

MEASUREMENT OF THE LIFETIME DIFFERENCE
AND CP-VIOLATING PHASE IN $B_s \rightarrow J/\psi \phi$
DECAYS

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Physik der Universität
Karlsruhe (TH)

genehmigte

DISSERTATION

von

Michael Milnik
aus Aschaffenburg

Tag der mündlichen Prüfung: 30.11.2007

Referent: Prof. Dr. M. Feindt, Institut für Experimentelle Kernphysik

Korreferent: Prof. Dr. G. Quast, Institut für Experimentelle Kernphysik

Contents

Introduction	1
1 Theory	5
1.1 Weak Interaction and the CKM Matrix	6
1.1.1 Unitarity Triangle	8
1.2 B Meson Lifetimes	10
1.3 Time Evolution of Neutral B Mesons	12
1.3.1 B_s Decays	14
1.3.2 Untagged B_s Decays	15
1.3.3 No CP Violation in Untagged B_s Decays	16
1.3.4 Angular Analysis	16
1.4 Predictions	17
2 The Experiment	19
2.1 Accelerator Complex	19
2.1.1 Proton Production	20
2.1.2 Anti-proton Production	21
2.1.3 Main Injector	21
2.1.4 The Tevatron	22
2.2 CDF II Detector	23
2.2.1 Tracking System	25
2.2.2 Particle Identification	26
2.2.3 Calorimeters	28
3 Candidate Selection	31
3.1 Trigger System	32
3.1.1 Level-1	32
3.1.2 Level-2	35
3.1.3 Level-3	35
3.2 Offline Reconstruction	36
3.2.1 Good Run Selection	36

3.2.2	Event Reconstruction	37
3.2.3	Loose Candidate Selection	38
3.3	Pre-selection	39
3.4	Monte Carlo Simulation	39
3.5	Neural Network Selection	41
3.5.1	B_s Neural Network	42
3.5.2	B_d Neural Network	46
3.5.3	Swap Suppression Neural Network	48
4	Fit Function	53
4.1	Parameter Estimation	54
4.2	Fit Model Components	55
4.2.1	Mass	55
4.2.2	Decay Time	56
4.2.3	Decay Time Uncertainty Distribution	58
4.2.4	Angular Distributions	58
4.2.5	Angular Acceptance	59
4.2.6	Acceptance Function Correction	60
4.2.7	B_d Acceptance	66
4.3	B_s Log Likelihood Function	66
4.3.1	Case of no CP Violation	70
4.4	B_d Log Likelihood Function	70
4.5	Fit Procedure	71
5	B_d Results	73
5.1	$K\pi$ S-wave Contribution	73
5.2	Fit Verification	75
5.3	Fit Projections	80
5.4	Likelihood Profiles	80
5.5	Results	83
6	Measurement of $\Delta\Gamma$ under SM Assumption	87
6.1	Fit Validation	87
6.2	Likelihood Profiles	88
6.3	Fit Projections	89
6.4	Results	89
7	Systematic Uncertainties	99
7.1	Signal Mass Model	99
7.2	Angular Model of the Background	100
7.3	Decay Time Resolution Model	100

7.4	Cross-Feed	100
7.5	B_d Candidates with Swapped Mass Assignment	102
7.6	SVX Alignment	102
7.7	S-Wave Contribution	103
7.8	Decay Time Uncertainty Distribution	103
7.9	Summary	103
8	Measurement of ϕ_s	105
8.1	Systematic Bias	105
8.2	P-Value Definition	108
8.3	Confidence Region	109
8.4	Results	110
9	Summary and Outlook	117
A	Trigger Paths	123
B	Neural Network Input Variables	125
C	Derivation of the fit function	127
D	Normalization	131

Introduction

Over the past decades the current theoretical description, the Standard Model of elementary particle physics, was solidified by many measurements as the basic theory describing fundamental particles and their interactions. It is extremely successful in explaining the high-precision data collected by experiments so far. The Standard Model includes several intrinsic parameters which have to be measured in experiments. Independent analyses of different physical processes can constrain those parameters. By combining those measurements physicists might be sensitive to physics beyond the Standard Model. If they are inconsistent it allows to get a hint on the theory that might supersede the Standard Model.

The goal of the analysis presented in this thesis is to measure some of these parameters in the B_s meson system. The B_s meson, consisting of an anti- b and s quark, is not a pure mass eigenstate, thus allowing a B_s meson to oscillate into its antiparticle via weak interacting processes. This is a general feature of any neutral meson. The history of meson mixing measurements is more than 50 years old. It was first observed in the kaon system [1]. The oscillation in the B_d system was measured very precisely by the B factories [2, 3], whereas the oscillation frequency of the B_s was measured with more than 5σ significance last year by CDF [4] and first evidence for mixing in the D^0 system was presented only this year [5, 6].

Besides the measurement of the oscillation frequency and therefore the mass difference of the mass eigenstates, the difference in lifetime of both states is also of major interest as it is connected to CP violation. The measurement of CP violation is important since it can explain the observation that matter dominates the universe and not antimatter. One way to measure CP violation is by separating the two mass eigenstates of the B_s in common decay modes for B_s and \bar{B}_s , like in $B_s \rightarrow J/\psi \phi$. Direct CP violation and mixing induced CP violation are both very small and difficult to measure at the moment for B_s , but common decay modes allow for CP violation in interference between decay with and without mixing, which might be large enough to be measurable. Although the expected uncertainty for measuring

the CP violating phase is large, a significant deviation from the Standard Model expectation might be observed, thus indicating new physics.

The separation of the two lifetime components of the two mass eigenstates in the B_s system can be exploited to perform a precise measurement of the lifetime difference and their mean lifetime. Any lifetime measurement in the B hadron sector is interesting for comparison with other B hadrons. In first order the heavy b quark dominates the decay time for any B hadron leading to a similar lifetime for all B hadrons. Precise measurements of B hadron lifetimes showed a hierarchy, which can be understood in the framework of Heavy Quark Expansion (HQE) theory. In Chapter 1 a short introduction is given to the underlying theory for mixing and HQE.

B_s mesons are a very rare kind of particles and are currently only produced at the Tevatron, except for test runs at KEK-B and CLEO. In their default operation the center of mass energy of the B factories is not sufficient to create B_s , but only B_d mesons. Therefore the B_d sector is studied with a great precision, whereas measurements in the B_s sector have rather large uncertainties. This started to change with upper and lower boundaries on Δm_s by DØ [7] and the very precise Δm_s measurement by CDF [4] last year. The measurement of the lifetime difference of the two B_s mass eigenstates was previously done by CDF with roughly 20% of the statistics compared to the analysis presented here. Recently, DØ published their result [8], which does agree with the Standard Model expectation very well, whereas the previous CDF measurement [9] had a surprisingly large lifetime difference, but was consistent with Standard Model prediction within the large statistical uncertainty.

The good performance of the Tevatron and the CDF II detector allowed us to redo the $\Delta\Gamma$ measurement with much larger statistics. Further neural network techniques are used to improve the purity of the selection. Since the $B_s \rightarrow J/\psi \phi$ decay used in this thesis is very easy to trigger, this analysis is based on one of the largest B_s meson collections so far and allows to make one of the most precise lifetime and lifetime difference measurements of the B_s . The accelerator complex, CDF II detector and the selection process will be presented in Chapters 2 and 3.

The lifetime is measured by utilizing the decay time distribution of reconstructed events. The expected B_s lifetime difference is of the order of ten percent, thus the distribution is a combination of two very similar exponential functions. To separate these better, the fact that the B_s mass eigenstates are nearly CP eigenstates can be exploited. The pseudoscalar to vector-vector, $P \rightarrow VV$, decay allows to make an angular analysis, which helps to distinguish the two different mass eigenstates and measure their lifetimes. To obtain the parameters of interest from the $B_s \rightarrow J/\psi \phi$ decay a

time dependent angular analysis is presented in Chapter 4.

To cross-check the analysis the $B_d \rightarrow J/\psi K^*$ decay can be studied, which is also a $P \rightarrow VV$ decay. Although CDF is not able to compete with the lifetime measurements of the B factories due to their huge amount of B_d events, the angular analysis of this decay will have competitive sensitivity for the amplitudes and strong phases. Chapter 5 will present the angular analysis of $B_d \rightarrow J/\psi K^*$ decay.

The B_s result assuming no CP violation will be presented in Chapter 6. In Chapter 7 the evaluation of the systematic uncertainties for both decays is summarized. Chapter 8 will present the results for the B_s decay including the CP violating phase. The last chapter will then summarize and discuss the results. It will also provide an outlook into the near future of similar studies.

Chapter 1

Theory

This chapter focuses on the theoretical understanding of the B meson decay process described in this thesis. As underlying theory we use the Standard Model of elementary particle physics. It describes the properties of elementary particles and the interactions between them.

There are six types of quarks and six different leptons plus their anti-particles. They are all fermions, typically grouped into three families. Their main properties are listed in Table 1.1. Leptons are spin-1/2 particles which interact weakly and, in case they have electric charge, also electromagnetically. Quarks have in addition color charge and therefore also participate in the strong interaction. In the Standard Model the interactions are mediated by gauge bosons, associated with the symmetries of the model

$$\begin{aligned}SU_C(3) &\rightarrow G_\mu^\alpha, \alpha = 1 - 8 \\SU_L(2) &\rightarrow W_\mu^\alpha, \alpha = 1, 2, 3 \\U_Y(1) &\rightarrow B_\mu,\end{aligned}$$

where G_μ^α corresponds to eight spin-1 gluons mediating strong interaction and the three W_μ^α and B_μ give rise to the four physical bosons W^\pm and Z^0 , the force carriers of the weak interaction, and γ , mediating electromagnetic interaction. The respective charges of the three symmetry groups are the weak hypercharge, weak isospin and color.

The only way for the lightest particle of each quark family to decay is via the charged current interaction of the W^\pm boson, which allows for a connection between different families. As a consequence only the quarks of the lightest family are stable and are the constituents of our world.

The B mesons, which are of interest here, are strongly bound states of a heavy anti- b quark and a light quark. Throughout this thesis we will denote the different types by B_q , where q describes the lighter quark. Usually B

name	category	symbol	el. charge [e]	mass [MeV/ c^2]
electron	lepton	e	-1	0.51
electron neutrino	lepton	ν_e	0	$\leq 2 \cdot 10^{-6}$
up quark	quark	u	$+\frac{2}{3}$	1.5 – 3.0
down quark	quark	d	$-\frac{1}{3}$	3 – 7
muon	lepton	μ	-1	105.66
muon neutrino	lepton	ν_μ	0	≤ 0.190
charm quark	quark	c	$+\frac{2}{3}$	$(1.25 \pm 0.09) \cdot 10^3$
strange quark	quark	s	$-\frac{1}{3}$	95 ± 25
tau	lepton	τ	-1	1777.0
tau neutrino	lepton	ν_τ	0	≤ 18.2
top quark	quark	t	$+\frac{2}{3}$	$(174.2 \pm 3.3) \cdot 10^3$
bottom quark	quark	b	$-\frac{1}{3}$	$(4.20 \pm 0.09) \cdot 10^3$

Table 1.1: The elementary fermionic particles (spin $s = \frac{1}{2}$) with their electric charge in units of the electron charge and their mass.

mesons are described by the spectator model, where the light quark is the spectator. The lightest B meson can decay only via the weak interaction, which is dominated by the $b \rightarrow Wq$ process, with a virtual W boson and a c or u quark. This allows to access the V_{cb} and V_{ub} elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements, which describes the transition between the quarks types. Further, due to higher order processes, B meson decays are sensitive also to V_{td} and V_{ts} , elements of the CKM matrix.

The determination of the individual parts of the CKM matrix is one of the major tasks of contemporary particle physics. The following sections will present how the observables, we are interested in, are measured in the experiment and can be related to the Standard Model parameters.

1.1 Weak Interaction and the CKM Matrix

As mentioned above, we are interested in the charged current weak interaction. The part of the electro-weak Lagrangian describing the charged current weak interaction can be written in terms of the mass eigenstates of quarks as

$$\mathcal{L}_{cc} = -\frac{g_2}{\sqrt{2}}(\bar{u}_L, \bar{c}_L, \bar{t}_L)\gamma^\mu V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W_\mu^\dagger + h.c., \quad (1.1)$$

where V_{CKM} is the 3×3 transformation matrix between the mass and weak eigenstates. The Transformation can be written as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (1.2)$$

where the primed quarks are the weak eigenstates and the non-primed quarks are mass eigenstates. Since the elements of the CKM matrix are complex numbers, the matrix contains $2 \cdot 3^2$ real parameters. Constrains on the CKM matrix lower the actual number of free parameters. In the Standard Model the CKM matrix has to satisfy the following constrains:

- The unitarity of CKM matrix. This means $V_{CKM}^\dagger V_{CKM} = 1$, which reduces the number of free parameters by nine.
- The Lagrangian is invariant under one global phase rotation, which reduces the number of free parameters by one.
- Finally there are four relative phases which can be removed by redefinition of the quark field. This reduces the number of free parameters describing the CKM matrix by four.

Usually the remaining four parameters are interpreted as three rotational angles and a complex phase. The current standard notation for the CKM matrix is then

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (1.3)$$

where $c_{ij} = \cos \Theta_{ij}$ and $s_{ij} = \sin \Theta_{ij}$, thus leading to the three angles $\Theta_{12}, \Theta_{23}, \Theta_{13}$ and the phase δ .

From experiments it is known that the following hierarchy holds: $s_{13} \ll s_{23} \ll s_{12} \ll 1$, which allows to make an approximation using the Wolfenstein parameterization

$$V_{CKM} \approx \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (1.4)$$

where $\lambda = \sin \Theta_{12} \approx 0.2$, $\eta = \sin \Theta_{13} \sin \delta_{13}/A\lambda^3$, $\rho = \sin \Theta_{13} \cos \delta_{13}/A\lambda^3$, $A = \sin \Theta_{23}/\lambda^2$.

1.1.1 Unitarity Triangle

The unitarity condition of the V_{CKM} can be expressed as

$$\sum_{i=1}^3 V_{ij} V_{ik}^* = \delta_{jk} = \sum_{i=1}^3 V_{ji} V_{ki}^* \quad j=1..3, k=1..3.$$

Six of these equations can be visualized as triangles in a complex plane, called unitary triangles. These equations are equivalent to the products of two out of three columns or rows. Leading to two groups of three triangles. Two of each group have the same shape, being only rotated in the complex plane. The three equations corresponding to the column products give rise to the three conditions

$$(V^\dagger V)_{21} : V_{us}^* V_{ud} + V_{cs}^* V_{cd} + V_{ts}^* V_{td} = 0 \quad (1.5)$$

$$(V^\dagger V)_{31} : V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} = 0 \quad (1.6)$$

$$(V^\dagger V)_{32} : V_{ub}^* V_{us} + V_{cb}^* V_{cs} + V_{tb}^* V_{ts} = 0. \quad (1.7)$$

The graphical representation for one of the unitary triangles is shown in Figure 1.1. Here we choose to illustrate the second condition (equation 1.6), which is often called ‘‘The Unitary Triangle’’, since it is best studied as the three sides are of the same order of magnitude ($\mathcal{O}(\lambda^3)$). This is not true for the other two unitary triangles, which are very flat and therefore more difficult to study. The angles of the triangle are defined as

$$\phi_1 = \beta = \arg \left(-\frac{V_{cb}^* V_{cd}}{V_{tb}^* V_{td}} \right) \quad (1.8)$$

$$\phi_2 = \alpha = \arg \left(-\frac{V_{tb}^* V_{td}}{V_{ub}^* V_{ud}} \right) \quad (1.9)$$

$$\phi_3 = \gamma = \arg \left(-\frac{V_{ub}^* V_{ud}}{V_{cb}^* V_{cd}} \right) \quad (1.10)$$

Different measurements of physical quantities are related to different CKM matrix elements. These parameters are not derived in the Standard Model, but are fundamental parameters which have to be obtained through experiments. One of the major goals of particle physics today is to measure these quantities. Over-constraining them will allow to check for consistent results between different types of processes. Further we can determine whether the triangles are closed, or not. In this way we test the unitarity of the CKM matrix and for new physics.

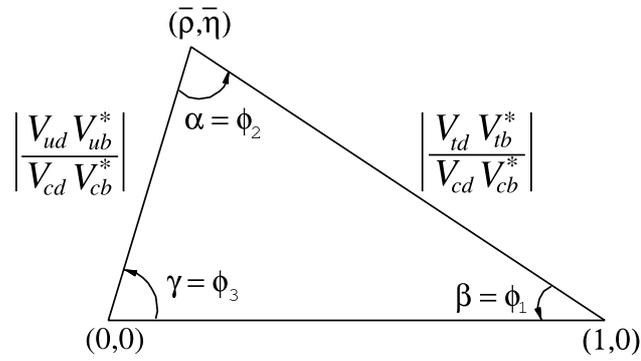


Figure 1.1: The unitarity triangle representing the product of the first and third column of the CKM matrix.

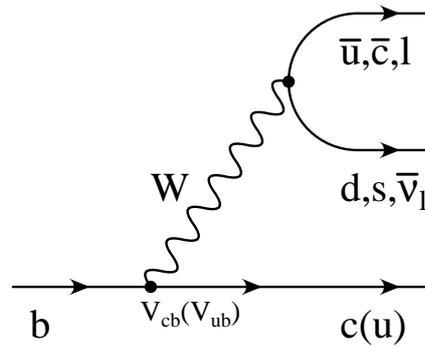


Figure 1.2: Feynman diagram describing the b quark decay in the Standard Model.

1.2 B Meson Lifetimes

The Standard Model allows the b quark to decay solely to a c or u quark via a virtual W boson, with the corresponding CKM matrix coupling elements (see Figure 1.2). The decay width for the hadronic b -decay in the Standard Model is given by

$$\Gamma_{q_1\bar{q}_2}(b \rightarrow q) = \frac{3G_F^2 m_b^5}{192\pi^3} |V_{q_1 q_2}|^2 |V_{qb}|^2 F(\epsilon_q), \quad (1.11)$$

where G_F is the Fermi coupling constant and $F(\epsilon_q)$ is the phase space factor for the decay. Thus measurements of the lifetime $\tau = 1/\Gamma$ have a direct connection to the fundamental Standard Model CKM parameters.

Now the issue here is, that we do not observe free b quarks, but they are always bound in hadronic states with other quarks. Due to interaction of the b quark, the lifetime of the hadron is not the same as for the free quark.

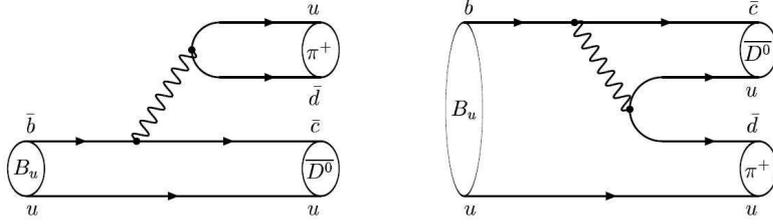
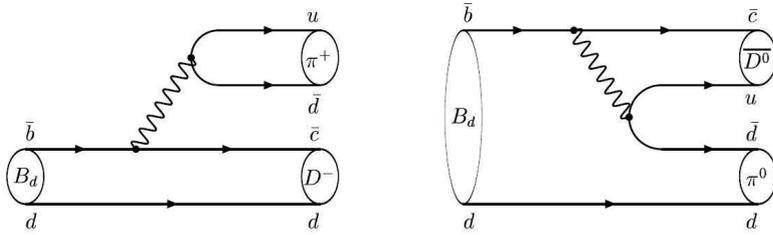
B mesons can be to first order described using the spectator model, where a heavy quark (in our case the b) is bound to the lighter “spectator” quark. In the spectator model, the heavy quark dominates the lifetime and the B meson decay can therefore be described with good precision by the decay of the free quark. A direct consequence is, that all B mesons would have the same lifetime, which does not agree with observed hierarchy

$$\tau_{B_c} < \tau_{B_s} \approx \tau_{B_d} < \tau_{B_u}. \quad (1.12)$$

The differences of the B meson lifetimes can be understood using Heavy Quark Expansion (HQE) theory of the Standard Model [10, 11, 12]. HQE is based on QCD principles, where the decay rate of the b mesons is expressed in an expansion series of $1/m_b$. Leading order calculations in HQE reproduce the common lifetime for B mesons, originating from the free b quark lifetime, as expected in the spectator model. Terms of order $1/m_b^2$ can be neglected, since they are higher order corrections associated with the b -quark alone. Therefore, they are similar for any B meson. However, terms of order $1/m_b^3$ describe effects between the two constituent quarks in the meson. For the B_d and B_u the most important processes are shown in Figures 1.3 to 1.5.

The influence of Pauli Interference (PI) on the hadron decay is depicted in Figures 1.3 and 1.4. For B_u the decay products of the external and color suppressed internal weak W decay are the same and allow for interference. In contrast, the decays of the B_d have different final states which does not allow for such an effect. Since the interference for the B_u is destructive, its lifetime gets larger compared to decays without Pauli Interference.

A second type of contribution, which exist only for some mesons, is the Weak Annihilation (WA) possible only for charged B mesons and allows for

Figure 1.3: Feynman diagrams describing Pauli Interference for B_u mesons.Figure 1.4: The similar Feynman diagrams for B_d mesons compared to B_u (see Figure 1.3). Since the final state is different there is no Pauli Interference of the two B_d decay modes.

an additional way to decay, which reduces the lifetime for B_u and B_c (see figure 1.5).

Weak Exchange (WE) on the other hand exists only for neutral B mesons and baryons (see figure 1.5). However, the Weak Exchange decay for mesons is helicity suppressed, since the cd (cs) spin is defined by the zero spin of the B_d (B_s). For the $q\bar{q}$ pair the helicity has to be opposite, which suppresses this decay mode for mesons.

In general in HQE the hard physics processes are summed up in the so called Willson Coefficients. They can be calculated using perturbation theory. The remaining soft process must be calculated by other means. This allows to make predictions of the different B hadron lifetimes. This technique can also be used for c hadron lifetime predictions, although the expansion is done for $1/m_c$ instead of $1/m_b$, which is worse.

The current status of theory and experiments show quantitatively good agreement. For this thesis especially the theoretical predictions which include B_s and B_d lifetimes [13, 10] are of interest

$$\tau_{B_u}/\tau_{B_d} = 1.053 \pm 0.023 \quad (1.13)$$

$$\tau_{B_s}/\tau_{B_d} = 1.00 \pm 0.01. \quad (1.14)$$

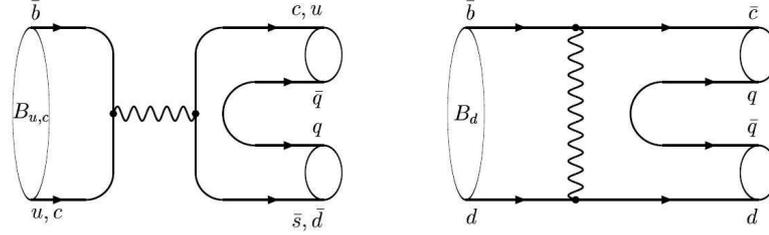


Figure 1.5: Feynman diagrams describing the Weak Annihilation (left) and Weak Exchange (right). Weak Annihilation is an additional way to decay for charged B mesons, while Weak Exchange contributes only to neutral B meson and baryon decays.

From the discussion above it is clear that there is no process which would cause a significant difference between the B_d and B_s lifetimes. Therefore, we expect to obtain a similar lifetime for B_s and B_d mesons. The B_d lifetime is very precisely measured by the B factories and will allow us to check whether our measurement of the B_s lifetime is consistent with it.

1.3 Time Evolution of Neutral B Mesons

The b quark can be combined to a neutral B meson with an s or d quarks, forming either a B_s or B_d meson. Charged B mesons are formed with u and c quarks. In this chapter we discuss the time evolution of neutral B_d and B_s mesons. The theoretical background is common for both mesons and we will just say B meson until we go into specific details of the B_s meson state.

Without any weak interaction there would be no connection between the $|B\rangle$ and $|\bar{B}\rangle$ states. However, weak interaction enables quarks to change its flavor, which allows for transitions between the two states. This behaviour is called B meson mixing [14, 15]. Although we introduced this for B mesons, mixing is allowed for any neutral meson state. This process can be nicely illustrated using Feynman diagrams (see Figure 1.6 for the lowest order mixing diagrams).

Due to the mixing any pure state will evolve into a quantum superposition, $a|B\rangle + b|\bar{B}\rangle$, after a while. Assuming the Wigner-Weisskopf approximation, the time evolution of the strong interacting $B - \bar{B}$ eigenstates is described by a Schrödinger equation $i\psi = \begin{pmatrix} a \\ b \end{pmatrix}$,

$$i\frac{\partial}{\partial t}\psi = (\mathbf{M} - \frac{i}{2}\mathbf{\Gamma})\psi, \quad (1.15)$$

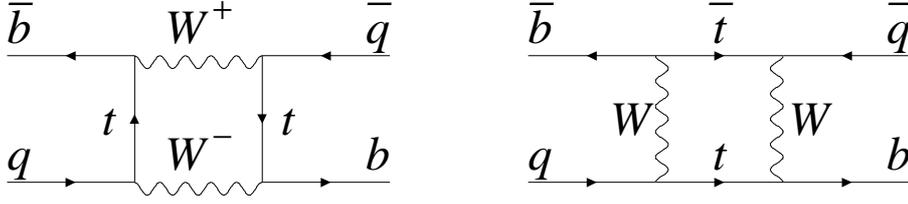


Figure 1.6: Lowest order Feynman diagrams illustrating B meson mixing. Due to their shape they are often called box diagrams. The t quark can be replaced by u and c , but the t quark dominates the loop process, since the amplitude is proportional to the square of the quark mass.

where

$$(\mathbf{M} - \frac{i}{2}\mathbf{\Gamma}) = \begin{pmatrix} m_{11} - \frac{i}{2}\Gamma_{11} & m_{12} - \frac{i}{2}\Gamma_{12} \\ m_{12}^* - \frac{i}{2}\Gamma_{12}^* & m_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}. \quad (1.16)$$

CPT invariance requires $m_{11} = m_{22} = m$ and $\Gamma_{11} = \Gamma_{22} = \Gamma$. The heavy and light mass eigenstate can be written as

$$|B_{L,H}\rangle = p|B^0\rangle \pm q|\bar{B}^0\rangle, \quad |p|^2 + |q|^2 = 1. \quad (1.17)$$

The eigenvalues are

$$\lambda_{L,H} = (m - \frac{i}{2}\Gamma) \pm \frac{q}{p}(M_{12} - \frac{i}{2}\Gamma_{12}) \quad (1.18)$$

with

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} = \frac{V_{tb}V_{td}^*}{V_{tb}^*V_{td}}. \quad (1.19)$$

The time evolution of the mass eigenstates are then

$$|B_{L,H}(t)\rangle = |B_{L,H}(0)\rangle e^{-i\lambda_{L,H}t} = |B_{L,H}(0)\rangle e^{-iM_{L,H}t - \frac{1}{2}\Gamma_{L,H}t}, \quad (1.20)$$

where $M_{L,H} = \text{Re}(\lambda_{L,H})$ and $\Gamma_{L,H} = -2\text{Im}(\lambda_{L,H})$. This can be translated into the time evolution of the flavor eigenstates

$$|B(t)\rangle = g_+(t)|B\rangle + \frac{q}{p}g_-(t)|\bar{B}\rangle \quad (1.21)$$

$$|\bar{B}(t)\rangle = \frac{p}{q}g_-(t)|B\rangle + g_+(t)|\bar{B}\rangle, \quad (1.22)$$

where $g_{\pm}(t) = \frac{1}{2}(e^{-i\lambda_L t} \pm e^{-i\lambda_H t})$.

Now we can calculate the time-dependent decay rates for flavour specific modes

$$|\langle B|\bar{B}(t)\rangle|^2 = \left|\frac{p}{q}\right|^2 |g_-(t)|^2 \quad (1.23)$$

$$|\langle \bar{B}|B(t)\rangle|^2 = \left|\frac{q}{p}\right|^2 |g_-(t)|^2 \quad (1.24)$$

$$|\langle B|B(t)\rangle|^2 = |g_+(t)|^2 \quad (1.25)$$

$$|\langle \bar{B}|\bar{B}(t)\rangle|^2 = |g_+(t)|^2, \quad (1.26)$$

where

$$|g_{\pm}(t)|^2 = \frac{1}{2} \left[\cosh\left(\frac{\Delta\Gamma t}{2}\right) \pm \cos(\Delta M t) \right] e^{-\Gamma t},$$

$$\Delta M = M_H - M_L,$$

$$\Delta\Gamma = \Gamma_L - \Gamma_H.$$

1.3.1 B_s Decays

For the understanding of the CP-violation in the B_s decays we follow the derivation from reference [15]. From the time evolution we get the following decay rates for the two CP final states f and \bar{f}

$$\begin{aligned} \Gamma[B_s(t) \rightarrow f] &= |A_f|^2 \left\{ |g_+(t)|^2 + |\lambda_f|^2 |g_-(t)|^2 + 2\text{Re}[\lambda_f g_+^*(t) g_-(t)] \right\}, \\ \Gamma[B_s(t) \rightarrow \bar{f}] &= |\bar{A}_{\bar{f}}|^2 \left| \frac{q}{p} \right|^2 \left\{ |g_-(t)|^2 + |\lambda_{\bar{f}}^{-1}|^2 |g_+(t)|^2 + 2\text{Re}[\lambda_{\bar{f}}^{-1} g_+(t) g_-^*(t)] \right\}, \\ \Gamma[\bar{B}_s(t) \rightarrow f] &= |A_f|^2 \left| \frac{p}{q} \right|^2 \left\{ |g_-(t)|^2 + |\lambda_f|^2 |g_+(t)|^2 + 2\text{Re}[\lambda_f g_+(t) g_-^*(t)] \right\}, \\ \Gamma[\bar{B}_s(t) \rightarrow \bar{f}] &= |\bar{A}_{\bar{f}}|^2 \left\{ |g_+(t)|^2 + |\lambda_{\bar{f}}^{-1}|^2 |g_-(t)|^2 + 2\text{Re}[\lambda_{\bar{f}}^{-1} g_+^*(t) g_-(t)] \right\}. \end{aligned} \quad (1.27)$$

In case when the decay amplitudes are dominated by a single weak phase. Then

$$|A_f| = |\bar{A}_{\bar{f}}|, \quad |A_{\bar{f}}| = |\bar{A}_f|,$$

and

$$\lambda = \frac{q \langle f|\bar{B}_s(t)\rangle}{p \langle f|B_s(t)\rangle}, \quad \bar{\lambda} = \frac{p \langle \bar{f}|\bar{B}_s(t)\rangle}{q \langle \bar{f}|B_s(t)\rangle}.$$

λ can be expressed in terms of related CKM matrix elements. Using the definition $\beta_s = \arg(-\frac{V_{cd}V_{cb}^*}{V_{tb}^*V_{ts}})$ it translates to

$$\lambda = \frac{q \langle f | \bar{B}_s(t) \rangle}{p \langle f | B_s(t) \rangle} = \frac{V_{tb}V_{ts}^* V_{cs}V_{cb}^*}{V_{tb}^*V_{ts} V_{cs}^*V_{cb}} = \frac{V_{cd}V_{cb}^* V_{tb}V_{td}^*}{V_{tb}^*V_{ts} V_{cs}^*V_{cb}}, \quad (1.28)$$

$$\text{Re}(\lambda) = \text{Re} \left(\frac{V_{cd}V_{cb}^* V_{tb}V_{td}^*}{V_{tb}^*V_{ts} V_{cs}^*V_{cb}} \right) = \cos(2\beta_s), \quad (1.29)$$

$$\text{Im}(\lambda) = \text{Im} \left(\frac{V_{cd}V_{cb}^* V_{tb}V_{td}^*}{V_{tb}^*V_{ts} V_{cs}^*V_{cb}} \right) = \sin(2\beta_s). \quad (1.30)$$

If we assume no CP violation in decay

$$|\langle f | B_s \rangle| = |\langle \bar{f} | \bar{B}_s \rangle| \quad (1.31)$$

and no CP violation in mixing

$$|q/p| \neq 1, \quad (1.32)$$

so that CP violation only comes in via interference between mixing and decay

$$|\langle f | \bar{B}_s \rightarrow B_s \rangle + \langle f | B_s \rangle| \neq |\langle \bar{f} | B_s \rightarrow \bar{B}_s \rangle + \langle \bar{f} | \bar{B}_s \rangle| \quad (1.33)$$

Since the final state $|f\rangle$ and $|\bar{f}\rangle$ with $|\bar{f}\rangle = CP|f\rangle$, are CP eigenstates, $\lambda_{fCP} = \eta_f = \pm 1$. Equation (1.27) can then be rewritten for this case as follows

$$\begin{aligned} \Gamma[B_s(t) \rightarrow f] &= |A_f|^2 e^{-\Gamma_s t} \{ \cosh(\Delta\Gamma_s t/2) \\ &\quad - \cos(2\beta_s) \sinh(\Delta\Gamma_s t/2) - \sin(2\beta_s) \sin(\Delta m_s t) \}, \\ \Gamma[B_s(t) \rightarrow \bar{f}] &= |A_f|^2 e^{-\Gamma_s t} \{ \cosh(\Delta\Gamma_s t/2) \\ &\quad + \cos(2\beta_s) \sinh(\Delta\Gamma_s t/2) + \sin(2\beta_s) \sin(\Delta m_s t) \}, \\ \Gamma[\bar{B}_s(t) \rightarrow f] &= |A_f|^2 e^{-\Gamma_s t} \{ \cosh(\Delta\Gamma_s t/2) \\ &\quad - \cos(2\beta_s) \sinh(\Delta\Gamma_s t/2) + \sin(2\beta_s) \sin(\Delta m_s t) \}, \\ \Gamma[\bar{B}_s(t) \rightarrow \bar{f}] &= |A_f|^2 e^{-\Gamma_s t} \{ \cosh(\Delta\Gamma_s t/2) \\ &\quad + \cos(2\beta_s) \sinh(\Delta\Gamma_s t/2) - \sin(2\beta_s) \sin(\Delta m_s t) \}. \end{aligned} \quad (1.34)$$

1.3.2 Untagged B_s Decays

For an untagged sample, meaning that we do not attempt to identify the b quark flavor of the B meson at production time, the formulas described

above change slightly. Assuming that the same number of B_s and \bar{B}_s are produced, the decay rates for $B_s(t)$ and $\bar{B}_s(t)$ into a given final state are added

$$\begin{aligned}\Gamma[\overset{(-)}{B}_s(t) \rightarrow f] &= |A_f|^2 e^{-\Gamma_s t} \{ \cosh(\Delta\Gamma_s t/2) - \cos(2\beta_s) \sinh(\Delta\Gamma_s t/2) \}, \\ \Gamma[\overset{(-)}{B}_s(t) \rightarrow \bar{f}] &= |A_f|^2 e^{-\Gamma_s t} \{ \cosh(\Delta\Gamma_s t/2) + \cos(2\beta_s) \sinh(\Delta\Gamma_s t/2) \}.\end{aligned}\tag{1.35}$$

The terms depending on the mass difference Δm_s cancel each other out. But an untagged analysis still allows to determine the lifetime difference $\Delta\Gamma$ and the weak phase β_s .

1.3.3 No CP Violation in Untagged B_s Decays

In the Standard Model the approximation of no CP Violation for B_s ($\beta_s = 0$) is close to the expectation. Looking at $B_{s(L,H)}$ we see that for $\frac{q}{p} = -e^{-2i\beta_s}$ the two mass eigenstates are CP eigenstates. Thus any information about the CP state when the B_s decays can help to separate the two mass eigenstates. For the $B_s \rightarrow J/\psi \phi$ decay the angular distribution provides such information as described in the next section.

In contrast for B_d , where $|\beta| \gg 0$, $\frac{q}{p} \neq 1$, the mass eigenstates are no CP eigenstates. Therefore, it is not used to separate the $B_{d(L,H)}$. Since the lifetime difference for B_d is expected to be small, we fit only the mean lifetime for B_d and set $\Delta\Gamma = 0$.

1.3.4 Angular Analysis

Equation 1.35 describes the behaviour of a decay to a single CP eigenstates. The $B_s \rightarrow J/\psi \phi$ decay studied in this analysis is a mixture of CP eigenstates. To be more precise, the decay products are two vector mesons, which can have three independent final states. These three states can be written in terms of different angular momenta $L = 0, 1, 2$. Alternatively we can write the decay amplitudes in components longitudinal or transverse to the direction of motion of the vector meson, which are perpendicular or parallel to each other. These states have different amplitudes $A_0, A_\perp, A_\parallel$ ($|A_0|^2 + |A_\perp|^2 + |A_\parallel|^2 = 1$), where the first and third are CP even final state and the second is CP odd. All three angular momentum states have different polarizations, which determine their angular distributions.

We analyze the angular distributions in the coordinate system of the J/ψ rest frame, where the ϕ direction defines the x-axis and the KK decay plane

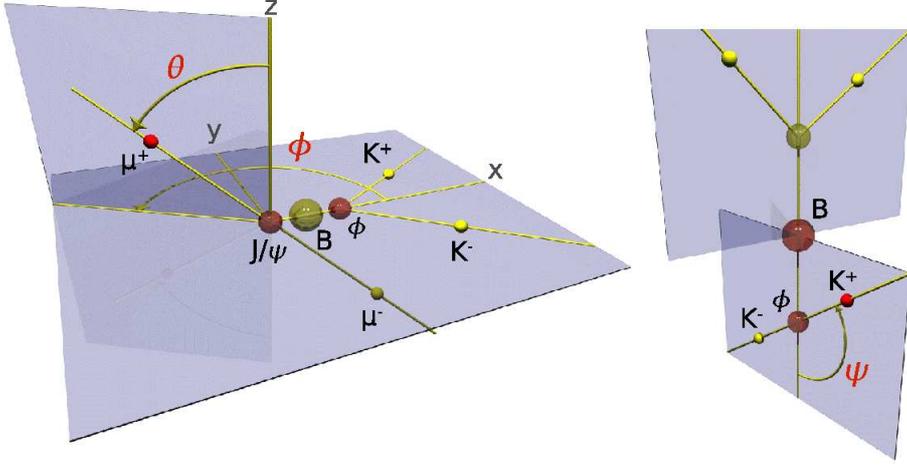


Figure 1.7: Definition of the angles describing the $B_s \rightarrow J/\psi \phi$ decay.

the y-axis, with $p_Y(K^+) > 0$. The z-axis is perpendicular to the xy-plane. The decay can be described using the following three different angles (see figure 1.7) [16]:

- Θ : angle between the μ^+ and the z-axis in the J/ψ rest frame.
- ϕ : angle between the μ^+ projection on the xy-plane and the x-axis in the J/ψ rest frame.
- Ψ : angle between the K^+ and the negative B direction in the ϕ rest frame.

