Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

IEKP-KA/2008-24

Measurement of the Branching Fractions of $B_s \to D_s^{+(*)} D_s^{-(*)}$ Meson Decays at CDF–II

Dominik Emmanuel Horn

Diplomarbeit

an der Fakultät für Physik
der Universität Karlsruhe

Referent: Prof. Dr. M. Feindt
Korreferent: Prof. Dr. T. Müller

Institut für Experimentelle Kernphysik

August 2008

Deutsche Kurzfassung

Mit dem Versuch, in der Natur ablaufende Vorgänge zu verstehen und wissenschaftlich zu beschreiben, war stets das Bestreben verbunden, auf Grundlage empirischer Erfahrungen und Experimente erstellte Beschreibungsversuche der Natur zu harmonisieren und auf ein einheitliches naturwissenschaftliches Modell zurück zu führen. Das *Standardmodell der Elementarteilchenphysik* stellt einen extrem erfolgreichen Meilenstein auf dem Weg zu einem vereinheitlichten Modell dar. In den 60er und 70er Jahren des 20. Jahrhunderts in der Sprache der Quantenfeldtheorie aufgestellt, war und ist es Gegenstand einer Vielzahl aufwendiger Experimente. Bis zum heutigen Tag wurden darin eine Vielzahl der Vorraussagen des Standardmodells präzise bestätigt. Daher wird das Standardmodell heute gemeinhin als das beste Modell angesehen, das die Eigenschaften der Grundbausteine der Materie und der zwischen diesen stattfindenden Wechselwirkungen konsistent zu beschreiben vermag. Von den vier bekannten fundamentalen Wechselwirkungen – die starke, die schwache, die elektromagnetische und die gravitative Wechselwirkung – konnte die Gravitation jedoch bislang nicht ein allgemeingültiges Modell integriert werden.

Gemäß des Standardmodells sind die fundamentalen Bausteine der Materie *Fermionen*. Diese Elementarteilchen werden hinsichtlich ihrer Möglichkeiten, zu interagieren, in zwei Klassen eingeteilt – *Quarks* und *Leptonen* – die wiederum entsprechend einer bestimmten Ladung, genannt *flavor*, in jeweils sechs verschiedene Arten kategorisiert werden. Im Quark–Sektor sind dies die *flavors* up, down, charm, strange, top und bottom. Leptonen setzen sich aus den drei elektrisch geladenen Elementarteilchen Elektron, Myon und Tau sowie den drei neutralen Leptonen Elektron–, Myon– und Tauneutrino zusammen. Die Wechselwirkungen, an denen Quarks und Leptonen teilnehmen, werden durch Austauschteilchen, den sog. *Bosonen* vermittelt.

Quarks kommen niemals in isolierter, sondern stets in gebundener Form vor. Diese gebundenen Zustände werden als *Hadronen* bezeichnet, wobei bis heute ausschließlich aus drei Quarks bestehende Zustände (*Baryonen*) und aus einem Quark und seinem Antiteilchen, dem Antiquark, aufgebaute Zustände (*Mesonen*) beobachtet wurden. Das Proton ist das einzige Hadron, von dem angenommen wird, dass es stabil ist. Alle anderen Hadronen zerfallen nach einer gewissen Zeitdauer ohne äußeren Einfluss in weitere Hadronen und Leptonen. Hierfür gibt es eine Vielzahl an Zerfallsmöglichkeiten. Die Eigenschaften von Teilchen und deren Zerfallsprodukte zu studieren ist ein essentielles Instrument, um Vorhersagen zu überprüfen und Parameter des Standardmodells

präzise zu bestimmen. Einige der Teilchen und Teilchenzerfälle, auch – oder gerade – solche bei hohen Energien, können durch Messungen an der natürlichen kosmischen Strahlung studiert werden. Um seltene Teilchen bei hinreichend hoher Statistik untersuchen zu können, müssen enorme experimentelle Anstrengungen unternommen werden, da diese Teilchen künstlich erzeugt werden müssen. Dies geschieht in sog. Teilchenbeschleunigern, in denen Elektronen, Wasserstoffkerne oder Ionen auf hohe Energien beschleunigt und mit einem ruhenden Target oder ebenfalls beschleunigten Teilchen zur Kollision gebracht werden. Aus der Bewegungsenergie werden neue Teilchen erzeugt, deren Zerfallsprodukte mit Hilfe komplexer Detektoren nachgewiesen und identifiziert werden.

Das *Tevatron* am *Fermi National Accelerator Laboratory* (Fermilab), etwa 70 km westlich von Chicago im US–Bundesstaat Illinois gelegen, ist eine derartige experimentelle Anlage. In diesem Ringbeschleuniger, der einen Radius von 1 km aufweist, werden Protonen (p) und Antiprotonen ($\bar{p}$) entgegengesetzt durch eine Strahlachse geführt, vermittels supraleitender Magnetspulen auf eine Energie von jeweils 980 GeV/c^2 beschleunigt und an zwei Wechselwirkungspunkten bei einer Schwerpunktsenergie von 1.96 TeV/c^2 zur Kollision gebracht. Bis zum Start des *Large Hadron Collider* (LHC) am CERN (Conseil Européen pour la Recherche Nucléaire) in wenigen Wochen ist dies die höchste an einem erdgebundenen Teilchenbeschleuniger realisierte Wechselwirkungsenergie. An einem der Kollisionspunkte befindet sich der *Collider Detector at Fermilab* (CDF). Bei diesem etwa 12 m hohen, in guter Näherung zylinderförmigen Universaldetektor sind die Detektorkomponenten für die Impulsmessung, Spurerkennung, Kalorimetrie sowie Myonerkennung von innen nach außen schalenförmig um den nominellen Wechselwirkungspunkt angeordnet. Um den Datenfluss zu reduzieren, ist online ein dreistufiges Triggersystem implementiert, das anhand vordefinierter Kriterien entscheidet, welches Ereignis gespeichert wird. Spurrekonstruktion und Teilchenidentifikation werden durch Anwendung verschiedener, in Offline–Software enthaltener Algorithmen realisiert. Die vorliegende Analyse greift auf Daten zurück, die im Zeitraum Februar 2002 bis Januar 2007 am CDF–Detektor in der zweiten Ausbaustufe (CDF–II) genommen wurden. Neben experimentellen Daten sind für viele Analyseschritte simulierte Daten von großer Wichtigkeit. Diese werden mit sog. Monte–Carlo–Methoden erzeugt. Beschleunigungsvorgang, Detektorkomponenten, Datennahme und Ereignisrekonstruktion werden in den Kapiteln 2 und 4 der vorliegenden Arbeit diskutiert.

B–Mesonen, also Quark–Antiquark–Zustände, die aus einem bottom Quark b und einem leichteren Quark aufgebaut sind, haben im Zusammenhang mit der Überprüfung von Standardmodellvorhersagen in jüngerer Zeit große Bedeutung erlangt. Gegenstand der vorliegenden Arbeit ist das B_s Meson, das neben dem Anti–bottom Quark $\bar{b}$ das strange Quark s enthält. Wie von theoretischer Seite gezeigt wurde [1], könnte dieses Meson die Möglichkeit bieten, physikalische Effekte über das Standardmodell hinaus zu beobachten.

Diese Vermutung ist durch die inhärenten physikalischen Eigenschaften des B_s–Mesons motiviert. Wie bei allen neutralen Mesonen findet auch beim B_s–Meson eine

permanente Oszillation zwischen den Flavor–Eigenzuständen B_s und $\bar{B}_s$ statt. Diese B_s–$\bar{B}_s$–Mischung ist darauf zurückzuführen, dass die quantenmechanischen Flavor–Eigenzustände B_s und $\bar{B}_s$ nicht mit den Massen–Eigenzuständen B_s^L und B_s^H zusammenfallen. Vielmehr werden diese Massenzustände durch Überlagerung der Flavor–Eigenzustände gebildet. Die Massendifferenz der Masseigenzustände bestimmt dabei die Frequenz der Oszillation, die 2006 erstmals durch die CDF–II–Kollaboration vermessen werden konnte [2]. Im Rahmen des Standardmodells wird angenommen, dass die B_s–$\bar{B}_s$–Mischung eine große Zerfallsbreitendifferenz $\Delta\Gamma$ verursacht. Ferner sind die Zustände B_s und $\bar{B}_s$ nicht invariant gegenüber der aufeinanderfolgenden Wirkung des Ladungs–Operators C (*charge*) und des Paritäts–Operators P (*parity*). Wohl aber sind B_s^{gerade} und $B_s^{ungerade}$, bestimmte Überlagerungen der Flavor–Eigenzustände B_s und $\bar{B}_s$, Eigenzustände von CP. Der Zusammenhang zwischen den relativen CP– und Massenzerfallsbreiten ist dabei gegeben durch $\Delta\Gamma/\Gamma = \Delta\Gamma_{CP}/\Gamma \cdot \cos\phi$, wobei ϕ als die Phase zwischen B_s–Mischung und –Zerfall identifiziert wird. Gemäß des Standardmodells ist die Phase ϕ verschwindend klein, was mit einer verschwindenden CP–Verletzung in der Interferenz von Mischung und Zerfall im B_s–$\bar{B}_s$–System korrespondiert. Diese Art der CP–Verletzung beschreibt das Phänomen, dass sich die Betragsamplitude der Interferenz eines gemischten B_s–Mesonzerfalls mit einem ungemischten B_s–Mesonzerfall in einen CP–Eigenzustand sich von der eines $\bar{B}_s$ unterscheidet. Im Standardmodell, mit $\phi = 0$, fallen die relativen Zerfallsbreiten $\Delta\Gamma/\Gamma$ und $\Delta\Gamma_{CP}/\Gamma$ zusammen. Der Nachweis einer nicht–verschwindenden Phase ϕ wäre ein unmissverständliches Zeichen für neue Physik jenseits des Standardmodells.

Es gibt unterschiedliche Ansätze, sich der aufgezeigten Problemstellung zu nähern und mögliche physikalische Szenarien im B_s–System einzuschränken. Eine Möglichkeit, die aufgezeigt wurde [1], ist die direkte Schätzung einer unteren Schranke für $\Delta\Gamma_{CP}/\Gamma$ vermittels einer Messung des Verzweigungsverhältnisses des Zerfalls $B_s \to D_s^+ D_s^-$. Diese direkte Abschätzung wird dadurch ermöglicht, dass der Endzustand $D_s^+ D_s^-$ ein reiner CP–gerader Eigenzustand ist. Es wird zudem angenommen, dass dieser Zerfall in hohem Maße zur Zerfallsbreitendifferenz im B_s–System beiträgt. Kapitel 3 gibt eine kurze theoretische Einführung in den aufgezeigten Problemkreis.

Der Zerfall $B_s \to D_s^+ D_s^-$ auf Grundlage einer exklusiven Rekonstruktion wurde erstmals von der CDF–II-Kollaboration im Jahr 2006 beobachtet [3]. Dabei wurde ein Verzweigungsverhältnis von $Br[B_s \to D_s^+ D_s^-] = (9.4^{+4.4}_{-4.2}) \times 10^{-3}$ gemessen. Die Berechnung erfolgte über die Messung des relativen Verzweigungsverhältnisses $f_{B_s} = f_s/f_d \cdot Br[B_s \to D_s^+ D_s^-]/Br[B^0 \to D_s^+ D^-]$, mit dem s– zu d Quark–Produktionsverhältnis f_s/f_d. Für das Verzweigungsverhältnis $Br[B^0 \to D_s^+ D^-]$ wurde der aktuelle PDG–Wert (*Particle Data Group*, [4]) eingesetzt.

Die Messung des relativen Verzweigungsverhältnisses f_{B_s} und daraus die Berechnung des Verzweigungsverhältnis $B_s \to D_s^+ D_s^-$ war auch Ziel der vorliegenden Analyse. Wie im Falle der bereits existierenden Messung werden die Mesonzerfälle $B_s \to D_s^+ D_s^-$ und $B^0 \to D_s^+ D^-$ aus den hadronischen Zerfällen $D_s^+ \to \phi\pi^+$, $D_s \to K^{0*} K^+$ und $D_s \to \pi^+\pi^-\pi^-$ bzw. $D \to K^+\pi^-\pi^-$ rekonstruiert, wobei $\phi \to K^+K^-$ und $K^{0*} \to K^+\pi^-$. Allerdings werden nun bei der Rekonstruktion von $D_s^+ D_s^-$ auch

Kombinationen ohne $\phi\pi^+$ im Endzustand herangezogen. Für die Selektion von Signalereignissen in den sechs B_s– bzw. drei untersuchten B^0–Kanälen werden nach der Anwendung weicher Vorschnitte künstliche neuronale Netze eingesetzt. Die Anzahl der Signalereignisse wird mit Hilfe eines Fits des kanalspezifischen invarianten Massenspektrums auf Grundlage der erweiterten ungebinnten *Maximum–Likelihood–Methode* bestimmt. Alle ungebinnten Maximum–Likelihood–Fits werden innerhalb des *Fitter Framework*, einer flexiblen Software–Umgebung für unterschiedliche Typen von Datenanpassungsprozessen, durchgeführt. Die hierfür vorgenommenen Schritte, insbesondere das Training der neuronalen Netze und die Aufstellung der Fit–Modelle, sowie die gemessenen Ereigniszahlen präsentiert Kapitel 4. Die Analyse der drei B_s–Zerfallskanäle, die nicht über $D_s \to \phi\pi$ rekonstruiert werden, erwies sich aufgrund der Beimischung unterschiedlicher Reflektionen als sehr schwierig. Daher wurden diese bei den folgenden Analyseschritten nicht weiter berücksichtigt. Unter Reflektionen versteht man Beiträge fremder, aber hinsichtlich der Zerfallstopologie ähnlicher Moden, die infolge einer falschen Massenhypothese bei der Zerfallsrekonstruktion das gegebene Massenspektrum verunreinigen.

Das Fitter–Framework erlaubt es, mehrere Massenspektren mit der Maximum–Likelihood–Methode simultan zu fitten und hierbei bestimmte Fitparameter durch identische Namensgebung zu teilen. Diese Möglichkeit wird bei der Parameterschätzung von f_{B_s} für unterschiedliche Kombinationen von jeweils einem $B_s \to D_s^+ D_s^-$– und einem $B^0 \to D_s^+ D^-$–Kanal ausgenutzt (siehe Kaptitel 5). Durch Mittelung der aus den Simultan–Fits resultierenden Stichproben $f_{B_s}^i$ und durch Einsetzen der Werte für f_s/f_d und $Br[B^0 \to D_s^+ D^-]$ ergibt sich:

$$\begin{aligned} Br[B_s \to D_s^+ D_s^-] &= \{10.5 \pm 1.2(Stat) \pm 0.8(Br) \pm 1.2(Br_{Corr}) \\ &\quad \pm 1.6(f_s/f_d) \pm 2.1(Br_{B^0})\} \times 10^{-3} \\ &= (10.5 \pm 3.2) \times 10^{-3} \end{aligned}$$

Innerhalb der Unsicherheiten stimmt dieses Ergebnis mit dem bereits existierenden Messwert $Br[B_s \to D_s^+ D_s^-] = (9.4^{+4.4}_{-4.2}) \times 10^{-3}$ [3] gut überein. Jedoch enthalten die angegebenen Messfehler der vorliegenden Arbeit nur statistische Unsicherheiten (inklusive der Unsicherheiten durch die kombinierten Rekonstruktions– und Selektionseffizienzen) sowie Unsicherheiten bereits gemessener Verzweigungsverhältnisse. Eine gründliche Untersuchung systematischer Effekte sowie der bereits erwähnten Reflektionen in einigen Zerfallsmoden wird zentraler Bestandteil einer zukünftigen Analyse sein müssen. Ferner könnte die Präzision des Verzweigungsverhältnisses $Br[B_s \to D_s^+ D_s^-]$ von einem alternativen Normierungskanal in $f_{B_s} = f_s/f_d \cdot Br[B_s \to D_s^+ D_s^-]/Br[B^0 \to D_s^+ D^-]$ profitieren. Vorstellbar wären hier beispielsweise die Zerfälle $B_s \to D_s \pi$ oder $B_s \to \pi\pi\pi$.

Im Zuge der Analyse wurden ferner die relativen Verzweigungsverhältnisse $Br[B_s \to D_s^{(*)+} D_s^{(*)-}]/Br[B_s \to D_s^+ D_s^-]$ und $Br[B^0 \to D_s^{(*)+} D^{(*)-}]/Br[B^0 \to D_s^+ D^-]$ gemessen. Die Zerfälle $B_s \to D_s^{(*)+} D_s^{(*)-}$ bzw. $B^0 \to D_s^{(*)+} D^{(*)-}$ erscheinen in den Massenspektren als sog. Satelliten–Peaks etwa ein bis zwei Pion–Massen unterhalb der invari-

anten B_s– bzw. B^0–Masse und werden als partiell rekonstruierte Zerfälle bezeichnet, da das Photon bzw. das neutrale Pion aus dem $D^*_{(s)}$–Zerfall im Detektor nicht erkannt wird. Die Messergebnisse für $Br[B_s \to D_s^{(*)+} D_s^{(*)-}]/Br[B_s \to D_s^+ D_s^-]$ decken sich innerhalb der Fehler gut mit den im PDG angegebenen Werten, die Messung der relativen Verzweigungsverhältnisse $Br[B_s \to D_s^{(*)+} D_s^{(*)-}]/Br[B_s \to D_s^+ D_s^-]$ stellt die erste dieser Art im B_s–Mesonsektor dar. Es wurden folgende Werte ermittelt:

$$\frac{Br\,(B_s \to D_s^{+*} D_s^-)}{Br\,(B_s \to D_s^+ D_s^-)} = 2.87 \pm 0.52$$

$$\frac{Br\,(B_s \to D_s^{+*} D_s^{-*})}{Br\,(B_s \to D_s^+ D_s^-)} = 3.44 \pm 0.71$$

$$\frac{Br\,(B^0 \to D_s^+ D^{-*})}{Br\,(B^0 \to D_s^+ D^-)} = 0.91 \pm 0.14$$

$$\frac{Br\,(B^0 \to D_s^{+*} D^-)}{Br\,(B^0 \to D_s^+ D^-)} = 1.12 \pm 0.07$$

$$\frac{Br\,(B^0 \to D_s^{+*} D^{-*})}{Br\,(B^0 \to D_s^+ D^-)} = 2.94 \pm 0.20$$

Um einen Vergleich zwischen den drei B^0– und den für das B_s–System aufgeführten Verzweigungsverhältnissen zu gestatten, können die relativen Verzweigungsverhältnisse der Zerfallsmoden $B^0 \to D_s^+ D^{-*}$ und $B^0 \to D_s^{+*} D^-$ in grober Vereinfachung zu einem Verzweigungsverhältnis zusammengefasst werden. In dieser Näherung fallen die relativen Verzweigungsverhältnisse $B_s \to D_s^{+*} D_s^{(-*)}$– und $B^0 \to D_s^{(+*)} D^{(-*)}$–Zerfällen sehr ähnlich aus. Zukünftige Messungen mit höherer Statistik werden sicher zu einer Verbesserung der Präzision der Messung von Verzweigungsverältnissen im B_s–Sektor beitragen.

Abschließend wird als Gegenprobe zu Methode und Messwerten ein alternatives Fit–Verfahren vorgeschlagen, das die gegebenen Informationen auf eine leicht abweichende Art und Weise nutzt. Eine einfache Modifikation der Fitfunktionen für die Massenspektren der Zerfälle $B_s \to D_s^+ D_s^-$ und $B^0 \to D_s^+ D^-$ erlaubt es, alle für die Messung von Verzweigungsverhältnisen berücksichtigten B_s– und B^0–Zerfallskanälle in einem einzigen simultanen Massenfit zu vereinigen und damit die in dieser Arbeit bestimmten Größen auf Grundlage einer Messung zu überprüfen. Hierbei werden die zuvor gewonnenen Messergebnisse teils reproduzert oder zumindest gut bestätigt. Die alternative Methode soll jedoch in erster Linie die Durchführbarkeit eines derartigen Simultan–Fits belegen und als Vorschlag eines alternativen Ansatzes verstanden werden. Systematische Effekte und insbesondere Korrelationen zwischen Eingangsparametern werden weiterer Verfeinerungen der Methode bedürfen.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The variety of phenomena occurring in the world surrounding us always has been stirring up curiosity of men. What is matter composed of, which are the elementary forces appearing in the universe, ever have been the fundamental questions human beings sought answers for. Based on empirical observations of nature and on experiments becoming more and more complex in the course of time, a variety of models concerning the structure of matter have been conceived. The idea of matter consisting of elementary constituents that cannot be further subdivided, which were called *Atoms* (gr. $\tau\varepsilon\mu\nu\varepsilon\iota\nu$ – to cut, $\alpha\tau o\mu o\varsigma$ – not cuttable), already emerged in the ancient Greece.

Today we know that Atoms are far from being the smallest constituents of matter. The *atomic nucleus*, which is orbited by electrons, is built up of *nucleons* – protons and neutrons – which in turn are composed of even smaller particles, the *quarks*. According to our current understanding, quarks and *leptons*, which are referred to as *fermions*, are the elementary particles matter consists of. Quarks bear electric charge and appear in six different species, called *flavors*, which are referred to as up, down, charm, strange, top and bottom. The other class of elementary particles, the leptons, appear in three charged flavors – electron, myon and tau – and three electrically neutral flavors, called electron, myon and tau neutrino. The elementary particles interact with each other by exchanging particles, called *bosons*. The properties of these elementary particles and the interactions among them are described in the Standard Model of elementary particle physics which was developed in the sixties and seventies of the 20th century. Since then, the Standard Model has been subject to various experimental tests and a wealth of the predictions proved to be correct. Therefore, the Standard Model is commonly accepted as a very successful theory that consistently describes the properties of particles and three out of the four fundamental interactions – the strong, the weak and the electromagnetic interaction – known today. Up to now, the forth elementary interaction, gravitation, has not been integrated into one all–embracing theory yet.

Quarks occur in bound states only, called *hadrons*, which in turn are categorized into *baryons* consisting of three quarks, and *mesons* which are composed of one quark and its quark anti–partner, called anti–quark. Apart from the proton, hadrons are

not stable but decay into different hadrons or leptons. For these decay modes there exists a multitude of possibilities. Measuring the properties and decay products of particles provides the opportunity to test the predictions and determine parameters of the Standard Model, or to seek for new physical contributions beyond this theoretical framework. For this task, enormous experimental efforts are required because most of the particles or decay events occur only rarely and have to be generated in artificial earth–bound particle accelerators operating at high energies. The *Tevatron* at the *Fermi National Accelerator Laboratory* (Fermilab) in Illinois, USA, is such an artificial particle accelerator. Inside this ring accelerator protons and antiprotons circulate in opposite directions and collide at two interaction points at a center of mass energy of 1.96 TeV/c^2, resulting in the production of new particles whose decay products are measured by two distinct particle detector devices, namely the *DØ* detector and the *Collider Detector at Fermilab* (CDF). This analysis uses data taken during operation period Run–II of the CDF experiment. Chapter 2 will outline the basic functions of the Tevatron ring accelerator and the individual detector components of the CDF–II detector.

The particle being subject to this thesis is the neutral B_s meson, which consists of the quark anti–partner of the bottom quark b, called $\bar{b}$ and one strange quark s. As all neutral mesons the B_s has the property to oscillate into its antiparticle $\bar{B}_s$ and vice versa, a phenomenon which is referred to as B_s–$\bar{B}_s$ meson mixing. The frequency of this mixing process, which has been measured by the CDF–II collaboration in 2006 [2], is determined by the mass difference of the two mass eigenstates B_s^L and B_s^H which are two distinct superpositions of the flavor eigenstates B_s and $\bar{B}_s$. In the theoretical framework of the Standard Model, the CP eigenstates B_s^{even} and B_s^{odd} are introduced which are superpositions of the flavor eigenstates and invariant under the consecutive charge (C) and parity (P) transformation, except for the sign of the eigenvalue in the case of B_s^{odd}. In the Standard Model, the relative decay width difference of the mass eigenstates, $\Delta\Gamma/\Gamma$, equals the relative CP width difference, $\Delta\Gamma_{CP}/\Gamma$, multiplied by the factor $\cos\phi$. The angle ϕ is identified as the phase between B_s mixing and decay. In the Standard Model the phase ϕ is expected to be close to zero, which corresponds to vanishing CP violation in the interference between mixing and decay. In this special case CP violation describes the phenomenon that the absolute amplitude of the interference between a mixed and an unmixed B_s decay into a CP final state does not equal the corresponding one of the $\bar{B}_s$ decay into a final state having the opposite CP content. A vanishing phase ϕ has the consequence that $\Delta\Gamma/\Gamma$ and $\Delta\Gamma_{CP}/\Gamma$ on the one hand, and the CP and mass eigenstates on the other hand coincide. However, the observation of a sizable CP violating phase ϕ in the B_s–$B\bar{B}_s$ system would clearly indicate physics beyond the Standard Model.

There are different approaches [1] that allow to make inferences to one of the quantities of the relation $\Delta\Gamma/\Gamma = \Delta\Gamma_{CP}/\Gamma \cdot \cos\phi$. One important input is given by the measurement of $\Delta\Gamma_{CP}/\Gamma$. It has been shown [1] that the measurement of the branching fraction of the decay $B_s \to D_s^+ D_s^-$ provides for a direct estimation of the lower bound of $\Delta\Gamma_{CP}/\Gamma$. Directly inferring a lower bound for $\Delta\Gamma_{CP}/\Gamma$ is possible

because $D_s^+ D_s^-$ is a purely CP even state. Furthermore, the decay $B_s \to D_s^+ D_s^-$ is expected to place the main contribution to the decay width difference in the B_s–$\bar{B}_s$ system which is predicted to be sizable in the Standard Model. The underlying theoretical framework needed for a basic understanding will be discussed in Chapter 3.

In this analysis $Br[B_s \to D_s^+ D_s^-]$ is determined by measuring the relative branching fraction $Br[B_s \to D_s^+ D_s^-]/Br[B^0 \to D_s^+ D^-]$. B^0 denotes the neutral B_d meson which consists of one $\bar{b}$ and a d quark. The decay $B_s \to D_s^+ D_s^-$ is reconstructed in six hadronic decay channels, resulting from the different combinations of the decays $D_s \to \phi\pi$, $D_s \to K^{0*}K$ and $D_s \to \pi\pi\pi$. $B^0 \to D_s^+ D^-$ is analyzed in three decay modes where the D_s meson decays the same way and the D meson is reconstructed from the final state $K\pi\pi$.

The first observation of the exclusive decay $B_s \to D_s^+ D_s^-$ was reported by the CDF collaboration in the end of 2006 [3]. This measurement applied cut based signal selection techniques. For the selection of B_s and B^0 mesons this thesis makes extensive use of artificial neural networks. Particle reconstruction and selection techniques as well as the number of signal events obtained in the studied decay channels will be presented in Chapter 4. Chapter 5 will focus on the method applied to measure $Br[B_s \to D_s^+ D_s^-]$ by means of a combined decay channel fit in mass space. As a byproduct of the fit on the invariant mass spectra further relative branching fractions in $B_s \to D_s^{(*)+} D_s^{(*)-}$ and $B^0 \to D_s^{(*)+} D^{(*)-}$ decays are measured. The last Chapter will briefly summarize the results of this analysis and provide an outlook on possible measures that could improve the precision of the $Br[B_s \to D_s^+ D_s^-]$ measurement in the future.

Chapter 2

Experimental Setup

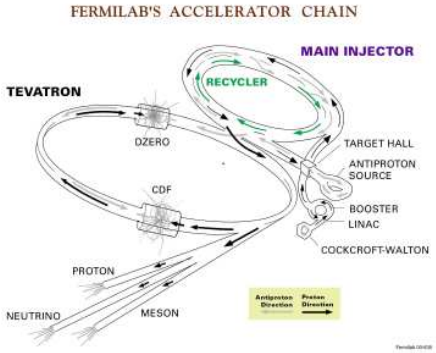
2.1 The CDF–II Experiment at Fermilab

Studies in the field of high energy physics in general and in terms of heavy meson physics in particular require eminent experimental endeavors since requested events usually occur only rarely. Data with reasonable statistics are provided by means of artificial earth–bound particle accelerators and detector devices. The presented analysis of the $B_s \to D_s D_s$ meson decay uses data provided by the CDF–II experiment at the *Fermi National Accelerator Laboratory* (Fermilab), located in Batavia 70 km west of Chicago (Illinois, USA). The Fermilab is host to the *Tevatron*, a $p\bar{p}$–collider which accelerates proton and antiprotons to a center of mass energy of $\sqrt{s} = 1.96$ TeV/c^2, currently representing the highest accessible collision energy achieved by an earth–bound particle accelerator. At the Tevatron p and $\bar{p}$ bunches interact at two collision points where the two particle detectors *DØ* and the *Collider Detector at Fermilab* (CDF) are situated. Figure 2.1(a) shows an aerial photo of the Fermilab facilities.

The Tevatron commenced operations in 1985 with a center of mass energy of 1.8 TeV/c^2. Run I was terminated after an operation time of 11 years. After the shutdown in 1996 major upgrades where applied to the accelerator complex and the two detectors DØ and CDF. The second period of operation, called Run II, was started in the end of 2001.

While primarily designed for being capable of performing studies in the high p_T sector, the research facilities DØ and CDF at Fermilab provided and still provide important insights into the field of heavy flavor physics. The discovery of the top quark (1995) [5, 6], complementing the picture in the quark sector and thus fortifying the pillars of the Standard Model, represents one of the famous achievements. Recent important achievements comprise the measurement of B_s oscillations in 2006 [2] and the observation of single top quark production in 2007 [7, 8].

(a) Aerial photo of the Fermilab accelerator complex



(b) Scheme of the accelerator complex

Figure 2.1: The accelerator complex at Fermilab.

2.2 The Accelerator Complex

This section outlines the technical devices and procedures used to produce protons and antiprotons, accelerate them to the final beam energy of 980 GeV/c^2 and realize the collision of the $p\bar{p}$ bunches at the desired interaction points. All these procedures are implemented within the Fermilab accelerator complex by means of sophisticated techniques. In figure 2.1(b) a schematic sketch of the particle production and accelerator chain at Fermilab is shown.

2.2.1 Proton Production and Pre–Acceleration

The starting point of the acceleration process is the *Cockcroft–Walton Pre–Accelerator*. Inside this device hydrogen gas is ionized by surface ionization effects. The negatively charged H^- ions are accelerated to a kinetic energy of 750 keV/c^2 by means of a positive high voltage field and injected into the *LINAC*, a linear accelerator consisting of cavities with oscillating radio frequency (RF) fields. Having traversed a distance of 150 m inside the LINAC the ions reach their final energy of 400 MeV/c^2 and are arranged into bunches. Before entering the *Booster* the hydrogen bunches pass a carbon foil which strips off the electrons from the H^- ions, leaving bunches only consisting of bare protons.

The *Booster*, a circular synchrotron with a diameter of approximately 150 m, represents the next stage of the accelerator chain. With each revolution the protons gain an increment of 750 keV/c^2 in kinetic energy as a result of a series of kicks from RF cavities. Additionally the intensity of the proton beam is gradually increased by repeatedly injecting further protons. After up to 20,000 revolutions the protons attain an energy of 8 GeV/c^2 and are transferred to the *Main Injector*.

2.2.2 The Main Injector

This 3 km circumference synchrotron covers two essential functions: the acceleration of protons coming from the *Booster* on the one hand, and the production and acceleration of antiprotons on the other hand. Therefore the proton beam is split up into two beams. The first one, consisting of protons with a kinetic energy of 150 GeV/c^2, is directly routed into the *Tevatron*. The second one, comprising protons with an energy of 120 GeV/c^2, are extracted from the *Main Injector* and routed towards the *Antiproton Source*. Here the protons collide with a nickel target, generating a shower of secondary particles mainly consisting of protons and pions which are focused into a beam by a lithium lens. The particles of interest, the antiprotons, are only produced to a very small extent though, leading to high production times for the desired quantity of antiproton bunches containing a sufficient number ($\approx 5 \times 10^{10}$) of particles. Due to this limiting factor to a high luminosity operation, the accelerator complex features several techniques and devices to maximize antiproton yields.

The separation of the antiprotons from other particle species is accomplished by a pulsed magnet working with a magnetic field strength of 760 T, selecting 8 GeV/c^2 antiprotons. Because of the arrangement in particle bunches the extracted antiprotons exhibit a large spread in energy. Debunching of the antiproton beam, corresponding to an expansion in space, and thus reducing the spread in momentum space is performed inside the *Debuncher*. Following the debunching process the antiprotons are piled up and stored inside the *Accumulator Synchrotron*. In order to minimize the antiproton loss rate, stochastic and electron cooling techniques are employed to cool down the antiproton beam, thereby confining the beam to a smaller volume in phase space.

Once enough antiprotons are accumulated, they are directed back contrariwise to the rotational direction of the proton beam into the *Main Injector* and accelerated to an energy of 150 GeV/c^2.

The latest technical upgrade to the Main Injector ring was the *Antiproton Recycler* in 2004. Originally destined as a means of collecting antiprotons after a colliding store and recycling them for the next cycle it now functions as an additional storage of antiprotons coming from the *Accumulator*, thus allowing for an optimal Accumulator operation efficiency.

2.2.3 The Tevatron

Having reached their final pre-acceleration energy of 150 GeV/c^2 in the *Main Injector*, both the p and $\bar{p}$ beams are fed into the *Tevatron*, representing the starting point for the last phase of the acceleration process. As soon as 36 proton and 36 antiproton bunches are circulating inside the *Tevatron*, the particles are accelerated to an energy of 980 GeV/c^2, corresponding to a center of mass energy of $\sqrt{s} = 1.96$ TeV/c^2 at a $p\bar{p}$ collision. The proton bunch typically consists of 3×10^{11}, the antiproton bunch contains 9×10^{10} particles. Since the bunch configuration results in 72 bunch crossing regions inside the 6 km circumference Tevatron, the facility features mechanisms to minimize the

bunch crossing rate at 70 of these unwanted interaction points where no experimental devices are situated. This is achieved by directing the opposite running beams on different helix pathways. Maximizing the number of collisions at the locations of the detector DØ and CDF is obtained by magnetic quadrupole focusing techniques. Since each collision reduces the number of particles, after a certain time period of operation, called store, the beams are reinforced by injecting additional protons and antiprotons from the Main Injector.

2.2.4 Accelerator Performance

To assess the overall performance of an accelerator a parameter quantifying the performance, given by the so called luminosity $\mathcal{L}$, is introduced. Since $\mathcal{L}$ directly infers the protons and antiprotons interaction rate, the layout and the configuration of an accelerator are always geared towards maximizing this quantity while providing for the desired center of mass energy – in case of high center of mass energies a non–trivial task. The relation between $\mathcal{L}$ and accelerator parameters reads as follows:

$$\mathcal{L} = f \cdot \frac{N_b N_p N_{\bar{p}}}{2\pi \left(\sigma_p^2 + \sigma_{\bar{p}}^2\right)} \cdot F\left(\frac{\sigma_l}{\beta^*}\right) \tag{2.1}$$

Here f denotes the revolution frequency, N_b the number of bunches, N_p and $N_{\bar{p}}$ the number of protons and antiprotons per bunch, σ_p and $\sigma_{\bar{p}}$ the average transverse elongation of the proton and antiproton bunches respectively. F is a form factor describing geometric properties of a bunch. The instantaneous luminosity $\mathcal{L}$ is also linked to the interaction rate

$$\dot{N} = \sigma_{int}\mathcal{L}, \tag{2.2}$$

where the cross section σ_{int} directly corresponds to the probability for the interaction of two particles. The overall performance in a certain accelerator runtime is obtained by integrating the luminosity over the time of the given data taking period. Thus, integrating the above equation yields an estimation for the number of events for the process of interest:

$$N = \sigma_{int} \cdot \int \mathcal{L}\mathrm{d}t \tag{2.3}$$

Figures 2.2 and 2.3 [9] show the development of the integrated luminosity and initial luminosities obtained since the start of Tevatron Run II. According to the shown statistics, an integrated luminosity of approximately 4 fb^{-1} has been delivered to the CDF–II detector so far. Until time of writing, $3.16 \times 10^{32} s^{-1}$ cm^{-1} depicts the record in the initial peak luminosity at CDF–II.

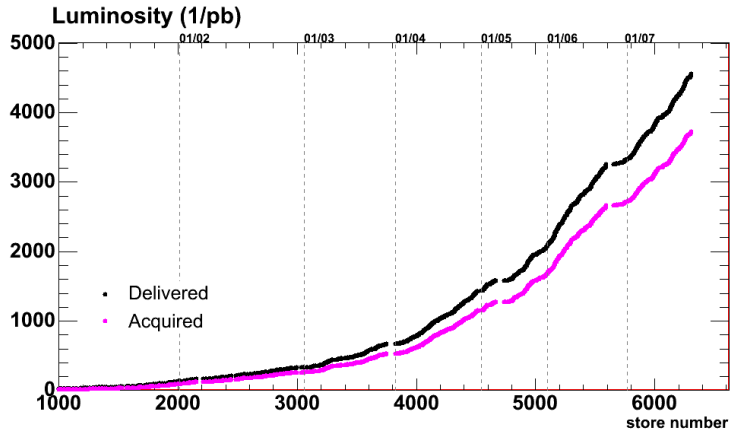


Figure 2.2: Time dependent development of the integrated luminosity per store since the beginning of Tevatron Run II.

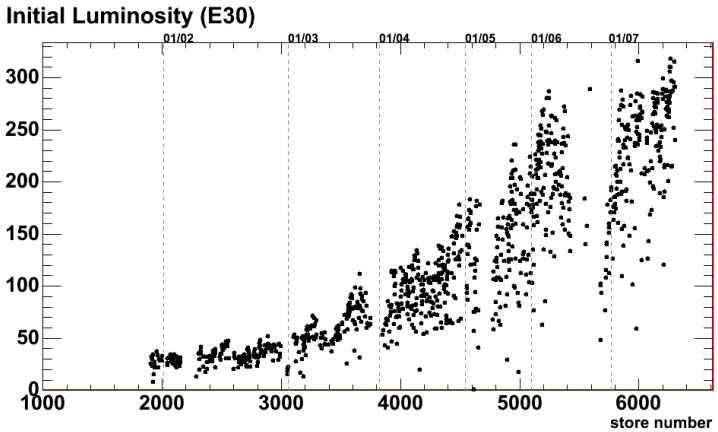
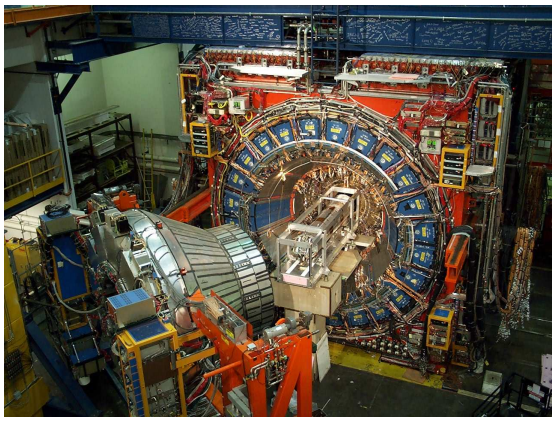
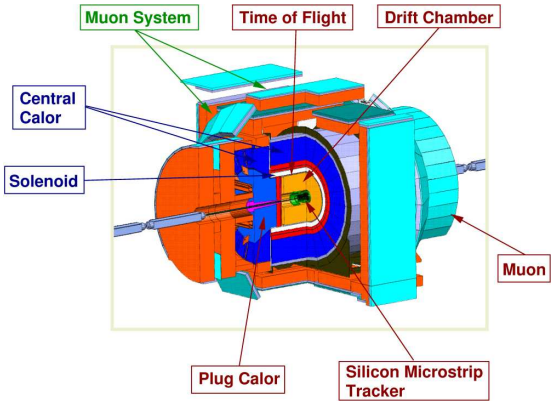


Figure 2.3: Development of the Tevatron Run II initial luminosities.



(a) Photograph of the opened CDF–II detector



(b) Schematic view

Figure 2.4: The CDF–II detector. The individual detector components are outlined throughout the following sections.

2.3 The CDF Detector

2.3.1 Overview

The *Collider Detector at Fermilab* [10], situated at interaction position $B0$ of the Tevatron, is a multipurpose collider detector designed for a broad range of high energy physics regimes. The photograph in figure 2.4(a) and the schematic depiction of the CDF detector in figure 2.4(b) indicate the overall cylindrical symmetry typical to collider detectors in general. Due to the equivalent proton and antiproton energies CDF in addition exhibits a forward–backward symmetry.

A superconducting solenoid generating a 1.4 Tesla magnetic field orientated parallel to the beamline encloses the innermost detector layers, containing the tracking system and the time-of-flight detector (TOF) [11]. Outside the solenoid, electromagnetic and hadronic calorimeters are situated. The outermost part of the detector houses the muon tracking chambers. Overall, the detector scales 12 m in length and 12 m in height.

