DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of

doctor rerum naturalium

presented by

Dipl.-Phys. Christoph Langenbruch

born in Münster, Germany

Oral examination: November 09$^{\text{th}}$, 2011

Measurement of the $\mathrm{B_s^0}$ mixing phase in the decay $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$ with the LHCb experiment

Referees: Prof. Dr. Ulrich Uwer

Prof. Dr. Tilman Plehn

Abstract

This work presents a method to extract the CP-violating phase ϕ_s using the B_s^0 meson decay $B_s^0 \to J/\psi\,\phi$. The determination of ϕ_s is one of the key measurements of the physics program of the LHCb experiment. The phase ϕ_s arises from the interference between B_s^0 mixing and the decay and is predicted by the Standard Model to be $\phi_s = (-0.0363 \pm 0.0017)$ rad [1]. Physics processes beyond the Standard Model can affect the B_s^0 mixing amplitude and lead to deviations from this prediction. The precision measurement of ϕ_s therefore constitutes an indirect search for New Physics.

In the scope of this thesis an unbinned maximum likelihood method to extract ϕ_s has been implemented. Since the decay $B_s^0 \to J/\psi\,\phi$ is a pseudoscalar to vector-vector transition the extraction method depends on both decay time and decay angles. Acceptance and resolution effects are also taken into account. The procedure is successfully tested on fully simulated events and fast simulated data. It is used to analyze data taken by the LHCb experiment in 2010 which correspond to an integrated luminosity of $36\,\mathrm{pb}^{-1}$. With a flavor blind analysis the decay width difference $\Delta\Gamma_s$ is determined to be $\Delta\Gamma_s = (0.084 \pm 0.112_{\mathrm{stat.}} \pm 0.014_{\mathrm{syst.}})\ \mathrm{ps}^{-1}$ under the assumption of no CP violation. Using a B_s^0 production flavor dependent analysis two-dimensional confidence regions for ϕ_s and $\Delta\Gamma_s$ are extracted. The Standard Model prediction is found to agree within 1.0 standard deviations with the data. In one dimension the confidence interval for ϕ_s at 68.3% confidence level is determined to be $[-2.78, -0.39]$ rad.

Kurzfassung

In der vorliegenden Arbeit wird ein Algorithmus zur Bestimmung der CP-verletzenden phase ϕ_s im Zerfall $B_s^0 \to J/\psi\,\phi$ vorgestellt. Die Bestimmung von ϕ_s stellt eine der Kernanalysen des Physik Programms des LHCb Experiments dar. Die Phase ϕ_s tritt in der Interferenz zwischen B_s^0 Mischung und Zerfall auf und wird im Standardmodell zu $\phi_s = (-0{,}0363 \pm 0{,}0017)$ rad vorhergesagt [1]. Physikalische Prozesse jenseits des Standardmodells können die B_s^0 Mischungsamplitude modifizieren und so zu Abweichungen von dieser Vorhersage führen. Die Messung von ϕ_s stellt daher eine indirekte Suche nach Neuer Physik dar.

Im Kontext dieser Arbeit wurde ein Maximum Likelihood Fit zur Bestimmung von ϕ_s entwickelt. Da der Zerfall $B_s^0 \to J/\psi\,\phi$ einen Übergang eines pseudoskalaren- in zwei Vektorteilchen darstellt hängt der Algorithmus sowohl von der Zerfallszeit als auch den Zerfallswinkeln ab. Akzeptanz- und Auflösungseffekte werden ebenfalls berücksichtigt. Der Algorithmus wurde erfolgreich auf simulierten Daten getestet. Der Algorithmus wird verwendet um Daten zu analysieren welche vom LHCb Experiment im Jahr 2010 aufgezeichnet wurden und einer integrierten Luminosität von $36\,\mathrm{pb}^{-1}$ entsprechen. Ohne den Flavor des B_s^0 Mesons im Anfangszustand zu berücksichtigen wird der Zerfallsbreitenunterschied $\Delta\Gamma_s$ zu $\Delta\Gamma_s = (0{,}084 \pm 0{,}112_{\mathrm{stat.}} \pm 0{,}014_{\mathrm{syst.}})\ \mathrm{ps}^{-1}$ bestimmt, wobei der Effekt möglicher CP Verletzung vernachlässigt wurde. Mittels einer Analyse welche Informationen über den Flavor des B_s^0 Mesons im Anfangszustand verwendet werden zweidimensionale Konfidenzregionen für ϕ_s and $\Delta\Gamma_s$ bestimmt. Die Standardmodellvorhersage stimmt mit den Daten innerhalb von 1,0 Standardabweichungen überein. In einer Dimension ergibt sich für ϕ_s das Konfidenzintervall $[-2{,}78, -0{,}39]$ rad für ein Konfidenzniveau von 68,3%.

Contents

Introduction

The Standard Model of particle physics in its current form was developed in the 1960s [2, 3, 4]. Over the years it has been tested in numerous experiments and so far it has been very successful in the description of all available data from particle physics experiments[1]. But even though the predictions made by the Standard Model have been confirmed with astonishing precision some open questions remain

- The predominance of matter over antimatter in the universe is not understood. Since matter and antimatter are believed to have been produced in equal amounts in the big bang it is unclear why antimatter in our universe is practically absent. CP violation[2] is one of the prerequisites for the matter abundance seen today [5].

- Cosmological observations show that only 4% of the energy-density of the universe originates from baryonic matter which is described by the Standard Model. The remaining 96% are given by dark matter (20%) and dark energy (76%) [6, 7]. The Standard Model contains no viable dark matter candidate [8].

- Particle masses in the Standard Model are generated by the mechanism of spontaneous symmetry breaking. The Higgs boson [9, 10, 11] resulting from this mechanism is the only particle in the Standard Model which has not been observed yet.

The Large Hadron Collider (LHC) was constructed to find answers to these questions. As a proton-proton collider with a nominal center of mass energy of 14 TeV[3] and a design luminosity of $1 \cdot 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ it is the most powerful particle accelerator to date. Four large experiments are located around the interaction points. Two experiments, CMS and ATLAS are general purpose detectors performing direct searches for the Higgs boson and signatures of particles beyond the Standard Model. The other two experiments are specialized detectors built to study heavy ion collisions (ALICE) and to perform precision measurements of the decay properties of b hadrons (LHCb).

The LHCb experiment profits from the large $\mathrm{b\bar{b}}$ production cross section $\sigma_{\mathrm{b\bar{b}}} \sim 500\,\mu\mathrm{b}$ at $\sqrt{s} = 14\,\mathrm{TeV}$. At nominal luminosity[4] the Large Hadron Collider will produce about 10^{12} $\mathrm{b\bar{b}}$ pairs at the LHCb interaction point per year. LHCb searches for effects beyond the Standard Model (New Physics) by performing precision measurements of loop mediated processes. Loop mediated processes are heavily suppressed in the Standard Model. They are therefore promising candidates to see comparatively large effects from

[1] With the exception of the observation of non-zero neutrino masses.
[2] C is the operator which inverts charge, P the parity operator which inverts space $\vec{x} \to -\vec{x}$
[3] Currently the LHC operates at $\sqrt{s} = 7\,\mathrm{TeV}$.
[4] The nominal luminosity at the LHCb interaction point is $2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

physics processes beyond the Standard Model which are not necessarily suppressed. These possible new contributions can have a large effect on the branching ratios of rare decays and introduce new sources of CP violation.

This thesis presents the determination of the CP violating phase ϕ_s using the B_s^0 meson decay $B_s^0 \to J/\psi\,\phi$. The phase ϕ_s originates from interference between the direct decay and the decay after the B_s^0 has oscillated. The mixing of a B_s^0 meson into its antiparticle $\bar{B}_s^0$ is a flavor changing neutral current process (FCNC) which is forbidden at tree level in the Standard Model. The transition therefore occurs only at loop level via so called box-diagrams (Figure 1.3). B_s^0 mixing is determined by three parameters, which will be introduced in detail in the next chapter, the mixing frequency Δm_s, the decay width difference $\Delta\Gamma_s$ and the mixing phase ϕ_s. The phase ϕ_s is of particular interest since it is well predicted in the Standard Model $\phi_s = -0.0363 \pm 0.0017$ [1] so that large deviations from this prediction would constitute an unambiguous sign of New Physics.

One of the key components of the analysis of the decay $B_s^0 \to J/\psi\,\phi$ was implemented in the course of this thesis, namely the angular and time dependent unbinned maximum likelihood fit to extract ϕ_s from data taken by the LHCb detector. Parameter estimation via the unbinned maximum likelihood technique requires the signal and background components in the data to be modeled by probability density functions depending on the parameters of interest. The signal component is modeled according to the differential decay rate of the signal decay given by theory. The background contribution is modeled empirically. Both components depend on the B_s^0 decay time, the decay angles and the reconstructed B_s^0 mass. In a minimization process the parameters for which the measured data sample becomes most likely are then determined. The analysis can be performed either in a flavor blind way or depending on the B_s^0 production flavor with the former being only minimally sensitive to ϕ_s. The decay time resolution is taken into account. Decay time and angular acceptance effects due to detector geometry and selection are also included. The implemented code was rigorously tested using simulated data. It was used to perform a flavor blind analysis to extract $\Delta\Gamma_s$ under the assumption of no CP violation. A B_s^0 production flavor dependent analysis was performed and confidence contours in the ϕ_s-$\Delta\Gamma_s$ parameter space were deduced from the data. Both measurements will be discussed in this thesis.

The thesis is organized in eight chapters following this introduction. Chapter 1 introduces the theoretical framework of the Standard Model and gives a short overview of CP violation focusing on the signal decay $B_s^0 \to J/\psi\,\phi$. Chapter 2 describes the LHCb detector and its subsystems. The selection of the signal decay and reconstruction effects that need to be accounted for in the extraction of ϕ_s are discussed in Chapter 3. Chapter 4 first summarizes the method of parameter estimation via unbinned maximum likelihood fits and the determination of confidence intervals. Furthermore it discusses the description of the data via probability density functions which is necessary for the determination of the physics parameters. Chapter 5 discusses the fit validation on simulated events generated using the full detector simulation and via a fast simulation technique. The extraction of $\Delta\Gamma_s$ using a flavor blind study under the assumption of no CP violation is detailed in Chapter 6. Chapter 7 covers the determination of

confidence regions for ϕ_s and $\Delta\Gamma_s$ exploiting the determination of the B_s^0 production flavor. Chapter 8 finally summarizes the results and gives a short outlook on the determination of ϕ_s with the 2011 data sample.

1. Theory

This chapter gives a short overview of the Standard Model focusing on the flavor sector. The CKM mechanism and the phenomenon of CP violation caused by a non-trivial CKM phase are discussed in detail. The mixing of neutral B mesons is introduced and a formalism to describe both mixing and decay of B mesons is presented. The decay $B_s^0 \to J/\psi\,\phi$ is introduced as an example of time dependent CP violation. In this decay the CP violating phase ϕ_s arises due to the interference between direct decay and decay after B_s^0-$\bar{B}_s^0$ mixing. Finally the current experimental status of the phase ϕ_s is presented and the effect of possible contributions from processes beyond the Standard Model on ϕ_s is discussed.

1.1. The Standard Model

The aim of Particle Physics is to understand the laws of nature at the smallest scales. The Standard Model of Particle Physics (SM) [2, 3, 4] describes the fundamental building blocks of matter and their interactions. It is a relativistic renormalizable quantum field theory encompassing all the current knowledge in the field of Particle Physics. The Standard Model combines the theory of the strong interactions, Quantum Chromo Dynamics (QCD) with the Glashow-Salam-Weinberg electroweak theory. It is is well tested and, with only few exceptions, very successfull in describing the experimental observations. An enormous amount of literature exists on the topic of the Standard Model, therefore only the main concepts will be summarized below. For more detailed review articles on which this overview is based see [12, 13, 14, 15].

The symmetry group of the Standard Model

Symmetries are of central importance in physics since they imply, as Noether's theorem states, conservation laws. An example would be Quantum Electro Dynamics (QED) where the symmetry of the Lagrangian under a global phase rotation of the wave function results in conservation of the electric charge. Promoting the global symmetry to a local symmetry gives rise to the gauge boson of QED, the photon, and its interaction with the matter fields via covariant derivatives.

The matter particles in the Standard Model are described as fermionic spin 1/2 fields. The 12 fermions of the Standard Model are further subdivided in two classes, quarks and leptons in three generations each, which are given in Table 1.1. The Standard Model Lagrangian is characterized by its invariance under local gauge transformations of the symmetry group $\mathrm{SU}(3)_\mathrm{C} \otimes \mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{U}(1)_\mathrm{Y}$. This gauge symmetry fixes the interactions of the theory. In particular it results in the mediators of the forces, the spin 1 gauge bosons of the Standard Model.

Quantum Chromo Dynamics

$\mathrm{SU}(3)_\mathrm{C}$ is the gauge group of QCD, where the index C denotes color. The gauge bosons of QCD are the 8 gluons, each corresponding to one of the generators of $\mathrm{SU}(3)_\mathrm{C}$. Only the quarks, denoted in Table 1.1a, take part in the strong interaction as the leptons do not carry color charge. The strong force is the interaction responsible for the binding of quarks in mesons and baryons and the binding of protons and neutrons in nuclei. The gluons also carry color charge themselves, a consequence of the non-abelian structure of $\mathrm{SU}(3)_\mathrm{C}$, which results in three and four gluon interactions. This also gives rise to a strong dependency of the effective coupling on the transfered four momentum q^2, the coupling is said to be "running" with q^2. For large q^2 the coupling is small which is known as "asymptotic freedom". The opposite is true for small q^2 where the coupling rises with the inverse logarithm of q^2. This behaviour is known as confinement and is the reason for the absence of free quarks in nature. The gluon self-coupling also implies that the strong force only has comparatively short range even though the gluons are massless.

The electroweak sector

$\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{U}(1)_\mathrm{Y}$ is the gauge group of the unified electroweak theory. The gauge bosons resulting from invariance of the Lagrangian under the gauge transformations are the electroweak gauge bosons W_i^μ and B^μ. The L in $\mathrm{SU}(2)_\mathrm{L}$ denotes that the gauge bosons W_i^μ only couple to left-handed weak isospin doublets. The gauge boson B^μ on the other hand couples to both the left-handed doublets as well as the right-handed isospin singlets carrying weak hypercharge denoted by Y. Both quarks and leptons take part in the electroweak interaction organized as left-handed weak isospin doublets and right-handed isospin singlets as shown in Table 1.3. The commutators of the $\mathrm{SU}(2)_\mathrm{L}$ generators are, similar to the $\mathrm{SU}(3)_\mathrm{C}$ generators, nonzero which leads to W_i^μ self-interaction. The more familiar charged weak bosons W^+ and W^- are given by linear combinations of W_1^μ and W_2^μ. The neutral Z^0 boson and the photon are linear combinations of W_3^μ and B^μ, connected via the weak mixing angle. The gauge bosons of the Standard Model are summarily listed in Table 1.2.

The photon is the mediator of the electromagnetic force which is responsible for binding atoms and molecules together. It couples to electric charge Q and is itself electrically neutral. Since the photon is massless the electromagnetic force has long range. The electromagnetic interaction has, as has the strong interaction, pure vector character.

The exchange bosons of the weak interaction are the charged weak bosons W^+ and W^- and the neutral Z^0. Interactions via the charged weak force are the only processes in the SM that can change flavor and are thus responsible for all flavor changing processes such as β decay. Unlike the electromagnetic and the strong interaction, the weak interaction is not invariant under the Parity operation P due to the different couplings of left- and right-handed fields.

Quarks		
Generation	Particle	Mass
1^{st}	$\begin{pmatrix} u \\ d \end{pmatrix}$	$1.5 - 3.3\,\text{MeV}$ $3.5 - 6.0\,\text{MeV}$
2^{nd}	$\begin{pmatrix} c \\ s \end{pmatrix}$	$1.27^{+0.07}_{-0.11}\,\text{GeV}$ $104^{+26}_{-34}\,\text{MeV}$
3^{rd}	$\begin{pmatrix} t \\ b \end{pmatrix}$	$171.2 \pm 2.1\,\text{GeV}$ $4.20^{+0.17}_{-0.07}\,\text{GeV}$

(a)

Leptons		
Generation	Particle	Mass
1^{st}	$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}$	$511.0\,\text{keV}$ $< 2\,\text{eV}$
2^{nd}	$\begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}$	$105.7\,\text{MeV}$ $< 2\,\text{eV}$
3^{rd}	$\begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}$	$1776.8\,\text{MeV}$ $< 2\,\text{eV}$

(b)

Table 1.1.: (a) Quarks and (b) leptons in the Standard Model. Values for the quark and lepton masses are compiled from [16].

Spontaneous symmetry breaking

From experimental observation it is well known that the weak bosons $W^{\pm}$ and Z^0, unlike the photon, are massive. In fact the masses of all massive particles in the Standard Model, quarks and charged leptons as well as the massive gauge bosons, are generated by the mechanism of spontaneous symmetry breaking, also known as the Higgs-mechanism. In its minimal version an isospin doublet Φ of two complex scalar fields with the Higgs-potential $V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2$ with $\mu^2 < 0$ and $\lambda > 0$ is introduced in the Standard Model Lagrangian. This potential leads to a non-symmetric ground state, breaking the $\text{SU}(2)_\text{L} \otimes \text{U}(1)_\text{Y}$ electroweak symmetry down to the $\text{U}(1)_\text{Q}$ electric charge symmetry. The heavy gauge bosons acquire mass via the covariant derivatives in the kinematic term $(D^\mu\Phi)^\dagger D_\mu\Phi$. Fermion masses are introduced via Yukawa couplings to Φ. The Yukawa couplings are in general not diagonal. Diagonalizing the Yukawa couplings for the quarks leads to their mass eigenstates given in Table 1.1a. The mismatch between mass and weak eigenstates is the reason for the introduction of the Cabibbo-Kobayashi-Maskawa (CKM) matrix which connects the mass and weak eigenstates.

The Higgs-mechanism predicts one observable massive spin 0 particle, the Higgs-boson. The Higgs-boson is the only particle in the Standard Model which has not yet been discovered and the search for it is one of the main motivations for the construction of the Large Hadron Collider.

1.2. The Flavor Sector of the Standard Model

1.2.1. The CKM mechanism

The charged weak interaction is the only interaction in the SM that can change flavor. The charged current part of the Standard Model Lagrangian wich governs the charged

Gauge bosons		
boson	mass	couples to
photon γ	0	el. charge
gluons $G^{i=1...8}$	0	color
W^+	80.4 GeV	weak charge
W^-	80.4 GeV	weak charge
Z^0	91.2 GeV	weak & el. charge

Table 1.2.: Gauge bosons in the Standard Model. The masses of the heavy gauge bosons are compiled from [16].

Fermions						
Quarks			Q	T	T_3	Y
$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$+\frac{2}{3}$ $-\frac{1}{3}$	$\frac{1}{2}$	$+\frac{1}{2}$ $-\frac{1}{2}$	$+\frac{1}{3}$ $+\frac{1}{3}$
u_R	c_R	t_R	$+\frac{2}{3}$	0	0	$+\frac{4}{3}$
d_R	s_R	b_R	$-\frac{1}{3}$	0	0	$-\frac{2}{3}$
Leptons			Q	T	T_3	Y
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	0 -1	$\frac{1}{2}$	$+\frac{1}{2}$ $-\frac{1}{2}$	-1 -1
e_R	μ_R	τ_R	-1	0	0	-2

Table 1.3.: Quarks and leptons are organized in left-handed doublets and right handed singlets of the weak isospin T. The dashed quarks d', s' and b' are the weak eigenstates which are connected to the mass eigenstates in Table 1.1a via the CKM matrix (see equation 1.2). Y denotes the weak hypercharge which is connected to the electric charge Q and the third component of the weak isospin T_3 via $Y = 2(Q - T_3)$. Table from [17].

weak interaction is given by

$$\mathcal{L}_{\mathrm{CC}} = -\frac{g}{2\sqrt{2}}\Big[\bar{u}_i W_\mu^+ \gamma^\mu (1-\gamma^5) d_i' + \bar{d}_i' W_\mu^- \gamma^\mu (1-\gamma^5) u_i\Big], \tag{1.1}$$

where the dashed quark fields denote the weak eigenstates. The weak eigenstates are constructed by rotating the mass eigenstates with the Cabbibo-Kobayashi-Maskawa (CKM) matrix $\mathrm{V_{CKM}}$ [18, 19], a complex unitary 3×3 matrix. It is customary to only rotate the down-type quarks (d, s, b) to

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathrm{V_{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} \mathrm{V_{ud}} & \mathrm{V_{us}} & \mathrm{V_{ub}} \\ \mathrm{V_{cd}} & \mathrm{V_{cs}} & \mathrm{V_{cb}} \\ \mathrm{V_{td}} & \mathrm{V_{ts}} & \mathrm{V_{tb}} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.2}$$

As $\mathrm{V_{CKM}}$ is unitary, $\mathrm{V_{CKM}V_{CKM}^\dagger} = 1$, the 18 parameters which describe a complex 3×3 matrix reduce to nine. Five parameters can be absorbed as unobservable quark phases which leaves four independent parameters. The standard parametrization of $\mathrm{V_{CKM}}$ [20] is characterized by three Euler angles Θ_{12}, Θ_{13}, Θ_{23} and the phase δ. Using the shortcuts $c_{ij} = \cos\Theta_{ij}$ and $s_{ij} = \sin\Theta_{ij}$ $\mathrm{V_{CKM}}$ is given by

$$\begin{aligned} \mathrm{V_{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}. \end{aligned} \tag{1.3}$$

A common parametrization of $\mathrm{V_{CKM}}$ which reflects the hierarchy of the matrix elements is the Wolfenstein parametrization [21]. It uses the real parameters λ, A, ρ and η, with η being responsible for the imaginary part of the entries in $\mathrm{V_{CKM}}$. It can be obtained by using the definitions

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

in the standard parametrization. The parameter $\lambda \approx 0.23$ has the role of an expansion parameter which simplifies the estimation of the size of the elements of $\mathrm{V_{CKM}}$. The parametrization up to order λ^4 is

$$\mathrm{V_{CKM}} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 + \frac{1}{2}A\lambda^4(1 - 2(\rho + i\eta)) & 1 - \frac{1}{2}A^4\lambda^4 \end{pmatrix} \text{ [22].}$$

From this parametrization it is immediately clear that the diagonal elements are large ~ 1, whereas the off-diagonal elements which are responsible for transitions between different generations are smaller with $V_{us}, V_{cd} \sim \lambda$, $V_{cb}, V_{ts} \sim \lambda^2$ and $V_{ub}, V_{td} \sim \lambda^3$. The imaginary part relative to the magnitude of the matrix element is largest for V_{ub}. Up to order λ^4 only V_{ub}, V_{td} and V_{ts} have an imaginary component.

1.2.2. CP violation

Applying the CP transformation to the charged current part of the Standard Model Lagrangian given by

$$\mathcal{L}_{CC} = -\frac{g}{2\sqrt{2}}\Big[(V_{CKM})_{ij}\,\bar{u}_i W_\mu^+ \gamma^\mu(1-\gamma^5)d_j + (V_{CKM}^*)_{ij}\,\bar{d}_j W_\mu^- \gamma^\mu(1-\gamma^5)u_i\Big]$$

results in the CP conjugated Lagrangian

$$\mathcal{L}_{CC}^{CP} = -\frac{g}{2\sqrt{2}}\Big[(V_{CKM})_{ij}\,\bar{d}_j W_\mu^- \gamma^\mu(1-\gamma^5)u_i + (V_{CKM}^*)_{ij}\,\bar{u}_i W_\mu^+ \gamma^\mu(1-\gamma^5)d_j\Big]$$ [22].

Both expressions are identical if $(V_{CKM}^*)_{ij} = (V_{CKM})_{ij}$, i.e. all CKM matrix elements are real. A non-trivial phase δ (see equation 1.3) inducing complex CKM matrix elements can therefore lead to CP violation.

In nature the CP symmetry is indeed violated. CP violation was first discovered in 1964 in neutral K decays [23]. More recently CP violation has also been observed in the B meson sector by the BaBar and Belle collaborations [24, 25].

1.2.3. The unitarity triangles

The unitarity condition $V_{CKM}V_{CKM}^\dagger = 1$ results in three equations for the off-diagonal elements[1]:

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \quad (1.4)$$

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

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (1.6)$$

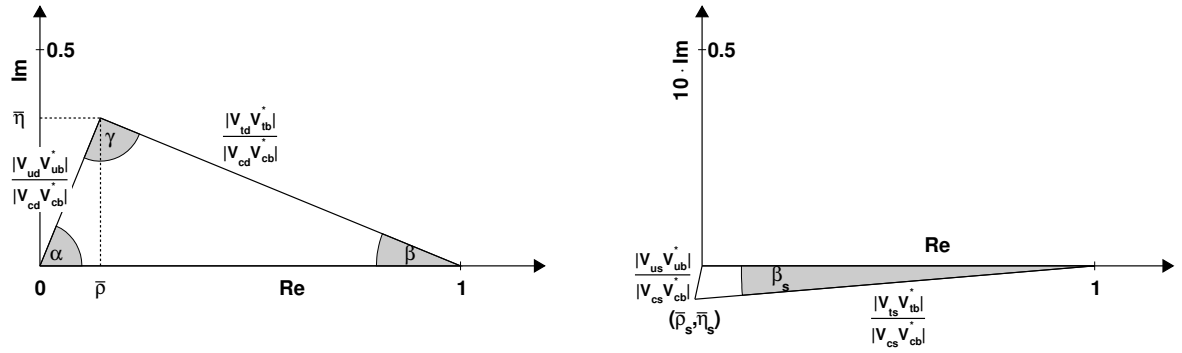
Since on the left-hand side of the equations are sums of three complex numbers the conditions can be visualized as triangles in the complex plane. Usually one side of the three triangles is normalized to coincide with the real axis which results in the equations

$$\frac{V_{ud}V_{us}^*}{V_{cd}V_{cs}^*} + \frac{V_{cd}V_{cs}^*}{V_{cd}V_{cs}^*} + \frac{V_{td}V_{ts}^*}{V_{cd}V_{cs}^*} = 0 \quad (1.7)$$

$$\frac{V_{us}V_{ub}^*}{V_{cs}V_{cb}^*} + \frac{V_{cs}V_{cb}^*}{V_{cs}V_{cb}^*} + \frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*} = 0 \quad (1.8)$$

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + \frac{V_{cd}V_{cb}^*}{V_{cd}V_{cb}^*} + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0 \quad (1.9)$$

[1] The other three off-diagonal elements in the unitarity equation $V_{CKM}V_{CKM}^\dagger = 1$ result in three equations which are just the compex conjugates of equations 1.4, 1.5 and 1.6. Keep in mind that each one of the equations 1.4, 1.5 and 1.6 gives two conditions, one for the real part and one for the imaginary part.



(a) B_d^0 triangle

(b) B_s^0 triangle

Figure 1.1.: Illustration of (a) the B_d^0 triangle and (b) the B_s^0 triangle in the complex plane. The B_s^0 triangle is very flat because the angle β_s is very small. Note that the imaginary axis for this plot is scaled by a factor 10.

An illustration of the unitarity triangle resulting from equation 1.9 is given in Figure 1.1a. It is also called the "B_d^0 triangle", since its sides and angles are accessible through B_d^0 decays. The vertices of the triangle are, due to the normalization, $(0,0)$, $(0,1)$ and $(\bar{\rho},\bar{\eta})$. The angles at the vertices are defined as

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \tag{1.10}$$

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{1.11}$$

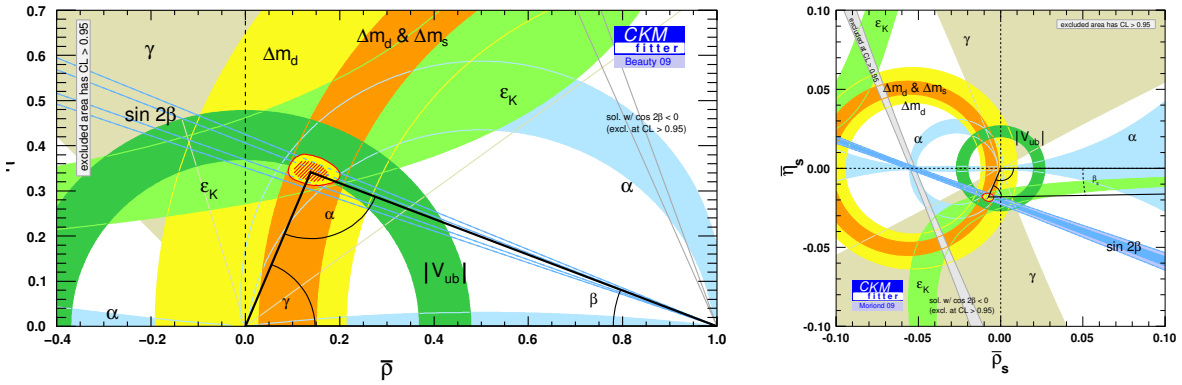
$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right). \tag{1.12}$$

Figure 1.1b gives the "B_s^0 unitarity triangle" resulting from equation 1.8. The apex of the triangle is located at $(\bar{\rho}_s,\bar{\eta}_s)$. This unitarity triangle is nearly flat with the small angle β_s at the vertex $(1,0)$ given by

$$\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right). \tag{1.13}$$

1.2.4. Experimental status of the unitarity triangles

The Standard Model does not predict the CKM matrix elements. A central aim of flavor physics is to overconstrain the CKM triangles to test the unitarity conditions. Figures 1.2a and 1.2b show the unitarity triangles resulting from a global fit of the



(a) B_d^0 triangle (b) B_s^0 triangle

Figure 1.2.: (a) B_d^0 triangle and (b) B_s^0 triangle in the complex plane from a global CKM fit. Both plots are from [1].

CKM matrix elements in the Standard Model. The lengths of the sides of the unitarity triangles are determined by the absolute values of CKM matrix elements.

- $|V_{ub}|$ and $|V_{cb}|$ determine the length of the left side of the B_d^0 triangle. Both quantities can be accessed in semileptonic B decays. $|V_{ub}|$ can be extracted from semileptonic decays to light mesons, e.g. $B \to \pi\ell\nu$. $|V_{cb}|$ can be accessed using semileptonic B decays to charm, e.g. $B \to D\ell\nu$.

- The right side of the B_d^0 triangle is determined by $|V_{td}|$ and $|V_{tb}|$ which can be constrained using B_d^0-$\bar{B}_d^0$ and B_s^0-$\bar{B}_s^0$ mixing.

The angles of the unitarity triangles are directly related to complex CKM matrix elements. They can therefore be accessed by precision measurements of CP-violation.

- The angle β appears in B_d^0 mixing. The precise measurement of $\sin 2\beta$ was one of the main goal of the experiments BaBar and Belle and confirmed the existence of CP violation in the B sector for the first time [24, 25]. Both experiments determine β using the "gold plated" decay channel $B_d^0 \to J/\psi\, K_S^0$. The most recent combined result is

  $$\sin 2\beta = (0.671 \pm 0.023) \text{ rad [16].}$$

  This combination represents the most precise angular constraint entering the fit. It is given by the dark blue band in Figure 1.2.

- The angle β_s appears in B_s^0 mixing. The current experimental situation is discussed in detail in section 1.3.4. The determination of $\phi_s = -2\beta_s$[2] using the decay channel $B_s^0 \to J/\psi\, \phi$ is one of the main physics goals of the LHCb experiment and the topic of this thesis.

[2] In the Standard Model

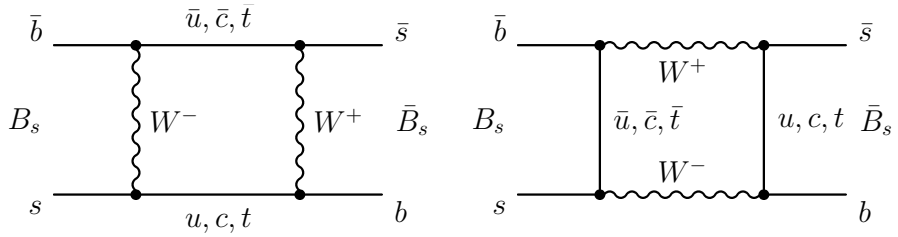


Figure 1.3.: B_s^0 mixing in the Standard Model.

1.2.5. Mixing of neutral B mesons

The process of B meson mixing describes transitions between the flavor eigenstates B_q^0 and $\bar{B}_q^0$. The flavor content of these states is given by $\bar{b}q$ and $b\bar{q}$ respectively. Oscillations between B_q^0 and $\bar{B}_q^0$ therefore constitute flavor changing neutral currents (FCNCs) with $\Delta b = 2$. In the Standard Model FCNCs are forbidden at tree level. FCNCs can however be induced by the charged weak interaction at loop level. Figure 1.3 shows the Feynman diagrams which contribute to B_s^0 mixing in the Standard Model. The diagram which dominates the mixing is the contribution from the top quark. This is caused by the strong GIM suppression (Glashow, Iliopoulos, Maiani) of the contribution from the up and charm quark $\propto (m_c^2 - m_u^2)/M_W^2$. The top contribution on the other hand experiences no suppression because $(m_t^2 - m_c^2)/M_W^2 \sim 1$.

Since it only occurs at loop level B mixing is suppressed in the Standard Model. In addition B mixing is CKM suppressed due to off-diagonal CKM matrix elements. For B_s^0 mixing the CKM matrix element entering is $|V_{ts}|$, for B_d^0 mixing the suppression is even larger with $|V_{td}|$ resulting in slower mixing. Possible contributions to B meson mixing from processes beyond the Standard Model which do not necessarily have to be suppressed are discussed in section 1.3.5.

The mixing of neutral B_d^0 mesons was first observed by the ARGUS collaboration [26]. The currently best value for the mixing frequency in the B_d^0 system is $\Delta m_d = (0.507 \pm 0.004)\,\mathrm{ps}^{-1}$ [27]. In 2006 mixing of neutral B_s^0 mesons was established by the Tevatron [28]. The CDF experiment determined the B_s^0 mixing frequency to be $\Delta m_s = (17.77 \pm 0.12)\,\mathrm{ps}^{-1}$ [28]. Using data taken in 2010 the LHCb experiment measures $\Delta m_s = (17.63 \pm 0.11_{\mathrm{stat.}} \pm 0.03_{\mathrm{syst.}})\,\mathrm{ps}^{-1}$ [29] which is well compatible with this value.

1.2.6. Mixing phenomenology

This short overview of the phenomenology of B mixing and decay broadly follows the more detailed review articles [22, 30, 31, 32, 33]. Even though the formalism developed below describes neutral B mesons B_q^0 it is also valid for K and D mesons.

Decays of neutral mesons B_q^0 and their CP conjugates $\bar{\mathrm{B}}_\mathrm{q}^0$ into final states f and $\bar{f}$ are described by the decay amplitudes

$$\begin{aligned} A_f &= \langle f|\mathcal{H}|\mathrm{B}_\mathrm{q}^0\rangle \\ \bar{A}_f &= \langle f|\mathcal{H}|\bar{\mathrm{B}}_\mathrm{q}^0\rangle \\ A_{\bar{f}} &= \langle \bar{f}|\mathcal{H}|\mathrm{B}_\mathrm{q}^0\rangle \\ \bar{A}_{\bar{f}} &= \langle \bar{f}|\mathcal{H}|\bar{\mathrm{B}}_\mathrm{q}^0\rangle. \end{aligned}$$

Neutral B-mesons transform under CP operation according to

$$\begin{aligned} \mathrm{CP}|\mathrm{B}_\mathrm{q}^0\rangle &= -|\bar{\mathrm{B}}_\mathrm{q}^0\rangle \text{ and} \\ \mathrm{CP}|\bar{\mathrm{B}}_\mathrm{q}^0\rangle &= -|\mathrm{B}_\mathrm{q}^0\rangle \end{aligned}$$

where an arbitrary non-physical phase factor has been omitted. If the final state f is a CP eigenstate f and $\bar{f}$ are connected via the CP operation according to

$$\begin{aligned} \mathrm{CP}|f\rangle &= \eta_f|\bar{f}\rangle \text{ and} \\ \mathrm{CP}|\bar{f}\rangle &= \eta_f|f\rangle \end{aligned}$$

with $\eta_f = \pm 1$ denoting the CP eigenvalue of the final state.

The time developement of the flavor eigenstates $|\mathrm{B}_\mathrm{q}^0\rangle$ and $|\bar{\mathrm{B}}_\mathrm{q}^0\rangle$ is given by the phenomenological Schrödinger equation,

$$\begin{aligned} i\frac{\partial}{\partial t}\begin{pmatrix}|\mathrm{B}_\mathrm{q}^0\rangle \\ |\bar{\mathrm{B}}_\mathrm{q}^0\rangle\end{pmatrix} &= \begin{pmatrix}\mathrm{M} - \frac{i}{2}\Gamma\end{pmatrix}\begin{pmatrix}|\mathrm{B}_\mathrm{q}^0\rangle \\ |\bar{\mathrm{B}}_\mathrm{q}^0\rangle\end{pmatrix} \\ &= \begin{pmatrix}\mathrm{M}_{11} - \frac{\mathrm{i}}{2}\Gamma_{11} & \mathrm{M}_{12} - \frac{\mathrm{i}}{2}\Gamma_{12} \\ \mathrm{M}_{21} - \frac{\mathrm{i}}{2}\Gamma_{21} & \mathrm{M}_{22} - \frac{\mathrm{i}}{2}\Gamma_{22}\end{pmatrix}\begin{pmatrix}|\mathrm{B}_\mathrm{q}^0\rangle \\ |\bar{\mathrm{B}}_\mathrm{q}^0\rangle\end{pmatrix}, \end{aligned}$$

where the Hamiltonian is constructed from two hermitean matrices, the mass matrix M and the decay matrix Γ. The mass matrix describes B_q^0-$\bar{\mathrm{B}}_\mathrm{q}^0$ mixing which was discussed in section 1.2.5. In the Standard Model the off-diagonal terms M_{12} and M_{12}^* stem from the flavor changing $\Delta b = 2$ processes given in Figure 1.3. Due to the hermeticity of M and Γ the off-diagonal elements are complex conjugates, $\mathrm{M}_{21} = \mathrm{M}_{12}^*$ and $\Gamma_{21} = \Gamma_{12}^*$. Additionaly, CPT invariance gives $\Gamma = \Gamma_{11} = \Gamma_{22}$ and $\mathrm{M} = \mathrm{M}_{11} = \mathrm{M}_{22}$ so that the Hamiltonian simplifies to

$$i\frac{\partial}{\partial t}\begin{pmatrix}|\mathrm{B}_\mathrm{q}^0\rangle \\ |\bar{\mathrm{B}}_\mathrm{q}^0\rangle\end{pmatrix} = \begin{pmatrix}\mathrm{M} - \frac{i}{2}\Gamma & \mathrm{M}_{12} - \frac{\mathrm{i}}{2}\Gamma_{12} \\ \mathrm{M}_{12}^* - \frac{\mathrm{i}}{2}\Gamma_{12}^* & \mathrm{M} - \frac{i}{2}\Gamma\end{pmatrix}\begin{pmatrix}|\mathrm{B}_\mathrm{q}^0\rangle \\ |\bar{\mathrm{B}}_\mathrm{q}^0\rangle\end{pmatrix}. \tag{1.14}$$

Diagonalizing the Hamiltonian in equation 1.14 leads to the mass eigenstates $|\mathrm{B_L}\rangle$ and $|\mathrm{B_H}\rangle$.

$$\begin{aligned} |\mathrm{B_L}\rangle &= p|\mathrm{B}_\mathrm{q}^0\rangle + q|\bar{\mathrm{B}}_\mathrm{q}^0\rangle \\ |\mathrm{B_H}\rangle &= p|\mathrm{B}_\mathrm{q}^0\rangle - q|\bar{\mathrm{B}}_\mathrm{q}^0\rangle \end{aligned}$$

with $|p|^2 + |q|^2 = 1$. $|\mathrm{B_L}\rangle$ denotes the lighter and $|\mathrm{B_H}\rangle$ the heavier mass eigenstate. The mass eigenstates develop in time with

$$\begin{aligned} |\mathrm{B_L}(t)\rangle &= e^{-i\mathrm{M_L t}} e^{-\frac{\Gamma_\mathrm{L}}{2}t} |\mathrm{B_L}\rangle \\ |\mathrm{B_H}(t)\rangle &= e^{-i\mathrm{M_H t}} e^{-\frac{\Gamma_\mathrm{H}}{2}t} |\mathrm{B_H}\rangle, \end{aligned}$$

where $\mathrm{M_L}$ and $\mathrm{M_H}$ denote the masses and Γ_L and Γ_H the decay widths of $|\mathrm{B_L}\rangle$ and $|\mathrm{B_H}\rangle$. The diagonalization procedure relates $\mathrm{M_{L/H}}$ and $\Gamma_\mathrm{L/H}$ to the elements of the Hamiltonian in equation 1.14

$$\begin{aligned} \mathrm{M_{L/H}} - \frac{\mathrm{i}}{2}\Gamma_\mathrm{L/H} &= M - \frac{i}{2}\Gamma \mp \sqrt{\left(\mathrm{M_{12}} - \frac{\mathrm{i}}{2}\Gamma_{12}\right)\left(\mathrm{M_{12}^*} - \frac{\mathrm{i}}{2}\Gamma_{12}^*\right)} \\ &= M - \frac{i}{2}\Gamma \mp \sqrt{|\mathrm{M_{12}}|^2 - \frac{1}{4}|\Gamma_{12}|^2 - i\,|\mathrm{M_{12}}|\,|\Gamma_{12}|\cos(\phi_\Gamma - \phi_\mathrm{M})} \end{aligned} \quad (1.15)$$

with the phases $\phi_\Gamma = \arg\Gamma_{12}$ and $\phi_\mathrm{M} = \arg\mathrm{M_{12}}$. The masses and decay widths of the mass eigenstates are related to Γ and M via

$$\begin{aligned} \Gamma &= \frac{\Gamma_\mathrm{L} + \Gamma_\mathrm{H}}{2} \\ M &= \frac{\mathrm{M_L} + \mathrm{M_H}}{2}. \end{aligned}$$

In the following the mass difference Δm and decay width difference $\Delta\Gamma$ is defined as

$$\begin{aligned} \Delta m &= \mathrm{M_H} - \mathrm{M_L} \\ \Delta\Gamma &= \Gamma_\mathrm{L} - \Gamma_\mathrm{H}. \end{aligned}$$

Using equation 1.15 the following relations for Δm and $\Delta\Gamma$ can be derived:

$$\Delta m^2 - \frac{1}{4}\Delta\Gamma^2 = 4\,|\mathrm{M_{12}}|^2 - |\Gamma_{12}|^2 \quad (1.16)$$

$$\Delta m \Delta\Gamma = -4\,|\mathrm{M_{12}}|\,|\Gamma_{12}|\cos(\phi_\Gamma - \phi_\mathrm{M}). \quad (1.17)$$

In both the $\mathrm{B_d^0}$ and the $\mathrm{B_s^0}$ system experimental evidence shows $\Delta m \gg \Delta\Gamma$ which translates to $|\mathrm{M_{12}}| \gg |\Gamma_{12}|$. Neglecting $\Delta\Gamma$ and $|\Gamma_{12}|^2$ in equation 1.16 results in

$$\Delta m = 2\,|\mathrm{M_{12}}|. \quad (1.18)$$

Inserting this in equation 1.17 gives

$$\Delta\Gamma = -2\,|\Gamma_{12}|\cos(\phi_\Gamma - \phi_\mathrm{M}). \quad (1.19)$$

The coefficients of the eigenstates are determined to be

$$\frac{q}{p} = -\sqrt{\frac{\mathrm{M_{12}^*} - \frac{\mathrm{i}}{2}\Gamma_{12}^*}{\mathrm{M_{12}} - \frac{\mathrm{i}}{2}\Gamma_{12}}}.$$

When studying the B_d^0 and the B_s^0 system this expression can be expanded in $|\Gamma_{12}| / |M_{12}|$ which gives

$$\begin{aligned}\frac{q}{p} &= -\sqrt{e^{-2i\phi_M}\frac{1-\frac{i}{2}\frac{|\Gamma_{12}|}{|M_{12}|}e^{-i\phi_\Gamma+i\phi_M}}{1-\frac{i}{2}\frac{|\Gamma_{12}|}{|M_{12}|}e^{+i\phi_\Gamma-i\phi_M}}} \\ &= -e^{-i\phi_M}\left[1-\frac{1}{2}\sin\left(\phi_\Gamma-\phi_M\right)\frac{|\Gamma_{12}|}{|M_{12}|}+\mathcal{O}\left(\frac{|\Gamma_{12}|^2}{|M_{12}|^2}\right)\right].\end{aligned}$$

The time developement for $|B_q^0\rangle$ and $|\bar{B}_q^0\rangle$ is given by

$$\begin{aligned}|B_q^0(t)\rangle &= \frac{1}{2p}\left(|B_L(t)\rangle + |B_H(t)\rangle\right) \\ |\bar{B}_q^0(t)\rangle &= \frac{1}{2q}\left(|B_L(t)\rangle - |B_H(t)\rangle\right)\end{aligned}$$

inserting $|B_L(t)\rangle$ and $|B_H(t)\rangle$ results in

$$\begin{aligned}|B_q^0(t)\rangle &= \frac{1}{2}\left(e^{-iM_L t}e^{-\frac{\Gamma_L}{2}t}+e^{-iM_H t}e^{-\frac{\Gamma_H}{2}t}\right)|B_q^0\rangle + \frac{q}{2p}\left(e^{-iM_L t}e^{-\frac{\Gamma_L}{2}t}-e^{-iM_H t}e^{-\frac{\Gamma_H}{2}t}\right)|\bar{B}_q^0\rangle \\ &= g_+(t)\,|B_q^0\rangle + \frac{q}{p}g_-(t)\,|\bar{B}_q^0\rangle \\ |\bar{B}_q^0(t)\rangle &= \frac{p}{2q}\left(e^{-iM_L t}e^{-\frac{\Gamma_L}{2}t}-e^{-iM_H t}e^{-\frac{\Gamma_H}{2}t}\right)|B_q^0\rangle + \frac{1}{2}\left(e^{-iM_L t}e^{-\frac{\Gamma_L}{2}t}+e^{-iM_H t}e^{-\frac{\Gamma_H}{2}t}\right)|\bar{B}_q^0\rangle \\ &= \frac{p}{q}g_-(t)\,|B_q^0\rangle + g_+(t)\,|\bar{B}_q^0\rangle\end{aligned}$$

with

$$g_\pm(t) = \frac{1}{2}\left(e^{-iM_L t}e^{-\frac{\Gamma_L}{2}t} \pm e^{-iM_H t}e^{-\frac{\Gamma_H}{2}t}\right).$$

For the determination of time dependent decay rates in the following it is convenient to explicitly calculate some combinations of these terms

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

To simplify the expressions it is further useful to define one central quantity for CP violation,

$$\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}.$$

While the phases of both $\bar{A}_f/A_f$ and q/p are convention dependent, the phase of λ_f is a measurable physical quantity. The time dependent decay rate for the decay of a produced B_q^0 to the final state f is

$$\begin{aligned}
\frac{d\Gamma\left(\mathrm{B}_\mathrm{q}^0 \to f\right)}{dt\mathcal{N}_f} &= \left|\langle f|\mathrm{B}_\mathrm{q}^0(t)\rangle\right|^2 \\
&= \left|g_+(t)\,A_f + \frac{q}{p}g_-(t)\,\bar{A}_f\right|^2 \\
&= |A_f|^2\left[|g_+(t)|^2 + |\lambda_f|^2\,|g_-(t)|^2 + \lambda_f^* g_+(t)\,g_-^*(t) + \lambda_f g_+^*(t)\,g_-(t)\right] \\
&= \frac{1}{2}|A_f|^2 e^{-\Gamma t}\left[\left(1+|\lambda_f|^2\right)\cosh\frac{\Delta\Gamma}{2}t + \left(1-|\lambda_f|^2\right)\cos\Delta m t\right. \\
&\qquad \left. -2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re\lambda_f - 2\sin(\Delta m t)\,\Im\lambda_f\right]. \qquad (1.20)
\end{aligned}$$

Similarly the decay rate for a $\bar{\mathrm{B}}_\mathrm{q}^0$ to decay to f is given by

$$\begin{aligned}
\frac{d\Gamma\left(\bar{\mathrm{B}}_\mathrm{q}^0 \to f\right)}{dt\mathcal{N}_f} &= \left|\langle f|\bar{\mathrm{B}}_\mathrm{q}^0(t)\rangle\right|^2 \\
&= \left|\frac{p}{q}g_-(t)\,A_f + g_+(t)\,\bar{A}_f\right|^2 \\
&= |A_f|^2\left|\frac{p}{q}\right|^2\left[|g_-(t)|^2 + |\lambda_f|^2\,|g_+(t)|^2 + \lambda_f^* g_+^*(t)\,g_-(t) + \lambda_f g_+(t)\,g_-^*(t)\right] \\
&= \frac{1}{2}\left|\frac{p}{q}\right|^2|A_f|^2 e^{-\Gamma t}\left[\left(1+|\lambda_f|^2\right)\cosh\frac{\Delta\Gamma}{2}t - \left(1-|\lambda_f|^2\right)\cos\Delta m t\right. \\
&\qquad \left. -2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re\lambda_f + 2\sin(\Delta m t)\,\Im\lambda_f\right] \qquad (1.21)
\end{aligned}$$

where $\mathcal{N}_f$ is a normalization factor. The decay rates to the CP conjugated final state $\bar{f}$ are given by the above expressions when substituting $A_f \to A_{\bar{f}}$, $\bar{A}_f \to \bar{A}_{\bar{f}}$ and $\lambda_f \to \lambda_{\bar{f}} = q/p\,\bar{A}_{\bar{f}}/A_{\bar{f}}$.

1.2.7. Types of CP violation

CP violation in B meson decays can be caused by three different mechanisms

- **CP violation in decay**
  This is the only type of CP violation possible for decays of charged mesons. It occurs when $\left|\bar{A}_{\bar{f}}/A_f\right| \neq 1$, i. e. the amplitudes for the process $B \to f$ and its CP conjugate $\bar{B} \to \bar{f}$ differ. CP violation then manifests itself as asymmetry

$$\begin{aligned}
\mathcal{A}_\mathrm{CP}^{dir} &= \frac{\Gamma\left(\mathrm{B}^- \to f^-\right) - \Gamma\left(\mathrm{B}^+ \to f^+\right)}{\Gamma\left(\mathrm{B}^- \to f^-\right) + \Gamma\left(\mathrm{B}^+ \to f^+\right)} \\
&= \frac{\left|\bar{A}_{\bar{f}}\right|^2 - |A_f|^2}{\left|\bar{A}_{\bar{f}}\right|^2 + |A_f|^2} = \frac{\left|\bar{A}_{\bar{f}}/A_f\right|^2 - 1}{\left|\bar{A}_{\bar{f}}/A_f\right|^2 + 1}. \qquad (1.22)
\end{aligned}$$

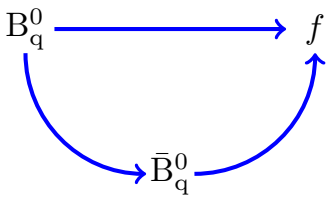


Figure 1.4.: Interference between decay and decay after mixing.

An example for this kind of CP asymmetry is the decay $B_d^0 \to K^+\pi^-$. In this decay the tree level amplitude interferes with penguin amplitudes leading to a CP asymmetry of $\sim 10\%$.

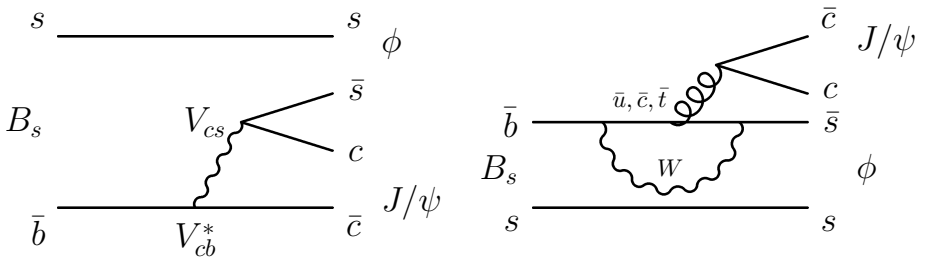
- **CP violation in mixing**
  CP violation in mixing occurs when $|q/p| \neq 1$. In this case the probability for a B_q^0 to transition into a $\bar{B}_q^0$, $\mathcal{P}(B_q^0 \to \bar{B}_q^0)$, differs from the probability for the CP conjugated process, $\mathcal{P}(\bar{B}_q^0 \to B_q^0)$. The resulting asymmetry assuming no direct CP violation, i. .e. $A_f = \bar{A}_{\bar{f}}$ and $A_{\bar{f}} = \bar{A}_f = 0$, is given by

$$\begin{aligned} \mathcal{A}_{\mathrm{CP}}^{mix} &= \frac{\Gamma\left(\bar{B}_q^0 \to f\right) - \Gamma\left(B_q^0 \to \overline{f}\right)}{\Gamma\left(\bar{B}_q^0 \to f\right) + \Gamma\left(B_q^0 \to \overline{f}\right)} \\ &= \frac{\left|\frac{p}{q} g_-(t)\, A_f\right|^2 - \left|\frac{q}{p} g_-(t)\, \bar{A}_{\bar{f}}\right|^2}{\left|\frac{p}{q} g_-(t)\, A_f\right|^2 + \left|\frac{q}{p} g_-(t)\, \bar{A}_{\bar{f}}\right|^2} = \frac{1 - \left|\frac{q}{p}\right|^4}{1 + \left|\frac{q}{p}\right|^4}. \end{aligned} \tag{1.23}$$

  CP violation in mixing can be studied using semileptonic decays $B_q^0 \to X\ell^+\nu$. Events where the B_q^0 has mixed before decaying semileptonically result in "wrong sign" decays containing a ℓ^- in the final state. Under the assumption of equal production of B_q^0 and $\bar{B}_q^0$ $\mathcal{P}(B_q^0 \to \bar{B}_q^0) \neq \mathcal{P}(\bar{B}_q^0 \to B_q^0)$ will lead to an asymmetry in the observed number of ℓ^+ and ℓ^- given by $\mathcal{A}_{\mathrm{CP}}^{mix}$.

- **Mixing induced CP violation**
  Mixing induced CP violation can occur when the direct decay $B_q^0 \to f$ interferes with mixing from B_q^0 to $\bar{B}_q^0$ followed by the decay $\bar{B}_q^0 \to f$. This situation is illustrated in Figure 1.4. If the term λ_f in equations 1.20 and 1.21 has a non-trivial phase, i. e. $\Im(\lambda_f) = \Im\left(q/p\, \bar{A}_f/A_f\right) \neq 0$, this gives rise to the time dependent CP asymmetry

$$\begin{aligned} \mathcal{A}_{\mathrm{CP}}(t) &= \frac{\Gamma\big(\bar{B}_q^0 \to f\big)(t) - \Gamma\big(B_q^0 \to \overline{f}\big)(t)}{\Gamma\big(\bar{B}_q^0 \to f\big)(t) + \Gamma\big(B_q^0 \to \overline{f}\big)(t)} \\ &= \frac{-\left(1 - |\lambda_f|^2\right)\cos(\Delta m t) + 2\sin(\Delta m t)\,\Im\lambda_f}{\left(1 + |\lambda_f|^2\right)\cosh\left(\frac{\Delta\Gamma}{2}t\right) - 2\sinh\left(\frac{\Delta\Gamma}{2}t\right)\Re\lambda_f}. \end{aligned} \tag{1.24}$$



(a) Tree contribution (b) Penguin contribution

Figure 1.5.: Both (a) tree and (b) penguin processes can contribute to the decay $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$, but the tree decay amplitude dominates.

