Search for the Higgs Boson with the CDF experiment at the Tevatron

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

DISSERTATION

von

Dipl.-Phys. Martin Hennecke
aus Arnsberg/Westfalen

Tag der mündlichen Prüfung: 10.06.2005
Referent: Prof. Dr. Wim de Boer, Institut für Experimentelle Kernphysik
Koreferent: Prof. Dr. Michael Feindt, Institut für Experimentelle Kernphysik

Abstract

A search for a low-mass SM Higgs-Boson in the channel $WH \rightarrow l\nu b\bar{b}$ has been performed using neural networks. The data were taken by the CDF experiment at the p-$\bar{p}$ collider Tevatron from 2000-2003, corresponding to in integrated luminosity of $\mathcal{L}_{int} = 162$ pb^{-1} at a CMS-energy of $\sqrt{s} = 1.96$ TeV. 95% confidence level upper limits are set on $\sigma \times BR$, the product of the production cross section times the Branching ratio, as a function of the Higgs boson mass. Cross sections above 8 pb are excluded for six different Higgs masses between 110 GeV/c^2 and 150 GeV/c^2. The requirred integrated luminosities for a 95% $C.L.$ exclusion, 3σ evidence and 5σ discovery are calculated.

Zusammenfassung

Eine Suche nach dem leichten SM Higgs-Boson wurde mit neuronalen Netzen durchgeführt. Die Daten wurden mit dem CDF-Experiment am $p\bar{p}$-Beschleuniger Tevatron von 2000-2003 aufgezeichnet, und entsprechen einer integrierten Luminosität von $\mathcal{L}_{int} = 162$ pb^{-1}, bei einer Schwerpunktsenergie von $\sqrt{s} = 1.96$ TeV. Bei einer Vertrauensgrenze von 95% werden obere Grenzen auf $\sigma \times BR$, dem Produkt von Produktions-Wirkungsquerschnitt und Verzweigungsverhältnis, als Funktion der Higgs-Masse gesetzt. Wirkungsquerschnitte oberhalb 8 pb werden für sechs verschiedene Higgs-Massen zwischen 110 GeV/c^2 and 150 GeV/c^2 ausgeschlossen. Die benötigten integrierten Luminositäten für einen 95% $C.L.$-Ausschluß, eine 3σ-Evidenz und eine 5σ-Entdeckung werden berechnet.

Contents

Deutsche Zusammenfassung

Einleitung

Die vorliegende Arbeit behandelt die Suche nach einem leichten Standard-Modell Higgs-Boson mit dem CDF-Experiment am Proton-Antiproton Beschleuniger Tevatron. Untersucht wurde die assoziierte Produktion in Verbindung mit einem W-Boson im Kanal $WH \to l\nu b\bar{b}$. Hierbei zerfällt das W-Boson leptonisch[1] und das Higgs in ein Paar schwerer b-Quarks. Die betrachteten sechs Higgs-Massen liegen im Bereich zwischen 110 GeV/c^2 und 150 GeV/c^2 in Schritten von 10 GeV/c^2 sowie bei 115 GeV/c^2. Die Analyse beruht auf Daten aus den Jahren 2000 bis 2003, entsprechend einer integrierten Luminosität von $\mathcal{L}_{int} = 162$ pb^{-1}. Die Schwerpunktsenergie des Tevatron betrug in dieser Zeit $\sqrt{s} = 1.96$ TeV.

Das Tevatron

Der $p\bar{p}$-Beschleuniger Tevatron befindet sich am Fermi National Accelerator Laboratory (FNAL), ca. 60 km westlich von Chicago. Bei einer Schwerpunktsenergie von 1.96 TeV ist er der zur Zeit höchstenergetische Collider der Welt. Der Umfang beträgt ca. 6 km. Zwei gegeneinander beschleunigte Teilchenstrahlen kollidieren an zwei Wechselwirkungszonen. An einer dieser Stellen befindet sich das CDF II-Experiment. Um Protonen und Antiprotonen bis auf eine Energie von etwa 1 TeV zu beschleunigen, ist ein System verschiedener Vorbeschleuniger nötig. Jeder erhöht die Teilchenenergie bis zur minimalen Einschußenergie des nächsten Systems. Abb. 1 auf Seite 9 zeigt links eine schematische Darstellung des Tevatron und seiner Vorbeschleuniger.

Der CDF II-Detektor

Das CDF[2]-Experiment ist ein sog. Multi-Purpose-Detektor. Verschiedene Detektor-Systeme sind um den Wechselwirkungspunkt in mehreren Lagen angeordnet. Abb. 1 zeigt rechts die Anordnung der unterschiedlichen Komponenten. Das Tracking-System befindet sich am nächsten zum Strahlrohr und besteht aus 3 Silizium-Detektoren. Diese werden

[1] W-Zerfälle in τ-Leptonen werden in dieser Arbeit, aufgrund des großen Untergrund, nicht berücksichtigt.

[2] Collider Detector at Fermilab

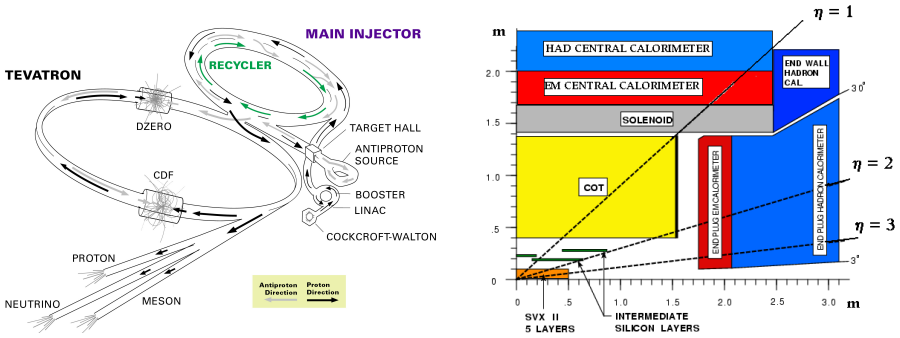


Abbildung 1: Schematische Darstellung des Tevatron mit seinen Vorbeschleunigern (links) und Längsschnitt durch den CDF II Detektor (rechts)[1].

von einer Driftkammer, der COT[3], umschlossen. Außerhalb der COT befindet sich eine supraleitende Solenoid-Spule, die ein magnetisches Feld von 1.4 T erzeugt. Ein Time-Of-Flight-System befindet sich zwischen Solenoid und der COT. Als nächstes schließt sich das Kalorimeter an, das einen hadronischen und elektromagnetischen Teil hat. Myonen-Kammern stellen die äußersten Detektoren dar. Vier unterschiedliche Subsysteme umschließen fast vollständig die inneren Komponenten.

Theoretische Grundlagen

Die wichtigsten Prozesse der Higgs-Erzeugung am Tevatron sind die Gluon-Gluon Fusion und die assoziierte Produktion des Higgs-Boson in Verbindung mit einem W- oder Z-Boson. Abbildung 2 auf Seite 10 zeigt links die NLO[4]-Wirkungsquerschnitte für alle drei Prozesse. Die Gluon-Gluon Fusion hat den größten Wirkungsquerschnitt, allerdings ist hier beim Zerfall $H \rightarrow b\bar{b}$ die Untergrundsituation ungünstig. Bei den WH- und ZH-Reaktionen bildet ein Quark-Antiquark-Paar ein Vektorboson, das anschließend ein Higgs-Teilchen abstrahlt. Das entsprechende Feynman-Diagramm ist in Abb. 3 dargestellt. Der Querschnitt für den WH-Prozess liegt etwa einen Faktor zwei über dem ZH-Prozess und wird im folgenden weiter betrachtet. Die Verzweigungsverhältnisse des Standardmodell-Higgs sind in Abb. 2 auf der rechten Seite dargestellt. Das größte Verzweigungsverhältnis bis zu Massen von ca. 140 GeV/c^2 hat der Zerfall in zwei b-Quarks, $H \rightarrow b\bar{b}$. Für die Suche nach leichten Higgs-Bosonen bietet sich daher der Prozess $WH \rightarrow l\nu b\bar{b}$ an.

[3]Central Outer Tracker
[4]Next-to-Leading Order

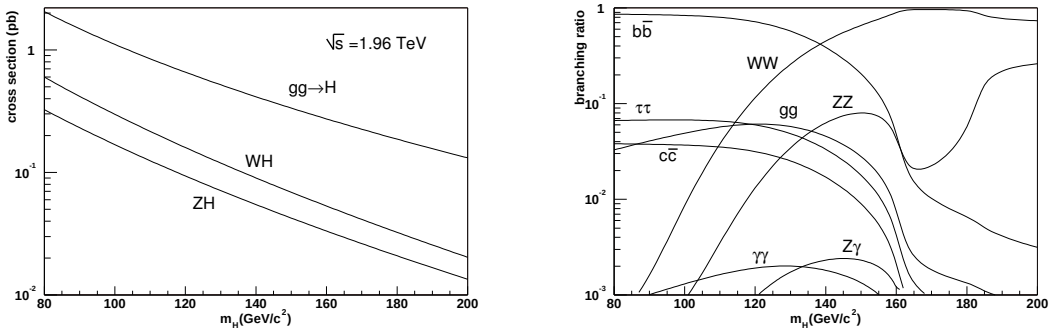


Abbildung 2: Produktions-Wirkungsquerschnitte für Gluon-Fusion und assoziierte Produktion bei 1.96 TeV (links) und Verzweigungsverhältnisse (rechts) des Standardmodel Higgs-Boson.

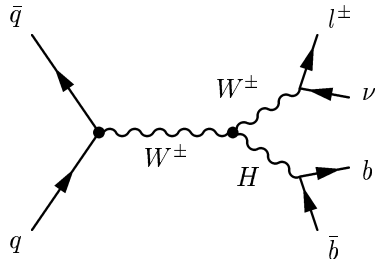


Abbildung 3: Feynman-Diagramm für den Signalprozess $WH \rightarrow l\nu b\bar{b}$.

Untergrundprozesse für diese Reaktion sind solche, in denen gleiche oder ähnliche Endzuständen auftreten, d.h. ein Lepton mit hohem Transversalimpuls, fehlende Energie durch Neutrinos und b-Quark-Jets. In dieser Analyse wurden folgende Untergründe berücksichtigt: Produktion von Top-Quarks (einzeln und paarweise), W + Jets, Paarproduktion von Eichbosonen (WW, WZ und WZ), $Z \rightarrow \tau\tau$ und QCD[5]-Untergrund. Exemplarische Feynman-Diagramme für diese Prozesse sind in den Abb. 3.14 und 3.15 auf den Seiten 58 und 59 angegeben.

Neuronale Netze

Neuronale Netze entstammen der Forschung über künstliche Intelligenz. Um die Lernfähigkeit und Fehlertoleranz biologischer Systeme auf Computer zu übetragen, begann man mit der Modellierung einzelner Nervenzellen. Abb. 4 zeigt links das mathematische Modell eines sog. Netzwerkknoten nach McCulloch und Pitts. Zuerst wird die gewichtete Summe der Eingabewerte berechnet und der Schwellenwert subtrahiert. Bei positivem Ergebnis ist der Ausgabewert 1, ansonsten 0. Heute benutzt man als sog. Ak-

[5]Quantum Chromodynamics

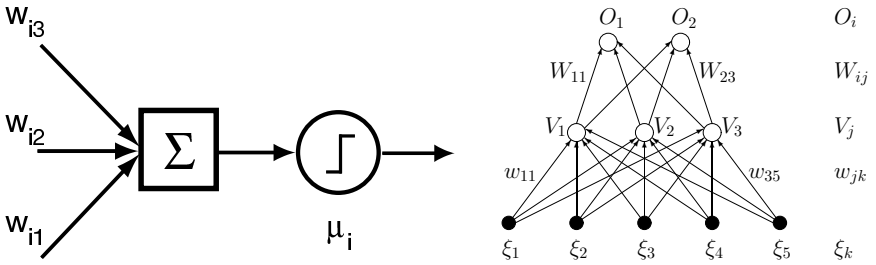


Abbildung 4: Einzelner Netzwerk-Knoten im McCulloch-Pitts-Model (links) und ein 3-lagiges Netzwerk (rechts) [2].

tivierungsfunktion meistens eine Sigmoid-Funktion der Form

$$a \cdot tanh(bx) = a\,\frac{e^{bx} - e^{-bx}}{e^{bx} + e^{-bx}}.$$

Dies hat den Vorteil, daß die Ausgabewerte kontinuierlich zwischen -1 und +1 liegen. Aus Einzelnen dieser 'Nodes' lassen sich mehrlagige Netze aufbauen (siehe Abb. 4 rechts). Der Prozess des Netzwerk-Training besteht aus dem Anpassen der Gewichte, sodaß die Differenz zwischen Ausgabe und Zielwert minimal wird. Die quadratische Fehlerfunktion E ist gegeben durch

$$E[\vec{w}] = \frac{1}{2}\sum_i \sum_\mu [\zeta_i^\mu - O_i^\mu(\vec{w})]^2 \tag{1}$$

mit ζ_i^μ als dem gewünschten Training-Target und $O_i^\mu(\vec{w})$ als aktueller Netzwerk-Ausgabe.

Die Analyse

Die gesamte Analyse besteht aus zwei Teilen. In einem ersten Schritt wird eine Vorselektion durchgeführt um einen erste, grobe Trennung von Signal und Untergrund zu erreichen. Diese basiert auf konventionellen Schnitten. In einem zweiten Schritt werden neuronale Netze eingesetzt, um eine weitere Verbesserung des Signal-zu-Untergrund Verhältnis zu erreichen.

Vorselektion

Die Schnitte der Vorselektion sind angepasst an die zu erwartende Signatur des Signalprozess und lauten zusammengefasst wie folgt:

- ein "tight" Lepton mit Transversalimpuls $P_t > 20$ GeV (Standard CDF Elektron- und Myonselektion),
- fehlende Transversalenergie $\not\!\!E_t > 20$ GeV,
- zwei oder 3 "tight" Jets mit transversaler Energie $E_t > 15$ GeV, $|\eta| < 2.0$,

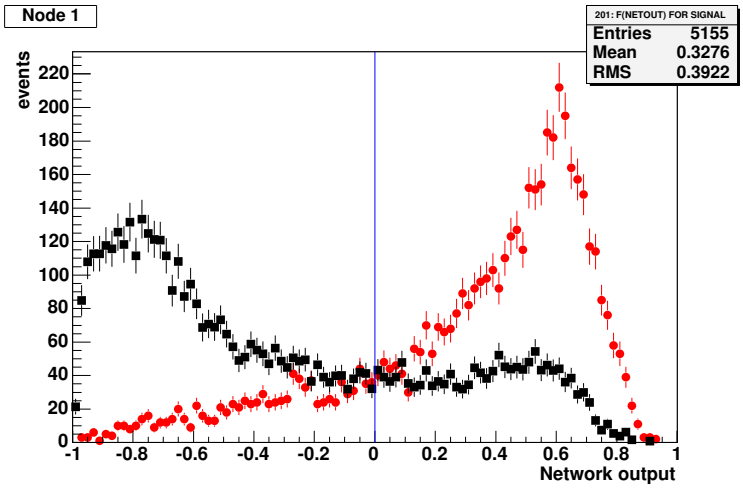


Abbildung 5: Ausgabewert des Netzwerk für Signal (rot) und Background (schwarz), gewichtet im Verhältnis 1:1. Die Higgsmasse des Signalprozess war $m_H = 120$ GeV/c^2.

- mindestens ein b-tag sowie
- Z^0-Veto, Cosmic-Veto und Konversions-Veto.

Anwendung Neuronaler Netze

Das benutzte Netzwerkpaket NeuroBayes® wurde von Prof. Dr. Michael Feindt am Institut für Experimentelle Kernphysik[6] entwickelt[7]. Die Netzwerk-Topologie war stets 3-lagig. In der ersten Lage, auch Eingabe-Lage genannt, befand sich für jede Eingabevariable ein Netzwerkknoten[8]. Die Zahl der Knoten in der zweiten Lage war etwas größer als in der ersten Lage. In der dritten Lage, der Ausgabe-Lage, befand sich stets nur ein Netzwerkknoten.

Der Ausgabewert der neuronalen Netze (jeweils eins pro untersuchter Higgsmasse) lag somit immer zwischen -1 (untergrundartig) und +1 (signalartig). Der Schnitt auf diesen Ausgabewert (dargestellt in Abb. 5) wird so gewählt, daß sich das beste a-priori Limit aus den Monte-Carlo Simulationen ergibt. Anschließend wird für die invariante Masse des 2-Jet-System ein Massenfenster um die zu untersuchende Higgsmasse gelegt. Die Grenzen für dieses Massenfenster werden wieder so gesetzt, daß sich das beste erreichbare Limit einstellt.

[6]Das IEKP ist Teil des Centrum für Elementarteilchenphysik und Astroteilchenphysik CETA.

[7]NeuroBayes® und <phi-t>® sind eingetragene Warenzeichen der Physics Information Technologies GmbH [3].

[8]Zusätzlich gibt es noch einen sogenannten Bias-Knoten in der ersten Lage.

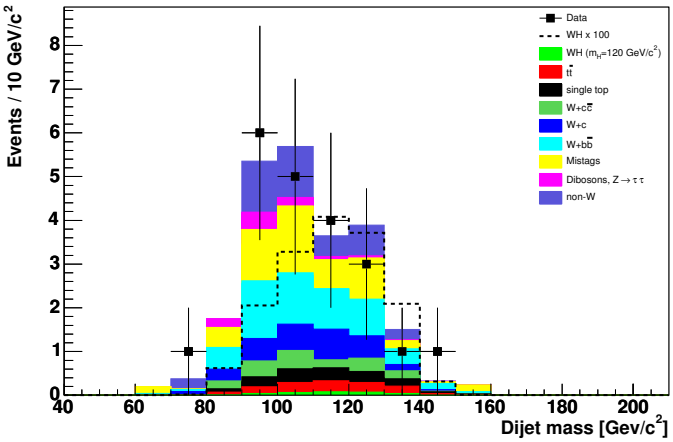


Abbildung 6: Invariante Masse des 2-Jet Systems nach dem Netzwerk-Schnitt. Das neuronale Netz wurde bei einer Higgsmasse von $m_H = 120$ GeV/c^2 trainiert. Die punktierte Linie zeigt den zu erwartenden Signalpeak bei hundertfachem Wirkungsquerschnitt.

Ergebnisse

Die erwarteten Ereigniszahlen nach der Vorselektion sind in Tabelle 5.10 auf Seite 93 gezeigt. Die entsprechenden Zahlen nach Anwendung des neuronalen Netzes sowie nach dem Schnitt auf die invariante Masse des 2-Jet-Systems sind in den Tabellen 6.1 und 6.2 auf Seite 110 angegeben. Die Daten sind unter berücksichtigung der Fehler in guter Übereinstimmung mit dem zu erwartenden Untergrund des Standardmodell. Abbildung 6 zeigt die Verteilung der invariante Masse des 2-Jet-Systems nach dem Schnitt auf die Netzwerkausgabe. Die Zahl der Events im Massenfenster 80 GeV/c^2 $\leq m_H \leq$ 140 GeV/c^2 entspricht den Zahlen in Tabelle 6.2 und ist Grundlage für die Berechnung des Limits im Fall $m_H = 120$ GeV/c^2. Verteilungen für die restlichen sechs Higgsmassen sind auf Seite 112 dargestellt.

In Abbildung 7 zeigen die farbig markierten Graphen die zu erwartende Sensitivität nach jedem Analyseschritt. Die quadratischen Markierungen zeigen das Datenlimit, beruhend auf den Zahlen aus Tabelle 6.2 von Seite 110. Für die offenen Marker wurden systematischen Fehler nicht einbezogen. Die gefüllten Quadrate zeigen die endgültigen Ausschlußgrenzen unter Beachtung der Fehler. Diese liegen daher stets über der ersten Kurve. Wirkungsquerschnitte von etwa 8 pb sind durch diese Analyse ausgeschlossen. Das Limit liegt zu hoch, um eine untere Massengrenze für das Higgs setzen zu können.

Aus den Zahlen von Tabelle 6.2 auf Seite 110 lässt sich die benötigte Luminosität für einen Ausschluß des Standardmodell-Higgs, für ein 3σ-Limit oder eine 5σ-Entdeckung berechnen. Diese sind in Abbildung 8 dargestellt. Die schraffierten Flächen zeigen das Ergebnis dieser Arbeit und geben die benötigte Luminosität für den Zerfallskanal $WH \rightarrow l\nu b\bar{b}$. Die anderen Bänder zeigen die Resultate einer Studie aus dem Jahr 1998 [4]. Wie in der

Vergleichsstudie wurden keine systematischen Unsicherheiten berücksichtigt. Die Breiten der Kurven ergeben sich durch ein hochskalieren der unteren Grenzen um 30%.

Wie in Abb. 8 gezeigt, beträgt die intergrierte Luminosität bei Kombination von WH und ZH ca. 70 fb^{-1} für eine 3-σ Entdeckung eines Standarmodell-Higgsboson von 115 GeV. Higgsmassen bis 114.1 GeV sind durch die LEP-Experiments ausgeschlossen [5]. Diese 70 fb^{-1} liegen über der zu vom Tevatron in Run II zu erwartenden Luminosität von ca. 9 fb^{-1} (vgl. Seite 23). Die hier bestimmte Luminosität für eine 3-σ Entdeckung liegt ca. um einen Faktor 25 über der Schätzung der Higgs Working Group und sind hauptsächlich dadurch begründet, daß in dieser Arbeit wesentlich mehr Untergundprozesse mitberücksichtigt und gemessene Effizienzen benutzt wurden (vgl. Kapitel 6.3.1 auf Seite 117).

Ausblick

Eine Verbesserung des hier bestimmten Higgs-Limits ließe sich am einfachsten mit einem grösseren Datensatz erreichen. Weiter 40 pb^{-1} an Daten stehen schon jetzt zur Verfügung, konnten aber für diese Arbeit nicht mehr berücksichtigt werden.

Aus technischer Sicht sind natürlich verbesserte Algorithmen zu nennen. Zur Berechnung der Ausschlußgrenzen ließe sich statt eines einfachen 'event counting' z.B. ein Fit der invarianten Massenverteilung durchführen. Spezielle Energiekorrekturen für b-Quark-Jets könnten die Breite des Higgs-Massenpeak verringern. Eine höhere b-tagging Effizienz wäre z.B. ebenfalls mit Hilfe neuronaler Netze denkbar [6]. Eine Erhöhung der Signal-Akzeptanz ließe sich durch einen vergrösserten η-Bereich erzielen. Die Berücksichtigung von Elektronen aus den Vorwärts-Detektoren wäre hier möglich. Die Kombination verschiedener Kanäle, auch mit den Ergebnissen des D0-Experiment, würde ebenfalls die Perspektiven der Higgs-Physik am Tevatron stark verbessern.

Sollte das Higgs-Boson am Tevatron nicht gefunden werden, wird die Suche am Large Hadron Collider (LHC) fortgesetzt. Dieser Proton-Proton-Beschleuniger befindet sich zur Zeit im Bau und wird vorrausichtlich im Jahr 2007 in Betrieb gehen. Zwei Multi-Purpose-Experimente, CMS und ATLAS, werden sich mit der Suche nach neuer Physik beschäftigen. Aufgrund der hohen Luminosität und Schwerpunktsenergie von 14 TeV bestehen gute Erfolgsausssichten.

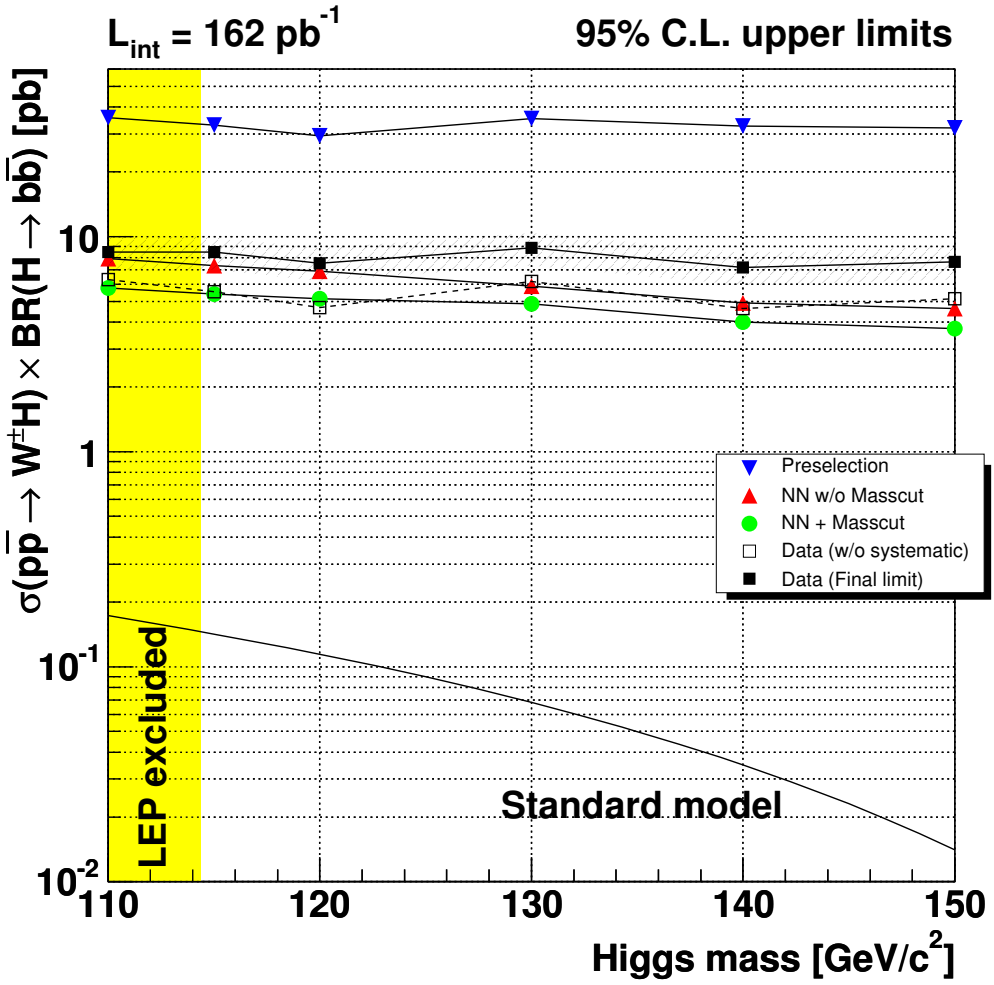


Abbildung 7: Ausschlußgrenzen für den WH Produktionswirkungsquerschnitt multipliziert mit dem Verzweigungsverhältnis $BR(H \to b\bar{b})$ als Funktion der Higgs-Boson Masse. Die Vertrauensgrenze beträgt 95%. Die gelbe Fläche überdeckt den von LEP ausgeschlossenen Massenbereich bis $m_H = 114.1$ GeV/c^2. Die drei farbig markierten Kurven zeigen die erwartete Sensitivität der Vorselektion und nach Anwendung des neuronalen Netzes, ermittelt aus Monte-Carlo Simulationen. Die offenen Quadrate zeigen das Daten-Limit ohne Berücksichtigung der systematischen Fehler. Das endgültige Limit, mit Berücksichtigung der Fehler, zeigen die geschlossenen Quadrate. Im unteren Teil sieht man den NLO-Wirkungsquerschnitt des Standardmodell für den betrachteten Kanal.

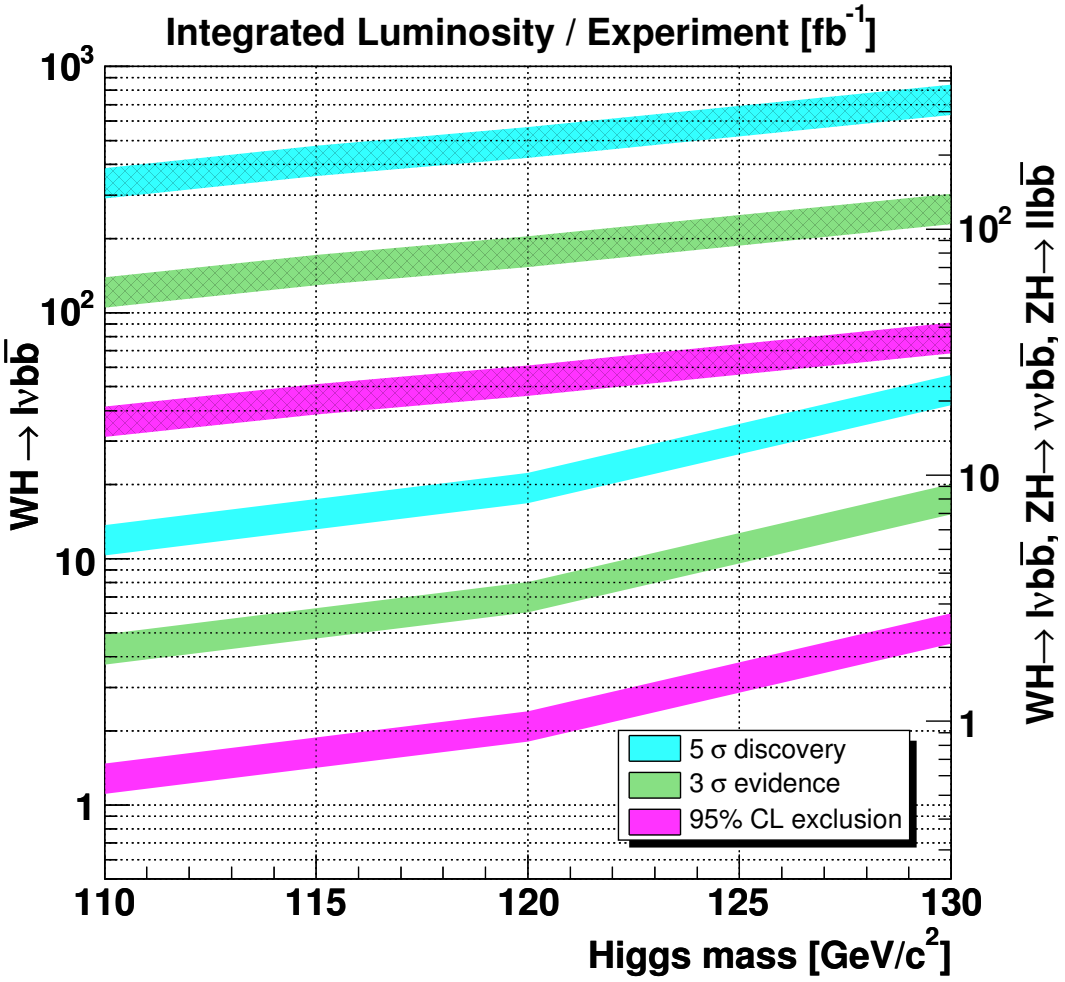


Abbildung 8: Notwendige integrierte Luminositäten für einen 95% *C.L.*-Ausschluß, eine 3σ-Evidenz und eine 5σ-Entdeckung. Die drei unteren Kurven sind die Ergebnisse der Higgs Working Group Studie aus dem Jahr 1998 für den WH Kanal [4]. Die drei oberen schraffierten Kurven stellen das Ergebnis dieser Arbeit dar. Die linke Skala gibt die Luminosität für den WH Kanal an. Die rechte Skala zeigt die Luminosität für eine Kombination des WH mit den ZH Kanälen. Diese liegt ca. einen Faktor zwei niedriger als beim WH Kanal allein. Dies liegt am besseren Verhältnis $S/\sqrt{B}$, das sich durch die Kombination der Kanäle ergeben würde.

'Das Unbeschreibliche, hier ist's
getan.'

Faust: Der Tragödie zweiter Teil
Johann Wolfgang von Goethe

Chapter 1

Introduction

Physics is the only discipline having excellent theories[1]. Its three frontiers are the physics of the "infinitely" big (cosmology), the "infinitely" complicated (chaotic systems) and the "infinitely" small (particle physics). The aim of the latter one is a theory of matter which unifies all known forces to just one interaction. Apart from being simple, it should give a good quantitative description of all observable phenomena. The current theory of particle physics is the standard model. It assumes matter to consist of few elementary particles, an idea which was first introduced by Democritus[2].

1.1 The Standard Model

Today we believe matter to consists of fundamental fermions[3], i. e. quarks and leptons. Due to their properties both are classified in three families (cf. table 1.1 on the following page). The scalar Higgs particle [8, 9] is predicted by the standard model but all searches have been unsuccessful so far. It is introduced by the mechanism of spontaneous symmetry breaking and could explain the generation of particle masses.

Although the standard model has proven to be successful in describing experimental data there are still many open questions. One is the problem of gravity which up to now is not included in the standard model. Other open questions are:

- Why is the charge ratio of quarks and charged leptons exactly 1/3 and why is the number of both lepton and quark generations equal to three ?
- There is no explanation for the symmetry of quarks and leptons w. r. t. the electroweak interaction.
- The theory relies on 18 free parameters which have to be taken from measurements. They cannot be calculated from the standard model in the first place.

[1] In [7] R. Penrose classifies all known theories as being excellent, useful or unproven.
[2] Pre-socratic greek philosopher, 460-370 BC.
[3] Fermions are particles with half integer spin. Particles with integer spin are called bosons.

Leptons (Spin 1/2)	$\begin{pmatrix}\nu_e\\e^-\end{pmatrix}_L \quad \begin{pmatrix}\nu_\mu\\\mu^-\end{pmatrix}_L \quad \begin{pmatrix}\nu_\tau\\\tau^-\end{pmatrix}_L \quad e_R^-,\ \mu_R^-,\ \tau_R^-$
Quarks (Spin 1/2)	$\begin{pmatrix}u\\d\end{pmatrix}_L \quad \begin{pmatrix}c\\s\end{pmatrix}_L \quad \begin{pmatrix}t\\b\end{pmatrix}_L \quad u_R,\ d_R,\ c_R,\ s_R,\ t_R,\ b_R$
Gauge Bosons (Spin 1)	γ, Z^0, $W^\pm$, 8 gluons g
Scalar (Spin 0)	*Higgs*

Table 1.1: The particles of the standard model. All but the Higgs boson have been discovered so far.

- The scalar Higgs particle gives rise to the so called hierarchy problem. This term refers to the instability of the Higgs mass against radiative corrections. A Higgs mass in the order of 1 TeV/c^2 requires a fine tuning of two parameters over 24 significant digits. It is hard to believe that this fine tuning is realised in nature.

Various theories exist which try to solve the problems of the standard model. They all have in common the fact that they contain the standard model as a low energy approximation.

1.2 Motivation

The main objective of this study is to extent previous searches for the Higgs boson at CDF using the new Run II dataset as well as new tools. These tools are so called Neural Networks. They are an interesting alternative to conventional purely cut based analyses. A network developed by Prof. M. Feindt (NeuroBayes®) is being used to classify signal and background processes in this search.

The reaction investigated in this study is the associated production of light SM Higgs-Bosons in the channel $WH \to l\nu b\bar{b}$ because search prospects are best in this production and decay channel. Six different Higgs masses between 110 GeV/c^2 and 150 GeV/c^2 are being tested. The analysis is based on data taken from 2000 to 2003, corresponding to a total integrated luminosity of $\mathcal{L}_{int} = 162\,\mathrm{pb}^{-1}$.

1.3 Overview

The thesis is organised as follows: Chapter 2 on the next page gives an overview of the $p\bar{p}$-collider Tevatron and the CDF detector. In chapter 3 on page 37 the Standard Model of particle physics is presented. Higgs production as well as the corresponding background processes are emphasised. The fourth chapter on page 61 describes the theory of neural networks. Chapter 5 on page 81 explains the analysis itself. In chapter 6 upper limits are set on $\sigma \times BR$, the product of the production cross section times the branching ratio as a function of the Higgs boson mass, starting from page 109. The last chapter on page 123 summarises and gives an outlook on how to further improve the analysis in the future.

'Don't be too proud of this
technological terror.'
Star Wars, Episode IV
Darth Vader

Chapter 2

Tevatron and CDF

2.1 The Tevatron

The Tevatron collider is located at the Fermi National Accelerator Laboratory (FNAL) approximately 60 km west of Chicago/USA. Fermilab is one of the mayor national US laboratories for high energy physics. The Tevatron is a proton-antiproton accelerator with a centre-of-mass energy of 1.96 TeV and currently the most energetic collider in the world. Fig. 2.1 on the following page shows an aerial view of the facility. It has a circumference of about 6 km and the two counter-rotating particle beams collide head on at two interaction points. This is where the multi-purpose detectors, CDF and D0, are located. In order to accelerate protons and antiprotons to an energy of almost 1 TeV, a system of various pre-accelerators is needed. Each one increases the particle energy to the minimum injection energy for the next device in the accelerator chain.

2.1.1 Accelerator chain

The Fermilab accelerator complex is depicted in fig. 2.2 on page 21. The acceleration process starts with the production of H^- ions by adding an electron to hydrogen atoms in an ion source. A Cockroft-Walton accelerator [10] increases the energy of the H^- ions from 0 to 750 KeV. Next the hydrogen ions are sent into a LINAC[1] which is about 150 m long. It boosts the H^- energy from 750 KeV to 400 MeV. Prior to entering the next machine the ions pass through a carbon foil to strip off the two electrons from the proton. The next step of acceleration is performed by the Booster, a synchrotron with a radius of about 75 m. After the proton energy is raised from 400 MeV to 8 GeV they are transferred into the Main Injector (MI). The Main Injector accepts 8 GeV protons and anti-protons from either the Booster, the anti-proton accumulator or the Recycler. It can accelerate the p- and $\bar{p}$-beams to 150 GeV and inject them into the Tevatron.

The Tevatron receives protons and anti-protons from the Main Injector at 150 GeV and accelerates them to their final energy of 980 GeV. Typically the two beams consist of 36 bunches with 180×10^9 protons/bunch for the p-beam and 12×10^9 anti-protons/bunch

[1] Linear Accelerator

Figure 2.1: Aerial view of the Fermilab site. The circular structure in the background is the inner maintenance road of the Tevatron. The one in the foreground shows the outer maintenance road for the Main Injector and the Recycler which are both located in the same tunnel. The main building and the meson area can be seen in the upper left corner.

for the $\bar{p}$-beam. After the ramping[2] is complete collisions are initiated at the B0 and D0 interaction regions. Stores[3] are kept for typically 16 h while more anti-protons are made for the next shot[4]. It takes 10-16 h to create enough anti-protons for a shot. Hence the $\bar{p}$-production is a limiting factor for the Tevatron luminosity.

2.1.2 Anti-proton production

Anti-proton production is accomplished by extracting 120 GeV protons from the Main Injector and directing them onto a nickel target . The protons striking the target produce anti-protons as well as many other secondary particles in the proton-nucleus interaction. A Lithium lens focuses these particles and a bend magnet selects negative particles around 8 GeV. Particles other than protons decay away and only anti-protons are left in the beam transferred to the anti-proton ring.

The anti-proton ring consists of two parts, the Debuncher and the Accumulator, both having a triangular shape. In the Debuncher particles enter with a narrow time and broad energy spread. The RF[5] is phased such that high energy particles are decelerated and low energy particles are accelerated resulting in a narrow energy but broad time

[2]To excite a magnet with a time dependent excitation current.

[3]To inject circulating beam into an accelerator and keep it there for long periods of time.

[4]The injection of protons and anti-protons into the Tevatron in preparation for colliding beams operation.

[5]Radio Frequency

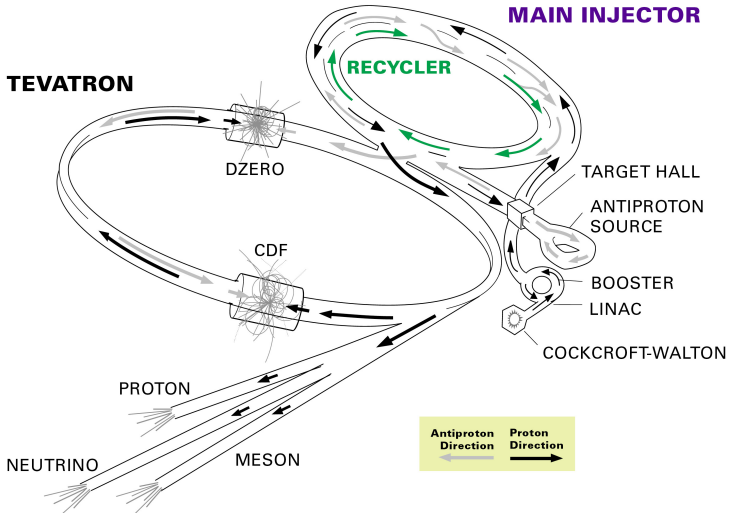


Figure 2.2: Schematic view of the Tevatron accelerator complex.

spread. The beam having a bunch structure in the beginning of the process has been "de-bunched". In addition, the $\bar{p}$-emittance is reduced by stochastic cooling. Pickups detect deviations from the ideal particle orbit which are used to kick the orbit back to nominal values. This reduces the transverse emittance in a statistical way. Finally the anti-protons are transferred to the Accumulator. Here the anti-protons are stacked with typical rates of about $7 \times 10^{11} \bar{p}$/h up to a maximum of about 120×10^{10} $\bar{p}$'s. At that point they are transferred to the Main Injector where they are accelerated from 8 GeV to 150 GeV before being sent to the Tevatron.

2.1.3 Recycler

The Recycler is an 8 GeV anti-proton storage ring installed near the ceiling of the Main Injector tunnel. Its di- and quadrupoles[6] are made out of permanent magnets. Hence the particle energy is fixed to a constant value. The original goal of the Recycler was

- to store anti-protons from the accumulator, thereby increasing the total anti-proton production capacity and
- to recover anti-protons from a Tevatron store for use in subsequent stores.

Due to technical problems the latter goal has been abandoned. For the stacking rate planned, the difference in integrated luminosity with and without recycling is only about 10% [11]. As of 2003, the Recycler is not being used in standard operation. Due to a vacuum incident in January 2003 the Recycler has not been commissioned and the planning for its use has been delayed. The commissioning phase is supposed to start in calendar year 2004. It is planned to include an electron cooler into the Recycler ring. The cooling of ion beams by a co-moving low emittance electron beam is a well established

[6] A magnet consisting of four poles, used for focusing beams of particles.

technique for nuclear physics facilities [12, 13]. This project, however, is the first attempt at achieving medium energy cooling. Previous cooling systems were built at an order of magnitude lower beam energy. If successful, the implementation of electron cooling will allow very large antiproton stacks accumulated in the Recycler Ring to be transferred to the Tevatron with small longitudinal emittance.

2.1.4 Luminosity

The most important quantity characterising a collider, apart from its centre-of-mass energy, is the instantaneous Luminosity $\mathcal{L}$. Together with the cross-section for a particular physics process it defines the event rate

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma \tag{2.1}$$

at which particle interactions occur. The integrated Luminosity $\mathcal{L}_{int}$ is just the time integral of $\mathcal{L}$:

$$\mathcal{L}_{int} = \int \mathcal{L}\, dt. \tag{2.2}$$

It is usually measured in pb^{-1} which stands for inverse picobarn[7]. At the Tevatron the instantaneous Luminosity is given by

$$\mathcal{L} = \frac{N_p N_{\bar{p}} B f_0}{4\pi\sigma_x\sigma_y} \tag{2.3}$$

where N_p and $N_{\bar{p}}$ are the numbers of protons and anti-protons in a particle bunch, B is the number of bunches and f_0 is the revolution frequency of the beam ($\approx$ 50 KHz). σ_x and σ_y characterise the width of the Gaussian beam profile in x and y. Typically the number of particle bunches is 36 for both particle types resulting in a bunch-crossing time of 132 ns or an interaction frequency of 40 MHz. The average number of interactions per bunch-crossing is about 2. Fig. 2.3 on the facing page shows the average number of interactions per bunch crossing $\bar{N}$ as a function of luminosity.

Table 2.1 on page 25 shows a list of various Tevatron parameters for both Run I and Run II [14]. The performance goals have been met in a few runs so far. Fig. 2.4 on page 24 shows the initial Luminosity per store. The record is about $10 \times 10^{31}\mathrm{cm}^2\,\mathrm{s}^{-1}$. The integrated Luminosity $\mathcal{L}_{int}$ of Run II can be seen in fig. 2.5 on page 24. Until July 2004 about 650 pb^{-1} have been delivered by the Fermilab beams division. The average data taking efficiency of CDF has been around 80% resulting in approximately 520 pb^{-1} written to tape. These data are available for physics analysis. Requirring a certain data quality further reduces the amount of usable data.

[7] The unit barn is defined as $10^{-28}\mathrm{m}^2$.

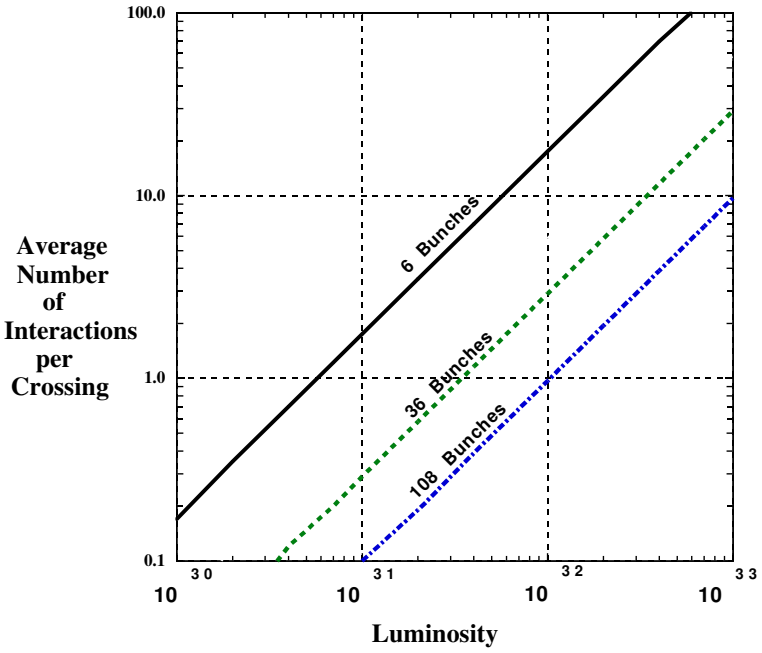


Figure 2.3: Average number of interactions per bunch crossing for 6, 36 and 108 bunches vs. instantaneous luminosity. The three graphs represent mean values of a Poisson distribution.

2.1.5 Outlook

The scope of the Fermilab Tevatron program has been reviewed by the DOE[8] [15] in October 2002. Plans for operating the Tevatron with 132 nsec bunch spacing rather than the present 396 ns spacing have been dropped from the project scope. One reason to skip the luminosity upgrade plans to operate the Tevatron with 136 bunches/beam and a bunch crossing time of 96 ns was the large effect of beam-beam interactions. The effect was found to be much stronger than anticipated. In summer 2003 the Fermilab directorate has worked out a new Run II Luminosity Upgrade Plan which defines two Luminosity projections through Fiscal Year 2009:

- a base projection of 4.4 fb^{-1} and
- a design projection of 8.6 fb^{-1}.

However both base and design projection assume successful integration of electron cooling in the Recycler. This represents a significant uncertainty since an electron cooling system at such high energies has never been built before [16].

