Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Suche nach dem Higgs-Boson im WH-Kanal mit dem CDF-II-Experiment

Thomas Peiffer

Diplomarbeit

bei Prof. Dr. Th. Müller
Institut für Experimentelle Kernphysik

Korreferent Priv.-Doz. Dr. W. Wagner
Institut für Experimentelle Kernphysik

an der Fakultät für Physik
der Universität Karlsruhe (TH)

15. Januar 2008

Zusammenfassung

In der vorliegenden Analyse wird eine Suche nach dem Higgs-Teilchen, dem letzten unentdeckten Bestandteil des Standardmodells der Teilchenphysik, präsentiert. Die grundlegende Idee hinter dem Standardmodell ist das Prinzip der lokalen Eichinvarianz, das besagt, dass jeder physikalische Prozess unabhängig von der Wahl des Koordinatensystems und der Parametrisierung in jedem Raum-Zeit-Punkt ist. Das Higgs-Boson dient im Standardmodell der Vereinheitlichung der schwachen mit der elektromagnetischen Wechselwirkung sowie der Erzeugung von Massen aller weiteren Teilchen, ohne dabei die Eichsymmetrie zu verletzen.
Die heutige Untergrenze für die Masse des Higgs-Bosons, die von den Experimenten am Large Electron Positron Collider (LEP) bestimmt wurde, liegt bei ungefähr 115 GeV/c^2, wohingegen Messungen der Massen des top-Quarks und des W-Bosons eine nur geringfügig höhere Masse bevorzugen. Um Teilchen mit solch hoher Masse zu erzeugen, bedarf es der Kollision leichterer Teilchen mit einer Schwerpunktsenergie, die mindestens der Ruhemasse der zu erzeugenden neuen Teilchen entspricht. Der derzeit einzige Beschleuniger, der die Energie aufbringen kann, um Higgs-Bosonen zu erzeugen, ist der Tevatron-Beschleuniger am Fermi National Accelerator Laboratory (Fermilab) nahe Chicago, USA. Am Tevatron werden Protonen und Antiprotonen auf eine Energie von jeweils 980 GeV beschleunigt und zur Kollision gebracht. Die Produkte dieser Teilchenkollisionen werden mit den beiden Experimenten CDF II und DØ aufgezeichnet. In dieser Analyse werden die von CDF II gesammelten Daten zur Suche nach dem Higgs-Teilchen verwendet.
Es gibt mehrere mögliche Kanäle zur Erzeugung von Higgs-Bosonen in Proton-Antiproton-Kollisionen. Im hier betrachteten Kanal wird die assoziierte Produktion des Higgs-Teilchens mit einem W-Boson betrachtet, was am Tevatron für eine Higgs-Masse im Bereich von 120 GeV/c^2 der Kanal mit der höchsten erwarteten Sensitivität ist. Zum $p\bar{p} \to WH \to l\nu b\bar{b}$ Prozess existiert eine Reihe an weiteren physikalischen Prozessen, die denselben Endzustand aufweisen, beziehungsweise dieselbe Ereignissignatur im Detektor hervorrufen. Um ein Modell sowohl für das erwartete Higgs-Signal als auch für die verschiedenen Untergründe zu erhalten, werden so genannte Monte-Carlo-Generatoren verwendet, welche auf Zufallszahlen basierende Teilchenreaktionen simulieren. Andere Untergründe, wie zum Beispiel rein quantenchromodynamische Prozesse (QCD-Untergrund), werden aus Daten modelliert.
Für den W-Higgs-Prozess erwartet man im Detektor die folgende Signatur: Zwei hadronische Jets, die aus den beiden b-Quarks, den Zerfallsprodukten des Higgs-Bosons, hervorgehen, ein geladenes Lepton sowie fehlende Energiedeposition in Transversalrichtung zur Strahlachse. Letzteres wird durch das im Detektor nicht nachweisbare Neutrino hervorgerufen. Auf Trigger-Ebene werden bereits nur solche Ereignisse selektiert, die eine derartige Signatur aufweisen. In der weiteren Selektion werden weitere Schnitte angewendet, um die Zahl der Untergrundereignisse zu reduzieren. Dazu gehört die Forderung nach mindestens einer so genannten b-Markierung bei mindestens einem Jet. Weitere Schnitte wurden optimiert, um den QCD-Untergrund

zu reduzieren. Durch das in dieser Analyse entwickelte QCD-Veto gelingt es, die Zahl der QCD-Ereignisse soweit zu reduzieren, dass die meisten kinematischen Ereignisgrößen durch die Untergrundmodelle beschrieben werden, ohne dabei die Zahl der Signalereignisse zu stark zu verringern. Die Zahl der erwarteten W-Higgs-Ereignisse ergibt sich als Produkt aus dem theoretisch vorhergesagten Wirkungsquerschnitt, der integrierten Luminosität (einem Maß für die untersuchte Datenmenge) und der Effizienz, welche die Akzeptanzen der einzelnen Schnitte berücksichtigt. Für eine Masse des Higgs-Bosons von 120 GeV/c^2 erwartet man in der untersuchten Datenmenge von 1,9 fb^{-1} 3,5 Signalereignisse bei gleichzeitig rund 1400 Untergrundereignissen.
Um die Sensitivität weiter zu erhöhen, werden in der vorgestellten Analyse Neuronale Netze verwendet. Ein Neuronales Netz ist ein Algorithmus, der nach dem Vorbild menschlicher Gehirne modelliert wurde, um zum Beispiel zwei Klassen von Ereignissen anhand verschiedener Merkmale voneinander zu trennen. Hier werden Neuronale Netze auf Datensätzen trainiert, die sowohl Higgs-Ereignisse als auch die verschiedenen Untergründe enthalten. Während des Lernprozesses ist dem Algorithmus die Kategorie jedes Ereignisses, also Signal oder Untergrund, bekannt, und er versucht, die einzelnen Informationen der Eingabevariablen zu einer einzelnen Größe zu kombinieren. Angewendet auf den echten Daten, liefert diese Größe eine Wahrscheinlichkeit dafür, ob ein Ereigniss eher untergrund- oder signalartig ist.
Als Eingabevariablen werden verschiedene kinematische Größen verwendet, die alle auf ihre Modellierung durch die Untergrundmodelle hin überprüft wurden. Es werden nur solche Variablen verwendet, die durch die Untergrundmodelle gut beschrieben werden. Die wichtigsten Variablen, die eine Trennung von Higgs- von Untergrundereignissen ermöglichen, sind die rekonstruierte Masse des Higgs-Bosons, welches aus den beiden gemessenen Jets rekonstruiert wird, sowie die Ausgabe des Flavor-Separators, ebenfalls eines Neuronalen Netzes, welches Jets, die ein b-Quark enthalten, von solchen mit leichteren Quarks unterscheiden kann. Es werden separate Netzwerke für die Suche nach dem Higgs-Boson trainiert sowohl für unterschiedliche angenommene Massen des Higgs-Teilchens als auch für verschiedene Teilmengen des selektierten Datensatzes, bei denen entweder genau ein beziehungsweise zwei Jets als Jets, die ein b-Quark enthalten, identifiziert wurden. Untersucht wird ein Massenbereich des Higgs-Bosons von 110 bis 130 GeV/c^2.
Die Ausgabeverteilungen der Neuronalen Netze werden in einer Likelihood-Anpassung dazu verwendet, den Anteil an W-Higgs-Ereignissen in dem beobachteten Datensatz zu bestimmen. Die erwartete Sensitivität wird durch die Anwendung der Likelihood-Anpassung an in Pseudoexperimenten simulierte Datensätze bestimmt. Für eine signifikante Beobachtung reicht der analysierte Datensatz nicht aus. Aus der Unsicherheit auf das erhaltene Anpassungsresultat lässt sich allerdings durch Integration der Likelihood-Funktion eine obere Grenze auf den Wirkungsquerschnitt des W-Higgs-Prozesses bestimmen. Die erwarteten Obergrenzen liegen in etwa einen Faktor zehn über den theoretisch vorhergesagten Wirkungsquerschnitten, welche je nach Masse des Higgs-Bosons zwischen 0,06 und 0,16 pb liegen.
Die erwarteten Obergrenzen werden bestimmt für zwei Hypothesen: Die erste Hypothese basiert auf der Annnahme, dass in den untersuchten Daten Higgs-Ereignisse

