Search for Orbitally Excited B Mesons with the CDF II Detector

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

von der Fakultät für Physik der
Universität Karlsruhe (TH)

genehmigte

DISSERTATION

von

Dipl.-Phys. Andreas Gessler
aus Dettelbach

Tag der mündlichen Prüfung:	17. Juli 2009
Referent:	Prof. Dr. M. Feindt
Korreferent:	Prof. Dr. G. Quast

Zusammenfassung

Einleitung

Auf dem Gebiet der Teilchenphysik wurden in den vergangenen Jahrzehnten große Anstrengungen unternommen, die Struktur der Materie und deren Naturgesetze zu erforschen. Für ein besseres Verständnis dieser Gesetze sind zunehmend größere experimentelle Apparate notwendig, da eine ausreichend hohe Energiedichte erzeugt werden muss. So haben heutige Teilchenbeschleuniger, die Strukturen kleiner als den milliardsten Teil eines Atomkernes auflösen können, mehrere Kilometer Umfang. Die Aufgabe der Teilchenphysik besteht nun hauptsächlich darin, mit Hilfe von Apparaten wie den Teilchenbeschleunigern theoretische Vorhersagen experimentell zu bestätigen, die Zahlenwerte der freien Parameter dieser Theorien zu bestimmen und Hinweise auf neue Physik jenseits der etablierten Theorien zu finden.

Standardmodell der Teilchenphysik

Das Standardmodell der Teilchenphysik beschreibt die Elementarteilchen, aus denen die uns vertraute Materie zusammengesetzt ist. Darüberhinaus beinhaltet es drei fundamentale Wechselwirkungen, die unter diesen Teilchen wirken, wobei die Gravitation als vierte Wechselwirkung in der Allgemeinen Relativitätstheorie beschrieben wird und nicht Teil des Standardmodells ist.

Die elementaren Teilchen sind innerhalb des Standardmodells in Fermionen und Bosonen unterteilt, deren Eigenschaften durch Quantenzahlen festgelegt werden. Beispiele solcher Quantenzahlen sind die elektrische Ladung oder der Spin. Die Fermionen besitzen einen halbzahligen Spin, die Bosonen einen ganzzahligen Spin.

Insgesamt gibt es zwölf elementare Fermionen. Sie sind unterteilt in Quarks und Leptonen und werden typischerweise in drei Familien gruppiert:

$$\begin{array}{llll} \text{Familie} & 1. & 2. & 3. \\ \text{Quarks} & \begin{pmatrix} u \\ d \end{pmatrix} & \begin{pmatrix} c \\ s \end{pmatrix} & \begin{pmatrix} t \\ b \end{pmatrix} \\ \text{Leptonen} & \begin{pmatrix} e \\ \nu_e \end{pmatrix} & \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} & \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} \end{array}$$

Die Teilchen der ersten Familie besitzen die geringsten Massen. Gemäß den Energieerhaltungssätzen zerfallen Teilchen stets in leichtere, sofern der Zerfall keine anderen Erhaltungssätze verletzt. Demzufolge besteht die gewöhnliche Materie, der wir im Alltag begegnen, ausschließlich aus Teilchen der ersten Familie.

Die drei fundamentalen Wechselwirkungen unter den Teilchen heißen elektromagnetische, schwache und starke Wechselwirkung und werden innerhalb des Standardmodells über den Austausch verschiedener Bosonen erklärt, wobei jeder Wechselwirkung ihre charakteristischen Bosonen zugeordnet werden.

Das Boson der elektromagnetischen Wechselwirkung ist das masselose Photon. Durch die fehlende Masse ist die Reichweite dieser Wechselwirkung unbegrenzt.

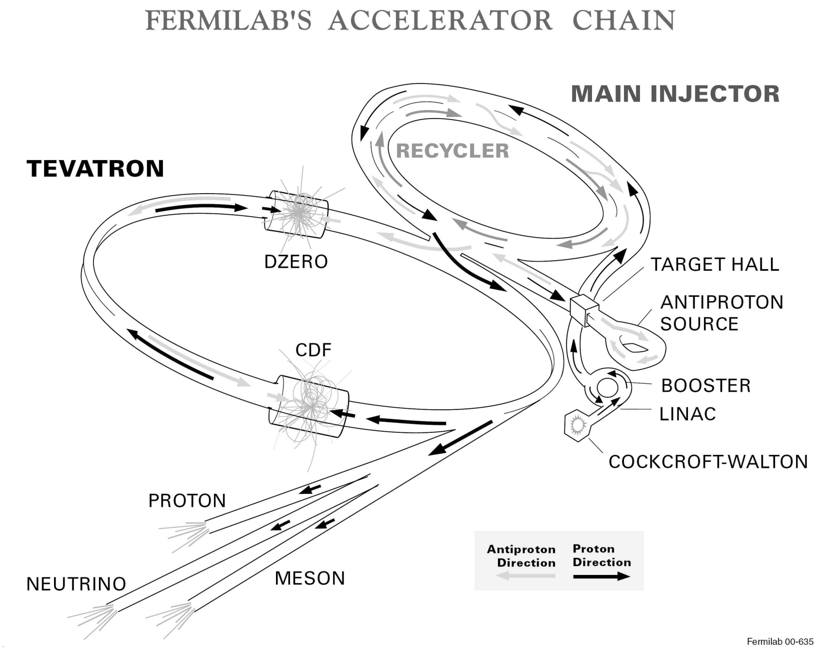
Die schwache Wechselwirkung beruht auf den geladenen $W^{\pm}$-Bosonen und dem ungeladenen Z^0-Boson. Aufgrund der schwachen Wechselwirkung sind Übergänge innerhalb und zwischen den einzelnen Teilchenfamilien möglich.

Die Bosonen der starken Wechselwirkung sind die acht Gluonen, die gebundene Zustände zwischen den Quarks hervorbringen. Bei diesen wird unterschieden zwischen Mesonen, die aus einem Quark-Antiquark-Paar bestehen und Baryonen, die aus drei Quarks bestehen. Da sich Gluonen wechselseitig beeinflussen, können Quarks nicht voneinander getrennt werden beziehungsweise können keine ungebundenen Quarks erzeugt werden. Die hierfür benötigte Energie nimmt mit dem Abstand zu, daher ist es energetisch günstiger, ab einer potentiellen Energie von zwei Quarkmassen ein neues Quark-Antiquark-Paar zu erzeugen.

Beim Erforschen der starken Wechselwirkung ist die Nichtexistenz von freien Quarks ein Nachteil, weshalb hier in Analogie zum Wasserstoffatom die Mesonen untersucht werden, die aus einem schweren und einem leichten Quark bestehen. Für das Verstehen der starken Wechselwirkung spielen diese Mesonen eine ähnlich wichtige Rolle wie das Wasserstoffatom bei der elektromagnetischen Wechselwirkung. Das schwere Quark übernimmt dabei also die Rolle des Protons, das leichte Quark die des Elektrons.

Analog zur Entstehung der Hyperfeinstruktur aufgrund der Spin-Bahn-Kopplung des Elektrons mit dem Proton ergibt die Spin-Bahn-Kopplung des leichten und des schweren Quarks eine Massenaufspaltung der orbital angeregten Zustände der Mesonen.

In der vorliegenden Arbeit werden B^{**}- und B_s^{**}-Mesonen in Zerfällen nach $B^{**} \rightarrow B^{(*)+}\pi^-$ und nach $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ untersucht. Diese Mesonen bestehen jeweils aus einem schweren b-Quark und einem leichten u-Quark beziehungsweise s-Quark. Aufgrund des Bahndrehimpulses von $L = 1$ ergeben sich durch die oben erwähnte Spin-Bahn-Kopplung vier angeregte Zustände: $B^*_{(s)0}$, $B^*_{(s)1}$, $B_{(s)1}$ und $B^*_{(s)2}$. Diese vier Zustände werden unter dem Namen $B^{**}_{(s)}$ zusammengefasst.


FERMILAB'S ACCELERATOR CHAIN

Fermilab 00-635

Abbildung 0.1: Skizze der Beschleuniger des Tevatrons.

Experimenteller Aufbau

Die Daten, die in der vorliegenden Arbeit verwendet werden, wurden mit Hilfe des am Tevatron stehenden CDF-II-Detektors aufgezeichnet. Das Tevatron ist ein Teilchenbeschleuniger des Fermi National Accelerator Laboratory, kurz Fermilab, 50 Kilometer westlich von Chicago, Illinois (USA). In ihm werden Protonen und Antiprotonen gegenläufig beschleunigt und an zwei Punkten zur Kollision gebracht. An diesen beiden Wechselwirkungspunkten stehen die Experimente CDF und DØ.

Das Tevatron hat eine Schwerpunktenergie von $1,96$ TeV/c^2 und um die Protonen und Antiprotonen auf die dafür nötige Energie zu beschleunigen, durchlaufen sie eine Kaskade von Vorbeschleunigern, deren Funktionsschema in Abbildung 0.1 zu sehen ist.

Die Protonen werden aus gasförmigem Wasserstoff gewonnen und mit Hilfe der Vorbeschleuniger auf ihre Endgeschwindigkeit gebracht. Ein Teil wird dazu direkt ins Tevatron eingeleitet, die anderen werden zum Erzeugen der Antiprotonen verwendet. Hierfür wird der Protonenstrahl auf einen Nickelblock geschossen. Die dabei erzeugten Antiprotonen werden gesammelt und ebenfalls in das Tevatron eingeleitet.

Abbildung 0.2 zeigt den CDF-II-Detektor. Dieser ist ein typischer Vielzweckdetektor der Teilchenphysik. Er hat eine zylindersymmetrische Geometrie und ist spiegelsymmetrisch bezüglich des Wechselwirkungspunktes. Der Detektor setzt sich aus

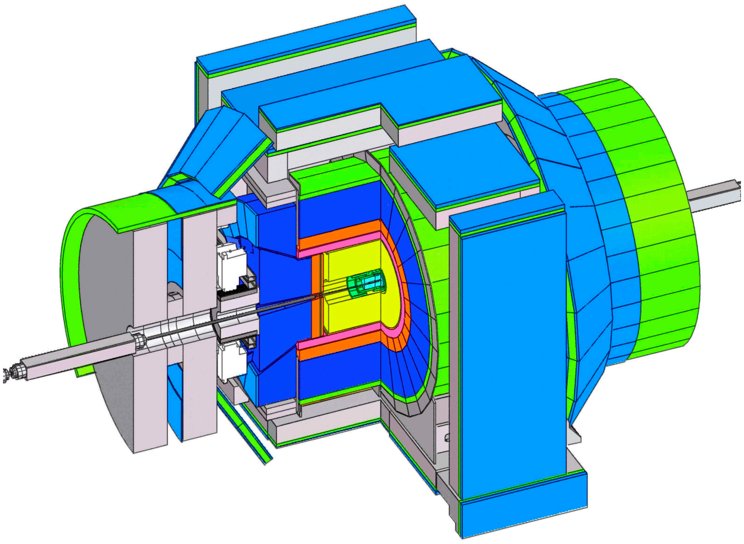


Abbildung 0.2: Skizze des CDF-II-Detektors.

Komponenten zur Spurrekonstruktion und Teilchenidentifikation, sowie Kalorimetern und Myonenkammern zusammen. Die Spurdetektoren sind von einer supraleitenden Spule umgeben, die ein Magnetfeld von $1,4\ \mathrm{T}$ parallel zum Strahlrohr erzeugt. Dadurch werden die Spuren der geladenen Teilchen gekrümmt, womit deren Impuls bestimmt werden kann.

Aufgrund der hohen Kollisionrate der Protonen und Antiprotonen ist es nicht möglich, alle detektierten Ereignisse aufzuzeichnen. Darüberhinaus sind nur wenige dieser Ereignisse überhaupt von physikalischem Interesse. Deswegen verfügt der CDF-II-Detektor über ein komplexes System von Filtern, welches die Vorselektion interessanter Ereignisse nahezu instantan bewerkstelligt.

Die Datensätze, die im Rahmen dieser Arbeit verwendet werden, basieren vollständig auf der Selektion des Two-Track-Triggers und des Di-Muon-Triggers. Jener identifiziert Ereignisse, deren Spuren einen großen Stoßparameter haben, was auf einen großen Abstand zwischen Erzeugungs- und Zerfallsort hindeutet, wie es für B-Mesonen der Fall ist. Der Di-Muon-Trigger selektiert Ereignisse, die mindestens zwei Myonen erhalten, welche einem J/ψ-Zerfall zugeordnet werden können.

Neuronale Netze

In der Regel besteht der weitaus größte Anteil der Ereignisse in den aufgezeichneten Rohdaten aus Untergrundereignissen, die von den Signalereignissen getrennt werden müssen. Zur Klassifizierung der Ereignisse werden innerhalb dieser Arbeit neurona-

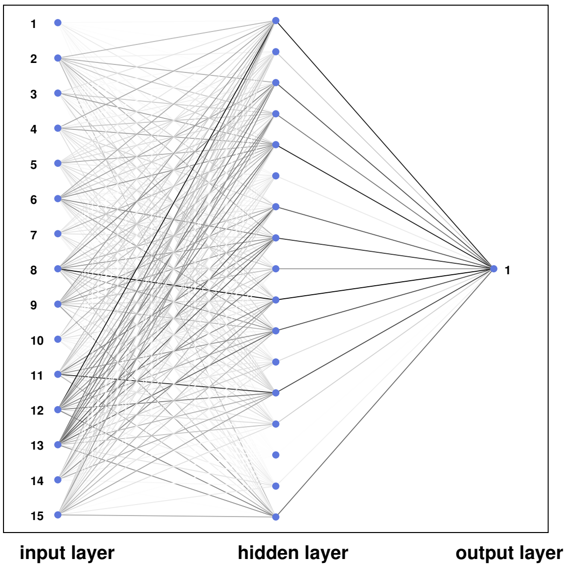


Abbildung 0.3: Topologie eines dreilagigen, vorwärtsgekoppelten neuronalen Netzes. Die Stärke der Verbindunslinien zwischen den Neuronen ist proportional zu deren Signifikanz für das Netzwerk.

le Netze eingesetzt. Diese besitzen gegenüber rein schnittbasierten Analysen den Vorteil, dass sie Korrelationen zwischen den verwendeten Variablen berücksichtigen, indem sie die Informationen aller Selektionsgrößen auf eine einzige Ausgangsvariable abbilden. Für die endgültige Selektion wird ein Schnitt auf diese eine Ausgangsvariable angewendet.

In dieser Arbeit werden ausschließlich neuronale Netze aus dem NeuroBayes®-Paket verwendet, das an der Universität Karlsruhe entwickelt wurde. Es handelt sich hierbei um ein dreilagiges, vorwärtsgekoppeltes Netzwerk, das über einen zusätzlichen Algorithmus verfügt, der die Daten in einem Arbeitsschritt vor dem eigentlichen neuronalen Netz aufbereitet und die Signalsignifikanz entscheidend verbessert.

Neuronale Netze bestehen im Allgemeinen aus Neuronen, die hier in drei Schichten angeordnet sind. Abbildung 0.3 zeigt exemplarisch die Topologie eines solchen Netzes. In der Eingabeschicht wird jeder verwendeten Variablen ein Neuron zugeordnet, das wiederum mit den Neuronen der mittleren Schicht verbunden ist. Die Ausgabeschicht besteht aus lediglich einem Neuron, das für eine binäre Klassifizierung ausreichend ist. Bei einem vorwärtsgekoppelten neuronalen Netz sind die Neuronen einer Schicht ausschließlich mit den Neuronen der nachfolgenden Schicht verbunden.

Bevor ein neuronales Netz zur Trennung von Signal- und Untergrundereignissen eingesetzt werden kann, muss es anhand von Musterdaten, die sowohl signalartige als auch untergrundartige Ereignisse enthalten, trainiert werden. Hierbei werden simulierte Ereignisse als Signalmuster verwendet. Da bisherige Modelle die komplexen Zusammenhänge des Untergrundes, wie er an Teilchenbeschleunigern auftritt, nicht mit der benötigten Genauigkeit beschreiben können, werden als Muster für Untergrundereignisse reale Daten aus einem Massenbereich genommen, in dem keine oder äußerst wenige Signalereignisse zu erwarten sind.

Rekonstruktion und Selektion der Signalereignisse

In dieser Arbeit werden B^{**}- und B_s^{**}-Mesonen in den Zerfällen $B^{**} \to B^{(*)+}\pi^-$ und $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ untersucht, die unter Verwendung verschiedener Zerfallskanäle der B^+- und B_s-Mesonen rekonstruiert werden. Für die Rekonstruktion der B^{**}-Mesonen werden folgende Zerfällskanäle verwendet:

$$\begin{array}{lll}
B^{**} \to B^{(*)+}\pi^- & B^+ \to J/\psi K^+ & J/\psi \to \mu^+\mu^- \\
 & B^+ \to \bar{D}^0\pi^+ & \bar{D}^0 \to K^+\pi^- \\
 & B^+ \to \bar{D}^0\,3\pi^\pm & \bar{D}^0 \to K^+\pi^-
\end{array}$$

Für die Rekonstruktion der B_s^{**}-Mesonen sind es folgende Zerfällskanäle:

$$\begin{array}{lll}
B_s^{**} \to B_s^{(*)}\pi^+\pi^- & B_s \to D_s^-\pi^+ & D_s^- \to \bar{K}^*K^- \\
 & & D_s^- \to \phi\pi^- \\
 & & D_s^- \to 3\pi^\pm \\
 & B_s \to D_s^- 3\pi^\pm & D_s^- \to \bar{K}^*K^- \\
 & & D_s^- \to \phi\pi^- \\
 & & D_s^- \to 3\pi^\pm
\end{array}$$

Sowohl die B^*- als auch die B_s^*-Mesonen gehen unter Aussendung eines Photons in den Grundzustand $B_{(s)}^* \to B_{(s)}\gamma$ über, wobei das dabei erzeugte Photon mit dem CDF-II-Detektor nicht direkt nachweisbar ist, was zu einer Verschiebung der $B_{(s)}^*$-Massen führt.

Die Rekonstruktion der Mesonen beginnt am Ende der jeweiligen Zerfallskette bei den stabilen Teilchen. In diesem Zusammenhang bedeutet stabil, dass diese Teilchen eine ausreichend lange Lebensdauer besitzen, um den Detektor vollständig passieren zu können. Ausgehend von diesen Teilchen werden zunächst die D-Mesonen beziehungsweise die J/ψ-Mesonen rekonstruiert, indem die Viererimpulse der Zerfallsprodukte addiert werden. Anschließend werden diesen Kandidaten weitere Spuren durch Addition der Viererimpulse hinzugefügt, um so die B-Mesonen und schließlich die $B_{(s)}^{**}$-Mesonen zu rekonstruieren.

Die Selektion der Signalkandidaten wird in zwei Schritten durchgeführt. Im ersten werden neuronale Netze zur Selektion der B^+- und B_s-Kandidaten eingesetzt. Diese

Netze werden mit Hilfe von simulierten Signal- und realen Untergrundereignissen, die dem oberen Massenseitenband entnommen sind, trainiert. Als Eingabevariablen werden Größen verwendet, die für die Zerfallstopologie charakteristisch sind. Dazu gehören zum einen kinematische Variablen wie Zerfallslänge, Zerfallswinkel und Stoßparameter und zum anderen Güteparameter, die Aussagen über die Qualität der gefundenen Spuren und deren Konsistenz machen. Im darauffolgenden zweiten Schritt werden neuronale Netze zur Selektion der B^{**}- und B_s^{**}-Kandidaten trainiert. Die Klassifizierungsinformation der neuronalen Netze zur Selektion der B-Mesonen ist dabei eine zusätzliche Eingangsvariable zu den oben bereits erwähnten zerfallstypischen Größen.

Zur endgültigen Selektion der Signalkandidaten wird ein Schnitt auf die Ausgangsgröße des zweiten Netzes gemacht. Dieser Schnitt wird so gewählt, dass er die Signalsignifikanz maximiert, die sich aus der Zahl der Signal- und Untergrundereignisse innerhalb des Signalbereiches berechnen lässt.

Bestimmung der B^{**}-Eigenschaften

Zur Messung der Masse der schmalen Zustände des B^{**}-Mesons wird der Verlauf der Q-Werteverteilung parametrisiert. Der Q-Wert stellt die Massendifferenz zwischen dem Meson und dessen Zerfallsprodukten dar: $Q = m(B^{**}) - m(B^+) - m_\pi$. Die Likelihood-Funktion zur Beschreibung der Q-Verteilung beinhaltet hierbei jeweils Komponenten zur Modellierung des Signal- und des Untergrundanteils. Die Signalfunktion umfasst die drei Zerfälle $B_2^* \to B\pi$, $B_2^* \to B^*\pi$ und $B_1 \to B^*\pi$. Die Funktion jedes einzelnen Signals setzt sich aus der Faltung einer Breit-Wigner-Funktion mit dem Modell der Detektorauflösung zusammen. Die Parametrisierung der Untergrundverteilung berücksichtigt zwei verschiedene Untergrundursachen. Zum einen umfasst der Untergrund B_s^{**}-Mesonen aus dem ähnlichen Zerfall $B_s^{**} \to B^+K^-$, bei denen zur Rekonstruktion aufgrund mangelnder Eindeutigkeit die Masse des Pions statt der des Kaons verwendet wurde. Zum anderen besteht der Untergrund aus vermeintlichen Signalereignissen, die aufgrund der Vielzahl an verwendeten Spuren und deren zahlreichen Kombinationsmöglichkeiten innerhalb des Rekonstruktionsprozesses entstehen.

Durch die Minimierung der Likelihood-Funktion mit Hilfe der Maximum-Likelihood-Methode werden deren Parameter bestimmt. Abbildung 0.4 zeigt das Ergebnis dieser Minimierung, bei dem ein deutliches Signal zu erkennen ist. Aus dieser Minimierung ergeben sich die Q-Werte für die schmalen Zustände der B^{**}-Mesonen und die Breite des B_2^*. Über die Weltmittelwerte [2] für die Massen des B^+ und des geladenen Pions werden die absoluten Massen der schmalen Zustände berechnet. Die Unsicherheiten der Weltmittelwerte gehen dabei quadratisch in den systematischen

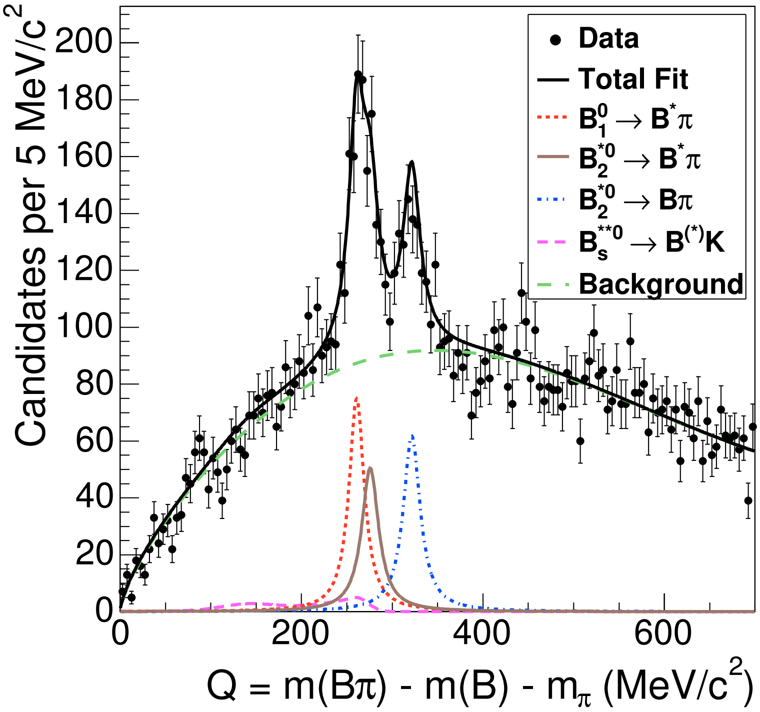


Abbildung 0.4: Parametrisierung der B^{**} Q-Verteilung. Die Daten beinhalten den Zerfall $B^{**} \to B^{(*)+}\pi^-$ mit den Zerfallskanälen $B^+ \to J/\psi K^+$, $B^+ \to \bar{D}^0\pi^+$ und $B^+ \to \bar{D}^0\,3\pi^\pm$.

Fehler ein. Die gemessenen Signalparameter sind:

$$\begin{aligned} m(B_2^*) &= 5740.2^{+1.7}_{-1.8}\,(\text{stat})\ {}^{+0.9}_{-0.8}\,(\text{syst})\ \text{MeV/c}^2, \\ m(B_1) &= 5725.3^{+1.6}_{-2.2}\,(\text{stat})\ {}^{+1.4}_{-1.5}\,(\text{syst})\ \text{MeV/c}^2, \\ \Gamma(B_2^*) &= 22.7^{+3.8}_{-3.2}\,(\text{stat})\ {}^{+3.2}_{-10.2}\,(\text{syst})\ \text{MeV/c}^2. \end{aligned}$$

Bestimmung der B_s^{**}-Eigenschaften

Zur Bestimmung des Verzweigungsverhältnisses der schmalen Zustände der B_s^{**}-Mesonen wird im Rahmen dieser Arbeit die Maximum-Likelihood-Methode verwendet. Die Likelihood-Funktion zur Beschreibung der Q-Verteilung der B_s^{**}-Mesonen enthält jeweils eine Komponente für den Signal- und den Untergrundanteil. Das Signal wird mit Hilfe einer Gaußverteilung parametrisiert, deren Breite auf den Wert der Detektorauflösung festgesetzt wird. Da der Untergrund hauptsächlich kombina-

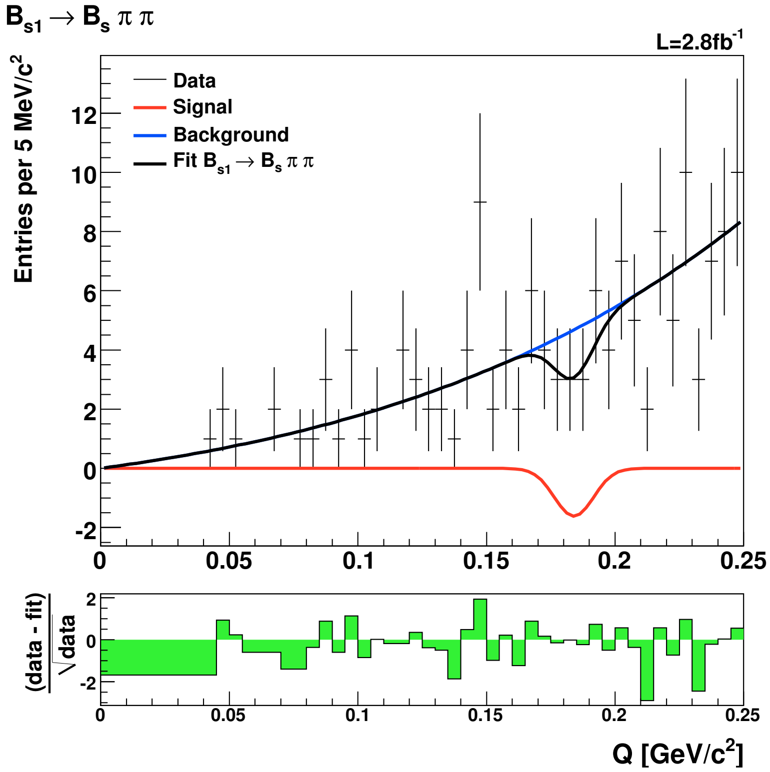


Abbildung 0.5: Parametrisierung der Q-Verteilung des B_{s1}-Zustandes in Zerfällen nach $B_s^{**} \to B_s\pi^+\pi^-$.

torischen Ursprungs ist, wird dieser mittels einer Exponentialverteilung beschrieben, die mit einem linearen Term multipliziert wird, um zu gewährleisten, dass der Untergrundanteil im Ursprung verschwindet.

Zur Bestimmung der Signalparameter wird die Likelihood-Funktion mit Hilfe der Maximum-Likelihood-Methode minimiert. Für die beiden schmalen Zustände B_{s1} und B_{s2}^* werden getrennte Minimierungen durchgeführt deren Massen aufgrund einer früheren CDF-Messung bekannt sind [1]. Bei der Durchführung der Minimierung werden die Mittelwerte der Gaußverteilungen auf diese Werte festgesetzt. Die Abbildungen 0.5 und 0.6 zeigen die Minimierungen der beiden Parametrisierungen. Da für keinen der schmalen Zustände ein signifikantes Signal zu beobachten ist, wird ein Limit auf das Verzweigungsverhältnis dieser beiden Zustände berechnet.

Die Berechnung des Limits beruht auf dem Bayesschen Ansatz unter Annahme einer gleichförmig verteilten A-Priori-Wahrscheinlichkeit. Da die Produktionsraten zur Erzeugung der B_s^{**}-Mesonen innerhalb des CDF-Experiments nicht bekannt sind, wird das Limit relativ bezüglich des bereits beobachteten Zerfalls $B_s^{**} \to B^+K^-$ gemessen. Das obere Limit ist durch das Verzweigungsverhältnis bestimmt, für welches

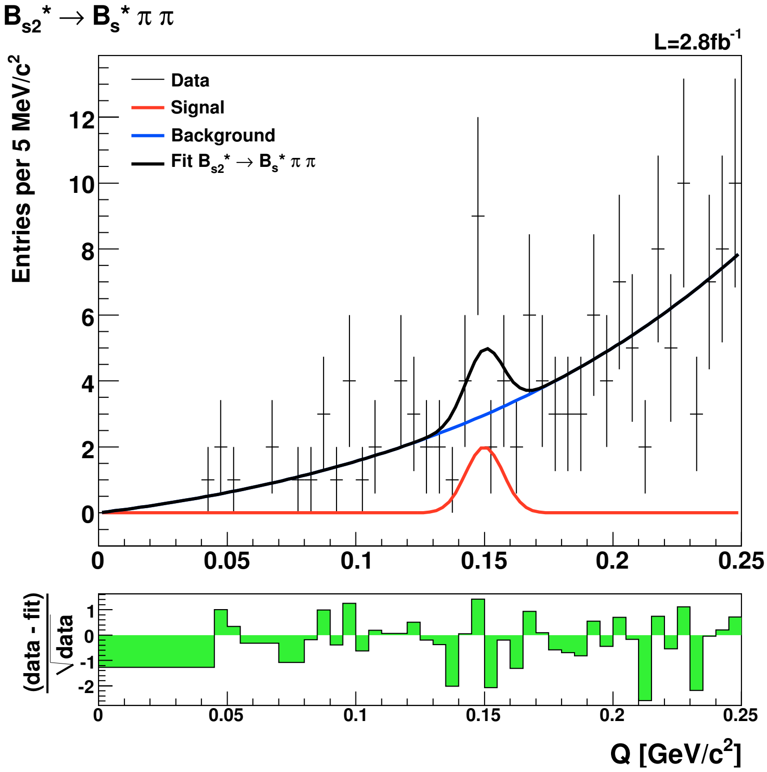


Abbildung 0.6: Parametrisierung der Q-Verteilung des B_{s2}^*-Zustandes in Zerfällen nach $B_s^{**} \to B_s * \pi^+\pi^-$.

das Integral der A-Posteriori-Wahrscheinlichkeit den Wert 0.95 annimmt. Für einen Vertrauensbereich von 95% werden für die oberen Grenzen folgende Werte gemessen:

$$\frac{\mathrm{BR}(B_{s1} \to B_s\pi\pi)}{\mathrm{BR}(B_{s1} \to B_s\pi\pi) + \mathrm{BR}(B_{s1} \to BK)} < 0.54,$$

$$\frac{\mathrm{BR}(B_{s2}^* \to B_s\pi\pi)}{\mathrm{BR}(B_{s2}^* \to B_s\pi\pi) + \mathrm{BR}(B_{s2}^* \to BK)} < 0.53.$$

Die hohen Werte dieser Limits sind dadurch bedingt, dass verschiedene Annahmen, die innerhalb der Minimierungsfunktion getroffen werden, bereits mit großen Unsicherheiten behaftet sind. Dazu zählen zum Beispiel die Verzweigungsverhältnisse der B_s-Zerfälle sowie die Anzahl der Signalereignisse des Referenzkanals $B_s^{**} \to B^+K^-$, die jeweils relative Unsicherheiten im Bereich von 30 bis 40 Prozent haben.

Resümee

In der vorliegenden Arbeit wurden B^{**}- und B_s^{**}-Mesonen in den Zerfällen $B^{**} \to B^{(*)+}\pi^-$ und $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ untersucht. Hierbei wurden verschiedene Zerfallskanäle der B^+- und B_s-Mesonen berücksichtigt, um die B^{**}- und B_s^{**}-Mesonen zu rekonstruieren.

Die Methode zur Selektion der Signalereignisse wurde mit Hilfe von neuronalen Netzen optimiert. Diese besitzen gegenüber rein schnittbasierten Selektionen den Vorteil, dass sie Korrelationen unter den verwendeten Variablen berücksichtigen, und somit zu einem verbesserten Verhältnis von Signal zu Untergrund führen.

Durch die Maximum-Likelihood-Methode wurden die Massen der schmalen Zustände des B^{**}-Mesons sowie die Breite des B_2^*-Zustandes gemessen. Es handelt sich dabei um die bisher präziseste Messung der B_1- und B_2^*-Massen und die erste Messung der B_2^*-Breite. Die Genauigkeit der gemessenen Werte ist durch den statistischen Fehler begrenzt.

Für die schmalen Zustände des B_s^{**}-Mesons wurde kein signifikantes Signal beobachtet. Aus diesem Grund wurde gemäß dem Bayesschen Ansatz ein oberes Limit auf die Verzweigungsverhältnisse der schmalen B_{s1}- und B_{s2}^*-Zustände berechnet.

Die Analysen der B^{**}- und B_s^{**}-Mesonen leisten einen weiteren Beitrag zum besseren Verständnis des Standardmodells. Die erneute Durchführung der beiden Messungen mit größeren Datensätzen wird zur weiteren Reduzierung des statistischen Fehlers führen und könnte die Beobachtung eines Signals im Zerfall $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ möglich machen. Mit einer ausreichend verbesserten Auflösung des Experiments wäre es außerdem möglich, die Signale der drei B^{**}-Zerfälle $B_1 \to B^*\pi$, $B_2^* \to B^*\pi$ und $B_2^* \to B\pi$ zu unterscheiden.

In dieser Arbeit wurden die schmalen Zustände der B^{**}- und B_s^{**}-Mesonen untersucht. Darüberhinaus ist es für ein vollständiges Verständnis der orbital angeregten B- und B_s-Mesonen unerlässlich, die breiten Zustände ebenfalls zu untersuchen, die jedoch mit den gängigen Analysestrategien äußerst schwer vom kombinatorischen Untergrund zu unterscheiden sind.

Contents

List of Figures

List of Tables

1 Introduction

Over the past decades, physicists put much effort into getting a deeper knowledge about the structure of matter and the forces forming it. On the way of getting a better understanding of nature it is necessary to build up larger instruments operating at high energy densities to dig deeper into the substructure of matter. Today, the particle accelerators used to resolve structures smaller than a billionth of a nucleus have circumferences of several kilometres.

The elementary particles forming the matter surrounding us are explained by the standard model of particle physics. The standard model also describes the fundamental interactions amongst these particles, except for gravity which is treated by general relativity. According to the current state of knowledge there are three families of elementary particles grouped into quarks and leptons. On the basis of the strong interaction compound particles are composed of quarks being classified into baryons and mesons.

The standard model is a powerful theory being capable of predicting many parameters and properties. However, the standard model has several intrinsic free parameters whose values cannot be derived theoretically. It is the aim of particle physics to confirm the standard model predictions, to determine the free parameters experimentally and to find evidence for new physics beyond the standard model.

Quantum chromodynamics is the theory describing the above mentioned strong interaction. Quarks occur only in bound states and cannot be observed as unbound free particles according to quantum chromodynamics. This phenomenon is a major drawback in studying quantum chromodynamics. Thus, a good way of doing research in the field of quantum chromodynamics is studying mesons consisting of heavy and light quarks in analogy to the hydrogen atom in quantum electrodynamics. Indeed, heavy-light quark mesons play a similarly important role in understanding quantum chromodynamics as the hydrogen atom did for quantum electrodynamics. In this analogy, the heavy quark takes the role of the nucleus and the light quark plays the part of the electron.

In quantum electrodynamics the hyperfine splitting among members of the same multiplets arises from the spin-orbit coupling of the nucleus and the electron. In a similar way mass splitting emerges among orbitally excited heavy-light mesons based on spin-orbit coupling.

The analysis presented in this thesis performs a search for B^{**} and B_s^{**} mesons in decays of $B^{**} \to B^{(*)+}\pi^-$ and $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$, respectively. B^{**} and B_s^{**} are orbitally excited ($L = 1$) mesons consisting of a heavy b quark and light u or s quark. In doing so, three exclusive decay modes of the B^+ meson are used as well as six exclusive decay modes of the B_s. Exclusive refers the fact that all decay particles within a distinct decay channel are fully reconstructed. The decay width of the excited states depends on the transition mechanism. Consequently, the two hyperfine multiplets have either broad or narrow decay width. The narrow B_s^{**} states have been already seen in decays of $B_s^{**} \to B^+K^-$ [1]. But if the broad states of the B_s^{**} mesons are below the threshold for decays into BK the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ is a good possibility of observing them.

The data used in this thesis was collected with the CDF II detector at the Tevatron starting in February 2002. The data samples correspond to an integrated luminosity of 1.7 fb^{-1} and 2.8 fb^{-1} for the B^{**} and B_s^{**} sample, respectively. The Tevatron is a proton antiproton collider located at the Fermi National Accelerator Laboratory which is situated near to the west of Chicago, Illinois. It runs at a center of mass energy of 1.96 TeV and is currently the world's largest operating particle accelerator. The particles collide at two interaction points where the CDF II and DØ experiments are located.

The selection of the signal candidates is based on neural networks. Neural networks have the advantage of taking into account correlations among the input variables in contrast to a purely cut based selection. The selection is optimised to maximise the signal significance being generally defined as a function of the number of signal and background events.

From the unbinned maximum likelihood fit to the Q value distribution the masses of the narrow B^{**} states are determined as well as the width of the B_2^*. The Q value is the mass difference between a particle and its decay particles. For decays of $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ no significant signal is observed. Therefore, a Bayesian limit is set on the branching fractions of the narrow B_s^{**} states.

2 Theoretical Overview

2.1 The Standard Model

The standard model is the comprehensive underlying theory of particle physics. From today's point of view, it is the most thoroughly tested theory in physics so far but despite its great success it is not capable of answering all the open questions in the field of particle physics.

The standard model classifies the elementary particles into fermions and bosons. The interactions among them are based on three fundamental forces, the electromagnetic, strong and weak force. Gravitation, the fourth known force in nature is not incorporated in the standard model and is treated by general relativity. The properties of the elementary particles are described by quantum numbers. The electromagnetic charge or the spin of a particle are examples of such quantum numbers. Each of the particles in the standard model has an associated antiparticle whose quantum numbers are derived by applying the charge conjugation transformation meaning that all charge-like quantum numbers have opposite signs.

Fermions are half-integer spin particles comprising quarks and leptons. The twelve fundamental types of fermions are typically grouped into three families:

Family	1.	2.	3.
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$
Leptons	$\begin{pmatrix} e \\ \nu_e \end{pmatrix}$	$\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}$	$\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$

The laws of nature postulate particles decaying into lighter particles unless the decay is forbidden by conservation laws otherwise. Therefore, the ordinary matter we meet in everyday life is formed by particles of the first family only as they have the smallest masses. Particles of higher families are produced solely in high-energy interactions as it is the case at the Tevatron collider, for instance. Moreover, the six quarks occur only in bound states of the strong force and not as free particles except for the t quark which is not long-lived enough to form bound states. These bound states are called hadrons and they consist of either three quarks, called baryons, or

Category	Name	Symbol	Charge [e]	Mass [MeV/c^2]
Quark	up quark	u	$+\frac{2}{3}$	1.5 to 3.3
	down quark	d	$-\frac{1}{3}$	3.5 to 6.0
	charm quark	c	$+\frac{2}{3}$	$(1.27^{+0.07}_{-0.11}) \cdot 10^3$
	strange quark	s	$-\frac{1}{3}$	104^{+26}_{-34}
	top quark	t	$+\frac{2}{3}$	$(171.2 \pm 2.1) \cdot 10^3$
	bottom quark	b	$-\frac{1}{3}$	$4.20^{+0.17}_{-0.07} \cdot 10^3$
Lepton	electron	e	-1	0.51
	electron neutrino	ν_e	0	$\leq 2 \cdot 10^{-6}$
	muon	μ	-1	105.66
	muon neutrino	ν_μ	0	≤ 0.19
	tau lepton	τ	-1	1776.84 ± 0.17
	tau neutrino	ν_τ	0	≤ 18.2

Table 2.1: Overview of the elementary fermions of the standard model with their electric charges and masses [2].

Interaction	Name	Symbol	Charge [e]	Mass [GeV/c^2]
Electromagnetic	Photon	γ	0	0
Weak	W boson	$W^\pm$	± 1	80.398 ± 0.025
Weak	Z boson	Z^0	0	91.1876 ± 0.0021
Strong	Gluon	g	0	0

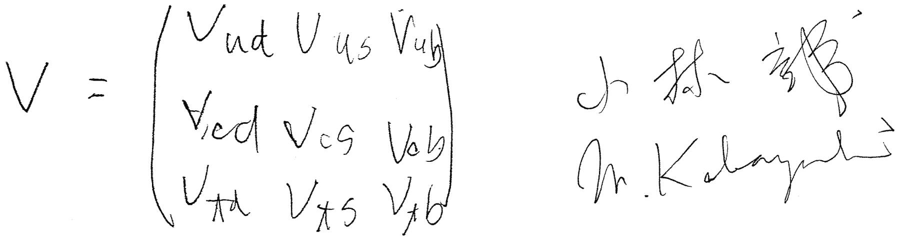
Table 2.2: The elementary gauge bosons of the standard model with their electric charges and masses [2].

of quark-antiquark combinations, called mesons. Table 2.1 gives a summary of the elementary fermions.

The interactions between the particles in the standard model occur by the exchange of bosons [3]. Bosons are particles with integer spin and each of the fundamental forces has its distinct bosons which are listed in table 2.2. Besides that, each force has an associated charge and only particles carrying the specific charge interact with that force. Mathematically, the particles of the standard model are described by quantum field theories and the interactions among them are described by gauge theories.

The theory describing the electromagnetic force is quantum electrodynamics (QED). Its gauge boson is the massless and chargeless photon. The nonexistent mass of the photon causes the electromagnetic force to be of infinite range. All elementary fermions except of the neutrinos carry electromagnetic charge, as well as the $W^\pm$ bosons.

The strong force is described by quantum chromodynamics (QCD) having eight different exchange bosons. These so-called gluons are massless particles differing only



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

Figure 2.1: Autograph of the Cabibbo-Kobayashi-Maskawa matrix written by Makoto Kobayashi [5].

in the colour quantum number. Colour is the term for the charge of the strong force having the possible values red, green and blue. Quarks are the only fermions having colour charge. The fact that gluons carry colour charge themselves yields a potential of QCD being very different from that of QED resulting in a short range force. The strong force forms bound states among the quarks and is therefore the interaction in question for the orbitally excited B mesons studied in this analysis. For this reason, quantum chromodynamics will be discussed more in detail in section 2.2.1.

The weak interaction is very different to the electromagnetic and strong force, particularly at low energies. Firstly, it cannot form any bound states and secondly, it is the only force being able to change the flavour, that is, the quark type by the charged current interaction of the $W^{\pm}$ boson. Hence, no transition would be possible between the three different fermion families without the weak interaction. The third boson of the weak force is the neutral Z^0 boson. In terms of the weak interaction the eigenstates of the quarks differ from the mass eigenstates which are those in the QED and QCD Lagrangian function. The transformation between these different states is described by the Cabibbo-Kobayashi-Maskawa matrix (CKM matrix) [4, 6] seen in figure 2.1. The entries in the CKM matrix are complex numbers constrained by the unitarity of the matrix and by convention to actually four free parameters commonly interpreted as three angles and one complex phase. The determination of these parameters is a major field of research in particle physics at present. All fermions including the neutrinos interact weakly.

Despite the fact that the electromagnetic and the weak force seem to be of a very different nature especially at low energies, at high energies they can be unified on the basis of the formalism of the electroweak theory, developed by Glashow, Salam and Weinberg [7, 8, 9]. Thereby, the $W^{\pm}$ bosons form a triplet of the newly introduced weak isospin together with the newly introduced W^0. The Z^0 and the photon are linear combinations of the W^0 and an additional new singlet, the B^0. The splitting up of the electroweak force into the weak and electromagnetic force below the unification energy is explained by a postulated symmetry breaking mechanism. The Higgs mechanism is the most popular explanation within this context yielding the last so far unobserved elementary particle of the standard model, the Higgs boson [10, 11]. Finding the Higgs particle is another major field of research of particle physics where great efforts are made.

According to the current understanding of the standard model, it might also be possible to merge all three forces into one single theory at very high energies of about 10^{16} GeV [3]. At such high energies transitions between quarks and leptons are possible. This approach of unification is called the grand unification theory (GUT).

2.2 Bound Quark States

In the standard model of elementary particle physics, bound states among the six different quark types, so-called hadrons, are possible. These states are formed on the basis of the strong interaction described by the formalism of quantum chromodynamics. Among these hadrons, two different types of quark combinations have been observed, so far: mesons and baryons.

Mesons are hadrons consisting of a quark and an antiquark. Hence, the total spin S of the meson can either be $S = 0$ or $S = 1$. Let L be the orbital angular momentum between quarks, the total angular momentum J of the meson is

$$J = L \oplus S \tag{2.1}$$

whereas the parity P and the charge conjugation C of a meson consisting of quarks of the same flavour are defined as

$$P = (-1)^{L+1} \quad \text{and} \quad C = (-1)^{J}. \tag{2.2}$$

Hadrons comprising three quarks are called baryons. In contrast to the mesons, the parity of the baryons is defined as

$$P = (-1)^{L} \tag{2.3}$$

due to the convention of the proton parity to be +1.

The orbitally excited mesons being studied in this analysis are mesons consisting of a light and a heavy quark having an angular momentum of $L = 1$. The quantum numbers of their excited states can be derived in the limit of quantum chromodynamics where the mass of the heavy quark is considered being infinite.

2.2.1 Quantum Chromodynamics

Quantum chromodynamics is the gauge field theory describing the strong interaction in the standard model of elementary particles. The associated charge of the strong force is called colour.

The colour charge was introduced in consequence of the observation of the Δ^{++} baryon, a light quarks baryon having spin $J = 3/2$. Fermions are subject to the Pauli exclusion principle claiming that fermions cannot be in the same quantum state at the same time and position. The Δ^{++} baryon consists of three u quarks having the same spin orientation. Since it has no angular momentum and positive parity its wave function would be totally symmetric. Hence, it violates the Pauli exclusion principle unless an additional quantum number is attached to the quarks. This quantum number is the colour charge.

The colour charge takes either the values red, green or blue whereas colour-anticolour combinations and combinations of all three colours result in a colourless state. Since only colour singlet states have been observed in nature so far, mesons are supposed to be linear combinations of colour-anticolour states and baryons are linear combinations of quarks having all three possible colours.