By studying the angular distribution, it is possible to statistically distinguish the different CP eigenstates of the B_s .

1.4 Predictions

In this analysis the two untagged decays $B_s \rightarrow J/\psi \phi$ and $B_d \rightarrow J/\psi K^*$ are analysed. We want to measure the lifetime of the B_d and B_s . Using an angular analysis we will also be able to separate the two CP eigenstates for B_s , allowing us to measure the lifetime difference and CP violating phase $\phi_s = 2\beta_s$. This will allow us to check HQE predictions for the lifetimes and will help to constrain the CKM matrix.

Current theoretical predictions and measurements for the most interesting parameters are

Quantity	Prediction	
τ_{B_u}/τ_{B_d}	1.053 ± 0.023	[13]
τ_{B_s}/τ_{B_d}	1.00 ± 0.01	[10]
τ_{B_d}	$458.7 \mu\text{m}$	[17]
$\Delta\Gamma$	$0.096 \pm 0.039 \text{ ps}^{-1}$	[18]
ϕ_s	$0.24^\circ \pm 0.08^\circ$	[18]

Chapter 2

The Experiment

Since the B meson of interest do not occur in our environment they have to be produced artificially for study. They are created in high energy particle collisions. Accelerators are used to obtain those high energy particles, and complex detectors are used for studying their collisions. This thesis uses data collected by the CDF II (Collider Detector at Fermilab) detector, which is used to reconstruct the decay products of the collision. The Fermilab Tevatron particle accelerator complex provides high energetic proton– anti-proton collisions. Both are situated at the Fermi National Accelerator Laboratory, Fermilab, located in Batavia near Chicago, Illinois (USA).

The Tevatron started with RUN I, where first collisions were produced and detected in 1985 at a center-of-mass energy of $\sqrt{s} = 1.8$ TeV. CDF is a general purpose detector, which has taken data with the integrated luminosity of 90 pb^{-1} in the first eleven years of operation.

Since the shutdown in 1996, the Tevatron and its detectors CDF and DØ have undergone major upgrades for RUN II, which started at the end of 2001. The analysis presented here uses up to 1.7 fb^{-1} of data taken during RUN II.

2.1 Accelerator Complex

The Fermilab accelerator is an advanced accelerator complex, which is schematically illustrated in Figure 2.1. It consists of several systems to obtain protons, create anti-protons, accelerate both types of particles and produce collisions at the interaction points inside the detectors.

FERMILAB'S ACCELERATOR CHAIN

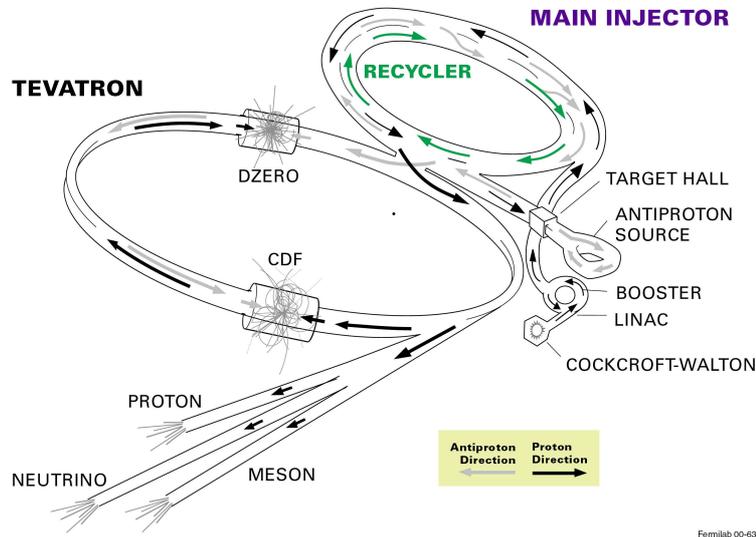


Figure 2.1: Fermilab's accelerator complex for RUN II.

2.1.1 Proton Production

The first stage of acceleration is achieved using a Cockcroft-Walton pre-accelerator. Negatively charged hydrogen ions are produced with electric discharges into a hydrogen gas. The electrostatic field in the Cockcroft-Walton chamber accelerates the negative ions. After filtering out anything except the H^- , they are introduced into a 150 m long linear accelerator (Linac) [19], where they are accelerated from an energy of approximately 750 keV using drift tubes to reach 116 MeV followed by RF cavities, which allow for the final energy of 400 MeV.

At the entry to the Booster [20, 21], a synchrotron with a diameter of about 150 m, the negatively charged 400 MeV hydrogen ions are directed through a carbon foil, which strips off the electrons, leaving bare protons. In the Booster the protons are accelerated to an energy of 8 GeV using RF cavities [22].

The last acceleration stage before entering the Tevatron is taking place in the Main Injector (MI) [20].

2.1.2 Anti-proton Production

Compared to protons, anti-protons are difficult to obtain and their production is the main limiting factors for the Tevatron performance. In the previous chapter it was described how to get protons into the Main Injector. For the anti-proton production protons are accelerated to 120 GeV in the Main Injector and are sent to a nickel target. The collisions produces a shower of secondary particles, which is focused by a lithium lens [23]. Using magnetic fields, 8 GeV anti-protons are selected and sent to the Debuncher [20, 23]. Here the momentum spread of the beam is reduced by extending it in space by bunch rotation and adiabatic de-bunching is performed [20, 23].

Since the anti-proton production rate is low several techniques are used to minimize their loss. In general the anti-proton beam should be confined to a small phase space volume, which is equivalent to a cold beam. To achieve this, stochastic [24, 25] and electron cooling are used [26, 27, 28].

From the Debuncher the anti-protons are extracted into the Accumulator [20, 23], which is a synchrotron with a mean radius of 75 m. Here, as the name implies, the anti-protons can be collected for hours at an energy of 8 GeV and are continuously cooled down using stochastic cooling.

The anti-protons are transferred to the Recycler [20, 29], which is located in the same tunnel as the Main Injector. Originally the Recycler was meant to accept and store the anti-protons after a colliding store (period of colliding beams in the Tevatron) for the next cycle along with new anti-protons from the Antiproton Source. These plans did not work out and the Recycler now stores the anti-protons allowing to operate the Accumulator at its best efficiency and cools the anti-protons further using stochastic cooling as well as electron cooling.

2.1.3 Main Injector

The Main Injector is designed to operate in different modes for different purposes. As mentioned before it is capable of accepting protons from the Booster and anti-protons from the Recycler and Accumulator. As the final pre-acceleration step both beam types can be accelerated to 150 GeV for the injection into the Tevatron. Besides anti-proton production, the proton beam with an energy of 120 GeV is also used for other experiments at the Fermilab.

2.1.4 The Tevatron

The final acceleration is achieved by the Tevatron [20], a collider with a circumference of about six kilometers. There, the protons and anti-protons get their final energy of 0.98 TeV each, which provides a center of mass energy of 1.96 TeV. Both beams share the same beam pipe but circulate in opposite directions. Until LHC startup, which is planned for the near future, it is the highest energy hadron collider in the world. It is the first accelerator using super-conducting magnets, necessary to achieve such high energies for the given radius.

The beam is split in three trains, each containing 12 bunches. The gaps between the trains allow the kicker magnets to ramp up in order to abort the beam by dumping it into graphite beam dumps. Once the Tevatron is loaded with 36 proton and 36 anti-proton bunches the particles are accelerated to the maximum energy. Starting with 150 GeV protons and anti-protons the acceleration to 980 GeV takes only a few seconds. Having two times 36 bunches circulating in opposite directions there are 72 regions where bunch crossing occur. Two of these regions are placed close to the center of the two detectors, CDF II and DØ. At these special two points additional effort is made to maximize the number of collisions of protons and anti-protons by focusing both beams using quadrupole magnets. On the other 70 crossing regions the interaction is minimized by having proton beam circulating on one strand of a helix and the anti-protons on the other.

At the interaction regions the approximate instantaneous luminosity is given by

$$L = \frac{nfN_pN_{\bar{p}}}{4\pi\sigma_x\sigma_y},$$

where n is the number of bunches, f the revolution frequency, $N_p(N_{\bar{p}})$ is the number of protons (anti-protons) per bunch and σ_x and σ_y are the average transverse widths of the bunches. The instantaneous luminosity decreases over time since particles are lost and the long range interaction at the crossing regions heat up the beam, leading to an increase in σ_x and σ_y . By integrating L over time we get the amount of collected data.

Figure 2.2 shows the development of the initial luminosity over time, whereas Figure 2.3 shows the integrated luminosity from the beginning of RUN II until recently. The mean efficiency of data taking is $> 80\%$. The data used for this analysis corresponds to an integrated luminosity of up to 1.7 fb^{-1} .

As was shown so far, to create collisions at this high center of mass energy a lot of effort has to be made. To make inferences of the physical processes,

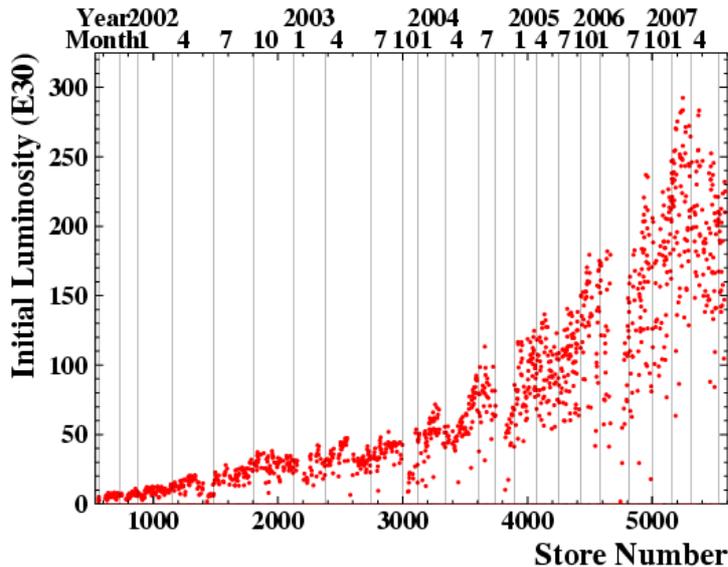


Figure 2.2: The peak initial luminosity in RUN II over time.

detectors have to be build to analyse the outcomes of the interactions.

2.2 CDF II Detector

The CDF II detector [30] is a solenoidal general purpose detector located at the interaction point B0 of the Tevatron. It is build to allow for very broad physics program, which go from relatively low energetic charm and bottom physics to high energy top and Higgs physics as well as direct searches for physics beyond Standard Model

The detector scheme is shown in Figure 2.4. It is designed to have an azimuthal and forward-backward symmetry. The inner part consists of a tracking system for charged particles, followed by the time-of-flight detector, calorimeters and muon detectors on the outside.

The detector is described using a coordinate system where the polar angle θ is measured from the positive z direction (proton direction, which points east at the location of the CDF II detector), the azimuthal angle ϕ is measured from the Tevatron plane. Instead of θ the Lorentz invariant pseudo-rapidity η is commonly used, $\eta = -\ln(\tan(\frac{\theta}{2}))$.

The following short description focuses on the central region only, $|\eta| < 1$, and does not explain anything about the end cap detectors. It uses information of the central tracking system and particle identification from the energy

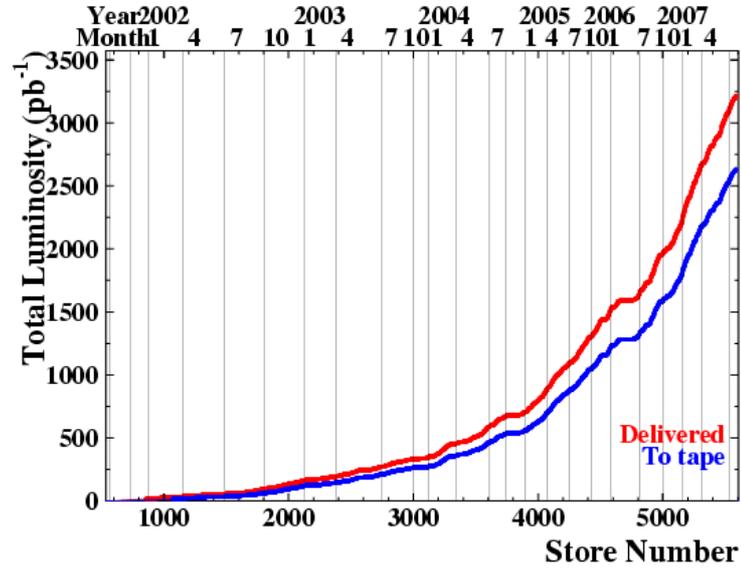


Figure 2.3: Integrated luminosity during RUN II over time. In red the delivered luminosity by the Tevatron and in blue the recorded luminosity by CDF II detector.

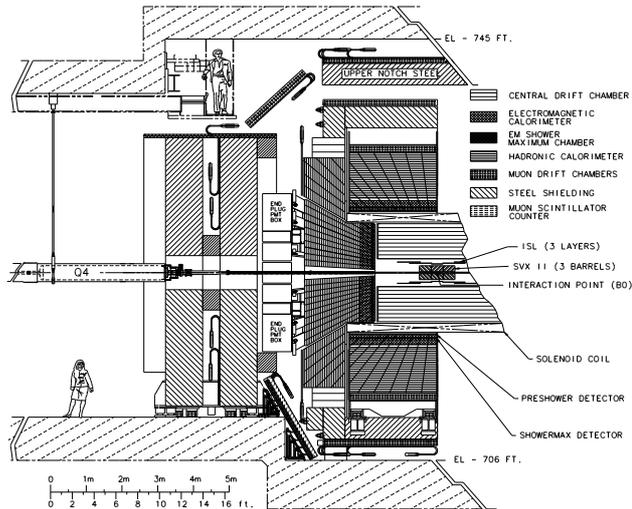


Figure 2.4: Elevation view of one half of the CDF II detector.

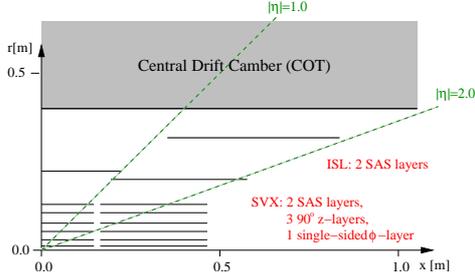


Figure 2.5: R/z cross view of the inner tracking system.

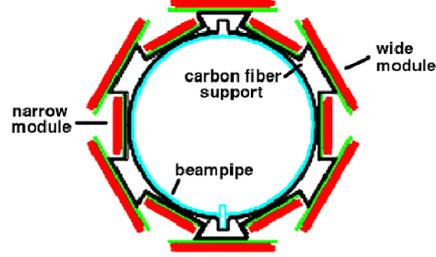


Figure 2.6: Schematic view of Layer 00.

deposit in the drift chamber, time-of-flight and muon detectors.

2.2.1 Tracking System

The tracking system of the CDF II detector is placed inside a super-conducting solenoid, which provides a uniform magnetic field of up to 1.4 T along the detector axis to bend the tracks of charged particles to helices.

All charged particles ionize the passed material. This ionization happens near the trajectory of the particle. By using detectors which are capable of measuring ionization, the path of a charged particle can be reconstructed from the deposited energy. This process is usually called tracking.

The CDF tracking system consists of two different detector systems. Next to the beam pipe is a silicon micro strip detector, which is surrounded by an open-cell drift chamber. The silicon detector provides the necessary resolution near the interaction point, where the track density is rather large. The drift chamber on the other hand provides excellent tracking at larger radius.

Silicon Detector

The silicon detector itself consists of several sub-detectors. For a schematic view see Figure 2.5. On the outside it starts with the Intermediate Silicon Layers (ISL) which covers a radial range of 20 to 28 cm. It has an acceptance in $|\eta| < 1.9$. It is followed by the Silicon Vertex Detector (SVX II), which covers a radial range of 2.5 to 10.7 cm with an $|\eta|$ acceptance up to 2. And finally the Layer 00 (L00) at a radius of about 1.5 cm.

For a precise lifetime measurement of B mesons with a typical lifetime of $\approx 450\mu\text{m}$, one needs a very good track resolution near the beam pipe. Layer 00 [31], which is mounted directly on the beam pipe, provides a measurement as close as possible to the interaction point. The whole silicon detector consists of radiation hard micro strip detectors.

Layer 00 consists of two overlapping single-sided layers of silicon at radii 1.35 cm and 1.62 cm with different widths (see Figure 2.6). In z it has a total length of 94 cm, which corresponds to $|\eta| < 4$.

The Silicon Vertex Detector [32] is separated in three cylindrical barrels, each 29 cm long. Each barrel consists of 5 layers of double-sided silicon strip detectors located at radii between 2.5 and 10.7 cm. Layers 0 (innermost), 1 and 3 combine a $r - \phi$ measurement on one side with a 90° stereo ($r - z$) on the other side. Layers 2 and 4 (outermost) combine an $r - \phi$ measurement with a second measurement rotated not by 90° but only by 1.2° (small angle stereo, SAS).

The outermost silicon detector system is the Intermediate Silicon Layers detector [33] (see Figure 2.5). It consists of a single layer at a radius of 22 cm in the central region of $|\eta| < 1$, and two layers in the plug region ($1 < |\eta| < 2$) at radii of 20 cm and 29 cm. The whole Intermediate Silicon Layers detector is made of double sided small angle stereo micro strip detectors, but only every second strip is read out to reduce the amount of data. This affects the single hit resolution compared to the silicon detector, which is constructed in the same way.

The whole silicon detector provides an impact parameter resolution of about $40 \mu\text{m}$, which already includes the $30 \mu\text{m}$ contribution from the beam width. The z_0 resolution is roughly $70 \mu\text{m}$. Although it is possible to reconstruct tracks using only the silicon detector, the tracking quality is greatly improved by using tracks that were already found in the drift chamber.

Drift Chamber

The second tracking system is a cylindrical multi-wire open cell drift chamber called Central Outer Tracker (COT) [34] which covers the radii from 44 to 132 cm and a region $|\eta| \leq 1$. The design goal for the COT was to ensure that the maximum drift time is less than the 132 ns bunch spacing. It is designed to operate at 100 ns by reducing the maximum drift distance and using a gas mixture with a large drift velocity (50:50 Argon-Ethane). Although it has a much poorer direction and position resolution ($140 \mu\text{m}$) than the Inner Tracker, it provides a much better momentum resolution ($\sigma(p_T)/p_t^2 = 0.0015c/\text{GeV}$) thanks to the larger radial extension, and a higher purity due to lower track density compared to the silicon detector.

2.2.2 Particle Identification

There are different types of stable particles created during collisions. The attribute stable is used for any kind of particle with a lifetime long enough to

pass through the detector. These are protons, kaons, pions, muons, electrons and neutrinos. Only the track of the charged ones can be measured. For the analysis it is very helpful to distinguish those kinds of particles.

The CDF II detector has several systems dedicated to particle identification. In the central region there are from the inside outwards the Time of Flight System (ToF) [35], electromagnetic and hadronic calorimeter and the muon chambers. Additionally the energy deposit in the drift chamber is used for particle separation.

For the analysis presented here the muon chambers are crucial for the J/ψ selection. Further we need a good separation between pions and kaons to reduce the background for the $\phi \rightarrow KK$ decay and to distinguish the decay products for the $K^* \rightarrow \pi K$. The ToF provides good separation power for low energetic particle and is complemented by the energy deposit measurement in the drift chamber at higher momenta.

Time of Flight

The Time of Flight System, which was fitted in the space between the drift chamber and the solenoid (see Figure 2.7), consists of 216 scintillating bars along the length of the drift chamber. Those bars cover the full ϕ range and roughly $|\eta| < 1$. The ToF measures the arrival time T_m for a particle. This can be combined with the momentum measurement from the COT to derive the mass of the particle by using the relation

$$m = \frac{p}{c} \sqrt{\frac{(ct)^2}{L^2} - 1},$$

where L is the path length and $t = T_m - T_0$ is time difference of the measured time and the production time T_0 . The ToF time resolution is ≈ 100 ps which allows a particle separation between kaons and pions above 2σ for $p < 1.6$ GeV/c. In Figure 2.8 the separation power for the different particle types depending on the momentum is shown. It can be seen that the separation power is better for low energetic particles.

This is a powerful tool for background rejection and as such of vital importance for this analysis. The loss in separation power is complemented with the measurement of the specific energy-loss in the drift chamber.

Specific Energy-Loss (dE/dx)

When traversing the drift chamber, charged particle loss energy primarily due to ionization and atomic excitation. The energy loss is described by the Bethe-Bloch equation and depends on the velocity $\beta = v/c$ of the particle.

Thus a particle identification is possible. The amount of deposited energy is not measured directly, but is determined from the width of the pulses on the sense wire. The width Δt is logarithmically proportional to the charge deposit and therefore to the energy loss

$$\Delta t \propto \log Q \propto \frac{dE}{dx}.$$

Figure 2.8 also shows the discriminating power versus the momentum for the energy loss measurement in the drift chamber. It supplements the particle identification from the Time of Flight measurement at higher momentum reasonably well.

Muon Chambers

Muons are about 200 time heavier than electrons and thus do radiate $200^2 = 40000$ times less bremsstrahlung. Further they do not interact strongly with the atomic nuclei of the detector. Therefore muons can transverse much more material than any other charged particle. Since most other particles are absorbed by the calorimeter, muon chambers are typically mounted on the outside of the detector. This provides a very pure muon candidate selection.

The muon system [36] at the CDF II detector is a set of scintillators, drift tubes and steel absorber, used for the detection of muons above $p_t \approx 1.5 \text{ GeV}/c$. The muon systems are not able to take data within two bunch crossings, but the low occupancy of the muon chambers allows integration over several bunch crossings.

CDF II has four different muon detectors, covering a region of $|\eta| < 2$. The track segment measured in the muon chambers is matched to a track in the drift chamber to obtain a good track resolution. Due to the screening of the calorimeters, the muon identification is very pure. The remaining background consists of a few kaons and pions reaching the muon chambers, muons coming from outside and from the in-flight decay of pions to muons. Especially the last one is difficult to remove, since it is a real muon coming from the direction of the interaction point.

Table 2.1 gives an overview of the different muon systems. The CMP is an upgrade in the very central region on top of the CMU to reduce the misidentification rate, whereas the CMX extends the $|\eta|$ range up to 1.

2.2.3 Calorimeters

The information derived from the calorimeters is not used for this analysis since it is optimized for much higher energies and jets. In context of this

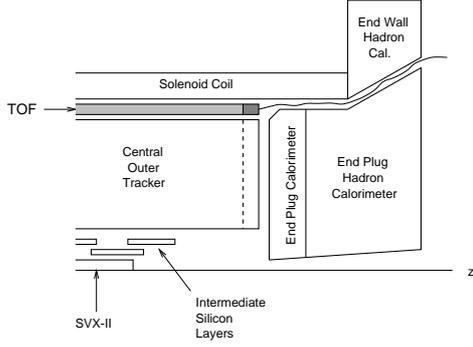


Figure 2.7: Location of the Time of Flight detector system.

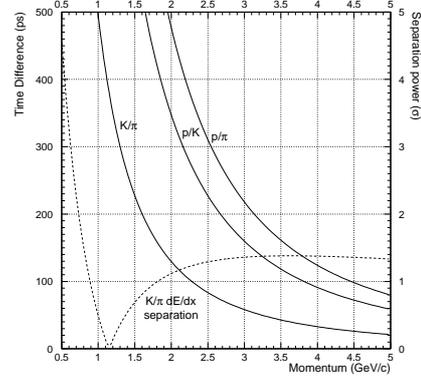


Figure 2.8: Time of flight differences as a function of particle type and momentum. The dashed line indicates the COT dE/dx separation power for K/π .

	CMU	CMP	CMX	IMU
$ \eta $	≤ 0.6	≤ 0.6	0.6 – 1.0	1.0 – 2.0
drift tube length	226	640	180	363
max. drift time [ns]	800	1400	1400	80
pion absorption length	5.5	7.8	6.2	6.2-20
min. muon pt [GeV]	1.4	2.2	1.4	1.4-2.0

Table 2.1: Design parameters of the different muon systems.

analysis it is mainly important as absorber in front of the muon chambers and is described here shortly for completeness.

The solenoid and the tracking volume are surrounded by the calorimeters, designed to measure the energy of photons, electrons and jets. Measuring everything else, it allows to determine the missing energy associated to neutrinos. There are altogether five calorimeter systems: central electromagnetic calorimeters, central hadron calorimeters, end-wall hadron calorimeters, end-plug electromagnetic and hadron calorimeters, covering the whole azimuth range and the pseudo rapidity up to $|\eta| = 3.64$.

CEM The scintillator based Central Electromagnetic Calorimeter (CEM) is segmented into 48 independent wedge modules. It is constructed of two rings of 24 wedges each that make contact at $z = 0$. Each wedge subtends 15° in azimuth and approximately one unit in η .

CES The Central Electromagnetic Shower (CES) detector is a set of cathode strips and anode wires. It measures the energy distribution at the electron maximum transverse shower position. The stripes and wires find the position of the shower peak within ~ 2 mm in z and x . The energy profile of the shower is measured using the 5 strips and wires neighboring the peak. This includes more than 99 % of the shower. The CES anode wires provides measurements in ϕ direction and cathode strips in η direction. Inside each chamber the wires run along the z direction are split in the middle at $z = 121.2$ cm. There are 32 wires at low $|z|$ ($0.2 < |z| < 121.2$ cm) and 32 at high $|z|$ ($121.2 < |z| < 239.6$ cm) for a total of 64 in one wedge. There are 69 strips at low $|z|$ and 59 at high $|z|$. The perpendicular distance to the beam line is 184 cm.

CPR The Central Pre-Radiator (CPR) chamber system consists of a set of chambers placed between the solenoid coil and the CEM. Its distance to the beam line is 168 cm. Inside each chamber the wires are running along the z axis. There are 16 wires at low $|z|$ ($7.9 < |z| < 119.7$ cm) and 32 at high $|z|$ ($123.5 < |z| < 235.3$ cm).

CHA/WHA The hadronic calorimeters are directly behind the electromagnetic calorimeter. Similar to the electromagnetic one they consist of alternating layers of scintillator and steel. For more detailed information about the calorimeters see references [37, 38, 39].

Chapter 3

Candidate Selection

This chapter focuses on the description of the selection process of the B candidates, which are used for the time dependent angular analysis. The Tevatron produces collisions with the interaction rate of 2.5 MHz.

The complicated structure of the CDF II detectors allows for different measurements in different sections of the detector. Because of the following reasons this analysis uses only the central region ($|\eta| < 1$) of the CDF II detector:

- The Time of Flight detector, which is very important for the separation between kaons and pions, is covering only the central region of the detector. In case of the B_s decay, this helps to improve the purity of the signal. For the B_d decay it is crucial, since the definition of the angles describing the decay depends on the identification of the kaon.
- The di-muon triggers select only CMU ($|\eta| < 0.6$) and CMX ($|\eta| < 1$) muon candidates. Since the angular acceptance is crucial for the later analysis, we need candidates which are well defined and understood in the simulation. Therefore, only those di-muon candidates from CMU and CMX that fired the trigger are considered.
- To pass all measurement layers in the COT a particle has to be in the $|\eta|$ region below 1. For a precise momentum resolution it is important to have as many hits in the COT as possible. The mass resolution of reconstructed particles is largely influenced by the momentum resolution. Having a good mass resolution improves the signal to background ratio.

Taking all these effects into account the central region is the optimal region for this analysis.

This chapter is structured in the same way the final candidate selection is obtained. Although the trigger is partially implemented in hardware and directly connected to the detector, it is the first step in the selection process and therefore described in this chapter. The second part describes the CDF software used for reconstructing the events. Some rough candidate pre-selection is described in section three, followed by a short interlude about the simulated Monte Carlo events. In the last section we describe the final candidate selection obtained by using neural networks.

3.1 Trigger System

At the Tevatron, collisions at the interaction point occur at a rate of 2.5 MHz. Most of these events are background events, which are not interesting for the physics program of CDF. A large fraction of these are minimum bias events, with proton anti-proton scattering processes. At the Tevatron, the production rate for these events is four orders of magnitudes higher than for bottom quarks. Accepting every event would lead to a data flow of ≈ 0.5 TB/s. This enormous amount is impossible to store. The task of the triggers is the selection of the physically interesting events and writing only these to tape for further processing.

The trigger system at CDF is organized in three different levels. The structure and data flow is illustrated in Figure 3.1. Figure 3.2 illustrates the trigger subsystems of Level-1 and Level-2. The individual levels are described in the following sections. For the analysis presented here, the di-muon trigger is used and the description will focus on this one.

3.1.1 Level-1

The first trigger level uses information from all detector components but the silicon vertex detector (see Figure 3.2). At this level, three independent systems reconstruct basic physical objects:

- L1Cal: calorimeter objects
- L1Muon: muon candidates
- L1Track: track reconstruction by the eXtremely Fast Tracker (XFT) [40]. Additionally tracks are matched to hits in the muon chamber and energy deposit in the calorimeters.

Information from these three components is used to make a decision whether an event is passed to the next level. Level-1 needs roughly $5.5 \mu\text{s}$ to make a

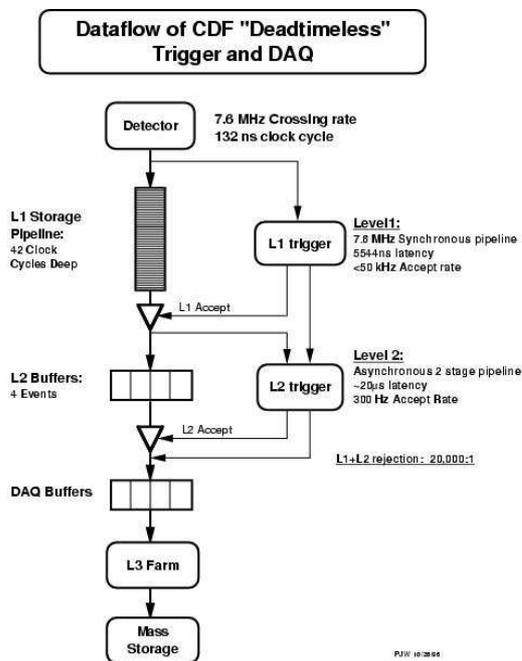


Figure 3.1: View of the data flow of the CDF II trigger system.

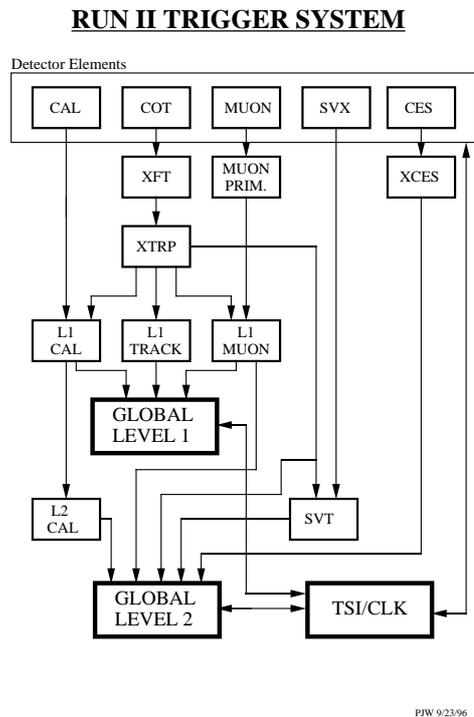


Figure 3.2: Diagram of the Level 1 and 2 trigger components of the CDF II trigger system.

decision, which is longer than the time between two bunches. Therefore, the information is stored in a 42 clock cycle deep pipeline. It is a synchronous system implemented in custom designed hardware which reads in a new event and makes a decision for an earlier event in the pipeline at the same time. Thus being able to handle the 2.5 MHz event rate. The acceptance rate of the first trigger level is below 50 kHz.

Different Level-1 triggers looking for different characteristics exist simultaneously. The di-muon trigger is one of them. During operation of CDF in the last years the di-muon trigger was constantly revised and optimized. Therefore the data used in this analyses consists of events with different characteristics. The different trigger paths can be divided in two main groups: CMU-CMU triggers, where both muons are found in the very central region, and CMU-CMX triggers, where one muon is detected in the range covered by the CMX detector. The di-muon trigger checks several things to determine whether an event fulfills the following criteria.

- The transverse momentum and ϕ information is taken from the XFT tracks matched to hits in the muon chambers. This is extrapolated to the inner radius of the muon system. Due to the uncertainty of the extrapolation coming from multiple scattering, a window covering the 3σ area in ϕ is determined and called the footprint.
- A tower in the muon system has to agree which at least one footprint. Then the tower is called a muon tower. A tower requires at least one of its stacks to have a Level-1 muon track segment. A stack are four cells of scintillator stacked on top of each other. It is required that at least cells 1 and 3 or 2 and 4 have hits which coincide during the stub gate width of 396 ns.
- For the di-muon trigger, two muon towers are needed, which are either at different sides (east and west) of the detector, or at least separated by two muon towers, where the gap between two wedges is treated as a tower, as well.

Further there are the following criteria, which are encoded in the trigger name:

PTx: p_t of the XFT track needs to be greater than x.

DPHIx: The difference in ϕ of the two muons must be smaller than x.

OPPQ: Both muons need to have opposite charge.

xMTy: The transverse mass for the two muons must be between x and y.

PSx, DPS: The trigger is either pre-scaled by the factor x or dynamically pre-scaled.

Some of the triggers are pre-scaled, which means that only one out of n events that fulfill the trigger requirements is sent to Level-2, where n is the pre-scale factor. This prevents the overall trigger system from exceeding the maximal possible event rate at different luminosities. The dynamic pre-scale is an improvement of the fixed pre-scale, that allows to change the factor dynamically to allow the optimal performance at all times and luminosities.

3.1.2 Level-2

In contrast to Level-1, Level-2 is an asynchronous system processing events accepted before. It extends the information from Level-1 by adding data from the central calorimeter (CES) and the $r - \phi$ strips of the silicon detector.

The CES information is processed by the L2CAL subsystem looking for energy tower clusters. The XSEC system matches XFT track primitives extrapolated to the CES radius with energy deposits above certain thresholds in the calorimeter, in order to identify electrons. The third Level-2 component is the Silicon Vertex Trigger (SVT) [41], which identifies displaced secondary vertices, which is very important for studying long lived b-hadrons. The SVT extends XFT tracks inside the silicon detector by adding $r - \phi$ hits. This improves the ϕ_0 and p_t resolution and adds the track impact parameter d_0 .

Level-2 needs roughly $20 \mu\text{s}$ to process a given event. A buffer of 4 events reduces the risk of losing interesting events coming from Level-1. The output rate of Level-2 is about several hundred Hz. The full list of trigger paths used in this analysis is listed in appendix A. For each trigger the details can be derived from the name of the trigger. If an event is fulfilling all criteria, it is accepted and passed on to Level-2.

3.1.3 Level-3

Level-3 is implemented in software running on a farm of Linux PCs. In case of a Level-2 accept, the entire detector is read out from the buffers of each detector component. The so called Event Builder takes care of the proper order of all event fragments. The converter node creates the event record, which is the complete event information stored in a format suitable to be analysed by the CDF software.

The event record is submitted to one of the processors in Level-3. Having the complete information of the detector and more time for processing, the full COT and SVX offline track reconstruction is performed. The Level-1

and Level-2 trigger decisions are confirmed with improved resolution for p_t , ϕ , z_0 , $\cos \Theta$ and d_0 .

The triggers used in this analysis require during Level-3 that the invariant mass of the reconstructed di-muon candidate is nearby the J/ψ mass. An accepted event is stored on tape for later analysis. The output rate of Level-3 is of the order of 100 Hz.

3.2 Offline Reconstruction

The CDF II experiment uses a common software framework developed in C++. As mentioned above, the data written out by the Level-3 trigger is stored in a format to be analyzed by the CDF II offline software.

The framework is organized in packages, which allows a modular approach, to fit different analysis strategies. The packages are structured in a way, that independent physics groups can develop their independent software without interference. The modules of the framework can be ordered in a way, that the output from the first module is available for the later ones. Thus a hierarchy starting from low level modules, which for example reconstruct tracks, to higher level physical objects like B candidates is easily implemented.

The next sections will describe some important part of the reconstruction of the B mesons.

3.2.1 Good Run Selection

The experiments CDF II and $D\bar{O}$ are run by independent collaborations. This implies that the two experiments and the beam division have to optimize the global performance of the experimental setup. This leads to situations where the collisions are carried on, although on of both experiments does not perform optimal. In such a situation the CDF II detector does not stop taking data, but will continue to record events as long as there are collisions. Different physics groups do rely primarily on different parts of the detector. Therefore even if some parts of the detector do not work, some groups might consider the data useful.

A simple example are the muon chambers. In case they do not work properly, the analysis presented here could not use the data taken during this period. On the other hand do many analysis not need them and would not be affected. This implies that for any run period used in this analysis it should be checked that the data was taken under stable condition with all necessary components fully functional.

To make this feasible, the individual parts of the experimental setup are monitored constantly during data taking and marked appropriately. For different systems, there exist several good run flags, which can be checked for each run in the offline analysis.

The B group at CDF agreed on a set of requirements necessary for any analysis. They can be summarized as:

- A single run should have more than 10 nb^{-1} with a trigger table approved for physics.
- All three trigger levels worked properly and there were no problems with the data processing.
- The Cerenkov luminosity counters were working properly.
- The silicon detector, muon chambers and drift chamber were working properly.

