DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola-University of Heidelberg, Germany

for the degree of

doctor rerum naturalium

presented by

Dipl.-Phys. Sebastian Wandernoth

born in Ludwigshafen am Rhein, Germany

Oral examination: February 04th, 2014

Measurement of the B_s^0–$\overline{B}_s^0$ Oscillation Frequency at LHCb using $1\,\text{fb}^{-1}$ of data taken in 2011

Referees:

Prof. Dr. Stephanie Hansmann-Menzemer

Prof. Dr. Hans-Christian Schultz-Coulon

Abstract

In this thesis the measurement of the B_s^0–$\overline{B}_s^0$ oscillation frequency, Δm_s, in the decay $B_s^0 \to D_s^- \pi^+$ is presented. It is an important step to prove that the LHCb experiment is able to resolve the fast B_s^0–$\overline{B}_s^0$ oscillation and, thus, can perform precision measurements in the B_s^0 system to search for physics beyond the Standard Model. Examples for such searches are the measurement of the weak mixing phase, ϕ_s, in the decays $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to J/\psi\,\pi^+\pi^-$ [1] and the CKM-angle, γ, in the decay $B_s^0 \to D_s^\pm K^\mp$ [2]. For the measurement of Δm_s presented in this thesis a data set, accumulated by the LHCb experiment at the CERN LHC in 2011, is used. The data sample corresponds to an integrated luminosity of $1\,\text{fb}^{-1}$ and about 34,000 signal B_s^0 candidates are reconstructed with an average decay time resolution of 45 fs. To determine the production flavour of the B_s^0 candidates both opposite and same side kaon flavour tagging algorithms are used with effective tagging efficiencies $\varepsilon_{\text{eff}}^{\text{OST}} = 2.4 \pm 0.4\%$ and $\varepsilon_{\text{eff}}^{\text{SST}} = 1.2 \pm 0.3\%$. Both, the excellent decay time resolution and performance by the flavour tagging algorithms are prerequisits for a precision measurement of Δm_s. The oscillation frequency is measured to be $\Delta m_s = 17.768 \pm 0.023\ (\text{stat}) \pm 0.006\ (\text{syst})\ \text{ps}^{-1}$, which is the world's best measurement of this quantity.

Kurzfassung

In dieser Arbeit wird die Messung der Oszillationsfrequenz im B_s^0–$\overline{B}_s^0$ System Δm_s im Zerfall $B_s^0 \to D_s^- \pi^+$ vorgestellt. Diese Messung stellt eine wichtige Voraussetzung für die Suche nach Neuen Physik Phänomenen im B_s^0-mixing Sektor am LHCb Experiment dar, wie zum Beispiel die Messung der schwachen *CP*-verletzenden Phase ϕ_s in den Zerfällen $B_s^0 \to J/\psi\,\phi$ und $B_s^0 \to J/\psi\,\pi^+\pi^-$ [1] oder des CKM-Winkels γ im Zerfall $B_s^0 \to D_s^\pm K^\mp$ [2]. Die Messung der Oszillationsfrequenz Δm_s demonstriert, dass das LHCb Experiment in der Lage ist die schnelle B_s^0–$\overline{B}_s^0$ Oszillation aufzulösen und Präzisionsmessungen im B_s^0-System durchzuführen. Für die Messung der Oszillationsfrequenz wurde der Datensatz benutzt, der im Jahr 2011 vom LHCb Experiment gesammelt wurde. Dieser entspricht einer integrierten Luminosität von $1\,\text{fb}^{-1}$ und enthält ca. rekonstruierte 34.000 B_s^0-Signalkandidaten. Die mittlere Eigenzeitauflösung beträgt 45 fs. Um den Produktionsflavour der B_s^0-Mesonen zu bestimmen werden sowohl Opposite-Side- als auch Same-Side-Kaon-Flavourtaggingalgorithmen benutzt mit effektiven Taggingleistungen $\varepsilon_{\text{eff}}^{\text{OST}} = 2.4 \pm 0.4\%$ und $\varepsilon_{\text{eff}}^{\text{SST}} = 1.2 \pm 0.3\%$. Sowohl die exzellente Eigenzeitauflösung als auch die ausgezeichnete Leistung der Flavourtaggingalgorithmen sind Voraussetzungen für die Präzisionsmessung von Δm_s. Der gemessene Wert der Oszillationsfrequenz beträgt $\Delta m_s = 17.768 \pm 0.023\ (\text{stat}) \pm 0.006\ (\text{syst})\ \text{ps}^{-1}$, welches die weltbeste Messung dieser Größe darstellt.

Contents

CHAPTER 1

Introduction

The Standard Model of particle physics [3–5] describes the elementary particles which are the fundamental building blocks of matter and their interactions to the best of today's knowledge. It has been extensively tested over the last decades in laboratory experiments and the results agree well with the predictions made by the theory. However there are several experimental observations that are not answered in the Standard Model, such as:

Baryon asymmetry: Why does our universe consist mainly of matter and not in equal parts of matter and antimatter as predicted by the theory [6, 7]?

Dark Matter/Energy: What is the origin of Dark Matter and Dark Energy, which make up 27% and 68% of the universe, respectively [8–10]?

Additionally there are conceptual problems of the Standard Model, for example:

Gravity: Why can the gravitational force and the theory of General relativity not be included in the Standard Model [11, 12]?

Hierarchy problem: Why is the measured Higgs-Boson mass so low compared to the Planck scale [11, 13, 14]?

There are two conceptually different approaches to look for effects from physics beyond the Standard Model (BSM): direct and indirect searches. Direct searches aim at directly producing new heavy particles and are, thus, performed at the highest possible energies. This approach is currently mostly pursued by the ATLAS and CMS collaboration at the LHC.

In indirect searches, precision measurements of observables in quantum loop processes are performed. Any deviation from the theoretical predictions can be attributed to physics beyond the Standard Model. Especially accurate theoretical predictions can be made in the quark sector of the Standard Model, which therefore is an excellent place for indirect searches. Transitions in the quark sector are described by the so-called CKM-mechanism introduced by Kobayashi and Maskawa in 1973 [15] (see Section 2.2). Especially decays from b hadrons have been used to study the flavour sector of the Standard Model. b quarks can decay into all other quark species, with exception of the t quark, which gives analyses of b hadron decays access to a large variety of independent ways to measure CKM parameters.

Dedicated b-physics experiments include the "B-factories" BaBar and Belle which were e^+e^- accellerators with a center-of-mass energy of a $b\bar{b}$ resonance to produce large amounts of B mesons. These experiments were the key in establishing the CKM-mechanism which resulted in the Nobel Prize for Kobayashi and Maskawa in 2008. LHCb is the next-generation B-physics experiment which now exploits the B_s^0 system for indirect searches for physics beyond the Standard Model.

Most of the produced data samples at the "B-factories" were limited to the light B meson species (B^0, B^+). The heavier B_s^0 mesons are only produced in large amounts at hadron colliders like the Tevatron and especially now at the LHC. Despite its short running time, the LHCb experiment at the LHC had in 2011 already collected a large sample of B_s^0 mesons.

Neutral B mesons pose a particularly interesting probe of the Standard Model since these particles can oscillate into their antiparticles via quantum loop processes (B mixing, see Section 2.3) which are susceptible to BSM contributions. While the lighter B^0 mesons have been studied extensively at the B-factories time-dependent CP violation observables in the B_s^0 meson system like for example the weak mixing phase ϕ_s [1] and the CKM angle γ in the decay $B_s^0 \to D_s^{\pm} K^{\mp}$ [2] still leave room for BSM contributions. Both analyses must resolve the B_s^0–$\overline{B}_s^0$ oscillation with a frequency Δm_s which is 35 times larger than in the B^0 meson system. Therefore, the measurement of the oscillation frequency Δm_s which is presented in this thesis, is an important benchmark for the LHCb experiment to prove that the fast B_s^0 oscillation can be resolved and precision measurements in the B_s^0-mixing sector are possible at LHCb.

In 2006 the D0 collaboration gave a first two-sided limit on the B_s^0–$\overline{B}_s^0$ oscillation frequency, Δm_s, at $17\,\mathrm{ps}^{-1} < \Delta m_s < 21\,\mathrm{ps}^{-1}$ at a 90% confidence level [16]. Shortly afterwards the CDF collaboration published a measurement of $\Delta m_s = 17.77 \pm 0.10(\mathrm{stat}) \pm 0.07(\mathrm{syst})\,\mathrm{ps}^{-1}$ [17].

The analysis presented in this thesis describes the measurement of Δm_s using $B_s^0 \to D_s^- \pi^+$ decays at LHCb which is the "Golden Channel" to measure B_s^0

oscillations. Even with a small data sample corresponding to the first $37\,\text{pb}^{-1}$ of data taken at the LHC in 2010, the LHCb experiment has been able to give a comparable result [18] to the measurement by the CDF collaboration [17]. An update with $340\,\text{pb}^{-1}$ lead to a preliminary result [19] and the final measurement utilizing the full dataset taken by LHCb during 2011 corresponding to $1\,\text{fb}^{-1}$ integrated luminosity has been published in 2013 [20] and is described in detail in this thesis. In each round of the analysis it has been refined and optimized to achieve the best possible sensitivity for Δm_s leading to the world's best measurement of this quantity. The author of this thesis has been the main contributor to all these three analysis updates.

This thesis is organized in the following way: In Chapter 2, an outline of the theoretical background of flavour physics in the Standard Model is given and in Chapter 3 an overview of the LHCb experiment is presented with the individual subsystems and software framework. The analysis strategy is summarized in Chapter 4 and the signal selection for the decay $B_s^0 \to D_s^- \pi^+$ in Chapter 5. Chapters 6 and 7 address the main experimental challenges for this analysis, namely flavour tagging and the decay time resolution and acceptance. The two dimensions of the fit with their probability density functions are explained in Chapters 8 and 9. The validation of the fit procedure is briefly described in Chapter 10 and in Chapter 11 the systematic uncertainties on this measurement are summarized. Finally, Chapter 12 gives a short summary and conclusion.

CHAPTER 2

Theory

In this chapter the theoretical background of B_s^0–$\overline{B}_s^0$ mixing is provided. First, the Standard Model of particle physics (SM) is briefly summarized [3–5]. Then, a more detailed description of the quark flavour sector and the CKM formalism is given, the mixing of neutral B mesons is introduced and the formalism describing both mixing and decay is discussed. Finally, the theoretical prediction for the B_s^0–$\overline{B}_s^0$ oscillation frequency, Δm_s, is given.

2.1. The Standard Model of Particle Physics

The Standard Model of particle physics is a relativistic, renormalizable quantum field theory which combines the theory of strong interactions, Quantum Chromo Dynamics (QCD), with the electroweak theory by Glashow-Salam-Weinberg [3–5]. Thus it is a unified theory describing three of the four fundamental forces in nature, namely the electromagnetic force, the weak force and the strong force. The gravitational force has not yet been incorporated in the Standard Model. There is a tremendous amount of literature available on the Standard Model. The brief overview presented in this section is based on [11, 21, 22].

The fundamental particles that build up the known matter in the universe consist of spin 1/2 fields. In total there are 12 elementary fermion fields which are grouped according to the charges they carry. Six of the fermions carry color charge (quarks) which makes them able to interact via the strong interaction. The remaining six

Quarks

Generation	Flavour	Electric Charge	Mass
1st	u	2/3 e	$2.3^{+0.7}_{-0.5}$ MeV/c^2
	d	-1/3 e	$4.8^{+0.7}_{-0.3}$ MeV/c^2
2nd	c	2/3 e	1.275 ± 0.025 GeV/c^2
	s	-1/3 e	95 ± 5 MeV/c^2
3rd	t	2/3 e	$173.07 \pm 0.52 \pm 0.72$ GeV/c^2
	b	-1/3 e	4.18 ± 0.03 GeV/c^2

Table 2.1.: *Quarks of the Standard Model. Masses are taken from [23].*

Leptons

Generation	Flavour	Electric Charge	Mass
1st	e^-	$-e$	511 keV/c^2
	ν_e	0	< 2 eV/c^2
2nd	μ^-	$-e$	106 MeV/c^2
	ν_μ	0	< 2 eV/c^2
3rd	τ^-	$-e$	1776.82 ± 0.16 MeV/c^2
	ν_τ	0	< 2 eV/c^2

Table 2.2.: *Leptons of the Standard Model. Masses are taken from [23].*

particles that do not carry color charge are called leptons. The fermion fields are divided into three generations with the same characteristics except for their masses. Each generation includes one lepton that carries one unit of elementary electric charge (e^-, μ^-, τ^-) and one uncharged lepton, called neutrino (ν_e, ν_μ, ν_τ). Each quark generation consists of one up-type quark (u, c, t), with electric charge 2/3 and one down-type quark (d, s, b), with electric charge -1/3. The fermionic fields of the Standard Model are summarized in Tables 2.1 and 2.2.

The Standard Model Lagrangian is invariant under local gauge transformations of the symmetry group $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)}_Y$. This symmetry introduces the interactions between the fermionic fields mediated by spin-1 gauge bosons. The massless photon, γ, and gluons, g, are mediators for the electromagnetic force

Bosons

Boson	Spin	Mass	couples to
Photon γ	1	0	el. charge
Gluon g	1	0	color
W^+	1	$80.4\,\mathrm{GeV}/c^2$	weak isospin
W^-	1	$80.4\,\mathrm{GeV}/c^2$	weak isospin
Z^0	1	$91.2\,\mathrm{GeV}/c^2$	weak isospin & el. charge
H^0	0	$125.9\pm0.4\,\mathrm{GeV}/c^2$	mass

Table 2.3.: *Bosons of the Standard Model including gauge bosons mediating the forces and the Higgs boson. Masses are taken from [23].*

and strong force, respectively. The massive $W^\pm$ and Z^0 bosons mediate the weak interaction. The scalar Higgs boson H^0 is not a mediator of a fundamental force but the consequence of the spontaneous breaking of the electroweak symmetry, which will be described in Section 2.1.3. The fundamental bosons of the Standard Model are listed in Table 2.3.

2.1.1. Quantum Chromo Dynamics

Quantum Chromo Dynamics is the theory describing strong interactions. $\mathrm{SU(3)_C}$ is the gauge group of QCD, where the index C denotes the charge color which is carried only by quarks. The color charge can take the quantum numbers red, green and blue, as well as the corresponding anti-colors. The gauge bosons are the 8 generators of $\mathrm{SU(3)_C}$, which are called gluons. They carry color charge as well which results in three and four gluon self-interactions.

These gluon self-interactions give rise to a large dependence of the coupling constant of the strong interaction, α_s, on the transferred four-momentum q^2. The coupling becomes weak for large q^2, which is an effect that is called "asymptotic freedom". It becomes very strong for low values of q^2 which gives rise to an effect called "confinement". It is the reason why quarks exist only in bound states in nature, so-called hadrons. There are two known types of hadrons. Mesons consist of a quark–antiquark pair and baryons consist of three quarks or three antiquarks. The lightest baryons are the constituents of atomic nuclei: protons and neutrons. Due to the self-interaction of gluons the strong force has a very short range, even though gluons are massless.

2.1.2. Electroweak Sector of the Standard Model

$\mathrm{SU(2)_L} \otimes \mathrm{U(1)}_Y$ is the gauge group of the unified electroweak theory. It describes electromagnetic and weak interactions. The electroweak gauge bosons are the three W bosons W_i^μ and a boson B^μ. The charges to which the electroweak gauge bosons couple are the weak isospin, T, with third component T_3 and the hypercharge, $Y = Q - T_3$, with the electric charge Q.

The fermionic fields are divided into doublets of one up-type and the corresponding down-type quark, which carry weak isospin (left-handed). Similar doublets exist for charged leptons and the corresponding neutrinos. Singlets which have weak isospin $T = 0$ exist as well (right-handed) for all particles except neutrinos. For antiparticles the right-handed fields are weak isospin doublets and the left-handed fields are the singlets. The index L in $\mathrm{SU(2)_L}$ denotes that the W_i^μ bosons only couple to the left-handed particle doublets and right-handed antiparticle doublets of the weak isospin. B^μ bosons couple to particles carrying hypercharge, Y, independent of the weak isospin.

The $\mathrm{SU(2)_L} \otimes \mathrm{U(1)}_Y$ symmetry is broken by the Higgs-mechanism, which will be discussed in Section 2.1.3. One result of this is, that the physical mediators of the weak and electromagnetic force are not the W_i^μ and B^μ bosons, but linear combinations of those.

The electromagnetic interaction is mediated via the photon which is a linear combination of the W_3^μ and the B^μ bosons. It couples to the electric charge Q, but doesn't carry charge itself. This together with the fact that the photon is massless, implies that the range of the electromagnetic force is very large.

The weak interaction is divided into the charged and neutral currents. The charged current is mediated by the $W^\pm$ bosons which are linear combinations of the W_1^μ and W_2^μ bosons. The neutral current of the weak interaction is mediated by the Z^0 boson which is like the photon, a linear combination of the W_3^μ and the B^μ bosons. Since the mediators of the weak interaction are massive they are rather short ranged interactions.

Due to the fact that $W^\pm$ and Z^0 bosons couple differently to left- and right-handed particles, the weak interaction does not conserve Parity, P, which is the transformation of spatial inversion. The charged current of the weak interaction is also the only possible interaction in the Standard Model between quarks from different generations. Flavour changing neutral currents (FCNC) are forbidden in the Standard Model in tree-level processes.

2.1.3. Electroweak Symmetry Breaking

Experimental measurements have shown that the mediators of the weak interaction, the $W^{\pm}$ and Z^0 bosons, are massive. However separate mass terms for these particles in the Standard Model Lagrangian would violate local gauge invariance. Thus, another mechanism has to be responsible for the masses of these particles as well as the massive fermions. In the Standard Model this mechanism is the Higgs-mechanism which introduces an isospin doublet, $\boldsymbol{\phi}$, of two complex scalar fields with the potential

$$V(\boldsymbol{\phi}) = \mu^2 \boldsymbol{\phi}^\dagger \boldsymbol{\phi} + \lambda(\boldsymbol{\phi}^\dagger \boldsymbol{\phi}) \tag{2.1}$$

with $\mu^2 < 0$ and $\lambda > 0$.

This potential leads to a non-zero vacuum expectation value which spontaneously breaks the electroweak $\mathrm{SU(2)_L} \otimes \mathrm{U(1)}_Y$ symmetry down to the unbroken $\mathrm{U(1)}_Q$ electric charge symmetry. Thus, the photon is left massless. The massive gauge bosons acquire their mass via the covariant derivatives in the kinetic Higgs term of the Standard Model Lagrangian $(D^\mu \boldsymbol{\phi})^\dagger D_\mu \boldsymbol{\phi}$. As a consequence also the fermionic fields acquire mass via the Yukawa terms which will be discussed in more detail in Section 2.2.

The Higgs-mechanism aside from introducing masses to the elementary particles predicts one massive spin-0 particle, the Higgs-boson. Until recently the Higgs-boson was the only particle in the Standard Model that had not yet been observed. However, both the ATLAS [24] and CMS [25] collaborations published in 2012 the observation of a Higgs-like particle.

2.2. The Flavour sector of the Standard Model and CKM formalism

As mentioned in the previous section the fermionic fields in the Standard Model acquire mass through the so-called Yukawa terms in the Lagrangian. For the quark fields this can be written as

$$\mathcal{L}_\mathrm{Y} = -\frac{v}{\sqrt{2}} \left(\bar{d'}_\mathrm{L} Y_d d'_\mathrm{R} + \bar{u'}_\mathrm{L} Y_u u'_\mathrm{R} \right) + h.c. \quad , \tag{2.2}$$

where the Higgs field, $\boldsymbol{\phi}$, is replaced by its vacuum expectation value v. d' and u' denote the weak eigenstates of the down-type (d, s, b) and up-type (u, c, t) quarks. $\bar{d'}$ and $\bar{u'}$ are the weak eigenstates of the corresponding antiparticles.

The Yukawa matrices Y_d and Y_u have non-zero diagonal elements which results in the weak eigenstates not being identical to the physical mass eigenstates. To derive

the mass eigenstates, the Yukawa matrices are diagonalized to mass matrices M_d and M_u using unitary transformations V_{L}^d, V_{L}^u, V_{R}^d and V_{R}^u

$$M_d = \frac{v}{\sqrt{2}} V_{\mathrm{L}}^d Y_d (V_{\mathrm{R}}^d)^\dagger \qquad M_u = \frac{v}{\sqrt{2}} V_{\mathrm{L}}^u Y_u (V_{\mathrm{R}}^u)^\dagger. \tag{2.3}$$

These unitary matrices are hence absorbed into the quark fields transforming the weak eigenstates d' and u' into the mass eigenstates d and u.

$$q_L = V_{\mathrm{L}}^q q'_{\mathrm{L}} \ , \quad q_R = V_{\mathrm{R}}^q q'_{\mathrm{R}} \ , \quad \text{with } q = d, u \tag{2.4}$$

These transformations leave all parts of the Standard Model Lagrangian unchanged except for the term describing the charged current of the weak interaction since it is the only interaction combining up-type with down-type quarks.

$$\mathcal{L}_{\mathrm{CC}} = -\frac{g}{2\sqrt{2}} \left[\bar{u}'_{\mathrm{L}} W_\mu^+ \gamma^\mu (1-\gamma^5) d'_{\mathrm{L}} \ + \ \bar{d}'_{\mathrm{L}} W_\mu^- \gamma^\mu (1-\gamma^5) u'_{\mathrm{L}} \right] . \tag{2.5}$$

Here the transformation from the weak eigenstates to the mass eigenstates leads to an additional factor[1] $V_{\mathrm{L}}^u V_{\mathrm{L}}^{d\dagger}$ which is called the Cabbibo-Kobayashi-Maskawa (CKM) matrix [15]

$$V_{\mathrm{CKM}} = V_{\mathrm{L}}^u V_{\mathrm{L}}^{d\dagger}, \tag{2.6}$$

which transforms Equation 2.5 to

$$\mathcal{L}_{\mathrm{CC}} = -\frac{g}{2\sqrt{2}} \left[\bar{u}_{\mathrm{L}} W_\mu^+ \gamma^\mu (1-\gamma^5) V_{\mathrm{CKM}} d_{\mathrm{L}} \ + \ \bar{d}_{\mathrm{L}} W_\mu^- \gamma^\mu (1-\gamma^5) V_{\mathrm{CKM}}^\dagger u_{\mathrm{L}} \right] . \tag{2.7}$$

The CKM-matrix is a complex, unitary 3×3 matrix relating the left-handed down-type quarks with the left-handed up-type quarks

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} . \tag{2.8}$$

Since the matrix elements V_{ij} are complex numbers, this results in 18 free parameters. Due to V_{CKM} being unitary, nine of these parameters are fixed. Five additional parameters are absorbed in unobservable, relative phases between the quark fields. Thus, V_{CKM} can be parametrized by the four remaining parameters, using three Euler angles, θ_{12}, θ_{23}, and θ_{13} and one phase δ. Using the abbreviations $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$ the standard parametrization of the CKM-matrix is given by [23]

$$\begin{aligned} V_{\mathrm{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \end{aligned} \tag{2.9}$$

[1] In the case of the neutral current linking up-type quarks to up-type quarks this factor would be $V_{\mathrm{L}}^u V_{\mathrm{L}}^{u\dagger}$ which is equal to $\mathbb{1}$ due to the unitarity of V_{L}^u.

Another common parametrization of the CKM-matrix is the so-called Wolfenstein parametrization [26]. It illustrates better the hierarchy of the individual matrix elements. It uses the real parameters λ, A, ρ and η with the definitions

$$\begin{aligned} \lambda &= s_{12} \\ A\lambda^2 &= s_{23} \\ A\lambda^3(\rho + i\eta) &= s_{13}e^{i\delta}. \end{aligned} \tag{2.10}$$

The parameter $\lambda \approx 0.23$ is used as expansion parameter to estimate the size of the matrix elements. The parametrization up to order λ^3 is given by

$$V_{\mathrm{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.11}$$

The diagonal elements of the matrix are $\mathcal{O}(1)$ which define transitions between quarks within the same generation. The matrix elements describing transitions between the first two generations are suppressed by a factor λ, between the second and third generation by λ^2 and between the first and third generation by λ^3.

The parameter η describes the imaginary part of the CKM-matrix. It is also responsible for the effect of *CP* violation in the Standard Model. *CP* transforms processes for particles into those for antiparticles by applying a charge conjugation (C) and a parity transformation (P) which is a spatial inversion of all fields[2]. Applying the *CP* transformation to the charged current Lagrangian gives

$$\mathcal{L}_{\mathrm{CC}} = -\frac{g}{2\sqrt{2}} \left[\bar{d}_{\mathrm{L}} W^-_\mu \gamma^\mu (1 - \gamma^5) V^T_{\mathrm{CKM}} u_{\mathrm{L}} \; + \; \bar{u}_{\mathrm{L}} W^+_\mu \gamma^\mu (1 - \gamma^5) V^*_{\mathrm{CKM}} d_{\mathrm{L}} \right] . \tag{2.12}$$

which is identical to Equation 2.7, if $V_{\mathrm{CKM}} = V^*_{\mathrm{CKM}}$, *i.e.* if all components of V_{CKM} were real. Therefore, the non-zero parameter η is the unique source of *CP* violation in the Standard Model.

Due to the unitarity condition of the CKM-matrix ($V^\dagger_{\mathrm{CKM}} V_{\mathrm{CKM}} = \mathbb{1}$), multiplying the third row with the first column gives the relation

$$V_{ud}V^*_{ub} + V_{cd}V^*_{cb} + V_{td}V^*_{tb} = 0. \tag{2.13}$$

Five similar relations can be derived, but from now on this relation is used as an example to illustrate the derivation of CKM-matrix elements. Since these are complex numbers, Equation 2.13 can be interpreted as a triangle in the complex plane. The corresponding angles of the triangle are

$$\alpha \equiv \arg\left(-\frac{V_{td}V^*_{tb}}{V_{ud}V^*_{ub}}\right) \quad \beta \equiv \arg\left(-\frac{V_{cd}V^*_{cb}}{V_{td}V^*_{tb}}\right) \quad \gamma \equiv \arg\left(-\frac{V_{ud}V^*_{ub}}{V_{cd}V^*_{cb}}\right). \tag{2.14}$$

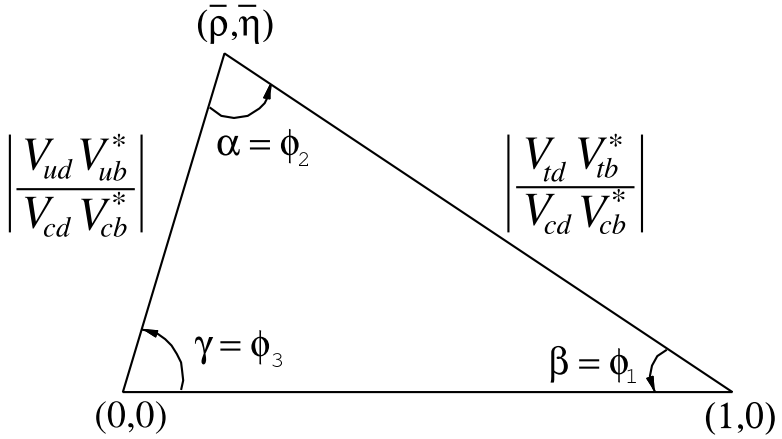


Figure 2.1.: *Unitarity triangle from the first row and third column of the CKM matrix. Taken from [23]*

Figure 2.1 shows this triangle with one side normalized to $V_{cd}V_{cb}^*$. The coordinates of the apex of the triangle are $\overline{\rho}$ and $\overline{\eta}$, which are related to the Wolfenstein parameters ρ and η as

$$\overline{\rho} = \rho\left(1 - \frac{\lambda^2}{2}\right) \qquad \text{and} \qquad \overline{\eta} = \eta\left(1 - \frac{\lambda^2}{2}\right) . \tag{2.15}$$

The parameters of the CKM-matrix are not predicted by the Standard Model. However, they are experimentally accessible. Since all sides as well as the angles of the triangle can be measured, it is overconstrained. Figure 2.2 shows the current experimental status of the measurement of the apex of this unitarity triangle. It can be seen that all measurements agree well and the apex is well determined. The Wolfenstein parameters are measured to be

$$\begin{aligned} \lambda &= 0.22535 \pm 0.00065 , & A &= 0.817 \pm 0.015 \\ \overline{\rho} &= 0.136 \pm 0.018 , & \overline{\eta} &= 0.348 \pm 0.014 \end{aligned} \tag{2.16}$$

[2]To transform particles into their antiparticles the parity transformation must be included because $W^{\pm}$ and Z^0 bosons couple similarly to left-handed particles and right-handed antiparticles

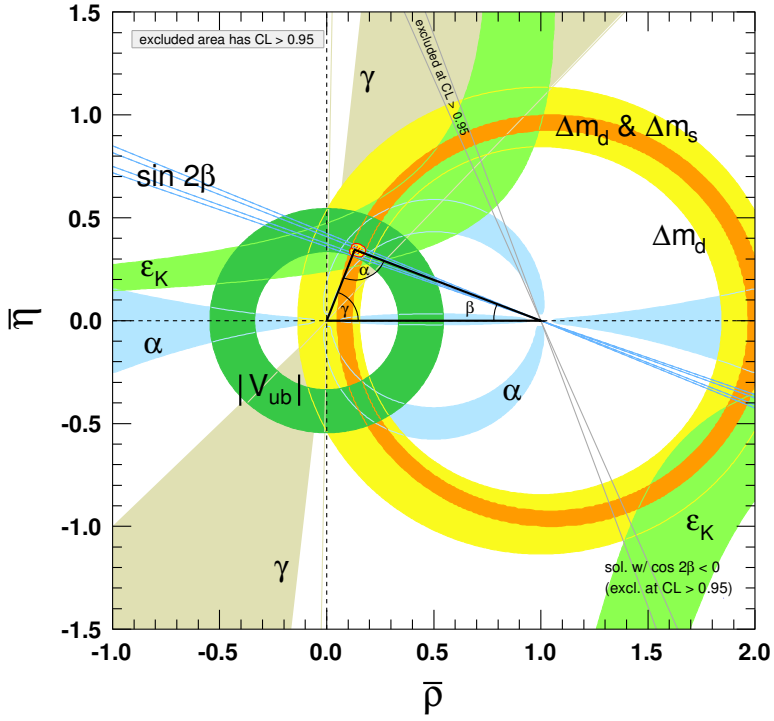


Figure 2.2.: *Experimental status of the unitarity triangle defined in Equation 2.13. Shaded areas correspond to 95% Confidence Intervals of the single measurements. Taken from [23]*

2.3. Neutral B meson mixing

This brief overview of neutral B meson mixing is based on [27–29]. In these references more detailed introductions into the mixing phenomenology are given.

2.3.1. Introduction to neutral B meson mixing

As mentioned in Section 2.1.1 quarks only exist bound in hadrons. Neutral B mesons contain one b quark and one d or s quark respectively. The phenomenological description of mixing applies to both neutral B mesons in the same way. Thus, they are denoted as B_q^0 where q stands for either d or s.

The process of B meson mixing describes transitions between the flavour eigenstates $B_q^0 = |\bar{b}q\rangle$ to $\overline{B}_q^0 = |b\bar{q}\rangle$ and vice versa. This requires transitions of the form $b \to d, s$. However, as mentioned in Section 2.1.2, flavour changing neutral currents (FCNC) are forbidden at tree-level in the Standard Model, but allowed in higher order processes, like loop diagrams (see Figure 2.3).

Therefore, the processes for B meson mixing in the Standard Model are given by

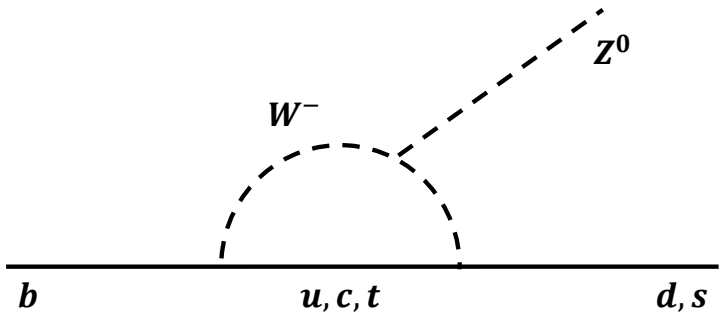


Figure 2.3.: *Feynman diagram of the flavour changing neutral current (FCNC) in a loop process, which is allowed in the Standard Model.*

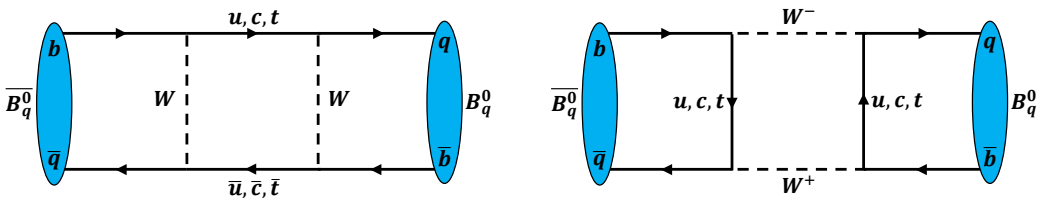


Figure 2.4.: *Dominant Feynman diagrams for neutral B meson mixing in the Standard Model. Here, q stands for either an s quark or a d quark.*

the exchange of two $W^{\pm}$ bosons. Figure 2.4 shows the corresponding Feynman diagrams. For the up-type quarks (u, c, t) in these so-called box-diagrams, the top quark contribution is by far the largest. This is due to the GIM suppression (Glashow, Iliopoulos, Maiani) [30]. The amplitude of the sum of the box diagrams with the three up-type quarks is proportional to the term

$$\mathcal{A} \propto m_u^2 V_{uq} V_{ub}^* + m_c^2 V_{cq} V_{cb}^* + m_t^2 V_{tq} V_{tb}^* \tag{2.17}$$

If all quark masses were identical the three processes would completely cancel because the amplitude would be proportional to

$$V_{uq} V_{ub}^* + V_{cq} V_{cb}^* + V_{tq} V_{tb}^* = 0. \tag{2.18}$$

which for $q = d$ and $q = s$ corresponds to unitarity relations of the CKM-matrix. In comparison to the top quark mass, the charm- and up quark masses are very small and nearly identical ($m_u \approx m_c \ll m_t$). With this assumption Equation 2.17 becomes

$$\mathcal{A} \propto m_u^2 \underbrace{(V_{uq} V_{ub}^* + V_{cq} V_{cb}^* + V_{tq} V_{tb}^*)}_{=0} + \underbrace{(m_t^2 - m_u^2)}_{=m_t^2} V_{tb} V_{ts}^* \tag{2.19}$$

Since B mixing processes are only allowed in loop processes, they are suppressed in the Standard Model by the coupling constants at the vertices. An additional suppression is given by the CKM matrix element $|V_{ts}|^2$ in case of B_s^0–$\overline{B}_s^0$ mixing and $|V_{td}|^2$ in case of B^0–$\overline{B}^0$ mixing. This results in an about thirty times larger mixing frequency in case of B_s^0 mesons compared to B^0 mesons.

2.3.2. Phenomenology of neutral B meson mixing

Defining the flavour eigenstates $|B_q^0\rangle = |\bar{b}q\rangle$ and $|\overline{B}_q^0\rangle = |b\bar{q}\rangle$ for neutral B mesons, they transform under the CP operation as

$$CP|B_q^0\rangle = -|\overline{B}_q^0\rangle \quad \text{and} \quad CP|\overline{B}_q^0\rangle = -|B_q^0\rangle \tag{2.20}$$

The time development of these states is given by the phenomenological Schroedinger equation for 2-state systems

$$\begin{aligned} i\frac{\partial}{\partial t}\begin{pmatrix} |B_q^0\rangle \\ |\overline{B}_q^0\rangle \end{pmatrix} &= \left(\boldsymbol{M} - i\frac{\boldsymbol{\Gamma}}{2}\right)\begin{pmatrix} |B_q^0\rangle \\ |\overline{B}_q^0\rangle \end{pmatrix} \\ &= \begin{pmatrix} M_{11} - i\frac{\Gamma_{11}}{2} & M_{12} - i\frac{\Gamma_{12}}{2} \\ M_{21} - i\frac{\Gamma_{21}}{2} & M_{22} - i\frac{\Gamma_{22}}{2} \end{pmatrix}\begin{pmatrix} |B_q^0\rangle \\ |\overline{B}_q^0\rangle \end{pmatrix} \end{aligned} \tag{2.21}$$

with two hermitian matrices, the mass matrix, $\boldsymbol{M}$, and the decay width matrix, $\boldsymbol{\Gamma}$. The mass matrix describes the B mixing processes introduced in Section 2.3.1.

The off-diagonal elements of the decay width matrix describe long distance effects in which the quarks in the "box" of Figure 2.4 are produced on-shell.

Since $\boldsymbol{M}$ and $\boldsymbol{\Gamma}$ are hermitian, the off-diagonal elements are complex conjugates ($M_{21} = M_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$). Furthermore, CPT invariance gives $M_{11} = M_{22} = M$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$ so that Equation 2.21 becomes

$$i\frac{\partial}{\partial t}\begin{pmatrix}|B_q^0\rangle\\|\overline{B}_q^0\rangle\end{pmatrix} = \begin{pmatrix}M - i\frac{\Gamma}{2} & M_{12} - i\frac{\Gamma_{12}}{2}\\ M_{12}^* - i\frac{\Gamma_{12}^*}{2} & M - i\frac{\Gamma}{2}\end{pmatrix}\begin{pmatrix}|B_q^0\rangle\\|\overline{B}_q^0\rangle\end{pmatrix} \tag{2.22}$$

Diagonalizing the matrix leads to the mass eigenstates $|B_\mathrm{L}\rangle$ and $|B_\mathrm{H}\rangle$ which are linear combinations of the flavour eigenstates $|B_q^0\rangle$ and $|\overline{B}_q^0\rangle$

$$\begin{aligned}|B_\mathrm{L}\rangle &= p|B_q^0\rangle + q|\overline{B}_q^0\rangle\ ,\\ |B_\mathrm{H}\rangle &= p|B_q^0\rangle - q|\overline{B}_q^0\rangle\ ,\end{aligned} \tag{2.23}$$

with $|p|^2 + |q|^2 = 1$. The indices L and H denote the lighter and heavier mass eigenstate, respectively. After diagonalization the mass eigenvalues are given by $(M_\mathrm{L} - i\frac{\Gamma_\mathrm{L}}{2})$ and $(M_\mathrm{H} - i\frac{\Gamma_\mathrm{H}}{2})$ and the time evolution is given by the relations

$$\begin{aligned}|B_\mathrm{L}(t)\rangle &= e^{-(iM_\mathrm{L} + \frac{\Gamma_\mathrm{L}}{2})t}|B_\mathrm{L}\rangle\ ,\\ |B_\mathrm{H}(t)\rangle &= e^{-(iM_\mathrm{H} + \frac{\Gamma_\mathrm{H}}{2})t}|B_\mathrm{H}\rangle\ ,\end{aligned} \tag{2.24}$$

with M_L, M_H, Γ_L, and Γ_H the masses and decay widths of the mass eigenstates. These are related to the parameters in the Hamiltonian in Equation 2.22 by

$$\begin{aligned}M_{\mathrm{L/H}} - i\frac{\Gamma_{\mathrm{L/H}}}{2} =& M - i\frac{\Gamma}{2} \mp \sqrt{\left(M_{12} - i\frac{\Gamma_{12}}{2}\right)\left(M_{12}^* - i\frac{\Gamma_{12}^*}{2}\right)}\ ,\\ =& M - i\frac{\Gamma}{2}\\ &\mp \sqrt{|M_{12}|^2 - 1/4|\Gamma_{12}|^2 - i|M_{12}||\Gamma_{12}|\cos(\phi_{12})}\ ,\end{aligned} \tag{2.25}$$

with the phase $\phi_{12} = \arg -\frac{\Gamma_{12}}{M_{12}}$. Furthermore, the masses and decay widths in the diagonal elements in the Hamiltonian in Equation 2.22 are the arithmetic means of the masses and decay widths of the mass eigenstates:

$$\begin{aligned}M &= \frac{M_\mathrm{H} + M_\mathrm{L}}{2}\ ,\\ \Gamma &= \frac{\Gamma_\mathrm{H} + \Gamma_\mathrm{L}}{2}\ .\end{aligned} \tag{2.26}$$

Using Equation 2.25 and the definitions

$$\Delta m = M_\mathrm{H} - M_\mathrm{L} \qquad \text{and} \qquad \Delta\Gamma = \Gamma_\mathrm{L} - \Gamma_\mathrm{H} \tag{2.27}$$

for the mass difference Δm and the decay width difference $\Delta\Gamma$ of the light and heavy mass eigenstates, the following realtions can be derived

$$\begin{aligned} \Delta m^2 - \frac{1}{4}\Delta\Gamma^2 &= 4|M_{12}|^2 - |\Gamma_{12}|^2 \,, \\ \Delta m \Delta\Gamma &= -4|M_{12}||\Gamma_{12}|\cos(\phi_{12}) \,. \end{aligned} \tag{2.28}$$

It has been shown experimentally that the short-distance effects in the diagrams in Figure 2.4 are dominant over the long distance effects of on-shell particles ($|M_{12}| \gg |\Gamma_{12}|$) which translates to $\Delta m \gg \Delta\Gamma$. This simplifies Equations 2.28 to

$$\Delta m = 2|M_{12}| \,, \tag{2.29}$$

$$\Delta\Gamma = -2|\Gamma_{12}|\cos(\phi_{12}) \,. \tag{2.30}$$

Furthermore, the coefficients of the flavour eigenstates q and p are related to the matrix elements through

$$\begin{aligned} \frac{q}{p} &= -\sqrt{\frac{M_{12}^* - i\frac{\Gamma_{12}^*}{2}}{M_{12} - i\frac{\Gamma_{12}}{2}}} \,, \\ &= -\sqrt{e^{-2i\phi_M}\frac{1 - i\frac{|\Gamma_{12}|}{2|M_{12}|}e^{-i(\phi_{12})}}{1 - i\frac{|\Gamma_{12}|}{2|M_{12}|}e^{+i(\phi_{12})}}} \,, \end{aligned} \tag{2.31}$$

with the phase $\phi_M = \arg M_{12}$. Expanding this in $\frac{|\Gamma_{12}|}{|M_{12}|}$ leads to

$$\frac{q}{p} = -e^{-i\phi_M}\left[1 - \frac{1}{2}\sin(\phi_{12})\frac{|\Gamma_{12}|}{|M_{12}|} + \mathcal{O}\left(\frac{\Gamma_{12}^2}{M_{12}^2}\right)\right] \,. \tag{2.32}$$

The time evolution of the flavour eigenstates using Equation 2.23 is given by

$$\begin{aligned} |B_q^0(t)\rangle &= \frac{1}{2p}\left(|B_{\mathrm{L}}(t)\rangle + |B_{\mathrm{H}}(t)\rangle\right) \,, \\ |\overline{B}_q^0(t)\rangle &= \frac{1}{2q}\left(|B_{\mathrm{L}}(t)\rangle - |B_{\mathrm{H}}(t)\rangle\right) \,. \end{aligned} \tag{2.33}$$

Using the identities for the mass eigenstates from Equations 2.23 and 2.24 this leads to

$$\begin{aligned} |B_q^0(t)\rangle &= g_+(t)|B_q^0\rangle + \frac{q}{p}g_-(t)|\overline{B}_q^0\rangle \,, \\ |\overline{B}_q^0(t)\rangle &= \frac{p}{q}g_-(t)|B_q^0\rangle + g_+(t)|\overline{B}_q^0\rangle \,. \end{aligned} \tag{2.34}$$

with the coefficients

$$g_\pm(t) = \frac{1}{2}\left(e^{-iM_{\mathrm{L}}t}e^{-\frac{\Gamma_{\mathrm{L}}}{2}t} \pm e^{-iM_{\mathrm{H}}t}e^{-\frac{\Gamma_{\mathrm{H}}}{2}t}\right) \tag{2.35}$$

For later calculations, it is useful to calculate the following identities for these coefficients

$$\begin{aligned}
|g_+(t)|^2 &= \frac{1}{2}e^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) + \cos\left(\Delta m t\right)\right) \\
|g_-(t)|^2 &= \frac{1}{2}e^{-\Gamma t}\left(\cosh\left(\frac{\Delta\Gamma}{2}t\right) - \cos\left(\Delta m t\right)\right) \\
g_+(t)g_-^*(t) &= \frac{1}{2}e^{-\Gamma t}\left(-\sinh\left(\frac{\Delta\Gamma}{2}t\right) - \sin\left(\Delta m t\right)\right) \\
g_+^*(t)g_-(t) &= \frac{1}{2}e^{-\Gamma t}\left(-\sinh\left(\frac{\Delta\Gamma}{2}t\right) + \sin\left(\Delta m t\right)\right)
\end{aligned} \tag{2.36}$$

Looking at decays of B_q^0 and $\overline{B}_q^0$ mesons to a final state f and the CP conjugate $\overline{f}$, the decay amplitudes are given by

$$\begin{aligned}
A_f &= \langle f|\mathcal{H}|B_q^0\rangle \\
\overline{A}_f &= \langle f|\mathcal{H}|\overline{B}_q^0\rangle \\
A_{\bar{f}} &= \langle \overline{f}|\mathcal{H}|B_q^0\rangle \\
\overline{A}_{\bar{f}} &= \langle \overline{f}|\mathcal{H}|\overline{B}_q^0\rangle \ .
\end{aligned} \tag{2.37}$$

For convenience the CP violation parameter λ_f is introduced

$$\lambda_f = \frac{q}{p}\frac{\overline{A}_f}{A_f} \ . \tag{2.38}$$

The time-dependent decay rates for the decay of meson that was produced as a B_q^0 at time $t = 0$ into the final state f is given by

$$\begin{aligned}
\frac{d\Gamma(B_q^0 \to f)}{dt} &= \left|\langle f|B_q^0(t)\rangle\right|^2 \\
&= \left|g_+(t)A_f + \frac{q}{p}g_-(t)\overline{A}_f\right|^2 \\
&= |A_f|^2\left(|g_+(t)|^2 + |\lambda_f|^2|g_-(t)|^2 + \lambda_f^* g_+(t)g_-^*(t) + \lambda_f g_+^*(t)g_-(t)\right) \\
&= \frac{1}{2}|A_f|^2 e^{-\Gamma t}\Bigg[\left(1 + |\lambda_f|^2\right)\cosh\left(\frac{\Delta\Gamma}{2}t\right) + \left(1 - |\lambda_f|^2\right)\cos\left(\Delta m t\right) \\
&\qquad - 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re(\lambda_f) - 2\sin\left(\Delta m t\right)\Im(\lambda_f)\Bigg] \ .
\end{aligned} \tag{2.39}$$

In the last step the identities of Equation 2.36 are used.

Similarly the decay rate for a meson that was a $\overline{B}^0_q$ at production into the final state f is given by

$$\begin{aligned}\frac{d\Gamma(\overline{B}^0_q \to f)}{dt} &= \left|\langle f|\overline{B}^0_q(t)\rangle\right|^2 \\ &= \left|\frac{p}{q}g_-(t)A_f + g_+(t)\overline{A}_f\right|^2 \\ &= |A_f|^2\left|\frac{p}{q}\right|^2\left(|g_-(t)|^2 + |\lambda_f|^2|g_+(t)|^2 + \lambda_f^* g_+^*(t)g_-(t) + \lambda_f g_+(t)g_-^*(t)\right) \\ &= \frac{1}{2}\left|\frac{p}{q}\right|^2|A_f|^2 e^{-\Gamma t}\Bigg[\left(1+|\lambda_f|^2\right)\cosh\left(\frac{\Delta\Gamma}{2}t\right) - \left(1-|\lambda_f|^2\right)\cos\left(\Delta m t\right) \\ &\qquad - 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re(\lambda_f) + 2\sin\left(\Delta m t\right)\Im(\lambda_f)\Bigg].\end{aligned} \tag{2.40}$$

The expressions for the decays to the CP conjugate final state $\overline{f}$ are identical to Equations 2.39 and 2.40 when substituting the index f by $\overline{f}$, with $\lambda_{\overline{f}} = \frac{q}{p}\frac{A_{\overline{f}}}{\overline{A}_{\overline{f}}}$.

The decay used in this analysis of the measurement of Δm_s is $B^0_s \to D^-_s\pi^+$. $D^-_s\pi^+$ is a flavour specific final state, meaning that only decays $B^0_s \to D^-_s\pi^+$ and $\overline{B}^0_s \to D^+_s\pi^-$ are allowed. This results in $A_{\overline{f}} = \overline{A}_f = 0$ for this decay channel. Additionally, since the decay is completely dominated by the tree-level process (see Figure 2.5) and there is no other sizeable contribution from other processes, which could interfere to cause CP violation[3]. This leads to $\left|\frac{A_f}{\overline{A}_{\overline{f}}}\right| = 1$.

Additionally, in the Standard Model the CP violation due to different probabilities for $(B^0_s \to \overline{B}^0_s)$ and $(\overline{B}^0_s \to B^0_s)$ is expected to be negligible[4] [31, 32]. The reason for this is that the only diagrams contributing to neutral B meson mixing are shown in Figure 2.4. The only possible interference that could cause CP violation would be due to the individual up-type quarks in the box. Due to the GIM mechanism described in Equation 2.19, the contribution from u and c quarks are highly suppressed. Therefore, also $\left|\frac{q}{p}\right| = 1$ can be assumed. This simplifies the differential decay rates considerably

$$\begin{aligned}\frac{d\Gamma(B^0_s \to D^-_s\pi^+)}{dt} &= \frac{d\Gamma(\overline{B}^0_s \to D^+_s\pi^-)}{dt} \\ &= \frac{1}{2}|A_f|^2 e^{-\Gamma t}\left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) + \cos\left(\Delta m_s t\right)\right],\end{aligned} \tag{2.41}$$

[3] In the literature this effect is referred to as *CP violation in decay*

[4] In the literature this effect is referred to as *CP violation in mixing*

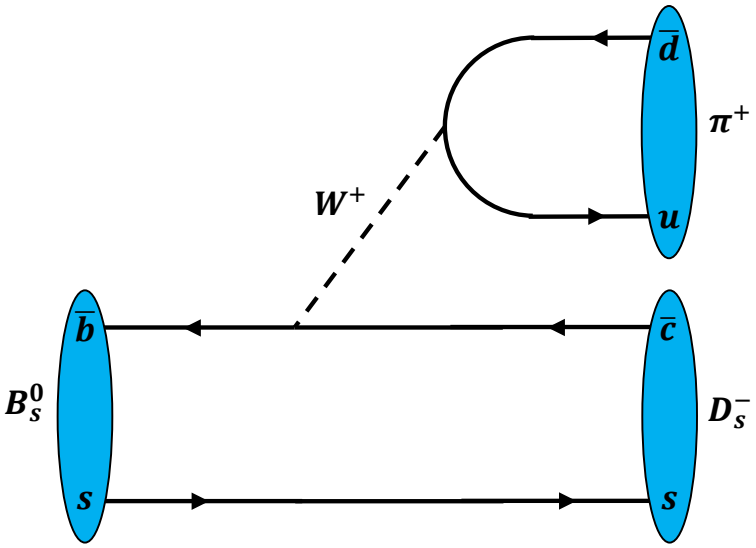


Figure 2.5.: *Feynman diagram for the decay $B_s^0 \to D_s^- \pi^+$ that is used in this analysis. The only sizeable contribution to this decay is the shown tree-level diagram.*

for decays where the production flavour and decay flavour of the B_s^0 meson are identical. Simlarly

$$\begin{aligned} \frac{d\Gamma(\overline{B}_s^0 \to D_s^- \pi^+)}{dt} &= \frac{d\Gamma(B_s^0 \to D_s^+ \pi^-)}{dt} \\ &= \frac{1}{2} |A_f|^2 e^{-\Gamma t} \left[\cosh\left(\frac{\Delta\Gamma_s}{2} t\right) - \cos\left(\Delta m_s t\right) \right] \end{aligned} \tag{2.42}$$

give the differential decay rate for decays where the flavour at decay is different from the flavour at production of the B_s^0 meson.