For $\Delta\Gamma = 0$ and $|\lambda_f| = 1$ the asymmetry simplifies to

$$\mathcal{A}_{\mathrm{CP}}(t) = \sin(\Delta m t)\,\Im\lambda_f .$$

An example for this type of CP violation is the measurement of $\sin 2\beta_{\mathrm{d}}$ in the decay $\mathrm{B_d^0} \to \mathrm{J/\psi\,K_S^0}$. The analogous "golden mode" in the $\mathrm{B_s^0}$ system is given by the decay $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$ which is the focus of this thesis. This thesis therefore constitutes a measurement of a time dependent CP asymmetry induced by mixing.

1.3. The CP-violating phase ϕ_s in the decay $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$

The decay $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$ implies a $b \to c\bar{c}s$ quark transition as shown in Figure 1.5. The total amplitude $A_{\mathrm{J/\psi}\,\phi}$ for the transition of a $\mathrm{B_s^0}$ meson into the final state $\mathrm{J/\psi}\,\phi$ is given by the sum

$$\begin{aligned} A_{\mathrm{J/\psi\phi}} &= \mathrm{V_{cs}V_{cb}^*}T + \mathrm{V_{us}V_{ub}^*}P_u + \mathrm{V_{cs}V_{cb}^*}P_c + \mathrm{V_{ts}V_{tb}^*}P_t \\ &= \mathrm{V_{cs}V_{cb}^*}\left(T + P_c - P_t\right) + \mathrm{V_{us}V_{ub}^*}\left(P_u - P_t\right), \end{aligned}$$

where T denotes the tree amplitude (Figure 1.5a) and P_q with $q = u, c, t$ the corresponding penguin amplitudes (Figure 1.5b). Neglecting the terms proportional to $\mathrm{V_{us}V_{ub}^*}$ the amplitude ratio $\bar{A}_f/A_f$ is given by

$$\begin{aligned} \frac{\bar{A}_f}{A_f} &= -\eta_{\mathrm{J/\psi\phi}}\frac{\mathrm{V_{cb}V_{cs}^*}}{\mathrm{V_{cb}^*V_{cs}}} \\ &= -\eta_{\mathrm{J/\psi\phi}}e^{2i\phi_{\mathrm{D}}}, \end{aligned}$$

with the decay phase $\phi_{\mathrm{D}} = \arg\left(\mathrm{V_{cb}V_{cs}^*}\right)$ and the CP eigenvalue of the final state denoted by $\eta_{\mathrm{J/\psi\phi}}$. The same final state $\mathrm{J/\psi}\,\phi$ can be reached if the $\mathrm{B_s^0}$ first mixes to a $\mathrm{\bar{B}_s^0}$ (illustrated in Figure 1.3) and then decays. The ratio q/p is given by

$$\begin{aligned} \frac{q}{p} &= -\frac{\mathrm{V_{ts}V_{tb}^*}}{\mathrm{V_{ts}^*V_{tb}}} \\ &= -e^{-i\phi_{\mathrm{M}}} \end{aligned}$$

with the mixing phase $\phi_{\mathrm{M}} = -2\arg\left(\mathrm{V_{ts}V_{tb}^{*}}\right)$. Combining $\bar{A}_f/A_f$ and q/p to λ_f gives

$$\begin{aligned}\lambda_f &= \eta_{\mathrm{J}/\psi\phi}\frac{\mathrm{V_{cb}V_{cs}^{*}}}{\mathrm{V_{cb}^{*}V_{cs}}}\frac{\mathrm{V_{ts}V_{tb}^{*}}}{\mathrm{V_{ts}^{*}V_{tb}}}. \\ &= \eta_{\mathrm{J}/\psi\phi}e^{+2i\phi_{\mathrm{D}}-i\phi_{\mathrm{M}}} \\ &= \eta_{\mathrm{J}/\psi\phi}e^{-i\phi_s}\end{aligned}$$

with $\phi_s = -\arg\left(\eta_{\mathrm{J}/\psi\phi}\lambda_f\right)$. The phase ϕ_s can be related to the angle β_{s} in the $\mathrm{B_s^0}$ unitarity triangle via the relation $\phi_s = -2\beta_{\mathrm{s}}$ with $\beta_{\mathrm{s}} = \arg\left(-\mathrm{V_{ts}V_{tb}^{*}}/\left(\mathrm{V_{cs}V_{cb}^{*}}\right)\right)$ as defined in equation 1.13. While the mixing and decay phases depend on phase conventions and are not observable, the phase ϕ_s is a measurable physical quantity.

Looking at the Wolfenstein parametrization it is obvious that the the angle β_{s} will be much smaller than than β since $\mathrm{V_{ts}}$ is real up to $\mathcal{O}\left(\lambda^3\right)$ whereas $\mathrm{V_{td}}$ is already complex at $\mathcal{O}\left(\lambda^3\right)$. The theoretical prediction for β_{s} is indeed very small $\beta_{\mathrm{s}} = 0.01817^{+0.00087}_{-0.00083}\,\mathrm{rad}$ [1] which results in the precise Standard Model prediction $\phi_s = -0.0363 \pm 0.0017\,\mathrm{rad}$.

In contrast to the $\mathrm{B_d^0}$ system, the $\mathrm{B_s^0}$ system posseses a significant decay width difference $\Delta\Gamma_{\mathrm{s}}$, the SM predicts $\Delta\Gamma_{\mathrm{s}} = (0.087 \pm 0.021)\,\mathrm{ps}^{-1}$ [34], which means that the decay width difference may not be neglected.

In summary the time dependent mixing induced CP asymmetry for the signal decay $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ is given by

$$\mathcal{A}_{\mathrm{CP}}(t) = \frac{-\eta_{\mathrm{J}/\psi\phi}\sin\left(\phi_s\right)\sin\left(\Delta m_{\mathrm{s}}t\right)}{\cosh\left(\frac{\Delta\Gamma_{\mathrm{s}}}{2}t\right) - \eta_{\mathrm{J}/\psi\phi}\cos\left(\phi_s\right)\sinh\left(\frac{\Delta\Gamma_{\mathrm{s}}}{2}t\right)}. \tag{1.25}$$

1.3.1. Angular analysis

Since both J/ψ and ϕ are vector mesons, the decay $\mathrm{B_s^0} \to \mathrm{J}/\psi\,\phi$ is a pseudoscalar to vector-vector transition ($P \to VV$). This allows the final state mesons to have relative angular momentum l which leads to different CP eigenvalues of the final state depending on l,

$$\begin{aligned}\eta_{\mathrm{J}/\psi\phi} &= \eta_{\mathrm{J}/\psi}^{\mathrm{CP}}\eta_{\phi}^{\mathrm{CP}}\left(-1\right)^l \\ &= \left(-1\right)^l.\end{aligned}$$

To decouple the CP-odd ($\eta_{\mathrm{J}/\psi\phi} = -1$) and CP-even ($\eta_{\mathrm{J}/\psi\phi} = +1$) components statistically an analysis of the angular distributions is necessary. The dependence of the differential decay rates on the decay angles is expressed in the so called transversity base which is shown in Figure 1.6 and defines the transversity angles $\cos\theta$, φ and $\cos\psi$.

The angles θ and φ are defined in the J/ψ rest frame. The flight direction of the ϕ meson defines the x-axis, the ϕ decay products K^+K^- the $x-y$ plane. The direction of the y-axis is chosen such that $p_y\left(K^+\right) > 0$. The angle θ is defined as the angle between μ^+ flight direction relative to the z-axis, while φ is defined as the angle between the x-axis and the projection of the μ^+ direction onto the $x-y$ plane. The angle ψ is defined in the ϕ rest frame. It is the angle of the K^+ relative to the negative flight direction of the J/ψ.

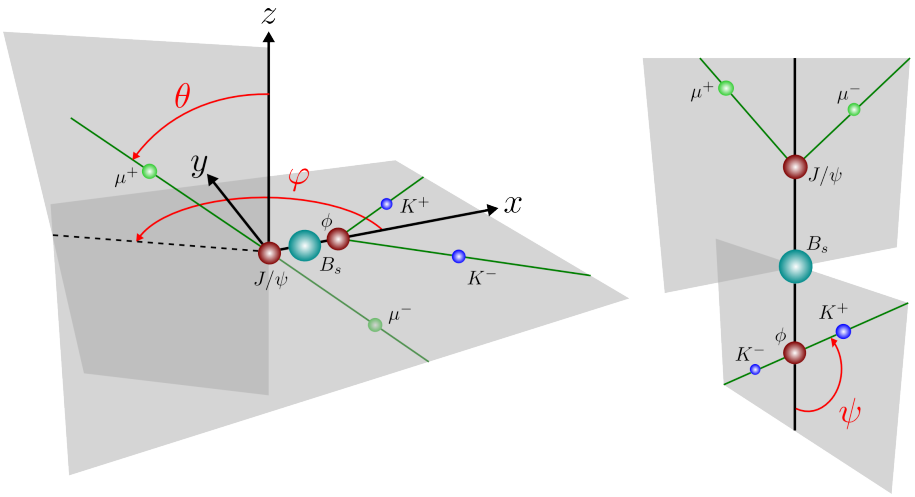


Figure 1.6.: The transversity angles. Figure from [35].

1.3.2. Differential decay rates

The time and angular dependent differential decay rate for a B_s^0 meson produced at $t = 0$ to decay to the final state $J/\psi\phi$ is given by

$$\frac{d\Gamma\left(B_s^0 \to J/\psi\,\phi\right)}{dt\,d\Omega} \propto \sum_{i=1}^{6} A_i\left(t\right) \cdot f_i\left(\cos\theta, \varphi, \cos\psi\right), \tag{1.26}$$

where A_i are six combinations of the three complex transversity amplitudes $A_0(t)$, $A_\perp(t)$ and $A_\parallel(t)$ [36]. The transversity amplitudes correspond to different linear polarization states of the vector mesons J/ψ and ϕ with respect to their direction of motion. $A_0(t)$ corresponds to longitudinal polarization of the vector mesons. The transversity amplitudes $A_\parallel(t)$ and $A_\perp(t)$ correspond to transverse polarization of the vector mesons, in the former case the polarization states of J/ψ and ϕ are parallel in the latter case perpendicular to each other.

The relative phases of the transversity amplitudes at $t = 0$ are denoted as $\delta_0 = \arg\left(A_0(0)\right)$, $\delta_\parallel = arg\left(A_\parallel(0)\right)$ and $\delta_\perp = \arg\left(A_\perp(0)\right)$. The phases arise from strong final state interactions and are difficult to predict in theory. As they are invariant under CP they are also called "strong phases". One of these phases can be chosen freely since only the relative phases are important, in this thesis the convention $\delta_0 = 0$ is used. The magnitudes of the transversity amplitudes fullfill the normalization condition $|A_0(0)|^2 + |A_\perp(0)|^2 + \left|A_\parallel(0)\right|^2 = 1$. The angular dependent terms $f_i\left(\cos\theta, \varphi, \cos\psi\right)$ are given in [31] and summarized in Table 1.4.

The prescription used for produced $\bar{B}_s^0$ mesons at $t = 0$ is very similar and uses $\bar{A}_0(t)$, $\bar{A}_\perp(t)$ and $\bar{A}_\parallel(t)$

$$\frac{d\Gamma\left(\bar{B}_s^0 \to J/\psi\,\phi\right)}{dt\,d\Omega} \propto \sum_{i=1}^{6} \bar{A}_i\left(t\right) \cdot f_i\left(\cos\theta, \varphi, \cos\psi\right) \tag{1.27}$$

with the same angular dependent terms $f_i\left(\cos\theta, \varphi, \cos\psi\right)$. The time dependent amplitude combinations $A_i\left(\mathrm{t}\right)$ are given in [31]. For a $\mathrm{B_s^0}$ meson they are defined as

$$
\begin{aligned}
|A_0(\mathrm{t})|^2 &= |A_0(0)|^2 e^{-\Gamma_\mathrm{s} t}\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \sin\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\right] \\
\left|A_\parallel(\mathrm{t})\right|^2 &= \left|A_\parallel(0)\right|^2 e^{-\Gamma_\mathrm{s} t}\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \sin\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\right] \\
|A_\perp(\mathrm{t})|^2 &= |A_\perp(0)|^2 e^{-\Gamma_\mathrm{s} t}\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \sin\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\right] \\
\Im(A_\parallel^*(\mathrm{t})A_\perp(\mathrm{t})) &= \left|A_\parallel(0)\right| |A_\perp(0)| e^{-\Gamma_\mathrm{s} t}\Bigg[-\cos\left(\delta_\perp - \delta_\parallel\right)\sin\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) \\
&\quad + \sin\left(\delta_\perp - \delta_\parallel\right)\cos\left(\Delta m_\mathrm{s} t\right) - \cos\left(\delta_\perp - \delta_\parallel\right)\cos\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\Bigg] \\
\Re(A_0^*(\mathrm{t})A_\parallel(\mathrm{t})) &= |A_0(0)| \left|A_\parallel(0)\right| e^{-\Gamma_\mathrm{s} t}\cos\delta_\parallel\Bigg[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) \\
&\quad + \sin\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\Bigg] \\
\Im(A_0^*(\mathrm{t})A_\perp(\mathrm{t})) &= |A_0(0)| |A_\perp(0)| e^{-\Gamma_\mathrm{s} t}\Bigg[-\cos\delta_\perp \sin\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) \\
&\quad + \sin\delta_\perp \cos\left(\Delta m_\mathrm{s} t\right) - \cos\delta_\perp \cos\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\Bigg]
\end{aligned}
$$

The time dependent amplitudes for a $\mathrm{\bar{B}_s^0}$ are found by changing the signs in front of all terms proportional to $\sin\left(\Delta m_\mathrm{s} t\right)$ or $\cos\left(\Delta m_\mathrm{s} t\right)$ which results in

$$
\begin{aligned}
\left|\bar{A}_0(\mathrm{t})\right|^2 &= \left|\bar{A}_0(0)\right|^2 e^{-\Gamma_\mathrm{s} t}\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \sin\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\right] \\
\left|\bar{A}_\parallel(\mathrm{t})\right|^2 &= \left|\bar{A}_\parallel(0)\right|^2 e^{-\Gamma_\mathrm{s} t}\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \sin\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\right] \\
\left|\bar{A}_\perp(\mathrm{t})\right|^2 &= \left|\bar{A}_\perp(0)\right|^2 e^{-\Gamma_\mathrm{s} t}\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \sin\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\right]
\end{aligned}
$$

i	$A_i\,(t)$	$\bar{A}_i\,(t)$	$f_i\,(\cos\theta, \varphi, \cos\psi)$
1	$\lvert A_0(t)\rvert^2$	$\lvert \bar{A}_0(t)\rvert^2$	$\frac{9}{32\pi} 2\cos^2\psi\left(1-\sin^2\theta\cos^2\varphi\right)$
2	$\lvert A_\parallel(t)\rvert^2$	$\lvert \bar{A}_\parallel(t)\rvert^2$	$\frac{9}{32\pi}\sin^2\psi\left(1-\sin^2\theta\sin^2\varphi\right)$
3	$\lvert A_\perp(t)\rvert^2$	$\lvert \bar{A}_\perp(t)\rvert^2$	$\frac{9}{32\pi}\sin^2\psi\sin^2\theta$
4	$\Im\left(A_\parallel^*(t)A_\perp(t)\right)$	$\Im\left(\bar{A}_\parallel^*(t)\bar{A}_\perp(t)\right)$	$-\frac{9}{32\pi}\sin^2\psi\sin 2\theta\sin\varphi$
5	$\Re\left(A_0^*(t)A_\parallel(t)\right)$	$\Re\left(\bar{A}_0^*(t)\bar{A}_\parallel(t)\right)$	$\frac{9}{32\pi\sqrt{2}}\sin 2\psi\sin^2\theta\sin 2\varphi$
6	$\Im\left(A_0^*(t)A_\perp(t)\right)$	$\Im\left(\bar{A}_0^*(t)\bar{A}_\perp(t)\right)$	$\frac{9}{32\pi\sqrt{2}}\sin 2\psi\sin 2\theta\cos\varphi$

Table 1.4.: Angular dependent terms $f_i\,(\cos\theta, \varphi, \cos\psi)$ for the P-wave [31].

$$
\begin{aligned}
\Im(\bar{A}_\parallel^*(t)\bar{A}_\perp(t)) &= \left|\bar{A}_\parallel(0)\right|\left|\bar{A}_\perp(0)\right| e^{-\Gamma_s t}\Bigg[-\cos\left(\delta_\perp-\delta_\parallel\right)\sin\phi_s\sinh\left(\frac{\Delta\Gamma_s}{2}t\right) \\
&\quad -\sin\left(\delta_\perp-\delta_\parallel\right)\cos\left(\Delta m_s t\right)+\cos\left(\delta_\perp-\delta_\parallel\right)\cos\phi_s\sin\left(\Delta m_s t\right)\Bigg] \\
\Re(\bar{A}_0^*(t)\bar{A}_\parallel(t)) &= \left|\bar{A}_0(0)\right|\left|\bar{A}_\parallel(0)\right| e^{-\Gamma_s t}\cos\delta_\parallel\Bigg[\cosh\left(\frac{\Delta\Gamma_s}{2}t\right)-\cos\phi_s\sinh\left(\frac{\Delta\Gamma_s}{2}t\right) \\
&\quad -\sin\phi_s\sin\left(\Delta m_s t\right)\Bigg] \\
\Im(\bar{A}_0^*(t)\bar{A}_\perp(t)) &= \left|\bar{A}_0(0)\right|\left|\bar{A}_\perp(0)\right| e^{-\Gamma_s t}\Bigg[-\cos\delta_\perp\sin\phi_s\sinh\left(\frac{\Delta\Gamma_s}{2}t\right) \\
&\quad -\sin\delta_\perp\cos\left(\Delta m_s t\right)+\cos\delta_\perp\cos\phi_s\sin\left(\Delta m_s t\right)\Bigg]
\end{aligned}
$$

The decay rates for the decay $B_s^0 \to J/\psi\,\phi$ given above exhibit an interesting two-fold symmetry under the substitutions

$$
\begin{aligned}
\phi_s &\to \pi-\phi_s \\
\Delta\Gamma_s &\to -\Delta\Gamma_s \\
\delta_\parallel &\to -\delta_\parallel \\
\delta_\perp &\to \pi-\delta_\perp .
\end{aligned}
\tag{1.28}
$$

It is therefore expected to find two solutions when extracting the physics parameters from the data.

1.3.3. S-wave contribution

The differential decay rates given in the previous section assume that the detected K^+K^- system in the final state originates solely from the decay of the ϕ resonance. Since the ϕ is a vector meson the K^+K^- system is in a P-wave configuration. In addition to this contribution there is however the possibility that the detected K^+K^- results from a non-resonant $\ell = 0$ contribution or decay of the $f_0(980)$ which is a scalar meson [37].

In both cases the K^+K^- system would be in an S-wave configuration. These decays represent an irreducible contribution to the final state $\mu^+\mu^-K^+K^-$ of the signal decay. The contribution to the differential decay rates can be described by the introduction of the S-wave amplitude $A_s(t)$ with phase δ_s. The S-wave amplitude can also interfere with the P-wave amplitudes so equations 1.26 and 1.27 need to be modified [38] to

$$\frac{\mathrm{d}\Gamma\left(\mathrm{B_s^0} \to \mathrm{J}\!/\!\psi\, K^+K^-\right)}{\mathrm{d}t\mathrm{d}\Omega} \propto \sum_{i=1}^{10} A_i\left(\mathrm{t}\right) \cdot f_i\left(\cos\theta, \varphi, \cos\psi\right) \tag{1.29}$$

$$\frac{\mathrm{d}\Gamma\left(\bar{\mathrm{B}}_\mathrm{s}^0 \to \mathrm{J}\!/\!\psi\, K^+K^-\right)}{\mathrm{d}t\mathrm{d}\Omega} \propto \sum_{i=1}^{10} \bar{A}_i\left(\mathrm{t}\right) \cdot f_i\left(\cos\theta, \varphi, \cos\psi\right). \tag{1.30}$$

The additional amplitude combinations $A_{i=7\ldots10}$ are given by [38] as

$$\begin{aligned}
|A_\mathrm{s}(\mathrm{t})|^2 &= |A_\mathrm{s}(0)|^2\, e^{-\Gamma_\mathrm{s}t}\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \sin\phi_s \sin\left(\Delta m_\mathrm{s}t\right)\right] \\
\Re\left(A_\mathrm{s}^*(\mathrm{t})A_\parallel(\mathrm{t})\right) &= |A_\mathrm{s}(0)|\left|A_\parallel(0)\right| e^{-\Gamma_\mathrm{s}t}\Bigg[-\sin\left(\delta_\parallel - \delta_\mathrm{s}\right)\sin\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) \\
&\quad - \sin\left(\delta_\parallel - \delta_\mathrm{s}\right)\cos\phi_s \sin\left(\Delta m_\mathrm{s}t\right) + \cos\left(\delta_\parallel - \delta_\mathrm{s}\right)\cos\left(\Delta m_\mathrm{s}t\right)\Bigg] \\
\Im\left(A_\mathrm{s}^*(\mathrm{t})A_\perp(\mathrm{t})\right) &= |A_\perp(0)|\,|A_\mathrm{s}(0)|\, e^{-\Gamma_\mathrm{s}t}\sin\left(\delta_\perp - \delta_\mathrm{s}\right)\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right)\right. \\
&\quad \left. - \sin\phi_s \sin\left(\Delta m_\mathrm{s}t\right)\right] \\
\Re\left(A_\mathrm{s}^*(\mathrm{t})A_0(\mathrm{t})\right) &= |A_0(0)|\,|A_\mathrm{s}(0)|\, e^{-\Gamma_\mathrm{s}t}\Bigg[\sin\delta_\mathrm{s}\sin\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \sin\delta_\mathrm{s}\cos\phi_s \sin\left(\Delta m_\mathrm{s}t\right) \\
&\quad + \cos\delta_\mathrm{s}\cos\left(\Delta m_\mathrm{s}t\right)\Bigg]
\end{aligned}$$

i	$A_i\,(\mathrm{t})$	$\bar{A}_i\,(\mathrm{t})$	$f_i\,(\cos\theta, \varphi, \cos\psi)$
7	$A_\mathrm{s}(\mathrm{t})$	$\bar{A}_\mathrm{s}(\mathrm{t})$	$\frac{3}{32\pi} 2\left(1-\sin^2\theta\cos^2\varphi\right)$
8	$\Re\left(A_\mathrm{s}^*(\mathrm{t})A_\parallel(\mathrm{t})\right)$	$\Re\left(\bar{A}_\mathrm{s}^*(\mathrm{t})\bar{A}_\parallel(\mathrm{t})\right)$	$\frac{3}{32\pi}\sqrt{6}\sin^2\theta\sin\psi\sin 2\varphi$
9	$\Im\left(A_\mathrm{s}^*(\mathrm{t})A_\perp(\mathrm{t})\right)$	$\Im\left(\bar{A}_\mathrm{s}^*(\mathrm{t})\bar{A}_\perp(\mathrm{t})\right)$	$\frac{3}{32\pi}\sqrt{6}\sin 2\theta\sin\psi\cos\varphi$
10	$\Re\left(A_\mathrm{s}^*(\mathrm{t})A_0(\mathrm{t})\right)$	$\Re\left(\bar{A}_\mathrm{s}^*(\mathrm{t})\bar{A}_0(\mathrm{t})\right)$	$\frac{3}{32\pi}4\sqrt{3}\cos\psi\left(1-\sin^2\theta\cos^2\varphi\right)$

Table 1.5.: Angular dependent terms $f_i\,(\cos\theta, \varphi, \cos\psi)$ for the S-wave [38].

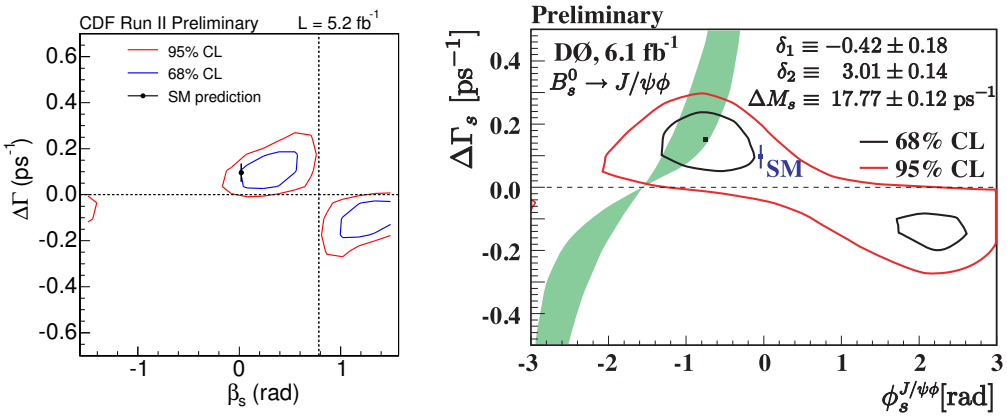
and the combinations $\bar{A}_{i=7\ldots10}$ as

$$
\begin{aligned}
\left|\bar{A}_\mathrm{s}(\mathrm{t})\right|^2 &= \left|\bar{A}_\mathrm{s}(0)\right|^2 e^{-\Gamma_\mathrm{s} t}\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \cos\phi_s \sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \sin\phi_s \sin\left(\Delta m_\mathrm{s} t\right)\right] \\
\Re\left(\bar{A}_\mathrm{s}^*(\mathrm{t})\bar{A}_\parallel(\mathrm{t})\right) &= \left|\bar{A}_\mathrm{s}(0)\right|\left|\bar{A}_\parallel(0)\right| e^{-\Gamma_\mathrm{s} t}\left[-\sin\left(\delta_\parallel-\delta_\mathrm{s}\right)\sin\phi_s\sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right)\right. \\
&\quad \left. + \sin\left(\delta_\parallel-\delta_\mathrm{s}\right)\cos\phi_s\sin\left(\Delta m_\mathrm{s} t\right) - \cos\left(\delta_\parallel-\delta_\mathrm{s}\right)\cos\left(\Delta m_\mathrm{s} t\right)\right] \\
\Im\left(\bar{A}_\mathrm{s}^*(\mathrm{t})\bar{A}_\perp(\mathrm{t})\right) &= \left|\bar{A}_\perp(0)\right|\left|\bar{A}_\mathrm{s}(0)\right| e^{-\Gamma_\mathrm{s} t}\sin\left(\delta_\perp-\delta_\mathrm{s}\right)\left[\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) + \cos\phi_s\sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right)\right. \\
&\quad \left. + \sin\phi_s\sin\left(\Delta m_\mathrm{s} t\right)\right] \\
\Re\left(\bar{A}_\mathrm{s}^*(\mathrm{t})\bar{A}_0(\mathrm{t})\right) &= \left|\bar{A}_0(0)\right|\left|\bar{A}_\mathrm{s}(0)\right| e^{-\Gamma_\mathrm{s} t}\left[\sin\delta_\mathrm{s}\sin\phi_s\sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \sin\delta_\mathrm{s}\cos\phi_s\sin\left(\Delta m_\mathrm{s} t\right)\right. \\
&\quad \left. - \cos\delta_\mathrm{s}\cos\left(\Delta m_\mathrm{s} t\right)\right].
\end{aligned}
$$

In presence of an S-wave contribution the magnitudes of the amplitudes need to fullfill the new normalization condition $|A_0(0)|^2 + |A_\perp(0)|^2 + \left|A_\parallel(0)\right|^2 + |A_\mathrm{s}(0)|^2 = 1$. The angular dependent terms $f_{i=7\ldots10}$ are given in Table 1.5. The CDF collaboration has determined the S-wave fraction to be smaller than 6.7% at the 95% confidence level [39].

1.3.4. Current experimental status of ϕ_s

The currently most precise determination of ϕ_s was performed by the Tevatron experiments CDF [39] and DØ [40]. Figure 1.7a gives confidence regions in the β_s-$\Delta\Gamma_\mathrm{s}$ parameter space determined by the CDF experiment on a data set corresponding to an integrated luminosity of $5.2\,\mathrm{fb}^{-1}$. Note that CDF chose to quote confidence contours depending on β_s which is related to ϕ_s via $\beta_\mathrm{s} = -\phi_s/2$. Figure 1.7b gives the measurement of ϕ_s performed by the DØ experiment. Here the contours are given in the ϕ_s-$\Delta\Gamma_\mathrm{s}$ parameter space. Both experiments favor negative values for ϕ_s, the p-value of the Standard Model hypothesis corresponds to ~ 1 standard deviations in both cases.



(a) β_s-$\Delta\Gamma_s$ plane from CDF[39] (b) ϕ_s-$\Delta\Gamma_s$ plane from DØ [40]

Figure 1.7.: Current experimental status of ϕ_s from (a) the CDF and (b) DØ experiments. Shown are coverage corrected confidence regions in the ϕ_s (or β_s) and $\Delta\Gamma_s$ parameter space. Figures from [39] and [40].

1.3.5. Beyond the Standard Model contributions to B mixing

Physics processes beyond the Standard Model (New Physics) can manifest themselves by introducing additional contributions to B^0_d and B^0_s mixing. This would affect the off-diagonal matrix elements M_{12} in the Hamiltonian in equation 1.14. Figure 1.8 gives a possible contribution through gluino diagrams [41]. For an overview of New Physics scenarios that can introduce additional contributions to B mixing see [42].

It is convenient to parametrize possible New Physics contributions to B mixing in a model independent way. In [42] complex factors Δ_d and Δ_s are introduced to parameterize the effect on M^d_{12} and M^s_{12} according to

$$\begin{aligned} M^d_{12} &= M^{d,SM}_{12}\Delta_d \\ M^s_{12} &= M^{s,SM}_{12}\Delta_s \text{ with} \\ \Delta_d &= |\Delta_d| e^{i\phi^\Delta_d} \text{ and} \\ \Delta_s &= |\Delta_s| e^{i\phi^\Delta_s}. \end{aligned}$$

The Standard Model hypothesis is of course $\Delta_d = \Delta_s = 1$. If the phase ϕ^Δ_s is nonzero this would directly modify the phase of λ_f leading to the measurement of

$$\phi_s = \phi^{SM}_s + \phi^\Delta_s.$$

Assuming the independence of Δ_d and Δ_s a global fit of the CKM matrix elements and the parameters Δ_d and Δ_s describing New Physics contributions in the B sector was performed in [42]. Figure 1.9 shows the results of the fit for the parameters Δ_d and Δ_s in the complex plane. The fit disfavors the Standard Model prediction for B^0_d mixing, $\Delta_d = 1$, by 2.7 standard deviations. For B^0_s mixing the deviation from the

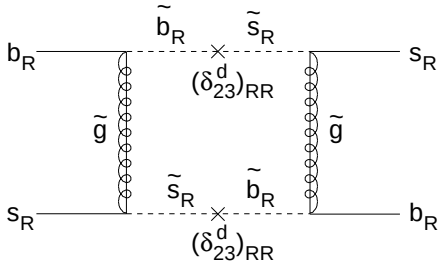
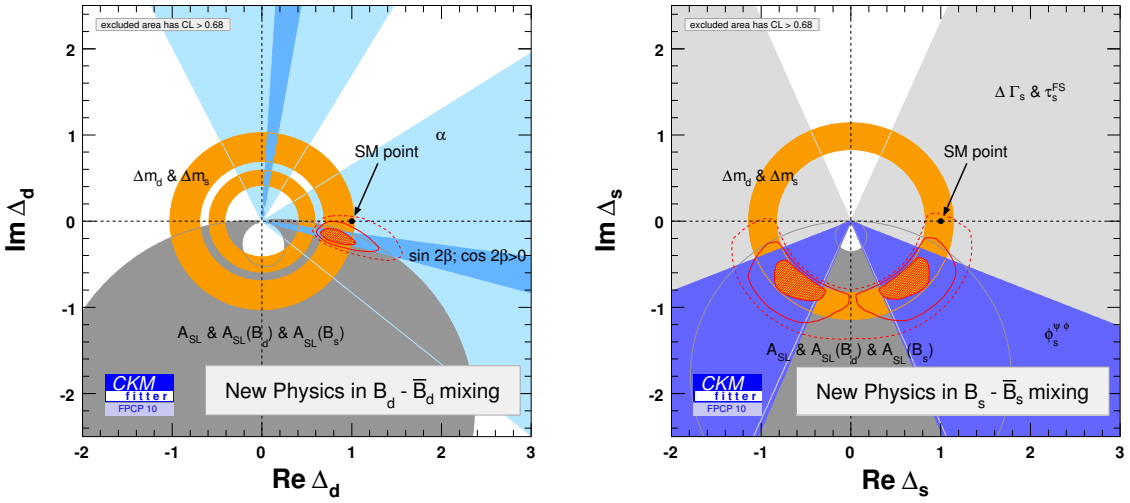


Figure 1.8.: Example for a possible contribution to B_s^0 mixing through gluino diagrams. Figure from [41].

Standard Model value $\Delta_s = 1$ corresponds to 2.7σ as well. The combined probability of $\Delta_d = \Delta_s = 1$ is found to be equivalent to 3.6 standard deviations [42]. These hints at possible problems of the Standard Model description of B mixing represent a strong motivation for a more precise determination of ϕ_s.

1.4. B-meson Production at LHCb

B-hadrons at the LHC are produced in the hadronization of $b\bar{b}$ quark pairs. $b\bar{b}$ pairs are produced via the strong interaction, through gluon and parton fusion. This is shown in Figure 1.10 where the Feynman graphs of the leading order production channels are given. The production cross section for $b\bar{b}$ pairs is strongly dependent on the polar angle to the beam axis. The reason for this are the, compared with the beam energy, low mass of the $b\bar{b}$ system in conjuntion with the parton distribution functions of quarks and gluons in the colliding protons. Since the quark and the antiquark (or the two gluons) which produce the $b\bar{b}$ system can carry very different momentum fractions of the proton, the $b\bar{b}$ system will likely be boosted in either forward or backward direction. Figure 1.11 shows the polar angles of the $b\bar{b}$ quarks as simulated by the PYTHIA event generator [43]. The forward peaking $b\bar{b}$ production cross section determines the geometry of the LHCb detector as forward spectrometer. Figure 1.12 shows the total $b\bar{b}$ production cross sections for different SM processes depending on the center of mass energy $\sqrt{s}$. The cited calculation shows $\sigma_{b\bar{b}} = 633\,\mu b$ at a nominal center of mass energy of 14 TeV. LHCb studies of Monte Carlo simulated events use conservative values of $\sigma_{b\bar{b}} = 500\,\mu b$ at $\sqrt{s} = 14\,\text{TeV}$ and $\sigma_{b\bar{b}} = 250\,\mu b$ at $\sqrt{s} = 7\,\text{TeV}$. A measurement using J/ψ mesons from b-hadron decays confirms this prediction measuring a total $b\bar{b}$ cross section of $\sigma_{b\bar{b}} = (288 \pm 4_{\text{stat}} \pm 48_{\text{syst}})\,\mu b$ [44] using data taken with the LHCb detector in 2010 at a center of mass energy of $\sqrt{s} = 7\,\text{TeV}$.



(a) B_d^0 sector (b) B_s^0 sector

Figure 1.9.: Global fit of possible New Physics contributions to (a) B_d^0 mixing and (b) B_s^0 mixing. Figure is from [1].

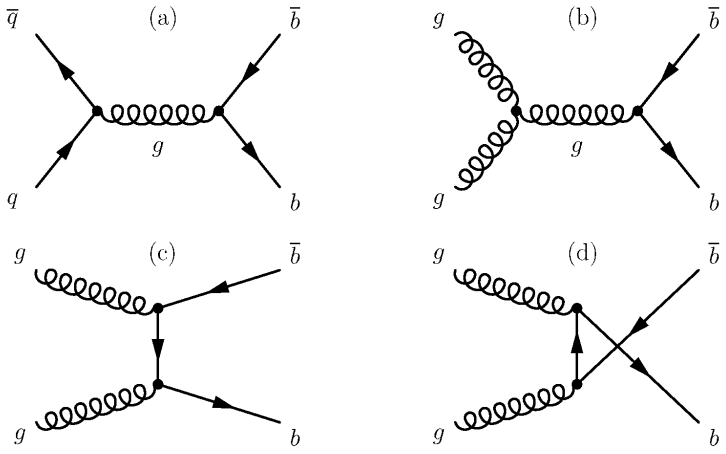


Figure 1.10.: Leading order $b\bar{b}$ production processes. Figure from [45].

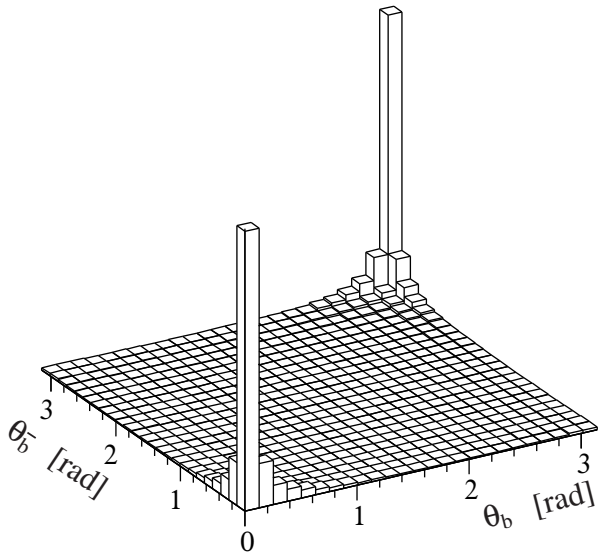


Figure 1.11.: Polar angles of the b and $\overline{\text{b}}$ quark from a simulation using the PYTHIA event generator [43]. Figure from [46].

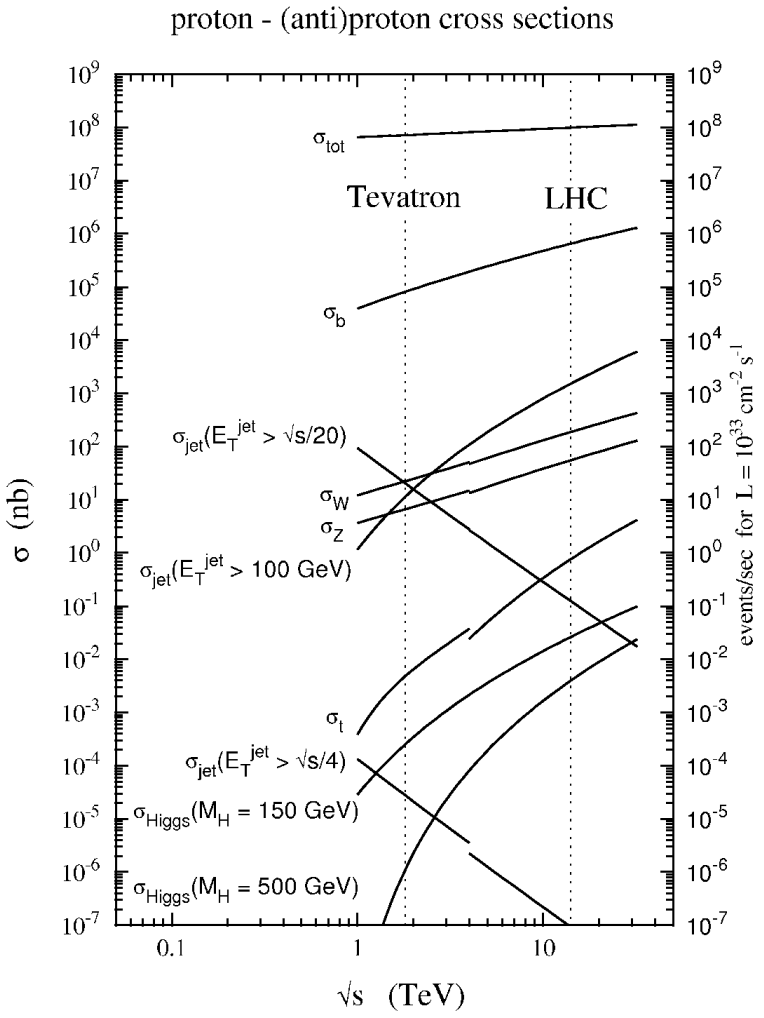


Figure 1.12.: Cross sections for different SM processes depending on $\sqrt{s}$. At $\sqrt{s} = 14\,\text{TeV}$ the values are $\sigma_{\text{tot}} = 99.4\,\text{mb}$, $\sigma_{\text{b}\bar{\text{b}}} = 633\,\mu\text{b}$. Figure from [47].

2. The LHCb experiment

To be able to access the CP violating phase ϕ_s large numbers of B_s^0 mesons are necessary. With a large $b\bar{b}$ cross section of $\sigma_{b\bar{b}} \sim 500\,\mu b$[1] the Large Hadron Collider (LHC) is well suited to study B decays. Key quantities of the LHC will be briefly introduced in the first part of this chapter. The LHCb detector which is located at one of the interaction points of the LHC is used to reconstruct the signal decay $B_s^0 \to J/\psi\,\phi$. The discussion of the LHCb detector and its subsystems constitutes the main part of this chapter. In addition a short overview of the LHCb software environment is given.

2.1. The Large Hadron Collider

The Large Hadron Collider (LHC) [48], located near Geneva on the Swiss-French border is a proton-proton collider. With a design energy of $2 \times 7\,\text{TeV}$ and a design luminosity of $\mathcal{L} = 1 \cdot 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ it is the most powerful particle accelerator to date. At nominal configuration 2808 bunches per beam, with $\sim 10^{11}$ protons each, are accelerated to 7 TeV in the 27 km long tunnel and brought to collision at the four interaction points. The magnetic fields to hold the beams in orbit are supplied by superconducting magnets cooled down to 1.9 K and operating at a nominal magnetic field strength of 8.34 T. An overview of the experimental site is shown in Figure 2.1. The four interaction points house two general purpose detectors, ATLAS and CMS, the ALICE detector which is specialized on heavy ion collisions and the LHCb experiment which studies the decays of B and D mesons. The proton bunches are designed to cross every 25 ns at each of the interaction points which leads to a bunch crossing rate of $R_{\text{crossing}} = 40\,\text{MHz}$. With the nominal luminosity of $\mathcal{L} = 1 \cdot 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ and a total pp cross section of $\sim 100\,\text{mb}$ the numbers of interactions per collision of two bunches can be calculated according to

$$N = \sigma_{\text{tot}} \int \mathcal{L}\text{dt}.$$

This results in a mean number of

$$\langle n \rangle = \frac{\sigma_{\text{tot}} \mathcal{L}}{R_{\text{crossing}}} = 25$$

interactions per bunch crossing[2].

During the startup in 2010 the LHC operated at a lower center of mass energy of $\sqrt{s} = 7\,\text{TeV}$ with up to ~ 400 bunches per beam and 150 ns bunch spacing. Using this configuration the LHC reached peak instantaneous luminosities of up to $\sim 2 \cdot 10^{32}\,\text{cm}^{-2}\text{s}^{-1}$.

[1] At 14 TeV.
[2] Poisson distributed

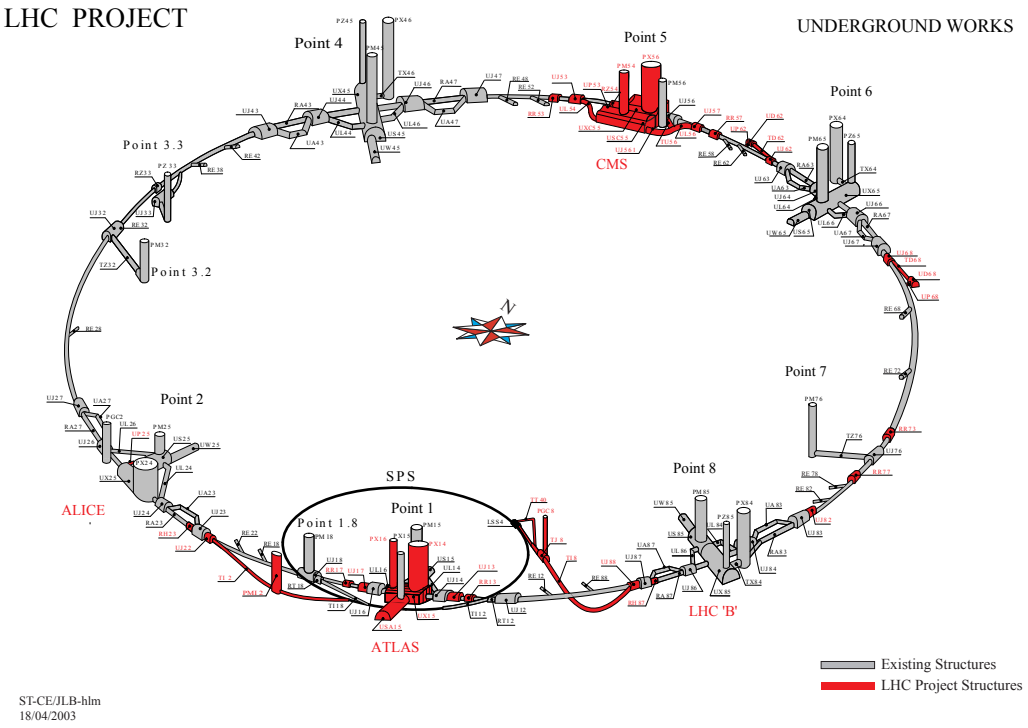


Figure 2.1.: Overview over the Large Hadron Collider. The four main experiments are Atlas (Point 1), Alice (Point 2), CMS (Point 5) and LHCb (Point 8). Figure from [49]

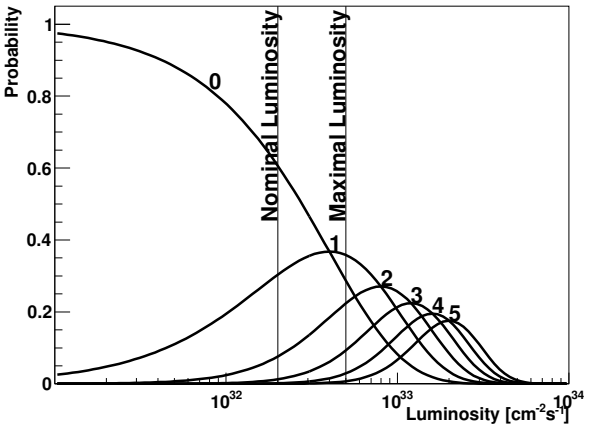


Figure 2.2.: Probabilities for different numbers of interactions per bunch crossing, shown are the probabilities for up to 5 interactions per bunch crossing. The nominal luminosity at LHCb was chosen to be $2 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, the maximal possible luminosity for which the detector was designed is $5 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The plot shows that there are predominantly one or zero interaction per bunch crossing at nominal luminosity.

2.2. The LHCb Experiment

The LHCb experiment is specialized for the study of B mesons and CP violation in the B meson sector. The high number of interactions per collision of the proton bunches severely impacts the ability to associate the B-decay vertices to the correct production vertices. Therefore the LHCb experiment has decided to run at a lower luminosity than the nominal $1 \cdot 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This is possible since the luminosity at the LHCb interaction point can be tuned by changing the beam focus. Figure 2.2 shows the poisson probabilities for different numbers of interactions per crossing depending on the chosen luminosity. It has been decided to run at a nominal luminosity of $2 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ where there are predominantly either one or zero interactions per bunch crossing, but in principle the detectors are designed to be able to run at luminosities of up to $5 \cdot 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This decision also positively impacts track resolution and lowers the occupancy of the detectors.

The specialization on B mesons determines the detector geometry. As discussed in section 1.4 the B mesons are produced predominantly in forward direction. Therefore LHCb is designed as a single-arm forward spectrometer with an acceptance of $10 - 300\,\mathrm{mrad}$ in the bending plane and $10 - 250\,\mathrm{mrad}$ vertical to the bending plane. An overview of the detector is given in Figure 2.3. The different subdetectors are, starting at the left closest to the interaction point

- The Vertex Locator (Velo), LHCb's silicon vertex tracker, which is built around the pp interaction point. The Velo resolves the primary interaction vertex and decay vertices of longlived particles with high precision.

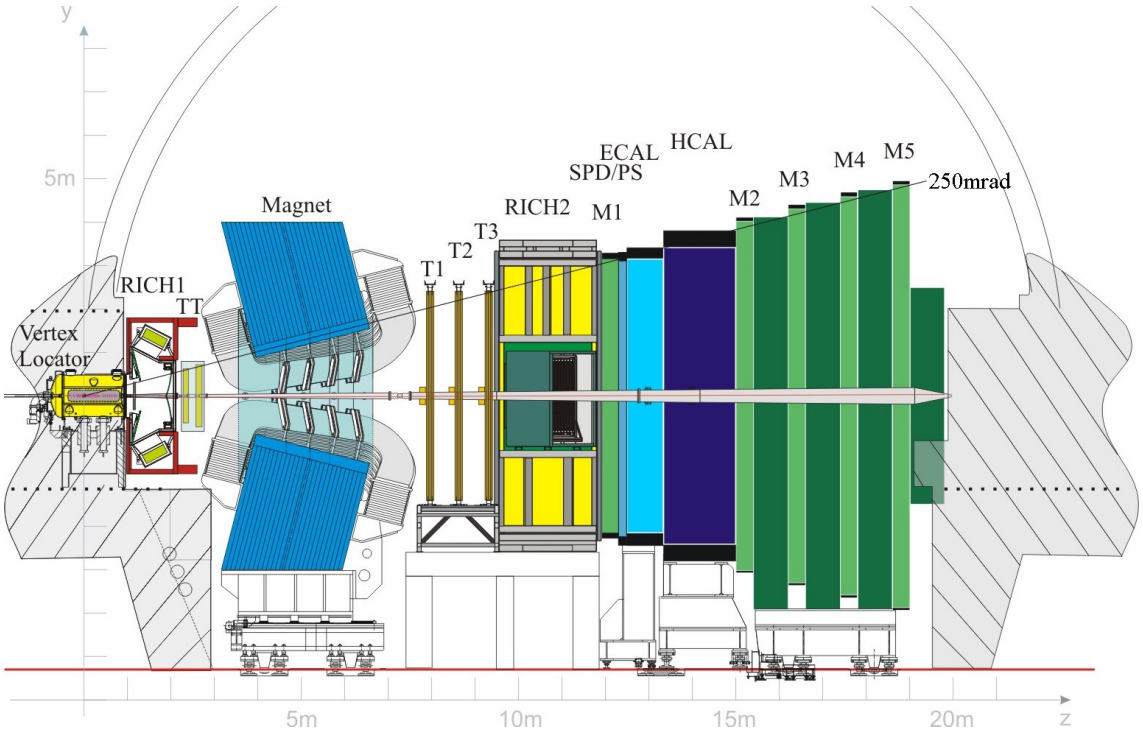


Figure 2.3.: A side view of the LHCb detector, Figure from [50]. The pp collisions occur to the very left in LHCb's silicon vertex tracker, the Vertex Locator (Velo). Downstream of the Velo the first ring imaging Cherenkov (RICH) detector and another tracking station (TT) are placed in front of the magnet. The main tracking stations, consisting of Inner and Outer Tracker are located behind the magnet. Finally particles traverse the second RICH, RICH2, the calorimetry (ECAL+HCAL) and the muon chambers.

- A first ring imaging Cherenkov detector (RICH) for particle identification follows behind the Velo.
- The Tracker Turicensis (TT), a first tracking station, is located in front of the magnet.
- The main tracker consists of three tracking stations located behind the magnet. The Outer Tracker (OT) is implemented as a straw tube gas detector and covers the largest fraction of the acceptance. The Inner Tracker (IT) covers the high occupancy region close to the beam pipe and uses silicon strip sensors.
- The RICH2 detector, again a ring imaging Cherenkov detector, is located behind the main tracker and used for particle identification.
- The calorimetry consists of two subsystems, the electromagnetic (ECAL) and hadronic (HCAL) calorimeter. The calorimeters measure the energy deposited by electromagnetic and hadronic showers respectively. In addition they constitute an important part of LHCb's trigger system (see section 2.3).
- The rightmost subsystem are the muon chambers called M1-M5 in Figure 2.3. The muon chambers are used for muon identification and essential for the muon trigger system.

All subsystems will be briefly reviewed in the next sections. A more detailed overview of the LHCb detector is given in [50, 51].

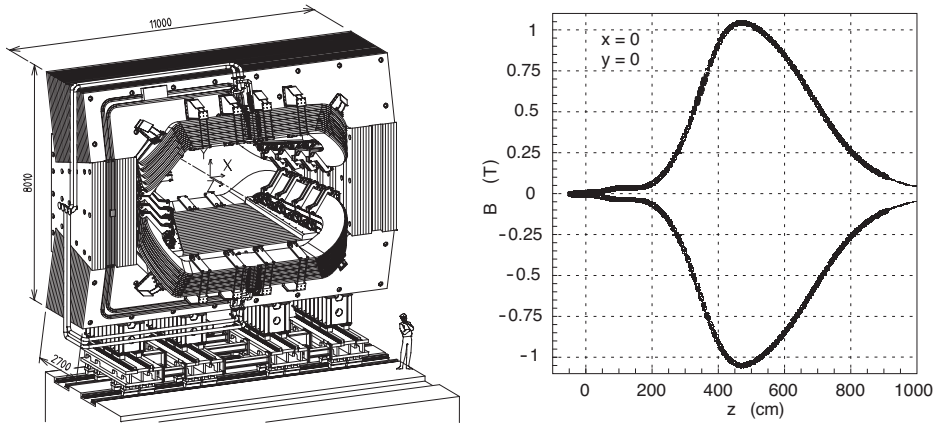
2.2.1. Spectrometer and tracking system

The Magnet

The spectrometer uses the fact that charged particles are bent by magnetic fields to determine their momenta. In the LHCb experiment the magnetic field is provided by the LHCb dipole magnet. The LHCb dipole is a warm magnet design. The geometry of the iron yoke and magnet coils is given in Figure 2.4a. The integrated field over 10 m is $\int B \mathrm{d}\ell = 4\,\mathrm{Tm}$, the y component of the magnetic field depending on the z coordinate is shown in Figure 2.4b. The particles are predominantly bent in the x-z plane, the components of the magnetic magnetic field orthogonal to the y axis are small. The total weight of the magnet is about 1600 t.

The Vertex Locator

The LHCb Vertex Locator (Velo) is the detector closest to the interaction point. Its purpose is to precisely measure tracks originating from the interaction vertex and possible secondary vertices originating from weak b or c quark decays. It comprises 21 silicon modules both left and right alongside the beam axis, as shown in Figure 2.5. Each of the modules is equipped with two sensors located on opposites sides of each module, one providing a measurement in radial direction r and one an azimuthal measurement



(a) (b)

Figure 2.4.: (a) The LHCb magnet and (b) the y component of the magnetic field of the dipole. The direction of the magnetic field can be inverted. Figures from [50].

in φ direction. In addition 2 pileup-stations each consisting of two r sensors are located upstream which can be used in the trigger. Both r and φ sensors are 300 μm thick with variable pitch. The r sensors are concentric semi-circles with a minimal pitch of 40 μm closest to the beam where the occupancy is highest which gradually increases to 102 μm at the outer edge of the sensor. The φ sensors are subdivided in two regions, an inner region and outer region as illustrated in Figure 2.6. Their pitch varies from 38 μm to 97 μm. The average occupancy expected at nominal luminosity is about 1%.

Since the precision of the determination of primary and secondary vertices is influenced by extrapolating the measured tracks the Velo modules are placed as close to the beam as possible. The Velo is separated from the beam vacuum only by a thin aluminum foil which keeps the amount of material particles have to traverse before passing through the sensors small. During data taking this foil is only 5 mm away from the beam. The distance of the active sensors of the Velo modules to the beam is 8 mm. Since the aperture of the beam during ramp up of the beam energy can be larger than this distance, the Velo is able to move out into the shadow of the beam pipe while the LHC is filled and the beam energy is ramped up. Only when the LHC signals stable running conditions the Velo is moved into its nominal position. Figure 2.6 shows that the Velo sensors from the two sides overlap when the Velo is closed.