Exploiting the azimuthal and forward–backward symmetry, particle tracking with CDF–II is performed using an admixture of a polar and a cylindrical coordinate system that is defined in the following way:

- The point of the nominal $p\bar{p}$ interaction marks the *point of origin* of the coordinate system
- The positive direction in z is determined by the proton beam direction
- r measures the distance from the beamline
- The azimuthal angle ϕ is measured upwards within the x–y–plane where the x axis lies in the accelerator plane, radially directing outwards off the center

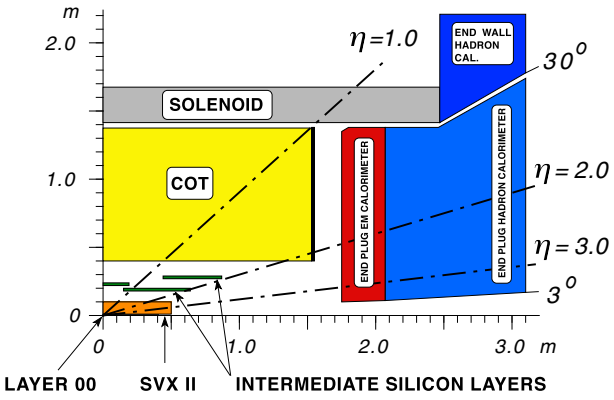


Figure 2.5: r–z view of the tracking system.

of the ring; the Cartesian coordinate y is pointing upright with respect to the accelerator plane

- θ represents the polar angle lying in the y–z–plane. In the field of experimental particle physics the *pseudorapidity* η, which is an approximation of the Lorentz invariant rapidity $\Theta = \tanh^{-1}(v/c)$, is commonly used: $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$

The following sections briefly discuss the basic features and components of the CDF–II detector system. Figure 2.4(b) illustrates the positions of the several subsystems in CDF–II.

2.3.2 Tracking System

Reconstructing particle tracks is performed by exploiting a particle's interaction with the detector material situated inside the superconducting solenoid: When passing through matter a charged particle causes ionization of the material leaving an energy deposit close to the trajectory of the particle. Therefore, measuring the deposited energy by means of a suitable detector enables the reconstruction of the particle track. In addition, bending of the particle track in the magnetic field allows to derive the momentum of the particle.

Due to the radius dependent track densities the CDF–II tracking system consists of two detector components: a silicon detector [12] for high resolution tracking in the immediate vicinity of the interaction point, and a drift chamber, the *Central Outer Tracker* (COT) [13], surrounding the silicon tracker. The silicon detector comprises several sub-systems, namely the *Layer 00* (L00), the *Silicon Vertex Detector* (SVX II) and the *Intermediate Silicon Layer* (ISL). The detector section containing the devices for particle tracking is schematically depicted in figure 2.5.

Silicon Detector

The cylindrically built device, which consists of radiation hard micro strip detectors, radially ranges from 1.35 cm to 28 cm, covering a pseudorapidity range $|\eta| < 2$. Starting from the innermost region the Silicon Detector comprises the following components: (for a schematic view in the $r - \phi$ plane see figure 2.6):

- *Layer 00* (L00):
  In the course of the Run II upgrade *Layer 00* was directly mounted to the beam pipe, thus improving the precision of track measurements and impact parameter resolution. It consists of two overlapping single-sided layers of silicon, each comprising six narrow and six wide modules at radii 1.35 cm and 1.62 cm respectively. L00 has a total length of 94 cm in z direction, resulting in an η coverage of $|\eta| \leq 4$.

- *Silicon Vertex Detector* (SVX II):
  SVX II, a double-sided micro strip detector ranging from $r = 2.1$ cm to $r = 17.3$ cm, basically consists of three barrel like devices, each of them 29 cm in length. The overall geometry of SVX II allows for a track reconstruction with a maximum pseudorapidity of $|\eta| = 2$. Due to the special layout of the five layers inside each device, a high resolution of the impact parameter and the azimuthal angle ϕ of particle tracks is achieved: On one side of each of the layers the silicon microstrips are axially aligned for measuring the r and ϕ coordinate. In addition, layers 0, 1 and 3 feature orthogonally attached stereo strips on the other side, supplementing the $r - \phi$ measurement with a z-measurement, whereas the strips on the other side of layers 2 and 4 are tilted by 1.2 degrees only (SAS, *small angle stereo*), thereby ensuring a reduction of combinatorial ambiguities of multiple hit events in the Silicon Detector.

- *Intermediate Silicon Layer* (ISL):
  The outermost component of the silicon detector system is the 28 cm long *Intermediate Silicon Layer*, which radially ranges from $r = 20$ cm to $r = 28$ cm. The main function of the ISL is to act as a link between SVX II and COT by combining tracks measured in both detector components. It consists of three silicon microstrip layers designed in the same way as layers 2 and 4 of the SVX II. One layer positioned in the central region at $r = 22$ cm covers $|\eta| < 1$, whereas two of the three layers are located at $r = 20$ cm and $r = 28$ cm covering the forward/backward region ($1 < |\eta| < 2$).

Drift Chamber

Situated between $r = 44$ cm and $r = 137$ cm, the *Central Outer Tracker* (COT) represents the outermost part of the tracking system, covering a pseudorapidity range

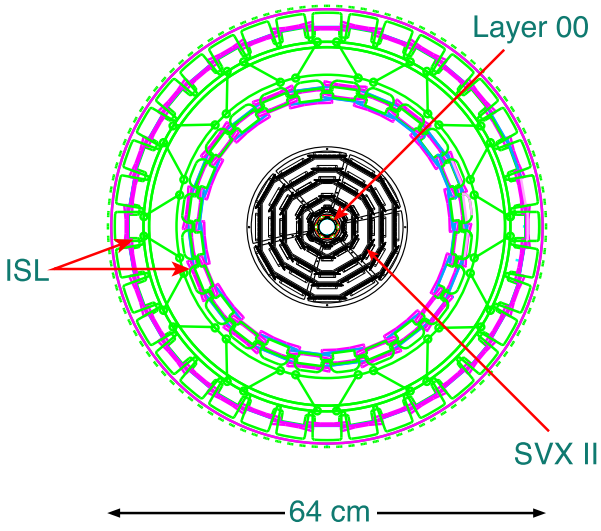


Figure 2.6: The three sub–systems of the Silicon Detector viewed in the $r-\phi$ plane.

$|\eta| < 2$. The cylindrical device contains 96 wire layers in total, arranged into eight superlayers. One half of the superlayers consists of wires running in z-direction destined for the $r-\phi$–measurement (axial superlayers), the wires of the other four superlayers are slightly tilted by 2 degrees with respect to the beamline (stereo superlayers), providing for an additional z–measurement. The medium being ionized by a passing particle inside the drift chamber is a gas mixture composed of Ar-Ethane-CF_4 having a mixing ratio of 50:35:15. The readout of the electric signal in the sense wire caused by the ionization induced electron is performed by an ASDQ (amplifier, shaper, discriminator and charge encoding) chip. Potential mismatching of tracks originating from different beam crossings to a single collision event is avoided by the short drift time of about 100 ns. In addition to the spatial measurement, the COT is suitable for measuring the particle specific energy loss dE/dx, thus supplementing the particle identification discussed in the next section. Direction and position resolution are much poorer than in the inner tracking system, however, due to its high volume expansion allowing long particle tracks, momentum resolution is superior to the silicon detector. Furthermore, the purity of the track reconstruction is better thanks to the smaller density of tracks, which comes as a natural consequence of the higher radial displacement with respect to the primary interaction point.

Track Reconstruction

Based on the response behavior of the detector components so far discussed, several algorithms provided by the CDF–II offline software are applied to reconstruct particle tracks from the available information.

The *COT Standalone Tracking* procedure performs a three-dimensional helix fit [14] to COT hits exclusively. In a next step, these reconstructed tracks are extrapo-

lated to smaller radii and compared to valid hits inside the Inner Tracker (*Outside–In Tracking*) with the criterion for a valid track being at least three axial silicon hits. However, since this procedure does not account for particle tracks possessing too low momenta or occurring at $|\eta| \geq 2$ as a consequence of the limited COT coverage, further reconstruction mechanisms need to be applied. Therefore, the *Silicon Standalone Algorithm* (SISA) [15] performs a helix fit using silicon hits not coinciding with the hits obtained from inwardly extrapolated COT tracks. Instead of requiring two 3D silicon hits as the SISA does, the *Silicon Forward Tracking* (FWD) algorithm reconstructs tracks from one 3D hit and one r-ϕ hit, where in addition a minimum transverse momentum of 0.8 GeV/c^2 is demanded. The *Inside-Out Tracking* (IO) algorithm [16] just inverts the OI mode of operation: Hits inside the silicon detector are extrapolated to larger radii and then the attempt is made to link them to COT hits.

2.3.3 Particle Identification

The capability of identifying pions, kaons, protons, electrons or muons is essential for many kinds of physics analyses done at CDF–II. This particularly holds true in case of the presented analysis, where B^0 or B_s mesons are reconstructed using different hadronic decay channels. By combining several detector measurements the CDF offline software is capable of providing a probability for a given particle identity associated to a measured track. Devices and techniques used for particle identification are briefly discussed below.

Central Detector Devices

Installed at a radius of $r = 140$ cm with respect to the beam pipe, the *Time of Flight* detector (TOF) represents the outermost device enclosed by the superconducting magnetic coil. It consists of 216 plastic scintillator bars, each 280 cm long. At the end of each scintillator a photomultiplier is attached for signal readout.

Particle identification is carried out by calculating the invariant mass

$$m = \frac{p}{c}\sqrt{\frac{(ct)^2}{L^2} - 1}, \tag{2.4}$$

where L is the path length of the particle measured between the primary vertex and the position of a scintillator hit inside the TOF. $t = T_{TOF} - T_0$ is the time difference between detection time T_{TOF} and production time T_0. Several inputs are needed for utilizing this relation: First of all, the position of the primary vertex of an event needs to be known. For this, the TOF signal is matched to the corresponding track in the COT and in a next step the position of the primary vertex is reconstructed. p is taken from the momentum measurement performed in the COT.

Another contribution to the particle identification process comes from the property of a charged particle to lose energy when passing through matter, essentially caused

by ionization processes and atomic excitation. The specific energy loss dE/dx for a given $\beta\gamma = p/m$ is described by the Bethe-Bloch equation, thus allowing for particle identification by measuring dE/dx in the COT. The amount of deposited energy is assessed by exploiting the relation

$$\Delta t \propto \log Q \propto \frac{dE}{dx}. \tag{2.5}$$

Here Δt denotes the pulse width of a COT signal and Q is the total charge deposited in the COT.

Muon Detector System

Since muons are much heavier than electrons and do not interact strongly with the atomic nuclei of the detector material, they are capable of passing through a much higher volume of matter than other charged particles. A detector signal in the muon chambers [17], situated in the outermost region of the CDF detector and thus benefitting from the screening effect of the calorimeters, provides a very strong indication for a muon candidate. There are however some events being misreconstructed as muons. Such an event is called a muon fake event. The muon fake rates are rather small, which is due to a low number of kaons and pions reaching the muon chambers. There is also a background component consisting of real muons, either coming from outside the detector or stemming from pion or kaon decays. A muon track segment, called stub, measured in one the muon chambers is matched to a track of the COT drift chamber. The muon detector system at CDF consists of four detectors, made up of scintillators, drift tubes and steel absorbers. The individual components exhibit different geometries leading to different η coverages, with the maximum pseudorapidity coverage being $|\eta| = 1.5$. Since the particle reconstruction and identification procedures used for candidate selection in the given analysis do not utilize the functionality of the muon detector system, an in–depth discussion of these devices is not necessary.

2.3.4 Calorimeters

The calorimetry system [18, 19, 20], located in a radial range between the solenoid and the muon detector system, complements the functionality of the CDF detector systems by measuring the deposited energy of high energy photons, electrons and hadronic jets. The electromagnetic (hadronic) calorimeters consist of alternating layers of scintillators and lead (iron) as absorbing material. The calorimetry system, grouped into five independent calorimeters, in total covers a pseudorapidity range $|\eta| \leq 3.4$. These individual detector components are: *Central Electromagnetic Calorimeter* (CEM, $|\eta| \leq 1.1$), the *Central Hadronic Calorimeter* (CHA, $|\eta| \leq 0.9$), the *Endwall Hadronic Calorimeter* (WHA, $0.8 \leq |\eta| \leq 1.2$), the *Plug Electromagnetic Calorimeter* (PEM, $1.1 \leq |\eta| \leq 3.6$) and the *Plug Hadronic Calorimeter* (PHA, $1.2 \leq |\eta| \leq 3.6$).

2.3.5 Trigger System

As a consequence of the high bunch crossing rate of 2.5 MHz at Tevatron Run II, storing every single collision event would inevitably generate an enormous amount of data which would need to be managed. With today's computer and mass storage capabilities in place these data are not uncontrollable, however, since most of the collision events are just combinatorial background, one is simply not interested in keeping every single detector event. Therefore preselection measures based on predefined criteria need to be in place in order to reduce data drastically, allowing for a more effective data management and extraction of events which are interesting for later physical studies. At CDF–II this is implemented by means of a three level *trigger* system (see figure 2.7 for illustration) which evaluates information coming from the various detector components. The decision to discard an event or not successively made by each trigger subsystem (Level) is based on trigger tables containing *trigger paths*, which in turn incorporate rules defining requirements for an event. The *Two Track Trigger* (TTT) is a prominent example for the content of such a trigger path. Since it is of particular importance for studies where B mesons are reconstructed on an exclusive basis, as performed in the generation of data used for the presented analysis, a short description of the TTT will be given in the next sub–section. The design and functionality of the three trigger levels is described in [21] and briefly outlined in [22, 23].

Two Track Trigger

The *Two Track Trigger* selects a track pair displaced from the primary interaction vertex and requires this event to possess the following attributes:

- *Level 1*
  At least two XFT tracks (*Extremely Fast Tracker*, [24]) each having a minimum transverse momentum $p_T > 2$ GeV/c and a total transverse momentum of $p_{T,1} + p_{T,2} > 5.5$ GeV/c. The azimuthal angle between the two tracks is required to be $\Delta\phi_{1,2} < 135°$.

- *Level 2*
  A minimum of two SVT tracks (*Silicon Vertex Trigger*, [25]) matching the XFT tracks ($\chi^2_{SVT} < 25$) and each possessing $p_T > 2$ GeV/c, an impact parameter d_0 in the range 120 μm and 1 mm and decay length $L_{xy} > 200$ μm.

- *Level 3*
  The SVT tracks have to match the tracks in the COT. The additional requirements are: $2° < \Delta\phi_{1,2} < 90°$, 80 μm$< d_0 <$ 1 mm, $L_{xy} > 200$ μm.

Additional requirements are implemented by the Two Track Trigger subpaths. These are necessary in order to adjust data taking to different luminosity scenarios.

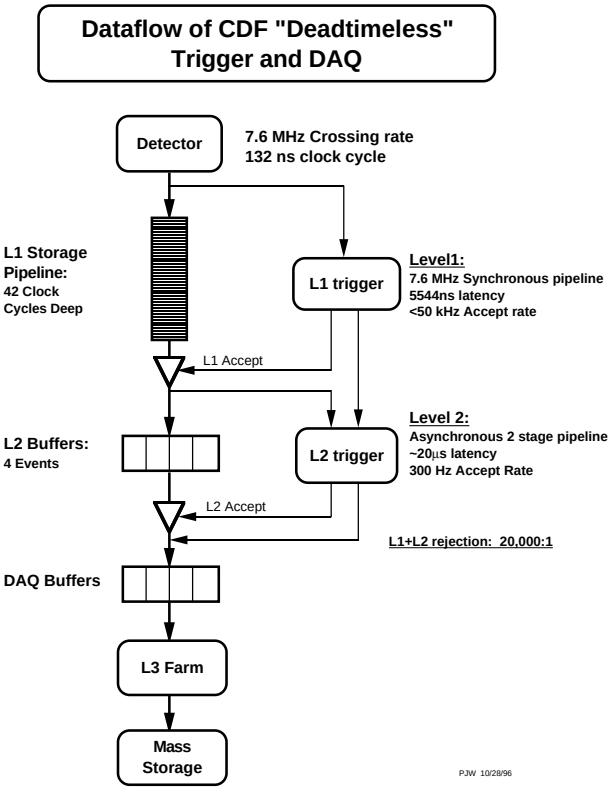


Figure 2.7: The three level trigger system at CDF–II.

- `B_LOWPT`: Low p_T scenario: At least two SVT tracks with a minimum total transverse momentum of $p_{T,1} + p_{T,2} > 4.0$ GeV/c.
- `B_CHARM`: Medium p_T scenario: Requires at least two oppositely charged SVT tracks with $p_{T,1} + p_{T,2} > 5.5$ GeV/c.
- `B_HIGHPT`: High p_T scenario: At least two oppositely charged SVT tracks with $p_{T,1} + p_{T,2} > 6.5$ GeV/c.

In order to scale down the amount of data in periods of high luminosity, a dynamically adjusted quantity p, called *prescale factor*, is introduced. Prescaling of the trigger paths means that only one out of $N = 1/p$ events having met the trigger requirements is accepted. This provides for more free bandwidth for other trigger paths by decreasing the deadtime of the trigger system. The prescale factors are stored in a database to be accessible in offline software.

Chapter 3

Theoretical Overview: B_S Meson Physics

In order to understand how and to which extent an analysis of the decay $B_s \to D_s D_s$ can help to make inferences to physics in the B_s meson sector, this chapter aims to motivate the given study by outlining the underlying theoretical framework. In the first place a brief introduction into the Standard Model of particle physics is given.

3.1 Introduction

Introduced in the 1960's, the *Standard Model* (SM) of elementary particle physics [26, 27, 28, 29], which since then has been and still is subject to numerous tests and cross–checks, up to now is commonly accepted as the best theory to consistently describe the properties of particles and three out of the four fundamental interactions known today. In mathematical parlance the description of the Standard Model corresponds to the local symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. The strong, electromagnetic and weak interaction between spin 1/2 *fermions* are mediated by the exchange of spin 1 particles, called *bosons*. The exchange particles are associated with gauge fields of the respective symmetry groups:

- $SU(3) \quad \to \quad G_\mu^\alpha, \quad \alpha = 1...8$
  The gauge fields G_μ^α represent the eight physical particles mediating the strong interaction, called gluons.

- $SU(2) \quad \to \quad W_\mu^\alpha, \quad \alpha = 1...3$

- $U(1) \quad \to \quad B_\mu$
  The gauge fields W_μ^α and B_μ are associated with two charged bosons, $W^\pm$, and one neutral boson, Z^0, being responsible for the exchange of the weak interaction, and the massless photon, γ, mediating the electromagnetic interaction.

$$\begin{pmatrix}\nu_e\\ e\end{pmatrix} \quad \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix} \quad \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix} \qquad \begin{pmatrix}u\\ d\end{pmatrix} \quad \begin{pmatrix}c\\ s\end{pmatrix} \quad \begin{pmatrix}t\\ b\end{pmatrix}$$

Table 3.1: The elementary fermionic particles of the Standard Model, leptons (left) and quarks (right). Each lepton and quark in addition has its antipartner $\bar{l}$ and $\bar{q}$ respectively.

All fermions share the capability to participate in the weak, and if charged, in the electromagnetic interaction. Fermions are divided into *quarks* and *leptons*, each being classified into three families or six flavors respectively. In addition, each fermion possesses its anti-partner, giving rise to 12 elementary particles in the fermion and 12 elementary particles in the quark sector (see table 3.1). Quarks belonging to a given family are subdivided into *up* and *down* type quarks. In consequence of featuring another quantum number, called color charge, quarks additionally take part in strong interactions. Due to self–interaction of the gluons, a result of carrying color themselves, quarks never occur isolated, a phenomenon usually referred to as confinement. To complete the rudimentarily drawn picture of the Standard Model, the bound states being formed by quarks are called *hadrons*, whereupon hadrons consisting of three quarks are referred to as *baryons*, those composed of one quark and one antiquark are identified as *mesons*.

Mesons consisting of the bottom quark b and a light quark in general are referred to as B mesons, where B_q is a commonly used denotation for a particular B meson with q denoting the lighter quark. Thus, the neutral B_s ($\bar{B}_s$) meson is composed of a $\bar{b}$ (b) and the lighter s ($\bar{s}$) quark. In order to understand the underlying physics in the B_s–$\bar{B}_s$ system, which are governed by charged current weak interactions, some additional insights into the mechanism of quark transitions are needed.

3.2 CKM Matrix and Unitarity Triangle

As mentioned above, the weak interaction is mediated by $W^\pm$ and Z^0 gauge bosons. The only process allowing for flavor changes is given by charged current interactions, corresponding to the exchange of $W^\pm$ bosons. In the quark sector, the transition between different quark flavors is hence realized by a $W^\pm$ boson interaction coupling to the physical up and down type quarks[1]. The particular coupling strengths are given by the elements of the 3×3 *Cabibbo-Kobayashi-Maskawa* (CKM) matrix $\mathbf{V}_{CKM}$ [30]. The transformation between the weak eigenstates q' and the mass eigenstates q, which

[1] To put it correctly, in the Standard Model the $W^\pm$ gauge boson couples to so called left handed fermions q_l, l_l only, a property describing the *chirality* of a given particle

do not coincide, can be written as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{3.1}$$

The 18 free real parameters of the complex matrix $\mathbf{V}_{CKM}$ are reduced to nine if the matrix fulfills the unitarity requirement $\mathbf{V}_{CKM}\,\mathbf{V}_{CKM}^{\dagger} = \mathbf{1}$, corresponding to the existence of three quark families in the Standard Model. Five of these parameters are phases containing no physical information. Thus, the amount of parameters can be further reduced to four by means of arbitrary phase conventions. These parameters are three rotation angles and one complex phase, giving rise to the common CKM parameterization:

$$\mathbf{V}_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} , \tag{3.2}$$

where s_{ij} and c_{ij} stand for $\sin\theta_{ij}$ and $\cos\theta_{ij}$ respectively, θ_{ij} are the rotation angles and δ is the complex phase.

For the above parameterization $\mathbf{V}_{CKM}$ holds the experimentally known hierarchy $s_{13} \ll s_{23} \ll s_{12} \ll 1$. This motivates a parameterization which exploits this hierarchy. By setting

- $\lambda = \sin\theta_{12} \approx 0.2$
- $\eta = \sin\theta_{13}\sin\delta_{13}/A\lambda^3$
- $\rho = \sin\theta_{13}\cos\delta_{13}/A\lambda^3$
- $A = \sin\theta_{23}/\lambda^2$

and expanding the relations in powers of λ up to the third order, the CKM matrix in the more frequently used *Wolfenstein* parameterization is obtained:

$$\mathbf{V}_{CKM} \cong \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right) \tag{3.3}$$

From this parameterization it can be seen that the diagonal elements V_{ii} are close to 1, whereas $|V_{us}| \simeq |V_{cd}| \approx 0.2$. The remaining matrix elements in contrast are very small, of the order 10^{-3}.

The CKM matrix exhibits a nice feature offering the possibility to interprete the underlying physics in a graphical way if one applies the unitarity conditions

$$\sum_{i=1}^{3} V_{ij}V_{ik}^{*} = \delta_{jk}, \quad \text{and} \quad \sum_{j=1}^{3} V_{ij}V_{kj}^{*} = \delta_{ik}, \quad k = 1,2,3 \tag{3.4}$$

to $\mathbf{V}_{CKM}$, leading to six vanishing equations resulting from multiplying two out of the three columns or rows respectively, out of which the combination of column one and three is the most commonly used one:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{3.5}$$

Dividing each side of (3.5) by $A\lambda^3 \approx V_{cd}V_{cb}^*$, and using the Wolfenstein parameterization leads to

$$(\bar{\rho} + i\bar{\eta}) - 1 + (1 - \bar{\rho} - i\bar{\eta}) = 0. \tag{3.6}$$

This depicts one of six possible graphical representations of unitarity triangles in a complex plane, where the parameters belonging to the chosen one are best studied and constrained, since the lengths of the sides are of the same order of magnitude. The coordinates of the corners of the triangle shown in figure 3.1 are (0, 0), (1, 0) and $(\bar{\rho}, \bar{\eta})$, where $\bar{\rho}$ and $\bar{\eta}$ are defined as

$$\bar{\rho} = \left(1 - \lambda^2/2\right)\rho, \quad \bar{\eta} = \left(1 - \lambda^2/2\right)\eta \tag{3.7}$$

The angles of the unitarity triangle are related to the elements of the CKM matrix in the following way:

$$\phi_1 = \beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) \tag{3.8}$$

$$\phi_2 = \alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) \tag{3.9}$$

$$\phi_3 = \gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) \tag{3.10}$$

Overconstraining the parameters of the unitarity triangle by independent experimental measurements is an important goal of flavor physics because this provides an essential opportunity for validating the Standard Model or making constraints to new physics.

For the discussion of physics in the B_s–$\bar{B}_s$ system however, the unitarity requirement based on the product of the second and third column of $\mathbf{V}_{CKM}$ is needed:

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \tag{3.11}$$

This again gives rise to another possible representation of the unitarity triangle. Since the angle $\phi_s = \beta_s = \arg\left(-V_{ts}V_{tb}^*/V_{cs}V_{cb}^*\right)$ is expected to be close to zero in the Standard Model, the resulting unitarity triangle is degenerated. This prediction in turn is equivalent to a vanishing CP violation in the B_s–$\bar{B}_s$ system as the angle ϕ_s is identified as the phase being responsible for CP violation.

In the next section the formal framework needed for motivating the goal of this analysis is introduced. The brief description and the definitions of the various quantities closely follow the discussion in [1], where the theoretical framework and the experimental possibilities for the search of new physics in B_s decays are pointed out in detail. Hence, only the most important aspects are outlined while mathematical derivations are blanked out completely throughout the following brief discussion.

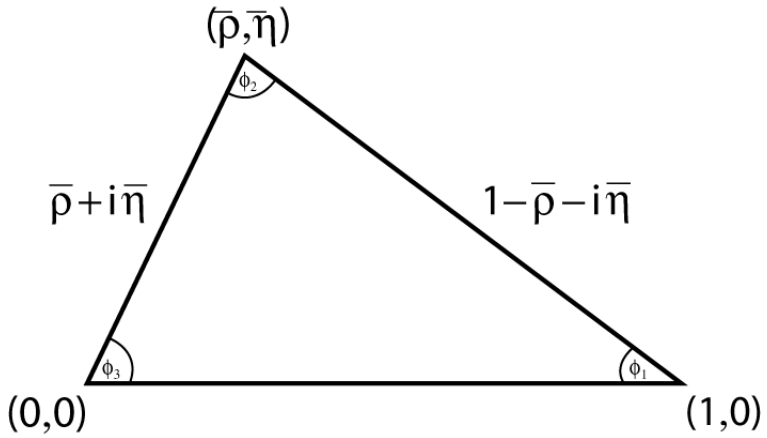


Figure 3.1: The unitarity triangle as one of the possible graphical representations of the unitarity condition for $\mathbf{V}_{CKM}$.

3.3 B_s Meson Mixing and Decay

3.3.1 Overview

The flavor eigenstates B_s and $\bar{B}_s$ are not invariant under the consecutive application of the charge (C) and parity (P) transformation, in fact the eigenvalue problem of CP reads as follows:

$$CP\,|B_s\rangle \;=\; -\left|\bar{B}_s\right\rangle \tag{3.12}$$

Thus, B_s and $\bar{B}_s$ are not eigenstates of CP, yet their linear combinations

$$|B_s^{even}\rangle \;=\; \frac{1}{\sqrt{2}}\left(|B_s\rangle \;-\; \left|\bar{B}_s\right\rangle\right), \quad \text{and} \quad \left|B_s^{odd}\right\rangle \;=\; \frac{1}{\sqrt{2}}\left(|B_s\rangle \;+\; \left|\bar{B}_s\right\rangle\right) \tag{3.13}$$

Given the CP even and odd final states f_{CP+} and f_{CP+} originating from a B_s–$\bar{B}_s$ decay, one introduces the CP width difference $\Delta\Gamma_{CP} \;=\; \Gamma\left(B_s \;\rightarrow\; f_{CP+}\right) - \Gamma\left(B_s \;\rightarrow\; f_{CP-}\right) \;=\; \Gamma_{even} - \Gamma_{odd}$. This is theoretically predicted to be non–vanishing in the B_s–$\bar{B}_s$ system, corresponding to the fact that certain final states stemming from B_s–$\bar{B}_s$ decays are not accessible due to their CP content: precisely, more CP even than CP odd states are realized in B_s decays. Furthermore the weak eigenstates do not coincide with the mass eigenstates which again are linear combinations of B_s and $\bar{B}_s$:

$$\left|B_s^L\right\rangle \;=\; p\,|B_s\rangle + q\left|\bar{B}_s\right\rangle, \quad \text{and} \quad \left|B_s^H\right\rangle \;=\; p\,|B_s\rangle - q\left|\bar{B}_s\right\rangle \tag{3.14}$$

where p and q are required to fulfill the normalization condition $|p|^2 + |q|^2 = 1$. The mass and width difference between the light and heavy mass eigenstates are defined as

$$\Delta m \;=\; M_H \;-\; M_L, \quad \Delta\Gamma \;=\; \Gamma_L \;-\; \Gamma_H, \quad \Gamma \;=\; \frac{1}{\tau_{B_s}} \;=\; \frac{\Gamma_H + \Gamma_L}{2}. \tag{3.15}$$

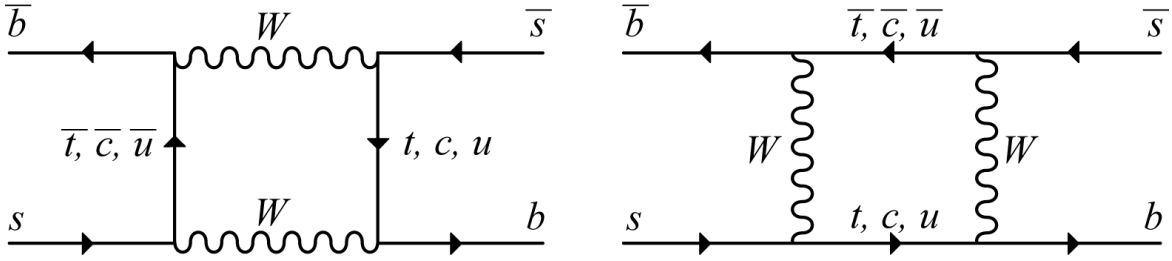


Figure 3.2: Lowest order Feynman diagrams, so called box diagrams, illustrating B_s mixing. The main contribution to the loop processes is induced by the t quark.

As described above, on quark level the weak interaction allows the transition between different flavors. Since quarks are the basic material for mesons, the underlying mechanism also enables a $|B_s\rangle$ state to change into a $|\bar{B}_s\rangle$ state and vice versa. This phenomenon is referred to as B_s meson mixing. The time evolution of the B_s–$\bar{B}_s$ system is governed by the Schrödinger equation:

$$i\frac{d}{dt}\begin{pmatrix}|B_s(t)\rangle\\|\bar{B}_s(t)\rangle\end{pmatrix} = \left(\mathbf{M}-\frac{i}{2}\mathbf{\Gamma}\right)\begin{pmatrix}|B_s(t)\rangle\\|\bar{B}_s(t)\rangle\end{pmatrix} \tag{3.16}$$

Here the Hamiltonian

$$\mathbf{H} = \left(\mathbf{M}-\frac{i}{2}\mathbf{\Gamma}\right) = \begin{pmatrix}M_{11}-\frac{i}{2}\Gamma_{11} & M_{12}-\frac{i}{2}\Gamma_{12}\\ M_{12}^*-\frac{i}{2}\Gamma_{12}^* & M_{22}-\frac{i}{2}\Gamma_{22}\end{pmatrix} \tag{3.17}$$

is composed of the mass matrix $\mathbf{M} = \mathbf{M}^\dagger$ and the decay matrix $\mathbf{\Gamma} = \mathbf{\Gamma}^\dagger$, where the off-diagonal elements $M_{12} = M_{21}^*$ and $\Gamma_{12} = \Gamma_{21}^*$ govern B_s–$\bar{B}_s$ mixing and mixed decay. In the Standard Model, mixing is realized by second order transitions of weak interaction with the b ($\bar{b}$) quark going into an s ($\bar{s}$) and vice versa by W^+ and W^- exchange. These loop processes, that are commonly illustrated in box diagrams (see figure 3.2), are strongly dominated by the *top* quark with small contributions stemming from u and c quarks. Hence, the phase ϕ_M of the complex 2×2 mass matrix $\mathbf{M}$ reads $\phi_M = \arg(M_{12}) = \arg(V_{tb}V_{ts}^*)^2$ in the Standard Model. The oscillations that were found to be very rapid in the B_s–$\bar{B}_s$ case ($\Delta m = 17.77 \pm 0.10 \pm 0.07$ ps^{-1}) [2] compete with tree level decays, whereas the leading contribution to Γ_{12} comes from the strongly CKM favored decay $b \to c\bar{c}s$ (see figure 3.3 for illustration). $b \to c\bar{u}d$ has a larger phase space than $b \to c\bar{c}s$ though, however it can not form a CP eigenstate contrary to a $c\bar{c}$ final state. The doubly Cabibbo–suppressed contributions proportional to $V_{ub}V_{us}^*$ can safely be neglected. Therefore, the tree level decay $b \to c\bar{c}s$ is the preferred mode for a joint B_s–$\bar{B}_s$ decay. Because $\arg(V_{cb}V_{cs}^*)^2$ is close to the mixing phase $\arg(V_{tb}V_{ts}^*)^2$, the Standard Model expectation for the phase between B_s mixing and decay

$$\phi = \phi_M - \phi_\Gamma = \arg(M_{12}) - \arg(-\Gamma_{12}) \tag{3.18}$$

is close to zero, of the order 10^{-2}, corresponding to vanishing CP violation in the interference between mixing and decay. In this case, CP violation describes the phenomenon that the absolute amplitude of the interference between a mixed and an

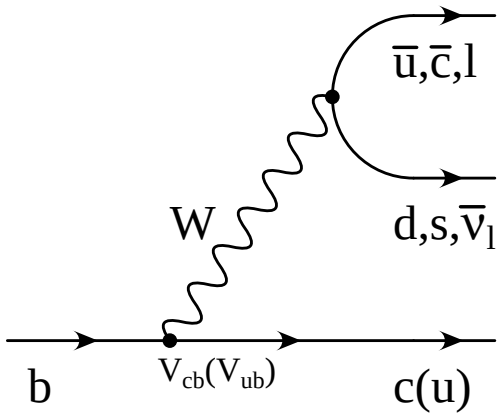


Figure 3.3: Feynman diagram showing tree level b decays in the Standard Model. The main contribution stems from $b \to c\bar{c}s$.

unmixed B_s decay into a CP final state f does not equal the one of the $\bar{B}_s$ into a opposite CP final state $\bar{f}$:

$$\left|\left\langle f \mid \bar{B}_s \to B_s \right\rangle + \left\langle f \mid B_s \right\rangle\right| \neq \left|\left\langle \bar{f} \mid B_s \to \bar{B}_s \right\rangle + \left\langle \bar{f} \mid \bar{B}_s \right\rangle\right| \tag{3.19}$$

Thus, the detection of a sizeable CP violating phase in the B_s–$\bar{B}_s$ system would clearly point at physics beyond the Standard Model. As the main contribution to the second order loop processes in B_s–$\bar{B}_s$ mixing originates from the t quark and $|V_{ts}|$ is small, new physics could easily compete with the Standard Model contributions to M_{12}, offering an opportunity for hypothetical new physics effects to be observable.

As shown in [1], one can derive

$$\Delta\Gamma \;=\; \Delta\Gamma_{CP}\cos\phi, \quad \text{with}\,\Delta\Gamma_{CP} \;\equiv\; 2\,|\Gamma_{12}| \;=\; \Gamma\left(B_s^{even}\right) \;-\; \Gamma\left(B_s^{odd}\right) \tag{3.20}$$

and

$$|B_L\rangle \;=\; \frac{1+e^{i\phi}}{2}\,|B_s^{even}\rangle - \frac{1-e^{i\phi}}{2}\,\left|B_s^{odd}\right\rangle \tag{3.21}$$

$$|B_H\rangle \;=\; -\frac{1-e^{i\phi}}{2}\,|B_s^{even}\rangle + \frac{1+e^{i\phi}}{2}\,\left|B_s^{odd}\right\rangle. \tag{3.22}$$

Therefore, in the Standard Model case where ϕ equals zero, $\Delta\Gamma = \Delta\Gamma_{CP}$ and the CP eigenstates coincide with the mass eigenstates. This allows for the separation of the mass eigenstates by measuring the CP state at the time of the B_s decay. However, with a no–vanishing phase, $\Delta\Gamma$ would be diminished by a factor of $\cos\phi$. Since $\Delta\Gamma$ is identified as $2\,|\Gamma_{12}|$ which, as stated above, is strongly governed by CKM favored Standard Model physics, the CP decay width difference $\Delta\Gamma_{CP}$ itself is insensitive to new physics.

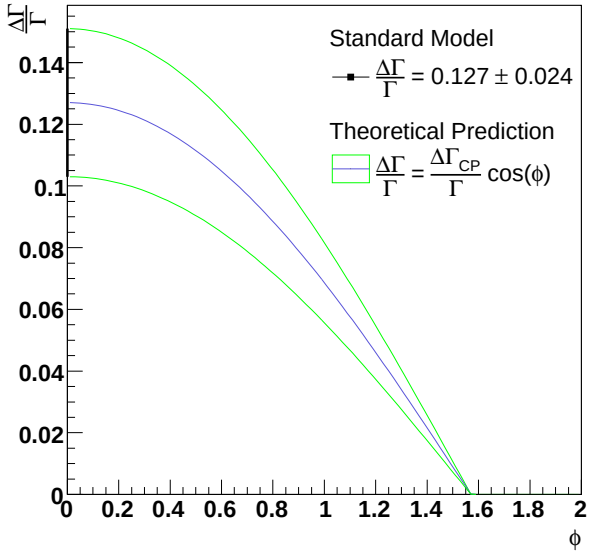


Figure 3.4: Illustration of $\Delta\Gamma/\Gamma = \Delta\Gamma_{CP}/\Gamma \cdot \cos\phi$ in the first quadrant. The black line shows the Standard Model expectation for $\phi = 0$.

3.3.2 Determination of $\Delta\Gamma_{CP}/\Gamma$

Assuming an untagged sample of B_s decays, meaning that one does not determine the b quark flavor of the B_s at production time, and further assuming the final state f to be a purely CP even product of a $b \to c\bar{c}s$ tree level decay, it can be shown [1] to good approximation that

$$2\,Br\,[f_{CP+}] \cong \frac{\Delta\Gamma_{CP}}{\Gamma}. \tag{3.23}$$

To put it crudely, by counting events in a particular decay mode one is able to directly infer the relative CP decay width difference $\Delta\Gamma_{CP}/\Gamma$ and, with the presumption of no new physics contributions, $\Delta\Gamma/\Gamma$. In the non Standard Model case of $\phi \neq 0$ however, a measurement of $\Delta\Gamma_{CP}/\Gamma$ provides an important input to constrain the relation $\Delta\Gamma/\Gamma = \Delta\Gamma_{CP}/\Gamma \cdot \cos\phi$ which is illustrated in figure 3.4. The decay $B_s \to D_s^+ D_s^-$ is considered to be an optimal candidate for an analysis of this kind: The final state $D_s^+ D_s^-$ is realized via the CKM favored tree level decay $b \to c\bar{c}s$, while to good approximation the s quark is supposed to act as a spectator quark. The Feynman diagram of this decay is depicted in figure 3.5. Moreover, $B_s \to D_s^+ D_s^-$ is believed to yield the main contribution to the decay width difference in the B_s–$\bar{B}_s$ system, making this decay mode notably interesting. $D_s^+ D_s^-$ is a single, purely CP even final state with a total angular momentum of $L = 0$. Therefore, an angular analysis of the final state to distinguish the different CP states, as it was performed in case of the $B_s \to J/\psi\phi$ analysis [22], where the final state has the same quark content, is not required.

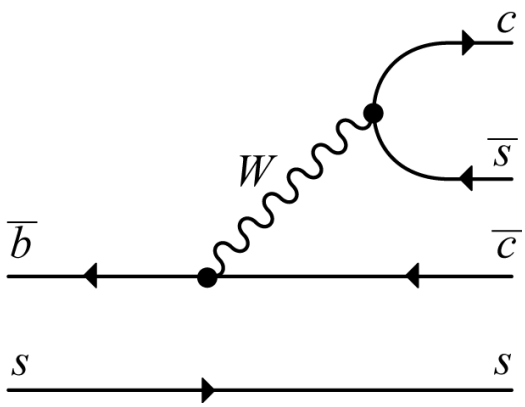


Figure 3.5: Feynman diagram illustrating the decay of a neutral B_s ($\bar{b}s$) into the charged mesons D_s^+ ($c\bar{s}$) and D_s^- ($\bar{c}s$).

