Department of Physics and Astronomy

University of Heidelberg

Diploma thesis
in Physics

submitted by
Alexander Bien
born in Heidelberg, Germany

April 2010

Simulation study for the determination of the CP violating B_s mixing phase Φ_s with LHCb

This diploma thesis has been carried out by Alexander Bien
at the
Physical Institute
under the supervision of
Prof. Dr. Ulrich Uwer

Abstract: Simulation study for the determination of the CP violating B_s mixing phase Φ_s with LHCb

In the Standard Model of particle physics the CP violating B_S^0 mixing phase Φ_s arising in the decay $B_S^0 \to J/\psi\phi$ can be predicted very precisely. Possible new particles can contribute to the mixing phase and lead to deviations from the Standard Model value. Thus, the precise measurement of Φ_s in B_S^0 decays is an exact method to indirectly search for new phenomena ("New Physics") and one of the key goals of LHCb. The mixing phase is determined by fits to four-dimensional distributions of measured quantities. At very low statistics, corresponding to integrated luminosities of 10 to 50 pb^{-1}, the assumption of Gaussian errors is not valid any more. Therefore, a method proposed by Feldman and Cousins is applied which is able to determine confidence levels at low statistics correctly. In this thesis, the method is motivated and the implementation for the decay $B_S^0 \to J/\psi\phi$ is explained. Using simulated data one can conclude that while the method is inevitable for very low statistics it becomes less important with increasing integrated luminosity (above 100 to 150 pb^{-1}). In addition, in this thesis the influence of different characteristic detector properties (tagging power, lifetime resolution) on the determination of Φ_s is investigated. The change of these properties in the case of an increased interaction rate is studied as well.

Kurzfassung: Simulationsstudie zur Bestimmung der CP verletzenden B_s-Mischungsphase Φ_s mit LHCb

Im Standardmodell der Teilchenphysik wird die CP verletzende B_S^0-Mischungsphase Φ_s im Zerfall $B_S^0 \to J/\psi\phi$ sehr genau vorausgesagt. Mögliche neue Teilchen können zur Mischungsphase beitragen und so zu Abweichungen vom Standardmodellwert führen. Daher stellt die exakte Messung von Φ_s in B_S^0-Zerfällen eine exzellente Methode zur indirekten Suche nach neuen Phänomenen („Neue Physik") dar und ist eines der Hauptziele von LHCb. Die Mischungsphase wird durch Anpassungsrechnungen an vierdimensionale Messgrößenverteilungen gewonnen. Bei sehr niedriger Ereignisstatistik, entsprechend einer integrierten Luminosität von 10 bis 50 pb^{-1}, ist die Annahme gauß'scher Fehler hierbei nicht mehr gültig. Deshalb wird eine von Feldman und Cousins vorgeschlagene Methode angewandt, die in der Lage ist, Vertrauensbereiche bei niedriger Statistik korrekt zu bestimmen. In dieser Arbeit wird die Methode motiviert und die Implementierung für den Zerfall $B_S^0 \to J/\psi\phi$ erklärt. Zudem wird anhand simulierter Daten belegt, dass die Methode zwar bei sehr niedriger Statistik unumgänglich ist, sie aber mit zunehmender integrierter Luminosität weniger wichtig wird (ab etwa 100 bis 150 pb^{-1}). Weiterhin wird in der Arbeit der Einfluss verschiedener Detektorkenngrößen (Tagging Power, Zeitauflösung) auf die Bestimmung von Φ_s untersucht. Die Änderung dieser Kenngrößen im Falle einer höheren Wechselwirkungsrate wird ebenfalls studiert.

Contents

Introduction

Beginning in the 1960s a theoretical framework was developed that includes all known particles and their interactions [1][2][3]: the so-called Standard Model of particle physics. Today the Standard Model is in excellent accordance with experiments, only one piece, the so-called Higgs boson, has not been discovered yet. There are, however, several observations that cannot be explained within the framework of the Standard Model. One of these open questions is the existence of Dark Matter. Cosmological observations show that almost 25% of the energy content in the universe consists of this unknown form of matter [4].

Another important observation that cannot be explained by the Standard Model is the matter-antimatter asymmetry in the universe. Obviously, the universe is made of matter and not of antimatter resulting in an imbalance between baryons and antibaryons. In 1967, the Russian physicist Andrei Sakharov named three minimum conditions for this asymmetry to occur [5]: Baryon number violation, C and CP violation and deviation from thermal equilibrium.

CP violation (where C stands for charge conjugation and P for parity transformation) was discovered around 50 years ago [6]. There is, however, the common understanding that the amount of CP violation provided by the Standard Model is not sufficient to explain the matter-antimatter asymmetry. Numerous theories beyond the Standard Model have been developed that are able to provide new sources of CP violation. Since in the Standard Model several observables related to CP violation can be predicted very precisely, even small experimental deviations can indicate new phenomena ("New Physics").

To examine such CP violating effects and possible deviations from Standard Model predictions, especially in the decay channels of B hadrons, the LHCb detector was built. It is one of the four large experiments at the Large Hadron Collider near Geneva where protons will collide at a centre-of-mass energy of $\sqrt{s} = 14$ TeV. The first proton-proton collisions have been observed in autumn 2009.

In the Standard Model the weak eigenstates of the quarks are not equivalent to the mass eigenstates. The transformation between the two sets of eigenstates is described by the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix [7]. The matrix can be represented by three rotation angles and one phase. This phase is responsible for the CP violation in the Standard Model. The CKM mechanism also predicts CP violation in the neutral $B_{d,s}$ meson system. In this system one can measure an additional phase due to CP violation in the interference between mixing and decay. At LHCb these effects are mainly studied in the "golden channel" $B_S^0 \rightarrow J/\psi(\mu\mu)\phi(KK)$. For this case the phase describing CP violation in the interference between mixing and decay is called B_S mixing phase Φ_S. At the Tevatron Φ_S has already been measured but the results still have a large statistical uncertainty [8] [9]. At LHCb one will achieve a much higher accuracy and be able to improve the existing results and possibly discover signs of New Physics.

In the early phase of data taking the extraction of the B_S mixing phase Φ_S needs special consideration. With only few events available the error determination in an analysis with many free parameters is extremely difficult. Standard likelihood methods lead to an underestimation of the errors. Instead, a method by Feldman and Cousins [11] is applied in this thesis. This method provides the correct confidence regions by construction. Of particular importance for the extraction of Φ_S especially at low statistics are the flavour tagging performance and the lifetime resolution.

The outline of the thesis is as follows: In Chapter 1 a theoretical introduction is given. The main focus is on CP violation in the B_S system, in particular for the decay $B_S^0 \rightarrow J/\psi\phi$. The LHCb experiment is presented in Chapter 2 with special emphasis on flavour tagging and lifetime resolution. In Chapter 3 the analysed data sets and the standard selection strategy for the decay channel $B_S^0 \rightarrow J/\psi\phi$ are explained. Key quantities like mass resolution and trigger efficiencies are studied for two different bunch filling scenarios to investigate the possibility to operate at a higher than nominal luminosity without significant loss in reconstruction power. In Chapter 4 the afore-mentioned Feldman-Cousins method is motivated and the implementation for the decay $B_S^0 \rightarrow J/\psi\phi$ is explained. Using toy experiments Chapter 5 examines the question how quantities like flavour tagging performance, lifetime resolution and the number of signal events influence the sensitivity on Φ_S and the shape of the confidence regions. Chapter 6 is dedicated to the extraction of Φ_S from fully simulated Monte Carlo data corresponding to integrated luminosities up to $50 \, \mathrm{pb}^{-1}$. The thesis ends with a short conclusion.

Theoretical Background

In this chapter the Standard Model of particle physics is briefly summarised. In addition, the theoretical background for CP violation in the neutral B meson system is discussed, with special emphasis on the decay channel $B_S^0 \rightarrow J/\psi\phi$.

1.1 The Standard Model of Particle Physics

The so-called Standard Model of particle physics comprises the current knowledge of fundamental particles and their interactions [12]. The fundamental matter particles, quarks and leptons, are fermions and carry spin 1/2. There are three different generations which are listed in Table 1.1. In the Standard Model, the quarks and charged leptons are massive whereas the neutrinos are massless. Recent experiments have, however, shown that neutrinos have small non-zero masses (see e.g. [13]). Corresponding to these fermions there are the respective antifermions with identical mass but opposite electric charge and colour.

		Quarks			Leptons	
generation	type	charge [e]	mass	type	charge [e]	mass
I	u	$+\frac{2}{3}$	1.5 - 3.3 MeV	ν_e	0	$\leq$ 2 eV
	d	$-\frac{1}{3}$	3.5 - 6.0 MeV	e	-1	511.0 keV
II	c	$+\frac{2}{3}$	$\sim$ 1.27 GeV	ν_μ	0	$\leq$ 2 eV
	s	$-\frac{1}{3}$	$\sim$ 104 MeV	μ	-1	105.7 MeV
III	t	$+\frac{2}{3}$	$\sim$ 171.2 GeV	ν_τ	0	$\leq$ 2 eV
	b	$-\frac{1}{3}$	$\sim$ 4.2 GeV	τ	-1	1.78 GeV

Table 1.1: *Fundamental fermions in the Standard Model [14]*

The three fundamental interactions covered by the Standard Model, the electromagnetic, the weak and the strong interaction, are all mediated by vector bosons. An overview can be found in Table 1.2. Gluons, the bosons of the strong interaction, as well as W and Z bosons, mediating the weak interaction,

interaction	couples to	boson	mass [GeV]
strong	colour	8 gluons (g)	0
e.m.	electric charge	photon (γ)	0
weak	weak charge	$W^\pm$, Z^0	80.4, 91.2

Table 1.2: *Vector bosons mediating the three interactions in particle physics [12], [14]*

carry charge (strong and weak charge respectively) and hence, interact with each other. Photons, responsible for electromagnetic interactions, however, are uncharged particles and thus are not self-coupling. Only the bosons of the weak interaction are massive (of the order of 100 GeV).

In the Standard Model the electromagnetic and the weak interaction are unified into one electroweak interaction. To achieve this a new particle, the so-called Higgs boson, has to be introduced. The Higgs field is responsible for the spontaneous symmetry breaking between the massive weak vector bosons and the massless photon. After symmetry breaking also the fundamental fermions acquire their masses through the Higgs field. The Higgs boson is the only particle within the Standard Model that has not been found yet experimentally [15].

So far, there has been a very good agreement between the assumptions of the Standard Model and experiments, there are, however, still unsolved problems and questions which have led to numerous theories beyond the Standard Model. One of the main problems is that the Standard Model cannot provide for all the CP violation that is needed to understand the matter-antimatter asymmetry in the universe. Furthermore, no candidate for Dark Matter has been found yet. Various theoretical models like Supersymmetry try to give explanations.

1.2 The CKM mechanism

In the quark sector, the electroweak eigenstates (d', s', b') of the quarks are not equivalent to the mass eigenstates (d, s, b) [7] [16]. Hence, one quark can be transformed into one another. The transformation between the two sets of eigenstates can be described by a unitary matrix, the Cabibbo-Kobayashi-Maskawa (CKM) matrix [7]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

Since the matrix elements can, in principle, be complex there are 18 parameters describing the CKM matrix. However, the unitarity condition $V_{CKM}(V_{CKM})^\dagger = \mathbb{1}$ reduces the number of free parameters to nine and five of them can be absorbed into unobservable phases of the quark fields. So, there are four parameters left, three rotation angles and one phase. This phase is the only source of CP viola-

tion in the Standard Model. CP violation means that the combined symmetry operation of transforming a particle into its antiparticle (charge inversion C) and of inverting its spatial coordinates (parity transformation P) is broken.

There are many parametrisations of the CKM matrix. One of the most common ones is the so-called Wolfenstein parametrisation [17] in which each matrix element can be written as a power series in the parameter $\lambda = |V_{us}| = 0.225$ and which reflects the hierarchy of the matrix:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.2}$$

Due to the unitarity of the matrix there are six "triangle" relations between its elements, e.g. the b-d relation

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.3}$$

This equation can be illustrated as a triangle in the complex plane. Usually, one side, $V_{cd}V_{cb}^*$, is normalised to one and one defines $\overline{\rho} + i\overline{\eta} \equiv -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}$ [18]. The rescaled triangle can be seen in Figure 1.1.
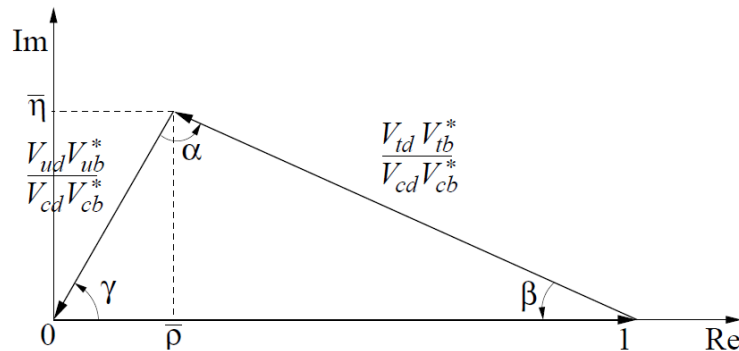

Figure 1.1: *The unitarity triangle from equation (1.3) with one side normalised to one [19]*

The three angles used in the figure are defined as

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{1.4}$$

Figure 1.2 shows the current experimental status (as of 2009) of the unitarity triangle [20]. In the plot the area dashed in red corresponds to the 68% confidence level for the position of the triangle's apex. In the Standard Model one expects a closed triangle with the apex at $(\overline{\rho}, \overline{\eta})$. Any significant experimental
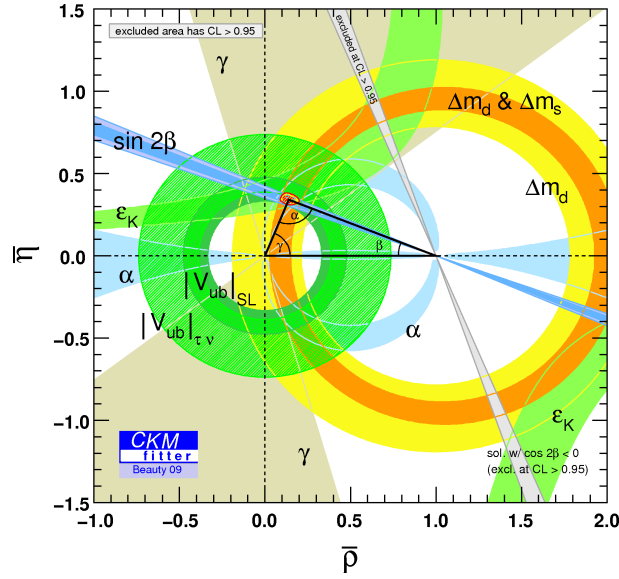
Figure 1.2: *Experimental status of the unitarity triangle [20]*

deviation would imply additional flavour violation not explained by the Standard Model. From Figure 1.2 one can conclude that the bulk of flavour violation in the Standard Model is actually described by the CKM mechanism and that new sources are heavily constrained.

The global CKM fit includes measurements of $|V_{ub}|$ from inclusive $B \to X_u l^- \overline{\nu_l}$ and exclusive $B \to \pi l \nu_l$ decays and measurements of ϵ_k from indirect CP violation in the system of neutral kaons (all green), as well as measurements of the mass difference in the B_d^0 and B_s^0 meson systems (orange) and of $\sin 2\beta$ (blue).

Another unitarity triangle relation, relevant for the study of the B_s^0 meson system, can be obtained by replacing the d with an s quark in equation (1.3):

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{1.5}$$

The first side is much shorter than the other two. Consequently, the opposing angle

$$\beta_s \equiv \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right) \tag{1.6}$$

is very small. The measurement of β_s is one way to search for New Physics [18].

1.3 Mixing induced CP violation in $B_S^0 \to J/\psi\phi$

B_q^0-$\overline{B}_q^0$ mixing (q = s,d) is defined as the transition between the two flavour eigenstates $\left|B_q^0\right\rangle$ and $|\overline{B}_q^0\rangle$. In the Standard Model the mixing is caused by weak interactions which can be described by the box diagrams in Figure 1.3 [18].
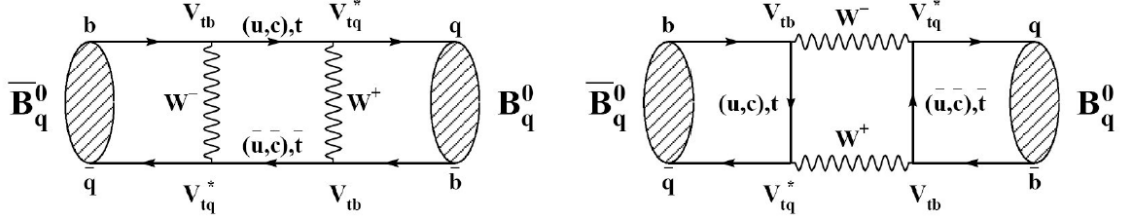
Figure 1.3: *Feynman diagrams describing B-mixing within the Standard Model [23]*

The B_q^0-$\overline{B}_q^0$ mixing processes induce oscillations between B_q^0 and $\overline{B}_q^0$. In other words, an initial B_q^0 or $\overline{B}_q^0$ can evolve into a superposition of a B_q^0 and a $\overline{B}_q^0$. The time evolution of the flavour eigenstates $|B_q^0\rangle$ and $|\overline{B}_q^0\rangle$ can be described by the Schrödinger equation

$$i\frac{d}{dt}\left(\begin{array}{c} |\,B_q^0\,\rangle \\ |\,\overline{B}_q^0\,\rangle \end{array}\right) = \left(\mathbf{M} - i\frac{\boldsymbol{\Gamma}}{2}\right)\left(\begin{array}{c} |\,B_q^0\,\rangle \\ |\,\overline{B}_q^0\,\rangle \end{array}\right) \tag{1.7}$$

where $\mathbf{M}$ and $\boldsymbol{\Gamma}$ are 2×2 hermitian matrices representing the mass and the decay width of the B_q^0 and the $\overline{B}_q^0$. Due to the quark mixing the mass eigenstates of the B meson are not equal to the flavour eigenstates $|B_q^0\rangle$ and $|\overline{B}_q^0\rangle$. The mass eigenstates are the eigenvectors of $\mathbf{M} - i\boldsymbol{\Gamma}/2$ and can be expressed as a linear combination of the flavour eigenstates:

$$|B_L\rangle = p\,|B_q^0\rangle + q\,|\,\overline{B}_q^0\,\rangle \tag{1.8}$$

$$|B_H\rangle = p\,|B_q^0\rangle - q\,|\,\overline{B}_q^0\,\rangle \tag{1.9}$$

where the complex coefficients p and q obey the normalisation condition $|p|^2 + |q|^2 = 1$.

The mass and width differences of the B meson eigenstates are defined as

$$\Delta M_q = M_H - M_L, \qquad \Delta\Gamma_q = \Gamma_H - \Gamma_L \tag{1.10}$$

whereas the average mass and width can be written as

$$M_q = \frac{M_H + M_L}{2}, \qquad \Gamma_q = \frac{\Gamma_H + \Gamma_L}{2}. \tag{1.11}$$

From now on only the case of B_s mixing is considered (i.e. q=s). The latest experimental result for the respective mass difference is $\Delta M_s = (17.77 \pm 0.12)$ ps^{-1} [21].

By calculating the eigenvectors and eigenvalues of $\mathbf{M} - i\boldsymbol{\Gamma}/2$ the ratio q/p can be expressed in terms of the off-diagonal matrix elements M_{12} and Γ_{12}:

$$\frac{q}{p} = -\frac{2M_{12}^* - i\Gamma_{12}^*}{\Delta M + \frac{i}{2}\Delta\Gamma} \tag{1.12}$$

The ratio q/p represents the mixing between the two flavour eigenstates which can be seen from equations (1.8) and (1.9). The decay into a final state f is described by the amplitudes $A_f = \langle\, f \,|\, B_s^0 \,\rangle$ and $\overline{A}_f = \langle\, f \,|\, \overline{B}_s^0 \,\rangle$. In the scope of this thesis $f = J/\psi\phi$ is considered. The B_s mixing phase Φ_s is then defined as

$$\Phi_s = -\arg\left(\frac{q}{p}\frac{\overline{A}_f}{A_f}\right). \tag{1.13}$$

This expression can be interpreted as the relative phase between q/p and $\frac{\overline{A}_f}{A_f}$, i.e. between mixing and decay. If $\frac{q}{p}\frac{\overline{A}_f}{A_f} \neq \pm 1$, then CP is violated. This is called CP violation in the interference between decays with and without mixing because it is resulting from the interference between $B_s^0 \to f$ and $B_s^0 \to \overline{B}_s^0 \to f$. In the following (except for the conclusion) the mixing phase is denoted simply as Φ.

New Physics only effects M_{12} but not Γ_{12}. The mixing phase can be expressed as the sum of its Standard Model value and a possible New Physics contribution:

$$\Phi = \Phi^{SM} + \Phi^{NP} \tag{1.14}$$

The Standard Model part of the mixing phase is linked to the angle β_s defined in equation (1.6):

$$\Phi^{SM} = -2\beta_s \tag{1.15}$$

A more detailed description of B_s^0-mixing can be found in [22].

The CP violating mixing phase Φ is measured by analysing the decay $B_s^0 \to J/\psi(\mu\mu)\phi(KK)$. The Standard Model value of Φ can be calculated with high precision [23]:

$$\Phi^{SM} = -0.0368 \pm 0.0017 \tag{1.16}$$

Therefore, the decay mode $B_s^0 \to J/\psi\phi$ is one of the "golden channels" to measure effects of New Physics on CP observables. To precisely determine Φ in $B_s^0 \to J/\psi\phi$ is one of the key measurements of LHCb.

