Department of Physics and Astronomy

University of Heidelberg

Diploma thesis
in Physics

submitted by
Sebastian Wandernoth
born in Ludwigshafen am Rhein/Germany

July 2009

Measurement of the $\mathrm{B_s^0}$-$\overline{\mathrm{B}}_\mathrm{s}^0$ Mixing Frequency and Calibration of the Same Side Tagger at LHCb

This diploma thesis has been carried out by Sebastian Wandernoth at the Physical Institute
under the supervision of
Prof. Dr. Stephanie Hansmann-Menzemer

Kurzfassung

Der Large Hadron Collider LHC wird mit 10^{12} produzierten $b\bar{b}$ Paaren die weltgrößte B-Fabrik sein. Dies macht ihn zu einem idealen Standort für LHCb einem dediziertes B-Physik Experiment, das sich hauptsächlich mit Teilchen-Antiteilchen Asymmetrien beschäftigt. Wegen seiner hohen Statistik ist der Kanal $B_s^0 \to D_s^- \pi^+$ am besten geeignet für Studien der Mischung im B_s-System. Bei nomineller Luminosität von $L = 2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ werden innerhalb eines Jahres $2.5 \cdot 10^5$ Ereignisse rekonstruierter $B_s^0 \to D_s^- \pi^+$ Zerfälle erwartet. In dieser Arbeit wird gezeigt, dass dies um $3.6 \cdot 10^5$ Zerfälle der Kanäle $B_s^0 \to D_s^{*-} \pi^+$ und $B_s^0 \to D_s^- \rho^+$ erweitert werden kann, die ebenfalls wie $B_s^0 \to D_s^- \pi^+$ rekonstruiert werden. Neben dem Signal wurden in dieser Studie auch die wichtigsten Beiträge zum Hintergrund untersucht. Die erwartete statistische Sensitivität auf die Mischungsfrequenz Δm_s beträgt $\sigma(\Delta m_s) = 0.006\,\mathrm{ps}^{-1}$.
Eine der wichtigsten Analysen bei LHCb ist die Studie der Parameter der CP Verletzung im Zerfall $B_s^0 \to J/\psi\phi$. Hierfür sind sogenannte Flavor Tagging Algorithmen notwendig um den Flavor des B_s^0 bei der Produktion zu bestimmen. Der Same Side Tagger muss auf Daten in einem B_s^0 Referenzkanal kalibriert werden und $B_s^0 \to D_s^- \pi^+$ ist aufgrund der hohen Statistik am besten hierfür geeignet. Die statistische Unsicherheit für diese Kalibration wird durch das Einbeziehen der Zerfallsmoden $B_s^0 \to D_s^{*-} \pi^+$ und $B_s^0 \to D_s^- \rho^+$ um einen Faktor $\sqrt{1.7}$ erhöht.

Abstract

The Large Hadron Collider will be the largest B-factory in the world with 10^{12} produced $b\bar{b}$ pairs per year. Therefore, it provides the ideal setting for the LHCb experiment which is a dedicated B-physics experiment with its main focus on the measurement of particle anti-particle asymmetries. Because of its high statistics, the channel $B_s^0 \to D_s^- \pi^+$ is most suitable for studies of the oscillation in the B_s system. Within one nominal year of data taking at LHCb $2.5 \cdot 10^5$ reconstructed signal events are expected. In addition to that $3.6 \cdot 10^5$ events of the decay modes $B_s^0 \to D_s^{*-} \pi^+$ and $B_s^0 \to D_s^- \rho^+$ can also be used as signal. A study of the main background contributions has been performed on simulated data. The expected statistical sensitivity to the oscillation frequency was determined to be $\sigma(\Delta m_s) = 0.006\,\mathrm{ps}^{-1}$.
One of the main analyses at LHCb is the study of the parameters of CP violation in the decay $B_s^0 \to J/\psi\phi$. For this purpose so-called flavor tagging algorithms are necessary to determine the flavor of the B_s^0 at production. The same side kaon tagger needs to be calibrated on data in a B_s^0 reference channel and $B_s^0 \to D_s^- \pi^+$ is because of its high statistics the most adequate to operate as such. By adding the decay modes $B_s^0 \to D_s^{*-} \pi^+$ and $B_s^0 \to D_s^- \rho^+$ the statistical power for the determination of the mistag rate ω is improved by a factor 1.7.

Contents

Chapter 1

Introduction

The Standard Model of particle physics has been supported by many experiments, but there are still some unanswered questions like the existence of dark matter or the baryon asymmetry. There are two main approaches to the search of physics beyond the Standard Model. On the one hand there are direct searches for new particles, on the other hand there are searches for indirect effects of virtual particles in loop processes. The LHCb experiment follows the latter approach. There are several reasons why b-physics is such a strong tool for exploring these indirect searches.
First of all the large lifetime makes events that contain B-mesons leave a clear signature, *i.e.* they are more easily detectable. Secondly the big mass difference between the b-quark and the second quark in the meson makes it possible to use approximations for QCD processes that cannot be described by perturbation theory. Another reason is that the most interesting loop processes are strongly suppressed so that the small effects of the New Physics particles can be seen better. In b-decays there are several of these rare processes. The LHC is especially interesting from a B-physics point of view because of the large cross section of b-pair production at the energy at which the LHC will be operating. Every 150th proton-proton collision produces a $b\overline{b}$ pair which leads to 10^{12} produced $b\overline{b}$ pairs per year.
Neutral B-mesons like the B^0 and the B_s^0 can oscillate into their anti particles. The oscillation frequency Δm_s of the B_s^0 has been measured first by the CDF [1] and the DØ collaboration [2] at Tevatron. This thesis deals with the tools that are needed to describe this oscillation and measure the oscillation frequency Δm_s in the decay channel $B_s^0 \to D_s^- \pi^+$. These tools are the so-called flavor tagging algorithms on the one side and an unbinned maximum likelihood fitter on the other side. The flavor tagging algorithms are used to extract whether the B-meson was a B_s^0 or a $\overline{B}_s^0$ when it was produced.
One of the major analyses of the LHCb experiment is the measurement of the

CP violation[1] in the $B_s^0 \to J/\psi(\mu\mu)\phi(KK)$ decay. One necessary information for this analysis is the knowledge whether a B_s^0 or a $\overline{B}_s^0$ decayed into the final state $J/\psi\phi$. To obtain that information the aforementioned flavor tagging algorithms are used. These have to be properly calibrated on reference channels of which the decay channel $B_s^0 \to D_s^- \pi^+$ is the most adequate B_s^0 channel. In this thesis it will be shown how the calibration of the same side kaon tagger will be performed in this decay channel.
Chapter 2 of this thesis is an introduction to the theoretical background of the CKM mechanism and B-physics. In chapter 3, an overview of the detector of the LHCb experiment is presented. The framework of the analysis is displayed in chapter 4. The flavor tagging algorithms mentioned above will be explained in detail in chapter 5. The unbinned maximum likelihood fitter will be described in chapter 6.
During one year of nominal runtime of the LHC $2\,\mathrm{fb}^{-1}$ of integrated luminosity is expected for the LHCb experiment. Chapter 7 will describe what the expected statistical uncertainties on the measurement of Δm_s are with the statistics corresponding to $2\,\mathrm{fb}^{-1}$. Furthermore, the expected sensitivity to the calibration of the flavor tagging algorithms will be shown with the statistics corresponding to one nominal year of runtime at LHCb.
In chapter 8, a short summary of the findings in this analysis will be given.

[1] C stands for charge conjugation in which the particle is transformed into its anti-particle and P stands for the parity transformation which describes a spatial inversion

Chapter 2

Theory

2.1 Standard Model

According to today's state of knowledge there are four fundamental interactions (strong, weak, electromagnetic and gravitational). The Standard Model combines the first three of them in three gauge theories with the gauge groups $SU(3) \otimes SU(2) \otimes U(1)$. It describes the fundamental particles of which all matter is made up. Over the last 30 years it has been extremely successful in describing numerous experiments, e.g. the discovery of the W- and Z-bosons.
The point-like elementary particles are divided into two groups according to their spin. On the one hand there are fermions which are spin-$\frac{1}{2}$ particles, on the other

	I	II	III		
Quarks	u	c	t	γ	Bosons
	d	s	b	$W^{\pm}$	
Leptons	e	μ	τ	Z^0	
	ν_e	ν_μ	ν_τ	g	

Table 2.1: *Standard Model of particle physics - Three generations of fermions from lightest (left column) to heaviest (right column) and force carriers in the utmost right column*

hand there are bosons which have spin-1 and describe the interactions between the fermions (see Table 2.1). The electromagnetic and weak force are mediated by the photon γ, the $W^{\pm}$ and the Z^0. The carrier of the strong force is the gluon g. The fermions are again divided into quarks and leptons which both interact via the weak force. Except for the neutrinos all fermions interact electromagnetically.
There are 2 types of quarks. The up-type quarks have electric charge of $\frac{2}{3}$ and down-type quarks $-\frac{1}{3}$. Additionally quarks carry color charge, i.e. they interact via the strong force. The potential of the strong interaction rises with the dis-

tance between the quarks. That is the reason why quarks never exist on their own but always form so called hadrons that are color neutral to the outside, e.g. the proton or the neutron. Both quarks and leptons are grouped together in pairs, the so called generations, of which only the lightest is stable. Aside from these particles there are the anti-particles of the fermions which are identical to their partners except for a flipped sign in their electric charge.
One of the most important symmetries in the Standard Model is the CP symmetry where C stands for charge conjugation which transforms a particle into its anti-particle. P stands for the parity transformation which is a spatial inversion. The CP symmetry describes the fact that particles and anti-particles follow the same physical processes and have the same branching ratios. If there is an asymmetry between their decay rates the CP symmetry is broken. By measuring these asymmetries very accurately one gets a strong indicator for new physics beyond the Standard Model.

2.2 CKM Mechanism

The weak eigenstates of the quarks are not identical to their mass eigenstates. That is why it is possible for quarks to transform into quarks of a different flavor. The transformation from the mass eigenstates q to the weak eigenstates q' can be described by a unitary 3x3 matrix, the so called Cabbibo-Kobayashi-Maskawa (CKM) matrix [4]:

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

Since the diagonal elements of the CKM matrix are close to one, the transition within the quark families are most probable and transitions between generations are suppressed. These transitions are mediated through the charged current by exchanging $W^{\pm}$ bosons. So an up-type quark has to transform into a down-type one and vice versa.
A complex 3x3 matrix has 18 free parameters. Because of the CKM matrix being unitary that number is reduced to 9. By redefinition of the quark fields the number of free parameters is further reduced to 4, namely three real angles and one complex phase. This phase together with the different quark masses give the prerequisite for CP-violation in the Standard Model.
One approximation of the CKM matrix is the Wolfenstein parameterization [5] with the real parameters λ, A, ρ, and η to the order λ^3:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{2.2}$$

The complex entries in V_{ub} and V_{td} describe the above mentioned phase that is responsible for the CP-violation.
By using the unitarity relation

$$\sum_k = V_{ik}V_{kj}^* = \delta_{ij} \tag{2.3}$$

one can obtain equations that can be displayed as triangles in the complex plane like:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{2.4}$$

Normalizing (2.4) to $V_{cd}V_{cb}^*$ and defining the angles

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{2.5}$$

gives the unitarity triangle (Fig:2.1).

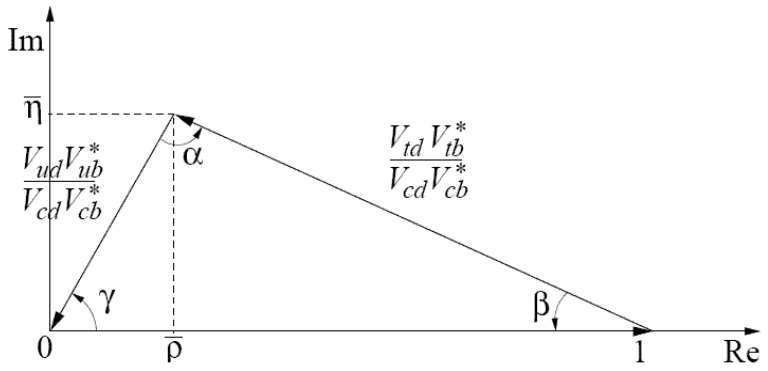


Figure 2.1: *Unitarity triangle from the first row and third column of the CKM matrix[3]*

Figure 2.2 shows the most recent results of the measurements of this unitarity triangle [7]. Precise measurements of these quantities give a very good handle on finding new physics because the theoretical uncertainties in the Standard Model are very small. So any significant discrepancy is a strong indicator for the existence of a process that is not yet described in the Standard Model.

2.3 B-Physics

2.3.1 B-Mesons

One of the big advantages of B-physics is that B-mesons have a large lifetime (see Table 2.2) which makes them leave a distinct signature. That is due to

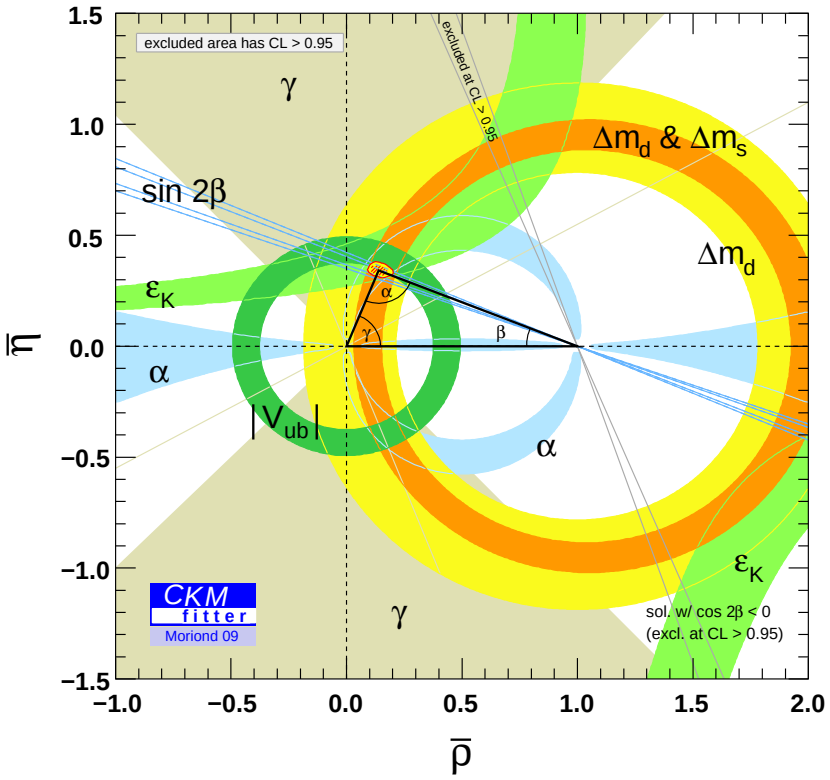
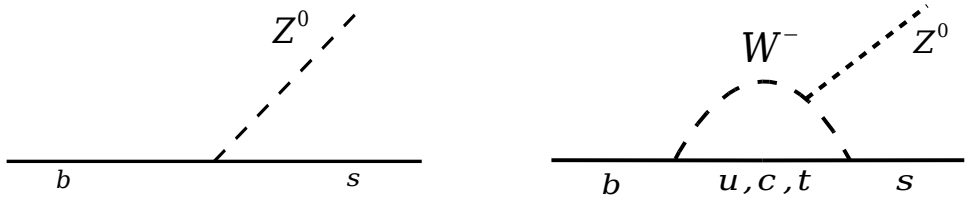


Figure 2.2: *Latest measurement of the unitarity triangle as of march 2009 [7]*

the fact that B-mesons aside from the Υ resonances carry the quantum number beautyness which is conserved under the strong interaction. Therefore they can only decay via the weak interaction. Here only decays into the lighter charm and up quarks are possible which are cabibbo suppressed. Cabibbo suppression refers to the smaller CKM matrix elements aside from the diagonal. Therefore these quark transitions are less probable to happen.

Additionally the so-called GIM suppression exists which makes b decays even less probable. The flavor change via the charged current, *i.e.* interaction via a W-boson is cabbibo suppressed. Flavor changing neutral currents are possible beyond the tree level in the Standard Model. This means that only in loop processes like in Figure 2.3(b) a transition from a b-quark to another down-type quark is possible. The sum over the decay amplitudes is proportional to term

$$m_u^2 V_{ub} V_{us} + m_c^2 V_{cb} V_{cs} + m_t^2 V_{tb} V_{ts}. \tag{2.6}$$



(a) FCNC in tree decays (b) FCNC in loop decays

Figure 2.3: *Feynman diagram of the flavor changing neutral current (FCNC) in a tree process (a), which is forbidden in the Standard Model, and in a loop process (b).*

Therefore if all quark masses were the same the process would be completely forbidden because of the unitarity relation

$$V_{ub}V_{us} + V_{cb}V_{cs} + V_{tb}V_{ts} = 0. \tag{2.7}$$

In comparison to the top mass the up and charm quark masses are tiny and nearly the same ($m_c \approx m_u \ll m_t$). Equation (2.6) becomes

$$m_u^2 \underbrace{(V_{ub}V_{us} + V_{cb}V_{cs} + V_{tb}V_{ts})}_{=0} + \underbrace{(m_t^2 - m_u^2)}_{=m_t^2} V_{ts}V_{tb} \tag{2.8}$$

and therefore the processes with a charm or up quark in the loop cancel and only the top quark remains.

2.3.2 B-Mixing

In this analysis the main focus is on neutral B-mesons of which there are two types with the quark content of

$$\begin{aligned} |\mathrm{B}^0\rangle &= |\overline{\mathrm{b}}\mathrm{d}\rangle, & |\overline{\mathrm{B}}^0\rangle &= |\mathrm{b}\overline{\mathrm{d}}\rangle, \\ |\mathrm{B}_\mathrm{s}^0\rangle &= |\overline{\mathrm{b}}\mathrm{s}\rangle, & |\overline{\mathrm{B}}_\mathrm{s}^0\rangle &= |\mathrm{b}\overline{\mathrm{s}}\rangle. \end{aligned} \tag{2.9}$$

B^0 and $\overline{\mathrm{B}}^0$ are anti particles of each other. Together they form the B_d system. Respectively B_s^0 and $\overline{\mathrm{B}}_\mathrm{s}^0$ form the B_s system. Their most important physical quantities are given in Table 2.2.
From now on to simplify indexing everything will take place in the B_s system. However all processes are as valid in the B_d system. One important characteristic of these neutral B-mesons is that they can transform into their antiparticles, the so called oscillation. This occurs because the mass eigenstates of the B-mesons

	B_d system	B_s system
mass [MeV]	5279.53 ± 0.33	5366.3 ± 0.6
mean lifetime [ps]	1.530 ± 0.009	$1.470^{+0.026}_{-0.027}$
$\Delta\Gamma$[ps]	Re(z) = -0.007 $\pm$ 0.004	$0.062^{+0.034}_{-0.037}$
$\Delta m[\hbar\,\mathrm{ps}^{-1}]$	0.507 ± 0.033	17.77 ± 0.17

Table 2.2: *Physical quantities of B-mesons, taken from [6]*

are different from their flavor eigenstates given in 2.9 The mass eigenstates are given as:

$$\begin{aligned} |\mathrm{B_L}\rangle &= p|\mathrm{B_s^0}\rangle + q|\overline{\mathrm{B}}_\mathrm{s}^0\rangle \qquad \text{and} \\ |\mathrm{B_H}\rangle &= p|\mathrm{B_s^0}\rangle - q|\overline{\mathrm{B}}_\mathrm{s}^0\rangle, \end{aligned} \tag{2.10}$$

with $|p|^2 + |q|^2 = 1$. Because of the different mass of $\mathrm{B_L}$ and $\mathrm{B_H}$ it is possible for the flavor eigenstates $\mathrm{B_s^0}$ and $\overline{\mathrm{B}}_\mathrm{s}^0$ to transform into each other (see Figure 2.4). The

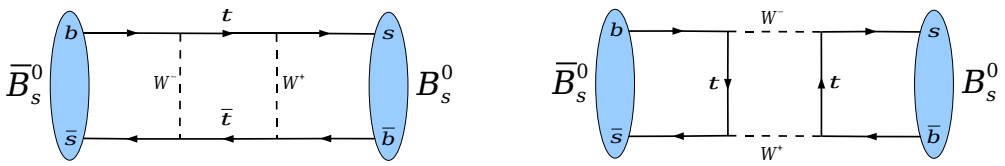


Figure 2.4: *Dominant Feynman diagrams for* $\mathrm{B_s^0}$ *mixing*

reason that there are only contributions from a top quark in the box diagrams and not from charm or up quarks is the above explained GIM mechanism. The frequency of this oscillation is proportional to the difference in mass Δm_s of $\mathrm{B_H}$ and $\mathrm{B_L}$.

2.3.3 CP Violation in the B_s system

The CP violation that was introduced at the end of section 2.1 is the main focus of many analyses in the LHCb experiment. There are three possibilities of CP violation in the B_s system:

- **CP Violation in mixing** means that the probability for $\mathrm{B_s^0} \to \overline{\mathrm{B}}_\mathrm{s}^0$ is different from the probability for $\overline{\mathrm{B}}_\mathrm{s}^0 \to \mathrm{B_s^0}$. This is the case if $|\frac{q}{p}| \neq 1$ in

(2.10). However this source of CP violation is predicted to be very small [8].

- **CP Violation in decay** means that the decay amplitude into the same final state f $A_f = \langle f|\mathrm{B_s^0}\rangle$ is different from the amplitude $\bar{A}_f = \langle f|\overline{\mathrm{B}}_\mathrm{s}^0\rangle$. This means that the probability for $\mathrm{B_s^0} \to f$ is different from the probability for $\overline{\mathrm{B}}_\mathrm{s}^0 \to f$. The size of the expected CP asymmetry strongly depends on what decay is analyzed. There are several studies in the LHCb experiment that examine this type of CP violation.

- **CP Violation in the interference between mixing and decay** means that the CP symmetry is broken because of a phase difference between $\mathrm{B_s^0}$ mixing and decay. One of the analyses within the LHCb experiment that deals with this source of CP violation is the study of the decay $\mathrm{B_s^0} \to \mathrm{J}\!/\!\psi(\mu\mu)\phi(\mathrm{KK})$ which will be presented briefly now.

In this decay the $\mathrm{B_s^0}$ can either decay directly or first oscillate into a $\overline{\mathrm{B}}_\mathrm{s}^0$ and then decay. This introduces a mixing phase Φ.

$$\Phi = -\arg\left(\frac{q}{p}\frac{A_f}{\bar{A}_f}\right) \tag{2.11}$$

which is in the Standard Model equal to

$$\Phi = -2\beta_\mathrm{s} = -2\arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right), \tag{2.12}$$

with β_s being the angle in the corresponding unitarity triangle. Its theoretical prediction in the Standard Model is:

$$-2\beta_\mathrm{s} = -0.0360^{+0.0020}_{-0.0016}\,\mathrm{rad} \qquad [8]. \tag{2.13}$$

Because of the small theoretical uncertainties any significant deviation from that predicted value is a clear sign for a contribution from New Physics. A detailed description of the analysis of the channel $\mathrm{B_s^0} \to \mathrm{J}\!/\!\psi\phi$ in the LHCb collaboration is given in [9].