enthalten sind, bei der zweiten Hypothese nimmt man an, dass kein Higgs-Boson im untersuchten Massenbereich existiert. Die auf dem CDF-II-Datensatz gemessenen oberen Grenzen sind kompatibel mit den aus den Pseudoexperimenten erwarteten Grenzen auf den Higgs-Produktionswirkungsquerschnitt für beide Hypothesen. Abbildung 1 fasst die Resultate dieser Analyse zusammen. In dem untersuchten Massenbereich ergeben sich Obergrenzen auf den Wirkungsquerschnitt zwischen 1,3 und 1,5 pb.

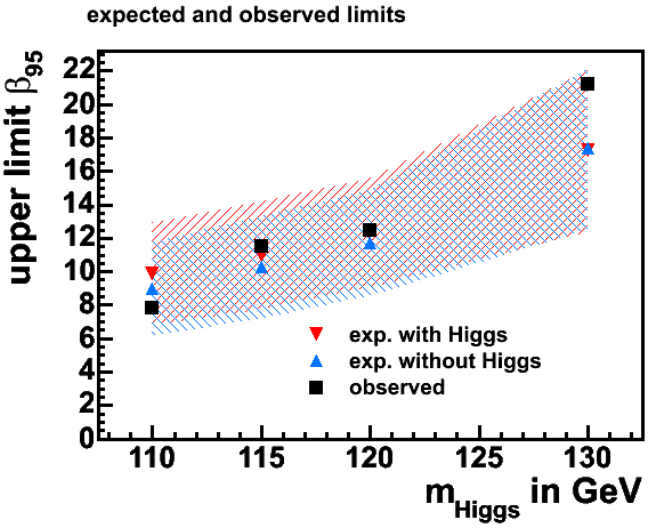

Abbildung 1: Die erwarteten und beobachteten Oberegrenzen β_{95} als Vielfache der vorhergesagten Wirkungsquerschnitte als Funktion der Higgs-Masse. Die gestreiften Bereiche markieren die $\pm 1\sigma$ Fehler.

Universität Karlsruhe (TH)
Research University • founded 1825

Search for the Higgs Boson in the WH Channel with the CDF II Experiment

Thomas Peiffer

Diploma Thesis

at the Faculty for Physics
of the University of Karlsruhe (TH)

Referee: Prof. Dr. Th. Müller
Institut für Experimentelle Kernphysik

Co-Referee: Priv.-Doz. Dr. W. Wagner
Institut für Experimentelle Kernphysik

January 15, 2008

Introduction

Today nearly all observable phenomena in particle interactions, covering a huge scale of energies, can be described by the Standard Model of particle physics. The idea of the Standard model is the principle of symmetries, which keep physical processes invariant. This principle is not a new feature introduced by the Standard Model. Already classical mechanics is based on symmetries. For instance a physical process is independent of the time and the space, where it occurs. This effect can be interpreted as such a symmetry of nature. The idea of physics to provide a quantitative, mathematical description of nature would become much more complicated or perhaps impracticable, if the laws of nature would not respect some basic symmetries. The basic symmetry of the Standard Model is a $SU(3) \times SU(2) \times U(1)$, where each subgroup corresponds to a rotation in a more abstract space to describe the three fundamental forces of the strong, the weak and the electromagnetic interactions. The postulation of this symmetry and the observations in the weak interaction made it necessary to introduce an additional particle, the Higgs boson, into the Standard Model. This particle is the last undiscovered particle of the Standard Model. A discovery of the Higgs boson would be a confirmation of the Standard Model, while the absence of this particle would pose the question, if the symmetry group of the Standard Model is realized as a fundamental symmetry of nature.
To find the last undiscovered piece of the Standard Model, the collisions of high-energetic particles have to be analyzed. The results of the experiments at the Large Electron Positron collider (LEP) at CERN give an lower limit to the mass of the Higgs boson of approximately 115 GeV/c^2. On the other hand the measurements of the masses of the top quark and the W boson prefer a mass not much higher than this lower bound. To reach higher mass regions, the energy in the particle interactions has to be increased in order to produce massive particles like the Higgs. In this analysis data taken with the CDF II experiment at the Tevatron collider at the Fermi National Accelerator Laboratory is used to find limits to the production cross section of the Higgs boson. At the Tevatron protons and antiprotons are accelerated to 980 GeV each and are brought to collisions. The scattering products are recorded by the CDF detector, a multi-purpose detector, which provides the energy and track measurement of electrons, muons and hadronic jets.
In this thesis a search for the Higgs boson in associated production with a W boson and a decay of the Higgs boson into a pair of bottom and anti-bottom quark using the CDF II data is presented. The associated Higgs production is the most sensitive Higgs production channel for an assumed Higgs mass of approximately 120 GeV/c^2. Nevertheless, many background processes occurring in proton antiproton collisions have cross sections, which are several orders of magnitude larger than the Higgs production cross section. To enrich the fraction of potential signal events in the data, several selection criteria are developed. The expected fraction of Higgs events in the data sample is still less than 1% after applying these selection cuts, wherefore a more powerful discrimination method is needed.