[8]Department of Energy

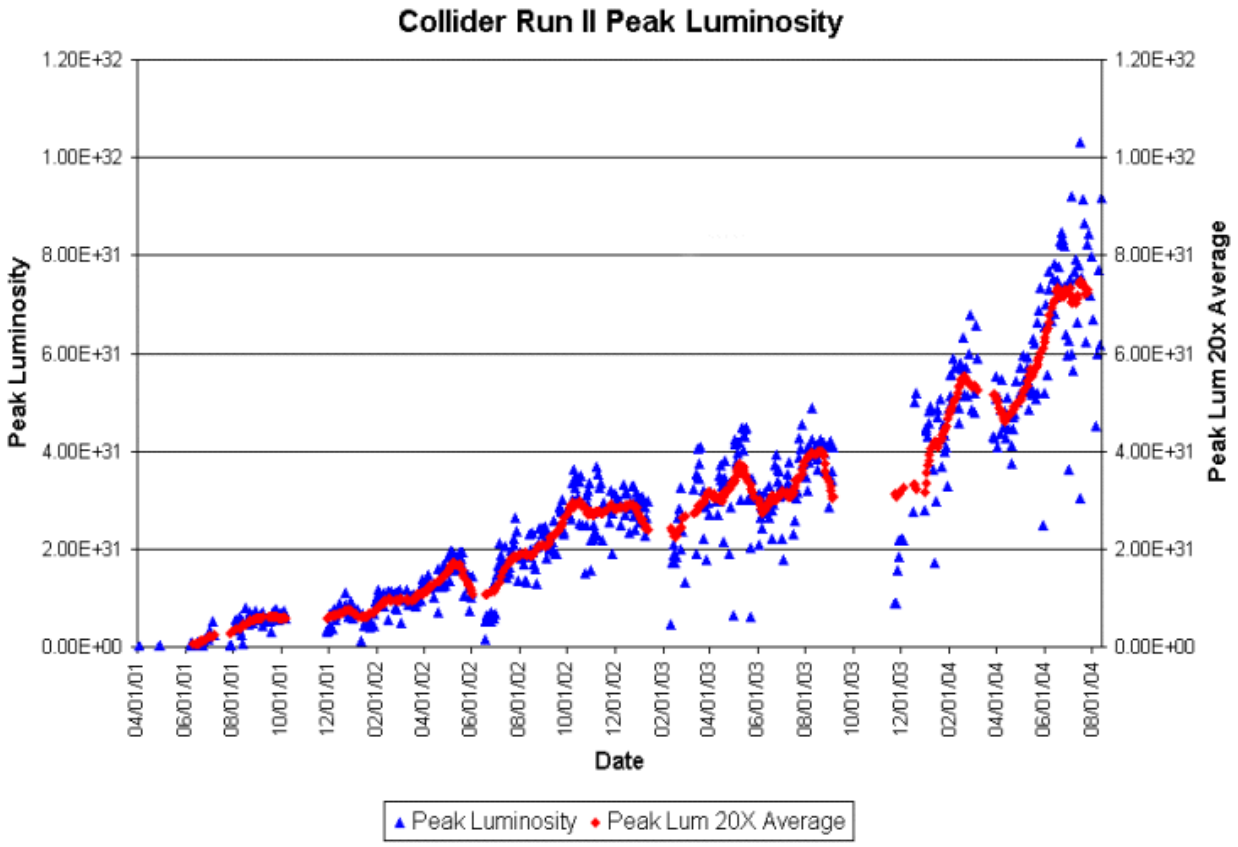


Figure 2.4: Initial peak luminosity per store in Run II. The record luminosity achieved in July 2004 is about $10 \times 10^{31} \mathrm{cm}^2\,\mathrm{s}^{-1}$. Intervals without data points represent shutdown periods for maintenance work.

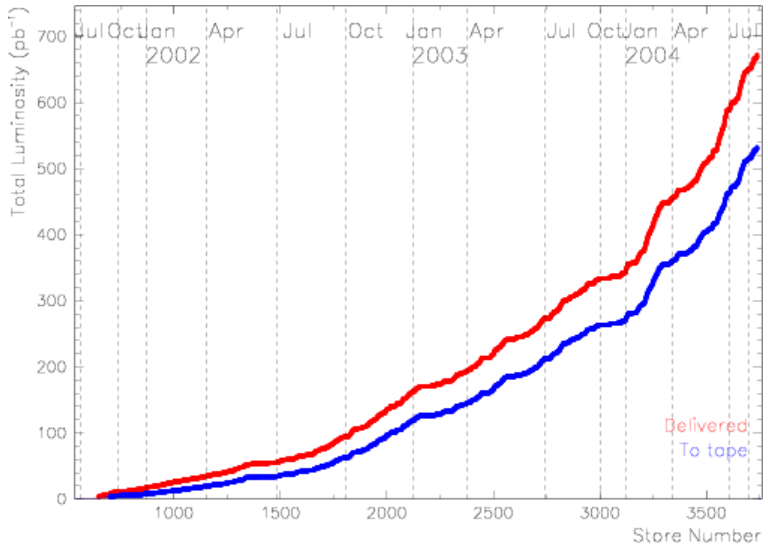


Figure 2.5: Integrated luminosity vs. store number. The upper curve shows the luminosity delivered by the Tevatron. The lower curve represents $\mathcal{L}_{int}$ for data recorded by CDF.

Run II Parameter List	RUN Ib (1993-95) (6x6)	Run II (36x36)	Unit
Protons/bunch	$2.3 * 10^{11}$	$2.7 * 10^{11}$	
Antiprotons/bunch	$5.5 * 10^{10}$	$3.0 * 10^{10}$	
Total Antiprotons	$3.3 * 10^{11}$	$1.1 * 10^{12}$	
Pbar Production Rate	$6.0 * 10^{10}$	$1.0 * 10^{11}$	hr^{-1}
Proton emittance	23π	20π	mm*mrad
Antiproton emittance	13π	15π	mm*mrad
β^*	35	35	cm
Energy	900	1000	GeV
Antiproton Bunches	6	36	
Bunch length (rms)	0.60	0.37	m
Crossing Angle	0	0	μrad
Typical Luminosity	$0.16 * 10^{31}$	$0.86 * 10^{32}$	$\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$
Integrated Luminosity	3.2	17.3	pb^{-1}/week
Bunch Spacing	~ 3500	396	ns
Interactions/crossing	2.5	2.3	

Table 2.1: List of Tevatron machine parameters for Run Ib and Run II. The Run II numbers are design values. The quantities for Run Ib represent measured data.

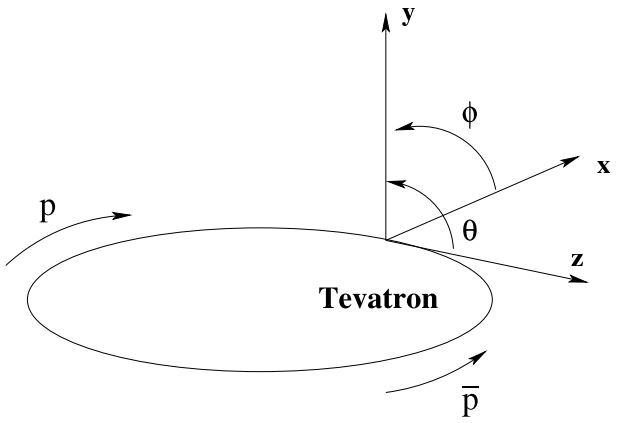


Figure 2.6: The CDF coordinate system [1]. The proton beam defines the positive z-axis with $\theta = 0$. ϕ is measured from the Tevatron plane.

2.2 The CDF II Detector

The CDF experiment is a multi-purpose particle detector. It has been built and is maintained by an international collaboration of about 600 physicists from 50 institutes in 11 countries. Its size is about 15 m × 10 m × 10 m and it weighs approx. 5000 tons.

The coordinate system used by CDF is depicted in fig. 2.6. The z-axis is defined by the direction of the particle beam. Protons are moving into the positive z-direction. The polar angle θ is measured from there. Hence $\theta = 0$ for the protons and $\theta = \pi$ for the anti-proton beam. However it is common to describe polar angles by the pseudo-rapidity η[9] which is defined by:

$$\eta = -ln(tan\frac{\theta}{2}). \tag{2.4}$$

The azimuthal angle ϕ is measured from the Tevatron plane. Particles having $\phi = 0$ and $\theta = \frac{\pi}{2}$ point away from the centre of the accelerator ring.

The layout of the CDF experiment follows common design principles for this kind of detector, i. e. it has forward-backward as well as spherical symmetry around the beam pipe. Different detector systems are placed around the interaction region in various layers (cf. fig. 2.7 and 2.8). The tracking system, being closest to the beam pipe, consists out of silicon detectors in the centre surrounded by the COT, a gas drift chamber. The COT itself is surrounded by a large superconducting solenoid which creates a magnetic field of 1.4 T. A Time-Of-Flight system is placed between the solenoid and the COT. Its main purpose is particle identification. The next detector is the calorimeter which is divided into an electromagnetic and a hadronic part. The outermost detector components are the muon chambers. Four different subsystems almost completely surround the inner detector systems.

The following sections describe the different detector components in more detail. A complete description of the CDF II detector can be found elsewhere [17].

[9] Jets described in the $\eta - \phi$ space look the same, regardless of their η.

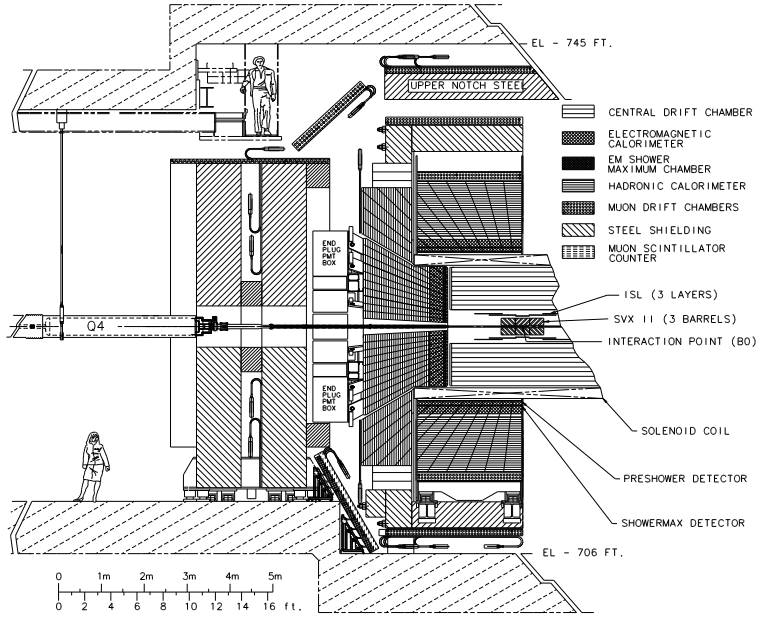


Figure 2.7: Elevation view of the CDF II detector.

2.2.1 Tracking System

Silicon Detectors

The complete silicon detector consists out of three different sub-systems: L00, SVX II and the ISL. In total, eight layers of silicon surround the beam pipe, ranging from $r = 1.35$ cm to a radius of $r = 28$ cm and length from 90 cm to almost two meters. The total detector area is 6 m^2 with 722,000 readout channels.

Layer 00: Layer 00 is the innermost layer of silicon. It is directly glued on to the beryllium beam pipe in a hexagonal shape. However the detector elements don't provide complete ϕ-coverage. L00 is supposed to strongly improve the tracking resolution of CDF because of its small distance to the interaction region. Unfortunately, due to high electronic noise it hasn't been used so far.

Silicon Vertex Detector (SVX II): The SVX II consists of five layers of double sided detectors at radii between 2.4 cm and 10.7 cm. In total it is 96 cm long and covers rapidities up to $|\eta| \leq 2$. The layers are assembled in three cylindrical barrels.

Intermediate Silicon Layers (ISL): In order to link the silicon hits from the SVX II and the tracks from the COT a third silicon sub-detector has been added to the design: the ISL. In the central region a single layer of silicon is placed at a radius of 22 cm. In the region $1.0 \leq |\eta| \leq 2.0$ two layers are placed at radii of 20 and 28 cm as can be seen in Fig. 2.8 on the next page.

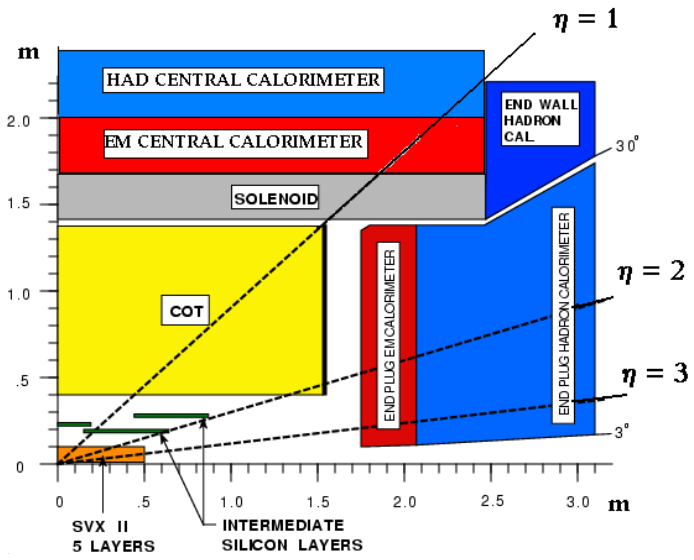


Figure 2.8: Longitudinal view of the CDF II detector [1]. The Time-Of-Flight system which is located between the COT and the solenoid is not shown.

Central Outer Tracker (COT)

The COT is a cylindrical open-cell drift chamber. Its inner and outer radii are 44 cm and 132 cm and the length of the active region is 310 cm. Hence it covers the region $|\eta| \leq 1.0$. One mayor design goal of the COT was to achieve a maximum drift time less than 132 ns which was supposed to be the bunch crossing time in Run IIb. The COT is designed to find charged particle tracks with transverse momenta as low as $p_t = 400$ MeV. It is segmented into 4 axial and 4 stereo superlayers. Each super-layer consists of 12 sense wires alternated with 13 potential wires which shape the field within the cell. In total 96 measurements can be made in the radial direction. Argon-Ethane (50:50) is used as drift gas. Although it has a much poorer position and direction resolution than the silicon detectors it provides a much better momentum resolution. This is due to the greater radial extension and a higher purity due to a lower track density w. r. t. the silicon. Using both the silicon detectors and the COT the overall momentum resolution for charged particles is $\delta p_t / P_t^2 \leq 0.1\%/\text{GeV/c}$.

2.2.2 Calorimeters

The CDF calorimeters are designed to accurately measure particle energies of electrons, photons and hadrons. In total there are five different calorimeter subsystems installed in CDF: the central EM[10] and Hadron calorimeters, the End-Plug EM and Hadron calorimeters (PEM and PHA) and the End-Wall Hadron calorimeter (WHA). Fig. 2.8 on the facing page shows their location. The CDF calorimeters provide complete ϕ-coverage and η-coverage up to $|\eta| \leq 3.64$ and are segmented such that they form a projective tower geometry which points to the interaction region. The central and end wall calorimeters are made of two halves, referred to as east and west arcs. Each half consists out of 24 wedges. A single wedge covers 15° in ϕ and is subdivided into 10 towers of 0.1 units in η. Fig. 2.9 on the next page shows one of the wedges for the central calorimeter.

The CDF calorimeters are so called sampling calorimeters. Several layers of active detection material are interspaced with layers of absorption material. The electromagnetic calorimeters use lead as absorption material and the hadronic ones use iron. Scintillator material has been chosen as active material. Table 2.2 on the following page summarises some properties of the CDF calorimeters. Their thickness is given in terms of X_0[11] and λ[12].

Both EM-calorimeters have pre-shower (CPR[13]) and stereo shower maximum detectors (CES[14]) to improve their spacial resolution.

Detector	η Range	Active medium	Thickness	Energy Resolution
CEM	$\lvert\eta\rvert \leq 1.1$	polystyrene scintillator	19 X_0, 1 λ	$13.7\%/\sqrt{E_t} \oplus 2\%$
PEM	$1.1 \leq \lvert\eta\rvert \leq 3.64$	proportional chambers	21 X_0, 1 λ	$16\%/\sqrt{E} \oplus 1\%$
CHA	$\lvert\eta\rvert \leq 0.9$	acrylic scintillator	4.5 λ	$50\%/\sqrt{E_t} \oplus 2\%$
WHA	$0.7 \leq \lvert\eta\rvert \leq 1.3$	acrylic scintillator	4.5 λ	$75\%/\sqrt{E} \oplus 4\%$
PHA	$1.2 \leq \lvert\eta\rvert \leq 3.64$	proportional chambers	7 λ	$80\%/\sqrt{E} \oplus 5\%$

Table 2.2: Summary of CDF calorimeter properties. The $\oplus$ signifies that the constant term is added in quadrature. The resolutions are given for energies measured in GeV. They apply for incident electrons and photons in the case of the EM calorimeters. For the hadronic calorimeters the quoted values apply to incident isolated pions.

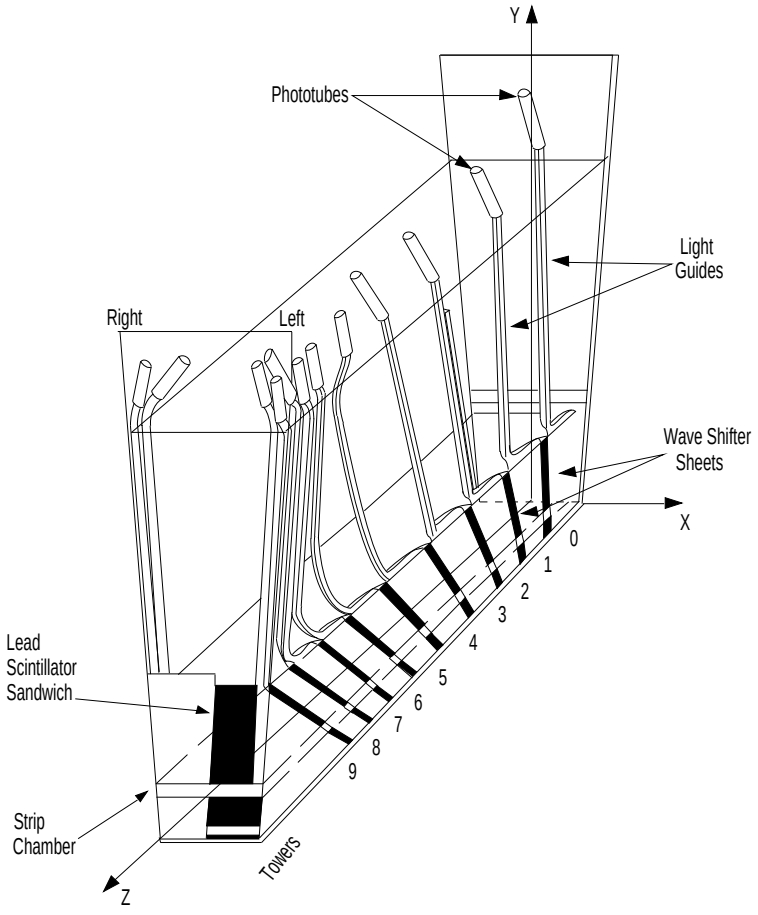


Figure 2.9: A wedge of the central calorimeter showing the ten towers in η. The lower part is the electromagnetic calorimeter and includes a strip chamber, the CES. The upper part is the hadronic calorimeter.

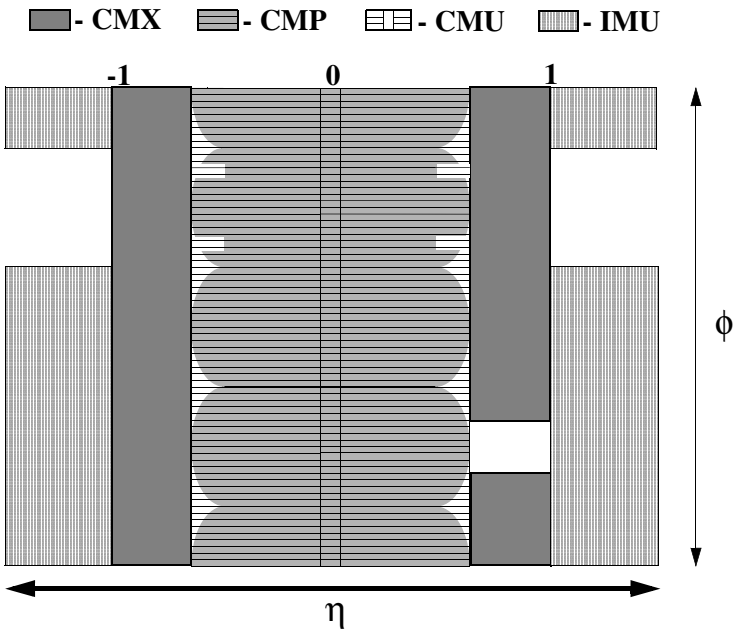


Figure 2.10: $\eta - \phi$ coverage of the CDF muon system. The CMU and CMP are overlapping in the central region.

2.2.3 Muon Detectors

The central calorimeters act as a hadron absorber for the Central Muon Upgrade (CMU). It consists of four layers of drift chambers located outside the central hadronic calorimeter. It covers 84% of the solid angle for the pseudorapidity interval $|\eta| \leq 0.6$ and can be reached by muons with a transverse momentum greater than 1.4 GeV. In 1992 the system was upgraded by adding 0.6 m of steel behind the CMU and additional four layers of drift chambers behind the steel. This new system is called CMP[15]. For $|\eta| \leq 0.6$ the CMP covers 63% of the solid angle while both systems overlap in 53% of the solid angle. In addition, the pseudo-rapidity range of $0.6 \leq |\eta| \leq 1.0$ is covered by the CMX[16] to 71% of the solid angle. Fig. 2.10 shows the $\eta - \phi$ coverage for the different systems.

The changes for Run II in the muon systems represent incremental improvements. New Chambers are added to the CMP and CMX systems to close gaps in the azimuthal coverage and the shielding is improved. The forward muon system is replaced with the IMU[17], covering from $1.0 \leq |\eta| \leq 1.5$. Table 2.3 on the following page gives an overview of the different muon systems.

[10]Electromagnetic
[11]Radiation length, usually measured in $\mathrm{g\,cm^{-2}}$. It is the mean distance over which a high energy electron losses all but 1/e of its energy.
[12]Nuclear interaction length. The mean free path between inelastic interactions, measured in $\mathrm{g\,cm^{-2}}$.
[13]Central Pre-Radiator
[14]Central Electromagnetic strip/wire gas chamber
[15]Central Muon Upgrade
[16]Central Muon Extension
[17]Intermediate Muon System

	CMU	CMP/CSP	CMX/CSX	IMU
Pseudo-rapidity coverage	$\|\eta\| \leq 0.6$	$\|\eta\| \leq 0.6$	$0.6 \leq \|\eta\| \leq 1.0$	$1.0 \leq \|\eta\| \leq 1.5$
Total counters		269	324	864
Min p_T of detectable μ	1.4 GeV	2.2 GeV	1.4 GeV	$1.4 - 2.0$ GeV

Table 2.3: Design parameters of the CDF II muon detectors.

2.2.4 Other Detectors

CLC

The CLC[18] consists of two modules which are located in the so-called "3-degree holes" inside the CDF plug calorimeters which cover the $3.7 \leq |\eta| \leq 4.7$ pseudo-rapidity range. Each detector module is made of 48 thin, long, conical, gas-filled Čherenkov counters. These counters are arranged around the beam pipe in three concentric layers with 16 counters each and point to the centre of the interaction region. Isobutane is used as radiator for it has one of the largest refractive indices for commonly available gases (1.00143) and good transparency for photons in the ultraviolet region where most of the Čherenkov light is emitted. The CLC monitors the average number of inelastic $p\bar{p}$ interactions by measuring the number of particles and their arrival time in each bunch crossing. For these primary particles efficient PMTs[19] collect about 100 photoelectrons with good amplitude and time resolution [18].

Time Of Flight (TOF)

Between the COT and the solenoid the Time-Of-Flight system is installed. It consists out of scintillator panels which provide both timing and amplitude information. The timing resolution is 100 ps. The detector covers the central region up to $\eta \leq 1.1$ and will be capable of identifying kaons from pions by their flight time difference.

Forward Detectors

Beam Shower Counters (BSC): The BSC can detect particles originating from the interaction point at very small angles ($5.5 \leq |\eta| \leq 7.5$). It will be used to study single diffraction (SD) and double-pomeron exchange (DPE) processes. In addition to their use in the forward physics program these detectors can be used for beam loss measurements of the Tevatron. The BSC system consists of four stations on the West and East side of CDF. All stations are located along the beam pipe, at increasing distances from the IP as one goes from BSC-1 to BSC-4. The stations are made of two scintillator counters. The scintillator material is SCSN-81 and has a thickness of 1/4" for the BSC-1 and 3/8" for the other stations. It is preceded by a 3/8" thick lead plate to convert photons. Each counter is viewed by its own PMT [19]. Fig. 2.11 on the next page shows where the different forward detector systems are located.

[18] Čerenkov Luminosity Counter
[19] Photo Multiplier Tubes

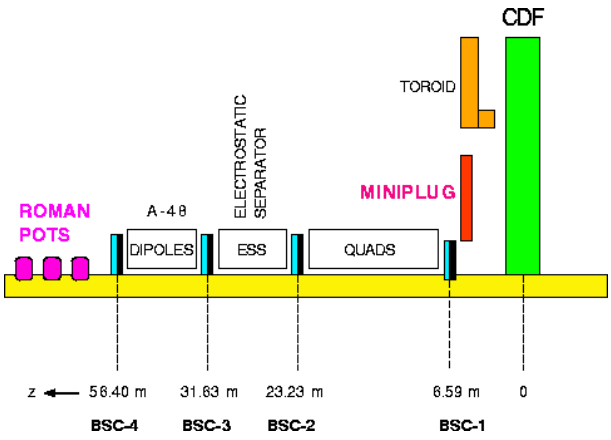


Figure 2.11: Location of the CDF forward detectors.

Miniplug calorimeters (MP): The MiniPlug calorimeters measure the energy and lateral position of particles in the forward region. They extend the pseudo-rapidity region covered by the Plug calorimeters to the beam pipe ($3.6 \leq |\eta| \leq 5.2$). They consist of lead and liquid scintillator read out by wavelength shifting (WLS) fibres perpendicular to the lead plates and parallel to the beam pipe. This pixel-type tower-less geometry is suitable for "calorimetric tracking". The MiniPlug energy resolution for electrons is given by $\sigma/E = 18\%/\sqrt{E}$ where E is the incident particle energy in GeV [20].

Roman Pots: The three Roman Pot stations [21] are located at about 57 meters from the Interaction Point, and approximately one meter apart from each other. They consist of a total of 240 scintillator fibre channels and of 3 scintillator counters.

2.2.5 Data Acquisition and Trigger

A schematic view of the CDF DAQ[20] and trigger system is given in fig. 2.13 on page 35. The trigger plays an important role to efficiently extract the most interesting physics events from the large number of minimum bias and background events and to reduce the amount of data to a reasonable volume. A huge rejection already at trigger level is essential to retrieve the high statistics needed for the search for new physics.

The CDF trigger is a three level system. The time available for event processing increases in each level of the trigger which permits the use of an increasing amount of information to either accept or reject an event. While Level-1 and Level-2 triggers are based on only parts of the detector information, the Level-3 triggers makes use of the complete event data. A signal is defined as an event where a variable (for instance the energy in the calorimeter) lies above a certain trigger threshold. A list of quantities that can be cut on at the different trigger levels is given in [22]. L1 and L2 are hardware triggers while L3 is a software trigger. An optimised version of the reconstruction executable is running on a Linux PC farm with about 100 nodes. The design processing rates for Level-1, 2 and 3 are 50 kHz, 300 Hz and 50 Hz respectively. The typical event size is about 250-300 kB.

[20]Data Acquisition

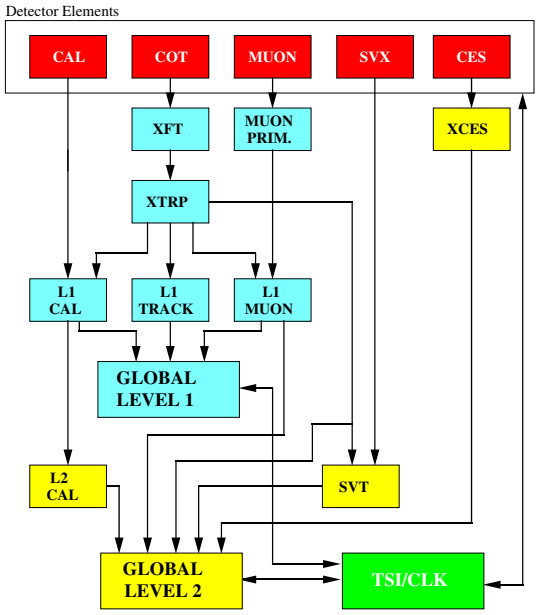


Figure 2.12: Functional block diagram of the CDF L1 and L2 trigger system.

The L1 triggers base their decisions on information of the calorimeters, the muon system, the forward detectors and the drift chamber (see fig. 2.12). The XFT[21] reconstructs r/ϕ tracks in the COT with a transverse momentum resolution of $\delta p_t/p_t^2 = 0.01651$ GeV^{-1} and an angular resolution of 5.1 mrad.

An important feature of Level 2 is the SVT[22]. It adds silicon r/ϕ hits to the L1 XFT tracks. This allows to select events with two tracks having an impact parameter larger than 120 μm in order to identify secondary vertices[23]. This will make a large number of important processes involving the hadronic decays of bottom hadrons accessible. This is of special interest for Higgs physics since for low m_h the Higgs boson predominantly decays into two bottom quarks.

Full event reconstruction takes place on the L3 trigger farm and hence a wide variety of requirements can be imposed on the events passing L3 [23]. Computing power on the order of one second on a Pentium II CPU[24] is available per event.

Events passing the final trigger level belong to a certain trigger path. Each "path" is a unique combination of L1, L2 and L3 triggers. The trigger decisions are combined via a logical "AND". Many paths combined by a logical "OR" can be used to feed a single dataset. The data is written to approximately 20 streams and stored on tape. After reprocessing the events they are split up into more specific datasets. During measurements the data quality is monitored online [24].

[21] Extremely Fast Tracker
[22] Silicon Vertex Tracker
[23] A displaced vertex wrt. the primary vertex.
[24] Central Processing Unit

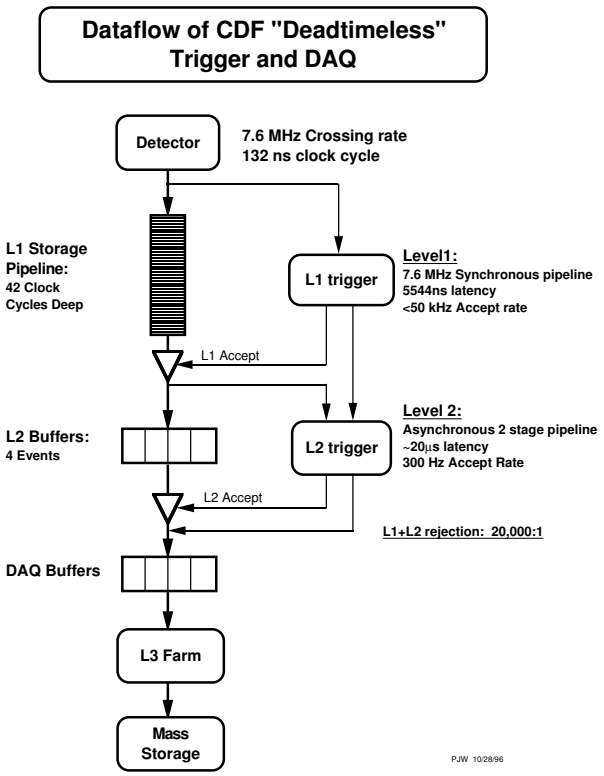


Figure 2.13: Data flow schematic of the three level pipelined and buffered trigger system.

'... the Higgs boson may be just
around the corner.'
Electroweak symmetry breaking
and the Higgs sector
Chris Quigg

Chapter 3

Theoretical Foundations

The first part of this chapter gives a brief introduction into the current theory of particle physics. For a more detailed coverage of the topic the reader is referred to standard textbooks like [25, 26, 27, 28]. Review articles and collections of experimental results are published regularly by the Particle Data Group [29]. Section 3.2 describes Higgs production at the Tevatron and section 3.3 deals with the various background types.

3.1 The Standard Model

3.1.1 Bosons and Fermions

In the Standard Model all particles are classified as being fermions or bosons. Both classes are characterised by their spin. Fermions carry half integer spin, i. e. the values of the spin quantum number are $n + \frac{1}{2}$ with n being a positive integer including zero. Bosons have integer spin and act as exchange particles for the fundamental forces. Fermions are the building blocks of the matter surrounding us.

Today we know four different fundamental interactions: the strong and the weak interaction, electromagnetism and gravitation. Since gravitation is the weakest of all forces and there is no renormalisable theory describing it, it will not be considered any further. Its effects in particle physics are negligible because of the small particle masses involved. In the Standard Model strong interactions are described by QCD[1]. The weak and the electromagnetic force have been unified by the electroweak theory from Glashow, Salam and Weinberg [30, 31].

The fermions are divided into two subgroups: leptons and quarks, each of which have similar characteristics. They exist in three different families or generations. The particle properties of different generations are similar while masses tend to increase with increasing generation number. The fermions of the Standard Model and some of their characteristic properties are listed in table 3.1. It gives the electric charge, isospin and weak hypercharge. The weak isospin I_3 groups together the fermions participating in the weak interaction. The first group are left-handed isodoublets with isospin $\pm\frac{1}{2}$ denoted by L. The second

[1] Quantum Cromo Dynamics

	Generation			Q	I_3	Y
	1	2	3			
Leptons	$\begin{pmatrix}\nu_e\\e^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\\mu^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\\tau^-\end{pmatrix}_L$	$\begin{pmatrix}0\\-1\end{pmatrix}$	$\begin{pmatrix}+\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}$	-1
	e_R	μ_R	τ_R	-1	0	-2
Quarks	$\begin{pmatrix}u\\d'\end{pmatrix}_L$	$\begin{pmatrix}c\\s'\end{pmatrix}_L$	$\begin{pmatrix}t\\b'\end{pmatrix}_L$	$\begin{pmatrix}+\frac{2}{3}\\-\frac{1}{3}\end{pmatrix}$	$\begin{pmatrix}+\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}$	$-\frac{1}{3}$
	u_R	c_R	t_R	$+\frac{2}{3}$	0	$+\frac{4}{3}$
	d_R	s_R	b_R	$-\frac{1}{3}$	0	$-\frac{2}{3}$

Table 3.1: Fermions of the Standard Model. The listed properties are the electric charge Q, the third component of the weak isospin I_3 and the weak hypercharge Y. Subscripts L and R indicate the chirality. For each particle there is a corresponding anti-fermion with equal mass and multiplicative quantum numbers but opposite additive quantum numbers.

Interaction	Boson	Spin	el. Charge	Mass [GeV/c^2]
electromag.	Photon γ	1	0	0
weak	Z^0	1	0	91.2
weak	$W^\pm$	1	±1	80.4
strong	8 Gluons g	1	0	0

Table 3.2: Bosons of the Standard Model. In the SM with minimal Higgs sector a neutral scalar Higgs Boson has to be added.

group consists of right-handed singlets with isospin 0 and index R. In addition to the electric charge quarks carry colour charge, while anti-quarks carry anti-colour. The values of the colour charge can either be red (R), green (G) or blue (B).

The force carriers of the Standard Model are listed in table 3.2. The exchange particle of the electromagnetic interaction described by QED[2] is the massless photon. It couples to electric charges while being neutral itself. The field quanta of the weak interaction are the $W^\pm$ and the Z^0 Boson with masses of 80.4 and 91.2 GeV/c^2 respectively. They couple to all leptons and quarks. The strong interaction is mediated by massless gluons which carry colour charge. They couple to quarks and unlike photons to each other.

The scalar Higgs particle [8, 9] is postulated by the standard model but all searches have been unsuccessful so far. It is introduced by the mechanism of spontaneous symmetry breaking and explains the generation of particle masses.

[2]Quantum Electro Dynamics

3.1.2 The gauge principle

An important concept within the Standard Model is the so called gauge principle. Gauge invariance implies that physics is invariant under phase transformations. One can distinguish two kinds of transformations. The first kind are local phase transformations, the second one are global ones. Global gauge transformations change a wave function $\Psi(x)$ to $\Psi'(x)$ by applying a phase factor $e^{i\alpha}$, with α being a constant phase:

$$\Psi(x) \rightarrow \Psi'(x) = e^{i\alpha}\Psi(x). \tag{3.1}$$

For local gauge transformations the phase factor depends on the local space and time coordinates. For instance in QED such a transformation is given by

$$\Psi' = e^{iq\chi(x)}\Psi \tag{3.2}$$

where q is the electric charge. Applying equation 3.2 to the Dirac equation for a particle in free space

$$(i\gamma^\mu\partial_\mu - m)\Psi(x) = 0 \tag{3.3}$$

one gets

$$\begin{aligned}(i\gamma^\mu\partial_\mu - m)\Psi'(x) &= -q\gamma^\mu\partial_\mu\chi(x)\Psi'(x) &(3.4)\\ &= q\gamma^\mu A'_\mu\Psi'(x) &(3.5)\end{aligned}$$

which describes a particle in an electromagnetic field. For now the negative gradient of the scalar function $\chi(x)$ has been identified as being the transformed vector potential A'_μ. Obviously eqns. 3.3 and 3.5 are not equivalent and hence violate local gauge invariance. In order to regain it one has to replace the derivative ∂_μ with the covariant derivative D_μ defined by

$$D_\mu = \partial_\mu + iqA_\mu \tag{3.6}$$

As a result the vector potential now transforms in the following way:

$$A'_\mu = A_\mu - \partial_\mu\chi(x). \tag{3.7}$$

Inserting equation 3.6 into the Dirac equation 3.3 results in an expression that is invariant under local phase transformations, i. e.

$$(i\gamma^\mu D_\mu - m)\Psi(x) = 0 \tag{3.8}$$

gives

$$(i\gamma^\mu\partial_\mu - m)\Psi(x) = q\gamma^\mu A_\mu\Psi(x). \tag{3.9}$$

Comparing equation 3.5 with 3.9 it is obvious that both formulae are equivalent. That is, one can change one into the other by replacing $\Psi(x)$ and A_μ with the transformed versions $\Psi'(x)$ and A'_μ.

The above example is taken from QED where the vector field A_μ is massless. In general it is possible to achieve local gauge invariance for massless vector fields. The same is not true for massive vector fields. However, as will be shown in section 3.1.6 on page 46, one can conserve local gauge invariance by the mechanism of spontaneous symmetry breaking.

3.1.3 The Strong Interaction

The strong interaction of quarks and gluons is described by QCD. As was already mentioned on page 38, both quarks and gluons carry colour charge. Unlike leptons quarks don't occur as free, single particle. They are always bound in a two- or three-quark state. The systems consisting of a quark-anti-quark pair ($q\bar{q}$) are called mesons. Particles made up of three quarks (qqq) are called baryons. In both cases the net colour charge of the combined object is 'white' because the quarks in the meson carry colour and anti-colour and the constituents of the baryons are coloured red, green and blue. The fact that quarks only appear in colourless bound states explains why there are no free quarks as well as the absence of qq and $qqqq$ states. However particles with quark content $q\bar{q}q\bar{q}$ or $q\bar{q}qqq$ (pentaquarks) are allowed within this model and may have been observed recently by different experiments [32, 33, 34].

In QCD the Lagrange function for a free quark can be written as

$$\mathcal{L} = \bar{q}_j(i\gamma^\mu\partial_\mu - m)q_j \tag{3.10}$$

where th q_j ($j = 1, 2, 3$) are the quark colour fields. This Lagrangian has to be invariant under the non-abelian local gauge transformation $SU(3)_C$. It is given by

$$q_j(x) \rightarrow e^{\alpha^a(x)T_a}q_j(x). \tag{3.11}$$

The $\alpha^a(x)$ ($a = 1, ..., 8$) are the group -parameters and the T_a are the eight generators of the group. These generators are linearly independent, hermitian 3×3 matrices with trace equal to zero. They don't commute because of the non-abelian character of $SU(3)_C$. As in QED one has to introduce a covariant derivative to conserve local gauge invariance. In equation 3.10 ∂_μ has to be replaced by

$$D_\mu = \partial_\mu + ig_sT_aG^a_\mu. \tag{3.12}$$

Here g_s is a coupling constant and the G^a_μ represent gauge fields which are closely related to the eight gluons. Equation 3.12 leads to the following transformation of the G^a_μ:

$$G^a_\mu \rightarrow G^{a\prime}_\mu = G^a_\mu - \frac{1}{g_s}\partial_\mu\alpha^a - f^a_{bc}\alpha^bG^c_\mu. \tag{3.13}$$

Equation 3.13 and its QED equivalent 3.7 on the preceding page have the same form, except for the last term $f^a_{bc}\alpha^bG^c_\mu$, where the f^a_{bc} are the structure constants of $SU(3)_C$. It prevents the gauge fields from not commuting and hence express the non-abelian character of $SU(3)_C$.

The gauge invariant Lagrangian is now given by

$$\mathcal{L} = \bar{q}_j(i\gamma^\mu\partial_\mu - m)q_j - g(\bar{q}_j\gamma^\mu T_aq_j)G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{3.14}$$

where the last term has to be added to take into account gluon interactions and kinematics. The field tensor $G^a_{\mu\nu}$ is given by

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - gf^a_{bc}G^b_\mu G^c_\nu. \tag{3.15}$$

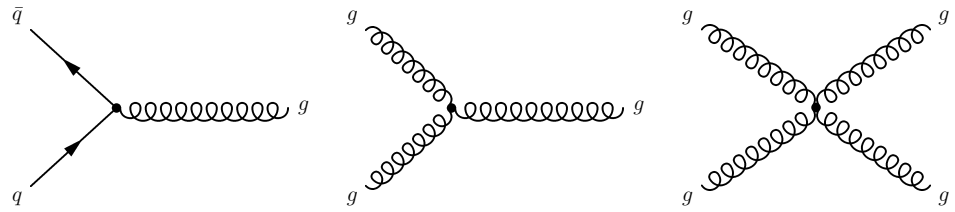


Figure 3.1: Gluon interactions in QCD. Shown are Feynman diagrams for gluon radiation off a quark (left), a triple-gluon vertex (middle) and a four-gluon vertex (right). The latter two are representing the gluon self interaction.

Equation 3.14 contains the self-interaction of gluons. The possible quark-gluon and gluon-gluon interactions are depicted in fig. 3.1. The three- and four-gluon vertices give rise to the aforementioned effect of quark confinement that forbids the existence of free colour-charged particles. They are also responsible for the running of the strong coupling constant α_s, that will be discussed in section 3.1.5 on page 44. One can rewrite eqn. 3.14 using symbolic shortcuts in order to clarify the meaning of the different terms. One gets

$$\mathcal{L} = q\bar{q} + g_s q\bar{q}G + G^2 + g_s G^3 + g_s G^4 \tag{3.16}$$

where the different expressions have the following meaning:

- $q\bar{q}$: kinematics of free quarks
- $g_s q\bar{q}G$: quark-gluon coupling
- G^2: kinematics of free gluons
- $g_s G^3$: triple-gluon vertex (gluon self interaction)
- $g_s G^4$: four-gluon vertex (gluon self interaction).

3.1.4 The Weak Interaction

In order to formulate a theory describing weak interactions, which are for instance responsible for radioactive decays, one has to consider some experimental results:

- The weak interaction violates parity, i. e. left-handed particles are preferred. This was shown for the first time by the Wu experiment [35].
- The weak interaction has a very short range.
- There are neutral and charged weak interactions. β-decays are a prominent example for the so called changed current reactions. Neutral currents have been discovered at CERN[4] about 30 years ago.

[4]Counseil Européenne pour la Recherche Nucléaire

In order to explain the finite range of the weak force one has to assume massive exchange particles. However, we will assume massless bosons for now and deal with this problem in section 3.1.6.

First a modified form of the phase transition (eqn. 3.2) is introduced:

$$\Psi'(x) = e^{i\frac{g}{2}\tau\beta(x)}\Psi(x). \tag{3.17}$$

Here g, the coupling of the weak interaction, has replaced the electric charge q and the Pauli matrices $\tau = (\tau_1, \tau_2, \tau_3)$ are introduced because the transformation should be unitary and hermitian with a trace equal to zero. The above equation represents a local $SU(2)$ phase transition. As in the previous sections one has to introduce a covariant derivative which, in the case of the weak force, has the following form:

$$D_\mu = \partial_\mu + i\frac{g}{2}W_\mu^a\tau_a. \tag{3.18}$$

As before, the weak field W^μ changes and transforms by applying the local phase transition. One gets

$$W_\mu^a \rightarrow W_\mu'^a = W_\mu^a - \partial_\mu\beta^a - g(\epsilon_{abc}\beta^b W_\mu^c), \tag{3.19}$$

being similar to the equations found in the previous sections. The term $-g(\epsilon_{abc}\beta^b W_\mu^c)$ is a cross-product and results from the non-abelian character of $SU(2)$. It prevents eqn. 3.19 from commuting.

In order to unify the weak and the electromagnetic interaction one needs a new group:

$$SU(2)_L \times U(1)_Y. \tag{3.20}$$

With this group fermions are divided in left-handed doublets and right handed singlets, as shown in table 3.1 on page 38. This is based on the fact that the charged current of the weak interaction is only coupling to left-handed particles. The singlets do not take part in charged weak interactions. The chirality, indicated by L and R in table 3.1, is defined by the eigenvalues of the two projection operators P_L and P_R defined by

$$P_L = \frac{1-\gamma^5}{2} \tag{3.21}$$

$$P_R = \frac{1+\gamma^5}{2}. \tag{3.22}$$

The electromagnetic interaction does not distinguish between left and right handed particles.

As can be seen from equation 3.20, $U(1)_Y$ is used for the product with $SU(2)_L$ and not $U(1)_{QED}$. The use of $U(1)_{QED}$ would lead to charged leptons and neutrinos having the same electric charge, in contradiction to experimental observations. To avoid this problem Q is replaced by another quantity called hypercharge Y. The relation between Q (the generator of $U(1)_{QED}$) and Y is given by the Gell-Mann-Nishijima law

$$Q = I_3 + \frac{Y}{2} \tag{3.23}$$

where I_3 (the generator of $SU(2)_L$) is the third component of the weak isospin.

Because of the new groups $SU(2)_L \times U(1)_Y$ one has to modify equation 3.18 and add the term $i\frac{g'}{2}B_\mu Y$ leading to

$$D_\mu = \partial_\mu + i\frac{g'}{2}B_\mu Y + i\frac{g}{2}W_\mu^a \tau_a. \tag{3.24}$$

Here the coupling constants for $SU(2)_L$ and $U(1)_Y$ are given by g and g' respectively. The gauge invariant Lagrangian of the weak interaction is now given by

$$\begin{aligned} \mathcal{L} &= \bar{\Psi}\gamma^\mu(i\partial_\mu - \frac{g}{2}W_\mu^a\tau_a - g'\frac{Y}{2}B_\mu)\Psi & (3.25)\\ &\quad -\frac{1}{4}W_a^{\mu\nu}W_{\mu\nu}^a - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} & (3.26)\end{aligned}$$

The first part (3.25) describes the kinematics of fermions and their coupling to gauge bosons. The second part (3.26) describes the kinematics of bosons. The field strength tensors $W_{\mu\nu}^a$ and $B_{\mu\nu}$ are given by

$$\begin{aligned} W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g\epsilon_{abc}W_\mu^b W_n u^c, & (3.27)\\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. & (3.28)\end{aligned}$$

From experiments one knows that charged $W^\pm$ bosons and neutral Z^0 bosons are mediating the weak interaction while the neutral photon is the field quantum of QED. These particles are linear combinations of the W_μ^a and B_μ vector fields of $SU(2)_L \times U(1)_Y$ and hence show the close relation between the weak interaction and QED. They are given by

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \tag{3.29}$$

and

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \tag{3.30}$$

where θ_W is the Weinberg angle and A_μ represents the photon field. The sine and cosine of θ_W are related to g and g' via

$$\begin{aligned} \sin\theta_W &= \frac{g'}{\sqrt{g^2+g'^2}}, & (3.31)\\ \cos\theta_W &= \frac{g}{\sqrt{g^2+g'^2}}. & (3.32)\end{aligned}$$

As one can see from the matrix in eqn. 3.30, A_μ and Z_μ are orthogonal states.