In theories beyond the Standard Model numerous diagrams can, in principle, contribute to the mixing phase and hence, change the above value. As an example, one of the most popular theories of New Physics is the so called **M**inimal **S**upersymmetric **S**tandard **M**odel (MSSM, see [24]). This theory predicts e.g. two Higgs doublets that can contribute to the mixing of B mesons [25]. A more extensive discussion of New Physics contributions would be beyond the scope of this thesis. Instead, more details can be, for example, found in the theory papers [25], [26] and [27].

In the following the quantities needed to measure CP violation in the decay $B_s^0 \to J/\psi\phi$ are discussed [23]. In this decay mode a pseudo-scalar particle decays into two vector particles. Hence, there are three possible final states with relative orbital angular momenta $l = 0, 1, 2$. The CP eigenvalue of the final state depends on the relative angular momentum:

$$CP\,|J/\psi\phi\rangle = (-1)^l\,|J/\psi\phi\rangle \tag{1.17}$$

In general, the final state is an admixture of CP-even ($l = 0, 2$) and CP-odd ($l = 1$) states which can be separated by means of an angular analysis. In the so-called transversity formalism, for example, the amplitudes at $t = 0$, $A_0(0)$ and $A_\parallel(0)$, are CP-even whereas $A_\perp(0)$ is CP-odd [23]. Within this formalism, the final states can be described by three angles $\Omega = \{\theta, \varphi, \psi\}$ which are shown in Figure 1.4.
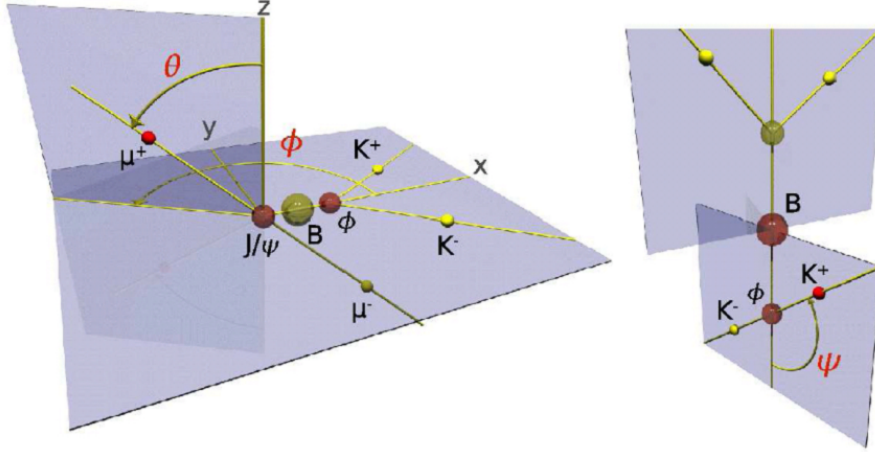

Figure 1.4: *The three decay product angles (highlighted in red) in the transversity formalism [23]*

In the J/ψ rest frame, the polar and azimuthal angles θ and φ describe the direction of the μ^+. The angle ψ is defined in the ϕ rest frame and describes the angle between the momenta $\vec{p}(K^+)$ and $-\vec{p}(J/\psi)$. To describe the B_S^0 decay rates one introduces four parameters:

- Two independent amplitudes: $|A_\perp(0)|^2$ and $|A_\parallel(0)|^2$ with the convention $|A_\perp(0)|^2 + |A_\parallel(0)|^2 + |A_0(0)|^2 = 1$.

- Two independent CP conserving (strong) phases ($\delta_0 = 0$ by convention): $\delta_\parallel = \arg(A_\parallel(0)A_0^*(0))$ and $\delta_\perp = \arg(A_\perp(0)A_0^*(0))$.

In total eight physics parameters describe the decay $B_S^0 \to J/\psi\phi$: Φ, $\Delta\Gamma_S$, Γ_S, ΔM_S, $\delta_\parallel$, $\delta_\perp$, $|A_\perp(0)|^2$ and $|A_\parallel(0)|^2$.
The differential decay rate for the B_S^0 is given by (see [23])

$$\frac{\mathrm{d}^4\Gamma(B_S^0 \to J/\psi\phi)}{\mathrm{d}t\,\mathrm{d}\cos\theta\,\mathrm{d}\varphi\,\mathrm{d}\cos\psi} \equiv \frac{\mathrm{d}^4\Gamma}{\mathrm{d}t\mathrm{d}\Omega} \propto \sum_{k=1}^{6} h_k(t)f_k(\Omega). \tag{1.18}$$

The formulae for the $\overline{B}_S^0$ are completely analogous and are not written down explicitly.

The functions $h_k(t)$ and $f_k(\Omega)$ introduced in the equation above are defined in Table 1.3.

k	$h_k(t)$	$f_k(\Omega)$
1	$\left|A_0(t)\right|^2$	$2\cos^2\psi(1-\sin^2\theta\cos^2\varphi)$
2	$\left|A_\parallel(t)\right|^2$	$\sin^2\psi(1-\sin^2\theta\sin^2\varphi)$
3	$\left|A_\perp(t)\right|^2$	$\sin^2\psi\sin^2\theta$
4	$\Im\left\{A_\parallel^*(t)A_\perp(t)\right\}$	$-\sin^2\psi\sin 2\theta\sin\varphi$
5	$\Re\left\{A_0^*(t)A_\parallel(t)\right\}$	$\frac{1}{\sqrt{2}}\sin 2\psi\sin^2\theta\sin 2\varphi$
6	$\Im\left\{A_0^*(t)A_\perp(t)\right\}$	$\frac{1}{\sqrt{2}}\sin 2\psi\sin 2\theta\cos\varphi$

Table 1.3: *Definitions of the functions from equation (1.18) [23]*

The time dependent amplitudes can be written down explicitly. Accordingly, the expressions for the $\overline{B}_S^0$ are given by the CP conjugated functions $\overline{h}_k(t)$ whereas the angular terms do not change.

$$\left|A_0(t)\right|^2 = \left|A_0(0)\right|^2\,e^{-\Gamma_st}\left[\cosh\left(\frac{\Delta\Gamma_st}{2}\right)-\cos\Phi\sinh\left(\frac{\Delta\Gamma_st}{2}\right)+\sin\Phi\sin(\Delta m_st)\right],$$

$$\left|A_\parallel(t)\right|^2 = \left|A_\parallel(0)\right|^2\,e^{-\Gamma_st}\left[\cosh\left(\frac{\Delta\Gamma_st}{2}\right)-\cos\Phi\sinh\left(\frac{\Delta\Gamma_st}{2}\right)+\sin\Phi\sin(\Delta m_st)\right],$$

$$\left|A_\perp(t)\right|^2 = \left|A_\perp(0)\right|^2\,e^{-\Gamma_st}\left[\cosh\left(\frac{\Delta\Gamma_st}{2}\right)+\cos\Phi\sinh\left(\frac{\Delta\Gamma_st}{2}\right)-\sin\Phi\sin(\Delta m_st)\right],$$

$$\Im\left\{A_\parallel^*(t)A_\perp(t)\right\} = \left|A_\parallel(0)\right|\left|A_\perp(0)\right|\,e^{-\Gamma_st}\left[-\cos(\delta_\perp-\delta_\parallel)\sin\Phi\sinh\left(\frac{\Delta\Gamma_st}{2}\right)\right.$$
$$\left.+\sin(\delta_\perp-\delta_\parallel)\cos(\Delta m_st)-\cos(\delta_\perp-\delta_\parallel)\cos\Phi\sin(\Delta m_st)\right],$$

$$\Re\left\{A_0^*(t)A_\parallel(t)\right\} = \left|A_0(0)\right|\left|A_\parallel(0)\right|\cos\delta_\parallel e^{-\Gamma_st}\left[\cosh\left(\frac{\Delta\Gamma_st}{2}\right)-\cos\Phi\sinh\left(\frac{\Delta\Gamma_st}{2}\right)\right.$$
$$\left.+\sin\Phi\sin(\Delta m_st)\right]\text{ and}$$

$$\Im\left\{A_0^*(t)A_\perp(t)\right\} = \left|A_0(0)\right|\left|A_\perp(0)\right|\,e^{-\Gamma_st}\left[-\cos\delta_\perp\sin\Phi\sinh\left(\frac{\Delta\Gamma_st}{2}\right)\right.$$
$$\left.+\sin\delta_\perp\cos(\Delta m_st)-\cos\delta_\perp\cos\Phi\sin(\Delta m_st)\right].$$

There is a two-fold ambiguity in the decay rates that is caused by an exact symmetry: The decay rates are invariant under the four simultaneous transformations

$$\Phi \longleftrightarrow \pi-\Phi,$$
$$\Delta\Gamma_S \longleftrightarrow -\Delta\Gamma_S,$$
$$\delta_\parallel \longleftrightarrow 2\pi-\delta_\parallel \text{ and}$$
$$\delta_\perp \longleftrightarrow \pi-\delta_\perp.$$

The CP violating phase Φ can be extracted from several terms in the decay rate which can be seen from the time dependent amplitudes. As the Standard Model value of Φ is very close to 0, basically all the information about the phase comes from the terms containing $\sin\Phi$ ($\cos\Phi$ close to 1). In four out of the six amplitude terms $\sin\Phi$ is multiplied by $\sin(\Delta m_{\mathrm{S}}t)$. This means almost all the information on Φ is accessed via the amplitude of the oscillation in the time distribution.

Flavour tagging means measuring whether the decay particle was produced as B_s^0 or $\overline{B}_s^0$. The $\sin(\Delta m_{\mathrm{S}}t)$ terms have opposite signs in the case of a B_s^0 and $\overline{B}_s^0$ respectively and thus in the case of an untagged analysis where one cannot resolve the flavour of the B-meson, these terms cancel out. Therefore, almost complete cancellation of the $\sin\Phi$ information – which is crucial to the extraction of Φ – occurs. This shows that flavour tagging is an extremely important tool to measure the B_S mixing phase Φ.

Apart from the CP violating phase, the difference in the decay width $\Delta\Gamma_{\mathrm{S}}$ (in the following denoted simply as $\Delta\Gamma$) is important, especially in the view of stating confidence intervals for Φ (see end of the section). Unlike Φ, $\Delta\Gamma$ is not significantly affected by New Physics because it is dominated by the CKM favoured $b \to c\bar{c}s$ tree-level decays [22] which are not affected by New Physics. Latest theoretical calculations predict a value for $\Delta\Gamma$ in the Standard Model of (see [22])

$$\Delta\Gamma = (0.096 \pm 0.039) \text{ ps}^{-1}. \tag{1.19}$$

As can be seen from the time dependent amplitudes, $\Delta\Gamma$ can be extracted from the sinh and cosh terms. These terms are time dependent themselves and hence, the sensitivity to $\Delta\Gamma$ varies with time.

At the Tevatron, a proton-antiproton collider at the Fermi National Accelerator Laboratory near Chicago, a tagged analysis of the decay $B_s^0 \to J/\psi\phi$ has already been performed ([8] [28]). The two collaborations at the Tevatron, CDF and DØ, have determined confidence levels for the mixing phase Φ. CDF alone have recently updated their result with roughly 3200 signal events corresponding to 2.8 fb^{-1} of data [8]. As point estimates are meaningless due to non-Gaussian errors, the confidence region is quoted in the $\Delta\Gamma - \Phi$ plane (as the main physics parameters of interest). From this the interval

$$\Phi = [-2.58, -0.56] \text{ at } 68\% \text{ confidence level} \tag{1.20}$$

is obtained as a result for the mixing phase [8]. A deviation from the Standard Model of 1.8σ is measured. The current status (winter 2009) of the combined CDF and DØ measurements can be seen in Figure 1.5. These results, too, show a deviation of Φ from the Standard Model value of nearly 2σ (95% contour). The green band indicates the region allowed in New Physics models.

As can be seen from Figure 1.5 the statistical uncertainties in the determination of Φ are large. At LHCb one will achieve a much higher accuracy and be able to check whether the deviation from the Standard Model as measured by CDF and
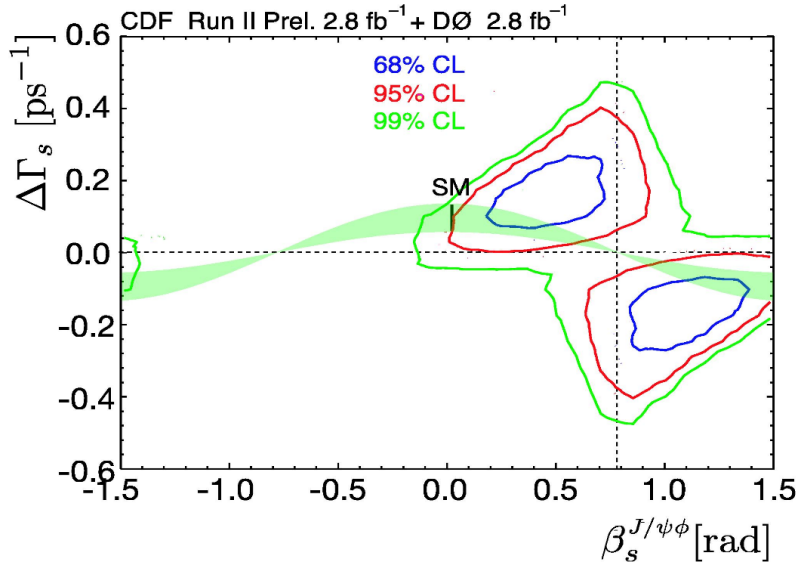
Figure 1.5: *Current status of the combined CDF and DØ measurements. The graph shows the 68, 95 and 99% confidence regions in the $\Delta\Gamma$-β_S plane (Note that $\Delta\Gamma \equiv \Delta\Gamma_{\mathrm{S}}$). The measured value of $\Phi = -2\beta_S$ has a deviation from the Standard Model value of nearly 2σ. The region allowed in New Physics models is also shown (green band) [9].*

DØ holds and thus is, indeed, a sign for New Physics. The discovery of physics beyond the Standard Model would open a new and fascinating branch of both experimental and theoretical particle physics because a deviation of Φ from the Standard Model value would only be an indirect measurement of New Physics. One could not directly conclude to the actual nature of the physics beyond the Standard Model.

A short introduction to basic statistical definitions follows in Section 4.1.

The LHCb experiment

The **L**arge **H**adron **C**ollider **b**eauty (LHCb) experiment is one of the four main experiments operated at the **L**arge **H**adron **C**ollider (LHC) in Geneva. After a short introduction to the LHC itself, the LHCb experiment is described in detail. Flavour tagging, time resolution and the trigger system are particularly emphasised.

2.1 The Large Hadron Collider

The LHC [29] is a proton-proton accelerator at the European Organisation for Nuclear Research (CERN, **C**onseil **E**uropéen pour la **R**echerche **N**ucléaire) in Geneva. It has a circumference of 27 km (for an overview see Figure 2.1).
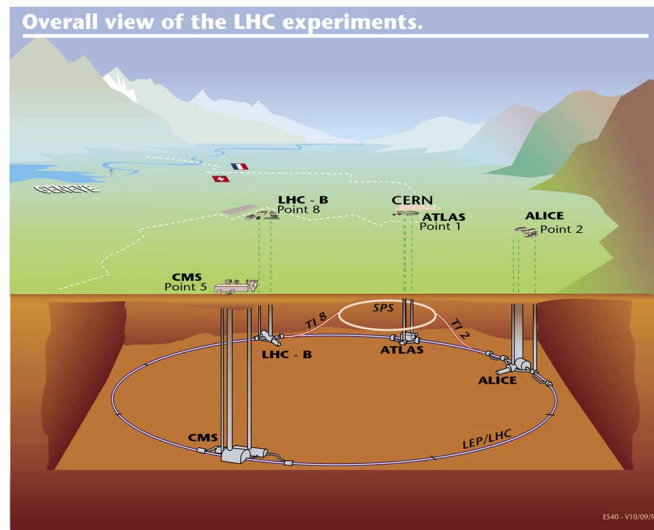

Figure 2.1: *Schematic view of the LHC and the four main experiments ALICE, ATLAS, CMS and LHCb [30]*

The machine design foresees the proton beams to be accelerated up to an energy of 7 TeV in two opposite directions. Each of the oppositely rotating beams consists of 2808 bunches with a spacing of 25 ns, corresponding to a distance of 7.5 metres. Every bunch contains 10^{11} protons. Before injecting the two beams into the LHC, the protons run through a chain of pre-accelerators. The two main pre-accelerating components are the Proton Synchroton (PS) that accelerates the protons up to an energy of 26 GeV and the Super Proton Synchroton (SPS) with a final energy of 450 GeV which also corresponds to the LHC injection energy [31]. The final acceleration to 7 TeV takes about 20 minutes. At the four interaction points, the experiments ALICE, ATLAS, CMS and LHCb, the two beams are focussed and brought to collision. The first proton-proton collisions were observed in November 2009. ALICE will analyse in addition lead-lead collisions in separate LHC runs.

2.2   The LHCb detector

While ATLAS and CMS are multi-purpose experiments and ALICE is built to examine quark-gluon plasma, the LHCb detector is optimised to record and analyse B decays. At the LHC centre-of-mass energy the dominant production mechanism of b and $\bar{b}$ quarks is gluon-gluon fusion. Due to the small mass of these b quarks gluons from a wide momentum range contribute to the production of $b\bar{b}$ pairs. It is very unlikely that both gluons carry the same proton momentum fraction x. Different x-values lead to a boost of the $b\bar{b}$ pairs in beam direction. Hence, the b and anti-b hadrons predominantly fly either in forward or backward direction. This can be seen in Figure 2.2 where the angle between a $b/\bar{b}$ pair produced in a proton-proton interaction and the beam direction is visualised [32].
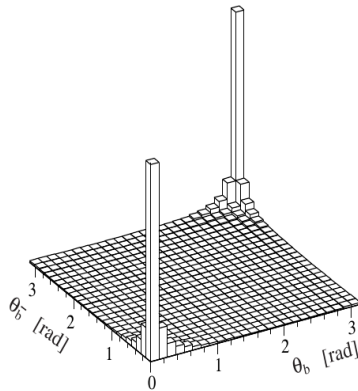

Figure 2.2: *Angle between a $b/\bar{b}$ quark produced in a proton-proton interaction and the beam direction [32]*

Due to this fact that B and $\overline{\text{B}}$ mesons are mainly produced in the same forward or backward cone the LHCb detector is built as a single arm forward spectrometer with an angular coverage from 10 to 300 mrad in the bending plane (x-z plane) and from 10 to 250 mrad in the non-bending plane (y-z plane). A schematic view of the detector can be seen in Figure 2.3 (from [33]).
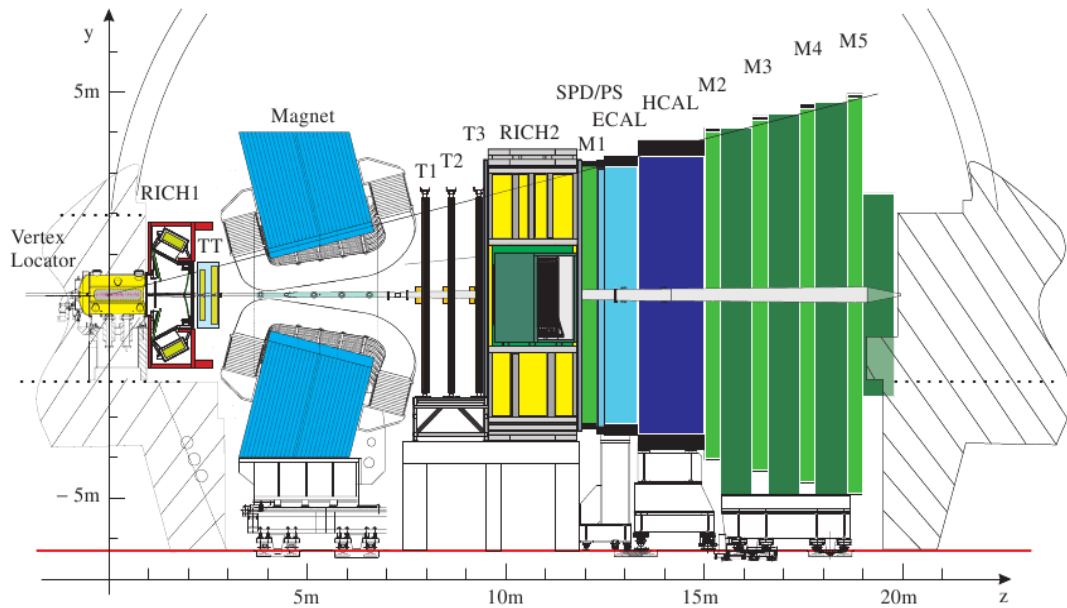

Figure 2.3: *Schematic view of the LHCb detector in the y-z plane (picture from [33])*

Contrary to the ATLAS and CMS experiments the proton beams are only weakly focussed and the nominal luminosity of $\mathcal{L} = 2 \cdot 10^{32}$ cm^{-2}s^{-1} is 100 times smaller than for the above-mentioned multi-purpose detectors (for further explanation see Chapter 3). In the following the main parts of the detector are briefly discussed.

2.2.1 Track Reconstruction

A powerful track reconstruction system is required to precisely measure the trajectories and momenta of the particles. In a magnetic field (integrated field strength at LHCb: $\int B \, \mathrm{d}l = 4.2$ Tm) charged particles are bent depending on their momentum. Thus, the measurement of the slope of the track in front of and behind the magnet allows for the determination of the particle momentum. The tracking system is comprised of the VErtex LOcator (VELO) system and

the Trigger Tracker (TT) installed before and three tracking stations (T1, T2, T3) situated behind the dipole magnet [33].