Quantum chromodynamics was developed in the style of quantum electrodynamics describing the force between colour-charged particles based on the exchange of massless spin-1 gauge bosons, the gluons. The nine different colour-anticolour combinations of the gluons yield a colour octet and a colour singlet. However, since the colour singlet is totally symmetric it cannot interact with other colour-charged states. Hence, the colour singlet has no physical meaning and reduces the number of gluons to eight.

In quantum field theory, the strength of an interaction between particles is governed by a coupling constant. After careful consideration, it turns out that this coupling constant depends on the scale being typical for the interaction process, Q^2. In most cases, this is the momentum transfer squared. In quantum chromodynamics, the coupling constant has a strong dependence on Q^2 due to the self-carrying of colour by the gluons. At first order perturbation theory, the Q^2 dependence of the coupling constant α_s has the form

$$\alpha_s\left(Q^2\right) = \frac{4\pi}{\left(11 - \frac{2}{3}n_F\right)\ln\left(Q^2/\Lambda_{\mathrm{QCD}}^2\right)} \tag{2.4}$$

where n_F is the number of quark flavours to be considered in the interaction process, that is, quarks having masses smaller than Q. The parameter Λ_{QCD} is the fundamental parameter of quantum chromodynamics, usually referred to as the asymptotic scale parameter. It is the only free parameter in quantum chromodynamics and has to be determined in experiments.

According to the de Broglie relation between wave length λ and momentum $\vec{p}$

$$\lambda = \frac{h}{|\vec{p}|} = \frac{h}{\sqrt{Q^2}} \tag{2.5}$$

the momentum transfer Q^2 dependence can be interpreted in terms of a distance dependence. Deduced from equation (2.4), the coupling of the strong interaction vanishes for short distances and high momenta, respectively:

$$Q^2 \to \infty \quad \Rightarrow \quad \alpha_s \to 0. \tag{2.6}$$

In this limit, known as the asymptotic freedom, the quarks are quasi-free since they are bound rather weakly due to the colour-shielding of the gluons.

On the other side, at large distances and lower momenta, the coupling of the strong force increases dramatically towards infinity:

$$Q^2 \to 0 \quad \Rightarrow \quad \alpha_s \to \infty. \tag{2.7}$$

In this so-called confinement region, increasing energy is required to separate quarks out from bound states. If the potential energy exceeds twice a quark mass, the total energy is minimised by producing a quark-antiquark pair instead. This process is called fragmentation and explains why no free quarks have been observed.

Symmetries in Quantum Chromodynamics

Perturbative predictions in quantum chromodynamics are only possible for small coupling constants $\alpha_s \ll 1$ being true for small distances where $Q^2 \gg \Lambda^2_{\mathrm{QCD}}$. In the non-perturbative region there are only a few cases where predictions can be made mainly on the basis of symmetries arising from quantum chromodynamics [12].

- **Isospin symmetry**

  Historically, isospin symmetry was the first symmetry of QCD discovered. The isospin is a quantum number describing states composed of u and d quarks whereas the u and the d quark themselves form an isospin doublet. Accordingly, the proton and the neutron represent an isospin doublet as well. Thereby they can be interpreted as two different representations of the nucleon.

  Isospin symmetry is not an exact symmetry, though. The approximation arises from the fact that the mass difference $(m_d - m_u)$ between the d and the u quark mass is small against Λ_{QCD}. Examples for isospin multiplets are:

$$\begin{pmatrix} p \\ n \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix} .$$

- **Chiral symmetry**

  Chirality is a rather abstract concept of quantum theory. For massless particles it is identical to the helicity being the projection of the spin onto the direction of the momentum. Chiral symmetry emerges because the u and d quark masses are small compared to Λ_{QCD}. Despite the fact that chiral symmetry is spontaneously broken, predictions in terms of the chiral perturbative theory are possible.

- **Heavy quark symmetry**

  If the mass of a quark gets sufficiently large, that is $m_Q \gg \Lambda_{\mathrm{QCD}}$, the heavy quark is not affected by the light quarks it is bound to. From this symmetry an effective theory of quantum chromodynamics can be derived, referred to as Heavy Quark Effective Theory.

2.3 Excited B Mesons

The Heavy Quark Effective Theory (HQET) is a limit of quantum chromodynamics based on the heavy quark symmetry. In this limit, predictions of mesons consisting of pairs of heavy and light quarks $Q\bar{q}$ are simplified, provided that the heavy quark mass is large compared to the asymptotic scale parameter: $m_Q \gg \Lambda_{\mathrm{QCD}}$. In this case, the momentum exchange among the heavy and light quark is noticeably smaller than the heavy quark mass. Hence, the recoil is negligible and in the limit of $m_Q \to \infty$ the heavy quark can be treated as a static source of an electromagnetic and a colour field. In this limit, the interactions of the light quark are independent of the heavy quark. The spectrum of the meson is therefore determined by the excitations of the light quark only and would be perfectly degenerate for infinitely heavy quark masses.

In reality the heavy quark mass is finite and the Heavy Quark Effective Theory provides a formalism for corrections in powers of $\Lambda_{\mathrm{QCD}}/m_Q$. These corrections yield a spectrum of the meson similar to the spectrum of the hydrogen atom. Indeed, the heavy-light quark mesons play a similar role in quantum chromodynamics as the hydrogen atom in quantum electrodynamics. In analogy, the mass splitting of the excited states is called hyperfine structure. It is proportional to the heavy quark chromomagnetic moment μ_Q:

$$\mu_Q = \frac{g}{2m_Q}. \tag{2.8}$$

The mesons studied in this thesis are the orbitally excited states of the B and B_s mesons containing a heavy $\bar{b}$ antiquark and a light d or s quark, respectively. The charge conjugated mesons $\bar{B}$ ($\bar{d}b$) and $\bar{B}_s$ ($\bar{s}b$) are always included within this analysis unless explicitly stated otherwise. The quantum numbers of the excited states are now derived with respect to the B and B_s mesons as follows. Thereby, the nomenclature of $B_{(s)}$ refers to both mesons at the same time.

In the ground state, with orbital angular momentum $L = 0$, the total angular momentum j_q of the light quark is $j_q = 1/2$. Thus, the total angular momentum of the meson can either be $J = 0$ or $J = 1$. The $J^P = 0^-$ state is a pseudoscalar meson, called $B_{(s)}$, and the $J^P = 1^-$ state is a vector meson, referred to as $B^*_{(s)}$. The decay of the $B^*_{(s)}$ is purely electromagnetic by photon emission because the marginal mass difference between the $B^*_{(s)}$ and the $B_{(s)}$ is insufficient to produce even the lightest hadrons:

$$B^*_{(s)} \to B_{(s)}\ \gamma. \tag{2.9}$$

For the lowest possible orbital angular momentum $L = 1$, the spin s_q of the light quark couples with the orbital angular momentum:

$$j_q = |L \pm s_q|\,. \tag{2.10}$$

Here, j_q is the total angular momentum of the light quark taking the values of either $j_q = 1/2$ or $j_q = 3/2$. Since the mass of the heavy quark is much larger than the light quark mass the quantum numbers are independent of the heavy quark. Therefore, j_q is a good quantum number and it couples with the spin s_Q of the heavy quark to the total angular momentum

$$J = |j_q \pm s_Q| \tag{2.11}$$

forming two isodoublets in terms of j_q being altogether four excited states with $J = 0, 1, 1$ and $J = 2$. The states in the first doublet are called $B^*_{(s)0}$ and $B^*_{(s)1}$ and the states in the second doublet are called $B_{(s)1}$ and $B^*_{(s)2}$. Together, these four states are collectively referred to as $B^{**}_{(s)}$ mesons and they are summarised in table 2.3.

Actually, the naming convention in particle physics stated in the “Review of Particle Physics” [2] does not use a star for $J = 1$ states. Instead, the convention reads to write the mass of the state in parentheses to the symbol. Since merely the masses of the $B^{**}_{(s)}$ states with $j_q = 3/2$ are known the stars will still be used throughout this thesis to avoid ambiguities among the $J = 1$ states.

In general, the B^{**} mesons decay into B or B^* mesons by emitting a pion. In the following sections the B and B^* mesons are referred to as $B^{(*)}$. However, the decay

State	j_q	J^P	Transition	Shape	Decay mode
B_0^*	$\frac{1}{2}$	0^+	S-wave	broad	$B\pi$
B_1^*	$\frac{1}{2}$	1^+	S-wave	broad	$B^*\pi$
B_1	$\frac{3}{2}$	1^+	D-wave	narrow	$B^*\pi$
B_2^*	$\frac{3}{2}$	2^+	D-wave	narrow	$B\pi$, $B^*\pi$

Table 2.3: Summary of the four orbitally excited B^{**} states and their properties. The nomenclature of the different states is purely historical and does not follow a consistent scheme.

$B_1 \to B\pi$ is forbidden by angular momentum and parity conservation. The same is true for the corresponding decays of the B_s^{**} mesons into $B_s^{(*)}\pi$ due to isospin conservation. Therefore, the B_s^{**} mesons decay into BK. Furthermore, the excited B^{**} and B_s^{**} states can also decay by emitting two pions. Here, both the decays $B^{**} \to B\pi\pi$ and $B_s^{**} \to B_s\pi\pi$ are possible.

Figure 2.2 shows the schematic term diagram of the B^{**} transitions via pion emission. For mesons of the isodoublets having $j_q = 3/2$ only D-wave transitions are allowed. Consequently, they are expected to have a very narrow decay width compared to the typical strong decay width. In contrast, the isodoublets with $j_q = 1/2$ are supposed to have a rather broad decay width of about 100 MeV/c^2 [13] due to the allowed S-wave transition. Therefore, these broads states are very difficult to distinguish from the combinatorial background in experimental particle detectors.

2.4 Theoretical Predictions and Experimental Results

Predictions concerning the mass and width of the excited B^{**} and B_s^{**} mesons are possible in terms of the Heavy Quark Effective Theory. Table 2.4 summarises some mass predictions on the basis of Heavy Quark Effective Theory applying different models to describe the motion of the light quark. These models include next to leading order heavy quark expansions [14], non-relativistic models for the light quark [15, 16, 17], fully relativistic treatment of the light quark [18, 20], and a variant of the MIT bag model [19]. In this context, two of the theoretical calculations of the masses also predict the width of some excited states, as listed in table 2.5.

Very recently, predictions about the decay width of the B_s^{**} states in decays of $B_s^{**} \to B_s\pi\pi$ have been published taking into account the latest results from CDF [1] and DØ [33]. As seen in table 2.6, partial decay widths in the range of a few keV/c^2 are predicted.

In recent years, a lot of effort was put into finding evidence for orbitally excited states in the D, D_s, B and B_s sector. Concerning excited B mesons, first evidence has been seen at LEP [22, 23, 24, 28]. However, due to statistical limitation the

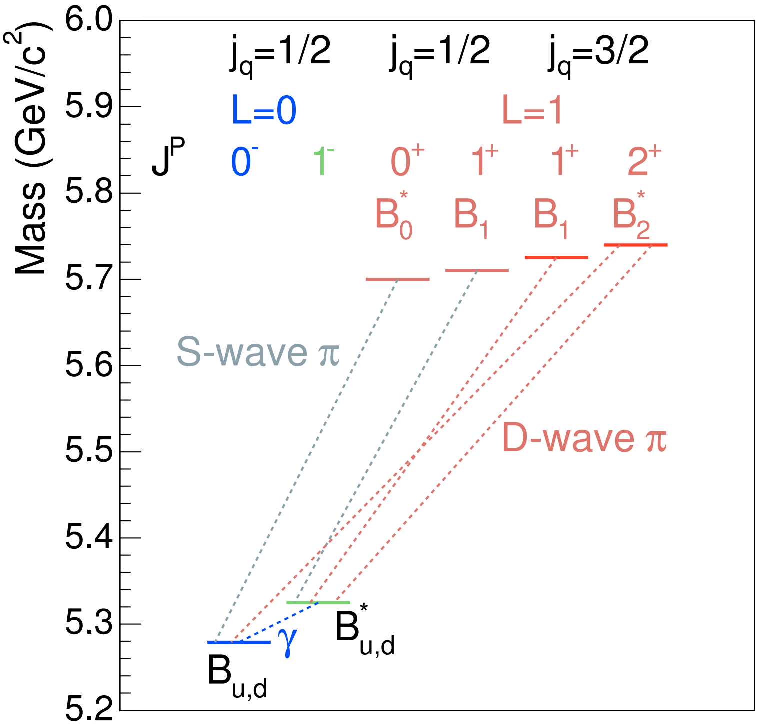


Figure 2.2: Term diagram of the B^{**} mesons. The diagram for B_s^{**} mesons looks in principle the same except that B_s^{**} mesons decay into B mesons by emitting a kaon. The decays where two pions are emitted are omitted for clarity.

Theory	$m(B_0^*)$	$m(B_1^*)$	$m(B_1)$	$m(B_2^*)$	Year
Ref. [14]	–	–	5.780	5.794	1995
Ref. [15, 16]	5.650	5.650	5.759	5.771	1993, 1994
Ref. [17]	5.870	5.875	5.700	5.715	1998
Ref. [18]	5.738	5.757	5.757	5.733	1999
Ref. [19]	5.592	5.671	5.623	5.637	1999
Ref. [20]	5.92	5.649	5.720	5.737	2007

Table 2.4: Mass predictions for the four B^{**} states with different models used for describing the motion of the light quark. The values are given in GeV/c^2.

	$\Gamma(B_0^*)$	$\Gamma(B_1^*)$	$\Gamma(B_1)$	$\Gamma(B_2^*)$
Ref. [14]	–	–	–	16 ± 6
Ref. [19]	141	139	20	29

Table 2.5: Predictions of the widths of the B^{**} states. The values are given in MeV/c^2.

Decay mode	$\Gamma_{\pi^+\pi^-}$ [keV]	$\Gamma_{\pi^0\pi^0}$ [keV]
$B_{s1} \to B_s\pi\pi$	2.5–6.1	1.4–3.4
$B_{s1} \to B_s^*\pi\pi$	0.9–2.2	0.5–1.3
$B_{s2}^* \to B_s^*\pi\pi$	18.8–46	10.8–26.6

Table 2.6: Predictions of the partial decay widths of the B_s^{**} states [21].

narrow B^{**} states could not be separated from each other. Afterwards, narrow states were measured by the DELPHI Collaboration [29, 30, 31] and the DØ experiment [25, 26]. Moreover, CDF also measured this states previously in 374 pb^{-1} of data [27].

In the B_s sector however, far less observations have been made. Historically, the OPAL Collaboration [28] has first seen one excited state of the B_s which was later confirmed by the DELPHI Collaboration [29, 30, 31] and recently also observed by DØ [32]. By then, merely one single B_s state was observed. But in the meantime, in 2006, the CDF Collaboration could measure the B_{s1} and B_{s2}^* states separately as two single peaks in decays of $B_s^{**} \to B^+K^-$ [1].

3 Experimental Setup

The Tevatron is a proton antiproton accelerator at the Fermi National Accelerator Laboratory (Fermilab) located about 70 km west of Chicago in Batavia, Illinois (USA). It accelerates protons and antiprotons in order to produce collisions at two interaction points, where the experiments CDF and DØ are located, respectively. The Tevatron has a radius of 1 km and reaches a center of mass energy of 1.96 TeV and it is the world's largest collider currently in operation before it will be superseded by the Large Hadron Collider at the CERN.

The Tevatron started operation in 1985, but after eleven years of operation the accelerator and the experiments were upgraded in 1996. During this first period, also referred to as Run I, data has been taken with a final dataset having an integrated luminosity of 90 pb^{-1}. A major achievement of Run I was the discovery of the top quark.

After five years of upgrading the accelerator and the experiments, the Tevatron restarted operation in 2001 with Run II [35]. Figure 3.1 shows the integrated luminosity delivered and recorded within this period until 2009.

3.1 Accelerator Complex

Figure 3.2 shows a sketch of the accelerator complex at the Fermilab used for accelerating protons and antiprotons [34]. Beginning from the Cockcroft-Walton accelerator the protons pass a chain of several successive accelerators before they are injected into the Tevatron, the main accelerator, where they are accelerated together with the antiprotons to the final energy of 980 GeV before they are collided. The antiproton production is done at an intermediate step of the proton acceleration.

3.1.1 Proton Source

Negatively charged hydrogen ions are accelerated in the Cockcroft-Walton accelerator to 750 keV using high voltage fields. Afterwards they are passed to the LINAC [36, 37], a 150 m long linear accelerator where the ions are accelerated to 400 MeV by radio frequency resonators. After the LINAC the ions enter the Booster [38], the first synchrotron in the accelerator chain. Upon entrance they pass

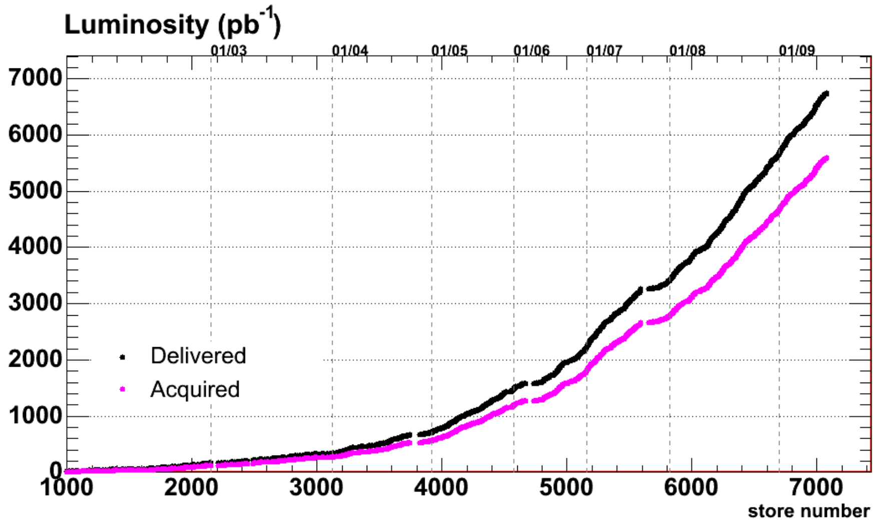


Figure 3.1: The integrated luminosity delivered by the Tevatron (black) and recorded by the CDF II detector (purple).

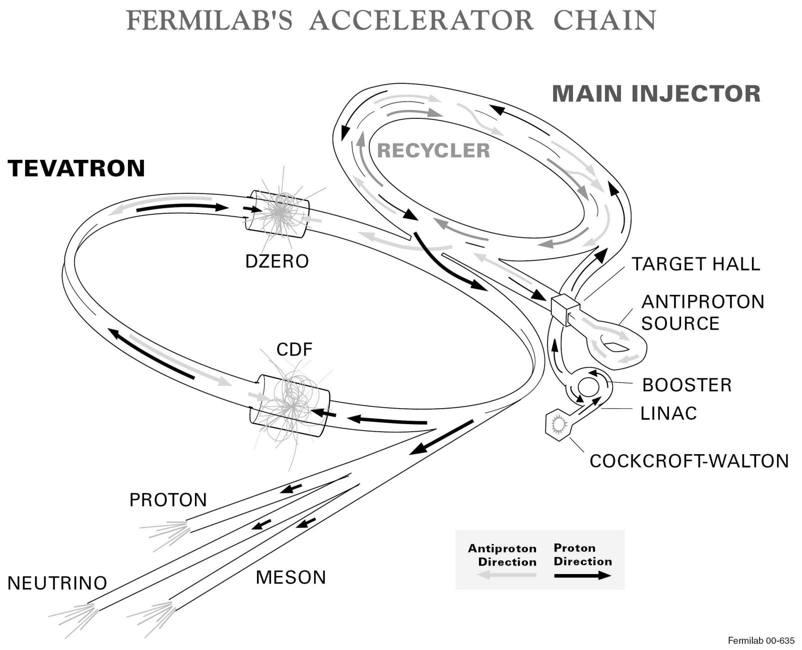


Figure 3.2: The accelerator chain at Fermilab.

through a carbon foil where their electrons are stripped off. The remaining protons are accelerated further to 8 GeV using radio frequency cavities.

The last preaccelerator the protons pass is the Main Injector [44] having several operation modes. It can either accelerate the protons to 120 GeV for the antiproton production, the so-called stacking, or it can accelerate the protons to 150 GeV for injection into the Tevatron. The 120 GeV protons are also used for fixed target experiments located at the Fermilab.

3.1.2 Antiproton Source

The antiproton production is the limiting factor for the Tevatron performance since it takes up to 20 hours to produce one stack of antiprotons. For the antiproton production the proton beam coming from the Main Injector is guided onto a nickel target. Striking the target produces all kinds of secondary particles including antiprotons with an efficiency of approximately $20 \cdot 10^{-6}$. The antiprotons with the energy of 8 GeV/c^2 which are separated from the other secondary particles by pulsed magnets and focused by a lithium lens are gathered and sent to the Debuncher.

The Debuncher, which has a mean radius of 90 m, is the first synchrotron within the antiproton production chain. It rotates and cools the antiproton bunches stochastically before they enter the second synchrotron, the Accumulator having a mean radius of 75 m. The Accumulator uses stochastic cooling [39, 40] to reduce the wide spread in kinematic energy. At this stage, the antiprotons are still at an energy of 8 GeV. In the Recycler [43], which is in the same tunnel as the Main Injector, the antiprotons are cooled further using stochastic and electron cooling [41, 42]. The Recycler is also the place where the antiprotons are stored.

3.1.3 Tevatron

The Tevatron, a synchrotron with 6 km circumference, represents the last stage in the accelerator chain. It is the world's first synchrotron using superconducting magnets which are essential for beam optics at such high energies. The superconducting coils are made of a niobium-titanium alloy and they are kept at a temperature of approximately 4 K. The Tevatron accelerates the protons and antiprotons up from 150 GeV to 980 GeV in a few seconds corresponding to a final center of mass energy of 1.96 TeV when the particles collide.

The beam is split into three trains of twelve bunches with a bunch revolution time of 21 μs. The empty space between trains, the abort gap, corresponds to a time of 2.6 μs. This time is used for ramping up the kicker magnets to dump the beam safely into graphite blocks. Since the particles diffuse out of the bunches it is important to keep the abort gap clean. Therefore a low energetic electron beam is directed into the gap which excites the particles between the bunches to high amplitudes until they get lost and thus keeping the gaps clean.

The period between beam initialisation and beam abort is called store. When a new store starts, the preaccelerated protons and antiprotons are injected during a period, called shot setup, into the Tevatron. The store duration averages several hours. For the proton injection seven bunches of protons from the Booster are transfered to the Main Injector and accelerated to 150 GeV before they are merged to two bunches and shot into the Tevatron. This procedure is repeated 36 times to have 36 proton bunches in the Tevatron. Thereafter four bunches of antiprotons are transferred from the Recycler to the Main Injector where they are accelerated to 150 GeV before they are shot into the Tevatron. This procedure is repeated nine times. The bunches collide at two intersection points where the CDF and DØ experiments are located, respectively.

To prevent the detectors from being damaged, stable beam conditions are required. Thus, the beam has to be fine tuned before data can be taken implying focusing the beam to a smaller transverse size (low beta squeeze) for gaining luminosity and removing the beam halo by ramping stainless steel collimators close to the beam (scraping).

3.1.4 Luminosity

The instantaneous luminosity $\mathcal{L}$ is a measure for the performance of an accelerator. It can be written as

$$\mathcal{L} = n \cdot f \cdot \frac{N_p \, N_{\bar{p}}}{4\pi \, \sigma_x \, \sigma_y} \tag{3.1}$$

where n is the number of bunches, f is the revolution frequency, $N_{p(\bar{p})}$ is the number of protons (antiprotons) per bunch and $\sigma_{x,y}$ is the average transverse width of the bunches. The luminosity is measured in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Since particles get lost due to collisions, the luminosity decreases over the time and has its maximum at the beginning of a store. The integrated luminosity over time yields an estimate of the amount of collected data. For a given cross section the number of produced events N can be calculated using

$$N = \sigma \int \mathcal{L} \, \mathrm{d}t. \tag{3.2}$$

The CDF experiment, described in section 3.2, determines the instantaneous luminosity with the aid of Cherenkov luminosity counters (CLC) [45] located around the beam pipe in the forward and backward region. They cover a a polar angle range from 2.8° to 1.0°. The counters measure the number of inelastic $p\bar{p}$ collisions per bunch crossing by the amount of detected Cherenkov light caused by primary particles in the gas-filled Cherenkov detectors.

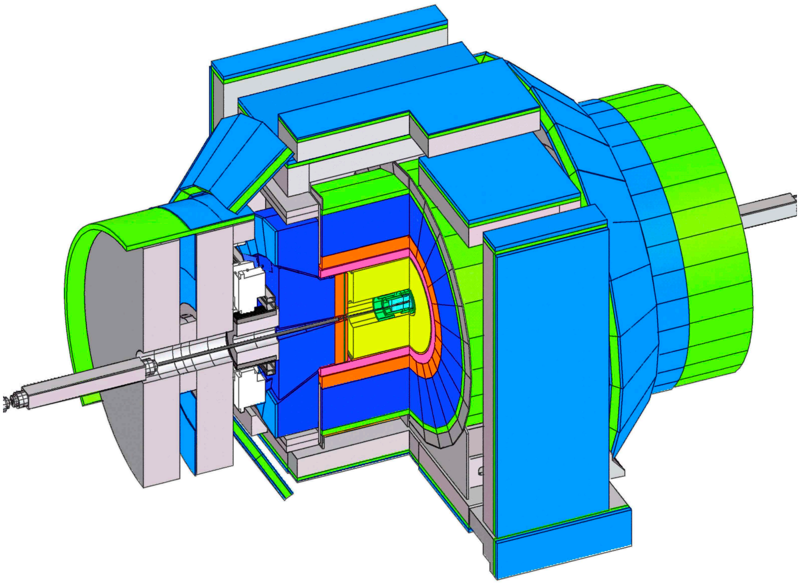


Figure 3.3: The CDF II detector.

3.2 Collider Detector at Fermilab

The CDF II detector [46] is a multipurpose collider detector with cylindrical and forward-backward symmetry as seen in figure 3.3. It consists of a tracking system, a time-of-flight subdetector, calorimeters and muon chambers. The tracking system is located in a superconducting solenoid generating a magnetic field of 1.4 T, oriented parallel to the beam pipe. Figure 3.4 shows an elevation view of the detector.

The origin of the coordinate system of the detector is the nominal particle interaction point. The positive direction of the z axis is defined by the direction of the proton beam pointing from the west to the east. In the Cartesian representation of the coordinates the x axis is in the horizontal plane pointing outwards from the center of the collider and the y axis is perpendicular to the horizontal plane pointing upwards. Most often coordinates are used which are defined in the following way:

- r coordinate
  Distance to the center of the beam pipe.

- ϕ coordinate
  The azimuthal angle is measured perpendicular to the horizontal plane.

- θ coordinate
  The polar angle is measured from the positive z direction.

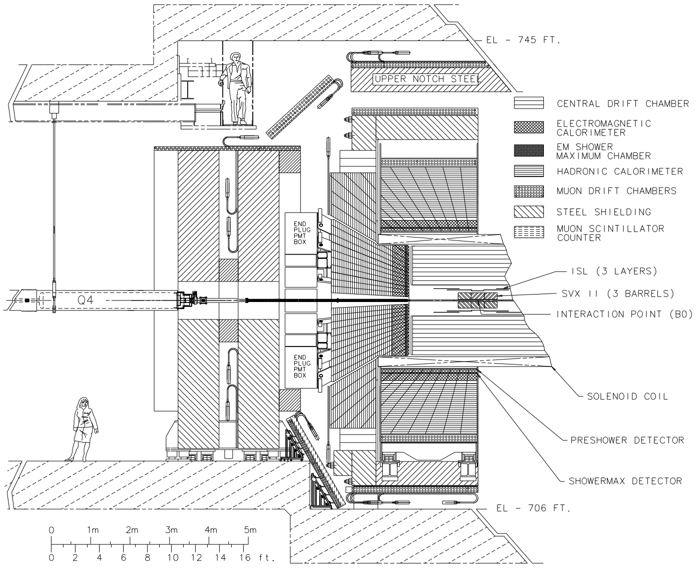


Figure 3.4: Elevation view of the CDF II detector.

Within this context the pseudorapidity η is often mentioned being derived from the polar angle θ as

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{3.3}$$

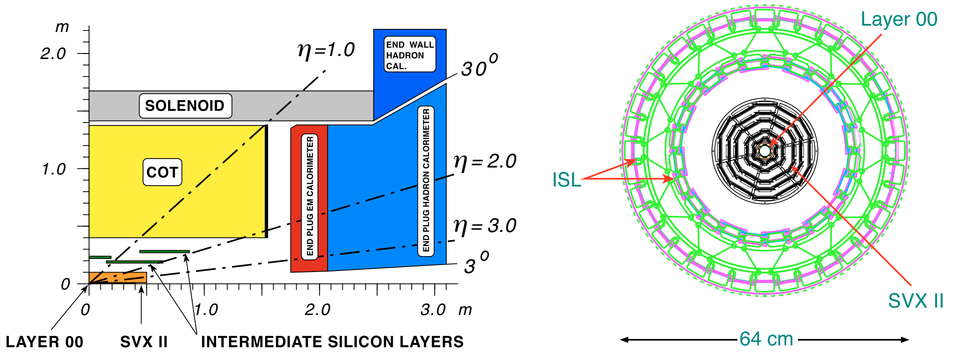
The pseudorapidity is very handy since the difference in η of two particles having negligible masses is independent from any Lorentz boosts along the beam pipe.

3.2.1 Tracking System

The tracking system consists of two subdetectors for detecting charged particles and measuring their momenta and displacements with respect to the primary vertex. The innermost part of the tracking system is the silicon detector with its subsystems Silicon Vertex Detector (SVXII), Intermediate Silicon Layer (ISL) and Layer00 (L00). The outer part is the Central Outer Tracker (COT) surrounded by a superconducting solenoid. Figure 3.5(a) shows a sketch of the tracking system.

Figure 3.5(b) shows an endview of the silicon detector which covers a radial range from 1.6 cm to 28 cm with an acceptance of $|\eta| < 2$. The subcomponents are:

- **Layer00 (L00)**
  Layer00 [47] was added during the upgrade for Run II. It is made of radiation hard silicon microstrips mounted directly onto the beam pipe consisting of two overlapping layers at a radial position of 1.35 cm and 1.62 cm, respectively. The microstrips deliver information about track hits in the r-ϕ plane.



(a) The CDF II tracking system and the plug calorimeters.
(b) The silicon microstrip detector.

Figure 3.5: CDF II tracking system.

- **Silicon Vertex Detector (SVXII)**
  The SVXII [48] comprises five layers of double-sided microstrips at a radial range from 2.1 cm up to 17.3 cm. The microstrips on one side are in z direction whereas the microstrips of the other side are either perpendicular for layer 0, 1 and 3 or with a small stereo angle of 1.28° for layer 2 and 4 in order to provide a measurement of the z coordinate.

- **Intermediate Silicon Layer (ISL)**
  The ISL [49] consists of three layers. The central layer is at the radial position of 22 cm with an acceptance of $|\eta| < 1$ and the two layers in the forward/backward region of the detector are at a radial position of 20 cm and 28 cm, respectively, with an acceptance of $1 < |\eta| < 2$. The ISL is made of double-sided microstrips, which are parallel to the z axis on one side and with a small stereo angle on the other side. The ISL is designed for linking tracks from the silicon detector with a good spatial resolution to track hits found in the COT having a worse spatial resolution.

The silicon detector provides a very precise measurement of the track's impact parameter and the ϕ coordinate yielding an accurate 3-dimensional track reconstruction with a high spatial resolution.

The Central Outer Tracker (COT) [50] is a cylindrical open-cell drift chamber outside the silicon detector covering a radial range from 33 cm to 143 cm. It is filled with an Argon-Ethane-CF$_4$ gas mixture and it has 96 wire layers grouped into eight superlayers. Four superlayers, the axial superlayers, have wires in z direction whereas the other four superlayers, the stereo superlayers, have wires tilted by two degrees with respect to the z direction. The outermost superlayer has an acceptance of $|\eta| < 1$. Particles going through the drift chamber ionise the gas mixture and the voltage bias applied to the wires causes the electrons to drift to the sense wires where they induce an electric signal. For geometrical reasons the track density in the COT is smaller than in the silicon detector. Since the COT can have a maximum of 96 hits per track it yields a high precision measurement of the transverse momentum in the r-ϕ plane.

3.2.2 Time of Flight Detector

The time of flight detector [51] is cylindrical around the COT but still within the superconducting solenoid. It comprises 216 scintillators and photomultipliers for measuring the particle's time of flight between the interaction point and the time of flight detector. In conjunction with the momentum measurement from the COT the particle mass can be determined by

$$m = \frac{p}{c}\sqrt{\frac{(ct)^2}{L^2} - 1} \tag{3.4}$$

where L is the path length and t is the time difference between the production time and the arrival time at the detector. Since the resolution of the time of flight detector is in the order of magnitude of picoseconds the separation power between pions and kaons based on their masses is about two standard deviations for momenta below 1.6 GeV/c.

3.2.3 Particle Identification

In order to obtain a clean signal sample, it is conducive being able to separate the long-lived particles, namely protons, kaons, pions, muons, electrons which are long-lived enough to pass through the whole detector. The time of flight measurement provides already a good particle identification but only for low momenta. Thus, the particle identification is improved by taking into account the particle's specific energy loss in the COT.

Charged particles passing the COT ionise the gas mixture in the drift chamber. The pulse width Δt on the readout chips is logarithmically proportional to the charge deposit Q and therefore proportional to the energy loss $\mathrm{d}E/\mathrm{d}x$:

$$\Delta t \sim \log Q \sim \frac{\mathrm{d}E}{\mathrm{d}x}. \tag{3.5}$$

Figure 3.6 shows the separation power for different particle combinations using the time of flight measurement for kaon/pion, proton/kaon and proton/pion separation and the separation power of the COT $\mathrm{d}E/\mathrm{d}x$ measurement.

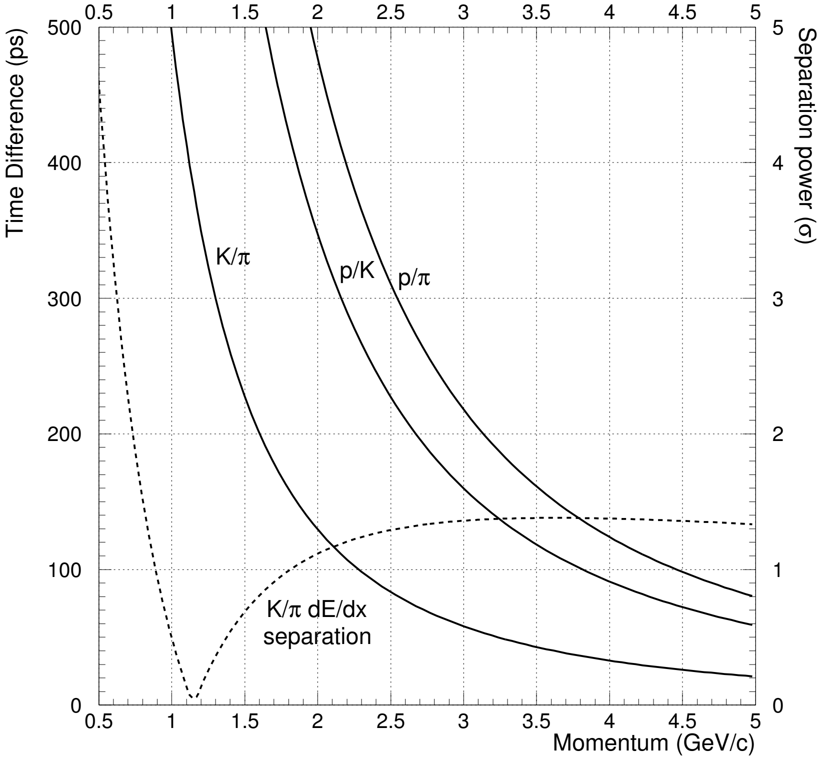


Figure 3.6: Time of flight difference between different particle species (solid line) and the separation power for K/π using COT $\mathrm{d}E/\mathrm{d}x$ (dashed line) [53].

3.2.4 Calorimeters

The five calorimeters [52] of the CDF II detector are outside the superconducting solenoid. The central and end-plug calorimeters are electromagnetic [54] and hadronic calorimeters, whereas the end-wall calorimeter is a hadronic calorimeter only [55]. Altogether they have an acceptance of $|\eta| < 3.6$ and measure the energy of high energetic electrons, photons and hadronic jets. Since they are optimised especially for high energy physics the calorimeters are not important in context of this analysis. A detailed description is given in the references.

3.2.5 Muon System

The outmost part of the CDF II detector is the muon system [56, 57]. Since electrons, photons and hadrons are usually absorbed in the calorimeters the muon system provides a very good muon identification. The muon system consists of four subdetectors including drift chambers, scintillators and steel absorber. The Central Muon Detector (CMU) covers the central part having a range of $|\eta| < 0.6$. The Central Muon Upgrade (CMP) has the same coverage as the CMU for the purpose of reducing the misidentification rate of muons in the CMU. The Central Muon Extension (CMX) enlarges the coverage of the muon system to $|\eta| < 1.0$. The different parts of the Intermediate Muon System (IMU) which are mounted on both ends of the CDF II detector extends once more the muon coverage to $|\eta| < 1.5$.

3.2.6 Trigger System

The Tevatron runs with a collision rate of 2.5 MHz. That is, every 0.4 μs data of approximately 250 kB accumulates yielding a total data rate of 625 GB/s assuming an interaction at every collision. It is neither efficient nor possible to store such an amount of data considering that most of the collected events are background events anyway. Hence, the CDF II detector has a sophisticated trigger system to make a decision almost in real time whether an event is worth to be recorded or can be left out [58].

The trigger system comprises a three level system making successive decisions and thereby reducing the data rate from 2.5 MHz to about 100 Hz. Accepted events are finally stored on tape and ready to be prepared by the various reconstruction algorithms for the different analyses. In order to be fast enough reducing the data rate down to 50 kHz and 300 Hz, respectively, the Level 1 trigger is entirely hardware implemented and Level 2 [59] is a hardware-software implementation. In contrast, Level 3 trigger is a pure software implementation running on a Linux PC farm. Figure 3.7 shows a sketch of the data flow in the CDF II trigger system.

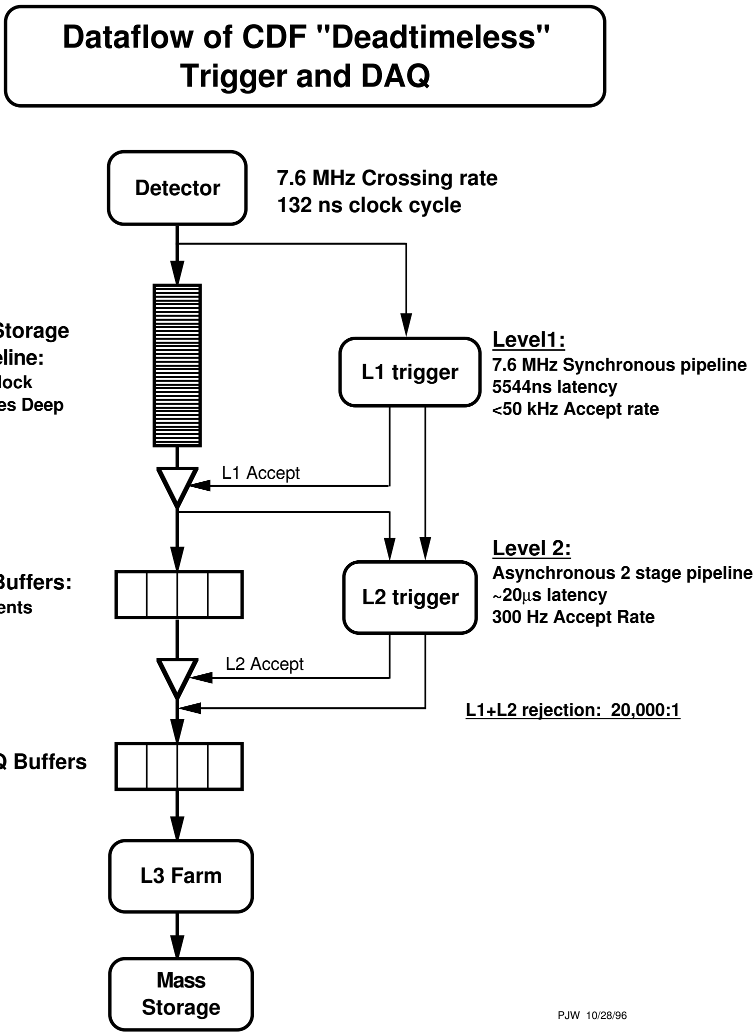


Figure 3.7: Block diagram of the data flow in the CDF II trigger and data acquisition system.

Level 1 trigger (L1)

Data in the detector is digitised and read out after every bunch crossing. While the Level 1 trigger decision takes place the data is stored in the L1 pipelines for further read out and use for Level 2 trigger. The Level 1 trigger is a synchronous trigger meaning its decisions are in time with the bunch crossing rate. For this analysis the trigger decision is mainly based on the eXtremely Fast Tracker (XFT) [60] reconstructing the transverse momentum p_t and the azimuthal angle ϕ of tracks in the COT. In addition, Level 1 uses also information from the muon chambers and the calorimeter towers to find muons, electrons and photons.

Level 2 trigger (L2)

If an event is accepted by the Level 1 trigger, it is passed to Level 2 where it is again evaluated by taking into account more information about the event provided by different detector components. Since a Level 2 decision takes slightly more time, Level 2 works asynchronously and its buffer can therefore hold up to four events. The Level 2 trigger divides into a two stage pipeline, where the first stage consists of dedicated hardware and the second stage comprises programmable processors. The core functionality for B-physics of L2 comes from the connection with the Silicon Vertex Tracker (SVT) [61]. The SVT searches for coincidences between tracks from the XFT and hits in the silicon vertex detector by comparing data using a set of precalculated patterns. Linking information from different detector components together improves the spatial resolution and thereby the precision of L2. If an event is accepted, it is passed on to Level 3.

Level 3 trigger (L3)

In order to confirm the L1 and L2 decision, Level 3 trigger runs similar algorithms as in the offline reconstruction but with different parameters. Thus, it has an improved resolution of the transverse momentum p_t, the impact parameter d_0 and other track parameters compared to L1 and L2. If an event is finally accepted by Level 3, it is stored on disks as an intermediate step before it is written to tape ready for the full offline reconstruction.

Two Track Trigger

The data of all decay channels investigated in this analysis having no J/ψ as an intermediate state was collected by the Two Track Trigger which is based on the L1 and L2 track processors with a requirement for a minimal transverse momentum p_t. The requirements are:

- **Level 1**
  At least two XFT oppositely charged tracks having a transverse momentum of $p_t > 2$ GeV/c each and the angle between the two tracks has to be $\Delta\phi^{(1,2)} < 135°$.

- **Level 2**
  At least two SVT tracks matching the XFT tracks ($\chi^2_{SVT} < 25$) having $p_t > 2$ GeV/c each with the impact parameter 120 μm $< |d_0| <$ 1 mm and the decay length $L_{xy} > 200$ μm.

- **Level 3**
  The SVT tracks have to match COT tracks, and cuts on $2° < \Delta\phi^{(1,2)} < 90°$, 80 μm $< |d_0| <$ 1 mm and $L_{xy} > 200$ μm are applied.

Besides these requirements additional trigger subpaths ask for further criteria:

- **L2_B_LOWPT**
  At least two SVT tracks having $p_t^{(1)} + p_t^{(2)} > 4.0$ GeV/c.

- **L2_B_CHARM**
  At least two oppositely charged SVT tracks having $p_t^{(1)} + p_t^{(2)} > 5.5$ GeV/c.

- **L2_B_HIGHPT**
  At least two oppositely charged SVT tracks having $p_t^{(1)} + p_t^{(2)} > 6.5$ GeV/c.

Di-Muon Trigger

The data of the decay $B^+ \to J/\psi K^+$ used in this analysis is collected with the di-muon trigger. As the name already indicates the di-muon trigger looks for pairs of muons coming from a J/ψ. At Level 1 the trigger asks for tracks in the XFT having a transverse momentum and a ϕ coordinate which match hits in the muon system. For Level 2 energy clusters in the calorimeters above a certain threshold are required. Finally, at Level 3, the mass of the J/ψ candidates has to be nearby the world average value.

4 Statistical Methods

4.1 Parameter Estimation

Taking data in high energy physics is comparable of measuring repeatedly the outcome of an experiment. Ideally measurements are repeatable but in nature repeated measurements do not yield exactly the same values. Depending on the measurement this can be understood as a result of thermal fluctuations, quantum effects and the finite resolution of the experiment, respectively, inducing a randomly distributed difference between the measured values and the truth.

Hence, the necessity is obvious of having an appropriate method for determining parameters and their uncertainties in the context of a given hypothesis. The different methods of parameter estimation are characterised by a set of criteria [62]. In the following $\hat{\vec{a}}$ is the set of estimated parameters and $\vec{a}_0$ refers the truth. The criteria are:

- **Consistency**
  In the limit of infinite statistics $\hat{\vec{a}}$ converges to the truth: $\lim_{n\to\infty} \hat{\vec{a}} = \vec{a}_0$.
- **No bias**
  The expectation value of $\hat{\vec{a}}$ is the truth: $E[\hat{\vec{a}}] = \vec{a}_0$.
- **Efficiency**
  The variance of $\hat{\vec{a}}$ is as small as possible.
- **Robustness**
  The method is robust against defective data or insufficient hypotheses.

A well-established method for estimating parameters is the maximum likelihood method [62, 2]. Since it is widely used in this analysis for determining parameters, it will be described in the next sections.

4.1.1 The Maximum Likelihood Method

Let $\vec{x}_1$, $\vec{x}_2$, …, $\vec{x}_n$ be a set of n independent measurements of the variable $\vec{x}$ whose multi-dimensional probability density function $f(\vec{x}|\vec{a})$ is known except for the parameters $\vec{a}$. Assuming the measurements being statistically independent the likelihood function L can be written as:

$$L(\vec{a}) = f(\vec{x}_1|\vec{a}) \cdot f(\vec{x}_2|\vec{a}) \cdot \ldots \cdot f(\vec{x}_n|\vec{a}) = \prod_{i=1}^{n} f(\vec{x}_i|\vec{a}). \tag{4.1}$$

The likelihood function is a joint probability density function for the data but not a probability density function for the parameters. It can be interpreted as a probability to get the measured data for a given set of parameters. Hence, the best estimators $\hat{\vec{a}}$ for the parameters $\vec{a}$ are those maximising the likelihood function. During the maximising procedure it is required that the likelihood function is normalised for any parameters $\vec{a}$ at any iteration step to find the maximum:

$$\int f(\vec{x}|\vec{a})\,\mathrm{d}\vec{x} = 1 \qquad \text{for all } \vec{a}. \tag{4.2}$$

To avoid numerical instabilities and since most of the statistic packages are optimised for minimising, the negative logarithm of the likelihood function is commonly used, known as the negative log likelihood $\mathcal{L}$:

$$\mathcal{L}(\vec{a}) = -\ln L(\vec{a}) = -\sum_{i=1}^{n} \ln f(\vec{x}_i|\vec{a}). \tag{4.3}$$

In general, the maximum likelihood method is consistent. In the limit of infinite statistics it is also unbiased and efficient. Drawbacks are the high computing power consumption and the assumed knowledge of the probability density function

In the asymptotic limit $n \to \infty$ the likelihood approaches Gaussian and the negative log likelihood function approaches parabolic shape. It can be expanded around its minimum $\hat{\vec{a}}$ where the first derivation vanishes ($\frac{\partial \mathcal{L}}{\partial \vec{a}} = 0$):

$$\begin{aligned} \mathcal{L}(\vec{a}) &= \mathcal{L}(\hat{\vec{a}}) + \frac{1}{2}\sum_{i,j} \frac{\partial^2 \mathcal{L}}{\partial a_i\, \partial a_j}(a_i - \hat{a}_i)(a_j - \hat{a}_j) + \ldots && (4.4)\\ &= \mathcal{L}(\hat{\vec{a}}) + \frac{1}{2}\sum_{i,j} G_{ij}(a_i - \hat{a}_i)(a_j - \hat{a}_j) + \ldots && (4.5) \end{aligned}$$

where $\vec{a} = (a_1, a_2, \ldots)$ and $\mathbf{G}$ is the inverse of the covariance matrix $\mathbf{V}^{-1}$ being identical to the Hessian matrix $\mathbf{H}$ in the asymptotic case. In the non-asymptotic case the Hessian matrix $\mathbf{H}$ is a good approximation for $\mathbf{V}^{-1}$.

