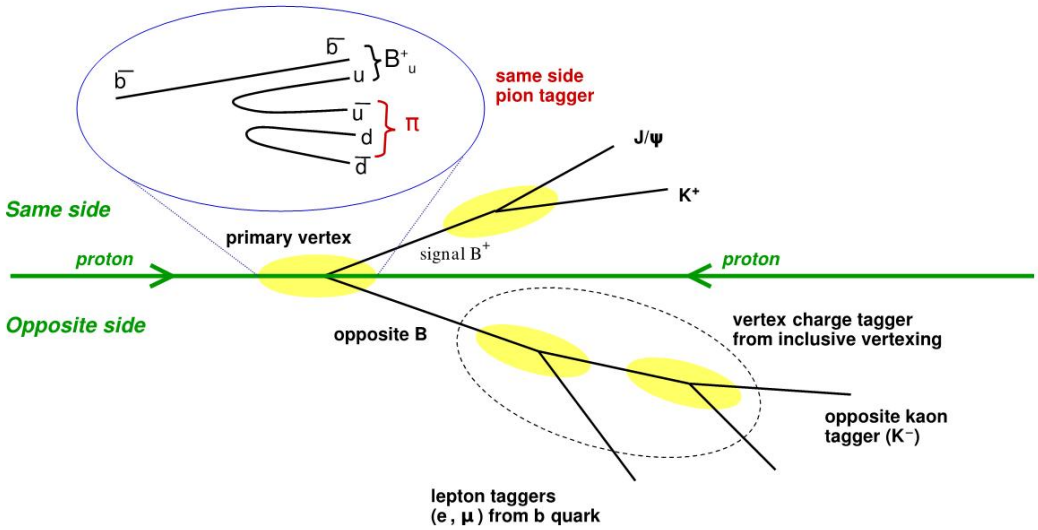


Figure 3.9: *Tagging process at the LHCb experiment [22].*

creation is called "flavour tagging". It uses several algorithms, also called "taggers".
B mesons in the LHCb detector are mostly produced in the primary interaction as $b\bar{b}$ pairs. Through hadronisation two B mesons are created, of which one is reconstructed as the signal B. The other B meson, the tagging B can be used to determine the initial flavour of the signal B. Figure 3.9 schematically shows the tagging process used at LHCb. Flavour tagging generally can be split in two categories, the same-side (SS) and opposite-side (OS) tagging algorithms. The two SS taggers (SS pion and SS kaon) use the charged pions and kaons that can be produced together with the B mesons in hadronisation. There are four OS tagging algorithms used at the LHCb experiment:

- **OS electron tagger:** this tagging algorithm uses the electron originating from the semileptonic decay of the tagging B. This algorithm assumes, that no oscillation occurred for the tagging B. The flavour of the tagging B and in that way the signal B can directly be determined from the electron charge.

- **OS muon tagger:** the muon tagging algorithm works analogously to the OS electron tagger. As muons are in general easier to detect than electrons, due to their longer lifetime and lower interaction rate with matter, this tagging algorithm is more efficient than the former.

- **OS kaon tagger:** this tagging algorithm uses the decay chain $b \to c \to s$ of the opposite B. When a charged kaon is created through this decay chain of the tagging B the flavour of the signal B can be directly determined from the charge of the kaon.

- **OS vertex charge tagger:** the vertex charge tagger uses an inclusive reconstruction of the tagging B decay vertex to determine the initial flavour of the signal B. To reconstruct this vertex, all particles are rejected that are part of the signal decay.

Tagger category	p_0	p_1	$\langle\eta\rangle$
Muon	0.309 ± 0.004	1.20 ± 0.06	0.304
Electron	0.306 ± 0.006	0.974 ± 0.09	0.346
Kaon	0.393 ± 0.002	0.706 ± 0.04	0.354
Vertex Charge	0.404 ± 0.002	0.84 ± 0.03	0.362

Table 3.1: *Calibration parameters for the opposite-side taggers using a $B^+ \to J/\psi\ K^+$ signal sample corresponding to an integrated luminosity of 1.0 fb^{-1} [22].*

All of the tagging algorithms suffer from limited efficiencies and wrong tagging decisions. These are caused by oscillations of the tagging B meson between creation and decay, wrong sign lepton tags from semileptonic D meson decays (from the decay chain $b \to c \to \ell$), wrong particle identifications and additional particles in the event. These limitations are expressed by two quantities, the tagging efficiency ϵ_{tag} and the mistag probability ω. The efficiency ϵ_{tag} is the fraction of events tagged by the respective algorithm out of all events and can be determined directly by the respective tag decisions. Measuring ω is possible in flavour-specific decays. For charged B mesons the decay is "self-tagging" as no oscillation is possible, while for neutral B mesons ω can be extracted from decay time measurements. Measurements of mixing frequencies are strongly affected by these quantities, the statistical power of this kind of analyses is directly proportional to the effective tagging power $\epsilon_{eff} = \epsilon_{tag}(1-2\omega)^2$ [22]. However, the statistical power for mixing measurements can be significantly increased by using a per-event mistag probability instead of a global parameter. This per-event mistag is calculated from neural net algorithms which exploit the dependency of the mistag probability on several geometric and kinematic parameters. The predicted mistag η is separately calculated for each event from these algorithms. As the neural net algorithms are trained on simulations and as data collection might differ between different data samples, the predicted mistag might not describe the data correctly. To account for that, the predicted mistag η is calibrated using a control sample. This calibration is done using the signal decay $B^+ \to J/\psi K^+$, for which the correct flavour is known from the kaon charge and the true mistag probability ω can thus be measured. The difference between predicted mistag and true mistag probability is expressed by the correction function

$$\omega(\eta) = p_0 + p_1(\eta - \langle\eta\rangle), \tag{3.2}$$

were $\langle\eta\rangle$ is the average η of the calibration sample and p_0 and p_1 fit parameters. These parameters are evaluated for each tagging algorithm separately by fitting a first order polynomial [22]. The results of the calibration of the OS taggers are shown in table 3.1. For a correctly calibrated sample the parameters would take the values $p_1 \approx 1$ and $p_0 \approx \langle\eta\rangle$. These calibration parameters are used to correct the predicted mistag to describe ω correctly on data. This was checked for several control channels [23]. The calibration worked correctly for all the tested control channels, i.e. the calibration parameters were in all cases $p_1 \approx 1$ and $p_0 \approx \langle\eta\rangle$.

3.5 Data samples

This analysis uses the data collected by the LHCb detector during the 2011 run. It corresponds to an integrated luminosity of $1.0\,\text{fb}^{-1}$ at a center-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$. This data sample consists of all events that successfully pass the trigger and all reconstruction and selection steps that are applied to select decays $B^0 \to D^- \mu^+ X$. The full selection applied to the data sample is explained in detail in chapter 5.
Besides the sample of real data, this analysis uses samples of simulated events to study systematic effects. 5 million events within the acceptance of the LHCb detector were generated each for the signal decay $B^0 \to D^- \mu^+ X$ and the important background decay $B^+ \to D^- \mu^+ X$. They were generated with the generator Pythia 6 [24] and the EvtGen library [25]. The detection process was further simulated with the Geant4 package [26]. The physics parameters used for the generation of these simulation samples are shown in table 3.2. Both of these simulation samples consist of a number of different decay modes, where the fractions of each individual channel were generated according to the known branching ratios. Tables 3.3 and 3.4 show the branching ratios of the most abundant decay channels used for the creation of the respective samples.

parameter	**generated value**
Δm_d	$0.507\,\text{ps}^{-1}$
Γ_{B^0}	$0.656\,\text{ps}^{-1}$
Γ_{B^+}	$0.609\,\text{ps}^{-1}$

Table 3.2: *Parameters used in the generation of the used simulation samples.*

process	**branching ratio**
$B^0 \to D^-\ \mu^+\ \nu_\mu$	2.17%
$B^0 \to D^{*-}\ (\to D^-\ \pi^0)\ \mu^+\ \nu_\mu$	1.62%
$B^0 \to D^-\ \tau^+\ \nu_\tau \to D^-\ \mu^+\ \nu_\mu\ X$	0.19%
$B^0 \to D_1^-\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.18%
$B^0 \to D_2^{*-}\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.16%
$B^0 \to D_0^{*-}\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.14%
$B^0 \to D^-\ \pi^+\ \pi^-\ \mu^+\ \nu_\mu$	0.11%
overall $B^0 \to D^-\ \mu^+\ X$	4.95%

Table 3.3: *Branching ratios used in the generation of the signal simulation sample.*

process	**branching ratio**
$B^+ \to \overline{D}_0^{*0}\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.263%
$B^+ \to \overline{D}_2^{*0}\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.230%
$B^+ \to \overline{D}_1^{0}\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.208%
$B^+ \to D^-\ \pi^-\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.090%
$B^+ \to D^{*-}\ \pi^-\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.071%
$B^+ \to \overline{D}_1^{\prime 0}\ \mu^+\ \nu_\mu \to D^-\ \mu^+\ \nu_\mu\ X$	0.069%
overall $B^+ \to D^-\ \mu^+\ X$	0.974%

Table 3.4: *Branching ratios used in the generation of the B^+ background simulation sample.*

Chapter 4

Analysis strategy

The analysis presented in this thesis is an important part of the measurement of the asymmetry in semileptonic B^0 decays a^d_{sl}. As discussed in section 2.3.2 this asymmetry is a way to probe CP violation appearing in neutral meson mixing. The decay channels used in the a^d_{sl} analysis are $B^0 \to D^{*-}\mu^+X$ and $B^0 \to D^-\mu^+X$, where the D^{*-} decays via $D^{*-} \to \overline{D}^0\pi^-$ and $\overline{D}^0 \to K^+\pi^-$ and the D^- via $D^- \to K^+\pi^-\pi^-$, respectively. The X stands for any kind of undetected particles, which is at least a muon neutrino, but might also contain further neutrinos and pions. Analogical to the general definition of semileptonic asymmetries in Formula (2.18), a^d_{sl} is defined as

$$a^d_{sl} = \frac{\Gamma(\overline{B}^0_{phys}(t) \to D^{(*)-}\mu^+\nu_\mu) - \Gamma(B^0_{phys}(t) \to D^{(*)+}\mu^-\overline{\nu}_\mu)}{\Gamma(\overline{B}^0_{phys}(t) \to D^{(*)-}\mu^+\nu_\mu) + \Gamma(B^0_{phys}(t) \to D^{(*)+}\mu^-\overline{\nu}_\mu)}. \tag{4.1}$$

The theoretical prediction of this value is $a^d_{sl} = (-0.041 \pm 0.006)\%$ [27]. The current average value of existing measurements is $a^d_{sl} = (-0.03 \pm 0.21)\%$ [28]. The uncertainty obtained so far is significantly higher than the predicted value itself. Therefore, reaching a high precision is an important goal of the asymmetry analysis.
However, to measure a^d_{sl} directly, the flavour of the B meson at production needs to be determined. This reduces the statistical power available for the analysis significantly (see section 3.4). Thus an untagged analysis is performed, which to first order reduces statistical power only by a factor of two, shown in Formula (4.3). In this untagged analysis the decay time dependent number of decays into final states containing a μ^+ ($N(f,t)$) or into the CP conjugate states ($N(\overline{f},t)$) is measured. The value that can then be observed is the untagged asymmetry

$$A_{meas}(t) = \frac{N(f,t) - N(\overline{f},t)}{N(f,t) + N(\overline{f},t)}. \tag{4.2}$$

It is connected to a^d_{sl} to first order by

$$A_{meas}(t) \approx A_D + \frac{a^d_{sl}}{2} + \left(A_P - \frac{a^d_{sl}}{2}\right)\cos(\Delta m_d t), \tag{4.3}$$

where A_D is the detection asymmetry regarding the final states and A_P is the $B^0/\overline{B}^0$ production asymmetry. The former is defined as

$$A_D = \frac{\epsilon(f) - \epsilon(\overline{f})}{\epsilon(f) + \epsilon(\overline{f})},$$

where $\epsilon(f, \overline{f})$ is the efficiency of detecting the final states f and $\overline{f}$, respectively. A_P is defined as

$$A_P = \frac{P(B^0) - P(\overline{B}^0)}{P(B^0) + P(\overline{B}^0)},$$

where $P(B^0, \overline{B}^0)$ are the probabilities to create a B^0 or $\overline{B}^0$ in the primary pp collision. A_D appears due to the fact that the detector consists of matter and therefore interacts differently with particles and antiparticles. The production asymmetry directly originates from the matter-antimatter asymmetry of the available quarks in the primary interaction, as the LHC is a proton-proton collider.
For this time-dependent analysis, the decay time t of a particle in its rest frame, often also referred to as "proper time", is needed. It can be calculated from its decay time in the lab system t_{lab} and the Lorentz boost factor γ of the particle rest frame with respect to the lab system:

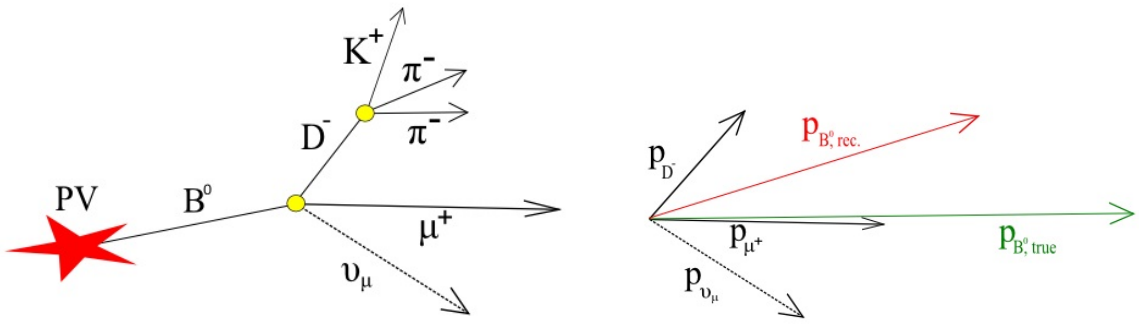
$$t = \frac{t_{lab}}{\gamma}.$$

Using the particle's flight distance in the lab system L, its momentum $p = |\vec{p}|$ and its invariant mass M, this can be written as

$$t = \frac{L \cdot M}{p}. \tag{4.4}$$

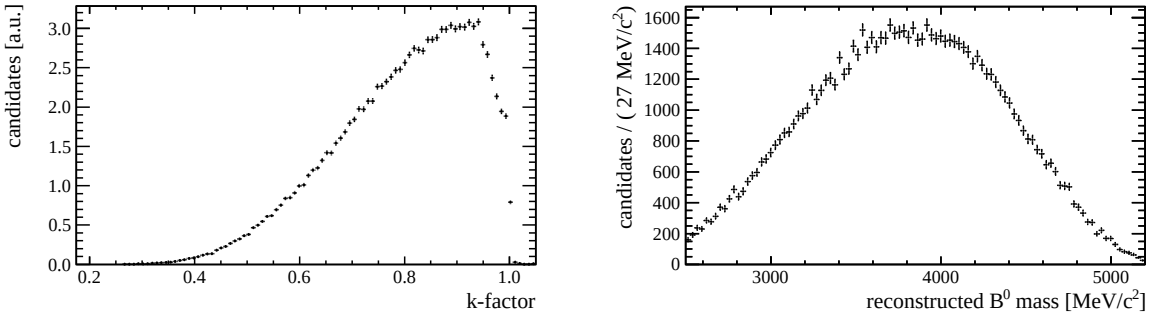
Here, the relations $t_{lab} = L/v$ and $v \cdot \gamma = p/M$ are used. Figure 4.1a schematically shows the semileptonic decay of a B^0 meson. The neutrinos originating from B^0 decays are not detectable by the LHCb detector. Thus the B^0 momentum p, which is calculated from the momenta of the final state particles, is not measured precisely, as illustrated in Fig. 4.1b. A method used by former analyses [29] to account for this is the use of a so-called k-factor. It uses simulated events to compare the reconstructed B^0 momentum to the generated true value to calculate the k-factor

$$k = \frac{p_{meas}}{p_{true}}. \tag{4.5}$$



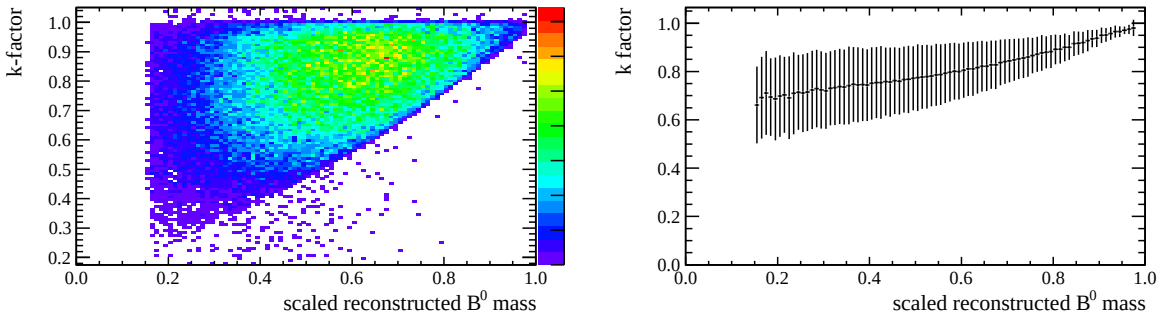
(a) (b)

Figure 4.1: *Schematical depiction of a semileptonic B^0 decay (a). The non-detection of the ν_μ leads to a reduction in the reconstructed B^0 momentum (b).*



(a) **(b)**

Figure 4.2: *k-factor distribution (a) and as a function of the scaled reconstructed B^0 mass (a)obtained from simulated $B^0 \to D^- \mu^+ X$ signal decays.*



(a) **(b)**

Figure 4.3: *k-factor distribution for the signal decay as a function of the scaled B^0 mass (a) and in bins of the scaled B^0 mass (b). Errors in latter reflect the RMS of the k-factor distribution in each bin.*

Figure 4.2a shows the distribution of the k-factor for simulated signal decays. As the invariant mass of the B^0 meson is reconstructed from the momentum four-vectors of its visible daughter particles, this measured mass is reduced with respect to the true B^0 meson mass in semileptonic decays. The distribution of the reconstructed B^0 mass is shown in Fig. 4.2b. This difference between reconstructed and nominal B^0 mass is correlated to the k-factor, as it is dependent on the reconstructed momentum. Figure 4.3a shows the k-factor distribution as a function of the scaled reconstructed B^0 mass. This scaled mass n is defined as

$$n = \frac{M_{B^0,meas} - M_{D^-,PDG} - M_{\mu^+,PDG}}{M_{B^0,PDG} - M_{D^-,PDG} - M_{\mu^+,PDG}},$$

where the nominal masses according to the PDG [30] are used: $M_{B^0,PDG} = 5279.58\,\mathrm{MeV}/c^2$, $M_{D^-,PDG} = 1869.62\,\mathrm{MeV}/c^2$ and $M_{\mu^+,PDG} = 105.66\,\mathrm{MeV}/c^2$. This leads to $n = 1$ if the reconstructed B^0 mass is equal to its nominal mass and $n = 0$ if it is equal to the sum of the invariant D^- and μ^+ masses. This spectrum can be used to determine the k-factor dependency on the scaled B^0 mass. In order to model the dependency of the k-factor on n, the k-factor distributions are determined in bins of n. The mean of each distribution is used as the value of the k-factor of the respective bin and the RMS of the k-factor distribution as the error of this value. The distribution derived that way is shown in Fig. 4.3b. This distribution is well described by a second order polynomial. Changing to fourth order polynomial does not increase the fit quality. Figure 4.4 shows both of these fits. This model can be used to correct the measured decay time with a k-factor depending on the measured scaled B^0 mass n:

$$t = \frac{L \cdot M_{B^0,PDG}}{p_{meas}} \cdot k(n) \tag{4.6}$$

However, using this k-factor correction leads to biases of the decay time, as in each of the reconstructed mass bins the k-factor distribution is asymmetric. Figure 4.5 shows the k-factor distribution for low and high reconstructed B^0 masses. Each of these distributions is asymmetric, similar to the global distribution and the mean does not account for this effect. The decay time bias introduced by the correction might lead to further issues when trying to fit for Δm_d and asymmetry properties.
To avoid possible decay time biases from this correction this analysis uses an alternative

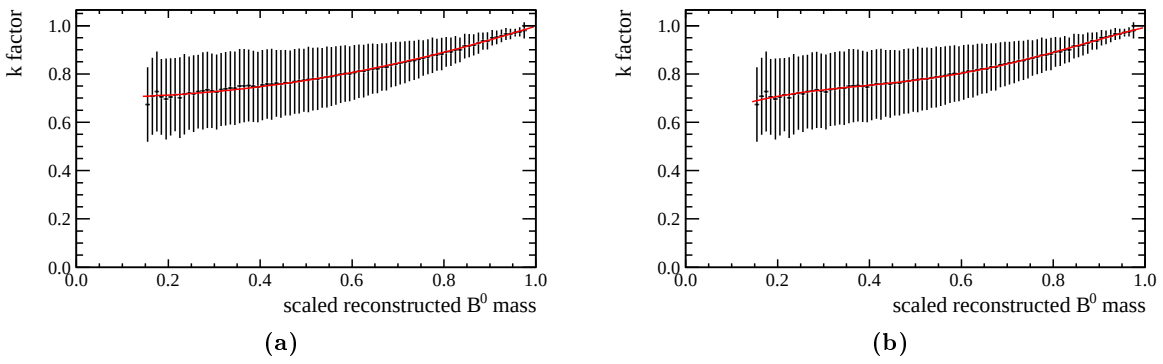
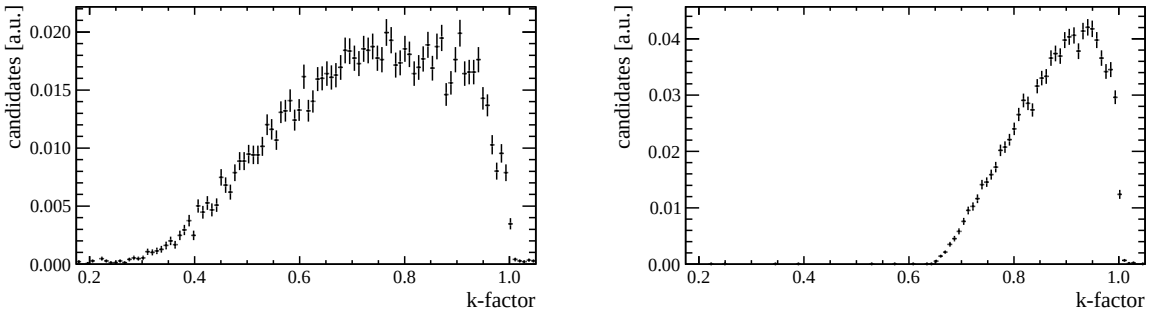