The Tracker Turicensis

The Tracker Turicensis (TT) is a silicon microstrip detector covering the full acceptance of the experiment located just before the magnet. It consists of four layers in an x-u-v-x arrangement with vertical x layers and stereo layers u and v which are rotated by $-5°$ and $+5°$ with respect to the vertical axis. The first two layers, x-u are separated from the

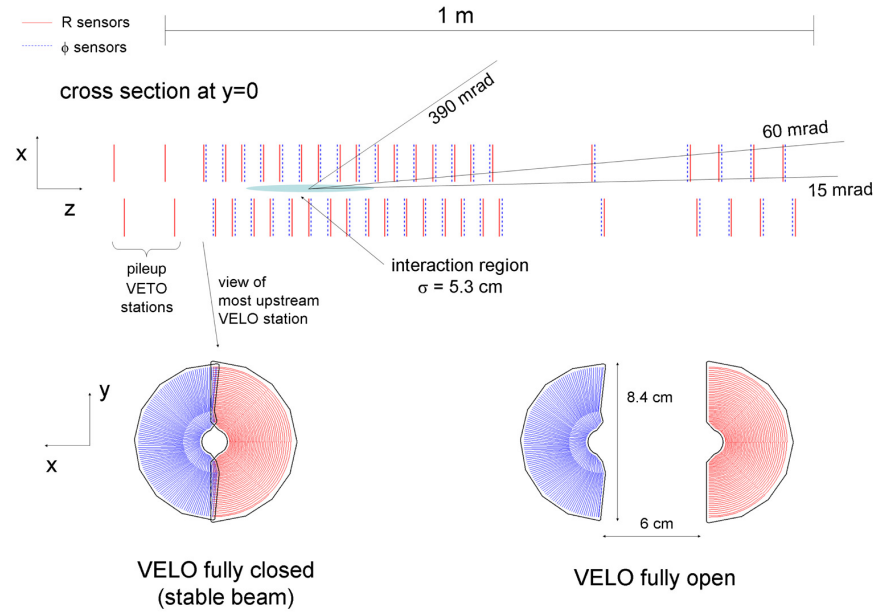


Figure 2.5.: The vertex locator consists of 21 modules left and right alongside the beam axis. It is used for the determination of primary interaction vertices and possible secondary vertices originating from weak decays of b or c quarks. Figure from [50].

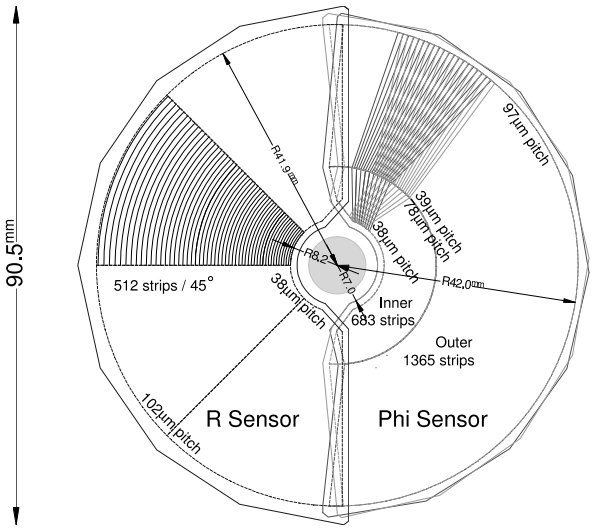


Figure 2.6.: Every Velo module is equipped with two sensors, one provides measurements in r, the other in φ direction. When closed both sides of the Velo overlap. Figure from [50].

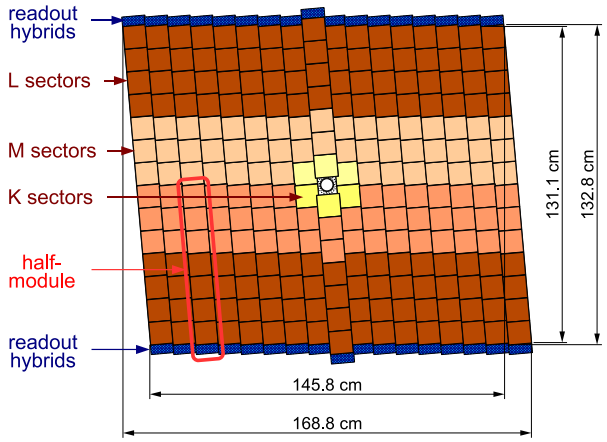


Figure 2.7.: Layout of the u layer of the TT. The u layer is rotated by $-5°$ with respect to the vertical axis. Figure from [50].

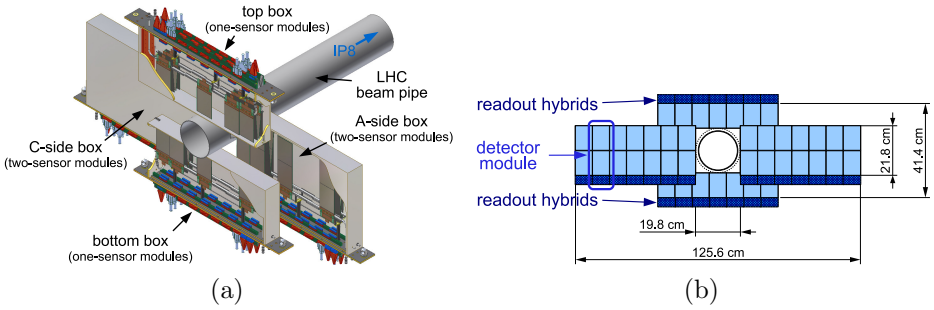
second two layers y-x by 27 cm. Since the charged particle density expected in regions close to beampipe (up to $5 \times 10^{-2}\,\mathrm{cm}^{-2}$) is much higher than at the outer edges of the detector (about $5 \times 10^{-4}\,\mathrm{cm}^{-2}$) the detector is segmented in different readout sections. Figure 2.7 shows the layout of a detector layer. A TT layer is made up of half-modules which are themselves assembled from seven silicon sensors. The silicon sensors are 9.44 cm long, 9.46 cm wide and have a pitch of 183 μm. The readout electronics is located at the bottom and top of the detector outside the LHCb acceptance to minimize multiple scattering. The single hit resolution of the TT is about 50 μm, the strip occupancies do not exceed a few percent.

The Inner Tracker

Further downstream, behind the magnet, the Inner Tracker (IT) is covering the region close to the beampipe. Similar to the TT, the IT is a silicon strip detector. The IT also has arranged its four layers per station in x-u-v-x fashion with the stereo layers u and v rotated by $\mp 5°$. It consists of four boxes arranged around the beampipe in three stations. One station of the IT is shown in Figure 2.8a. Figure 2.8b shows the organization of the silicon sensors inside the detector boxes. The silicon sensors are 11 cm long and 7.6 cm wide with a strip pitch of 198 μm. In total the IT has an active area of 4.0 m^2 with 129k readout channels. Despite the high charged particle densities of up to $1.5 \cdot 10^{-2}\,\mathrm{cm}^{-2}$ the hit occupancy is expected to be in the order of a few percent. Like the TT, the IT has a single hit resolution of about 50 μm.

The Outer Tracker

The Outer Tracker (OT) is a straw tube detector which consists of three stations located behind the magnet. It covers the large outer area of the LHCb acceptance with an active



(a) (b)

Figure 2.8.: (a) The inner tracker comprises three tracking stations each consisting of four boxes arranged around the beam pipe. (b) Shows one x layer of the IT. Figures from [50].

area of 6 m × 5 m. Every OT station consists of four layers in x-u-v-x configuration, where the u and v layers are rotated with respect to the x layers by $\mp 5°$. The layers are split in two parts left and right to the beampipe and supported by aluminum support structures, the so called C-frames. Two half-layers are carried by one C-frame which can be independently moved on rails to facilitate access to the OT modules and the frontend electronics. Figure 2.9 shows the mechanical support structure for the C-frames, the bridge, with two C-frames, one to the left and one to the right of the beampipe in the opened position. In total twelve C-frames hold the OT detector modules. Two types of modules exist, full (F) and short (S) modules. The F modules are 4850 mm long and split into an upper and a lower half. The S modules are used above and below the beampipe and have about half the length of an F module. An OT half layer consists of seven F and four S modules. In total the OT comprises 168 full and 96 short modules corresponding to about 55000 active channels.

Both types of modules consist of two staggered layers of 64 strawtubes each with inner diameters of 4.9 mm. A cross-section of a module is given in Figure 2.10. The straws are wound from two foils, the inner foil made of 40 μm Kapton-XC[36], the outer foil using 37.5 μm Kapton-aluminum. As anode wire a 25.4 μm thick gold-plated tungsten wire is used. One of the central goals of the design of the OT was to keep the material budget as low as possible to minimize multiple scattering effects. The material of the OT in the LHCb acceptance, consisting of 12 layers of modules, is equivalent to only 9.6% of one radiation length.

As counting gas a mixture of Argon/CO_2 with a ratio of 70/30% was chosen. Figure 2.11 shows the results of a study performed at a 6 GeV electron test beam at the DESY II facility in 2005. The straws show efficiencies larger than 99% near the center of the straw with dropping efficiencies at the edges. The R-t relation shown in Figure 2.11 is used to determine the distance of the throughgoing particle from the wire from the drift time of the ionization clusters. For single cells a position resolution of less than 200 μm is reached for nominal operation voltages of 1550 V.

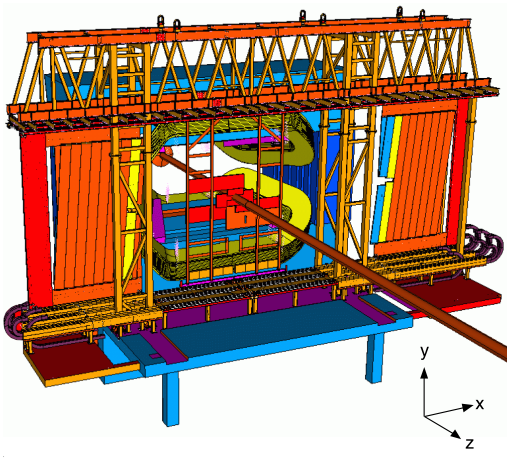
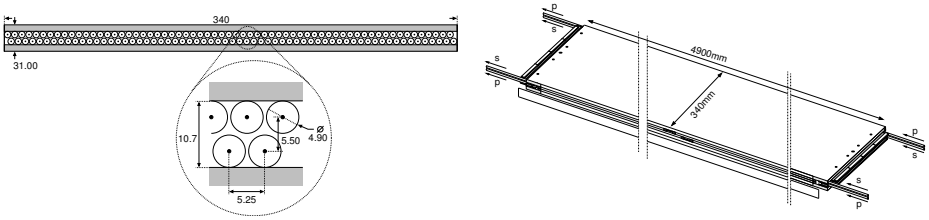
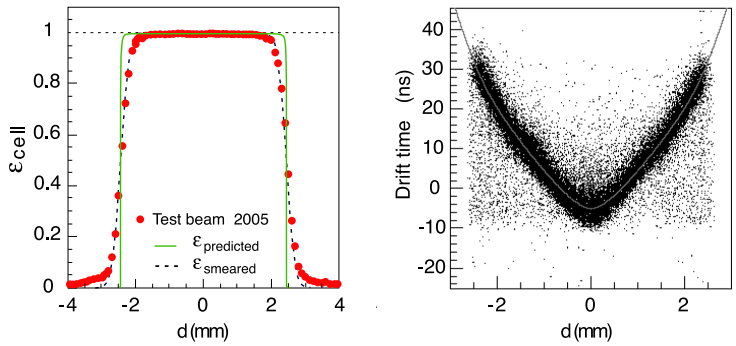


Figure 2.9.: The bridge of the Outer Tracker located behind the magnet. Two Outer Tracker C-frames and one Inner Tracker station are shown. The two C-frames are in the retracted position. Figure from [50].



(a) Cross section of an OT module (b) Full module

Figure 2.10.: Figure 2.10a shows a cross section through an OT module. A module consists of two staggered layers of 64 straw tubes each. Figure 2.10b shows a full module.



(a) Efficiency over the straw (b) R-t relation

Figure 2.11.: (a) Efficiency for a track to be detected when it is passing through a strawtube depending on the distance to the wire. (b) R-t relation which is used to derive the distance of the track from the wire using the measured drift time. Figures from [50].

Outer Tracker frontend electronics

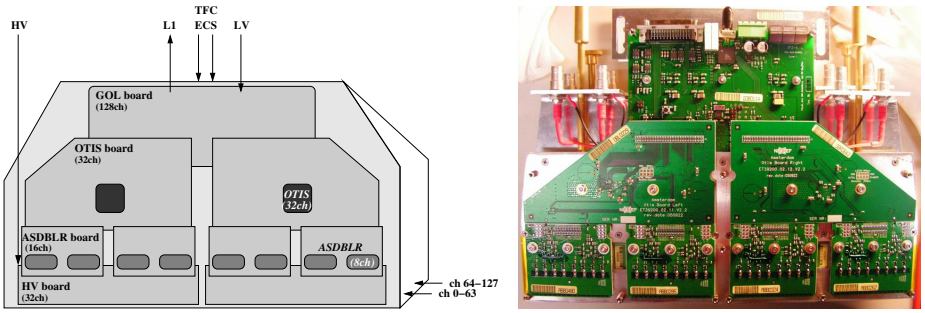
During my thesis I was involved in the maintenance and improvement of the Outer Tracker detector control software which is responsible for the configuration and monitoring of the OT electronics. Therefore the readout of the OT modules will be briefly summarized below.

The frontend electronics [52] for the readout of the OT modules are located on the top and bottom of the modules. Their design is shown in Figure 2.12. The signal is first processed by the ASDBLR chips (Amplifier, Shaper, Discriminator with Base Line Restoration) which provide amplification and perform a comparison with a variable threshold. Two ASDBLR boards equipped each with two ASDBLR Asics are connected to the OTIS board housing the OTIS Asic. The OTIS is a TDC (Time to Digital Converter) for digitization of the drift times. After digitization the drift times are stored in the OTIS pipeline and, on a positive trigger decision, they are written to the derandomizer buffer. Four OTIS boards are connected to one GOL/AUX board which houses the GOL (Gigabit Optical Link) chip responsible for serialization of the data and transfer to the TELL1 level 1 buffer board via optical link.

18 frontends in the upper and lower half of a single C-frame are connected to one control box which distributes the fast and slow control signals to the frontends. The fast control signals (trigger, bunch clock) are derived from the TFC (Timing and Fast Control) fiber provided from the TFC system. The slow control (configuration, monitoring) is realized via the SPECS [53] protocol.

Track reconstruction

The task of the track reconstruction is to determine tracks and their momenta from measurements (hits) of charged particles traversing the different tracking subsystems. It



(a) Frontend schematics (b) Opened frontend box

Figure 2.12.: Design schematics of (a) an outer tracker frontend and (b) photograph of an opened frontend box. Figures from [50].

is convenient to define different track types depending on the subdetectors contributing to their reconstruction

- **Velo tracks** consist of only hits in the Velo modules.
- **T-tracks** are track segments consisting of hits only in the tracking stations after the magnet.
- **Downstream tracks** are tracks built from hits in the TT and the tracking stations after the magnet. They are only useful in rare circumstances, one example being the reconstruction of longlived K_S^0 mesons which decay into $\pi^+\pi^-$.
- **Upstream tracks** with hits only in the Velo and the TT. They are mostly low momentum tracks bent out of the acceptance of the OT by the magnet.
- **Long tracks** are the tracks most commonly used by the physics analyses. They consist of hits in the Velo and the tracking stations and give the best momentum resolution due to the long lever arm.

To find hits belonging to a track pattern recognition algorithms are applied. These algorithms reconstruct track segments in the Velo and the T-Stations. The resulting Velo and T-tracks are then used to find long tracks. For this purpose two track finding algorithms are used, forward tracking and track matching.

- **Forward tracking** starts with a reconstructed Velo segment and a hit in the tracking stations. The algorithm then searches for additional hits in a windows of interest around the extrapolated track. If the resulting track candidates satisfy certain quality criteria they are chosen as long tracks. Finally hits in the TT are picked up if they are close to the track.
- **Track matching** tries to match a Velo segment with a T-track by extrapolating both to the bending plane inside the magnet to see whether they actually belong together. TT hits are again added afterwards if they are close to the track.

Two main figures describe the performance of the track finding, the track finding efficiency and the ghost rate. Ghost tracks are defined as tracks where less than 70% of the hits originate from a real particle. Typical numbers for long tracks with momenta larger than 10 GeV are track finding efficiencies of 94% with ghost rates of about 9% [50].

The measurements in the different subdetectors are fitted using a Kalman filter to gain a momentum estimate for the track. The quality of the track is given by the track fit χ^2_{track} divided by the number of degrees of freedom. The momentum resolution reached by the tracking system depends on the particle momentum. For long tracks resolutions of

$$\frac{\delta p}{p} = 0.35\% \text{ to } 0.55\%$$

are found for track momenta from 10 GeV to about 130 GeV [50].

2.2.2. Particle Identification

The LHCb experiment uses ring imaging Cherenkov counters (RICH) to distinguish between different charged particle species. When charged particles traverse material with refraction index n with a higher velocity than the velocity of light in the material ($c' = c/n$) they emit photons under the Cherenkov angle $\theta_{\text{Cherenkov}}$ with

$$\cos\theta_{\text{Cherenkov}} = \frac{1}{n\beta}$$

where $\beta = v/c$ is the velocity of the charged particle. Together with the momentum measurement of the particle performed by the tracking system this allows to calculate the particle's mass and therefore to determine the particle type. Figure 2.13 shows the Cherenkov angle depending on particle momentum for different particle species and different radiators.

To guarantee a good K-π separation over the full momentum range from 2 to 100 GeV LHCb uses three different radiators in two Cherenkov detectors. One of the two detectors, the RICH 1 (Figure 2.14a), is located before the magnet. It uses two radiators, silica aerogel ($n = 1.03$) and C_4F_{10} ($n = 1.0014$), and is responsible for particle identification at low momenta from about 1 to 60 GeV. The emitted Cherenkov photons are reflected by mirrors onto Hybrid Photo Detectors (HPDs) which detect photons in the wavelength range of 200-600 nm and measure their spatial position.

The second RICH detector, the RICH2 (Figure 2.14b), is located behind the magnet and tracking stations and is responsible for the identification of particles with momenta from 15 GeV up to 100 GeV and beyond. It contains CF_4 ($n = 1.0005$) as a gas radiator. Whereas the RICH 1 detector covers the full LHCb acceptance of $25 - 300$ mrad ($25 - 250$ mrad) in the horizontal (vertical) plane, RICH2 covers only an angular acceptance of $15 - 120$ mrad ($15 - 100$ mrad) horizontally (vertically).

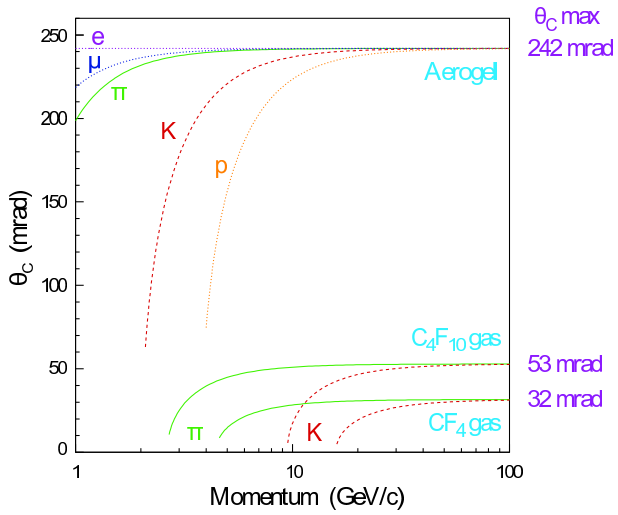


Figure 2.13.: Cherenkov angle versus particle momentum for different particle species and the three different radiators used in the RICH detectors. Figure from [50].

Charged hadron identification

A likelihood approach is used to identify different types of particles from the measured Cherenkov photons. For every track the Cherenkov angle for all the possible particle species is calculated (π, K, p, etc.) using the known track momentum. A likelihood is created by calculating the probability for every detected Cherenkov photon to have been emitted under the reconstructed angle given the specific track and the particle hypothesis. Of particular interest for many analyses is the difference between the logarithm of the likelihoods for different particle species, an example would be the variable

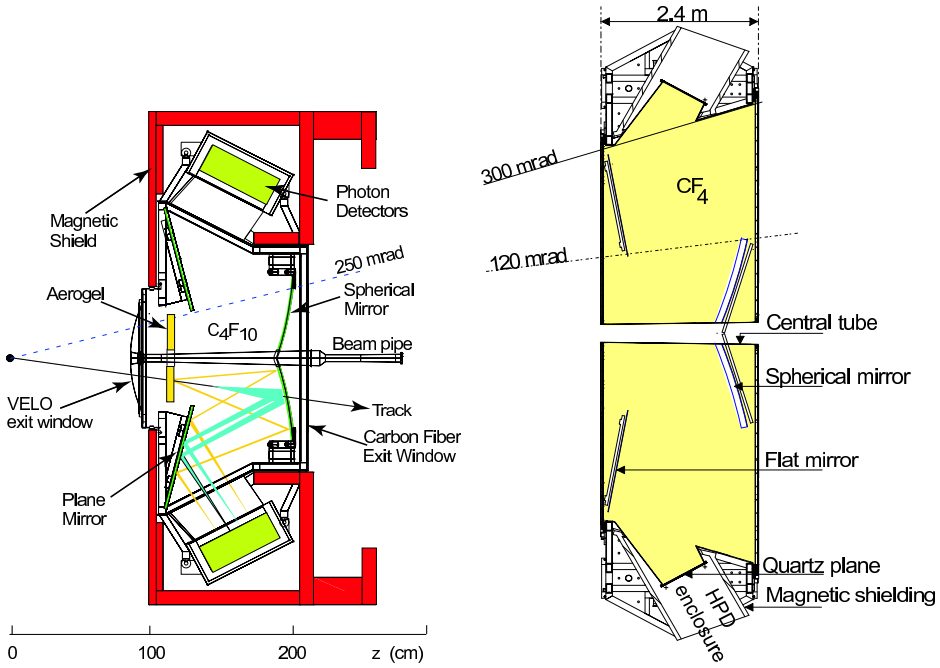
$$\Delta \ln \mathcal{L}_{K\pi} \quad = \quad \ln \mathcal{L}_K - \ln \mathcal{L}_\pi$$

which is used to separate pions from kaons. This variable is important for the study of final states with charged kaons, where, without particle identification, one would be overwhelmed by the pion background.

The RICH detectors provide an excellent K-π separation over the full momentum range of 2-100 GeV with an average kaon identification efficiency of 95% and a π-as-K misidentification rate of 5% [50].

2.2.3. Calorimetry

The LHCb calorimeter system measures the size and the position of energy depositions and performs electron, photon and hadron identification. Furthermore the calorimetry is used in the first level (L0) trigger. This puts stringent requirements on the readout speed of the calorimeters, since the L0 decision needs to occur after only 4 μs.

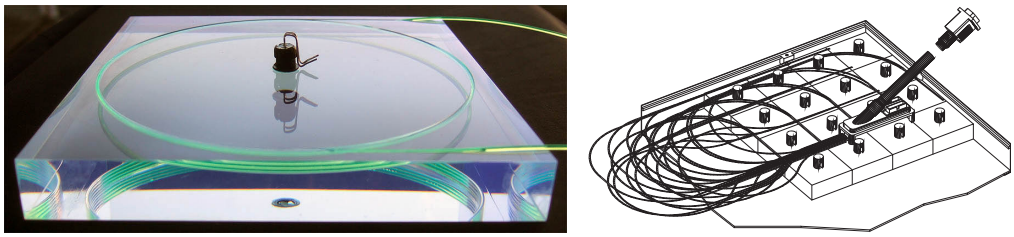


(a) RICH 1 (b) RICH 2

Figure 2.14.: Schematic view of (a) RICH 1 and (b) RICH 2. Figures from [50].

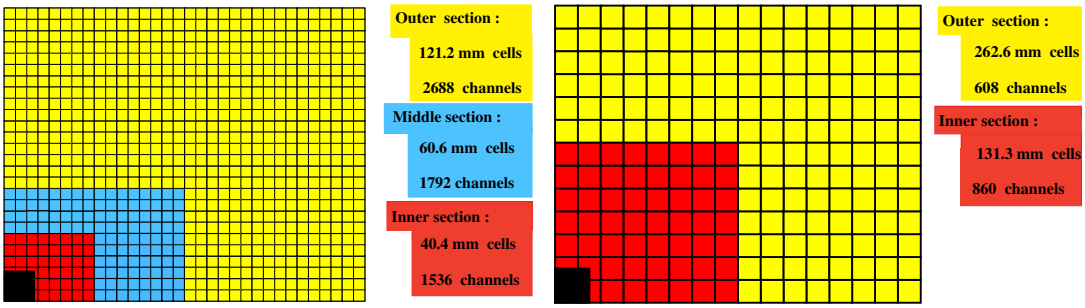
Charged incident particles and photons can deposit energy in the electromagnetic calorimeter by producing electromagnetic showers via bremsstrahlung and pair production. Hadrons can produce hadronic showers in the hadronic calorimeter. The particles of the shower produced in the absorber material traverse active scintillating material which they are able to excite. Photons emitted during the de-excitation process can then be read out via wavelength shifting fibers and detected with photomultiplier tubes (PMTs). LHCb's calorimeter system consists of the following subsystems:

- The Scintillating Pad Detector (SPD) used for the separation of electrons and photons.
- The Preshower (PS) follows the SPD after 15 mm of lead absorber. It is used to distinguish between electrons and charged pions.
- The electromagnetic calorimeter (ECAL) which measures the energy and position of showers from electrons and photons.
- The hadronic calorimeter (HCAL) which measures the energy and shower shape of hadrons.



(a) Scintillating pad (b) SPD/PS module box

Figure 2.15.: (a) Scintillating pad and (b) fiber routing in an inner SPD/PS module box. Figures from [50].



(a) SPD/PS and ECAL sections (b) HCAL sections

Figure 2.16.: Different detector sections of the (a) SPD/PS, ECAL and (b) HCAL. The corresponding segmentation can be found in Table 2.1. Figures from [50].

Scintillating pad detector and preshower detectors

SPD and PS are located behind the first muon station M1 on the two opposite sides of a 15 mm thick lead absorber. Both detectors are constructed out of scintillating pads as shown in Figure 2.15. The detectors are subdivided in three different regions with different segmentation. Figure 2.16a shows the three different regions, the segmentation decreases from 4 cm × 4 cm large pads in the inner region to 12 cm × 12 cm large pads in the outer region as given in Table 2.1.

The purpose of the SPD detector is to separate electrons from photons. The principle behind the electron-photon separation is that while electrons will give a signal in the SPD the electrically neutral photons will not. The photon as electron misidentification rate is found to be below 3%.

The PS detector is built for electron-pion separation. The deposited energy for 50 GeV electrons and pions is shown in Figure 2.17 in arbitrary units. In tests with e/π beams the PS showed pion rejection rates of 99.6%, 99.6%, and 99.7% with electron retention rates of 91%, 92% and 97% for particles with 10 GeV, 20 GeV and 50 GeV respectively.

detector	depth [cm]	interaction length X_0/λ_I [%]	segmentation [cm^2] inner section	middle section	outer section
SPD	18	2.5/0.1	4.04×4.04	6.06×6.06	12.12×12.12
PS			4.04×4.04	6.06×6.06	12.12×12.12
ECAL	42	25/1.1	4.04×4.04	6.06×6.06	12.12×12.12
HCAL	165	−/5.6	13.13×13.13	−	26.26×26.26

Table 2.1.: The segmentation of the subdetectors of the calorimeter system. SPD, PS and ECAL have identical segmentation.

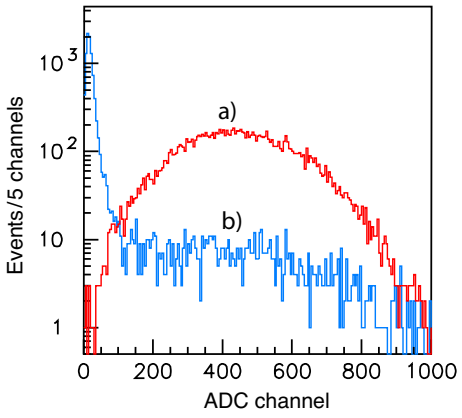
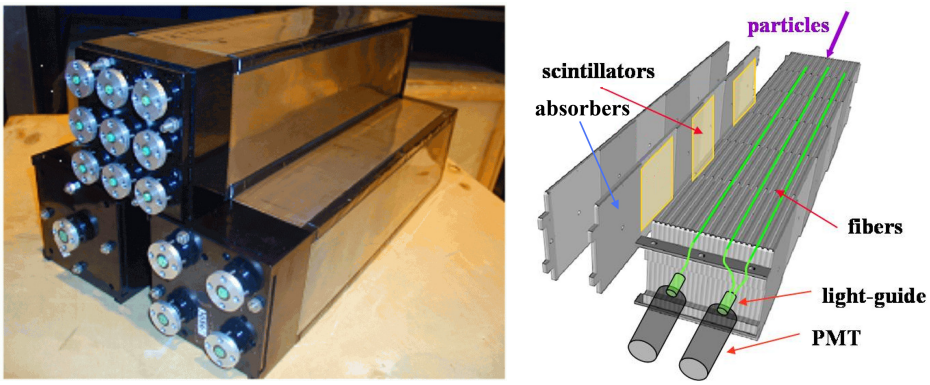


Figure 2.17.: Energy deposition (in arbitrary units) in the PS detector for (a) 50 GeV electrons and (b) 50 GeV pions. Figure from [50].



(a) ECAL modules (b) HCAL module

Figure 2.18.: (a) ECAL modules for the three different detector regions. (b) schematics of an HCAL module. Figures from [50].

Electromagnetic calorimeter

The electromagnetic calorimeter is located behind the PS and built as a so called "shashlik" calorimeter. It is a sampling structure built from alternating layers of 2 mm lead absorber and 4 mm scintillator material read out via wavelength shifting fibers. The segmentation of the ECAL is the same as the one used for the SPD/PS, see Table 2.1, to have a finer granularity in the regions close to the beam pipe where the particle density is high. Three different types of modules are produced for this purpose as shown in Figure 2.18a. The ECAL has a length corresponding to 25 radiation lengths X_0 and 1.1 hadronic interaction lengths λ_I. The energy resolution of the ECAL is given by

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E(\,\mathrm{GeV})}} \oplus 1\%,$$

where the resolution terms on the right hand side should be added in quadrature.

Hadronic calorimeter

The HCAL is located after the ECAL and uses an iron/scintillator sampling structure as shown in Figure 2.18b. The segmentation of the HCAL is coarser than for the ECAL with the cells having a size of about 13 cm × 13 cm in the inner and 26 cm × 26 cm in the outer region, see Table 2.1. In total the HCAL has a depth corresponding to 5.6 hadronic interaction lengths λ_I. The energy resolution of the HCAL is given by

$$\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E(\,\mathrm{GeV})}} \oplus 10\%.$$

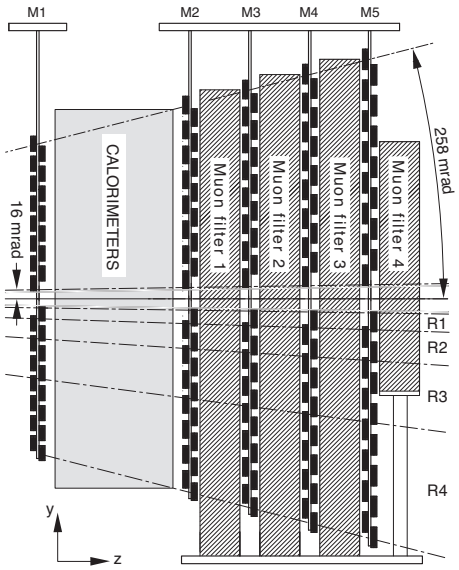


Figure 2.19.: Overview over the five muon stations. Figure from [50].

	M1	M2	M3	M4	M5
R1 pad size [cm^2]	1×2.5	0.63×3.1	0.67×3.4	2.9×3.6	3.1×3.9
R2 pad size [cm^2]	2×5	1.25×6.3	1.35×6.8	5.8×7.3	6.2×7.7
R3 pad size [cm^2]	4×10	2.5×12.5	2.7×13.5	11.6×14.5	12.4×15.5
R4 pad size [cm^2]	8×20	5×25	5.4×27	23.1×29	24.8×30.9

Table 2.2.: Sizes of the pads for the four regions in the muon stations M1-M5. Numbers are compiled from [50].

2.2.4. The Muon chambers

The muon chambers are essential for muon identification and triggering of B meson decays containing muons in the final state. The LHCb muon system consists of five muon stations (M1-M5), see Figure 2.19. M1 is located upstream of the calorimeters to improve the p_T resolution in the muon trigger by minimizing uncertainties caused by multiple scattering in the calorimeter material. 80 cm thick iron absorbers are placed between the muon stations M2-M5 located after the calorimeter to filter out all particles except muons. To traverse all muon chambers and interleaved absorbers muons need to have a minimum momentum of 6 GeV. Each of the muon stations is divided in four regions R1-R4 with finer segmentation in the regions with higher particle multiplicity close to the beam pipe. Figure 2.20 shows the segmentation of M1. The segmentation over the four regions R1-R2-R3-R4 scales as 1:2:4:8. The specific pad dimensions are given in Table 2.2.

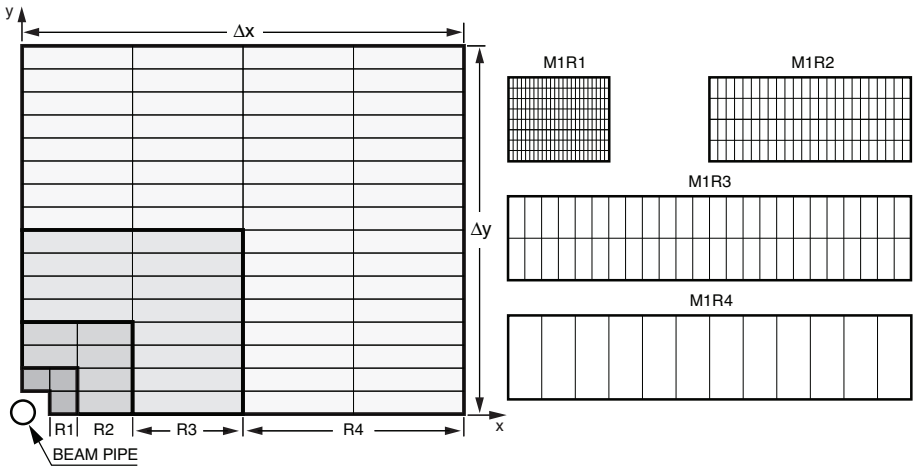


Figure 2.20.: Segmentation of the muon station M1. Figure from [50].

All regions use multiwire proportional chambers (MWPC) except region R1 in M1 which uses triple-GEM detectors since the MWPCs would not meet the requirements for radiation tolerance in this region of high particle flux. Both detector types are able to collect the signal in less than 20 ns with an efficiency larger than 95%.

Muon identification

A track which is considered as a muon candidate is extrapolated into the muon system. To be confirmed as muon candidate a certain number of hits in the muon chambers need to be found in a field of interest around the track. In addition a likelihood is created using the distance of the hits from the extrapolated track under both the muon and pion hypothesis. The difference between the logarithms of the likelihoods for the μ and π particle hypotheses,

$$\Delta \ln \mathcal{L}_{\mu\pi} \quad = \quad \ln \mathcal{L}_{\mu} - \ln \mathcal{L}_{\pi}$$

can be used to differentiate between the two particle types. LHCb's muon system exhibits a high muon identification efficiency of 95% and a π-as-μ misidentification rate of 3% [50].

2.3. The LHCb Trigger system

LHCb uses a multi-level trigger system to reduce the bunch crossing rate of 40 MHz to a data rate of about 2.2 kHz which is written to tape for later analysis. It therefore needs to reject events that are of no interest for the LHCb physics program but at the same time retain events containing decays of B and D mesons. The LHCb trigger system consists of three levels called Level-0 (L0), High Level Trigger 1 (HLT1) and High Level Trigger 2 (HLT2). The subsequent reduction of rates enables the trigger

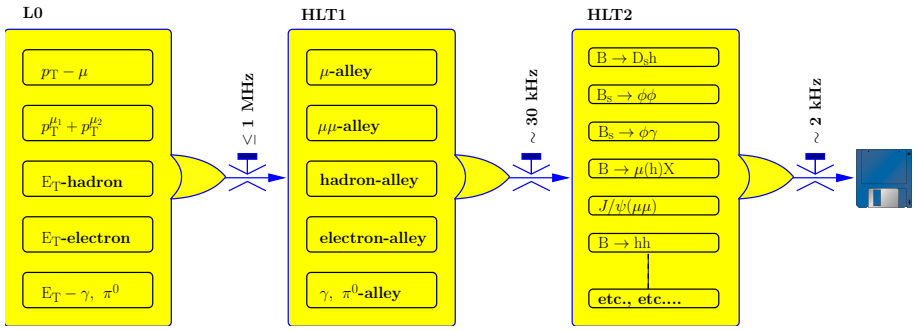


Figure 2.21.: The three trigger levels of the LHCb trigger system. Figure from [50].

stages to use more and more refined reconstruction mechanisms to decide whether an event is of interest or can be discarded. An overview of the three trigger levels is given in Figure 2.21. The trigger system is designed to be very configurable to be able to adapt to different running conditions. A unique Trigger Configuration Key (TCK) is assigned to identify every trigger configuration.

2.3.1. Level 0 Trigger

The first trigger level is called L0. The L0 trigger is completely implemented in hardware to be able to cope with the high interaction rate. It takes advantage of the somewhat higher transverse momenta of particles from B decays relative to QCD events which in general show a "softer" p_T distribution. The L0 trigger consists of three subsystems, the calorimeter trigger, the muon trigger and the pileup system.

- The calorimeter trigger sums up the transverse energy E_T of an array of 2×2 calorimeter cells to form calorimeter clusters. A particle hypothesis (e, γ or hadron) is assigned to the clusters using information from the SPD, PS, the ECAL and the HCAL. The trigger then selects the hadron, electron and photon cluster with the largest transverse energy in the event.

- The muon trigger tries to reconstruct, for each quadrant of the detector, the two muons with the largest transverse momentum p_T. Assuming the muon tracks originate from the interaction point, the p_T of the candidates can be estimated using the slope in the first two muon stations[3]. The muon standalone reconstruction exhibits a momentum resolution of $\sim 25\%$.

- The pileup system upstream of the Velo can be used to veto events with high backwards activity indicating multiple interactions[4].

The L0 decision unit combines the different trigger subsystems and gives the final trigger decision. The L0 trigger reduces the rate to 1 MHz with which the detectors are read

[3]The effect of the magnetic field is described by an effective bending plane.
[4]The pile-up system is currently not used in the trigger decision.

out. The buffer size of the readout electronics fixes the latency of the L0 trigger to 4 μs. Since the time of flight of particles and cable lengths, as well as various other delays need to be taken into account the L0 decision needs to be taken in only 2 μs.

2.3.2. High Level Trigger

The High Level Trigger is implemented in software (written in C++) and runs on a large computing cluster called the event filter farm. It further reduces the event rate from 1 MHz, the L0 output rate, to 2.2 kHz which is written to disk. The High Level Trigger consists of two stages, the HLT1 and the HLT2.

- The HLT1 performs a partial event reconstruction. It tries to confirm the L0 decision by reconstructing tracks in the Velo and the tracking stations corresponding to the candidates found by the L0 trigger. This L0 confirmation reduces the event rate to ~ 30 kHz.
- The HLT2 fully reconstructs tracks in the event in a manner very similar to the offline procedure. It performs various inclusive and exclusive selections. The inclusive selections try to partially reconstruct the final states of B decays, an example would be the reconstruction of b-hadron decays to (J/ψ + X). Exclusive selections are used for $B \to$ hadron decays where the final state is fully reconstructed. The HLT2 reduces the event rate to 2.2 kHz.

2.4. The LHCb software framework

Several software packages are needed to perform an analysis with the data taken by the LHCb detector.

- **Brunel**
  Data that was triggered and written to disk by the event filter farm is available as so called raw data files. The first step in the analysis is to reconstruct physical quantities from the measurements of the detector subsystems. The Brunel software package [54] performs this reconstruction step. Brunel performs track finding and fitting. It also links tracks to the available particle identification information extracted from the calorimeter, muon and RICH subdetectors. Furthermore Brunel reconstructs the energy of electromagnetic and hadronic showers using the calorimetry. At the end of the reconstruction step Data Summary Tape (DST) files are produced on which all further analyses is based.
- **DaVinci**
  DaVinci [55] is the analysis software package of LHCb built on the Gaudi [56] framework. DaVinci applies particle hypotheses to tracks and combines particles to reconstruct signal decays.

  To make the large amounts of data more manageable for physics analysis preselections are performed on the reconstructed data. The reconstruction of the signal

decays and the preselection is performed using the DaVinci analysis software. This analysis step is also called "stripping". Similar preselections are combined to form so called streams for different physics topics. Examples for streams provided by the stripping are the *hadron stream* which selects hadronic decays of B mesons, the *charm stream* which selects decays of D mesons and the *dimuon stream* which is optimized for B decays containing two muons in the final state.

The offline analysis is performed on the appropriate physics stream, again using the DaVinci analysis software. The full selection is applied using the signal candidates reconstructed in the stripping stage.

- **Dirac**
  Due to the large quantities of data that need to be processed at LHCb these tasks are performed on the Grid using the Dirac workload and data management system [57].

2.5. Generation of simulated events at LHCb

2.5.1. LHCb Monte Carlo Simulation Framework

To asses the feasibility of physics studies and to devise and test appropriate analysis strategies simulation of the physical processes occurring at LHCb is essential. The method of Monte Carlo (MC) simulation is used to model both the particles emerging from the proton-proton collision as well as their interaction with material. The full MC simulation is performed using the LHCb simulation framework which is based on the Gaudi framework and consists of several steps

- **Gauss**
  Monte Carlo simulated events are generated using the Gauss [58] software package. For the simulation of the proton-proton collision Gauss uses the Pythia [43] program, a general purpose MC event generator. Pythia contains routines for the hard processes (examples are the b production processes given in section 1.4) in the proton-proton collision as well as initial and final state radiation and subsequent hadronization. The decay of heavy B mesons is performed using the EvtGen [59] package which is specialized on the physics of B decays. In contrast to Pythia EvtGen does not use probabilities but the quantum mechanical amplitudes to correctly account for the numerous interference effects in decay and mixing of B mesons. The particles produced in the proton-proton collision finally decay to longlived particles which traverse the detector. The simulation of their electromagnetic and hadronic interaction with material is performed by the Geant [60] software.

- **Boole**
  The output of the detector simulation stage is used as input for Boole [61] which simulates the detector response to hits deposited in the detector by traversing particles.

- **Moore**
  MOORE [62] is the high level trigger application of LHCb. It can be run on Monte Carlo data in practically identical fashion as on the event filter farm, accurately replicating the behavior of the trigger on real data.

- **Brunel**
  The reconstruction of the events is performed the same way as for real data using the BRUNEL software package. For simulated data BRUNEL additionally performs the associations (matching) of reconstructed tracks and calorimeter clusters to the originally generated particles.

- **DaVinci**
  The analysis is performed using the physics analysis software, DAVINCI, analogous to the processing of real data

2.5.2. Simulated signal events

The Monte Carlo simulated signal events used for this thesis contain one $B_s^0 \to J/\psi\,[\mu^+\mu^-]\phi\,[K^+K^-]$ signal decay in every event. To achieve this, repeated hadronization of the generated $b\bar{b}$ pair is performed until a B_s^0 is found. The B_s^0 is then forced to decay into the signal channel. For the generation of the signal decay EVTGEN uses the time dependent decay amplitudes given in section 1.3. To save computational costs the event is only kept if the final state particles of the signal decay lie in the acceptance of the LHCb detector ($10\,\text{mrad} < \Theta < 400\,\text{mrad}$). The parameters used in the Monte Carlo generation are summarized in Table 2.3.

Besides the obvious use of the simulated signal events to test the extraction of the B_s^0 mixing phase and to optimize the selection criteria the simulated events are also used for the determination of possible acceptance effects due to detector geometry, reconstruction and selection.

Physics Parameter	Value
ϕ_s	-0.70 rad
Γ_{s}	$0.6793\,\mathrm{ps}^{-1}$
$\Delta\Gamma_{\mathrm{s}}$	$0.0600\,\mathrm{ps}^{-1}$
$\mathrm{m}_{\mathrm{B}_{\mathrm{s}}^0}$	5366.3 MeV
$\lvert A_0(0)\rvert^2$	0.60
$\lvert A_\parallel(0)\rvert^2$	0.24
$\lvert A_\perp(0)\rvert^2$	0.16
$\delta_\parallel$	2.50 rad
$\delta_\perp$	-0.17 rad
Δm_{s}	$17.8\,\mathrm{ps}^{-1}$

Table 2.3.: Parameters used in the generation of the Monte Carlo simulated signal events. The value for ϕ_s is about twenty times larger than what is expected in the Standard Model. This is the so called "New Physics" scenario. In total $2 \cdot 10^6$ simulated signal events are generated.

3. Reconstruction and selection of the decay $B_s^0 \to J/\psi\,\phi$

To perform an analysis of the decay $B_s^0 \to J/\psi\,\phi$ it is necessary to select a clean sample of signal events. The first selection stage is performed by the LHCb trigger system which enriches the B meson content of the data sample. In addition several offline selection criteria are applied to the data exploiting particle identification information, information on the reconstructed particle tracks and the kinematics of the signal decay. Both trigger and offline selection criteria are discussed in this chapter.

Furthermore the chapter outlines the reconstruction of the B_s^0 decay time and the transversity angles and the reconstruction effects that need to be accounted for when trying to determine ϕ_s. In addition the determination of the production flavor of the B_s^0 meson (flavor tagging) is introduced which is of central importance for the determination of ϕ_s.

3.1. Data sample

The data used in this analysis corresponds to an integrated luminosity of $36\,\text{pb}^{-1}$. It was taken with the LHCb detector in 2010 with the LHC running at a center of mass energy of $\sqrt{s} = 7\,\text{TeV}$. The data is reconstructed using the LHCb reconstruction software BRUNEL[1]. The analysis is performed within the analysis framework DAVINCI[2].

3.2. Triggering the signal decay

With two muons in the final state the muon trigger lines are used to select the signal decay $B_s^0 \to J/\psi\,[\mu^+\mu^-]\phi\,[K^+K^-]$. There are two L0 muon lines of interest for this analysis, the single-muon and the L0 dimuon lines. The single-muon L0 line searches for a single muon track with moderately high p_T whereas the dimuon line searches for two muon tracks with a combined $p_T = p_T(\mu_1) + p_T(\mu_2)$ larger than 1 GeV. Only candidates where the muons from the reconstructed B_s^0 signal decay triggered the event are used. These events are called TOS events (Triggered On Signal). The High Level Trigger for muons consists of two distinctive types of trigger lines, the *lifetime biased* and the *lifetime unbiased* muon lines. Both are used in the analysis and are discussed in detail below. A single trigger configuration[3] was used for the majority of the running period in 2010.

[1] BRUNEL version v37r8p4
[2] DAVINCI version v26r3
[3] Defined by the TCK `0x002E002A`

Lifetime unbiased trigger lines

There are two *lifetime unbiased* trigger lines in HLT1[4]. The first line requires a muon candidate with a well reconstructed track and a transverse momentum of $p_T > 1.8\,\text{GeV}$. The second line is able to recover muons with lower p_T by requiring the sum of the p_T of the muons to exceed 1 GeV. The invariant mass of the dimuon system is required to be larger than 2.5 GeV and the closest distance between two muon tracks[5] needs to be smaller than 0.5 mm.

In the second stage of the HLT (HLT2) a J/ψ meson is reconstructed from two muon tracks which are fit to a common vertex[6]. The HLT2 trigger line[7] requires this vertex fit to be of good quality. The χ^2 of the vertex fit χ^2_{Vtx} is required to be smaller than 25. It furthermore cuts on the invariant mass of the dimuon system. Finally the reconstructed J/ψ is required to have a minimum p_T exploiting the comparably large mass of the B_s^0 meson.

A full list of the criteria used in the *lifetime unbiased* HLT1 and HLT2 trigger lines is given in Table 3.1. The events passing the *lifetime unbiased* trigger lines are referred to as the *lifetime unbiased* data sample.

Lifetime biased trigger lines

The *lifetime biased* triggers used in the analysis consists of two trigger lines[8]. A full list of the criteria used in the *lifetime biased* trigger lines is given in Table 3.2. The main difference compared to the *lifetime unbiased* trigger lines is the requirement of a minimum Impact Parameter (IP) of the muon tracks. The IP is defined as the minimum distance of the track to the nearest primary vertex. By requiring IP > 0.1 mm the *lifetime biased* trigger lines exploit the large lifetimes of the B mesons. The same HLT2 selection as for the *unbiased* trigger lines is used if events pass the *lifetime biased* trigger lines in HLT1. The events passing the *lifetime biased* trigger lines are referred to as the *lifetime biased* data sample.

3.3. Selection of the signal events

The main objectives of the event selection are as follows:

[4] Hlt1SingleMuonNoIP and Hlt1DiMuonNoIPL0

[5] Distance Of Closest Approach, or DOCA, of the two tracks

[6] The vertex fit determines the position of the common decay vertex using the method of least squares fitting. As input of the vertex fit the estimates of the track parameters and the associated covariance matrices are provided by the tracking algorithms. The vertex fit minimizes χ^2_{Vtx} which is given by $\chi^2_{\text{Vtx}} = \sum_i (\vec{m}_i - \vec{m}(\vec{x})_i)^T V_i^{-1} (\vec{m}_i - \vec{m}(\vec{x})_i)$. In this expression $\vec{m}_i$ denotes the measured track parameters and V_i the covariance matrix for track i. $\vec{m}(\vec{x})_i$ gives the fit model for track i depending on the common decay vertex $\vec{x}$.

[7] The Hlt2UnbiasedJPsi HLT2 line is used in this analysis.

[8] Hlt1TrackAllL0 and Hlt1TrackMuon

[9] The IsMuon criterion requires for muons with $3\,\text{GeV} < \text{p}_\mu < 6\,\text{GeV}$ hits in M2+M3, for $6\,\text{GeV} < \text{p}_\mu < 10\,\text{GeV}$ hits in M2+M3+(M4 or M5) and for $10\,\text{GeV} < \text{p}_\mu$ hits in M2+M3+M4+M5.

	Hlt1SingleMuonNoIP	Hlt1DiMuonNoIPL0	Hlt2UnbiasedJPsi
L0	L0-Muon	L0-DiMuon	-
$p_T(\mu)$	$> 1.8\,\text{GeV}$	-	$> 500\,\text{MeV}$
$p(\mu)$	$> 10\,\text{GeV}$	$> 10\,\text{GeV}$	-
$\chi^2_{\text{Track}}/\text{ndof}(\mu)$	< 10	< 10	-
$p_T(\mu_1) + p_T(\mu_2)$	-	$> 1\,\text{GeV}$	-
Dimuon DOCA	-	$< 0.5\,\text{mm}$	-
Dimuon mass	-	$> 2.5\,\text{GeV}$	$> 2977\,\text{MeV}$ $< 3211\,\text{MeV}$
Dimuon χ^2_{Vtx}	-	-	< 25

Table 3.1.: HLT selection criteria for trigger configuration key (TCK) `0x002E002A` which was used for the majority of the data taking period in 2010. Shown are the *lifetime unbiased* trigger lines used in the analysis. Table from [63].

	Hlt1TrackAllL0	Hlt1TrackMuon
L0	L0-Muon	L0-Muon OR L0-DiMuon
IP	$> 0.11\,\text{mm}$	$> 0.11\,\text{mm}$
# velo hits	> 9	> 9
# missed velo hits	< 3	< 3
$p_T(\mu)$	$> 1.85\,\text{GeV}$	$> 0.8\,\text{GeV}$
$p(\mu)$	$> 13.3\,\text{GeV}$	$> 8\,\text{GeV}$
IP χ^2	> 34	> 25
IsMuon[9]	-	True

Table 3.2.: HLT selection criteria for trigger configuration key (TCK) `0x002E002A` which was used for the majority of the data taking period in 2010. Shown are the *lifetime biased* trigger lines used in the analysis. The impact parameter and impact parameter χ^2 of the muon track is determined with respect to the nearest primary vertex. Table from [63].

$J/\psi \to \mu\mu$ selection		
Criterion	Stripping	Offline selection
$\Delta \ln \mathcal{L}_{\mu\pi}$	> 0	> 0
$\chi^2_{\text{Track}}/\text{ndof}(\mu)$	< 5	< 4
$\min(p_T(\mu^+), p_T(\mu^-))$	-	$> 0.5\,\text{GeV}$
$\chi^2_{\text{Vtx}}(J/\psi)$	< 16	< 11
$\lvert M(\mu^+\mu^-) - M^{\text{PDG}}_{\text{J}/\psi}\rvert$	$< 80\,\text{MeV}$	$< 80\,\text{MeV}$
$\lvert M(\mu^+\mu^-) - M^{\text{PDG}}_{\text{J}/\psi}\rvert/\sigma_{M(\mu^+\mu^-)}$	-	1.4×3

Table 3.3.: Common J/ψ selection criteria. Given are the values used in the stripping as well as in the offline selection. Table from [63].

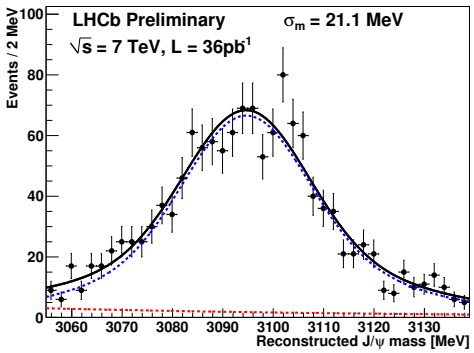
- Maximize the signal event yield.
- Minimize the contamination by background processes mimicking the signal decay.
- Minimize kinematic distortions due to the event selection. The method to determine the physics parameters needs to correctly account for these acceptance effects. This issue is discussed in more detail later in sections 3.6.2 and 4.5.7.

The requirements of both the preselection in the stripping stage as well as the offline selection are given in Tables 3.3 and 3.4 . The selection was developed in [63] together with the selections of other $B \to \text{J}/\psi\,\text{X}$ decays.

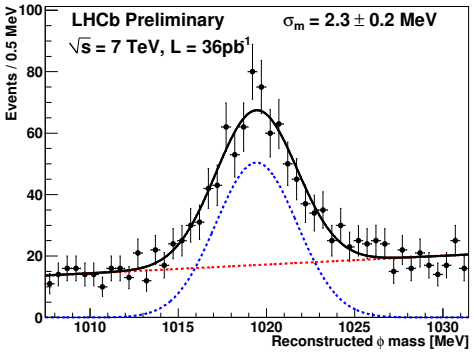
J/ψ reconstruction

Information from the muon stations is used to select well reconstructed muons ($\Delta\mathcal{L}_{\mu\pi} > 0$) with an associated well fitted track ($\chi^2_{\text{Track}}/\text{ndof} < 4$) and a minimum transverse momentum of $p_T > 0.5\,\text{GeV}$. Two muon tracks are combined to form a J/ψ meson. As the two muons are supposed to originate from the same vertex a vertex fit is performed and the vertex χ^2_{Vtx} is determined. If the muons did indeed originate from a common point the quantity χ^2_{Vtx} should be distributed according to a χ^2 distribution. Only events for which $\chi^2_{\text{Vtx}}/\text{ndof}$ is smaller than 11 are accepted. The invariant mass of the dimuon system $M(\mu^+\mu^-)$ is required to be close to the J/ψ mass given by $M^{\text{PDG}}_{\text{J}/\psi}$. An additional criterion takes the uncertainty of the reconstructed J/ψ mass, $\sigma_{M(\mu^+\mu^-)}$, into account. This uncertainty is determined by the resolution of the momentum estimates for the tracks. J/ψ candidates are only accepted if the reconstructed J/ψ mass is no further than $1.4 \cdot 3 \times \sigma_{M(\mu^+\mu^-)}$[10] away from $M^{\text{PDG}}_{\text{J}/\psi}$. Figure 3.1a gives the distribution of the reconstructed J/ψ mass.