Having only one parameter a the expanded likelihood function has the form

$$\mathcal{L}(a) = \mathcal{L}(\hat{a}) + \frac{1}{2}\frac{\partial^2 \mathcal{L}}{\partial a^2}(a - \hat{a})^2 + \ldots \tag{4.6}$$

and hence

$$L(a) \cong \text{const} \cdot \exp\left(-\frac{1}{2}\frac{\partial^2 \mathcal{L}}{\partial a^2}(a-\hat{a})^2\right) = \text{const} \cdot \exp\left(-\frac{(a-\hat{a})^2}{2\sigma}\right) \tag{4.7}$$

yielding in comparison with the Gaussian distribution for the asymptotic limit

$$\sigma(\hat{a}) = \left(\left.\frac{\partial^2 \mathcal{L}}{\partial a^2}\right|_{\hat{a}}\right)^{-\frac{1}{2}}. \tag{4.8}$$

Due to the parabolic shape of the log likelihood function its second derivation is constant and the value of $\mathcal{L}$ around the minimum at $a = \hat{a} \pm r \cdot \sigma$ is

$$\mathcal{L}(\hat{a} \pm r \cdot \sigma) = \mathcal{L}(\hat{a}) + \frac{1}{2}r^2 \tag{4.9}$$

where r is the number of standard deviations. Equation (4.9) can be used to calculate the standard deviation of $\hat{a}$ which is simply the position where $\mathcal{L}$ is $\frac{1}{2}$ above the minimum. This is also true if the negative log likelihood function is not of parabolic shape and the errors are asymmetric thereby. For probability density functions with more than one parameter the standard deviation and the possibly asymmetric errors are defined in a similar way

$$\mathcal{L}_{min}(\hat{a_i} + \sigma_R) = \mathcal{L}(\hat{\vec{a}}) + \frac{1}{2}, \tag{4.10}$$

$$\mathcal{L}_{min}(\hat{a_i} - \sigma_L) = \mathcal{L}(\hat{\vec{a}}) + \frac{1}{2} \tag{4.11}$$

where $\mathcal{L}_{min}(a_i)$ is the minimum of $\mathcal{L}$ for all other parameters a_j except for $i = j$.

4.1.2 The Maximum Likelihood Method for Binned Data

Mostly, measured data is represented in a histogram especially when the number of measurements n is large. The x axis of the histogram is divided into N bins containing n_j events in bin j. The number of bin entries follows a Poisson distribution with the expectation value μ_j

$$P(n_j|\mu_j) = \frac{\mu_j^{n_j} e^{-\mu_j}}{n_j!} \tag{4.12}$$

where the expectation value μ_j can be obtained by multiplying the integrated probability density function in every bin with the number n of events in the sample

$$\mu_j(a) = n \int_j f(x|a)\,\mathrm{d}x \approx n \cdot f(x_c|a) \cdot \Delta x \tag{4.13}$$

where x_c is the x value of the bin mean and Δx is the bin width. Hence, the negative log likelihood functions reads

$$\mathcal{L}(a) = -\sum_{j=1}^{J} \ln \left(\frac{\mu_j^{n_j} e^{-\mu_j}}{n_j!} \right) = -\sum_{j=1}^{J} \ln n_j \mu_j + \sum_{j=1}^{J} \mu_j + \sum_{j=1}^{J} \ln(n_j!). \tag{4.14}$$

Since the last term of the negative log likelihood function is constant and therefore does not affect the minimisation, it can be omitted.

For a large number of expected events μ is much larger than 1 and the Poisson distribution can be approximated by Gaussian distribution with variance $\sigma_j^2 = \mu_j$

$$P(n_j|\mu) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(n_j-\mu_j)^2}{2\sigma_j^2}} \tag{4.15}$$

yielding a negative log likelihood function of

$$\mathcal{L}(a) = -\sum_{j=1}^{J} \ln \left(\frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(n_j-\mu_j)^2}{2\sigma_j^2}} \right) = \frac{1}{2} \sum_{j=1}^{J} \frac{(n_j - \mu_j)^2}{\mu_j} + \text{const.} \tag{4.16}$$

On the basis of a Gaussian distribution $2\mathcal{L}(a)$ follows a χ^2 distribution with k degrees of freedom determined by the number of bins N and the number of free parameters N_a in the likelihood function: $k = N - N_a$. The ratio χ^2/k is a figure of merit for the fit performance and should be approximately 1 if the probability density function $f(x|a)$ describes the data well.

4.1.3 The Extended Maximum Likelihood Method

The use of the maximum likelihood method is the determination of parameters based on a normalised probability density function. But sometimes the average number of events is also a parameter which is to be determined. Instead of the normalised probability density function $f(x|\vec{a})$ a different function $g(x|\vec{a}) = N \cdot f(x|\vec{a})$ is used whose integral is equal to the number of events

$$\int g(x|\vec{a})\,\mathrm{d}x = N. \tag{4.17}$$

The extended negative log likelihood function is now defined as

$$\mathcal{L}(\vec{a}) = -\sum_{i=1}^{n} \ln\left(g(x_i|\vec{a})\right) + \int g(x|\vec{a})\,\mathrm{d}x \tag{4.18}$$

which can be deduced by multiplying the likelihood function $L(\vec{a})$ with the Poisson probability $P = \frac{N^n \cdot e^{-N}}{n!}$ yielding the negative log likelihood function

$$\mathcal{L}(\vec{a}) = -\sum_{i=1}^{n} \ln\left(f(x_i|\vec{a})\right) - n\ln N + N + \ln n!. \tag{4.19}$$

Evidently, this formula is the same as equation (4.18) except for the last term which is constant and can therefore be left out since it does not affect the minimisation procedure.

4.2 Artificial Neural Networks

Classifying data into signal and background samples plays a decisive role in high energy physics. The classification is commonly done by applying cuts on different discriminating variables. However, cutting on a set of variables does not take into account any correlations among those variables and the result is optimal for well-separable problems only. In contrast, mapping the set of discriminating variables to one single variable is a far more efficient approach. The classification is thereby done by cutting on this single variable only. Artificial neural networks fulfill the requirements for such a map function

$$f : \mathbb{R}^n \xrightarrow{\text{neural network}} \mathbb{R}. \tag{4.20}$$

Neural networks are the basis for the selection algorithms used for classifying into signal and background in this analysis.

4.2.1 NeuroBayes

All neural networks used in the analysis are based on the NeuroBayes® package [63] originally developed by Michael Feindt at the University of Karlsruhe and now maintained by Phi-T® Physics Information Technologies GmbH [64]. NeuroBayes® comprises a three layer feed forward neural network with a sophisticated preprocessing performed before the data is put into the network.

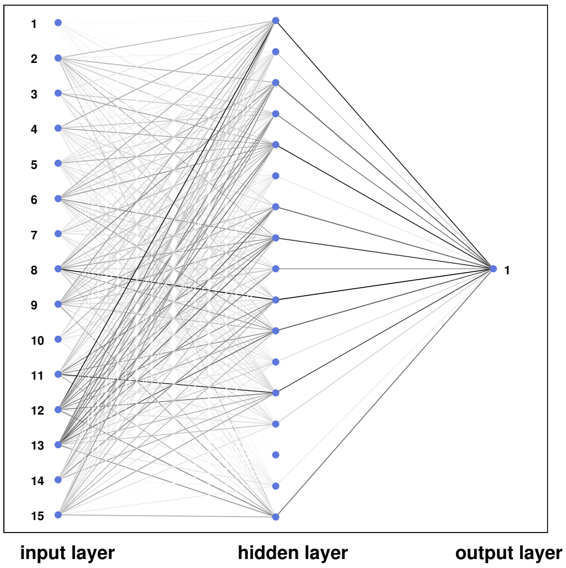


Figure 4.1: Example of the topology of a three layer feed forward neural network. The intensity of the connecting lines between nodes reflects their importance for the neural network.

4.2.2 Topology

An artificial neural network consists of nodes grouped in multiple layers. Each node has several inputs and exactly one output. The nodes are linked together among each other following a certain pattern depending on the network topology. A three layer feed forward network is one possible network topology and forms the basis for the neural networks used in this analysis. It comprises three layers of nodes: the input layer, the intermediate layer and the output layer. The information flow of a feed forward network strictly goes from the inputs to the output. That is, the outputs of the nodes in one layer are exclusively connected to the inputs of the nodes in the subsequent layer. Nodes in a layer are neither connected among each other nor are their outputs connected to nodes in the preceding layers. Figure 4.1 shows an example of the topology of a three layer feed forward neural network.

Each input variable of the neural network corresponds to one node in the input layer. The number of nodes in the intermediate layer is arbitrary to some extent, but too many nodes favour the neural network to memorise and too few nodes do not provide enough degrees of freedom being sufficient for the neural network to learn enough. For binary decisions like separating events into signal and background samples the output layer requires only one node.

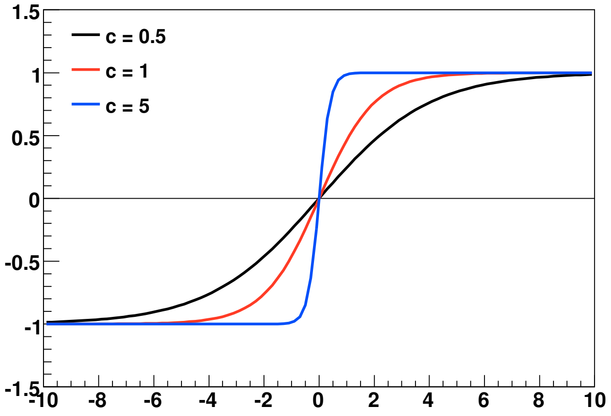


Figure 4.2: The sigmoid function, described by equation (4.22), plotted for different parameters c.

The underlying idea is the same for all nodes. The information x_i given on the n inputs of a node are weighted by the weight w_i, added up, and the sum is put into an activation function S which fires the node output y

$$y = S\left(\sum_i^n (w_i\, x_i - \mu)\right) \tag{4.21}$$

where μ is a bias threshold of the node which shifts the weighted sum to the most sensitive part of the activation function.

The activation function can be interpreted as a threshold for the node's input. A commonly used activation function is the sigmoid function

$$S(t) = \frac{2}{1 + e^{-c \cdot t}} - 1 \tag{4.22}$$

which is plotted in figure 4.2. It maps the interval $(-\infty, +\infty)$ to $(-1, +1)$. The sigmoid function is very sensitive for small values around zero but goes rather quick to saturation for large absolute values.

4.2.3 Training

A neural network needs to be trained before it can be applied to real data samples. For the training two classes of event samples are required from which the neural network can learn. One sample comprises pure signal events whereas the other sample comprises background events. During the training the weights of the nodes are adjusted iteratively until the neural network output matches the classification of the training samples. At the beginning of the training the weights of the neural network are randomly Gaussian distributed around zero with width one.

The training progress is evaluated by a cost function measuring the discrepancy between the neural network output and the target, that is, the classification of the training samples. NeuroBayes® uses usually the entropy as the cost function

$$E(\vec{w}) = \sum_{j}^{N} \ln\left(\frac{1}{2}\left(1 + A_j(\vec{w})\, s_j + \epsilon\right)\right) \tag{4.23}$$

where N is the number of events in the training sample, A_j the output of node j depending on the weights $\vec{w}$, s_j indicates whether the event is signal ($s_j = 1$) or background ($s_j = -1$) and ϵ is a regularisation constant. Although the regularisation constant vanishes after a few training iterations it is implemented to avoid numerical problems at the beginning of the training.

The cost function is a multi dimensional function in the parameter space of the weights. During the training the cost function in NeuroBayes® is minimised with the aid of gradient descent where the change of the weights Δw_j is proportional to the gradient of the cost function and the step width η_j:

$$\Delta w_j = -\eta_j \frac{\partial E(\vec{w})}{\partial w_j} \tag{4.24}$$

where the step width is adjusted by estimating the diagonal elements of the Hessian matrix of the cost function.

If the entropy is used as cost function and the neural network is well trained, the neural network output can be directly interpreted as signal probability [63].

4.2.4 Preprocessing and Pruning

To simplify the search for a minimum during the training, NeuroBayes® preprocesses the input data before it is fed into the neural network, referred to as the global preprocessing. The input distributions are flattened, scaled down to the interval $[-1, +1]$, and transformed to a Gaussian distribution with mean zero and width one to be out of saturation of the sigmoid function. Furthermore, the variables are decorrelated by transforming the covariance matrix to unit matrix so that the input variables are independent in first order. The transformation of the covariance matrix is performed by the Jacobi method [62].

Besides the global preprocessing the input variables can also be preprocessed individually. In doing so the distributions of the input variables are often replaced by spline fits suppressing statistical fluctuations. Sometimes it is also helpful to split discrete distributions into classes or to apply a delta function to variables which are not filled for all events.

The preprocessing algorithms of NeuroBayes® also calculate the significance of each input variable. The calculation is performed iteratively by removing one variable at a time from the set of n input variables and calculating the loss of total correlation to the target. Finally, the variable with the smallest loss is taken out and the procedure is repeated for the new set of $(n-1)$ input variables. Thus, after n iterations the significances of all input variables are determined and it is now possible to remove insignificant inputs in order to only keep the most significant variables in the neural network.

4.2.5 Regularisation

Regularisation of the neural network is done in favour of optimising the neural network training and avoiding overtraining. It reduces the number of free parameters by removing insignificant connections among the nodes bringing an improved signal to noise ratio in the data sample after the classification.

Another way of enhancing the training is to add an additional term to the cost function E, known as weight decay

$$E \rightarrow \tilde{E} = E + \frac{\tau}{2} \sum_i w_i^2. \tag{4.25}$$

The new cost function $\tilde{E}$ yields a neural network with smaller weights since the weights would decay exponentially in a training without any input data. The weight decay method suppresses oscillations around minima and reduces dependencies on random fluctuations. Consequently, the neural network is more robust.

4.2.6 Bayes Theorem

NeuroBayes® interprets the network output by the Bayes theorem considering conditional probabilities

$$P(H\,|D\,) = \frac{P(D\,|H\,) \cdot P(H)}{P(D)} \tag{4.26}$$

where D refers to data described by a hypothesis H. The Bayes theorem yields the a posteriori probability $P(H\,|D\,)$ for the hypothesis modelling the data correctly. $P(H)$ is the data independent a priori probability of the hypothesis H and $P(D)$ is the probability to measure the data D. $P(D\,|H\,)$ is the likelihood to measure the data D assuming the theory H.

The output of NeuroBayes® can be directly interpreted as a posteriori probability provided that the signal to background ratio is the same in the training sample and in the classification sample. However, if the ratio is different the neural network output can be transformed into the signal probability assuming that the ratio in the classification sample is known.

4.2.7 Training Result

The separation power of a neural network can be measured by two quantities, the purity and the efficiency.

- Purity P

$$P = \frac{N_S\left(n_{\text{out}} > n_{\text{cut}}\right)}{N_S\left(n_{\text{out}} > n_{\text{cut}}\right) + N_B\left(n_{\text{out}} > n_{\text{cut}}\right)} \tag{4.27}$$

- Efficiency ϵ

$$\epsilon = \frac{N_S\left(n_{\text{out}} > n_{\text{cut}}\right)}{N_S} \tag{4.28}$$

Here, n_{cut} is an selected cut on the network output n_{out} and N_S is the total number of signal events. $N_S\left(n_{\text{out}} > n_{\text{cut}}\right)$ are the remaining signal events after the cut on the network output is applied. The nomenclature for the background events N_B is analogous.

The purity states the signal to noise ratio of the selected events considered being signal events and the efficiency is the ratio of selected signal events to the total number of signal events. The aim is therefore to increase both purity and efficiency.

5 Candidate Selection

5.1 Candidate Reconstruction

In this thesis, a search is performed for B^{**} and B_s^{**} mesons in decays of $B^{**} \to B^{(*)+}\pi^-$ and $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$, respectively. The charge conjugated decays are implicitly included. To reconstruct the B^{**} candidates, three independent decay channels of the B^+ meson are used:

$$\begin{array}{lll}
B^{**} \to B^{(*)+}\pi^- & B^+ \to J/\psi K^+ & J/\psi \to \mu^+\mu^- \\
 & B^+ \to \bar{D}^0\pi^+ & \bar{D}^0 \to K^+\pi^- \\
 & B^+ \to \bar{D}^0\,3\pi^\pm & \bar{D}^0 \to K^+\pi^-
\end{array}$$

The B_s^{**} candidates are reconstructed using six different decay channels of the B_s meson. These decay modes are:

$$\begin{array}{lll}
B_s^{**} \to B_s^{(*)}\pi^+\pi^- & B_s \to D_s^-\pi^+ & D_s^- \to \bar{K}^*K^- \\
 & & D_s^- \to \phi\pi^- \\
 & & D_s^- \to 3\pi^\pm \\
 & B_s \to D_s^- 3\pi^\pm & D_s^- \to \bar{K}^*K^- \\
 & & D_s^- \to \phi\pi^- \\
 & & D_s^- \to 3\pi^\pm
\end{array}$$

The B^* and B_s^* mesons decay into $B\gamma$ and $B_s\gamma$, respectively. Since the photon cannot be detected by the CDF II detector the reconstructed masses of the B^* and B_s^* mesons are shifted downwards by the photon energy.

The candidate selection is based on a chain of neural networks. The first neural network in this chain selects either the B^+ or B_s candidate, respectively. The second neural network selects subsequently the B^{**} or B_s^{**} candidate. Before the neural networks can be applied, they have to be trained on a training sample containing signal and background. Here, Monte Carlo samples are used as signal pattern and data events are taken as background pattern.

Regarding the reconstruction and selection of the B^+ and B_s mesons, code from previous analyses being conceptually the same is reused here. This applies to the selection code for the B^+ meson in decays of $B^+ \to J/\psi K^+$ and $B^+ \to \bar{D}^0\pi^+$ which was developed in the analysis of B_s^{**} mesons in decays of $B_s^{**} \to B^+K^-$ [1]. It also applies to the selection code for the six B_s decays modes being implemented

Dataset	Trigger		Integrated luminosity [fb^{-1}]	
	Di-muon	Two track	Per dataset	Sum
0d	xpmmgd	h77jf0	0.6	0.6
0h	xpmmgh	h77jg0	0.5	1.0
0i (until Feb. 2006)	xpmmhi	h77jm0		
0i (after Feb. 2006)	xpmmhi	xbhdii	0.7	1.7

Table 5.1: Summary of the BStntuples datasets used for the B^{**} data samples.

in context of the B_s mixing analysis [65, 66, 67]. However, the development of the code for reconstructing and selecting the B^{**} and B_s^{**} candidates is part of this thesis as well as the reconstruction and selection code on the B^+ meson level for $B^+ \rightarrow \bar{D}^0\, 3\pi^\pm$ with $\bar{D}^0 \rightarrow K^+\pi^-$. The reconstruction code used in this thesis is based on the official CDF software version 6.1.4.

5.1.1 B^{**} Data Samples

The data samples for the B^{**} analysis are collected by the di-muon trigger and the two track trigger. To have small and handy data files for the analysis, flat ntuples are skimmed from the officially provided BStntuples. These BStntuples are centrally produced and are suited for many different analyses. Thus, they contain much more information than needed here yielding a nonpractically large file size.

The analysis uses data being collected between February 2002 and November 2006 corresponding to an integrated luminosity of 1.7 fb^{-1}. The data samples contain only runs selected by the good run list for B-physics. The good run list is provided by the DQM group [68] and comprises the run numbers of approved runs. Table 5.1 summarises the used dataset for the B^{**} data samples.

5.1.2 B^{**} Monte Carlo Samples

There are two sets of Monte Carlo samples used in the B^{**} analysis. The first one models decays of $B^+ \rightarrow \bar{D}^0\, 3\pi^\pm$ with $\bar{D}^0 \rightarrow K^+\pi^-$ and the second set comprises decays of $B^{**} \rightarrow B^{(*)+}\pi^-$ with $B^+ \rightarrow J/\psi K^+$, $B^+ \rightarrow \bar{D}^0\pi^+$ and $B^+ \rightarrow \bar{D}^0\, 3\pi^\pm$. The samples are signal-only Monte Carlo samples, that is they only model signal events without underlaying events and do not contain background events.

The Monte Carlo sample for the decay $B^+ \rightarrow \bar{D}^0\, 3\pi^\pm$ was generated with the BGenerator event generator [70]. It was produced by the B_s mixing group in context of the B_s oscillation analysis. The sample is available in form of BStntuples from which flat ntuples were skimmed for this analysis using the same skimming code as for the data samples.

The B^{**} Monte Carlo samples were exclusively generated for this analysis on the basis of the BGenerator event generator. A custom module was attached to the generator in order to have the B^{**} mass flatly distributed over the mass range being searched by this analysis. The decays of the generated particles were calculated using the EvtGen [71] package followed by a full detector and trigger simulation. Afterwards, the events were reconstructed by applying the standard reconstruction software including the BottomMods [69] package with the same settings as used for data reconstruction. The B^{**} Monte Carlo samples contain decays into $B^+\pi^-$ as well as into $B^{*+}\pi^-$. The complete decay tables of the samples are given in the appendix C.

Monte Carlo Event Reweighting

The Monte Carlo samples for the different decay modes of the B^{**} mesons are generated having a flat mass distribution. To avoid biasing the neural network during training the Monte Carlo events are reweighted to the same Q value distribution as in data. A distinct weight w_i is applied to each Monte Carlo event i having the Q value Q_i. The weights are calculated as

$$w_i(Q_i) = \frac{f_{\text{data}}(Q_i)}{f_{\text{MC}}(Q_i)} \tag{5.1}$$

where $f_{\text{data}}(Q)$ and $f_{\text{MC}}(Q)$ are fourth order polynomial functions describing the Q value distribution in data and in the Monte Carlo sample, respectively.

For the Monte Carlo sample an additional fourth order polynomial function is added below 45 MeV/c^2 to describe the bump in the Q value distribution. This bump arises from the fact that above 45 MeV/c^2 half of the events decay into $B^{*+}\pi^-$ which are shifted by $m(B^*)-m(B) = 45$ MeV/c^2 due to the unreconstructed photon from the decay $B^* \to B\gamma$. Below 45 MeV/c^2 all events decay into $B^+\pi^-$ and stay around the generated mass.

The function parameters for f_{data} and f_{MC} are determined by fits to the corresponding Q value distributions shown in figures 5.1 and 5.2. For each decay channel, a different set of parameters is obtained and used to calculate the weights for the events. Figure 5.3 displays the distribution of the weights for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$. The weights for the other decays $B^+ \to \bar{D}^0\pi^+$ and $B^+ \to \bar{D}^0\, 3\pi^\pm$ look very similar.

5.1.3 B_s^{**} Data Samples

The data samples used in the B_s^{**} analysis are collected by the two track trigger. From the two track trigger data, custom made BStntuples were created applying a

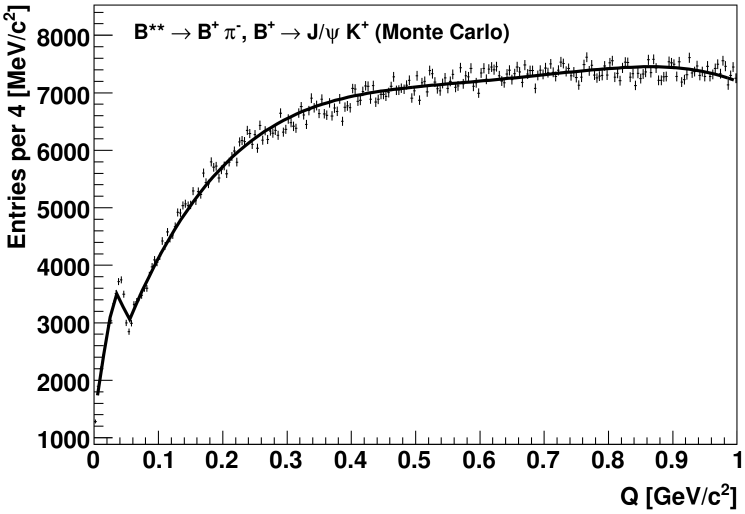


Figure 5.1: Fit to the Q value of the $B^{**} \rightarrow B^{(*)+}\pi^-$ Monte Carlo sample with $B^+ \rightarrow J/\psi K^+$ after having applied the preselection.

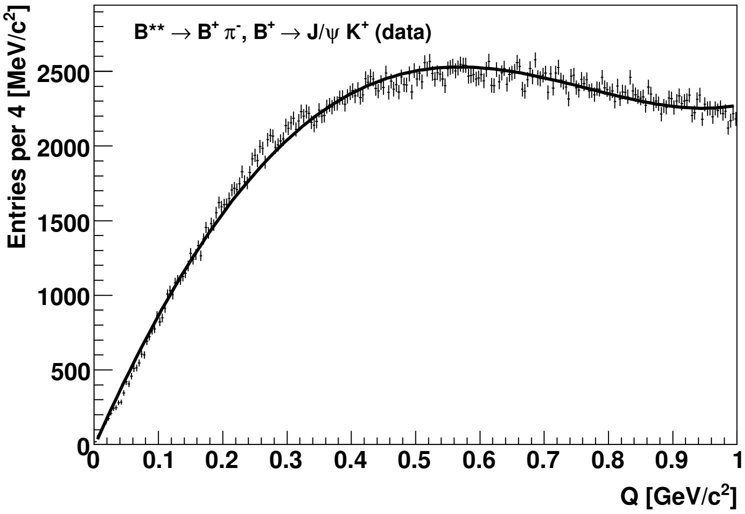


Figure 5.2: Fit to the Q value of the $B^{**} \rightarrow B^{(*)+}\pi^-$ datas ample with $B^+ \rightarrow J/\psi K^+$ after having applied the preselection.

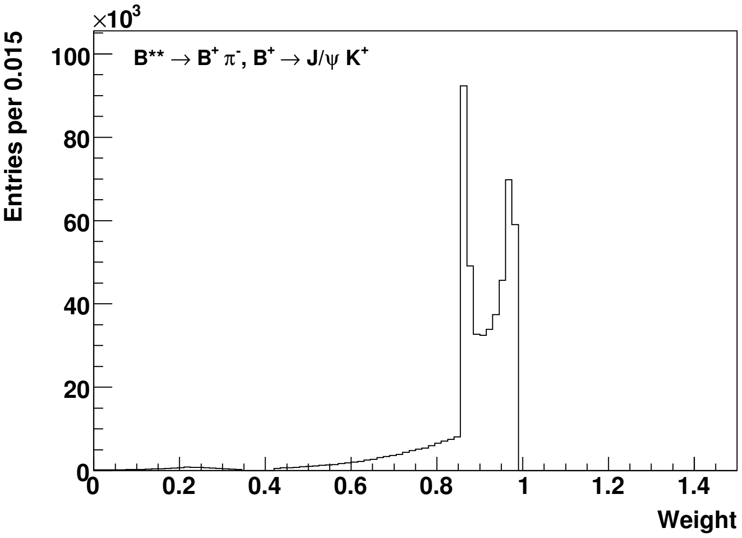


Figure 5.3: Distribution of the Monte Carlo weights for the $B^{**} \rightarrow B^{(*)+}\pi^-$ sample with $B^+ \rightarrow J/\psi K^+$.

full vertex fit to the $B_s\pi\pi$ combinations. The BStntuples were skimmed afterwards to obtain small and handy flat ntuples for the analysis. The data samples use as much data as available at the time they were created, that is the amount of data collected between February 2002 and February 2008. Only runs selected by the good run list for B-physics are put into the samples. Table 5.2 gives an overview of the used datasets corresponding to a total integrated luminosity of 2.8 fb^{-1}.

Dataset	Integrated luminosity [fb^{-1}]	
	Per dataset	Sum
0d	0.3	0.3
0h	0.4	0.7
0i	0.6	1.3
0j	1.0	2.3
0k	0.5	2.8

Table 5.2: Summary of the datasets used for the B_s^{**} data samples. The integrated luminosity takes into account the good run list for B-physics.

5.1.4 B_s^{**} Monte Carlo Samples

The B_s^{**} Monte Carlo samples are generated exclusively for this analysis on the basis of the BGenerator event generator. The simulation was done in a similar way as for the B^{**} Monte Carlo samples. That is, a custom module was attached to the generator in order to have the B_s^{**} mass distributed flatly. The decays of the generated particles were computed using the EvtGen package followed by a full detector and trigger simulation. Afterwards, the events are reconstructed with the same reconstruction software as used for the data reconstruction. The Monte Carlo samples contain decays into $B_s\pi^+\pi^-$ as well as into $B_s^*\pi^+\pi^-$. The complete decay tables are given in the appendix C.

Monte Carlo Event Reweighting

The low statistics of the data samples in each B_s^{**} decay mode makes it impossible to train independent neural networks for selecting the B_s^{**} candidates in each decay channel as it is described in section 5.2. Therefore, both the data samples and the Monte Carlo samples are combined in each case and one single neural network is trained to select the B_s^{**}. In order to avoid biasing the neural network training the Monte Carlo events are reweighted taking into account two physical conditions. Firstly, the Monte Carlo events are reweighted to the same Q value distribution as in data. Secondly, the mixture of decay channels is required in the combined Monte Carlo sample to corresponds to the mixture in data. Hence, the individual weight w_i being applied to each Monte Carlo event i having the Q value Q_i is

$$w_i(Q_i) = \frac{f_{\text{data}}(Q_i)}{f_{\text{MC}}(Q_i)} \cdot \frac{\text{BR}_{\text{PDG}}}{\text{BR}_{\text{MC}}}. \tag{5.2}$$

Here, $f_{\text{data}}(Q)$ and $f_{\text{MC}}(Q)$ are functions describing the Q value distributions in data and in Monte Carlo, respectively. Their ratio corrects for the Q value distributions. These functions are fourth order polynomials whose parameters are obtained from different fits to the Q value in data and in the Monte Carlo samples as seen in figures 5.4 and 5.5.

For the Monte Carlo sample an additional fourth order polynomial function is used below 60 MeV/c^2 to describe the bump in the Q value distribution. Similar to the B^{**} Monte Carlo sample, half of the events generated above 47 MeV/c^2 decay into $B_s^*\pi\pi$ which are shifted by $m(B_s^*) - m(B_s) = 47$ MeV/c^2 due to the undetected photon from the decay $B_s^* \to B_s\gamma$. The events below 47 MeV/c^2 decay only into $B_s\pi\pi$ and stay around the generated mass.

For each decay channel a distinct set of parameters is used. The second fraction in equation (5.2) corrects for the right channel mixture. Thereby, BR_{PDG} is the world average branching ratio [2] of the B_s decay in question and BR_{MC} is the corresponding

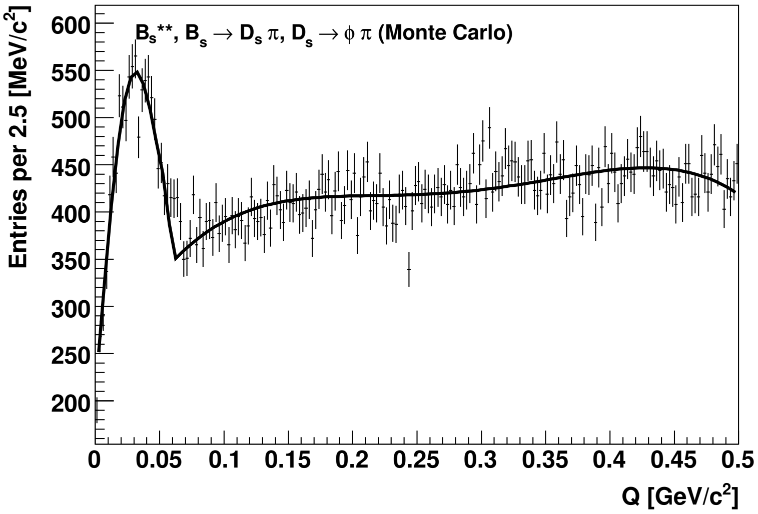


Figure 5.4: Fit to the Q value of the $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ Monte Carlo sample with $B_s \rightarrow D_s^-\pi^+$, $D_s^- \rightarrow \phi\pi^-$ after having applied the preselection.

branching ratio implemented in the decay tables being used to generate the Monte Carlo sample. Figure 5.6 displays the distribution of the Monte Carlo weights.

5.1.5 Candidate Reconstruction

The data samples used in this thesis are based on custom made and official BStntuples containing the B^{**} and B_s^{**} decays in question. Both types of BStntuples are conceptually the same and therefore the reconstruction of the different B^{**} and B_s^{**} mesons is done in a similar way. The nomenclature of $B_{(s)}$ in this section here refers to the B^+ and B_s mesons at the same time. The same is true for $B_{(s)}^{**}$.

The reconstruction process starts with the reconstruction of stable particles by fitting tracks and applying mass hypotheses to these tracks. The stable particles are combined to form higher level objects like D and B mesons. By adding further tracks to them the $B_{(s)}^{**}$ mesons are reconstructed finally. Since the reconstruction is conceptually the same for each decay mode it will be explained for the decay $B^{**} \rightarrow B^{(*)+}\pi^-$ with $B^+ \rightarrow J/\psi K^+$ as an example.

Two muons having opposite charge signs are combined in a vertex fit to form the J/ψ. In the fit the world average values is used for the muon mass. The so-formed J/ψ is then combined with a further track assumed to be a kaon yielding the B^+ candidate.

In order to reconstruct the $B_{(s)}^{**}$ candidates, tracks applying the pion mass hy-

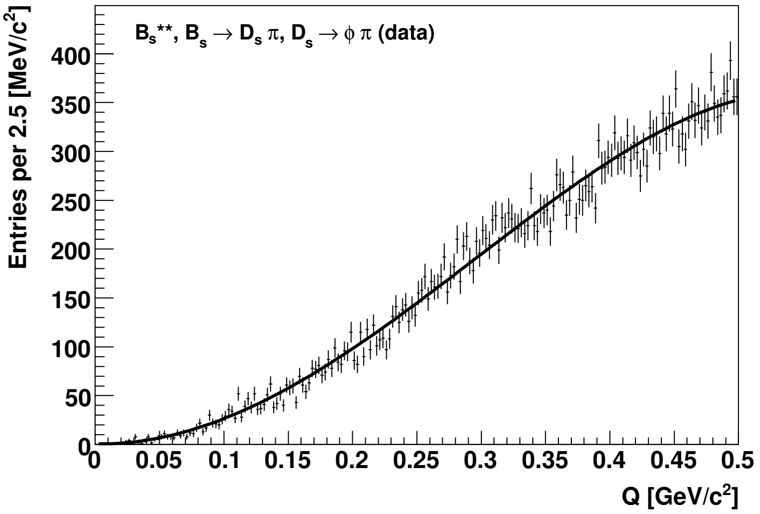


Figure 5.5: Fit to the Q value of the $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ data sample with $B_s \rightarrow D_s^-\pi^+$, $D_s^- \rightarrow \phi\pi^-$ after having applied the preselection.

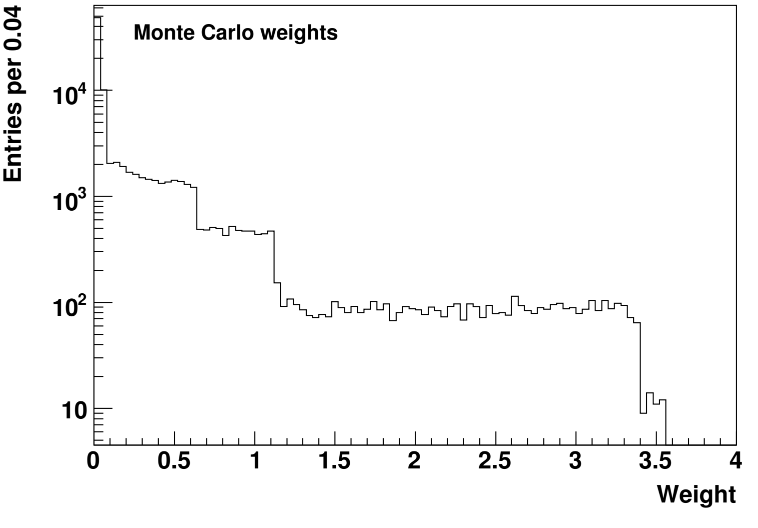


Figure 5.6: Distribution of the Monte Carlo weights for the $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ sample including all six decay modes. The steps arise from the correction factors for the right B_s decay channel mixture.

pothesis are combined with $B_{(s)}$ candidates. The invariant mass of the so-formed $B^{**}_{(s)}$ meson is calculated by the four-momenta of the $B_{(s)}$ candidate and the pion track(s).

To keep the background level as low as possible on this stage, a full vertex fit to the so far formed combinations is highly preferable. A vertex fit is a kinematic fit with appropriate topology constraints. However, a full vertex fit is a very time consuming step of the reconstruction procedure. Therefore, it is not applied to the B^{**} data samples having the advantage of a simplified data handling which allows adding new data rather quickly.

Among the decay modes of the $B^{**}_s \to B^{(*)}_s \pi^+\pi^-$ there are final states possible with up to eight pions. Hence, a higher combinatorial background level is expected here. For this reason, a vertex fit is performed on the $B_s\pi^+\pi^-$ combinations to reduce the background level. Of course, the computing time of data reconstruction increases and adding new data will take more time.

Minimal Pion Momentum

The two pions from the decay $B^{**}_s \to B^{(*)}_s \pi^+\pi^-$ are expect to have low momenta. This is due to the low Q value of the B^{**}_s meson. Generator level Monte Carlo samples were used for estimating the expected pion momentum distribution. The samples were produced based on the EvtGen package [71] assuming the different masses 5.8, 5.84, 5.88 and 5.9 GeV/c^2 for the B^{**}_s. For each sample 100000 events were generated. Figure 5.7 shows the minimal pion momentum distributions of these samples for the pions from the decay $B^{**}_s \to B^{(*)}_s \pi^+\pi^-$.

For CDF analyses an established quality cut is usually put on the transverse pion momenta at 400 MeV/c. Below this value the background level rises drastically and moreover the Monte Carlo samples used for modelling the signal are no more reliable. Consequently the cut at 400 MeV/c is also applied to the B^{**}_s data samples although the transverse pion momentum distribution is mainly located below 400 MeV/c.

5.2 Outline of the Candidate Selection

The candidate selection of the B^{**} and B^{**}_s mesons in decays of $B^{**} \to B^{(*)+}\pi^-$ and $B^{**}_s \to B^{(*)}_s \pi^+\pi^-$, respectively, is based on a chain of neural networks organised in two steps. At the first step, for each B^+ and B_s decay mode a distinct neural network is applied to select the B^+ and B_s signal candidates. At the second step, neural networks are trained to select the B^{**} and B^{**}_s candidates. For the B^{**} selection, three independent neural networks are used, one single network for each decay channel. Due to the insufficient low statistics in the B^{**}_s data samples as listed in table 5.3 it is not possible to have one single neural network for each B^{**}_s decay

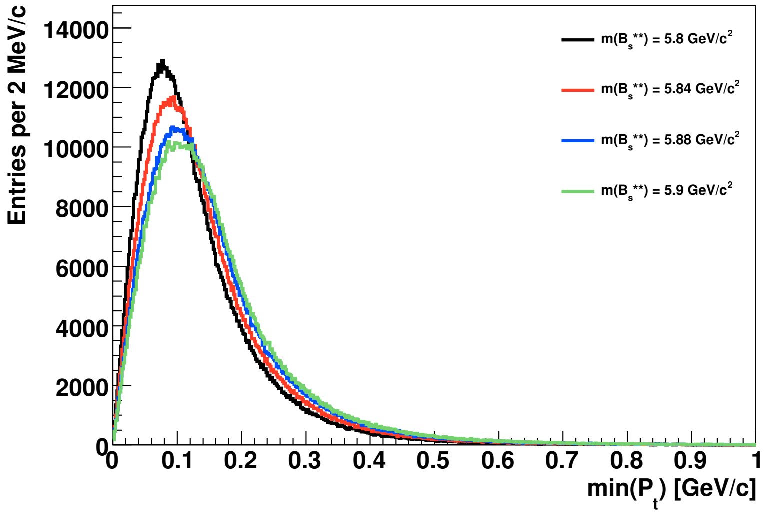


Figure 5.7: Minimal transverse pion momentum of the pions from the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$. The distributions are generated assuming different masses for the B_s^{**}.

mode. For the neural network training, however, a rather large number of events is necessary to properly adapt the weights of all nodes in the neural network. For this reason, one single neural network is trained acting on the combined data sample of all six decay modes.

On the level of the B^+ and B_s meson selection, there is no optimised cut applied on the output of the corresponding neural network, except for a soft precut in order to remove a large amount of background events. The outputs of the B^+ and B_s neural networks are rather used as inputs to the corresponding neural network for

$B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ Decay		Monte Carlo	Data
$B_s \to D_s^-\pi^+$	$D_s^- \to \bar{K}^*K^-$	1395	2624
	$D_s^- \to \phi\pi^-$	26201	783
	$D_s^- \to 3\pi^\pm$	23024	1777
$B_s \to D_s^- 3\pi^\pm$	$D_s^- \to \bar{K}^*K^-$	7857	6736
	$D_s^- \to \phi\pi^-$	16926	3208
	$D_s^- \to 3\pi^\pm$	15484	9410
	Sum	90887	24538

Table 5.3: Number of Monte Carlo and data events of the B_s^{**} samples after having applied the preselection.

selecting the B^{**} and B_s^{**} candidates, respectively. Thus, the B^{**} and B_s^{**} neural networks act on events preselected by the B^+ and B_s neural networks. Consequently, the second B^{**} and B_s^{**} neural networks are the main selection tools and the final candidate selection is done by cutting on the output of these neural networks.

All neural networks used in this analysis are taken from the NeuroBayes® package [63, 64]. The networks are trained on Monte Carlo events as signal pattern and real data events as background pattern. For simulating background events, the full quark production and hadronisation process has to be modelled. However, the details of these processes are not understood with sufficient precision. So far, background models are therefore not capable of describing the background satisfactorily enough. A better way of getting background events is achieved by taking data events from the mass sidebands. This course of action saves computing time and yields samples containing all sorts of background events at the same time.

At each stage of the selection procedure, a soft preselection is done before the corresponding neural network is applied. These high signal efficiency requirements reduce the background level by removing a large part of background events but only a marginal fraction of signal events.

In order to evaluate the performance of the neural networks, so-called quality plots are produced during the network training. The distribution of the signal and background as a function of the neural network output demonstrates the selection power of the neural network. The purity of the neural network is also shown for each training which should be a linear function of the network output. Both are characteristics of well-trained neural networks. These plots together with detailed information about the different neural networks developed in the context of this thesis are presented in the following sections.

For each neural network the correlation matrix is also presented showing that there are correlations between the input variables. These correlations emphasise the fact that an improved signal and background separation can be expected by taking into account correlations among the input variables in the selection procedure which holds by using neural networks. The variables shown in the correlation matrix have the same order as they are put into the neural network but with a different index number. The index i of each column or row in the correlation matrix corresponds to the variable $(i-1)$ in the variable list. Index 1 in the matrix denotes the target variable.

The list of input variables given for each neural network also contains the significance of every input variable for the neural network. Based on this significance the rank of the variables is also listed. The symbols for the variables and the naming scheme of the particles are defined in appendix A.

Variable	Unit	Preselection requirement
$L_{xy}/\sigma_{L_{xy}}(B^+)$		> 7.5
$\lvert d_0(B^+)\rvert$	[cm]	> 0.0075
$\ell_{xy}/\sigma_{\ell_{xy}}(D^0)$		> -4
$p_t(D^0)$	[GeV/c]	> 1.5
$\lvert d_0(D^0)\rvert$	[cm]	> 0.002
$p_t(K_{D^0})$	[MeV/c]	> 400
—		Correct charge combination of the π_{D^0} and K_{D^0}

Table 5.4: Preselection requirements for $B^+ \rightarrow \bar{D}^0\, 3\pi^\pm$.

5.3 B^{**} Candidate Selection

5.3.1 Selection of the B^+ with $B^+ \rightarrow \bar{D}^0\, 3\pi^\pm$

The selection of the B^+ using $B^+ \rightarrow \bar{D}^0\, 3\pi^\pm$ starts by applying soft preselection requirements to the data samples which are listed in table 5.4.

For the training of the neural network Monte Carlo events are used as signal pattern and data from the upper B^+ mass sideband in the range from 5.325 GeV/c^2 to 5.395 GeV/c^2 is used as background pattern. More satisfying would be to use data from both the lower and upper sideband as background pattern, but since the lower sideband also contains partially reconstructed B^+ mesons, only the upper sideband can be considered containing rather pure background events. The neural network inputs are given in table 5.5.

The quality plots obtained from the neural network training demonstrate the training success. Figure 5.8 shows the distributions of the signal and background events and the purity of the neural network is shown in figure 5.9. The correlation matrix of the input variables is plotted in figure 5.10.

The invariant mass distribution of the B^+ candidates selected by requiring the neural network output larger than 0.9 is shown in figure 5.11.

5.3.2 Selection of the $B^{**} \rightarrow B^{(*)+}\pi^-$

For the selection of the B^{**} candidates three different neural networks are used for the B^+ decays $B^+ \rightarrow J/\psi K^+$, $B^+ \rightarrow \bar{D}^0\pi^+$ and $B^+ \rightarrow \bar{D}^0\, 3\pi^\pm$. The preselection requirements applied before the neural network training are listed in table 5.6.