The data fulfilling these criteria correspond to roughly 1.7 fb^{-1} . The amount of data for the B_d decay is lower, due to the following criterion. The calibration of the particle identification via energy loss in the drift chamber is very useful for both decays to suppress background. The data above 1.3 fb^{-1} is not perfectly calibrated, which implies a small inefficiency for the B_s decay in the event selection. For the B_d decay the calibration is crucial for the determination of the acceptance due to an additional background. This peculiarity for B_d will be explained in more detail later in this chapter and does not allow to take data above 1.3 fb^{-1} at the moment.

3.2.2 Event Reconstruction

For this analysis the BottomMods package [42, 43] was used to reconstruct both, $B_s \rightarrow J/\psi \phi$ and $B_d \rightarrow J/\psi K^*$, decays. The lowest order Feynman diagrams of the B meson decays are shown in Figure 3.3. The three intermediate resonant states are reconstructed in the following decay modes

- $J/\psi \rightarrow \mu^+ \mu^-$
- $\phi \rightarrow K^+ K^-$
- $K^* \rightarrow K^\pm \pi^\mp$.

The J/ψ reconstruction is the same for both B mesons, but the decay products of the second intermediate meson are different. Before reconstructing the B meson from their decay products, all tracks are refitted using

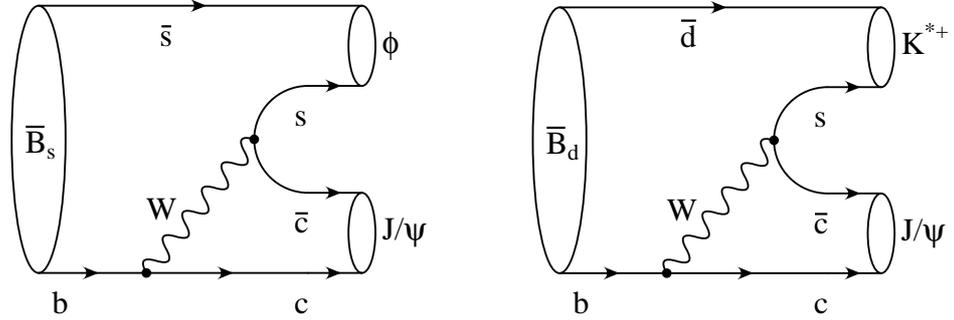


Figure 3.3: Feynman graph describing the $B_s \rightarrow J/\psi \phi$ and $B_d \rightarrow J/\psi K^*$ decays.

the KalmanFitter and the appropriate rest mass hypothesis for the particle hypothesis of the final state. For muons the candidate selection is rather pure due to the low backgrounds in the muon chambers, whereas for pions and kaons any track is reconstructed using both particle hypotheses. This causes a lot of combinatorial background, which should be reduced as much as possible.

The B meson candidates are reconstructed backwards from the stable particles by combining particles originating from the same resonances. As described in Chapter 2, a muon candidate needs to have a track segment in the muon chambers associated with a track in the drift chamber, which is required to have a transverse momentum of at least $1.5 \text{ GeV}/c$. Two muon candidates with opposite charge are combined in a vertex fit to form a J/ψ candidate. The world average mass is used for the muon candidates in the vertex fit. We use di-kaon pairs for the ϕ reconstruction and use the kaon world average mass in the vertex fit. Both kaons are required to have different charge. In the similar way the K^* is obtained by doing a vertex fit using a kaon and pion candidate with opposite charge. For the reconstruction of the B meson the J/ψ is then combined using the partner meson ϕ or K^* to get either a B_s or B_d candidate. In the vertex fit of the B meson, the mass of the J/ψ candidate is set to the world average value for the J/ψ , whereas for the ϕ or K^* mass the reconstructed value from the previous fit is used.

3.2.3 Loose Candidate Selection

The reconstruction above provides a very loose candidate selection, which is dominated by background events. To keep the combinatorial background low, without losing too much signal events, the mass windows, opposite charge

requirements and minimal transverse momenta of the decay products are used.

The decay chain is reproduced in software, using the modular ansatz of the CDF II software framework. The surviving B meson candidates are stored using the B Stntuple framework [44]. Each reconstructed stable and unstable particle for each event is stored, including the vertex fit information. The individual information for each particle and the vertex fit results can then be accessed in the next analysis steps.

3.3 Pre-selection

An important step for most analyses in particle physics is an event selection of a reasonable size, which allows fast development cycles in the final analysis. Although the Stntuple described in the previous section allows to make a full analysis, we use an additional step to reduce the data sample size. Whereas the Stntuple framework allows for a complicated file structure, e.g. a single pion instance can be part of different B meson candidates in an event, we convert the files into flat ntuples, which means that for each candidate all decay particles and vertex information are stored independently in flat ROOT ntuples.

This simple structure allows for fast changes in the next analysis steps. Another set of preselecting cuts are implemented on this level. These cuts are mainly on kinematic and vertex-fit variables and sometimes just confirm requirements of the level before.

Basic quality cuts on the different particle candidates are summarized in table 3.1. Independent of the particle hypothesis, each track needs to have a minimal transverse momentum of 0.4 GeV/c.

The cuts presented in this section are only very loose requirements to reduce the combinatorial background. The final candidate selection is done using a neural network, which transforms all available information into a single discriminating variable. The neural network will be discussed after a short interlude about simulated decays.

3.4 Monte Carlo Simulation

Monte Carlo samples used in this analysis are generated using standard CDF full simulation in the release 6.1.4mc. Event generation starts with BGenerator to produce a single B_s (B_d) meson, which is then decayed using the EvtGen package [45]. All particles are decayed according to phase space. This

Cuts	B_d	B_s	
Run number	<226196		
K or π :			
p_t	> 0.4	> 0.4	GeV/c
Muon:			
Di-muon pair triggered	yes	yes	
p_t	> 1.5	> 1.5	GeV/c
J/ψ ($M_{J/\psi} = 3.0969$ GeV) [17]:			
$ M_{\mu\mu} - M_{J/\psi}^{PDG} $	≤ 80	≤ 80	MeV/c ²
K^* ($M_{K^*} = 0.89166$ GeV) [17]:			
p_t	> 2.0		GeV/c
$ M_{K\pi} - M_{K^*}^{PDG} $	≤ 80		MeV/c ²
ϕ ($M_\phi = 1.01946$ GeV) [17]:			
p_t		> 1.0	GeV/c
$ M_{KK} - M_\phi^{PDG} $		≤ 100	MeV/c ²
B Meson:			
$\sigma_{c\tau}$	< 150	< 150	μm
p_t	> 4.0	> 4.0	GeV/c

Table 3.1: Summary of loose cuts during B_d and B_s Meson reconstruction

allows us to re-weight the sample according to any decay model depending on our needs. After the decay, the full simulation of the CDF detector [46] and di-muon J/ψ trigger simulation follows. Simulated events are then reconstructed by standard CDF offline production and stored at GridKa Karlsruhe into SAM [47].

Both samples are further processed by the same pre-cuts and analysis steps described in the sections before. This guarantees a good agreement between the simulated events and the signal events on data. Simulated events are of major importance in this analysis for two reasons. First they are used in the training of the neural networks described in the next section, second they are essential for deriving detector and reconstruction efficiency, which will be described in chapter 4.

3.5 Neural Network Selection

For the final selection of the B candidates we use a neural network to separate signal from background. Using this tool, correlations between variables are taken into account and the information of the different variables is combined in an optimal way into a single discriminating variable.

We choose the NeuroBayes [48] package, since it is one of the most advanced implementations of a neural network. It combines a three-layer feed forward network with a very robust and sophisticated pre-processing. Figure 3.4 is a schematic representation of such a three-layer network topology. Each node in the input layer corresponds to one input variable. The hidden layer connects all input nodes and output nodes. Their number should be of the same order as the input nodes. In our case only a binary decision for the classification of signal and background events is needed. This requires only a single output node ($[-1,+1]$).

Nodes in different layers of the network are connected with weights. For a single node, these weights w_{ij} of the input nodes x_i have to be combined and the output of the node must be calculated. Figure 3.5 shows a graphical representation for a single node. The input is combined by a biased weighted sum

$$a_j(\vec{x}) = \sum_i w_{ij}x_i + \mu_j,$$

where μ_j is needed for the threshold, introduced by the additional bias node in the input layer. The output of a single node is mapped from the possible input interval $[-\text{inf}, +\text{inf}]$ to $[-1, +1]$ by the sigmoid function

$$S(\vec{x}) = \frac{2}{1 + \exp(a(\vec{x}))} - 1.$$

The network topology is defined by the number of input variables, output nodes and hidden nodes, but the weights between the layers are unknown and have to be determined. This is done during the training of the network. Using a training sample of known classes, the weights of the network can be optimized in order to get as close as possible to the truth of the training sample. The optimization is done by minimizing a cost function, which describes the difference of the current network output and the true result. NeuroBayes uses an entropy function as cost function

$$E = \sum_{i=1}^N -\ln \left(\frac{1}{2}(1 + o_i t_i) \right),$$

where o_i is the network output, t_i the target value and N the number of training samples. In the multidimensional space, spanned by the weights, a back-propagation algorithm is used to find a minimum.

Hitherto, the description of NeuroBayes is similar to any feed-forward network, but might be different in some detail. One unique feature of NeuroBayes is its sophisticated preprocessing for each input variable. The preprocessing helps the minimization step of the cost function described before. It transforms and de-correlates the input variables and allows the use of discrete input variables and variables that are not always defined [49].

Further regularization schemes are implemented to avoid over-training and to improve the networks generalization capabilities. This is done by removing weights which get insignificant and by adding the sum of squared weights to the cost function. The later causing the optimization to prefer smaller weights, stabilising the training.

The next sections will explain individually for each neural network used in this analyses

- the datasets used for the training,
- the variables chosen as input and
- the result of each training.

3.5.1 B_s Neural Network

The aim of this neural network is to distinguish B_s meson decays from background events. We assume that the background events in the B_s mass region have a similar properties as near the B_s mass. Since we can take events from both sides of the mass peak, we can assume that these events will on average

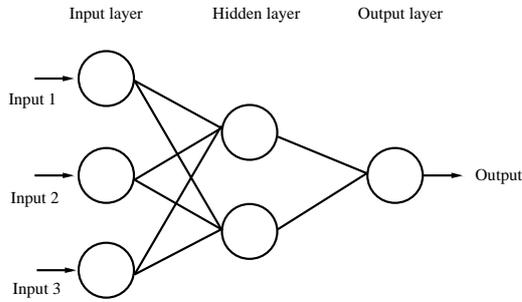


Figure 3.4: Schematic view of a three layer feed forward network.

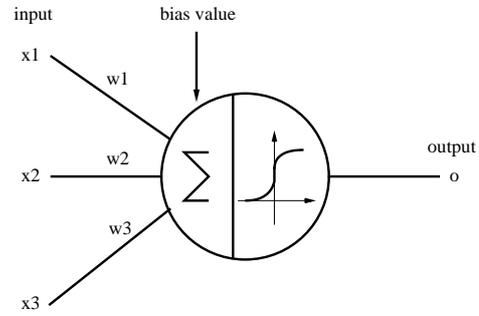


Figure 3.5: Illustration of a single node and its mathematical transfer functions.

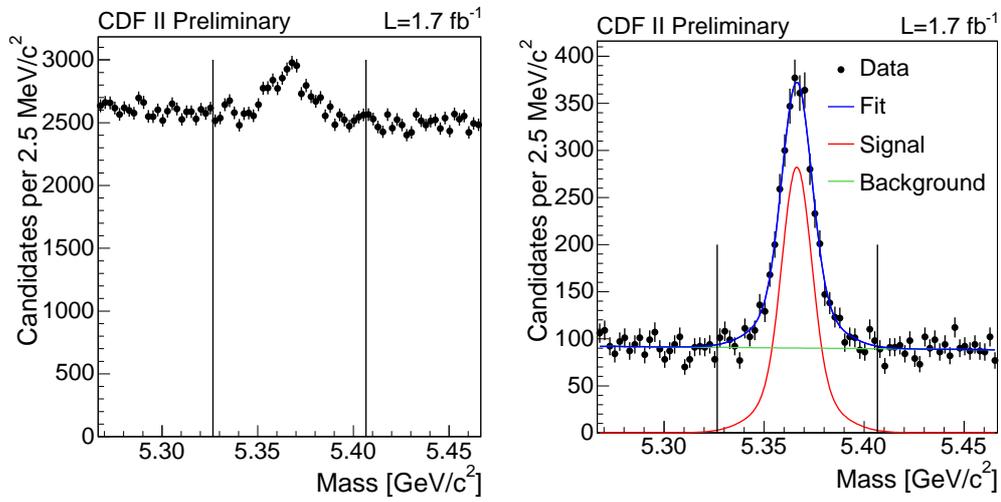


Figure 3.6: Mass distribution of the B_s candidates after the pre-selection (left) and the final neural network selection (right). The lines denote the sidebands.

represent the topology of the background events inside the B_s meson mass region. Thus a representative collection of background events for training can be taken from the lower and upper sidebands in the mass distribution (5.2666-5.3266 GeV/ c^2 and 5.4066-5.4666 GeV/ c^2). The mass distribution after the pre-selection, and before the network selection, is shown in Figure 3.6.

Since the signal mass region consists of a mixture of signal and background events, we use simulated B_s meson decays as signal for the training of the neural network. Fifty thousand signal events and fifty thousand background events from the sidebands were used for the training of the network.

As input for the neural network many different discriminating variables are available. Since we want to make a time-dependent fit, we cannot cut on or use the decay time of the B meson without sculpting the decay time distribution, we want to measure. Also the mass of the meson candidate is not available, since we use different mass regions for the signal and background events for training. Besides these two restrictions a large variety is available. Typical examples are transverse momenta, masses of decay products, quality of vertex fits and variables identifying particle types (PID). For a detailed overview and the correlations between the variables, see table 3.2. For the PID we use the likelihood ratio which combines the energy deposit measurement in the drift chamber and the time of flight measurement, originally developed for the same side kaon tagger [50]. For the muon identification we employ the likelihood discriminant used by the soft muon tagger [51].

The results of the training can be seen in Figure 3.7. The network separates signal (+1) from background (-1) very well, and the purity is linearly dependent on the network output. Any deviation from the linear dependence would mean a non optimal behaviour of the neural network. This network is then used to make the final candidate selection, by imposing a cut on the network output.

To get an unbiased criterion of the optimal cut on the network output, we estimated for different cuts on the network output the significance, $N_S/\sqrt{N_S + N_B}$, on data in the mass region ± 20 MeV/ c^2 around the signal peak. N_B is the number of background events calculated from the background parameters determined in a mass fit. N_S is the total number of events in the ± 20 MeV/ c^2 mass region minus N_B .

Figure 3.8 shows the distribution of the significance versus the cut on the network output. Since it is relatively flat around its maximum, the exact value of the cut has no strong influence on the significance. We choose a cut of 0.6 for B_s for the final selection. For the ± 20 MeV/ c^2 mass window around the B_s mass we obtain a significance of $S/\sqrt{(S + B)} = 37.7$ for the final selection. The mass distribution of the final selection is shown

#	Variable B_s	B_d
1	training target	←
3	$p_t(B)$	←
10	$p_t(J/\psi)$	←
5	$p_t(\phi)$	$p_t(K^*)$
8	$p_t(K_1)$	$p_t(K)$
9	$p_t(K_2)$	$p_t(\pi)$
15	$\max(p_t(\mu^+), p_t(\mu^-))$	←
16	$\min(p_t(\mu^+), p_t(\mu^-))$	←
12	$ M_{\mu\mu} - M_{J/\psi}^{PDG} $	←
7	$ M_{KK} $	$ M_{K\pi} $
4	$Prob(\chi^2)(B_s)$	←
11	$Prob(\chi^2)(J/\psi)$	←
6	$Prob(\chi^2)(\phi)$	(K^*)
2	$\chi_{r\phi}^2(B)$	←
17	$PID_K(K^+)$	$PID_K(K)$
18	$PID_\pi(K^+)$	$PID_\pi(K)$
19	$PID_K(K^-)$	$PID_K(\pi)$
20	$PID_\pi(K^-)$	$PID_\pi(\pi)$
13	$\max(\text{lh}(\mu^+), \text{lh}(\mu^-))$	←
14	$\min(\text{lh}(\mu^+), \text{lh}(\mu^-))$	←

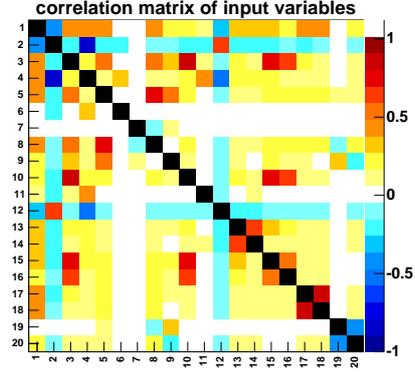
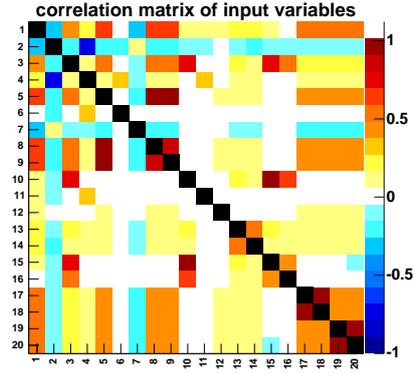


Table 3.2: Overview of the variables (left) and their correlations (right) used for training of the B_s (top) and B_d (bottom) neural network. The arrow indicates common variables of both decays. For a detailed description of the individual variables see Appendix B.

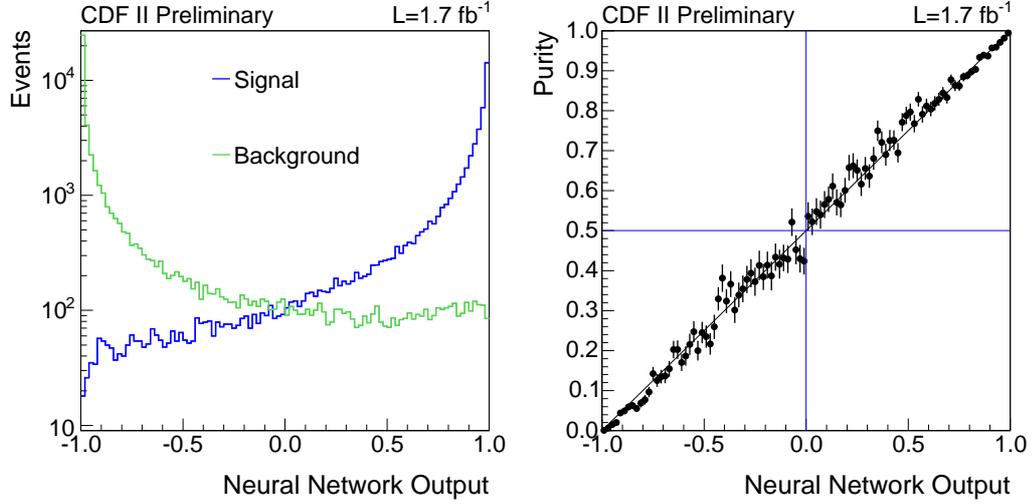


Figure 3.7: Network output distribution (left) for the simulated B_s mesons (blue) and background events from the sidebands (green). The purity on the training sample as a function of network output (right).

in Figure 3.6. Compared to the selection before the neural network, the number of background events is reduced by a factor of ≈ 25 with roughly half of the signal being kept. Since the neural network was trained using simulated signal decays and background events from data, any discrepancy in the simulation might influence the training result. The neural network might have learned to distinguish real data from simulated events and not B_s decay from background. To check that this is not the case we can compare the neural network output distribution for the simulated signal and the signal from data, using sideband subtracted data in the signal region. Figure 3.9 shows the good agreement for the two distributions.

3.5.2 B_d Neural Network

The neural network selection procedure for the B_d candidates is very similar to the B_s , described in the section before. The neural network was trained using a combination of reconstructed data events from the sidebands and simulated decays for signal. The lower and upper sidebands of the mass windows for the background events are different, due to the lower mass of the B_d : 5.13-5.2094 GeV/c^2 and 5.3294-5.43 GeV/c^2 .

The variables used for the training are very similar and are also listed in Table 3.2. For the result of the training see Figure 3.10. As before, the neural network is able to separate signal from background very well and the neural

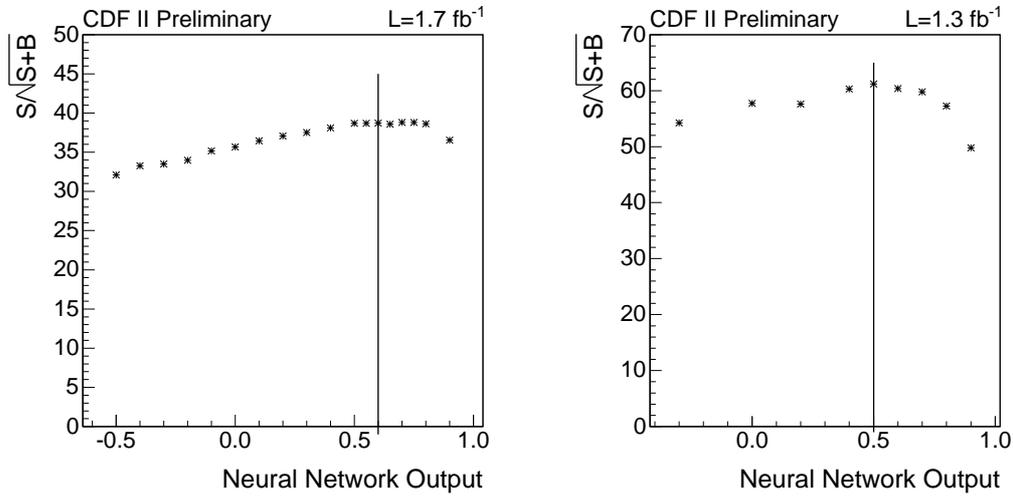


Figure 3.8: Signal significance as a function of the cut on the network output for B_s . (left) and B_d (right). The line denotes the cut used for the final selection.

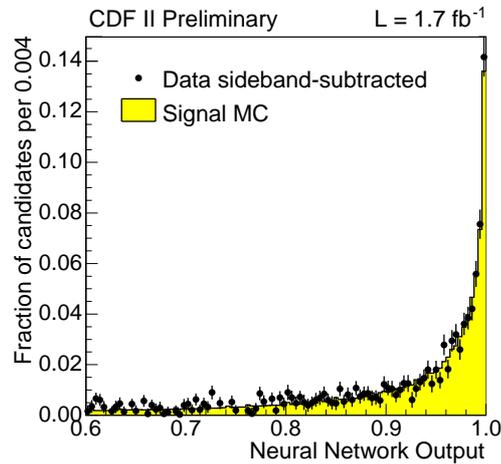


Figure 3.9: The B_s neural network output distribution for simulated signal events and sideband subtracted data in the signal region.

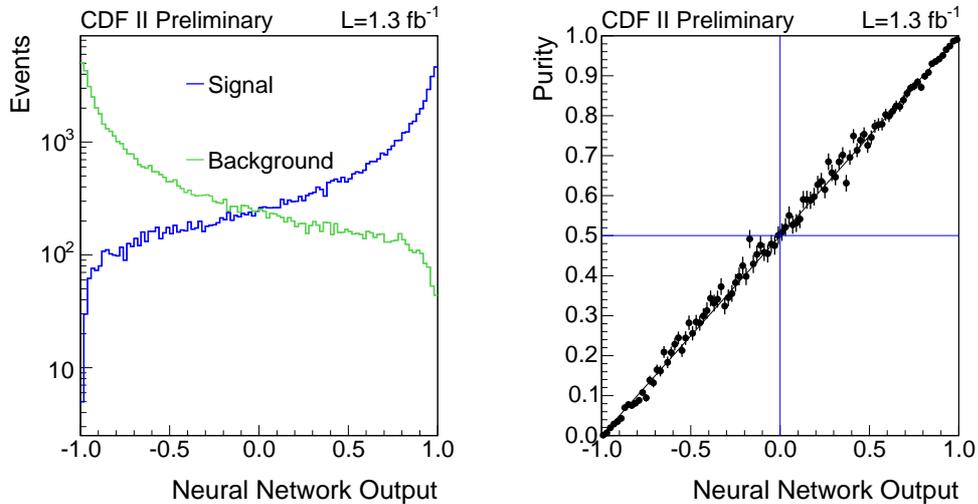


Figure 3.10: Network output distribution (left) for the simulated B_d mesons (blue) and background events from the sidebands (green). The purity on the training sample as a function of network output (right).

network output is near the optimum, since the purity is a linear function of the network output.

3.5.3 Swap Suppression Neural Network

In the B_d decay the association of the K^* decay products is ambiguous. Due to this there are B candidates reconstructed with swapped pion-kaon hypothesis. From previous studies at CDF, we expect that roughly 10% [9] of the signal candidates are reconstructed with swapped mass assignment. B_d events reconstructed with the wrong mass hypothesis have several severe effects on the distributions. Although the underlying true angular distribution is correct, the angles describing the decay of these events are calculated with the wrong particle as K reference. Further the transverse momenta and reconstructed masses are calculated with the wrong mass hypothesis. Another difference is the shape of the acceptance of these candidate, which will be introduced in the next chapter.

Since there is a large fraction of these events, they have to be taken into account. There are two possible ways. Either they have to be parameterized and included in the fit function, where the swapped and un-swapped angles can be calculated, or, to treat them as systematic uncertainty, they have to be removed from the selection as good as possible.

Although we can include the swapped component in the fit model, there

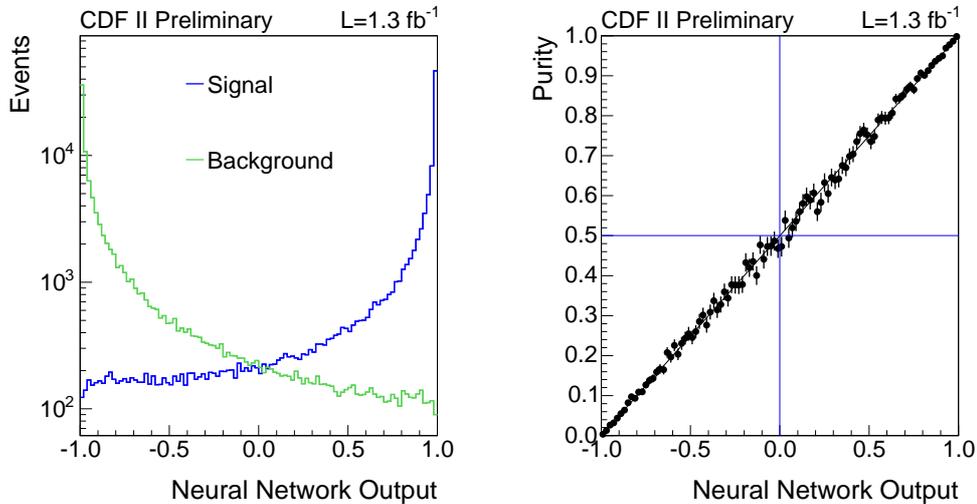


Figure 3.11: Network output distribution (left) for the simulated B_d mesons (blue) and background events from the sidebands (green) with swapped kaon pion assignment and the purity on the training sample as a function of network output (right).

are several uncertainties which will result in large systematic uncertainties. Besides the very different acceptance distribution, the mass distribution of the swapped candidates is broader but at the same place as the signal. This means we have to rely on the mass model from simulated events since it cannot be determined reliably in a fit on data. The same is true for the fraction of swapped candidates. Therefore, we decided to follow the second procedure by training a dedicated neural network to distinguish swapped B_d candidates and remove them from our data sample.

To train the network we used the same simulated events for signal and background as before, once reconstructed with the right mass assignment and a second time using the wrong assignment. The input variables and their correlations are listed in table 3.3. Besides the particle identification variables, the reconstructed K^* mass is a very important input.

The network output distribution of the training sample can be seen in Figure 3.11. The network can identify swapped candidates very well and is able to reduce the amount of swapped candidates to roughly half a percent, estimated on the simulated events, at a purity of fifty percent using a neural network cut of 0.8. We are then able to fit the data without any swapped component and treat its influence as systematic uncertainty.

Since we have two neural networks for the B_d selection the selection procedure is slightly different compared to B_s . First we use the precut of 0.8

#	Variable B_d Swap
1	training target
2	$p_t(K)$
3	$\sigma(p_t(K))$
4	$p_t(\pi)$
5	$\sigma(p_t(\pi))$
6	$ M_{K\pi} $
7	$PID_K(K)$
8	$PID_\pi(K)$
9	$PID_\mu(K)$
10	$PID_e(K)$
11	$PID_p(K)$
12	$PID_K(\pi)$
13	$PID_\pi(\pi)$
14	$PID_\mu(\pi)$
15	$PID_e(\pi)$
16	$PID_p(\pi)$

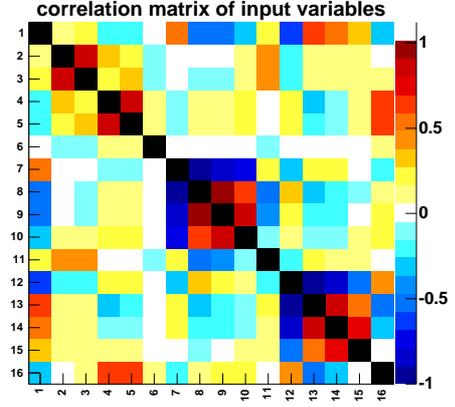


Table 3.3: Overview of the variables (left) and their correlations (right) used for training of the swap suppression neural network. PID is the likelihood ratio for different particle hypotheses. Similar variables are described in Appendix B.

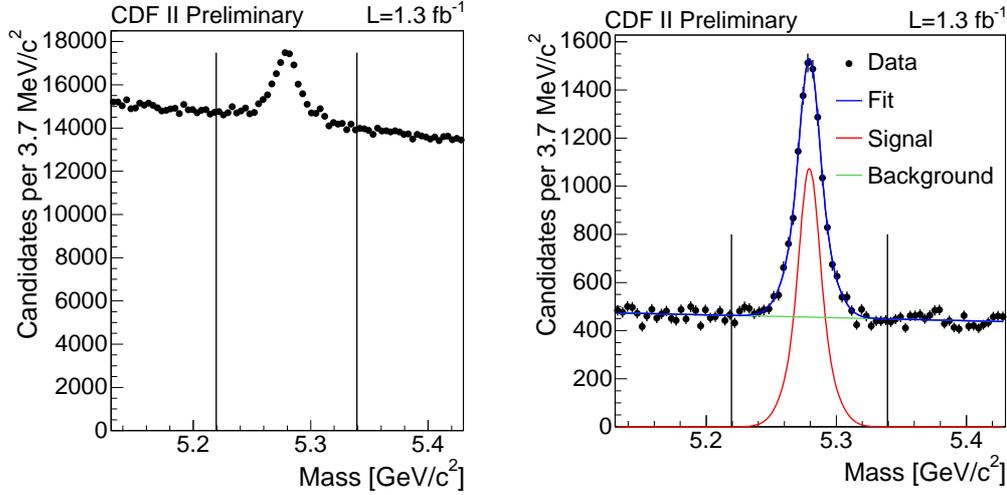


Figure 3.12: Mass distribution of the B_d candidates after the pre-selection (left) and the final neural network selection (right). The lines denote the sidebands.

on the swap suppression network to suppress the swapped fraction to very low values and optimized then in a similar way as for the B_s by maximizing the significance. We choose to use > 0.5 for the cut on the neural network output for the B_d candidates. The selection after both neural network cuts is shown in Figure 3.12. For the significance in the $20 \text{ MeV}/c$ mass region, $N_S/\sqrt{(N_S + N_B)(\pm 20 \text{ MeV}/c^2)}$, we obtain 61.8. The background is significantly reduced with a large fraction of signal events remaining. Similar as for B_s we check that the neural network learned to distinguish B_d decays from background by comparing the neural network output distribution for the simulated signal and the signal from data, using sideband subtracted data in the signal region. Figure 3.13 shows the good agreement for the two distributions.

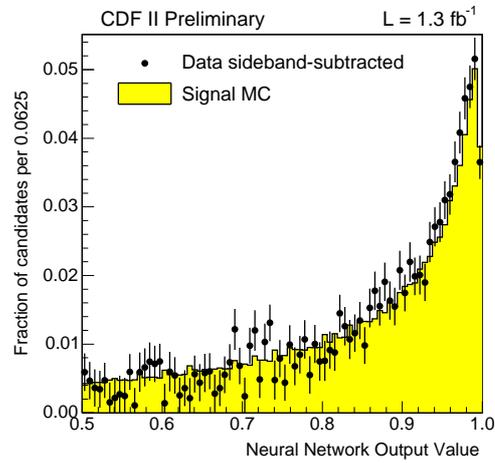


Figure 3.13: The B_d neural network output distribution for simulated signal events and sideband subtracted data in the signal region.

Chapter 4

Fit Function

In Chapter 1 we derived, starting from the Standard Model, formulas describing the B meson decay. In this chapter, we will continue this course and discuss all parts necessary to obtain the probability density function describing the data. The signal functions, resolution and acceptance are given by theory and experiment, whereas we are going to take an empirical approach for the background description. We will slightly change our parameterization by using $\phi_s = 2\beta_s$ instead of β_s from now on.

The decay is described only by a few variables, which have to be derived from the data. The CDF II detector described in Chapter 2 consists of many subsystems, providing many measurements of very different quantities. In the previous Chapter 3 we discussed the reconstruction of physical objects from the detector measurements, which allowed us to make a very good selection of the interesting events. The likelihood function only depends on some of these reconstructed physical observables:

m the mass of the B meson,

ct the decay time of the B meson,

σ_{ct} the estimated decay time uncertainty of the B meson,

$\vec{\omega} = (\psi, \phi, \theta)$ three angles describing the decay, defined in section 1.3.4,

ξ describing the tag of the events.

In the next section, we will explain shortly the maximum likelihood method used to estimate the parameters. The second section will discuss each individual variable and the way its shape is parameterized, which will also contain detector effects. We will then combine these in the following two sections independently for B_s and B_d to obtain the complete probability

density function. The last section will then discuss some details of the fitting procedure and its implementation.

4.1 Parameter Estimation

For the final selection, we can generalize the data from the experiment as n independent measurements, $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$, of the multidimensional variable $\vec{x} = (m, ct, \sigma_{ct}, \vec{\omega}, \xi)$. The purpose of the parameter estimation is to deduce one or more parameters of the model from these measurements. Due to limited statistics and the uncertainty of the input variables, the estimated parameters will be determined only with some precision. Having a probability density function $f(\vec{x}|\vec{a})$ describing the data, where \vec{a} are all the parameters, we can use the maximum likelihood method [52] to extract the best estimate of the parameters \vec{a} and their uncertainties.

From the probability density function and the n measurements we can construct the likelihood function

$$L(\vec{a}) = f(\vec{x}_1|\vec{a}) \cdot f(\vec{x}_2|\vec{a}) \cdot \dots \cdot f(\vec{x}_n|\vec{a}) = \prod_{i=1}^n f(\vec{x}_i|\vec{a}).$$

The maximum likelihood method returns the best estimate of parameters \hat{a} that maximizes $L(\vec{a})$, having the given measurements \vec{x}_i . $L(\hat{a})$ is equivalent to the statement that for those parameter values \hat{a} , the probability to observe the given measurements \vec{x}_i is maximized. Due to practical reasons, usually the negative log likelihood function

$$F(\vec{a}) = - \sum_{i=1}^n \ln f(\vec{x}_i|\vec{a})$$

is minimized instead. One necessity of the maximum likelihood method is that the probability density function must be normalized for any set of parameter values

$$\int f(\vec{x}_i|\vec{a}) d\vec{x} = 1.$$

The normalization of the likelihood function is not always trivial. It will be discussed in more detail later.

For any parameter estimation method, there are several criteria which should be fulfilled. If \vec{a}_0 are the true values then the method is

consistent: For infinite statistics the estimated parameters will be the true parameters, $\lim_{n \rightarrow \infty} \hat{a} = \vec{a}_0$.

unbiased: The expected parameters are the true parameters, $E[\hat{a}] = \vec{a}_0$, where $E[\hat{a}]$ is the expectation of \hat{a} .

efficient: The fit returns the smallest possible uncertainties of \hat{a} .

robust: The fit is not affected by wrong data and wrong assumptions.

In general the maximum likelihood method is consistent, but it is only unbiased and efficient in case of infinite statistics [52].

4.2 Fit Model Components

The probability density function can be factorized into mass, decay time and angular function, if they are uncorrelated. In addition we have two kinds of classes of events, signal and background, which have different shapes in some of the variables. Since the signal and background distributions are two independent components, the probability density function can be written as

$$P = f_s P_S(m, c\tau, \sigma_{c\tau}, \vec{\omega}, \xi) + (1 - f_s) P_B(m, c\tau, \sigma_{c\tau}, \vec{\omega}), \quad (4.1)$$

where f_s is the fraction of signal events in the sample of n measurements. In case the different variables are not correlated, the probability density function can be factorized. In our case the probability density function gets the form

$$P = f_s X_S(m) \cdot Y_S(c\tau, \sigma_{c\tau}, \vec{\omega}, \xi) \cdot Y'_S(\sigma_{ct}) \\ + (1 - f_s) X_B(m) \cdot Y_B(c\tau, \sigma_{c\tau}) \cdot Y'_B(\sigma_{ct}) \cdot Z_B(\vec{\omega}). \quad (4.2)$$

We have written the formula already in a the very specific way useful for the analysis. We omitted the parameters of each part here for compactness, but will introduce them when discussing the individual parts of the probability density function in the next sections.

In general we choose an empirical approach for the background description. By selecting pure background events from the sideband, we can choose a set of functions that describes the background. The same description is then used for the background in the signal region, as well.

4.2.1 Mass

Any hadronic state has a natural width, depending on the decay. We analyze weak decays, which means that the natural width is very small. The mass distribution is convoluted with the resolution function. The shape of the mass

distribution is dominated by the detector resolution, which is approximately Gaussian. Figures 3.6 and 3.12 show the mass distribution for the two decays of interest. A single Gaussian seems not to be sufficient to describe the mass shape properly, therefore we use two Gaussian for the signal part. Both Gaussians use the same mean M but have independent widths σ_m and $s_m\sigma_m$. The corresponding normalized signal probability density function is

$$X_S(m|M, \sigma, f_m, s_m) = (1 - f_m) \frac{1}{\sqrt{2\pi}\sigma_m} e^{-\frac{(m-M)^2}{2\sigma_m^2}} + f_m \frac{1}{\sqrt{2\pi}(s_m\sigma_m)} e^{-\frac{(m-M)^2}{2(s_m\sigma_m)^2}}, \quad (4.3)$$

where f_m is the fraction between the two Gaussians.