- The Vertex Locator is built around the interaction point and allows for the precise measurement of primary as well as displaced secondary vertices (see Section 2.4). The 21 VELO stations consist of two types of semi-circular silicon sensors to measure both the r as well as the ϕ (azimuthal angle) coordinates of the track with respect to the beam. The cylindrical design was chosen over a rectilinear geometry to allow for the fast reconstruction of tracks and vertices in the LHCb trigger. The VELO is moveable and in the fully closed position has a distance of 5 mm to the beam axis. The spatial resolution of the reconstructed primary vertex is 42 μm in z-direction and 10 μm in x- and y-direction. A more detailed discussion can be for example found in [34].

- The Trigger Tracker is used for track reconstruction as well as for the momentum measurement of low-energetic particles. In contrast to the tracking stations the TT is able to detect these highly-bent low energy tracks. The Trigger Tracker consists of four layers of silicon sensors. The two outer layers are orientated in y-direction whereas the two in between are tilted by ± 5 degrees with respect to the y-axis. This allows for a spatial resolution in the x-direction as well. More information is available in [35].

- The third component of the tracking system is the Main Tracker consisting of the Inner Tracker (IT) and the Outer Tracker (OT) is divided into three stations T1, T2 and T3. Each of the IT stations comprises four detector boxes distributed around the beam pipe and is covered by silicon microstrip detectors. Similarly to the TT two of the detector layers are vertically oriented whereas the other two are tilted by an angle of ± 5 degrees. The outer region is covered by the OT. The OT consists of four layers per tracking station arranged in a similar geometry as the IT and is composed of straw-tube drift chambers. The tubes with an inner diameter of 5 mm are filled with an Ar-CO_2 gas mixture. The drift time is below 50 ns, the spatial resolution of the OT in x-direction is around 200 μm. In addition, it has a very high efficiency of 98%. For more information on IT and OT see [36] and [37].

2.2.2   Particle Identification

In addition to the reconstruction of charged particle tracks, particle identification is an integral element for the LHCb detector. Especially the separation between kaons and pions is essential. The information from the two RICH detectors, the calorimeters and the muon chambers allows to assign a relative probability for a particle hypothesis to the reconstructed particle tracks (see next subsection).

- The two Ring Imaging Cherenkov Detectors (RICH1 and RICH2) are based on the so-called Cherenkov effect: Energetic electrically charged particles emit electro-magnetic radiation in a medium of refractive index n if their velocity v is larger than the speed of light in the medium $c' = \frac{c}{n}$ where c is the speed of light in vacuum. The radiation is emitted on a cone of opening angle θ_c which can be calculated via

$$\cos\theta_c = \frac{c'}{v} = \frac{1}{\beta n} \tag{2.1}$$

with $\beta = \frac{v}{c}$. With the measurement of the opening angle of the emitted light cone and the momentum from track reconstruction a particle hypothesis can be made. The first detector (RICH1) is situated before the magnet to identify particles in a momentum range up to 50 GeV. It contains both aerogel and C_4F_{10} radiators. The second Cherenkov detector (RICH2) is placed behind the magnet and contains only a CF_4 radiator to analyse tracks with momenta between 50 and 100 GeV in the innermost part of the detector. Both detectors are installed in regions with low magnetic field so that the tracks are not bent significantly while passing through the radiators. A more detailed discussion can be found in [38].

- The calorimeter system consists of four parts to separate electrons/positrons, photons and hadrons as well as to measure energies and positions.
The first component is the Scintillator Pad Detector (SPD) consisting of scintillators with a thickness of 15 mm, situated directly behind the first muon chamber. A signal is only triggered by charged particles and therefore, the SPD can be used to distinguish between electrons/positrons and photons.
In a 12 mm thick lead wall after the SPD an electromagnetic shower is induced that is analysed in the Pre-Shower Detector (PS). The PS is located directly behind the lead wall and also consists of scintillator pads. Since hadronic showers are only produced in the hadronic calorimeter the PS can be used to separate electrons/positrons from hadrons.
In the Electromagnetic CALorimeter (ECAL) electro-magnetic shower energies can be measured with a resolution of

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus 1.5\% \tag{2.2}$$

where $\oplus$ denotes quadratic summation (energy E given in GeV). This resolution is obtained by using alternating layers of 2 mm lead sheets and 4 mm thick scintillator plates and wavelength shifting fibres. Electromagnetic showers are induced in the lead layers and are then sampled by the scintillators.

Behind the ECAL a Hadronic CALorimeter (HCAL) is located. The energy resolution for hadrons is significantly worse than for the electromagnetic showers and is determined to be:

$$\frac{\sigma(E)}{E} = \frac{80\%}{\sqrt{E}} \oplus 10\%. \tag{2.3}$$

The HCAL contains in turns layers of 16 mm thick iron and 4 mm thick scintillators. Hadrons induce hadronic showers in the iron plates that are detected by the scintillators. The HCAL has an overall thickness of 1.2 metres. For a more extensive discussion of the calorimeter system see [39].

- The identification of muons is a fundamental part of the experiment because muons are in the final states of many B-decays sensitive to CP violation and of rare decays. The muon system consists of five chambers (M1 to M5) where M1 is in front of the calorimeter to obtain a better track estimation of the muon trigger and M2 to M5 behind the calorimeter. The chambers are composed of multi-wire proportional chambers (MWPC) and gas electron multipliers (GEM) in the inner part. To shield the muon chambers from hadrons there are 80 cm thick iron absorbers separating the chambers M2 to M5. To shield from muons coming from the opposite direction there is an additional iron absorber behind M5. To pass all five chambers a muon is required to have a minimum momentum of around 6 GeV. As the track density increases towards the beam pipe a higher granularity is chosen in that region. For more information see [40].

2.2.3 Particle Hypothesis

Using the particle identification detectors a particle hypothesis can be assigned to every track [41]. With the measurements of the Cherenkov opening angle θ_c in the RICH detectors and of the momentum from track reconstruction a mass hypothesis can be made via

$$\cos\theta_c = \sqrt{1 + \left(\frac{m}{p}\right)^2} \tag{2.4}$$

which can be solved for m. With additional information from the calorimeter and the muon chambers, a likelihood value $\mathcal{L}(X)$ is connected to the mass hypothesis of a particle X which is compared to the likelihood value for a pion $\mathcal{L}(\pi)$. A relative particle hypothesis ("delta log-likelihood") can then be calculated as follows

$$\Delta\mathcal{L}_{X\pi} = (\ln\mathcal{L}(X) - \ln\mathcal{L}(\pi)) = \ln\left(\frac{\mathcal{L}(X)}{\mathcal{L}(\pi)}\right) \tag{2.5}$$

If $\Delta\mathcal{L}_{X\pi} > 0$, a charged track results more likely from a particle X than from a pion and vice versa.

2.3 Flavour Tagging at LHCb

As will be discussed later flavour tagging, i.e. the identification of the initial flavour of the B-meson, is crucial to measure CP violation in channels like $B_S^0 \rightarrow J/\psi\phi$ [42]. The two basic quantities of flavour tagging are the tagging efficiency ϵ and the the mistag rate ω:

$$\epsilon = \frac{N_R + N_W}{N_R + N_W + N_U}, \quad \omega = \frac{N_W}{N_R + N_W} \tag{2.6}$$

where N_R is the number of correctly tagged, N_W the number of incorrectly tagged and N_U the number of untagged events. With these quantities one can define the dilution $D = 1 - 2\omega$ and the effective tagging power $\epsilon_{eff} = \epsilon D^2$ [1]. A schematic overview of different methods of flavour tagging of B-mesons is shown in Figure 2.4.
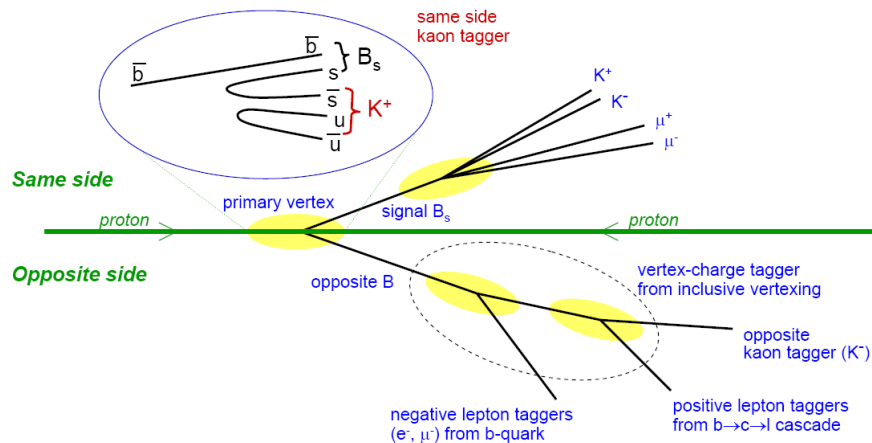

Figure 2.4: *Schematic view of the B-meson flavour tagging mechanisms at LHCb [42]*

One cannot conclude from the B meson decay products to the flavour of the B meson at creation because it may have oscillated before it decays. Consequently, this information has to be obtained from the rest of the event. As in proton-proton collisions quarks are produced in pairs, there are, in principle, two different categories of flavour tagging [42][43][44]: Opposite-side (OS) tagging is to determine the flavour of the reconstructed B meson from the other B hadron produced in the event. This other B hadron is also called the companion B hadron. On the other hand, same-side (SS) tagging is to infer the flavour of the initial b quark from correlations in the fragmentation decay chain of the reconstructed B meson.

---

[1]The statistical weight of N events with $\epsilon_{eff} = \epsilon D^2$ for the CP analysis is that of $\epsilon_{eff} N$.

These flavour tagging techniques rely on the charge of certain particles.
For OS tagging one would ideally reconstruct all decay modes of the companion B
hadron. As this would be particularly difficult to implement, generic signatures,
so-called tags, are used instead. The charge of these tags is directly correlated to
the companion B hadron's flavour. At LHCb this can be done in three different
ways. One can

- measure the charge of an electron or muon from a semileptonic B decay
  (B $\rightarrow$ $l\nu_l$X). This is called *electron tagging* and *muon tagging* respectively.
  If, for example, in such a case a $\bar{\text{b}}$ quark is transformed into a $\bar{\text{c}}$ quark by
  exchanging a W$^+$ boson, then a positively charged lepton is produced which
  allows to tag the flavour of the companion B hadron.

- measure the charge of a kaon from b $\rightarrow$ c $\rightarrow$ s cascade decays (B $\rightarrow$ DX
  $\rightarrow$ KX). This charge is directly related to the flavour of the companion B
  hadron. The method is called *OS kaon tagging*.

- inclusively reconstruct the decay vertex of the companion B hadron. The
  flavour of the companion B hadron can then be concluded from the vertex
  charge (sum of the charges of all tracks originating in the decay vertex).
  This is called *vertex charge tagging*.

SS taggers take advantage from the correlation in the fragmentation decay chain
of the signal B itself. If, for example, the fragmentation of a $\bar{\text{b}}$ quark produces a
B$_\text{S}^0$($\bar{\text{b}}$s), an extra $\bar{\text{s}}$ quark is available (see Figure 2.4). In some cases this additional
quark hadronises into a K$^+$($\bar{\text{s}}$u). The charge of this kaon then implies the flavour
of the B meson at production. This is called *SS kaon tagging*. Apart from
applying cuts on quantities like the momentum and the transverse momentum of
the kaon, a cut on the impact parameter (IP) significance, defined as IP/σ_{IP} [2],
is used to reject kaons that do not originate from the primary vertex. A more
detailed discussion on flavour tagging can be found in [42], [43] and [44].
For the channel B$_\text{S}^0$ $\rightarrow$ J/$\psi\phi$, typical mistag rates of the combined taggers are
of the order of 35% and combined tagging efficiencies are determined to be at
roughly 55%. The exact numbers (overall and for every sub-tagging mechanism
explained in this section) are discussed in more detail in Chapter 3.

2.4   Lifetime resolution at LHCb

The second key issue for the analysis of B$_\text{S}^0$ $\rightarrow$ J/$\psi(\mu\mu)\phi$(KK) events is the
precise determination of the B$_\text{S}^0$ lifetime [45]. With a lifetime of roughly 1.5 ps
the B$_\text{S}^0$ is a relatively long-lived particle and a secondary vertex can be resolved

---

[2]The impact parameter of a track is defined as the distance of closest approach of the track
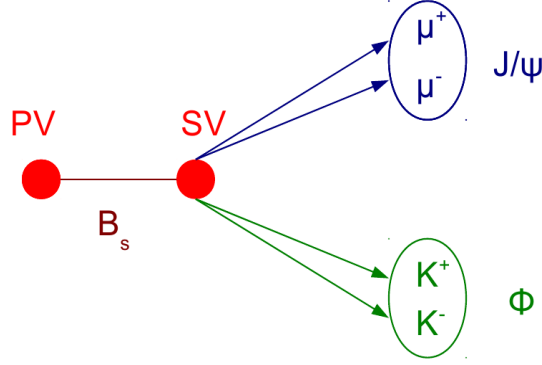to the primary vertex.

Figure 2.5: $B_s^0 \to J/\psi(\mu\mu)\phi(KK)$ *decay topology. The* B_s^0 *is produced in the primary vertex (PV) and decays at the secondary vertex (SV) into a* J/ψ *and a* ϕ *which themselves decay into a muon and a kaon pair respectively.*

in the detector (flight length of about 1 cm). The decay topology can be seen in Figure 2.5.

The proper time is determined by measuring the distance between the primary vertex and the decay vertex and the momentum of the B meson. Technically this is done by a constrained χ^2 fit [46]. The lifetime can then simply be calculated by (originating from $p = \gamma M\beta$)

$$t(B_s^0) = \left(\vec{d} \cdot \frac{\vec{p}}{|\vec{p}|} \right) \frac{M_{B_s}}{|\vec{p}|\, c} \tag{2.7}$$

where $\vec{d}$ is the decay length vector, $\vec{p}$ the momentum vector and M_{B_s} the mass of the B_s^0. To resolve the fast B_s^0 oscillations (frequency of $\frac{\Delta M_s}{2\pi} = 2.83$ ps^{-1}) and thus to determine the mixing phase Φ with sufficient accuracy an excellent proper time resolution is required. In the case of signal events a good approximation for the lifetime resolution is the sum of two Gaussians [45]:

$$R(t) = f_{\sigma_{ct,1}} \cdot G(t, \mu_{ct,1}, \sigma_{ct,1}) + (1 - f_{\sigma_{ct,1}}) \cdot G(t, \mu_{ct,2}, \sigma_{ct,2}) \tag{2.8}$$

where $G(\mu, \sigma)$ is a Gaussian with mean μ and width σ. Monte Carlo studies show [45] that for $B_s^0 \to J/\psi\phi$ signal events a total proper time resolution of

$$\sigma_{ct} = \sqrt{f_{\sigma_{ct,1}} \cdot \sigma_{ct,1}^2 + (1 - f_{\sigma_{ct,1}}) \cdot \sigma_{ct,2}^2} = 38.9 \ \text{fs} \tag{2.9}$$

can be achieved.

In the case of background (for the different categories of background see Chapter 3) only a single Gaussian is assumed as lifetime resolution model (further explanations in Chapter 4).

As will be discussed in Chapter 5 the lifetime resolution is, alongside with the flavour tagging performance, the key quantity in order to precisely determine the B_S mixing phase Φ.

2.5   The Trigger system of LHCb

As for every modern high energy physics experiment the amount of data which has to be processed is enormous. To reduce the actual output rate from 40 MHz (rate of proton-proton collisions) to 2 kHz a trigger system is used. It consists of two steps, firsty the hardware based Level 0 Trigger (L0) and secondly, the software based High Level Trigger (HLT) [47].

The L0 trigger is designed to reduce the rate from 40 MHz to 1 MHz. It uses information from the Muon Chambers to reconstruct particles with high transverse momentum (muon decision) and information from the calorimeters to select particles with a high transverse energy (calorimeter decision). The L0 muon trigger requires a candidate with a minimum transverse momentum of 1.3 GeV, the dimuon trigger two candidates with a total minimum transverse momentum of 1.5 GeV. The L0 hadron trigger is issued if the transverse energy of a hadronic cluster exceeds 3.5 GeV whereas the L0 electromagnetic trigger requires a minimum transverse energy E_t of 2.6 GeV for electron candidates and an E_t of 2.3 GeV for a photon candidate. The hardware trigger also has the possibility to veto very busy events and particles with more than one primary vertex. The L0 trigger has a total latency of $4\mu s$ [47].

The rate of 1 MHz is transferred to the Event Filter Farm (EFF) where the software trigger is running. The EFF consists of 1000 16-core computing nodes. The software trigger itself consists of two stages, the HLT1 and the HLT2. The purpose of the HLT1 is to reduce the background by partially reconstructing the events and verifying the L0 decision. Consequently, an output rate of 30 kHz is passed to the HLT2 where a full event reconstruction is performed using inclusive selections of B-candidates. The final output rate is 2 kHz [47].

As shown in the next chapter the total trigger efficiency is estimated to be 90% for $B_S^0 \to J/\psi\phi$ events based on a selection of Monte Carlo simulated events.

Analysis of Monte Carlo simulated $B_S^0 \rightarrow J/\psi\phi$ events

In this chapter the standard LHCb selection of $B_S^0 \rightarrow J/\psi\phi$ events (with $J/\psi \rightarrow \mu^+\mu^-$ and $\phi \rightarrow K^+K^-$) is discussed. Two different Monte Carlo event samples corresponding to two different instantaneous luminosities are studied: The first sample with one and the second one with three proton-proton interactions per bunch crossing. For both samples a few key quantities such as the mass and proper time resolution are examined. Furthermore, the influence of the higher number of proton-proton interactions per bunch crossing on tagging and trigger efficiencies is analysed.

3.1   Introduction

The ATLAS and CMS experiments are designed to run at a very high instantaneous luminosity. The average number ν of proton-proton interactions per bunch crossing at the design value will be of the order of 20 [48]. This high number, however, would make the reconstruction of B hadrons and the unambiguous association of a primary vertex very difficult, if not impossible. Consequently, for LHCb a much lower design luminosity which results only in an average of $\nu = 1$ proton-proton interactions per bunch crossing is chosen. LHCb is operating at a luminosity 100 times smaller than the nominal luminosity planned for ATLAS and CMS. In 2009, different bunch filling scenarios were discussed. To achieve an increased integrated luminosity, LHCb has investigated the possibility to run also at a higher instantaneous luminosity. Within this thesis it is checked whether the detector performance, especially the tagging efficiency, significantly decreases if the number of proton-proton interactions per bunch crossing is increased or if it is still possible to keep up a reasonable performance even at higher interaction rates. In section 3.3 the performance of two luminosity scenarios is compared: the standard $\nu = 1$ case and the case where ν is increased to 3.

29

3.2 Selection strategy for $B_S^0 \to J/\psi\phi$ events

In this section the selection of $B_S^0 \to J/\psi(\mu\mu)\phi(KK)$ Monte Carlo events is explained [3]. At the LHC centre-of-mass energy the two b-hadrons are mainly produced in the same forward or backward region: In 40% of the cases at least one B hadron is produced in a cone of roughly 400 mrad. This fact allows for a cut on the angle between the beam axis and the direction of the B_S^0 decay products, Θ, such that

$$10 \text{ mrad} < \Theta < 400 \text{ mrad}, \tag{3.1}$$

which reduces the simulation time of the Monte Carlo sample significantly. The selection strategy can be summarised by three main goals [49]:

- Maximisation of the signal yield

- Minimisation of distortions in lifetime and angular acceptances for the signal channel as well as for the two main control channels $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$.

- Common selection of b-hadrons for signal and control channels in terms of their momentum distribution. This allows to apply tagging performances determined on the control channels to the signal sample.

The Monte Carlo events are part of the LHCb 'Monte Carlo 2009' (MC09) series, generated with the LHCb software Gauss v37r3p1 [50] and Boole v18r1 [51] and reconstructed with Brunel v34r7 [52].[4] The actual analysis is carried out with DaVinci v24r4 [53].

The B_S^0 "physics parameters" describing the B_S^0 decays that are used in the generation of the events are given in Table 3.1. Note that by convention the phase δ_0 is set to 0 and that for the polarisation amplitudes the normalisation relation $|A_\perp(0)|^2 + |A_\parallel(0)|^2 + |A_0(0)|^2 = 1$ is valid.

Parameter	Value
B_S^0 mass	5366.3 MeV
Γ_S	0.6793 ps^{-1}
Δm_S	17.8 ps^{-1}
$\Delta\Gamma_S$	0.06 ps^{-1}
$|A_\perp(0)|^2$	0.16
$|A_0(0)|^2$	0.6

---

[3]At the LHC B hadrons are produced in pairs.

[4]The actual names of the data samples are MC09-Beam5TeV-VeloClosed-MagDown-Nu1-MC09-Sim04Reco02-withTruth for one and MC09-Beam5TeV-VeloClosed-MagDown-Nu3-MC09-Sim04Reco02-withTruth for three pp interactions per bunch crossing. The event type for both samples is 13144002-BsJpsiphi,mm=CPV,DecProdCut.

$\delta_\parallel$	2.50 rad
$\delta_\perp$	-0.17 rad
ϕ	-0.04 rad

Table 3.1: B_S^0 *decay parameters used in the generation of the MC09 sample (modified from [49])*

The selection criteria for the $B_S^0 \to J/\psi\phi$ analysis as optimised in [49] are listed in Table 3.2. They can be divided into three categories, the J/ψ selection, the ϕ selection and the B_S^0 selection.