For the neural network training Monte Carlo samples are used as signal pattern and data events taken from the Q value range from 0 to 1.0 GeV/c^2 are used as background pattern. After having applied the preselection requirements the data sample is still dominated by background as seen in figures 5.12, 5.13 and 5.14. Therefore, the full Q value range can be well used as background pattern for the neural network

Index	Significance	Rank	Variable
1	33	7	$\lvert d_0(B^+)\rvert$
2	136.5	2	$P(B^+)$
3	2.1	29	$L_{xy}/\sigma_{L_{xy}}(B^+)$
4	15.1	17	$p_t(D^0)$
5	5.5	27	$\left\lvert d_0(D^0)\right\rvert$
6	8.8	25	$P(D^0)$
7	18.5	12	$L_{xy}/\sigma_{L_{xy}}(D^0)$
8	27.9	9	$\ell_{xy}/\sigma_{\ell_{xy}}(D^0)$
9	28.4	8	$p_t(\pi_{B^+}^{(1)})$
10	8	26	$p_t(\pi_{B^+}^{(2)})$
11	15.9	15	$p_t(\pi_{B^+}^{(3)})$
12	43.1	6	$d_0^{lts}/\sigma_{d_0}(\pi_{B^+}^{(1)})$
13	77.3	3	$d_0^{lts}/\sigma_{d_0}(\pi_{B^+}^{(2)})$
14	53.2	4	$d_0^{lts}/\sigma_{d_0}(\pi_{B^+}^{(3)})$
15	19.6	13	$\theta^*(\pi_{B^+}^{(1)})$
16	19.3	14	$\theta^*(\pi_{B^+}^{(2)})$
17	13.6	18	$\theta^*(\pi_{B^+}^{(3)})$
18	8.8	23	$\theta_{\mathrm{hel}}(\pi_{D^0})$
19	9.3	24	$p_t(K_{D^0};<)$
20	9.3	22	$p_t(K_{D^0};>)$
21	50	5	$L_K(K_{D^0};\mathrm{TOF})$
22	21.8	11	$L_K(K_{D^0};\overline{\mathrm{TOF}})$
23	4	28	$m\left(\pi_{B^+}^{(1)},\pi_{B^+}^{(2)}\right)$
24	23.6	10	$m_{\min}\left(\pi_{B^+}^{(1,2)},\pi_{B^+}^{(3)}\right)$
25	15.5	16	$m_{\max}\left(\pi_{B^+}^{(1,2)},\pi_{B^+}^{(3)}\right)$
26	10.1	20	$\min_{B^+}(\lvert d_0\rvert)$
27	11.3	19	$\max_{B^+}(\lvert d_0\rvert)$
28	9.8	21	θ_{a_1}
29	232.3	1	$m_{\pi^+\pi^+\pi^-}$

Table 5.5: List of input variables of the neural network for $B^+ \to \bar{D}^0\,3\pi^{\pm}$.

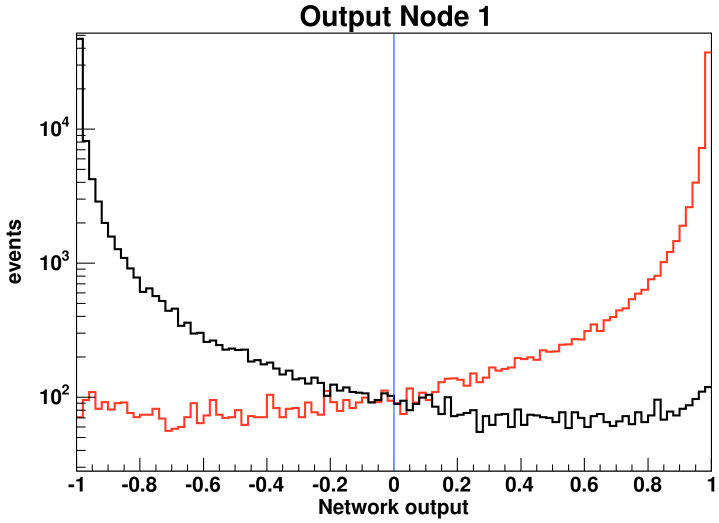


Figure 5.8: The distribution of the neural network output for signal (red) and background (black) events for the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$ using the training sample.

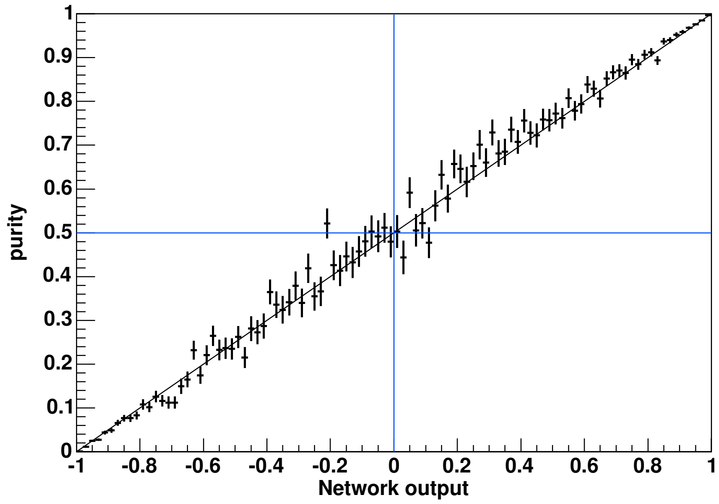


Figure 5.9: The purity as a function of the neural network output for the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$ using the training sample.

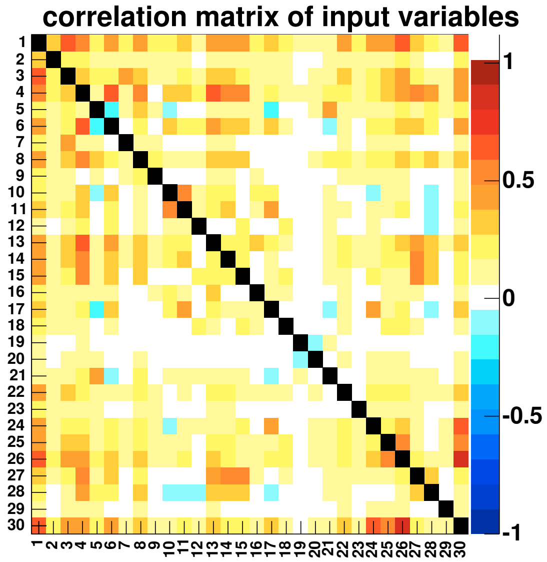


Figure 5.10: The correlation matrix of the input variables of the neural network for the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$.

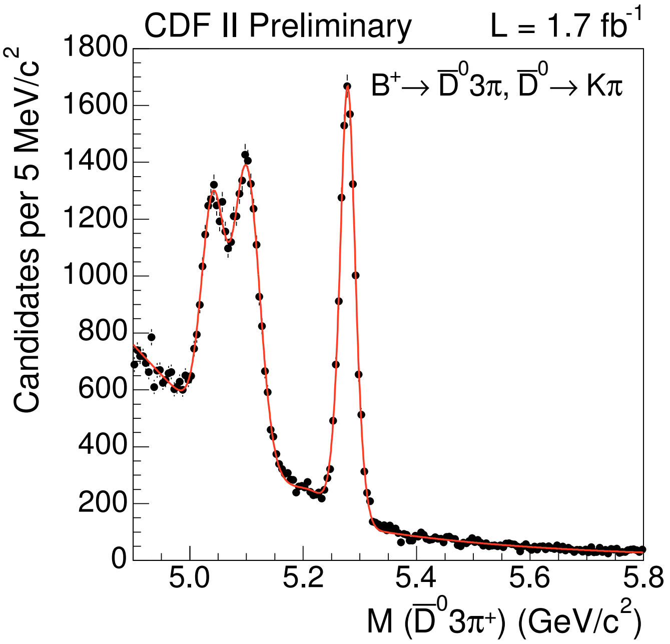


Figure 5.11: The invariant mass distribution of the B^+ in decays of $B^+ \to \bar{D}^0\, 3\pi^\pm$ selected by the best cut on the neural network output of 0.9.

Variable	Unit	Preselection requirement		
		$B^+ \to J/\psi K^+$	$B^+ \to \bar{D}^0\pi^+$	$B^+ \to \bar{D}^0\, 3\pi^\pm$
$n_{\mathrm{NN}}(B^+)$		> 0.5	> -0.5	> -0.5
$m(B^+)$	[GeV/c^2]	$5.24 - 5.31$	$5.24 - 5.31$	$5.24 - 5.35$
$\lvert d_0(B^+)\rvert$	[cm]	—	< 0.07	< 0.01
$L_{xy}/\sigma_{L_{xy}}(B^+)$		—	> 8	> 7
$\lvert d_0(D^0)\rvert$	[cm]	—	< 0.1	< 0.2
$\ell_{xy}/\sigma_{\ell_{xy}}(D^0)$		—	> -4	> 4
$p_t(\pi_{B^{**}})$	[GeV/c]	> 0.4	—	—
$d_0^{lts}/\sigma_{d_0}(\pi_{B^+})$		—	> -5	–
Q	[GeV/c^2]	—	$0 - 1.0$	$0 - 1.0$
—		Correct charge combination		

Table 5.6: Preselection requirements for $B^{**} \to B^{(*)+}\pi^-$ in the different B^+ decay channels.

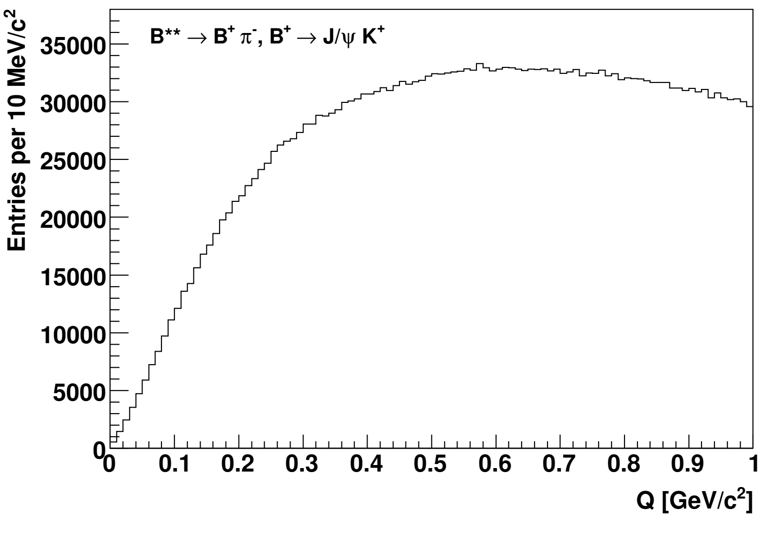


Figure 5.12: The Q value distribution from the B^{**} data sample with $B^+ \to J/\psi K^+$ being used for the neural network training as background pattern.

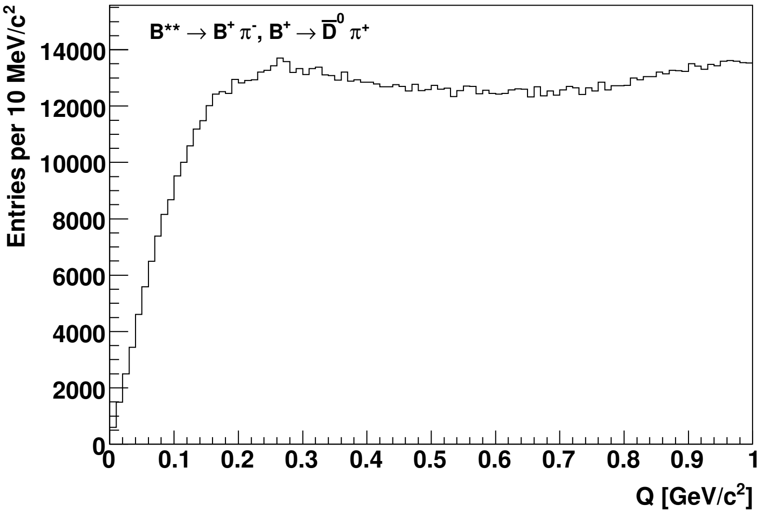


Figure 5.13: The Q value distribution from the B^{**} data sample with $B^+ \to \bar{D}^0\pi^+$ being used for the neural network training as background pattern.

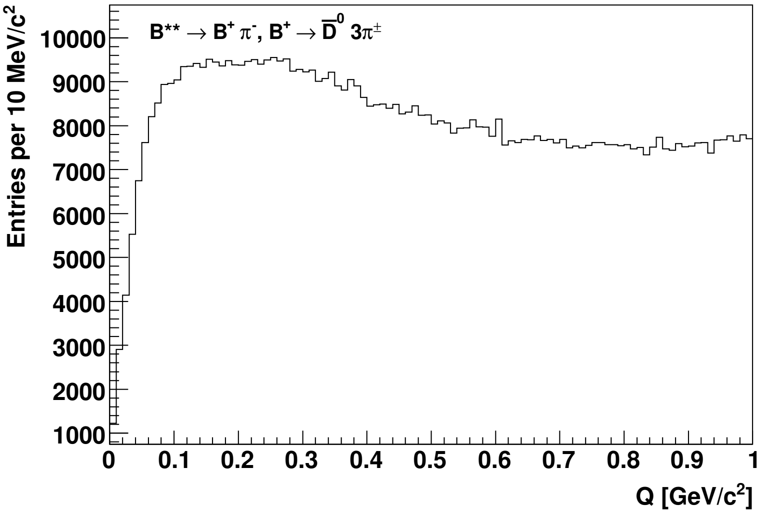


Figure 5.14: The Q value distribution from the B^{**} data sample with $B^+ \to \bar{D}^0\,3\pi^\pm$ being used for the neural network training as background pattern.

Index	Significance	Rank	Used	Variable
1	34.6	9	✓	$\lvert d_0(B^+)\rvert$
2	40.3	7	✓	$P(B^+)$
3	14.9	14	✓	$L_{xy}/\sigma_{L_{xy}}(B^+)$
4	21.3	11	✓	$\lvert d_0(J/\psi)\rvert$
5	11.3	15	✓	$d_0/\sigma_{d_0}(J/\psi)$
6	6.6	18	✓	$\ell_{xy}/\sigma_{\ell_{xy}}(J/\psi)$
7	36.7	8	✓	$p_t(K^+)$
8	87.1	5	✓	$\lvert m(B^+) - m_{\mathrm{PDG}}\rvert$
9	17.1	12	✓	$p_t(B^{**})$
10	34.2	10	✓	$\theta^*(\pi_{B^{**}})$
11	7.4	17	✓	$\theta^*(B^+)$
12	459.4	1	✓	$n_{\mathrm{NN}}(B^+)$
13	290.8	2	✓	$p_t(\pi_{B^{**}})$
14	185.9	3	✓	$d_0/\sigma_{d_0}(\pi_{B^{**}})$
15	63.5	6	✓	$\lvert d_0(\pi_{B^{**}})\rvert$
16	3.8	19	✓	$\eta(B^{**})$
17	8.5	16	✓	$\Delta R(B^{**}, \pi_{B^{**}})$
18	16.1	13	✓	$L_\pi(\pi_{B^{**}}; \mathrm{TOF})$
19	123.6	4	✓	$L_\pi(\pi_{B^{**}}; \overline{\mathrm{TOF}})$

Table 5.7: List of input variables of the neural network for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$.

training. The Monte Carlo events are reweighted to the same Q value distribution as in data in order to avoid biasing the training as described in section 5.1.2.

Tables 5.7, 5.8 and 5.9 list the input variables of the different neural networks. Variables which have a significance less than 3σ are automatically removed by the preprocessing algorithms of NeuroBayes® and are not used in the neural network.

The quality plots produced during the trainings indicate the success of the network trainings. Figures 5.15, 5.16 and 5.17 show the signal and background distributions and figures 5.18, 5.19 and 5.20 display the purity as a function of the neural network output. In figures 5.21, 5.22 and 5.23 the correlation matrices are plotted.

5.3.3 Cut Optimisation

The selection of the B^{**} candidates is done by applying cuts on two quantities: The output of the neural network trained for the specific decay channel and the number of candidates per event.

In order to make a decision about the best cut on the network output the signal significance S_{NN} has to be considered being a function of the cut on the network

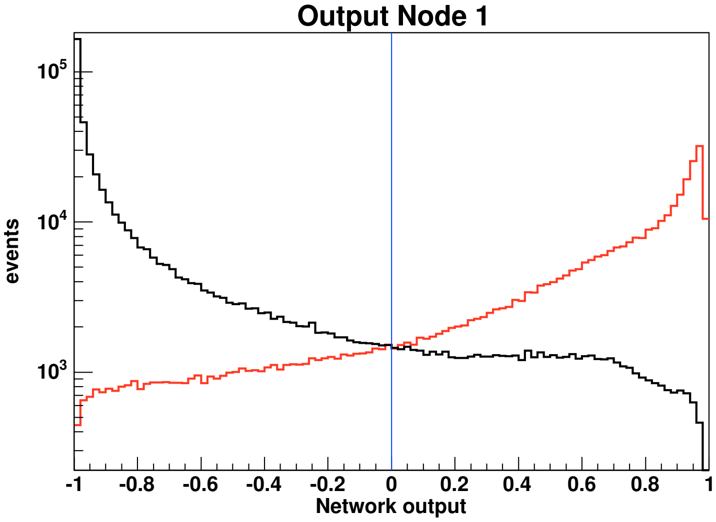


Figure 5.15: The distribution of the neural network output for signal (red) and background (black) events for the decay $B^{**} \rightarrow B^{(*)+}\pi^-$ with $B^+ \rightarrow J/\psi K^+$ using the training sample.

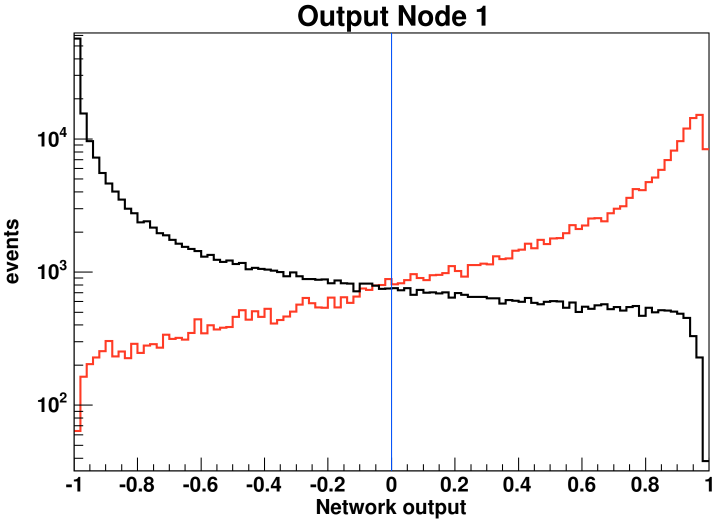


Figure 5.16: The distribution of the neural network output for signal (red) and background (black) events for the decay $B^{**} \rightarrow B^{(*)+}\pi^-$ with $B^+ \rightarrow \bar{D}^0\pi^+$ using the training sample.

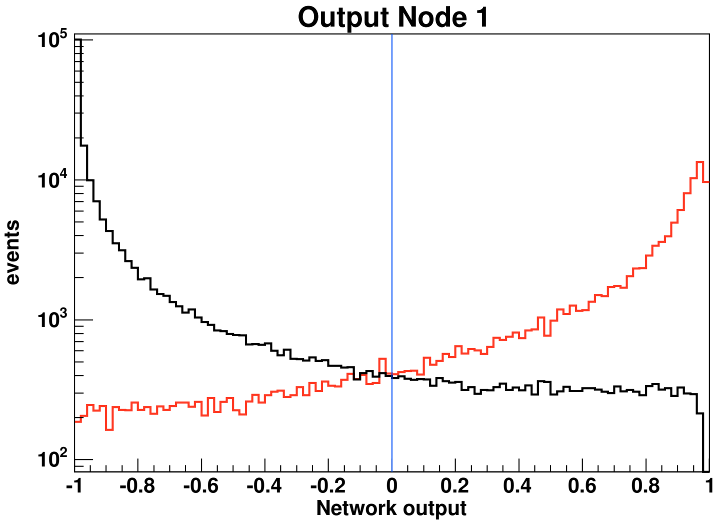


Figure 5.17: The distribution of the neural network output for signal (red) and background (black) events for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\, 3\pi^\pm$ using the training sample.

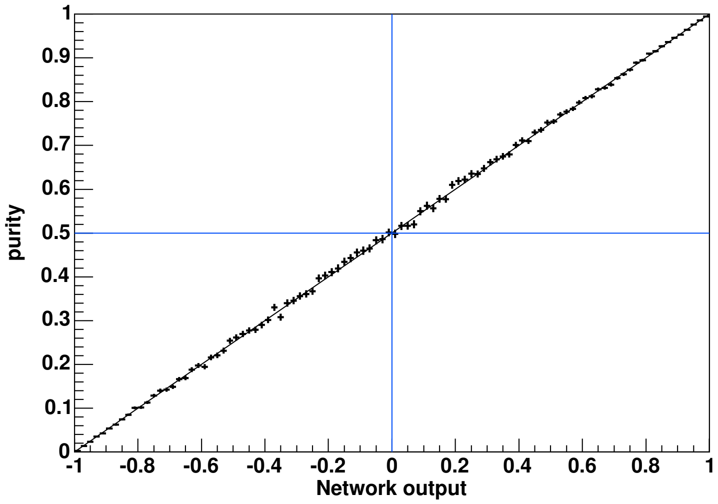


Figure 5.18: The purity as a function of the neural network output for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$ using the training sample.

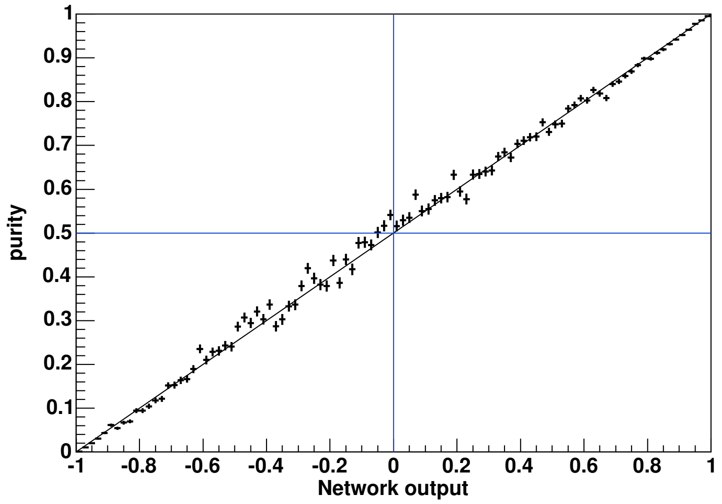


Figure 5.19: The purity as a function of the neural network output for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\pi^+$ using the training sample.

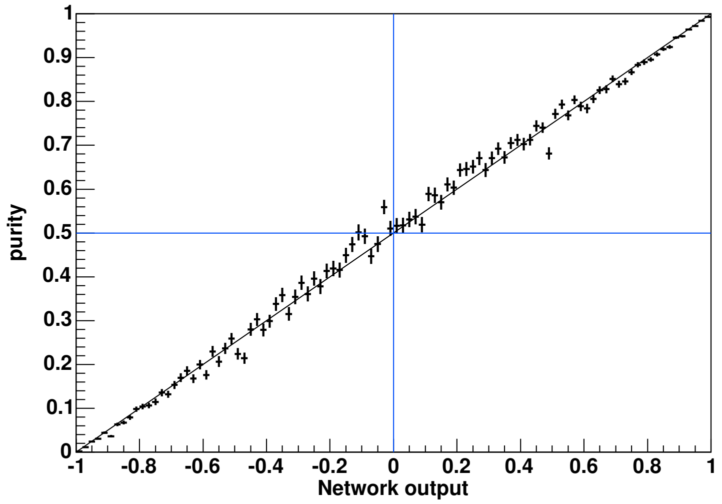


Figure 5.20: The purity as a function of the neural network output for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\,3\pi^{\pm}$ using the training sample.

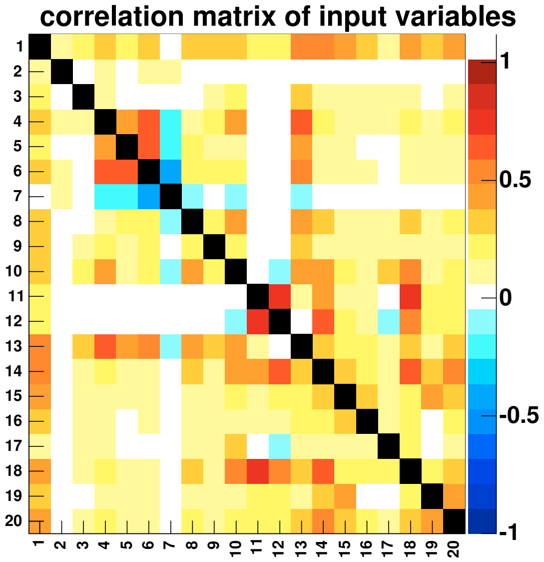


Figure 5.21: Correlation matrix of the neural network for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$.

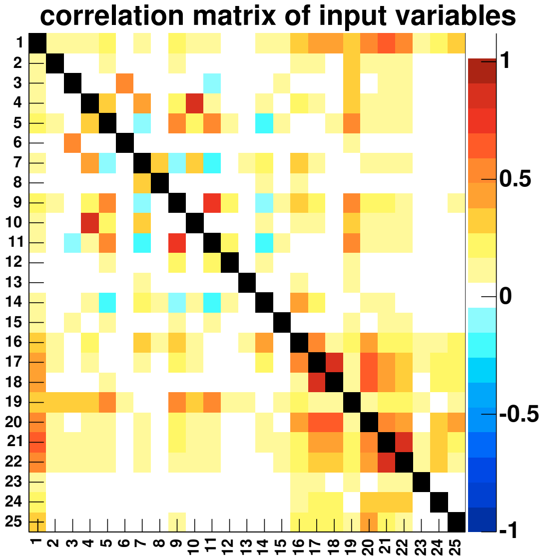


Figure 5.22: Correlation matrix of the neural network for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\pi^+$.

Index	Significance	Rank	Used	Variable
1	9.7	18	✓	$\lvert d_0(B^+)\rvert$
2	12.5	17	✓	$P(B^+)$
3	46.7	5	✓	$L_{xy}/\sigma_{L_{xy}}(B^+)$
4	16.4	13	✓	$\left\lvert d_0(D^0)\right\rvert$
5	7.5	20	✓	$P(D^0)$
6	31.7	10	✓	$L_{xy}/\sigma_{L_{xy}}(D^0)$
7	5.5	21	✓	$\ell_{xy}/\sigma_{\ell_{xy}}(D^0)$
8	40.4	9	✓	$p_t(\pi_{B^+})$
9	20.3	4	✓	$d_0^{lts}/\sigma_{d_0}(\pi_{B^+})$
10	8.2	19	✓	$\theta^*(\pi_{B^+})$
11	0.4	24	—	$\theta_{\text{hel}}(\pi_{D^0})$
12	1.6	23	—	$p_t(K_{D^0};<)$
13	3.8	22	✓	$p_t(K_{D^0};>)$
14	23.2	11	✓	$\lvert m(B^+)-m_{\text{PDG}}\rvert$
15	12.9	16	✓	$p_t(B^{**})$
16	46.6	6	✓	$\Delta R(B^+,\pi_{B^{**}})$
17	11.1	15	✓	$\theta^*(\pi_{B^{**}})$
18	88.1	3	✓	$n_{\text{NN}}(B^+)$
19	182	2	✓	$p_t(\pi_{B^{**}})$
20	41.4	8	✓	$\lvert d_0(\pi_{B^{**}})\rvert$
21	315.4	1	✓	$d_0/\sigma_{d_0}(\pi_{B^{**}})$
22	21	12	✓	$\eta(B^{**})$
23	12.1	14	✓	$L_\pi(\pi_{B^{**}};\text{TOF})$
24	47.1	7	✓	$L_\pi(\pi_{B^{**}};\overline{\text{TOF}})$

Table 5.8: List of input variables of the neural network for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\pi^+$.

output:

$$S_{\text{NN}} = \frac{N_{\text{MC}}^{(\text{cand})}\left(n_{\text{out}} > n_{\text{cut}}\right)}{\sqrt{N_{\text{data}}^{(\text{cand})}\left(n_{\text{out}} > n_{\text{cut}}\right)}}. \tag{5.3}$$

Here, $N_{\text{MC}}^{(\text{cand})}$ is the number of candidates in the Monte Carlo sample at a given cut n_{cut} on the network output n_{out}. Analogically, $N_{\text{data}}^{(\text{cand})}$ is the number of candidates selected in the data sample with the same cut on the network output. The best cut on the network output maximises the significance. The number of Monte Carlo events is regarded to be proportional to the number of signal events in data whereas the number of data events is the sum of background and signal events in data independently of the actual ratio. Since the neural networks for selecting the B^{**} candidates in different decay channels are independent, the optimisation of the selection cut is also performed independently for each decay mode. In order to calculate the significance, Monte Carlo and data events are counted in a Q value window where the signals are expected. Therefore, a Q value range from 0.2 GeV/c^2 to 0.4 GeV/c^2 is

Index	Significance	Rank	Used	Variable
1	3.9	23	✓	$\lvert d_0(B^+)\rvert$
2	5.6	19	✓	$P(B^+)$
3	6.9	15	✓	$L_{xy}/\sigma_{L_{xy}}(B^+)$
4	9	14	✓	$\big\lvert d_0(D^0)\big\rvert$
5	0.3	30	—	$P(D^0)$
6	15.5	9	✓	$L_{xy}/\sigma_{L_{xy}}(D^0)$
7	5	20	✓	$\ell_{xy}/\sigma_{\ell_{xy}}(D^0)$
8	2.1	28	—	$p_t(\pi^{(1)}_{B^+})$
9	3.3	25	✓	$d_0^{lts}/\sigma_{d_0}(\pi^{(1)}_{B^+})$
10	10.7	11	✓	$\theta^*(\pi^{(1)}_{B^+})$
11	2.9	26	—	$p_t(\pi^{(2)}_{B^+})$
12	4.4	21	✓	$d_0^{lts}/\sigma_{d_0}(\pi^{(2)}_{B^+})$
13	7.3	13	✓	$\theta^*(\pi^{(2)}_{B^+})$
14	0.7	29	—	$p_t(\pi^{(3)}_{B^+})$
15	8.2	16	✓	$d_0^{lts}/\sigma_{d_0}(\pi^{(3)}_{B^+})$
16	5.8	18	✓	$\theta^*(\pi^{(3)}_{B^+})$
17	12.7	10	✓	$\theta_{\mathrm{hel}}(\pi_{D^0})$
18	4.3	22	✓	$p_t(K_{D^0};<)$
19	6.4	17	✓	$p_t(K_{D^0};>)$
20	92.5	3	✓	$\lvert m(B^+) - m_{\mathrm{PDG}}\rvert$
21	1.9	27	—	$p_t(B^{**})$
22	27.1	7	✓	$\Delta R(B^+, \pi_{B^{**}})$
23	9.3	12	✓	$\theta^*(\pi_{B^{**}})$
24	184.6	2	✓	$n_{\mathrm{NN}}(B^+)$
25	88	4	✓	$p_t(\pi_{B^{**}})$
26	311.1	1	✓	$\lvert d_0(\pi_{B^{**}})\rvert$
27	32.5	6	✓	$d_0/\sigma_{d_0}(\pi_{B^{**}})$
28	21.3	8	✓	$\eta(B^{**})$
29	3.3	24	✓	$L_\pi(\pi_{B^{**}};\mathrm{TOF})$
30	37.2	5	✓	$L_\pi(\pi_{B^{**}};\overline{\mathrm{TOF}})$

Table 5.9: List of input variables of the neural network for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\,3\pi^\pm$.

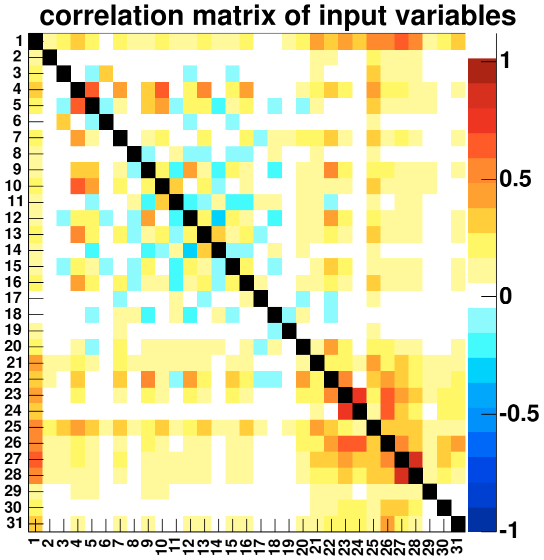


Figure 5.23: Correlation matrix of the neural network for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\, 3\pi^\pm$.

chosen for all three decay channels. Within this window, the Monte Carlo and data events are counted at different cuts on the output of the neural network.

Figures 5.24, 5.25 and 5.26 show the significance for each decay channel as a function of the cut on the network output. Based on these the best cuts were determined to be 0.5 for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$, 0.7 for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\pi^+$ and 0.75 for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\, 3\pi^\pm$.

For measuring the masses of the B^{**} states a fit on the combined data of all three decay channels is performed. Therefore, a cut optimisation is required maximising the overall significance. This significance is a function of all three neural network outputs. Thus, the cut on the neural network outputs are optimised simultaneously when the overall significance is maximised being defined by

$$S_{\mathrm{NN}} = \frac{\sum_i \left(w_i \cdot N^{(\mathrm{cand})}_{\mathrm{MC},i}\, (n_{\mathrm{out},i} > n_{\mathrm{cut},i}) \right)}{\sqrt{\sum_i N^{(\mathrm{cand})}_{\mathrm{data},i}\, (n_{\mathrm{out},i} > n_{\mathrm{cut},i})}}. \tag{5.4}$$

Here, $N^{(\mathrm{cand})}_{\mathrm{MC},i}$ and $N^{(\mathrm{cand})}_{\mathrm{data}}$ have the same meaning as in equation (5.3). The parameter i runs over all three decay modes. To take into account the ratio of generated Monte Carlo events among the different decay channels the number of Monte Carlo events is weighted by the factor w_i. This weight is the ratio of the number of Monte Carlo

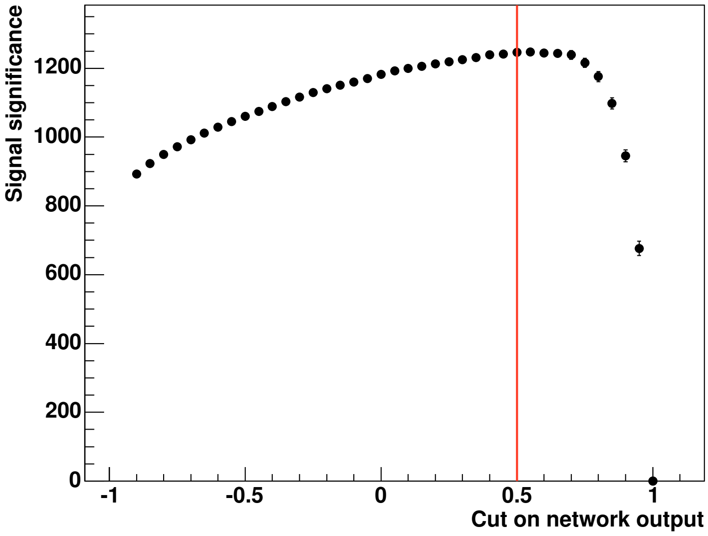


Figure 5.24: The significance over the cut on the network output for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$. The red line denotes where the selection cut is put.

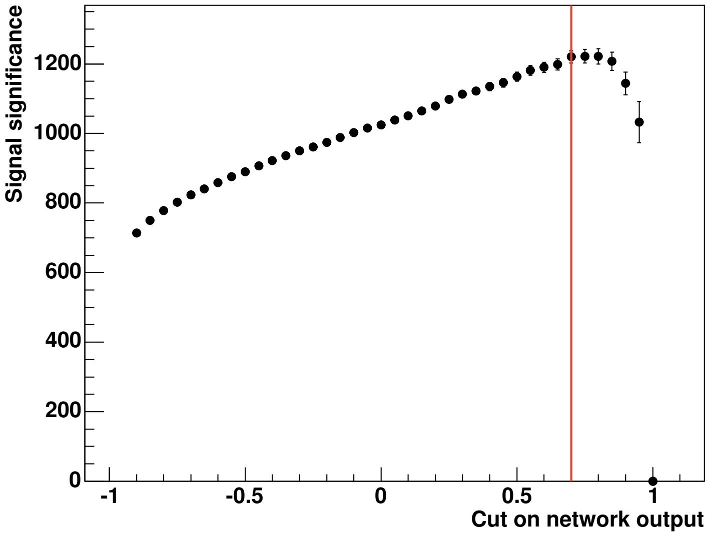


Figure 5.25: The significance over the cut on the network output for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\pi^+$. The red line denotes where the selection cut is put.

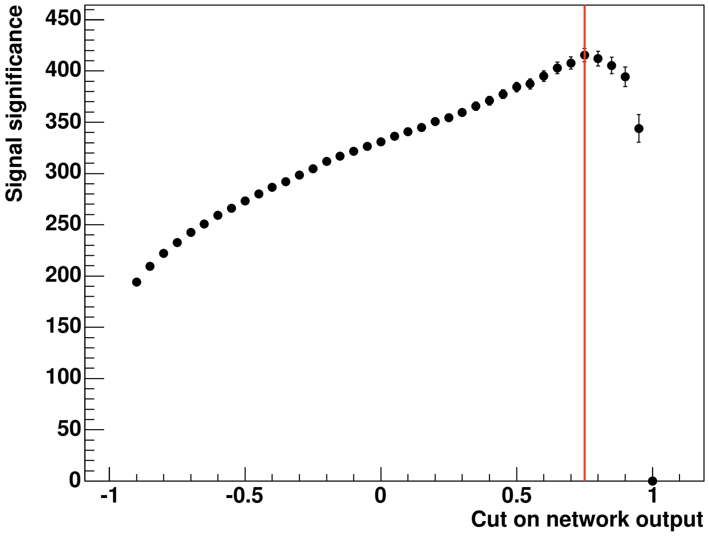


Figure 5.26: The significance over the cut on the network output for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\,3\pi^\pm$. The red line denotes where the selection cut is put.

events and signal events at the network cut $n_{\mathrm{out}} > 0$ in each decay channel. For the significance calculation, Monte Carlo and data events are counted in a Q value window from 0.2 GeV/c^2 to 0.4 GeV/c^2. In order to calculate the significance a three-dimensional scan in steps of 0.05 on the three different neural network outputs is done.

Figure 5.27 shows the combined significance with different cuts on one neural network output while the cuts on the other two network outputs are kept at their best values. The best selection cuts are determined to be 0.85 for the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$, 0.7 for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\pi^+$ and 0.75 for $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\,3\pi^\pm$.

For the final selection, a cut is also applied on the number of candidates per event. As there can be at most two B^{**} candidates per event it is more likely to find a real B^{**} meson in events having a low number of candidates per event. Since the number of candidates is not modelled in the Monte Carlo samples, the selection cut on this quantity is applied after the significance optimisation. For this reason, there is no way to optimise this cut to its best value. Instead of that, the cut is put at a rather arbitrary but unbiased value since it was done a priori without looking at the obtained signals. For the B^{**} selection, the number of candidates per event has to be less than 6. Figure 5.28 displays the Q value distribution for all three decay channels after applying the final selection cuts obtained from the overall significance

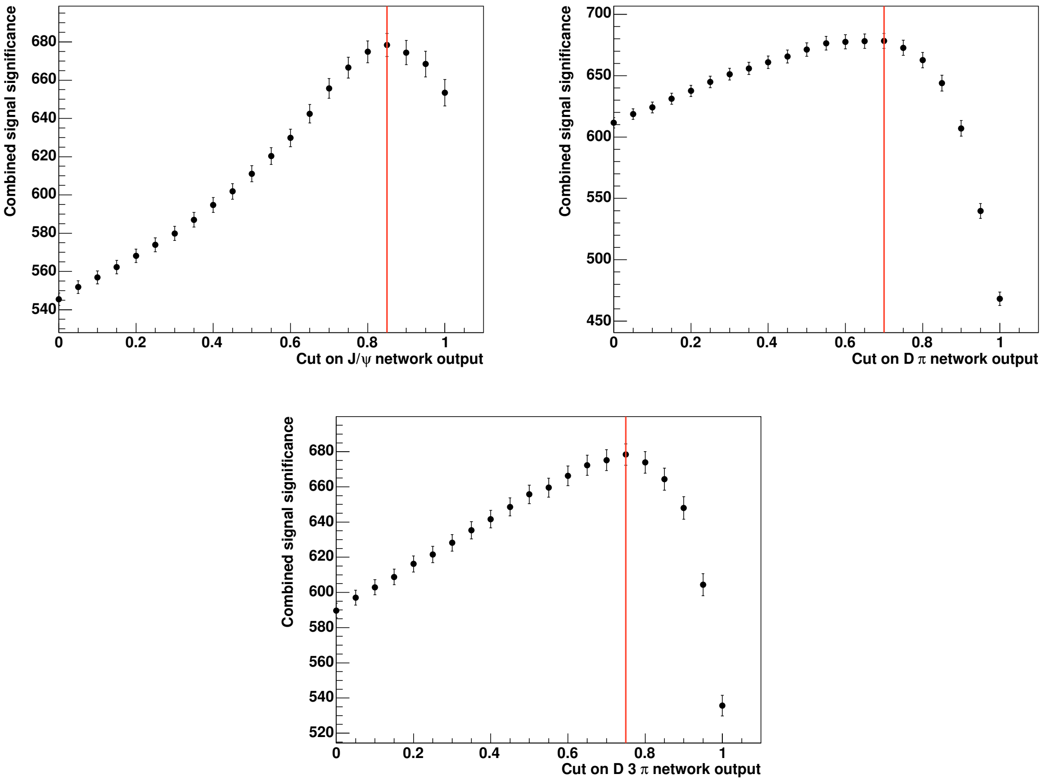


Figure 5.27: The combined significance with different cuts larger than zero on one network output while the cuts on the other network outputs are kept at their best values. The red line in each plot denotes where the selection cut is put for the combined significance.

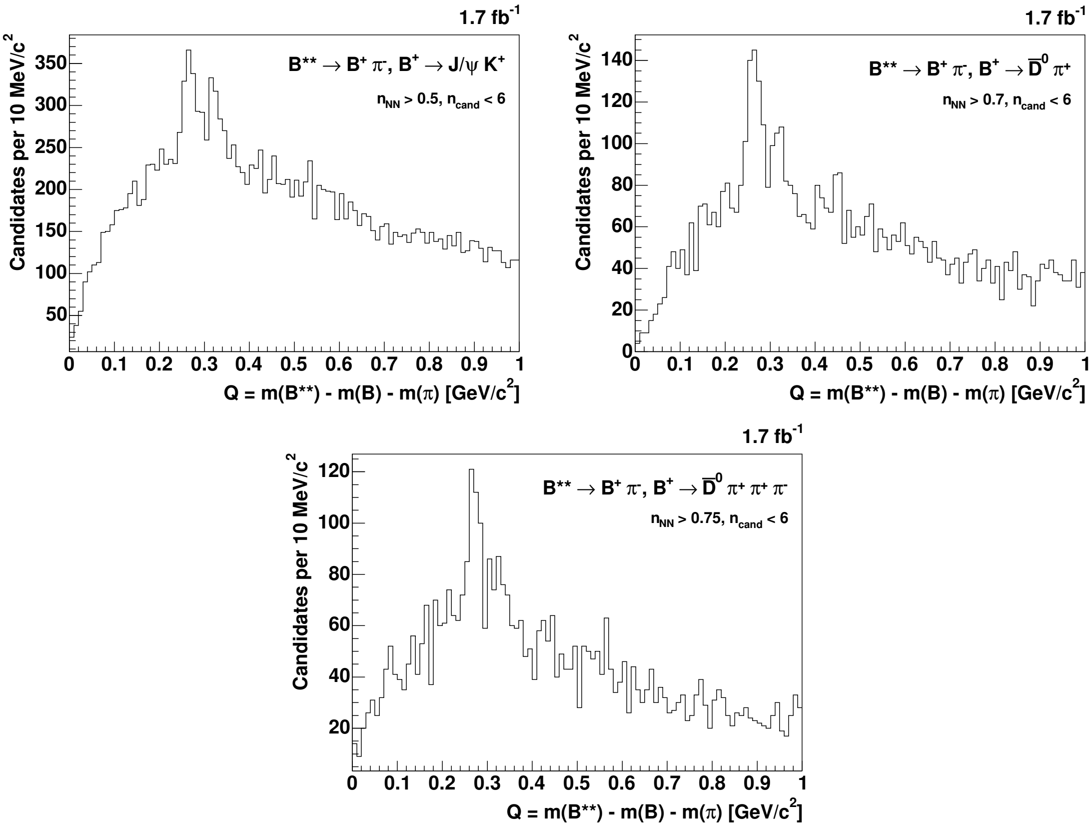


Figure 5.28: The Q value distribution of the B^{**} candidates of the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$, $B^+ \to \bar{D}^0\pi^+$ and $B^+ \to \bar{D}^0\,3\pi^{\pm}$. The distributions are obtained after applying the final selection cuts.

optimisation. It shows two clean signal peaks for the B_1 and B_2^* states.

5.3.4 Selection Stability

In order to check the stability of the B^{**} candidate selection, a set of tests is done verifying that the signal is not an artefact of the specific selection. The plots produced by these checks can be found in figures 5.29, 5.30 and 5.31 for the three different decay channels. As reference, the Q value distribution of the final selection is plotted in figure 5.28.

For the first check, the Q value distributions of the three decay channels are plotted with different cuts on the neural network output which is shown in the top row of each figure. The used network cuts are ±0.1 around the cut being applied for the final candidate selection. The requirement on the number of candidates per event remains unchanged and is less than 6.

Furthermore, the Q value distributions are plotted with different cuts on the number of candidates per event as seen in the middle row of the figures. The chosen cuts on this quantity are ± 1 around the cut which is applied in the final candidate selection. The cut on the network output remains at the best value used in the final selection.

For the last crosscheck, the Q value distributions are plotted with different binnings of 5 MeV/c^2 and 15 MeV/c^2 to ensure that the seen peaks are no artefacts of the binning. These plots are shown on the bottom row. In these binning-checks the cut on the neural network output and the number of candidates per event in every decay channel is the same as for the final selection.

To sum up, the Q value distribution of the B^{**} signal candidates is indeed stable regarding different cuts on the network output, the number of candidates per event and different binnings. All plots show clean signal peaks which can be interpreted as the two narrow states of the B^{**}.

5.3.5 Wrong Sign Candidates

Another cross check can be done by looking at the Q value distributions of the wrong sign candidates. Wrong sign candidates are B^{**} candidates which are combinations of $B^+\pi^+$ instead of $B^+\pi^-$. Consequently, these B^{**} mesons have a total electric charge being not equal zero. There is no signal expected in data samples containing only wrong sign candidates.

Figures 5.32, 5.33 and 5.34 show the Q value distribution of the wrong sign candidates with different cuts on the network output. In these plots the cut on the number of right sign candidates per event is kept at the same value as for the final signal selection. That is, the number of right sign candidates has to be less than 6.

In addition a different selection of wrong sign candidates is done to be closer to the logic of the final signal selection of right sign candidates. Thereby, the number of wrong sign candidates per event has to be less than 6 instead of the number of right sign candidates per event. The resulting Q value distributions can be seen in figures 5.35, 5.36 and 5.37.

In summary, none of the distributions showing the wrong sign candidates gives evidence for a fake signal which would be a sign for a biased neural network training and consequently a biased signal selection.