Quark mixing

Experimental observations have shown that the weak eigenstates of the down-type quarks (d, s, b) are not the same as the mass eigenstates. Charged weak currents do not conserve quarks flavour because the weak eigenstates are linear combinations of the mass eigenstates. Neural currents, on the other hand, do conserve flavour. The weak and mass

$\lvert V_{ud}\rvert = 0.9742 - 0.9757$	$\lvert V_{us}\rvert = 0.2190 - 0.2260$	$\lvert V_{ub}\rvert = 0.0020 - 0.0050$
$\lvert V_{cd}\rvert = 0.2190 - 0.2250$	$\lvert V_{cs}\rvert = 0.9734 - 0.9749$	$\lvert V_{cb}\rvert = 0.0370 - 0.0430$
$\lvert V_{td}\rvert = 0.0040 - 0.0140$	$\lvert V_{ts}\rvert = 0.0350 - 0.0430$	$\lvert V_{tb}\rvert = 0.9990 - 0.9993$

Table 3.3: 95% confidence level limits for the absolute values of the CKM matrix elements [29].

eigenstates of the d-, s- and b-quarks are connected via the Cabbibo-Kobayashi-Maskawa matrix V_{CKM}:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{L.} \tag{3.33}$$

The matrix elements of the CKM matrix are complex numbers, however flavour mixing only depends on the absolute value of the matrix elements. The non-zero phase in some of the off-diagonal elements gives rise to CP violating decays. Experimentally the effect of flavour mixing can be observed, for instance, in the decay of B-mesons. These would be stable in the absence of mixed flavour states. However the CKM matrix allows the decay chain

$$b \to c \to s \to u. \tag{3.34}$$

Because the matrix element V_{cb} is small compared to the on-diagonal elements the decay $b \to c$ is suppressed and leads to a relatively long lifetime of the B-meson. This long lifetime is useful for the experimental identification of B-jets because it gives rise to so called secondary vertices. The matrix elements of V_{CKM} can not be calculated from the Standard Model. They have to be measured experimentally. The current 95% confidence level limits for the absolute values of the matrix elements are given in table 3.3. Because of the unitarity of the CKM matrix, the matrix elements are not independent. For a complete description three angles and one phase are sufficient.

3.1.5 Running coupling constants

The coupling constants of the Standard Model are not really constant, as one would expect from the nomenclature, but depend on the energy or 4-momentum transfer Q^2. This behaviour is caused by vacuum polarisation, i. e. the fact that the vacuum is filled with virtual pairs of particles and anti-particles. In the presence of a field these particles orient themselves like dipoles and either decrease (screening) or increase (anti-screening) the effective charge. These effects can be described by radiative corrections shown in fig. 3.2 and 3.3 on the next page. Because the integrals that have to be solved are divergent one uses the technique of renormalisation, i. e. the divergences are absorbed in the renormalised quantities like the electric charge. Hence one has an effective charge e depending on Q^2 and the bare charge e_0. The relation between the two is given by

$$e(Q^2) = e_0^2(1 + \frac{e_0^2}{12\pi^2} ln \frac{|Q^2|}{m^2}). \tag{3.35}$$

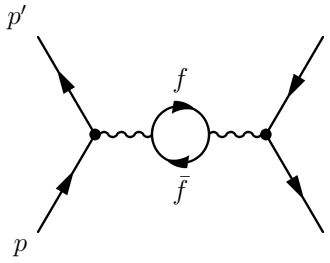


Figure 3.2: Feynman graph for QED vacuum polarisation in lowest order. The photon splits into a charged fermion-antifermion pair thereby reducing the effective charge between the scattered particles. The 4-momentum transfer Q is given by the difference $Q = p - p'$.

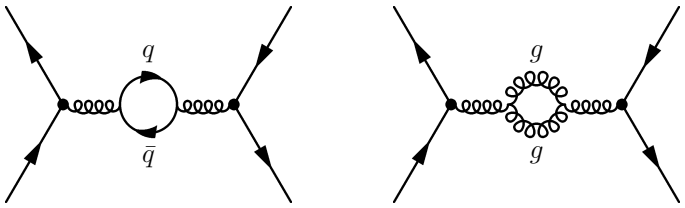


Figure 3.3: Feynman graphs for QCD vacuum polarisation. The gluon splits either into a $q\bar{q}$ pair (left) or a gluon-gluon pair (right).

This Q^2-dependence of the charge can also be interpreted as a Q^2-dependence of α:

$$\alpha(Q^2) = \frac{\alpha}{1 - \frac{\alpha}{3\pi} ln \frac{Q^2}{m_e^2}}. \tag{3.36}$$

In QED screening causes the fine-structure constant α to decrease with Q^2. Only fermion-antifermion pairs contribute to the radiative corrections. There is no self-coupling between the photons. Fig. 3.2 shows the corresponding Feynman diagram. The Q^2 dependence of α is given by

$$\alpha(Q^2) = \frac{\alpha}{1 - \frac{\alpha}{3\pi} ln \frac{Q^2}{m_e^2}} \tag{3.37}$$

with

$$\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137.04}. \tag{3.38}$$

At LEP[5] with CMS[6]-energies of 90 GeV, α decreases by $\sim 6\%$ and is given by $\sim \frac{1}{128}$.

In QCD the situation is more complicated. The colour charge of the quarks causes a screening effect as in QED. The corresponding dipoles consist of $q\bar{q}$ pairs. But the first effect is dominated by anti-screening. It is caused by the fact that the gluons themselves carry colour charge. They self-interact with each other. This reduces the effective colour

[5]Large Electron Positron collider
[6]Center of Mass System

charge the closer one gets to the true charge since less gluons contribute to the observed charge. The effects of quark confinement and asymptotic freedom are consequences of this behaviour. Quark confinement describes the fact that quarks can not be observed as free particles. They are confined in a bound states with other quarks because α_s increases with increasing distance. On the other hand bound quarks behave like free particles. Due to the short distances among them α_s is small. The Q^2 dependence of α_s is given by

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \frac{\alpha_s(\mu_R^2)(33-2N_f)}{12\pi} ln \frac{|Q^2|}{\mu_R^2}} \tag{3.39}$$

where n_f is the number of quark flavours and μ_R is the renormalisation scale[7]. This equation allows to transform α_s measurements made at $|Q^2| = \mu_R^2$ to other values of $|Q^2|$. To compare α_s values from experimental measurements it is commonplace to quote $\alpha_s(M_Z)$. Often the value of $\sqrt{|Q^2|}$ where the denominator of eqn. 3.39 vanishes is called Λ, the QCD cutoff-parameter. Using this definition, eqn. 3.39 becomes

$$\alpha_s(Q^2) = \frac{4\pi}{(11 - \frac{2N_f}{3}) ln \frac{Q^2}{\Lambda^2}}. \tag{3.40}$$

From this relation one can see the effect of asymptotic freedom since

$$\lim_{|Q^2| \to \infty} \alpha_s(|Q^2|) = 0 \tag{3.41}$$

as long as $N_f \leq 16$, i. e. the number of quark flavours is limited to 16. For the process of quark-quark scattering high values of Q^2 are equivalent to small distances. At first sight it looks like eqn. 3.40 can describe quark confinement as well since $\alpha_s \to \infty$ for $|Q^2| \to \Lambda^2$. However it was obtained using perturbation theory and is only valid for α_s values smaller than 1.

Another consequence of renormalisation is the running of the quark masses. Like α_s, the quark mass depends on Q^2, i. e. $m = m(Q^2)$. Usually the pole mass[8] is referred to as the physical mass. The running masses decrease with Q^2. Since the coupling constants have different values and Q^2 dependence one could imagine a unification at high Q^2, i. e. all coupling constants become equally strong. Unfortunately this unification is not possible within the standard model, as shown in fig. 3.4 on the next page, but becomes possible in the MSSM[9].

3.1.6 The Higgs Mechanism

As has already been mentioned in section 3.1.2, local gauge invariance for massive vector fields can be achieved by using the principle of spontaneous symmetry breaking[10]. One assumes all particles as being massless in the beginning. Mass is generated by introducing

[7] Usually this is given by the momentum transfer of the scattering experiment.

[8] The mass at the pole of the quark propagator.

[9] Minimal Supersymmetric Standard Model

[10] This mechanism was originally developed in solid state physics as a relativistic extension of the Ginzburg-Landau-Model of superconductivity [37].

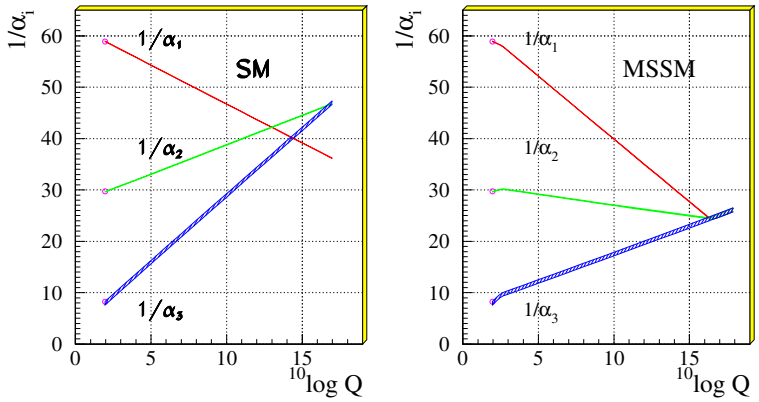


Figure 3.4: Running of the coupling constants. The coupling constants $\alpha_1 = g$, $\alpha_2 = \alpha$ and $\alpha_3 = \alpha_s$ should unify for large values of Q^2. In the SM (left) the unification is not possible but can be achieved in the MSSM (right), a supersymmetric extension of the Standard Model [36].

an additional scalar background field. The interaction of the massless particles with the field gives mass to the gauge bosons. This field is called the Higgs field [8].

The Higgs field is chosen to be a scalar, complex field with two components,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1^+ + i\phi_2^+ \\ \phi_1^0 + i\phi_2^0 \end{pmatrix} \tag{3.42}$$

with $I_3 = \frac{1}{2}$ and $Y = 1$, leading to a neutral field as one can see from eqn. 3.23 on page 42. ϕ^+ and ϕ^0 represent a charged and a neutral component respectively. The $SU(2) \times U(1)$ invariant Lagrangian is given by

$$\mathcal{L}_{Higgs} = (\partial_\mu \Phi)^\dagger \partial^\mu \Phi - V(\Phi^\dagger \Phi). \tag{3.43}$$

It contains only even terms to preserve the symmetry w. r. t. $\Phi = 0$[11]. The potential V is given by

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \tag{3.44}$$

Depending on the choice of μ^2 one either gets a solution with non-trivial minima or one with a minimum at $\Phi = 0$, as shown in fig. 3.5 on the next page. For $\mu^2 < 0$, V has minima at

$$|\Phi| = \frac{v}{\sqrt{2}} \tag{3.45}$$

with

$$v = \frac{\mu}{\lambda} \tag{3.46}$$

and an arbitrary phase $\theta = 0$. By fixing the phase to a certain value the symmetry of the system, the phase invariance, is broken. This is called spontaneous symmetry breaking. Since the vacuum is neutral one can set the first component in eqn. 3.42 to zero and gets

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{3.47}$$

[11] Terms of order $(\Phi^\dagger \Phi)^4$ are not included because they would lead to renormalisation problems.

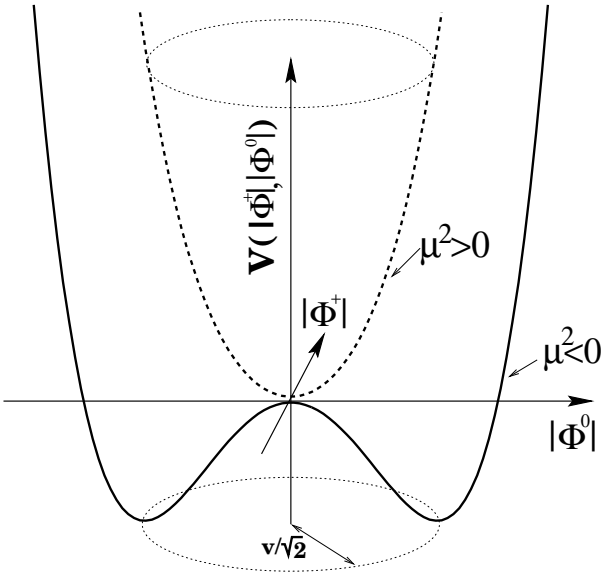


Figure 3.5: The Higgs potential $V(\Phi)$ for $\mu^2 > 0$ and $\mu^2 < 0$ [38]. For $\mu^2 > 0$, the symmetry is unbroken and V only has the trivial minimum at zero. For $\mu^2 < 0$ the symmetry is broken and the potential is minimal if $|\Phi^+|^2 + |\Phi^0|^2 = \frac{v^2}{2}$. The shape of V now resembles that of a mexican hat.

for the minimum of V. Fluctuations around the minimum can be parametrised via

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta + i\zeta \end{pmatrix}. \tag{3.48}$$

Inserting this into eqn.3.43 one gets

$$V(\phi\phi) = \mu^2 \left[-\frac{1}{4}v^2 + \eta^2 + \frac{\eta\zeta^2}{v} + \frac{\eta^2\zeta^2}{2v^2} + \mathcal{O}(\eta^2) + \mathcal{O}(\eta^4) \right] \tag{3.49}$$

For the following calculations it is sufficient to take only the two first terms. Inserting the above potential and 3.48 into the Lagrangian 3.43 one obtains for the leading terms

$$\mathcal{L} = \left[\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2 \right] + \frac{1}{2}(\partial_\mu \zeta)(\partial^\mu \zeta) + ... \tag{3.50}$$

The first term in brackets represents the Klein-Gordon equation for a massive scalar particle η of mass $m_\eta = \sqrt{2}\mu$. One identifies this as being the Higgs particle. The second term describes the massless scalar particle ζ, the Goldstone boson.

Gauge Boson masses

To give mass to gauge bosons an extra term, the gauge field Lagrangian $\mathcal{L}_{gauge}$, has to be added to eqn. 3.43, i. e.

$$\mathcal{L} = \mathcal{L}_{Higgs} + \mathcal{L}_{gauge} \tag{3.51}$$

with $\mathcal{L}_{gauge}$ being

$$\mathcal{L}_{gauge} = -\frac{1}{4}W^i_{\mu\nu}W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{3.52}$$

where W and B are the field strength tensors defined in eqn. 3.27 and 3.28 on page 43. In order to conserve gauge invariance one uses the covariant derivative 3.6 on page 39 and inserts it into the Lagrangian 3.51. The kinetic part of the gauge invariant Lagrangian is now given by

$$\mathcal{L}_{kin} = \frac{1}{4}[(gW^i_\mu\tau_i + g'B_\mu)\phi]^\dagger[(gW^{\mu i}\tau_i + g'B^\mu)\Phi]. \tag{3.53}$$

Replacing W and B with the physical fields A_μ, Z_μ and $W^\pm_\mu$ one finally gets

$$\begin{aligned}
\mathcal{L} &= [\frac{1}{2}(\partial^\mu\eta)(\partial_\mu\eta) - \mu^2\eta^2] && (3.54)\\
&\quad -\frac{1}{4}W^i_{\mu\nu}W^{i\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} && (3.55)\\
&\quad +\frac{1}{2}\frac{g^2\nu^2}{4}((W^+_\mu)^2 + (W^-_\mu)^2) && (3.56)\\
&\quad +\frac{1}{2}\frac{g^2\nu^2}{4\cos^2\theta_W}Z^2_\mu. && (3.57)
\end{aligned}$$

The first term (3.54) is the already known Klein-Gordon equation for a massive Higgs particle with mass $m_H = \sqrt{2}\mu$. Equation 3.55 is the gauge field Lagrangian $\mathcal{L}_{gauge}$, and eqns. 3.56 and 3.57 are mass terms. The masses of the Z^0 and $W^\pm$ bosons can now be identified as being

$$m_W = \frac{g\nu}{2} \tag{3.58}$$

and

$$\begin{aligned}
m_Z &= \frac{g\nu}{2\cos\theta_W} && (3.59)\\
&= \frac{m_W}{\cos\theta_W}. && (3.60)
\end{aligned}$$

The photon field A_μ does not acquire a mass term:

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}}(gB_\mu + g'W_{3\mu}). \tag{3.61}$$

The vacuum expectation value of the Higgs field v can not be calculated within the Standard Model. However it can be obtained from measuring the Fermi constant G_F and is given by

$$v = \frac{1}{\sqrt{2}\sqrt{G_F}} \approx 246\,\text{GeV}. \tag{3.62}$$

Giving mass to Fermions

In order to give mass to fermions one introduces new couplings to the Standard Model, the Yukawa couplings g^f_Y, which are different for every charged lepton and quark [39]. The corresponding Lagrangian (describing electrons in this case) is given by

$$\mathcal{L}_{Yukawa} = -g^e_Y[\bar{e}_R\Phi^\dagger\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L + \begin{pmatrix}\bar{\nu}_{\bar{e}}\\ e\end{pmatrix}_L\Phi e_R]. \tag{3.63}$$

Inserting the vacuum expectation value for Φ, eqn. 3.47, one obtains

$$\mathcal{L}_{Yukawa} = -\frac{g_Y^e}{\sqrt{2}}[\bar{e}_R(0,\nu)\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L + \begin{pmatrix}\bar{\nu}_e\\ \bar{e}\end{pmatrix}_L\begin{pmatrix}0\\ \nu\end{pmatrix}e_R] \tag{3.64}$$

$$= -\frac{g_Y^e}{\sqrt{2}}[\bar{e}_L e_R + \bar{e}_R e_L] \tag{3.65}$$

$$= -\frac{g_Y^e}{\sqrt{2}}\bar{e}e \tag{3.66}$$

where $\frac{g_Y^e}{\sqrt{2}}$ has to be identified as being the electron mass m_e. The Yukawa couplings for all fermions and quarks have to be chosen such that their masses agree with experimental measurements. Because the fermion masses are proportional to g_Y^f, the Higgs preferably couples to heavy fermions. Neutrinos are massless in the model. They do not appear in $\mathcal{L}_{Yukawa}$ because the charged part of the vacuum expectation value in eqn. 3.47 is zero[12].

The Higgs Boson mass

As has been shown in the previous section, the mass of the Higgs boson is given by $m_H = \sqrt{2}\mu$. Unfortunately the parameter μ can not be calculated in the SM and hence its exact value is unknown. The mass can only be estimated by theoretical assumptions and experimental searches so far.

[12]The fact that neutrino experiments have found evidence for neutrino oscillations is a clear sign for neutrinos being massive particles [40]. However they are still considered massless in the Standard Model.

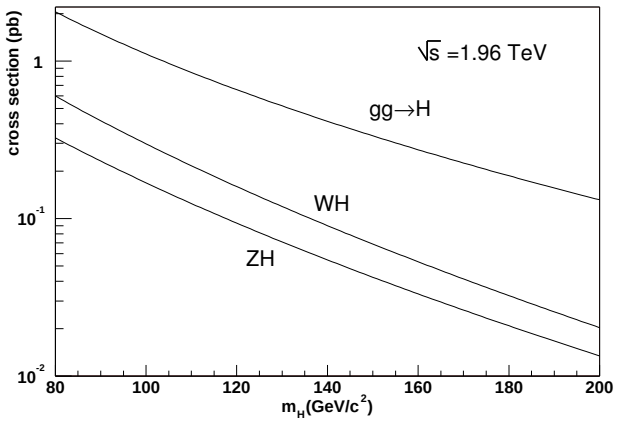


Figure 3.6: Higgs production cross sections at NLO for gluon fusion and associated production at 1.96 TeV.

3.2 Higgs production and decay at the Tevatron

3.2.1 Production Processes

The most important processes of Higgs production at the Tevatron are the associated production in conjunction with a W or Z boson[13] and the process of gluon-gluon fusion. The NLO[14] cross-sections for these processes are shown in fig. 3.6[15] and LO[16] Feynman diagrams are depicted in fig. 3.7.

Gluon-gluon fusion has the largest cross-section and takes place via a one-loop graph where the Higgs couples to a virtual quark pair. The triangle loop is dominated by the exchange of t- and b-quarks[17]. Cross-sections are in the range 1.0 - 0.1 pb for Higgs masses between 100 GeV/c^2 and 200 GeV/c^2.

In the WH and ZH reactions a $q\bar{q}$-pair produces a W or Z boson followed by the emission of a Higgs particle. The cross-section for $q\bar{q} \to WH$ is in the range 0.3 - 0.03 pb for 100 GeV/c^2 $\leq m_H \leq$ 200 GeV/c^2. For $q\bar{q} \to ZH$ the cross-section in the same mass range is lower by a factor of about two.

[13]Higgs production at the Tevatron via the WH and ZH processed was discussed for the first time in [41].

[14]Next-to-Leading Order

[15]Plots were made with HIGLU (v2.1), HDECAY (v3.0) [42] and V2HV. The source code for all programs is available from [43]. The input parameters are documented in appendix A on page 125. Tables with and NLO cross-sections, branching fractions and the total decay width from 80 GeV/c^2 to 200 GeV/c^2 can be found there as well.

[16]Leading Order

[17]This is also the case at the Large Hadron Collider (LHC).

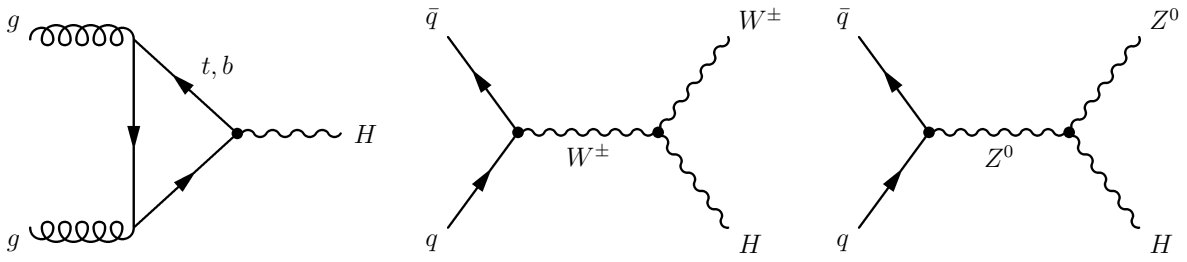


Figure 3.7: Feynman diagrams for the Higgs production processes $gg \to H$, WH and ZH.

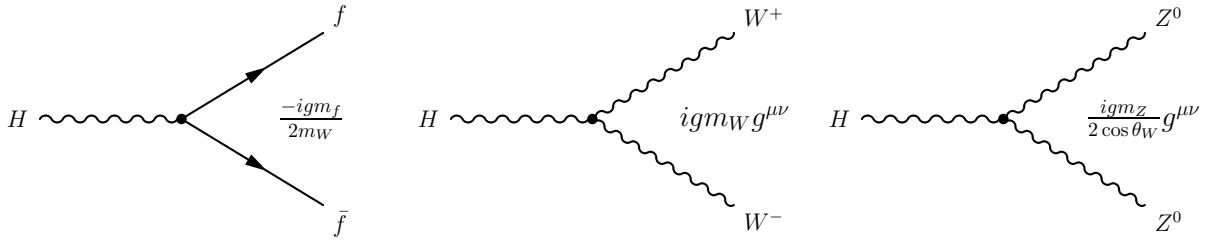


Figure 3.8: Higgs boson decays at tree level and their couplings.

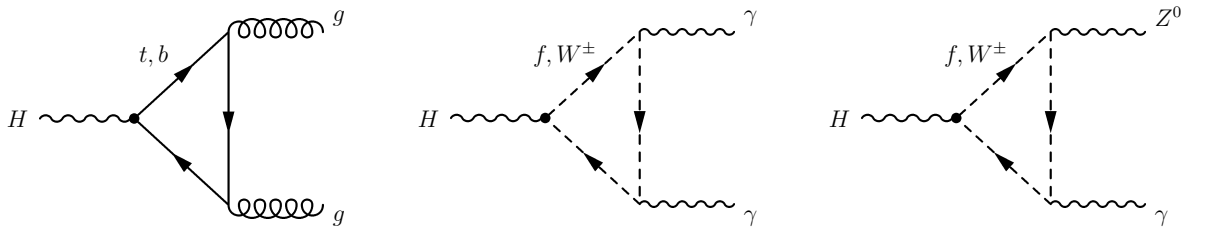


Figure 3.9: Higgs decays via loop diagrams.

3.2.2 Higgs Decays

The branching ratios of the dominant decay modes of the Standard Model Higgs as function of its mass are shown in fig. 3.10[18]. Feynman diagrams for all those processes are given in figs. 3.8 and 3.9. Fig. 3.11 shows the total decay width.

$H \to f\bar{f}$

The dominant decay mode for Higgs masses below roughly 140 GeV/c^2 is $H \to b\bar{b}$, the decay into a pair of b-quarks. For 100 GeV/c^2 $\leq m_H \leq$ 200 GeV/c^2 it is in the range of 80% to 35%. The reason is that the coupling is proportional to the particle mass m as can be seen from fig. 3.8. The decay into top-quark pairs $H \to t\bar{t}$ is not possible in the Higgs mass range discussed here. Other branching fractions in fig. 3.10, where the Higgs decays into fermion pairs are $H \to \tau\tau$ and $H \to c\bar{c}$.

$H \to gg$

This decay is of the same order of magnitude in the mass region of interest. The reason is the large top Yukawa coupling and the colour factor. The partial width is of interest because it determines the production cross-section for the process $gg \to H$ discussed on page 51.

[18]C.f. footnote 15 on page 51.

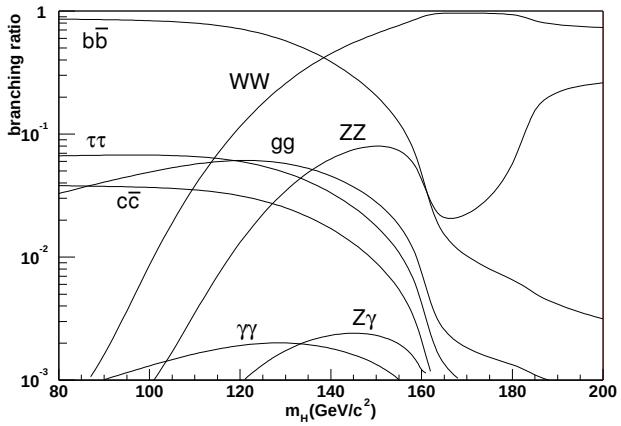


Figure 3.10: Branching ratios of the SM Higgs boson.

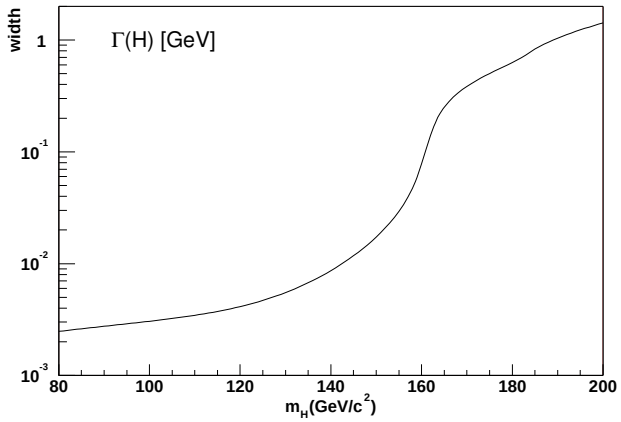


Figure 3.11: Total decay width of the Higgs boson in GeV.

$H \to WW, ZZ$

For Higgs masses above roughly 140 GeV/c^2 the decay $H \to WW$ is the dominant one. At least one of the W-bosons is produced off-shell. The decay $H \to ZZ$ is less important because the branching ratio is much smaller (cf. fig. 3.10).

$H \to Z\gamma, \gamma\gamma$

The partial widths for these two processes are the smallest of all processes discussed so far. Feynman diagrams are shown in fig. 3.9 where either a charged fermion or a W-boson runs in the triangle loop. The Higgs can not couple to the photon directly because it is neutral.

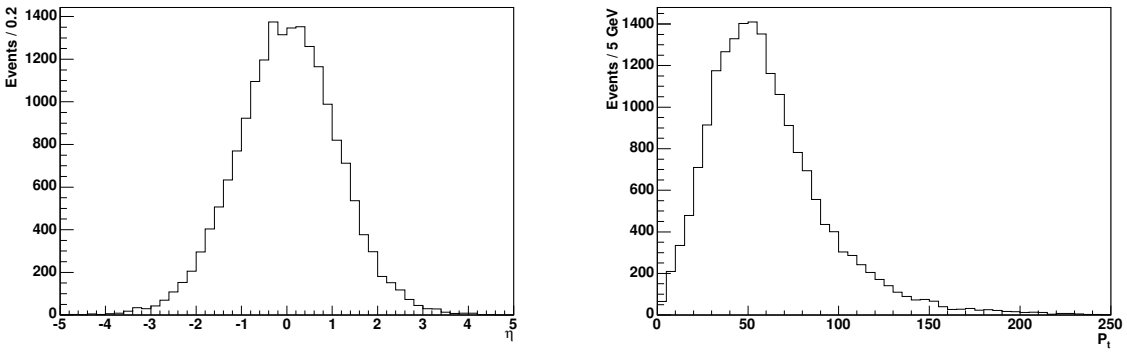


Figure 3.12: Generator level plots of η (left) and p_t (right) for b-quarks originating from H decays. The Higgs mass was set to $m_H = 120$ GeV/c^2.

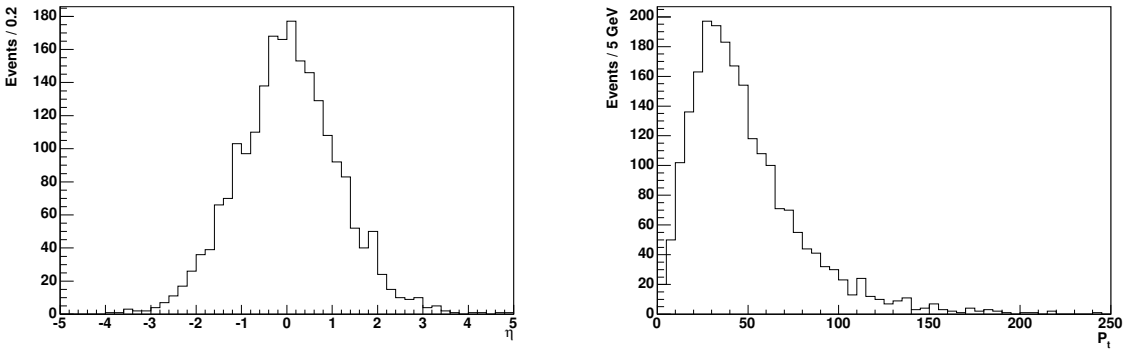


Figure 3.13: Generator level plots of η (left) and p_t (right) for electrons and muons originating from the W decays in $WH \to l\nu b\bar{b}$. The Higgs mass was set to $m_H = 120$ GeV/c^2.

3.2.3 The channel $WH \to l\nu b\bar{b}$

The most promising channel for low mass Higgs searches (i. e. $m_H \leq 140$ GeV/c^2) at the Tevatron is given by $q\bar{q} \to WH \to l\nu b\bar{b}$. A quark-antiquark-pair produces at W-boson which radiates a Higgs boson. The W then decays into a lepton and a neutrino and the Higgs goes to $b\bar{b}$. The Feynman diagram for the process is shown in fig. 3.14 on page 58 in the top left corner. The signature of the signal events can be read off the graph. One expects two jets, a high-p_t lepton and missing energy from the neutrino. Both the b-jets and the leptons are expected to lie predominantly within $|\eta| \leq 2.0$, i. e. in the central region of the CDF detector, as can be seen from figs. 3.12 and 3.13.

This channel is the subject of the study presented in the following chapters.

Since the production process $gg \to H$ as a much larger cross-section one might be tempted to look at the channel $gg \to H \to b\bar{b}$. Unfortunately this is not promising because of the large $b\bar{b}$ background that can be expected from QCD processes.

3.2.4 The channel $ZH \to ll\,b\bar{b}$

Already the channel $ZH \to ll\,b\bar{b}$ is by far less attractive to look at when compared to $WH \to l\nu b\bar{b}$. First, the NLO cross-section for the ZH process is almost a factor two lower than σ_{WH}, namely

$$\begin{aligned} \sigma_{WH,NLO} &= 0.1597 && (3.67)\\ \sigma_{ZH,NLO} &= 0.0936 && (3.68)\\ & && (3.69) \end{aligned}$$

for $m_H = 120$ GeV/c^2. Also the branching fractions for $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ are roughly a factor three lower compared to $W \to e\nu$ and $W \to \mu\nu$. The numbers from the PDG[19] [44] are

$$\begin{aligned} BR(Z \to e^+e^-) &= 0.033 && (3.70)\\ BR(Z \to \mu\mu) &= 0.033 && (3.71)\\ BR(W \to e\nu) &= 0.1056 && (3.72)\\ BR(W \to \mu\nu) &= 0.105\ . && (3.73)\\ & && (3.74) \end{aligned}$$

With an electron identification efficiency of $\epsilon \approx 0.85$ one finds that the ratio of $\sigma \times BR$ for WH and ZH is about 45:

$$\begin{aligned} & \frac{\sigma_{WH} \cdot BR(W \to e,\mu)}{\sigma_{ZH} \cdot BR(Z \to e,\mu) \cdot (1-\epsilon)} && (3.75)\\ = & \frac{0.1597 \cdot 0.21}{0.0936 \cdot 0.066 \cdot (1-0.85)} && (3.76)\\ = & \ 46.15\ . && (3.77) \end{aligned}$$

Hence the ZH channel where the Z decays leptonically and the H goes to b-quarks has not been looked at in this study.

[19] Particle Data Group

3.3 Background processes

Example diagrams for the background processes considered in this study are given in figs. 3.14 and 3.15. As can be seen from these figures almost all backgrounds can be expected to have the same signature as the signal itself. That is, a lepton from W-decay, one or more real b-quarks and missing energy from neutrinos. Even worse, the s-channel single-top process, $Wb\bar{b}$ and WZ can have exactly the same final state $l\nu\, b\bar{b}$. Hence one can expect the discrimination between events originating from the signal and the background to be a difficult task.

Top production

The top-quark almost always decays into b and W since $V_{tb} \approx 1$. Leptonic decays of the W-boson can provide a lepton and neutrino. In all reactions at least two b-quarks are produced as well.

W + heavy flavour

In the $W + b\bar{b}, c\bar{c}$ reaction a W is produced from a $q\bar{q}$-pair. In addition one of the incoming quarks radiates off a gluon that splits into a $b\bar{b}$ or $c\bar{c}$ quark pair. The c-quarks can accidentally be tagged as b-jets and pass the event selection. The same statements also holds for the $W + c$ process.

QCD

QCD processes have to be considered because of their large cross-section. The probability that a QCD multi-jet event contains b-quarks is small. However the high cross-section for QDC can compensate this and contribute to the background expectation.

Dibosons

The processes WW, WZ and ZZ yield final states with leptons and jets. In all cases one of the bosons decays leptonically and one undergoes a hadronic decay. The ZZ does not have a neutrino in the final state. However, if one of the leptons is not reconstructed in the detector, the signature is consistent with the signal process.

$Z \to \tau\tau$

This process has to be considered because a hadronically decaying tau can satisfy the jet requirements of the analysis and a tau decaying to a lepton can survive the lepton cuts. Missing energy is caused by several neutrinos.

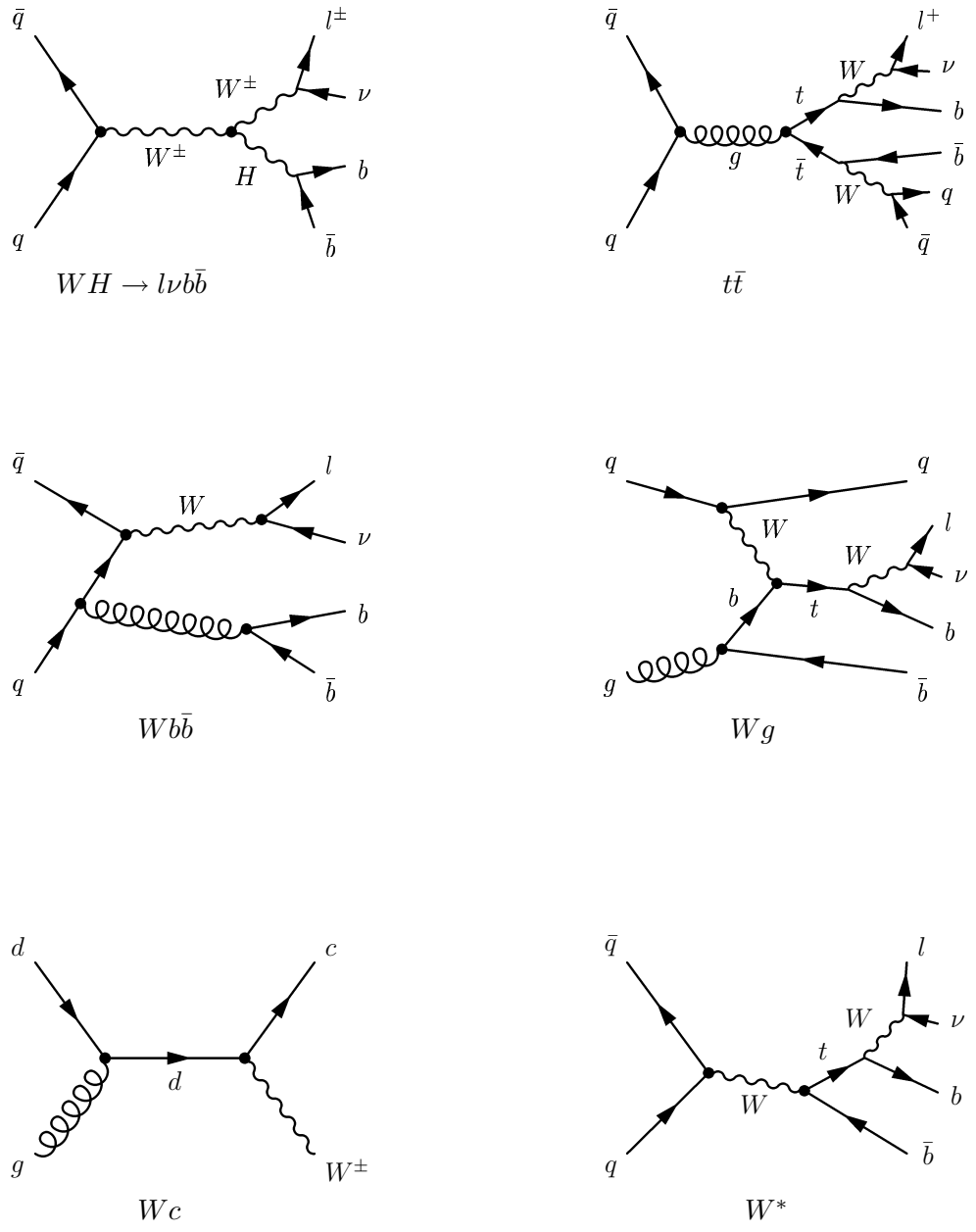


Figure 3.14: Feynman diagrams for the signal (top left) and background processes. Graphs for top-pair and single-top production are shown on the right. W^* denotes the s-channel and Wg the t-channel of the single-top process.

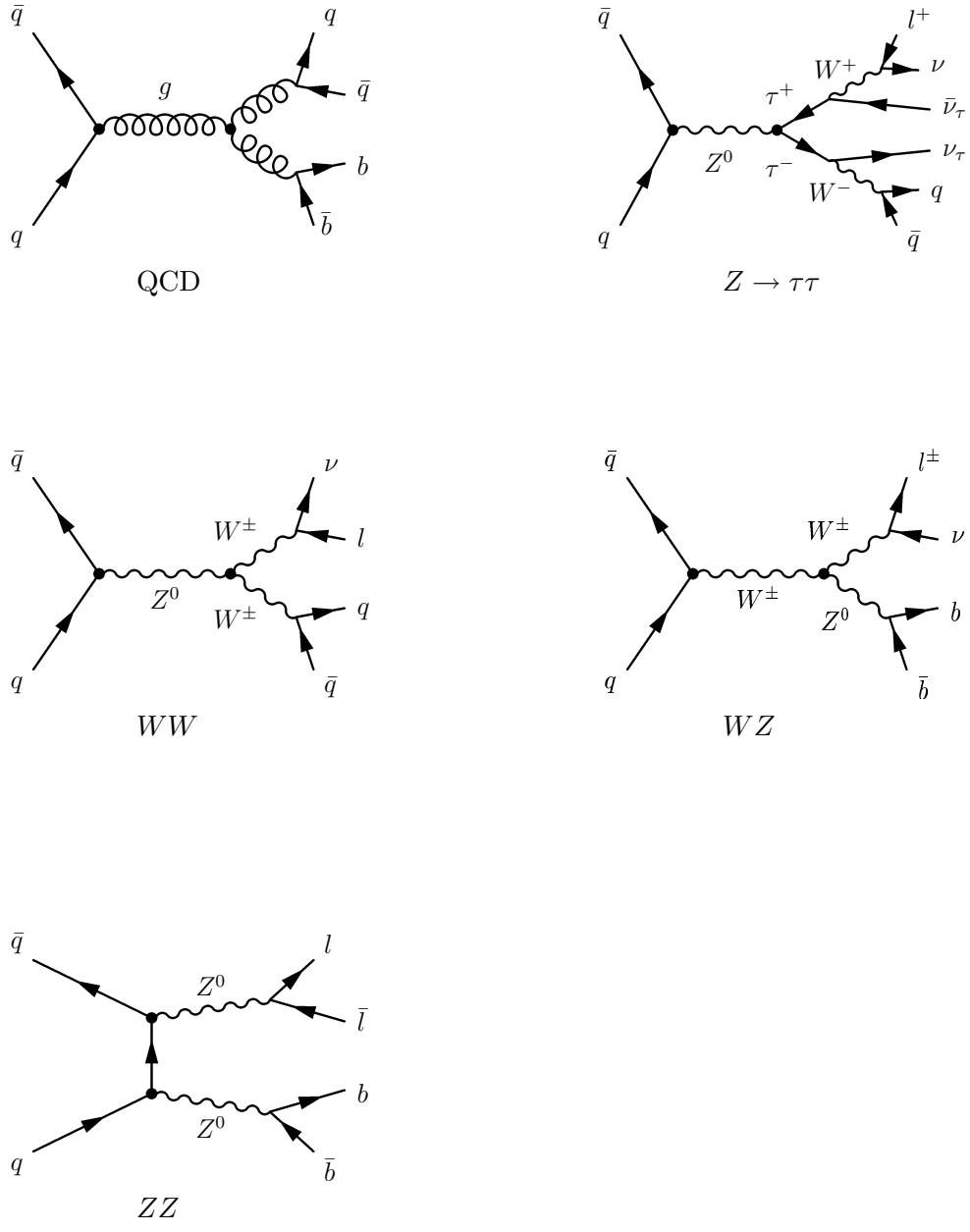


Figure 3.15: Example Feynman diagrams for a QCD process and $Z \to \tau\tau$ as well as the di-boson production processes WW, WZ and ZZ.

'There is nothing particularly
magical about multi-layer
neural networks.'
Pattern Classification
R.O. Duda et al.

Chapter 4

Neural Networks

This chapter describes neural networks and the underlying theoretical concepts. After presenting network basics in chapter 4.1 and network training in chapter 4.2, some advanced methods of neural computation are explained in sections 4.3 and 4.4. The last part on page 78 gives a short overview on the NeuroBayes® package. A more detailed description of neural networks can be found in various textbooks [45, 46, 47, 48].

4.1 Modelling neurons

In the recent past there has been a growing interest in neural networks[1] (NN). They are being successfully applied in in a wide range of scientific disciplines such as finance, medicine, engineering and physics. One can attribute their success to a few key factors such as modelling extremely complex non-linear functions and their ease of use. Neural networks grew out of research in artificial intelligence when people attempted to mimic the fault-tolerance and capacity to learn of biological systems with machines. It was felt that these functions could only be modelled by copying the functionality of the low level structures of the brain.

4.1.1 Biological Neurons

The basic building block of the brain is called neuron. About 10^{10} neurons can be found in the human brain. They are massively interconnected with an average of several thousand connections per neuron. A neuron is a specialised nerve cell that propagates electrochemical signals. It consists of the cell body or soma, small fibres called dendrites and a single long fibre extending from the cell body, the axon. Fig. 4.1 on the next page shows a schematic drawing of a typical neuron. The dendrites are fibres to which other neurons connect via synapses and are transferring the incoming electrical signals from other nerve cells to the cell body. The dendrites are not directly attached to other neurons.

The axon represents the outgoing fibre that branches into strands and sub-strands. At the end of these are the transmitting ends called synapses. The receiving ends of these

[1] Also called Artificial Neural Networks (ANN).

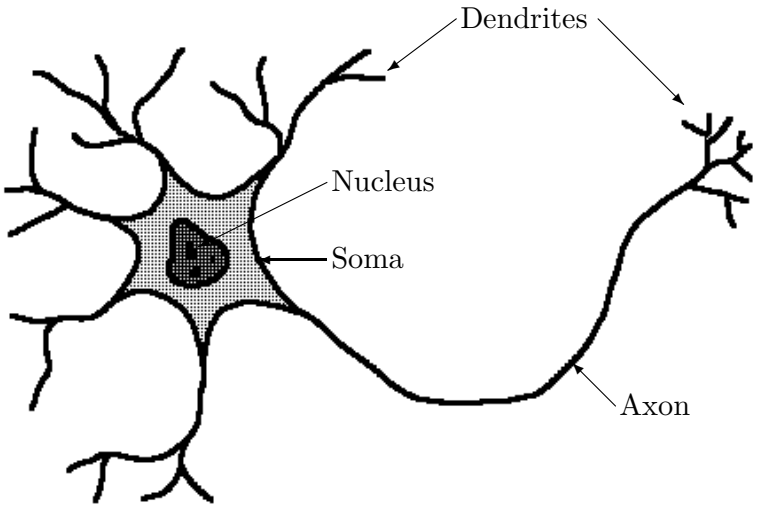


Figure 4.1: Schematic picture of a single neuron. The cell body or soma contains the nucleus. The dendrites directly connected to the soma receive electrochemical signal from other nerve cells. The axon, splitting up into sub-branches, transmits the output signal to other neurons [49].

junctions on other cells can be found both on the dendrites and on the cell bodies themselves. There is no direct connection between neurons. Electrical signals are passed via chemicals called neuro-transmitters. The neuro-transmitters raise or lower the electrical potential of the neuron. Once the potential exceeds a certain threshold an electrical pulse of fixed duration and strength is send to the axon and branches out to other cells. After firing, the neuron cannot be activated again until the refracting time has elapsed.

4.1.2 The Mathematical Model

In order to simulate neural systems on a computer one has to model mathematically the basic function of a single biological neuron. These building blocks for artificial neural networks are called nodes or units, as well as neurons[2]. The basic functionality of a single node is:

- It receives a number of inputs. Each input comes via a connection that has a strength or weight corresponding to synaptic efficacy. The weighted sum of the inputs is formed. To get the activation of the node the threshold is subtracted.
- The activation signal is passed through an activation function to compute the output of the node.

The first corresponding mathematical model was developed by McCulloch and Pitts in the 1940's. Fig. 4.2 on the facing page shows a schematic of the basic principle. In this

[2]All terms will be used in this study.

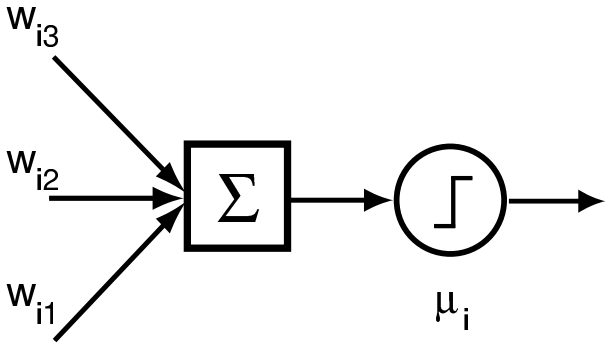


Figure 4.2: Visual representation of a single node in the McCulloch-Pitts model. First, the weighted sum of the input values is calculated and second, the threshold μ_i is subtracted. If the result is positive the output value is 1, otherwise it is 0 [2].

model the neural function is described by

$$n_i(t+1) = \Theta(\sum_j w_{ij} n_j(t) - \mu_i). \tag{4.1}$$

where $\Theta(x)$ is the Heavyside- or step-function defined by

$$\Theta(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The n_i is the node output of node i representing its state: 1 = fired, 0 = not fired. The variable t denotes a certain discrete time-step. The numbers w_{ij} are the weights and represent the strength of the connection between node i and node j. These weights can also be negative and hence act like an "inhibitory synapse" in biology. Positive weights correspond to "excitatory synapses". If a weight is zero no connection between node i and j has been established. Finally, μ_i is the threshold for node i. The weighted sum of the inputs has to extend this value for the neuron to fire.