[10]The factor 1.4 accounts for an underestimation of the uncertainty of the J/ψ mass reconstruction.



(a) Reconstructed J/ψ mass



(b) Reconstructed ϕ mass

Figure 3.1.: Distribution of (a) the reconstructed J/ψ mass and (b) the reconstructed ϕ mass of $B_s^0 \to J/\psi\,\phi$ candidates triggered by the *lifetime unbiased* trigger lines after the full selection.

ϕ reconstruction

For the reconstruction of the $\phi \to K^+K^-$ decay particle identification information provided by the RICH detectors is used. Two tracks from particles of opposite charge which are well reconstructed ($\chi^2_{\text{track}}/\text{ndof} < 4$) and identified as kaons ($\Delta\mathcal{L}_{K\pi} > 0$) are fit to a common vertex to form a ϕ meson. The reconstructed meson is required to have a transverse momentum of $p_T > 1\,\text{GeV}$ again utilizing the large mass of the B_s^0 meson. Subsequently a cut on the invariant mass of the reconstructed K^+K^- system and the quality of the vertex fit is applied. Figure 3.1b gives the reconstructed ϕ mass.

B_s^0 reconstruction

The reconstructed J/ψ and ϕ mesons are combined to form B_s^0 candidates. The fit of the common B_s^0 decay vertex is required to fulfill $\chi^2_{\text{Vtx}}/\text{ndof} < 10$. The reconstructed B_s^0 mass needs to be in the mass range $[5100, 5550]\,\text{MeV}$. Furthermore the B_s^0 impact parameter with respect to the primary vertex is constrained to ensure that the reconstructed B_s^0 originated from the primary vertex.

In addition to the vertex fit which is used to determine the χ^2_{Vtx} discussed in the last paragraph a global kinematic fit of the decay tree is performed [64]. The global kinematic fit parameterizes the full decay tree in terms of particle momenta, vertex positions and decay lengths. The decay tree is fit using least squares minimization taking into account four-momentum conservation at each vertex. The measurements of the track momenta and positions as well as their covariance matrices are used as input to the fit and the uncertainties are correctly propagated. The global kinematic fit also constrains the B_s^0 candidate to originate from the primary vertex.

To determine the reconstructed B_s^0 mass which is used for signal and background separation an additional constraint is applied. The mass of the dimuon system is fixed

$\phi \to K^+K^-$ selection		
Criterion	Stripping	Offline selection
$\Delta \ln \mathcal{L}_{K\pi}$	> -2	> 0
$\chi^2_{\text{Track}}/\text{ndof}(K)$	< 5	< 4
$p_T(\phi)$	$> 1\,\text{GeV}$	$> 1\,\text{GeV}$
$M(\phi)$	$\in [980, 1050]\,\text{MeV}$	$\in [1008, 1032]\,\text{MeV}$
$\chi^2_{\text{Vtx}}(\phi)$	< 16	< 16

$B_s \to J/\psi\,\phi$ selection		
Criterion	Stripping	Offline selection
$M(B_s)$	$\in [5100, 5550]\,\text{MeV}$	$\in [5200, 5550]\,\text{MeV}$
$\chi^2_{\text{Vtx}}/\text{ndof}(B_s)$	< 10	< 10
$\chi^2_{\text{DTF(B+PV)}}/\text{ndof}(B_s)$	-	< 5
$\chi^2_{\text{IP}}/\text{ndof}(B_s)$	-	< 25

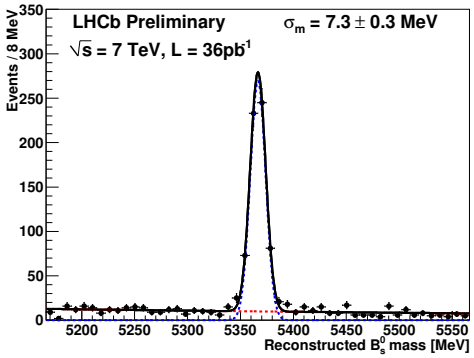
Table 3.4.: Selection criteria for the signal decay $\text{B}^0_\text{s} \to \text{J}/\psi\,\phi$. The J/$\psi$ selection is given in Table 3.3. Shown are the selection criteria used in the stripping as well as in the offline selection. Table from [63].

to the true J/ψ mass given by $M^{\text{PDG}}_{\text{J}/\psi}$. Figure 3.2a shows the reconstructed B^0_s mass for the *lifetime unbiased* data sample after the full selection with the mass constraint applied. To illustrate the effect of the J/ψ mass constraint Figure 3.2b shows the reconstructed B^0_s mass without using the J/ψ mass constraint. The mass resolution in this case decreases from 7.3 MeV to 18.4 MeV. Considering the background levels applying the J/ψ mass constraint leads to an increase of the signal to background ratio from 4.6 to 11.2 (evaluated $3\sigma_m$ around the B^0_s mass peak). Figure 3.3 shows, in addition, the reconstructed B^0_s mass for both the *lifetime unbiased* and the *lifetime biased* data sample. In both cases the J/ψ mass constraint is applied. The mass peaks are clearly defined and the observed background contribution is very low.

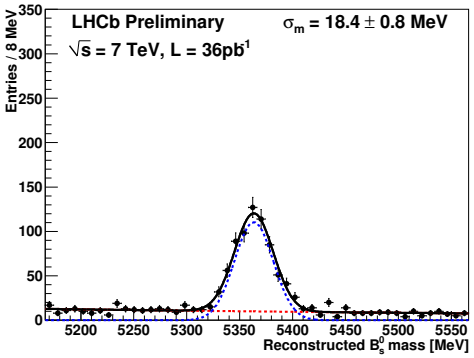
The χ^2 of the global kinematic fit without J/ψ mass constraint, denoted as $\chi^2_{\text{DTF(B+PV)}}$, is used as selection criterion to ensure that the reconstructed signal candidate is compatible with the decay of a B^0_s meson originating from the primary vertex. In the case of multiple candidates per event the candidate with the smallest $\chi^2_{\text{DTF(B+PV)}}$ is chosen.

B^0_s decay time

The decay time of the B^0_s, often also referred to as the B^0_s proper time, is defined as the time after which the B^0_s meson decays in its own rest frame. The decay time can be

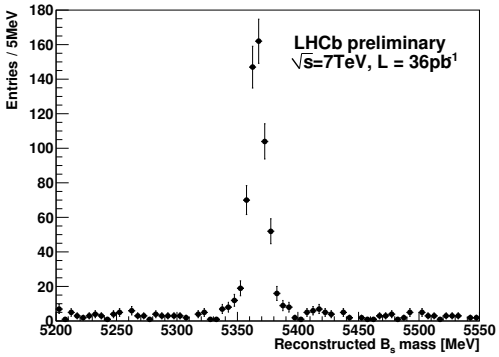


(a) J/ψ mass constraint applied

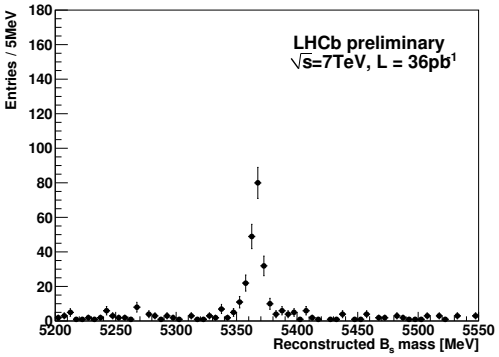


(b) J/ψ mass constraint not applied

Figure 3.2.: The reconstructed B_s^0 mass peak (a) with and (b) without the J/ψ mass constraint for *lifetime unbiased* data.

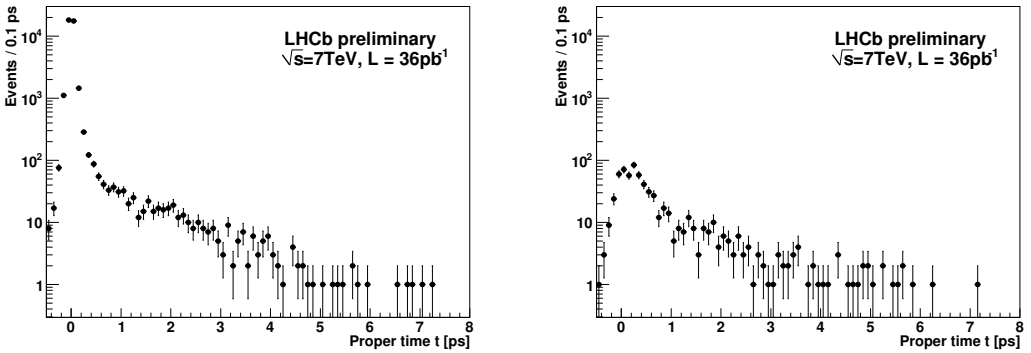


(a) Unbiased events



(b) Biased events

Figure 3.3.: Distribution of the reconstructed B_s^0 mass of $B_s^0 \to J/\psi\,\phi$ candidates triggered by (a) the *unbiased* and (b) the *biased* trigger lines after the full selection.



(a) Unbiased events (b) Biased events

Figure 3.4.: Proper time distribution of $B_s^0 \to J/\psi\,\phi$ candidates triggered by (a) the *unbiased* and (b) the *biased* trigger lines after the full selection with the exception of the proper time criterion $t > 0.3\,\text{ps}$.

calculated using the decay length $\vec{d}$ after which the B_s^0 decays via

$$t = \frac{|\vec{d}|}{\beta\gamma} = \frac{|\vec{d}|}{|\vec{p}|} m_{B_s^0},$$

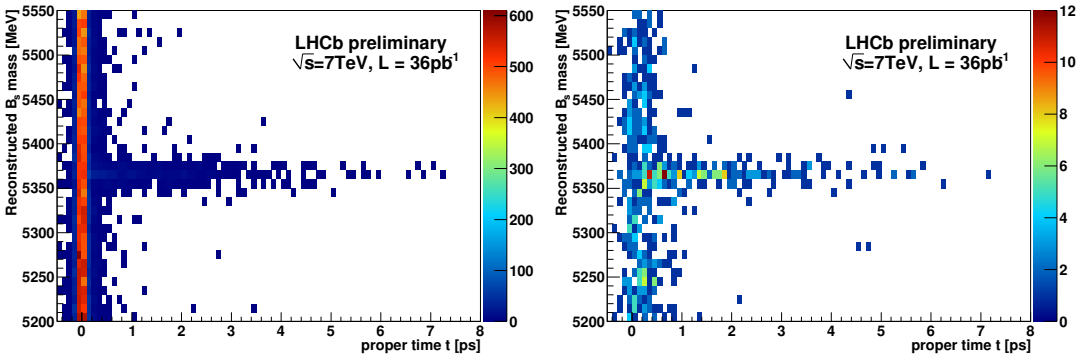
where $m_{B_s^0}$ denotes the reconstructed B_s^0 mass. The decay length $\vec{d}$ is defined as the distance between the primary (production) vertex $\vec{PV}$ and the secondary (decay) vertex $\vec{SV}$. Using $\vec{d} = \vec{SV} - \vec{PV}$ the proper time is given by

$$t = \frac{\vec{p}_{B_s^0} \cdot (\vec{SV} - \vec{PV})}{|\vec{p}_{B_s^0}|^2} m_{B_s^0}.$$

In this analysis the decay time is extracted using the global kinematic fit described in the last section since it properly propagates the tracking uncertainties. No J/ψ mass constraint is applied since such a constraint could introduce correlations between the reconstructed B_s^0 mass and the decay time. Figure 3.4 shows the decay time for the *lifetime unbiased* and the *lifetime biased* data sample. The very different distributions for small decay times can be explained by the large correlation of the decay time with the IP of the muon tracks that are required to be larger than 0.11 mm for the *lifetime biased* trigger lines. This criterion clearly cuts into the proper time distribution.

Primary vertex association

For events in which several primary vertices are reconstructed the association of a wrong primary vertex as production vertex of the B_s^0 meson would lead to a misreconstruction of the decay time. To prevent this the IP of the B_s^0 with respect to all primary vertices in the event is calculated. If there is more than one primary vertex for which the χ^2_{IP} is smaller than 50 the B_s^0 candidate is rejected.



(a) Unbiased events (b) Biased events

Figure 3.5.: Proper time distribution versus the reconstructed B_s^0 mass of $B_s^0 \to J/\psi\,\phi$ candidates triggered by (a) the *unbiased* and (b) the *biased* trigger lines after the full selection with the exception of the proper time criterion $t > 0.3\,\text{ps}$.

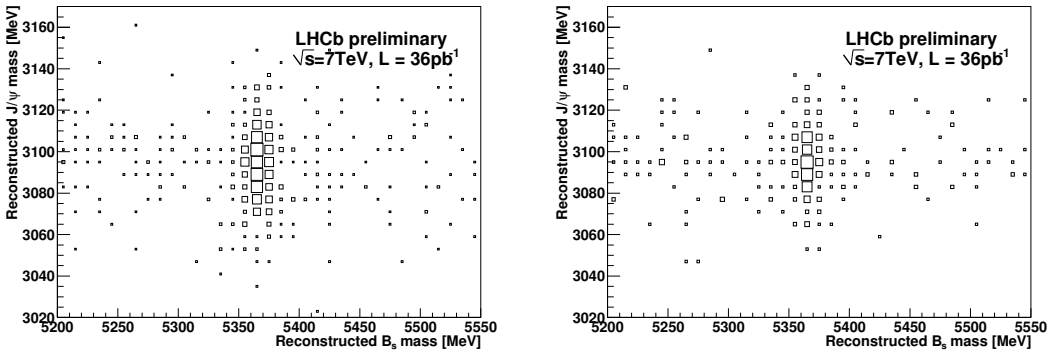
Background

Figure 3.5 shows the decay time versus the reconstructed B_s^0 mass for the *lifetime unbiased* and the *biased* data sample. A clear enhancement around the B_s^0 mass can be seen which constitutes the $B_s^0 \to J/\psi\,\phi$ signal. As is expected for the signal decay from longlived B_s^0 mesons ($\tau(B_s^0) = 1.472^{+0.024}_{-0.026}\,\text{ps}$ [27]) the B_s^0 candidates can reach large decay times. In contrast to this a large number of candidates around $t = 0$ in Figure 3.5a show no discernible enhancement around the B_s^0 mass. These candidates constitute combinatoric background from tracks originating from the primary vertex. In the following this background is referred to as "prompt background". To reject this type of background which overwhelms the signal component for small decay times it was decided to require $t > 0.3\,\text{ps}$ for the analysis. The loss of sensitivity on the physics parameters due to this requirement is negligible. However it reduces both the data set and also the computing requirements for the extraction of the physics parameters by a factor of ~ 36.

Figure 3.6 shows the *lifetime unbiased* and *biased* data sample in the two-dimensional plane of the reconstructed B_s^0 mass and J/ψ mass. Practically no candidates are found in the J/ψ mass sidebands, i. e. the remaining background events in the B_s^0 mass sidebands mainly originate from true J/ψ decays with additional random tracks to form the B_s^0 meson.

3.4. Efficiency of trigger and selection

The efficiency of the trigger and the offline selection on signal events is determined in a study using fully simulated signal events. To lower the computational costs of the event generation simulated events are only accepted if all decay products of the signal decay



(a) Unbiased events (b) Biased events

Figure 3.6.: Reconstructed B_s^0 mass versus the reconstructed J/ψ mass of $B_s^0 \to J/\psi\,\phi$ candidates triggered by the (a) unbiased and (b) biased trigger lines after the full selection. The cut on the proper time of $t > 0.3\,\text{ps}$ has also been applied.

lie in the acceptance of the LHCb detector ($10\,\text{mrad} < \Theta < 400\,\text{mrad}$). This requirement was found to have an efficiency of

$$\epsilon_{\text{accept}} \;=\; (18.14 \pm 0.04)\%\ [65].$$

Two million signal events were produced within the LHCb acceptance and the full signal selection without trigger requirements is applied. The resulting selection efficiency is determined to be

$$\epsilon_{\text{sel}} \;=\; 14.50\%.$$

On the fully selected signal sample the following trigger efficiencies are found

$$\begin{aligned}
\epsilon_{\text{L0}} &= 91.88\% \\
\epsilon_{\text{L0}} \cdot \epsilon^{\text{HLT}}_{\text{biased}} &= 56.55\% \\
\epsilon_{\text{L0}} \cdot \epsilon^{\text{HLT}}_{\text{unbiased}} &= 60.84\% \\
\epsilon_{\text{L0}} \cdot \epsilon^{\text{HLT}}_{\text{biased|unbiased}} &= 74.77\%.
\end{aligned}$$

The resulting total efficiency for a $B_s^0 \to J/\psi\,[\mu^+\mu^-]\phi\,[K^+K^-]$ signal decay to be triggered, reconstructed and selected is therefore

$$\begin{aligned}
\epsilon_{\text{tot}} &= \epsilon_{\text{accept}} \cdot \epsilon_{\text{sel}} \cdot \epsilon_{\text{L0}} \cdot \epsilon^{\text{HLT}}_{\text{biased|unbiased}} \\
&= 1.97\%.
\end{aligned}$$

3.5. Signal yield

The expected signal yield can be calculated with the formula

$$\mathcal{S} = \mathcal{L}_{\text{int}} \cdot \sigma_{\text{b}\bar{\text{b}}} \cdot 2 \cdot f_s \cdot \mathcal{B}_{\text{vis}}(B_s^0 \to J/\psi\,\left[\mu^+\mu^-\right]\phi\,\left[K^+K^-\right]) \cdot \epsilon_{\text{tot}}.$$

	candidates	candidates ($t > 0.3\,\mathrm{ps}$)	signal yield	signal yield ($t > 0.3\,\mathrm{ps}$)
unbiased only	38225	250	230 ± 53	161 ± 13
biased only	653	345	208 ± 16	196 ± 15
both	1123	521	398 ± 22	400 ± 20
total	40001	1116	836 ± 60	757 ± 28

Table 3.5.: Number of candidates (including background) and signal yields from a one-dimensional fit to the reconstructed $\mathrm{B_s^0}$ mass. The number of candidates and the signal yield are given for both the full proper time range $[-1, 14]\,\mathrm{ps}$ and for proper times larger than 0.3 ps. Table from [35].

The integrated luminosity $\mathcal{L}_{\mathrm{int}}$ taken in 2010 is $36\,\mathrm{pb}^{-1}$. The $\mathrm{b\bar{b}}$ cross section was measured in [44] to be $\sigma_{\mathrm{b\bar{b}}} = (288 \pm 4 \pm 48)\,\mu\mathrm{b}$. The factor 2 is needed because a pair of b quarks is produced. f_s gives the hadronization probability for a $\overline{\mathrm{b}}$ (b) quark to form a $\mathrm{B_s^0}$ ($\mathrm{\bar{B}_s^0}$) meson. The hadronization probability is measured to be $f_s = (10.7 \pm 1.2)\%$ [27]. $\mathcal{B}_{\mathrm{vis}}(\mathrm{B_s^0} \to \mathrm{J/\psi}\,[\mu^+\mu^-]\phi\,[\mathrm{K^+K^-}])$ denotes the visible branching fraction which is the probability for a $\mathrm{B_s^0}$ to decay into the reconstructed final state $\mu^+\mu^- K^+ K^-$. The visible branching fraction is therefore the product of the probability of the $\mathrm{B_s^0}$ meson to decay into $\mathrm{J/\psi}\,\phi$ given by $\mathcal{B}(\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi) = (1.3 \pm 0.4) \cdot 10^{-3}$ [27], the probability of the $\mathrm{J/\psi}$ to decay to two muons $\mathcal{B}(\mathrm{J/\psi} \to \mu^+\mu^-) = (5.93 \pm 0.06) \cdot 10^{-2}$ [27] and the probability for a ϕ meson to decay to two kaons $\mathcal{B}(\phi \to K^+K^-) = (48.9 \pm 0.5) \cdot 10^{-2}$ [27]. This results in a visible branching fraction of $\mathcal{B}_{\mathrm{vis}}(\mathrm{B_s^0} \to \mathrm{J/\psi}\,[\mu^+\mu^-]\phi\,[\mathrm{K^+K^-}]) = (3.77 \pm 1.16) \cdot 10^{-5}$. Putting everything together the expected signal yield for 2010 data is found to be

$$\begin{aligned} \mathcal{S} &= 36\,\mathrm{pb}^{-1} \cdot 288\,\mu\mathrm{b} \cdot 2 \cdot 0.107 \cdot 3.77 \cdot 10^{-5} \cdot 1.97 \cdot 10^{-2} \\ &= 1645 \pm 605 \end{aligned}$$

where the error results from the large uncertainty of the visible branching fraction, the $\mathrm{b\bar{b}}$ cross section and the hadronization fraction f_s. The signal yield found in 2010 data after the full selection is given in Table 3.5. The predicted signal yield from simulation and the yield found in data are still compatible considering the large statistical uncertainties.

3.6. Proper time reconstruction

3.6.1. Proper time resolution

The measurement of the proper time of the $\mathrm{B_s^0}$ is limited by detector resolution and multiple scattering which results in a finite proper time resolution. It is instructive to review the influence of the proper time resolution on the time dependent CP asymmetry in the interference between decay and mixing and decay. For a $\mathrm{B_s^0}$ meson decaying into

a CP eigenstate f the CP asymmetry is given by equation 1.25

$$\begin{aligned}\mathcal{A}_{\mathrm{CP}}(t) &= \frac{\Gamma(\bar{\mathrm{B}}_\mathrm{s}^0 \to f)(t) - \Gamma(\mathrm{B}_\mathrm{s}^0 \to f)(t)}{\Gamma(\bar{\mathrm{B}}_\mathrm{s}^0 \to f)(t) + \Gamma(\mathrm{B}_\mathrm{s}^0 \to f)(t)} \\ &= \frac{-\eta_f \sin\phi_s \sin(\Delta m_\mathrm{s} t)}{\cosh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right) - \eta_f|\cos\phi_s|\sinh\left(\frac{\Delta\Gamma_\mathrm{s}}{2}t\right)},\end{aligned}$$

where direct CP violation is neglected. Setting $\Delta\Gamma_\mathrm{s} = 0$ for simplicity and selecting a final state with definite CP value (e. g. J/$\psi\,f_0$) the asymmetry becomes[11]

$$\mathcal{A}_{\mathrm{CP}}(t) = \sin(\Delta m_\mathrm{s} t)\sin\phi_s.$$

When the proper time t is determined with a finite resolution σ_t the oscillation $\sim \sin(\Delta m_\mathrm{s} t)$ is diluted. To determine the effect of a proper time resolution model $\mathcal{R}(t)$ on the proper time dependent function $f(t)$ the convolution integral

$$(f \otimes \mathcal{R})(t) = \int_{-\infty}^{+\infty} f(t')\mathcal{R}(t-t')\mathrm{d}t'$$

needs to be solved. For the proper time dependent CP asymmetry $\mathcal{A}_{\mathrm{CP}}(t)$ and a Gaussian resolution model this integral is given by

$$\begin{aligned}\mathcal{A}_{\mathrm{CP}}(t) \otimes \frac{1}{\sqrt{2\pi}\sigma_\mathrm{t}} e^{-\frac{t^2}{2{\sigma_\mathrm{t}}^2}} &= \int_{-\infty}^{+\infty} \sin\phi_s \sin(\Delta m_\mathrm{s} t') \frac{1}{\sqrt{2\pi}\sigma_\mathrm{t}} e^{-\frac{(t-t')^2}{2{\sigma_\mathrm{t}}^2}} \mathrm{d}t' \\ &= e^{-\frac{\Delta m_\mathrm{s}^2 {\sigma_\mathrm{t}}^2}{2}} \sin\phi_s \sin(\Delta m_\mathrm{s} t)\,.\end{aligned}$$

The observed CP asymmetry reduces to

$$\mathcal{A}_{\mathrm{meas}} = \mathcal{D}_{\mathrm{res}}\mathcal{A}_{\mathrm{CP}}$$

with the dilution factor $\mathcal{D}_{\mathrm{res}} = \exp(-\Delta m_\mathrm{s}^2 {\sigma_\mathrm{t}}^2/2)$. Using a proper time resolution model which does not accurately describe the data will lead to a bias for the extracted asymmetry. The determination of the proper time resolution is therefore very important for the extraction of ϕ_s. A method to extract the proper time resolution from data is discussed in the next section.

The statistical error of the measured asymmetry $\mathcal{A}_{\mathrm{meas}}$ is given by binomial statistics

$$\sigma(\mathcal{A}_{\mathrm{meas}}) = \sqrt{\frac{1-\mathcal{A}_{\mathrm{meas}}^2}{N}}.$$

Propagating the statistical uncertainty from the measured quantity $\mathcal{A}_{\mathrm{meas}}$ to the theoretical asymmetry $\mathcal{A}_{\mathrm{CP}}$ related to $\sin\phi_s$ results in

$$\sigma(\mathcal{A}_{\mathrm{CP}}) = \frac{1}{\mathcal{D}_{\mathrm{res}}}\sqrt{\frac{1-\mathcal{A}_{\mathrm{meas}}^2}{N}},$$

where possible background pollution is neglected. The statistical power of a data sample consisting of N events with proper time resolution σ_t is the same as the statistical power of a data sample containing $\mathcal{D}_{\mathrm{res}}^2 \cdot \mathrm{N}$ events with perfect proper time reconstruction.

[11] The same derivation qualitatively also applies for the signal decay $\mathrm{B}_\mathrm{s}^0 \to \mathrm{J}/\psi\,\phi$ with the additional complication of the angular analysis to separate CP-even and CP-odd amplitudes.

Per event proper time resolution

As discussed previously the proper time is determined in the global kinematic fit. In addition to the proper time this fit also provides an estimate for the proper time uncertainty for each candidate. Using this per event estimate for the proper time uncertainty is in principle beneficial because events for which the proper time was determined more precisely get a higher weight in the fit. The uncertainty estimate σ_t returned by the kinematic fit can possibly over-or underestimate the proper time resolution due to imperfect description of material in the detector. To correct for this effect an additional scaling factor s_t is introduced and the product $s_t\sigma_t$ is used instead of σ_t. The scaling factor can be determined from a fit to prompt background events. For 2010 data a scaling factor of ~ 1.5 is found.

The statistical power of the per event resolution model is given by

$$\left\langle \mathcal{D}^2_{\text{res}} \right\rangle = \frac{1}{N} \sum_{e=1}^{N} e^{-\Delta m_s^2 s_t^2 \sigma_{t,e}^2}.$$

Using this formula the gain in statistical power using the proper time resolution for every event is found to be smaller than 6%. This results in a sensitivity gain for ϕ_s which is less than 3%.

Determination of the proper time resolution using prompt background events

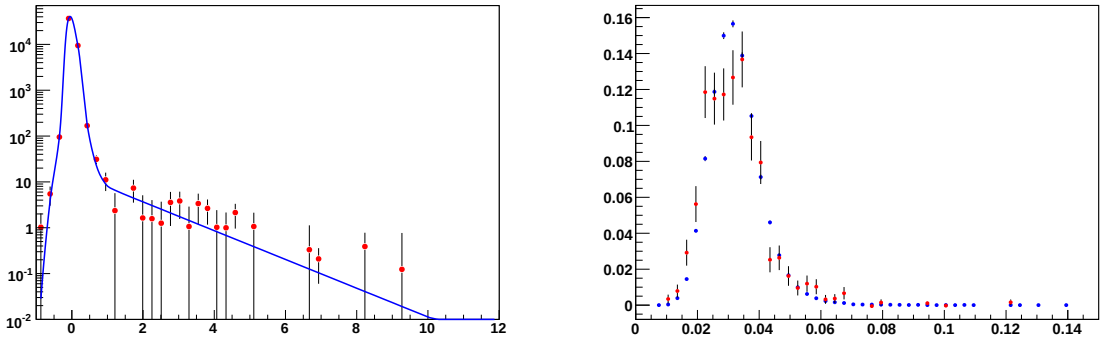
A way to determine the proper time resolution without relying on Monte Carlo simulation is to use the prompt background events. Since the prompt background consists mainly of true J/ψ mesons originating from the primary vertex their true proper time distribution is given by a delta distribution centered around $t = 0$. The reconstructed proper time distribution of the prompt background then reflects the proper time resolution model.

Figure 3.7a gives the proper time distribution of the background component extracted from data using the sPlot technique [66][12]. The large peak around $t = 0$ constitutes the prompt background which is used for the determination of the resolution parameters. The prompt background is described by a triple Gaussian model with widths $\sigma^{\text{sig}}_{\text{t,1}}$, $\sigma^{\text{sig}}_{\text{t,2}}$ and $\sigma^{\text{sig}}_{\text{t,3}}$ and relative fractions $f^{\text{sig}}_{\text{t,1}}$ and $f^{\text{sig}}_{\text{t,2}}$, the exact functional form is given in equation 4.10. The events with large positive lifetimes in Figure 3.7a are longlived background events. They are modeled according to a double exponential and not used for the determination of the resolution parameters. To extract the resolution parameters the sFit technique [67] is used. The resolution parameters determined in [63] are given in Table 3.6.

To be able to use this background resolution model for the signal decays the assumption that background and signal component have identical proper time resolution needs to be confirmed. Figure 3.7b shows the per event proper time error for both the signal and background component. The sPlot technique [66] is used to extract the two components from data[13]. No large difference in proper time resolution is seen for signal

[12]The reconstructed B_s^0 mass is used to assign background and signal weights to every event. Using these weights the background distribution can be extracted from data.

[13]The reconstructed B_s^0 mass is used as discriminating variable.



(a) Background proper time distribution

(b) Per event proper time error estimate

Figure 3.7.: (a) The proper time distribution for background events is dominated by the large prompt background component. The background component is extracted using the sPlot technique [66]. (b) Distribution of the per event proper time error estimate for background (blue) and signal (red) events. Both components are extracted using the sPlot technique. The distributions are statistically compatible which justifies using the resolution model determined from prompt background events for the signal component. Figures from [63].

and background component. This justifies using the resolution model determined from the prompt background for the signal component.

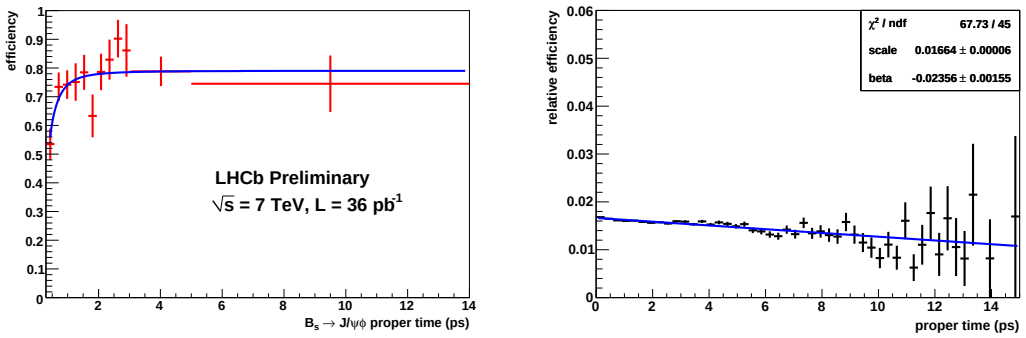
3.6.2. Proper time acceptance

Lifetime biased trigger lines

Figure 3.8a shows the proper time acceptance of the *lifetime biased* trigger lines for $B_s^0 \to J/\psi\,\phi$ events on data. The efficiency of the *lifetime biased* events is determined relative to events which are triggered by both the *lifetime biased* and *lifetime unbiased* trigger lines. Events triggered by both lines simultaneously should have the same trigger

Resolution Parameter	Value
$\sigma_{t,1}^{sig}$	0.0337 ps
$\sigma_{t,2}^{sig}$	0.0646 ps
$\sigma_{t,3}^{sig}$	0.183 ps
$f_{t,1}^{sig}$	0.527
$f_{t,2}^{sig}$	0.964

Table 3.6.: Resolution parameters extracted from prompt background events in data [63].



(a) Lifetime biased trigger lines (b) Lifetime unbiased trigger lines

Figure 3.8.: (a) The efficiency of the *biased* relative to the *unbiased* trigger lines on data. Figure from [35]. (b) Proper time acceptance caused by a reconstruction effect. The efficiency is determined using simulated signal events. Figure from [63].

and selection efficiency as events triggered only by the *lifetime unbiased* line. This assumption is confirmed in [35].

To determine the efficiency of the *lifetime biased* trigger lines with respect to the *lifetime unbiased* events the sPlot technique [66] is used to extract the $B_s^0 \to J/\psi\,\phi$ signal component from both data samples using the reconstructed B_s^0 mass. The functional form

$$\epsilon'(t)_{\text{Biased}} \quad = \quad n\frac{(at)^c}{1+(at)^c}$$

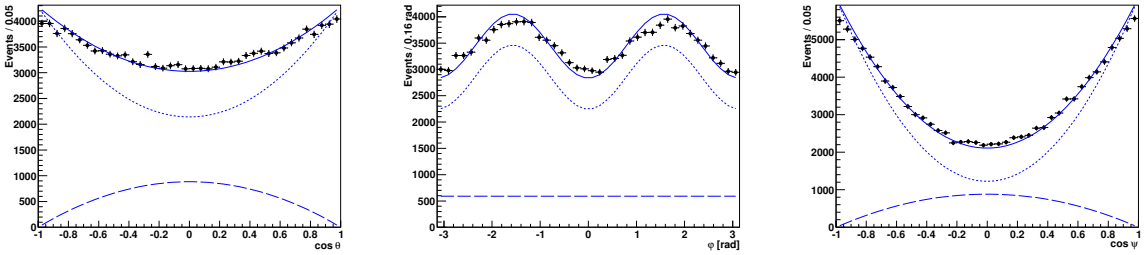
is used as parameterization of the proper time dependent efficiency for the *lifetime biased* trigger lines. From a fit to the data the parameters are determined to be $a = (3.23 \pm 0.63)\ \text{ps}^{-1}$ and $c = 2.48 \pm 1.07$ [35].

Reconstruction efficiency

In [63] a study of fully simulated signal events revealed a reconstruction effect which results in a proper time dependent acceptance. Figure 3.8b shows that the proper time acceptance declines linearly with the proper time. The acceptance effect is therefore parameterized as

$$\epsilon(t)_{\text{Unbiased}} = 1 + \beta t \tag{3.1}$$

with $\beta = (-0.025 \pm 0.025)\,\text{ps}^{-1}$ [63]. A study of the origin of this effect is still ongoing. At least part of the effect is caused by an inefficiency of the pattern recognition if the radial distance of the secondary vertex to the beamline is large ($> 1\,\text{mm}$). The large radial distance of the B_s^0 decay vertex is correlated with large proper times, leading



(a) Transversity angle $\cos\theta$ (b) Transversity angle φ (c) Transversity angle $\cos\psi$

Figure 3.9.: The transversity angles from Monte Carlo simulated signal events overlayed with the angular distributions from theory, calculated according to the values used in the event generation. The solid blue line denotes the signal model from theory, the dotted and dashed blue lines denote the CP-even and CP-odd components respectively. Clear discrepancies in the angles are seen which need to be corrected for when the physics parameters are determined from data.

to a proper time dependent efficiency. For this analysis the full size of the correction parameter is used as systematic uncertainty. The value of $\Delta\Gamma_s$ determined in chapter 6 is unaffected by this conservative estimate.

Since the time dependent efficiency of the *lifetime biased* trigger lines was determined with respect to the *lifetime unbiased* events the linear reconstruction efficiency also needs to be applied to the *lifetime biased* events resulting in

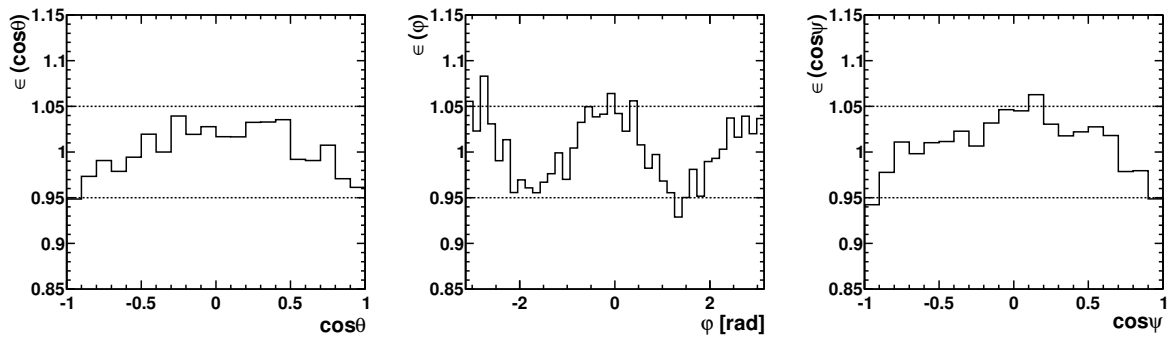
$$\epsilon(t)_{\text{Biased}} \quad = \quad (1+\beta t)\frac{(at)^c}{1+(at)^c} \tag{3.2}$$

with $\beta = (-0.025 \pm 0.025)\,\text{ps}^{-1}$, $a = (3.23 \pm 0.63)\ \text{ps}^{-1}$ and $c = 2.48 \pm 1.07$.

3.7. Reconstruction of the transversity angles

The transversity angles $\Omega = \{\cos\theta, \varphi, \cos\psi\}$ are reconstructed according to the prescription given in Figure 1.6. Figure 3.9 gives the reconstructed transversity angles for a Monte Carlo simulated signal sample. In addition to the angular distribution of the reconstructed simulated signal events the expected distributions from theory are shown as solid blue line. The expectation from theory is calculated using the parameters which were used in the generation of the simulated signal events. Clear discrepancies between the reconstructed data and the theoretical distributions are observed. Figure 3.10 shows the resulting angular dependent acceptance (Plots are normalized). The relative variation of the angular acceptance is of the order of $\pm 5\%$.

As discussed in section 1.3.1 the angular distributions are necessary to statistically disentangle the CP-even from the CP-odd component in the final state. Distortions



(a) $\epsilon(\cos\theta)$ (b) $\epsilon(\varphi)$ (c) $\epsilon(\cos\psi)$

Figure 3.10.: Angular efficiencies depending on the three transversity angles (a) $\cos\theta$, (b) φ and (c) $\cos\psi$. The figures show the efficiency histogram determined from Monte Carlo simulated signal events. The dashed lines denote relative variations of the acceptance of $\pm 5\%$. The y-scale was deliberately set to 1 to be able to easier read off the relative size of the acceptance effect.

of the angular distributions can therefore lead to deviations of the physics parameters determined in the angular analysis. In particular the transversity amplitudes, giving the CP-odd and CP-even fractions in data are affected by angular distortions. For the extraction of the physics parameters from data it is therefore important to model this angular acceptance effect in the fitting method. Section 4.5.7 describes the implementation of the three-dimensional efficiency $\epsilon(\Omega)$ in detail.

Studies using generator level Monte Carlo simulated signal events without any selection cuts revealed that the angular acceptance effect is mainly caused by the polar angle acceptance of the LHCb spectrometer and implicit momentum cuts for muons and kaons [68]. Since the description of the angular acceptance effect is taken from simulated signal events it is mandatory that these distributions are well described in the simulation. To verify this assumption the simulated signal events are compared with the signal component from data extracted using the sideband subtraction technique[14]. Figure 3.11 shows the polar angle distributions and Figure 3.12 gives the momenta of the final state particles. The red points denote simulated signal events, the black points give the signal component extracted from data. Reasonable agreement within the statistics is seen for all distributions with the possible exception of the muon momenta.

The systematic uncertainty due to a possible incorrect description of the muon momenta is studied in sections 6.2 and 7.2. The simulated events are reweighted according to the muon momenta and a new acceptance correction is derived. Repeating the fit using this acceptance correction gives an estimate of the size of the systematic effect.

[14]The sideband subtraction technique extracts the signal component from data by subtracting events from the B_s^0 mass sidebands from the signal region. The sideband events that are subtracted are

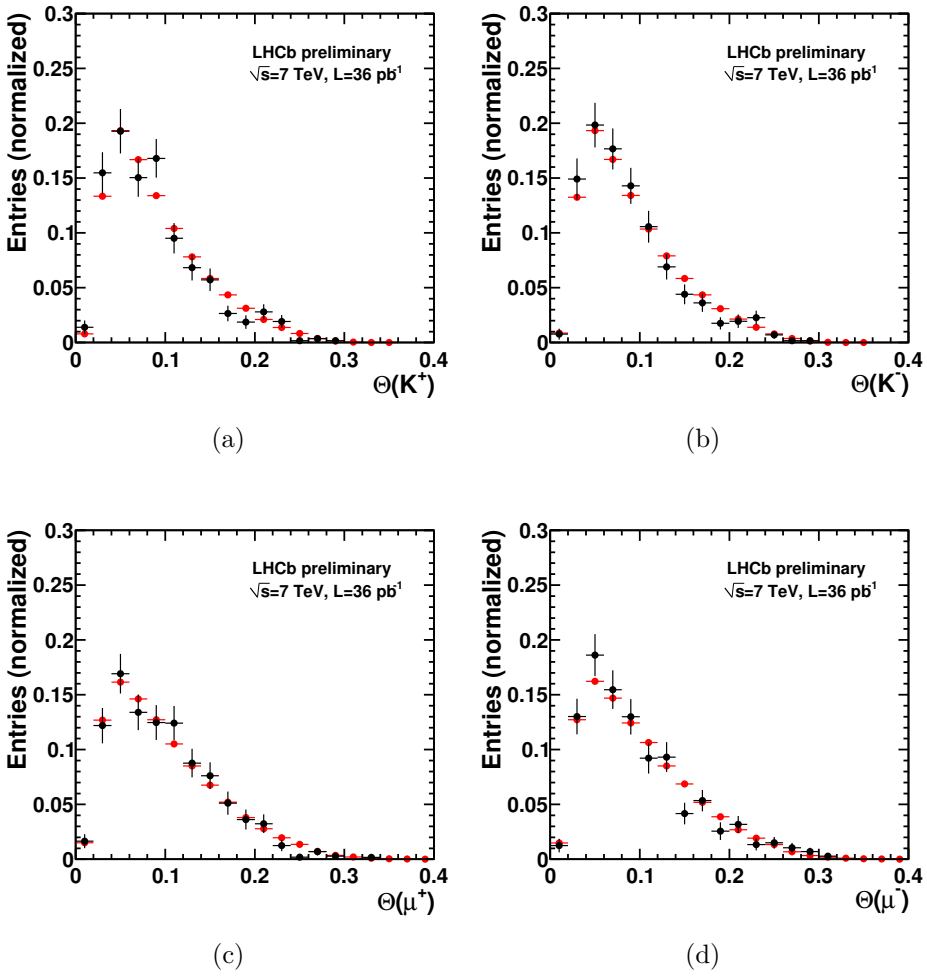


Figure 3.11.: Polar angle distributions of the final state particles for simulated signal events (red) and sideband subtracted data (black). The distributions show good agreement.

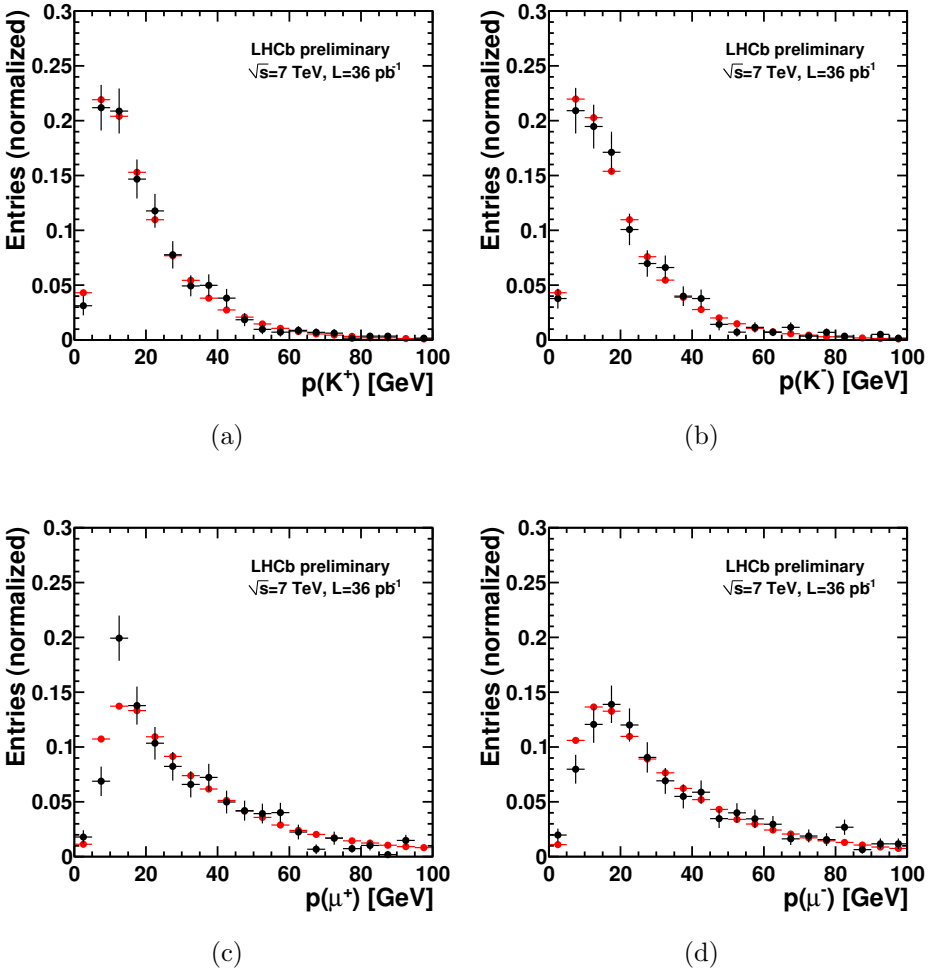


Figure 3.12.: Momentum distributions of the final state particles for simulated signal events (red) and sideband subtracted data (black). The distributions show reasonably good agreement with the possible exception of the muon momenta. This possible discrepancy is explored as one of the systematic uncertainties in section 6.2.

3.8. Determination of the B_s^0 production flavor

The determination of the B_s^0 production flavor is essential for the extraction of ϕ_s and is called "Flavor Tagging". There are two approaches to determine the production flavor of a reconstructed B_s^0 meson, *same side flavor tagging* and *opposite side flavor tagging*. Both approaches are illustrated in Figure 3.13.

- *Same side flavor tagging* exploits the B_s^0 hadronization. Besides the $\bar{b}$-quark a signal B_s^0 meson also contains an s-quark. The s-quark itself was created together with its antiparticle in the hadronization process. This $\bar{s}$-quark may pick up a light quark and form a K^+ kinematically close to the B_s^0 meson. For a $\bar{B}_s^0$ system a negatively charged K^- would be the result of the hadronization process. Searching for charged kaons in the vicinity of the signal B_s^0 therefore can give information on the production flavor of the B_s^0. This type of flavor tagging is called *same side kaon tagging* (SSK). For the B_d^0 system pions instead of kaons can be used as same side tagging particles as indicated in the upper part of Figure 3.13.

- *Opposite side tagging* algorithms use the b-quark which was created together with the $\bar{b}$-quark contained in the signal B_s^0 meson, as is illustrated in the lower part of Figure 3.13. This b-quark can decay semileptonically in a $b \to c\ell^+\nu_\ell$ transition. The lepton charge then can be used to determine the production flavor of the signal B. Further down the decay chain the c quark produced in the decay mentioned above can hadronize and produce a charmed meson which then decays preferably into a kaon. The kaon charge can also be used by the tagging algorithms. A final algorithm to determine the B_s^0 production flavor exploits the vertex charge of the secondary vertex. The algorithm sums up the charges of the tracks originating from the secondary vertex weighted by the track momenta. The combination of the results of all tagging algorithms is referred to as *opposite side tag*.

The main quantities describing a tagging algorithm are the tagging efficiency ϵ_{tag} and the mistag probability ω_{tag}. The tagging efficiency gives the ratio of events for which the tagging algorithms are able to deliver a decision, the mistag probability gives the ratio of events for which this decision is wrong.

Tagging dilution

It is again instructive to study the time dependent CP asymmetry from equation 1.25. For an imperfect determination of the B_s^0 production flavor the CP asymmetry

$$\mathcal{A}_{\text{CP}}(t) = \frac{\Gamma(\bar{B}_s^0 \to f)(t) - \Gamma(B_s^0 \to f)(t)}{\Gamma(\bar{B}_s^0 \to f)(t) + \Gamma(B_s^0 \to f)(t)}$$

weighted according to the background fraction in the signal region.

is diluted to

$$\begin{aligned}\mathcal{A}_{\text{meas}} &= \frac{(1-\omega_{\text{tag}})\Gamma(\bar{B}_s^0 \to f) + \omega_{\text{tag}}\Gamma(B_s^0 \to f) - (1-\omega_{\text{tag}})\Gamma(B_s^0 \to f) - \omega_{\text{tag}}\Gamma(\bar{B}_s^0 \to f)}{(1-\omega_{\text{tag}})\Gamma(\bar{B}_s^0 \to f) + \omega_{\text{tag}}\Gamma(B_s^0 \to f) + (1-\omega_{\text{tag}})\Gamma(B_s^0 \to f) + \omega_{\text{tag}}\Gamma(\bar{B}_s^0 \to f)} \\ &= \mathcal{D}_{\text{tag}}\mathcal{A}_{\text{CP}}\end{aligned}$$

with the tagging dilution $\mathcal{D}_{\text{tag}} = (1 - 2\omega_{\text{tag}})$. Together with the proper time resolution, the observed asymmetry becomes

$$\mathcal{A}_{\text{meas}} = \mathcal{D}_{\text{tag}}\mathcal{D}_{\text{res}}\mathcal{A}_{\text{CP}}.$$

The statistical uncertainty on the CP asymmetry is given by

$$\sigma(\mathcal{A}_{\text{CP}}) = \frac{1}{\mathcal{D}_{\text{tag}}\mathcal{D}_{\text{res}}}\sqrt{\frac{1-\mathcal{A}_{\text{meas}}^2}{\epsilon_{\text{tag}}N}}$$

where the tagging efficiency ϵ_{tag} is included to take into account that the tagging algorithms are not able to determine the flavor tag for every event. From the perspective of the flavor tagging the statistical power of the data sample is determined by the effective tagging power defined as

$$\epsilon_{\text{tag}}^{\text{eff}} = \epsilon_{\text{tag}}\mathcal{D}_{\text{tag}}^2.$$

A data sample with N signal events and tagging power $\epsilon_{\text{tag}}^{\text{eff}}$ has the same statistical power as a data sample with $\epsilon_{\text{tag}}^{\text{eff}} \cdot \text{N}$ signal events and perfect knowledge of the B_s^0 production flavor for every event.

Per event mistag

Besides the tagging decision the tagging algorithms also return an estimated mistag probability for every event depending on kinematic properties of the event. This constitutes additional information which is beneficial for the extraction of ϕ_s since candidates for which the determination of the production flavor is more reliable are given a larger weight. The effective tagging power using per event mistag is given by

$$\epsilon_{\text{tag}}\left\langle\mathcal{D}_{\text{tag}}^2\right\rangle = \frac{1}{N}\sum_{\text{tagged } e}(1-2\omega_{\text{tag,e}})^2,$$

where the sum only includes tagged events. On 2010 data an increase in effective tagging power of $\sim 23\%$ is observed when using per event mistag. This results in an estimated gain in sensitivity on $\sin\phi_s$ of $\sim 11\%$.

3.8.1. Calibration of the B_s^0 production flavor determination

To be able to use the per event mistag given by the tagging algorithms it is necessary to calibrate the algorithms on data. The calibration of the mistag of the *opposite side*

Parameter	Value
p_0	0.338 ± 0.012
p_1	1.01 ± 0.12
$\langle\omega_{\rm tag}\rangle$	0.339

Table 3.7.: The parameters resulting from the tagging calibration using $B^+ \to J/\psi\,K^+$ decays [69].

taggers can be performed using $B^+ \to J/\psi\,K^+$ decays [69]. Since charged $B^\pm$ mesons do not oscillate the correct production flavor can be determined from the kaon charge in the final state $J/\psi\,K^\pm$. The decay is said to be "self-tagging". In the calibration procedure the estimated mistag $\omega_{\rm tag}^{\rm est}$ given by the tagging algorithms is compared with the mistag probability $\omega_{\rm tag}$ measured in the calibration channel. A linear dependence between $\omega_{\rm tag}^{\rm est}$ and $\omega_{\rm tag}$ is observed in data. In [69] the parameterization

$$\omega_{\rm tag} \;=\; p_0 + p_1 \cdot (\omega_{\rm tag}^{\rm est} - \langle\omega_{\rm tag}^{\rm est}\rangle)$$

is used with the calibration parameters p_0 and p_1 and the mean estimated mistag $\langle\omega_{\rm tag}^{\rm est}\rangle$. This parameterization of the linear dependence was chosen to minimize correlations between p_0 and p_1. The calibration procedure indeed only finds a small negative correlation of -0.05 which can be neglected. For a calibrated data sample the calibration parameters are expected to be compatible with $p_0 = \langle\omega_{\rm tag}^{\rm est}\rangle$ and $p_1 = 1$. The values for p_0 and p_1 determined in [69] on ~ 11000 $B^+ \to J/\psi\,K^+$ events are given in Table 3.7. The effective tagging power for the *opposite side taggers* determined from $B^+ \to J/\psi\,K^+$ events in [69] is $(1.97 \pm 0.31)\,\%$.

To calibrate the *same side tagging* the decay $B_s^0 \to D_s^\mp \pi^\pm$ can be used[15]. While the flavor of the final state $D_s^\mp\,\pi^\pm$ can be extracted from the pion charge the neutral B_s^0 meson can oscillate before decay. To determine the B_s^0 production flavor a time dependent analysis is therefore necessary. The 2010 data contain too few $B_s^0 \to D_s^\mp\,\pi^\pm$ events to allow a meaningful calibration via a time-dependent analysis. Therefore this analysis only uses the *opposite side tagging* algorithms. With the data sample LHCb will collect in 2011 a calibration of the *same side tagging* algorithms will be possible and an increase of the sensitivity of the analysis by an estimated 30% to 50% is expected.

[15]The decay $B^+ \to J/\psi\,K^+$ can not be used since the $B^\pm$ does not contain an s-quark.

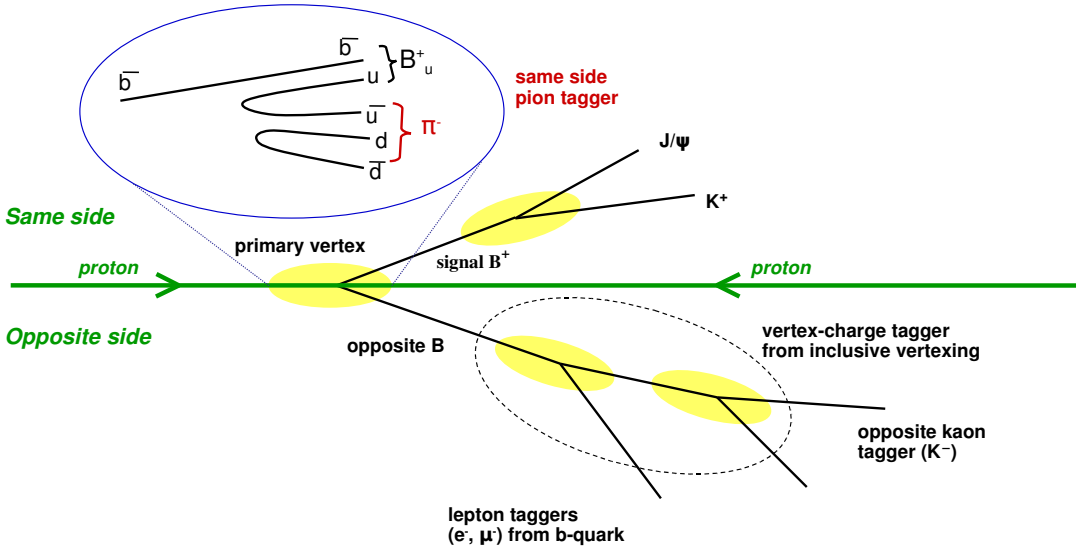


Figure 3.13.: Illustration of the flavor tagging methods. Sketched are both *same side* and *opposite side* flavor tagging. Figure from [69].