2.4. Theoretical prediction of Δm_s

The theoretical predictions for Δm_s have very large uncertainties. The reason for this is that M_{12} is proportional to

$$M_{12} \propto (V_{tq}^* V_{tb})^2 B_{B_q^0} f_{B_q^0}^2 \tag{2.43}$$

with the decay constant $f_{B_q^0}$ and the so-called bag parameter $B_{B_q^0}$ which parametrize the non-perturbative matrix element. Latice QCD calculations give [33]

$$f_{B_s^0} = 231 \pm 15\,\text{MeV} \qquad B_{B_s^0} = 0.841 \pm 0.020 \tag{2.44}$$

which has leads to a relative uncertainty on the factor $\sqrt{B_{B_q^0}} f_{B_q^0}$ of about 14%. This leads to a theoretical prediction of Δm_s of [34]

$$\Delta m_s = 17.3 \pm 2.6\,\mathrm{ps}^{-1} \tag{2.45}$$

It has been found useful to use the ratio $\frac{\Delta m_d}{\Delta m_s}$ to constrain the CKM unitarity triangle in Figure 2.2. In this ratio many hadronic uncertainties cancel and the theoretical prediction is proportional to the factor ξ

$$\xi = \frac{f_{B_s^0}\sqrt{B_{B_s^0}}}{f_{B^0}\sqrt{B_{B^0}}} = 1.22 \pm 0.04 \tag{2.46}$$

which has a relative uncertainty of about 3% [33]. Other computations give similar results also with a few percent relative uncertainty on ξ [35, 36].

Due to the large theoretical uncertainties on Δm_s, this analysis is not suited to look for contributions from physics beyond the Standard Model directly, until the lattice QCD calculations are improved considerably. However, the measurement is an important benchmark measurement proving that very precise measurements in the B_s^0 mixing system can be performed at the LHCb experiment. Furthermore, the measured value of Δm_s is used as an input for the measurement of other observables in the B_s^0 system, like the weak mixing phase ϕ_s in the decays $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to J/\psi\,\pi^+\pi^-$ which has significantly smaller theoretical uncertainties or the measurement of decay time-dependent *CP* observables in the decay $B_s^0 \to D_s^{\pm} K^{\mp}$ to measure the CKM angle γ which is experimentally well constraint by the measurements of the apex of the unitarity triangle in Figure 2.2.

CHAPTER 3

The LHCb Experiment

The LHCb (**L**arge **H**adron **C**ollider **b**eauty) experiment is one of the four large experiments at the Large Hadron Collider (LHC) at CERN. Its design is dedicated to perform precision measurements of *CP* violation in *b*- and *c*-hadron decays, as well as searches for rare B meson decays.

In the following chapter the running conditions of the LHC and the production mechanism of b quarks at the LHC are briefly discussed. The mode of operation and technical details of the LHCb detector and its subsystems are summarized. In the end, the trigger system of LHCb are described and the measured and simulated data sets used in this analysis are discussed.

3.1. The Large Hadron Collider

The Large Hadron Collider is a proton-proton collider (pp)[1] located at CERN in Geneva. The proton beams are accelerated in a 27 km long underground tunnel under the Swiss-French border and then brought to collision at four interaction points. At these collision points the four major LHC experiments are located: ATLAS (**A** **T**oriodal **L**HC **A**pparatu**S**), CMS (**C**ompact **M**uon **S**olenoid), LHCb and ALICE (**A** **L**arge **I**on **C**ollider **E**xperiment) (see Figure 3.1).

ATLAS and CMS are multi-purpose detectors designed for the search of the Higgs

[1] For a short period, the LHC also collided lead ions ($PbPb$).

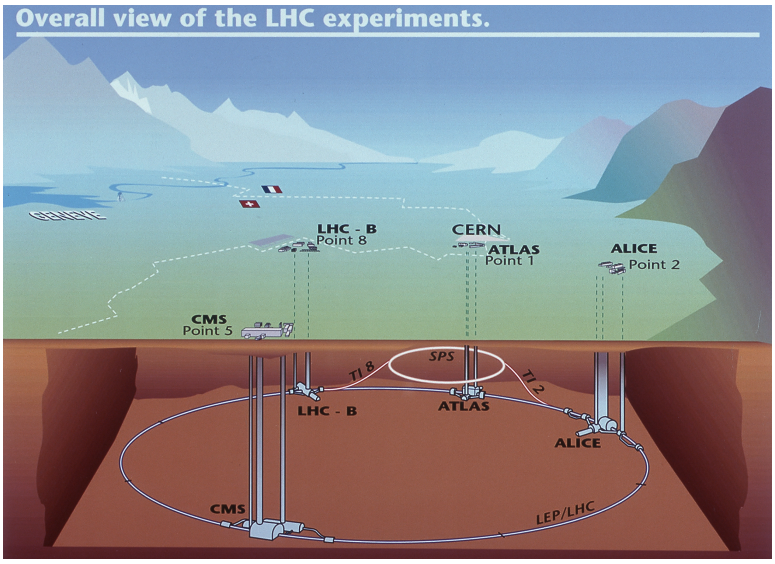


Figure 3.1.: *Schematic view of the Large Hadron Collider with the four major experiments: ATLAS, CMS, LHCb and ALICE. Taken from [37]*

boson and new heavy particles produced at high energies. LHCb is a dedicated b- and c-physics experiment and ALICE is specialized in heavy-ion physics.

The design center of mass energy for pp collisions at the LHC is $\sqrt{s} = 14\,\text{TeV}$ with a design luminosity of $\mathcal{L} = 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. The proton beams are divided in 2808 bunches with $\sim 10^{11}$ protons each. The collisions of the bunches occur with a spacing of 25 ns which leads to an interaction rate of 40 MHz.

In 2011, during the data-taking period for the sample used in this analysis, the LHC was operated at a center of mass energy of $\sqrt{\text{s}} = 7\,\text{TeV}$. The LHCb experiment operates at a lower instantaneous luminosity than the maximum that the LHC can deliver because otherwise the particle density in the LHCb detector would be too large. For this purpose the proton beams are slightly defocussed before collision. This has the advantage that the instantaneous luminosity can be held constant over the whole data taking process even though the proton beams constantly loose their energy. In 2011 the LHCb experiment was operated with an instantaneous luminosity of $\mathcal{L} = 4 \cdot 10^{32}\,\text{cm}^{-2}\,\text{s}^{-1}$.

The total integrated luminosity recorded in 2011 by the LHCb experiment amounts to $1\,\text{fb}^{-1}$.

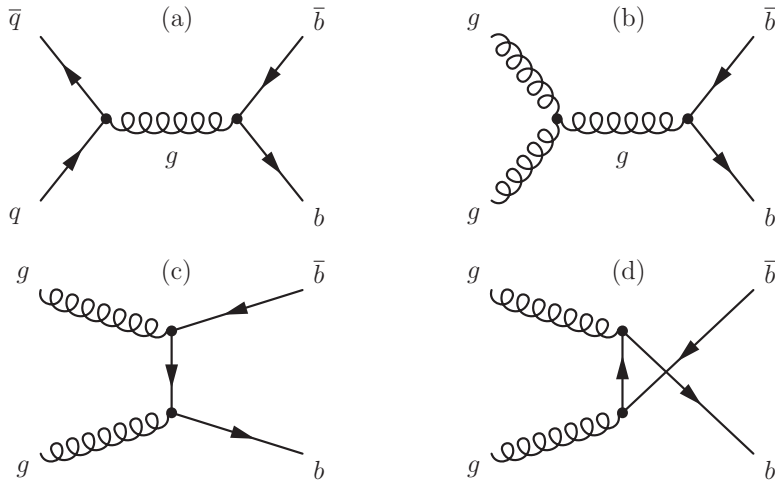


Figure 3.2.: *Feynman diagrams of the dominant processes for b quark production at the LHC. Taken from [38]*

3.2. Production of b hadrons at the LHC

In collisions at the LHC, b quarks are predominantly produced in $b\bar{b}$ pairs through gluon fusion or quark–antiquark annihilation (see Figure 3.2). At LHC energies the parton density functions of the proton are dominated by gluons. Thus, the gluon fusion processes in Figure 3.2 are dominant.

Gluons have a broad momentum spectrum within the protons. Since the $b\bar{b}$ production threshold is small compared to the center of mass energy, it is very likely that the $b\bar{b}$ pair is produced by gluons with very different momenta. This results in the $b\bar{b}$ pair being highly boosted in the forward or backward direction along the beam line. Figure 3.3 shows the polar angle distribution for the b- and $\bar{b}$ quark. It can be seen that they are mostly produced in the same direction close to the beam line. Therefore, for the LHCb experiment the geometry of a single-arm forward spectrometer is chosen. The geometrical acceptance of the LHCb detector in Figure 3.3 is marked in red. In total $\sim 25\%$ of all produced $b\bar{b}$ pairs end up in the LHCb detector.

The total $b\bar{b}$ cross section in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ has been measured at LHCb using semileptonic b hadron decays. The result is [40]

$$\sigma(pp \to b\bar{b}X) = 284 \pm 20 \pm 49\,\mu\text{b}\,. \tag{3.1}$$

In 2011 an integrated luminosity of $\mathcal{L}_{\text{int}} = 1\,\text{fb}^{-1}$ was recorded which leads to a total of

$$\sigma(pp \to b\bar{b}X) \cdot \mathcal{L}_{\text{int}} \cdot 25\% \approx 7 \cdot 10^{10} \tag{3.2}$$

$b\bar{b}$ pairs produced within the LHCb detector acceptance.

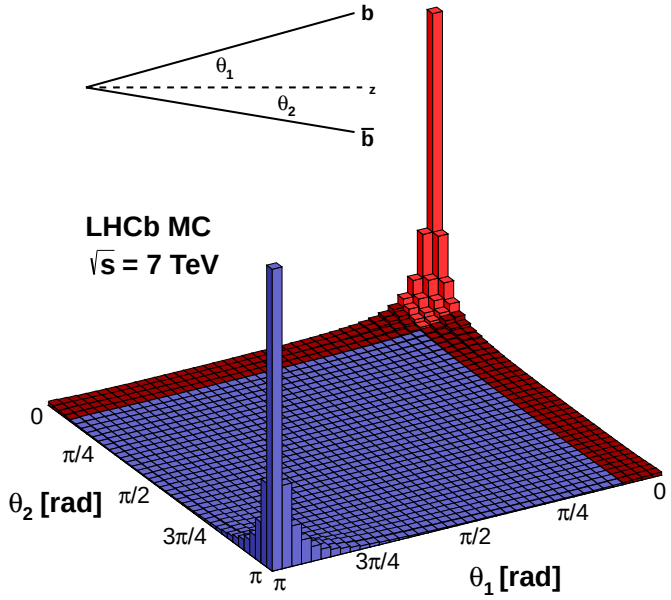


Figure 3.3.: *Simulated distribution of the polar angles θ_1 and θ_2 of the produced $b\bar{b}$ pair. The LHCb geometrical acceptance is marked in red. Taken from [39]*

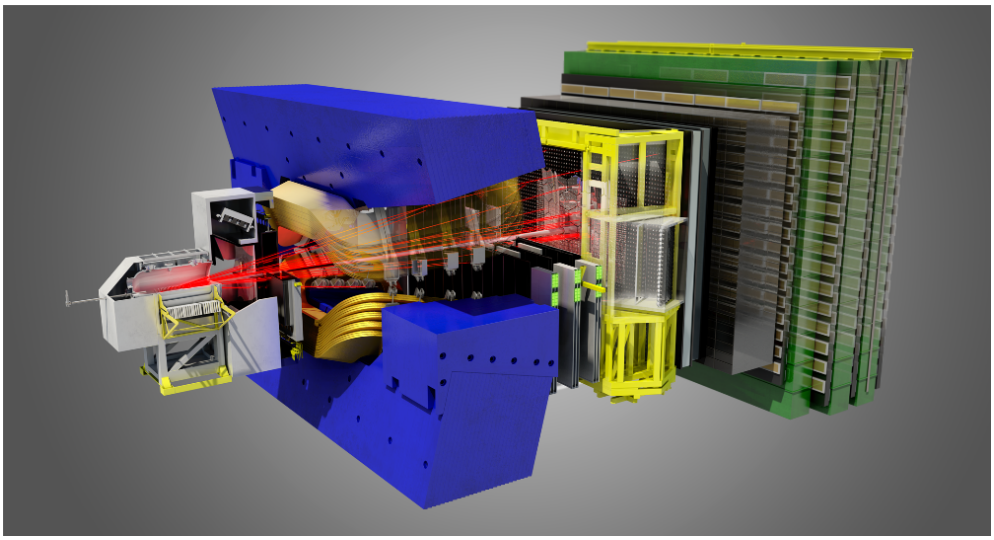


Figure 3.4.: *Simulated 3d-view of the LHCb detector. Taken from [39]*

3.3. The LHCb Detector

As mentioned in the previous section, the LHCb detector is a one-armed forward spectrometer designed to maximize coverage of produced $b\bar{b}$ pairs. It covers a pseudorapidity range of $2 < \eta < 5$. The coordinate system of the detector is defined by the beam line (z-direction). The dipole magnet bends charged paticle tracks in the xz-plane. The vertical y-direction is not affected by the magnetic field.

Figure 3.4 shows a simulated view of the LHCb detector.

The LHCb detector took data during Run 1 of the LHC in the years 2010, 2011 and 2012 with a data taking efficiency of more than 90% (see Figure 3.5). Out of the recorded data more than 99% are regarded good and can be used for physics analyses.

The individual subsystems of the detector can be divided in two categories (see Figure 3.6): track reconstruction and particle identification. The track reconstruction subsystems include the Vertex detector (VELO) surrounding the interaction point, a small tracking station (Tracker Turicensis, TT) before the dipole magnet and the main tracking stations (T1–T3) after the magnet.

The subdetectors for particle identification include two Cherenkov detectors (RICH1 and RICH2) before and behind the magnet to distinguish different hadron species (charged kaons, pions and protons). Additionally, the calorimetry system consists of a scintillating pad detector (SPD), a pre-shower detector (PS), and an electromagnetic (ECAL) and hadronic (HCAL) calorimeter. Finally the muon

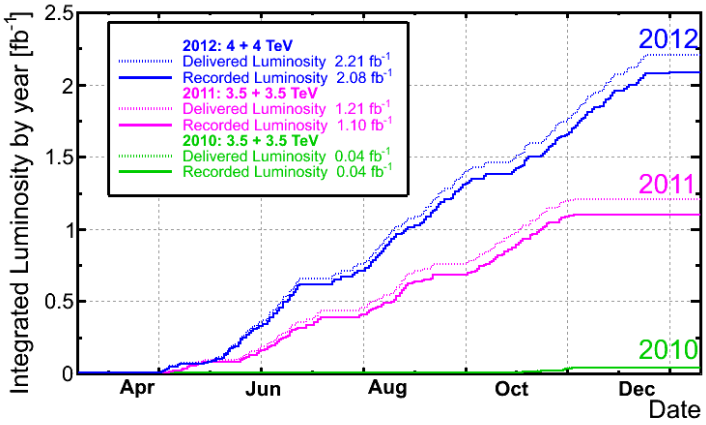


Figure 3.5.: *Delivered and recorded integrated luminosity, $\mathcal{L}_{int}$, during the years 2010–2012. Taken from [39]*

stations (M1–M5) identify tracks from charged muons.

The individual detector components are described in detail in [41] and will be summarized in the following sections. Figure 3.6 shows a schematic cross section through the LHCb detector showing the individual subsystems.

3.4. Track Reconstruction

The purpose of the track reconstruction detectors is to measure the trajectories (tracks) of charged particles and their momenta. Charged particles are bent in the xz-plane by the magnetic field of the dipole magnet. The position measurements before the magnet (VELO and TT) are combined with the measurements after the magnet (T1–T3) to form the trajectory. The percentage of working detector channels was $> 99\%$ over the whole data taking period and the track reconstruction efficiency is about 96%.

3.4.1. Dipole Magnet

The momentum of charged particles is measured by how much the trajectory is bent by the magnetic field. The LHCb dipole magnet is a non-superconducting magnet. The shapes of the magnet coils and the iron yoke are shown in Figure 3.7(a). The integrated magnetic field over $l = 10\,\mathrm{m}$ in z-direction is

$$\int Bdl = 4\mathrm{Tm} \tag{3.3}$$

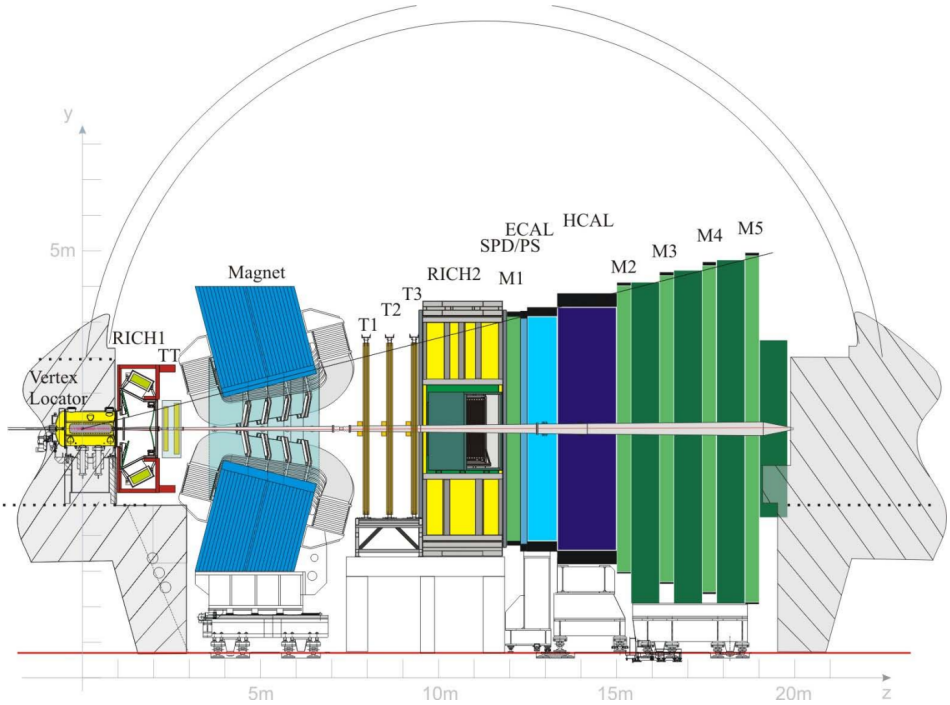


Figure 3.6.: *Schematic cross-section through the LHCb detector showing the yz-plane. The individual subsystems are described in sections 3.4 and 3.5. Taken from [42]*

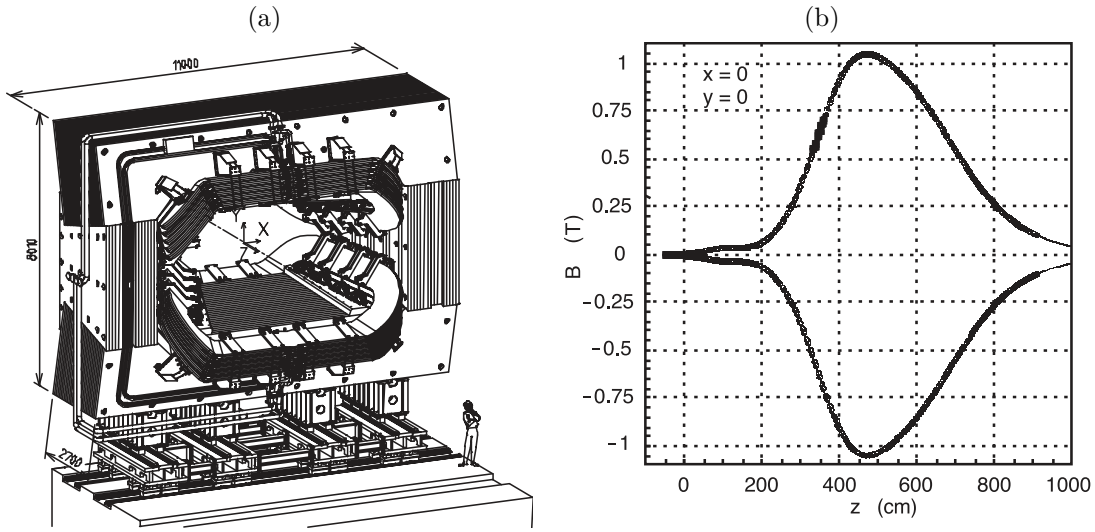


Figure 3.7.: *(a) Geometry of the LHCb dipole magnet. (b) Field strength of the magnetic field as a function of z-position. Taken from [41].*

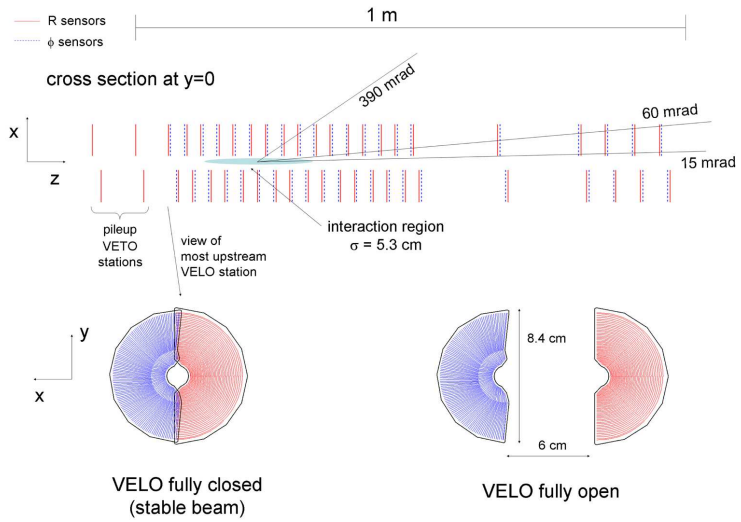


Figure 3.8.: *Geometry of the LHCb vertex locator (VELO). It consists of 21 half-circular shaped stations along the beam line. Until beam conditions are stable, the modules are retracted from the nominal position. Taken from [41].*

The field is designed to have the field lines in y-direction with minimal components in x-direction. Thus, tracks from charged particles are not bent in y-direction. Figure 3.7(b) shows the field strength as a function of the z-position. The polarity of the LHCb magnet can be reverted which results in positively charged particles occupying regions of the detector where before the negatively charged particles were bent to. Thus many detector asymmetries are canceled in combining data taken with magnet polarity upward (MagUp) and magnet polarity downward (MagDown). The polartity was regularly switched during data taking.

3.4.2. Vertex Locator

The Vertex locator (VELO) is the subdetector surrounding the interaction point. Its purpose is to measure precisely the tracks of charged particles produced in the pp collisions and distinguish particles from secondary vertices from prompt particles that were produced directly in the primary interaction.

The VELO comprises 21 stations with half-circular shaped modules located left and right of the beam line (see Figure 3.8). Each module consists of two sensors, one providing a measurement in radial direction (r) and the other one in azimuthal direction (ϕ). Additionally, two pile-up stations are located before the main sensors consisting of two r-sensors each. They are used in the trigger.

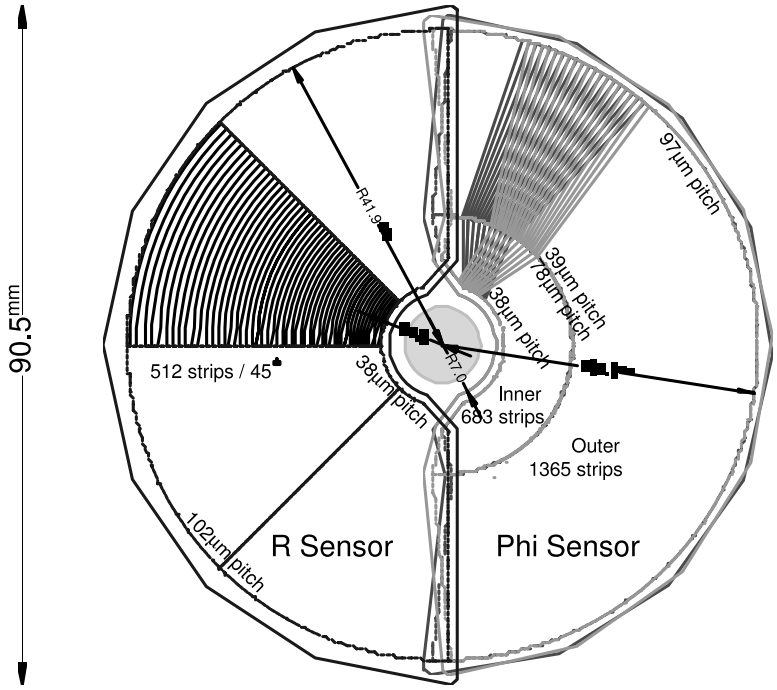


Figure 3.9.: *Geometry of VELO r- and ϕ-sensors which together make up one VELO module. Taken from [41].*

The VELO sensors are silicon strip detectors. The strips of the r-sensors are concentric semi-circles with a minimum pitch of 40 µm, close to the beam, where the occupancy is highest. The pitch increases to 102 µm at the outer edge of the sensors. For the ϕ-sensors the silicon strips are arranged in radial shape, but divided into two regions. In the inner region the pitch is 38 µm and in the outer region it increases to 97 µm. Figure 3.9 illustrates the geometry of the r- and ϕ-sensors.

The precision on the position of vertices depends on the extrapolation of the measured tracks to the point of origin. Therefore, the VELO modules are placed as close to the beam as possible. Rather than mounting the modules on top of the beam-pipe, the beam-pipe is removed and the active region of the VELO sensors are in nominal position only 8 mm away from the beam. During acceleration of the beam the beam position can change more than these 8 mm. Therefore, the VELO modules are retracted from their nominal position by 30 mm while the LHC is filled and the beams are accelerated to nominal energy (see Figure 3.8). Only when the LHC operators declare that the beam conditions are stable, the VELO modules are moved into nominal position for data taking.

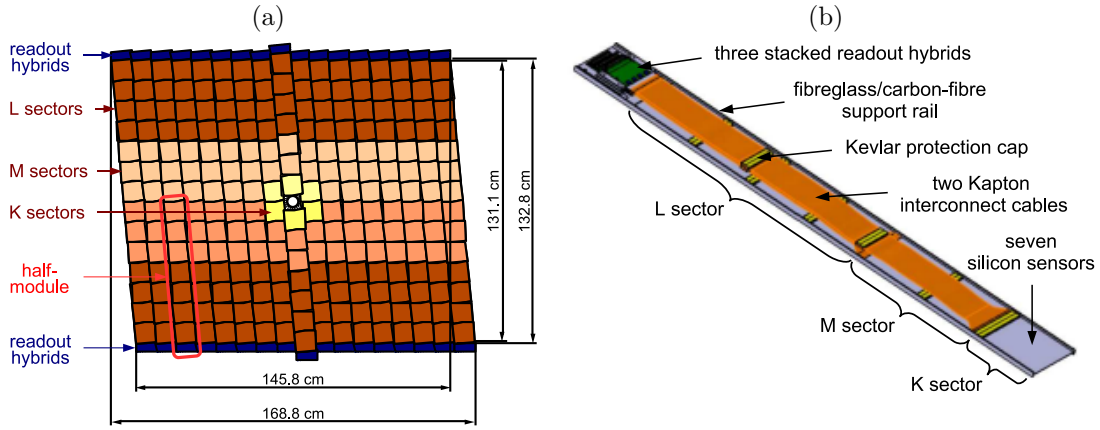


Figure 3.10.: *(a) Layout of one of the layers of the LHCb Tracker Turicensis (TT) that is tilted by 5°. (b) Layout of TT half module. Taken from [41]*

3.4.3. Tracker Turicensis

The Tracker Turicensis (TT) is a 150 cm wide and 130 cm high silicon microstrip detector located just before the magnet. It consists of four layers out of which two are tilted by 5° to be able to perform three dimensional measurements. The design is chosen to have the best possible resolution in the bending plane of the magnet (xz-plane).

Figure 3.10(a) shows the geometry of the TT modules. The detector is divided into different readout sections, since the occupancy close to the beam is much higher than in the outer regions. The layers of the TT are built up from half modules which themselves consist of seven silicon sensors (see Figure 3.10(b)). Each sensor is 9.44 cm long, 9.46 cm wide and has a pitch of 183 µm. The single hit resolution of the TT is about 50 µm.

3.4.4. Inner Tracker

The main tracking stations (T1–T3) are divided into the Inner tracker (IT) and the Outer tracker (OT). The IT is a silicon strip detector located in the high occupancy region close to the beam pipe. It is about 120 cm wide and 40 cm high.

Each of the three IT-stations comprises four layers of silicon strip sensors of which similar to the TT the middle two are tilted by 5°. Each station consists of four boxes arranged around the beam pipe. The layout of one IT station is shown in Figure 3.11(a) and the arrangement of the sensors in one layer inside one of the boxes is illustrated in Figure 3.11(b).

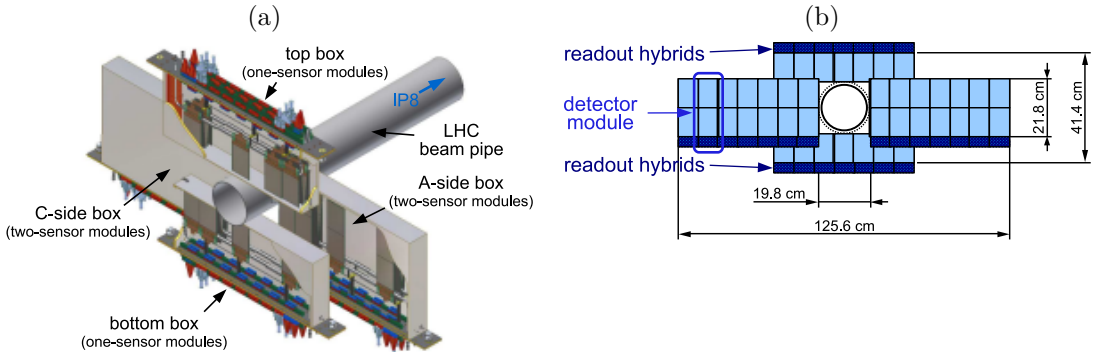


Figure 3.11.: *(a) Layout of one of the Inner Tracker (IT) stations. (b) Layout of the sensors in one of the IT layers. Taken from [41]*

The silicon sensors are 7.6 cm wide and 11 cm long, with a pitch of about 200 µm. The sensors in the boxes above and below the beam pipe are 320 µm thick and the sensors in the boxes on the left/right are 410 µm thick. This geometry was chosen to limit the maximum occupancy to a few percent. The single hit resolution of the IT is about 50 µm as of the TT.

3.4.5. Outer Tracker

The Outer Tracker (OT) is the second subdetector of the main tracking stations (T1–T3). In contrast to the other tracking subsystems it is not silicon-based, but a drift tube detector. It covers the large area in the xy-plane outside the IT acceptance (see Figure 3.12(a)). It is about 6 m wide and 5 m high.

Like in case of the IT the three tracking stations consist of four OT layers each. They are as well arranged with the middle two layers tilted by 5° to allow three dimensional measurements. Each layer consists of modules, each containing a double layer of straw-tubes (see Figure 3.12(b)). The straws are filled with a mixture of Argon/CO_2/O_2 with the proportions 70%/28.5%/1.5%. In the center of the straw an anode wire made of gold coated tungsten is located. The straws have a pitch of 5.25 mm with an inner diameter of 4.9 mm.

The single hit resolution of the OT is 200 µm in x-direction.

3.4.6. Track reconstruction algorithms

The track reconstruction software used in LHCb is called Brunel. It constructs four different track types depending on the sub detectors that contributed:

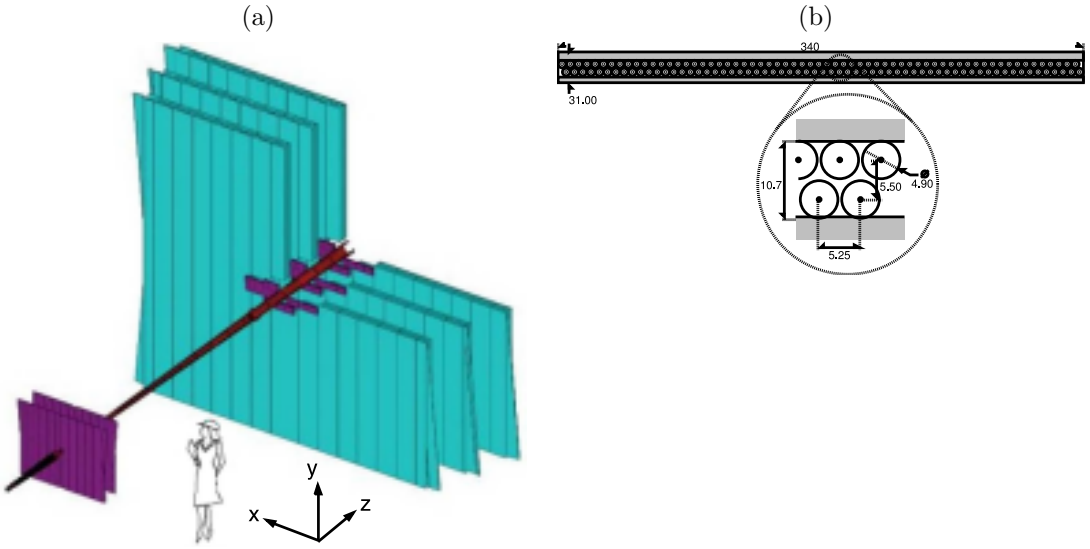


Figure 3.12.: *(a) Layout of the Outer Tracker (OT) stations (light blue). The TT and IT modules are indicated in purple. (b) Layout of the straw-tubes inside an OT module. Taken from [41]*

- **VELO tracks** contain only information from the vertex locator. Since there is no magnetic field present in the VELO, these are straight line tracks. They are used to build the primary vertices.
- **Long tracks** are the default tracks used in LHCb analyses. They combine the information from all tracking detectors (VELO, TT[2] and T1–T3). They have the best momentum resolution and the smallest contribution from fake tracks (ghosts).
- **Upstream tracks** have only information from the VELO and TT. These are usually low momentum tracks from particles that are bent out of the acceptance of the main tracking stations by the magnetic field. Since the magnetic field reaches into the TT acceptance for these tracks a momentum estimate is possible.
- **Downstream tracks** contain only information from the TT and the main tracking stations (T1–T3). These tracks are mainly used to reconstruct long-lived particles like K^0_S or Λ which decay outside of the VELO acceptance.
- **T-tracks** are track segments built only from hits in the main tracking stations. These tracks are used in the process of building Long tracks.

In the Brunel software, the first step are pattern recognition algorithms to find hits belonging to a track. These algorithms reconstruct track segments in the

[2]Hits in the TT are added if they are found, but are not necessary to form a Long track.

VELO and the main tracking stations. These tracks are used to build Long tracks. Two different algorithms are used for the Long track reconstruction.

- **Forward tracking** starts with a VELO track and extrapolates it to the main tracking stations. Hits in the main tracking stations are searched in a window around the extrapolated VELO track. If certain quality criteria are sufficiently satisfied, the resulting track is chosen as a Long track. Finally, if available, hits in the TT are added to the Long track.
- **Track matching** also starts with a VELO track, but searches for matching T-tracks. Both track segments are extrapolated inside the magnet and if they match together, they are combined to a Long track. Again the hits in the TT are added after the matching is finalized.

The two Long track finding algorithms have a large overlap of $\sim 90\%$. After the track reconstruction, duplicate tracks (clones) are removed. Then the remaining tracks are fitted using a Kalman Filter [43] which takes into account effects from multiple scattering and energy loss.

One measure of the quality of the fitted track is the fit χ^2 over the degrees of freedom, χ^2/ndf.

3.5. Particle Identification

The ability to distinguish different particle species (particle identification, PID) is crucial for the reconstruction of B meson decays at the LHC. Especially, separating charged $\pi^\pm$ from other particle species ($K^\pm$, $\mu^\pm$, $e^\pm$, p) is challenging because there are are seven times more $\pi^\pm$ than charged particles from other species. Particularly for purely hadronic final states, like used in this analysis ($B_s^0 \to D_s^- \pi^+$), PID poses a considerable challenge. Several subdetectors provide information for particle identification which is combined to a common particle hypothesis.

3.5.1. Ring Imaging CHerenkov detector

Ring Imaging CHerenkov detectors (RICH) take advantage of the so-called Cherenkov-effect. Charged particles emit photons if they traverse a medium that has refractive index n with a velocity higher than the speed of light in the material ($c' = \frac{c}{n}$). This light is emitted in the shape of a cone with the Cherenkov angle θ_{Ch} as opening angle:

$$\cos\theta_{\mathrm{Ch}} = \frac{c'}{v} = \frac{1}{n\beta}. \tag{3.4}$$

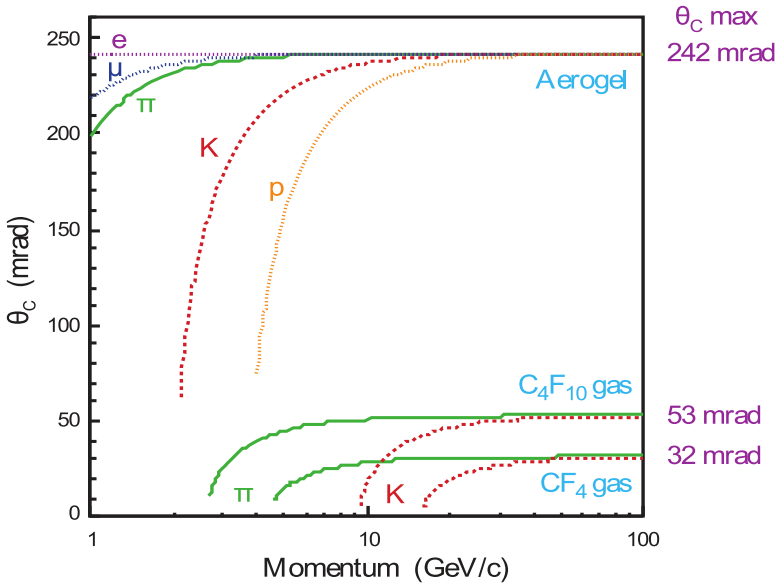


Figure 3.13.: *Cherenkov angles for different particle species traversing the three different materials used in the LHCb RICH detectors (Aerogel,* C_4F_{10} *and* CF_4*) plotted against the particle momenta. Taken from [41].*

with $\beta = v/c$, the velocity of the particle. Combining the measurement of θ_{Ch} with the measured momentum from the tracking system, described in Section 3.4, allows to calculate the particle's mass.

Figure 3.13 shows the Cherenkov angles for different particle species over a large momentum range for the three different radiator materials used in the LHCb RICH detectors. It can be seen that the Aerogel is better for particle separation at low momenta, whereas the CF_4 gas is better suited to separate high momentum $K^{\pm}$ and $\pi^{\pm}$.

To ensure a good K–π separation over the whole momentum range these three radiators are used in two RICH detectors in the LHCb detector. The first, RICH1 (see Figure 3.14(a)), is located before the magnet and uses the radiators, an aerogel with refractive index $n = 1.03$ and C_4F_{10} with $n = 1.0014$ to cover a momentum range from about 1 GeV/c to 60 GeV/c. The emitted Cherenkov photons are reflected and focussed through a system of mirrors onto Hybrid Photo Detectors (HPD) to form the shape of a circle. By measuring the radius of the circle the Cherenkov angle is determined. The HPDs are sensitive to photons with a wavelength of 200–600 nm. RICH1 covers the full LHCb angular acceptance of 25–300 mrad in x-direction and 25–250 mrad in y-direction.

The second RICH detector, RICH2 (see Figure 3.14(b)), is located behind the main tracking stations. Its purpose is to identify charged particles with a momentum

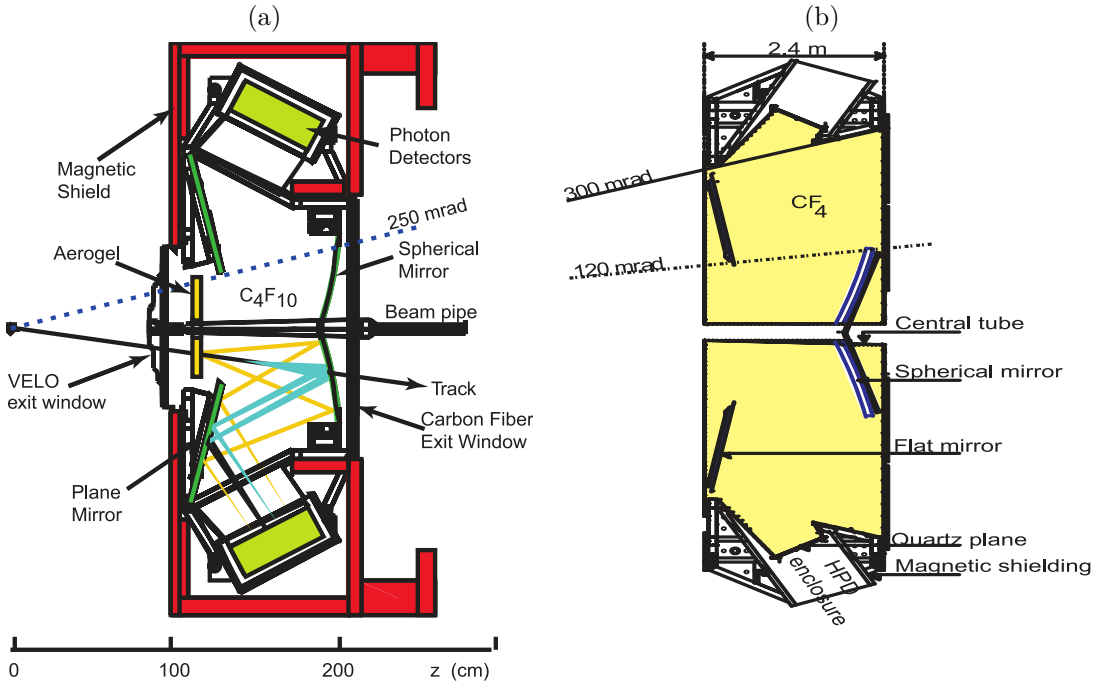


Figure 3.14.: *Layout of the Ring Imaging CHerenkov detectors, RICH1 (a) and RICH2 (b). Taken from [41]*

from 15 GeV/c to 100 GeV/c. It uses CF_4 with a refractive index of $n = 1.0005$ as radiator. RICH2 covers an angular acceptance of 15–120 mrad in x- and 15–100 mrad in y-direction.

The efficiency to identify charged kaons is 95% for a probability of 5% to misidentify a $\pi^{\pm}$ as a $K^{\pm}$.

3.5.2. Calorimetry

The purpose of the calorimeter system is to measure the amount and position of energy depositions of particles. It is the only subsystem of the LHCb detector that is sensitive to neutral particles like photons. The calorimetry system is used in the first stage of the LHCb trigger (L0) which will be described in Section 3.6. This imposes strong requirements on the speed of the readout of the calorimetry system since the L0 decision has to occur after only 4 µs.

The LHCb calorimetry system consists of several subsystems which have the same working principle. The particles first traverse an absorber material which induces particle showers. The particles in the shower pass through scintillating material. The photons emitted from the scintillators are read out via wavelength shifting

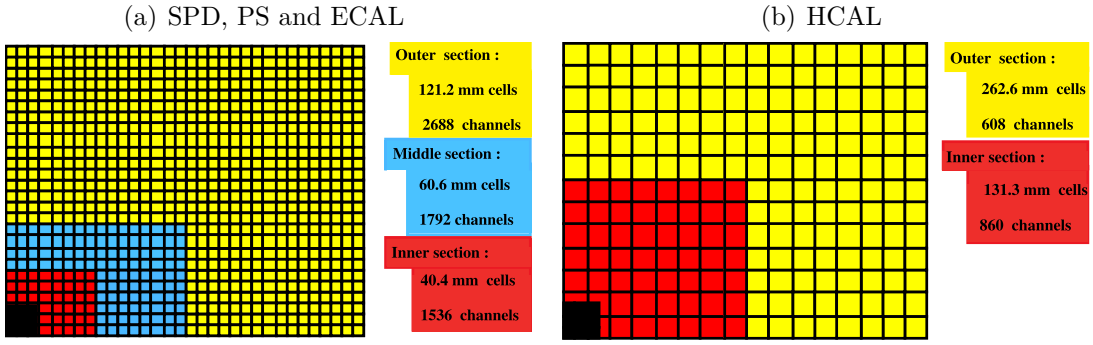


Figure 3.15.: *Granularity for the different detector regions of the SPD, PS, and ECAL (a) and of the HCAL (b). Taken from [41]*

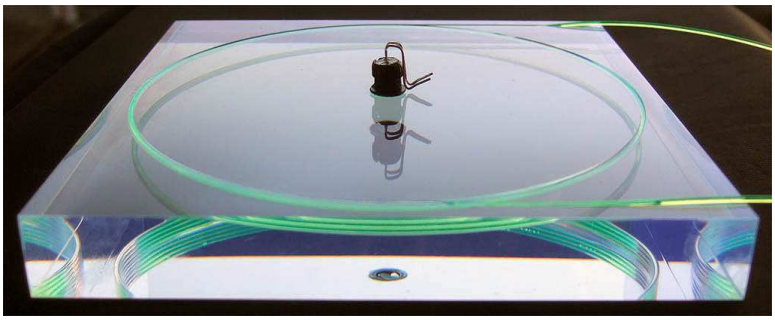


Figure 3.16.: *Scintillator pad used in the SPD and PS with the wavelength shifting fiber layout. Taken from [41].*

fibers and detected with photomultiplier tubes. By collecting all photons induced by the particle shower, the energy of the initial particle can be calculated.

The calorimeter subsystems include the Scintillating pad detector (SPD), the preshower detector (PS) and the electromagnetic (ECAL) and hadronic (HCAL) calorimeters. The granularity of the sensors becomes finer with proximity to the beam pipe. Figure 3.15 shows the different regions for the individual calorimeter subsystems.

Scintillating Pad Detector and Preshower Detector

The first subdetector of the LHCb calorimetry system are the Scintillating pad detector (SPD) and preshower detector (PS). They are located behind the first muon station (M1) which will be discussed in Section 3.5.3. The SPD and PS are separated by a 15 mm thick lead absorber. Both detectors use scintillating pads as shown in Figure 3.16. Figure 3.15(a) shows the granularity for the different detector regions for both the SPD and PS.

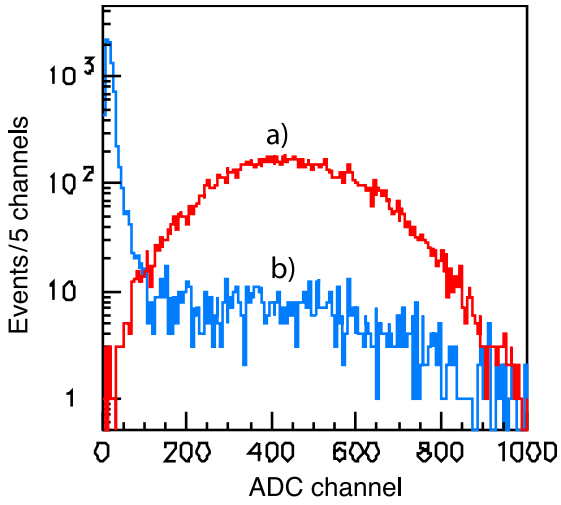


Figure 3.17.: *Energy depositions in the PS detector for 50* GeV/c *electrons (a) and* $\pi^{\pm}$ *(b) are shown in arbitrary units. Taken from [41].*

The purpose of the SPD is to distinguish electrons from photons, since the electrically neutral photons do not induce a signal in the scintillating pads. The misidentification rate of photons as electrons is found to be below 3%.

The lead absorber in front of the PS detector intiates showers from electrons and photons. The purpose of the PS detector is to distinguish between electrons and $\pi^{\pm}$. The energy depositions for 50 GeV/c electrons and $\pi^{\pm}$ are shown in Figure 3.17 in arbitrary units. In this momentum region the PS shows an electron identification efficiency of 97% while rejecting 99.7% of the $\pi^{\pm}$.

Electromagnetic calorimeter

The electromagnetic calorimeter, ECAL, detects particle showers from photons and electrons. It is has a so-called "shashlik" structure, which means that it is a sampling calorimeter with a stack of alternating layers of 2 mm thick lead absorber and 4 mm thick scintillator material. The scintillators are again read out using wavelength shifting fibers.

Three different types of ECAL modules were produced (see Figure 3.18(a)) for the different granularity regions shown in Figure 3.15(a). The length of the ECAL corresponds to 25 radiation lengths X_0 and 1.1 hadronic interaction lengths λ_{I}. The energy resolution of the ECAL is

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{10\%}{\sqrt{\mathrm{E[GeV]}}} \oplus 1\% \tag{3.5}$$

(a) ECAL module (b) HCAL module

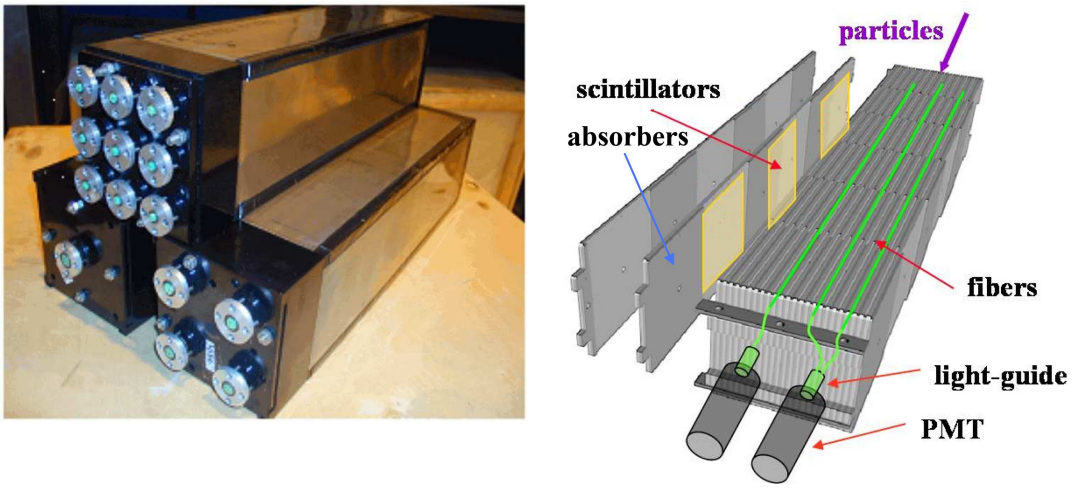


Figure 3.18.: *(a) Three ECAL modules for the different detector regions described in Figure 3.15(a). (b) schematic view of an HCAL module. Taken from [41]*

where the symbol $\oplus$ denotes the summation in quadrature. The efficiency to identify electrons is $\sim 90\%$ for a 5% $\pi^{\pm} \to e^{\pm}$ misidentification probability.

Hadronic calorimeter

The hadronic calorimeter, HCAL, detects particle showers from hadrons and is located behind the ECAL. It is a sampling calorimeter using alternating iron absorbers and scintillator material. The special feature of the HCAL is the orientation of the scintillating tiles parallel to the beam pipe. They are in longitudinal direction intersected with iron absorbers which in thickness correspond to one hadronic interaction length in steel λ_{I}. In transverse direction the active scintillator tiles are separated by 1 cm thick iron tiles.

Figure 3.18(b) illustrates an HCAL module. The total length of absorber and scintillating material corresponds to 5.6 hadronic interaction lengths. The energy resolution of the HCAL is

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{80\%}{\sqrt{\mathrm{E[GeV]}}} \oplus 10\% \tag{3.6}$$

where again the symbol $\oplus$ denotes the summation in quadrature.

