Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

IEKP-KA/2007-03

Fitting Procedure for the Determination of the B_s Mixing Frequency

Jan Fraenkle

Anpassungsrechnung zur Messung der B_s-Oszillationsfrequenz

Diplomarbeit

an der Fakultät für Physik
der Universität Karlsruhe

Referent: Prof. Dr. M. Feindt
Korreferent: Prof. Dr. G. Quast

Institut für Experimentelle Kernphysik

Januar 2007

Deutsche Zusammenfassung

Einführung

In der heute etablierten Theorie des Standardmodells der Elementarteilchenphysik werden die Eigenschaften der Elementarteilchen und ihre Wechselwirkung untereinander beschrieben. Es gibt darin zwei verschiedene Arten von Elementarteilchen, Quarks und Leptonen. Quarks kommen in der Natur nicht als freie Teilchen vor, sondern nur als gebundene Zustände (Hadronen). Baryonen sind Teilchen aus drei Quarks, und Teilchen, die aus einem Quark und einem Antiquark bestehen, werden Mesonen genannt. Zu jedem Teilchen gibt es im Standardmodell ein Antiteilchen mit zwar entgegengesetzten Quantenzahlen, aber gleicher Masse wie das zugehörige Teilchen.
Einige Mesonen zeigen ein interessantes Verhalten, sie oszillieren, d.h. sie können sich in ihre Antiteilchen umwandeln und umgekehrt. Man geht davon aus, dass alle neutralen Mesonen, die nicht ihr eigenes Antiteilchen sind, oszillieren können. Es wird dadurch erkärt, dass die Masseneigenzustände nicht mit den Eigenzuständen der schwachen Wechselwirkung der Quarks übereinstimmen. Die Verknüpfung zwischen den Masseneigenzuständen mit den Eigenzuständen der schwachen Wechselwirkung ist durch die CKM-Matrix gegeben. Die Frequenz ist proportional zur Massendifferenz der Masseneigenzustände Δm. Eine Meson-Antimeson Oszillation wurde erstmals im Jahre 1956 im $K^0 - \bar{K}^0$ System [3] experimentell nachgewiesen. Die Massendifferenz im $D^0 - \bar{D}^0$ System wird als sehr klein erwartet und ist dadurch kaum zu messen, da die Teilchen zu schnell zerfallen. Die erste Evidenz für B-Meson-Oszillationen wurde 1987 von den Experimenten ARGUS [5] und UA1 [4] beobachtet. Die Oszillationsfrequenz von $B^0 - \bar{B}^0$ ist sehr präzise bestimmt und der Weltmittelwert liegt bei $\Delta m_d = 0.507 \pm 0.005\, ps^{-1}$ [24]. Die Frequenz der $B_s - \bar{B}_s$ Oszillationen wird deutlich größer vorhergesagt und die experimentell bestimmte untere Grenze lag 2005 bei $\Delta m_s > 16.6\, ps^{-1}$, was die direkte zeitaufgelöste Messung erschwert. Während dem Run II des CDF-Experiments sollten jedoch genügend Daten gesammelt werden können, um ausreichende Statistik für die direkte Messung der Oszillationsfrequenz von B_s-Mesonen zu haben. Diese Messung ist eines der Hauptziele des CDF-Experiments. In der Kollaboration und speziell in unserer Arbeitsgruppe wurde mittlerweile vieles an Vorarbeit geleistet, was die Messung der Oszillationsfrequenz möglich machen sollte. Darunter fällt eine optimierte Signalselektion durch die Neuronale Netztechnik, NeuroBayes®, sowie ein verbesserter Flavor-Tagging-Algorithmus, beruhend auf der gleichen Technik.
Das Ziel dieser Arbeit ist nun die Entwicklung einer Software zur Bestimmung der B_s Oszillationsfrequenz Δm_s durch Kombination der verfügbaren Ergebnissen mit Hilfe eines ungebinnten Maximum-Likelihood-Fits.

Das CDF-Experiment

Die verwendeten Daten in dieser Arbeit wurden mit Hilfe des CDF II-Detektors gesammelt. Der Detektor steht, wie auch das DØ Experiment, am Tevatron Beschleuniger im Fermi National Accelerator Laboratory (Fermilab) in Batavia bei Chicago (Illinois).

Das Tevatron ist ein kreisförmiger Proton-Antiproton Beschleuniger mit einem Radius von 1 km und einer Schwerpunktsenergie von 1,96 TeV. An einem der beiden Wechselwirkungspunkten befindet sich der CDF-Detektor. Er ist ein Mehrzweckteilchendetektor und findet für viele physikalische Fragestellungen Verwendung. Der Detektor, in dessen Zentrum der Wechselwirkungspunkt liegt, ist zylindersymmetrisch aufgebaut. Die innersten Komponenten sind die Spurdetektoren, der Siliziumvertexdetektor und die Spurkammer COT. Sie dienen zum Nachweis von Spuren geladener Teilchen. Zusammen mit dem supraleitenden Solenoidmagneten, der die Teilchen auf eine Kreisbahn ablenkt, kann man den Transversalimpuls bestimmen. Zwischen den Spurkammern und dem Solenoidmagenten befindet sich der zur Teilchenidentifikation verwendete Flugzeitmesser. Außerhalb des Magenten befinden sich die elektromagnetischen und hadronischen Kalorimeter, die die Energien von Teilchen, bzw. Jets bei hohen Energien messen. Das Myonsystem des Detektors umgibt das Ganze mit seinen Spurkammern und Szintillatoren zum Nachweis von Myonen. Durch hohe Wechselwirkungsraten bei CDF II ist es unmöglich, jedoch auch nicht wünschenswert, alle resultierenden Ereignisse zu speichern. Um das hohe Datenaufkommen zu reduzieren, ist ein Filter notwendig, was durch einen mehrstufigen Trigger realisiert ist. Das komplexe System selektiert die Ereignisse und ist auf Hardware- und Softwareebene implementiert. Die für diese Arbeit relevanten Daten wurden mit dem sogenannten Zweispurtrigger vorselektiert.

Ereignisrekonstruktion und Signalselektion

Nach der Vorselektion durch den Zweispurtrigger und dem Speichern der Daten wird versucht aus den detektierten Zerfallsprodukten das jeweilige eventuell zugrunde liegende B-Meson zu rekonstruieren. Die Rekonstruktion eines B_s-Zerfalls in einem exklusiven Endzustand basiert auf der Spurrekonstruktion der Zerfallsprodukte, wobei man dazu sukzessive die Zerfallvertizes aller Zwischenzustände rekonstruiert. Es wird also im Falle von $B_s \to D_s\pi$, $D_s \to \phi\pi$ damit begonnen, den Zerfall $D_s \to \phi\pi$ zu rekonstruieren, indem die Kombination der Spuren durch einen Vertexfit auf die Hypothese überprüft wird, ob die Spuren einen gemeinsamen Ursprung haben und ob die invariante Masse der Spurkombination der Ruhemasse des D_s-Mesons entspricht. Als nächster Schritt wird überprüft, ob die D_s-Meson-Kandidaten zusammen mit einer weiteren Spur mit dem Zerfall eines B_s-Mesons kompatibel sind.

Trotz der Vorselektion durch den Trigger und der Verwerfung einiger Zerfälle als Nicht-B_s-Zerfälle durch den Rekonstruktionsalgorithmus, erfüllen viele Ereignisse alle Bedingungen und werden als Signal behandelt, obwohl sie einen anderen Ursprung als einen B_s-Zerfall haben. Einige davon erfüllen durch zufällige Kombinationen von Spuren die Bedingungen bei anderen wurde eine falsche Teilchenhypothese angenommen, d.h. Nicht-B_s-Teilchen werden als solche interpretiert. Nach der B_s-Rekonstruktion enthält der Datensatz immer noch hauptsächlich Untergrundereignisse, richtig rekonstruierte B_s-Zerfälle machen nur einen relativ kleinen Teil aus. Durch eine Signalselektion mit Hilfe von NeuroBayes können weitere Untergrundereignisse verworfen werden, was zur Verbesserung der Signalsignifikanz $\frac{n_S}{\sqrt{n_S+n_B}}$[1] führt. Da die Signalsignifikanz proportional zur Signifikanz der Messung der B_s-Oszillationsfrequenz ist, hat die Signalselektion einen direkten Einfluss auf das Ergebniss.

Bestimmung der Oszillationsfrequenz Δm_s

Die Bestimmung der B_s-Oszillationsfrequenz Δm_s erfordert die Kombination einiger relevanter Ereignisinformationen in einem ungebinnten Maximum-Likelihood-Fit mit Δm_s als Fitparameter. Die Methode erfordert die Beschreibung der Verteilungen der invarianten Masse m und der Zerfallsdauern t von Ereignissen, die im Datensatz enthalten sind. Darüber hinaus muss die Information, ob sich ein Teilchen umgewandelt hat ($\xi = -1$) oder nicht ($\xi = 1$), durch den Taggingalgorithmus ermittelt werden und für jedes Teilchen zur Verfügung gestellt werden.

[1] n_S, n_B, Anzahl der Signal-, bzw. der Untergrundereignissen

Die Wahrscheinlichkeitsdichte $P(m, t, \xi, \sigma_t)$ berücksichtigt diese Ereignisinformationen und ist gegeben durch

$$P(m, t, \sigma_t, \xi) = (1 - f_B) \cdot P_S(m) \cdot P_S(t, \sigma_t, \xi) + f_B \cdot P_B(m) \cdot P_B(t, \xi)$$

Der Parameter f_B beschreibt den Anteil von Untergrundeignissen von allen Ereignissen im Datensatz. $P_B(t)$ ist die Parametrisierung der Lebensdauerverteilung der Untergrundereignisse. Der Teil $(1 - f_B) \cdot P_S(m) + f_B \cdot P_B(m)$ parametrisiert das invariante Massenspektrum der B_s-Kandidaten. Der Ausdruck $P_S(t, \sigma_t, \xi)$ beschreibt die Oszillation und die Zerfallsdauerverteilung von Signalereignissen und die Wahrscheinlichkeitsdichte ist gegeben durch

$$P_S(t, \xi) = \frac{1}{N_S(t, \xi, \sigma_t)} \left(\frac{1 + \xi D \cos(\Delta m t)}{1 + |\xi|} \frac{1}{\tau} e^{\frac{t}{\tau}} \right) \otimes G(t - t', \sigma_t) \cdot \epsilon(t)$$

wobei $N_S(t, \xi, \sigma_t)$ für die Normierung verantwortlich ist, D (Dilution) ist ein Maß für die Sicherheit der Taggingentscheidung, t ist die gemessene Zerfallszeit und σ_t die dazugehörige Auflösung. Die Faltung des Oszillations- und Zerfallsterms mit einer Gauß-Funktion beschreibt die endliche Auflösung der Zerfallsdauermessung. Die Multiplikation mit der Effizienzfunktion $\epsilon(t)$ wird der eingeschränkten Akzeptanz des Detektors gerecht. Durch eine Parameterschätzung mit Hilfe der Maximum-Likelihood-Methode wird die Lebensdauer τ oder die Oszillationsfrequenz Δm_s bestimmt. Die komplette Likelihood-Funktion hat allerdings pro Zerfallskanal mehr als 50 Parameter, die alle bestimmt werden müssen, was durch die Endlichkeit der Anzahl der verfügbaren Ereignissen nicht in einem Fit geht. Zur Bestimmung der meisten dieser Parameter werden vor dem ungebinnten Maximum-Likelihood-Fit einige Verteilungen, wie das invariante Massenspektrum, die Effizienzverteilung und Lebensdauerverteilungen von Untergründen mittels gebinnten Fits parametrisiert. Zur Verbesserung der Signifikanz des Ergebnisses werden zur Analyse alle Zerfallskanäle parallel heran gezogen. Das Framework bietet die Möglichkeit die meisten dieser Parameter mittels gebinnter Fits für jeden Zerfallskanal einzeln zu bestimmen und dann die Parameter für den ungebinnten Fit zusammen zu führen und zu bestimmen welcher Parameter für jeden Kanal individuell oder in allen Kanälen gleich sein soll.

Im Falle limitierter Statistik, d.h. bei einer kleinen verfügbaren Menge an Ereignissen, die zur Bestimmung der Oszillationsfrequenz herangezogen werden können, wird der ungebinnte Maximum-Likelihood-Fit für Δm_s im Allgemeinen nicht konvergieren und es ist schwierig, Grenzen für Δm_s anzugeben. Anstatt einer direkten Bestimmung der Frequenz, die als freier Fitparameter in der Likelihood-Funktion vorkommt, gibt es einen alternativen Ansatz. Der Amplitudenscan basiert auf der Idee einer Fourier-Transformation der Signaloszillationen. Dafür muss im Oszillationsterm die Amplitude A als zusätzlicher Parameter eingeführt werden.

$$\frac{1 + \xi D \mathbf{A} \cos(\Delta m t)}{1 + |\xi|}$$

Beim Amplitudenscan ist dann die Amplitude A ein freier Fitparameter, während Δm_s fest ist. Der Fit wird jedoch für verschiedene Werte von Δm_s wiederholt und jeweils die Amplitude A bestimmt. Für den wahren Wert von Δm_s ist der erwartete Wert für A gleich 1 und $A = 0$ für alle anderen Werte der Oszillationsfrequenz.

Der Amplitudenscan (Abbildung 1) ist ein vorläufiges Ergebnis dieser Analyse. Er zeigt eine Evidenz eines Oszillationssignals bei $\Delta m_s \approx 18\ ps^{-1}$, da die Amplitude A mit 1 kompatibel ist und 0 mit mehr als 3σ ausschließt. Die Sensitivität diser Analyse liegt bei 24.8 ps^{-1}. Für diesen Amplitudenscan wurden nur die Ereignisse aus einem von sechs Zerfallskanälen heran gezogen. Der Kanal $B_s \to D_s \pi$, $D_s \to \phi \pi$ ist der wichtigste, denn er enthält ca. 40% der Gesamtstatistik und hat die beste Signalsignifikanz. Das Ergebnis dieser vorläufigen Analyse ergab eine B_s-Oszillationsfrequenz von

$$\Delta m_s = 18.32 \pm 0.13\ ps^{-1}$$

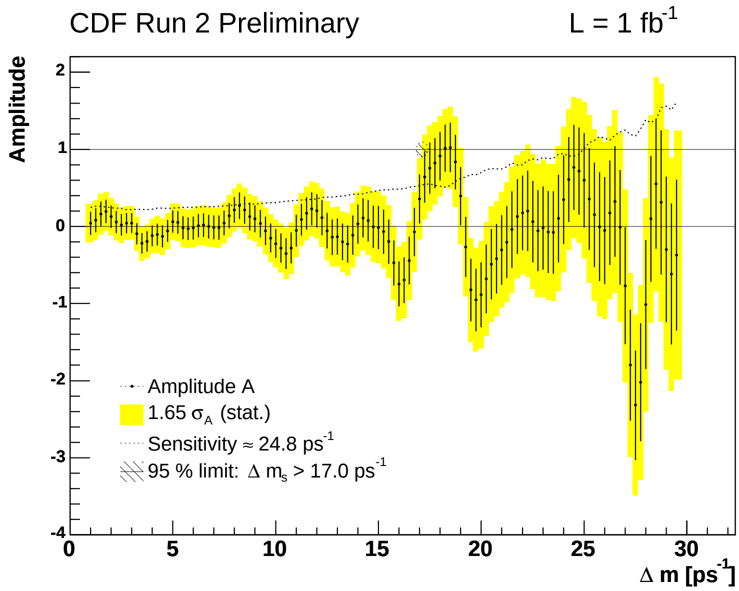

Figure 1: Amplitudenscan des B_s-Oszillationssignals im Zerfallskanal $B_s \rightarrow D_s\pi$, $D_s \rightarrow \phi\pi$.

Zusammenfassung und Ausblick

Die Entwicklung dieses Frameworks zur Ermittlung der B_s-Oszillationsfrequenz ist ein wichtiger Schritt, um angemessen von vergangenen Arbeiten in unserer Gruppe zu profitieren. Das beinhaltet die Optimierung der Signalselektion in den exklusiven Zerfallskanälen des B_s-Mesons, sowie die Entwicklung eines kombinierten Tagging-Algorithmus mit der doppelten Tagging-Power des bisherigen. Zusammen mit den jetzt verfügbaren Daten, die einer integrierten Luminosität von $1\ fb^{-1}$ entsprechen, erlauben die Methoden eine direkte zeitabhängige Messung der Oszillationsfrquenz Δm_s.

Bisher gibt es noch kein Ergebnis von unserer Gruppe mit einem optimal arbeitenden Flavor-Tagger und mit der vollen Statistik aus allen sechs exklusiven Kanälen. Lediglich ein Amplitudenscan mit den Daten aus dem Zerfallskanal $B_s \rightarrow D_s\pi$, $D_s \rightarrow \phi\pi$ ist in Abbildung 1 gezeigt. Eine Evidenz für ein Oszillationssignal bei $\Delta m_s = 18.32 \pm 0.13\ ps^{-1}$ ist durch die Konsistenz der Amplitude A mit 1 und mit mehr als 3σ Entfernung von 0 gegeben. Die Sensitivität der Messung liegt bei $24.8\ ps^{-1}$.

Vor kurzem veröffentlichte die CDF-Kollaboration eine direkte Messung von Δm_s. Das Ergebnis dieser Analyse ist mit $\Delta m_s = 17.77 \pm 0.10(stat) \pm 0.07(syst) ps^{-1}$ [34] sehr präzise und gilt mit einer Signifikanz von mehr als 5σ als sicher. Die Wahrscheinlichkeit, dass eine Untergrundfluktuation das Signal hervorrief, beträgt nur $8 \cdot 10^{-8}$.

Das Ziel in der Zukunft wird sein, die Signifikanz der aktuellen CDF-Messung zu verbessern und die Oszillationsfrequenz der $B_s - \bar{B}_s$-Oszillationen noch etwas genauer zu bestimmen. Eine weitere Anwendung ist die Kalibrierung des neuen Tagging-Algorithmus nach der Oszillationsmessung, da dieser auch für die Messung der CP-Verletzung im B_s-System notwendig ist.

Contents

List of Figures

Introduction

The whole universe as we know is made of only a few different elementary particles. These particles themselves and combinations of them are manifested as matter. But besides matter, also antimatter is existing. It was suggested by P. Dirac in 1928 by the development of his relativistic, quantum mechanical equation for electrons [1], known as the Dirac equation. The experimental confirmation was done by the discovery of the positron by C.D. Anderson in the year 1932 [2]. Not only the electron has an antimatter partner, but all elementary particles have their corresponding antiparticles. The quantum numbers of antiparticles are opposite to the quantum numbers of matter particles, but they have the same mass. A particle and its antiparticle will annihilate, e.g. into two photons when they meet.
The currently known elementary particles and their interactions are described in the Standard Model of particle physics. Quarks and leptons are divided into three generations and virtually all matter in our world is build up from two of the light quarks and electrons. The particles belonging to the other generations are not stable and existed only in a short time after the big bang or are artificially produced in particle accelerators.
Quarks do not exist as free particles, but build up bound states, called hadrons. Baryons like protons or neutrons are particles made up of three quarks. Particles made up of a quark and an antiquark are called mesons. Some of these mesons show an interesting behavior, they oscillate, i.e. they transform themselves into their own antiparticle. This is explained in the Standard Model by the mass difference Δm of the two mass eigenstates and by the fact that the flavor eigenstates are a linear combination of the mass eigenstates. Particle-antiparticle oscillations are expected for all neutral mesons being not their own antiparticles. The mass difference Δm is proportional to the mixing frequency and can be determined by a time resolved measurement, it is a very important and nontrivial test of the well established Standard Model. The first observation of particle-antiparticle oscillations was in the $K^0 - \bar{K}^0$ system in 1956 [3]. The mass difference in the $D^0 - \bar{D}^0$ system is predicted to be very small, i.e. the oscillation frequency is small compared to the lifetime and thus hard to measure. In 1987 the first evidence for neutral B meson mixing was observed by two experiments namely ARGUS [5] and UA1 [4]. Since then the oscillation frequency of $B^0 - \bar{B}^0$ is measured very precisely and the world average is $\Delta m_d = 0.507 \pm 0.005\ ps^{-1}$ [24]. In case of $B_s - \bar{B}_s$, ($s\bar{b} - \bar{s}b$) oscillation, the oscillation frequency is predicted much higher. The experimental expectation for Δm_s before the CDF measurements in 2005 was $\Delta m_s = 18.3^{+6.5}_{-1.5}\ ps^{-1}$and the experimental lower limit of Δm_s from 2005 is $\Delta m_s > 16.6\ ps^{-1}$. So the time resolved analysis and direct measurement of Δm_s is more difficult.
During Run II of the CDF experiment it should be possible to collect enough data for the Δm_s measurement and was thus one of the main aims of the experiment. A lot of work is done in the meantime to improve the B_s selection with the help of neural networks [18, 19, 20] and the development of a powerful tagging algorithm [21]. Final measurement of Δm_s is realized by an unbinned maximum likelihood fit in mass, lifetime and flavor tagging space of B_s candidates. The content of this diploma thesis is the development and implementation of parts of a framework doing the measurement.

This thesis starts with the theoretical introduction in chapter 1 which gives a review over the relations of CKM matrix, unitarity triangle and meson-antimeson oscillation. Chapter 2 presents the CDF experiment placed at Fermilab in Batavia near Chicago (Illinois). The operating mode of the accelerator Tevatron and the CDF detector is introduced. The analysis outline in chapter 3 contains necessary informations about the input to the fitter framework. The main chapter 4 presents everything about the composition of the fitter framework and the procedure inside. Chapter 5 deals with the extraction of the mixing frequency Δm_s and presents the result obtained until now. Finally chapter 6 gives a conclusion and an outlook.

Chapter 1

Theoretical Overview

1.1 The Standard Model

The Standard Model of particle physics is a theory describing fundamental particles and their interactions.

The world consists of three families of leptons and quarks. Altogether there are six different flavors of leptons and six different flavors of quarks. They all have spin $\frac{1}{2}$ and thus are fermions.

There are three different types of interactions between the particles of the Standard Model. Each interaction is mediated by an exchange particle called gauge boson. The exchange particle of the electromagnetic interaction is γ, the ones of the weak interaction are $W^{\pm}$ and Z^0 and the ones of the strong interaction are eight different gluons.

The quarks take part in the strong and electromagnetic interaction but also in the weak interaction. The coupling of the charged $W^{\pm}$ bosons connects fermions of different families. So quarks with the flavor up, charm and top are able to transmute in a quark with the flavor down, strange or bottom under emission of a W^{+} boson. The process is described by the CKM-matrix where the square of the absolute value of the matrix elements is the transition probability.

1.2 The CKM-Matrix and Unitarity Triangle

The Cabibbo-Kobayashi-Maskawa-Matrix (CKM-Matrix) [7, 8] is a unitary matrix describing the transformation between quark mass eigenstates and flavor eigenstates. In its most general form it can be written as

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

This is the most general notation of the CKM-Matrix with three quark generations. V_{xx} are complex numbers corresponding to 18 free real parameters. The unitarity requirement $V_{CKM}V_{CKM}^{\dagger} = 1$ reduces 9 of the 18 free parameters. 5 are eliminated through 5 unobservable arbitrary complex phases [9]. The remaining 4 parameters can be seen as 3 rotation angles and one complex phase which is a implication for CP-violation.

A convenient parameterization of the CKM-matrix is the Wolfenstein parameterization [10]. The CKM-matrix is expanded in a small parameter $\lambda \approx 0.22$. The four free parameters are λ, A, ρ, and η and can be interpreted geometrically.

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

Some of the Wolfenstein parameters ($A = 0.82 \pm 0.04$, $\lambda = 0.223 \pm 0.002$) are determined quite well by measurements of some CKM matrix elements in different experiments [11, 12, 13]. The complex phase is described by parameters η and ρ. The special case $\eta = 0$ denotes a real CKM-matrix and no CP-violation.