J/ψ selection
μ minimum $\Delta \ln \mathcal{L}_{\mu\pi} > -5$
μ track maximum $\chi^2_{\text{track}}/\text{nDoF} < 5$
μ minimum $p_T > 500$
J/ψ $\chi^2_{\text{vtx}}/\text{nDoF} < 6$
J/ψ $p_T > 1$ GeV
$\lvert M(\mu\mu) - M(\text{J}/\psi)\rvert < 42$ MeV (3σ)
ϕ selection
K minimum $\Delta \ln \mathcal{L}_{\text{K}\pi} > 0$
K $p > 2$ GeV
K track maximum $\chi^2_{\text{track}}/\text{nDoF} < 10$
ϕ $\chi^2_{\text{vtx}}/\text{nDoF} < 20$
ϕ $p_T > 1$ GeV
$\lvert M(\text{KK}) - M(\phi)\rvert < 12$ MeV (3σ)
B_S^0 selection
B_S^0 $\chi^2_{\text{vtx}}/\text{nDoF} < 5$
$\lvert M(B_S^0) - M(\text{J}/\psi\phi)\rvert < 300$ MeV
$\text{IPS}(B_S^0) < 5$

Table 3.2: *Selection criteria for the decay $B_S^0 \to J/\psi\phi$ [49]*

J/ψ selection

For the J/ψ selection, the most powerful selection quantity is the particle ID variable $\Delta \ln \mathcal{L}_{\mu\pi}$ (for the definition of delta log-likelihood see Chapter 2) which is chosen to be larger than -5. This allows to separate real muons from hadrons (mostly pions, misidentified as muons). The cut reduces the background by roughly 50%. As this selection criterion has barely an effect on rejecting hadrons that decay before the Muon Chambers (which is the largest contribution to hadrons misidentified as muons), a cut on the χ^2 of the track reconstruction and the transverse

momentum of the muon is applied. Similarly, cuts on the transverse momentum and the χ^2 of the vertex reconstruction of the J/ψ are applied. J/ψ candidates are chosen in such a way that the invariant mass lies within a 3σ (±42 MeV) window with respect to the nominal value (3097 MeV).

Just for illustration purposes, in Figure 3.1 the reconstructed J/ψ mass for the $\nu = 1$ sample is plotted. The plot relies on truthmatched quantities. For the exact definition of truthmatched and further information on the data set see section 3.3.
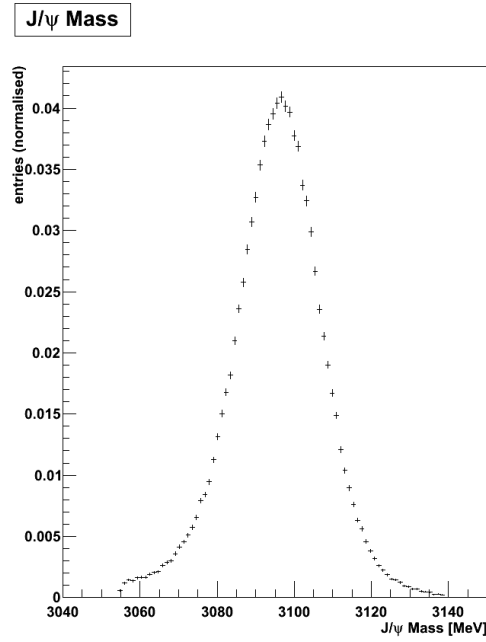
Figure 3.1: *Reconstructed J/ψ mass for $\nu = 1$. The distribution is normalised to an area of 1.*

ϕ selection

Only events surviving the previous selection are used as input for the ϕ selection. This helps to partly remove combinatorial background from prompt (directly in the proton-proton interaction produced) pions and kaons. To identify the kaons from the decay of ϕ mesons a cut on $\Delta \ln \mathcal{L}_{K\pi}$ is applied. Furthermore, a cut on the kaon momentum and χ^2 of the track reconstruction is required. To avoid biases of the angular acceptances only a p_T cut on ϕ itself is chosen instead of a p_T cut on the ϕ decay products. Also, a cut on the quality of the ϕ vertex reconstruction is applied by the condition $\chi^2_{\mathrm{vtx}}/\mathrm{nDoF} < 20$. Finally, the ϕ candidates must have a mass within the 3σ (±12 MeV) region around the nominal value (1019 MeV). In Figure 3.2 the reconstructed ϕ mass for the $\nu = 1$ sample is plotted (again truthmatched distribution).
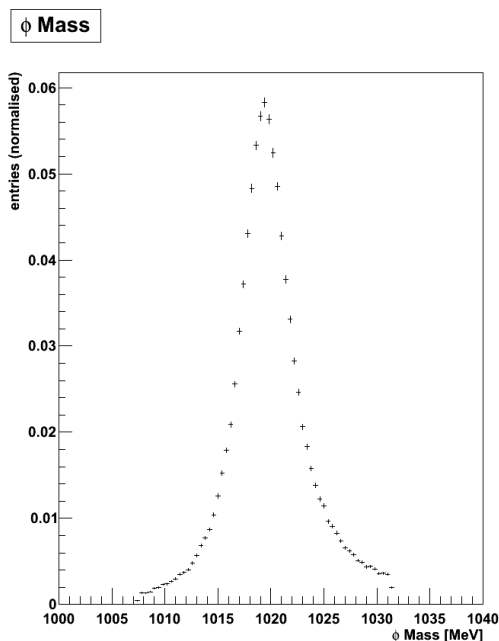
Figure 3.2: *Reconstructed ϕ mass for $\nu = 1$. The distribution is normalised to an area of 1.*

B_S^0 selection

After applying the J/ψ and the ϕ selection only three cuts are used for the B_S^0 selection: a cut on the χ^2 of the B_S^0 decay vertex reconstruction and on the IP significance (for a definition see Chapter 2) with respect to the primary vertex and a mass window for the B_S^0 of ± 300 MeV. The reconstructed B_S^0 mass is shown in Figure 3.3 (truthmatched distribution). A more detailed discussion of the mass distribution can be found in section 3.3.

Event yield

To calculate the expected annual event yields one first has to consider the envolved branching fractions: The B_S^0 decays with a branching ratio $BR(B_S^0 \to J/\psi\phi) = (9.3 \pm 3.3) \cdot 10^{-4}$ [49] into a J/ψ and a ϕ. The decay of the J/ψ into a muon pair has a branching fraction of $BR(J/\psi \to \mu\mu) = (5.93 \pm 0.06) \cdot 10^{-2}$ [49] whereas the ϕ decays into a kaon pair with the ratio $BR(\phi \to KK) = (49.2 \pm 0.6) \cdot 10^{-2}$ [49]. The product of these three branching ratios gives the total visible branching ratio BR_{vis}. Consequently, the annual event yield can be computed according to (see [49])

$$\mathcal{S} = \mathcal{L}_{int} \cdot \sigma_{b\bar{b}} \cdot 2 \cdot f_B \cdot BR_{vis} \cdot \epsilon_{tot}, \qquad (3.2)$$
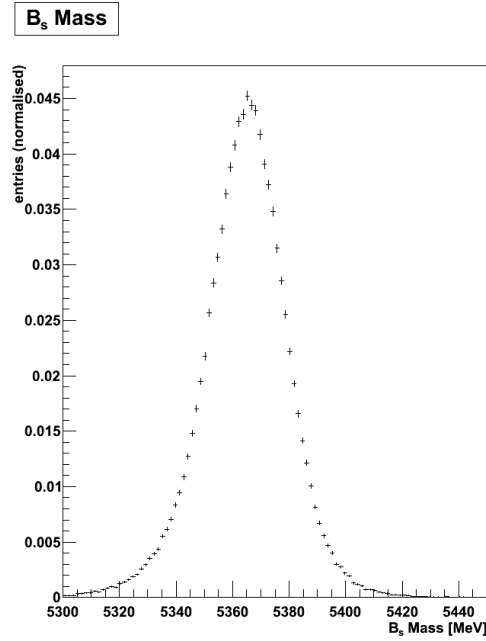
Figure 3.3: *Reconstructed* B_S^0 *mass for* $\nu = 1$. *The distribution is normalised to an area of 1.*

where $\mathcal{L}_{int} = 2$ fb^{-1} is the integrated luminosity for one nominal year of data taking (10^7 s) at a luminosity of $L \sim 2 \cdot 10^{32}$ cm^{-2}s^{-1} on average, and $\sigma_{b\bar{b}} = 500$ μb is the b$\bar{b}$ cross section at 14 TeV. The factor two accounts for pair production. $f_B = 11.0\%$ is the hadronisation fraction for b $\to$ B$_S^0$, BR$_{vis} = 2.71 \cdot 10^{-5}$ the total visible branching ratio and $\epsilon_{tot} = 1.96\%$ the total efficiency after selection and triggers. This gives a yield per nominal year of roughly 117 000 events. Note that the numbers are taken from [49] where a data set from the 'Data Challenge 2006' and not from the 'Monte Carlo 2009' production was used and that the uncertainties especially for $\sigma_{b\bar{b}}$ and BR$_{vis}$ are large (of the order of 40%). The number for the annual event yield can only be a rough estimate. These numbers are also used in the Feldman Cousins analysis explained and discussed in the following chapters.

The b$\bar{b}$ cross section depends, of course, on the centre-of-mass energy. For the calculation above a centre-of-mass energy of 14 TeV has been assumed resulting in a cross section of $\sigma_{b\bar{b}} = 500$ μb. At 10 TeV, however, the cross section is reduced to 357 μb (71% compared to 500 μb) and to 219 μb at 7 TeV (reduction to 44% with respect to the nominal cross section). The expected annual event yield drops accordingly.

3.3 Signal distributions, tagging and trigger efficiencies

To test the quality of the reconstruction an original particle created in the simulation has to be associated to every reconstructed particle. These particles are called truthmatched. All distributions and numbers presented in this section rely on truthmatched quantities. The $\nu = 1$ sample comprises roughly 213 000 and the $\nu = 3$ sample circa 177 000 reconstructed and truthmatched B_S^0 or $\overline{B}_S^0$ particles. The exact numbers can be found in Table 3.3.

quantity	$\nu = 1$	$\nu = 3$
number of generated events	1 515 052	1 302 611
number of reconstructed $B_S^0/\overline{B}_S^0$	226 412	188 776
offline reconstruction efficiency	$(14.9 \pm 0.1)\%$	$(14.5 \pm 0.1)\%$
number of thruthmatched $B_S^0/\overline{B}_S^0$	213 410	176 553
fraction of non-associated $B_S^0/\overline{B}_S^0$ candidates	$(5.5 \pm 0.1)\%$	$(6.3 \pm 0.1)\%$

Table 3.3: *Reconstruction performance for the $\nu = 1$ and the $\nu = 3$ sample*

The main performance quantities are, firstly, the offline reconstruction efficiency which is defined as $\epsilon_{off} = \dfrac{\text{number of reconstructed } B_S^0/\overline{B}_S^0}{\text{number of generated events}}$ (assuming that there is one $B_S^0/\overline{B}_S^0$ per event). This number decreases slightly for $\nu = 3$. The second main performance quantity is the fraction of wrongly associated (non-truthmatched) B_S^0 or $\overline{B}_S^0$ candidates. This latter number is significantly higher for the $\nu = 3$ sample. This means that the the reconstruction is not as pure as for the $\nu = 1$ case.
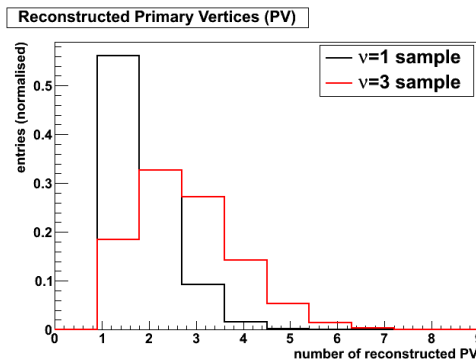

Figure 3.4: *Number of reconstructed primary vertices. Both distributions are normalised to an area of 1.*

The number of reconstructed primary vertices is shown in Figure 3.4. The difference between the two samples ($\nu = 1$ and $\nu = 3$) is as expected. Note that "zero" bin entries are not included.

Two quantities which are of importance for the analysis are the resolution of the reconstructed $\mathrm{B_S^0}$ mass and the resolution of the reconstructed $\mathrm{B_S^0}$ lifetime. Both distributions can be seen in Figure 3.5.
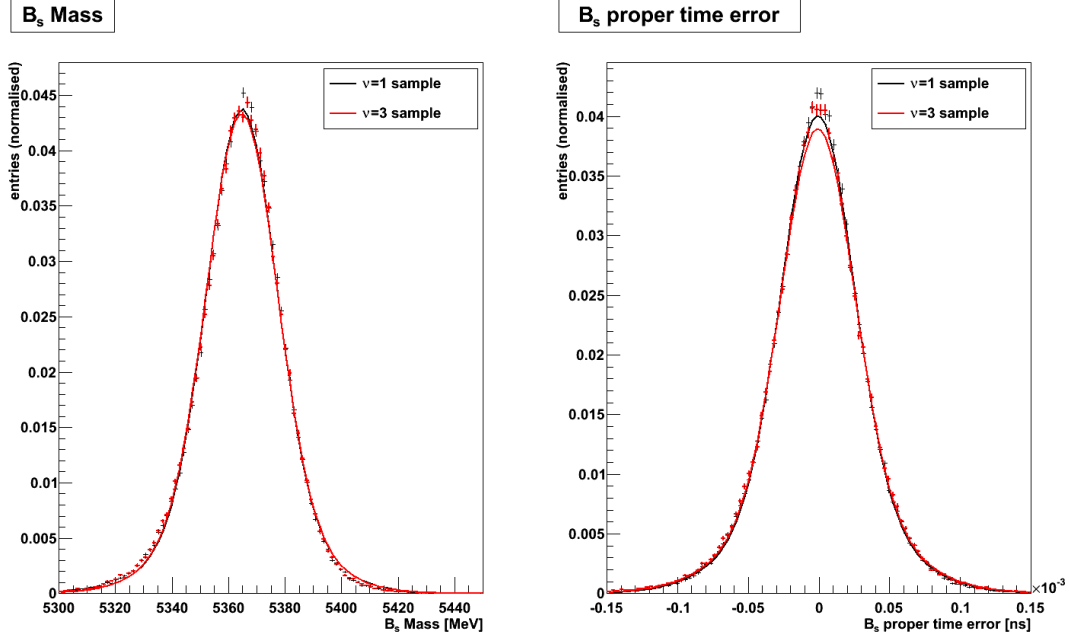

Figure 3.5: $\mathrm{B_S^0}$ *mass and lifetime error distribution for the $\nu = 1$ and the $\nu = 3$ sample using two Gaussians. Both distributions are normalised to an area of 1.*

For the mass distribution the sum G_D of two Gaussians G_1 and G_2 is fitted to the data points such that

$$G_D = f_{\sigma_{m,1}} \cdot G_1(m, \sigma_{m,1}) + (1 - f_{\sigma_{m,1}}) \cdot G_2(m, \sigma_{m,2}) \qquad (3.3)$$

where $\sigma_{m,1/2}$ are the widths of the two Gaussians and $f_{m,1}$ the weight of the first Gaussian. m is the mean reconstructed mass which should be the same for the two Gaussians. The total resolution is usually calculated as the quadratic average $\sigma_m = \sqrt{f_{m,1}\sigma_{m,1}^2 + (1 - f_{m,1})\sigma_{m,2}^2}$. The results

- $m = 5364.7$ MeV, $\sigma_m = 15.3$ MeV for $\nu = 1$ and

- $m = 5364.8$ MeV, $\sigma_m = 15.4$ MeV for $\nu = 3$

show that the number of proton-proton interactions per bunch crossing has practically no influence on the mass resolution of the $\mathrm{B_S^0}$.

For calculating the resolution of the B_s^0 lifetime (proper time) the lifetime error $t_{err} = t(B_s^0) - t(B_{s,\,MC}^0)$ is computed where $t(B_s^0)$ is the reconstructed lifetime and $t(B_{s,\,MC}^0)$ the one of the corresponding Monte Carlo particle. As already described in Chapter 2 the lifetime is calculated by

$$t(B_s^0) = \left(\vec{d} \cdot \frac{\vec{p}}{|\vec{p}|} \right) \frac{M_{B_s}}{|\vec{p}|\,c}. \tag{3.4}$$

Applying the same procedure as for the mass resolution the lifetime error is fitted using two Gaussians. Accordingly, one obtains

- mean $t_{err} = -0.70$ fs (lifetime bias), $\sigma_{ct} = 34.3$ fs for $\nu = 1$ and

- mean $t_{err} = -0.74$ fs (lifetime bias), $\sigma_{ct} = 35.4$ fs for $\nu = 3$.

With increasing number of proton-proton interactions the resolution is slightly degrading, the performance for $\nu = 3$, however, is not significantly worse.

Two issues remain to be addressed. First of all it has to be checked whether there is an influence on the reconstructed B_S^0 lifetime if there is more than one primary vertex in the event. In the case of more than one primary vertex there is the possibility that a wrong primary vertex with respect to the B_S^0 is reconstructed. If this happens frequently, there might be an effect visible in the B_S^0 lifetime distribution. As the average number of reconstructed primary vertices is higher for the $\nu = 3$ sample, such an effect should be more pronounced for the high luminosity sample. To check for any differences between the two samples the lifetime error distribution in Figure 3.5 has to be considered once more. In Figure 3.5 (right) an effect is not visible. The widths $\sigma_{ct,1}$ and $\sigma_{ct,2}$ of the two Gaussians fitted to each of the lifetime error distributions are determined to be

- $\sigma_{ct,1} = 48.81 \pm 0.08$ fs; $\sigma_{ct,2} = 24.91 \pm 0.08$ fs for $\nu = 1$ and

- $\sigma_{ct,1} = 50.72 \pm 0.08$ fs; $\sigma_{ct,2} = 25.47 \pm 0.08$ fs for $\nu = 3$.

The weight of the first Gaussian is 0.32 ± 0.01 ($\nu = 1$) and 0.31 ± 0.01 ($\nu = 3$) respectively. Apart from the fact that the lifetime resolution is slightly worse for the $\nu = 3$ sample there is no difference between the two samples. If there are any, wrongly determined primary vertices have no impact on the extraction of the B_S^0 lifetime.

The second issue that has to be considered is the problem of combinatoric background. As in this section only truthmatched distributions are shown, there is basically no combinatoric background visible in the B_S^0 mass spectrum (Figure 3.5, left). This is because in this section only a signal sample is used. Unfortunately, the corresponding background samples were not generated for $\nu = 3$. And even if the non-associated candidates are included, it has been varified that there is no significant difference visible. However, the fraction of non-associated B_S^0 or $\overline{B}_S^0$

candidates from Table 3.3 might give a hint. At $\nu = 3$ this fraction is by 15% larger compared to the $\nu = 1$ sample. On the other hand a much higher luminosity is available in the case of $\nu = 3$ which compensates for the larger combinatoric background.

Distributions for daughter particles

To check that the distributions of key quantities of the daughter particles remain unchanged between the two luminosity samples, the J/ψ momentum and the reconstructed ϕ mass are plotted in Figure 3.6. Again, no significant difference can be seen.
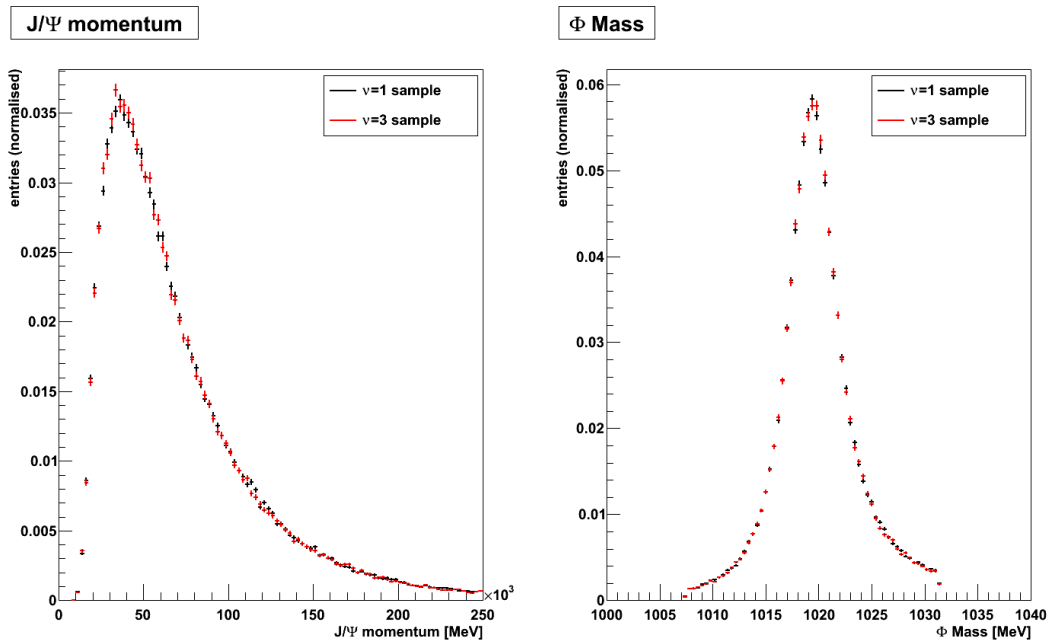
Figure 3.6: *Key quantities of the daughter particles: J/ψ momentum and ϕ mass for $\nu = 1$ and $\nu = 3$. Both distributions are normalised to an area of 1.*