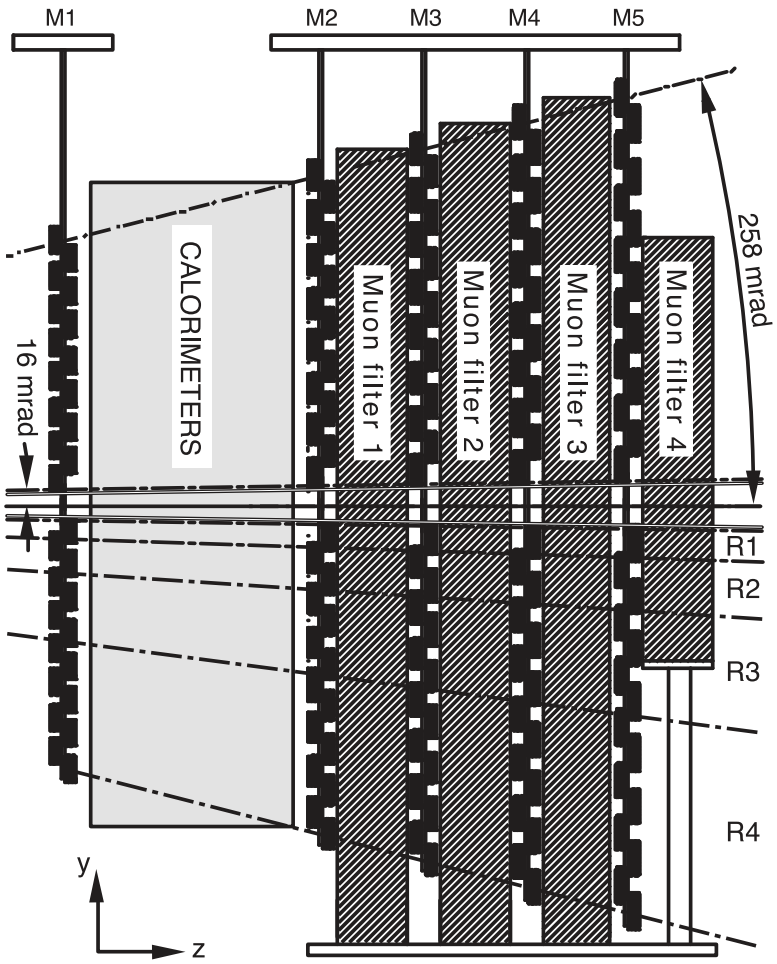


Figure 3.19.: *Schematic view of the LHCb muon system. Taken from [41].*

3.5.3. Muon Chambers

The LHCb muon stations (M1–M5) are crucial for identifying B meson decays with muons in the final state. Furthermore, they are used in the first stage of the trigger (L0). The muon system consists of five stations (M1–M5), as shown in Figure 3.19.

The first station (M1) is located before the calorimetry system and its purpose is to improve the momentum resolution in the muon trigger by using a measurement before multiple scattering occurs in the calorimeter. Between the stations M2–M5 80 cm thick iron absorbers are placed to filter out all particles that are not muons. These absorbers introduce an implicit momentum requirement on muons of 6 GeV/c to traverse the whole detector until M5.

Similar to the calorimeters the muon stations are divided into regions of different

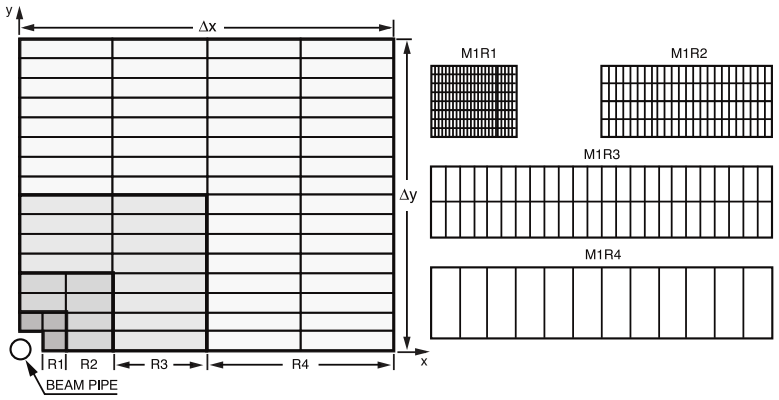


Figure 3.20.: *Schematic view of the different detector regions in M1. Taken from [41].*

granularity depending on their proximity to the beam pipe. Figure 3.20 shows the different regions for M1. In M2 and M3 the number of pad columns per chamber is double than in M1, while in M4 and M5 it is half. The number of pad rows per chamber is the same in all stations.

In all regions the pads consist of multiwire proportional chambers (MWPC) except of R1 in M1. Here, gas electron multiplier (GEM) detectors are used since MWPCs are too susceptible to radiation damage to be operated in the high particle flux in this region. The angular acceptance of the muon chambers reaches from 20 mrad to 306 mrad in the xz-plane and from 16 mrad to 258 mrad in the yz-plane. The muon identification efficiency is $\sim 97\%$ for a 1–3% $\pi^{\pm} \to \mu^{\pm}$ misidentification probability.

3.6. Trigger System

The purpose of the LHCb trigger system is to reduce the event-rate from the nominal 40 MHz bunch crossing rate to about 5 kHz at which the data is recorded for physics analyses. With an instantaneous luminosity of $\mathcal{L} = 4 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and the $b\bar{b}$ cross section in Equation 3.1, a production of about 30,000 $b\bar{b}$ pairs per second in the LHCb acceptance can be calculated. Since the rate at which data can be recorded is limited to 5 kHz, it is not enough to just identify $b\bar{b}$ events in the trigger. The b hadron decays that are interesting for the physics analyses have to be distinguished from other b hadron decays already at trigger stage.

The LHCb trigger is a three stage system consisting of one hardware trigger, called Level0 (L0), and two software trigger stages, called High-Level-Trigger (HLT1 and HLT2). Figure 3.21 illustrates the trigger scheme. The trigger system is designed to be adjustable to different running conditions. The trigger configuration and

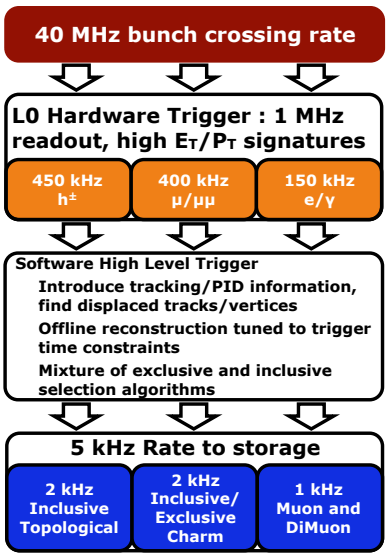


Figure 3.21.: *Overview of the LHCb trigger scheme. The event rate is reduced by the hardware stage (L0) from 40* MHz *to 1* MHz*. The software stage reduces it further to 5* kHz*. Taken from [39].*

performance for the data taking during 2011 are described in [44] and will be briefly summarized in this section.

3.6.1. Hardware Trigger

The first stage of the LHCb trigger is completely implemented in hardware and called Level0 (L0). It is divided into two parts: The L0-muon trigger and the L0-calorimeter trigger.

The **L0-calorimeter** trigger uses information from the calorimetry system described in Section 3.5.2 (SPD, PS, ECAL and HCAL). It selects events with high transverse energy, E_{T}, depositions in 2×2 cells, which is defined as

$$E_{\mathrm{T}} = \sum_{i=1}^{4} E_i \sin\theta_i, \tag{3.7}$$

with E_i being the energy deposition in cell i. The angle θ_i is defined as the angle between the beam axis and the connection between the primary interaction point and the center of cell i. Candidates with high E_{T} are selected and a particle hypothesis (L0-Hadron, L0-Photon, L0-Electron) is assigned.

The **L0-muon** trigger reconstructs in each quadrant of the detector the two muon candidates with the largest transverse momentum p_{T}. The muon candidates' p_{T} is estimated assuming they originate from the primary interaction point and using the slope in the xz-plane in M1 and M2. The event is accepted if either a requirement on the p_{T} of the candidate with the largest p_{T} (L0-Muon) or the combination of the two largest p_{T} candidates (L0-DiMuon) is met.

3.6.2. Software Trigger

The software stage of the LHCb trigger is called High Level Trigger (HLT). It runs on a large computing cluster with about 26,000 CPUs. The HLT works in two consecutive steps (HLT1 and HLT2).

In the HLT1 a partial event reconstruction is performed based on the candidates passed on by the L0 decision. A fast track reconstruction in the VELO and main tracking stations is performed to find displaced vertices. The purpose of the HLT1 is to reduce the event rate to 30 kHz.

The HLT2 fully reconstructs the whole event which is done in a similar manner to the offline reconstruction, but with concessions to the timing constraints. In this step PID information is added and the output bandwidth is divided in 131 trigger lines. A large portion of the recorded output is dedicated to inclusive topological trigger lines which attempt to reconstruct all b-hadron decays with two or more charged particles in the final state and a displaced vertex. This analysis uses events triggered by the topological trigger lines.

The rest of the bandwidth is divided into inclusive and exclusive trigger lines selecting c-hadron decays and inclusive and exclusive Single- and Di-Muon trigger lines. The combined trigger efficiency for the hardware and software stages is about 90% for signal decays containing two muons and about 30% for multi-body hadronic final states.

3.7. The LHCb Software Framework

To perform an analysis, the data taken by the LHCb experiment has to be processed by two software packages.

- **Brunel**
  The first step to process the raw data, is track reconstruction. This is performed by the Brunel software package [45]. Aside from the track fit described in Section 3.4.6, Brunel also uses PID information from the RICH, calorimeter system and the muon chambers, combines it and

assigns it to the tracks. Furthermore the electromagnetic and hadronic showers in the calorimeters are reconstructed and the energy is added to the particle candidates. The result of this reconstruction step are so-called Data Summary Tape (DST) files on which the LHCb analysis software DaVinci is run.

- **DaVinci**
  The analysis software package DaVinci [46] is built on the Gaudi framework [47]. During the analysis process particle hypotheses are added to the tracks and the stable particles are combined to reconstruct the signal decays for the analysis. Within this process the decay tree fitter (DTF) [48] uses the particle tracks as inputs and fits simultaneously all vertices specified in the signal decay chain.

  Various selection criteria are applied in two steps. First there is a central processing of all LHCb data in which a preselection is applied. This step is called Stripping. The reason is to reduce the data analysed in each single physics analysis to a manageable amount. The second step is applying the final offline selection that is unique for each analysis. Both steps are performed using the DaVinci analysis software and will be discussed in detail for the signal $B_s^0 \to D_s^- \pi^+$ decays used in this analysis in Chapter 5.

Simulated data sets are produced using the following software packages which are all built on the Gaudi framework

- **Gauss**
  Simulated events are generated using the Gauss [49] software package. pp collisions are simulated using Pythia 6.4 [50] with a special tuning for LHCb [51]. The decays of hadronic particles is simulated using EvtGen [52] in which final state radiation is generated using Photos [53]. The interaction of particles with the detector material is simulated by Geant4 [54], [55].

- **Boole**
  The output of the simulated interaction of the particles with the detector is digitized by the Boole [56] software which simulates the detector response to the signals induced in the individual subdetectors.

- **Moore**
  The Moore [57] software package emulates the signals of the trigger on the detector responses in the same fashion as on data.

- **Brunel and DaVinci**
  All individual steps of the track reconstruction and analysis software are performed in the same way as on data.

3.8. Data sets used

The data set used in this analysis consists of all data of pp collisions at a center of mass energy of $\sqrt{s} = 7\,\text{TeV}$ taken by the LHCb experiment in 2011. It corresponds to 1 fb^{-1} of integrated luminosity.

Additionally simulated data samples are used for several exclusive decay modes. In these samples is made sure that the specified signal decay is present in each event and the stable decay products are within the LHCb acceptance.

CHAPTER 4

Analysis setup

This chapter gives an overview of the analysis strategy for the measurement of Δm_s. In Section 4.2 the Maximum Likelihood formalism is explained which is used by the fit algorithm in this analysis.

4.1. Analysis strategy

The goal of this analysis is to provide a measurement of the B_s^0–$\overline{B}_s^0$ oscillation frequency Δm_s. This analysis is performed decay-time-dependent. The decay time t of the particle describes the time between production and decay of the particle in its rest frame. It is measured as:

$$t = \frac{(\vec{\ell} \cdot \vec{p})m}{|\vec{p}|^2}, \tag{4.1}$$

with $\vec{\ell}$ being the vector from the production vertex of the particle to its decay vertex, m being its reconstructed invariant mass and $\vec{p}$ being its reconstructed three-momentum. It can also be calculated as

$$t = \frac{|\vec{\ell}|}{\beta\gamma} \tag{4.2}$$

with $\beta = \frac{v}{c}$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ being the usual relativistic quantities describing the Lorentz boost.

The decay time resolution is limited by two effects, the uncertainty of the momentum and the uncertainty of the measured flight distance. LHCb achieves an excellent decay time resolution because both contributions are small. The relative momentum resolution is $\frac{\Delta p}{p}$ is 0.4% for low momentum tracks ($< 5\,\mathrm{GeV}/c$) and 0.6% for high momentum tracks ($> 100\,\mathrm{GeV}/c$), provided by the tracking system described in Section 3.4. The effect on the decay time resolution from the uncertainty on the flight distance $|\vec{\ell}|$ is dominated by the accuracy with which the position of the secondary vertex can be measured. Due to the strong boost of the B_s^0 mesons produced at the LHC, the B_s^0 decay vertex is well separated from the primary interaction point. The treatment of the decay time resolution in this analysis will be discussed in detail in section 7.1.

The second challenge for this analysis is that the signal B_s^0 candidates have to be sorted into two categories, namely if they were *mixed* or *unmixed*. Mixed, in this case, refers to candidates that were produced as B_s^0 mesons, but decayed as $\overline{B}_s^0$ mesons or vice versa. It is irrelevant for this definition how often the meson changed flavour (B_s^0 or $\overline{B}_s^0$) in between, only that the flavour at production is different from the flavour at decay. Unmixed candidates respectively are those which have the same flavour at production and decay. The decay $B_s^0 \to D_s^- \pi^+$ is chosen because it is self-tagging. This means that the flavour at decay is directly determined by the charge of the decay products, *i.e.* the $D_s^\mp$ and the $\pi^\pm$. The pion produced in association with the $D_s^\mp$ will in the following be referred to as bachelor pion. The probability to assign the wrong charge to both reconstructed particles is low enough to be neglected. The flavour at production is provided by so called flavour tagging algorithms which will be described in detail in Chapter 6. Since the decay channel is self-tagging, the decision by these algorithms give directly the decision, q, whether the B_s^0 candidate was mixed or unmixed.

The final challenge for this analysis is to provide a sample of $B_s^0 \to D_s^- \pi^+$ signal candidates with large statistic and high purity. The decay $B_s^0 \to D_s^- \pi^+$ is reconstructed in five different D_s^- decay modes namely $D_s^- \to \phi(K^+K^-)\pi^-$, $D_s^- \to K^{*0}(K^+\pi^-)K^-$, $D_s^- \to K^+K^-\pi^-$[1], $D_s^- \to K^-\pi^+\pi^-$ and $D_s^- \to \pi^-\pi^+\pi^-$. The signal selection used to maximize statistics is discussed in Section 5.

4.2. Unbinned Maximum Likelihood Fit

The fit algorithm used in this analysis is based on the unbinned maximum likelihood technique. It estimates a set of unknown parameters $\vec{\lambda} = \{\lambda_1, \lambda_2, ...\}$ from a set of measured observables $\vec{X} = \{X_1, X_2, ...\}$ assuming a theoretical distribution $f(\vec{\lambda}; \vec{X})$. In this analysis the measured observables $\vec{X} = \{m, t, q\}$ consist of the

[1] If there is no intermediate resonance specified "$D_s^- \to K^+K^-\pi^-$" will refer to only the non-resonant part from now on

reconstructed invariant mass, m, of the B_s^0 candidate, its reconstructed decay time, t, which is defined in Equation 4.1 and the decision, q, given by the flavour tagging algorithms. The probability density function (PDF) is defined as the distribution $f(\vec{\lambda};\vec{X})$ normalized over the whole range of $\vec{X}$

$$\mathcal{P}(\vec{\lambda};\vec{X}_i) = \frac{f(\vec{\lambda};\vec{X}_i)}{\int f(\vec{\lambda};\vec{X}'_i)d\vec{X}'_i}. \tag{4.3}$$

It gives, for a certain set of parameters, $\vec{\lambda}$, the probability to measure in a single event, i, the data, $\vec{X}_i$, within an interval, $\Delta\vec{X}_i$. The likelihood function, $\mathcal{L}$, is the product of the PDFs of all reconstructed candidates.

$$\mathcal{L}(\vec{\lambda};\vec{X}) = \prod_i^{\text{candidates}} \mathcal{P}(\vec{\lambda};\vec{X}_i) \tag{4.4}$$

The likelihood function gives a measure for the probability for a certain parameter set, $\vec{\lambda}$, to measure the data set

$$\vec{X} = \bigcup_i \vec{X}_i. \tag{4.5}$$

The fit algorithm maximizes the likelihood by varying the parameters $\vec{\lambda}$. In this way, the set, $\vec{\lambda}_{\text{max}}$, for which the probability to measure the given data set, $\vec{X}$, is maximal, is found.

The word "unbinned" refers to the fact that the data sets $\vec{X}_i$ are not filled in bins before the fit is performed, but every event is used individually by the fit. Therefore, no information is lost.

The PDF used in this analysis is divided into two parts. One part describes the contributions from the signal components, $\mathcal{P}_{\text{sig}}$, and another describes the background components, $\mathcal{P}_{\text{bkg}}$, which will be listed in Section 5.8. They are combined in the following way:

$$\mathcal{P}(\vec{\lambda};\vec{X}) = f_{\text{sig}} \cdot \mathcal{P}_{\text{sig}}(\vec{\lambda}_{\text{sig}}; m, t, q) + (1 - f_{\text{sig}}) \cdot \mathcal{P}_{\text{bkg}}(\vec{\lambda}_{\text{bkg}}; m, t, q), \tag{4.6}$$

with the fraction of signal events

$$f_{\text{sig}} = \frac{N_{\text{sig}}}{N_{\text{sig}} + N_{\text{bkg}}}. \tag{4.7}$$

The union of the parameter sets $\vec{\lambda}_{\text{sig}}$ and $\vec{\lambda}_{\text{bkg}}$ gives the total set $\vec{\lambda}$

$$\vec{\lambda} = \vec{\lambda}_{\text{sig}} \cup \vec{\lambda}_{\text{bkg}} \tag{4.8}$$

The invariant mass, m, is assumed to be uncorrelated to the other variables, t and q. Thus a factorization ansatz is chosen for the PDFs describing the mass and mixing distributions

$$\mathcal{P}_J = \mathcal{P}_{J;m} \cdot \mathcal{P}_{J;t,q}, \tag{4.9}$$

where J stands for the individual signal/background contribution. The mixing PDF, $\mathcal{P}_{J;t,q}$, is two dimensional in decay time and the discrete flavour tagging decision. The mass PDFs for the individual components will be explained in detail in Chapter 8. The mixing PDFs with the respective parameters $\vec{\lambda}$ including Δm_s will be discussed in Chapter 9.

In this analysis, several event-by-event quantities are used. One example is an estimate for the decay time resolution σ_t. It has been shown that for each of those variables an additional PDF has to be included as a conditional probability in the fit to avoid biased results [58]. Therefore, the complete PDF for each signal/background contribution is

$$\mathcal{P}_J = \mathcal{P}_{J;m} \cdot \mathcal{P}_{J;t,q} \cdot \mathcal{P}_{J;\sigma_t} \cdot \mathcal{P}_{J;\eta_{\mathrm{OST}}} \cdot \mathcal{P}_{J;\eta_{\mathrm{SST}}}. \tag{4.10}$$

More information on the event-by-event uncertainty estimate σ_t and mistag estimates η_{OST} and η_{SST} will be given in Chapters 7 and 6.

CHAPTER 5

Signal selection for the decay $B_s^0 \to D_s^- \pi^+$

As mentioned in Section 3.7 the selection of signal $B_s^0 \to D_s^- \pi^+$ candidates is performed sequentially. The first stage is the selection in the trigger before the events are stored. The next step is a preselection that is performed centrally for all LHCb analyses. This step is called Stripping. The last step is the final offline signal selection that was optimized by the individual analysis to achieve the best signal significance.

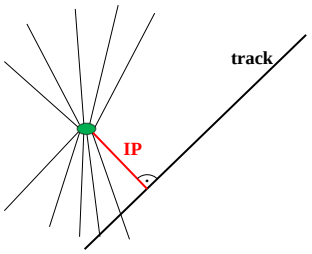
There are other analyses in LHCb using the decay $B_s^0 \to D_s^- \pi^+$ like for example the measurement of *CP* violation in the decay $B_s^0 \to D_s^\pm K^\mp$ [2]. For comparability reasons the signal selection was tried to be kept as close as possible among these analyses.

5.1. Variables used in the signal selection

In this section, the most important variables used in the signal selection are introduced and explained via small diagrams. Additionally, the signal and background distributions for these variables from simulated data are given.

The signal selection of the decay $B_s^0 \to D_s^- \pi^+$ is challenging since the final state is purely hadronic. Many physics analyses rely on final states including muons since muons leave clear signatures in the detector. For this analysis, however, a high statistics decay channel with a flavour specific final state is required. Furthermore, the final state has to be fully reconstructed to be able to use the B_s^0 invariant

(a) Impact parameter



(b) IP distribution

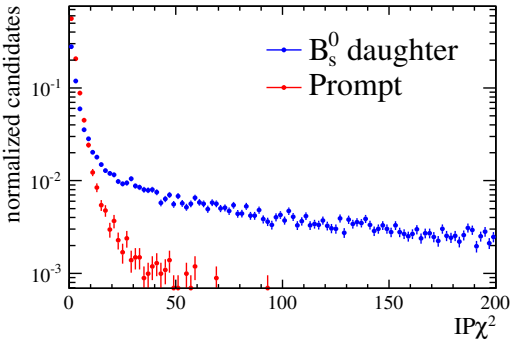


Figure 5.1.: *(a) Diagram for the definition of the impact parameter (IP) and (b) distribution of the IP χ^2 for B_s^0 daughter particles (blue) and prompt particles from the PV (red)*

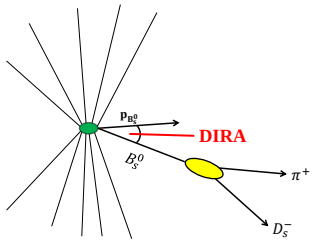
mass to distinguish between signal and background. This excludes semileponic decays, which are also self-tagging and have the largest Branching Ratio, and the only possible choice is $B_s^0 \to D_s^- \pi^+$. Therefore, many variables are necessary to separate signal from background.

One of the most powerful variables is the impact parameter (IP) which is the minimum distance between the reconstructed track and the primary vertex (see Figure 5.1(a)). It is used to distinguish between particles coming from the primary interaction or short-lived resonances and daughter particles from long-lived particles like B mesons. Daughter particles of B mesons will, on average, have larger IPs than particles that were produced directly at the primary vertex. This can be seen in Figure 5.1(b) which shows the IPχ^2 distributions for B_s^0 daughter particles (red) and prompt particles (blue). The χ^2 of the impact parameter shows a better separation than the IP itself.

Another useful geometrical selection criterion is the direction angle (DIRA). It is the angle between the flight direction, given by the connection between production vertex and decay vertex of the particle, and its reconstructed momentum (see Figure 5.2(a)). It is a good measure to see if all particles were reconstructed correctly, *i.e.* if all reconstructed particles were produced in the same decay and their momenta were also reconstructed correctly. Figure 5.2(b) shows the distribution of the cosine of the DIRA for B_s^0 mesons and for background.

Another way to take advantage of the long lifetime of B mesons in the selection is cutting directly on the flight distance (FD) of the decay particles (see Figure 5.3(a)). For this selection, the radial flight distance (RFD) of the B_s^0 candidate, which is the flight distance in the xy-plane, is used. In combination with the other

(a) DIRA



(b) DIRA distribution

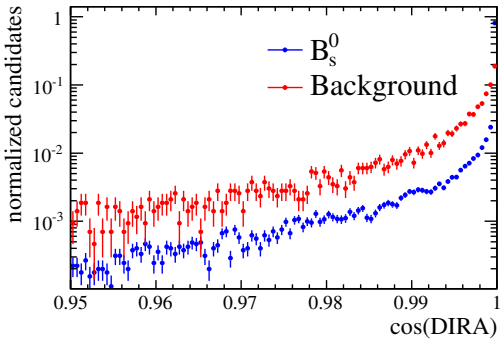


Figure 5.2.: *(a) Diagram for the definition of the direction angle (DIRA) and (b) distribution of the cosine of the DIRA for B_s^0 mesons (blue) and background (red)*

variables used in the selection it shows a better separation between signal and background than the three dimensional flight distance. In Figure 5.3(b) the radial flight distance for B_s^0 mesons and for background is shown.

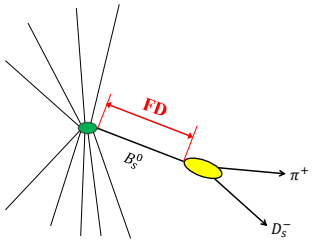
For the bachelor $\pi^\pm$ variables like the transverse momentum (p_T), *i.e.* the momentum in the xy-plane, and the polar angle, θ, between the reconstructed momentum and the z-axis are used in the selection (see Figure 5.4(a)). The distributions of the p_T and the cosine of θ for bachelor $\pi^\pm$ and background are given in Figures 5.4(b) and 5.4(c).

To decide if reconstructed particles originate from a common vertex, a small χ^2/ndf of the vertex fit is required.

Due to timing constraints in the software trigger, no full event reconstruction is performed. Also the decay tree fitter (DTF) is not run in the trigger stage. Therefore, not all of the aforementioned variables can be used to select B_s^0 candidates. Thus, other variables with similar characteristics have to be used like for example instead of a vertex χ^2/ndf the distance of closest approach (DOCA) between two tracks is used (see Figure 5.5(a)). Figure 5.5(b) shows the distributions of these DOCA for signal and background tracks. Signal refers to tracks originating from the same vertex and background refers to particles coming from different vertices.

To distinguish pions, kaons, protons, and muons, particle identification (PID) information from the RICH detector, the calorimeters, and the muon stations are combined. In the LHCb software these are given by the difference of the logarithms of their respective likelihoods to give a measure of the likeliness of the particle being one species relative to it being another. These Delta Log Likelihoods (DLL)

(a) Flight distance



(b) RFD distribution

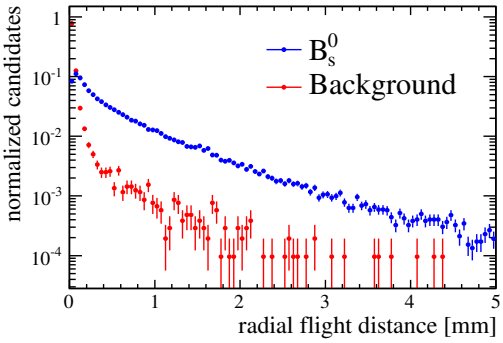
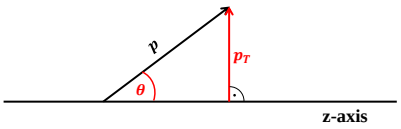
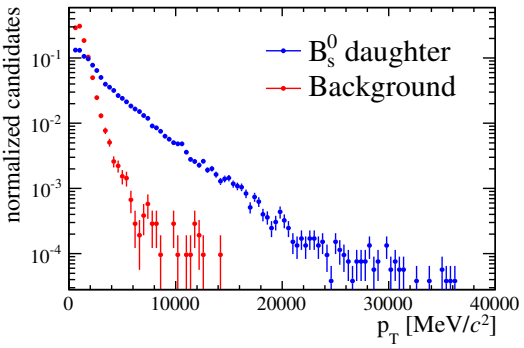


Figure 5.3.: *(a) Diagram for the definition of the flight distance (FD) and (b) distribution of the radial flight distance for B_s^0 mesons (blue) and background (red)*

(a) p_T and θ



(b) p_T distribution



(c) $\cos\theta$ distribution

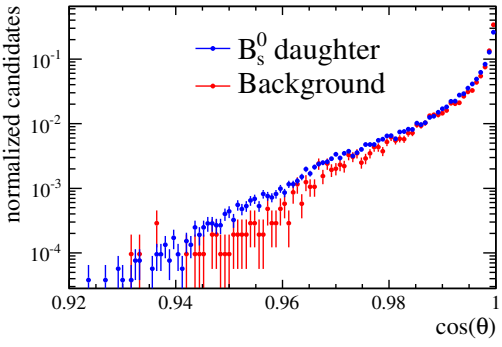
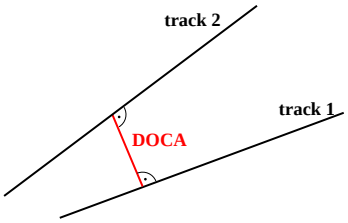


Figure 5.4.: *(a) Diagram for the definition of the transverse momentum (p_T) and the polar angle (θ). Distributions of the p_T (b) and $\cos\theta$ for bachelor $\pi^\pm$ (blue) and background (red)*

(a) Distance of closest approach



(b) DOCA distribution

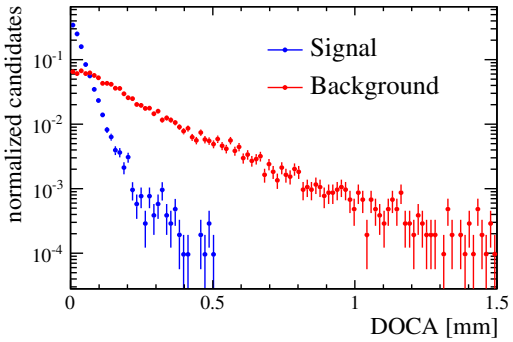


Figure 5.5.: *(a) Diagram for the definition of the distance of closest approach (DOCA) between two tracks and (b) distribution of the DOCA for tracks coming from the same vertex (blue) and tracks coming from different vertices (red)*

are always given relative to the pion hypothesis which is by far the most common particle species in the detector.

$$\Delta \ln \mathcal{L}_{\mathrm{X}\pi} = (\ln \mathcal{L}(\mathrm{X}) - \ln \mathcal{L}(\pi)) \tag{5.1}$$

5.2. Multivariate analysis tools

In the various stages of the signal selection, boosted decision trees (BDT) are used to discriminate signal from background. BDTs are so called multivariate analysis tools, as they combine multiple variables which have discriminating power, and compute one response from them. The response from this multivariate classifier contains all information of the single input variables, but it is combined in a more efficient way than applying single rectangular cuts.

Therefore, a signal selection including multivariate tools usually results in a better performance than simple cut-based selections. The improved performance is shown in Figure 5.6 where the background levels in the nominal signal selection used in this analysis including a BDT and a cut-based selection used in a previous analysis of Δm_s using $B_s^0 \to D_s^- \pi^+$ decays at LHCb [19] are compared. For better comparison only the $D_s^- \to K^+K^-\pi^-$ decay modes have been used. The previous analysis was performed on a smaller data sample. Thus, the absolute number of candidates should not be compared, but only the signal to background ratio. The individual background components will be discussed in detail in Section 5.8.

Multivariate analysis tools are required to be “trained”. Clean samples of signal

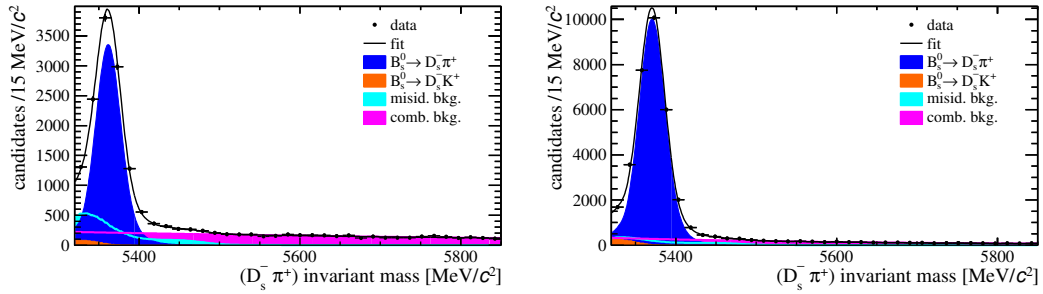


Figure 5.6.: *Comparison of the background level of a cutbased selection (a) and a selection including multivariate tools (b). For comparison only the $D_s^- \to K^+K^-\pi^-$ decay modes have been used.*

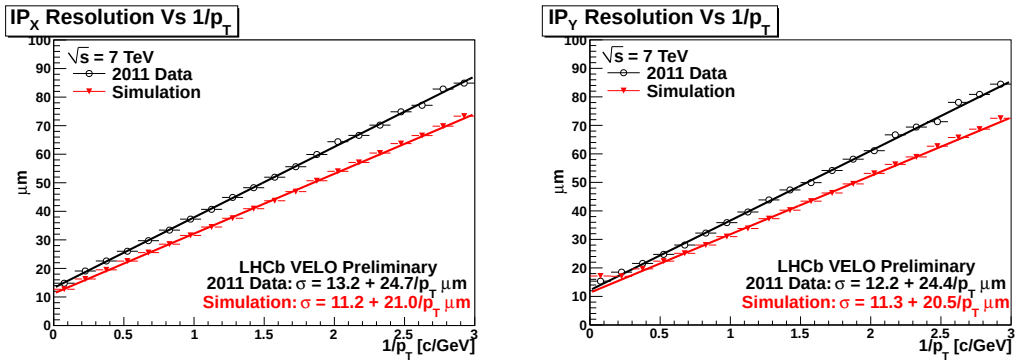


Figure 5.7.: *Impact parameter resolution in x-direction (a) and y-direction (b) against 1./p_{T} of the track, for data and simulation. Taken from [39]*

and background must be provided for the algorithms to find the best way to separate the two using the input variables. These samples can be taken from simulated data, which has the advantage that it is known whether the event is signal or background. The disadvantage is that the simulation does not always reproduce the real data. One known example is that the IP resolution in data is not well reproduced in simulation (see Figure 5.7). The alternative is to take the training samples directly from data. In this case the disadvantage is that the samples are not pure signal/background samples. Also, if they are taken from control channels, it is not guaranteed that they represent the data used in the analysis completely.

Other commonly used multivariate analysis tools include Fisher discriminants, Neural networks and decision trees. For the offline selection, the gradient boosted decision tree (BDTG) shows the best performance and is therefore chosen for the final selection.

5.2.1. Boosted decision trees

Decision trees are sequences of binary splits of the data sample. In each step the variable with the largest separation power is chosen. Then the data is split according to a cut placed on this variable at the value with the best separation which is defined by the training algorithm. Then the resulting two samples are evaluated in the same way. This procedure is repeated until either all resulting samples are split completely into signal and background or the statistics are too small to decide (see Figure 5.8). These final data samples are called "leafs" and are classified to be either signal-like or background-like.

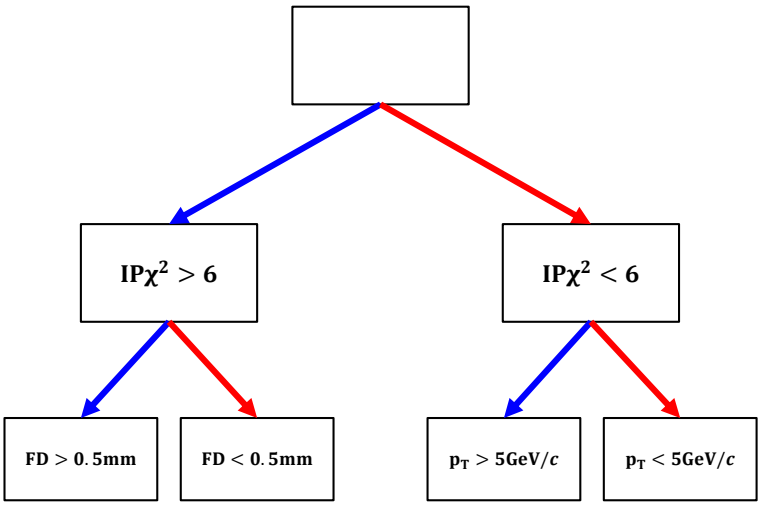


Figure 5.8.: *Diagram illustrating the structure of a decision tree. Blue arrows indicate signal-like branches, red arrows background-like branches.*

Decision trees are powerful classifiers, but unstable. Their separation power depend on the statistical fluctuations of the input samples. Therefore they are made more robust by the use of so-called boosting. Boosting is a procedure in which the optimization procedure is repeated and, in each iteration, events which are wrongly assigned to a leaf, *i.e.* a background event ends up on a signal-like leaf or vice versa, are given larger weights. In this way a large number of decision trees are built and the final response is calculated by looping over all of the trees for each event and add a weight of +1 if the event ended up on a signal leaf and -1 if it ended up in a background leaf. The response is normalized to the total number of trees. This stabilizes the procedure considerably.

Several boosting procedures exist. The two most commonly used are AdaBoost and Gradient Boost. They are described in detail in [59]. The gradient boosting is

found to give the better performance for this analysis and has the advantage that it is more robust against outliers. Therefore it is chosen for the offline selection step.

5.3. Trigger selection

As explained in Section 3.6, the LHCb trigger is a two stage system. In this analysis, no additional selection is applied in the first stage, the L0 trigger components [41]. All events passing this stage are used.

For the second stage, the software trigger (HLT), events used in this analysis are restricted to those which pass certain trigger lines. In addition, it is not only required that the corresponding trigger line selected the event but also that it was activated by tracks belonging to the offline signal candidate. These events are called triggered-on-signal (TOS). Since the composition of trigger lines changed over the course of data taking in 2011, a higher compatibility between data and simulated events is achieved by limiting the analysis to the trigger lines described in this section.

In the HLT1 stage, only events that triggered the HLT1TrackAllL0Line [60] were selected. The cuts applied in this trigger line are listed in Table 5.1.

Due to timing reasons, not all track segments built in the VELO are extrapolated to the tracking stations. Only track segments that pass requirements on the impact parameter, number of hits on the track, and number of missed hits are used. The missed hits variable is given by the difference between the number of hits on the track and the number hits expected which is given by a straight line extrapolation for the track and comparing it to the active region of the VELO. On the full tracks, selection criteria based on the number of hits, momentum , p, and transverse momentum, p_{T}, the χ^2 of the track fit divided by the number of degrees of freedom, and IPχ^2 are applied.

The HLT2 lines used, are comprised of topological trigger lines that select 2-, 3- and 4-body decays and a trigger line that reconstructs ϕ resonances exclusively in the channel $\phi \to K^+K^-$. The topological trigger lines, which are called HLT2Topo2(3,4)BodyBBDTLine [61], apply a very loose cut-based preselection and then use a boosted decision tree to select decays of B mesons. For the preselection, tracks which satisfy certain criteria on their IPχ^2, track χ^2 and momenta are used as input. These tracks are combined to form vertices which must fulfill requirements on the sum of the transverse momenta of the tracks, their invariant mass, their distance of closest approach (DOCA) and their flight distance (FD) χ^2 with respect to the PV. These requirements are shown in Table 5.2.

A boosted decision tree is used in the last step of the trigger selection ($\mathrm{BDT}_{\mathrm{trig}}$).

Cuts on VELO tracks	
IP	$> 100\,\mu\text{m}$
Number of hits	> 9
Number of missed hits	< 3
Cuts on full tracks	
Number of hits	> 16
p_{T}	$> 1700\,\text{MeV}/c$
p	$> 10\,\text{GeV}/c$
χ^2/ndf of track fit	< 2
IP χ^2	> 16

Table 5.1.: *Selection of candidates that pass HLT1TrackAllL0Line.*

Cuts on input tracks	
IP χ^2	> 4
track χ^2	< 3
Transverse momentum p_T^{min}	$> 500\,\text{MeV}/c$
momentum	$> 5\,\text{GeV}/c$
Cuts on 2-,3-,4-body combinations	
$\sum \lvert p_T \rvert$	$> 3, 4, 4\,\text{GeV}/c$
invariant mass	$< 7\,\text{GeV}/c^2$
DOCA	$< 0.2\,\text{mm}$
FDχ^2	> 100
BDT_{trig}	$> 0.4, 0.4, 0.3$

Table 5.2.: *Selection of candidates that pass HLT2Topo2(3,4)BodyBBDTLine.*

Cuts on $K^\pm$	
IPχ^2	> 6
track χ^2	< 5
DOCA	$< 0.2\,\text{mm}$
${p_\text{T}}^{min}$	$> 800\,\text{MeV}/c$
Cuts on ϕ	
$\lvert m - m_{PDG}\rvert$	$< 20\,\text{MeV}/c^2$
p_T	$< 1800\,\text{MeV}/c$
vtxχ^2	< 20

Table 5.3.: *Selection of candidates that pass HLT2IncPhiLine.*

The input variables are the sum of the momenta, $\sum |p_T|$, and the minimum of the transverse momenta of the tracks p_T^{min}, the invariant mass, the corrected mass, DOCA, the candidate IPχ^2 and the flight distance χ^2. The corrected mass takes the missing p_T into account assuming the particle originates from the primary vertex.

$$m_{corr} = \sqrt{m^2 + |p_T^{miss}|} + |p_T^{miss}|, \tag{5.2}$$

where m is the invariant mass of the B candidate. The final cut, depending on whether it is a 2-, 3- or 4-body decay candidate is given in Table 5.2 as well. More details on the topological trigger lines can be found in [61].

Since one of the signal decay channels contains the ϕ resonance, the HLT2IncPhiLine adds signal events to the analysis as well. The selection of the $\phi \to K^+K^-$ candidates is given in Table 5.3. It uses the IPχ^2, the track χ^2, DOCA and p_T^{min} of the final state particles and the invariant mass, p_T and χ^2 of the vertex fit of the ϕ candidate.

5.4. Stripping selection

The candidates that pass the trigger selection described in the Section 5.3 are recorded. On these events, a loose preselection that is performed within the central processing of all data taken by LHCb. This analysis uses the preselection that was applied during the Stripping17 processing of the data.

The selection is listed in Table 5.4. There are also cuts applied on the reconstructed B_s^0, D_s^- and bachelor π^+ candidates.

The cuts on the B_s^0 candidate include the reconstructed invariant mass, $m(B_s^0)$, the χ^2 of the B_s^0-vertex fit from the decay tree fitter (DTF) [48] as well as the impact parameter of the B_s^0, the cosine of the direction angle (cos DIRA), the

Cuts on B_s^0 candidate	
$m(B_s^0)$	$[4.75, 7.0]\,\mathrm{GeV}/c^2$
vtxχ^2	< 10
IPχ^2	< 25
cos DIRA	> 0.999
t	$> 0.2\,\mathrm{ps}$
$\sum p_\mathrm{T}$	$> 5\,\mathrm{GeV}/c$
$\mathrm{BDT}_\mathrm{strip}$	> 0.05
IsTopo cut (Boolean flag if following cuts are passed)	
track χ^2	< 3
p_T	$> 500\,\mathrm{MeV}/c$
p	$> 5\,\mathrm{GeV}/c$
Cuts on D_s^- candidate	
$\lvert m(D_s^-) - m(D_s^-)_{PDG}\rvert$	$< 100\,\mathrm{MeV}/c^2$
vtxχ^2	< 10
vtx distance wrt. PV χ^2	> 36
cos DIRA	> 0
$\sum p_\mathrm{T}$	$> 1.8\,\mathrm{GeV}/c$
at least 1 daughter: IsTopo	True
Cuts on bachelor π^+	
IsTopo	True
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$	< 20

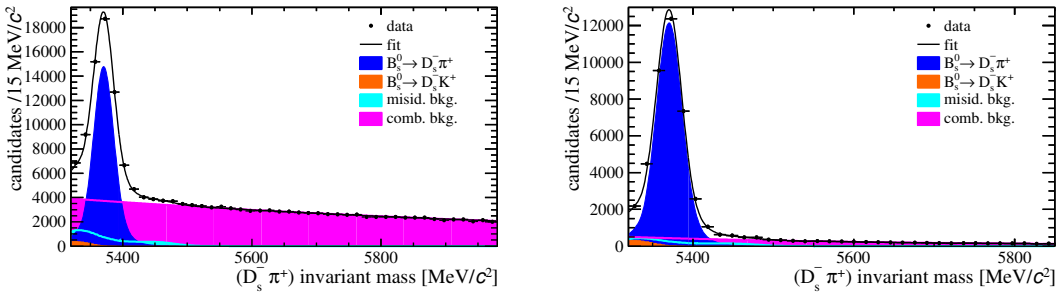
Table 5.4.: *Stripping selection of $B_s^0 \to D_s^- \pi^+$ candidates.*

reconstructed decay time, t, and the sum of the transverse momenta, $\sum p_\mathrm{T}$, of the B_s^0 daughter tracks. Additionally a boosted decision tree ($\mathrm{BDT}_\mathrm{strip}$) is used. The input variables are the B_s^0 transverse momentum and flight distance χ^2 as well as the sum of vertex χ^2 from the DTF, summed over the B_s^0 and D_s^- vertices.

The cuts on the D_s^- candidate include, as well, requirements on the reconstructed invariant mass, the χ^2 of the vertex fit from the DTF, as well as of the distance between the D_s^- decay vertex and the primary vertex. Additionally, it is required to pass cuts on the cos DIRA, and the sum of the transverse momenta, $\sum p_\mathrm{T}$, of its daughters. At least one of the daughter tracks is required to pass the *isTopo* cut which places harder requirements on track χ^2, transverse momentum and momentum. The bachelor π^+ is required to pass the *isTopo* cut as well as a loose cut on the PID separating variable $\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$.

The background level after this step of the selection is already quite moderate, as shown in Fig 5.9(a). The ratio of number of signal candidates to num-

ber of background candidates, in a tight window around the B_s^0 mass [23] of [5320,5450] MeV/c^2, is $S/B = 1.4$. This is further improved in the final step of the selection.



(a) Stripping selection

(b) Offline selection

Figure 5.9.: *Invariant mass distributions for B_s^0 candidates after the Stripping selection step (a) and the complete offline selection (b)*

5.5. BDT in offline selection

BDTs are used in all stages of the selection process (trigger, Stripping and offline). For this analysis, the BDTs that were provided by LHCb were used. The training of the BDTG[1] used in the offline selection will be presented in this section. To be consistent with publications of the LHCb collaboration with the latest results on Δm_s [20] the official LHCb BDTG is used in this analysis.

For the BDTG used in the offline signal selection, the training samples were taken from data. To obtain the signal/background sample, all requirements described in Section 5.4, are applied and then a fit to the mass distribution is performed. From that fit, weights are calculated according to how likely an event is signal or background, using the sPlot method [62]. By applying these weights, the background is statistically subtracted from the signal and vice versa. This procedure has the disadvantage that the background is only subtracted statistically and there is some polution of background events in the signal sample remaining. Another disadvantage is that, if the same sample is used for training the BDT and the final analysis, there is a chance of "overtraining", *i.e.* being sensitive to statistical fluctuations in the signal/background samples. If the BDT is overtrained the estimated signal efficiency does not reflect the real separation power of the BDT. However, since in this analysis the absolute value of the signal efficiency is not of interest, overtraining is not regarded as an issue.

[1] The G refers to the fact that this boosted decision tree uses the Gradient boost algorithm as described in [59]

B_s^0 variables
$\mathrm{IP}\chi^2$
cos DIRA
RFD
vtxχ^2/ndf
decay time χ^2/ndf
Bachelor π^+ variables
$\mathrm{IP}\chi^2$
p_T
cosθ
D_s^- variables
$\mathrm{IP}\chi^2$
cos DIRA wrt. primary vtx
cos DIRA wrt. B_s^0 vtx
RFD
vtxχ^2/ndf
stable particle variables
Minimum(D_s^- daughters p_T)
Minimum(D_s^- daughters $\mathrm{IP}\chi^2$)
Maximum(all stable particles: track ghost probability)

Table 5.5.: *Input variables for the boosted decision tree used in the offline selection (BDTG).*

The input variables for the BDTG are listed in Table 5.5. They include the radial flight distance, RFD, and decay time χ^2/ndf of the B_s^0 candidate, the $\mathrm{IP}\chi^2$ of the B_s^0 daughters, the cosine of the direction angle with respect to the PV and the RFD of the D_s^- which all take advantage of the long lifetime and, thus, the long flight distance of the B_s^0 meson. Additionally, variables like direction angle and vertex χ^2/ndf of the B_s^0 and D_s^- candidates and the $\mathrm{IP}\chi^2$ of the B_s^0 add information about the decay topology to the BDTG. Kinematic variables like the p_T of the bachelor π^+ and the D_s^+ daughters also add to the separation power of the BDTG. Other input variables are the angle between the bachelor π^+ and the beam axis, and a track ghost probability of the stable particles.

5.6. Offline signal selection

The candidates passing the Stripping selection are then used as input for the final offline signal selection. The largest separation power between signal and

background is given by the offline BDT described in Section 5.5. The cuts are summarized in Table 5.6. They include tighter invariant mass windows for the reconstructed B_s^0 and D_s^- candidates. Additionally, the trigger lines of the HLT2 stage described in Section 5.3 are required to be triggered on signal (TOS). For the bachelor π^+, the PID requirement is tightened with repect to the Stripping selection.

In the offline selection, different criteria are applied for the individual D_s^- decay modes. For the $D_s^- \to K^+K^-\pi^-$ modes, a cut to remove events from the decay $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$, where the $\overline{\Lambda}_c^-$ decays into $K^+\overline{p}\pi^-$, is placed: The D_s^- candidate is reconstructed using the proton mass hypothesis for the K^- and if the resulting invariant mass lies within a window around the nominal Λ_c^+ mass, a harder PID requirement is applied to the K^- candidate to reduce the possibility of $\overline{p} \to K^-$ misidentification. A similar requirement is placed for the $D_s^- \to K^-\pi^+\pi^-$ mode.

For the $D_s^- \to \phi(K^+K^-)\pi^-$ and $D_s^- \to K^{*0}(K^+\pi^-)K^-$ modes, cuts on windows around the invariant masses of the corresponding resonances are placed. In case of the $D_s^- \to K^{*0}(K^+\pi^-)K^-$ and the non-resonant $D_s^- \to K^+K^-\pi^-$ modes the PID requirements on the D_s^- daughters are tightened further. These cuts are harder for the same-sign K than for the opposite-sign K. The terms "same-sign" and "opposite-sign" refer to the charge of the K with respect to the D_s charge. To reject background contributions from semileponic B_s^0 decays in the $D_s^- \to K^-\pi^+\pi^-$ decay mode the bachelor pion is required not to fulfill the ISMUON requirement which is a momentum dependent requirement on the presence of hits in the muon stations. Additionally, there are loose PID requirements placed on the $\pi^\pm$ and a hard PID requirement on the K candidate. Also, there is a cut on the χ^2 of the flight distance of the D_s^- candidate with respect to the B_s^0 decay vertex to reject background contributions from other B_s^0 decays.

Finally, for the $D_s^- \to \pi^-\pi^+\pi^-$ decay mode, there is also a loose PID requirement on all reconstructed D_s^+ daughter particles as well as requirements on the invariant masses of the combinations of two of the D_s^+ daughters. In this way, signal enriched regions in the Dalitz plane are taken advantage of. For example the f^0 resonance is selected.

After all selection criteria are applied the remaining background level is very low, as shown in Fig 5.9(b). The signal to background ratio in the reconstructed B_s^0 mass window [5320,5450] MeV/c^2 is $S/B = 7.4$.