The unitarity constraints of the CKM-matrix give a rise to the equation

$$V_{CKM}V_{CKM}^{\dagger} = V_{CKM}^{\dagger}V_{CKM} = 1$$

Subsequently follows the constraints $V_{ik}^{*}V_{ij} = \delta_{kj}$ which can be applied on different rows and columns. The most interesting combinations are the following equations

$$V_{ub}^{*}V_{ud} + V_{cb}^{*}V_{cd} + V_{tb}^{*}V_{td} = 0 \tag{1.1}$$

$$V_{td}^{*}V_{ud} + V_{ts}^{*}V_{cs} + V_{tb}^{*}V_{cb} = 0 \tag{1.2}$$

These equations can be interpreted geometrically as a triangle in the complex plane. All sides of the two unitarity triangles described by equations 1.1 and 1.2 are of the same order of magnitude.

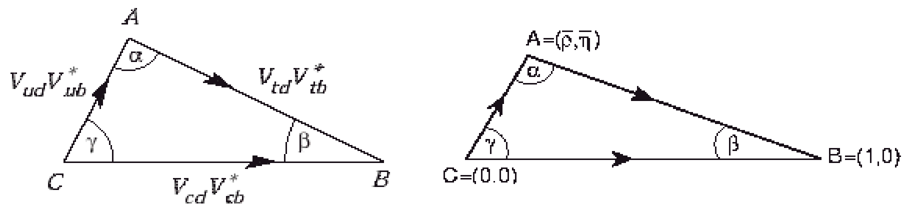

Figure 1.1: The unitarity triangle in a standard form (left) and in Wolfenstein parameterization rescaled by $1/V_{cd}V_{cb}^{*}$ (right).

The sides of the unitarity triangle (figure 1.1) can also be described by Wolfenstein parameters. The triangle is rescaled by $1/V_{cd}V_{cb}^{*}$ so that the coordinates of the unitarity triangle result in $(0,0)$, $(1,0)$ and $(\bar{\rho}, \bar{\eta})$, where ($\bar{\rho}$ and $\bar{\eta}$) can be written in Wolfenstein parameters:

$$\bar{\rho} = \left(1-\lambda^2\right)\rho$$

$$\bar{\eta} = \left(1-\lambda^2\right)\eta$$

The determination of all properties of the unitarity triangle can be done experimentally. It is one of the most important issues in particle physics to be solved in recent years. Behind the properties of the unitarity triangle are different factors describing the Standard Model so that the constraining of the unitarity triangle is a mayor test of the Standard Model. Finding out experimentally that the unitarity triangle is not closed would be a hint for new physics beyond the Standard Model. The experimental status of the unitarity triangle is shown in figure 1.2.

1.3 Neutral B-Meson Mixing

Neutral mesons are able to pass into their own antiparticles. In the Standard Model it is connected to the fact that the mass eigenstates and the eigenstates of weak interaction are not the same but

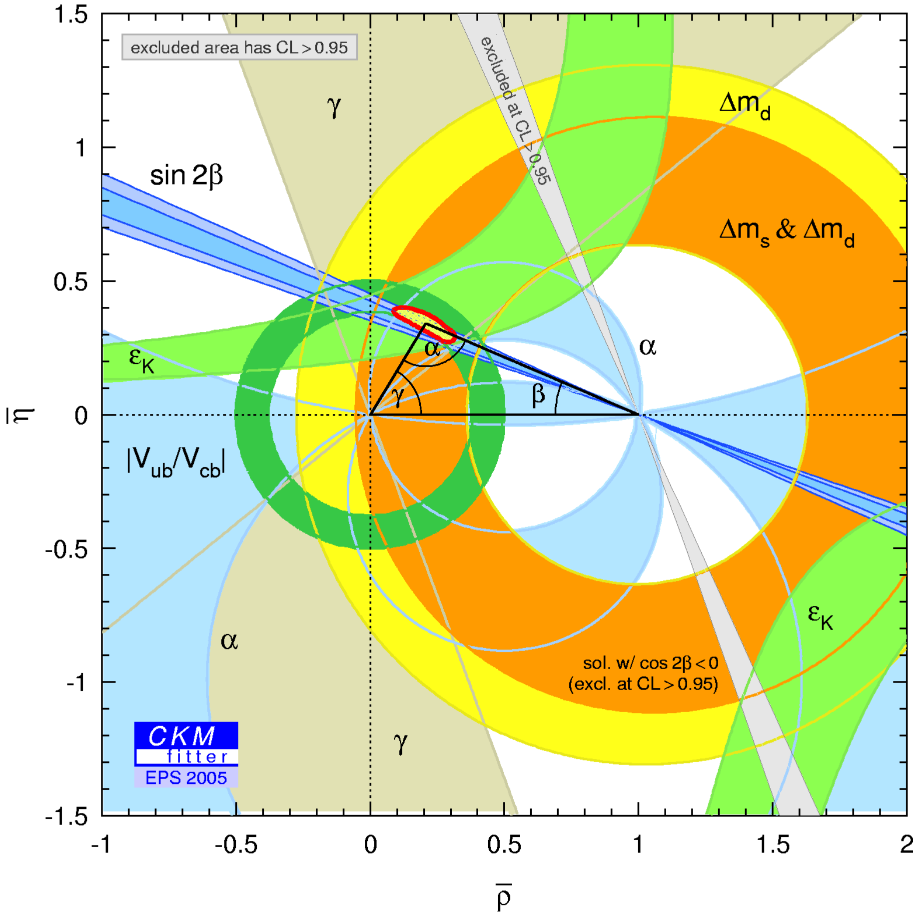

Figure 1.2: Experimental status of constraints on the parameters describing the unitarity triangle without actual CDF measurement of Δm_s [11].

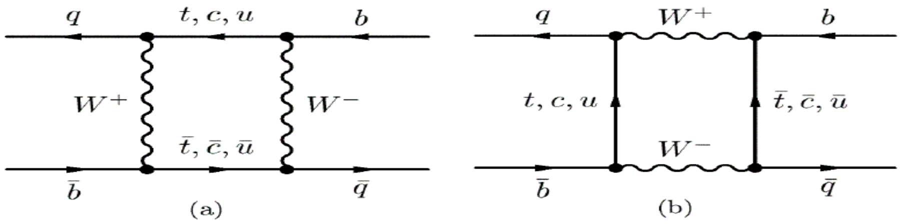

Figure 1.3: Lowest order Feynman diagrams for $p_q \rightarrow \bar{p}\bar{q}$ show the interactions during mixing process.

the eigenstates of weak interaction can be written as a linear combination of the mass eigenstates. Particle-antiparticle mixing is expected for all neutral mesons which are not their own antiparticle, i.e. $K^0\bar{K}^0$, $D^0\bar{D}^0$, $B^0_d\bar{B}^0_d$ and $B^0_s\bar{B}^0_s$. It is observed in all these cases except $\bar{D}^0$.

The particle-antiparticle transformation in first order is shown in two lowest order Feynman diagrams for $p_q \rightarrow \bar{p}\bar{q}$ transitions (figure 1.3). Δm_q can be calculated in first order from these diagrams. The index q represent either a d or a s quark depend on having B^0_s or B^0_d mixing.

$$\Delta m_q = \frac{G_F^2}{6\pi^2}\eta_B m_{B_q} f^2_{B_q} m^2_W S(x_t)|V^*_{tb}V_{tq}|$$

The two matrix elements $V^*_{tb}V_{tq}$ can be written in Wolfenstein parameters showing that a constraint on Δm_q describes approximately a circle around (1,0) in the $(\bar{\rho}, \bar{\eta})$ plane.

$$V^*_{tb}V_{tq} = \lambda^6 A^2[(1-\bar{\rho})^2 + \bar{\eta}^2] + \mathcal{O}(\lambda^{10})$$

It is useful to determine the fraction of Δm_s and Δm_d because the ratio of the prefactors can be calculated more precisely than separately.

$$\frac{\Delta m_s}{\Delta m_d} = \xi \frac{m_{B_s}}{m_{B^0}} \left|\frac{V_{ts}}{V_{td}}\right|^2$$

After calculating Δm_q out of the Feynman diagram and connecting with CKM-matrix elements, the time dependence of the mixing has to be introduced. Beginning with the flavor eigenstates of B^0_s-meson:

$$|B^0_q\rangle = \frac{1}{\sqrt{2}}\left(|B_{H,q}\rangle + |B_{L,q}\rangle\right)$$

$$|\bar{B}^0_q\rangle = \frac{1}{\sqrt{2}}\left(|B_{H,q}\rangle - |B_{L,q}\rangle\right)$$

$|B^0_q\rangle$ and $|\bar{B}^0_q\rangle$ are the flavor eigenstates and $|B_{H,q}\rangle$ is the heavy and $|B_{L,q}\rangle$ the light mass eigenstate.

To know what happens at the time $t > 0$, the time evolution of the eigenstates has to be done. A non-relativistic quantum mechanical description is given by the time dependent Schrödinger equation

$$H|B\rangle = i\frac{\partial}{\partial t}|B\rangle.$$

H is given by:

$$H = M + \frac{i}{2}\Gamma = \begin{pmatrix} M_{11} + \frac{i}{2}\Gamma_{11} & M_{12} + \frac{i}{2}\Gamma_{12} \\ M^*_{12} + \frac{i}{2}\Gamma^*_{12} & M^*_{22} + \frac{i}{2}\Gamma^*_{22} \end{pmatrix}.$$

M_{xx} stands for mass and Γ_{xx} for decay width.
The time evolution of the flavor eigenstates can be written as:

$$|B_q^0(t)\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\left(m_1-\frac{i}{2}\Gamma_1\right)t}|B_{L,q}\rangle + e^{-i\left(m_2-\frac{i}{2}\Gamma_2\right)t}|B_{H,q}\rangle\right)$$

$$|\bar{B}_q^0(t)\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\left(m_1-\frac{i}{2}\Gamma_1\right)t}|B_{L,q}\rangle - e^{-i\left(m_2-\frac{i}{2}\Gamma_2\right)t}|B_{H,q}\rangle\right)$$

The probability for an initially pure B_q^0 ($\bar{B}_q^0$) to decay as $\bar{B}_q^0$ (B_q^0) is

$$P = |\langle B_q^0|\bar{B}_q^0(t)\rangle|^2 = |\langle \bar{B}_q^0|B_q^0(t)\rangle|^2 = \frac{1}{4}\left(e^{-\Gamma_1 t} + e^{-\Gamma_2 t} + 2e^{-\bar{\Gamma}t}cos(\Delta mt)\right)$$

with $\bar{\Gamma} = \frac{\Gamma_1+\Gamma_2}{2}$, $\Delta\Gamma = \Gamma_2 - \Gamma_1$ and $\Delta m = m_2 - m_1$. The two flavor eigenstates are superpositions of the two mass eigenstates (short- and long-lived components) having different decay widths Γ_1 and Γ_2. In case of B_s^0 the experimental result is given by $\Delta\Gamma/\bar{\Gamma} = 0.31^{+0.11}_{-0.13}$ [24]. The ratio $\Delta\Gamma/\bar{\Gamma}$ is in case of B_d^0 significantly smaller and can be neglected in the B_d^0-mixing analysis [14].

In case of B_s, the mixing probability is given by

$$P(t) = \frac{1}{2}e^{-\Gamma_s t}\left(cosh\left(\frac{\Delta\Gamma_s t}{2}\right) - cos(\Delta m_s t)\right) \tag{1.3}$$

At the beginning of this work, in January 2006, the B_s mixing frequency was not determined in a direct time dependent measurement yet. There is a world limit $\Delta m_s > 16.6\ ps^{-1}$ at 95% confidence level [35] resulting from combination of different measurements. This combination includes contributions from ALEPH, DELPHI, OPAL, SLD and CDF. The included CDF result from October 2005 based upon 355 pb^{-1} reports a limit of $\Delta m_s > 8.6\ ps^{-1}$ at 95% confidence level [36]. The experimental status of the unitarity triangle based on these measurements is shown in figure 1.2.

The intention of this work is the development of a framework for a time dependent measurement of the B_s mixing frequency Δm_s. Using 1 fb^{-1} of data, an efficient signal selection [18, 20] and an efficient b flavor tagger [21] promise a precise measurement of Δm_s with high significance.

Chapter 2

The CDF Experiment

2.1 The Tevatron

The Tevatron is a proton-antiproton collider with the world's highest center of mass energy until LHC will start in 2007. It is located at the Fermi National Accelerator Laboratory (Fermilab) in Batavia/Illinois. The main accelerator ring has a radius of 1 km and accelerates protons and antiprotons in opposite direction up to 980 GeV beam energy. The collisions take place at two interaction points where the experiments DØ and CDF are placed. The data used in this work are collected with the CDF II detector.

2.1.1 The Accelerator Chain

The accelerator chain is divided in different preaccelerating stages and the final acceleration in the main ring, the Tevatron. A schematic view can be seen in figure 2.1.

In the first step hydrogen gas is ionized to H^- and accelerated to a kinetic energy of 750 keV inside the Cockroft-Walton device. A linear accelerator (Linac) brings the hydrogen ions subsequently up to a kinetic energy of 400 MeV. This happens in a distance of 150 m using oscillating RF fields. At the end of Linac ions are grouped into bunches. At this point the hydrogen is focused on a carbon foil which wipes off the electrons, that only bare protons are left. After accelerating protons in a synchrotron, called Booster, to the kinetic energy of 8 GeV, the protons are injected into the last stage of preacceleration, the Main Injector, a 3 km circumference synchrotron. It is used for two important tasks, on the one hand for accelerating the protons up to 150 GeV and on the other hand for creating antiprotons by directing 120 GeV protons on a nickel target. During that procedure $\sim$20 antiprotons per 10^6 protons are produced. To separate these antiprotons from the background consisting mainly of protons, pions and neutrons, a pulsed magnet and a lithium lens are used to focus the beam in the Debuncher. These antiprotons have an average kinetic energy of about 8 GeV with a wide spread around this value. So they have to be cooled down with stochastic cooling to reduce this spread in the energy spectrum of the antiprotons in the Accumulator Synchrotron which also stores the antiprotons.

If there are enough antiprotons available, protons and antiprotons are accelerated up to the energy of 150 GeV in the main injector and are sent into the Tevatron in opposite directions. In the main accelerator ring both beams are accelerated up to a kinetic energy of 0.98 TeV, so that the maximal accomplishable center of mass energy is 1.96 TeV. After a period of collisions, called store, the remaining antiprotons are stored and cooled down in the Recycler, to increase the number of available antiprotons $\bar{p}$ for the next store. This is a useful issue because $\sim$ 75% of the antiprotons survive a store and the production rate is a limiting factor of the luminosity.

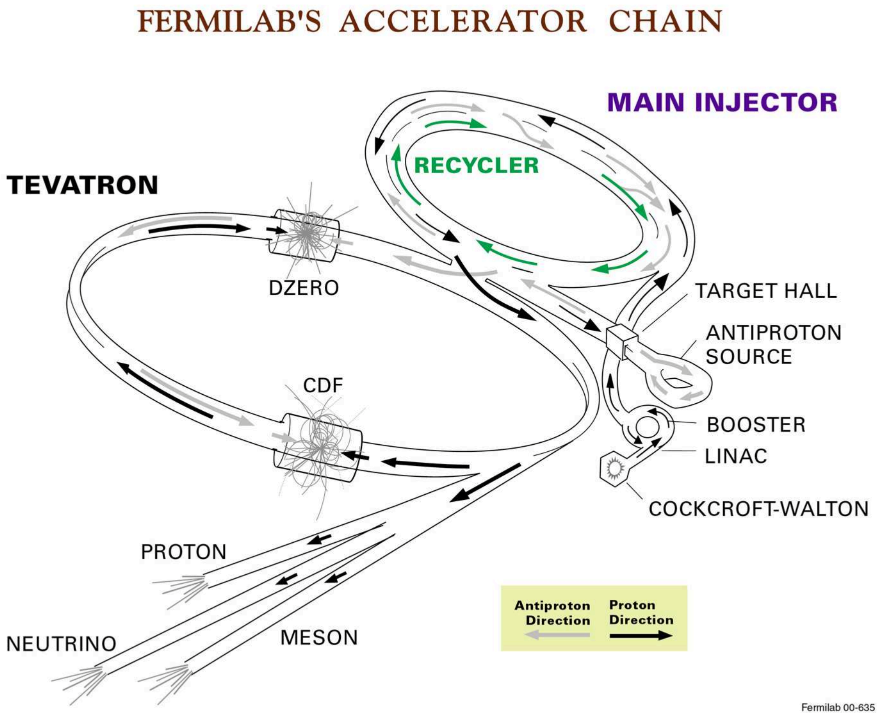

Figure 2.1: The Fermilab accelerator complex

2.2 The CDF II Detector

The CDF II detector (Collider Detector at Fermilab II) is a multi-purpose collider detector [15]. It is designed to detect and measure properties of particles generated during $p\bar{p}$-collisions.

It consists of a vertexing and tracking system, particle identification, a superconducting solenoid, calorimeters and muons chambers. These units are arranged in layers cylindrically symmetric with respect to the beamline as it can be seen in figure 2.2.

2.2.1 The Coordinate System

To simplify the description of the detector, a general right-handed coordinate system is introduced. The point of origin of that coordinate system is lying in the center of the detector. Two descriptions of the coordinate system are common. It can be described through Cartesian and polar coordinates. The positive z direction is along the beam line in the proton direction, the y-axis points vertically upward and the x-axis points radially outwards in the horizontal plane.
The polar coordinates are ϕ called azimuth angle, measured from the plane defined by the Tevatron and is lying in the x-y-plane and θ called polar angle which is measured from positive z-axis, it is lying in the x-z-plane.

η is another important quantity used instead of θ itself, it is defined by $\eta = -ln(tan(\theta/2))$ and is called Pseudorapidity.

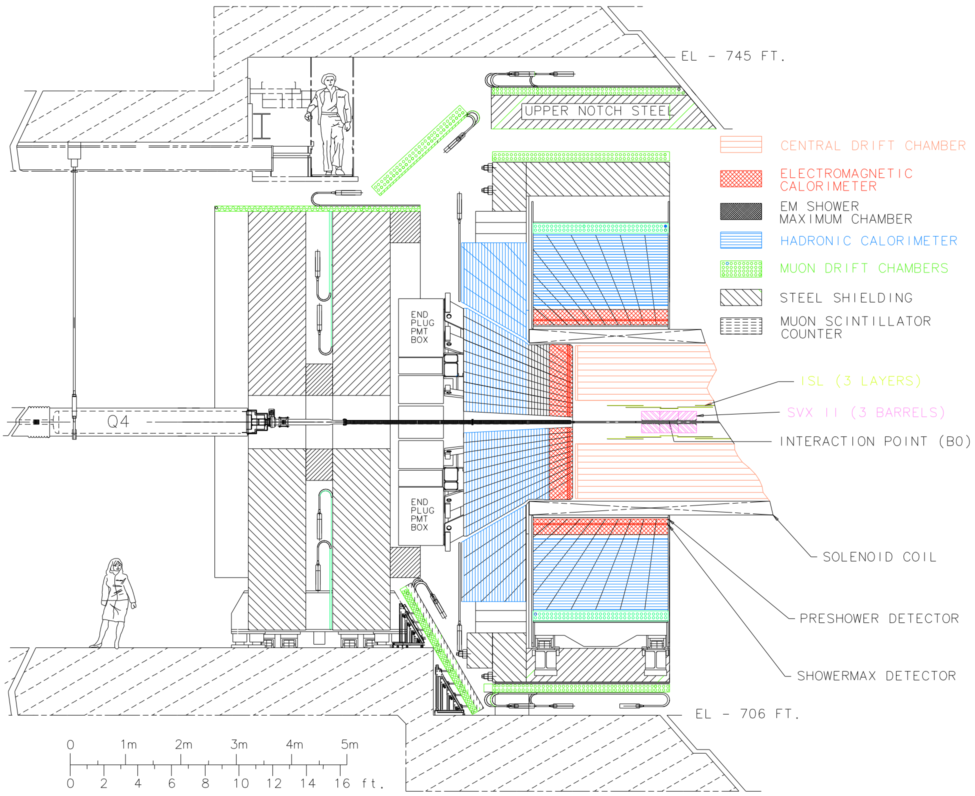

Figure 2.2: Elevation view of one half of the CDF II detector

2.2.2 The Tracking System

The tracking system is in principle made up of two different trackers. They all have in common that they can be used to measure momenta and displacement with respect to the primary vertex of charged particles. The main difference between them is the mode of operation. One part of the tracking system is the silicon tracker consisting of the Layer00 (L00), the Intermediate Silicon Layer (ISL) and the Silicon Vertex Detector (SVX II). The other part is the Central Outer Tracker (COT), a cylindrical drift chamber. The complete tracking system is inside the 1.4 T magnetic field generated by the superconducting solenoid essential for measuring the momenta of the particles. The tracking system together with the angle coverage is schematically shown in figure 2.3.

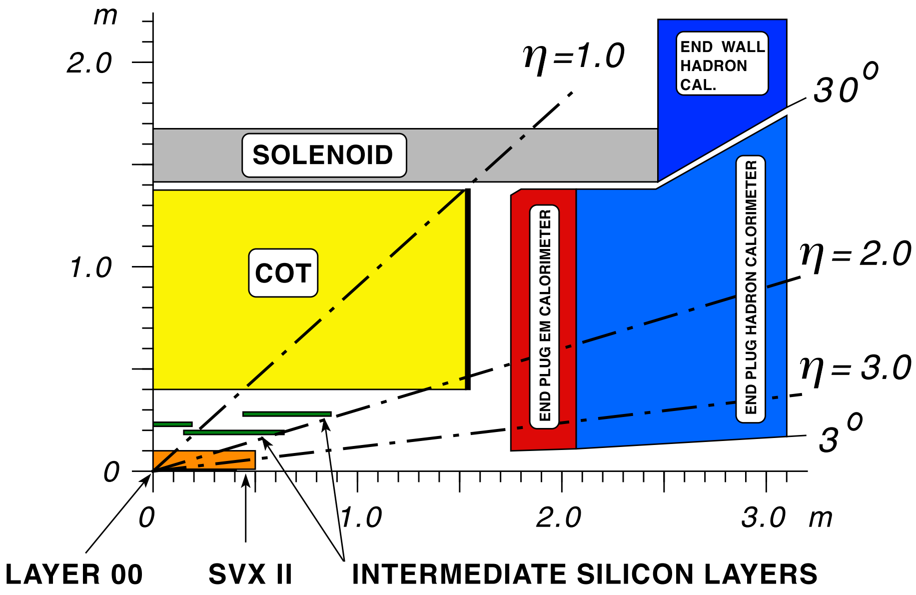

Figure 2.3: Schematic view of the CDF II tracking system

The complete construction of the silicon tracking system of the detector is shown in figure 2.4. The detector component closest to the beam line is L00 with a radial range from r=1.6 cm to r=2.1 cm. It consists of one-sided silicon strip detectors and is mounted directly on the beam pipe. Outside L00 follows SVX II with a radial range from r=2.45 to r=10.6 cm and an acceptance up to $|\eta| < 2.0$. It consists of three barrels arranged along the z-axis. Each of them has 5 layers made up of 12 angular segments. Every layer is made up of double-sided silicon strip detectors which allow a three dimensional track reconstruction because they are rotated by a certain angle. Between SVX II and COT another part of the silicon tracking system is located. It helps in linking the tracks measured in both detectors. This part called ISL (Intermediate Silicon Layer) consists of two components. The central layer is located at r=22 cm and covers $|\eta| < 1.0$. The forward/backward layers are located at r=20 cm and at r=28 cm and cover the region $1.0 < \eta < 2.0$.

The complete silicon tracking system device is especially useful to measure the impact parameter d_0 and the azimuthal angle ϕ very precisely. Together with the location at small radii high precise measurements of the secondary vertices of long-lived particles are possible.

The second part of the the tracking system is the COT working in a completely different mode of operation. It is a cylindrical drift chamber covering the radial range from r=33 cm up to r=143 cm and has an optimal acceptance in the range $|\eta| < 1$. The COT is made up of 96 wire layers grouped into 8 superlayers. 4 of the 8 superlayers are axial i.e. the wires run in z-direction, the other 4 superlayers are stereo layers with wires tilted by an angle of 2 degrees which allows measuring the z-coordinate.