The background is described using a set of linear functions. Two independent linear functions are used for the prompt and non-prompt background in the mass space. The combined normalized background probability density function for the mass is then

$$\begin{aligned} X_B(m|a, a_{NP}, f_{NP}) &= (1 - f_{NP})X_{bg}^P(m|a) + f_{NP}X_{bg}^{NP}(m|a_{NP}) \\ &= (1 - f_{NP})am + \frac{1 - \frac{a}{2}(M_{max}^2 - M_{min}^2)}{M_{max} - M_{min}} + \\ &\quad f_{NP}a_{NP}m + \frac{1 - \frac{a_{NP}}{2}(M_{max}^2 - M_{min}^2)}{M_{max} - M_{min}}, \end{aligned}$$

where a and a_{NP} are the slopes of the lines, f_{NP} the fraction of the non prompt background and M_{min} and M_{max} are the upper and lower boundaries of the mass window (5.2666-5.4666 GeV/ c^2 for B_s and 5.13-5.43 GeV/ c^2 for B_d).

Although the aim of this analysis is not a mass measurement, the fit of the mass distribution is very important. It provides a very good prior probability, whether an event is more likely signal or background. The signal fraction parameter is dominated by the influence of the mass distribution. The CDF II detector has a very precise momentum measurement due to the large drift chamber. Since the detector resolution governs the width, it allows to have narrow mass resonances of ≈ 10 MeV/ c^2 for B decays leading to less background.

4.2.2 Decay Time

Throughout this thesis the decay time is measured as a distance by multiplying the measured time until the decay with c . The decay of a particle is governed by an exponential law, but to describe the observed distribution, an exponential function is not enough. We also have to take into account

the detector resolution, which is of the order of several ten micrometers. Thus, the distribution we observe is an exponential function convoluted with a Gaussian. The width of the Gaussian, corresponding to the estimated uncertainty on the decay time σ_{ct} , is calculated on an event-by-event basis for each candidate multiplied with an global scale factor S_{ct} . The probability density function for the decay time is then given by

$$Y(ct, \sigma_{ct}|c\tau, S_{ct}) = E(ct|c\tau) \otimes G(ct, \sigma_{ct}|S_{ct}), \quad (4.4)$$

where

$$E(ct|c\tau) = \Theta(ct) \frac{1}{c\tau} e^{-\frac{ct}{c\tau}} \quad (4.5)$$

$$G(ct, \sigma_{ct}|S_{ct}) = \frac{1}{\sqrt{2\pi} S_{ct} \sigma_{ct}} e^{-\frac{(ct)^2}{2(S_{ct} \sigma_{ct})^2}}. \quad (4.6)$$

All decay time distributions in this thesis are convoluted with the Gaussian resolution in the same way. For the two mass eigenstates of the signal we have two different lifetime parameters: $c\tau_H$ and $c\tau_L$. The implemented decay time probability density function for the signal will be described later in more detail.

Similar to the approach of the mass distribution of the background, its decay time distribution is described using an empirical set of functions. For an example of the decay time distribution of the sidebands see Figures 5.5 or 6.7. The dominant part of the background distribution is a Gaussian peak around zero, describing prompt decays. The event-by-event decay time uncertainty allows us to compose the central Gaussian out of the summation of the individual Gaussian distributions. We observe that the composed function is too narrow to describe the data, since the estimated decay time uncertainty is too optimistic. This is compensated by introducing the scale factor S_{ct} . Assuming that this scale factor is on average the same for all events, we are able to determine it by the prompt background distribution. Further the distribution has a negative and a larger positive tail, which are not described well with the single Gaussian. Both tails are subsumed under non-prompt background. An often used parameterization for the complete background distribution is a Gaussian around zero to describe the prompt component and three exponentials for the rest, one to describe the negative tail ($f_-, c\tau_-$), one for the shortlived positive component ($f_+, c\tau_+$) and another one for the longlived positive component ($f_{++}, c\tau_{++}$). The decay time probability density function for background becomes

$$\begin{aligned}
& Y_B(ct, \sigma_{ct} | f_-, f_+, f_{++}, c\tau_-, c\tau_+, c\tau_{++}, S_{ct}) \\
&= (1 - f_{NP})Y_B^P(ct, \sigma_{ct} | S_{ct}) \\
&\quad + f_{NP}Y_B^{NP}(ct, \sigma_{ct} | f_-, f_+, f_{++}, c\tau_-, c\tau_+, c\tau_{++}, S_{ct}) \\
&= (1 - f_- - f_+ - f_{++})G(ct, \sigma_{ct} | S_{ct}) \\
&\quad + f_-E(-ct | c\tau_-) \otimes G(ct, \sigma_{ct} | S_{ct}) \\
&\quad + f_+E(ct | c\tau_+) \otimes G(ct, \sigma_{ct} | S_{ct}) \\
&\quad + f_{++}E(ct | c\tau_{++}) \otimes G(ct, \sigma_{ct} | S_{ct}), \tag{4.7}
\end{aligned}$$

where the width of the Gaussian is also scaled with the global scale factor S_{ct} , f_x is the fraction and $c\tau_x$ the lifetime of the given background component. Note that $f_- + f_+ + f_{++}$ is equal to the non-prompt fraction f_{NP} .

4.2.3 Decay Time Uncertainty Distribution

Comparing the decay time uncertainty distribution for signal and background in data, we observe that they are not the same. The background distribution can easily be obtained from the sidebands, whereas the signal distribution is derived by subtracting the distribution of the sidebands from the distribution in the signal region. If those distributions are different for signal and background, the decay time fit has to take the decay time uncertainty distribution into account [53]. The signal and background probability density function is multiplied by $Y'_S(\sigma_{ct})$ and $Y'_B(\sigma_{ct})$ respectively. The distributions are shown in Figure 4.1. These are normalized distributions of the decay time uncertainty, used directly in the likelihood function by taking the content of the bin for a given $\sigma_{c\tau}$ from the appropriate histogram.

4.2.4 Angular Distributions

To investigate angular distributions we use the transversity basis, as introduced in section 1.3.4, to describe the decay into two vector mesons. We will denote the three angles by $\vec{\omega} = \{\theta, \phi, \psi\}$.

The angular distribution for the signal is discussed separately for B_s and B_d later in this chapter, since it is not independent of the decay time distribution. Here we focus on the description of the background using an empirical set of functions. In first order we assume that the three angles are independent of each other, but we allow for correlation by adding the three functions, $f_4 - f_6$, which describe the interference terms in the signal model. We use

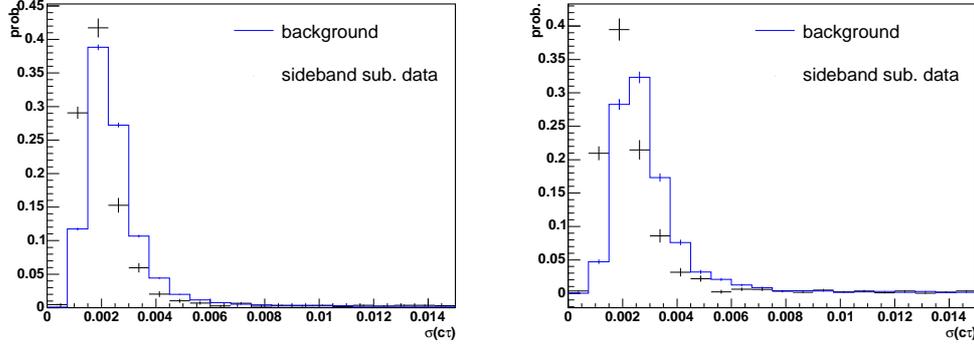


Figure 4.1: Normalized decay time uncertainty distributions for B_d (left) and B_s (right). The background (blue) is taken from the sidebands. Signal (black) is the sideband subtracted signal region.

the following empirical function for the background

$$\begin{aligned}
Z_B(\vec{\omega}|z_{0,\phi}, z_{1,\phi}, z_{2,\phi}, z_{1,\theta}, z_{2,\theta}, z_{1,\psi}, z_{2,\psi}, z_{3,\psi}, z_{4,\psi}, z_{5,\psi}, z_4, z_5, z_6) = \\
& \frac{(1 + z_{1,\phi} \cos(2\phi + z_{0,\phi}) + z_{2,\phi} \cos^2(2\phi + z_{0,\phi}))}{\pi(2 + z_{2,\phi})} \times \\
& \frac{1 - z_{1,\theta} \cos^2 \theta + z_{2,\theta} \cos^4 \theta}{2 - z_{1,\theta} 2/3 + z_{2,\theta} 2/5} \times \\
& \left(\frac{1 + z_{1,\psi} \cos(\psi) + z_{2,\psi} \cos^2(\psi) + z_{3,\psi} \cos^3(\psi)}{2 + 2/3 z_{2,\psi} + 2/5 z_{4,\psi}} \right. \\
& \quad \left. + \frac{z_{4,\psi} \cos^4(\psi) + z_{5,\psi} \cos^5(\psi)}{2 + 2/3 z_{2,\psi} + 2/5 z_{4,\psi}} \right) \\
& - z_4 \frac{9}{32\pi} \sin^2 \psi \sin 2\theta \sin \phi \\
& + z_5 \frac{9}{32\pi} \frac{1}{\sqrt{2}} \sin 2\psi \sin^2 \theta \sin 2\phi \\
& + z_6 \frac{9}{32\pi} \frac{1}{\sqrt{2}} \sin 2\psi \sin 2\theta \cos \phi. \tag{4.8}
\end{aligned}$$

This description is based on data from the sidebands and any detector and selection effects are already included in this empirical description. The good agreement of the fit for the sidebands can be seen in Figures 5.5 and 6.7.

4.2.5 Angular Acceptance

With a perfect detector and without any selection of the candidates we could assume to have homogeneous acceptance in the detector and it would be

easy to normalize the probability density function. Since this is not the case, we have to check and correct those distributions being affected by the detector geometry and the selection process. The J/ψ trigger has a uniform acceptance when measuring decay times. To keep the decay time distribution unaffected by the selection, we did not use any decay time related quantities for the soft candidate pre-selection or included them in the neural network. This is not the case for the angles $\vec{\omega}$ describing the decay. Figure 4.2 shows the distributions of the three angles for simulated B_s events after reconstruction. These events were generated using a flat distribution in all three angles. The effect from the detector and selection process on the angular distributions is too significant to be left alone and must be incorporated into the fit function. Since the distributions were generated flat, the acceptance function can be interpreted as probability to find an event at a position in the $\vec{\omega}$ space. It can be incorporated in the probability density function by multiplying the angular functions with the three dimensional acceptance function $\epsilon(\vec{\omega})$.

The acceptance is stored in a three dimensional histogram. Both, the angular distribution and the acceptance function, are normalized independently, thus their product might not be normalized anymore. The difficulty is now to keep the function used for fitting normalized for all possible values of the different parameters. The integral of the angular functions can be calculated analytically. After integrating over the volume of a single bin of the three dimensional acceptance histogram, the normalization factor can be calculated by summing over all bins. The analytic normalization of the probability density function and the acceptance is described in more detail in the appendix D. The normalization factor Z_n needs to be calculated whenever a parameter changes, but the normalization factor can be written in a way that the major part of the computation needs to be done only once (see equation D.3). The time dependent angular distribution of the signal including the acceptance $A(\vec{\omega})$ is then

$$\begin{aligned}
 Y_S(\vec{\omega}, t, \sigma_{ct} | \Gamma_L, \Gamma_H, |A_0(0)|, |A_{||}(0)|, |A_{\perp}(0)|, \delta_{\perp}, \delta_{||}, \phi_s, S_{ct}) \\
 = \frac{d^4 P(\vec{\omega}, t)}{d\vec{\omega} dt} \times A(\vec{\omega}) / Z_n \quad (4.9)
 \end{aligned}$$

4.2.6 Acceptance Function Correction

The signal only simulation (see section 3.4) used to estimate the acceptance function is not in perfect agreement with the data used in the analysis. The run period in the simulation does only cover the period up to 760 pb^{-1} . Further, there are always small differences in the simulation compared to

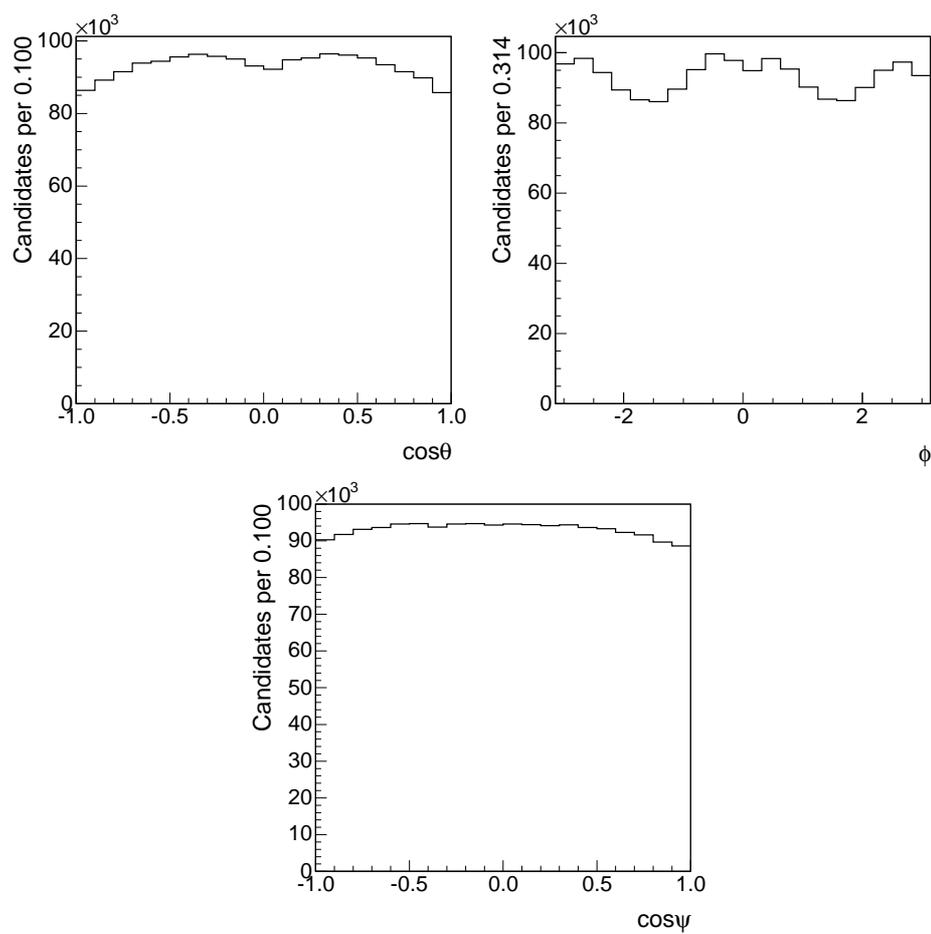


Figure 4.2: The distributions of the three angles describing the $B_s \rightarrow J/\psi \phi$ candidates after reconstruction of the simulated events. All three distributions were generated flat.

the data. We correct for these effects by comparing different variables and reweighting the simulated data before deriving the shape of the acceptance. The distribution of the data in the signal region is obtained by subtracting the background using the sidebands in the mass space.

We correct the simulated events for differences in some variables to obtain the acceptance histogram.

- Different ratio of di-muon pairs triggered by CMU-CMU and CMU-CMX triggers. This will influence the η distribution of the muons which might have different acceptance.
- Over time the minimal transverse momentum requirement changed and different trigger paths with different transverse momentum thresholds operate at the same time. To compensate this we order the candidates in three different classes of muon p_T : both muons with $p_T > 3$ GeV/c, both with $p_T > 2$ GeV/c but at least one below $p_T < 3$ GeV/c, everything else (at least on muon with $p_T < 2$ GeV/c).
- The B meson transverse momentum spectrum is taken from a histogram, which might slightly differ from the data. Therefore, we correct for the p_T of the B .

There are different ways how to correct a sample in different variables. If the variables are uncorrelated, we could just estimate the correction factor for each distribution, but in our case all three variables are correlated in some unknown way. To estimate the influence of this correction we implemented several methods to correct the acceptance function.

- Neglecting the correlation of the variables we just estimate the weights independently and re-weight according to the product of the weights.
- A consecutive approach, where we first correct for the trigger ratio, then re-weight using the di-muon transverse momentum classes and finally take the B meson transverse momentum spectrum into account. The weights of the second and third step is only calculated after correcting for the previous effects.
- The third method was a training of a neural network to estimate the individual weight for each event. Using a neural network we take possible correlations between the variables used for reweighting into account in an optimal way. We used more quantities as input variables than for the default method. Those are the transverse momentum and η variables of the B , ϕ , J/ψ and muons. To train the network with sideband

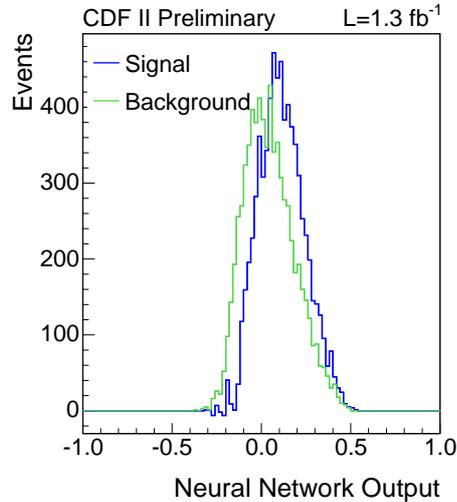


Figure 4.3: Network output distributions for the sideband subtracted data/signal (blue) and simulated/background events (green).

subtracted events we used the special feature of NeuroBayes that allows to train with negative training weights. By using data from the sidebands with positive training weights for background and additionally in the signal region with negative training weights we can train against the simulated data and let the network estimate the difference. The training with opposite weights is similar to create distributions with sideband subtracted background. The weight which accounts for the data-simulation difference is then calculated by: $P(\text{NN})/(1 - P(\text{NN}))$, where the network output, $\text{NN} \in [-1, 1]$, is transformed into a probability $P(\text{NN}) = (\text{NN} + 1)/2$. The result of the training can be seen in Figure 4.3. That both distributions are close to zero shows that in first order their agreement is quite good.

Although the neural network approach is in general the best way to correct for the disagreement between data and simulation, it might actually introduce some uncertainties. The issue is not the method itself, but the signal only simulation used to derive the acceptance. Since it was generated flat in the angular space, it might have some differences to the data which should not be corrected for. This is actually true for all of the reweighting procedures. Since all of the correction methods improve the agreement between data and simulation significantly and produce similar results, we take the remaining differences into account by treating them as systematic uncertainty. This is discussed later in Chapter 7. As can be seen in Figures 4.4 and 4.5 the data agrees well with the simulation after the reweighting.

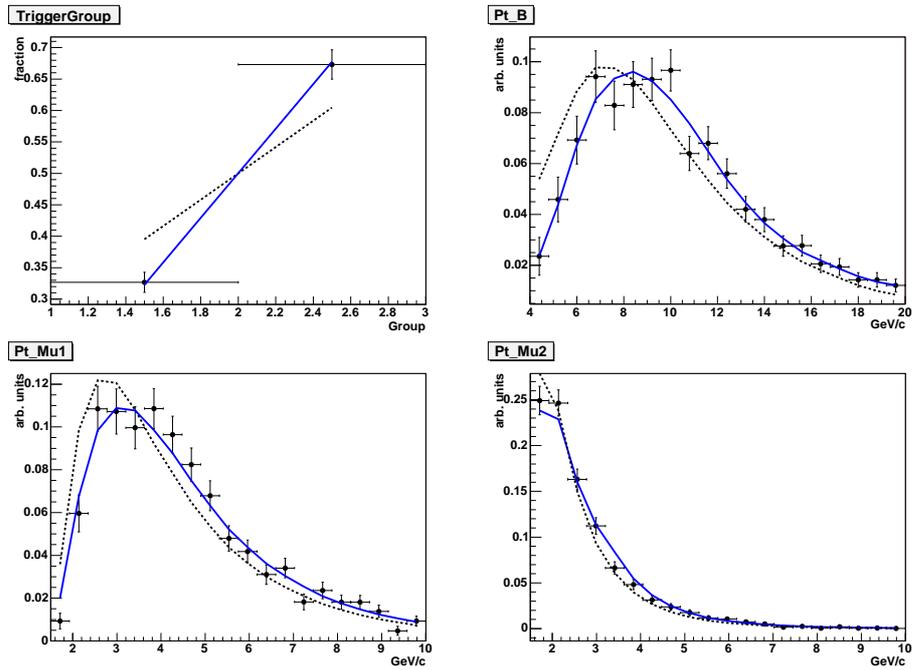


Figure 4.4: Comparison of sideband subtracted B_s data distribution (dots) with simulation before reweighting (dotted black line) and after reweighting (solid blue line). Shown are the fractions of CMX/CMU triggers (top left), transverse momentum of the B_s (top right) and the transverse momentum of the higher/lower p_T muon (bottom left/right).

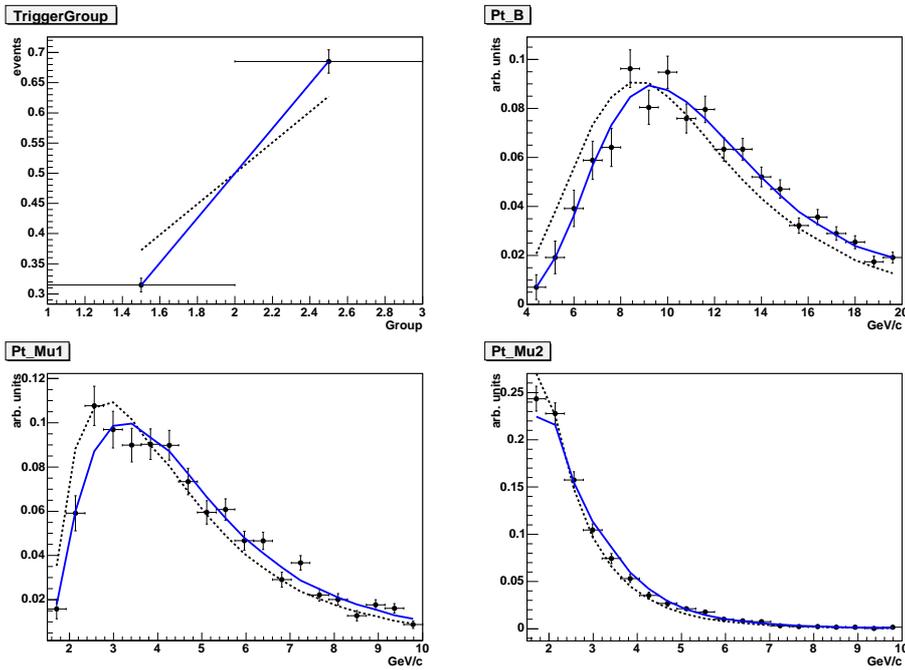


Figure 4.5: Comparison of sideband subtracted B_d data distribution (dots) with simulation before reweighting (dotted black line) and after reweighting (solid blue line). Shown are the fractions of CMX/CMU triggers (top left), transverse momentum of the B_d (top right) and the transverse momentum of the higher/lower p_T muon (bottom left/right).

4.2.7 B_d Acceptance

Similar to the B_s decay the B_d acceptance is affected by the detector acceptance and reconstruction. Again we use a simulated Monte Carlo sample which was generated with flat angular distribution to derive the acceptance. Figure 4.6 shows the acceptance of these angles for simulated B_d events after reconstruction. The angular function is then multiplied with the three dimensional acceptance and normalized.

The most striking difference to the B_s acceptance is the shape of the $\cos \psi$ distribution. The ψ angle is measured between the kaon from the K^* decay and the negative direction of the B meson in the K^* rest frame. There are two reasons for the observed differences. First, the mass difference of the kaon and pion of the K^* decay leads to an asymmetry in the distribution. The acceptance disfavors large values of $\cos \psi$, which is equivalent to the kaon going in the direction of the original K^* . This leads to a low momentum for the pion for those candidates, which reduces their reconstruction probability.

The other major difference is the ditch for $\cos \psi$ around -0.5. The acceptance for B_d candidates with swapped kaon-pion mass hypothesis is very different. Due to the wrongly assigned particle type the $\cos \psi$ distribution peaks around the value -0.5. Since we use a swap suppression network to attempt to suppress swapped candidates we remove part of the properly reconstructed B_d meson candidates in that $\cos \psi$ region.

4.3 B_s Log Likelihood Function

For the B_s the parameterization of the angular background is slightly modified. Since the $\cos(\psi)$ distribution is rather flat, we only use the first two parameters of lowest order and fix the other, $z_{3,\psi}$, $z_{4,\psi}$ and $z_{5,\psi}$, to zero. Also the ϕ offset $z_{0,\phi}$ is fixed to zero in the case of B_s .

As described in Chapter 1, we observe different CP final states, which have different angular momenta. Using the linear polarization amplitudes $A_0(t)$, $A_{||}(t)$ and $A_{\perp}(t)$ introduced in section 1.3.4, we can write the time dependent angular distribution as combinations of the bilinear decay amplitudes [54] as

$$\begin{aligned} \frac{d^4 P(\vec{\omega}, t)}{d\vec{\omega} dt} &\propto |A_0(0)|^2 f_1 \mathcal{T}_+ + |A_{||}(0)|^2 f_2 \mathcal{T}_+ + \\ &|A_{\perp}(0)|^2 f_3 \mathcal{T}_- + |A_0(0)||A_{||}(0)| f_5 \cos(\delta_{||}) \mathcal{T}_+ + \\ &|A_{||}(0)||A_{\perp}(0)| f_4 \cos(\delta_{\perp} - \delta_{||}) \sin \phi_s (e^{-\Gamma_H t} - e^{-\Gamma_L t})/2 + \\ &|A_0(0)||A_{\perp}(0)| f_6 \cos(\delta_{\perp}) \sin \phi_s (e^{-\Gamma_H t} - e^{-\Gamma_L t})/2, \quad (4.10) \end{aligned}$$

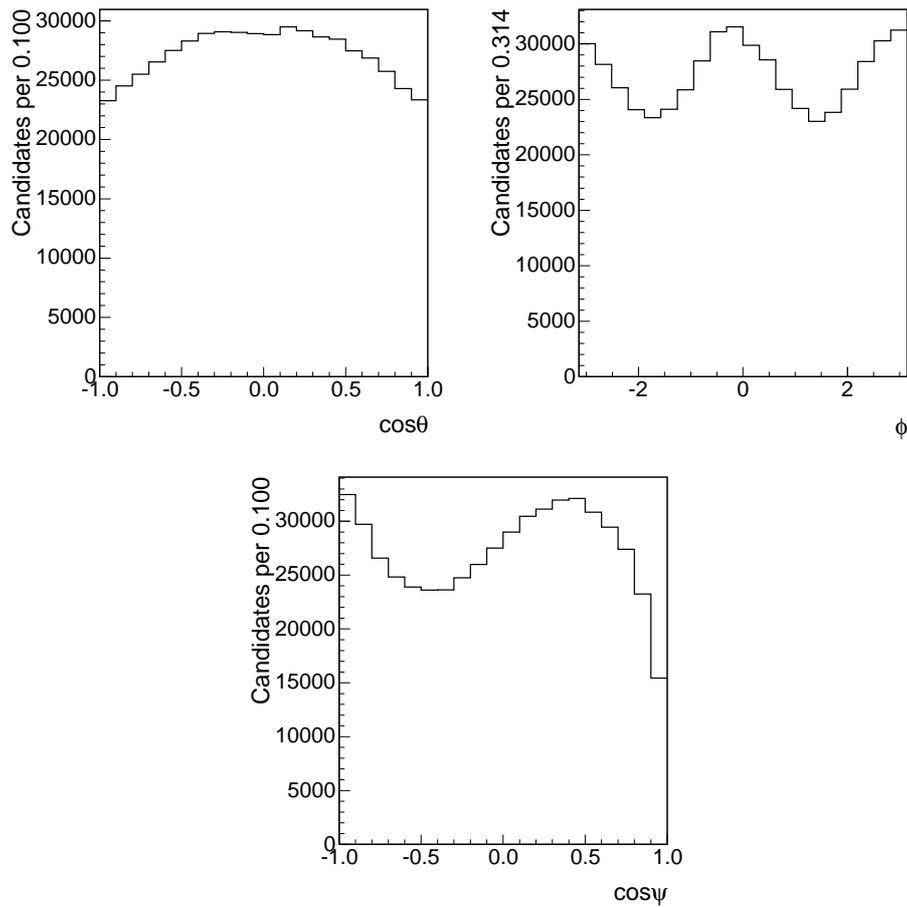


Figure 4.6: The distributions of the three angles describing the $B_d \rightarrow J/\psi K^*$ candidates after reconstruction of the simulated events without swapped candidates. All three distributions were generated flat.

where

$$\begin{aligned}
\mathcal{T}_{\pm} &= ((1 \pm \cos \phi_s)e^{-\Gamma_L t} + (1 \mp \cos \phi_s)e^{-\Gamma_H t})/2 \\
f_1(\vec{\omega}) &= \frac{9}{32\pi} 2 \cos^2 \psi (1 - \sin^2 \theta \cos^2 \phi) \\
f_2(\vec{\omega}) &= \frac{9}{32\pi} \sin^2 \psi (1 - \sin^2 \theta \sin^2 \phi) \\
f_3(\vec{\omega}) &= \frac{9}{32\pi} \sin^2 \psi \sin^2 \theta \\
f_4(\vec{\omega}) &= -\frac{9}{32\pi} \sin^2 \psi \sin 2\theta \sin \phi \\
f_5(\vec{\omega}) &= \frac{9}{32\pi} \frac{1}{\sqrt{2}} \sin 2\psi \sin^2 \theta \sin 2\phi \\
f_6(\vec{\omega}) &= \frac{9}{32\pi} \frac{1}{\sqrt{2}} \sin 2\psi \sin 2\theta \cos \phi
\end{aligned} \tag{4.11}$$

Note that the time dependent angular distribution is invariant under the transformations

$$\phi_s \rightarrow -\phi_s, \quad \delta_{\perp} \rightarrow \delta_{\perp} + \pi \tag{4.12}$$

and

$$\Delta\Gamma \rightarrow -\Delta\Gamma, \quad \phi_s \rightarrow \phi_s + \pi. \tag{4.13}$$

Because of this four fold ambiguity the measurement is insensitive to the signs of ϕ_s and $\Delta\Gamma$. In the following we will only show one of the four solutions. The other three equally likely solutions can be obtained by the given transformations.

For the background the detector acceptance was included in the empirical description. The signal shape is given by theory, but the detector itself and the selection process will introduce an angular acceptance function $A(\vec{\omega})$. We try to keep the acceptance as uniform as possible, but we have to take it into account. Thus the time dependent angular probability density function becomes

$$Y_S(c\tau, \sigma_{c\tau}, \vec{\omega}, \xi) \cdot Y'_S(\sigma_{ct}) = \frac{d^4 P(\vec{\omega}, t)}{d\vec{\omega} dt} \otimes G(ct, \sigma_{ct} | S_{ct}) A(\vec{\omega}) / Z_n Y'_S(\sigma_{ct}). \tag{4.14}$$

All the necessary parts of the likelihood function are now introduced, and we can combine them. Whereas the probability density function for prompt and non prompt background factorizes in mass, decay time and angular space, and is therefore just the product of three probability density functions, this is not the case for the signal. As described in section 4.2.5 we have to

normalize the time dependent angular part for each change in parameters. Starting from equation 4.1, the complete $\ln \mathcal{L}$ of the time-dependent angular analysis is

$$\begin{aligned}
\ln \mathcal{L} &= \sum_{j=1}^N \ln [f_s X_S(m_j) Y'_S(\sigma_{ct_j}) Z_S(\vec{\omega}_j, t_j, \sigma_{ct_j}) + \\
&\quad (1 - f_s) [(1 - f_{NP}) X_B^P(m_j) Y_B^P(ct_j, \sigma_{ct_j}) + \\
&\quad f_{NP} X_B^{NP}(m_j) Y_B^{NP}(ct_j, \sigma_{ct_j})] Y'_B(\sigma_{ct_j}) Z_B(\vec{\omega}_j)] \\
&= \sum_{j=1}^N \ln [f_s X_S(m_j) \frac{d^4 P(\vec{\omega}, t)}{d\vec{\omega} dt} \otimes G(ct, \sigma_{ct} | S_{ct}) \times A(\vec{\omega}) / Z_n + \\
&\quad (1 - f_s) [(1 - f_{NP}) X_B^P(m_j) Y_B^P(ct_j, \sigma_{ct_j}) \\
&\quad + f_{NP} X_B^{NP}(m_j) Y_B^{NP}(ct_j, \sigma_{ct_j})] Y'_B(\sigma_{ct_j}) Z_B(\vec{\omega}_j)], \tag{4.15}
\end{aligned}$$

where the individual part are defined in equation 4.3 to 4.8 and $Y'_B(\sigma_{ct_j})$ is shown in Figure 4.1 and $A(\vec{\omega})$ in Figure 4.2.

We then minimize $-2 \ln \mathcal{L}$ using MINUIT [55]. In the actual implementation of the fit we did not use exactly the parameters described above but the linear combination of the interesting parameters to obtain directly the quantities of highest interest. Instead of Γ_H and Γ_L we used the decay width difference $\Delta\Gamma = \Gamma_H - \Gamma_L$ and the mean lifetime $c\tau_s = 2c/(\Gamma_H + \Gamma_L)$ as fit parameters.

$$\Gamma = \frac{c}{c\tau_s} \tag{4.16}$$

$$c\tau_L = \frac{c\tau_s}{1 + 0.5\Delta\Gamma\tau_s} \tag{4.17}$$

$$c\tau_H = \frac{c\tau_s}{1 - 0.5\Delta\Gamma\tau_s} \tag{4.18}$$

$$\tag{4.19}$$

Since the sum of the amplitudes should be one, one of them is given by the other two. Further the amplitude parameterization is changed to guarantee that the amplitudes stays in physically allowed range. This stabilizes the minimization procedure, since no unphysical values are allowed. The parameters needed can be calculated from the fit result by the transformations

$$A_{\parallel} = \sqrt{1 - A_0^2 A'_{\parallel}} \tag{4.20}$$

$$A_{\perp} = \sqrt{1 - A_0^2 - A_{\parallel}^2}. \tag{4.21}$$

4.3.1 Case of no CP Violation

Assuming the Standard Model, where ϕ_s is of the order of one degree we can fix $\phi_s = 0$. In such a case the fit function simplifies and \mathcal{T}_\pm in equation 4.11 and 4.10 become

$$\begin{aligned}\mathcal{T}_+ &= e^{-\Gamma_L t} \\ \mathcal{T}_- &= e^{-\Gamma_H t}\end{aligned}$$

In addition the interference terms between CP-even and CP-odd eigenstates drops out. Thus equation 4.10 gets

$$\begin{aligned}\frac{d^4 P(\vec{\omega}, t)}{d\vec{\omega} dt} &\propto |A_0(0)|^2 f_1 e^{-\Gamma_L t} + |A_{||}(0)|^2 f_2 e^{-\Gamma_L t} + \\ &\quad |A_{\perp}(0)|^2 f_3 e^{-\Gamma_H t} + |A_0(0)||A_{||}(0)| f_5 \cos(\delta_{||}) e^{-\Gamma_L t} \\ &\propto [|A_0(0)|^2 f_1 + |A_{||}(0)|^2 f_2 + |A_0(0)||A_{||}(0)| f_5 \cos(\delta_{||})] e^{-\Gamma_L t} + \\ &\quad |A_{\perp}(0)|^2 f_3 e^{-\Gamma_H t}\end{aligned}\tag{4.22}$$

Comparing equations 4.10 and 4.22 we see that the second one does not depend on the strong phase δ_{\perp} anymore, which therefore cannot be estimated.

4.4 B_d Log Likelihood Function

To describe the decay $B_d \rightarrow J/\psi K^*$ we can use the same definition for the angles as before. But there are few differences and complications. The lifetime difference in the B_d system is estimated to be very small, $(3.0 \pm 1.2)10^{-3}$ [56], and measurements so far agree with zero. In the case of B_s , the mass eigenstates are close to CP eigenstates, which is not the case for B_d . Further $B_d \rightarrow J/\psi K^*$ is not a common decay mode, but a flavour-specific decay. Taking these effects together, the time-dependence for an initially produced $B_d(\bar{B}_d)$ gets [56]

$$0.5[\cosh(\Delta\Gamma t/2) \pm \cos(\Delta m_d t)]e^{-\Gamma t}.\tag{4.23}$$

We assume the same number of B_d and \bar{B}_d and expanding in $\Delta\Gamma$. When ignoring terms quadratic in $\Delta\Gamma$ the time dependent angular distribution is

$$\begin{aligned}
\frac{d^4P(\vec{\omega}, t, \xi)}{d\vec{\omega}dt} &\propto [|A_0(0)|^2 f_1(\vec{\omega}) + |A_{||}(0)|^2 f_2(\vec{\omega}) + \\
&|A_{\perp}(0)|^2 f_3(\vec{\omega}) + \text{Re}(A_0^*(0)A_{||}(0))f_5(\vec{\omega}) \pm \\
&\text{Im}(A_{||}^*(0)A_{\perp}(0))f_4(\vec{\omega}) \pm \\
&\text{Im}(A_0^*(0)A_{\perp}(0))f_6(\vec{\omega})] \times e^{-\Gamma_{dt}} \\
&= \{|A_0(0)|^2 f_1(\vec{\omega}) + |A_{||}(0)|^2 f_2(\vec{\omega}) + \\
&|A_{\perp}(0)|^2 f_3(\vec{\omega}) + |A_0(0)||A_{||}(0)|f_5(\vec{\omega}) \cos(\delta_{||}) \pm \\
&|A_{||}(0)||A_{\perp}(0)|f_4(\vec{\omega})[\sin(\delta_{\perp} - \delta_{||})] \pm \\
&|A_0(0)||A_{\perp}(0)|f_6(\vec{\omega})[\sin(\delta_{\perp})]\} \times e^{-\Gamma_{dt}}, \quad (4.24)
\end{aligned}$$

where the upper (lower) arithmetic operator denotes the flavor specific decay of a B_d (\bar{B}_d) to $K^+\pi^-$ ($K^-\pi^+$). The angular function are defined above. The likelihood is put together in the same way as for B_s and looks the same as equation 4.15. The implementation of both likelihoods is actually done only once, by implementing the most general likelihood function, described in the next section.