Figure 4.4: *Fits of the k-factor dependency on the scaled B^0 mass using a second (a) and fourth (b) order polynomial model*



(a) **(b)**

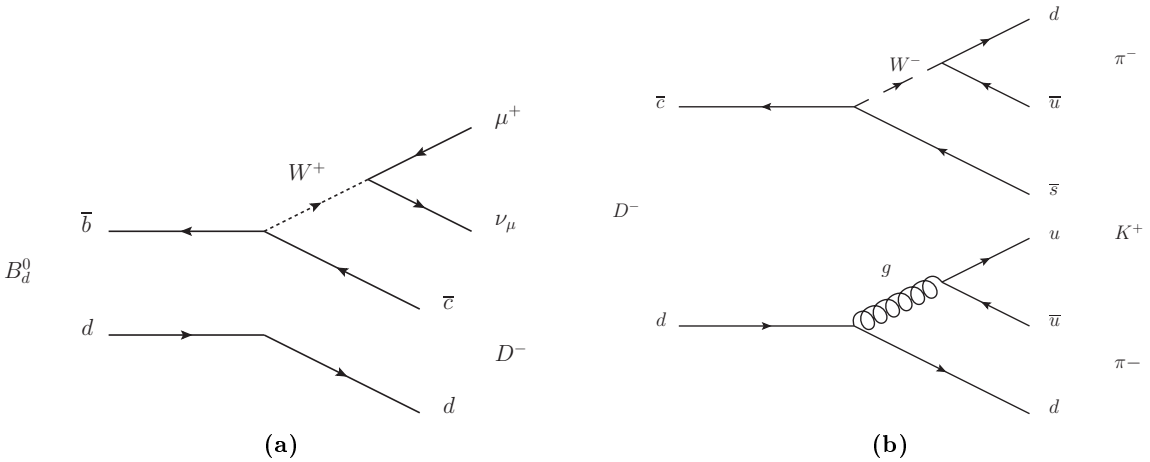
Figure 4.5: *k-factor distributions in for $n < 0.16$ (a) and $n > 0.69$ (b).*

approach to account for the missing momentum. Instead of correcting the decay time with a k-factor depending on the scaled B^0 mass n, a global k-factor distribution (without dependence on n) is used. With this distribution the probability density functions ($\mathcal{PDF}$) used in the fit for Δm_d and a_{sl}^d will be transformed to describe the distribution of the uncorrected decay time

$$t_{meas} = \frac{L \cdot M_{B^0,PDG}}{p_{meas}}. \tag{4.7}$$

This transformation is achieved by a numerical convolution of the function describing the true decay time with the k-factor distribution derived from simulated events.

The goal of the analysis presented in this thesis is to measure the B^0 mixing frequency Δm_d with this alternative method to test its validity. Furthermore, it is supposed to solidify the knowledge of Δm_d, as its value is needed to extract a_{sl}^d from the measured asymmetry A_{meas}. It is one of the first analyses of semileptonic B^0 decays in LHCb.



(a) **(b)**

Figure 4.6: *Feynman diagrams for the decays $B^0 \to D^- \mu^+ X$ (a) and $D^- \to K^+ \pi^- \pi^-$ (b).*

The decay channel used is $B^0 \to D^- \mu^+ X$, where the D^- further decays via $D^- \to K^+\pi^-\pi^-$ and its CP conjugate. This channel provides a large number of events but compared to $B^0 \to D^{*-}\mu^+ X$ it has the disadvantage that it has a large background. Figure 4.6 shows the leading order Feynman diagrams contributing to these decays. Decays where the B^0 decays to the D^- via an intermediate D^* meson are also considered signal. The mixing frequency Δm_d can be extracted from a time-dependent fit, when distinguishing between mixed and unmixed B mesons. B mesons that were created as a B^0 in the primary collision and decay as a $\overline{B}^0$ are called mixed, those that decay as a B^0 are called unmixed. The flavour of the B when decaying is immediately identified from the muon charge in semileptonic decays. Through flavour tagging (see section 3.4) one can determine the flavour of the B meson when produced. Neglecting decay time resolution and acceptance effects and assuming perfect tagging, the decay time distribution of mixed and unmixed events are described by the probability distributions

$$\begin{aligned} \mathcal{PDF}_{unmixed}(t) &\propto e^{-\Gamma t} \cdot (1 + \cos(\Delta m_d t)), \\ \mathcal{PDF}_{mixed}(t) &\propto e^{-\Gamma t} \cdot (1 - \cos(\Delta m_d t)), \end{aligned} \tag{4.8}$$

where Γ is the B^0 mean decay width $(1/\tau)$. The alternative approach to treat the missing B^0 momentum will transform the $\mathcal{PDF}$ to allow for fitting the uncorrected decay time spectrum. To be able to do this transformation, it is critical to check if the k-factor distribution is dependent on the B^0 decay time. Therefore, the mean of the k-factor distribution is plotted in bins of the B^0 decay time and fitted with a first order polynomial, shown in Fig. 4.7. The fit result is compatible with the k-factor being independent of the B^0 decay time.
Taking into account realistic conditions, decay time resolution, acceptance and tagging performance also have to be considered. This will change the fit $\mathcal{PDF}$ which will be discussed in detail in chapter 6.
There have been several measurements of Δm_d so far, the best single measurement to date was performed by the LHCb collaboration using the decays $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi K^{*0}$, resulting in a value of $\Delta m_d = (0.5156 \pm 0.0051(stat.) \pm 0.0033(syst.))\,\mathrm{ps}^{-1}$ [31].
The current world average value according to the Particle Data Group is $\Delta m_d = (0.510 \pm 0.004)\,\mathrm{ps}^{-1}$ [30].

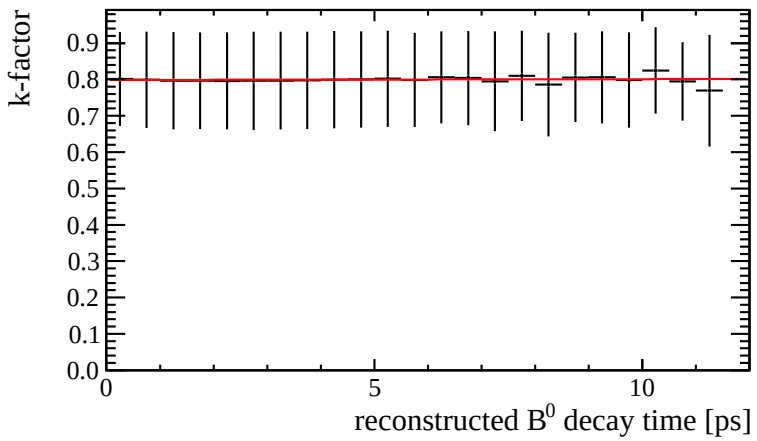


Figure 4.7: *Average k-factors in bins of the B^0 decay time, fitted with a first order polynomial.*

Chapter 5

Event selection

This chapter describes the selection criteria that are used to select signal decays and reduce the contribution from background events and furthermore reject events that might introduce biases in the asymmetry measurement. The signal decay channel used in this analysis is $B^0 \to D^- \mu^+ \nu_\mu$, where the D^- further decays via $D^- \to K^+ \pi^- \pi^-$. Decays with intermediate excited D states originating from the B^0 and decaying into the D^- are also considered as signal but the D^* mesons are not explicitly reconstructed. Figure 5.1 shows one of these decay modes. As an abbreviation, the signal decay modes will be referred to as $B^0 \to D^- \mu^+ X$ further on. The respective CP conjugated modes are also included. The final state particles detected in this decay are one μ^+, one K^+ and two π^-. As all of the particles originating from the D^- decay are detected, it can be fully reconstructed. Therefore, a fit on the invariant D^- mass spectrum is performed to evaluate the number of signal candidates. Background events peaking in the D^- mass (see sections 5.3.1 and 5.3.3) are indistinguishable from signal events when using this mass spectrum only. To account for that the term "signal candidates" and the respective yields will include those peaking background events. Using the sPlot technique [32], the outcome of the fit on the invariant D^- mass spectrum can be used to calculate event weights which allow to derive signal and background distributions in other variables. These can then be used to discriminate signal candidates from background.
To describe the reduced background levels achieved by the respective selection steps the signal to background ratio S/B will be given after each step.

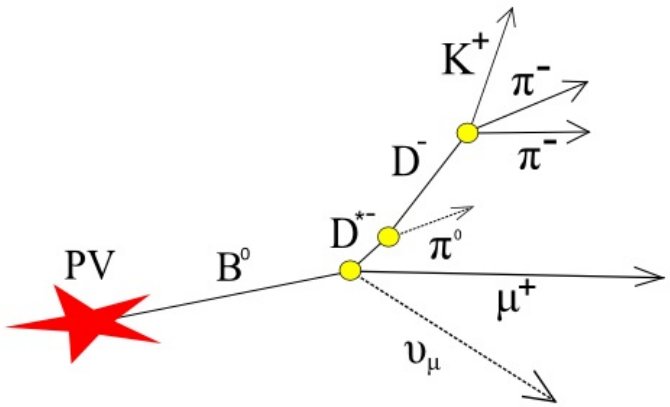


Figure 5.1: *Schematical drawing of one the decay modes that are part of the signal decay. The dashed lines represent particles missed in the reconstruction.*

5.1 Stripping selection

To reduce the data size for a given event topology under study, a so-called stripping selection is applied. In the stripping selection, decay candidates are reconstructed and cuts are applied, to significantly reduce background. The stripping line used in this analysis is called Strippingb2DpMuXB2DMuNuXLine, using the Stripping17b version. The cuts applied to the data in this stripping line are shown in table 5.1. This selection contains standard demands like minimum momentum p and transverse momentum p_{T} requirements. In the event reconstruction, tracks and vertices are fitted. From fits χ^2 values are derived, which are a measure for compatibility of this fit with the measured values. The B^0 has to originate from the primary vertex (PV). Therefore it is required to have a low impact parameter (IP, the smallest perpendicular distance between a particle's path and the PV). This requirement is fulfilled by a cut on the χ^2_{IP}, which is the increase of the vertex fit χ^2 when adding the respective particle track to the fit of the primary vertex. It is a measure for the impact parameter significance $(IP/\sigma_{IP})^2$, as it behaves similar to it. As the final state particles and the D^- are required to **not** originate from the PV, they need to have a χ^2_{IP} above a certain threshold. Furthermore, good qualities are demanded for the track reconstruction of the final state particles (χ^2_{tr}) and the fit of the decay vertices of the B^0 and D^- mesons (χ^2_{vtx}).
The DLL variable describes the particle identification and corresponds to the **D**ifference in the **L**ogarithmic **L**ikelihood between two particle hypotheses. These likelihoods $\mathcal{L}$ are the outputs of software algorithms taking information from **p**article **id**entification (PID) parts of the detector. A $\mathrm{DLL}_{K\pi} = \log \frac{\mathcal{L}_K}{\mathcal{L}_\pi} > 0$ corresponds to the likelihood for the respective particle being a kaon is higher than being a pion. Different cuts on the DLL are applied on the final state particles to ensure correct particle identification.

Kaon		**Pions**		**Muon**	
p	> 2.0 GeV/c	p	> 2.0 GeV/c	p	> 3.0 GeV/c
p_{T}	> 300 MeV/c	p_{T}	> 300 MeV/c	p_{T}	> 800 MeV/c
χ^2_{tr}/ndf	< 4	χ^2_{tr}/ndf	< 4	χ^2_{tr}/ndf	< 4
χ^2_{IP}	> 9	χ^2_{IP}	> 9	χ^2_{IP}	> 9
$\mathrm{DLL}_{K\pi}$	> 4	$\mathrm{DLL}_{\pi K}$	> -10	$\mathrm{DLL}_{\mu\pi}$	> 0

D	
$\lvert \mathrm{M}(D) - \mathrm{M}_{PDG}(D) \rvert$	< 80 MeV/c^2
$\sum p_{\mathrm{T}}$ D daughters	> 1800 MeV/c
χ^2_{DOCA} D daughters	< 20
χ^2_{vtx}/ndf	< 6
χ^2_{dist} D vertex - PV	> 100
cos α	> 0.99

B	
$\mathrm{M}(B)$	$\in [2.5,6.0]$ GeV/c^2
χ^2_{vtx}/ndf	< 6
cos α	> 0.999

Table 5.1: *Stripping 17b selections for line Strippingb2DpMuXB2DMuNuXLine*

The direction angle α used in the stripping selection is the angle between the reconstructed momentum vector of a particle and the direction of its flight path, which is the vector between its production and decay vertex. Constraining $\cos\alpha$ close to 1 corresponds to demanding a high agreement of these two directions, which one would expect for real B^0 mesons.
In semileptonic B^0 decays a neutrino is leaving the detector undetected, thus the B^0 mass cannot be fully reconstructed. This is why the constraint on the B^0 mass is loose and there is no physically motivated model to describe the shape of its distribution.
In order to calculate the number of signal candidates, a fit is performed on the invariant D^- mass spectrum. To describe this spectrum the sum of two Gaussian distributions is used to model the signal, while an exponential function is used for the background. The **p**robability **d**ensity **f**unction ($\mathcal{PDF}$) used in this fit consists of a signal and background component:

$$\mathcal{PDF}(m_{K\pi\pi}) = f_{sig} \cdot \mathcal{PDF}_{sig}(m_{K\pi\pi}) + (1 - f_{sig}) \cdot \mathcal{PDF}_{bkg}(m_{K\pi\pi}), \tag{5.1}$$

where $\mathcal{PDF}_{sig}$ and $\mathcal{PDF}_{bkg}$ are defined as:

$$\begin{aligned}\mathcal{PDF}_{sig}(m_{K\pi\pi}) =& f_m \cdot \text{Gaussian}(m_{K\pi\pi}|\mu_{sig}, \sigma_{m,1}) \\ &+ (1 - f_m) \cdot \text{Gaussian}(m_{K\pi\pi}|\mu_{sig}, \sigma_{m,2}),\end{aligned} \tag{5.2}$$

$$\mathcal{PDF}_{bkg}(m_{K\pi\pi}) = \lambda_{bkg} \frac{e^{-\lambda_{bkg}\, m_{K\pi\pi}}}{e^{-\lambda_{bkg}\, m_{min}} - e^{-\lambda_{bkg}\, m_{max}}}. \tag{5.3}$$

Figure 5.2 shows the fitted D^- mass spectrum after the stripping selection is applied. Of the 7.95 million events passing the stripping, 4.315 million events are considered signal candidates, leading to a signal to background ratio in the full invariant D^- mass range of 1.186:1.

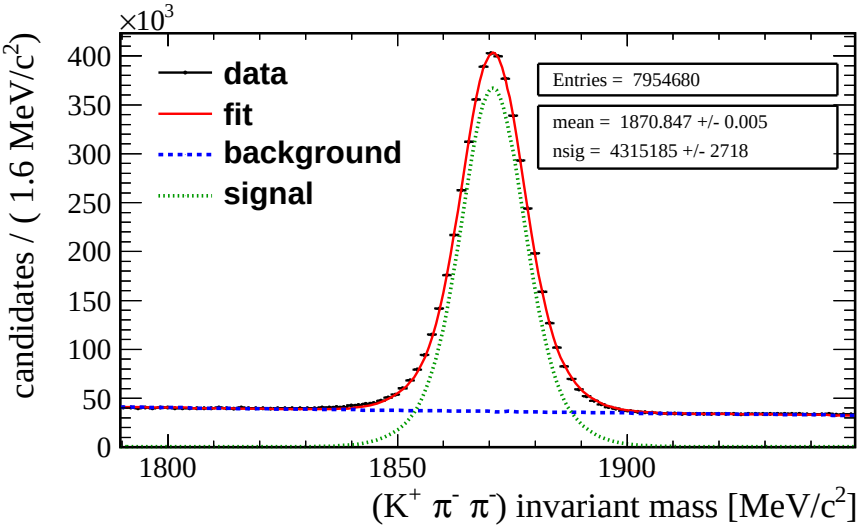


Figure 5.2: *D mass spectrum after the stripping selection. The red (solid) line shows the overall fit function, blue (dashed) the sideband background and green (dotted) the signal peak.*

5.2 Trigger line selection

At the triggering process of the LHCb detector, certain requirements on the transverse momentum p_T and energy E_T have to be met by reconstructed particles. The detection asymmetry of pions and kaons has been found to be dependent on their p_T. Therefore, these asymmetries might be biased depending on which requirements are demanded by the respective trigger decisions. Misalignments in the detector components used by the trigger can lead to additional detection asymmetries.
Depending on the trigger decision that was fulfilled by an event, the respective trigger line tag is assigned to it. These trigger line tags can be used in the offline analysis to select only events that met the requirements of specific trigger lines. The biases on asymmetries that are introduced by the individual trigger lines differ. For some of these trigger lines biases introduced on the asymmetries are known from former analyses. Therefore, only events that pass certain of trigger lines are selected in this analysis. The respective trigger lines are shown in table 5.2. Each event has to have passed at least one of the shown trigger lines per trigger stage (L0, HLT1, HLT2). A description of the LHCb trigger system can be found in section 3.3. As the trigger line selection is applied only to events that passed the stripping selection, numbers giving in the following always refer only to events that match the signal topology and do not reflect general efficiencies of these trigger lines. Most of the events passing the stripping selection (81.1%) are triggered by the L0 Muon trigger. Muons are comparably easy to trigger, due to their long lifetime and low interaction rate with matter. Adding the L0 Hadron trigger line would only add a very low amount of events but additional complications. It was therefore decided not to use the hadron lines.
While the L0 Muon trigger only demands minimum p_T values for particles passing the muon chambers, the HLT1 lines are used to confirm the L0 decisions. Close to all (97.5%) of these events pass the HLT1 confirmation. In the HLT1 tracks of the particles that triggered

Trigger line	Signal events passed	Fraction passed
after stripping	4.32M	100.0 %
L0Muon on μ	3.50M	81.1 %
L0 combined	3.50M	81.1 %
Hlt1TrackMuon on μ	2.76M	78.9 %
Hlt1TrackAllL0 on B^0	3.11M	88.9 %
HLT 1 combined	3.41M	97.5 %
Hlt2TopoMu2BodyBBDT on B^0	1.68M	49.1 %
Hlt2TopoMu3BodyBBDT on B^0	2.13M	62.5 %
Hlt2TopoMu4BodyBBDT on B^0	1.42M	41.7 %
Hlt2SingleMuon on μ	0.63M	18.4 %
HLT 2 combined	2.84M	83.3 %
combined trigger	2.84M	65.9 %

Table 5.2: *Trigger lines used in this analysis. The fractions reflect the number of signal events that are selected by the stripping selection and triggered the respective trigger line. HLT 1 (HLT 2) fractions are given compared to L0 (HLT 1) combined levels.*

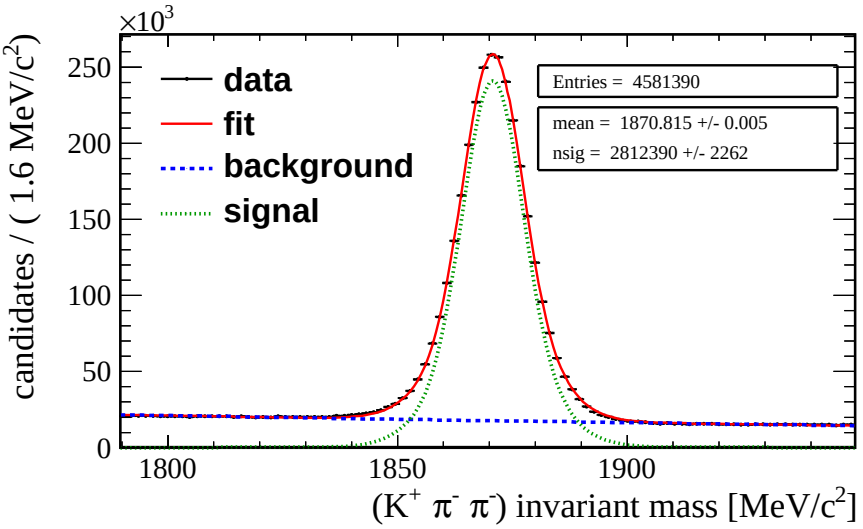


Figure 5.3: *D mass spectrum after the trigger line and TCK selection.*

the L0 decision are reconstructed and are required to pass certain p_{T} thresholds. Also demands on the primary vertex of the event have to be fulfilled. The Hlt1TrackMuon requires the reconstruction of the track of the μ triggering the L0 Muon line, while Hlt1TrackAllL0 requires all particles associated with the decay of the B^0 to trigger their respective L0 trigger. The HLT2 triggers use the output of HLT1 and are able to reconstruct detailed decays. The Hlt2TopoMuNBodyBBDT trigger lines make use of a so-called "Boosted Decision Tree" (BDT), which is a method of multivariate analyses. It triggers only on events that fulfill certain kinematical and geometric requirements associated with the decay of a B meson to N particles, of which one has to be a μ. The most effective of those Hlt2TopoMuNBodyBBDT trigger lines is the one that is supposed to trigger on decays to three bodies. The decay investigated in this analysis is exactly this type of decay. The trigger lines for topological 2- and 4-body decays are also selected, to account for non-reconstructed particles and wrongly associated tracks. The Hlt2SingleMuon trigger line requires at least one single muon reconstructed in the event. At least one of the selected HLT2 trigger lines is triggered from most of the events passing HLT1 (83.3%). By selecting only events that pass the described trigger lines, the number of events used in this analysis is reduced to 65.9%.
The requirements of the high level triggers are configured by the so-called trigger configuration keys (TCK). These keys define the p_{T} requirements and prescale factors. Further information regarding the TCKs can be found at [33]. As mentioned, detection asymmetries are dependent on the p_{T} of the respective particle. Thus different p_{T} thresholds in the trigger lines might introduce additional asymmetries for the subsamples, taken with different TCKs. Therefore only TCKs are selected that have similar p_{T} requirements and prescale rates. Using this selection reduces the number of signal events only by an additional 1.1%.
Figure 5.3 shows the fit of the invariant D^- mass spectrum after applying the trigger line and TCK selection. Out of the 4.58 million events 2.81 million are considered signal, leading to an increased signal to background ratio of 1.590:1.