For the final discrimination between Higgs signal events and the remaining background neural networks are trained. A neural network is an algorithm, which is modelled according to the structure of neurons in a human brain. It combines the information of several inputs to get a more powerful separation variable. In this analysis neural networks are trained to distinguish signal-like events from the backgrounds, using kinematic variables of particles measured in the detector as inputs.
For the training of the neural networks computer simulated collision events for signal and background are used. In a following step the output distributions for these events are compared with the real data, applying a maximum log-likelihood fit. From this fit we obtain an upper limit to the cross section of the associated Higgs production.
In the first chapter of this thesis we give an overview over the theory of the Higgs mechanism. In the following chapter the Tevatron and the CDF experiment are described. The third chapter covers the different backgrounds of the W-Higgs process and their modeling with Monte Carlo event generators. In chapter four we give a summary of the applied selection criteria and their development. Also an estiamtion of the expected number of signal and background events is presented. In chapter five the basic functionality of neural networks is explained. The neural network analysis and the results of the likelihood fit are presented in chapter six. The last chapter gives a summary of this analysis and an outlook of the Higgs search at the Tevatron.

Contents

1 Theory of the Higgs Mechanism **3**
1.1 Lagrangian Formalism of Field Theory . . . 3
1.2 Local Gauge Symmetries and Conservation Laws . . . 4
1.3 Spontaneous Symmetry Breaking and Goldstone Theorem . . . 7
1.4 The Higgs Mechanism . . . 7
1.5 Yukawa Sector: Giving Masses to Fermions . . . 11
1.6 Constraints to the Higgs Mass . . . 13
1.7 Production and Decay Modes at the Tevatron . . . 14

2 The Experimental Site **19**
2.1 The Accelerator System . . . 19
2.2 The CDF II Experiment . . . 23
2.2.1 Tracking System . . . 24
2.2.2 Calorimetry . . . 25
2.2.3 Muon System . . . 27
2.3 Trigger System . . . 27

3 Background and Signal Modeling **31**
3.1 The Backgrounds of the W-Higgs Production . . . 31
3.2 Monte Carlo Event Generators . . . 32
3.3 Monte Carlo and Data Based Models . . . 36
3.3.1 Heavy Flavor Overlap Removal . . . 37

4 Selection of Events **39**
4.1 Lepton Selection . . . 39
4.2 Jet Selection . . . 40
4.2.1 b Tagging . . . 41
4.3 Missing Transverse Energy . . . 42
4.4 Further Selection Cuts . . . 43
4.5 QCD Veto . . . 43
4.6 Signal Expectation . . . 52
4.7 Background Estimation . . . 54
4.8 Event Reconstruction . . . 55

5 Neural Networks **59**
5.1 Basics of Neural Networks 59
5.2 Training the Network 61
5.3 Methods for an Efficient Training 62
5.3.1 Preprocessing 62
5.3.2 Automatic Selection of Relevant Variables 63
5.3.3 Optimization during the Training 63
5.4 The Network Output 63
5.5 Application: The Neural Network Flavor Separator 64

6 Measurement of upper Limits on the Higgs Cross Section **67**
6.1 Neural Network Analysis 67
6.2 Likelihood Fit 84
6.2.1 Systematic Uncertainties 86
6.2.2 Ensemble Tests 89
6.3 Results of the Likelihood Fit 93

7 Conclusions **95**
7.1 Summary 95
7.2 Results of Recent Analyses 96
7.3 Outlook 97

Chapter 1

Theory of the Higgs Mechanism

The current theory of particle physics is the so called Standard Model. The Standard Model has been developed about 40 years ago, but up to now it describes nearly all observed experimental data very accurately. One essential part of these model is the Higgs mechanism. The Higgs mechanism generates the masses of all particles by introducing a new field, which couples to all particles proportional to their masses. There is one remaining degree of freedom of this field, which should lead to an observable particle. This particle is the Higgs boson and the last undiscovered particle predicted by the Standard Model.
The basis of many field theories such as the Standard Model, but also of all supersymmetric models, is the principle of local gauge invariance, which says, that the physical observation does not depend on the choice of a theoretical parametrization of the process at each point in space and time. In the following sections a short overview how to construct gauge invariant theories is presented. From this we will conclude the necessity of a mechanism to create mass terms for the weak gauge bosons in the Standard Model. In the last sections the current theoretical and experimental limits and predictions of the parameters of the Higgs boson are presented.

1.1 Lagrangian Formalism of Field Theory

The Lagrangian formalism is an advantageous way to describe an elementary field theory. Especially for the discussion of symmetries and conservation laws the Lagrangian field theory is essential.
The Lagrangian density $\mathcal{L}$, often simply called Lagrangian, is a function depending on the field ϕ, the four dimensional space-time x_μ and the derivatives of the field $\frac{\partial \phi}{\partial x_\mu}$:

$$\mathcal{L} = \mathcal{L}\left(\phi, \frac{\partial \phi}{\partial x_\mu}, x_\mu\right). \tag{1.1}$$

The relativistic wave equation for an arbitrary field ϕ is gained by solving the Euler-Lagrange equation,which leads to the equation of motion for the field ϕ:

$$\frac{\partial}{\partial x_\mu}\left(\frac{\partial \mathcal{L}}{\partial\left(\partial\phi/\partial x_\mu\right)}\right)-\frac{\partial \mathcal{L}}{\partial \phi}=0. \tag{1.2}$$

For example, the Lagrangian of a real spin-0 field with mass m is given by:

$$\mathcal{L}=\frac{1}{2}\left(\partial_\mu\phi\right)\left(\partial^\mu\phi\right)-\frac{1}{2}m^2\phi^2. \tag{1.3}$$

By inserting equation 1.3 into the Euler-Lagrange equation 1.2 we obtain the Klein-Gordon equation:

$$\partial_\mu\partial^\mu\phi+m^2\phi=0. \tag{1.4}$$

The Lagrangian of a fermion field ψ, for example an electron, is given by:

$$\mathcal{L}=\bar{\psi}\left(i\gamma^\mu\partial_\mu-m\right)\psi, \tag{1.5}$$

where γ^μ are the Dirac matrices and $\bar{\psi}=\psi^\dagger\gamma^0$. In this Lagrangian ψ and $\bar{\psi}$ are treated as independent variables. Applying the Euler-Lagrange equation 1.2 for the conjugated field $\bar{\psi}$ we obtain the Dirac equation:

$$i\gamma^\mu\partial_\mu\psi-m\psi=0. \tag{1.6}$$

If we consider the derivatives of the field ψ in equation 1.2, we find the conjugated Dirac equation:

$$i\partial_\mu\bar{\psi}\gamma^\mu+m\bar{\psi}=0, \tag{1.7}$$

which is equivalent to equation 1.6 due to $\gamma^0\gamma^\mu=-\gamma^\mu\gamma^0$.