5.7. $B_s^0 \to D_s^\pm K^\mp$ signal contribution

After the signal selection is applied, there is a contribution left from $B_s^0 \to D_s^\pm K^\mp$ decays where the bachelor $K^\mp$ is misidentified as a $\pi^\mp$. This decay is considered

Cuts on B_s^0 candidate	
$m_{B_s^0}$	[5320, 5980] MeV/c^2
BDTG	> 0.3
Cuts on D_s^- candidate	
$m_{D_s^-}$	[1944, 1990] MeV/c^2
Trigger requirements	
Triggered on signal (TOS)	Hlt2Topo2(3,4)BodyBBDT OR Hlt2IncPhi
Cuts on bachelor π^+	
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ for pions	< 10
Cuts specific for the $D_s^- \to K^+K^-\pi^-$ modes	
If $\lvert m(D_s^-)_{\overline{\Lambda}_c^- hypo} - m(\overline{\Lambda}_c^-)_{PDG}\rvert < 50$ MeV/c^2:	
$\Delta \ln \mathcal{L}_{\mathrm{Kp}}$ for same-sign K	> 0
Cuts specific for $D_s^- \to \phi(K^+K^-)\pi^-$	
$\lvert m(\phi) - m(\phi)_{PDG}\rvert$	< 12 MeV/c^2
Cuts specific for $D_s^- \to K^{*0}(K^+\pi^-)K^-$	
$\lvert m(K^*) - m(K^*)_{PDG}\rvert$	< 70 MeV/c^2
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ for same-sign K	> 5
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ for opposite-sign K	> 0
Cuts specific for $D_s^- \to K^+K^-\pi^-$ non-resonant	
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ for same-sign K	> 5
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ for opposite-sign K	> 0
Cuts specific for non-resonant $D_s^- \to K^-\pi^+\pi^-$	
ISMUON for bachelor π^+	False
D_s^- FDχ^2 wrt. B_s^0 vertex	> 4
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ for $\pi^\pm$	< 5
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ for K	> 5
If $\lvert m(D_s^-)_{\overline{\Lambda}_c^- hypo} - m(\overline{\Lambda}_c^-)_{PDG}\rvert < 24$ MeV/c^2:	
$\Delta \ln \mathcal{L}_{\mathrm{Kp}}$ for same-sign π	> 0
Cuts specific for $D_s^- \to \pi^-\pi^+\pi^-$	
$\Delta \ln \mathcal{L}_{\mathrm{K}\pi}$ for $\pi^\pm$	< 5
Either | $\lvert m(\pi^+\pi^-) - m(f^0)\rvert$	< 30 MeV/c^2
Or | $m(\pi^+\pi^-)$	> 700 MeV/c^2
Or | If $m(\pi^+\pi_1^-) < 900$ MeV/c^2: $m(\pi^+\pi_2^-)$	< 1000 MeV/c^2

Table 5.6.: *Offline selection applied additional to the Stripping17 selection of the $B_s^0 \to D_s^-\pi^+$ candidates.*

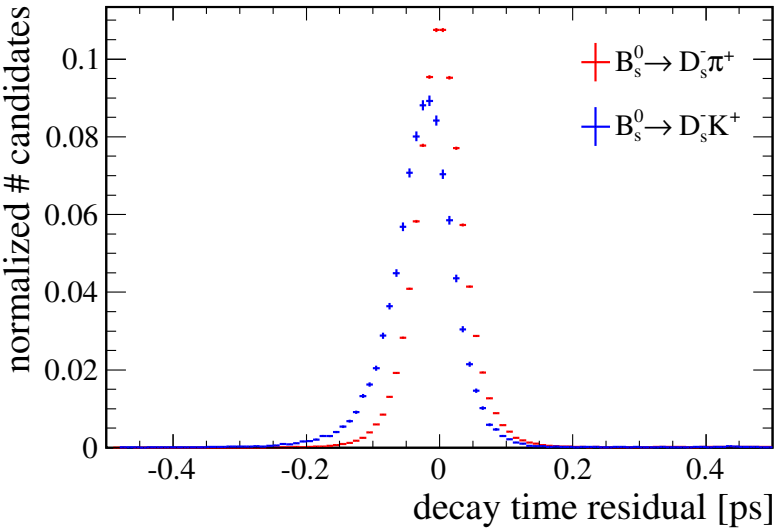


Figure 5.10.: *Distribution of B_s^0 decay time residuals for $B_s^0 \to D_s^- \pi^+$ events (red) and $B_s^0 \to D_s^\pm K^\mp$ events (blue).*

signal in the decay time analysis. Since the mass hypothesis for the bachelor is wrong, the reconstructed momentum and invariant mass of the B_s^0 candidate are wrong as well. The effect on the reconstructed decay time has been studied using simulated data (see Fig. 5.10). It can be seen that the distributions of the decay time residuals which is the difference between the true, simulated decay time and the reconstructed decay time are slightly different from the $B_s^0 \to D_s^- \pi^+$ signal candidates. The distribution is wider and shifted towards lower values. The effect of ignoring this different resolution for $B_s^0 \to D_s^\pm K^\mp$ events has been studied using pseudo-experiments. It shows no effect on the measurement of Δm_s. Therefore it is concluded that it is justified to treat these events as signal in the decay time analysis.

5.8. Background Composition

For this analysis, several sources of backgrounds have been considered. Aside from combinatorial background, which arises from the combination of four random tracks that pass the signal selection, there are two sources of physics backgrounds present. "Physics" refers to the fact that these are real b hadron decays in which one or several particles were misidentified and make the reconstruction of those decays mistakenly consistent with the $B_s^0 \to D_s^- \pi^+$ hypothesis. These backgrounds are henceforth referred to as "misid." background. Additionally, there are contributions from partially reconstructed decays, but they are rejected by an invariant mass requirement on the B_s^0 candidate.

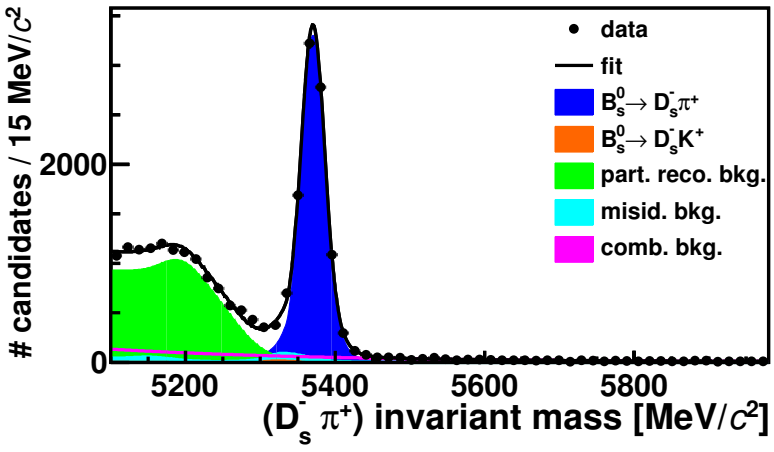


Figure 5.11.: *Distribution of the reconstructed B_s^0 invariant mass in a wide mass window. At the low invariant masses the contribution from partially reconstructed decays can be seen (green).*

5.8.1. Partially reconstructed decays

The decay $B_s^0 \to D_s^- \pi^+$ is possible via higher resonances of the D_s^- and π^+ mesons, like for example the decays $B_s^0 \to D_s^{*-} \pi^+$ and $B_s^0 \to D_s^- \rho^+$, with $D_s^{*-} \to D_s^- \gamma$ and $\rho^+ \to \pi^+ \pi^0$. The neutral particles are not reconstructed, which is why the contributions from these decays are called "partially reconstructed". Due to the missing momentum of the neutral particle, the reconstructed B_s^0 invariant mass is lower than the nominal B_s^0 meson mass (see Figure 5.11).

The partially reconstructed modes are still B_s^0 meson decays, which show the same mixing behaviour as $B_s^0 \to D_s^- \pi^+$ signal decays, and could in principle be included as signal as shown in [63]. However, the gain in statistical significance is small and the decay time distributions for these decays are hard to model due to the missing momentum of the particle that was not reconstructed. Therefore, it has been decided to place a hard cut at the lower edge of the invariant mass signal region at $m > 5320\,\mathrm{MeV}/c^2$ to remove these background contributions.

5.8.2. Combinatorial background

The first source of background considered in this analysis is combinatorial. It results from the random combination of tracks in the event that do not originate from a common particle decay. It is also possible that a real D_s^- meson is combined with a random π^+ which happens to have the correct kinematic properties to make the resulting B_s^0 candidate pass the selection. This background does not pose a problem for this analysis since it is expected to be well described by a smooth,

non-peaking function in the B_s^0 invariant mass distribution. In the selection, the B_s^0 invariant mass window is chosen in a way that selected events with an invariant mass higher than the range of the signal peak (high-mass sideband) are a clean sample of combinatorial background events. This is used to find a fit function parametrization to model this background distribution in the decay time.

5.8.3. Physics backgrounds

The physics backgrounds considered in this analysis are so called misid. backgrounds. This means that they are resulting from b hadron decays with identical decay products as the signal $B_s^0 \to D_s^- \pi^+$ decay, except for one particle. This particle is then misidentified so that the candidate is reconstructed as a signal decay. In principle these decays could be suppressed by harder cuts on the PID variables $\Delta \ln \mathcal{L}_{\mathrm{XX}}$, but due to the limited separation power of these variables, too many signal B_s^0 candidates would be lost. Thus, there are still contributions from these backgrounds present in the final dataset.

$B^0 \to D^- \pi^+$ background

The first contribution to the misid. backgrounds originates from the decay $B^0 \to D^- \pi^+$ with $D^- \to K^+ \pi^- \pi^-$. For $B_s^0 \to D_s^- \pi^+$ decays with $D_s^- \to K^+ K^- \pi^-$, one of the π^- in the decay of the D^- meson can be misidentified as a K^-. Therefore, the reconstructed D^- invariant mass is higher than the nominal D^- mass and falls into the D_s^- invariant mass window of the selection. To reject this background, a harder PID requirement is placed on the "same-sign" K which is the one that is potentially a misidentified π.

In case of $B_s^0 \to D_s^- \pi^+$ decays with $D_s^- \to K^- \pi^+ \pi^-$, this background is present as well due to the double misidentification of a K and a π meson. This mode has the highest background level and the lowest signal yield among the D_s^- decay modes considered in this analysis. Therefore, it was decided to keep the PID requirements loose to maximize the signal yield accepting the higher contribution from this background.

Contributions from this background are assumed not to be present in the sample containing $B_s^0 \to D_s^- \pi^+$ decays with $D_s^- \to \pi^- \pi^+ \pi^-$, since in this case the K^+ meson is misidentified as a π^+ the reconstructed D^- invariant mass is shifted to lower values due to assigning the wrong mass hypothesis to the K^+. Therefore, these candidates do not pass the D_s^- mass cut.

$\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ background

The last misid. background considered in this analysis is the decay $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ where the $\overline{\Lambda}_c^-$ decays into $K^+ \overline{p} \pi^-$. Due to the veto-cuts described in Section 5.6 the contribution from this background is low. Again, in case of the $D_s^- \to \pi^- \pi^+ \pi^-$ mode, the contribution from this background is neglected since the same arguments hold as for the $B^0 \to D^- \pi^+$ misid. background.

CHAPTER 6

Flavour Tagging

As mentioned in Section 4.1, one of the challenges in this analysis is to determine whether the B_s^0 candidate was a B_s^0 or a $\overline{B}_s^0$ at the time of its production, to divide the signal sample into *mixed* and *unmixed* candidates. This information is obtained by so-called flavour tagging algorithms, or short: flavour taggers. There are two conceptionally different approaches how to determine the production flavour of the B_s^0 candidate.

The so-called opposite side taggers (OST) take advantage of the fact that, in pp collisions, b quarks are predominantly produced in $b\overline{b}$ pairs (see Figure 3.2). Therefore, there must be a second b hadron in addition to the signal B_s^0 candidate with the opposite b flavour produced in the primary interaction. Due to the strong boost in z-direction most of the time both b hadrons are in the geometrical acceptance of the LHCb detector (see Figure 3.3). By partially reconstructing this second B hadron the flavour of the signal B_s^0 can be determined.

The so-called same side tagging algorithm (SST) uses the fact that the net strangeness in the pp collisions is zero. Therefore, the s quark used in the hadronization of the signal B_s^0 must originate from an $s\overline{s}$ pair produced from the vacuum. Analogous to the opposite side tagging, the second s quark in the event has the opposite flavour to the s quark in the signal B_s^0 meson. Therefore, by reconstructing the hadron containing the second s quark in the vicinity of the signal B_s^0, the flavour of the signal B_s^0 can be determined.

The flavour tagging algorithms perform independently of the decay products of the reconstructed B_s^0 candidate. Therefore, it is assumed that the performance in

$B_s^0 \to D_s^\pm K^\mp$ decays is the same as for signal decays.

6.1. Flavour tagging quantities

The flavour tagging algorithms estimate whether a B_s^0 candidate was, at production, a B_s^0 or $\overline{B}_s^0$ meson. This response is combined with the charge of the reconstructed decay particles to form the mixing decision q. The latter decision is relevant for this analysis, stating if the signal B_s^0 is mixed ($q = +1$) or unmixed ($q = -1$).

Flavour tagging algorithms do not work perfectly. The performance of these algorithms can be described by two quantities: firstly the tagging efficiency

$$\varepsilon_{\text{tag}} = \frac{\mathrm{N_R} + \mathrm{N_W}}{\mathrm{N_R} + \mathrm{N_W} + \mathrm{N_U}}, \tag{6.1}$$

is calculated using the number of correctly tagged candidates $\mathrm{N_R}$, wrongly tagged candidates $\mathrm{N_W}$ and untagged candidates $\mathrm{N_U}$. The tagging efficiency is the fraction of candidates for which the algorithms give a decision. Depending on the algorithm, ε_{tag} usually lies between 15% and 30%. The second benchmark for the algorithms is the so-called mistag probability

$$\omega = \frac{\mathrm{N_W}}{\mathrm{N_R} + \mathrm{N_W}}, \tag{6.2}$$

which gives the fraction of tagged candidates for which the given tagging decision is wrong. This quantity usually lies between 30% and 40% at hadron colliders. The two quantities, ε_{tag} and ω, can be combined to be the tagging power or effective tagging efficiency

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}}(1 - 2\omega)^2 = \varepsilon_{\text{tag}}\mathcal{D}^2, \tag{6.3}$$

with the dilution factor

$$\mathcal{D} = 1 - 2\omega. \tag{6.4}$$

The flavour tagging algorithms were optimized to measure decay-time-dependent *CP* asymmetries of processes in which B_s^0 or $\overline{B}_s^0$ mesons decay into a final state f. These have the form

$$\mathcal{A}(t) = \frac{\Gamma(\overline{B}_s^0 \to f)(t) - \Gamma(B_s^0 \to f)(t)}{\Gamma(\overline{B}_s^0 \to f)(t) + \Gamma(B_s^0 \to f)(t)} = \frac{\overline{\Gamma}(t) - \Gamma(t)}{\overline{\Gamma}(t) + \Gamma(t)}, \tag{6.5}$$

with the time dependent decay rates Γ for B_s^0 and $\overline{\Gamma}$ for $\overline{B}_s^0$. Including the effects of wrong tagging decisions, the measured decay rates are

$$\Gamma_m(t) = (1-\omega)\Gamma(t) + \omega\overline{\Gamma}(t) \tag{6.6}$$

$$\overline{\Gamma}_m(t) = (1-\omega)\overline{\Gamma}(t) + \omega\Gamma(t). \tag{6.7}$$

which leads to the measured asymmetry

$$\mathcal{A}_m = \frac{\overline{\Gamma}_m(t) - \Gamma_m(t)}{\overline{\Gamma}_m(t) + \Gamma_m(t)} = \mathcal{D}\mathcal{A}. \tag{6.8}$$

with the uncertainty

$$\sigma_{\mathcal{A}_m} = \mathcal{D}\sigma_{\mathcal{A}} \tag{6.9}$$

Using Gaussian error propagation

$$\sigma^2_{\mathcal{A}_m} = \left(\frac{\partial \mathcal{A}_m}{\partial \Gamma_m}\right)^2 \sigma^2_{\Gamma_m} + \left(\frac{\partial \mathcal{A}_m}{\partial \overline{\Gamma}_m}\right)^2 \sigma^2_{\overline{\Gamma}_m} \,, \tag{6.10}$$

and assuming Poisson distributions for $\overline{\Gamma}_m$ and Γ_m

$$\sigma_{\Gamma_m} = \sqrt{\Gamma_m} \qquad \text{and} \qquad \sigma_{\overline{\Gamma}_m} = \sqrt{\overline{\Gamma}_m} \,, \tag{6.11}$$

the uncertainty on the measured asymmetry is given by

$$\sigma^2_{\mathcal{A}_m} = \frac{4\Gamma_m \overline{\Gamma}_m}{(\Gamma_m + \overline{\Gamma}_m)^3} = \frac{1 - \mathcal{A}^2_m}{\Gamma_m + \overline{\Gamma}_m}, \tag{6.12}$$

using $1 - \mathcal{A}^2_m = \frac{4\Gamma_m \overline{\Gamma}_m}{(\Gamma_m + \overline{\Gamma}_m)^2}$ in the second step. With $N^{\text{tag}} = \Gamma_m + \overline{\Gamma}_m = \varepsilon_{\text{tag}} N$ (the number of events with tagging information available), the uncertainty becomes

$$\sigma^2_{\mathcal{A}_m} = \frac{1 - \mathcal{A}^2_m}{\varepsilon_{\text{tag}} N}. \tag{6.13}$$

Using Equation 6.9 and 6.8, the uncertainty on the true asymmetry is given by

$$\sigma^2_{\mathcal{A}} = \frac{1 - \mathcal{D}^2\mathcal{A}^2}{\varepsilon_{\text{tag}} N \mathcal{D}^2} = \frac{1}{\varepsilon_{\text{tag}} N} \left(\frac{1}{\mathcal{D}^2} - \mathcal{A}^2\right). \tag{6.14}$$

$\mathcal{D}$is typically 20–40% at hadron colliders and the decay-time-dependent *CP* asymmetries are usually of the order of a few percent. Thus, the assumption $\frac{1}{\mathcal{D}^2} \gg \mathcal{A}^2$ can be made and the uncertainty $\sigma_{\mathcal{A}}$ becomes

$$\sigma_{\mathcal{A}} = \frac{1}{\sqrt{\varepsilon_{\text{tag}} \mathcal{D}^2 N}} = \frac{1}{\sqrt{\varepsilon_{\text{eff}} N}}. \tag{6.15}$$

Thus, the effective tagging efficiency, ε_{eff}, is the factor by which the statistical power of a data sample in a measurement of a decay-time-dependent *CP* asymmetry is reduced due to imperfect tagging.

6.2. Opposite Side tagging algorithms

As mentioned before, the opposite side taggers take advantage of the predominant $b\bar{b}$ quark pair production by partially reconstructing the second b hadron, called B_{opp}, in the event [64]. The particles used in the reconstruction of the signal decay are excluded from the selection of the tagging candidates.

There are two different approaches to determine the B_{opp} flavour. The first is to reconstruct single particles from the B_{opp} decay chain, like muons and electrons from semileptonic decays, and kaons from $b \to c \to s$ transitions. The second approach is to inclusively reconstruct the B_{opp} decay vertex and sum the charges of all tracks belonging to the vertex. The former are called single particle taggers, the latter, vertex charge tagger. Both are described in detail in the following sections. Figure 6.1 shows the principle of the opposite side tagging algorithms.

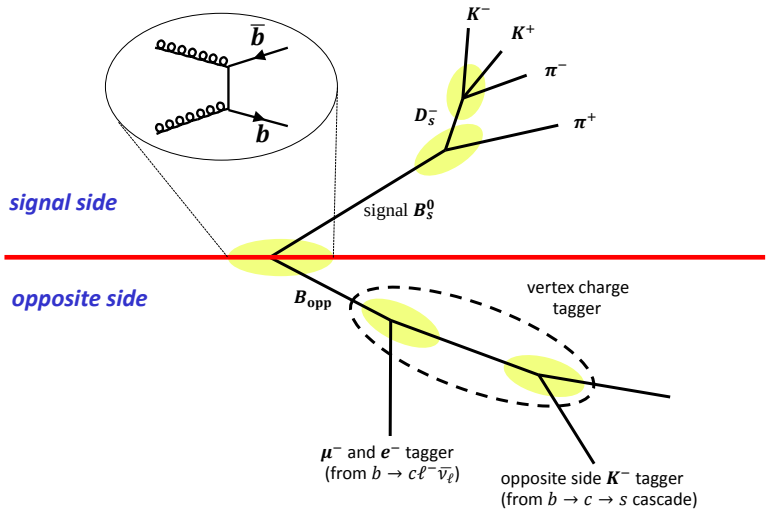


Figure 6.1.: *The diagram shows the principle of the opposite side tagging algorithms. The produced $b\bar{b}$ pair is divided into signal side and opposite side. The opposite side B hadron is partially reconstructed and from the charge of the reconstructed particles the flavour of the signal B_s^0 at production is determined.*

For the opposite side taggers, it is assumed that the B_{opp} flavour at decay is the same as at production. Thus by determining the flavour of the B_{opp} at the decay the signal B_s^0 flavour at production can be identified. This assumption leads to an intrinsic dilution, if the B_{opp} is neutral and changed flavour between production and decay.

6.2.1. Single particle taggers

The single particle tagging algorithms attempt to determine the flavour of the B_{opp} hadron (and thus the flavour of the signal B_s^0 candidate at production) by reconstructing single daughter tracks which are directly linked to the B_{opp} decay flavour. Tracks taken into account as B_{opp} daughter tracks are called tagging candidates. To reject background contributions from particles from the primary interaction only tagging candidates with a large IPχ^2 are considered. There are three individual single particle taggers used at LHCb.

First there is the opposite side kaon tagger which reconstructs candidates from the $b \to c \to s$ transition. Thus, a K^- corresponds to a b quark in the B_{opp} hadron and respectively a K^+ corresponds to a $\overline{b}$ in the B_{opp} hadron. The $K^\pm$ candidates are selected using requirements on the PID information to distinguish them from $\pi^\pm$.

Secondly, there are two lepton taggers which reconstruct candidates from semileptonic $b \to c\ell^-\overline{\nu}_\ell$ decays. Similarly to the opposite side kaon tagger the charge of the reconstructed lepton (muon or electron) identifies the decay flavour of the B_{opp} candidate. The contribution from semileptonic $b \to c \to s\ell^+\nu_\ell$ decays are reduced by placing a minimum p_{T} cut on the lepton tagging candidate.

Muons are selected by requiring hits in the muon chambers and RICH information. Additionally, misidentified $\pi^\pm$ which share hits with true muons are rejected. This cut is especially useful in signal decays containing muons in the final state. Electron candidates are selected using information from the electromagnetic calorimeter. Furthermore, electron candidates are required to be in the acceptance of the hadronic calorimeter to reject contributions from $\pi^\pm$. They are also required to have deposited less than a maximum ionization charge in the VELO to reject contributions from photon conversions.

The lepton taggers are naturally limited by the Branching Fraction of semileptonic B decays which is about 10%. They have usually a low tagging efficiency ε_{tag} while having also a low mistag probability ω and, thus, a high purity.

6.2.2. Vertex Charge tagger

The vertex charge tagger is based on the inclusive reconstruction of the B_{opp} decay vertex. The algorithm starts by forming a seed two-track vertex that is not compatible with the PV by requiring large IPχ^2 for both tracks. Additional requirements on kinematic and geometric variables are applied to ensure that the selected tracks originate from a b hadron decay [65]. Afterwards, more tracks are added to the seed vertex if they are not compatible with the PV (large IPχ^2 with

respect to PV) and at the same time are compatible with the seed vertex (small IP and DOCA with respect to the seed vertex).

After all tracks in the event that pass the required selection are added to the B_{opp} vertex, the tagging decision is made based on the weighted vertex charge which is calculated as

$$\mathcal{Q}_{\text{vtx}} = \frac{\sum_i {p_{\text{T}}}^{\kappa}(i)\mathcal{Q}_i}{\sum_i {p_{\text{T}}}^{\kappa}(i)}. \tag{6.16}$$

Each track included in the B_{opp} vertex is weighted by its transverse momentum to the power of a calibration parameter κ. The value for this parameter κ that maximizes ε_{eff} of this algorithm has been found to be $\kappa = 0.4$ [65].

For the tagging decision only candidates with $|\mathcal{Q}_{\text{vtx}}| > 0.25$ are considered. The resulting tagging decision is equal to the sign of $\mathcal{Q}_{\text{vtx}}$.

Typically this algorithm has a high tagging efficiency of about 15% while also having a large mistag probability ω.

6.3. Same Side Kaon tagger

The same side kaon tagger determines the signal B_s^0 flavour by exploiting its hadronization. Since the s quark present in the signal B_s^0 must be created from the vacuum there has to be an $\overline{s}$ quark present in the event. In about 50% of the cases this $\overline{s}$ quark forms a charged kaon. Figure 6.2 shows the corresponding Feynman diagram.

Contrary to the opposite side tagging algorithms, in this case, tracks from the PV are selected by requiring small IPχ^2. This also minimizes the correlation between opposite and same side taggers since they are using samples of tracks with almost no overlap. However, the small IPχ^2 poses the largest challenge for this tagger at hadron colliders. Finding the correct kaon is difficult because the track density around the PV is high.

To select the correct $K^{\pm}$ from the underlying event, only tracks which are close to the signal B_s^0 candidate in phase space are selected by requiring small differences in polar angle, ϕ, and pseudo-rapidity. Low momentum background is rejected by a minimum requirement on the tagging candidate's p_{T}. Background contributions from misidentified $\pi^{\pm}$ are reduced by hard requirements on the PID variable $\Delta \ln \mathcal{L}_{\text{K}\pi}$.

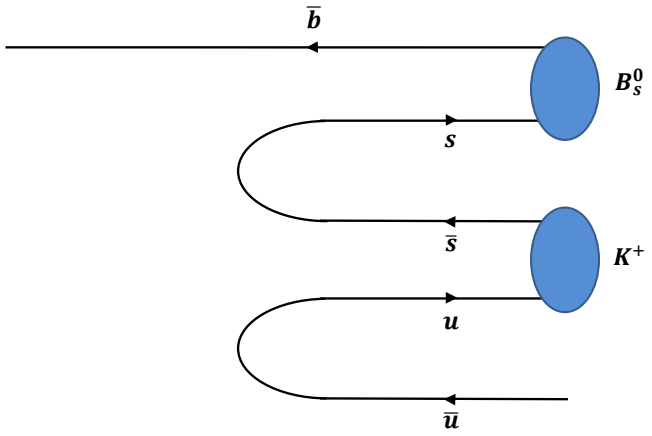


Figure 6.2.: *Feynman diagram of the fragmentation chain exploited in the same side kaon tagger.*

6.4. Physical sources of mistags

There are several reasons for wrong tagging decisions of the opposite side tagging algorithms. Firstly, if the second b quark hadronizes into a neutral B meson, which happens in about 50% of the cases, the B_{opp} meson can oscillate which will lead to a wrong tagging decision. The fraction of B^0 mesons in all B_{opp} is about 40%. These still hold some information because only 19% change flavour before decay. This is different for the quickly oscillating B_s^0 mesons. They make out 10% of all B_{opp}, but half of them change flavour before decaying. Therefore they hold no information about the original B_{opp} flavour. This is relevant for all opposite side taggers.

Secondly, for the single particle taggers, if a track from the underlying event is selected, this track will have a random charge leading in 50% of the cases to a wrong tagging decision. For the lepton taggers, if the B_{opp} meson decays through a $b \to c \to s\ell^+$ transition instead of direct semileptonic decay, the resulting lepton will have the wrong charge. This will also lead to a wrong tagging decision.

If additional tracks, which did not originate from the B_{opp} decay, are included in the vertex charge tagger, this might lead to a wrong tagging decision. This can also be true if one true daughter particle from the B_{opp} hadron is missed in the reconstruction.

The main source for wrong tagging decisions by the same side kaon tagger is if

the associated s quark does not form a charged kaon. Then the particle selected by the algorithm will give a wrong tagging decision in 50% of the cases.

Table 6.1 shows the performance of the individual opposite side taggers in the control channel $B^+ \to J/\psi\, K^+$ taken from [64]. For the combination of the opposite side taggers events with poor mistags have been removed to increase the effective efficiency ε_{eff}. Furthermore, the responses from the individual opposite side taggers are correlated. Thus, the efficiency is smaller than the sum of the single taggers.

Taggers	ε_{tag} [%]	ω [%]	ε_{eff} [%]
μ	4.8±0.1	29.9±0.7	0.77±0.07
e	2.2±0.1	33.2±1.1	0.25±0.04
K	11.6±0.1	38.3±0.5	0.63±0.06
Q_{vtx}	15.1±0.1	40.0±0.4	0.60±0.06
OS average	17.8±0.1	34.6±0.4	1.69±0.10

Table 6.1.: *Tagging performance of the opposite side tagging algorithms in $B^+ \to J/\psi\, K^+$ decays. Taken from [64].*

The same side kaon tagger has a tagging efficiency of $\varepsilon_{\text{tag}} = 15.8 \pm 0.3$ and an effective tagging efficiency of $\varepsilon_{\text{eff}} = 1.5 \pm 0.4$ in $B_s^0 \to D_s^- \pi^+$ events. These results are taken from [66].

6.5. Calibration of flavour taggers

The individual flavour tagging algorithms give, in addition to the tagging decision, q, also an estimate, η, for the mistag probability for each B_s^0 candidate. This estimate is the output of a neural net with several event variables as input. The estimate has to be calibrated to represent the actual mistag probability ω. This is done in external analyses using high statistics control channels. Since the performance of the opposite side taggers is largely independent of the signal decay mode, it is possible to calibrate them in non-oscillating B^+ decay modes like $B^+ \to J/\psi\, K^+$ [64]. The advantage is that the flavour at decay is identical to the flavour at production for these B^+ candidates. Thus it can be directly determined if the tagging decision was correct by means of the charges of the decay products.

For the relation between the estimated mistag probability η and the actual mistag probability ω a linear dependency is assumed.

$$\omega = p_0 + (\eta - \langle\eta\rangle) \cdot p_1. \tag{6.17}$$

This function is chosen to have the most simple parametrization with minimal correlation between the calibration parameters p_1 and p_0. Figure 6.3, taken

from [64], shows the calibration for $B^+ \to J/\psi K^+$ decays. It can be seen that the linear parametrization describes the relation between estimated mistag η and actual mistag ω well.

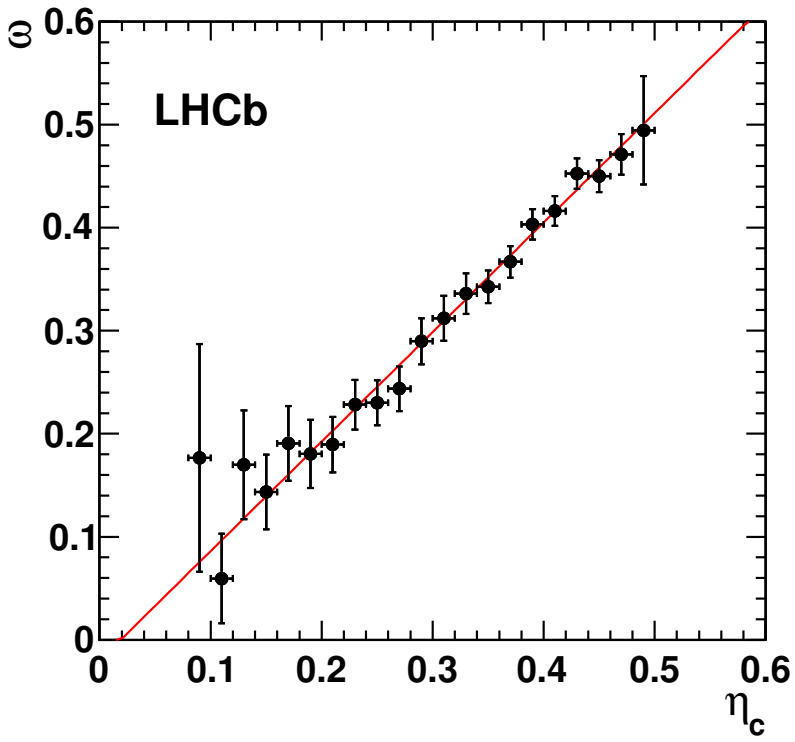


Figure 6.3.: *Actual mistag probability ω plotted as a function of the estimated mistag probability η. Taken from [64].*

The same side kaon tagger depends on the signal fragmentation and thus has to be calibrated in a B_s^0 decay mode. This means that the B_s^0–$\overline{B}_s^0$ oscillation has to be fitted and from the fitted amplitude of the oscillation the calibration parameters are determined. The B_s^0 decay mode that is suited best for this is $B_s^0 \to D_s^- \pi^+$ [66]. Since the data sample used for this calibration has a large overlap with the sample used in this analysis it was decided not to constrain the calibration parameters p_0 and p_1 to the values from the external calibration, but to include them in the PDFs described in Section 9.1 and leave them free in the fit. For consistency, the same is done for the calibration parameters of the opposite side taggers. However, these parameters are closely related to the amplitude of the oscillation which is orthogonal to the frequency. Therefore, they are uncorrelated to Δm_s and leaving them free does not influence the final result (see Appendix D).

6.6. Combination of flavour taggers

The information from the decisions, q_i, and mistag estimates, η_i, of the individual flavour tagging algorithms are combined into a single decision q and estimate η. This is done in a way that no information is lost from the individual taggers and they are assumed to be uncorrelated. This assumption is true for all flavour tagging algorithms except for the vertex charge tagger which is found to be correlated with the single particle opposite side taggers. This correlation is ignored and therefore, even though the individual taggers were calibrated, the combination of opposite side taggers has to be calibrated as well.

The combination is only relevant if more than one algorithm gave a tagging decision q. In this case their predicted η_i are combined using the following formula:

$$p(q=-1) = \prod_i \left(\frac{1+q_i}{2} - q_i(1-\omega_i)\right), \tag{6.18}$$

$$p(q=+1) = \prod_i \left(\frac{1-q_i}{2} + q_i(1-\omega_i)\right), \tag{6.19}$$

with ω_i being the calibrated mistag probability of the individual tagger

$$\omega_i = p_{0,i} + (\eta_i - \langle\eta_i\rangle) + p_{1,i}, \tag{6.20}$$

The quantities $p(q=-1)$ and $p(q=+1)$ give measures of how likely a candidate was unmixed or mixed, respectively. These measures have to be normalized to be probabilities:

$$\begin{aligned} \mathcal{P}(q=-1) &= \frac{p(q=-1)}{p(q=-1)+p(q=+1)}, \\ \mathcal{P}(q=+1) &= 1-\mathcal{P}(q=-1). \end{aligned} \tag{6.21}$$

The combined decision, q, is made according to which probability is larger:

$$\begin{aligned} q &= +1 \qquad \text{for } \mathcal{P}(q=+1) > \mathcal{P}(q=-1) \\ q &= -1 \qquad \text{for } \mathcal{P}(q=-1) > \mathcal{P}(q=+1) \end{aligned} \tag{6.22}$$

The combined estimated mistag probability is set to $\eta = 1-\mathcal{P}$ of the corresponding combined tagging decision. At the time this analysis was conducted, the calibration of the opposite side taggers has been completed, but the same side kaon tagger has not yet been calibrated. Thus for the opposite side taggers the combination was used, but the same side kaon tagger was used separately. The tagging decisions q^{OST} and q^{SST} and estimated mistag probabilities η_{OST} and η_{SST} of the combined opposite side taggers and the same side kaon taggers are combined in the way explained above and the calibration parameters p_0^{OST}, p_0^{SST}, p_1^{OST} and p_1^{SST} are free parameters in the fit.

6.7. Flavour Tagging for background events

The flavour tagging decisions are designed to be meaningful only in case of a real $B_s^0 \to D_s^- \pi^+$ signal decay. However, this doesn't mean that the decisions are completely random in case of background decays. This has to be taken into account for relative normalization purposes of the samples tagged as mixed or unmixed.

6.7.1. Combinatorial background

Since the combinatorial background consists of random combination of particles not originating from a single decay, the flavour tagging algorithms are assumed to give random responses without information for this background contribution. However, the fraction of tagged events do not have to be the same as for signal decays. Since combinatorial background candidates that pass the signal selection are more likely to occur in events with a high particle density and such events are also more likely to produce a tagging candidate, the tagging efficiencies for combinatorial background candidates are expected to be larger than for signal candidates. To check this, additional tagging efficiencies for the opposite side taggers $\varepsilon^{\mathrm{OST}}_{\mathrm{tag;comb}}$ and the same side kaon tagger $\varepsilon^{\mathrm{SST}}_{\mathrm{tag;comb}}$ are free parameters in the fit.

Additionally, a parameter is introduced to allow for different numbers of candidates which are tagged as mixed or unmixed. These asymmetry parameters, $\omega^{\mathrm{OST}}_{\mathrm{comb}}$ and $\omega^{\mathrm{SST}}_{\mathrm{comb}}$, for opposite and same side taggers, are, as well, free parameters in the fit, but are expected to be compatible with 0.5.

6.7.2. B^0 background

The background contribution from $B^0 \to D^- \pi^+$ decays must be treated differently for candidates tagged by the opposite side taggers and candidates tagged by the same side kaon tagger.

For the opposite side taggers this background is treated as signal since B^0 mesons contain real b quarks and it has a flavour specific final state. Therefore, the opposite side flavour tagging algorithms are assumed to be able to distinguish between mixed and unmixed decays. However B^0 mesons oscillate with a lower frequency than B_s^0 mesons. This is taken into account in the mixing PDF of this background.

For the same side kaon tagger the tagging behavior of this background is assumed to be random because B^0 mesons do not contain an s quark and thus there cannot

be a charged kaon from the fragmentation chain. This background consequently is treated in the same way as the combinatorial background, with floating parameters $\varepsilon_{B^0}^{\mathrm{SST}}$ and $\omega_{B^0}^{\mathrm{SST}}$ for the efficiency and tagging asymmetry. As for the combinatorial background, the asymmetry parameter, $\omega_{B^0}^{\mathrm{SST}}$, is expected to be compatible with 0.5.

6.7.3. Λ_b^0 background

The background contribution from $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays are assumed to have the same tagging behavior as the $B^0 \to D^- \pi^+$ decays. For the opposite side taggers it has the same tagging behavior as the signal sample. The only difference is that Λ_b^0 baryons do not oscillate into their antiparticles. Therefore they are all unmixed and the only way they are tagged as mixed is in case of a wrong decision.

For the same side kaon tagger the same argument as for the $B^0 \to D^- \pi^+$ background is valid and the tagging behavior is assumed to be random. Like for the other background contributions, the parameters $\varepsilon_{\Lambda_b^0}^{\mathrm{SST}}$ and $\omega_{\Lambda_b^0}^{\mathrm{SST}}$ are free parameters in the fit, with $\omega_{\Lambda_b^0}^{\mathrm{SST}}$ expected to be compatible with 0.5.

CHAPTER 7

Decay Time Resolution and Acceptance

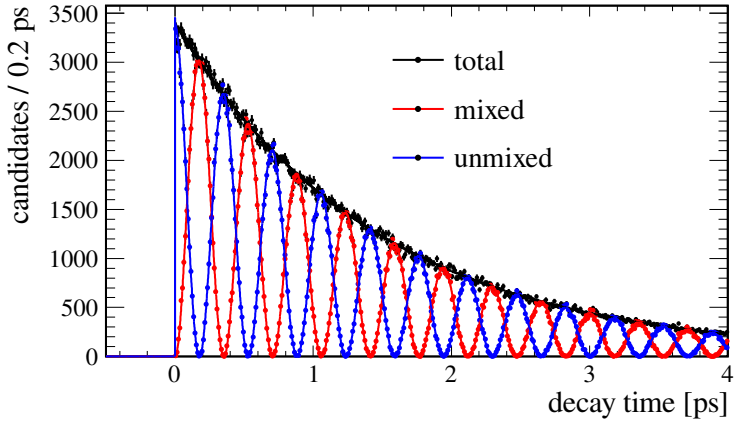
When performing a decay-time-dependent analysis, there are two important issues: the decay time resolution and the decay time acceptance. The decay time resolution dilutes the decay time distribution so that the amplitude of the measured oscillation is reduced. Therefore, an excellent decay time resolution is required to resolve the fast B_s^0 oscillation. This is provided by three effects:

- The B_s^0 particles are highly boosted in the forward direction. This helps to separate secondary vertices from the PV.
- The amount of material the particles will traverse is minimized to reduce effects from multiple scattering.
- The excellent vertex resolution of the VELO detector at LHCb (The IP resolution of the VELO in data is 13 µm in x-direction and 12 µm in y-direction for tracks with high p_{T}, see Figure 5.7)

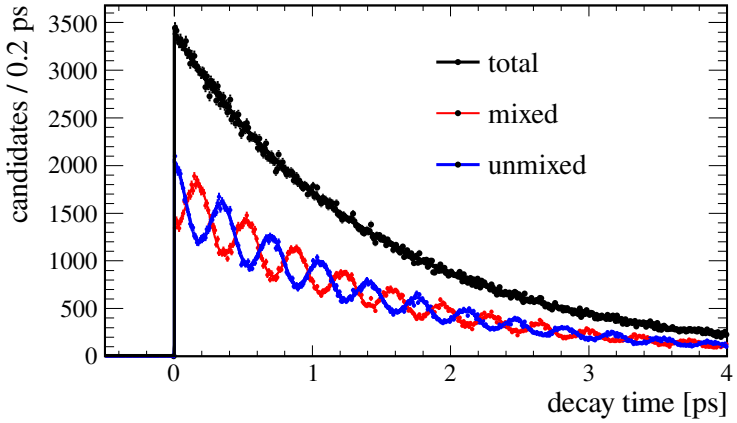
Historically, this analysis was used not only to measure the B_s^0 oscillation frequency Δm_s, but also to calibrate the same side kaon tagger. The decay time resolution as well as the mistag probability is closely related to the amplitude of the oscillation. Figure 7.1 shows the effects on the measured oscillation due to imperfect flavour tagging and decay time resolution. Figure 7.1(a) shows the theoretical mixing distribution. It can be seen that the amplitude of the oscillation is maximal. Figure 7.1(b) shows the effect from realistic flavour tagging[1]. Due to wrong tagging decisions mixed candidates are counted as unmixed and vice versa, dampening

[1]For this simulation the same tagging performance as measured in data was used

(a) Perfect tagging, perfect resolution



(b) Realistic tagging, perfect resolution



(c) Perfect tagging, realistic resolution

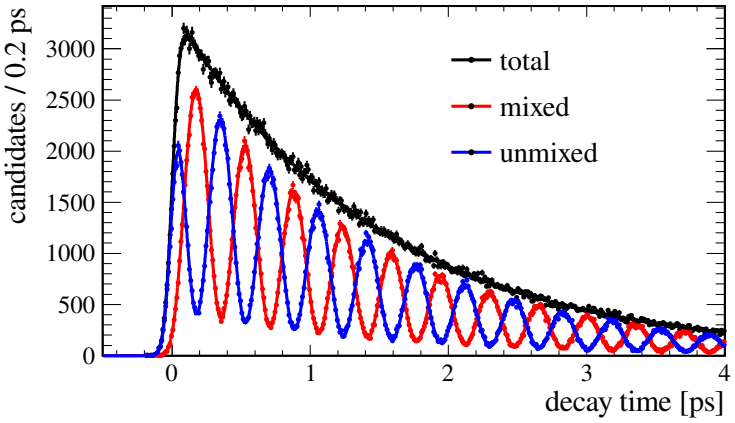


Figure 7.1.: *Simulation showing the effects of flavour tagging and decay time resolution on the measured B_s^0–$\overline{B}_s^0$ oscillation. (a) shows the true oscillation with maximal amplitude. (b) shows the effect of realistic tagging and (c) shows the effect of a realistic decay time resolution on the measured oscillation.*

the oscillation amplitude. The amount of the dampening is equal to the dilution factor $\mathcal{D}$, defined in Equation 6.4.

Figure 7.1(c) shows the effect of a realistic decay time resolution[2]. Due to resolution effects the measured decay time is slightly smaller or larger than the true decay time. This also results in an effective dampening of the amplitude by a dilution factor $\mathcal{D}_{\text{res}}$. The resolution model is used to convolute the theoretical distribution

$$\mathcal{S}(t_{\text{true}}) \propto \cos \Delta m_s t_{\text{true}}. \tag{7.1}$$

A mathematical convolution of a function f with another function g is defined as

$$f(t) \otimes g(t) \equiv \int_{-\infty}^{\infty} f(t) g(t - t') dt' \tag{7.2}$$

Assuming a Gaussian resolution model with mean zero and width σ_{res}

$$\mathcal{S}(t_{\text{true}}) \otimes \mathcal{G}(0, \sigma_{\text{res}}; t) \propto \mathcal{D}_{\text{res}} \cdot \cos \Delta m_s t \tag{7.3}$$

with the dilution factor

$$\mathcal{D}_{\text{res}} = e^{\frac{-\Delta m_s^2 \cdot \sigma_{\text{res}}^2}{2}} \tag{7.4}$$

The measured average decay time resolution in data is about 45 fs. This leads to a dampening factor of $\mathcal{D}_{\text{res}} = 0.73$.

Since both the parameters describing the flavour tagging and the decay time resolution are related to the measured amplitude, it is not possible to fit for both at the same time, and the parameters describing the decay time resolution have to be fixed. To be able to accurately determine the parameters describing the tagging performance, the resolution has to be known very precisely. This is why it is not possible to rely on simulated data and in Section 7.1 the calibration of the decay time resolution model using a data driven study is described. Even though the precise knowledge of the decay time resolution has only a secondary effect on the measurement of Δm_s the same method is used in this analysis.

The second issue for the analysis, the decay time acceptance, describes the distortion of the measured decay time distribution due to the signal selection which was described in Section 5. The determination of its parametrization from simulated data is described in Section 7.2.

[2]For this simulation a resolution of 45 fs, as measured in data, was used.

7.1. Calibration of the Decay Time Resolution

The decay time resolution has to be modelled correctly in the fit for Δm_s to accurately describe the B_s^0–$\overline{B}_s^0$ oscillation. As mentioned in Section 3.7 the decay tree fitter (DTF) [48] takes the particle tracks and fits simultaneously all vertices specified in the decay chain. In this analysis an event-by-event estimate of the uncertainty on the measured decay time provided by the DTF is used. It is calculated using the uncertainties on the particle momenta calculated in the track fit and the uncertainty on the position of the B_s^0 decay estimated in the vertex fit. The latter is the dominant effect. The data driven study described in the following sections has been used to calibrate this event-by-event uncertainty estimate.

7.1.1. Principle of the Calibration Method

The method to calibrate the decay time resolution takes advantage of the fact that the resolution is dominated by the uncertainty on the position of the secondary vertex, *i.e.* the B_s^0 decay vertex. The uncertainty on the position of the primary interaction vertex is much smaller because there are many more tracks coming from the primary interaction that are used to fit this vertex. In contrast to that, there are only four tracks that contribute to the fit of the B_s^0 decay vertex.

The event-by-event error estimate coming from the algorithm that fits the B_s^0 decay time already propagates the uncertainties coming from the track fit of the final state particles. Other detector resolution effects cannot be taken into account, such as imperfect detector alignment and uncertainties on the amount of material the decay particles traverse which leads to multiple scattering. Therefore, the resulting uncertainty estimate must be calibrated on data.

This calibration is performed using reconstructed $D_s^- \to K^+K^-\pi^-$ candidates that do not come from any B meson decay, but from the primary interaction. These D_s^- candidates are then combined with random π^+ from the primary interaction to form so-called fake B_s^0 candidates. These fake B_s^0 candidates have, by construction, decay time $t = 0$, because they were built using only particles from the primary interaction.

Figure 7.2 illustrates the calibration method. In the left diagram, a normally reconstructed signal decay is shown. The B_s^0 meson is produced in the primary vertex and propagates oscillating until it decays into a D_s^- meson and a pion. In the right diagram the fake decay used in the calibration is shown. It is identical to the right part of the real signal decay, but there was no B_s^0 in the first place and the secondary vertex is identical to the primary vertex. Thus, the reconstructed decay time is equal to zero, aside from the resolution effects.

The selection criteria for these fake B_s^0 candidates are chosen to be as close to

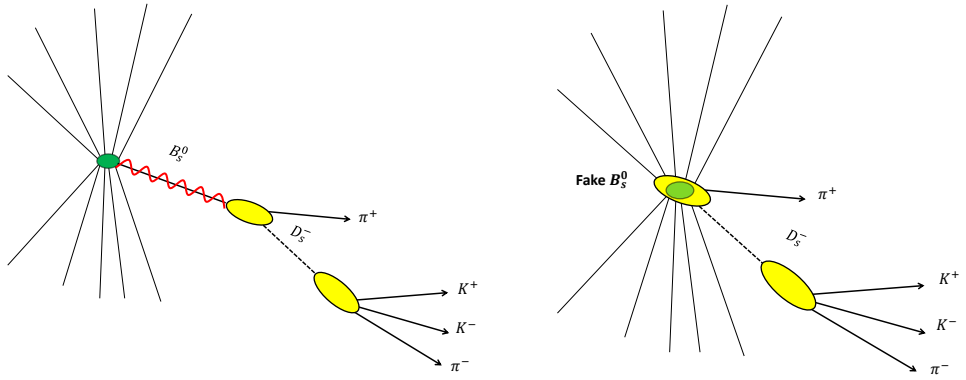


Figure 7.2.: *Diagram of the method used to calibrate the decay time resolution. The left figure shows a normally reconstructed signal $B_s^0 \to D_s^- \pi^+$ decay. The right figure shows a $D_s^- \to K^+K^-\pi^-$ decay from the primary interaction combined with a random π^+ from the primary interaction.*

the signal selection described in Section 5, without applying any cuts that would influence the fake B_s^0 decay time distribution. Examples for such cuts are given in Section 7.2.

It would be ideal to apply all cuts that do not influence the decay time spectrum while omitting the others. Unfortunately, this is not possible since these variables are used in the BDT and cannot be easily ommitted. A different Stripping selection using only variables that leave the B_s^0 decay time spectrum unchanged was available. Therefore, it has to be checked whether the differences in the selections have any influence on the B_s^0 decay time resolution.

Figure 7.3 shows the signal B_s^0 decay time residuals on simulated data for the nominal offline selection (left figure) and the selection used for the calibration (right figure). Both distributions are fitted with Gaussian distribution. It can be seen that the two distributions have the same width. Thus it is concluded that the differences in the selections do not influence the measured decay time resolution and it is justified to use the calibration sample to calibrate the decay time resolution.

There are several other crucial points for the calibration of the decay time resolution. For the fake B_s^0 candidates, it has to be ensured that the sample on which the calibration is performed contains only real D_s^- mesons. This is achieved by statistically subtracting the background from particles other than real D_s^-. A fit to the invariant mass of the D_s^- candidates is performed and weights are computed using the sPlot method [62], as in the training of the BDT for the selection (see Section 5.5).