Using the decay $B_s \rightarrow D_s^+ D_s^-$, relation (3.23) becomes:

$$2\,Br\left[D_s^{+(*)} D_s^{-(*)}\right] \cong \frac{\Delta\Gamma_{CP}}{\Gamma}. \tag{3.24}$$

However, the final states consisting of a combination of one or two excited D_s mesons give rise to some concern regarding the exact validity of the equation above, because the CP content of $D_s^{+(*)} D_s^{-(*)}$ is a priori not known. It is commonly argued that with some theoretical input $D_s^{+(*)} D_s^{-(*)}$ can be regarded as predominantly CP even in good approximation. The arguments [31] read as follows:

- The decay $B_s^{odd} \rightarrow D_s^+ D_s^{-(*)}$ is forbidden if $m_c \rightarrow \infty$ is set and terms of the order $1/N_{color}$ are neglected. Furthermore in the untagged decay of $B_s^{un} \rightarrow D_s^{+(*)} D_s^{-(*)}$ the transition to the final state is an S–wave.

- In the Shifman–Voloshin (SV) [32] limit $m_c \rightarrow \infty$ and with $m_b - 2m_c \rightarrow 0$, $\Delta\Gamma_{CP}$ is saturated by $\Gamma(B_s^{un} \rightarrow D_s^{+(*)} D_s^{-(*)})$. In addition, with $N_{color} \rightarrow \infty$ and the SV limit, $2\Gamma(B_s^{un} \rightarrow D_s^{+(*)} D_s^{-(*)})$ equals $\Delta\Gamma_{CP}$.

Within these theoretical limits $\Gamma_{CP}^{odd}\left(B_s^{un} \rightarrow D_s^{+(*)} D_s^{-(*)}\right)$ vanishes and relation (3.24) also holds true for the excited final states. Recent measurements however indicate that in the B_d^0 system the final state $D^{+*} D^{-*}$ possesses CP odd contributions of the order $12 - 14$ percent. At this moment one can not safely proof that this also holds true for B_s, but it cannot be excluded that CP odd components neither are negligible in case of $B_s \rightarrow D_s^{+(*)} D_s^{-(*)}$. Additionally it has been argued [1] that the corrections to the adopted limits may be numerically sizeable. Besides, in the SV limit no multibody decays with more than two final states are considered. Since an exclusive selection of the purely CP even final state $B_s \rightarrow D_s^+ D_s^-$ alone would underestimate the CP

content, one may use the measurement of its branching fraction for setting a lower bound to the CP width difference,

$$\frac{\Delta\Gamma_{CP}}{\Gamma} \geq 2Br\left[D_s^+ D_s^-\right], \tag{3.25}$$

as it was performed in the previous CDF measurement [3] where the first observation of $B_s \to D_s^+ D_s^-$ by means of an exclusive reconstruction was reported.

3.4 Existing Measurements

There exist several measurements related to the decay $B_s \to D_s^+ D_s^-$. In 2000, ALEPH reported evidence for this decay [33]. There, D_s^* mesons were reconstructed by evaluating correlations between two detected ϕ mesons. The inclusive branching fraction was quoted as $Br[B_s \to D_s^{+(*)} D_s^{-(*)}] = 0.23 \pm 0.10^{+0.19}_{-0.09}$, pointing to $\Delta\Gamma_{CP}/\Gamma = 0.26^{+0.30}_{-0.15}$.

In addition, DØ recently updated [34] their inclusive measurement of $Br[B_s \to D_s^{+(*)} D_s^{-(*)}]$ using data corresponding to 2.8 fb^{-1}. The analysis based on the reconstruction of the semileptonic decays $D_s^* \to D_s\gamma(\pi^0)$, $D_s \to \phi\mu\nu$. With $Br[B_s \to D_s^{+(*)} D_s^{-(*)}] = 0.042 \pm 0.015(stat) \pm 0.017(syst)$ and by assuming the inclusive final state to be predominantly CP even, $\Delta\Gamma/\Gamma = 0.088 \pm 0.030 \pm 0.036$ in the Standard Model scenario was derived.

The observation of the decay $B_s \to D_s^+ D_s^-$ on the basis of an exclusive measurement was reported by the CDF collaboration in the end of 2006 [3]. Accounting for the fact that $B_s \to D_s^{*+} D_s^-$ lacks a defined CP content and $B_s \to D_s^{*+} D_s^{*-}$ is not fully CP even, the measurement of $Br[B_s \to D_s^+ D_s^-] = 0.094^{+0.044}_{-0.042}$ was used to obtain the lower bound $\Delta\Gamma_{CP}/\Gamma > 0.012$ at 95% C.L.

Chapter 4

Event Reconstruction and Candidate Selection

This chapter focuses upon the first stage of this analysis, comprising the candidate reconstruction performed by the CDF–II offline software as well as the pre–selection and final selection of B_s and B^0 mesons. Since artificial neural networks are of primary importance both for the pre– and final selection, a brief introduction to this topic is given first.

4.1 Artificial Neural Networks

A frequently used procedure to extract signal events from a given data sample composed of different event classes is to cut on several representative variables. Then an iterative cut optimization is performed in order to find the cut values yielding an optimal ratio between number of signal events and background. However, this cut optimization procedure does not account for correlations between variables, potentially leading to a significant loss of signal candidates.

An alternative approach is to perform a transformation mapping the n dimensional space spanned by a given set of variables onto a single scalar in such a way that this variable combines all information and correlations contained in the input variables. This variable therefore can be used as a cut variable instead of cutting on a combination of several variables. A mapping of this kind can be realized by artificial neural networks (ANN):

$$f : \mathbb{R}^n \xrightarrow{ANN} \mathbb{R} \tag{4.1}$$

The basic element of artificial neural networks is the artificial *neuron* or *node* which operates analogous to the biological neuron in the brain. This specialized cell is composed of an *input structure* (the dendrites), a cell body and an *output structure* (the axon). The neurons communicate among each other by exchanging electrical impulses via the dendrites and the output synapses of the axon. If the amplitude of the electrical signal received via the dendrites exceeds a certain threshold, the neuron is triggered

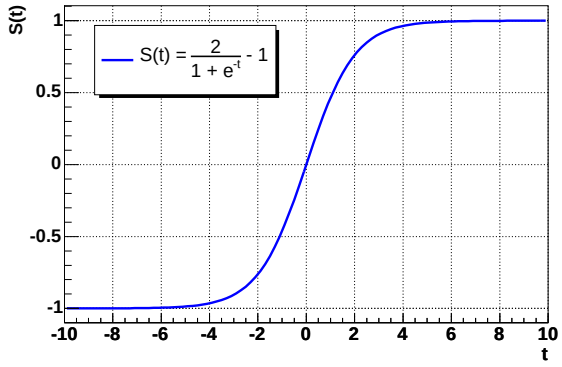


Figure 4.1: A possible representation of the sigmoid function.

and fires an electrochemical signal along the axon. The intensity of the electrical signal is ruled by the synaptic efficiency.

The artificial neuron, displayed in figure 4.2(a), imitates the layout and basic function of the biological neuron. The dendrites are substituted by an input vector $\vec{x}$, whereas the intensities of the individual inputs are weighted by factors w_{ij} corresponding to the synaptic efficiencies.

The output of the neuron is activated if the weighted sum of the inputs exceeds a certain threshold, commonly modeled by the sigmoid function

$$S(t) = \frac{2}{1 + e^{-t}} - 1, \tag{4.2}$$

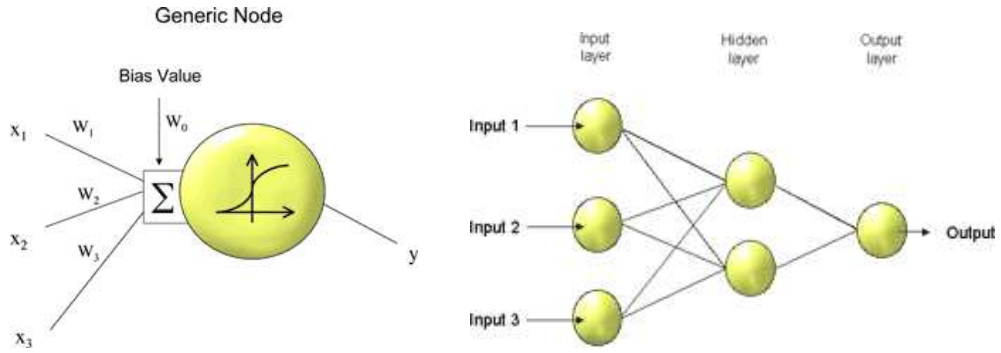
which is symmetric with respect to the point of origin and maps $]-\infty, +\infty[$ to $[-1, 1]$ (see figure 4.1). Hence, the output o_j of the artificial neuron j with n inputs x_i linked to the neuron can be expressed by

$$o_j = S\left(\sum_i^n w_{ij} x_i - \mu_j\right), \tag{4.3}$$

where μ_{ij} is an additional bias controlling the signal threshold of the neuron.

4.1.1 Feed Forward Networks

The configuration of the artificial neurons within a neural network gives rise to several different networks topologies. Since for the purpose of this analysis only the three–layered *feed forward network* with a single output is of interest, the following brief discussion will focus on this specific network topology. The denotation *feed forward* means that information is transferred from the n input nodes x_i via the m nodes y_j of the *hidden layer* to the single output node o in one direction only. Figure 4.2(b) shows a simple representation of a three layer feed forward network. The number



(a) Layout of an artificial neuron (b) A simple three layer feed forward network

Figure 4.2: The artificial neuron, also called node, represents the basic element for any neural network.

of input nodes corresponds to the number of variables that are assumed to contain important information for solving the separation problem. Normally an additional bias node with a constant value of 1 is added to the input layer. The number of nodes contained in the hidden layer is allowed to be arbitrary. However, with too many hidden nodes a network can tend to learn specific features by heart, losing the capability to generalize. In the opposite case, if the number of nodes in the hidden layer is too small, the network might not be able to cover and process all information sufficiently. As a general guideline, the number of hidden nodes is chosen to be similar to the quantity of input variables.

Combining the considerations discussed so far, the output of a three layer feed forward network with n input nodes, m hidden layer nodes and a single output node o is calculated by

$$o = S\left(\sum_{j}^{m} v_j S\left(\sum_{i}^{n} w_{ij}x_i - \mu_j\right)\right), \tag{4.4}$$

where v_j denotes the weight of the connection of a hidden layer j to the output node.

4.1.2 Neural Network Training

For the application of neural networks historical or simulated data with known classifications need to be available. These data are used to train a neural network allowing the network to comprehend certain patterns and thus gain the capability to separate events with a particular classification out of a non–classified data sample. To put it precisely, a network training aims at determining the weights of the individual connections in such a way that the value of the network output, corresponding to a certain classification, coincides with the known classification of the training event. This is performed by iteratively adjusting the weights. After each iterative step the difference between the true and the calculated network output value is compared by calculating a *cost function* E. One possible representation of a cost function E is the sum over

the squared differences between the network outputs o_i and the target values t_i

$$E = \sum_{i}^{N} \left(o_i\left(\vec{w}\right) - t_i\right)^2, \tag{4.5}$$

where N denotes the number of training events and $\vec{w}$ the vector of weights. Another possibility is given by the *entropy function*

$$E = \sum_{i}^{N} -\ln\left(\frac{1}{2}\left(1 + o_i\left(\vec{w}\right) \cdot t_i\right)\right). \tag{4.6}$$

The entropy function used in the NeuroBayes® package (see next section) contains an additional summand $\tau \cdot \frac{1}{2}\sum_i w_i^2 = \tau \cdot P$, called the *penalty term*, in order to suppress oscillations around minima and to keep the training in the learning region. The purpose of the penalty term is to punish the network for high weights, resulting in weight decay. This procedure represents one possible *regularization* method to optimize the training of a neural network.

Within the training procedure the particular set of weights is chosen which minimizes the cost function in the multidimensional parameter space. For the minimization of the cost function the method of steepest descent [35] can be used. Here the change Δw_{ij} of the weights is oppositely proportional to the respective gradient,

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}}, \tag{4.7}$$

where the step size η is the constant of proportionality. In NeuroBayes® η is adjusted automatically for each weight and for every iteration.

4.1.3 NeuroBayes®

All neural networks applied within this analysis are based on the NeuroBayes® package [36] providing tools suitable for the prediction of probability density functions and for binary classification problems. For the latter type of problems three layer feed forward networks with a single output node are used. NeuroBayes® was developed at the University of Karlsruhe and is also adopted beyond physical problems at the spin off company <phi-t>®. In the following subsections two of the main NeuroBayes® features will be outlined in a nutshell.

The Bayesian Approach

As its name indicates NeuroBayes® makes use of Bayes' theorem [35]:

$$P(A\,|B) = \frac{P(B\,|A) \cdot P(A)}{P(B)} \tag{4.8}$$

where the conditional probability of A assuming a given B, $P(A\,|B)$, is connected to the conditional probability of B assuming a given A, $P(B\,|A)$. $P(A)$ denotes the *a priori* probability to measure A which is independent of the prior probability $P(B)$. If one identifies A as a hypothesis and B as observed data, the Bayesian theorem provides an *a posteriori* probability $P(A\,|B)$ for the hypothesis to hold true for a given event.

Assuming that the ratio of signal and background is the same for training data and the data to be classified, in the case of a binary decision problem the output of a well trained NeuroBayes® network can be interpreted as a Bayesian a posteriori probability.

Preprocessing

Prior to the neural network training process where the optimal configuration of weights is determined, the input variables are processed in a way that locating the minimum of the cost function proceeds more effective. For this purpose, the NeuroBayes® package features a powerful and sophisticated variable preprocessing that provides a wide range of global and individual preprocessing options. First, the distributions of input variables are flattened and mapped onto the interval $[-1, +1]$ and transformed into Gaussian distributions with mean 0 and width 1. In order to dissolve correlations between input variables, the covariance matrix is diagonalized by a series of rotations. The minimization process benefits from a decorrelated set of input variables significantly.

By means of several methods that can be activated by individual preprocessing options, or *preprocessing flags*, certain features of the input variables can be treated individually. A method often applied throughout this analysis is to fit and replace a distribution of a variable with *spline functions*, suppressing fluctuations of the distribution that might cause overtraining of the network.

Yet another important feature of the preprocessing is the iterative calculation of the significances of the input variables. Starting from an initial set of N variables, the total correlation to the *target* is calculated. Then one training variable after the other is omitted and for each set of variables the loss of total correlation is computed. The variable causing the smallest loss of information is then eliminated so that the new set of variables entering the next iteration contains $N-1$ variables. By repeating this procedure up to the point where there is no variable left one obtains a list of correlations to the target and of the significances for the complete set of input variables. The correlations between the variables can be illustrated in a graphical representation of the correlation matrix, which is given in an analysis file at the end of each training: Here the coefficients $\rho_{ij} = c_{ij}/\sigma_i\sigma_j$ of the covariance matrix c_{ij} are displayed color–coded according to their magnitude.

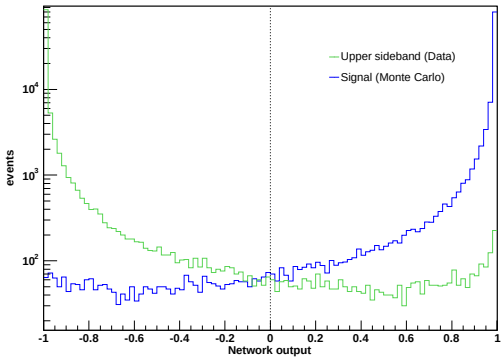


Figure 4.3: Example of a neural network classification of the training data. The network output -1 stands for pure background, $+1$ for signal.

Evaluating the Neural Network Performance

The analysis file, which is generated subsequent to a training, contains comprehensive information about preprocessing results for each individual variable and further plots indicating the quality and potential performance of a neural network.

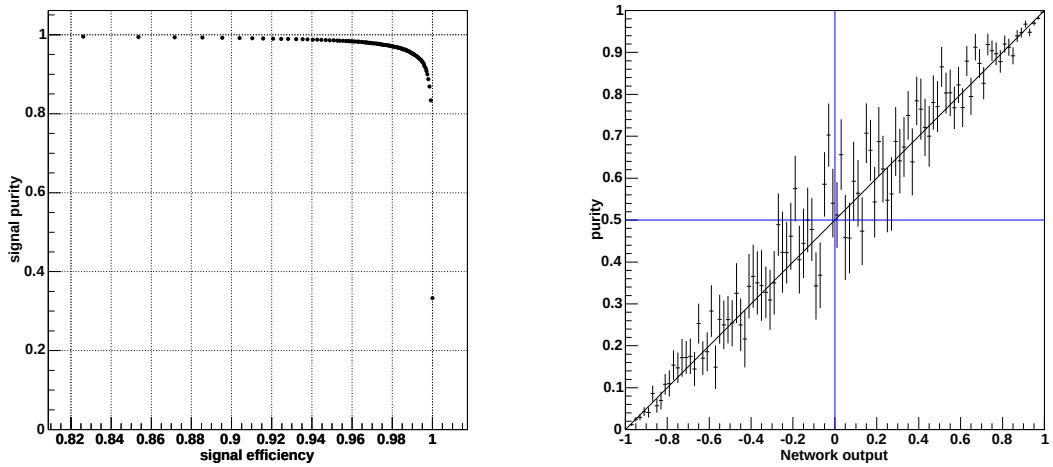
Figure 4.3 shows an example of the network classification of training data. Here the network output ranges between $[-1, +1]$ corresponding to the result of the binary classification of the training data where -1 stands for background and $+1$ for signal.

Introducing the quantities *purity* P and *efficiency* ϵ provides further indicators for the qualitiy and performance of a neural network:

$$P\,(\text{nnout}) = \frac{N_S\,(> \text{nnout})}{(N_S + N_B)\,(>\text{nnout})} \tag{4.9}$$

$$\epsilon\,(\text{nnout}) = \frac{N_S\,(> \text{nnout})}{N_S}. \tag{4.10}$$

Here, N_S and N_B denote the signal and background events respectively obtained for a given cut on the network output *nnout*. Figures 4.4(a) and 4.4(b) exemplarily show possible graphical representations of the purity, once plotted against the efficiency and once against the network output. The output can be mapped onto the interval $[0, 1]$ and, in the case of an optimal trained neural network, it can be interpreted as a probability. This assumption for a given network training is allowed if the purity in figure 4.4(b) lies on the diagonal. In the purity–efficiency plot the data point being closest to the coordinate $(1, 1)$ can be used to select an optimal working point. Using figure 4.4(b) this can be translated into a value of the neural network output which one may prefer to cut on.



(a) Purity–efficiency plot

(b) The purity plotted against the neural network output

Figure 4.4: Examples of NeuroBayes® plots indicating the performance of a neural network.

4.2 Data Samples

This study uses data taken between February 2002 and January 2007 during the CDF Run–II operation period. The full dataset is split up into several subsets, namely `xbhd0d`, `xbhd0h`, `xbhd0i` and `xbhd0j` (CDF internal notation), recorded during different data acquisition periods. From the `xbhd0j` dataset only data up to the run number 233111 enter the analysis, equivalent to total statistics corresponding to an integrated luminosity of 1.7 fb^{-1}. In the next sections the steps performed to generate and prepare the data samples being available for this analysis are outlined.

4.2.1 Event Reconstruction

The starting point of the generation of the used data is the *Two Track Trigger* (TTT), which is an essential part of the CDF–II data acquisition system for many analyses in the sector of B meson physics. Every event candidate collected in the given dataset was selected by the TTT. The predefined trigger paths of the TTT have already been outlined in section 2.3.5. For these triggered events raw tracks are reconstructed by means of several tracking algorithms implemented in the CDF–II offline software (see brief discussion in section 2.3.2).

The next step comprises the reconstruction of exclusive B meson decays performed with the `BottomMods` [37, 38] software package. Starting from simple objects, like reconstructed tracks, the modular–built software reconstructs high level objects, like

High level objects		Low level objects	
$B_s \rightarrow$	$D_s^+ \rightarrow$	$\phi\pi^+ \rightarrow$	$K^+K^-\pi^+$
	$D_s^- \rightarrow$	$\phi\pi^- \rightarrow$	$K^+K^-\pi^-$
$B_s \rightarrow$	$D_s^+ \rightarrow$	$\phi\pi^+ \rightarrow$	$K^+K^-\pi^+$
	$D_s^- \rightarrow$	$K^{0*}K^- \rightarrow$	$K^+\pi^-K^-$
$B_s \rightarrow$	$D_s^+ \rightarrow$	$\phi\pi^+ \rightarrow$	$K^+K^-\pi^+$
	$D_s^- \rightarrow$		$\pi^+\pi^-\pi^-$
$B_s \rightarrow$	$D_s^+ \rightarrow$	$K^{0*}K^+ \rightarrow$	$K^+\pi^-K^+$
	$D_s^- \rightarrow$	$K^{0*}K^- \rightarrow$	$K^+\pi^-K^-$
$B_s \rightarrow$	$D_s^+ \rightarrow$	$K^{0*}K^+ \rightarrow$	$K^+\pi^-K^+$
	$D_s^- \rightarrow$		$\pi^+\pi^-\pi^-$
$B_s \rightarrow$	$D_s^+ \rightarrow$		$\pi^+\pi^-\pi^+$
	$D_s^- \rightarrow$		$\pi^+\pi^-\pi^-$

Table 4.1: Decay chains of the studied $B_s \rightarrow D_s^+ D_s^-$ decays. The highest level objects are the exclusively reconstructed B_s mesons, the lowest level objects represent the stable particles, pions and kaons, where the reconstruction chain starts from.

B mesons, by inverting the sequence of the decay chain. The output of each module serves as input to the next one with the last module providing a list of candidates for a particular meson. The bottom–up reconstruction with `BottomMods` comprises the following steps:

- All tracks are required to have a successful helix fit.
- A collection of pion and kaon candidates is obtained by refitting the selected tracks with either a pion or kaon mass hypothesis.
- The collection of stable particle candidates is used to reconstruct a collection of unstable particles of a given event, which in turn can be used to reconstruct unstable particles at a higher level of the decay hierarchy.
- Dependent on the decay topology, the highest level object is obtained by combining a certain collection of unstable and stable particles of the previous reconstruction steps.

Throughout this thesis six hadronic B_s and three hadronic B^0 decay modes are analyzed (see tables 4.1 and 4.2 for details). All of these decay channels have six charged tracks in the final state and most of the channels share a similar decay topology with two secondary resonances decaying into two particles each. Therefore the discussion of the meson reconstruction is exemplarily given for the decay $B_s \rightarrow D_s^+(\phi\pi^+)D_s^-(K^{0*}K^-)$ only.

High level objects			Low level objects
$B^0 \rightarrow$	$D_s^+ \rightarrow$	$\phi\pi^+ \rightarrow$	$K^+K^-\pi^+$
	$D^- \rightarrow$		$K^-\pi^+\pi^-$
$B^0 \rightarrow$	$D_s^+ \rightarrow$	$K^{0*}K^+ \rightarrow$	$K^+\pi^-K^+$
	$D^- \rightarrow$		$K^-\pi^+\pi^-$
$B^0 \rightarrow$	$D_s^+ \rightarrow$		$\pi^+\pi^-\pi^+$
	$D^- \rightarrow$		$K^-\pi^+\pi^-$

Table 4.2: Decay chains of $B^0 \rightarrow D_s^+ D^-$ decays being subject to this analysis.

The reconstruction of a B_s in the chosen decay mode starts with the selection of a ϕ candidate. For this, as a first step two oppositely charged tracks assumed to be kaons are combined. However, one has to account for the fact that most of these combinations do not originate from a real ϕ decay but are just random and therefore are treated as background. To reject a large amount of these background events, a full vertex fit is performed and from that a candidate's invariant mass based on the sum of the four momenta of the kaon track candidates is calculated. The calculation of the vertices, invariant masses, particle four momenta and the corresponding covariance matrices is done with the `CTVMFT` [39] software package. Since the computation needed for a full vertex fit is rather time–consuming, prior to performing the vertex fit some loose cuts and particularly a soft preselection on the estimated raw invariant mass of a track pair are applied. A selected track pair is accepted as a ϕ candidate if the vertex fit was successful and the invariant mass lies in the window 1.005 GeV/c^2 $< M_\phi <$ 1.035 GeV/c^2. In order to form a K^{0*} candidate for the other decay branch, two oppositely charged tracks are assumed to stem from a pion and a kaon. This candidate is only accepted if the invariant mass after a successful vertex fit does fulfill the requirement 0.837 GeV/c^2 $< M_{K^{0*}} <$ 0.947 GeV/c^2. Adding one further charged track to each decay branch, which is supposed to be a pion in the ϕ case and a kaon in the K^{0*} case, a D_s^+ and a D_s^- candidate is formed. In the following vertex fit the tracks belonging to the selected candidates for the decays $D_s^+ \rightarrow \phi\pi^+$ and $D^- \rightarrow K^{0*}K^-$ are required to come from one common vertex each. All the loosely matching combinations of charged D_s meson candidates allowing for the reconstruction of a $B_s \rightarrow D_s^+ D_s^-$ are accepted if the vertex fits succeeded and their respective invariant masses lie within 1.87 and 2.07 GeV/c^2 corresponding to a mass window of +100 MeV/c^2 around the D_s mass of 1.968 GeV/c^2 published by the *Particle Data Group* (PDG) [4]. For performing the B_s vertex fit the D_s mass is constrained to its PDG value. At this stage, in addition the B_s candidates have to fulfill the specifications of the Two Track Trigger and the requirements of the `B_LOWPT` `B_CHARM` and `B_HIGHPT` trigger subpaths.

4.2.2 *B Stntuples* and *flat Ntuples*

The data produced by the reconstruction of B_s or B^0 candidates are stored in a special data container, called *B Stntuple* [40], which was custom–built for the purpose of storing data of this kind and performing studies in the field of B meson physics. *B Stntuples* hold all the vertex fit information obtained in every stage of the reconstruction chain and for each accepted meson candidate and decay particle. However, the format is very consuming in terms of mass storage ($\mathcal{O}$(TB)) and requires to run the CDF software to use the data for physics studies. Therefore, in the course of the analysis presented in [41] a data format was developed, called *flat Ntuples*, which allowed to reduce the amount of data significantly by streamlining the data structure and exclusively storing data belonging to one single decay mode of interest. In addition, working with *flat Ntuples* benefits from the fact that they can directly be analyzed by using the widely spread `ROOT` software framework [42] which is an essential tool for this analysis as well. Henceforth, *flat B Stntuples* will be mostly just referred to as *Ntuples*.

4.2.3 Monte Carlo Simulations

In addition to experimental data realistic simulations reflecting the physical behavior of true B_s and B^0 events in the studied decay modes are an essential input. These simulated data, that are generated by Monte Carlo (MC) techniques, are needed for various steps in the analysis.

Monte Carlo Ntuple production comprises several successive stages, starting with the creation of B mesons based on NLO calculations[1] [43] by using the event generator `BGenerator` (BGen)[44]. In contrast to the widely–used PYTHIA [45] event generator which creates $b\bar{b}$ pairs, BGen only simulates single B mesons without the anti–bottom quark and fragmentation products. Fragmentation processes are implemented via the Peterson fragmentation function [46]. For the simulation of the different B decay chains the package `EvtGen` [47, 48] is used. In order to reflect effects encountered in the experimental measurement and data taking processes additional simulation packages are needed. The simulation of the detector response is implemented with the `cdfSim` [49] package, the Two Track Trigger simulation is performed by using `TRGSim++` [50].

The MC datasets are generated in such a way that they possess the same Ntuple data structure as the experimental datasets. Hence, the same reconstruction and Ntuple production mechanisms can be adopted.

4.3 Candidate Selection Strategy

Since the Ntuples used within this analysis still occupy several GB in terms of disk storage, corresponding to the fact that most of the B_s and B^0 meson candidates

[1] *Next to Leading Order*

stem from random combinatorics, the datasets have to be further reduced by placing additional constraints to the candidate attributes. This is achieved by applying loose pre–cuts on certain variables stored in the Ntuple data structure so that obvious background events are sorted out. This enables a much smoother operation of artificial Neural Networks which are used for the final classification and selection of $B_s \to D_s D_s$ and $B^0 \to D D_s$ decay events.

In order to improve the track quality, all the Ntuples originating from experimental data and likewise the respective MC Ntuples datasets share the following common pre–cuts per track:

- Number of COT stereo hits ≥ 10
- Number of COT axial hits ≥ 10
- Number of Silicon Stereo + Silicon axial hits ≥ 3
- Minimum track transverse momentum $p_T \geq 0.35$ GeV/c
- A maximum two–dimensional $\chi^2_{r\phi} < 40$ in the B_s vertex fit

Moreover, candidates reconstructed from tracks with wrong charge combinations are rejected. The charge constraints depend on the final pion and kaon states in the respective decay mode and are listed in tables 4.3 and 4.4 where all the applied pre–selection cuts are given.

As far as the decay $B_s \to D_s^+ D_s^-$ is concerned, only very few events are expected to be observed in experimental data which is due to the small branching fractions of the studied hadronic D_s decay modes. Therefore, the pre–cuts are chosen to be rather conservative. The selection of variables and cut values is mainly based upon the results of a first neural network training performed for each channel. From that, the flattened distributions of input variables which are generated in the course of the variable pre–processing are studied individually. Some variables exhibit a distinct separation of signal and background. A cut value can be estimated in such a way that regions of the flattened distribution containing a vanishing number of signal events are cut off. To illustrate this procedure, figure 4.5(a) shows the flattened distribution of the quantity $|d_0\,(B_s)|$ used as an input variable in the network training for the channel $B_s \to D_s(\phi\pi) D_s(\phi\pi)$. From this distribution a cut sorting out obvious background is estimated. In figure 4.5(b) the original distribution of $d_0\,(B_s)$ is shown after the chosen cut has been applied. By using this pre–cut the long tails to the left and right of the simulated signal peak (MC), corresponding to non–B_s meson events with a large displacement with respect to the primary vertex, are cut off. This plot also illustrates the pursued strategy of applying soft pre–cuts only. Besides, in the case of the exemplified decay channel $B_s \to D_s(\phi\pi) D_s(\phi\pi)$ one is not interested in placing too hard cuts because a sufficient amount of background has to be conserved for performing the neural network training (see section 4.4.1). The large efficiency of background reduction already achieved by soft pre–cuts is caused by the fact that

$D \to K\pi\pi$, $D_s \to \phi\pi$, $\phi \to KK$	
$q(\pi_{D_s}) + q(K_D) + q(\pi^1_D) + q(\pi^2_D)$	$\stackrel{!}{=} 0$
$\lvert d_0(B^0)\rvert$	< 0.02 cm
$L_{xy}/\sigma_{L_{xy}}(B^0)$	> -2

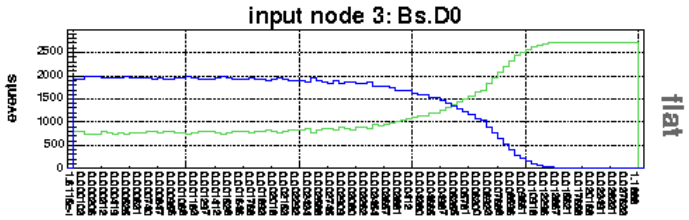
$D \to K\pi\pi$, $D_s \to K^{0*}K$, $K^{0*} \to K\pi$	
$q(K_{D_s}) + q(K_D) + q(\pi^1_D) + q(\pi^2_D)$	$\stackrel{!}{=} 0$
$p_T(D_s)$	> 1.75 GeV/c
$\lvert d_0(B^0)\rvert$	< 0.02 cm
$L_{xy}/\sigma_{L_{xy}}(B^0)$	> -2

$D \to K\pi\pi$, $D_s \to \pi\pi\pi$	
$p_T(D_s)$	> 1.2 GeV/c
$\lvert d_0(B^0)\rvert$	< 0.02 cm
$L_{xy}/\sigma_{L_{xy}}(B^0)$	> -1.3
$L_{xy}/\sigma_{L_{xy}}(D_s)$	> 3.3
$L_{xy}/\sigma_{L_{xy}}(D)$	> 3.3

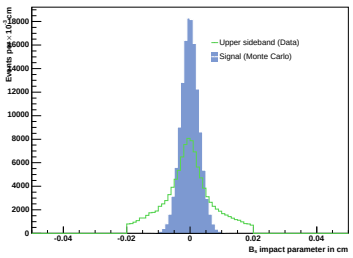
Table 4.3: Applied pre–cuts for the studied $B^0 \to D^+ D_s^-$ decays. See section A in the Appendix for the definition of variables.

$D_s^{1,2} \to \phi\pi,\ \phi \to KK$		$D_s^1 \to \phi\pi,\ D_s^2 \to K^{0*}K,\ K^{0*} \to K\pi$	
$q\left(\pi_{D_s^1}\right) + q\left(\pi_{D_s^2}\right)$	$\stackrel{!}{=} 0$	$q\left(\pi_{D_s^1}\right) + q\left(K_{D_s^2}\right)$	$\stackrel{!}{=} 0$
$p_T\left(B_s\right)$	> 5.5 GeV/c	$p_T\left(B_s\right)$	> 3.5 GeV/c
$p_T\left(D_s^2\right)$	> 0.5 GeV/c	$p_T\left(D_s^1\right)$	> 0.5 GeV/c
$\lvert d_0\left(B_s\right)\rvert$	< 0.02 cm	$\lvert d_0\left(B_s\right)\rvert$	< 0.02 cm
$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	> -2.3	$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	> -2.5
$L_{xy}/\sigma_{L_{xy}}\left(D_s^1\right)$	> 0.0	$L_{xy}/\sigma_{L_{xy}}\left(D_s^1\right)$	> -1.2
$L_{xy}/\sigma_{L_{xy}}\left(D_s^2\right)$	> 0.0		
$D_s^1 \to \phi\pi,\ D_s^2 \to \pi\pi\pi,\ \phi \to KK$		$D_s^{1,2} \to K^{0*}K,\ K^{0*} \to K\pi$	
$q\left(\pi_{D_s^1}\right) + \sum_{i=1}^{3} q^i\left(\pi^i_{D_s^2}\right)$	$\stackrel{!}{=} 0$	$q\left(K_{D_s^1}\right) + q\left(K_{D_s^2}\right)$	$\stackrel{!}{=} 0$
$\lvert d_0\left(B_s\right)\rvert$	< 0.02 cm	$p_T\left(D_s^2\right)$	> 1.0 GeV/c
$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	> -2.8	$\lvert d_0\left(B_s\right)\rvert$	< 0.011 cm
nnout	> -0.995	$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	> -1.8
$D_s^1 \to K^{0*},\ D_s^2 \to \pi\pi\pi,\ K^{0*} \to K\pi$			$D_s^{1,2} \to \pi\pi\pi$
$q\left(K_{D_s^1}\right) + \sum_{i=1}^{3} q^i\left(\pi^i_{D_s^2}\right)$	$\stackrel{!}{=} 0$	$\sum_{j=1}^{2}\sum_{i=1}^{3} q^i\left(\pi^i_{D_s^j}\right)$	$\stackrel{!}{=} 0$
$p_T\left(D_s^1\right)$	> 1.0 GeV/c	$\lvert d_0\left(B_s\right)\rvert$	< 0.02 cm
$\lvert d_0\left(B_s\right)\rvert$	< 0.02 cm	$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	> -4
$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	> -2.8	*nnout*	> -0.97
nnout	> -0.99		

Table 4.4: Applied pre–cuts for the studied $B_s \to D_s^+ D_s^-$ decays that are used for preparing data for the next stage of candidate selection. See section A in the Appendix for the definition of variables.



(a) Flattened distribution of $|d_0\,(B_s)|$



(b) $d_0\,(B_s)$ distribution after pre–cut

Flattened distribution of the absolute value of the impact parameter, $|d_0\,(B_s)|$ (left), used for the $B_s \to D_s(\phi\pi)D_s(\phi\pi)$ network. Original distribution of the impact parameter after a soft pre–cut of $|d_0\,(B_s)| < 0.02$ cm has been applied (right).

both the D_s mesons, which are narrow states, decay into a ϕ meson also being a narrow state. Therefore signal and background are well separable.

For decay channels with small expected signal yields and where no sufficient background suppression by means of pre–cuts could be achieved, a very loose pre–cut on a network output based on a first training was applied. In doing so a dramatic background reduction was observed while signal efficiency in Monte Carlo was larger than 99% in all three affected channels.

Subsequent to the pre–selection, artificial neural networks are employed in the next stage of the selection of signal events. A brief introduction to theory and utilization of artificial neural networks was already given at the beginning of this chapter. Neural network training results are shown in section 4.4.2 and in section B of the Appendix. Finally, the observed signal yields obtained by the exclusive reconstruction and selection of B_s and B^0 meson will be presented in section 4.5.

4.4 Exclusive B_s and B^0 Meson Selection

Measuring the branching fraction $Br\,[B_s \to D_s^+ D_s^-]$ requires a sufficient amount of signal events both in the decays $B_s \to D_s^+ D_s^-$ and $B^0 \to D^+ D_s^-$. A quantity describing the performance of an event selection in a data sample is the *Significance* $S = \frac{N_S}{\sqrt{N_S+N_B}}$ where signal events N_S and background events N_B are taken from a predefined region. Therefore, in many analyses this quantity is the optimized figure of merit to obtain a good candidate selection to be used for further analytical steps. This analysis makes use of the NeuroBayes® package for data classification and candidate selection. The general ideas that were briefly discussed before are now specified in more detail.

4.4.1 NeuroBayes® Network Training and Data Classification

In order to perform a training of artificial neural network, a set of variables sufficiently representing the information contained in the training patterns needs to be in place.

Composition of Training Data Samples

Training patterns comprise two distinctive data samples which are a priori known to be signal or background events. Signal events are generated by means of Monte Carlo simulations. The other sample has to reflect the random combinatorial behavior of non–signal events. Since current physical models used in Monte Carlo simulations are not able to model the complex quark production and hadronization processes sufficiently, background samples are taken from real data. To put it more precisely, usually one uses data from the upper and lower sidebands of the invariant mass spectrum that are assumed to contain a vanishing amount of signal events, but are as close as possible to the invariant mass of the signal. However, in the case of $B_s \to D_s^+ D_s^-$ and $B^0 \to D^+ D_s^-$, background data sampling is restricted to the upper sideband because the lower mass region also contains a huge amount of signal like decay events, mainly partially reconstructed B_s and B^0. These will be discussed in section 4.5 in more detail. Since the upper sideband is very similar to the signal region in kinematical respects, it is assumed to provide a good description of the background in the signal region. For all the studied decay channels of $B_s \to D_s^+ D_s^-$ ($B^0 \to D^+ D_s^-$) background data was extracted from an upper sideband starting at 5.45 GeV/c^2 (5.35 GeV/c^2). The upper limit of the sideband depends on the amount of background data desired for a given network training and thus varies from channel to channel. Figure 4.5 shows an example of a typical composition of training data samples.

Prior to serving as an input for network training, Monte Carlo and background data are prepared by likewise applying the set of pre–cuts specified in section 4.3.

Selection of Input Variables

The selection of input variables assumed to provide maximum information needed for an effective discrimination between signal and background events is critical to the network performance. There are different strategies for determining an optimal set of variables. First, one might start with a small set that is empirically known to provide for a good discrimination for a given problem. For this, variables used in a cut based analysis certainly form a good starting point. Further variables are added one by one to the original set, and for each set of variables a network training including preprocessing is performed. Variables not adding significant information or exhibiting large correlations to other variables are sorted out again. This procedure is repeated until no further variables are found adding any discriminating power. Here, one of the basic rules is to keep the set of input variables as small as possible in order to improve the network's generalization capabilities and prevent it from learning features by heart.

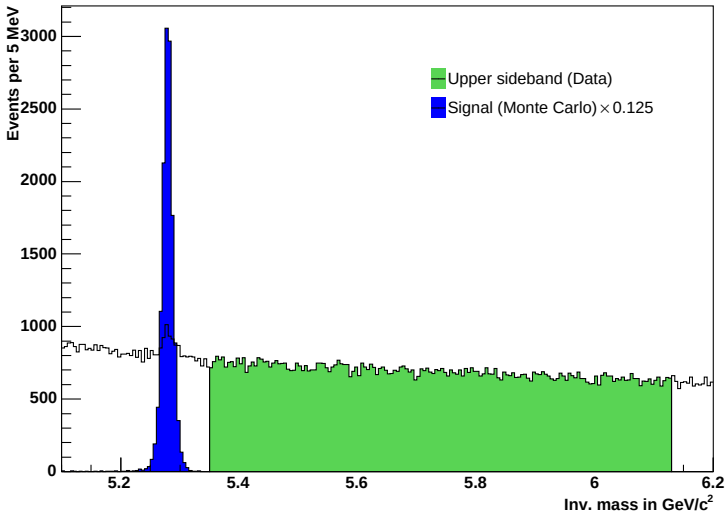


Figure 4.5: Composition of data typically used for a neural network training. The samples are composed of simulated signal events (Monte Carlo) and background extracted from the upper sideband of the invariant mass spectrum. The example represents the training data used for the channel $B^0 \to D^+ D_s^- \to \phi\pi K\pi\pi$. Please note that the representation does not display data true to scale.