1.2 Local Gauge Symmetries and Conservation Laws

One fundamental principle of quantum field theory is the postulation of symmetries. From Noether's theorem we know, that every symmetry induces a conserved quantity. For example time invariance of a physical process leads to the conservation of energy, the invariance under rotation in space leads to the conservation of angular momentum. In the words of field theory: To every transformation of a field ϕ or of the space-time coordinates x_μ which let the Lagrangian invariant there exists a conserved quantity, called Noether current j_μ, with

$$\partial_\mu j^\mu=0. \tag{1.8}$$

With the conservation of the current also the Noether charge

$$Q:=\int d^3\vec{x}j^0 \tag{1.9}$$

is time invariant. This can be seen by using equation 1.8 and the Gaussian theorem:

$$\partial_t Q = \int d^3\vec{x}\partial_t j^0 = -\int d^3\vec{x}\nabla\vec{j} = -\oint d\vec{S}\vec{j} = 0, \tag{1.10}$$

if we assume a localized current with $j \to 0$ for $x_i \to \infty$.
As an example, the Dirac-Lagrangian 1.5 is invariant under a global $U(1)$ transformation, given by

$$\psi(x) \to e^{i\lambda}\psi(x), \tag{1.11}$$

with an arbitrary phase λ. It is possible to show (by demanding $\delta\mathcal{L} = 0$ for an infinitesimal transformation $\psi \to (1+i\lambda)\psi$) that the Noether current of this quantity is:

$$j^\mu = -e\bar{\psi}\gamma^\mu\psi, \tag{1.12}$$

where $-e$ is a constant chosen in such a way that the charge Q from equation 1.9 corresponds to the electric charge of the electron.
Up to now we only talked about global gauge symmetries. If the phase of the transformation λ depends on the space-time coordinates, i.e. $\lambda = \lambda(x)$, we speak about local gauge transformations. A general $SU(N)$ transformation, where $SU(N)$ is the set of all unitary $N \times N$ matrices with determinate 1, is then given by:

$$\psi(x) \to U(x)\psi(x), \tag{1.13}$$

where $U \in SU(N)$. Every element of $SU(N)$ can be written as:

$$U(x) = e^{i\lambda_a(x)T^a}, \tag{1.14}$$

where the T^a are the generators of the group. They form the Lie algebra of $SU(N)$, which means that the T^a are related as follows:

$$\left[T^a, T^b\right] = if^{abc}T^c. \tag{1.15}$$

Here the f^{abc} are the structure constants. The number of linear independent T^a is $N^2 - 1$. For example in $SU(2)$ we can choose the representation $T^a = \frac{\sigma^a}{2}$, where σ^a are the Pauli matrices. Then the structure constants are given by the epsilon tensor, $f^{abc} = \epsilon^{abc}$.
The concept of local gauge symmetry is one of the most successful theoretical models. The $SU(N)$ gauge theories are sometimes called Yang-Mills theories, since Yang and Mills were the first ones using this concept to explain the strong isospin in 1954 [4]. They treated the strong interaction as an $SU(2)$ theory with a dublett consisting of proton and neutron as elementary particles and the pion as gauge boson. Although it turned out that this was only an effective and not a fundamental theory, the concept of local gauge symmetry survived and is applied to describe the electroweak interactions ($SU(2)$) as well as the quantum chromodynamic ($SU(3)$) very successfully.
However, the Lagrangian of fermions given in equation 1.5 is not invariant under a local gauge transformation (eq. 1.13), since terms of the form $\partial_\mu\lambda_a(x)$ are gained. For that reason it is necessary to modify the Lagrangian by the introduction of a

gauge field A^a_μ. The gauge field goes into the covariant derivative, which replaces the ordinary derivative:

$$\partial_\mu \to D_\mu := \partial_\mu - igA^a_\mu T^a, \tag{1.16}$$

so that D_μ transforms as $D_\mu \to UD_\mu U^+$. The introduction of the covariant derivative leads to an additional term in the Lagrangian, which corresponds to a coupling of the particles ψ to the gauge fields A^a_μ. This term is given by $g\left(\bar{\psi}\gamma^\mu T^a\psi\right)A^a_\mu$. The strength of this coupling is parametrized by the coupling constant g, which is a free parameter of the theory. Now the Lagrangian is invariant under the gauge transformation. To take also the kinetic energy of the gauge field A_μ into account, we add the local gauge invariant term $\mathcal{L}_{\text{gauge}} = -\frac{1}{4}F^a_{\mu\nu}F^{\mu\nu,a}$ to the lagrangian. Here $F^a_{\mu\nu}$ is the field strength tensor defined as:

$$F^a_{\mu\nu} := \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu. \tag{1.17}$$

It is possible to add this term, since we still find the same wave equation, if we apply the Euler-Lagrange equation 1.2. The $F_{\mu\nu}$ term just vanishes. The full gauge invariant Lagrangian is:

$$\mathcal{L} = \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi + g\left(\bar{\psi}\gamma^\mu T^a\psi\right)A^a_\mu - \frac{1}{4}F^a_{\mu\nu}F^{\mu\nu,a}. \tag{1.18}$$

In summary, we introduced the fields A_μ to preserve the gauge invariance. In nature several particles, which can be described as a gauge particle of a local gauge symmetry, have been observed. For instance the photon is the gauge field of a $U(1)$ transformation, the gluons are the particles, that correspond to an $SU(3)$ invariance. In the weak interaction we also observe gauge fields, the W and the Z boson. But in contrast to the photon and the gluons, the W and the Z bosons are massive particles.

A mass term for a gauge boson with mass m contributes to the Lagrangian in form of a term like $\frac{m^2}{2}A^a_\mu A^{\mu,a}$. It turns out, that such a term is not gauge invariant under local gauge transformations. Only if $\partial_\mu\lambda_a = 0$, i.e. for global transformations, the mass term is invariant under an $SU(N)$ transformation. It can be shown that for an infinitesimal transformation $U = 1 + i\lambda_a T^a$ the gauge field transformation reads as follows:

$$A^a_\mu \to A^a_\mu + \lambda_c f^{abc}A^b_\mu + \frac{1}{g}\partial_\mu\lambda_a. \tag{1.19}$$

This leads to a transformation of the mass term:

$$\frac{m^2}{2}A^a_\mu A^{\mu,a} \to \frac{m^2}{2}\left(A^a_\mu A^{\mu,a} + \frac{2}{g}A^{\mu,a}\partial_\mu\lambda_a + O(\lambda_a^2)\right), \tag{1.20}$$

which is only invariant for $m = 0$ under a local gauge transformation. This theoretical prediction is in contradiction to the observation of massive gauge bosons in the weak interaction. To solve this problem without giving up the very successful concept of local gauge invariance, the Higgs mechanism has been introduced.