In modern implementations of the McCulloch-Pitts model the Heavyside function in equation 4.1 is replaced by a continuous activation function. In general, these functions as well as their first derivative are monotonic. Often sigmoid functions[3] are used which map the interval $(-\infty, \infty)$ to $[-1, 1]$. The saturation is particularly desirable when the output is meant to represent a probability[4]. The non-linear mapping is the essential feature of neural networks. It enables the network to learn non-linear relations between different properties. Other properties that an activation function should have are continuity and smoothness of the function itself and its first derivative as well as linearity for small values of x. This will allow the system to implement a linear model if necessary. A sigmoid that is a popular activation function is given by the hyperbolic tangent

$$a \cdot tanh(bx) = a \frac{e^{bx} - e^{-bx}}{e^{bx} + e^{-bx}}$$

where a and b are free parameters. An example with $a = b = 1$ is depicted in fig. 4.3.

[3]The term sigmoid means "S-shaped".

[4]This choice can also be motivated by the Bayes theorem . See [48] for further details.

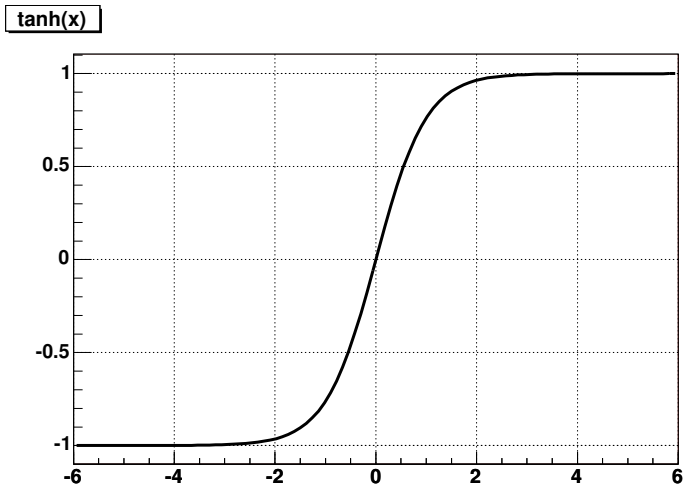


Figure 4.3: An example sigmoid function: the hyperbolic tangent $f(x) = tanh(x)$.

4.1.3 Network topologies

In order to build up neural networks from single nodes one has to choose a way of arranging nodes with each other, i. e. one has to decide on the network topology. The most basic networks consist of two layers, called input and output layer. However the input layer is not really neural at all, these nodes simply serve to introduce the values of the input variables. Each layer may consist of one or more nodes. Each node in the input layer is connected to each node in the output layer via a weight w_{ik}.

This vanilla topology can be extended by introducing a third layer between the input- and output layer. Fig. 4.8 on page 70 shows an example of of such a three-layer network. The second layer is usually called hidden layer because its nodes are not directly connected to the outside and hence are not seen by the user. Again all nodes in a certain layer are connected to all nodes in the next and/or previous layer. The data is "fed" to the input layer, multiplied by the weights w_{jk} and processed by the hidden layer to give its output. The outputs of layer two are transferred to layer three where the final output is calculated. This whole process corresponds to a "forward flow of information". The nodes of the network only rely on the output of nodes in the previous layer. This type of network is thus called a "feed-forward network".

The introduction of a bias node in the first layer may increase the performance of the network significantly. It is implemented as an extra node in the input layer with a constant input value of -1. It is connected to all nodes in the next layer and can be used to modify the threshold values of those nodes.

Recurrent networks are an extension to the feed-forward networks. Each node can be connected to any other node in the network or even to itself. The increased number of connections results in a higher number of weights. However it has been shown that recurrent networks give a large improvement in performance over normal feed-forward networks for a number of problems such as pattern completion. The concept of layers is somewhat obsolete here. For instance a certain node may serve as both input and output

node.

The network topology plays an important role for the neural net classification and the optimal architecture will depend upon the problem at hand. There is no general choice for the number of layers and nodes. However some guiding principles have evolved over time. The number of output nodes can easily be determined by the application the network is used for. For a binary classification ("yes-or-no decision") only one output node is necessary in the last layer. Several ones are needed for other tasks. The number of nodes in the hidden layer(s) is somewhat arbitrary. When choosing too few nodes there might be too few free parameters to allow for a good separation between the patterns. The network will not be able to learn all features of the presented input and the generalisation capabilities will be limited. With too many nodes present in the hidden layer the network might pick up noise in the training data. If the network is capable of pruning, i. e. it can eliminate redundant connections during the learning process itself, the number of nodes in the hidden layer(s) is not critical. As a rule of thumb the number of nodes in the hidden layer should be a bit higher than the number of nodes in the input layer. Obviously the number of nodes in the input layer is given by the number of input variables if no bias node is added.

As has been shown by Kolmogorov and others any continuous function from input to output can be implemented in a three-layer net, given a sufficient number of hidden units, proper nonlinearities and weights. Fig. 4.4 gives an example for the decision boundaries of a two- and three-layer network with two input nodes.

4.2 Network training

Training a network is the process of changing the weights such that the residual between the output and the target is minimised[5]. Usually the training process starts with an untrained network, i. e. the initial weights are chosen randomly. Training patterns are presented to the input layer and passed through the net. The network output is compared to the target value, i. e. the wanted network output. The difference between the two corresponds to an error. This error function[6] is a scalar function of the weights and minimal when the network outputs match the desired output values. Hence it is a measure for how far the network output is away from the true value. Adjusting the network weights such that the error function becomes minimal is the goal of the training process. The quadratic error function E is defined by

$$E[\vec{w}] = \frac{1}{2} \sum_i \sum_\mu [\zeta_i^\mu - O_i^\mu(\vec{w})]^2 \tag{4.2}$$

where ζ_i^μ is the desired output or training target and $O_i^\mu(\vec{w})$ is the current network output [7]. The weights w_{pq}[8] form the vector $\vec{w}$. The index i runs over the number of output nodes and μ over the input patterns.

[5] Only supervised learning will be discussed here.

[6] The error function is also called cost or criterion function.

[7] Cf. fig. 4.7 on page 69 and fig. 4.8 on page 70.

[8] Note that w_{pq} stands for any weight in the network, regardless of the layers it connects to.

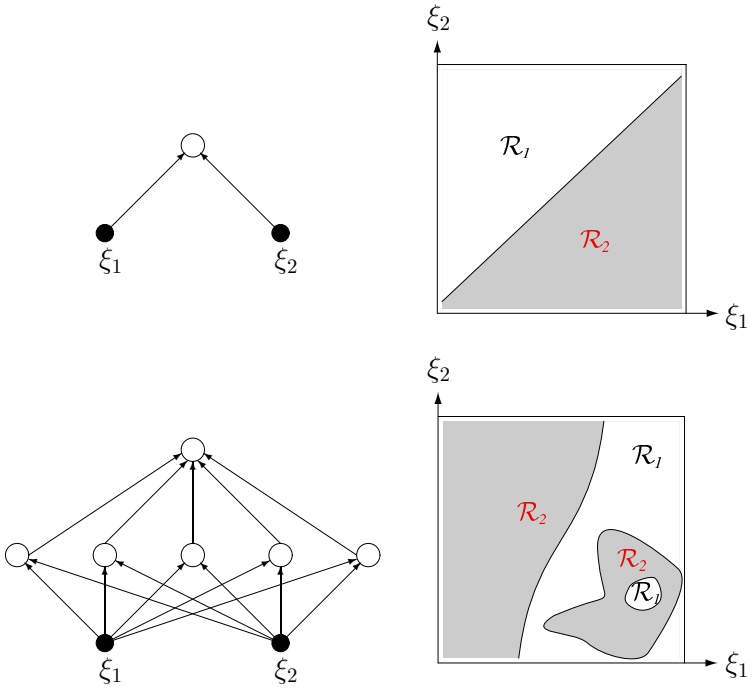


Figure 4.4: Decision boundaries for a two- and three-layer neural network with two input nodes. The two-layer network (upper plots) can only implement a linear decision boundary. Given an adequate number of hidden units, three- and higher-layer networks can implement arbitrary decision boundaries. The decision regions need not be convex or simply connected (lower plots) [45, 50].

Unfortunately the surface of the error function can be complicated and is in general unknown. Thus it is difficult to find the global minimum that would provide the optimal network performance, i. e. $E(\vec{w}) = 0$. In most cases the training process will end up in a local minimum.

Gradient Descent

A straightforward algorithm to calculate the weight changes needed for minimisation of the error function is called gradient descent. Here the weight change Δw_{pq} is proportional to the gradient of the error function and given by

$$\Delta w_{pq} = -\eta \frac{\partial E}{\partial w_{pq}}. \tag{4.3}$$

This procedure can be geometrically interpreted as a descent along the steepest direction of the error function. The learning rate η should be chosen such that the algorithm converges quickly. If η is chosen to small it will take long to converge and if it is too large the algorithm might oscillate wildly. Fig. 4.5 shows an example for a one-dimensional quadratic error function. Having a separate learning rate for each weight is advantageous because the weight change can be adjusted to compensate for a large or small derivate $\frac{\partial E}{\partial w_{pq}}$.

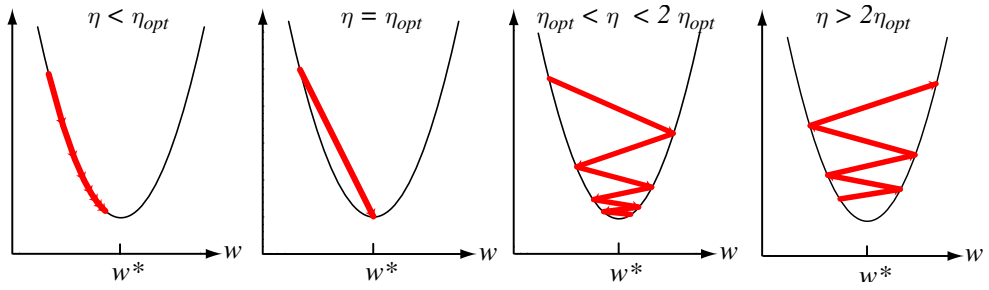


Figure 4.5: Gradient descent in a one dimensional quadratic error function with different learning rates. If $\eta < \eta_{opt}$, the optimal learning rate, convergence is guaranteed but the training can be needlessly slow. For $\eta = \eta_{opt}$ the minimum w^* of the error function is found in a single learning step. If $\eta_{opt} < \eta < 2\eta_{opt}$ the system oscillates but will converge finally. However the learning process is slow. If $\eta > 2\eta_{opt}$ the algorithm diverges [45, 50].

Initialisation of Weights

As can be seen from equation 4.3 on the preceding page one cannot set the initial weights to 0 because the weight changes would also be 0 and learning could not take place. It is commonplace to choose the initial weights randomly from a single distribution to help ensure uniform learning. Because the preprocessing described in section 4.3.1 on page 73 gives both positive and negative input values on average, the initial weights should be positive and negative as well. The weights should be small to avoid saturation of the nodes. However if the weights are chosen to be too small the sigmoidal activation function is linear which can lead to slow training. This suggests that the summed inputs to the nodes should be of order unity (cf. eqns. 4.4 and 4.6 on page 69).

Training protocols

The most common learning protocols for the supervised training of neural networks are stochastic, batch and online learning and all of them have their strengths and drawbacks[9]. In stochastic training, patterns are chosen randomly from the training set. In batch mode all patterns are presented to the network before the weights are being updated. In the online mode, training patterns are fed into the network one at a time and the weights are changed immediately. Hence there is no need of computer memory for storing patterns. It is also possible to present several patterns before calculating new weights. Another method is to let each node choose independently when to change its weights.

Shuffling the order of the input samples once all patterns have been presented keeps the network from learning correlations due to a specific order of the inputs.

[9] A forth protocol called "learning with queries" where the network output is used to select new training protocols is rarely used.

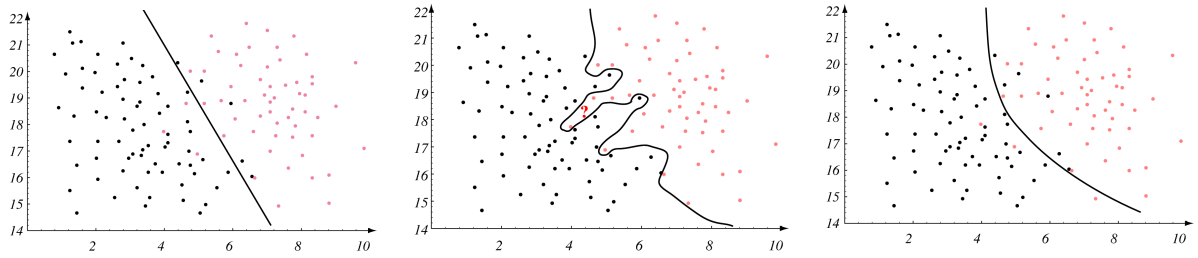


Figure 4.6: Finding the optimal decision boundary in a two-dimensional input space. The dark and the light dots can be distinguished by just using a linear decision boundary (left plot) which is obviously suboptimal. The overall classification error will be large. An overtrained network will lead to an overly complex model of the decision boundary (middle plot). While this can lead to a perfect classification of the training data the generalisation capabilities will be poor. The benchmark point marked ? most likely belongs to the class of dark dots but the complex model leads to a misclassification. The optimal boundary (right plot) represents a tradeoff between the performance on the training patterns and simplicity of the model, hence giving the highest accuracy on future data [45, 50].

Stopping criterion

The network training always has to stop at some point because excessive training can lead to poor generalisation capabilities, i. e. the network is overtrained. It implements a complex decision boundary tuned to the specific training data rather than the general properties of the underlying distributions (see fig. 4.6). The simplest stopping criterion is to end the training when the change in the error function $E[\vec{w}]$ is smaller than some preset value θ. A more favourable approach is to use a small fraction of the training patterns as validation set. Once the error function for the validation sample starts to increase the training should be stopped because the goal of the network training is a low generalisation error as sketched in fig. 4.7 on the next page.

4.2.1 Backpropagation

Backpropagation is a common method for supervised training of multilayer neural networks. It allows to calculate an effective error for each hidden unit and thus to derive learning rates for all weights. The training process is divided into two steps. First, a training pattern is presented to the input nodes and passed through the network to yield output values from the output nodes. Second, the error is calculated for each node starting with the output layer and propagated backwards. Here we consider the three-layer network shown in fig. 4.8 on the following page. The figure shows the notational conventions that will be used in the calculation of the weight changes. The index i always refers to an output unit, j to a hidden one and k to an input unit. The nodes themselves are denoted O_i (output nodes) and V_j (hidden notes). The input patterns presented to the first layer nodes are called ξ_k, the desired output values or targets of the third layer are named ζ_i. There are connections w_{jk} from the input to the hidden units and W_{ij} from

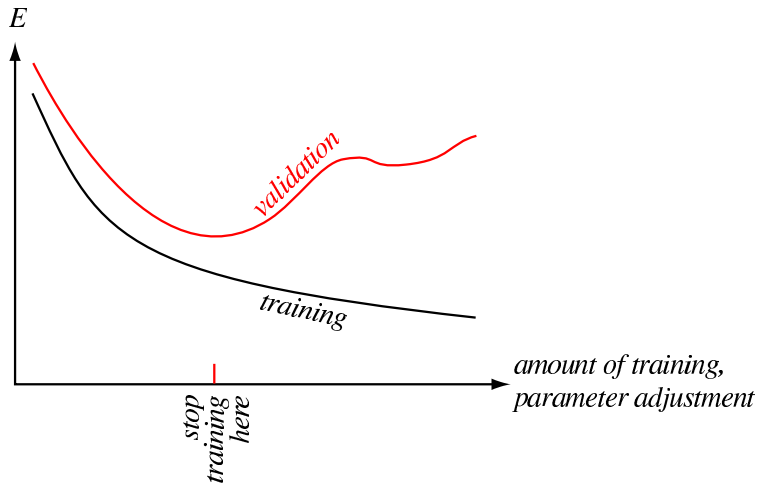


Figure 4.7: Stopping the network training by means of a validation sample. The training data can be split into two parts. The first one (e. g. 90%) is used as standard training set. The other (e. g. 10%) is the validation set and meant to represent the generalisation task. For most problems the training error decreases monotonically during training (black curve). The error on the validation set (red curve) decreases first, but than starts to increase and indicates that the network may be overfitting the data. The training should be stopped at the first minimum of the validation error [45, 50].

the hidden to the output nodes. Different patterns are labelled by a superscript μ, i. e. the input value for node k is ξ_k^μ when pattern μ is being presented.

In order to calculate the weight change Δw_{pq} for each node in the network one has to calculate the error function as a function of all weights first. Second, the derivatives as defined by the gradient descent method have to be computed. We start by calculating the output values O_i^μ for the nodes in the last layer.

Given an input pattern μ the hidden node j receives the input

$$h_j^\mu = \sum_k w_{jk} \xi_k^\mu \tag{4.4}$$

from all nodes in the first layer. The output value of this node is simply

$$V_j^\mu = g(h_j^\mu) = g(\sum_k w_{jk} \xi_k^\mu) \tag{4.5}$$

where $g(x)$ is the activation function of the node. Node i in the output layer thus receives

$$h_i^\mu = \sum_j W_{ij} V_j^\mu = \sum_j W_{ij} g(\sum_k w_{jk} \xi_k^\mu) \tag{4.6}$$

and produces the output value

$$O_i^\mu = g(h_i^\mu) = g(\sum_j W_{ij} g(\sum_k w_{jk} \xi_k^\mu)). \tag{4.7}$$

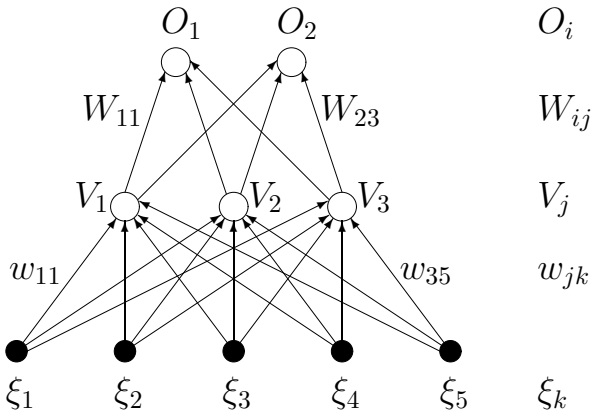


Figure 4.8: A three-layer network showing the notation for nodes and weights used in subsequent calculations. The index i refers to an output unit, j to a hidden one and k to an input unit. The nodes are denoted O_i (output nodes), V_j (hidden notes) and ξ_k (input nodes). Connections w_{jk} connect the input to the hidden units and W_{ij} the hidden to the output nodes.

The error function 4.2

$$E[\vec{w}] = \frac{1}{2}\sum_{i,\mu}[\zeta_i^\mu - O_i^\mu]^2 \tag{4.8}$$

now becomes

$$\begin{aligned} E[\vec{w}] &= \frac{1}{2}\sum_{i,\mu}[\zeta_i^\mu - g(h_i^\mu)]^2 &(4.9)\\ &= \frac{1}{2}\sum_{i,\mu}[\zeta_i^\mu - g(\sum_j W_{ij}V_j^\mu)]^2 &(4.10)\\ &= \frac{1}{2}\sum_{i,\mu}[\zeta_i^\mu - g(\sum_j W_{ij}\, g(\sum_k w_{jk}\xi_k^\mu))]^2. &(4.11)\end{aligned}$$

Applying the gradient descent rule from page 66 for the hidden-to-output connections one gets

$$\begin{aligned} \Delta W_{ij} = -\eta\frac{\partial E}{\partial W_{ij}} &= -\eta\sum_\mu[\zeta_i^\mu - O_i^\mu]\, g'(h_i^\mu)V_i^\mu &(4.12)\\ &= -\eta\sum_\mu \delta_i^\mu V_j^\mu &(4.13)\end{aligned}$$

with

$$\delta_i^\mu = g'(h_i^\mu)[\zeta_i^\mu - O_i^\mu]. \tag{4.14}$$

Here $g'(x)$ is the first derivate of the activation function. To calculate the weight changes Δw_{jk} for the input-to-hidden connections the chain rule has to be applied to get the derivatives w. r. t. the w_{jk}'s. One obtains

$$\Delta w_{jk} = -\eta\,\frac{\partial E}{\partial w_{jk}} \tag{4.15}$$

$$= -\eta \sum_{\mu} \frac{\partial E}{\partial V_j^\mu} \frac{\partial V_j^\mu}{\partial w_{jk}} \tag{4.16}$$

$$= -\eta \sum_{i,\mu} \underbrace{[\zeta_i^\mu - O_i^\mu] g'(h_i^\mu)}_{\delta_i^\mu} W_{ij} g'(h_j^\mu) \xi_k^\mu \tag{4.17}$$

$$= -\eta \sum_{i,\mu} \delta_i^\mu \, W_{ij} \, g'(h_j^\mu) \, \xi_k^\mu \tag{4.18}$$

$$= -\eta \sum_{i,\mu} \delta_j^\mu \, \xi_k^\mu \tag{4.19}$$

with

$$\delta_j^\mu = g'(h_j^\mu) \sum_i W_{ij} \delta_i^\mu. \tag{4.20}$$

It is obvious that eqn. 4.13 on the preceding page has the same form as eqn. 4.19 but with a different definition of the δ's. As can be seen from eqn. 4.20, δ_j^μ depends upon δ_i^μ that is needed to compute Δw_{ij}. This backward propagation of errors gave the method its name: backpropagation. In general the backpropagation update rule always has the form

$$\Delta w_{pq} = -\eta \sum_{patterns} \delta_{output} \times V_{input} \tag{4.21}$$

where the output and input refer to the ends of the connection concerned and V is the activation from a hidden unit or the real input. The definition of the δ's depends on the layer concerned. For the output layer it is given by eqn. 4.14 on the facing page. For other layers it is given by an equation like 4.20 and can be calculated by further application of the chain rule.

Testing network performance

After minimising the error function through the network training the output values of a single node should lie on a straight line, as shown in fig. 4.9 on the following page. Therefore it can be used to check the performance of the network. It is straightforward to show the linear correlation with the network purity.

The contribution of just a single output node to the error function is simply given by 4.2 on page 65 by omitting the summation over the index i. One obtains

$$E_i = \frac{1}{2} \sum_{\mu} [\zeta_i^\mu - O_i^\mu]^2. \tag{4.22}$$

The mean error of a single output node i is hence given by

$$\overline{E_i} = \frac{1}{2} \cdot \frac{1}{S+B} \sum_{\mu} [\zeta_i^\mu - O_i^\mu]^2 \tag{4.23}$$

since the number of training samples is the sum $S + B$ of signal and background events. Now the sum over the input patterns is split into two parts: the S signal samples with a

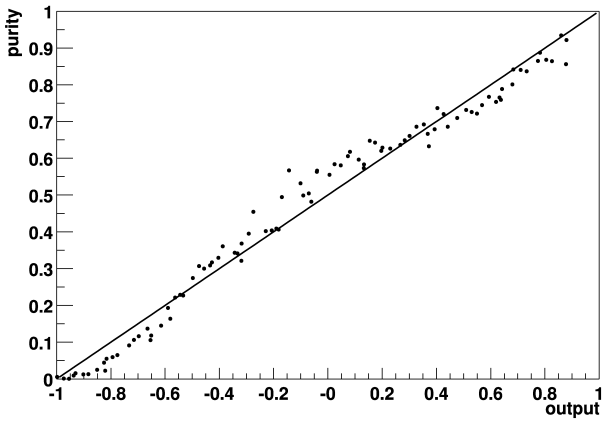


Figure 4.9: Signal purity as function of network output. The purity $\frac{S}{S+B}$ of a trained neural network (black points) as function of the network output almost shows a linear behaviour indicating that the training has been successful. The straight line represents the ideal case of a perfect network where $P = \frac{O+1}{2}$ [51, 52].

target value of +1 and the B background samples with a target value of -1. The mean error of an arbitrary output node for events with output O now becomes

$$\overline{E} = \frac{1}{2} \cdot \frac{1}{S+B} \left[S \, (1-O)^2 + B \, (-1-O)^2 \right] \tag{4.24}$$

$$= \frac{1}{2} \left[P \, (1-O)^2 + (1-P) \, (-1-O)^2 \right] \tag{4.25}$$

where P denotes the purity defined by

$$P = \frac{S}{S+B}. \tag{4.26}$$

The network training is optimal when $\overline{E}$ is minimal, i. e.

$$\frac{\partial \overline{E}}{\partial O} = 0. \tag{4.27}$$

Inserting eqn. 4.25 into 4.27 leads to

$$P = \frac{O+1}{2}. \tag{4.28}$$

Hence, for a well trained network the signal purity P as function of the network output O should have a linear behaviour.

4.3 Improving Network Performance

4.3.1 Preprocessing

The use of preprocessing often greatly improves the performance of a neural net. The process of finding the global minimum of the error function is significantly simplified if the input variables are transformed appropriately. This is because the sensitivity of the activation function is limited to a small input range around zero. Although the input can be in any range, there is a saturation effect for large input values.

Scaling

In order to avoid saturation effects, input values have to be scaled to a range that is appropriate for the network. Typically, variables are scaled linearly. However in some circumstances non-linear scaling may be appropriate[10].

Variable Transformation

Another way of preparing data for a neural network is to transform the variables such that they are equally distributed. The transformation is performed by integrating the input distribution. If the probability density distribution of a variable is given by $f(x_i)$ the transformation can be achieved by

$$y(x_i) = \int_{-\infty}^{x_i} \frac{1}{N} f(x_i')\, dx_i' \tag{4.29}$$

where N is a normalisation factor to restrict y to the interval $[0, 1]$. Fig. 4.10 on the following page explains the procedure.

Decorrelation

The above transformations only treat single variables without taking into account any possible correlations. Uncorrelated input variables have the advantage that the weights can be minimised independently which is easier than minimising all weights at the same time. Inputs can be decorrelated by diagonalising their covariance matrix. One gets new input variables which are linear combinations of the raw inputs. By dividing these new variables by the square root of the corresponding eigenvalue, the input vector space can be transformed to a hypersphere with a standard deviation of 1, centred around the origin.

For shape reconstruction the degeneracy of the new covariance matrix allows an arbitrary rotation of the basis in the input vector space. The basis can be rotated such that the first variable contains all information on the first moment (i. e. the mean) of the target distribution, the second variable all information on the second moment and so on.

[10] For instance, if an input variable is distributed exponentially one might take the logarithm.

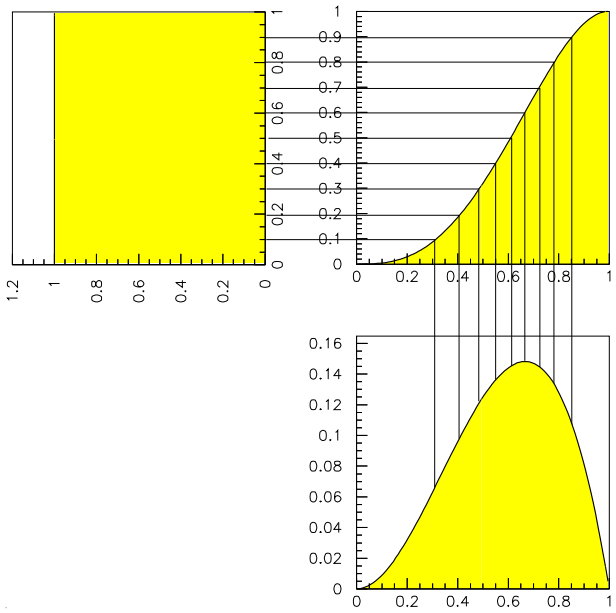


Figure 4.10: Equalising a distribution by use of the integrated function. The function $f(x_i)$ (lower plot) is integrated according to equation 4.29 on the preceding page leading to $y(x_i)$ (upper right). The output distribution (upper left) is flat between 0 and 1 and has 10 equidistant bins [53].

4.3.2 Regularisation

In order to avoid oscillatory behaviour of the network one can add a penalty term to the cost function:

$$E[\vec{w}] \rightarrow \tilde{E}[\vec{w}] = E[\vec{w}] + \lambda \cdot P[\vec{w}] \tag{4.30}$$

where P denotes the penalty term and λ is a free parameter. The new cost function $\tilde{E}[\vec{w}]$ not only depends on the classical training error but also on the complexity of the model. It penalises highly complex networks, i. e. networks with high weights. Searching for the minimum of the new error function is to balance the error on the training set with complexity. The parameter λ is adjusted to impose the regularisation more or less strongly. If λ is chosen too large the penalty term will dominate the error function. An easy way of implementing regularisation is to add a term based on the squared network weights:

$$P[\vec{w}] = \sum_{i,\mu} w_{pq}^2. \tag{4.31}$$

This approach is called "formal stabilisation" and will lead to a network with small weights.

4.3.3 Momentum Term

Error surfaces often have plateaus, i. e. regions where the slope $\frac{\partial E}{\partial \vec{w}}$ is very small. These can arise when there are too many weights and thus the error function depends only weakly

upon any of them. To allow the network to learn more quickly one can add a momentum term to the learning rule (cf. eqn. 4.3 on page 66)

$$\Delta w_{pq}(t+1) = -\eta \frac{\partial E}{\partial w_{pq}} + \alpha \cdot w_{pq}(t) \tag{4.32}$$

where α has to be positive and less than 1 for stability. Doing this one alters the learning rule such that some fraction α of the previous weight update is included in the current weight change. By increasing stability the momentum term can speed up the learning process.

4.3.4 Weight Decay and Pruning

Another approach for keeping the network weights small, apart from regularisation, is the weight decay method. The basic idea is to decay all weights during the training. After each weight update every weight is decreased according to

$$w_{pq}^{new} = w_{pq}^{old} \cdot (1 - \epsilon) \tag{4.33}$$

where $0 < \epsilon < 1$[11]. However it is also possible to shrink the weights exponentially. Giving all connections this tendency to decay to zero they will disappear unless reinforced by the network. Hence the network is forced to increase the weights it considers to be important. In the end the system achieves a balance between pattern error and some measure of the overall weight. Sometimes it might be advantageous to prune away connections whose weights are too small. This pruning technique, where weights are set to zero once they fall below a certain threshold value, can help the network to optimise its own topology. If all connections to a certain node are pruned away obviously the whole node is removed from the system. In the beginning of the training process the threshold should be set to a value close to zero and increase during the training process. This keeps the network from "destroying" itself right away since the initial weights might be small.

4.4 Advanced methods

4.4.1 The Hessian Matrix

The gradient descent algorithm described in section 4.2 on page 66 is one of the basic optimisation techniques. However it is not very efficient and more powerful techniques are available. Instead of just using the first-order derivatives one can also use higher order information. Here the second-order derivative is used to improve the finding of a minimum of the cost function.

Expanding the error function $E(\vec{w})$ in a Taylor series at the current point $\vec{w_0}$ one gets

$$E(\vec{w}) = E_0 + (\vec{w} - \vec{w_0}) \cdot \nabla E(\vec{w_0}) + \frac{1}{2}(\vec{w} - \vec{w_0}) \cdot H \cdot (\vec{w} - \vec{w_0}) + ... \tag{4.34}$$

[11] This is equivalent to regularisation (cf. section 4.3.2).

where H is the Hessian matrix

$$H_{pq} = \frac{\partial^2 E}{\partial w_p \, \partial w_q} \tag{4.35}$$

evaluated at $\vec{w_0}$[12]. Differentiating eqn. 4.34 gives a similar expression for the gradient:

$$\nabla E(\vec{w}) = \nabla E(\vec{w_0}) + H \cdot (\vec{w} - \vec{w_0}) + \ldots \, . \tag{4.36}$$

To find the minimum of $E(\vec{w})$ where $\nabla E(\vec{w}) = 0$, one can set eqn. 4.36 to zero, ignoring the higher order terms:

$$\nabla E(\vec{w_0}) + H \cdot (\vec{w} - \vec{w_0}) = 0. \tag{4.37}$$

Rewriting the last equation gives

$$\vec{w} = \vec{w_0} - H^{-1} \nabla E(\vec{w_0}) \tag{4.38}$$

as an estimate for the location of the minimum. Comparing this equation with the gradient descent rule (eqn. 4.3 on page 66) one sees that the new optimal learning rate is given by

$$\eta_{opt} = H^{-1}. \tag{4.39}$$

The above equations are just approximations in case the error function is not quadratic and they have to be used iteratively. However they improve the bare gradient descent method discussed earlier. The drawback of using this new method is that for a network having N weights the algorithm requires to compute, store and invert the $N \times N$ Hessian matrix. This makes the algorithm impractical for all but small problems. Fortunately several methods exist to compute some properties of the Hessian matrix without knowing all matrix elements. They can be used to get an estimate of the optimal learning rate. One should keep in mind that the algorithm doesn't converge necessarily for non-quadratic error surfaces. But more sophisticated methods, such as conjugate gradient descent, exist.

4.4.2 Conjugate Gradient Descent

An improved version of the gradient descent algorithm is conjugate gradient descent. It employs a series of line searches in weight space. In a line search starting from $\vec{w_0}$ in direction $\vec{d}$ one stays on the line

$$\vec{w} = \vec{w_0} + \lambda \cdot \vec{d} \tag{4.40}$$

and chooses λ to minimise the error function. The first search direction is simply given by

$$\vec{d} = -\nabla E(\vec{w_0}). \tag{4.41}$$

Subsequent search directions are calculated according to

$$\vec{d}^{new} = -\nabla E^{new} + \beta \cdot d^{old} \tag{4.42}$$

for some value of β. This new vector $\vec{d}$ is a compromise between the gradient direction and the previous search direction. The value of β is chosen such that the new search direction

[12]Note that the vector $\vec{w}$ is used for a point in the weight space. Specifying $\vec{w}$ corresponds to specifying all the network weights.

spoils as little as possible the minimisation achieved by the previous search step. The old and new search directions fulfill

$$\vec{d}^{old} \cdot H \cdot \vec{d}^{new} = 0 \tag{4.43}$$

where $\vec{d}^{old}$ and $\vec{d}^{new}$ are then said to be conjugate. The value of β can be calculated with the Polak-Ribiere rule:

$$\beta = \frac{(\nabla E^{new} - \nabla E^{old}) \cdot \nabla E^{new}}{(\nabla E^{old})^2}. \tag{4.44}$$

As can be seen from eqn. 4.42 and 4.44 no explicit knowledge of the Hessian matrix is needed for this method.

In summary, the conjugate gradient method (using the Polak-Ribiere variant) consists of a succession of line minimisations along directions given by eqn. 4.42 and 4.44. For quadratic error functions in n dimensions it finds the minimum in exactly n steps, i. e. the computation time depends only linearly on the dimension of the problem. However conjugate gradient descent requires batch training, i. e. all input patterns have to be presented to the network before weight changes are calculated.

4.4.3 Alternative cost functions

The quadratic error function (cf. eqn. 4.2 on page 65) is a common cost function because it is easy to compute. Nevertheless other error functions can be beneficial in some cases. In principal any differentiable function of the output and target values can be used, provided it is minimised when the network outputs equal the target values for all input patterns. One popular alternative is based on relative entropy. Given two discrete probability distributions $p(x)$ and $q(x)$ over the same variable x, the relative entropy, or Kullback-Leibler distance, is a measure of the "distance" between these two distributions:

$$d = \sum_x q(x)\, ln \frac{q(x)}{p(x)}. \tag{4.45}$$

Based on this relative entropy is the following cost function that has been suggested by several authors:

$$E = \sum_{i,\mu} [\, \frac{1}{2}\,(1 + \zeta_i^\mu)\, log \frac{1 + \xi_i^\mu}{1 + O_i^\mu} + \frac{1}{2}\,(1 - \zeta_i^\mu)\, log \frac{1 - \zeta_i^\mu}{1 - O_i^\mu}]. \tag{4.46}$$

Here the term $\frac{1}{2}(1 + O_i^\mu)$ has to be interpreted as the probability that the output value of unit i is true. $O_i^\mu = -1$ means definitely false and $O_i^\mu = +1$ means definitely true. Similarly $\frac{1}{2}(1 + \zeta_i^\mu)$ is interpreted as a target set of probabilities.

Like the quadratic error function 4.2 on page 65, equation 4.46 is always positive except when $O_i^\mu = \zeta_i^\mu$ for all i and μ, where $E = 0$. Its qualitative advantage is that it diverges if the output of one unit saturates at the wrong extreme. The quadratic measure just approaches a constant in that case and hence the learning can float around on a relatively flat plateau of E for a long time.

4.5 NeuroBayes®

NeuroBayes® (NB) is a new, powerful software package for data analysis initially developed at the University of Karlsruhe by Prof. M. Feindt[13]. It uses a robust and easy-to-use algorithm and is based on a feed-forward network and the Bayesian interpretation of the network output. In order to improve the network performance, input values are preprocessed using variable transformation and decorrelation. NeuroBayes® has already been used extensively for B-physics analyses within the DELPHI-experiment. Detailed information on the algorithm and its usage can be found in [52], [54] and [55].

The NeuroBayes® package has been used throughout in the analysis presented in chapters 5 and 6 where it was applied to solve binary classification problems. It allows to change many parameters of the network training. Among others, some of the available options and features are:

- various types of regularisation and preprocessing,
- the number of training iterations,
- pruning,
- automatic variable selection,
- two types of loss-functions,
- change of the momentum term,
- change of the learn speed and
- training with weights.

All of these features can influence the performance of the network. Some points have to be explained in more detail since they were used in the analysis.

- Automatic variable selection:
  NeuroBayes® can automatically determine the relevance for a variable to improve the separation of signal and background and provides a ranking table. Such an example ranking list is given in table 4.1. It is possible to cut on the σ value and exclude variables below the threshold from the training by using the automatic variable selection. Possible cut values range from $0.5\,\sigma$ to $4.5\,\sigma$ in steps of $0.5\,\sigma$. Although information is lost, the cut can help the network to learn and improve the performance.

- Momentum:
  A momentum can be specified for the training. It can help the neural network to get out of local minima[14]. The momentum term may lie in the interval $[0.0, 1.0]$. The default value is 0.0.

[13] NeuroBayes® and <phi-t>® are registered trademarks of Physics Information Technologies [3].
[14] Cf. page 74, section 4.3.3.

Rank	Variable	Correlation	σ
1	$Q \cdot \eta$	28.5	42.1
2	A	27.5	40.7
3	$E_t^{j_2}$	9.53	14.1
4	$\Delta E_t^{j_1 j_2}$	8.33	12.3
5	P_t^{Lep}	6.59	9.73
6	$M_{b\bar{b}}$	4.71	6.97
7	$E_t^{j_2}$	4.16	6.14
8	$E_t^{j_1}$	3.24	4.78
9	$M_{l\nu b\bar{b}}$	3.46	5.11
10	$\not{E}_t$	3.24	4.79
11	C	1.80	2.66
12	ΔR_{lj_2}	1.39	2.06
13	$M_{l\nu b}$	1.28	1.89

Table 4.1: Example ranking list. The column denoted correlation shows the correlation coefficient to the target value and σ the significance. These ranking tables are automatically provided by the NeuroBayes®-Teacher and written to the log file.

- Learn speed:
  A multiplicative factor by which the learning speed calculated by NeuroBayes® is multiplied. Thus, the network will learn faster but might not learn as well as with a low learning speed. As default a speed-factor of 1.0 (i. e. learning speed is not increased) is used.

- Training with weights:
  Weights can be assigned to the input patterns, i. e. events can be weighted by any real valued number in the range [0.0,1.0]. A weight of 1.0 means that the event is taken as it is. A training pattern with a weight of 0.0 would be ignored from the training process.

The optimisation of some of these parameters for the Higgs search described in chapter 5 was the subject of a dedicated study that is documented in section 5.6.3 on page 103.

'The sum of many small things
is a big thing.'
G. Altarelli

Chapter 5

Data analysis

In this chapter the analysis of the $WH \to l\nu b\bar{b}$ channel is described. In some cases parts of the more detailed material can be found in appendix C and D on pages 139 and 143 respectively.

After some preliminary remarks in section 5.1, section 5.2 describes the data sample and Monte Carlo simulations (MC) that have been used to model signal and background processes. The preselection cuts for the baseline event selection are given in section 5.3 on page 85 with the expected event numbers listed in section 5.4 on page 93. Section 5.5 on page 95 deals with sources of systematic errors. The last step in the analysis, the application of the neural networks, is described starting from page 100 in section 5.6. The final results are presented in chapter 6 on page 109.

5.1 Introduction

The analysis consists of two major parts. As a first step a set of simple preselection cuts is applied in order to have a first rough reduction of the background. In a second step neural networks are trained to further improve the signal to noise ratio. The final network output for each event lies between -1 (background like) and +1 (signal like). The cut on the output value discriminating between signal and background is chosen such that it provides the best a priori limit. Subsequently a mass window is chosen for the invariant mass of the dijet system. Again it is optimised to provide the best achievable limit. Counting the events for both data and background in the mass window provides the limit on the production cross-section for $p\bar{p} \to WH$ times the branching ratio $H \to b\bar{b}$.

5.2 Data and Monte Carlo samples

As has already been explained in section 3.3 on page 57, various background processes have to be taken into account. These are listed in table 5.1. In order to reduce confusion about the nomenclature in subsequent sections the naming scheme used in this study is given as well, if necessary. The signal processes are denoted as WH with the corresponding Higgs

Physics process	Alternative nomenclature
Top-quark production	
Top pair production	$t\bar{t}$
single Top	Wg, t-channel
single Top	W^*, s-channel
W + heavy flavour	
$Wb\bar{b}$	Wbb
$Wc\bar{c}$	Wcc
Wc	
W + light flavours	
$W2p$	
Di-boson production	
WW	
WZ	
ZZ	
Other	
$Z \rightarrow \tau\tau$	Ztt
QCD processes	non-W

Table 5.1: List of background processes and their nomenclature.

mass given in GeV/c^2. Six different mass values around 120 GeV/c^2 are considered in this study: 110 GeV/c^2 up to 150 GeV/c^2 in steps of 10 GeV/c^2 as well as 115 GeV/c^2. The last value lies close to the final LEP Higgs mass limit of 114.1 GeV/c^2.

Higgs mass	Generator	Dataset ID	Events
100 GeV/c^2	A+H	N/A	310K
110 GeV/c^2	A+H	hewo3w	469K
115 GeV/c^2	A+H	hewo0w	450K
120 GeV/c^2	A+H	hexo1w	477K
130 GeV/c^2	A+H	hexo2w	455K
140 GeV/c^2	A+H	N/A	342K
150 GeV/c^2	A+H	N/A	378K

Table 5.2: List of Monte Carlo signal samples. All samples were generated with Alpgen [63] and the subsequent fragmentation was done with Herwig [62]. All W and H decays were allowed in the generation.

5.2.1 Data sample and Luminosity

This analysis uses the full high-p_t dataset taken until September 7^{th}, 2003. The runs are defined by the silicon good run list (version 4, December 4^{th}, 2003) of the DQM group [56]. The good runs from 141544 to 157715 correspond to an integrated luminosity of $\mathcal{L}_{int} = 161.57 \pm 9.5\,\mathrm{pb}^{-1}$ for both the CEM and CMUP detectors [1]. The CMX luminosity adds up to $149.94 \pm 8.8\,\mathrm{pb}^{-1}$. The offline luminosity has been corrected by a factor 1.019 [57] and assigned a 6% uncertainty [58]. Official TopNtuples have been used produced with offline version 4.11.1 [59, 60]. Silicon tracks have been refit using the standard alignment and refitter from 4.11.1.

5.2.2 Monte Carlo samples

The Monte Carlo samples used in this study were created with Pythia [61], Herwig [62] and Alpgen [63]. Table 5.2 lists the signal and table 5.3 the background samples together with their size and CDF dataset ID. The QCD sample was created out of the real physics data. The isolation cut was reversed w. r. t. the preselection, i. e. the isolation had to be greater than 0.1. Furthermore the b-tag criterion was relaxed. Instead of requiring a SECVTX[2] b-tag, a taggability criterion [64] was applied demanding

- uncorrected $E_t > 10$ GeV,
- $\eta < 2.4$ and
- $N_{trk}^{good} \geq 2$ (see [65] for the definition of good tracks.).

[1] Runs 164844, 164870, 164871 and 164872 were excluded because of CSL problems. Runs 163463 and 163474 were removed because of a bad beamline.

[2] Secondary Vertex. A CDF b-tagging algorithm based on the reconstruction of secondary vertices.

Process	Generator	Dataset ID	Events	Remarks
Top				
$t\bar{t}$	Herwig	ttopli	378K	$m_t = 175$ GeV/c^2
W^*	Pythia	ttop0s	512K	s-channel
Wg	Pythia	ttop1s	440K	t-channel
W + heavy flavour				
$Wb\bar{b}$ (WevBB0p)	A+H	atop40	233K	
$Wb\bar{b}$ (WmvBB0p)	A+H	atop46	220K	
$Wb\bar{b}$ (WtvBB0p)	A+H	atop4c	219K	
$Wc\bar{c}$ (WevCC0p)	A+H	atop43	252K	
$Wc\bar{c}$ (WmvCC0p)	A+H	atop49	293K	
$Wc\bar{c}$ (WtvCC0p)	A+H	atop4f	304K	
Wc (WevC0p)	A+H	atop0w	293K	
Wc (WmvC0p)	A+H	atop3w	179K	
$Wc1p$ (WevC1p)	A+H	atop1w	290K	
$Wc1p$ (WmvC1p)	A+H	atop4w	264K	
W + light flavours				
$W2p$ (Wev2p)	A+H	atop02	189K	
$W2p$ (Wmv2p)	A+H	atop08	260K	
$W2p$ (Wtv2p)	A+H	atop2e	293K	
Di - boson				
ZZ (ZZ0p)	A+H	atop0z	223K	
WZ (WZ0p)	A+H	atop0y	191K	
WW (WW0p)	A+H	atop4x	944K	
Other				
$Z \rightarrow \tau\tau$	Pythia	zewk1t	497K	

Table 5.3: List of Monte Carlo background samples. As in table 5.2 on the page before, samples marked 'A+H' were generated with Alpgen [63] and the subsequent fragmentation was done with Herwig [62]. The nomenclature in brackets is the one used by the ALPGEN interface [66].

- 1 tight Lepton with $P_t > 20$ GeV (Standard e and μ selection)
- $\not\!\!E_t >$ 20 GeV (corrected)
- 2 tight Jets with $E_t > 15$ GeV, $|\eta| < 2.0$ (corrected)
- at least one b-tag
- Z veto, Cosmic veto, Conversion veto

Table 5.4: Overview of preselection cuts.

Variable	Cut value	TopNtuple variable
Global quantities		
Z veto	applied	smry->fTopEventClass & 0x8 != 0
Cosmic veto	applied	smry->fTopEventClass & 0x20 != 0
Missing E_t	≥ 20 GeV	
# tight Leptons	1	
$\lvert z_{lepton} - z_{eventvertex} \rvert$	≤ 5 cm	ele->TrkZ0 / myon->Z0 / zvertex
Jets		
E_t	≥ 15 GeV	
η	< 2	Eta
# of Jets	2	
# of SECVTX tags	≥ 1	

Table 5.5: Preselection cuts for Jets and global event quantities.

5.3 Preselection Cuts

The expected signature for associated Higgs production in the channel $WH \to l\nu b\bar{b}$ was already discussed in section 3.2.3 on page 55. One expects two b-jets, a high-p_t lepton and missing energy from the neutrino. In order to have a first reduction of the background the preselection cuts have been chosen to reflect the phenomenology. Only W decays to $e^{\pm}\nu$ and $\mu^{\pm}\nu$ are considered. Tau decays were omitted because of the difficult background situation. In a nutshell, the main cuts are listed in table 5.4. The complete set of selection cuts is listed in tables 5.5, 5.6 and 5.7 and explained in more detail in the subsequent sections. They largely follow the baseline event selection from [67] which itself is based on [68].

5.3.1 Jets and missing E_t

The cuts for the jet selection, missing E_t and other global event quantities are given in table 5.5. Jets are reconstructed using the JETCLU[3] algorithm. They are identified

[3]Jet Cluster. A CDF jet reconstruction algorithm.