Within this thesis however, the contrary procedure to this bottom–up approach is used. A large set of training variables is fed into the training preprocessing and a first network training is performed. This set is composed of variables that might hold any kind of relevant information for the given classification problem, including variables used in a rectangular cut analysis. Now, on the basis of the variable preprocessing and the performance of the resulting neural network, single or groups of variables being highly correlated to others or holding only very small correlation to the target are removed. This procedure of sorting out variables and performing network trainings is iterated until removing further variables would result in a loss of significance being too high. The minimal significance required for the least significant training variable depends on the maximum significances of the variables providing for the majority of separation power and can hence vary from problem to problem.

The final sets of training variables for the particular decay channels are listed in Appendix B. For nearly all of the decays the input variable containing the bulk of information is the decay length divided by its uncertainty, $L_{xy}/\sigma_{L_{xy}}(B)$, of the B^0 or the B_s meson respectively, followed by the decay length significance $L_{xy}/\sigma_{L_{xy}}(D)$ of the D or D_s meson. Further variables extensively contributing to the discrimination between signal and background are variables containing information concerning the quality of the vertex fit in the reconstruction, $\chi^2_{r\phi}(B)$, the impact parameter $|d_0(B)|$ and various transverse momenta p_t. In addition to these variables shared by all studied

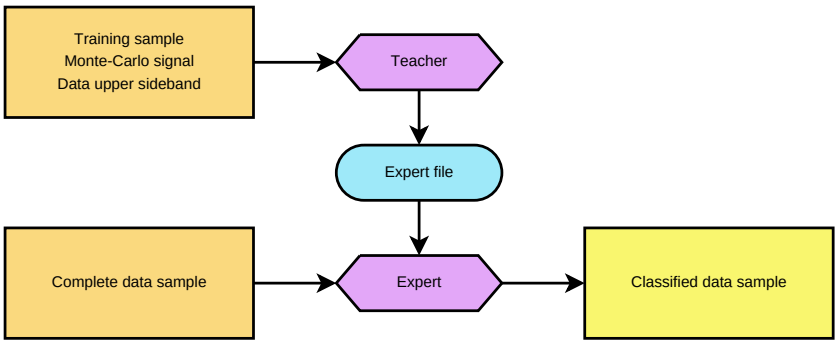


Figure 4.6: Diagram of the data classification process with NeuroBayes®.

decay channels, each neural network contains variables specific to the particular decay topology. For example, in the case of the channels where the D_s decays into three pions, the invariant masses of two of the three pions, $m_{\pi^i\pi^j}$, are introduced. The distributions of these variables contain resonances reflecting the intermediate meson states contained in the $D_s \to \pi\pi\pi$ decay.

NeuroBayes® Program Flow

Without going into detail, the program sequence when applying NeuroBayes® is briefly outlined. In order to perpare and run a neural network training, a program called *Teacher* is used to specify the set of input variables and the signal and background data samples. Furthermore one is allowed to set global and individual preprocessing flags in order to treat special features of the several input variables and to assign weights to training events. The training result, called *Expertise*, is stored in the *Expert File* holding all relevant information. As a last step, the *Expert* performs the classification of the dataset by exploiting the training information contained in the *Expertise*. In doing so, a variable containing the neural network output representing the classification of a given event is written onto the dataset. Figure 4.6 schematically illustrates the mode of operation.

4.4.2 Training Results

In this section the result of one neural network training is shown exemplarily. The training results of the networks trained for the remaining decay channels are presented in Appendix B. The compilation of results is confined to the essential input and output information: First, the variables used for the network are listed ordered by their significances. The correlations between the variables are illustrated by the color–coded correlation matrix. Finally, two of the characteristic graphs indicating a neural network's quality and performance are given: In the first one the linearity of the purity as a function of the network output is tested, in the second one a network's capability to separate between signal and background events is verified on the training data

Rank	Name	Significance [σ]	Index
-	Target	-	1
1	$L_{xy}/\sigma_{L_{xy}}(D_s^2)$	349.27	8
2	$L_{xy}/\sigma_{L_{xy}}(B_s)$	104.68	2
3	$\texttt{PID.R}_\pi(K^2_{\phi^2})$	83.52	24
4	$\chi^2_{r\phi}(B_s)$	57.52	4
5	$\lvert d_0(B_s)\rvert$	45.84	3
6	$\min(p_t)$	37.17	26
7	$\theta_{hel}(K^1_{D_s^2})$	33.56	21
8	$\texttt{PID.R}_K(K^1_{\phi^2})$	30.05	22
9	$\mathrm{m}(\phi_{D_s^2})$	28.38	11
10	$L_{xy}/\sigma_{L_{xy}}(D_s^1)$	25.38	6
11	$\min(d^0/\sigma_{d_0})$	18.79	27
12	$\texttt{PID.R}_K(\pi_{D_s^2})$	17.16	17
13	$\lvert d_0(D_s^2)\rvert$	16.41	9
14	$\theta_{hel}(K^1_{D_s^1})$	15.93	18
15	$\chi^2_{3D}(B_s)$	14.01	5
16	$\mathrm{m}(\phi^1)$	13.24	10
17	$\chi^2_{3D}(D_s^1)$	13.16	7
18	$d_0^{lts}/\sigma_{d_0}(\pi_{D_s^2})$	12.57	16
19	$\chi^2_{3D}(\phi_{D_s^2})$	12.10	13
20	$L_{xy}(D_s^2 \leftarrow \phi^2)$	11.56	25
21	$\lvert d_0(\pi_{D_s^2})\rvert$	11.43	15
22	$\texttt{PID.R}_K(\pi_{D_s^1})$	11.26	14
23	$d_0^{lts}/\sigma_{d_0}(K^1_{\phi^2})$	10.32	20
24	$\texttt{PID.R}_K(K^2_{\phi^1})$	10.16	19
25	$\texttt{PID.pullTof}_\pi(K^1_{\phi^2})$	8.78	23
26	$L_{xy}/\sigma_{L_{xy}}(\phi_{D_s^2})$	8.54	12

Table 4.5: The input variables of the $Bs \to D_s^1(\phi^1\pi)D_s^2(\phi^2\pi)$ network. $K^{1,2}$ denotes the first or second kaon from a ϕ meson decay. See Appendix A for variable definitions.

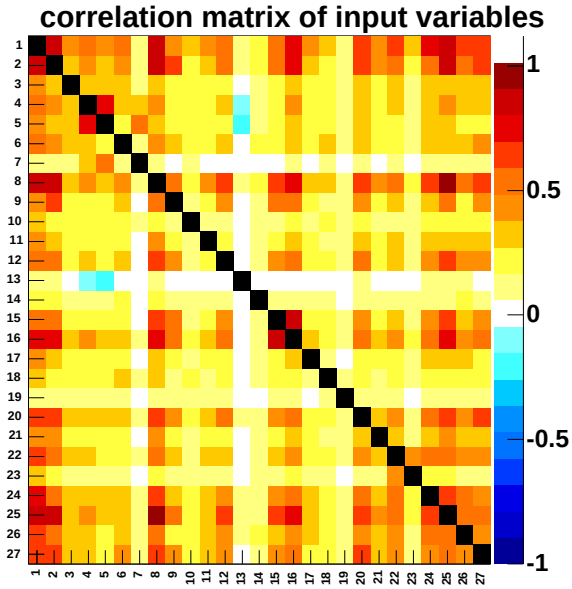
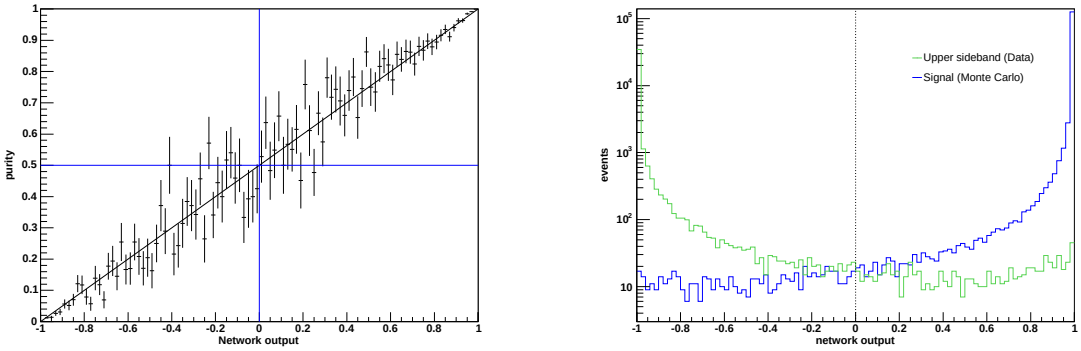


Figure 4.7: Correlation matrix of the input variables used for the decay $B_s \rightarrow D_s\,(\phi\pi)\,D_s\,(\phi\pi)$.



(a) Purity over network output.

(b) Network classification of training data.

Figure 4.8: Two of the training graphs indicating the quality and the performance of the network trained for the channel $Bs \rightarrow D_s\,(\rightarrow \phi\pi)\,D_s\,(\rightarrow \phi\pi)$.

sample.

4.5 Determination of Signal Yields

By running the *Expert* on the unclassified datasets the additional variable *nnout* holding an event by event classification encoded in a real number in the range $[-1, 1]$ is written onto the Ntuples. The selection of signal candidates is done by performing a cut on *nnout*, providing for an effective elimination of the vast amount of background events. A good candidate selection can be achieved by maximizing the significance $S = \frac{N_S}{\sqrt{N_S+N_B}}$ obtained after performing a fit the invariant mass spectrum resulting from a particular cut on the network output. For this, a statistical method for parameter estimation and a robust ansatz for the fit function need to be in place.

4.5.1 Extended Maximum Likelihood Method

For all fits to experimental data performed throughout this analysis the *unbinned extended maximum likelihood* method is applied [35]. In principle one is also allowed to use the *maximum likelihood* method on *binned* data which is faster in terms of computing time and sufficient in the case of high statistics. However, due to the small signal yields expected for the $B_s \to D_s D_s$ channels, treating data not on an event–by–event but only on a binned basis would not be advantageous in terms of precision. In the *unbinned maximum likelihood* method, considering a set of n independent measurements (or events) of the variable $\vec{x}$ which follows a normalized probability density $f(\vec{x} \mid \vec{a})$, an unknown set of parameters $\vec{a}$ is determined by maximizing the joint probability density function, called *likelihood function*

$$L(\vec{a}) = \prod_{i=1}^{n} f(\vec{x_i} \mid \vec{a}) \tag{4.11}$$

by setting

$$\frac{\partial L(a_k)}{\partial a_k} = 0. \tag{4.12}$$

For numerical and computational reasons the utilization of the *negative log likelihood function* $F(\vec{a})$ is more advantageous:

$$F(\vec{a}) = -\ln L(\vec{a}) = -\sum_{i=1}^{n} \ln f(\vec{x_i} \mid \vec{a}) \tag{4.13}$$

In order to find the set of parameters $\vec{a}$ maximizing the likelihood function $L(\vec{a})$, the negative log likelihood function $F(\vec{a})$ needs to be minimized by requiring:

$$\frac{\partial F(a_k)}{\partial a_k} = 0. \tag{4.14}$$

There are statistical problems where the average number of signal events by itself is a parameter to be determined. For this purpose, instead of a normalized probability density $f(\vec{x} \mid \vec{a})$ a function $g(\vec{x} \mid \vec{a})$ is used, where the integral over $g(\vec{x} \mid \vec{a})$ in a measuring range Ω corresponds to the quantity of expected events N:

$$N = \int_{\Omega} g(\vec{x} \mid \vec{a}) \, \mathrm{d}x \tag{4.15}$$

Thus, in the extended maximum likelihood method the minimum of the negative log likelihood function

$$F(\vec{a}) = -\sum_{i=1}^{n} \ln g(\vec{x_i} \mid \vec{a}) + N \tag{4.16}$$

in the multi-dimensional parameters space has to be determined. Within this analysis the fitting procedure using the unbinned maximum likelihood method is carried out in a special software environment, called the *fitter framework*. The fitter framework is built up in a modular way, hence providing a flexible structure suitable for different types of fitting procedures. The fit function, parameter management and certain features of the fit method are implemented by including an additional module, or `ModeClass`, which in turn is run by a steering executable. The `ModeClass` represents the interface to the `Fitter`, which performs the minimization process by using the program `MINUIT` and passes the determined parameter values back to the `ModeClass` object. The fitter framework provides the possibility for including several mode classes and sharing certain parameters among them, thus allowing for performing simultaneous fits on multiple decay channels. The software architecture of the fitter framework is outlined in [23] and [51].

4.5.2 General Features of the Fit Model

The ansatz for the negative log likelihood function being implemented in the mode class is set up in such a way that it is valid for all the studied $B_s \to D_s D_s$ and $B^0 \to D D_s$ decay channels respectively. In order to formulate a valid model, the critical features of the invariant mass spectrum in the analyzed range need to be understood and taken into account, leading to a fit model being built up of several individual template functions. The contributions encountered in the invariant mass spectra entering the fit function can roughly be classified into three categories:

- **Signal**
  Fully reconstructed B_s or B^0 meson characterized by a sharp resonant peak.

- **Combinatorial background**
  Random combinations of tracks not originating from the decay of a B meson that happened to pass the reconstruction and selection requirements.

- **Physics background**
  Events sharing several common features with true signal events. This category can be classified into two sub–types:

  Partially reconstructed events
  Events originating from real mesons where one neutral decay particle had not been reconstructed.

  Reflections and crosstalks from other decays
  Fully reconstructed decays of mesons that happen to occur in the signal range if a wrong particle hypothesis has been assigned to one of the particles in the final state, hence leading to the misreconstruction of another decay mode belonging to the same meson or originating from a different, but similar meson decay. This phenomena are usually referred to as *reflections* or *crosstalks*.

The different fit model contributions belonging to one of the categories quoted above are discussed in more detail for the different B_s and B^0 channels separately. To provide for a correct implementation, the various features require dedicated examinations mainly being based upon comprehensive Monte Carlo studies. However, due to constraints in terms of time and lack of Monte Carlo data available for the presented analysis, not all features of the discussed physical effects could be included in the fit model for all channels. If not quoted separately, physical background contributions were not implemented by means of an additional template but generally considered as background and hence included in the general background fit template function.

Since the huge number of physical background contributions encountered in the studied decay channels already has been discussed extensively in [3], the following discussion is confined to the templates contained in the fit model used in this analysis.

4.5.3 $B_s \to D_s D_s$ Fit Function

The model for performing an unbinned fit to the invariant mass spectra makes use of the following templates:

Signal Template

The shape of the signal peak is modeled by a normalized combination of two Gaussian distributions with differing widths but sharing a common mean value. With i being the index for the particular final state, the probability density for the signal reads as:

$$P^i_{sig} = f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i\sigma^j_1\right) + \left(1 - f^i_{\mathcal{G}1}\right)\mathcal{G}\left(m \mid \mu^i\sigma^i_2\right) \tag{4.17}$$

For each channel the widths and the mean of the double Gaussians are determined by a binned fit to a signal Monte Carlo sample (See section C in the Appendix). Due to the large number of parameters in the overall fit function, these parameters, as all shape parameters determined in MC pre–fits, are kept fixed for the final unbinned fit.

Combinatorial Background Template Function

Combinatorial background and further background contributions not being explicitly considered in separate templates are well described by a single exponential function:

$$P^i_{bg} = a^i \exp\left(m \mid b^i\right) \tag{4.18}$$

The shape parameters of this template are not determined by any pre–fitting and kept free in the global unbinned fit.

Templates for Partially Reconstructed Events

In addition to combinatorial background, the lower sideband region is dominated by $B_s \to D_s^{+(*)} D_s^{-(*)}$ decays which arise in the $B_s \to D_s D_s$ invariant mass spectrum as so called satellite peaks stemming from $D_s^{+*} \to D_s^+ \gamma\,(\pi^0)$, where the missing neutral π^0 or γ can not be reconstructed. Therefore these decays are referred to as partially reconstructed events. The main contribution comes from $D_s^{+*} \to D_s^+ \gamma$.

- $B_s \to D_s^{+*} D_s^-$
  For this decay a template made up of three Gaussians each having its own mean and width is used. The small admixture of $D_s^{+*} \to D_s^+ \pi^0$ produces a smooth double peak structure.

$$P^i_{p1} = f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i_1\sigma^i_1\right) + f^i_{\mathcal{G}2}\mathcal{G}\left(m \mid \mu^i_2\sigma^i_2\right) + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2}\right)\mathcal{G}\left(m \mid \mu^i_3\sigma^i_3\right) \tag{4.19}$$

- $B_s \to D_s^{+*} D_s^{-*}$
  This reflection generates a broad, featureless shape which is also fitted by a sum of three Gaussians:

$$P^i_{p2} = f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i_1\sigma^i_1\right) + f^i_{\mathcal{G}2}\mathcal{G}\left(m \mid \mu^i_2\sigma^i_2\right) + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2}\right)\mathcal{G}\left(m \mid \mu^i_3\sigma^i_3\right) \tag{4.20}$$

The shape parameters and fractions of the individual single Gaussians are fixed after performing the binned fit on Monte Carlo.

Reflection Templates

- The decay $B_s \to D_s^+\,(\phi\pi^+)\,D_s^-\,(K^{0*}K^-)$ has a strong reflection from $B^0 \to D_s^+\,(\phi\pi^+)\;D^-(K^-\pi^+\pi^-)$, which arises if one pion in the final state happens to be misreconstructed as a kaon, leading to a misreconstruction of a B^0 as a B_s meson. To tackle this reflection $B^0 \to D_s^+\,(\phi\pi^+)\,D^-\,(K^-\pi^+\pi^-)$ Monte Carlo samples were reconstructed as $B_s \to D_s^+\,(\phi\pi^+)\,D_s^-\,(K^{0*}K^-)$ and then subjected to the same pre–selection and final selection (neural network cut) measures as the polluted $B_s \to D_s^+ D_s^-$ counterpart. The Monte Carlo signal template was fitted with a double Gaussian:

$$P^i_{reflsig} = f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i\sigma^i_1\right) + \left(1 - f^i_{\mathcal{G}1}\right)\mathcal{G}\left(m \mid \mu^i\sigma^i_2\right) \tag{4.21}$$

The width and the fraction resulting from the MC pre–fit were fixed, the mean was allowed to float in the final unbinned fit to test the validity of the reflection model.

This reflection causes additional contributions, which arise due to the partial reconstruction of the B^0 meson because of neutral pions or photons coming from $D_s^{+(*)}D^{-(*)}$ decays being lost. Each of these contributions is fitted by means of a triple Gaussian and will be discussed together with the $B^0 \to D_s^+ D^-$ templates in the next section.

$$P^i_{rp1,2,3} = f^i_{\mathcal{G}1}\mathcal{G}\left(m \mid \mu^i_1\sigma^i_1\right) + f^i_{\mathcal{G}2}\mathcal{G}\left(m \mid \mu^i_2\sigma^i_2\right) + \left(1 - f^i_{\mathcal{G}1} - f^i_{\mathcal{G}2}\right)\mathcal{G}\left(m \mid \mu^i_3\sigma^i_3\right) \tag{4.22}$$

- $B_s \to D_s^+ \left(\phi\pi^+\right) D_s^- \left(\pi^+\pi^-\pi^-\right)$ has a faint reflection from $B_s \to D_s^+ \left(\phi\pi^+\right) \pi^+\pi^-\pi^-$ where the three pions falsely were reconstructed to directly come from the D_s^- vertex. Handling this reflection would require Monte Carlo and reconstructed real data for $B_s \to D_s^+ \left(\phi\pi^+\right) \pi^+\pi^-\pi^-$ to estimate a scale factor between Monte Carlo and real data. Applying this scale factor on the $B_s \to D_s^+\pi^+\pi^-\pi^-$ Monte Carlo, reconstructed as $B_s \to D_s^+ D_s^-$, would provide a correct normalization of a template obtained from this Monte Carlo. This procedure was performed in the course of the former CDF analysis [3] and the contribution stemming from this reflection was found to be small. Due to lack of Monte Carlo to reproduce the correct treatment, this reflection was neglected throughout this thesis.

- The situation gets even more complicated when looking at the three remaining decay modes $B_s \to D_s^+ \left(K^{0*}K^+\right) D_s^- \left(K^{0*}K^-\right)$, $B_s \to D_s^+ \left(K^{0*}K^{-+}\right) D_s^- \left(\pi^+\pi^-\pi^-\right)$ and $B_s \to D_s^+ \left(\pi^+\pi^-\pi^+\right) D_s^- \left(\pi^+\pi^-\pi^-\right)$ where the reflection modes originating from a wrong pion or kaon mass hypothesis are admixed. Developing approaches for eliminating or correctly handling these pollutions require thorough and comprehensive Monte Carlo studies which have not have been conducted yet.

The results of the binned fits for all the discussed templates can be found in section C in the Appendix. The combination of each listed template function gives rise to the following negative log likelihood function being used as an ansatz for fitting the $B_s \to D_s D_s$ invariant mass spectrum:

$$\begin{aligned} F^i(\vec{a}) \;= -2\sum_{k=1}^{n} \ln \;& [N^i_1 P^i_{sig}(m_k \mid \vec{a_{sig}}) + N^i_2 P^i_{p1}(m_k \mid \vec{a}_{p1}) + N^i_3 P^i_{p2}(m_k \mid \vec{a}_{p1}) \\ & + N^i_4 P^i_{rsig}(m_k \mid \vec{a}_{rsig1}) + N^i_5 P^i_{rp1}(m_k \mid \vec{a}_{rp1}) + N^i_6 P^i_{rp2}(m_k \mid \vec{a}_{rp2}) \\ & + N^i_7 P^i_{rp3}(m_k \mid \vec{a}_{rp3}) + N^i_8 P^i_{bg}(m_k \mid \vec{a}_{bg})] \\ & + 2\sum_{l=1}^{8} N^i_l, \end{aligned}$$

The factor of 2 in front of the sum is convention and i corresponds to one of the six studies B_s decay modes. The factors N^i_l in front of each probability density function

are the numbers of events of the given component. For modes where no reflection components were introduced, these are switched off by setting $N_4^i = N_5^i = N_6^i = N_7^i = 0$.

Prior to committing the log likelihood function to the fitting process, the function is reparameterized by relating the number of partially reconstructed events to the respective fully reconstructed signal events. This is motivated by the fact that one prefers not to have the absolute numbers of partially reconstructed events as fit parameters, but the ratios of these with respect to the B_s signal peak. For this, one exploits the relations

$$N_2^i = f_2 N_1^i \epsilon_{21}^i \tag{4.23}$$
$$N_3^i = f_3 N_1^i \epsilon_{31}^i. \tag{4.24}$$

ϵ_{lm}^i are the ratios of the combined reconstruction and selection efficiencies extracted from Monte Carlo (see section 4.5.5 for details),

$$\epsilon_{21}^i = \frac{\epsilon_2^i \left(B_s \to D_s^{+*} D_s^-\right)}{\epsilon_1^i \left(B_s \to D_s^+ D_s^-\right)} \tag{4.25}$$
$$\epsilon_{31}^i = \frac{\epsilon_3^i \left(B_s \to D_s^{+*} D_s^{-*}\right)}{\epsilon_1^i \left(B_s \to D_s^+ D_s^-\right)}. \tag{4.26}$$

f_2, f_3 are identified as the ratios of branching fractions

$$f_2 = f_{D_s^* D_s} = \frac{Br\left(B_s \to D_s^{+*} D_s^-\right)}{Br\left(B_s \to D_s^+ D_s^-\right)} \tag{4.27}$$
$$f_3 = f_{D_s^* D_s^*} = \frac{Br\left(B_s \to D_s^{+*} D_s^{-*}\right)}{Br\left(B_s \to D_s^+ D_s^-\right)}, \tag{4.28}$$

which directly serve as fit parameters in the rewritten negative log likelihood function now having the form:

$$\begin{aligned} F^i(\vec{a}^i) = -2 \textstyle\sum_{k=1}^n \ln \; & [N_1^i P_{sig}^i(m_k \mid \vec{a}_{sig}^i) + f_2 N_1^i \epsilon_{21}^i P_{p1}^i(m_k \mid \vec{a}_{p2}^i) \\ & + f_3 N_1^i \epsilon_{31}^i P_{p3}^i(m_k \mid \vec{a}_{p3}^i) + N_4^i P_{rsig1}^i(m_k \mid \vec{a}_{rsig1}) \\ & + f_5 N_4^i \epsilon_{54}^i P_{p1}^i P_{rp1}^i(m_k \mid \vec{a}_{rp1}^i) + f_6 N_4^i \epsilon_{64}^i P_{p1}^i P_{rp2}^i(m_k \mid \vec{a}_{rp2}^i) \\ & + f_7 N_4^i \epsilon_{74}^i P_{p1}^i P_{rp3}^i(m_k \mid \vec{a}_{rp3}^i) + N_8^i P_{bg}^i(m_k \mid \vec{a}_{bg}^i)] \\ & + 2[N_1^i(f_2 \epsilon_{21}^i + f_3 \epsilon_{31}^i) + N_4^i(f_5 \epsilon_{54}^i + f_6 \epsilon_{64}^i + f_7 \epsilon_{74}^i) + N_{bg}^i]. \end{aligned}$$

4.5.4 $B^0 \to D_s D$ Fit Function

As far as $B^0 \to D_s D$ is concerned, the discussion for setting up the log likelihood function and the individual templates involved closely follows the argumentation in the last section. The decay $B^0 \to D_s D$ additionally was studied because it serves as the normalization mode for the calculation of the relative branching fraction $Br[B_s \to$

$D_sD_s]/Br[B^0 \to D_sD]$. The templates for *signal* and for *combinatorial background* are the same as for $B_s \to D_sD_s$, namely the sum of two *Gaussian* distributions, having a common mean value, and a *single exponential* respectively. The topology of the lower sideband region of the B^0 invariant mass spectrum is very similar to that of the B_s meson, hence dominated by several satellite peaks originating from partially reconstructed $B^0 \to D_s^{(*)}D^{(*)}$ decays. Furthermore, there are some faint crosstalk contributions from other decay modes [3] as well, which are not accounted for in the B^0 fit model though, either due to being negligible or due to lack of Monte Carlo required for detailed studies.

Templates for Partially Reconstructed Events

The decays $B^0 \to D_s^{+*}D^{-*}$ produce a series of satellite peaks shifted to smaller masses approximately by one to two pion masses with respect to the $B^0 \to D_s^+D^-$ signal peak. There are three possible combinations:

- $B^0 \to D^{-*}D_s^+$
  The D^{-*} can decay either into D^- and a lost $\gamma\,(\pi^0)$, or into $D^0\pi^-$ which is not reconstructed. The latter effect has to be accounted for by the overall reconstruction efficiency. $B^0 \to D^{+*}D_s^-$ creates a distinct double peak structure. This results from the D^{+*} polarization in the B^0 decay. Therefore, the angle between the π^0 and the momentum helicity of the D^* follows a $\cos^2\theta$ distribution which corresponds to a preferred pion release direction in the direction of the D^{+*} or in opposition to it. The shape is fitted by a combination of three Gaussians.

- $B^0 \to D_s^{+*}D^-$
  The D_s^{+*} decay is dominated by $D_s^{+*} \to D_s^+\gamma$, $D_s^{+*} \to D_s^+\pi^0$ accounts for 5.8% only. The shape does not feature any distinctive characteristics and is fitted by the sum of three Gaussian distributions.

- $B^0 \to D_s^{+*}D^{-*}$
  For the decay of the B^0 into an excited D_s^{+*} and D^{*-} also a triple Gaussian is used. This mode generates a wide bump located about two pion masses below the $B^0 \to D_s^+D^-$ signal.

Combining these input information, the log likelihood function for the jth B^0 mode is set up analogously to the B_s case:

$$\begin{aligned} F^j(\vec{b}^j) \;= -2\sum_{k=1}^{n} \ln \;& [N_1^j P_{sig}^j(m_k \mid \vec{b}_{sig}^j) + N_2^j P_{p1}^j(m_k \mid \vec{b}_{p1}^j) + N_3^j P_{p2}^j(m_k \mid \vec{b}_{p2}^j) \\ & + N_4^j P_{p3}^j(m_k \mid \vec{b}_{p3}^j) + N_5^i P_{bg}^j(m_k \mid \vec{b}_{bg}^j)] \\ & + 2\sum_{l=1}^{5} N_l^j . \end{aligned}$$

In order to link the number of partially reconstructed events to the number of signal events, the likelihood function is modified by substituting

$$N_2^j = f_2 N_1^j \epsilon_{21}^j \quad (4.29)$$
$$N_3^j = f_3 N_1^j \epsilon_{31}^j \quad (4.30)$$
$$N_4^j = f_4 N_1^j \epsilon_{41}^j. \quad (4.31)$$

Here, f_2, f_3 and f_4 again are equivalent to the ratio of branching fractions:

$$f_2 = f_{D_s D^*} = \frac{Br\,(B^0 \to D_s^+ D^{-*})}{Br\,(B^0 \to D_s^+ D^-)} \quad (4.32)$$
$$f_3 = f_{D_s^* D} = \frac{Br\,(B^0 \to D_s^{+*} D^-)}{Br\,(B^0 \to D_s^+ D^-)} \quad (4.33)$$
$$f_4 = f_{D_s^* D^*} = \frac{Br\,(B^0 \to D_s^{+*} D^{-*})}{Br\,(B^0 \to D_s^+ D^-)} \quad (4.34)$$

The efficiency ratios ϵ_{lm}^j are defined analogously to their B_s pendants in section 4.5.3. Finally, with these modifications the negative log likelihood function entering the fitting procedure of the jth $B^0 \to D_s^+ D^-$ mode reads as follows:

$$\begin{aligned} F^j(\vec{b}^j) = -2 \textstyle\sum_{k=1}^n \ln \; & [N_1^j P_{sig}^j(m_k \mid \vec{b}_{sig}^j) + f_2 N_1^j \epsilon_{21}^j P_{p1}^j(m_k \mid \vec{b}_{p1}^j) \\ & + f_3 N_1^j \epsilon_{31}^j P_{p2}^j(m_k \mid \vec{b}_{p2}^j) + f_4 N_1^j \epsilon_{41}^j P_{p3}^j(m_k \mid \vec{b}_{p3}^j) \\ & + N_{bg}^i P_{bg}^j(m_k \mid \vec{b}_{bg}^j)] \\ & + 2[N_1^j (f_2 \epsilon_{21}^j + f_3 \epsilon_{31}^j + f_4 \epsilon_{41}^j) + N_{bg}^j]. \end{aligned}$$

4.5.5 Efficiencies

The combined reconstruction and selection efficiency $\epsilon_k^{i,j}$ for a certain decay sub–type k (e.g. fully or partial reconstructed decay) in the ith (jth) $B_s \to D_s D_s$ ($B^0 \to D_s D$) decay mode is extracted from Monte Carlo by calculating

$$\epsilon_k^{i,j} = \frac{N_{rec}^{i,j,k}|_{MC}}{f_k^{i,j} N_{tot}^{B_s,B^0}|_{MC}}, \quad (4.35)$$

where $N_{rec}^{i,j,k}$ is the number of events of a particular decay mode determined by a fit to the respective Monte Carlo sample, $f_k^{i,j}$ is the pre–set branching fraction for this specific decay in Monte Carlo and N_{tot} denotes the total inclusive number of simulated $B_s \to D_s D_s$ or $B^0 \to D_s D$ events. This relation needs some modification due to a systematic inconsistency between data and Monte Carlo: The fraction of candidates triggered by Trigger 1, 2 or 3 with respect to the sum of Trigger 1, 2 or 3 events slightly differs in the real and the corresponding simulated datasets. This needs to be taken into account when one makes use of efficiencies. Therefore, in this analysis an

effective efficiency $\tilde{\epsilon}$ is calculated by forming the weighted sum (indices from above are omitted):

$$\tilde{\epsilon} = \frac{\sum_{t=1}^{3} w_t \epsilon_t}{\sum_{t=1}^{3} w_t}. \tag{4.36}$$

Here, ϵ_t is the trigger dependent efficiency extracted from the fraction of events in Monte Carlo having this specific trigger flag only, and w_t denotes the corresponding re–weighting factor obtained by:

$$w_t = \frac{N_t^{data}/N^{data}}{N_t^{mc}/N^{mc}} \tag{4.37}$$

4.5.6 Compilation of Fit Results

This section presents the outcomes of the extended unbinned likelihood fitting procedures for each individual studied B_s and B^0 decay mode by using the fit functions and the assumptions declared in the last sections. Until time of writing there are no results available on the basis of a cut based selection using a 1.7 fb^{-1} dataset which the presented numbers of signal events could be compared with. The cuts applied on the neural network output to obtain significant signal selections not always were geared towards a maximized significance but rather towards conserving signal events, because in some channels obtaining the highest significance would have required a very hard cut, resulting in a substantial loss of signal events, while the gain in significance was marginal. Therefore, in these cases softer network cuts leading to a sufficient significance level were preferred.

The signal yields obtained by the invariant mass fits are compiled in tables 4.6 and 4.7. The presented number of signal events were obtained by integrating over the broad invariant mass signal range $[5.32, 5.42]$ GeV/c^2 for the B_s and $[5.23, 5.33]$ GeV/c^2 for the B^0 meson decays.

Final state	cut on nnout	N_S	N_B	N_S/N_B	$N_S/\sqrt{N_S+N_B}$
$\phi\pi\ \phi\pi$	0.90	44.64	27.60	1.62	5.25
$\phi\pi\ K^{0*}K$	0.76	44.23	164.23	0.27	3.06
$\phi\pi\ \pi\pi\pi$	0.88	38.87	98.92	0.39	3.31
$K^{0*}K\ K^{0*}K$	0.90	49.35	96.67	0.51	4.08
$K^{0*}K\ \pi\pi\pi$	0.92	31.85	77.77	0.41	3.04
$\pi\pi\pi\ \pi\pi\pi$	0.90	11.01	58.83	0.19	1.32

Table 4.6: Signal yields observed in the six studied $B_s \to D_s D_s$ decay modes. The numbers obtained in the decay modes 4 – 6 (grey) should be regarded as very preliminary results.

Within the ensemble of studied B_s and B^0 decay modes, the decays $B_s \to D_s(\phi\pi)D_s(\phi\pi)$ and $B^0 \to D_s(\phi\pi)D(K\pi\pi)$ are the 'golden' decay modes having the maximum signal significance. This is because the ϕ meson is a very narrow state, thus allowing

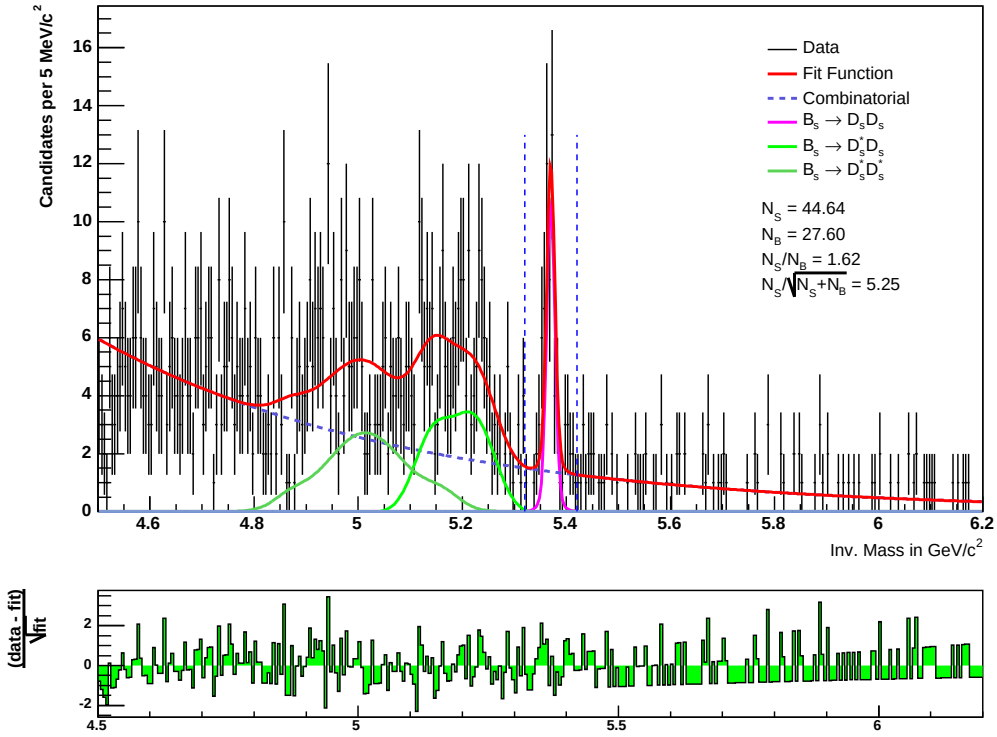


Figure 4.9: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(\phi\pi)\,D_s\,(\phi\pi)$.

Final state	cut on nnout	N_S	N_B	N_S/N_B	$N_S/\sqrt{N_S+N_B}$
$\phi\pi\ K\pi\pi$	0.64	962.65	368.63	2.61	26.58
$K^{0*}K\ K\pi\pi$	0.84	533.02	622.93	0.86	15.68
$\pi\pi\pi\ K\pi\pi$	0.92	583.15	942.74	0.91	16.66

Table 4.7: Signal yields observed in the three studied $B^0 \to D_sD$ decay modes.

for efficiently discriminating signal from combinatorial background. In contrast, signal selection in the decay channel $B_s \to D_s^+\,(\pi^+\pi^-\pi^+)\,D_s^-\,(\pi^+\pi^-\pi^-)$ is very difficult since the combinatorial background is very high in this mode. This is caused by the huge number of possible combinations arising from the six pion tracks in the final state. As discussed in section 4.5.3, the decay modes $B_s \to D_s^+\,(K^{0*}K^+)\,D_s^-\,(K^{0*}K^-)$, $B_s \to D_s^+\,(K^{0*}K^{-+})\,D_s^-\,(\pi^+\pi^-\pi^-)$ and $B_s \to D_s^+\,(\pi^+\pi^-\pi^+)\,D_s^-\,(\pi^+\pi^-\pi^-)$ feature an admixture of several reflections polluting the invariant mass spectra. Hence, the signal yields quoted for these channels can differ from the true number of signal events to an unknown and possibly large extent. This topic will require additional thorough examinations on the basis of comprehensive Monte Carlo studies that have not been conducted yet. Therefore these channels have not been accounted for in the calculation of branching fractions presented in the next chapter.

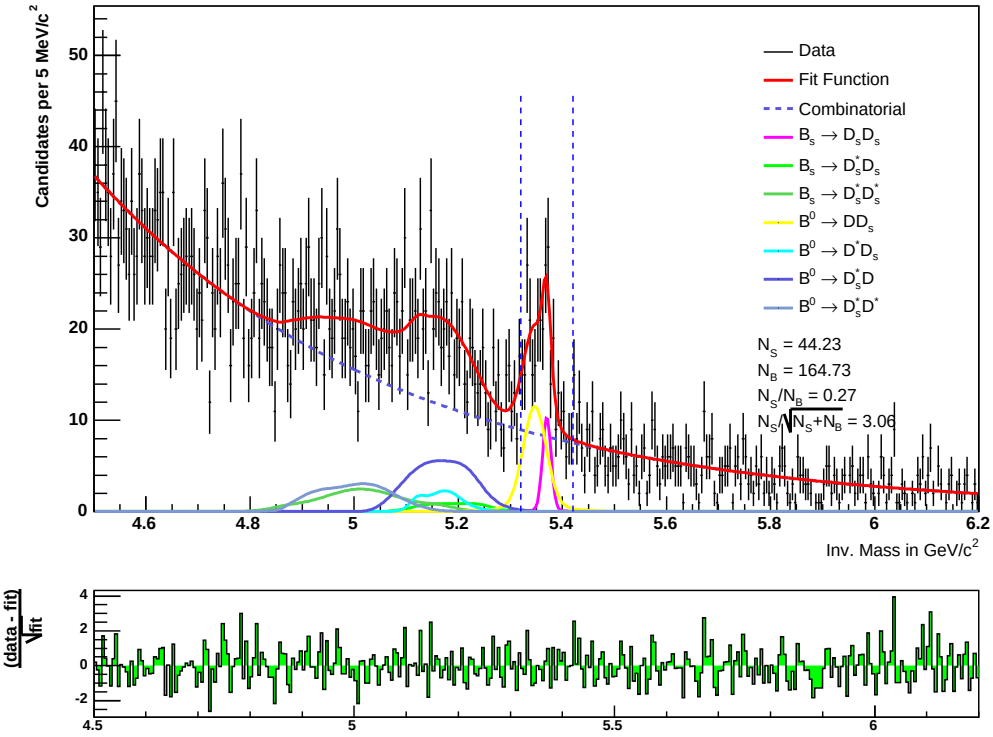


Figure 4.10: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(\phi\pi)\,D_s\,(K^{0*}K)$.