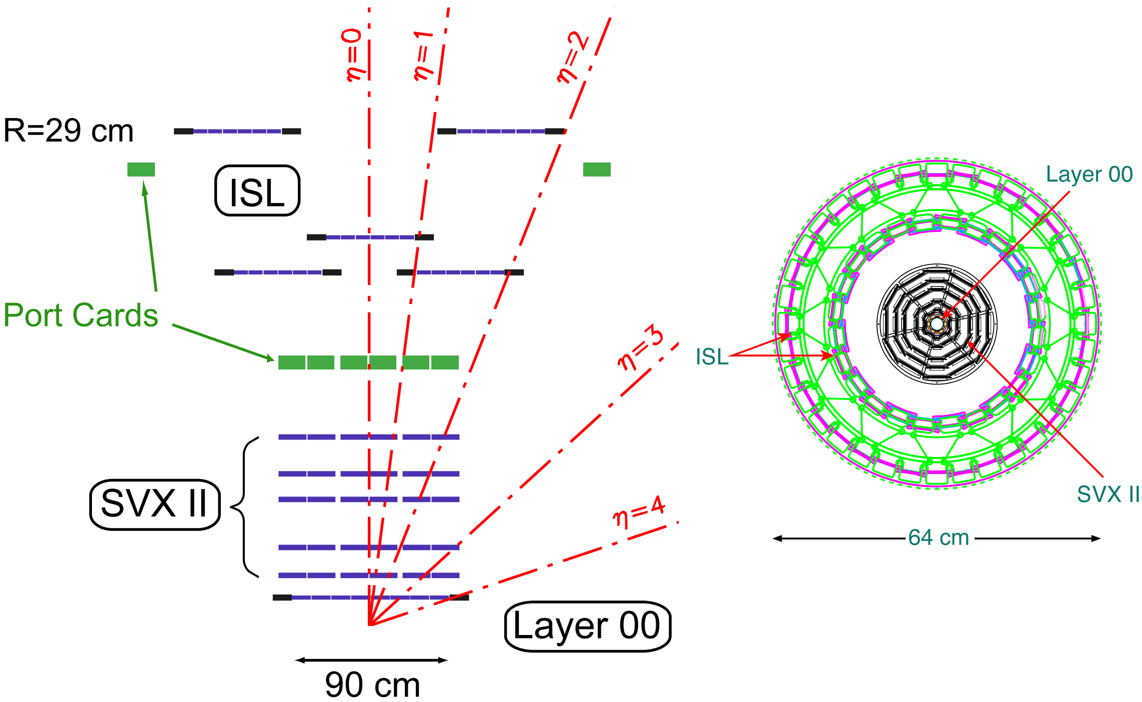

Figure 2.4: Two views of the silicon system of CDF II. The sidewise view with different layers and the η coverage of SVX (left) and the frontal view (right).

In a drift chamber the particles move through and ionize the gas inside. Electrons drift towards the sense wires and induce an electric signal read out by an electronic circuit. Thanks to the large radius, the momentum is measured precisely with the COT. The other important quantity measured with the COT is the energy loss $\frac{dE}{dx}$ used for particle identification.

2.2.3 The Time Of Flight Detector

The Time Of Flight Detector (TOF) is used to measure the time a particle needs to travel from the interaction point to the TOF. It is located outside the COT at the radius r=140 cm and consists of 216 scintillator bars.

To determine the time of flight $T = T_{meas} - t_0$ one needs besides T_{meas} also the time of interaction t_0 which is obtained by matching a reconstructed COT track to a TOF signal. The use of these measurements is the identification of the measured particles by determining the mass with respect to the momentum and the time of flight.

2.2.4 The Calorimeters

Calorimeters are used to measure the kinetic energy of particles. The calorimeter system consists of different calorimeters like central and end-plug electromagnetic and hadronic calorimeters and the end-wall hadronic calorimeter. They cover the range between $-3.6 < \eta < 3.6$ and $0 < \phi < 2\pi$. The assignment of calorimeters is absorbing the kinetic energy by interaction of crossing particles. The calorimeters are used to measure energy of high energy electrons, photons and hadronic jets. They are shown as red and blue areas in figure 2.5.

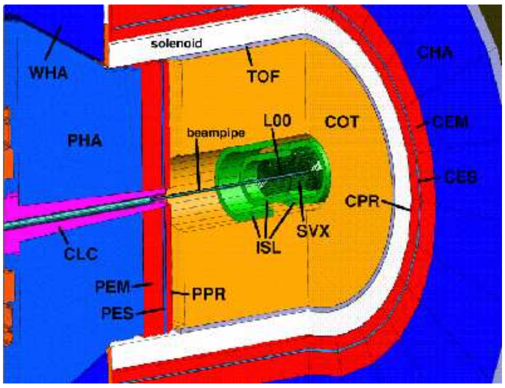

Figure 2.5: The calorimeters of CDF II detector

2.2.5 The Muon System

The muon system is located outside the calorimetry system of the detector, shown in figure 2.6. It consists of different drift chambers and scintillators. There are three drift chambers the central muon detector, central muon upgrade and intermediated muon detector. In addition there is the muon extension detector which is a combined drift chamber and scintillator counter. Most of the particles detected in the muon system, i.e. they are not absorbed in the calorimeters are muons. Some kaons and pions also survive calorimeters and can produce muon fake rates at the level of few percent.

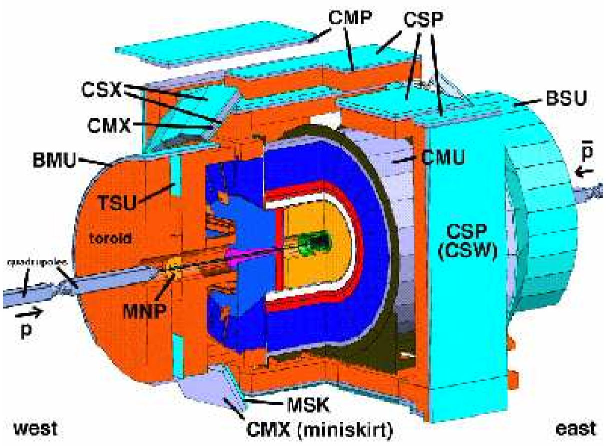

Figure 2.6: The muon system of CDF II detector

2.2.6 The Trigger System

Working with collider detectors cause the problem that a huge amount of collision occur every second and it is not possible to store all of them. So the aim is to make a preselection and store only interesting ones. At CDF it is done by a three level trigger system. These three trigger levels decide consecutively whether a event is sent to the next trigger level and is finally stored or not. This data flow is shown in figure 2.7. The first two trigger levels are implemented in hardware and use different detector devices while the third trigger level is implemented in software running

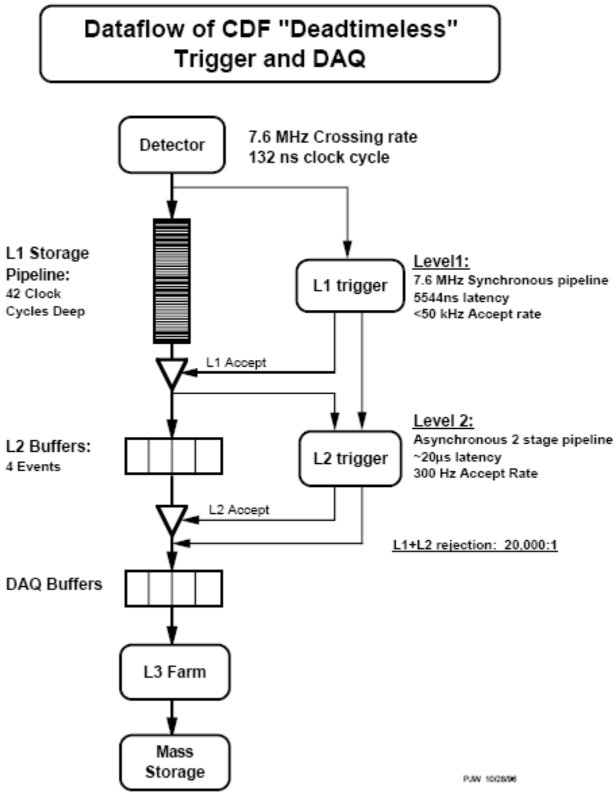

Figure 2.7: Data-flow and data acquisition of the CDF II trigger system

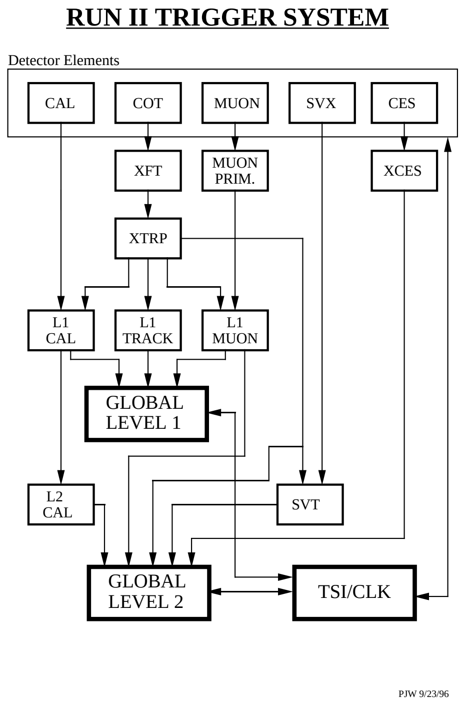

Figure 2.8: Detector elements and the first two trigger levels

on a computer farm. Each trigger has his own criteria and combinations of criteria, called trigger paths, for deciding. The trigger system reduces the amount of data by a factor of ~ 20000.

One aim of dividing the trigger system in three different levels is avoiding or at least reducing dead time by dividing the available data from the detector, shown in figure 2.8. The level one trigger uses only data from calorimeter, muon chamber and muon scintillators and the central tracking chamber. These informations are combined in various different ways and are the basis to make a very fast decision. If the event is accepted by level one trigger, it is passed to the level two trigger. There the informations are combined in the first stage with data from the CES strip chambers and the silicon vertex detector. In the second stage data from the calorimeters, track informations, muon data and SVX data is used to make the decision. At the level three system, the different sections of data are combined to form a whole event the first time. It acts as a filter and reduces the number of events that need to be written to disk and it is the first stage of the event reconstruction.

Chapter 3

Analysis Outline

3.1 Data Selection

3.1.1 Datasets

The data used for the this thesis have been taken by the CDF II detector between February 2002 and February 2006. The integrated luminosity of almost 1 fb^{-1} is shared across three datasets, see table 3.1. The accumulation of the data was done by a trigger called Two Track Trigger.

xbhd0d [1]	Feb 2002 - Aug 2004	341 pb^{-1}
xbhd0h [1]	Dec 2004 - Sep 2005	397 pb^{-1}
xbhd0i [1]	Oct 2005 - Feb 2006	253 pb^{-1}
integrated total luminosity		991 pb^{-1}

Table 3.1: Datasets collected between 2002 and 2006 with the Two Track Trigger and a total integrated luminosity of almost 1 fb^{-1}.

3.1.2 The Two Track Trigger

All analysis in this work are using data which is collected by using the so called Two Track Trigger which is defined by a set of trigger paths (see 2.2.6). The trigger requirements of the Two Track Trigger are given by:

- **Level 1:** at least two oppositely charged XFT [2] tracks with $p_t > 2\ GeV/c$, an aggregate transversal momentum $p_{t,1}+p_{t,2} > 5.5\ GeV$ and an angle between the tracks with $\Delta\phi < 135°$
- **Level 2:** at least two SVT [3] tracks matching to XFT tracks ($\chi^2_{SVT} < 25$) with $p_t > 2\ GeV/c$ and an impact parameter in the range $100\mu m < |d_0| < 1mm$
- **Level 3:** the SVT tracks match to COT tracks and a confirmation of the p_t and impact parameter requirements; the angle between tracks $2° < \Delta\phi < 90°$ and the decay length in the r-ϕ-plane projected on p_t (L_{xy}) greater than $200\mu m$

[1] CDF internal denotation of the data set
[2] Extremely Fast Tracker.
[3] Silicon Vertex Trigger, [17]

3.1.3 Event reconstruction

The offline event reconstruction is done after collecting data with the two track trigger and storing it. Tracks are reconstructed out of tracking informations from COT and SVX. The B_s-mesons are reconstructed using the BottomMods [22] software package. It is highly modular and starts with selecting appropriate tracks and ends with the reconstruction of higher level objects, like reconstructed B mesons. At the beginning, tracks with a successful helix fit are refitted with either a pion or kaon mass hypothesis taking multiple scattering and energy loss into account. They are stored internally as stable particle candidates. The next step is the reconstruction of unstable particles in the opposite order as the decays took place. In case of $B_s \to D_s\pi$ and $D_s \to K^*K$, starting from stable particles π and K, the unstable particles D_s, K^* and subsequently B_s are reconstructed.

3.1.4 Decay Channels

The limiting factor of the B_s mixing measurement is the available statistics. So every available event should be used for the analysis. Here only the exclusive hadronic decay channels of B_s mesons are included because of the much better decay time resolution than in the semileptonic ones. The decay channel $B_s \to D_s\pi$, $D_s \to \phi\pi$ is the most important one with the largest branching ratio and the most efficient reconstruction containing almost 40 % of the statistics. All decay channels used in this work are enumerated in table 3.2.

$B_s \to D_s\pi$	$D_s \to \phi\pi$
	$D_s \to K^*K$
	$D_s \to \pi\pi\pi$
$B_s \to D_s\pi\pi\pi$	$D_s \to \phi\pi$
	$D_s \to K^*K$
	$D_s \to \pi\pi\pi$

Table 3.2: The decay channels of B_s mesons used in this work.

3.1.5 Monte Carlo Samples

Obtaining samples of simulated B_s events, two different Monte Carlo generators are used [29]. In case of BGEN [28] as Monte Carlo generator, each event consists exclusively of a b-hadron without an opposite b quark or any fragmentation tracks. The PYTHIA generator [27] simulates both, b-hadrons and the fragmentation tracks. For producing B mesons decays, the event generator EvtGen program package [31, 30] is used. The detector simulation is done by cdfSim [32] where the output has the same structure as real data. The following Two Track Trigger simulation is done by TrigSim++ with the help of svtsim [33]. The reconstruction and selection is done in the same way as the reconstruction and selection of real data.

These Monte Carlo samples are used for the parameterization of mass[1] and decay time[2] distributions of each single component contained in the samples of B_s candidates. The different decays which have to be taken into account are listed in table 3.3. Besides the Monte Carlo samples listed in the table, a signal Monte Carlo sample is used for the determination of the lifetime efficiency curve[3] and also for the parameterization of the invariant mass distribution of the signal.

[1] see section 4.3 Invariant Mass Spectrum
[2] see section 4.4.3 Lifetime Background Distributions
[3] see section 4.4.2 The Efficiency Curve

	$B_s \to D_s\pi$			$B_s \to D_s 3\pi$		
	$D_s \to \phi\pi$	$D_s \to K^*K$	$D_s \to 3\pi$	$D_s \to \phi\pi$	$D_s \to K^*K$	$D_s \to 3\pi$
$\Lambda_B \to \Lambda_c^+\pi^-$		$\surd$	$\surd$			
$\Lambda_B \to \Lambda_c^+ a_1$					$\surd$	$\surd$
$B^0 \to D\pi$		$\surd$				
$B^0 \to D3\pi$					$\surd$	
$B_s \to D_s K$	$\surd$	$\surd$	$\surd$			
$B_s \to D_s K\pi\pi$				$\surd$	$\surd$	$\surd$
$B_s \to D_s X$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$	$\surd$

Table 3.3: Overview of Monte Carlo samples that are taken into account for describing the mass and decay time distribution of the six different exclusively reconstructed B_s decay modes.

3.1.6 Optimization of Signal Selection

Even after the event reconstruction with the BottomMods package, besides real B_s mesons also background events are left in the dataset. To improve the signal selection, i.e. getting a cleaner signal peak, a neural network selection is done. The commitment of NeuroBayes ® [23] distinguishing between signal and background events gave rise to a large improvement of the signal significance $\frac{n_S}{\sqrt{n_S+n_B}}$ where n_S is the number of signal events and n_B the number of background events inside the signal region defined in the invariant mass range between 5.32 and 5.42 GeV. Due to the fact that the significance of the mixing signal is proportional to the signal significance, this improvement has a direct influence to the final result.

$$S = \frac{n_S}{\sqrt{n_S + n_B}} \sqrt{\frac{\epsilon D^2}{2}} \cdot e^{-\sigma_t^2 \Delta m_s^2/2} \tag{3.1}$$

where S is the size of the expected amplitude of the oscillation signal in terms of its standard deviations, ϵD^2 is the tagging power, introduced in 3.3 and σ_t is the decay time resolution of the B_s meson.
More details are described in the works of Andreas Gessler [20], Philipp Mack [19] and Christian Dörr [18]. They are the ones who optimize the signal selection for the six exclusively reconstructed B_s decay channels.

3.2 The B_s Decay Time Measurement

The observable decay time is calculated from known quantities resulting from the candidate reconstruction. The time needed to fly from the primary vertex to the decay vertex is the decay time of a particle. Taking into account the geometrical relations, shown in figure 3.1, the proper decay time of the B_s meson in its rest frame is

$$t = L_{xy} \frac{m_B}{c \cdot p_T(B)}$$

where L_{xy} is the projection of the decay length in the x-y plane, m_B is the B_s hadron mass for which 5.3675 GeV [24] is taken, c denotes the velocity of light and p_t the projection of the transverse momentum in the x-y plane.

The uncertainty of the decay time can be calculated with the help of error propagation. The uncertainty of t has two contributions, namely the decay length resolution and the momentum resolution.

$$\sigma_t = \sqrt{\left(\frac{m_B}{cp_t}\sigma_{L_{xy}}\right)^2 + \left(\frac{m_B L_{xy}}{cp_t^2}\sigma_{p_t}\right)^2} = \sqrt{\left(\frac{m_B}{cp_t}\sigma_{L_{xy}}\right)^2 + \left(\frac{t}{p_t}\sigma_{p_t}\right)^2}$$

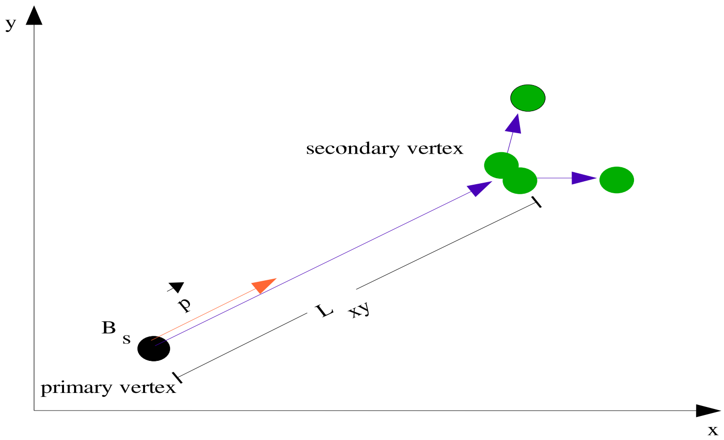

Figure 3.1: Geometrical relations of the B_s decay.

In case of exclusively reconstructed hadronic decays the momentum resolution σ_{p_t} is negligible compared to the L_{xy} resolution $\sigma_{p_{xy}}$. Then the upper expression can be reduced to

$$\sigma_t = \frac{\sigma_{L_{xy}} m_B}{c p_t}$$

In case of semileptonic decays the resolution of σ_{p_t} is dominating because of the missing neutrino momentum and cannot be neglected.

3.3 Flavor Tagging

For the analysis of B_s oscillations it is essential to know whether the B_s meson has mixed or not. This has be to found out by comparing the flavor of the B_s meson at the production time and at the decay time. The flavor of the B_s meson at the decay time is identified by the B_s decay products. The knowledge about the flavor of the B_s meson at the production time is obtained from the flavor tagging algorithm.

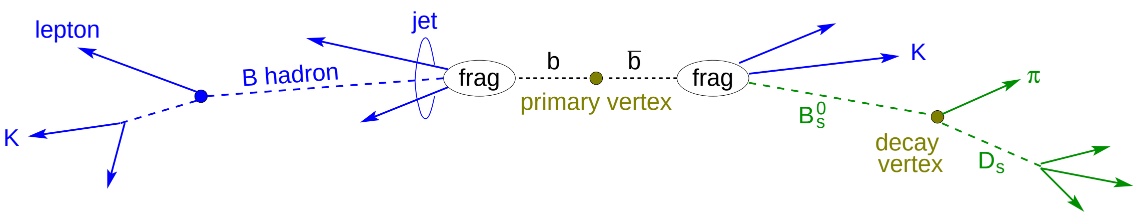

Figure 3.2: Illustration of a $b\bar{b}$ production at the primary vertex. The $\bar{b}$ and a s quark produce a B_s decaying into a D_s and a π at the same side. At the opposite side, the b quark produces another B hadron decaying semileptonically.

The tagging algorithms can be classified into two kinds of taggers, the same side (SST) and the opposite side tagger (OST). The idea of the SST is the identification of the flavor of the studied B_s meson using information from the fragmentation of the b quark to B_s hadron. Another particle besides the B_s meson is created at the same side because forming a $\bar{B}_s$ meson from a b quark, a $\bar{s}$ quark is needed, too. It is created by pulling a $s\bar{s}$ pair from vacuum. The s quark ends up in a kaon with high probability and the charge of the kaon contains information about the flavor of the B_s meson. The OST tries to determine the flavor of the opposite side b quark. In figure 3.2, the same side is on the right of the primary vertex and in this case the side of the $\bar{b}$. The opposite side is on the left of the primary vertex and the side of the b. Both types of taggers can be used basically always because in most cases an opposite b quark exists. Single b quarks can be produced by electroweak processes with much lower cross sections that it can be neglected here.

The major difficulty in tagging is that the decisions are not unambiguous in most cases leading to wrong or no tagging decisions. The quality of the tagger can be expressed by two quantities, the dilution D and efficiency ϵ. The efficiency ϵ specifies the number of cases whether a tagging decision is reached or not and it is defined by

$$\epsilon = \frac{N_{RS} + N_{WS}}{N_{RS} + N_{WS} + N_{NT}}$$

where N_{NT} is the number of events where the tagger gives no decision. N_{RS} is the number of right and N_{WS} is the number of wrong decisions.
The dilution specifies the number of correct decisions with respect to the number of wrong decisions. It is defined by

$$D = \frac{N_{RS} - N_{WS}}{N_{RS} + N_{WS}}$$

These two quantities define the tagging power T as

$$T = \epsilon D^2$$

It defines the effective statistics of the analysis. The product of the real number of signal events with T is equivalent for the analysis to the number of signal events with a perfect tagger. It has influence to the significance, see equation 3.1.

The tagging algorithm used for this analysis is a combined tagger (SST and OST) developed by Andreas Schmidt [21]. The available flavor informations are combined to one single probability variable by a neural network algorithm to consider the correlations between the tagging information sources.

3.4 The Asymmetry Term

The time dependent expression describing the oscillation of a B_s meson into a $\bar{B}_s$ and vice versa has to be determined out of equation 1.3. In this analysis the approximation $\Delta\Gamma = 0$ is used and gives rise to $P(t) = \frac{1}{2}e^{-\Gamma_s t}\,(1 - \cos(\Delta m_s t))$ for the B_s mixing probability. Experimental result of $\Delta\Gamma$ is consistent with 0 and simplifies thus the expression. The number of mesons having mixed or not are given by the following expressions

$$\begin{aligned} N_{unmixed}(t) &= \frac{N_0}{2} e^{-\Gamma_s t}\,(1 + \cos(\Delta m_s t)) \\ N_{mixed}(t) &= \frac{N_0}{2} e^{-\Gamma_s t}\,(1 - \cos(\Delta m_s t)) \end{aligned}$$

where Γ_s is the decay width and N_0 the number of B_s mesons produced at $t = 0$. The asymmetry term I which is the expression that can be measured, expresses the time dependent difference of the number of mixed or unmixed events.

$$I(t) = \frac{N_{unmixed}(t) - N_{mixed}(t)}{N_{unmixed}(t) + N_{mixed}(t)} = \cos(\Delta m_s t) \tag{3.2}$$

Chapter 4

The Fitter Framework

The estimate of the B_s mixing frequency Δm_s requires a sophisticated combination of relevant event information. The provisioning of data is not the main topic in this thesis but is mentioned in chapter 3. The extraction of Δm_s is done by an unbinned maximum likelihood fit inside the fitter framework introduced in this chapter. A general introduction in the maximum likelihood method is given in appendix C.2.