Chapter 3

The LHCb Experiment

The European nuclear research center CERN[1] in Geneva has become the largest organization for particle physics in the world. It was founded in 1954 and since then some of the most important discoveries of particle physics were made at CERN like for example the discovery of the $W^{\pm}$ and Z^0 bosons.
The most recent project at CERN is the Large Hadron Collider (LHC) which is supposed to start operation in fall 2009. One of the four big experiments at the LHC is LHCb, the Large Hadron Collider beauty experiment. Here the unprecedented center of mass energy of 14 TeV will be used to perform precision measurements in the B-sector and searches for signs of new physics beyond the Standard Model. In this chapter after a short introduction of the accelerator and the other experiments the individual components of the LHCb detector will be presented.

3.1 The Large Hadron Collider

The LHC is a 27 km long underground proton-proton synchrotron. The protons are accelerated to receive an energy of 7 TeV and then assembled in so called bunches of 10^{11} protons each. Per direction 2808 of these bunches will be in the accelerator at a time. At the interaction points of the four big experiments (see Figure 3.1) the two proton beams are focused and brought together so that every 25ns (40MHz) two bunches collide which leads to a maximum luminosity of $L = 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.
The four most important experiments at the LHC are ATLAS[2], CMS[3], ALICE[4] and LHCb. They are designed for once to do precision measurements of important quantities in the Standard Model, but also to search for signs of new physics.

[1] Conseil Européen pour la Recherche Nucléaire
[2] A Toroidal LHC ApparatuS
[3] Compact Muon Solenoid
[4] A Large Ion Collider Experiment

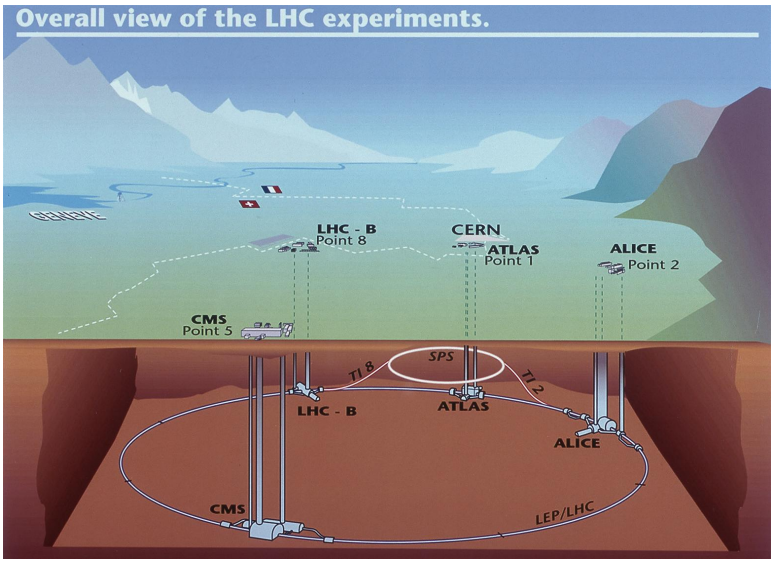


Figure 3.1: *Schematic view of the LHC tunnel with the four big experiments ATLAS, ALICE, CMS and LHCb. Taken from [10].*

The main focus of the ATLAS and CMS experiments is to measure the Higgs boson and search for supersymmetric particles. The Higgs boson is the last particle predicted by the Standard Model that has not been seen yet in any experiment. Supersymmetric particles are part of theories beyond the Standard Model and candidates for Dark Matter which plays an important role in cosmology.
The aim of the ALICE experiment is to examine a state of matter called the quark-gluon plasma in which quarks and gluons are deconfined. It is assumed that this state existed shortly after the big bang. The quark-gluon plasma manifests itself at immense temperatures and densities. This is the reason why for the main experiments ALICE will have collisions of heavy ions (Pb-nuclei) instead of protons.
The LHCb experiment has its main focus on precision measurements in rare decays of B-mesons. Especially the violation of CP symmetry is going to be examined within the framework of the LHCb experiment. Comparison to the predicted Standard Model values can give signs of new physics. The LHCb detector is optimized for a luminosity of $L = 2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ because for higher luminosities there would be too many tracks in the high pseudo-rapidity region that is examined. Therefore the beams are defocused before collision to lower the luminosity. The $\mathrm{b\bar{b}}$ cross section at 14 TeV center of mass energy is assumed to be around 500 μbarn. This leads to 10^{12} produced $\mathrm{b\bar{b}}$ pairs per year (assuming a runtime of 10^7 seconds per year) which makes the LHC the largest B-factory in the world.

3.2 The LHCb Detector

The detector of the LHCb experiment is a one-armed forward spectrometer (see Figure 3.2). The reason that this design was chosen is that $b\bar{b}$ pairs are produced strongly boosted into either the forward or backward direction (see Figure 3.3(a)). This occurs because at the center of mass energy at the LHC the Bjorken variable x which describes the fraction of the total momentum that the single partons

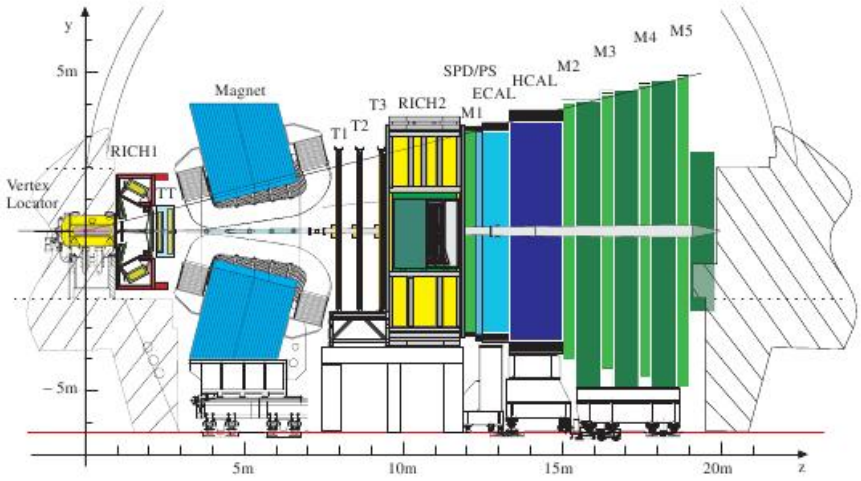


Figure 3.2: *Schematic side view of the individual components of the LHCb detector. Taken from [11]*

carry is about 10^{-3}. Therefore, the proton-proton interactions are dominated by gluon fusion processes. Figure 3.3(b) shows the parton density functions inside the proton plotted against the Bjorken variable x. To produce a $b\bar{b}$ pair an energy of 10 GeV is necessary. Therefore, a low energy gluon and a medium energy gluon suffice to produce it and the combined momentum is not equal to 0. So there is a large number of possible combinatorics for the gluons which results in the $b\bar{b}$ pair being boosted in forward or backward direction (see Figure 3.3(a)).
In case of a high mass particle like the Higgs ($m > 140\,\text{GeV}$) only high energy gluons have enough energy to produce it which makes the problem a symmetrical one where the system is not heavily boosted in beam direction. This is the reason why the ATLAS and CMS detectors have a cylindrical shape around the beam pipe instead of a forward spectrometer.
The LHCb detector consists of the individual components displayed in Figure 3.2. These detector components can be divided according to their use in two categories. On the one hand the vertex locator, trigger tracker and the tracking stations are used for track reconstruction. On the other hand RICH, calorimeters and muon chambers are used for particle identification. In the following part each of these components will be presented.

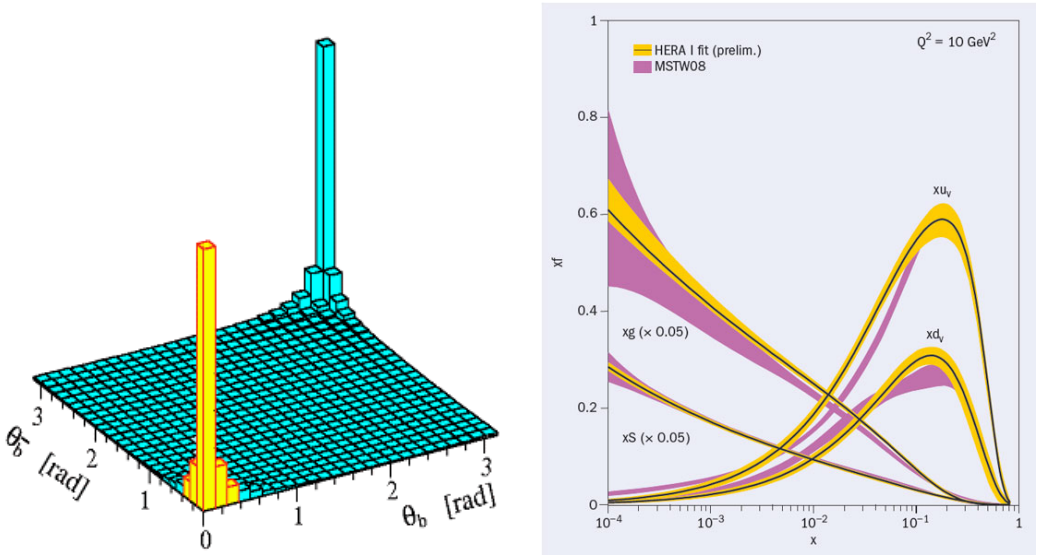


Figure 3.3: *Polar angles θ of the* b- *and the* $\overline{\text{b}}$*-hadrons (a), taken from [12]. Parton density function for the proton (b), taken from [13]*

3.2.1 Track Reconstruction

Between the trigger tracker (TT) and the tracking stations (T1, T2, T3) is the LHCb magnet [14]. The magnetic field amounts to about 1 Tesla and the integrated magnetic field ($\int Bdl$) is about 4Tm. This magnetic field is used to determine the momentum of the particle by comparing the tracks before and after the magnet. The LHCb coordinate system has its z-axis along the beam. The detector covers an angular range of 10 - 250 mrad in y-direction (vertical) and 10 - 300 mrad in x-direction (horizontal direction of deflection by the magnetic field). Thereby about 25% of the produced B-mesons can be reconstructed at LHCb.

Vertex Locator

The vertex locator (VeLo) is located directly around the interaction point. It consists of 21 Stations with two semi-circular silicon modules [15]. There are two different kinds of these modules which are arranged alternately along the beam. The so-called R-sensors are used to determine the radial distance from the beam and the ϕ-sensors are used to determine the azimuthal angle of the track.
The modules have a small overlap in x-direction and have a distance of about 2 mm in z-direction. The whole VeLo is located in a vacuum which is needed because there is no beam pipe around the interaction point at LHCb to allow the modules to be very close to the beam (5mm). This is a fundamental difference to the other LHC experiments. The modules can be moved away from the beam

during the time until the beam is completely focused to avoid damaging them. The spatial resolution of the primary vertex is about 42 μm in z-direction and 10 μm in x- and y-direction.

Trigger Tracker

The trigger tracker (TT) is located between the RICH1 and the magnet. Aside from the track reconstruction it is used to determine the momentum of very low energetic particles. These are deflected so much by the magnetic field that they are bent out of the acceptance of the tracking stations. Furthermore, it is used for the reconstruction of very long-living particles like the K^0_S which decay after the VeLo.
The trigger tracker is composed of silicon strips. The first and the last of the four layers are arranged parallel to the y-axis and the inner two layers are tilted by $\pm 5\%$ to achieve a spatial resolution in x-direction as well as in y-direction.

Tracking Stations

The tracking stations T1, T2 and T3 are made up of the inner tracker (IT) and the outer tracker (OT). The inner tracker is for the reconstruction of tracks with small polar angle θ. It consists of silicon strips which cover 2% of the total acceptance, but about 20% of all tracks [16].
The outer track is composed of straw tubes which are 4.9 mm in diameter and 2.5 m long. In the middle of these straws is an anode wire and they are filled with a mixture of 30% CO_2 and 70% argon to function as a drift chamber detector. With this mixture a maximal drift time of 43 ns is achieved and so two consecutive proton-proton collisions (25ns) can only be resolved by widening the read-out window to 75ns [17]. This is an additional challenge for the tracking software because the so-called spill-over from the previous event has to be separated from the hits of the actual event.
Both the inner and outer tracker have like the trigger tracker layers that are tilted around the z-axis to achieve a spatial resolution in x- and y-direction. The spatial resolution in x-direction within the IT will be around 50 μm in every layer and in the outer tracker about 200 μm.

3.2.2 Particle Identification

Because of the high number of tracks in one event (≈ 30) it is important to have a good particle identification (PID) system to reject background. Especially the separation of kaons and pions constitutes a challenge. The particle identification system of LHCb which consists of two RICH detectors, the calorimeter system and the muon chambers is able to assign different probabilities to a reconstructed track according to different particle hypotheses.

Ring Imaging CHerenkov detector

The ring imaging cherenkov detector (RICH) consists of two detectors, one between the VeLo and the trigger tracker and one between the tracking stations and the calorimeters. The reason for that is to be able to also receive PID information for the low energetic particles which are bent out of the detector acceptance by the magnetic field. The RICH1 detector can separate kaons and pions with momentum smaller than 50 GeV and the RICH2 is designed to separate particles with momentum between 50 GeV and 100 GeV[18].
The principle of the RICH detectors originates in the Cherenkov effect according to which high energetic particles emit light if they traverse through a medium with refraction index n with a velocity v which is higher than the speed of light in this medium $c' = \frac{c}{n}$. The light is emitted in the shape of a cone with opening angle θ_{C}

$$\cos\theta_{\mathrm{C}} = \frac{c'}{v} = \frac{1}{\beta n}. \tag{3.1}$$

By projecting the light cone onto hybrid photo detectors the radii of the resulting rings give information over the particles velocity. Together with the momentum information from the tracking system a mass hypothesis can be made.

Calorimeters

The main task of the calorimeter system is the separation between electrons and photons or electrons and hadrons as well as measuring their energies [19]. The first layer of the calorimeter system which the particles have to pass is the scintillator pad detector (SPD). It consists of 15mm thick scintillators. Only charged particles leave a signal in here and so a separation between electrons and photons can be established.
Afterwards the particles pass a 12mm thick lead layer in which they induce electromagnetic showers, *i.e.* cascades of secondary particles which are detected by the scintillators of the pre-shower detector (PS). Since the hadronic showers are induced later in the hadronic calorimeter this detector can be used to separate electrons and hadrons.
The electronic calorimeter (ECAL) detects particle showers induced by electrons and photons. It is the only detector layer that is able to detect the non-charged photons. It consists of alternating layers of 2mm thick lead tiles and 4mm thick scintillators. The passing particles induce showers in the lead by means of pair production or bremsstrahlung which are measured by the scintillators. The energy resolution in the ECAL is

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus 1.5\%, \tag{3.2}$$

where $\oplus$ indicates a quadratic summation.
The final layer of the calorimeter system is the hadronic calorimeter (HCAL). It

is composed of alternating layers of 16mm thick iron tiles and 4mm thick scintillators. It is designed to measure the energy of hadrons, which induce showers in the iron, which are then detected by the scintillators. The energy resolution of the HCAL is

$$\frac{\sigma(E)}{E} = \frac{80\%}{\sqrt{E}} \oplus 10\%, \tag{3.3}$$

The hadrons deposit most of their energy in the HCAL which is why they do not leave signals in the muon chambers.

Muon Chambers

The muon chambers are used mainly for muon identification, but also for the reconstruction of tracks. The first muon chamber M1 is located in front of the calorimeter system to improve the spatial resolution and the other four M2 - M5 are positioned behind the calorimeters. Muons are the only particles which traverse through the whole detector and reach the muon chambers. These are composed of multi-wire proportional chambers (MWPC) and gas electron multiplier (GEM) [20].
In between the chambers are 80mm thick lead tiles to shield from hadrons. Additionally there is a lead tile behind the last muon chamber (M5) to shield from muons coming from the other direction. To pass through all five chambers a muon requires a momentum of about 6 GeV. Since the the track density increases towards the beam pipe a higher granularity in this region was chosen. Also is the resolution in x-direction higher than in y-direction to achieve a better momentum resolution. With this design of the muon chambers about 20% of all muons produced in semi-leptonic B-decays are within the detector acceptance.

3.2.3 Relative Particle Hypothesis

Using the detector components presented in the previous section it is possible to assign a relative particle hypothesis to a track $\Delta \ln \mathcal{L}_{\mathrm{X}\pi}$ (delta log-likelihood). The cherenkov angle θ_{C} that is measured in the RICH can be used to make a mass hypothesis using

$$\cos\theta_{\mathrm{C}} = \sqrt{1 + \left(\frac{m}{p}\right)^2}. \tag{3.4}$$

Together with the information from the calorimeters and the muon chambers a likelihood value can be assigned to the mass hypothesis. Since most of the tracks originate from pions this likelihood is normalized relatively to that of a pion $\mathcal{L}(\pi)$

$$\Delta \ln \mathcal{L}_{\mathrm{X}\pi} = (\ln \mathcal{L}(\mathrm{X}) - \ln \mathcal{L}(\pi)) = \ln\left(\frac{\mathcal{L}(\mathrm{X})}{\mathcal{L}(\pi)}\right) \tag{3.5}$$

This means if $\Delta \ln \mathcal{L}_{\mathrm{X}\pi} > 0$ it is more probable that the particle was a particle X than a pion.

3.2.4 Trigger System

One of the challenges of the LHCb experiments is handling the huge amounts of data and selecting the interesting events. As mentioned before proton-proton collisions occur with a rate of 40MHz at the LHC. The maximum rate of the data that can be written on tape is 2kHz. So the trigger system has to select online out of all interactions the events which are most likely to contain the decays that are analyzed at LHCb.
The LHCb trigger system is a 2-stage system consisting of a hardware trigger (Level-0 trigger) and a software trigger (High Level Trigger).

Hardware Trigger

The hardware level-0 trigger (L0) is designed to reduce the read-out rate from 40MHz to 1MHz. To do that the L0 trigger uses information from the muon chambers, the calorimeters and the VeLo. It selects events with high transverse momentum in the muon chambers and high transverse energy in the calorimeters. Furthermore it is able to veto events which are particularly busy, *i.e.* they have a high number of tracks, or events which have multiple primary vertices.

Software Trigger

The second stage of the trigger system is the high level software trigger (HLT). A computer farm of 1000 16-core computing nodes is used to carry out the calculations needed by the software trigger. It is divided into two parts HLT1 and HLT2. The HLT1 reduces the read-out rate from 1MHz to 30kHz. It performs a partial reconstruction of the event using seeds from the L0 trigger to establish a region of interest where the reconstruction is carried out.
Then the HLT2 picks up the events passing the HLT1 and performs a full track reconstruction of the event. Several inclusive selections of B-decays are performed to reduce the rate to the target output rate of 2kHz.

Chapter 4

Setup of the Analysis

4.1 Overview

The goal of this analysis is to provide the tools to measure the quantities of B-mixing which was presented in section 2.3.2. The fast oscillation of the B_s^0 meson was first discovered by the LEP experiment at CERN [21]. The oscillation frequency Δm_s was first measured by the CDF collaboration [1] and the DØ collaboration [2].
The aim of this analysis is to extract Δm_s from the reconstruction of the decay $B_s^0 \to D_s^-(K^+K^-\pi^-)\pi^+$. To do that it is crucial to know whether the B-meson has mixed before it decayed. Here having mixed means that the production flavor of the B_s^0 is different from its decay flavor. It does not matter how often it has oscillated in between.
The advantage of the decay channel $B_s^0 \to D_s^-\pi^+$ is that the decay flavor of the B-meson is given directly by the charge of the bachelor $\pi^\pm$ since a B_s^0 always decays into $D_s^-\pi^+$ and a $\overline{B}_s^0$ into $D_s^+\pi^-$. The flavor at production is determined by so-called flavor tagging algorithms which will be presented in section 5.
One difficulty that one is faced with conducting this analysis is that due to the fast oscillation the resolution of the proper time has to be excellent. The proper time t of a particle is the time between its production and decay in its rest frame. It is defined as:

$$t = \frac{FD \cdot m}{p}, \tag{4.1}$$

with the flight distance FD, which is the distance from the production vertex to the decay vertex of the particle, the mass m and the momentum p of the B_s^0. The uncertainty in the proper time is dominated by the spatial resolution of the reconstructed vertices, *i.e.* the flight distance. In Figure 4.1 is shown what the proper time resolution for this decay looks like. It follows a Double Gaussian distribution:

$$DG(\mu, f_\sigma, \sigma_1, \sigma_2; x) = f_\sigma \cdot G(\mu, \sigma_1; x) + (1 - f_\sigma) \cdot G(\mu, \sigma_2; x), \tag{4.2}$$

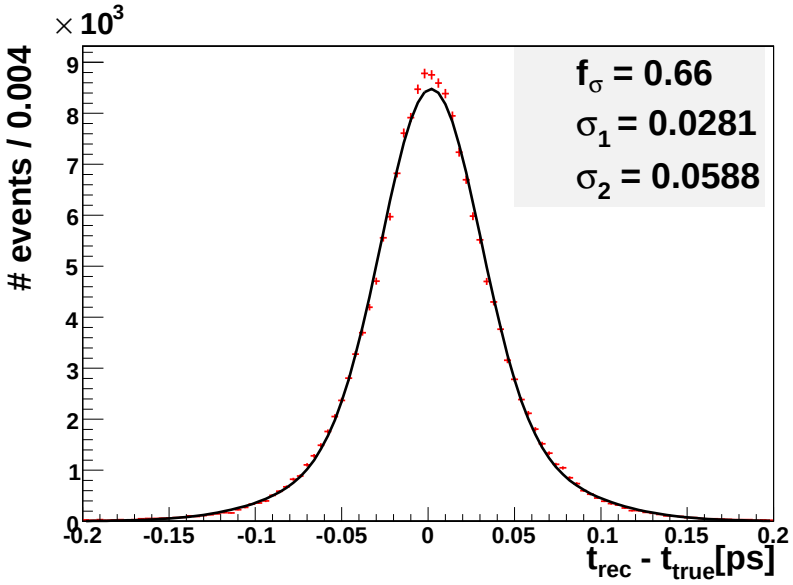


Figure 4.1: *Proper time resolution σ_t extracted from Monte Carlo simulated data. Plotted is the difference between reconstructed proper time and true proper time of the B_s^0. Fitted is a Double Gaussian distribution*

which leads to a total proper time resolution of

$$\sigma_t := \sqrt{f_\sigma \cdot \sigma_1^2 + (1 - f_\sigma) \cdot \sigma_2^2} = 0.0412\,\text{ps}. \tag{4.3}$$

In the B_d system the oscillation is so slow that the uncertainty of the measured proper time does not disturb the determination of the oscillation frequency Δm. In the B_s-system however the wavelength of the oscillation and the proper time resolution are of the same order of magnitude so it has a significant effect on the measurement of mixing parameters.
Another difficulty that presents itself when examining a B_s^0 decay channel is the lower statistics. b-quarks form only in about 11% of the cases a B_s^0(see Table 4.1. So compared to a B^0 or B^+ channel there is about a factor 4 less statistics available. Even though the channel $B_s^0 \to D_s^- \pi^+$ has one of the largest branching ratios among all B_s^0 channels this problem still exists.
To enhance the available statistics in this analysis the mass window around the B_s^0 mass is widened so that the partially reconstructed decay modes $B_s^0 \to D_s^{*-} \pi^+$ and $B_s^0 \to D_s^- \rho^+$ can be included as signal. These decays are called partially reconstructed because one misses a neutral particle in the reconstruction. In case of $B_s^0 \to D_s^{*-} \pi^+$ the $D_s^{*\pm}$ decays into $D_s^\pm \gamma$ and the γ is lost because neutral

particles are not reconstructed in this analysis. In case of the $B_s^0 \to D_s^- \rho^+$ decay mode the $\rho^\pm$ decays into $\pi^\pm \pi^0$ and the π^0 is lost. Because of the missing particle in the reconstruction the reconstructed mass is shifted toward the lower mass sideband. In Figure 4.2 the mass distribution of these three signal decay modes is displayed. By including these partially reconstructed decays the statistics is enlarged by a factor of about 2.4. In the following sections of this chapter it will be presented what selection cuts were applied and which signal and background contributions are expected.