1.3 Spontaneous Symmetry Breaking and Goldstone Theorem

The Higgs mechanism uses an idea first developed in solid states physics, the spontaneous symmetry breaking. If the Lagrangian $\mathcal{L}$ is symmetric with respect to the group G, but the ground state of the vacuum $|0\rangle$ is not invariant under a transformation from G, G is called a spontaneously broken symmetry.
The **Goldstone theorem** says, that if the Lagrangian $\mathcal{L}$ is symmetric according to a group G_1 and $|0\rangle$ has the symmetry group $G_2 \subset G_1$, i.e G_1 is spontaneously broken to G_2, there exist $M := \mathrm{dimG}_1 - \mathrm{dimG}_2$ massless bosons, the so called Goldstone bosons. A proof for the Goldstone theorem can be found for instance in [6].
If $\mathcal{L} = \mathcal{L}(\Phi)$ with $\Phi = (\phi_1, \ldots, \phi_N)$ and at least one vacuum expectation value v_j of the field ϕ_j is not equal to zero,

$$v_j := \langle 0|\phi_j|0\rangle \neq 0, \tag{1.21}$$

every symmetry, which transforms the field ϕ_j non trivially, is spontaneously broken. This can be concluded by considering an infinitesimal transformation:

$$\phi_j \to \phi_j + i\lambda_a T^a_{jk}\phi_k. \tag{1.22}$$

If this symmetry is not broken, we must have:

$$\langle 0|\phi_j|0\rangle \stackrel{!}{=} \langle 0|\phi_j|0\rangle + i\lambda_a T^a_{jk}\langle 0|\phi_k|0\rangle \tag{1.23}$$

and therefore $v_k = \langle 0|\phi_k|0\rangle \stackrel{!}{=} 0$, since (for an irreducible representation T^a and at least one $\lambda_a \neq 0$) $\lambda_a T^a$ is not singular. To avoid that also the $SU(2)$ Lorentz symmetry, which describes the theory of relativity, is broken, the field ϕ_j has to be a scalar particle, since a violation of the Lorentz symmetry has not been observed in nature.

1.4 The Higgs Mechanism

In the electroweak section of the Standard Model the Higgs mechanism realizes a spontaneous symmetry breaking of the group $SU(2)_L \times U(1)_Y$, which unifies the electromagnetic and the weak force, to the subgroup $U(1)_{e.m.}$. Here $SU(2)_L$ is the gauge group of the weak interaction, L indicates, that only left handed fermions participate in weak interactions. Right handed fermions are singlets with respect to $SU(2)_L$. The corresponding quantum number is called weak isospin $(\vec{I}^2, I_3)$. The quantum number of $U(1)_Y$ is the so called hypercharge Y.
We introduce the spin-0 field Φ with the Lagrangian

$$\mathcal{L}_{\text{Higgs}} = (\partial_\mu \Phi^\dagger)(\partial^\mu \Phi) - V(\Phi), \tag{1.24}$$

where $V(\Phi)$ is the potential of the self interaction of the field Φ. The potential for the Higgs mechanism is parametrized using the two parameters μ and λ as follows:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda(\Phi^\dagger \Phi)^2. \tag{1.25}$$

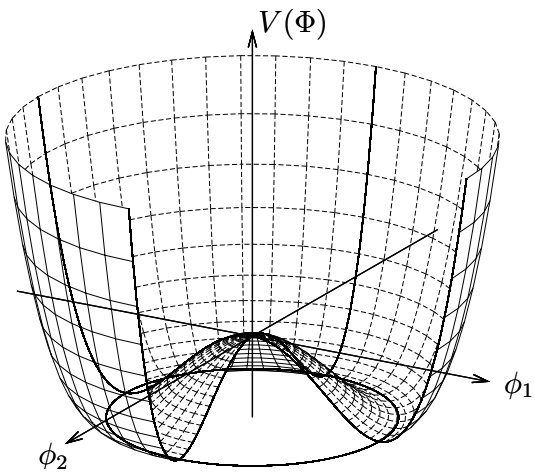

Figure 1.1: The potential of the Higgs field: The set of minima of the potential is a circle. By choosing one minimum to be the vacuum expectation value of the Higgs field, the symmetry of the field is broken [7].

The potential of the Higgs field is illustrated in figure 1.1. To break the $SU(2)_L \times U(1)_Y$ symmetry to $U(1)_{\text{e.m.}}$ with a non vanishing vacuum expectation value $\langle 0|\Phi|0\rangle$, Φ has to be at least an $SU(2)$ dublet:

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}, \tag{1.26}$$

with two complex scalar fields ϕ^+ and ϕ^0. Due to the invariance of $\mathcal{L}$ under an $SU(2)$ transformation we can choose the vacuum expectation value without loss of generality to $\begin{pmatrix} 0 \\ v \end{pmatrix}$ with $v \neq 0$, $v \in \mathbb{R}$.

- The hypercharge operator Y of $U(1)_Y$ transforms the field Φ as:

$$\Phi \to e^{i\varphi Y}\Phi, \qquad Y\Phi = y\Phi. \tag{1.27}$$

  For $y \neq 0$ the $U(1)_Y$ is broken by the field. We can choose the arbitrary value $y = 1$.

- As mentioned above, the field Φ is an $SU(2)$ dublet, and therefore the isospin operators act as:

$$\vec{I}^2\Phi = \frac{1}{2}\left(\frac{1}{2}+1\right)\Phi, \qquad I_3\Phi = \frac{1}{2}\begin{pmatrix} \phi^+ \\ -\phi^0 \end{pmatrix}. \tag{1.28}$$

We define the operator $Q := I_3 + \frac{Y}{2}$ and the transformation

$$\Phi \to e^{i\varphi Q}\Phi. \tag{1.29}$$

The vacuum expectation value of Φ is invariant under 1.29. This can be seen by the transformation of Φ under 1.29:

$$e^{i\varphi Q}\Phi = e^{i\varphi\left(I_3+\frac{Y}{2}\right)}\begin{pmatrix}\phi^+(x)\\ \phi^0(x)\end{pmatrix} = \begin{pmatrix} e^{i\varphi}\phi^+(x)\\ \phi^0(x)\end{pmatrix} \tag{1.30}$$

and $\langle 0|\Phi|0\rangle = \begin{pmatrix}0\\ v\end{pmatrix}$. For this reason the transformation 1.29 corresponds to the unbroken $U(1)_{\text{e.m.}}$ symmetry and Q can be identified as the operator of the electric charge. Applying Q to the fields ϕ^+ and ϕ^0 gives:

$$Q\phi^+ = 1\cdot\phi^+, \qquad Q\phi^0 = 0\cdot\phi^0. \tag{1.31}$$

Going back to the potential $V(\Phi)$ from equation 1.25 and demanding $\mu^2 < 0$, $\lambda > 0$, we find a minimum $\neq 0$. This minimum of the potential corresponds to the vacuum expectation value Φ, i.e. the value of the field in the minimum Φ_0 is equal to $\begin{pmatrix}0\\ v\end{pmatrix}$:

$$0 = \left.\frac{\partial V(\Phi)}{\partial\Phi^\dagger}\right|_{\Phi=\Phi_0} = -\mu^2\Phi_0 + 2\lambda(\Phi_0^\dagger\Phi_0)\Phi_0 = (-\mu^2 + 2\lambda v^2)\begin{pmatrix}0\\ v\end{pmatrix}. \tag{1.32}$$