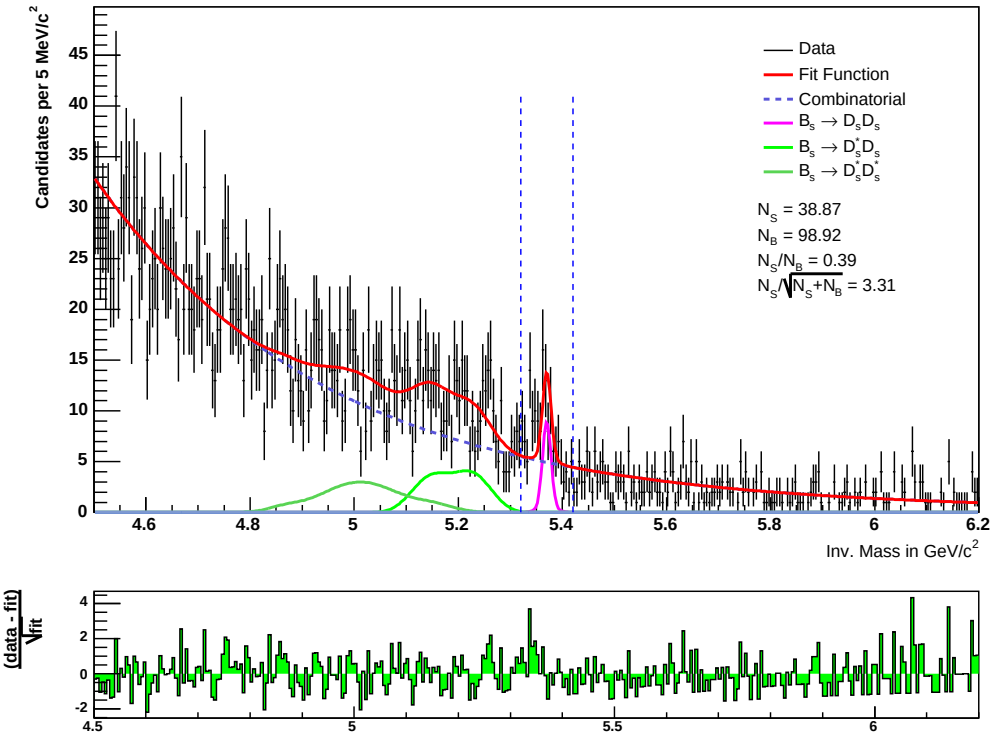


Figure 4.11: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(\phi\pi)\,D_s\,(\pi\pi\pi)$.

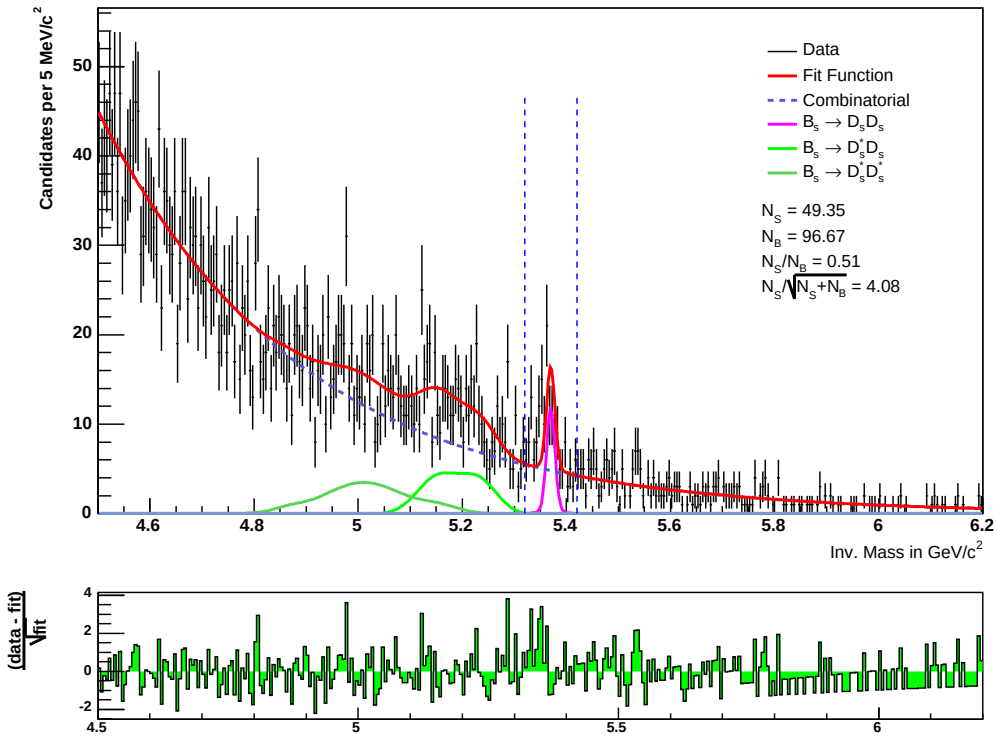


Figure 4.12: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(K^{0*}K)\,D_s\,(K^{0*}K)$.

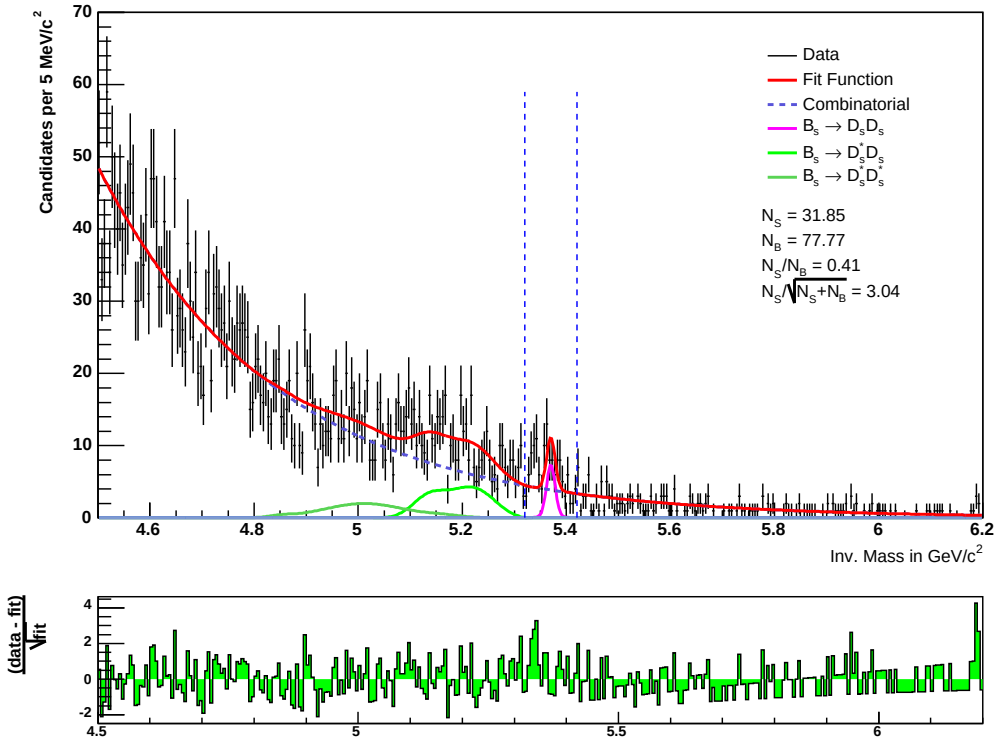


Figure 4.13: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(K^{0*}K)\,D_s\,(\pi\pi\pi)$.

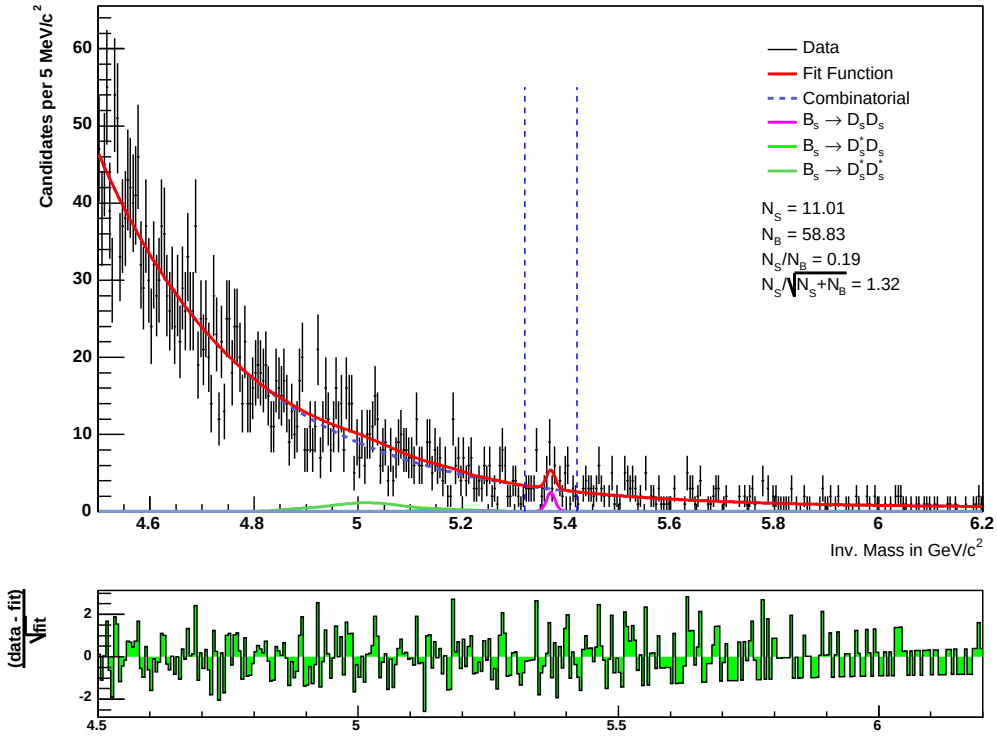


Figure 4.14: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(\pi\pi\pi)\,D_s\,(\pi\pi\pi)$.

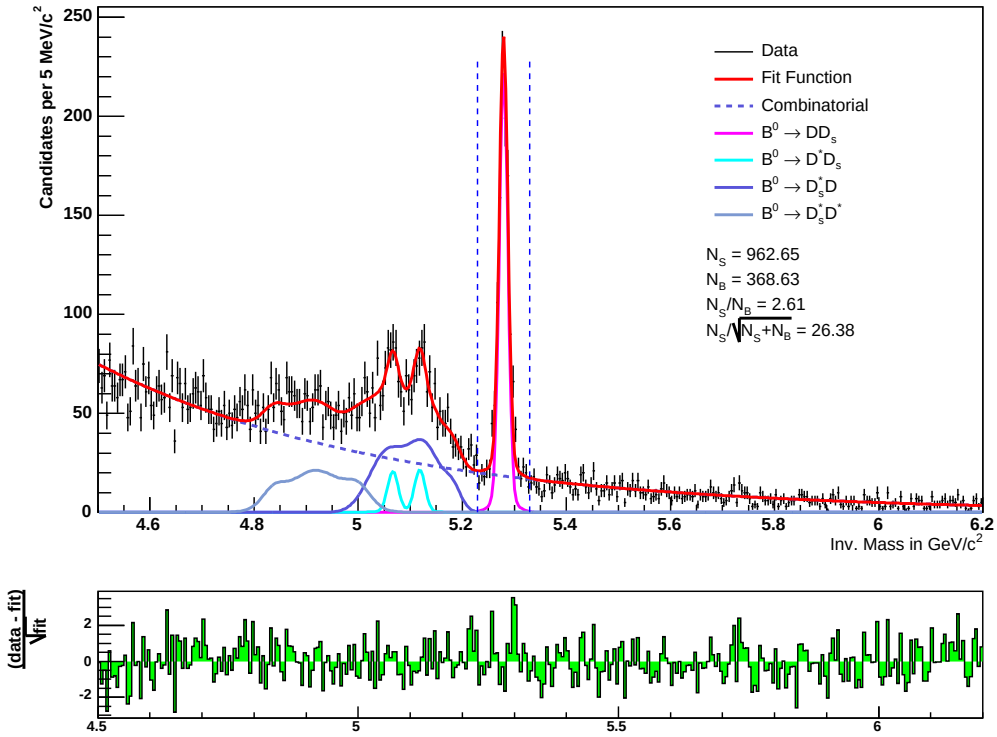


Figure 4.15: Unbinned fit to the invariant mass spectrum of $B^0 \to D_s\,(\phi\pi)\,D\,(K\pi\pi)$.

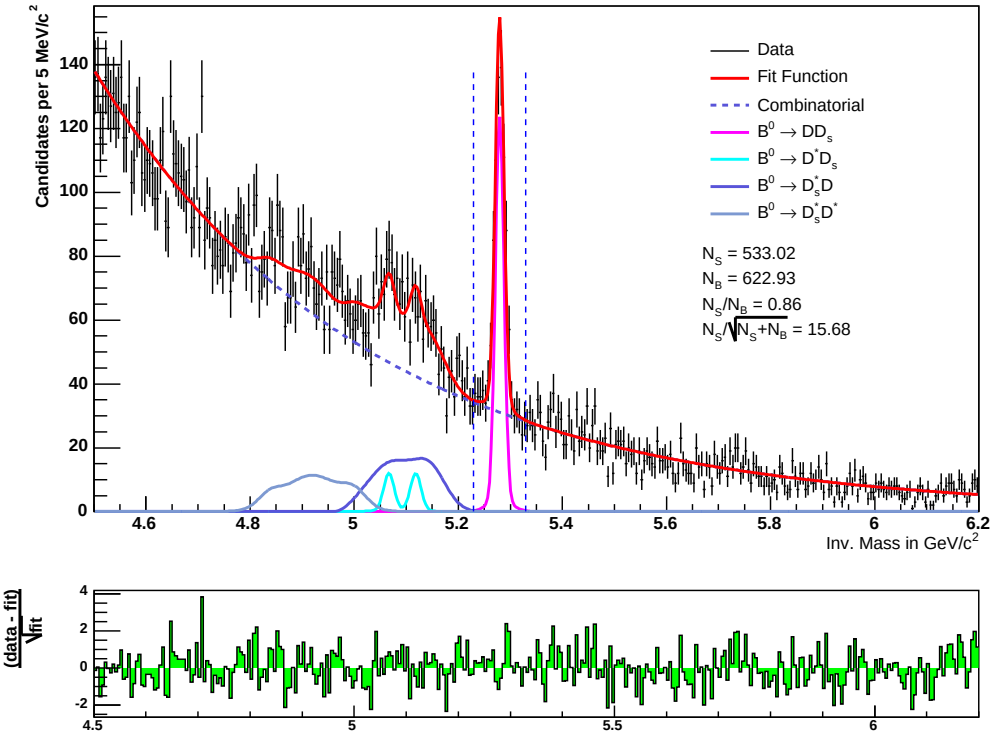


Figure 4.16: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(K^{0*}K)\,D_s\,(\pi\pi\pi)$.

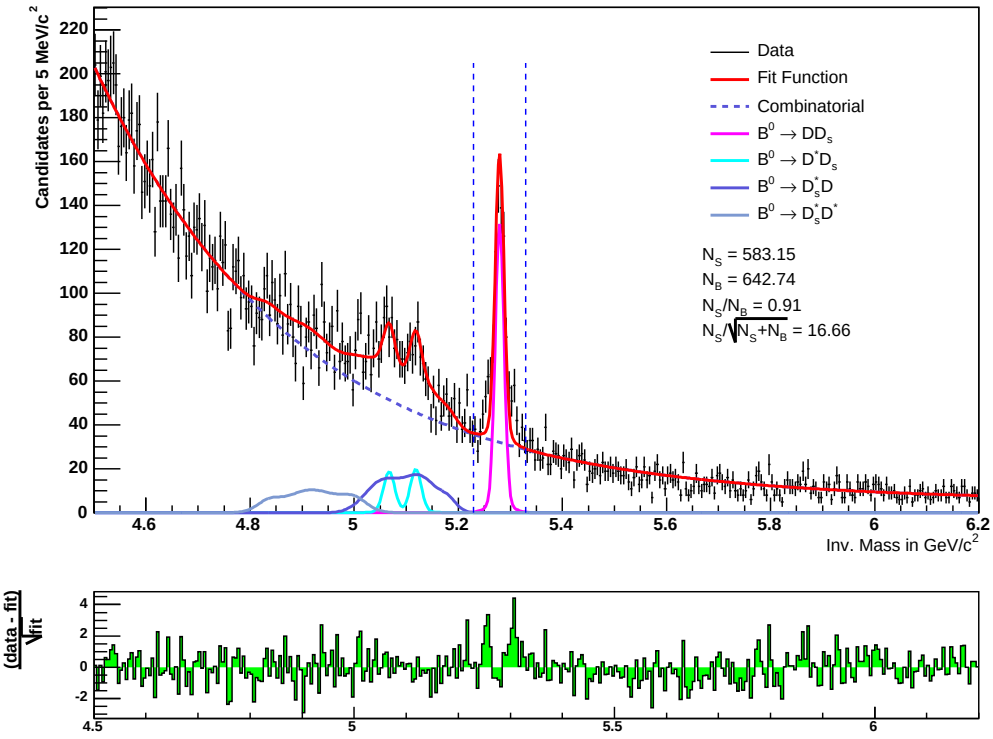


Figure 4.17: Unbinned fit to the invariant mass spectrum of $B^0 \to D_s\,(\pi\pi\pi)\,D\,(K\pi\pi)$.

4.5.7 Checking the Mass Dependence

In principle one cannot exclude the possibility that choosing data from the upper sideband region as the background training sample and Monte Carlo events as the signal sample might cause unwanted effects in data classification when using a neural network. If some of the training variables exhibit a significant dependance on the invariant mass, a neural network can tend to learn to distinguish B mesons not mainly upon the physical properties provided by the input variables, but upon the particular mass range where training events are taken from. In the worst case this could lead to an overestimation of the number of events in the signal region and a disproportionate suppression of events in the invariant mass region where data were taken from as the background training sample.

In order to check if there exists a mass dependance in a neural network classification, one can manipulate a data sample in such a way that it is known to be purely composed of background events. If a network cut using the same neural network that was utilized for the classification of the data sample consisting of background and signal events produces a peak in the signal region, a mass dependance of the neural network is very likely.

This check was performed for the decay channel $B_s \rightarrow D_s(\phi\pi)D_s(\phi\pi)$. To generate a sample purely consisting of background events, the charge pre–cut requirement quoted in table 4.4 was inverted, resulting in exclusively selecting the wrong sign combinations of the particle tracks. This background sample was then classified using the same neural network (see training quality and performance graphs given in section 4.4.2). After applying the same network cut of $nnout > 0.9$, the invariant mass spectrum was fitted using the same fit model discussed in section 4.5.3. The outcome of the unbinned mass fit is shown in figure 4.18.

The fit results in a signal yield of 1 event in the signal region. As an additional cross check the same invariant mass spectrum is re–fitted with a non–signal hypothesis, corresponding to a fit using an exponential as background template function only.

Based on a comparison of the residuals, the 'peak' within the signal region should be regarded as an accidental fluctuation also occurring in some other bins of the histogram. Thus, the neural network is not able to generate a signal peak in a background data sample. Hence, a mass dependent learning effect of the network can be excluded. Since for each studied channel a very similar set of neural network input variables essentially provided the bulk of discriminating power, this check was performed for the channel $B_s \rightarrow D_s\,(\phi\pi)\,D_s\,(\phi\pi)$ only.

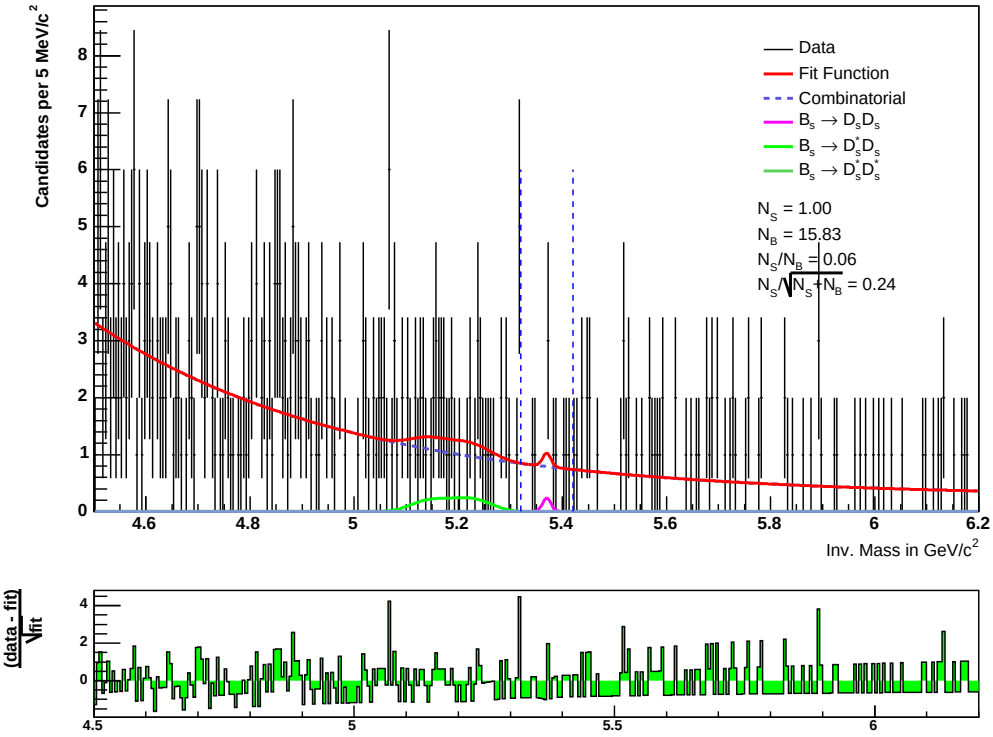


Figure 4.18: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(\phi\pi)\,D_s\,(\phi\pi)$ exclusively containing background events.

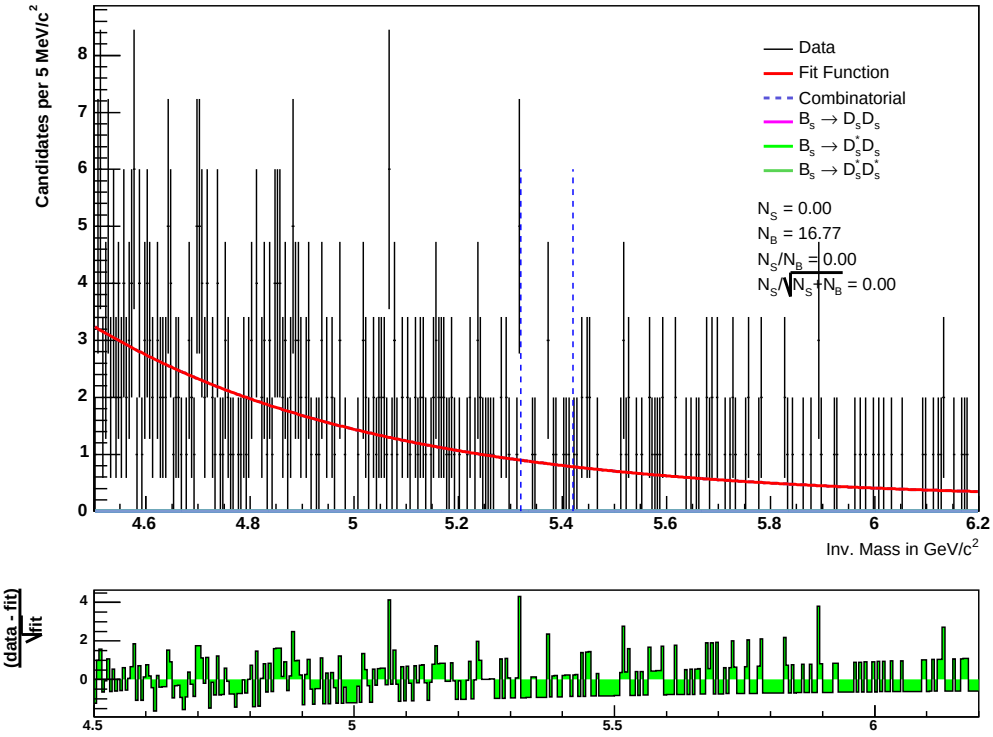


Figure 4.19: Unbinned fit to the invariant mass spectrum of $B_s \to D_s\,(\phi\pi)\,D_s\,(\phi\pi)$ exclusively containing background events. For the fit a non–signal hypothesis was assumed.

Chapter 5

Extraction of Branching Fractions

Beyond studying the branching fraction $Br\,[B_s \to D_s^+ D_s^-]$, as a byproduct of the fitting procedure the presented analysis provides the opportunity to determine additional quantities, namely the relative branching fractions $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$, $f_{D_* D_s^*}$, $f_{D_s^* D}$ and $f_{D_s^* D^*}$. Therefore, prior to presenting the analysis of $Br\,[B_s \to D_s^+ D_s^-]$ as the central element of this analysis, the extraction of these relative branching fractions is outlined first.

5.1 Measurement of $Br[B_s \to D_s^{+(*)} D_s^{-(*)}]$ and $Br[B^0 \to D^{+(*)} D_s^{-(*)}]$

5.1.1 Approach

The $B_s \to D_s^+ D_s^-$ modes 4 – 6 are henceforth not incorporated in the measurement of branching fractions since the fit model currently applied does not account for the non–trivial mixture of reflections in these channels (see Section 4.5.3 for details). The relations between the number of signal events in $B_s \to D_s^{+(*)} D_s^{-(*)}$ and $B^0 \to D^{+(*)} D_s^{-(*)}$ decays and the relative branching fractions already have been quoted in equations (4.23), (4.24) and (4.32) – (4.34). Evaluating these formula by inserting the numbers of signal events obtained in three out the six studied decay modes of $B_s \to D_s^+ D_s^-$ and $B^0 \to D^+ D_s^-$ respectively would lead to three separate values for each relative branching fraction that need to be combined for the calculation of a final result.

In this analysis however, a more holistic approach exploiting a certain feature of the fitter framework that allows to pass several `ModeClass` objects to the `Fitter` is chosen. By doing so, a synchronous fit to the mass spectra of several decay channels can be performed. Merging the `ModeClass` objects results in summing up the individual negative log likelihood functions $F^i(\vec{a}_i)$ and $F^j(\vec{b}_j)$, $i, j = 1, 2, 3$, specified for the

different $B_s \to D_s^+ D_s^-$ and $B^0 \to D^+ D_s^-$ decay modes, giving rise to

$$F(\vec{A}) = \sum_{i=1}^{3} F^i(\vec{a}_i) \tag{5.1}$$

for the combination of the incorporated B_s modes, and

$$F(\vec{B}) = \sum_{j=1}^{3} F^j(\vec{b}_j) \tag{5.2}$$

for the combined B^0 modes. The resulting parameter vectors are the sum of the parameter vectors belonging to each single log likelihood function, $\vec{A} = \sum_{i=1}^{3} \vec{a}_i$ and $\vec{B} = \sum_{i=1}^{3} \vec{b}_j$. Within the fitter framework, fit parameters are identified by their denominations, hence the elements of the parameter vectors $\vec{A}$ and $\vec{B}$ are discriminable by their individual naming. As the relative branching fractions from a physical point of view have to be independent of a particular $B_s \to D_s^+ D_s^-$ and $B^0 \to D^+ D_s^-$ decay mode, the fit parameters $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$ and $f_{D^* D_s^*}$, $f_{D_s^* D}$, $f_{D_s^* D^*}$ can be shared among the dedicated combination of B_s and B^0 `ModeClass` objects entering the fitter. Combining several decay modes into one fitting process is referred to as a *simultaneous fit*.

Since the amount of fitting parameters being managed by the fitter rises nearly linearly with the number of mode objects, their starting values need to be constrained within some reasonable limits, otherwise the fit might not converge due to having too much freedom in the multi–dimensional parameter space. In addition, reasonable starting values have to be set. This is achieved by using the parameter values obtained in the fits to the individual decay modes (see section 4.5) as starting values. The shape parameters that were determined by means of binned fits to Monte Carlo samples in the first stage of the fitting process are still kept fixed. Figure 5.1 shows a diagram illustrating the functional principle of the simultaneous fit for the parameter estimation for the relative branching fractions.

5.1.2 Results for $B^0 \to D^{+(*)} D_s^{-(*)}$

The outcomes of the simultaneous fit to the three $B^0 \to D^{+(*)} D_s^{-(*)}$ decay modes are shown in figures D.4 – D.6, Appendix D. The fit yields the following values for $f_{D^* D_s^*}$, $f_{D_s^* D}$ and $f_{D_s^* D^*}$:

$$\frac{Br\,(B^0 \to D_s^+ D^{-*})}{Br\,(B^0 \to D_s^+ D^-)} = 0.91 \pm 0.14 \tag{5.3}$$

$$\frac{Br\,(B^0 \to D_s^{+*} D^-)}{Br\,(B^0 \to D_s^+ D^-)} = 1.12 \pm 0.07 \tag{5.4}$$

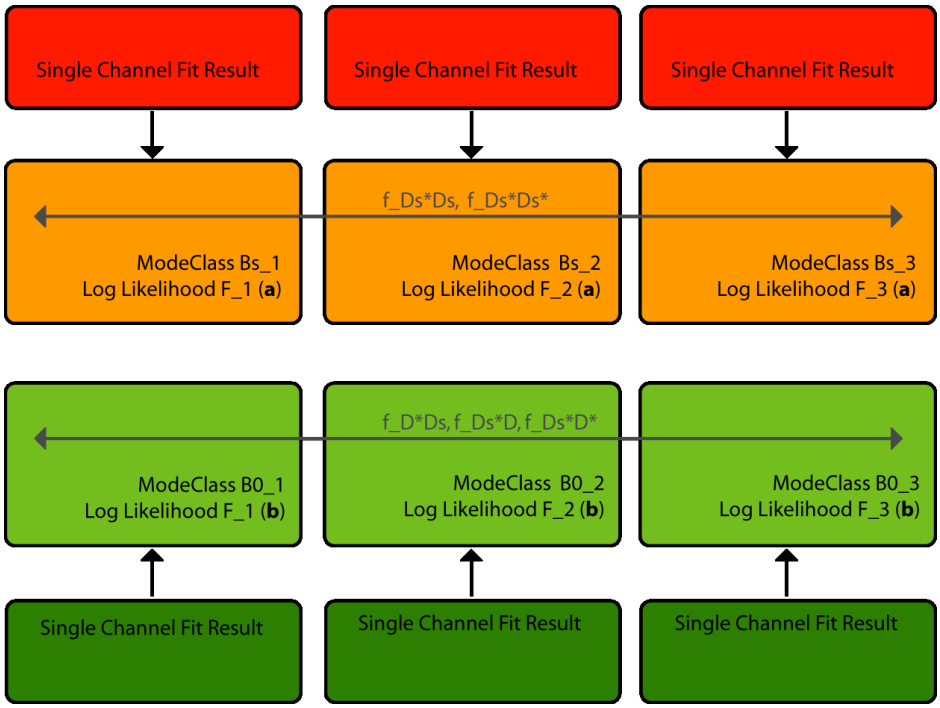


Figure 5.1: Scheme of the simultaneous fit method for the parameter estimation of relative branching fractions in $B_s \to D_s^{+(*)} D_s^{-(*)}$ and $B^0 \to D^{+(*)} D_s^{-(*)}$ decays. The results of the single channel fits are used to set parameter start values. The relative branching fractions $f_{D_s^* D}$, $f_{D_s^* D^*}$ enter as fit parameters shared among the B_s modes, $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$ and $f_{D^* D_s^*}$ are shared among the B^0 modes.

$$\frac{Br\,(B^0 \to D_s^{+*} D^{-*})}{Br\,(B^0 \to D_s^{+} D^{-})} = 2.94 \pm 0.20 \tag{5.5}$$

Since a thorough study of systematic uncertainties has not been conducted yet, the quoted uncertainties incorporate statistical uncertainties and the uncertainties stemming from the efficiencies only. There are no errors due to branching fractions since the mode specific branching fractions $Br[D_s \to \phi\pi, \phi \to KK]$, $Br[D_s \to K^*K, K^* \to K\pi]$, $Br[D_s \to \pi\pi\pi]$ and $Br[D \to K\pi\pi]$ cancel out.

Comparison to PDG

Based upon the absolute branching fractions of $B^0 \to D^{+(*)} D_s^{-(*)}$ decays given in the PDG 2007 (partial update for 2008) [4], the following relative branching fractions are obtained:

$$\frac{Br\,(B^0 \to D_s^{+} D^{-*})}{Br\,(B^0 \to D_s^{+} D^{-})} = 1.23 \pm 0.30 \tag{5.6}$$

$$\frac{Br\,(B^0 \to D_s^{+*} D^{-})}{Br\,(B^0 \to D_s^{+} D^{-})} = 1.14 \pm 0.34 \tag{5.7}$$

$$\frac{Br\,(B^0 \to D_s^{+*} D^{-*})}{Br\,(B^0 \to D_s^{+} D^{-})} = 2.72 \pm 0.58 \tag{5.8}$$

Within the uncertainties – which are underestimated in this analysis because systematic effects were not accounted for – this measurement agrees well with the PDG figures.

5.1.3 Results for $B_s \to D_s^{+(*)} D_s^{-(*)}$

The following numbers for $f_{D_s^* D_s}$ and $f_{D_s^* D_s^*}$ are obtained by performing a simultaneous fit to three out of the six studied B_s decay channels (see figures D.1 – D.3)

$$\frac{Br\,(B_s \to D_s^{*} D_s)}{Br\,(B_s \to D_s D_s)} = 2.87 \pm 0.52 \tag{5.9}$$

$$\frac{Br\,(B_s \to D_s^{*} D_s^{*})}{Br\,(B_s \to D_s D_s)} = 3.44 \pm 0.71 \tag{5.10}$$

Until time of writing there are no published results available which the quoted results could be compared with. As in the case of $B^0 \to D^{+(*)} D_s^{-(*)}$ no study of systematic uncertainties has been performed in the course of this analysis.

5.2 Measurement of $Br[B_s \to D_s^+ D_s^-]$

5.2.1 Method

Before discussing the computational techniques used for measuring $Br[B_s \to D_s^+ D_s^-]$, the underlying formal context is briefly outlined.

Given the number of signal events $N^i_{B_s}$ of a $B_s \to D_s^+ D_s^- \to \Phi^i$ decay, where Φ^i is a particular final state, the branching fraction $Br[B_s \to D_s^+ D_s^-]$ is defined via

$$N^i_{B_s} = N^{tot}_{B_s} Br[B_s \to D_s^+ D_s^-] Br[D_s^+ D_s^- \to \Phi^i] \epsilon^i, \tag{5.11}$$

where $N^{tot}_{B_s}$ is the total number of produced B_s mesons at the Tevatron, $Br[D_s^+ D_s^- \to \Phi^i]$ denotes the branching fraction of the final state Φ^i and ϵ^i is the combined reconstruction and selection efficiency extracted from Monte Carlo. However, to suppress systematics the branching fraction $Br[B_s \to D_s^+ D_s^-]$ is measured via the ratio of $Br[B_s \to D_s^+ D_s^-]$ to the branching fraction of another decay mode. In this analysis $B^0 \to D^+ D_s^-$ is chosen as the normalization mode, where $Br[B^0 \to D_s^+ D^-]$ is defined analogously to (5.11):

$$N^j_{B^0} = N^{tot}_{B^0} Br[B^0 \to D_s^+ D^-] Br[D_s^+ D^- \to \Phi^j] \epsilon^j. \tag{5.12}$$

Here again $N^{tot}_{B^0}$ is the total amount of generated B^0 at the Tevatron and ϵ^j the combined reconstruction and selection efficiency for the decay mode j. From (5.11) and (5.12) one derives that

$$N^i_{B_s} = f^{ij}_{B_s} N^j_{B^0} \frac{Br[D_s^+ D_s^- \to \Phi^i]}{Br[D_s^+ D^- \to \Phi^j]} \epsilon^{ij}, \tag{5.13}$$

where $\epsilon^{ij} = \epsilon^i / \epsilon^j$. f_{B_s} is defined by

$$f_{B_s} = \frac{f_s}{f_d} \frac{Br[B_s \to D_s^+ D_s^-]}{Br[B^0 \to D_s^+ D^-]}. \tag{5.14}$$

f_s/f_d is the production ratio of s and d quarks, having the PDG [4] value $f_s/f_d = 0.259 \pm 0.038$. Hence, evaluating equations (5.13) and (5.14) by inserting the number of signal events $N^i_{B_s}$ and $N^j_{B^0}$ allows to calculate $Br[B_s \to D_s^+ D_s^-]$. For this purpose, in addition the branching fraction of $B^0 \to D_s^+ D^-$ has to be known. In the PDG the following value is quoted:

$$Br[B^0 \to D_s^+ D^-] = 0.0065 \pm 0.0013 \tag{5.15}$$

The choice which of the studied decay modes to use for the calculation of f_{B_s} in principle is completely arbitrary. However, one prefers to combine certain $B_s \to D_s^+ D_s^-$ and $B^0 \to D_s^+ D^-$ decay modes in a such a way that systematic effects stemming from

the sizable uncertainties of the involved final state branching fractions are minimized. The final state branching fractions are taken from the PDG:

$$Br[D_s^+ \to \phi\pi^+, \phi \to K^+K^-] = 0.022 \pm 0.0020 \quad (5.16)$$
$$Br[D_s^+ \to K^{0*}K^+, K^{0*} \to K^+\pi^-] = 0.025 \pm 0.005 \quad (5.17)$$
$$Br[D_s^+ \to \pi^+\pi^-\pi^+] = 0.0124 \pm 0.0020 \quad (5.18)$$
$$Br[D^+ \to K^+\pi^-\pi^+] = 0.0951 \pm 0.0034 \quad (5.19)$$

By using only a particular B_s decay mode i and a certain B^0 decay mode j, one of the final state branching fractions cancels out. Using the first three out of the six studied $B_s \to D_s^+ D_s^-$ modes and the three $B^0 \to D^+ D_s^-$ modes, these combinations are in detail:

- $$N_{B_s}^1 = f_{B_s}^{11} N_{B^0}^1 \frac{Br[D_s \to \phi\pi]Br[D_s \to \phi\pi]}{Br[D \to K\pi\pi]Br[D_s \to \phi\pi]}\epsilon^{11} = f_{B_s}^{11} N_{B^0}^1 \frac{Br[D_s \to \phi\pi]}{Br[D \to K\pi\pi]}\epsilon^{11}$$

- $$N_{B_s}^2 = f_{B_s}^{22} N_{B^0}^2 \frac{Br[D_s \to \phi\pi]Br[D_s \to K^*K]}{Br[D \to K\pi\pi]Br[D_s \to K^*K]}\epsilon^{22} = f_{B_s}^{22} N_{B^0}^2 \frac{Br[D_s \to \phi\pi]}{Br[D \to K\pi\pi]}\epsilon^{22}$$

- $$N_{B_s}^3 = f_{B_s}^{33} N_{B^0}^3 \frac{Br[D_s \to \phi\pi]Br[D_s \to \pi\pi\pi]}{Br[D \to K\pi\pi]Br[D_s \to \pi\pi\pi]}\epsilon^{33} = f_{B_s}^{33} N_{B^0}^3 \frac{Br[D_s \to \phi\pi]}{Br[D \to K\pi\pi]}\epsilon^{33}$$

- $$N_{B_s}^2 = f_{B_s}^{21} N_{B^0}^1 \frac{Br[D_s \to \phi\pi]Br[D_s \to K^*K]}{Br[D \to K\pi\pi]Br[D_s \to \phi\pi]}\epsilon^{21} = f_{B_s}^{21} N_{B^0}^1 \frac{Br[D_s \to K^*K]}{Br[D \to K\pi\pi]}\epsilon^{21}$$

- $$N_{B_s}^3 = f_{B_s}^{31} N_{B^0}^1 \frac{Br[D_s \to \phi\pi]Br[D_s \to \pi\pi\pi]}{Br[D \to K\pi\pi]Br[D_s \to \phi\pi]}\epsilon^{31} = f_{B_s}^{31} N_{B^0}^1 \frac{Br[D_s \to \pi\pi\pi]}{Br[D \to K\pi\pi]}\epsilon^{31}$$

Out of these combinations, the smallest systematic uncertainty originating from the branching fraction term is introduced by the first three formula. Considering the different relative branching fraction uncertainties quoted in 5.16, the last two combinations are not preferable and not considered in the following steps.