4.1 Fitting Procedure

The determination of the mixing frequency bases upon the time dependent asymmetry

$$I(t) = \frac{N_{unmixed}(t) - N_{mixed}(t)}{N_{unmixed}(t) + N_{mixed}(t)} = \cos(\Delta m_s t)$$

where the numbers of particles having mixed or not at a certain time are taken into account. A binned fit is not feasible in case of limited statistics because of the division of data into bins on the time axis leads to an additional smearing. The large oscillation frequency requires many bins giving rise to statistical bin-by-bin fluctuations. The solution and practical implementation is the introduction of the probability density function of the proper decay time and the individual information for each meson having mixed ($\xi = -1$) or not ($\xi = 1$). The expression for this pdf without normalization is given by

$$P_S(t, \xi) = \left(\frac{1 + \xi D \cos(\Delta m_s t)}{1 + |\xi|} \frac{1}{\tau} e^{\frac{t}{\tau}} \right) \otimes G(t - t', \sigma_t) \cdot \epsilon(t)$$

where the mixing and lifetime term are convoluted by a Gaussian to take the finite resolution of the decay time measurement into account. The limited acceptance of the detector is considered by the multiplication with the efficiency function $\epsilon(t)$. This means that besides the mixing description also the decay time distribution has to be understood and parameterized. The description of the lifetime space has to distinguish between signal and background. The background fraction is a defined as $f_B = \frac{n_B}{n_S + n_B}$ where n_B is the number of background and n_S the number signal events. It is the same parameter in mass and lifetime space of the B_s meson sample and can be determined much better in mass space. So the description and understanding of the invariant mass spectrum is required.

The likelihood function describing mass and lifetime space has about 50 to 60 parameters, depending on the considered channel. So in case of limited statistics there is no chance to estimate all these parameters simultaneously and most of them have to be fixed. That is why several binned maximum likelihood fits are done before to estimate them either to fix them or at least to chose good starting points.

Figure 4.1 shows an overview of the described fit procedure. The left column contains the fits in mass space and the right one the fits in lifetime space. From top to bottom of the flow chart, the chronological order of the fits is mentioned. The order is considerable because some fits use information of previous fits. All information extracted from the binned fits flow together in the unbinned fit where the lifetime τ or the mixing frequency Δm_s of B_s mesons can be determined.

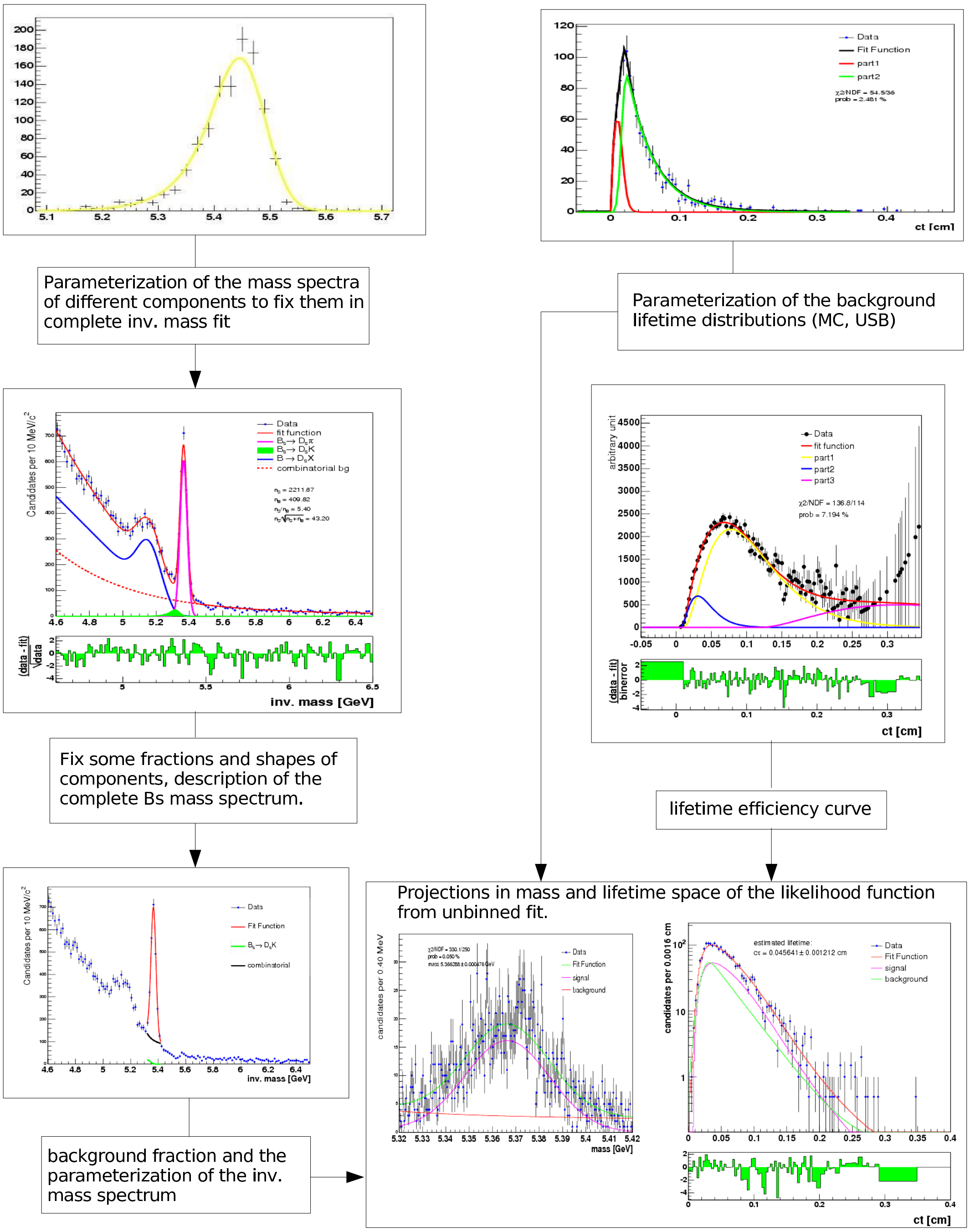

Figure 4.1: Overview of the fit procedure, flowchart which shows the chronological order of the fits in mass space on the left side and the fits in lifetime space on the right.

4.2 Software Architecture

The fitter framework is designed in a way that several decay channels can be used simultaneously for the mixing analysis. It is flexible for adding decay modes easily and be able to change things relating to one or more modes. The possibility consists to decide which fit parameters should be shared by only one or more modes without too much effort.

A general overview over the software architecture of the fitter framework is shown in figure 4.2. Each decay mode is represented inside the framework by its own Mode class containing all infor-

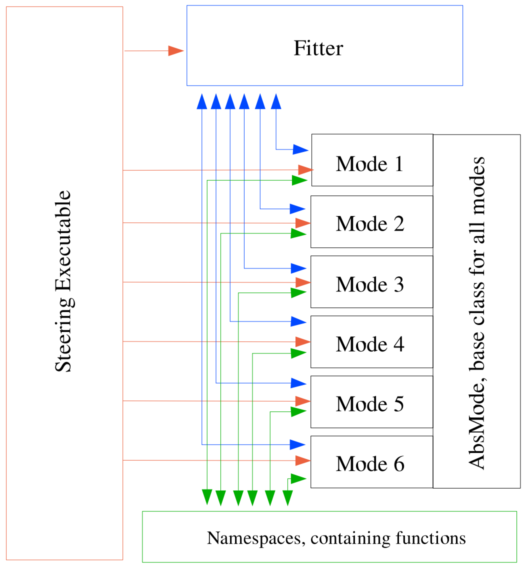

Figure 4.2: Overview of the software architecture of the whole fitting framework.

mation like the data and the method to calculate the likelihood. Also all the binned fits necessary for reducing the effective number of parameters in the unbinned fit are done inside the mode classes. The functions taken for the calculation of the likelihood are similar for all modes and are therefore provided centrally in some namespaces. AbsMode is the base class of all the mode classes providing a common interface for the Fitter which is the core of the framework. Inside the Fitter class, the unbinned fit takes place after including the relevant modes. It gets the information from all involved mode classes and minimizes the negative log-likelihood function bearing in mind that the parameters are identified by their names, further information about the Fitter can be found in section 4.5.3. The minimization algorithm from the TMinuit package from “ROOT” [25] is used.

4.3 The Invariant Mass Spectrum

The invariant mass spectrum of B_s mesons (an example is shown in 4.3) contains a sample of B_s candidates of a special decay mode fulfilling all reconstruction requirements and passing the preselection cuts. Only some of the candidates are real B_s mesons, the rest is background. The reconstructed and real B_s mesons form the signal peak described by a Gaussian distribution with the mean at the position of the B_s mass ($\approx 5.36GeV$). The reason why the background distributions besides the signal have to be considered is the presence of background in the signal region (5.32-5.42 GeV). Understanding the background distribution is the only way to predict the fraction of signal to background or the background fraction $f_B = \frac{n_B}{n_S+n_B}$ where n_S is the number of signal events and n_B the number of background events in the signal region.

Three different kinds of background are contained in the sample of B_s meson candidates. The reasons why the reconstructed tracks end up in the invariant mass spectrum of B_s mesons are essentially different. The first type of background events are the partially reconstructed B_s mesons where one or more particles are missed in the reconstruction. They are responsible for a rise in the lower invariant mass spectrum parameterized by a straight line plus a Gaussian. The combinatorial background results from random combinations of tracks passing the reconstruction and the selection cuts. It is spread over the complete invariant mass spectrum and is described by an exponential function plus a constant. The third type of background events is the reconstruction of the exclusive final state of a different species of b-hadrons where a wrong particle hypothesis is assigned to one or more particles in the decay chain. Examples of such particles reconstructed as B_s mesons are Λ_b or B^0 mesons. These backgrounds are mostly described by a Gaussian or an asymmetric function like an exponential convoluted with a Gaussian.

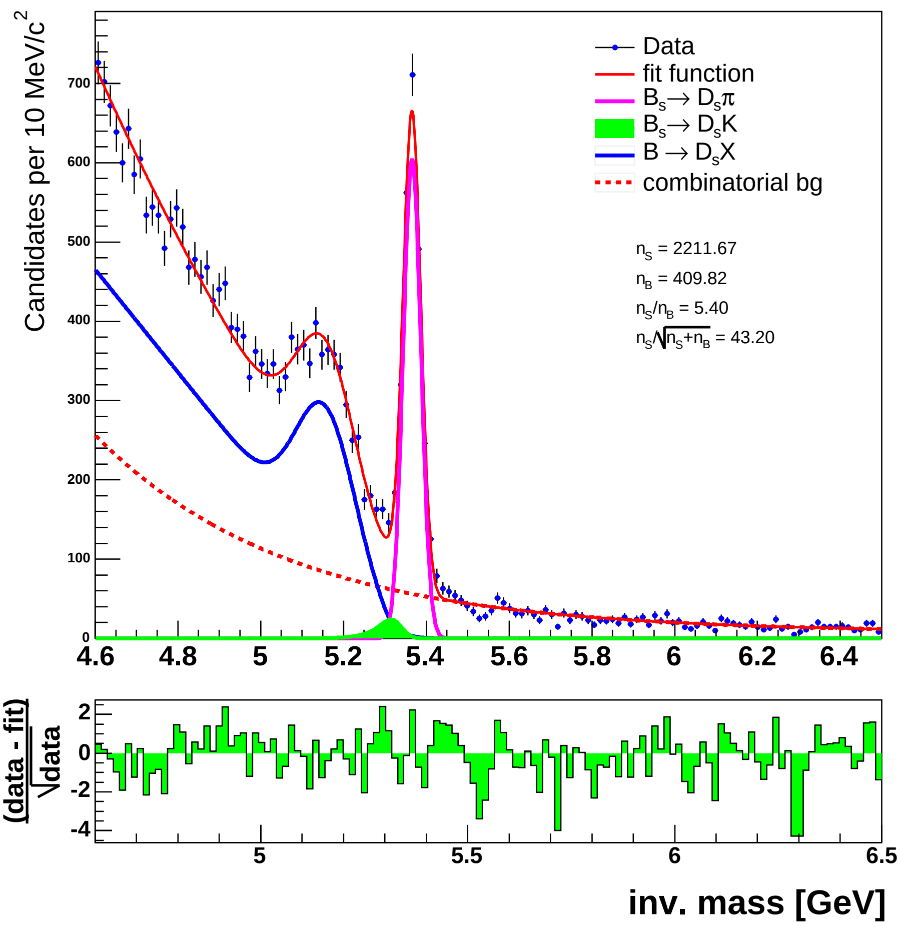

Figure 4.3: Invariant mass spectrum of B_s of the decay channel $B_s \rightarrow D_s\pi$, $D_s \rightarrow \phi\pi$ with a cut on the neural network output at 0.76.

4.3.1 Fitting the Invariant Mass Spectrum

The invariant mass spectrum of the B_s mesons is fitted twice during the fitting procedure. Once the full spectrum is fitted to determine background shapes and fractions of the different components of the spectrum in the complete relevant mass range. The events in the lower and in the upper mass region and of course the events in the signal region are taken into account. Shapes of components reaching from the lower region to the upper region or at least to the signal region are determined. The second fit is done in a narrower range which has to be the same range as the unbinned fit later because the calculation of the background fraction $f_B = \frac{n_B}{n_S+n_B}$ is done there.

The fit function of the invariant mass distribution is a sum of functions describing several components. It is very important to take care of the normalization of the likelihood function if using the maximum likelihood method. Every single component of the complete likelihood function is normalized to a special value, usually 1. To derive the complete likelihood function, the weighted sum of always two of the components is calculated where the weights are the prior probabilities for each single component. They are always summed up pairwisely that a binary tree structure emerges. Figure 4.4 shows an example of the tree structure of the mass function. It is the most general one and is used in the channels $B_s \to D_s(3)\pi$, $D_s \to K^*K$.

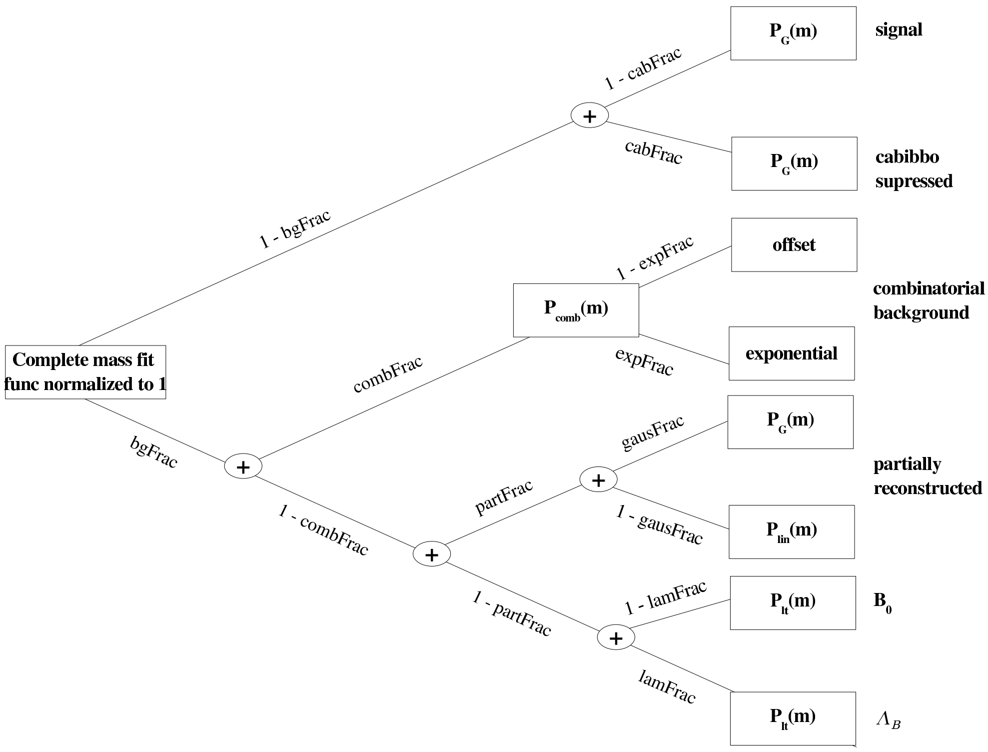

Figure 4.4: Binary tree structure of the function describing the invariant mass spectrum of the channels $B_s \to D_s(3)\pi$, $D_s \to 3\pi$. The fractions are weights of the summands and have values between 0 and 1 to guaranty the normalization. The analytic expressions of the functions P_{xx} can be found in appendix A.

The analytic description of the invariant mass spectrum of the $B_s \to D_s\pi$, $D_s \to \phi\pi$ channel can be written as:

$$\begin{aligned} P(m) =& N \cdot ((1 - f_B) \cdot ((1 - f_{cab}) \cdot P_{cabibbo}(m) + f_{cab} \cdot P_{signal}(m)) \\ &+ f_B \cdot ((1 - f_{comb}) \cdot ((1 - f_{gaus}) \cdot P_{lin}(m) + f_{gaus} \cdot P_{gaus}(m)) \\ &+ f_{comb} \cdot P_{comb}) \end{aligned}$$

It is a typical example although some backgrounds like Λ_b or B^0 are not included because they play no role in this channel because of small branching ratios. For reasons of simplicity this example is chosen. The analytical description of the function shown in figure 4.4 is a much longer expression abandoned here. Each function P_{xx} is a normalized function describing one single background or signal component in mass space. Which function is used for different components can be looked up in table 4.2. The components for different channels are shown in table 4.1.

	$B_s \to D_s\pi$, $B_s \to D_s\pi\pi\pi$		
	$D_s \to \phi\pi$	$D_s \to K^*K$	$D_s \to \pi\pi\pi$
combinatorial background	$\surd$	$\surd$	$\surd$
Λ_b		$\surd$	$\surd$
B^0		$\surd$	$\surd$
partially reconstructed	$\surd$		$\surd$
partially reconstructed (lin.)		$\surd$	
partially rec. containing D_s^*		$\surd$	
partially rec. containing ρ		$\surd$	
signal	$\surd$	$\surd$	$\surd$
cabibbo suppressed $B_s \to D_sK$	$\surd$	$\surd$	$\surd$

Table 4.1: Overview of the components contained in the different channels

	mass function	lifetime function
combinatorial background	$P_{comb}(m)$	$P_{al}(t)$
Λ_b	$P_{lt}(m)$	$P_l(t)$
B^0	$P_{lt}(m)$	$P_l(t)$
partially reconstructed	$b \cdot P_G(m) + (1-b) \cdot P_{lin}(m)$	
partially reconstructed (lin)	$P_{lin}(m)$	
partially rec. containing D_s^*	$P_G(m)$	
partially rec. containing ρ	$P_{lt}(m)$	
signal	$P_G(m)$	$P_S(t)$
Cabibbo suppressed	$P_{lt}(m)$	

Table 4.2: The different functions parameterizing the mass and lifetime distributions are itemized in this table. The analytic expressions of these functions and their normalizations can be looked up in appendix A.

Some of these components are determined with the help of an adequate sample of simulated events because mass distributions of different components are available separately in case of Monte Carlo samples. Fitting the different distributions separately and fixing the parameters helps to reduce the number of free parameters in the complete mass fit.

4.4 Lifetime Space

The lifetime distribution of the sample of B_s meson candidates contains in combination with the tagger the mixing information. So the understanding and modeling of the lifetime distribution is of fundamental meaning. Therefore knowledge about the lifetime distributions of signal and background is necessary. For lifetime distributions of different background components, Monte Carlo samples and the lifetime distribution of the upper side band are taken. The probability density function describing the signal distribution is a smeared exponential multiplied with an efficiency function. The lifetime of the B_s meson is a parameter in the probability density function and can be estimated by fitting the distribution if background and efficiency is fixed before. In case of the mixing analysis, the probability density function is extended by the mixing terms and the rest of the model stays the same.

4.4.1 Probability Density Function

Lifetime

For the measured decay time distribution of a particle, we would expect an exponential distribution $P_l(t) = \frac{1}{\tau}e^{-\frac{t}{\tau}}$. But we have to consider some effects following from the finite precision of the measurement and the limited acceptance of the detector. The finite precision causes a different probability density distribution of the lifetime described by an exponential convoluted with a Gaussian $P_s(t, \sigma_{t,i}) = \frac{1}{\tau}e^{-\frac{t}{\tau}} \otimes G(t-t', \sigma_{t,i})$. The width of the Gaussian is the decay time resolution σ_t. The acceptance of the CDF detector is constituted in the trigger preselection and analysis cuts and depends on the decay time of the observed B_s meson.
Altogether the distribution of the measured decay time including all occurred effects is modeled by the following function

$$P_S(t) = \frac{1}{\tau}e^{-\frac{t}{\tau}} \otimes G(t-t', \sigma_{t,i}) \cdot \epsilon(t) \tag{4.1}$$

where $\epsilon(t)$ is the acceptance or efficiency function of the detector.
The calculated convolution integral and the normalization of the function 4.1 can be found in A.2.

Mixing

The probability density function describing the time dependent evolution of the $B - \bar{B}$ oscillation signal consists of several components. The asymmetry term 3.2 including the $cos(\Delta mt)$ modulation of the mixing signal and the tagging decision ξ. The dilution D has also to be take into account to give consideration to wrong tagger decisions. It is given for each tagged event and expresses the certainty of the tagging decision. Also the mixing term is smeared by a Gaussian because of the finite resolution of the decay time measurement. Another component in the probability density function is the lifetime term describing the decay of B_s mesons.

$$P_S(t, \xi) = \frac{1}{N_S(t, \xi, \sigma_t)} \left(\frac{1 + \xi D \cos(\Delta mt)}{1 + |\xi|} \frac{1}{\tau} e^{\frac{t}{\tau}} \right) \otimes G(t-t', \sigma_t) \cdot \epsilon(t) \tag{4.2}$$

The parameterization of the lifetime distribution of the background is a phenomenological description $P_{bg}(t)$ multiplied by a term taking possible background flavor asymmetries into account. It is globally described by the dilution-like fit parameter D_{bg}. The probability density function is given by

$$\tilde{P}_B(t, \xi) = \frac{1 + \xi \cdot D_{bg}}{1 + |\xi|} \cdot P_B(t) \tag{4.3}$$

where ξ is the already known tagging decision and D_{bg} the background dilution.

4.4.2 The Efficiency Curve

The introduction of the efficiency is required due to the limited acceptance of the detector depending on proper decay time. It can be determined by the ratio of the measured decay time distribution after trigger and cuts and the expected distribution with perfect acceptance.

$$\epsilon(t) = \frac{g(t) \text{ after trigger and cuts}}{\sum_i \frac{1}{\tau}e^{-\frac{t}{\tau}} \otimes G(t-t', \sigma_{t,i})} \tag{4.4}$$

The decay time distribution $g(t)$ is obtained by applying the same signal selection on a sample of simulated signal events where the trigger bias is taken into account. For calculating the denominator of equation 4.4, a sum of the expected distribution over the same simulated signal events with their individual resolutions $\sigma_{t,i}$ is calculated. The resolutions $\sigma_{t,i}$ are the uncertainties of the measured decay time in the sample of simulated events.