4. Technique to determine ϕ_s in the decay $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$

Fitting methods are an invaluable tool for the determination of physics parameters from a given data sample. These parameter estimation techniques fit a theoretical prediction which depends on the parameters of interest to the data. By varying the parameter values the fit determines the set of parameters for which the prediction best fits the data. Two classes of fitting methods exist, binned and unbinned techniques.

Binned techniques subdivide the data sample in bins of the measured quantities and compare the predicted number of events for every bin with the observed number of events. Binning leads to a loss of information since variations within one bin can not be resolved. The two main binned methods are the χ^2 fit (method of least squares) and the binned maximum likelihood technique.

Unbinned methods use the measured quantities directly and therefore do not suffer from a loss of information due to binning. They are however computationally more expensive and do not offer a simple measure for the goodness of fit. For the determination of ϕ_s this analysis uses an unbinned maximum likelihood fit. The method of unbinned maximum likelihood fits for parameter estimation is discussed in detail below. In addition the determination of errors or rather confidence intervals is discussed in this chapter. The last part of this chapter details the description of the data using probability density functions (PDFs). The signal component is modeled according to the differential decay rates of the signal decay given by theory, the background component is modeled empirically using the $\mathrm{B_s^0}$ mass sidebands.

4.1. The principle of maximum likelihood

4.1.1. General remarks

To determine the physics parameters of interest from the data an unbinned maximum likelihood fit is used. Several steps, listed below, are generally needed when performing this parameter estimation technique:

1. One or more quantities $\vec{x}$ are measured. Usually multiple measurements of the quantities $\vec{x}$ exist. These measurements are denoted as $\vec{x}_e$ with the index $e = 1 \ldots \mathrm{N}$.

2. There exists a theoretical prediction for the probability to measure $\vec{x}$ depending on one or more parameters $\vec{\lambda}$ that need to be determined. This prediction for the distribution of $\vec{x}$ is called probability density function (PDF) and is written as $\mathcal{P}(\vec{x};\vec{\lambda})$. It is normalized to one, $\int \mathcal{P}(\vec{x};\vec{\lambda})\mathrm{d}\vec{x} = 1$, because the probability for a measurement to return any possible result is given by 1.

3. The parameters $\vec{\lambda}$ are varied to maximize the probability $\mathcal{L} = \prod_{e=1}^{\mathrm{N}} \mathcal{P}(\vec{x}_e;\vec{\lambda})$ to draw the data set that has been measured in step 1 which constitutes a maximization

problem. This procedure finds the parameters for which it is most likely to obtain the drawn data set. For numerical reason one usually minimizes the negative logarithm of the likelihood

$$-\ln\mathcal{L} \;=\; -\sum_{e=1}^{\mathrm{N}} \ln\mathcal{P}(\vec{x}_e;\vec{\lambda})$$

and makes use of minimization algorithms like MINUIT [70].

For large data samples the maximum likelihood estimator is consistent, i. e. it converges to the true value $\vec{\lambda}$ in the limit $\mathrm{N}\to\infty$, and unbiased, which means that its expectation value is equal to the true $\vec{\lambda}$. It is also efficient, i. e. the variance of the maximum likelihood estimator is small[1]. However for small data samples the parameter estimation via unbinned maximum likelihood fits can be biased. For more details on parameter estimation using maximum likelihood fits see [71].

4.1.2. Resolutions of measured quantities

Possible complications arise due to the fact that the quantities $\vec{x}$ are usually only measured with a certain resolution $\vec{\sigma}$. This modifies the probability to observe an event with certain measured quantities and therefore needs to be included in the PDF $\mathcal{P}(\vec{x};\vec{\lambda})$. To describe resolutions, the original PDF has to be convoluted with a resolution model. Assuming a Gaussian resolution model for the measured quantity x_i this results in the convolution integral

$$\begin{aligned}\mathcal{P}(\vec{x};\vec{\lambda})\otimes\mathcal{R}(x_i;\sigma_i) &= \int \mathcal{P}(x_1,..,x_i',..,x_{\mathrm{M}};\vec{\lambda})\mathcal{R}(x_i-x_i';\sigma_i)\mathrm{d}x_i' \\ &= \int \mathcal{P}(x_1,..,x_i',..,x_{\mathrm{M}};\vec{\lambda})\frac{1}{\sqrt{2\pi}\sigma_i}e^{-\frac{(x_i-x_i')^2}{2\sigma_i^2}}\mathrm{d}x_i'.\end{aligned}$$

Section 4.5.6 discusses the treatment of the detector resolution effects in the analysis of $\mathrm{B_s^0}\to \mathrm{J/}\psi\,\phi$.

4.1.3. Acceptance effects

Furthermore, the distributions from theory can be modified by the detector acceptance and the event selection. This needs to be included in the PDF $\mathcal{P}(\vec{x};\vec{\lambda})$ via a so called acceptance function $\epsilon(\vec{x})$ which modifies the PDF further

$$\mathcal{P}(\vec{x};\vec{\lambda}) \;\to\; \epsilon(\vec{x})\mathcal{P}(\vec{x};\vec{\lambda}).$$

Special care needs to be taken for the correct normalization of the PDF when including acceptance effects, $\int\epsilon(\vec{x})\mathcal{P}(\vec{x};\vec{\lambda})\mathrm{d}\vec{x}=1$ must still hold. The implications for the $\mathrm{B_s^0}\to \mathrm{J/}\psi\,\phi$ analysis are discussed in section 4.5.7.

[1] Equal to the minimum variance bound $-1/\langle\frac{\mathrm{d}^2\ln\mathcal{L}}{\mathrm{d}\lambda^2}\rangle$

4.2. Confidence intervals

4.2.1. Neyman construction

As previously mentioned the PDF $\mathcal{P}(\vec{x};\vec{\lambda})$ denotes the probability to measure $\vec{x}$ given a parameter set $\vec{\lambda}$. It is essential to not confuse this with the probability $\mathcal{P}(\vec{\lambda};\vec{x})$ for a parameter set $\vec{\lambda}$ given a measurement $\vec{x}$. The Bayesian approach [72] to this problem is to use the Bayesian theorem $\mathcal{P}(A;B)\mathcal{P}(B) = \mathcal{P}(B;A)\mathcal{P}(A)$ to obtain $\mathcal{P}(\vec{\lambda};\vec{x})$. Note that this formula requires knowledge on the prior $\mathcal{P}(\vec{\lambda})$. With the help of the Bayesian theorem confidence intervals (or "credible intervals") for the parameter λ can easily be obtained. Assuming an interval $[\lambda_1, \lambda_2]$ it will contain the true value λ with probability

$$\alpha \;=\; \int_{\lambda_1}^{\lambda_2} \mathcal{P}(\lambda;x)\mathrm{d}\lambda = \frac{\int_{\lambda_1}^{\lambda_2} \mathcal{P}(x;\lambda)\mathcal{P}(\lambda)\mathrm{d}\lambda}{\int \mathcal{P}(x;\lambda')\mathcal{P}(\lambda')\mathrm{d}\lambda'}.$$

The frequentist approach does not claim that the true parameter is contained in the confidence interval determined from the data with a certain probability. Instead confidence intervals with coverage α obtained via frequentist methods can be interpreted as follows: If the experiment is repeated a large number of times in identical conditions a fraction α of the determined confidence intervals contain the true value λ. If the interval does contain the true value λ in a larger (smaller) fraction of all experiments the confidence interval is said to overcover (undercover). Intervals that overcover, also called "conservative" intervals, reduce the sensitivity of the experiment and should be avoided. Intervals that undercover are even more problematic since they result in an underestimated uncertainty of the parameter and therefore overestimate the sensitivity of the experiment.

Classical frequentist confidence intervals are obtained using the Neyman construction [73]. As illustration Figure 4.1a shows the construction of confidence belts for one measured quantity x and a parameter λ. For every possible parameter value λ intervals $[x_1, x_2]$ are selected such that $\int_{x_1}^{x_2} \mathcal{P}(x;\lambda)\mathrm{d}x = \alpha$. The confidence interval for the parameter λ can then be read of vertically and contains all values λ for which the measured quantity x_0 intersects with the intervals $[x_1, x_2]$ as illustrated by Figure 4.1b.

There is a certain freedom in what points x to include in the interval $[x_1, x_2]$. Upper limits on λ result from choosing intervals $[x_1, +\infty]$ with

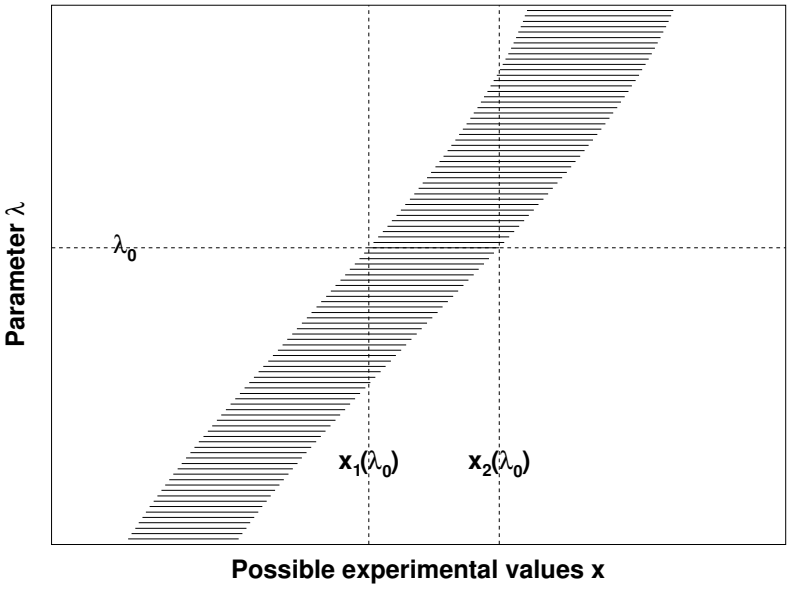
$$\int_{x_1}^{+\infty} \mathcal{P}(x;\lambda)\mathrm{d}x = \alpha$$

and lower limits from the interval choice $[-\infty, x_2]$

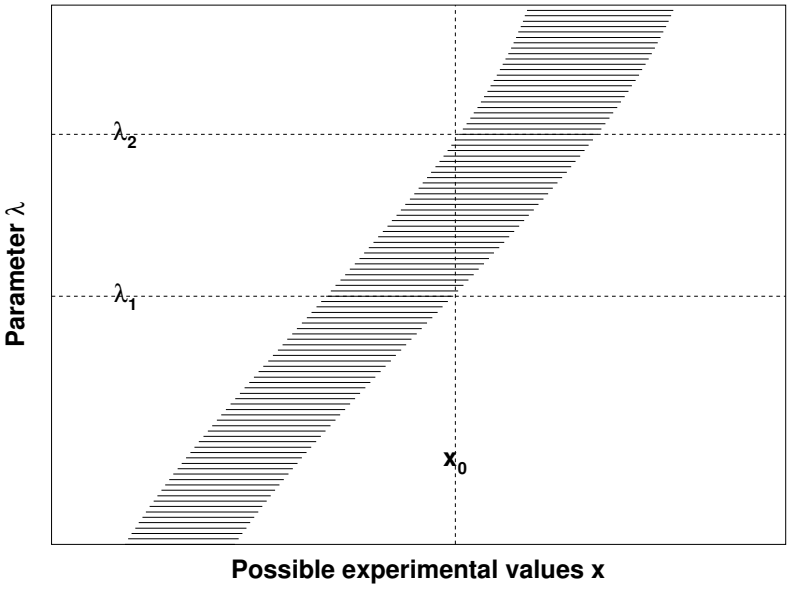
$$\int_{-\infty}^{x_2} \mathcal{P}(x;\lambda)\mathrm{d}x = \alpha.$$

Another possibility is to choose central confidence intervals $[x_1, x_2]$ with

$$\int_{-\infty}^{x_1} \mathcal{P}(x;\lambda)\mathrm{d}x = \int_{x_2}^{\infty} \mathcal{P}(x;\lambda)\mathrm{d}x = 1 - \frac{\alpha}{2}.$$



(a) Construction of the confidence belts



(b) Resulting confidence interval $[\lambda_1, \lambda_2]$

Figure 4.1.: Illustration of the Neyman construction of confidence belts. (a) The intervals are constructed horizontally including points x to an interval $[x_1, x_2]$ such that $\int_{x_1}^{x_2} \mathcal{P}(x;\lambda)\mathrm{d}x = \alpha$. This procedure is done for all possible λ. (b) The confidence interval for λ is read off vertically. For a measurement x_0 it includes all values of λ for which the interval $[x_1(\lambda), x_2(\lambda)]$ contains the measurement x_0.

Feldman and Cousins observed that choosing, depending on the data, whether to quote a one-sided limit or a two-sided interval (flip-flopping) can lead to undercoverage [74]. They proposed an ordering principle based on likelihood ratios which eliminates this pathology. This likelihood ratio ordering, known as the "Feldman-Cousins method", is the method of choice for the determination of confidence intervals for ϕ_s.

4.2.2. Feldman-Cousins method

The construction of confidence intervals depends on the choice which points to include in the confidence interval $[x_1, x_2]$ to reach the desired coverage. In [74] Feldman and Cousins proposed to add points x to the interval in descending order of their likelihood ratio

$$R = \frac{\mathcal{P}(x;\lambda)}{\mathcal{P}(x;\lambda_{\text{best}})}, \tag{4.1}$$

where λ_{best} is the parameter for which $\mathcal{P}(x;\lambda)$ is maximal[2]. The Feldman-Cousins procedure is computation intensive since it requires an additional maximization for every point x to determine λ_{best}. Confidence intervals can be created by iterating over all possible values of x and adding them to the confidence interval according to the ordering principle given above until the desired coverage is reached. The combination of all values of λ for which the measurement x_0 is contained in the confidence intervals $[x_1(\lambda), x_2(\lambda)]$ gives the frequentist confidence interval for λ.

A more efficient way to determine the confidence interval for λ is to use simulated events (see also section 5.1.1). For every fixed λ the PDF $\mathcal{P}(x;\lambda)$ is used to generate a large number N of simulated data samples x_i. The fraction of simulated data sets for which a larger likelihood ratio $R_{\text{toy,i}}$ is found than for data, i. e.

$$R_{\text{toy,i}} = \frac{\mathcal{P}(x_i;\lambda)}{\mathcal{P}(x_i;\lambda_{\text{best}})} \quad > \quad R_{\text{data}} = \frac{\mathcal{P}(x_0;\lambda)}{\mathcal{P}(x_0;\lambda_{\text{best}})},$$

gives the confidence level of the shortest interval containing the measurement x_0[3]. The determination of the confidence levels of intervals containing the measurement x_0 is repeated for a grid of all possible parameter values λ. The conjunction of all grid points in λ with a confidence level of 90% or smaller results in the 90% confidence interval for λ.

[2] For simplicity $\mathcal{P}(x;\lambda)$ is assumed to be one dimensional and to depend only on the parameter λ. In addition the data sample is assumed to consists of a single event, for multiple measurements x_e the ratio in equation 4.1 should read $R = \prod_e \mathcal{P}(x_e;\lambda) / \prod_e \mathcal{P}(x_e;\lambda_{\text{best}})$ instead. For a large number of events it is numerically beneficial to use the logarithm of the likelihood ratio $\ln R = \sum_e \ln \mathcal{P}(x_e;\lambda) - \sum_e \ln \mathcal{P}(x_e;\lambda_{\text{best}})$.

[3] To understand why confidence intervals can be constructed using simulated data drawn from the PDF it might be instructive to first use the simpler classical ordering rule to include points in the interval $[x_1, x_2]$ according to their probability given by the PDF $\mathcal{P}(x;\lambda)$. Assuming a given parameter λ a large number of simulated data sets x_i are generated. If their probabilities $\mathcal{P}(x_i;\lambda)$ are larger than $\mathcal{P}(x_0;\lambda)$ of the data in 90% of the cases x_0 belongs to the 90% confidence interval.

4.2.3. Likelihood ratio method

A common approximate method for the determination of confidence intervals is the likelihood ratio method, often also referred to as likelihood scan [75]. This method evaluates error intervals by searching for the points where the negative logarithm of the likelihood rises by a certain value.

Using a Gaussian PDF it is easy to derive the confidence levels which correspond to certain values of $-2\Delta \ln \mathcal{L}$. The integral

$$\int_{-a}^{+a} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \mathrm{d}x \;=\; \mathrm{erf}\left(\frac{a}{\sqrt{2}}\right)$$

gives the probability of a measurement to fall in the interval $[\lambda - a, \lambda + a]$ given parameter λ. Therefore the confidence interval for λ with confidence level $\mathrm{erf}\left(a/\sqrt{2}\right)$ is $[x_0 - a, x_0 + a]$, given a measurement x_0. The difference between the negative logarithm of the likelihood for the points $x_0 - a$ and $x_0 + a$ relative to the minimum of the negative logarithm of the likelihood is given by

$$\begin{aligned} -\Delta \ln \mathcal{L} &= \ln \mathcal{L}_{\max} - \ln \mathcal{L} \\ &= \frac{a^2}{2}. \end{aligned}$$

The points where $-2\Delta \ln \mathcal{L}$ rises by 1 therefore give a confidence interval with coverage $\mathrm{erf}\left(1/\sqrt{2}\right) = 68.3\%$. Table 4.1 gives the coverages corresponding to $-2\Delta \ln \mathcal{L}$ in one and two dimensions.

It should be noted that, in the case of low statistics, this assignment is generally not valid. Reference [75] gives a classification of PDFs for which the method yields confidence intervals with correct coverage. Since, due to the central limit theorem, every likelihood becomes Gaussian for large statistics the likelihood ratio method can also be useful for non-Gaussian PDFs in the asymptotic limit. In practice the likelihood is evaluated over a grid in the parameter(s) of interest. When the likelihood depends on additional parameters it needs to be minimized with respect to these parameters at each grid point. Unlike the method of parabolic errors discussed in section 4.2.4 the likelihood ratio method is able to return asymmetric error intervals[4].

4.2.4. Parabolic errors

Another frequently used approximate method for error estimation is the inverse second derivative of the negative logarithm of the likelihood with respect to the parameters of interest. The error on parameter λ in the limit of large statistics is given by

$$\sigma(\lambda) \;=\; \sqrt{\left(-\frac{\mathrm{d}^2 \ln \mathcal{L}}{\mathrm{d}\lambda^2}\right)^{-1}} \tag{4.2}$$

[4] The minimization package MINUIT uses the likelihood ratio technique when using the MINOS algorithm for error estimation.

$-2\Delta \ln \mathcal{L}$	C.L. 1D	C.L. 2D
1	68.3%	39.3%
4	95.4%	86.5%
9	99.7%	98.9%

(a)

C.L.	$-2\Delta \ln \mathcal{L}$ 1D	$-2\Delta \ln \mathcal{L}$ 2D
68.3%	1.00	2.30
90.0%	2.71	4.61
95.0%	3.84	5.99

(b)

Table 4.1.: (a) Confidence levels corresponding to different values of $-2\Delta \ln \mathcal{L}$ for Gaussian PDFs in one and two dimensions. For one dimension the equation relating C.L. with $-2\Delta \ln \mathcal{L}$ is C.L. = erf $\left(\sqrt{-2\Delta \ln \mathcal{L}}/\sqrt{2}\right)$. For two dimensions C.L. $= 1 - \exp\left(-(-2\Delta \ln \mathcal{L})/2\right)$ is found after a short calculation. (b) $-2\Delta \ln \mathcal{L}$ corresponding to different confidence levels.

which is derived for example in [71] using the central limit theorem. The minimization package MINUIT uses this error estimation technique when performing a HESSE step after the minimization. One obvious disadvantage of the method is that it will only yield symmetric error intervals. Furthermore it will generally give incorrect coverage for low statistics.

4.3. Goodness of fit

Unbinned maximum likelihood fits do not return a measure for the goodness of fit. While binned χ^2 fits return the minimized χ^2 which should be distributed according to a χ^2 distribution with the given numbers of degrees of freedom a fit result obtained from an unbinned maximum likelihood fit will not return a criterion of how well the data is described. One-dimensional projections of the data and the fitted PDF can give a first idea of the quality of the description of the data. However a good description of the data in the projections does not necessarily mean that the multidimensional distributions are well described. In [76] several methods to determine a measure of the goodness of fit are described.

The most well known technique is the binned χ^2 method. For this analysis however the χ^2 method is not useful since the PDF used is five-dimensional (proper time, reconstructed B_s^0 mass and the threee transversity angles). For 10 bins per dimension this results in 10^5 bins in total and only few (if any) entries in every single bin. The χ^2 method is however only applicable for large numbers of entries per bin.

Therefore a point-to-point dissimilarity test as described in section 3.3 of [76] is used to determine how well the data is described by the fitted PDF. To compare the distribution of the true PDF from which the n_D data events were drawn and the fitted PDF the fitted PDF is used to generate a large amount $n_{MC} \gg n_D$ of simulated events. A weighting function $\psi(\vec{x}, \vec{x}') = -\ln(|\vec{x} - \vec{x}'|)$ is used to calculate the distance of two events to be able to give a measure for the similarity of the data sample and the sample of events generated by the fitted PDF. The T statistics giving the correlation between the data

samples is determined according to

$$T = \frac{1}{n_D(n_D-1)}\sum_{i,j>1}^{n_D}\psi(\vec{x}_i^D,\vec{x}_j^D) - \frac{1}{n_D n_{MC}}\sum_{i,j}^{n_D,n_{MC}}\psi(\vec{x}_i^D,\vec{x}_j^{MC})\ [76].$$

To obtain the distribution of the test statistics T for the case that the PDF is described correctly a permutation test is performed. For the permutation test data and simulated events are combined. Random events of the combination are selected and taken as new data sample containing n_D events. The remaining events are taken as simulated events and the T statistic is calculated. This procedure is repeated multiple times and a distribution of T values is obtained. The p-value is given by the fraction of permutations for which the T value is larger than the T statistics obtained from the real data sample. The hypothesis that the fitted PDF is identical to the true PDF can be rejected at confidence level α if the p value is larger than $1-\alpha$.

4.4. Description of the decay $\mathrm{B_s^0} \to \mathrm{J/\psi\,\phi}$

The physics parameters of interest, in the following denoted as $\vec{\lambda}_{\mathrm{Phys}}$, are extracted using an unbinned maximum likelihood fit to the measured B_s mass, the proper time t and the transversity angles $\Omega = \{\cos\theta, \varphi, \cos\psi\}$. For the tagged analysis the tagging decision q ($q = +1$ if the tagging algorithm determines the $\mathrm{B_s^0}$ candidate to be a $\mathrm{B_s^0}$ at production, $q = -1$ if it is determined to be a $\bar{\mathrm{B}}_\mathrm{s}^0$ or $q = 0$ if no decision can be reached) and the per event mistag probability ω_{tag} are used additionally.

In addition to the physics parameters $\vec{\lambda}_{\mathrm{Phys}}$ more parameters are required to describe the data for the decay $\mathrm{B_s^0} \to \mathrm{J/\psi\,\phi}$ accurately. Resolution parameters $\vec{\lambda}_{\mathrm{Det}}$ are needed to model the limited detector resolution and background parameters $\vec{\lambda}_{\mathrm{Bkg}}$ to describe the background components. Both will be discussed in detail in the following sections.

The negative logarithm of the likelihood which needs to be minimized with respect to the parameters can be written as

$$-\ln\mathcal{L} = -\sum_{e=1}^{N}\ln\mathcal{P}(\{m,t,\Omega,q,\omega_{\mathrm{tag}}\}_e;\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}},\vec{\lambda}_{\mathrm{Bkg}}) \tag{4.3}$$

The PDF $\mathcal{P}$ is composed of the signal component $\mathcal{S}$ and the background component $\mathcal{B}$

$$\mathcal{P} = f_{\mathrm{sig}}\mathcal{S}(m,t,\Omega,q,\omega_{\mathrm{tag}};\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}}) + (1-f_{\mathrm{sig}})\mathcal{B}(m,t,\Omega,q,\omega_{\mathrm{tag}};\vec{\lambda}_{\mathrm{Det}},\vec{\lambda}_{\mathrm{Bkg}}),$$

with the signal fraction f_{sig}.

4.5. Signal description

The signal component $\mathcal{S}(m,t,\Omega,q,\omega_{\mathrm{tag}};\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}})$ separates in a proper time and angular dependent and a mass dependent part

$$\mathcal{S}(m,t,\Omega,q,\omega_{\mathrm{tag}};\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}}) = \mathcal{S}(m;\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}})\mathcal{S}(t,\Omega,q,\omega_{\mathrm{tag}};\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}}).$$

4.5.1. Mass dependence

The reconstructed B_s mass (see Figure 3.3) is used for separation of signal and background events. The signal component is modeled as a single Gaussian

$$\mathcal{S}(\mathrm{m};\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}}) \;=\; \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{m}}^{\mathrm{sig}}} e^{-\frac{(\mathrm{m}-\mathrm{m}_{\mathrm{B}_{\mathrm{s}}^0})^2}{2\sigma_{\mathrm{m}}^{\mathrm{sig}\,2}}} / \mathrm{C}_{\mathrm{m}}^{\mathrm{sig}}$$

with width $\sigma_{\mathrm{m}}^{\mathrm{sig}}$. For fits over a limited mass range the normalization factor $\mathrm{C}_{\mathrm{m}}^{\mathrm{sig}}$ is important to ensure the proper normalization $\int_{m_{\min}}^{m_{\max}} \mathcal{S}(m)\mathrm{dm} = 1$. It is given by

$$\mathrm{C}_{\mathrm{m}}^{\mathrm{sig}} \;=\; \left[\frac{1}{2}\mathrm{erf}\left(\frac{\mathrm{m}_{\max}}{\sqrt{2}\sigma_{\mathrm{m}}^{\mathrm{sig}}} - \frac{\mathrm{m}_{\mathrm{B}_{\mathrm{s}}^0}}{\sqrt{2}\sigma_{\mathrm{m}}^{\mathrm{sig}}}\right) - \frac{1}{2}\mathrm{erf}\left(\frac{\mathrm{m}_{\min}}{\sqrt{2}\sigma_{\mathrm{m}}^{\mathrm{sig}}} - \frac{\mathrm{m}_{\mathrm{B}_{\mathrm{s}}^0}}{\sqrt{2}\sigma_{\mathrm{m}}^{\mathrm{sig}}}\right)\right].$$

4.5.2. Angular and proper time dependence with tagging information

The time and angular dependent differential decay rates of $\mathrm{B}_{\mathrm{s}}^0$ and $\bar{\mathrm{B}}_{\mathrm{s}}^0$ mesons to the final state J/$\psi\,\phi$ are given by theory (see equations 1.26 and 1.27 in section 1.3.2). The differential decay rates depend on a priori unknown quantities, the physics parameters

$$\vec{\lambda}_{\mathrm{Phys}} \;=\; \left\{\phi_s, \Gamma_{\mathrm{s}}, \Delta\Gamma_{\mathrm{s}}, |A_0(0)|^2, |A_\perp(0)|^2, \delta_\parallel, \delta_\perp, \Delta m_{\mathrm{s}}\right\}. \tag{4.4}$$

The extraction of these parameters, in particular of ϕ_s and $\Delta\Gamma_{\mathrm{s}}$ is the main goal of this analysis. Please note that the decay amplitude $|A_\parallel(0)|^2$ does not appear as a free physics parameter since it is given by the normalization condition $|A_0(0)|^2 + \left|A_\parallel(0)\right|^2 + |A_\perp(0)|^2 = 1$ as discussed in section 1.3.2.

With perfect knowledge of the $\mathrm{B}_{\mathrm{s}}^0$ production flavor it would be possible to simply use the differential decay rates in equations 1.26 and 1.27 as PDFs after normalizing them properly. Due to the mistag probability ω_{tag} the proper time and angular dependent signal PDF needs to be modified. Accounting for ω_{tag} the signal component is given by

$$\begin{aligned}
\mathcal{S}(\mathrm{t},\Omega,\mathrm{q},\omega_{\mathrm{tag}};\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}}) \;=\; & \left[\epsilon(t,\Omega)\frac{1+\mathrm{q}(1-2\omega_{\mathrm{tag}})}{2}\frac{\mathrm{d}\Gamma(\mathrm{B}_{\mathrm{s}}^0 \to \mathrm{J}/\psi\phi)}{\mathrm{d}\Omega\,\mathrm{dt}}\right. \\
& \left. +\epsilon(t,\Omega)\frac{1-\mathrm{q}(1-2\omega_{\mathrm{tag}})}{2}\frac{\mathrm{d}\Gamma(\bar{\mathrm{B}}_{\mathrm{s}}^0 \to \mathrm{J}/\psi\phi)}{\mathrm{d}\Omega\,\mathrm{dt}}\right] / \mathrm{C}_{\Omega,\mathrm{t}}^{\mathrm{sig}} \qquad (4.5) \\
= \; & \left[\epsilon(t,\Omega)\frac{1+\mathrm{q}(1-2\omega_{\mathrm{tag}})}{2}\sum_{i=1}^{6} A_i\,(\mathrm{t})\,f_i\,(\cos\theta,\varphi,\cos\psi)\right. \\
& \left. +\epsilon(t,\Omega)\frac{1-\mathrm{q}(1-2\omega_{\mathrm{tag}})}{2}\sum_{i=1}^{6} \bar{A}_i\,(\mathrm{t})\,f_i\,(\cos\theta,\varphi,\cos\psi)\right] / \mathrm{C}_{\Omega,\mathrm{t}}^{\mathrm{sig}}.
\end{aligned}$$

where $A_i\,(\mathrm{t})$, $\bar{A}_i\,(\mathrm{t})$ and $f_i\,(\cos\theta,\varphi,\cos\psi)$ are defined as in section 1.3.2 and $\epsilon(t,\Omega)$ denotes a possible angular and proper time dependent acceptance. Note that for perfect

tagging with $\omega_{\rm tag} = 0$ the above would result in the differential decay rates for $\mathrm{B_s^0}$ or $\mathrm{\bar{B}_s^0}$ mentioned previously. The normalization factor $\mathrm{C}^{\rm sig}_{\Omega,\mathrm{t}}$ is given by

$$\begin{aligned}
\mathrm{C}^{\rm sig}_{\Omega,\mathrm{t}} &= \int \epsilon(t,\Omega)\left(\frac{1}{2}\frac{\mathrm{d}\Gamma(\mathrm{B_s^0}\to \mathrm{J}/\psi\phi)}{\mathrm{d}\Omega\,\mathrm{d}t} + \frac{1}{2}\frac{\mathrm{d}\Gamma(\mathrm{\bar{B}_s^0}\to \mathrm{J}/\psi\phi)}{\mathrm{d}\Omega\,\mathrm{d}t}\right)\mathrm{d}\Omega\,\mathrm{d}t \\
&= \frac{1}{2}\sum_{i=1}^{6}\int \epsilon(t,\Omega) A_i\,(\mathrm{t})\, f_i\,(\cos\theta,\varphi,\cos\psi)\,\mathrm{d}\Omega\,\mathrm{d}t \\
&+ \frac{1}{2}\sum_{i=1}^{6}\int \epsilon(t,\Omega) \bar{A}_i\,(\mathrm{t})\, f_i\,(\cos\theta,\varphi,\cos\psi)\,\mathrm{d}\Omega\,\mathrm{d}t \qquad (4.6)
\end{aligned}$$

which assumes equal number of produced $\mathrm{B_s^0}$ and $\mathrm{\bar{B}_s^0}$. The terms $f_i\,(\cos\theta,\varphi,\cos\psi)$ do not depend on the parameters to be determined in the fit. Therefore the angular part of the normalization integral,

$$\xi_i = \int \epsilon(t,\Omega) f_i\,(\cos\theta,\varphi,\cos\psi)\,\mathrm{d}\Omega,$$

can be calculated beforehand which is computationally beneficial. For no acceptance correction ($\epsilon(t,\Omega) = 1$) the angular integration can easily be performed using the $f_i\,(\cos\theta,\varphi,\cos\psi)$ terms found in Table 1.4. The resulting terms ξ_i are $\xi_1 = \xi_2 = \xi_3 = 1$ and $\xi_4 = \xi_5 = \xi_6 = 0$.

Per event mistag probability $\omega_{\rm tag}$

In this analysis the mistag probability $\omega_{\rm tag}$ is included by using the per event estimate from the tagging algorithms. In section 3.8 this was shown to result in a significant improvement in sensitivity for $\sin\phi_s$ compared to using global mistag rate. To account for different $\omega_{\rm tag}$ distributions of signal and background events additional terms $\mathcal{P}_{\rm sig}(\omega_{\rm tag})$ and $\mathcal{P}_{\rm bkg}(\omega_{\rm tag})$ need to be introduced which give the probability to find a certain $\omega_{\rm tag}$ for signal and background respectively:

$$\begin{aligned}
\mathcal{P} &= f_{\rm sig}\mathcal{S}(m,t,\Omega,q;\vec{\lambda}_{\rm Phys},\vec{\lambda}_{\rm Det})\mathcal{P}_{\rm sig}(\omega_{\rm tag}) \\
&\quad +(1-f_{\rm sig})\mathcal{B}(m,t,\Omega;\vec{\lambda}_{\rm Bkg},\vec{\lambda}_{\rm Det})\mathcal{P}_{\rm bkg}(\omega_{\rm tag}). \qquad (4.7)
\end{aligned}$$

Figure 4.2 shows the distribution of the tagging dilution $\mathcal{D}_{\rm tag} = 1-2\omega_{\rm tag}$ for data samples representative of signal and background events. The background sample was obtained by using events from the $\mathrm{B_s^0}$ mass sidebands $[5200, 5321.67]\,\mathrm{MeV} \cup [5411.67, 5550]\,\mathrm{MeV}$. The signal data sample was constructed by performing a sideband subtraction i. e. removing sidebands events from the signal region $[5321.67, 5411.67]\,\mathrm{MeV}$ weighted according to the background fraction. The fit uses the distributions in Figure 4.2 as probabilities $\mathcal{P}_{\rm sig}(\omega_{\rm tag})$ and $\mathcal{P}_{\rm bkg}(\omega_{\rm tag})$ to find a certain estimated mistag $\omega_{\rm tag}$ for signal and background events.

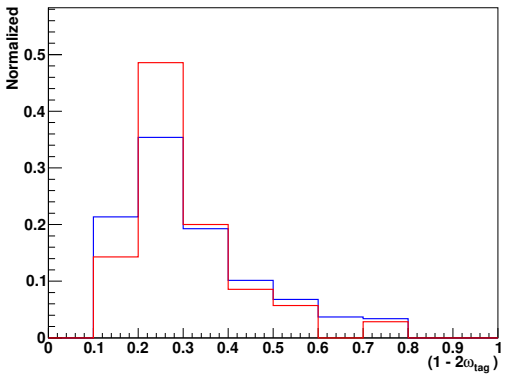


Figure 4.2.: Distribution of the tagging dilution $\mathcal{D}_{\text{tag}} = 1 - 2\omega_{\text{tag}}$ for sideband subtracted signal (blue) and sideband (red) data. The histograms are normalized to 1.

4.5.3. Angular and proper time dependence neglecting tagging information

Neglecting the information on the initial B_s^0 flavor it is still possible to determine a subset of the parameters in equation 4.4. Choosing $q = 0$ in equation 4.5 drops all terms in the PDF proportional to $\sin(\Delta m_\text{s} t)$ or $\cos(\Delta m_\text{s} t)$ resulting in

$$\begin{aligned}
\mathcal{S}(\text{t}, \Omega; \vec{\lambda}_{\text{Phys}}, \vec{\lambda}_{\text{Det}}) &= \epsilon(t,\Omega) e^{-\Gamma_\text{s} t} \cdot \\
&\Bigg[f_1(\Omega) \left|A_0(0)\right|^2 \left(\cosh\frac{\Delta\Gamma_\text{s}}{2}t - \cos\phi_s \sinh\frac{\Delta\Gamma_\text{s}}{2}t\right) \\
&+ f_2(\Omega) \left|A_\parallel(0)\right|^2 \left(\cosh\frac{\Delta\Gamma_\text{s}}{2}t - \cos\phi_s \sinh\frac{\Delta\Gamma_\text{s}}{2}t\right) \\
&+ f_3(\Omega) \left|A_\perp(0)\right|^2 \left(\cosh\frac{\Delta\Gamma_\text{s}}{2}t + \cos\phi_s \sinh\frac{\Delta\Gamma_\text{s}}{2}t\right) \\
&+ f_4(\Omega) \left|A_\parallel(0)\right| \left|A_\perp(0)\right| \left(-\cos(\delta_\perp - \delta_\parallel) \sin\phi_s \sinh\frac{\Delta\Gamma_\text{s}}{2}t\right) \\
&+ f_5(\Omega) \left|A_0(0)\right| \left|A_\parallel(0)\right| \cos\delta_\parallel \left(\cosh\frac{\Delta\Gamma_\text{s}}{2}t - \cos\phi_s \sinh\frac{\Delta\Gamma_\text{s}}{2}t\right) \\
&+ f_6(\Omega) \left|A_0(0)\right| \left|A_\perp(0)\right| \left(-\cos\delta_\perp \sin\phi_s \sinh\frac{\Delta\Gamma_\text{s}}{2}t\right)\Bigg] / \text{C}^{\text{sig}}_{\Omega,\text{t}}. \quad (4.8)
\end{aligned}$$

For the untagged analysis the set of observable physics parameters therefore becomes

$$\vec{\lambda}^{\text{Untagged}}_{\text{Phys}} = \left\{\phi_s, \Gamma_\text{s}, \Delta\Gamma_\text{s}, |A_0(0)|^2, |A_\perp(0)|^2, \delta_\parallel, \delta_\perp\right\}.$$

With the additional assumption of no time dependent CP violation, equivalent to $\phi_s = 0$ which is close to the value predicted by the Standard Model, the analysis further

simplifies. In this case the signal PDF reduces to

$$\begin{aligned}
\mathcal{S}(\mathrm{t},\Omega;\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}}) &= \epsilon(t,\Omega)e^{-\Gamma_{\mathrm{s}}t}\cdot \\
&\Bigg[f_1(\Omega)\left|A_0(0)\right|^2 e^{-\frac{\Delta\Gamma_{\mathrm{s}}}{2}t} + f_2(\Omega)\left|A_{\parallel}(0)\right|^2 e^{-\frac{\Delta\Gamma_{\mathrm{s}}}{2}t} + f_3(\Omega)\left|A_{\perp}(0)\right|^2 e^{+\frac{\Delta\Gamma_{\mathrm{s}}}{2}t} \\
&+ f_5(\Omega)\left|A_0(0)\right|\left|A_{\parallel}(0)\right|\cos\delta_{\parallel} e^{-\frac{\Delta\Gamma_{\mathrm{s}}}{2}t}\Bigg]/\mathrm{C}^{\mathrm{sig}}_{\Omega,\mathrm{t}} \qquad (4.9)\\
&= \epsilon(t,\Omega)e^{-\Gamma_{\mathrm{s}}t}\cdot \\
&\Bigg[e^{-\Gamma_{\mathrm{L}}t}\left(f_1(\Omega)\left|A_0(0)\right|^2 + f_2(\Omega)\left|A_{\parallel}(0)\right|^2 + f_5(\Omega)\left|A_0(0)\right|\left|A_{\parallel}(0)\right|\cos\delta_{\parallel}\right) \\
&+ e^{-\Gamma_{\mathrm{H}}t}\left(f_3(\Omega)\left|A_{\perp}(0)\right|^2\right)\Bigg]/\mathrm{C}^{\mathrm{sig}}_{\Omega,\mathrm{t}}.
\end{aligned}$$

This expression is no longer sensitive to the strong phase $\delta_{\perp}$. Only the dependence on the strong phase $\delta_{\parallel}$ or, more accurately, on $\cos\delta_{\parallel}$ remains. The above expression also illustrates that in the case of no CP-violation the CP even component (proportional to $|A_0(0)|^2$ and $\left|A_{\parallel}(0)\right|^2$) is equivalent to the light mass eigenstate whereas the time development of the CP odd component (proportional to $|A_{\perp}(0)|^2$) is given by the heavy mass eigenstate. The remaining physics parameters are

$$\vec{\lambda}^{\phi_s=0}_{\mathrm{Phys}} = \left\{\Gamma_{\mathrm{s}}, \Delta\Gamma_{\mathrm{s}}, |A_0(0)|^2, |A_{\perp}(0)|^2, \delta_{\parallel}\right\}.$$

4.5.4. Symmetry of the signal description

While the differential decay rates for $\mathrm{B_s^0}$ and $\bar{\mathrm{B}}_{\mathrm{s}}^0$ exhibit the two-fold symmetry given in equation 1.28 the untagged PDF in equation 4.8 is symmetric under the exchanges

$$\begin{array}{rcl} \phi_s & \to & -\phi_s \\ \Delta\Gamma_{\mathrm{s}} & \to & -\Delta\Gamma_{\mathrm{s}} \end{array} \quad \text{and} \quad \begin{array}{rcl} \phi_s & \to & -\phi_s \\ \delta_{\parallel} & \to & -\delta_{\parallel} \\ \delta_{\perp} & \to & \pi - \delta_{\perp}. \end{array}$$

resulting in a four-fold symmetry. It is therefore expected to find four minima for $-2\ln\mathcal{L}$ in the ϕ_s and $\Delta\Gamma_{\mathrm{s}}$ parameter space when the $\mathrm{B_s^0}$ production flavor is neglected.

4.5.5. S-wave contribution

To account for a possible S-wave contribution to the K^+K^- final state which was discussed in section 1.3.3 it is necessary to modify equations 4.5, 4.8 and 4.9. In equation 4.5 the sum over six terms needs to be expanded to include the four additional S-wave terms from section 1.3.3

$$\begin{aligned}
\mathcal{S}(\mathrm{t},\Omega,\mathrm{q},\omega_{\mathrm{tag}};\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}}) &= \Bigg[\epsilon(t,\Omega)\frac{1+\mathrm{q}(1-2\omega_{\mathrm{tag}})}{2}\sum_{i=1}^{10} A_i\left(\mathrm{t}\right) f_i\left(\cos\theta,\varphi,\cos\psi\right) \\
&\quad +\epsilon(t,\Omega)\frac{1-\mathrm{q}(1-2\omega_{\mathrm{tag}})}{2}\sum_{i=1}^{10} \bar{A}_i\left(\mathrm{t}\right) f_i\left(\cos\theta,\varphi,\cos\psi\right)\Bigg]/\mathrm{C}^{\mathrm{sig}}_{\Omega,\mathrm{t}}.
\end{aligned}$$

The normalization condition which needs to be respected by the amplitudes in this case is $|A_0(0)|^2 + |A_\parallel(0)|^2 + |A_\perp(0)|^2 + |A_\mathrm{s}(0)|^2 = 1$. The physics parameters are

$$\vec{\lambda}^{\text{S-wave}}_{\text{Phys}} = \left\{\phi_s, \Gamma_\mathrm{s}, \Delta\Gamma_\mathrm{s}, |A_0(0)|^2, |A_\perp(0)|^2, |A_\mathrm{s}(0)|^2, \delta_\parallel, \delta_\perp, \delta_\mathrm{s}, \Delta m_\mathrm{s}\right\}.$$

To determine the correct normalization $\mathrm{C}^{\mathrm{sig}}_{\Omega,\mathrm{t}}$ the additional S-wave terms need to be included as well.

The PDF for the untagged analysis can be derived from the above equation by setting $q = 0$. If the analysis is performed under the assumption of no time dependent CP violation, ϕ_s needs to be set to zero as well.

4.5.6. Proper time resolution

The proper time resolution of the detector needs to be modeled and considered in the PDF to correctly describe the data. For this analysis a triple Gaussian resolution model was chosen which describes the prompt background events well (see Figure 3.7a in section 3.6.1).

$$\begin{aligned} G(t;\vec{\lambda}_{\mathrm{Det}}) &= f^{\mathrm{sig}}_{\mathrm{t},1}\frac{1}{\sqrt{2\pi}\sigma^{\mathrm{sig}}_{\mathrm{t},1}}e^{-\frac{t^2}{2\sigma^{\mathrm{sig}\,2}_{\mathrm{t},1}}} \\ &+ (1-f^{\mathrm{sig}}_{\mathrm{t},1})f^{\mathrm{sig}}_{\mathrm{t},2}\frac{1}{\sqrt{2\pi}\sigma^{\mathrm{sig}}_{\mathrm{t},2}}e^{-\frac{t^2}{2\sigma^{\mathrm{sig}\,2}_{\mathrm{t},2}}} \\ &+ (1-f^{\mathrm{sig}}_{\mathrm{t},1})(1-f^{\mathrm{sig}}_{\mathrm{t},2})\frac{1}{\sqrt{2\pi}\sigma^{\mathrm{sig}}_{\mathrm{t},3}}e^{-\frac{t^2}{2\sigma^{\mathrm{sig}\,2}_{\mathrm{t},3}}}. \end{aligned} \tag{4.10}$$

The resolution parameters used in this analysis are given in Table 3.6. They were determined using prompt background events as discussed in section 3.6.1.

To correctly account for the proper time resolution all time dependent terms in the signal PDF need to be convoluted with this model. The convolution integrals that need to be calculated for this analysis are

$$\int G(t-t',\vec{\lambda}_{\mathrm{Det}})\Theta(t')e^{-\Gamma t'}\mathrm{d}t',$$
$$\int G(t-t',\vec{\lambda}_{\mathrm{Det}})\Theta(t')\sin(\Delta m_\mathrm{s} t')e^{-\Gamma t'}\mathrm{d}t', \text{ and}$$
$$\int G(t-t',\vec{\lambda}_{\mathrm{Det}})\Theta(t')\cos(\Delta m_\mathrm{s} t')e^{-\Gamma t'}\mathrm{d}t'.$$

All three integrations are performed analytically.

Per event proper time resolution

This analysis does not use the per event estimate of the proper time uncertainty since it was shown in section 3.6.1 to result in only a small increase in sensitivity on $\sin\phi_s$ compared to a global resolution model.

The algorithm developed in the scope of this thesis is however able to use the per event resolution σ_t which might be useful for future analyses. In this case scaling parameters are introduced to correct for possible underestimation of the proper time uncertainty provided by the global kinematic fit as discussed in section 3.6.1. Analogous to the case of the per event mistag probability additional terms must be included in the PDF describing the probability to find certain values of σ_t for signal and background events

$$\begin{aligned}\mathcal{P} &= f_{\text{sig}}\mathcal{S}(m,t,\sigma_t,\Omega,q,\omega_{\text{tag}};\vec{\lambda}_{\text{Phys}},\vec{\lambda}_{\text{Det}})\mathcal{P}_{\text{sig}}(\sigma_t)\\ &\quad +(1-f_{\text{sig}})\mathcal{B}(m,t,\sigma_t,\Omega;\vec{\lambda}_{\text{Bkg}},\vec{\lambda}_{\text{Det}})\mathcal{P}_{\text{bkg}}(\sigma_t).\end{aligned}$$

4.5.7. Angular and proper time acceptances

In equation 4.5 an additional factor $\epsilon(t,\Omega)$ was introduced to account for possible time and angular dependent acceptance effects. For this analysis it is assumed that the acceptance separates in the proper time and the transversity angles $\epsilon(t,\Omega) = \epsilon(t)\epsilon(\Omega)$. This assumption seems sensible since the two main reasons for acceptance effects, the lifetime bias due to IP cuts in the *lifetime biased* trigger and the angular acceptance due to detector geometry are unrelated.

Angular acceptance

As discussed in section 3.7 the detector geometry and selection criteria distort the angular distributions of the particles in the final state for the signal decay $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$. This effect has been studied using fully simulated signal events. Figure 3.10 shows the acceptance depending on the three transversity angles $\Omega = \{\cos\theta, \varphi, \cos\psi\}$.

There are multiple ways the angular acceptances can be treated when using the fitting code developed for this thesis:

- **Three-dimensional histogram:** The number of reconstructed and selected simulated signal events is filled into a three-dimensional histogram and divided by the number of expected events in the specific bin.

  $$\epsilon(\Omega) = \mathcal{H}(\Omega) = \frac{\mathrm{N_{Selected}}(\Omega)}{\mathrm{N_{Theory}}(\Omega)}.$$

  The angular integrals ξ_i are performed analytically before the fit.

  $$\xi_i = \sum_{\text{All bins j}} \int \mathcal{H}(\Omega) f_i(\Omega)\mathrm{d}\Omega_{\mathrm{j}}$$

  The advantage of this method is that it is relatively easy to implement. However binning effects due to the limited Monte Carlo statistics might occur. The limited statistics of the simulated signal events also limits the number of bins per dimension and thus the ability to resolve possible structures in the acceptance. Figure 4.3 shows the projection of the acceptance histogram.

- **Parameterization using Legendre Polynomials:** To reduce possible binning effects the acceptance effect can be described by an analytic parameterization. Legendre polynomials are chosen as basis functions for the description of the efficiency resulting in a sum of combinations of Legendre polynomials,

$$\epsilon(\Omega) \quad = \quad \sum_{jkl} c_{jkl} P_j(\cos\theta) P_k(\varphi) P_l(\cos\psi), \tag{4.11}$$

with the indices j, k and l denoting the degree of the Legendre polynomials. The coefficients c_{jkl} are determined in a binned maximum likelihood fit to a three-dimensional efficiency histogram. Only a limited number of combinations of Legendre polynomials is used[5]. The integrals ξ_i are determined analytically

$$\xi_i \quad = \quad \sum_{jkl} c_{jkl} \int f_i(\Omega) P_j(\cos\theta) P_k(\varphi) P_l(\cos\psi) \mathrm{d}\Omega$$

Figure 4.3 shows the acceptance parameterization using Legendre polynomials as solid blue line. The acceptance effect is well described by the parameterization.

- **Normalization weights:** The angular integrals ξ_i can also be determined in an unbinned fashion. The concept of this method is to go to the limit of infinitesimally small bin sizes in the angular integration

$$\begin{aligned} \xi_i \quad &= \quad \int f_i(\Omega) \epsilon(\Omega) \mathrm{d}\Omega \\ &= \quad \sum_{\text{bin j}} f_i(\Omega_j) \epsilon(\Omega_j) \mathrm{d}\Omega_j. \end{aligned}$$

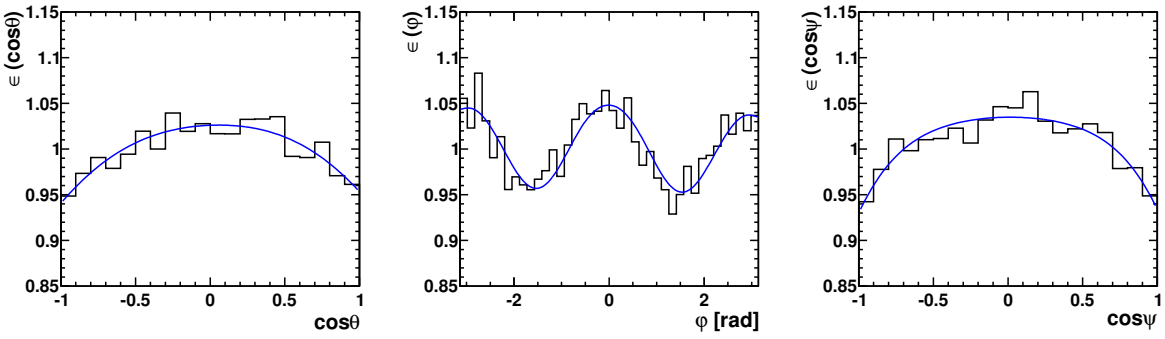
Using

$$\epsilon(\Omega_j) \quad = \quad \sum_{\text{events e in bin j}} \frac{1}{\mathcal{P}(\Omega_e)}$$

as efficiency and performing the limit results in

$$\xi_i \quad = \quad \sum_{\text{all events e}} \frac{f_i(\Omega_e)}{\int \mathcal{P}(t, \Omega_e) \mathrm{dt}}.$$

[5]In this analysis the maximum order of Legendre polynomials describing the acceptance is $j_{\max} = 4$ for $\cos\theta$, $k_{\max} = 8$ for φ and $l_{\max} = 4$ for $\cos\psi$. These orders were chosen to be able to accurately describe the shape of the acceptance effect in the angular projections. The sum $j + k + l$ is required to be smaller or equal to 12. This limit on the possible combinations of higher order polynomials makes the fit of the coefficients feasible. Within these requirements all combinations of even orders are allowed. Polynomials with odd order are only allowed in combinations of up to linear order. This last requirement stems from the fact that the acceptances in the channel $\mathrm{B_s^0} \to \mathrm{J/\psi}\,\phi$ are largely even.



(a) $\epsilon(\cos\theta)$ (b) $\epsilon(\varphi)$ (c) $\epsilon(\cos\psi)$

Figure 4.3.: Projections of the three-dimensional angular efficiency on the three transversity angles (a) $\cos\theta$, (b) φ and (c) $\cos\psi$. The figures show the efficiency histogram determined from Monte Carlo simulated signal events. In addition the analytic parameterization of the acceptance effect using Legendre polynomials is shown as solid blue line.

This method has the advantage of using the available simulated signal events optimally, however for both the fit and also plotting procedure the acceptance still needs to be modeled via a histogram or an analytic parameterization[6].

The integrals ξ_i gained from the three different methods are given in Table 4.2. Excellent agreement between the three different methods is observed. For this analysis it was decided to use the parameterization via Legendre polynomials as nominal method.

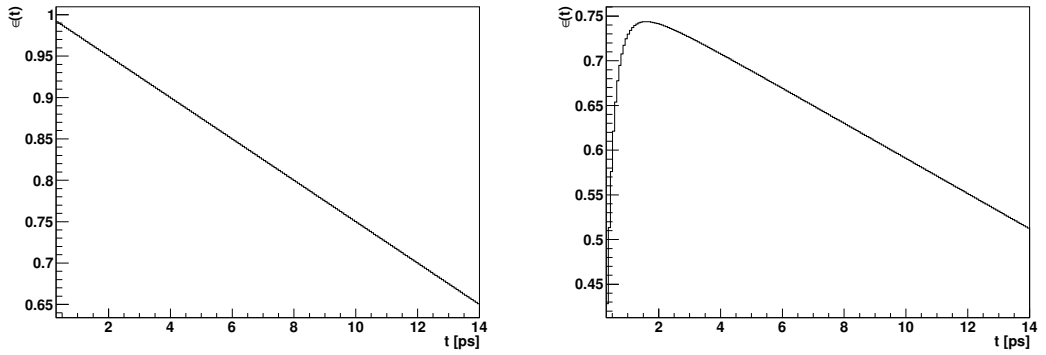
Proper time acceptance

As discussed in section 3.6.2 both the *lifetime unbiased* and the *lifetime biased* data samples exhibit proper time dependent efficiencies. They are accounted for in the fit using one-dimensional histograms $\epsilon(t)_{\text{Biased}} = \mathcal{H}(t)_{\text{Biased}}$ and $\epsilon(t)_{\text{Unbiased}} = \mathcal{H}(t)_{\text{Unbiased}}$ given in Figure 4.4. Both histograms consist of 200 bins over the timerange $[0.3, 14]$ ps and are filled using the analytic parameterizations given in section 3.6.2. To calculate the normalization constant $C^{\text{sig}}_{\Omega,\text{t}}$ of the signal component the necessary proper time integration is performed analytically for every bin.