Simultaneous Fitting Procedure

Following the principle outlined in section 5.1, a simultaneous fit is conducted by merging one fitter `ModeClass` object B_s^i with one $B^{0,j}$ object according to the advantageous combinations where one of the two branching fractions included in $Br[D_s^+ D_{(s)}^- \to \Phi^{(i)j}]$ (equation (5.13)) is eliminated. This necessitates a slight modification of the log likelihood function $F^i(\vec{a}_i)$: In order to use f_{Bs} as a direct fit parameter, the number of signal events of a certain B_s decay channel i, N_1^i, is replaced by the identity specified in (5.13). By doing so, the number of $B^{0,j}$ signal events $N_{B^0}^j$ is shared among each

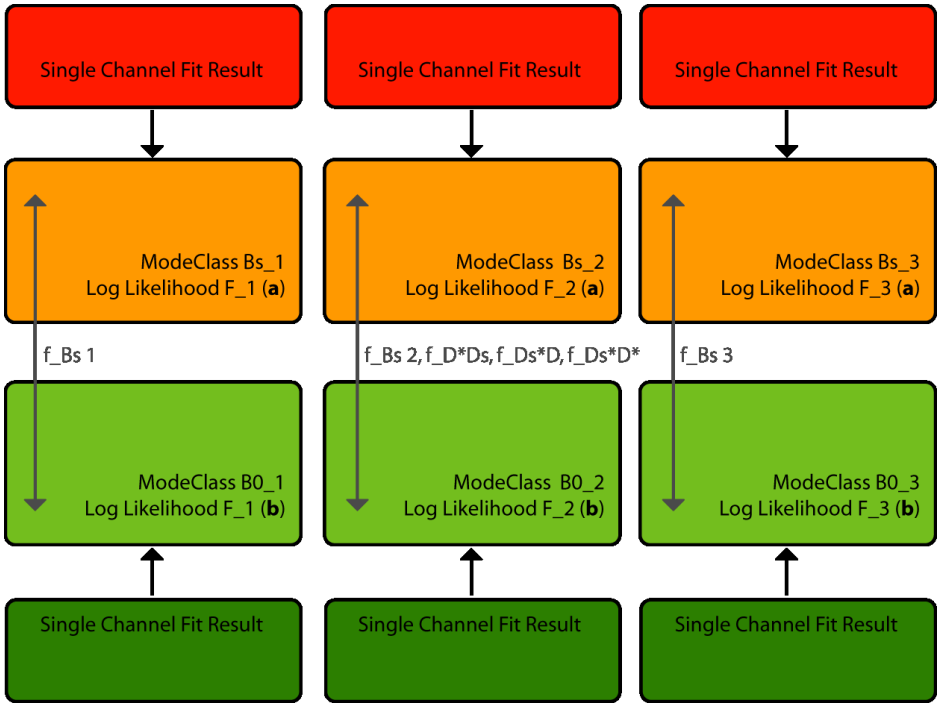


Figure 5.2: Scheme of the simultaneous fit method for the parameter estimation of f_{B_s}. The results of the single channel fits are used for setting parameter start values.

advantageous combination of one B_s and one B^0 mode class object. In addition, for B_s mode 2 the parameters $f_{D^* D_s^*}$, $f_{D_s^* D}$ are shared with the respective $B^{0,j}$ mode. The method is illustrated in figure 5.2.

The uncertainties $\boldsymbol{\sigma}_{br}$ of the branching fractions involved are passed to the fitter by adding the Gaussian constraints

$$\frac{(\mathbf{p}_{br} - \bar{\mathbf{p}}_{br})^2}{\boldsymbol{\sigma}^2_{p_{br}}} \tag{5.20}$$

to the log likelihood function. Here $\mathbf{p}_{br}$ are the branching fraction parameters. By doing so these parameters are varied according to their systematic uncertainties while being strongly constrained to their mean value.

Individual Fit Results

The invariant mass plots resulting from the three simultaneous fits are shown in figures D.7(a) – D.9(b), Appendix D. The three measurements of f_{B_s} yield the following numbers:

$$f_{B_s}^{11} = 0.46 \pm 0.08(stat) \pm 0.04(Br) = 0.46 \pm 0.09 \tag{5.21}$$
$$f_{B_s}^{22} = 0.37 \pm 0.11(stat) \pm 0.10(Br) = 0.37 \pm 0.15 \tag{5.22}$$
$$f_{B_s}^{33} = 0.32 \pm 0.07(stat) \pm 0.10(Br) = 0.32 \pm 0.12 \tag{5.23}$$

The individual uncertainties were obtained by repeating the simultaneous fits with the branching fraction uncertainties set to zero. In the next section, the values are combined to one single result for f_{B_s}, finally allowing for calculating $Br[B_s \to D_s^+ D_s^-]$.

5.2.2 Combined Branching Fraction Result

The results can be combined rather easily since $Br[D_s \to \phi\pi]/Br[D \to K\pi\pi]$ entering relation (5.13) is a common factor for all three combinations and is thus 100% correlated. This correlation has to be taken into account when calculating a combined result for f_{B_s}.

Averaging $f_{B_s}^{11}$, $f_{B_s}^{22}$ and $f_{B_s}^{33}$ yields:

$$f_{B_s} = \frac{f_s}{f_d}\frac{Br[B_s \to D_s^+ D_s^-]}{Br[B^0 \to D_s^+ D^-]} = 0.42 \pm 0.05(Stat) \pm 0.03(Br) \pm 0.05(Br_{Corr}) = 0.42 \pm 0.08$$

Br_{Corr} denotes the uncertainty due to the correlation of branching fractions. The given value compares to $f_{B_s} = 0.37^{+0.10}_{-0.08}(Stat)^{+0.03}_{-0.04}(Syst) \pm 0.01(Br) \pm 0.05(Br_{Corr})$ measured in the former CDF analysis [3] where data corresponding to 355 pb^{-1} were used. The effect of the larger amount of statistics available for the given measurement is visible in the statistical uncertainty. Using (5.14) one calculates:

$$\begin{aligned}\frac{Br[B_s \to D_s^+ D_s^-]}{Br[B^0 \to D_s^+ D^-]} &= 1.62 \pm 0.19(stat) \pm 0.12(Br) \pm 0.19(Br_{Corr}) \pm 0.24(f_s/f_d) \\ &= 1.62 \pm 0.38.\end{aligned}$$

Finally, inserting the PDG value for $Br[B^0 \to D_s^+ D^-]$, the result for the branching fraction of $B_s \to D_s^+ D_s^-$ reads:

$$\begin{aligned}\frac{\Gamma(D_s^+ D_s^-)}{\Gamma_{total}} = Br[B_s \to D_s^+ D_s^-] &= \{10.5 \pm 1.2(Stat) \pm 0.8(Br) \pm 1.2(Br_{Corr}) \\ &\quad \pm 1.6(f_s/f_d) \pm 2.1(Br_{B^0})\} \times 10^{-3} \\ &= (10.5 \pm 3.2) \times 10^{-3}\end{aligned}$$

Albeit slightly higher, within the uncertainties this result is in good agreement with the measured branching fraction $Br[B_s \to D_s^+ D_s^-] = (9.4^{+4.4}_{-4.2}) \times 10^{-3}$ obtained in the previous CDF measurement.

Exploiting (3.25), the measured branching fraction can be used to estimate a lower bound for the CP width difference $\Delta\Gamma_{CP}/\Gamma$:

$$\frac{\Delta\Gamma_{CP}}{\Gamma} \geq 0.021 \pm 0.006$$

Assuming the Standard Model expectation $\Delta\Gamma_{CP}/\Gamma = \Delta\Gamma/\Gamma$, this lower bound is in agreement with the theoretical expectation $\Delta\Gamma/\Gamma = 0.127 \pm 0.024$.

5.2.3 Alternative Approach

As a cross check to the results presented so far an additional possibility that allows the determination of the relative branching fraction f_{B_s} by means of a single measurement is presented. Before outlining the slightly different ansatz, it is stressed that the method needs further refinement since not all aspects were studied sufficiently. These will be pointed out in the course of this section.

The approach is based on the same formal context initially set up in section 5.2.1. There, equation (5.13),

$$N_{B_s}^i = f_{B_s}^{ij} N_{B^0}^j \frac{Br[D_s^+ D_s^- \to \Phi^i]}{Br[D_s^+ D^- \to \Phi^j]} \epsilon^{ij},$$

served as the starting point for the measurement of f_{B_s} by using it as a direct fitting parameter in a simultaneous fit. However, for obtaining one final figure for f_{B_s} the different results from several simultaneous fits using different combinations of `ModeClass` objects (or decay modes) need to be combined. Therefore, performing one single simultaneous fit using all six modes involved at the same time would be more convenient. Using (5.13) this intention fails due the fact that the mode specific parameter $N_{B^0}^j$ inherently can not be shared among the B^0 modes. This restriction can be resolved easily by replacing the parameter N_1^j in each log likelihood function $F^j(\vec{b}_j)$ by the identity (5.12), resulting in

$$N_{B^0}^j = N_{B^0}^{tot} Br[B^0 \to D_s^+ D^-] Br[D_s^+ D^- \to \Phi^j] \epsilon^j.$$

As a next step, this is inserted into equation (5.13), giving rise to

$$N_{B_s}^i = f_{B_s} N_{B^0}^{tot} Br[B^0 \to D_s^+ D^-] Br[D_s^+ D^- \to \Phi^j] \epsilon^i. \tag{5.24}$$

Instead of (5.12) this identity now is used to express the parameter N_1^i in $F^i(\vec{a}_i)$. By using the modified fit functions $\tilde{F}^i(\vec{a}_i)$ and $\tilde{F}^j(\vec{b}_j)$ as an ansatz for the fitting process, one is allowed to simultaneously pass all three relevant B_s and B^0 modes into one fitter, whereas the parameter f_{B_s} is shared among the B_s modes and $N_{B^0}^{tot}$ among all six mode objects. In addition, the parameters $f_{D_s^* D_s}$, $f_{D_s^* D_s^*}$, $f_{D^* D_s^*}$, $f_{D_s^* D}$ and $f_{D_s^* D^*}$ again are shared among their respective mode class objects in order to test the validity of this method. For all the parameters the results from the combined B_s and B^0 fits respectively were used as start values and again some reasonable limits were set. The principle of this global simultaneous fit method is shown in a graphical way in figure 5.3.

As mentioned at the beginning of this section, there are however some issues that need to be considered. First of all, the benefit one gains from more statistics in the global simultaneous fit is damped by the fact that in addition to the advantageous combinations of the decay modes (see section 5.2.1) also disadvantageous combinations are added, leading to a larger systematic error due to branching fraction uncertainties. In addition, the pool of parameters might hold correlations that need to be accounted

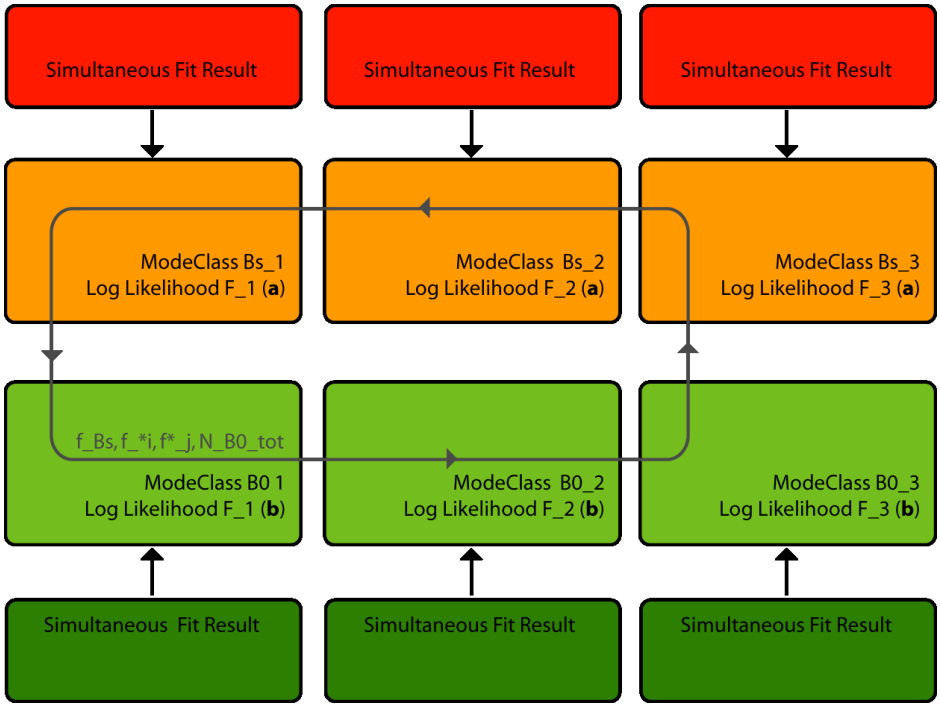


Figure 5.3: Scheme of the alternative simultaneous fit method for parameter estimation of f_{B_s}. Three B_s and three B^0 modes are simultaneously combined and fitted. The results of the respective B_s and B^0 simultaneous fits are used for setting start values of the parameters which are allowed to float within reasonable limits.

for. Ignoring these correlations could introduce a statistical bias of the final result for f_{B_s}. Therefore, instead of simple Gaussian constraints (equation (5.20)) the term

$$(\mathbf{p} - \bar{\mathbf{p}})^T \mathbf{V} (\mathbf{p} - \bar{\mathbf{p}}), \tag{5.25}$$

with the correlation matrix $\mathbf{V}$ and the vectors of constrained parameters $\mathbf{p}$, needs to be added to the overall log likelihood function. This has not been implemented yet, therefore the result presented in the next section should be seen as a cross check on the measurement of f_{B_s} only.

Global Simultaneous Fit Result

The six invariant mass spectra resulting from the simultaneous fit are shown in figures D.10(a) – D.12(b). The fit results for the relative branching fractions

$$\frac{Br\,(B^0 \to D_s D^*)}{Br\,(B^0 \to D_s D)} = 0.91 \pm 0.14 \tag{5.26}$$

$$\frac{Br\,(B^0 \to D_s^* D)}{Br\,(B^0 \to D_s D)} = 1.17 \pm 0.07 \tag{5.27}$$

$$\frac{Br\,(B^0 \to D_s^* D^*)}{Br\,(B^0 \to D_s D)} = 2.94 \pm 0.19 \tag{5.28}$$

$$\frac{Br\,(B_s \to D_s^* D_s)}{Br\,(B_s \to D_s D_s)} = 2.81 \pm 0.48 \tag{5.29}$$

$$\frac{Br\,(B_s \to D_s^* D_s^*)}{Br\,(B_s \to D_s D_s)} = 3.53 \pm 0.67 \tag{5.30}$$

show a good agreement with the figures from equations (5.3) – (5.5), (5.9) and (5.10). Considering the fact that $f_{D^* D_s^*}$, $f_{D_s^* D}$ and $f_{D_s^* D^*}$ are not only shared among all the B^0 modes but in addition are shared with B_s decay mode 2 to account for the reflection from B^0 decay mode 1, the consistency of the quoted numbers can be seen as an indicator for the general fit model to be valid.

Furthermore, the measurement of f_{B_s} resulting from this alternative approach yields:

$$f_{B_s} = 0.42 \pm 0.08. \tag{5.31}$$

This result exactly reproduces the combined result calculated in section 5.2.2. This is, however, no big surprise since essentially the same input information is exploited

in a slightly differing manner. Looking at the combined uncertainty, it seems that the benefit from more statstics is compensated by the larger systematic uncertainties stemming from the disadvantageous combinations where none of the D_s branching fractions cancels out.

In spite of the concerns that have been outlined before, the global simultaneous fit method for the simultaneous determination of branching fractions could be an interesting alternative to the method presented in section 5.2.2 if some improvements to the fit model are introduced that account for delicate effects which were not accounted for so far, like correlations between input parameters for instance.

Chapter 6

Conclusion and Outlook

Heavy flavor physics, particular when it comes to mesons containing bottom quarks, currently is a notable arena of efforts in the field of particle physics. Studies of the phenomenology of B meson decay and mixing give important insights into the mechanisms of weak interaction in the Standard Model, and in addition provide the opportunity to search for new physics beyond the current commonly accepted theoretical framework. Since CP violation is predicted to vanish in the B_s–$\bar{B}_s$ system, evidence for a sizeable CP–violating phase in B_s–$\bar{B}_s$ mixing would clearly indicate new physics beyond the Standard Model.

Knowledge of the branching fraction of the decay $B_s \to D_s^+ D_s^-$ provides an important input to make inferences to this topic. First of all, this decay is believed to place the main contribution to the decay width difference in the B_s system which is theoretically predicted to be sizable. Moreover, the final state $D_s^+ D_s^-$ is a purely even eigenstate of CP, allowing to directly relate the branching fraction to the CP width difference in the B_s system. Measuring the CP width difference alone might not deliver evidence for a sizable CP violating phase in B_s–$\bar{B}_s$, however it can be regarded as an important parameter to narrow down possible scenarios.

The goal of this thesis was to provide a new measurement of the branching fraction $Br[B_s \to D_s^+ D_s^-]$, based on data corresponding to 1.7 fb^{-1} provided by the CDF–II experiment at Fermilab. The first step comprised the optimization of the selection of exclusively reconstructed B_s signal events in $B_s \to D_s^+ D_s^-$ decay modes with the D_s meson decaying into $\phi\pi$, $K^{*0}K$ and $\pi\pi\pi$. For signal selection artificial neural networks were applied which provide a powerful tool for this task since the information contained in the input variables are mapped onto a single discriminating variable. In doing so, additionally correlations between variables are taken into account. Observed signal yields were obtained by performing fits to the invariant mass spectra of each studied channel by means of the extended unbinned maximum likelihood method. For all fitting procedures the fitter framework was utilized throughout this analysis. The fitter framework is a special software environment providing a flexible structure suitable for different types of fit methods.

The branching fraction was not measured directly but via the ratio $Br[B_s \to$

$D_s^+D_s^-]/Br[B^0 \to D_s^+D^-]$. For this, in addition three hadronic decay modes of the decay $B^0 \to D_s^+D^-$ have been analyzed. Exploiting a certain feature of the fitter framework that allows to combine several decay modes within one fitting process, which is then called simultaneous fit, two congeneric methods for determining $Br[B_s \to D_s^+D_s^-]/Br[B^0 \to D_s^+D^-]$ were developed and applied. At this stage of the analysis not all studied B_s modes were accounted for: The three decay modes where there is no $\phi\pi$ in the final state were omitted since the applied fit model failed in describing data sufficiently. This is because these modes comprise a mixture of different reflections originating from the misreconstruction of similar meson decays.

By performing a simultaneous fit for each of the combinations $[B_s \to D_s(\phi\pi)D_s(\phi\pi)$, $B^0 \to D_s(\phi\pi)D(K\pi\pi)]$, $[B_s \to D_s(\phi\pi)D_s(K^{0*}K)$, $B^0 \to D_s(K^{0*}K)D(K\pi\pi)]$ and $[B_s \to D_s(\phi\pi)D_s(\pi\pi\pi)$, $B^0 \to D_s(\pi\pi\pi)D(K\pi\pi)]$ and averaging the three results, the following value for the branching fraction of $B_s \to D_s^+D_s^-$ was obtained:

$$\begin{aligned} Br[B_s \to D_s^+D_s^-] &= \{10.5 \pm 1.2(Stat) \pm 0.8(Br) \pm 1.2(Br_{Corr}) \\ &\quad \pm 1.6(f_s/f_d) \pm 2.1(Br_{B^0})\} \times 10^{-3} \\ &= (10.5 \pm 3.2) \times 10^{-3} \end{aligned}$$

This measurement is in good agreement with the result $(9.4^{+4.4}_{-4.2}) \times 10^{-3}$ obtained in a former CDF measurement which used 355 pb^{-1} of data [3].

The uncertainties quoted in this thesis comprise statistical errors and systematic uncertainties stemming from branching fractions only. The small uncertainties introduced by the combined reconstruction and selection efficiencies are included in the statistical error. It can be seen that uncertainties originating from branching fractions and the quark production ratio f_s/f_d still place a large contribution to the overall uncertainty which is not significantly smaller compared to the former measurement.

As a byproduct of the application of the simultaneous fit method the relative branching fractions of different combinations of excited $D^*_{(s)}$ meson final states with respect to the ground states were measured. The results for the different combinations of $Br[B^0 \to D_s^{(*)+}D^{(*)-}]/Br[B^0 \to D_s^+D^-]$ agree well with the values published by the Particle Data Group [4]. The measurement of $Br[B_s \to D_s^{(*)+}D_s^{(*)-}]/Br[B_s \to D_s^+D_s^-]$ represents the first of this kind in the B_s system. The results obtained throughout this analysis are

$$\frac{Br\,(B_s \to D_s^{+*}D_s^-)}{Br\,(B_s \to D_s^+D_s^-)} = 2.87 \pm 0.52$$

$$\frac{Br\,(B_s \to D_s^{+*}D_s^{-*})}{Br\,(B_s \to D_s^+D_s^-)} = 3.44 \pm 0.71$$

$$\frac{Br\,(B^0 \to D_s^+D^{-*})}{Br\,(B^0 \to D_s^+D^-)} = 0.91 \pm 0.14$$

$$\frac{Br\,(B^0 \to D_s^{+*} D^-)}{Br\,(B^0 \to D_s^+ D^-)} = 1.12 \pm 0.07$$

$$\frac{Br\,(B^0 \to D_s^{+*} D^{-*})}{Br\,(B^0 \to D_s^+ D^-)} = 2.94 \pm 0.20$$

Adding up the relative branching fractions of the B^0 decays where one of the two $D_{(s)}$ mesons is excited allows a rough comparison between the relative branching fractions in the B_s and B^0 sector. Albeit shifted to some higher values, within the large uncertainties the relative branching fractions $Br[B_s \to D_s^* D_s^{(*)}]/Br[B_s \to D_s D_s]$ are very similar to those in $B^0 \to D_s^{(*)} D^{(*)}$ decays.

There are several possibilities and opportunities to advance the measurement of $Br[B_s \to D_s^+ D_s^-]$ in the future. First, with up to 4 fb^{-1} of CDF data on tape the accuracy of a future analysis will significantly benefit from the excess of statistics. In addition, a thorough understanding of all reflections and crosstalks occurring in the reconstruction of $B_0 \to D_s^+ D_s^-$ decays would allow to develop a more precise fit model, particularly for the three decay modes that were omitted in this analysis. Incorporating these final states into a measurement of $Br[B^0 \to D_s^+ D_s^-]$ would provide some additional statistics. Furthermore, a possibility to correctly account for correlations of the input parameters in the fit model was briefly outlined in Chapter 5 already. Moreover, since using the decay $B^0 \to D_s^+ D^-$ introduces the additional systematic uncertainty stemming from the quark production ratio f_s/f_d, it is worth considering to prefer a different normalization channel for obtaining $Br[B_s \to D_s^+ D_s^-]$ by a relative branching fraction measurement. The branching fractions of the decays $B_s \to D_s \pi$ or $B_s \to \pi\pi\pi$ are possible candidates for this task.

Finally, the presented methods utilizing simultaneous fitting procedures could be modified to directly measure the inclusive branching fraction $Br[B_s \to D_s^{(*)+} D_s^{(*)-}]$ by using this quantity as a fit parameter in a simultaneous fit. Provided that a more precise understanding of the CP content of these excited final states is achieved in future studies, a measurement of $Br[B_s \to D_s^{(*)+} D_s^{(*)-}]$ could be used to provide for a better estimation of the CP decay width difference in the B_s–$\bar{B}_s$ system.

Appendix A

Definition of Variables

$L_{xy}(P)$ Decay length of particle P

$\sigma_{L_{xy}}(P)$ Error of the decay length of particle P

$L_{xy}/\sigma_{L_{xy}}(P)$ Significance of the decay length of particle P

$L_{xy}(P \leftarrow C)$ Decay length of child particle C with respect to parent Particle P

$|d_0|(P)$ Absolute value of the impact parameter of particle P

$d_0/\sigma_{d_0}(P)$ Impact parameter significance of particle P

$\min d_0/\sigma_{d_0}$ Minimum of the impact parameter significances of all particle tracks in the final state

$d_0^{lts}/\sigma_{d_0}(P)$ Lifetime signed impact parameter of particle P

$\min d_0(C_P^n)$ Minimum of the impact parameters of the n child particles C^n being decay products of parent particle P

$\max d_0(C_P^n)$ Maximum of the impact parameters of the n final state particles C^n being decay products of parent particle P

$\chi^2_{3D}(P)$ Three dimensional χ^2 of the vertex fit for particle P

$\chi^2_{r\phi}(P)$ Two dimensional χ^2 of the vertex fit for particle P

$p_t(P)$ Transverse Momentum of Particle P

$\min(p_t)$ Minimum of the transverse momenta of all particle tracks

$\mathrm{m}(P)$ Invariant mass of particle P

$\mathrm{m}_{C_P^i,C_P^j}$ Invariant mass of the four momenta of child particles $C_{i,j}$ being decay products of parent particle P

$\min\left(m_{C_P^i,C_P^j}, m_{C_P^j,C_P^k}\right)$ Minimum of the two invariant masses $\mathrm{m}_{C_P^i,C_P^j}$ and $\mathrm{m}_{C_P^j,C_P^k}$

$\max\left(m_{C_P^i,C_P^j}, m_{C_P^j,C_P^k}\right)$ Maximum of the two invariant masses $\mathrm{m}_{C_P^i,C_P^j}$ and $\mathrm{m}_{C_P^j,C_P^k}$

$\theta_{hel}\left(C_P^i\right)$ Cosine of the helicity angle between the three momentum of the parent particle P and the three momentum of the ith child particle C^i

$\texttt{PID.R}_F\left(C_P^i\right)$ Particle Identification Variable: Likelihood ratio for the identification of the final state particle F given for the child particle C coming from the parent particle P

$\texttt{PID.pullTof}_F\left(C_P^i\right)$ Particle Identification Variable: Pull to the likelihood ratio for the identification of the final state particle F given for the child particle C coming from the parent particle P

$q\left(C_P^i\right)$ Electrical charge of the ith child particle C coming from the parent particle P

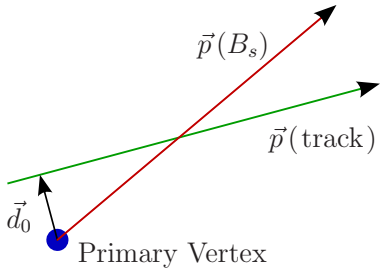


Figure A.1: Definition of the lifetime signed impact parameter for a track reconstructed to come from a B_s decay: $d_0^{lts} = \left|\vec{d}_0\right| \text{ signum } \left(\vec{d}_0 \cdot \vec{p}(B_s)\right)$.

Appendix B

Compilation of Neural Network Training Results

B_s Decay Mode 1: $B_s \to D_s(\phi\pi)\, D_s(\phi\pi)\,, \phi \to KK$

Rank	Name	Significance [σ]	Index
-	Target	-	1
1	$L_{xy}/\sigma_{L_{xy}}(D_s^2)$	349.27	8
2	$L_{xy}/\sigma_{L_{xy}}(B_s)$	104.68	2
3	$\texttt{PID.R}_\pi(K^2_{\phi^2})$	83.52	24
4	$\chi^2_{r\phi}(B_s)$	57.52	4
5	$\lvert d_0(B_s)\rvert$	45.84	3
6	$\min(p_t)$	37.17	26
7	$\theta_{hel}(K^1_{D_s^2})$	33.56	21
8	$\texttt{PID.R}_K(K^1_{\phi^2})$	30.05	22
9	$\mathrm{m}(\phi_{D_s^2})$	28.38	11
10	$L_{xy}/\sigma_{L_{xy}}(D_s^1)$	25.38	6
11	$\min(d^0/\sigma_{d_0})$	18.79	27
12	$\texttt{PID.R}_K(\pi_{D_s^2})$	17.16	17
13	$\lvert d_0(D_s^2)\rvert$	16.41	9
14	$\theta_{hel}(K^1_{D_s^1})$	15.93	18
15	$\chi^2_{3D}(B_s)$	14.01	5
16	$\mathrm{m}(\phi^1)$	13.24	10
17	$\chi^2_{3D}(D_s^1)$	13.16	7
18	$d_0^{lts}/\sigma_{d_0}(\pi_{D_s^2})$	12.57	16
19	$\chi^2_{3D}(\phi_{D_s^2})$	12.10	13
20	$L_{xy}(D_s^2 \leftarrow \phi^2)$	11.56	25
21	$\lvert d_0(\pi_{D_s^2})\rvert$	11.43	15
22	$\texttt{PID.R}_K(\pi_{D_s^1})$	11.26	14
23	$d_0^{lts}/\sigma_{d_0}(K^1_{\phi^2})$	10.32	20
24	$\texttt{PID.R}_K(K^2_{\phi^1})$	10.16	19
25	$\texttt{PID.pullTof}_\pi(K^1_{\phi^2})$	8.78	23
26	$L_{xy}/\sigma_{L_{xy}}(\phi_{D_s^2})$	8.54	12

Table B.1: The input variables of the $Bs \to D_s^1(\phi^1\pi)D_s^2(\phi^2\pi)$ network. $K^{1,2}$ denotes the first or second kaon from a ϕ meson decay. See Appendix A for variable definitions.

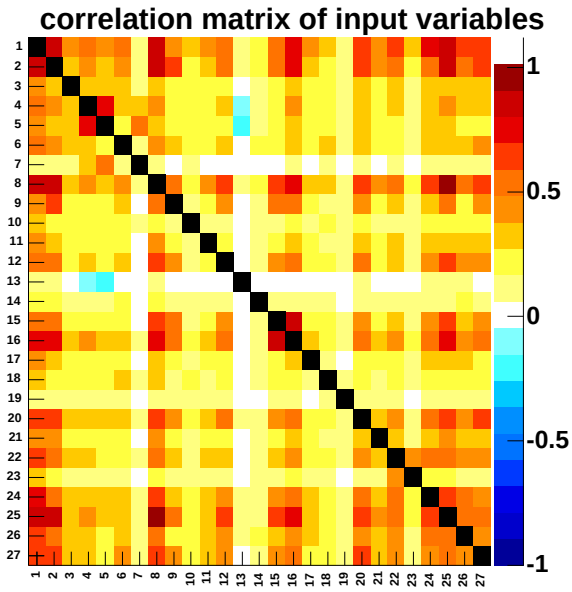
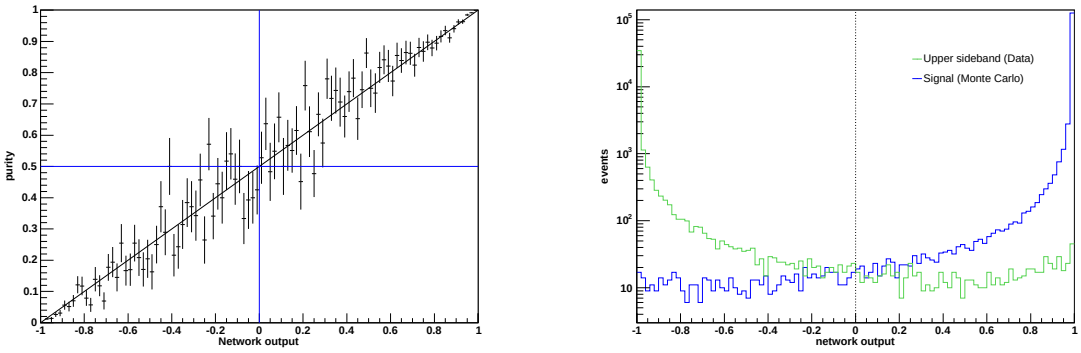


Figure B.1: Correlation matrix of the input variables used for the decay $B_s \rightarrow D_s\,(\phi\pi)\,D_s\,(\phi\pi)$.



(a) Purity over network output.

(b) Network classification of training data.

Figure B.2: Two of the training graphs indicating the performance of the network trained for the channel $Bs \rightarrow D_s\,(\phi\pi)\,D_s\,(\phi\pi)$.

B_s **Decay Mode 2:** $B_s \to D_s\left(\phi\pi\right) D_s\left(K^{0*}K\right), \phi \to KK, K^{0*} \to K\pi$

Rank	**Name**	**Significance [σ]**	**Index**
-	Target	-	1
1	$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	371.41	2
2	$\chi^2_{r\phi}\left(B_s\right)$	113.41	4
3	$L_{xy}/\sigma_{L_{xy}}\left(D^1_s\right)$	90.33	7
4	$\left\|d_0\left(B_s\right)\right\|$	61.70	4
5	$d_0^{lts}/\sigma_{d_0}\left(K_{D^2_s}\right)$	47.75	17
6	$\texttt{PID.R}_K\left(K^2_{\phi^1}\right)$	43.51	20
7	$\theta_{hel}\left(\pi_{K^{0*}}\right)$	36.58	25
8	$\min(p_t)$	36.09	28
9	$\texttt{PID.R}_K\left(K_{K^{0*}}\right)$	30.96	22
10	$\theta_{hel}\left(K^1_\phi\right)$	29.28	18
11	$\mathrm{m}(\phi)$	28.08	12
12	$\min\left(d^0/\sigma_{d_0}\right)$	26.52	29
13	$p_t\left(K_{D^2_s}\right)$	19.28	16
14	$\chi^2_{3D}\left(B_s\right)$	18.08	6
15	$L_{xy}/\sigma_{L_{xy}}\left(D^2_s\right)$	16.13	9
16	$\texttt{PID.R}_K\left(K^1_\phi\right)$	15.68	19
17	$\texttt{PID.R}_K\left(\pi_{K^{0*}}\right)$	15.52	26
18	$\left\|d_0\left(K_{K^{0*}}\right)\right\|$	15.51	21
19	$p_t\left(\pi_{K^{0*}}\right)$	14.04	23
20	$L_{xy}\left(B_s \leftarrow D^2_s\right)$	13.70	23
21	$\mathrm{m}\left(K^{0*}\right)$	11.45	15
22	$d_0^{lts}/\sigma_{d_0}\left(\pi_{K^{0*}}\right)$	10.90	24
23	$p_t\left(B_s\right)$	10.46	5
24	$\left\|d_0\left(D^2_s\right)\right\|$	9.49	10
25	$\chi^2_{3D}\left(\phi\right)$	8.23	14
26	$L_{xy}/\sigma_{L_{xy}}\left(\phi\right)$	8.13	13
27	$\chi^2_{r\phi}\left(D^1_s\right)$	5.21	8
28	$\chi^2_{3D}\left(\phi\right)$	5.30	11

Table B.2: The input variables of the $Bs \to D^1_s\left(\to \phi\pi\right) D^2_s\left(\to K^{0*}K\right)$ network. See Appendix A for variable definitions.

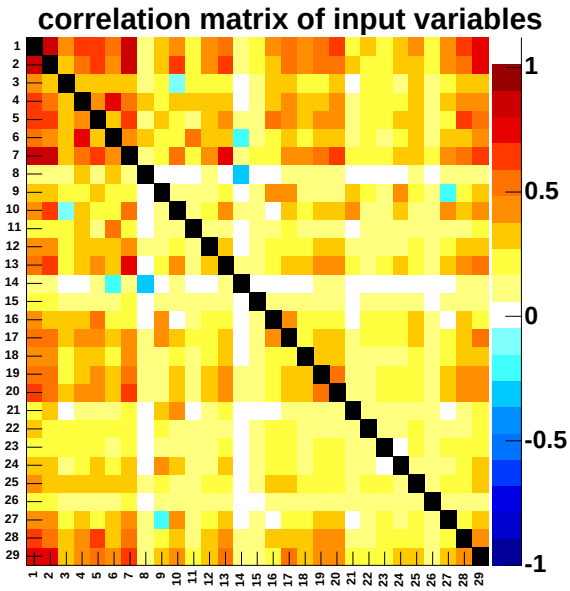
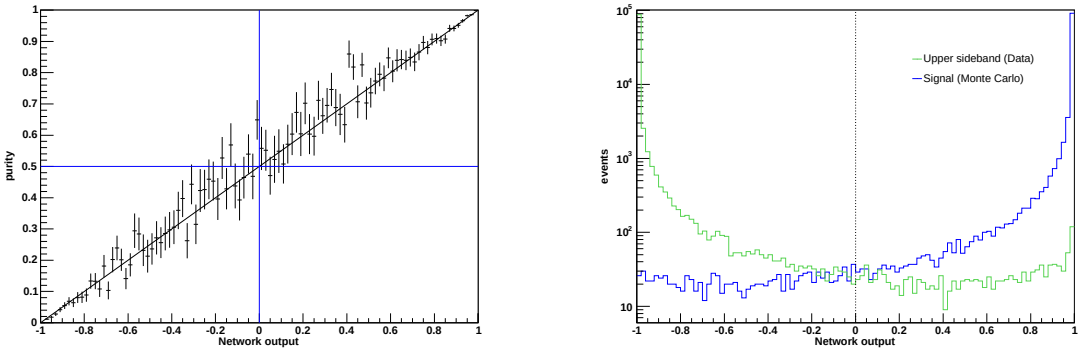


Figure B.3: Correlation matrix of the input variables used for the decay $B_s \rightarrow D_s\,(\phi\pi)\,D_s\,(K^{0*}K)$.



(a) Purity over network output

(b) Network classification of training data

Figure B.4: Training graphs indicating the performance of the network trained for the channel $Bs \rightarrow D_s\,(\phi\pi)\,D_s\,(K^{0*}K)$.

B_s Decay Mode 3: $B_s \to D_s(\phi\pi)\, D_s(\pi\pi\pi)\,, \phi \to KK$

Rank	Name	Significance [σ]	Index
-	Target	-	1
1	$L_{xy}/\sigma_{L_{xy}}(B_s)$	309.99	2
2	$\chi^2_{3D}(B_s)$	114.24	5
3	$L_{xy}/\sigma_{L_{xy}}(D_s^2)$	84.57	8
4	$p_t(D_s^1)$	69.93	7
5	$\lvert d_0(B_s)\rvert$	67.03	3
6	$L_{xy}(B_s \leftarrow D_s^2)$	54.62	23
7	$p_t(D_s^2)$	50.98	10
8	$L_{xy}/\sigma_{L_{xy}}(D_s^1)$	49.47	6
9	$\texttt{PID.R}_K(K_\phi^2)$	47.69	21
10	$\min(d^0/\sigma_{d_0})$	37.14	25
11	$\theta_{hel}(K_\phi^1)$	34.39	18
12	$\max(m_{\pi^1\pi^2}, m_{\pi^2\pi^3})$	33.08	29
13	$\mathrm{m}(\phi)$	32.74	11
14	$\texttt{PID.R}_K(\pi^1_{D_s^2})$	31.50	15
15	$\texttt{PID.R}_K(K^1_{\phi^1})$	29.06	19
16	$\texttt{PID.R}_\pi(\pi^3_{D_s^2})$	28.06	17
17	$\chi^2_{r\phi}(B_s)$	25.24	4
18	$L_{xy}(B_s \leftarrow D_s^1)$	24.43	22
19	$\texttt{PID.R}_K(\pi^2_{D_s^2})$	22.96	16
20	$\lvert d_0(D_s^2)\rvert$	21.36	9
21	$\min(m_{\pi^1\pi^2}, m_{\pi^2\pi^3})$	19.89	27
22	$\chi^2_{3D}(\phi)$	17.47	13
23	$\min(p_t)$	16.30	24
24	$m_{\pi^1\pi^3}$	15.82	30
25	$\texttt{PID.R}_\pi(\pi_{D_s^1})$	14.37	14
26	$\mathrm{m}(D_s^2)$	14.13	31
27	$L_{xy}/\sigma_{L_{xy}}(\phi_{D_s^1})$	11.72	12
28	$\texttt{PID.pullTof}_\pi(K^1_{\phi^1})$	9.47	20
29	$\max\, d^0(\pi^{1,2,3}_{D_s^2})$	7.07	28
30	$\min\, d^0(\pi^{1,2,3}_{D_s^2})$	4.63	26

Table B.3: Input variables of the $Bs \to D_s^1(\phi\pi)D_s^2(\pi^1\pi^2\pi^3)$ network

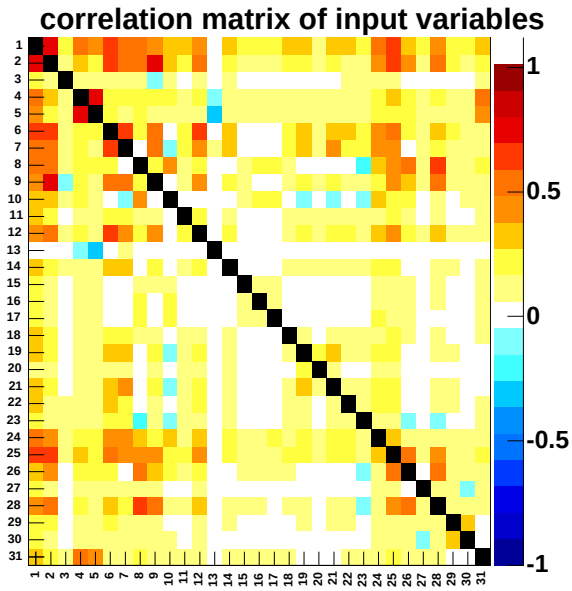
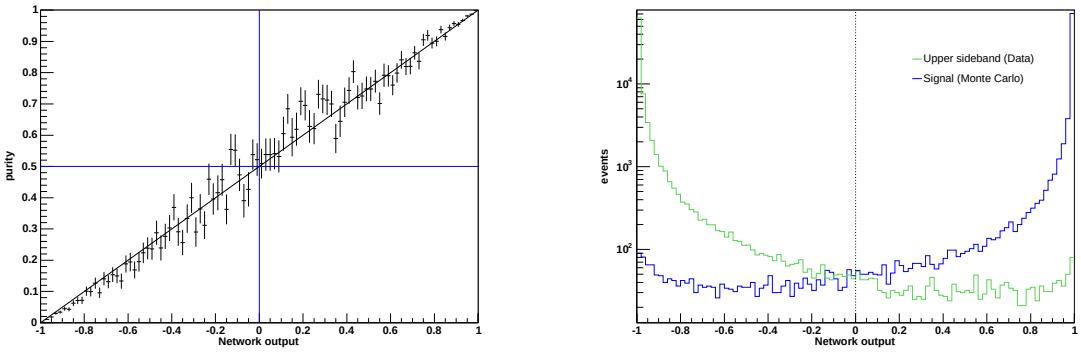


Figure B.5: Correlation matrix of the input variables used for the decay $B_s \rightarrow D_s\,(\phi\pi)\,D_s\,(\pi\pi\pi)$.