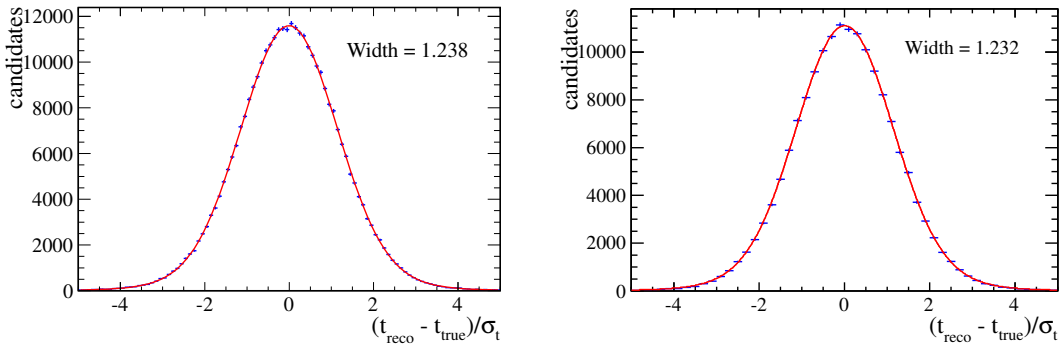


Figure 7.3.: *Decay time residuals for signal B_s^0 candidates from simulated data comparing the nominal signal selection (left) with the selection used for the calibration of the decay time resolution (right).*

This sample of true D_s^- candidates has to be cleaned from D_s^- mesons that were not *prompt* (they were not produced in the primary interaction), but are daughter particles from long-lived particles, like b hadrons. The reason for this is that in these cases the resulting fake B_s^0 candidates do not have decay time $t = 0$ and they must not be used to determine the resolution. To disentangle the prompt D_s^- candidates from the non-prompt D_s^- candidates the D_s^- $log(IP)$ distribution is used. Again, the sample is reweighted using the same method to subtract the long-lived background.

Finally the distribution of the decay time pull is fitted, which is defined as the decay time divided by the event-by-event error estimate, σ_t^i. If the event-by-event error estimate described the real uncertainty on the reconstructed B_s^0 decay time accurately, this distribution would be a Gaussian distribution with mean $\mu = 0$ and width $\sigma = 1$. A width $\sigma = S_{\sigma_t}$ which is greater than one indicates that the error estimate underestimates the real uncertainty. Scaling the error estimate by the fitted width $(S_{\sigma_t} \cdot \sigma_t^i)$ yields to the desired pull distribution.

Due to limited statistics in data this study is actually performed using $D^- \to K^+\pi^-\pi^-$ passing the same selection instead of $D_s^- \to h^-h^+h^-$ decays. $D^- \to K^+\pi^-\pi^-$ decays are kinematically and topologically very similar to $D_s^- \to h^-h^+h^-$ decays. Figure 7.4 shows the $log(IP)$ and decay time pull distributions for $D^- \to K^+\pi^-\pi^-$ and $D_s^- \to K^+K^-\pi^-$ candidates in simulated data. There is no difference visible and it is concluded that it is justified to use $D^- \to K^+\pi^-\pi^-$ decays for the calibration of the B_s^0 decay time resolution.

In the following sections each step required for the calibration of the decay time resolution is be discussed in detail.

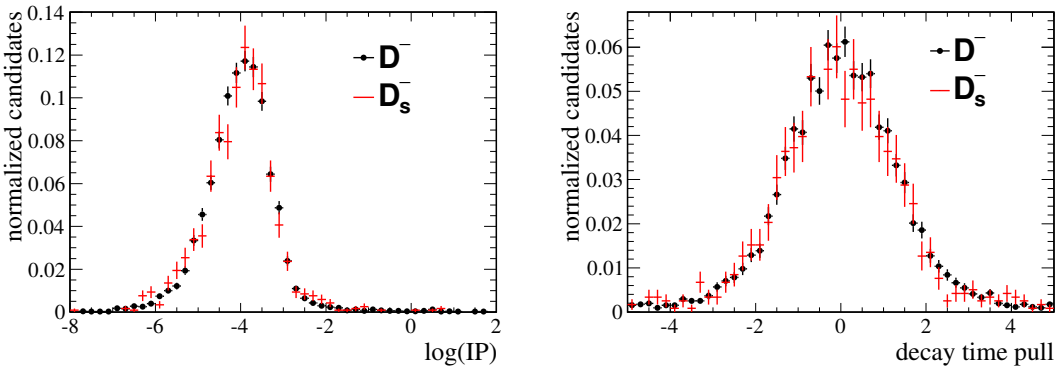


Figure 7.4.: *log(IP) (left) and decay time pull (right) distribution for $D^- \to K^+\pi^-\pi^-$ (red) and $D_s^- \to K^+K^-\pi^-$ (black) candidates in simulated data.*

7.1.2. Fit of the $D^- \to K^+\pi^-\pi^-$ invariant mass distribution

As described in the previous section the first step for the calibration of the decay time resolution is fitting the invariant mass distribution of the $D^- \to K^+\pi^-\pi^-$ candidates to separate signal from combinatorial background.

The fit model to describe the signal candidates in the invariant mass distribution is given by the sum of two Gaussian distributions with a common mean

$$\begin{aligned}\mathcal{P}_{sig,m_{D_s^-}} =& f_{m_{D_s^-}} \cdot \mathcal{G}(\mu_{m_{D_s^-}}, \sigma_{1,m_{D_s^-}}; m_{D_s^-}) \\ &+ (1 - f_{m_{D_s^-}}) \cdot \mathcal{G}(\mu_{m_{D_s^-}}, \sigma_{2,m_{D_s^-}}; m_{D_s^-}).\end{aligned} \tag{7.5}$$

The combinatorial background is described by a first order polynomial with the slope parameter $\alpha_{m_{D_s^-}}$. The fit is performed as a binned extended maximum likelihood fit. Therefore, the yields for signal N_{sig} and background N_{bkg} are fit parameters as well. The fitted invariant mass distribution is shown in Figure 7.5. From this fit result, weights are calculated using the sPlot technique [62] to subtract the combinatorial background from the signal. The weights are applied to the data set in the consecutive steps of the calibration. The fit results and the distribution of the calculated weights are given in Appendix A.

7.1.3. Fit of the D^- $log(IP)$ distribution

As described in Section 7.1.1, the second step of the calibration of the decay time resolution is to select only those $D^- \to K^+\pi^-\pi^-$ candidates which were produced

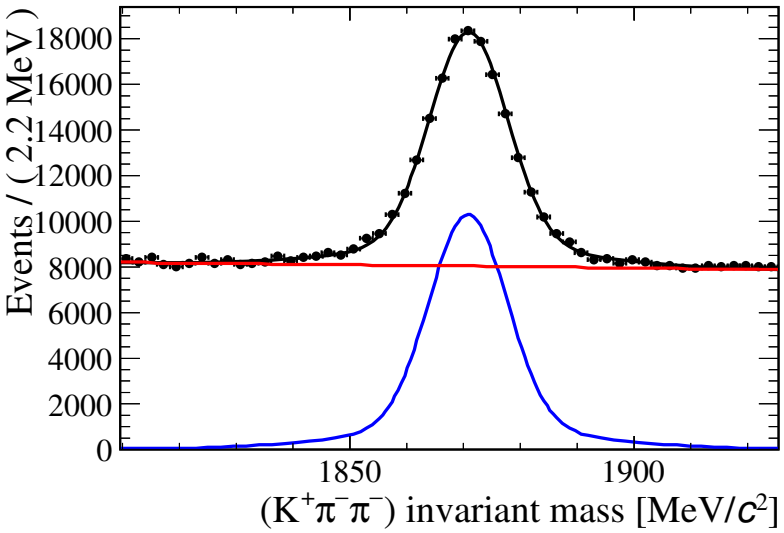


Figure 7.5.: *Distribution of the invariant mass of the $D^- \to K^+\pi^-\pi^-$ candidates used for the calibration of the B_s^0 decay time resolution. The red curve describes the fitted background distribution and the blue curve describes the signal distribution.*

directly in the primary interaction. From now on, these candidates will be referred to as "prompt". The variable used to separate these candidates from those that were part of a decay chain of a long-lived particle, like a b hadron, is the logarithm of the impact parameter, $log(IP)$, of these candidates. The IP gives a measure of how well the reconstructed momentum of the D^- meson points back to the primary vertex, PV. In this case the logarithm of the IP is chosen.

The shape of the $log(IP)$ distribution for prompt and non-prompt D_s^- candidates are parametrized using the so-called Bukin PDF [67]. The complete parametrization of the Bukin PDF is given in Appendix A.

The shape of the prompt D^- and those from long-lived particles are fitted to simulated data. For this study, a sample of 20 million $D^- \to K^+\pi^-\pi^-$ was used. After selection, about 10,000 prompt and 6,700 non-prompt candidates remain. In Figure 7.6, the distributions together with the fitted functions are shown.

For fit stability reasons, the shape parameters ξ, ρ_1 and ρ_2 are fixed to the values obtained on simulated data for the fit on data. Only the peak position parameters, x_p, and the widths, σ_p, are left floating in the fit, since the simulated data are not expected to reproduce these. Figure 7.7 shows the $log(IP)$ distribution of the D^- candidates on data as well as the fitted functions for the prompt component (green, dotted) and the non-prompt component (red, dashed). The figure shows that the distribution is described well by the fitted function.

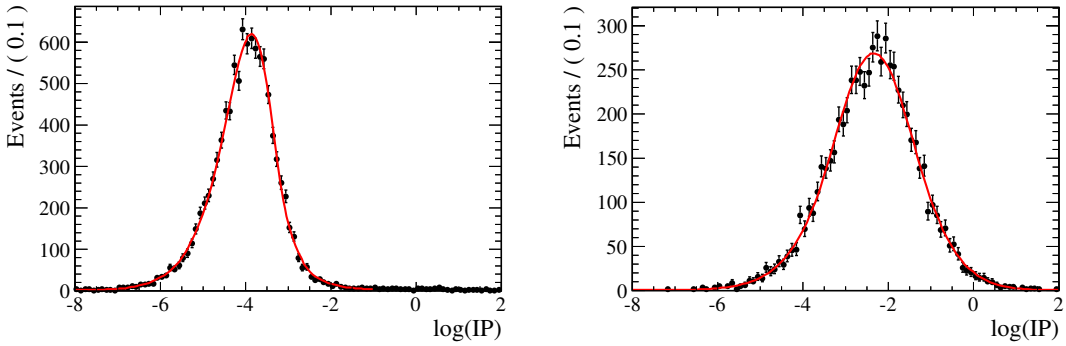


Figure 7.6.: *$log(IP)$ distributions for prompt (left) and non-prompt (right) $D^- \to K^+\pi^-\pi^-$ candidates. The fitted Bukin PDFs are overlaid.*

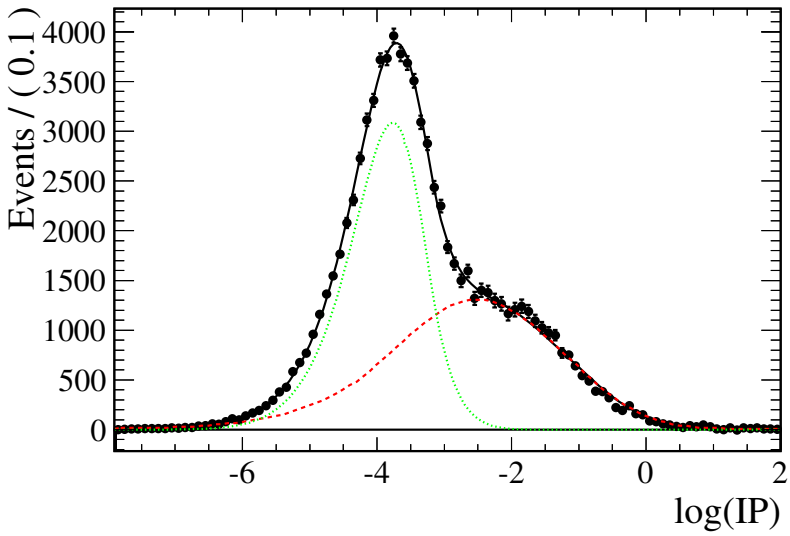


Figure 7.7.: *Weighted $log(IP)$ distribution for $D^- \to K^+\pi^-\pi^-$ candidates, that passed the selection. The fitted PDFs are shown for prompt (green, dotted) and non-prompt (red, dashed) components.*

Similar to, the fit of the invariant mass distribution, weights are calculated using the sPlot technique [62], from the fit to the $log(IP)$ distribution of the D^- candidates to subtract the non-prompt component. Again, the fit results and the distribution of the calculated weights are given in Appendix A. Applying these weights to the data subtracts effectively the non-prompt component. Thus, in the final step, the fit to the distribution of the decay time pull, the fake B_s^0 candidates represent candidates with real decay time $t = 0$.

7.1.4. Fit to the decay time pull of the fake B_s^0 candidates

The resolution model describing the relation between the measured decay time and the real decay time is assumed to be a Gaussian distribution.

$$t_{meas} = \mathcal{G}(0, \sigma_t; t_{true}). \tag{7.6}$$

The final step of the calibration of the decay time resolution is the determination of the proportionality factor S_{σ_t} between the event-by-event estimate σ_t^i and the actual resolution σ_t

$$\sigma_t = S_{\sigma_t} \cdot \sigma_t^i. \tag{7.7}$$

Therefore, the weights determined in the previous sections are applied to the decay time pull of the fake B_s^0 candidates. The distribution is shown in Figure 7.8. It can be seen that it is well described by a Gaussian distribution. The fitted width corresponds directly to the scale factor S_{σ_t}. The fitted value is $S_{\sigma_t} = 1.36$.

7.1.5. Dependence of S_{σ_t} on kinematic variables

The final step for the calibration of the decay time resolution is the determination of the uncertainty on S_{σ_t}. The statistical uncertainty on the fitted width in Figure 7.8 is negligible. The dominant effect is found to be the dependence of S_{σ_t} on the kinematics of the fake B_s^0 candidates. Even if S_{σ_t} is found to depend on these variables, the average scale factor could accurately describe the real resolution of the signal $B_s^0 \to D_s^- \pi^+$ candidates. To ensure this the fake B_s^0 candidates used for the calibration would have to show the same distributions in these variables as the signal B_s^0 candidates.

The variables used for this cross-check are the transverse momentum p_{T}, the momentum p and the χ^2 of the vertex fit of the B_s^0 candidates because they show the largest dependence of S_{σ_t}. The distributions of these three variables for signal candidates and the fake B_s^0 candidates from simulated data are shown in Figure 7.9. It can be seen that the distributions do not agree with each other. Therefore, to determine the systematic effect of a wrong scale factor, S_{σ_t}, on the

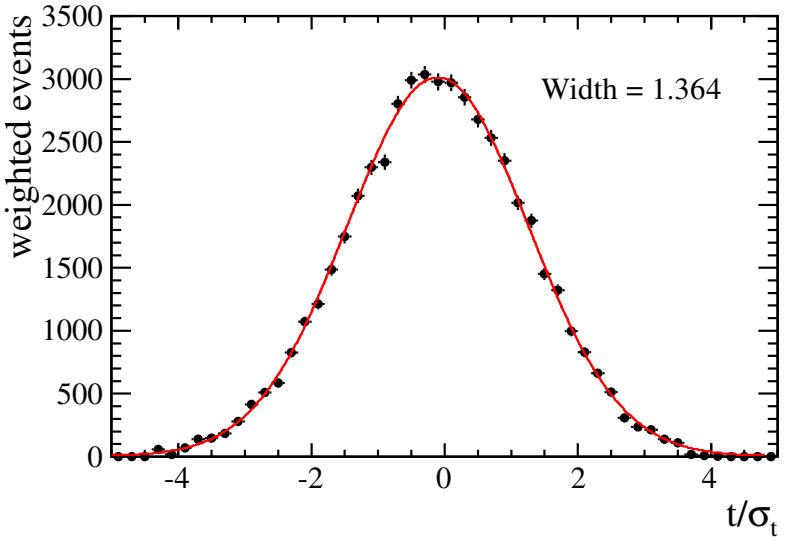


Figure 7.8.: *Distribution of the decay time pull of the fake B_s^0 candidates used for the calibration of the decay time resolution. The fit function is a Gaussian distribution.*

measurement of Δm_s, the full range of the variation of S_{σ_t} as a function of these variables is taken into account.

The three steps of the calibration (invariant mass fit, $log(IP)$ fit, decay time pull fit) are performed in bins of B_s^0 vtx χ^2, p_{T} and p, respectively. The bins are chosen in a way that there are roughly the same number of candidates in each bin. The results for S_{σ_t} as function of these variables are shown in Figure 7.10. The fitted values vary in the range between [1.25, 1.45]. This range is chosen to assign the systematic uncertainty on Δm_s, as will be discussed in Section 11.

7.1.6. Decay time resolution on simulated data

As a cross-check, the method to calibrate the decay time resolution is applied to simulated data. The distribution of the decay time pull of the fake B_s^0 candidates built from prompt $D^- \to K^+\pi^-\pi^-$ candidates is shown in Figure 7.11 on the left. On the right in Figure 7.11 the decay time pull for signal $B_s^0 \to D_s^-\pi^+$ candidates is shown.

It can be seen that the two distributions have slightly different widths. The reason for this are the different kinematic distributions mentioned in Section 7.1.5. The widths of the two distributions are well in agreement within the 10% uncertainty assigned to this effect.

Secondly, the single Gaussian distribution does not describe the decay time pull

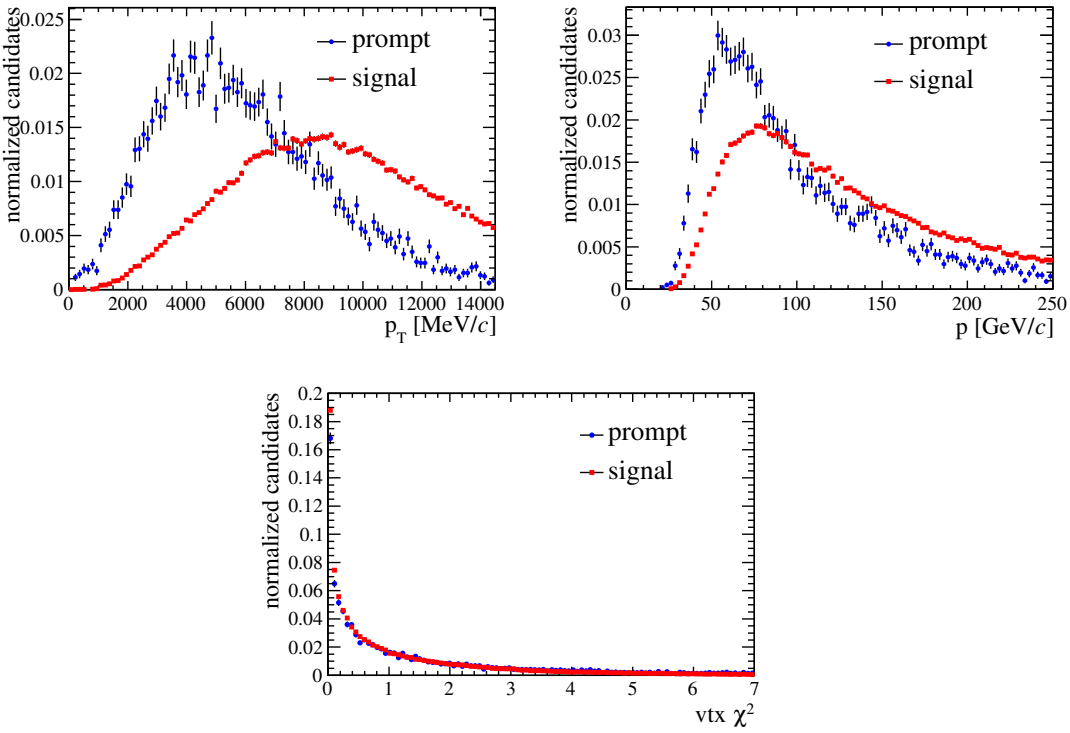


Figure 7.9.: *Distributions of p_T (upper left), p (upper right) and vertex χ^2 for signal $B_s^0 \to D_s^- \pi^+$ (red squares) and fake B_s^0 candidates used in the calibration of the decay time resolution (blue, circles) from simulated data.*

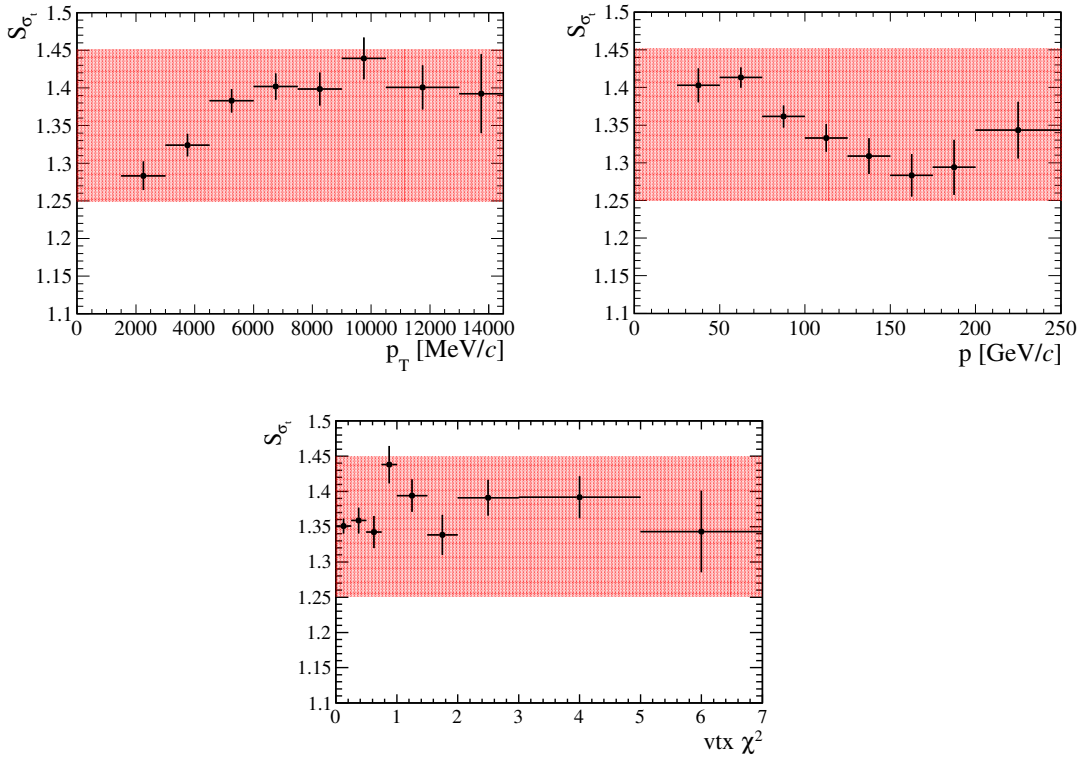


Figure 7.10.: *Fitted values of S_{σ_t} in bins of p_T (upper left), p (upper right) and vertex χ^2.*

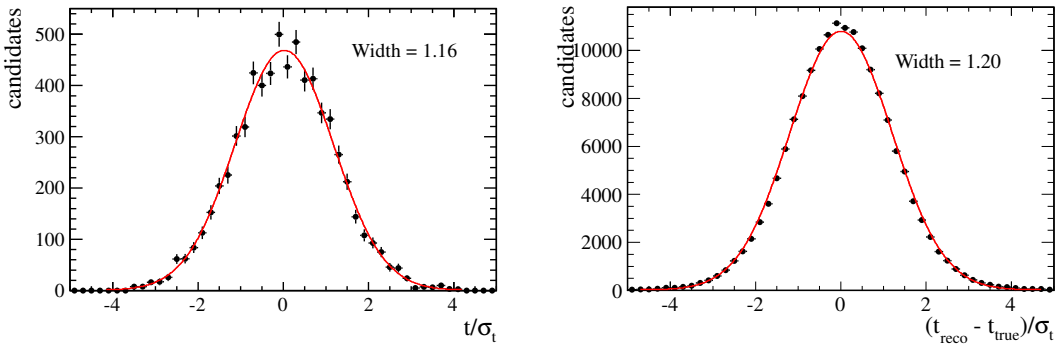


Figure 7.11.: *Decay time pull distributions for prompt $D^- \to K^+\pi^-\pi^-$ candidates (left) and $B_s^0 \to D_s^-\pi^+$ signal candidates (right) in simulated data. The distributions are fitted with a Gaussian distribution.*

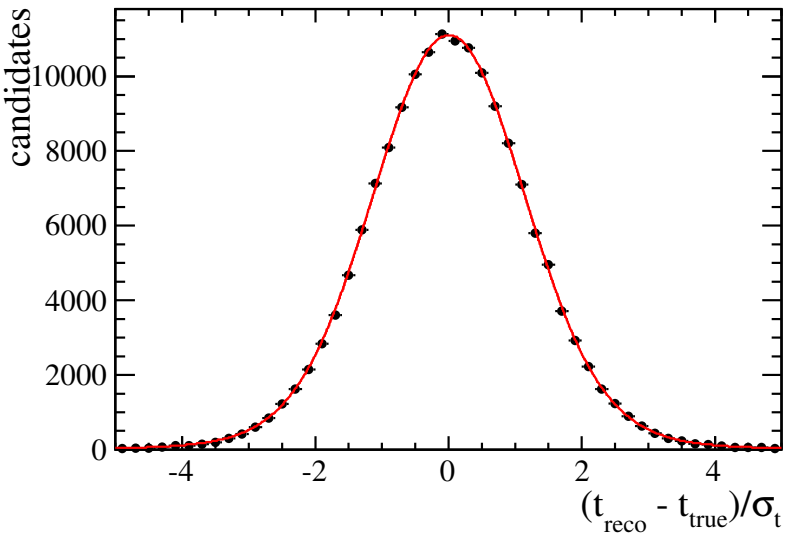


Figure 7.12.: *Distribution of the decay time pull of the signal $B_s^0 \to D_s^- \pi^+$ candidates from simulated data. The fit function is the sum of two Gaussian distributions.*

for the $B_s^0 \to D_s^- \pi^+$ candidates well as can be seen in Figure 7.11, the reduced χ^2 of the fit is $\chi^2/ndf = 14.6$. A more accurate description is given by the sum of two Gaussian distributions instead (see Figure 7.12). The reduced χ^2 of this fit is $\chi^2/ndf = 1.2$. The second, wider Gaussian distribution has a contribution of about 10% and a width that is about double the size of the inner one. Since the decay time pull distribution of the prompt D^- is described well by a single Gaussian distribution in data, this will be the nominal resolution model in the fit for Δm_s. The systematic effect of adding a second Gaussian with 10% contribution and double the width to the nominal resolution model will be studied in Chapter 11.

7.2. Decay Time Acceptance

Several of the selection criteria that are applied in the signal selection distort the decay time distribution of the B_s^0 signal candidates. The reason for this is that the efficiency of those cuts depends on the B_s^0 decay time. The efficiency of a cut is defined by the number of B_s^0 candidates passing the cut over the number of candidates before the cut was applied. One example for such a cut is the IPχ^2 of the B_s^0 and D_s^- daughters. These cuts are more efficient for B_s^0 candidates that decayed at a later time than for very shortlived ones. In the case of the shortlived B_s^0, the daughter tracks will be more likely to point to the PV. Other examples are the radial flight distance of the B_s^0, the direction angle cosDIRA, and the χ^2 of the distance between the PV and the secondary vertex. A second effect is

introduced due to the fact that also the reconstruction of tracks has an efficiency that depends on the decay time of the B_s^0.

Due to these efficiencies, the decay time distribution of the signal B_s^0 candidates cannot be modelled by a simple exponential. An acceptance function has to be taken into account. This acceptance describes the decay time dependence of the overall signal selection efficiency. It is obtained from Monte Carlo simulated data by dividing the decay time distribution of the B_s^0 signal candidates that pass the selection by a theory distribution which represents the reconstructed decay time distribution of the candidates before the selection.

$$\mathcal{E}_t(t) = \frac{S_{\text{reco}}(t)}{S_{\text{theo}}(t)} \tag{7.8}$$

The theory distribution is the one given in Equations 2.41 and 2.42 to describe the time evolution of the B_s^0 candidates without the part describing the oscillation. The acceptance is assumed to be the same for mixed and unmixed candidates. To take into account resolution effects due to the reconstruction, this distribution is convoluted with a Gaussian distribution using the same per-event error estimate σ_t^i and scale factor S_{σ_t} as determined on simulated data in Section 7.1.6.

$$S_{\text{theo}}(t_{\text{rec}}) = e^{-\Gamma_s t_{\text{true}}} \cosh \frac{\Delta\Gamma_s}{2} t_{\text{true}} \otimes G(0; S_{\sigma_t}\sigma_t) \tag{7.9}$$

This acceptance function could also be determined as a function of the true decay time (without the convolution), since the true, simulated decay time is available for the selected events. It has been found that the overall probability density function is analytically integrable for an acceptance as function of the reconstructed decay time (see Appendix B). This achieves a considerable reduction in computing time. Therefore, the acceptance is chosen as a function of the reconstructed decay time rather than the true one.

The resulting acceptance histogram is shown in Figure 7.13. The decay time acceptance should not be understood as an efficiency. For the acceptance, only the shape is relevant and not the overall scale which would be the case for an efficiency. Therefore the y-axis is given in arbitrary units. The function fitted to the histogram is parametrized as

$$\mathcal{E}_t(t) = (1 - e^{-\frac{(t-s_{\text{acc}})}{\alpha_{\text{acc}}}})(1 + \beta_{\text{acc}} t) \cdot \Theta(t - \zeta_{\text{acc}}) \tag{7.10}$$

with the parameters s_{acc}, ζ_{acc}, α_{acc} and β_{acc}. The parameters s_{acc} and α_{acc} describe the shape of the rising edge for low decay times. The parameter β_{acc} describes the decrease of efficiency for larger decay times parametrized as a linear drop-off. Finally the Heaviside step-function Θ describes the cut-off for decay times $t < \zeta_{\text{acc}}$ which are hard to model. In the analysis there is a hard cut placed at

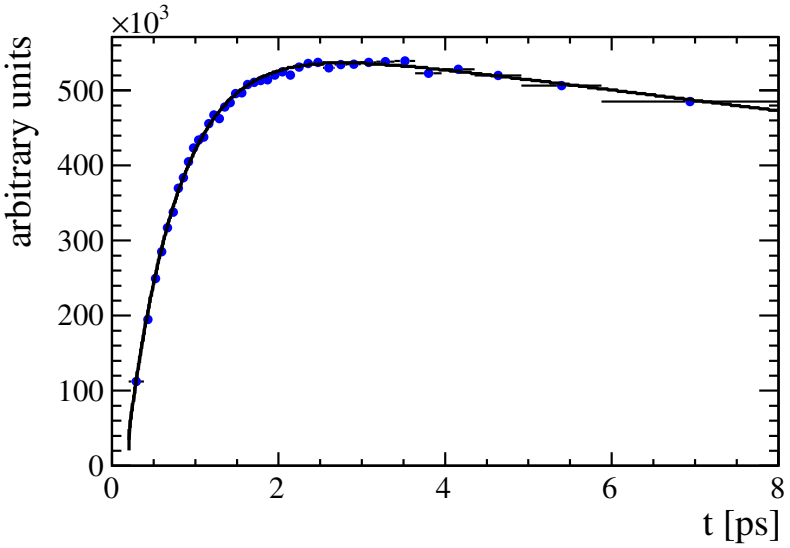


Figure 7.13.: *Decay time acceptance for B_s^0 candidates passing the signal selection. The fitted function is parametrized as shown in Equation 7.10. Fit results are summarized in Tab 7.1.*

Parameter	**Fit results**
α_{acc}	0.6329±0.0013 ps
β_{acc}	-0.0237±0.00024 1./ ps
s_{acc}	0.1571±0.0005 ps
ζ_{acc}	0.2 ps

Table 7.1.: *Fit results for decay time acceptance on simulated data. The parameter ζ_{acc} is fixed in the fit.*

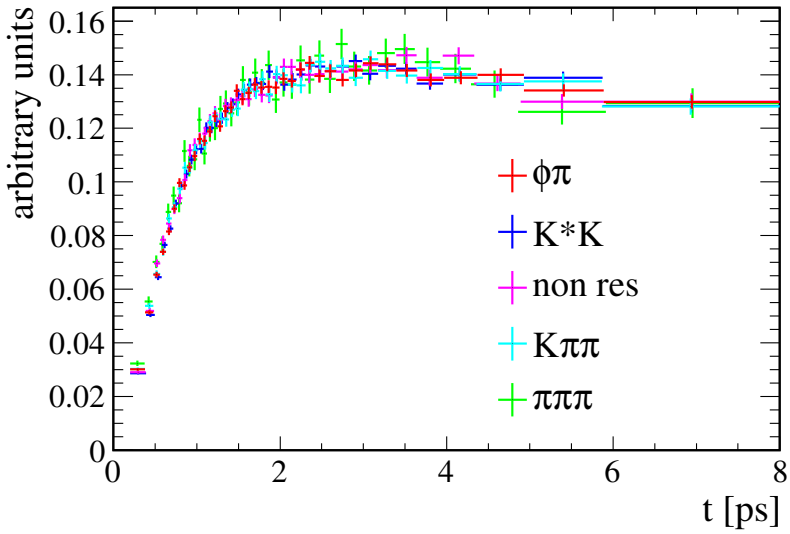


Figure 7.14.: *Decay time acceptance for B_s^0 candidates passing the signal selection for the individual D_s^- decay modes. The individual acceptances are normalized to the same area.*

$t = \zeta_{\text{acc}} = 0.2$. Therefore, it is fixed in the fit in Figure 7.13. The fitted parameters are summarized in Table 7.1.

The acceptance is assumed to be identical for the five different D_s^- decay modes. To justify this assumption, the decay time acceptance distributions for simulated data samples for each of the different D_s^- decay modes are plotted on top of each other in Figure 7.14. It can be seen that the shape of the acceptance is identical for all five decay modes. For the fit shown in Figure 7.13, the sum of the histograms in Figure 7.14 is used.

Also, for the $B_s^0 \to D_s^\pm K^\mp$ mode and the two misid backgrounds, the same acceptance function is used. For the misid backgrounds the simulated data samples were not large enough to compare the acceptances, but for the $B_s^0 \to D_s^\pm K^\mp$ mode the acceptance distribution is compared to the $B_s^0 \to D_s^- \pi^+$ signal in Figure 7.15. It can be seen that the acceptances are the same within the uncertainties. Therefore, it is concluded that it is justified to use the same acceptance parameters for the signal and backgrounds. Variation of the IP resolution in simulation to reproduce the one in data and different trigger selections lead to variations of the parameter α_{acc} of about 20%. The systematic effect of this variation on the measurement of Δm_s is described in Chapter 11.

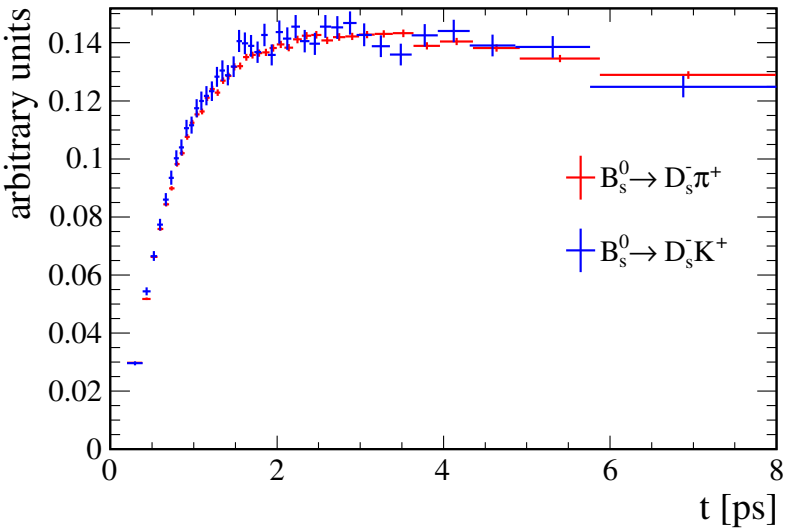


Figure 7.15.: *Decay time acceptance for B_s^0 candidates passing the signal selection for the $B_s^0 \to D_s^- \pi^+$ signal and the $B_s^0 \to D_s^\pm K^\mp$ mode. The individual acceptances are normalized to the same area.*

CHAPTER 8

Fit to the B_s^0 Mass Distribution

As discussed in Section 4.2, the physical probability density function (PDF) used in this analysis is divided into two parts which are considered independent. One is describing the invariant mass distribution of the reconstructed B_s^0 signal candidates and the second one is describing the mixing behaviour, namely the decay time distribution depending on the flavour tagging decision, q. In this chapter, the signal and background PDFs for the B_s^0 invariant mass distribution are described. The results for the parameters for these PDFs are given as determined in the combined fit in B_s^0 invariant mass, m, decay time, t, and flavour tagging decision, q.

As discussed in Section 5.8.1, there are a large number of background contributions from partially reconstructed decays present in the invariant mass region below the signal peak. These background contributions, in which one or more particles are not reconstructed, are difficult to model in the decay time distribution. Therefore, it has been decided to place a tight cut at the lower edge of the invariant mass signal region at $m > 5320\,\mathrm{MeV}/c^2$ to remove these background contributions.

8.1. Invariant Mass PDF for signal decays

In this section the PDFs describing the two signal contributions are given. The invariant mass PDF for signal is divided into one part describing the $B_s^0 \to D_s^- \pi^+$ decay, which includes most of the statistics, and the contribution from $B_s^0 \to D_s^\pm K^\mp$

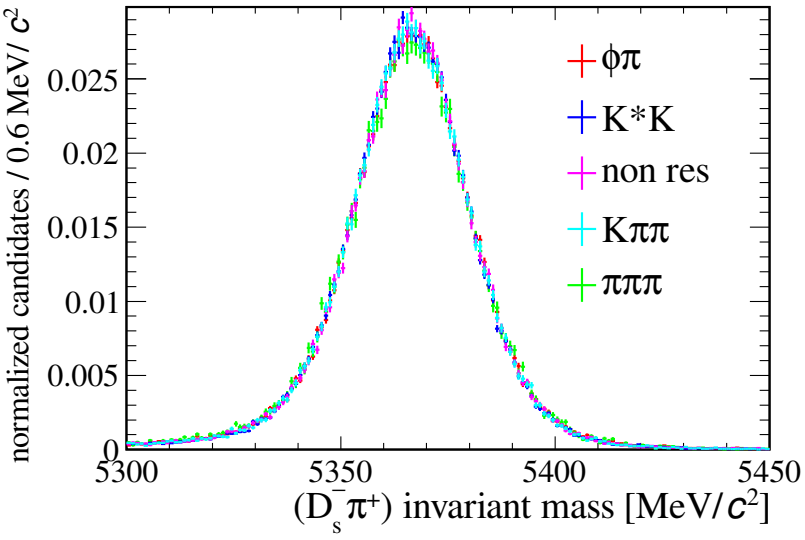


Figure 8.1.: *Invariant mass distributions for B_s^0 candidates for the individual D_s^- decay modes in simulated data. The individual distributions are normalized to the same area.*

decays, which have a different invariant mass distribution than $B_s^0 \to D_s^- \pi^+$ decays, but a signal-like mixing distribution. These two contributions are combined using their relative fraction f_{D_sK} as a fit parameter

$$\mathcal{P}_{\text{sig}} = f_{D_sK}\mathcal{P}_{D_sK} + (1 - f_{D_sK})\mathcal{P}_{D_s\pi}. \tag{8.1}$$

8.1.1. $B_s^0 \to D_s^- \pi^+$ invariant mass distribution

The invariant mass shape for the $B_s^0 \to D_s^- \pi^+$ signal candidates is motivated from simulated data. The invariant mass shapes for the five different D_s^- decay modes used in this analysis are assumed to be identical. This assumption is made because, in the calculation of the invariant mass of the B_s^0 candidate, the D_s^- invariant is constraint to the PDG value [23]. This should correct possible differences in the shape of the B_s^0 mass due to different phase space regions of the D_s^- daughters. As a cross-check, Figure 8.1 shows the invariant mass distributions of the B_s^0 candidates for the five D_s^- decay modes in simulated data. It can be seen that the distributions are identical and, thus, it is justified to use the same parameters for all five modes.

As parametrization for the signal shape, the sum of two Crystal Ball (CB) functions [68] has been chosen. It is an asymmetric function with 4 parameters defined as

$$CB(\alpha, n, \mu, \sigma; x) = N \cdot \begin{cases} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\mu}{\sigma})^{-n}, & \text{for } \frac{x-\mu}{\sigma} \leqslant -\alpha, \end{cases} \tag{8.2}$$

with A and B defined as

$$\begin{aligned} A &= \left(\frac{n}{|\alpha|}\right)^n \cdot e^{-\frac{|\alpha|^2}{2}} \quad \text{and} \\ B &= \frac{n}{|\alpha|} - |\alpha| \,. \end{aligned} \tag{8.3}$$

The parameter μ describes the peak position, σ describes the width of the distribution and α and n describe the asymmetric tail.

To be used as a PDF the CB function has to be normalized. The normalization factor N is given by

$$N = \frac{1}{\sigma(C+D)}, \tag{8.4}$$

with

$$\begin{aligned} C &= \frac{n}{|\alpha|} \cdot \frac{1}{n-1} \cdot e^{-\frac{|\alpha|^2}{2}} \quad \text{and} \\ D &= \sqrt{\frac{\pi}{2}} \left(1 + \text{erf}\left(\frac{|\alpha|}{\sqrt{2}}\right)\right). \end{aligned} \tag{8.5}$$

where erf is the error function[1] For the invariant mass signal shape the peak position of the two CB functions is required to be the same. The full invariant mass PDF for the $B_s^0 \to D_s^- \pi^+$ signal decays is then given by

$$\mathcal{P}_{B_s^0 \to D_s^- \pi^+; m} = f_m \cdot CB(\alpha_m^1, n_m^1, m_{B_s^0}, \sigma_m^1) + (1 - f_m) \cdot CB(\alpha_m^2, n_m^2, m_{B_s^0}, \sigma_m^2), \tag{8.6}$$

with f_m the relative fraction of the two CB functions.

Figure 8.2 shows the fitted distribution for simulated data. Table 8.1 summarizes the results for the fitted parameters. In the final fit using data, all of these parameters, except for the peak position $m_{B_s^0}$ and the two widths σ_m^1 and σ_m^2, are fixed to the results in Table 8.1. The systematic effect on Δm_s of an alternative mass model of the sum of two Gaussian distributions is studied in Chapter 11

8.1.2. $B_s^0 \to D_s^\pm K^\mp$ mass distribution

The template for the $B_s^0 \to D_s^\pm K^\mp$ signal contribution is taken from simulated data. The full signal selection described in Chapter 5 is applied and approximately

[1] The error function is defined as

$$\text{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

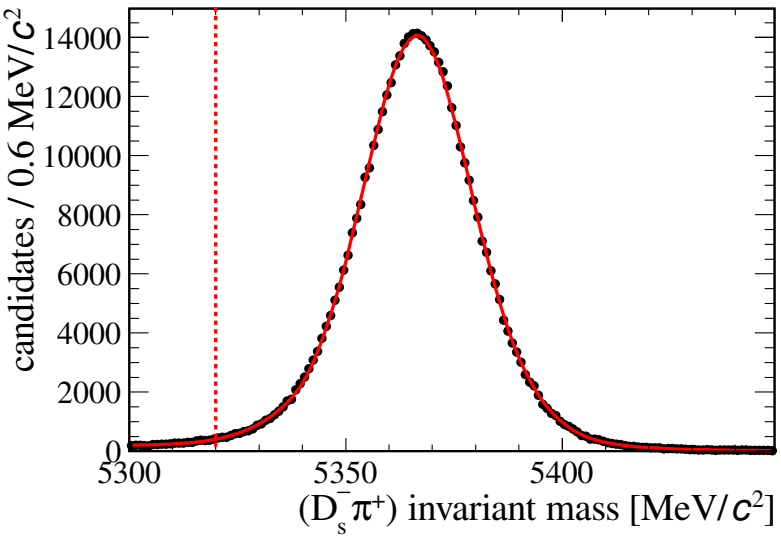


Figure 8.2.: *Invariant mass distributions for B_s^0 candidates in simulated data. The fitted function is the sum of two CB functions as defined in Equation 8.6. The fit results are summarized in Table 8.1. The dashed (red) line indicates the invariant mass cut applied in the signal selection.*

parameter	**Fit results**
$m_{B_s^0}$ [MeV/c^2]	5366.50±0.02
σ_m^1 [MeV/c^2]	12.47±0.13
σ_m^2 [MeV/c^2]	21.9±0.3
f_m	0.78±0.02
α_m^1	2.13±0.06
α_m^2	-2.07±0.04
n_m^1	1.10±0.06
n_m^2	5.8±0.4

Table 8.1.: *Fit results for the invariant mass distribution of the $B_s^0 \to D_s^- \pi^+$ candidates from simulated data. The fit function is the sum of two CB functions as defined in Equation 8.6.*

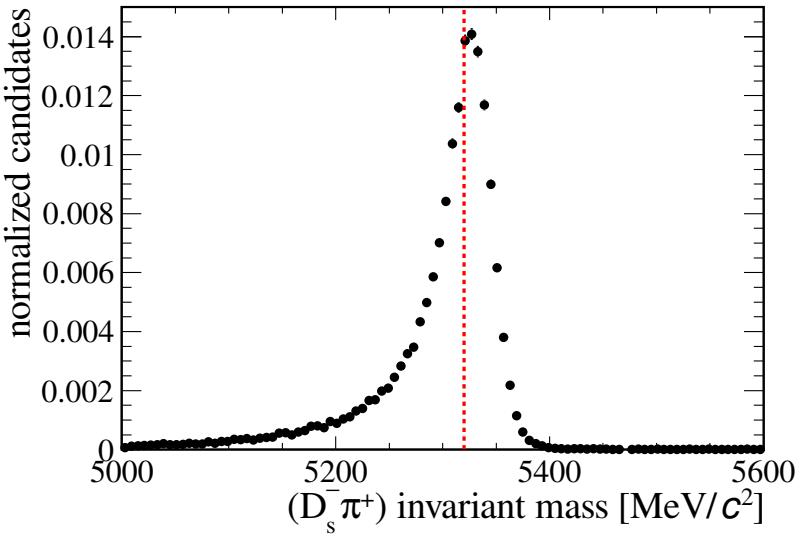


Figure 8.3.: *Invariant mass distributions for $B_s^0 \to D_s^\pm K^\mp$ candidates reconstructed as $B_s^0 \to D_s^- \pi^+$ in simulated data. The dashed (red) line indicates the invariant mass cut applied in the signal selection.*

55,000 candidates remain. The invariant mass distribution of the candidates that were reconstructed under the $B_s^0 \to D_s^- \pi^+$ hypothesis is shown in Figure 8.3.

The histogram is normalized to have an integral equal to one. This histogram could in principle be used directly as the invariant mass PDF $\mathcal{P}_{D_s K;m}$ in the fit. However, it has been found that, due to imperfect momentum calibration, the reconstructed invariant mass spectrum is shifted by 3.9 MeV/c^2 compared to the simulated distributions. Therefore, the histogram is shifted by this amount to be compatible with the reconstructed B_s^0 mass in data. The histogram in Figure 8.3 already includes this shift.

It should be noted that for PDFs the normalization has to be calculated in the range of the selected data. The nominal invariant mass window of the selection is [5320, 5980] MeV/c^2. Therefore, the histogram used as mass PDF is normalized so that the integral in this mass range is equal to one. Another complication with histograms arises if the edge of the mass window does not coincide with a bin edge. In such a case the area of the excluded rectangle must also be excluded from the calculation of the normalization. These two effects are taken into account for this histogram as well as every other histogram that is later used as a PDF in the fit.

8.2. Invariant Mass PDFs for background decays

The templates for the mass distributions of the background components described in Section 5.8 are taken from different sources, both simulated and real data. The mass templates for the misid. background contributions from $B^0 \to D^-\pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ are taken from [2]. The shape of the combinatorial background is fitted directly to data. In this section the PDFs describing these three background contributions are given.

8.2.1. $B^0 \to D^-\pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ mass distribution

The mass distributions for the misid. background contributions from the decays $B^0 \to D^-\pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ have been provided by an external analysis of *CP* violation in the decay $B_s^0 \to D_s^\pm K^\mp$ [2]. They have been obtained from data as follows: A clean sample of reconstructed $B^0 \to D^-\pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ candidates has been selected by reconstructing them with their respective correct mass hypotheses and applying tight invariant mass cuts. Since the performance of the PID cuts depend on the momentum of the corresponding particle, these cuts can influence the resulting invariant mass shape. Thus, calibration functions obtained from D^* decays have been used to reweight the momenta of the misidentified particles. The samples have then been reconstructed a second time using the $B_s^0 \to D_s^- \pi^+$ mass hypotheses to obtain the invariant mass shapes.

The resulting mass distributions are shown in Figure 8.4. Similar to the $B_s^0 \to D_s^\pm K^\mp$ signal mass distribution these are normalized histograms which are shifted by $3.9\,\mathrm{MeV}/c^2$ compared to the simulated distributions.

8.2.2. Combinatorial background mass distribution

The combinatorial background mass distribution is fitted directly to data. It is exponentially distributed, thus the mass PDF is given by

$$\mathcal{P}_{\mathrm{comb};m} = \frac{1}{\mathcal{N}} e^{-\alpha_{\mathrm{comb}} m}, \tag{8.7}$$

with the normalization factor $\mathcal{N}$

$$\mathcal{N} = \int_{m_{\mathrm{min}}}^{m_{\mathrm{max}}} e^{-\alpha_{\mathrm{comb}} m} dm = \frac{1}{\alpha_{\mathrm{comb}}} \cdot \left(e^{-\alpha_{\mathrm{comb}} m_{\mathrm{min}}} - e^{-\alpha_{\mathrm{comb}} m_{\mathrm{max}}} \right). \tag{8.8}$$

B_s^0 mesons are the heaviest B mesons produced in large quantities. Therefore, contributions from partially reconstructed other B meson decays are all located

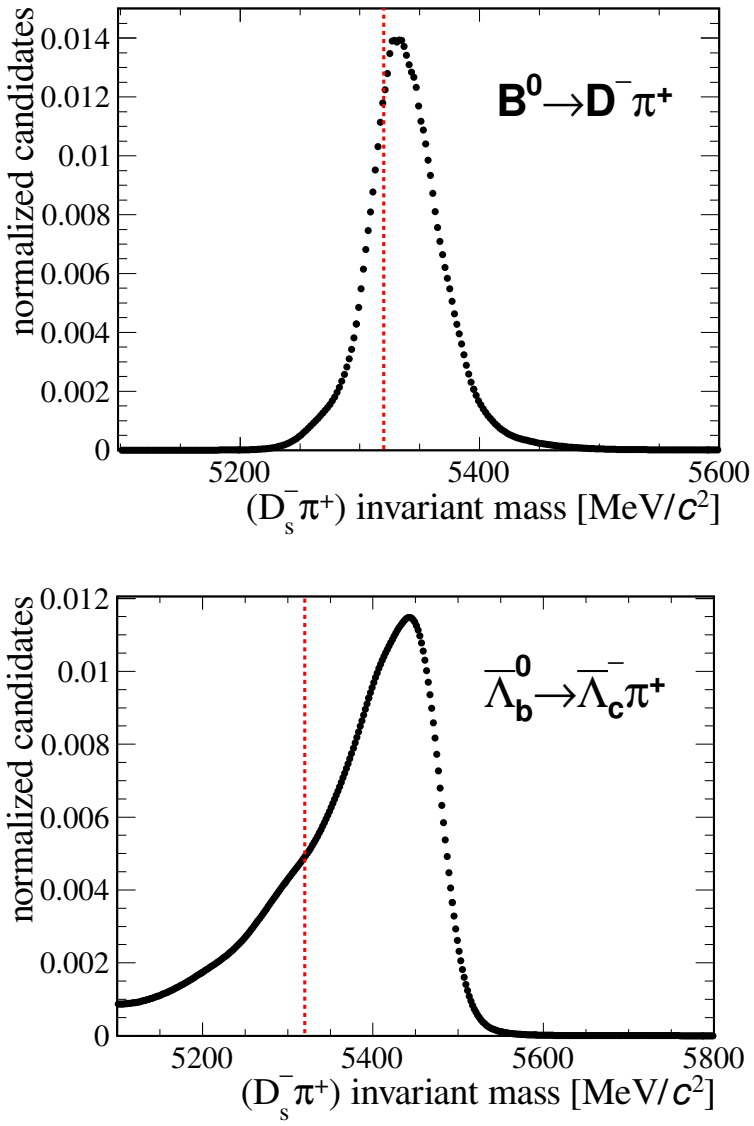


Figure 8.4.: *Invariant mass distributions for $B^0 \to D^- \pi^+$ decays (top) and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays (bottom) reconstructed as $B_s^0 \to D_s^- \pi^+$ [2]. The dashed (red) line indicates the invariant mass cut applied in the signal selection.*

in the invariant mass region below the $B_s^0 \to D_s^- \pi^+$ signal peak. In the high-mass sideband region there are only contributions from combinatorial backgrounds present. Thus, the slope parameter α_{comb} can be determined accurately in the mass fit independent of the models for the other background contributions. The slope and the amount of this contribution is highly dependent on the selection criteria applied. Therefore, both parameters α_{comb} and the fraction f_{comb} are floating individually for the five D_s^- decay modes.