5.3 Background sources

This section describes the different background categories encountered in this analysis as well as the criteria used to reduce them or, if irreducible, the way the background events are treated later on in the lifetime fit (see Chapter 6).

5.3.1 Partially reconstructed semileptonic B^+ decays

The background from partially reconstructed B^+ decays is the most prominent background in this analysis. These events are as well the ones that are most difficult to deal with. Decays contributing to this background events are $B^+ \to D^{(*)-}\mu^+\nu_\mu\pi^+$ and $B^+ \to D_0^{*0}(\to D^-\pi^+)\mu^+\nu_\mu$ (the D_0^{*0} here represents any kind of neutral excited D state). The D^- further decays via $D^- \to K^+\pi^-\pi^-$. If the π^+ is missed in the reconstruction these decays are both selected as signal. Figure 5.4 shows two examples for these decay modes. As the D^- is fully reconstructed the shape of the invariant D^- mass spectrum is the same as for the signal decay. Therefore, one can not discriminate events from this background from signal candidates by using a fit to the D^- mass spectrum. As the lifetimes of the B^0 (PDG value: $\tau_{B^0} = 1.519 \pm 0.007\,\mathrm{ps}^{-1}$) and B^+ ($\tau_{B^+} = 1.641 \pm 0.008\,\mathrm{ps}^{-1}$) are similar, discrimination using this variable is close to impossible as well. To check the efficiency of the stripping selection on signal events and the B^+ background studies were performed using the simulation samples introduced in section 3.5. For signal and B^+ background 5 million events were generated each. In case of the signal sample 271514 events pass stripping, corresponding to a stripping efficiency $\epsilon_{strip,B^0} = (5.43 \pm 0.01)\%$. For the B^+ background sample 228862 events pass stripping, resulting in a stripping efficiency $\epsilon_{strip,B^+} = (4.57 \pm 0.01)\%$, which is not significantly lower than it is for the signal sample. The combined branching ratio of all decay modes contributing to the signal decay ($\mathcal{BR}_{B^0} = (4.95 \pm 0.11)\%$) is significantly higher than for the B^+ background ($\mathcal{BR}_{B^+} = (0.974 \pm 0.054)\%$), according to the PDG values [30]. For the estimation

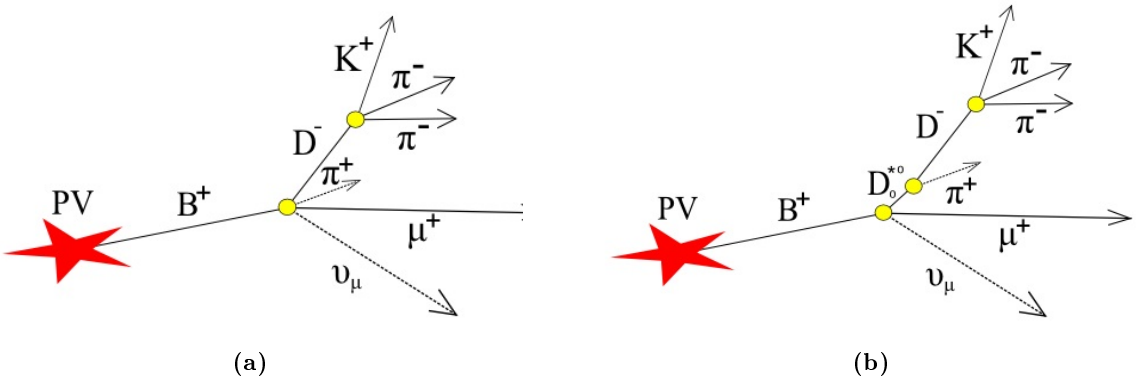


Figure 5.4: *Schematical drawings of two of the decay modes contributing to the B^+ background. Particles that are missed in the reconstruction are represented by dashed lines.*

of these combined branching ratios the decay $D^- \to K^+\pi^-\pi^-$ was ignored, as it is part of both the signal and B^+ background decay chains. From this the fraction of B^+ background events compared to the signal yield that remain in the data sample after stripping can be calculated:

$$\frac{f_{B^+}}{f_{B^0}} = \frac{\epsilon_{strip,B^+}}{\epsilon_{strip,B^0}} \cdot \frac{\mathcal{BR}_{B^+}}{\mathcal{BR}_{B^0}} = (16.59 \pm 0.99)\%, \tag{5.4}$$

where it was assumed that the amount of B^+ and B^0 mesons created in the pp collision is the same. The error is dominated by the uncertainties of the branching ratios taken from the particle data group. If all other peaking backgrounds can be neglected ($f_{B^+} + f_{B^0} = f_{peak}$), the contribution to all events peaking in the invariant D^- mass spectrum can be estimated to $\frac{f_{B^+}}{f_{peak}} = (14.23 \pm 0.73)$ %. Applying cuts might change the B^+ background fraction. It will therefore be revisited at the end of this chapter.

Due to the additional missing charged particle the most prominent difference between signal decays and this background is the k-factor distribution. As shown in fig. 5.5, the k-factor for the B^+ background is lower than for signal, as expected. This also reflects in a slightly lower reconstructed B^0 mass, shown in fig. 5.6. A good measure to describe the effectiveness of a cut is the signal significance $S/\sqrt{S+B}$. Searching for a cut value by optimising it ensures significant statistical power while reducing background. No cut based on the difference in reconstructed B^0 mass between signal and background increases $S/\sqrt{S+B}$ according to studies on simulated events. Therefore, no such cut will be applied.

As no cut was found to significantly reduce the B^+ background it will be included as a component in the decay time fit from which the fraction of B^+ background will be determined (see next chapter).

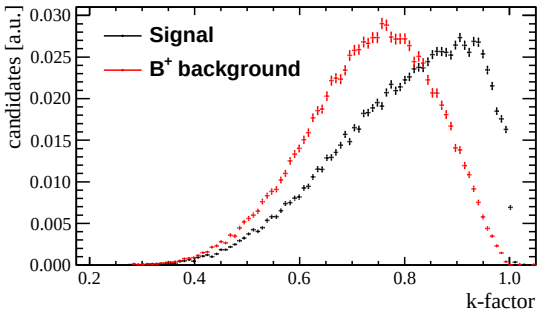


Figure 5.5: *Comparison of the k-factor distributions of signal (black) and B^+ background (red).*

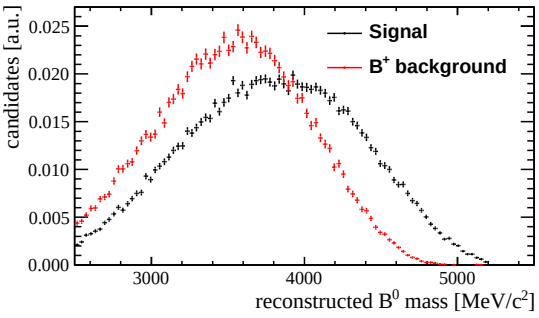


Figure 5.6: *Comparison of the reconstructed B^0 mass distributions of signal (black) and B^+ background (red).*

5.3.2 Fully reconstructed $B^0 \to D^-\pi^+$ decays

The background contribution from fully reconstructed B^0 decays is rather small, due to the PID selection criteria applied in the stripping selection. Looking at the spectrum of the reconstructed B^0 mass (fig. 5.7), one can see a small peak at the nominal B^0 mass (≈ 5280 MeV/c^2). The source of these events are $B^0 \to D^-\pi^+$ decays where the π^+ is wrongly identified as a μ^+. The reconstructed B^0 mass of semileptonic decays should in general be lower than the nominal

B^0 mass. The easiest way to remove these background events is to cut on the reconstructed B^0 mass. In order to remove all contributions from fully reconstructed B^0 decays all events with $m_{B^0,reco} > 5.2\,\mathrm{GeV}/c^2$ are rejected. This leaves the signal to background ratio nearly unchanged, it increases to 1.605:1 by applying this cut.

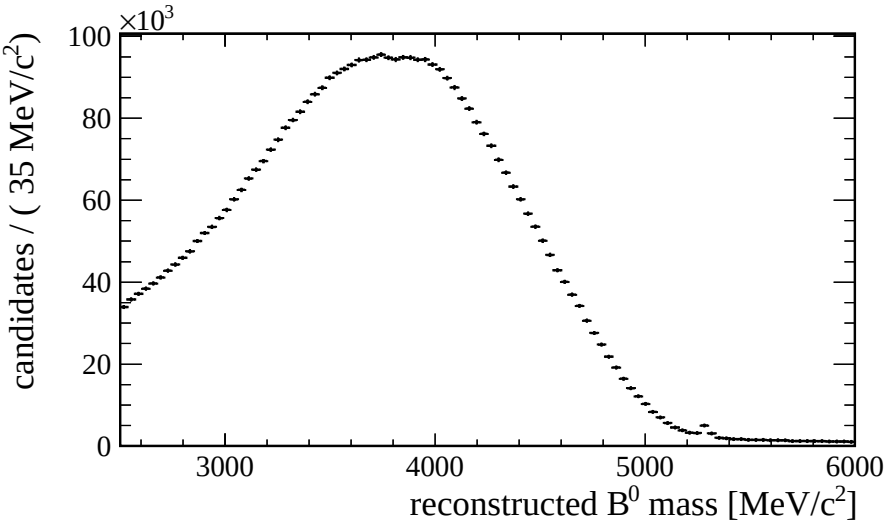


Figure 5.7: *Reconstructed B^0 mass spectrum after, a small peak of fully reconstructed B^0 decays is detected at $\approx$ 5280 MeV/c^2.*

5.3.3 Prompt $D^- \to K^+\pi^-\pi^-$ decays

Another contribution to the background peaking in the invariant D^- mass spectrum originates from D^- mesons created in the primary collision which are combined with a random μ^+. As the D^- is fully reconstructed, its shape in the D^- mass spectrum is the same as it is for the signal. Thus such events can not be discriminated via a fit of this variable from signal decays. As the prompt D^- mesons originate directly from the primary vertex (PV), they typically have a smaller impact parameter compared to D^- mesons from signal or B^+ decays, where the B decay vertex is displaced from the PV. This can be exploited by using the distribution of the logarithm of the χ^2_{IP} of the D^- meson. As the prompt D^- background is peaking in the D^- mass, the sPlot method [32] is used to derive the spectrum of $\log\chi^2_{\mathrm{IP}}$ of the D^- mesons for events with a real $D^- \to K^+\pi^-\pi^-$ decay. The distribution of these events is shown in Fig. 5.8. On first sight, there seems to be only a sharp peak for non-prompt D mesons and no distinct structure for low values of $\log\chi^2_{\mathrm{IP}}$. However, the contribution of prompt D decays to the events peaking in the invariant D^- mass is only expected to be small (1 to 2%) according to simulated events.
Besides the correct combination of a D^- with a μ^+, there is also a dataset available containing only candidates where the D meson and μ have the same charge and can therefore not be originating from a B^0 decay. As the prompt D^- mesons should mostly be randomly combined with the muons with which they are reconstructed to form a B^0 meson, one would expect this sample to have a very high contribution of prompt D^- mesons. The distribution of peaking events in $\log\chi^2_{\mathrm{IP}}$ of the D mesons is shown in fig. 5.9. As expected, one can see a clear peak

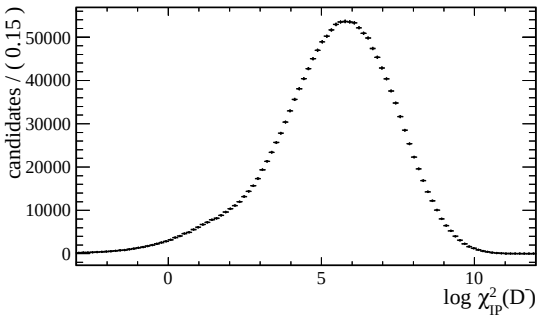


Figure 5.8: $\log \chi^2_{\mathrm{IP}}(D)$ *for events with the correct charge combination of B daughters.*

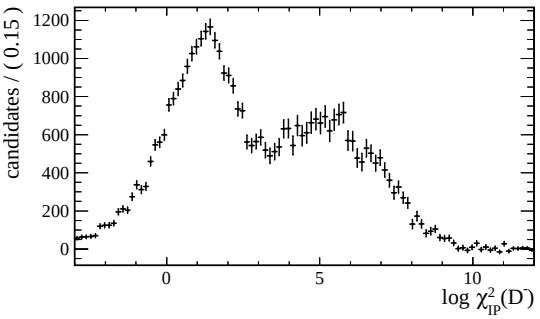


Figure 5.9: $\log \chi^2_{\mathrm{IP}}(D)$ *for events with the wrong charge combination of B daughters.*

at low $\log \chi^2_{\mathrm{IP}}$ values and as well a secondary peak of non-prompt events, as seen in the correct charge combination sample. This distribution will be used to determine the fraction of prompt D^- decays in the data sample. To model the shape, a bifurcated Gaussian is used, which is a Gaussian with different variances on each side of the mean:

$$\mathrm{BiGaussian}(x|\mu,\sigma_L,\sigma_R) = \sqrt{\frac{2}{\pi(\sigma_L+\sigma_R)^2}} \cdot \begin{cases} \exp(-\frac{(x-\mu)^2}{2\sigma_L^2}), \text{for } x < \mu \\ \exp(-\frac{(x-\mu)^2}{2\sigma_R^2}), \text{for } x \geqslant \mu \end{cases} \tag{5.5}$$

The probability distribution used to fit the whole distribution consists of a signal and prompt background component

$$\mathcal{PDF}(\log \chi^2_{\mathrm{IP}}) = f_{sig} \cdot \mathcal{PDF}_{sig}(\log \chi^2_{\mathrm{IP}}) + (1 - f_{sig}) \cdot \mathcal{PDF}_{prompt}(\log \chi^2_{\mathrm{IP}}) \tag{5.6}$$

where $\mathcal{PDF}_{sig}$ and $\mathcal{PDF}_{prompt}$ each consist of two bifurcated Gaussians:

$$\begin{aligned} \mathcal{PDF}_{sig}(\log \chi^2_{\mathrm{IP}}) =& f_{12} \cdot \mathrm{BiFurGauss}(\log \chi^2_{\mathrm{IP}}|\mu_{sig,1},\sigma_{sig,L1},\sigma_{sig,R1}) \\ &+ (1-f_{12}) \cdot \mathrm{BiFurGauss}(\log \chi^2_{\mathrm{IP}}|\mu_{sig,2},\sigma_{sig,L2},\sigma_{sig,R2}) \qquad (5.7) \\ \mathcal{PDF}_{prompt}(\log \chi^2_{\mathrm{IP}}) =& f_{34} \cdot \mathrm{BiFurGauss}(\log \chi^2_{\mathrm{IP}}|\mu_{prompt},\sigma_{prompt,L1},\sigma_{prompt,R1}) \\ &+ (1-f_{34}) \cdot \mathrm{BiFurGauss}(\log \chi^2_{\mathrm{IP}}|\mu_{prompt},\sigma_{prompt,L2},\sigma_{prompt,R2}) \qquad (5.8) \end{aligned}$$

As this fit function has many parameters, floating all of them in the same fit leads to high correlations. Therefore, the shape of the signal distribution is fixed from a fit on simulated signal events. Figure 5.10 shows the fit on these simulated events, the results are shown in table 5.3. This shape (all parameters of $\mathcal{PDF}_{sig}$) is used to fit the $\log \chi^2_{\mathrm{IP}}$ distribution of data events peaking in the invariant D^- mass, leaving all parameters of $\mathcal{PDF}_{prompt}$ and f_{sig} floating. Figure 5.11 shows the result of this fit, while the parameters are given in Tab. 5.4. The fraction of prompt D mesons in the invariant D^- mass peak is derived to be $(3.32 \pm 0.06)\%$. The uncertainty on this fraction is the statistical uncertainty as an output of the fit algorithm. It does not reflect the fact that the non-prompt shape is completely fixed, which leads to uncertainties caused by possible disagreements between data and simulation. As a crosscheck whether the shape fixed from simulation models the non-prompt D^- events well and how much this affects the prompt fraction, the fit is repeated with all parameters

Fit parameter	value
f_{12}	0.570 ± 0.073
$\mu_{sig,1}$	5.661 ± 0.068
$\mu_{sig,2}$	7.102 ± 0.118
$\sigma_{sig,L1}$	1.705 ± 0.064
$\sigma_{sig,L2}$	2.966 ± 0.106
$\sigma_{sig,R1}$	1.352 ± 0.072
$\sigma_{sig,R2}$	1.154 ± 0.036

Table 5.3: *Fit results for the* $\log \chi^2_{\mathrm{IP}}$ *fit on simulated signal events.*

Fit parameter	value
f_{sig}	0.9668 ± 0.0006
f_{34}	0.450 ± 0.029
μ_{prompt}	1.334 ± 0.031
$\sigma_{prompt,L1}$	2.470 ± 0.058
$\sigma_{prompt,L2}$	1.306 ± 0.036
$\sigma_{prompt,R1}$	2.230 ± 0.164
$\sigma_{prompt,R2}$	0.560 ± 0.029

Table 5.4: *Fit results for the* $\log \chi^2_{\mathrm{IP}}$ *fit on on data with the signal shape fixed from simulation.*

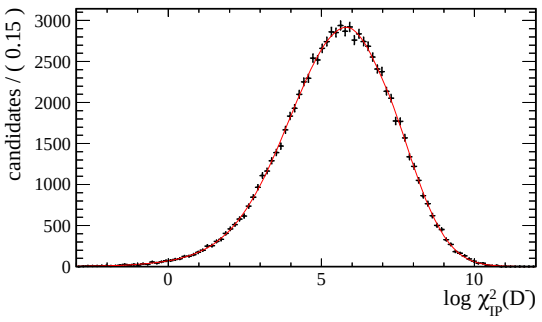


Figure 5.10: *Fitted signal Monte Carlo distribution of* $\log \chi^2_{\mathrm{IP}}(D)$.

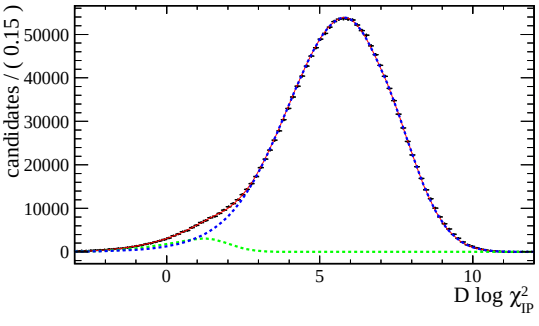


Figure 5.11: *Fitted distribution of* $\log \chi^2_{\mathrm{IP}}(D)$ *for events with the correct charge combination of* B *daughters.*

floating. To ensure convergence, the means and variances of of the non-prompt shape are constrained within a ±20% window around the value derived from simulation. The result of this fit is shown in Fig. 5.12 and the parameters in Tab. 5.5. In that case, the fraction of prompt D^- events was calculated to $(4.47 \pm 0.06)\%$, were the difference originates in the smaller tail of the non-prompt distribution towards lower $\log \chi^2_{\mathrm{IP}}$ values. However, the fit results in a non positive-definite error matrix, indicating problems which might be a result of the highly correlated parameters. Fixing the shape for non-prompt D^- decays from simulated events (Fig. 5.10) does not result in these problems, therefore this result will be used. The

Fit parameter	value
f_{sig}	0.9553 ± 0.0006
f_{12}	0.703 ± 0.032
$\mu_{sig,1}$	5.720 ± 0.005
$\mu_{sig,2}$	7.332 ± 0.009
$\sigma_{sig,L1}$	1.896 ± 0.006
$\sigma_{sig,L2}$	2.428 ± 0.018
$\sigma_{sig,R1}$	1.353 ± 0.005
$\sigma_{sig,R2}$	1.098 ± 0.005
f_{34}	0.758 ± 0.687
μ_{prompt}	1.412 ± 0.026
$\sigma_{prompt,L1}$	1.618 ± 0.006
$\sigma_{prompt,L2}$	1.618 ± 0.061
$\sigma_{prompt,R1}$	0.753 ± 0.009
$\sigma_{prompt,R2}$	0.761 ± 0.181

Table 5.5: *Fit results for the* $\log \chi^2_{\mathrm{IP}}$ *fit on on data with all parameters floating.*

reconstructed B^0 lifetime of this background is rather hard to model, mainly because of the lack of a sufficiently large simulated data sample. So for this analysis it was decided to reduce the prompt D background to a level were it can be neglected in the B^0 lifetime fit.
There are two obvious variables which would help to reduce the number of prompt D^- mesons: the reconstructed B lifetime and the $\log \chi^2_{\mathrm{IP}}$ of the D meson. A reconstructed B lifetime cut

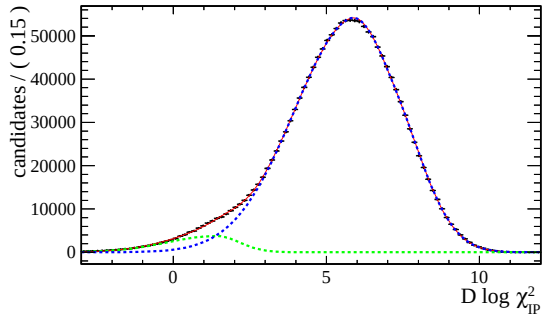


Figure 5.12: *Fitted distribution of* $\log \chi^2_{\mathrm{IP}}(D)$ *with all fit parameters floating.*

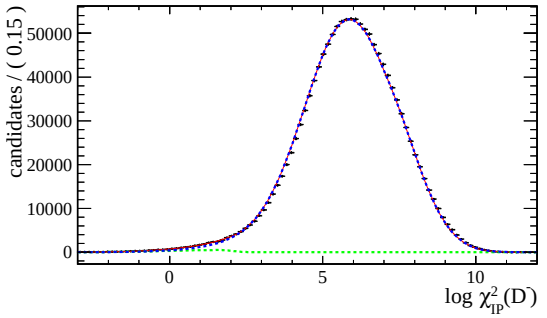


Figure 5.13: *Fitted distribution of* $\log \chi^2_{\mathrm{IP}}(D)$ *with a* B *decay time cut at 0.8* ps.

at 0.8 ps leads to a prompt D fraction of $(0.629 \pm 0.008)\%$ (see fig. 5.13), while reducing the signal yield by 15.1% according to the simulated signal sample.
To achieve a reduction of the prompt D fraction to 0.5% with a cut on $\log \chi^2_{\mathrm{IP}}(D)$, one has to cut at a value of 2.1. This reduces the signal only by 4.05%. This cut is therefore used in this analysis. With this cut applied, 4.312 million events pass the selection of which 2.622 million are considered signal. This is a slight reduction of the signal to background ratio to $S/B = 1.551 : 1$. This originates from the prompt D mesons peaking in the invariant D^- mass spectrum and thus contributing to the "signal" yield when derived from the mass fit.