[6]It has been suggested that the efficiency term in front of the signal PDF can be neglected since it does not depend on the parameters and therefore just constitutes an additional constant term in the sum of logarithms. However this is only possible if there are no background events present. This is not the case for the analysis of $B_s^0 \to J/\psi\,\phi$ unless the background events are subtracted via some statistical method (e. g. the sFit method [67]).

	Histogram	Parameterization	Unbinned
ξ_1	0.977	0.977	0.976
ξ_2	1.030	1.030	1.031
ξ_3	1.037	1.037	1.038
ξ_4	0.003	0.005	0.003
ξ_5	-0.001	0.000	-0.001
ξ_6	0.001	0.002	0.001
ξ_7	0.997	0.997	0.997
ξ_8	-0.003	-0.005	-0.004
ξ_9	-0.002	0.000	-0.002
ξ_{10}	0.001	0.002	0.001

Table 4.2.: Integrated terms $\xi_i = \int \epsilon(\Omega) f_i(\Omega) \mathrm{d}\Omega$ for different methods to describe the angular acceptance. The terms $\xi_7, .., \xi_{10}$ are only needed for an analysis including an S-wave component. Excellent agreement between the three different methods to describe the angular acceptance effect is observed.



(a) Lifetime unbiased (b) Lifetime biased

Figure 4.4.: The proper time acceptance histograms used for (a) the *lifetime unbiased* and (b) *lifetime biased* data sample. Each histogram consists of 200 bins in the proper time. Note that the absolute scale is arbitrary since it will be absorbed in the normalization constant $\mathrm{C}^{\mathrm{sig}}_{\Omega,\mathrm{t}}$.

4.6. Background description

Background events mimicking the signal decay in this analysis can be classified as two different types

- **Prompt background** either consists of the combination of four random tracks from the primary vertex or originates from true prompt J/ψ mesons from the primary vertex combined with two additional random tracks.
- **Longlived background** is either caused by misreconstructed non-signal $B \to J/\psi\, X$ decays or by other misidentified non-J/ψ longlived decays (e. g. B $\to$ hadrons).

The background in this analysis is therefore modeled as two components, a prompt component $\mathcal{B}_{\mathrm{Pr}}$ and a longlived component $\mathcal{B}_{\mathrm{LL}}$ according to

$$\mathcal{B}(m,t,\Omega;\vec{\lambda}_{\mathrm{Bkg}},\vec{\lambda}_{\mathrm{Det}}) \;=\; f_{\mathrm{Pr}}\mathcal{B}_{\mathrm{Pr}}(m,t,\Omega;\vec{\lambda}_{\mathrm{Bkg}},\vec{\lambda}_{\mathrm{Det}}) + (1-f_{\mathrm{Pr}})\mathcal{B}_{\mathrm{LL}}(m,t,\Omega;\vec{\lambda}_{\mathrm{Bkg}},\vec{\lambda}_{\mathrm{Det}}),$$

where the parameter f_{Pr} gives the fraction of prompt background events. The proper time cut $t > 0.3\,\mathrm{ps}$ rejects most of the prompt background, for the nominal analysis f_{Pr} is therefore set to $f_{\mathrm{Pr}} = 0$. The PDFs describing the background components are modeled empirically using events in the $\mathrm{B_s^0}$ mass sidebands defined as the regions $[5200, 5321.67]\,\mathrm{MeV} \cup [5411.67, 5550]\,\mathrm{MeV}$.

4.6.1. Longlived background

The longlived background component $\mathcal{B}_{\mathrm{LL}}$ factorizes in a mass, a time and an angular dependent part

$$\mathcal{B}_{\mathrm{LL}}(m,t,\Omega;\vec{\lambda}_{\mathrm{Bkg}},\vec{\lambda}_{\mathrm{Det}}) \;=\; \mathcal{B}_{\mathrm{LL}}(m;\vec{\lambda}_{\mathrm{Bkg}})\mathcal{B}_{\mathrm{LL}}(t;\vec{\lambda}_{\mathrm{Bkg}},\vec{\lambda}_{\mathrm{Det}})\mathcal{B}_{\mathrm{LL}}(\Omega).$$

Mass dependence

The mass dependence of the background components is modeled as a single exponential

$$\mathcal{B}_{\mathrm{LL}}(m;\vec{\lambda}_{\mathrm{Bkg}}) \;=\; e^{-\alpha_{\mathrm{m}}^{\mathrm{LL}} m}/\mathrm{C}_{\mathrm{m}}^{\mathrm{LL}}$$

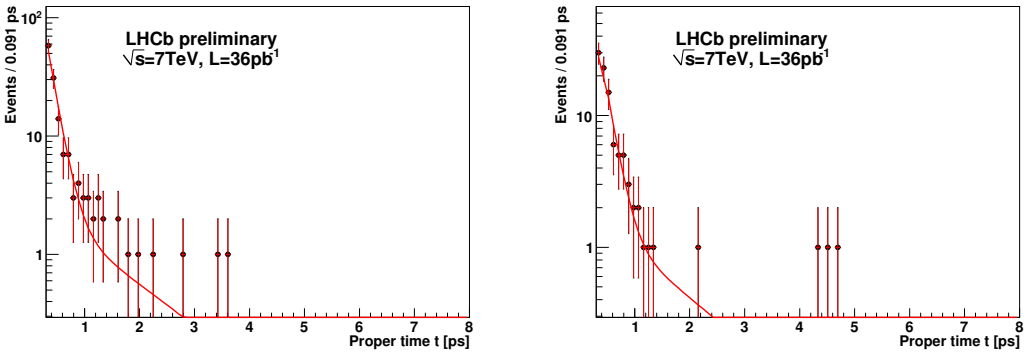
with normalization

$$\mathrm{C}_{\mathrm{m}}^{\mathrm{LL}} \;=\; \frac{1}{\alpha_{\mathrm{m}}^{\mathrm{LL}}}\left(e^{-\alpha_{\mathrm{m}}^{\mathrm{LL}} m_{\mathrm{min}}} - e^{-\alpha_{\mathrm{m}}^{\mathrm{LL}} m_{\mathrm{max}}}\right)$$

Proper time dependence

Two exponentials are used for the modeling of the proper time dependence of the longlived background

$$\mathcal{B}_{\mathrm{LL}}(t;\vec{\lambda}_{\mathrm{Bkg}},\vec{\lambda}_{\mathrm{Det}}) \;=\; \epsilon(t)\left[f_{\tau_1}^{\mathrm{LL}}\tau_1^{\mathrm{LL}} e^{-\frac{t}{\tau_1^{\mathrm{LL}}}} + (1-f_{\tau_1}^{\mathrm{LL}})\tau_2^{\mathrm{LL}} e^{-\frac{t}{\tau_2^{\mathrm{LL}}}}\right]. \qquad (4.12)$$



(a) Lifetime unbiased (b) Lifetime biased

Figure 4.5.: The proper time distribution of the longlived background for both the (a) *lifetime unbiased* and (b) *lifetime biased* data sample. Overlayed is the fitted distribution. Only the B_s^0 mass sidebands $[5200, 5321.67]\,\text{MeV} \cup [5411.67, 5550]\,\text{MeV}$ are shown.

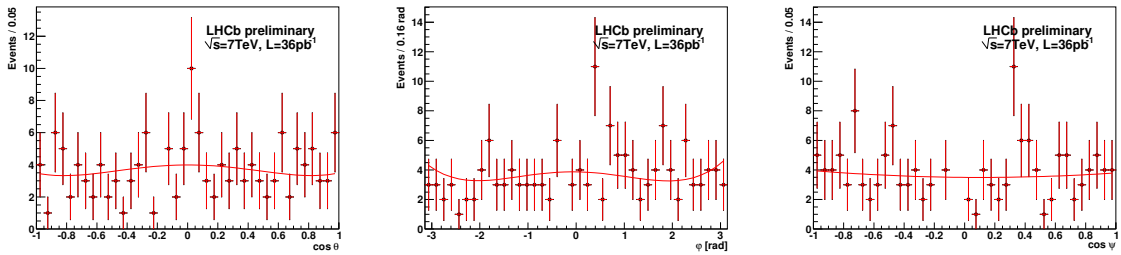
The same time dependent efficiencies $\epsilon(t)$ applied to the signal are also applied to the longlived background component. Furthermore the time dependent part is convoluted with the same triple Gaussian resolution model as the signal component. Figure 4.5 shows the proper time distribution for events in the B_s^0 mass sideband after the full selection for both the *lifetime biased* and *lifetime unbiased* data sample. Overlayed is the fitted proper time distribution for the background component. For both *lifetime unbiased* and *biased* data sample the same parameters τ_1^{LL} and τ_2^{LL} are used. Good agreement with data is observed.

Angular dependence

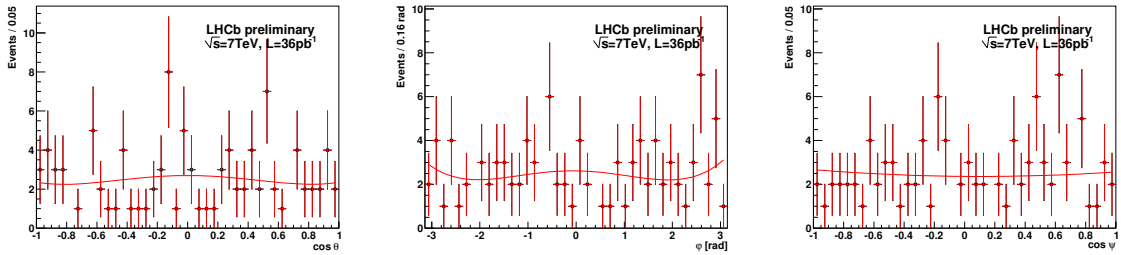
The angular dependence of the background component is modeled using a three-dimensional analytic parameterization. Correlations between the transversity angles $\cos\theta$, φ and $\cos\psi$ can therefore be described. Legendre polynomials are used as basis functions resulting in the parameterization

$$\mathcal{B}_{\text{LL}}(\Omega) \quad = \quad \sum_{ijk} \alpha_{ijk} P_i(\cos\theta) P_j(\varphi) P_k(\cos\psi),$$

which is similar to the analytic description of the angular acceptance effect given in equation 4.11. The coefficients α_{ijk} are determined from a fit to the B_s^0 mass sidebands, including also prompt events to obtain a sufficiently large number of background events. The description uses all combinations of Legendre polynomials up to order $i = 4$ for $P_i(\cos\theta)$, $j = 4$ for $P_j(\varphi)$ and $k = 2$ for $P_i(\cos\psi)$. An overlay of the fitted background distribution over the sideband data projected on the transversity angles is given in



(a) Lifetime unbiased



(b) Lifetime biased

Figure 4.6.: The angular distributions of the longlived background for both the (a) *lifetime unbiased* and (b) *lifetime biased* data sample. Overlayed are the fitted distributions. Only the B_s^0 mass sidebands $[5200, 5321.67]\,\text{MeV} \cup [5411.67, 5550]\,\text{MeV}$ are shown. The full selection is applied.

Figure 4.6. The parameterization and the data from the B_s^0 mass sidebands agree within the very limited statistics.

4.6.2. Prompt background

For the nominal fit it was decided to require $t > 0.3\,\text{ps}$ to reject the prompt background events. However section 6.1.1 gives the results of a cross-check study where the fit was performed over the full lifetime range. In this case the prompt background is described according to the prescription given below. As the longlived background the prompt background separates in a mass, a proper time and an angular dependent part

$$\mathcal{B}_{\text{Pr}}(m, t, \Omega; \vec{\lambda}_{\text{Bkg}}, \vec{\lambda}_{\text{Det}}) \;=\; \mathcal{B}_{\text{Pr}}(m; \vec{\lambda}_{\text{Bkg}})\mathcal{B}_{\text{Pr}}(t; \vec{\lambda}_{\text{Bkg}}, \vec{\lambda}_{\text{Det}})\mathcal{B}_{\text{Pr}}(\Omega).$$

Proper time dependence

The proper time distribution of the prompt background is essential for the determination of the proper time resolution model as was discussed in section 3.6.1. The proper time distribution of the prompt background is described as a triple Gaussian (see equation 4.5.6) using the same parameters as the signal resolution model.

Mass dependence

Analogous to the longlived background the mass dependence of the prompt background is modeled as a single exponential

$$\mathcal{B}_{\mathrm{Pr}}(m;\vec{\lambda}_{\mathrm{Bkg}}) \;=\; e^{-\alpha_{\mathrm{m}}^{\mathrm{Pr}} m}/\mathrm{C}_{\mathrm{m}}^{\mathrm{Pr}}$$

Angular dependence

The angular dependence uses the same parameterization with Legendre polynomials as the longlived background

$$\mathcal{B}_{\mathrm{Pr}}(\Omega) \;=\; \sum_{ijk} \alpha_{ijk} P_i(\cos\theta) P_j(\varphi) P_k(\cos\psi).$$

The configuration and determination of the coefficients α_{ijk} is identical to the description of the longlived background.

4.7. Details of the fitting method

4.7.1. Implementation details

The fitting algorithm for the extraction of ϕ_s which is presented in this thesis is implemented using the C++ programming language and the MINUIT minimization package [70]. In the minimization procedure precision is preferred over speed[7]. After the minimization the errors are estimated using the Gaussian approximation given in equation 4.2[8]. It is also possible to use the likelihood ratio method discussed in section 4.2.3 to obtain asymmetric error intervals[9], but this comes at a prize for the performance.

When evaluating the likelihood the sum of logarithms of the probabilities will soon be very large compared to the logarithm of the probability for a single event. This is problematic since adding numbers of very different sizes inevitably leads to a loss of numerical precision. To solve this problem the mean logarithm for the probability of an event is calculated before minimization at the parameter start values according to $-2\ln\mathcal{P}_{\mathrm{mean}} = -2\sum_{e=1}^{\mathrm{N}} \ln\mathcal{P}/\mathrm{N}$. For every $\mathrm{B_s^0}$ candidate this number is subtracted from the logarithm of its probability. Since the resulting terms in the sum fluctuate around 0 the total sum will be of the same size as the current term. This noticeably improves the numerical accuracy of the fit.

4.7.2. Constrained parameters

Parameters which have been measured precisely in previous experiments or in different decay channels can be constrained to the measured value with a Gaussian error. In this

[7] The fitting algorithm uses the MIGRAD minimization technique with strategy 2 which allows MINUIT to use more function calls to improve the reliability of the minimization.
[8] The HESSE algorithm is used for the error estimation.
[9] MINOS is used in this case.

case the negative logarithm of the likelihood is modified to

$$\begin{aligned} -2\ln\mathcal{L} &= -2\sum_{e=1}^{\mathrm{N}} \ln\mathcal{P}(\{\mathrm{t},\Omega,\mathrm{m},\mathrm{q},\omega_{\mathrm{tag}}\}_e;\vec{\lambda}) \\ &+ \sum_{i}^{\text{all constraints}} \frac{(\lambda_i-\lambda_i^{\mathrm{PDG}})^2}{\sigma_i^{\mathrm{PDG}\,2}} \end{aligned}$$

where λ_i denotes the parameters which have been measured before to be λ_i^{PDG}. The error of this measurement is given by σ_i^{PDG}.

In this analysis a Gaussian constraint is applied on the mixing frequency Δm_{s} which was determined previously by the CDF collaboration to be 17.77 ± 0.12 [28]. The measurement performed by the LHCb collaboration on data taken in 2010 results in $\Delta m_{\mathrm{s}} = (17.63 \pm 0.11_{\mathrm{stat.}} \pm 0.03_{\mathrm{syst.}})\,\mathrm{ps}^{-1}$ [29] which is well compatible with this value. Additionally the flavor tagging parameters p_0 and p_1, see section 3.8.1, are constrained to the values that were found from the calibration procedure, $p_0 = 0.338 \pm 0.012$ and $p_1 = 1.01 \pm 0.12$. This correctly propagates the systematic uncertainty of the tagging calibration to the result.

4.7.3. Simultaneous fit

For a maximal sensitivity on the physics parameters all available signal candidates need to be used in the analysis. For the $\mathrm{B_s^0}$ production flavor dependent analysis to extract ϕ_s it is therefore necessary to perform a simultaneous fit to both the *lifetime biased* and the *lifetime unbiased* data set. The complete likelihood in this case is given by the sum

$$\begin{aligned} -2\ln\mathcal{L} &= -2\ln\mathcal{L}_{\mathrm{Unbiased}} - 2\ln\mathcal{L}_{\mathrm{Biased}} \\ &= -2\sum_{\mathrm{Unbiased\ e}} \ln\mathcal{P}(\{m,t,\Omega,q,\omega_{\mathrm{tag}}\}_e;\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}},\vec{\lambda}_{\mathrm{Bkg}}^{\mathrm{Unbiased}}) \\ &\quad -2\sum_{\mathrm{Biased\ e}} \ln\mathcal{P}(\{m,t,\Omega,q,\omega_{\mathrm{tag}}\}_e;\vec{\lambda}_{\mathrm{Phys}},\vec{\lambda}_{\mathrm{Det}},\vec{\lambda}_{\mathrm{Bkg}}^{\mathrm{Biased}}) \end{aligned}$$

The physics parameters $\vec{\lambda}_{\mathrm{Phys}}$ and resolution parameters $\vec{\lambda}_{\mathrm{Det}}$ used in the fit are identical for both data samples. Besides the different proper time acceptances the data samples also show different levels of background pollution. For this reason two different signal fractions, $f_{\mathrm{sig}}^{\mathrm{Biased}}$ and $f_{\mathrm{sig}}^{\mathrm{Unbiased}}$ are used. In addition the parameters that describe the distribution of the longlived background events in the reconstructed $\mathrm{B_s^0}$ mass, $\alpha_{\mathrm{m,Biased}}^{\mathrm{LL}}$ and $\alpha_{\mathrm{m,Unbiased}}^{\mathrm{LL}}$, are allowed to vary independently. All other parameters are identical.

5. Validation of the fitting algorithm and performance estimates for ϕ_s

Before ϕ_s can be determined on data the fitting method has to be validated. The fitter validation ensures that both the parameter values as well as the parameter errors are determined correctly.

Large amounts of simulated events need to be generated for the fit validation. A fast generation method is therefore implemented using the rejection sampling method [77]. Both the principle of the rejection sampling method and its implementation used to generate simulated events are discussed in the first part of this chapter. Using the simulated events provided by the fast generation the fitting method is validated in the second part of this chapter. In addition to the validation of the fitting method the expected sensitivity to the extracted parameters is determined. The last part of the chapter discusses the validation of the fitting method using fully simulated signal events. Using fully simulated signal events cross-checks the implementation of the signal decay[1]. In addition it ensures that the fitting method takes reconstruction effects correctly into account.

5.1. Fast event simulation

5.1.1. The rejection sampling method

To thoroughly test the fitting method more simulated events are needed than can be fully simulated in a reasonable timeframe. Of the order of 10^8 simulated events are used for a typical fit validation study. To generate this amount of simulated events the rejection sampling method [77] is used[2].

The principle of the rejection sampling method is illustrated in Figure 5.1. The goal of the rejection sampling method is to generate events x_i according to a given PDF $\mathcal{P}(x)$ (given as blue solid line in Figure 5.1) which can have a complicated dependence on x. The generation procedure using the rejection sampling method proceeds as follows: A random number r_1 is drawn from the uniform distribution $]0,1]$. Scaling r_1 to the available range of x results in a value for x_i. A second random number r_2 is drawn from the uniform distribution $]0,1]$. The event x_i is accepted if

$$r_2 \cdot \mathrm{C}_{\max} < \mathcal{P}(x_i), \tag{5.1}$$

[1] Strictly speaking it checks that the description of the signal decay in the EvtGen event generator and the unbinned maximum likelihood fit is identical. This is an important cross-check since the fast event generation and the unbinned maximum likelihood fit use largely the same code for the description of the signal decay.

[2] Rejection sampling is also known as the accept-reject method

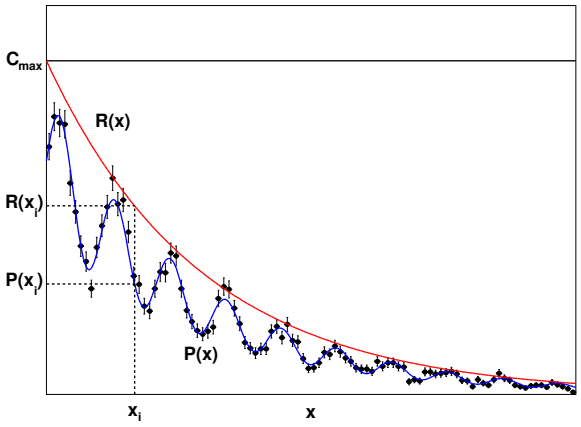


Figure 5.1.: Illustration of the rejection sampling method. The goal of the rejection sampling method is to generate events x_i according to the PDF $\mathcal{P}(x)$ which is denoted by the blue curve. The function $\mathcal{R}(x)$ given in red is an enveloping function for $\mathcal{P}(x)$ with $\mathcal{R}(x) > \mathcal{P}(x)$ for all x. The black dots denote 10000 events generated according to the PDF $\mathcal{P}(x)$ using the rejection sampling method.

otherwise it is rejected. The constant $\mathrm{C}_{\mathrm{max}}$ in equation 5.1 needs to fulfill the condition $\mathrm{C}_{\mathrm{max}} > \mathcal{P}(x)$ for all values of x.

It is apparent that this method of event generation is not very efficient if the PDF $\mathcal{P}(x)$ is quickly falling in x. An example would be the PDF used in Figure 5.1 which has an exponentially falling behavior $\propto \exp(-x/\tau_x)$. The algorithm would produce events uniformly in x but events with large x_i are only rarely accepted. In this case using an exponential enveloping function $\mathcal{R}(x)$ with the property $\mathcal{R}(x) > \mathcal{P}(x)$ for all x can be much more efficient. In Figure 5.1 an exponential enveloping function fulfilling this requirement is given as solid red line. Assuming events x_i can be generated according to the distribution $\mathcal{R}(x)$[3] and let r_2 be a random number from the interval $]0, 1]$ events x_i are accepted if

$$r_2 \cdot \mathcal{R}(x_i) < \mathcal{P}(x_i).$$

The number of operations needed for the generation of a certain number of events scales with the integral of the enveloping function. Using an exponential rather than a constant $\mathrm{C}_{\mathrm{max}}$ is therefore computationally beneficial, especially for large ranges x.

[3]To generate events according to an exponential distribution the simple formula $x_i = -\tau_x \cdot \ln(r_1)$ can be used where r_1 is a random number drawn from the uniform distribution in $]0, 1]$.

5.1.2. Fast simulation of signal events

The rejection sampling method described in the last section is used for the generation of simulated signal events. The PDF $\mathcal{P}(\mathrm{t}, \Omega)$ is given by the differential decay rates 1.26 and 1.27 in section 1.3.2. Events are generated according to flat distributions in the transversity angles and an exponential distribution in the proper time. The enveloping function $\mathcal{R}(\mathrm{t}, \Omega)$ according to which the events are generated is therefore given by

$$\mathcal{R}(\mathrm{t}, \Omega) \quad = \quad ce^{-\alpha t}$$

with $\alpha < \Gamma_{\mathrm{L}}$ and a sufficiently large constant c such that $\mathcal{R}(\mathrm{t}, \Omega) > \mathcal{P}(\mathrm{t}, \Omega)$ for all possible values of the proper time and the transversity angles. Candidates drawn from the distribution $\mathcal{R}(\mathrm{t}, \Omega)$ are accepted if

$$r_2 \mathcal{R}(\mathrm{t}, \Omega) \quad < \quad \mathcal{P}(\mathrm{t}, \Omega) = \epsilon(t, \Omega) \frac{\mathrm{d}\Gamma(\mathrm{B_s^0}/\bar{\mathrm{B}}_\mathrm{s}^0 \to \mathrm{J}/\psi\, \phi)}{\mathrm{d}t \mathrm{d}\Omega}.$$

where r_2 denotes a random number drawn from the uniform distribution in $]0, 1]$ and $\epsilon(t, \Omega)$ describes a possible proper time and angular dependent acceptance effect. Using an exponential enveloping function instead of a flat distribution in the proper time is a factor ~ 20 more efficient.

The fast generation tries to reproduce the data taken with LHCb as accurately as possible. Resolution effects are included in the generation by smearing the proper time according to the resolution model after the signal candidate is accepted. The triple Gaussian resolution model extracted from the prompt background events in 2010 data is used (Table 3.6 gives the resolution parameters). Depending on whether *lifetime biased* or *unbiased* events are generated the appropriate proper time dependent efficiency is used. The angular acceptance is determined from fully simulated signal events as described in section 3.7. Tagging decision and mistag probability ω_{tag} are generated per event. Both tagging quantities are generated according to the tagging efficiency and mistag distribution found from sideband subtracted 2010 data (Figure 4.2). The mass dependence of the signal decay is generated according to a single Gaussian, the mass resolution is taken from the fit to the 2010 data.

5.1.3. Fast simulation of background events

Generating events according to the background description in section 4.6 is simpler than to generate signal events because it is assumed that the proper time and angular dependent components of the background PDF factorize. The proper time distribution of the longlived background is generated according to the double exponential in equation 4.12. Background parameters, i. e. the background fractions and lifetimes are taken from a fit to the 2010 data. Proper time resolution for background events is implemented using the same triple Gaussian resolution model as for the signal component. The transversity angles Ω are generated according to the analytic parameterization described in section 4.6.1. Angular and proper time projections of the background PDFs used for the generation of background events are shown in Figures 4.5 and 4.6. The same proper

time dependent acceptance as for signal events is applied. Tagging of background events is simulated according to the tagging efficiency of sideband events in 2010 data. The tagging decision of the background events is chosen randomly (50 : 50). The mistag probability ω_{tag} for background events is drawn from the mistag distribution extracted from the sidebands of the 2010 data (Figure 4.2).

5.2. Fit validation and sensitivity studies

The fast generation of signal and background events described in section 5.1.2 and 5.1.3 respectively allows the simulation of large event samples. Multiple data samples of simulated events are generated and subsequently fit[4]. The parameters determined in the fit of these simulated data samples should be compatible with the parameter values used in the event generation. For a quantitative test that the fit determines both the parameter values and errors correctly pull distributions are used. The pull for parameter p is defined as

$$\text{pull} \;=\; \frac{p_{\text{fit}} - p_{\text{true}}}{\sigma_{\text{fit}}(p)}.$$

Combining the pull values for parameter p resulting from the fit of all simulated data samples gives the pull distribution for parameter p. The pull distribution should be described by a Gaussian centered around 0[5] with a width of 1[6]. To extract the mean and the width of pull distributions unbinned maximum likelihood fits are performed using single Gaussians.

The simulated events can also be used to determine the sensitivity to the extracted parameters, i. e. the expected statistical uncertainty for a certain data sample size. The sensitivity can be determined from the widths of the distributions of the fitted parameter values. Unbinned maximum likelihood fits of single Gaussians are used to extract the sensitivity of the fit to the parameters.

5.2.1. Configuration of the fit validation studies

The fit validation studies performed for this thesis consist of 1000 samples of simulated events each corresponding to an integrated luminosity of $2\,\text{fb}^{-1}$ which is the integrated luminosity for one nominal year of data-taking at LHCb. The physics parameters used in the generation of the simulated signal events are given in Table 5.1.

There are two approaches to the analysis of the decay $\text{B}_\text{s}^0 \to \text{J}/\psi\,\phi$. Neglecting information on the B_s^0 production flavor an analysis to extract $\Delta\Gamma_\text{s}$ under the assumption of no time dependent CP violation ($\phi_s = 0$) can be performed. The flavor blind analysis using the 2010 data is described in chapter 6. It is performed using only the *lifetime unbiased* events. The validation of the flavor blind fit is discussed in section 5.2.2.

[4]This is often referred to as a *toy study*.
[5]The parameter estimation is said to be unbiased.
[6]The estimation of the parameter uncertainty is correct.

Parameter	Generator value
ϕ_s[rad]	0 (no CPV) or -0.0363 (SM) or -0.70 (NP)
Γ_s	$0.6793\,\mathrm{ps}^{-1}$
$\Delta\Gamma_s$	$0.0600\,\mathrm{ps}^{-1}$
$\lvert A_0(0)\rvert^2$	0.60
$\lvert A_\parallel(0)\rvert^2$	0.24
$\lvert A_\perp(0)\rvert^2$	0.16
$\delta_\parallel$	2.50 rad
$\delta_\perp$	-0.17 rad
Δm_s	$17.8\,\mathrm{ps}^{-1}$

Table 5.1.: Physics parameters used for the fast simulation of signal events.

To extract both $\Delta\Gamma_s$ and ϕ_s a B_s^0 production flavor dependent analysis is necessary. The B_s^0 production flavor dependent analysis on data is described in chapter 7. It uses the maximum available information from the 2010 data by performing a simultaneous fit to both the *lifetime biased* and *unbiased* data sample. For the validation of the B_s^0 production flavor dependent fit two different scenarios are studied. The first scenario is the Standard Model scenario which uses $\phi_s = -0.0363$ rad in the generation of simulated signal events. The second scenario assumes a large value of ϕ_s caused by possible physics processes beyond the Standard Model. This is the New Physics scenario which uses $\phi_s = -0.70$ rad in the event generation. In section 5.2.3 the tagged fit is validated for both scenarios.

5.2.2. Validation of the flavor blind fit

Table 5.2 summarizes the result of the flavor blind analysis under the assumption $\phi_s = 0$[7]. Figure 5.3 gives the distribution of the fitted physics parameters. The pull distributions are given in Figure 5.2. The pull distributions are well described by a single Gaussian. The mean values of the fitted single Gaussians are within 3σ of 0, their respective widths are within 3σ of 1. This confirms that the flavor blind fit is able to determine the parameter values and errors correctly. For simulated data corresponding to $2\,\mathrm{fb}^{-1}$ a sensitivity of $0.021\,\mathrm{ps}^{-1}$ to $\Delta\Gamma_s$ is found. The mean correlations between the physics parameters are given in Table 5.3. As expected from the PDF given by equation 4.9 moderate to large correlations are seen between $\Delta\Gamma_s$, Γ_s and the transversity amplitudes.

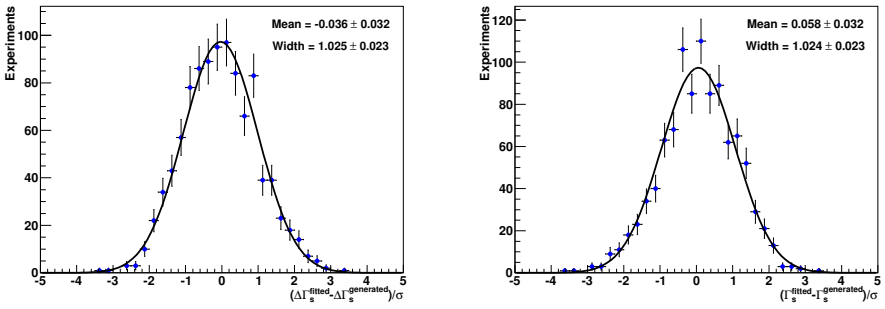
[7]For the influence of values $\phi_s \neq 0$ on the untagged fit result see section 6.3.

	Sensitivity	Pull mean	Pull width
Γ_s	0.008 ± 0.000	0.06 ± 0.03	1.02 ± 0.02
$\Delta\Gamma_s$	0.021 ± 0.000	-0.04 ± 0.03	1.03 ± 0.02
$\lvert A_\perp \rvert^2$	0.008 ± 0.000	0.05 ± 0.03	1.02 ± 0.02
$\lvert A_0 \rvert^2$	0.005 ± 0.000	-0.07 ± 0.03	1.02 ± 0.02
$\cos\delta_\parallel$	0.027 ± 0.001	-0.03 ± 0.03	1.05 ± 0.02

Table 5.2.: Results of the flavor blind *toy study* under the assumption of no CP-violation ($\phi_s = 0\,\text{rad}$). 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\text{fb}^{-1}$.

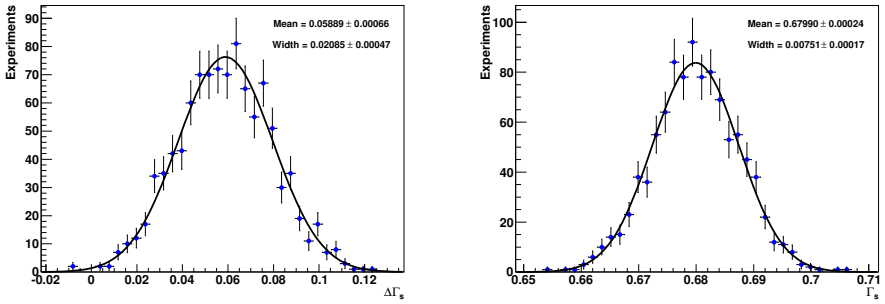
	Γ_s	$\Delta\Gamma_s$	$\lvert A_\perp \rvert^2$	$\lvert A_0 \rvert^2$	$\cos\delta_\parallel$
Γ_s	1	**−0.82**	**0.61**	**−0.56**	−0.05
$\Delta\Gamma_s$		1	**−0.69**	**0.66**	0.04
$\lvert A_\perp \rvert^2$			1	**−0.58**	−0.34
$\lvert A_0 \rvert^2$				1	-
$\cos\delta_\parallel$					1

Table 5.3.: Mean correlations of the physics parameters for the flavor blind configuration. Absolute correlations larger than 0.5 are given in bold, correlations smaller than 0.005 are omitted. Since the correlations depend on the data sample the RMS of the correlations is given in Table A.1.



(a) $\Delta\Gamma_s$ pull distribution (b) Γ_s pull distribution

Figure 5.2.: Pull distributions for the flavor blind *toy study* with no CP violation. 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\text{fb}^{-1}$. The remaining distributions are given in Figure A.1.



(a) $\Delta\Gamma_s$ fitted parameter values (b) Γ_s fitted parameter values

Figure 5.3.: Distributions of the fitted parameter values for the flavor blind *toy study* with no CP violation. 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\mathrm{fb}^{-1}$. The remaining distributions are given in Figure A.2.

5.2.3. Validation of the B_s^0 production flavor dependent fit

Standard Model scenario

Table 5.4 gives the results of the study using the Standard Model configuration with $\phi_s = -0.0363\,\mathrm{rad}$. Figure 5.5 shows the distributions of the fitted physics parameters. They are well described by a single Gaussian. The widths of the Gaussians give the sensitivity values quoted in Table 5.4. The expected sensitivity on ϕ_s and $\Delta\Gamma_s$ for one nominal year of data is $0.097\,\mathrm{rad}$ and $0.017\,\mathrm{ps}^{-1}$ respectively. The pull distributions are shown in Figure 5.4. They are well described by single Gaussians. The means of the pull distributions are within 3σ of 0, their respective widths are within 3σ of 1. This confirms that the unbinned maximum likelihood fit is unbiased and estimates the parameter uncertainties correctly.

Table 5.5 gives the mean correlations of the extracted physics parameters. As is expected, large correlations between Γ_s and $\Delta\Gamma_s$ as well as the transversity amplitudes are observed. These correlations are an inherent property of the probability density functions (equations 1.26 and 1.27) for the description of the signal decay. In the Standard-Model like configuration ϕ_s exhibits only very small correlations with the other physics parameter.

New Physics scenario

The distribution of the fitted parameter values for the New Physics scenario is given in Figure 5.7. The pull distributions are given in Figure 5.6. Both fitted parameter values and pull distributions can be described by single Gaussians. The pull means are within 3σ of 0, their widths within 3σ of 1 which validates the fit for the case of a large CP violating phase ϕ_s.

	Sensitivity	Pull mean	Pull width
Γ_s	0.006 ± 0.000	0.05 ± 0.03	1.02 ± 0.02
$\Delta\Gamma_\mathrm{s}$	0.018 ± 0.000	-0.02 ± 0.03	1.02 ± 0.02
ϕ_s	0.095 ± 0.002	0.04 ± 0.03	1.03 ± 0.02
$\lvert A_\perp \rvert^2$	0.007 ± 0.000	0.05 ± 0.03	1.00 ± 0.02
$\lvert A_0 \rvert^2$	0.005 ± 0.000	-0.04 ± 0.03	0.99 ± 0.02
$\delta_\parallel$	0.036 ± 0.001	0.03 ± 0.03	0.99 ± 0.02
$\delta_\perp$	0.310 ± 0.007	-0.04 ± 0.03	1.00 ± 0.02
Δm_s	0.103 ± 0.002	-0.01 ± 0.03	1.04 ± 0.02

Table 5.4.: Results of the *toy study* for the Standard Model scenario with $\phi_s = -0.0363\,\mathrm{rad}$. 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\mathrm{fb}^{-1}$.

The largest difference between tables 5.6 and 5.4 is the sensitivity on Δm_s which increases from $0.107\,\mathrm{ps}^{-1}$ to $0.057\,\mathrm{ps}^{-1}$. This change can be understood by the fact that the main sensitivity on Δm_s is given by terms proportional to $\sin\phi_s \sin(\Delta m_\mathrm{s} t)$. If the amplitude of this oscillation given by $\sin\phi_s$ is small it is naturally difficult to resolve the oscillation frequency. The sensitivity on ϕ_s is slightly reduced to 0.11 rad. The reason for this becomes apparent when comparing the correlations in Table 5.7 with the correlation coefficients in the Standard Model study in Table 5.5. The larger correlations allow the fit to absorb changes of the likelihood originating from different values for ϕ_s by adjusting other parameters thus leading to a larger uncertainty for ϕ_s.

	Γ_s	$\Delta\Gamma_s$	ϕ_s	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel$	$\delta_\perp$	Δm_s
Γ_s	1	**−0.82**	−0.04	**0.61**	**−0.56**	0.05	-	-
$\Delta\Gamma_s$		1	0.03	**−0.70**	**0.67**	−0.04	-	-
ϕ_s			1	−0.03	0.03	-	0.01	-
$\lvert A_\perp\rvert^2$				1	**−0.59**	0.33	0.02	-
$\lvert A_0\rvert^2$					1	-	-	-
$\delta_\parallel$						1	0.06	-
$\delta_\perp$							1	**0.59**
Δm_s								1

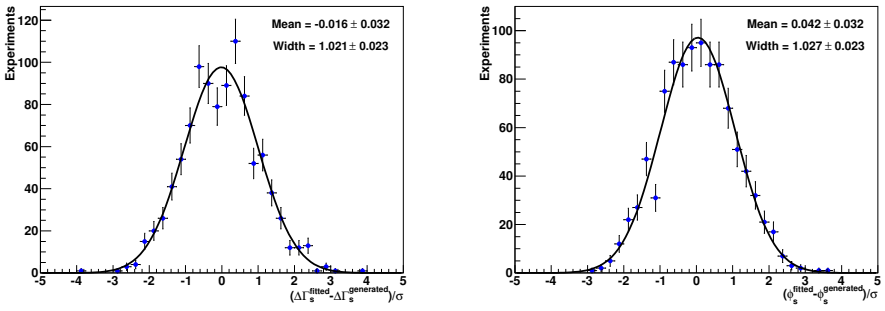
Table 5.5.: Mean correlations of the physics parameters for the Standard Model scenario with $\phi_s = -0.0363\,\text{rad}$. Absolute correlations larger than 0.5 are given in bold, correlations smaller than 0.005 are omitted. Since the correlations depend on the data sample the RMS of the correlations is given in Table A.1b.

	Sensitivity	Pull mean	Pull width
Γ_s	0.005 ± 0.000	0.09 ± 0.03	1.03 ± 0.02
$\Delta\Gamma_s$	0.016 ± 0.000	-0.09 ± 0.03	0.97 ± 0.02
ϕ_s	0.114 ± 0.003	0.07 ± 0.03	0.98 ± 0.02
$\lvert A_\perp\rvert^2$	0.006 ± 0.000	0.06 ± 0.03	1.00 ± 0.02
$\lvert A_0\rvert^2$	0.004 ± 0.000	-0.04 ± 0.03	1.02 ± 0.02
$\delta_\parallel$	0.036 ± 0.001	0.09 ± 0.03	0.99 ± 0.02
$\delta_\perp$	0.278 ± 0.006	-0.06 ± 0.03	1.03 ± 0.02
Δm_s	0.057 ± 0.001	-0.05 ± 0.03	1.02 ± 0.02

Table 5.6.: Results of the *toy study* for the New Physics scenario with $\phi_s = -0.70\,\text{rad}$. 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\text{fb}^{-1}$.

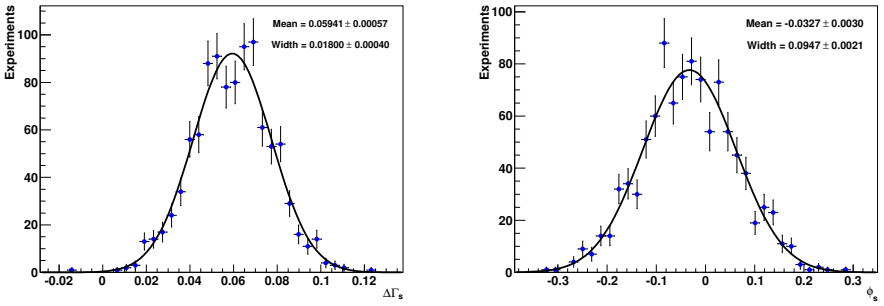
	Γ_s	$\Delta\Gamma_s$	ϕ_s	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel$	$\delta_\perp$	Δm_s
Γ_s	1	**−0.66**	−0.39	0.49	−0.44	0.05	−0.01	0.01
$\Delta\Gamma_s$		1	0.13	**−0.58**	**0.56**	−0.04	0.01	-
ϕ_s			1	−0.28	0.25	−0.02	0.01	−0.02
$\lvert A_\perp\rvert^2$				1	**−0.52**	0.36	0.02	-
$\lvert A_0\rvert^2$					1	-	0.01	-
$\delta_\parallel$						1	0.08	-
$\delta_\perp$							1	0.29
Δm_s								1

Table 5.7.: Mean correlations of the physics parameters for the New Physics scenario with $\phi_s = -0.70\,\text{rad}$. Absolute correlations larger than 0.5 are given in bold, correlations smaller than 0.005 are omitted. Since the correlations depend on the data sample the RMS of the correlations is given in Table A.1c.



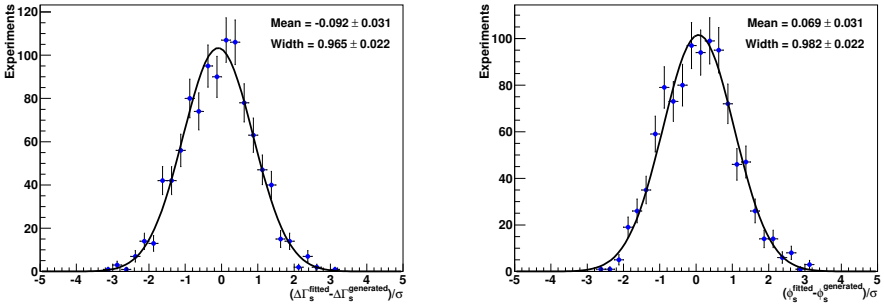
(a) $\Delta\Gamma_s$ pull distribution (b) ϕ_s pull distribution

Figure 5.4.: Pull distributions for the Standard Model scenario ($\phi_s = -0.0363\,\text{rad}$). 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\text{fb}^{-1}$. The remaining distributions are given in Figure A.3.



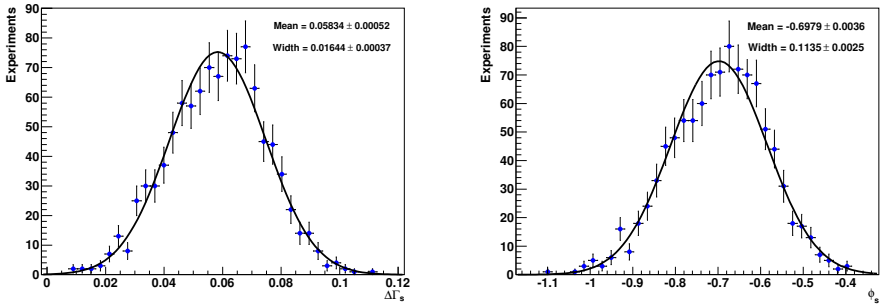
(a) $\Delta\Gamma_s$ fitted parameter values (b) ϕ_s fitted parameter values

Figure 5.5.: Distributions of the fitted parameter values for the Standard Model scenario ($\phi_s = -0.0363\,\text{rad}$). 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\text{fb}^{-1}$. The remaining distributions are given in Figure A.4.



(a) $\Delta\Gamma_s$ pull distribution (b) ϕ_s pull distribution

Figure 5.6.: Pull distributions for the New Physics scenario ($\phi_s = -0.70\,\text{rad}$). 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\text{fb}^{-1}$. The remaining distributions are given in Figure A.5.



(a) $\Delta\Gamma_s$ fitted parameter values (b) ϕ_s fitted parameter values

Figure 5.7.: Distributions of the fitted parameter values for the New Physics scenario ($\phi_s = -0.70\,\text{rad}$). 1000 *toy* data sets are generated and fit. Each data set corresponds to an integrated luminosity of $2\,\text{fb}^{-1}$. The remaining distributions are given in Figure A.6.

5.3. Fit of fully simulated signal events

Fully simulated signal events generated as described in section 2.5 are used to validate the fitting method. If the description of the signal decay is implemented correctly the maximum likelihood fit should extract the parameters given in Table 2.3 that were used in the generation of the simulated events. After the full selection is applied the simulated data sample contains 178k signal candidates. This amount of simulated signal events corresponds to roughly $10\,\mathrm{fb}^{-1}$ of real data[8]. The fit to the fully simulated events is performed using a triple Gaussian proper time resolution model. The resolution parameters are extracted from a fit to the known $t_{\mathrm{true}} - t_{\mathrm{rec}}$ distribution of the simulated events and given in Table 5.8. Only the information provided by the opposite side taggers is used in the fit.

Table 5.9 gives the resulting fitted parameters and their deviations from the values used in the generation of the simulated events. All parameters with the exception of Γ_{s} are reproduced with less than one standard deviation ($1\sigma_{\mathrm{stat}}$) from the value used in the generation. This confirms the correct implementation of the signal description used in the unbinned maximum likelihood fit.

The deviation of Γ_{s} from the generated value is $-2.9\sigma_{stat}$. As discussed in section 3.6.2 a proper time dependent acceptance correction $\epsilon(t) = 1 + \beta t$ was found originating from a reconstruction effect. In [63] the correction factor β was determined to be $\beta = -0.025$. This value of β slightly overestimates the acceptance effect. In section 6.2.2 and 7.2.3 it will be shown that the lifetime acceptance $\propto (1 + \beta t)$ only affects the value of Γ_{s}. The other physics parameters are not affected. Using the size of β as systematic uncertainty fully accounts for the overcorrection. Therefore this thesis uses the value $\beta = -0.025 \pm 0.025$ for the proper time acceptance effect.

Figure 5.8 shows the projection of the fitted signal PDF on the reconstructed $\mathrm{B}_{\mathrm{s}}^{0}$ mass, proper time and the transversity angles. Angular acceptance effects were taken into account in the fit, which is apparent when comparing the angular description of Figure 3.9 with Figure 5.8. The description of the angular distributions in the latter case is excellent[9]. The correlation matrix for the fit is given in Table 5.10.

[8] Extrapolated from 628 lifetime unbiased signal events in the 2010 data corresponding to $36\,\mathrm{pb}^{-1}$.

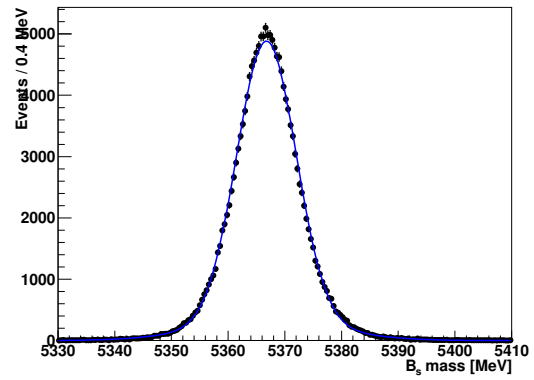
[9] Note that the same simulated events were used to determine the angular acceptances and to perform the fit.

Parameter	Value
$f_{\mathrm{t},1}^{\mathrm{sig}}$	0.314
$f_{\mathrm{t},2}^{\mathrm{sig}}$	0.821
$\sigma_{\mathrm{t},1}^{\mathrm{sig}}$	23.6 fs
$\sigma_{\mathrm{t},2}^{\mathrm{sig}}$	39.4 fs
$\sigma_{\mathrm{t},3}^{\mathrm{sig}}$	71.9 fs

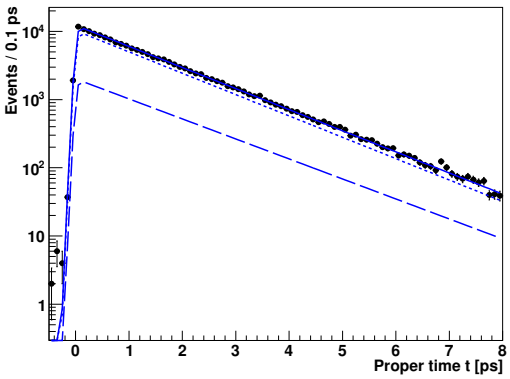
Table 5.8.: Resolution parameters determined using fully simulated signal events.

Parameter	Result	Deviation from generation [σ_{stat}]
Γ_{s}	$(0.6718 \pm 0.0026)\,\mathrm{ps}^{-1}$	-2.9
$\Delta\Gamma_{\mathrm{s}}$	$(0.0603 \pm 0.0089)\,\mathrm{ps}^{-1}$	0.0
ϕ_{s}	$(-0.728 \pm 0.045)\,\mathrm{rad}$	-0.6
$\lvert A_{\perp}\rvert^2$	0.1580 ± 0.0027	-0.7
$\lvert A_0\rvert^2$	0.5991 ± 0.0019	-0.5
$\delta_{\parallel}$	$(2.500 \pm 0.017)\,\mathrm{rad}$	-0.0
$\delta_{\perp}$	$(-0.27 \pm 0.11)\,\mathrm{rad}$	-0.9
Δm_{s}	$(17.803 \pm 0.028)\,\mathrm{ps}^{-1}$	0.1

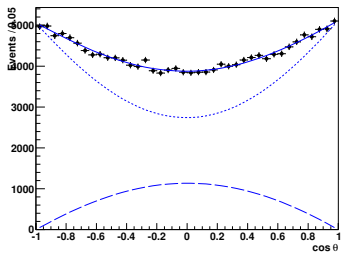
Table 5.9.: Physics parameters extracted from a data sample of fully simulated signal events corresponding to ca. $\sim 10\,\mathrm{fb}^{-1}$. All physics parameters with the exception of Γ_{s} are extracted with a deviation of less than $1\sigma_{\mathrm{stat.}}$ with respect to the values used in the event generation.



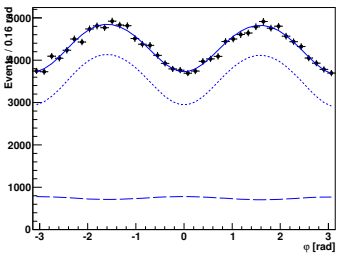
(a) Reconstructed B_s^0 mass



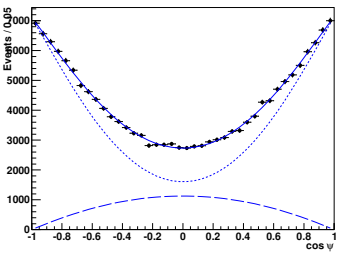
(b) Proper time t



(c) Transversity angle $\cos\theta$



(d) Transversity angle φ



(e) Transversity angle $\cos\psi$

Figure 5.8.: Projection of the fitted PDF on (a) the reconstructed B_s^0 mass, (b) the proper time and (c)-(e) the transversity angles. The dotted line gives the CP-even part of the signal PDF, the dashed part denotes the CP-odd part. A good description of the simulated events is observed. Compared with Figure 3.9 the improvement of the description of the transversity angles due to the angular dependent acceptance correction is clearly visible.

	Γ_s	$\Delta\Gamma_s$	ϕ_s	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel$	$\delta_\perp$	Δm_s	m_{B_s}	$f^{sig}_{m,1}$	$\sigma^{sig}_{m,1}$	$\sigma^{sig}_{m,2}$
Γ_s	1.00	**-0.68**	-0.33	0.44	-0.37	0.05	-0.01	-	-	-	-	-
$\Delta\Gamma_s$		1.00	0.08	**-0.51**	0.48	-0.04	0.01	0.01	-	-	-	-
ϕ_s			1.00	-0.24	0.18	-0.05	0.02	-0.03	-	-	-	-
$\lvert A_\perp\rvert^2$				1.00	-0.44	0.38	0.04	0.01	-	-	-	-
$\lvert A_0\rvert^2$					1.00	-0.01	0.01	-	-	-	-	-
$\delta_\parallel$						1.00	0.10	-	-	-	-	-
$\delta_\perp$							1.00	0.20	-	-	-	-
Δm_s								1.00	-	-	-	-
m_{B_s}									1.00	-	-	-0.01
$f^{sig}_{m,1}$										1.00	**0.82**	**0.87**
$\sigma^{sig}_{m,1}$											1.00	**0.68**
$\sigma^{sig}_{m,2}$												1.00

Table 5.10.: Parameter correlations extracted from a data sample of fully simulated signal events corresponding to ca. $\sim 10\,\mathrm{fb}^{-1}$. Correlations larger than 0.5 are given in bold.

6. Flavor blind analysis to determine $\Delta\Gamma_s$

The analysis of the decay $B_s^0 \to J/\psi\,\phi$ is conducted in two subsequent steps. First an analysis that does not use information about the B_s^0 production flavor (flavor blind or untagged analysis) is performed. This was necessary since the calibration of the flavor tagging algorithms on channels such as $B^+ \to J/\psi\,K^+$ is challenging, thus flavor tagging information was not immediately usable after data taking. The untagged analysis of the decay $B_s^0 \to J/\psi\,\phi$ is described in this chapter. Only the *lifetime unbiased* data sample is used for this study. It will become apparent that the untagged analysis can not constrain ϕ_s using the amount of data taken by LHCb in 2010. Instead the primary aim of the untagged analysis is to determine $\Delta\Gamma_s$ under the assumption of no CP violation, i. e. $\phi_s = 0$.

The second step of the analysis of the decay $B_s^0 \to J/\psi\,\phi$ is to perform a B_s^0 production flavor dependent fit which is able to constrain ϕ_s. This tagged analysis is described in chapter 7.

6.1. Determination of $\Delta\Gamma_s$ under the assumption of no CP violation

The PDF used for the parameter estimation ignoring the information on the initial B_s^0 flavor and assuming no CP violation is given by equation 4.9. The physics parameters are given by

$$\vec{\lambda}_{\text{Phys}}^{\phi_s=0} = \left\{\Gamma_s, \Delta\Gamma_s, |A_0(0)|^2, |A_\perp(0)|^2, \delta_\parallel\right\}. \tag{6.1}$$

The strong phase $\delta_\parallel$ only appears as $\cos\delta_\parallel$ in the PDF. An alternative parameterization is therefore

$$\vec{\lambda}_{\text{Phys}}^{\phi_s=0} = \left\{\Gamma_s, \Delta\Gamma_s, |A_0(0)|^2, |A_\perp(0)|^2, \cos\delta_\parallel\right\}. \tag{6.2}$$

Table 6.1 gives the fit results for both parameterizations using the *lifetime unbiased* data set consisting of 771 signal candidates. All parameters show good agreement for the two different parameterizations. Using the first parameterization the extracted value for the strong phase $\delta_\parallel$ is exactly $+\pi$. This value for $\delta_\parallel$ minimizes the value of $\cos\delta_\parallel$ resulting in $\cos\delta_\parallel = -1$. With the second parameterization the fit determines $\cos\delta_\parallel$ to be $\cos\delta_\parallel = (-1.237 \pm 0.274)$. This is, probably due to a statistical fluctuation in the data, smaller than $\cos\pi = -1$ which is the smallest value accessible for $\cos\delta_\parallel$ when using the parameterization 6.1. The implicit constraint $\cos\delta_\parallel \in [-1, +1]$ applied when using the first parameterization prevents the fit from reaching the best description of the data. This also results in an unreliable error estimate for $\delta_\parallel$. Therefore the parameterization 6.2

with $\cos\delta_\parallel$ as physics parameter has been chosen as nominal fit configuration for the untagged analysis.