Angular distributions

Finally, the distributions of the B_S^0 transversity angles $\cos\psi$, $\cos\theta$ and φ, defined in Chapter 1, are plotted in Figure 3.7. Within the errors there is no difference between the two samples visible. For both $\cos\theta$ distributions the bin around $\cos\theta \approx -0.6$ is significantly lower than expected. The reason for this reconstruction inefficiency could not be fully clarified. It might, however, be an artefact.
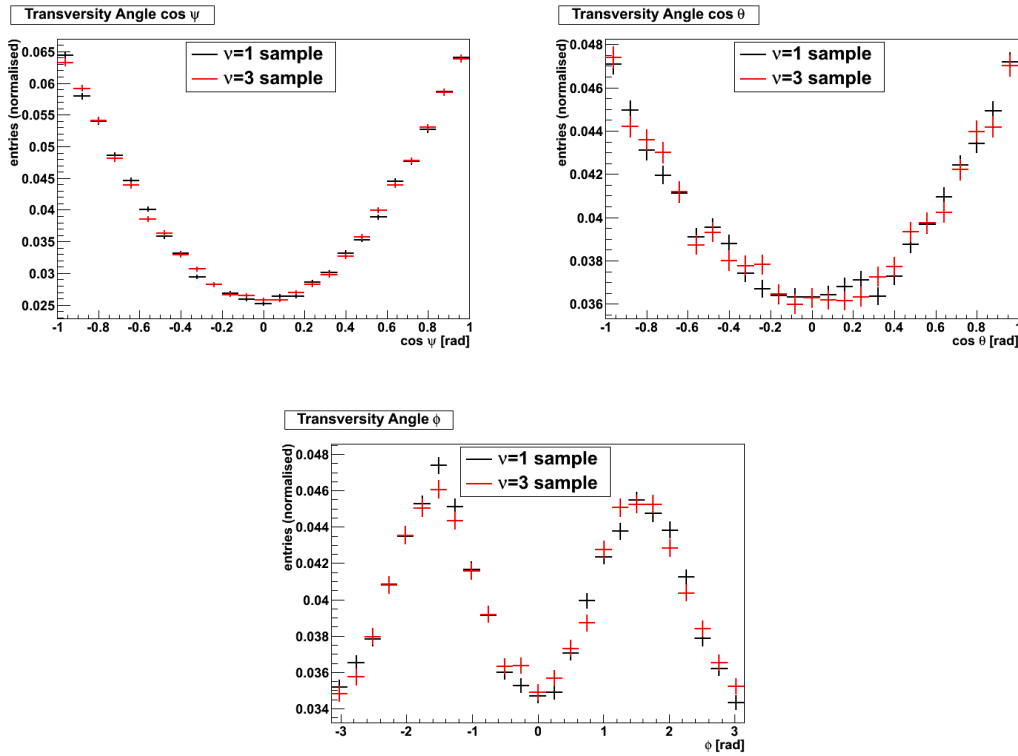
Figure 3.7: *Distributions of the* B_s^0 *transversity angles* $\cos\psi$, $\cos\theta$ *and* φ. *The distributions are normalised to 1. Note that the y-axis is zero-suppressed.*

Tagging and trigger efficiencies

Flavour tagging is very crucial in order to extract the mixing phase from the data. Therefore, it should be checked that the tagging performance for $\nu = 3$ is not significantly worse than for the $\nu = 1$ sample. To do this, the performance of the tagging algorithms was studied for both cases. The tagging efficiency (in %) and the mistag probability (in %) are listed in Table 3.4 for the two samples. These numbers show that the increase to $\nu = 3$ has, indeed, an impact on the tagging performance. The average mistag rate for the combination of the five taggers increases. Mainly responsible is the poorer performance of the OS kaon and the vertex tagger. The increase in ω is accompanied by a higher tagging efficiency. Since ϵ only contributes linearly to $\epsilon_{eff} = \epsilon(1-\omega)^2$, the mistag rate is the crucial quantity and consequently, the effective tagging power of the combination of the five taggers is decreasing from 4.6% for $\nu = 1$ to 3.6% in the $\nu = 3$ case. This corresponds to a loss of roughly 20%. However, an increase in the number of proton-proton interactions per bunch crossing from 1 to 3 corresponds to a luminosity increase by a factor of 3 (assuming the same running conditions). This in a way compensates for the loss of effective tagging power.

	$\nu = 1$		$\nu = 3$	
	$\omega[\%]$	$\epsilon[\%]$	$\omega[\%]$	$\epsilon[\%]$
OS kaon tagger	34.0±0.2	17.3±0.1	36.6±0.3	18.4±0.1
SS kaon tagger	36.1±0.1	27.3±0.1	36.2±0.1	26.2±0.1
electron tagger	33.7±0.5	3.8±0.05	34.3±0.6	3.7±0.04
muon tagger	32.9±0.4	6.1±0.05	33.7±0.5	6.2±0.06
vertex tagger	41.8±0.2	42.8±0.1	43.5±0.1	49.9±0.1
Combination of all five tagger categories				
average ω and ϵ	35.7±0.1	56.3±0.1	37.6±0.1	59.6±0.1
average dilution D	28.6±0.3		24.7±0.3	
average eff. tagging power ϵ_{eff}	4.6±0.1		3.6±0.1	

Table 3.4: *Mistag fractions ω and tagging efficiencies ϵ for $\nu = 1$, $\nu = 3$ (in %)*

For the trigger rates a similar conclusion can be drawn. In Table 3.5 the trigger efficiencies for all three triggers (L0, HLT1, HLT2) are listed.[5] The change in the efficiencies for the L0 and the HLT2 trigger are barely significant, there is, however, a drop in the HLT1 efficiency. This drop can be explained in the following way: With increasing number of proton-proton interactions per bunch crossing and therefore, luminosity the trigger rate is expected to go up. On the other hand, however, the trigger thresholds remain the same. This leads to inefficiencies. The numbers, however, show that this effect does not play a significant role.

To summarise, an increase of the number of proton-proton interactions per bunch crossing from 1 to 3 does not significantly worsen the detector performance. As far as the analysis of simulated $B_s^0 \to J/\psi\phi$ events shows, the operation of the detector at a higher luminosity is possible.

Trigger efficiency		$\nu = 1$	$\nu = 3$
L0:	$\dfrac{\text{\# reconstructed and L0-passed } B_s^0/\overline{B}_s^0}{\text{\# reconstructed } B_s^0/\overline{B}_s^0}$	99.7 ± 0.01	99.8 ± 0.01
HLT1:	$\dfrac{\text{\# reconstructed and L0+HLT1-passed } B_s^0/\overline{B}_s^0}{\text{\# reconstructed and L0-passed } B_s^0/\overline{B}_s^0}$	92.2 ± 0.05	91.5 ± 0.06
HLT2:	$\dfrac{\text{\# reconstructed and L0+HLT-passed } B_s^0/\overline{B}_s^0}{\text{\# reconstructed and L0+HLT1-passed } B_s^0/\overline{B}_s^0}$	97.7 ± 0.03	97.4 ± 0.03

Table 3.5: *Trigger performance for $\nu = 1$, $\nu = 3$ (in %)*

---

[5]For the HLT the trigger scenario 'Physics10000Vis1000L040Hlt1EffectiveHlt2Jul09' has been used.

3.4 Background and acceptances

For completeness, the main sources of background and the behaviour of detector acceptances is briefly discussed in this section. As mentioned above a specific study of the background for $\nu = 1$ and $\nu = 3$ was not possible as the corresponding background samples were not generated.

Background

A realistic determination of background distributions from Monte Carlo simulations is difficult due to the restricted amount of events. The data set closest to reality, the so-called minimum bias (MB) sample, only corresponds to an integrated luminosity of roughly 10^{-3} pb^{-1}. Nonetheless, a rough estimate of signal to background ratios can be done using the MB as well as other samples. A precise background analysis on data can be done using sidebands of the mass distributions. There are two main categories of background (for more information see [49]):

- Prompt background comprises all reconstructed $B_S^0 \to J/\psi\phi$ events where all tracks originate in the primary vertex. Using the minimum bias sample, one can see that the dominant prompt background component is the combination of a prompt $J/\psi \to \mu\mu$ with two prompt kaons to a B_S^0 candidate. Therefore, the more abundant Monte Carlo sample of J/ψ inclusive events is used to analyse the prompt background contribution. After removing events with at least one track from long-lived hadrons the background to signal ratio for the prompt background is then calculated by interpolating the background evaluated in an enlarged mass window around the B_S^0 mass (±300 MeV) into the signal region (±50 MeV). In [49] a background to signal ratio of

$$\frac{B_{\text{prompt}}}{S} = (1.6 \pm 0.6).\tag{3.5}$$

  is obtained.

- Long-lived background comprises all reconstructed $B_S^0 \to J/\psi\phi$ events where at least one particle used for reconstruction originates from a b-hadron. Its properties are evaluated with $b\bar{b}$ inclusive and $B_q^0 \to J\psi X$ Monte Carlo samples. The proper time, mass and angular distributions are studied with the $B_q^0 \to J\psi X$ sample. The $b\bar{b}$ sample is used for an estimation of the long-lived background to signal ratio. In a mass window of ±50 MeV, the ratio according to [49] is

$$\frac{B_{\text{longlived}}}{S} = (0.51 \pm 0.08).\tag{3.6}$$

In the case of J/ψ particles the background is mainly composed of $J/\psi \to \mu\mu$ events from b decays and random combinations of muons coming from

pions and kaons whereas for ϕ particles the main background contribution are pions misidentified as kaons and protons. This explaines the choice of the $B_q^0 \rightarrow J\psi X$ sample to study the key distributions of the long-lived background.
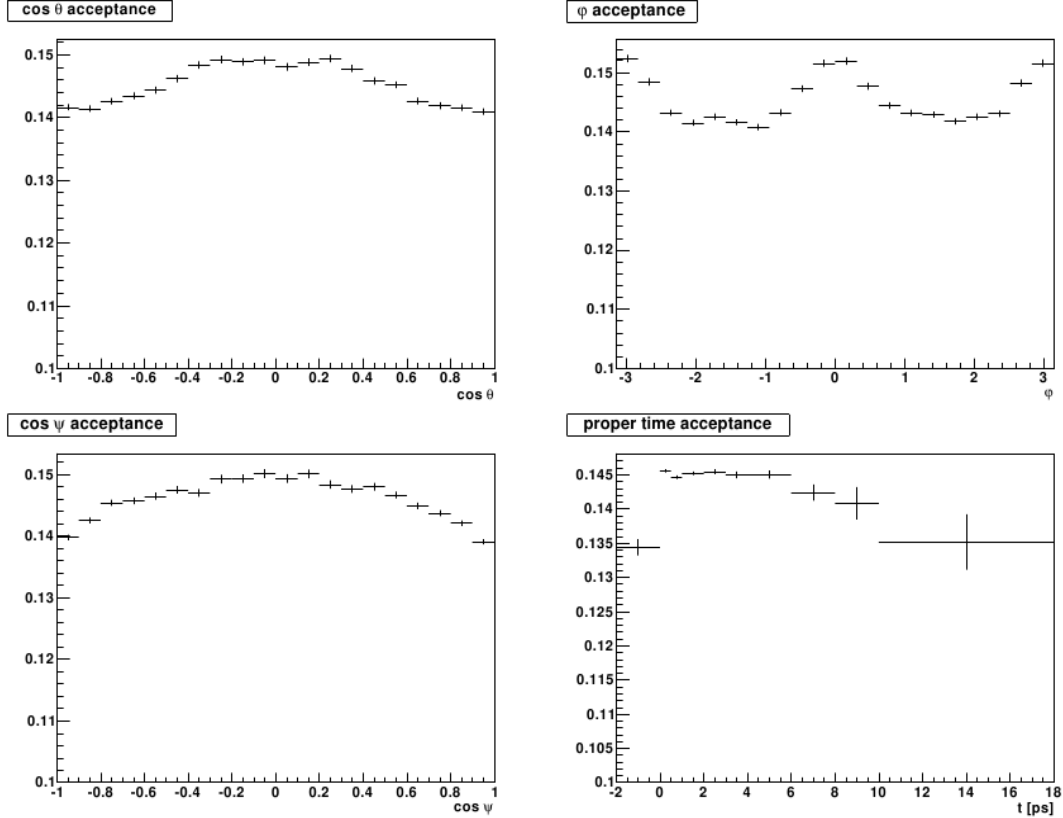
Acceptances

Figure 3.8: *Projections of the acceptances determined in [54]*

In addition, proper time and angular acceptances play an important role: Like every other detector the LHCb detector as well is not perfect. Therefore, the measured distributions in the transversity angles and the lifetime are strongly dependent on detector inefficiencies and geometrical effects. These effects have to be considered and compensated for. The method described in [54] uses four-dimensional histograms for the calculation. The acceptance is then simply the number of accepted Monte Carlo simulated events in the specific bin divided by the theoretical expectation:

$$\epsilon(t_i, \cos\theta_j, \varphi_k, \cos\psi_l) = \frac{\text{Accepted MC}(t_i, \cos\theta_j, \varphi_k, \cos\psi_l)}{\text{Theory}(t_i, \cos\theta_j, \varphi_k, \cos\psi_l) \otimes G_D(\sigma_1^t, \sigma_2^t)} \tag{3.7}$$

where the convolution with the sum of two Gaussians $G_D(\sigma_1^t, \sigma_2^t)$ accounts for proper time resolution effects. In the denominator one cannot use the number of generated events as usual because a cut on the angle between the beam axis and the direction of the decay products has been applied for the signal sample (see section 3.2), i.e. while producing the events the geometry of the detector has already been taken into account. Using the number of generated events would therefore distort the acceptance distributions. An extensive analysis of acceptances has been done in [54]. To get a rough idea how these acceptances are distributed the corresponding projections can be seen in Figure 3.8.

From the distributions in Figure 3.8 one can see that the deviations of the acceptances from being flat are smaller than 10%. The main reason for these deviations is that all four particle tracks (two muons, two kaons) are required to be reconstructed within the LHCb geometrical acceptance.

Extracting the B_s mixing phase for low event statistics

The determination of the mixing phase Φ from $B_s^0 \to J/\psi\phi$ decays is one of the key measurements of the LHCb experiment and allows for the indirect search for New Physics. Especially in the early phase of data taking it is essential to think about the correct determination of the errors and confidence levels in order to correctly interpret the data. While for high statistics the usual assumption of Gaussian errors is valid this cannot be, in general, assumed for a low statistics sample. The Feldman Cousins method [11] is able to correctly determine confidence regions in the case of low statistics. In this chapter the Feldman Cousins method is motivated and explained. In addition, it is investigated whether the Feldman Cousins approach is needed for the analysis of early data at LHCb or if a classical method to calculate confidence intervals is sufficient. The focus is on low event statistics corresponding to integrated luminosities of 10 to 50 pb^{-1}. The chapter starts with a short introduction to basic statistical definitions.

4.1 Basic statistical definitions

There are several ways to define *probability* [55]. In an empirical definition one considers an observation which is performed n times under the same conditions. The observations shall be independent from each other. If a property $\mathcal{A}$ is observed k times, then the probability for $\mathcal{A}$ to occur is (see [55])

$$W(\mathcal{A}) = \lim_{n \to \infty} \frac{k}{n}. \tag{4.1}$$

A *continuous random variable* can assume a continuum of different values [55]. It is the outcome of a measurement or observation and can be described by a real number x. The probability that the measurement of x gives a value $a \le x \le b$ is defined as

$$W(a \le x \le b) = \int_a^b \mathcal{P}(x)\mathrm{d}x \tag{4.2}$$

where $\mathcal{P}(x)$ is the probability density function (PDF) for the variable x. The PDF has to obey the relations

$$\mathcal{P}(x) \geq 0, \qquad \int_{-\infty}^{\infty} \mathcal{P}(x)\mathrm{d}x = 1. \tag{4.3}$$

Using a *maximum likelihood method* one or several unknown parameters λ of a theoretical distribution can be determined. One considers n measurements of the random variable x where x can be a single variable or a vector of variables [55]. It is assumed that the underlying PDF $\mathcal{P}(x|\lambda)$ is known. By applying the maximum likelihood method one can determine a best estimate for the parameters λ. The likelihood function is defined as

$$\mathcal{L}'(\lambda) = \prod_{i=1}^{n} \mathcal{P}(x_i|\lambda). \tag{4.4}$$

The best estimate for λ is that value that maximises $\mathcal{L}'(\lambda)$, i.e. the probability to obtain the observed set of values x_i. It is important that the normalisation condition

$$\int \mathcal{P}(x|\lambda)\mathrm{d}x = 1 \quad \text{for every } \lambda \tag{4.5}$$

is obeyed. Usually, this maximisation is done numerically by varying the unknown parameters until the likelihood function is maximised. This procedure is called fitting. During the fitting process certain unknown parameters can be fixed, i.e. they are not varied. This can be done in cases where the specific parameter is well-known from theory or previous experiments. Due to numerical reasons not the product over the PDFs is maximised but instead the negative logarithmic likelihood is minimised:

$$\mathcal{L}(\lambda) = -\sum_{i}^{n} \ln \mathcal{P}(x_i|\lambda). \tag{4.6}$$

4.2   Motivation

In 2007 CDF was able to extract the mixing phase Φ based on a tagged analysis of $\mathrm{B}_\mathrm{S}^0 \to \mathrm{J}/\psi\phi$ events for the first time (tagging is crucial for the extraction of Φ, see Chapter 1) [10]. When calculating confidence levels in the $\Delta\Gamma$-Φ plane (as shown in Figure 1.5), the CDF collaboration discovered that the standard method using two-dimensional likelihood profile scans (which is explained in Section 4.4 and is referred to as "classical method") leads to undercoverage [10][56]. This means that the obtained confidence contour does not result in the correct probability coverage but is smaller than desired (e.g. the 95% contour shows in fact a confidence level smaller than 95%). Consequently, instead of using the classical
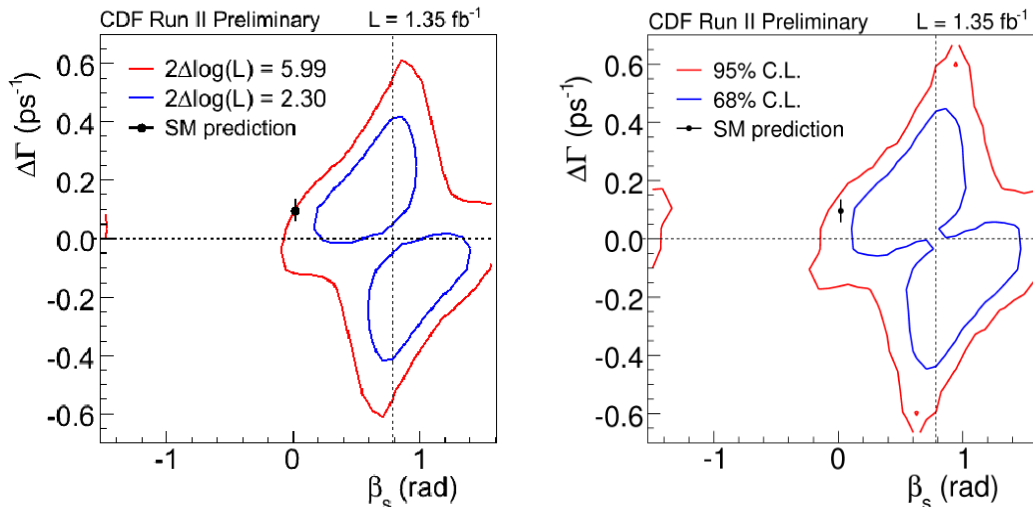
Figure 4.1: *CDF Run II results: Two-dimensional likelihood profile corresponding to 68% and 95% confidence levels (left) and respective confidence regions obtained by the Feldman Cousins method (right) [56]. Note that $2\Delta log(\mathcal{L})$=2.30 corresponds to a confidence level of 68% and $2\Delta log(\mathcal{L})$=5.99 to a confidence level of 95%.*

method, the Feldman Cousins method, providing correct coverage by construction, has been applied (see Section 4.5). To illustrate the difference between the two approaches the results for both methods are shown in Figure 4.1. The graphs are from CDF Run II with an integrated luminosity of 1.35 fb^{-1} [56]. Note that $\Phi^{SM} = -2\beta_S$.

The difference between the two plots is clearly visible. On the left plot the Standard Model point lies exactly on the 95% contour [6] whereas on the Feldman Cousins plot it is well inside the red contour. In principle, this undercoverage can be understood: The CDF results are based on about 2000 signal events only (corresponding to roughly 35 pb^{-1} at LHCb). At low statistics the errors are not Gaussian and the corresponding likelihood profiles are highly non-parabolic and have multiple not well-separated minima (see Figure 4.1). Under such conditions one cannot, at least a priori, expect the likelihood contour to give the correct confidence regions. In fact, it is known that the classical method yields under-coverage in certain circumstances [57].

It is worth checking whether the CDF observations are also seen in the case of simulated low statistics LHCb data and in which cases the Feldman Cousins method has to be used instead of the classical likelihood scan method.

---

[6]For the two-dimensional likelihood scan the 95% confidence level corresponds to the contour defined by $2\Delta log(\mathcal{L})$=5.99. For further explanation see Section 4.4

4.3   The unbinned likelihood fit for the extraction of the mixing phase

To extract the mixing phase and to apply the Feldman Cousins method an unbinned likelihood fit described in [54] is used. As it is an integral part of the Feldman Cousins analysis, this fitter framework is briefly summarised.

As described in Section 4.1 the likelihood function is defined as the product of the probability density functions $\mathcal{P}$ for every single event i. As explained above not the product over the PDFs is maximised but instead the negative logarithmic likelihood is minimised:

$$\mathcal{L} = - \sum_{i}^{all\ events} \ln\left(\mathcal{P}_i(m, t, \cos\theta, \varphi, \cos\psi, q | \vec{\lambda})\right) \tag{4.7}$$

The event probability depends on the reconstructed B_S^0 mass m, the proper time t, the transversity angles $(\cos\theta, \varphi, \cos\psi)$ and the tagging decision q (+1 for a B_S^0, -1 for a $\overline{B}_S^0$ and 0 for untagged events). $\vec{\lambda}$ denotes all parameters that are determined by the fit. The minimisation is performed using MINUIT (see [58]). The key ingredients of the likelihood parametrisation are as follows

- The total PDF $\mathcal{P}_i$ for every event has of two parts, the signal and the background part with f_{sig} denoting the signal fraction:

$$\mathcal{P}_i = f_{sig} \cdot \mathcal{S}_i(m, t, \cos\theta, \varphi, \cos\psi, q) + (1 - f_{sig}) \cdot \mathcal{B}_i(m, t, \cos\theta, \varphi, \cos\psi, q) \tag{4.8}$$

- Assuming that the mass does not depend on other quantities, the mass-dependent term in the signal PDF can be factorised. Then, the signal mass distribution can be described by a sum of two Gaussians. The relevant parameters are the ratio between the two Gaussians $f_{\sigma_{m,1}}$, the widths of the Gaussians $\sigma_{m,1}$ and $\sigma_{m,2}$ and the mean of both Gaussians m_{B_s}.