5.3.4 $\Lambda_b^0 \to \Lambda_c^- \mu^+ \nu_\mu$ decays

A significant fraction (about 10% [34]) of the b quarks produced in the LHCb detector hadronize to a Λ_b^0 baryon. About 10% of the Λ_b^0 baryons decay semileptonically via $\Lambda_b^0 \to \Lambda_c^- \mu^+ \nu_\mu$ where about 5% of the produced Λ_c^+ baryons decay further via $\Lambda_c^- \to \overline{p} K^+ \pi^-$. For events where the $\overline{p}$ is misidentified as a π^-, the reconstructed final state is exactly the same as the signal decay. As no real D^- is reconstructed, this background is not peaking in the D^- mass spectrum.
To check the size of this background, the proton mass hypothesis was checked for each of the reconstructed π^-. To do this, the energy of the energy-momentum-vector of the respective π^- was increased via $E' = \sqrt{E^2 - m_\pi^2 + m_p^2}$ and afterwards the energy-momentum-vector of the $K^+ \pi^- \pi^-$ system reconstructed. This way the invariant mass (corresponding to the absolute value of the energy-momentum-vector) of this system is shifted to higher values. Figure 5.14 shows the new $K^+ \overline{p} \pi^-$ invariant mass distribution. There is a prominent peak at the Λ_c^+ mass (≈ 2286 MeV/c^2) originating from $\Lambda_c^- \to \overline{p} K^+ \pi^-$ decays. The easiest way to remove this background would be to reject all events in a mass window around this peak in the $K^+ \overline{p} \pi^-$ spectrum. Evaluating the effect on signal and background with and without this cut applied shows that it removes more signal than background. Therefore, this mass window cut

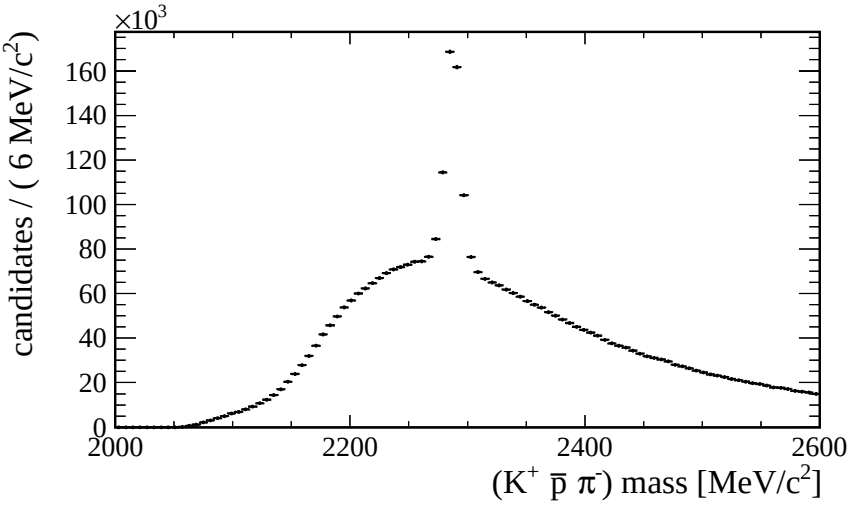


Figure 5.14: *$K^+ \pi^- \pi^-$ mass spectrum after applying the p mass hypothesis on one π^-.*

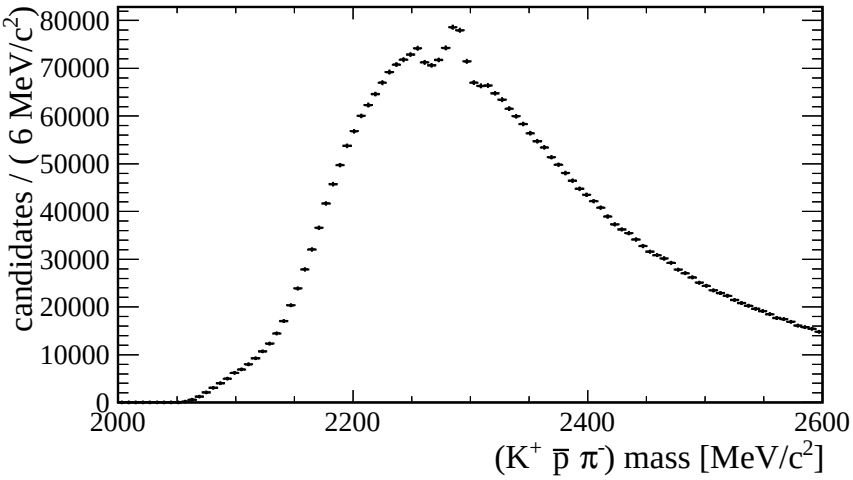


Figure 5.15: *$K^+\,\pi^-\,\pi^-$ mass spectrum after applying the Λ_c^- veto as described in the text.*

is combined with information of the pion $\mathrm{DLL}_{p\pi}$: all events are rejected, which have a shifted invariant mass inside a window around the Λ_c^+ peak ($\mathrm{m}_{K^+\overline{p}\pi^-} \in [2260{,}2310]$ MeV/c^2) and a π^- with a high likelihood of being a proton ($\mathrm{DLL}_{p\pi} > 10$). Using this combined veto removes most of the background along with some signal events, as can be seen in fig. 5.15. A small peak is still left, but it is small enough to neglect this background further in the analysis. After this cut the signal to background ratio to $S/B = 1.845 : 1$.

5.3.5 $B^0 \to J/\psi(\to \mu^+\mu^-)X$ decays

Another background related to improper particle identification is resulting from decays $B^0 \to J/\psi\, X$, where the J/ψ further decays via $J/\psi \to \mu^+\,\mu^-$. In this case the μ^- from the J/ψ decay can be misidentified as a π^-. As the K^{*0} almost exclusively decays to a $K^+\,\pi^-$ pair, the final state looks exactly like the signal decay. Checking the contribution from this background was done in a similar manner as for the Λ_b^0 background. The muon mass was given to one of the reconstructed π^-, but in that case only the invariant mass spectrum of the $\pi^-\,\mu^+$ system was investigated, to check for a peak at the J/ψ mass (≈ 3097 MeV/c^2). As shown in fig. 5.16 such a peak is observed, being pretty sharp. Figure 5.17 shows a zoomed in version of this spectrum, where one can see that the peak is clearly defined. To remove contributions from this background a approach similar to the one used for the Λ_b^0 background is chosen. A veto is applied for events inside a mass window around the J/ψ mass ($\mathrm{m}_{\pi^-\,\mu\mathrm{m}} \in [3070{,}3150]$ MeV/c^2) combined with the π^- having the ISMUON tag. This ISMUON tag is assigned to particles that are passing certain muon stations depending on their momentum. As a final state π^- of the signal decay might decay in flight via $\pi^- \to \mu^-\,\overline{\nu}_\mu$ it can as well have this tag. The result of applying this veto is shown in fig. 5.18, where the $m_{\mu^+\pi^-}$ distribution now shows no peak anymore. After this cut background originating from $B^0 \to J/\psi\, K^{*0}$ can be neglected further on. With this cut applied the signal to background ratio increases to $S/B = 1.937 : 1$.

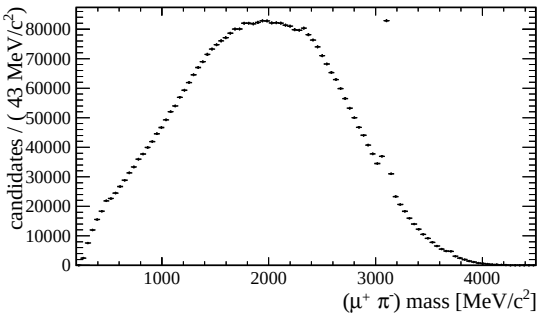


Figure 5.16: π^- μ^+ *mass spectrum after applying the* μ *mass hypothesis on one* π^-.

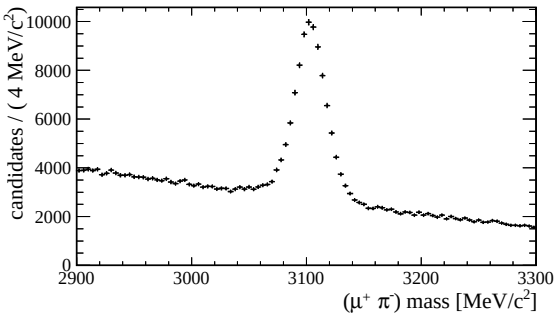


Figure 5.17: π^- μ^+ *mass spectrum after applying the* μ *mass hypothesis on one* π^-, *zoomed in.*

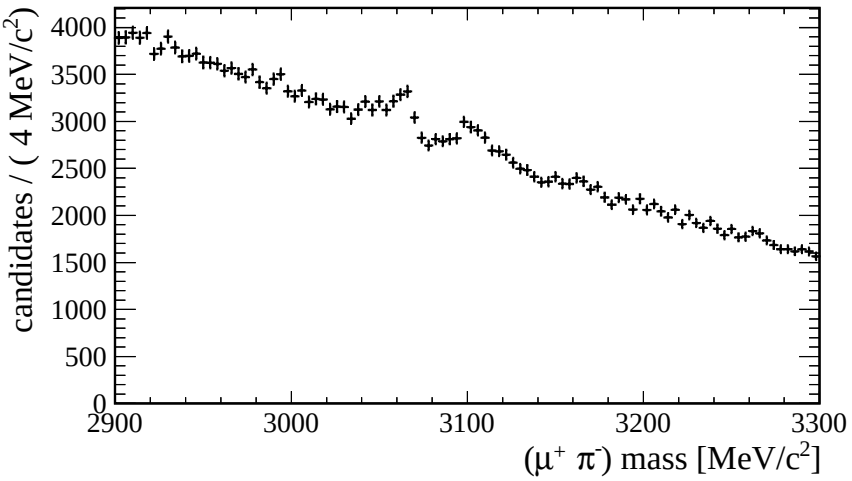


Figure 5.18: π^- μ^+ *mass spectrum after applying the* J/ψ *veto.*

5.3.6 Sideband background

Besides the mentioned physical backgrounds there is as well background left which is flat in the invariant D^- mass sidebands. This background is assumed to be combinatorial, composed of random combinations of final state particles, that are not associated to a specific decay. However, this background has a significant component with an oscillating behaviour in the candidate B^0 reconstructed decay time (see chapter 6). It will be referred to as sideband background. This background does not peak in the D mass, or any of the shifted mass spectra discussed before. An exponential function describes its behaviour in the invariant D^- mass spectrum sufficiently well. The sPlot technique is used to separate the background on a statistical basis from signal decays. It is therefore possible to search for variables discriminating these sideband backgrounds. Selection criteria found in this way and the effects on the sideband background level will be described in the additional analysis selection in section 5.4.

5.4 Additional analysis selection

In addition to the selection described so far (stripping, trigger line and TCK selection and cuts to reduce physical backgrounds) there are additional cuts applied in the analysis selection.
In the a_{sl}^d analysis a D calibration sample will be used to fix parts of the detection asymmetry. To use this, certain values have to cancel in the comparison of data and the D calibration sample. These values depend on kinematics, which is why the cuts on $\mathrm{DLL}_{K\pi}$ and p_T of the K^+ from stripping are tightened further to the levels used in the calibration sample: the kaon $\mathrm{DLL}_{K\pi}$ is required to be > 7 (>4 in stripping) and its $p_\mathrm{T} > 500$ MeV/c (> 300 MeV/c).

The only additional variable which was found to increase the signal significance $S/\sqrt{S+B}$ was the reconstructed decay time of the D^- meson:

$$\tau_{D^-} = \frac{L_{D^-} \cdot M_{PDG,D^-}}{p_{D^-}},$$

where the flight distance L_{D^-} is calculated as the distance between the reconstructed decay vertices of the B^0 and D^- mesons. M_{PDG,D^-} is the PDG value of the D^- mass of 1869.62 MeV/c^2 and p_{D^-} is the absolute value of the measured momentum of the D^- meson. Figure 5.19 shows the distribution of τ_{D^-} separated into signal candidates and combinatorial background via sPlot.
As shown in this figure, the background peaks at low reconstructed D^- lifetimes. The signal significance $S/\sqrt{S+B}$ is increased the most by cutting at $\tau_{D^-} < 0.1$ ps.
This concludes all the selection criteria used on the data sample for this analysis. After applying these cuts, 2.322 million signal candidates out of 3.060 million events are left, leading to a signal significance $S/\sqrt{S+B}$ of 1328 and a S/B ratio of 3.150 : 1. The reduced signal significance compared to the status after trigger line and TCK selection (1383) originates from

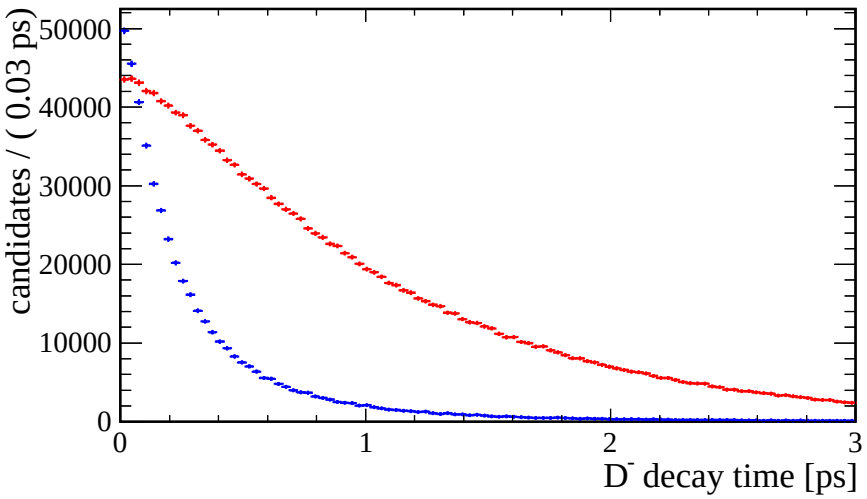


Figure 5.19: *Reconstructed D^- lifetime distributions for signal candidates (red) and combinatorial background (blue), separated via sPlot.*

the fact, that some of the backgrounds that were reduced peak in the invariant D^- mass, which is why they are considered signal in $S/\sqrt{S+B}$. However, the signal to background ratio has significantly improved compared to the beginning of this chapter, leading to a very clean signal, as can be seen in fig. 5.20. To check for the fraction of B^+ background (see section 5.3.1) in the events peaking in the invariant D^- mass, once again simulated events are used. Out of 5M events generated for each signal and B^+ background 84812 (B^0) and 62689 (B^+) pass all selection criteria. This leads to efficiencies of $\epsilon_{B^0,sel} = (1.70 \pm 0.01)\%$ and $\epsilon_{B^+,sel} = (1.25 \pm 0.01)\%$, respectively. Using formula 5.4, this leads to an estimated B^+ fraction in the signal peak of

$$\frac{f_{B^+}}{f_{peak}} = (12.70 \pm 0.66)\%. \tag{5.9}$$

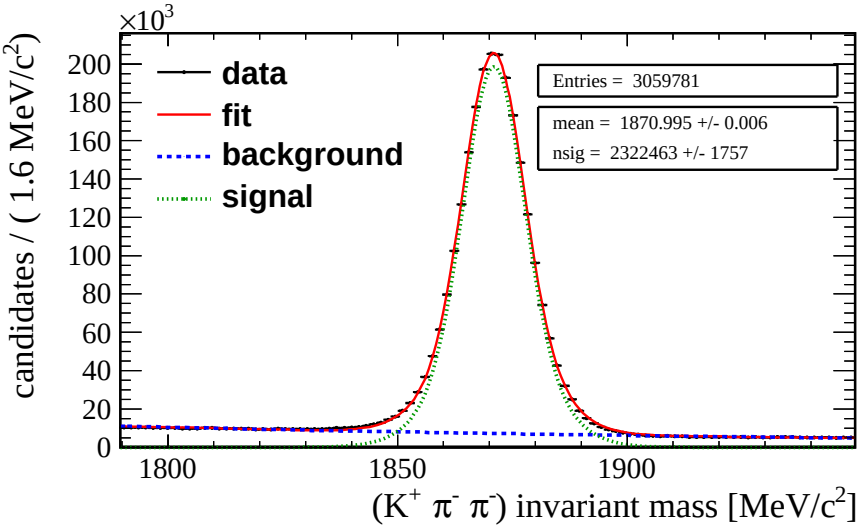


Figure 5.20: *Fitted D mass spectrum with the full selection cuts applied.*

Chapter 6

Fitting procedure

This chapter describes the probability density functions ($\mathcal{PDF}$) used to fit for Δm_d as well as the parameters extracted from this fit. The fit is performed three-dimensional in the variables uncorrected B^0 decay time t_{meas}, mixing decision q and invariant D^- mass $m_{K\pi\pi}$. The $B^0-\overline{B}^0$ mixing frequency Δm_d is extracted from the tagged decay time fit, while the invariant D^- mass spectrum is used to discriminate between signal events and sideband background.
The $\mathcal{PDF}$ used for the fit in the invariant D^- mass spectrum is identical to the one used in the optimisation of the selection, which was described previously in chapter 5. It consists of one component for the peaking events $\mathcal{PDF}_{peak}$ and one for the sideband background $\mathcal{PDF}_{SB}$. The peak is described by a the sum of two Gaussian distributions, while the latter is modeled by an exponential function:

$$\begin{aligned}
\mathcal{PDF}(m_{K\pi\pi}) &= (1-f_{SB})\cdot\mathcal{PDF}_{peak}(m_{K\pi\pi}) + f_{SB}\cdot\mathcal{PDF}_{SB}(m_{K\pi\pi}),\\
\mathcal{PDF}_{peak}(m_{K\pi\pi}) &= f_m\cdot\text{Gaussian}(m_{K\pi\pi}|\mu_{peak},\sigma_{m,1})\\
&\quad + (1-f_m)\cdot\text{Gaussian}(m_{K\pi\pi}|\mu_{peak},\sigma_{m,2}),\\
\mathcal{PDF}_{SB}(m_{K\pi\pi}) &= \lambda_{SB}\frac{e^{-\lambda_{SB}\,m_{K\pi\pi}}}{e^{-\lambda_{SB}\,m_{min}}-e^{-\lambda_{SB}\,m_{max}}}.
\end{aligned} \tag{6.1}$$

The parameter f_{SB} denotes the fraction of sideband background events.
The decay time $\mathcal{PDF}$ consists of three components. One to describe the D^- mass sideband background and two to describe events peaking in the D^- mass. These two components are describing signal candidates and background events originating from B^+ decays, as they were described in section 5.3.1. Each of those are described by a different decay time $\mathcal{PDF}$.
The tagged decay time $\mathcal{PDF}$ is described by

$$\mathcal{PDF}(t,q) = \frac{\Gamma_d}{2}e^{-\Gamma_d t}\cdot(1-q\,\cos(\Delta m_d\,t)). \tag{6.2}$$

The mixing decision q is equal to $+1$ for B^0 mesons that have oscillated to a $\overline{B}^0$ before decaying and equal to -1 for those that decayed as a B^0. If the flavour of the B^0 meson at production could not be tagged, the mixing decision is equal to 0. The non-oscillating B^+ background is not mixing and is described by

$$\mathcal{PDF}_{B^+}(t) = \Gamma_u e^{-\Gamma_u t}. \tag{6.3}$$

Γ_d and Γ_u are the mean decay widths of the B^0 and B^+ mesons, respectively.
Due to the method chosen to correct for the missing B^0 momentum, in this analysis the decay time $\mathcal{PDF}$s have to be transformed to the space of the uncorrected decay time. The PDFs describing the sideband background are introduced directly as functions of the uncorrected

decay time, as there is no way to describe the relation of its reconstructed and true decay time. It was found to have an oscillating component, so the sideband background is not purely combinatorial but also consists of other B^0 decays. The decay time distribution of D^- mass sideband events is therefore described by the following $\mathcal{PDF}$

$$\mathcal{PDF}_{SB}(t_{meas}, q) = \frac{\Gamma_{SB}}{2}\, e^{-\Gamma_{SB}\, t_{meas}} \cdot (1 - q\, \cos(\Delta m_{SB}\, t_{meas})). \tag{6.4}$$

The sideband mixing frequency extracted from the lower D^- mass sideband has the value $\Delta m_{SB} = (0.2607 \pm 0.0148)\,\mathrm{ps}^{-1}$. The data is incompatible with a non-oscillating model. The frequency Δm_{SB} will be left floating in the final fit along with the mean decay width Γ_{SB}. The name Δm_{SB} is chosen for consistency reasons. The frequency of the sideband oscillation is not associated with any mass difference.
Realistic decay time resolution, acceptance and tagging performances influence the $\mathcal{PDF}$ as well. The following sections will describe those changes and present the final decay time $\mathcal{PDF}$. Furthermore, a validation of the fitting technique via studies using simulated pseudo-experiments will be shown.