(a) Purity over network output

(b) Network classification of training data

Figure B.6: Training graphs indicating the performance of the network trained for the channel $Bs \rightarrow D_s\,(\phi\pi)\,D_s\,(\pi\pi\pi)$.

B_s Decay Mode 4: $B_s \to D_s\left(K^{0*}K\right) D_s\left(K^{0*}K\right), K^{0*} \to K\pi$

Rank	Name	Significance [σ]	Index
-	Target	-	1
1	$\chi^2_{r\phi}\left(B_s\right)$	219.59	4
2	$p_t\left(D_s^2\right)$	120.12	9
3	$d_0^{lts}/\sigma_{d_0}\left(K_{D_s^2}\right)$	77.67	18
4	$\min\left(d^0/\sigma_{d_0}\right)$	50.50	31
5	$\theta_{hel}\left(\pi_{K^{0*,1}}\right)$	37.05	26
6	$\texttt{PID.R}_K\left(K_{K^{0*}}\right)$	34.08	21
7	$\theta_{hel}\left(\pi_{K^{0*,2}}\right)$	31.83	23
8	$L_{xy}\left(B_s \leftarrow D_s^2\right)$	29.34	29
9	$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	28.36	2
10	$\chi^2_{3D}\left(B_s\right)$	24.28	5
11	$L_{xy}/\sigma_{L_{xy}}\left(D_s^1\right)$	23.47	6
12	$L_{xy}\left(B_s \leftarrow D_s^1\right)$	21.80	28
13	$\theta_{hel}\left(K_{D_s^2}\right)$	20.62	19
14	$\left\lvert d_0\left(B_s\right)\right\rvert$	18.28	3
15	$L_{xy}\left(D_s^2 \leftarrow K^{0*,2}\right)$	17.02	30
16	$d_0^{lts}/\sigma_{d_0}\left(K_{D_s^1}\right)$	15.25	16
17	$\texttt{PID.pullTof}_\pi\left(K_{D_s^2}\right)$	15.05	20
18	$p_t\left(D_s^1\right)$	14.55	7
19	$L_{xy}/\sigma_{L_{xy}}\left(D_s^2\right)$	13.68	8
20	$\texttt{PID.R}_K\left(K_{D_s^1}\right)$	13.52	17
21	$m_{K^{0*,2}}$	11.94	13
22	$p_t\left(K^{0*,2}\right)$	11.19	15
23	$\left\lvert d_0\left(K^{0*,1}\right)\right\rvert$	10.95	12
24	$\texttt{PID.R}_K\left(\pi_{K^{0*,1}}\right)$	11.14	27
25	$\texttt{PID.R}_K\left(K_{K^{0*,1}}\right)$	10.06	24
26	$L_{xy}/\sigma_{L_{xy}}\left(K^{0*,2}\right)$	9.63	14
27	$m_{D_s^2}$	9.63	10
28	$m_{K^{0*,1}}$	9.63	11
29	$\min(p_t)$	9.34	31
30	$\texttt{PID.pullTof}_K\left(K^{0*,22}\right)$	8.45	22
31	$p_t\left(\pi_{K^{0*}}\right)$	8.13	25

Table B.4: Input variables of the $Bs \to D_s^1\left(K^{0*,1}K\right) D_s^2\left(K^{0*,2}K\right)$ network.

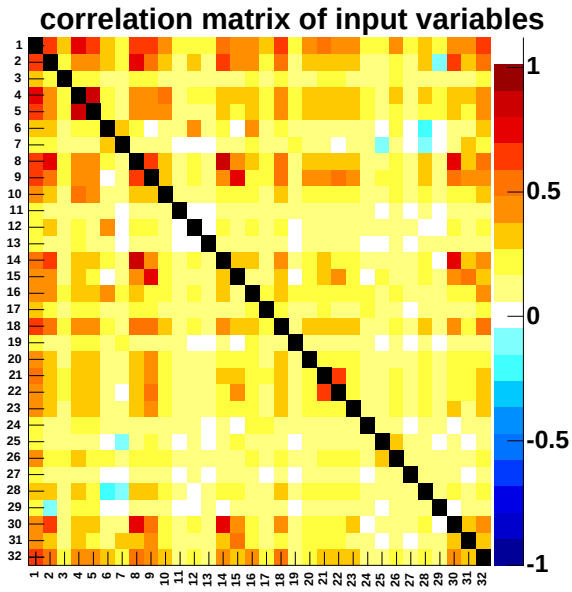
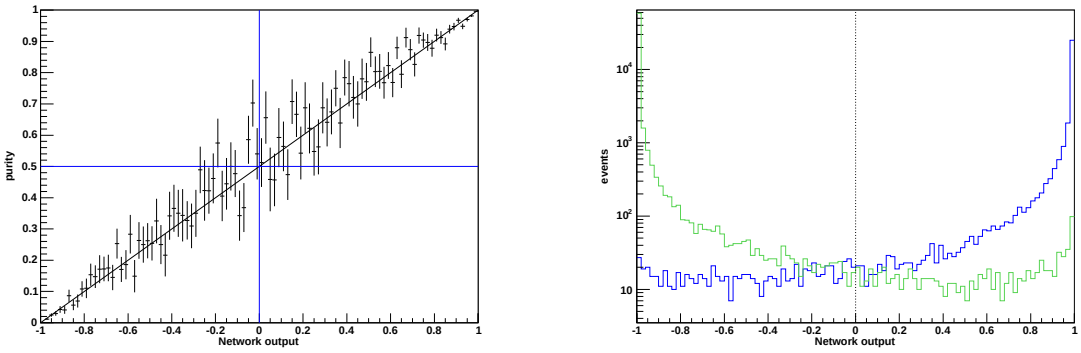


Figure B.7: Correlation matrix of the input variables used for the decay $B_s \rightarrow D_s\,(K^{0*}K)\,D_s\,(K^{0*}K)$.



(a) Purity over network output

(b) Network classification of training data

Figure B.8: Training graphs indicating the performance of the network trained for the channel $Bs \rightarrow D_s\,(K^{0*}K)\,D_s\,(K^{0*}K)$.

B_s Decay Mode 5: $B_s \to D_s\left(K^{0*}K\right) D_s\left(\pi\pi\pi\right), K^{0*} \to K\pi$

Rank	Name	Significance $[\sigma]$	Index
-	Target	-	1
1	$L_{xy}/\sigma_{L_{xy}}\left(D_s^2\right)$	143.86	9
2	$p_t\left(D_s^1\right)$	98.67	8
3	$\chi^2_{3D}\left(B_s\right)$	85.77	6
4	$\min\left(d^0/\sigma_{d_0}\right)$	47.56	28
5	$\lvert d_0\left(B_s\right)\rvert$	42.50	3
6	$L_{xy}\left(B_s \leftarrow D_s^1\right)$	33.89	25
7	$L_{xy}\left(B_s \leftarrow D_s^2\right)$	32.31	26
8	$\theta_{hel}\left(K_{D_s^1}\right)$	29.91	14
9	$\theta_{hel}\left(\pi_{K^{0*,1}}\right)$	29.19	19
10	$p_t\left(D_s^2\right)$	26.83	10
11	$L_{xy}/\sigma_{L_{xy}}\left(D_s^1\right)$	25.68	7
12	$\texttt{PID.R}_K\left(K_{K^{0*,1}}\right)$	23.96	18
13	$\max(m_{\pi^1\pi^2}, m_{\pi^2\pi^3})$	21.79	30
14	$\chi^2_{r\phi}\left(B_s\right)$	20.40	4
15	$\texttt{PID.R}_\pi\left(\pi^3\right)$	18.74	24
16	$\min(m_{\pi^1\pi^2}, m_{\pi^2\pi^3})$	18.46	29
17	$\texttt{PID.R}_K\left(\pi^1\right)$	17.13	21
18	$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	16.07	2
19	$L_{xy}\left(D_s^1 \leftarrow K^{0*}\right)$	14.84	27
20	$\texttt{PID.R}_K\left(K_{D_s}1\right)$	15.53	15
21	$d_0^{lts}/\sigma_{d_0}\left(K_{D_s^1}\right)$	13.86	13
22	$\texttt{PID.R}_K\left(\pi^2\right)$	13.36	22
23	$m_{K^{0*}}$	12.74	11
24	$m_{\pi^1\pi^3}$	12.27	31
25	$\texttt{PID.R}_K\left(\pi_{K^{*0}}\right)$	11.98	20
26	$\texttt{PID.pullTof}_K\left(K_{D_s^1}\right)$	11.82	16
27	$p_t\left(K_{D_s^1}\right)$	7.37	12
28	$p_t\left(B_s\right)$	6.91	5
29	$p_t\left(\pi^3\right)$	8.13	23
30	$m_{D_s^2}$	7.61	32
31	$\theta_{hel}\left(\pi_{K^{0*}}\right)$	7.40	19
32	$\lvert d_0\left(K_{K^{0*}}\right)\rvert$	5.74	17

Table B.5: Input variables of the $Bs \to D_s^1\left(K^{0*}K\right) D_s^2\left(\pi^1\pi^2\pi^3\right)$ network.

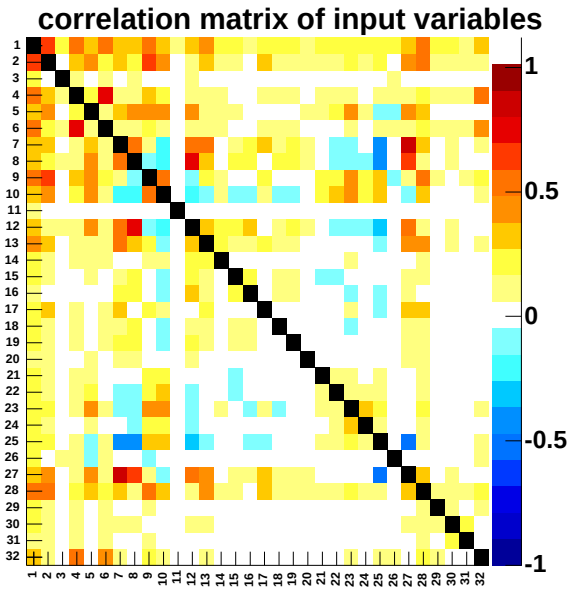
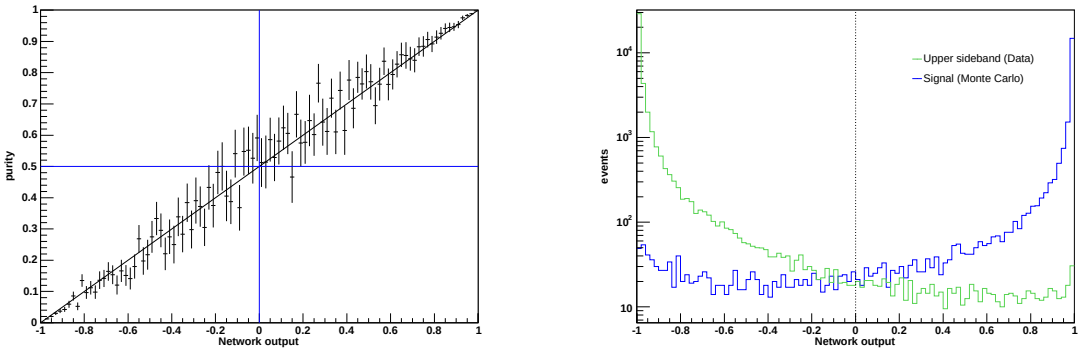


Figure B.9: Correlation matrix of the input variables used for the decay $B_s \rightarrow D_s\,(K^{0*}K)\,D_s\,(\pi\pi\pi)$.



(a) Purity over network output

(b) Network classification of training data

Figure B.10: Training graphs indicating the performance of the network trained for the channel $Bs \rightarrow D_s\,(K^{0*}K)\,D_s\,(\pi\pi\pi)$.

B_s Decay Mode 6: $B_s \to D_s(\pi\pi\pi)\, D_s(\pi\pi\pi)$

Rank	Name	Significance [σ]	Index
-	Target	-	1
1	$L_{xy}/\sigma_{L_{xy}}\left(D_s^2\right)$	132.78	8
2	$\chi^2_{3D}\left(\phi_{B_s}\right)$	69.19	5
3	$\min d^0/\sigma_{d_0}$	46.54	23
4	$\lvert d_0\left(B_s\right)\rvert$	36.30	3
7	$L_{xy}\left(B_s \leftarrow D_s^1\right)$	27.64	21
5	$p_t\left(D_s^2\right)$	27.58	10
6	$L_{xy}/\sigma_{L_{xy}}\left(D_s^1\right)$	26.50	6
8	$\max(m_{\pi^1\pi^2}, m_{\pi^2\pi^3})_{\,D_s^1}$	20.34	30
9	$\texttt{PID.R}_\pi\left(\pi^1_{D_s^1}\right)$	19.72	12
10	$L_{xy}\left(B_s \leftarrow D_s^2\right)$	15.85	22
11	$L_{xy}/\sigma_{L_{xy}}\left(B_s\right)$	17.32	2
12	$m_{\pi^1\pi^2}\left(D_s^2\right)$	16.06	25
13	$m_{\pi^2\pi^3}\left(D_s^2\right)$	14.96	27
14	$\texttt{PID.R}_\pi\left(\pi^3_{D_s^2}\right)$	14.95	20
15	$\min(m_{\pi^1\pi^2}, m_{\pi^2\pi^3})_{\,D_s^1}$	14.46	29
17	$m_{D_s^1}$	14.33	32
18	$\texttt{PID.R}_K\left(\pi^2_{D_s^1}\right)$	13.71	13
16	$m_{D_s^2}$	13.03	28
19	$\texttt{PID.R}_\pi\left(\pi^1_{D_s^2}\right)$	13.13	17
20	$m_{\pi^1\pi^3}\left(D_s^2\right)$	12.67	26
21	$\texttt{PID.R}_\pi\left(\pi^3_{D_s^1}\right)$	11.68	15
22	$m_{\pi^1\pi^3}\left(D_s^1\right)$	10.30	31
23	$p_t\left(\pi^3_{D_s^1}\right)$	10.27	14
24	$\chi^2_{r\phi}\left(D_s^2\right)$	9.68	9
25	$\theta_{hel}\left(\pi^1_{D_s^2}\right)$	9.27	16
26	$\texttt{PID.R}_\pi\left(\pi^2_{D_s^2}\right)$	9.26	18
27	$\left\lvert d_0\left(D_s^1\right)\right\rvert$	6.13	7
28	$\chi^2_{r\phi}\left(B_s\right)$	6.06	4
29	$\max d^0\left(\pi_{D_s^2}\right)$	5.10	24
30	$d_0^{lts}/\sigma_{d_0}\left(\pi^3_{D_s^2}\right)$	4.26	19
31	$ct_{D_s^2}$	4.22	11

Table B.6: Input variables of the $Bs \to D_s^1(\pi^1\pi^2\pi^3)\, D_s^2(\to \pi^1\pi^2\pi^3)$ network.

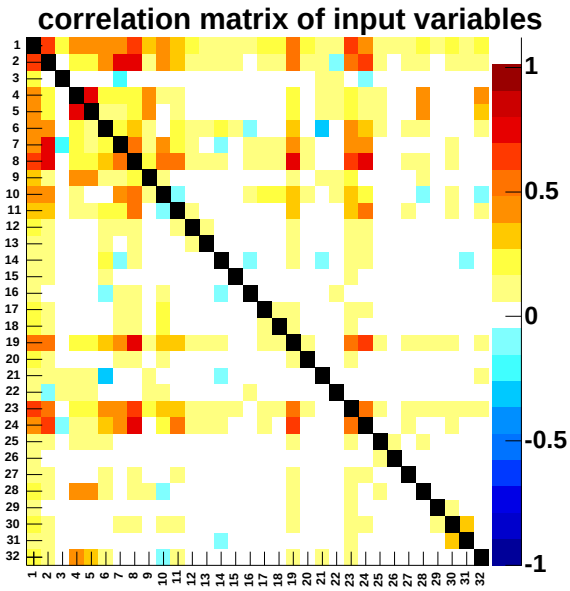
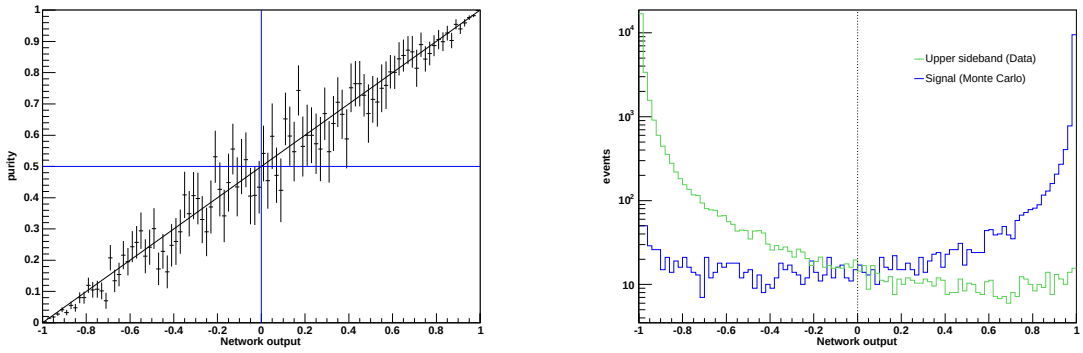


Figure B.11: Correlation matrix of the input variables used for the decay $B_s \rightarrow D_s\,(\pi\pi\pi)\,D_s\,(\pi\pi\pi)$.



(a) Purity over network output

(b) Network classification of training data

Figure B.12: Training graphs indicating the performance of the network trained for the channel $Bs \rightarrow D_s\,(\pi\pi\pi)\,D_s\,(\pi\pi\pi)$.

B^0 Decay Mode 1: $B^0 \to D\,(K\pi\pi)\,D_s\,(\phi\pi)$

Rank	Name	Significance [σ]	Index
-	Target	-	1
1	$L_{xy}/\sigma_{L_{xy}}\,(B)$	311.90	2
2	$\chi^2_{r\phi}\,(B)$	178.55	5
3	$\min(p_t)$	130.23	27
4	$\min\left(d^0/\sigma_{d_0}\right)$	86.02	28
5	$\lvert d_0\,(B)\rvert$	68.71	3
6	$p_t\,(D_s)$	55.09	8
7	$L_{xy}/\sigma_{L_{xy}}\,(D)$	55.21	9
8	$\texttt{PID.R}_K\left(K^2_\phi\right)$	47.55	20
9	$L_{xy}\,(B \leftarrow D_s)$	37.42	24
10	$\mathrm{m}(\phi)$	35.98	12
11	$p_t\,(D)$	35.50	11
12	$\theta_{hel}\left(K^1_\phi\right)$	43.05	18
13	$\chi^2_{r\phi}\,(B)$	28.90	4
14	$\texttt{PID.R}_K\left(\pi^1_D\right)$	24.91	21
15	$L_{xy}\,(B \leftarrow D)$	22.97	25
16	$\texttt{PID.R}_K\left(K^1_\phi\right)$	22.08	19
17	$\texttt{PID.R}_K\left(\pi^2_D\right)$	17.17	23
18	$\chi^2_{3D}\,(D)$	13.30	10
19	$\lvert d_0\,(K_D)\rvert$	11.50	17
20	$p_t\left(\pi^2_D\right)$	10.93	22
21	$\chi^2_{3D}\,(D_s)$	10.29	7
22	$L_{xy}\,(D_s \leftarrow \phi)$	9.90	26
23	$\theta_{hel}\,(\pi_{D_s})$	8.99	15
24	$p_t\,(K_D)$	8.99	16
25	$\chi^2_{r\phi}\,(\phi)$	6.99	13
26	$L_{xy}/\sigma_{L_{xy}}\,(D_s)$	5.59	6
27	$d_0^{lts}/\sigma_{d_0}\,(\pi_{D_s})$	4.46	14

Table B.7: Input variables of the $B^0 \to D\,(K\pi^1\pi^2)\,D_s\,(\phi\pi)$ network.

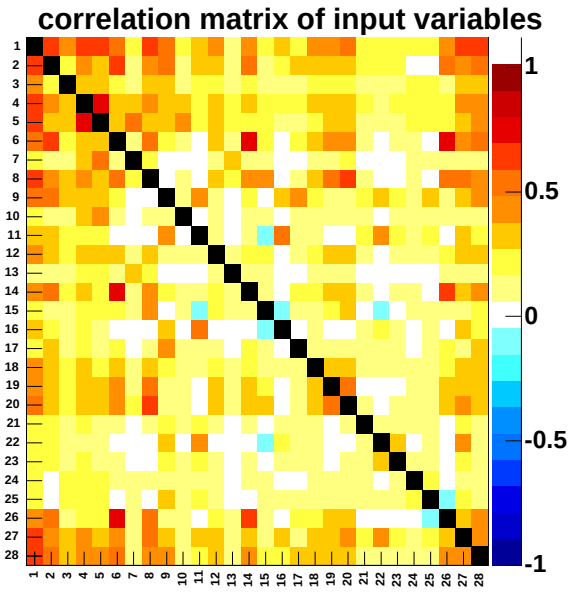
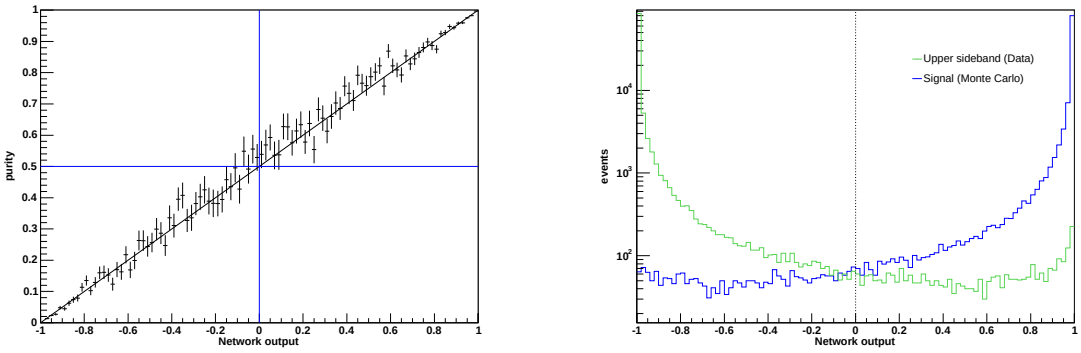


Figure B.13: Correlation matrix of the input variables used for the decay $B^0 \rightarrow D\,(K\pi\pi)\,D_s\,(\phi\pi)$.



(a) Purity over network output

(b) Network classification of training data

Figure B.14: Training graphs indicating the performance of the network trained for the channel $B^0 \rightarrow D\,(K\pi\pi)\,D_s\,(\phi\phi)$.

B^0 Decay Mode 2: $B^0 \to D\,(K\pi\pi)\,D_s\,(K^{0*}K)$

Rank	Name	Significance [σ]	Index
-	Target	-	1
1	$\chi^2_{r\phi}\,(B)$	191.10	4
2	$L_{xy}/\sigma_{L_{xy}}\,(D)$	92.30	10
3	$p_t\,(D_s)$	78.08	9
4	$d_0^{lts}/\sigma_{d_0}\,(K_{D_s})$	51.21	15
5	$\lvert d_0\,(B)\rvert$	37.45	3
6	$\theta_{hel}\,(\pi_{K^{0*}})$	31.08	22
7	$L_{xy}/\sigma_{L_{xy}}\,(B)$	27.22	2
8	$L_{xy}\,(B \leftarrow D)$	28.12	29
9	$p_t\,(B)$	25.82	6
10	$\texttt{PID.R}_K\left(K^1_{K^{0*}}\right)$	22.89	20
11	$\min\left(d^0/\sigma_{d_0}\right)$	21.00	31
12	$p_t\,(K_D)$	17.02	17
13	$\theta_{hel}\,(K_{D_s})$	15.58	16
14	$\texttt{PID.R}_\pi\left(\pi^1_D\right)$	13.41	25
15	$\mathrm{m}\left(K^{0*}\right)$	12.38	12
16	$\texttt{PID.R}_K\left(\pi^2_D\right)$	10.90	28
17	$p_t\,(B)$	11.28	5
18	$\texttt{PID.R}_K\,(\pi_{K^{0*}})$	9.90	23
19	$p_t\left(\pi^1_D\right)$	9.39	24
20	$L_{xy}/\sigma_{L_{xy}}\,(D_s)$	9.51	7
21	$p_t\left(K^{0*}\right)$	8.55	14
22	$\lvert d_0\,(K_{K^{0*}})\rvert$	8.78	19
23	$\lvert d_0\,(\pi_{K^{0*}})\rvert$	8.42	21
24	$\chi^2_{3D}\,(D)$	7.26	11
25	$\chi^2_{3D}\,(D_s)$	7.60	8
26	$\left\lvert d_0\left(\pi^2_D\right)\right\rvert$	7.32	26
27	$L_{xy}/\sigma_{L_{xy}}\left(K^{0*}\right)$	6.49	13
28	$L_{xy}\left(D_s \leftarrow K^{0*}\right)$	5.48	30
29	$\lvert d_0\,(K_D)\rvert$	4.30	18
30	$d_0^{lts}/\sigma_{d_0}\left(\pi^2_D\right)$	2.78	27

Table B.8: Input variables of the $B^0 \to D\,(K\pi^1\pi^2)\,D_s\,(K^{0*}K)$ network.

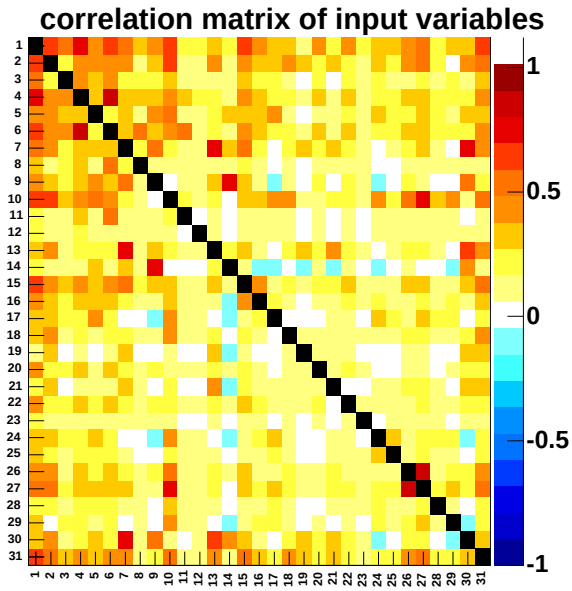
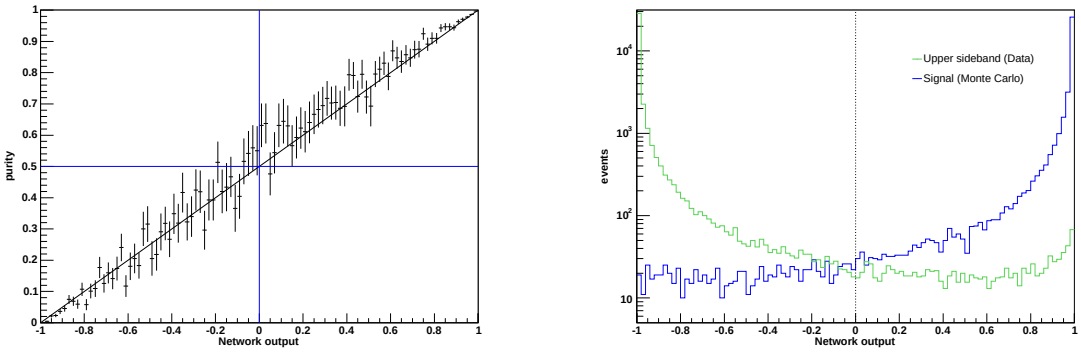


Figure B.15: Correlation matrix of the input variables used for the decay $B^0 \rightarrow D\,(K\pi\pi)\,D_s\,(K^{0*})$.



(a) Purity over network output

(b) Network classification of training data

Figure B.16: Training graphs indicating the performance of the network trained for the channel $B^0 \rightarrow D\,(K\pi\pi)\,D_s\,(K^{0*})$.

B^0 Decay Mode 3: $B^0 \to D\,(K\pi\pi)\,D_s\,(\pi\pi\pi)$

Rank	Name	Significance [σ]	Index
-	Target	-	1
1	$\chi^2_{3D}\,(B)$	176.20	6
2	$\min\left(d^0/\sigma_{d_0}\right)$	107.33	25
3	$L_{xy}/\sigma_{L_{xy}}\,(D_s)$	56.14	7
4	$p_t\,(D)$	50.12	12
5	$\lvert d_0\,(B)\rvert$	41.16	3
6	$L_{xy}/\sigma_{L_{xy}}\,(D)$	31.04	10
7	$p_t\,(D_s)$	27.38	8
8	$\max(m_{\pi^1\pi^2}, m_{\pi^2\pi^3})$	24.66	28
9	$\chi^2_{r\phi}\,(B)$	23.10	4
10	$L_{xy}\,(B \leftarrow D_s)$	20.60	9
11	$\texttt{PID.R}_\pi\left(\pi^1_{D_s}\right)$	15.93	15
12	$\min(m_{\pi^1\pi^2}, m_{\pi^2\pi^3})$	15.68	26
13	$\mathrm{m}^2(D_s)$	11.11	30
14	$\texttt{PID.R}_\pi\left(\pi^3_{D_s}\right)$	11.00	19
15	$L_{xy}/\sigma_{L_{xy}}\,(B)$	10.99	2
16	$L_{xy}\,(B \leftarrow D)$	10.74	13
17	$\texttt{PID.R}_\pi\left(\pi^2_{D_s}\right)$	10.50	16
18	$\texttt{PID.R}_\pi\left(\pi^1_{D}\right)$	9.97	22
19	$m_{\pi^1\pi^3}$	9.59	29
20	$\theta_{hel}\left(\pi^1_{D_s}\right)$	8.92	14
21	$\max\left(d^0\left(\pi_{D_s}\right)\right)$	7.65	27
22	$p_t\left(K^D\right)$	7.40	20
23	$\texttt{PID.R}_\pi\left(\pi^2_{D}\right)$	7.20	23
24	$p_t\,(B)$	6.88	5
25	$d_0^{lts}/\sigma_{d_0}\left(\pi^3_{D_s}\right)$	6.00	17
26	$\chi^2_{3D}\,(D)$	5.63	11
27	$\lvert d_0\,(K_D)\rvert$	5.34	21
28	$\min(p_t)$	4.83	24
29	$\theta_{hel}\left(\pi^3_{D_s}\right)$	4.47	18

Table B.9: Input variables of the $B^0 \to D\,(K\pi^1\pi^2)\,D_s\,(\pi^1\pi^2\pi^3)$ network.

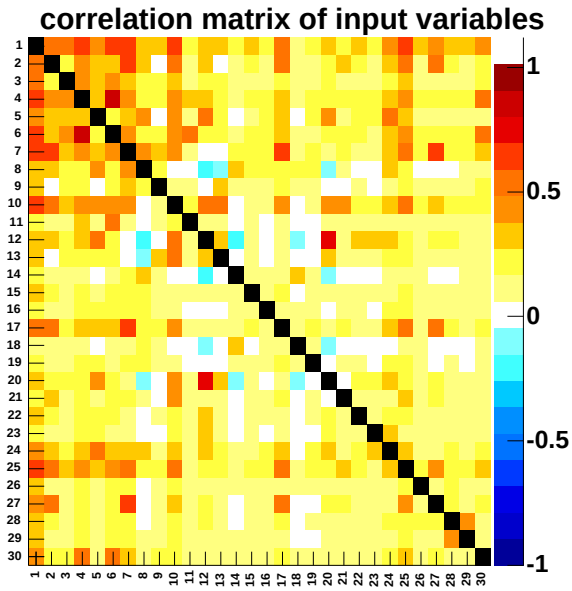
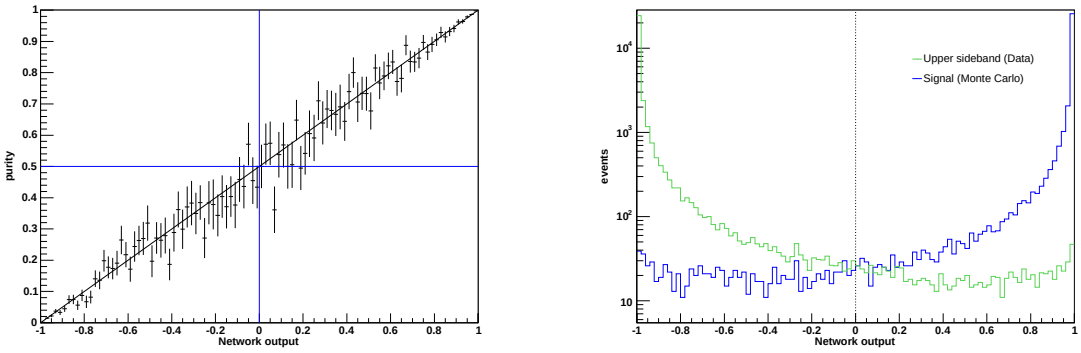


Figure B.17: Correlation matrix of the input variables used for the decay $B^0 \to D\,(K\pi\pi)\,D_s\,(\pi\pi\pi)$.



(a) Purity over network output

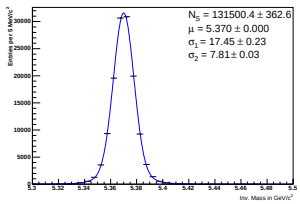
(b) Network classification of training data

Figure B.18: Training graphs indicating the performance of the network trained for the channel $B^0 \to D\,(K\pi\pi)\,D_s\,(\pi\pi\pi)$.

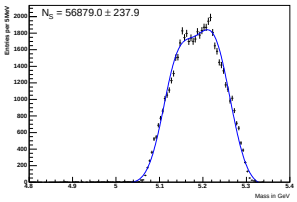
Appendix C

Compilation of MC Fit Templates

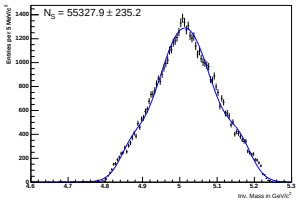
C.1 Templates for $B_s \to D_s(\phi\pi)D_s(\phi\pi)$



(a) $B_s \to D_s(\phi\pi)D_s(\phi\pi)$ fit template



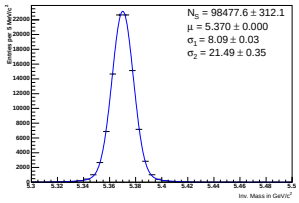
(b) $B_s \to D_s^*(\phi\pi)D_s(\phi\pi)$ fit template



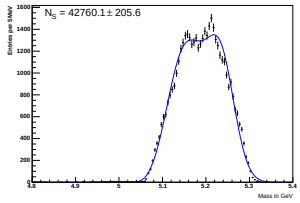
(c) $B_s \to D_s^*(\phi\pi)D_s^*(\phi\pi)$ fit template

Figure C.1: MC fit templates for the decay $B_s \to D_s(\phi\pi)D_s(\phi\pi)$: signal (left), partially reconstructed modes (middle and right).

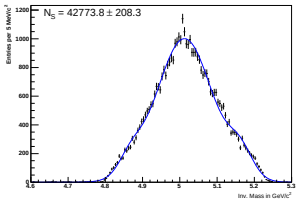
C.2 Templates for $B_s \to D_s(\phi\pi)D_s(K^*K)$



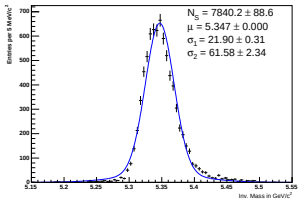
(a) $B_s \to D_s(\phi\pi)D_s(K^*K)$



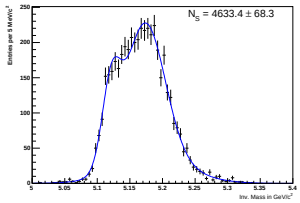
(b) $B_s \to D_s^*(\phi\pi)D_s(K^*K)$



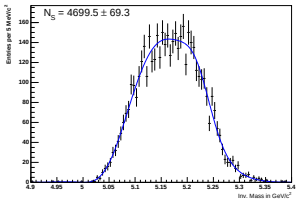
(c) $B_s \to D_s^*(\phi\pi)D_s^*(K^*K)$



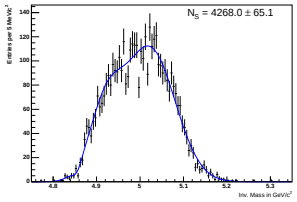
(d) $B^0 \to D(K\pi\pi)D_s(\phi\pi)$ Reflection



(e) $B^0 \to D^*(K\pi\pi)D_s(\phi\pi)$ Reflection



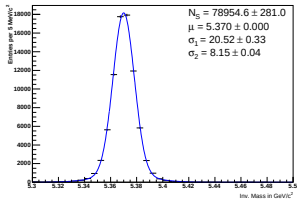
(f) $B^0 \to D_s^*(\phi\pi)D(K\pi\pi)$ Reflection



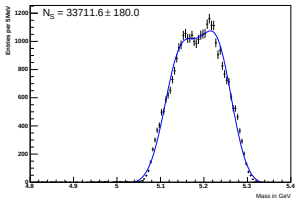
(g) $B^0 \to D_s^*(\phi\pi)D^*(K\pi\pi)$ Reflection

Figure C.2: MC fit templates for the decay $B_s \to D_s(\phi\pi)D_s(K^*K)$: signal, partially reconstructed modes, reflection modes (from left to right and top to bottom).

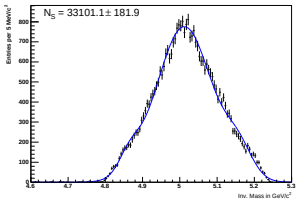
C.3 Templates for $B_s \to D_s(\phi\pi)D_s(\pi\pi\pi)$



(a) $B_s \to D_s(\phi\pi)D_s(\pi\pi\pi)$



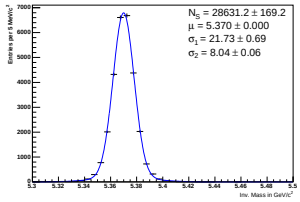
(b) $B_s \to D_s^*(\phi\pi)D_s(\pi\pi\pi)$



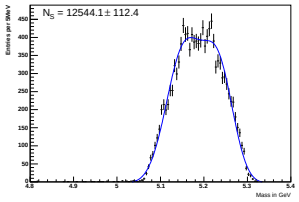
(c) $B_s \to D_s^*(\phi\pi)D_s^*(\pi\pi\pi)$

Figure C.3: MC fit templates for the decay $B_s \to D_s(\phi\pi)D_s(\pi\pi\pi)$: signal (left), partially reconstructed modes (middle and right).

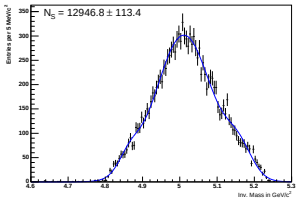
C.4 Templates for $B_s \to D_s(K^*K)D_s(K^*K)$



(a) $B_s \to D_s(K^*K)D_s(K^*K)$



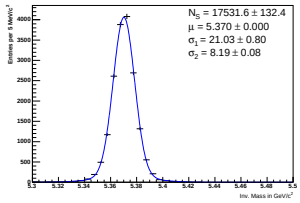
(b) $B_s \to D_s^*(K^*K)D_s(K^*K)$



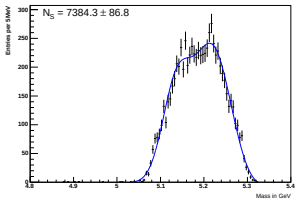
(c) $B_s \to D_s^*(K^*K)D_s^*(K^*K)$

Figure C.4: MC fit templates for the decay $B_s \to D_s(K^*K)D_s(K^*K)$: signal (left), partially reconstructed modes (middle and right).