8.3. Results of the invariant mass parameters in data

The full PDF including all signal and background components is given by

$$\begin{aligned}\mathcal{P} =& f_{\text{sig}} \left(f_{D_s K} \mathcal{P}_{B_s^0 \to D_s^\pm K^\mp} + (1 - f_{D_s K}) \mathcal{P}_{B_s^0 \to D_s^- \pi^+} \right) \\ &+ (1 - f_{\text{sig}}) \Big(f_{\text{comb}} \mathcal{P}_{comb} \\ &\quad + (1 - f_{\text{comb}}) \left(f_{\Lambda_b^0} \mathcal{P}_{\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+} + (1 - f_{\Lambda_b^0}) \mathcal{P}_{B^0 \to D^- \pi^+} \right) \Big)\end{aligned} \tag{8.9}$$

with the fractions f_{sig}, $f_{D_s K}$, f_{comb} and $f_{\Lambda_b^0}$.

As discussed in Section 4.2, the individual PDFs each consist of the five components describing the mass, $\mathcal{P}_{J;m}$, the decay time and tagging decision, $\mathcal{P}_{J;t,q}$, the decay time error estimate, $\mathcal{P}_{J;\sigma_t}$, and the mistag estimates for the opposite side tagger, $\mathcal{P}_{J;\eta_{\text{OST}}}$, and the same side tagger, $\mathcal{P}_{J;\eta_{\text{SST}}}$. However, the largest separation power between the individual signal/background components comes from the mass PDFs. Therefore, the fractions will be listed as parameters of the invariant mass distribution.

The shape of the mass templates of the signal contribution from $B_s^0 \to D_s^\pm K^\mp$ decays and the misid. backgrounds described in Sections 8.1.2 and 8.2.1 are fixed since histograms are used. Only their relative fractions are free parameters in the fit. The strategy for the simultaneous fit of the five D_s^- decay modes is to have as many parameters as possible to be shared among the modes. The common fit parameters of the invariant mass description are the signal peak position, $m_{B_s^0}$, and widths, σ_m^1 and σ_m^2, and the fraction, $f_{D_s K}$, of $B_s^0 \to D_s^\pm K^\mp$ candidates compared to the $B_s^0 \to D_s^- \pi^+$ signal candidates.

In the mass range $[5320, 5980]\,\text{MeV}/c^2$ the shapes of the $B^0 \to D^- \pi^+$ background and the $B_s^0 \to D_s^\pm K^\mp$ background are very similar. Therefore, it is not possible to fit both fractions individually at the same time. Thus, the $B_s^0 \to D_s^\pm K^\mp$ fraction is obtained from an additional mass fit to the sample of $B_s^0 \to D_s^- \pi^+$ decays with $D_s^- \to \pi^- \pi^+ \pi^-$. As explained in Section 5.8.3, in this D_s^- decay mode there is no contribution from $B^0 \to D^- \pi^+$ background and only $B_s^0 \to D_s^\pm K^\mp$ background

and combinatorial background are present. The fraction f_{D_sK} is extracted from this mass fit and then, since it is assumed to be independent of the D_s^- decay, fixed in the fit of all decay modes. Using only the $D_s^- \to \pi^-\pi^+\pi^-$ mode to fit f_{D_sK} results in a large uncertainty on the fitted fraction $f_{D_sK} = 0.019 \pm 0.010\%$. The variation of f_{D_sK} within these uncertainties is part of the systematic studies on the measurement of Δm_s.

The parameters f_{sig}, f_{comb}, $f_{\Lambda_b^0}$ and α_{comb} are floating independently in the five decay modes with except for the aforementioned decay mode $D_s^- \to \pi^-\pi^+\pi^-$ in which f_{comb} is set to 1.0 to set the contribution by misid. backgrounds to zero. The fit results of the invariant mass parameters are summarized in Table 8.2. The invariant mass distributions with the fitted PDFs are shown in Figure 8.5 for the different D_s^- decay modes.

Table 8.3 lists the fitted number of signal candidates given by the total number of candidates used in the fit multiplied by the corresponding fractions. The correlations between the different fractions are taken into account in the uncertainties quoted.

To justify the exponential model used to describe the combinatorial background the invariant mass distributions are shown on a logarithmic scale in Figure 8.6. It can be seen that the high-mass sidebands are well described by the exponential function.

Figure 8.7 shows the mass distribution for all candidates passing the signal selection and the sum of the PDFs for the five decay modes.

Common parameters among the five D_s^- decay modes	
$m_{B_s^0}[\,\mathrm{MeV}/c^2]$	5370.44±0.13
$\sigma_m^1[\,\mathrm{MeV}/c^2]$	15.23±0.13
$\sigma_m^2[\,\mathrm{MeV}/c^2]$	27.96±0.87
f_{D_sK}	0.019±0.0 (fixed)
Parameters of the $D_s^- \to \phi(K^+K^-)\pi^-$ decay	
f_{sig}	0.8502±0.0075
f_{comb}	0.826±0.045
$f_{\Lambda_b^0}$	0.46±0.13
α_{comb}	0.00238±0.00020
Parameters of the $D_s^- \to K^{*0}(K^+\pi^-)K^-$ decay	
f_{sig}	0.8748±0.0085
f_{comb}	0.748±0.058
$f_{\Lambda_b^0}$	0.27±0.11
α_{comb}	0.00279±0.00029
Parameters of the $D_s^- \to K^+K^-\pi^-$ non-resonant decay	
f_{sig}	0.6066±0.0089
f_{comb}	0.700±0.024
$f_{\Lambda_b^0}$	0.560±0.043
α_{comb}	0.00300±0.00017
Parameters of the $D_s^- \to K^-\pi^+\pi^-$ decay	
f_{sig}	0.446±0.013
f_{comb}	0.849±0.026
$f_{\Lambda_b^0}$	0.460±0.090
α_{comb}	0.00275±0.00018
Parameters of the $D_s^- \to \pi^-\pi^+\pi^-$ decay	
f_{sig}	0.6310±0.0081
f_{comb}	1.0±0.0 (fixed)
α_{comb}	0.00287±0.00014

Table 8.2.: *Results of the mass parameters in the fit of the five $B_s^0 \to D_s^-\pi^+$ decay modes.*

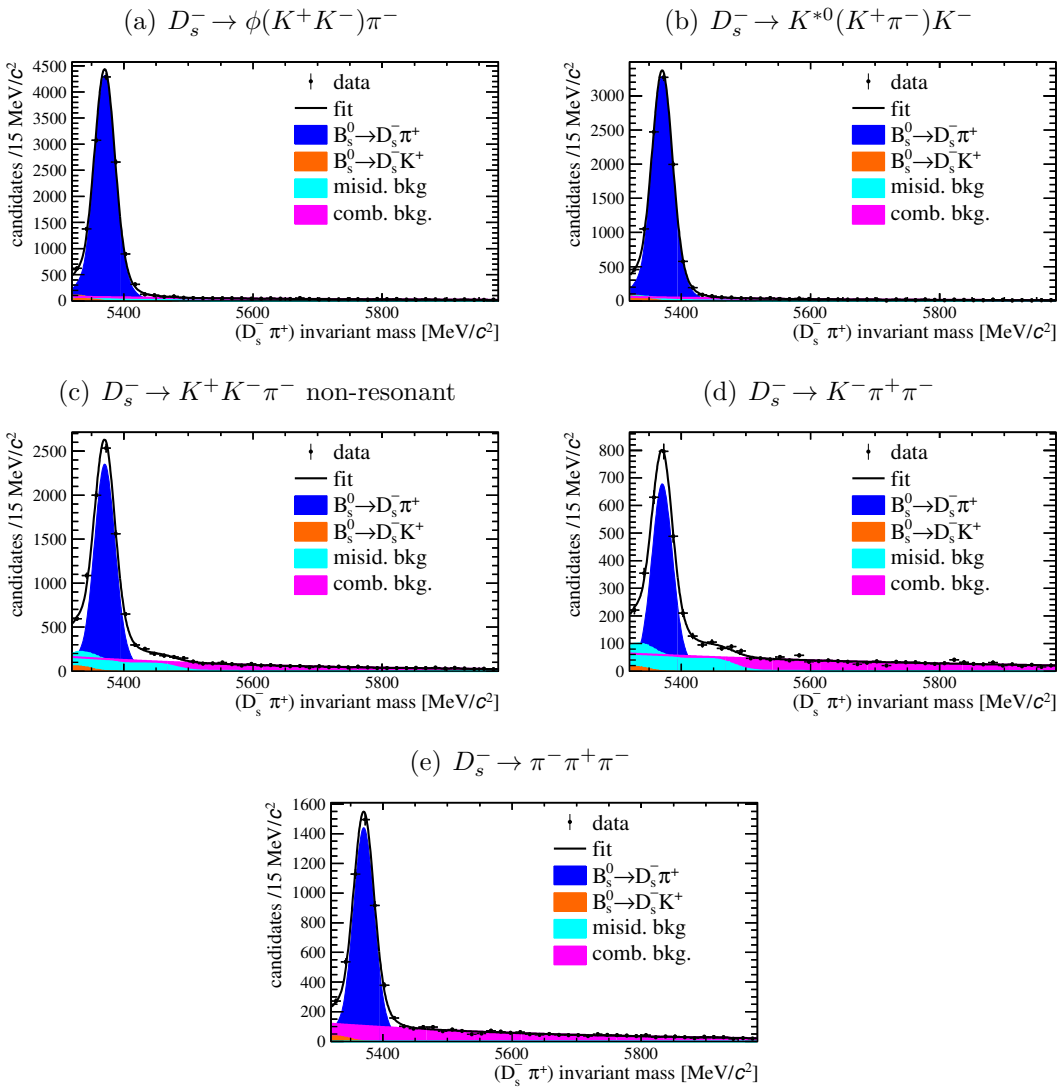


Figure 8.5.: *Projection of the fitted mass distribution in (a) the $D_s^- \to \phi(K^+K^-)\pi^-$ decay, (b) the $D_s^- \to K^{*0}(K^+\pi^-)K^-$ decay, (c) the $D_s^- \to K^+K^-\pi^-$ non-resonant decay, (d) the $D_s^- \to K^-\pi^+\pi^-$ decay and (e) the $D_s^- \to \pi^-\pi^+\pi^-$ decay in data.*

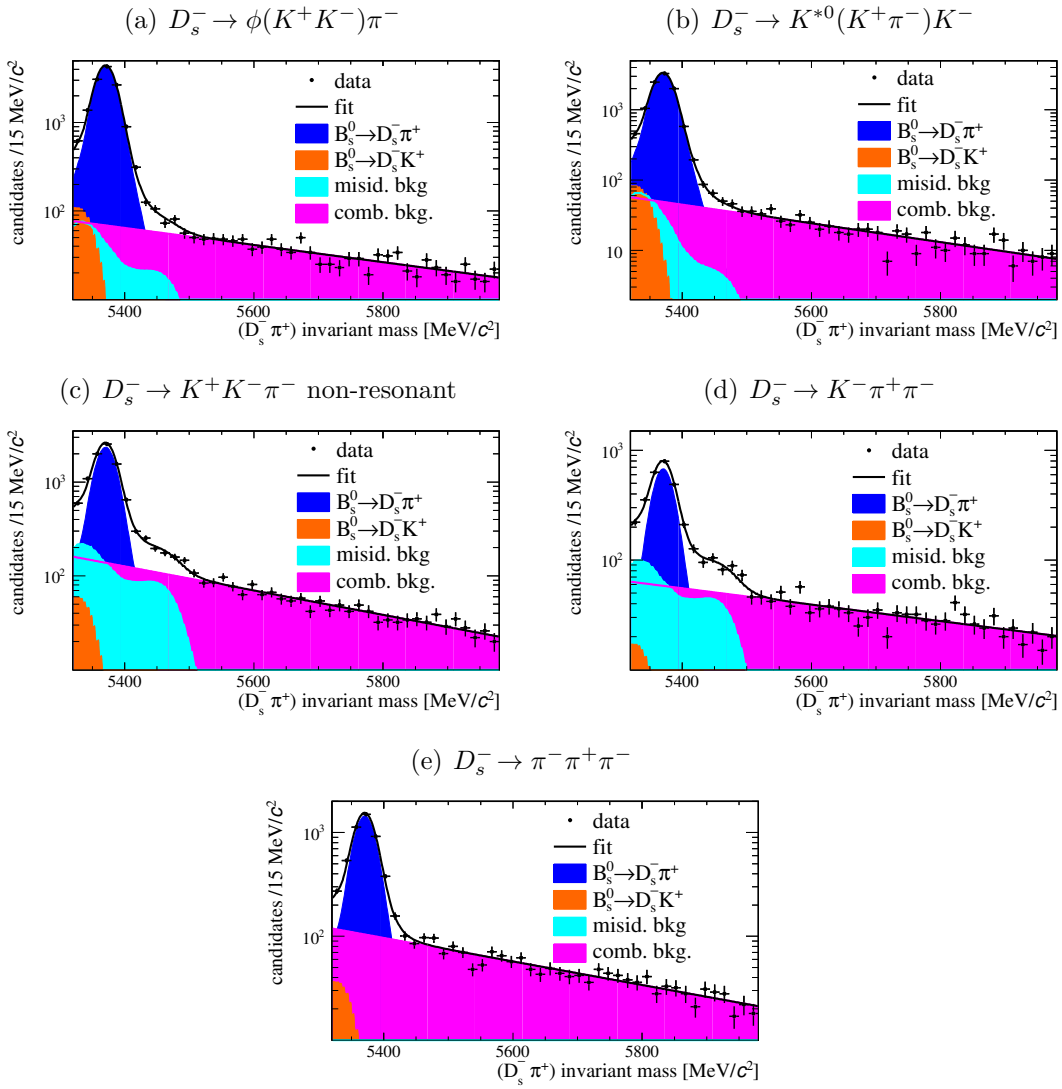


Figure 8.6.: *Projection of the fitted mass distribution in (a) the $D_s^- \to \phi(K^+K^-)\pi^-$ decay, (b) the $D_s^- \to K^{*0}(K^+\pi^-)K^-$ decay, (c) the $D_s^- \to K^+K^-\pi^-$ non-resonant decay, (d) the $D_s^- \to K^-\pi^+\pi^-$ decay and (e) the $D_s^- \to \pi^-\pi^+\pi^-$ decay in data on a logarithmic scale.*

D_s^- decay mode	$B_s^0 \to D_s^- \pi^+$ signal yield	$B_s^0 \to D_s^\pm K^\mp$ signal yield
$D_s^- \to \phi(K^+K^-)\pi^-$	12250±110	240±130
$D_s^- \to K^{*0}(K^+\pi^-)K^-$	9320±90	180±90
$D_s^- \to K^+K^-\pi^-$ non-resonant	6700±100	130±70
$D_s^- \to K^-\pi^+\pi^-$	1870±60	40±20
$D_s^- \to \pi^-\pi^+\pi^-$	4000±50	80±40
Total	34140±190	670±360

Table 8.3.: *Number of B_s^0 signal candidates in the data. The uncertainties on the $B_s^0 \to D_s^\pm K^\mp$ yields result from the fit to the $D_s^- \to \pi^-\pi^+\pi^-$ sample alone.*

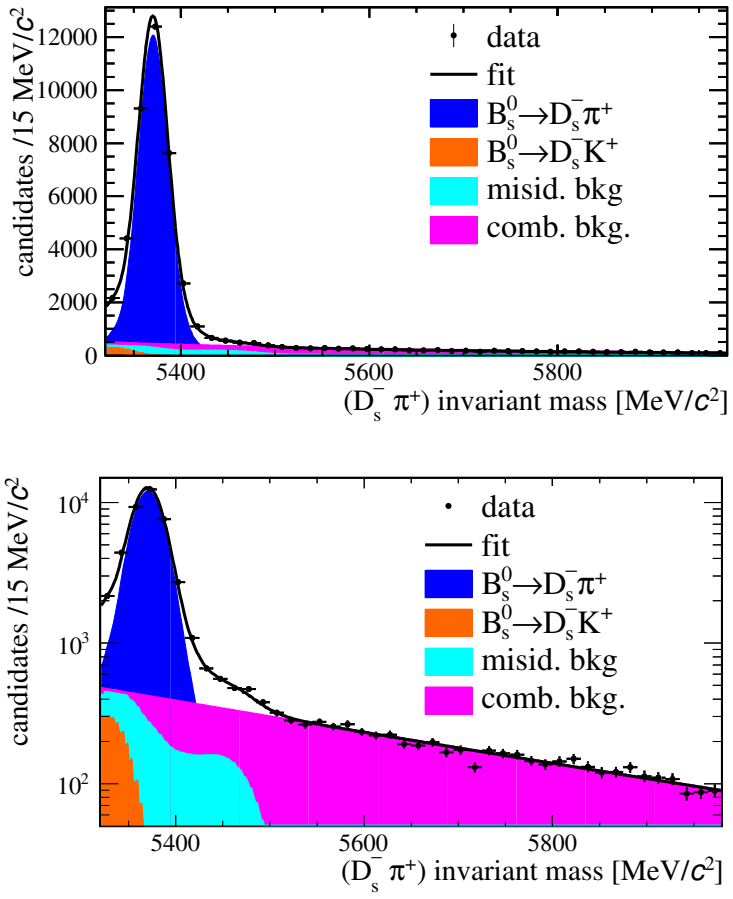


Figure 8.7.: *Projection of the fitted mass distributions for the sum of the five D_s^- decay modes in normal scale (top) and logarithmic scale (bottom).*

CHAPTER 9

Fit to the B_s^0 Mixing Distribution

The $B_s^0 \to D_s^- \pi^+$ mixing distribution depends on the B_s^0 decay time, t, and the flavour tagging decision, q. The mixing PDF is the second part of the full PDF describing the physics parameters in this analysis. In this chapter the mixing PDFs for signal and background components are described and the results for the floating parameters in the mixing PDF are given for the combined fit in B_s^0 invariant mass, decay time and flavour tagging decision. The physical quantities of the signal mixing PDF include the B_s^0–$\overline{B}_s^0$ oscillation frequency Δm_s.

The factorization ansatz for the combination of the invariant mass PDF and the mixing PDF is only justified if there is no correlation between the invariant mass and the mixing variables, t and q. To test this, a fit of the invariant mass distribution for low decay time ($t < 2\,\text{ps}$) and high decay time ($t > 2\,\text{ps}$) is performed[1].

For the signal PDF, the signal mass resolution between the two fits is compared in Table 9.1. Since the parameters describing the two widths (σ_m^1 and σ_m^2) of the signal CB functions are highly correlated, one is fixed to the nominal fit result and the other width is floating.

It can be seen that the signal mass resolution is compatible within the uncertainties for high and low decay times. Thus, it is concluded that there is no sizeable correlation between these two dimensions for signal decays.

[1] The cut at 2 ps was chosen to split the sample in two parts with approximately the same statistics.

parameter		$> 2\,\text{ps}$	$< 2\,\text{ps}$
σ_m^1		15.15±0.18	15.25±0.17
$\alpha_{\text{comb}}^{\phi\pi^-}$	$[1/\,\text{MeV}/c^2\,]$	0.0024±0.0005	0.0021±0.0003
$\alpha_{\text{comb}}^{K^{*0}K^-}$	$[1/\,\text{MeV}/c^2\,]$	0.0039±0.0007	0.0028±0.0004
$\alpha_{\text{comb}}^{K^+K^-\pi^-}$	$[1/\,\text{MeV}/c^2\,]$	0.0042±0.0005	0.0027±0.0002
$\alpha_{\text{comb}}^{K^-\pi^-\pi^+}$	$[1/\,\text{MeV}/c^2\,]$	0.0025±0.0006	0.0016±0.0003
$\alpha_{\text{comb}}^{\pi^+\pi^-\pi^+}$	$[1/\,\text{MeV}/c^2\,]$	0.0042±0.0005	0.0024±0.0002

Table 9.1.: *Results of the signal mass resolution for events with high ($> 2\,\text{ps}$) and low ($< 2\,\text{ps}$) decay time.*

For the background PDF, the shape parameters of the exponential describing the combinatorial background are also compared in Table 9.1. The slope parameters α_{comb} in the high decay time sample are significantly larger than for low decay times. This effect has, however, a negligible effect on the measurement of Δm_s, as will be shown in a systematic study described in Chapter 11.

9.1. Mixing PDF for signal decays

As described in Section 8.1, the signal contributions in the fit consist of $B_s^0 \to D_s^- \pi^+$ decays and $B_s^0 \to D_s^\pm K^\mp$ decays. In this section the PDFs describing these two signal contributions are given.

9.1.1. $B_s^0 \to D_s^- \pi^+$ mixing distribution

The mixing PDF for $B_s^0 \to D_s^- \pi^+$ signal candidates is the theory distribution from Equations 2.41 and 2.42 in Section 2.3

$$\mathcal{P}_{D_s\pi;t,q} \propto e^{-\Gamma_s t} \cdot \begin{cases} \cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - \cos\left(\Delta m_s t\right) & \text{for mixed } B_s^0 \text{ cand.} \\ \cosh\left(\frac{\Delta\Gamma_s}{2}t\right) + \cos\left(\Delta m_s t\right) & \text{for unmixed } B_s^0 \text{ cand.}\,, \end{cases} \tag{9.1}$$

including as physics parameters the decay width, Γ_s, the decay width difference, $\Delta\Gamma_s$, and the oscillation frequency, Δm_s. In the fit, Γ_s is fixed to the world average ($\Gamma_s = 0.6596 \pm 0.0073\,\text{ps}^{-1}$ [23]) and $\Delta\Gamma_s$ is fixed to the latest result by LHCb ($\Delta\Gamma_s = 0.106 \pm 0.011(\text{stat}) \pm 0.007(\text{syst})\,\text{ps}^{-1}$ [1]). The systematic effect of wrong values for Γ_s abd $\Delta\Gamma_s$ are studied in Section 11. Using the combined flavour tagging decision for opposite and same side taggers (see Section 6.6) $q = +1$ for mixed events and $q = -1$ for unmixed events, Equation 9.1 can be written as

$$\mathcal{P}_{D_s\pi;t,q} \propto e^{-\Gamma_s t} \cdot \left[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - q\cos\left(\Delta m_s t\right)\right]. \tag{9.2}$$

To include the finite decay time resolution, the PDF is convoluted with a Gaussian distribution with a mean equal to zero and width equal to the calibrated error estimate, $\sigma = S_{\sigma_t} \cdot \sigma_t$, described in Section 7.1. This resolution is assumed to be independent from the tagging decision, q. This gives

$$\mathcal{P}_{D_s\pi;t,q} \propto \left(e^{-\Gamma_s t} \cdot \left[\cosh\left(\frac{\Delta\Gamma_s}{2} t\right) - q \cos\left(\Delta m_s t\right) \right] \right) \otimes \mathcal{G}(0, S_{\sigma_t}\sigma_t; t). \tag{9.3}$$

The scale factor is fixed to the value obtained in the calibratrion, $S_{\sigma_t} = 1.36$. In the next step, the mixing PDF is multiplied by the decay time acceptance, $\mathcal{E}_t(t)$, described in Section 7.2. Similar to the resolution, it is assumed to be identical for different tagging decisions

$$\begin{aligned} \mathcal{P}_{D_s\pi;t,q} \propto \Bigg[& \left(e^{-\Gamma_s t} \cdot \left[\cosh\left(\frac{\Delta\Gamma_s}{2} t\right) - q \cos\left(\Delta m_s t\right) \right] \right) \\ & \otimes \mathcal{G}(0, S_{\sigma_t}\sigma_t; t) \Bigg] \cdot \mathcal{E}_t(t). \end{aligned} \tag{9.4}$$

The acceptance parameters β_{acc}, s_{acc} and ζ_{acc} are fixed in the fit to the values obtained from simulated data (see Table 7.1). The parameter α_{acc} describes the shape of the rising edge for low decay times. To describe the data accurately, it is left floating.

Imperfect tagging is taken into account in the decay time PDF by the dilution factor $\mathcal{D} = (1 - 2\omega)$ and the tagging efficiencies $\varepsilon_{\text{sig}}^{\text{OST}}$ and $\varepsilon_{\text{sig}}^{\text{SST}}$. The final mixing PDF for $B_s^0 \to D_s^- \pi^+$ signal candidates is:

$$\begin{aligned} \mathcal{P}_{D_s\pi;t,q} \propto \Bigg[& \left(e^{-\Gamma_s t} \cdot \left[\cosh\left(\frac{\Delta\Gamma_s}{2} t\right) - q \cdot (1 - 2\omega) \cdot \cos\left(\Delta m_s t\right) \right] \right) \\ & \otimes \mathcal{G}(0, S_{\sigma_t}\sigma_t; t) \Bigg] \cdot \mathcal{E}_t(t) \cdot \mathcal{N}_{\text{tag}}^{\text{OST}} \cdot \mathcal{N}_{\text{tag}}^{\text{SST}}. \end{aligned} \tag{9.5}$$

with the combined mistag probability ω which is calibrated using

$$\omega_i = p_{0,i} + (\eta_i - \langle \eta_i \rangle) + p_{1,i} \ , \tag{9.6}$$

and combined as described in Section 6.6. $\mathcal{N}_{\text{tag}}^{\text{OST}}$ is the relative normalization factor for the flavour tagging dimension

$$\mathcal{N}_{\text{tag}}^{\text{OST}} = \frac{|q^{\text{OST}}|\varepsilon_{\text{sig}}^{\text{OST}}}{2} \ + \ (1 - |q^{\text{OST}}|) \cdot (1 - \varepsilon_{\text{sig}}^{\text{OST}}), \tag{9.7}$$

which is equal to $\frac{\varepsilon_{\text{sig}}^{\text{OST}}}{2}$ for tagged events ($q^{\text{OST}} = \pm 1$) and $(1 - \varepsilon_{\text{sig}}^{\text{OST}})$ for untagged events ($q^{\text{OST}} = 0$). Analogously, the normalization factor for the same side kaon

tagger is defined as

$$\mathcal{N}_{\text{tag}}^{\text{SST}} = \frac{|q^{\text{SST}}|\varepsilon_{\text{sig}}^{\text{SST}}}{2} + (1 - |q^{\text{SST}}|) \cdot (1 - \varepsilon_{\text{sig}}^{\text{SST}}). \tag{9.8}$$

Equation 9.5 describes the dependencies of the signal mixing distribution. It has to be normalized to be used as a PDF. The derivation of the full normalization factor $\mathcal{N}$ is described in Appendix B

9.1.2. $B_s^0 \to D_s^\pm K^\mp$ mixing distribution

The mixing distribution of $B_s^0 \to D_s^\pm K^\mp$ decays is similar to the $B_s^0 \to D_s^- \pi^+$ mixing distribution. However $B_s^0 \to D_s^\pm K^\mp$ have an ambiguous final state, meaning that B_s^0 mesons can decay both into $D_s^- K^+$ and $D_s^+ K^-$. Hence in contrast to $B_s^0 \to D_s^- \pi^+$ decays $B_s^0 \to D_s^\pm K^\mp$ decays are not self-tagging, *i.e.* the charge of the final state particles do not give directly the decay flavour of the B_s^0. Therefore, the mixing PDF for $B_s^0 \to D_s^\pm K^\mp$ decays has to include the charge q_K of the bachelor kaon in addition to the tagging decision q.

Another consequence of the ambiguous final state is that time dependent *CP* violation is possible in this decay[2]. It is described by three coefficients, C, D_{eff}, and S_{eff} which are defined as follows

$$C = \frac{1 - \lambda^2}{1 + \lambda^2} \tag{9.9}$$

with λ being the ratio of the amplitudes of the two Feynman diagrams of $B_s^0 \to D_s^- K^+$ and $B_s^0 \to D_s^+ K^-$. It is set to $\lambda = 0.372$ [23].
D_{eff} and S_{eff} depend on the charge of the bachelor kaon $q_K = +1, -1$ for K^+ and K^- respectively

$$D_{\text{eff}} = \frac{1}{2}\left((1 + q_K)D_f + (1 - q_K)D_{\bar{f}}\right), \tag{9.10}$$

$$S_{\text{eff}} = \frac{1}{2}\left((1 + q_K)S_f + (1 - q_K)S_{\bar{f}}\right), \tag{9.11}$$

with the *CP* violation parameters D_f, $D_{\bar{f}}$, S_f and $S_{\bar{f}}$

$$D_f = \frac{2\lambda \cos(\delta_s - (\gamma + \phi_s))}{1 + \lambda^2}, \quad D_{\bar{f}} = \frac{2\lambda \cos(\delta_s + (\gamma + \phi_s))}{1 + \lambda^2} \tag{9.12}$$

$$S_f = \frac{2\lambda \sin(\delta_s - (\gamma + \phi_s))}{1 + \lambda^2}, \quad S_{\bar{f}} = \frac{2\lambda \sin(\delta_s + (\gamma + \phi_s))}{1 + \lambda^2}, \tag{9.13}$$

[2]This makes the PDF considerably more complicated. Since the fraction from $B_s^0 \to D_s^\pm K^\mp$ decays is found to be only 2%, there is no big effect expected on Δm_s from simplifying the mixing PDF and ignoring the ambiguous final state. However, to be consistent with other analyses in LHCb [2], the full description of the $B_s^0 \to D_s^\pm K^\mp$ mixing PDF is included.

which depend on the strong phase δ_s, the CKM-angle γ and the weak phase ϕ_s. These parameters are fixed in the fit to

$$\delta_s = 20^\circ, \;\; \gamma = 60^\circ \;\text{ and }\; \phi_s = -0.02. \tag{9.14}$$

Variations of these parameters showed no effect on the measured value of Δm_s.

With the three parameters C, D_{eff} and S_{eff} the final mixing PDF of $B_s^0 \to D_s^\pm K^\mp$ candidates becomes

$$\begin{aligned}\mathcal{P}_{D_sK;t,q} \propto \Bigg[\Bigg(e^{-\Gamma_s t}\cdot\Bigg[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right) - D_{\text{eff}}\sinh\left(\frac{\Delta\Gamma_s}{2}t\right) \\ - q\cdot(1-2\omega)\cdot\left(C\cos\left(\Delta m_s t\right) - S_{\text{eff}}\sin\left(\Delta m_s t\right)\right)\Bigg]\Bigg) \\ \otimes\,\mathcal{G}(0, S_{\sigma_t}\sigma_t; t)\Bigg]\cdot\mathcal{E}_t(t)\cdot\mathcal{N}_{\text{tag}}^{\text{OST}}\cdot\mathcal{N}_{\text{tag}}^{\text{SST}},\end{aligned} \tag{9.15}$$

using the same decay time resolution and acceptance as for the $B_s^0 \to D_s^-\pi^+$ decays. With $C = 1$, $D_{\text{eff}} = 0$ and $S_{\text{eff}} = 0$ this PDF would be identical to the mixing PDF of $B_s^0 \to D_s^-\pi^+$ decays.

9.2. Mixing PDFs for background decays

In this section the mixing PDFs for the background contributions are described. The PDFs for the $B^0 \to D^-\pi^+$ and $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^-\pi^+$ contributions are motivated from the physical decay time distributions and the flavour tagging behaviour described in Section 6.7. The shape of the decay time distribution of the combinatorial background is motivated from the high-mass sidebands and obtained from a fit to data.

9.2.1. $B^0 \to D^-\pi^+$ mixing distribution

The decay time distribution of $B^0 \to D^-\pi^+$ decays is similar to the $B_s^0 \to D_s^-\pi^+$ signal distribution. The same acceptance and resolution model as for the signal decays are used. The mixing frequency of B^0 mesons is $\Delta m_d = 0.510 \pm 0.004\,\text{ps}^{-1}$, which is about 35 times smaller than Δm_s and the decay width is $\Gamma_d = 0.658 \pm 0.003\,\text{ps}^{-1}$, which is slightly smaller than the B_s^0 decay width Γ_s. The third difference in the parameters is the decay width difference $\Delta\Gamma_d$ which is consistent with zero in the case of B^0 mesons [23]. Thus, the hyperbolic cosine term in the PDF can be assumed to be equal to one.

As described in Section 6.7.2, only the opposite side flavour tagging algorithms give any information about the mixing state of B^0 background candidates. Thus,

the B^0 oscillation can only be resolved by the opposite side taggers. The resulting PDF is

$$\begin{aligned}\mathcal{P}_{B^0;t,q} \propto \Big[\Big(e^{-\Gamma_d t}\cdot\Big[1-q^{\mathrm{OST}}(1-2\omega^{\mathrm{OST}})\cdot\cos\left(\Delta m_d t\right)\Big]\Big)\otimes\mathcal{G}(0,S_{\sigma_t}\sigma_t;t)\Big] \\ \cdot\,\mathcal{E}_t(t)\cdot\mathcal{N}_{\mathrm{tag}}^{\mathrm{OST}},\end{aligned} \tag{9.16}$$

with the calibrated mistag probability of the opposite side taggers, ω^{OST}

$$\omega^{\mathrm{OST}} = p_0^{\mathrm{OST}} + (\eta^{\mathrm{OST}} - \langle\eta^{\mathrm{OST}}\rangle) + p_1^{\mathrm{OST}}\ , \tag{9.17}$$

and with the same relative normalization factor $\mathcal{N}_{\mathrm{tag}}^{\mathrm{OST}}$ as for the signal decays.

Since in the hadronization of B^0 mesons no s quark is involved, the same side kaon tagger is not expected to yield any information about the production flavour of B^0 mesons and, hence, the mixing state of the B^0 background candidate. Nevertheless, the response given by the same side kaon tagger has to be taken into account for the relative normalization.

Therefore, the mixing PDF is multiplied by a factor $\mathcal{N}_{\mathrm{tag}B^0}^{\mathrm{SST}}$ defined as

$$\mathcal{N}_{\mathrm{tag}B^0}^{\mathrm{SST}} = q^{\mathrm{SST}}\cdot\left(\omega_{B^0}^{\mathrm{SST}} + \frac{(q^{\mathrm{SST}}-1)}{2}\right)\frac{\varepsilon_{B^0}^{\mathrm{SST}}}{2} + (1-|q^{\mathrm{SST}}|)\cdot(1-\varepsilon_{B^0}^{\mathrm{SST}}) \tag{9.18}$$

which is equal to $\omega_{B^0}^{\mathrm{SST}}\frac{\varepsilon_{B^0}^{\mathrm{SST}}}{2}$ for events tagged as mixed ($q^{\mathrm{SST}}=+1$). It is equal to $(1-\omega_{B^0}^{\mathrm{SST}})\frac{\varepsilon_{B^0}^{\mathrm{SST}}}{2}$ for events tagged as unmixed ($q^{\mathrm{SST}}=-1$) and $(1-\varepsilon_{B^0}^{\mathrm{SST}})$ for untagged events ($q^{\mathrm{SST}}=0$).

The parameters Δm_d, Γ_d and $\Delta\Gamma_d$ are fixed in the fit to their nominal values. The additional parameters from this background contribution which are floating in the fit are the tagging efficiency $\varepsilon_{B^0}^{\mathrm{SST}}$ and the asymmetry parameter $\omega_{B^0}^{\mathrm{SST}}$. Every decision from the same side kaon tagger in B^0 decays is based on the wrong assumption of a charged kaon produced in association with the signal B meson. Therefore, the same side kaon tagging behaviour is expected to be random and the parameter $\omega_{B^0}^{\mathrm{SST}}$ is expected to be consistent with 0.5, which will be checked.

9.2.2. $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ mixing distribution

The mixing PDF for $\overline{\Lambda}_b^0 \to \overline{\Lambda}_c^- \pi^+$ decays is the same as the PDF for $B^0 \to D^- \pi^+$ events except for the fact that Λ_b^0 baryons do not oscillate. Thus the decay time distribution of this background contribution is a simple exponential. As for the $B^0 \to D^- \pi^+$ decays, the same decay time acceptance and resolution model as for signal candidates is assumed. The decay width $\Gamma_{\Lambda_b^0}$ is fixed to the nominal value ($0.700 \pm 0.012\,\mathrm{ps}^{-1}$ [23]).

As described in Section 6.7.3 only the opposite side taggers are assumed to carry information about the production flavour of the Λ_b^0 background candidates. Due to the absence of oscillation in this decay all Λ_b^0 baryons that were tagged correctly have tagging decision $q^{\mathrm{OST}} = -1$ and all events tagged wrongly have $q^{\mathrm{OST}} = +1$. Thus the final PDF is

$$\mathcal{P}_{\Lambda_b^0;t,q} \propto \left(e^{-\Gamma_{\Lambda_b^0} t} \otimes \mathcal{G}(0, S_{\sigma_t}\sigma_t; t) \right) \cdot \mathcal{E}_t(t) \cdot \mathcal{N}^{\mathrm{OST}}_{\mathrm{tag}\Lambda_b^0} \cdot \mathcal{N}^{\mathrm{SST}}_{\mathrm{tag}\Lambda_b^0}, \tag{9.19}$$

with the normalization factor

$$\mathcal{N}^{\mathrm{OST}}_{\mathrm{tag}\Lambda_b^0} = q^{\mathrm{OST}} \cdot \left(\frac{(q^{\mathrm{OST}}+1)}{2} - \omega^{\mathrm{OST}} \right) \frac{\varepsilon^{\mathrm{OST}}_{\mathrm{sig}}}{2} + (1 - |q^{\mathrm{OST}}|) \cdot (1 - \varepsilon^{\mathrm{OST}}_{\mathrm{sig}}). \tag{9.20}$$

The normalization factor for the same side kaon tagger $\mathcal{N}^{\mathrm{SST}}_{\mathrm{tag}\Lambda_b^0}$ is defined analogously to $\mathcal{N}^{\mathrm{SST}}_{\mathrm{tag}B^0}$ with its parameters $\varepsilon^{\mathrm{SST}}_{\Lambda_b^0}$ and $\omega^{\mathrm{SST}}_{\Lambda_b^0}$

$$\begin{aligned} \mathcal{N}^{\mathrm{SST}}_{\mathrm{tag}\Lambda_b^0} =& q^{\mathrm{SST}} \cdot \left(\omega^{\mathrm{SST}}_{\Lambda_b^0} + \frac{(q^{\mathrm{SST}}-1)}{2} \right) \frac{\varepsilon^{\mathrm{SST}}_{\Lambda_b^0}}{2} \\ &+ (1 - |q^{\mathrm{SST}}|) \cdot (1 - \varepsilon^{\mathrm{SST}}_{\Lambda_b^0}). \end{aligned} \tag{9.21}$$

The additional parameters from this background contribution which are floating in the fit are the tagging efficiency $\varepsilon^{\mathrm{SST}}_{\Lambda_b^0}$ and the asymmetry parameter $\omega^{\mathrm{SST}}_{\Lambda_b^0}$. As for B^0 meson decays, the same side kaon tagging behaviour is expected to be random and the parameter $\omega^{\mathrm{SST}}_{\Lambda_b^0}$ is expected to be consistent with 0.5, which will be checked.

9.2.3. Combinatorial background mixing distribution

The mixing PDF of combinatorial background candidates is determined from the B_s^0 candidates in the high-mass sideband ($m > 5500\,\mathrm{MeV}/c^2$). The response for both the opposite side and same side tagging algorithms is assumed to be random. Therefore, the only contribution from the flavour tagging algorithms to the mixing PDF of combinatorial background candidates are the normalization factors $\mathcal{N}^{\mathrm{OST}}_{\mathrm{tag;comb}}$ and $\mathcal{N}^{\mathrm{SST}}_{\mathrm{tag;comb}}$

$$\mathcal{P}_{\mathrm{comb};t,q} = \mathcal{P}_{\mathrm{comb};t} \cdot \mathcal{N}^{\mathrm{OST}}_{\mathrm{tag;comb}} \cdot \mathcal{N}^{\mathrm{SST}}_{\mathrm{tag;comb}} \tag{9.22}$$

which are defined as:

$$\begin{aligned} \mathcal{N}^{\mathrm{OST}}_{\mathrm{tag;comb}} =& q^{\mathrm{OST}} \cdot \left(\omega^{\mathrm{OST}}_{\mathrm{comb}} + \frac{(q^{\mathrm{OST}}-1)}{2} \right) \frac{\varepsilon^{\mathrm{OST}}_{\mathrm{tag;comb}}}{2} \\ &+ (1 - |q^{\mathrm{OST}}|) \cdot (1 - \varepsilon^{\mathrm{OST}}_{\mathrm{tag;comb}}) \end{aligned} \tag{9.23}$$

and

$$\begin{aligned}\mathcal{N}^{\mathrm{SST}}_{\mathrm{tag;comb}} =& q^{\mathrm{SST}} \cdot \left(\omega^{\mathrm{SST}}_{\mathrm{comb}} + \frac{(q^{\mathrm{SST}}-1)}{2}\right) \frac{\varepsilon^{\mathrm{SST}}_{\mathrm{tag;comb}}}{2} \\ &+ (1-|q^{\mathrm{SST}}|)\cdot(1-\varepsilon^{\mathrm{SST}}_{\mathrm{tag;comb}}).\end{aligned} \tag{9.24}$$

with the efficiencies, $\varepsilon^{\mathrm{OST}}_{\mathrm{tag;comb}}$ and $\varepsilon^{\mathrm{SST}}_{\mathrm{tag;comb}}$, and the asymmetry parameters, $\omega^{\mathrm{OST}}_{\mathrm{comb}}$ and $\omega^{\mathrm{SST}}_{\mathrm{comb}}$, which are free parameters in the fit. Since the tagging behaviour is expected to be random, also the parameters $\omega^{\mathrm{OST}}_{\mathrm{comb}}$ and $\omega^{\mathrm{SST}}_{\mathrm{comb}}$ are expected to be consistent with 0.5. Additionally since combinatorial background candidates are random combination of particles in the event, they are more likely to pass the signal selection in events with large activity which result in many particles in the detector. Such events have also a higher probability to find a tagging candidate. Therefore, the parameters $\varepsilon^{\mathrm{OST}}_{\mathrm{tag;comb}}$ and $\varepsilon^{\mathrm{SST}}_{\mathrm{tag;comb}}$ are expected to be larger than the corresponding signal parameters.

The decay time distribution for events in the high-mass sideband is given in Figure 9.1 for the five D_s^- decay modes. In each histogram the fitted distribution is parametrized as

$$\mathcal{P}_{\mathrm{comb};t} \propto (t-a^{\mathrm{DT}}_{\mathrm{comb}})^2 \cdot (f^{\mathrm{DT}}_{\mathrm{comb}} e^{-\alpha^{\mathrm{DT}}_{\mathrm{comb}} t} + (1-f^{\mathrm{DT}}_{\mathrm{comb}}) e^{-\beta^{\mathrm{DT}}_{\mathrm{comb}} t}) \tag{9.25}$$

Due to large correlations between the parameters $f^{\mathrm{DT}}_{\mathrm{comb}}$, $\alpha^{\mathrm{DT}}_{\mathrm{comb}}$ and $\beta^{\mathrm{DT}}_{\mathrm{comb}}$ these parameters cannot be fitted at the same time. Therefore, it was decided to fix $\alpha^{\mathrm{DT}}_{\mathrm{comb}}$ to $0.4 \cdot \beta^{\mathrm{DT}}_{\mathrm{comb}}$ and $f^{\mathrm{DT}}_{\mathrm{comb}}$ to 0.02. These are the values obtained from a fit to the sum of the three $D_s^- \to K^+K^-\pi^-$ modes and the $D_s^- \to K^-\pi^+\pi^-$ mode. The $D_s^- \to \pi^-\pi^+\pi^-$ mode is found to be incompatible to the other modes due to the different selection criteria (see Figure 9.1(f)). The fits shown in Figure 9.1 have $a^{\mathrm{DT}}_{\mathrm{comb}}$ and $\beta^{\mathrm{DT}}_{\mathrm{comb}}$ as free individual parameters. In the final fit on data, the floating parameters $a^{\mathrm{DT}}_{\mathrm{comb}}$ and $\beta^{\mathrm{DT}}_{\mathrm{comb}}$ are common among the four aforementioned modes and the $D_s^- \to \pi^-\pi^+\pi^-$ decay mode has separate parameters $a^{\mathrm{DT}}_{\mathrm{comb};3\pi}$ and $\beta^{\mathrm{DT}}_{\mathrm{comb};3\pi}$.

9.3. PDFs for event-by-event quantities

Three event-by-event variables are included in the mixing PDFs: the decay time error estimate, σ_t, and the mistag probability estimates, η_{OST} and η_{SST}. As mentioned in Section 4.2, it has been shown [58] that additional PDFs must be included as conditional probabilities to avoid biased results if event-by-event variables are used in the PDFs. These PDFs are the normalized distributions of the event-by-event variables for signal and background components. Possible biases in the fit results only occur if an event-by-event variable is distributed

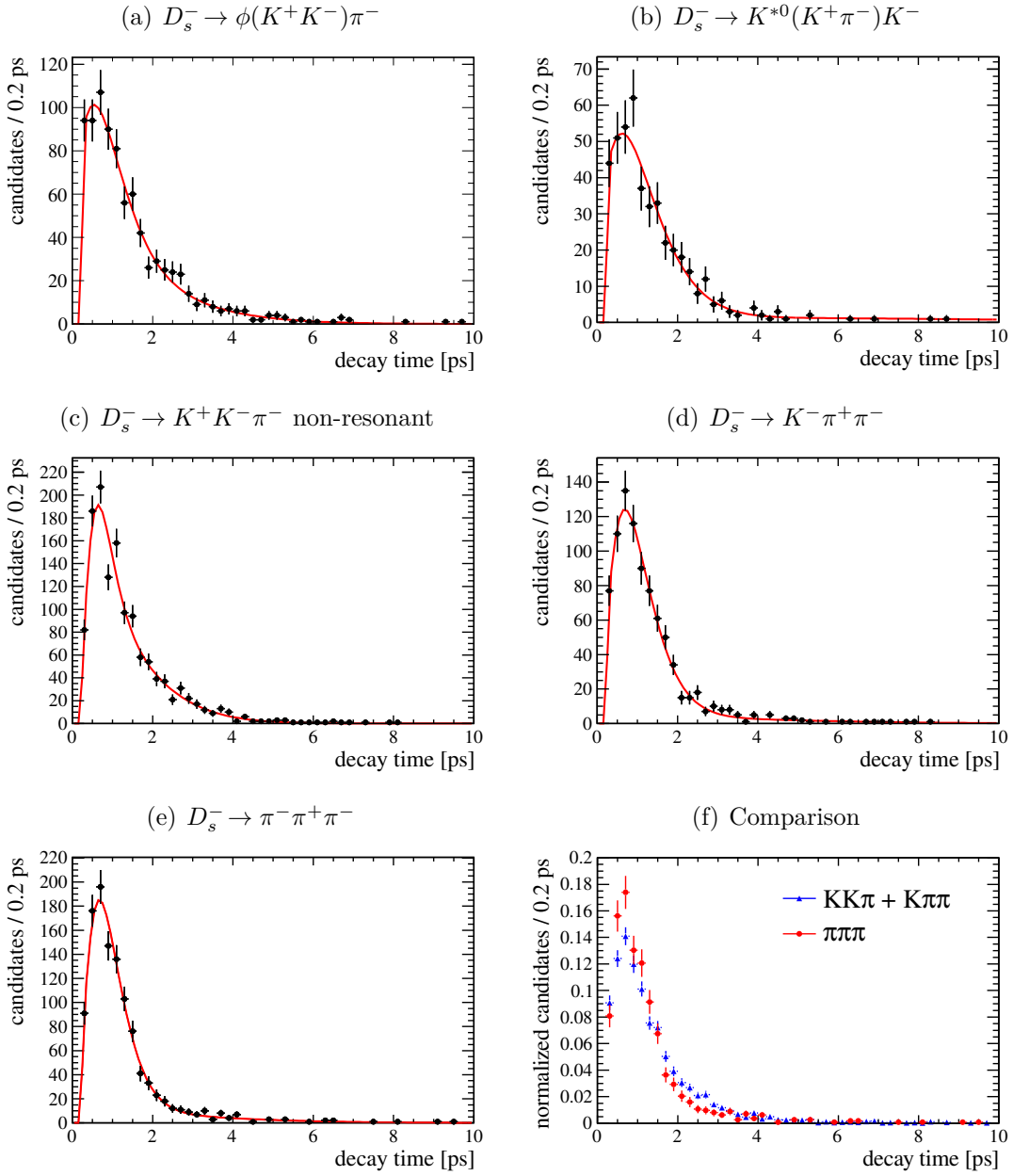


Figure 9.1.: *Decay time distribution for candidates in the high-mass sideband [5500, 5980]*MeV/c^2 *in (a) the* $D_s^- \to \phi(K^+K^-)\pi^-$ *decay, (b) the* $D_s^- \to K^{*0}(K^+\pi^-)K^-$ *decay, (c) the* $D_s^- \to K^+K^-\pi^-$ *non-resonant decay, (d) the* $D_s^- \to K^-\pi^+\pi^-$ *decay and (e) the* $D_s^- \to \pi^-\pi^+\pi^-$ *decay in data. (f) shows the comparison of the* $D_s^- \to \pi^-\pi^+\pi^-$ *decay mode with the sum of the other* D_s^- *decay modes.*

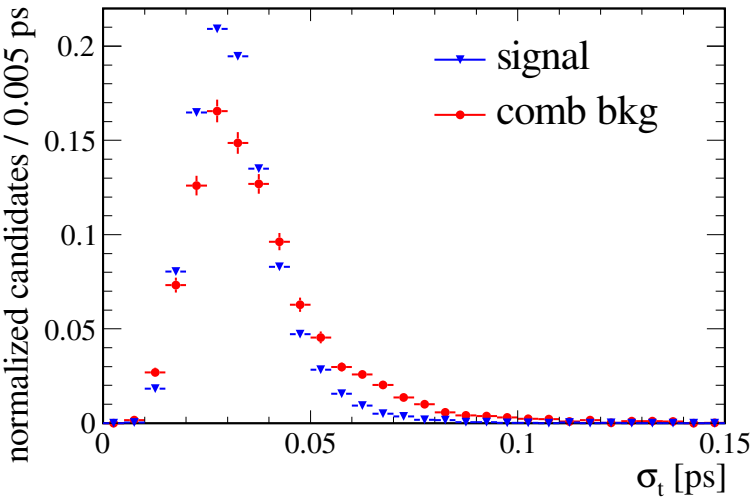


Figure 9.2.: *Normalized distributions of the decay time resolution estimate, σ_t, used in the fit for signal (blue triangles) and combinatorial background (red circles) candidates in data. The background candidates are obtained from the high-mass sideband [5500, 5980]*MeV/c^2.