6.1 Unbinned maximum-likelihood fit

The simultaneous fit of the reconstructed B^0 decay time and invariant D^- mass is performed unbinned, therefore each signal candidate is treated individually. This unbinned fit applies a maximum-likelihood-technique. This fitting technique uses an i-dimensional probability density function $\mathcal{PDF}(\vec{x}|\vec{a})$. This $\mathcal{PDF}$ depends on a set of n parameters $\vec{a} = (a_1, ..., a_n)$ and describes a number of i measured values $\vec{x} = (x_1, ..., x_i)$ for each event. The $\mathcal{PDF}$ has to be normalised such that

$$\int_{\Omega} \mathcal{PDF}(\vec{x}|\vec{a}) d\vec{x} = 1 \quad \forall \vec{a}, \tag{6.5}$$

where Ω is the possible range of $\vec{x}$. With a set of N measured events $\vec{x}_1, ..., \vec{x}_N$ one can define the likelihood function

$$\mathcal{L}(\vec{a}) = \mathcal{PDF}(\vec{x}_1|\vec{a}) \cdot ... \cdot \mathcal{PDF}(\vec{x}_N|\vec{a}) = \prod_{j=1}^{N} \mathcal{PDF}(\vec{x}_j|\vec{a}). \tag{6.6}$$

This likelihood is a function of the parameters $\vec{a}$ and describes the probability for this set of N events to yield the measured values. According to the maximum-likelihood-principle, the best estimate of the true values of the parameters $\vec{a}$ is the value $\hat{\vec{a}}$ for which $\mathcal{L}(\vec{a})$ is maximal. Due to computational reasons, in practice usually the negative log-likelihood-function is used:

$$\mathcal{F}(\vec{a}) = -ln\mathcal{L}(\vec{a}) = -\sum_{j=1}^{N} ln\mathcal{PDF}(\vec{x}_j|\vec{a}). \tag{6.7}$$

As the logarithm is a strictly monotonic function, the maximum of $\mathcal{L}(\vec{a})$ corresponds to the minimum of $\mathcal{F}(\vec{a})$. Minimising $\mathcal{F}$ gives the constraint

$$\frac{\partial \mathcal{F}(\vec{a})}{\partial a_k} = 0, \tag{6.8}$$

for every parameter a_k $(k = 1, ..., n)$.

The minimisation in form of a fit is performed within a C++ framework developed by the LHCb group of the Physikalisches Institut Heidelberg [35] using the minimisation procedure Minuit [36].
In this analysis, three values are measured for each event: the mixing decision q, the reconstructed B^0 decay time t_{meas} and invariant D^- mass $m_{K\pi\pi}$. Therefore, the used $\mathcal{PDF}$ is 3-dimensional. The mixing decision $q = 1, -1$ or 0 corresponding to the three discrete possible states ("mixed", "unmixed" and "untagged") will be directly included in a common $\mathcal{PDF}(t_{meas}, q)$ with the B^0 decay time. The following chapters will introduce these $\mathcal{PDF}$s and their parameters, which will be extracted from the final fit.

6.2 Effects of flavour tagging

To distinguish between mixed and unmixed B^0 decays, flavour tagging algorithms are used (see section 3.4). Only events that have a tagging decision are contributing to the statistical power of the Δm_d analysis. Furthermore, the probability to assign a wrong flavour tag, the mistag probability ω, leads to a decrease in statistical power. The effective tagging power $\epsilon_{eff} = \epsilon_{tag}(1 - 2\,\omega)^2$ is combining these two effects. It can be significantly increased by using the per-event mistag prediction η, which is calculated from several geometric and kinematic quantities. The predicted mistag η is connected to the mistag probability ω via a linear dependency derived from an external calibration in control channels. This dependency is described by

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

where p_0 and p_1 are the calibration parameters and $\langle\eta\rangle$ is the average η of the calibration sample.

The B^+ background discussed in section 5.3.1 cannot be discriminated from signal decays via a fit to the B^0 lifetime or to the invariant D^- mass. However, this background does not oscillate, so all contributions to candidates that are tagged as "mixed" originate from mistags. Therefore, with the knowledge of the tagging parameters, this can be used to fit for the fraction of B^+ background events peaking in the D^- mass.
Studies of simulated signal and the B^+ background events for the investigated decay showed large differences in the tagging performance. Table 6.1 shows tagging efficiency, mistag prob-

	signal			B^+ **background**		
Tagging algorithm	ϵ_{tag}[%]	ω[%]	ϵ_{eff}[%]	ϵ_{tag}[%]	ω[%]	ϵ_{eff}[%]
Muon	5.34 ± 0.06	30.3 ± 0.6	0.83 ± 0.05	5.61 ± 0.08	31.4 ± 0.7	0.78 ± 0.06
Electron	2.99 ± 0.05	34.7 ± 0.8	0.28 ± 0.03	2.80 ± 0.06	36.0 ± 1.0	0.22 ± 0.03
Kaon	17.22 ± 0.11	35.3 ± 0.3	1.50 ± 0.07	18.44 ± 0.13	38.7 ± 0.4	0.94 ± 0.06
Vertex charge	19.50 ± 0.11	36.5 ± 0.3	1.43 ± 0.07	23.37 ± 0.14	49.5 ± 0.4	0.00 ± 0.00
OS combined	33.88 ± 0.14	34.6 ± 0.2	3.20 ± 0.10	38.04 ± 0.17	39.8 ± 0.3	1.57 ± 0.08
SS pion	17.81 ± 0.11	37.5 ± 0.3	1.11 ± 0.06	24.50 ± 0.11	43.5 ± 0.4	0.42 ± 0.04

Table 6.1: *Signal and B^+ background tagging performances [37].*

ability and tagging powers split by the different flavour tagging algorithms for signal and B^+ background as investigated for the decay $B^0 \to D^{*-}\mu^+X$. For the decay $B^0 \to D^-\mu^+X$ the effects are comparable but were not studied in detail. Efficiency and mistag observed in the B^+ background sample are significantly larger than in the signal sample for the vertex charge and the same-side pion tagger. This is caused by the π^+ that is missed in the signal-side reconstruction. In the vertex charge tagging algorithms it is associated with the secondary vertex in about 28% of the reconstructed background decays. In the same-side pion tagger this pion is selected in about 25% of the background decays. In both cases, this leads to the assignment of a wrong tag. The opposite-side kaon tagger seems as well to perform differently on signal and B^+ background, yet the reason for this has not been found. Using only events that have a tagging decision cancels out any difference in tagging efficiencies. This reduces the number of selected events from 3.06 million to 1.08 million, which is acceptable as the untagged events do not contribute to the sensitivity on measuring Δm_d. They would only contribute to the fit to the invariant D^- mass and the shape of the decay time acceptance. The muon and electron taggers seem to have similar mistag probabilities within uncertainties for both signal and B^+ background, therefore only events tagged by one of these taggers will be used in this analysis.
This choice reduces the number of selected events further from 1.08 million to 227581, as the leptonic taggers have an efficiency of 7.44%. Alternatively, the leptonic taggers could only be used to fit for the B^+ background fraction and be fixed in a fit where all taggers are used. Then again, the different tagging efficiencies of signal and B^+ background have to be considered. These can only be evaluated from simulated events, leading to additional uncertainties. Therefore, due to time constraint of this thesis, the lower statistical power available from the leptonic taggers is the preferred way for this analysis, while the alternative approach might be used in future analyses.
Considering the tagging parameters, the $\mathcal{PDF}$ describing the decay time distribution of the signal decay changes to

$$\mathcal{PDF}_{B^0}(t,q) = \frac{\Gamma_d}{2} e^{-\Gamma_d t} \cdot (1 - q\,(1 - 2\,\omega(\eta)) \cos(\Delta m_d\, t)). \tag{6.9}$$

Due to mistags, the physically non-mixing B^+ background events can as well be assigned with the "mixed" tag. Therefore, it is described by two $\mathcal{PDF}$s, one for each mixing decision:

$$\mathcal{PDF}_{B^+,q=-1}(t) = (1 - \omega(\eta))\,\Gamma_u e^{-\Gamma_u t},$$
$$\mathcal{PDF}_{B^+,q=+1}(t) = \omega(\eta)\,\Gamma_u e^{-\Gamma_u t},$$

which can be written as one combined function

$$\mathcal{PDF}_{B^+}(t,q) = \left(\frac{1}{2}(1-q) + q\,\omega(\eta)\right) \Gamma_u e^{-\Gamma_u t}. \tag{6.10}$$

The parameter $\omega(\eta)$ is the mistag probability of the respective event calculated from the predicted mistag η and the calibration parameters which were given earlier in Tab. 3.1. With these calibration parameters applied, η should give the correct mistag probability ω for each event. However, possible backgrounds that are not considered might change the mistag probability. To account for that a recalibration will be done while fitting, fixing the calibration parameter p_1 to 1 and leaving p_0 floating.

The mistag probability of the sideband background will be described by a global parameter ω_{SB}, which is floating in the fit. With this the sideband decay time $\mathcal{PDF}$ is given by

$$\mathcal{PDF}_{SB}(t_{meas}, q) = \frac{\Gamma_{SB}}{2} e^{-\Gamma_{SB}\, t_{meas}} \cdot (1 - q\,(1 - 2\,\omega_{SB}) \cos(\Delta m_{SB}\, t_{meas})). \tag{6.11}$$

However, by using the per-event mistag probability η, a new variable is introduced. To account for this, the probability of an event to be categorised into sideband, B^+ background and signal has to additionally be weighted by the respective $\mathcal{PDF}(\eta)$ [38]. However, the distributions of η are compatible for all three fit components, as shown in Fig. 6.1. Due to this, $\mathcal{PDF}(\eta)$ is a common multiplication factor in all components and thus does not have an effect on the fit mimimisation. Therefore, it will not be used in the final fit.

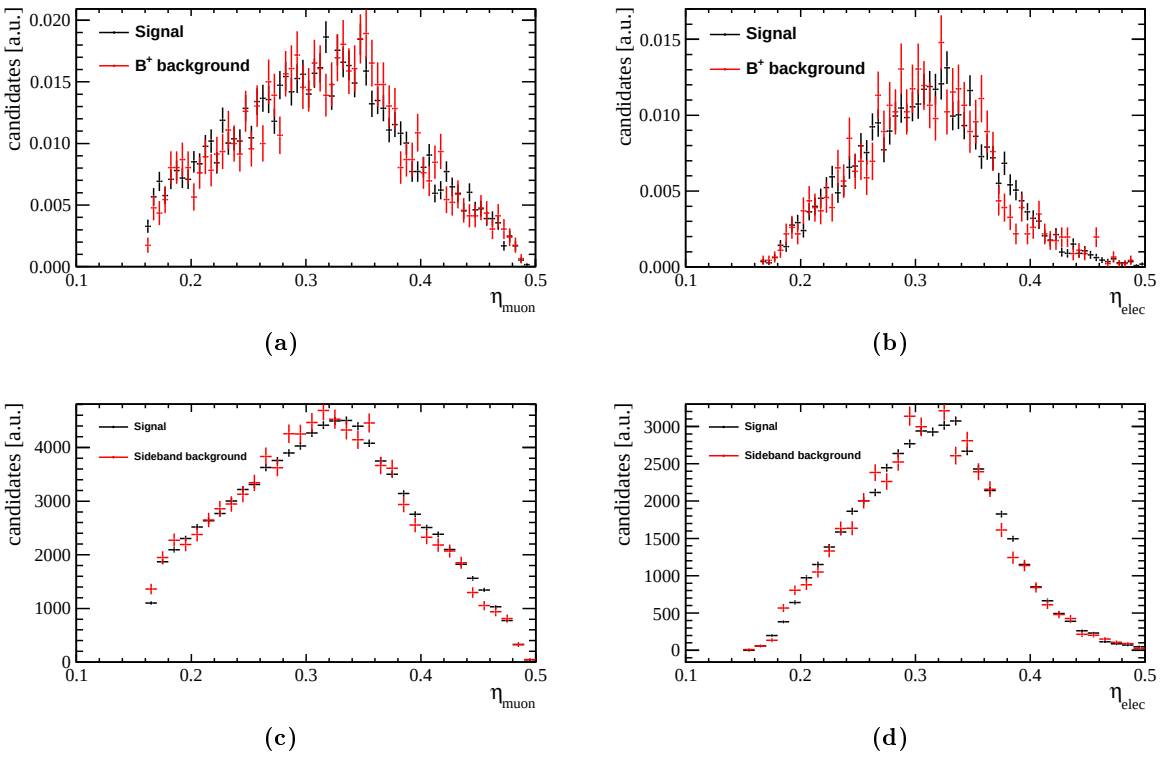


Figure 6.1: *(a) and (b) show the comparison of the η distributions for signal and B^+ background derived from simulation. (c) and (d) show the comparison of the η distributions for sideband background and peaking events in the invariant D^- mass on data.*

6.3 Missing momentum correction

The main difference of this analysis compared to other analyses measuring Δm_d in semileptonic B^0 decays is the treatment of the missing B^0 momentum. Instead of correcting the decay time with an average k-factor, the $\mathcal{PDF}$ used in the fit will be transformed to describe the

uncorrected decay time. This uncorrected, measured decay time t_{meas} is defined as

$$t_{meas} = \frac{L \cdot M_{B^0,PDG}}{p_{meas}}, \tag{6.12}$$

where L is the flight distance of the B^0 in the LHCb detector before decaying, $M_{B^0,PDG}$ its nominal mass according to PDG data [30] and p_{meas} its reconstructed momentum. As described in Chapter 4 the fraction of reconstructed and true momentum is expressed by the k-factor (Formula (4.5)). Simulated signal events are used to derive the distributions of the k-factor for signal and B^+ background individually. Figure 6.2 shows these k-factor distributions. As besides the neutrino also at least one π^+ is not reconstructed in case of B^+ background the k-factor is in general lower for it than for the signal decay.
Assuming these distributions describe the reduction of the reconstructed momentum on data correctly, these distributions are used to change the decay time $\mathcal{PDF}$ to a function of the reconstructed decay time t_{meas}. The k-factor distributions are normalised and each separated into 20 bins. Higher amounts of bins were also tested but did not result in an increase of sensitivity. Each of these bins with has a k-factor k_j $(j = 1, ..., 20)$ and a weight A_j, where the weights are normalised before the fit such that $\sum_{j=1}^{20} A_j = 1$. The measured decay time depending on the k-factor k_j is calculated by $t = t_{meas} \cdot k_j$. Using this, the measured decay time $\mathcal{PDF}$ is achieved via a numerical convolution with the k-factor distribution. This numerical convolution corresponds to a weighted summation, transforming the signal $\mathcal{PDF}$ into

$$\mathcal{PDF}_{B^0}(t_{meas}, q) = \sum_{j=1}^{20} A_{j,B^0} \frac{\Gamma_d \, k_j}{2} \, e^{-\Gamma_d \, t_{meas} \, k_j} \, (1 - q \, (1 - 2 \, \omega(\eta)) \cos(\Delta m_d \, t_{meas} \, k_j)). \tag{6.13}$$

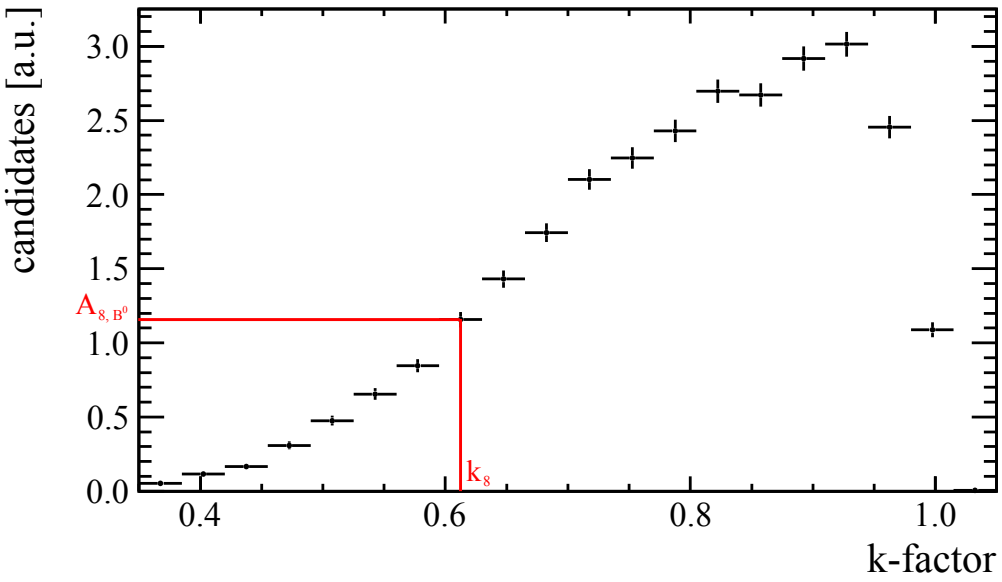
The B^+ background decay time distribution is accordingly described by

$$\mathcal{PDF}_{B^+}(t_{meas}, q) = \left(\frac{1}{2}(1 - q) + q \, \omega(\eta) \right) \sum_{j=1}^{20} A_{j,B^+} \, \Gamma_u \, k_j \, e^{-\Gamma_u \, t_{meas} \, k_j}. \tag{6.14}$$

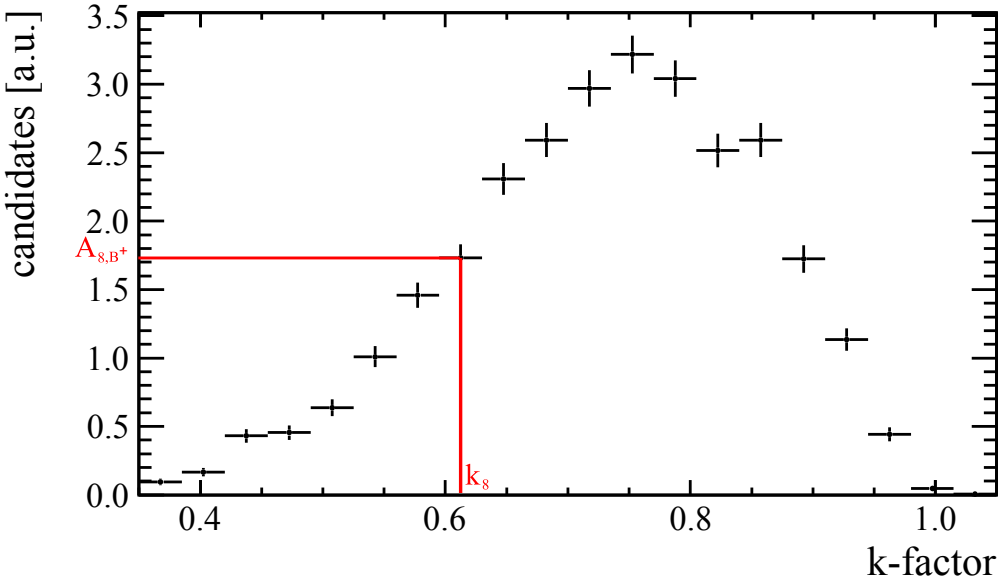
The coefficients A_{j,B^0} and A_{j,B^+} are the weights of a certain k-factor k_j derived from the respective histogram.