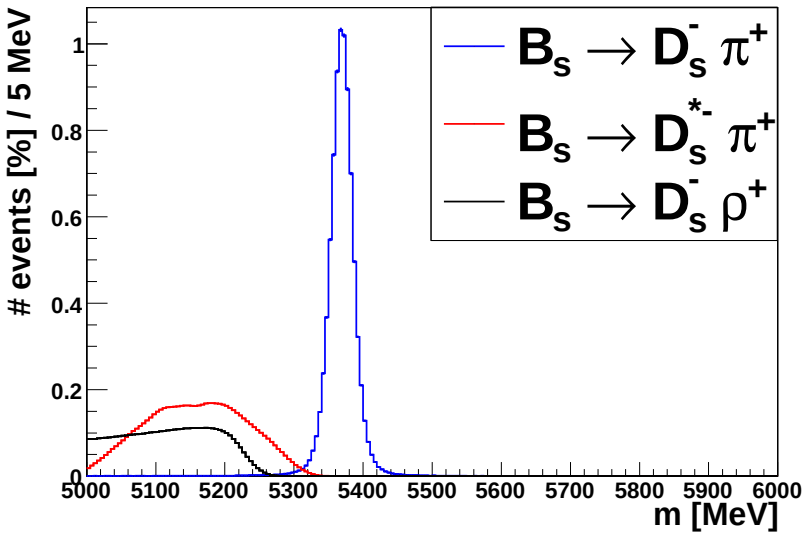


Figure 4.2: *Distribution of the reconstructed mass of the* B_s^0 *meson for fully reconstructed* $B_s^0 \to D_s^- \pi^+$ *decays (blue) and partially reconstructed* $B_s^0 \to D_s^{*-} \pi^+$ *decays (red) and* $B_s^0 \to D_s^- \rho^+$ *decays (black).*

4.2 Selection for the Decay $B_s^0 \to D_s^- \pi^+$

Since data from the LHC is not available yet, the analysis depends on data samples that are generated by Monte Carlo simulation. For the LHCb experiment Pythia is used for generating these samples. First the primary collision between two protons is simulated. Then a full simulation of the detector is performed by use of the Geant 4 [22] software including simulated hits in all detector layers. Afterwards the LHCb analysis software is used to reconstruct the tracks of the

simulated particles and put together the whole decay tree.
In this section the standard selection cuts (see [23][1]) that were used in this analysis will be presented. The simulated data samples that were used in this analysis were produced within the framework of the LHCb Monte Carlo Data Challenge 06. A complete list of the samples that were used for the different decay modes is provided in Table A.1 in the appendix.
The following formula is used for the computation of the number of events expected within a certain integrated luminosity:

$$N_{evts} = L_{int} \cdot \sigma_{\mathrm{b\bar{b}}} \cdot f_B \cdot 2 \cdot BR_{vis} \cdot \varepsilon_{gen} \cdot \varepsilon_{sel} \cdot \varepsilon_{L0} \tag{4.4}$$

where L_{int} is the integrated luminosity which is for 1 year of nominal runtime equal to $2\,\mathrm{fb}^{-1}$. $\sigma_{\mathrm{b\bar{b}}}$ is the cross section for the $\mathrm{b\bar{b}}$ production. In the DC06 production it is set to $700\,\mu$barn. f_B describes the fraction of b-quarks that hadronize into a specific B-meson (see Table 4.1). The factor 2 derives from the fact that b-quarks are always produced in $\mathrm{b\bar{b}}$ pairs so in every event there are two of them. BR_{vis} stands for the visible branching ratio, *i.e.* the product of the branching ratios of all decays that are required in the analysis. In the case of the signal decays for example it is the product of the Branching Ratio for $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$

Hadron	$f_B[\%]$
$\mathrm{B^0}$	39.9 ± 1.1
$\mathrm{B^+}$	39.9 ± 1.1
$\mathrm{B_s^0}$	11.0 ± 1.2
Λ_b	09.2 ± 1.9

Table 4.1: *Hadronization of b-quarks (taken from [6])*

($= 2.6 \cdot 10^{-3}$[6]) and for the decay $\mathrm{D_s^-} \to \mathrm{K^+K^-}\pi^-$ ($= 4.9 \cdot 10^{-2}$[6]). ε_{gen} is the generator cut efficiency which describes the fraction of events that are rejected due to cuts that were applied already on generator level like for example the so-called DecProdCut which is applied to almost all of the data samples used in this analysis. In this cut a generated event is only accepted if the decay products, *i.e.* the stable particles (in this case Kaons and Pions), are between 10 mrad and 400 mrad around the z-axis which roughly corresponds to the geometrical acceptance of the detector. ε_{sel} is the selection efficiency which is defined as the number of events passing the selection over all generated events. ε_{L0} is the L0 trigger efficiency which is the number of selected events that would also have passed the L0 trigger over all events that passed the standard selection.
The individual cuts of the standard selection [23] are displayed in Table 4.2.

[1] An updated version of the standard selection has recently been presented in [24], but at the time that this analysis has been conducted it had not been available yet.

All particle momenta	> 2000 MeV/c
All π DLL $K - \pi$	< 5
$D_s^\pm$ K daughters DLL $K - \pi$	> 0
$D_s^\pm$ K daughters DLL $K - p$	$>$ -10
$D_s^\pm$ daughter p_T	> 300 MeV/c
$D_s^\pm$ p_T	> 2000 MeV/c
$D_s^\pm$ daughter IP χ^2	> 9
$D_s^\pm$ IP χ^2	> 9
$D_s^\pm$ mass	$\pm$ 21 MeV/c^2
$D_s^\pm$ vertex χ^2	< 15
$D_s^\pm$ FSPV χ^2	> 100
B_s^0 tight mass	$\pm$ 50 MeV/c^2
B_s^0 vertex χ^2	< 10
B_s^0 IP χ^2	< 16
B_s^0 FSPV χ^2	> 6.25
B_s^0 $\cos\theta$	> 0.9999

Table 4.2: *Selection cuts from [23]*

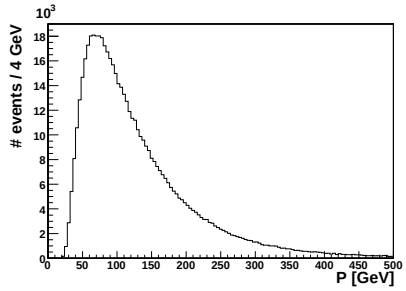
The different quantities that are cut on in Table 4.2 are:

- **DLL $K - \pi$ separation**
  DLL stands for Delta Log Likelihood. The two Ring-Imaging Cherenkov (RICH) detectors that were described in chapter 3 give a probability for the reconstruction of a track under certain particle hypotheses in the form of a loglikelihood. The difference of two of these loglikelihoods (DLL) is a measure for the probability of one particle hypothesis relative to another. In this case it compares the π hypothesis to the K hypothesis, but it can be formed between any two mass hypotheses.

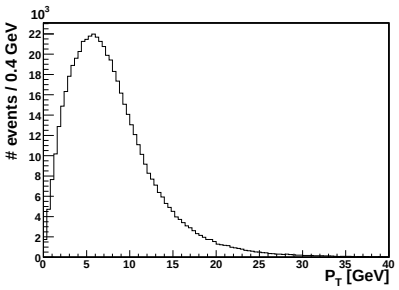
- **Impact parameter**
  The impact parameter (IP) of a track is defined by the perpendicular distance of the track to the primary vertex. In Figure 4.4 there is a schematic display of the impact parameter definition. The impact parameter χ^2 is defined as the square of the IP significance which is the IP divided by its uncertainty.

$$\mathrm{IP}\chi^2 = \left(\frac{\mathrm{IP}}{\sigma_{\mathrm{IP}}}\right)^2 \tag{4.5}$$

- **Transverse momentum $\mathbf{p_T}$**
  The transverse momentum p_T is defined as the portion of the momentum perpendicular to the z-axis in the laboratory system (beam axis). In Figure 4.3 the distributions of the reconstructed momentum p and transverse momentum p_T are shown.



(a) B_s^0 momentum p



(b) B_s^0 transverse momentum p_T

Figure 4.3: *Distributions of the reconstructed momentum (a) and transverse momentum (b) of the* B_s^0

- **Vertex χ^2**
  The vertex[2] χ^2 is the value that is minimized by the vertex fitter. It is defined as the squared sum of the perpendicular distances between a vertex and all tracks assigned to it. The uncertainties in these distances are included in the calculation.

- **FSPV**
  The minimum flight separation from the primary vertex (FSPV) is defined as the spatial separation of a particle to its primary vertex. If there is more than one primary vertex in the event, the one with the smallest IP χ^2 with respect to the B_s^0 is chosen. The principle of this variable is also displayed in Figure 4.4. The cut is performed on the χ^2 of the FSPV which is defined in a similar way as equation (4.5).

- **Flight direction angle** $\cos\theta$
  The angle θ is defined as the angle between the momentum and the flight direction of the B_s^0. The flight direction is the vector joining the particle's production vertex and decay vertex (see Figure 4.4 for a schematic display). Again, in case of more than one primary vertex in the event, the one with the smallest IP χ^2 with respect to the B_s^0 is chosen.

Using equation (4.4) one is left with 6.4 million signal events in $2\,\mathrm{fb}^{-1}$ integrated luminosity before selection. After applying these cuts 570229 events are selected which corresponds to a selection efficiency of $\varepsilon_{sel} = 8.9\%$. After the L0 trigger one obtains 249310 events which corresponds to a L0 efficiency of $\varepsilon_{L0} = 43.3\%$. These numbers are obtained from a signal sample of $B_s^0 \to D_s^- \pi^+$ decays (see Table A.1).

[2]vertex here always refers to the decay vertex of a particle

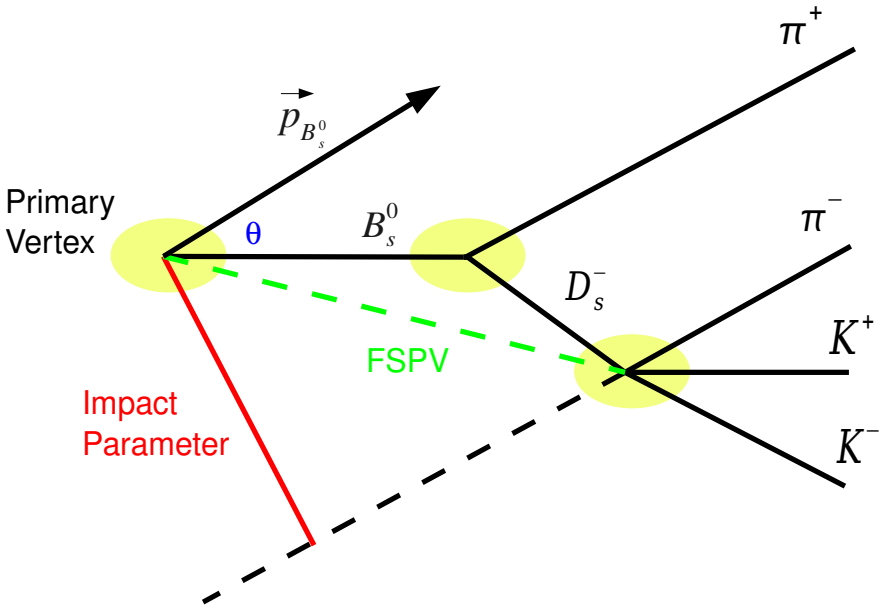


Figure 4.4: *Schematic display of the cut variables impact parameter (red), FSPV (green) and flight direction angle θ (blue)*

4.3 The Partially Reconstructed Modes $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$

To include the partially reconstructed decay modes $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$as signal into the analysis the mass cut described in Table 4.2 is widened to

$$5000\,\mathrm{MeV}/c^2 < \mathrm{m} < 6000\,\mathrm{MeV}/c^2. \tag{4.6}$$

Table 4.3 shows the expected annual yield for the fully and partially reconstructed decay modes. By the time this analysis was conducted there was no full simulation Monte Carlo data sample for the decay $B_s^0 \to D_s^-\rho^+$ available that corresponds to 2 fb^{-1}. The selection efficiency ε_{sel} that was used to calculate the number of reconstructed events is estimated from the relative number of reconstructed $B_s^0 \to D_s^-\rho^+$ decays to fully reconstructed decays in the inclusive $B_s^0 \to D_s^-X$ sample[3]. For the L0 trigger efficiency the same as for the $B_s^0 \to D_s^{*-}\pi^+$ decay was assumed ($\varepsilon_{L0} = 42.8\%$). Since the Branching Ratios of the modes $B_s^0 \to D_s^-\rho^+$ and $B_s^0 \to D_s^{*-}\pi^+$ have not been measured yet their yields are computed from the fractions relative to the fully reconstructed mode $B_s^0 \to D_s^-\pi^+$. The used fractions

[3] An inclusive data sample does not have the requirement that one specific decay is included, but that every event has to contain at least one $B_s^0 \to D_s^-X$ decay

are identical to those used by the CDF collaboration [25] in their analysis of mixing in the decay $B_s^0 \to D_s^- \pi^+$.

decay mode	Branching Ratio [10^{-3}]	annual yield [10^6 events]	after reconstruction	after L0
$B_s^0 \to D_s^- \pi^+$	2.6	6.38	570229	249310
$B_s^0 \to D_s^- \rho^+$	7.3	16.74	328693	140960
$B_s^0 \to D_s^{*-} \pi^+$	2.7	6.62	512271	219688

Table 4.3: *Expected yield in 2*$\,\text{fb}^{-1}$ *for fully and partially reconstructed modes*

4.4 Background Composition

For the background studies a sample of 851000 $b\bar{b}$ inclusive events were examined. 12 B_s^0 candidates passed the selection cuts. Within these 12 events there were 11 which also contained a $D_s^\pm$-meson. The 12th event was a event in which a B^+ interacted with material and produced a lot of charged particles three of which were combined with a pion from the continuum to form a fake B_s^0. Therefore it was decided to assume that the main physical background is composed of events that contain a real $D_s^\pm$. A DC06 Monte Carlo sample (see Table A.1 in the appendix) of 12.7 million events of all known decay trees with a $D_s^\pm$ in it were used to evaluate what the background is composed of. Additionally a reflection of the decay $B^0 \to D^- \pi^+$ was examined. A reflection is a decay in which one particle is misidentified and therefore it passes the selection.
The main source for the later separation of signal and background are the different mass distributions. Therefore the background mass distributions will briefly be presented here (see Figure 4.5 and 4.6) and the functions fitted to them will be described in detail in section 6.2.1. The expected number of events of the individual background components are listed in Table 4.4 according to whether the tight mass cut or the wide one was chosen. The numbers in the following descriptions of each component all refer to the wide mass cut.

4.4.1 Inclusive $D_s^\pm$ Background

The background that contains a real Ds meson has a total number of expected events in $2\,\text{fb}^{-1}$ of 198645 ± 4168. This corresponds to a background fraction of $B/(S+B) = 32.6\% \pm 0.6\%$. These numbers are derived by comparing the number of background events in the sample "23263001 - incl_Ds=DecProdCut"to the 704 signal events in it. It consists of B^0 decays which oscillate with the frequency Δm_d into its anti-particle and non oscillating background from B^+ and Λ_b decays.

bkg component	tight mass cut	wide mass cut
Λ_b bkg	24835 ± 1384	88806 ± 2349
B^+ bkg	5731 ± 347	39729 ± 1150
total non oscillating bkg	30567 ± 1661	128535 ± 3143
B^0 bkg from incl. $D_s^\pm$	5731 ± 347	70110 ± 1920
total incl. $D_s^\pm$ bkg	36298 ± 1921	198645 ± 4168
$B^0 \to D^-\pi^+$ reflection	16946 ± 646	26847 ± 813
total background	53244 ± 2567	225492 ± 4981

Table 4.4: *Expected number of events of the different background components according to whether the tight mass cut(50* MeV *around* B_s^0 *mass) or the wide mass cut (*$5000\,\text{MeV} < m < 6000\,\text{MeV}$*) was chosen.*

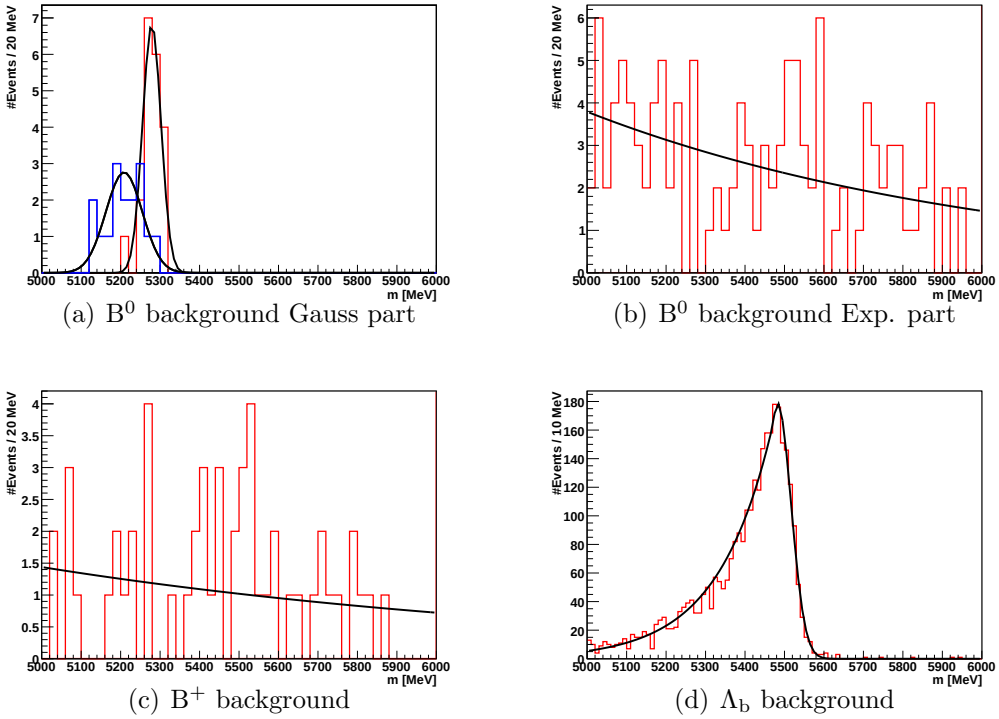
- The Λ_b background which represents 44.7% ± 1.3% of the inclusive $D_s^\pm$ background is composed mostly of a reflection of the decay $\Lambda_b \to D_s^- p^+$ where the proton is misidentified as a pion. Its mass distribution is shown in Fig 4.5(d). To obtain it a sample of 57661 $\Lambda_b \to D_s^- p^+$ decays was examined out of which 3300 Λ_b passed the selection cuts. The function that is used to describe the histogram is an exponential for the left edge and a Gaussian for the right edge.

- The B^+ background contributes 20.0% ± 0.8% to the inclusive $D_s^\pm$ background. There was no special decay identifiable that overweighs in the 68 events that pass the selection. So to model the mass distribution an exponential was assumed (see Fig 4.5(c)). It also has contributions in the mass region above the B_s^0 mass because it is not a reflection of an exclusive decay where only one particle is misidentified, but the real D_s is often combined with a pion from the continuum which makes such a high reconstructed mass possible.

- The B^0 part of this background which represents the remaining 35.3%±1.2% consists of 26% ± 4% completely reconstructed $B^0 \to D_s^+\pi^-$ decays which show a Gaussian peak at the B^0 mass. The decay $B^0 \to D_s^+K^-$ contributes 12% ± 3% to that part of the background and is modelled by a second, broader Gaussian below the B^0 mass (see Figure 4.5(a). The rest of the B^0 part is described by an exponential (see Figure 4.5(b)).

4.4.2 Bd Reflection Background

This background results from the decay $B^0 \to D^-\pi^+$ where the D^- meson decays into $D^- \to K^+\pi^-\pi^-$. If one of the pions is misidentified as a Kaon the D^- mass



(a) B^0 background Gauss part

(b) B^0 background Exp. part

(c) B^+ background

(d) Λ_b background

Figure 4.5: *Mass distributions of inclusive* D_s *background for (a)* $B^0 \to D_s^+ \pi^-$ *(red) and* $B^0 \to D_s^+ K^-$ *(blue) decays, (b) other* B^0 *decays, (c)* B^+ *decays and (d)* Λ_b *decays.*

is shifted under the D_s mass peak making it pass the selection. This results in a peaking background in the B_s^0 mass. The mass distribution is shown in Fig 4.6. It is fitted with a Double Gaussian in which the core Gaussian contributes 95%. Examining a Monte Carlo data sample with 1 million events of this decay leads to an annual yield using formula (4.4) of about 27000 events after the L0 trigger (see Table 6.2) which makes the background fraction $B/(S+B) = 4.2\% \pm 0.1\%$.

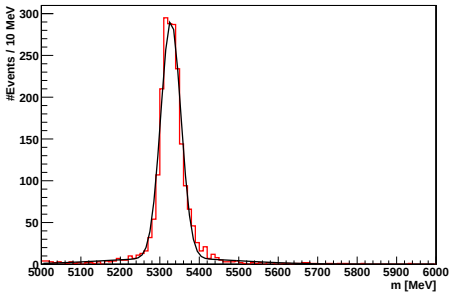


Figure 4.6: *Mass distribution of the* $B^0 \to D^- \pi^+$ *reflection*

Chapter 5

Flavor Tagging

As mentioned in the previous section the knowledge of the production flavor of the B_s^0 is necessary to determine whether it had mixed before it decayed. This is essential information for the mixing analysis because it divides the data sample in two parts (mixed and unmixed) and from these two parts the oscillation frequency Δm_s can be extracted. To find out the flavor at production so-called flavor tagging algorithms are used which will be presented in this section.
The main idea that they take advantage of is that in the collisions at LHC quarks are mainly produced in pairs. So for every b-quark in an event there is a second b-quark of the opposite flavor. This means that there are two ways to determine the production flavor of the reconstructed signal B-meson which has decayed into $D_s^{\pm}\pi^{\mp}$. Either one uses methods to determine that flavor directly on the signal side (same side) or one uses methods to find out the flavor of the partner of the signal b-quark (opposite side). Because of the strong Lorentz boost both the same side and the opposite side particles are close to each other in phase space and can be reconstructed in the LHCb detector. A schematic display of the individual taggers is given in Figure 5.1. They will be introduced in the later parts of this section.