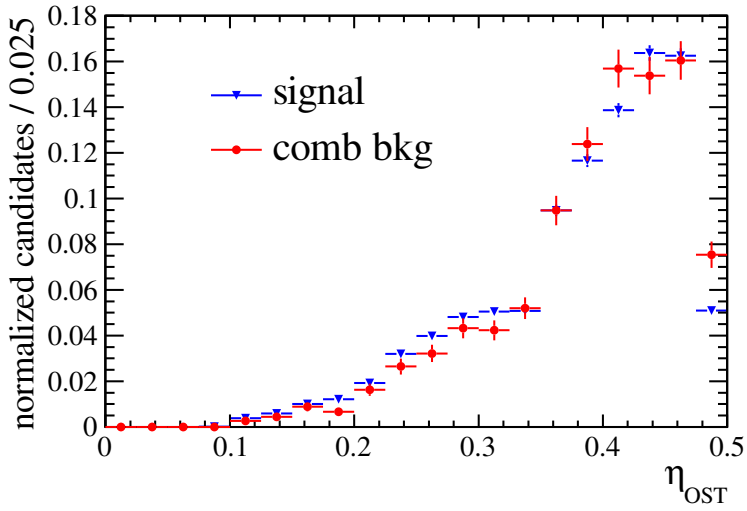
differently for the signal components and the background components and the conditional probabilities are not included.

For this analysis the distributions $\mathcal{P}_{X;\sigma_t}$, $\mathcal{P}_{X;\eta_{\mathrm{OST}}}$ and $\mathcal{P}_{X;\eta_{\mathrm{SST}}}$ are taken from data. For the combinatorial background they are taken from the high-mass sideband. For the signal contributions, the combinatorial background is statistically subtracted from the signal region. For the misid. backgrounds, the indivdual event-by-event variables are treated differently. The decay time resolution estimate distribution and the opposite side tagger mistag estimate distribution are assumed to be identical for the misid. backgrounds and signal. The distribution of the same side kaon tagger mistag estimate is assumed to be identical for the misid. backgrounds and the combinatorial background.

Therefore, for each variable there is one signal and one background PDF. They are shown in Figures 9.2 and 9.3. It can be seen that the σ_t distributions for signal and background are significantly different, whereas for the mistag estimates they are very similar. It should be noted that as for the flavour tagging decision is expected to be random in background events[3], the predicted mistag probability is also wrong in these decays. It is the output of a neural net with random numbers as input. Therefore, also the distribution of these predicted mistag fractions should not be interpreted in any physical way. The point of including the distribution as

[3]Background refers in this context to non-B-meson decays in case of the opposite side tagger and non-B_s^0-meson decays in case of the same side kaon tagger

(a) estimated mistag probability η_{OST}



(b) estimated mistag probability η_{SST}

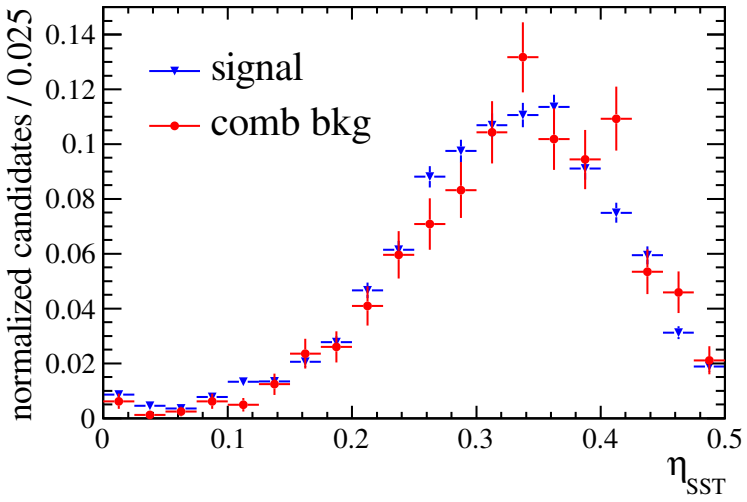


Figure 9.3.: *Normalized distributions of the mistag estimate given by the opposite side taggers, η_{OST}, and the same side kaon tagger, η_{SST}, used in the fit for signal (blue triangles) and background (red circles) candidates in data. The combinatorial background candidates are obtained from the high-mass sideband [5500, 5980]*MeV/c^2.

conditional PDFs, is just that if they are different from the signal distribution, the result on Δm_s would be biased purely due to mathematical reasons of using the maximum likelihood technique.

The assumptions for the misid. backgrounds are based on simulated data samples. However, the statistics of the samples are limited and the behaviour of flavour taggers are also not too reliable on simulation. Therefore, it has been decided to conservatively estimate the systematic effect by completely neglecting them (see Chapter 11).

9.4. Results of the mixing parameters in data

The measurement of Δm_s was performed "blinded", *i.e.* throughout the analysis process the result of Δm_s was shifted by an unknown value. The reason to do this, is to not introduce any biases in the fit results if the result does (not) agree with the expectations. The whole process of the analysis should be finalized including all cross checks before the final result of Δm_s is revealed. After "unblinding" the result of the analysis must not be modified.

The final results for the free parameters in the mixing PDFs for the signal and background components are summarized in Table 9.2. It can be seen that, as expected, for the parameters describing the tagging behaviour of the background contributions, $\omega^{\mathrm{OST}}_{\mathrm{comb}}$, $\omega^{\mathrm{SST}}_{\mathrm{comb}}$, $\omega^{\mathrm{SST}}_{B^0}$, and $\omega^{\mathrm{SST}}_{\Lambda^0_b}$, the results are compatible with 0.5. If these parameters would not be compatible with 0.5, this would point to another background contribution that is not described by the PDFs. Additionally, it can be seen that, as expected, the tagging efficiencies for the combinatorial background candidates, $\varepsilon^{\mathrm{OST}}_{\mathrm{tag;comb}}$ and $\varepsilon^{\mathrm{SST}}_{\mathrm{tag;comb}}$, are indeed larger than the corresponding signal parameters, $\varepsilon^{\mathrm{OST}}_{\mathrm{sig}}$ and $\varepsilon^{\mathrm{SST}}_{\mathrm{sig}}$. So it can be summarized that the results of the tagging parameters of the background are reassuring that the backgrounds are correctly described.

Figure 9.4 shows the decay time distributions of the five D_s^- decay modes together with the projections of the fit results. It can be seen that the decay time distributions are well reproduced by the fitted functions. For illustration purposes these distributions are plotted only for events in the signal region $m \in [5320, 5450]\,\mathrm{MeV}/c^2$. Figure 9.5 shows the decay time distribution for all candidates in the signal invariant mass window that pass the signal selection described in Chapter 5 and the sum of the decay time PDFs for the five D_s^- decay modes.

As a cross check, the fit is performed using only the opposite side taggers and only the same side kaon tagger. This does not check the result on Δm_s, since lower quality of flavour tagging has no impact on Δm_s. It only influences parameters

Parameters of the signal mixing PDF	
Δm_s [ps^{-1}]	17.768±0.023
$\alpha_{\rm acc}$ [ps]	0.771±0.016
$\varepsilon_{\rm sig}^{\rm OST}$	0.3957±0.0026
$\varepsilon_{\rm sig}^{\rm SST}$	0.1518±0.0022
$p_1^{\rm OST}$	0.96±0.11
$p_0^{\rm OST}$	0.4105±0.0093
$p_1^{\rm SST}$	0.58±0.14
$p_0^{\rm SST}$	0.377±0.013
Parameters of the combinatorial background mixing PDF	
$a_{\rm comb}^{\rm DT}$ [ps]	−0.056±0.021
$\beta_{\rm comb}^{\rm DT}$ [ps^{-1}]	2.872±0.065
$a_{\rm comb;3\pi}^{\rm DT}$ [ps]	0.069±0.019
$\beta_{\rm comb;3\pi}^{\rm DT}$ [ps^{-1}]	3.85±0.13
$\varepsilon_{\rm tag;comb}^{\rm OST}$	0.4962±0.0060
$\varepsilon_{\rm tag;comb}^{\rm SST}$	0.1752±0.0048
$\omega_{\rm comb}^{\rm OST}$	0.5042±0.0084
$\omega_{\rm comb}^{\rm SST}$	0.490±0.015
Parameters of the misid. backgrounds mixing PDFs	
$\varepsilon_{B^0}^{\rm SST}$	0.121±0.034
$\omega_{B^0}^{\rm SST}$	0.27±0.15
$\varepsilon_{\Lambda_b^0}^{\rm SST}$	0.155±0.022
$\omega_{\Lambda_b^0}^{\rm SST}$	0.471±0.077

Table 9.2.: *Results of the mixing parameters in the fit of the five $B_s^0 \to D_s^- \pi^+$ decay modes using both opposite and same side kaon tagging algorithms.*

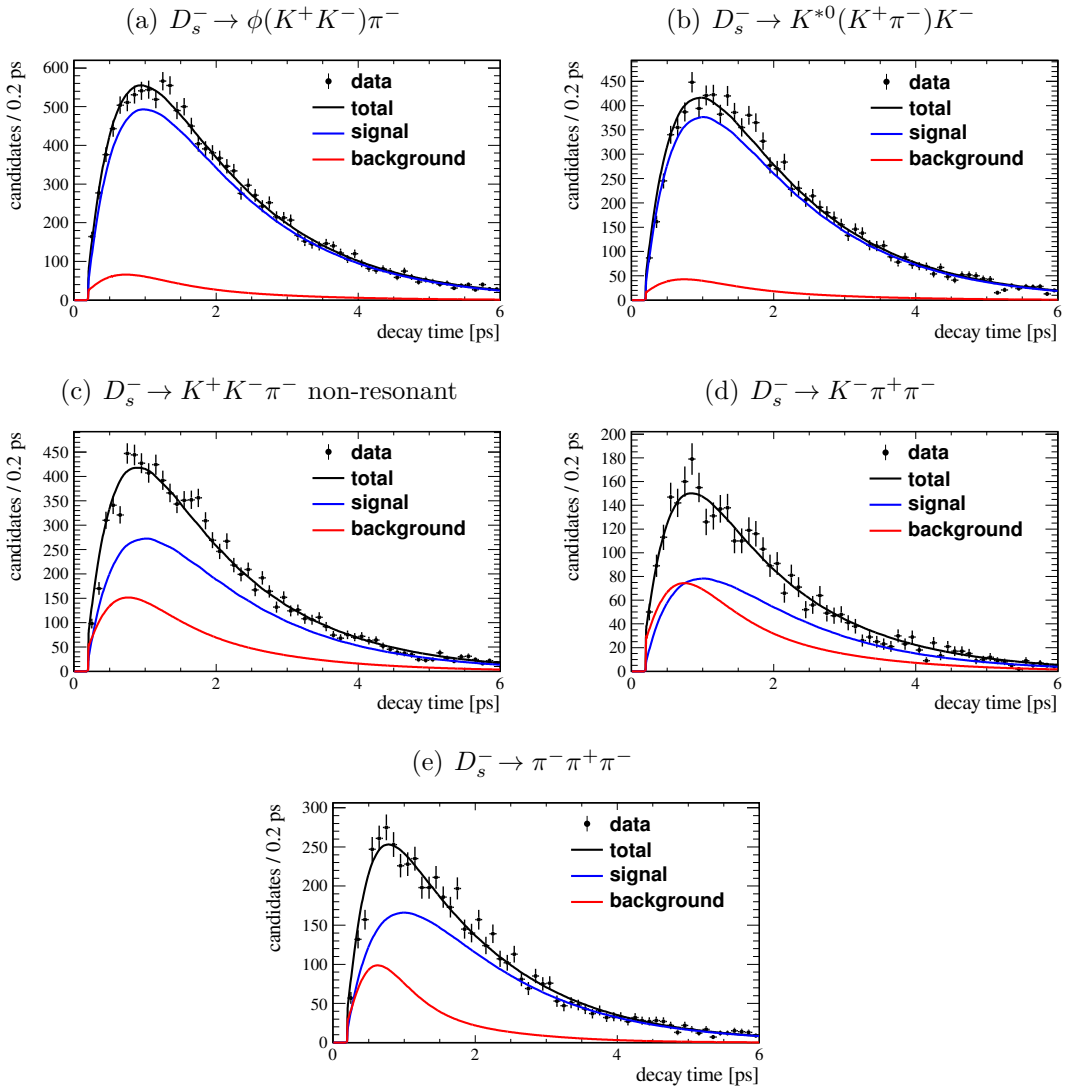


Figure 9.4.: *Projection of the fitted decay time distribution in (a) the $D_s^- \to \phi(K^+K^-)\pi^-$ decay, (b) the $D_s^- \to K^{*0}(K^+\pi^-)K^-$ decay, (c) the $D_s^- \to K^+K^-\pi^-$ non-resonant decay, (d) the $D_s^- \to K^-\pi^+\pi^-$ decay and (e) the $D_s^- \to \pi^-\pi^+\pi^-$ decay in data.*

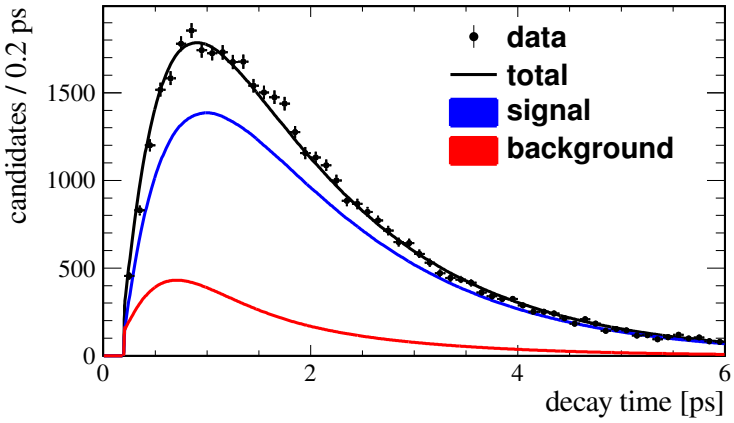


Figure 9.5.: *Projection of the fitted decay time distributions for the sum of the five D_s^- decay modes.*

that are related to the amplitude of the oscillation and not the frequency. So it should be seen as a cross check that nothing goes wrong in the progress of combining the opposite and same side kaon taggers. The results for the fit using only the opposite side taggers are summarized in Table 9.3 and the results for the fit using only the same side kaon tagger in Table 9.4. It can be seen that the results for all tagging parameters are in good agreement.

In Figure 9.6 the mixing distributions for tagged events are shown for all candidates in the invariant mass signal window. Overlaid are the fit projections on the mixing dimension of decay time and flavour tagging decision for the three fits. Figure 9.6(a) shows the nominal fit using both the opposite side tagging algorithms and the same side kaon tagger. Figures 9.6(b) and 9.6(c) show the fits using only the opposite side tagger and only the same side kaon tagger respectively. It can be seen that for all three fits the oscillation can be resolved over many periods.

The effective tagging efficiency ε_{eff}, defined in Equation 6.3, for each tagger is computed as

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} \cdot \frac{1}{N_{\text{sig}}} \sum_{i}^{\text{signal candidates}} \mathcal{D}_i^2 = \varepsilon_{\text{tag}} \cdot \frac{1}{N_{\text{sig}}} \sum_{i}^{\text{signal candidates}} (1 - 2\omega_i)^2. \tag{9.26}$$

with the number of signal candidates N_{sig} and the calibrated per-event mistag probabilities $\omega_i = p_0 + (\eta_i - \langle\eta\rangle) \cdot p_1$. The sum over the signal candidates is performed by statistically subtracting the background. The result is

$$\varepsilon_{\text{eff}}^{\text{OST}} = 2.4 \pm 0.4\% \tag{9.27}$$

for the opposite side taggers and

$$\varepsilon_{\text{eff}}^{\text{SST}} = 1.2 \pm 0.3\% \tag{9.28}$$

for the same side kaon tagger.

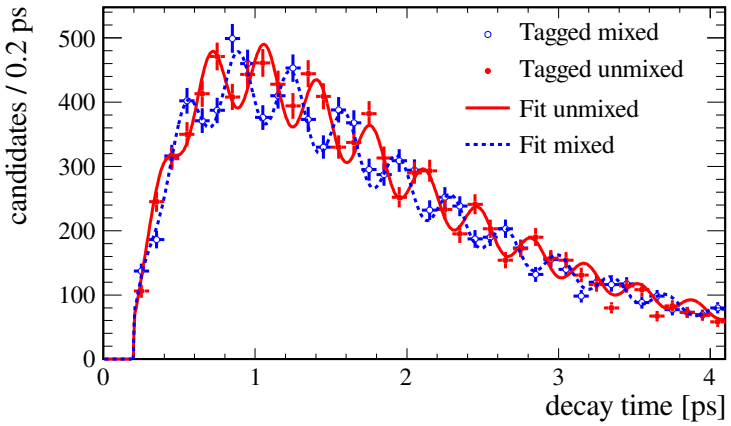
Parameters of the signal mixing PDF	
Δm_s [ps^{-1}]	17.781±0.027
$\alpha_{\rm acc}$ [ps]	0.772±0.016
$\varepsilon_{\rm sig}^{\rm OST}$	0.3957±0.0026
$p_1^{\rm OST}$	0.98±0.11
$p_0^{\rm OST}$	0.4102±0.0092
Parameters of the combinatorial background mixing PDF	
$a_{\rm comb}^{\rm DT}$ [ps]	−0.058±0.021
$\beta_{\rm comb}^{\rm DT}$ [ps^{-1}]	2.866±0.064
$a_{{\rm comb};3\pi}^{\rm DT}$ [ps]	0.068±0.019
$\beta_{{\rm comb};3\pi}^{\rm DT}$ [ps^{-1}]	3.83±0.13
$\varepsilon_{\rm tag;comb}^{\rm OST}$	0.4959±0.0060
$\omega_{\rm comb}^{\rm OST}$	0.5054±0.0084

Table 9.3.: *Results of the mixing parameters in the fit of the five $B_s^0 \to D_s^- \pi^+$ decay modes using only opposite side tagging algorithms.*

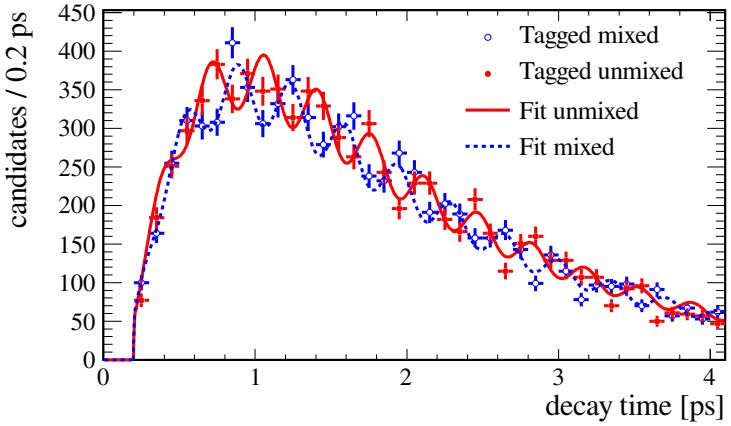
Parameters of the signal mixing PDF	
Δm_s [ps^{-1}]	17.745±0.042
$\alpha_{\rm acc}$ [ps]	0.769±0.016
$\varepsilon_{\rm sig}^{\rm SST}$	0.1519±0.0022
$p_1^{\rm SST}$	0.59±0.14
$p_0^{\rm SST}$	0.376±0.013
Parameters of the combinatorial background mixing PDF	
$a_{\rm comb}^{\rm DT}$ [ps]	−0.058±0.021
$\beta_{\rm comb}^{\rm DT}$ [ps^{-1}]	2.866±0.065
$a_{{\rm comb};3\pi}^{\rm DT}$ [ps]	0.068±0.019
$\beta_{{\rm comb};3\pi}^{\rm DT}$ [ps^{-1}]	3.84±0.13
$\varepsilon_{\rm tag;comb}^{\rm SST}$	0.1751±0.0048
$\omega_{\rm comb}^{\rm SST}$	0.487±0.015
Parameters of the misid. backgrounds mixing PDFs	
$\varepsilon_{B^0}^{\rm SST}$	0.128±0.033
$\omega_{B^0}^{\rm SST}$	0.31±0.13
$\varepsilon_{\Lambda_b^0}^{\rm SST}$	0.150±0.022
$\omega_{\Lambda_b^0}^{\rm SST}$	0.490±0.079

Table 9.4.: *Results of the mixing parameters in the fit of the five $B_s^0 \to D_s^- \pi^+$ decay modes using only the same side kaon tagging algorithm.*

(a) Nominal Fit



(b) only OST



(c) only SST

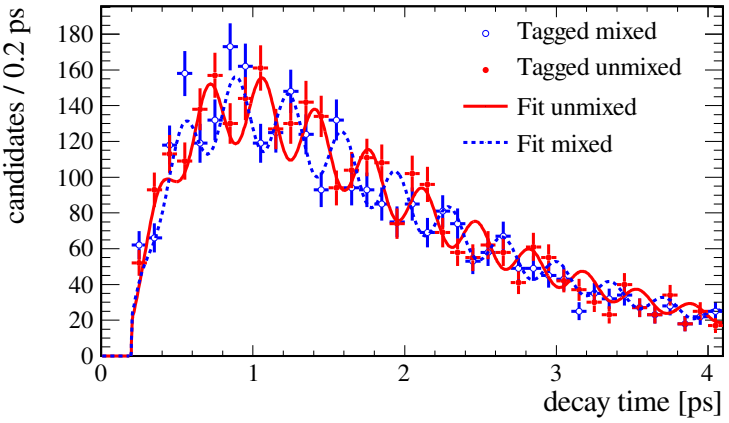


Figure 9.6.: *Projection of the fitted mixing distributions for tagged candidates from the sum of the five D_s^- decay modes in the nominal fit using opposite (OST) and same side kaon (SST) taggers (top), using only the opposite side taggers (middle) and using only the same side kaon tagger (bottom).*

CHAPTER 10

Validation of the Fit Procedure

In this chapter a consistency check to validate the fit procedure is presented. The goal is to show that all PDFs described in Chapters 8 and 9 are implemented correctly in the fit algorithm and the fitter gives correct results for Δm_s with the correct uncertainty estimate.

The fit procedure is validated using pseudo-experiments, so-called "toy-experiments". In these toy-experiments values for the observables:

$$m, t, q, \sigma_t, \eta_{\mathrm{OST}} \text{ and } \eta_{\mathrm{SST}}$$

are generated according to the PDFs used in the fit. For the generation of these distributions random number generators are used. This way, the resulting data samples have the assumed distribution with just statistical fluctuations. This generated data sample is then fitted in the same way the real data sample is fitted which is described in chapters 8 and 9. The fit results of this pseudo-experiment

$$\vec{\lambda}^{\mathrm{Fit}} = \{\lambda_1^{\mathrm{Fit}}, \lambda_2^{\mathrm{Fit}}, ...\}$$

is saved together with the uncertainty estimate on each fitted parameter, σ_{λ_i}, and the value used in the generation,

$$\vec{\lambda}^{\mathrm{Gen}} = \{\lambda_1^{\mathrm{Gen}}, \lambda_2^{\mathrm{Gen}}, ...\}.$$

Using these quantities the pull of each parameter is calculated as

$$\mathrm{pull}_i = \frac{\lambda_i^{\mathrm{Fit}} - \lambda_i^{\mathrm{Gen}}}{\sigma_{\lambda_i}}. \tag{10.1}$$

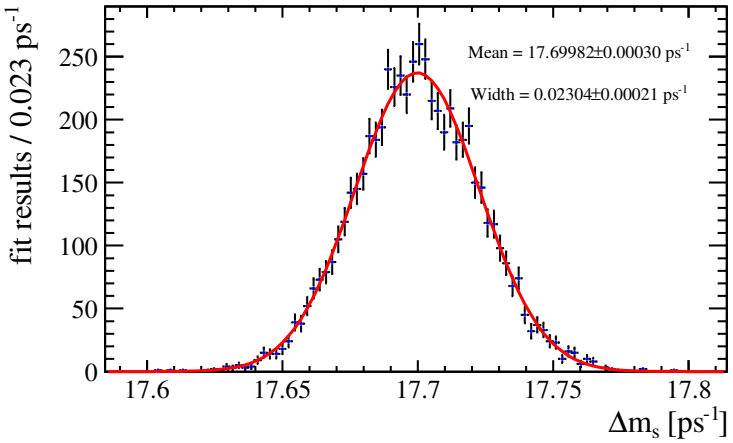


Figure 10.1.: *Results of 6000 toy-experiments. The distribution of the fit results for Δm_s fitted with a Gaussian distribution are shown. The generated value has been set to Δm_s =17.7ps^{-1}.*

This procedure is repeated six thousand times. If the fit procedure works properly the fit results should follow a Gaussian distribution centered around the generated values. Additionally, the uncertainty estimate given by the fit represents the actual uncertainty of the fitted parameter. If these two conditions are fulfilled, the pulls of the fitted parameters follow a Gaussian distribution with mean $\mu = 0$ and width $\sigma = 1$.

In each toy experiment, the number of generated events is set to the number of events available in data. The fit results of the fit on data for all free parameters are used for these parameters in the generation.

Figure 10.1 shows the distribution of the fitted values for Δm_s in the toy-experiments. Figure 10.2 shows the distribution of the uncertainty estimates given by the fit. Finally, Figure 10.3 shows the pull distribution of Δm_s.

It can be seen that the pull distribution is well described by a Gaussian with a mean compatible with zero and a width compatible with one. The statistical uncertainty given by the fit on data (0.023 ps^{-1}, see Table 9.2) is very close to the mean of the distribution of uncertainties in Figure 10.2. This is reassuring that the uncertainty given by the fit on data represents the real statistical uncertainty on this measurement of Δm_s.

This method uses the same physical models in generation and fit. It is meant as a consistency check that the models are implemented correctly in the fit algorithm. This method does not probe effects due to wrong underlying theory models. Therefore, in the next chapter additional systematic effects are studied.

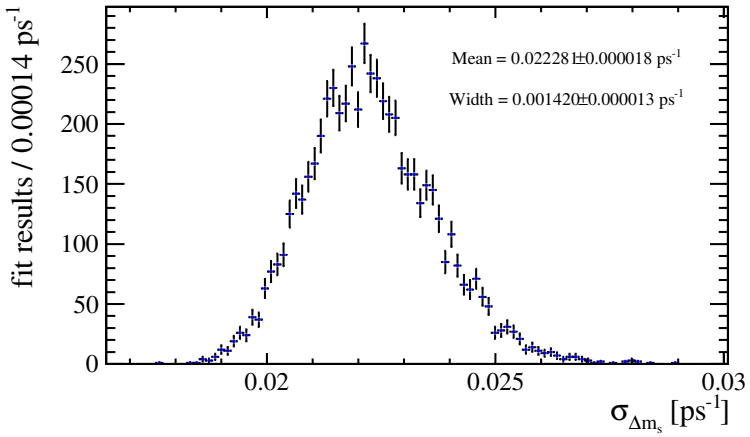


Figure 10.2.: *Results of 6000 toy-experiments. The distribution of the uncertainty estimate on Δm_s given by the fit algorithm are shown.*

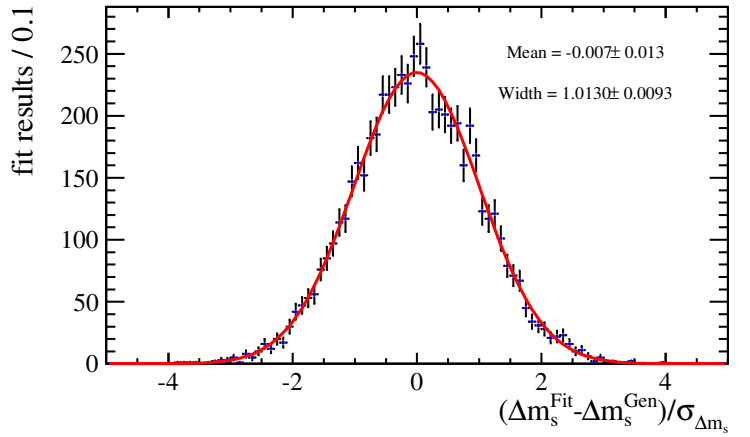


Figure 10.3.: *Results of 6000 toy-experiments. The pull distribution for Δm_s as defined in Eq 10.1 is shown. The distribution is fitted with a Gaussian distribution.*

CHAPTER 11

Systematic Studies

In this chapter the studies of systematic uncertainties on the measurement of Δm_s are summarized. Systematic uncertainties are all effects that can bias the result of Δm_s which are not due to the statistical limitation of the data sample used.

One source for such an effect is fixing some of the fit parameters to values obtained from external studies with finite precision. In this study the decay width Γ_s and the decay width difference $\Delta\Gamma_s$ are such fixed parameters.

Another possible source of systematic effects is wrong modelling of the signal or background components in the PDFs. For the signal components the models studied are the decay time acceptance shape, the decay time resolution model and an alternative mass model. For the mass templates of the background components histograms are used. To evaluate the influence of this model, it has been decided to fix the background fraction to values within the fitted uncertainties. Additionally, the effect of the additional PDFs including the conditional probabilities for the event-by-event variables is studied.

A third possible systematic uncertainty arises from the uncertainties on the measured quantities themselves. For example if detector or reconstruction effects bias the measured decay time and the fit functions do not take this into account, the measured value for Δm_s will be shifted from the real value. The effect of such a systematic shift seen on simulated data is investigated. Additionally, the effect from uncertainties on the calibration of the length scale of the detector and the scale of the measured momentum is studied.

The results of these studies are divided into two categories:

- cross-checks where no effect is observed and, thus, no systematic is assigned,
- observed effects with a resulting systematic.

11.1. Cross-checks

These systematic effects are studied using pseudo-experiments as described in Chapter 10. The same number of events as available in data are generated, but with different distributions than used in the fit on data. Then these data samples are fitted with the nominal distributions. The distribution of fit results for Δm_s is compared to the generated value. The uncertainty on the mean of the fitted values of Δm_s for 500 pseudo-experiments is $0.0004\,\mathrm{ps}^{-1}$ which corresponds to 5% of the statistical uncertainty. Therefore, if the mean of the resulting distribution of the fitted values for Δm_s minus the generated value ($\Delta_{\Delta m_s} = \Delta m_s^{\mathrm{Fit}} - \Delta m_s^{\mathrm{Gen}}$) is compatible with zero within 2σ, no systematic effect is assigned. Table 11.1 summarizes the observed shifts $\Delta_{\Delta m_s}$ for the individual cross-checks

11.1.1. Decay width Γ_s and decay width difference $\Delta\Gamma_s$

The value for the B_s^0 decay width Γ_s is fixed to the world average [23]

$$\Gamma_s = 0.6596 \pm 0.0073\,\mathrm{ps}^{-1} \,.$$

The value for the B_s^0 decay width difference $\Delta\Gamma_s$ is fixed to the latest value measured by LHCb [1]

$$\Delta\Gamma_s = 0.106 \pm 0.011(\mathrm{stat}) \pm 0.007(\mathrm{syst})\,\mathrm{ps}^{-1} \,.$$

The systematic effect of fixing Γ_s and $\Delta\Gamma_s$ to wrong values is estimated using four sets of 500 pseudo-experiments. The data samples are generated using values for Γ_s or $\Delta\Gamma_s$ which are one standard deviation[1] above or below the nominal value, respectively.

The resulting distributions of fitted values for $\Delta_{\Delta m_s}^{\Gamma_s/\Delta\Gamma_s}$ are compatible with zero (see Table 11.1). Thus, it is concluded that there is no systematic effect on the measurement of Δm_s from the uncertainty on Γ_s and $\Delta\Gamma_s$.

[1] In case of $\Delta\Gamma_s$ the quadratic sum of statistical and systematic uncertainty was taken.

Systematic	**observed shift $\Delta_{\Delta m_s}$ [ps^{-1}]**
$\Delta^{\Gamma_s-\sigma}_{\Delta m_s}$	-0.0003±0.0004
$\Delta^{\Gamma_s+\sigma}_{\Delta m_s}$	0.0004±0.0004
$\Delta^{\Delta\Gamma_s-\sigma}_{\Delta m_s}$	0.0000±0.0004
$\Delta^{\Delta\Gamma_s+\sigma}_{\Delta m_s}$	-0.0006±0.0004
$\Delta^{\alpha_{\rm acc}+20\%}_{\Delta m_s}$	-0.0005±0.0004
$\Delta^{\alpha_{\rm acc}-20\%}_{\Delta m_s}$	0.0003±0.0004
$\Delta^{S_{\sigma_t}=1.25}_{\Delta m_s}$	0.0003±0.0004
$\Delta^{S_{\sigma_t}=1.45}_{\Delta m_s}$	0.0000±0.0004
$\Delta^{DG,\ \rm reso}_{\Delta m_s}$	-0.0004±0.0004
$\Delta^{DG,\ \rm mass}_{\Delta m_s}$	-0.0001±0.0004
$\Delta^{f_{\rm sig}+3\sigma}_{\Delta m_s}$	0.0007±0.0004
$\Delta^{f_{\rm sig}-3\sigma}_{\Delta m_s}$	-0.0002±0.0004
$\Delta^{f_{D_sK}=0}_{\Delta m_s}$	-0.0006±0.0004
$\Delta^{f_{D_sK}=4\%}_{\Delta m_s}$	0.0000±0.0004
$\Delta^{t+0.2\,\rm fs}_{\Delta m_s}$	-0.0004±0.0004
$\Delta^{t-0.2\,\rm fs}_{\Delta m_s}$	-0.0002±0.0004
$\Delta^{\rm per\text{-}ev}_{\Delta m_s}$	0.0002±0.0004
$\Delta^{\rm corr}_{\Delta m_s}$	-0.0001±0.0004

Table 11.1.: *Mean of the distribution of the observed deviations of the fitted values for Δm_s from the generated values , $\Delta_{\Delta m_s}$, for the individual cross-checks presented in Section 11.1.*

11.1.2. Decay time acceptance

The effect of the decay time acceptance $\mathcal{E}_t(t)$ on the measurement of Δm_s is expected to be negligible since it is the same factor multiplied to both the mixed signal PDF $\mathcal{P}_{D_s\pi;t,q=+1}$ and the unmixed PDF $\mathcal{P}_{D_s\pi;t,q=-1}$. The measurement of Δm_s is however sensitive to the difference between the two PDFs. Therefore, no systematic effect due to a wrong decay time acceptance is expected. This is, however, only true for the fast B_s^0–$\overline{B}_s^0$ oscillation in which the number of candidates that are mixed/unmixed are the same. The measurement of the slow B^0–$\overline{B}^0$ oscillation frequency Δm_d is sensitive also to the overall number of mixed/unmixed candidates and, thus, also to acceptance effects.

This is checked using two sets of 500 pseudo-experiments in which the data samples were generated with the parameter α_{acc} varied $\pm 20\%$ around the value fitted in data. The uncertainty of $\pm 20\%$ includes the variation seen in simulated data when adjusting the IP resolution to better reflect data and for different trigger scenarios. The parameter α_{acc} was chosen because it describes the rising edge of the acceptance curve at low decay times and this is where most of the background is located.

As expected the distributions of deviations $\Delta_{\Delta m_s}^{\alpha_{\text{acc}}}$ from the generated value are compatible with zero. Therefore, no systematic effect on the measurement of Δm_s is assigned due to a wrongly determined decay time acceptance.

11.1.3. Decay time resolution

In Section 7.1 the determination of the decay time resolution in a data-driven study is described. The resulting uncertainty on the calibration parameter S_{σ_t} is $\pm 10\%$. The systematic effect of this uncertainty on the measurement of Δm_s is evaluated using two sets of 500 pseudo-experiments. The data samples are generated using $S_{\sigma_t} = 1.25$ and $S_{\sigma_t} = 1.45$, respectively.

The distribution of $\Delta_{\Delta m_s}^{S_{\sigma_t}}$ are compatible with zero. Thus, it is concluded that the systematic effect on the measurement of Δm_s due to the uncertainty of S_{σ_t} is negligible.

11.1.4. Decay time resolution model

It has been seen that the single Gaussian model used in the nominal fit to describe the decay time resolution, does not completely describe the decay time pull distribution in simulation (see Section 7.1.6). Therefore, the effect of neglecting a contribution of 10% of a second Gaussian distribution with width, that is twice

the nominal width, in the decay time resolution is studied using 500 pseudo-experiment.

The distribution of deviations of the fitted values for Δm_s, $\Delta_{\Delta m_s}^{DG,\,\mathrm{reso}}$ are compatible with zero. Thus, it is concluded that the systematic effect on the measurement of Δm_s due to the decay time resolution model is negligible.

11.1.5. Mass model

The effect of a different invariant mass model for the signal distribution is evaluated by generating data-sets with the sum of two Gaussian distributions as signal PDF instead of the sum of two Crystal Ball functions which is used in the nominal fit. In the fit the nominal PDF is used. This pseudo-experiment is as well repeated 500 times.

The results for Δm_s are compatible with the nominal value (see $\Delta_{\Delta m_s}^{DG,\,\mathrm{mass}}$ in Table 11.1). Thus, it is concluded that the systematic effect on the measurement of Δm_s due to the signal mass model is negligible.

11.1.6. Background yield

Since the signal fractions are free fit parameters, the systematic uncertainty due to higher or lower background contributions than fitted should be already included in the statistical uncertainty. However this does not account for the fact that the invariant mass shapes for the individual background components that were taken from simulated data could potentially not represent the real distributions.

This would result in a wrong signal fraction. Therefore, two sets of 500 pseudo-experiments were performed in which the signal fraction is fixed to the nominal value. In the generation a value of three times the statistical uncertainty larger/smaller than the nominal value was used. This should be a conservative estimation for the systematic effect of a wrongly determined background yield.

The resulting distributions of $\Delta_{\Delta m_s}^{f_{\mathrm{sig}}}$ are compatible with zero. Thus, no systematic effect due to a wrong background yield is assigned to the measurement of Δm_s.

11.1.7. $B_s^0 \to D_s^{\pm} K^{\mp}$ yield

As discussed in Section 8.3 the fraction of the contribution from $B_s^0 \to D_s^{\pm} K^{\mp}$ decays to the signal, $f_{D_s K}$, is fixed in the nominal fit. The systematic effect from the uncertainty on $f_{D_s K}$ on the measurement of Δm_s is evaluated by generating data samples with $f_{D_s K}$ set to 4% and zero respectively.

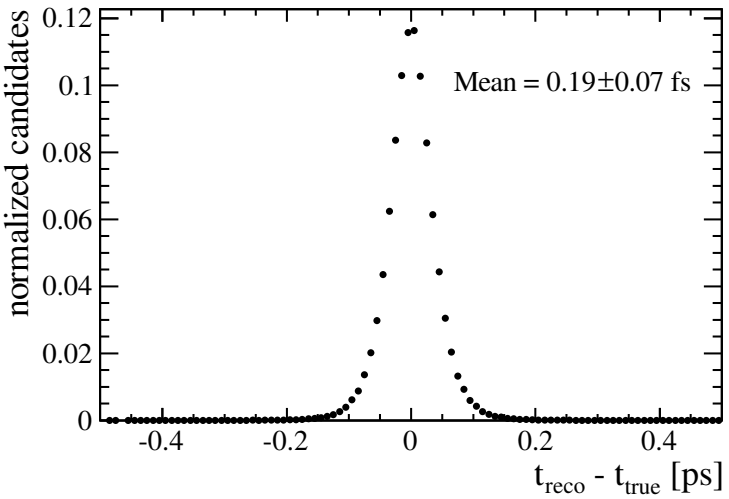


Figure 11.1.: *Difference between reconstructed decay time and true decay time of signal $B_s^0 \to D_s^- \pi^+$ candidates in simulated data after selection.*

The resulting distributions of deviations from the generated value $\Delta_{\Delta m_s}^{f_{D_s K}}$ are compatible with zero. Thus, no systematic effect due to a wrong $B_s^0 \to D_s^{\pm} K^{\mp}$ yield is is assigned.

11.1.8. Decay time bias

In simulated data, the distribution of the reconstructed decay time minus the true decay time, $t_{\text{reco}} - t_{\text{true}}$, after the whole signal selection is applied, is not exactly centered around zero, but slightly shifted to positive values. The observed shift amounts to 0.2 fs (see Figure 11.1).

The systematic effect of such a shift on the measurement of Δm_s is evaluated using 500 pseudo-experiments. The generated decay time is shifted by 0.2 fs and the nominal PDFs are used in the fit. The distributions of fit results for Δm_s in these pseudo-experiments, $\Delta_{\Delta m_s}^{t}$, are compatible with the nominal value. Thus, no systematic due to a decay time bias is assigned.

11.1.9. Decay time uncertainty and predicted η distributions

As mentioned in Section 9.3, it has been shown in [58] that the distributions for event-by-event variables must be included in the PDF as conditional probabilities to avoid biased fit results. The effect of wrongly determined distributions of the

decay time uncertainty estimate σ_t and the predicted mistag probabilities η_{OST} and η_{SST} is determined using pseudo-experiments.

In 500 pseudo-experiments data samples are generated using the nominal σ_t, η_{OST} and η_{SST} distributions. In the fit these distributions are set to be the same for signal and background which is identical to ignoring them completely. The effect on the distribution of the fit results for Δm_s, $\Delta^{\text{per-ev}}_{\Delta m_s}$ is negligible.

11.1.10. Correlation between invariant mass and decay time

As shown in Chapter 9 there is no sizeable correlation between the invariant mass, m, and the decay time, t, for signal decays. However, Table 9.1 shows a correlation between those two variables for combinatorial background candidates. The effect of this correlation is studied, using pseudo-experiments. The combinatorial background candidates are generated according to an exponential distribution with a slope parameter α_{comb}. If the candidate's generated decay time is larger than $t = 2\,\text{ps}$, the invariant mass is generated using a slope parameter which is twice as large as the nominal value

$$\alpha^{t>2\,\text{ps}}_{\text{comb}} = 2 \cdot \alpha^{t<2\,\text{ps}}_{\text{comb}} \tag{11.1}$$

This is a conservative estimation since the largest difference between the two parameters seen in data is a factor 1.75. In the fit, the parameters α_{comb} are floating for all five D_s^- decay modes. The observed shifts, $\Delta^{\text{corr}}_{\Delta m_s}$, are compatible with zero. Thus, no systematic due to the correlation between invariant mass and decay time is assigned.

11.2. z-scale uncertainty

As defined in Equation 4.1 the decay time is proportional to the measured flight distance of the B_s^0 candidate, given by the distance between its production and decay vertex. The accuracy to measure this distance depends on the precision to which the positions of the VELO modules along the beam direction (z-axis) are known.

There are two contributions to this uncertainty. The first is the knowledge of the overall length of the VELO. If the length of the VELO was in reality 1% larger than in the reference frame, this would result in a 1% shorter measured decay time. The precision of the position of the modules along the z-direction describes the overall scaling of the length of the VELO. This has been evaluated in a survey during assembly of the VELO. The measured deviations from the

nominal positions of the modules have been found to be 100 μm of the baseplate over the whole length of the VELO (1 m) [69]. This results in a z-scale uncertainty of

$$\sigma_{\text{survey}} = \frac{100 \times 10^{-3} \text{ mm}}{1000 \text{ mm}} = 0.01\%. \tag{11.2}$$

The second contribution to this systematic effect is due to a possible relative shift of the individual modules within the VELO. This has been evaluated using track based alignment [43, 70, 71] which is dominated by the uncertainty of the relative z-position of the first two modules that were hit by the particle. Only the first two modules are considered because the following modules are weighted down in the track fit due to multiple scattering and don't contribute as much to the track precision. This estimate is conservative since it assumes infinite multiple scattering and in reality the third and fourth modules do improve the precision.

In track based alignment the deviation of the z-position of single modules from their nominal positions are statistically distributed with a width of 20 μm. The systematic effect of the z-scale is in principle given by these 20 μm divided by the spacing between two modules (30 mm). However, not always the same two modules are hit first by the signal tracks. Therefore, this statistical spread is averaged over all VELO modules that are hit first by the signal tracks. If all modules contributed the same to the signal sample, the uncertainty would simply be scaled by $\sqrt{N_{\text{modules}}}$.

Figure 11.2 shows the distribution of the z-position of the first hit on the signal tracks. It can be seen that this distribution is not unitary. In the calculation of the RMS (100 mm) of this distribution, the weighted sum of the individual z-positions is taken into account. Thus, the uncertainty on the z-scale due to track based alignment is given by

$$\sigma_{\text{track}} = \frac{20 \times 10^{-3} \text{ mm}}{100 \text{ mm}} = 0.02\%. \tag{11.3}$$

The uncertainties from the two contributions are added in quadrature to obtain the final uncertainty on the z-scale

$$\sigma_{z\text{-scale}} = 0.022\%. \tag{11.4}$$

A relative uncertainty on the decay time results in a relative uncertainty on Δm_s. Thus, a systematic uncertainty on Δm_s due to the z-scale uncertainty of $\pm 0.004\,\text{ps}^{-1}$ is assigned.

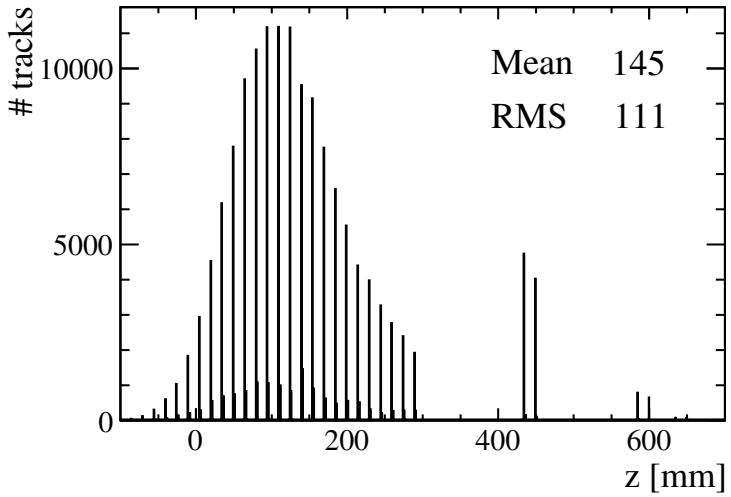


Figure 11.2.: *z-position of the first hit on each track used in this analysis.*

11.3. Momentum scale uncertainty

As defined in Equation 4.1 the decay time is proportional to the measured momentum, p, over the measured invariant mass, m, of the B_s^0 candidate. Therefore, the effect of the accuracy to which the momentum is measured at LHCb has to be evaluated. The uncertainty on the momentum scale has been found in an independent study, using mass measurements of well-known resonances. The measured deviations from the nominal values [23] are found to be within 0.15%.

However, since the momentum scale uncertainty influences both the measured momentum and the invariant mass of the B_s^0 candidates, the effect partially cancels. The uncertainty on the measured decay time can be estimated using

$$\frac{2m_{K^\pm}^2 + 2m_{\pi^\pm}^2}{m_{B_s^0}^2} \times 0.15\% = 0.003\%. \tag{11.5}$$

The full derivation of this formula is given in Appendix E. One simplification made in the derivation of this formula, $m_i/p_i \ll 1$, does not hold for all B_s^0 daughter tracks. Therefore, this systematic effect is underestimated.

To determine the systematic effect of the momentum scale uncertainty correctly, a study on simulated data was performed. In this study the three-momenta of the stable particles in the decay chain were scaled by a factor $\alpha = 1.0015$ and then the four-momenta were added to compute the B_s^0 momentum and invariant mass.

Figure 11.3 shows the distribution of the relative shift of the decay time computed using the scaled momenta compared to the true decay time of the B_s^0 candidates.

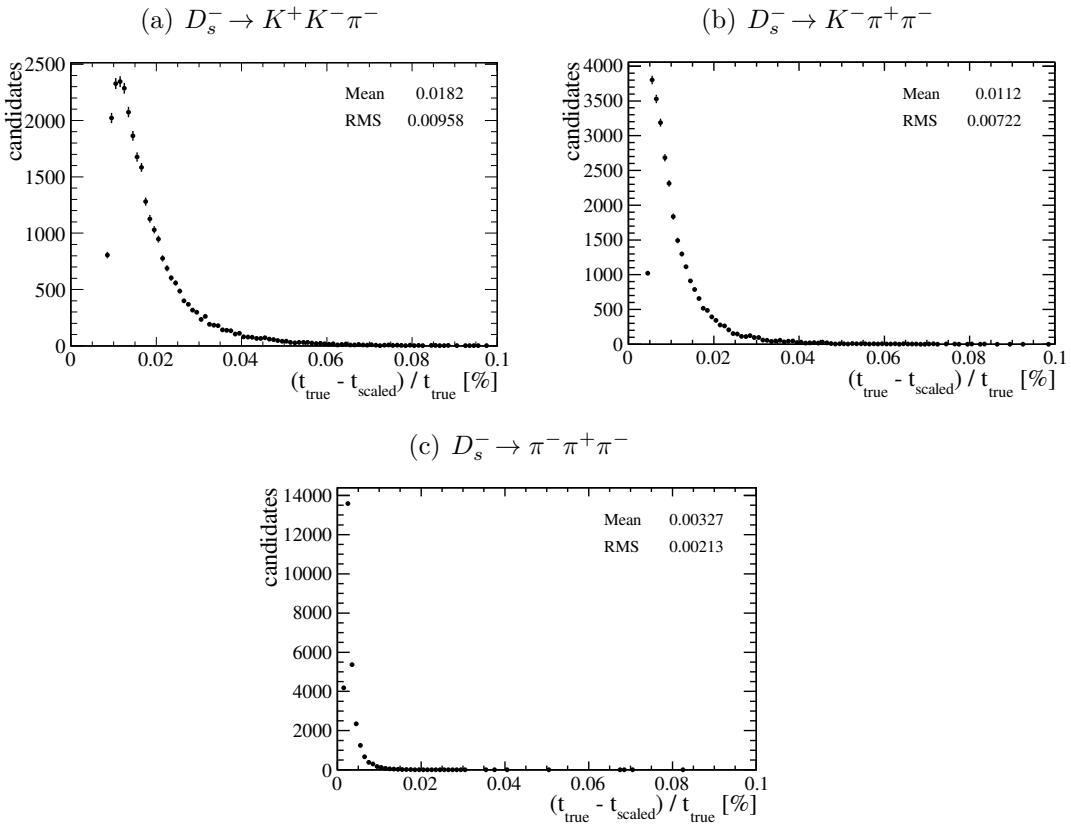


Figure 11.3.: *Results of the study on simulated data to assess the size of the effect of the uncertainty of the momentum scale on the decay time. Shown for $D_s^- \to K^+K^-\pi^-$ decays (Upper left plot), $D_s^- \to K^-\pi^+\pi^-$ decays (Upper right plot) and $D_s^- \to \pi^-\pi^+\pi^-$ decays (lower plot).*

It can be seen that, as expected, the effect depends on the mass of the stable particles. For the $D_s^- \to \pi^-\pi^+\pi^-$ decay mode the assumption $m_i/p_i \ll 1$ is given and the mean of the distribution is compatible with the result using Equation 11.5. However, the distributions for the other D_s^- decay modes are shifted to higher values. The shift is largest for $D_s^- \to K^+K^-\pi^-$ decays. Since these decays also constitutes the largest part of the data set, it was decided to conservatively use the observed shift of

$$\frac{\Delta t}{t} = 0.02\% \tag{11.6}$$

for the $D_s^- \to K^+K^-\pi^-$ mode as systematic effect on Δm_s. The relative uncertainty on the measured momentum results in a relative uncertainty on Δm_s. Thus, a systematic uncertainty on Δm_s due to the momentum-scale uncertainty of $\pm 0.004\,\mathrm{ps}^{-1}$ is assigned.

11.4. Summary of systematic studies

In this section the different contributions to the systematic uncertainty on the measurement of Δm_s have been presented. They are listed in Table 11.2 as well as the quadratic sum which is the total systematic uncertainty assigned to this analysis.