The correlation matrix for this configuration is given in Table 6.2. Correlations larger than 0.50 are given in bold, correlations smaller than 0.005 are omitted. As expected, moderately large correlations between the transversity amplitudes and Γ_s and $\Delta\Gamma_s$ are seen. The other correlation coefficients, with the exception of the proper time background model, are relatively small.

Figure 6.1 shows the fitted PDF and the data projected on the reconstructed B_s^0 mass, the proper time and the angles in the transversity base. No large deviations of the fitted model from the data are seen in the one-dimensional projections. For a quantitative statement about the goodness of fit the point-to-point dissimilarity method described in section 4.3 is employed. Figure 6.2 shows the T distribution resulting from 200 permutations. A fraction of 37% of the permutations has a larger T value than the true data set. The hypothesis that the fitted PDF corresponds to the true underlying PDF from which the data was drawn can therefore be rejected with a confidence level of only 63%. This means that the data is indeed well described by the fitted PDF.

Figure 6.3 shows the negative logarithm of the likelihood for the physics parameters. The negative logarithm of the likelihood is minimized with respect to all other parameters so that the uncertainty for the parameter is given by the points where $-2\ln\mathcal{L}$ rises by one unit. This is the likelihood ratio method (or likelihood scan) described in section 4.2.3. Close to the minimum the distributions are approximately parabolic justifying the assumption of parabolic errors[1]. A two-dimensional likelihood scan for Γ_s and $\Delta\Gamma_s$ is given in Figure 6.4 showing the correlation between the two parameters. The resulting confidence regions show very regular behavior, similar to two correlated Gaussian variables.

6.1.1. Determination of $\Delta\Gamma_s$ using the full lifetime range

The fit described in the last section was performed using only events with $t > 0.3\,\text{ps}$. This requirement rejects the prompt background events at small proper times which is computationally beneficial, in particular when using computation intensive methods as the method proposed by Feldman and Cousins [74]. However when using a simple maximum likelihood technique the analysis of all events without the proper time requirement is feasible.

The determination of $\Delta\Gamma_s$ over the full proper time range serves as an excellent crosscheck for the stability of the fit result even if no gain in sensitivity on $\Delta\Gamma_s$ is expected. Figure 6.5 gives the projection of the PDF for the maximum likelihood fit over the full range on the measured quantities. The projection on the proper time clearly shows the large prompt background peak, which is modeled as triple Gaussian in addition to the longlived background and signal components. The large number of prompt background events is also evident in the mass projection where the signal to background ratio is reduced. The maximum order of the Legendre polynomials used to

[1] In the MINUIT fit HESSE is used for error estimation

Parameter	Result
$\Gamma_s[\,\mathrm{ps}^{-1}]$	0.6792 ± 0.0360
$\Delta\Gamma_s[\,\mathrm{ps}^{-1}]$	0.077 ± 0.119
$\lvert A_\perp\rvert^2$	0.2634 ± 0.0555
$\lvert A_0\rvert^2$	0.5283 ± 0.0402
$\delta_\parallel[\,\mathrm{rad}]$	3.141 ± 0.523
$m_{B_s}[\,\mathrm{MeV}]$	5366.400 ± 0.332
$\sigma_{m,1}^{sig}[\,\mathrm{MeV}]$	7.490 ± 0.270
$\alpha_m^{LL}[\,\mathrm{MeV}^{-1}]$	0.000941 ± 0.000691
f_{sig}	0.7342 ± 0.0172
$f_{\tau,1}^{LL}$	0.6999 ± 0.0722
$\tau_1^{LL}[\,\mathrm{ps}]$	0.1255 ± 0.0195
$\tau_2^{LL}[\,\mathrm{ps}]$	0.867 ± 0.181

(a)

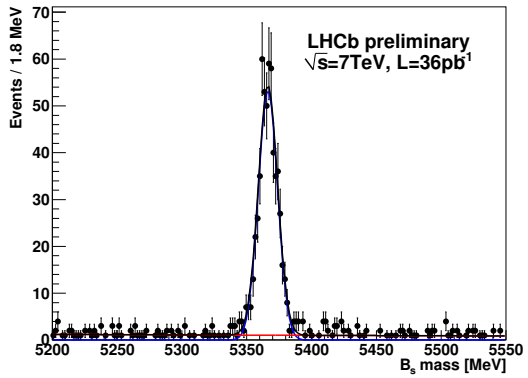
Parameter	Result
$\Gamma_s[\,\mathrm{ps}^{-1}]$	0.6797 ± 0.0342
$\Delta\Gamma_s[\,\mathrm{ps}^{-1}]$	0.084 ± 0.112
$\lvert A_\perp\rvert^2$	0.2786 ± 0.0572
$\lvert A_0\rvert^2$	0.5323 ± 0.0403
$\cos\delta_\parallel$	-1.237 ± 0.274
$m_{B_s}[\,\mathrm{MeV}]$	5366.396 ± 0.332
$\sigma_{m,1}^{sig}[\,\mathrm{MeV}]$	7.484 ± 0.269
$\alpha_m^{LL}[\,\mathrm{MeV}^{-1}]$	0.000939 ± 0.000690
f_{sig}	0.7339 ± 0.0172
$f_{\tau,1}^{LL}$	0.6993 ± 0.0721
$\tau_1^{LL}[\,\mathrm{ps}]$	0.1254 ± 0.0195
$\tau_2^{LL}[\,\mathrm{ps}]$	0.868 ± 0.181

(b)

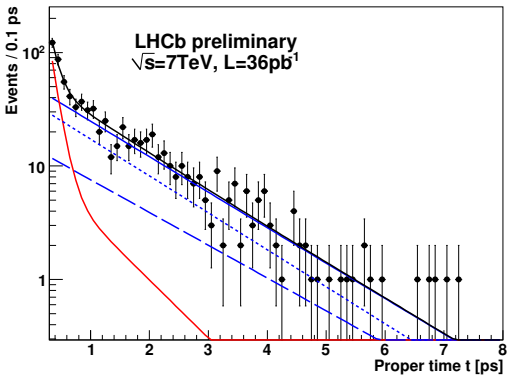
Table 6.1.: Results of an untagged fit to the *lifetime unbiased* data sample using two different parameterizations. (a) gives the results using $\delta_\parallel$ as physics parameter while (b) uses $\cos\delta_\parallel$ for the parameterization. The value returned for $\delta_\parallel$ using the first parameterization is exactly $+\pi$. This is consistent with the value of $\cos\delta_\parallel$ which is extracted using the second parameterization. $\cos\delta_\parallel$ is lower than -1 but statistically still compatible with it. The uncertainties given are statistical and determined by MINUITs HESSE step.

	Γ_s	$\Delta\Gamma_s$	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\cos\delta_\parallel$	m_{B_s}	$\sigma_{m,1}^{sig}$	α_m^{LL}	f_{sig}	$f_{\tau,1}^{LL}$	τ_1^{LL}	τ_2^{LL}
Γ_s	1.00	-0.43	0.38	-0.32	-0.04	0.01	0.02	-	0.05	-	-	0.02
$\Delta\Gamma_s$		1.00	**-0.68**	**0.68**	-0.05	-	-0.01	-	-0.01	-0.01	-	-
$\lvert A_\perp\rvert^2$			1.00	**-0.58**	-0.36	-	0.01	-	0.01	-	-	-
$\lvert A_0\rvert^2$				1.00	-0.12	-	-0.03	-	-0.03	-0.01	-	-0.01
$\cos\delta_\parallel$					1.00	0.01	0.03	-	0.02	0.01	-	-
m_{B_s}						1.00	0.01	0.01	0.01	-	0.01	-
$\sigma_{m,1}^{sig}$							1.00	-	0.14	0.04	-0.02	-0.08
α_m^{LL}								1.00	-	-	-	-
f_{sig}									1.00	0.02	-0.02	-0.03
$f_{\tau,1}^{LL}$										1.00	**0.63**	**0.61**
τ_1^{LL}											1.00	0.46
τ_2^{LL}												1.00

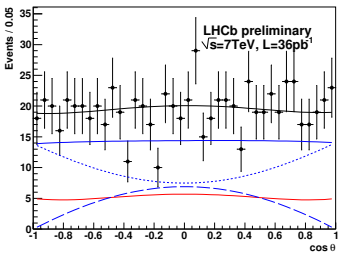
Table 6.2.: The correlations of the parameters in the nominal untagged fit to the *lifetime unbiased* data. Correlations larger than 0.50 are given in bold, correlations smaller than 0.005 are omitted.



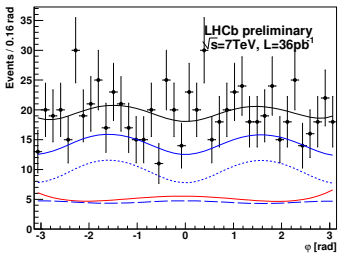
(a) Reconstructed B_s^0 mass



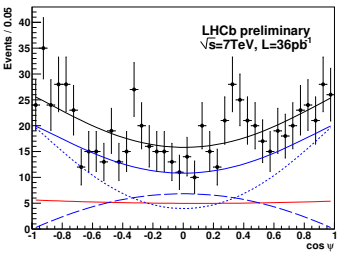
(b) Proper time t



(c) Transversity angle $\cos\theta$



(d) Transversity angle φ



(e) Transversity angle $\cos\psi$

Figure 6.1.: Projection of the fitted PDF of the untagged analysis on the measured quantities. Pictured are the reconstructed B_s^0 mass, the proper time and the three transversity angles $\Omega = \{\cos\theta, \varphi, \cos\psi\}$. The complete PDF is given as black line, the signal and background component are denoted by the blue and red lines respectively. The signal component is further subdivided into the CP-even (blue, dotted) part $\propto |A_0|^2$, $|A_\parallel|^2$ and the CP-odd (blue, dashed) part $\propto |A_\perp|^2$. Good agreement of the data with the fitted PDF is seen for all distributions.

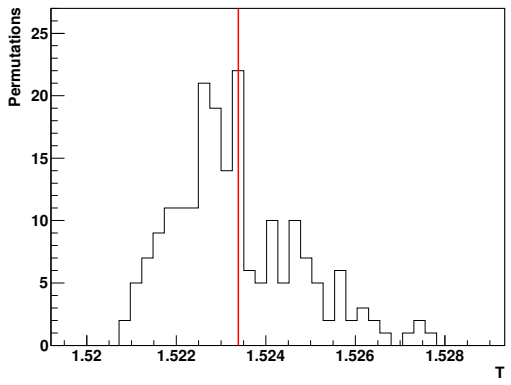


Figure 6.2.: T distribution from the point-to-point dissimilarity test for the untagged analysis. The red line denotes the T statistics of the data compared to a sample of $n_{\mathrm{MC}} = 20 \cdot n_{\mathrm{D}}$ simulated events generated according to the fitted PDF. The black histogram gives the T distribution of data samples of n_{D} events which are randomly drawn from the combination of real data and simulated events. 200 of these permutations are performed. The fraction of permutations with a larger T than the T statistics of the data gives a p-value of 37%.

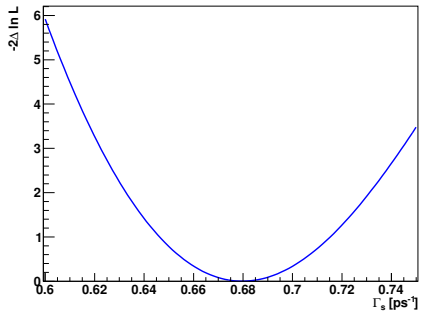
describe the angular background distributions was increased with respect to the nominal fit. This ensures an adequate angular description of the large background component.

The fit results are given in Table 6.3. Of particular interest is the extracted value of $\Delta\Gamma_{\mathrm{s}}$ which does not change significantly with respect to Table 6.1b. Also the other physics parameters show remarkably stable behavior. Together with very similar statistical uncertainties of the physics parameters this illustrates that the parameters are predominantly determined from signal candidates with larger proper times. Compared to the nominal fit in section 6.1 additional parameters are needed to describe the time and mass dependence of the prompt background component. In the proper time the prompt background is described by a triple Gaussian, as was detailed in section 4.6.2. The extracted parameters agree well with the triple Gaussian resolution model which is used in the nominal fit and was determined using the prompt background events in [63].

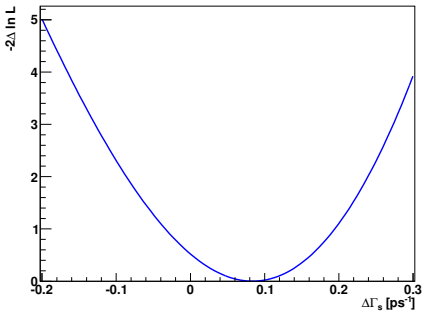
The correlation matrix for the fit over the full time range is given in Table B.2. The entries are similar to the correlations found for the nominal fit with $t > 0.3\,\mathrm{ps}$ in Table 6.2. Large correlations between the parameters of the triple Gaussian proper time resolution model are observed, but they are uncorrelated with the physics parameters.

6.2. Systematic uncertainties

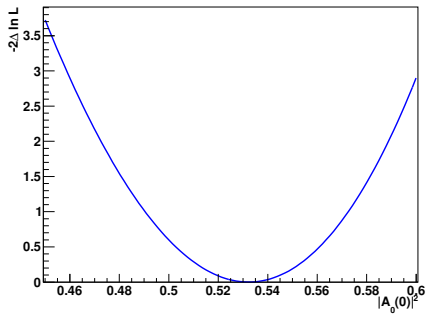
Besides the statistical uncertainties of the maximum likelihood fit systematic uncertainties can influence the result. With the low 2010 statistics the statistical uncertainties are



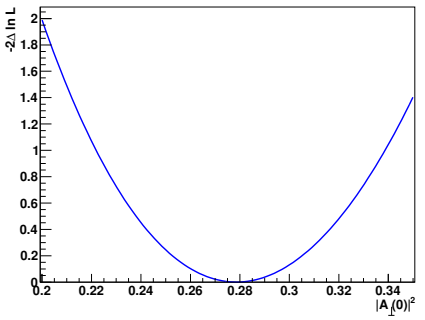
(a) $-2\ln\mathcal{L}\left(\Gamma_s\right)$



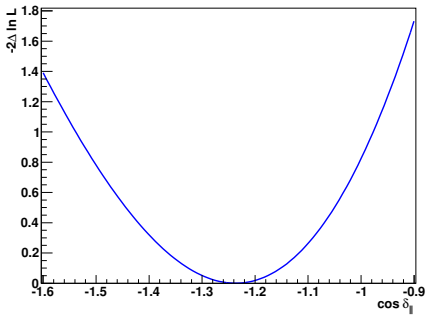
(b) $-2\ln\mathcal{L}\left(\Delta\Gamma_s\right)$



(c) $-2\ln\mathcal{L}\left(|A_0(0)|^2\right)$



(d) $-2\ln\mathcal{L}\left(|A_\perp(0)|^2\right)$



(e) $-2\ln\mathcal{L}\left(\cos\delta_\parallel\right)$

Figure 6.3.: One-dimensional likelihood scans for the physics parameters. The negative logarithm of the likelihood is minimized with respect to all other parameters. For Gaussian PDFs the resulting dependence on the parameter is parabolic, in the vicinity of the minimum this is approximately true for the physics parameters.

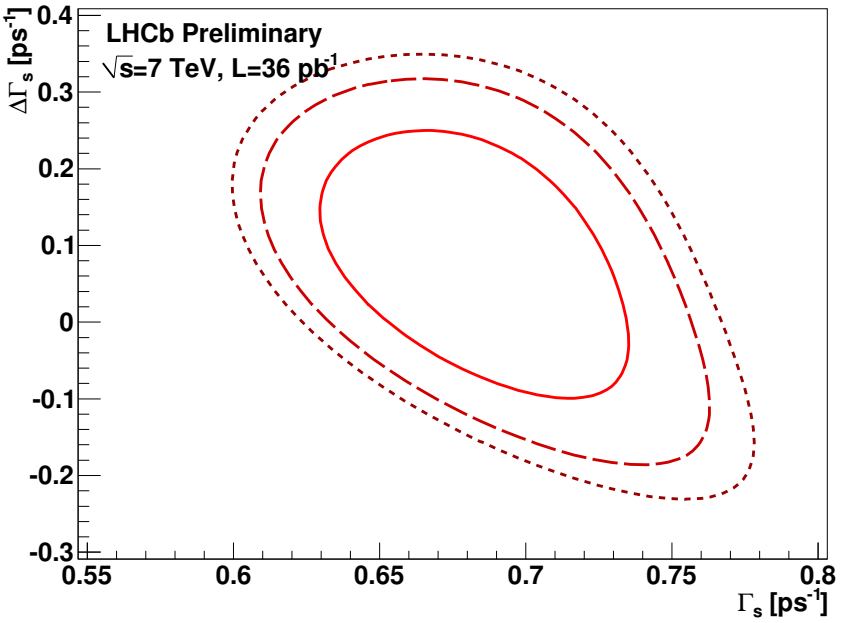
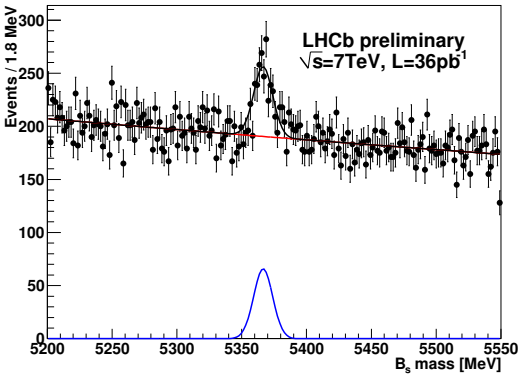
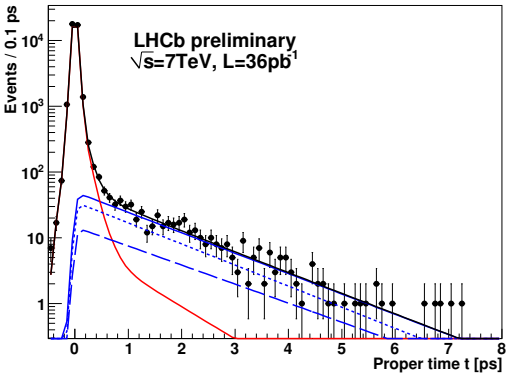


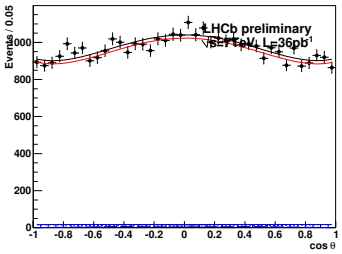
Figure 6.4.: Result of a two-dimensional likelihood scan for Γ_s and $\Delta\Gamma_s$. The resulting contours correspond to confidence levels of 68.3% (solid), 90% (dashed) and 95% (dotted) determined via the likelihood ratio method (section 4.2.3). The contours clearly show the negative correlation between the two parameters.



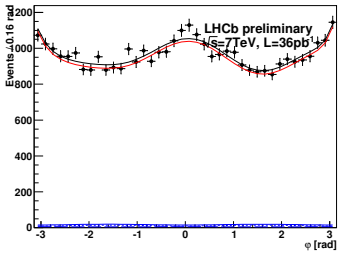
(a) Reconstructed B_s^0 mass



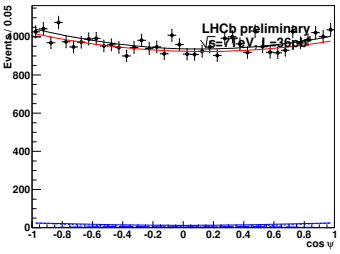
(b) Proper time t



(c) Transversity angle $\cos\theta$



(d) Transversity angle φ



(e) Transversity angle $\cos\psi$

Figure 6.5.: Projection of the fitted PDF on the measured quantities for the fit over the full proper time range. The fitted PDF describes the data reasonably well.

Parameter	Result
$\Gamma_{\mathrm{s}}[\,\mathrm{ps}^{-1}]$	0.675 ± 0.033
$\Delta\Gamma_{\mathrm{s}}[\,\mathrm{ps}^{-1}]$	0.071 ± 0.105
$\lvert A_{\perp}\rvert^2$	0.284 ± 0.053
$\lvert A_0\rvert^2$	0.536 ± 0.038
$\cos\delta_{\parallel}$	-1.249 ± 0.271
$m_{B_{\mathrm{s}}}[\,\mathrm{MeV}]$	5366.538 ± 0.318
$\sigma^{sig}_{\mathrm{m},1}[\,\mathrm{MeV}]$	7.396 ± 0.261
$\alpha^{\mathrm{P}r}_{\mathrm{m}}[\,\mathrm{MeV}^{-1}]$	0.0004853 ± 0.0000535
$\alpha^{LL}_{\mathrm{m}}[\,\mathrm{MeV}^{-1}]$	0.001120 ± 0.000545
f_{sig}	0.018011 ± 0.000748
$f_{\mathrm{P}r}$	0.97004 ± 0.00300
$f^{\mathrm{L}L}_{\tau,1}$	0.9319 ± 0.0184
$\tau^{\mathrm{L}L}_{1}[\,\mathrm{ps}]$	0.1381 ± 0.0115
$\tau^{\mathrm{L}L}_{2}[\,\mathrm{ps}]$	0.907 ± 0.181
$f^{sig}_{\mathrm{t},1}$	0.4940 ± 0.0561
$f^{sig}_{\mathrm{t},2}$	0.9387 ± 0.0111
$\sigma^{sig}_{\mathrm{t},1}[\,\mathrm{ps}]$	0.03428 ± 0.00132
$\sigma^{sig}_{\mathrm{t},2}[\,\mathrm{ps}]$	0.06090 ± 0.00246
$\sigma^{sig}_{\mathrm{t},3}[\,\mathrm{ps}]$	0.1462 ± 0.0103

Table 6.3.: Results of the untagged fit over the full proper time range. Compared with the nominal fit result for $t > 0.3\,\mathrm{ps}$ given in Table 6.1b the results are remarkably stable.

Parameter	Result	Abs. deviation
Γ_s	0.6798 ± 0.0342	$+0.0001$
$\Delta\Gamma_s$	0.084 ± 0.112	-
$\lvert A_\perp \rvert^2$	0.2786 ± 0.0571	-
$\lvert A_0 \rvert^2$	0.5323 ± 0.0403	-
$\cos\delta_\parallel$	-1.237 ± 0.274	-

Table 6.4.: Fit results with a proper time resolution model which is 50% worse than the nominal model.

expected to dominate. However the systematic uncertainties still need to be estimated.

To estimate the influence of a systematic uncertainty on the physics parameters the fit is repeated after applying the systematic variation. The fit result is compared with the nominal result and the deviation is used as systematic uncertainty. A more thorough approach would be to evaluate the systematic uncertainties using large amounts of simulated data. However since the systematic uncertainties at this stage are comparably small the simpler approach described above should give correct results in first approximation.

6.2.1. Proper time resolution

The proper time resolution model used in the nominal fit is a triple Gaussian as described in section 4.5.6. As previously mentioned the resolution parameters were determined in a different analysis [63]. The resolution parameters can only be extracted with a certain uncertainty. To account for a possible incorrect description of the proper time resolution the widths of the three Gaussians $\sigma_{t,1}^{sig}$, $\sigma_{t,2}^{sig}$ and $\sigma_{t,3}^{sig}$ are scaled with a factor 1.5. This is a rather conservative estimate and will overestimate the uncertainty of the resolution parameters which in reality is about 10%. However, since the untagged analysis does not need to resolve the fast B_s^0 oscillation the effect is expected to be negligible. This expectation is confirmed when comparing the result extracted using this resolution model with the nominal result. Table 6.4 gives the result and the absolute deviations of the physics parameters with respect to the nominal fit. As expected the deviations are negligible.

6.2.2. Proper time acceptance

As discussed in section 3.6.2 the event reconstruction exhibits a non-flat proper time dependent efficiency. In the fit this effect is fully accounted for by parameterizing the acceptance correction as linear function of the proper time. The full size of the acceptance correction is used as systematic uncertainty. The extracted parameters when neglecting the proper time acceptance are given in Table 6.5. As expected, mainly Γ_s is affected by the proper time acceptance. The effect on all other parameters is negligible.

Parameter	Result	Abs. deviation
Γ_s	0.7070 ± 0.0342	$+0.0273$
$\Delta\Gamma_s$	0.083 ± 0.112	-0.001
$\lvert A_\perp \rvert^2$	0.2789 ± 0.0571	$+0.0003$
$\lvert A_0 \rvert^2$	0.5321 ± 0.0403	-0.0002
$\cos\delta_\parallel$	-1.237 ± 0.274	-

Table 6.5.: Fit results under the assumption of a flat proper time acceptance.

6.2.3. Angular acceptance

Two sources of systematic uncertainties are studied that can arise from the description of the angular acceptance effect using simulated events. The first possible systematic effect arises from the limited statistics of events generated in the full simulation. To estimate this effect the number of accepted events in the histogram used to derive the acceptance parameterization is varied according to the statistical uncertainty. The acceptance parameterization is then rederived and the parameters are extracted with this varied acceptance description. This procedure is repeated four times and the largest deviations of the physics parameters are taken as systematic uncertainties[2]. The largest deviations are given in Table 6.6a.

As discussed in section 3.7 the angular acceptance effect is caused by the polar-angle acceptance of the detector and the implicit momentum cuts on the final state particles. Any discrepancies in the description of these quantities in the full simulation would result in incorrect acceptance descriptions. To check that the full simulation describes these quantities correctly a comparison with data is performed. To compare simulated signal events with the signal component in data the sideband subtraction technique is used[3]. Figures 3.11 and 3.12 show good agreement for sideband subtracted data (black) and fully simulated signal events (red) in all distributions with the possible exception of the muon momenta. To estimate the size of this effect the simulated events are reweighted according to their muon momenta and the acceptance correction is determined. The resulting new acceptance correction is used to determine the physics parameters. The deviations from the nominal fit values are used as an estimate of the systematic uncertainty and given in Table 6.6b.

[2] To obtain a more accurate estimate of the systematic effect the acceptances should be varied more often. Considering the large statistical uncertainties and the small observed deviations in the four variations performed four repetitions are sufficient.

[3] Events from the B_s^0 mass sidebands $[5200, 5321.67]\,\text{MeV} \cup [5411.67, 5550]\,\text{MeV}$ are subtracted from the signal region $[5321.67, 5411.67]\,\text{MeV}$ according to the background fraction.

Parameter	Largest abs. deviation
Γ_s	-0.0002
$\Delta\Gamma_s$	$+0.001$
$\lvert A_\perp\rvert^2$	-0.0028
$\lvert A_0\rvert^2$	-0.0021
$\cos\delta_\parallel$	0.027

(a)

Parameter	Result	Abs. deviation
Γ_s	0.6797 ± 0.0342	-
$\Delta\Gamma_s$	0.084 ± 0.112	-
$\lvert A_\perp\rvert^2$	0.2782 ± 0.0571	-0.0004
$\lvert A_0\rvert^2$	0.5329 ± 0.0403	$+0.0006$
$\cos\delta_\parallel$	-1.233 ± 0.274	$+0.004$

(b)

Table 6.6.: Systematic uncertainties due to the description of the angular acceptance effect. (a) gives the systematic uncertainty due to the limited Monte Carlo statistics. To estimate the effect the simulated signal events were varied four times within their statistical uncertainty. The largest deviations resulting from the four different acceptance corrections are taken as systematic uncertainties for the parameters. Table B.1 gives the detailed results of the four variations.
(b) gives the systematic effect from reweighting the fully simulated signal events according to the muon momentum.

6.2.4. Background modeling

In the nominal fit the angular dependence of the background component is described by a combination of Legendre polynomials. The precision of this description is however limited by the low number of events in the B_s^0 mass sidebands where the Legendre coefficients are determined. To estimate the systematic uncertainty of the background description due to the limited statistics of the 2010 data the nominal description is compared with the assumption of flat angular distributions of the background. The resulting parameters and their deviations from the nominal fit result are given in Table 6.7.

Parameter	Result	Abs. deviation
Γ_s	0.6800 ± 0.0342	$+0.0003$
$\Delta\Gamma_s$	0.082 ± 0.112	-0.002
$\lvert A_\perp\rvert^2$	0.2800 ± 0.0575	$+0.0014$
$\lvert A_0\rvert^2$	0.5308 ± 0.0403	-0.0015
$\cos\delta_\parallel$	-1.231 ± 0.274	$+0.006$

Table 6.7.: Fit results under the assumption of a flat angular distribution of the background component.

Parameter	Result	Abs. deviation
Γ_s	0.6801 ± 0.0343	$+0.0004$
$\Delta\Gamma_s$	0.087 ± 0.113	$+0.003$
$\lvert A_\perp \rvert^2$	0.2733 ± 0.0569	-0.0053
$\lvert A_0 \rvert^2$	0.5310 ± 0.0400	-0.0013
$\cos\delta_\parallel$	-1.201 ± 0.266	$+0.036$

Table 6.8.: Fit results using a double Gaussian to describe the signal mass peak.

6.2.5. Mass model

The nominal fit describes the signal component in the reconstructed B_s^0 mass as a single Gaussian. To estimate the systematic uncertainty due to a possibly incorrect modeling of the mass distribution a double Gaussian mass model is used instead. The extracted parameters and their deviations from the nominal result are given in Table 6.8.

6.2.6. S-wave contribution

The transition of the B_s^0 to the final state $\mu^+\mu^- K^+K^-$ usually occurs via the J/ψ and ϕ resonances. It is however also possible that the K^+K^- does not originate from the ϕ resonance but instead from an additional nonresonant $\ell = 0$ contribution or via the $f_0(980)$ resonance. This contribution is called "S-wave" since the resulting K^+K^- system is in an S-wave configuration ($\ell = 0$) as was already discussed in sections 1.3.3 and 4.5.5. In the nominal fit a possible S-wave contribution is neglected. To estimate the systematic uncertainty this causes when the S-wave is in fact present the fit is repeated with the largest S-wave contribution currently allowed by experimental data, $|A_s(0)|^2 = 0.067$ at the 95% confidence level [39]. Note that the PDF depends on an additional strong phase difference when an S-wave component is present, namely $\sin\left(\delta_\parallel - \delta_s\right)$. The resulting parameter values are given in Table 6.9.

This estimation of the systematic uncertainties which result from neglecting the S-wave contribution is conservative since it assumes the largest S-wave amplitude allowed by the CDF experiment. Ideally the S-wave contribution should be included by fitting for both the S-wave phase and the amplitude. However with the small number of signal candidates in the *lifetime unbiased* 2010 data sample this leads to fit instabilities.

6.2.7. Summary of the systematic uncertainties

A summary of the systematic uncertainties is given in Table 6.10. The systematic uncertainty on Γ_s is, as expected, dominated by the proper time acceptance. For all other parameters the largest systematic uncertainty stems from the neglected S-wave. The total systematic error is evaluated by adding the individual contributions in quadrature. Overall the systematic uncertainties are smaller than the statistical uncertainties.

Parameter	Result	Abs. deviation
Γ_s	0.6854 ± 0.0323	$+0.0057$
$\Delta\Gamma_s$	0.071 ± 0.122	-0.013
$\lvert A_\perp \rvert^2$	0.2670 ± 0.0537	-0.0116
$\lvert A_0 \rvert^2$	0.4973 ± 0.0408	-0.035
$\cos\delta_\parallel$	-1.340 ± 0.315	-0.103
$\sin(\delta_\perp - \delta_s)$	-0.122 ± 0.330	-

Table 6.9.: Fit results under the assumption of a 6.7% S-wave contribution and deviation of the parameters with respect to the nominal fit.

Syst. uncertainty	Γ_s	$\Delta\Gamma_s$	$\lvert A_\perp \rvert^2$	$\lvert A_0 \rvert^2$	$\cos\delta_\parallel$
Proper time resolution model	0.0001	–	–	–	–
Proper time acceptance	0.0273	0.001	0.0003	0.0002	–
Angular acceptance MC statistics	0.0002	0.001	0.0028	0.0021	0.027
Angular acceptance MC description	–	–	0.0004	0.0006	0.004
Angular background description	0.0003	0.002	0.0014	0.0015	0.006
Mass model	0.0004	0.003	0.0053	0.0013	0.036
S-wave contribution	0.0057	0.013	0.0116	0.035	0.103
Quadratic sum Σ	0.0279	0.0136	0.0131	0.0351	0.113

Table 6.10.: Overview of the different sources of systematic uncertainties. The total systematic error is evaluated by adding the individual contributions in quadrature.

Future analyses at LHCb will use a much larger data sample (data corresponding to more than $500\,\mathrm{pb}^{-1}$ have been taken until August 2011) and will therefore be able to significantly improve the measurement of the physics parameters.

6.3. Accounting for possible CP violation without using the tagging information

Equation 4.8 shows that the PDF has a weak dependence on ϕ_s even when no tagging information is used to distinguish the two B_s^0 production flavors. To determine the sensitivity on this parameter in the untagged analysis a Feldman-Cousins study [74] is performed. Confidence regions are constructed in the parameter space of the two physics parameters ϕ_s and $\Delta\Gamma_s$. Technically the analysis is performed on a 40×40 grid in ϕ_s and $\Delta\Gamma_s$. For every grid point the confidence level of the point is evaluated by performing 1000 toy experiments and using the likelihood ratio ordering principle proposed by

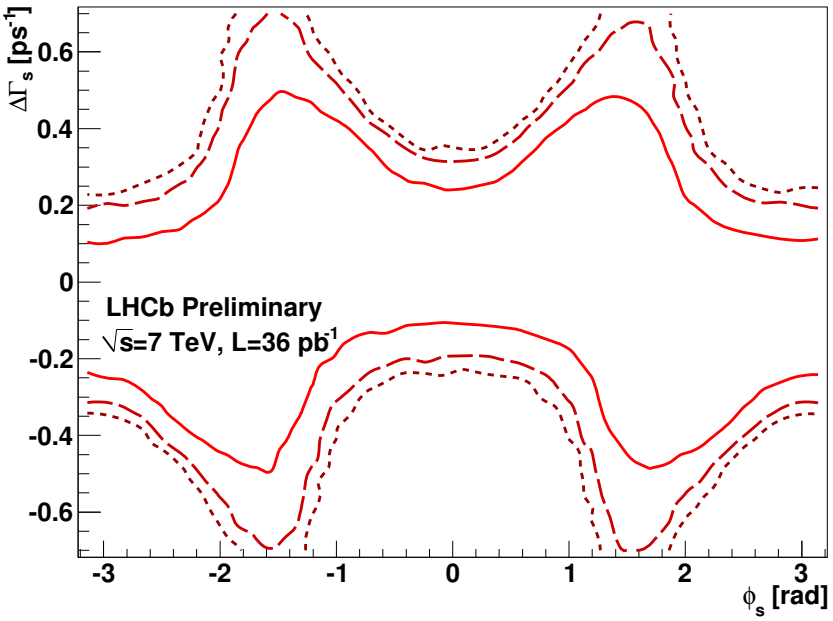


Figure 6.6.: Confidence contours determined by a Feldman-Cousins study in the ϕ_s- $\Delta\Gamma_s$ parameter space without using information on the B_s^0 production flavor. The confidence contours correspond to the 68.3% (solid), 90% (dashed) and 95% (dotted) confidence levels.

Feldman-Cousins which was discussed in detail in section 4.2.2. The resulting confidence regions for confidence levels of 68.3%, 90% and 95% are given in Figure 6.6. The figure shows a four-fold symmetry as expected from the discussion in 4.5.4. From the confidence contours it is clear that with the small size of the data sample taken in 2010 ϕ_s can not be constrained by a flavor-blind analysis. For a determination of ϕ_s the use of tagging information to distinguish between the two B_s^0 production flavors is therefore critical and described in detail in the next chapter.

6.4. Summary of the untagged result

The physics parameters determined in the untagged analysis are given in Table 6.11. The parameters are extracted under the assumption of no time dependent CP violation ($\phi_s = 0$) and using data taken by the LHCb detector on 2010. The analysis is performed on the *lifetime unbiased* data sample which contains 771 signal candidates over the full range of the reconstructed B_s^0 mass [5200 MeV, 5550 MeV]. The most interesting physics quantity is the decay width difference in the B_s^0 system $\Delta\Gamma_s$ which is determined to be $\Delta\Gamma_s = (0.084 \pm 0.112_{\text{stat.}} \pm 0.014_{\text{syst.}})$ ps^{-1}. This is in excellent agreement with

Parameter	Result $\pm$ stat. $\pm$ syst.
$\Gamma_s[\,\mathrm{ps}^{-1}]$	$0.6797 \pm 0.0342 \pm 0.0279$
$\Delta\Gamma_s[\,\mathrm{ps}^{-1}]$	$0.084 \pm 0.112 \pm 0.014$
$\lvert A_\perp\rvert^2$	$0.2786 \pm 0.0572 \pm 0.0131$
$\lvert A_0\rvert^2$	$0.5323 \pm 0.0403 \pm 0.0351$
$\cos\delta_\parallel$	$-1.237 \pm 0.274 \pm 0.113$

Table 6.11.: Summary of the results of the untagged analysis of the data taken by LHCb in 2010. The study is performed under the assumption $\phi_s = 0$.

the Standard Model prediction $\Delta\Gamma_s = (0.087 \pm 0.021)\,\mathrm{ps}^{-1}$ [34]. It also agrees with the currently best measurement of $\Delta\Gamma_s$ performed by the CDF collaboration $\Delta\Gamma_s = (0.075 \pm 0.035_{\mathrm{stat.}} \pm 0.01_{\mathrm{syst.}})\ \mathrm{ps}^{-1}$ [39]. The measurement using the LHCb data is not yet competitive with the CDF measurement due to the larger statistical uncertainty. Considering that this analysis was performed on data corresponding to only $36\,\mathrm{pb}^{-1}$ this is expected. Furthermore, the systematic uncertainties which have been estimated conservatively are already comparable to the systematic uncertainties quoted by CDF. With an expected integrated luminosity of up to $1\,\mathrm{fb}^{-1}$ in 2011 future measurements will profit from a much larger data sample. This also simplifies the determination of the systematic uncertainties. One example would be the S-wave contribution, which could be included in the nominal fit and extracted from the data. Therefore LHCb will be able to perform the most precise measurement of $\Delta\Gamma_s$ with the 2011 data.

7. Determination of ϕ_s using information on the $\mathrm{B_s^0}$ production flavor

This chapter presents the determination of ϕ_s and $\Delta\Gamma_\mathrm{s}$ using information on the initial $\mathrm{B_s^0}$ production flavor. Exploiting the largest possible event statistics using both the *lifetime biased* and *unbiased* data it will be shown that it is possible to constrain ϕ_s and $\Delta\Gamma_\mathrm{s}$ when using tagging information.

7.1. Two-dimensional constraints on ϕ_s and $\Delta\Gamma_\mathrm{s}$

When using information on the initial $\mathrm{B_s^0}$ flavor the PDF used for the signal component is given by equation 4.5 and depends on the physics parameters

$$\vec{\lambda}_{\mathrm{Phys}} \;=\; \left\{\phi_s, \Gamma_\mathrm{s}, \Delta\Gamma_\mathrm{s}, |A_0(0)|^2, |A_\perp(0)|^2, \delta_\parallel, \delta_\perp, \Delta m_\mathrm{s}\right\}.$$

The fit minimizes the negative logarithm of the likelihood (equation 4.3) by variation of the parameters and thereby determines the physics parameters for which the data becomes most likely. These parameter values are given in Table 7.1. Note that no uncertainties are given since an error estimate via the methods described in section 4.2.3 and 4.2.4 is only possible when the logarithm of the likelihood becomes parabolic. This is not the case as will become clear later in this section.

Using the fitted parameter values in Table 7.1 the PDF and data are projected on the measured quantities. Figure 7.1 shows the projection on the proper time. The projections on the reconstructed $\mathrm{B_s^0}$ mass and the transversity angles are given in Figures 7.2 and 7.3 respectively. All projections show the *lifetime biased* and *lifetime unbiased* data samples separately. Considering the low statistics good agreement between data and the fitted PDF is observed in all one-dimensional projections. Since the PDF is five-dimensional this is of course only a first qualitative hint that the data is well described by the fitted PDF. For a quantitative conclusion on the goodness of fit the point-to-point dissimilarity method described in section 4.3 is employed. Figure 7.4 shows the T distribution resulting from 200 permutations. A fraction of 27% of the permutations has a larger T value than the true data set. The hypothesis that the fitted PDF corresponds to the true underlying PDF from which the data was drawn can be rejected with a confidence level of 73%. The description of the data by the fitted PDF is therefore adequate.

To determine the confidence regions for ϕ_s and $\Delta\Gamma_\mathrm{s}$ the Feldman-Cousins method [74] as described in section 4.2.2 is used. In contrast to the simpler likelihood ratio method described in section 4.2.3 the Feldman-Cousins method guarantees correct coverage regardless of the shape of the likelihood. To determine the confidence regions 1000 simulated data sets are generated at every point on a 40×40 grid in ϕ_s and $\Delta\Gamma_\mathrm{s}$. The simulated events are fit twice. In the first fit the parameters ϕ_s and $\Delta\Gamma_\mathrm{s}$ are fixed to the

Parameter	Result
Γ_s	$0.690\,\mathrm{ps}^{-1}$
$\Delta\Gamma_s$	$0.064\,\mathrm{ps}^{-1}$
ϕ_s	$-1.13\,\mathrm{rad}$
$\lvert A_\perp\rvert^2$	0.290
$\lvert A_0\rvert^2$	0.524
$\delta_\parallel$	$3.29\,\mathrm{rad}$
$\delta_\perp$	$0.19\,\mathrm{rad}$
Δm_s	$17.75\,\mathrm{ps}^{-1}$

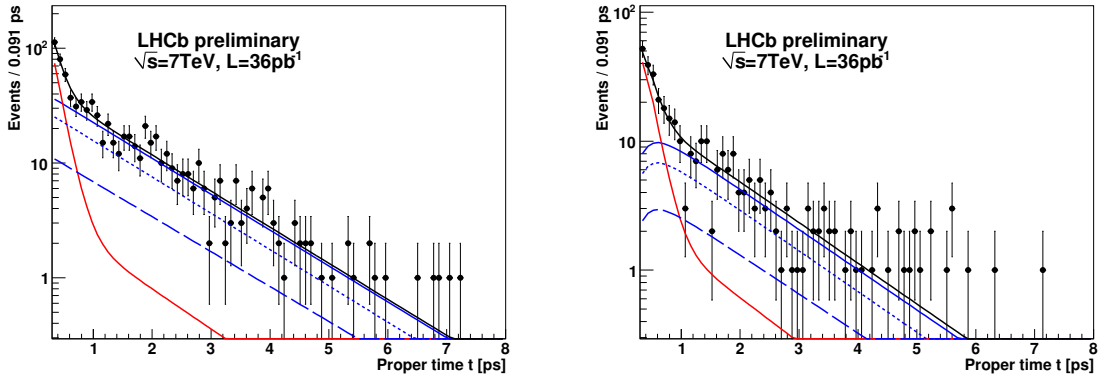
Table 7.1.: Fitted physics parameters for the tagged analysis using both the *lifetime biased* and the *lifetime unbiased* data sample. No error estimates are given since the negative logarithm of the likelihood is not parabolic.

specific grid point at which the simulated events are generated. The second fit allows for ϕ_s and $\Delta\Gamma_s$ to vary freely. The result are two negative logarithms of likelihoods, their difference is the logarithm of the likelihood ratio for the specific toy data set. The same procedure is performed for the 2010 data sample resulting in a likelihood ratio for the data at the specific grid point. As discussed in section 4.2.2 the confidence level of a certain grid point is given by the fraction of toy data sets that have a larger likelihood ratio than the data.

Figure 7.5 shows the confidence regions resulting from the Feldman-Cousins study. Three confidence regions are given corresponding to confidence levels of 68.3%, 90% and 95% respectively. Using the information on the B_s^0 production flavor some points in the ϕ_s-$\Delta\Gamma_s$ parameter space can already be excluded. The black dot denotes the Standard Model point. The probability for the Standard Model point corresponds to 1.0 standard deviations. The data is thus compatible with the Standard Model prediction.

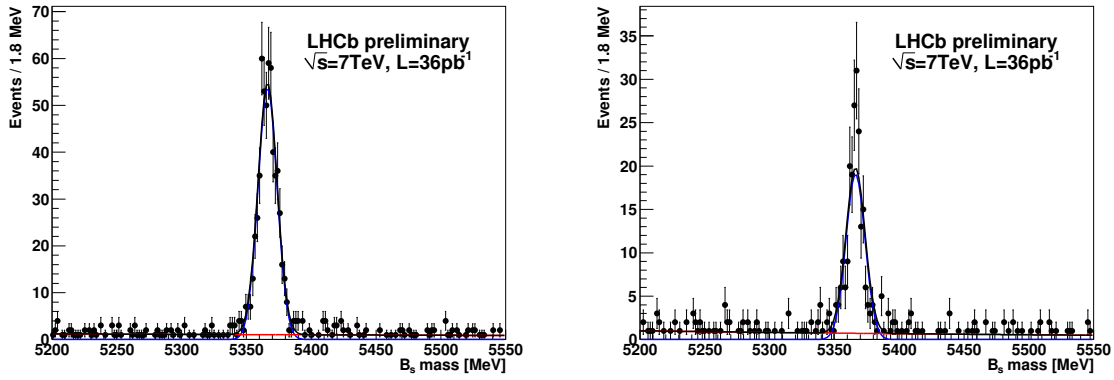
7.2. Evaluation of systematic uncertainties

The confidence contours presented in the last section are still quite large due to the limited number of signal events. To estimate the size of systematic effects confidence contours in the ϕ_s-$\Delta\Gamma_s$ parameter space are determined while varying the systematically limited parameters. The resulting contours are compared with the nominal confidence contours. The confidence contours are produced using the likelihood ratio method described in section 4.2.3. This approach has the advantage of being computationally much less expensive than the Feldman-Cousins method. While the likelihood ratio method does not necessarily give the correct coverage, the relative changes of the confidence contours will give a good estimate of the size of the systematic effects.



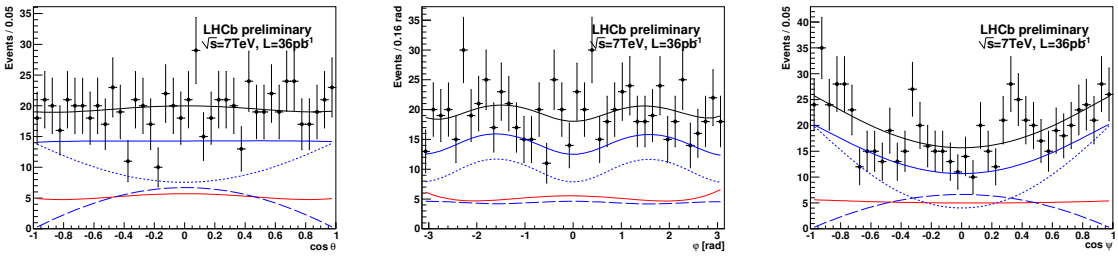
(a) Lifetime unbiased (b) Lifetime biased

Figure 7.1.: Distribution of the proper time for the *lifetime biased* and the *lifetime unbiased* data sample. The solid black line denotes the projection of the fitted PDF on the proper time. The signal component is given in blue, the background component in red. The CP-even part of the signal is denoted as blue dotted line, the CP-odd part is given by the blue dashed line. For the *lifetime biased* case the proper time acceptance affecting low proper times is clearly visible.

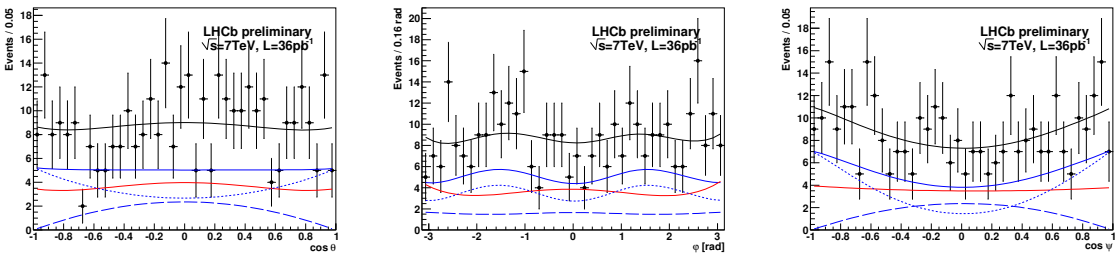


(a) Lifetime unbiased (b) Lifetime biased

Figure 7.2.: Distribution of the reconstructed B_s^0 mass for the *lifetime biased* and the *lifetime unbiased* data sample. The solid black line denotes the projection of the fitted PDF on the reconstructed B_s^0 mass. The signal component is given in blue, the background component in red.



(a) Lifetime unbiased



(b) Lifetime biased

Figure 7.3.: Distributions of the transversity angles (from left to right: $\cos\theta$, φ and $\cos\psi$) for the *lifetime biased* and the *lifetime unbiased* data sample. The solid black line denotes the projection of the fitted PDF on the transversity angles. The signal component is given in blue, the background component in red. The CP-even part of the signal is denoted as blue dotted line, the CP-odd part is given by the blue dashed line.

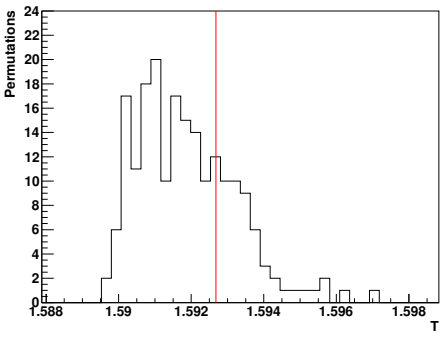


Figure 7.4.: T distribution from the point-to-point dissimilarity test for the tagged analysis. The red line denotes the T statistics of the data compared to a sample of $n_{\mathrm{MC}} = 20 \cdot n_{\mathrm{D}}$ simulated events generated according to the fitted PDF. The black histogram gives the T distribution of data samples of n_{D} events which are randomly drawn from the combination of real data and simulated events. 200 of these permutations are performed. The fraction of permutations with a larger T than the T statistics of the data gives a p-value of 27%.

7.2.1. Flavor tagging calibration

The flavor tagging calibration described in section 3.8.1 determines the parameters p_0 and p_1 which are used to obtain a calibrated mistag probability ω_{tag} for the analysis. The parameters are given in Table 3.7. The calibration procedure can only determine these parameters with a limited precision due to the limited event numbers of the calibration channels. A Gaussian constraint as discussed in section 4.7.2 is applied on the tagging parameters to account for the uncertainty of the tagging calibration. The systematic uncertainty due to the tagging calibration is therefore already included in the nominal result given in Figure 7.5.

7.2.2. Proper time resolution

To estimate the effect of an incorrect proper time resolution model the fit is repeated with the assumption of a 10% worse resolution than used in the nominal fit. The parameters $\sigma_{\mathrm{t},1}^{\mathrm{sig}}$, $\sigma_{\mathrm{t},2}^{\mathrm{sig}}$ and $\sigma_{\mathrm{t},3}^{\mathrm{sig}}$ are scaled with a factor 1.1. The resulting confidence contours are given in red in Figure 7.6a. For comparison, the nominal confidence contours are given in black. Compared to the statistical uncertainty the systematic change due to the different proper time resolution model is negligible.

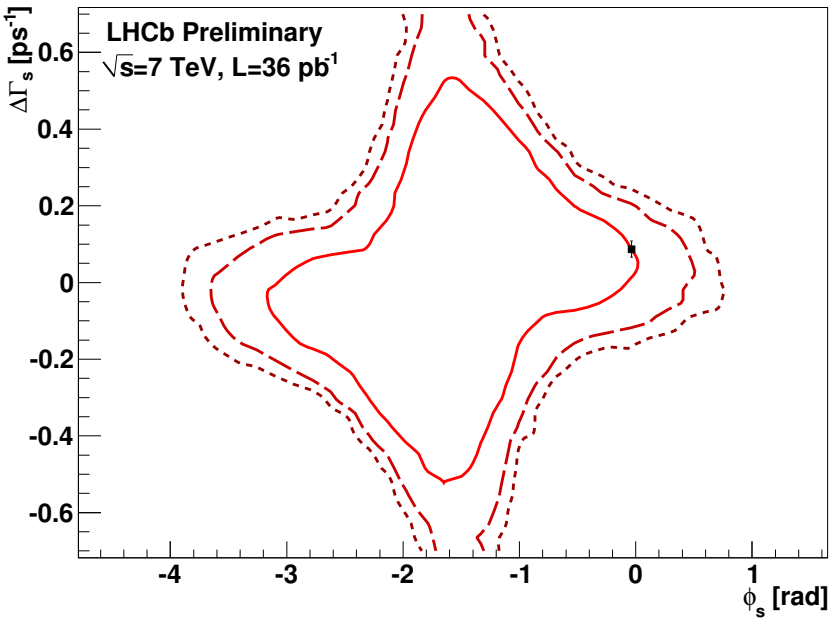


Figure 7.5.: Confidence contours in the ϕ_s-$\Delta\Gamma_s$ parameter space determined by the Feldman-Cousins method using tagging information on the B_s^0 production flavor. The confidence contours correspond to confidence levels of 68.3% (solid), 90% (dashed) and 95% (dotted) respectively. The black dot denotes the Standard Model prediction of $(-0.0363\,\text{rad}, 0.087\,\text{ps}^{-1})$.

7.2.3. Proper time acceptance

In section 3.6.2 a proper time dependent efficiency caused by a reconstruction effect was discussed. In the fit the effect is parameterized as $\epsilon(t) = (1 + \beta t)$ with $\beta = -0.025$. As systematic uncertainty the full size of β is taken. Figure 7.6b gives the confidence contours resulting under the assumption of $\beta = 0$. The difference with respect to the nominal contours is negligible.

7.2.4. Angular acceptance

Two sources of systematic uncertainties related to an incorrect description of the angular acceptance effect are studied. A questionable description of the muon momenta in data was observed for the simulated signal events (see Figure 3.12). The acceptance correction is therefore rederived from simulated signal events that were reweighted according to their muon momenta. The resulting confidence contours are given in Figure 7.7e.

Another possible source of systematic uncertainty is the limited number of simulated signal events used to determine the acceptance effect. To study this effect the number of simulated signal events in every bin of the acceptance histogram is varied within the statistical uncertainty. The acceptance parametrization is then rederived using the changed acceptance histogram. This is repeated four times which results in four slightly different acceptance corrections[1]. The confidence contours when using these acceptance corrections are given in Figures 7.7a to 7.7d. Both systematic effects are negligible compared to the large statistical uncertainty.

7.2.5. Background modeling

To estimate the effect of a possible incorrect angular description of the background component the analysis is performed using a flat angular description. Hardly any difference to the nominal three-dimensional parametrization is observed for the resulting contours in Figure 7.6c. This is a consequence of the very low background levels after the full selection.

7.2.6. Mass model

Instead of the nominal single Gaussian to describe the B_s^0 signal mass peak a double Gaussian mass model can also be used. The resulting confidence contours are given in Figure 7.6d. A small difference with respect ot the nominal contours is seen, however the systematic effect is negligible compared to the statistical uncertainty.

[1] As already mentioned in chapter 6.2.3 the estimation of the size of the systematic effect would be more precise when performing more variations. Considering the large statistical uncertainties and the small observed systematic deviations four repetitions are sufficient.