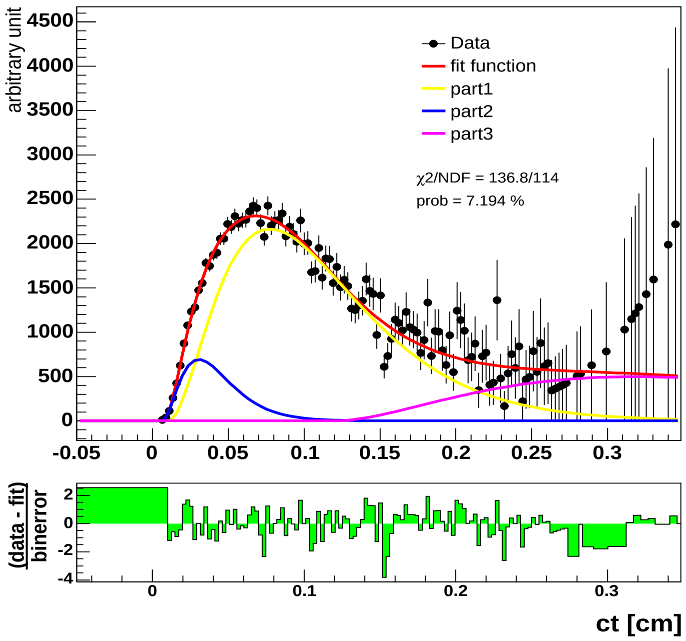

Figure 4.5: Example of an efficiency distribution from the channel $B_s \to D_s\pi$, $D_s \to \phi\pi$ and its parameterization $\epsilon(t)$ together with the three components the function consisting of.

The efficiency distribution (figure 4.5) is described phenomenologically by the function given by

$$\epsilon(t) = \sum_{j=0}^{2} N_{\epsilon,j} \cdot e^{-\frac{t}{\tau_j}} (-\beta_j + t)^2 \cdot \theta(t - \beta_j). \tag{4.5}$$

An advantage of this parameterization is the ability of analytical integration of equation 4.1 and 4.2, see A.2. It is necessary for a fast normalization done very often during the fit. The construction of $\epsilon(t)$ as a sum of three similar functions makes the parameterization very flexible and the analytical integratability is given. But it is also responsible for high correlations between some fit parameters. So it is very hard for the fit algorithm to find a reasonable minimum. Adequate starting values are thus necessary for a good fit result. To find such good starting values, the fit is repeated several times with random starting values and the best fit is taken at the end. These values are not completely random and are constrained in a reasonable interval and the following starting values can not differ arbitrarily from the current value estimated by the fit before.

The method taken for the fit here, is the χ^2 fit instead of the maximum likelihood method. In case of binned maximum likelihood method C.4 implemented in the fit algorithm of ROOT [25], the entries of bins are expected to be Poisson distributed like the bin contents of a frequency distribution. But the efficiency distribution is calculated by a quotient and the bin contents can not be assumed as Poisson distributed.

4.4.3 Lifetime Background Distributions

Before fitting the complete lifetime distribution of the B_s meson candidates after reconstruction and preselection, the different background lifetime components have to be determined. The weighted sum of all background lifetime components together with the signal parameterization is the description of the complete lifetime distribution.

The lifetime distributions of B^0 (figure 4.6), Λ_b and other physical background shapes are parameterized by the following function

$$\begin{aligned} P_l(t) &= N_l \cdot e^{-\frac{t}{\tau}} \otimes G(t,\mu,\sigma) \\ &= \frac{1}{2\tau} \cdot N_l \cdot \mathrm{Erfc}\left[\frac{\frac{\sigma}{\tau} - \frac{t-\mu}{\sigma}}{\sqrt{2}}\right] \cdot e^{\frac{1}{2}\frac{\sigma^2}{\tau^2} - \frac{t-\mu}{\tau}} \end{aligned} \tag{4.6}$$

It is a phenomenological description for asymmetric distributions where the maximum of the function can be moved by varying μ. The parameters have no special meaning and τ for example is not the lifetime of the related particle. The normalization factor N_l is calculated in A.2. normalization.

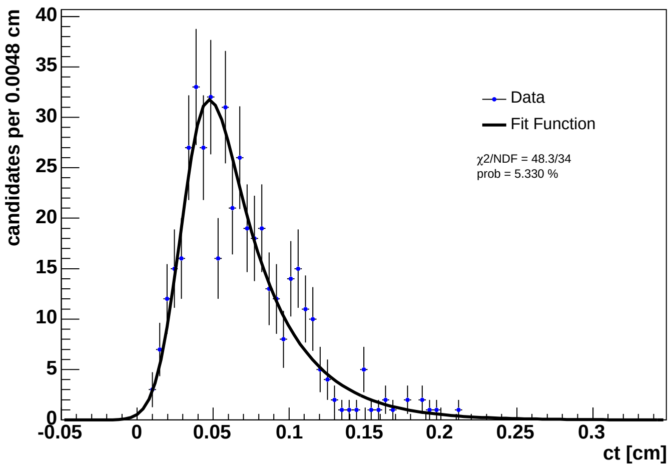

Figure 4.6: Simulated B^0 lifetime distribution of channel $B_s \to D_s\pi$, $D_s \to \phi\pi$

The lifetime distribution of the combinatorial background (figure 4.7) is parameterized by the function

$$P_{al}(t) = \sum_{j=1}^{2} A_j \cdot \frac{1}{2\tau_j} e^{\frac{\mu_j - t + \frac{\sigma_j^2}{2\tau_j}}{\tau_j}} \mathrm{Erfc}\left[\frac{\sigma_j^2 - (-\mu_j + t)\tau_j}{\sqrt{2}\sigma_j\tau_j}\right]. \tag{4.7}$$

The composition of this function enables the analytical integration (see A.2) but the correlations of the parameters require the "random fit" used also for the efficiency fit, see section 4.4.2.

All the parameters describing the lifetime distribution of the background are kept fix in the unbinned lifetime fit in which the signal parameters are estimated.

4.5 The Unbinned Fit

For measuring the B_s mixing frequency, individual information $\vec{x} = (m, t, \sigma_t, \xi, D)$ of each event are used where m is the reconstructed mass of the B_s meson, t is the measured decay time and σ_t

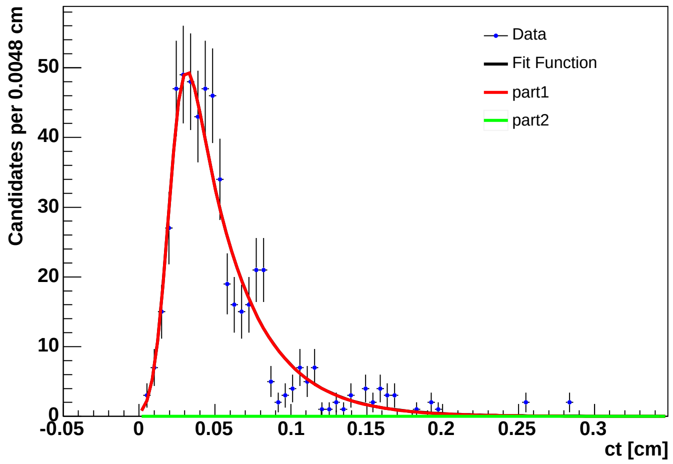

Figure 4.7: Lifetime distribution of the upper side band of channel $B_s \to D_s\pi$, $D_s \to \phi\pi$

is the resolution of t. The tagging decision ξ and tagging dilution D is the result of the tagging algorithm, see section 3.3. It is a discrete variable with the values $\xi = \{-1, 0, 1\}$ for {mixed, no decision, not mixed} and is given for each event. An unbinned fit is the only way to consider the individual information of each event.
The unbinned fit can be executed in mass, lifetime space or as a two dimensional fit in mass and lifetime space. The fraction of background over signal plus background f_B is a common parameter in both spaces. This parameter can only be estimated in mass space because of the similar and confusable shape of signal and background distribution in lifetime space.

4.5.1 The Likelihood

The likelihood L [26] for given events $\vec{x}_i$ is

$$\begin{aligned} L \equiv \prod_{i=1}^{N} P(\vec{x}_i) &= \prod_{i=1}^{N} P(\vec{x}_i, S \cup B) \\ &= \prod_{i=1}^{N} P(\vec{x}_i, S) + P(\vec{x}_i, B) \\ &= \prod_{i=1}^{N} P(\vec{x}_i|S) \cdot P(S) + P(\vec{x}_i|B) \cdot P(B) \end{aligned}$$

where S denotes signal and B background. $P(\vec{x})$ and $P(\vec{x}, ...)$ are probability density distributions and are therefore normalized to one.

$$\int P(\vec{x})d\vec{x} = \int P(\vec{x}|S)d\vec{x} = \int P(\vec{x}|B)d\vec{x} = 1$$

$P(S)$ and $P(B)$ are a priori probabilities that the event $\vec{x}$ is signal or background, so $P(S)+P(B) = 1$.

$$f_B := P(B) \text{ and } P(S) = 1 - f_B$$

The factor f_B is the already known background fraction $f_B = \frac{n_B}{n_S+n_B}$.

In case of an unbinned mass fit, the probability density function $P(\vec{x}, S)$ and $P(\vec{x}, B)$ are only mass dependent and the likelihood is given by

$$P(m_i) = (1 - f_B) \cdot P_S(m_i) + f_B \cdot P_B(m_i)$$

The exact composition of $P_S(m)$ and $P_B(m)$ is dependent which channel is taken into account and can be checked up in section 4.3.1.

For an unbinned lifetime fit, the probability density function in lifetime space is constructed analogous to the probability density function for mass space.

$$P(t_i) = (1 - f_B) \cdot P_S(t_i, \sigma_{t,i}) + f_B \cdot P_B(t_i, \sigma_{t,i})$$

$P_S(t, \sigma_t)$ is given by the equation

$$P_S(t, \sigma_t) = \frac{1}{\tau} e^{-\frac{t}{\tau}} \otimes G(t - t', \sigma_t) \cdot \epsilon(t)$$

which is already introduced in section 3.2 and details can be seen in A.2. $P_B(t)$ is the description of the lifetime distribution of the background (see section 4.4.3) which is a combination of different background components. Depending on the channel, besides the combinatorial background $P_{al}(t)$ also physical backgrounds have been taken into account. An example for $P_B(t)$ is given by equation 4.8 and the background templates which were used can be extracted out of table 4.1. The partially reconstructed background component is not considered in the description of the lifetime space because most of the partially reconstructed component is truncated in the narrow mass range.

$$P_B(t_i) = (1 - f_{B_0} - f_\Lambda) \cdot P_{al}(t_i) + f_{B_0} \cdot P_{B_0}(t_i) + f_\Lambda \cdot P_\Lambda(t_i) \tag{4.8}$$

In case of the mixing analysis, the expressions for $P_S(t, \sigma_t)$ and $P_B(t)$ have to be exchanged with the expressions including the mixing terms. So they are given by

$$P_S(t, \sigma_t, \xi) = \frac{1}{N_S(t, \xi, \sigma_t)} \left(\frac{1 + \xi D \cos(\Delta m t)}{1 + |\xi|} \frac{1}{\tau} e^{\frac{t}{\tau}} \right) \otimes G(t - t', \sigma_t) \cdot \epsilon(t) \tag{4.9}$$

$$P_B(t, \xi) = \frac{1 + \xi \cdot D_{bg}}{1 + |\xi|} \cdot P_{bg}(t) \tag{4.10}$$

as they are already introduced in section 4.4.1.

Usually the unbinned maximum likelihood fit is done simultaneously in mass and lifetime space because of the additional information introduced by the mass terms. The value of the probability density function in mass space multiplied with the lifetime term contains information of the probability whether the event is signal or background. The likelihood in this case is given by

$$\begin{aligned} L &= \prod_{i=1}^{N} P(\vec{x}_i|S) \cdot P(S) + P(\vec{x}_i|B) \cdot P(B) \\ &= \prod_{i=1}^{N} (1 - f_B) \cdot P_S(m_i) \cdot P_S(t_i, \sigma_{t,i}, \xi) + f_B \cdot P_B(m_i) \cdot P_B(t_i, \xi) \end{aligned} \tag{4.11}$$

where $P_S(t, \sigma_t, \xi)$ and $P_B(t, \xi)$ can contain the lifetime expression with or without the mixing term.

The projection of the likelihood in mass and lifetime space respectively is shown in figure 4.8 and 4.9. The figures show the projections of the likelihood function in mass and lifetime space together with data of the $B_s \to D_s\pi$, $D_s \to \phi\pi$ decay channel.

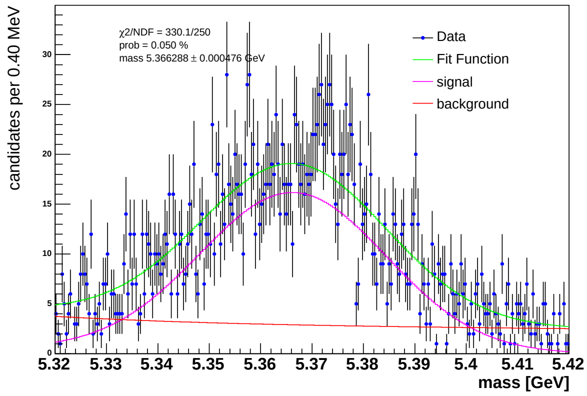

Figure 4.8: Projection of the likelihood function in mass space fitted to data of the $B_s \to D_s\pi$, $D_s \to \phi\pi$ decay channel

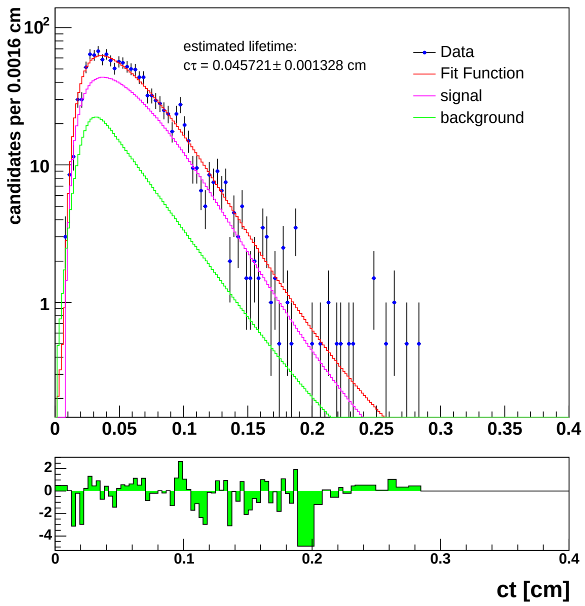

Figure 4.9: Projection of the likelihood in the lifetime space with logarithmic scale. The measured lifetime of the B_s mesons is $457.21 \pm 13.28\mu m$.

4.5.2 Combined Fit

The unbinned fit described in this chapter can be done for each regarded exclusive reconstructed decay mode. The mass spectra and the lifetime distributions are slightly different for each channel but some parameters like Δm or the lifetime τ of the B_s meson are in common in all channels. So each channel needs to be considered separately and individual fit functions have to be taken in to account. All binned fits have to be done for each channel and the results have to be provided to the unbinned fit. To increase statistics and use all available information of the channels, the unbinned fit has to take into account the data of all channels. It is done by multiplying the likelihoods of all channels, see equation 4.12.

$$L(\vec{x}, \vec{a}, \Delta m, \tau) = \prod_{j}^{N} L_j(\vec{x}, \vec{a}, m, \Delta m, \tau, A) \tag{4.12}$$

M denotes the number of channels. L_j is the likelihood of one channel, L the likelihood of all channels. $\vec{x}$ are the measured values and $\vec{a}$ are the parameters of channel j. Δm, τ and m are the parameters all modes have in common.

4.5.3 Fitter

The realization of such a combined fit requires a sophisticated handling of the fit parameters. Each mode has its own parameter list with individual and common parameters. By adding up the lists, the overlaps are taken into account automatically without spending much effort.

The class called “Fitter” is the core of the fitter framework and does the minimization of the negative log-likelihood function of the unbinned fit.
For doing the unbinned fit, an object of the type “Fitter” needs the information which kind of fit (mass, lifetime, mass-lifetime,...) has to be done and which modes have to be included. Inside the Fitter basically two global parameter lists plus one parameter list per channel are existing. One global list includes only the parameters playing a role in the current kind of fit (mass parameters in mass fit, lifetime parameters in lifetime fit,...). This list is provided to the minimizing algorithm and all parameters in this list which are not fixed are varied to minimize the negative log-likelihood function. The parameters in this list are identified by their index. The parameter list of each included channel contains the parameters in the sequence needed by the likelihood. They are identified by their name but they do not have the proper values given by the minimizing algorithm. The other global list includes all parameters of all modes relating the indices and names of the parameters. The minimization package used in the Fitter is the TMinuit class of the ROOT package [25]. It is a well tested algorithm and was originally implemented in Fortran and is converted to a C++ class. It minimizes a given function by varying the parameters which are members of the TMinuit parameter list with respect to the given parameter limits. The values of the parameters in the lists of the likelihood functions have to be updated on the actual values out of the TMinuit parameter list. In the Fitter it is done by the alignment of the likelihood parameter lists with the TMinuit parameter list with the help of the global list containing all parameters.

4.6 Verification of the Fitter Framework

The complexity of the complete fitter framework necessitates a verification whether it is working properly. An implicit phenomenological test where the well known input is compared to the fit results. This is done by a Toy Monte Carlo study generating distributions which are fitted afterwards. The parameters of the input distributions are known and can be compared to the results estimated by the fits.

4.6.1 Verification

Such Toy Monte Carlo experiments are usually done several times (order of magnitude 1000) obtaining a statistical significant sample of each estimated parameter. Statistical fluctuations cause slightly different values for the estimated parameters but the pulls of them should be normally distributed around 0 with $\sigma = 1$. So each bias caused by the fitter framework should be observed by a look on the pulls. For two important fit parameters (mass and lifetime of B_s mesons) the pulls are calculated for each toy experiment. Figure 4.10 and 4.11 show the distributions together with a fitted Gaussian.

The verification of the framework is done in the $B_s \to D_s\pi$, $D_s \to \phi\pi$ decay mode, i.e. the components combinatorial background, partially reconstructed background, cabibbo suppressed and signal are taken into account in the simulation. About 20000 events are generated, it is conform with the number of events in data of this channel.

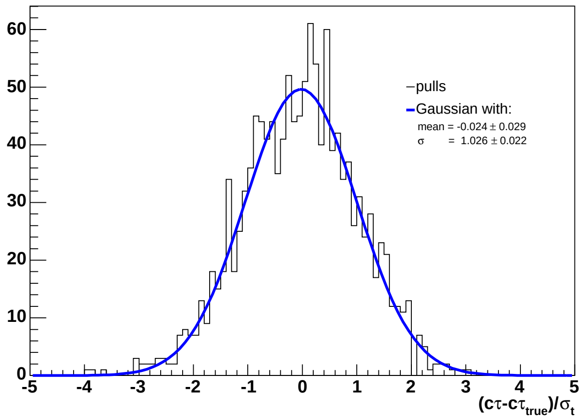

Figure 4.10: Pulls of $c\tau$ of approximately 1300 Toy Monte Carlo experiments.

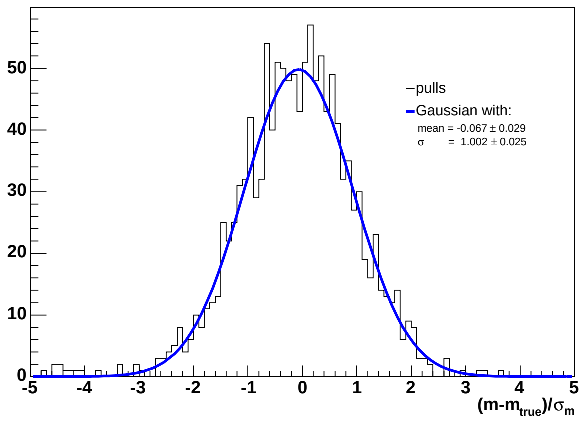

Figure 4.11: Distribution of pulls of the mass m parameterized by a Gaussian.

Parameter	Mean	RMS
$c\tau$	-0.024 ± 0.029	1.026 ± 0.022
m	-0.067 ± 0.029	1.002 ± 0.025

Table 4.3: Pulls of fit parameters

The pulls are well consistent with unit Gaussian in both cases, see fit parameters in table 4.3.

Chapter 5

Determination of Δm_s

5.1 The Amplitude Scan

The measurement of the mixing frequency Δm_s can be done by performing the unbinned maximum likelihood fit on a data sample of B_s candidates as it is described in chapter 4. Therefore the probability density function 4.9 has to be taken for the description of the signal lifetime space and Δm_s is a free fit parameter. In case of low statistics, i.e. a small number of available B_s events, the unbinned fit will not converge in general, and a derivation of limits for Δm_s will be difficult. Instead of a direct determination of Δm_s by the fit, an alternative approach, called amplitude scan, can be chosen. It is based on the idea of performing a Fourier transformation of the oscillation signal [37]. The term describing the oscillation signal has to be amended by the amplitude A.

$$\frac{1 + \xi \mathbf{A} D \cos(\Delta m t)}{1 + |\xi|} \tag{5.1}$$

The amplitude A is a free parameter in the unbinned maximum likelihood fit when Δm_s is fixed. Such fits are repeated for different values of Δm_s so that the whole spectrum is be scanned. The expected value of the amplitude A for the correct assumption of Δm_s is compatible with 1 and compatible with 0 for all other values of Δm_s.

The analysis provides an easy way to evaluate the sensitivity and a lower limit for Δm_s. The sensitivity of the analysis is defined as the value of the frequency for which a measured amplitude $A = 0$ would imply the exclusion of $A = 1$ at the desired confidence level. The degree of exclusion of a given frequency in the scan, for which the measured amplitude and associated uncertainty are A and σ_A, is given by [38]

$$\frac{1}{\sqrt{2\pi}\sigma_A} \int_{-\infty}^{1} e^{-\frac{(x-A)^2}{2\sigma_A^2}} \, dx$$

The exclusion and sensitivity conditions are given as follows for a confidence level of 95%:

$$\begin{aligned} A + 1.645 \cdot \sigma_A &< 1 \quad \text{95\% C.L. exclusion condition} \\ 1.645 \cdot \sigma_A &= 1 \quad \text{95\% C.L. sensitivity condition} \end{aligned}$$

The exclusion limit is defined as the largest frequency value below which all frequencies are excluded.

5.2 Result

The amplitude scan shown in figure 5.1 is a tentative result of the mixing analysis done by this fitter framework. Data of the most important decay channel ($B_s \to D_s\pi$, $D_s \to \phi\pi$) is used because it has the best signal significance and it provides about 40% of the statistics compared to the other five modes. The amplitude scan is also one of the first applications of a preliminary version of the new combined flavor tagging algorithm [21].

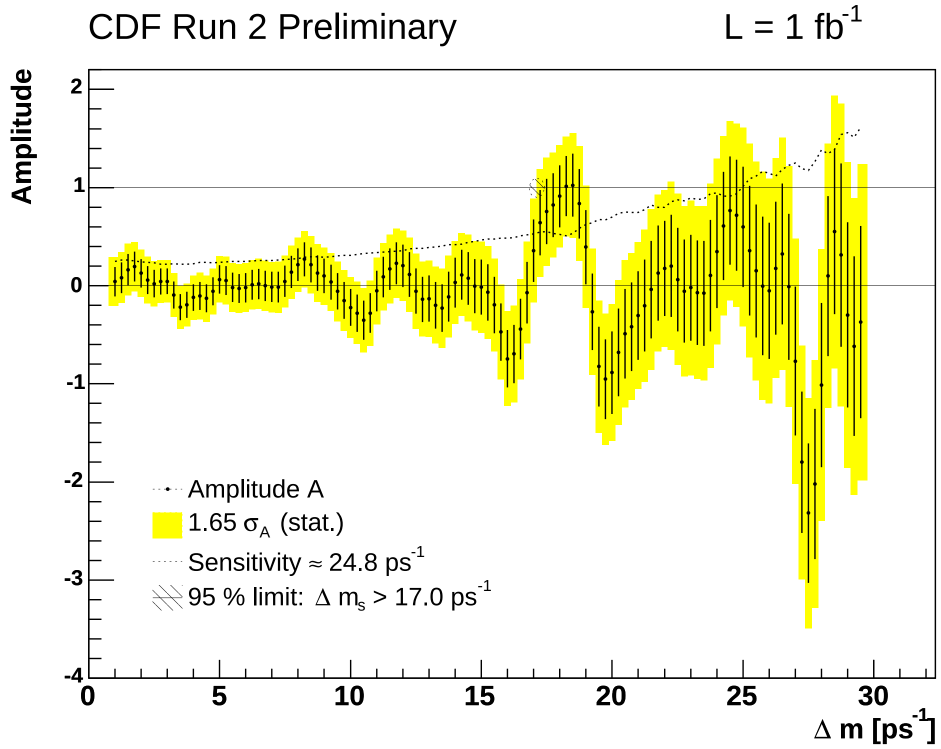

Figure 5.1: Amplitude scan

The significance curve is the dotted line and leads to a sensitivity of 24.8 ps^-1. In a region of Δm_s around 18 ps^{-1}, the amplitude A is consistent with 1 and more than 3 σ away from 0 which gives evidence for a mixing signal. The projection of the log-likelihood ratio (see figure 5.2) has its global minimum at $\Delta m_s = 18.32\, ps^{-1}$. So the value of Δm_s obtained form this analysis is

$$\Delta m_s = 18.32 \pm 0.13\ (stat.)\ ps^{-1}$$

which can be interpreted as a tentative result.