6.4 Decay time resolution

Measured physical quantities always suffer from resolution effects due the limited precision of the instruments. These limitations affect the decay time measurement in the LHCb detector. The effect of the decay time resolution will be evaluated by using simulated events. The decay time as defined in Formula (6.12) depends on two measured quantities, the flight distance L and the measured momentum p_{meas}. The measured flight distance L is compared to its true value L_{true} for each simulated event to estimate the magnitude of resolution effects. Due to the missing neutrino the measured B^0 momentum is already lowered compared to its true value, so evaluating its momentum resolution can not be done in such a way. However, any resolution effects on the momentum is already included in the k-factor as it is the fraction of the measured to the true momentum. Therefore, it will be neglected in the decay time

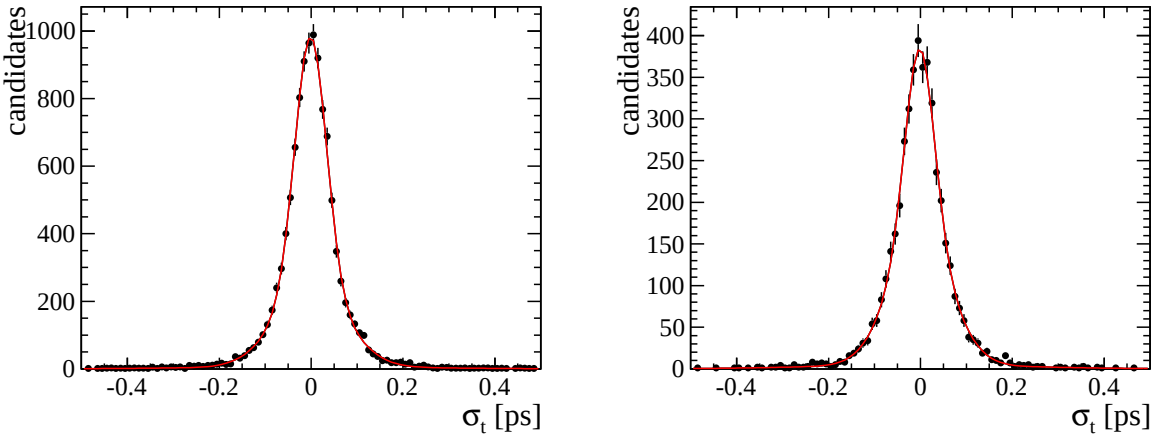


(a)



(b)

Figure 6.2: *k-factor distributions for the signal decay (a) and B^+ background (b) used in the fit. Each of the 20 k-factors k_j is assigned a weight A_j for each signal and B^+ background.*



(a) (b)

Figure 6.3: *Decay time resolutions for the signal decay (a) and B^+ background (b), both fitted with the sum of three Gaussian distributions.*

resolution.
For each simulated event the effect of the decay time resolution is calculated as

$$\sigma_t = \frac{(L - L_{true}) \cdot M_{PDG,B^0}}{p_{true}}. \tag{6.15}$$

The distributions of this quantity are shown in fig. 6.3 both for a simulation sample of signal events and simulated B^+ background events, respectively. They are modeled by the sum of three Gaussian distributions, which describes them well:

$$\begin{aligned}\mathcal{PDF}_{reso}(\sigma_t) = {} & f_{res,1} \cdot \text{Gaussian}(\sigma_t|\mu_{res}, \sigma_{res,1}) \\ & + (1 - f_{res,1}) \cdot f_{res,2} \cdot \text{Gaussian}(\sigma_t|\mu_{res}, \sigma_{res,2}) \\ & + (1 - f_{res,1}) \cdot (1 - f_{res,2}) \cdot \text{Gaussian}(\sigma_t|\mu_{res}, \sigma_{res,3}).\end{aligned} \tag{6.16}$$

Fit parameter	signal	B^+ background
μ_{res}	(-1.027 ± 0.475) fs	(-1.507 ± 0.789) fs
$\sigma_{res,1}$	$(3.402 \pm 0.121)10^{-2}$ ps	$(6.581 \pm 0.466)10^{-2}$ ps
$\sigma_{res,2}$	$(7.602 \pm 0.385)10^{-2}$ ps	$(3.136 \pm 0.289)10^{-2}$ ps
$\sigma_{res,3}$	$(24.23 \pm 4.44)10^{-2}$ ps	$(22.22 \pm 5.17)10^{-2}$ ps
$f_{res,1}$	0.7623 ± 0.0323	0.3888 ± 0.0738
$f_{res,2}$	0.9746 ± 0.0087	0.9794 ± 0.0109
σ_{eff}	$(5.075 \pm 0.414)10^{-2}$ ps	$(5.380 \pm 0.702)10^{-2}$ ps

Table 6.2: *Fit results of the decay time resolutions derived from simulated signal and B^+ background events.*

The fit results are shown in table 6.2. The effective resolution σ_{eff} is calculated by

$$\sigma_{eff} = \sqrt{f_{res,1}\,\sigma^2_{res,1} + (1 - f_{res,1})\,f_{res,2}\,\sigma^2_{res,2} + (1 - f_{res,1})\,(1 - f_{res,2})\,\sigma^2_{res,3}}. \tag{6.17}$$

The effective resolutions of signal and B^+ background are compatible with each other. This is expected, as resolution effects usually originate from technical limitations of the detector or reconstruction. Both signal and background events should be affected by these effects in the same way. Therefore, the same resolution model will be used in all the decay time $\mathcal{PDF}$s.
The decay time resolution has to be accounted for in the decay time $\mathcal{PDF}$s by means of an analytical convolution with the resolution function derived from the signal fit. Each $\mathcal{PDF}$ is convoluted with $\mathcal{R}(t)$,

$$\mathcal{PDF}'_i(t_{meas}) = \mathcal{PDF}_i(t_{meas}) \otimes_t \mathcal{R}(t) = \int_{-\infty}^{\infty} \mathcal{PDF}_i(t)\mathcal{R}(t_{meas} - t)dt. \tag{6.18}$$

The resolution function $\mathcal{R}(t)$ used for the convolution is given by

$$\begin{aligned}\mathcal{R}(t) = {} & f_{res,1} \cdot \text{Gaussian}(t|\mu_{res}, \sigma_{res,1}) \\ & + (1 - f_{res,1}) \cdot f_{res,2} \cdot \text{Gaussian}(t|\mu_{res}, \sigma_{res,2}) \\ & + (1 - f_{res,1}) \cdot (1 - f_{res,2}) \cdot \text{Gaussian}(t|\mu_{res}, \sigma_{res,3}),\end{aligned} \tag{6.19}$$

where the parameters are taken from the fit on the resolution distribution of the signal.
As a convolution is linear per definition, convolving any of the $\mathcal{PDF}$s with $\mathcal{R}(t)$ can be split in two general convolutions

$$e^{-\lambda\, t_{meas}} \quad \otimes_t \text{Gaussian}(t|\mu, \sigma)$$

and

$$\left(e^{-\lambda\, t_{meas}} \cdot \cos(\omega\, t_{meas})\right) \otimes_t \text{Gaussian}(t|\mu, \sigma).$$

The execution of these convolutions can be found in Appendix A. As an abbreviation, the convolution with the decay time resolution will be denoted as $\otimes_t \mathcal{R}(t)$ when describing $\mathcal{PDF}$s in the following sections.

6.5 Decay time acceptance

An important effect that has to be considered when fitting decay time distributions is the decay time acceptance. Due to selection criteria applied to the data sample the B^0 decay time is biased. For example, the daughter particles of the B^0 are required to have a certain displacement from the primary vertex to reject background. The specific requirements applied in this analysis can be found in chapter 5. This leads to the rejection of short-lived B^0 mesons, as their decay vertex is close to the primary vertex and thus their daughter particles have a

(a) (b)

Figure 6.4: *Particles originating from short-lived B^0 mesons (a) have in general lower impact parameters than for long-lived B^0 mesons (b). A cut on the impact parameter thus removes short-lived B^0 and introduces a bias of the measured decay time.*

small displacement from it. Figure 6.4 illustrates this effect. This decay time acceptance is defined as the time-dependent ratio of the number of reconstructed B^0 decays after applying the selection criteria compared to the number of B^0 decays without any selection applied

$$A(t_{meas}) = \frac{\#\text{selected reconstructed decays}(t_{meas})}{\#\text{all decays}(t_{meas})}. \tag{6.20}$$

The decay time fit is performed in the space of the uncorrected, reconstructed decay time t_{meas} of the B^0 instead of its true decay time. Therefore, the acceptance has to be evaluated for t_{meas} to use it as a multiplicative factor in the decay time $\mathcal{PDF}$.
As an estimation of the shape of the decay time acceptance, simulated events are used for signal and B^+ background. The reconstructed decay time distribution of these samples is compared to the distribution that is expected from an exponential decay when smeared with decay time resolution and k-factor distribution. The decay time resolution model and k-factor distribution derived in the former sections are used for this evaluation. The decay time acceptance as a function of the reconstructed B^0 decay time is defined as

$$A(t_{meas}) = \frac{\#\text{selected reconstructed decays}(t_{meas})}{\#\text{events from the distribution}\left[(\Gamma e^{-\Gamma t}) \otimes_k F(k) \otimes_t \mathcal{R}(t)\right](t_{meas})}, \tag{6.21}$$

where $F(k)$ is the k-factor distribution and $\mathcal{R}(t)$ the decay time resolution. The convolution with the k-factor distribution is in fact a numerical convolution with a k-factor histogram divided into 100 bins. While the numerator contains the reconstructed decay time distributions derived from simulation samples and is therefore limited by the number of events generated in these samples, the denominator can be created in an arbitrary size. To avoid statistical uncertainties originating from the denominator, one million events are generated for each signal and B^+ background. This is significantly larger than the number of events passing the selection applied to the simulation samples ($\mathcal{O}(10k)$). The decay time acceptances derived in this way are shown in fig. 6.5 for signal and B^+ background. The shapes for both signal and background are comparable and can be modeled well by the function

$$A(t_{meas}) = a_0 \left(1 - e^{-\frac{t_{meas} - t_{shift}}{\alpha}}\right) \cdot (1 + \beta\, t_{meas}), \tag{6.22}$$

where α, β and t_{shift} are parameters of the fit. The parameter a_0 is a normalisation factor, accounting for the different number of events contained in the numerator and denominator.

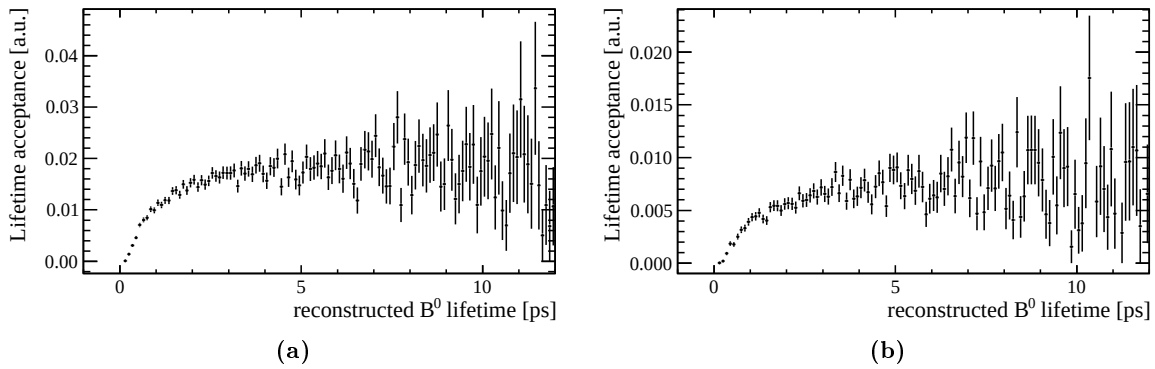


Figure 6.5: *Decay time acceptances for the signal decay (a) and B^+ background (b)*

To avoid the technical problem of negative values of this function, a cut has to be performed on B^0 decay times smaller than t_{shift}. The final decay time cut selected to account for this will be discussed at the end of this section. Figure 6.6 shows the signal acceptance fitted with this function in the decay time range $t_{meas} \in [0.2, 12.0]$ ps. The respective values of the fit parameters are shown in table 6.3. The value of β is smaller than its error and compatible with 0. It therefore seems to be an unnecessary degree of freedom and will be fixed to 0.

The acceptance is used as one common multiplication factor in the decay time fit for all flavour tagging states (mixed and unmixed). To account for disagreement between data and simulation, α will be left floating in the decay time fit. The parameter t_{shift} will be fixed to the value derived from simulation, as it is highly correlated with α. The effect on the value of Δm_d will be evaluated as a systematic uncertainty in Chapter 8. The B^+ background fraction as well as the individual acceptances affect the relative amplitude of signal and B^+ background in the final $\mathcal{PDF}$. In order to fit for the fraction of B^+ background present in

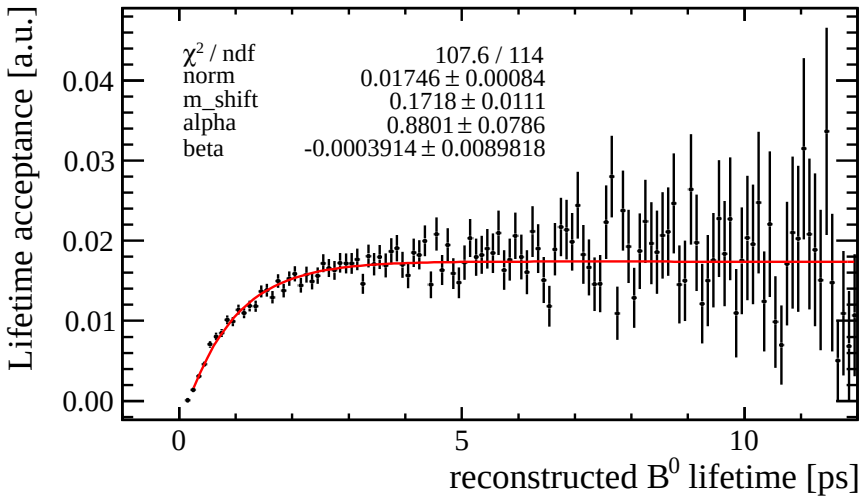


Figure 6.6: *Fitted acceptance of the signal Monte Carlo sample.*

Fit parameter	value
a_0	0.01746 ± 0.00084
t_{shift}	(0.1718 ± 0.0111) ps
α	(0.8801 ± 0.0786) ps
β	$(-0.0003914 \pm 0.0089818)$ ps^{-1}

Table 6.3: *Fit results for the decay time acceptance of the signal sample.*

the invariant D^- mass peak, its decay time acceptance cannot be varied independently of the signal acceptance. However, this can be solved by leaving the parameter α floating only for the signal acceptance and fixing the B^+ acceptance by the ratio of the two acceptances. The ratio of signal and B^+ background acceptances is modeled well by a first order polynomial

$$\frac{A_{B^+}}{A_{B^0}}(t_{meas}) = p_0 + p_1\, t_{meas}. \tag{6.23}$$

Figure 6.7 shows the acceptance ratio for simulated data fitted with this polynomial. The result of this fit is shown in table 6.4. As the slope parameter p_1 is compatible with 0, the same acceptance function will be used to describe signal and B^+ background. The value of p_0 is a scale factor resulting from the different amount of events simulated for signal and B^+ background. The decay time acceptance appears as a multiplication factor in the final $\mathcal{PDF}$.

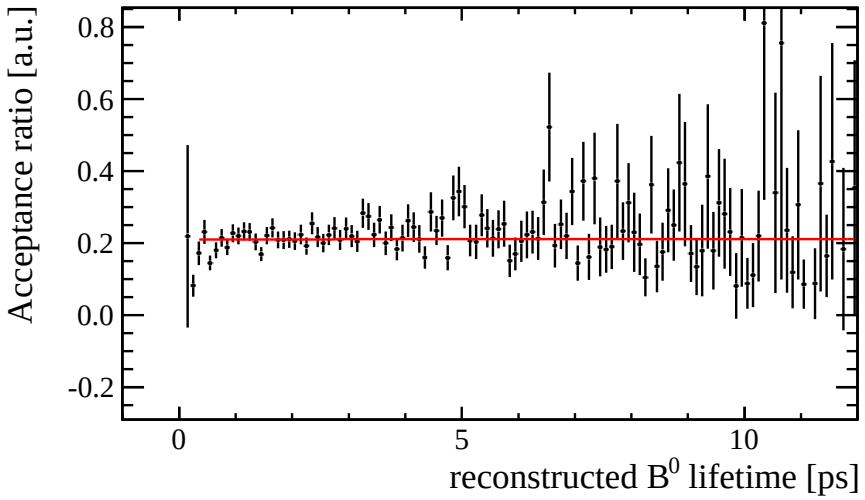


Figure 6.7: *B^+ to signal decay time acceptance ratio derived from simulation, fitted with a first order polynomial.*

Fit parameter	value
p_0	0.21059 ± 0.00635
p_1	(0.000021 ± 0.001721) ps^{-1}

Table 6.4: *Fit results for the ratio of B^+ to signal decay time acceptances.*

Thus this factor will be canceled by normalisation and will therefore be neglected.
The acceptance of the sideband background is not investigated using simulated events as it can be extracted from data simultaneously with its decay time from the D^- mass sidebands. To investigate the acceptance of the sideband background, data events of the lower sideband will be used. The acceptance is well modeled by an exponential growth, similar to what is used for signal and B^+ background. As the acceptance is investigated using data it is extracted from a simplified decay time $\mathcal{PDF}$ of the sideband background using no tagging information

$$\mathcal{PDF}_{SB}(t_{meas}) = e^{-\Gamma_{SB}\, t_{meas}} \left(1 - e^{-\frac{t_{meas}-t_{shift,SB}}{\alpha_{SB}}}\right), \tag{6.24}$$

where the second term describes the decay time acceptance of the sideband. The decay time distribution of the lower D^- mass sideband ($m_{K\pi\pi} < 1830\,\mathrm{MeV}/c^2$) fitted with this function is shown in Fig. 6.8. The parameters derived from this fit are shown in table 6.5. Similar to the signal acceptance the parameters α_{SB} and $t_{shift,SB}$ are highly correlated, thus the latter will be fixed while α_{SB} and Γ_{SB} will be left free to extract more accurate values using the whole invariant D^- mass spectrum.
$t_{shift,SB}$ is larger than t_{shift} derived for signal and background. In order for the $\mathcal{PDF}$ not to return negative values, all events of the data sample with a decay time lower than 0.3 ps will be rejected. This reduces the number of selected candidates further from 227581 to 225538.

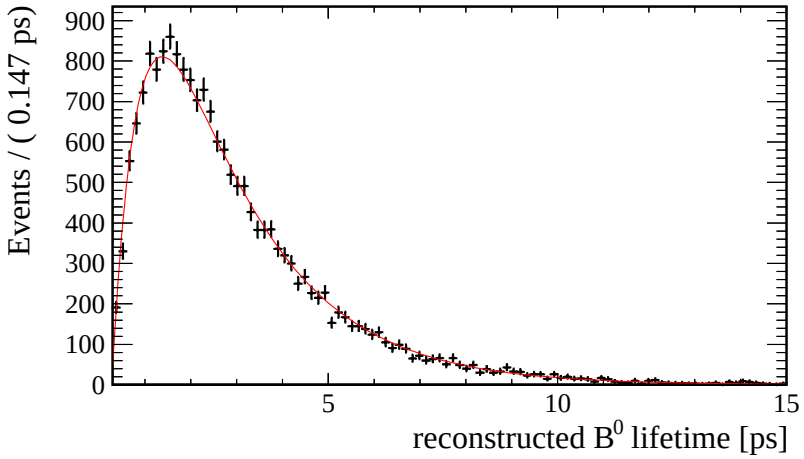


Figure 6.8: *Fitted reconstructed decay time of the lower D^- mass sideband.*

Fit parameter	value
Γ_{SB}	$(0.48660 \pm 0.005289)\,\mathrm{ps}^{-1}$
$t_{shift,SB}$	$(0.275871 \pm 0.006231)\,\mathrm{ps}$
α_{SB}	$(0.96482 \pm 0.05100)\,\mathrm{ps}^{-1}$

Table 6.5: *Fit results for the simultaneous fit of decay time and acceptance in the upper D^- mass sideband.*

6.6 Final fit $\mathcal{PDF}$

With all the extensions discussed in the previous sections, the $\mathcal{PDF}$s to describe the distribution of the reconstructed uncorrected decay time t_{meas} are given by

$$\begin{aligned}
\mathcal{PDF}_{B^0}(t_{meas}, q) &= \left[\left(\sum_{j=1}^{20} A_{j,B^0} \frac{\Gamma_d\, k_j}{2}\, e^{-\Gamma_d\, t_{meas}\, k_j} \left(1 - q\,(1 - 2\,\omega(\eta)) \cos(\Delta m_d\, t_{meas}\, k_j)\right) \right) \otimes_t \mathcal{R}(t) \right] \\
&\quad \cdot \left(1 - e^{-\frac{t_{meas} - t_{shift}}{\alpha}}\right) (1 + \beta\, t_{meas}), \\
\mathcal{PDF}_{B^+}(t_{meas}, q) &= \left[\left(\left(\frac{1}{2}(1 - q) + q\,\omega(\eta)\right) \sum_{j=1}^{20} A_{j,B^+}\, \Gamma_u\, k_j\, e^{-\Gamma_u\, t_{meas}\, k_j} \right) \otimes_t \mathcal{R}(t) \right] \\
&\quad \cdot \left(1 - e^{-\frac{t_{meas} - t_{shift}}{\alpha}}\right) (1 + \beta\, t_{meas}),
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{PDF}_{SB}(t_{meas}, q) &= \left[\left(\frac{\Gamma_{SB}}{2} e^{-\Gamma_{SB}\, t_{meas}} \cdot \left(1 - q\,(1 - 2\,\omega_{SB}) \cos(\Delta m_{SB}\, t_{meas})\right) \right) \otimes_t \mathcal{R}(t) \right] \\
&\quad \cdot \left(1 - e^{-\frac{t_{meas} - t_{shift,SB}}{\alpha_{SB}}}\right). \qquad (6.25)
\end{aligned}$$

Calibrated per-event mistag and true mistag probability might differ due to reasons stated in section 6.2. Therefore the per-event mistag probability used in the fit is defined by

$$\omega(\eta) = \eta - \langle \eta \rangle + p_0, \qquad (6.26)$$

where p_0 is a free parameter for both electron and muon taggers individually. The fit to data is performed simultaneously in three dimensions: mixing decision, invariant D^- mass and reconstructed B^0 decay time. The overall fit $\mathcal{PDF}$ as a function of mixing decision q, invariant D^- mass and reconstructed B^0 decay time is defined as

$$\begin{aligned}
\mathcal{PDF}(m_{K\pi\pi}, t_{meas}, q) = f_{sig} \cdot (1 - f_{SB}) &\quad \cdot \mathcal{PDF}_{peak}(m_{K\pi\pi}) \cdot \mathcal{PDF}_{B^0}(t_{meas}, q) \\
+ (1 - f_{sig}) \cdot (1 - f_{SB}) &\quad \cdot \mathcal{PDF}_{peak}(m_{K\pi\pi}) \cdot \mathcal{PDF}_{B^+}(t_{meas}, q) \\
+ f_{SB} &\quad \cdot \mathcal{PDF}_{SB}(m_{K\pi\pi}) \cdot \mathcal{PDF}_{SB}(t_{meas}, q). \qquad (6.27)
\end{aligned}$$

A list of all parameters of this description and whether they are fixed or left floating in the fit is shown in table 6.6. The mean decay widths Γ_d and Γ_u are fixed from the values of the respective mean lifetimes according to the PDG [30], as known at the time of this thesis. With 11 parameters fixed as described in previous sections, 15 parameters remain floating in the final fit.

category	parameter name	treatment
general	f_{sig}	floating
	f_{SB}	floating
reconstructed D^- mass	f_m	floating
	μ_{peak}	floating
	$\sigma_{m,1}$	floating
	$\sigma_{m,2}$	floating
	λ_{SB}	floating
mean decay widths	Γ_d	fixed to $0.60938\,\mathrm{ps}^{-1}$
	Γ_u	fixed to $0.65833\,\mathrm{ps}^{-1}$
	Γ_{SB}	floating
decay time acceptance	t_{shift}	fixed to 0.1718 ps
	$t_{shift,SB}$	fixed to 0.2759 ps
	α	floating
	α_{SB}	floating
	β	fixed to $0\,\mathrm{ps}^{-1}$
decay time resolution	μ_{res}	fixed to -1.027 fs
	$\sigma_{res,1}$	fixed to 0.03402 ps
	$\sigma_{res,1}$	fixed to 0.07602 ps
	$\sigma_{res,1}$	fixed to 0.2423 ps
	$f_{res,1}$	fixed to 0.7623
	$f_{res,2}$	fixed to 0.9746
mixing and tagging	Δm_{SB}	floating
	ω_{SB}	floating
	$p_{0,elec}$	floating
	$p_{0,muon}$	floating
	Δm_d	floating