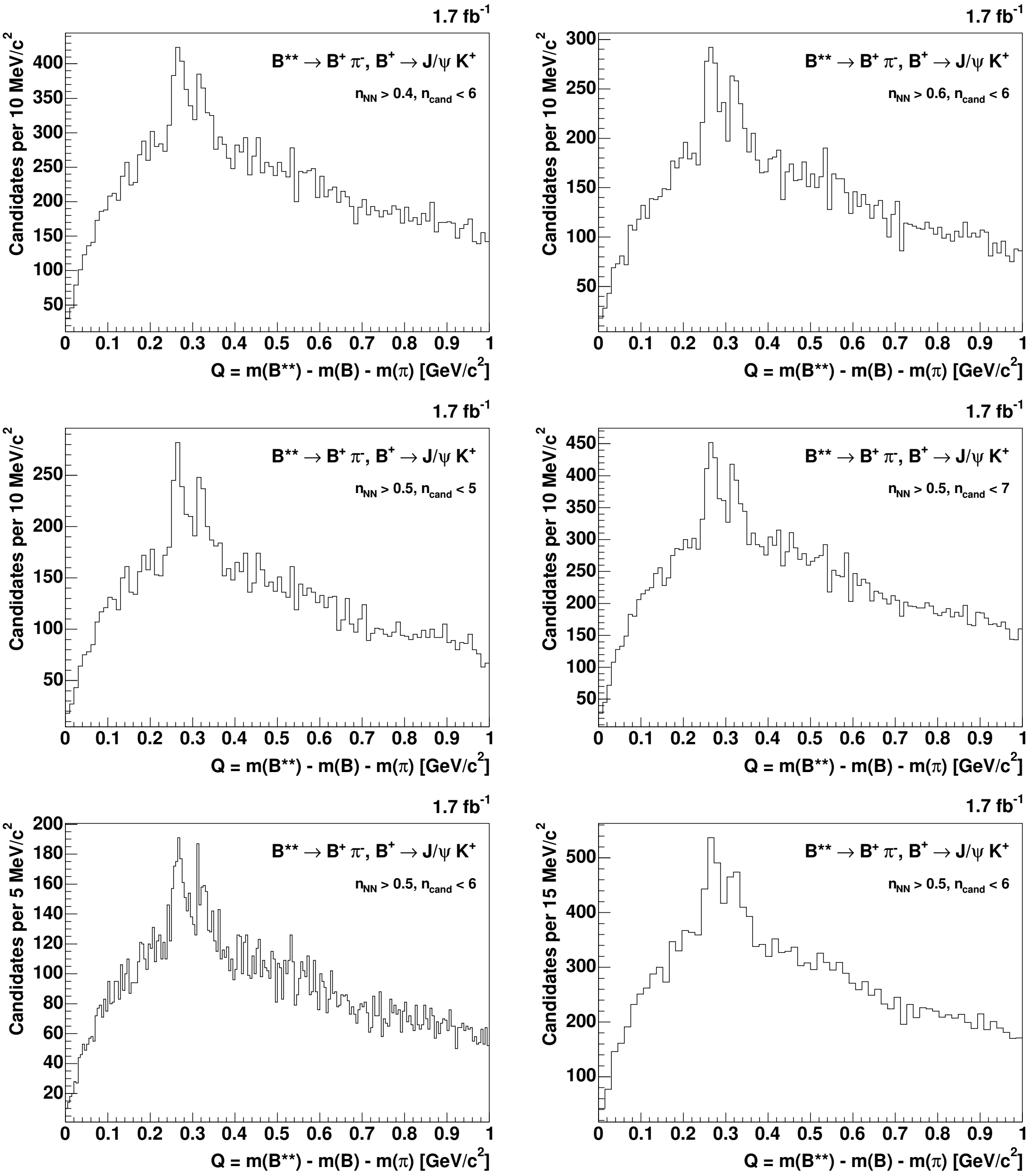


Figure 5.29: The Q value distribution of the B^{**} candidates in the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$ selected by cuts being slightly different from the final selection cuts and with different binnings.

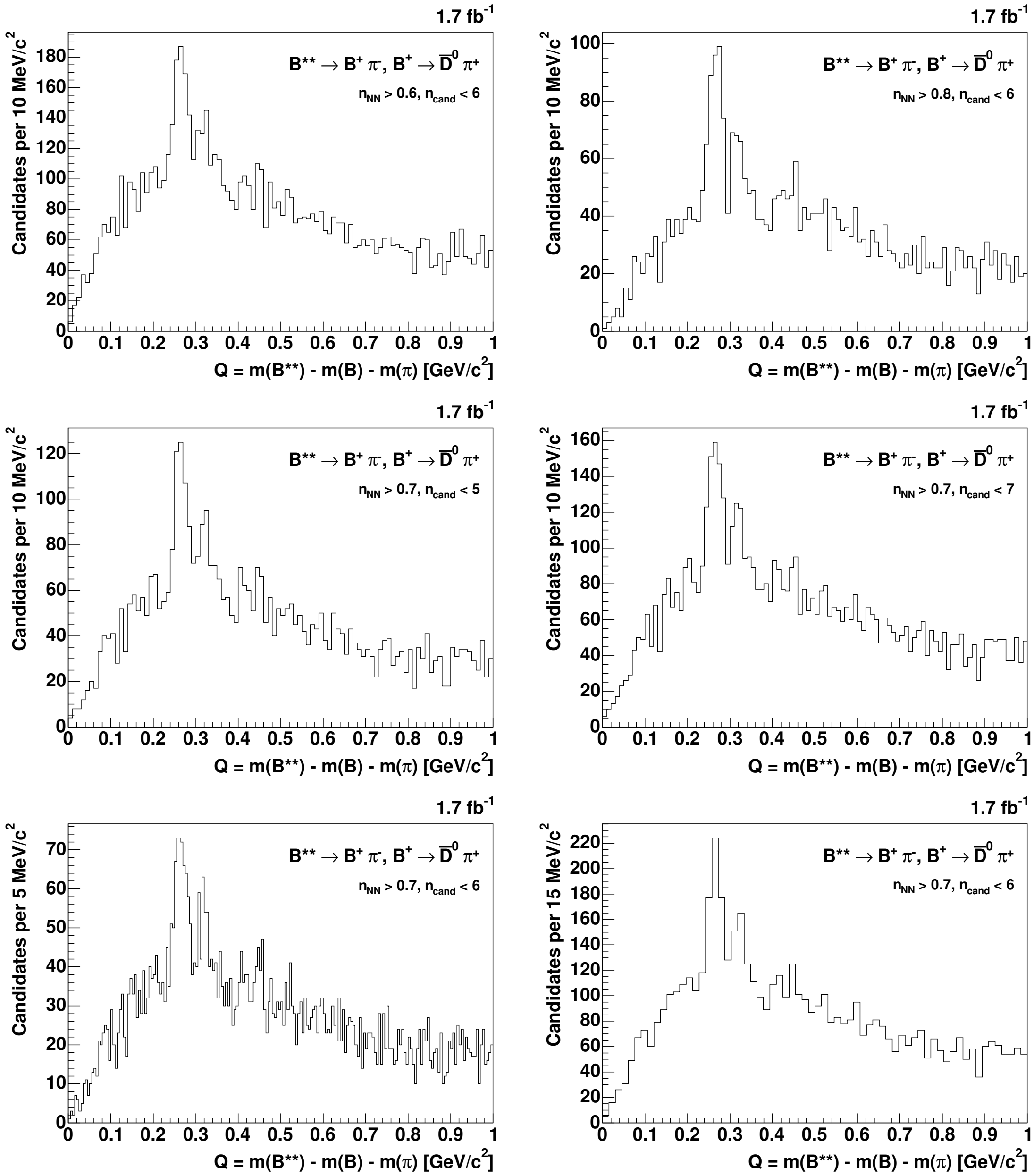


Figure 5.30: The Q value distribution of the B^{**} candidates in the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\pi^+$ selected by cuts being slightly different from the final selection cuts and with different binnings.

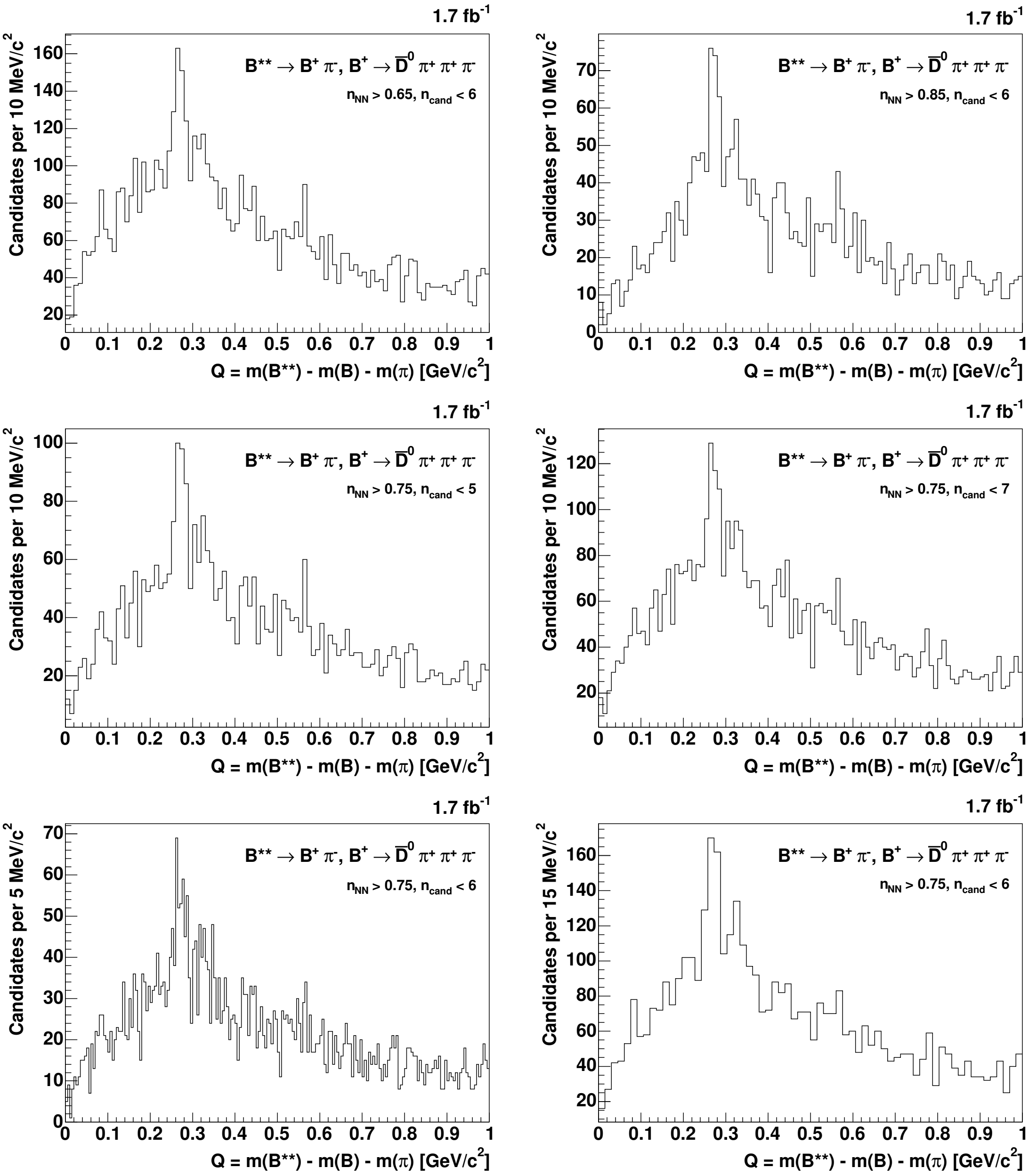


Figure 5.31: The Q value distribution of the B^{**} candidates in the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to \bar{D}^0\, 3\pi^\pm$ selected by cuts being slightly different from the final selection cuts and with different binnings.

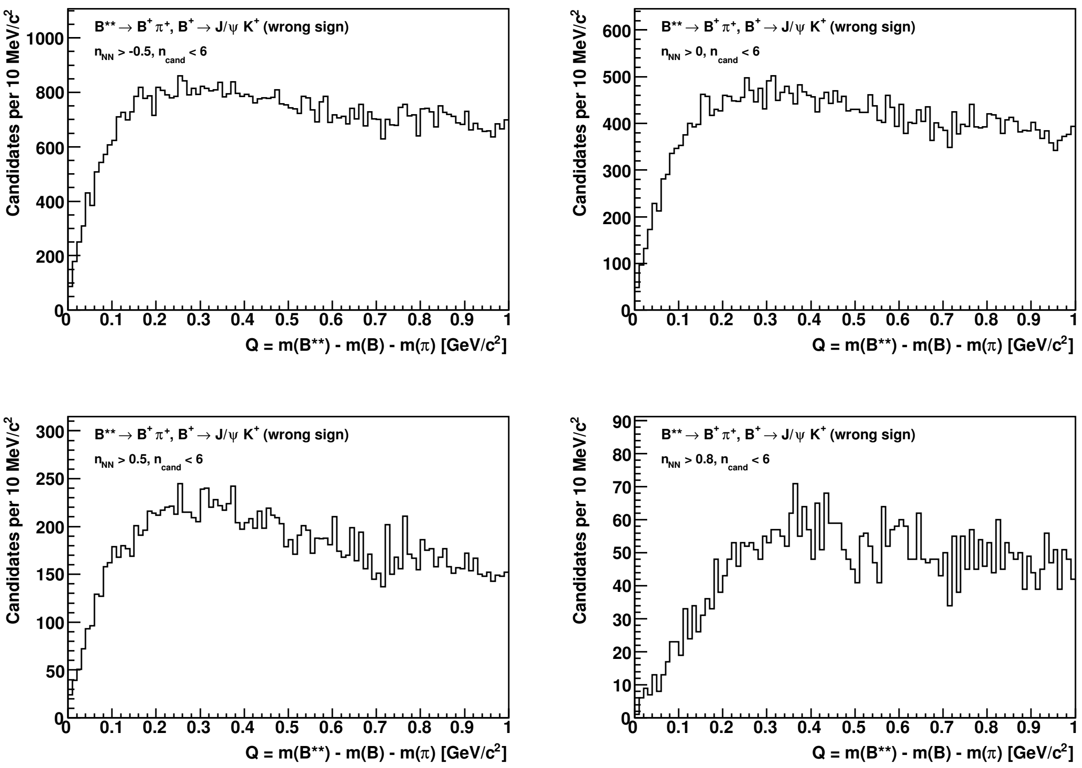


Figure 5.32: The Q value distribution of the wrong sign B^{**} candidates of the decay $B^{**} \rightarrow B^+\pi^+$ with $B^+ \rightarrow J/\psi K^+$ selected by different cuts on the network output. The selections also require the number of right sign candidates to be less than 6.

Variable	Unit	Preselection requirement
$n_{\mathrm{NN}}(B_s)$		> 0
$\lvert d_0(B_s)\rvert$	[cm]	> 0.008
$L_{xy}/\sigma_{L_{xy}}(B_s)$		> 4
$m(B_s)$	[GeV/c^2]	$5.2 - 5.5$
$L_{xy}/\sigma_{L_{xy}}(D_s)$		> 2
Q	[GeV/c]	< 0.25
—		Each π from the B_s^{**} must have hits in the silicon detectors

Table 5.10: Preselection requirements for $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$.

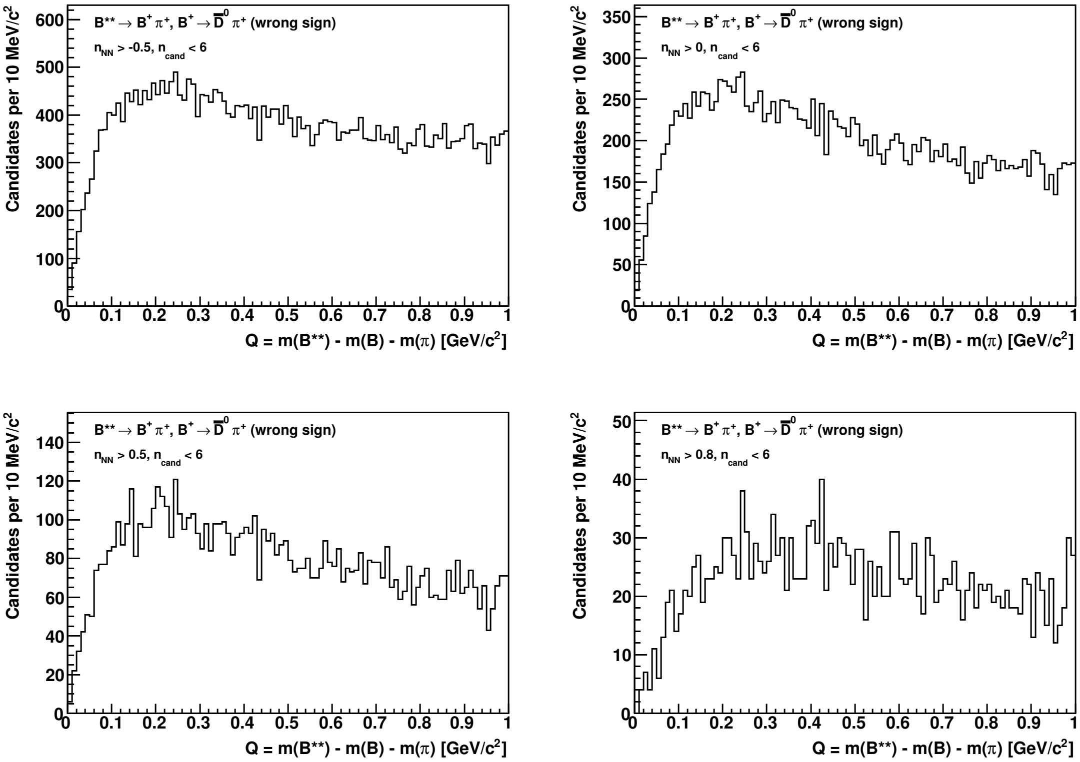


Figure 5.33: The Q value distribution of the wrong sign B^{**} candidates of the decay $B^{**} \to B^+\pi^+$ with $B^+ \to \bar{D}^0\pi^+$ selected by different cuts on the network output. The selections also require the number of right sign candidates to be less than 6.

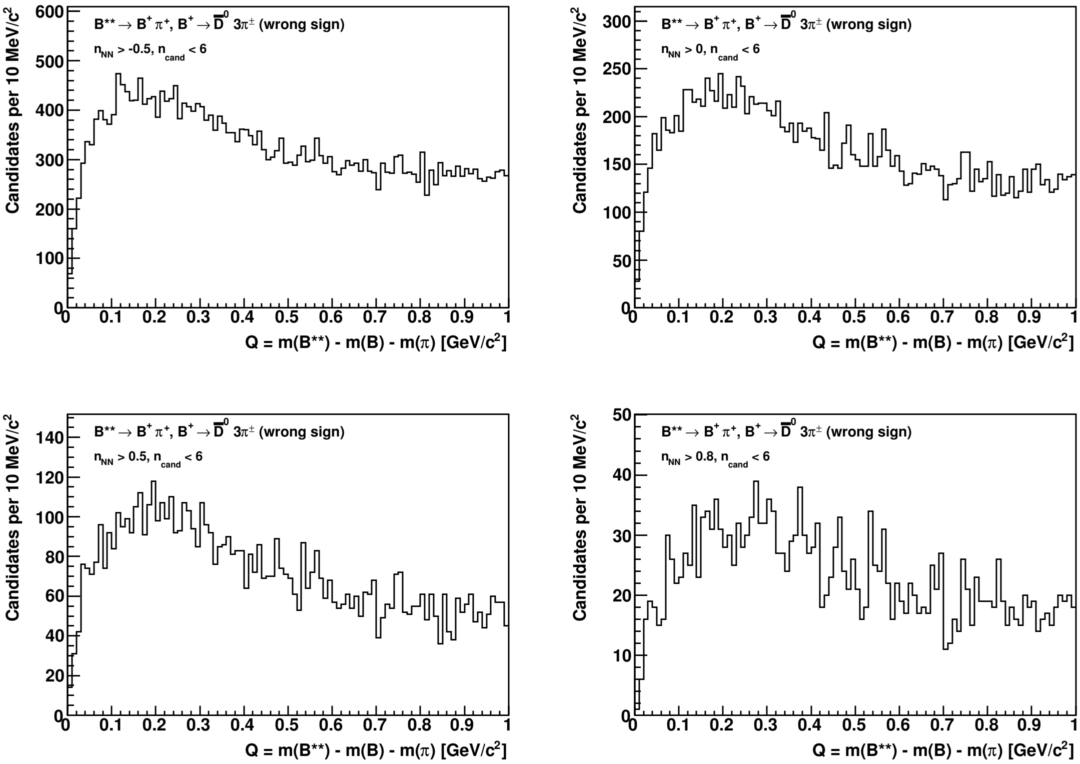


Figure 5.34: The Q value distribution of the wrong sign B^{**} candidates of the decay $B^{**} \to B^+\pi^+$ with $B^+ \to \bar{D}^0\, 3\pi^\pm$ selected by different cuts on the network output. The selections also require the number of right sign candidates to be less than 6.

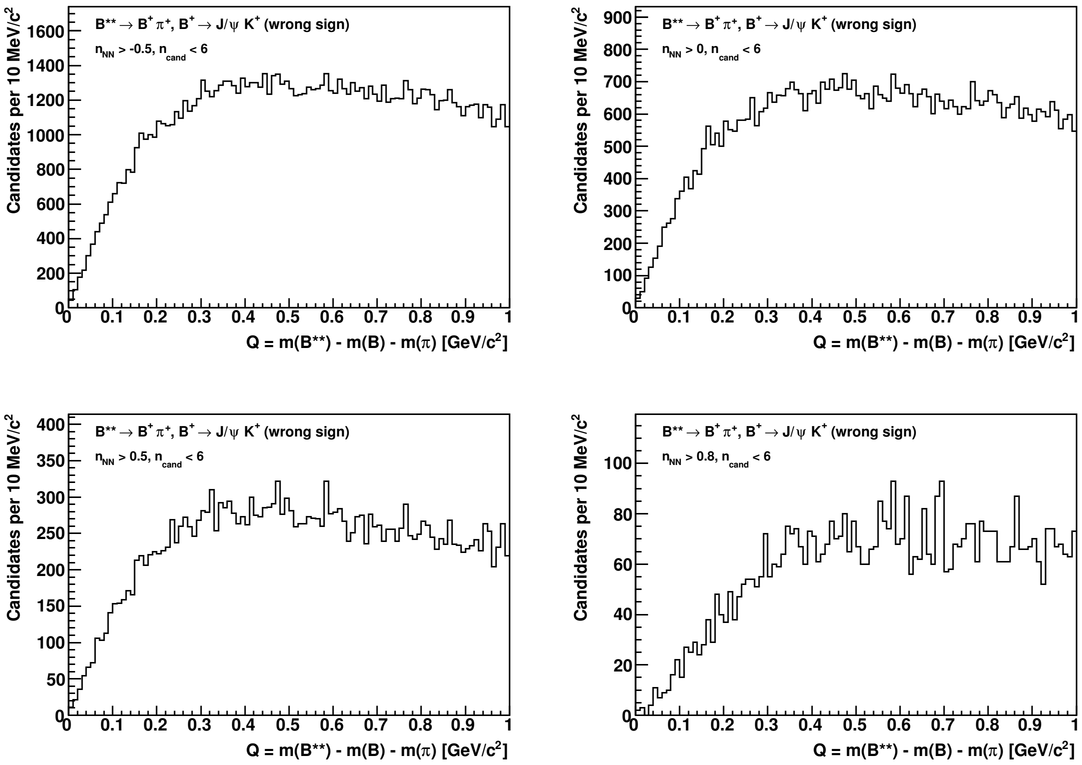


Figure 5.35: The Q value distribution of the wrong sign B^{**} candidates of the decay $B^{**} \rightarrow B^+\pi^+$ with $B^+ \rightarrow J/\psi K^+$ selected by different cuts on the network output. The selections also require the number of wrong sign candidates to be less than 6.

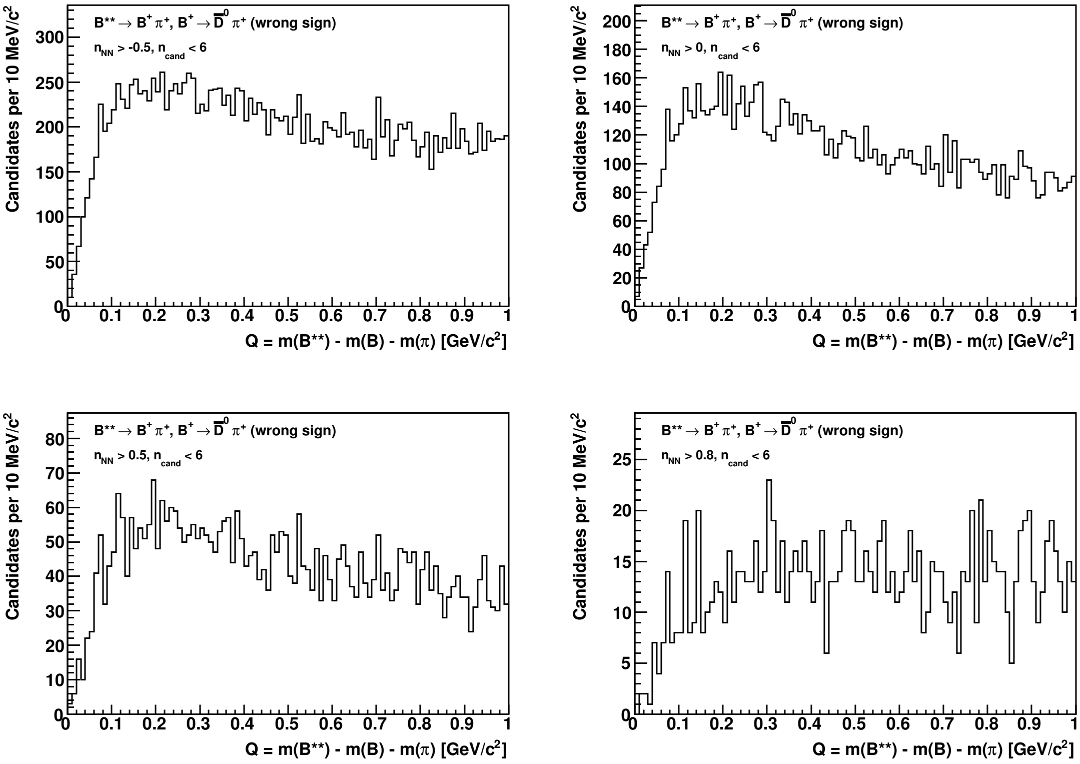


Figure 5.36: The Q value distribution of the wrong sign B^{**} candidates of the decay $B^{**} \to B^+\pi^+$ with $B^+ \to \bar{D}^0\pi^+$ selected by different cuts on the network output. The selections also require the number of wrong sign candidates to be less than 6.

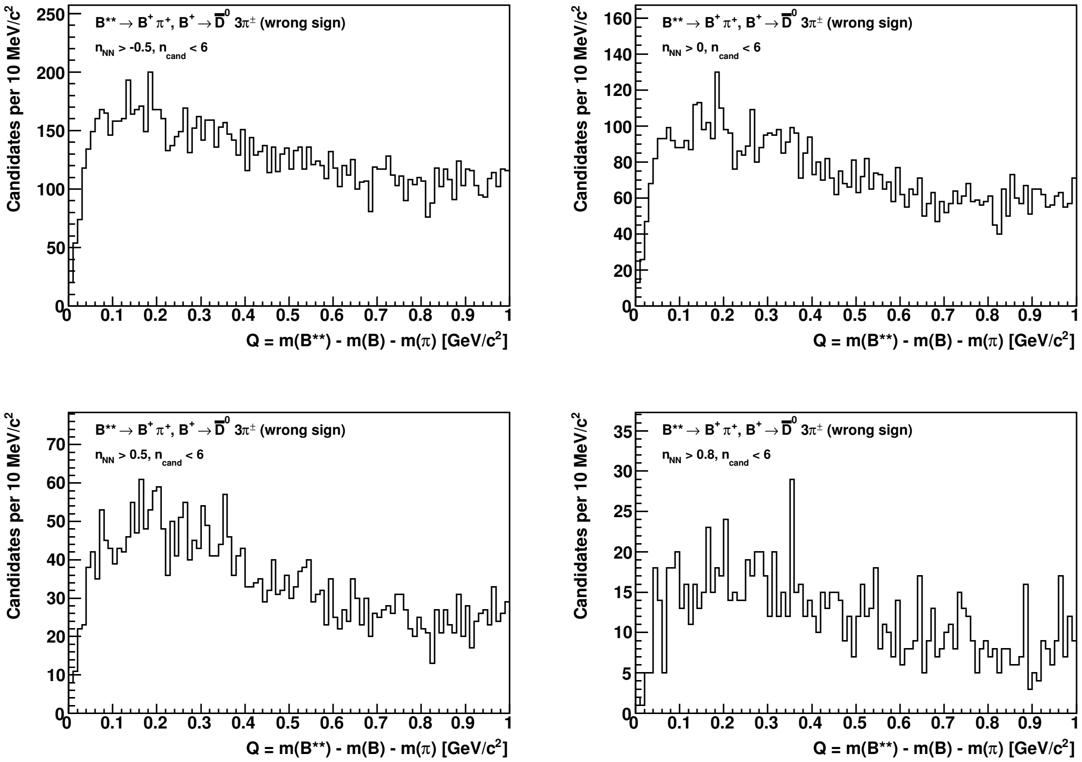


Figure 5.37: The Q value distribution of the wrong sign B^{**} candidates of the decay $B^{**} \to B^+\pi^+$ with $B^+ \to \bar{D}^0\, 3\pi^\pm$ selected by different cuts on the network output. The selections also require the number of wrong sign candidates to be less than 6.

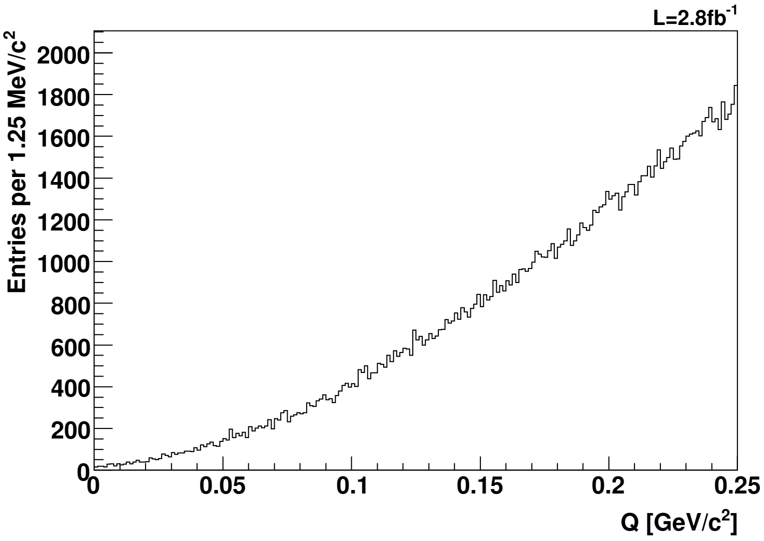


Figure 5.38: The Q value distribution of the combined B_s^{**} data sample including all six B_s decay channels. This distribution is used for the neural network training as background pattern.

5.4 B_s^{**} Candidate Selection

5.4.1 Selection of the $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$

The selection of the B_s^{**} candidates is done in a very similar way to the B^{**} selection. The main difference is that for the B_s^{**} selection only one single neural network is used for all six B_s decay channels. This is because of the low statistics in every decay mode which does not allow to train a neural network for each decay channel separately. The B_s^{**} selection starts by applying soft preselection requirements which remove a large number of background events but only a marginal part of the signal events. These requirements are listed in table 5.10.

For the neural network training Monte Carlo events are used as signal pattern and data in the Q value range from 0 to 0.25 GeV/c^2 is used as background pattern. Despite having applied the preselection requirements, the data sample is still dominated by background events as figure 5.38 shows. Therefore, the data sample can be well used for the neural network training as background pattern.

To avoid biasing the neural network training the Monte Carlo events are reweighted to the same Q value distribution as in data. Furthermore, the Monte Carlo events are reweighted in order to correct for the mixture of the different B_s decay channels in the combined Monte Carlo sample. The details of the Monte Carlo reweight-

Index	Significance	Rank	Used	Variable
1	4.1	20	✓	$\lvert d_0(B_s^{**})\rvert$
2	3.5	22	✓	$\chi^2_{r\phi}(B_s^{**})$
3	12.1	10	✓	$L_{xy}/\sigma_{L_{xy}}(B_s^{**})$
4	10	13	✓	$L_{xy}(B_s)$
5	3.5	12	✓	$L_{xy}/\sigma_{L_{xy}}(B_s)$
6	146.9	1	✓	$m(B_s)$
7	15.1	9	✓	$p_t(D_s)$
8	2.9	26	—	$\eta(D_s)$
9	2.5	24	—	$\chi^2_{r\phi}(D_s)$
10	2.9	25	—	$P(D_s)$
11	4.4	19	✓	$L_{xy}(D_s)$
12	5.1	16	✓	$\lvert d_0(D_s)\rvert$
13	18.6	8	✓	$L_{xy}/\sigma_{L_{xy}}(D_s)$
14	3.8	18	✓	$\ell_{xy}(D_s)$
15	9.9	11	✓	$\theta^*(D_s)$
16	2.4	28	—	$d_0^{lts}/\sigma_{d_0}(\pi_{B_s})$
17	1.9	29	—	$\lvert d_0(\pi_{B_s})\rvert$
18	1.5	31	—	$\left\lvert d_0(\pi_{B_s}^{(1)})\right\rvert$
19	1.8	30	—	$\left\lvert d_0(\pi_{B_s}^{(2)})\right\rvert$
20	2.3	27	—	$\left\lvert d_0(\pi_{B_s}^{(3)})\right\rvert$
21	5.7	15	✓	$\min_{B_s^{**}}(p_t)$
22	3.2	23	✓	$\max_{B_s^{**}}(p_t)$
23	4.8	17	✓	$\min_{B_s^{**}}(\lvert d_0\rvert)$
24	37.3	4	✓	$\max_{B_s^{**}}(\lvert d_0\rvert)$
25	25.3	7	✓	n_{Dcy}
26	24.9	6	✓	$n_{\mathrm{NN}}(B_s \to D_s^- \pi^+, D_s^- \to \bar{K}^* K^-)$
27	4	21	✓	$n_{\mathrm{NN}}(B_s \to D_s^- \pi^+, D_s^- \to \phi\pi^-)$
28	9.6	14	✓	$n_{\mathrm{NN}}(B_s \to D_s^- \pi^+, D_s^- \to 3\pi^\pm)$
29	44.9	3	✓	$n_{\mathrm{NN}}(B_s \to D_s^- 3\pi^\pm, D_s^- \to \bar{K}^* K^-)$
30	26.5	5	✓	$n_{\mathrm{NN}}(B_s \to D_s^- 3\pi^\pm, D_s^- \to \phi\pi^-)$
31	40.5	2	✓	$n_{\mathrm{NN}}(B_s \to D_s^- 3\pi^\pm, D_s^- \to 3\pi^\pm)$
32	0.6	32	—	Q

Table 5.11: List of input variables of the neural network for $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$.

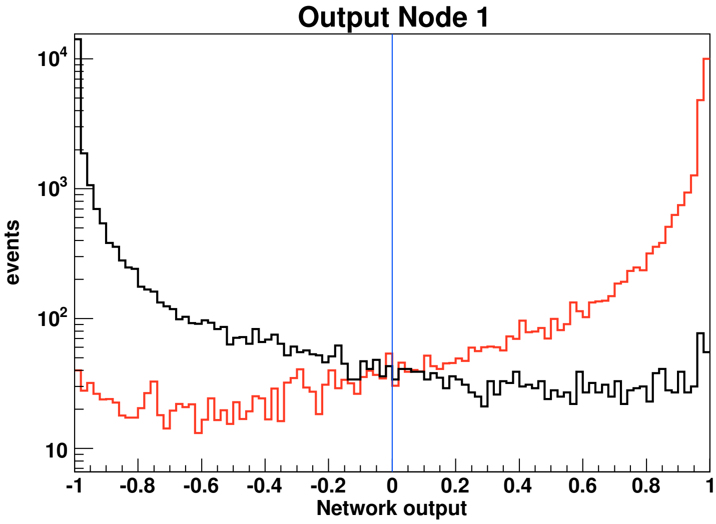


Figure 5.39: The distribution of the neural network output for signal (red) and background (black) events for the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ using the training sample.

ing are described in section 5.1.4. The inputs of the neural network are listed in table 5.11.

The preprocessing algorithms of NeuroBayes® keep only the most significant variables. If the significance of a variable is below 3σ the variable is pruned and not used in the neural network. The Q value is put into the network as a crosscheck to ensure the correctness of the Monte Carlo reweighting. Since the neural network cannot learn anything from the Q value distribution it is pruned which is an indication of a successful Monte Carlo reweighting.

The quality plots obtained by the neural network training can be evaluated in order to prove the success of the training. Figure 5.39 shows the separation power of the neural network for signal and background. In figure 5.40 the purity of the neural network is plotted as a function of the network output. Figure 5.41 presents the correlation matrix of the input variables.

5.4.2 Cut Optimisation

The final selection of the B_s^{**} is done by cutting on the output of the neural network trained to select the B_s^{**} candidates. The best cut on the neural network output is supposed to maximise the significance. The significance S_{NN} for the B_s^{**} selection is of the form

$$S_{\mathrm{NN}} = \frac{N_{\mathrm{MC}}^{(\mathrm{cand})}\left(n_{\mathrm{out}} > n_{\mathrm{cut}}\right)}{1.5 + \sqrt{N_{\mathrm{data}}^{(\mathrm{cand})}\left(n_{\mathrm{out}} > n_{\mathrm{cut}}\right)}} \tag{5.5}$$

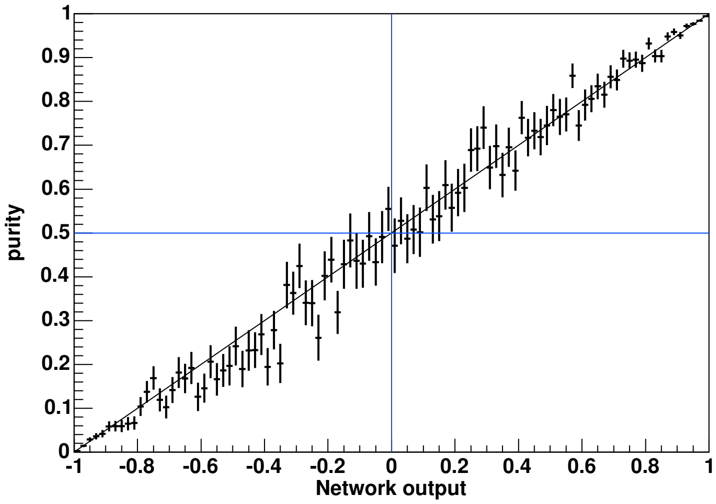


Figure 5.40: The purity as a function of the neural network output for the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ using the training sample.

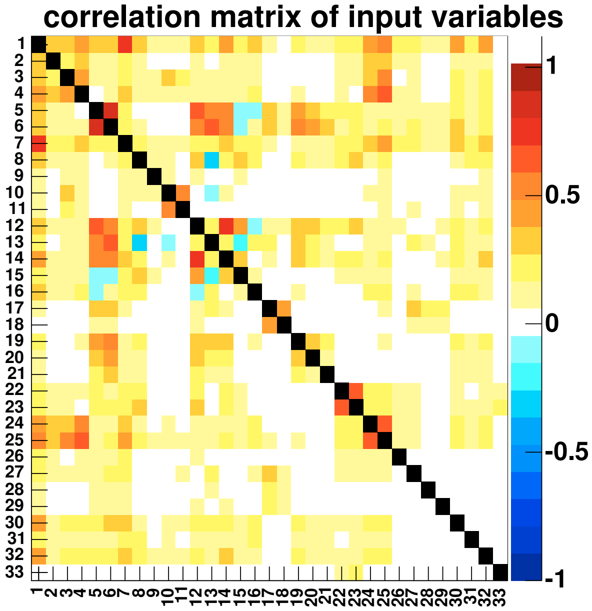


Figure 5.41: The correlation matrix of the neural network for the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$.

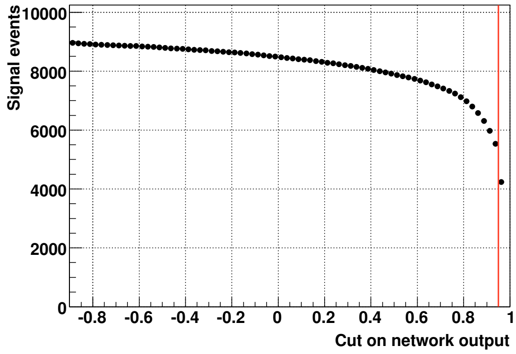


Figure 5.42: The number of signal events in the B_s^{**} sample over the cut on the network output. The red line denotes where the selection cut is put.

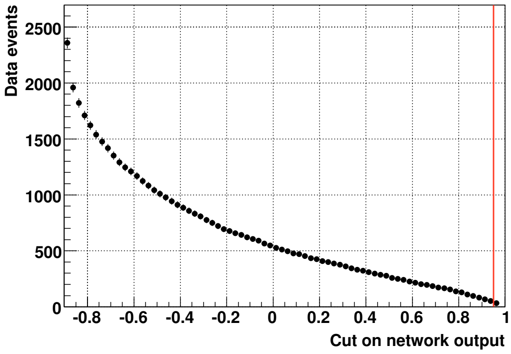


Figure 5.43: The number of data events in the B_s^{**} sample over the cut on the network output. The red line denotes where the selection cut is put.

where $N_{\mathrm{MC}}^{(\mathrm{cand})}$ is the number of Monte Carlo events selected at a given cut n_{cut} on the neural network output n_{out} and $N_{\mathrm{data}}^{(\mathrm{cand})}$ is the number of data events at the same cut on the neural network output. The number of Monte Carlo events is regarded to be proportional to the number of signal events and the number of data events is the sum of background events and signal events independently of the actual ratio. Since merely a low signal is expected, a constant summand is added to the denominator. This summand of 1.5 arises from the fact that the formula is optimised towards a 3σ observation [73]. For a search optimisation the significance formula written in equation (5.3) would tend to cut on the neural network output as hard as possible without really maximising the significance. Therefore, the significance definition has to be adapted when searching for new phenomena.

In order to calculate the significance, Monte Carlo and data events are counted at a given cut on the network output in a Q value range where the signal peaks are expected. The search window in Q was therefore chosen from 0.134 GeV/c^2 to 0.2 GeV/c^2. For increasingly harder cuts on the network output, the number of data events decreases faster than the number of signal events, as shown in figures 5.42 and 5.43. Figures 5.44 and 5.45 show the significance over the cut on the network output. Based on these plots the selection cut on the neural network output is determined to be 0.95. The Q value distribution of the B_s^{**} candidates selected by this cut is given in figure 5.46. Since no signal peak is visible a Bayesian limit is set on the branching ratios of the narrow B_s^{**} states.

In contrast to the selection of the B^{**} candidates there is no cut applied on the number of candidates per event for the B_s^{**} selection. Setting a limit on the branching ratios requires the knowledge of the signal efficiency. But since the number

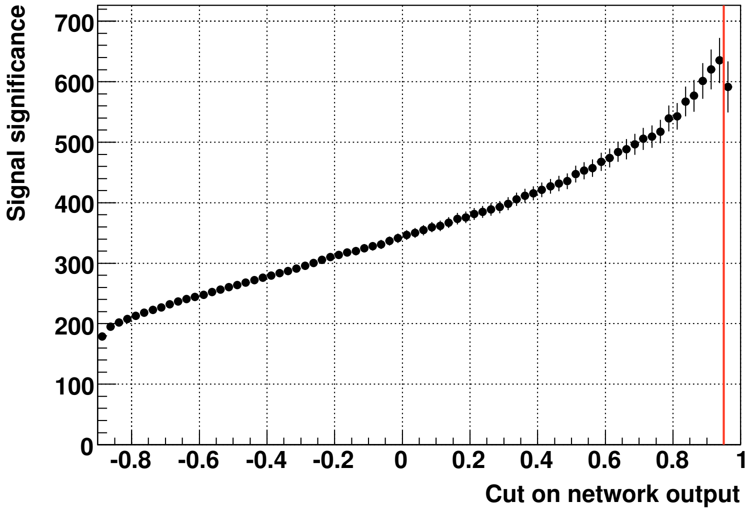


Figure 5.44: The significance S_{NN} of the decay $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ over the cut on the neural network output . The red line denotes where the selection cut is put.

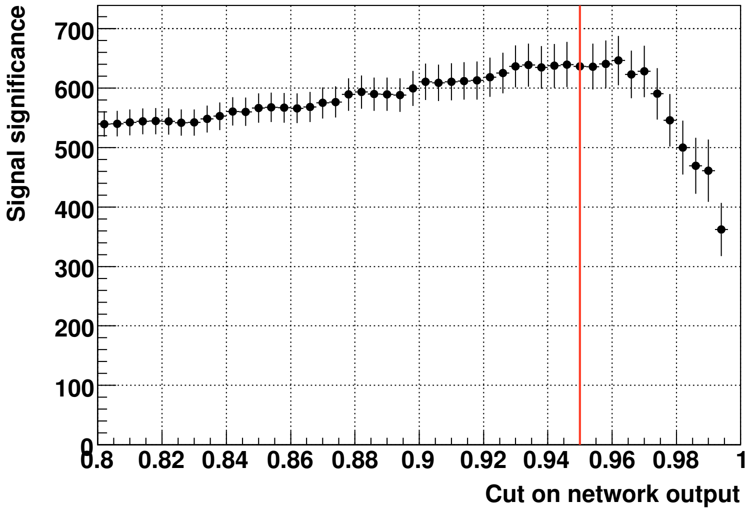


Figure 5.45: A more detailed plot of the significance S_{NN} of the decay $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ over the cut on the neural network output for cuts larger than 0.8. The red line denotes where the selection cut is put

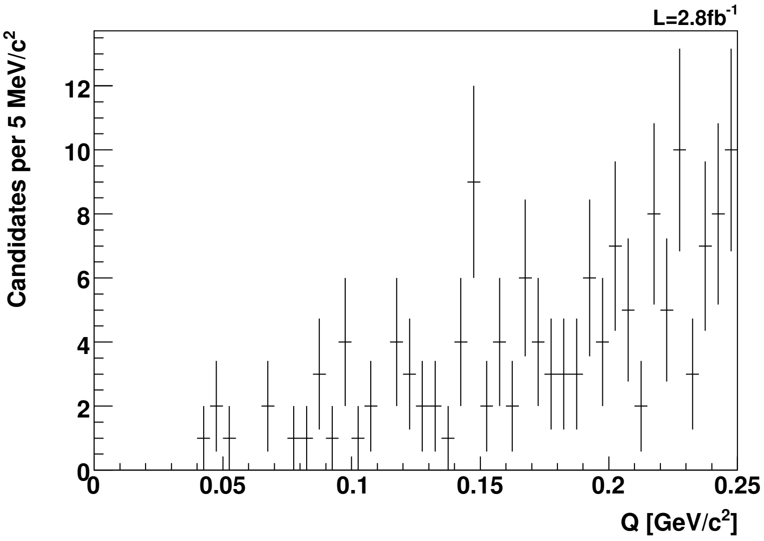


Figure 5.46: The Q value distribution of the B_s^{**} candidates selected by the final selection cut on the output of the neural network of 0.95. The expected Q values for the B_{s1} is 184 MeV/c^2 and 150 MeV/c^2 for the B_{s2}^*.

of candidates cannot be modelled by the Monte Carlo samples it is not possible to determine the efficiency of the cut on the number of candidates.