5.1 Tagging Quantities

The tagging algorithms take several quantities like kinematic variables, impact parameter, PID information, etc. as input. Then a neural net is used to evaluate this information and give out a decision. This tagging decision indicates whether the particle is tagged as a B_s^0 (decision = 1), a $\overline{B}_s^0$ (decision = -1) or the tagger cannot give a decision (decision = 0).
In reality these algorithms do not work perfectly, *i.e.* there is a probability that they give back the wrong flavor. This is equivalent to saying that the amplitude of the measured mixing asymmetry A_{mixing} is not equal to one. (see Figure 5.2).

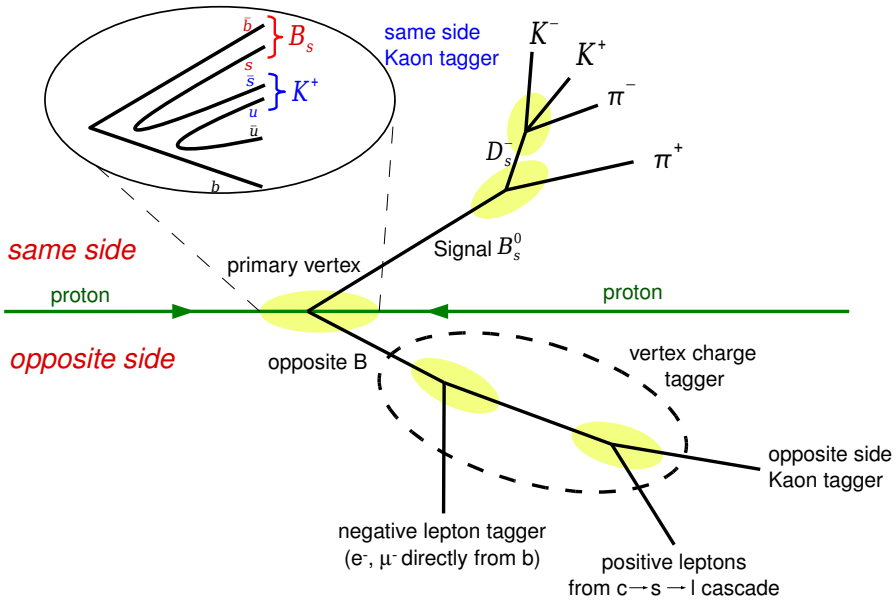


Figure 5.1: *Schematic display of the different sources for the flavor tagging algorithms to determine the production flavor of a B-meson, Picture as in [29]*

A_{mixing} is defined as:

$$A_{mixing}(t) = \frac{N_{unmixed}(t) - N_{mixed}(t)}{N_{unmixed}(t) + N_{mixed}(t)}, \tag{5.1}$$

with $N_{mixed}(t)$ being the number of B_s^0 that decay after a certain proper time t and have mixed before. $N_{unmixed}(t)$ is the number of B_s^0 that do not have mixed before decaying. The two quantities that describe the tagging power of these algorithms are the tagging efficiency ε_{tag} and the mistag probability ω which are defined as:

$$\varepsilon_{\text{tag}} = \frac{N_R + N_W}{N_R + N_W + N_U}, \qquad \omega = \frac{N_W}{N_R + N_W}, \tag{5.2}$$

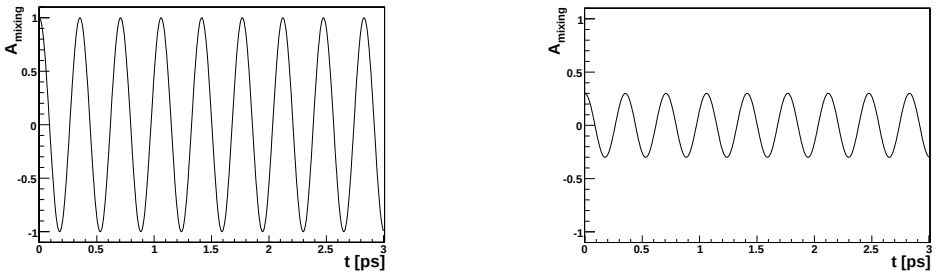
where N_R, N_W and N_U are the number of events that are tagged correctly, tagged incorrectly and untagged (decision = 0) respectively. ε_{tag} and ω can be combined to the effective tagging efficiency:

$$\varepsilon_{\text{eff}} = \varepsilon_{\text{tag}} D^2 = \varepsilon_{\text{tag}}(1 - 2\omega)^2, \tag{5.3}$$

introducing the dilution factor D which is the amplitude of the measured oscillation in Figure 5.2.

$$A_{measured} = D \cdot A_{mixing} \tag{5.4}$$

D is also the quantity that is the target of the calibration in chapter 7.2.



(a) perfect tagging (b) realistic tagging

Figure 5.2: *Mixing asymmetry A_{mixing} for the ideal case of perfect tagging (a) and for realistic tagging (b).*

The effective tagging efficiency ε_{eff} is chosen to describe the tagging power of the algorithms because it represents exactly the fraction with which the total number of signal events has to be multiplied to receive the effective statistics. This means that if you have for example 100000 signal events and your effective tagging efficiency is $\varepsilon_{\text{eff}} = 10\%$ then your statistical power is the same as for 10000 signal events with a perfect tagger. Table 5.1 shows the effective tagging

	ε_{eff} [%]	ε_{tag} [%]	$\overline{\omega}_{\text{true}}$ [%]
e	0.39 ± 0.03	2.95 ± 0.03	31.8 ± 0.5
μ	1.43 ± 0.05	9.10 ± 0.06	30.1 ± 0.3
OS Kaon	1.69 ± 0.06	18.04 ± 0.08	34.7 ± 0.2
Vertex Charge	1.46 ± 0.05	46.88 ± 0.10	41.2 ± 0.1
SS Kaon	3.46 ± 0.08	29.12 ± 0.09	32.8 ± 0.2
combined	6.22 ± 0.10	60.33 ± 0.10	33.9 ± 0.1

Table 5.1: *Tagging performance of the individual tagging algorithms on Monte Carlo data corresponding to 2*fb^{-1} *of* $\text{B}_\text{s}^0 \to \text{D}_\text{s}^- \pi^+$ *decays. Uncertainties are statistical*

efficiency ε_{eff}, the tagging efficiency ε_{tag} and average mistag rate $\overline{\omega}_{\text{true}}$ for the individual tagging algorithms and for the combination of all taggers which will be explained later on. These numbers are taken from the DC06 Monte Carlo data sample of the fully reconstructed decay mode, but they are consistent with those from the $\text{B}_\text{s}^0 \to \text{D}_\text{s}^{*-} \pi^+$ sample. In the following sections the ideas of the individual tagging algorithms will be presented together with a selection of the most important cuts that were applied. A complete list with the exact cuts used by the algorithms can be found in [29]. In this analysis the algorithms were used with these standard configurations only.

5.2 Opposite Side Taggers

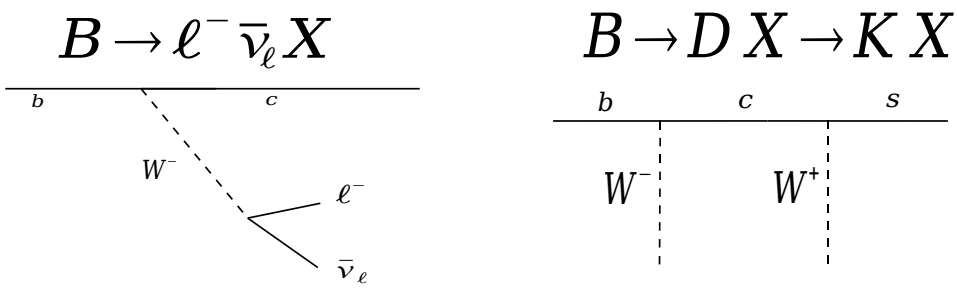
The opposite side taggers use flavor specific decays to determine the flavor of the opposite b-quark which is the aforementioned partner of the signal b. They are further distinguished according to the method or particle they use to identify the flavor of the b.

5.2.1 Lepton Taggers

The Lepton taggers consist of the muon tagger and the electron tagger. Both use the charge of the lepton in semileptonic decays of the opposite B-hadron to tag its flavor. They take advantage from the fact that b-quarks when they are transformed into up type quarks like the charm emit a W^- boson which can decay into a negatively charged lepton (see Figure 5.3(a)). For both taggers a reconstructed momentum of $p > 5\,\mathrm{GeV}/c$ is required.
In case of the muon tagger a transverse momentum of $p_\mathrm{T} > 1.2\,\mathrm{GeV}/c$ is required to reduce the contribution from $\mathrm{b} \to \mathrm{c} \to \ell$ cascade which would give a wrong tag due to the opposite sign of the charge of the muon. Additionally there is a cut on the muon $\mathrm{DLL}(\mu - \pi) > -3$ and for additional separation of muons and pions an algorithm called "Non Shared Hits" (NSH) is run. This algorithm tries to reject pions that are close to a true muon track and may share the same hits with it. Therefore they might be falsely identified as muons and could be picked up by the tagger.
The electron tagger requires a transverse momentum of $p_\mathrm{T} > 1\,\mathrm{GeV}/c$. In addition the ionization charge deposited in the silicon layers of the VeLo is cut on to reduce background contributions coming from photon conversions near the interaction point and misidentified hadrons. The last variable that is used is the ratio E/p of the energy measured in the electro magnetic calorimeter over the particle's momentum. Here a cut of $E/p > 0.8$ is applied. For the separation of electrons from pions a PID cut of $\mathrm{DLL}(e - \pi) > 3$ is applied.

5.2.2 Opposite Side Kaon Tagger

The opposite side kaon tagger uses a charged kaon produced in a $\mathrm{b} \to \mathrm{c} \to \mathrm{s}$ decay chain to determine the flavor of the B-meson (see 5.3(b)). It requires a momentum of $p > 3\,\mathrm{GeV}/c$ and $p_\mathrm{T} > 0.4\,\mathrm{GeV}/c$. Moreover, a cut on the IP significance with respect to the primary vertex of $\mathrm{IP}/\sigma_\mathrm{IP} > 3.5$ is implemented to ensure that the kaon originates from a longlived particle like a B meson. In addition PID cuts of $\mathrm{DLL}(K - \pi) > 3$ and $\mathrm{DLL}(K - p) > -4$ are applied.



(a) lepton tagger (b) opposite side kaon tagger

Figure 5.3: *(a) semileptonic decay of a b-quark that is used by the lepton taggers. (b)* $\mathrm{b} \to c \to s$ *decay chain of a b-quark used by the OS kaon tagger*

5.2.3 Vertex Charge Tagger

The vertex charge tagger uses an inclusive reconstruction of the secondary vertex (the B-hadron decay vertex) to determine the charge of the opposite B-hadron and thereby tag the flavor of the opposite b-quark. The sum of the charges of all tracks originating in this vertex is the charge of the B. A detailed description of how this reconstruction is executed is given in [28] and [29].

5.3 Same Side Tagger

The same side tagger takes advantage of the fact that in the primary interaction most of the quarks are produced in pairs. So if a strange quark forms a $\mathrm{B_s^0}$ an anti-strange quark should be close by to form a hadron. This hadron results in 50% of the cases in a charged kaon[1].
The tagger requires the kaon to have a momentum of $p > 4\,\mathrm{GeV}/c$ and $p_\mathrm{T} > 0.4\,\mathrm{GeV}/c$. Furthermore, an IP significance cut of $\mathrm{IP}/\sigma_\mathrm{IP} < 2.5$ is applied to reject kaons that do not come from the primary vertex. For the same side kaon tagger the same PID cuts as for the opposite side kaon tagger are implemented. To ensure that the kaon and the signal B fly roughly into the same direction there are also cuts on the difference in pseudo-rapidity of $|\Delta\eta| < 1$ and difference in azimuthal angle ϕ of $|\Delta\phi| < 1.1\,\mathrm{rad}$. One also cuts on mass difference between the $\mathrm{B_s^0K}$ system and the reconstructed $\mathrm{B_s^0}$.

[1]Respectively if one analyses a $\mathrm{B^0}$ channel the second d-quark forms a charged pion.

5.4 Combination of Taggers

Since the taggers use different methods to make their decision it is possible that more than one of them give a tagging decision for one B-candidate. In that case the individual decisions are combined to form a global decision. To achieve that the probability of the B-meson containing a b-quark $\mathcal{P}(\mathrm{b})$ is calculated as:

$$\mathcal{P}(\mathrm{b}) = \frac{p(\mathrm{b})}{p(\mathrm{b}) + p(\overline{\mathrm{b}})} \qquad \text{and} \qquad \mathcal{P}(\overline{\mathrm{b}}) = 1 - \mathcal{P}(\mathrm{b}) \tag{5.5}$$

$$p(\mathrm{b}) = \prod_i \left(\frac{1 - q_i}{2} + q_i p_i \right) \qquad \text{and} \qquad p(\overline{\mathrm{b}}) = \prod_i \left(\frac{1 + q_i}{2} - q_i p_i \right). \tag{5.6}$$

$q_i = \pm 1$ is the tagging decision of the i-th tagger and $p_i = 1 - \omega_i$ is the probability that the i-th tagger's decision is correct. Those events with $0.45 < \mathcal{P}(\mathrm{b}) < 0.55$ are declared untagged.
Using this method to combine the output of the tagging algorithms assumes that the taggers are uncorrelated because the single weighted decisions from the individual taggers are simply multiplied. This assumption is only a first approximation because especially the vertex charge tagger is correlated to the other opposite side taggers. That is because if there is a high momentum track that one of the other opposite side taggers used there is a significantly higher possibility that the vertex charge tagger is able to make a decision as well.
A correlation could also exist between the same side tagger and the opposite side taggers because same and opposite side are so close to each other in phase space, the same tracks could be picked up by both taggers. The particle identification system of LHCb described in section 3.2.2 is able to separate kaons from leptons. Therefore the only single particle tagger that could have a correlation with the same side kaon tagger is the opposite side kaon tagger.
To avoid this, different impact parameter cuts were applied for the tracks used in the same and the opposite side tagger. For the opposite side tagger a cut of $\mathrm{IP}\chi^2 > 12.25$ was applied to make sure that the kaons are coming from longlived decay particles like a B-hadron. For the same side taggers the cut is at $\mathrm{IP}\chi^2 < 9$ to take into account mostly the kaons from the hadronization of the primary interaction. So two independent samples of tracks were used to rule out any correlation between the two taggers. For the vertex charge tagger a cut at $\mathrm{IP}\chi^2 > 4$ is applied to reduce correlations with the same side tagger as well.
The tagging power parameters $\varepsilon_{\mathrm{eff}}$, $\varepsilon_{\mathrm{tag}}$ and $\overline{\omega}_{\mathrm{true}}$ for the combination of the taggers are shown in the last row of Table 5.1.

5.5 Tagging Behavior of the Background

To do an analysis in which one has to distinguish between mixed and unmixed B-mesons it is crucial to understand the tagging behavior of the background components. So one must examine what the individual background components are likely to be tagged as.

5.5.1 Λ_b Background

The Λ_b contains a b-quark and decays according to $\Lambda_b \rightarrow D_s^- p^+$. Because of the charge of the D-meson it is always reconstructed as a B_s^0 meson which contains the opposite quark flavor content, *i.e.* a $\overline{b}$-quark. This leads to the fact that a Λ_b is always reconstructed as a B_s^0 and a $\overline{\Lambda}_b$ always as a $\overline{B}_s^0$. That is important for the tagging analysis because it means that all Λ_b should be tagged as mixed B_s^0 if the tagging worked perfectly. Since there is a mistag probability ω some of the Λ_b are tagged as unmixed B_s^0 mesons. The mistag probability for these Λ_b is about 33% which is the same as for the signal B_s^0 which indicates that the tagging algorithms work equally well on Λ_b background and on signal.

5.5.2 B^+ Background

B^+ mesons decay into positively charged D_s and are therefore always reconstructed as $\overline{B}_s^0$. So they should be tagged mainly as a mixed B_s^0. This is true for the opposite side taggers which give a mistag probability of $\omega = 46\% \pm 9\%$. The same side tagger unfortunately reverses that ratio which is understandable if one considers that the B^+ decays mostly into $D_s^+ D^0$ and the D^0 gives a positively charged track next to the negative pion. That track can be interpreted by the sameside kaon tagger as a K^+ which makes it tag the B as a $\overline{B}_s^0$. So there is a high mistag probability given for the sameside kaon tagger if the event is a B^+ event.
Since there was no sufficiently large Monte Carlo data sample of these B^+ decays available this tagging behavior could not be examined in more detail. Instead, the same side tagging behavior of the B^+ background was assumed to be the same as for the signal, but in the description of the proper time distribution for signal and background (see section 6.2) different parameters for ε_{tag} and ω are implemented. Furthermore systematic studies were performed to examine the consequences of a wrong mistag probability on the determination of Δm_s and for the calibration of the same side tagger. These will presented in the sections 7.3.2 and 7.3.1.

5.5.3 B^0 Background

B^0 mesons in the inclusive $D_s^{\pm}$ background all decay into D_s^+ meaning that they are as well reconstructed as B_s^0 mesons of the opposite quark flavor content. On the other hand the B^0 from the $B^0 \to D^-\pi^+$ reflection are all reconstructed as B_s^0 mesons of the same quark flavor content. These two opposing effects have to be monitored because in case of the tight mass cut the contribution from the reflection is larger and for the wide mass cut the contribution from the inclusive $D_s^{\pm}$ background outweighs the reflection. Furthermore, the two components do not necessarily have the same mistag probability. For this analysis the B^0 background is regarded to have the same mistag rate ω as the signal and should be reconstructed as mixed B_s^0 in 50% of the cases. To account for possible biases introduced by wrongly determined tagging behavior of the background, systematic studies as for the B^+ background are presented in sections 7.3.2 and 7.3.1.
An additional difficulty with this background is that B^0 can oscillate into their antiparticles which was also included in the model for the description of this background.

Chapter 6

Unbinned Maximum Likelihood Fitter

A maximum likelihood fitter can be used to extract a set of unknown parameters $\vec{\lambda} = \{\lambda_1, \lambda_2, ...\}$ from a theoretical distribution $f(\vec{\lambda}; \vec{X})$. In this case $\vec{X}$ is the measured data $\vec{X} = \{m, t, q\}$, with the reconstructed mass m, proper time t and the tagging decision q. So the known measured quantities $\vec{X}$ are used to determine the underlying parameters $\vec{\lambda}$. By normalizing $f(\vec{\lambda}; \vec{X})$ as:

$$P(\vec{\lambda}; \vec{X}) = \frac{f(\vec{\lambda}; \vec{X})}{\int f(\vec{\lambda}; \vec{X}')d\vec{X}'} \tag{6.1}$$

one is left with a probability density function (pdf) $P(\vec{\lambda}; \vec{X})$. It gives for a certain set of $\vec{\lambda}$ the probability to measure in a single experiment the data $\vec{X}$ within an interval $d\vec{X}$. The likelihood function $\mathcal{L}$ is the product over the pdf of every event.

$$\mathcal{L} = \prod_{l}^{\text{all events}} P(\vec{\lambda}; \vec{X}_l) \tag{6.2}$$

The likelihood function is a function of the parameter set $\vec{\lambda}$. It gives a measure for the probability for a certain set $\vec{\lambda}$ to measure the data set $\vec{X}_l$. The fitter maximizes this probability by variation of the parameters in such a way that it finds the set $\vec{\lambda}$ for which the probability to measure the given data set $\vec{X}_l$ is maximal. The word 'unbinned' in unbinned maximum likelihood fitter refers to the fact that the data sets $\vec{X}_l$ are not binned in a histogram, but every single event is used individually by the fit. Therefore, there is no loss of information due to binning effects.
In this analysis the likelihood function consists of two independent parts. One part describes the signal and one for the background components. The measured data set $\vec{X}_l$ for each event l contains the reconstructed mass m_l and proper time

t_l as well as the tagging decision q_l.

$$\mathcal{L} = \prod_{l}^{\text{all events}} f_{\text{sig}} P_{\text{sig}}(\vec{\lambda}, m_l, t_l, q_l) + (1 - f_{\text{sig}}) P_{\text{bkg}}(\vec{\lambda}, m_l, t_l, q_l) \tag{6.3}$$

f_{sig} is the fraction of signal events:

$$f_{\text{sig}} = \frac{N_{\text{sig}}}{N_{\text{sig}} + N_{\text{bkg}}} \tag{6.4}$$

with N_{sig} being the number of signal events and N_{bkg} the number of background events, respectively. The pdfs P_{sig} and P_{bkg} are going to be described in detail in the following sections together with their individual parameters $\vec{\lambda}$.

6.1 The Signal PDF

One of the assumptions made when using a maximum likelihood fitter is that the underlying theory distributions, *i.e.* the probability density functions, describe the real distributions correctly. So it is essential for the extraction of the parameters $\vec{\lambda}$ to know these distributions beforehand. For this analysis the distributions are taken from the Monte Carlo simulated data samples described in A.1. The signal pdf P_{sig} consists of a term for the fully reconstructed decay mode $P_{\mathrm{B_s^0 \to D_s^- \pi^+}}$ and one for each partially reconstructed mode$P_{\mathrm{B_s^0 \to D_s^{*-} \pi^+}}$ and $P_{\mathrm{B_s^0 \to D_s^- \rho^+}}$.

$$P_{\text{sig}} = f_{\text{fully}} P_{\mathrm{B_s^0 \to D_s^- \pi^+}} + (1 - f_{\text{fully}})(f_{\mathrm{B_s^0 \to D_s^{*-} \pi^+}} P_{\mathrm{B_s^0 \to D_s^{*-} \pi^+}} + (1 - f_{\mathrm{B_s^0 \to D_s^{*-} \pi^+}}) P_{\mathrm{B_s^0 \to D_s^- \rho^+}}) \tag{6.5}$$

f_{fully} describes the fraction of fully reconstructed signal decays relative to all signal decays. $f_{\mathrm{B_s^0 \to D_s^{*-} \pi^+}}$ is the fraction of $\mathrm{B_s^0 \to D_s^{*-} \pi^+}$ decays within the partially reconstructed decays (see Table 6.1).