Table 6.6: *List of parameters of the complete fit $\mathcal{PDF}$ and their treatment in the fit.*

6.7 Fit validation

It is necessary to ensure that the $\mathcal{PDF}$s derived in this chapter are correctly implemented and that no correlations between floating parameters lead to any biases in the physics parameters extracted. To validate the functionality of the fitting process, pseudo-experiments are performed. In these pseudo-experiments 50000 events are generated with the previously described $\mathcal{PDF}$s and parameters. This is a lower number than the number of events in the data sample (225538), but should be sufficient to show potential issues of the fit. When fitting the pseudo-data the same parameters are fixed as it is the case for data. The events generated this way are then fitted the same way as the data sample is treated. This pseudo-experiment is repeated 600 times. If the fitting process works properly, the values of the parameters extracted from these fits should be normal distributed. The mean values of these distributions was extracted by fitting a Gaussian distribution to each of them. The results of this procedure are shown in table 6.7. All means are compatible with the generated values. Another quantity that might show possible biases of the fit is the pull p of the fitted parameters. For a parameter A the pull p_A defined as

$$p_A = \frac{A_{fitted} - A_{generated}}{\sigma_A}, \tag{6.28}$$

where σ_A is the fit error of the respective parameter. In a correctly working fit, the distributions of these pulls for all parameters should be Gaussian, with a mean compatible with 0 and a standard deviation of 1. Figure 6.9 shows the pull distributions for all parameters of the nominal fit. No biases are observed from the investigation of the pulls.

parameter	**generated value**	**mean of fitted values**
f_{sig}	0.87	0.872 ± 0.0008
f_{SB}	0.25	0.25 ± 0.0001
f_m	0.3	0.312 ± 0.004
μ_{peak}	$1870\,\mathrm{MeV}/c^2$	$(1870 \pm 0.002)\,\mathrm{MeV}/c^2$
$\sigma_{m,1}$	$10\,\mathrm{MeV}/c^2$	$(10.1 \pm 0.03)\,\mathrm{MeV}/c^2$
$\sigma_{m,2}$	$6.5\,\mathrm{MeV}/c^2$	$(6.47 \pm 0.01)\,\mathrm{MeV}/c^2$
λ_{SB}	$5(\,\mathrm{GeV}/c^2)^{-1}$	$(5 \pm 0.008)(\,\mathrm{GeV}/c^2)^{-1}$
Γ_{SB}	$0.449\,\mathrm{ps}^{-1}$	$(0.449 \pm 0.0003)\,\mathrm{ps}^{-1}$
α	$0.8682\,\mathrm{ps}^{-1}$	$(0.869 \pm 0.0008)\,\mathrm{ps}^{-1}$
α_{SB}	$0.7428\,\mathrm{ps}^{-1}$	$(0.745 \pm 0.002)\,\mathrm{ps}^{-1}$
Δm_{SB}	$0.2\,\mathrm{ps}^{-1}$	$(0.199 \pm 0.0007)\,\mathrm{ps}^{-1}$
ω_{SB}	0.4	0.4 ± 0.0003
$p_{0,elec}$	0.307	0.307 ± 0.0003
$p_{0,muon}$	0.315	0.315 ± 0.0002
Δm_d	$0.510\,\mathrm{ps}^{-1}$	$(0.509 \pm 0.0005)\,\mathrm{ps}^{-1}$

Table 6.7: *Generated and fitted values for the free fit parameters derived from pseudo-experiments.*

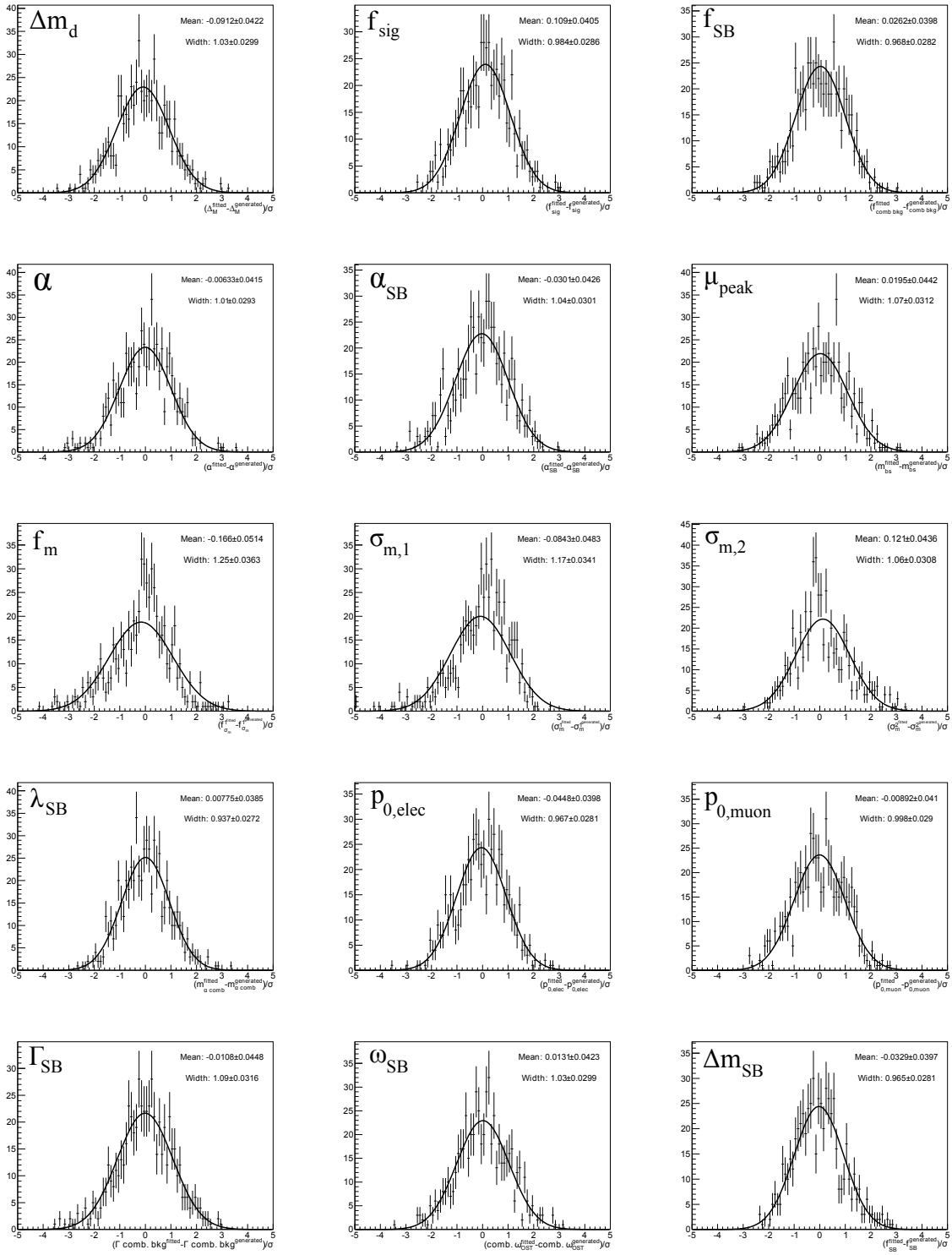


Figure 6.9: *Distributions of the pulls for all fit parameters.*

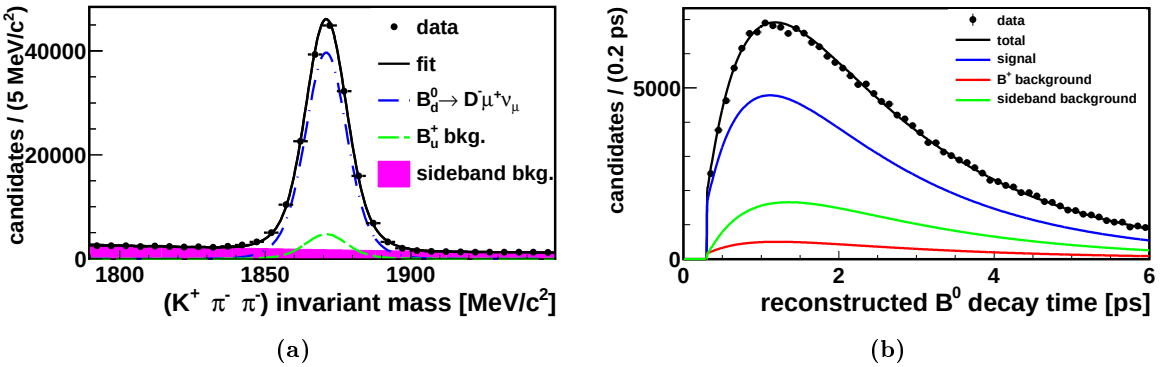
Chapter 7

Results

The results of the final fit to the 225538 signal candidates are shown in table 7.1. Figure 7.2 shows the fitted reconstructed D^- mass and B^0 lifetime distributions. The expectations are well met for most parameters. The acceptance parameter α differs significantly from what was derived from simulation, showing large disagreement from the acceptance derived from simulated events. As the parameter t_{shift} was fixed from the investigation of the acceptance of simulated events, this disagreement might affect the outcome of the fit. To evaluate the magnitude of this effect, it will be evaluated as a systematic uncertainty as presented in the next chapter.
The tagging parameters $p_{0,elec}$ and $p_{0,muon}$ are significantly higher than the average predicted mistags ($\langle\eta\rangle_{elec} \approx 30.9\%$ and $\langle\eta\rangle_{muon} \approx 31.1\%$). This might be caused by backgrounds that have not been considered so far. However, leaving the p_0 parameters floating compensates for this effect. This results in an increase of the statistical uncertainty evaluated from the minimisation procedure. Further investigation in this regard is needed, as the knowledge of the tagging performance has a significant impact on the sensitivity in measuring Δm_d. The effects of uncertainties of this knowledge on the value of Δm is evaluated in the next chapter. The fraction of B^+ background in the D^- mass peak $\frac{f_{B^+}}{f_{peak}} = 1 - f_{sig} = (10.64 \pm 1.06)\%$ is

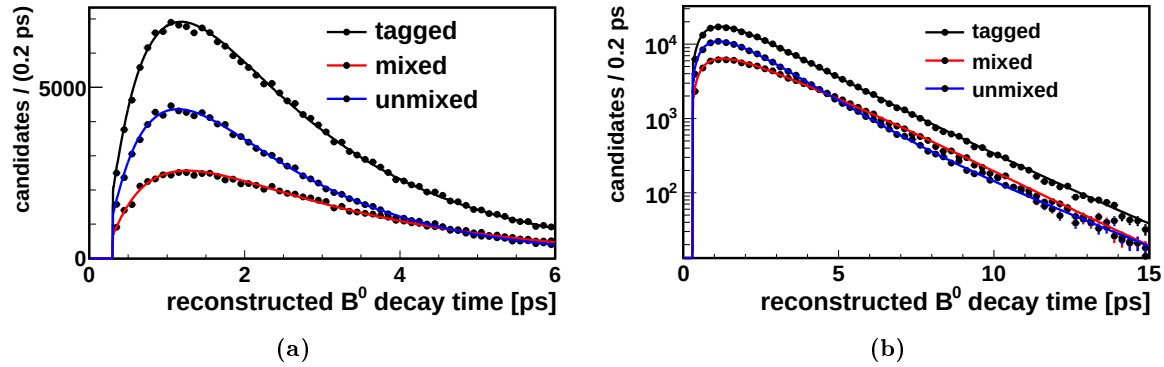
parameter	result
μ_{peak}	$(1870.9725 \pm 0.0215)\,\mathrm{MeV}/c^2$
f_m	0.3121 ± 0.0279
$\sigma_{m,1}$	$(11.418 \pm 0.319)\,\mathrm{MeV}/c^2$
$\sigma_{m,2}$	$(6.5306 \pm 0.0884)\,\mathrm{MeV}/c^2$
λ_{SB}	$(5.3849 \pm 0.0957)(\,\mathrm{GeV}/c^2)^{-1}$
f_{sig}	0.8936 ± 0.0106
f_{SB}	0.25618 ± 0.00137
Γ_{SB}	$(0.48155 \pm 0.00373)\,\mathrm{ps}^{-1}$
Δm_{SB}	$(0.2548 \pm 0.0112)\,\mathrm{ps}^{-1}$
α	$(0.75306 \pm 0.00868)\,\mathrm{ps}$
α_{SB}	$(0.9004 \pm 0.0279)\,\mathrm{ps}$
Δm_d	$(0.51030 \pm 0.00653)\,\mathrm{ps}^{-1}$
ω_{SB}	0.39513 ± 0.00338
$p_{0,muon}$	0.33545 ± 0.00238
$p_{0,elec}$	0.35088 ± 0.00332

Table 7.1: *Results of the nominal fit.*



(a) (b)

Figure 7.1: *Fitted distributions of the reconstructed invariant D^- mass (a) and reconstructed B^0 decay time (b). The projections of the fit components are also shown.*



(a) (b)

Figure 7.2: *Fitted distributions of the decay time fit separated into mixing decisions in linear (a) and logarithmic scale (b).*

slightly lower than what would have been expected from simulation ($\frac{f_{B^+}}{f_{peak}} = (12.69 \pm 0.66)\%$). However, both values are compatible within their errors.
Figure 7.1a shows the fitted invariant D^- mass spectrum and projections of the individual fit components. The fit describes the distribution well and the sideband background shows a flat behaviour in this spectrum, leading to a clear separation from peaking events.
The fitted reconstructed B^0 decay time distribution is shown in Fig. 7.1b, where also projections of the fit components are shown. Figure 7.2 shows the reconstructed B^0 decay time distribution separated into the mixing decisions "mixed" and "unmixed". The fit projection for those two decisions shows that both distributions are well described by the fit.

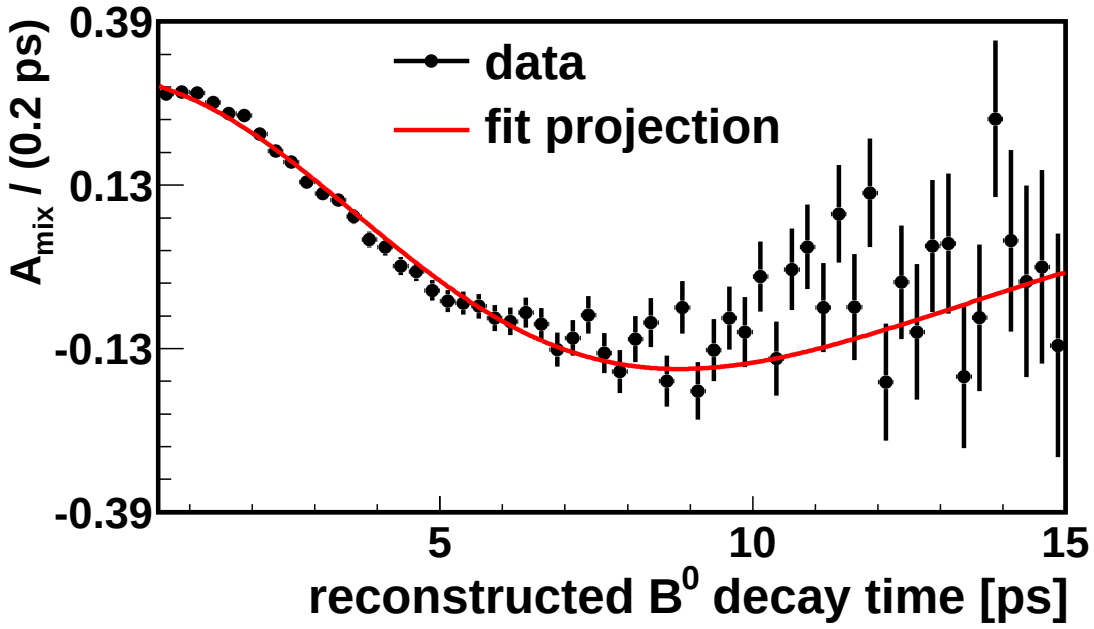


Figure 7.3: *Mixing asymmetry and projection of the result of the fit to the reconstructed B^0 decay time.*

A quantity which shows the mixing behaviour better than the lifetime distributions is the mixing asymmetry A_{mix}. It is calculated by

$$A_{mix} = \frac{N(unmixed) - N(mixed)}{N(unmixed) + N(mixed)}(t_{meas}), \tag{7.1}$$

where $N(mixed/unmixed)(t_{meas})$ is the number of mixed/unmixed candidates with the reconstructed B^0 decay time t_{meas}. Figure 7.3 shows the distribution of this quantity. The projection of the fit describes this spectrum sufficiently well. This is a projection of the $\mathcal{PDF}$ describing the reconstructed B^0 decay time and not a fit to the A_{mix} spectrum. The distribution of A_{mix} does not start at a value of 1 due to mistags. At $t_{meas} \approx 9\,\text{ps}$ most $B^0/\overline{B}^0$ mesons have oscillated into the respective antiparticle and the mixing asymmetry reaches its lowest point. It does not go back up to the initial value due to the contribution from the sideband background. This background is large fraction of the whole data sample ($\approx 25\%$), has a longer lifetime than the signal and oscillates with a lower frequency. Due to this the background becomes more dominant a high t_{meas} and the slower oscillation interferes with the B^0 mixing.

The correlations between the fit parameters as reported by the minimisation procedure Minuit [36] is shown in table 7.2. All correlations with values larger than 0.5 are printed in bold. The parameters describing the invariant D^- distribution of signal events and B^+ background are highly correlated. This is expected, as this distribution is described by the sum of two Gaussians, resulting in a certain ambiguity. Systematic uncertainties induced by the choice of the mass model will be discussed in the next chapter. Furthermore there is a high correlation between the acceptance parameter α_{SB} and the mean decay width of the sideband background.

However, the studies of pseudo-experiments showed no biases of these parameters.
There also is a significant correlation observed between the mixing frequency Δm_d and the fraction of B^+ background. In the B^0 sector in average only about 0.7 oscillations occur before a B^0 meson decays. Thus there is a high overlap between the amplitude of the decay and the oscillation. The studies using pseudo-experiments described in 6.7 however showed, that this correlation does not lead to a bias of Δm_d or the B^+ background fraction.

The result of this analysis is a measurement of the $B^0 - \overline{B}^0$ mixing frequency

$$\boxed{\Delta m_d = (0.5103 \pm 0.0065(\text{stat.}) \pm 0.0046(\text{syst.}))\,\text{ps}^{-1}}. \tag{7.2}$$

This is in excellent agreement with the current world average value according to the PDG of $\Delta m_d = (0.510 \pm 0.004)\ \text{ps}^{-1}$ [30]. The estimation of the systematic uncertainty used in this result will be treated in the next chapter.

	μ_{peak}	f_m	$\sigma_{m,1}$	$\sigma_{m,2}$	λ_{SB}	f_{sig}	f_{SB}	Γ_{SB}	Δm_{SB}	α	α_{SB}	Δm_d	ω_{SB}	$p_{0,muon}$	$p_{0,elec}$
μ_{peak}	1.00	-	-	-0.01	0.06	-	0.02	-	-	-	-	-	-	-	-
f_m		1.00	**-0.96**	**-0.96**	-0.09	-	0.32	-0.02	0.02	-	-0.05	-	-0.02	0.01	0.01
$\sigma_{m,1}$			1.00	**0.88**	0.12	-	-0.43	0.02	-0.03	-	0.06	0.01	0.03	-0.01	-0.01
$\sigma_{m,2}$				1.00	0.08	-	-0.29	0.02	-0.02	-	0.04	-	0.02	-0.01	-0.01
λ_{SB}					1.00	-	-0.17	-	-0.01	0.01	0.01	-	0.01	-	-
f_{sig}						1.00	-	-0.08	-0.13	0.43	-0.05	**-0.67**	0.07	0.14	0.07
f_{SB}							1.00	-0.01	0.03	-0.02	-0.04	-	-0.02	0.01	0.01
Γ_{SB}								1.00	0.01	-0.04	**0.79**	0.05	-0.01	-0.01	-
Δm_{SB}									1.00	-0.06	0.01	-0.03	-0.39	0.06	0.04
α										1.00	-0.12	-0.29	0.03	0.06	0.03
α_{SB}											1.00	0.03	-	-0.01	-
Δm_d												1.00	-0.06	-0.24	-0.11
ω_{SB}													1.00	-0.12	-0.09
$p_{0,muon}$														1.00	0.04
$p_{0,elec}$															1.00

Table 7.2: *Correlation matrix of the fit parameters, elements below the diagonal are not shown.*

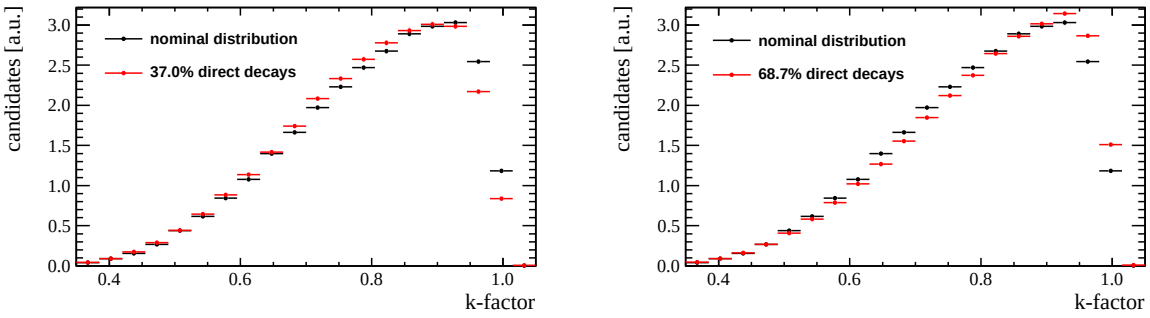
Chapter 8

Systematic uncertainties

In this chapter possible sources of systematic uncertainties will be discussed and their magnitude is estimated. The uncertainties will account for general modeling choices as well as effects caused by the determination of parameters from simulation.