- The signal PDF is of the form of equation 1.18. Furthermore, it has to be considered that tagging is not perfect. To account for this a dilution term is added to the signal PDF (for the definition of dilution see Chapter 2). The angular and time dependent part can be written as

$$\mathcal{S'}_i(t, \cos\theta, \varphi, \cos\psi, q) \propto \frac{1 + qD}{2} \sum_{k=1}^{6} h_k(t) \cdot f_k(\cos\theta, \varphi, \cos\psi) + \\ \frac{1 - qD}{2} \sum_{k=1}^{6} \overline{h}_k(t) \cdot f_k(\cos\theta, \varphi, \cos\psi) \tag{4.9}$$

where D is the dilution and $h_k(t)$, $\overline{h}_k(t)$ and $f_k(\cos\theta, \varphi, \cos\psi)$ are the terms defined in Chapter 1.

- As explained in Chapter 3 there are two categories of background. The prompt background has a peak at a proper time of $t = 0$. The longlived background can be described by two exponentials with lifetimes $\tau_{ll,1}$ and $\tau_{ll,2}$ (fraction of the longlived part is f_{ll}). The mass distributions of the two types of background are modelled as exponentials with different decay constants α.

- Due to the limited detector resolution the proper time PDF has to be convoluted with a resolution function for the proper time. As described in Chapter 2 this is done with two Gaussians in case of the signal part (see equation 2.8). For the background components one Gaussian of width $\sigma_{ct,prompt}$ and $\sigma_{ct,longlived}$ respectively is used. Angular resolution effects play no role and are not considered.

A fit can be performed either on "toy experiments" where only theoretical distributions of data points are generated with a random number generator according to the model explained above and physical processes like interactions with the detector are neglected, or on fully simulated Monte Carlo samples of the kind described in Chapter 3. The input values for the toy generator and the fits are summarised in Table 4.1.

Acceptances and angular resolutions are not taken into account. There are two different scenarios considered for the value of the mixing phase Φ, the Standard Model case and a New Physics scenario. In the Standard Model the value of Φ is -0.0368. The value in the New Physics scenario of Φ is 20 times larger than the Standard Model value and therefore, is close to the most likely value of the mixing phase according to the measurement by CDF.

Parameter Category	Model / Input Values
B_s mass	$m_{B_s} = 5369.6$ MeV Sum of two Gaussians for signal: $f_{\sigma_{m,1}} = 0.74$ $\sigma_{m,1} = 13.2$ MeV $\sigma_{m,2} = 22.5$ MeV Exponential shape for background: $\alpha_{m,prompt} = 0.0006$ MeV $\alpha_{m,longlived} = 0.001$ MeV
Amplitudes	$|A_0|^2 = 0.56$ $|A_\perp|^2 = 0.233$
Strong phases	$\delta_\parallel = -2.93$ $\delta_\perp = +2.91$
Mixing phase	Standard Model: $\Phi^{SM} = -0.0368$ New Physics: $\Phi^{NP} = -0.736 = 20 \cdot \Phi^{SM}$

Event fractions	For mass window ± 50 MeV and 2 fb^{-1} $N_{signal} = 117$k $N_{background,prompt} = 210.6$k $N_{background,longlived} = 58.5$k $N_{total} = 386.1$k
B$_s$ mixing quantities	$\Delta M = 17.77$ ps^{-1} $\Delta \Gamma = 0.049$ ps^{-1} $\Gamma = 0.68$ ps^{-1}
Proper time	Sum of two Gaussians for signal $\delta(0)$ for prompt background Two exponentials for longlived background: $f_{ll} = 0.22$ $\tau_{ll,1} = 1114$ fs $\tau_{ll,2} = 161$ fs
Proper time resolutions	Sum of two Gaussians for signal: $f_{\sigma_{ct,1}} = 0.85$ $\sigma_{ct,1} = 31.5$ fs $\sigma_{ct,2} = 66.7$ fs $\sigma_{ct,total} = 38.9$ fs Gaussian shape for background: $\sigma_{ct,prompt} = 44.0$ fs; $\sigma_{ct,prompt} = 66.0$ fs
Tagging quantities	$\omega_{signal} = 0.334$ $\epsilon_{signal} = 0.564$ $\epsilon_{background,prompt} = 0.3$; $\epsilon_{background,longlived} = 0.62$

Table 4.1: *Input values for the simulated toy experiments*

4.4   Two-dimensional likelihood scan

Before introducing the Feldman Cousins method in this section the standard two-dimensional likelihood profile scan ("classical method", see Section 4.2) is explained. It consists of three steps:

Step 1

The first step is a fit to data where all parameters are floating. Note that for the time being a single set of toy events is used as "data set". In Chapter 6 a sample of fully simulated Monte Carlo events serves as "data set". The fit results are the best fit values $\widehat{\Delta \Gamma}$, $\widehat{\Phi}$ and $\widehat{\xi}$ where ξ denotes all other parameters except $\Delta \Gamma$ and Φ, the defining parameters for the two-dimensional contour plots. The

parameters ξ are called nuisance parameters. After having performed the fit, the value of the likelihood function corresponding to the best fit values is calculated.

Step 2

In a second step a grid in the $\Delta\Gamma$-Φ plane is constructed. Unless stated otherwise, this is a $40{\times}40$ grid ($\Phi \in [-\frac{\pi}{2}, +\frac{\pi}{2}]$ and $\Delta\Gamma \in [-0.3, 0.5]$ in the Standard Model case, $\Phi \in [-\pi, +\frac{\pi}{3}]$ and $\Delta\Gamma \in [-0.5, 0.5]$ in the New Physics case). Then, for every grid point a fit to data with $\Delta\Gamma$ and Φ fixed to the grid value and the nuisance parameters floating is performed and again the corresponding likelihood value is computed.

Step 3

Having calculated the two different likelihood values (the one of the global fit and the other one separately for every grid point) one can determine for every grid point the logarithmic ratio of the two likelihoods:

$$LR_{data} = -2 \cdot \log \frac{\mathcal{L}(\Phi^*, \Delta\Gamma^*, \widehat{\xi})}{\mathcal{L}(\widehat{\Phi}, \widehat{\Delta\Gamma}, \widehat{\xi})}, \tag{4.10}$$

where $\widehat{\ }$ denotes that the parameter is kept freely in the fit and $*$ that the parameter is fixed to the grid value. The factor of 2 is convention.

This procedure corresponds to the classical two-dimensional likelihood scan. To every confidence level a respective likelihood ratio is assigned according to Table 4.2 (m stands for the number of estimated parameters). Note that in this case one has two-dimensional confidence regions. This means that a confidence level of 95% corresponds to a likelihood ratio of 5.99. As mentioned before the classical method can lead to undercoverage which has been observed by CDF (see the two plots of Figure 4.1). Although a rigorous mathematical argumentation

C.L. (%)	$m = 1$	$m = 2$	$m = 3$
68.27	1.00	2.30	3.53
90.00	2.71	4.61	6.25
95.00	3.84	5.99	7.82
95.45	4.00	6.18	8.03
99.00	6.63	9.21	11.34
99.73	9.00	11.83	14.16

Table 4.2: *Values of 2Δlog($\mathcal{L}$) corresponding to a coverage probability of C.L. for the joint estimation of m parameters. [14]*

cannot be given, a qualitative explanation can be found: At low statistics it is likely that the calculated minimum of the likelihood function is not at the true position. Therefore, the determination of the confidence levels by the fixed values in Table 4.2 does not give the correct coverage. The Feldman Cousins approach, however, does calculate the correct coverage by construction.

4.5   The Feldman Cousins method

The original motivation for Feldman and Cousins to develop the method described in [11] was the experimental search for neutrino oscillations. For this case the two authors compare their technique, known as Feldman Cousins method, with three alternative classical approaches. They come to the conclusion that only their method fulfills the three properties needed for a meaningful calculation of confidence regions: the method applied has to give proper coverage, has to be powerful in the sense that it can distinguish between a likely and an unlikely value (of Φ in the case of this thesis) and has to give physically useful results. The authors suggest to apply the Feldman Cousins method quite generally in cases where the confidence regions have a particularly complicated structure and the PDF has multiple local minima [11]. In this sense, the Feldman Cousins method can also be used to determine the two-dimensional confidence region in the $\Delta\Gamma$-Φ plane for $B_S^0 \rightarrow J/\psi\phi$ decays [56]. Using the fitter framework described in the previous section the Feldman Cousins method has been implemented in the following way. After having performed **steps 1 to 3** explained in the previous section one has to proceed with

Step 4

At every grid point a significant amount of toy experiments is generated with the values in Table 4.1 as generation parameters except for $\Delta\Gamma$ and Φ for which the values of the respective grid point is taken. Unless stated otherwise, 250 toy sets are generated. The procedure that follows is almost identical to the one described above for the calculation of LR_{data}. For every toy experiment and at every grid point one again fits twice: once with all parameters floating and once with $\Delta\Gamma$ and Φ fixed to the grid value (nuisance parameters floating). The starting values for the fits are the best fit values mentioned in the previous section. In the same way as above, a likelihood ratio, LR_{toy}, can be computed. This means for every grid point one gets a LR_{toy} distribution of 250 values, corresponding to the 250 toy experiments generated at every gridpoint.

As one can easily imagine, the required computing resources for calculating confidence levels with the Feldman Cousins method is enormous. In the case of a $40{\times}40$ grid and 250 toy sets per grid point 400 000 toy samples have to be gen-
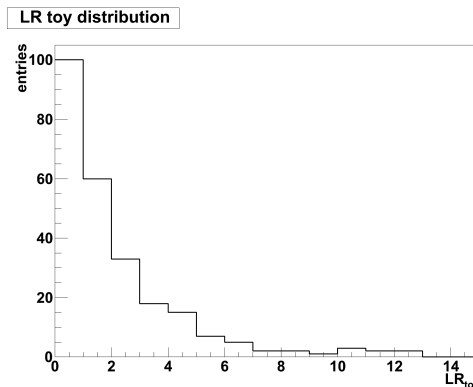
Figure 4.2: *Distribution of 250 toy likelihood ratios for the grid point* $(\Phi, \Delta\Gamma) = (-0.03, 0.05)$

erated and 800 000 fits (two per toy set and grid point) have to be performed. Therefore, only four nuisance parameters are used: $\delta_{\|}$, $\delta_{\perp}$, $|A_0|^2$ and $|A_{\perp}|^2$. All other parameters are fixed in the fits. Unless stated otherwise, this is the case for every contour plot presented in this thesis. Even with very low statistics of only 20 pb^{-1} and only these four nuisance parameters, the procedure for the whole grid and one of the 250 toy sets takes 4 hours.

To illustrate how a distribution of LR_{toy} looks like one can see a distribution of the 250 likelihood ratios for the grid point $(\Phi, \Delta\Gamma) = (-0.03, 0.05)$ in Figure 4.2. The total number of events generated in each toy experiment is 3861 (1170 signal events and 2691 background events) which corresponds to 20 pb^{-1} (1/100 of a nominal year).

The key point of the Feldman Cousins method is the calculation of the confidence levels. To do this the p-value p for every grid point is calculated. The p-value is the fraction of likelihood ratios LR_{toy} that exceed the likelihood ratio LR_{data} obtained from "dat" for the specific grid point. The corresponding confidence level is $1 - p$. In the case of the LR_{toy} distribution in Figure 4.2 the LR_{data} is determined to be 0.143. Consequently, in 237 out of 250 cases LR_{toy} exceeds LR_{data} for this specific grid point which leads to a p-value of 0.948 or a confidence level of $1 - p = 0.052$.

In Figure 4.3 two examples for a p-value distribution can be seen. For every grid point the p-value has been calculated as explained in the previous paragraph. The number of events generated in each toy experiment is again 3861 corresponding to 20 pb^{-1}. For the left plot the Standard Model (SM) value of Φ ($\Phi^{SM} = -0.0368$) has been used as input parameter in the toy generation and the fit, for the right plot a possible New Physics (NP) value ($\Phi^{NP} = -0.736 = 20 \cdot \Phi_{SM}$) has been used.

From such p-value distributions a confidence level can be determined for every grid point by calculating $1 - p$. These in such a way determined confidence levels
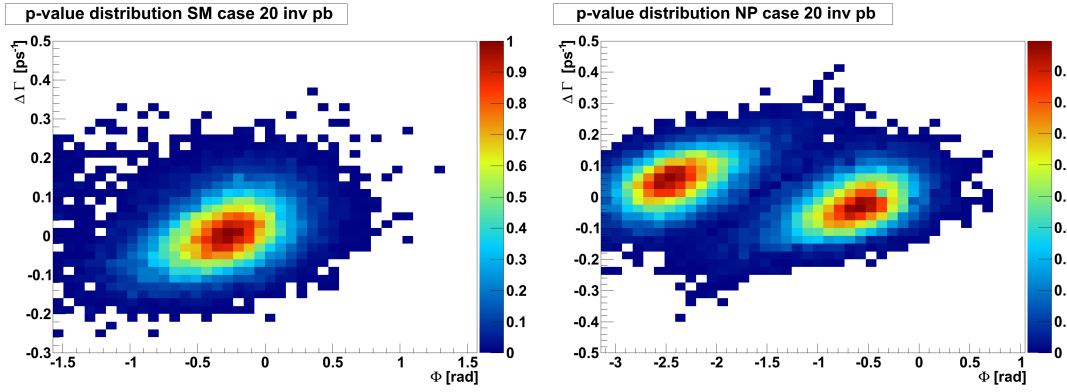
Figure 4.3: *p-value distributions corresponding to 20 pb^{-1}. A SM input value for Φ (left) and a NP input value (right) has been used respectively.*
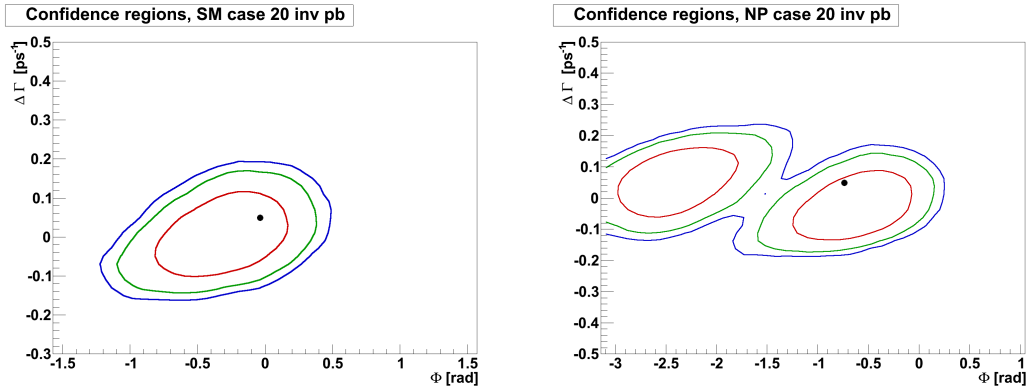
Figure 4.4: *68% (—), 90% (—) and 95% (—) contours for SM (left) and NP (right) input value of Φ. The dot denotes the production values.*

allow to draw confidence contours. Figure 4.4 shows the 68% (—), the 90% (—) and the 95% (—) contours corresponding to the p-value distributions plotted in Figure 4.3. The dot denotes the production values of the toy samples. Note that these graphs are based on an arbitrarily chosen "data set" and therefore can only visualise the qualitative behaviour of the confidence contours. The effect of different toy data sets is discussed in more detail in the next chapter. As mentioned above, in this chapter a "data set" is just a single set of toy experiments and not a sample of fully simulated Monte Carlo events.

One can see immediately that the sensitivity on the SM value of Φ is much better than on the NP value. To explain this one has to consider the decay rate which is discussed in Section 1.3. There it is mentioned that the CP violating phase can only be extracted in terms of sin Φ and cos Φ. In the SM case the mixing phase is close to 0. This means that cos Φ is nearly 1 and cannot be used for the extraction of Φ. On the other hand sin Φ is close to 0 which allows to extract the

mixing phase very well. In the NP case both $\sin\Phi$ and $\cos\Phi$ are of the order of 0.7. This means that the sensitivity on Φ becomes significantly worse than for the SM case. One can also look at the correlations between Φ and $\Delta\Gamma$. In [54] one finds that the correlation between the two parameters is 0.01 in the case of the SM value of Φ and 0.12 in the NP case. This leads to a reduced sensitivity and the resolution decreases.

As expected from the two-fold ambiguity described in Section 1.3, there are two solutions. As the two solutions in the SM case are well-separated, only one solution is shown in Figures 4.3 and 4.4.

Testing the method

The Feldman Cousins method is constructed in such a way that it automatically gives the right coverage. To test that the method is implemented correctly it has to be checked that it gives, indeed, the correct coverage. To do this a Feldman Cousins analysis has been applied to 1000 different "data sets". Every "data set" has been generated with the Standard Model value of Φ. This, however, requires a huge amount of CPU time because for every "data set" a full Feldman Cousins analysis has to be performed. Even with only four nuisance parameters this is not possible. Thus several changes have to be made for each of the 1000 experiments. Firstly, the test has only been done for 20 pb^{-1}. Secondly, only a 20×20 grid in the region $\Phi \in [-1.5, 1.0]$ and $\Delta\Gamma \in [-0.3, 0.3]$ is used. Thirdly, only 100 instead of 250 toy experiments per grid point are performed to calculate the p-values. And fourthly, no nuisance parameters are used, i.e. only $\Delta\Gamma$ and Φ are floating in half of the fits (and also fixed in the other half).

For every of the 1000 experiments it has been checked whether the production values for Φ and $\Delta\Gamma$ are within the 90% contour. If the Feldman Cousins method works correctly, this should be the case in 900 of the 1000 experiments. The measured coverage for the nominal 90% contour has been determined to be

$$(88.7 \pm 1.0)\%.$$

Within the errors one can conclude that the Feldman Cousins method is implemented correctly. Tests for different coverage probabilities and the New Physics value for Φ show similar results.

4.6 Undercoverage at low statistics

The key question of this analysis is whether at low statistics the problem of undercoverage explained in Section 4.2 is so significant that the Feldman Cousins method has to be applied to determine the contour levels. If the difference to the two-dimensional likelihood scan was negligible, a significant amount of CPU time could be saved by applying the classical method only. Especially for data
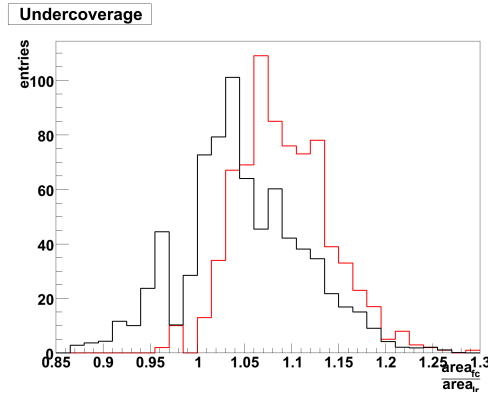
Figure 4.5: *Distribution of the ratio defined in equation 4.11 for 750 different toy experiments at 20 (——) and 50 pb^{-1} (——), SM case*

samples corresponding to more than 50 pb^{-1} the application of the Feldman Cousins method becomes very difficult.

There are different methods to check for undercoverage. One method has been proposed by CDF (see e.g. [56]). However, the method is not explained clearly and therefore, it could not be fully understood. Instead, an own technique has been developed: The size of the contours calculated by Feldman Cousins and by the classical two-dimensional likelihood scan is compared by calculating the ratio

$$\frac{\text{Area of the 90\% contour determined by Feldman Cousins}}{\text{Area of the 90\% contour determined by likelihood scan}}. \qquad (4.11)$$

This method has the advantage that by dividing the two areas statistical fluctuations cancel out. The same procedure is executed for a significant amount of "data sets" (750 in this case). As the Feldman Cousins method is constructed such that the corresponding confidence region has the right coverage (see previous section), a ratio significantly higher than 1 would imply an undercoverage of the contour calculated by the likelihood scan. As this procedure again requires a huge amount of CPU time, the same changes as explained at the end of the previous section ("Testing the method") have been applied. However, the same procedure has been not only performed for 20 pb^{-1} but also for 50 pb^{-1}.

In Figure 4.5 one can see the distribution of the ratio defined in equation 4.11 for the 750 different experiments at 20 (——) and 50 pb^{-1} (——). The scans are performed in the region $\Phi \in [-1.5, 1.0]$ and $\Delta\Gamma \in [-0.3, 0.3]$ for the Standard Model value of Φ. A spot test in the NP case has shown similar distributions.

One can see that an undercoverage occurs when using a two-dimensional likelihood scan at low statistics. At 20 pb^{-1} the mean ratio is 1.092±0.002 and it decreases to 1.047±0.002 at 50 pb^{-1}. As expected the undercoverage becomes less significant for increasing statistics. For 50 pb^{-1} the undercoverage is barely significant at all. From a linear extrapolation one would expect to get a ratio of

1 between 100 to 150 pb^{-1}.

However, one remark has to be made. A wide range of different degrees of under-coverage is possible for a single experiment. As one can see from Figure 4.5 the Feldman Cousins contour can be more than 20% larger but also slightly smaller than the corresponding likelihood contour (which would correspond to an over-coverage). So, the reason for the large discrepancy between the Feldman Cousins and the likelihood scanning method at CDF might be that their sample represents the unlucky case where the ratio in equation (4.11) is very large.