From equation 1.32 follows $v = \frac{\mu}{\sqrt{2\lambda}}$. We can parametrize the fluctuation of the field around the minimum using four real fields θ_1, θ_2, θ_3, and h:

$$\Phi(x) = e^{\frac{i}{\sqrt{2}v}\sigma_j\theta_j(x)}\begin{pmatrix}0\\ v + \frac{h(x)}{\sqrt{2}}\end{pmatrix}. \tag{1.33}$$

The fields θ_j are the predicted massless Goldstone bosons. The field h is the Higgs field. The mass of the Higgs field can be determined by plugging the expansion 1.33 into the Lagrangian 1.24 and by collecting all terms in h^2. The multiplication factor of this term is $\frac{m_h^2}{2}$. We find

$$m_h = 2\sqrt{\lambda}v \neq 0. \tag{1.34}$$

Up to now we neglected the gauge fields of the $SU(2)_L \times U(1)_Y$ symmetry in the Lagrangian. We introduce the gauge fields B_μ for the $U(1)_Y$ and W_μ^a, $a = 1, 2, 3$ for the $SU(2)_L$. The derivative in 1.24 has to be replaced by the covariant derivative:

$$D_\mu = \partial_\mu - ig_1 Y B_\mu - ig_2\frac{\sigma_a}{2}W_\mu^a. \tag{1.35}$$

Introducing these gauge fields the Lagrangian becomes locally invariant under the $SU(2)_L \times U(1)_Y$ symmetry. From the $SU(2)$ invariance we can now gauge the Goldstone bosons "away", since $\frac{i}{\sqrt{2}v}\sigma_j\theta_j(x)$ corresponds to a local $SU(2)$ transformation. Only the Higgs field h remains. This possibility of choosing the gauge such that we dispose of the massless Goldstone bosons, which are not observable, is the key point

of the Higgs mechanism.

If we put the covariant derivative 1.35 and the field $\Phi = \begin{pmatrix} 0 \\ v + \frac{h(x)}{\sqrt{2}} \end{pmatrix}$ into the Lagrangian, we find coupling terms of the Higgs field and the gauge bosons W_μ and B_μ. The interesting part of the Lagrangian reads as follows:

$$\begin{aligned} &\left|\left(-i\tfrac{g_1}{2}B_\mu - ig_2\tfrac{\sigma_a}{2}W^a_\mu\right)\Phi_0\right|^2 \\ &= (\tfrac{1}{\sqrt{2}}vg_2)^2 W^+_\mu W^{-\mu} + \tfrac{1}{4}v^2\left(W^3_\mu, B_\mu\right)\begin{pmatrix} g_2^2 & -g_1g_2 \\ -g_1g_2 & g_1^2 \end{pmatrix}\begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix}. \end{aligned} \tag{1.36}$$

Here we replaced the fields W^1_μ and W^2_μ by:

$$\begin{aligned} W^-_\mu &:= \tfrac{1}{\sqrt{2}}\left(W^1_\mu + iW^2_\mu\right) \\ W^+_\mu &:= \tfrac{1}{\sqrt{2}}\left(W^1_\mu - iW^2_\mu\right). \end{aligned} \tag{1.37}$$

The fields $W^\pm_\mu$ are eigenvectors of I_3 with the eigenvalues ± 1. They do not carry hypercharge Y. Since $Q = I_3 + \frac{Y}{2}$ the electric charge of these physically observable W bosons is ± 1.

The term 1.36 is exactly the mass term for the gauge bosons of the weak interaction, but now we kept the local gauge invariance of our Lagrangian. This procedure is called the Higgs mechanism. From 1.36 we find the mass of the W bosons to be

$$m_W = \frac{g_2 v}{\sqrt{2}}. \tag{1.38}$$

We also see from equation 1.36, that B_μ and W^3_μ are no mass eigenstates. To find observable particles we have to diagonalize the matrix in equation 1.36. The fields which satisfy this requirement are:

$$\begin{aligned} A_\mu &:= \tfrac{g_1 W^3_\mu + g_2 B_\mu}{\sqrt{g_1^1 + g_2^2}} \\ Z_\mu &:= \tfrac{g_2 W^3_\mu - g_1 B_\mu}{\sqrt{g_1^1 + g_2^2}}. \end{aligned} \tag{1.39}$$

The remaining mass term of A_μ and Z_μ in the Lagrangian is

$$\frac{1}{2}m_A^2 A_\mu A^\mu + \frac{1}{2}m_Z^2 Z_\mu Z^\mu \qquad \text{with } m_A = 0 \quad \text{and } m_Z = \frac{v}{\sqrt{2}}\sqrt{g_1^2 + g_2^2}. \tag{1.40}$$

The field A^μ describes the massless photon, while the field Z^μ is the observed uncharged Z boson.

An alternative way to write the transformation 1.39 is a $U(1)$ rotation by introducing the weak mixing or Weinberg angle θ_W. Then the transformation for the gauge bosons becomes

$$\begin{aligned} A_\mu &= \cos\theta_W B_\mu + \sin\theta_W W^3_\mu \\ Z_\mu &= -\sin\theta_W B_\mu + \cos\theta_W W^3_\mu. \end{aligned} \tag{1.41}$$

Comparison of equations 1.39, 1.40 and 1.41 leads to the following relations:

$$\tan\theta_W = \frac{g_1}{g_2} \tag{1.42}$$

$$\cos\theta_W = \frac{m_W}{m_Z}. \tag{1.43}$$

With θ_W we can also find a relation between the couplings of the $SU(2)_L \times U(1)_Y$ and coupling of the $U(1)_{e.m.}$, called e:

$$e = g_1 \cos\theta_W = g_2 \sin\theta_W. \tag{1.44}$$

The masses of the W and Z bosons as well as the electromagnetic coupling e have been measured [8].

$$m_W = 80.4\,\text{GeV} \qquad m_Z = 91.19\,\text{GeV} \tag{1.45}$$

From these measured quantities it is possible to calculate the weak mixing angle θ_W and the vacuum expectation value v of the Higgs field.

$$v = 174\,\text{GeV} \qquad \sin^2\theta_W = 0.23 \tag{1.46}$$

Unfortunately it is not possible to predict the mass of the Higgs boson, since in equation 1.34 the parameter λ goes into the relation of v and m_h. λ is a free parameter of the theory, which is not predicted by the Standard Model.