6.1.1 PDF for Fully Reconstructed $\mathrm{B_s^0 \to D_s^- \pi^+}$ Decays

Each of the P_{X} terms where X stands for the different signal decay modes is subdivided into 2 parts. This is the case because the fit is performed in two independent parameter spaces. One part describes the distribution of the measured mass ($P_{\text{X}}^m \equiv P_{\text{X}}^m(\vec{\lambda}; m)$) and one describes the proper time distribution ($P_{\text{X}}^t \equiv P_{\text{X}}^t(\vec{\lambda}; t, q)$). The two parts are considered independent so they factorize into:

$$P_{\text{X}}(\vec{\lambda}; m, t, q) = P_{\text{X}}^m \cdot P_{\text{X}}^t \tag{6.6}$$

The full parameter set $\vec{\lambda}_{\text{sig}}$ for the signal part is displayed in Table 6.1 together with their values that were extracted from DC06 Monte Carlo data. The single entries in the Table will be explained in the following sections.

Description	Parameter	Value	Unit
Fractions			
Signal fraction (S/(S+B))	f_{sig}	0.68	
Fraction of fully reconstructed decays	f_{fully}	0.419	
Fraction of $\mathrm{B_s^0} \to \mathrm{D_s^{*-}}\pi^+$	$f_{\mathrm{B_s^0}\to\mathrm{D_s^{*-}}\pi^+}$	0.599	
Detector Resolution parameters			
Width of first mass Gaussian	σ_{m1}	15.0	MeV
Width of second mass Gaussian	σ_{m2}	37.4	MeV
Fraction of first mass Gaussian	f_{σ_m}	0.87	
Width of first proper time Gaussian	σ_{t1}	0.0281	ps
Width of second proper time Gaussian	σ_{t2}	0.0588	ps
Fraction of first proper time Gaussian	f_{σ_t}	0.66	
Physics parameters			
$\mathrm{B_s^0}$mass	$\mathrm{m_{B_s^0}}$	5369.6	MeV
Decay width	$\Gamma_{\mathrm{s}}(=1/\tau)$	0.6845	ps^{-1}
Difference in decay widths	$\Delta\Gamma_{\mathrm{s}}$	0.06845	ps^{-1}
Oscillation Frequency	Δm_{s}	20.00	ps^{-1}
Flavor Tagging parameters			
Tagging efficiency	ε_{tag}	0.573	
Tagging dilution	$D(=1-2\omega)$	0.318	

Table 6.1: *Parameter set $\vec{\lambda}_{sig}$ for the signal probability density function with the parameter values extracted from the DC06 Monte Carlo data sample*

Mass PDF for Fully Reconstructed $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ Decays

The mass distribution P_{X}^m for the fully reconstructed decay mode is in theory described by a narrow Breit-Wigner peak with the width Γ, but it is completely dominated by the detector resolution. So the reconstructed mass is described by a double Gaussian distribution as defined in equation (4.2) (see Figure 4.2). This leads to:

$$P^m_{\mathrm{B_s^0}\to\mathrm{D_s^-}\pi^+} = DG(\mathrm{m_{B_s^0}}, f_{\sigma_m}, \sigma_{m1}, \sigma_{m2}; m)$$

$\mathrm{m_{B_s^0}}$, σ_{m1}, σ_{m2} and f_{σ_m} are part of the parameter set $\vec{\lambda}_{\text{sig}}$ (see Table 6.1). The distribution of the reconstructed mass taken from the Monte Carlo data sample is displayed in Figure 6.1 together with the fitted values for parameters of the double Gaussian. The contribution from the first Gaussian f_{σ_m} and the different widths of the two Gaussian distributions σ_{m1} and σ_{m2} describe the detector resolution. $\mathrm{m_{B_s^0}}$ is the true mass of the $\mathrm{B_s^0}$ meson.

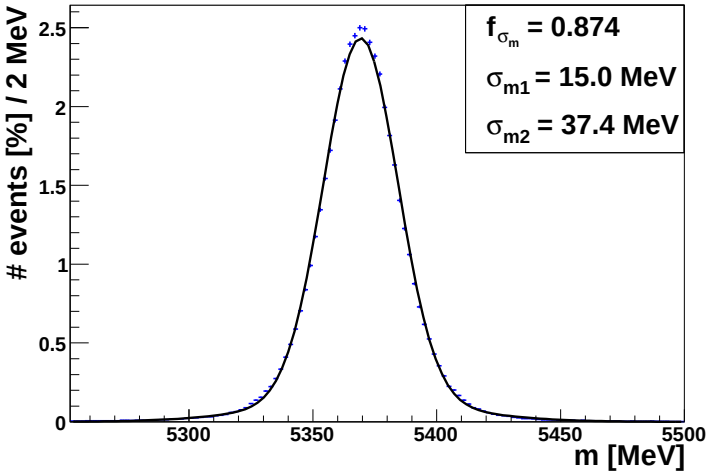


Figure 6.1: *Mass template for fully reconstructed* $\mathrm{B_s^0} \rightarrow \mathrm{D_s^-}\pi^+$ *decays*

Proper Time PDF for Fully Reconstructed $\mathrm{B_s^0} \rightarrow \mathrm{D_s^-}\pi^+$ Decays

In principle, the proper time pdf P_{X}^t is modelled as the sum of an exponential term P_{exp} and a term that describes the oscillation P_{cos}.

$$P_{\mathrm{theo}}^t = P_{\mathrm{exp}} \pm P_{\mathrm{cos}} \tag{6.7}$$

$$P_{\mathrm{exp}} = e^{-\Gamma_{\mathrm{s}} t} \cdot \cosh\left(\frac{\Delta\Gamma_{\mathrm{s}}}{2} t\right) \qquad \text{and} \tag{6.8}$$

$$P_{\mathrm{cos}} = e^{-\Gamma_{\mathrm{s}} t} \cdot \cos(\Delta m_{\mathrm{s}} t). \tag{6.9}$$

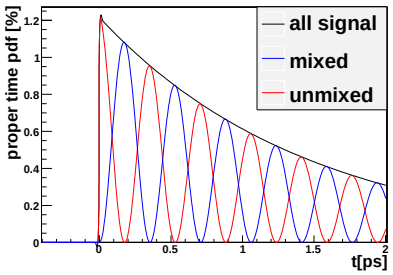
To account for the finite tagging power of the algorithms ($\varepsilon_{\mathrm{tag}}$, ω) P_{theo}^t becomes:

$$P_{\mathrm{theo}}^t = \begin{cases} \frac{1}{2}\varepsilon_{\mathrm{tag}}(P_{\mathrm{exp}} - q(1-2\omega)P_{\mathrm{cos}}) & (q = \pm 1), \\ (1-\varepsilon_{\mathrm{tag}})P_{\mathrm{exp}} & (q = 0). \end{cases} \tag{6.10}$$

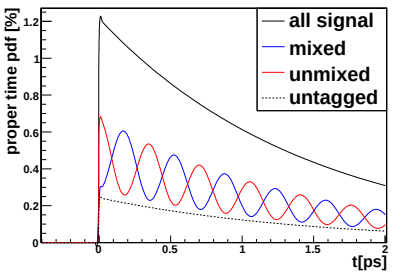
with the tagging decision $q = +1$ for mixed events and $q = -1$ for unmixed events. This theoretical distribution is further convoluted with a double Gaussian (defined in equation (4.2)) to account for the detector resolution and in the end multiplied with the proper time acceptance function. This leads to the final proper time pdf:

$$P_{\mathrm{fully}}^t = \frac{1}{\mathcal{N}}[P_{\mathrm{theo}}^{t'} \otimes DG(t', f_{\sigma_t}, \sigma_{t1}, \sigma_{t2}; t)] \cdot \varepsilon_t(t) \tag{6.11}$$

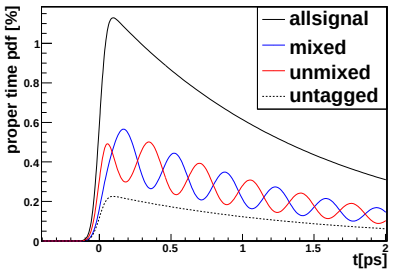
The normalization $\mathcal{N}$ has to be performed only for the exponential part P_{exp} of the pdf. The reason for that is that only the sum over all events has to be normalized and it is not expected that there is a difference in numbers between mixed and unmixed events so the $P_{\cos}$ terms should cancel each other. Its exact form is discussed in the appendix B. The proper time acceptance function $\varepsilon_t(t)$ will be described in the following section.
The parameters f_{σ_t}, σ_{t1} and σ_{t2} in this pdf are for the detector resolution, ε_{tag} and ω are for the tagging performance (see Table 6.1). The physics parameters in the proper time pdf are Γ_s (or respectively the lifetime $\tau = \frac{1}{\Gamma_s}$), $\Delta\Gamma_s$ and Δm_s. In Figure 6.2 are the individual stages of the proper time pdf shown. Figure 6.2(a) shows the pure theory distribution from equation (6.7). Figure 6.2(b) shows the effect that the finite tagging power has on the pdf (equivalent to equation (6.10)). In Figures 6.2(c) and 6.2(d) are the effects of the convolution for the detector resolution and the proper time acceptance function $\varepsilon_t(t)$ displayed (final stage of equation (6.11)). It can clearly be seen that the proper time resolution has a direct influence on the amplitude of the oscillation. Therefore, if the mistag



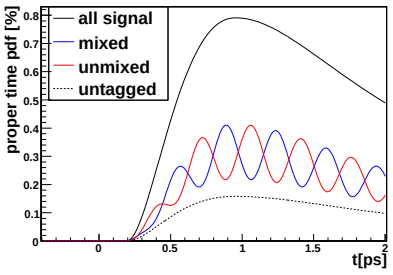
(a) proper time pdf



(b) with tagging effects



(c) with resolution effects



(d) with acceptance

Figure 6.2: *Proper time probability density function for signal events. Model of the pure theory distribution (a), tagging effects included (b), proper time resolution effects included (c) and proper time acceptance (see section 6.1.2) included (d)*

probability is fitted (see section 7.2) one has to know the proper time resolution very accurately.

6.1.2 Proper Time Acceptance

The selection used for this analysis which was presented in section 4.2 is modifying the proper time distribution. The IPχ^2 cut on the pions and kaons and the FSPV cuts are the main cause for this. This leads to the fact that B_s^0 that live longer have a higher probability to be reconstructed than the very shortlived ones. Accordingly, the proper time acceptance function is defined as:

$$\varepsilon_t(t) = \frac{S_{\text{reco}}(t)}{S_{\text{theo}}(t)}, \tag{6.12}$$

where S_{reco} describes the distribution of the reconstructed proper time of the B_s^0 and S_{theo} is the theory distribution:

$$S_{\text{theo}} = \left(e^{-\Gamma \cdot t} \cdot \cosh \frac{\Delta\Gamma}{2} t \right) \otimes G(t, \sigma_{t,i}), \tag{6.13}$$

which takes into account the difference in decay widths $\Delta\Gamma$ and convolutes the resulting exponential with a Gaussian to simulate the detector resolution. This is done so that the acceptance is a function of the reconstructed proper time and not the true one which makes it easier to include it into the proper time pdf. The subscript i of $\sigma_{t,i}$ indicates that it is an event-per-event property. That means that for every event the corresponding proper time error $(t_{\text{reco}} - t_{\text{true}})$ is taken as width for the Gaussian $\sigma_{t,i}$ to provide a simulation which is as realistic as possible. The result is shown in Figure 6.3.
Here the fitted function is:

$$\varepsilon_{ct}(t) = \sum_{i=1}^{3} a_i \cdot (t - \zeta_i)^2 \cdot e^{-\frac{t}{\tau_i}} \cdot \Theta(t - \zeta_i), \tag{6.14}$$

where $\Theta(t-\zeta_i)$ is the Heaviside step function. ε_{ct} has nine free parameter (a_i, ζ_i, τ_i, with $i = 1, 2, 3$). It is used because with that choice the proper time pdf in the fitter is analytically integrable (see [26] for more information). Here it is fitted in the region from $t = 0$ ps to $t = 10$ ps because in the proper time region above there is no more information given in the histogram. It is plotted until $t = 20$ ps to show that it does not diverge.

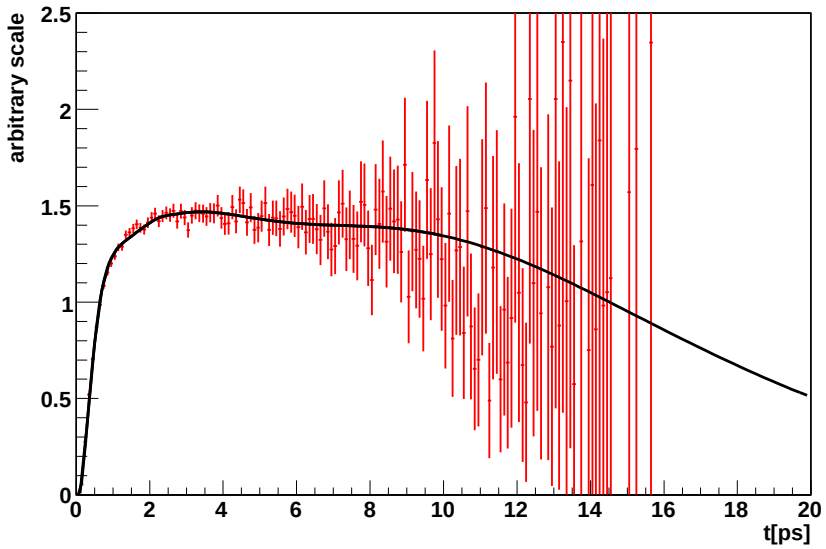
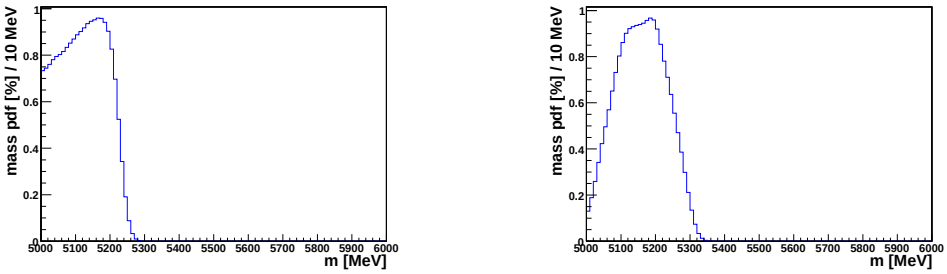


Figure 6.3: *Proper time acceptance fitted with an analytically integrable function taken from [26]*

6.1.3 PDF for Partially Reconstructed $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$ Decays

Mass PDF for Partially Reconstructed Decays

For the mass probability density function of the partially reconstructed $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$ decays the histograms in Figure 6.4(a) and 6.4(b) are normalized to 1 and read in. These histograms were obtained from the Monte Carlo data samples described in A.1.
The shift towards lower masses is due to the missing neutral particle in the reconstruction. Both decays are reconstructed as $B_s^0 \to D_s^-\pi^+$ decays and in $B_s^0 \to D_s^{*-}(D_s^-\gamma)\pi^+$ a photon is missed and in $B_s^0 \to D_s^-\rho^+(\pi^0\pi^+)$ a π^0 is not reconstructed. Therefore the invariant mass of the $D_s^\pm\pi^\mp$ system is too low and the reconstructed mass of the B_s^0 is shifted.



(a) $B_s^0 \to D_s^- \rho^+$ (b) $B_s^0 \to D_s^{*-} \pi^+$

Figure 6.4: *Mass templates for partially reconstructed* $B_s^0 \to D_s^{*-}\pi^+$ *(a) and* $B_s^0 \to D_s^- \rho^+$ *(b)*

Proper Time PDF for Partially Reconstructed Decays

Due to the missing particle in the reconstruction the momentum of the B_s^0 is reconstructed too low as well. For the proper time pdf of the fitter it is important to correct the reconstructed momentum of the B_s^0 for the missing momentum of the undetected particle because the proper time is inversely proportional to the momentum. The problem is that in the unbinned maximum likelihood fitter every single event is examined and it is of course unknown what momentum the missing particle had. So for every event the distribution of the so called k-factor, which is defined in (6.15) as

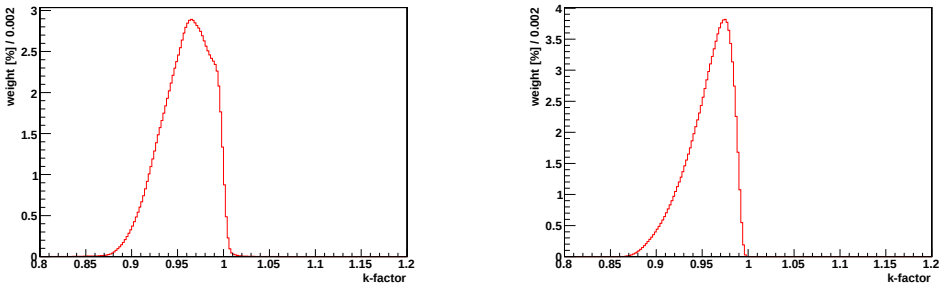
$$k = \frac{p^{rec}_{B_s^0}}{p^{true}_{B_s^0}}, \tag{6.15}$$

is convoluted into the proper time pdf. The k-factor distributions $\mathcal{F}(k)$ are given in Figure 6.5(a) for $B_s^0 \to D_s^{*-}\pi^+$ and in 6.5(b) for $B_s^0 \to D_s^- \rho^+$.

The reason that for the $B_s^0 \to D_s^{*-}\pi^+$ mode a reconstructed momentum larger than the true one is possible are resolution effects and that unlike in case of the $B_s^0 \to D_s^- \rho^+$ the missing neutral particle (γ) has no mass.
Otherwise the same proper time pdf was used as in case of the fully reconstructed decay mode. So the full pdf is given as:

$$P^t_{\text{partially}} = \frac{1}{\mathcal{N}}[P^{t'}_{\text{theo}} \otimes \mathcal{F}(k) \otimes DG(t', f_{\sigma_t}, \sigma_{t1}\sigma_{t2}; t)] \cdot \varepsilon_t(t) \tag{6.16}$$

Also, the same proper time acceptance function $\varepsilon_t(t)$ was used as for the fully reconstructed decay mode. The explicit form of the proper time pdf and the normalization factor $\mathcal{N}$ are described in detail in the appendix B.



(a) $B_s^0 \to D_s^{*-}\pi^+$ (b) $B_s^0 \to D_s^-\rho^+$

Figure 6.5: *k-factor distributions for the partially reconstructed decay modes* $B_s^0 \to D_s^{*-}\pi^+$ *(a) and* $B_s^0 \to D_s^-\rho^+$ *(b)*

6.2 Background PDF

The different background contributions were introduced in section 4.4. The structure for the background probability density function P_{bkg} is similar to P_{sig}. There are also different parts for the individual background components weighted with their respective fractions like the fully and partially reconstructed decay modes in the signal case.

$$\begin{aligned} P_{\text{bkg}} =& f_{\text{incl.D}_s^\pm}(f_{\text{non osc}}P_{\text{non osc}} + (1 - f_{\text{non osc}})P_{(\text{incl.D}_s^\pm)\text{B}^0}) \\ &+ (1 - f_{\text{incl.D}_s^\pm})P_{\text{B}^0\text{refl}} \end{aligned} \tag{6.17}$$

The different P_{X} have the same two components for mass and proper time as described in equation (6.6). Here the X stands for the different background contributions. The relative fractions of the background components are defined as:

$$f_{\text{bkg}} = \frac{N_{\text{bkg}}}{N_{\text{all}}} \quad , \quad f_{\text{incl.D}_s^\pm} = \frac{N_{\text{inclD}_s^\pm}}{N_{\text{bkg}}},$$
$$f_{\text{non osc}} = \frac{N_{\text{nonosc}}}{N_{\text{inclD}_s^\pm}} \quad , \quad f_{\text{B}^+} = \frac{N_{\text{B}^+}}{N_{\text{nonosc}}},$$

with N_{all}, N_{bkg}, $N_{\text{inclD}_s^\pm}$, N_{nonosc} and N_{B^+} being the number of all events, background events, inclusive Ds background events, non oscillating background events and B^+ background events respectively.
The different values for these fractions according to whether the tight mass cut or the wide one was used is given in Table 6.2. These numbers are obtained from the same DC06 Monte Carlo data samples as the expected annual yields in Table 4.4.
In the framework of this diploma thesis a combinatorical background has not

Fractions	tight mass cut	wide mass cut
f_{bkg}	0.24±0.02	0.37±0.01
$f_{\text{incl.D}_s^\pm}$	0.68±0.03	0.89±0.01
$f_{\text{non osc}}$	0.84±0.02	0.65±0.01
f_{B^+}	0.18±0.02	0.31±0.01

Table 6.2: *Fractions of the different background components according to whether the tight mass cut(50* MeV *around* B_s^0 *mass) or the wide mass cut (*5000 MeV $<$ m $<$ 6000 MeV*) was chosen.*

yet been included in the description of the background. However for the analysis on real data this background will be extracted from the mass side bands. For the flavor tagging there is no special behavior expected for this background ($\omega = 50\%$).