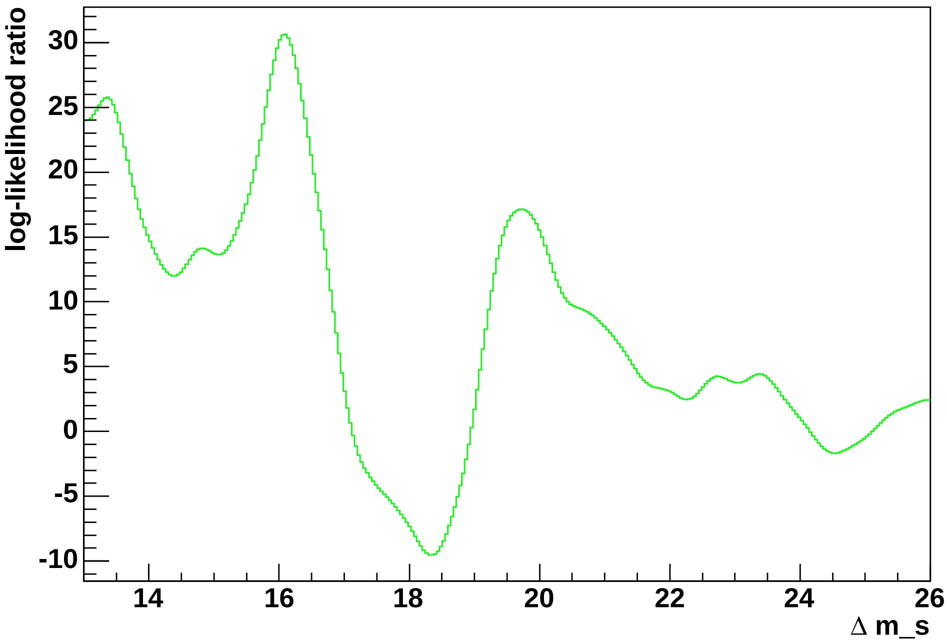

Figure 5.2: Log-likelihood ratio projection

Chapter 6

Conclusion and Outlook

The development of the fitter framework is an important task to benefit from previous work in our group. An improvement of the signal selection was performed for six exclusively reconstructed B_s decay modes with sophisticated neural network technology (NeuroBayes®[23]). Also the combination of same side and opposite side taggers with the same technology yields a doubling of the tagging power. These improvements and the now available data corresponding to the integrated luminosity of 1 fb^{-1} allow the direct time dependent measurement of the mixing frequency Δm_s of the $B_s - \bar{B}_s$ oscillations. The frequency can be extracted with an unbinned maximum likelihood fit in mass, lifetime and tagging space of the sample of B_s candidates. Therefore the invariant mass spectrum and the lifetime distribution have to be understood and parameterized. The probability density function describing the mass and lifetime distribution of one decay channel has more than 50 parameters which have to be determined. In case of limited statistics, it is not possible to determine them simultaneously. So several fits in mass and lifetime space have to be executed to estimate most of the parameters and fix them in the final unbinned fit. To benefit from all available B_s candidates, a combined fit including different decay modes is done and leads to higher significance of the measurement.

Up to now, no result with an optimally working tagger and with the statistics of all six exclusive reconstructed decay modes is prepared. Up to this write-up, an amplitude scan with only one mode is available as it is shown in section 5.2. Evidence for a mixing signal at $18.32 \pm 0.13\, ps^{-1}$ is given with a significance of more than 3σ and the sensitivity of the measurement is $24.8\, ps^{-1}$.

Very recently, the CDF collaboration published the direct measurement of Δm_s [34]. The result under oscillation hypothesis is given by $\Delta m_s = 17.77 \pm 0.10\, (stat) \pm 0.07\, (syst)\, ps^{-1}$. The measurement is already very precise and counts, with a signal significance of more than 5σ, also as safe and observed. The probability of a background fluctuation producing the signal is still merely $8 \cdot 10^{-8}$. Together with $\Delta m_d = (0.507 \pm 0.005)\, ps^{-1}$, $m(B^0) = 5.2794 \pm 0.0005\, GeV$, $m(B_s) = 5.3696 \pm 0.0024\, GeV$ and $\xi = 1.21^{+0.047}_{-0.035}$, the absolute ratio $\left|\frac{V_{td}}{V_{ts}}\right|$ is given by $\left|\frac{V_{td}}{V_{ts}}\right| = 0.2060 \pm 0.0007(\Delta m_s)^{+0.0081}_{-0.0060}(\Delta m_d + theor)$. The current experimental status of the constraints of the unitarity triangle is shown in figure 6.1.

Further steps in the near future are the verification of the framework also for the determination of Δm_s with toy Monte Carlo simulations as it is already done for the determination of $c\tau$. The aim and the application for the fitter framework in future is the improvement of the recent official CDF result in significance and accuracy with the help of the doubled tagging power and a signal selection with higher signal significance. The second application is the calibration of the new tagging algorithm which is required for the measurement of CP-violation in the B_s system.

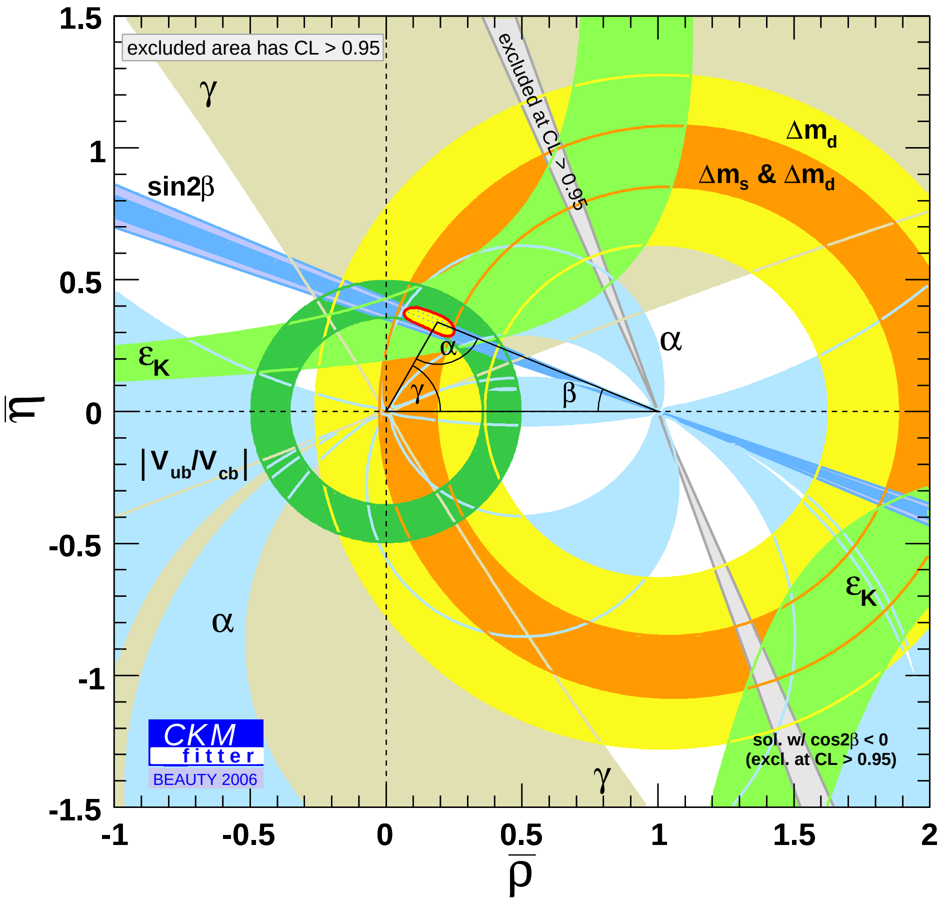

Figure 6.1: Experimental status of constraints on the parameters describing the unitarity triangle with the actual CDF measurement [34]

Appendix A

Analytic Expressions

A.1 Mass Functions

A Gaussian distribution is taken to describe different mass templates in the B_s invariant mass spectrum, for example the signal mass distribution.

$$P_G(m) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(-\mu+m)^2}{2\sigma^2}} \tag{A.1}$$

$$AD_G(m) = -\frac{1}{2} \cdot Erf\left[\frac{\mu - m}{\sqrt{2}\sigma}\right] \tag{A.2}$$

The antiderivative is used to normalize the function in the likelihood fit.

Some mass distributions (B^0, Λ_B) have a shape very similar to a lifetime distribution. So this function is used to describe mass templates phenomenologically and it has nothing to do with the lifetime of a particle itself. It is an exponential function convoluted with a Gaussian.

$$P_{lt}(m) = N_{lt}\frac{1}{2\tau}e^{\frac{\sigma^2}{2\tau^2}-\frac{\mu-m}{\tau}}\left(1 - Erf\left[\frac{-\frac{\mu-m}{\sigma}+\frac{\sigma}{\tau}}{\sqrt{2}}\right]\right) \tag{A.3}$$

Again the antiderivative is used to normalize the function.

$$AD_{lt}(m) = -\frac{1}{2\tau}e^{\frac{\sigma^2-2\mu\tau+2\tau m}{2\tau^2}}\left(e^{-\frac{(\sigma^2+(-\mu+m)\tau)^2}{2\sigma^2\tau^2}}\sqrt{\frac{2}{\pi}}\sigma - t\right.$$
$$\left. + \frac{(\sigma^2 + (-\mu+m)\tau)Erf\left[\frac{\sigma^2+(-\mu+m)\tau}{\sqrt{2}\sigma\tau}\right]}{\tau}\right) \tag{A.4}$$

The normalization constant is given by

$$N_{lt} = \frac{1}{AD_{lt}(m_{high}) - AD_{lt}(m_{low})}$$

The shape of partially reconstructed background in the mass space is parameterized by a straight line plus a Gaussian. The function of the straight line is defined in A.5.

$$P_{lin}(m) = 2 \cdot \frac{p_1 - m}{(p_1 - p_2)^2} \tag{A.5}$$

p_1 is the position where the line meets the m-axis and p_2 is the lowest value of the considered spectrum. This means that the line is defined through its root and the area below it.

The combinatorial background is also described phenomenologically by an exponential function plus a constant.

$$P_{comb}(m) = N_1 \cdot (1 - f_0) \cdot N_{exp} \cdot e^{-\lambda(m-m_e)} + N_2 \cdot f_0 \tag{A.6}$$

In consideration of N_1 and N_2 P_{comb} is normalized to one.

$$N_1 = \frac{1}{\lambda} \cdot e^{\lambda(m_e - m_{low})} - \frac{1}{\lambda} \cdot e^{\lambda(m_e - m_{high})}$$

$$N_2 = m_{high} - m_{low}$$

A.2 Lifetime Functions

The exponential is smeared by a Gaussian, describing the measured decay time of a particle with finite resolution and the lifetime τ. The mean of the Gaussian is 0 and each parameter has a physical meaning. The normalization factor is not specified because this function is multiplied with the efficiency function before normalizing.

$$\begin{aligned} P_\tau(t) &= e^{-\frac{t}{\tau}} \otimes G(t, \sigma) \\ &= \int e^{-\frac{t'}{\tau}} \cdot G(t - t', \sigma) dt' \\ &= \frac{1}{2\tau} e^{\frac{-t + \frac{\sigma^2}{2\tau}}{\tau}} \operatorname{Erfc}\left[\frac{\sigma^2 - t\tau}{\sqrt{2}\sigma\tau}\right] \end{aligned} \tag{A.7}$$

The sign $\otimes$ is used as abbreviation for the convolution integral.

The function $P_l(t)$ parameterizes the lifetime distribution of different physical backgrounds. It is a convolution of an exponential function and a Gaussian with mean μ and is a more general case than A.7.

$$\begin{aligned} P_l(t) &= N_l \cdot e^{-\frac{t}{\tau}} \otimes G(t, \mu, \sigma) \\ &= \frac{1}{2\tau} \cdot N_l \cdot \operatorname{Erfc}\left[\frac{\frac{\sigma}{\tau} - \frac{t-\mu}{\sigma}}{\sqrt{2}}\right] \cdot e^{\frac{1}{2}\frac{\sigma^2}{\tau^2} - \frac{t-\mu}{\tau}} \end{aligned} \tag{A.8}$$

The antiderivative of A.8 is given by

$$AD_l(t) = \frac{1}{2} e^{-\frac{t}{\tau}} \left(e^{\frac{t}{\tau}} \operatorname{Erf}\left[\frac{-\mu + t}{\sqrt{2}\sigma_t}\right] + e^{\frac{\sigma_t^2 + 2\mu\tau}{2\tau^2}} \left(-1 + \operatorname{Erf}\left[\frac{\sigma_t^2 + \mu\tau - t\tau}{\sqrt{2}\sigma_t\tau}\right]\right)\right) \tag{A.9}$$

and the normalization constant by

$$N_l = \frac{1}{AD_l(t_{max}) - AD_l(t_{min})} \tag{A.10}$$

The so called “antilife” function is an phenomenological description of the combinatorial background in the B_s lifetime spectrum. None of the parameters has a special meaning in this case. The function is a sum of two equal functions with different parameters.

$$P_{al}(t) = \sum_{j=1}^{2} A_j \cdot \frac{1}{2\tau_j} e^{\frac{\mu_j - t + \frac{\sigma_j^2}{2\tau_j}}{\tau_j}} \mathrm{Erfc}\left[\frac{\sigma_j^2 - (-\mu_j + t)\tau_j}{\sqrt{2}\sigma_j \tau_j}\right] \tag{A.11}$$

Each summand is multiplied with A_j normalizing the function. The two coefficients are given by

$$A_1 = \frac{w}{AD_{al,1}(t_{max}) - AD_{al,1}(t_{min})}$$
$$A_2 = \frac{1-w}{AD_{al,2}(t_{max}) - AD_{al,2}(t_{min})}$$

Parameter w weights the two parts of the complete function and $AD_{al,j}$ is the antiderivative of one summand of the function A.11.

$$AD_{al,j}(t) = \frac{1}{2}\left(\mathrm{Erf}\left[\frac{-\mu_j + t}{\sqrt{2}\sigma_j}\right] - e^{\frac{\sigma_j^2 + 2\mu_j\tau_j - 2t\tau_j}{2\tau_j^2}} \mathrm{Erfc}\left[\frac{\sigma_j^2 + (\mu_j - t)\tau_j}{\sqrt{2}\sigma_j\tau_j}\right]\right) \tag{A.12}$$

The efficiency distribution is parameterized by $\epsilon(t)$.

$$\epsilon(t) = \sum_{j=0}^{2} N_{\epsilon,j} \cdot e^{-\frac{t}{\tau_j}} (-\beta_j + t)^2 \cdot \theta(t - \beta_j) \tag{A.13}$$

The antiderivative of one summand of the efficiency function is given by A.14 and is used to normalize the function.

$$\begin{aligned} AD_{eff,j}(t) = {} & e^{-\frac{\beta_j}{\tau}} \tau(-e^{\frac{\beta_j - t}{\tau}}((\beta_j - t)^2 - 2(\beta_j - t)\tau_j + 2\tau^2) \\ & + (-2\tau^2 + e^{\frac{\beta_j - t}{\tau}}((\beta_j - t)^2 - 2(\beta_j - t)\tau + 2\tau^2))\theta[\beta_j - t]) \end{aligned} \tag{A.14}$$

$N_{\epsilon,j}$ are the normalization constants of the efficiency function.

$$N_{\epsilon,0} = \frac{w_1}{AD_{eff,0}(t_{high}) - AD_{eff,0}(t_{low})}$$

$$N_{\epsilon,1} = \frac{w_2}{AD_{eff,1}(t_{high}) - AD_{eff,1}(t_{low}))}$$

$$N_{\epsilon,2} = \frac{1 - w_1 - w_2}{AD_{eff,2}(t_{high}) - AD_{eff,2}(t_{low})}$$

The function $P_S(t)$ describing the measured lifetime distribution is a product of the smeared exponential and the efficiency function.

$$P_S(t) = N_S \cdot P_\tau(t) \cdot \epsilon(t) \tag{A.15}$$
$$= \sum_{j=0}^{2} P_{S,j}(t) \tag{A.16}$$

To normalize this function, the antiderivative $AD_S(t)$ is used. It is not the antiderivative of the whole $P_S(t)$ function but the antiderivative of

$$P_{S,j}(t) = N_{S,j} \cdot \frac{1}{2\tau} e^{\frac{-t + \frac{\sigma^2}{2\tau}}{\tau}} \mathrm{Erfc}\left[\frac{\sigma^2 - t\tau}{\sqrt{2}\sigma\tau}\right] \cdot e^{-\frac{t}{\tau_j}} (-\beta_j + t)^2 \cdot \theta(t - \beta_j)$$

$$\begin{aligned}
AD_{S,j}(t) = \frac{1}{2\sqrt{\pi}\tau_j(\tau+\tau_j)^3} & \Bigg(e^{-\frac{1}{2}t(\frac{t}{\sigma^2}+\frac{2}{\tau}+\frac{4}{\tau_j})} \Bigg(e^{\frac{t^2+\frac{\sigma^4}{\tau_j^2}}{2\sigma^2}+t(\frac{1}{\tau}+\frac{2}{\tau_j})} \sqrt{\pi}(\sigma^4(\tau+\tau_j)^2 \\
& + \sigma^2\tau_j(\tau+\tau_j)(-\tau\tau_j+\tau_j^2+2\tau\beta_j+2\tau_j\beta_j) + \tau_i^2(\tau_i^2\beta_j^2+2\tau\tau_j\beta_j(-\tau_j+\beta_j) \\
& + \tau^2(2\tau_i^2-2\tau_j\beta_j+\beta_j^2)))\mathrm{Erf}\left[\frac{\sigma^2+t\tau_j}{\sqrt{2}\sigma\tau_j}\right] - \tau_j \Bigg(\sqrt{2}e^{t(\frac{1}{\tau}+\frac{1}{\tau_j})}\sigma(\tau+\tau_j)(-\sigma^2(\tau+\tau_j) \\
& + \tau_j(2\tau\tau_j+t(\tau+\tau_j)-2\tau\beta_j-2\tau_j\beta_j)) + e^{\frac{1}{2}(\frac{t^2}{\sigma^2}+\frac{\sigma^2}{\tau^2}+\frac{2t}{\tau_j})}\sqrt{\pi}\tau_j(t^2(\tau+\tau_j)^2+\tau_j^2\beta_j^2 \\
& + 2\tau\tau_j\beta_j(-\tau_j+\beta_j)+2t(\tau+\tau_j)(\tau(\tau_j-\beta_j)-\tau_j\beta_j) \\
& + \tau^2(2\tau_j^2-2\tau_j\beta_j+\beta_j^2))\mathrm{Erfc}\left[\frac{\sigma^2-t\tau}{\sqrt{2}\sigma\tau}\right]\Bigg)\Bigg)\Bigg)
\end{aligned} \tag{A.17}$$

and

$$\begin{aligned}
N_{S,0} &= \frac{w_1}{AD_{S,0}(t_{high})-AD_{S,0}(t_{low})} \\
N_{S,1} &= \frac{w_2}{AD_{S,1}(t_{high})-AD_{S,1}(t_{low}))} \\
N_{S,2} &= \frac{1-w_1-w_2}{AD_{S,2}(t_{high})-AD_{S,2}(t_{low})}
\end{aligned}$$

A.3 Mixing Functions

The probability density function describing a mixing signal is given by A.18. It is multiplied with the efficiency function $\epsilon(t)$ known from the lifetime description, equation A.13.

$$P_S(t,\xi) = \frac{1}{N_S(t)} \cdot \left(\frac{1+\xi AD\cos(\Delta m t)}{1+|\xi|}\frac{1}{\tau}e^{-\frac{t}{\tau}}\right) \otimes G(t-t',\sigma_t)\cdot\epsilon(t) \tag{A.18}$$

A denotes the amplitude factor, D is the dilution and ξ is the tagging decision ($\xi = 1$ for not mixed, $\xi = -1$ for mixed, $\xi = 0$ for no decision). For calculating the convolution, the specified integral is solved between the borders 0 and ∞ because the real lifetime t' can not be negative.

$$\begin{aligned}
f(t,\xi) &= \left(\frac{1+\xi AD\cos(\Delta m t)}{1+|\xi|}\frac{1}{\tau}e^{-\frac{t}{\tau}}\right) \otimes G(t-t',\sigma_t) \\
&= \int_0^\infty \left(\frac{1+\xi AD\cos(\Delta m t)}{1+|\xi|}\frac{1}{\tau}e^{-\frac{t}{\tau}}\right)\cdot G(t-t',\sigma_t)dt'
\end{aligned} \tag{A.19}$$

The solution of the convolution is given in equation A.20, where Re denotes the real part of the complex expression because here it is easier to use $Re[e^{-i\Delta m t}]$ instead of $\cos(\Delta m t)$.

$$\begin{aligned}
f(t,\xi) = \frac{1}{2\tau(|\xi|+1)}e^{-\frac{t-\frac{\sigma_t^2}{2\tau}}{\tau}} & \left(\mathrm{Erfc}\left[\frac{\sigma_t^2-t\tau}{\sqrt{2}\sigma_t\tau}\right]\right. \\
& \left. + \xi AD e^{-\frac{1}{2}\sigma_t^2\Delta m^2}\mathrm{Re}\left[e^{-i\Delta m\left(t-\frac{\sigma_t^2}{\tau}\right)}\mathrm{Erfc}\left[\frac{\sigma_t^2-t\tau}{\sqrt{2}\sigma_t\tau}+i\frac{\Delta m\sigma_t}{\sqrt{2}}\right]\right]\right)
\end{aligned} \tag{A.20}$$

To use the likelihood method for fitting, $P_S(t)$ has to be normalized analytically because of a much better performance. Therefore the factor N_S has to be calculated.

$$N_S = \int_{-\infty}^{\infty} f(t,\xi)\cdot\epsilon(t)dt \tag{A.21}$$

The integral can be divided in different parts which can be added up after integration. The efficiency function A.13 can be written in the form

$$\epsilon(t) = \sum_{j=0}^{2}(d_j + c_j t + b_j t^2)e^{-\alpha_j t}\Theta\left(t + \frac{c_j}{2b_j}\right) \tag{A.22}$$

where $\alpha_j = 1/\tau_j$, $b_j = N_{\epsilon,j}$, $c_j = -2b_j\beta_j$ and $d_j = a_j\beta_j^2$. The function A.20 can be written as a sum of two functions

$$f(t,\xi) = f_1(t,\xi) + f_2(t,\xi) \tag{A.23}$$

where

$$f_1(t,\xi) = \frac{1}{2\tau(|\xi|+1)}e^{-\frac{t-\frac{\sigma_t^2}{2\tau}}{\tau}}\mathrm{Erfc}\left[\frac{\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau}\right] \tag{A.24}$$

$$f_2(t,\xi) = \frac{1}{2\tau(|\xi|+1)}e^{-\frac{t-\frac{\sigma_t^2}{2\tau}}{\tau}}\xi ADe^{-\frac{1}{2}\sigma_t^2\Delta m^2}\mathrm{Re}\left[e^{-i\Delta m\left(t-\frac{\sigma_t^2}{\tau}\right)}\mathrm{Erfc}\left[\frac{\sigma_t^2 - t\tau}{\sqrt{2}\sigma_t\tau} + i\frac{\Delta m\sigma_t}{\sqrt{2}}\right]\right] \tag{A.25}$$