1.5 Yukawa Sector: Giving Masses to Fermions

Observation of weak interactions lead to the conclusion, that only left-handed fermions, i.e. states with negative eigenvalues of the helicity operator, which projects the direction of the spin to the direction of the momentum, interact weakly. Therefore, we can collect left-handed states into $SU(2)_L$ doublets, whereas the right-handed particles remain in $SU(2)_L$ singlets. The leptons of the Standard Model can be summarized in the following form ($j = e, \mu, \tau$ indicates the three lepton families):

$$\begin{aligned} L_j &= \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau_L \end{pmatrix} \\ l_j &= e_R, \mu_R, \tau_R. \end{aligned} \tag{1.47}$$

The Standard Model only predicts massless left-handed neutrinos. Since the neutrino has been measured to be not massless, a neutrino state has a very small probability to be found in a right-handed state, because for a massive particle propagating with a velocity less than the speed of light, we can always find a coordinate system wherein the neutrino becomes right-handed. The origin of neutrino masses has not finally become clear.

An ordinary mass term for the leptons would be of the form $-m_j L_j l_j$, but it turns out, that this term is also not invariant under a local $SU(2)$ transformation, since L_j transforms as a doublet, whereas l_j transforms as a singlet. A solution for this issue can be constructed by introducing a coupling of the Higgs field Φ to the fermion fields. Instead of an ordinary mass term we can add the following term to the Lagrangian:

$$\mathcal{L}_Y^l = -\sum_{j=e,\mu,\tau} y_j \bar{L}_j \Phi l_j + h.c. \tag{1.48}$$

Here y_j are the Yukawa couplings, which are arbitrary chosen constants. A more arbitrary way to write the Yukawa term is replacing the constants y_j by a matrix Y_{jk}, which mixes the three lepton families. However, it can be shown, that we can always find a diagonalized term like in 1.48, since the leptonic sector of the Standard Model has a global $U(3) \times U(3)$ flavor symmetry.

If we insert the field $\Phi = \begin{pmatrix} 0 \\ v + \frac{h(x)}{\sqrt{2}} \end{pmatrix}$ into equation 1.48, we find the mass term for the charged leptons to be a part of $\mathcal{L}_Y^l$:

$$-y_e v \bar{e}_L e_R - y_\mu v \bar{\mu}_L \mu_R - y_\tau v \bar{\tau}_L \tau_R + h.c. \tag{1.49}$$

The terms including the neutrinos ν_j do no longer appear in this equation. The masses for the fermions are now given by $m_j = y_j v$. So since the Yukawa coupling is a free parameter, the Yukawa mechanism gives no prediction of the masses of the leptons.

The other terms, which remain in the Lagrangian 1.48, correspond to couplings between the real Higgs field h and the lepton fields:

$$-\frac{m_e}{\sqrt{2}v} \bar{e}_L e_R h - \frac{m_\mu}{\sqrt{2}v} \bar{\mu}_L \mu_R h - \frac{m_\tau}{\sqrt{2}v} \bar{\tau}_L \tau_R h + h.c. \tag{1.50}$$

Since $v = 174$ GeV, the couplings of the leptons to the Higgs field are very small and proportional to the lepton masses. Only the $h\tau\tau$ and the $h\mu\mu$ couplings might be observable at future experiments like the LHC.

In a similar way we can also find mass terms for the quark sector by coupling the quark fields to the field Φ. The left- and right-handed quark fields can be arranged similar to the leptons ($j = 1, 2, 3$ indicates the three quark generations):

$$\begin{aligned} Q_j &= \begin{pmatrix} u_L \\ d'_L \end{pmatrix}, \begin{pmatrix} c_L \\ s'_L \end{pmatrix}, \begin{pmatrix} t_L \\ b'_L \end{pmatrix} \\ u_{Rj} &= u_R, c_R, t_R \\ d'_{Rj} &= d'_R, s'_R, b'_R. \end{aligned} \tag{1.51}$$

As it will turn out later, by convention the weak eigenstates of the down-type quarks are not identical to their mass eigenstates. To distinguish the mass and the weak eigenstates, the eigenstates of the weak $SU(2)_L$ are denoted as d', and the mass eigenstates as d. The mass term and the coupling to the Higgs field h for the down-type quarks is generated via:

$$\mathcal{L}_Y^d = -\sum_{j,k=1}^{3} \bar{Q}_j \Phi Y_{jk}^d d'_{Rk} + h.c. \tag{1.52}$$

With the choice of the field $\Phi = \begin{pmatrix} 0 \\ v + \frac{h(x)}{\sqrt{2}} \end{pmatrix}$ we find mass terms with the mass matrix for the down-type quarks:

$$M_{jk}^d := v Y_{jk}^d \tag{1.53}$$

If we want to create masses for the up-type quarks, we have to introduce a coupling to the conjugated field Φ^c:

$$\Phi^c := -i\sigma_2\Phi^* = \begin{pmatrix} v + \frac{h(x)}{\sqrt{2}} \\ 0 \end{pmatrix}. \tag{1.54}$$

Φ^c transforms identically to Φ under $SU(2)$, so we can write the Yukawa coupling for the up-type quarks as:

$$\mathcal{L}_Y^u = -\sum_{j,k=1}^{3} \bar{Q}_j \Phi^c Y_{jk}^u u_{Rk} + h.c. \tag{1.55}$$

Again we find the mass terms for the up-type quarks with the mass matrix:

$$M_{jk}^u := vY_{jk}^u \tag{1.56}$$

Unfortunately it turns out that M^u and M^d cannot be diagonalized simultaneously. That means, that in general the mass eigenstates cannot be chosen identically to the eigenstates of the $SU(2)_L$. Without loss of generality we can diagonalize the up-type mass matrix, $M^u = \mathrm{diag}(m_u, m_c, m_t)$. To diagonalize M^d at the same time, the down-type quarks have to be transformed. This is possible by introducing the unitary 3×3 Cabbibo-Kobajashi-Maskawa matrix V^{CKM}:

$$d'_{Lj} \to d_{Lj} := \sum_{k=1}^{3} V_{jk}^{\mathrm{CKM}} d'_{Lk}. \tag{1.57}$$

The CKM matrix is constructed in such a way, that in the basis of the transformed down-type states the mass matrix becomes diagonal. Now the Yukawa Lagrangian of the quark sector becomes:

$$\mathcal{L}_Y^q = \mathcal{L}_Y^u + \mathcal{L}_Y^d = -\sum_{j=1}^{3} m_{d,j}\bar{d}_{Lj}d_{Rj}\left(1 + \frac{h}{\sqrt{2}v}\right) - \sum_{j=1}^{3} m_{u,j}\bar{u}_{Lj}u_{Rj}\left(1 + \frac{h}{\sqrt{2}v}\right) + h.c. \tag{1.58}$$

Analogously to the leptonic sector we find the couplings of the quarks to the Higgs field h to be proportional to the quark masses. Here the most important couplings of the Higgs boson are the couplings to t and b quarks, which carry the highest masses.
Again the Yukawa couplings Y_{jk}^u and Y_{jk}^d as well as the components of the CKM matrix are free parameters of the theory. Therefore, there is again no prediction of the fermion masses possible in the Standard Model.