Central Electron Variable	Cut value	TopNtuple variable
Fiducial and CEM	$= 1$	Fiducial
E_t	≥ 20.0 GeV	Et
P_t	≥ 10.0 GeV/c	TrkPt
E/P (if $p_t \leq 50$ GeV/c)	≤ 2.0	EP
E_{had}/E_{em}	$\leq 0.055 + 0.00045 \cdot E$	Hadem
L_{shr}	≤ 0.2	LshrTrk
$\lvert\Delta z\rvert$	≤ 3.0 cm	DeltaZ
$Q_{track} * \Delta x$	≥ -3.0 cm, ≤ 1.5 cm	Charge * DeltaX
χ^2_{strip}	≤ 10.0	StripChi2
$\lvert z_0\rvert$	≤ 60.0 cm	TrkZ0
Good COT axial segments	≥ 3	TrkAxSeg
Good COT stereo segments	≥ 3	TrkStSeg
Isolation	≤ 0.1	Isolation
Conversion	false	Conversion

Table 5.6: Preselection cuts for CEM electrons.

as a cluster of energy within a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \leq 0.4$ using the event z-vertex. Their energies are corrected at level 4 using the jet energy corrections defined in ref. [69]. A short list of jet correction levels with explanations can be found in [70]. The number of required jets is two and at least one of the jets has to have a SECVTX b-tag [65]. Larry Nodulman's curvature correction for muons is applied in the calculation of missing E_t. The events had to pass the ELECTRON_CENTRAL_18 trigger for CEM electrons, the MUON_CMUP18 trigger for CMUP muons and the MUON_CMX18 trigger for CMX muons. The check for cosmic rays is only done for muon events. The cosmic ray flag from the CosmicFinderModule [71] is checked to be zero. A global Z veto flag is checked to remove Z candidates. Events are skipped when the tight lepton and a second object forms an invariant mass that lies within a window around the Z mass. The default window is [76,106] GeV/c^2. A detailed description of the Z veto can be found in ref. [72].

5.3.2 Electrons

The cuts for the electron selection are listed in table 5.6. Most of the cuts represent the standard CDF electron selection for electrons in the CEM calorimeter and have been developed by the CDF Electron Task Force [73]. The variable definition and the meaning of the separate cuts are as follows:

- E_t:
  The transverse electromagnetic energy deposited by the electron in the CEM is calculated as the EM cluster energy multiplied by $\sin\theta$. θ is defined as the polar angle provided by the best COT track pointing to the electromagnetic cluster (cf. fig. 2.6 on page 26). Details of the EM cluster algorithm can be found in ref. [74].

- P_t:
  The transverse momentum of the COT beam constrained track as measured using the track curvature in the COT within the magnetic field.

- E/P:
  The ratio of the EM cluster transverse energy to the COT track transverse momentum.

- E_{had}/E_{em}:
  The ratio of the hadronic (CHA + WHA) calorimeter energy to the electromagnetic (CEM) calorimeter energy for a cluster.

- L_{shr}:
  The lateral shower profile for electrons. The variable compares the energy in CEM towers adjacent to the seed tower for data and test beam electrons [75].

- $|\Delta z|$:
  The distance in the r-z plane between the extrapolated beam constrained COT track and the best matching CES cluster.

- $Q_{track} * \Delta x$:
  The distance in the r-ϕ plane between between the extrapolated beam constrained COT track and the best matching CES cluster times the charge of the track.

- χ^2_{strip}:
  The χ^2 comparison of the CES shower profile in the r-z view with the same profile extracted from test beam electrons.

- $|z_0|$:
  The z intersection of the track with the beam axis in the r-z plane.

- Track quality cuts:
  The electron associated track must have passed through 3 axial and 3 stereo superlayers (SL) of the COT to ensure a high track quality.

- Isolation:
  The energy in a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \leq 0.4$ around the electron cluster excluding the electron cluster divided by the energy in the electron cluster. The isolation variable is corrected for calorimeter leakage [76, 77].

- Conversion:
  For each electron a conversion flag is tested. A values of 1 means the electron is flagged as coming from a conversion $\gamma \rightarrow e^+e^-$. Trident events where a conversion is caused by a bremsstrahlung photon are still considered good events. In these cases the conversion flag has a value of -2 [68].

5.3.3 Muons

Table 5.7 on the following page shows the cuts for the muon selection. They mostly

Central Muon Variable	Cut value	TopNtuple variable
Region	CMUP or CMX	
P_t	≥ 20.0 GeV/c	Pt
E_{em}	$\leq max(2, 2 + 0.0115 \cdot (p - 100))$ GeV	EmEnergy
E_{had}	$\leq max(6, 6 + 0.0280 \cdot (p - 100))$ GeV	HadEnergy
CMU $\lvert\Delta x\rvert$	≤ 3.0 cm	CmuDx
CMP $\lvert\Delta x\rvert$	≤ 5.0 cm	CmpDx
CMX $\lvert\Delta x\rvert$	≤ 6.0 cm	CmxDx
$\lvert z_0\rvert$	≤ 60.0 cm	Z0
$\lvert d_0\rvert$ if no Si hits	≤ 0.2 cm	D0
$\lvert d_0\rvert$ if Si hits	≤ 0.02 cm	D0
Good COT axial segments	≥ 3	TrkAxSeg
Good COT stereo segments	≥ 3	TrkStSeg
Isolation	≤ 0.1	Isolation
COT exit radius (CMX only)	≥ 140 cm	

Table 5.7: Preselection cuts for CMUP and CMX muons.

represent the recommended cuts by the muon group [78]. The variables have the following definition:

- P_t:
  Transverse momentum of the muon track P_t. The cut value ensures a high trigger efficiency.

- E_{em}:
  The energy deposited in the EM calorimeter by the muon. The low cut selects minimum ionising particles (MIP's).

- E_{had}:
  The energy deposited in the hadronic calorimeter by the muon. Again the cut is chosen to select MIP's.

- $|\Delta x|$:
  The three different Δx cuts for the CMU, CMP and CMX ensure a good match in the r-ϕ plane between extrapolated tracks and muon stubs.

- $|z_0|$:
  The cut on $|z_0|$ enforces calorimeter projectivity, i. e. the calorimeter towers can only be projected on to the vertex within a range of -60 cm $\leq z_0 \leq$ + 60 cm.

- $|d_0|$:
  The impact parameter of the muon has two different cut values. They depend on tracks having silicon hits or not. The purpose of the cut is cosmic ray (CR) rejection.

- Track quality cuts:
  As was the case for electrons, the muon track must have passed through 3 axial and 3 stereo superlayers of the COT to ensure a high track quality.

- Isolation:
  The muon isolation has to be smaller than 0.1 in order to reject QCD events.

- COT exit radius (CMX only):
  The level 1 CMX18 trigger requires that a track has hits in all four COT superlayers. However, it is geometrically possible for a CMX muon not to pass all four layers. In order to treat data and Monte Carlo in a consistent way, one removes that part of the acceptance and demands that the COT exit radius of a track belonging to a CMX muon has to be greater than 140 cm. This cut guarantees that the muon intersects all four COT superlayers [79].

5.4 Event Estimation

5.4.1 Acceptance

For simple event counting the number of expected events for a certain physics process may be written as

$$N = \sigma \cdot BR \cdot \epsilon. \tag{5.1}$$

The total event detection efficiency can be defined as

$$\epsilon = \epsilon_{z_0} \cdot \epsilon_{geom} \cdot \epsilon_{lepton\ id} \cdot \epsilon_{iso} \cdot \epsilon_{met} \cdot \epsilon_{veto} \cdot \epsilon_{jet} \cdot \epsilon_{b-tag} \cdot \epsilon_{trig} \tag{5.2}$$

where ϵ_{z_0} is the efficiency of the $|z_0| \leq 60$ cm cut, ϵ_{geom} is the efficiency of the lepton to be in the fiducial region, $\epsilon_{lepton\ id}$ is the efficiency to identify a lepton and ϵ_{iso} is the efficiency of the isolation cut. The efficiency of the missing E_t cut is given by ϵ_{met} and ϵ_{veto} is the efficiency of the conversion, cosmic and Z^0 rejections. ϵ_{jet} gives the percentage of events fulfilling the jet requirements from table 5.5 and ϵ_{b-tag} is the efficiency to b-tag at least one jet in the event. Finally ϵ_{trig} is the trigger efficiency for identifying high P_t leptons.

For the calculation of the signal acceptance and the Monte Carlo derived backgrounds several scale factors have to be taken into account and applied to the simulation. This is done in order to correct for various over- and under-efficiencies in the Monte Carlo, i. e. to calibrate the simulation to the data.

A common scale factor for all lepton types is that for the z-vertex cut. It has been measured from data and found to be $\epsilon_{z_0} = 0.948 \pm 0.003$ [80]. The scale factors for lepton identification efficiencies are 0.965 ± 0.006 [81], 0.939 ± 0.007 and 1.014 ± 0.007 [82] for CEM, CMUP and CMX, respectively. The scale factors for CMUP and CMX muon reconstruction efficiencies are 0.945 ± 0.0006 and 0.992 ± 0.003 [82]. The trigger efficiency for CEM electrons is given by 0.9656 ± 0.0006 [83]. For CMUP and CMX muons the trigger efficiencies are 0.887 ± 0.007 and 0.954 ± 0.006 [82]. The scale factor for the b-tagging efficiency has been measured using a b-enriched electron sample and found to be 0.82 ± 0.06 [84]. The total b-tag efficiency averaged over E_t is given by 0.2403 and 0.2924 for Data and Monte Carlo respectively. Fig. 5.1 on the next page shows the efficiency and the scale factor as function of jet E_t.

The combined systematic error of these corrections is asssumed to be 10% in the calculation of the total systematic error in chapter 5.5.6 on page 98.

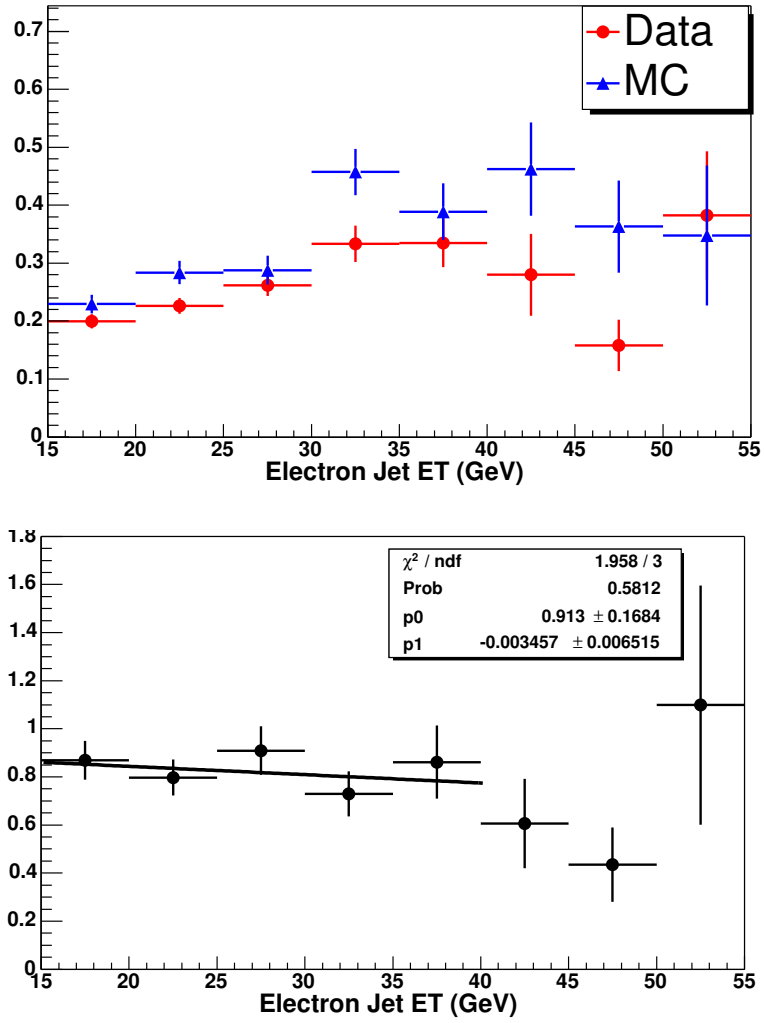


Figure 5.1: Efficiency to tag a b-jet as function of jet E_t in data and Monte Carlo (top) and the data/MC scale factor (bottom) [84]. The average tagging efficiencies are 0.2403 (data) and 0.2934 (MC) leading to a sale factor of 0.82 .

5.4.2 Signal process

The number of events passing each cut for the process $WH \rightarrow l\nu b\bar{b}$ with a Higgs mass of $m_H = 120$ GeV/c^2 is summarised in table 5.8 on the next page. Tables with numbers for the other six Higgs masses can be found in Appendix C on page 139. In all tables the number of double-tags is given for informational purpose only. The acceptance is calculated as the ratio of events passing the z-vertex cut (denoted N_{obsv}) and $N_{single-tag}$, the number of events having at least one SECVTX b-tag. This number has to be corrected to take into account the effects described in the previous section. The corrected acceptance for all Higgs masses is listed in table 5.9 on the next page and shown in fig. 5.2 on page 92. It rises almost linearly from 1.42 ± 0.14 to 1.83 ± 0.18 as m_H increases from 110 GeV/c^2 to 150 GeV/c^2. In order to get the right event estimate the various scale factors listed in section 5.4.1 have to be applied to the samples. Furthermore it is based on the measured luminosities presented in section 5.2.1 on page 83.

Cut	CEM	CMUP	CMX	Total
Initial sample				477500
N_{obsv}				462414
N_{geom}	125627	19125	9119	147930
$N_{leptonid}$	22912	13186	5289	41289
N_{iso}	20870	11992	4796	37611
N_{met}	18571	10565	4220	33316
N_{diveto}	18396	10503	4197	33058
N_{zveto}	18269	10464	4185	32880
$N_{convveto}$	17551	10075	4033	31626
N_{jet}	9867	5754	2274	17877
$N_{single-tag}$	5678	3320	1323	10310
$N_{double-tag}$	1456	809	331	2593
Acceptance	0.0123	0.00718	0.00286	0.0223

Table 5.8: Number of CEM, CMUP and CMX events after each cut. The Higgs mass was set to $m_H = 120$ GeV/c^2.

Process	N_{obs}	$N_{single-tag}$	Accept.	Corr. Accept.
$WH110$	454940	9444	2.08%	1.49%
$WH115$	436430	9274	2.12%	1.53%
$WH120$	462414	10310	2.23%	1.61%
$WH130$	441022	10061	2.28%	1.64%
$WH140$	330274	8154	2.47%	1.77%
$WH150$	365165	9252	2.53%	1.82%

Table 5.9: Uncorrected and corrected acceptance for all Higgs masses.

5.4.3 Background processes

Mistags and W+Heavy Flavour

The $Wb\bar{b}$ background has exactly the same final state as the signal process and hence is an irreducible background. Events originating from $Wc\bar{c}$ can survive the selection requirements due to the mis-identification of a c-jet. Background numbers shown in table 5.10 are taken from [67] and scaled accordingly. The numbers are determined from both Monte Carlo and data. The contribution from heavy flavour production in W+jet events is obtained from measurements of the heavy flavour event fraction in these events and their tagging efficiency. Their overall rate is normalised to the data [85, 86].

non-W Background (QCD)

As in the previous section, the background estimate for QCD background is taken from [67]. The missing $\not\!\!E_t$ vs. isolation method has been used to calculate the expected event

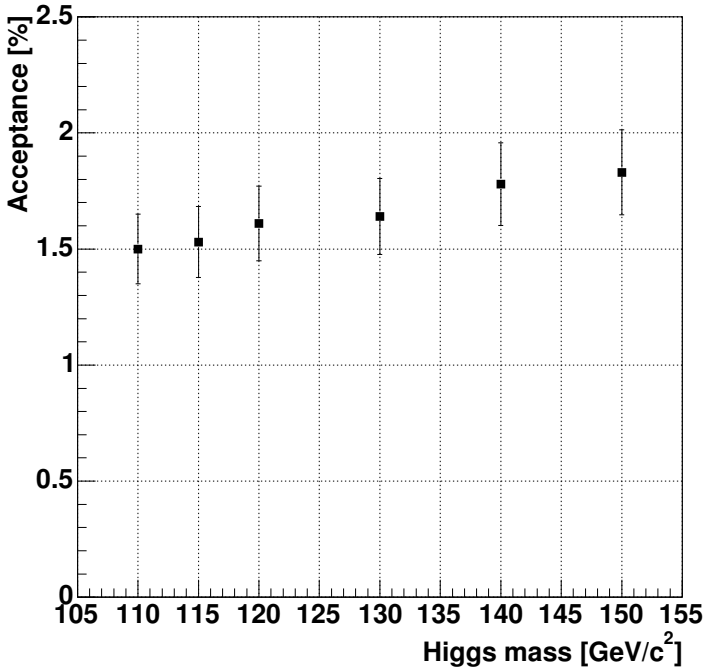


Figure 5.2: The total corrected $WH \to l\nu b\bar{b}$ acceptance as function of the Higgs boson mass.

numbers. A detailed description of the method can be found in ref. [87]. The background estimate is given in table 5.10.

Single top Background

Single top production in both the s-channel and the t-channel contributes to the background. See table 5.10 for the estimated number of events.

Diboson background and $Z \to \tau\tau$

The diboson background consists of WW, WZ and ZZ production where one of the bosons decays to leptons and the other one decays into jets with c- and b-quarks. In the case of $Z \to \tau\tau$, one of the τ's decays leptonically and the other one decays having three or more charged tracks. These so called 3-prong decays could be mis-identified as a jet with a displaced vertex due to the τ lifetime. The expected contribution for the four background sources is given in table 5.10 on the facing page.

5.4.4 MC derived Background

The normalisation for the top-pair cross-section has been measured by the CDF experiment. A value of $\sigma_{t\bar{t}} = 5.6 \pm 1.6$ pb is assumed [88]. As was the case for the signal

Process	
$WH110$	0.52 ± 0.09
$WH115$	0.45 ± 0.078
$WH120$	0.41 ± 0.071
$WH130$	0.31 ± 0.054
$WH140$	0.27 ± 0.047
$WH150$	0.20 ± 0.035
Mistags	17.0 ± 2.4
$Wb\bar{b}$	22.5 ± 6.5
$Wc\bar{c}$	8.0 ± 2.2
Wc	7.7 ± 2.0
non-W/QCD	10.5 ± 1.9
$t\bar{t}$	7.38 ± 1.14
single Top($Wg + W^*$)	4.6 ± 0.5
$WW/WZ/ZZ/Z \to \tau\tau$	2.5 ± 0.4
Total background	80.18 ± 9.9
Data	73

Table 5.10: Number of expected events for signal, background and data after the preselection. Numbers are based on an integrated luminosity of $\mathcal{L}_{int} = 162\,\mathrm{pb}^{-1}$. In the diboson processes one boson decays leptonically with the other boson decaying into hadrons. The error on the number of signal events is taken from table 5.16 on page 98. The statistical error for the signal has been neglected due to the high Monte Carlo statistics of about ten thousand events for each mass.

samples, the different scale factors listed in section 5.4.1 have to be considered as well as the luminosities. The acceptance table can be found in appendix C on page 142.

5.4.5 Summary of Event Estimation

Table 5.10 gives the number of observed data events and summarises the background and signal estimation for all processes after the preselection. The expectation for the various backgrounds is in reasonable agreement with the CDF data.

5.5 Systematic Errors

This section discusses the systematic uncertainties on the signal normalisation. Systematic errors are studied for the signal process $WH \to l\nu b\bar{b}$ with $m_H = 120$ GeV/c^2. These errors are also applied to the Monte Carlo samples for the remaining six Higgs masses. Uncertainties for the $t\bar{t}$ background are discussed in section 5.5.7 on page 98. One has to consider several sources of uncertainties, i. e.

- the uncertainty due to the jet energy scale,
- the uncertainty due to the modelling of initial and final state radiation,
- the uncertainty due to the choice of parton distribution functions,
- the uncertainty due to the signal generators Alpgen and Herwig versus other signal generators,
- the uncertainty on the top quark mass (for the $t\bar{t}$ background),
- the uncertainty on the b-tagging efficiency and
- the uncertainty on the trigger and lepton ID efficiency.

To study these systematic uncertainties various Higgs samples with $m_H = 120$ GeV/c^2 were generated:

- a sample with ISR switched off,
- samples with different PDF's
- and a sample using Pythia as generator rather than Alpgen in conjunction with Herwig.

These samples are listed in table 5.11 on the following page. The remaining systematic errors could be determined by using the standard MC samples listed in table 5.2 on page 83.

5.5.1 Jet energy scale

The uncertainty on the jet energy scale (JES) was investigated by using the member function in the JetEnergyCorrection class which varies the jet energy scale by one sigma [89, 69]. Instructions for its usage are documented in [90]. The resulting deviations are given in table 5.12 on the following page in percent. To cover the normalisation uncertainty due to the jet energy scale a systematic error of 2.67% is assigned to the WH process. This covers the maximum changes seen in table 5.12.

Higgs mass	Generator	Events	Remarks
120 GeV/c^2	Pythia	100K	Generator
120 GeV/c^2	Pythia	100K	ISR off
120 GeV/c^2	A+H	100K	CTEQ4L
120 GeV/c^2	A+H	100K	CTEQ6L
120 GeV/c^2	A+H	100K	MRST99
120 GeV/c^2	A+H	100K	MRTSLO02

Table 5.11: List of Monte Carlo systematic samples. Samples marked 'A+H' were generated with Alpgen [63] and the subsequent fragmentation was done with Herwig [62].

Process	Variation	CEM	CMUP	CMX	Total
$WH120$	$+1\sigma$	+2.67%	+2.22%	+2.31%	+2.45%
	-1σ	-0.11%	+0.12%	+0.18%	0.00%

Table 5.12: The uncertainty due to the jet energy scale.

5.5.2 Radiation modelling

At present there exists no realistic way to estimate uncertainties due to the modelling of ISR[4] and FSR[5] radiation in the Monte Carlo. Realistic would be to use two different models which represent a reasonable range. Instead Monte Carlo events were ISR was completely switched off were generated. The effect of ISR is taken to be half the absolute difference between the nominal efficiency and the no-ISR efficiency:

$$\Delta\epsilon^{ISR} = \frac{\epsilon^{noISR} - \epsilon^{0}}{2\epsilon^{0}}. \tag{5.3}$$

The FSR uncertainty was determined from the no-ISR sample. Reconstructed jets were matched with gluons using the Monte Carlo truth. When the gluon and the jet had an η-ϕ-distance of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \leq 0.4$ the jet was assumed to originate from the gluon. This way the acceptance change due to FSR could be determined. It is given by

$$\Delta\epsilon^{ISR} = \frac{\epsilon^{noISR,noFSR} - \epsilon^{noISR}}{2\epsilon^{noISR}}. \tag{5.4}$$

Acceptance deviations due to ISR and FSR are listed in table 5.13 on the next page.

5.5.3 Top quark mass

It is assumed that the uncertainty on to top quark mass m_t has no impact on the WH acceptance. However the effect is considered for the $t\bar{t}$ background on page 98.

[4]Initial State Radiation
[5]Final State Radiation

Process	Radiation type	Deviation
$WH120$	ISR	+5.71 %
	FSR	-4.63 %

Table 5.13: Relative deviation due to ISR and FSR.

Process	PDF	Deviation
$WH120$	CTEQ5L (default)	0.00%
	CTEQ4L	+1.75%
	CTEQ6L	-3.15%
	MRST99	-8.74%
	MRST02LO	+0.35%

Table 5.14: Relative deviation due to the choice of PDF.

5.5.4 Parton Distribution Functions

To evaluate the uncertainty associated with the choice of a specific PDF[6] several sets were investigated. The results are given in table 5.14. The maximum deviation (MRST99) from the standard PDF set (CTEQ5L) was taken with

$$\Delta\epsilon^{PDF} = \frac{\epsilon^{newPDF} - \epsilon^{0}}{\epsilon^{0}}. \tag{5.5}$$

5.5.5 Signal and Background generators

Systematic errors can arise from the Monte Carlo generator used for creating the signal and background samples. The different programs use different models (for instance for the hadronisation of jets) or are based on different cross-section calculations. The shifts imposed on the signal are listed in table 5.15 and are given by

$$\Delta\epsilon^{Gen} = \frac{\epsilon^{newGen} - \epsilon^{0}}{\epsilon^{0}}. \tag{5.6}$$

Process	Generator	Deviation
$WH120$	Alpgen+Herwig (default)	0.0 %
	Pythia	-5.15 %

Table 5.15: Relative deviation for the signal caused by different event generators.

[6]Parton Distribution Functions

Source	Efficiency change $WH120$
Jet E_t $+1\sigma$	2.67 %
Jet E_t -1σ	0.0 %
ISR	5.71 %
FSR	4.63 %
PDF	8.74 %
Generator	5.15 %
Top mass	0.0 %
ϵ_{evt}	10 %
Luminosity	6%
Quadratic sum	17.32%

Table 5.16: Summary table of systematic uncertainties on the signal normalisation. ϵ_{evt} is the error on the acceptance (cf. section 5.4.1 on page 89).

Generator	Events	Remarks
Herwig	100K	$m_t = 170GeV$
Herwig	100K	$m_t = 180GeV$
Pythia	100K	Generator
Pythia	100K	ISR off
Pythia	100K	CTEQ3L
Pythia	100K	GRV94L
Pythia	100K	MRST72
Pythia	100K	MRST75

Table 5.17: List of $t\bar{t}$ systematic samples.

5.5.6 Summary of Systematic Uncertainties

Table 5.16 summarises the systematic uncertainties on the signal normalisation. ϵ_{evt} is the total error on the acceptance and represents a conservative estimate. The quadratic sum of 17.32% is applied to all six Higgs masses.

5.5.7 $t\bar{t}$ Systematics

To determine systematic uncertainties for the Monte Carlo derived $t\bar{t}$ background several MC samples had to be generated. These are listed in table 5.17. In addition to the error sources mentioned in the previous section, the effect of a varying top mass has to be taken into account. As in the signal case, remaining systematic errors could be determined by using the standard MC sample. The errors are determined in the same way as it was done for the signal process. The results are summarised in table 5.18 on the facing page.

Source	Efficiency change $t\bar{t}$
Jet E_t $+1\sigma$	8.0%
Jet E_t -1σ	0.28%
ISR	3.8%
FSR	1.72%
PDF	4.5%
Generator	1.4%
Top mass	0.85%
ϵ_{evt}	10%
Luminosity	6%
Quadratic sum	15.51%

Table 5.18: Systematic uncertainties for the $t\bar{t}$ background.

5.6 Application of the Neural Network

The NeuroBayes® package consists out of two major parts: the NeuroBayes®-Teacher and the NeuroBayes®-Expert. The network training is performed by calling the NeuroBayes®-Teacher. This part sets up the network topology and training parameters and performs the actual training (cf. chapter 4.2 on page 65). When the training process is finished a file is written (called the "expertise") that contains all the information (e. g. weights between the nodes) to perform an analysis on new data. This analysis is done via the NeuroBayes®-Expert. It reads in the expertise file and classifies unknown events.

The Monte Carlo samples listed in tables 5.2 and 5.3 on page 83 and 84 were divided into two parts with each part containing 50% of the events. This was done in order to have two independent sets of MC data. The first one was used for the network training with the NB-Teacher, the second one was used for measuring the network performance with the NB-Expert.

Three-layer networks have been used throughout in the analysis. The number of nodes in the first, or input layer, was equal to the number of input variables[7]. Different sets of variables have been tested as will be explained in section 5.6.3. The number of nodes in the second layer was varied as well, however it was always larger than the number of nodes in layer one. In all cases there was only one output node in the third layer. It would have been possible to use one output node for each background type[8]. But for binary classification problems such a network topology is not superior to one with just a single node in the last layer [91].

5.6.1 Network input variables

The following list of quantities has been considered as network input variables[9].

- η_{Lep}, the pseudo-rapidity of the lepton defined by $-\ln\tan(\frac{\theta}{2})$.
- P_t^{Lep}, the transverse momentum of the lepton given by $P_t = P\sin\theta$.
- $\not{E}_t$, the missing transverse Energy.
- $E_t^{j_1}$ and $E_t^{j_2}$, the transverse energy of the first and second jet w. r. t. the z-axis where $E_t = E\sin\theta$.
- $\Delta E_t^{j_1 j_2}$, the difference between the scalar E_t of the first and second jet.
- H_t, the scalar sum of the P_t of the jets.

[7] In addition, NeuroBayes® introduces a so-called bias-node in the first layer.

[8] This setup has been used in ref. [2].

[9] The variable `jpbRPpos` from the JetProb algorithm [92] has been looked at as well and the background reduction could be improved. Unfortunately this variable was not properly filled in the reconstructed Ntuples for the muon data sample and hence dropped from the input variable list.

- $\Delta R_{j_1 j_2}$, $\Delta R_{l j_1}$, $\Delta R_{l j_2}$, the difference $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between jet 1 and jet 2, jet 1 and the lepton as well as jet 2 and the lepton.

- Sphericity S, aplanarity A and centrality C[10].

- $Q \cdot \eta$, with Q being the lepton charge (in units of the elementary charge e) and η being the pseudo-rapidity of the b-tagged jet in the event. If there are multiple b-tagged jets, the jet which the maximum $Q \cdot \eta$ is taken [94].

- $M_{l\nu b}$, the reconstructed top-quark mass [94]. When reconstructing $M_{l\nu b}$ the z-component of the neutrino is unknown but can be calculated up to a two-fold uncertainty using the following kinematic constraints:

$$p_\mu(W) = p_\mu(l) + p_\mu(\nu) \qquad\qquad p_\mu(\nu)\, p^\mu(\nu) = 0. \tag{5.7}$$

  Solving for $p_z(\nu)$ gives:

$$\begin{aligned} p_z(\nu) &= \frac{\kappa p_z(l)}{E^2(l) - p_z^2(l)} \pm \frac{1}{2(E^2(l) - p_z^2(l))} && (5.8)\\ &\quad \cdot\sqrt{(2\kappa p_z(l))^2 - 4\,(E^2(l)p_t^2(\nu) - \kappa^2)\cdot(E^2(l) - p_z^2(l))} && (5.9)\\ \text{with } \kappa &= 0.5\cdot(M_W^2 - m_l^2) + cos(\phi_l - \phi_\nu)\cdot p_t(l)\, p_t(\nu). && (5.10) \end{aligned}$$

  The missing energy $\not{E}_t$ is used as $p_t(\nu)$. Out of the two solutions the one with the smallest absolute value is choosen. If p_z turns out to be complex with non-zero imaginary part only the real part of p_z is used. The neutrino energy is calculated according to

$$E(\nu) = \sqrt{(\not{E}_t{}^2 + p_z^2)}. \tag{5.11}$$

- $\sum \eta_{WH}$, the sum of η_W and η_H, the pseudo-rapidities of the reconstructed W- and Higgs-Boson.

- $\Delta\eta_{WH}$, the difference of η_W and η_H.

- $M_{b\bar{b}}$, the invariant mass of the reconstructed Higgs, i. e. the invariant mass of the first and second jet.

- $M_{l\nu b\bar{b}}$, the invariant mass of the first two jets, the lepton and the reconstructed neutrino.

- δ_{12}, the event shape variable from the k_T jet algorithm [95].

Some of these variables need to be explained in more detail.

- C, the centrality, is defined by $C = \frac{\sum |E_t|}{\sum E}$ where E is the jet energy and E_t the jet transverse energy.

[10] These quantities are already provided in the TopNtuple [93].

- The sphericity S describes the shape of an event which can be sphere like, ellipsoid-like or rod-like. It is defined as $\frac{3}{2}(Q_1 + Q_2)$ where Q_1 and Q_2 are the smaller ones of the three eigenvalues of the sphericity tensor, i. e.

$$S = \frac{3}{2}(Q_1 + Q_2) = \frac{3}{2}\frac{\mathrm{Min}(\sum_k (P_\perp^k)^2)}{\sum_k |P^k|^2}. \tag{5.12}$$

  The sphericity tensor S_{ij} itself is given by

$$S_{ij} = \frac{\sum_{k=1}^{N} P_i^k P_j^k}{\sum_{k=1}^{N} |P^k|^2} \tag{5.13}$$

  where i and j are the usual x, y, z coordinate axis and the sum is over all the N observed particles. It is defined such that the sum of its eigenvalues is one, i. e. $Q_1 + Q_2 + Q_3 = 1$ [96]. For an event with two collinear back-to-back jets one of the eigenvalues will be much larger than the others and hence S will be quite small. On the other hand, an isotropic event will yield $Q_1 \approx Q_2 \approx Q_3 \approx 1/3$ and thus $S \approx 1$.

- The aplanarity A is given by $A = \frac{3}{2}Q_1$.

- The event shape variable δ_{12} from the k_T algorithm [95, 97] provides a measure at which an event changes from from having $n+1$ jets to having n jets and is given by

$$\delta_{i,j} = \mathrm{Min}\left(\frac{\mathrm{Min}(E_i^2, E_j^2)(1 - \cos\alpha_{i,j})}{E^2}\right), \quad i, j \in (n, n+1). \tag{5.14}$$

5.6.2 Network Training with weights

As was already mentioned in chapter 4.5 on page 79, NeuroBayes® allows to perform the network training with weights, i. e. a certain degree of acceptance is assigned to a training pattern.

Before the NB-Teacher was used to train a network for a particular Higgs mass, weights were calculated for the different background samples. No weights were applied to the signal samples during the training. The background events were weighted such that their total number was equal to the number of accepted signal events, i. e. the ratio was 1:1. Furthermore the background events were normalised w. r. t. to their expected number after the preselection. The net effect is a background sample having a composition one would expect in real data. This was done to improve the network learning since the task of the network was to perform a binary classification, i. e. to discriminate between the signal and one background class. Hence it does make sense to have the signal and the background class weighted 1:1 for the training process. Details on the weight calculation can be found in appendix B on page 135.

The weights for the NB-expert were calculated such that the signal and background numbers equal those of the preselection.

Rank	Variable	Correlation	σ
1	$E_t^{j_1}$	11.8	14.5
2	δ_{12}	17.4	21.4
3	P_t^{Lep}	7.69	9.47
4	$M_{l\nu b\bar{b}}$	13.9	17.2
5	$Q \times \eta$	12.5	15.3
6	$\not{E}_t$	7.92	9.75
7	$M_{b\bar{b}}$	3.74	4.61
8	ΔR_{j1j2}	3.47	4.27
9	$M_{l\nu b}$	2.68	3.29
10	$E_t^{j_2}$	1.84	2.26
11	ΔR_{lj_2}	1.84	2.26
12	η_{Lep}	1.48	1.82
13	A	1.22	1.51
14	$\Delta E_t^{j_1 j_2}$	1.1	1.36
15	$\delta\eta_{WH}$	0.74	0.91
16	$\sum \eta_{WH}$	0.914	1.12
17	ΔR_{lj_1}	0.672	0.82
18	S	0.581	0.71
19	C	0.512	0.63
20	H_t	0.315	0.38

Table 5.19: Variable ranking table for all input variables considered. The column denoted correlation shows the correlation coefficient to the target value and σ the significance. The Higgs mass of the signal sample was 120 GeV/c^2.

5.6.3 Optimisation study

NeuroBayes® allows to automatically determine the separation power of the input variables (cf. table 4.1 on page 79). In a first step, all variables descibed in the previous section were presented to the network. The ranking list shown in table 5.19 was obtained.

Starting from this ranking list different sets of input variables were defined starting with the four best variables. For the subsequent sets one variable was added at a time with the last one having 10 variables[11]. The different scenarios are listed in table 5.20 on the next page and named 'A' to 'G'.

These input variable scenarios were tested with various different parameter settings for the network training. The parameters that have been changed in order to find the combination that provides the best separation of signal and background as well as good linearity of the network output vs. purity are[12]:

- the number of training iterations,

[11] The NeuroBayes® feature of automatically choosing the best variables above a certain σ-cut could not be used, due to a bug in the C++ interface to the network package.

[12] Cf. figures 5.4 and 5.5 on page 107.

Scenario	N_{var}	Variables
A	4	$E_t^{j_1}$, δ_{12}, P_t^{Lep}, $M_{l\nu b\bar{b}}$
B	5	$E_t^{j_1}$, δ_{12}, P_t^{Lep}, $M_{l\nu b\bar{b}}$, $Q \cdot \eta$
C	6	$E_t^{j_1}$, δ_{12}, P_t^{Lep}, $M_{l\nu b\bar{b}}$, $Q \cdot \eta$, $\not{E}_t$
D	7	$E_t^{j_1}$, δ_{12}, P_t^{Lep}, $M_{l\nu b\bar{b}}$, $Q \cdot \eta$, $\not{E}_t$, $M_{b\bar{b}}$
E	8	$E_t^{j_1}$, δ_{12}, P_t^{Lep}, $M_{l\nu b\bar{b}}$, $Q \cdot \eta$, $\not{E}_t$, $M_{b\bar{b}}$, ΔR_{j1j2}
F	9	$E_t^{j_1}$, δ_{12}, P_t^{Lep}, $M_{l\nu b\bar{b}}$, $Q \cdot \eta$, $\not{E}_t$, $M_{b\bar{b}}$, ΔR_{j1j2}, $M_{l\nu b}$
G	10	$E_t^{j_1}$, δ_{12}, P_t^{Lep}, $M_{l\nu b\bar{b}}$, $Q \cdot \eta$, $\not{E}_t$, $M_{b\bar{b}}$, ΔR_{j1j2}, $M_{l\nu b}$, $E_t^{j_2}$

Table 5.20: List of input variable scenarios. In each case the variables are given in the order in which they were present to the network.

Parameter	Values
Iterations	100, 1k, 10k, 20k, 40k, 50k, 100k
N_{update}	100, 200
Nodes in 2^{nd} layer	$N_{var}+1$, $N_{var}+2$, $N_{var}+3$
Learn speed	1, 20, 50, 100, 200
Input variable scenarios	A, B, C, D, D, E, F, G

Table 5.21: Parameter list of the optimisation study.

- the number of events after which the network weights were updated,
- the number of nodes in the 2^{nd} layer and
- the learn speed.

Table 5.21 summarises the different parameter values.

5.6.4 Final Network Parameters and performance

In this section the network and its performance are described for $m_H = 120$ GeV/c^2. Plots and tables for the remaining Higgs masses can be found in the Appendix D on page 143.

The variation of the parameters in table 5.21 gave the best separation using scenario E and 100k iteration. The remaining network settings are shown in table 5.22 with option 12 used as a global preprocessing flag[13].

The error for the training and test sample and the regularisation loss function are shown in fig. 5.3 on page 106. They are decreasing throughout the learning process as they should. This is also the case for the other masses (cf. fig. D.3 and D.4 on page 147). Table 5.23 shows the correlation matrix of the network input variables in percent.

Fig. 5.4 on page 106 shows the network output for signal and background, weighted 1:1. One can see that the network was able to separate both event classes. On page 71 in

[13]Individual preprocessing has not been used in this study.

Parameter	
Input variable scenario	E (8 variables)
Preprocessing	12
Test sample fraction	20%
Iterations	100k
Task	'CLA'
Regularisation	'REG'
Loss function	'ENTROPY'
Def_Shape	'OFF'
N_{update}	100
Nodes in 2^{nd} layer	10
Learn speed	20
Maxlearn	1

Table 5.22: Final network parameter list. The same network parameters were used for all Higgs masses.

	T	$E_t^{j_1}$	δ_{12}	P_t^{Lep}	$M_{l\nu b\bar{b}}$	$Q \times \eta$	$\not{E}_t$	$M_{b\bar{b}}$
T	100	-	-	-	-	-	-	-
$E_t^{j_1}$	11.8	100	-	-	-	-	-	-
δ_{12}	6.3	65.7	100	-	-	-	-	-
P_t^{Lep}	8.2	22.7	25	100	-	-	-	-
$M_{l\nu b\bar{b}}$	3.2	70.4	67.2	54.3	100	-	-	-
$Q \times \eta$	-13	-0.5	0.6	1	-0.8	100	-	-
$\not{E}_t$	-6.4	24.9	25.5	-16.2	5.3	-1.4	100	-
$M_{b\bar{b}}$	6.3	60	46.1	-3	57.9	-2	7.8	100

Table 5.23: Correlation matrix of network input variables in percent for $m_H = 120$ GeV/c^2. Elements above the diagonal are left out because the matrix is symmetric. T denotes the network target value.

section 4.2.1 it has been shown that the purity as function of the network output should be linear for a perfectly trained network. Fig. 5.5 shows the corresponding plot for the network trained with $m_H = 120$ GeV/c^2. The efficiency vs. purity for the Neurobayes teacher is given in fig. 5.6. Fig. 5.7 on page 108 shows the same plot for the NB-expert. The final network cut values and mass windows can be found in table 5.24.

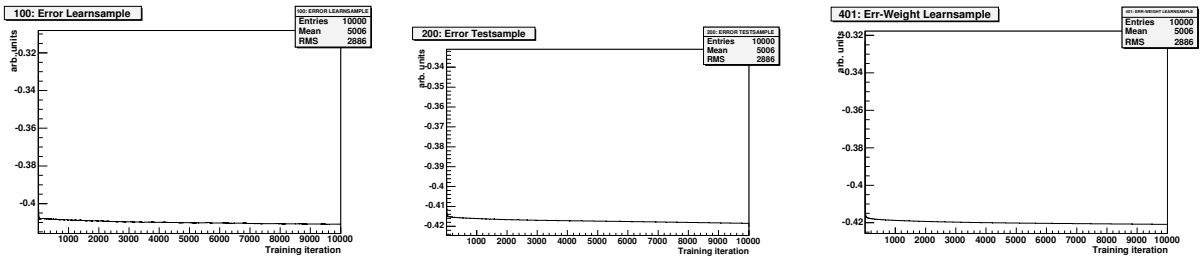


Figure 5.3: Error for the training (left) and test sample (middle) and the regularisation loss function (right) for a Higgs boson mass of $m_H = 120$ GeV/c^2. The latter one is the quantity that is actually minimised by the network. Twenty percent of the training sample was used as a test sample to independently determine the network error. The three quantities decrease throughout the network training as they should (cf. fig. 4.7 on page 69). The decreasing error for the test sample indicates that the network is not overtrainied.

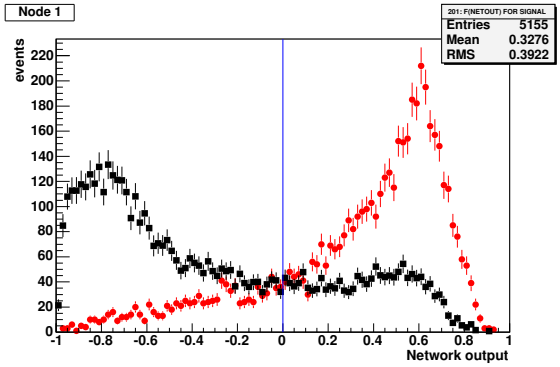


Figure 5.4: Network output of the NB-teacher for signal (red) and background (black). Signal and background are weighted 1:1. The Higgs mass of the signal was 120 GeV/c^2.

Higgs mass [GeV/c^2]	110	115	120	130	140	150
Network cut	0.0	0.0	0.0	0.0	0.0	0.0
m_{low} [GeV/c^2]	70	75	80	90	100	110
m_{high} [GeV/c^2]	150	155	160	170	180	190

Table 5.24: Values of the network output and invariant mass cut for all Higgs masses.

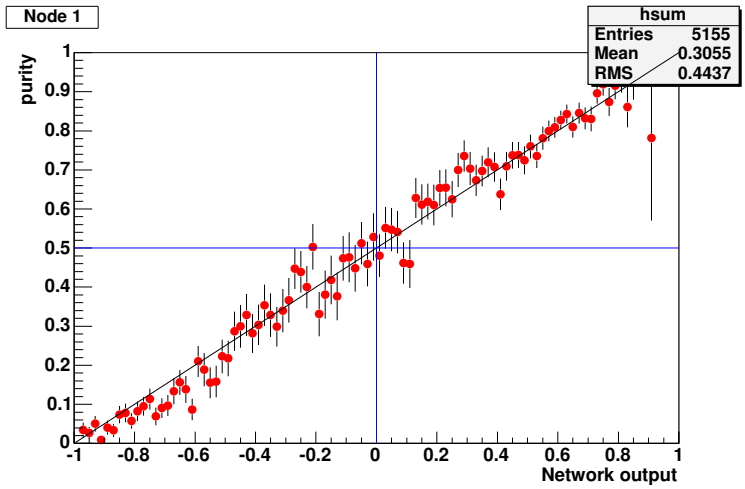


Figure 5.5: Network output vs. purity of the NB-teacher for a Higgs boson mass of $m_H = 120$ GeV/c^2 (cf. fig. 4.9 on page 72).

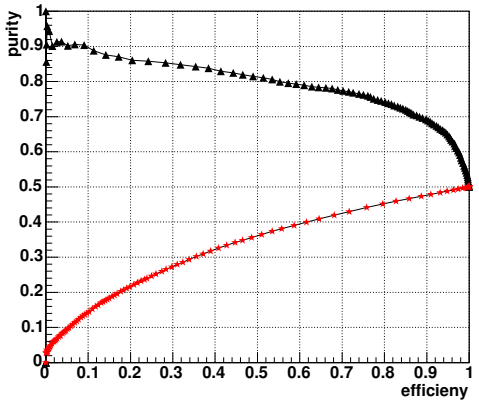


Figure 5.6: Efficiency vs. purity of the NB-teacher for signal (upper curve) and background (lower curve). The signal mass of the Higgs boson was 120 GeV/c^2. An efficiency of 100% corresponds to a cut value of -1 on the network output. Hence the number of signal and background events are equal and the purity is 50% for both.

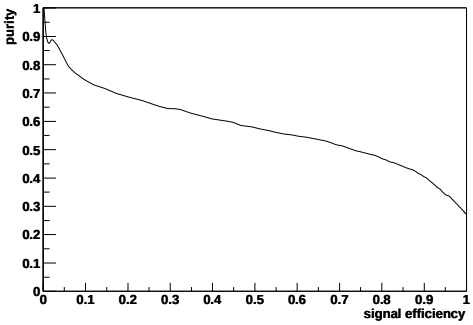


Figure 5.7: Efficiency vs. purity of the NB-expert for the signal. The signal mass of the Higgs boson was 120 GeV/c^2.

'It wants to be found.'
The Lord of the Rings
J.R.R Tolkien

Chapter 6

Results

The event expectation after the preselection cuts has already been shown in table 5.10 on page 93. The next step in the analysis is the usage of the neural networks. The optimized network parameters are listed in table 5.22 on page 105. Fig. 6.1 shows on the left the network output for 120 GeV/c^2. The remaining distributions are summarized in fig. 6.2 on page 111. The distributions for the signal processes scaled up by a factor of 100 are shown by the dashed line. As expected the signal tends to have high network output values whereas the background tends to a low network output. The values for the cuts made on these distributions are given in table 5.24 on page 106. In all cases the cut value was 0.0 . Table 6.1 on the next page shows the resulting event numbers.

The invariant mass spectra of events surviving the network cut are given in fig. 6.1 on the right for 120 GeV/c^2 and in fig. 6.3 on page 112. These spectra always lie in the region of the signal masses, i. e. the network is able to select the events in the mass region of interest. Although the expected number of signal events decreases due to the network cut the signal mass peak shown by the dashed line is retained. Cutting on these distributions is the last step of the analysis. The upper and lower limits of the mass window are shown in table 5.24. The event numbers obtained after this last cut are given in table 6.2 on the next page and are the basis for the exclusion limits calculated in the next section.