What does this imply for the LHCb analysis? First of all, at very low statistics (10 to 20 pb^{-1}) a Feldman Cousins analysis is inevitable in order to get the right coverage. With increasing data samples it becomes less important and beyond around 100 to 150 pb^{-1} (corresponding to 5850 and 8775 signal events respectively) the classical likelihood profile method is likely to be sufficient. 100 pb^{-1}, however, corresponds to 1/20 of a nominal year only and one can hope to achieve this integrated luminosity with the first physics run that is scheduled to end in autumn 2011. A final assessment, however, can only be made with the first data coming in.

To illustrate the difference between the two methods to calculate confidence regions, in Figure 4.6 the result of both methods (Feldman Cousins on the left and the likelihood scan on the right) is shown for one specific "data set". (for 20 pb^{-1} and a NP value of Φ). Again the 68% (—), the 90% (—) and the 95% (—) contour is plotted. The triangle denotes the SM point, the dot the NP point.
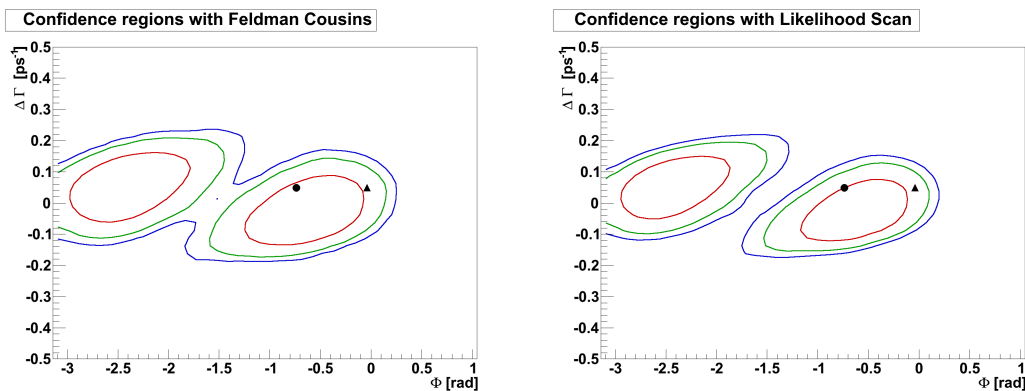
Figure 4.6: *68% (—), 90% (—) and 95% (—) contours with Feldman Cousins (left) and a likelihood scan (right) at 20 pb^{-1} and for the NP value of Φ) for one specific toy experiment as data set. The triangle denotes the SM point, the dot the NP value.*

The effect for this specific toy sample is very small. The 95% contours cor- responding to the two solutions are connected in the case of the Feldman Cousins analysis and well-separated in the case of the likelihood scan but the confidence level of the SM value does not change significantly. One can imagine that in the 50 pb^{-1} case this change would be barely visible.

In the next chapter the systematic uncertainties are examined. It is shown how the contours behave when changing key parameters like proper time resolution and tagging power.

Sensitivity studies for the extraction of the mixing phase

After having introduced the Feldman Cousins method and analysed the first $\Delta\Gamma$-Φ contour plots it is interesting to study how the sensitivity on the mixing phase evolves when changing key parameters like statistics, proper time resolution, mistag rate and data set. In the early phase of data taking it is not clear that one can achieve the nominal proper time resolution and mistag rate drawn from Monte Carlo studies. Therefore, it is important to check how the performance changes if these properties are, in fact, worse than expected. Note that the event yields are based on the calculations in Section 3.2, i.e. on the assumption of a centre-of-mass energy of 14 TeV.

5.1 Statistics

The first quantity that is discussed in this chapter is the number of events. Figure 5.1 shows the 95% contours for 10 (—), 20 (—) and 50 pb^{-1} (—), for the SM value of Φ as well as for the NP value. To ensure compatibility, for all the contours the same random seed [7] has been used in the event generator. The input values can be found in Table 4.1. As expected the contours get significantly smaller (i.e. Φ can be resolved better) with increasing statistics. In the NP case, however, one can see a change in orientation between the three contours: The two solutions get tilted with decreasing statistics. This effect, however, is dependent on the data set. Studying another data set this shift in orientation has been less pronounced. Thus, it is attributed to statistical fluctuations.

Obviously the LHCb results improve with more statistics. It is, however, very promising to see that even with statistics as low as 20 pb^{-1} (1170 signal events) a very clear contour can be produced. So, even within the first months of data taking one can be looking forward to observing interesting results.

---

[7]A seed or random seed is a number used to initialise a random number generator
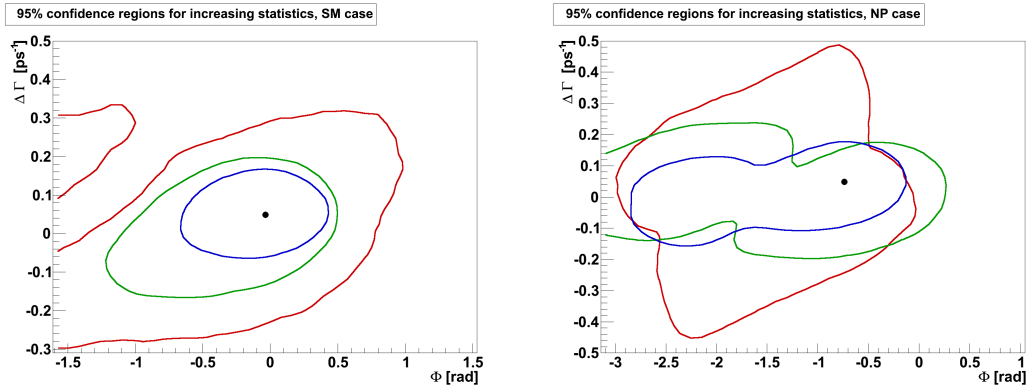
Figure 5.1: *95% contours corresponding to 10 pb^{-1} (—), 20 pb^{-1} (—) and 50 pb^{-1} (—) for SM (left) and NP (right) input value of Φ. The dot denotes the production values.*

To see how the confidence regions evolve for much higher statistics, Figure 5.2 shows the 95% contours for 200 pb^{-1}, for the SM value (left) and the NP value (right) of Φ, again using the same random seed as before.
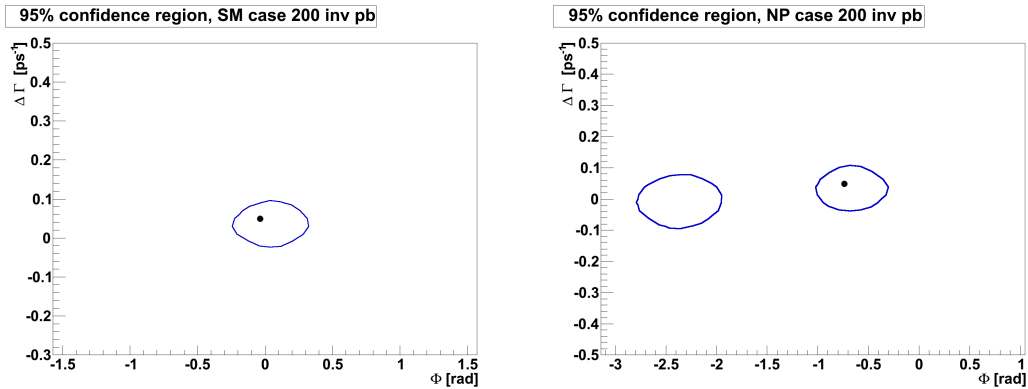
Figure 5.2: *95% contours corr. to 200 pb^{-1} for the SM (left) and the NP (right) value of Φ. The dot denotes the production values.*

The figure shows very small and nearly perfectly elliptical contours. In the NP case the two contours are well-separated. An integrated luminosity of 200 pb^{-1} only corresponds to 1/10 of a nominal year (11700 signal events).

5.2   Proper time resolution

One of the most important detector quantities for the extraction of the B$_S^0$ mixing phase is the proper time resolution σ_{ct}. As already mentioned in the previous chapter, a model using two Gaussians is assumed for the signal part and a single Gaussian for the prompt and longlived background. The nominal quantities are

listed in Table 4.1.

In Figure 5.3 one can see how the 95% contours change if the proper time resolutions are by 20 (—) and 50% (—) larger than the respective nominal values (—). In the case of the signal proper time resolution this would correspond to 38.9, 46.7 and 58.4 fs. The number of events corresponds to 20 pb^{-1}. The same random seed has been used for the generation of the data sets.
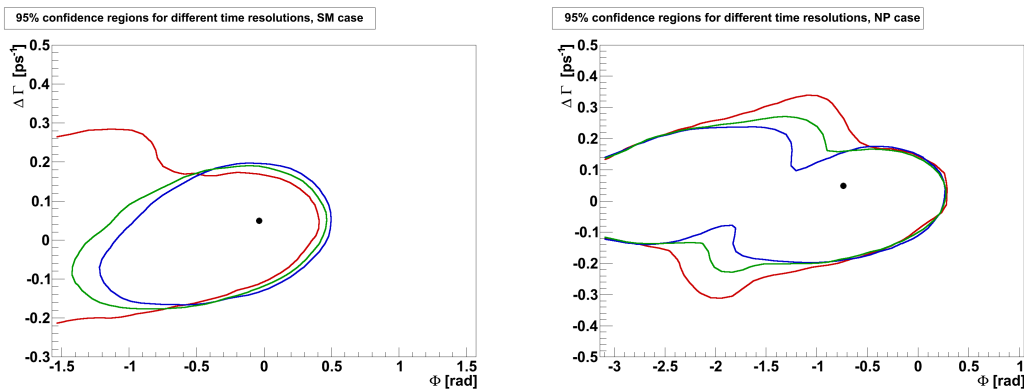
Figure 5.3: *95% contours for different proper time resolutions for the SM (left) and NP (right) case: nominal value (—), 20% worse (—) and 50% worse (—) than nominal. The dot denotes the production values.*

It is obvious that the proper time resolution has quite a significant impact on the resolution of the contours. In the SM case it can be seen that the two 95% contours representing the two solutions can be separated for nominal conditions and for a resolution 20% larger than the nominal value. This is not the case any more for a σ_{ct} which is by 50% larger. Despite this dependency it can be, however, stated that a loss of 20% compared to the nominal proper time resolution (which might be a realistic scenario in the early phase of data taking) does not significantly damage the ability to produce clear contours (see the small difference between the blue and the green curve in both figures).

5.3   Mistagging

As already mentioned before the mistag rate ω plays an important role for the sensitivity on the mixing phase. It is directly related to the effective tagging power $\epsilon_{eff} = \epsilon \cdot (1 - 2\omega)^2$ as shown in Chapter 2 (ϵ is the tagging efficiency). The statistical weight of N events with an effective tagging power of ϵ_{eff} is only that of $\epsilon_{eff}N$. Thus it is essential to optimise the tagging performance.

In Figure 5.4 for both the SM and the NP case three 95% contours (again at 20 pb^{-1} and for the same input values and the same seed as above) are shown: one

for nominal conditions ($\omega = 0.334$, $\epsilon_{eff} = 6.2\%$, —), one for $\omega = 0.367$ ($\epsilon_{eff} = 4.0\%$, —) and one for $\omega = 0.406$ ($\epsilon_{eff} = 2.0\%$, —). The tagging efficiency ϵ remains unchanged at the nominal value.
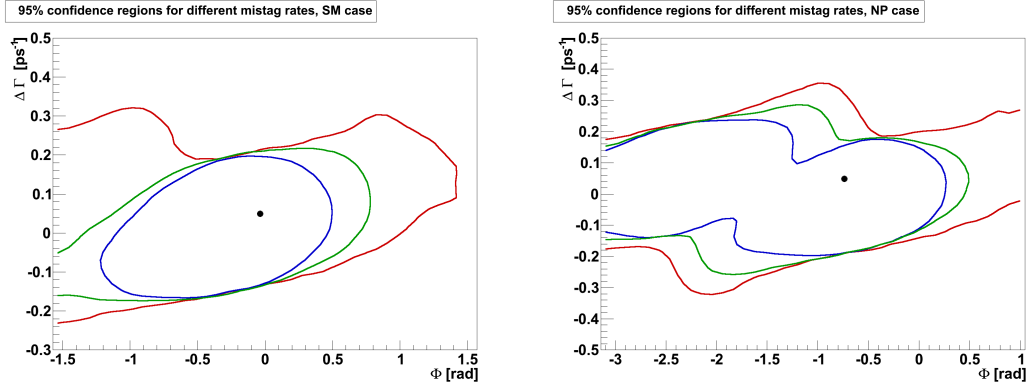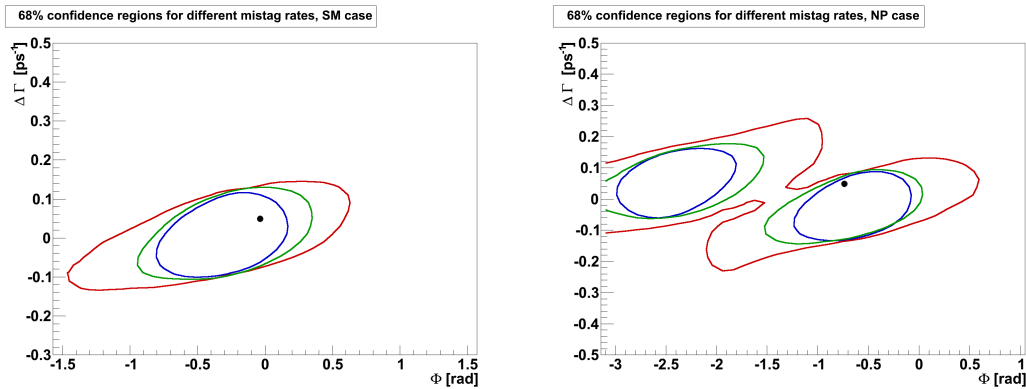

Figure 5.4: *95% contours for different mistag rates corresponding to an effective tagging power of 6.2% (nominal, —), 4.0% (—) and 2.0% (—) in the SM and NP case. The dot denotes the production values.*

One can clearly see that the mistag rate has a large influence on the contours as it is the case for the proper time resolution. Without a good tagging performance the extraction of the contours can only be done with a significant loss in resolution. Additionally, the 68% contours corresponding to the results in Figure 5.4 are shown in Figure 5.5 (SM case on the left, NP case on the right). These plots show a similar dependency as the 95% contours above.

Figure 5.5: *68% contours for different mistag rates corresponding to an effective tagging power of 6.2% (nominal, —), 4.0% (—) and 2.0% (—) in the SM and NP case. The dot denotes the production values.*

5.4 Toy data set

The effect of a different data set has already been mentioned in a previous section. It has to be pointed out again that a set of toy experiments is used as a "data set". By using a different seed for the random number generator in the $B_S^0 \rightarrow J/\psi\phi$ event generator a different toy data set can be produced and analysed. Figure 5.6 shows the effect of using different data sets at a luminosity of 50 pb^{-1} (corresponding to 2925 signal events) in the case of a NP input value of Φ. Plotted are the 68% (—), 90% (—) and 95% (—) contours.
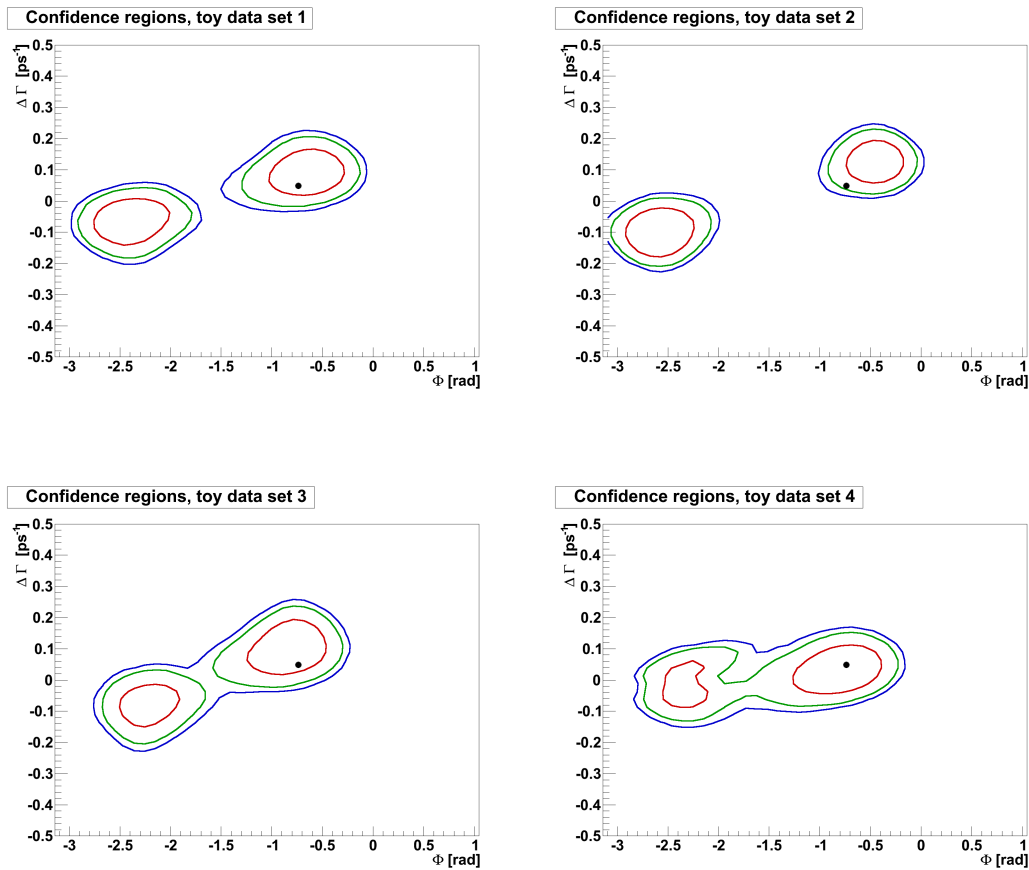
Figure 5.6: *68% (—), 90% (—) and 95% (—) contours for four different data sets at 50 pb^{-1} for the NP value of Φ. The dot denotes the production values.*

One can see that the resolution of the contours has a large dependency on the data set. The two solutions can be completely connected or well-separated. Some contours (like the one on the top right) have almost parabolic contours – the Feldman Cousins method would probably not be necessary – whereas others (e.g. bottom right) are rather oddly shaped and not parabolic at all. This shows how unpredictable the exact quantitative outcome of the analysis on data is.

5.5   Reproducing CDF results

It is an important cross-check to try to reproduce some of the CDF results. This is quite a difficult task because the CDF data sets are, of course, not available. Nevertheless using the information from several CDF notes ([10] and [56]) some interesting observations can be made.

One major difference between the LHCb and the CDF detector is the proper time resolution. As stated in [10] the signal proper time resolution at CDF is roughly 100 fs compared to circa 39 fs at LHCb (estimated from Monte Carlo samples). Other quantities are not expected to differ significantly. What the tagging is concerned the effective tagging power at CDF is still 4.8% (corresponding to a confidence contour between the blue and the green contour in Figure 5.4) and thus relatively close to the nominal LHCb value.

It can be assumed that for the same number of events the resolution of the mixing phase Φ extracted from simulated LHCb data is better than for the CDF sample. It has been tried to reproduce the CDF contour plot presented in [10] and [56] and to compare it to the corresponding LHCb confidence region. It is, however, evident that the CDF analysis cannot be reproduced exactly. To reproduce the CDF results the following changes to Table 4.1 (with the NP value of Φ) have been made (extracted e.g. from [10] and [56]):

- Difference in decay width $\Delta\Gamma = 0.15$

- 2000 signal events with a signal fraction $f_{sig} = 0.26$ (CDF: 1.35 fb^{-1})

- Proper time resolution of $\sigma_{ct} = 100$ fs (single Gaussian resolution model)

- Mistag rate $\omega = 0.345$, tagging efficiency $\epsilon = 0.5$

- Consequently, the effective tagging power is $\epsilon_{eff} = \epsilon \cdot (1 - 2\omega)^2 = 4.8\%$

- Decay width Γ, mixing frequency ΔM are additional nuisance parameters

For the corresponding LHCb plot the same values as in Table 4.1 have been used (NP case), except for the difference in the decay width: $\Delta\Gamma = 0.15$ as for the CDF plot to ensure that in both cases the same physics parameters are used. Γ and ΔM are additional nuisance parameters. The number of signal events is the same as in the CDF case (LHCb: 35 pb^{-1}). In Figure 5.6 the original CDF plot (top), the reproduced CDF plot (bottom left) and the corresponding LHCb plot (bottom right) are shown with the 68% contour (—) and the 95% contour (—). Although the CDF parameters and procedure are not exactly known, the contours can be retrieved quite accurately. Due to the good time resolution at LHCb one can expect a better sensitivity than at CDF as long as the detector quantities are close to the nominal conditions. Note also that 2000 signal events correspond to an integrated luminosity of only roughly 35 pb^{-1}. This again shows that one can expect interesting results already in the early phase of data taking.
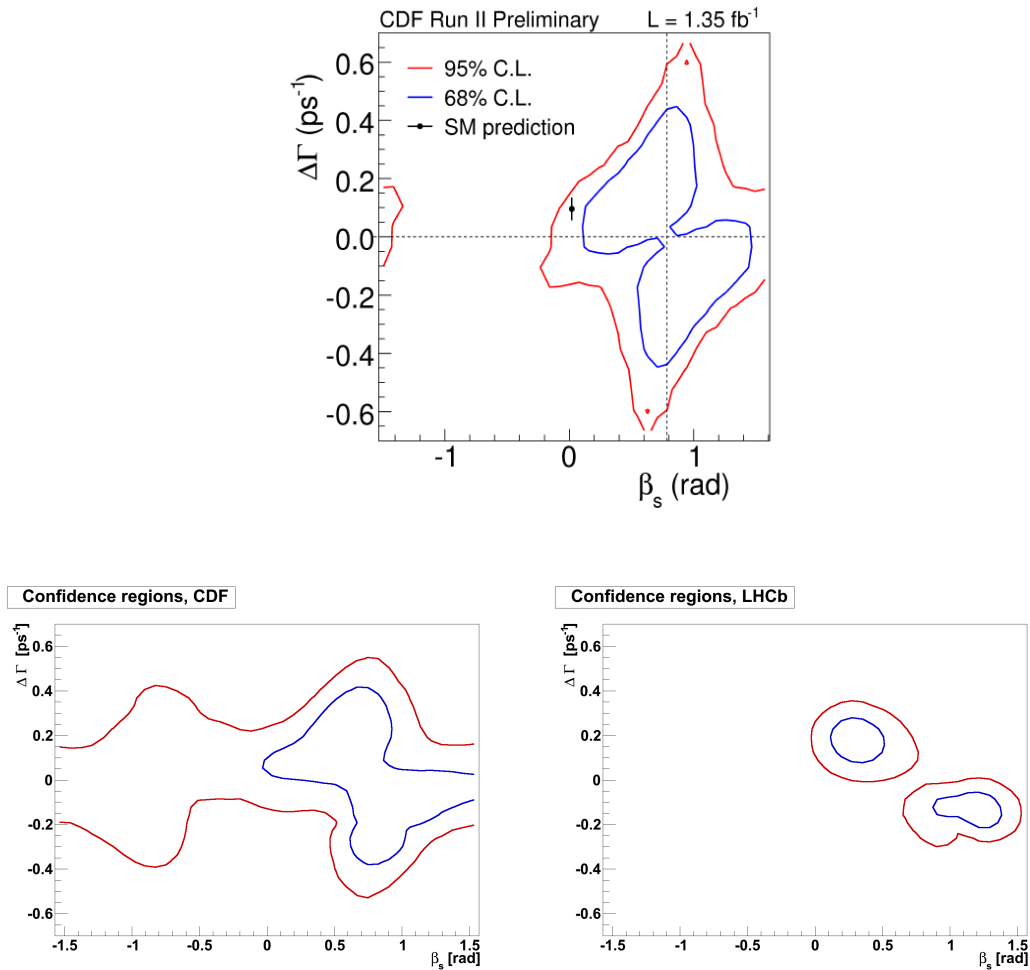
Figure 5.6: *Original CDF plot (top, [56]), reproduced CDF plot (bottom left) and corresponding LHCb plot (bottom right) showing the 68% contour (—) and the 95% contour (—) respectively. Note that $\Phi^{SM} = -2\beta_{\mathrm{S}}$.*

Up to now only toy experiments have been used. These are very useful to get a first impression on the subject of the Feldman Cousins analysis and to study the sensitivity on the mixing phase. It is, however, important to apply the method also on fully simulated Monte Carlo data. This is done in the following chapter for both the MC09 signal sample discussed in Chapter 3 and an additional sample generated with a NP value of the mixing phase Φ.