The dominating contributions to the systematic uncertainty on Δm_s originate from the z-scale and momentum uncertainties, which will be the long term limits for this analysis.

systematic	$\Delta_{\Delta m_s}[\,\mathrm{ps}^{-1}\,]$
Γ_s	–
$\Delta\Gamma_s$	–
decay time acceptance	–
decay time resolution	–
decay time resolution model	–
mass model	–
background yield	–
$B_s^0 \to D_s^{\pm} K^{\mp}$ yield	–
σ_t and predicted η PDFs	–
decay time bias	–
mass/decay time correlation	–
z-scale	0.004
momentum scale	0.004
total systematic	0.006

Table 11.2.: *Summary of the systematic uncertainties on Δm_s and their quadratic sum. The first rows summarize the cross checks described in Section 11.1 where no effect is observed and therefore no systematic is assigned. The dominant systematics are the z-scale and momentum scale.*

CHAPTER 12

Summary and Conclusion

In this thesis, the analysis to measure the B_s^0–$\overline{B}_s^0$ oscillation frequency Δm_s has been presented. A sample of 34,000 signal candidates of the decay $B_s^0 \to D_s^- \pi^+$ has been selected reconstructing five different D_s^- decay modes, with very little background contribution in the signal region ($S/B = 7.4$). In this dataset, which was taken by the LHCb experiment in 2011 corresponding to an integrated luminosity of $\mathcal{L} = 1\,\text{fb}^{-1}$, the distributions of the B_s^0 invariant mass, decay time and flavour tagging decision were fitted to obtain Δm_s.

For this analysis the two main challenges pose the determination of the production flavour of the B_s^0 meson and the decay time resolution. The production flavour is obtained using both opposite and same side kaon flavour tagging algorithms. The effective tagging efficiencies are $\varepsilon_{\text{eff}}^{\text{OST}} = 2.4 \pm 0.4\%$ and $\varepsilon_{\text{eff}}^{\text{SST}} = 1.2 \pm 0.3\%$ which represent the statistical reduction of the data sample due to imperfect flavour tagging. The decay time resolution is calibrated on data finding an average decay time uncertainty of 45 fs.

Various systematic effects have been studied and most have been found negligible compared to the statistical uncertainty. The dominant systematic uncertainties in this analysis are the uncertainties on the scales with which lengths and momenta can be measured at the LHCb experiment.

The final result for Δm_s is

$$\Delta m_s = 17.768 \pm 0.023(\text{stat}) \pm 0.006(\text{syst})\,\text{ps}^{-1} \tag{12.1}$$

which is the most precise measurement of this quantity to date. Its combined

uncertainty (statistical and systematic) is a factor three smaller than the current world average [23]. The result is well in agreement with the previous result by the CDF collaboration [17] and with the first published result by LHCb [18] which used a statistically independent sample. Over the individual rounds, the analysis was more and more refined and the systematic uncertainties have been reevaluated and reduced from $0.024\,\mathrm{ps}^{-1}$ to $0.006\,\mathrm{ps}^{-1}$.

The result is also well in agreement with the theoretical prediction which suffer from poor uncertainties due to insufficient lattice QCD calculations [33]. Therefore, it is not feasible to look for deviations from the theoretical predictions until these lattice QCD calculations are more accurate. However, The B_s^0–$\overline{B}_s^0$ mixing sector is still an excellent place to look for physics contributions beyond the Standard Model [72]. Other observables like the weak phase ϕ_s in the decays $B_s^0 \to J/\psi\,\phi$ and $B_s^0 \to J/\psi\,\pi^+\pi^-$ [1] have much better theoretical predictions. These analyses must resolve the fast B_s^0 oscillation, which is why the measurement of the oscillation frequency Δm_s is a prerequisite to prove that LHCb is able to perform precision measurements in the B_s^0 mixing sector. The excellent decay time resolution and flavour tagging performance provided by the LHCb experiment lead to this precise result of Δm_s and Figure 9.6 which is the first visualization of the B_s^0–$\overline{B}_s^0$ oscillation over so many periods ever shown.

The analysis presented in this thesis was also the base for the optimization and calibration of the same side kaon tagger in data, which will be the long-term application of this analysis. In recent studies this resulted in an improvement of the same side kaon tagger performance by about 56% [73].

This measurement of Δm_s is still statistically limited. Adding the dataset taken by LHCb in 2012 which corresponds to $2\,\mathrm{fb}^{-1}$ and using the optimized same side kaon tagger with the improved performance will improve the uncertainty on Δm_s significantly.

CHAPTER 13

Acknowledgements

I would like to thank first and foremost my two supervisors Stephanie Hansmann-Menzemer and Ulrich Uwer. They guided me through the whole process as a PhD student and gave me the opportunity to constantly improve my analysis and presentation skills by providing constructive critcism. The best examples are probably my presentations for which I always receive extensive critcism and the result is always a big improvement. Also, I would like to thank Hans-Christian Schultz-Coulon for agreeing to be the second referee for my thesis.

Next, I would like to express my gratitude to my colleagues which helped bring this document into shape. First to mention here is Paul Seyfert who read the whole document and provided me more than once a useful second view on the matter at hand. I also want to thank Blake Leverington who showed corners of the English language to me which I had no idea about. Additionally I received many useful comments from Francesca Dordei, Lucia Grillo, Georg Krocker and Sascha Stahl. This includes only the people who helped me to write this document, but of course during my time as a PhD student I had many fruitful discussions with all colleagues in the Heidelberg LHCb group.

Another person I would like to point out is my former colleague Christian Linn, with whom I developed the fitting algorithm, used for the measurement of Δm_s (Heidelberg Fitter). He was always able to provide the most useful ideas how to improve the algorithm. In this context, I also want to mention Christoph Langenbruch, who wrote the first version of the Heidelberg Fitter without which my analysis would not have been possible.

I also want to say thank you to my colleagues at CERN. The members of the B2OCTDSubWG (B to open Charm time-dependent sub-working group) of LHCb and especially Vladimir (Vava) Gligorov have provided me a lot of deeper insight into my analysis and have always been eager to help with any issues. Vava has always treated me with the utmost respect and loyalty and I want to thank him for that.

Finally, I want to express my gratitude to my family who has always given me the support which made my studies possible in the first place.

The largest support in my life has been my future wife Kathrin which has always been there for me. She has distracted me when I needed to clear my head and gave me freedom when I had to work long nights. But mostly she has given me the security that I can always rely on her. I love you, Kathrin.

Last and least, I would like to thank NESPRESSO for providing the caffeine without which nothing of this would have been possible.

References

[1] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP violation and the B^0_s meson decay width difference with $B^0_s \to J/\psi K^+K^-$ and $B^0_s \to J/\psi\pi^+\pi^-$ decays*, Phys. Rev. **D87** (2013) 112010, arXiv:1304.2600.

[2] LHCb collaboration, *Measurement of the time-dependent CP-violation parameters in $B^0_s \to D^\mp_s K^\pm$*, LHCb-CONF-2012-029.

[3] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961), no. 4 579 .

[4] A. Salam and J. Ward, *Electromagnetic and weak interactions*, Physics Letters **13** (1964) 168.

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

[6] A. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[7] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. D **51** (1995) 379.

[8] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: evidence, candidates and constraints*, Physics Reports **405** (2005), no. 5 – 6 279 .

[9] J. A. Frieman, M. S. Turner, and D. Huterer, *Dark energy and the accelerating universe*, Annual Review of Astronomy and Astrophysics **46** (2008), no. 1 385.

[10] Planck Collaboration, P. Ade *et al.*, *Planck 2013 results. I. Overview of products and scientific results*, arXiv:1303.5062.

[11] I. H. Brock, ed., *Physics at the Terascale*, Wiley-VCH, Weinheim, 1. aufl. ed., 2011.

[12] A. Einstein, *Die grundlage der allgemeinen relativitätstheorie*, Annalen der Physik **354** (1916), no. 7 769.

[13] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, *The Hierarchy problem and new dimensions at a millimeter*, Phys. Lett. **B429** (1998) 263, arXiv:hep-ph/9803315.

[14] M. Drees, *An Introduction to supersymmetry*, arXiv:hep-ph/9611409.

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

[16] D0 Collaboration, V. Abazov *et al.*, *First direct two-sided bound on the B_s^0 oscillation frequency*, Phys. Rev. Lett. **97** (2006) 021802, arXiv:hep-ex/0603029.

[17] CDF Collaboration, A. Abulencia *et al.*, *Measurement of the b 0 s âĂŞ $\overline{\mathrm{B}}_{\mathrm{s}}^{0}$ oscillation frequency*, Phys. Rev. Lett. **97** (2006) 062003.

[18] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the B 0 s âĂŞ $\overline{\mathrm{B}}_{\mathrm{s}}^{0}$ oscillation frequency Δm_s in $B_s^0 \to D_s^-(3)\pi$ decays*, Phys. Lett. **B709** (2012) 177, arXiv:1112.4311.

[19] 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.

[20] R. Aaij *et al.*, *Precision measurement of the B_s^0– $\overline{B}_s^0$ oscillation frequency with the decay $B_s^0 \to D_s^-\pi^+$*, New Journal of Physics **15** (2013), no. 5 053021.

[21] G. ALTARELLI, *The Standard model of particle physics*, arXiv:hep-ph/0510281.

[22] O. Nachtmann, *Phänomene und Konzepte der Elementarteilchenphysik: mit 40 Tabellen*, Friedrich Vieweg und Sohn, 1986.

[23] Particle Data Group, J. Beringer *et al.*, *Review of particle physics*, Phys. Rev. **D86** (2012) 010001.

[24] G. Aad *et al.*, *Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc*, Physics Letters B **716** (2012), no. 1 1 .

[25] S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc*, Physics Letters B **716** (2012), no. 1 30 .

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

[27] P. Harrison and H. Quinn, *The Babar Physics Book: Physics at an Asymmetric B Factory*, SLAC report, Stanford Linear Acceleration Center, 1998.

[28] K. Anikeev *et al.*, *B physics at the Tevatron: Run II and beyond*, `arXiv:hep-ph/0201071`.

[29] I. Dunietz, R. Fleischer, and U. Nierste, *In pursuit of new physics with B_s decays*, Phys. Rev. D **63** (2001) 114015.

[30] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[31] M. Beneke, G. Buchalla, A. Lenz, and U. Nierste, *CP asymmetry in flavour-specific B decays beyond leading logarithms*, Physics Letters B **576** (2003), no. 1 – 2 173 .

[32] M. Ciuchini *et al.*, *Lifetime differences and CP violation parameters of neutral B mesons at the next-to-leading order in QCD*, Journal of High Energy Physics **2003** (2003), no. 08 031.

[33] CKMfitter Group, A. Lenz *et al.*, *Anatomy of new physics in b-$\overline{B}$ mixing*, Phys. Rev. D **83** (2011) 036004.

[34] A. Lenz and U. Nierste, *Numerical Updates of Lifetimes and Mixing Parameters of B Mesons*, `arXiv:1102.4274`.

[35] A. Bazavov *et al.*, *Neutral B-meson mixing from three-flavor lattice QCD: Determination of the SU(3)-breaking ratio ξ*, Phys. Rev. **D86** (2012) 034503, `arXiv:1205.7013`.

[36] J. Laiho, E. Lunghi, and R. S. Van de Water, *Lattice QCD inputs to the CKM unitarity triangle analysis*, Phys. Rev. **D81** (2010) 034503, `arXiv:0910.2928`.

[37] A. Team, *The four main LHC experiments*, CERN document server, https://cds.cern.ch/record/40525, Jun, 1999.

[38] J. Nardulli, *Reconstruction of two-body B decays in LHCb*, PhD thesis, Vrije U. Amsterdam, Amsterdam, 2007, Presented on 04 Oct 2007.

[39] `http://lhcb.web.cern.ch/lhcb/speakersbureau/excel/default.html`.

[40] LHCb Collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s}$=7 TeV in the forward region*, Phys. Lett. B **694** (2010) 209, Comments: 11 pages, 5 figures, submitted to Physics Lett. B.

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

[42] R. Lindner, *LHCb layout_2. LHCb schema_2*, LHCb Collection., Feb, 2008.

[43] W. Hulsbergen, *The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **600** (2009), no. 2 471 .

[44] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, J. Instrum. **8** (2012) P04022. 31 p.

[45] LHCb Collaboration, *The Brunel project*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/`.

[46] LHCb Collaboration, *The DaVinci project*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/`.

[47] M. Clemencic *et al.*, *Recent developments in the LHCb software framework gaudi*, Journal of Physics: Conference Series **219** (2010), no. 4 042006, `http://stacks.iop.org/1742-6596/219/i=4/a=042006`.

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

[49] LHCb Collaboration, *The Gauss project*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/`.

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

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

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

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

[54] GEANT4 collaboration, S. Agostinelli *et al.*, *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[55] GEANT4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[56] LHCb Collaboration, *The Boole project*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/`.

[57] LHCb Collaboration, *The Moore project*, `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/`.

[58] G. Punzi, *Comments on likelihood fits with variable resolution*, eConf **C030908** (2003) WELT002, `arXiv:physics/0401045`.

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

[60] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003.

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

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

[63] S. Wandernoth, *Measurement of the B_s^0–$\overline{B}_s^0$ Mixing Frequence and Calibration of the Same Side Tagger at LHCb*, Master's thesis, Heidelberg University., June, 2009.

[64] R. Aaij *et al.*, *Opposite-side flavour tagging of B mesons at the LHCb experiment*, Eur. Phys. J. C **72** (2012) 2022. 28 p, Comments: 22 pages, 9 figures.

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

[66] LHCb collaboration, *Optimization and calibration of the same-side kaon tagging algorithm using hadronic B_s^0 decays in 2011 data*, LHCb-CONF-2012-033.

[67] `http://root.cern.ch/root/html/RooBukinPdf.html`.

[68] 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.

[69] LHCb VELO group, *Performance of the LHCb VELO*, To be submitted to JINST.

[70] S. Viret, C. Parkes, and M. Gersabeck, *Alignment procedure of the LHCb vertex detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **596** (2008), no. 2 157 .

[71] S. Borghi *et al.*, *First spatial alignment of the LHCb VELO and analysis of beam absorber collision data*, Nucl. Instrum. Methods Phys. Res. , A **618** (2010), no. LHCb-DP-2010-001. CERN-LHCb-DP-2010-001 108.

[72] W. Hulsbergen, *Constraining new physics in B_s^0 meson mixing*, Mod. Phys. Lett. **A28** (2013) 1330023, arXiv:1306.6474.

[73] G. A. Krocker, *Development and calibration of a same side kaon tagging algorithm and measurement of the B_s^0–$\bar{B}_s^0$ oscillation frequency Δm_s at the LHCb experiment*, PhD thesis, Heidelberg U, Sep, 2013, Presented 20 Nov 2013.

APPENDIX A

Additional Information on the calibration of the decay time resolution

In this appendix the fit results and resulting distribution of weights are given. The fitted parameters from the D^- mass fit in Figure 7.5 are sumarized in Table A.1 and the distribution of the resulting weights is shown in Figure A.1. It can be seen that the weights are negative for events outside the invariant mass signal region to subtract the background.

The Bukin PDF used in the fit of the D^- $log(IP)$ distribution is defined as

$$\mathcal{P}(x; x_p, \sigma_p, \xi, \rho_i) = A_p e^{\left(\frac{\xi\sqrt{\xi^2+1}(x-x_1)\sqrt{2\ln 2}}{\sigma_p\left(\sqrt{\xi^2+1}-\xi\right)^2 \ln\left(\sqrt{\xi^2+1}+\xi\right)} + \rho_i \left(\frac{x-x_i}{x_p-x_i}\right)^2 - \ln 2 \right)}, \tag{A.1}$$

where $\rho_i = \rho_1$ and $x_i = x_1$ for $x < x_1$, and $\rho_i = \rho_2$ and $x_i = x_2$ for $x \geq x_2$ with

$$x_{1,2} = x_p + \sigma_p \sqrt{2\ln 2}\left(\frac{\xi}{\sqrt{\xi+1}} \mp 1\right) \tag{A.2}$$

while in the central region ($x_1 < x < x_2$) it is parametrized as

$$\mathcal{P}(x; x_p, \sigma_p, \xi, \rho) = e^{-\ln 2 \left(\frac{\ln(1+4\xi\sqrt{\xi^2+1}\frac{x-x_p}{\sigma 2\sqrt{2\ln 2}})}{\ln(1+2\xi(\xi-\sqrt{\xi^2+1}))} \right)}. \tag{A.3}$$

The parameters x_p and σ_p describe the peak position and the FWHM. The parameter ξ describes the asymmetry of the distribution is and the parameters ρ_1 and ρ_2 decribe the left and right tail, respectively.

Parameter	**Fit results**
$\mu_{m_{D_s^-}}$	1870.953±0.0055 MeV/c^2
$\sigma_{1,m_{D_s^-}}$	12.2±1.2 MeV/c^2
$\sigma_{1,m_{D_s^-}}$	6.08±0.25 MeV/c^2
$f_{m_{D_s^-}}$	0.386±0.072
$\alpha_{m_{D_s^-}}$	0.0495±0.0013
N_{sig}	51898±6600
N_{bkg}	656000±83000

Table A.1.: *Fit results for the invariant mass distribution of the $D^- \to K^+\pi^-\pi^-$ candidates used for the calibration of the decay time resolution.*

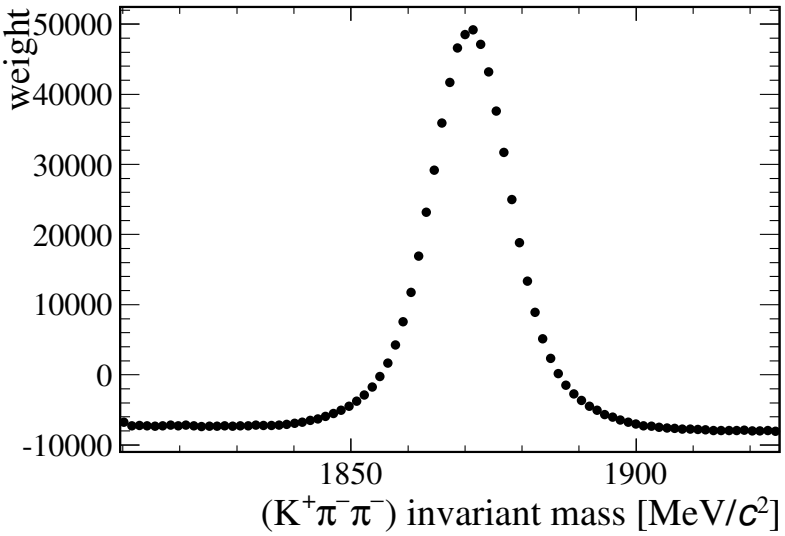


Figure A.1.: *Distribution of the weights obtained from the fit to the $D^- \to K^+\pi^-\pi^-$ invariant mass. These weights are applied to the sample to subtract the background for the next steps of the calibration.*

Parameter	**Prompt D^-**	**Non-prompt D^-**
x_p	-3.847±0.015	-2.330±0.032
σ_p	0.5993±0.0079	0.963±0.016
ξ	-0.145±0.016	-0.006±0.023
ρ_1	-0.229±0.027	-0.177±0.057
ρ_2	-0.359±0.026	-1.22±0.46

Table A.2.: *Fit results for the $log(IP)$ distribution of the prompt and non-prompt $D^- \to K^+\pi^-\pi^-$ candidates from simulated data.*

Parameter	**Prompt D^-**	**Non-prompt D^-**
x_p	-3.767±0.0043	-2.503±0.016
σ_p	0.549±0.0044	1.25±0.07
ξ	-0.145	-0.006
ρ_1	-0.229	-0.177
ρ_2	-0.359	-1.22

Table A.3.: *Fit results for the $log(IP)$ distribution of the prompt and non-prompt $D^- \to K^+\pi^-\pi^-$ candidates from data. The shape parameter ξ, ρ_1 and ρ_2 have been fixed from the fit to simulated data.*

Table A.2 summarizes the results for the fitted Bukin PDFs to the $log(IP)$ distributions in Figure 7.6 of prompt and non-prompt D^- candidates in simulated data. The fit results for the fit of the $log(IP)$ distribution in data, shown in Figure 7.7, are summarized in Table A.3.

The distribution of the weights obtained from this fit is shown in Figure A.2. It can be seen that the weights have negative values for the region dominated by the non-prompt component and positive values for low $log(IP)$.

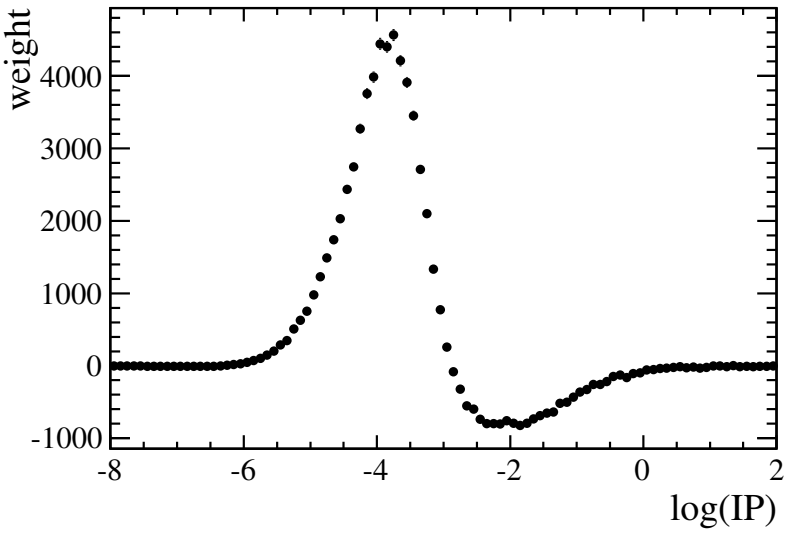


Figure A.2.: *Distribution of the weights obtained from the fit to the D^- $log(IP)$ distribution. These weights are applied to the sample to subtract the contribution of non-prompt D^- for the final step of the calibration.*

APPENDIX B

Normalization of the Signal Decay Time PDF

The signal decay time PDF is definend in Equation 9.5 as[1]

$$\mathcal{P}_{D_s\pi;t,q} \propto \Bigg[\left(e^{-\Gamma_s t} \cdot \left[\cosh\left(\frac{\Delta\Gamma_s}{2} t \right) - q \cdot (1-2\omega) \cdot \cos\left(\Delta m_s t\right) \right] \right) \otimes \mathcal{G}(0, S_{\sigma_t}\sigma_t; t) \Bigg] \cdot \mathcal{E}_t(t) \cdot \mathcal{N}_{\text{tag}}^{\text{OST}} \cdot \mathcal{N}_{\text{tag}}^{\text{SST}}.$$

The PDF has to be normalized so that the integral is equal to one. Therefore, the normalization factor $\mathcal{N}$ is the integral

$$\mathcal{N} = \int \mathcal{P}_{D_s\pi;t,q} dt \; dq \; . \tag{B.1}$$

The integral over the discrete flavour tagging decision is performed explicitly by summing over the possible values ($q \in \{+1; -1; 0\}$)

$$\int \mathcal{P}_{D_s\pi;t;q} dt \; dq = \int \sum_{q}^{+1;-1;0} \mathcal{P}_{D_s\pi;t,q} dt \; . \tag{B.2}$$

Due to this sum the cosine terms in the mixing PDF completely cancel and the relative normalization factors $\mathcal{N}_{\text{tag}}^{\text{OST}}$ and $\mathcal{N}_{\text{tag}}^{\text{SST}}$ add up to one. This leaves the

[1] Only positive true decay times are physical. Therefore the exponential $e^{\Gamma_s t}$ implies the Heaviside Function $\Theta(t)$

normalization factor $\mathcal{N}$ to be:

$$\begin{aligned}\mathcal{N} &= \int \left[e^{-\Gamma_s t} \cdot \cosh\left(\frac{\Delta\Gamma_s}{2} t\right)\right] \otimes \mathcal{G}(0, S_{\sigma_t}\sigma_t; t) \cdot \mathcal{E}_t(t) dt \\ &= \int 0.5 \left[e^{-\left(\Gamma_s + \frac{\Delta\Gamma_s}{2}\right)t} + e^{-\left(\Gamma_s - \frac{\Delta\Gamma_s}{2}\right)t}\right] \otimes \mathcal{G}(0, S_{\sigma_t}\sigma_t; t) \cdot \mathcal{E}_t(t) dt \ .\end{aligned} \tag{B.3}$$

Using the linearity of the integral and the convolution this can be reduced to two terms of the form

$$\int e^{-ct} \otimes \mathcal{G}(0, \sigma_t; t) \cdot \mathcal{E}_t(t) dt \tag{B.4}$$

Involving the simple convolution of an exponential function with slope c starting at $t = 0$ with a Gaussian with a mean equal to zero and a width equal to σ_t which is given by

$$e^{-ct} \otimes \mathcal{G}(0, \sigma_t; t) = \frac{c}{2} e^{-c\left(t - \frac{c\sigma_t^2}{2}\right)} \cdot \mathrm{Erfc}\left(\frac{c\sigma_t^2 - t}{\sqrt{2}\sigma_t}\right) \tag{B.5}$$

with the complementary error function $\mathrm{Erfc}(z) \equiv \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-u^2} du$. So Equation B.3 becomes

$$\mathcal{N} = \int_{-\infty}^{\infty} \sum_{i=1}^{2} \frac{c_i}{2} e^{-c_i\left(t - \frac{c_i\sigma_t^2}{2}\right)} \cdot \mathrm{Erfc}\left(\frac{c_i\sigma_t^2 - t}{\sqrt{2}\sigma_t}\right) \cdot \mathcal{E}_t(t) \tag{B.6}$$

with the coefficients $c_i = \Gamma_s \pm \Delta\Gamma_s/2$. For simplicity, from now on only one coefficient c will be considered. With the decay time acceptance function defined in Equation 7.10 as

$$\mathcal{E}_t(t) = (1 - e^{-\frac{(t - s_{\mathrm{acc}})}{\alpha_{\mathrm{acc}}}})(1 + \beta_{\mathrm{acc}} t) \cdot \Theta(t - \zeta_{\mathrm{acc}}),$$

the normalization of the signal decay time PDF is given by

$$\begin{aligned}\mathcal{N} = \frac{c}{2b} \Bigg[& \left(a_0 + \frac{a_1}{bc}\left(1 - \frac{bc\sigma_t^2}{\alpha_{\mathrm{acc}}}\right)\right) \\ & \cdot \left(2 - \mathrm{Erfc}\left(-\frac{1}{\sqrt{2}} \cdot \left(\frac{\zeta_{\mathrm{acc}}}{\sigma_t} + \frac{\sigma_t}{\alpha_{\mathrm{acc}}}\right)\right)\right) \cdot e^{\frac{\sigma_t^2}{2\alpha_{\mathrm{acc}}^2}} \\ & + \left(a_0 + \frac{a_1}{bc}(1 + bc\zeta_{\mathrm{acc}}\right) \\ & \cdot \mathrm{Erfc}\left(-\frac{1}{\sqrt{2}} \cdot \left(\frac{\zeta_{\mathrm{acc}}}{\sigma_t} - c\sigma_t\right)\right) \cdot e^{\left(-bc\zeta_{\mathrm{acc}} + \frac{\sigma_t^2 c^2}{2}\right)} \\ & + a_1\sigma_t\sqrt{\frac{2}{\pi}} \cdot e^{-c\zeta_{\mathrm{acc}} \cdot \left(\frac{1}{c\alpha_{\mathrm{acc}}} + \frac{\zeta_{\mathrm{acc}}}{2c\sigma_t^2}\right)} \Bigg]\end{aligned} \tag{B.7}$$

with the coefficients

$$\begin{aligned} a_0 &= e^{s_{\mathrm{acc}}/\alpha_{\mathrm{acc}}} \\ a_1 &= e^{s_{\mathrm{acc}}/\alpha_{\mathrm{acc}}} \cdot \beta_{\mathrm{acc}} \\ b &= 1 + \frac{1}{c\alpha_{\mathrm{acc}}} \end{aligned} \tag{B.8}$$

APPENDIX C

Results from validation of fit procedure

In this appendix the results from the pseudo-experiments described in Chapter 10 are given. Table C.1 shows the means and widths of the pull distributions as defined in Equation 10.1 for all free parameters in the fit. It can be seen that the means are all compatible with zero and the widths are compatible with one. This shows that the PDFs are correctly implemented in the fit algorithm.

parameter	pull mean	pull width
$m_{B_s^0}$	0.017±0.013	0.998±0.010
σ_m^1	−0.004±0.013	0.994±0.010
σ_m^2	−0.027±0.013	1.006±0.010
$f_{\rm sig}^{\phi\pi}$	−0.020±0.013	1.035±0.010
$f_{\rm comb}^{\phi\pi}$	−0.045±0.014	1.028±0.010
$f_{\Lambda_b^0}^{\phi\pi}$	−0.014±0.012	0.985±0.010
$a_{\rm comb}^{\rm DT}$	0.034±0.013	1.012±0.010
$\beta_{\rm comb}^{\rm DT}$	−0.002±0.013	1.000±0.010
$\alpha_{\rm acc}$	0.004±0.013	1.003±0.010
$\alpha_{\rm comb}^{\phi\pi}$	0.009±0.013	1.000±0.010
Δm_s	−0.017±0.013	1.023±0.010
$\varepsilon_{\rm sig}^{\rm OST}$	−0.014±0.013	0.992±0.010
$\varepsilon_{\rm sig}^{\rm SST}$	−0.008±0.013	0.999±0.010
$\omega_{\rm comb}^{\rm OST}$	−0.006±0.013	0.993±0.010

$\omega_{\text{comb}}^{\text{SST}}$	0.037 ± 0.013	1.012 ± 0.010
$\varepsilon_{\text{tag;comb}}^{\text{OST}}$	0.011 ± 0.013	1.009 ± 0.010
$\varepsilon_{\text{tag;comb}}^{\text{SST}}$	0.000 ± 0.013	1.000 ± 0.010
$\omega_{\Lambda_b^0}^{\text{SST}}$	-0.012 ± 0.013	0.991 ± 0.010
$\varepsilon_{\Lambda_b^0}^{\text{SST}}$	-0.038 ± 0.013	1.015 ± 0.010
$\omega_{B^0}^{\text{SST}}$	0.023 ± 0.013	0.980 ± 0.010
$\varepsilon_{B^0}^{\text{SST}}$	-0.014 ± 0.013	1.005 ± 0.010
p_1^{SST}	0.010 ± 0.013	0.974 ± 0.010
p_1^{OST}	0.015 ± 0.013	1.001 ± 0.010
p_0^{OST}	-0.011 ± 0.013	0.982 ± 0.010
p_0^{SST}	-0.024 ± 0.013	1.002 ± 0.010
$f_{\text{sig}}^{K^{*0}K}$	-0.021 ± 0.013	1.019 ± 0.010
$f_{\text{comb}}^{K^{*0}K}$	-0.008 ± 0.013	1.036 ± 0.010
$f_{\Lambda_b^0}^{K^{*0}K}$	-0.022 ± 0.012	0.983 ± 0.010
$\alpha_{\text{comb}}^{K^{*0}K}$	0.028 ± 0.013	0.995 ± 0.010
$f_{\text{sig}}^{KK\pi}$	-0.020 ± 0.013	0.990 ± 0.010
$f_{\text{comb}}^{KK\pi}$	-0.019 ± 0.013	1.010 ± 0.010
$f_{\Lambda_b^0}^{KK\pi}$	-0.020 ± 0.013	0.998 ± 0.010
$\alpha_{\text{comb}}^{KK\pi}$	0.004 ± 0.013	1.003 ± 0.010
$f_{\text{sig}}^{K\pi\pi}$	-0.003 ± 0.013	1.009 ± 0.010
$f_{\text{comb}}^{K\pi\pi}$	-0.000 ± 0.013	1.015 ± 0.010
$f_{\Lambda_b^0}^{K\pi\pi}$	0.037 ± 0.013	0.978 ± 0.010
$\alpha_{\text{comb}}^{K\pi\pi}$	0.003 ± 0.013	1.020 ± 0.010
$f_{\text{sig}}^{\pi\pi\pi}$	-0.017 ± 0.013	1.003 ± 0.010
$a_{\text{comb};3\pi}^{\text{DT}}$	0.025 ± 0.013	1.026 ± 0.010
$\beta_{\text{comb};3\pi}^{\text{DT}}$	-0.017 ± 0.013	1.015 ± 0.010
$\alpha_{\text{comb}}^{\pi\pi\pi}$	0.030 ± 0.013	1.000 ± 0.010

Table C.1.: *Result of 6000 pseudo-experiments to validate the fit procedure. Mean and width of the pull distributions for all free parameters are given.*

APPENDIX D

Correlation matrix of the nominal fit

Tables D.1 and D.2 show the full correlation matrix from the nominal fit to obtain Δm_s using both opposite side and same side kaon taggers. It shows how much the individual parameters are intertwined in the fit. The correlation coefficients ρ_{ij} are defined as

$$\rho_{ij} = \mathrm{Cov}(i,j) \cdot \sigma_i \cdot \sigma_j, \tag{D.1}$$

with the covariance $\mathrm{Cov}(i,j)$ and the variances σ of the individual parameters i and j. The coefficients ρ_{ij} range from -1 to 1.

The row and columns of Δm_s are highlighted. It can be seen that there are barely any correlations present between Δm_s and all other free parameters in the fit.

	$m_{B_s^0}$	σ_m^1	σ_m^2	$f_{\text{sig}}^{\phi\pi}$	$f_{\text{comb}}^{\phi\pi}$	$f_{\Lambda_b^0}^{\phi\pi}$	$a_{\text{comb}}^{\text{DT}}$	$\beta_{\text{comb}}^{\text{DT}}$	α_{acc}	$\alpha_{\text{comb}}^{\phi\pi}$	Δm_s	$\varepsilon_{\text{sig}}^{\text{OST}}$	$\varepsilon_{\text{sig}}^{\text{SST}}$	$\omega_{\text{comb}}^{\text{OST}}$	$\omega_{\text{comb}}^{\text{SST}}$	$\varepsilon_{\text{tag;comb}}^{\text{OST}}$	$\varepsilon_{\text{tag;comb}}^{\text{SST}}$	$\omega_{\Lambda_b^0}^{\text{SST}}$	$\varepsilon_{\Lambda_b^0}^{\text{SST}}$	$\omega_{B^0}^{\text{SST}}$	$\varepsilon_{B^0}^{\text{SST}}$
$m_{B_s^0}$	1.00	-0.27	-0.19	-0.33	-0.32	-0.36	-	-	-	-	-	-	0.05	0.01	-	-	-	-	-	0.06	-0.03
σ_m^1		1.00	-0.15	0.22	0.20	0.27	-	-	-0.01	-0.01	-	-	-0.03	-	-	-	-	-	-0.01	-0.04	0.03
σ_m^2			1.00	0.45	0.42	0.05	0.01	-0.01	-0.04	-0.02	-0.01	-	-0.04	-0.03	-0.01	0.02	-0.01	0.01	0.06	-0.04	0.02
$f_{\text{sig}}^{\phi\pi}$				1.00	0.79	0.62	0.01	-0.01	-0.03	-0.04	-0.01	-	-0.05	-0.02	-0.01	0.01	-	-0.01	0.03	-0.01	-
$f_{\text{comb}}^{\phi\pi}$					1.00	0.55	-	0.02	-	0.32	-0.01	-0.01	-0.05	-0.03	-0.02	-0.01	-0.01	-	0.02	-0.01	-0.01
$f_{\Lambda_b^0}^{\phi\pi}$						1.00	-	-	-	-0.03	-	-	-0.04	-	-	0.01	-	-0.04	0.01	0.02	-0.02
$a_{\text{comb}}^{\text{DT}}$							1.00	-0.88	0.02	-	-	-0.01	-	-	-	-	-	-0.01	-	-0.01	-0.01
$\beta_{\text{comb}}^{\text{DT}}$								1.00	-0.07	0.03	-	-	-	-	-	-	-0.01	0.01	0.01	-	0.01
α_{acc}									1.00	0.05	-	-0.01	-	-	-	-0.01	0.01	-	-0.02	-0.01	-0.01
$\alpha_{\text{comb}}^{\phi\pi}$										1.00	-	-0.03	-	-0.01	-0.02	-0.01	-0.01	0.02	-	-	-0.01
Δm_s											1.00	-	-	-	-	-	-	-	-	-	-
$\varepsilon_{\text{sig}}^{\text{OST}}$												1.00	-	0.01	-	-0.16	-	-	-	-	0.01
$\varepsilon_{\text{sig}}^{\text{SST}}$													1.00	-	-	-	-0.05	-	-0.07	-0.13	-0.36
$\omega_{\text{comb}}^{\text{OST}}$														1.00	0.01	-	-0.01	-	0.01	-	-
$\omega_{\text{comb}}^{\text{SST}}$															1.00	0.01	-0.01	-0.28	0.01	-0.07	-
$\varepsilon_{\text{tag;comb}}^{\text{OST}}$																1.00	0.01	-0.01	-	-	-
$\varepsilon_{\text{tag;comb}}^{\text{SST}}$																	1.00	-	-0.28	-0.02	-0.05
$\omega_{\Lambda_b^0}^{\text{SST}}$																		1.00	-	-0.24	0.01
$\varepsilon_{\Lambda_b^0}^{\text{SST}}$																			1.00	-0.08	-0.21
$\omega_{B^0}^{\text{SST}}$																				1.00	0.36
$\varepsilon_{B^0}^{\text{SST}}$																					1.00
p_1^{OST}																					
p_1^{SST}																					
p_0^{OST}																					
p_0^{SST}																					
$f_{\text{sig}}^{K^{*0}K}$																					
$f_{\text{comb}}^{K^{*0}K}$																					
$f_{\Lambda_b^0}^{K^{*0}K}$																					
$\alpha_{\text{comb}}^{K^{*0}K}$																					
$f_{\text{sig}}^{KK\pi}$																					
$f_{\text{comb}}^{KK\pi}$																					
$f_{\Lambda_b^0}^{KK\pi}$																					
$\alpha_{\text{comb}}^{KK\pi}$																					
$f_{\text{sig}}^{K\pi\pi}$																					
$f_{\text{comb}}^{K\pi\pi}$																					
$f_{\Lambda_b^0}^{K\pi\pi}$																					
$\alpha_{\text{comb}}^{K\pi\pi}$																					
$f_{\text{sig}}^{\pi\pi\pi}$																					
$a_{\text{comb;}3\pi}^{\text{DT}}$																					
$\beta_{\text{comb;}3\pi}^{\text{DT}}$																					
$\alpha_{\text{comb}}^{\pi\pi\pi}$																					

Table D.1.: *Correlation matrix for the nominal fit using both SST and OST.*

	p_1^{OST}	p_1^{SST}	p_0^{OST}	p_0^{SST}	$f_{\mathrm{sig}}^{K^{*0}K}$	$f_{\mathrm{comb}}^{K^{*0}K}$	$f_{\Lambda_b^0}^{K^{*0}K}$	$\alpha_{\mathrm{comb}}^{K^{*0}K}$	$f_{\mathrm{sig}}^{KK\pi}$	$f_{\mathrm{comb}}^{KK\pi}$	$f_{\Lambda_b^0}^{KK\pi}$	$\alpha_{\mathrm{comb}}^{KK\pi}$	$f_{\mathrm{sig}}^{K\pi\pi}$	$f_{\mathrm{comb}}^{K\pi\pi}$	$f_{\Lambda_b^0}^{K\pi\pi}$	$\alpha_{\mathrm{comb}}^{K\pi\pi}$	$f_{\mathrm{sig}}^{\pi\pi\pi}$	$a_{\mathrm{comb};3\pi}^{\mathrm{DT}}$	$\beta_{\mathrm{comb};3\pi}^{\mathrm{DT}}$	$\alpha_{\mathrm{comb}}^{\pi\pi\pi}$
$m_{B_s^0}$	0.01	0.01	-0.01	-0.02	-0.33	-0.31	-0.25	-0.01	-0.24	-0.17	-0.26	-0.01	-0.13	-0.09	-0.13	-	-0.06	-	0.01	0.06
σ_m^1	-0.01	-0.01	-	0.01	0.20	0.18	0.21	-	0.19	0.12	0.18	-0.01	0.10	0.07	0.09	-0.01	0.05	-0.01	-	-0.05
σ_m^2	-0.01	-0.01	0.01	0.02	0.41	0.38	-0.16	-	0.31	0.22	0.01	-	0.15	0.10	-	-0.01	0.21	0.03	-0.07	-0.16
$f_{\mathrm{sig}}^{\phi\pi}$	-0.01	-0.02	0.01	0.03	0.29	0.27	0.02	-	0.23	0.16	0.09	-	0.11	0.08	0.05	-	0.12	0.01	-0.04	-0.10
$f_{\mathrm{comb}}^{\phi\pi}$	-0.01	-0.01	0.01	0.03	0.27	0.25	0.01	-	0.21	0.16	0.09	0.01	0.11	0.08	0.04	-	0.11	0.01	-0.03	-0.09
$f_{\Lambda_b^0}^{\phi\pi}$	-0.01	-0.01	0.01	0.02	0.14	0.13	0.11	-	0.11	0.08	0.11	-	0.06	0.04	0.06	-	0.03	-	-0.01	-0.03
$a_{\mathrm{comb}}^{\mathrm{DT}}$	-	-	-	-	-	0.02	-0.02	0.03	0.01	0.06	-0.01	0.06	-	0.02	-0.01	0.02	-	-	-	-
$\beta_{\mathrm{comb}}^{\mathrm{DT}}$	-	-	-	-	-	-	0.01	-	-0.01	-	-0.01	-	-	0.01	-	0.01	-	-	-	-
α_{acc}	-	-	-0.01	-	-0.03	-	-	0.04	-0.03	0.05	-	0.07	-0.02	0.04	-0.01	0.06	-0.06	0.01	-0.03	0.05
$\alpha_{\mathrm{comb}}^{\phi\pi}$	-	-	-0.01	-	-0.01	-	-	0.01	-0.01	0.01	-	0.01	-	0.01	-	0.01	-0.01	-	-	0.01
Δm_s	0.04	-	-0.01	0.02	-	-	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-
$\varepsilon_{\mathrm{sig}}^{\mathrm{OST}}$	-	-	0.01	-	-	-0.01	0.01	-0.02	0.01	-0.03	-	-0.04	-	-0.02	-	-0.02	0.01	-	-	-0.01
$\varepsilon_{\mathrm{sig}}^{\mathrm{SST}}$	-	-0.01	-	0.02	-0.11	-0.11	-0.07	-0.02	-0.03	-0.01	-0.01	0.01	-	0.01	0.01	0.01	-0.01	-	-	-
$\omega_{\mathrm{comb}}^{\mathrm{OST}}$	-0.02	-	-0.08	-	-0.02	-0.02	0.01	-0.01	-0.02	-0.03	-	-0.02	-0.01	-0.03	-	-0.02	-0.01	-	-	-
$\omega_{\mathrm{comb}}^{\mathrm{SST}}$	-	-	-	-	-	-	0.01	-	-	0.01	-0.01	0.01	-0.01	-0.02	0.01	-0.01	0.01	-	-	-0.01
$\varepsilon_{\mathrm{tag;comb}}^{\mathrm{OST}}$	-	-	-0.01	-	0.01	-0.01	0.01	-0.02	-	-0.04	0.01	-0.03	-	-0.03	0.02	-0.03	0.02	-	-	-0.02
$\varepsilon_{\mathrm{tag;comb}}^{\mathrm{SST}}$	-	-	0.01	-0.01	-0.01	-0.04	0.03	-0.03	-	-	-	-	-	-	-	-	-	-	-	-
$\omega_{\Lambda_b^0}^{\mathrm{SST}}$	-0.01	-	-	0.01	0.02	0.03	-	0.01	-	-0.04	0.03	-0.04	-	0.03	-0.04	0.03	-	-	-	-
$\varepsilon_{\Lambda_b^0}^{\mathrm{SST}}$	-	-	-0.02	-	0.01	0.04	-0.11	0.05	0.02	-0.02	0.01	-0.03	0.01	-0.01	0.03	-0.01	0.01	-	-	-0.01
$\omega_{B^0}^{\mathrm{SST}}$	0.01	-	0.02	-0.02	-0.06	-0.07	-0.03	-0.02	-0.13	-0.12	-0.16	-0.04	-0.06	-0.08	-0.06	-0.04	-0.01	-	-	0.01
$\varepsilon_{B^0}^{\mathrm{SST}}$	-0.01	0.02	0.01	-0.04	0.15	0.14	0.14	0.01	-0.02	-0.04	-0.05	-0.03	-0.03	-0.05	-0.07	-0.04	-	-	-	-
p_1^{OST}	1.00	-	0.31	-	-0.03	-0.03	-0.02	-	-	-	-0.01	0.01	-	0.01	-	0.01	-0.01	-	-	-
p_1^{SST}		1.00	-	0.09	-0.01	-0.01	-0.01	-	-0.01	-0.01	-0.01	-	-	-	-	-	-	-	-	-
p_0^{OST}			1.00	0.02	0.01	0.01	-	-	-	-0.02	0.01	-0.02	0.01	-0.01	-	-0.01	0.01	-	-	-0.01
p_0^{SST}				1.00	0.02	0.02	-	-	0.02	0.01	0.01	-	0.02	0.01	0.01	-	0.01	-	-	-0.01
$f_{\mathrm{sig}}^{K^{*0}K}$					1.00	0.79	0.38	-0.02	0.21	0.15	0.10	-	0.10	0.07	0.04	-0.01	0.10	0.01	-0.03	-0.09
$f_{\mathrm{comb}}^{K^{*0}K}$						1.00	0.18	0.36	0.20	0.14	0.09	0.01	0.09	0.07	0.03	-	0.09	0.01	-0.03	-0.08
$f_{\Lambda_b^0}^{K^{*0}K}$							1.00	-0.20	0.02	0.01	0.09	-0.01	0.01	-	0.03	-	-0.02	-0.01	0.01	0.01
$\alpha_{\mathrm{comb}}^{K^{*0}K}$								1.00	-	0.01	-	0.01	-	0.01	-	0.01	-	-	-	-
$f_{\mathrm{sig}}^{KK\pi}$									1.00	0.49	0.46	-0.03	0.09	0.06	0.04	-	0.08	0.01	-0.03	-0.07
$f_{\mathrm{comb}}^{KK\pi}$										1.00	0.20	0.56	0.06	0.06	0.03	0.02	0.05	0.01	-0.02	-0.04
$f_{\Lambda_b^0}^{KK\pi}$											1.00	-0.06	0.05	0.04	0.05	-	0.02	-	-	-0.02
$\alpha_{\mathrm{comb}}^{KK\pi}$												1.00	-	0.02	-	0.02	-0.01	-	-	-
$f_{\mathrm{sig}}^{K\pi\pi}$													1.00	0.43	0.37	-0.06	0.04	-	-0.01	-0.03
$f_{\mathrm{comb}}^{K\pi\pi}$														1.00	0.09	0.53	0.02	-	-0.01	-0.02
$f_{\Lambda_b^0}^{K\pi\pi}$															1.00	-0.11	0.01	-	-	-0.01
$\alpha_{\mathrm{comb}}^{K\pi\pi}$																1.00	-0.01	-	-	0.01
$f_{\mathrm{sig}}^{\pi\pi\pi}$																	1.00	-	-0.07	-0.41
$a_{\mathrm{comb};3\pi}^{\mathrm{DT}}$																		1.00	-0.84	-
$\beta_{\mathrm{comb};3\pi}^{\mathrm{DT}}$																			1.00	0.06
$\alpha_{\mathrm{comb}}^{\pi\pi\pi}$																				1.00

Table D.2.: *Correlation matrix for the nominal fit using both SST and OST.*

APPENDIX E

Momentum scale systematic

For a simple two-body decay the mass of the B-meson is given by:

$$m^2_{B^0_s} = m^2_1 + m^2_2 + 2p_1p_2\sqrt{1 + \frac{m^2_1}{p^2_1}\frac{m^2_2}{p^2_2} + \frac{m^2_1}{p^2_1} + \frac{m^2_2}{p^2_2}} - 2p_1p_2cos(\theta_{12}). \tag{E.1}$$

Using $\frac{m_{1,2}}{p_{1,2}} << 1$ one receives:

$$m^2_{B^0_s} = m^2_1 + m^2_2 + 2p_1p_2(1 - cos(\theta_{12})). \tag{E.2}$$

Now a scaling of the momenta of $p_{1,2} \to (1-\alpha)p_{1,2}$ is considered.

$$m'^2_{B^0_s} = m^2_1 + m^2_2 + 2(1-\alpha)^2p_1p_2(1 - cos(\theta_{12})) \tag{E.3}$$

Factoring out $(1-\alpha)^2$ leads to

$$m'^2_{B^0_s} = (1-\alpha)^2\left(\frac{1}{(1-\alpha)^2}(m^2_1 + m^2_2) + 2p_1p_2(1 - cos(\theta_{12}))\right). \tag{E.4}$$

Taking the Taylor expansion of $\frac{1}{(1-\alpha)}$ in α until the linear term and neglecting quadratic terms in α:

$$\frac{1}{(1-\alpha)^2} \to (1+\alpha)^2 \to 1 + 2\alpha \tag{E.5}$$

one gets:

$$m'^2_{B^0_s} = (1-\alpha)^2\left((1+2\alpha)(m^2_1 + m^2_2) + 2p_1p_2(1 - cos(\theta_{12}))\right). \tag{E.6}$$

Inserting a zero $((m_1^2+m_2^2)-(m_1^2+m_2^2))$ gives:

$$m'^2_{B_s^0} = (1-\alpha)^2\left(((1+2\alpha)-1)\,(m_1^2+m_2^2)+m_{B_s^0}\right). \tag{E.7}$$

For the proper decay time we need the uncertainty on $\frac{m_{B_s^0}}{p_{B_s^0}}$. With the momentum of the B_s^0

$$p'_{B_s^0} = (1-\alpha)p_{B_s^0} \tag{E.8}$$

we get:

$$\frac{m'^2_{B_s^0}}{p'^2_{B_s^0}} = \frac{(1-\alpha)^2\left(2\alpha(m_1^2+m_2^2)+m_{B_s^0}\right)}{(1-\alpha)^2 p_{B_s^0}}. \tag{E.9}$$

This leaves us with:

$$\frac{m'^2_{B_s^0}}{p'^2_{B_s^0}} = \frac{2\alpha(m_1^2+m_2^2)}{p^2_{B_s^0}} + \frac{m^2_{B_s^0}}{p^2_{B_s^0}}. \tag{E.10}$$

Inserting a one $(\frac{m^2_{B_s^0}}{m^2_{B_s^0}})$ we have in the end:

$$\frac{m'^2_{B_s^0}}{p'^2_{B_s^0}} = \left(2\alpha\frac{(m_1^2+m_2^2)}{m^2_{B_s^0}}+1\right)\frac{m^2_{B_s^0}}{p^2_{B_s^0}}. \tag{E.11}$$

By inserting another negigible factor quadratic in α of the form $\left(\alpha\frac{m_1^2+m_2^2}{m^2_{B_s^0}}\right)^2$ we get:

$$\frac{m'^2_{B_s^0}}{p'^2_{B_s^0}} = \left(1+\alpha\frac{(m_1^2+m_2^2)}{m^2_{B_s^0}}\right)^2\frac{m^2_{B_s^0}}{p^2_{B_s^0}} \tag{E.12}$$

or

$$\frac{m'_{B_s^0}}{p'_{B_s^0}} = \left(1+\alpha\frac{(m_1^2+m_2^2)}{m^2_{B_s^0}}\right)\frac{m_{B_s^0}}{p_{B_s^0}} \tag{E.13}$$

multiplying with the flight distance the proper decay time is given by:

$$\frac{t'}{t} = \left(1+\alpha\frac{(m_1^2+m_2^2)}{m^2_{B_s^0}}\right) \tag{E.14}$$

So with α being the relative uncertainty of the momentum, we get a relative uncertainty of the proper decay time that has the form:

$$\frac{\Delta_t}{t} = \frac{(m_1^2+m_2^2)}{m^2_{B_s^0}}\frac{\Delta_p}{p}. \tag{E.15}$$

Since the D_s^+ mass constraint is not used for the computation of the proper decay time in principle a four-body decay is given. The above shown derivation holds in the same way for more-body decays. Therefore the final uncertainty is

$$\frac{\Delta_t}{t} = \frac{(2m_K^2+2m_\pi^2)}{m^2_{B_s^0}}\frac{\Delta_p}{p}. \tag{E.16}$$