6 B^{**} Fit Description

6.1 Q Value Resolution of the B^{**} Signal

An unbinned maximum likelihood fit is performed to measure the properties of the narrow B^{**} states. The signal peaks in the likelihood function are described by non-relativistic Breit-Wigner distributions convoluted with the detector resolution model. The Q value resolution of the signal is determined using the same Monte Carlo samples as the ones used for the neural network training. Since the Monte Carlo samples are generated with a flat mass distribution the Q dependence of the resolution can also be studied.

To study the signal resolution the residual Q value is plotted for different ranges of the Q value. The residual Q value is the difference between the generated and reconstructed Q value. The residual distribution is fitted with a double Gaussian for each $B^{**} \to B\pi$ and $B^{**} \to B^{*}\pi$ decay. The double Gaussian has the form

$$\mathcal{G}(Q) = \frac{N \cdot f_1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(Q-Q_0)^2}{2\sigma_1^2}} + \frac{N \cdot (1-f_1)}{\sqrt{2\pi}\sigma_2} e^{-\frac{(Q-Q_0)^2}{2\sigma_2^2}} \tag{6.1}$$

where Q_0 is the mean of the peak and σ_1, σ_2 are the width of the two Gaussians. The factor f_1 is the fraction of the Gaussian with width σ_1. For a better behaviour of the fit the fraction is fixed to $f_1 = 0.2$ as it describes the data well. In figure 6.1 an example of one of these fits is shown.

Figures 6.2 and 6.3 show the Q value resolution as a function of the measured Q value. Since the Q value resolution among the decay modes is consistent, the B^{**} data samples can be put together for the final maximum likelihood fit. To determine the Q value resolution of the combined data, the Monte Carlo samples are merged and fitted again. The resolution as a function of the measured Q value of the combined sample can be seen in figure 6.4. To extract the Q value resolution at distinct Q values the Q dependence is parametrised with a linear function of the form

$$\sigma(Q) = a + b \cdot Q. \tag{6.2}$$

The parametrisation is done in the range from $Q = 125\ \mathrm{MeV/c^2}$ to $675\ \mathrm{MeV/c^2}$ and the obtained parameters are listed in table 6.1.

In a preliminary fit to the Q value distribution of the B^{**} states the resolution

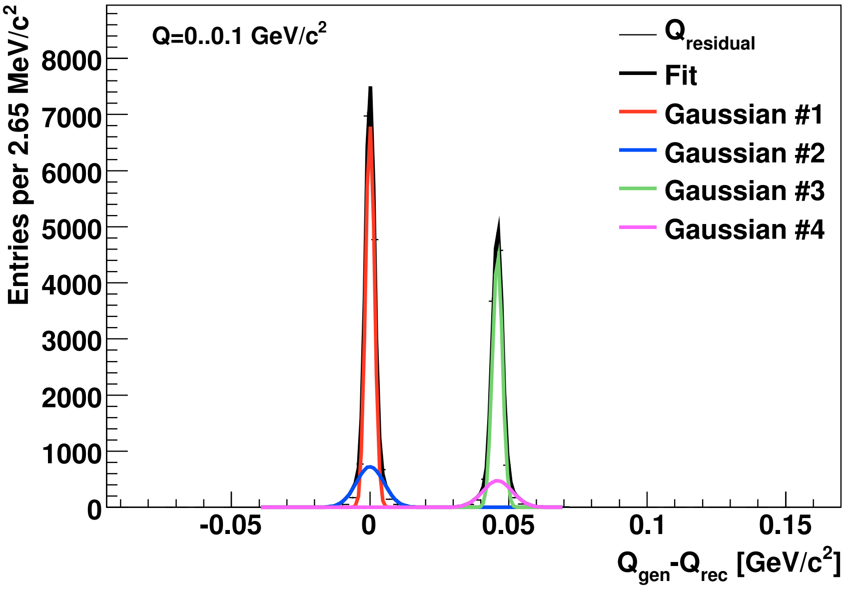


Figure 6.1: Residual Q value distribution in the Monte Carlo sample of the decay $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$. Having two peaks arises from the fact that the photon from the decay $B^* \to B\gamma$ is not detected by the CDF II detector.

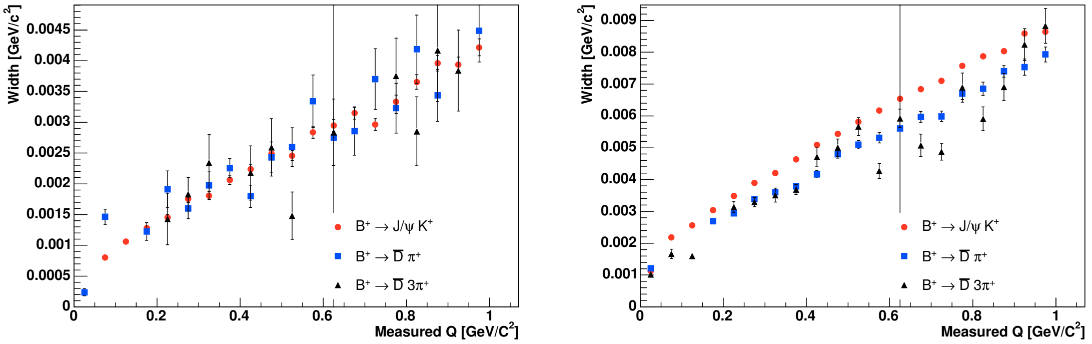


Figure 6.2: Q value resolution of the $B^{**} \to B\pi$ decays as a function of the measured Q value. The resolution is estimated using Monte Carlo samples and is parametrised by a double Gaussian. The width of the narrow Gaussian is shown on the left and the width of the broad one is shown on the right.

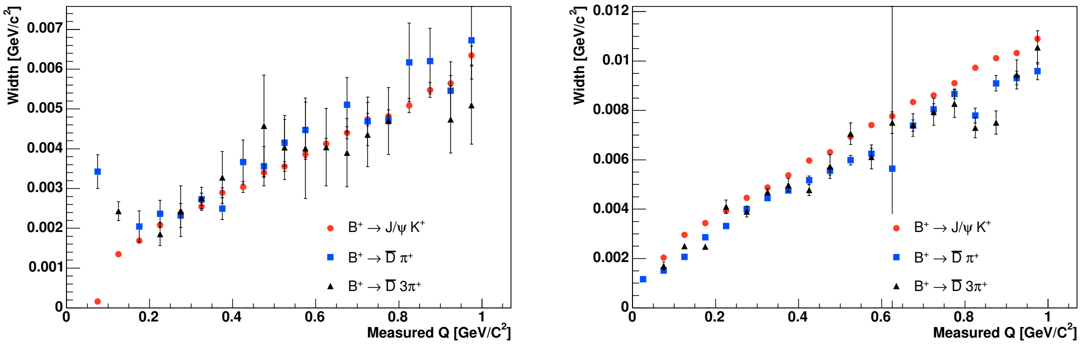


Figure 6.3: Q value resolution of the $B^{**} \to B^*\pi$ decays as a function of the measured Q value. The resolution is estimated using Monte Carlo samples and is parametrised by a double Gaussian. The width of the narrow Gaussian is shown on the left and the width of the broad one is shown on the right.

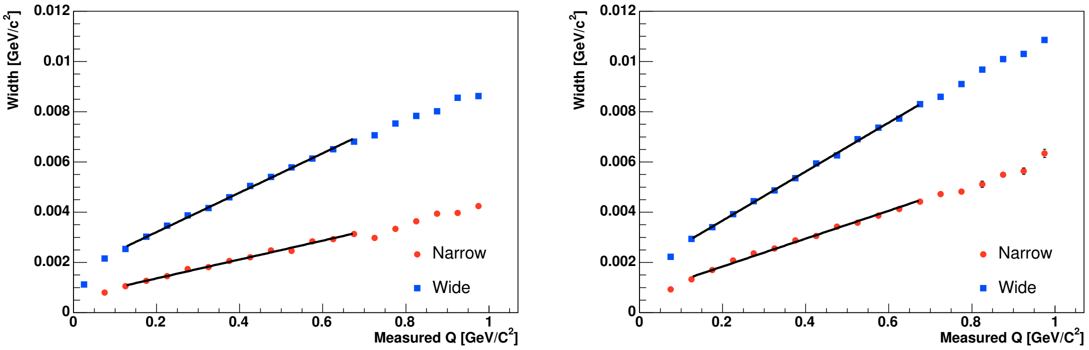


Figure 6.4: Q value resolution of $B^{**} \to B\pi$ (left) and $B^{**} \to B^*\pi$ (right) decays as a function of the measured Q value. The plots show the combination of all three B^+ decay modes.

	σ_1		σ_2	
Decay	a [10^{-5} GeV/c^2]	b [10^{-5}]	a [10^{-5} GeV/c^2]	b [10^{-5}]
$B^{**} \to B\pi$	61 ± 3	375 ± 8	164 ± 2	784 ± 5
$B^{**} \to B^*\pi$	73 ± 4	555 ± 11	172 ± 2	974 ± 6

Table 6.1: Parametrisation of the Q value resolution for the B^{**} signal as a function of the measured Q value.

Decay	Expected Q [MeV/c^2]	σ_1 [MeV/c^2]	σ_2 [MeV/c^2]
$B_1 \to B^*\pi$	263	2.18	4.27
$B_2^* \to B^*\pi$	275	2.26	4.40
$B_2^* \to B\pi$	321	1.82	4.16

Table 6.2: Q value resolution of the B^{**} states used in the final B^{**} fit. The expected Q values are determined by a preliminary fit to the Q value.

at the Q values where the signals are expected is put into the fit. From this fit the exact signal positions are obtained and the appropriate signal resolutions can be determined for the usage in the final fit, which are summarised in table 6.2.

6.2 Likelihood Function

In order to measure the masses of the narrow B^{**} states an unbinned maximum likelihood fit to the Q value is performed. However, the implementation of the fit was developed in cooperation with a research group of the CDF collaboration and is not part of this thesis [75]. For completeness the important parts of the fit are described here nonetheless.

The likelihood function of the fit comprises different components to describe the Q value distribution of the combined B^{**} data sample. The different components are:

- **Three signal peaks**

  Each of the three signal peaks $B_2^* \to B\pi$, $B_2^* \to B^*\pi$ and $B_1 \to B^*\pi$ is described by two Voigtians because of the detector resolution being modelled by a double Gaussian, as shown in section 6.1. The Voigtian function is a Breit-Wigner distribution convoluted with a Gaussian distribution. According to the mass and width predictions the peaks $B_2^* \to B^*\pi$ and $B_1 \to B^*\pi$ are expected to overlap. In the fit, the width of the $B_2^* \to B\pi$ peak is floating whereas the signal width of $B_2^* \to B^*\pi$ is fixed to the width of the $B_2^* \to B\pi$ peak. According to the prediction [14], the width of the B_1 is related to the B_2^* width:

  $$\frac{\Gamma(B_1)}{\Gamma(B_2^*)} = 0.9. \tag{6.3}$$

  This ratio is added to the fit with a Gaussian constraint assuming an uncertainty of 20% since there is no theoretical uncertainty quoted on this fraction.

  The positions of the $B_2^* \to B^*\pi$ and $B_2^* \to B\pi$ signal peaks are constraint to each other by the energy of the missing photon, $E(\gamma) = 45.78 \pm 0.35$ MeV/c^2 [2], from the decay $B^* \to B\gamma$. The uncertainty of the photon energy is put into the fit with a Gaussian constraint. The normalisation of the peak $B_2^* \to B^*\pi$ is kept relative to the $B_2^* \to B\pi$ peak and determined by reference [74]:

$$\frac{\mathrm{BR}\left(B_2^* \to B\pi\right)}{\mathrm{BR}\left(B_2^* \to B^*\pi\right)} = 1.1 \pm 0.3. \tag{6.4}$$

Again, the parameter is constricted with a Gaussian constraint.

- **B_s^{**} reflections**

  The background contribution caused by B_s^{**} reflections is treated separately in the fit function. B_s^{**} reflections are misreconstructed B_s^{**} mesons of the decay $B_s^{**} \to B^+K^-$ where the kaon is mistakenly assigned the pion mass. The shape of the reflections is derived from Monte Carlo samples simulating the decays $B_{s1} \to B^*K$ and $B_{s2}^* \to BK$. The shape of the B_s^{**} reflections is fixed in the final fit. The masses of the B_{s1} and B_{s2}^* states are taken from the $B_s^{**} \to B^+K^-$ analysis [1]. The B_s^{**} Monte Carlo samples are selected by the B^{**} neural network implemented in the context of this thesis. The shape of the B_{s1} is modelled by a single Gaussian function and the shape of the B_{s2}^* by a Crystall Ball function. The yield of the expected B_s^{**} reflections is estimated from Monte Carlo samples and free within a Gaussian constraint.

- **Hadronisation and combinatorial background**

  Further background components are taken into account in the likelihood function in addition to the B_s^{**} reflections. This type of background is modelled by a function of the form

$$f(Q) = \left(\frac{Q}{Q_{\mathrm{max}}}\right)^{\alpha} \cdot e^{-\frac{\alpha}{\gamma}\left(\left(\frac{Q}{Q_{\mathrm{max}}}\right)^{\gamma}-1\right)}. \tag{6.5}$$

  Here, Q_{max} is the value of Q for which $f(Q)$ has its maximum of $f(Q_{\mathrm{max}}) = 1$ for any α and γ. The parameters α and γ control the shape of the background from 0 to Q_{max} and above Q_{max}, respectively.

6.3 Separate Fits to B^{**} Decay Modes

Before the final fit to the combined data sample is performed, each of the three B^{**} data samples is fitted separately. In doing so, the signal consistency among the decay modes is checked. The signal significance of these data samples was optimised separately, as described in section 5.3.3. The individual fits to the different data samples use a simplified background shape where the background being caused by

Parameter	$B^+ \to J/\psi K^+$	$B^+ \to \bar{D}^0\pi^+$	$B^+ \to \bar{D}^0\,3\pi^\pm$
$Q(B_2^*)$ [MeV/c^2]	321 ± 3	320 ± 3	324 ± 3
$m(B_2^*) - m(B_1)$ [MeV/c^2]	13 ± 4	14 ± 3	17 ± 3
$\Gamma(B_2^*)$ [MeV/c^2]	35 ± 9	20 ± 6	21 ± 6
Number of $B_1 \to B^*\pi$ events	260 ± 102	208 ± 57	145 ± 49
Number of $B_2^* \to B\pi$ events	454 ± 78	127 ± 31	112 ± 24

Table 6.3: Signal parameters of the separate B^{**} fits to the three different decay modes.

B_s^{**} reflections is neglected since these fits are intended to yield merely a rough estimation of the signal Q values. The fit results are given in figure 6.5. These fits were done at a stage of the analysis where a data sample of merely 1.3 fb^{-1} was available for the $B^+ \to \bar{D}^0\pi^+$ and $B^+ \to \bar{D}^0\,3\pi^\pm$ decay channels. Table 6.3 summarises the signal parameters obtained from the fits.

The positions of the signal peaks are consistent for the different decay modes. The B_2^* is located around 322 MeV/c^2 in the Q value and the mass difference between B_1 and B_2^* is about 15 MeV/c^2. There is a discrepancy of the intrinsic width of the B_2^* which is still within the statistic uncertainty, though.

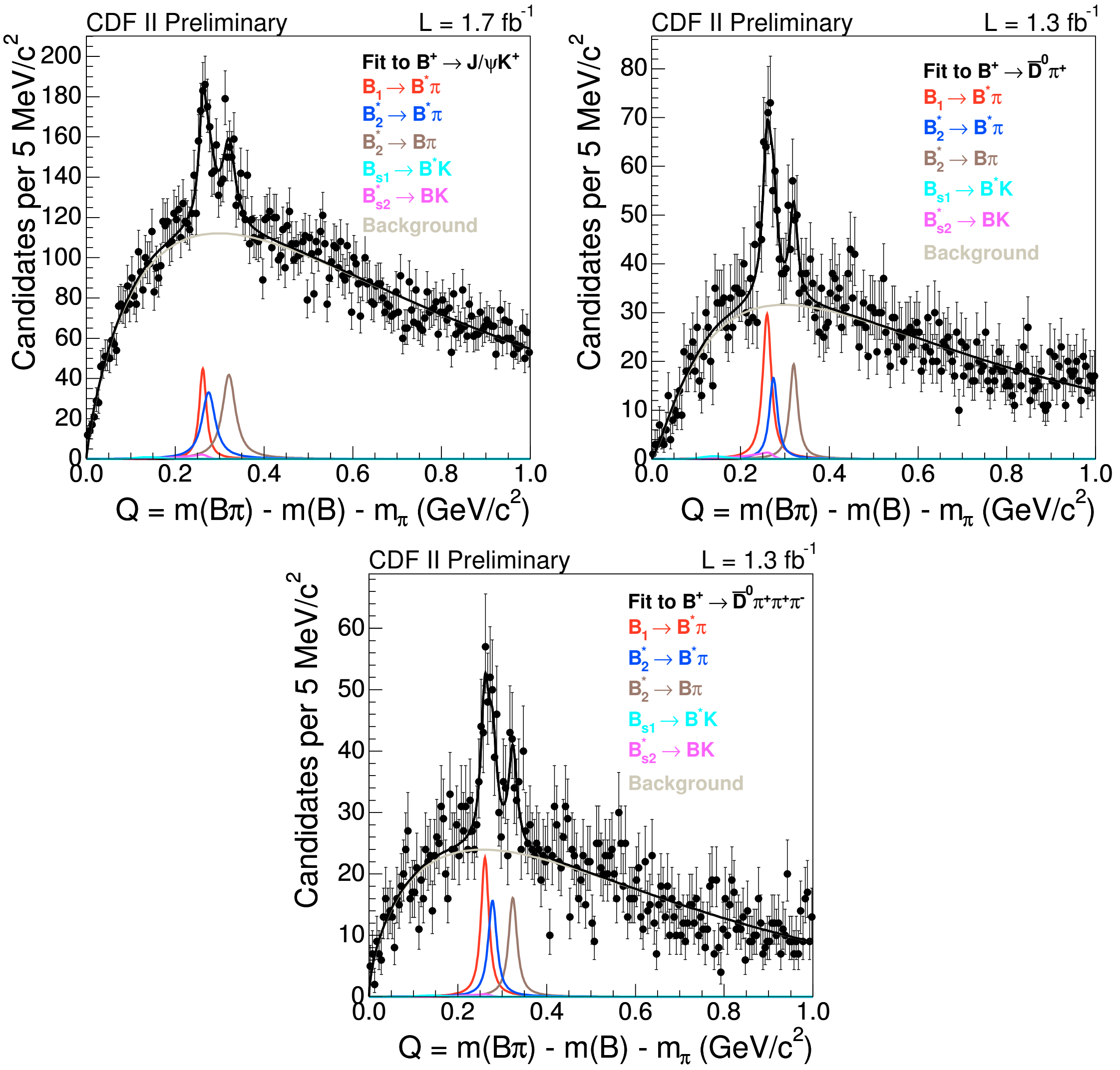


Figure 6.5: Results of the separate B^{**} fits to the three different decay modes. These fits are a cross-check to see if the signals are consistent in all three decay channels.

7 B_s^{**} Fit Description

7.1 Reconstruction Efficiencies

Due to the lack of a significant signal in the B_s^{**} data samples a Bayesian limit is set for the branching ratios of the narrow B_{s1} and B_{s2}^* states. The branching ratios are measured with respect to a reference decay since the absolute production rates of the B_s^{**} mesons are unknown. At CDF, the B_{s1} and B_{s2}^* states have already been measured in decays of $B_s^{**} \to B^+K^-$ with $B^+ \to \bar{D}^0\pi^+$ and $B^+ \to J/\psi K^+$ [1]. Since the data of the $B^+ \to J/\psi K^+$ decay was collected using a different trigger than for the $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ data samples the decay $B_s^{**} \to B^+K^-$ with $B^+ \to \bar{D}^0\pi^+$ was chosen as reference decay for measuring the branching ratios of the B_{s1} and B_{s2}^* states in decays to $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$. The Bayesian limit is calculated assuming a flat prior for the branching ratio.

7.1.1 Signal Efficiency

The signal efficiency is determined by studying the Monte Carlo samples that were used in the training of the neural network for selecting the B_s^{**} candidates. The efficiency ϵ is defined as the ratio of reconstructed ($N_{\mathrm{MC}}^{(\mathrm{reconstructed})}$) and generated ($N_{\mathrm{MC}}^{(\mathrm{generated})}$) Monte Carlo events:

$$\epsilon = \frac{N_{\mathrm{MC}}^{(\mathrm{reconstructed})}}{N_{\mathrm{MC}}^{(\mathrm{generated})}}. \tag{7.1}$$

The number of reconstructed Monte Carlo events is counted after applying the same selection procedure as for the data samples taking into account the proper branching ratios from the Monte Carlo generation. Table 7.1 shows the signal efficiencies of the different B_s decay channels. The signal efficiencies of the reference decay $B_s^{**} \to B^+K^-$ are determined in the same way as for the $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ decays. Table 7.2 lists the efficiencies of the reference channel.

7.1.2 Efficiency Uncertainty

For the decay $B_s^{**} \to B^+K^-$ the efficiency uncertainty is estimated by studying the ratio of Monte Carlo to signal events in data at different selection cuts around the

B_s^{**}	Decay		Q [MeV/c^2]	Selected	Efficiency [10^{-05}]
B_{s1}	$B_s \to D_s^-\pi^+$	$D_s^- \to \bar{K}^*K^-$	$176\ldots192$	52 ± 7.21	1.28 ± 0.178
		$D_s^- \to \phi\pi^-$		49 ± 7	1.21 ± 0.173
		$D_s^- \to 3\pi^\pm$		299 ± 17.3	7.38 ± 0.427
	$B_s \to D_s^- 3\pi^\pm$	$D_s^- \to \bar{K}^*K^-$	$176\ldots192$	351 ± 18.7	8.26 ± 0.441
		$D_s^- \to \phi\pi^-$		635 ± 25.2	14.9 ± 0.593
		$D_s^- \to 3\pi^\pm$		605 ± 24.6	14.2 ± 0.579
B_{s2}^*	$B_s \to D_s^-\pi^+$	$D_s^- \to \bar{K}^*K^-$	$138\ldots162$	60 ± 7.75	0.988 ± 0.128
		$D_s^- \to \phi\pi^-$		61 ± 7.81	1 ± 0.129
		$D_s^- \to 3\pi^\pm$		372 ± 19.3	6.12 ± 0.318
	$B_s \to D_s^- 3\pi^\pm$	$D_s^- \to \bar{K}^*K^-$	$138\ldots162$	517 ± 22.7	8.11 ± 0.357
		$D_s^- \to \phi\pi^-$		912 ± 30.2	14.3 ± 0.474
		$D_s^- \to 3\pi^\pm$		916 ± 30.3	14.4 ± 0.475

Table 7.1: Signal efficiencies of the B_{s1} and B_{s2}^* in decays of $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ estimated from the Monte Carlo samples.

	$B^+ \to \bar{D}^0\pi^+$	
	B_{s1}	B_{s2}^*
$\bar{D}^0 \to K^+\pi^-$	$1.4329 \cdot 10^{-03}$	$8.8928 \cdot 10^{-04}$

Table 7.2: Signal efficiencies of the B_{s1} and B_{s2}^* in decays of $B_s^{**} \to B^+K^-$ estimated from the Monte Carlo samples.

original selection cut used in the $B_s^{**} \to B^+K^-$ analysis [1]. The ratio of Monte Carlo to data events is:

$$\frac{N_{\mathrm{MC}}\left(n_{\mathrm{out}} > n_{\mathrm{cut}}\right)}{N_{\mathrm{data}}\left(n_{\mathrm{out}} > n_{\mathrm{cut}}\right)}. \tag{7.2}$$

For the selection of these events different cuts n_{cut} on the output n_{out} of the same neural network being used in the $B_s^{**} \to B^+K^-$ analysis are applied. Additionally, the numbers of signal events obtained from the data sample are sideband subtracted. To study the efficiency for the B_{s1} and B_{s2}^* separately, the events are selected in the corresponding Q value ranges having different sideband regions.

Figures 7.1 and 7.2 show the ratio as a function of the cut on the output of the neural network. The major part of the values are within a band of 10% uncertainty of the average value. Therefore, an uncertainty of 10% is assumed for the efficiency of the B_{s1} and B_{s2}^* in decays to $B_s^{**} \to B^+K^-$.

For lack of any clean signal in the $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ data sample the efficiency uncertainty of the B_{s1} and B_{s2}^* cannot be directly studied in this decay. Assuming the efficiency uncertainty being dominated by the efficiency uncertainty of the slow

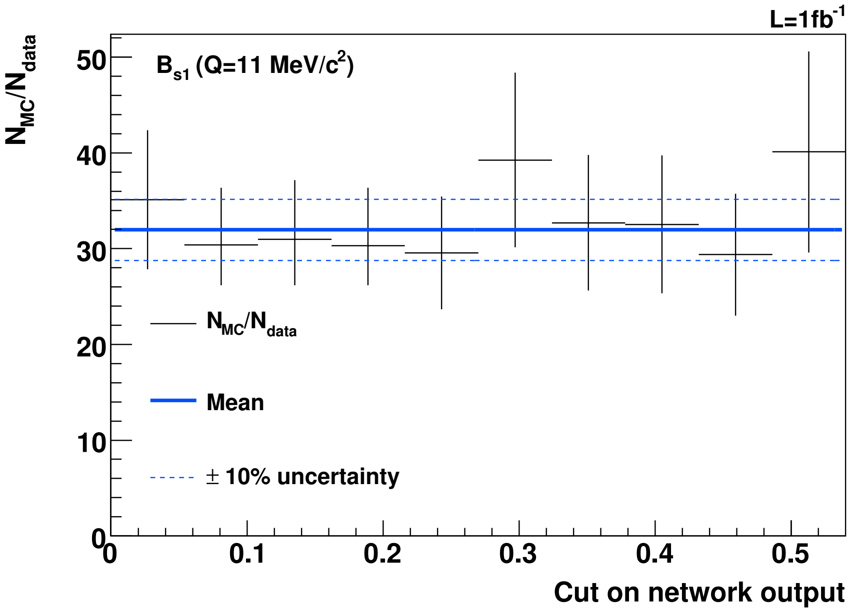


Figure 7.1: Ratio of the number of Monte Carlo to data events for the B_{s1} of the decay $B_s^{**} \to B^+K^-$ for different cuts on the output of the neural network.

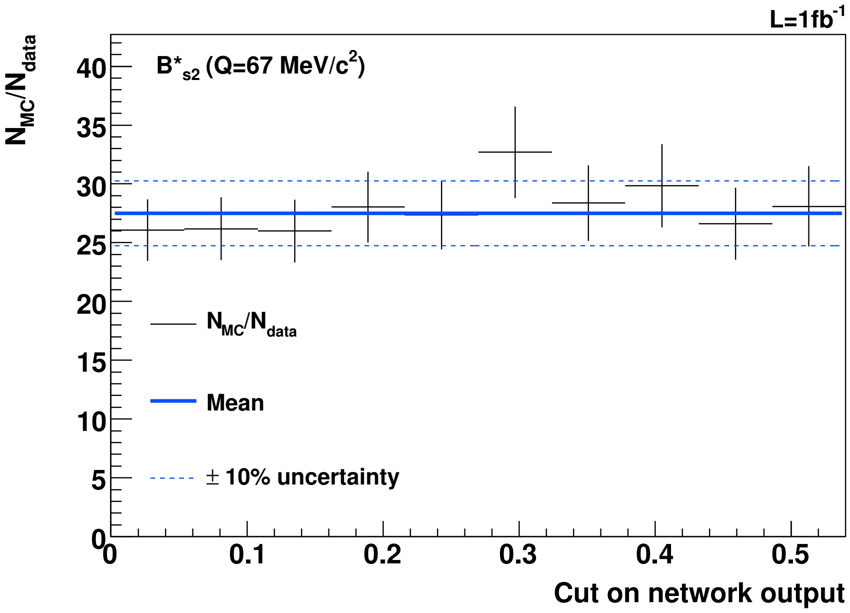


Figure 7.2: Ratio of the number of Monte Carlo to data events for the B_{s2}^* of the decay $B_s^{**} \to B^+K^-$ for different cuts on the output of the neural network.

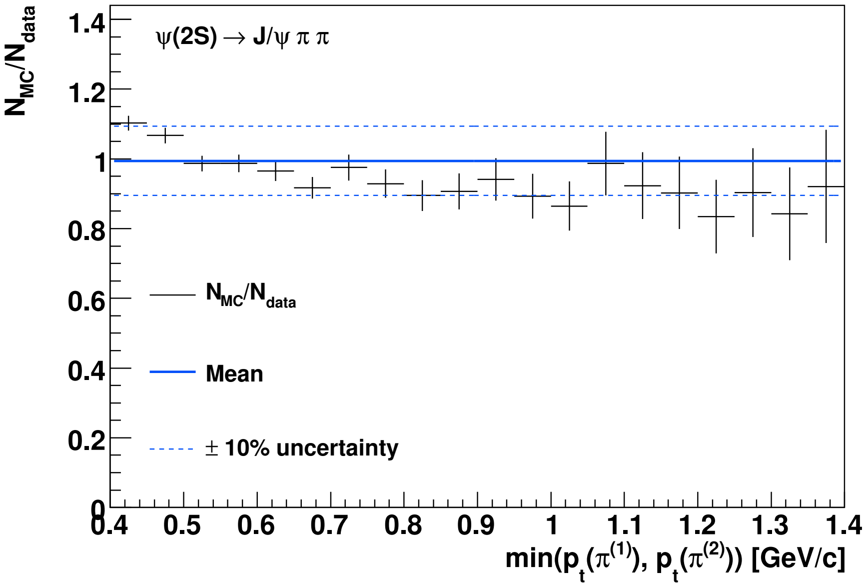


Figure 7.3: Ratio of the number of Monte Carlo to data events as a function of the minimal pion transverse momentum which is the minimum of the two transverse momenta of the pions coming from the decay $\psi(2S) \to J/\psi\pi\pi$. The distributions are normalised.

pions, the efficiency uncertainty can be estimated by studying an equivalent decay having similar final states and a similar decay topology. For this reason, the reference decay $\psi(2S) \to J/\psi\pi\pi$ was chosen whose data samples and selection procedure are taken from the $X(3872)$ analysis [78].

To study the efficiency uncertainty of the $\psi(2S) \to J/\psi\pi\pi$ sample the ratio of Monte Carlo and data events as a function of the minimal pion momentum is plotted. The minimal pion momentum is the minimal transverse momentum of the pions coming from the decay $\psi(2S) \to J/\psi\pi\pi$. The distributions of the transverse momentum are obtained by applying the selection procedure being used in the $X(3872)$ analysis and an additional sideband subtraction applied to the data sample. Figure 7.3 shows the momentum dependence of this ratio. Most of the values are within an uncertainty of 10%. Therefore, the signal efficiency uncertainty of the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ is estimated to be equally 10%.

7.2 Q Value Resolution of the B_s^{**} Signal

The Q value resolution is put into the fit as the width of the signal peaks. For estimating the Q value resolution the Monte Carlo sample used in the neural network training is studied. In order to estimate the resolution, events from the Monte Carlo sample are selected by applying the same selection algorithm as for the data samples, described in section 5.4. The difference between the measured and generated Q value of the selected events is plotted in histograms with 4 MeV/c^2 bin width. Each of the peaks of either the decay $B_s^{**} \to B_s\pi^+\pi^-$ and $B_s^{**} \to B_s^*\pi^+\pi^-$ is fitted using a double Gaussian function of the form

$$\mathcal{G}\left(Q\right) = \frac{N_1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(Q-Q_0)^2}{2\sigma_1}} + \frac{N_2}{\sqrt{2\pi}\sigma_2} e^{-\frac{(Q-Q_0)^2}{2\sigma_2}}. \tag{7.3}$$

Here, Q_0 is the common mean of the two Gaussian functions having the widths σ_1 and σ_2, respectively. N_1 and N_2 are the number of events of each single Gaussian component. For a better monotonous behaviour as a function of the measured Q value the ratio of the widths and numbers of events in each double Gaussian are fixed to have the following values:

$$N_2 = 0.33 \cdot N_1, \tag{7.4}$$

$$\sigma_2 = 2.63 \cdot \sigma_1. \tag{7.5}$$

The arbitrary factors of 0.33 and 2.63 are chosen since they describe the residual Q value distribution very well. In the appendix B these fits to the Monte Carlo residuals are shown.

Figures 7.4 and 7.5 show the Q value resolution for the different B_s^{**} decay channels. The Q value resolution at the Q values where the B_{s1} and B_{s2}^* signals are expected is shown in figures 7.6 and 7.7. Based on these measurements the widths of the signal peaks in the unbinned likelihood fit are set to the average values of the Q value resolutions:

$$\sigma_Q(B_{s1}) = 7.46 \text{ MeV/c}^2, \tag{7.6}$$

$$\sigma_Q(B_{s2}^*) = 7.14 \text{ MeV/c}^2. \tag{7.7}$$

7.3 Likelihood Function

In order to measure the branching ratio of the B_{s1} and B_{s2}^* in decays of $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ an extended unbinned likelihood fit to the Q value distribution of the data sample is performed. The used fitter is built using the MINUIT [76] minimiser

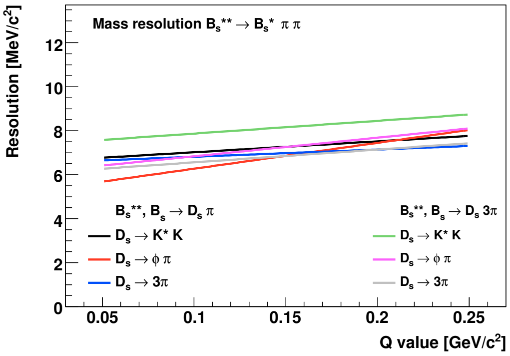


Figure 7.4: Q value resolution of the B_s^{**} signal for the decay $B_s^{**} \to B_s\pi^+\pi^-$ estimated from the Monte Carlo samples.

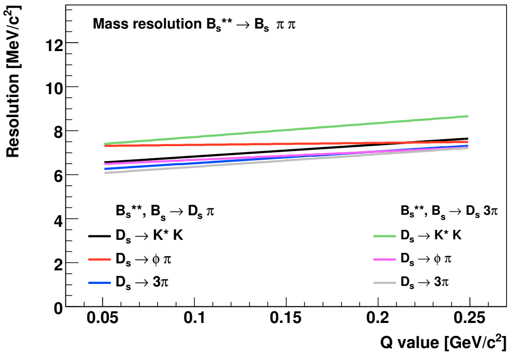


Figure 7.5: Q value resolution of the B_s^{**} signal for the decay $B_s^{**} \to B_s^{*}\pi^+\pi^-$ estimated from the Monte Carlo samples.

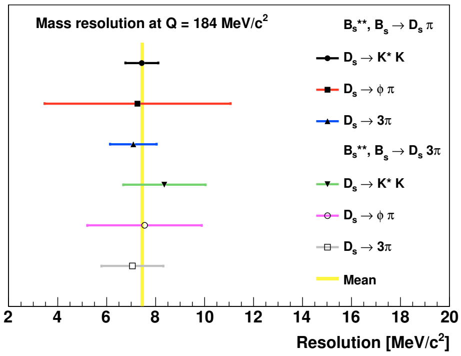


Figure 7.6: Q value resolution of the B_s^{**} signal for the decay $B_s^{**} \to B_s\pi^+\pi^-$ at the distinct Q value where the B_{s1} ($Q = 184$ MeV/c^2) is expected. The mass resolutions are estimated from Monte Carlo samples.

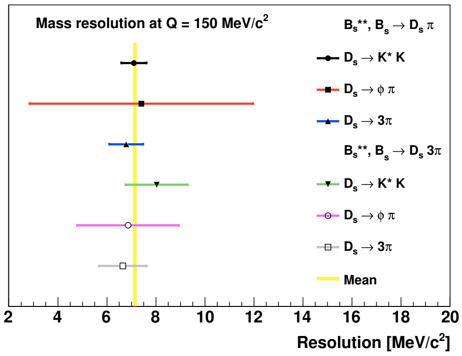


Figure 7.7: Q value resolution of the B_s^{**} signal for the decay $B_s^{**} \to B_s^{*}\pi^+\pi^-$ at the distinct Q value where the B_{s2}^{*} ($Q = 150$ MeV/c^2) is expected. The mass resolutions are estimated from Monte Carlo samples.

provided by the Root [77] package. The likelihood expression L of the fit function consists of a Gaussian signal component and an exponential background component:

$$L = \prod_i \left[\frac{2 \cdot \mathrm{BR}}{\frac{3}{2}\left(1 - \mathrm{BR}\right)} \cdot \widetilde{\mathrm{BR}} \cdot N'_{B_s^{**}} \cdot f_{sig}\left(Q_i\right) + N_{bkg} \cdot f_{bkg}\left(Q_i\right) \right]. \tag{7.8}$$

Here, the index i runs over all events in the data sample. The components of the likelihood functions in detail are:

- **Branching ratio BR**

  BR is the branching ratio of the B_s^{**} states and the factor

$$\frac{2 \cdot \mathrm{BR}}{\frac{3}{2}\left(1 - \mathrm{BR}\right)} \tag{7.9}$$

  is the relative branching ratio of the studied B_s^{**} decay. Based on the parameter BR a Bayesian limit is set on the branching ratio for the B_{s1} and B_{s2}^* in decays of $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$. The parameter BR is defined in the following way:

$$\mathrm{BR} = \frac{\frac{3}{2}\,\mathrm{BR}\left(B_s^{**} \to B_s\pi^+\pi^-\right)}{\frac{3}{2}\,\mathrm{BR}\left(B_s^{**} \to B_s\pi^+\pi^-\right) + 2\,\mathrm{BR}\left(B_s^{**} \to B^+K^-\right)}. \tag{7.10}$$

  Assuming that $B_s^{**} \to B_s\pi\pi$ and $B_s^{**} \to BK$ are the only decay modes of the B_s^{**} mesons this parameter is the absolute branching ratio for the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$. The advantage of its definition is the restriction to the interval $[0,1]$ avoiding divergences in the Bayesian limit calculation assuming a flat prior. The factors $\frac{3}{2}$ and 2 take into account the different ratios of decay probabilities into charged and uncharged particles considering the proper isospin selection rules:

$$\frac{\mathrm{BR}\left(B_s^{**} \to B_s\pi^+\pi^-\right)}{\mathrm{BR}\left(B_s^{**} \to B_s\pi^+\pi^-\right) + \mathrm{BR}\left(B_s^{**} \to B_s\pi^0\pi^0\right)} = \frac{2}{3}, \tag{7.11}$$

$$\frac{\mathrm{BR}\left(B_s^{**} \to B^+K^-\right)}{\mathrm{BR}\left(B_s^{**} \to B^+K^-\right) + \mathrm{BR}\left(B_s^{**} \to B^0K^0\right)} = \frac{1}{2}. \tag{7.12}$$

- **Branching ratio $\widetilde{\mathrm{BR}}$**

  $\widetilde{\mathrm{BR}}$ is the relative branching ratio of the B_s subdecays with respect to the reference decay taking into account the different efficiencies ϵ of the B_s and B

decay modes. It is defined by

$$\widetilde{\mathrm{BR}} = \frac{\sum_i \epsilon_i \mathrm{BR}_i}{\epsilon\left(B \to D\pi\right) \cdot \mathrm{BR}\left(B \to D\pi\right)} \tag{7.13}$$

where i runs over the six different B_s subdecays whose branching ratios are denoted by the parameter BR_i. Section 7.1 describes how the signal efficiencies of the subdecays are determined. For the branching ratios BR_i the world average values [2] are taken. The parameter $\widetilde{\mathrm{BR}}$ is free in the fit within a Gaussian constraint.

- **Signal events** $N'_{B_s^{**}}$

  Parameter $N'_{B_s^{**}}$ denotes the number of B_s^{**} signal candidates in the reference decay $B_s^{**} \to B^+K^-$. Section 7.5 describes how this number is obtained. This parameter is constricted with a Gaussian constraint in the fit.

- **Background events** N_{bkg}

  Number of background events in the $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ data sample.

- **Signal probability density function** f_{sig}

  The signal contribution of the likelihood expression is composed of a normalised Gaussian function

  $$f_{sig}\left(Q\right) = \mathcal{G}\left(Q; \mu_{\mathrm{Q}}, \sigma_{\mathrm{Q}}\right) \tag{7.14}$$

  having mean μ_{Q} and width σ_{Q}. In the fit procedure the mean and width of the signal component are kept fixed. For the signal width the mass resolution described in section 7.2 is taken since the intrinsic width of the narrow B_s^{**} states is negligible compared to the detector resolution [21]. The position of the signal peaks is determined by the known masses of the B_{s1} and B_{s2}^* states measured in the $B_s^{**} \to B^+K^-$ analysis [1]. Therefore, the $B_{s1} \to B_s\pi^+\pi^-$ signal mean is set to 184 MeV/c^2 and to 150 MeV/c^2 for the $B_{s2}^* \to B_s^*\pi^+\pi^-$ signal when fitting the $B_s\pi^+\pi^-$ samples. The B_{s2}^* signal is shifted downwards by the energy of the undetected photon from the B_s^* decay.

- **Background probability density function** f_{bkg}

  Since the background is mainly combinatorial background and should be zero at $Q = 0$ due to physical reasons it is modelled by a linear term times an exponential function

  $$f_{bkg}\left(Q\right) = \frac{Q \cdot \exp\left(\alpha \cdot Q\right)}{\int_{Q_{min}}^{Q_{max}} Q' \cdot \exp\left(\alpha \cdot Q'\right) \mathrm{d}Q'}. \tag{7.15}$$

> Here, α is the only free parameter intended to describe the background distribution. The background function is normalised within the fitting range from $Q_{min} = 0$ to $Q_{max} = 0.25$ GeV/c^2.

In the fit the negative log likelihood function $\mathcal{L}$ being defined as

$$\begin{aligned}\mathcal{L} &= -2 \cdot \ln L \\ &+ 2 \cdot \frac{2 \cdot \mathrm{BR}}{\frac{3}{2}(1-\mathrm{BR})} \cdot \widetilde{\mathrm{BR}} \cdot N'_{B_s^{**}} + 2 \cdot N_{bkg} \\ &+ \left(\frac{\widetilde{\mathrm{BR}} - \mu_{\widetilde{\mathrm{BR}}}}{\sigma_{\widetilde{\mathrm{BR}}}}\right)^2 + \left(\frac{N'_{B_s^{**}} - \mu_{N'_{B_s^{**}}}}{\sigma_{N'_{B_s^{**}}}}\right)^2\end{aligned} \tag{7.16}$$

is minimised. Here, the second line in equation (7.16) implies the extended part of the likelihood fit providing proper uncertainties. As above mentioned, the mean and width of the signal peak are kept fixed during the fit. The mean values of the signal peaks are taken from the CDF observation of orbitally excited B_s^{**} mesons in decays of $B_s^{**} \to B^+K^-$ [1] whereas the widths of the signal peaks are set to the Q value resolution, described in section 7.2. The number of signal candidates of the reference decay and the branching ratio $\widetilde{\mathrm{BR}}$ are floating in the fit but confined by Gaussian constraints implemented in the third line of equation (7.16). The number of signal candidates in the reference decay is again taken from reference [1] and quoted for completeness in table 7.3. Due to the Gaussian constraints the likelihood value of the fit gets increasingly worse as it moves the floating parameters far away from the values put initially into the fit to which they are constrained

7.4 Bayesian Limit

The Bayesian limit on the branching ratio of the narrow B_s^{**} states is calculated according to the formalism described in section 32.3.1 of the "Review of Particle Physics" [2]. In order to set the limit a fit to the data sample is performed as described in section 7.3 which is referred to as the free fit. In addition to the free fit a set of consecutive fits is done where the parameter BR for the total branching ratio is fixed to values in the interval $[0, 1)$. Each fit returns a negative log likelihood value $\mathcal{L}$ as defined in equation (7.16) which is transformed back into the likelihood value ΔL as

$$\Delta L = \exp\left(\frac{-\mathcal{L} + \mathcal{L}_0}{2}\right). \tag{7.17}$$

Here, $\mathcal{L}_0$ is the negative log likelihood value of the free fit. $\mathcal{L}_0$ is constant for the complete set of consecutive fits having a fixed value for the branching ratio. Thus, it scales ΔL by a constant factor which has no effects on the result but avoids numerical problems with extreme values of ΔL.

The Bayesian posterior probability density function is calculated from the likelihood value as defined by

$$p\left(\vec{a}\middle|\,\vec{x}\right) = \frac{\Delta L\left(\vec{x}\middle|\,\vec{a}\right)\pi\left(\vec{a}\right)}{\int \Delta L\left(\vec{x}\middle|\,\vec{a}'\right)\pi\left(\vec{a}'\right)\mathrm{d}\vec{a}'} \tag{7.18}$$

where $\pi\left(\vec{a}\right)$ is the prior probability density function which is assumed being flat:

$$\pi\left(\vec{a}\right) = 1 \qquad \text{for all } \vec{a}. \tag{7.19}$$

The denominator in equation (7.18) is merely intended to normalise the posterior probability density function. In the single parameter case a credibility interval can be determined containing a given fraction $(1-\beta)$ of the probability:

$$1-\beta = \int_{a_{\mathrm{lo}}}^{a_{\mathrm{up}}} p\left(a\middle|\,\vec{x}\right)\mathrm{d}a. \tag{7.20}$$

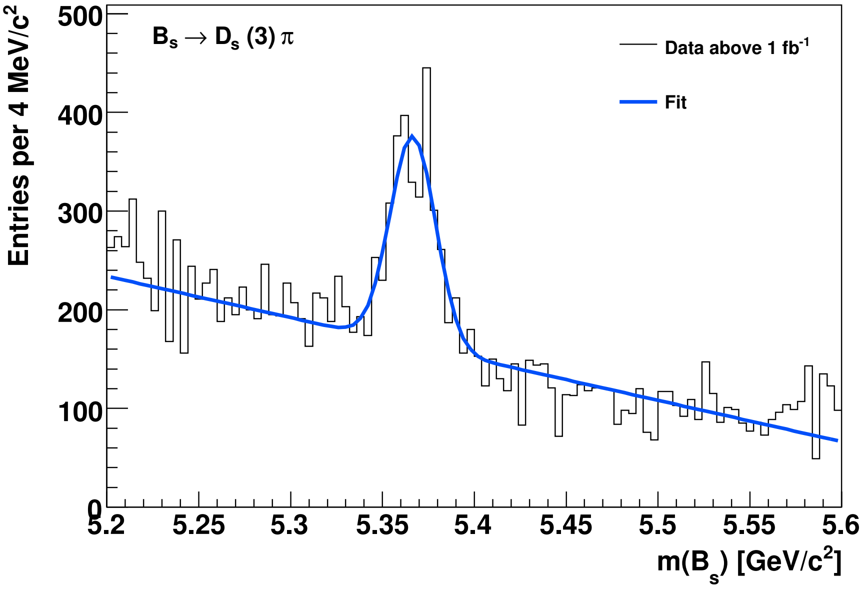
For the upper limit a_{up} on the branching ratios of the narrow B_s^{**} states the single parameter a is the total branching ratio $a = \mathrm{BR}$ and $a_{\mathrm{lo}} = 0$.