6.2.1 Mass PDF for the Background

The mass distributions of the individual background contributions have been presented in section 4.4. The probability density functions which describe the functions fitted to the histograms in Figures 4.5 and 4.6 will be introduced in this section.
For the non oscillating part the mass pdf is defined as:

$$P^m_{\text{non osc}} = f_{\text{B}^+}\frac{P_{\text{B}^+}}{\mathcal{N}_{\text{B}^+}} + (1 - f_{\text{B}^+})\frac{P_{\Lambda_\text{b}}}{\mathcal{N}_{\Lambda_\text{b}}} \tag{6.18}$$

with the B^+ contribution

$$P_{\text{B}^+} = e^{-s_{\text{B}^+}\cdot m + c_{\text{B}^+}} \tag{6.19}$$

and the Λ_b contribution

$$P_{\Lambda_\text{b}} = \begin{cases} e^{s_{\Lambda_\text{b}}\cdot m + c_{\Lambda_\text{b}}} & (\text{for m } < 5482), \\ e^{-\frac{(m-\mu_{\Lambda_\text{b}})^2}{2\sigma^2_{m,\Lambda_\text{b}}}} & (\text{for m } > 5482). \end{cases} \tag{6.20}$$

Since the mass is fitted in a finite window, the normalizations ($\mathcal{N}_X$) are performed by using the integral over a Gaussian distribution within finite borders:

$$\int_{\text{min}}^{\text{max}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{2}\left[\text{Erf}\left(\frac{\text{max}-\mu}{\sqrt{2}\sigma}\right) - \text{Erf}\left(\frac{\text{min}-\mu}{\sqrt{2}\sigma}\right)\right]\cdot\sqrt{2\pi}\sigma \tag{6.21}$$

In case of the B^0 part of the inclusive $\text{D}_s^\pm$ background the pdf is:

$$P^m_{(\text{incl.D}_s^\pm)\text{B}^0} = P^m_{\text{B}^0\to\text{D}_s^+\pi^-} + P^m_{\text{B}^0\to\text{D}_s^+\text{K}^-} + P^m_{\text{B}^0,\text{exp}} \tag{6.22}$$

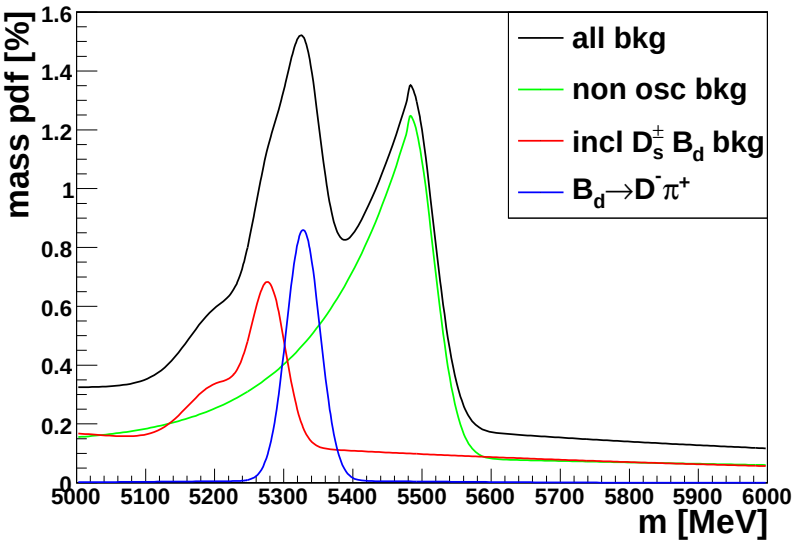


Figure 6.6: *Background mass pdf with non oscillating part (green),* B^0 *from inclusive* $D_s^{\pm}$ *background (red) and* $B^0 \to D^- \pi^+$ *reflection (blue). Total background pdf black continuous line*

with a Gaussian distribution for $P^m_{B^0 \to D_s^+ \pi^-}$ around the B^0 mass and with the width $\sigma_{B^0 \to D_s^+ \pi^-}$. $P^m_{B^0 \to D_s^+ K^-}$ is modelled with a Gaussian distribution as well with mean $\mu_{B^0 \to D_s^+ K^-}$ and width $\sigma_{B^0 \to D_s^+ \pi^-}$.

$$P^m_{B^0 \to D_s^+ \pi^-} = G(m_{B^0}, \sigma_{B^0 \to D_s^+ \pi^-}; m) \tag{6.23}$$

and

$$P^m_{B^0 \to D_s^+ K^-} = G(\mu_{B^0 \to D_s^+ K^-}, \sigma_{B^0 \to D_s^+ \pi^-}; m). \tag{6.24}$$

The rest of the B^0 part of the inclusive $D_s^{\pm}$ background is modelled with a simple exponential

$$P^m_{B^0,\mathrm{exp}} = e^{s_{B^0} \cdot m + c_{B^0}}. \tag{6.25}$$

The normalizations are also performed using (6.21).
The last part of the background is the $B^0 \to D^- \pi^+$ reflection. Its mass is described by a Double Gaussian distribution (as defined in (4.2)).

$$P^m_{B^0 \mathrm{refl}} = DG(\mu_{B^0 \mathrm{refl}}, f_{\sigma_m}^{B^0 \mathrm{refl}}, \sigma_{m1}^{B^0 \mathrm{refl}}, \sigma_{m2}^{B^0 \mathrm{refl}}; m) \tag{6.26}$$

In Figure 6.2.1 the individual components of the background mass pdf are displayed.

6.2.2 Proper Time PDF for the Background

In the proper time pdfs one only distinguishes between non-oscillating and B^0 background. The non-oscillating background is modelled by a simple exponential that is convoluted with a Double Gaussian. For reasons of simplicity the proper time resolution parameters for all background components are assumed to be the same as for the signal[1]. Furthermore the proper time acceptance is also assumed to be the same as in the signal case.
As described in section 6.1 for the proper time fit the sample is subdivided into three parts according to whether the event was tagged as mixed ($q = 1$), unmixed ($q = -1$) or untagged ($q = 0$). This has to be done for the background as well. Therefore the tagging behavior of the background was modelled according to the findings presented in section 5.5.
The non oscillating part of the background is modelled as a simple exponential. It should always be tagged as mixed and the same parameters ω and ε_{tag} were used as in the signal case. This leads to a theoretical distribution similar to the signal case:

$$P^t_{\text{theo, non osc}} = \begin{cases} \frac{1}{2}\varepsilon_{\text{tag}} \cdot \left(1 - q(1-2\omega)e^{-\Gamma_{\text{non osc}}t}\right) & (q = \pm 1), \\ (1-\varepsilon_{\text{tag}})e^{-\Gamma_{\text{non osc}}t} & (q = 0). \end{cases} \tag{6.27}$$

Adding the detector resolution effects and proper time acceptance $\varepsilon_t(t)$ one obtains

$$P^t_{\text{non osc}} = \frac{P^{t'}_{\text{theo, non osc}} \otimes DG(t', f^{\text{non osc}}_{\sigma_m}, \sigma^{\text{non osc}}_{m1}, \sigma^{\text{non osc}}_{m2}; t)}{\mathcal{N}^t_{\text{non osc}}}\, \varepsilon_t(t). \tag{6.28}$$

The slowly oscillating B^0 background is modelled exactly like the fully reconstructed signal part except for the difference in decay widths $\Delta\Gamma$ which is negligible in the B_d system and therefore the cosh term is approximated to 1. This leads to theoretical distribution for this background pdf of

$$P^t_{\text{theo,B}^0} = \begin{cases} \frac{1}{2}\varepsilon_{\text{tag}} e^{-\Gamma_{\text{d}}t}\left(1 - (2f^t_{\text{B}^0} - 1)q(1-2\omega)\cos(\Delta m_{\text{d}} t)\right) & (q = \pm 1), \\ (1-\varepsilon_{\text{tag}})P^{\text{B}^0}_{\text{exp}} & (q = 0). \end{cases} \tag{6.29}$$

Again after adding the detector resolution effects and proper time acceptance $\varepsilon_t(t)$ one is left with

$$P^t_{\text{B}^0} = \frac{1}{\mathcal{N}_{t,\text{B}^0}}[P^t_{\text{theo,B}^0} \otimes DG(t', f^{\text{B}^0}_{\sigma_m}, \sigma^{\text{B}^0}_{m1}, \sigma^{\text{B}^0}_{m2}; t)] \cdot \varepsilon_t(t).$$

[1] This is not self-evident since the background can have different distributions in the kinematic variables which influence the proper time resolution significantly. But due to lack of enough statistics for the background this choice for the proper time resolution is made, but the option of a different proper time resolution is also available

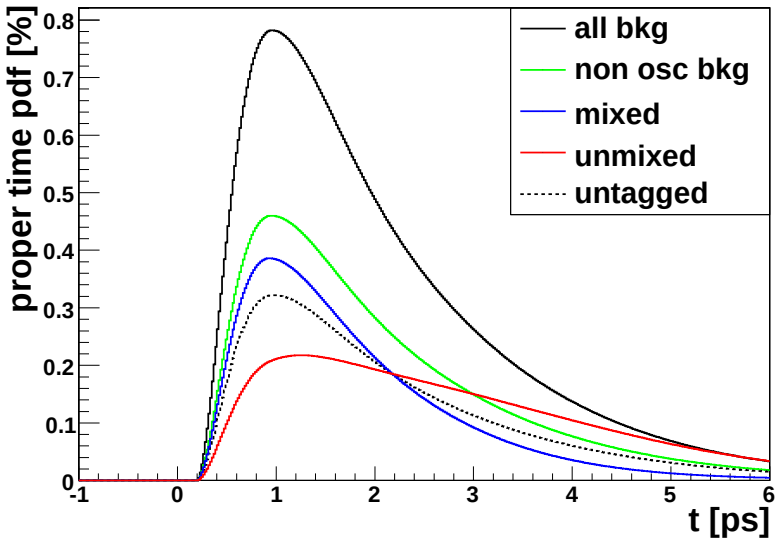


Figure 6.7: *Proper time pdf for the background with non oscillating part (green),* B^0 *reconstructed as mixed (blue), unmixed (red) or untagged (dashed). Total background pdf black continuous line*

The normalization is performed in the same way as for the signal pdf (see appendix B).
For the tagging behavior described in section 5.5 one has to distinguish between B^0 coming from inclusive $D_s^{\pm}$ background and those coming from the $B^0 \to D^-\pi^+$ reflection. The former should all be tagged as mixed and the latter as unmixed. Both use the identical proper time pdf, but with their own individual parameters for tagging ($\varepsilon_{\text{tag}}^{B^0 \to D^-\pi^+}$, $\varepsilon_{\text{tag}}^{(\text{incl.}D_s^{\pm})B^0}$, $\omega_{B^0 \to D^-\pi^+}$, $\omega_{\text{incl.}D_s^{\pm}B^0}$).
In Tables 6.3 and 6.4 is a complete list of the parameters for the background pdf listed. The values of these parameters are again from fits to the according distributions in the DC06 Monte Carlo data samples (see section 4.4. The background pdfs are not used to extract any of the parameters but to provide a better signal to background separation.

In Figure 6.7 the proper time pdf for the different background components is displayed. For this display only the B^0 contribution from the $B^0 \to D^-\pi^+$ reflection were plotted the values for ε_{tag}, ω were set to $\varepsilon_{\text{tag}} = 0.8, \omega = 0.2$ because otherwise all lines would have been on top of each other.

Description	Parameter	Value	Unit
Fractions			
Fraction of inclusive $D_s^{\pm}$ bkg	$f_{\mathrm{incl.}D_s^{\pm}}$	0.89	
Fraction of non oscillating bkg	$f_{\mathrm{non\ osc}}$	0.65	
Fraction of B^+ part in non oscillating bkg	f_{B^+}	0.30	
B^0 **background parameters**			
Mean of the Double Gaussian for B^0 refl.	$\mu_m^{B^0\mathrm{refl}}$	5328	MeV
Width of first mass Gaussian for B^0 refl.	$\sigma_{m1}^{B^0\mathrm{refl}}$	24.3	MeV
Width of second mass Gaussian for B^0 refl.	$\sigma_{m2}^{B^0\mathrm{refl}}$	204.8	MeV
Fraction of first mass Gaussian for B^0 refl.	$f_{\sigma_m}^{B^0\mathrm{refl}}$	0.74	
Width of $B^0 \to D_s^+\pi^-$ Gaussian	$\sigma_{B^0\to D_s^+\pi^-}$	24.1	MeV
Mean of $B^0 \to D_s^+\pi^-$ Gaussian (B^0 mass)	m_{B^0}	5279	MeV
Width of $B^0 \to D_s^+K^-$ Gaussian	$\sigma_{B^0\to D_s^+\pi^-}$	46.7	MeV
Mean of $B^0 \to D_s^+K^-$ Gaussian	$\mu_{B^0\to D_s^+K^-}$	5209	MeV
Slope of exp of B^0 part of inclusive $D_s^{\pm}$ bkg	$s_{\mathrm{incl.}D_s^{\pm}B^0}$	0.00109	MeV
Offset of exp of B^0 part of inclusive $D_s^{\pm}$ bkg	$c_{\mathrm{incl.}D_s^{\pm}B^0}$	6.273	MeV
Width of first proper time Gaussian	$\sigma_{m1}^{B^0}$	0.0281	ps
Width of second proper time Gaussian	$\sigma_{m2}^{B^0}$	0.0588	ps
Fraction of first proper time Gaussian	$f_{\sigma_m}^{B^0}$	0.66	
Non oscillating background parameters			
Slope of exponential of B^+ background	s_{B^+}	0.000689	MeV
Offset of exponential of B^+ background	c_{B^+}	3.808	MeV
Slope of exponential of Λ_b background	s_{Λ_b}	-0.00724	MeV
Offset of exponential of Λ_b background	c_{Λ_b}	-34.3	MeV
Width of Λ_b mass Gaussian	σ_{m,Λ_b}	34.3	MeV
Mean of Λ_b mass Gaussian	μ_{Λ_b}	5482	MeV
Width of first proper time Gaussian	$\sigma_{m1}^{\mathrm{non\ osc}}$	0.0281	ps
Width of second proper time Gaussian	$\sigma_{m2}^{\mathrm{non\ osc}}$	0.0588	ps
Fraction of first proper time Gaussian	$f_{\sigma_m}^{\mathrm{non\ osc}}$	0.66	

Table 6.3: *Parameter set $\vec{\lambda}_{bkg}$ for the background probability density function and the values used in this study (continued on the next page)*

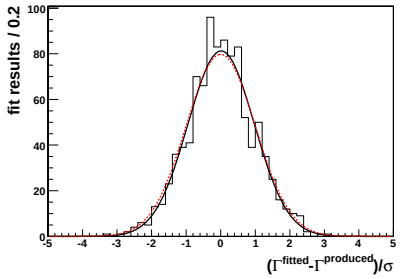
Description	Parameter	Value	Unit
Physics parameters			
Non oscillating background decay width	$\Gamma_{\text{non osc}}$	0.690	ps^{-1}
B^0 decay width	Γ_{d}	0.6510	ps^{-1}
B^0 oscillation frequency	Δm_{d}	0.5	ps^{-1}
Flavor Tagging parameters			
Tagging efficiency non oscillating	$\varepsilon_{\text{tag}}^{\text{non osc}}$	0.573	
Tagging dilution non oscillating	$D_{\text{non osc}}$	0.318	
Tagging efficiency incl. $\text{D}_\text{s}^\pm\text{B}^0$ background	$\varepsilon_{\text{tag}}^{(\text{incl.D}_\text{s}^\pm)\text{B}^0}$	0.573	
Tagging dilution incl. $\text{D}_\text{s}^\pm\text{B}^0$ background	$D_{\text{incl.D}_\text{s}^\pm\text{B}^0}$	0.318	
Tagging efficiency $\text{B}^0 \to \text{D}^-\pi^+$ reflection	$\varepsilon_{\text{tag}}^{\text{B}^0\to\text{D}^-\pi^+}$	0.573	
Tagging dilution $\text{B}^0 \to \text{D}^-\pi^+$ reflection	$D_{\text{B}^0\to\text{D}^-\pi^+}$	0.318	

Table 6.4: *(Continued from previous page) Parameter set $\vec{\lambda}_{bkg}$ for the background probability density function and the values used in this study*

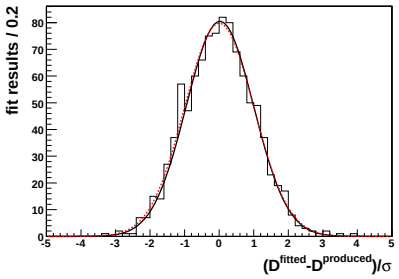
6.3 Validation of the Fit Procedure

To validate whether the fitter works correctly, a procedure called a Monte Carlo toy study is performed. In contrast to the full DC06 Monte Carlo simulation described in the beginning of chapter 4, in a toystudy not the proton-proton collision and the detector reaction to the particles is simulated, but simply the measured quantities. The probability density functions presented above are used to generate the data variables mass m, proper time t and tagging decision q for a single event. This procedure is repeated various times for the simulation of many events. Then the unbinned maximum likelihood fitter extracts the parameter set $\vec{\lambda}$ from these data sets $\vec{X} = \{m, t, q\}$.
Now the performance of the fitter under these conditions should be ideal because there is no systematic bias since the same distributions were used to generate the data sets and to fit on them. So if the fitter works properly and the pdfs are correctly normalized the extracted parameters should be distributed according to a Gaussian distribution around the input values. Furthermore, the fitter gives an error estimate on the fitted parameter. If that is correct as well, it should be identical to the width of the Gaussian mentioned above. These toy experiments are then repeated various times (runs) and the so called pull distributions of the parameters are filled into histograms (see Figure 6.8). The pull of a parameter λ with the error estimate σ_λ is defined as:

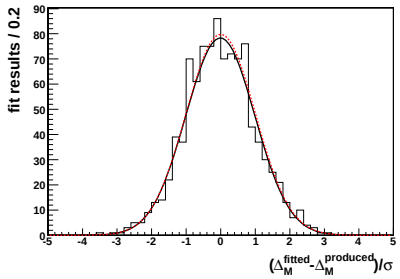
$$pull_\lambda = \frac{\lambda_{\text{fitted}} - \lambda_{\text{input}}}{\sigma_\lambda} \tag{6.30}$$



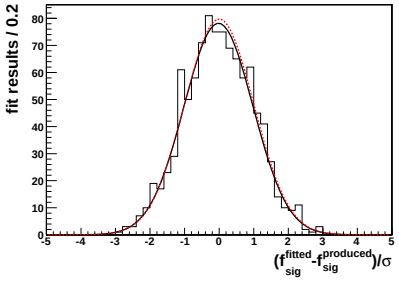
(a) Decay Width Γ_s



(b) Dilution ($D = 1 - 2\omega$)



(c) Oscillation Frequency Δm_s



(d) Signal Fraction f_{sig}

Figure 6.8: *Pull distributions as defined in* (6.30) *for parameters* Γ_s*(a),* D*(b),* Δm_s*(c) and* f_{sig}*(d). Dotted red line is Gaussian with mean* $\mu = 0$ *and width* $\sigma = 1$*. Black line is Gaussian fitted to pull histogram.*

The expected distribution of the $pull_\lambda$ is a Gaussian around $\mu = 0$ with the width $\sigma = 1$. In Figure 6.8, this is represented by the dotted red line. The black line is the Gaussian fitted to the pull histogram. It can be seen that the two Gaussian distributions agree for all displayed parameters. In Table 6.5, the fitted values for the mean and width of the pull distributions are displayed. These toy studies were performed for all parameters. In Figure 6.8 are those displayed which are interesting for the analysis. For these plots the wide mass cut was used with the

Parameter	χ^2/ndf	Mean	Width
Γ_s	0.914	0.013 ± 0.032	0.960 ± 0.025
D	0.644	0.023 ± 0.032	0.974 ± 0.022
Δm_s	0.711	-0.011 ± 0.032	0.998 ± 0.026
f_{sig}	0.827	0.019 ± 0.033	1.000 ± 0.024

Table 6.5: *Results from the fits of the pull distributions shown in Figure 6.8.*

respective fractions of the signal and background components. 94000 events per toy experiment were generated which corresponds to $0.2\,\mathrm{fb}^{-1}$ and 1000 runs were performed.
It can be seen that the distributions are well described by a Gaussian because the reduced χ^2 of the fits are approximately equal to 1. Furthermore the parameters of the fitted Gaussians agree within 2 σ with the expected ones. So in conclusion it can be said that the fitter works properly and the probability density functions are correctly normalized.

Chapter 7

Results

The aim of this study is on the one hand to give an estimate on what the statistical sensitivity to the measurement of Δm_{s} in the decay channel $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ is. On the other hand this study will show how accurately the tagging power of the same side kaon tagger can be determined on data in the decay $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$.
The results of these studies will be presented in this chapter.

7.1 Δm_{s} Analysis in the Channel $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$

In section 2.3.2 the principle of the oscillation of neutral B-mesons into their anti-particles was presented. It was shown that the frequency of this oscillation is given by the mass difference Δm of the mass eigenstates $\mathrm{B_L}$ and $\mathrm{B_H}$. In case of the B_s system this frequency Δm_{s} is very high which makes it hard to measure the quick oscillation precisely. A superb time and spatial resolution is indispensable to make a precision measurement of this oscillation. Both are given in the LHCb experiment and because of the large cross section for the production of b-quarks there will also within a reasonably short time be more statistics available than in the other experiments that have measured Δm_{s} so far.
In this section the planned analysis of Δm_{s} in the channel $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ is presented. First will be shown how sensitive this study is on Δm_{s} using the tight mass cut illustrated in 4.2. The second part is about using the wide mass cut to include partially reconstructed decay modes and thereby enhancing the available statistics.

7.1.1 Δm_{s} in Fully Reconstructed $\mathrm{B_s^0} \to \mathrm{D_s^-}\pi^+$ Decays

The tools of the analysis presented in chapters 5 and 6 are used to determine Δm_{s}. All flavor tagging algorithms (both same side and opposite side) are used to extract the information whether the $\mathrm{B_s^0}$ was mixed or unmixed. To find out how sensitive this analysis is to Δm_{s} a toy study with the full expected statistics

corresponding to $2\,\mathrm{fb}^{-1}$ is performed 10 times. In these toy experiments all parameters shown in Tables 6.1, 6.3 and 6.4 except for Δm_s were fixed. Then the mean of the error on Δm_s that is given out by the unbinned maximum likelihood fitter which was presented in section 6 is taken to estimate the statistical uncertainty on Δm_s after one nominal year of data taking. To get an idea of how wide the values are spread the standard deviation of these 10 results is calculated and given as an estimate for the uncertainty on $\sigma_{\Delta m_\mathrm{s}}$.
Table 6.2 shows the differences in the fraction of the individual signal components, overall signal fraction and total number of signal events depending on whether the tight or the wide mass cut was chosen. In this section the set in the left column for the tight cut is used. Using the signal and background fractions for the tight cut given in Tables 7.1 and 6.2

$$\sigma_{\Delta m_\mathrm{s}} = (0.0069 \pm 0.0001)\,\mathrm{ps}^{-1} \tag{7.1}$$

is obtained for the sensitivity on Δm_s after one nominal year of data taking at LHCb.

Parameter	tight mass cut	wide mass cut
f_fully	1.0	0.419
f_sig	0.76	0.68
# signal events	250000	610000

Table 7.1: *Fraction of the signal components f_{fully}, overall signal fraction f_{sig} and total number of signal events according to whether the tight mass cut(50* MeV *around* B_s^0 *mass) or the wide mass cut (*$5000\,\mathrm{MeV} < \mathrm{m} < 6000\,\mathrm{MeV}$*) was chosen.*

7.1.2 Δm_s in Fully and Partially Reconstructed Decays

For the inclusion of the partially reconstructed decay modes $\mathrm{B}_\mathrm{s}^0 \to \mathrm{D}_\mathrm{s}^{*-}\pi^+$ and $\mathrm{B}_\mathrm{s}^0 \to \mathrm{D}_\mathrm{s}^-\rho^+$ the mass window is widened to $5000\,\mathrm{MeV} < \mathrm{m} < 6000\,\mathrm{MeV}$. Thereby the statistic is enhanced by a factor 2.4 (see Table 7.1). Furthermore the partially reconstructed decays are missing a particle and therefore the reconstructed momentum of the B_s^0 is too low. This is somehow compensated by convoluting the k-factor into the proper time pdf (see section 6.1.3). However the momentum resolution is worsened by that and since the proper time depends directly on the momentum (see (4.1)) the sensitivity to Δm_s of the partially reconstructed decay modes are expected to be not as good as in case of the fully reconstructed.
Like in the previous section a toystudy with the statistics corresponding to 2 fb^{-1} is performed to obtain

$$\sigma_{\Delta m_\mathrm{s}} = (0.0062 \pm 0.0002)\,\mathrm{ps}^{-1}. \tag{7.2}$$

This gain corresponds to a factor of 1.2 in statistics.
The CDF collaboration has performed the most precise measurement of Δm_s so far [1]. Their result was

$$\Delta m_s = (17.77 \pm (0.10) \pm (0.07))\,\mathrm{ps}^{-1}. \tag{7.3}$$

So after the duration of one year of nominal runtime at LHCb the statistical sensitivity is reduced from $0.10\,\mathrm{ps}^{-1}$ to $0.006\,\mathrm{ps}^{-1}$. However the statistical uncertainties are no longer dominant in this region. The largest systematic effect in the study performed by the CDF collaboration originated from effects of the detector alignment which has a direct effect on the scaling of the proper time axis and therefore on Δm_s. This effect must be studied for the future analysis of Δm_s at LHCb.