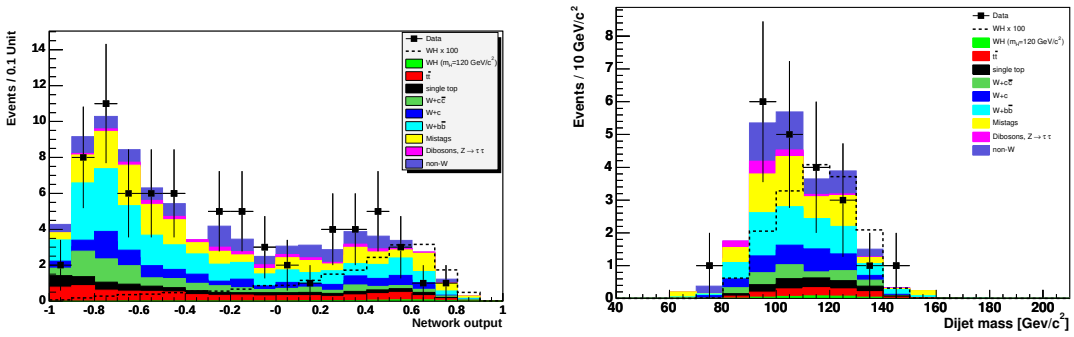


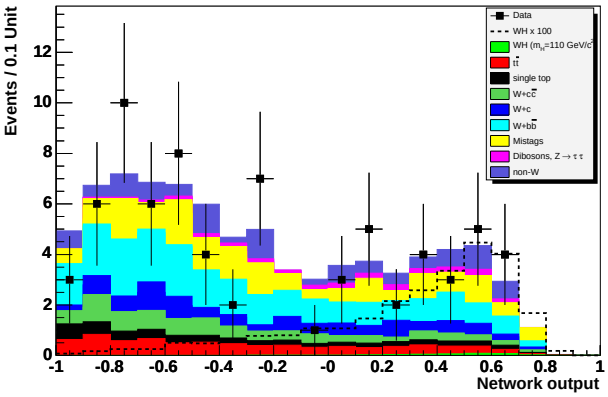
Figure 6.1: Network output (left) and invariant mass distribution (right) for $m_H = 120$ GeV/c^2. A cut on the network output of -0.2 was applied to obtain the dijet mass plot.

Process	Events					
m_H [GeV/c^2]	110	115	120	130	140	150
WH	0.42	0.36	0.32	0.26	0.21	0.17
$t\bar{t}$	2.13	2.28	2.28	2.65	2.81	3.042
W^*	0.66	0.67	0.67	0.72	0.71	0.705
Wg	0.71	0.72	0.72	0.78	0.87	0.87
$Wb\bar{b}$	6.32	5.77	5.37	5.04	4.19	3.72
$Wc\bar{c}$	1.9	1.78	1.68	1.45	1.2	1.08
Wc	3.26	3.12	2.86	2.67	2.07	1.68
Mistags	5.8	5.21	5.29	4.41	3.86	3.50
WW	0.044	0.037	0.031	0.017	0.01	0.004
WZ	0.7	0.58	0.48	0.36	0.2	0.12
ZZ	0.63	0.5	0.38	0.23	0.13	0.1
$Z \to \tau\tau$	0.086	0.12	0.086	0.1	0.086	0.086
QCD	4.43	3.97	3.97	4.2	3.27	2.57
Total background	26.68	24.76	23.8	22.64	19.4	17.48
Data	26	24	21	24	18	18

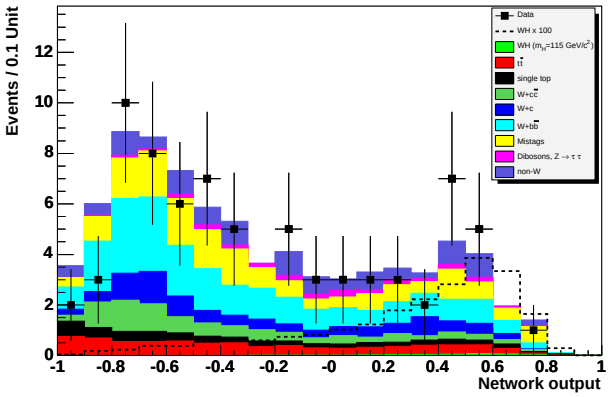
Table 6.1: Number of expected events after the application of the neural network. The network training parameters are given in table 5.21 on page 104. Table 5.24 on page 106 lists the cut values on the network output for each mass. At this point, no cut on the invariant mass distribution is applied.

Process	Events					
m_H [GeV/c^2]	110	115	120	130	140	150
WH	0.42	0.36	0.32	0.26	0.21	0.16
$t\bar{t}$	1.05	1.14	1.13	1.31	1.35	1.42
W^*	0.65	0.67	0.67	0.71	0.69	0.66
Wg	0.71	0.72	0.71	0.78	0.83	0.83
$Wb\bar{b}$	6.23	5.75	5.3	4.94	3.88	3.42
$Wc\bar{c}$	1.84	1.78	1.65	1.45	1.13	1.01
Wc	3.26	3.08	2.78	2.57	2.04	1.6
Mistags	5.65	5.21	5.13	4.4	3.58	3.14
WW	0.042	0.035	0.029	0.017	0.006	0.004
WZ	0.7	0.58	0.48	0.35	0.1553	0.08
ZZ	0.63	0.5	0.38	0.21	0.09	0.06
$Z\tau\tau$	0.086	0.12	0.086	0.1	0.086	0.086
QCD	4.43	3.97	3.73	4.2	3.27	2.57
Total background	25.28	23.55	22.07	21.03	17.1	14.87
Data	26	23	20	24	18	18

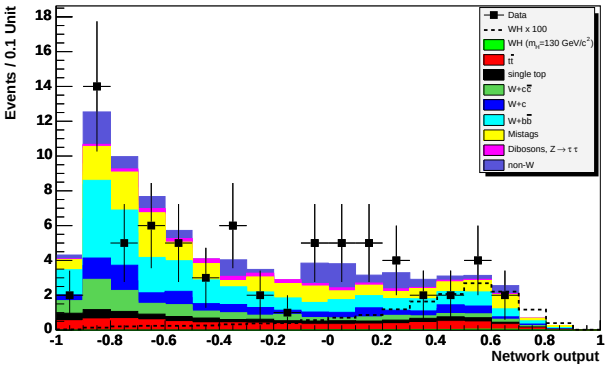
Table 6.2: Number of expected events after the application of the neural network and the cut on m_{bb}. The cuts on the network output value and the invariant mass distribution are shown in table 5.24 on page 106.



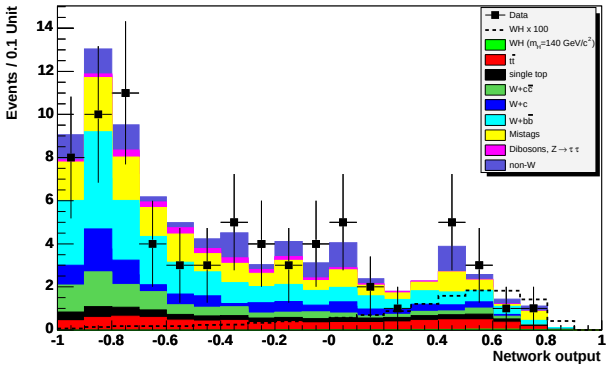
$m_H = 110$ GeV/c^2



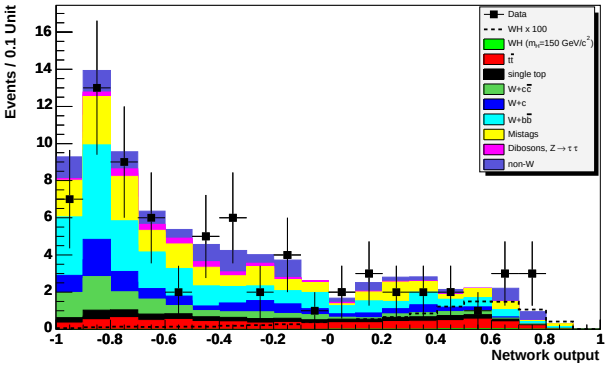
$m_H = 115$ GeV/c^2



$m_H = 130$ GeV/c^2

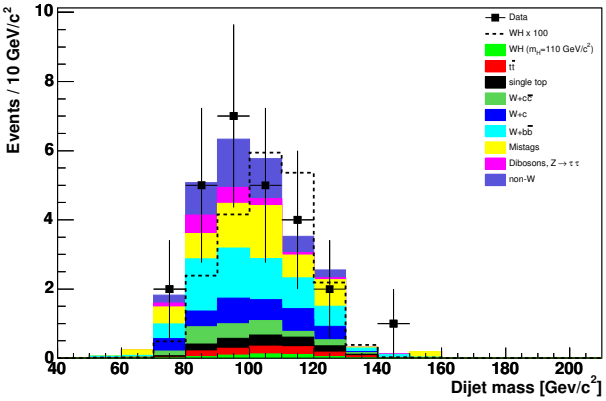


$m_H = 140$ GeV/c^2

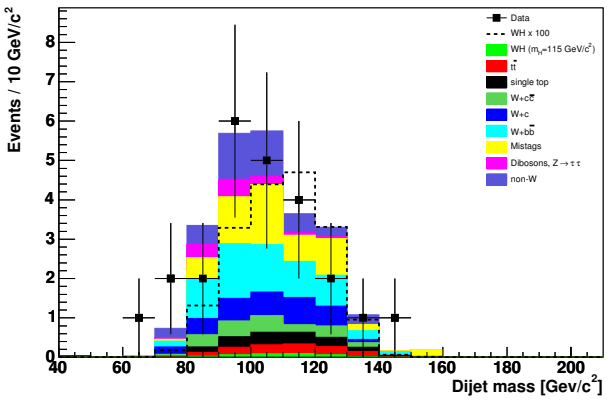


$m_H = 150$ GeV/c^2

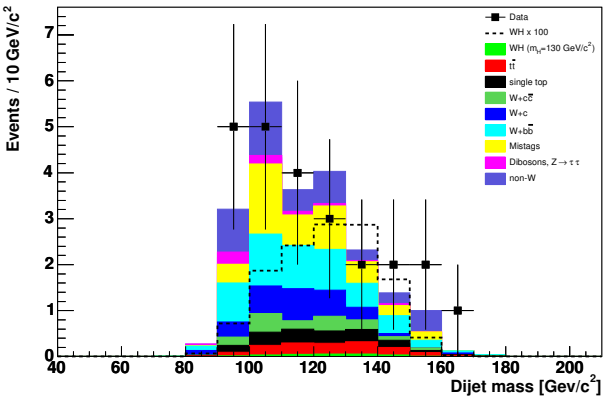
Figure 6.2: Network output for masses other than $m_H = 120$ GeV/c^2.



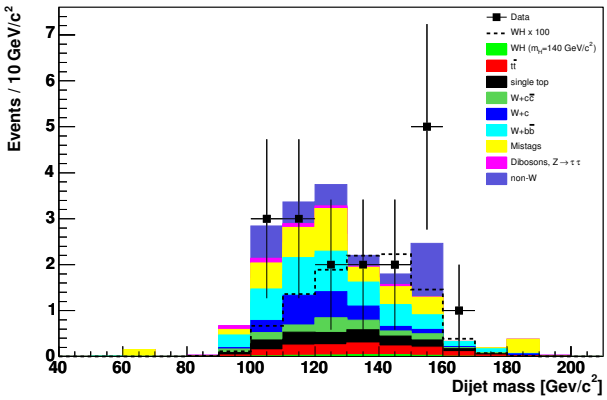
$m_H = 110$ GeV/c^2



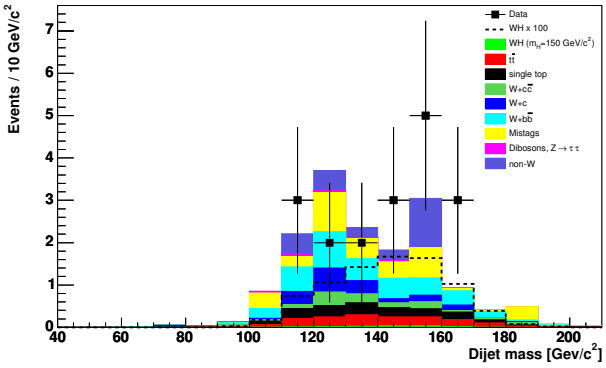
$m_H = 115$ GeV/c^2



$m_H = 130$ GeV/c^2



$m_H = 140$ GeV/c^2



$m_H = 150$ GeV/c^2

Figure 6.3: Invariant mass of the 2-jet system after the network cut for masses other than $m_H = 120$ GeV/c^2.

6.1 Limit calculation

Since no significant deviation from the standard model can be found in the mass spectrum, exclusion limits on the cross section of the WH process for Higgs bosons with a certain mass will be derived in this chapter. The method applied is described in ref. [98] and the source code for the limit calculation is available from [99].

6.1.1 Poisson statistics

The method to calculate limits herein is based on the Poisson distribution. Given a certain number of observed events n_0, the probability P for observing that number depends on the mean number μ of expected events and is given by

$$P(n_0, \mu) = \frac{\mu^{n_0} \cdot e^{-\mu}}{n_0!} \tag{6.1}$$

In particle searches one wants to determine the value of μ. Important properties of the Poisson distribution are [100]:

- the total probability is one, i. e. , $\sum_{n=0}^{\infty} P(n, \mu) = 1$,
- the mean number of events is $\langle n \rangle = \mu$ and
- the standard deviation is given by $\sigma = \sqrt{\mu}$.

The upper limit of μ is given by N[1], the value of μ for which there is some probability ϵ to observe n_0 or fewer events. The confidence level C.L. of this upper limit is given by $1 - \epsilon$. The value of ϵ is calculated by summing over the Poisson probabilities, i. e.

$$\epsilon = \sum_{n=0}^{n_0} P(n, \mu). \tag{6.2}$$

In order to calculate N, one varies μ until the desired value of ϵ is found, corresponding to the confidence level one wants.

If an average of μ_B background events is expected among the n_0 observed events, and if μ_B is known precisely one can calculate an upper limit N on the number of signal events present in the observed sample. Now N gives that value of μ_S, the mean number of signal events expected, for which the probability is $1-\epsilon$ that in a random experiment one would observe more than n_0 events and have $n_B \leq n_0$, where n_B is the number of background events in the sample. It can be calculated using

$$\epsilon = \frac{\sum_{n=0}^{n_0} P(n, \mu_B + N)}{\sum_{n=0}^{n_0} P(n, \mu_B)}. \tag{6.3}$$

N is varied until the proper value of ϵ is found.

[1] This value doesn't have to be an integer.

Higgs mass [GeV/c^2]	110	115	120	130	140	150
Preselection	35.67	33.02	29.36	35.44	32.73	32.03
NN w/o mass cut	7.89	7.35	6.89	5.88	4.92	4.63
NN + mass cut	5.75	5.40	5.13	4.86	4.00	3.73
Data w/o systematic	6.32	5.54	4.68	6.19	4.63	5.14
Data (final limit)	8.48	8.48	7.53	8.85	7.17	7.64

Table 6.3: Exclusion limit on the WH production cross section times the branching ratio $BR(H \to b\bar{b})$ as function of the Higgs boson mass.

6.1.2 Incorporating Uncertainties

Usually the number of data events and the expected background are not known precisely. Hence uncertainties have to be taken into account in the limit calculation.

Let's assume the the value of μ_B is known with a Gaussian uncertainty of σ_B and the acceptance A with an overall uncertainty of σ_A. The relative uncertainty on μ_S is given by σ_A/A. The Poisson upper limit N on μ_S can be defined as before, i. e. the value of the true μ_S for which one would observe more than n_0 events and have $n_B \leq n_0$. One determines the value of N such that

$$\epsilon = \frac{\sum_{n=0}^{n_0} \frac{1}{\sqrt{2\pi\sigma_N^2}} \int_0^\infty \int_0^\infty P(n, \mu'_B + \mu'_S) e^{-\frac{(\mu_B - \mu'_B)^2}{2\sigma_B^2}} e^{-\frac{(N - \mu'_S)^2}{2\sigma_N^2}} d\mu'_B d\mu'_S}{\sum_{n=0}^{n_0} \int_0^\infty P(n, \mu_B) e^{-\frac{(\mu_B - \mu'_B)^2}{2\sigma_B^2}} d\mu'_B} \tag{6.4}$$

with $\sigma_N = N\sigma_A/A$. This approach is applied in the Fortran program `poilim.f` that was used to calculate the limits presented in the next section.

6.2 Final Exclusion Limit

Limits are derived at six mass values ranging from 110 GeV/c^2 to 150 GeV/c^2 in steps of 10 GeV/c^2, as well as 115 GeV/c^2. This procedure is applied in order to map out the distribution. Counting the events for both data and background in the mass window provides the limit on the cross-section times the branching ratio. Fig. 6.4 on page 116 shows the exclusion limit on $\sigma \cdot BR$ at 95% confidence level. The corresponding numbers are shown in table 6.3.

Searches for the SM Higgs bosons at LEP have provided a lower mass limit and exclude Higgs masses below 114.1 GeV/c^2. In fig. 6.4 this area is shown in yellow. The three curves labeled with colored markers give the sensitivity from MC simulations after the preselection and the application of the neural network with and without a mass window around the signal mass. The empty squares show the data limits without systematic uncertainties. The final limits including systematics are given by the full black squares. The area above this curve is ruled out by this analysis. In total, cross sections of about 8 pb are excluded. For $m_H = 110$ GeV/c^2 the limit is too high by a factor of 35 to place

a lower limit on the Higgs mass. For higher masses this factor goes up since $\sigma \cdot BR$ decreases while the limit is almost constant.

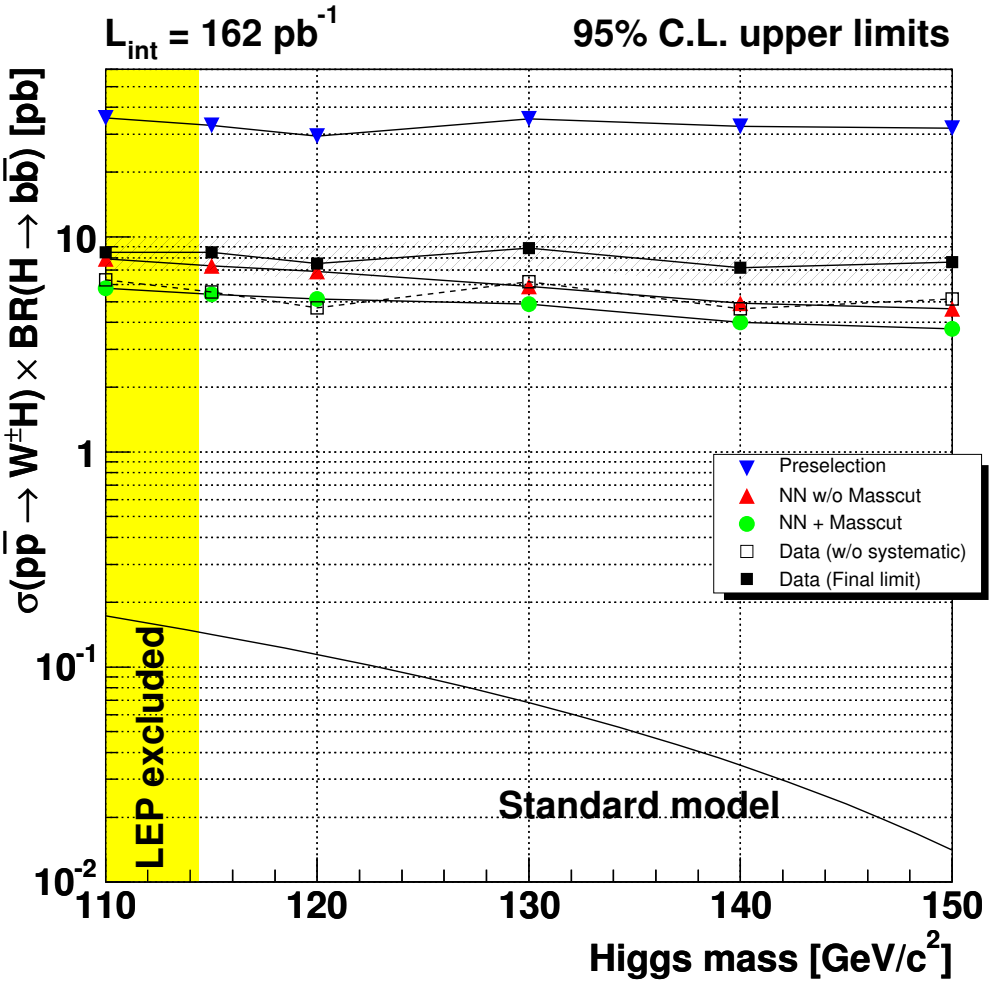


Figure 6.4: Exclusion limit on the WH production cross section times the branching ratio $BR(H \to b\bar{b})$ as function of the Higgs boson mass. The confidence level is 95%. The yellow area covers the mass region up to $m_H = 114.1$ GeV/c^2 that is excluded by LEP. The three curves labeled with colored markers show the expected sensitivity from MC after the preselection and the application of the neural network. The empty squares represent the data limit without systematic uncertainties. The final limit including systematic errors is given by the full black squares. The solid line at the bottom shows the standard model cross section for the process $WH \to l\nu b\bar{b}$.

Process	Events	
	Higgs report	This study
WH (Signal)	0.71	0.32
$t\bar{t}$	1.36	1.13
W^*	0.95	0.67
Wg	0.12	0.71
$Wb\bar{b}$	1.56	5.3
WZ	0.16	0.48
$\sum$	4.15	8.29
$Wc\bar{c}$	N/A	1.65
Wc	N/A	2.78
Mistags	N/A	5.13
WW	N/A	0.03
ZZ	N/A	0.38
$Z \rightarrow \tau\tau$	N/A	0.09
QCD	N/A	3.73
Total background	4.15	22.07

Table 6.4: Number of expected events from the Run II Higgs report and this study for $m_H = 120$ GeV/c^2 and a luminosity of $\mathcal{L}_{int} = 162$ pb^{-1}. The channels marked with N/A have not been looked at in the Run II report.

6.3 Comparison with other studies

6.3.1 Run II Higgs Report

In the year 1998 the Tevatron Higgs Working group studied the potential Higgs reach of the Tevatron in Run II [4]. Among other decay channels $WH \rightarrow l\nu b\bar{b}$ was studied by means of neural networks. The signal and background numbers for $m_H = 120$ GeV/c^2 are shown in table 6.4.

Only four background processes were considered in the Higgs report[2]. compared to 11 in this study. For these processes one sees differences in both the expected signal and the background event numbers that are discussed in the corresponding subsections on page 118 and 118. In general these differences are due to various assumptions that were made in the Run II study.

- Detector simulation:
  A special parameterized detector model was developed for the Workshop called 'SHW' instead of using the CDF and D0 detectors descriptions.

- b-tagging:
  The double b-tag efficiency was assumed to be in the range 30-40 %. This is higher than the current CDF single tag efficiency of 24% (cf. fig. 5.1 on page 90).

[2]Both W^* and Wg count as single-top production.

Identification	This study	Higgs report
Electron Trigger	CEM: 96.5 % [83]	100% [4, p.63]
Muon Trigger	CMX: 95.4 % [82] CMUP: 88.7 % [82]	100% [4, p.63]
B-tagging	27 % [84]	30-40% [4, p.64]

Table 6.5: Comparison of particle ID and trigger efficiencies. The numbers in square brackets are the corresponding references.

- Trigger efficiency:
  For most efficiencies a number of 100% was assumed in the Higgs report. A comparison with the actual efficiencies used herein is given in table 6.5.

- Cross sections:
  Some backgrounds differ due to different cross-section that were used in the Monte-Carlo generators. Others (especially $Wb\bar{b}$) were estimated using real data.

Signal process

The factor of almost two in the expected number of events between this study and the RunÍI Higgs report can be explained by the reasons mentioned above and are discussed here in more detail. The production cross-section for the WH process times the branching ratio into $l\nu b\bar{b}$ in the Run II report was assumed to be 0.045 pb. In this study a value of 0.035 pb was assumed for a Higgs mass of 120 GeV/c^2. This is due to different Monte-Carlo generators and the 25% difference is in good agreement with the difference in the Leading-Order and Next-to-Leading-Order cross sections from table A.1 on page 129. Another point is the difference in the b-tagging efficiency. Assuming an average of 27% for this study and 35% for the Run II report gives a 30% difference in the event.

Looking at table 6.5 assuming another 5% difference in trigger efficiency is reasonable. Table 6.6 summarizes the actual numbers and shows that multiplying these reduces the difference in signal numbers to 10%.

Backgrounds

The difference in the number of background events in table 6.4 is mostly given by the $Wb\bar{b}$ background. In the Run II study this background was simulated assuming a cross-section of 3.5 pb. Because this cross-section is difficult to calculate the background number for this study is measured from real data and was obtained using the so called 'Method 2' [85, 86].

Integrated luminosity

Based on the numbers from table 6.2 it is possible to calculate the integrated luminosities necessary for a 95% C.L. exclusion, a 3σ evidence and a 5σ discovery. The uncertainty

This study			Higgs report		
S	=	0.38	S^{HR}	=	0.713
σ	=	0.036	σ^{HR}	=	0.045
ϵ_{Trig}	=	0.95	ϵ^{HR}_{Trig}	=	1.00
ϵ_{b-tag}	=	0.27	ϵ^{HR}_{b-tag}	=	0.35

$$\left(\frac{S^{HR}}{S}\right)_{corrected} = \frac{S^{HR}}{S} \cdot \frac{\sigma}{\sigma^{HR}} \cdot \frac{\epsilon_{Trig}}{\epsilon^{HR}_{Trig}} \cdot \frac{\epsilon_{b-tag}}{\epsilon^{HR}_{b-tag}} \tag{6.5}$$

$$= \frac{0.38}{0.713} \cdot \frac{0.036}{0.045} \cdot \frac{0.95}{1} \cdot \frac{0.27}{35} \tag{6.6}$$

$$= 1.88 \cdot 0.8 \cdot 0.95 \cdot 0.77 \tag{6.7}$$

$$= 1.1 \tag{6.8}$$

Table 6.6: Comparison of signal event numbers. Multiplying the ratio S/S^{HR} of signal numbers before the $m_{b\bar{b}}$-cut by the different factors for trigger and b-tagging efficiency as well as the Monte-Carlo cross-sections reduces the original factor from 1.8 to 1.1. This justifies the difference in signal numbers between this study and the Higgs report.

on the background is assumed to be $\sqrt{B}$, the square root of the number of background events. For a 3σ evidence one requires

$$\frac{S_{final}}{\sqrt{B_{final}}} = 3. \tag{6.9}$$

Now one has to scale up the signal and background numbers by some factor f in order to obtain the desired significance. This factor is given by

$$\frac{S}{\sqrt{B}} = \frac{f \times S_{final}}{\sqrt{f \times B_{final}}} = \sqrt{f} \times \frac{S_{final}}{\sqrt{B_{final}}} = \sqrt{f} \times 3 \tag{6.10}$$

Hence the scale factor for a 3σ evidence is given by

$$f = \frac{1}{9} \cdot \frac{S^2}{B}. \tag{6.11}$$

For the ratio of two scale factors f_1 and f_2 one gets

$$\frac{f_1}{f_2} = \frac{\frac{S_1^2}{B_1}}{\frac{S_2^2}{B_2}} = \left(\frac{S_1}{S_2}\right)^2 \cdot \frac{B_2}{B_1} \tag{6.12}$$

Inserting the numbers from table 6.4 gives

$$\frac{f_1}{f_2} = \left(\frac{0.71}{0.32}\right)^2 \cdot \frac{22.07}{4.15} = 2.22^2 \cdot 5.32 = 26.18 \tag{6.13}$$

m_H [GeV/c^2]		110	120	130
95% C.L.	This study	31.27	45.83	68.40
	Higgs report	1.11	1.81	4.52
Ratio		28.17	25.32	15.13
3σ	This study	110.62	153.40	228.90
	Higgs report	3.71	6.04	15.12
Ratio		29.82	25.40	15.14
5σ	This study	290.61	426.00	635.82
	Higgs report	10.31	16.79	42.00
Ratio		28.19	25.37	15.14

Table 6.7: Integrated luminosities in fb^{-1} per experiment for a 95% C.L. exclusion, 3σ evidence and 5σ discovery. The numbers correspond to the lower edges of the colored bands in fig. 6.5 and are read off the left scale.

where f_1 denotes the scale factor for Higgs report numbers and f_2 the one for this study. The scale factors f differ by a factor of 26 for the 3σ and 120 GeV/c^2 case. This factor can also by read off the graph in fig. 6.5. Similar numbers apply for the other limits and masses. They can be seen in table 6.7.

Fig. 6.5 shows the integrated luminosities per experiment for a 95% C.L. exclusion, 3σ evidence and 5σ discovery. The lower bands are the results of the Run II Higgs report. The shaded curves show the results of this study. As in the Higgs report, no systematic uncertainties were considered and the values for the lower edges of the curves were scaled up by 30% to obtain their width. The scale on the left can be used for a direct comparison of the results, i. e. with the left scale one compares the results for the WH channel with the Run II result for the WH channel only.

However the curves can also be interpreted in a different way. When the WH channel is combined with the ZH channels one gains a factor of about two in the number of signal events. The right scale gives the required luminosities for this hypothetical combination of the results for the WH and ZH channels hence showing the potential improvement in the required luminosity.

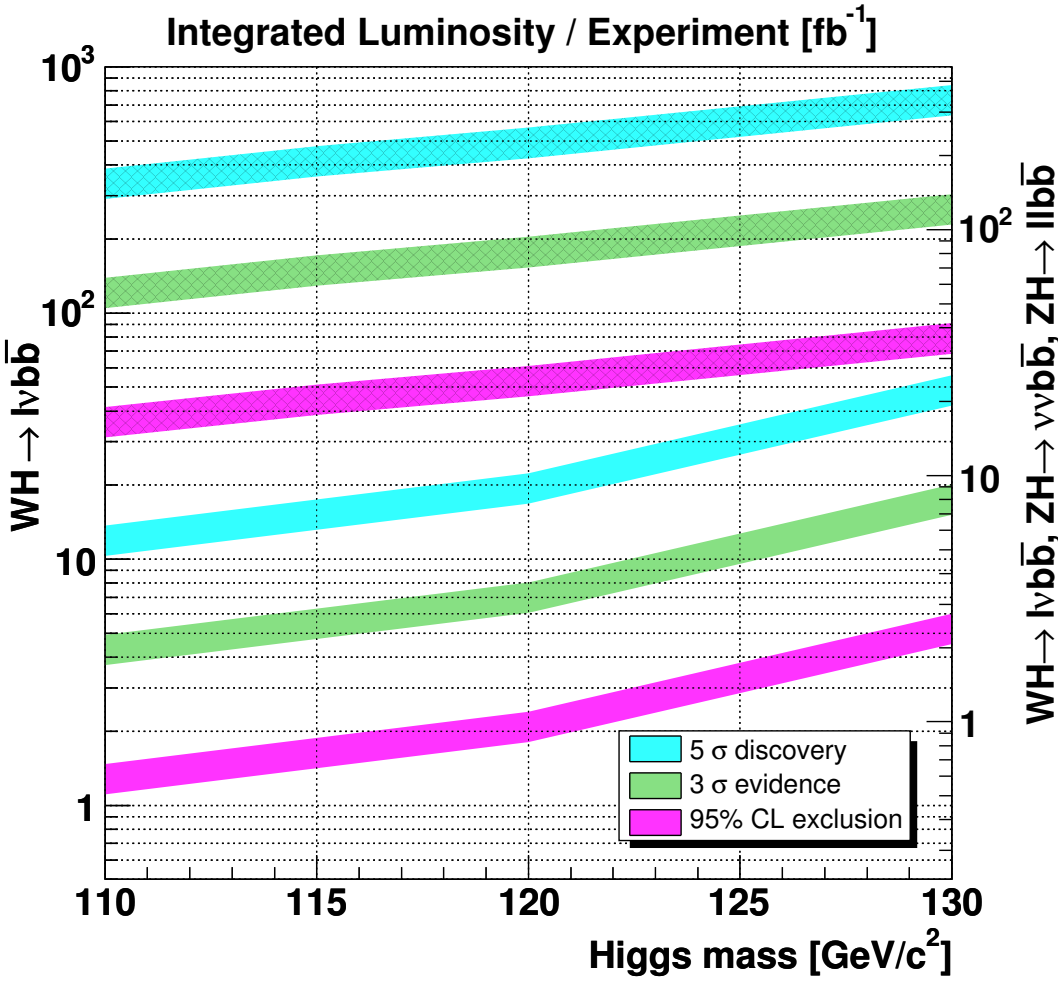


Figure 6.5: Integrated luminosities per experiment for a 95% C.L. exclusion, 3σ evidence and 5σ discovery. The three lower bands are the result of the Run II Higgs report. The three shaded upper curves show the result of this study. The left scale compares the results for the WH channel from this work with the Run II result for the WH channel only. The right scale gives the luminosity for a hypothetical combination of the WH channel with the ZH channels, hence decreasing the required luminosities by a factor of about two. This is due to the expected increase in $S/\sqrt{B}$ when combining the different channels.

'We will send word when we
may and some of us may yet
meet at times but I fear that we
shall not all be gathered
together ever again.'
The Lord of the Rings
J.R.R. Tolkien

Chapter 7

Conclusion and Outlook

7.1 Conclusion

A search for a low-mass SM Higgs-Boson in the channel $WH \rightarrow l\nu b\bar{b}$ has been performed using neural networks. The data was taken by the CDF experiment at the $p\bar{p}$ collider Tevatron from 2000-2003, corresponding to in integrated luminosity of $\mathcal{L}_{int} = 162$ pb^{-1} at a CMS-energy of $\sqrt{s} = 1.96$ TeV. A good agreement between the data and the Monte Carlo simulations is found for both event numbers and kinematic distributions. 95% confidence level upper limits were set on $\sigma \times BR$, the product of the production cross section times the branching ratio, as a function of the Higgs boson mass. Values larger than approx. 8 pb can be excluded for six different Higgs masses between 110 GeV/c^2 and 150 GeV/c^2. At this point the sensitivity is insufficient to place a lower limit on the Higgs boson mass since the derived cross-section limit lies above the Standard Model expectation.

As shown in fig. 6.5 combining WH with ZH would require an estimated luminosity of ~ 70 fb^{-1} for a 3 σ discovery of a SM Higgs at 115 GeV. Below 114.1 GeV SM Higgs masses are excluded by the LEP experiments [5]. These 70 fb^{-1} are unfortunately above the expected integrated luminosity of the Tevatron collider in the coming years, where at most a luminosity of ~ 9 fb^{-1} can be hoped for (cf. section 2.1.5 on page 23). This estimate for a 3σ discovery is roughly a factor 25 above the one estimated by the Tevatron Higgs working group mainly because in this thesis many more background channels were included and measured efficiencies were used.

7.2 Outlook

Future improvements on the limit can be achieved simply by using a larger dataset once it will be available. Furthermore, instead of calculating the exclusion limit based on event counting one could use advanced techniques. Fitting the shapes of the invariant mass spectra would be an example. In addition the analysis can be improved by using specific energy corrections for jets originating from b-quarks [101, 102]. This can reduce the width of the Higgs mass peak. A higher b-tagging efficiency would as well improve

the results. Neural networks can be applied for this purpose, a method that is under development at present [6]. Improving the detector alignment would also increase the b-tagging performance. An increased acceptance could be achieved be extending the η range in the detector. So called "phoenix electrons", i. e. electrons reconstructed in the CDF forward calorimeters, could be included. However, the background might rise faster than the signal in these regions and one has to check if the analysis would really benefit from this extension.

The channel considered in this thesis is not the only process that is worth looking into. The process ZH for instance, is investigated at present. Combining several search channels (also with the ones investigated by D0) would further enhance the Higgs prospects at the Tevatron.

If the Higgs boson is not going to be found at CDF and D0, the next place to look for it will be the LHC. The proton-proton collider is located in Geneva, Switzerland. Currently it is under construction and supposed to be operational in 2007. Two multi-purpose experiments, CMS and ATLAS, will be looking for new physics. Due to the high luminosity and a center-of-mass-energy of 14 TeV the prospects for a Higgs discovery are promising.

Appendix A

Cross section calculation

The tables presented in this section were made with HIGLU (v2.1), HDECAY (v3.0) [42] and V2HV. The source code for all programs is available from [43].

After presenting the input files for all three programs, tables A.1 on page 129 and A.2 on page 130 present the LO and NLO Higgs cross sections for the $gg \to H$, WH and ZH processes. Higgs branching ratios for $H \to b\bar{b}$, $\tau^+\tau^-$, $\mu^+\mu^-$, $s\bar{s}$ and $c\bar{c}$ are shown in tables A.3 and A.4 on pages 131 and 132, respectively . The last two tables, A.5 on page 133 and A.6 on page 134, list the Higgs branching ratios for $H \to gg$, $\gamma\gamma$, $Z\gamma$, WW and ZZ as well as the total decay width.

The following input file was used for HIGLU to calculate the $gg \to H$ cross section.

```
PROCESS:  0 = GG --> H         1 = H --> GG
=======
PROCESS  = 0
-----------------------------------------------------------------
COLLIDER:   0 = P P              1 = P PBAR
========
COLLIDER = 1
-----------------------------------------------------------------
TOTAL ENERGY: [TEV]
=============
ENERGY   = 1.96D0
-----------------------------------------------------------------
MODEL:    0 = SM   MSSM: 1 = SUBHPOLE  2 = SUBH
======                   3 = FEYNHIGGSFAST
MODEL    = 0
-----------------------------------------------------------------
TAN(BETA): (MSSM)
==========
TANBETA  = 3.D0
-----------------------------------------------------------------
COUPLINGS:  G_B = BOTTOM    G_T = TOP
==========  (MODEL = 0)
G_B      = 1.D0
G_T      = 1.D0
-----------------------------------------------------------------
QUARK MASSES: [GEV]
=============
```

```
M_SB(1)  = 0.190D0
M_C      = 1.5D0
M_B      = 5.0D0
M_T      = 175.D0
-----------------------------------------------------------------
HIGGS TYPE AND MASS [GEV]: 1 = HEAVY SCALAR 2 = PSEUDOSCALAR 3 = LIGHT SCALAR
=========================  INDIVIDU = 0: M_HIGGS = M_A
TYPE     = 3
INDIVIDU = 1
M_HIGGS  = DUMMYHIGGSMASS.D0
-----------------------------------------------------------------
SCALES: [GEV] MU = MU_1*M_HIGGS + MU_2:   RENORMALIZATION SCALE
=======        Q =  Q_1*M_HIGGS +  Q_2:     FACTORIZATION SCALE
MU_1     = 1.D0
MU_2     = 0.D0
Q_1      = 1.D0
Q_2      = 0.D0
-----------------------------------------------------------------
ORDER OF ALPHA_S:  1 = LO     2 = NLO
=================
LOOP     = 2
-----------------------------------------------------------------
DEFINITION OF ALPHA_S:  1 = ALPHA_S (M_Z)   2 = BY LAMBDA (N_F)
======================
CHOICE   = 1
-----------------------------------------------------------------
ALPHA_S (M_Z):
==============
ALPHA_S  = 0.116D0
-----------------------------------------------------------------
LAMBDA_NF: [GEV] (QCD SCALE)
==========
N_F      = 5
LAMBDA   = 0.226D0
-----------------------------------------------------------------
NUMBER OF EXTERNAL LIGHT FLAVORS: (FOR H --> GG)
=================================
N_EXT    = 5
-----------------------------------------------------------------
VEGAS:       ABSERR = ABSOLUTE ERROR
======       POINTS = NUMBER OF CALLS
             ITMAX  = NUMBER OF ITERATIONS
             PRINT  = PRINT OPTION FOR INTERMEDIATE VEGAS-OUTPUT
                      0            1             10
                      NO OUPUT     PRETTYPRINT   TABLE

ABSERR   = 0.D0
POINTS   = 1000
ITMAX    = 5
PRINT    = 10
-------------------------------------------------
STRUCTURE FUNCTIONS: STFUN: 0 = PDFLIB  1 = GRV
==================== SCHEME = FACTORIZATION SCHEME: 0 = MSBAR   1 = DIS
STFUN   =  1
SET     =  2
SCHEME  =  0
```

```
NGROUP  =  4
NSET    =  34
-------------------------------------------------------------
PARAMETERS: (FOR RAD. CORR. OF SUSY-COUPLINGS)
===========
            MZ   = Z-MASS [GEV]
            MSQ  = SQUARK-MASS [GEV]
GF       =  1.16639D-5
MZ       =  91.187D0
MW       =  80.41D0
MGLUINO  =  1000.D0
MSQ      =  1000.D0
MUR      =  1000.D0
MDR      =  1000.D0
M2       =  1000.D0
MU       =  100.D0
AU       =  0.D0
AD       =  0.D0
```

This input file was used for V2HV to calculate the WH and ZH cross sections.

```
MSSM     = 0
TGBET    = 3.D0
HIGGS    = 2
MA1      = 141.D0
MA2      = 200.D0
NA       = 60
Z=1/W=2  = 1
ENERGY   = 1960.D0
PP/PPBAR = 1

SCALE1   = 1.0D0
SCALE2   = 1.0D0
NSCALE   = 1
EPSILON  = 1.D-8

NGROUP   = 4
NSET     = 34

IPOINT   = 10000
ITERAT   = 5
NPRN     = 10

MC       = 1.5D0
MB       = 5.D0
MT       = 175.D0
NO       = 5
LAMBDA   = 0.202D0
LOOP     = 2
MW       = 80.41D0
MZ       = 91.187D0
SW2      = 0.2315D0
GF       = 1.16639D-5

MSQ      = 1000.D0
MUR      = 1000.D0
```

```
AT       = 0.D0
AB       = 0.D0
MU       = 0.D0
```

This last file was used for HDECAY to calculate the branching ratios.

```
HIGGS    = 0
MODEL    = 1
TGBET    = 30.D0
MABEG    = 80.D0
MAEND    = 200.D0
NMA      = 121
ALS(MZ)  = 0.118D0
MSBAR(1) = 0.190D0
MC       = 1.5D0
MB       = 5.D0
MT       = 175.D0
MTAU     = 1.7771D0
MMUON    = 0.105658389D0
1/ALPHA  = 137.0359895D0
GF       = 1.16639D-5
GAMW     = 2.080D0
GAMZ     = 2.490D0
MZ       = 91.187D0
MW       = 80.41D0
VUS      = 0.2205D0
VCB      = 0.04D0
VUB/VCB  = 0.08D0
MU       = 300.D0
M2       = 200.D0
MGLUINO  = 1000.D0
MSL1     = 1000.D0
MER1     = 1000.D0
MQL1     = 1000.D0
MUR1     = 1000.D0
MDR1     = 1000.D0
MSL      = 1000.D0
MER      = 1000.D0
MSQ      = 1000.D0
MUR      = 1000.D0
MDR      = 1000.D0
AL       = 1500.D0
AU       = 1500.D0
AD       = 1500.D0
NNLO (M) = 0
ON-SHELL = 0
ON-SH-WZ = 0
IPOLE    = 0
OFF-SUSY = 0
INDIDEC  = 0
NF-GG    = 5
IGOLD    = 0
MPLANCK  = 2.4D18
MGOLD    = 1.D-13
```

M(H)	gg->H		WH		ZH	
	LO	NLO	LO	NLO	LO	NLO
80	0.830001	2.05355	0.480773757	0.599457491	0.261416568	0.325369273
81	0.803386	1.98505	0.463219006	0.577603458	0.252427403	0.314224663
82	0.777919	1.91956	0.446250169	0.556607489	0.243720399	0.303432026
83	0.753533	1.85691	0.430071462	0.535923512	0.235401678	0.292783596
84	0.730169	1.79697	0.414567526	0.517404605	0.227397964	0.283295723
85	0.707769	1.73964	0.399704928	0.498990235	0.219718302	0.273744427
86	0.686280	1.68472	0.385579235	0.481343108	0.212404306	0.264668642
87	0.665655	1.63208	0.371934682	0.464766453	0.205331121	0.256043147
88	0.645848	1.58154	0.35900646	0.448551509	0.198621042	0.24760936
89	0.626815	1.53305	0.346283851	0.432856901	0.191971798	0.239434828
90	0.608519	1.48653	0.33450083	0.418060107	0.185830322	0.231742031
91	0.590920	1.44185	0.322878795	0.403793135	0.179746477	0.224259989
92	0.573986	1.39891	0.311699565	0.390324165	0.173876173	0.217244767
93	0.557682	1.35766	0.301285651	0.377190439	0.168395362	0.210326454
94	0.541980	1.31796	0.290959506	0.364391477	0.16296852	0.203607904
95	0.526849	1.27976	0.28109239	0.352206792	0.157753737	0.197170168
96	0.512265	1.24300	0.271734729	0.340477286	0.152813838	0.19100153
97	0.498200	1.20760	0.262698573	0.32940096	0.14802847	0.185177802
98	0.484631	1.17348	0.253979566	0.318407455	0.143400642	0.179308842
99	0.471537	1.14059	0.245574296	0.308101656	0.138925351	0.173834292
100	0.458895	1.10889	0.237614181	0.298115881	0.134690291	0.168526354
101	0.446686	1.07834	0.229772479	0.288311769	0.130494659	0.163287706
102	0.434891	1.04884	0.22236354	0.279200978	0.126547042	0.158439115
103	0.423492	1.02036	0.215110978	0.270222587	0.12264536	0.153651833
104	0.412472	0.992854	0.208302092	0.261693301	0.118988355	0.149073118
105	0.401816	0.966280	0.201637179	0.253343791	0.115400642	0.144565078
106	0.391508	0.940618	0.19526758	0.245576014	0.11196379	0.140393652
107	0.381534	0.915806	0.189077634	0.237786514	0.108620502	0.136185221
108	0.371881	0.891804	0.183060116	0.230436839	0.105360211	0.13222239
109	0.362535	0.868620	0.17734603	0.223232436	0.102262512	0.128328694
110	0.353484	0.846175	0.171868313	0.216400374	0.0992908441	0.124595566
111	0.344716	0.824386	0.166535516	0.209677165	0.096382366	0.120949748
112	0.336221	0.803482	0.161402057	0.203483114	0.0935893341	0.117616548
113	0.327988	0.783195	0.156484981	0.197168217	0.0909008737	0.114183408
114	0.320007	0.763523	0.151621588	0.191172866	0.0882454985	0.110898636
115	0.312268	0.744449	0.147030728	0.185380713	0.0857220165	0.107699712
116	0.304763	0.725949	0.142639383	0.179936725	0.0833108798	0.104740749
117	0.297482	0.708002	0.138295201	0.174441207	0.0809185769	0.101710111
118	0.290417	0.690641	0.134161621	0.169352293	0.0786383239	0.0989172889
119	0.283560	0.673833	0.130122147	0.164379814	0.0764129551	0.096196832
120	0.276905	0.657537	0.126281682	0.159655441	0.0742871213	0.0935814862
121	0.270442	0.641731	0.122555842	0.154832147	0.0722221474	0.0909105757
122	0.264167	0.626407	0.11894958	0.150393507	0.0702174621	0.0884566135
123	0.258071	0.611536	0.11546873	0.146020152	0.0682850375	0.0860456726
124	0.252149	0.597097	0.112096508	0.141749734	0.0664061715	0.0836615427
125	0.246395	0.583073	0.108844965	0.137769785	0.064592133	0.0814482402
126	0.240803	0.569458	0.105671336	0.133754418	0.0628162945	0.0792064279
127	0.235367	0.556235	0.102636353	0.129950636	0.061118135	0.0770812149
128	0.230082	0.543391	0.0996484329	0.126243611	0.0594367527	0.0749940401
129	0.224942	0.530915	0.0968240041	0.122653347	0.0578557072	0.0730069704
130	0.219944	0.518793	0.0940495355	0.11918354	0.0562930523	0.0710493387
131	0.215082	0.507012	0.0914382944	0.115948408	0.0548210525	0.0692395211
132	0.210352	0.495558	0.0888392156	0.112701057	0.053351846	0.0674170001
133	0.205749	0.484431	0.0863013497	0.109477813	0.05191627	0.0655893629
134	0.201270	0.473600	0.0839000031	0.106494913	0.0505558769	0.0639065688
135	0.196910	0.463075	0.0815563328	0.103560469	0.0492247531	0.0622455531
136	0.192666	0.452834	0.0792657853	0.100691224	0.0479229179	0.0606248942
137	0.188533	0.442874	0.0770944532	0.0979427599	0.0466879261	0.05906039
138	0.184508	0.433177	0.0749804513	0.0952743487	0.0454825841	0.05754519
139	0.180589	0.423757	0.0728943932	0.0926696608	0.044289569	0.0560540771
140	0.176771	0.414584	0.0709050579	0.0901599776	0.0431512223	0.0546280401

Table A.1: LO and NLO Higgs cross sections for $gg \to H$, $q\bar{q} \to WH$ and $q\bar{q} \to ZH$ for Higss masses from 80 GeV/c^2 to 140 GeV/c^2.