4.5 Fit Procedure

There are several minor implementation issues, which are discussed in this section. Instead of using two different fitters for B_d and B_s , it is possible to use the same likelihood function, but fixing some parameters or setting some variables in a way that it can be used for B_d and B_s .

Equation C.8 is the most general probability density function for a tagged sample. This function can be used for the fit in B_s with

- fixing the parameter $\delta_0 = 0$
- setting the product $\xi D = 0$.

Equation C.8 does then agree with equation 4.10. To fit B_d data the following is done

- fixing the parameter $\delta_0 = 0$
- fixing $\Delta\Gamma = 0$
- setting the variable $\xi D = \pm 1$ depending on the charge of the kaon from K^*

- fixing $\phi_s = 0$
- fixing $\Delta m = 0$.

Fixing $\Delta m = 0$ does not mean we fix the mixing frequency for B_d to zero, but has purely technical reasons to get an agreement between the two formulas. Equation C.8 then reduces to equation 4.24. This allows to use a single implementation for both analyses. Further the fit function is already prepared for future measurements using flavor tagged B_s samples.

When fitting such a complicated function, which depends on many parameters, it is necessary to start with reasonable parameter values. To get a good starting point for the final fit, a sequence of maximum likelihood fits is performed. In each step the parameters derived in the step before are used as starting values for the next step. First the background lifetime parameters are determined in a fit to the events in the mass sidebands. These sideband events are then used in the second step to fit the mass, decay time and angular distribution of the background. In the third step the mass of all events is fitted. The following fit of mass, lifetime and angles is the final fit for B_d . For B_s , two mass, decay time and angular fits of all events are performed. First ϕ_s and one strong phase is fixed to 0, which is equivalent to Standard Model expectation, then all parameters are kept floating.

Not only the values, but also the uncertainties are estimated by the fit. The default uncertainty calculation of Minuit uses the error matrix calculated in Migrad as approximation, which is valid in case the negative log likelihood function is parabolic around the minimum. For the final results, we use Minos to calculate the positive and negative error. Minos varies one parameters with respect to all others to find the parameter value for which the likelihood has a value $F(\hat{a}) + 1$. Therefore, Minos obtains the one sigma region above and below the minimum allowing to quote asymmetric and more precise uncertainties.

Chapter 5

B_d Results

As shown in the previous chapter and in Chapter 1, the fit function for B_d and B_s have large parts in common and therefore they are implemented as a single probability density function used for B_d and B_s . Thus, we are able to check our method and implementation of the angular fit and amplitude decomposition of the B_s decay by analyzing the independent B_d sample, which has more statistics. BaBar [57] and Belle [58] made similar angular analyses of $B_d \rightarrow J/\psi K^*$, allowing us to cross-check our result with their independent measurements.

Besides of the fit result, the next sections will present also a validation of the fitting framework. We will present pull distributions, likelihood scans and the fit projections. As the data is of high quality, we begin to be sensitive to small details of the signal description, which require to introduce the $K\pi$ S-wave component in the B_d description.

5.1 $K\pi$ S-wave Contribution

Up to now we concentrated the discussion of the B_d decay on the $B_d \rightarrow J/\psi K^*$ decay itself, the swap component, where pion and kaon hypothesis is swapped, and the empirical background parameterization. But when checking the projections of the fit results using the previously defined time dependent decay rate (equation 4.24), especially the $\cos\psi$ distribution (see Figure 5.1) does not seem to be described very well by the data anymore

The reason for the discrepancy is another contribution to the signal, which is not covered by the signal probability density function and not included in the background parameterization. We analyse the B_d decay to two vector mesons, but the decay $B_d \rightarrow J/\psi K\pi$ is also possible and previous studies [59] have shown, that it is not negligible. In principle the $K\pi$ can have any

integer spin, but the S and P waves are the main contributions. We can extend the differential decay rate to include the $K\pi$ S-wave contribution and the interference between the S and P wave. The signal probability density function for a B_d meson gets

$$\begin{aligned}
\frac{d^4 P(\vec{\omega}, t, \xi)}{d\vec{\omega} dt} &\propto \\
&= \cos^2 \lambda \{ [|A_0(0)|^2 f_1(\vec{\omega}) + |A_{||}(0)|^2 f_2(\vec{\omega}) \\
&+ |A_{\perp}(0)|^2 f_3(\vec{\omega}) + |A_0(0)||A_{||}(0)| f_5(\vec{\omega}) \cos(\delta_{||}) \\
&+ |A_{||}(0)||A_{\perp}(0)| f_4(\vec{\omega}) [\sin(\delta_{\perp} - \delta_{||})] \\
&+ |A_0(0)||A_{\perp}(0)| f_6(\vec{\omega}) [\sin(\delta_{\perp})] \} \\
&+ \sin^2 \lambda f_7(\vec{\omega}) \\
&+ 1/2 \sin 2\lambda (f_8(\vec{\omega}) \cos(\delta_{||} - \delta_S) |A_{||}| + f_9(\vec{\omega}) \sin(\delta_{\perp} - \delta_S) |A_{\perp}| \\
&+ f_{10}(\vec{\omega}) \cos(\delta_S) |A_0|) \} \\
&\times e^{-\Gamma_d t}, \tag{5.1}
\end{aligned}$$

where δ_s is another strong phase due to the S-wave component, $A_S = |A_S|e^{i\delta_s}$, and λ is proportional to its amplitude. It is defined as:

$$\begin{aligned}
\cos \lambda &= \frac{A_P}{\sqrt{A_P^2 + |A_S|^2}} \\
\sin \lambda &= \frac{|A_S|}{\sqrt{A_P^2 + |A_S|^2}},
\end{aligned}$$

where A_S is the amplitude of the $K\pi$ S-wave and A_P is the amplitude for the resonant K^* P-wave. For more details see reference [60]. The f functions are defined as

$$\begin{aligned}
f_7(\vec{\omega}) &= \frac{3}{32\pi} 2(1 - \sin^2 \theta \cos^2 \phi) \\
f_8(\vec{\omega}) &= -\frac{3}{32\pi} \sqrt{6} \sin \psi \sin^2 \theta \sin 2\phi \\
f_9(\vec{\omega}) &= \frac{3}{32\pi} \sqrt{6} \sin \psi \sin 2\theta \cos \phi \\
f_{10}(\vec{\omega}) &= \frac{3}{32\pi} 4\sqrt{3} \cos \psi (1 - \sin^2 \theta \cos^2 \phi). \tag{5.2}
\end{aligned}$$

In Figure 5.1 we can see the improvement in the fit projection when allowing for the $K\pi$ S-wave contribution in the fit function. The same behaviour was seen by the B factories [58, 60]. Similar to the swap component in the B_d signal, we can either include the S-wave in the fit function or add it to

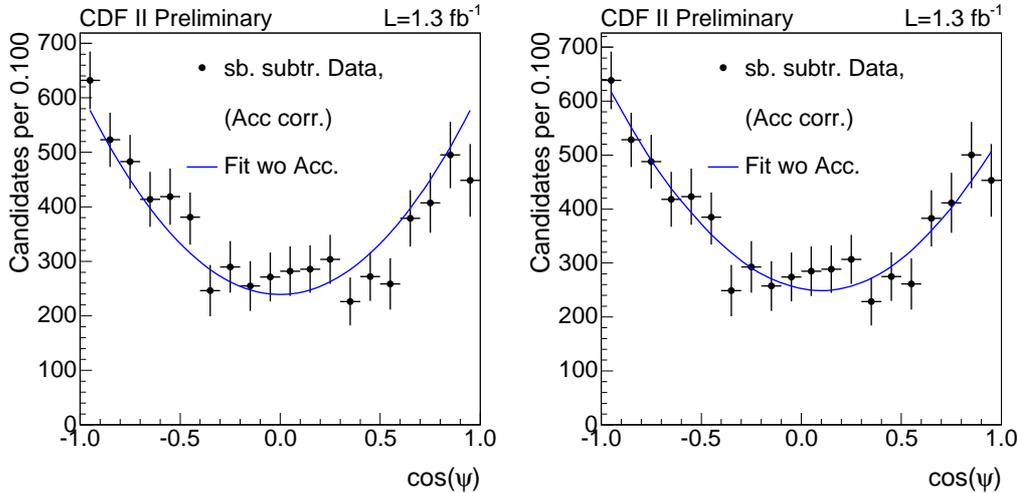


Figure 5.1: Sideband subtracted and acceptance corrected $\cos(\psi)$ fit projection for B_d without (left) and with (right) $K\pi$ S-wave contribution.

the systematic uncertainties. As this is a more complete description of the physics, we use the fit including the $K\pi$ S-wave as baseline fit and assign the difference to the fit without S-wave as systematic uncertainty.

5.2 Fit Verification

For an experiment there is typically only one set of data, which is used in the fit to derive the best estimate of the parameters. We do know the true values only by some uncertainty from other experiments or by theoretical predictions. Any sample is subject to statistical fluctuations, therefore we want to know whether the fit can reproduce the true parameter within its statistical uncertainties or not.

Typically this is done by so called toy studies using pseudo experiments. Each pseudo experiment consists of events which are generated from distributions used in the probability density function. For example, the mass distribution of the signal is simulated by producing random numbers with a Gaussian distribution. In a similar way we can generate decay time and angular distributions, which allows us to test the time dependent angular analysis we want to use on data. If we create such a pseudo experiment with infinite statistics, the fitter should reproduce the parameters we used for generating the distributions. If we want to have a more realistic understanding of the real data, we generate the same number of events we observe in the experiment. The single pseudo experiment will suffer from the same

statistical uncertainty as in data. If we generate hundreds of independent pseudo experiments, we can study the properties of the fit function. This can be conveniently studied by creating pull distributions of the pseudo experiments for each parameter derived from the fit. The pull for a parameter p is defined as

$$\frac{p_{fit} - p_{true}}{\sigma_{p_{fit}}}, \quad (5.3)$$

where p_{fit} and $\sigma_{p_{fit}}$ are the fit result of the parameter and its uncertainty for the given pseudo experiment, and p_{true} is the value used to produce the distribution. In the ideal case, if the fit gives the correct answer, the pull distributions of the pseudo experiments are Gaussians around zero with a width of one. If this is not the case, we cannot expect that the estimated parameter and its uncertainty from data will be correct.

The following pages show plots from pseudo experiment studies for $B_d \rightarrow J/\psi K^*$. There are four plots for each studied variable arranged in a square. The lower left plot shows the distribution of the fitted value. The input value is indicated by an arrow. The lower right plot shows the distribution of the uncertainty returned by the fit. The difference between fitted and input value is displayed in the upper right plot. Finally the upper left plot shows the distribution of the pulls.

1000 pseudo experiments are generated with the same number of events as in data. The input values of the parameters are listed in Table 5.1. Figures 5.2, 5.3 and 5.4 show the parameter distributions for the toy study. Table 5.2 summarizes the result of the pull distributions for the signal parameters. The mean and width of the Gaussian used to fit the pull distribution agree very well with the expectations.

parameter	value	parameter	value
f_s	0.143	f_{++}	0.069
M	5.28 GeV	f_+	0.077
σ	8.14 MeV	f_-	0.012
f_m	0.3	$c\tau_{++}$	0.042
s_m	2.0	$c\tau_+$	0.0062
$c\tau$	460 μm	$c\tau_-$	0.020
A_0	0.75	$z_{1,\phi}$	0.176
$\delta_{ }$	-2.9	$z_{2,\phi}$	-0.036
$A_{ }$	0.4	$z_{1,\theta}$	0.60
δ_{\perp}	2.9	$z_{2,\theta}$	0.33
a	-0.315	$z_{1,\psi}$	0.80
a_{NP}	-5.2	$z_{2,\psi}$	0.51
S_{ct}	1.37	$z_{3,\psi}$	-0.20
		$z_{4,\psi}$	-0.70
		$z_{5,\psi}$	-1.0
		$z_{0,\phi}$	0.2

Table 5.1: Parameter values used to create pull distributions for B_d .

Par.	Mean	signi.	Width	Fit prob.	Shift
$c\tau$	0.0527 ± 0.0326	1.61	1.03 ± 0.0231	0.0867	4.91e-05
$ A_0 $	0.0503 ± 0.0313	1.61	0.991 ± 0.0222	0.757	0.000238
$\delta_{ }$	0.0253 ± 0.0296	0.855	0.937 ± 0.0209	0.682	0.00103
$ A_{ } $	-0.0487 ± 0.0326	1.49	1.03 ± 0.023	0.0886	-0.00121
δ_{\perp}	-0.00433 ± 0.0311	0.139	0.984 ± 0.022	0.815	-0.00149

Table 5.2: Summary of the Gaussians fitted to the pull distributions for the individual parameters. The significance of the deviation of the mean is listed in the third column. The fit probability of the Gaussians are listed in the fifth and the shift of the mean of the parameter distribution to the value used for production is in the sixth column.

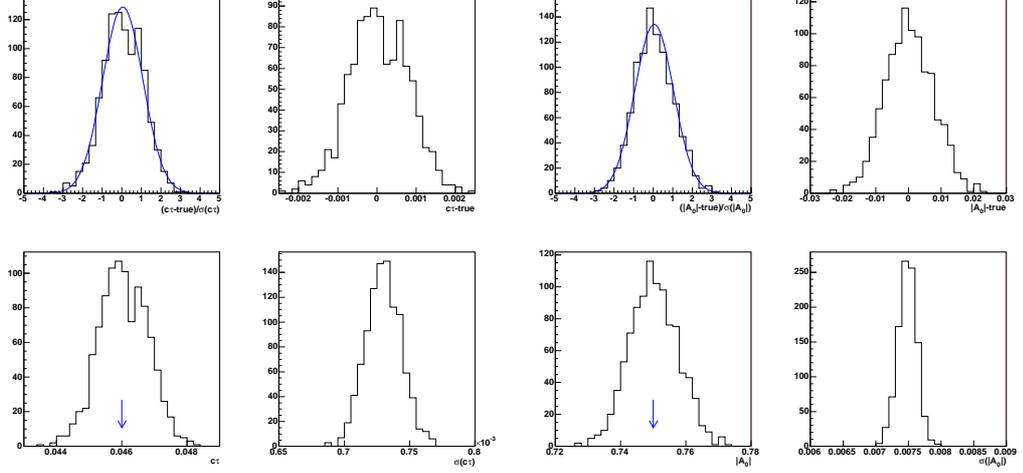


Figure 5.2: Distribution of pull (upper left), parameter (lower left), error (lower right) and parameter shifted with respect to the production value (upper right) for $c\tau$ (four left plots) and $|A_0|$ (four right plots) for B_d . The production value is denoted by the arrow.

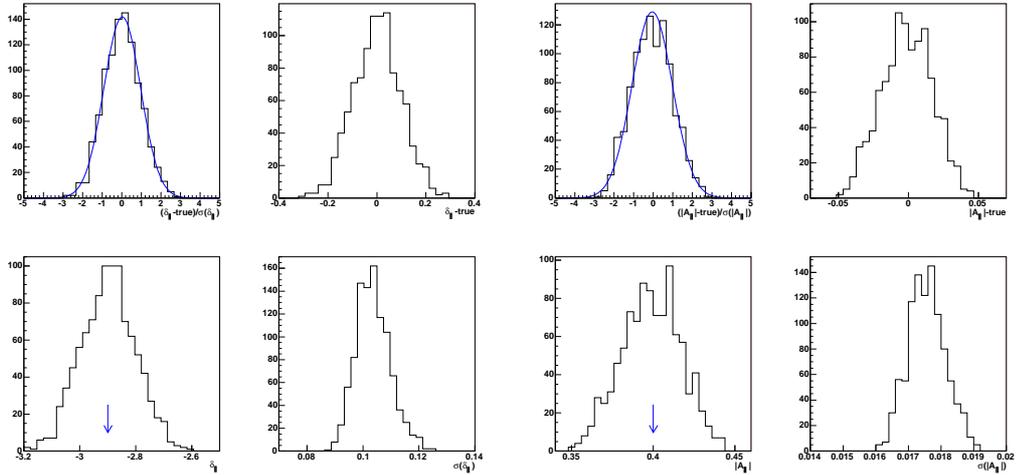


Figure 5.3: Distribution of pull (upper left), parameter (lower left), error (lower right) and parameter shifted with respect to the production value (upper right) of $\delta_{||}$ (four left plots) and $|A_{||}|$ (four right plots) for B_d . The production value is denoted by the arrow.

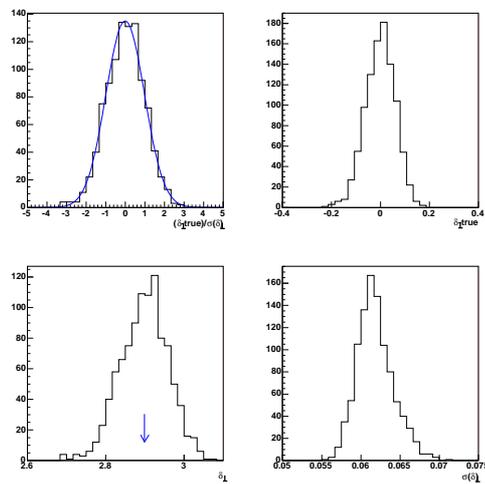


Figure 5.4: Distribution of pull (upper left), parameter (lower left), error (lower right) and parameter shifted with respect to the production value (upper right) of δ_{\perp} for B_d . The production value is denoted by the arrow.

5.3 Fit Projections

Another very important check is the agreement of the fit model with the derived parameters and the data. For a multidimensional fit, this cannot be easily visualized in the multidimensional space. Therefore, the projection for each of the variables is done by integrating over all other variables. These one dimensional distribution can be then easily compared to data.

As described in Chapter 4 the parameters of the background are estimated from the mass sidebands before the final fit. The empirical parameterization of the background can be checked using the projection of the sidebands. In Figure 5.5 the projections of the angles and decay time are shown. The fit does agree very well with the background. The mass distribution is not shown, since we remove the events in the signal mass region.

The final fit projections are shown for in Figure 5.6. The full fit results, as well as the signal (including the $K\pi$ S-wave contribution) and background part are plotted separately. Besides the fit projection comparison, there are more analytical ways to compare the agreement between data and the fit results. χ^2 , defined as

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - y_{t,i})^2}{y_i} \quad (5.4)$$

can be used to quantify the agreement of two distributions. Here y_i are the number of entries in bin i and $y_{t,i}$ is the prediction from the fit. The χ^2 value should be roughly equal to the number of bins, n , minus the number of free parameters, $\chi^2 \approx (n - n_{par.})$. The number of entries per bin is required to be greater than ten to allow for a meaningful result. In principle the χ^2 value should be calculated on the multidimensional space. If the likelihood factorizes, the independent distributions can be checked. The probability to observe such a χ^2 value or larger for the given number of degrees of freedom, ndf, is 98% ($\chi^2/ndf = 2561/2715$) for the angular function, 66% ($\chi^2/ndf = 80/86$) for the decay time distribution and 99% ($\chi^2/ndf = 47/71$) for the mass distribution.

5.4 Likelihood Profiles

Further the shape of the likelihood around the minimum should be studied. There are slightly different ways to study the likelihood around the minima. The most simple way is to vary the parameter of interest around the minima and calculate the value of the probability density function, keeping the rest of the parameters at their values at the global minimum. The more advanced

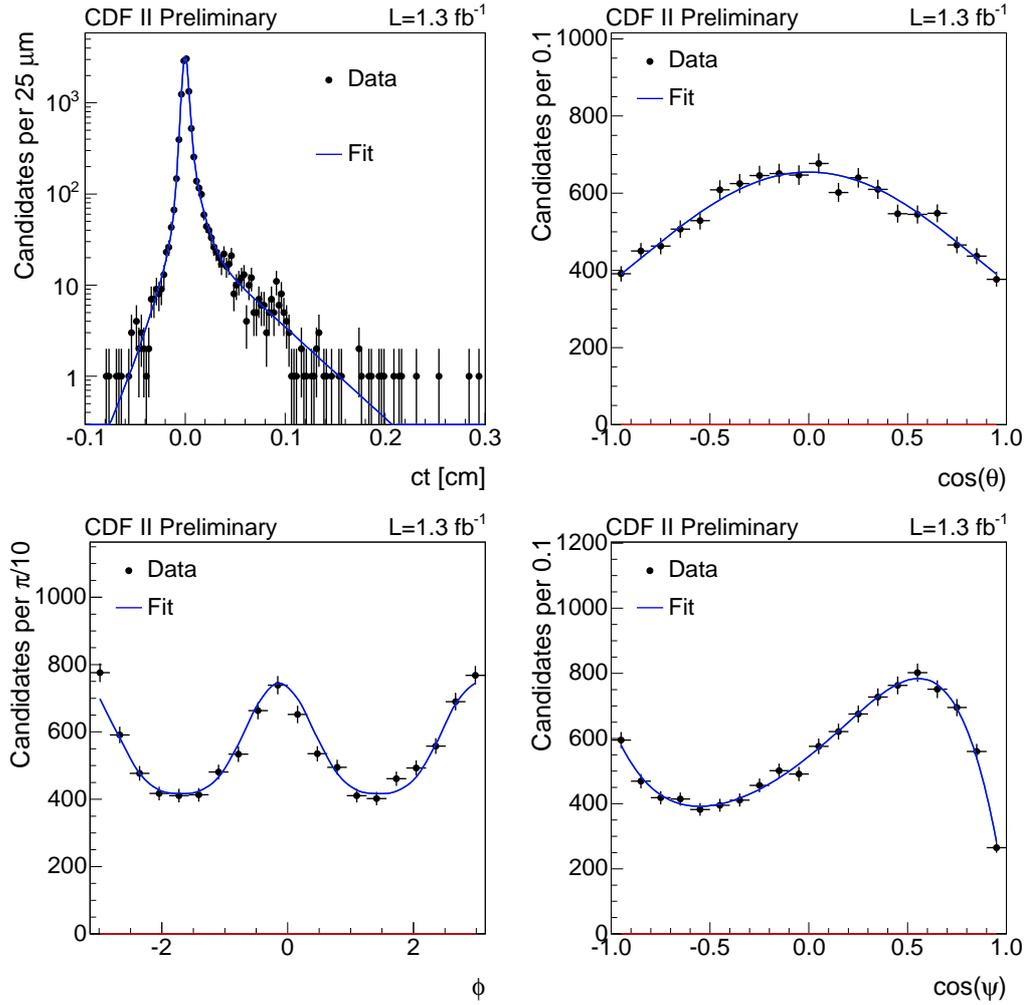


Figure 5.5: B_d decay time and angular projections of the upper and lower sidebands only.

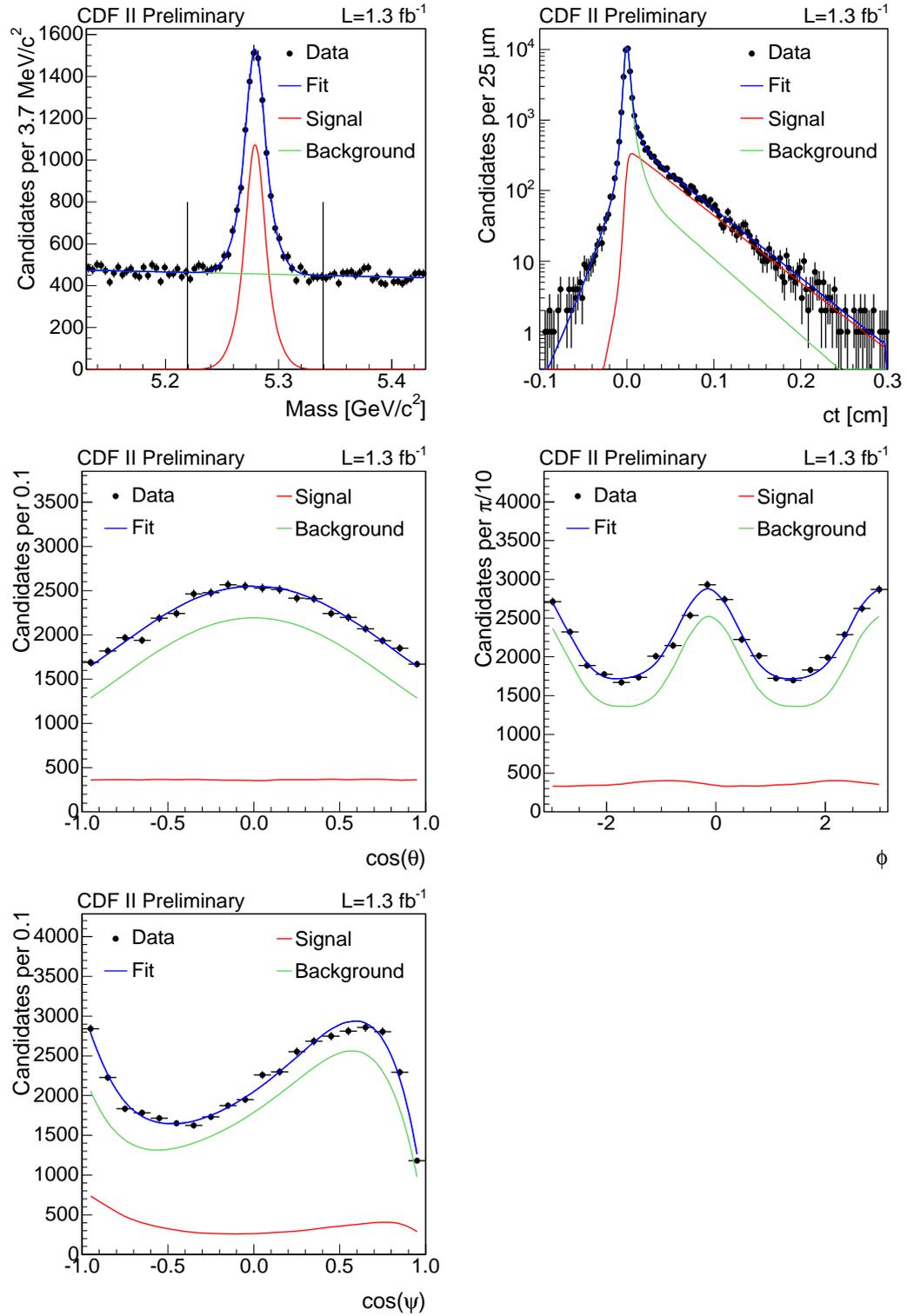


Figure 5.6: Projections of the results of the unbinned maximum likelihood fit of the B_d decay.

way would be to minimize for each point where the likelihood function is evaluated with respect to all other parameters. In case that a parameter is uncorrelated to all others, both methods will give the same results. For infinite statistics, the shape of the likelihood will be parabolic. With less statistics this might not be the case.

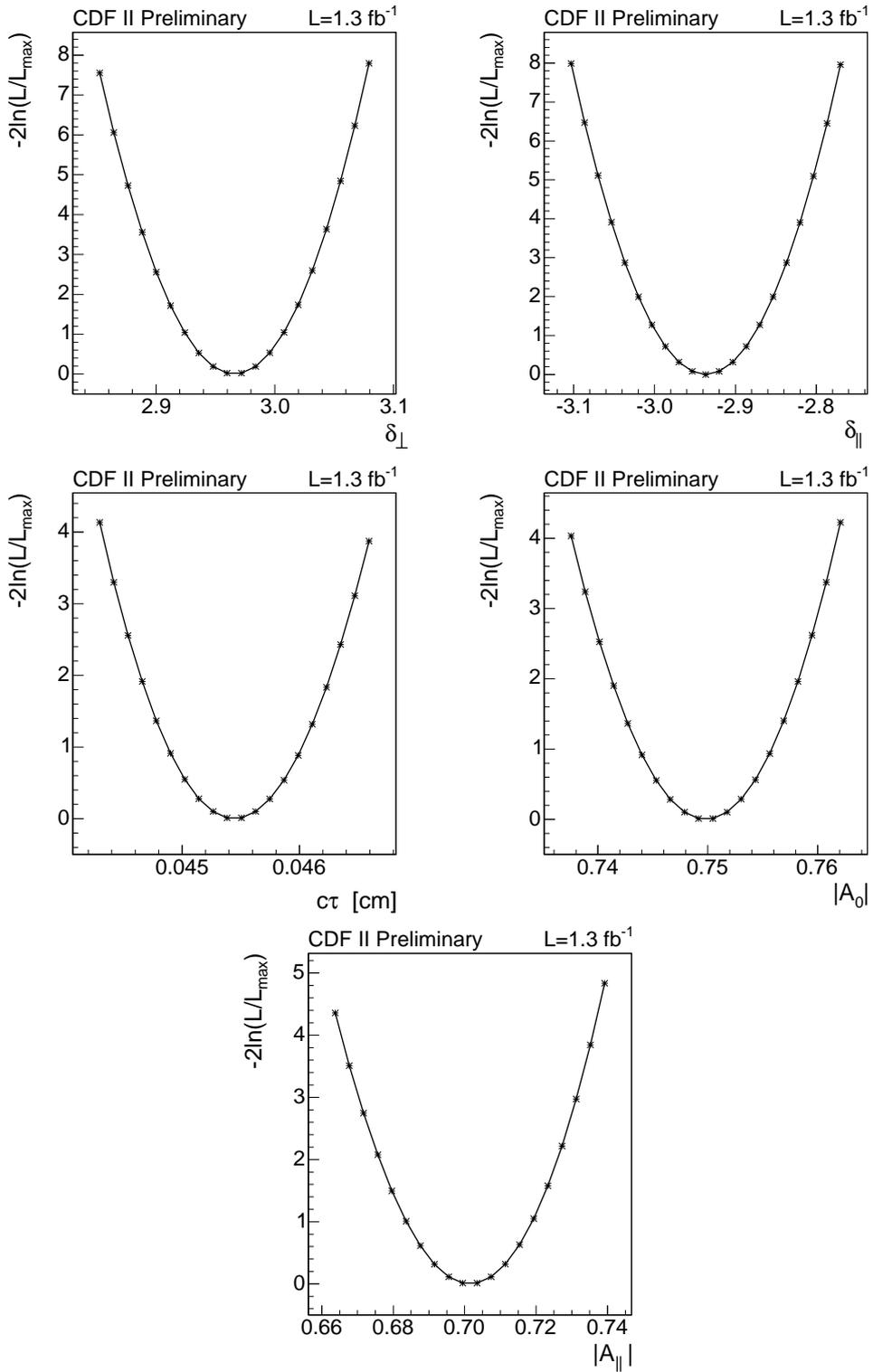
We scanned the most interesting signal parameters by calculating the likelihood values for 20 interpolation points in the two σ region around the parameter minimum. The shape of the likelihood function can be seen in Figure 5.7. On the y-axis the difference of the likelihood to the global minimum is shown. All distributions are close to the expected parabolic shape and show that the fit converged to a local minimum.

5.5 Results

None of the cross checks for the B_d decay shows any complications. They give us confidence that the fit result and its uncertainties are properly estimated by the likelihood method. The final results for all the B_d fit parameters are summarized in Table 5.3. They are in good agreement with previous measurements [57, 58]. The correlation matrix of the signal parameters of the final fit are listed in Table 5.4. The systematic uncertainties are discussed in Chapter 7. The fraction of $K\pi$ S-wave contribution is estimated to be $1.6 \pm 0.8\%$, which is similar to the estimation by the BaBar experiment [61].

We obtain a signal yield of $\approx 7567 \pm 98$ from 73726 events in the whole region. The fitted parameters for the signal are

$$\begin{aligned}
 c\tau_d &= 455 \pm 6 \mu\text{m} \\
 |A_0|^2 &= 0.564 \pm 0.009 \\
 |A_{||}|^2 &= 0.214 \pm 0.012 \\
 \delta_{||} &= -2.94 \pm 0.08 \\
 \delta_{\perp} &= 2.96 \pm 0.06
 \end{aligned} \tag{5.5}$$

Figure 5.7: B_d likelihood profiles for the signal parameters.

parameter	lt. bkg. fit	bkg. fit	mass fit	fit wo $K\pi$ S-w.	fit with $K\pi$ S-w.
f_s	-	-	0.1663 ± 0.0034	0.1660 ± 0.0021	0.1660 ± 0.0021
M	-	-	5.2793 ± 0.0002	5.2792 ± 0.0001	5.2792 ± 0.0001
σ	-	-	0.0071 ± 0.0007	0.0088 ± 0.0003	0.0088 ± 0.0003
f_m	-	-	0.582 ± 0.121	0.306 ± 0.042	0.306 ± 0.042
s_m	-	-	2.024 ± 0.131	-	-
$c\tau$ [μm]	-	-	-	454.7 ± 5.8	455.2 ± 5.8
$ A_0 $	-	-	-	0.755 ± 0.006	0.751 ± 0.006
δ_{\parallel}	-	-	-	-2.98 ± 0.087	-2.94 ± 0.083
$ A_{\parallel} $	-	-	-	0.697 ± 0.018	0.700 ± 0.019
δ_{\perp}	-	-	-	2.96 ± 0.06	2.96 ± 0.06
λ	-	-	-	-	0.123 ± 0.035
δ_s	-	-	-	-	2.27 ± 0.223
a	-	1.42 ± 0.35	-0.640 ± 0.243	-0.20 ± 0.24	-0.20 ± 0.24
a_{NP}	-	-2.16 ± 0.96	-	-4.13 ± 0.65	-4.11 ± 0.65
f_{++}	0.0546 ± 0.0049	0.0545 ± 0.0050	-	0.0542 ± 0.0032	0.0547 ± 0.0033
f_{+}	0.0977 ± 0.0073	0.0975 ± 0.0073	-	0.1016 ± 0.0045	0.1016 ± 0.0045
f_{-}	0.0183 ± 0.0022	0.0180 ± 0.0022	-	0.0184 ± 0.0012	0.0184 ± 0.0012
$c\tau_{++}$	0.0434 ± 0.0028	0.0434 ± 0.0029	-	0.0398 ± 0.0017	0.0397 ± 0.0017
$c\tau_{+}$	0.0073 ± 0.0008	0.0074 ± 0.0008	-	0.0067 ± 0.0004	0.0067 ± 0.0004
$c\tau_{-}$	0.0161 ± 0.0016	0.0162 ± 0.0017	-	0.0181 ± 0.0010	0.0184 ± 0.0012
S_{ct}	-	1.302 ± 0.013	-	1.294 ± 0.007	1.294 ± 0.007
$z_{1,\phi}$	-	0.325 ± 0.016	-	0.346 ± 0.009	0.346 ± 0.009
$z_{2,\phi}$	-	0.136 ± 0.030	-	0.134 ± 0.017	0.133 ± 0.017
$z_{1,\theta}$	-	0.569 ± 0.085	-	0.648 ± 0.047	0.649 ± 0.047
$z_{2,\theta}$	-	0.134 ± 0.097	-	0.228 ± 0.054	0.228 ± 0.054
$z_{1,\psi}$	-	0.893 ± 0.062	-	0.792 ± 0.036	0.800 ± 0.036
$z_{2,\psi}$	-	0.512 ± 0.107	-	0.527 ± 0.061	0.535 ± 0.061
$z_{3,\psi}$	-	-0.513 ± 0.225	-	-0.174 ± 0.131	-0.176 ± 0.132
$z_{4,\psi}$	-	-0.820 ± 0.117	-	-0.781 ± 0.067	-0.791 ± 0.067
$z_{5,\psi}$	-	-0.870 ± 0.194	-	-1.168 ± 0.113	-1.171 ± 0.114
$z_{0,\phi}$	-	0.266 ± 0.043	-	0.244 ± 0.022	0.241 ± 0.022

Table 5.3: Parameter values determined in a sequence of fits for B_d .

	$ A_0 $	$ A_{\parallel} $	δ_{\parallel}	δ_{\perp}	$c\tau$	δ_S	λ
$ A_0 $	1.000						
$ A_{\parallel} $	0.033	1.000					
δ_{\parallel}	0.048	0.188	1.000				
δ_{\perp}	-0.022	0.067	0.669	1.000			
$c\tau$	0.020	-0.001	0.004	0.001	1.000		
δ_S	-0.147	-0.325	-0.100	0.103	0.000	1.000	
λ	0.105	0.302	0.148	-0.076	0.006	-0.915	1.000

Table 5.4: Correlation matrix of the signal parameter of the final fit for B_d .

Chapter 6

Measurement of $\Delta\Gamma$ Assuming Validity of the Standard Model

In the Standard Model, the angle $2\beta_s = \phi_s$ of the unitary triangle is very small. This allows for the approximation to neglect the CP-violating phase ϕ_s in the fit. In this chapter we present the measurement of the B_s lifetime difference and mean lifetime using the simplified signal probability density function as described in section 4.3.1.

The structure of this chapter is very similar to the previous one. For details on definitions, see the appropriate section there. As in B_d case to validate the fit we present pull distributions, projections and the likelihood profiles for the signal parameters.

6.1 Fit Validation

1000 pseudo experiments are generated with the same number of events as in data. The input values of the parameters are chosen to be close to the values obtained on data. They are listed in Table 6.1.

Most pull distribution look very reasonable (see Figures 6.1, 6.2, 6.3 and 6.4). Table 6.2 summarizes the results of the Gaussian fit to the pull distributions. Only the mean of $c\tau$ is shifted by approximately 10% of the statistical uncertainty with more than three sigma significance. This bias seems to come from limited statistics, since the shift gets smaller when using a set of pseudo experiments with larger statistics.