7.2 Calibration of the Same Side Tagger

In section 2.3.3 the analysis of the CP violation in the $B_s^0 \to J/\psi(\mu\mu)\phi(KK)$ decay has been presented briefly. For this analysis the knowledge whether the B_s^0 has mixed or not is necessary as well. The calibration of the flavor tagging algorithms which are used to extract this information cannot be performed in this decay channel on real data. This is the case because both B_s^0 and $\overline{B}_s^0$ decay into the same final state $J/\psi(\mu\mu)\phi(KK)$. So the flavor of the B-meson at the decay cannot be determined from the decay products.
Therefore if one does not want to rely solely on Monte Carlo simulated data the calibration of the flavor tagging algorithms has to be performed in control channels that are self tagging. The opposite side taggers are considered uncorrelated to the variables on the signal side (see section 5.4). Therefore control channels of B^+ and B^0 decays can be used to calibrate them. For the same side kaon tagger the situation is different because the kaon is clearly correlated with the signal B_s^0. Due to this fact the calibration of the same side kaon tagger has to be performed in a B_s^0 decay channel.
This leads to several problems. First of all due to the fast oscillation the analysis has to be performed time dependently with a very small proper time resolution. In contrast to that there is no oscillation in case of a B^+ and a very slow oscillation in case of the B^0. Secondly the proper time resolution has to be known very accurately because it has a direct effect on the mixing amplitude (see Figure 6.2). So if one wants to extract the dilution factor D one has to know the proper time resolution exactly or include a systematic uncertainty. In section 7.3.1 such a systematic study will be presented.

7.2.1 Fit of a Constant Dilution

To extract the sensitivity to the dilution factor D, it is computed in a similar manner to the Δm_s analysis described in chapter 7.1. In this case as well a toy study over 10 runs with the full statistics corresponding to 2 fb^{-1} was performed. For the tagging parameters (ε_{tag} and ω) were the values for the same side tagger which are displayed in Table 5.1 used. In the fits all parameters except for the dilution factor were fixed. Table 7.2 shows the results of these fits for the tight and the loose mass cut. The improvement of the sensitivity by including the partially

	tight mass cut	loose mass cut
$\sigma_D[\%]$	0.732 ± 0.002	0.567 ± 0.001

Table 7.2: *Results of the sensitivity study of the determination of the dilution factor D of the same side tagger for the tight mass cut (left column) and the wide mass cut (right column)*

reconstructed decay modes corresponds to a gain of a factor 1.7 in statistics. This gain is larger than in case of the Δm_s measurement because the k-factor of the partially reconstructed decays corresponds to a scaling of the proper time axis. This has a direct effect on the frequency, but only an indirect effect on the amplitude of the oscillation. Therefore the sensitivity to the dilution factor is not worsened as much.
This result can be further improved by using a different method which will be described in detail in the following section.

7.2.2 Method of Fitting a Scale-Factor on an Event-per-Event Dilution

The flavor tagging algorithms that were presented in section 5 give out one more variable that has not been used in this analysis so far. Aside from the tagging decision q the taggers also give an estimate over the quality of the tag. This estimate has the form of a mistag probability ω and will be referred to as ω_{estimate} from now on. The idea for the use of this variable is to improve the effective statistic by giving the events that have a higher tagging quality (lower ω_{estimate}) a higher weight. This is the case because the average effective tagging efficiency $\overline{\varepsilon}_{\text{eff}}$ becomes

$$\overline{\varepsilon}_{\text{eff}} = \varepsilon_{\text{tag}}\overline{D}^2 \longrightarrow \sum_{j}^{\textit{all events}} \varepsilon_{\text{tag}} D_j^2 = \sum_{j}^{\textit{all events}} \varepsilon_{\text{tag}}(1-2\omega_{\text{estimate}})^2. \tag{7.4}$$

The subscript j indicates that the dilution factor $D = 1-2\omega$ is now an event-per-event quantity. Because the squared dilution factors of every event are summed over to compute ε_{eff} the events with a higher D (lower ω_{estimate}) receive a larger weight. To use this event based mistag probability it has to be made sure that it represents the real mistag fraction properly. Therefore it will be presented in the following sections how well these ω_{estimate} of the individual taggers describe their actual performance.

7.2.3 Performance of the Taggers using ω_{estimate}

The distributions of these ω_{estimate} are shown on the subfigures on left hand side of Figures 7.1 - 7.4 for each individual tagger in the $B_s^0 \to D_s^- \pi^+$ Monte Carlo data sample. Table 7.3 compares the effective efficiency from Table 5.1 using the average mistag probability $\overline{\omega}_{\text{true}}$ to ε_{eff} using the event-per-event quantity ω_{estimate}. The values in the right hand column were calculated using equation (7.4) and taking ω_{estimate} from the DC06 Monte Carlo sample of the fully reconstructed $B_s^0 \to D_s^- \pi^+$ decays. The numbers are without errors because the uncertainty of ω_{estimate} is unknown. It can be seen that the ε_{eff} for the same side kaon tagger improves from 3.46% to 4.10%. The individual improvements for the taggers differ

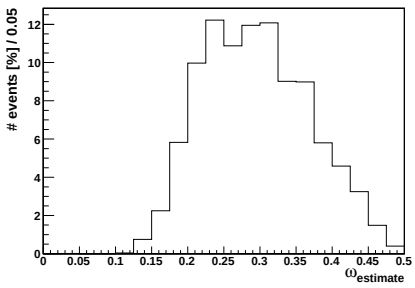
	$\varepsilon_{\text{eff}}[\%]$ using $\overline{\omega}_{\text{true}}$	$\varepsilon_{\text{eff}}[\%]$ using ω_{estimate}
e	0.39 ± 0.03	0.56
μ	1.43 ± 0.05	1.55
OS Kaon	1.69 ± 0.06	2.12
Vertex Charge	1.46 ± 0.05	1.46
SS Kaon	3.46 ± 0.08	4.10
combined	6.22 ± 0.10	9.92

Table 7.3: *Tagging performance of the individual tagging algorithms on Monte Carlo data corresponding to 2*fb^{-1} *of* $\text{B}_\text{s}^0 \rightarrow \text{D}_\text{s}^- \pi^+$ *decays. Comparison between average mistag probability* $\overline{\omega}_{true}$ *and event-per-event mistag probability* $\omega_{estimate}$.

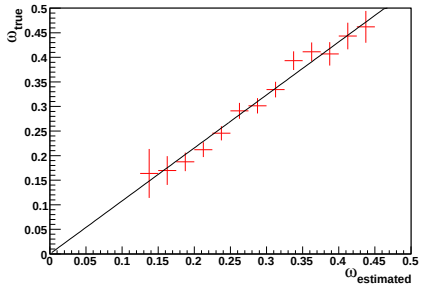
according to the spread in their distributions of ω_{estimate} and their average $\overline{\omega}_{\text{true}}$. This method is used in this analysis only for the calibration of the same side kaon tagger. The performance of the other tagging algorithms are only given for completeness. For more information please refer to [28] and [29].
On the right hand side of Figures 7.1 - 7.4 is shown how well the mistag probability is described by the estimated ω for each tagger. In every bin of ω_{estimate} it is counted how many events were tagged correctly and how many incorrectly. Fitted to the histograms are straight lines emerging from the origin. This model is an approximation in which one assumes that ω_{estimate} and ω_{true} are correlated by a global scale factor s_D.
In the ideal case of a perfectly calibrated tagging algorithm this scale factor would be equal to 1, *i.e.* the estimated mistag probability would exactly describe the true mistag rate. Unfortunately in reality this is not the case because ω_{estimate} is an output of a neural net that was trained on a different decay channel ($\text{B}_\text{s}^0 \rightarrow \text{J}/\psi(\mu\mu)\phi(\text{KK})$). If the distributions of the input variables for the neural net (for example kinematic variables like the B_s^0 momentum) are different for the $\text{B}_\text{s}^0 \rightarrow \text{D}_\text{s}^- \pi^+$ decay channel and the $\text{B}_\text{s}^0 \rightarrow \text{J}/\psi(\mu\mu)\phi(\text{KK})$ decay channel, then ω_{estimate} does not represent the tagger's performance accurately. Therefore the fitted straight lines in the subfigures on the right hand side of Figures 7.1 - 7.4 are either above the diagonal which represents a scale factor equal to 1 or below it. If s_D is larger than 1, it means that the tagger overestimates its power because the true ω is systematically higher than ω_{estimate}. If s_D is smaller than 1, the tagger underestimates its power respectively.

Lepton Taggers

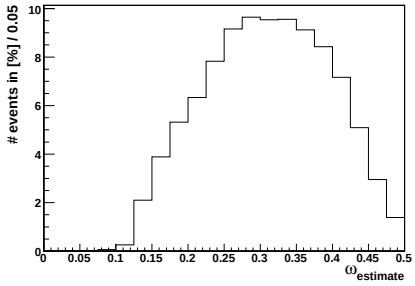
The performance of the lepton taggers can be evaluated by looking at Figure 7.1(a) and 7.1(c) and Table 5.1. It can be seen that ω_{estimate} is widely distributed for both the electron and the muon tagger over the range from 0.1 to 0.5. The $\overline{\omega}_{\text{true}}$ is about 30 percent for both lepton taggers. The main difference between the performance of the two taggers is the tagging efficiency ε_{tag}. The electron tagger gives only in about 3% of the events a decision and the muon tagger in about 9%. In Figure 7.1(b) can be seen that the electron tagger overestimates its tagging power. Figure 7.1(d) shows that the tagging power of the muon tagger is well described by its ω_{estimate}.



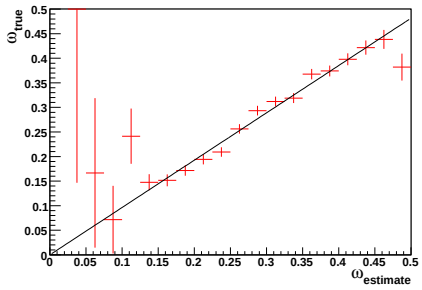
(a) Electron Tagger



(b) Electron Tagger



(c) Muon Tagger

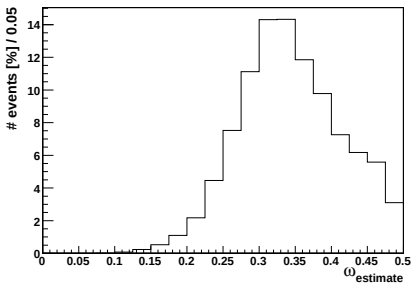


(d) Muon Tagger

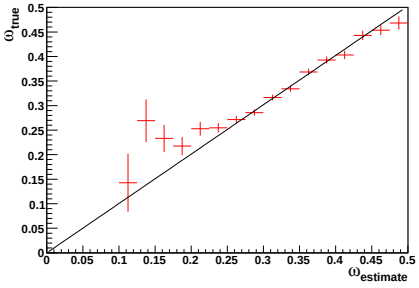
Figure 7.1: *Distributions of the estimated mistag probability $\omega_{estimate}$ for the lepton taggers (a) and (c). Comparison of estimated mistag probability and true mistag rate per bin for the lepton taggers (b) and (d).*

Opposite Side Kaon Tagger

Looking at Figure 7.2(a) and Table 5.1 one can see that the opposite side kaon tagger has a much better ε_{tag} in comparison with the lepton taggers, but a worse $\overline{\omega}_{\text{true}}$. ω_{estimate} is less spread for the OS Kaon tagger than for the lepton taggers. Figure 7.2(b) shows that for the opposite side kaon tagger the true mistag rate is well described by ω_{estimate}.



(a) OS Kaon Tagger

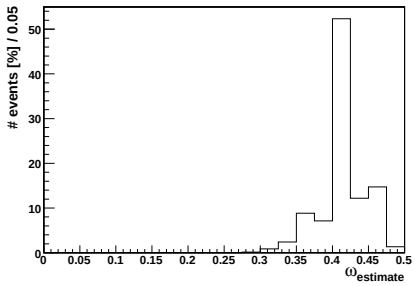


(b) OS Kaon Tagger

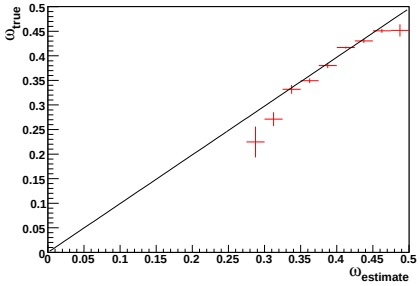
Figure 7.2: *Distributions of the estimated mistag probability $\omega_{estimate}$ for the opposite side kaon tagger (a). Comparison of estimated mistag probability and true mistag rate per bin for the opposite side kaon tagger (b).*

Vertex Charge Tagger

Figure 7.3(a) shows the distribution of ω_{estimate} for the vertex charge tagger. One can see that there is a narrow peak at about $\omega_{\text{estimate}} = 0.42$ that has more than 50% of all events tagged by this algorithm in it. This is due the feature that if all tracks have the same charge one discrete mistag probability of 0.42 is assigned to the tag. Table 5.1 reveals that the vertex charge tagger has a considerably higher tagging efficiency than all the other tagging algorithms, but because of the rather poor mistag rate the effective tagging efficiency is in the same region as the muon and opposite side kaon tagger.



(a) Vertex Charge Tagger

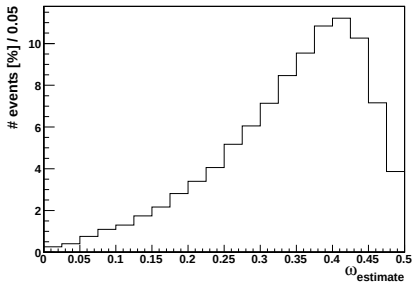


(b) Vertex Charge Tagger

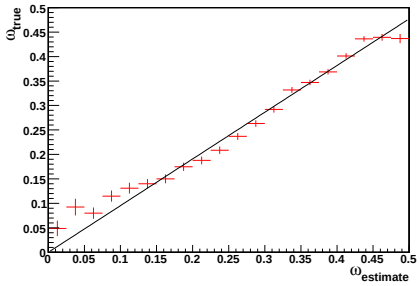
Figure 7.3: *Distributions of the estimated mistag probability $\omega_{estimate}$ for the vertex charge tagger (a). Comparison of estimated mistag probability and true mistag rate per bin for the vertex charge tagger (b).*

Same Side Kaon Tagger

Table 5.1 shows that the effective tagging power ε_{eff} of the same side kaon tagger is considerably higher than those of the opposite side taggers. This makes the same side tagger to a highly important tool in this analysis. One can see in Table 5.1 that it has a significantly high ε_{tag} while having a fairly low $\overline{\omega}_{\text{true}}$. In Figure 7.4(a) it is also illustrated that it is the only tagger which extends into the very low ω_{estimate} region (< 0.1). Figure 7.4(b) indicates that the same side kaon tagger even underestimates its tagging power a little.



(a) SS Kaon Tagger

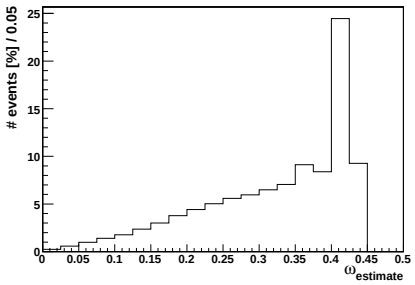


(b) SS Kaon Tagger

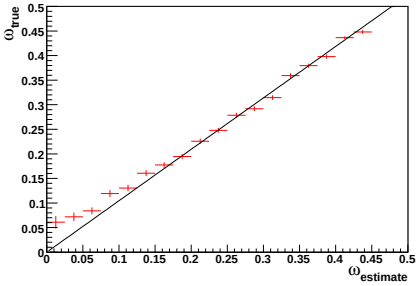
Figure 7.4: *Distributions of the estimated mistag probability $\omega_{estimate}$ for the same side kaon tagger (a). Comparison of estimated mistag probability and true mistag rate per bin for the same side kaon tagger (b).*

Combination of Taggers

In Figure 7.5(a) the distribution of the combined ω_{estimate} is displayed. The narrow peak from the vertex charge tagger is clearly visible but aside from that it is a smooth distribution that is cut off at $\omega_{\text{estimate}} = 0.45$. In Figure 7.5(b) one can see that the combination of the taggers also tends to overestimate its tagging power.



(a) Combination of all Taggers



(b) Combination of all Taggers

Figure 7.5: *Distributions of the estimated mistag probability $\omega_{estimate}$ for the combination of all taggers (a). Comparison of estimated mistag probability and true mistag rate per bin for the combination of all taggers (b).*

7.2.4 Calibration of the Same Side Tagger using the Event-per-Event ω_{estimate}

As mentioned before the goal of this analysis is to calibrate the same side tagger. Therefore the constant dilution factor in the proper time probability density function is replaced by the event-per-event dilution D_j multiplied with a scale factor which is the fitted parameter now.

$$D \longrightarrow s_D \cdot D_j \tag{7.5}$$

For this study again 10 runs with full statistics corresponding to $2\,\text{fb}^{-1}$ were performed. The result for the sensitivity of this scale factor is

$$\sigma_{s_D} = 1.55\% \pm 0.01\% \tag{7.6}$$

for the tight mass cut and

$$\sigma_{s_D} = 1.2\% \pm 0.01\% \tag{7.7}$$

if the wide mass cut is chosen to include the partially reconstructed modes.
This is the sensitivity if the whole signal sample is used. The problem that

occurs is that the scale factor may depend on different variables. Table 7.4 for example shows that if the sample is divided into low p_T and high p_T events and the scale factor is fitted for both samples separately, significantly different results are achieved.
To obtain these numbers a fit to the sample of fully simulated DC06 Monte Carlo data of the decays $B_s^0 \to D_s^- \pi^+$ and $B_s^0 \to D_s^{*-} \pi^+$ was performed. Only events

low p_T	high p_T
0.9527 ± 0.0158	0.8569 ± 0.0218

Table 7.4: *Fit of the scale factor on the event-per-event dilution in two different bins for the reconstructed* p_T *of the* B_s^0

that would have passed the L0 trigger were taken and they had to have been tagged by the same side kaon tagger as well. After these additional selections and the separation into low ($p_T < 8.8\,\mathrm{GeV}/c$) and high ($p_T > 8.8\,\mathrm{GeV}/c$) transverse momentum one is left with about 65000 events per sample.
If there are more of such dependencies, the scale factor has to be fitted in too many bins so that the statistical uncertainties become unbearable for the $B_s^0 \to J/\psi(\mu\mu)\phi(KK)$ analysis. This problem is not yet solved. One proposal for the future strategy in this analysis is not to fit the scale factors on real data in all of those bins. Instead the goal is to create a big enough Monte Carlo data sample, then fit in the different bins on that generated sample. Afterwards all these dependencies are taken as input for the fitter and one global scale factor is fitted with the sensitivity shown in Table 7.4. Then if everything is accounted for this global scale factor is identical with 1.
Another proposal is to include the partially reconstructed semi-leptonic decay modes $B_s^0 \to D_s^- \ell^+ \nu_\ell$ in which the neutrino is not reconstructed. Here there is a lot of statistics expected, but the expected distribution of the k-factor is broader which makes the momentum resolution worse in this decay. So it has to be examined whether the effective gain in statistics is enough to calibrate the scale-factors on data in all necessary bins.

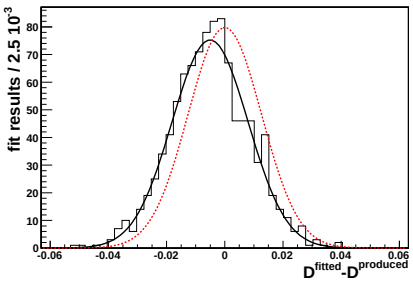
7.3 Systematic Studies

7.3.1 Systematic Effects on the Calibration of the Same Side Tagger

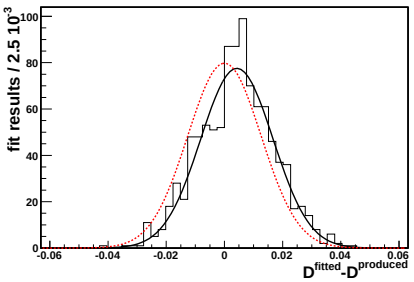
For the study of the tagging power of the same side kaon tagger systematic studies were performed in which the proper time resolution was varied over 5% and the background mistag probability over 10% and once for the total opposite (1-ω).

Proper Time Resolution

For this study toy experiments with 90000[1] events each were conducted and the effect of a 5% deviation of the proper time resolution was studied. To do that during the generation of the samples a value that was 5% higher or lower was used for the proper time resolution. In the fit the parameters for the proper time resolution were fixed to the default values.
In Figure 7.6 is shown what effect on the fit of a constant dilution factor D is caused by deviating the proper time resolution by 5%. These pull distributions are slightly different from the ones shown in section 6.3, because they are not

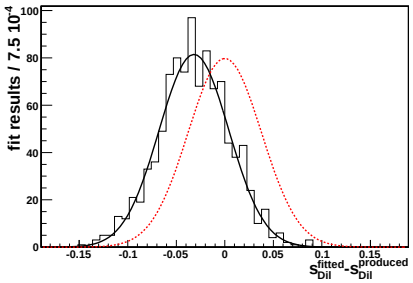


(a) Proper Time Resolution + 5%

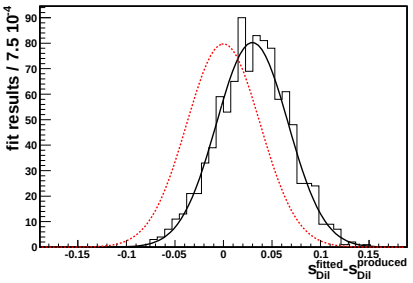


(b) Proper Time Resolution - 5%

Figure 7.6: *Systematic effect of a 5% wrongly determined proper time resolution on the fit of a constant dilution.*



(a) Proper Time Resolution + 5%



(b) Proper Time Resolution - 5%

Figure 7.7: *Systematic effect of a 5% wrongly determined proper time resolution on the fit of the scale factor on the event-per-event dilution.*

normalized to the error of the fit, but instead show the absolute systematic shift.

[1]The systematic effects are assumed to be independent from the amount of statistic used.