M(H)	gg->H		WH		ZH	
	LO	NLO	LO	NLO	LO	NLO
141	0.173051	0.405659	0.0689577704	0.0877618466	0.0420349026	0.0532544298
142	0.169427	0.396967	0.0670997295	0.0854328671	0.0409683275	0.0519261989
143	0.165895	0.388496	0.0652710108	0.0831103605	0.0399185279	0.0505984122
144	0.162452	0.380244	0.0635087082	0.0808323661	0.0389032134	0.0492882372
145	0.159097	0.372209	0.0617987299	0.0787434018	0.0379159745	0.0480901176
146	0.155826	0.364381	0.0601378216	0.0766675968	0.0369553122	0.0468930899
147	0.152636	0.356760	0.0585450811	0.0746169391	0.03603408	0.0457088959
148	0.149526	0.349338	0.0569862318	0.0726768477	0.0351311815	0.0445948028
149	0.146493	0.342098	0.055497978	0.0707931335	0.034268671	0.0435054519
150	0.143534	0.335045	0.0539996608	0.0688937775	0.0333955439	0.0424052498
151	0.140648	0.328167	0.0526125366	0.0671658752	0.0325899496	0.0414010003
152	0.137832	0.321466	0.0512203863	0.0653864917	0.0317773472	0.0403681379
153	0.135085	0.314933	0.0498618655	0.0637122885	0.0309823659	0.0393966054
154	0.132403	0.308562	0.0485969603	0.0620949264	0.0302409451	0.0384485241
155	0.129787	0.302344	0.0473124427	0.0605058153	0.0294904336	0.0375240938
156	0.127233	0.296280	0.0460799227	0.0589547111	0.0287665196	0.0366185361
157	0.124739	0.290364	0.0448995234	0.0574515113	0.0280739579	0.0357392379
158	0.122305	0.284593	0.0437472771	0.0559907116	0.0273963812	0.034884666
159	0.119929	0.278963	0.0426236988	0.0545689355	0.0267337747	0.0340499732
160	0.117608	0.273467	0.041526647	0.0531890444	0.0260858253	0.0332361099
161	0.115341	0.268101	0.0404860595	0.051862892	0.0254707025	0.032454493
162	0.113128	0.262863	0.0394395954	0.0505302426	0.0248488351	0.0316669658
163	0.110965	0.257756	0.0384489395	0.0492845057	0.0242651475	0.0309383782
164	0.108853	0.252769	0.0374647503	0.0480313982	0.0236785178	0.0301922672
165	0.106789	0.247901	0.0365375135	0.0468569943	0.0231271938	0.0294986549
166	0.104772	0.243146	0.0356179938	0.0456975435	0.0225792617	0.0288107949
167	0.102801	0.238499	0.0347335086	0.0445917096	0.0220517337	0.0281574257
168	0.100876	0.233958	0.0338633006	0.0434831186	0.0215315487	0.0274959843
169	0.989932E-01	0.229523	0.0330075669	0.0424077307	0.021018089	0.0268552956
170	0.971532E-01	0.225191	0.0321926738	0.0413754123	0.0205304138	0.0262397817
171	0.953544E-01	0.220960	0.0314042205	0.0403729982	0.0200581924	0.0256431832
172	0.935959E-01	0.216824	0.0306272016	0.0393720792	0.019591451	0.0250419632
173	0.918764E-01	0.212784	0.0298776471	0.0384337904	0.0191395998	0.0244827023
174	0.901951E-01	0.208830	0.0291508627	0.0374912744	0.0187006256	0.0239158245
175	0.885509E-01	0.204966	0.0284231234	0.0365828059	0.0182634804	0.0233719122
176	0.869428E-01	0.201190	0.0277340506	0.0357059749	0.0178462281	0.0228428375
177	0.853699E-01	0.197498	0.0270713267	0.0348660153	0.0174448875	0.0223372017
178	0.838314E-01	0.193887	0.026406129	0.0340128555	0.0170410211	0.0218201641
179	0.823263E-01	0.190358	0.0257707702	0.033215249	0.0166548045	0.0213398212
180	0.808537E-01	0.186906	0.0251440156	0.0324236727	0.0162744287	0.0208615584
181	0.794129E-01	0.183531	0.0245442898	0.0316710279	0.0159094296	0.0204063383
182	0.780031E-01	0.180230	0.0239578907	0.0309042537	0.0155520162	0.0199408799
183	0.766235E-01	0.176998	0.0233842739	0.0301841024	0.0152012654	0.0195036526
184	0.752733E-01	0.173837	0.022829791	0.0294748023	0.0148626138	0.0190719937
185	0.739518E-01	0.170747	0.0222865944	0.0287760586	0.0145299219	0.0186458767
186	0.726582E-01	0.167724	0.0217600207	0.028114092	0.0142071239	0.018243837
187	0.713920E-01	0.164764	0.0212419152	0.027452807	0.0138880923	0.0178389553
188	0.701523E-01	0.161869	0.0207424389	0.0268087022	0.0135814095	0.0174441605
189	0.689386E-01	0.159033	0.0202452988	0.0261869436	0.0132744161	0.017062279
190	0.677503E-01	0.156258	0.0197767506	0.0255719554	0.0129868182	0.0166892916
191	0.665867E-01	0.153543	0.0193094814	0.0249839776	0.0126974654	0.016325932
192	0.654472E-01	0.150885	0.018869956	0.0244272728	0.0124260965	0.0159836051
193	0.643312E-01	0.148283	0.0184280478	0.0238604251	0.0121517129	0.0156341818
194	0.632382E-01	0.145734	0.0179908097	0.02330046	0.0118810834	0.0152889938
195	0.621676E-01	0.143239	0.0175773268	0.0227743346	0.0116239764	0.0149649155
196	0.611189E-01	0.140796	0.0171706963	0.0222558481	0.0113711562	0.0146430175
197	0.600916E-01	0.138406	0.0167699438	0.021747679	0.0111216823	0.014327918
198	0.590851E-01	0.136064	0.0163892058	0.021257188	0.0108845522	0.0140247917
199	0.580990E-01	0.133772	0.0160160943	0.0207759979	0.01065177	0.0137273435
200	0.571328E-01	0.131526	0.0156440098	0.0203048179	0.0104187438	0.0134332184

Table A.2: LO and NLO Higgs cross sections for $gg \to H$, $q\bar{q} \to WH$ and $q\bar{q} \to ZH$ for Higgs masses from 141 GeV/c^2 to 200 GeV/c^2.

M(H)	bb	tau tau	muon muon	ss	cc
80	0.8604	0.6654E-01	0.2313E-03	0.5226E-03	0.3808E-01
81	0.8595	0.6664E-01	0.2316E-03	0.5220E-03	0.3803E-01
82	0.8587	0.6673E-01	0.2319E-03	0.5213E-03	0.3799E-01
83	0.8577	0.6682E-01	0.2322E-03	0.5207E-03	0.3794E-01
84	0.8568	0.6691E-01	0.2325E-03	0.5200E-03	0.3789E-01
85	0.8559	0.6699E-01	0.2327E-03	0.5194E-03	0.3784E-01
86	0.8549	0.6707E-01	0.2330E-03	0.5187E-03	0.3779E-01
87	0.8539	0.6714E-01	0.2332E-03	0.5180E-03	0.3774E-01
88	0.8529	0.6721E-01	0.2335E-03	0.5173E-03	0.3769E-01
89	0.8519	0.6728E-01	0.2337E-03	0.5166E-03	0.3764E-01
90	0.8508	0.6734E-01	0.2339E-03	0.5159E-03	0.3758E-01
91	0.8496	0.6739E-01	0.2341E-03	0.5151E-03	0.3752E-01
92	0.8484	0.6744E-01	0.2342E-03	0.5143E-03	0.3746E-01
93	0.8471	0.6748E-01	0.2343E-03	0.5135E-03	0.3740E-01
94	0.8458	0.6752E-01	0.2344E-03	0.5126E-03	0.3734E-01
95	0.8443	0.6754E-01	0.2345E-03	0.5116E-03	0.3727E-01
96	0.8427	0.6755E-01	0.2345E-03	0.5106E-03	0.3719E-01
97	0.8410	0.6755E-01	0.2345E-03	0.5096E-03	0.3711E-01
98	0.8392	0.6754E-01	0.2345E-03	0.5084E-03	0.3703E-01
99	0.8372	0.6751E-01	0.2344E-03	0.5071E-03	0.3694E-01
100	0.8350	0.6747E-01	0.2342E-03	0.5058E-03	0.3684E-01
101	0.8327	0.6740E-01	0.2340E-03	0.5043E-03	0.3673E-01
102	0.8301	0.6732E-01	0.2337E-03	0.5027E-03	0.3661E-01
103	0.8272	0.6722E-01	0.2333E-03	0.5009E-03	0.3648E-01
104	0.8241	0.6709E-01	0.2329E-03	0.4990E-03	0.3634E-01
105	0.8207	0.6693E-01	0.2323E-03	0.4969E-03	0.3618E-01
106	0.8170	0.6675E-01	0.2317E-03	0.4946E-03	0.3601E-01
107	0.8129	0.6654E-01	0.2309E-03	0.4921E-03	0.3583E-01
108	0.8084	0.6629E-01	0.2301E-03	0.4893E-03	0.3563E-01
109	0.8036	0.6600E-01	0.2291E-03	0.4863E-03	0.3541E-01
110	0.7982	0.6568E-01	0.2279E-03	0.4831E-03	0.3517E-01
111	0.7925	0.6532E-01	0.2267E-03	0.4796E-03	0.3492E-01
112	0.7862	0.6492E-01	0.2253E-03	0.4758E-03	0.3464E-01
113	0.7794	0.6447E-01	0.2237E-03	0.4716E-03	0.3434E-01
114	0.7721	0.6397E-01	0.2220E-03	0.4672E-03	0.3401E-01
115	0.7643	0.6342E-01	0.2201E-03	0.4624E-03	0.3366E-01
116	0.7558	0.6282E-01	0.2180E-03	0.4572E-03	0.3329E-01
117	0.7467	0.6217E-01	0.2157E-03	0.4517E-03	0.3289E-01
118	0.7371	0.6146E-01	0.2132E-03	0.4458E-03	0.3246E-01
119	0.7267	0.6070E-01	0.2106E-03	0.4396E-03	0.3200E-01
120	0.7158	0.5988E-01	0.2077E-03	0.4329E-03	0.3152E-01
121	0.7042	0.5900E-01	0.2047E-03	0.4259E-03	0.3100E-01
122	0.6919	0.5806E-01	0.2014E-03	0.4185E-03	0.3046E-01
123	0.6791	0.5707E-01	0.1980E-03	0.4106E-03	0.2989E-01
124	0.6655	0.5601E-01	0.1943E-03	0.4024E-03	0.2929E-01
125	0.6514	0.5491E-01	0.1905E-03	0.3939E-03	0.2867E-01
126	0.6366	0.5374E-01	0.1864E-03	0.3849E-03	0.2802E-01
127	0.6213	0.5253E-01	0.1822E-03	0.3757E-03	0.2734E-01
128	0.6054	0.5126E-01	0.1778E-03	0.3660E-03	0.2664E-01
129	0.5890	0.4994E-01	0.1732E-03	0.3561E-03	0.2592E-01
130	0.5722	0.4858E-01	0.1685E-03	0.3459E-03	0.2518E-01
131	0.5548	0.4718E-01	0.1636E-03	0.3354E-03	0.2441E-01
132	0.5370	0.4573E-01	0.1586E-03	0.3246E-03	0.2363E-01
133	0.5190	0.4426E-01	0.1535E-03	0.3137E-03	0.2283E-01
134	0.5006	0.4275E-01	0.1483E-03	0.3026E-03	0.2202E-01
135	0.4820	0.4121E-01	0.1429E-03	0.2913E-03	0.2120E-01
136	0.4631	0.3965E-01	0.1375E-03	0.2799E-03	0.2037E-01
137	0.4441	0.3808E-01	0.1321E-03	0.2684E-03	0.1953E-01
138	0.4251	0.3649E-01	0.1266E-03	0.2569E-03	0.1870E-01
139	0.4059	0.3490E-01	0.1210E-03	0.2453E-03	0.1785E-01
140	0.3867	0.3329E-01	0.1154E-03	0.2337E-03	0.1701E-01

Table A.3: Higgs branching ratios for $H \to b\bar{b}$, $\tau^+\tau^-$, $\mu^+\mu^-$, $s\bar{s}$ and $c\bar{c}$ for Higgs masses from 80 GeV/c^2 to 140 GeV/c^2.

M(H)	bb	tau tau	muon muon	ss	cc
141	0.3679	0.3171E-01	0.1100E-03	0.2223E-03	0.1618E-01
142	0.3486	0.3009E-01	0.1043E-03	0.2107E-03	0.1533E-01
143	0.3300	0.2852E-01	0.9891E-04	0.1994E-03	0.1451E-01
144	0.3113	0.2694E-01	0.9341E-04	0.1881E-03	0.1369E-01
145	0.2928	0.2537E-01	0.8798E-04	0.1769E-03	0.1288E-01
146	0.2747	0.2383E-01	0.8263E-04	0.1660E-03	0.1208E-01
147	0.2568	0.2230E-01	0.7733E-04	0.1551E-03	0.1129E-01
148	0.2391	0.2079E-01	0.7211E-04	0.1445E-03	0.1051E-01
149	0.2218	0.1931E-01	0.6697E-04	0.1340E-03	0.9751E-02
150	0.2047	0.1784E-01	0.6187E-04	0.1236E-03	0.8997E-02
151	0.1882	0.1643E-01	0.5696E-04	0.1137E-03	0.8273E-02
152	0.1714	0.1498E-01	0.5194E-04	0.1036E-03	0.7535E-02
153	0.1557	0.1362E-01	0.4723E-04	0.9405E-04	0.6843E-02
154	0.1395	0.1222E-01	0.4236E-04	0.8425E-04	0.6130E-02
155	0.1240	0.1088E-01	0.3771E-04	0.7492E-04	0.5451E-02
156	0.1085	0.9525E-02	0.3303E-04	0.6553E-04	0.4768E-02
157	0.9314E-01	0.8186E-02	0.2838E-04	0.5626E-04	0.4093E-02
158	0.7795E-01	0.6859E-02	0.2378E-04	0.4708E-04	0.3425E-02
159	0.6276E-01	0.5529E-02	0.1917E-04	0.3791E-04	0.2758E-02
160	0.4794E-01	0.4228E-02	0.1466E-04	0.2895E-04	0.2107E-02
161	0.3578E-01	0.3159E-02	0.1095E-04	0.2161E-04	0.1572E-02
162	0.2689E-01	0.2377E-02	0.8240E-05	0.1624E-04	0.1181E-02
163	0.2095E-01	0.1854E-02	0.6428E-05	0.1265E-04	0.9206E-03
164	0.1745E-01	0.1546E-02	0.5359E-05	0.1054E-04	0.7666E-03
165	0.1525E-01	0.1353E-02	0.4689E-05	0.9211E-05	0.6701E-03
166	0.1371E-01	0.1217E-02	0.4219E-05	0.8277E-05	0.6021E-03
167	0.1254E-01	0.1114E-02	0.3863E-05	0.7571E-05	0.5508E-03
168	0.1161E-01	0.1033E-02	0.3581E-05	0.7011E-05	0.5100E-03
169	0.1085E-01	0.9662E-03	0.3350E-05	0.6552E-05	0.4766E-03
170	0.1021E-01	0.9101E-03	0.3155E-05	0.6164E-05	0.4484E-03
171	0.9657E-02	0.8618E-03	0.2988E-05	0.5831E-05	0.4242E-03
172	0.9174E-02	0.8195E-03	0.2841E-05	0.5539E-05	0.4030E-03
173	0.8744E-02	0.7820E-03	0.2711E-05	0.5280E-05	0.3841E-03
174	0.8359E-02	0.7482E-03	0.2594E-05	0.5047E-05	0.3671E-03
175	0.8006E-02	0.7174E-03	0.2487E-05	0.4834E-05	0.3516E-03
176	0.7681E-02	0.6891E-03	0.2389E-05	0.4638E-05	0.3374E-03
177	0.7377E-02	0.6625E-03	0.2297E-05	0.4454E-05	0.3240E-03
178	0.7090E-02	0.6374E-03	0.2210E-05	0.4280E-05	0.3114E-03
179	0.6812E-02	0.6131E-03	0.2125E-05	0.4113E-05	0.2992E-03
180	0.6538E-02	0.5890E-03	0.2042E-05	0.3947E-05	0.2871E-03
181	0.6257E-02	0.5643E-03	0.1956E-05	0.3777E-05	0.2747E-03
182	0.5963E-02	0.5384E-03	0.1866E-05	0.3600E-05	0.2619E-03
183	0.5661E-02	0.5116E-03	0.1774E-05	0.3417E-05	0.2486E-03
184	0.5358E-02	0.4847E-03	0.1680E-05	0.3235E-05	0.2353E-03
185	0.5069E-02	0.4591E-03	0.1591E-05	0.3060E-05	0.2226E-03
186	0.4827E-02	0.4376E-03	0.1517E-05	0.2914E-05	0.2119E-03
187	0.4622E-02	0.4194E-03	0.1454E-05	0.2790E-05	0.2029E-03
188	0.4442E-02	0.4035E-03	0.1399E-05	0.2681E-05	0.1950E-03
189	0.4281E-02	0.3893E-03	0.1349E-05	0.2584E-05	0.1880E-03
190	0.4136E-02	0.3765E-03	0.1305E-05	0.2497E-05	0.1816E-03
191	0.4003E-02	0.3647E-03	0.1264E-05	0.2416E-05	0.1757E-03
192	0.3881E-02	0.3539E-03	0.1227E-05	0.2342E-05	0.1704E-03
193	0.3767E-02	0.3439E-03	0.1192E-05	0.2274E-05	0.1654E-03
194	0.3661E-02	0.3346E-03	0.1160E-05	0.2210E-05	0.1607E-03
195	0.3562E-02	0.3258E-03	0.1129E-05	0.2150E-05	0.1563E-03
196	0.3468E-02	0.3175E-03	0.1101E-05	0.2093E-05	0.1522E-03
197	0.3380E-02	0.3097E-03	0.1074E-05	0.2040E-05	0.1483E-03
198	0.3296E-02	0.3024E-03	0.1048E-05	0.1989E-05	0.1447E-03
199	0.3216E-02	0.2953E-03	0.1024E-05	0.1941E-05	0.1412E-03
200	0.3141E-02	0.2887E-03	0.1001E-05	0.1896E-05	0.1378E-03

Table A.4: Higgs branching ratios for $H \rightarrow b\bar{b}$, $\tau^+\tau^-$, $\mu^+\mu^-$, $s\bar{s}$ and $c\bar{c}$ for Higgs masses from 141 GeV/c^2 to 200 GeV/c^2.

M(H)	gg	gamma gamma	Z gamma	WW	ZZ	Width [GeV]
80	0.3283E-01	0.7568E-03	0.000	0.4881E-03	0.1320E-03	0.2476E-02
81	0.3357E-01	0.7795E-03	0.000	0.5419E-03	0.1450E-03	0.2503E-02
82	0.3431E-01	0.8028E-03	0.000	0.6022E-03	0.1591E-03	0.2531E-02
83	0.3507E-01	0.8265E-03	0.000	0.6702E-03	0.1745E-03	0.2558E-02
84	0.3584E-01	0.8507E-03	0.000	0.7475E-03	0.1914E-03	0.2586E-02
85	0.3661E-01	0.8753E-03	0.000	0.8366E-03	0.2098E-03	0.2613E-02
86	0.3740E-01	0.9005E-03	0.000	0.9411E-03	0.2300E-03	0.2641E-02
87	0.3819E-01	0.9262E-03	0.000	0.1064E-02	0.2520E-03	0.2668E-02
88	0.3899E-01	0.9524E-03	0.000	0.1212E-02	0.2761E-03	0.2696E-02
89	0.3981E-01	0.9791E-03	0.000	0.1392E-02	0.3025E-03	0.2724E-02
90	0.4063E-01	0.1006E-02	0.000	0.1610E-02	0.3314E-03	0.2752E-02
91	0.4145E-01	0.1034E-02	0.000	0.1874E-02	0.3632E-03	0.2781E-02
92	0.4229E-01	0.1062E-02	0.3655E-07	0.2201E-02	0.3981E-03	0.2809E-02
93	0.4313E-01	0.1091E-02	0.3970E-06	0.2594E-02	0.4367E-03	0.2838E-02
94	0.4397E-01	0.1120E-02	0.1453E-05	0.3069E-02	0.4795E-03	0.2867E-02
95	0.4482E-01	0.1150E-02	0.3547E-05	0.3647E-02	0.5272E-03	0.2897E-02
96	0.4568E-01	0.1180E-02	0.6994E-05	0.4332E-02	0.5810E-03	0.2927E-02
97	0.4653E-01	0.1211E-02	0.1208E-04	0.5150E-02	0.6423E-03	0.2957E-02
98	0.4739E-01	0.1242E-02	0.1908E-04	0.6120E-02	0.7128E-03	0.2988E-02
99	0.4824E-01	0.1274E-02	0.2824E-04	0.7255E-02	0.7943E-03	0.3020E-02
100	0.4909E-01	0.1306E-02	0.3978E-04	0.8585E-02	0.8900E-03	0.3052E-02
101	0.4994E-01	0.1338E-02	0.5390E-04	0.1012E-01	0.1003E-02	0.3086E-02
102	0.5077E-01	0.1371E-02	0.7080E-04	0.1190E-01	0.1136E-02	0.3120E-02
103	0.5160E-01	0.1403E-02	0.9063E-04	0.1394E-01	0.1293E-02	0.3155E-02
104	0.5242E-01	0.1436E-02	0.1136E-03	0.1627E-01	0.1480E-02	0.3192E-02
105	0.5322E-01	0.1469E-02	0.1397E-03	0.1891E-01	0.1700E-02	0.3230E-02
106	0.5401E-01	0.1502E-02	0.1692E-03	0.2190E-01	0.1957E-02	0.3270E-02
107	0.5477E-01	0.1535E-02	0.2021E-03	0.2526E-01	0.2261E-02	0.3311E-02
108	0.5551E-01	0.1568E-02	0.2384E-03	0.2901E-01	0.2616E-02	0.3355E-02
109	0.5622E-01	0.1601E-02	0.2783E-03	0.3320E-01	0.3024E-02	0.3400E-02
110	0.5690E-01	0.1633E-02	0.3218E-03	0.3784E-01	0.3498E-02	0.3448E-02
111	0.5754E-01	0.1665E-02	0.3689E-03	0.4297E-01	0.4042E-02	0.3499E-02
112	0.5814E-01	0.1697E-02	0.4194E-03	0.4861E-01	0.4661E-02	0.3552E-02
113	0.5870E-01	0.1727E-02	0.4735E-03	0.5479E-01	0.5365E-02	0.3609E-02
114	0.5921E-01	0.1757E-02	0.5309E-03	0.6154E-01	0.6158E-02	0.3669E-02
115	0.5967E-01	0.1786E-02	0.5917E-03	0.6888E-01	0.7048E-02	0.3733E-02
116	0.6007E-01	0.1814E-02	0.6556E-03	0.7684E-01	0.8041E-02	0.3802E-02
117	0.6041E-01	0.1840E-02	0.7224E-03	0.8543E-01	0.9143E-02	0.3875E-02
118	0.6069E-01	0.1865E-02	0.7921E-03	0.9467E-01	0.1036E-01	0.3953E-02
119	0.6089E-01	0.1889E-02	0.8643E-03	0.1046	0.1169E-01	0.4036E-02
120	0.6102E-01	0.1911E-02	0.9388E-03	0.1152	0.1315E-01	0.4126E-02
121	0.6108E-01	0.1931E-02	0.1015E-02	0.1264	0.1473E-01	0.4222E-02
122	0.6105E-01	0.1949E-02	0.1094E-02	0.1384	0.1644E-01	0.4326E-02
123	0.6094E-01	0.1964E-02	0.1173E-02	0.1510	0.1828E-01	0.4437E-02
124	0.6074E-01	0.1978E-02	0.1253E-02	0.1644	0.2024E-01	0.4557E-02
125	0.6046E-01	0.1989E-02	0.1334E-02	0.1783	0.2233E-01	0.4687E-02
126	0.6008E-01	0.1997E-02	0.1416E-02	0.1930	0.2454E-01	0.4826E-02
127	0.5961E-01	0.2003E-02	0.1496E-02	0.2083	0.2687E-01	0.4977E-02
128	0.5905E-01	0.2006E-02	0.1576E-02	0.2242	0.2932E-01	0.5140E-02
129	0.5839E-01	0.2006E-02	0.1655E-02	0.2407	0.3186E-01	0.5317E-02
130	0.5765E-01	0.2004E-02	0.1733E-02	0.2577	0.3450E-01	0.5508E-02
131	0.5681E-01	0.1998E-02	0.1808E-02	0.2753	0.3722E-01	0.5716E-02
132	0.5588E-01	0.1989E-02	0.1880E-02	0.2934	0.4001E-01	0.5942E-02
133	0.5487E-01	0.1977E-02	0.1949E-02	0.3118	0.4286E-01	0.6186E-02
134	0.5377E-01	0.1962E-02	0.2015E-02	0.3307	0.4575E-01	0.6453E-02
135	0.5259E-01	0.1944E-02	0.2077E-02	0.3499	0.4866E-01	0.6743E-02
136	0.5133E-01	0.1923E-02	0.2135E-02	0.3695	0.5158E-01	0.7060E-02
137	0.4999E-01	0.1899E-02	0.2188E-02	0.3893	0.5448E-01	0.7406E-02
138	0.4859E-01	0.1872E-02	0.2236E-02	0.4093	0.5736E-01	0.7784E-02
139	0.4712E-01	0.1843E-02	0.2279E-02	0.4295	0.6017E-01	0.8199E-02
140	0.4558E-01	0.1809E-02	0.2315E-02	0.4501	0.6288E-01	0.8657E-02

Table A.5: Higgs branching ratios for $H \to gg$, $\gamma\gamma$, $Z\gamma$, WW and ZZ for Higgs masses from 80 GeV/c^2 to 140 GeV/c^2. The last column shows to the total decay width in GeV. For $H \to Z\gamma$ the first 12 values are set to zero because of the limited precision of the calculation program.

M(H)	gg	gamma gamma	Z gamma	WW	ZZ	Width [GeV]
141	0.4402E-01	0.1775E-02	0.2348E-02	0.4702	0.6556E-01	0.9152E-02
142	0.4235E-01	0.1736E-02	0.2370E-02	0.4912	0.6801E-01	0.9715E-02
143	0.4070E-01	0.1696E-02	0.2390E-02	0.5115	0.7040E-01	0.1032E-01
144	0.3896E-01	0.1653E-02	0.2400E-02	0.5322	0.7255E-01	0.1100E-01
145	0.3720E-01	0.1607E-02	0.2404E-02	0.5529	0.7449E-01	0.1176E-01
146	0.3541E-01	0.1559E-02	0.2401E-02	0.5736	0.7621E-01	0.1261E-01
147	0.3359E-01	0.1509E-02	0.2390E-02	0.5943	0.7763E-01	0.1357E-01
148	0.3174E-01	0.1456E-02	0.2371E-02	0.6150	0.7876E-01	0.1465E-01
149	0.2988E-01	0.1401E-02	0.2344E-02	0.6358	0.7950E-01	0.1588E-01
150	0.2796E-01	0.1342E-02	0.2306E-02	0.6569	0.7982E-01	0.1731E-01
151	0.2609E-01	0.1283E-02	0.2263E-02	0.6775	0.7982E-01	0.1892E-01
152	0.2410E-01	0.1217E-02	0.2203E-02	0.6994	0.7902E-01	0.2089E-01
153	0.2220E-01	0.1153E-02	0.2141E-02	0.7202	0.7800E-01	0.2312E-01
154	0.2017E-01	0.1081E-02	0.2057E-02	0.7428	0.7592E-01	0.2595E-01
155	0.1819E-01	0.1008E-02	0.1966E-02	0.7650	0.7334E-01	0.2934E-01
156	0.1614E-01	0.9276E-03	0.1855E-02	0.7885	0.6970E-01	0.3371E-01
157	0.1405E-01	0.8420E-03	0.1727E-02	0.8129	0.6502E-01	0.3948E-01
158	0.1192E-01	0.7503E-03	0.1578E-02	0.8383	0.5910E-01	0.4741E-01
159	0.9731E-02	0.6502E-03	0.1405E-02	0.8653	0.5177E-01	0.5919E-01
160	0.7535E-02	0.5468E-03	0.1217E-02	0.8934	0.4295E-01	0.7789E-01
161	0.5701E-02	0.4925E-03	0.1143E-02	0.9172	0.3492E-01	0.1049
162	0.4343E-02	0.3732E-03	0.8807E-03	0.9354	0.2854E-01	0.1403
163	0.3430E-02	0.2931E-03	0.7027E-03	0.9476	0.2425E-01	0.1810
164	0.2895E-02	0.2460E-03	0.5986E-03	0.9545	0.2202E-01	0.2184
165	0.2565E-02	0.2167E-03	0.5350E-03	0.9584	0.2102E-01	0.2511
166	0.2336E-02	0.1962E-03	0.4911E-03	0.9608	0.2066E-01	0.2808
167	0.2165E-02	0.1808E-03	0.4586E-03	0.9623	0.2070E-01	0.3085
168	0.2031E-02	0.1686E-03	0.4332E-03	0.9632	0.2105E-01	0.3348
169	0.1923E-02	0.1586E-03	0.4127E-03	0.9636	0.2164E-01	0.3600
170	0.1834E-02	0.1502E-03	0.3957E-03	0.9636	0.2249E-01	0.3845
171	0.1757E-02	0.1430E-03	0.3811E-03	0.9632	0.2354E-01	0.4084
172	0.1691E-02	0.1367E-03	0.3684E-03	0.9625	0.2487E-01	0.4320
173	0.1633E-02	0.1311E-03	0.3572E-03	0.9614	0.2652E-01	0.4554
174	0.1581E-02	0.1260E-03	0.3471E-03	0.9601	0.2841E-01	0.4786
175	0.1534E-02	0.1214E-03	0.3378E-03	0.9580	0.3093E-01	0.5021
176	0.1491E-02	0.1171E-03	0.3293E-03	0.9555	0.3381E-01	0.5257
177	0.1450E-02	0.1131E-03	0.3211E-03	0.9521	0.3760E-01	0.5499
178	0.1411E-02	0.1092E-03	0.3131E-03	0.9477	0.4242E-01	0.5748
179	0.1373E-02	0.1055E-03	0.3052E-03	0.9418	0.4873E-01	0.6009
180	0.1335E-02	0.1017E-03	0.2970E-03	0.9334	0.5743E-01	0.6290
181	0.1293E-02	0.9783E-04	0.2881E-03	0.9215	0.6977E-01	0.6602
182	0.1248E-02	0.9367E-04	0.2782E-03	0.9050	0.8666E-01	0.6958
183	0.1200E-02	0.8932E-04	0.2675E-03	0.8843	0.1077	0.7363
184	0.1150E-02	0.8491E-04	0.2564E-03	0.8610	0.1315	0.7813
185	0.1101E-02	0.8067E-04	0.2455E-03	0.8371	0.1557	0.8295
186	0.1062E-02	0.7713E-04	0.2366E-03	0.8187	0.1745	0.8749
187	0.1029E-02	0.7414E-04	0.2291E-03	0.8045	0.1890	0.9178
188	0.1001E-02	0.7153E-04	0.2227E-03	0.7930	0.2007	0.9591
189	0.9769E-03	0.6920E-04	0.2170E-03	0.7835	0.2104	0.9994
190	0.9553E-03	0.6709E-04	0.2119E-03	0.7755	0.2185	1.039
191	0.9359E-03	0.6517E-04	0.2072E-03	0.7687	0.2255	1.078
192	0.9183E-03	0.6339E-04	0.2029E-03	0.7628	0.2316	1.117
193	0.9021E-03	0.6173E-04	0.1989E-03	0.7576	0.2369	1.155
194	0.8873E-03	0.6019E-04	0.1952E-03	0.7531	0.2416	1.194
195	0.8736E-03	0.5874E-04	0.1917E-03	0.7490	0.2458	1.232
196	0.8608E-03	0.5737E-04	0.1884E-03	0.7454	0.2495	1.271
197	0.8489E-03	0.5607E-04	0.1852E-03	0.7422	0.2529	1.309
198	0.8377E-03	0.5483E-04	0.1822E-03	0.7393	0.2559	1.348
199	0.8272E-03	0.5366E-04	0.1793E-03	0.7367	0.2586	1.387
200	0.8173E-03	0.5253E-04	0.1766E-03	0.7343	0.2611	1.426

Table A.6: Higgs branching ratios for $H \to gg$, $\gamma\gamma$, $Z\gamma$, WW and ZZ for Higgs masses from 141 GeV/c^2 to 200 GeV/c^2. The last column shows to the total decay width in GeV.

Appendix B

Weight calculation

In order to determine the correct weights that had to be applied to each event while running the network several effects were taken into account. They are described in the following subsections.

B.1 NB-teacher

Number of signal and background events

The weight w_S normalizes the number of Monte Carlo events for each background process to the number of Monte Carlo signal events, i. e.

$$w_S = \frac{N_{tot}^{signal}}{N_{tot}^{back}}. \tag{B.1}$$

N_{tot}^{signal} is the number of MC events for a given Higgs mass and N_{tot}^{back} is the number of MC events for a particular background process.

Cross-section, branching ratio and acceptance

The backgrounds are normalized the background with the biggest product of cross-section, branching ratio and acceptance, i. e.

$$w_\sigma = \frac{\sigma \cdot BR \cdot \mathcal{A}}{(\sigma \cdot BR \cdot \mathcal{A})_{max}}. \tag{B.2}$$

Summ of all background events

The weight w_Σ normalizes the weighted summ of all backgrounds to the number of signal events, i. e.

$$w_\Sigma = \frac{N_{tot}^{signal}}{N_{back}} \tag{B.3}$$

where N_{back} is given by

$$N_{back} = \sum_i w_S^i \cdot w_\sigma^i \cdot N_{tot}^i \tag{B.4}$$

and the index i runs of over the different backgrounds.

Final weights

The final weight that was applied during the network training is given by the product

$$w_{teach} = w_S \cdot w_\sigma \cdot w_\Sigma \tag{B.5}$$

and given in table B.1.

Process	N_{obs}	N_f	$w_S \cdot w_\sigma$	w_Σ	w_{teach}	N_{exp}	f_{lumi}
$WH120$	462414	10310	1	1	1	0.41	25146.3
$t\bar{t}$	365743	4221	0.604704	1.02518	0.619932	7.38	571.951
W^*	425541	8208	0.170034	1.02518	0.174316	2.07515	3955.38
Wg	482770	5326	0.217172	1.02518	0.222641	2.65043	2009.48
$Wbbe$	226125	2207	1.0	1.02518	1.02518	12.2043	180.838
$Wbb\mu$	213566	1664	0.797348	1.02518	0.817428	9.7311	170.998
$Wbb\tau$	667004	2026	0.0965909	1.02518	0.0990233	1.17883	1718.66
$Wcce$	244351	503	0.365657	1.02518	0.374865	4.46259	112.715
$Wcc\mu$	283361	453	0.284175	1.02518	0.291331	3.46816	130.617
$Wcc\tau$	669993	618	0.023569	1.02518	0.0241626	0.287644	2148.49
Wce	284217	320	0.428956	1.02518	0.439759	5.23512	61.1256
$Wc\mu$	321304	437	0.330909	1.02518	0.339242	4.03852	108.208
$Wce1p$	280676	794	0.170286	1.02518	0.174574	2.07823	382.056
$Wc\mu1p$	256436	560	0.130219	1.02518	0.133498	1.58923	352.371
$We\nu2p$	65706	56	0.5867	1.02518	0.601475	7.16028	7.82092
$W\mu\nu2p$	251635	106	0.499411	1.02518	0.511987	6.09497	17.3914
$W\tau\nu2p$	542703	97	0.0414983	1.02518	0.0425434	0.506459	191.526
WW	216547	104	0.00841751	1.02518	0.00862949	0.10273	1012.36
WZ	184934	425	0.0925926	1.02518	0.0949243	1.13003	376.096
ZZ	913775	2104	0.0925926	1.02518	0.0949243	1.13003	1861.9
$Z\tau\tau$	2105900	35	0.0252525	1.02518	0.0258885	0.30819	113.566
QCD	1390660	90	0.883838	1.02518	0.906096	10.7867	8.34365

Table B.1: Weights for the NB-teacher. N_{obs} is the number of events passing the z_{vertex}-cut of $|z_0| \leq 60$ cm and N_f the number of events passing the preselection cuts. The weights w_S, w_σ and w_Σ are explained in the text above. The final weight w_{teach} is the product of the three numbers. N_{exp} is the number of Monte Carlo events that are expected to be found in a data sample of 162 pb$^-$1. The factor f_{lumi} shows the fraction N_f/N_{exp} and describes how much more Monte Carlo luminosity was available than was available in the data. The final weight is given by w_{teach}.

Process	$w_{\mathcal{L}}$
$WH120$	4.00494e-05
$t\bar{t}$	0.00211161
W^*	0.000240671
Wg	0.00045498
$Wbbe$	0.00155441
$Wbb\mu$	0.00148549
$Wbb\tau$	0.00155845
$Wcce$	0.00227316
$Wcc\mu$	0.00175676
$Wcc\tau$	0.00024107
Wce	0.00841152
$Wc\mu$	0.01237
$Wce1p$	0.00516988
$Wc\mu1p$	0.00511266
$We\nu2p$	0.157047
$W\mu\nu2p$	0.103666
$W\tau\nu2p$	0.0968981
WW	8.2534e-05
WZ	0.000391735
ZZ	0.000516524
$Z\tau\tau$	0.00328132
QCD	0.0104621

Table B.2: Weights for the NB-expert.

B.2 NB-expert

For running the expert the events were weighted according the data luminosity. $w_{\mathcal{L}}$ weights the signal and the different background to the same CEM, CMUP and CMX luminosity, i. e.

$$w_{\mathcal{L}}^{det} = \frac{\mathcal{L}_{data}^{det}}{\mathcal{L}_{MC}}, \tag{B.6}$$

where *det* refers to a specific detector which is either the CEM, CMUP or CMX. The luminosity for the Monte Carlo is given by

$$\mathcal{L}_{MC} = \frac{N_{tot}}{\sigma \cdot BR \cdot \mathcal{A}}. \tag{B.7}$$

with σ being the cross-section, BR being the branching ratio and $\mathcal{A}$ being the acceptance. The weights are given in table B.2.

Appendix C

Acceptances

Cut	CEM	CMUP	CMX	Total
Initial sample				469837
N_{obsv}				454940
N_{geom}	114354	17866	8653	135713
$N_{leptonid}$	22213	12907	5151	40186
N_{iso}	20187	11750	4692	36587
N_{met}	17865	10273	4071	32171
N_{diveto}	17676	10234	4046	31922
N_{zveto}	17553	10192	4020	31731
$N_{convveto}$	16893	9847	3883	30590
N_{jet}	9173	5433	2116	16703
$N_{single-tag}$	5187	3073	1197	9444
$N_{double-tag}$	1311	763	315	2387
Acceptance	0.0114	0.00675	0.00263	0.0208

Table C.1: Number of CEM, CMUP and CMX events after each cut. The Higgs mass was set to $m_H = 110$ GeV/c^2.

Cut	CEM	CMUP	CMX	Total
Initial sample				450713
N_{obsv}				436430
N_{geom}	114274	17376	8408	134786
$N_{leptonid}$	21552	12394	4888	38749
N_{iso}	19534	11264	4420	35183
N_{met}	17327	9857	3870	31023
N_{diveto}	17158	9802	3848	30780
N_{zveto}	17045	9764	3830	30612
$N_{convveto}$	16375	9397	3688	29435
N_{jet}	8989	5189	1991	16153
$N_{single-tag}$	5164	2985	1133	9274
$N_{double-tag}$	1256	734	262	2251
Acceptance	0.0118	0.00684	0.0026	0.0212

Table C.2: Number of CEM, CMUP and CMX events after each cut. The Higgs mass was set to $m_H = 115$ GeV/c^2.

Cut	CEM	CMUP	CMX	Total
Initial sample				455446
N_{obsv}				441022
N_{geom}	128616	19182	8953	150084
$N_{leptonid}$	21893	12907	5013	39725
N_{iso}	19857	11781	4532	36130
N_{met}	17645	10322	3950	31880
N_{diveto}	17495	10272	3924	31655
N_{zveto}	17350	10205	3915	31435
$N_{convveto}$	16623	9841	3762	30191
N_{jet}	9655	5688	2166	17487
$N_{single-tag}$	5603	3222	1251	10061
$N_{double-tag}$	1407	819	303	2524
Acceptance	0.0127	0.00731	0.00284	0.0228

Table C.3: Number of CEM, CMUP and CMX events after each cut. The Higgs mass was set to $m_H = 130$ GeV/c^2.

Cut	CEM	CMUP	CMX	Total
Initial sample				342000
N_{obsv}				330274
N_{geom}	103962	16309	7246	121408
$N_{leptonid}$	17297	11032	4018	32276
N_{iso}	15658	10059	3649	29337
N_{met}	13627	8748	3142	25491
N_{diveto}	13484	8693	3117	25272
N_{zveto}	13380	8650	3101	25109
$N_{convveto}$	12773	8315	2984	24052
N_{jet}	7390	4781	1721	13880
$N_{single-tag}$	4298	2840	1022	8154
$N_{double-tag}$	1114	719	262	2093
Acceptance	0.013	0.0086	0.00309	0.0247

Table C.4: Number of CEM, CMUP and CMX events after each cut. The Higgs mass was set to $m_H = 140$ GeV/c^2.

Cut	CEM	CMUP	CMX	Total
Initial sample				378000
N_{obsv}				365165
N_{geom}	122300	18992	8309	141915
$N_{leptonid}$	19587	12422	4512	36419
N_{iso}	17653	11359	4088	33053
N_{met}	15396	9857	3558	28767
N_{diveto}	15230	9788	3534	28515
N_{zveto}	15068	9721	3520	28273
$N_{convveto}$	14376	9313	3378	27031
N_{jet}	8445	5419	1954	15792
$N_{single-tag}$	5026	3124	1117	9252
$N_{double-tag}$	1284	790	285	2357
Acceptance	0.0138	0.00856	0.00306	0.0253

Table C.5: Number of CEM, CMUP and CMX events after each cut. The Higgs mass was set to $m_H = 150$ GeV/c^2.

Cut	CEM	CMUP	CMX	Total
Initial sample				378471
N_{obsv}				365743
N_{geom}	160059	30584	13428	187427
$N_{leptonid}$	37029	22295	8226	65566
N_{iso}	33431	20258	7465	59396
N_{met}	30036	18137	6651	53223
N_{diveto}	26599	15995	5835	48266
N_{zveto}	25216	15425	5625	46111
$N_{convveto}$	23521	14500	5273	43147
N_{jet}	4336	2714	1034	8021
$N_{single-tag}$	2260	1425	576	4221
$N_{double-tag}$	448	264	102	805
Acceptance	0.00618	0.0039	0.00157	0.0115

Table C.6: Number of CEM, CMUP and CMX events after each cut for $t\bar{t}$. The top mass was set to $m_t = 175$ GeV/c^2.

Appendix D

Network performance

D.1 Input variables

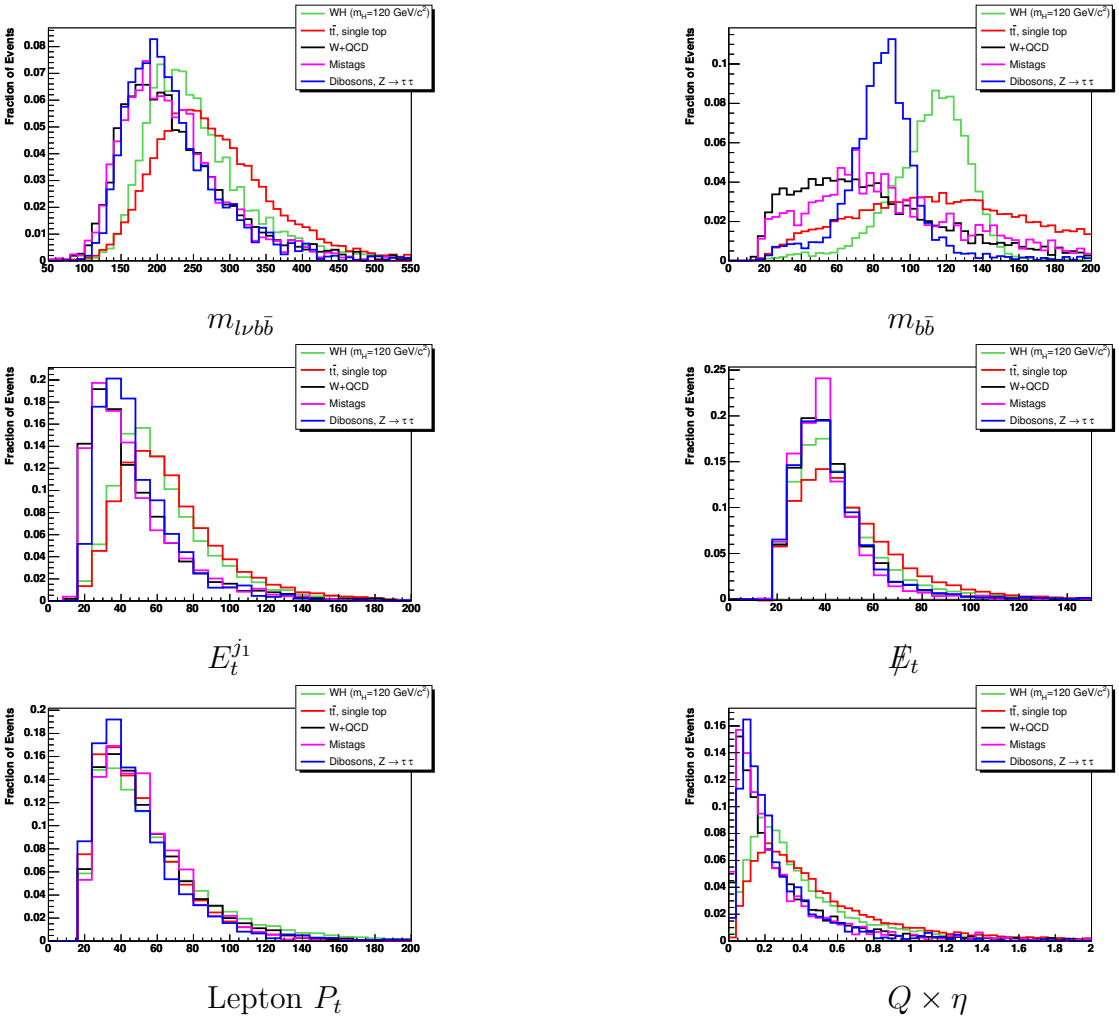


Figure D.1: Final input variables for a Higgs boson mass of 120 GeV/c^2. The integrals of all distributions are normalised to 1.