One of the parameters used to generate the pseudo experiments is actually rather different from what we observe on data. When changing $\delta_{||}$ to a value closer to what we obtain on data, $\delta_{||} = -2.7$, the pull distribution does not look acceptable anymore (see Figure 6.5). The distribution itself (lower left

parameter	value	parameter	value
f_s	0.33	f_{++}	0.023
M	5.366 GeV	f_+	0.062
σ	8.14 MeV	f_-	0.022
f_m	0.3	$c\tau_{++}$	0.052
s_m	2.0	$c\tau_+$	0.038
$c\tau$	440 μm	$c\tau_-$	0.013
$\Delta\Gamma$	0.13 ps^{-1}	S_{ct}	1.34
A_0	0.76	$z_{1,\phi}$	0.30
$\delta_{ }$	-2.0	$z_{2,\phi}$	0.26
$A_{ }$	0.45	$z_{1,\theta}$	0.5
δ_{\perp}	0	$z_{2,\theta}$	0.7
a	-0.53	$z_{1,\psi} = z_{2,\psi} = z_4 = z_5 = z_6$	0
a_{NP}	-4.13		

Table 6.1: Parameter values used to create pull distributions for B_d .

plot in Figure 6.5) shows a second peak structure at π and the peak around the produced value is actually shifted to lower values. The fit is not able to reproduce the signal parameters $\delta_{||}$ properly anymore. The influence of the discrepancy of $\delta_{||}$ on the other variables seems not to be significant, since the pull distribution are good for all other parameters (see Table 6.3). To understand the source of the issue of the $\delta_{||}$ pull distribution better we will take a look at the likelihood profiles and discuss the details in the following section.

6.2 Likelihood Profiles

Making one dimensional scans of the likelihood function around the global minimum revealed an interesting double structure for $\delta_{||}$ (see Figure 6.6). The $\delta_{||}$ likelihood profile shows a symmetry around π , which can be understood when looking at the probability density function (see equation 4.22). Since the strong phase is only in front of the cos-term of the f_5 function, the probability density function is symmetric around π . The symmetry point distorts the likelihood function and the fit does not determine the proper uncertainties with un-biased central values. In the previous section we have shown that the result of the fit is correct for the value of $\delta_{||} = 2.0$, thus indicating the proper working of the fit. Therefore we will not present a point estimate for $\delta_{||}$, but only its likelihood profile as a result.

Par.	Mean	signi.	Width	Fit prob.	Shift
f_s	0.0546 ± 0.032	1.71	0.999 ± 0.0226	0.876	-0.00196
M	0.0269 ± 0.0319	0.843	0.995 ± 0.0225	0.811	9.03e-06
$c\tau$	-0.103 ± 0.0328	3.12	1.02 ± 0.0232	0.589	-9.35e-05
$\Delta\Gamma$	0.0619 ± 0.0328	1.89	1.02 ± 0.0232	0.181	-0.000842
A_0	-0.0337 ± 0.0328	1.03	1.02 ± 0.0232	0.274	-0.000974
$\delta_{ }$	0.0306 ± 0.0302	1.01	0.943 ± 0.0214	0.122	-0.00816
$A_{ }$	-0.00992 ± 0.0321	0.309	1 ± 0.0227	0.173	-0.00189

Table 6.2: Summary of the Gaussians fitted to the pull distributions for the individual parameters. The significance of the deviation of the mean is listed in the third column. The fit probability of the Gaussians are listed in the fifth and the shift of the mean of the parameter distribution to the value used for production is in the sixth column. The pseudo experiments are produced using the numbers from Table 6.1.

6.3 Fit Projections

The projections of the fit using only the sidebands are shown in Figure 6.7. The background can be well parameterized using the empirical approach. Since the $\cos\psi$ distribution is simpler for the B_s decay than for B_d , only the first two parameters of the ψ distribution, $z_{1,\psi}$ and $z_{2,\psi}$, are used in the fit. The fit results for the background parameters are listed in Table 6.4.

The projections of the final fit for signal and background are shown in Figure 6.8. The signal and background parts are drawn independently. Also the CP even and CP odd components are illustrated. The agreement between data and fit is very good and with the current statistics, there is no indication of any missing component in the fit.

6.4 Results

Using the neural network we obtain a signal yield of 2506 ± 51 for B_s . The background description included some parameters, z_4, z_5 and z_6 , allowing for a similar correlation between the angles as in the signal. All these parameters turned out to be very small, indicating no such correlation in the background. Table 6.4 shows the parameters determined in each fit step for B_s . Due to the issues with the likelihood scan and pull distribution described in the sections before, there is no point estimate for $\delta_{||}$, but we present the likelihood scan instead. To estimate the goodness of the fit a χ^2 is determined for the

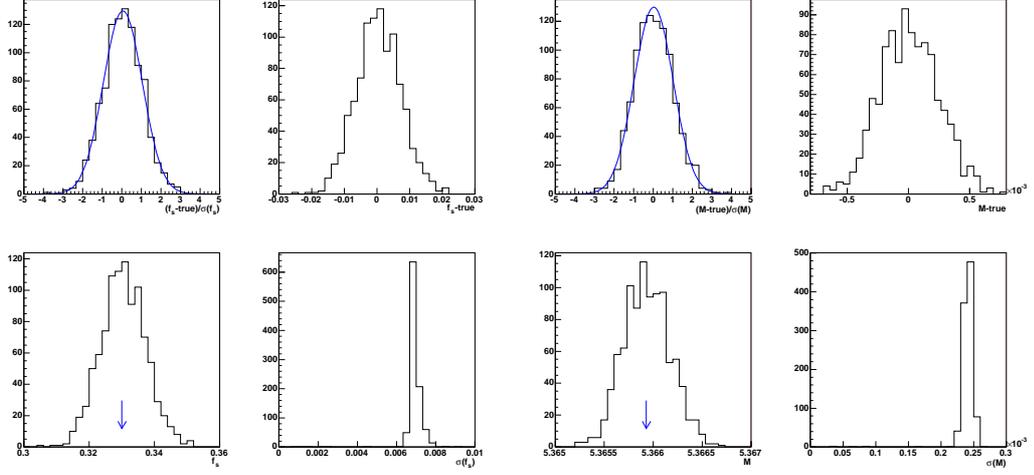


Figure 6.1: Distribution of pull (upper left), parameter (lower left), error (lower right) and parameter shifted with respect to the production value (upper right) for the signal fraction and mass for B_s . The production value is denoted by the arrow.

three-dimensional angular distribution using $20 \times 20 \times 20$ bins. Bins with low statistics are merged with neighboring bins until at least ten events are present in data and predicted by the fitted model. We obtain the probabilities of 34% ($\chi^2/\text{ndf} = 668.0/654$) for B_s with fixed $\phi_s = 0$ for the angular function, 70% ($\chi^2/\text{ndf} = 52/58$) for the decay time distribution and 96% ($\chi^2/\text{ndf} = 51/71$) for the mass distribution. We obtain the following values for the signal parameters

$$\begin{aligned}
 c\tau_s &= 456 \pm 13 \mu\text{m} \\
 \Delta\Gamma &= 0.076_{-0.063}^{+0.059} \text{ ps}^{-1} \\
 |A_0|^2 &= 0.530 \pm 0.021 \\
 |A_{||}|^2 &= 0.230 \pm 0.027
 \end{aligned}$$

The correlation matrix of the signal parameters is shown in Table 6.5.

Constrained Fit

An additional fit was done using a theoretical constraint. As described in section 1.2, using HQE we can estimate the lifetimes of B hadrons. One of the results of Heavy Quark Expansion theory is that the B_s and B_d lifetime is predicted to agree within 1%. Since the B_d lifetime is measured very precisely at B factories, we take it as an additional constraint for the B_s lifetime in

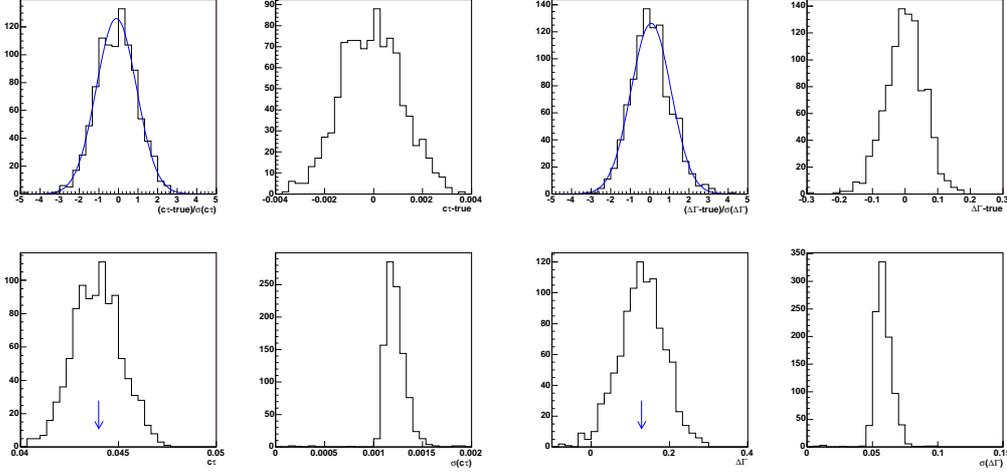


Figure 6.2: Distribution of pull (upper left), parameter (lower left), error (lower right) and parameter shifted with respect to the production value (upper right) of $c\tau$ and $\Delta\Gamma$ for B_s . The production value is denoted by the arrow.

the likelihood fit. Adding up the experimental and theoretical uncertainties we constrain the mean lifetime to $458.7 \pm 5.3 \mu\text{m}$.

The Gaussian constraint is included in the fit by adding the value of a normalized Gaussian for the parameter of interest to the existing log likelihood function, $\ln \mathcal{L}$,

$$\ln \mathcal{L}' = \ln \mathcal{L} + \ln G(c\tau | c\tau^{\text{constr}}, \sigma_{c\tau}^{\text{constr}}).$$

The Gaussian G is centered around $c\tau^{\text{constr}}$ and its width is given by $\sigma_{c\tau}^{\text{constr}}$. Using the B_d mean lifetime constraint the fit result is then

$$\begin{aligned} c\tau_s &= 458 \pm 5 \mu\text{m} \\ \Delta\Gamma &= 0.081 \pm 0.050 \text{ ps}^{-1} \\ |A_0|^2 &= 0.531 \pm 0.020 \\ |A_{||}|^2 &= 0.230 \pm 0.026 \end{aligned}$$

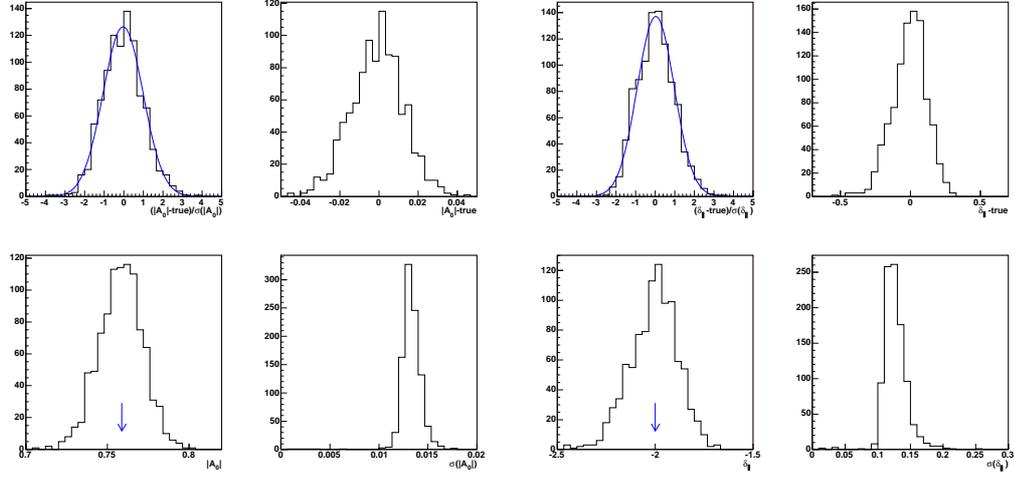


Figure 6.3: Distribution of pull (upper left), parameter (lower left), error (lower right) and parameter shifted with respect to the production value (upper right) of $|A_0|$ and δ_{\parallel} for B_s . The production value is denoted by the arrow.

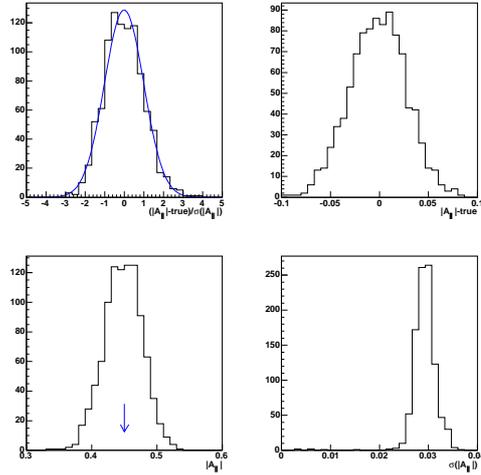


Figure 6.4: Distribution of pull (upper left), parameter (lower left), error (lower right) and parameter shifted with respect to the production value (upper right) of $|A_{\parallel}|$ for B_s . The production value is denoted by the arrow.

Par.	Mean	signi.	Width	Fit prob.	Shift
f_s	0.0142 ± 0.035	0.406	1.01 ± 0.0247	0.274	$9.78e-05$
M	0.0489 ± 0.0358	1.36	1.03 ± 0.0253	0.584	$7.63e-06$
$c\tau$	-0.101 ± 0.0346	2.92	0.999 ± 0.0245	0.724	-0.000127
$\Delta\Gamma$	-0.0234 ± 0.0337	0.695	0.972 ± 0.0238	0.207	-0.00506
A_0	-0.0346 ± 0.0341	1.01	0.984 ± 0.0241	0.28	-0.00108
$\delta_{ }$	-0.233 ± 0.0346	6.75	0.996 ± 0.0244	0	0.0312
$A_{ }$	0.0102 ± 0.0334	0.305	0.965 ± 0.0236	0.375	-0.0006

Table 6.3: Summary of the Gaussians fitted to the pull distributions for the individual parameters. The significance of the deviation of the mean is listed in the third column. The fit probability of the Gaussians are listed in the fifth and the shift of the mean of the parameter distribution to the value used for production is in the sixth column. The pseudo experiments are produced using $\delta_{||} = 2.7$ instead $\delta_{||} = 2.0$ from Table 6.1. The production value depending bias of $\delta_{||}$ illustrates the symmetry of the probability density function.

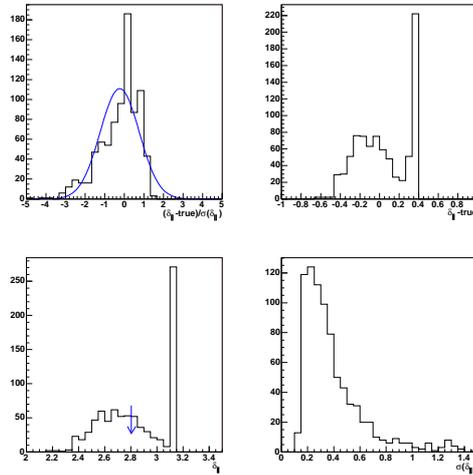


Figure 6.5: Distribution of pull (upper left), parameter (lower left), error (lower right) and parameter shifted with respect to the production value (upper right) distributions of $\delta_{||}$ for B_s . The production value is denoted by the arrow.

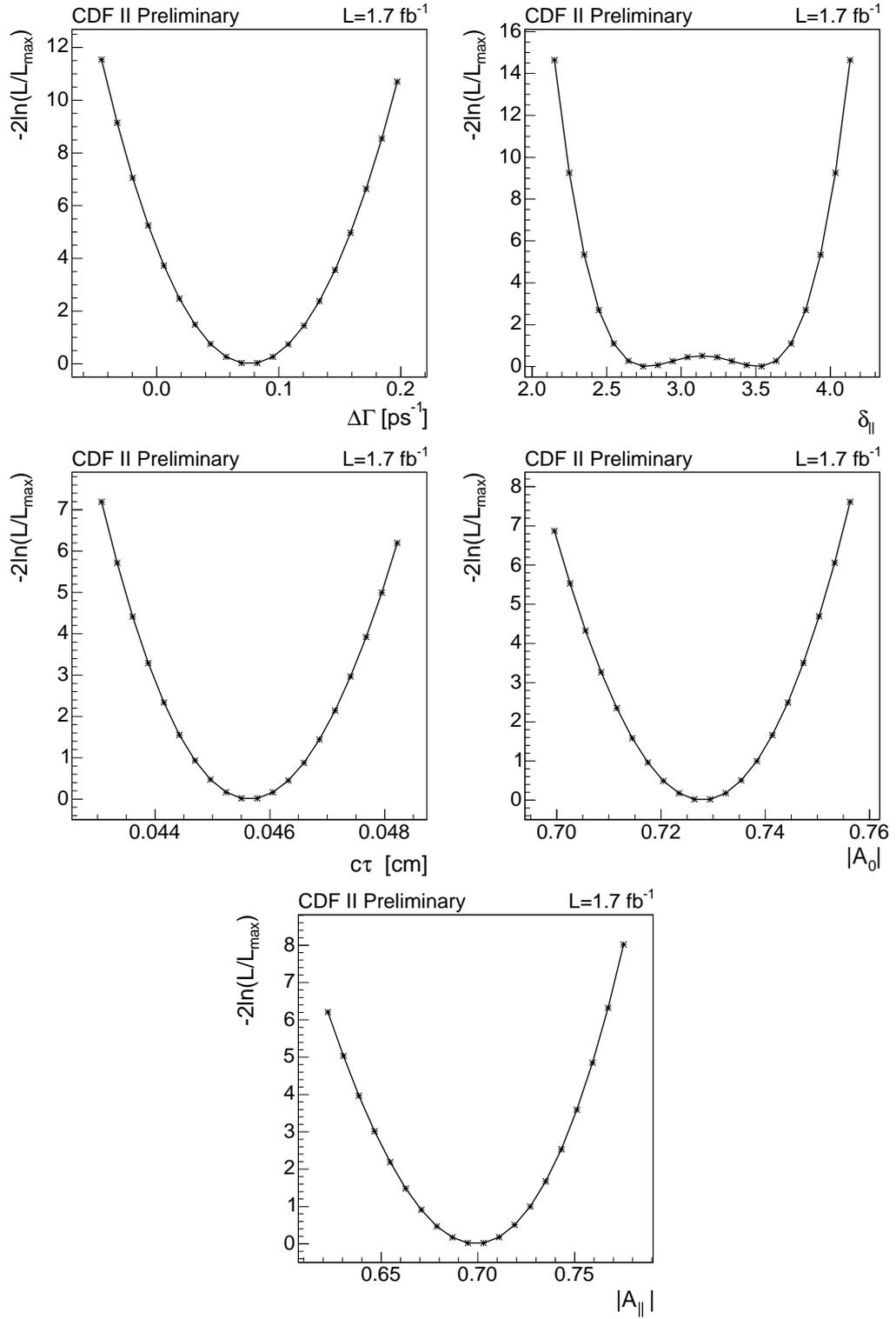


Figure 6.6: B_s likelihood profiles for the signal parameters.

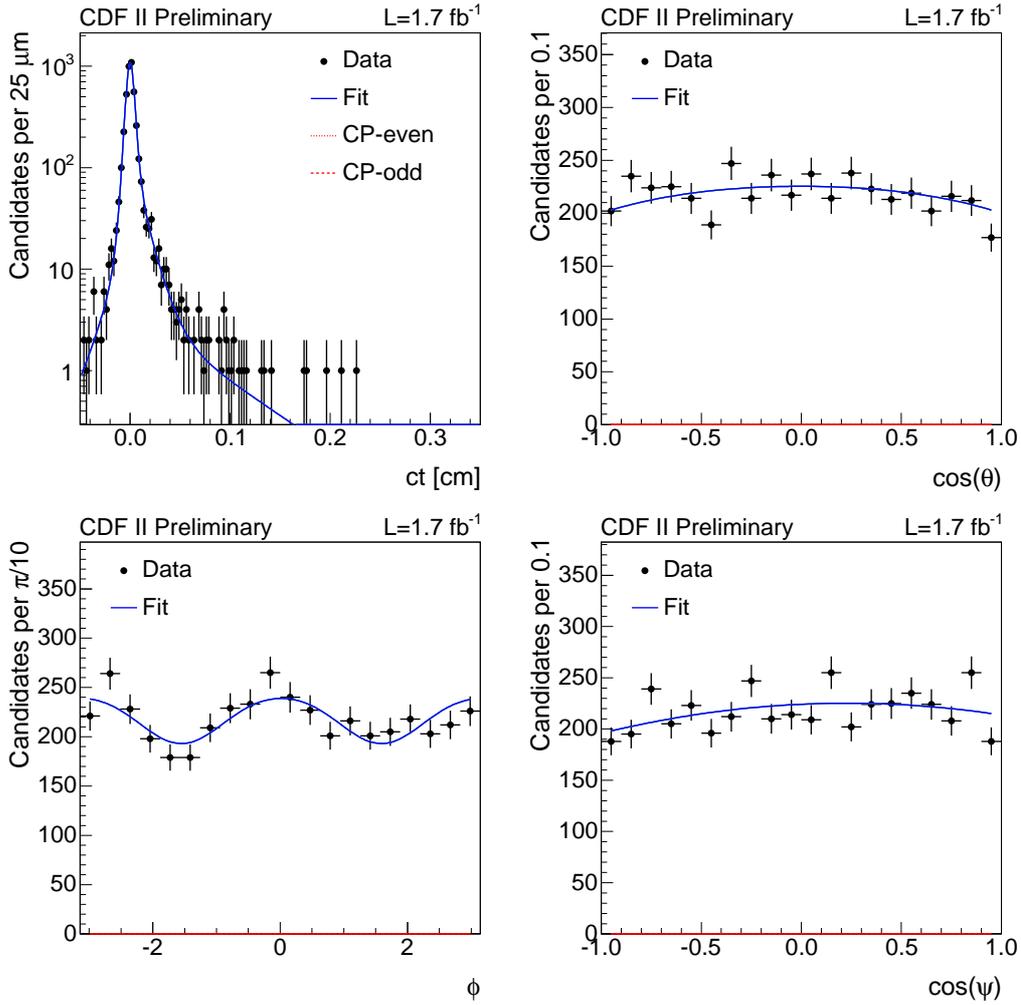


Figure 6.7: B_s lifetime and angular projections of the upper and lower sidebands only.

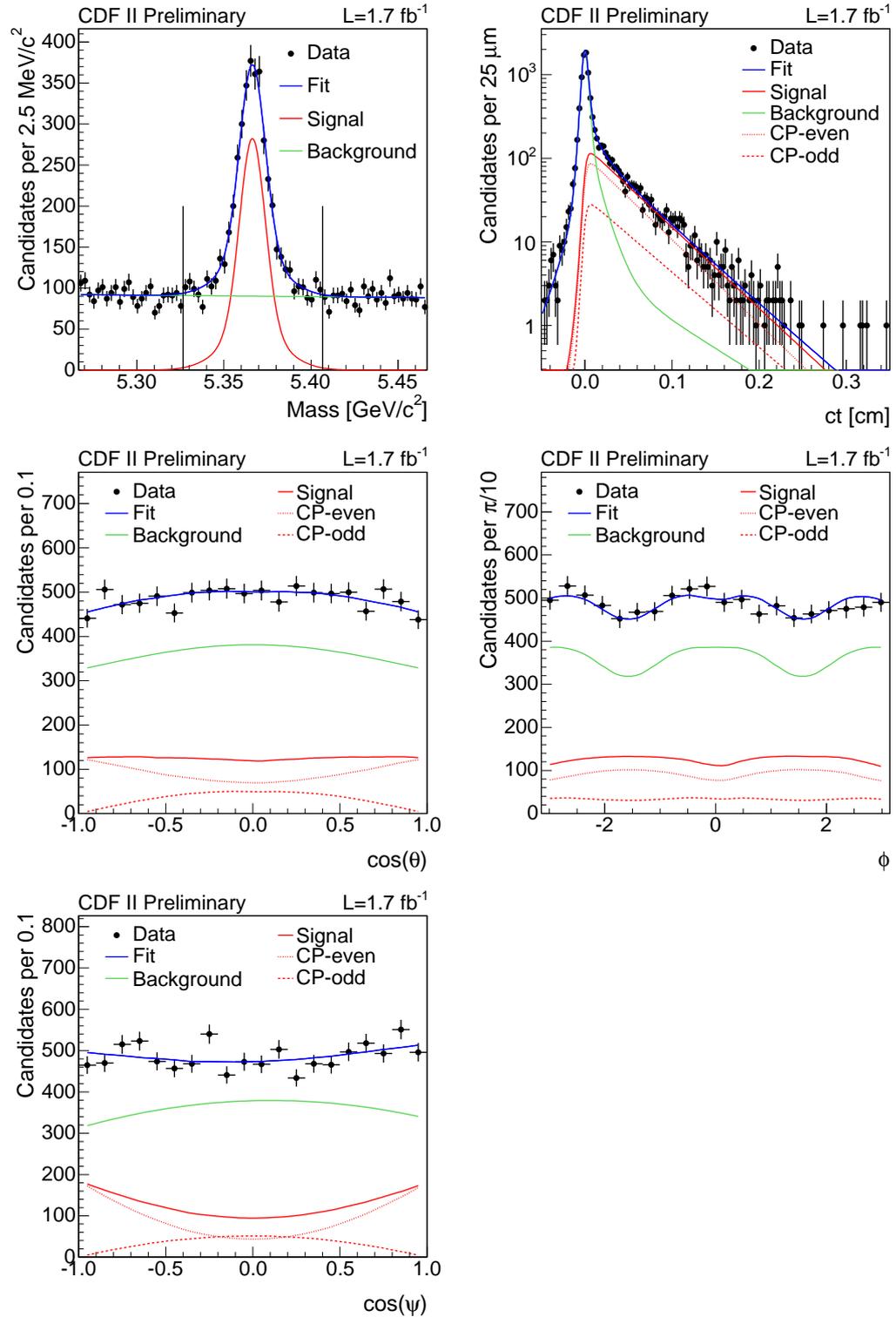


Figure 6.8: Projections of the results of the unbinned maximum likelihood fit of the B_s decay under the assumption of no CP violation, $\phi_s = 0$.

parameter	lf bkg. fit	bkg. fit	mass fit	fit with $\phi_s = 0$	constr. fit
f_s	-	-	0.2488 ± 0.0091	0.2581 ± 0.0053	0.2580 ± 0.0053
M	-	-	5.3663 ± 0.0003	5.3664 ± 0.0002	5.3664 ± 0.0002
σ	-	-	0.0067 ± 0.0012	0.0072 ± 0.0004	0.0072 ± 0.0004
f_m	-	-	0.42 ± 0.34	0.32 ± 0.07	0.32 ± 0.07
s_m	-	-	1.88 ± 0.40	$2.32^{+0.20}_{-0.16}$	2.32 ± 0.17
$c\tau$ [μm]	-	-	-	456.4 ± 13.0	458.4 ± 4.9
$\Delta\Gamma$ [ps^{-1}]	-	-	-	$0.076^{+0.059}_{-0.063}$	0.081 ± 0.050
$ A_0 $	-	-	-	$0.728^{+0.014}_{-0.015}$	0.729 ± 0.013
$\delta_{ }$	-	-	-	$2.77^{+0.98}_{-0.23}$	2.77 ± 0.31
$ A'_{ } $	-	-	-	$0.699^{+0.037}_{-0.039}$	0.701 ± 0.036
δ_{\perp}	-	-	-	-	-
ϕ_s	-	-	-	-	-
a	-	-0.97 ± 1.20	-0.86 ± 1.18	-0.86 ± 1.15	-0.86 ± 1.15
a_{NP}	-	-1.94 ± 3.89	-	$-2.10^{+3.93}_{-3.94}$	-2.10 ± 3.94
f_{++}	0.0978 ± 0.0092	0.0977 ± 0.0092	-	$0.0916^{+0.0080}_{-0.0079}$	0.0912 ± 0.0079
f_+	0.0218 ± 0.0061	0.0218 ± 0.0062	-	$0.0187^{+0.0063}_{-0.0054}$	0.0187 ± 0.0058
f_-	0.0180 ± 0.0035	0.0179 ± 0.0035	-	$0.0191^{+0.0033}_{-0.0029}$	0.0192 ± 0.0031
$c\tau_{++}$	0.0114 ± 0.0017	0.0114 ± 0.0017	-	$0.0120^{+0.0017}_{-0.0015}$	0.0120 ± 0.0016
$c\tau_+$	0.065 ± 0.012	0.0644 ± 0.0119	-	$0.0661^{+0.0146}_{-0.0105}$	0.0660 ± 0.0120
$c\tau_-$	0.0197 ± 0.0031	0.0197 ± 0.0032	-	$0.0176^{+0.0020}_{-0.0022}$	0.0176 ± 0.0022
S_{ct}	-	1.368 ± 0.022	-	$1.390^{+0.018}_{-0.017}$	1.390 ± 0.018
$z_{1,\phi}$	-	0.106 ± 0.022	-	0.093 ± 0.017	0.093 ± 0.017
$z_{2,\phi}$	-	-0.017 ± 0.047	-	$-0.051^{+0.033}_{-0.032}$	-0.051 ± 0.033
$z_{1,\theta}$	-	0.089 ± 0.169	-	$0.188^{+0.128}_{-0.133}$	0.188 ± 0.130
$z_{2,\theta}$	-	-0.023 ± 0.191	-	$0.039^{+0.146}_{-0.150}$	0.039 ± 0.147
$z_{1,\psi}$	-	0.040 ± 0.025	-	0.031 ± 0.020	0.031 ± 0.020
$z_{2,\psi}$	-	-0.087 ± 0.048	-	$-0.144^{+0.037}_{-0.036}$	-0.144 ± 0.037
$z_{0,\phi}$	-	-0.037 ± 0.218	-	$0.008^{+0.156}_{-0.159}$	0.008 ± 0.156
z_4	-	-0.0074 ± 0.0177	-	-0.0001 ± 0.0139	-0.0001 ± 0.0139
z_5	-	0.0005 ± 0.0254	-	$-0.0393^{+0.0205}_{-0.0204}$	-0.0393 ± 0.0205
z_6	-	0.0089 ± 0.0253	-	0.0185 ± 0.0201	0.0184 ± 0.0201

Table 6.4: Parameter values determined in a sequence of fits for B_s .

	$ A_0 $	$ A_{ } $	$\delta_{ }$	$\Delta\Gamma$	$c\tau$
$ A_0 $	1.000				
$ A_{ } $	0.229	1.000			
$\delta_{ }$	0.061	-0.430	1.000		
$\Delta\Gamma$	0.656	0.464	0.027	1.000	
$c\tau$	0.422	0.372	-0.037	0.601	1.000

Table 6.5: Correlation matrix of the signal parameters determined in the final unconstrained fit for B_s .

Chapter 7

Systematic Uncertainties

Every effect of the data that is not incorporated in the fit model might lead to systematic effects in the parameters estimated from the unbinned maximum likelihood fit. Using pseudo experiments we estimate the influence of each effect that might be or is in the data, individually on the signal parameters, by incorporating the effect in the simulation. In this chapter we present the different effects we studied for the B_d and B_s decay under the assumption of no CP violation ($\phi_s = 0$) in interference between decay with and without mixing. The analysis strategy for ϕ_s floating in the fit is very different and will be presented in the next chapter.

The parameter distributions of the pseudo experiments are used to estimate the systematic uncertainties. The parameter distributions obtained from the toy studies presented in the previous chapters are used as reference. When adding a systematic effect, pseudo experiments are produced to get a set of parameter and pull distribution for each parameter derived in the fit. The size of the systematic uncertainty is obtained by subtracting the mean of the references from the mean of the distributions derived when including the systematic effects. The numerical values of the uncertainties are listed in Table 7.1 and 7.2 for B_d and B_s , respectively.

7.1 Signal Mass Model

The signal mass distribution was modelled using a double Gaussian with a common mean. Although the model seems to describe the data very well, it might be that we missed an additional component. Due to the good agreement using the two Gaussian, we assume that an unknown third component of the signal description has less influence than the second Gaussian. Therefore, by producing pseudo experiments with a double Gaussian structure, but

fitting this allowing only a single Gaussian, the systematic effect of any third component should be smaller than what we see in this test. This was done for the B_d and B_s decays independently. The mass model seems to introduce only small uncertainties on the fit results.

7.2 Angular Model of the Background

The angular model of the background was tested by using two different fitters and quoting the differences as systematic effect. Two fits were done, one with the complete empirical model for the background, the other with a reduced model where some of the parameters were fixed to zero. Some of the polynomials were reduced by some order and the terms allowing for the correlations were disabled. This allows us to make a conservative estimation of any forgotten parameterization, or it allows us to estimate if we used too many parameters to model the background. In general we observe a good agreement between the background fit and the data, but we cannot know, whether we used too many parameters or used a model that is too simple.

7.3 Decay Time Resolution Model

As explained in section 4.2.2, the decay time distribution is not just an exponential, but it needs to be convoluted with the resolution function. A single Gaussian is used to model the decay time resolution function. But the negative tail in the decay time distribution of the background indicates that this might not be enough. The decay time distributions for the pseudo experiments are produced by simulating events following an exponential and convoluted with a Gaussian. To estimate the influence of an additional positive and negative longlived term we simulate events where we add exponential tails to the resolution function used to generate the pseudo experiments. The fraction and amount of exponential tails of signal events is taken from the fraction and lifetime of the negative lifetime component of the background. As we can see from Table 7.2 the decay time resolution model is the major contribution to the uncertainty of the B_s mean lifetime measurement.

7.4 Cross-Feed

The two decay channels, $B_d \rightarrow J/\psi K^*$ and $B_s \rightarrow J/\psi \phi$, analysed in this thesis are very similar in their final decay products. They only differ in a pion/kaon, which is combined with another kaon and two muons. This is

especially delicate, since the pion/kaon separation is not perfect. Thus it is very easy to misreconstruct one decay mode for the other. $B_d \rightarrow J/\psi K^*$ events can be reconstructed as B_s candidates by misidentifying the pion from the K^* decay as a kaon. Therefore the K^* is reconstructed as a ϕ leading to a fake B_s candidate. These candidates have by construction a mass that is close to the true B_s mass.

Taking simulated B_d events but reconstructing them as B_s candidates we can estimate the relative fraction of cross-feed that passes reconstruction. Starting from the same number of simulated B events for B_d and B_s , we obtain a relative reconstruction efficiency of $N_{B_d}/N_{B_s} \approx 0.35\%$. Thus starting from the same amount of B_d and B_s candidates a rather small fraction of the B_d candidates survive. To estimate the number of B_d events in the B_s sample we also have to take into account the relative production rate and branching ratios. At the Tevatron the relative production rate, f_d/f_s , for d and s quarks is roughly $39.6\%/10.4\% \approx 3.8$. Thus we will have about three to four times more B_d than B_s mesons. Further the different branching ratios will change the number of events for the specific decays we are interested in. The acceptance is already taken into account by the detector simulation.

Taking the numerical values from reference [17] we estimate the fraction of B_d in the B_s sample to be

$$\begin{aligned} \frac{N_{B_d}}{N_{B_s}} &\times \frac{f_d Br(B_d \rightarrow J/\psi K^*) \times Br(K^* \rightarrow K^\pm \pi^\mp)}{f_s Br(B_s \rightarrow J/\psi \phi) \times Br(\phi \rightarrow K^+ K^-)} = \\ 0.35\% &\times 3.8 \cdot \frac{1.3 \cdot 10^{-3} \cdot 0.67}{9.3 \cdot 10^{-4} \cdot 0.49} \approx 2.5\% \end{aligned} \quad (7.1)$$

We incorporated this in the systematic studies by exchanging three percent of the signal B_s candidates by B_d events for each pseudo experiment. For the mean lifetime we use the B_d world average of $458.7 \mu\text{m}$ [17]. The mass shape is taken from the simulated signal only sample. For the angular distributions we use the parameters determined in the fit of the $B_d \rightarrow J/\psi K^*$ sample.

The reverse misreconstruction is also possible. By mis-assigning the pion hypothesis to a kaon the ϕ of the B_s can be reconstructed as K^* leading to a B_d candidate. Taken into account the production fraction and branching ratios, the contamination in the B_d sample is one order of magnitude smaller and therefore neglected in this analysis.

7.5 B_d Candidates with Swapped Mass Assignment

The swapped candidates were already discussed in more detail in section 3.5.3. We estimate on simulated event, that we can reduce the fraction of swapped events to about half a percent using the swap suppression network. We can study their effect on the fit parameters by replacing some of the B_d events with swapped reconstructed decays, where the pion and kaon hypothesis is exchanged, in the pseudo experiments. Since the estimation of the fraction of those swapped reconstructed events is done on simulated events, we want to stay conservative and therefore exchange twice as much of the signal events with swapped candidates. From the simulated events we know that the mass distribution is wider than for properly reconstructed candidates. We simulate it to be three times as wide as the narrow signal component. The decay time is the same as for un-swapped B_d candidates and the angular distribution is taken from the simulated B_d events, which were reconstructed with swapped $K\pi$ assignment.

7.6 SVX Alignment

The position of the individual wavers of the silicon detector is known only within a limited precision. To study the influence on the fit, typically simulated events are produced, where the silicon detector is slightly changed in the detector simulation. In the previous CDF analysis of the decays presented here, the uncertainty on the the lifetime was estimated to be $\pm 1\mu\text{m}$ [9]. More recent studies at CDF [62] also estimated the uncertainty of the misalignment of the SVX detector. There it was concluded that the systematic uncertainty on the lifetime is about $2\mu\text{m}$. We assume to have similar uncertainties, as the other measurements, but we are not only interested in the lifetime we also have to estimate the influence on other parameters, as well. Therefore we use two sets of pseudo experiments which are generated with a lifetime changed by $\pm 2\mu\text{m}$ from the measured value. As the final systematic uncertainty on the individual parameters, we take the largest effect found for the positive and negative shift. For B_d it is the largest source of uncertainty for the measured lifetime.

7.7 S-Wave Contribution

In section 5.1 we introduced a B_d specific component in the fit, the S-Wave contribution. It describes the decay of the $B_d \rightarrow J/\psi K^*$ without the resonant K^* state. We obtained a S-Wave fraction of $1.6 \pm 0.8\%$, which does agree with zero within 2σ . Although this component is included in the fit, we take the difference between the fit with and without this component as additional systematic uncertainty. This effect is not considered for B_s decays since the ϕ mass width is much narrower and its mean is closer to the threshold than for the K^* . Thus the amount of S-Wave contribution for B_s is much smaller than for B_d .