Figure 7.7 shows the effect on the fit of a scale factor on the event-per-event dilution. The pull distributions in Figures 7.6 and 7.7 are clearly shifted from

Fit of a constant dilution		
Bias	χ^2/ndf	Mean [%]
$\sigma_t + 5\%$	0.84	-0.51±0.04
$\sigma_t - 5\%$	1.44	0.43±0.04
Fit of a scale factor on the event-per-event dilution		
Bias	χ^2/ndf	Mean [%]
$\sigma_t + 5\%$	1.19	-3.2±0.1
$\sigma_t - 5\%$	0.91	3.0±0.1

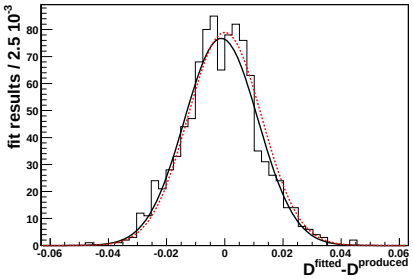
Table 7.5: *Results from the fits to the pull distributions of a constant dilution shown in Figure 7.6 and of a scale factor on the event-per-event dilution shown in Figure 7.7.*

their expected position (red dotted line). In case of the fit of a constant dilution a 5% deviation for the proper time resolution resolves in a systematic bias of about 0.5% and in case of the event-per-event dilution method the bias amounts to about 3%. These systematic uncertainties are added to the statistical uncertainties that were obtained in section 7.2.1 and 7.2.4.
This means that the second method increases the effective efficiency of the sample (see Table 7.3), but it is more sensitive to a wrongly determined proper time resolution.

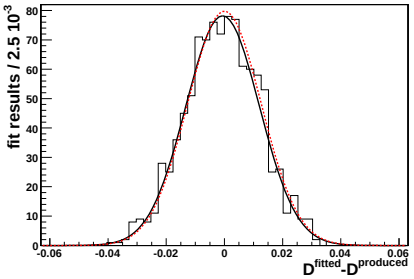
Tagging Power of the Background

The second systematic influence on the fit of the dilution that was examined is a wrongly determined tagging behavior of the background (see section 5.5). For this study the values for the mistag probability ω were varied for the background over ±10%. Also in one study the effect of a total opposite mistag probability ($\omega_{bkg} \rightarrow (1 - \omega_{bkg})$) in the background is examined. Like in the previous study the values were generated with the deviation and in the fit fixed to the default values.
Figures 7.8 and 7.9 show the pull distributions of the according toy experiments. Tables 7.6 and 7.6 show the results of the fitted Gaussians to these pull distributions. Like in the case of the proper time resolution the scale factor is more sensitive to the variation of the tagging behavior of the background than the fit of a constant dilution factor. For the fit of a constant dilution the variation by 10% leads to a systematic bias of about 0.1%. For the method of the event-per-event dilution the introduced systematic bias is about 0.5%.
If one assumes the complete opposite $(1 - \omega_{\mathrm{bkg}})$ this leads to a systematic bias of

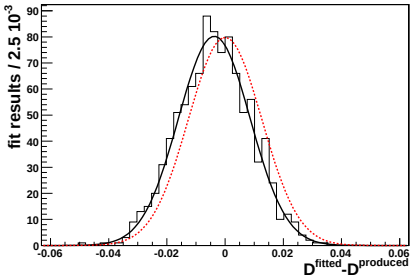
about 0.4% in case of the constant dilution and it is about 4% for the method of the event-per-event dilution.



(a) $\omega_{\mathrm{bkg}} + 10\%$



(b) $\omega_{\mathrm{bkg}} - 10\%$

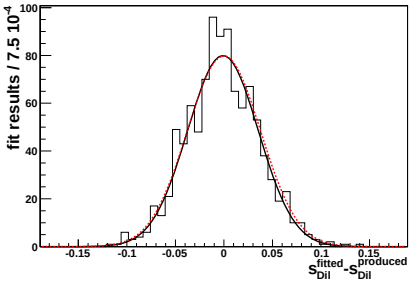


(c) $1 - \omega_{\mathrm{bkg}}$

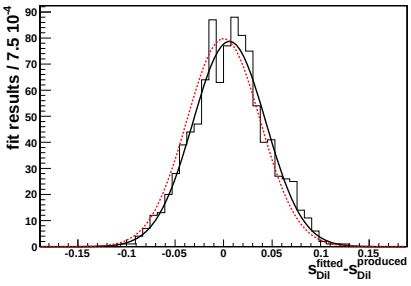
Figure 7.8: *Systematic effect of a 10% wrongly determined mistag probability of the background events (a) and (b) and the completely opposite tagging behavior ($\omega \rightarrow (1-\omega)$) (c) on the fit of a constant dilution.*

Fit of a constant dilution		
Bias	χ^2/ndf	Mean[%]
$\omega_{\mathrm{bkg}} + 10\%$	1.02	-0.12±0.04
$\omega_{\mathrm{bkg}} - 10\%$	0.87	-0.06±0.04
$(1 - \omega_{\mathrm{bkg}})$	0.69	-0.37±0.04

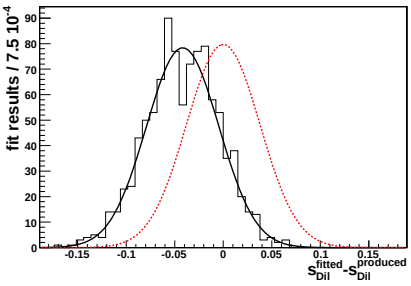
Table 7.6: *Results from the fits to the pull distributions of a constant dilution shown in Figure 7.8.*



(a) $\omega_{\text{bkg}} + 10\%$



(b) $\omega_{\text{bkg}} - 10\%$



(c) $1 - \omega_{\text{bkg}}$

Figure 7.9: *Systematic effect of a 10% wrongly determined mistag probability of the background events (a) and (b) and the completely opposite tagging behavior ($\omega \rightarrow (1-\omega)$) (c) on the fit of a scale factor to the event-per-event dilution.*

Fit of a scale factor on the event-per-event dilution		
Bias	χ^2/ndf	Mean[%]
$\omega_{\text{bkg}} + 10\%$	1.19	0.1±0.1
$\omega_{\text{bkg}} - 10\%$	0.78	0.6±0.1
$(1 - \omega_{\text{bkg}})$	1.14	-4.2±0.1

Table 7.7: *Results from the fits to the pull distributions of a scale factor on the event-per-event dilution shown in Figure 7.9.*

Combined Systematic Effect on the Fit of the Dilution

The systematic uncertainties from sections 7.3.1 and 7.3.1 are added quadratically. In case of a background mistag probability that is determined about 10% wrongly this leads to

$$\sigma_{syst} = 0.6\% \tag{7.8}$$

for a constant dilution $D = 35\%$ and

$$\sigma_{syst} = 3.04\% \tag{7.9}$$

for a scale factor $s_D = 120\%$.
In the worst case of the tagging behavior being the complete opposite of the assumed $(1 - \omega_{\text{bkg}})$ the systematic uncertainty of the constant dilution is

$$\sigma_{syst} = 0.64\% \tag{7.10}$$

and for a scale factor equal to $s_D = 1.2$ it is

$$\sigma_{syst} = 5\% \tag{7.11}$$

7.3.2 Systematic Effects on the Δm_{s} Analysis

For the Δm_{s} study the same systematic studies were performed as a sanity check. There is no systematic bias expected to be introduced by a wrongly determined proper time resolution or tagging power of the background because these directly effect the amplitude of the oscillation but should not have any effect on the frequency.

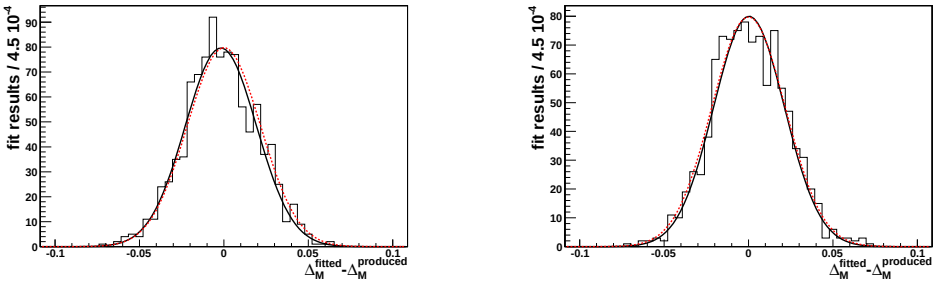
Proper Time Resolution

The systematic study for the Δm_{s} analysis is identical with that for the tagger calibration. Toy experiments with 90000 events each were conducted and the effect of a 5% deviation of the proper time resolution was studied. The results for 1000 runs of these toy experiments for both ±5% are displayed in Figure 7.10 It can be seen that as expected a 5% wrongly determined proper time resolution

Bias	χ^2/ndf	Mean	Width
$\sigma_t + 5\%$	1.13	-0.082 ± 0.032	0.972 ± 0.024
$\sigma_t - 5\%$	1.18	0.013 ± 0.031	0.963 ± 0.023

Table 7.8: *Results from the fits to the pull distributions of Δm_{s} shown in Figure 7.10.*

does not introduce a significant bias on Δm_{s}.



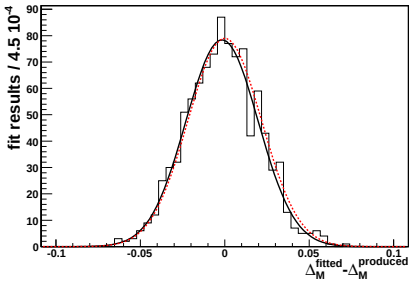
(a) Proper Time Resolution + 5%

(b) Proper Time Resolution - 5%

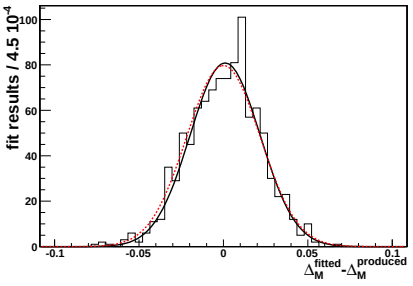
Figure 7.10: *Systematic effect of a 5% wrongly determined proper time resolution on the measurement of* Δm_{s}.

Tagging Power of the Background

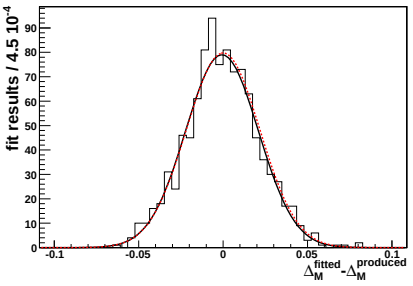
For this study like in case of the tagger calibration the values for the mistag probability ω were varied for the background over $\pm 10\%$ as well as the effect of a total opposite mistag probability ($\omega_{bkg} \rightarrow (1 - \omega_{bkg})$) in the background is examined. Figure 7.11 show the pull distributions of Δm_{s} with a deviation of ω_{bkg} of $\pm 10\%$ and the total opposite.
It can be seen that like expected neither a variation of the mistag probability of the background of 10% nor if one assumes the complete opposite tagging power of $(1 - \omega_{\mathrm{bkg}})$ introduces a significant bias to the measurement of Δm_{s} in this analysis.



(a) $\omega_{\mathrm{bkg}} + 10\%$



(b) $\omega_{\mathrm{bkg}} - 10\%$



(c) $1 - \omega_{\mathrm{bkg}}$

Figure 7.11: *Systematic effect of a 10% wrongly determined mistag probability of the background events (a) and (b) and the completely opposite tagging behavior ($\omega \to (1-\omega)$) (c) on the measurement of Δm_{s}.*

Bias	χ^2/ndf	Mean	Width
$\omega_{\mathrm{bkg}} + 10\%$	0.89	-0.075 $\pm$ 0.032	0.984 $\pm$ 0.025
$\omega_{\mathrm{bkg}} - 10\%$	1.31	0.046 $\pm$ 0.031	0.954 $\pm$ 0.024
$(1 - \omega_{\mathrm{bkg}})$	0.90	-0.031 $\pm$ 0.032	0.989 $\pm$ 0.026

Table 7.9: *Results from the fits to the pull distributions of Δm_{s} shown in Figure 7.11.*

Chapter 8

Summary and Conclusion

In this thesis the tools that are used to determine B_s^0 mixing parameters in the decay $B_s^0 \to D_s^- \pi^+$ were presented. After one year of nominal runtime an integrated luminosity of $\mathcal{L}_{\text{int}} = 2\,\text{fb}^{-1}$ is expected. This corresponds to about $2.5 \cdot 10^5$ $B_s^0 \to D_s^- \pi^+$ events. It has been shown that by including the about $3.6 \cdot 10^5$ $B_s^0 \to D_s^{*-} \pi^+$ and $B_s^0 \to D_s^- \rho^+$ events the available statistics can be improved.
An estimate for the statistical uncertainties on Δm_s and the mistag rate ω after one year of nominal data taking at LHCb was given. In the end, systematic studies were presented which examined the effects of uncertainties on the proper time resolution and tagging behavior of the background on the mistag rate and the event-per-event mistag rate. The final results are

$$\sigma_{\Delta m_s}(stat) = 0.006\,\text{ps}^{-1} \tag{8.1}$$

for the statistical sensitivity of Δm_s after one nominal year of LHCb runtime. The systematic uncertainties due to detector alignment are however expected to be larger than the presented statistical uncertainties.
For the calibration of the same side tagger the sensitivity of a constant dilution ($D = (1 - 2\omega)$) equal to 35% is expected to be

$$\sigma_D(stat) = 0.73\% \qquad \text{and} \qquad \sigma_D(syst) = 0.60\% \tag{8.2}$$

using $B_s^0 \to D_s^- \pi^+$ decays. The systematic uncertainties have been evaluated for a proper time resolution that is known with an accuracy of 5% and the tagging power of the background components with an accuracy of 10%. Here the statistical uncertainties are still larger than the estimated systematic ones. If the additional decay channels $B_s^0 \to D_s^{*-} \pi^+$ and $B_s^0 \to D_s^- \rho^+$ are included in the analysis the statistical sensitivity is improved to

$$\sigma_D(stat) = 0.57\%, \tag{8.3}$$

while systematics are expected to stay the same.
Using the method of the event-per-event dilution for the calibration of the same

side tagger one obtains for the sensitivity to a scale factor s_D

$$\sigma_{s_D}(stat) = 1.6\% \qquad \text{and} \qquad \sigma_{s_D}(syst) = 3.0\%. \tag{8.4}$$

However the dependency of the scale-factor on a range of kinematic variables like for example the B_s^0 momentum has to be evaluated. Therefore the statistical uncertainty in each of those bins is accordingly larger.
It has been shown that by including the additional decay modes $B_s^0 \to D_s^{*-}\pi^+$ and $B_s^0 \to D_s^-\rho^+$ the statistical uncertainty of the full sample corresponding to $2\,\text{fb}^{-1}$ is improved to

$$\sigma_{s_D}(stat) = 1.2\% \tag{8.5}$$

which corresponds to a gain in statistics of a factor 1.7.

Bibliography

[1] A. Abulencia *et al.* [CDF Collaboration], *Observation of B/s0 anti-B/s0 oscillations*, Phys. Rev. Lett. **97** (2006) 242003 [arXiv:hep-ex/0609040].

[2] V. M. Abazov *et al.* [D0 Collaboration], *First direct two-sided bound on the B_s^0 oscillation frequency*, Phys. Rev. Lett. **97** (2006) 021802 [arXiv:hep-ex/0603029].

[3] J. van Tilburg, *Track simulation and reconstruction at LHCb*, CERN-THESIS-2005-040, Amsterdam : NIKHEF, 2005.

[4] M. Kobayashi, T. Maskawa, *CP Violation In The Renormalizable Theory Of Weak Interaction*, Prog. Theor. Phys. **49** 652 (1973).

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

[6] Particle Data Group, *Particle Physics Booklet*, 2008

[7] CKMfitter Group (J. Charles *et al.*), Eur. Phys. J. C41, 1-131 (2005) [arXiv:hep-ph/0406184], updated results and plots available at: http://ckmfitter.in2p3.fr

[8] Amsler, C. *et al.*, Phys. Lett. B **667** , 1-1232 (2008).

[9] Albrecht, J. *et al.*, *Road map for the measurement of mixing induced CP violation in* $\mathrm{B_s^0} \to \mathrm{J}/\psi\phi$ *at LHCb*, LHCb/ROADMAP3-001 document in preparation, (2009).

[10] CERN Multimedia gallery, *http://multimedia-gallery.web.cern.ch/multimedia-gallery/*

[11] Antunes-Nobrega, R. *et al.*, *LHCb reoptimized detector design and performance : Technical Design Report*, CERN-LHCC-2003-030 ; LHCb-TDR-9

[12] Amato, S. *et al.*, *LHCb: Technical proposal*, CERN-LHCC-98-004

[13] CERN Courier archive, *http://cerncourier.com/cws/article/cern/35453/1/CCdis2_09_08*

[14] Amato, S. *et al.*, *LHCb magnet : Technical Design Report*, CERN-LHCC-2000-007 ; LHCb-TDR-1

[15] Barbosa-Marinho, P. R. *et al.*, *LHCb VELO (VErtex LOcator) : Technical Design Report*, CERN-LHCC-2001-011 ; LHCb-TDR-5

[16] Barbosa-Marinho, P. R. *et al.*, *LHCb inner tracker : Technical Design Report*, CERN-LHCC-2002-029 ; LHCb-TDR-8

[17] Barbosa-Marinho, P. R. *et al.*, *LHCb outer tracker : Technical Design Report*, CERN-LHCC-2001-024 ; LHCb-TDR-6

[18] Amato, S. *et al.*, *LHCb RICH : Technical Design Report*, CERN-LHCC-2000-037 ; LHCb-TDR-3

[19] Amato, S. *et al.*, *LHCb calorimeters : Technical Design Report*, CERN-LHCC-2000-036 ; LHCb-TDR-2

[20] Barbosa-Marinho, P. R. *et al.*, *LHCb muon system : Technical Design Report*, CERN-LHCC-2001-010 ; LHCb-TDR-4

[21] D. Buskulic *et al.* [ALEPH Collaboration], *An Investigation of B_d^0 and B_s^0 oscillation*, Phys. Lett. B **322** (1994) 441.

[22] S. Giani *et al.*, *Geant – Detector Description and Simulation Tool, CERN Program Library Long Writeup*, W5013 (1994).

[23] DaVinci v22r3, *$B2DPLUSXROOT/options v1r3p1*, 2008-12-02

[24] V.V.Gligorov, *Reconstruction of the decay modes* $\mathrm{B_d^0 \to D_d^{\pm}\pi^{\mp}}$, $\mathrm{B_s^0 \to D_s^-\pi^+}$ *and,* $\mathrm{B_s^0 \to D_s^{\pm}K^{\mp}}$ *at LHCb*, LHCb-PUB-2009-003

[25] K. Anikeev *et al.* [CDF B_s Mixing group], *B_s Oscillations in Fully and Partially Reconstructed* $\mathrm{B_s^0 \to D_s^-\pi^+}$ *Decays Using Opposite Side Tags and Same Side Tags in 1* $\mathrm{fb^{-1}}$, CDF/DOC/BOTTOM/CDFR/8347

[26] Nuno T. Leonardo, *Analysis of B_s flavor oscillations at CDF*, CDF/THESIS/CDF/PUBLIC/8609, Doctoral Thesis, MIT, September 2006

[27] DaVinci v22r3, *$DECFILESROOT/dkfiles/Bs_DsX=cocktail,DsInAcc.dec*, Gen/DecFiles v13r8

[28] M.Calvi *et al.*, *LHCb Flavour Tagging Performance*, LHCb 2003-115

[29] M.Calvi, O.Leroy, M.Musy, *Flavour Tagging Algorithms and Performance in LHCb*, LHCb 2007-058, Geneva: CERN, 2007

Appendix A

Monte Carlo Data Samples

Decay Mode	MC data sample
$B_s^0 \to D_s^- \pi^+$	"13264001 - Bs_Dspi=DecProdCut"
$B_s^0 \to D_s^{*-} \pi^+$	"13264201 - Bs_Dsstpi=DecProdCut"
$B_s^0 \to D_s^- \rho^+$	"13264401 - Bs_DsRho=DecProdCut"[1]
$B_s^0 \to D_s^- X$	"13410001 - Bs_DsX=cocktail,DsinAcc"[2]
inclusive $b\bar{b}$ decays	"10000000 - incl_b"
inclusive $D_s^\pm$ decays $B^0 \to D_s^\pm X$	"23263001 - incl_Ds=DecProdCut"
$\Lambda_b \to D_s^- p^+$	"15264001 - Lb_Dsp=DecProdCut"
$B^0 \to D^- \pi^+$	"11264001 - Bd_D-p+=DecProdCut"

Table A.1: *Monte Carlo data samples for the individual decays produced within the Data Challenge 06*

[1] Because of time issues there was no fully simulated MC data sample of this decay available. Therefore the events were generated without detector simulation and only the kinematic cuts were applied. Then the necessary distributions like for example the mass were extracted.

[2] The annex "cocktail" indicates that the sample does not include all possible $B_s^0 \to D_s^- X$ decays but a selection out of them (for an explicit list of the decays see [27]). "DsinAcc" means that the daughter particles of the D_s have to have an angle θ smaller than 400 mrad from the z-axis.

Appendix B

Normalization of the Signal Proper Time PDF

The normalization $\mathcal{N}$ of the proper time probability density function has to be performed only for the exponential part P_{exp} of equation (6.11) and (6.16). The reason for that is that only the sum over all events has to be normalized and it is not expected that there is a difference in numbers between mixed and unmixed events so the $P_{\cos}$ terms should cancel each other. From now on all equations are referring to the proper time pdf of the partially reconstructed decay modes $P^t_{\text{partially}}$. For the fully reconstructed decay mode the same formulas apply with the k-factor identical to 1. Without the $P_{\cos}$ term and with only one Gaussian distribution used for convolution equation (6.16) becomes

$$P^t_{\text{sig}} = \frac{1}{\mathcal{N}} \cdot \int \frac{k}{2\tau} e^{\left(-\frac{k}{\tau}(t-\frac{k\sigma_t^2}{2\tau})\right)} \cdot \text{Erfc}\left(\frac{k\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau}\right) \cdot \mathcal{F}(k)dk \cdot \varepsilon_t(t). \tag{B.1}$$

This leads to the normalization factor $\mathcal{N}$:

$$\mathcal{N}(\sigma_t) = \int \left[\int_{-\infty}^{+\infty} e^{-\frac{k}{\tau}t} \text{Erfc}\left(\frac{k\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau}\right)\varepsilon_t(t)\right] \cdot \frac{k}{2\tau} e^{\frac{k^2\sigma_t^2}{2\tau^2}} \cdot \mathcal{F}(k)dk \tag{B.2}$$

The function $\varepsilon_t(t)$ is chosen in such a way that the integration over the proper time t can be performed analytically (see [26]). The additional integration over the k-factor although also analytically possible is performed numerically everytime the parameters are changed because it depends on the proper time resolution σ_t. It has to be computed separately for the two resolutions of the Double Gaussian.

Erklärung

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

Heidelberg, den 02. Juli 2009 _____

-