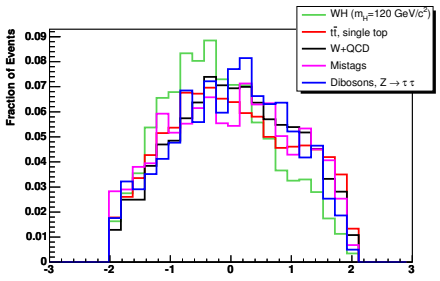
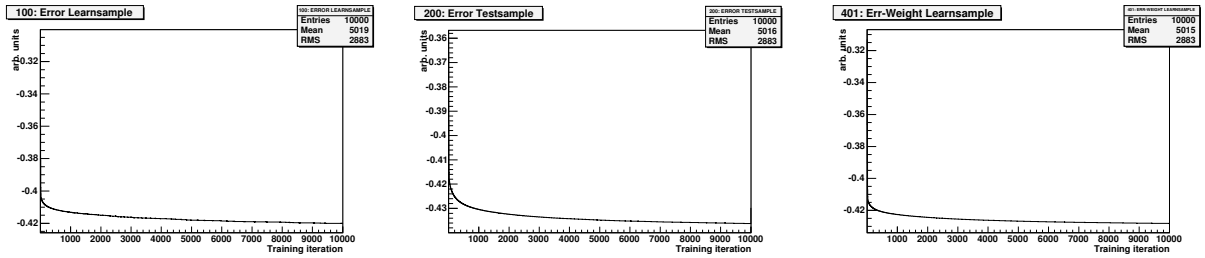
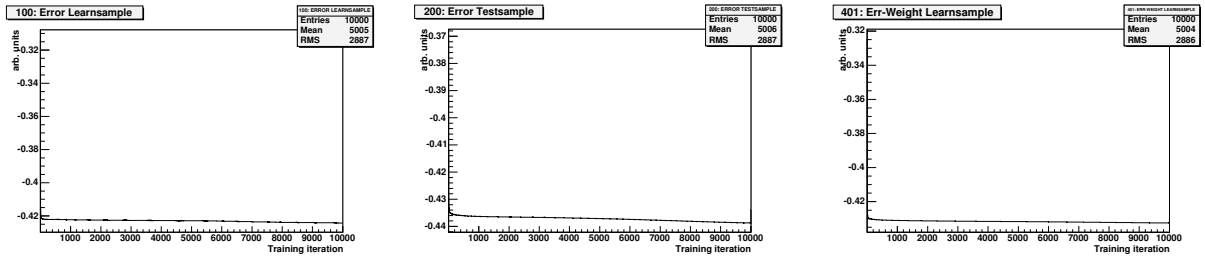


Figure D.2: Final input variable ($Q \times \eta$) for a Higgs boson mass of 120 GeV/c^2. The integrals of all distribution are normalised to 1.

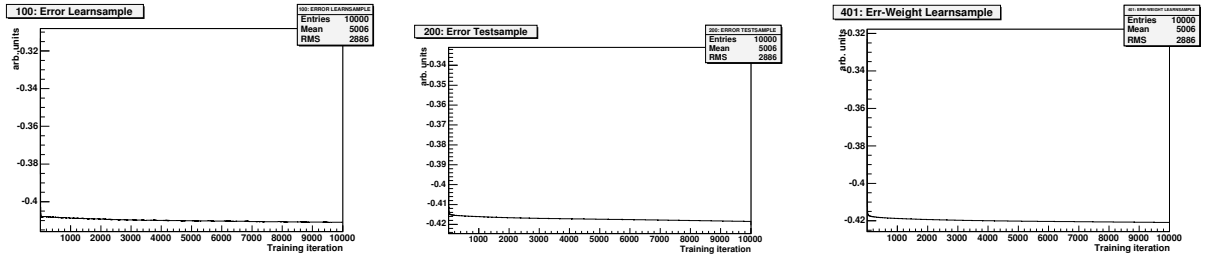
D.2 Network training and Output



$m_H = 110$ GeV/c^2

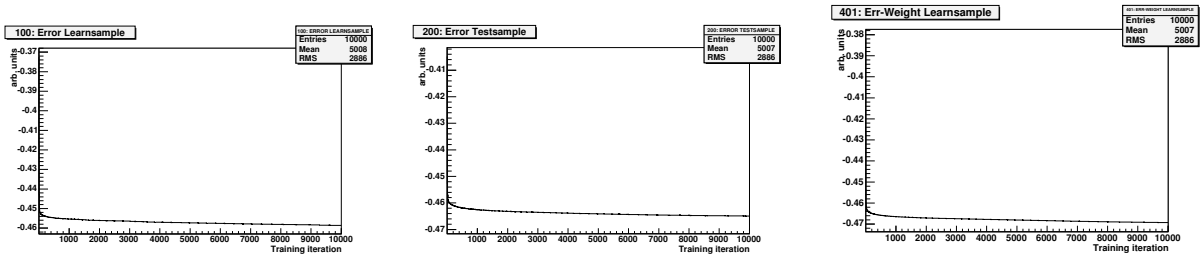


$m_H = 115$ GeV/c^2

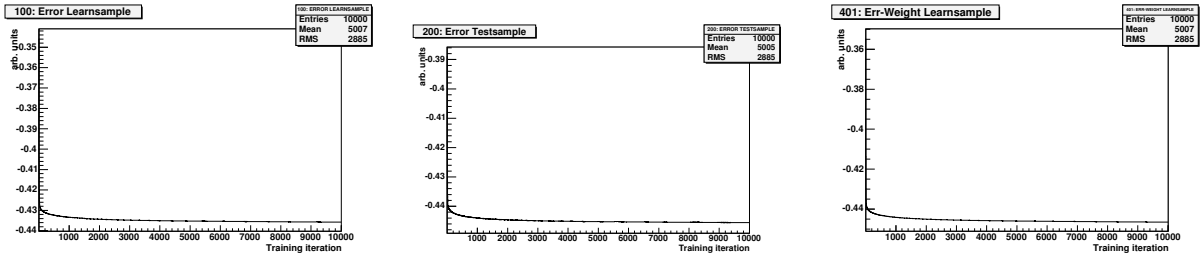


$m_H = 120$ GeV/c^2

Figure D.3: Error for the training (left) and test sample (middle) and the regularisation loss function (right) for Higgs boson masses of 110 GeV/c^2, 115 GeV/c^2 and 120 GeV/c^2. The latter one is the quantity that is actually minimised by the network.

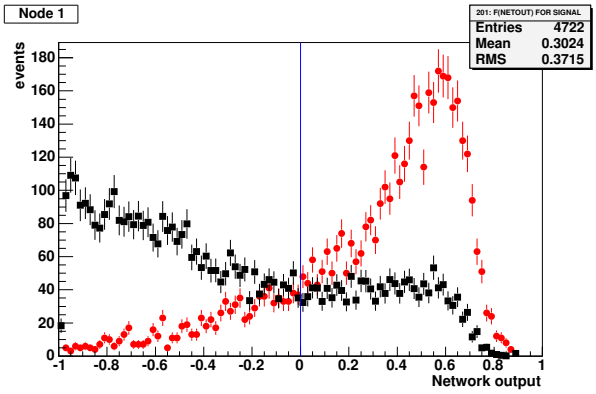


$m_H = 140$ GeV/c^2

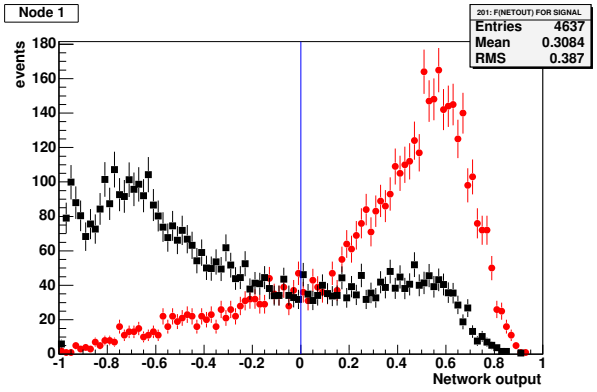


$m_H = 150$ GeV/c^2

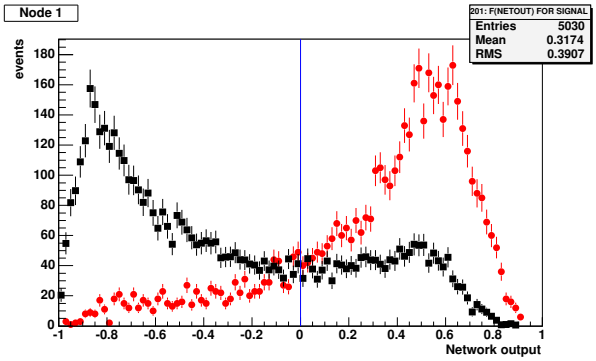
Figure D.4: Error for the training (left) and test sample (middle) and the regularisation loss function (right) for Higgs boson masses of 130 GeV/c^2, 140 GeV/c^2 and 150 GeV/c^2. The latter one is the quantity that is actually minimised by the network.



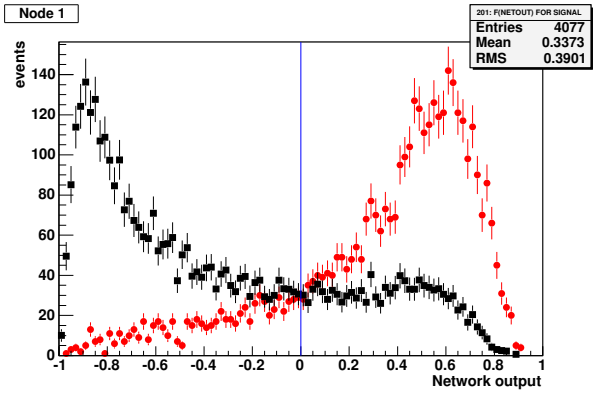
$m_H = 110$ GeV/c^2



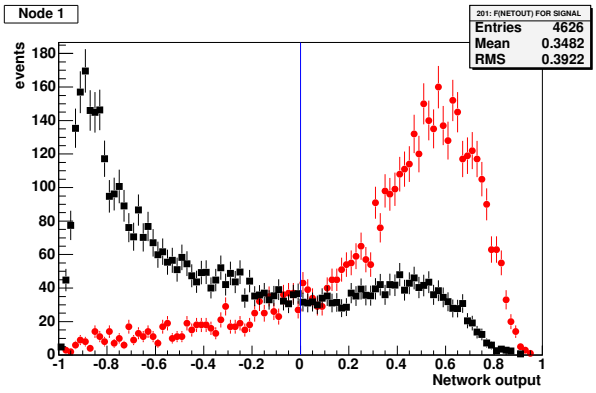
$m_H = 115$ GeV/c^2



$m_H = 130$ GeV/c^2

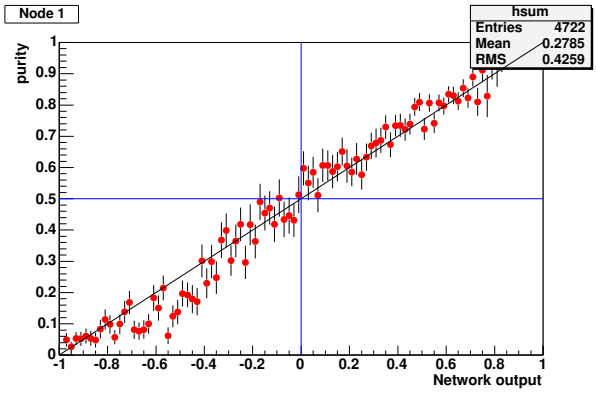


$m_H = 140$ GeV/c^2

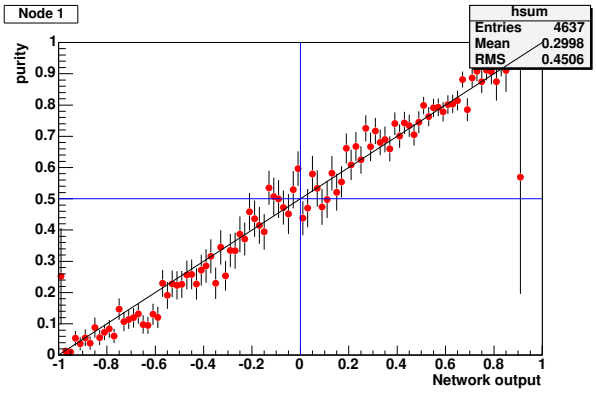


$m_H = 150$ GeV/c^2

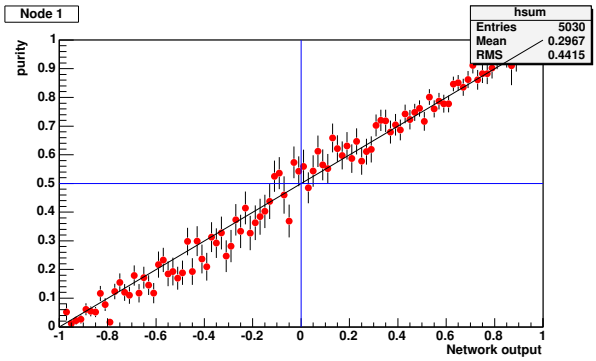
Figure D.5: Network outputs for signals (right, red) and background (left, black). Signal and background are weighted 1:1 for all six Higgs masses. The vertical line shows where the cut was made.



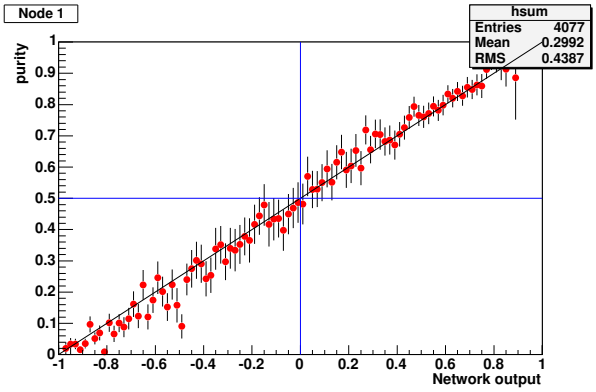
$m_H = 110$ GeV/c^2



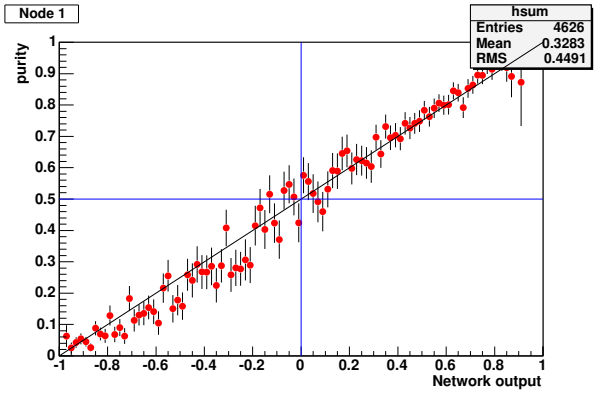
$m_H = 115$ GeV/c^2



$m_H = 130$ GeV/c^2

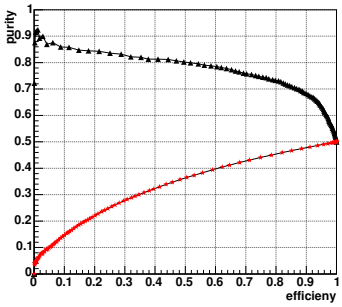


$m_H = 140$ GeV/c^2

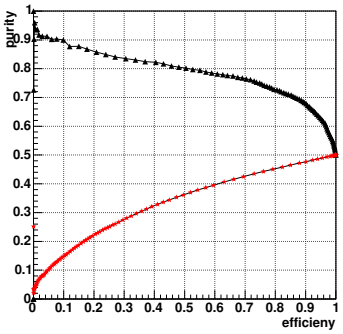


$m_H = 150$ GeV/c^2

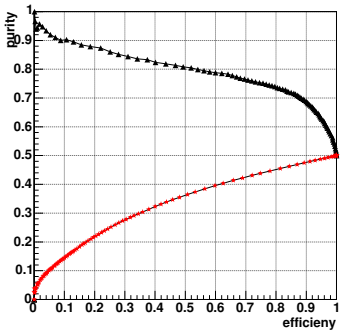
Figure D.6: Network output vs. purity for Higgs boson masses other than 120 GeV/c^2.



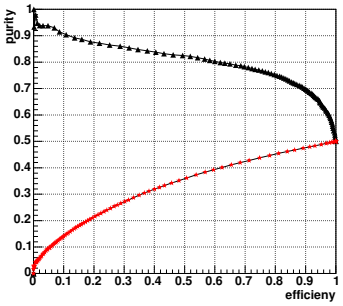
$m_H = 110$ GeV/c^2



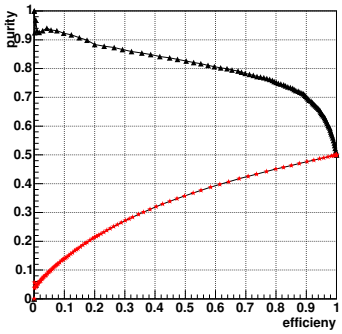
$m_H = 115$ GeV/c^2



$m_H = 130$ GeV/c^2

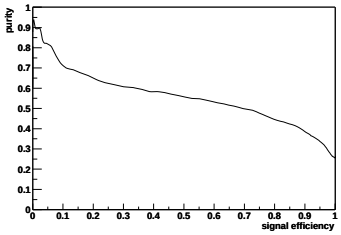


$m_H = 140$ GeV/c^2

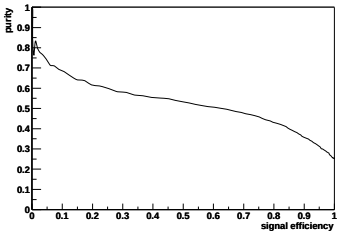


$m_H = 150$ GeV/c^2

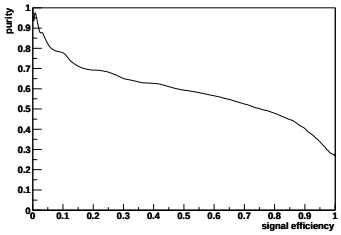
Figure D.7: Efficiency vs. purity for the NB-teacher for Higgs boson masses other than 120 GeV/c^2. The upper curves (black) show the distributions for the signal, the lower ones (red) for the background. An efficiency of 100% corresponds to a cut value of -1 on the network output. Hence the number of signal and background events are equal and the purity is 50% for both.



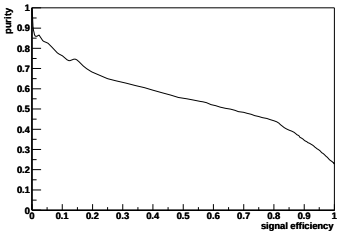
$m_H = 110$ GeV/c^2



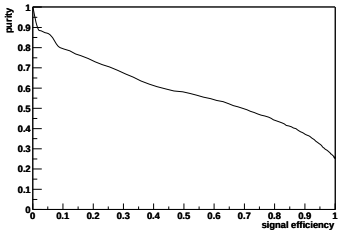
$m_H = 115$ GeV/c^2



$m_H = 130$ GeV/c^2



$m_H = 140$ GeV/c^2



$m_H = 150$ GeV/c^2

Figure D.8: Efficiency vs. purity for the NB-expert for Higgs boson masses other than 120 GeV/c^2. The curves show the distributions for the signal.

D.3 Correlation matrices

	T	$E_t^{j_1}$	δ_{12}	P_t^{Lep}	$M_{l\nu b\bar{b}}$	$Q \times \eta$	$\not{E}_t$	$M_{b\bar{b}}$
T	100	-	-	-	-	-	-	-
$E_t^{j_1}$	3.7	100	-	-	-	-	-	-
δ_{12}	-1.7	65.6	100	-	-	-	-	-
P_t^{Lep}	6.6	22.4	25	100	-	-	-	-
$M_{l\nu b\bar{b}}$	-2.9	70.1	67.2	54.4	100	-	-	-
$Q \times \eta$	-11.1	-0.3	0.9	0.7	0.6	100	-	-
$\not{E}_t$	-7.9	26.3	27.2	-14.1	7	-1.6	100	-
$M_{b\bar{b}}$	-0.8	60	45.6	-3.6	56.4	-1.4	7.6	100

Table D.1: Correlation matrix of network input variables in percent for a Higgs boson mass of $m_H = 110$ GeV/c^2.

	T	$E_t^{j_1}$	δ_{12}	P_t^{Lep}	$M_{l\nu b\bar{b}}$	$Q \times \eta$	$\not{E}_t$	$M_{b\bar{b}}$
T	100	-	-	-	-	-	-	-
$E_t^{j_1}$	9.5	100	-	-	-	-	-	-
δ_{12}	4.3	65.6	100	-	-	-	-	-
P_t^{Lep}	6.7	22.9	25.1	100	-	-	-	-
$M_{l\nu b\bar{b}}$	0.5	70.2	66.8	54	100	-	-	-
$Q \times \eta$	-12.5	-0.7	0.6	-0.3	-0.6	100	-	-
$\not{E}_t$	-5.7	26.7	27.6	-15.6	6.8	-1.8	100	-
$M_{b\bar{b}}$	3	60.1	45.7	-1.4	58.2	-1.8	8.1	100

Table D.2: Correlation matrix of network input variables in percent for a Higgs boson mass of $m_H = 115$ GeV/c^2.

	T	$E_t^{j_1}$	δ_{12}	P_t^{Lep}	$M_{l\nu b\bar{b}}$	$Q \times \eta$	$\not{E}_t$	$M_{b\bar{b}}$
T	100	-	-	-	-	-	-	-
$E_t^{j_1}$	19.8	100	-	-	-	-	-	-
δ_{12}	14.4	66.1	100	-	-	-	-	-
P_t^{Lep}	8.9	23	24.7	100	-	-	-	-
$M_{l\nu b\bar{b}}$	11.8	71	67.4	54.1	100	-	-	-
$Q \times \eta$	-11.6	-3.3	-2.2	0.1	-2	100	-	-
$\not{E}_t$	-2.8	26.4	27.2	-15.5	6.2	-1.5	100	-
$M_{b\bar{b}}$	14.2	61.4	48.5	-2	60	-3.3	7.5	100

Table D.3: Correlation matrix of network input variables in percent for a Higgs boson mass of $m_H = 130$ GeV/c^2.

	T	$E_t^{j_1}$	δ_{12}	P_t^{Lep}	$M_{l\nu b\bar{b}}$	$Q \times \eta$	$\not{E}_t$	$M_{b\bar{b}}$
T	100	-	-	-	-	-	-	-
$E_t^{j_1}$	28.3	100	-	-	-	-	-	-
δ_{12}	23.2	66.2	100	-	-	-	-	-
P_t^{Lep}	9.6	24.5	26.4	100	-	-	-	-
$M_{l\nu b\bar{b}}$	19.6	71.5	68.2	54	100	-	-	-
$Q \times \eta$	2.6	-1.4	-1.1	-2	1.5	100	-	-
$\not{E}_t$	-2.4	26.4	27.5	-13.9	6.1	-4.2	100	-
$M_{b\bar{b}}$	20.9	61.8	49.3	-1.5	60.9	-0.9	6.5	100

Table D.4: Correlation matrix of network input variables in percent for a Higgs boson mass of $m_H = 140$ GeV/c^2.

	T	$E_t^{j_1}$	δ_{12}	P_t^{Lep}	$M_{l\nu b\bar{b}}$	$Q \times \eta$	$\not{E}_t$	$M_{b\bar{b}}$
T	100	-	-	-	-	-	-	-
$E_t^{j_1}$	32.3	100	-	-	-	-	-	-
δ_{12}	27.4	66.5	100	-	-	-	-	-
P_t^{Lep}	10.2	24.7	26.5	100	-	-	-	-
$M_{l\nu b\bar{b}}$	25.3	72.6	69.2	53.4	100	-	-	-
$Q \times \eta$	3.6	0.4	0.6	-0.3	3.1	100	-	-
$\not{E}_t$	-0.5	27	28	-14.9	6.5	-4.8	100	-
$M_{b\bar{b}}$	26.5	62.9	51.1	-0.2	63	0.6	7.2	100

Table D.5: Correlation matrix of network input variables in percent for a Higgs boson mass of $m_H = 150$ GeV/c^2.

Glossary

barn *The unit barn is defined as* $10^{-28} m^2$. *Page 22*

λ *Nuclear interaction length. The mean free path between inelastic interactions, measured in* $g\,cm^{-2}$. *Page 29*

neuron *A single nerve cell, including its axons and dendrites. Page 61*

Poisson distribution *The definition of a poisson distribution is given in equation 6.1 on page 113. Page 23*

quadrupole *A magnet consisting of four poles, used for focusing beams of particles.* A quadrupole focuses the beam in one plane while defocussing it in the perpendicular direction. Hence a series of quadrupoles with alternating polarity is needed to result in a net focusing in both planes. *Page 21*

ramping *To excite a magnet with a time dependent excitation current.* A time varying excitation current for a magnet used either to track changing beam energy or to reduce average power dissipation by reducing the excitation when the magnetic field is not required. At Fermilab, the Main Ring and Tevatron bending magnetic field varies with the energy of the accelerating protons in order to confine them within the Main Ring beam pipe. *Page 20*

secondary vertex *A displaced vertex wrt. the primary vertex. Page 34*

shot *The injection of protons and anti-protons into the Tevatron in preparation for colliding beams operation. Page 20*

store *To inject circulating beam into an accelerator and keep it there for long periods of time.* In the Tevatron this means to inject protons and $\bar{p}$ bunches, cog the bunches to their proper collision points, ramp the Tevatron to 980 GeV and the collider detector experiments are taking data. In the antiproton source, it means that an antiproton stack is established. *Page 20*

synapse *The junction across which a nerve impulse passes from an axon terminal to a neuron or muscle cell. Page 61*

X_0 *Radiation length, usually measured in* $g\,cm^{-2}$. *It is the mean distance over which a high energy electron losses all but 1/e of its energy. Page 29*

List of Acronyms

ANN ..... Artificial Neural Network. See also NN. *Page 61*

ATLAS ..... A Toroidal LHC Apparatus. *Page 14*

BSC ..... Beam Shower Counters *Page 32*

CDF ..... Collider Detector at Fermilab *Page 26*

CERN ..... Counseil Européenne pour la Recherche Nucléaire *Page 41*

CES ..... Central Electromagnetic strip/wire gas chamber *Page 29*

CETA ..... Centrum für Elementarteilchenphysik und Astroteilchenphysik *Page 12*

CLC ..... Čerenkov Luminosity Counter *Page 32*

CMP ..... Central Muon Upgrade *Page 31*

CMS ..... Center of Mass System *Page 45*

CMS ..... Compact Muon Solenoid. An experiment for the Large Hadron Collider at CERN. *Page 14*

CMX ..... Central Muon Extension *Page 31*

COT ..... Central Outer Tracker *Page 29*

CPR ..... Central Pre-Radiator *Page 29*

CPU ..... Central Processing Unit *Page 34*

CR ..... Cosmic Ray *Page 88*

CSL ..... Consumer-Server/Logger *Page 83*

CTEQ ..... The Coordinated Theoretical-Experimental Project on QCD *Page 97*

D0 ..... An experiment located at the $p\bar{p}$-collider Tevatron, Fermilab. *Page 14*

DAQ ..... Data Acquisition *Page 33*

DOE ..... Department of Energy *Page 23*

DPE ..... Double-Pomeron Exchange *Page 32*

DQM ..... Data Quality Monitoring *Page 83*

EM ..... Electromagnetic *Page 29*

FNAL ..... Fermi National Accelerator Laboratory *Page 8*

FSR ..... Final State Radiation *Page 96*

IEKP ..... Institut für Experimentelle Kernphysik *Page 12*

IMU ..... Intermediate Muon System *Page 31*

IP ..... Interaction Point *Page 19*

ISR ..... Initial State Radiation *Page 96*

JES ..... Jet Energy Scale *Page 95*

JETCLU ..... Jet Cluster. A CDF jet reconstruction algorithm. *Page 85*

L00 ..... Layer 00 *Page 27*

L1 ..... Level-1 Trigger *Page 33*

L2 ..... Level-2 Trigger *Page 33*

L3 ..... Level-3 Trigger *Page 33*

LEP ..... Large Electron Positron collider *Page 45*

LHC ..... Large Hadron Collider *Page 14*

LINAC ..... Linear Accelerator *Page 19*

LO ..... Leading Order *Page 51*

MC ..... Monte Carlo. Monte Carlo methods are algorithms for solving various kinds of computational problems by using random numbers. *Page 81*

MI ..... Main Injector *Page 19*

MIP ..... Minimum Ionising Particle *Page 88*

MP ..... Miniplug Calorimeter *Page 33*

MRST ..... Martin, Roberts, Stirling, Thorne. The authors of a particular set of PDF's. *Page 97*

MSSM ..... Minimal Supersymmetric Standard Model *Page 46*

NB ..... NeuroBayes® *Page 78*

NLO ..... Next-to-Leading Order *Page 9*

NN ..... Neural Network. See also ANN. *Page 61*

PDF ..... Parton Distribution Functions *Page 97*

PDG ..... Particle Data Group *Page 56*

PEM ..... Plug Electromagnetic Calorimeter *Page 29*

PHA ..... Plug Hadron Calorimeter *Page 29*

PHI-T ..... Physics Information Technologies *Page 12*

PMT ..... Photo Multiplier Tube *Page 32*

QCD ..... Quantum Cromo Dynamics *Page 10*

QED ..... Quantum Electro Dynamics *Page 38*

RF ..... Radio Frequency *Page 20*

SD ..... Single Diffraction *Page 32*

SECVTX ..... Secondary Vertex. A CDF b-tagging algorithm based on the reconstruction of secondary vertices. *Page 83*

SL ..... Super Layer *Page 87*

SM ..... Standard Model *Page 38*

SVT ..... Silicon Vertex Tracker *Page 34*

SVX II ..... Silicon Vertex Detector II *Page 27*

TEVATRON . Fermilab's 2 TeV proton-antiproton accelerator, the world's highest-energy accelerator. *Page 8*

WHA ..... Endwall Hadron Calorimeter *Page 29*

WLS ..... Wavelength Shifting *Page 33*

XFT ..... Extremely Fast Tracker *Page 34*

Author Index

Page numbers given in *italics* refer to pages within the bibliography.

List of Figures

List of Tables

Bibliography

[1] Giulia Manca. A measurement of the ratio $R = \frac{\sigma_w \cdot br(w \to e^{\pm}\nu)}{\sigma_z \cdot br(z \to e^+e^-)}$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. 2003. CDF Note 6521. 9, 26, 28

[2] Christopher Neu. A Search for the Higgs Boson in Proton - Antiproton Collisions at $\sqrt{s} = 1.8$ TeV. 2003. CDF Note 6684. 11, 63, 100

[3] Physics Information Technologies GmbH, http://www.phi-t.de/. 12, 78

[4] Marcela Carena, J. S. Conway, H. E. Haber, J. D. Hobbs, et al. Report of the Tevatron Higgs Working Group. 2000. hep-ph/0010338. FERMILAB-CONF-00-279-T, FERMILAB-PUB-00-349 . 13, 16, 117, 118

[5] R. Barate et al. Search for the Standard Model Higgs Boson at LEP. *Phys. Lett.* **B 565** *(2003) 61*. hep-ex/0306033. 14, 123

[6] S. Richter. B-tagging with high p_t jets in CMS. Ph.D. thesis in preparation. 14, 124

[7] Roger Penrose. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford University Press, November 1989. ISBN 0198519737. 17

[8] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.* **A 12** *(1964) 132*. 17, 38, 47

[9] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.* **145** *(1966) 1156*. 17, 38

[10] J. D. Cockroft and E. T. S. Walton. Artificial Production of Fast protons. *Nature* **129** *(1932) 242*. 19

[11] U.S. Department of Energy Review of Accelerator Run II (http://www-bd.fnal.gov/doereview03/). The Run II Luminosity Upgrade at the Fermilab Tevatron: Project Plan and Resource-Loaded Schedule. June 2003. 21

[12] J. A. MacLachlan. Prospectus for an Electron Cooling System for the Recycler. November 1998. Fermilab-TM-2061. 22

[13] http://www-ap.fnal.gov/ecool/overview.html. 22

[14] http://www-bd.fnal.gov/runII/parameters.pdf. 22

[15] http://www-bd.fnal.gov/doereview02/. 23

[16] U.S. Department of Energy Review of Accelerator Run II (http://www-bd.fnal.gov/doereview03/). Department of Energy Assessment of the Run II Luminosity Plan at the Fermilab Tevatron. July 2003. Final report. 23

[17] The CDF Collaboration. The CDF II Detector Technical Design Report. November 1996. Fermilab-Pub-96-390-E. 26

[18] D. Acosta et al. The CDF Cherenkov Luminosity Monitor. *Nucl. Instrum. Meth.* **A461** *(2001) 540*. CDF Note 5559. 32

[19] The Beam Shower Counters for Run II. October 2000. CDF Note 5247. 32

[20] Stefano Lami. The CDF Miniplug Calorimeters. *Nucl. Instrum. Meth.* **A496** *(2003) 333*. hep-ex/0207017. CDF Note 5995. 33

[21] http://rock23.rockefeller.edu/~michgall/fd/fd_rps.html. 33

[22] http://www-cdf.fnal.gov/internal/upgrades/daq_trig/twg/tools/trigopts.ps. 33

[23] K. Anikeev, G. Bauer, I. K. Furic, D. Holmgren, A. Korn, I. Kravchenko, M. Mulhearn, P. Ngan, Ch. Paus, A. Rakitine, R. Rechenmacher, T. Shah, P. Sphicas, K. Sumorok, S. Tether, and J. Tseng. Event-Building and PC Farm based Level-3 Trigger at the CDF Experiment. 1999. CDF Note 5051. 34

[24] http://www-cdfonline.fnal.gov/consumer/home/consumer_home.html. 34

[25] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Perseus Press, September 1995. ISBN 0-201-503972. 37

[26] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories in Particle Physics*. Institut of Physics Publishing, 1996. ISBN 0-85274-328-9. 37

[27] Francis Halzen and Alan D. Martin. *Quarks & Leptons*. John Wiley & Sons, Inc., 1984. ISBN 0-471-88741-2. 37

[28] Christoph Berger. *Elementarteilchenphysik*. Springer Verlag, 2002. ISBN 3-540-41515-7. 37

[29] K. Hagiwara et al. Review of particle physics. *Phys. Rev.*, D66:010001, 2002. http://pdg.lbl.gov/. 37, 44

[30] S. L. Glashow. Partial symmetries of weak interactions. *Nucl. Phys.*, 22:579, 1961. 37

[31] J. Goldstone, A. Salam, and S. Weinberg. Broken symmetries. *Phys. Rev.*, 127:965, 1962. 37

[32] B. Aubert et al. Observation of a narrow meson decaying to D/s+ pi0 at a mass of 2.32-GeV/c**2. *Phys. Rev. Lett.* **90** *(2003) 242001*. hep-ex/0304021. 40

[33] D. Besson et al. Observation of a narrow resonance of mass 2.46-GeV/c**2 decaying to D/s*+ pi0 and confirmation of the D/sJ*(2317) state. *Phys. Rev.*, D68:032002, 2003. hep-ex/0305100. 40

[34] C. Alt et al. Observation of an exotic S = -2, Q = -2 baryon resonance in proton proton collisions at the CERN SPS. 2003. hep-ex/0310014. 40

[35] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.* **105** *(1957) 1413*. 41

[36] Ugo Amaldi, Wim de Boer, and Hermann Fürstenau. Comparison of grand unified theories with electroweak and strong coupling constants measured at LEP. *Phys. Lett.* **A B260** *(1991) 447*. CERN-PPE-91-44, IEKP-KA-91-01. 47

[37] V. L. Ginzburg and L. D. Landau. On the theory of superconductivity. *Zh. Eksp. Teor. Fiz.* **20** *(1950) 1064*. (In Russian). 46

[38] Ulrich Schwickerath. Suche nach neutralen Higgsbosonen im 4-Jet-Kanal mit dem DELPHI-Detektor am LEP. 1999. IEKP-KA/99-5. 48

[39] W. de Boer. Grand unified theories and supersymmetry in particle physics and cosmology. *Prog. Part. Nucl. Phys.* **33** *(1994) 201*. hep-ph/9402266. 49

[40] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Lett.* **A 81** *(1998) 1562*. hep-ex/9807003. 50

[41] A. Stange, William J. Marciano, and S. Willenbrock. Higgs bosons at the Fermilab Tevatron. *Phys. Rev.* **D49** *(1994) 1354*. hep-ph/9309294. 51

[42] Michael Spira. HIGLU and HDECAY: Programs for Higgs boson production at the LHC and Higgs boson decay widths. *Nucl. Instrum. Meth.* **A389** *(1997) 357*. hep-ph/9610350. 51, 125

[43] http://www.desy.de/~spira/. 51, 125

[44] S. Eidelman et al. Review of particle physics. *Phys. Lett.* **B B592** *(2004) 1*. 56

[45] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. John Wiley & Sons, Inc., 2000. ISBN 0-471-05669-3. 61, 66, 67, 68, 69, 176

[46] John Hertz, Anders Krogh, and Richard G. Palmer. *Introduction to the theory of Neural Computation*. Addison Wesley, 1991. ISBN 0-201-50395-6. 61

[47] Robert Hecht-Nielsen. *Neurocomputing*. Addison Wesley, 1989. ISBN 0-201-09355-3. 61

[48] Christopher M. Bishop. *Neural Networks for Patterns Recognition*. Oxford University Press, 1995. ISBN 0-19-853864-2. 61, 63

[49] http://www.ecs.soton.ac.uk/~kc/py301/amygdala.htm. 62

[50] Figures from [45] can be downloaded from http://rii.ricoh.com/~stork/DHS.html. 66, 67, 68, 69

[51] Bettina Hartmann. Searches for a 120 GeV Higgs Boson in the Channel $q\bar{q} \to WH \to e\nu_e b\bar{b}$ using neural networks. 2002. IEKP-KA/2002-7. 72

[52] Michael Feindt. A Neural Bayesian Estimator for Conditional Probability Densities. physics/0402093, 2004. IEKP-KA/04-05. 72, 78

[53] Zoltán Albrecht. Analyse von orbital angeregten B-Mesonen. 2003. IEKP-KA/2003-16. 74

[54] Physics Information Technologies. The NeuroBayes®User's Guide. March 2004. unpublished. 78

[55] C. Lecci. How to use NeuroBayes®in CDF software. October 2003. unpublished. 78

[56] http://www-cdf.fnal.gov/internal/dqm/dqm.html. 83

[57] S. Klimenko, J. Konigsberg, and T. Liss. Average cross-section of the inelastic PPbar scattering at 1800 TeV. 2003. CDF Note 6314. 83

[58] S. Klimenko, J. Konigsberg, S. M. Wang, D. Tsybychev, A. Sukhanov, S. Pronko, V. Necula, D. Acosta, G. Mitselmakher, A. Korytov, and G. Lungu. A first look at the CLC Luminosity Measurements. 2002. CDF Note 6052. 83

[59] E. Thomson. Description of data samples for Top Group for Summer 2003. 2003. CDF Note 6548. 83

[60] http://www-cdf.fnal.gov/internal/physics/top/topdata/newtopdata.html. 83

[61] Torbjörn Sjöstrand, Leif Lönnblad, and Stephen Mrenna. PYTHIA 6.2: Physics and manual. 2001. hep-ph/0108264. 83

[62] G. Corcella et al. HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *J. High Energy Phys.* **01** *(2001) 010*. hep-ph/0011363. 83, 84, 96

[63] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *J. High Energy Phys.* **07** *(2003) 001*. hep-ph/0206293. 83, 84, 96

[64] A. Dominguez, João Guimarães da Costa, A. Holloway, and S. Rappoccio. The SecVtx Tag Rate Matrix for the 2003 Summer Conferences. 2003. CDF Note 6533. 83

[65] C. Rott, H. Bachacou, K. Hoffman, A. Dominguez, W. Yao, D. Glenzinski, D. Bortoletto, M. Shochet, and F. delli Paoli. SecVtx Optimization for the 2003 Winter Conferences. 2003. CDF Note 6242. 83, 86

[66] C. Ferretti. A general interface to ALPGEN. 2003. CDF Note 6547. 84

[67] H. Bachacou, J. Nielsen, and W. Yao. Optimized Measurement of the t-tbar Production Cross Section in the SECVTX Tagged Sample. 2004. CDF Note 6902. 85, 91

[68] K. Bloom, K. Ebina, E. Thomson, and K. Yorita. Event Selection and Detection Efficiency for Winter 2004 Lepton+Jets Analyses. 2004. CDF Note 6844. 85, 87

[69] J. F. Arguin, A. A. Bhatti, C. A. Currat, A. Gibson, A. Galtieri, B. Heinemann, Latino G., and M. Shochet. Jet Corrections for Summer 2003 Analyses. 2003. CDF Note 6565. 86, 95

[70] http://fcdfwww.fnal.gov/internal/physics/top/jets/corrections.html. 86

[71] Anyes Taffard. Run II Cosmic Ray Tagger. 2003. CDF Note 6100. 86

[72] H. Bachacou, K. Bloom, K. Ebina, N. Goldschmidt, J. Kang, A. Messina, V. Necula, K. Sato, E. Thomson, R. Vilar, P. Wittich, and K. Yorita. $t\bar{t}$ Event Selection and Detection Efficiency for Winter 2003 Lepton+Jets Analyses. 2003. CDF Note 6084. 86

[73] http://fcdfwww.fnal.gov/internal/physics/top/r2leptons/etf/etf_main.html. 86

[74] Robert G. Wagner. Electron Identification for Run II: Algorithms. 2003. CDF Note 5456. 86

[75] Robert G. Wagner. Understanding and Using Lshr. 2003. CDF Note 6249. 87

[76] P. Wilson. Calorimeter Isolation and Lateral Shower Leakage for Photons and Electrons. 1998. CDF Note 4170. 87

[77] B. Heinemann and A. Wyatt. Correction for Leakage Energy in the Central and Plug Calorimeters in Run II. 2002. CDF Note 6167. 87

[78] http://fcdfwww.fnal.gov/internal/upgrades/computing/projects/reconstruction/muon/muon_software.html. 88

[79] V. Martin. High pt muons, recommended cuts and efficiencies for release 5.3.1. 2004. CDF Note 7031. 89

[80] A. Hocker and W. K. Sakumoto. Event $|Zvtx| < 60$ cm Cut Acceptance for Run II. 2004. CDF Note 6917. 89

[81] M. Coca, E. Halkiadakis, and S. Lockwitz. Central Electron Identification Efficiencies for the 200 pb^{-1} Run 2 Dataset. 2004. CDF Note 6580. 89

[82] V. Martin and L. Cerrito. Muon Cuts and Efficiencies for 4.11 Analyses. 2004. CDF Note 6825. 89, 118

[83] Y. K. Kim, J. Nielsen, L. Tompkins, and G. Veramendi. Trigger Efficiencies for High p_t Electrons. 2004. CDF Note 6234. 89, 118

[84] H. Bachacou, T. Wright, and W. Yao. Efficiency of SecVtx B-Tag Used for 2004 Winter Conferences. 2004. CDF Note 6904. 89, 90, 118

[85] H. Bachacou, J. Nielsen, and W. Yao. Preliminary Studies of Heavy Flavor Contents in W+jets. 2003. CDF Note 6570. 91, 118

[86] H. Bachacou, C. Ferretti, J. Nielsen, and W. Yao. Heavy Flavor Contributions to the SECVTX-tagged W+Jets Sample. 2004. CDF Note 7007. 91, 118

[87] H. Bachacou, J. Nielsen, and W. Yao. Non-W background in the lepton + jets SECVTX tagged sample. 2004. CDF Note 6559. 92

[88] D. Amidei, H. Bachacou, J. S. Miller, J. Nielsen, T. Wright, and W. Yao. Measurement of the t-tbar Cross Section using Lepton + Jets Events with Secondary Vertex Tagging. 2004. CDF Note 7138 (To be submitted to Phys. Rev. D). 92

[89] J. F. Arguin, A. A. Bhatti, C. A. Currat, R. Erbacher, A. Galtieri, R. Harris, B. Heinemann, H. Kim, Latino G., J. Lys, R. Madrak, M. Martinez, W. Sakumoto, M. Shochet, and M. Schmitt. Systematic Uncertainties Associated with Jet Corrections for Winter 2003 Conferences. 2003. CDF Note 6419. 95

[90] http://hep.physics.utoronto.ca/JeanFrancoisArguin/JetCorrDoc/SystUncert.html. 95

[91] Feindt, Michael. private communication. 100

[92] D. Acosta, V. Necula, P. Pfeiffer, M. Schmitt, A. Sukhanov, D. Tsybychev, and S. M. Wang. Introduction to Run II Jet Probability Heavy Flavor Tagger. 2003. CDF Note 6315. 100

[93] E. Thomson, B. Winer, R. Hughes, P. Koehn, J. Y. Chung, M. Coca, K. Bloom, E. James, A. Castro, A. Gresele, and N. Goldschmidt. Top and Electroweak Event Classification Module for CDF Run II. 2003. CDF Note 5947. 101

[94] C. Ciobanu, P. Savard, B. Stelzer, E. Thomson, W. Wagner, and T. Walter. Single Top Acceptances in Run II at CDF. 2003. CDF Note 6508. 101

[95] M. Erdmann, D. Hirschbühl, Y. Kemp, Müller T., H. Stadie, and W. Wagner. A study of single top quark production using event shape and rapidity observables. 2002. CDF Note 5976. 101, 102

[96] R. J. Barlow. Jets in high-energy interactions. *Rept. Prog. Phys.* **56** *(1993) 1067.* MAN-HEP-93-05. 102

[97] S. Catani, Yuri L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant K(t) clustering algorithms for hadron hadron collisions. *Nucl. Phys.* **B 406** *(1993) 187.* CERN-TH-6775-93, LU-TP-93-2. 102

[98] J. Conway and K. Maeshima. Poisson Upper Limits Incorporating Uncertainties in Acceptance and Background. 1998. CDF Note 4476. 113

[99] http://www-cdf.fnal.gov/physics/statistics/statistics_software.html. 113

[100] R. J. Barlow. *Statistics: A guide to the Use of Statistical Methods in the Physical Sciences*. John Wiley & Sons, 1989. ISBN 0-471-92294-3. 113

[101] S. Amerio, T. Dorigo, and I. Lazzizzera. Hyperballs User Guide. 2004. CDF Note 6910. 123

[102] S. Amerio, T. Dorigo, I. Lazzizzera, L. Scodellaro, J. Conway, T. Junk, J. Kroll, M. Martinez-Perez, F. Ukegawa, B. Winer, W. Yao, M. Hennecke, and P. McNamara. Narrowing the Higgs mass peak: the hyperball algorithm. 2003. CDF Note 6450. 123

Acknowledgements

Mein erster Dank gilt Herrn Prof. Dr. Wim de Boer, der es mir ermöglichte diese Arbeit am IEKP durchzuführen und so die Arbeit in internationalen Kollaborationen kennenzulernen. Herr Prof. Dr. Michael Feindt war so freundlich, das Netzwerkpaket NeuroBayes® zu entwickeln und das Koreferat zu übernehmen. Außerdem bedanke ich mich bei

- dem Graduiertenkolleg "Elementarteilchenphysik an Beschleunigern" und dem Graduiertenkolleg "Hochenergiephysik und Teilchenastrophysik" für die Gewährung eines Stipendiums und die finanzielle Unterstüzung,
- dem Land Baden-Würtemberg für die Gewährung eines Landes-GK-Stipendium,
- dem Deutschen Akademischen Austauschdienst DAAD für die Gewährung eines Stipendium während meiner Zeit in den USA,
- dem Institut für Experimentelle Kernphysik und Herrn Prof. Dr. Thomas Müller für die Entsendung ans Fermilab und die Weiterbeschäftigung nach Ablauf der Stipendien,
- der Universities Research Association URA für den Besuch in Washington D.C.,
- Prof. Dr. Gerd Schön für die Genehmigung diese Arbeit in Englisch zu verfassen,
- Dr. Thorsten Walter für nützliche Diskussionen und Tips zur Physik und zum Top-Ntuple,
- Dr. Cigdem Issever für den netten Zuspruch in schwierigen Zeiten,
- Dr. Andreas Korn für fur die Organisation des "physics round table",
- Dr. Carsten Rott für die vielen Sushi-Abende,
- Prof. Frank Würthwein für die nette Zusammenarbeit am Monte Carlo Projekt,
- Dr. Stefan Stonjek, der mich auf die Möglichkeiten von pdflatex hingewiesen hat,
- Dr. Andreas Heiss für die Generierung der NeuroBayes®-Lizenzen,
- Dr. Muge Karagoz Unel und Dr. Marcin Wolter für die Einladungen zum Kaffee,
- Andy Hass, Amber Jenkins, Jun Zhang und Reid Mumford für die Mitarbeit als GSA-Officer,

- Dr. Bettina Hartmann, die mit ihrer Diplomarbeit vorarbeiten zu dieser Studie geleistet hat,
- den SysAdmins des IEKP für die Instanthaltung des EKP*plus*-Clusters und der Desktops,
- den Sprudeleinkäufern Sven Schalla und Joachim Heuser
- und dem Kaffeeeinkäufer Dr. Markus Moch.

Frankfurt, Juli 2005 Martin Hennecke

Index