This leads to the normalization $N_S(t,\xi)$ which is calculated as a sum of $N_{ji}(t,\xi)$

$$N_S(t,\xi) = \sum_{j=0}^{2}\sum_{i=0}^{5}N_{ji}(t,\xi) \tag{A.26}$$

where j sums over the three components of $\epsilon(t)$ and i over the different summands arise from the product of $f(t)$ and $\epsilon(t)$. The N_{ji} are given by

$$\begin{aligned}
N_{j0}(t,\xi) &= \int d_j e^{-\alpha_j t}f_1(t,\xi)dt \\
N_{j1}(t,\xi) &= \int c_j t e^{-\alpha_j t}f_1(t,\xi)dt \\
N_{j2}(t,\xi) &= \int b_j t^2 e^{-\alpha_j t}f_1(t,\xi)dt \\
N_{j3}(t,\xi) &= \int d_j e^{-\alpha_j t}f_2(t,\xi)dt \\
N_{j4}(t,\xi) &= \int c_j t e^{-\alpha_j t}f_2(t,\xi)dt \\
N_{j5}(t,\xi) &= \int b_j t^2 e^{-\alpha_j t}f_2(t,\xi)dt
\end{aligned}$$

The analytical results of all the N_{ji} are given by the following equations.

$$\begin{aligned}
N_{j0} = \frac{d_j}{2(|\xi|+1)\tau(1+\alpha_j\tau)}&\left(e^{-(\alpha_j+1/\tau)t}\tau\left[e^{\alpha_j^2\sigma_t^2/2+\alpha_j t+t/\tau}\cdot\mathrm{Erf}\left(\frac{\alpha_j\sigma_t^2+t}{\sqrt{2}\sigma_t}\right)\right.\right. \\
&\left.\left.- e^{\frac{\sigma_t^2}{2\tau^2}}\cdot\mathrm{Erfc}\left(\frac{\sigma_t^2-\tau t}{\sqrt{2}\tau\sigma_t}\right)\right]\right)
\end{aligned} \tag{A.27}$$

$$N_{j1} = \frac{c_j \tau}{(|\xi|+1)\tau\sqrt{2\pi}\sigma_t(1+\sigma_t\tau)} \Bigg(- e^{-\alpha_j t - \frac{t^2}{2\sigma_t^2}} \sigma_t^2 - \alpha_j e^{\frac{\alpha_j^2 \sigma_t^2}{2}} \sqrt{\frac{2}{\pi}} \sigma_t^3 \mathrm{Erf}\left(\frac{\alpha_j \sigma_t^2 + \tau}{\sqrt{2}\sigma_t}\right)$$
$$+ \frac{1}{4} c_j \Bigg[\frac{\tau^2}{(1+\alpha_j\tau)^2} e^{\frac{\alpha_j^2\sigma_t^2}{2}} \mathrm{Erf}\left(\frac{\alpha_j\sigma_t^2+\tau}{\sqrt{2}\sigma_t}\right)$$
$$- \frac{\tau(\tau + t + \alpha_j \tau t)}{(1+\alpha_j\tau)^2} e^{\frac{\sigma_t^2}{2\tau^2} - \frac{1+(\alpha_j\tau)t}{\tau}} \mathrm{Erfc}\left(\frac{\sigma_t^2 - \tau t}{\sqrt{2}\tau\sigma_t}\right) \Bigg] \Bigg) \tag{A.28}$$

$$N_{j2} = \frac{1}{2(|\xi|+1)\tau(1+\alpha_j\tau)^2} \Bigg(b_j e^{-\alpha_j t - \frac{t^2}{2\sigma_t^2}} \sigma_t \tau \Bigg[- \sqrt{\frac{2}{\pi}} \left(-\alpha_j \sigma_t^2 + 2\tau - \alpha_j^2 \sigma_t^2 \tau + t + \alpha_j \tau t\right)$$
$$+ e^{\frac{(\alpha_j \sigma^2 + t)^2}{2\sigma_t^2}} \sigma_t \left(1 + \alpha_j^2 \sigma_t^2 - \alpha_j \tau + \alpha_j^3 \sigma_t^2 \tau\right) \mathrm{Erf}\left(\frac{\alpha_j \sigma_t^2 + \tau}{\sqrt{2}\sigma_t}\right) \Bigg] \Bigg)$$
$$+ \frac{1}{4\tau(1+\alpha_j\tau)^3} \Bigg(b_j \tau \Bigg[2 e^{\frac{\alpha_j^2 \sigma_t^2}{2}} \tau^2 \mathrm{Erf}\left(\frac{\alpha_j \sigma_t^2 + \tau}{\sqrt{2}\sigma_t}\right) - e^{\frac{\sigma_t^2 - 2(1+\alpha_j\tau)t}{2\tau^2}} \left(t^2 + 2\tau t(1+\alpha_j t)\right)$$
$$+ \tau^2(2 + 2\alpha_j t + \alpha_j^2 t^2) \mathrm{Erfc}\left(\frac{\sigma_t^2 - \tau t}{\sqrt{2}\tau\sigma_t}\right) \Bigg] \Bigg) \tag{A.29}$$

$$N_{j3} = -d_j \xi A D \cdot \mathrm{Re} \Bigg(\frac{1}{4\tau(1+\alpha_j\tau + i\tau\Delta m)^2 \sigma_t} - e^{-\frac{(1+2i\tau\Delta m + 2\alpha_j\tau(1+i\tau\Delta m))\sigma_t^2}{2\tau^2}}$$
$$\Bigg[e^{\frac{\Delta m^2 \sigma_t^2}{2} + \frac{(1+\alpha_j\tau+i\tau\Delta m)((1+i\tau\Delta m)\sigma_t^2 - \tau t)}{\tau^2}} \tau(1+\alpha_j\tau+i\tau\Delta m)$$
$$\sigma_t \mathrm{Erfc}\left(\frac{(1+i\tau\Delta m)\sigma_t^2 - \tau t)}{\sqrt{2}\tau\sigma_t}\right)$$
$$- e^{\frac{(1+\alpha_j\tau)(1+\alpha_j\tau+2i\tau\Delta m)\sigma_t^2}{\sigma_t\tau^2}} \sqrt{-\tau^2(1+\alpha_j\tau+i\tau\Delta m)^2\sigma_t^2}$$
$$\mathrm{Erfi}\left(\frac{\tau(1+\alpha_j\tau+i\tau\Delta m)(\alpha_j\sigma_t^2+t)}{\sqrt{2}\sqrt{-\tau^2(1+\alpha_j\tau+i\tau\Delta m)^2\sigma_t^2}}\right) \Bigg] \Bigg) \tag{A.30}$$

$$N_{j4} = c_j \xi A D \cdot \mathrm{Re} \Bigg(- \frac{i}{2(-i - i\alpha_j\tau + \tau\Delta m)\sqrt{2\pi}\sigma_t} \Bigg(- \sigma_t^2 e^{-\alpha_j t - \frac{t^2}{2\sigma_t^2}}$$
$$- \alpha_j e^{\frac{\alpha_j^2\sigma_t^2}{2}} \sqrt{\frac{\pi}{2}} \sigma_t^3 \mathrm{Erf}\left(\frac{\alpha_j\sigma^2 + t}{\sqrt{2}\sigma_t}\right) \Bigg) - \frac{1}{4(1+\alpha_j\tau+i\tau\Delta m)}$$
$$\Bigg[- e^{-\frac{(-i+\tau+\Delta m)^2\sigma_t^2 + 2\tau(1+\alpha_j\tau+i\tau\Delta m}{2\tau^2}} \Bigg(- \tau e^{\frac{(1+\alpha_j\tau+i\tau\Delta m)((-1+\alpha_j\tau-i\tau\Delta m)\sigma_t^2+2\tau t)}{2\tau^2}}$$
$$\mathrm{Erf}\left(\frac{\alpha_j\sigma_t^2 + t}{\sqrt{2}\sigma_t}\right) + (\tau + t + \alpha_j\tau t + i\tau\Delta m t))\mathrm{Erfc}\left(\frac{(1+i\tau\Delta m)\sigma_t^2 - \tau t}{\sqrt{2}\tau\sigma_t}\right) \Bigg) \Bigg] \Bigg) \tag{A.31}$$

$$
\begin{aligned}
N_{j5} = b_j \xi A D \cdot \mathrm{Re}\Bigg(& \frac{1}{4(1+\alpha_j\tau+i\tau\Delta m)^2\sqrt{2}}\Bigg[\sigma_t e^{-\alpha_j t-\frac{t^2}{2\sigma_t^2}}\Bigg(-\sqrt{2}\big(-\alpha_j\sigma_t^2+t \\
& +\tau(2-\alpha_j^2\sigma_t^2+i\Delta m t+\alpha_j(-i\Delta m\sigma_t^2+t))\big)+e^{\frac{(\alpha_j\sigma_t^2+t)^2}{2\sigma_t^2}}\sqrt{\pi}\sigma_t\big(1-\alpha_j\tau+i\tau\Delta m \\
& +\alpha_j^3\tau\sigma_t^2+\alpha_j^2(1+i\tau\Delta m)\sigma^2\big)\mathrm{Erf}\left(\frac{\alpha_j\sigma_t^2+t}{\sqrt{\pi}\sigma_t}\right)\Bigg)\Bigg] \\
& +\frac{1}{(1+\alpha_j\tau+i\tau\Delta m)^3}\Bigg[-e^{\frac{(-i+\tau\Delta m)^2\sigma_t^2+2\tau(1+\alpha_j\tau+i\tau\Delta m)t}{2\tau^2}}\tau\left(t^2+2\tau t(1+\alpha_j t+i\Delta m t)\right) \\
& +\tau^2\big(2+2i\Delta m t+\alpha_j^2 t^2-\Delta m^2 t^2+2\alpha t(1+i\Delta m t)\big)\mathrm{Erfc}\left(\frac{(1+i\tau\Delta m)\sigma_t^2-\tau t}{\sqrt{2}\tau\sigma_t}\right)\Bigg] \\
& -2e^{\frac{\alpha_j^2\sigma_t^2}{2}}\tau^3\sigma_t\mathrm{Erfi}\left(\frac{(1+\alpha_j\tau+i\tau\Delta m)^3(\alpha_j\sigma_t^2+t)}{\sqrt{2}\sqrt{-(1+\alpha_j\tau+i\tau\Delta m)^6\sigma_t^2}}\right)\Bigg)
\end{aligned} \tag{A.32}
$$

Erfi(z) denotes the complex error function $Erfi(z) = -iErf(iz)$. The evaluation of the error function Erf(z) with complex argument z is done by an asymptotic series expansion $w(iz) = e^{z^2}Erf(z)$ with

$$
w(iz) = \frac{1}{\sqrt{\pi}z}\left(1+\sum_{m=1}^{\infty}(-1)^m\frac{(2m-1)!!}{(2z^2)^m}\right) \tag{A.33}
$$

It is an advisable evaluation of $Erfi(z)$ because it is still much faster than numerical integration.

Appendix B

Compilation of Fit Results

This appendix shows the compilation of some fit results of distributions and their parameterizations described in chapter 4. The projections of the likelihood in mass and lifetime space of the likelihood function of the unbinned maximum likelihood fits are shown for each of the $B_s \rightarrow D_s\pi$ channels. Some plots of the binned fits in mass and lifetime space are shown before.

Binned Fits of the Invariant Mass Spectra

The invariant mass spectra for different decay modes with the parameterization is shown first. The fit functions which are taken and some explanations can be looked up in section 4.3.1.

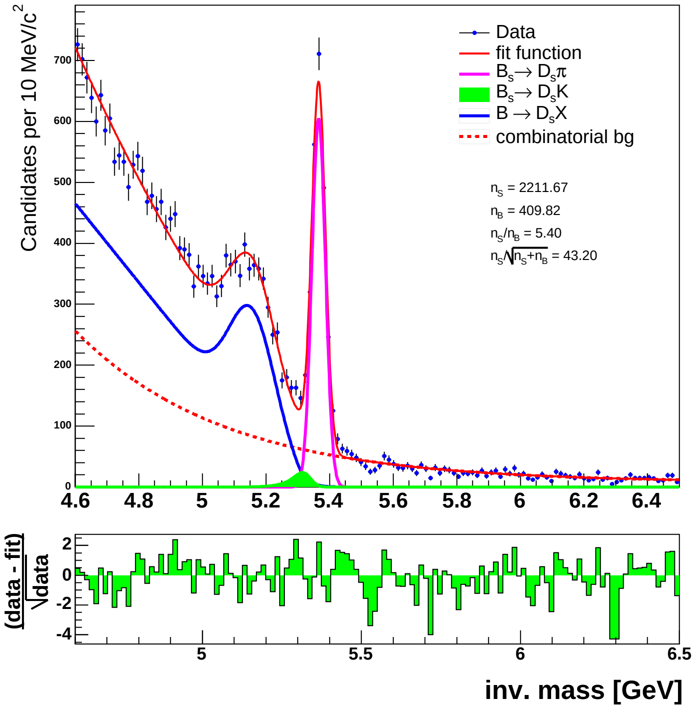

Figure B.1: Invariant mass spectrum of B_s in decay channel $B_s \to D_s\pi$, $D_s \to \phi\pi$

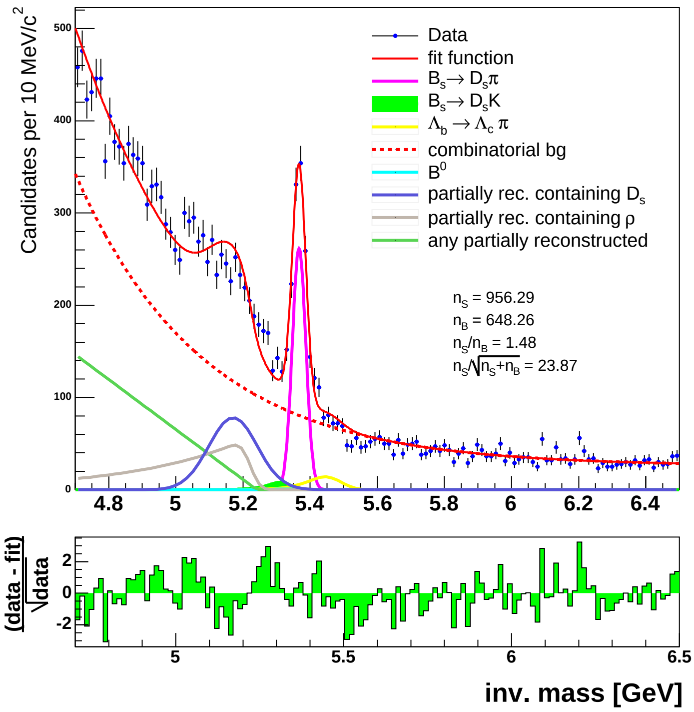

Figure B.2: Invariant mass spectrum of B_s in decay channel $B_s \to D_s\pi$, $D_s \to K^*K$

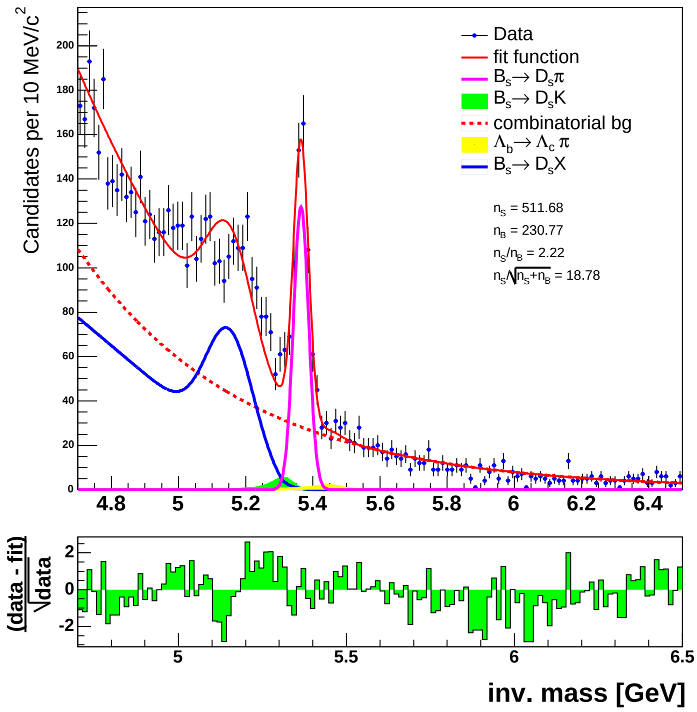

Figure B.3: Invariant mass spectrum of B_s in decay channel $B_s \to D_s\pi$, $D_s \to 3\pi$

Efficiency Curves

The lifetime efficiency distributions with their parameterizations for different decay channels are shown. The fit function which is taken is given by equation A.13.

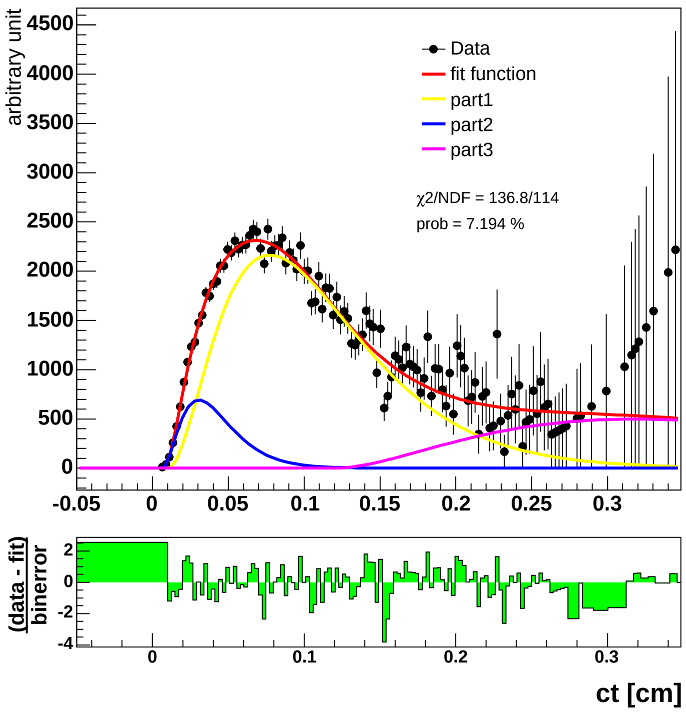

Figure B.4: Lifetime efficiency curve of the decay mode $B_s \to D_s\pi$, $D_s \to \phi\pi$

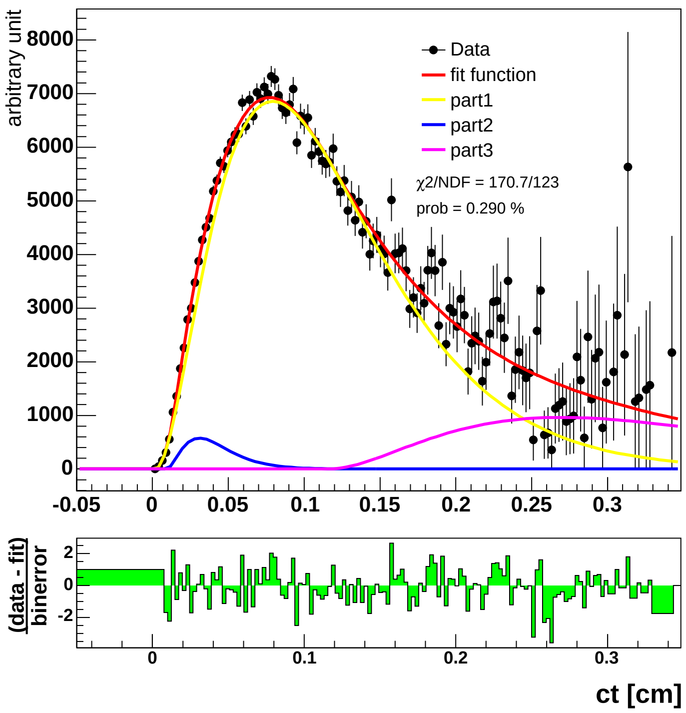

Figure B.5: Lifetime efficiency curve of the decay mode $B_s \to D_s\pi$, $D_s \to K^*K$

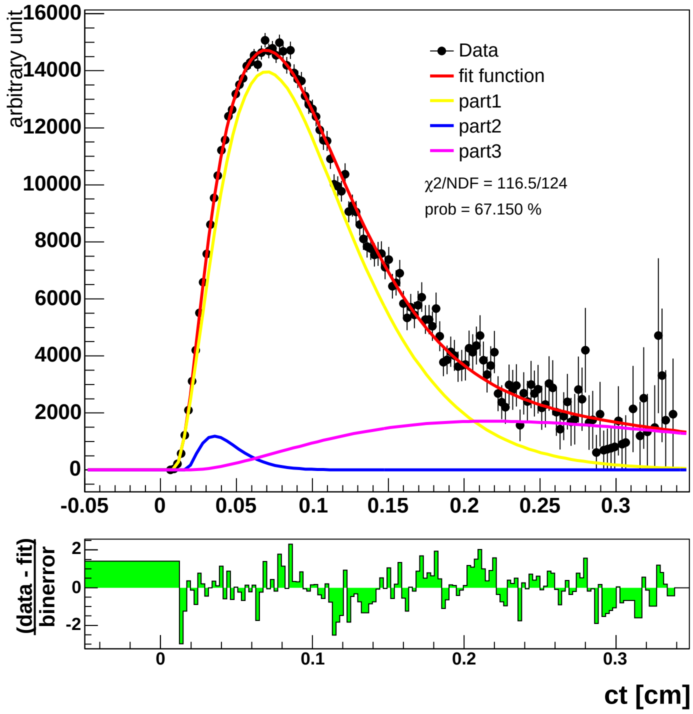

Figure B.6: Lifetime efficiency curve of the decay mode $B_s \to D_s\pi$, $D_s \to 3\pi$

Lifetime distributions of the Combinatorial Background

The lifetime distributions of the combinatorial background for different decay channels are parameterized by the function A.11.

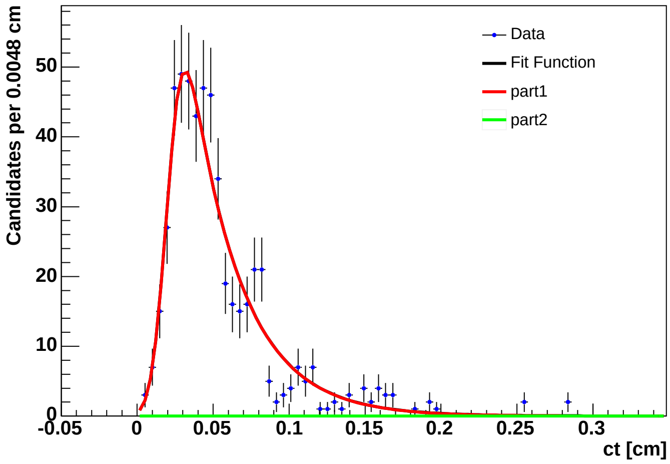

Figure B.7: Decay time distribution of the combinatorial background in the decay mode $B_s \to D_s\pi$, $D_s \to \phi\pi$

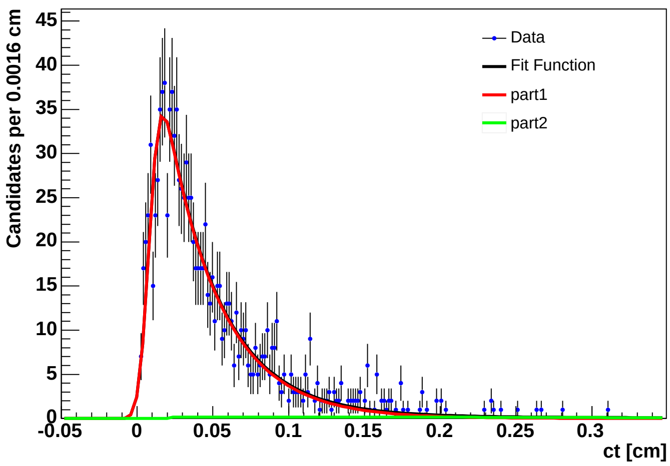

Figure B.8: Decay time distribution of the combinatorial background in the decay mode $B_s \to D_s\pi$, $D_s \to K^*K$

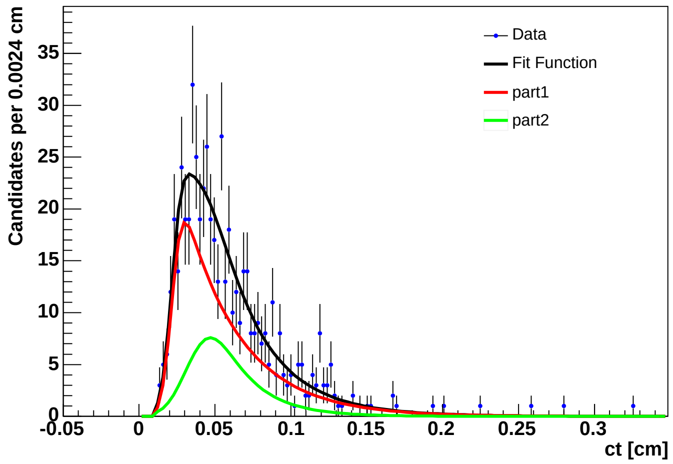

Figure B.9: Decay time distribution of the combinatorial background in the decay mode $B_s \to D_s\pi$, $D_s \to 3\pi$

Projections of the Likelihood in Mass and Lifetime Space

The projection of the complete likelihood function in mass space and lifetime space for different decay channels are presented in the following. The explanation and the used likelihood function can be found in section 4.5.