7.5 Candidates in the Reference Channel

The number of the signal events in the reference channel $B_s^{**} \to B^+K^-$ is taken from the $B_s^{**} \to B^+K^-$ analysis [1]. In this analysis a data sample of 1 fb^{-1} is used. Therefore, the measured event numbers and uncertainties are extrapolated in a way that they correspond to a data sample having 2.8 fb^{-1} as it is the case for the $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ data sample.

Assuming that the number of B_s^{**} mesons scales with the number of B_s mesons, the number of events given in the reference channel are multiplied by the ratio of B_s candidates in the full $B_s\pi^+\pi^-$ data sample of 2.8 fb^{-1} and below 1 fb^{-1} in the same data sample. The number of B_s mesons in the $B_s\pi^+\pi^-$ data sample is obtained from fits to the invariant mass of the B_s. The events used in the fit sample are selected by the same selection cuts as developed in context of the B_s mixing analysis [65, 66, 67]. In figure 7.8 the fits to the B_s^{**} subsample are shown. The signal is fitted with a Gaussian function and the background shape is modelled by a linear function. Table 7.3 gives the numbers of measured and extrapolated signal events in the reference channel.

Despite having almost three times the integrated luminosity the number of signal events are merely doubled. This is due to the fact that the CDF II experiment scales



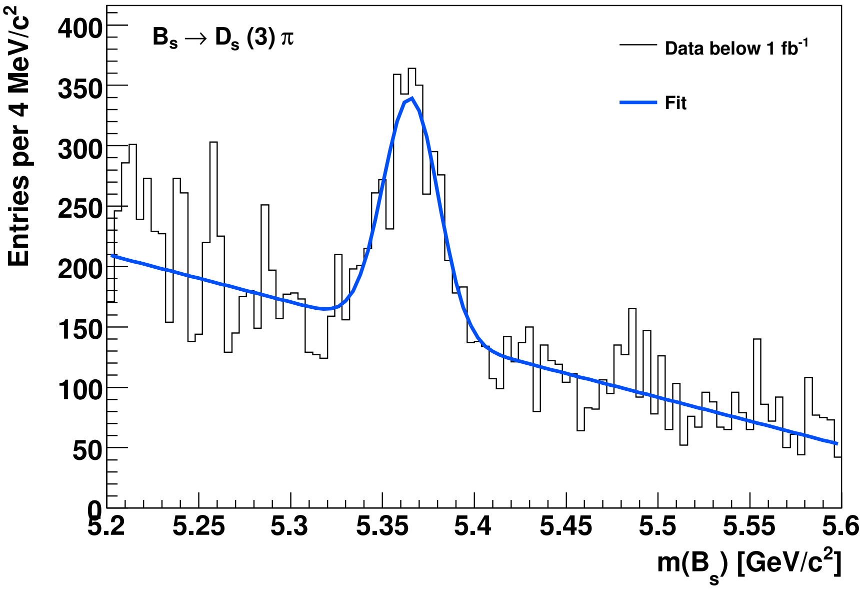


Figure 7.8: Fit to the invariant mass distribution of the B_s in different B_s^{**} data subsamples. The number of B_s candidates below and above 1.0 fb^{-1} are used to extrapolate the number of signal events in the reference channel.

$B_s^{**} \to B^+K^-$	measured in 1 fb^{-1}	extrapolated to 2.8 fb^{-1}
B_{s1}	20.66 ± 7.12	39.09 ± 13.47
B_{s2}^*	55.74 ± 19.2	105.46 ± 36.33

Table 7.3: Number of signal events in the reference channel $B_s^{**} \to B^+K^-$ with $B^+ \to \bar{D}^0\pi^+$, selected in 1 fb^{-1} of data and extrapolated to 2.8 fb^{-1}. The measured numbers are taken from the $B_s^{**} \to B^+K^-$ analysis [1].

down the triggers used for B-physics at high luminosities which happend for the last periods of data taking.

8 Results

In this chapter, the results of the unbinned maximum likelihood fits to the B^{**} and B_s^{**} data samples are presented. The fits are described in detail in chapter 6 and in chapter 7, respectively.

8.1 B^{**} Results

The masses of the narrow B^{**} states are determined by a likelihood fit to the Q value of the combined B^{**} data sample. In the same time, the width of the B_2^* state is also measured. The presentation of the results of the B^{**} analysis starts with describing the study of systematic uncertainties. Afterwards, the final results of the fit are presented.

8.1.1 Systematic Uncertainties

There are several sources of systematic uncertainties on the mass and width measurement. Their impact on the fit parameters is summarised in table 8.1. The different sources of uncertainties are:

- **Mass scale**

  Mass scale uncertainties arise from the calibration of the momentum scale. To determine the mass scale uncertainty the $\psi\left(2S\right)$ is reconstructed in decays of $\psi\left(2S\right) \rightarrow J/\psi\pi^+\pi^-$ having a similar Q value as the B^{**} decay. The difference between the measured $\psi\left(2S\right)$ mass and the world average value [2] is taken as mass scale uncertainty for the B^{**} measurement.

- **Signal efficiency**

  To determine the uncertainty due to the mass dependent signal efficiency Monte Carlo samples are generated. Each of these samples is fitted using the default fit model. The mass difference between generated and measured mass is plotted and fitted with a single Gaussian. The systematic uncertainty is determined by the mean of the difference distribution.

Source	$Q(B_2^*)$	$\Gamma(B_2^*)$	$m(B_2^*) - m(B_1)$	$Q(B_1)$
Mass scale	$+0.20$ -0.20	– –	$+0.003$ -0.003	$+0.20$ -0.20
Signal efficiency	$+0$ -0.03	$+0.40$ -0	$+0$ -0.30	$+0.40$ -0
Fit model	$+0.35$ -0.29	$+2.09$ -1.45	$+0.65$ -0.94	$+0.74$ -0.98
Fit bias and signal model	$+0$ -0.28	$+0.44$ -0	$+0$ -0.21	$+0.07$ -0.21
Background model	$+0.24$ -0	$+0$ -1.63	$+0.16$ -0	$+0$ -0.09
Detector resolution	$+0.001$ -0.001	$+0$ -0.40	$+0$ -0.003	$+0.005$ -0
Broad B^{**} states	$+0.65$ -0.51	$+2.27$ -9.92	$+0.89$ -0.95	$+0.96$ -0.92
Total	$+0.81$ -0.68	$+3.14$ -10.16	$+1.11$ -1.39	$+1.29$ -1.38

Table 8.1: Summary of the systematic uncertainties on the B^{**} signal parameters given in MeV/c^2.

- **Gaussian constraints**

  Assumptions are included in the fit as Gaussian constraints. Thus, the systematic uncertainties are already enclosed in the statistical error obtained from the fit. To separate the statistical and systematic error, different fits are made to the data. Fits having floating Gaussian constraints give the total error in the form:

  $$\sigma_{\text{total}} = \sqrt{\sigma_{\text{sys}}^2 + \sigma_{\text{stat}}^2}. \tag{8.1}$$

  Fits having fixed Gaussian constraints yield merely the statistical error. The systematic error can be determined on the basis of these two types of fits:

  $$\sigma_{\text{sys}} = \sqrt{\sigma_{\text{total}}^2 - \sigma_{\text{stat}}^2}. \tag{8.2}$$

- **Fit bias and signal model**

  The signal is modelled by a Breit-Wigner function convoluted with the detector resolution. To check the systematic uncertainties due to the signal model a set of Monte Carlo samples is generated according to a modified signal function. Here, the Breit-Wigner function used for the generation is multiplied by a phase space factor before convoluting. The Monte Carlo samples are fitted with the default Breit-Wigner function and the difference between the generated and measured masses is plotted. The means of these differences are taken as systematic uncertainties.

- **Background model**

  To study the systematic impact of the background model Monte Carlo samples are generated having an alternate background shape. This alternate background contains an additional part of a polynomial times an exponential function. Two types of fits are performed on these samples: The default fit having the default background component and the alternate fit having the alternate background shape. The systematic uncertainty is derived from the means of the differences between the default and the alternate fit.

- **Detector resolution**

  The mass resolution model enters the final fit as described in section 6.1. To estimate the systematic uncertainty Monte Carlo samples are generated with a Gaussian detector resolution whose width is increased by 20%. The samples are fitted and the difference in the Monte Carlo samples is plotted. Fits to the difference show that the shift is small compared to the uncertainty of the shift. Therefore, the uncertainty of the shift is taken as systematic uncertainty for the detector resolution.

- **Broad B^{**} states**

  Theoretical predictions about the position and width of the broad B^{**} states vary highly. To study the systematic impact caused by the broad states, the data sample is refitted using two additional Breit-Wigner functions for modelling the broad states. Different locations and widths are put into the Breit-Wigner functions of the broad states. From the fit to the data sample, two quantities are derived, called high value and low value. The high value is the central value of a parameter plus the positive uncertainty σ_+ calculated by

  $$\sigma_+ = \sqrt{\sigma^2_{(\text{broad states},+)} - \sigma^2_{(\text{baseline},+)}} \tag{8.3}$$

  where $\sigma_{(\text{broad states/baseline},+)}$ is the positive uncertainty of either the fit taking into account the broad states or the baseline fit. This formula is used to avoid double-counting the statistical errors. The low value is defined in a similar way using the negative uncertainty instead. The positive systematic uncertainty is given by the difference between the parameter value from the baseline fit and high value from the fit with broad states included. The negative systematic

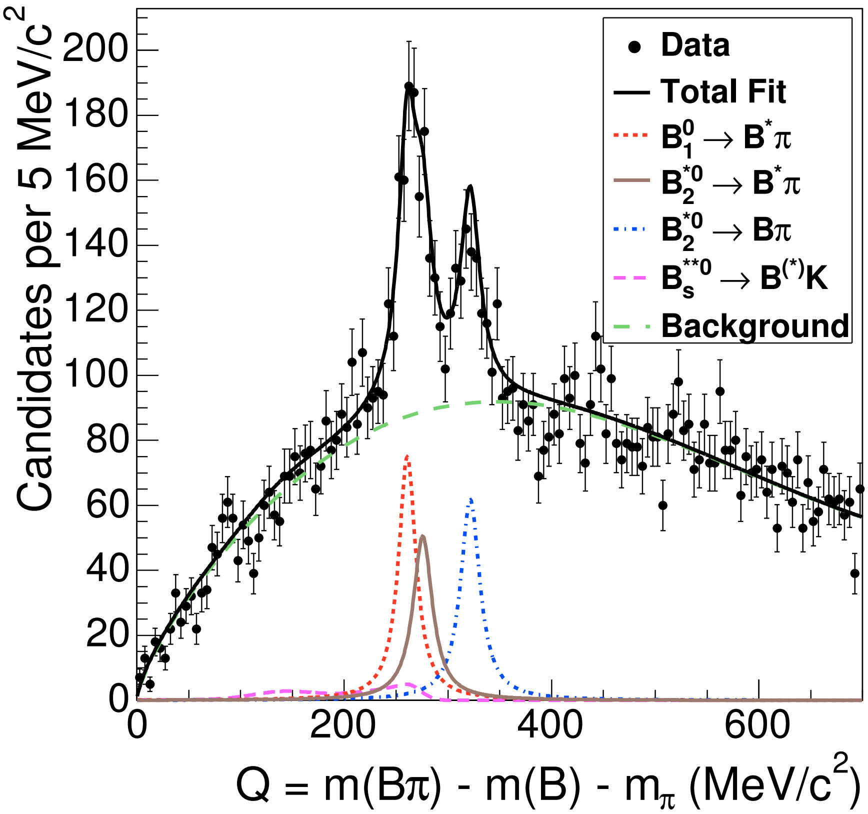


Figure 8.1: Result of the B^{**} fit combining all three decay modes $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$, $B^+ \to \bar{D}^0\pi^+$ and $B^+ \to \bar{D}^0\, 3\pi^\pm$.

uncertainty is derived analogously. For the finally quoted systematics the maximum of any tested combinations of the broads states is taken.

8.1.2 Fit Results for B^{**}

Figure 8.1 shows the fit to the combined B^{**} data sample which comprises the decays $B^{**} \to B^{(*)+}\pi^-$ with $B^+ \to J/\psi K^+$, $B^+ \to \bar{D}^0\pi^+$ and $B^+ \to \bar{D}^0\, 3\pi^\pm$. Table 8.2 summarises the parameter values obtained from the fit.

From the fit the following signal parameters are obtained:

$$Q(B_2^*) = 321.5^{+1.7}_{-1.8}\ \text{(stat)}\ ^{+0.9}_{-0.7}\ \text{(syst) MeV/c}^2, \tag{8.4}$$

$$m(B_2^*) - m(B_1) = 14.9^{+2.2}_{-2.5}\ \text{(stat)}\ ^{+1.2}_{-1.4}\ \text{(syst) MeV/c}^2, \tag{8.5}$$

$$\Gamma(B_2^*) = 22.7^{+3.8}_{-3.2}\ \text{(stat)}\ ^{+3.2}_{-10.2}\ \text{(syst) MeV/c}^2. \tag{8.6}$$

Parameter	Value	Parabolic Error	MINOS Errors
$Q(B_2^*)$ [MeV/c^2]	321.5	1.7	$(+1.7, -1.8)$
$\Gamma(B_2^*)$ [MeV/c^2]	22.7	3.8	$(+4.3, -3.5)$
$m(B_2^*) - m(B_1)$ [MeV/c^2]	14.9	2.4	$(+2.3, -2.7)$
$B_1 \to B^*\pi$ events	503	105	$(+114, -99)$
$B_2^* \to B\pi$ events	385	53	$(+55, -51)$
Frac. of B_2^* events	0.91	0.20	$(+0.20, -0.21)$
$E(\gamma)$ [MeV/c^2]	45.78	0.35	$(+0.35, -0.35)$
$\Gamma(B_1)/\Gamma(B_2^*)$	0.95	0.17	$(+0.17, -0.16)$
B_{s1} events	22	12	$(+12, -12)$
B_{s2}^* events	74	31	$(+31, -30)$
Bkg α	0.72	0.05	$(+0.05, -0.05)$
Bkg γ	2.21	0.24	$(+0.24, -0.23)$
Bkg $Q_{\max}$ [MeV/c^2]	335	9	$(+9, -8)$
Bkg events	10276	279	$(+289, -267)$
Extra Bkg events	2344	243	$(+226, -261)$

Table 8.2: Summary of the fit parameter values from the fit to the combined B^{**} data sample. The first five parameters are related to the signal and the next five parameters are Gaussian constraints. The last five parameters are used in the background parametrisation.

The absolute mass of the B_2^* can then be derived by adding the world average value [2] of the B^+ and the charged pion masses to the Q value of the B_2^*. The uncertainty of the B^+ and the charged pion mass are added in quadrature to the systematic uncertainty. The mass of the B_1 is obtained by subtracting the mass difference $m(B_2^*) - m(B_1)$ from the Q value of the B_2^* before adding the B^+ and the charged pion mass. The uncertainties on the world average values are again added to the systematic uncertainty. The absolute masses are calculated to be

$$m(B_2^*) = 5740.2^{+1.7}_{-1.8}\ \text{(stat)}\ ^{+0.9}_{-0.8}\ \text{(syst) MeV/c}^2, \tag{8.7}$$

$$m(B_1) = 5725.3^{+1.6}_{-2.2}\ \text{(stat)}\ ^{+1.4}_{-1.5}\ \text{(syst) MeV/c}^2. \tag{8.8}$$

Since the B^{**} production is not measured here, only the statistical uncertainties are given on the number of measured events with number of signal events to be:

$$N(B_2^* \to B^+\pi^-) = 385^{+48}_{-45}\ \text{(stat)}, \tag{8.9}$$

$$N(B_2^* \to B^{*+}\pi^-) = 351^{+48}_{-45}\ \text{(stat)}, \tag{8.10}$$

$$N(B_1 \to B^{*+}\pi^-) = 503^{+75}_{-68}\ \text{(stat)}. \tag{8.11}$$

Although not measured in this analysis, the ratio of the widths $\frac{\Gamma(B_1)}{\Gamma(B_2^*)}$ is consistent

Parameter	Value	Error	
BR	−1.791	4.639	
$\widetilde{\mathrm{BR}}$	0.183	0.0546	constrained
$N'_{B_s^{**}}$	39.09	13.5	constrained
μ_{Q}	0.184	fixed	$\left[\mathrm{GeV/c^2}\right]$
σ_{Q}	0.00746	fixed	$\left[\mathrm{GeV/c^2}\right]$
N_{bkg}	152.1	13.55	
α	4.235	1.496	

Table 8.3: Values of the fit parameters from the free fit for the B_{s1}.

Parameter	Value	Error	
BR	0.155	0.1074	
$\widetilde{\mathrm{BR}}$	0.282	0.0848	constrained
$N'_{B_s^{**}}$	105.5	36.49	constrained
μ_{Q}	0.15	fixed	$\left[\mathrm{GeV/c^2}\right]$
σ_{Q}	0.00714	fixed	$\left[\mathrm{GeV/c^2}\right]$
N_{bkg}	138.7	12.47	
α	4.724	1.617	

Table 8.4: Values of the fit parameters from the free fit for the B_{s2}^{*}.

with theoretical predictions of 0.9 ± 0.2 [14] as is the fraction of the branching ratios $\frac{\mathrm{BR}(B_2^* \to B\pi)}{BR(B_2^* \to B^*\pi)}$ with 1.1 ± 0.3 [79]. The measured signal parameters themselves are also consistent with theoretical predictions [18, 20].

In summary, this is the most precise mass measurement of the narrow B^{**} states and the first width measurement of the B_2^*. However, the mass measurement is still limited by statistics so further improvements from the CDF experiment are possible.

8.2 B_s^{**} Results

In this section, the results of the unbinned maximum likelihood fit to the Q value of the B_s^{**} data sample are given. The results are obtained by applying the fit function described in chapter 7 to the data samples selected as shown in chapter 5.

Tables 8.3 and 8.4 give the values of the parameters determined by the free fit, where only the mean and width of the signal peaks are fixed. The corresponding plots in figures 8.2 and 8.3 do not show evidence for a significant signal in the B_s^{**} data sample. Therefore, a Bayesian limit is set on the branching ratios of the B_{s1} and B_{s2}^{*} states. The Bayesian limit is calculated assuming a flat prior for the branching ratio.

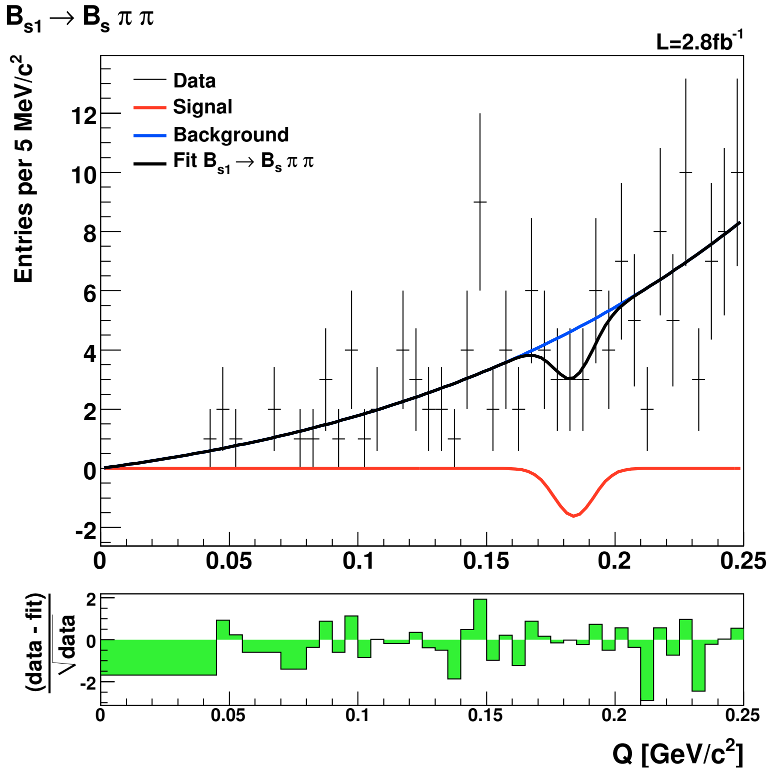


Figure 8.2: Result of the free fit for the B_{s1} in decays of $B_{s1} \to B_s\pi^+\pi^-$.

To see whether there are any other minima of the fit function, the negative log likelihood function is scanned for all floating parameters. After the free fit has converged, all fit parameters of the likelihood function are fixed except for one. The likelihood function is plotted as a function of this unfixed parameter. Figures 8.4 through 8.13 show the projection of the likelihood function for all parameters being floating in the free fit. The plots confirm that the fit converges well.

In order to set an upper limit on the branching ratios for either the B_{s1} and B_{s2}^* states the Bayesian approach is used assuming a flat prior for the branching ratio. The is consecutively refitted with different fixed values for the branching ratio BR. From these fits the fit posterior probability density as a function of the branching ratio BR is obtained and shown in figures 8.14 and 8.15.

The integrated posterior probability density function yields a direct value for the upper limit of the branching ratio for the B_{s1} and B_{s2}^*. The integral of the posterior probability density as a function of the branching ratio is plotted in figures 8.16 and 8.17. For the B_{s1} the upper limit is determined to be for a 95% credibility level 0.54 and for B_{s2}^* it is 0.53.

The rather loose limit on the branching ratios of the B_{s1} and B_{s2}^* states can be

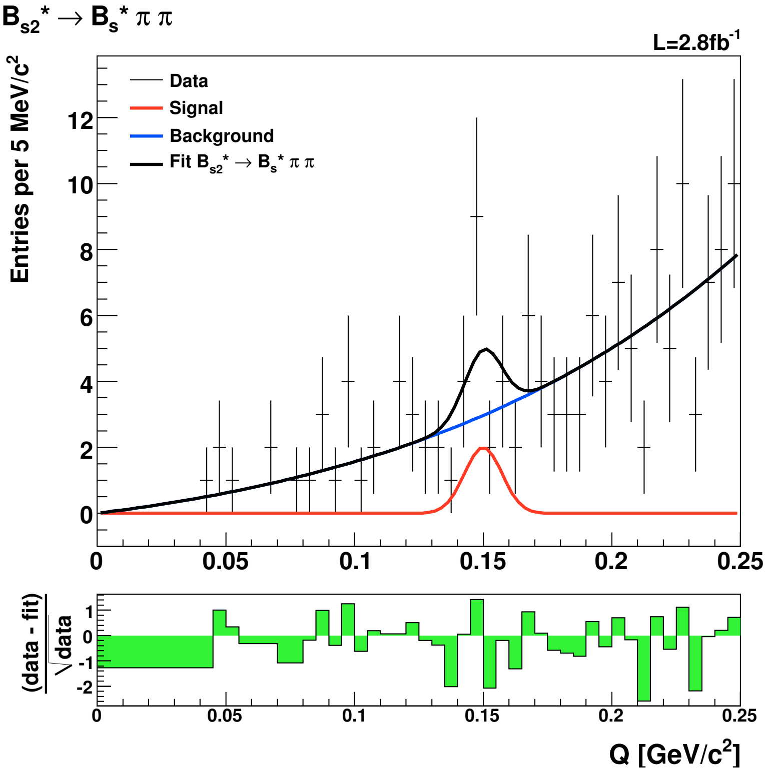


Figure 8.3: Result of the free fit for the B_{s2}^* in decays of $B_{s2}^* \to B_s^*\pi^+\pi^-$.

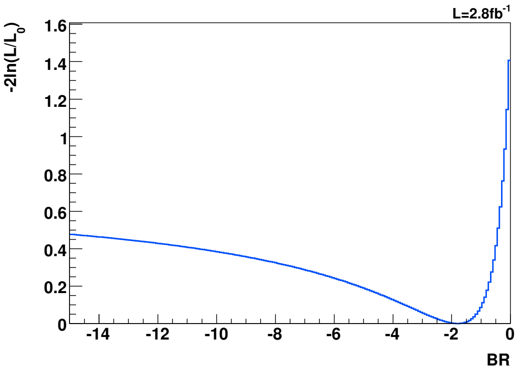


Figure 8.4: Negative log likelihood scan for the parameter BR in the free fit for the B_{s1}.

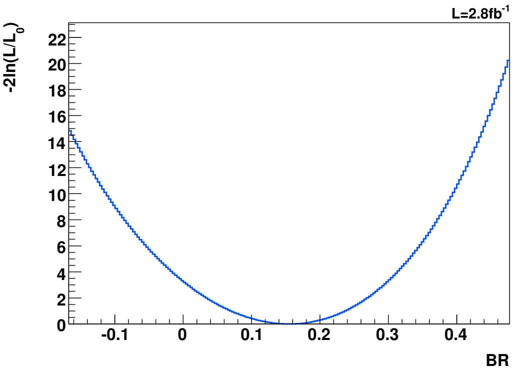


Figure 8.5: Negative log likelihood scan for the parameter BR in the free fit for the B_{s2}^*.

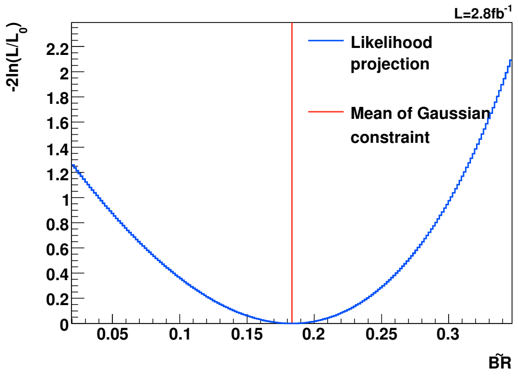


Figure 8.6: Negative log likelihood scan for the parameter $\widetilde{\mathrm{BR}}$ in the free fit for the B_{s1}.

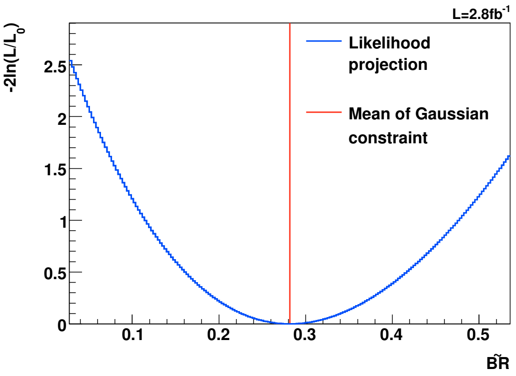


Figure 8.7: Negative log likelihood scan for the parameter $\widetilde{\mathrm{BR}}$ in the free fit for the B_{s2}^*.

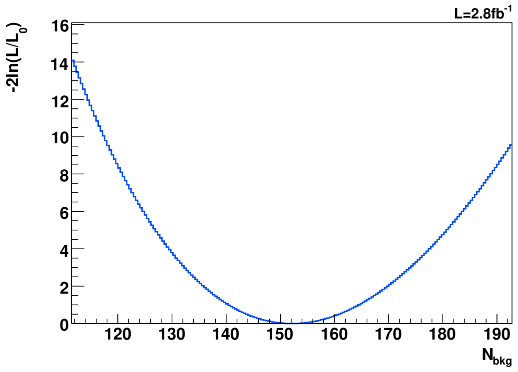


Figure 8.8: Negative log likelihood scan for the parameter N_{bkg} in the free fit for the B_{s1}.

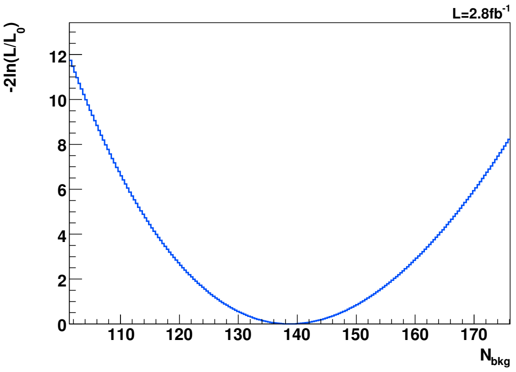


Figure 8.9: Negative log likelihood scan for the parameter N_{bkg} in the free fit for the B_{s2}^*.

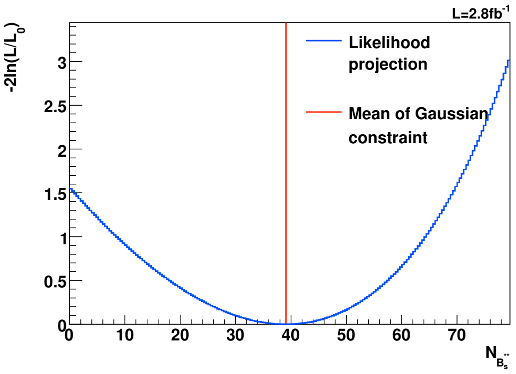


Figure 8.10: Negative log likelihood scan for the parameter $N'_{B_s^{**}}$ in the free fit for the B_{s1}.

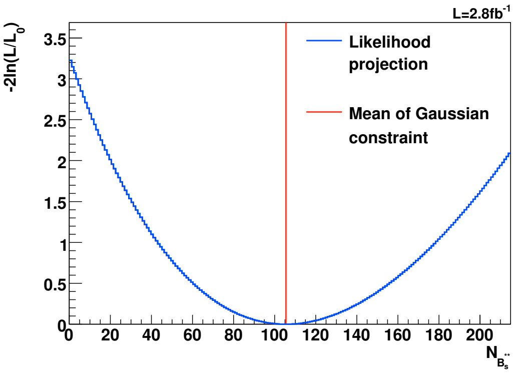


Figure 8.11: Negative log likelihood scan for the parameter $N'_{B_s^{**}}$ in the free fit for the B^*_{s2}.

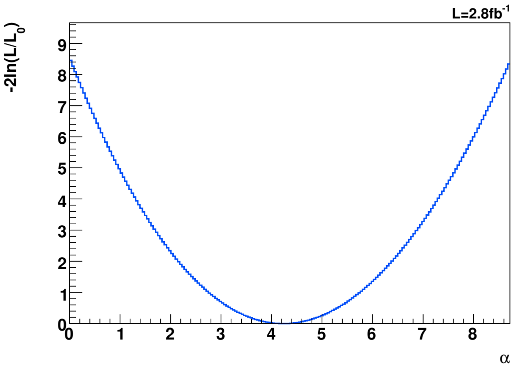


Figure 8.12: Negative log likelihood scan for the parameter α in the free fit for the B_{s1}.

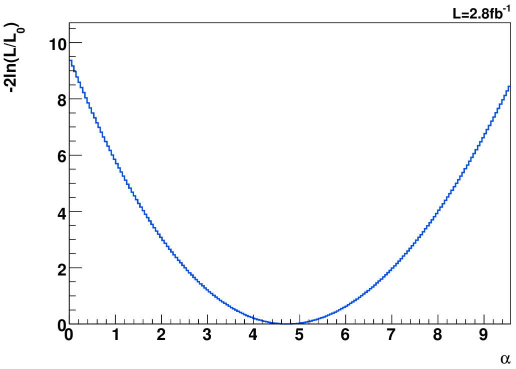


Figure 8.13: Negative log likelihood scan for the parameter α in the free fit for the B^*_{s2}.

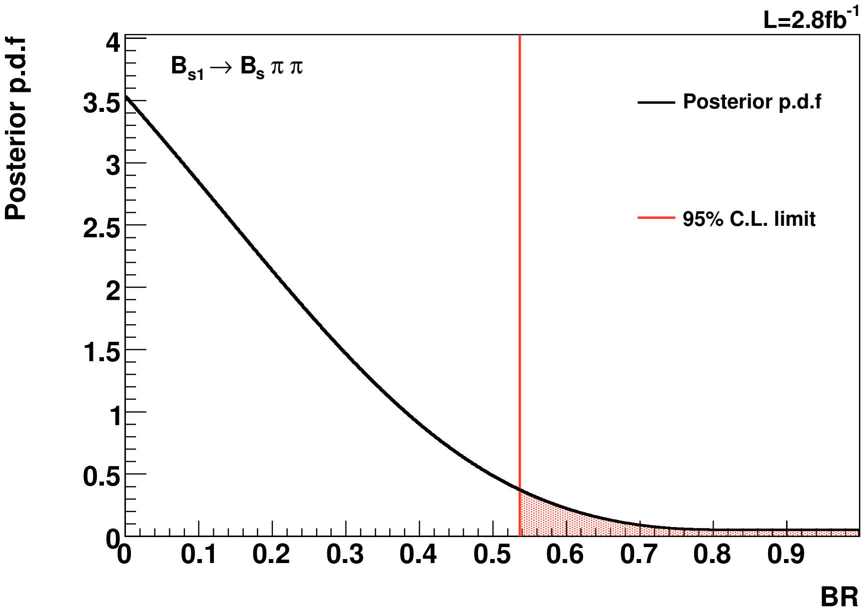


Figure 8.14: The posterior probability density function of fit for the B_{s1} as a function of the branching ratio BR.

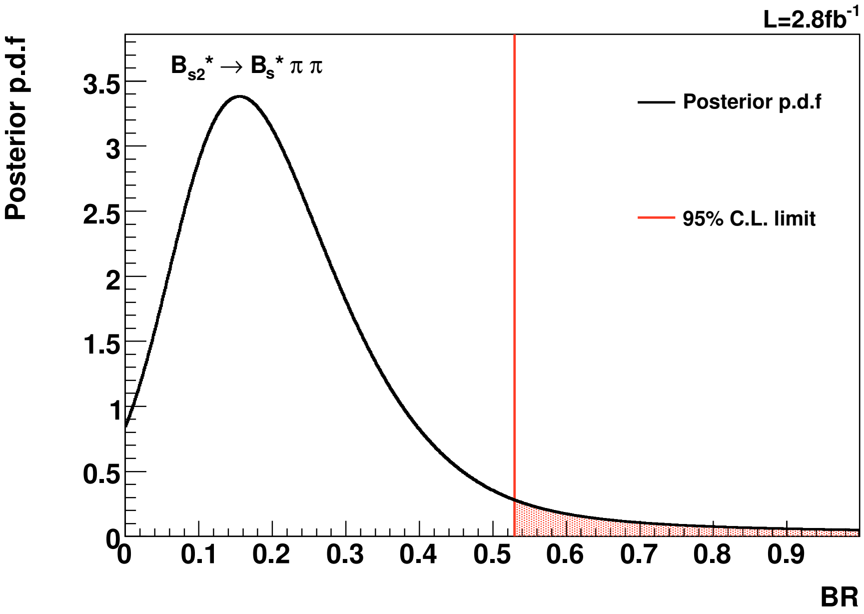


Figure 8.15: The posterior probability density function of fit for the B_{s2}^* as a function of the branching ratio BR.

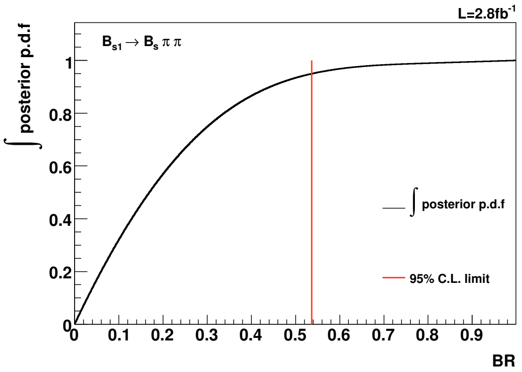


Figure 8.16: The integrated posterior probability density function the B_{s1} as a function of the branching ratio BR.

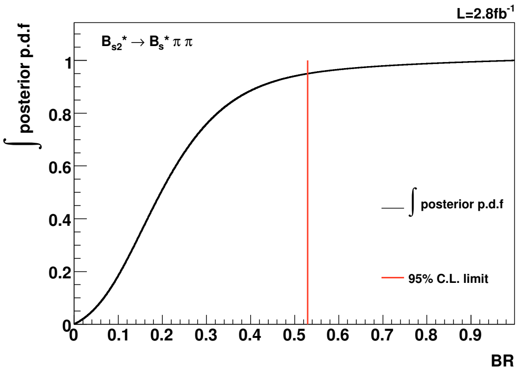


Figure 8.17: The integrated posterior probability density function of the B_{s2}^* as a function of the branching ratio BR.

Quantity	Relative uncertainty	Affected parameter	Reference
$\mathrm{BR}(B_s \to D_s\pi)$	28%	$\widetilde{\mathrm{BR}}$	[2]
$\mathrm{BR}(B_s \to D_s 3\pi)$	39%	$\widetilde{\mathrm{BR}}$	[2]
$N(B_s^{**} \to B^+K^-)$	31%	$N'_{B_s^{**}}$	[1]

Table 8.5: Quantities entering the fit function having large uncertainties. It is also listed which fit function parameters they affect and the source from where the values are obtained.

understood with regard to the assumptions put into the fit. The rather high uncertainties included in the fit as Gaussian constraints prevent the limit of being more stringent. Especially the uncertainties of the branching ratios of the B_s subdecays and the uncertainty of the number of candidates in the reference channel are quite large. Table 8.5 gives an overview about the quantities having the highest uncertainties which enter the fit.

9 Conclusion and Outlook

It was the aim of the analyses presented in this thesis to search for B^{**} and B_s^{**} mesons. Mesons consisting of a heavy and a light quark play a similar role for studying quantum chromodynamics as the hydrogen atom did for understanding quantum electrodynamics. Thereby, the heavy quark plays the role of the nucleus and the light quark takes the part of the electron. According to this analogy, a hyperfine mass splitting of the excited states emerges due to the spin-orbit coupling of the heavy and the light quark.

In this analysis, heavy-light mesons in decays of $B^{**} \to B^{(*)+}\pi^-$ and $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ were studied. To reconstruct the signal candidates three exclusive decay modes of the B^+ and six decay modes of the B_s were used. The data samples were collected with the CDF II detector located at the Fermi National Accelerator Laboratory in the vicinity of Chicago. The used data corresponds to an integrated luminosity of 1.7 fb^{-1} for the B^{**} sample and 2.8 fb^{-1} for the B_s^{**} sample.

The selection of the candidates is based on artificial neural networks. The networks were trained on simulated events for modelling the signal and real data events as background. In contrast to purely cut based selections, neural networks have the advantage of taking into account correlations among the input variables yielding a better signal to noise ratio. The selection was optimised to maximise the signal significance being generally defined as a function of the number of signal and of background events.

In the decays of $B^{**} \to B^{(*)+}\pi^-$ clean signals were seen which are interpreted as the two narrow B^{**} states. From the unbinned maximum likelihood fit to the Q value distribution the masses of the B_1 and B_2^* are determined to be

$$\begin{aligned} m(B_2^*) &= 5740.2^{+1.7}_{-1.8}\ (\text{stat})\ ^{+0.9}_{-0.8}\ (\text{syst})\ \text{MeV/c}^2, \\ m(B_1) &= 5725.3^{+1.6}_{-2.2}\ (\text{stat})\ ^{+1.4}_{-1.5}\ (\text{syst})\ \text{MeV/c}^2. \end{aligned}$$

The width of the B_2^* was measured to be

$$\Gamma(B_2^*) = 22.7^{+3.8}_{-3.2}\ (\text{stat})\ ^{+3.2}_{-10.2}\ (\text{syst})\ \text{MeV/c}^2.$$

This is the most precise mass measurement up to date for the narrow B^{**} states and the first measurement of the B_2^* width. However, the mass measurement is still limited by the statistical uncertainty so further improvements from the CDF experiment are possible.

In the decays of $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ no significant signal was observed. Therefore, a Bayesian limit was set on the branching ratio of the narrow B_s^{**} states. Since the B_s^{**} production rates are not known the branching ratio was measured relative to the reference decay $B_s^{**} \to B^+K^-$. The limit was derived by integrating the posterior probability density function of the unbinned maximum likelihood fit to the Q value distribution. The limits at a credibility level of 95% are determined to be

$$\frac{\mathrm{BR}(B_{s1} \to B_s\pi\pi)}{\mathrm{BR}(B_{s1} \to B_s\pi\pi) + \mathrm{BR}(B_{s1} \to BK)} < 0.54,$$
$$\frac{\mathrm{BR}(B_{s2}^* \to B_s\pi\pi)}{\mathrm{BR}(B_{s2}^* \to B_s^*\pi\pi) + \mathrm{BR}(B_{s2}^* \to BK)} < 0.53.$$

The limit is relatively weak as several assumptions being put into the fit to the B_s^{**} data samples have quite large uncertainties.

In summary, the analyses of the B^{**} and B_s^{**} mesons have yielded a new piece of puzzle to the understanding of quantum chromodynamics. Adding more data to the samples in future will decrease the statistical uncertainties. The appearance of a B_s^{**} signal in decays of $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ is still possible. With a sufficiently improved resolution of the experiment it would be also possible to distinguish the three B^{**} signal peaks of the decays $B_1 \to B^*\pi$, $B_2^* \to B^*\pi$ and $B_2^* \to B\pi$.

In this analysis a search for the narrow B^{**} and B_s^{**} states was performed. In order to complete the picture about orbitally excited B and B_s mesons, observations about the broad states are essential. Those states are of course hard to distinguish from the combinatorial background even with the aid of the currently available selection methods.

A Definitions

A.1 Particles

The symbols used for particles in chapter 5 are usually unambiguous within a single decay channel. If not, the symbol of the parent particle is put to the variable as an index. For instance, the $\pi_{B^{**}}$ is the pion from the decay $B^{**} \to B^{(*)+}\pi^-$.

Additionally, the three pions from the decays $B^+ \to \bar{D}^0\,3\pi^\pm$ and $B_s \to D_s^- 3\pi^\pm$ have superscript numbers, for example $\pi^{(1)}_{B^+}$, $\pi^{(2)}_{B^+}$ and $\pi^{(3)}_{B^+}$. These pions are sorted in a way that pion (1) and (2) have the same charge whereas the charge of pion (3) is different. Moreover, pion (1) has a higher transverse momentum than pion (2).

A.2 Variables

For the following definitions of the variable names used in chapter 5, the placeholders α, β and γ are used for aribtrary particles.

$p_t(\alpha)$	The transverse momentum of the particle α.
$\max_{B_s^{**}}(p_t)$	The maximum of the transverse momenta of the pions from the decay $B_s^{**} \to B_s\pi^+\pi^-$.
$\min_{B_s^{**}}(p_t)$	The minimum of the transverse momenta of the pions from the decay $B_s^{**} \to B_s\pi^+\pi^-$.
$p_t(\alpha; <)$	The transverse momentum of the α in the case it is smaller than 2.0 GeV/c.
$p_t(\alpha; >)$	The transverse momentum of the α in the case it is higher than 2.0 GeV/c.
$\lvert d_0(\alpha)\rvert$	The absolute value of the impact parameter of the α.
$d_0/\sigma_{d_0}(\alpha)$	The significance of the impact parameter of the α.
$d_0^{lts}/\sigma_{d_0}(\alpha)$	The lifetime signed impact parameter of the α.

$\max_{B_s^{**}}(\lvert d_0 \rvert)$	The maximum of the absolute values of the impact parameters of the pions from the decay $B_s^{**} \to B_s \pi^+ \pi^-$.
$\min_{B_s^{**}}(\lvert d_0 \rvert)$	The minimum of the absolute values of the impact parameters of the pions from the decay $B_s^{**} \to B_s \pi^+ \pi^-$.
$\max_{B^+}(\lvert d_0 \rvert)$	The maximum of the absolute values of the impact parameters of the pions from the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$.
$\min_{B^+}(\lvert d_0 \rvert)$	The minimum of the absolute values of the impact parameters of the pions from the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$.
$L_{xy}(\alpha)$	The transverse decay length of the α with respect to the primary vertex. The transverse decay length is the projection of the decay length to the x-y plane which is perpendicular to the beam axis.
$L_{xy}/\sigma_{L_{xy}}(\alpha)$	The significance of the transverse decay length of the α with respect to the primary vertex.
$\ell_{xy}(\alpha)$	The transverse decay length of the α with respect to the decay vertex of the parent particle.
$\ell_{xy}/\sigma_{\ell_{xy}}(\alpha)$	The significance of the transverse decay length of the α with respect to the decay vertex of the parent particle.
$\eta(\alpha)$	The pseudorapidity of the α.
$\Delta R(\alpha, \beta)$	$\Delta R(\alpha, \beta) = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}$, where η is the pseudorapidity and φ the azimuthal angle.
$\theta^*(\alpha)$	The cosine of the angle between the momentum of the α in rest frame of the the parent particle and the parent particle's momentum in the laboratory frame.
$\theta_{\text{hel}}(\alpha)$	The cosine of the helicity angle of the α which is defined as the angle between the α momentum in the rest frame of the particle and the momentum of the grandparent particle.
θ_{a_1}	The angle between the momentum of the a_1 and the normal vector of the plane spanned by the momenta of the two pions from the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$ having the same charge.
$P(\alpha)$	The fit probability of the α kinematic fit with appropriate topology constraints.

$\chi^2_{r\phi}(\alpha)$	The χ^2 of the two-dimensional vertex fit of the α.
$m(\alpha)$	The invariant mass of the α.
$\lvert m(\alpha) - m_{\mathrm{PDG}} \rvert$	The absolute value of the difference between the mass of the α and the world average value.
$m\left(\pi^{(1)}_{B^+}, \pi^{(2)}_{B^+}\right)$	The invariant mass of the two pions from the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$ having the same charge.
$m_{\mathrm{min}}\left(\pi^{(1,2)}_{B^+}, \pi^{(3)}_{B^+}\right)$	The invariant mass of the combination of two pions having different charges from the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$ yielding the smaller invariant mass
$m_{\mathrm{max}}\left(\pi^{(1,2)}_{B^+}, \pi^{(3)}_{B^+}\right)$	The invariant mass of the combination of two pions having different charges from the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$ yielding the higher invariant mass
$m_{\pi^+\pi^+\pi^-}$	The invariant mass of the three pions from the decay $B^+ \to \bar{D}^0\, 3\pi^\pm$ divided by the difference between the masses of the B^+ and the D.
Q	The Q value. For B^{**} decays it is defined as $Q = m(B^{**}) - m(B^+) - m(\pi)$ and for B_s^{**} decays it is $Q = m(B_s^{**}) - m(B_s) - 2\,m(\pi)$.
$L_\pi(\alpha; \mathrm{TOF})$	The PID likelihood ratio of the particle α under pion hypothesis in the case there is time of flight information. The PID is defined as $$\mathrm{PID} = \frac{P_{dE/dx}(h_i) \cdot P_{TOF}(h_i)}{\sum_{j \neq i} f_j \cdot P_{dE/dx(h_j)} \cdot P_{TOF}(h_j)} \tag{A.1}$$ where h_i is a given particle hypothesis, in the pion case $i = \pi$. $P_{dE/dx}$ and P_{TOF} are the probabilities from the energy loss measurement in the COT and the time of flight measurement, respectively. The factor f_j is 0.7 for pions, 0.2 for kaons and 0.1 for protons. The denominator sums over the remaining particle species.
$L_\pi(\alpha; \overline{\mathrm{TOF}})$	The PID likelihood ratio of the particle α under pion hypothesis in the case there is no time of flight information.
$L_K(\alpha; \mathrm{TOF})$	The PID likelihood ratio of the particle α under kaon hypothesis in the case there is time of flight information.