7.8 Decay Time Uncertainty Distribution

As has been shown in reference [53] the decay time uncertainty distributions have to be taken into account in a lifetime fit function if the distributions are different for signal and background. They have to be included in the likelihood function in case we want to measure a lifetime using the event-by-event decay time measurement. The fit can be done without the decay time uncertainty, but the result might be biased. For B_s we observed only a slight difference in the two fits with and without the distributions and therefore neglect the difference.

The change in the parameters for B_d was much larger when comparing those two fits. Therefore we decided to take the difference into account and add it as a systematic uncertainty. Since the fit without decay time uncertainty distributions is overestimating the effect, we used half of the observed difference as systematic uncertainty.

7.9 Summary

In this chapter we have described different effects that might affect the fit result in one direction or the other. The systematic uncertainties for the signal parameters are summarized in tables 7.1 and 7.2. The total uncertainty is obtained by adding individual contributions in quadrature. The distribution, the values are derived from, can be found in reference [63].

Sys. Eff. \ Parameter:	$c\tau[\mu m]$	$ A_0 $	$ A_{ } $	$\delta_{ }$	δ_{\perp}
Signal mass model	2	0.001	-	0.001	0.003
ct resolution model	3	-	0.002	0.005	0.002
K/ π swap	2	0.001	0.001	0.001	0.003
SVX alignment	4	-	-	0.001	0.001
Angular bg. model	2	0.005	0.004	-	-
S-Wave	1	0.006	0.004	0.031	0.002
$\sigma(ct)$ templates	1	0.005	0.002	0.006	0.003
Total	6	0.009	0.006	0.032	0.006

Table 7.1: Summary of systematic uncertainties assigned to the signal parameters of the B_d decay.

Sys. Eff. \ Parameter:	$c\tau[\mu m]$	$\Delta\Gamma$	$ A_0 $	$ A_{ } $	$\delta_{ }$
Signal mass model	-	0.001	0.001	0.002	0.004
ct resolution model	7	0.003	0.001	0.001	0.003
B_d cross feed	-	0.004	0.003	0.003	0.069
Binning	-	-	0.002	-	-
SVX alignment	1	0.002	-	-	0.001
Angular bg. model	2	0.004	0.006	0.002	0.072
Total	7	0.006	0.008	0.009	0.100

Table 7.2: Summary of systematic uncertainties assigned to the signal parameters of the B_s decay assuming $\phi_S = 0$.

Chapter 8

Measurement of ϕ_s

In the fit with floating ϕ_s we observe a systematic bias explained in the next section. As the bias is large and non-linear, we decided not to correct the central value and the uncertainty for this bias but to quote a p -value for compatibility with the Standard Model and to create a confidence region using likelihood ratios. In this way the bias is automatically taken into account by the pseudo experiments. These methods and the result are introduced in this chapter.

8.1 Systematic Bias

A maximum likelihood estimator is only guaranteed to be without any bias in the fit parameters for unlimited statistics (see also section 4.1). The amount of statistics needed for unbiased results depends on the probability density function and the value of the parameters. When fitting for $\Delta\Gamma$ under the assumption of $\phi_s = 0$ we had the first experience with a problematic variable, the strong phase $\delta_{||}$. There the value of the parameter is close to a symmetric point which together with the available statistics leads to a situation where two solutions are not well separated. For the fit of B_s with free ϕ_s we observe systematic biases when using pseudo experiments. In Figure 8.1 the fit results for 750 pseudo experiments are shown. The experiments were produced using $\phi_s = 0$ and $\Delta\Gamma = 0.1$ as input, which is close to the standard model expectation.

Naively we would expect that the parameter distribution from the pseudo experiments is centered around the production value and symmetric on both sides. Due to the four fold ambiguity described in section 4.3, all results are transformed to positive values of $\Delta\Gamma$ and negative ϕ_s . For $\Delta\Gamma$ the mean of the distribution tends to larger values with additional long tails leading to

an asymmetric shape. For ϕ_s the situation is a bit more complicated, since the distribution was produced at zero, but is strongly shifted away from zero. Due to the mentioned ambiguity, all values are flipped to the negative side. The shape of the distribution is a combined effect of statistical uncertainties and a bias to larger values. The bias explains the lack of fitted values at zero for ϕ_s and the surprisingly large absolute values of the parameter distribution, although the pseudo experiments were produced with $\phi_s = 0$.

Similar to the difficulties of the strong phase δ_{\parallel} for the Standard Model assumption fit, we can understand the behavior of the pseudo experiments by inspecting the probability density function. There are some difficulties when parameters tend toward some special values. As a reminder, the time dependent angular distribution, which was derived before (see equation 4.10), is

$$\begin{aligned} \frac{d^4 P(\vec{\omega}, t)}{d\vec{\omega} dt} &\propto |A_0(0)|^2 f_1(\vec{\omega})(\mathcal{T}_+) \\ &+ |A_{\parallel}(0)|^2 f_2(\vec{\omega})(\mathcal{T}_+) \\ &+ |A_{\perp}(0)|^2 f_3(\vec{\omega})(\mathcal{T}_-) \\ &+ |A_0(0)||A_{\parallel}(0)| f_5(\vec{\omega}) \cos(\delta_{\parallel})(\mathcal{T}_+) \\ &+ |A_{\parallel}(0)||A_{\perp}(0)| f_4(\vec{\omega}) [\cos(\delta_{\perp} - \delta_{\parallel}) \sin \phi_s (e^{-\Gamma_H t} - e^{-\Gamma_L t})/2] \\ &+ |A_0(0)||A_{\perp}(0)| f_6(\vec{\omega}) [\cos(\delta_{\perp}) \sin \phi_s (e^{-\Gamma_H t} - e^{-\Gamma_L t})/2], \quad (8.1) \end{aligned}$$

where

$$\mathcal{T}_{\pm} = ((1 \pm \cos \phi_s) e^{-\Gamma_L t} + (1 \mp \cos \phi_s) e^{-\Gamma_H t})/2.$$

Two parts of the function are zero for $\phi_s = 0$. This leads to an undetermined parameter δ_{\perp} and an effective loss of degree of freedoms of the probability density function. Similar if $\Delta\Gamma = 0.0 \text{ ps}^{-1}$ the lifetimes τ_H and τ_L have the same value and the difference of their exponentials gets zero. Then the two parameters ϕ_s and δ_{\perp} cannot be estimated anymore. Any statistical fluctuation in the sample will allow the likelihood estimator to find a better minimum when allowing these values from being non-zero, since there are more degrees of freedom to parameterize the data. This also means that with larger true ϕ_s the bias gets smaller because there is less room to improve the fit result by increasing the values. For $\Delta\Gamma$ the statistical uncertainty is comparable to the magnitude of the result, but for ϕ_s the uncertainty is much larger than the expected value. Due to the non-linearity of the ϕ_s bias, we are not able to determine the true value on data.

These complication in the fit function need special treatment, since the value and its uncertainty of the affected parameters will be biased. When

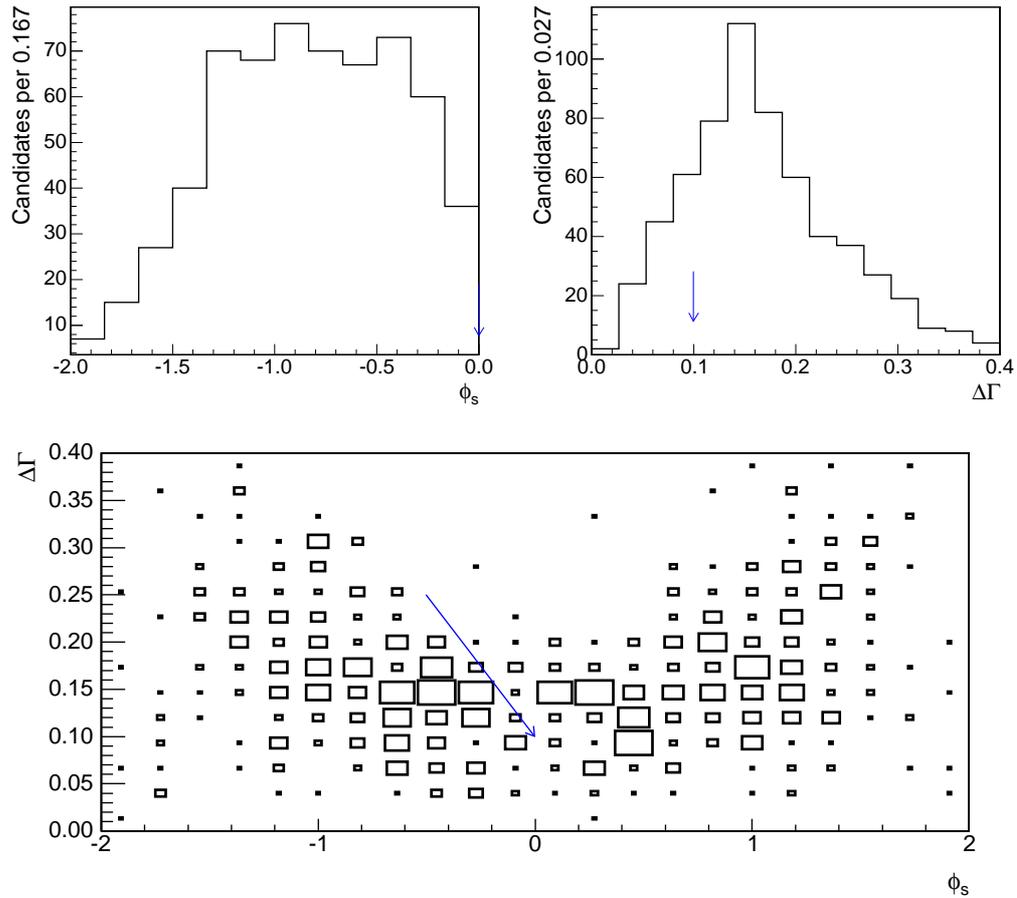


Figure 8.1: Fit result of ≈ 750 pseudo experiments for $\Delta\Gamma$ (top left) and ϕ_s (top right) with an assumption of $\phi_s = 0$ and $\Delta\Gamma = 0.1$ (blue arrows). The bottom plot, ϕ_s versus $\Delta\Gamma$, illustrates the bias away from zero for both parameters. ϕ_s and $\Delta\Gamma$ are significantly shifted in this case.

obtaining the result on data, it will be difficult to estimate the un-biased value. Therefore we check first the agreement of the Standard Model with our result on data and then extend this procedure to derive a confidence region in the $\Delta\Gamma - \phi_s$ plane corresponding to the fit result on data.

8.2 P-Value Definition

We need some way to express the agreement between the fit result on data and any specific model. We want to make a test of the specific hypothesis, that the true values are the ones given by the model. To quantify the compatibility of the fit result with any specific model we use a likelihood ratio R as suggested by Feldman and Cousins [64]. It is defined as

$$R(\Delta\Gamma, \phi_s) = \log \frac{\mathcal{L}(\hat{\Delta\Gamma}, \hat{\phi}_s, \hat{\theta})}{\mathcal{L}(\Delta\Gamma, \phi_s, \hat{\theta}')}, \quad (8.2)$$

where θ denotes all other fit parameters, except the ones tested $(\Delta\Gamma, \phi_s)$, and the hat indicates the values of parameters that minimize \mathcal{L} . Note that $\hat{\theta}$ are the parameter values that minimize \mathcal{L} for floating $\Delta\Gamma$ and ϕ_s and $\hat{\theta}'$ are the values at the minimum for $\Delta\Gamma$ and ϕ_s fixed to the values of the model of interest. The R value from data can then be compared to an R -value distribution obtained from pseudo experiments to calculate a probability for the agreement. The step-by-step procedure is the following:

- In the pseudo experiments that we generate to obtain the distribution of likelihood ratios for fixed values of $\Delta\Gamma$ and ϕ_s we have to choose values for all other parameters. Our best estimate for these parameters are the values obtained from a fit to data, where we fix $\Delta\Gamma$ and ϕ_s to the specific values we want to test. This procedure is known as plug-in method and allows us to determine all other parameters, θ , for a specific model.
- For the specific pair of $\Delta\Gamma$ and ϕ_s values that we want to investigate we generate pseudo experiments with the given $\Delta\Gamma$ and ϕ_s values and the other input parameters, θ , set to the values obtained from the fit to data in the previous step.
- We then compare the normalized distribution of R values from the pseudo experiments $f(R, \Delta\Gamma, \phi_s)$ to the value $R_{data}(\Delta\Gamma, \phi_s)$ from data to obtain the p -value, which is defined as

$$p(\Delta\Gamma, \phi_s) = \int_{R_{data}(\Delta\Gamma, \phi_s)}^{\infty} f(R, \Delta\Gamma, \phi_s) dR. \quad (8.3)$$

As there is no clear way how to incorporate the systematic uncertainties and as they are small compared to the statistical uncertainties, we only include the statistical uncertainty in the calculation of the p-values and the confidence region.

With this frequentist method we are able to determine the probability that we observe a result with $R \geq R_{data}$, if $\Delta\Gamma$ and ϕ_s are the values predicted by some model. We can extend this method to create a confidence region of the fit result on data instead of quoting results for single hypothesis.

8.3 Confidence Region

We use the same method as described in the previous section to define a confidence region that contains with 90% and 95% probability the true values of $\Delta\Gamma$ and ϕ_s . In order to construct such a region, for each pair on a $\Delta\Gamma$ - ϕ_s grid we generate pseudo experiments and determine the distribution $f(R, \Delta\Gamma, \phi_s)$ of likelihood ratios. The confidence region is then defined by the $\Delta\Gamma$ - ϕ_s pairs which satisfy the conditions

$$\begin{aligned} p(\Delta\Gamma, \phi_s) > 10\% &\Leftrightarrow (\Delta\Gamma, \phi_s) \text{ inside } 90\% \text{ C.L. region} \\ p(\Delta\Gamma, \phi_s) > 5\% &\Leftrightarrow (\Delta\Gamma, \phi_s) \text{ inside } 95\% \text{ C.L. region} \end{aligned} \quad (8.4)$$

The p-value calculation in the $\Delta\Gamma$ - ϕ_s plane is only an approximation, since we do not know whether the shape would change for a different value of the other parameter. The optimal thing to do, would be to create a confidence region on the multidimensional space of all parameters. But the vast number of test points in such a grid would be enormous and would require a lot of time and computational resources. On the other hand we assume that the confidence region will mainly depend on the two parameters of interest. To check this we modify some of the other parameters for several points to check, whether our assumption is legitimate. Further we checked that systematic effects do not strongly influence our results.

Plug-In Method Cross-Checks

To check that the plug-in method does not lead to a significant under- or over-coverage we generated pseudo experiments with varied values for $c\tau$, $\delta_{||}$ and δ_{\perp} , the parameters that show the strongest correlation to $\Delta\Gamma$ and ϕ_s . The R -distributions and the p-values obtained from this variation are shown for three different $\Delta\Gamma$ - ϕ_s pairs in Figures 8.2, 8.3 and 8.4. No strong dependence on the varied parameters is observed.

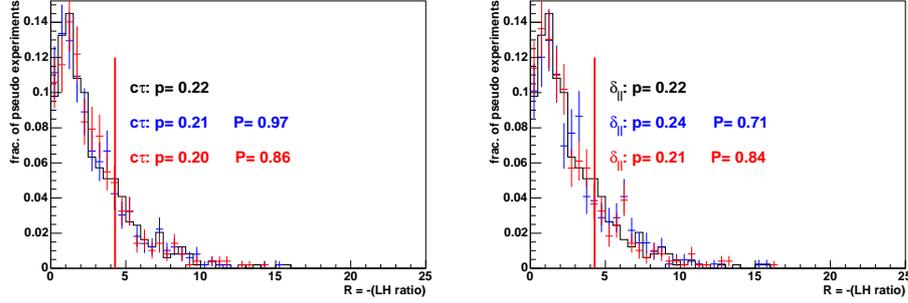


Figure 8.2: Influence of $c\tau$ and $\delta_{||}$ on the likelihood ratio and p -value for $\Delta\Gamma = 0.10 \text{ ps}^{-1}$ and $\phi_s = 0.0$ (δ_{\perp} was not considered, since for $\phi_s = 0.0$ δ_{\perp} is undetermined). The blue and red distributions are for input parameters which were shifted by $\pm 1\sigma$. p is the estimated probability and P denotes the probability to observe a value for χ^2 or larger for the agreement of the distributions for the original and shifted parameters.

A confidence region does only take into account statistical uncertainties. We checked that the whole shape is not strongly affected by systematic effects by comparing the p -value distribution of one point in the confidence region with simulated pseudo experiments including tails in the decay time uncertainty distributions and including cross feeds from B_0 as described in the section 7.3 and 7.4 respectively. In Figure 8.5 it is shown that there is no large discrepancy observed.

To keep the computational resources in a reasonable range, the fit of the toys at each point for the confidence region was done using a faster fit procedure than the one used for the final fits on data in Chapter 5 and 6. The background parameters were estimated from the sidebands and kept fixed when fitting the signal region. To confirm that this does not influence the confidence region, we compared for two point that there is no large discrepancy in the p -values (Figure 8.6).

8.4 Results

In this chapter we presented the method used to derive the agreement between our data and different models. Especially important here is the good agreement with the Standard Model of $p\text{-value(SM)} = 22\%$. We also scanned the full $\Delta\Gamma$ - ϕ_s plane to estimate the region which is compatible at 90% and 95% confidence level with the result we observe on data. The result is shown

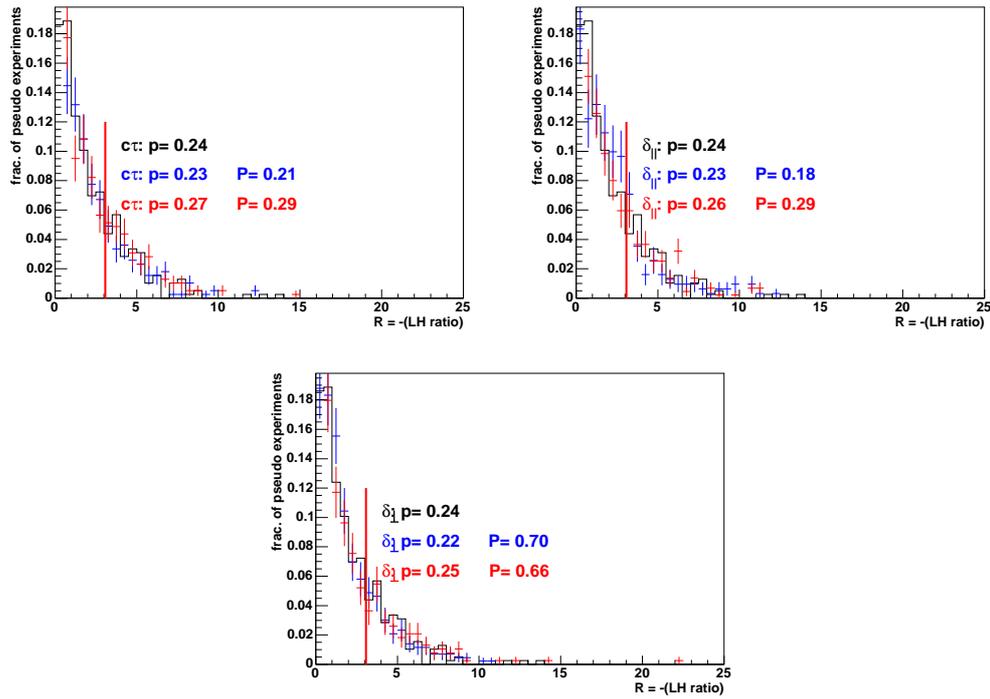


Figure 8.3: Influence of $c\tau$, $\delta_{||}$ and δ_{\perp} on the likelihood ratio and p -value for $\Delta\Gamma = 0.30 \text{ ps}^{-1}$ and $\phi_s = \pi/2$. The blue and red distributions are for input parameters which were shifted by $\pm 1\sigma$. p is the estimated probability and P denotes the probability to observe a value for χ^2 or larger for the agreement of the distributions for the original and shifted parameters.

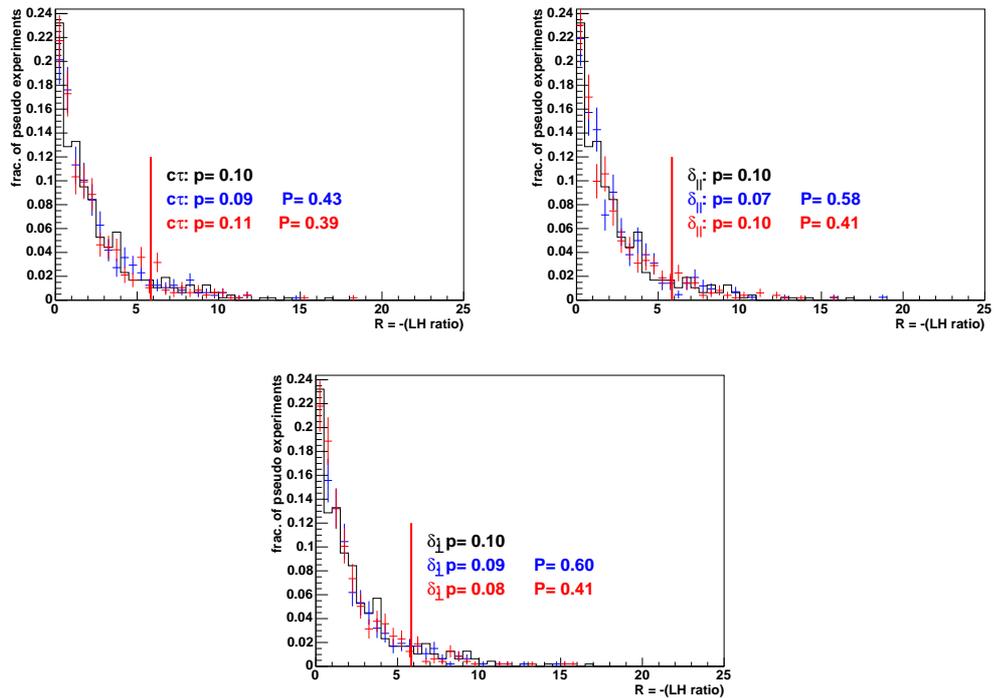


Figure 8.4: Influence of $c\tau$, $\delta_{||}$ and δ_{\perp} on the likelihood ratio and p -value for $\Delta\Gamma = 0.15 \text{ ps}^{-1}$ and $\phi_S = 3\pi/4$. The blue and red distributions are for input parameters which were shifted by $\pm 1\sigma$. p is the estimated probability and P denotes the probability to observe a value for χ^2 or larger for the agreement of the distributions for the original and shifted parameters.

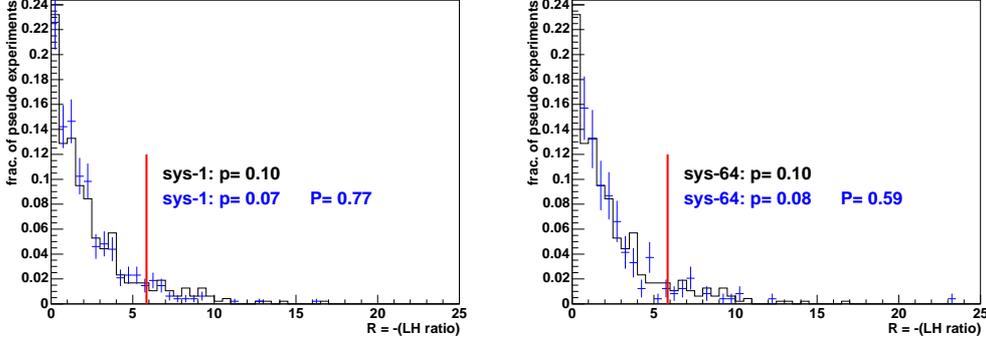


Figure 8.5: Differences of the p -value when producing pseudo experiments with and without specific systematic effects (left: resolution model, right: B_d cross feed) for: $\Delta\Gamma = 0.15 \text{ ps}^{-1}$, $\phi_s = 3\pi/4$. The blue and red distributions are for input parameters which were shifted by $\pm 1\sigma$. p is the estimated probability and P denotes the probability to observe a value for χ^2 or larger for the agreement of the distributions for the original and shifted parameters.

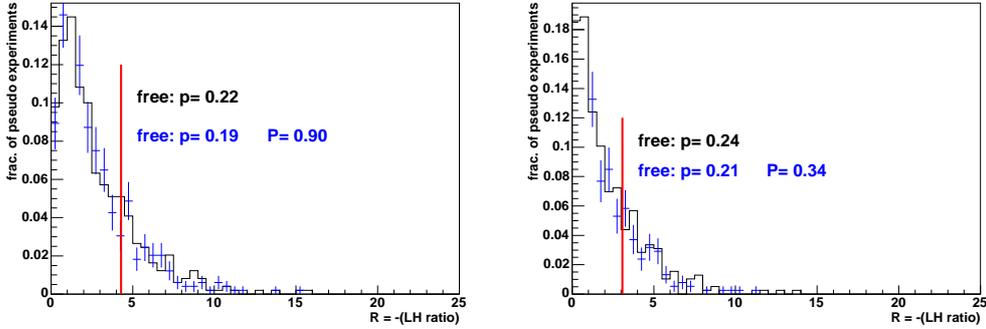


Figure 8.6: Differences of the p -value when letting all parameters float in the final fit and fixing the background parameters using sideband for two points: $\Delta\Gamma = 0.10 \text{ ps}^{-1}$, $\phi_s = 0.0$ and $\Delta\Gamma = 0.15 \text{ ps}^{-1}$, $\phi_s = 3\pi/4$. The blue and red distributions are for input parameters which were shifted by $\pm 1\sigma$. p is the estimated probability and P denotes the probability to observe a value for χ^2 or larger for the agreement of the distributions for the original and shifted parameters.

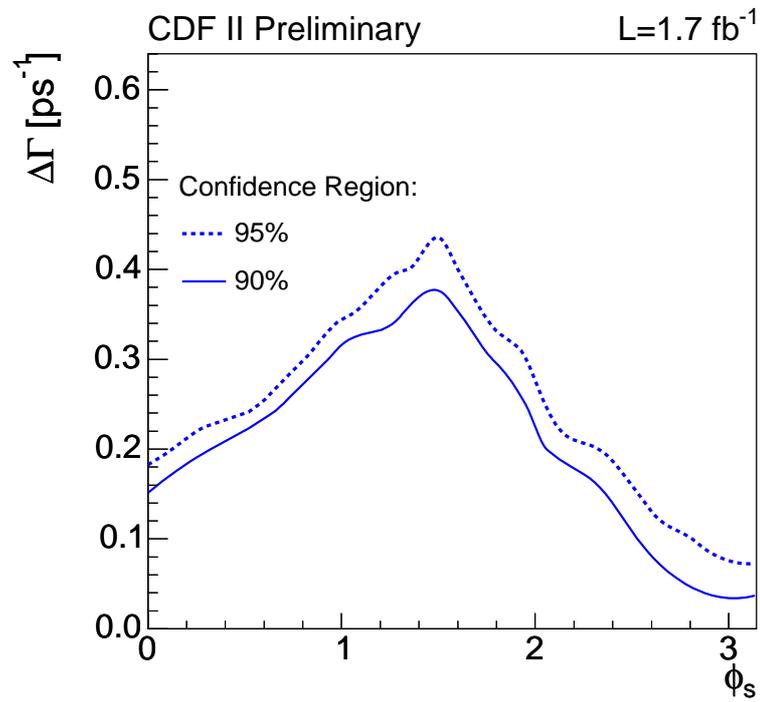


Figure 8.7: 90% and 95% confidence region in the ϕ_S - $\Delta\Gamma_S$ plane. Only the first quadrant is shown. The other three quadrants can be obtained via the transformations given in equations 4.12 and 4.13.

in Figure 8.7. It should be mentioned that the confidence region does not directly relate to a two dimensional likelihood scan. The confidence region takes into account any observed biases by testing the hypothesis for each single point.

Chapter 9

Summary and Outlook

Detailed studies of mesons containing bottom quarks allow precise tests of the Standard Model and provide unique opportunities to search for new phenomena beyond the current theoretical knowledge. The time dependent angular analysis of B meson decays provides an ideal proving ground for our current theoretical understanding. The Tevatron, a proton anti-proton collider at a center of mass energy of $\sqrt{s} = 1.96$ TeV, is currently the only place which provides sufficient energy to produce B_s mesons. The observed events used in this analysis are recorded by the CDF II experiment, one of the two detectors, designed to detect and measure the properties of particles produced in those collisions. The complex hadronic environment and the high collision rate need advanced trigger techniques to select the events of interest. One of the characteristic signatures of the two B meson decays of interest, $B_d \rightarrow J/\psi K^*$ and $B_s \rightarrow J/\psi \phi$, is the occurrence of two muons which can be exploited in the trigger. This analysis uses 1.7 fb^{-1} (1.3 fb^{-1} for B_d) of data preselected by the di-muon trigger. The goal of this analysis was to do measure the lifetime difference $\Delta\Gamma$ of the two mass eigenstates, B_{sH} and B_{sL} , the mean B_s lifetime $c\tau_s$ and the weak phase ϕ_s , using a time dependent angular analysis of the B_s decays. The lifetime difference does also depend on the weak phase ϕ_s , which originates from CP violation in interference between decays with and without mixing. In the Standard Model this phase is very small, allowing for the approximation of no CP violation. These parameters are of particular physical interest as they allow to constrain the possible contribution of new physics. The angular analysis was implemented as an unbinned maximum likelihood fit.

Multiple neural networks are used to improve the event selection of the B mesons. The neural networks allow to combine the input variables optimally into a single discriminating variable, which takes correlations between the variables into account. Simulated signal events and combinatorial back-

par.	this work	Belle [58]	BaBar [57]
$c\tau_d$	$454 \pm 6 \pm 6 \mu\text{m}$		
$ A_0 ^2$	$0.569 \pm 0.009 \pm 0.006$	$0.574 \pm 0.012 \pm 0.009$	$0.556 \pm 0.009 \pm 0.010$
$ A_{ } ^2$	$0.210 \pm 0.011 \pm 0.007$	$0.231 \pm 0.012 \pm 0.008$	$0.211 \pm 0.010 \pm 0.006$
$\delta_{ }$	$-2.990 \pm 0.09 \pm 0.01$	$-2.887 \pm 0.090 \pm 0.008$	$-2.93 \pm 0.08 \pm 0.04$
δ_{\perp}	$2.963 \pm 0.06 \pm 0.01$	$2.938 \pm 0.064 \pm 0.010$	$2.91 \pm 0.05 \pm 0.03$

Table 9.1: Our final result for the mean lifetime, the decay amplitudes and the strong phases for the B_d measurement, and the latest available results from the B factories. The uncertainties show competitive precision of CDF to dedicated B experiments.

ground events from data were used as training patterns. The final candidate selection provided a signal yield of roughly 7500 B_d and 2500 B_s events when optimizing the cut on the neural network output to maximize the significance, $N_S/\sqrt{N_S + N_B}$, where N_S and N_B are the number of signal and background events, respectively.

The $B_d \rightarrow J/\psi K^*$ decay channel was analyzed to cross-check the analysis strategy by comparing with the results from the B factories, which did very similar analyses of the same decay channel. The final result for the mean lifetime, the decay amplitudes and the strong phases are shown together with the latest available results from the B factories [57, 58] in Table 9.1. The results of the independent experiments are consistent within their uncertainties. Also the measured B_d lifetime is in very good agreement to the world average of $458.7 \pm 2.7 \mu\text{m}$ [17]. Compared to the previous CDF measurement [9], the B_d analysis was extended by the neural network for suppressing misreconstructed B_d events, where the kaon/pion particle hypothesis was swapped, by the additional non-resonant $K\pi$ S-wave B_d component and by increased statistics. The new result is compatible with the previous measurement with about 2.5 times smaller uncertainties.

The same method developed in this thesis can be used to investigate direct CP violation, since $B_d \rightarrow J/\psi K^*$ is a flavor specific decay mode where the charge of the kaon is directly correlated to the flavor of the b quark. Typically, direct CP violation is measured by an asymmetry in the total decay rate of each flavor. The angular analysis allows for a complementary measurement in the future by making two independent time dependent angular fits for B_d and \bar{B}_d separately.

Analyzing the $B_s \rightarrow J/\psi \phi$ decay channel assuming no CP violation, $\phi_s =$

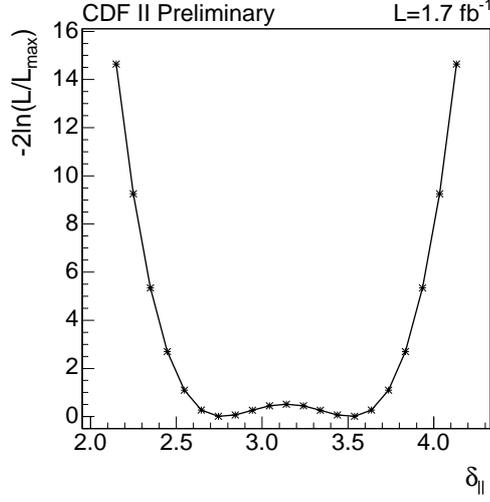


Figure 9.1: B_s likelihood profile for the strong phase under the assumption of no CP violation.

0, the following results for the physical parameters are obtained by the fit

$$\begin{aligned}
 c\tau_s &= 456 \pm 13(\text{stat.}) \pm 7(\text{sys.}) \mu\text{m} \\
 \Delta\Gamma &= 0.076^{+0.059}_{-0.063}(\text{stat.}) \pm 0.006(\text{sys.}) \text{ps}^{-1} \\
 |A_0|^2 &= 0.530 \pm 0.021(\text{stat.}) \pm 0.007(\text{sys.}) \\
 |A_{||}|^2 &= 0.230 \pm 0.027(\text{stat.}) \pm 0.009(\text{sys.}).
 \end{aligned}$$

The likelihood profile for the strong phase $\delta_{||}$ is shown in Figure 9.1. Having the current number of events, two equally allowed solutions are very close to each other. The non-parabolic shape can be understood from the symmetry of the strong phase in the probability density function around π . More statistics will allow to separate the two solutions and to give a numerical result for $\delta_{||}$ in the future.

In addition, a second fit was done where the B_s lifetime was constrained to be equal to the well measured B_d lifetime. The experimental uncertainty was added to the 1% theoretical uncertainty from Heavy Quark Expansion (HQE) theory. For the B_s fit assuming $\phi = 0$ and using the Gaussian lifetime constraint $c\tau_s = 458 \pm 5.3 \mu\text{m}$, the result is

$$\begin{aligned}
 c\tau_s &= 458 \pm 5(\text{stat.}) \pm 7(\text{sys.}) \mu\text{m} \\
 \Delta\Gamma &= 0.081 \pm 0.050(\text{stat.}) \pm 0.006(\text{sys.}) \text{ps}^{-1} \\
 |A_0|^2 &= 0.531 \pm 0.020(\text{stat.}) \pm 0.007(\text{sys.}) \\
 |A_{||}|^2 &= 0.230 \pm 0.026(\text{stat.}) \pm 0.009(\text{sys.}).
 \end{aligned}$$

The lifetime difference of the B_s mass eigenstates from the constrained and unconstrained fits are in very good agreement with Standard Model expectations. The measured mean lifetime for the unconstrained B_s fit is also in very good agreement with the current mean lifetime for B_d . The agreement between the mean lifetimes of the two neutral B mesons, B_s and B_d , is one of the predictions from HQE. The B_s lifetime and the lifetime difference are currently the most precise measurements of these quantities.

When fitting the floating weak phase, a bias to larger values in $\Delta\Gamma$ and ϕ_s is observed. The source of the bias is related to the smallness of the physical parameters of interest and the structure of the probability density function. Due to the limited number of events, the statistical uncertainties on those parameters are rather large and the maximum likelihood fit tends to increase those parameters, since it allows to describe any statistical fluctuation. To circumvent those difficulties probabilities are derived from likelihood ratios to construct a confidence region, which takes the bias into account. A likelihood ratio, following the procedure suggested by Feldman-Cousins [64], is used to test for the compatibility of a specific $\Delta\Gamma$ - ϕ_s pair with the data. Scanning many points in a grid in the $\Delta\Gamma$ - ϕ_s plane, a region is constructed which contains the true $\Delta\Gamma$ - ϕ_s values at a 90% or 95% confidence level. One of these points is the Standard Model assumption, where the interference induced weak phase is negligible and $\Delta\Gamma$ was set to 0.1 ps^{-1} . The probability to observe our result on data under the Standard Model assumption is 22%. The whole confidence region is shown in Figure 9.2. This figure also includes a band which denotes new physics models. Our result does not allow to exclude any of these models.

In conclusion, the currently most precise values for the mean lifetime and the lifetime difference of the two mass eigenstates of the B_s are obtained. The measurement of the CP violating phase is in good agreement with the Standard Model expectation.

The time dependent angular analysis can be extended to add the flavor tagging information of the b quark at production time for each event into the fit. One of the prerequisites for this is a good measurement of the B_s mixing frequency, done last year by CDF [4]. The result of the mixing measurement combined with the flavor tagging information will lead to a more precise measurement of the weak phase using the angular analysis and will allow to make more stringent constraints on physics beyond the Standard Model. This effort is currently ongoing by both collaborations DØ and CDF at the Tevatron and will provide interesting results and more insight into the B_s system in the near future.