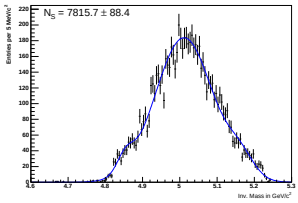
C.5 Templates for $B_s \to D_s(K^*K)D_s(\pi\pi\pi)$



(a) $B_s \to D_s(K^*K)D_s(\pi\pi\pi)$



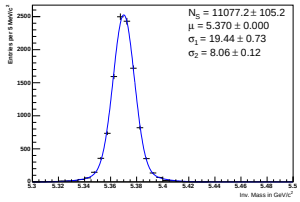
(b) $B_s \to D_s^*(K^*K)D_s(\pi\pi\pi)$



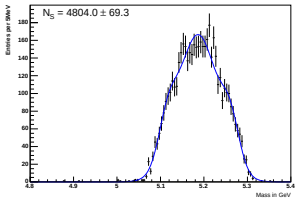
(c) $B_s \to D_s^*(K^*K)D_s^*(\pi\pi\pi)$

Figure C.5: MC fit templates for the decay $B_s \to D_s(K^*K)D_s(\pi\pi\pi)$: signal (left), partially reconstructed modes (middle and right).

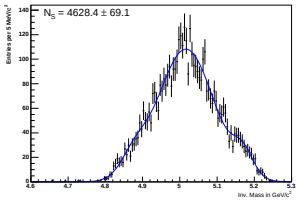
C.6 Templates for $B_s \to D_s(\pi\pi\pi)D_s(\pi\pi\pi)$



(a) $B_s \to D_s(\pi\pi\pi)D_s(\pi\pi\pi)$



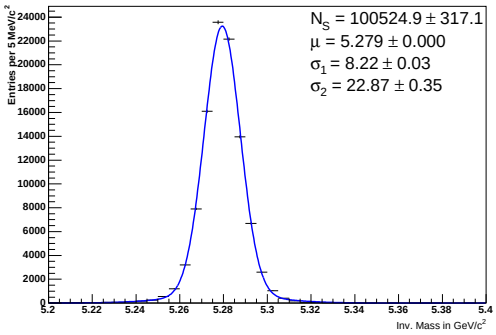
(b) $B_s \to D_s^*(\pi\pi\pi)D_s(\pi\pi\pi)$



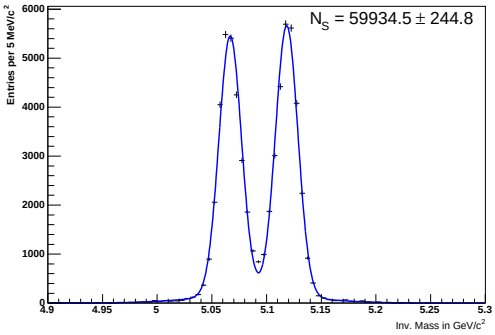
(c) $B_s \to D_s^*(\pi\pi\pi)D_s^*(\pi\pi\pi)$

Figure C.6: MC fit templates for the decay $B_s \to D_s(\pi\pi\pi)D_s(\pi\pi\pi)$: signal (left), partially reconstructed modes (middle and right).

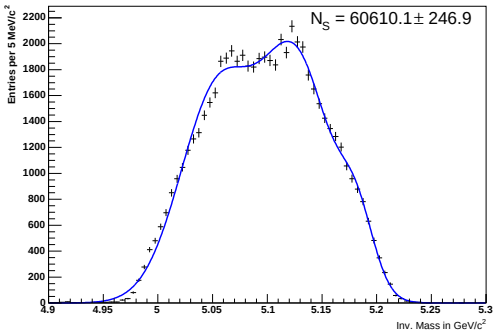
C.7 Templates for $B^0 \to D_s(\phi\pi)D(K\pi\pi)$



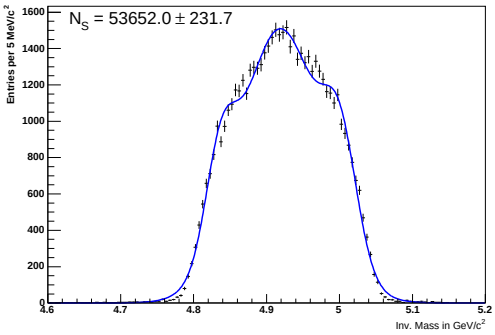
(a) $B^0 \to D_s(\phi\pi)D(K\pi\pi)$



(b) $B^0 \to D_s(\phi\pi)D^*(K\pi\pi)$



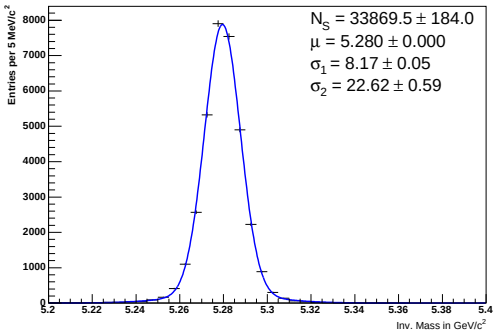
(c) $B^0 \to D_s^*(\phi\pi)D(K\pi\pi)$



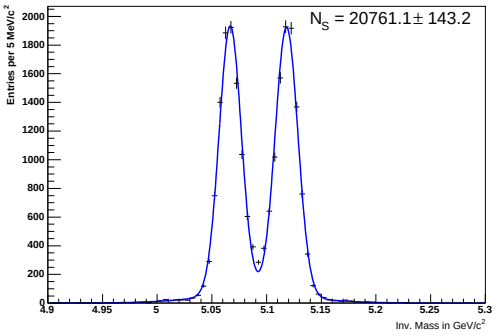
(d) $B^0 \to D_s^*(\phi\pi)D^*(K\pi\pi)$

Figure C.7: MC fit templates for the decay $B^0 \to D_s(\phi\pi)D(K\pi\pi)$: signal (top left), partially reconstructed decay modes.

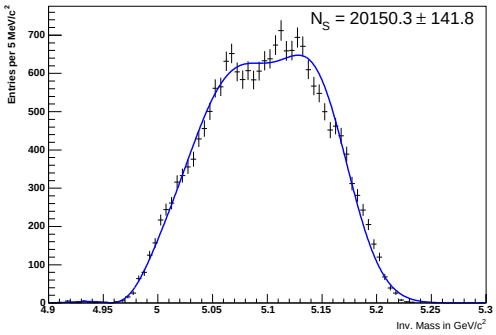
C.8 Templates for $B^0 \to D_s(K^*K)D(K\pi\pi)$



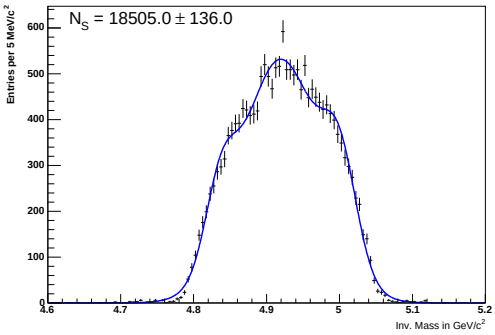
(a) $B^0 \to D_s(K^*K)D(K\pi\pi)$



(b) $B^0 \to D_s(\phi\pi)D^*(K\pi\pi)$



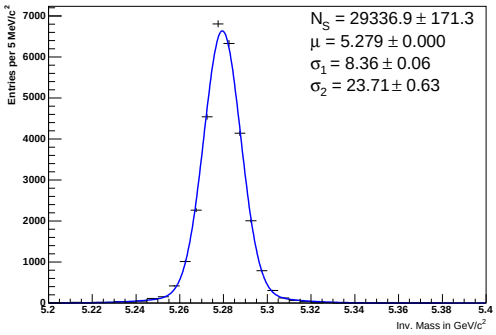
(c) $B^0 \to D_s^*(\phi\pi)D(K\pi\pi)$



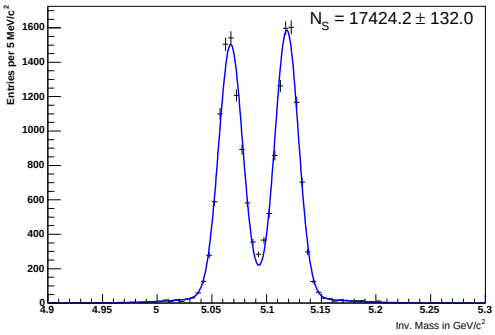
(d) $B^0 \to D_s^*(\phi\pi)D^*(K\pi\pi)$

Figure C.8: MC fit templates for the decay $B^0 \to D_s(K^*K)D(K\pi\pi)$: signal (top left), partially reconstructed decay modes.

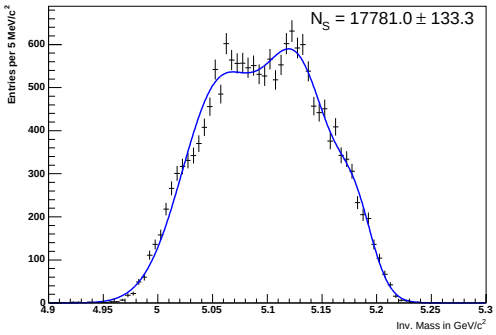
C.9 Templates for $B^0 \to D_s(\pi\pi\pi)D(K\pi\pi)$



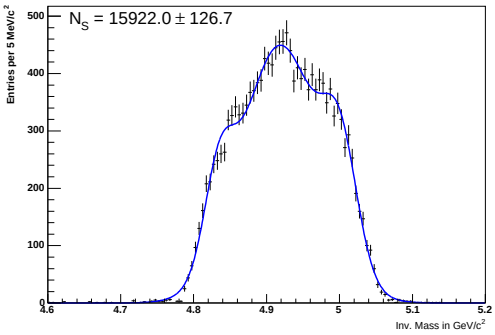
(a) $B^0 \to D_s(K^*K)D(K\pi\pi)$



(b) $B^0 \to D_s(\pi\pi\pi)D^*(K\pi\pi)$



(c) $B^0 \to D_s^*(\pi\pi\pi)D(K\pi\pi)$



(d) $B^0 \to D_s^*(\pi\pi\pi)D^*(K\pi\pi)$

Figure C.9: MC fit templates for the decay $B^0 \to D_s(K^*K)D(K\pi\pi)$: signal (top left), partially reconstructed decay modes.

Appendix D

Compilation of Simultaneous Fit Results

D.1 Simultaneous Mass Fit for the $f_{D_s^* D_s}$ and $f_{D_s^* D_s^*}$ Measurement

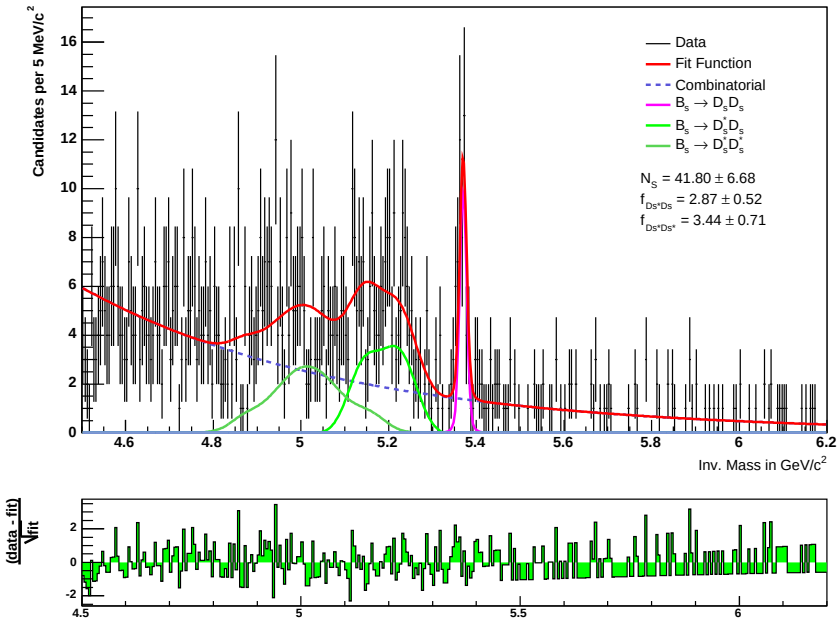


Figure D.1: Result of the simultaneous B_s mass fit for the decay $B_s \to D_s\,(\to \phi\pi)\, D_s\,(\to \phi\pi)$.

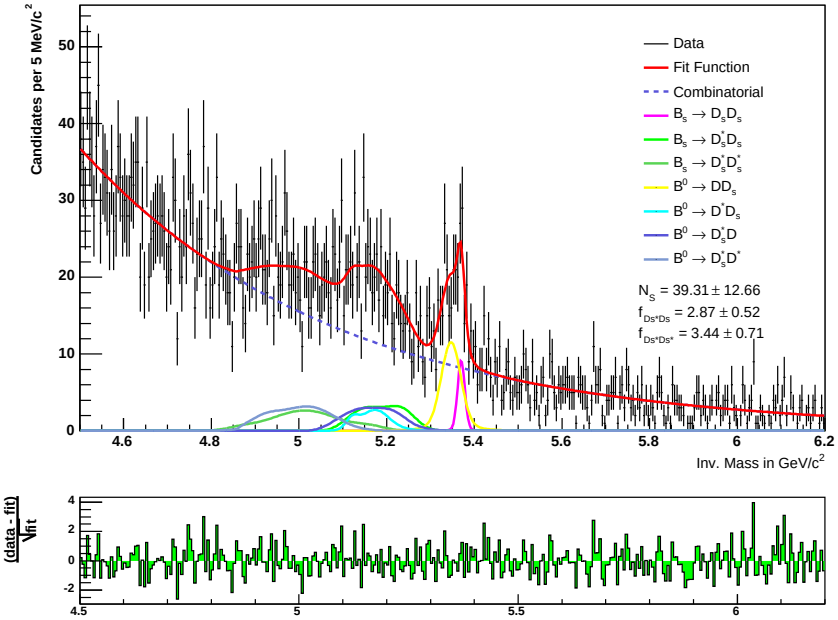


Figure D.2: Result of the simultaneous B_s mass fit for the decay $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to K^*K)$.

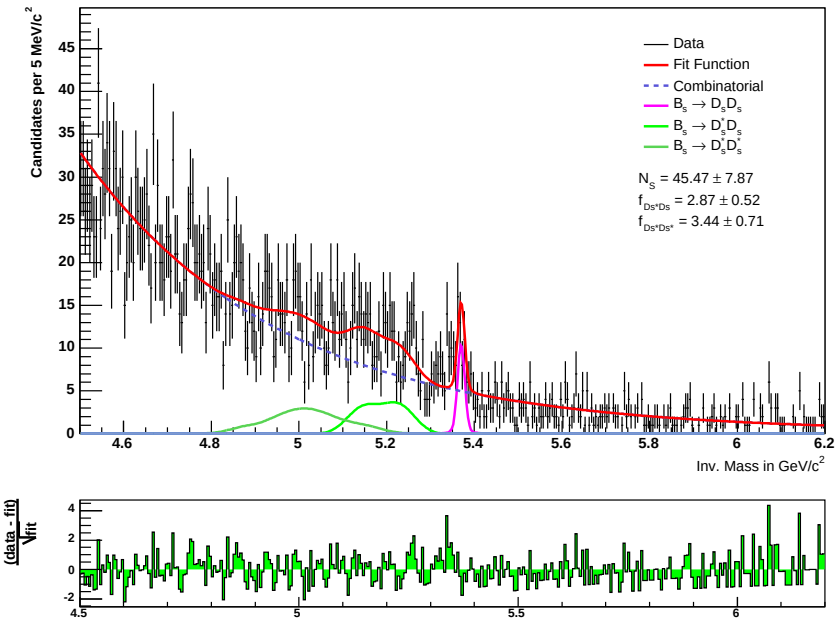


Figure D.3: Result of the simultaneous B_s mass fit for the decay $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to \pi\pi\pi)$.

D.2 Simultaneous Mass Fit for the $f_{D_*D_s^*}$, $f_{D_s^*D}$ and $f_{D_s^*D^*}$ Measurement

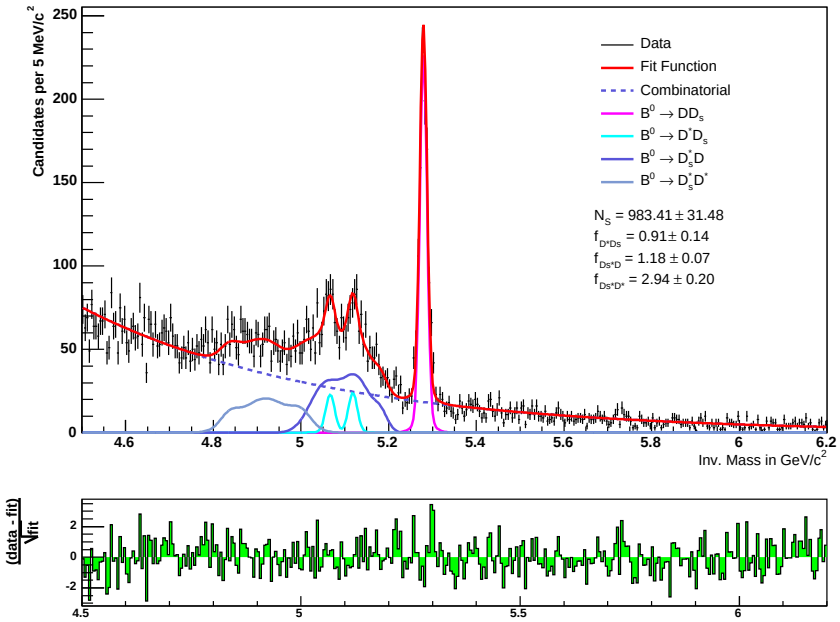


Figure D.4: Result of the simultaneous B^0 mass fit for the decay $B^0 \to D_s\,(\to \phi\pi)\,D\,(\to K\pi\pi)$.

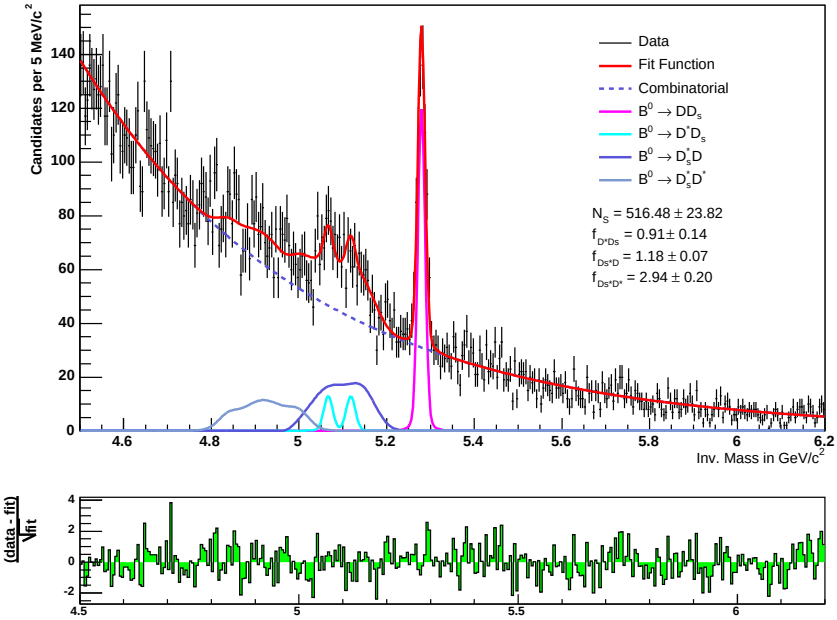


Figure D.5: Result of the simultaneous B^0 mass fit for the decay $B_s \to D_s\,(\to K^*)\,D_s\,(\to \pi\pi\pi)$.

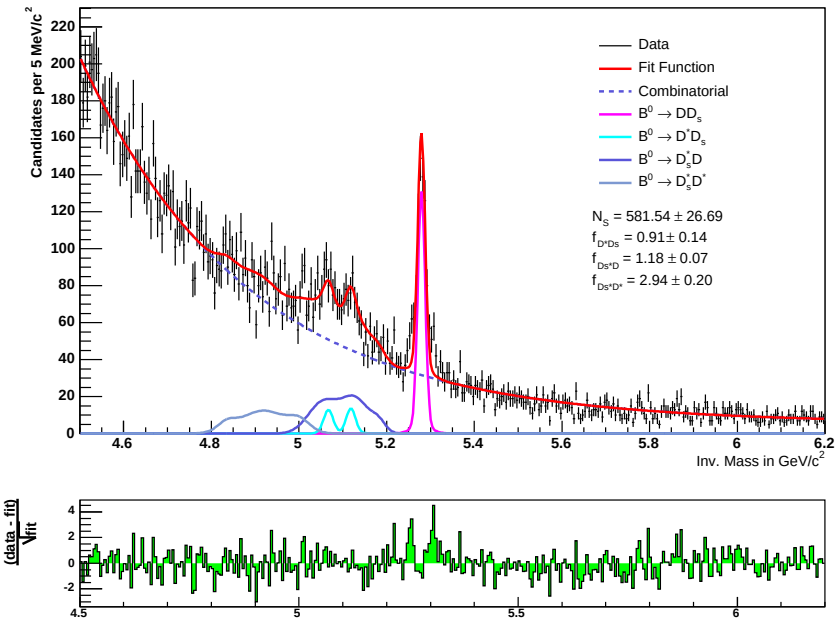
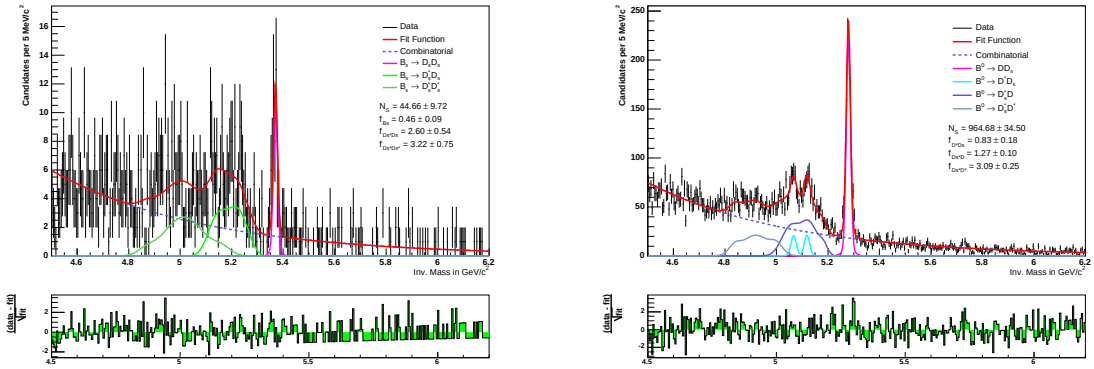


Figure D.6: Result of the simultaneous B^0 mass fit for the decay $B^0 \to D_s\,(\to \pi\pi\pi)\,D\,(\to K\pi\pi)$.

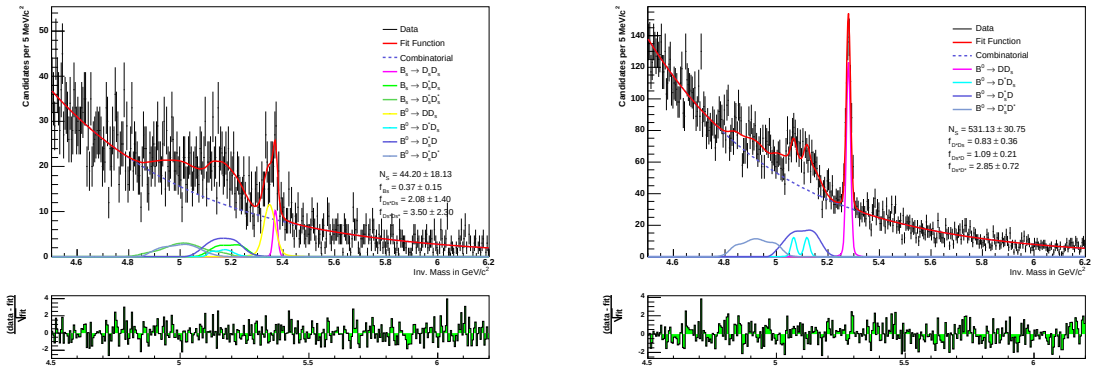
D.3 Simultaneous Mass Fits for the f_{B_s} Measurement

D.3.1 Combining One $B_s \to D_s D_s$ with One $B^0 \to D_s D$ Decay Mode



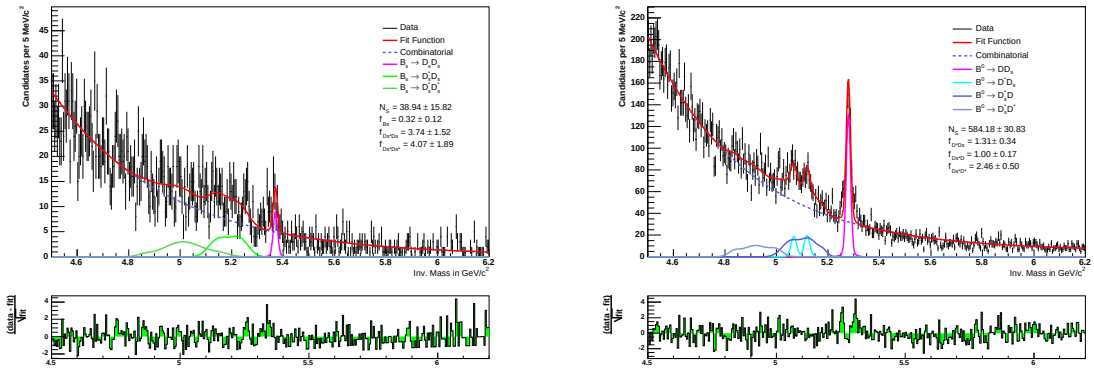
(a) $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to \phi\pi)$ (b) $B^0 \to D_s\,(\to \phi\pi)\,D\,(\to K\pi\pi)$

Figure D.7: Result of the simultaneous fit to the invariant mass spectra of the decays $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to \phi\pi)$ and $B^0 \to D_s\,(\to \phi\pi)\,D\,(\to K\pi\pi)$.



(a) $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to K^*K)$ (b) $B^0 \to D_s\,(\to K^*K)\,D\,(\to K\pi\pi)$

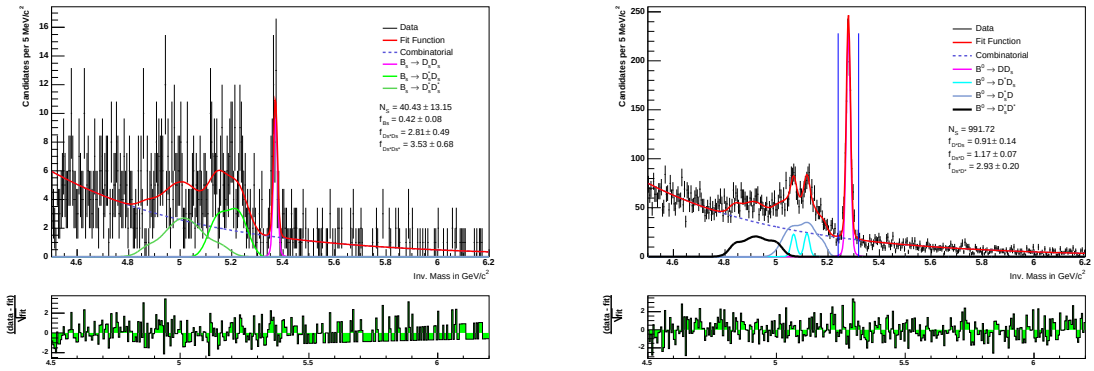
Figure D.8: Result of the simultaneous fit to the invariant mass spectra of the decays $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to K^*K)$ and $B^0 \to D_s\,(\to K^*K)\,D\,(\to K\pi\pi)$.



(a) $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to \pi\pi\pi)$ (b) $B^0 \to D_s\,(\to \pi\pi\pi)\,D\,(\to K\pi\pi)$

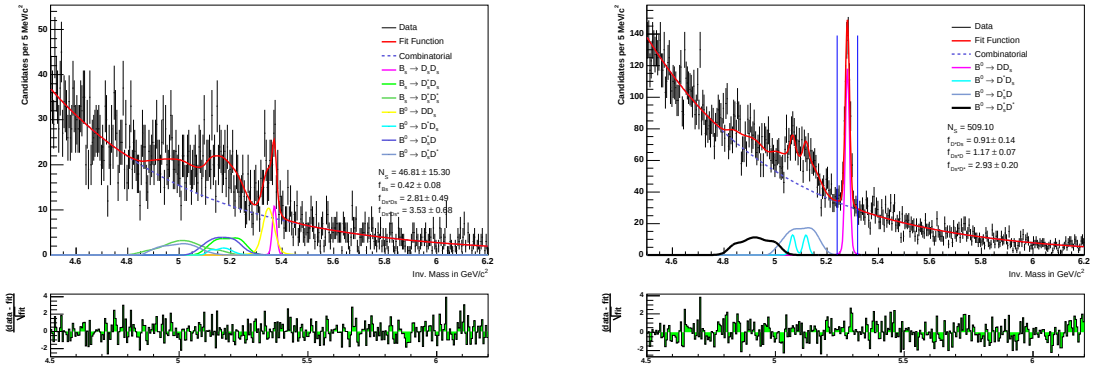
Figure D.9: Result of the simultaneous fit to the invariant mass spectra of the decays $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to \pi\pi\pi)$ and $B^0 \to D_s\,(\to \pi\pi\pi)\,D\,(\to K\pi\pi)$.

D.3.2 Combining all $B_s \to D_s D_s$ and $B^0 \to D_s D$ Decay Modes



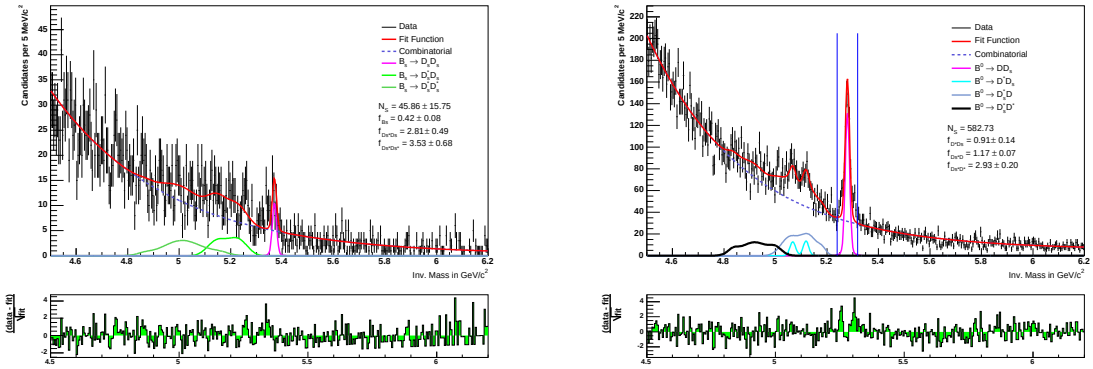
(a) $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to \phi\pi)$ (b) $B^0 \to D_s\,(\to \phi\pi)\,D\,(\to K\pi\pi)$

Figure D.10: Results of the simultaneous fit to the invariant mass spectra of all relevant B_s and B^0 decay modes. Here, the decays $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to \phi\pi)$ (left)and $B^0 \to D_s\,(\to \phi\pi)\,D\,(\to K\pi\pi)$ (right) are shown.



(a) $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to K^*K)$ (b) $B^0 \to D_s\,(\to K^*K)\,D\,(\to K\pi\pi)$

Figure D.11: Results of the simultaneous fit to the invariant mass spectra of all relevant B_s and B^0 decay modes. The decays $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to K^*K)$ (left)and $B^0 \to D_s\,(\to K^*K)\,D\,(\to K\pi\pi)$ (right) are shown.



(a) $B_s \to D_s\,(\to \phi\pi)\,D_s\,(\to \pi\pi\pi)$ (b) $B^0 \to D_s\,(\to \pi\pi\pi)\,D\,(\to K\pi\pi)$

Figure D.12: Results of the simultaneous fit to the invariant mass spectra of all relevant B_s and B^0 decay modes. The decays $B_s \to D_s\,(\to \pi\pi\pi)\,D_s\,(\to K^*K)$ (left)and $B^0 \to D_s\,(\to \pi\pi\pi)\,D\,(\to K\pi\pi)$ (right) are shown.

Bibliography

[1] I. Dunietz, R. Fleischer, U. Nierste. *Phys. Rev. D* **63**, 114015, 2001.

[2] The CDF Collaboration. *Observation of $B_s\bar{B}_s$ Oscillations*, *Phys. Rev. Lett.* 97 242003, 2006. [arXiv:hep-ex/0609040v1]

[3] B. Iyutin, C. Paus. *$\mathcal{B}(B_s^0 \to D_s^- D_s^+)/\mathcal{B}(B^0 \to D\_D_s^+)$ with $355pb^{-1}$ in Run II*, CDF/DOC/BOTTOM/CDFR/8085 Version 4.0, 2006.

[4] W.–M. Yao et al. (Particle Data Group) *J. of Physics* **G 33**, 1 (2006) and 2007 partial update for 2008. `http://pdg.lbl.gov`.

[5] The CDF Collaboration. *Observation of Top Quark Production in $p\bar{p}$ Collisions*, Fermilab-Pub-95/022E, 1995. CDF/PUB/TOP/PUBLIC/3040

[6] The D0 Collaboration. *Observation of the Top Quark*, Phys. Rev. Letters 74 2632, FERMILAB PUB-95/028-E, 1995.

[7] The D0 Collaboration. *Evidence for production of single top quarks and first direct measurement of $|V\_tb|$*, Phys. Rev. Lett. 98, 181802, 2007. [arXiv:hep-ex/0612052]

[8] The CDF Collaboration. *Combination of CDF Single Top Searches with 2.2 fb^{-1} of data*, CDF Note 9251, 2008. `http://www-cdf.fnal.gov/physics/new/top/public_singletop.html`

[9] `http://www-cdfonline.fnal.gov/ops/opshelp/stores/`

[10] The CDF II Collaboration.*The CDF-II Detector: Technical Design Report*, FERMILAB–Pub–96/390–E, 1996.

[11] D. Acosta et al. *A Time–of–Flight Detector in CDF–II*, Nucl. Instrum. Meth., A526:249-299, 2004.

[12] A. Sill et al. *CDF Run II Silicon Tracking Projects*, Nucl. Instrum. Meth. A447, 1, 2000.

[13] T. Affolder at al. *CDF Central Outer Tracker*, FERMILAB–Pub–03/355–E, 2004.

[14] P. Azzi, G. Busetto, P. Gatti, A. Ribon. *Histogram Tracking in the COT*, CDF/DOC/TRACKING/CDFR/5562, 2001.

[15] M. Feindt, S. Menzemer, K. Rinnert. *TrackingKal – A Tracking and Alignment Software for the CDFII Silicon Detector*, CDF/THESIS/TRACKING/PUBLIC/5968, 2002.

[16] Y. Huang, C. Hayes, A. Kotwal. *Inside–Out Tracking*, CDF/DOC/TRACKING/CDFR/6707, 2003.

[17] G. Ascoli et al. *CDF Central Muon Detector*, Nucl. Instrum. Meth., A268:33, 1988.

[18] L. Balka et al. *The CDF Central Electromagnetic Calorimeter*, Nucl. Instrum. Meth. A267:272, 1998.

[19] S.R. Hahn et al. *Calibration Systems for the CDF Central Electromagnetic Calorimeter*, Nucl. Instrumen. Meth. A267:351, 1988.

[20] S. Bertolucci et al. *The CDF Central and Endwall Hadron Calorimeter*, Nucl. Instrum. Meth., A267:301, 1988.

[21] H. Frisch et al. *Conceptual Design of a Deadtimeless Trigger for the CDF Trigger Upgrade*, CDF/DOC/TRIGGER/CDFR/2038, 1994.

[22] M. Milnik. *Measurement of the Lifetime Difference and the CP-Violating Phase in $B_s \to J/\psi\phi$ Decays*, PhD Thesis, IEKP-KA/2007-16, 2007.

[23] P. Mack. *Calibration of New Flavor Tagging Algorithms Using B_s Oscillations*, IEKP-KA/2007-10, 2007.

[24] E.J. Thomson et al. *Online Track Processor for the CDF Upgrade*, IEEE Transactions on Nuclear Science, Vol. 49, No. 3, 2002.

[25] B. Ashmanskas et al. *The CDF Silicon Vertex Trigger*, arXiv:physics/0306169v1, 2003.

[26] S. L. Glashow. *Nucl. Phys.*, 22:579, 1961.

[27] S. Weinberg. *Phys. Rev. Lett.*, 19:1264, 1967.

[28] A. Salam. in *Elementary Particle Theory*, 1968.

[29] D. J. Gross, F. Wilczek. *Phys. Rev.*, D8:3633-3652, 1973.

[30] M Kobayashi, T. Maskawa. *Prog. Theor. Phys.* **49**, 652, 1973.

[31] R. Aleksan, A. Le Yaouanc, L. Oliver. O. Pène, J. C. Raynal, *Phys. Lett.*, **B316** 567, 1993.

[32] M. A. Shifman, M. B. Voloshin. *Sov. J. Nucl. Phys.*, 47 511, 1988.

[33] R. Barate et al. (ALEPH Coll.). *Phys. Lett. B*, B 486 286, 2000.

[34] The D0 Collaboration. *Measurement of Br* $\left[D^{(*)}D^{(*)}\right]$ *and the Lifetime Difference in the* B_s^0 *system*, D0 Note 5651 CONF, 2008.

[35] V. Blobel, E. Lohrmann. *Statistische und numerische Methoden der Datenanalyse*, Teubner Taschenbücher Physik, 1971.

[36] M. Feindt. *A Neural Bayesian Estimator fir Conditional Probability Densities*, arXiv:physics/0402093V1, 2004.

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

[38] C. Paus et al. `http://cdfcodebrowser.fnal.gov/CdfCode/source/BottomMods`.

[39] J. Marriner. *Secondary Vertex Fit With Mass And Pointing Constraints (CTVMFT)*, 1993, CDF/DOC/SECVTX/PUBLIC/1996.

[40] I. Furic, B. Iyutin, C. Paus, *BottomMods – STNtuples for B physics*, 2005.

[41] C. Dörr. *Optimization of the Signal Selection of Exclusively Reconstructed Decays of* B_0 *and* B_s *Mesons at CDF–II*, PhD Thesis, IEKP-KA/2006-6, 2006.

[42] R. Brun, F. Rademakers. *ROOT – An Object Oriented Data Analysis Framework*, Nucl. Instrum. Meth. A389 81-86, 1997,
`http://root.cern.ch/`.

[43] P. Nason, S. Dawson, R. K. Ellis. *Nucl. Phys.* **B303** 607, 1998; *Nucl. Phys.* **B327** 49, 1989.

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

[45] T. Sjostrand, L. Lonnblad, S. Mra. *PYTHIA 6.2: Physics and Manual*, LU TP 01-21, hep-ph/0108264m 2001.

[46] C. Peterson et al. *Phys. Rev.* **D27** 105, 1983.

[47] W. Bell, J.P. Fernandez, L. Flores, F. Wuerthwein, R.J. Tesarek. *User Guide for EvtGen CDF*, CDF Note 5618.

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

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

[50] *The Trigger Simulation Project*, `http://ncdf70.fnal.gov:8001/trigsim/trgsim.html`.

[51] J. Fränkle. *Fitting Procedure for the Determination of the B_s Mixing Frequency*, IEKP-KA/2007-03, 2007.

Danksagung

An erster Stelle möchte ich Herrn Prof. Dr. Michael Feindt dafür danken, dass mir die Möglichkeit eröffnet wurde, am Institut für Experimentelle Kernphysik meine Diplomarbeit über ein sehr interessantes Thema anzufertigen. Mein Dank gilt auch Herrn Prof. Dr. Thomas Müller für die Übernahme des Korreferats.

Eine gute Betreuung ist für das Gelingen einer Diplomarbeit unerlässlich. In diesem Zusammenhang bedanke ich mich herzlich bei Dr. Michal Kreps, Dr. Thomas Kuhr und Prof. Dr. Michael Feindt für die hilfreichen Diskussionen und nützlichen Ratschläge. Für geduldige inhaltliche Erläuterungen und technische Hilfestellungen, die mir den Einstieg in die mir anfangs unbekannten Aspekte der Materie erheblich erleichterten, möchte ich nebst zuvor Genannten auch Andreas Gessler, Dr. Michael Milnik und Jan Morlock danken.

Herzlich bedanken möchte ich mich auch bei allen Kollegen, mit denen ich das Büro teilte, da stets für ein angenehmes, ruhiges und produktives Arbeitsklima gesorgt war. In diesem Kontext gilt mein Dank auch allen anderen Mitgliedern des EKP, insbesondere den mit administrativen Aufgaben betrauten Kollegen.

Bei Dr. Joachim Heuser und abermals Dr. Thomas Kuhr möchte ich mich herzlich für das Korrekturlesen meiner Arbeit und die kritischen aber stets hilfreichen Kommentare und Verbesserungsvorschläge bedanken.

In besonderem Maße bedanke ich mich bei allen, derer Unterstützung für mein Studium, und darüber hinaus, ich mir stets gewiss sein konnte: meinen Eltern Peter und Elisabeth Horn, meinen Geschwistern Judith, Pirmin und Simon Horn, meiner Freundin Virginia und allen Freunden, die die Zeit meines Studiums in jedweder Form bereicherten.