$L_K(\alpha;\overline{\mathrm{TOF}})$	The PID likelihood ratio of the particle α under kaon hypothesis in the case there is no time of flight information.
$n_{\mathrm{NN}}(\alpha)$	The output of the neural network used to select the α.
$n_{\mathrm{NN}}(\alpha \to \beta, \beta \to \gamma)$	The output of the neural network used to select the α in the decay of $\alpha \to \beta$ with $\beta \to \gamma$.
n_{Dcy}	The ordinal number of the B_s decay channel. Possible values are:

Value	Decay channel	
1	$B_s \to D_s^- \pi^+$	$D_s^- \to \bar{K}^* K^-$
2		$D_s^- \to \phi\pi^-$
3		$D_s^- \to 3\pi^\pm$
4	$B_s \to D_s^- 3\pi^\pm$	$D_s^- \to \bar{K}^* K^-$
5		$D_s^- \to \phi\pi^-$
6		$D_s^- \to 3\pi^\pm$

B Monte Carlo Residuals

The residual Q value distribution is used to determine the Q value resolution of the B_s^{**} signal, as described in section 7.2. Figures B.1 through B.6 show the residuals of the $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ Monte Carlo samples.

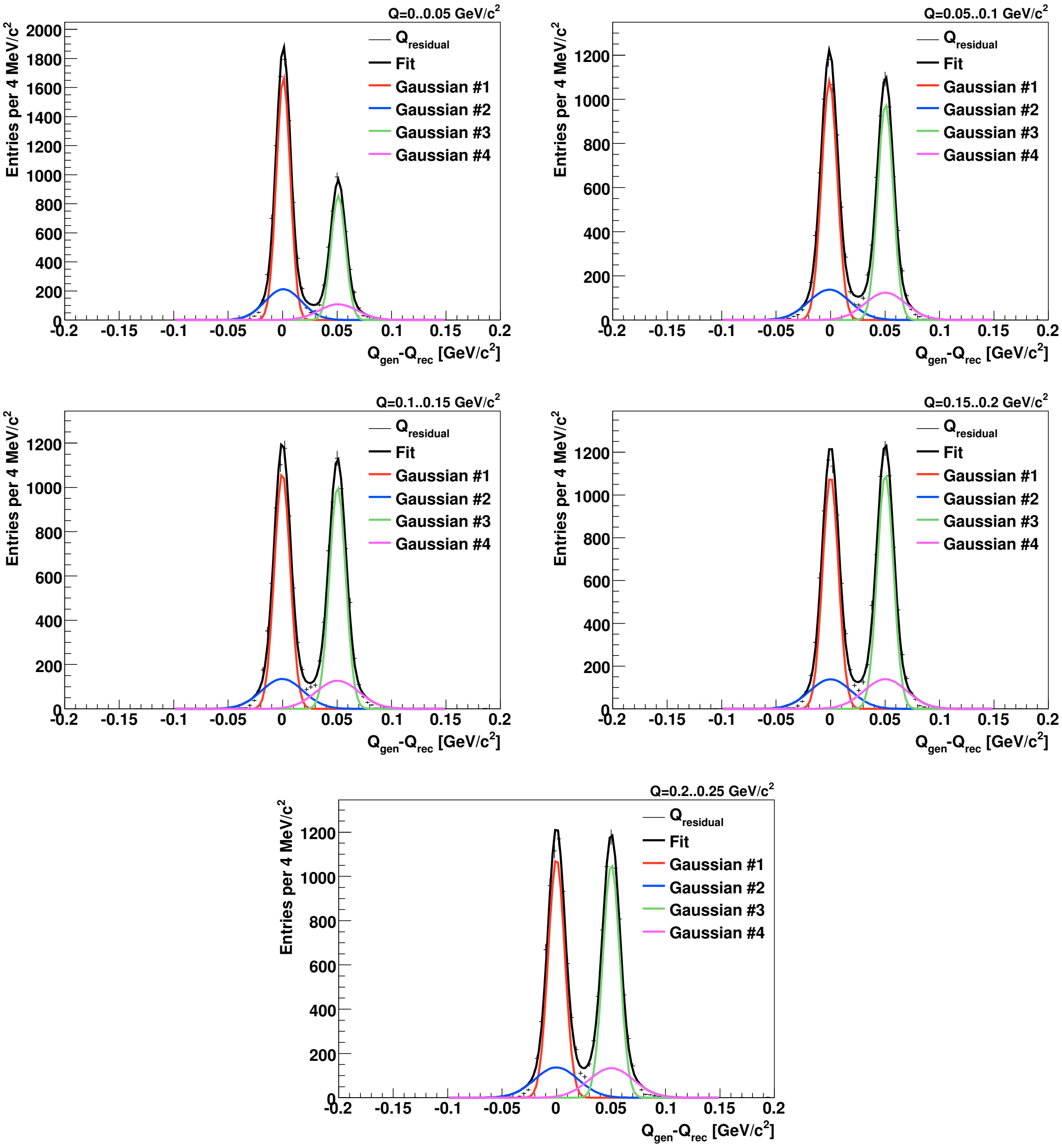


Figure B.1: Residual Q value distribution in the Monte Carlo sample of the decay $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ with $B_s \rightarrow D_s^-\pi^+$, $D_s^- \rightarrow \bar{K}^*K^-$ for different Q value ranges. Having two peaks arises from the fact that the photon from the decay $B_s^* \rightarrow B_s\gamma$ is not detected by the CDF II detector.

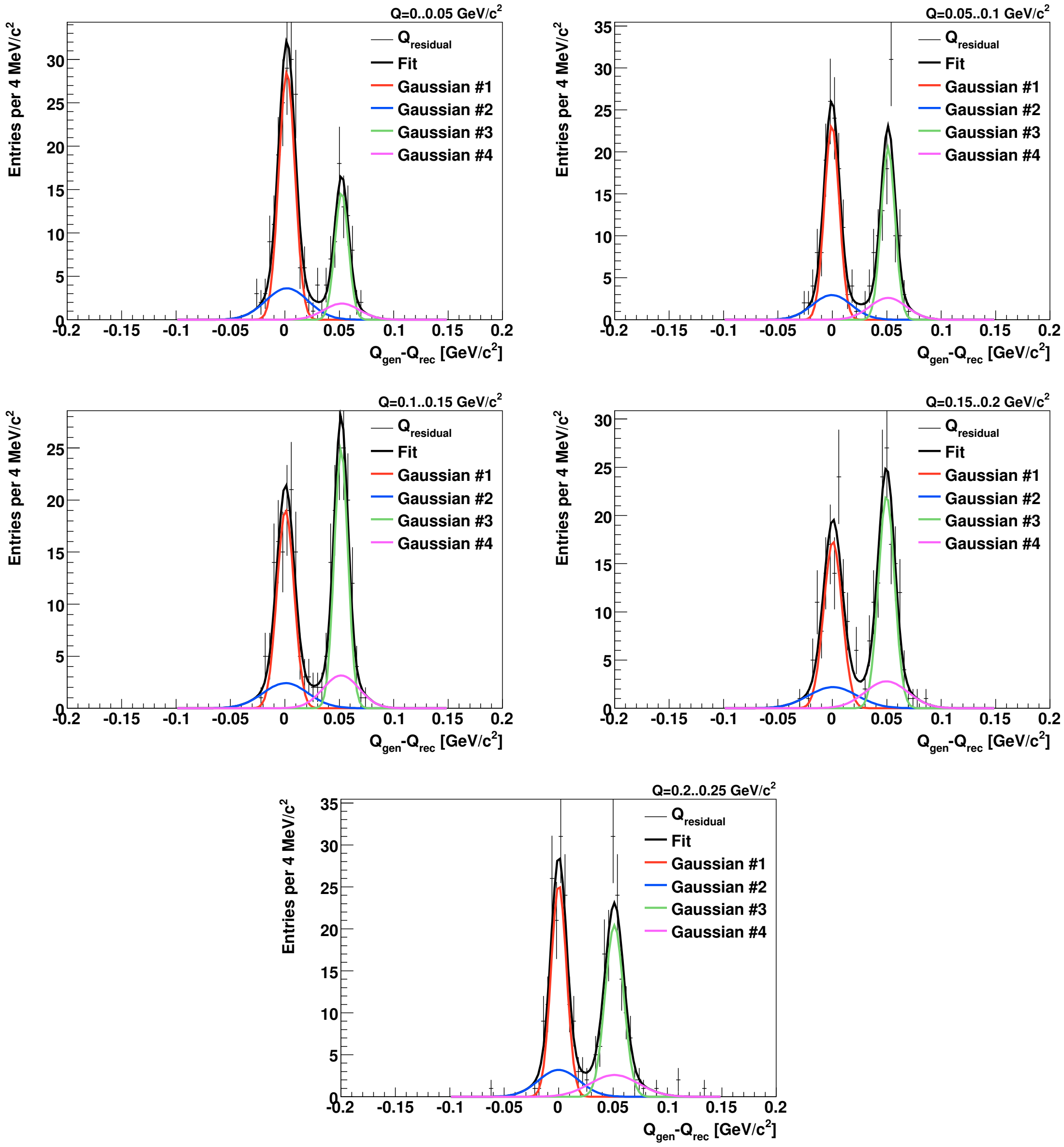


Figure B.2: Residual Q value distribution in the Monte Carlo sample of the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ with $B_s \to D_s^-\pi^+$, $D_s^- \to \phi\pi^-$ for different Q value ranges. Having two peaks arises from the fact that the photon from the decay $B_s^* \to B_s\gamma$ is not detected by the CDF II detector.

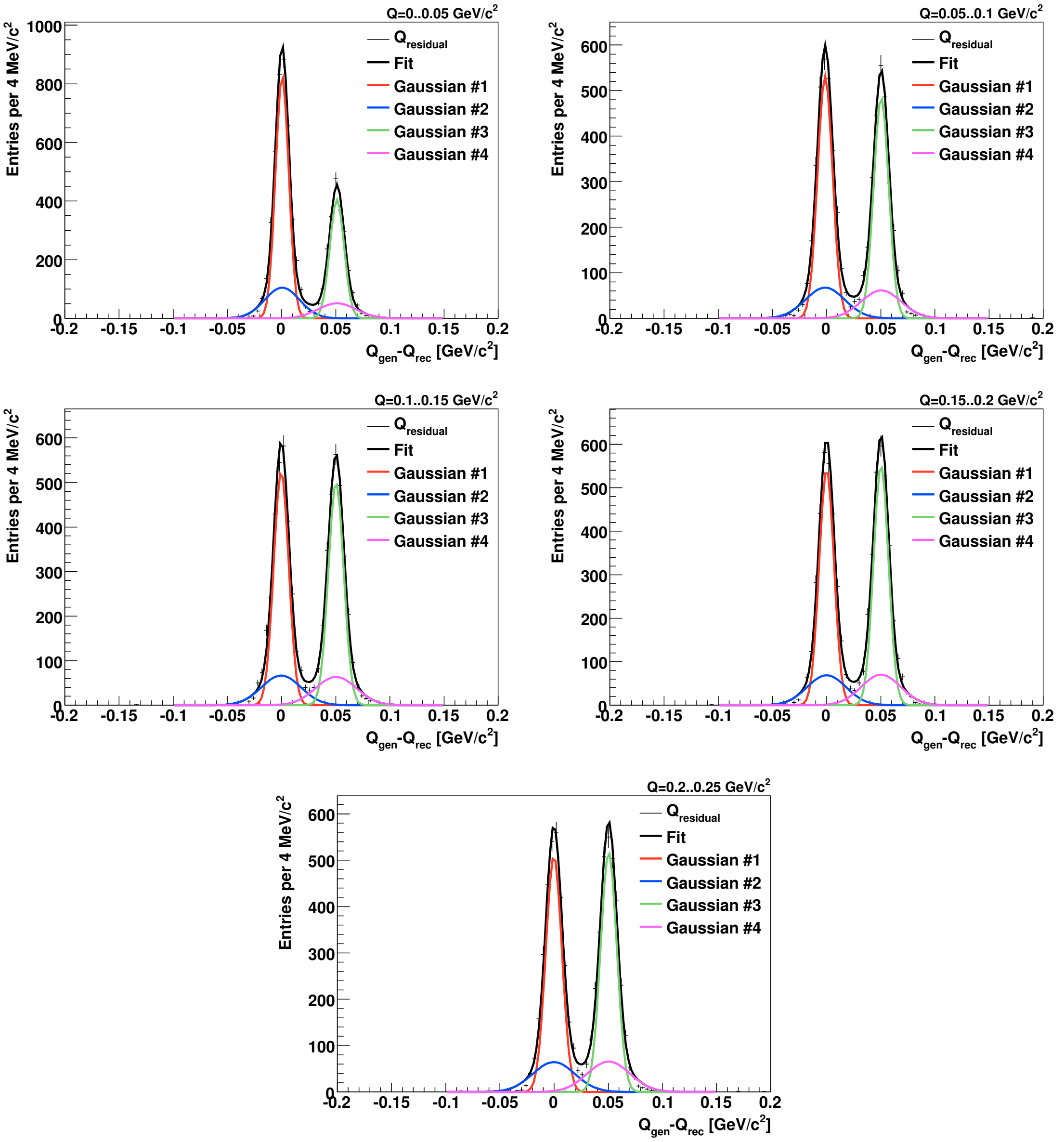


Figure B.3: Residual Q value distribution in the Monte Carlo sample of the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ with $B_s \to D_s^-\pi^+$, $D_s^- \to 3\pi^{\pm}$ for different Q value ranges. Having two peaks arises from the fact that the photon from the decay $B_s^* \to B_s\gamma$ is not detected by the CDF II detector.

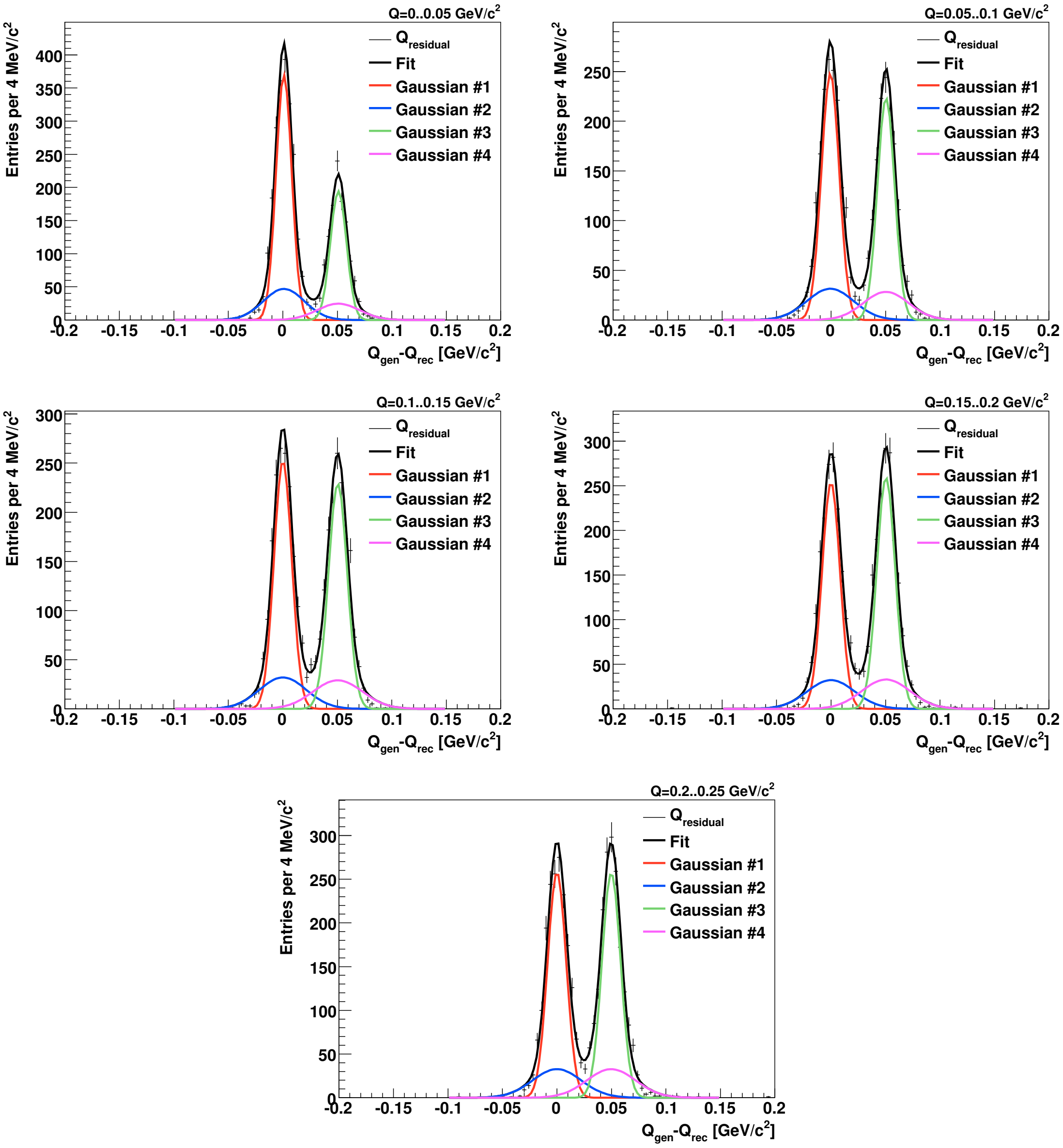


Figure B.4: Residual Q value distribution in the Monte Carlo sample of the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ with $B_s \to D_s^- 3\pi^\pm$, $D_s^- \to \bar{K}^* K^-$ for different Q value ranges. Having two peaks arises from the fact that the photon from the decay $B_s^* \to B_s\gamma$ is not detected by the CDF II detector.

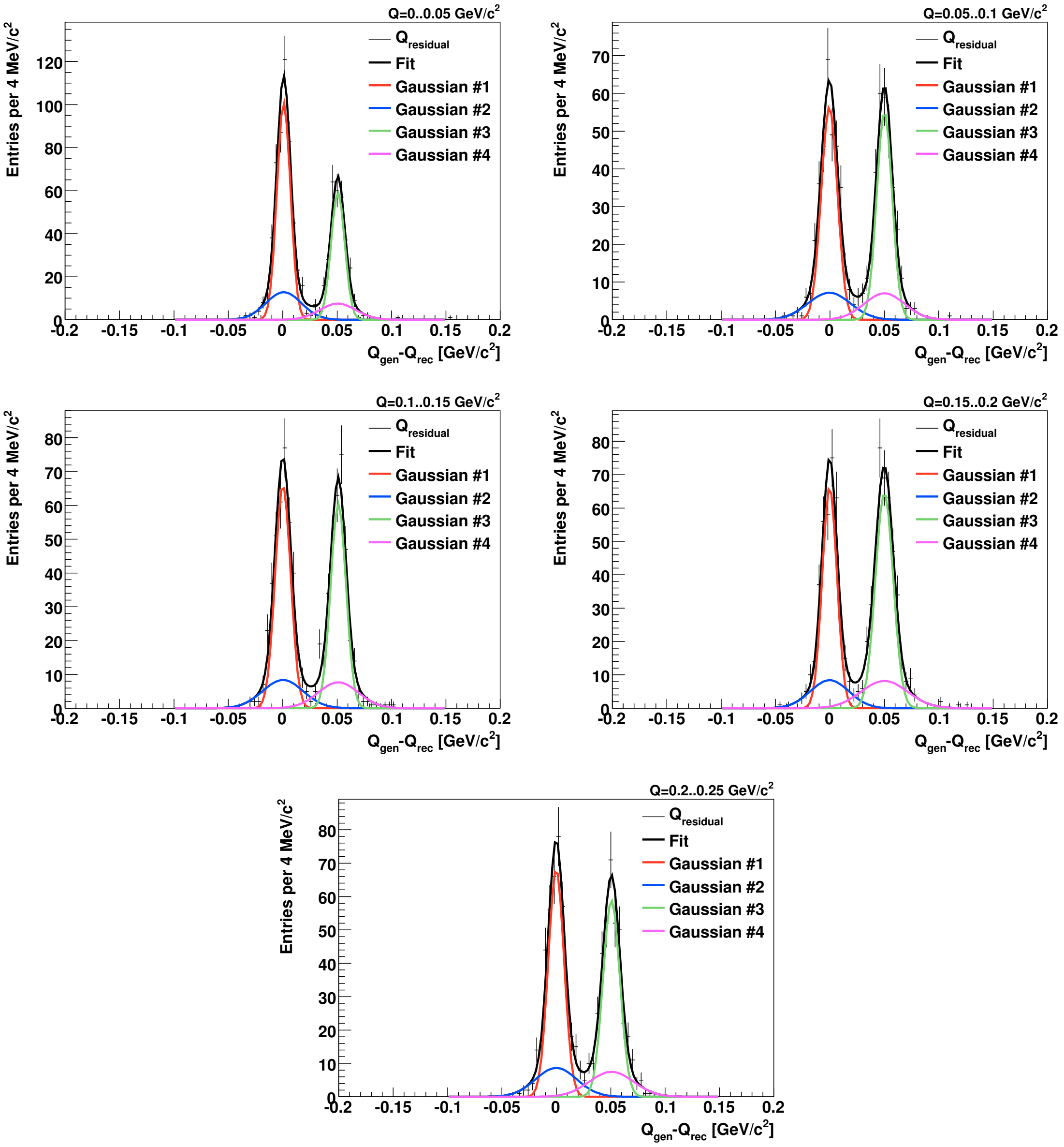


Figure B.5: Residual Q value distribution in the Monte Carlo sample of the decay $B_s^{**} \to B_s^{(*)}\pi^+\pi^-$ with $B_s \to D_s^- 3\pi^\pm$, $D_s^- \to \phi\pi^-$ for different Q value ranges. Having two peaks arises from the fact that the photon from the decay $B_s^* \to B_s\gamma$ is not detected by the CDF II detector.

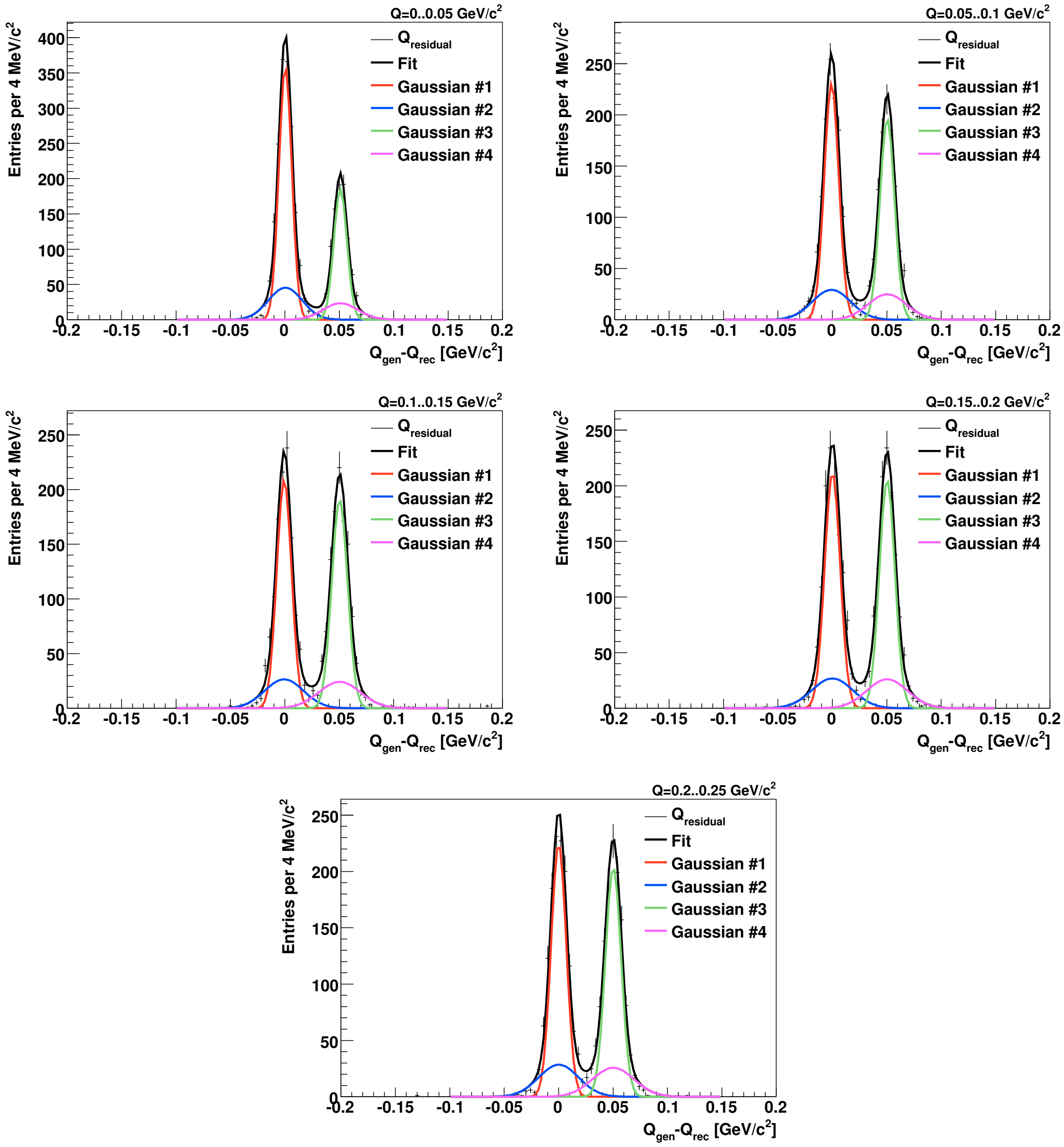


Figure B.6: Residual Q value distribution in the Monte Carlo sample of the decay $B_s^{**} \rightarrow B_s^{(*)}\pi^+\pi^-$ with $B_s \rightarrow D_s^- 3\pi^\pm$, $D_s^- \rightarrow 3\pi^\pm$ for different Q value ranges. Having two peaks arises from the fact that the photon from the decay $B_s^* \rightarrow B_s\gamma$ is not detected by the CDF II detector.

C Monte Carlo Decay Tables

C.1 Table for B^{**} Decays with $B^+ \rightarrow J/\psi K^+$

```
Alias    myB+     B+
Alias    myB-     B-
Alias    myB*+    B*+
Alias    myB*-    B*-
Alias    myJ/psi  J/psi


Decay    B_2*0
.2       myB+     pi-      PHSP;
.2       myB*+    pi-      PHSP;
Enddecay

Decay    anti-B_2*0
.2       myB-     pi+      PHSP;
.2       myB*-    pi+      PHSP;
Enddecay
#
Decay    myB*+
.2       myB+     gamma    PHSP;
Enddecay
#
Decay    myB*-
.2       myB-     gamma    PHSP;
Enddecay
#
#
Decay    myB+
.11      myJ/psi  K+       SVS;
Enddecay
#
Decay    myB-
.11      myJ/psi  K-       SVS;
Enddecay
#
Decay    myJ/psi
.11      mu+      mu-      PHOTOS  VLL;
Enddecay
#
End
```

C.2 Table for B^{**} Decays with $B^+ \rightarrow \bar{D}^0\,(3)\pi^{\pm}$

```
Alias    myB+     B+
Alias    myB-     B-
Alias    myB*+    B*+
```

```
Alias     myB*-     B*-

Decay     B_2*0
.2        myB*+     pi-       PHSP;
.2        myB+      pi-       PHSP;
Enddecay
#
Decay     anti-B_2*0
.2        myB*-     pi+       PHSP;
.2        myB-      pi+       PHSP;
Enddecay
#
#
Decay     myB*+
.2        myB+      gamma     PHSP;
Enddecay
#
Decay     myB*-
.2        myB-      gamma     PHSP;
Enddecay
#
#
Decay     myB+
0.0100    anti-D0   pi+       PHSP;
0.0090    a_1+      anti-D0   SVS;
0.0005    anti-D0   rho0      pi+       PHSP;
0.0005    anti-D0   pi-       pi+       pi+       PHSP;
Enddecay
#
Decay     myB-
0.0100    D0        pi-       PHSP;
0.0090    a_1-      D0        SVS;
0.0005    D0        rho0      pi-       PHSP;
0.0005    D0        pi+       pi-       pi-       PHSP;
Enddecay
#
#
Decay     D0
0.0383    K-        pi+       PHSP;
Enddecay
#
Decay     anti-D0
0.0383    K+        pi-       PHSP;
Enddecay
#
Decay     a_1+
0.4910    rho0      pi+       VVS_PWAVE           0.9091    0.0       0.0       0.0       -0.0909
0.0;
Enddecay
#
Decay     a_1-
0.4910    rho0      pi-       VVS_PWAVE           0.9091    0.0       0.0       0.0       -0.0909
0.0;
Enddecay
#
Decay     rho0
1.000     pi+       pi-       VSS;
Enddecay
#
End
```

C.3 Table for B_s^{**} Decays with $B_s \to D_s^- \pi^+$

```
Decay     B_s2*0
.2        B_s*0     pi+       pi-       PHSP;
.2        B_s0      pi+       pi-       PHSP;
Enddecay
#
Decay     anti-B_s2*0
.2        anti-B_s*0          pi-       pi+       PHSP;
.2        anti-B_s0           pi-       pi+       PHSP;
Enddecay
#
#
Decay     B_s*0
.2        B_s0      gamma     PHSP;
Enddecay
#
Decay     anti-B_s*0
.2        anti-B_s0           gamma     PHSP;
Enddecay
#
#
Decay     B_s0
0.0100    D_s-      pi+       PHSP;
0.0005    J/psi     phi       PHSP;
Enddecay
#
Decay     anti-B_s0
0.0100    D_s+      pi-       PHSP;
0.0005    J/psi     phi       PHSP;
Enddecay
#
Decay     D_s+
0.0247    phi       pi+       SVS;
0.0247    anti-K*0            K+        SVS;
0.0004    rho0      pi+       SVS;
0.0180    f_0       pi+       PHSP;
0.0023    f_2       pi+       PHSP;
0.0040    pi+       pi-       pi+       PHSP;
Enddecay
#
Decay     D_s-
0.0247    phi       pi-       SVS;
0.0247    K*0       K-        SVS;
0.0004    rho0      pi-       SVS;
0.0180    f_0       pi-       PHSP;
0.0023    f_2       pi-       PHSP;
0.0040    pi-       pi-       pi+       PHSP;
Enddecay
#
Decay     f_0
0.5200    pi+       pi-       PHSP;
Enddecay
#
Decay     f_2
0.5650    pi+       pi-       TSS;
Enddecay
#
Decay     phi
0.4910    K+        K-        VSS;
```

```
Enddecay
#
Decay   a_1+
0.6     rho0    pi+     VVS_PWAVE    1.0    0.0    0.0    0.0    -0.1    0.0;
Enddecay
#
Decay   a_1-
0.6     rho0    pi-     VVS_PWAVE    1.0    0.0    0.0    0.0    -0.1    0.0;
Enddecay
#
Decay   rho0
1.000   pi+     pi-     VSS;
Enddecay
#
Decay   K*0
0.6657  K+      pi-     VSS;
Enddecay
#
Decay   anti-K*0
0.6657  K-      pi+     VSS;
Enddecay
#
Decay   J/psi
1       mu+     mu-     PHSP;
Enddecay
#
Decay   phi
1       K+      K-      PHSP;
Enddecay
#
#
End
```

C.4 Table for B_s^{**} Decays with $B_s \to D_s^- 3\pi^\pm$

```
Decay   B_s2*0
.2      B_s*0   pi+     pi-     PHSP;
.2      B_s0    pi+     pi-     PHSP;
Enddecay
#
Decay   anti-B_s2*0
.2      anti-B_s*0      pi-     pi+     PHSP;
.2      anti-B_s0       pi-     pi+     PHSP;
Enddecay
#
#
Decay   B_s*0
.2      B_s0    gamma   PHSP;
Enddecay
#
Decay   anti-B_s*0
.2      anti-B_s0       gamma   PHSP;
Enddecay
#
#
Decay   B_s0
0.0090  a_1+    D_s-    SVS;
0.0005  D_s-    rho0    pi+     PHSP;
0.0005  D_s-    pi-     pi+     pi+     PHSP;
Enddecay
```

```
#
Decay    anti-B_s0
0.0090   a_1-     D_s+     SVS;
0.0005   D_s+     rho0     pi-      PHSP;
0.0005   D_s+     pi+      pi-      pi-      PHSP;
Enddecay
#
Decay    D_s+
0.0247   phi      pi+      SVS;
0.0247   anti-K*0          K+       SVS;
0.0004   rho0     pi+      SVS;
0.0180   f_0      pi+      PHSP;
0.0023   f_2      pi+      PHSP;
0.0040   pi+      pi-      pi+      PHSP;
Enddecay
#
Decay    D_s-
0.0247   phi      pi-      SVS;
0.0247   K*0      K-       SVS;
0.0004   rho0     pi-      SVS;
0.0180   f_0      pi-      PHSP;
0.0023   f_2      pi-      PHSP;
0.0040   pi-      pi-      pi+      PHSP;
Enddecay
#
Decay    f_0
0.5200   pi+      pi-      PHSP;
Enddecay
#
Decay    f_2
0.5650   pi+      pi-      TSS;
Enddecay
#
Decay    phi
0.4910   K+       K-       VSS;
Enddecay
#
Decay    a_1+
0.6      rho0     pi+      VVS_PWAVE         1.0      0.0      0.0      0.0      -0.1     0.0;
Enddecay
#
Decay    a_1-
0.6      rho0     pi-      VVS_PWAVE         1.0      0.0      0.0      0.0      -0.1     0.0;
Enddecay
#
#
Decay    rho0
1.000    pi+      pi-      VSS;
Enddecay
#
Decay    K*0
0.6657   K+       pi-      VSS;
Enddecay
#
Decay    anti-K*0
0.6657   K-       pi+      VSS;
Enddecay
#
End
```

Bibliography

[1] CDF Collaboration, “Observation of Orbitally Excited B_s Mesons”, Phys. Rev. Lett. **100**, 082001, 2008

[2] C. Amsler *et al.*, “Review of Particle Physics”, Physics Letters B **667**, 1, 2008

[3] B. Povh, K. Rith, C. Scholz, F. Zetsche, “Teilchen und Kerne”, Springer-Verlag, 2004

[4] N. Cabibbo, “Unitary symmetry and leptonic decays”, Phys. Rev. Lett. **10**, 531–532, 1963.

[5] Private communication with M. Kobayashi at “The 2007 Europhysics Conference on High Energy Physics” in Manchester, England

[6] M. Kobayashi, T. Maskawa, “CP violation in the renormalizable theory of weak interaction”, Prog. Theor. Phys. **49**, 652–657, 1973

[7] S. L. Glashow, “Partial symmetries of weak interactions”, Nucl. Phys. **22**, 579–588, 1961

[8] A. Salam, J. C. Ward, “Electromagnetic and weak interactions”, Phys. Lett. **13**, 168–171, 1964

[9] S. Weinberg, “A model of leptons”, Phys. Rev. Lett. **19**, 1264–1266, 1967

[10] P. W. Higgs, “Broken symmetries and the masses of gauge bosons”, Phys. Rev. Lett. **13**, 508–509, 1964

[11] P. W. Higgs, “Spontaneous symmetry breakdown without massless bosons”, Phys. Rev. **145**, 1156–1163, 1966

[12] “B Decays”, Revised 2nd Edition, World Scientific Publishing, 1994

[13] D. Vucinic, “Observation of Excited B Mesons in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV”, PhD thesis, Massachusetts Institute of Technology, CDF/THESIS/BOTTOM/PUBLIC/4817

[14] A. F. Falk, T. Mehen, “Excited heavy mesons beyond leading order in the heavy quark expansion”, Phys. Rev. D **53**, 231, 1996, [arXiv:hep-ph/9507311]

[15] E. J. Eichten, C. T. Hill, C. Quigg, “Properties of Orbitally Excited Heavy-Light Mesons”, Phys. Rev. Lett. **71**, 4116, 1993, [arXiv:hep-ph/9308337v1]

[16] E. J. Eichten, C. T. Hill, C. Quigg, “Orbitally excited heavy-light mesons revisited”, FERMILAB-CONF-94/118-T, 1994

[17] N. Isgur, “Spin-orbit inversion of excited heavy quark mesons”, Phys. Rev. D **57**, 4041, 1998

[18] D. Ebert, V. O. Galkin, R. N. Faustov, “Mass spectrum of orbitally and radially excited heavy-light mesons in the relativistic quark model”, Phys. Rev. D **57**, 5663, 1998, [Erratum-ibid. D **59**, 019902, 1999], [arXiv:hep-ph/9712318]

[19] A. H. Orsland, H. Hogaasen, “Strong and electromagnetic decays of excited heavy mesons”, Eur. Phys. J. C **9**, 503, 1999

[20] T. Matsuki, T. Morii, K. Sudoh, “New Heavy-Light Mesons $Q\bar{q}$”, Prog. Theor. Phys. **117**, 1077, 2007

[21] Zhi-Gang Luo, Xiao-Lin Chen, Xiang Liu, “$B_{s1}(5830)$ and $B_{s2}^*(5840)$”, arXiv:0901.0505v3

[22] ALEPH Collaboration, “Resonant Structure and Flavor-tagging in the $B\pi^\pm$ System Using Fully Reconstructed B Decays”, Contribution to the International Europhysics Conference on High Energy Physics, Brussels, Belgium, 1995

[23] DELPHI Collaboration, “Observation of Orbitally Excited B Mesons,” CERN-PPE/94-210, 1994

[24] DELPHI Collaboration, “Observation of Orbitally Excited B and B_s Mesons”, DELPHI 95-105 PHYS 540, 1995

[25] DØ Collaboration, “Study of Excited B-mesons”, DØ note 4517-CONF, 2004

[26] DØ Collaboration, “Study of Excited B-mesons”, DØ note 5026-CONF, 2006

[27] F. Bedeschi *et al.*, “Observation of orbitally excited ($L = 1$) B mesons”, CDF note 7820, 2005

[28] OPAL Collaboration, R. Akers *et al.*, Z. Phys. C **66**, 19, 1995

[29] DELPHI Collaboration, Z. Albrecht *et al.*, DELPHI 2004-025 CONF 700

[30] Z. Albrecht, “Analysis of Excited B-Mesons”, PhD-Thesis, IEKP-KA/03-16, Universität Karlsruhe, 2003

[31] M. Moch, “Study of B^*-Mesons and of Excited b-Hadron Properties”, PhD-Thesis, IEKP-KA/2004-16, Universität Karlsruhe, 2004

[32] DØ Collaboration, “First Direct Observation of B_{s2}^{*0} meson”, DØ note 5027-Conf

[33] DØ Collaboration, V.M. Abazov *et al.*, “Observation and Properties of the Orbitally Excited B_{s2}^* Meson”, Phys. Rev. Lett. **100**, 082002, 2008

[34] FNAL Accelerator Division, “Fermilab’s Chain of Accelerators”, `http://www-bd.fnal.gov/public/chain.html`.

[35] “Run II Handbook”, `http://www-bdnew.fnal.gov/pbar/run2b/Documents/RunII_handbook.pdf`

[36] Fermilab Accelerator Devision, “Operations Rookie Books”, `http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html`

[37] Fermilab Accelerator Devision, “LINAC Rookie Book”, `http://www-bdnew.fnal.gov/operations/rookie_books/LINAC_v2.pdf`

[38] Fermilab Accelerator Devision, “Booster Rookie Book”, `http://www-bdnew.fnal.gov/operations/rookie_books/Booster_V3_1.pdf`

[39] D. Mohl, G. Petrucci, L. Thorndahl, S. Van Der Meer, “Physics and Technique of Stochastic Cooling”, Phys. Rept. **58**, 73, 1980

[40] S. Van Der Meer, “Stochastic Cooling and the Accumulation of Anti-Protons”, Rev. Mod. Phys. **57**, 689, 1985

[41] S. Nagaitsev *et al.*, “Experimental Demonstration of Relativistic Electron Cooling”, Phys. Rev. Lett. **96**, 044801, 2006

[42] S. Nagaitsev *et al.*, “Status of the Fermilab electron cooling project”, Nucl. Instrum. Meth. A **532**, 275, 2004

[43] Fermilab Accelerator Devision, “Recycler Rookie Book”, `http://www-bdnew.fnal.gov/operations/rookie_books/Recycler_RB_v1.2.pdf`

[44] Fermilab Accelerator Devision, “Main Injector Rookie Book”, `http://www-bdnew.fnal.gov/operations/rookie_books/Main_Injector_v1.pdf`

[45] D. Acosta *et al.*, Nucl. Instrum. Meth. A **461**, 540, 2001

[46] CDF Collaboration, “The CDF II Detector Technical Design Report”, FERMILAB-Pub-96/390-E, 1996

[47] C. S. Hill *et al.*, “Operational experience and performance of the CDFII silicon detector”, Nucl. Instrum. Meth. A **530**, 1, 2004

[48] A. Sill, “CDF Run II silicon tracking projects”, Nucl. Instrum. Meth. A **447**, 1, 2000

[49] A. A. Affolder *et al.*, “Intermediate silicon layers detector for the CDF experiment”, Nucl. Instrum. Meth. A **453**, 84, 2000

[50] A. A. Affolder *et al.*, “CDF Central Outer Tracker”, Nucl. Instrum. Meth. A **526**, 249, 2004

[51] D. Acosta *et al.*, “A time-of-flight detector in CDF-II”, Nucl. Instrum. Meth. A **518**, 605, 2004

[52] S. Kuhlmann *et al.*, “The CDF calorimeter upgrade for Run IIb”, Nucl. Instrum. Meth. A **518**, 39, 2004

[53] S. Bailey, P. Maksimovic, “Prospects for Measuring γ with $B_s o D_s^{\mp} K^{\pm}$”, CDF/ANAL/BOTTOM/CDFR/4863

[54] L. Balka *et al.*, “The CDF Central Electromagnetic Calorimeter”, Nucl. Instrum. Meth. A **267**, 272, 1988

[55] S. Bertolucci *et al.*, “The CDF Central and Endwall Hadron Calorimeter”, Nucl. Instrum. Meth. A **267**, 301, 1988

[56] G. Ascoli *et al.*, “CDF Central Muon Detector”, Nucl. Instrum. Meth. A **268**, 33, 1988

[57] A. Artikov *et al.*, “Design and construction of new central and forward muon counters for CDF II”, Nucl. Instrum. Meth. A **538**, 358, 2005

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

[59] H. Ray, “Level 2 and the L1/L2 Trigger System for Non-Experts”, CDF/DOC/TRIGGER/PUBLIC/5860, 2002

[60] E. J. Thomson *et al.*, “Online track processor for the CDF upgrade”, IEEE Trans. Nucl. Sci., 49:1063-1070, 2002

[61] D. Lucchesi, "CDF – Secondary Vertex Trigger", Proceedings Beauty 2002 and CDF/PUB/TRIGGER/PUBLIC/6107, 2002

[62] V. Blobel, E. Lohrmann, "Statistische und numerische Methoden der Datenanalyse", B. G. Teubner Stuttgart Leibzig, 1998

[63] M. Feindt, "A Neural Bayesian Estimator for Conditional Probability Densities", physics/0402093, 2004, [arXiv:physics/0402093v1]

[64] Phi-T® Physics Information Technologies GmbH, `http://www.phi-t.de/`

[65] Ch. 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, Universität Karlsruhe, 2006

[66] P. Mack, "Calibration of New Flavor Tagging Algorithms using B_s Oscillations", PhD-Thesis, IEKP-KA/2007-10, Universität Karlsruhe, 2007

[67] A. Gessler, "Rekonstruktion der Zerfälle $B_s \to D_s\pi$ und $B_s \to D_s\pi\pi\pi$ mit dem CDF-II-Detektor", Diploma-Thesis, IEKP-KA/2006-12, Universität Karlsruhe, 2006

[68] CDF DQM Group, `http://www-cdf.fnal.gov/internal/dqm/goodrun/good.html`, Good run list

[69] C. Paus *et al.*, see CDF Code Browser `http://cdfkits.fnal.gov/CdfCode/source/BottomMods/`

[70] P. Murat, K. Anikeev, C. Paus, "Description of BGeneratorII", CDF Note 5092

[71] D. J. Lange, "The EvtGen particle decay simulation package", NIM A **462**, 152, 2001

[72] CDF Collaboration, "Study of orbitally excited ($L = 1$) B mesons", CDF note 8945, 2008

[73] Giovanni Punzi, "Sensitivity of searches for new signals and its optimization", arXiv:physics/0308063, 2003

[74] S. Godfrey, R. Kokoski, "The Properties of p Wave Mesons with One Heavy Quark", Phys. Rev. D **43**, 1679, 1991

[75] CDF Collaboration, "Measurement of Resonance Parameters of Orbitally Excited Narrow B^0 Mesons", Phys. Rev. Lett. **102**, 102003, 2009

[76] F. James, "Minuit, Function Minimization and Error Analysis", Cern Program Library Long Writeup D **506**, 1998

[77] R. Brun, F. Rademakers, "Root – An Object Oriented Data Analysis Framework", AIHENP'96 Workshop, Lausane, Nucl. Inst. & Meth. in Phys. Res. A **389**, 81-86, 1997 see `http://root.cern.ch/`

[78] CDF Collaboration, "Precision Measurement of the $X(3872)$ Mass in $J/\psi\pi^+\pi^-$ Decays", CDF note 9700, 2009

[79] A. Oyanguren (DELPHI Collaboration), Proceedings of 32nd International Conference on High-Energy Physics (ICHEP 04), Beijing, China, 16-22 Aug 2004. Hackensack, World Scientific, 2 vols, 2005

Danksagung

Keine Promotion könnte zustande kommen ohne die Unterstützung und den Einfluss von vielen Menschen, so dass ich an dieser Stelle nicht alle nennen kann.

Zunächst danke ich Herrn Professor Michael Feindt, als Doktorand in seiner Forschungsgruppe arbeiten zu dürfen, wodurch diese Arbeit überhaupt möglich wurde. Weiterer Dank gilt Herrn Professor Günter Quast für die Übernahme des Korreferates, Herrn Professor Dieter Zeppenfeld für die Betreuung als Mentor und Herrn Professor Thomas Müller, der mir als Ordinarius des Instituts für Experimentelle Kernphysik im Rahmen meiner Promotion zahlreiche Arbeitsaufenthalte am Fermilab sowie am CERN ermöglichte.

Für Korrekturen und Diskussionen rund um meine Arbeit danke ich vor allem Dr. Michal Kreps und Dr. Thomas Kuhr, deren tatkräftige Unterstützung und konstruktive Kritik die Dissertation in dieser Form gelingen ließen.

Außerdem danke ich Dr. Manuel Bä$\hbar$r, Daniel Zander, Sonja Sohns, Dominik Horn und Sebastian Neubauer, ohne deren emsiges Korrekturlesen nicht nur der eine oder andere Fehler unentdeckt geblieben wären.

Ermöglicht wurde die Dissertation durch ein Stipendium des *Graduiertenkollegs Hochenergiephysik und Teilchenphysik* der Deutschen Forschungsgemeinschaft. An dieser Stelle danke ich Michaela Gerstner, Isabelle Junge, Bärbel Bräunling und Waltraud Weißmann für den stetigen Kampf gegen die Bürokratie.

Maßgeblich zu dieser Arbeit beigetragen hat auch das fleißige Administratoren-Team des Instituts, dem ich hiermit ebenfalls danke. Dass im Arbeitsalltag nichts von dessen Wirken wahrgenommen wird, ist der beste Beleg für ihre hervorragende Arbeit.

Für die großartige Zeit zusammen in Raum 9-9 danke ich meinen Zimmerkollegen Dr. Thorsten Scheidle, Dr. Michael Milnik, Dominik Horn und Andreas Jäger.

Außerdem danke ich Herrn Makoto Kobayashi für die CKM-Matrix, Donald E. Knuth für TeX, Leslie Lamport für LaTeX und Linus Torvalds für Linux.