Feldman Cousins analysis with Monte Carlo simulated data

After having introduced how to calculate confidence levels for the mixing phase at low statistics and analysed the systematical uncertainties, in this chapter the Feldman Cousins method is applied to fully simulated Monte Carlo events. In contrast to toy experiments full Monte Carlo simulations include physical processes like decays and interactions with the detector.

In principle, the best sample to study would be the J/ψ inclusive sample mentioned in Chapter 3 because it contains prompt as well as longlived background and therefore is a realistic simulation of the first data. Unfortunately, a wrong model has been used when simulating the physics decay $B_S^0 \rightarrow J/\psi\phi$ events within the J/ψ inclusive sample. Therefore, this sample cannot be used.

The only alternative is the signal sample discussed in Chapter 3. An equivalent fully simulated sample with large CP violation (NP, $\Phi = -0.70$ rad) is also available [8]. Since these samples only contain signal events, a toy simulated background has been added.

It is clear that one cannot use the exact parameters from Table 4.1 used to generate the toy experiments as input values for fits to the data samples and the generation of the toy experiments. In a separate fit the mass and proper time resolution properties are determined. As both samples contain the truthmatched quantities as well (for the definition of truthmatched see Chapter 3), also the tagging quantities can be extracted. In the case of data there are different methods to extract the tagging quantities. For subsamples of 20 pb^{-1} and 50 pb^{-1} (1170 and 2925 signal events respectively) the values in Table 6.1 have been obtained. Due to the small amount of events only a single Gaussian has been used as resolution model for the proper time.

For the "physics parameters" the nominal values used in the generation which

---

[8]The name of the samples is MC09-Beam5TeV-VeloClosed-MagDown-Nu1-MC09-Sim04Reco02-withTruth. The event types are 13144002-BsJpsiphi,mm=CPV,DecProdCut and 13144006-BsJpsiphi,mm=LargeCPV,DecProdCut respectively.

Parameter	Standard Model Sample		New Physics Sample	
	$20\ \mathrm{pb}^{-1}$	$50\ \mathrm{pb}^{-1}$	$20\ \mathrm{pb}^{-1}$	$50\ \mathrm{pb}^{-1}$
m_{B_s}	5365.0 MeV	5364.8 Mev	5364.8 MeV	5364.6 MeV
$f_{\sigma_{m,1}}$	0.69	0.73	0.39	0.62
$\sigma_{m,1}$	10.8 MeV	11.6 MeV	10.0 MeV	11.1 MeV
$\sigma_{m,2}$	19.5 MeV	23.2 MeV	18.1 MeV	21.2 MeV
$\sigma_{m,total}$	14.1 MeV	15.6 MeV	15.5 MeV	15.7 MeV
$f_{\sigma_{ct,1}}$	1.0	1.0	1.0	1.0
$\sigma_{ct,total}$	35.7 fs	34.9 fs	34.4 fs	34.4 fs
ω	0.368	0.362	0.359	0.365
ϵ	0.571	0.583	0.588	0.588

Table 6.1: *Resolution parameters and tagging quantities determined from a 20 pb^{-1} / 50 pb^{-1} subsample of the two Monte Carlo data sets.*
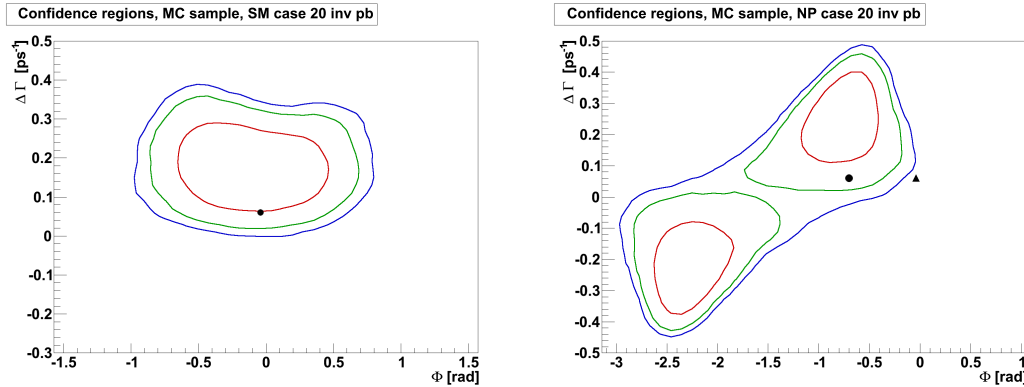
Figure 6.1: *Results of the Feldman Cousins analysis performed on a Monte Carlo sample corresponding to 20 pb^{-1} for a SM and a NP value of Φ, showing the 68% (—), the 90% (—) and the 95% (—) contour. The dot denotes the production values, the triangle the production values of the SM sample.*

are listed in Table 3.1 have been used. This is also the case for the only non-floating physics parameters, Γ and ΔM. For the generation of and the fit to the background events the same parameters as in Table 4.1 have been used.

The result of the Feldman Cousins analysis can be seen in Figure 6.1 (SM case left, NP case right), showing the 68% (—), the 90% (—) and the 95% (—) contour. The dot denotes the production values of the respective sample, the triangle the production values of the SM sample as a reference.

The results are similar to the contours shown in Chapter 4 and 5. In the SM case the values for $\Delta\Gamma$ and Φ are on the 68% contour and thus within a 1σ area around the most likely point. In the NP case the SM point is clearly outside
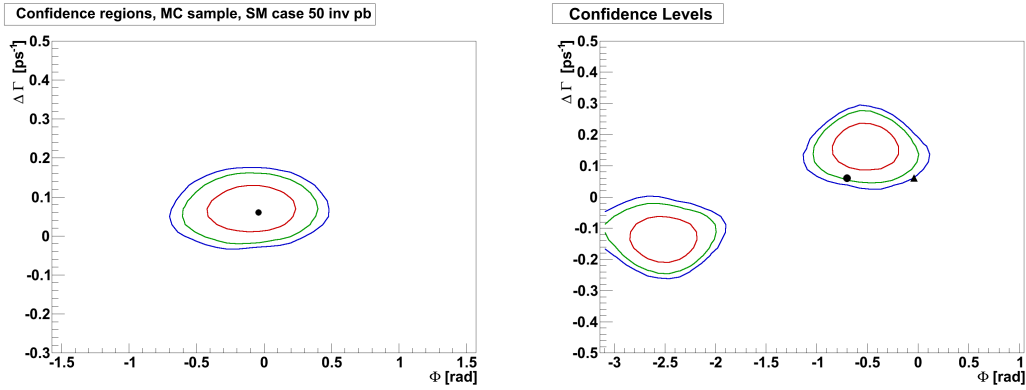
Figure 6.2: *Results of the Feldman Cousins analysis performed on a Monte Carlo sample corresponding to 50 pb^{-1} for a SM and a NP value of Φ, showing the 68% (—), the 90% (—) and the 95% (—) contour. The dot denotes the production values, the triangle the production values of the SM sample.*

the 95% confidence region whereas the production values are well inside the 90% contour. Note that the plot on the right and the combined CDF and DØ plot shown in Figure 1.5 are very similar.

In the same way, the contour levels are plotted for data samples corresponding to 50 pb^{-1}. The fit results for the mass and proper time resolution and the extracted tagging quantities are very close to the numbers for 20 pb^{-1}. They are listed in Table 6.1. The corresponding contour plots can be seen in Figure 6.2.

One can see that the contours are much smaller than the ones at 20 pb^{-1} which is expected for increasing statistics. In the SM case, the production values are now almost identical to the most likely point in the centre of the confidence regions shown in the left plot of Figure 6.2. In the NP case the two solutions are well-separated and the SM point is rejected with a confidence level of 95% whereas the production values are inside the 90% contour. The data sample, however, seems to have a small bias to larger values of $\Delta\Gamma$. In both cases the contours are almost parabolic which leads to the conclusion that the Feldman Cousins method is probably not needed for even larger data samples. This is consistent with the results discussed in Chapter 4.

It has to be pointed out that the results shown in this chapter are not a reliable prediction on how the contours drawn from data really look like. They can only be a guideline.

Conclusion

In this thesis the extraction of the B^0_S mixing phase Φ_S [9] arising in the decay $B^0_S \to J/\psi\phi$ has been studied for low event statistics. First of all, a selection strategy for the decay $B^0_S \to J/\psi\phi$ has been presented. The main goal of the selection is the maximisation of the signal yield. At an integrated luminosity of $\mathcal{L}_{int} = 2\text{fb}^{-1}$ (corresponding to a nominal year of data taking) one expects roughly 117 000 signal events (assuming a centre-of-mass energy of 14 TeV). The selection has been applied to two different simulated event samples, the first one representing nominal LHCb conditions, the second one a higher than nominal number of proton-proton interactions per bunch crossing. It has been shown that such an increase in the number of proton-proton interactions per bunch crossing does mainly influence the effective tagging power which is decreasing by roughly 20%. This, however, can be compensated with the higher instantaneous luminosity available at the increased number of proton-proton interactions per bunch crossing. Concerning the measurement of Φ_S in $B^0_S \to J/\psi\phi$ decays the operation of the detector at a higher than nominal luminosity seems possible.

The mixing phase is determined by fitting four-dimensional distributions with the theory prediction. At low statistics the extraction of the mixing phase Φ_S could suffer from complications. Errors are not Gaussian and the corresponding likelihood profiles are highly non-parabolic and have multiple not well-separated minima. It is shown in this thesis that using a standard two-dimensional likelihood scan to determine confidence levels leads to undercoverage of the predicted confidence regions. Instead, a method proposed by Feldman and Cousins, providing correct coverage by construction, is used. It is shown with simulated LHCb data that this method is, indeed, inevitable for very low statistics corresponding to integrated luminosities up to 50 pb^{-1}. With increasing statistics the results from both methods converge. Assuming nominal detector performance, the likelihood scan method is probably sufficient for integrated luminosities above 100 to 150 pb^{-1}.

---

[9]Note that for the sake of brevity in the body of the thesis (Chapter 1 to 6) the mixing phase is denoted as Φ and not as Φ_S.

Additionally, the influence of quantities like lifetime resolution, tagging power and number of events on the sensitivity on Φ_s and the shape of the confidence regions has been studied. Flavour tagging performance and time resolution are the crucial detector quantities that determine the sensitivity on Φ_s. Therefore, it is very important to achieve the nominal values especially in the early phase of data taking. Statistical fluctuations of the individual data sets could, however, lead to significant deviations from elliptical confidence regions and require the application of the Feldman Cousins method also for larger luminosities (more than 100 pb^{-1}). This fact makes it difficult to precisely predict the outcome of the analysis on data.

A Feldman Cousins analysis has also been performed using fully simulated Monte Carlo events corresponding to an integrated luminosity of 20 and 50 pb^{-1}. Two samples have been used, one generated with the Standard Model value of Φ_s and the other one with a possible New Physics value. It becomes evident that the resolution is largely dependent on the number of signal events. At 50 pb^{-1}, the contours are nearly elliptic meaning almost parabolic likelihood profiles.

Already very early (corresponding to luminosities of roughly 100 pb^{-1}) LHCb measurements will allow to either confirm the 2σ deviation from the Standard Model value of Φ_s seen by CDF and DØ or identify the deviation as a statistical fluctuation.

Bibliography

[1] S. L. Glashow; *Partial-symmetries of weak interactions*; Nucl. Phys. 22, 579 (1961)

[2] A. Salam and J. C. Ward; *Electromagnetic and weak interactions*; Phys. Lett. 13, 168 (1964)

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

[4] G. Bertone, D. Hooper, J. Silk; *Particle Dark Matter: Evidence, Candidates and Constraints*; arXiv:hep-ph/0404175v2 (2004)

[5] A. D. Sakharov; *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*; Journal of Exp. and Theoretical Physics 5: 2427 (1967)

[6] J. H. Christenson, J. W. Cronin, V. L. Fitch, R. Turlay; *Evidence for the 2π Decay of the K_2^0 Meson*; Phys. Rev. Lett. 13, 138 (1964)

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

[8] The CDF collaboration; *An updated measurement of the CP violating phase* $\Phi_{J/\psi\Phi}$; CDF/ANAL/BOTTOM/PUBLIC/9458 Version 1.0 (2008)

[9] D∅ webpage; more at `http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/B/B59/`

[10] T. Aaltonen et al.; *First Flavor-Tagged Determination of Bounds on Mixing-Induced CP Violation in* $B_S^0 \to J/\psi\phi$ *Decays*; CDF/PHYS/BOTTOM/PUBLIC/9102 Version 3.0 (2007)

[11] G. Feldman, R. Cousins; *Unified approach to the classical statistical analysis of small signals*; Phys. Rev. D 57, 7 (1998)

[12] Povh, Rith, Scholz, Zetsche; *Teilchen und Kerne*; Springer, 7. Auflage (2006)

[13] Y. Ashie et al. (The Super-Kamiokande Collaboration); *Evidence for an oscillatory signature in atmospheric neutrino oscillation*; arXiv:hep-ex/0404034v1 (2004)

[14] C. Amsler et al.; *Particle Physics Booklet*; Particle Data Group (July 2008)

[15] D. H. Perkins; *Introduction to High Energy Physics*; Cambridge University Press, 4th edition (2000)

[16] P. Ball; *The theory of CP-violation - in as much of a nutshell as will fit on 8 pages*; Frascati Physics Series Vol. XXXVI (2004)

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

[18] K. Anikeev et al.; *B Physics at the Tevatron: Run II and Beyond*; FERMILAB-Pub-01/197 (2001)

[19] J. Van Tilburg; *Track simulation and reconstruction in LHCb*; CERNTHESIS- 2005-020, (2005)

[20] CKMfitter Group (J. Charles et al.); *CP Violation and the CKM Matrix*; Eur. Phys. J. C41, 1-131 (2005) [hepph/0406184]; updated results and plots available at: `http://ckmfitter.in2p3.fr`

[21] A. Abulencia et al.; *Observation of* B_S^0*-* $\overline{B}_S^0$ *Oscillations*; Phys. Rev. Lett. 97, 242003 (2006)

[22] A. Lenz, U. Nierste; *Theoretical update on* B_S^0*-* $\overline{B}_S^0$*-Mixing*; hep-ph/0612167v3 (2007)

[23] J. Albrecht et al.; *Road map for the measurement of mixing induced CP violation in* $B_S^0 \rightarrow J/\psi\phi$ *at LHCb*; draft 9, LHCb 2009-xx (2009)

[24] J. Kalinowski; *SUSY theory review*; Acta Physica Polonica B, Vol. 38 (2007)

[25] S. Trine; *The MSSM Higgs sector and* B *-* $\overline{B}$ *mixing for large* $\tan\beta$; arXiv:0710.4955v1, hep-ph (2007)

[26] P. Ball; *Probing New Physics Through* B_S^0 *Mixing*; hep-ph/0703214v1 (2007)

[27] A. Lenz; *Unparticle phys. effects in* B_S^0 *Mixing*; Ph. Rev. D 76, 065006 (2007)

[28] Heavy Flavor Averaging Group; more at `http://www.slac.stanford.edu/xorg/hfag/`

[29] LHC design report, Volume 1, see `http://lhc.web.cern.ch/LHC/LHC-DesignReport.html`

[30] picture from CERN document server: Record #40525

[31] L. Evans, P. Bryant; *LHC Machine*; 2008 JINST 3 S08001 (2008)

[32] The LHCb collaboration; *LHCb Techn. Proposal*; CERN-LHCC-98-004 (1998)

[33] The LHCb collaboration; *The LHCb detector at the LHC*; 2008 JINST 3 S08005 (2008)

[34] The LHCb collaboration; *LHCb VELO (VErtex LOcator): Technical Design Report*; CERN-LHCC-2001-011 (2001)

[35] M. Needham, D. Volyanskyy; *Updated geometry description for the LHCb Trigger Tracker*; CERN-LHCb-2006-032 (2006)

[36] The LHCb collaboration; *LHCb inner tracker: Technical Design Report*; CERN-LHCC-2002-029 (2002)

[37] The LHCb collaboration; *LHCb outer tracker: Technical Design Report*; CERN-LHCC-2001-024 (2001)

[38] The LHCb collaboration; *LHCb RICH: Technical Design Report*; CERN-LHCC-2000-037 (2000)

[39] The LHCb collaboration; *LHCb calorimeters: Technical Design Report*; CERN-LHCC-2000-0036 (2000)

[40] The LHCb collaboration; *LHCb muon system: Technical Design Report*; CERN-LHCC-2001-010 (2001)

[41] R. Forty, O. Schneider; *RICH pattern recognition*; LHCb public note, LHCb 98-040 (1998)

[42] M. Calvi, O. Leroy, M. Musy; *Flavour Tagging Algorithms and Performances in LHCb*; LHCb public note, LHCb-2007-058 (2007)

[43] M. Calvi et al.; *LHCb Flavour Tagging Performance*, LHCb public note, LHCb 2003-115 (2003)

[44] J. W. Storey; *Flavour tagging with baryons and a study of two body Λ_b decays with the LHCb experiment*; CERN-THESIS-2008-26 (2008)

[45] A. Carbone, S. Vecchi; *Systematic studies on the B_S^0 proper time measurement for the determination of the mixing phase Φ_S in the channel $B_S^0 \to J/\psi\phi$*; LHCb public note, LHCb-2009-021 (2009)

[46] P. Vankov, G. Raven; *Proper-Time Resolution Modelling*; LHCb public note, LHCb-2007-055 (2007)

[47] The LHCb collaboration; *LHCb Trigger System Technical Design Report*; CERN LHCC 2003-031, LHCb TDR 10 (2003)

[48] U. Egede; *The search for a standard model Higgs at the LHC and electron identification using transition radiation in the ATLAS tracker*; CERN-THESIS-98-001 (1998)

[49] M. Calvi et al.; *Lifetime unbiased selection of* $B_S^0 \to J/\psi\phi$ *and related control channels*; LHCb public note, LHCb-2009-025 (2009)

[50] The LHCb Simulation Program; see `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/`

[51] The LHCb Digitisation Program; see `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/`

[52] The LHCb Reconstruction Program; see `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/`

[53] The LHCb Analysis Program; see `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/`

[54] C. Langenbruch, U. Uwer, S. Hansmann-Menzemer; *An unbinned likelihood fit for the measurement of* Φ_s *in the decay* $B_S^0 \to J/\psi\phi$; LHCb public note, LHCb-2009-028 (2009)

[55] V. Blobel, E. Lohrmann; *Statistische und numerische Methoden der Datenanalyse*; Teubner Studienbücher Physik (1998)

[56] C. Liu; *Search for CP violation in* $B_S^0 \to J/\psi\phi$; CDF PhD thesis, University of Pittsburgh (September 2008)

[57] W. Rolke et al.; *Limits and Confidence Intervals in the Presence of Nuisance Paramters*; arXiv:physics/04030599v5 (2009)

[58] F. James, M. Roos; *MINUIT, Function Minimization and Error Analysis*; CERN Program Library, D506

Acknowledgments

Erklärung

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

Heidelberg, den 20. April 2010 _____