8.1 k-factor determination

The k-factor distribution used to convolve the lifetime $\mathcal{PDF}$ to describe the reconstructed lifetime is based upon simulated decays. This simulation sample only consists to 52.88% of direct $B^0 \to D^- \mu^+ X$ decays, in the other cases the B^0 decays to the D^- via an intermediate D^* meson. In latter case, a neutral π is not reconstructed, which in general leads to a lower k-factor. The fraction of direct $B^0 \to D^- \mu^+ X$ decays contained in the simulation sample depends on the relative branching ratios compared to the other decays mentioned. This directly affects the k-factor distribution. The overall branching ratio of all decays of which the simulation sample is composed is $(4.95 \pm 0.11)\% \times \mathcal{BR}(D^- \to K^+\pi^-\pi^-)$, while the branching ratio of direct decays is $(2.18 \pm 0.12)\% \times \mathcal{BR}(D^- \to K^+\pi^-\pi^-)$. The uncertainties in these branching ratios leads to an uncertainty in the composition of the simulation sample. To investigate the magnitude of the effect of different sample compositions, the relative fraction of direct $B^0 \to D^- \mu^+ X$ decays is scaled by a factor of 1.3 and 0.7. That way, samples with a fraction of 68.7% and 37.0% direct decays are created, respectively. These are conservative values, which should as well account for uncertainties caused by differences between data and simulation caused by other effects. The k-factor distributions obtained this way are shown in fig. 8.1. These distributions are used each in the fit to determine their effect on Δm_d. The extracted value of Δm_d changes by $+0.0029\,\mathrm{ps}^{-1}$ when using the distribution with 37.0% direct decays and by -0.0016 ps^{-1} for the higher direct decay contribution. Effects of the binning scheme of the k-factor distributions are investigated by using 100 k-factor bins instead of 20. This results in a change of the extracted Δm_d value of $0.0001\,\mathrm{ps}^{-1}$. Compared to the effects caused by the composition of the simulation sample, this is negligible. The systematic assigned to uncertainties induced by the k-factor modeling is $\pm 0.0029\,\mathrm{ps}^{-1}$.



(a) (b)

Figure 8.1: *k-factor distributions with a low (a) and and high (b) direct $B^0 \to D^- \mu^+ X$ decay contribution compared to the distribution used in the nominal fit.*

8.2 B^+ background tagging performance

Signal and B^+ background showed to have different tagging performances, as discussed in section 6.2. As the fraction of B^+ background events peaking in the reconstructed D^- mass has to be determined from the fit, the tagging parameters cannot be determined for signal and B^+ background simultaneously. Therefore, only leptonic taggers are used in this analysis, which seem to perform comparable on both signal and background. To investigate the effect of a possible difference in tagging performance, the mistag probability of the background is changed by 3% with respect to the signal:

$$\omega_{B^+}(\eta) = \omega_{sig}(\eta) + 3\%,$$

where

$$\omega_{sig}(\eta) = \eta - \langle\eta\rangle + p_0.$$

This is done for electron and muon taggers simultaneously. With this change the whole fit is repeated, leaving $p_{0,elec}$ and $p_{0,muon}$ floating, as in the original fit. This changes the extracted value of Δm_d by $0.0024\,\mathrm{ps}^{-1}$, which is assigned as the systematic uncertainty for assuming the same tagging performances in signal and background.

8.3 Prompt D^- rejection

By rejecting all candidates where the $\log(\chi^2_{\mathrm{IP}})$ of the D^- is lower than 2.1, the contribution of D^- originating from the primary collision should be reduced to 0.5%, as discussed in section 5.3.3. To account for the effect of the remaining prompt D^- decays this cut will be loosened to an $\log(\chi^2_{\mathrm{IP}})$ value of 1.5, where the fraction of prompt D^- to signal decays is expected to be 1%. A refit of the data sample with the changed cut is used to evaluate the effect on Δm_d. As short-lived B^0 mostly also have a low value of $\log(\chi^2_{\mathrm{IP}})(D^-)$ the lifetime acceptance is reevaluated. The value of t_{shift}, which is fixed in the fit, changes that way to $t_{shift} = 0.1621\,\mathrm{ps}$.

The extracted Δm_d value derived from the fit differs from the nominal fit by $0.0017\,\mathrm{ps}^{-1}$. As a crosscheck, the $\log(\chi^2_{\mathrm{IP}})$ cut is tightened to a value of 3.05, where the prompt background fraction is expected to be 0.25%. This changes the extracted value of Δm_d by $0.0004\,\mathrm{ps}^{-1}$. The systematic uncertainty assigned to the prompt D^- treatment is $\pm 0.0017\,\mathrm{ps}^{-1}$.

8.4 Lifetime acceptance

The parameters modeling the lifetime acceptance of signal and B^+ background were derived from simulation. The acceptance parameter t_{shift} obtained from simulation was fixed in the nominal fit. From the lower D^- mass sideband the respective parameter $t_{shift,SB}$ was fixed as well. However, their values affect other parameters floating in the fit. To account for this, the fit is repeated where the values of the fixed acceptance parameters t_{shift} and $t_{shift,SB}$ are each individually changed by their errors. The results of these refits are shown in table 8.1. The effect of the changed $t_{shift,SB}$ can be neglected. The systematic uncertainty assigned to the fixing of acceptance parameters from simulation is $0.0016\,\mathrm{ps}^{-1}$.

new t_{shift} value	$\Delta m_d^{nominal} - \Delta m_d^{refit}$
$t_{shift} = 0.1829\,\mathrm{ps}$	-0.0015 ps^{-1}
$t_{shift} = 0.1607\,\mathrm{ps}$	0.0016 ps^{-1}
$t_{shift,SB} = 0.2821\,\mathrm{ps}$	0.00003 ps^{-1}
$t_{shift,SB} = 0.2697\,\mathrm{ps}$	-0.00004 ps^{-1}

Table 8.1: *Changes in Δm_d induced by different acceptance parameter values.*

8.5 Length and momentum scale

Scale uncertainties are induced by problems in the detector alignment and the calibration of the magnetic fields. This results in relative uncertainties of 0.1% in the length scale and 0.16% in the momentum scale [39]. Both the measured B^0 flight distance and momentum propagate these relative uncertainties linearly to the reconstructed B^0 lifetime and thus as well to Δm_d. Therefore, a relative uncertainty of 0.19% is assigned to Δm_d due to scale uncertainties. Using the result Δm_d obtained from the nominal fit this corresponds to a systematic uncertainty of $0.0010\,\mathrm{ps}^{-1}$.

8.6 Mass model

The D^- mass peak is modeled by a double Gaussian distribution in the nominal fit. To investigate uncertainties related to the specific mass model, the fit is repeated with using a

Crystal Ball function [40] and a Gaussian distribution with a common mean to model the D^- mass peak

$$\begin{aligned}\mathcal{PDF}_{peak}(m_{K\pi\pi}) =& f_{CB} \cdot \text{CrystalBall}(m_{K\pi\pi}|\mu_{peak}, \sigma_{CB}, \alpha_{CB}, n_{CB}) \\ &+ (1 - f_{CB}) \cdot \text{Gaussian}(m_{K\pi\pi}|\mu_{peak}, \sigma_{m,1}).\end{aligned}$$

The sideband model is not changed. The Crystal Ball parameter n_{CB} is fixed from a separate mass fit as it is highly correlated with α_{CB}. α_{CB} is describing the slope of the Crystal Ball tail and left floating in the fit. With this mass model the data can be well described. The extracted value of Δm_d changes by $0.0003\,\text{ps}^{-1}$, thus this is assigned as the systematic uncertainty of the mass model.

8.7 Summary of systematic uncertainties

Table 8.2 shows the list of systematic uncertainties and the respective magnitudes. The overall systematic uncertainty assigned to the measurement of Δm_d obtained in this analysis is $0.0046\,\text{ps}^{-1}$. This value is calculated from the sum of the squares of the individual systematic uncertainties. The dominant systematics in this analysis originate from the uncertainty of the k-factor description and the tagging performance of the B^+ background. The combined systematic uncertainty is smaller than the statistical uncertainty. Thus increasing the size of the dataset (for example by using in addition the data collected in 2012 by the LHCb experiment) or by the use of additional tagging algorithms might increase the sensitivity of this analysis significantly.

source	**systematic uncertainty**
k-factor description	$0.0029\,\text{ps}^{-1}$
B^+ tagging performance	$0.0024\,\text{ps}^{-1}$
prompt D^- rejection	$0.0017\,\text{ps}^{-1}$
scale uncertainties	$0.0010\,\text{ps}^{-1}$
lifetime acceptance	$0.0016\,\text{ps}^{-1}$
mass model	$0.0003\,\text{ps}^{-1}$
systematic combined	$0.0046\,\text{ps}^{-1}$
statistical uncertainty	$0.0065\,\text{ps}^{-1}$

Table 8.2: *Summary of uncertainties of* Δm_d.

Chapter 9

Conclusion

In this thesis the measurement of the $B^0 - \overline{B}^0$ mixing frequency Δm_d extracted from semileptonic $B^0 \to D^-(\to K^+\pi^-\pi^-)\mu^+X$ decays is presented. It is one of the first measurements of Δm_d using semileptonic decays at LHCb and uses a new approach to treat the missing reconstructed B^0 momentum.
To select signal events and reject different classes of background, a cut-based selection is applied. The selection increases the fraction of events peaking in the invariant D^- mass compared to flat background from 1.19 : 1 after the initial stripping selection to 3.15 : 1. The most prominent background originating from semileptonic $B^+ \to D^-\mu^+\nu_\mu X$ decays cannot be significantly reduced by any of these cuts.
The fit to data is performed in three dimensions: the mixing decision q, the reconstructed B^0 decay time t_{meas} and the reconstructed invariant D^- mass $m_{K\pi\pi}$. The probability density function used in the final fit consists of three components. These components describe the signal events, the B^+ background events and the sideband background events. While the invariant D^- mass distribution is the same for signal events and B^+ background they can be discriminated by exploiting the fact that the charged B mesons do not oscillate. However, problems in the flavour tagging performance in several tagging algorithms were discovered for this background category. To avoid these problems only the leptonic tagging algorithms are used in this analysis. This reduces the number of data events used in the fit by 79%.
The fit validity was checked by performing studies on a large set of pseudo-experiments. Fitting these successfully reproduced the values with which they were generated, showing the correct functionality of the fitting process. Model-dependent systematic uncertainties were separately investigated by repeating the fit to data with different assumptions applied.

The result obtained for the mixing frequency $\Delta m_d = (0.5103 \pm 0.0065(\text{stat.}) \pm 0.0046(\text{syst.}))\,\text{ps}^{-1}$ is in excellent agreement with former results and the world average value as provided by the Heavy Flavour Averaging Group [28]. The best single measurement of this quantity so far was obtained by the LHCb collaboration in 2012 using the decay channels $B^0 \to D^-\pi^+$ and $B^0 \to J/\psi\, K^{*0}$, where a value of $\Delta m_d = (0.5156 \pm 0.0051(\text{stat.}) \pm 0.0033(\text{syst.}))\,\text{ps}^{-1}$ was measured [31]. The uncertainties of the measurement presented in this thesis are slightly higher than the best single measurement, but there are several possible ways to further improve this result.

Two sources of systematic uncertainties are dominant in this study: the tagging performance of the B^+ background and the knowledge of the k-factor distribution.
To distinguish signal events from B^+ background events a detailed knowledge of the flavour tagging performance of both these categories is necessary. In the final fit used in this analysis

some tagging parameters were left floating, to account for non-considered background sources. A possible background source that is not considered in this analysis but discovered recently is the contribution from decays $B^0 \to D^+_{(s)} D^- X$ and $B^+ \to D^+_{(s)} \bar{D}^0 X$, where one of the D mesons decays semileptonically. From a first estimation the contribution of these decays is about 1 to 2% compared to the signal decay. Modeling the decay time distribution of this background and including it into the fit or searching for a cut rejecting the background are two possibilities to improve the fit result.
A known tagging performance and the knowledge that the B^+ background does not mix allows to determine its fraction. Many tagging algorithms performed differently on signal and B^+ background, thus only the leptonic tagging algorithms were used in this analysis. Additional studies of the performance of the remaining opposite-side taggers might allow to include them into the study. This would increase number of events used by this analysis by a factor of 5.
The k-factor distributions are obtained using simulated events. Uncertainty of the branching ratios of the contributing decay channels affects the k-factor distributions. The effect of this was roughly estimated by changing the contribution of direct $B^0 \to D^- \mu^+ \nu_\mu$ decays by $\pm 20\%$. This is a conservative approach which might be improved by further studies. Also more accurate measurements of the individual branching ratios might lower the uncertainty of the composition of the signal decay.
Studying these effects in detail might improve the statistical and systematic uncertainties significantly.

The result obtained in this analysis is supporting the new approach chosen to account for the missing reconstructed B^0 momentum in semileptonic decays. In the future it will be used in the measurement of the CP violating asymmetry a^d_{sl} which is currently ongoing. Although the measurement is not finished evaluations show promising results, competing with the most precise measurements of a^d_{sl}. Figure 7.3 shows a first projection of the mixing asymmetry. A publication of this asymmetry analysis and a more detailed analysis of Δm_d is planned for 2014.

Appendix A

Functions

This section contains mathematical functions used in this analysis, that were not fully explained previously.

The normalised Gaussian distribution, used in Formula (5.2) is defined as:

$$\text{Gaussian}(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp(-\frac{(x-\mu)^2}{2\sigma^2}) \tag{A.1}$$

The normalised Crystall Ball [40] distribution, used in Chapter 8 is defined as:

$$\text{CrystallBall}(x|\mu,\sigma,\alpha,n) = \frac{1}{N} \cdot \begin{cases} \exp(-\frac{(x-\mu)^2}{2\sigma^2}), & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) \cdot (\frac{n}{|\alpha|} - |\alpha| - \frac{x-\mu}{\sigma})^{-n}, & \text{for } \frac{x-\mu}{\sigma} \leqslant -\alpha \end{cases} \tag{A.2}$$

where

$$N = \sigma \cdot \left(\frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) + \sqrt{\frac{\pi}{2}}\left(1 + \text{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right)\right)$$

The normalised bifurcated Gaussian used in Formulas (5.7) and (5.8) is given as:

$$\text{BiGaussian}(x|\mu,\sigma_L,\sigma_R) = \sqrt{\frac{2}{\pi(\sigma_L+\sigma_R)^2}} \cdot \begin{cases} \exp(-\frac{(x-\mu)^2}{2\sigma_L^2}), \text{for } x < \mu \\ \exp(-\frac{(x-\mu)^2}{2\sigma_R^2}), \text{for } x \geqslant \mu \end{cases} \tag{A.3}$$

The convolution of a Gaussian with an exponential is described by

$$e^{-\lambda\, t_{meas}} \otimes_t \text{Gaussian}(t|\mu,\sigma) = \frac{\lambda}{2} e^{\frac{\lambda}{2}(2\mu+\lambda\sigma^2-2t_{meas})}\, \text{erfc}(\frac{\mu+\lambda\sigma^2-t_{meas}}{\sqrt{2}\sigma}). \tag{A.4}$$

The convolution of a Gaussian with the product of an exponential function and a cosine is described by

$$\left(e^{-\lambda\, t_{meas}} \cdot \cos(\omega\, t_{meas})\right) \otimes_t \text{Gaussian}(t|\mu,\sigma) =$$
$$\Re\Big(\frac{\lambda - i\omega}{2} e^{\frac{\lambda-i\omega}{2}(2\mu+\lambda\sigma^2-i\omega\sigma^2-2t_{meas})}\, \text{erfc}(\frac{\mu+\lambda\sigma^2-i\omega\sigma^2-t_{meas}}{\sqrt{2}\sigma})\Big). \tag{A.5}$$

List of Figures

Bibliography

[1] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367. 3

[2] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579. 3

[3] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264. 3

[4] E. D. Bloom *et al.*, *High-energy inelastic $e-p$ scattering at 6° and 10°*, Phys. Rev. Lett. **23** (1969) 930. 3

[5] J.-E. Augustin *et al.*, *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Physical Review Letters **33** (1974) 1406. 3

[6] J. J. Aubert *et al.*, *Experimental Observation of a Heavy Particle J*, Physical Review Letters **33** (1974) 1404. 3

[7] C. Campagnari and M. Franklin, *The discovery of the top quark*, Rev.Mod.Phys.69:137-212,1997, 1996. doi: 10.1103/RevModPhys.69.137. 3

[8] D. Collaboration, *Observation of tau neutrino interactions*, Phys.Lett.B504:218-224,2001, 2000. doi: 10.1016/S0370-2693(01)00307-0. 3

[9] T. A. Collaboration, *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, 2012. doi: 10.1016/j.physletb.2012.08.020. 3, 5

[10] G. Altarelli, *The standard model of particle physics*, 2005. 3

[11] G. Altarelli, *A qcd primer*, 2002. doi: 10.1063/1.1513677. 3

[12] G. Altarelli, *The standard electroweak theory and beyond*, 2000. 3

[13] P. Langacker, *Introduction to the standard model and electroweak physics*, 2009. 3

[14] `http://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg`. 4, 74

[15] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652. 5

[16] H. R. Quinn and P. F. Harrison, *The BaBar physics book: physics at an asymmetric B factory*, SLAC, Stanford, CA, 1998. 6

[17] `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/Material_for_Presentations.html`. 8, 12, 13, 15, 19, 74

[18] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970) 1285. 8

[19] A. D. Sakharov, *SPECIAL ISSUE: Violation of CP in variance, C asymmetry, and baryon asymmetry of the universe*, Soviet Physics Uspekhi **34** (1991) 392. 8

[20] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005. 14, 16, 17, 18, 74

[21] M. Adinolfi *et al.*, *Performance of the lhcb rich detector at the lhc*, 2012. 17

[22] M. Grabalosa and M. Musy, *Flavour Tagging developments within the LHCb experiment*, PhD thesis, Barcelona U., Mar, 2012, Presented 15 May 2012. 20, 21, 74

[23] T. Brambach et al., *Opposite side and same side pion tagging performances using 2011 data*, , LHCb-ANA-2012-005. 21

[24] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175`. 22

[25] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152. 22

[26] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250. 22

[27] A. Lenz, *Theoretical update of B-mixing and lifetimes*, `arXiv:1205.1444`. 23

[28] Heavy Flavor Averaging Group, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of early 2012*, `arXiv:1207.1158`, updated results and plots available at: `http://www.slac.stanford.edu/xorg/hfag/`. 23, 71

[29] R. Aaij *et al.*, *Observation of b^0_s-$\bar{B}^0_s$ mixing and measurement of mixing frequencies using semileptonic b decays*, Tech. Rep. arXiv:1308.1302. LHCB-PAPER-2013-036. CERN-PH-EP-2013-143, CERN, Geneva, Aug, 2013. 24

[30] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001. 26, 28, 34, 50, 58, 65

[31] LHCb collaboration, R. Aaij *et al.*, *Measurement of the B^0–$\overline{B}^0$ 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, arXiv:1210.6750. 28, 71

[32] M. Pivk and F. R. Le Diberder, *sPlot: a statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083. 29, 36

[33] https://twiki.cern.ch/twiki/bin/view/LHCb/TCK. 33

[34] LHCb collaboration, R. Aaij *et al.*, *Measurement of b hadron production fractions in 7 TeV pp collisions*, Phys. Rev. **D85** (2012) 032008, arXiv:1111.2357. 40

[35] C. Langenbruch and U. Uwer, *Measurement of the B_s^0 mixing phase in the decay $B_s^0 \to J/\psi\phi$ with the LHCb experiment*, PhD thesis, Heidelberg U., Heidelberg, 2011, Presented 09 Nov 2011. 47

[36] F. James and M. Roos, *Minuit: A system for function minimization and analysis of the parameter errors and correlation*, , Comput. Phys. Commun. **10**, 343-367. 47, 64

[37] M. Calvi and B. Khanji, *Tagging calibration with $B^0 \to \mu\ \nu_\mu$ background effects*, Presented in the LHCb Flavour Tagging Meeting, February, 2013. 47

[38] G. Punzi, *Comments on likelihood fits with variable resolution*, 2004. 49

[39] LHCb collaboration, *Measurement of Δm_s in the decay $B_s^0 \to D_s^-(K^+K^-\pi^-)\pi^+$ using opposite-side and same-side flavour tagging algorithms*, LHCb-CONF-2011-050. 69

[40] 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. 70, 73

Acknowledgements

Nur durch die zahlreiche Unterstützung vieler Menschen konnte diese Arbeit zustande kommen. Besonderer Dank gilt daher:

- Herr Prof. Dr. Ulrich Uwer, dem Betreuer dieser Arbeit. Er hatte stets konstruktive Ideen zur Verbesserung dieser Analyse.

- Herr Prof. Dr. André Schöning, der sich als Zweitkorrektor zur Verfügung stellt.

- Frau Prof. Dr. Stephanie Hansmann-Menzemer, die häufig mit Vorschlägen zum Fortschreiten der Analyse beitrug.

- Jeroen van Tilburg, Lucia Grillo und Sebastian Wandernoth, die mit mir im vergangenen Jahr intensiv zusammengearbeitet haben.

- Der gesamten LHCb Gruppe Heidelberg. Die Gruppentreffen waren immer sehr förderlich bei der Lösung von Problemstellungen. Viel wichtiger noch aber war die großartige freundschaftliche Atmosphäre in der Arbeitsgruppe, die das vergangene Jahr für mich nicht nur wissenschaftlich sondern vor allem auch menschlich besonders wertvoll machten.

- Der LHCb Kollaboration.

- Meiner Familie und meinen Freunden, die mich über meinen gesamten Werdegang hinweg unterstützt haben.

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 01.10.2013 .....