7.2.7. S-wave contribution

The inclusion of a possible S-wave component introduces two additional parameters into the fit, $|A_s(0)|^2$ and δ_s, as described in section 4.5.5. The resulting contours are given in Figure 7.6e. The small inner contours are artifacts caused by instabilities when fitting for the two additional physics parameters. Comparing the confidence contours in the ϕ_s-$\Delta\Gamma_s$ parameter space with the nominal fit no large difference is observed.

7.2.8. Summary of the systematic uncertainties

All systematic effects studied are negligible in size compared to the large statistical uncertainty and do not lead to a significant variation of the confidence contours. The systematic effects are therefore neglected at this stage.

7.3. One-dimensional confidence intervals for ϕ_s

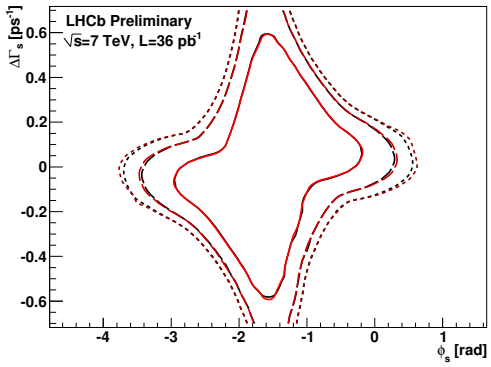
In addition to the two-dimensional confidence regions determined in section 7.1 one-dimensional confidence intervals for ϕ_s are determined using B_s^0 production flavor information. The Feldman-Cousins method is employed to determine confidence intervals with correct coverage. The Feldman-Cousins study is performed for 40 points in ϕ_s. At each point 1000 data samples of simulated events are generated and fitted. The result of the Feldman-Cousins study is given in Figure 7.8 which shows the confidence level depending on the value of ϕ_s. The confidence intervals corresponding to confidence levels of 68.3% and 90% are determined to be $[-2.78, -0.39]$ rad and $[-3.33, 0.21]$ rad respectively and are denoted by the vertical black lines. Since the systematic uncertainties were shown to be negligible in the two-dimensional ϕ_s-$\Delta\Gamma_s$ parameter space these confidence intervals for ϕ_s represent the one-dimensional tagged result[2].

7.4. Summary of the B_s^0 production flavor dependent analysis

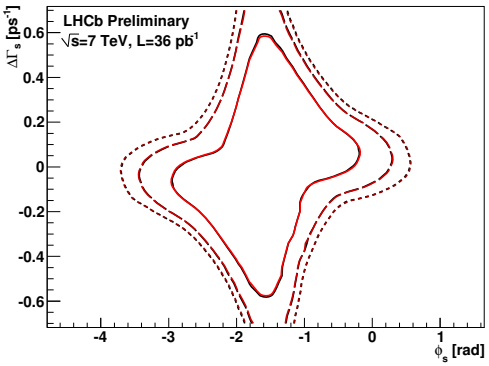
The two-dimensional constraints on ϕ_s and $\Delta\Gamma_s$ determined in the tagged analysis are shown in Figure 7.5. The confidence contours reflect the statistical error as well as the systematic uncertainty from the tagging calibration. All other systematic errors are neglected. The Standard Model prediction in the ϕ_s-$\Delta\Gamma_s$ parameter space is found to be in agreement with the data at the 1.0σ level (68.5% confidence level). In addition to the two-dimensional confidence regions in Figure 7.5 one-dimensional confidence intervals for ϕ_s are determined. The confidence intervals for ϕ_s corresponding to confidence levels of 68.3% and 90% are given by $[-2.78, -0.39]$ rad and $[-3.33, 0.21]$ rad respectively.

Compared with the latest measurements performed at the Tevatron which are given in Figure 1.7 good agreement is seen. All experiments favor negative ϕ_s. The confidence contours given in Figure 1.7 are somewhat smaller than in Figure 7.5 indicating that the

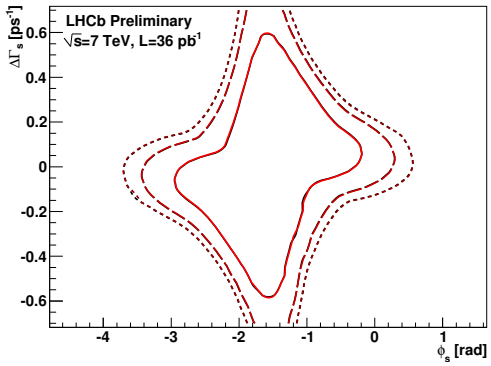
[2]Note that the systematic uncertainty due to the tagging calibration is included in the one-dimensional confidence intervals.



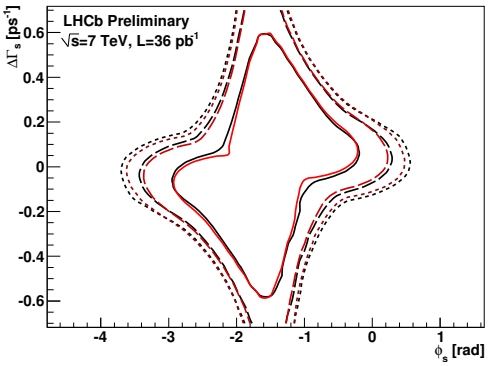
(a) Proper time resolution model



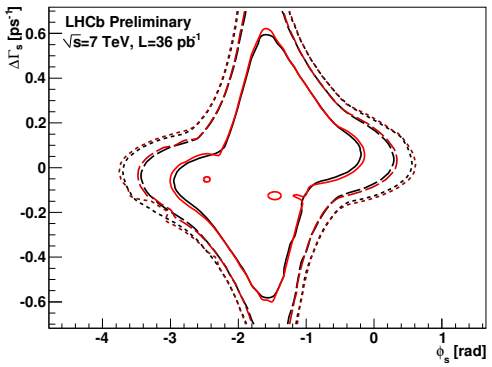
(b) Proper time acceptance



(c) Background model

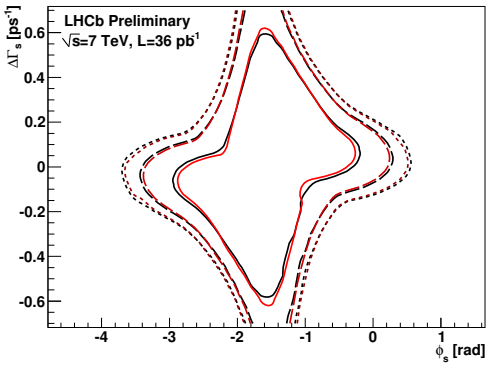


(d) Mass model

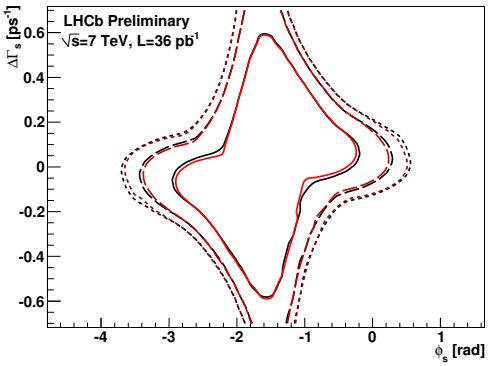


(e) S-wave

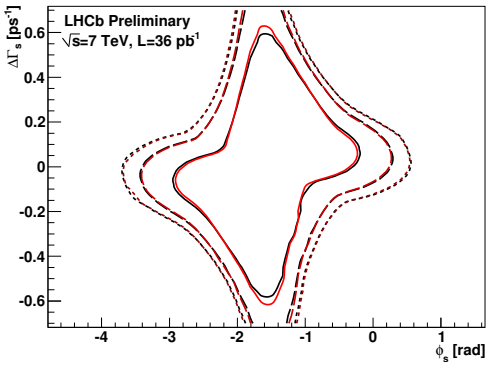
Figure 7.6.: Effect of systematic deviations on the confidence contours for the tagged fit. The red curves denote the contours after the systematic change while the black curves give the nominal confidence contours before the systematic change. The three contours given correspond to confidence levels of 68.3% (solid), 90% (dashed) and 95% (dotted) respectively. Systematic effects shown are (a) the proper time resolution model, (b) the proper time acceptance, (c) the background model, (d) the signal mass model and (e) the S-wave. The systematic changes are negligible compared to the statistical uncertainty for the 2010 data. The irregular shape of the contours when including a possible S-wave contribution shows that the fit with the two additional parameters is not completely stable with the available statistics.



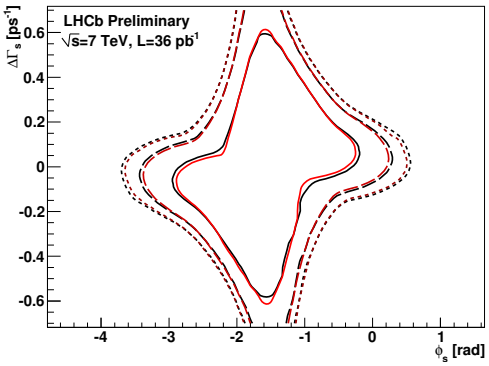
(a) Random variation within statistics



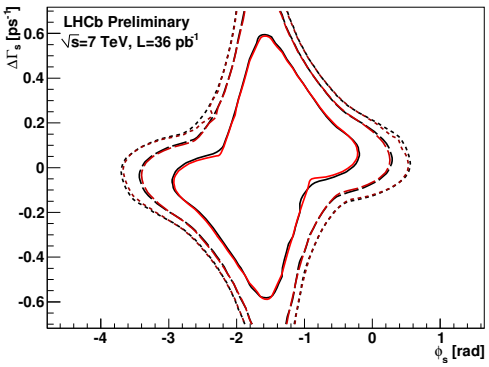
(b) Random variation within statistics



(c) Random variation within statistics



(d) Random variation within statistics



(e) Muon momentum reweighted

Figure 7.7.: Systematic deviations of the confidence contours due to the angular acceptance description. The red curves denote the confidence contours after the systematic change while the black curves give the nominal confidence contours before the systematic change. The three contours given correspond to confidence levels of 68.3% (solid), 90% (dashed) and 95% (dotted) respectively. Figures (a) to (d) show the effect of angular acceptance corrections that are determined after varying the simulated signal events within their statistics. Figure (e) gives the confidence contours resulting from an angular acceptance correction that was determined after reweighting the muon momenta.

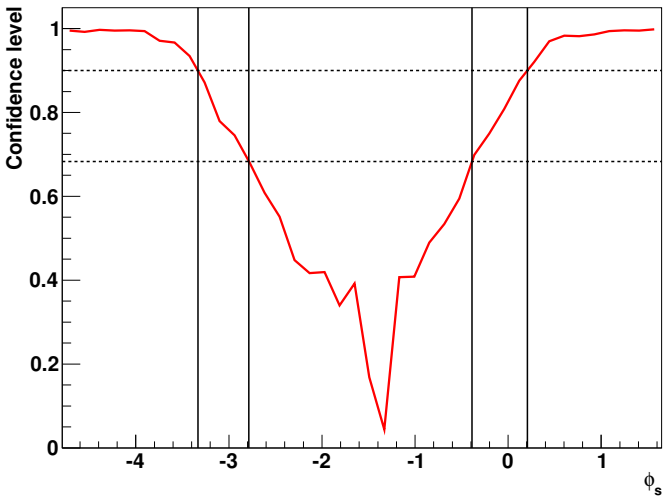


Figure 7.8.: Confidence levels depending on ϕ_s determined in a B_s^0 production flavor dependent one-dimensional Feldman-Cousins study. For each of the 40 points in ϕ_s 1000 data samples of simulated events are generated and fit. The vertical black lines denote the range of the 68.3% and the 90% confidence level intervals. They are given by $[-2.78, -0.39]$ rad and $[-3.33, 0.21]$ rad respectively.

LHCb experiment is not yet competitive with the data taken in 2010. Considering that both Tevatron experiments used data samples corresponding to much larger integrated luminosities (The CDF experiment used a data sample collected from 2002 to 2009 which corresponds to an integrated luminosity of $5.2\,\mathrm{fb}^{-1}$) this is not surprising. In this light the performance of the LHCb detector on the small data sample corresponding to only $36\,\mathrm{pb}^{-1}$ is quite impressive. This amount of data is meanwhile taken in one or two days.

7.5. Outlook on the analysis of data taken in 2011

Currently an analysis of the data taken by the LHCb detector in the first half of 2011 is underway. The integrated luminosity collected during this time correspond to nearly $400\,\mathrm{pb}^{-1}$. To illustrate what uncertainties can be expected for this data an analysis on simulated events is performed. Simulated events corresponding to $400\,\mathrm{pb}^{-1}$ are generated according to the fitted parameter values in Table 7.1. Tagging performance and time resolution are assumed to be identical to the values found for the data taken in 2010. The same is assumed for the acceptance effects and the background contribution. The confidence contours in the ϕ_s-$\Delta\Gamma_\mathrm{s}$ parameter space determined by the likelihood ratio method are given in Figure 7.9. Compared to the nominal result for 2010 data the confidence contours are much smaller. The two minima resulting from the inherent symmetry of the PDF are well separated for all three contours shown. Finding a similarly large deviation from the Standard Model prediction in the 2011 data would constitute an unambiguous sign of physics processes beyond the Standard Model.

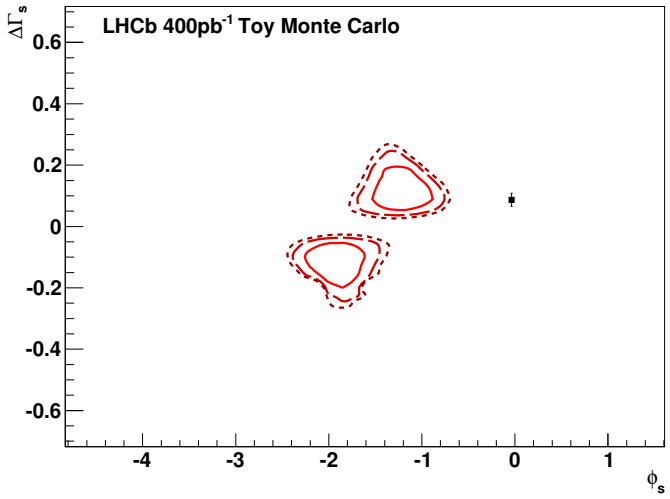


Figure 7.9.: Confidence contours in the ϕ_s-$\Delta\Gamma_s$ parameter space determined by the likelihood ratio method. Simulated events corresponding to an integrated luminosity of $400\,\mathrm{pb}^{-1}$ and generated according to the parameters given in Table 7.1 are used. The confidence contours correspond to confidence levels of 68.3% (solid), 90% (dashed) and 95% (dotted) respectively. An observation of similar contours on data would constitute an unambiguous sign of physics processes beyond the Standard Model.

8. Summary and conclusion

A method to determine the CP violating phase ϕ_s using the decay $B_s^0 \to J/\psi\,\phi$ was developed and is described in this thesis. The extraction method is implemented as a multidimensional unbinned maximum likelihood fit. The fit uses the reconstructed B_s^0 mass for background discrimination and the decay angles to separate CP-even and CP-odd decay amplitudes. Together with the reconstructed B_s^0 decay time this enables the multidimensional fit to perform a measurement of time dependent CP violation. Both flavor blind and B_s^0 production flavor dependent studies are possible with the latter having a significantly higher sensitivity to ϕ_s. The developed algorithm takes the decay time resolution into account. It also incorporates acceptance effects due to the detector geometry and selection, both for the decay time as well as for the decay angles. Multiple methods to account for the acceptance effects for both decay angles and decay time were implemented. To accurately describe the data also the background contribution was modeled. A sophisticated parameterization of the angular distributions of the background component was implemented.

The extraction method was tested on fully simulated data and successfully reproduces the physics parameters used in the generation of the simulated events. In addition a fast generation method to produce large amounts of simulated events quickly was implemented to ensure that the fitted parameter values and parameter errors are estimated correctly.

The developed unbinned maximum likelihood fit was used to analyze the data collected by the LHCb detector during the 2010 data taking period. A flavor blind analysis was performed to extract $\Delta\Gamma_s$, the decay width difference of the B_s^0 mass eigenstates. $\Delta\Gamma_s$ was determined to be $\Delta\Gamma_s = (0.084 \pm 0.112_{\text{stat.}} \pm 0.014_{\text{syst.}})\ \text{ps}^{-1}$. Systematic effects that were studied include the proper time resolution, angular and proper time acceptance, modeling of the reconstructed B_s^0 mass and the angular background description as well as a possible S-wave contribution. The extracted value for $\Delta\Gamma_s$ is in excellent agreement with the current best measurement performed by the CDF experiment, $\Delta\Gamma_s = (0.075 \pm 0.035_{\text{stat.}} \pm 0.01_{\text{syst.}})\,\text{ps}^{-1}$ [39].

Using information on the B_s^0 production flavor the extraction method implemented for this thesis was used to determine confidence regions for ϕ_s and $\Delta\Gamma_s$. The resulting confidence contours in the ϕ_s-$\Delta\Gamma_s$ parameter space are given in Figure 8.1 and constitute the main result of this thesis. They were determined using an advanced statistical method proposed by Feldman and Cousins [74] which guarantees correct coverage of the determined confidence regions. The probability of the Standard Model prediction is found to correspond to 1.0 standard deviations (68.5% confidence level). The systematic effect of the flavor tagging calibration is already included in this nominal result. Other systematic effects that were studied include the proper time and angular acceptance, the B_s^0 signal mass model and the angular acceptances as well as a possible S-wave contribution. They are found to have a negligible effect on the confidence regions

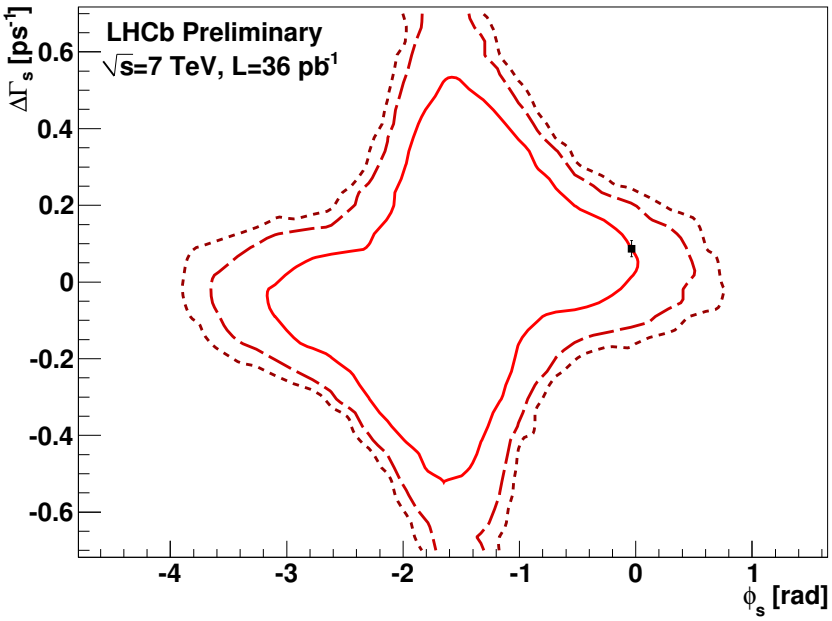


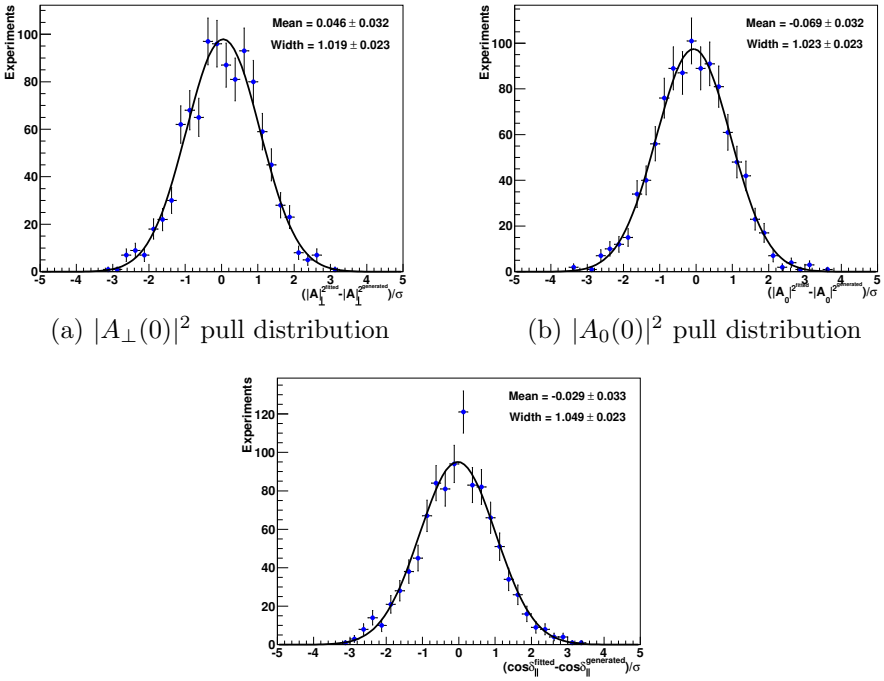
Figure 8.1.: Confidence contours in the ϕ_s-$\Delta\Gamma_s$ parameter space determined by the Feldman-Cousins method using tagging information on the B_s^0 production flavor. The confidence contours correspond to confidence levels of 68.3% (solid), 90% (dashed) and 95% (dotted) respectively. The black dot denotes the Standard Model prediction of $(-0.0363\,\text{rad}, 0.087\,\text{ps}^{-1})$.

compared to the large statistical uncertainty. In addition to the two dimensional confidence regions given by Figure 8.1 one-dimensional confidence intervals for ϕ_s are determined. The resulting confidence intervals for ϕ_s are given by $[-2.78, -0.39]\,\text{rad}$ and $[-3.33, 0.21]\,\text{rad}$ and correspond to confidence levels of 68.3% and 90% respectively.

The confidence regions given in Figure 8.1 are compatible with the latest measurements performed by the CDF and DØ collaborations which are given in Figure 1.7. All experiments favor negative values for ϕ_s. The confidence regions determined by LHCb are somewhat larger than the contours extracted by the CDF and DØ collaborations and therefore not competitive yet.

Currently (as of August 2011) a large effort is underway to analyze the data taken by the LHCb detector in the first half of 2011 which corresponds to an integrated luminosity of nearly $400\,\text{pb}^{-1}$. The extraction method developed for this thesis is one of the main algorithms that are used in the analysis which will be the worlds best measurement of ϕ_s and $\Delta\Gamma_s$.

A. Fit validation

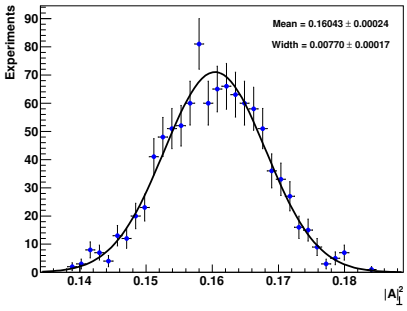


(a) $|A_\perp(0)|^2$ pull distribution

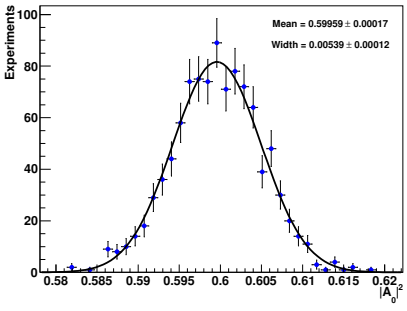
(b) $|A_0(0)|^2$ pull distribution

(c) $\cos\delta_\parallel$ pull distribution

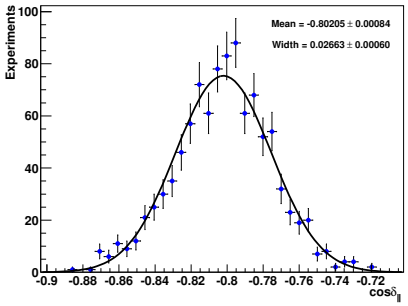
Figure A.1.: Pull distributions for the flavor blind *toy study* with no CP violation. 1000 *toy* data sets are generated and fit each corresponding to an integrated luminosity of $2\,\mathrm{fb}^{-1}$.



(a) $|A_\perp(0)|^2$ fitted parameter values

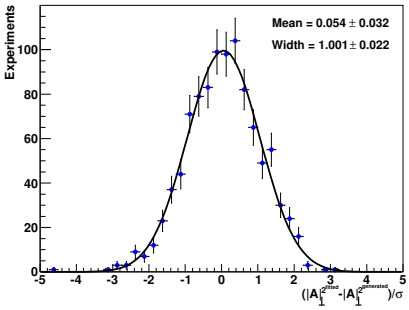


(b) $|A_0(0)|^2$ fitted parameter values

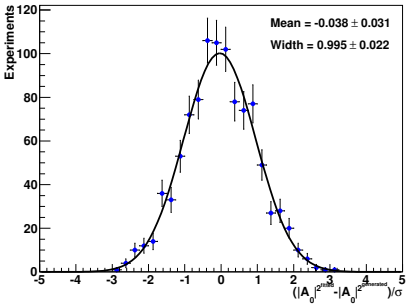


(c) $\cos\delta_\parallel$ fitted parameter values

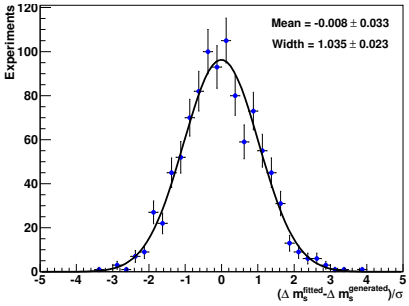
Figure A.2.: Distributions of the fitted parameter values for the flavor blind *toy study* with no CP violation. 1000 *toy* data sets are generated and fit each corresponding to an integrated luminosity of $2\,\mathrm{fb}^{-1}$.



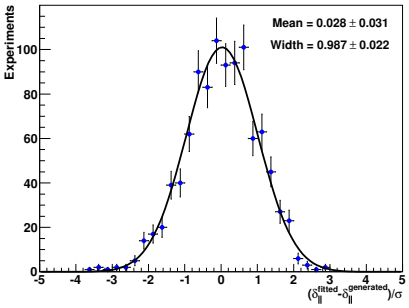
(a) $|A_\perp(0)|^2$ pull distribution



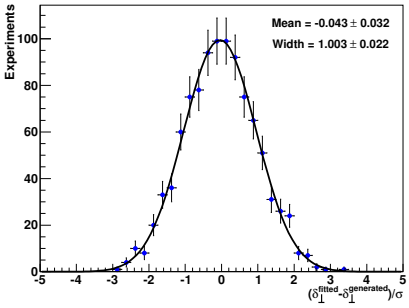
(b) $|A_0(0)|^2$ pull distribution



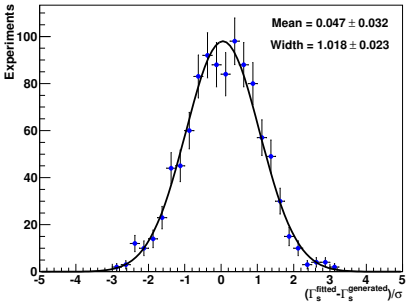
(c) Δm_s pull distribution



(d) $\delta_\parallel$ pull distribution

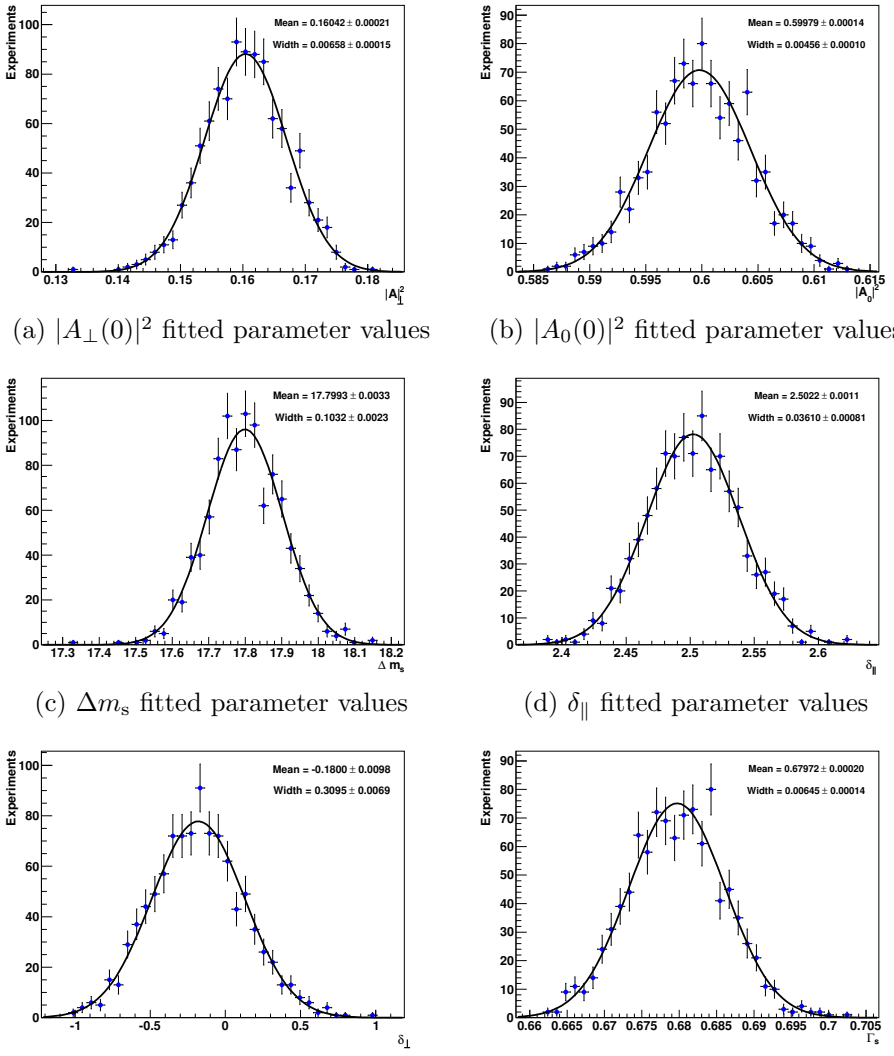


(e) $\delta_\perp$ pull distribution



(f) Γ_s pull distribution

Figure A.3.: Pull distributions for the Standard Model scenario ($\phi_s = -0.0363\,\mathrm{rad}$). 1000 *toy* data sets are generated and fit each corresponding to an integrated luminosity of $2\,\mathrm{fb}^{-1}$.



(a) $|A_\perp(0)|^2$ fitted parameter values

(b) $|A_0(0)|^2$ fitted parameter values

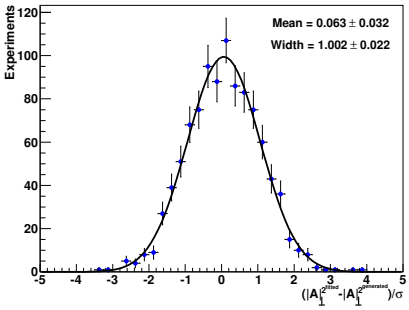
(c) Δm_s fitted parameter values

(d) $\delta_\parallel$ fitted parameter values

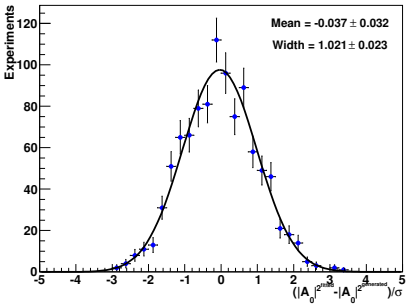
(e) $\delta_\perp$ fitted parameter values

(f) Γ_s fitted parameter values

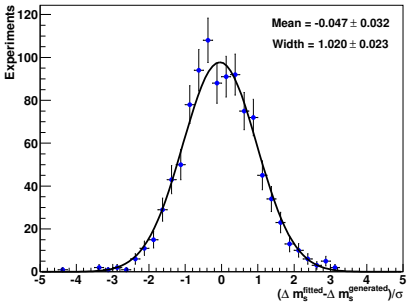
Figure A.4.: Distributions of the fitted parameter values for the Standard Model scenario ($\phi_s = -0.0363\,\mathrm{rad}$). 1000 *toy* data sets are generated and fit each corresponding to an integrated luminosity of $2\,\mathrm{fb}^{-1}$.



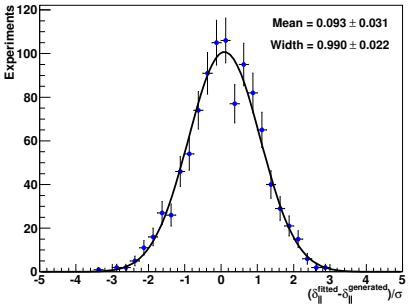
(a) $|A_\perp(0)|^2$ pull distribution



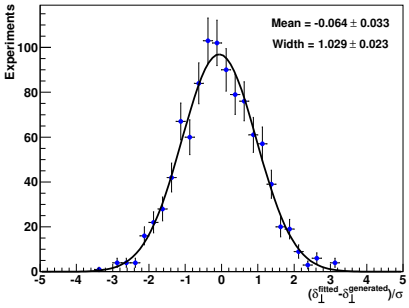
(b) $|A_0(0)|^2$ pull distribution



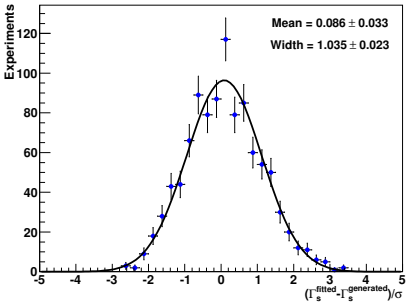
(c) Δm_s pull distribution



(d) $\delta_\parallel$ pull distribution

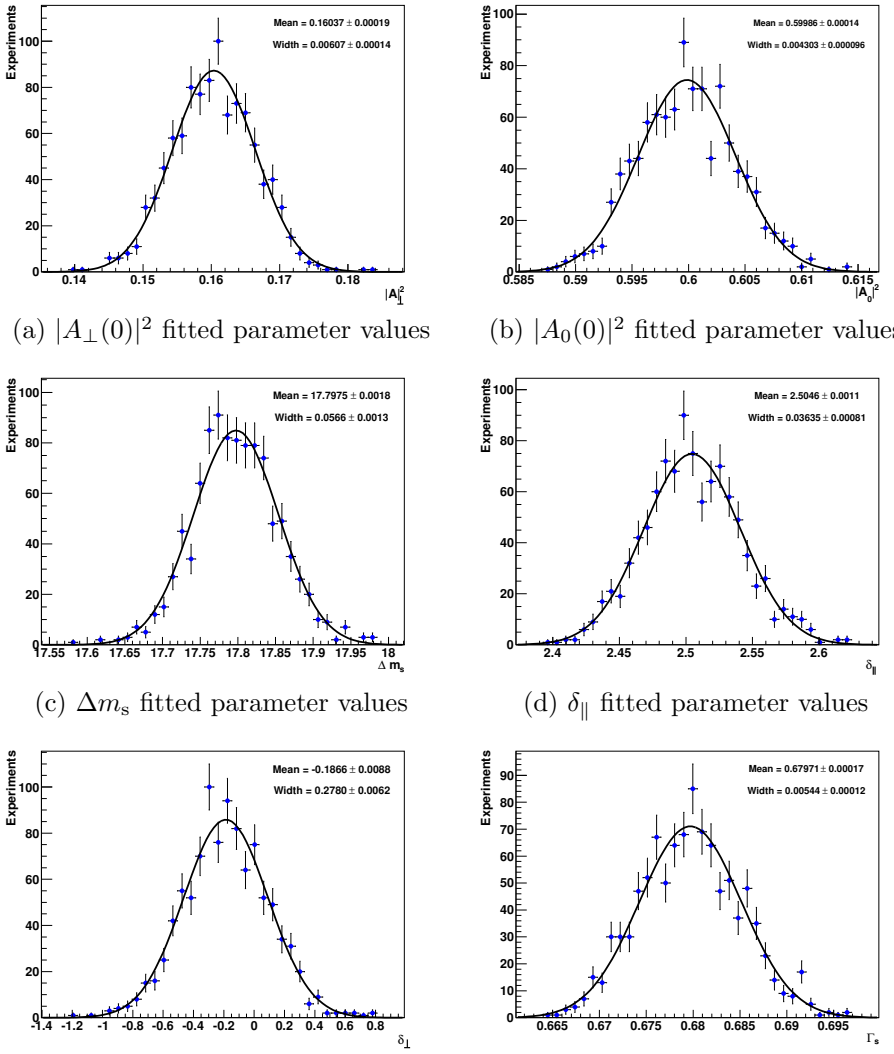


(e) $\delta_\perp$ pull distribution



(f) Γ_s pull distribution

Figure A.5.: Pull distributions for the New Physics scenario ($\phi_s = -0.70\,\mathrm{rad}$). 1000 *toy* data sets are generated and fit each corresponding to an integrated luminosity of $2\,\mathrm{fb}^{-1}$.



(a) $|A_\perp(0)|^2$ fitted parameter values

(b) $|A_0(0)|^2$ fitted parameter values

(c) Δm_s fitted parameter values

(d) $\delta_\parallel$ fitted parameter values

(e) $\delta_\perp$ fitted parameter values

(f) Γ_s fitted parameter values

Figure A.6.: Distributions of the fitted parameter values for the New Physics scenario ($\phi_s = -0.70\,\mathrm{rad}$). 1000 *toy* data sets are generated and fit each corresponding to an integrated luminosity of $2\,\mathrm{fb}^{-1}$.

	Γ_s	$\Delta\Gamma_s$	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\cos\delta_\parallel$
Γ_s	0	0.02	0.02	0.03	0.02
$\Delta\Gamma_s$		0	0.01	0.02	0.02
$\lvert A_\perp\rvert^2$			0	0.03	0.02
$\lvert A_0\rvert^2$				0	0.01
$\cos\delta_\parallel$					0

(a) RMS of the correlations for the untagged fit validation study.

	Γ_s	$\Delta\Gamma_s$	ϕ_s	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel$	$\delta_\perp$	Δm_s
Γ_s	0	0.02	0.10	0.01	0.02	0.02	0.02	0.02
$\Delta\Gamma_s$		0	0.10	0.01	0.02	0.02	0.02	0.03
ϕ_s			0	0.08	0.07	0.01	0.14	0.21
$\lvert A_\perp\rvert^2$				0	0.02	0.01	0.02	0.02
$\lvert A_0\rvert^2$					0	0.01	0.02	0.02
$\delta_\parallel$						0	0.02	0.01
$\delta_\perp$							0	0.07
Δm_s								0

(b) RMS of the correlations for the tagged fit validation study using the Standard Model scenario.

	Γ_s	$\Delta\Gamma_s$	ϕ_s	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\delta_\parallel$	$\delta_\perp$	Δm_s
Γ_s	0	0.08	0.11	0.05	0.05	0.02	0.08	0.03
$\Delta\Gamma_s$		0	0.10	0.05	0.05	0.02	0.13	0.04
ϕ_s			0	0.07	0.07	0.01	0.06	0.08
$\lvert A_\perp\rvert^2$				0	0.03	0.01	0.08	0.03
$\lvert A_0\rvert^2$					0	0.01	0.07	0.02
$\delta_\parallel$						0	0.02	0.01
$\delta_\perp$							0	0.10
Δm_s								0

(c) RMS of the correlations for the tagged fit validation study using the New Physics scenario.

Table A.1.: RMS of the correlations for (a) the untagged fit validation study, (b) the tagged fit validation study using the Standard Model scenario and (c) the tagged fit validation study using the New Physics scenario.

B. Flavor blind fit

Parameter	Result	Abs. deviation
Γ_s	0.6795 ± 0.0344	-0.0002
$\Delta\Gamma_s$	0.084 ± 0.113	-
$\lvert A_\perp \rvert^2$	0.2758 ± 0.0572	-0.0028
$\lvert A_0 \rvert^2$	0.5321 ± 0.0403	-0.0002
$\cos\delta_\parallel$	-1.210 ± 0.270	$+0.027$

(a)

Parameter	Result	Abs. deviation
Γ_s	0.6796 ± 0.0343	-0.0001
$\Delta\Gamma_s$	0.084 ± 0.113	-
$\lvert A_\perp \rvert^2$	0.2780 ± 0.0572	-0.0006
$\lvert A_0 \rvert^2$	0.5304 ± 0.0403	-0.0019
$\cos\delta_\parallel$	-1.227 ± 0.272	$+0.010$

(b)

Parameter	Result	Abs. deviation
Γ_s	0.6796 ± 0.0342	-0.0001
$\Delta\Gamma_s$	0.085 ± 0.112	$+0.001$
$\lvert A_\perp \rvert^2$	0.2792 ± 0.0573	$+0.0006$
$\lvert A_0 \rvert^2$	0.5302 ± 0.0403	-0.0021
$\cos\delta_\parallel$	-1.233 ± 0.273	$+0.004$

(c)

Parameter	Result	Abs. deviation
Γ_s	0.6798 ± 0.0342	$+0.0001$
$\Delta\Gamma_s$	0.084 ± 0.112	-
$\lvert A_\perp \rvert^2$	0.2798 ± 0.0573	$+0.0012$
$\lvert A_0 \rvert^2$	0.5309 ± 0.0404	-0.0014
$\cos\delta_\parallel$	-1.232 ± 0.275	$+0.005$

(d)

Table B.1.: Systematic uncertainties due to the description of the angular acceptance effect. (a)-(d) give the systematic effect due to the limited Monte Carlo statistics. To estimate the effect the Monte Carlo is varied within its statistical uncertainty four times and the largest deviation is taken as systematic uncertainty.

	Γ_s	$\Delta\Gamma_s$	$\lvert A_\perp\rvert^2$	$\lvert A_0\rvert^2$	$\cos\delta_\parallel$	m_{B_s}	$\sigma_{m,1}^{sig}$	α_m^{Pr}	α_m^{LL}	f_{sig}	f_{Pr}	$f_{\tau,1}^{LL}$	τ_1^{LL}	τ_2^{LL}	$f_{t,1}^{sig}$	$f_{t,2}^{sig}$	$\sigma_{t,1}^{sig}$	$\sigma_{t,2}^{sig}$	$\sigma_{t,3}^{sig}$
Γ_s	1.00	-0.46	0.38	-0.33	-0.02	0.02	0.03	-	-	0.13	0.03	-	-	0.02	-	-	-	-	-
$\Delta\Gamma_s$		1.00	-0.64	0.66	-0.05	-0.01	-0.02	-	-	-	-	-	-	-	-	-	-	-	-
$\lvert A_\perp\rvert^2$			1.00	-0.57	-0.38	-	0.01	-	-	-	-	-	-	-	-	-	-	-	-
$\lvert A_0\rvert^2$				1.00	-0.11	-0.01	-0.03	-	-	-0.02	-0.01	-0.01	-	-	-	-	-	-	-
$\cos\delta_\parallel$					1.00	0.02	0.01	-	-	0.02	0.01	-	-	-0.01	-	-	-	-	-
m_{B_s}						1.00	0.05	-	0.01	0.02	0.01	-	-	-	-	-	-	-	-
$\sigma_{m,1}^{sig}$							1.00	-	0.01	0.09	0.02	0.06	-0.02	-0.08	-	-	-	-	-
α_m^{Pr}								1.00	-0.22	-	0.04	-	0.03	0.01	-	-	-	0.01	-
α_m^{LL}									1.00	-	-0.01	-	-0.02	-0.01	-0.01	-0.01	-0.01	-0.02	-0.01
f_{sig}										1.00	0.04	0.01	-0.01	-0.01	-	-	-	-	-
f_{Pr}											1.00	-0.10	0.71	0.15	-0.03	-0.25	-0.01	0.01	-0.10
$f_{\tau,1}^{LL}$												1.00	0.28	0.63	0.01	0.07	0.01	0.01	-
τ_1^{LL}													1.00	0.38	-0.02	-0.15	-0.01	0.01	-0.16
τ_2^{LL}														1.00	-	-0.01	0.01	0.01	-0.02
$f_{t,1}^{sig}$															1.00	0.25	0.96	0.96	0.54
$f_{t,2}^{sig}$																1.00	0.16	0.44	0.81
$\sigma_{t,1}^{sig}$																	1.00	0.88	0.46
$\sigma_{t,2}^{sig}$																		1.00	0.66
$\sigma_{t,3}^{sig}$																			1.00

Table B.2.: Correlations for the untagged fit over the full proper time range. Correlations larger than 0.50 are given in bold, correlations smaller than 0.005 are omitted.

List of Figures

List of Tables

Bibliography

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

[2] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[3] Abdus Salam and John Clive Ward. Electromagnetic and weak interactions. *Phys. Lett.*, 13:168–171, 1964.

[4] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[5] A. D. Sakharov. Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967.

[6] Joshua Frieman, Michael Turner, and Dragan Huterer. Dark Energy and the Accelerating Universe. *Ann.Rev.Astron.Astrophys.*, 46:385–432, 2008. arXiv:astro-ph/0803.0982.

[7] M. Tegmark et al. Cosmological constraints from the SDSS luminous red galaxies. *Phys. Rev. D*, 74(12):123507, Dec 2006. arXiv:astro-ph/0608632.

[8] Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.*, 405:279–390, 2005. arXiv:hep-ph/0404175.

[9] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys.Rev.Lett.*, 13:321–322, 1964.

[10] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

[11] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13:585–587, 1964.

[12] Guido Altarelli. The Standard Model of Particle Physics. *To appear in the "Encyclopedia of Mathematical Physics", Elsevier*, (CERN-PH-TH/2005-206), 2005. arXiv:hep-ph/0510281v1.

[13] Guido Altarelli. A QCD Primer. *Lectures given at the 2001 European School of High-Energy Physics, Beatenberg, Switzerland and at the Pan American Advanced Studies Institute on "New States of Matter in Hadronic Interactions", January 2002,Campos do Jordao, Brazil*, (CERN-TH/2002-022), 2000. arXiv:hep-ph/0204179v1.

[14] Guido Altarelli. The Standard Electroweak Theory and Beyond. *Lectures given at the Nathiagali Summer School, Pakistan and the PSI Zuoz Summer School*, (CERN-TH/2000-291), 2000. arXiv:hep-ph/0011078v1.

[15] Paul Langacker. Introduction to the Standard Model and Electroweak Physics. *Lectures presented at TASI2008. 45 pages, 14 figures, 5 tables*, 2009. arXiv:hep-ph/0901.0241v1.

[16] C. Amsler et al. (Particle Data Group). Particle Data Book 2008. *Physics Letters B667*, 2008.

[17] D. Martin F. Halzen. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley and Sons, 1984.

[18] Nicola Cabbibo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–533, 1963.

[19] Toshihide Maskawa Makoto Kobayashi. CP-Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49(2):652–657, 1973.

[20] Wai-Yee Keung Ling-Lie Chau. Comments on the Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, 53:18021805, 1984.

[21] Lincoln Wolfenstein. Parametrization of the Kobayashi-Maskawa Matrix. *Phys. Rev. Lett.*, pages 1945–1947, 1983.

[22] P.F. Harrison et al. (BABAR Collaboration). *The BABAR physics book: Physics at an asymmetric B factory*. 1998. Papers from Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting), SLAC-504.

[23] J.H. Christenson et al. *Phys. Rev. Lett.*, 13:138, 1964.

[24] B. Aubert et al. (BABAR Collaboration). Observation of CP Violation in the B^0 Meson System. *Phys. Rev. Lett. 87, 091801*, 2001. arXiv:hep-ex/0107013.

[25] K. Abe et al. (Belle Collab.). Observation of Large CP Violation in the Neutral B Meson System. *Phys. Rev. Lett. 87, 091802*, 2001. arXiv:hep-ex/0107061.

[26] Argus Collaboration (H. Albrecht et al.). Observation of B0 - anti-B0 Mixing. *Phys.Lett.B*, 192, 1987.

[27] K. Nakamura et al. (Particle Data Group). The Review of Particle Physics. *J. Phys. G 37, 075021*, 2010.

[28] CDF Collaboration (A. Abulencia et al.). Observation of B_s^0 - $\bar{B}_s^0$ Oscillations. *Phys.Rev.Lett.*, 97, 2006. arXiv:hep-ex/0609040.

[29] The LHCb Collaboration. Measurement of $\Delta m_{\rm s}$ in the decay $\rm B_s^0 \to D_s(K^+K^-\pi^-)\pi^+$ and $\rm B_s^0 \to D_s(K^+K^-\pi^-)\pi^+\pi^+\pi^-$. LHCb-CONF-2011-005.

[30] K. Anikeev et al. B Physics at the Tevatron: Run II and Beyond. *FERMILAB-Pub-01/197*, 2002. arXiv:hep-ph/0201071.

[31] Ulrich Nierste Isard Dunietz, Robert Fleischer. In Pursuit of New Physics with B_s Decays. *Phys.Rev. D*, (63), 2001. arXiv:hep-ph/0012219.

[32] Yosef Nir. CP Violation - A New Era. (WIS/18/01-Aug-DPP), 2001. arXiv:hep-ph/0109090.

[33] U. Uwer G. Hiller. *Physics at the Terascale*, pages 163–186. WILEY-VCH, 2011.

[34] U. Nierste A. Lenz. Numerical updates of lifetimes and mixing parameters of B mesons. *Proceedings of the CKM workshop 2010 in Warwick*, 2010. arXiv:hep-ph/1102.4274.

[35] The LHCb Collaboration. Tagged time-dependent angular analysis of $B_s \to J/\psi\,\phi$ decays with the 2010 LHCb data. LHCb-CONF-2011-006.

[36] A. Dighe et al. Angular distributions and lifetime differences in $\mathrm{B_s^0} \to \mathrm{J\!/\!\psi}\,\phi$ decays. *Phys.Lett. B*, 369:144–150, 1996. arXiv:hep-ph/9511363.

[37] Yuehong Xie et. al. Determination of $2\beta_\mathrm{s}$ in $\mathrm{B_s^0} \to \mathrm{J\!/\!\psi} K^+K^-$ Decays in the Presence of a K^+K^- S-Wave Contribution. *JHEP 0909*, (074), 2009. arXiv:hep-ph/0908.3627.

[38] F. Azfar et. al. Formulae for the Analysis of the Flavor-Tagged Decay $\mathrm{B_s^0} \to \mathrm{J\!/\!\psi\phi}$. *JHEP 1011*, (158), 2010. arXiv:hep-ph/1008.4283.

[39] The CDF Collaboration. An Updated Measurement of the CP Violating Phase $\beta_s^{J/\psi\,\phi}$ in $B_s \to J/\psi\,\phi$ Decays using $5.2\,\mathrm{fb}^{-1}$ of Integrated Luminosity, 2010. CDF public note 10206.

[40] The DØ Collaboration. Updated measurement of the CP-violating phase ϕ_s using flavor-tagged decay $\mathrm{B_s^0} \to \mathrm{J\!/\!\psi}\,\phi$, 2010. Note 6098-CONF.

[41] R. Harnik et al. Atmospheric Neutrinos Can Make Beauty Strange. *Phys.Rev.D*, (69), 2004. arXiv:hep-ph/0212180v3.

[42] U. Nierste et al. A. Lenz. Anatomy of New Physics in $B - \bar{B}$ mixing. *Phys.Rev.D83*, 83, 2011. arXiv:hep-ph/1008.1593.

[43] S. Mrenna T. Sjöstrand and P. Skands. PYTHIA 6.4: Physics and manual. *JHEP 0605:026*, 2006. arXiv:hep-ph/0603175.

[44] LHCb collaboration. Measurement of J/ψ production in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$. *Eur.Phys.J.*, (C71):1645, 2011. arXiv:hep-ex/1103.0423.

[45] J. Nardulli. Reconstruction of two-body B decays in LHCb. *CERN-THESIS-2007-063*, 2007.

[46] CERN. Geneva. LHC Experiments Committee. LHCb : Technical Proposal. Technical Report CERN-LHCC-98-004, CERN, 1998.

[47] Michelangelo L. Mangano G. Altarelli. 1999 CERN Workshop on standard model physics (and more) at the LHC, CERN. *CERN-2000-004, May 2000. 546pp.*, 1999.

[48] Lyndon Evans and Philip Bryant (editors). LHC Machine. *JINST*, 3:S08001, 2008.

[49] Brüning et. al. LHC Design Report Volume II, The LHC Infrastructure and General Services. CERN-2004-003-V-2. - Geneva, http://lhc.web.cern.ch/lhc/.

[50] The LHCb Collaboration and A Augusto Alves Jr et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[51] The LHCb Collaboration. LHCb reoptimized detector design and performance : Technical Design Report, 2003. CERN-LHCC-2003-030 ; LHCb-TDR-9.

[52] A. Berkien et al. The LHCb Outer Tracker Front End Electronics. 2005. LHCb-2005-025 ; CERN-LHCb-2005-025.

[53] D. Breton and D. Charlet. SPECS: The serial protocol for the Experiment Control System of LHCb. 2003. LHCb-Note 2003-004.

[54] The LHCb collaboration. The Brunel project. Web page http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/.

[55] The LHCb collaboration. The DaVinci project. Web page http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/.

[56] Clemencic M. et al. Recent developments in the LHCb software framework Gaudi. *J.Phys.Conf.Ser.*, (219):042006, 2010.

[57] The LHCb collaboration. The Dirac project. Web page http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/dirac/.

[58] The LHCb collaboration. The Gauss project. Web page http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/.

[59] David J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152–155, 2001.

[60] J. Allison et al. Geant4 Developments and Applications. *IEEE Transactions on Nuclear Science*, 53 No. 1:270–278, 2006.

[61] The LHCb collaboration. The Boole project. Web page http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/.

[62] The LHCb collaboration. The Moore project. Web page http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore/.

[63] The LHCb Collaboration. b-hadron lifetime measurements with exclusive $B \to J/\psi\, X$ decays reconstructed in the 2010 data. LHCb-CONF-2011-001.

[64] Wouter D. Hulsbergen. Decay chain fitting with a Kalman filter. *Nucl. Instrum. Meth.*, A552:566–575, 2005. arXiv:physics/0503191.

[65] The LHCb Collaboration. Roadmap for selected key measurements of LHCb. 2010. LHCB-PUB-2009-029, arXiv:hep-ex/0912.4179.

[66] Francois R. Le Diberder Muriel Pivk. sPlot: a statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555:356–369, 2005. arXiv:physics/0402083.

[67] Y. Xie. sFit: a method for background subtraction in maximum likelihood fit. 2009. arXiv:physics/0905.0724.

[68] The LHCb Collaboration. Untagged angular analysis of $B_s \to J/\psi\, \phi$ and $B^0 \to J/\psi\, K^*$ with the 2010 data. LHCb-CONF-2011-002.

[69] The LHCb Collaboration. Optimization and Calibration of the Tagging performances using 2010 data. LHCb-CONF-2011-003.

[70] F. James. Minuit, Function Minimization and Error Analysis. *CERN long writeup D506.*

[71] R. J. Barlow. *Statistics, A Guide to the Use of Statistical Methods in the Physical Sciences.* John Wiley and Sons, 1989.

[72] T. Bayes. An essay towards solving a problem in the doctrine of chances. *Philos. Trans.*, 73:370–418, 1763.

[73] J. Neyman. Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. *Philos. Trans.*, 236:333–380, 1937.

[74] Robert D. Cousins Gary J. Feldman. A Unified Approach to the Classical Statistical Analysis of Small Signals. *Phys.Rev. D*, 57:3873–3889, 1998. arXiv:physics/9711021.

[75] Frank Porter. Interval Estimation using the Likelihood Function. *Nucl. Inst. Meth.*, A 368:793–803, 1996.

[76] Mike Williams. How good are your fits? Unbinned multivariate goodness-of-fit tests in high energy physics. *JINST*, 5, 2010. P09004, arXiv:hep-ex/1006.3019v2.

[77] J. von Neumann. Various techniques used in connection with random digits. *National Bureau of Standards, Applied Mathematics Series*, 12:36–38, 1951.