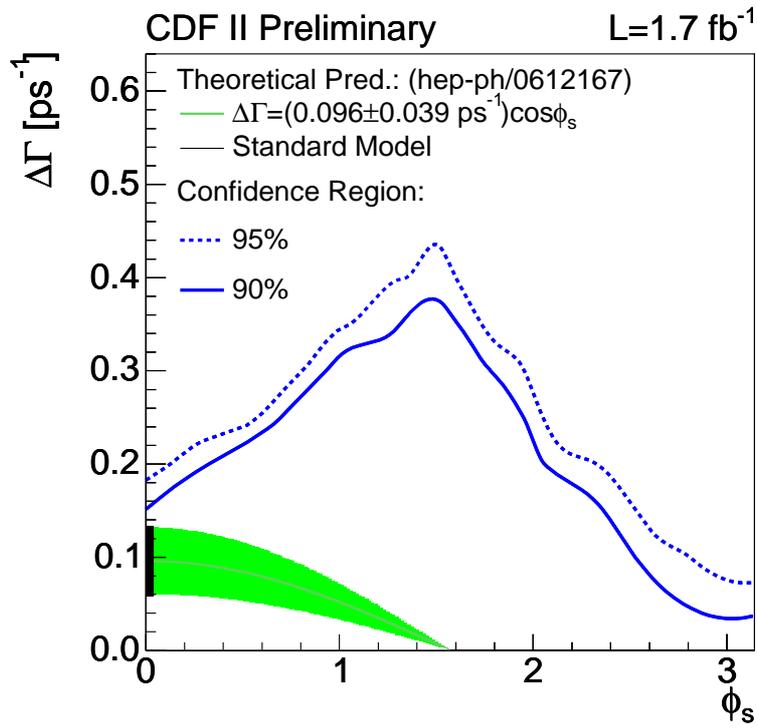


Figure 9.2: 90% and 95% confidence region in the ϕ_s - $\Delta\Gamma_s$ plane. Only the first quadrant is shown. The other three quadrants can be obtained via the transformations $\phi_s \rightarrow -\phi_s$ and $\Delta\Gamma \rightarrow -\Delta\Gamma$, $\phi_s \rightarrow \phi_s + \pi$.

Appendix A

Trigger Paths

L2 Trigger	L3 Trigger
L2.AUTO_L1.CMU1.5_PT1.5.&.CMX1.5_PT2	JPSL.CMU1.5.CMX20
L2.AUTO_L1.CMU1.5_PT1.5.&.CMX1.5_PT2_CSX	JPSL.CMU1.5.CMX22
L2.AUTO_L1.CMU1.5_PT1.5.&.CMX1.5_PT2_CSX_PS0	JPSL.CMU1.5.CMX23
L2.AUTO_L1.TWO.CMU1.5_PT1.5	JPSL.CMU1.5.CMX2_ALLPHI
L2.PS100.L1.CMU1.5_PT1.5.&.CMX1.5_PT2_CSX	JPSL.CMU1.5.CMX2_NOL2
L2.PS100.L1.TWO.CMU1.5_PT1.5	JPSL.CMU1.5.CMX2_NOL2
L2.PS10.L1.CMU1.5_PT1.5.&.CMX1.5_PT2_CSX	JPSL.CMU1.5.CMX2
L2.PS10.L1.TWO.CMU1.5_PT1.5	JPSL.CMU2.CMX2_PS10
L2.TWO.CMU1.5_PT1.5_DPFI120.OPPQ	JPSL.CMU2.CMX2_PS2
L2.TWO.CMU1.5_PT1.5_DPFI120.OPPQ_DPS	JPSL.CMU2.CMX2_PS50
L2.TWO.CMU1.5_PT2_DPFI120.OPPQ	JPSL.CMUCMU1.5
L2.TWO.CMU1.5_PT2_DPFI120.OPPQ_PS10	JPSL.CMUCMU1.5_ALLPHI
L2.AUTO_L1.CMUP6_PT4	JPSL.CMUCMU1.5_NOL2
L2.CMU1.5_PT1.5.&.CMX1.5_PT2_DPFI120.OPPQ	JPSL.CMUCMU2
L2.CMU1.5_PT1.5.&.CMX1.5_PT2_DPFI120.OPPQ_DPS	JPSL.CMUCMU2_PS10
L2.CMU1.5_PT1.7.&.CMU1.5_PT3.1.7MT7_DPS	JPSL.CMUCMU2_PS2
L2.CMU1.5_PT1.7.&.CMU1.5_PT3.1.7MT7_LUMI_185	JPSL.CMUCMU2_PS50
L2.CMU1.5_PT1.7.&.CMX1.5_PT3.1.7MT7_DPS	JPSL.CMUCMU_MT_DPS
L2.CMU1.5_PT1.7.&.CMX1.5_PT3.1.7MT7_LUMI_185	JPSL.CMUCMU_MT_LUMI_185
L2.CMU1.5_PT2.&.CMX1.5_PT2_DPFI120.OPPQ_PS10	JPSL.CMUCMX3.MT_DPS
L2.CMU1.5_PT2.&.CMX1.5_PT2_DPFI120.OPPQ_PS2	JPSL.CMUCMX3.MT_LUMI_185
L2.CMU1.5_PT2.&.CMX1.5_PT2_DPFI120.OPPQ_PS50	JPSL.CMUP4.CMU1.5
L2.CMUP6_PT8	JPSL.CMUP4.CMU1.5.L2_DPS
L2.CMUP6_PT8_DPS	JPSL.CMUP4.CMU.L2_DPS
L2.CMX1.5_PT2.&.CMU1.5_PT3.1.7MT7_DPS	JPSL.CMUP4.CMX2
L2.CMX1.5_PT2.&.CMU1.5_PT3.1.7MT7_LUMI_185	JPSL.CMUP4.CMX2.L2_DPS

L2_DPS_L1_CMUP6_PT4	JPSI_CMUP4_CMX_L2_DPS
L2_PS200_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX	JPSI_CMUPCMU_HIGHPT
L2_PS200_L1_TWO_CMU1.5_PT1.5	JPSI_CMUPCMU_HIGHPT_DPS
L2_PS500_L1_CMU1.5_PT1.5_&_CMX1.5_PT2_CSX	JPSI_CMUPCMX_HIGHPT
L2_RL20HZ_L1_CMUP6_PT4	JPSI_CMUPCMX_HIGHPT_DPS
L2_TRK8_L1_CMUP6_PT4	JPSI_CMXCMU3_MT_DPS
L2_TWO_CMU1.5_PT2_DPFI120_OPPQ_PS2	JPSI_CMXCMU3_MT_LUMI185
L2_TWO_CMU1.5_PT2_DPFI120_OPPQ_PS50	EXPRESS_JPSI_CMUCMU

Table A.1: Summary of triggers used in this analysis.

Appendix B

Neural Network Input Variables

#	Variable B_s	B_d	
1	training target	←	the target of the training: signal or background
3	$p_t(B)$	←	transverse momentum of the B meson (B_s or B_d)
10	$p_t(J/\psi)$	←	transverse momentum of the intermediate J/ψ resonance
5	$p_t(\phi)$	$p_t(K^*)$	transverse momentum of the intermediate ϕ or K^* resonance
8	$p_t(K_1)$	$p_t(K)$	transverse momentum of the one of the B_s kaons or of the kaon in B_d
9	$p_t(K_2)$	$p_t(\pi)$	transverse momentum of the other kaon or pion
15	$\max(p_t(\mu^+), p_t(\mu^-))$	←	maximal transverse momentum of the two muons
16	$\min(p_t(\mu^+), p_t(\mu^-))$	←	minimal transverse momentum of the two muons
12	$ M_{\mu\mu} - M_{J/\psi}^{PDG} $	←	mass difference of the di-muon resonance to the J/ψ world average mass
7	$ M_{KK} $	$ M_{K\pi} $	mass of the di-kaon (B_s) or kaon-pion (B_d) resonance
4	$Prob(\chi^2)(B_s)$	←	Probability of the B meson vertex fit

11	$Prob(\chi^2)(J/\psi)$	\leftarrow	Probability of the J/ψ resonance vertex fit
6	$Prob(\chi^2)(\phi)$	(K^*)	Probability of the ϕ or K^* resonance vertex fit
2	$\chi_{r\phi}^2(B)$	\leftarrow	$\chi_{r\phi}^2$ of the B meson vertex fit
17	$PID_K(K^+)$	$PID_K(K)$	likelihood ratio for kaon particle identification
18	$PID_\pi(K^+)$	$PID_\pi(K)$	likelihood ratio for pion particle identification
19	$PID_K(K^-)$	$PID_K(\pi)$	likelihood ratio for kaon particle identification
20	$PID_\pi(K^-)$	$PID_\pi(\pi)$	likelihood ratio for pion particle identification
13	$\max(\text{lh}(\mu^+), \text{lh}(\mu^-))$	\leftarrow	likelihood ratio for muon particle identification
14	$\min(\text{lh}(\mu^+), \text{lh}(\mu^-))$	\leftarrow	likelihood ratio for muon particle identification

Appendix C

Derivation of the fit function

The time evolution of the angular distribution for tagged B_s mesons is [16]

$$\begin{aligned}
\frac{d^3\Gamma(B_s(t))}{d\vec{\omega}} &\propto |A_0(t)|^2 f_1(\vec{\omega}) + |A_{||}(t)|^2 f_2(\vec{\omega}) \\
&+ |A_{\perp}(t)|^2 f_3(\vec{\omega}) + \text{Re}(A_0(t)^* A_{||}(t)) f_5(\vec{\omega}) \\
&+ \text{Im}(A_{||}(t)^* A_{\perp}(t)) f_4(\vec{\omega}) \\
&+ \text{Im}(A_0(t)^* A_{\perp}(t)) f_6(\vec{\omega}),
\end{aligned} \tag{C.1}$$

whereas for \bar{B}_s mesons it has the form

$$\begin{aligned}
\frac{d^3\Gamma(\bar{B}_s(t))}{d\vec{\omega}} &\propto |\bar{A}_0(t)|^2 f_1(\vec{\omega}) + |\bar{A}_{||}(t)|^2 f_2(\vec{\omega}) \\
&+ |\bar{A}_{\perp}(t)|^2 f_3(\vec{\omega}) + \text{Re}(\bar{A}_0^*(t) \bar{A}_{||}(t)) f_5(\vec{\omega}) \\
&+ \text{Im}(\bar{A}_{||}^*(t) \bar{A}_{\perp}(t)) f_4(\vec{\omega}) \\
&+ \text{Im}(\bar{A}_0^*(t) \bar{A}_{\perp}(t)) f_6(\vec{\omega}),
\end{aligned} \tag{C.2}$$

using [65]

$$\begin{aligned}
A_{\tilde{f}}^*(t) A_f(t) &= \langle (X_1 X_2)_{\tilde{f}} | H_{eff} | B_s \rangle^* \langle (X_1 X_2)_f | H_{eff} | B_s \rangle \\
&\times [|g_+(t)|^2 + \eta_{CP}^{\tilde{f}} e^{-i\phi_s} g_+(t) g_-^*(t) + \eta_{CP}^f e^{i\phi_s} g_+(t)^* g_-(t) + \eta_{CP}^{\tilde{f}} \eta_{CP}^f |g_-(t)|^2],
\end{aligned} \tag{C.3}$$

and

$$\begin{aligned}
\bar{A}_{\tilde{f}}^*(t) \bar{A}_f(t) &= \langle (X_1 X_2)_{\tilde{f}} | H_{eff} | B_s \rangle^* \langle (X_1 X_2)_f | H_{eff} | B_s \rangle \\
&\times [|g_-(t)|^2 + \eta_{CP}^{\tilde{f}} e^{-i\phi_s} g_+^*(t) g_-(t) + \eta_{CP}^f e^{i\phi_s} g_+(t) g_-^*(t) + \eta_{CP}^{\tilde{f}} \eta_{CP}^f |g_+(t)|^2],
\end{aligned} \tag{C.4}$$

where

$$\begin{aligned}
|g_{\pm}(t)|^2 &= \frac{1}{4}[e^{-\Gamma_L t} + e^{-\Gamma_L t} \pm 2e^{-\Gamma_L t} \cos(\Delta mt)] \\
g_+(t)g_-^*(t) &= \frac{1}{4}[e^{-\Gamma_L t} - e^{-\Gamma_L t} - 2ie^{-\Gamma_L t} \sin(\Delta mt)]. \quad (C.5)
\end{aligned}$$

A straight forward calculation gives the time evolution

$$\begin{aligned}
|A_0(t)|^2 &= \frac{|A_0(0)|^2}{2}[(1 + \cos(\phi_s))e^{-\Gamma_L t} + (1 - \cos(\phi_s))e^{-\Gamma_H t} \\
&\quad - 2e^{-\Gamma t} \sin(\Delta mt) \sin(\phi_s)] \\
|\bar{A}_0(t)|^2 &= \frac{|A_0(0)|^2}{2}[(1 + \cos(\phi_s))e^{-\Gamma_L t} + (1 - \cos(\phi_s))e^{-\Gamma_H t} \\
&\quad + 2e^{-\Gamma t} \sin(\Delta mt) \sin(\phi_s)] \\
|A_{||}(t)|^2 &= \frac{|A_{||}(0)|^2}{2}[(1 + \cos(\phi_s))e^{-\Gamma_L t} + (1 - \cos(\phi_s))e^{-\Gamma_H t} \\
&\quad - 2e^{-\Gamma t} \sin(\Delta mt) \sin(\phi_s)] \\
|\bar{A}_{||}(t)|^2 &= \frac{|A_{||}(0)|^2}{2}[(1 + \cos(\phi_s))e^{-\Gamma_L t} + (1 - \cos(\phi_s))e^{-\Gamma_H t} \\
&\quad + 2e^{-\Gamma t} \sin(\Delta mt) \sin(\phi_s)] \\
|A_{\perp}(t)|^2 &= \frac{|A_{\perp}(0)|^2}{2}[(1 - \cos(\phi_s))e^{-\Gamma_L t} + (1 + \cos(\phi_s))e^{-\Gamma_H t} \\
&\quad + 2e^{-\Gamma t} \sin(\Delta mt) \sin(\phi_s)] \\
|\bar{A}_{\perp}(t)|^2 &= \frac{|A_{\perp}(0)|^2}{2}[(1 - \cos(\phi_s))e^{-\Gamma_L t} + (1 + \cos(\phi_s))e^{-\Gamma_H t} \\
&\quad - 2e^{-\Gamma t} \sin(\Delta mt) \sin(\phi_s)]
\end{aligned} \quad (C.6)$$

$$\begin{aligned}
\text{Re}(A_0(t)^* A_{||}(t)) &= \frac{|A_0(0)||A_{||}(0)|}{2} \cos(\delta_0 + \delta_{||}) [(1 + \cos(\phi_s))e^{-\Gamma_L t} + \\
&\quad (1 - \cos(\phi_s))e^{-\Gamma_H t} - 2e^{-\Gamma t} \sin(\Delta mt) \sin(\phi_s)] \\
\text{Re}(\bar{A}_0(t)^* \bar{A}_{||}(t)) &= \frac{|A_0(0)||A_{||}(0)|}{2} \cos(\delta_0 + \delta_{||}) [(1 + \cos(\phi_s))e^{-\Gamma_L t} + \\
&\quad (1 - \cos(\phi_s))e^{-\Gamma_H t} + 2e^{-\Gamma t} \sin(\Delta mt) \sin(\phi_s)] \\
\text{Im}(A_{||}(t)^* A_{\perp}(t)) &= |A_{||}||A_{\perp}| [+e^{-\Gamma t} \cos(\Delta mt) \sin(\delta_{\perp} - \delta_{||}) \\
&\quad - 0.5 \sin(\phi_s) \cos(\delta_{\perp} - \delta_{||}) (e^{-\Gamma_L t} - e^{-\Gamma_H t}) \\
&\quad - e^{-\Gamma t} \sin(\Delta mt) \cos(\phi_s) \cos(\delta_{\perp} - \delta_{||})] \\
\text{Im}(\bar{A}_{||}(t)^* \bar{A}_{\perp}(t)) &= |A_{||}||A_{\perp}| [-e^{-\Gamma t} \cos(\Delta mt) \sin(\delta_{\perp} - \delta_{||}) \\
&\quad - 0.5 \sin(\phi_s) \cos(\delta_{\perp} - \delta_{||}) (e^{-\Gamma_L t} - e^{-\Gamma_H t}) \\
&\quad + e^{-\Gamma t} \sin(\Delta mt) \cos(\phi_s) \cos(\delta_{\perp} - \delta_{||})] \\
\text{Im}(A_0(t)^* A_{\perp}(t)) &= |A_0||A_{\perp}| [+e^{-\Gamma t} \cos(\Delta mt) \sin(\delta_{\perp} - \delta_0) \\
&\quad - 0.5 \sin(\phi_s) \cos(\delta_{\perp} - \delta_0) (e^{-\Gamma_L t} - e^{-\Gamma_H t}) \\
&\quad - e^{-\Gamma t} \sin(\Delta mt) \cos(\phi_s) \cos(\delta_{\perp} - \delta_0)] \\
\text{Im}(\bar{A}_0(t)^* \bar{A}_{\perp}(t)) &= |A_0||A_{\perp}| [-e^{-\Gamma t} \cos(\Delta mt) \sin(\delta_{\perp} - \delta_0) \\
&\quad - 0.5 \sin(\phi_s) \cos(\delta_{\perp} - \delta_0) (e^{-\Gamma_L t} - e^{-\Gamma_H t}) \\
&\quad + e^{-\Gamma t} \sin(\Delta mt) \cos(\phi_s) \cos(\delta_{\perp} - \delta_0)]
\end{aligned} \tag{C.7}$$

Taking the tagging decision $\xi \in \{-1, 0, 1\}$ and dilution $D = 2P - 1$ into account the probability density function gets

$$\begin{aligned}
\frac{d^4 P(\vec{\omega}, t)}{d\vec{\omega} dt} &\propto |A_0(0)|^2 f_1(\vec{\omega}) (\mathcal{T}_+ - (\xi DT)) \\
&+ |A_{||}(0)|^2 f_2(\vec{\omega}) (\mathcal{T}_+ - (\xi DT)) \\
&+ |A_{\perp}(0)|^2 f_3(\vec{\omega}) (\mathcal{T}_- + (\xi DT)) \\
&+ |A_0(0)||A_{||}(0)| f_5(\vec{\omega}) \cos(\delta_0 + \delta_{||}) (\mathcal{T}_+ - (\xi DT)) \\
&+ |A_{||}(0)||A_{\perp}(0)| f_4(\vec{\omega}) [\cos(\delta_{\perp} - \delta_{||}) \sin \phi_s (e^{-\Gamma_H t} - e^{-\Gamma_L t})/2 \\
&\quad + e^{-\Gamma t} \xi D (\cos(\Delta mt) \sin(\delta_{\perp} - \delta_{||}) \\
&\quad \quad - \sin(\Delta mt) \cos(\phi_s) \cos(\delta_{\perp} - \delta_{||}))] \\
&+ |A_0(0)||A_{\perp}(0)| f_6(\vec{\omega}) [\cos(\delta_{\perp} - \delta_0) \sin \phi_s (e^{-\Gamma_H t} - e^{-\Gamma_L t})/2 \\
&\quad + e^{-\Gamma t} \xi D (\cos(\Delta mt) \sin(\delta_{\perp} - \delta_0) \\
&\quad \quad - \sin(\Delta mt) \cos(\phi_s) \cos(\delta_{\perp} - \delta_0))]
\end{aligned} \tag{C.8}$$

where

$$\begin{aligned}\mathcal{T}_{\pm} &= ((1 \pm \cos \phi_s)e^{-\Gamma_L t} + (1 \mp \cos \phi_s)e^{-\Gamma_H t})/2 \\ \mathcal{T} &= e^{-\Gamma t} \sin(\Delta m t) \sin(\phi_s)\end{aligned}\tag{C.9}$$

In the experiment only the difference of the strong phases can be measured. This allows us to fix one of the strong phases. Throughout this thesis δ_0 is fixed to zero, $\delta_0 = 0$.

Appendix D

Normalization

The normalization factor for signal PDF multiplied with the acceptance has the following form

$$\begin{aligned}
P(\vec{\omega}, t) \times A(\vec{\omega}) \propto & A(\vec{\omega}) \times \int dt \int d\vec{\omega} \{ |A_0(0)|^2 f_1(\vec{\omega})(\mathcal{T}_+) + \\
& |A_{\parallel}(0)|^2 f_2(\vec{\omega})(\mathcal{T}_+) + \\
& |A_{\perp}(0)|^2 f_3(\vec{\omega})(\mathcal{T}_-) + \\
& |A_0(0)||A_{\parallel}(0)| f_5(\vec{\omega}) \cos(\delta_{\parallel})(\mathcal{T}_+) + \\
& |A_{\parallel}(0)||A_{\perp}(0)| f_4(\vec{\omega}) [\cos(\delta_{\perp} - \delta_{\parallel}) \sin \phi_s (e^{-\Gamma_H t} - e^{-\Gamma_L t})/2] + \\
& |A_0(0)||A_{\perp}(0)| f_6(\vec{\omega}) [\cos(\delta_{\perp}) \sin \phi_s (e^{-\Gamma_H t} - e^{-\Gamma_L t})/2] \}.
\end{aligned} \tag{D.1}$$

We can integrate out the time dependence analytically

$$\begin{aligned}
P(\vec{\omega}, t) \times A(\vec{\omega}) \propto & \int d\vec{\omega} \times A(i, j, k) \{ |A_0(0)|^2 f_1(\vec{\omega}) [(1 + \cos \phi_s) c\tau_L + \\
& (1 - \cos \phi_s) c\tau_H] + \\
& |A_{\parallel}(0)|^2 f_2(\vec{\omega}) [(1 + \cos \phi_s) c\tau_L + (1 - \cos \phi_s) c\tau_H] + \\
& |A_{\perp}(0)|^2 f_3(\vec{\omega}) [(1 - \cos \phi_s) c\tau_L + (1 + \cos \phi_s) c\tau_H] + \\
& |A_0(0)||A_{\parallel}(0)| f_5(\vec{\omega}) \cos(\delta_{\parallel}) [(1 + \cos \phi_s) c\tau_L + \\
& (1 - \cos \phi_s) c\tau_H] + \\
& |A_{\parallel}(0)||A_{\perp}(0)| f_4(\vec{\omega}) [\cos(\delta_{\perp} - \delta_{\parallel}) \sin \phi_s (c\tau_H - c\tau_L)/2] + \\
& |A_0(0)||A_{\perp}(0)| f_6(\vec{\omega}) [\cos(\delta_{\perp}) \sin \phi_s (c\tau_H - c\tau_L)/2] \} \\
& \times A(\vec{\omega}).
\end{aligned} \tag{D.2}$$

Since we use a three dimensional histogram $A(i, j, k)$ for the acceptance we can sum over the bins and integrate the angular functions over each bin

$$\begin{aligned}
& P(\vec{\omega}, t) \times A(\vec{\omega}) \propto \\
& \sum_{i,j,k} A(i, j, k) \times \{ |A_0(0)|^2 ((1 + \cos \phi_s) c\tau_L + (1 - \cos \phi_s) c\tau_H) \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_1(\vec{\omega}) + \\
& |A_{||}(0)|^2 ((1 + \cos \phi_s) c\tau_L + (1 - \cos \phi_s) c\tau_H) \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_2(\vec{\omega}) + \\
& |A_{\perp}(0)|^2 ((1 - \cos \phi_s) c\tau_L + (1 + \cos \phi_s) c\tau_H) \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_3(\vec{\omega}) + \\
& |A_0(0)| |A_{||}(0)| \cos(\delta_{||}) ((1 + \cos \phi_s) c\tau_L + (1 - \cos \phi_s) c\tau_H) \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_5(\vec{\omega}) + \\
& |A_{||}(0)| |A_{\perp}(0)| [\cos(\delta_{\perp} - \delta_{||}) \sin \phi_s (c\tau_H - c\tau_L) / 2] \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_4(\vec{\omega}) + \\
& |A_0(0)| |A_{\perp}(0)| [\cos(\delta_{\perp}) \sin \phi_s (c\tau_H - c\tau_L) / 2] \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_6(\vec{\omega}) \}, \tag{D.3}
\end{aligned}$$

where

$$\begin{aligned}
\int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_1(\vec{\omega}) &= [F_1]_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} = \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} \frac{9}{32\pi} 2 \cos^2 \psi (1 - \sin^2 \theta \cos^2 \phi) \\
\int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_2(\vec{\omega}) &= [F_2]_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} = \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} \frac{9}{32\pi} \sin^2 \psi (1 - \sin^2 \theta \sin^2 \phi) \\
\int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_3(\vec{\omega}) &= [F_3]_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} = \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} \frac{9}{32\pi} \sin^2 \psi \sin^2 \theta \\
\int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_4(\vec{\omega}) &= [F_4]_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} = \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} - \frac{9}{32\pi} \sin^2 \psi \sin 2\theta \sin \phi \\
\int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_5(\vec{\omega}) &= [F_5]_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} = \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} \frac{9}{32\pi} \frac{1}{\sqrt{2}} \sin 2\psi \sin^2 \theta \sin 2\phi \\
\int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} f_6(\vec{\omega}) &= [F_6]_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} = \int_{\vec{\omega}_{min}(i,j,k)}^{\vec{\omega}_{max}(i,j,k)} d\vec{\omega} \frac{9}{32\pi} \frac{1}{\sqrt{2}} \sin 2\psi \sin 2\theta \cos \phi. \tag{D.4}
\end{aligned}$$

The $F_1 \rightarrow F_6$ have to be calculated once at the beginning of the fit and keep their values. The summation over bins of the acceptance histogram has to be redone each time a parameter changes.

Acknowledgements

First I would like to thank my supervisor Professor Dr. Michael Feindt for accepting me as a PhD student in his group and offering me this very interesting topic. His guidance and many instructive discussions were very helpful for the success of this work.

I would also like to thank Professor Dr. Günter Quast for co-supervising this thesis and Professor Dr. Thomas Müller for his continuous work for everybody as head of the institute.

This work would not have been possible without the enormous support by Dr. Michal Kreps and Dr. Thomas Kuhr. Their advice and our discussions were very helpful and I am very grateful for that. Further I would like to thank Gian Piero di Giovanni and Khaldoun Makhoul, with whom we worked together for combining our measurements.

Special thanks go to all people at Fermilab, who helped to improve this analysis through their comments and suggestions. Especially I would like to mention the conveners and sub-conveners: Cheng-Ju Lin, Manfred Paulini, Sinead Farrington and Diego Tonelli.

I also thank all members of the Karlsruhe *B* group for their great working atmosphere and help. Especially I would like to mention Dr. Michal Kreps, Dr. Thomas Kuhr, Andreas Schmidt, Jan Morlock and Dr. Ulrich Kerzel for carefully reading and commenting on this manuscript and the associated CDF notes.

This work was supported by the scholarship of the “Graduiertenkolleg Hochenergiephysik und Teilchenphysik” promoted by the German Research Community and the Federal Ministry of Education and Research.

For the great time together, I would like to thank all my room mates with whom I shared office 9-9: Dr. Thorsten Scheidle, Dr. Ulrich Kerzel, Andreas Gessler and Dominik Horn.

Finally I would like to thank my friends and family, especially my wife Stephanie.

Bibliography

- [1] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman, and W. Chinowsky. *Observation of Long-Lived Neutral V Particles*. *Phys. Rev.*, 103(6):1901–1904, Sep 1956.
- [2] B. Aubert et al. *Measurement of $B^0 - \bar{B}^0$ flavor oscillations in hadronic B^0 decays*. *Phys. Rev. Lett.*, 88:221802, 2002.
- [3] T. Tomura et al. *Measurement of the oscillation frequency for $B^0 - \bar{B}^0$ mixing using hadronic B^0 decays*. *Phys. Lett.*, B542:207–215, 2002.
- [4] A. Abulencia et al. *Observation of $B_s^0 - \bar{B}_s^0$ oscillations*. *Phys. Rev. Lett.*, 97:242003, 2006.
- [5] M. Staric et al. *Evidence for $D^0 - \bar{D}^0$ Mixing*. *Phys. Rev. Lett.*, 98:211803, 2007.
- [6] B. Aubert et al. *Evidence for $D^0 - \bar{D}^0$ mixing*. *Phys. Rev. Lett.*, 98:211802, 2007.
- [7] D0 Collaboration. *Direct Limits on the B_s^0 Oscillation Frequency*. *Physical Review Letters*, 97(2):021802, 2006.
- [8] D0 Collaboration. *Lifetime Difference and CP-Violating Phase in the B_s^0 System*. *Physical Review Letters*, 98(12):121801, 2007.
- [9] D. Acosta et al. *Measurement of the lifetime difference between B_s mass eigenstates*. *Phys. Rev. Lett.*, 94:101803, 2005.
- [10] Ikaros I. Y. Bigi. *The QCD perspective on lifetimes of heavy flavor hadrons*. 1995. hep-ph/9508408.
- [11] E. Franco, V. Lubicz, F. Mescia, and C. Tarantino. *Lifetime ratios of beauty hadrons at the next-to-leading order in QCD*. *Nucl. Phys.*, B633:212–236, 2002.

- [12] Ikaros I. Y. Bigi. *The lifetimes of heavy flavour hadrons: A case study in quark hadron duality*. 1999. hep-ph/0001003.
- [13] Martin Beneke, Gerhard Buchalla, Christoph Greub, Alexander Lenz, and Ulrich Nierste. *The $B^+ - B_d^0$ lifetime difference beyond leading logarithms*. *Nucl. Phys.*, B639:389–407, 2002.
- [14] Isard Dunietz and Jonathan L. Rosner. *Time Dependent CP Violation Effects in $B^0 - \bar{B}^0$ Systems*. *Phys. Rev.*, D34:1404, 1986.
- [15] K. Anikeev et al. *B physics at the Tevatron: Run II and beyond*. 2001. hep-ph/0201071.
- [16] Amol S. Dighe, Isard Dunietz, Harry J. Lipkin, and Jonathan L. Rosner. *Angular distributions and lifetime differences in $B_s \rightarrow J/\psi\phi$ decays*. *Phys. Lett.*, B369:144–150, 1996.
- [17] W.-M. Yao et al. *Review of Particle Physics*. *Journal of Physics G*, 33:1+, 2006.
- [18] Alexander Lenz and Ulrich Nierste. *Theoretical update of $B_s - \bar{B}_s$ mixing*. *JHEP*, 06:072, 2007.
- [19] Fermilab Beams Division. *Fermilab Linac upgrade conceptual design revision 4A*. FERMILAB-LU-CONCEPTUAL-DESIGN <http://lss.fnal.gov/archive/linac/fermilab-lu-999.pdf>.
- [20] Fermilab Beams Division. *Run II handbook*. http://www-bdnew.fnal.gov/pbar/run2b/Documents/RunII_handbook.pdf.
- [21] Fermilab Beams Division. *Booster Rookie Book*. <http://www-bd.fnal.gov/proton/booster/rookie/realrookiebook.html>.
- [22] Fermilab Beams Division. *Operations Rookie Books*. http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html.
- [23] Fermilab Beams Division. *Antiproton Source Rookie Books*. http://www-bdnew.fnal.gov/operations/rookie_books/Pbar_V1_1.pdf.
- [24] D. Mohl, G. Petrucci, L. Thorndahl, and Simon Van Der Meer. *Physics and Technique of Stochastic Cooling*. *Phys. Rept.*, 58:73–119, 1980.

- [25] Simon Van Der Meer. *Stochastic cooling and the accumulation of anti-protons*. *Rev. Mod. Phys.*, 57:689–697, 1985.
- [26] S. Nagaitsev, A. Burov, K. Carlson, V. Dudnikov, B. Kramper, T. Kroc, J. Leibfritz, M. McGee, G. Saewert, C. W. Schmidt, A. Shemyakin, A. Warner, S. Seletsky, and V. Tupikov. *Status of the Fermilab electron cooling project*. *Nuclear Instruments and Methods in Physics Research A*, 532:275–278, October 2004.
- [27] S. Nagaitsev et al. *Experimental demonstration of relativistic electron cooling*. *Phys. Rev. Lett.*, 96:044801, 2006.
- [28] G. I. Budker. *An effective method of damping particle oscillations in proton and anti-proton storage rings*. *Sov. Atom. Energ.*, 22:438–440, 1967.
- [29] Fermilab Beams Division. *Recycler Rookie Book*.
http://www-bdnew.fnal.gov/operations/rookie_books/Recycler_RB_v1.pdf.
- [30] D. Acosta et al. *Measurement of the J/ψ meson and b -hadron production cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1960$ GeV*. *Phys. Rev.*, D71:032001, 2005.
- [31] Christopher S. Hill. *Operational experience and performance of the CDFII silicon detector*. *Nucl. Instrum. Meth.*, A530:1–6, 2004.
- [32] A. Sill. *CDF Run II silicon tracking projects*. *Nucl. Instrum. Meth.*, A447:1–8, 2000.
- [33] Anthony A. Affolder et al. *Intermediate silicon layers detector for the CDF experiment*. *Nucl. Instrum. Meth.*, A453:84–88, 2000.
- [34] Anthony A. Affolder et al. *CDF central outer tracker*. *Nucl. Instrum. Meth.*, A526:249–299, 2004.
- [35] D. Acosta et al. *A time-of-flight detector in CDF-II*. *Nucl. Instrum. Meth.*, A518:605–608, 2004.
- [36] G. Ascoli et al. *CDF Central Muon Detector*. *Nucl. Instrum. Meth.*, A268:33, 1988.
- [37] L. Balka et al. *The CDF Central Electromagnetic Calorimeter*. *Nucl. Instrum. Meth.*, A267:272, 1988.

- [38] S. R. Hahn et al. *Calibration Systems for the CDF Central Electromagnetic Calorimeter*. *Nucl. Instr. Meth.*, A267:351, 1988.
- [39] S. Bertolucci et al. *The CDF Central and Endwall Hadron Calorimeter*. *Nucl. Instrum. Meth.*, A267:301, 1988.
- [40] Evelyn J. Thomson et al. *Online track processor for the CDF upgrade*. *IEEE Trans. Nucl. Sci.*, 49:1063–1070, 2002.
- [41] Bill Ashmanskas et al. *The CDF silicon vertex trigger*. *Nucl. Instrum. Meth.*, A518:532–536, 2004.
- [42] C. Blocker, J. Boudreau, and Ch. Paus. *Common Tools for B Physics Analyses in Run II at CDF*. CDF internal note: 5735.
- [43] Ch. Paus et al.
<http://cdfkits.fnal.gov/CdfCode/source/BottomMods/>.
- [44] Ch. Paus et al. <http://mit1.fnal.gov/~paus/BStntuple/>.
- [45] D. J. Lange. *The EvtGen particle decay simulation package*. *Nucl. Instrum. Meth.*, A462:152–155, 2001.
- [46] *CDF Detector Simulation*.
http://www-cdf.fnal.gov/cdfsims/cdfsims_main.html.
- [47] *The full setup of the simulation can be found at*.
<http://www-ekp.physik.uni-karlsruhe.de/~kreps/analysis/MCProduction/>.
- [48] Michael Feindt. *A Neural Bayesian Estimator for Conditional Probability Densities*. *arxiv:physics/0402093*, 2004.
- [49] Christian Doerr, Michael Feindt, and Michal Kreps. *Signal Optimization for $B \rightarrow D\pi \rightarrow K\pi\pi$ using a Neural Network*, 2006. CDF internal note: 8123.
- [50] Bs Mixing group. *Same Side Kaon Tagging Studies in Fully Reconstructed Decays*, 2006. CDF public note: 8206.
- [51] G. Giurgiu, V. Tiwari, M. Paulini, and J. Russ. *Muon B Flavor Tagging - A Likelihood Approach*, 2005.
<http://www-cdf.fnal.gov/physics/new/bottom/072204.blessed-like-mu-tag/blessed-like-mu-tag.ps>.

- [52] Erich Lohrmann and Volker Blobel. *Statistische und numerische Methoden der Datenanalyse*. Teubner, Stuttgart, 1998.
- [53] G. Punzi. *Comments on likelihood fits with variable resolution*. arxiv: physics/0402045.
- [54] Isard Dunietz, Robert Fleischer, and Ulrich Nierste. *In pursuit of new physics with B_s decays*. *Phys. Rev.*, D63:114015, 2001.
- [55] F. James. *MINUIT. Function Minimization and Error Analysis*. <http://consult.cern.ch/writeup/minuit/>.
- [56] A. S. Dighe, T. Hurth, C. S. Kim, and T. Yoshikawa. *Measurement of the lifetime difference of B_d mesons: Possible and worthwhile?* *Nucl. Phys.*, B624:377–404, 2002.
- [57] B. Aubert et al. *Measurement of decay amplitudes of $B \rightarrow J/\psi K$, $B \rightarrow \psi(2S)K$ and $B \rightarrow \chi(c1)K$ with an angular analysis*. *Phys. Rev.*, D76:031102, 2007.
- [58] R. Itoh et al. *Studies of CP violation in $B \rightarrow J/\psi K^*$ decays*. *Phys. Rev. Lett.*, 95:091601, 2005.
- [59] D. Aston et al. *A Study of $K^- \pi^+$ Scattering in the Reaction $Kp \rightarrow K\pi^+n$ at 11 GeV/c*. *Nucl. Phys.*, B296:493, 1988.
- [60] The BABAR Collaboration: B. Aubert. *Ambiguity-Free Measurement of $\cos 2\beta$: Time-Integrated and Time-Dependent Angular Analyses of $B \rightarrow J/\psi K\pi$* . *Physical Review D*, 71:032005, 2005.
- [61] B. Aubert et al. *Measurement of the $B \rightarrow J/\psi K^*(892)$ decay amplitudes*. *Phys. Rev. Lett.*, 87:241801, 2001.
- [62] J. Boudreau, J. Cranshaw, L. Flores-Castillo, V. Papadimitriou, M. Spezziga, and K. Carrel. *Measurement of exclusive B lifetimes in the modes: $J\psi K^+$, $J\psi K^{*+}$, $J\psi K^{*0}$, and $J\psi K_s$* , 2004. CDF internal note: 6387.
- [63] Michael Feindt, Michal Kreps, Thomas Kuhr, and Michael Milnik. *Measurement of $\Delta\Gamma$ and ϕ_s in untagged $B_s \rightarrow J\psi\phi$ and $B_d \rightarrow J\psi K^*$ decays*, 2007. CDF internal note: 8753.
- [64] Gary J. Feldman and Robert D. Cousins. *A Unified approach to the classical statistical analysis of small signals*. *Phys. Rev.*, D57:3873–3889, 1998.

- [65] Robert Fleischer and Isard Dunietz. *CP violation and CKM phases from angular distributions for B_s decays into admixtures of CP eigenstates*. *Phys. Rev.*, D55:259–267, 1997.