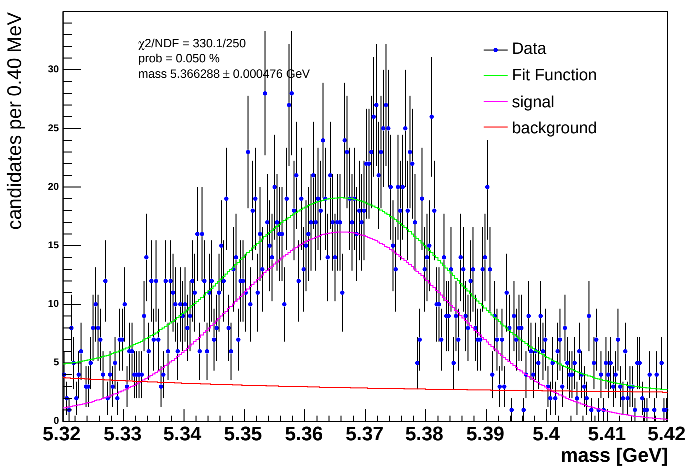

Figure B.10: Invariant mass spectrum in the signal range from 5.32 to 5.42 GeV of the decay mode $B_s \rightarrow D_s\pi$, $D_s \rightarrow \phi\pi$

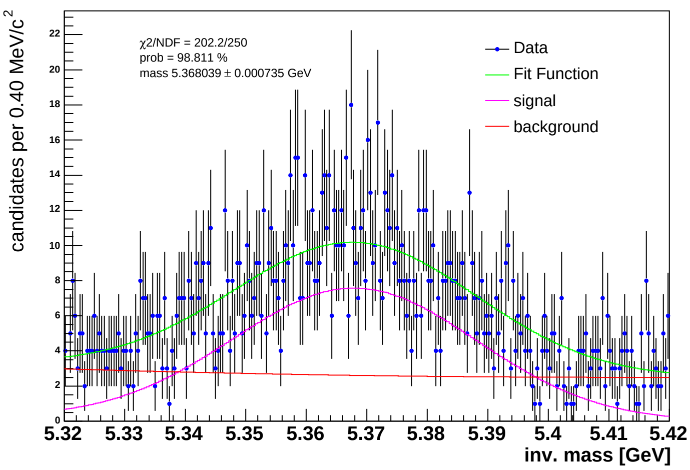

Figure B.11: Invariant mass spectrum in the signal range of the decay mode $B_s \rightarrow D_s\pi$, $D_s \rightarrow K^*K$

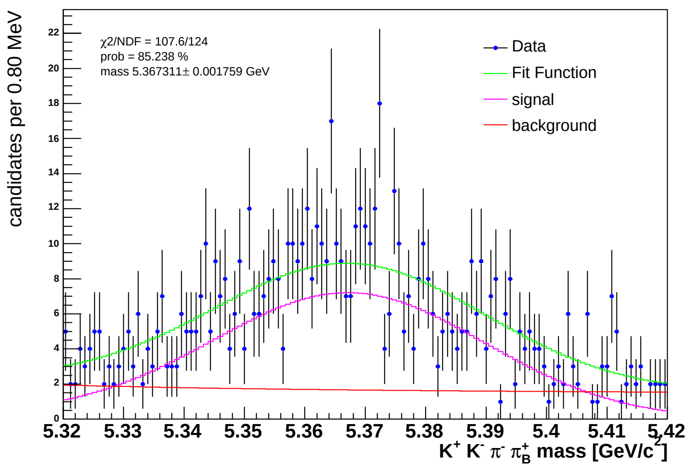

Figure B.12: Invariant mass spectrum in the signal range of the decay mode $B_s \to D_s\pi$, $D_s \to 3\pi$

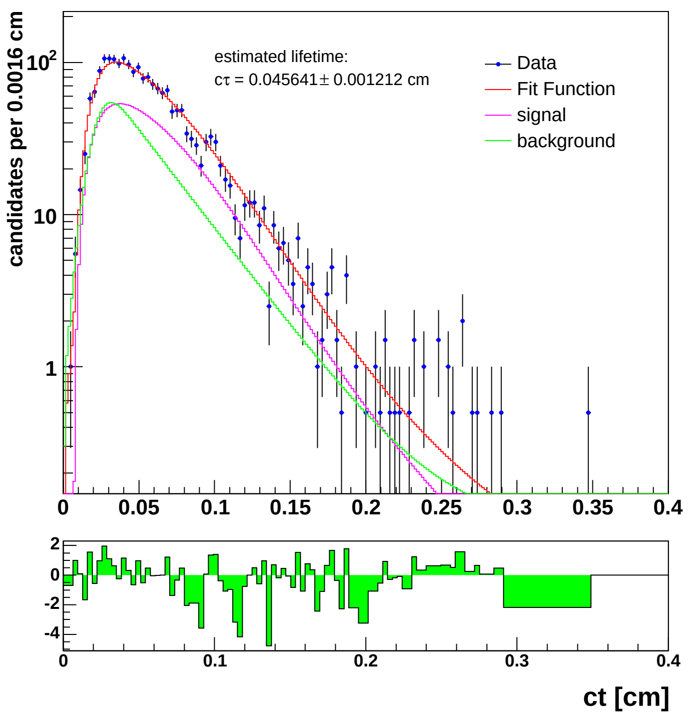

Figure B.13: Logarithmic scaled decay time spectrum of the decay mode $B_s \to D_s\pi$, $D_s \to \phi\pi$

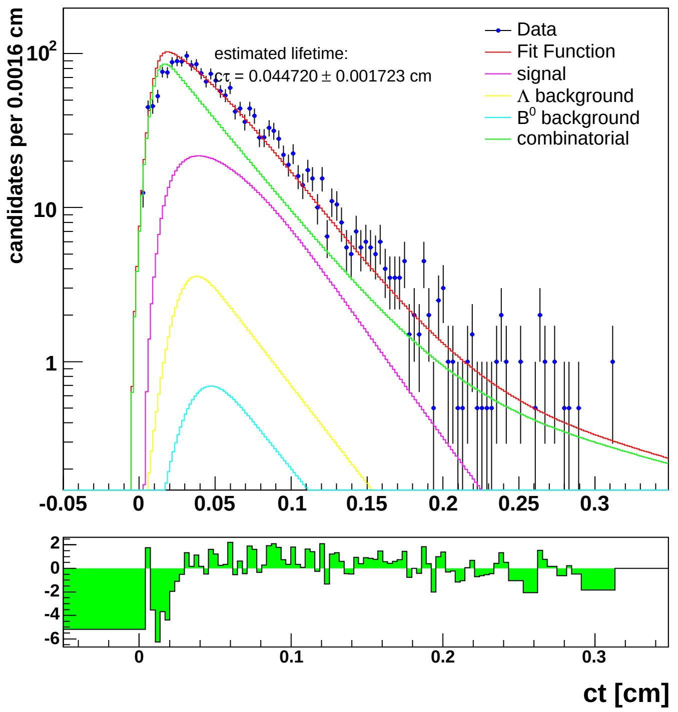

Figure B.14: Logarithmic scaled decay time spectrum of the decay mode $B_s \to D_s\pi$, $D_s \to K^*K$

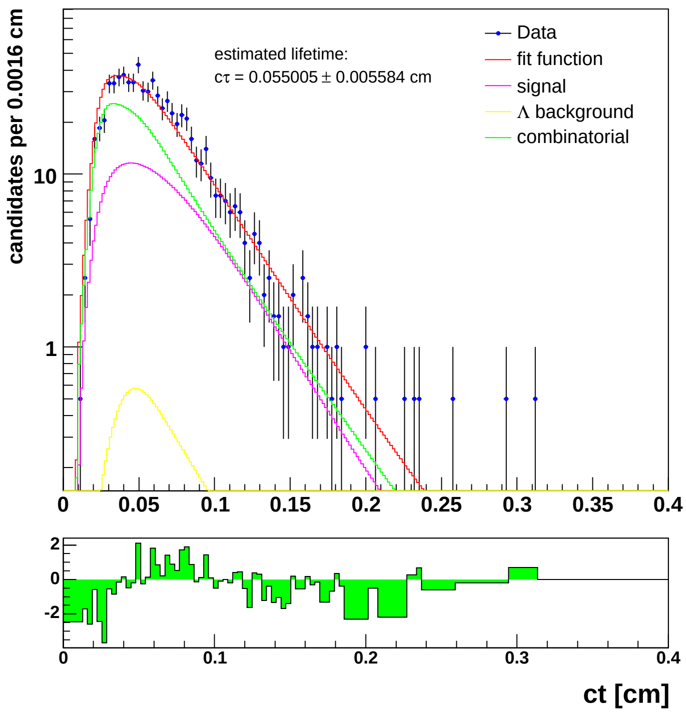

Figure B.15: Logarithmic scaled decay time spectrum of the decay mode $B_s \to D_s\pi$, $D_s \to 3\pi$

Appendix C

The Maximum Likelihood Method

The major task of this work is the estimation of parameters from measurements with errors. Therefore the maximum likelihood method [16] is used to estimate parameters of a probability density distribution.

C.1 Parameter Estimation

The estimation of parameters is a very general issue in science because of the typical situation: many measurements with errors of the same parameters but always different results. Different measurements generally do not give exactly the same results because of the finite instrument resolution and some other stochastic influences like quantum effects or thermal fluctuations. That is the reason for the statistical nature of experimental measurements. The frequency distribution of a sample of measured values behaves in case of $n \to \infty$ (n number of measurements) like the underlying probability density function multiplied with a normalization constant. The goal is to estimate the parameters of the probability density function in consideration of the errors of the measured data. It is also desirable that the method accomplishes the following criteria. It should be consistent ($\lim_{n\to\infty} \hat{a} = a_0$) if a_0 is the truth and $\hat{a}$ the estimate. It also should be unbiased in the sense that $E[\hat{a}] = a_0$, $E[\hat{a}]$ is the expectation of $\hat{a}$. Effectiveness and robustness play an eminent role, too.

Not all of these criteria can be satisfied simultaneously. The maximum likelihood method for example, is a very efficient estimation of parameters but is very unstable against choosing a wrong probability density function. Parameters can be biased and the likelihood method does not give any information about the fit quality. For achieving that, something like the χ^2 has to be calculated.

C.2 The Principle of the Maximum Likelihood Method

An essential assumption to use the maximum likelihood method is the knowledge about the (multidimensional) probability density function $f(\vec{x}|\vec{a})$ in $\vec{x}$ depending on a set of unknown parameters $\vec{a}$ which have to be estimated. With the set of n measurements $\vec{x}_i$, the likelihood function $L(\vec{a})$ is defined by

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

$L(\vec{a})$ can be interpreted as the probability to get these measured values $\vec{x}_i$ by a given choice of parameters $\vec{a}_i$. The best choice of parameters $\hat{\vec{a}}_i$ is the choice that maximizes the likelihood function $L(\vec{a})$ with the constraint that $f(\vec{x}|\vec{a})$ is normalized to a constant which is set w.l.o.g.[1] to one: $\int_{x_{min}}^{x_{max}} f(\vec{x}|\vec{a})dx = 1$, independent of $\vec{a}$.

The maximum of $L(\vec{a})$ can be found by requiring

$$\frac{\partial L(\vec{a})}{\partial a_k} = 0 \qquad \forall\, k$$

In practice $L(\vec{a})$ tends to very small values causing numerical instabilities because of finite accuracy of computers. Furthermore minimization algorithms are more common than maximization algorithms therefore usually the negative logarithm of the likelihood function is taken instead.

$$F(\vec{a}) = -\ln\left(L(\vec{a})\right)$$

As the logarithm is a monotonous function, the log-likelihood function has the extremum at the same place as the likelihood function.

C.3 Error Calculation

Often, especially in case of a large number of measurement $n \to \infty$, the likelihood function approaches a Gaussian distribution and the negative log-likelihood function $F(\vec{a})$ can be expanded around its minimum considering $\frac{\partial F(\vec{a})}{\partial a} = 0$ for $\vec{a} = \hat{\vec{a}}$.

$$\begin{aligned} F(a_1, a_2, ..., a_N) &\approx F(\hat{a}_1, \hat{a}_2, ..., \hat{a}_N) + \frac{1}{2}\sum_{i,k} \frac{\partial^2 F}{\partial a_i \cdot \partial a_k} (a_i - \hat{a}_i)(a_k - \hat{a}_k) + ... \\ &= F(\hat{a}_1, \hat{a}_2, ..., \hat{a}_N) + \frac{1}{2}\sum_{i,k} G_{ik} (a_i - \hat{a}_i)(a_k - \hat{a}_k) + ... \end{aligned}$$

$$\mathbf{V} = \mathbf{G}^{-1} \qquad with \qquad G_{ik} = \frac{\partial^2 F}{\partial a_i \cdot \partial a_k}$$

$\mathbf{G}$ has the form of a Hessian matrix and $\mathbf{V}$ can be interpreted as the covariance matrix in the asymptotic case, otherwise it is an approximation of the covariance matrix.

Due to the fact that the projection of the likelihood function on a single parameter space behaves like a Gaussian around the extremum, the log-likelihood function approaches a parabola. Subsequently it follows:

$$\frac{1}{(\sigma(\hat{a}))^2} = \left(\left.\frac{\partial^2 F}{\partial a^2}\right|_{\hat{a}} \right)$$

Where σ is the standard deviation of the approached Gaussian.

But also in the not asymptotic case the errors of each single parameter can be calculated considering the following. $F(\hat{\mathbf{a}})$ is the minimum concerning all parameters, then the (asymmetric) errors of a_i are defined:

$$F_{min}(\hat{a}_i + \sigma_r) = F(\hat{\mathbf{a}}) + \frac{1}{2}$$

$$F_{min}(\hat{a}_i - \sigma_l) = F(\hat{\mathbf{a}}) + \frac{1}{2}$$

σ is the distance one has to move along the x-axis until the negative log-likelihood function increased by $\frac{1}{2}$ compared to the minimum.

[1] without loss of generality

C.4 Binned Maximum Likelihood Fits

In case of many data points a histogram can be a adequate display format for a frequency distribution of measured data. The spectrum of occurred values is plotted on the x-axis which is divided in intervals called bins. The y-axis specifies the number of entries in each bin. A necessary task is to fit a probability distribution $f(x|a)$ on histogram content. The different bin contents are given by n_j with $j = 1, 2, ...J$ and $\sum_{j=1}^{J} n_j = n$ where n is the total number of entries. The numbers of entries in each bin are random numbers distributed according a Poisson distribution with the expectation μ_j. The probability density function of a Poisson distribution is

$$P(n_j|\mu_j) = \frac{\mu_j^{n_j} e^{-\mu_j}}{n_j!}$$

The expectation μ_j is determined for each bin by integrating the probability density function over the width of the bin and subsequent multiplying with n

$$\mu_j(a) = n \int_{bin\ j} f(x|a)dx$$

The negative log-likelihood function is given by

$$\begin{aligned} F(a) &= -\sum_{j=1}^{J} ln\left(\frac{\mu_j^{n_j} e^{-\mu_j}}{n_j!}\right) \\ &= -\sum_{j=1}^{J} n_j \ln \mu_j + \sum_{j=1}^{J} n_j \mu_j + \sum_{j=1}^{J} n_j \ln(n_j!) \end{aligned}$$

For $\mu >> 1$ and $\mu_j \approx n_j$ the Poisson distribution can be approximated by a Gaussian with the variance $\sigma_j^2 = \mu_j$. This has the advantage of an interpretable value in the minimum of the log-likelihood function transforming to

$$\begin{aligned} F(a) &= -\sum_{j=1}^{J} \ln\left(\frac{1}{\sqrt{2\pi}\sigma_j} e^{\frac{-(n_j-\mu_j)^2}{2\sigma_j^2}}\right) \\ &= \frac{1}{2}\sum_{j=1}^{J} \frac{(n_j - \mu_j)^2}{\mu_j} + const. \end{aligned}$$

In this case $2F(a)$ follows a χ^2-distribution with k degrees of freedom, i.e. the number of bins minus the number of fitted parameters. The value χ^2/k should be something around one for a reasonable fit. It is therefore an expedient information about the quality of the fit.

Bibliography

[1] P.A.M. Dirac, *The Quantum Theory of the Electron*, 1928, Proc. Roy. Soc. London A117, 610-624

[2] C.D. Anderson, *The positive electron.*, 1933, Phys. Rev. 43:491

[3] E.E. Lande et al., *Observation of Long-Lived Neutral V Particles*, 1956, Phys. Rev. 103:1901

[4] C. Albajar et al., *Search for $B^0 - \bar{B}^0$ Oscillations at the CERN $p - \bar{p}$ Collider*, 1987, Phys. Lett. B186:247

[5] H. Albrecht et al., *Observation of $B^0 - \bar{B}^0$ mixing*, 1987, Phys. Lett. B192:245

[6] The CKMfitter Group, www.slac.standford.edu/xorgckmfitter/ckm_results.html

[7] N. Cabibbo, *Unitarity Symmetry and Leptonic Decays*, 1963, Phys. Rev. Lett. 10

[8] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, 1973, Prog. Theor. Phys. 49, 652

[9] A. Rasin, *Diagonalization of Quark Mass Matrices and the Cabibbo-Kobayashi-Maskawa Matrix*, 1997, hep-ph/9708216

[10] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, 1983, Phys. Rev. Lett. 51

[11] The CKMFitter Group, *CP violation and the CKM matrix*, 2005, Eur. Phys. J. C41:1-131

[12] B. Aubert et al., *Measurement of the Cabibbo-Kobayashi-Maskawa Matrix Element $|V_{ub}|$ with $B \to \rho_e \nu$ Decays*, 2003, Phys. Rev. Lett. 90:181801

[13] S.B. Athar et al., *Study of the q^2-Dependence of $B \to \pi l \nu$ and $B \to \rho(\omega) l \nu$ Decay and Extraction of $|V_{ub}|$*, 2003, Phys. Rev. D68:072003

[14] M. Ciuchini, E. Franco, V. Lubicz, F. Mescia and C. Tarantino, *Lifetime differences and CP violation parameters of neutral B mesons at the next-to-leading order in QCD*, 2003, JHEP 08:031

[15] The CDF II Collabortation, *The CDF II Detector Technical Design Report*, 1996, FERMILAB-Pub-96/390-E

[16] V.Blobel and E.Lohrmann, *Statistische und numerische Methoden der Datenanlyse*, 1998, B.G. Teubner, Stuttgart

[17] D.Lucchesi, *The CDF II Secondary Vertex Trigger*, 2002, Proceedings of Beauty and CDF Note 6107.

[18] Christian Dörr, *Optimization of the Signal Selection of Exclusively Reconstructed Decays of B^0 and B_s Mesons at CDF II*, 2006, PhD thesis at the Institut für experimentelle Kernphysik at the University of Karlsruhe

[19] Philipp Mack *PhD thesis in process*

[20] Andreas Gessler, *Rekonstruktion der Zerfälle $B_s \rightarrow D_s\pi$ and $B_s \rightarrow D_s\pi\pi\pi$ mit dem CDF-II-Detektor*, 2006, Diploma thesis at the Institut für experimentelle Kernphysik at the University of Karlsruhe

[21] Andreas Schmidt, *Combined B Flavour Tagging with Artificial Neural Networks*, 2006, Diploma thesis at the Institut für experimentelle Kernphysik at the University of Karlsruhe

[22] C.Blocker, J.Boudreau, Ch. Paus, *Common Tools for B Physics Analyses in Run II at CDF*, 2002, CDF/DOC/BOTTOM/CDFR/5735

[23] M.Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, 2004, physics/0402093

[24] W.-M. Yao et al., *Journal of Physics*, G 33, 1 (2006)

[25] Rene Brun and Fons Rademakers, *ROOT - An Object-Oriented Data Analysis Fraemwork*, 1995-2006, http://root.cern.ch/

[26] N. Leonardo, *Methods for Likelihood Computation in Lifetime and Mixing Analyzes*, 2004, CDF Note 7385

[27] T. Sjostrand, L. Lonnblad and S. Mrenna, *Pythia 6.2: Physics and manual*, 2001, LU TP 01-21, hep-ph/0108264

[28] K. Anikeev, C. Paus, P. Murat, *Description of Bgenerator II*,1999, CDF/DOC/BOTTOM/CDFR/5092

[29] The B_s Mixing Group, *Monte Carlo Samples for Δm_s Measurement*, 2006

[30] The B_s Mixing Group, *Same Side Kaon Tagging Studies on Fully Reconstructed Decays*, 2005, CDF Note 7979

[31] D. J. Lange, *The Evtgen Particle Decay Simulation Package*, 2001, Nucl. Instrum. Meth., A462:152-155

[32] *The CDF Detector Simulation*, http://www-cdf.fnal.gov/cdfsim/cdfsim_main.html

[33] B.Ashmanskas and A. Cerri, http://cdfcodebrowser.fnal.gov/CdfCode/source/svtsim

[34] The CDF II Collabortation, *Observation of $B_s - \bar{B}_s$ Oscillations*, 2006, Phys. Rev. Lett. 97.242003

[35] Heavy Flavor Averaging Group, E.Barberio et al., hep-ex/0603003, update results available at http://www.slac.stanford.edu/xorg/hfag/

[36] The CDF Collaboration, *Study of $B_s \rightarrow D_s lX$ Oscillations in the Two-Track Trigger Data Sample*, 2005, CDF Note 7907

[37] H.-G.Moser and A.Roussarie, *Mathematical methods for $B^0 - \bar{B}^0$ oscillation analyses*, 1997, NIM A384:491-505

[38] Nuno T. Leonardo, *Analysis of B_s flavor oscillations at CDF*, 2006, PhD thesis at the MIT

Danksagung

Ich danke Herrn Prof. Dr. Michael Feindt, dass er mich in seiner Gruppe als Diplomand aufgenommen hat und für die hervorrangende Betreuung während dieser Zeit. Vielen Dank auch an Herrn Prof. Dr. Günter Quast für die Übernahme des Korreferats.

Außerdem danke ich allen Mitgliedern des Instituts für Experimentelle Kernphysik, vor allem aber den Leuten, mit denen ich das Büro teilte, für eine äußerst angenehme Arbeitsatmosphäre. Ein besonderer Dank geht an Dr. Michal Kreps, der die ganze Zeit ein offenes Ohr für kleinere und größere Problemchen hatte, Dr. Christian Dörr, der mir den Einstieg in die anfangs neue Thematik erleichterte, sowie Philipp Mack für die gute Zusammenarbeit und das Korrekturlesen dieser Arbeit.
Herzlichen Dank auch an das Admin-Team und allen Anderen, die sich an administrativen Aufgaben aller Art am Institut beteiligten.

Zum Schluß möchte ich meinen Eltern ganz besonders danken, die mich in der Zeit meines Studiums stets unterstützten. Meiner Freundin Clarissa danke ich für die ein oder andere Ablenkung und Überraschung in dieser Zeit.