1.6 Constraints to the Higgs Mass

As mentioned above, the mass of the Higgs boson is a free parameter within the Standard Model. Nevertheless it is possible to predict the mass range, where the

Higgs boson should appear. From the direct search for the Higgs boson at the Large Electron Positron Collider (LEP) a lower bound to the mass is given [9]:

$$m_h > 114.4\,\text{GeV at } 95\% \text{ C.L.} \tag{1.59}$$

From perturbation theory there exists an upper limit to the Higgs mass. If the parameter λ becomes greater than 10, some perturbative expansions would no longer converge. From this m_h should be less than about 1 TeV. Another constraint comes from the precise measurement of the masses of top quark and the weak gauge bosons. The top quark as well as the Higgs boson contribute by loop corrections to the propagators (and therefore their masses) of the W and Z bosons. We define the parameter Δr to collect effects of higher horder perturbation theory:

$$1 + \Delta r := \frac{m_W^2}{m_Z^2}(1 - \sin^2\theta_W). \tag{1.60}$$

It can be shown, that Δr is given by [10], [13]:

$$\Delta r = \frac{3G_F}{8\sqrt{2}\pi^2}m_t^2 + \frac{\sqrt{2}G_F}{16\pi^2}m_t^2\left[\frac{11}{3}\ln\left(\frac{m_h^2}{m_W^2}\right) + \ldots\right] + \ldots, \tag{1.61}$$

where $G_F = \frac{1}{2\sqrt{2}v^2}$ is the Fermi constant. In equation 1.61 we find a logarithmic dependence of m_h to Δr. The quantities m_t, m_W, and m_Z and therefore Δr have been measured very acurately at the Tevatron and LEP. This allows to apply a fit of the Higgs mass to these electroweak mesurements [9]. The result of this fit is $m_h = 76^{+33}_{-24}$ GeV and so an upper limit to the Higgs boson mass is

$$m_h < 144\,\text{GeV at } 95\% \text{ C.L.} \tag{1.62}$$

In figure 1.2 the $\Delta\chi^2$ distribution of the fit is shown. The yellow area assigns the mass region excluded by the LEP direct search. The black line shows the fit result without including the theoretical uncertanties, which are marked by the blue band. The minimum of this curve indicates the best fit value, the points on the curve with $\Delta\chi^2 = 1$ mark the one sigma errors of the fitted mean.

1.7 Production and Decay Modes at the Tevatron

At the Tevatron protons and antiprotons are collided at a center of mass energy of 1.96 TeV. The proton consists of quarks and gluons, so these are the particles which can contribute to the Higgs production. The most important production modes are the gluon-gluon fusion and the associated Higgs production in conjunction with a W or Z boson.

The gluon-gluon fusion (see figure 1.3) is the production process with the largest cross section (figure 1.5, [10], [12]). The gluons couple via a fermion loop to the Higgs boson. In the loop there are contributions from all fermions, but the most relevant contributions come from t and b quarks, since the coupling of the Higgs boson is proportional to the fermion masses (see equation 1.58). The predicted cross section

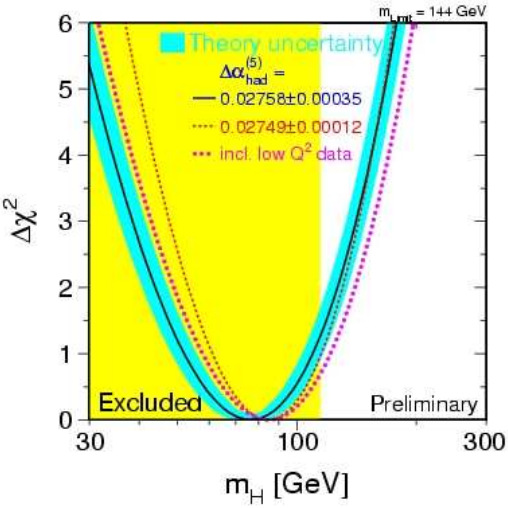

Figure 1.2: Fit of the Higgs mass to the electroweak data

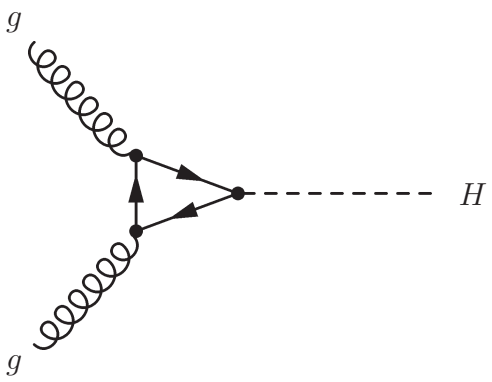

Figure 1.3: gluon-gluon fusion

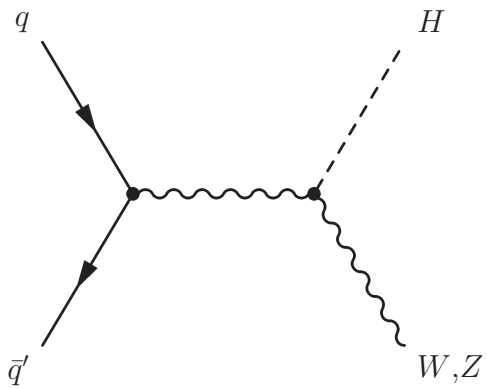

Figure 1.4: WH,ZH

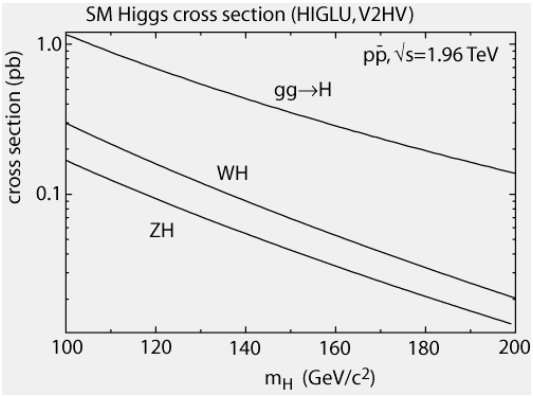

Figure 1.5: cross sections of the Higgs production at Tevatron

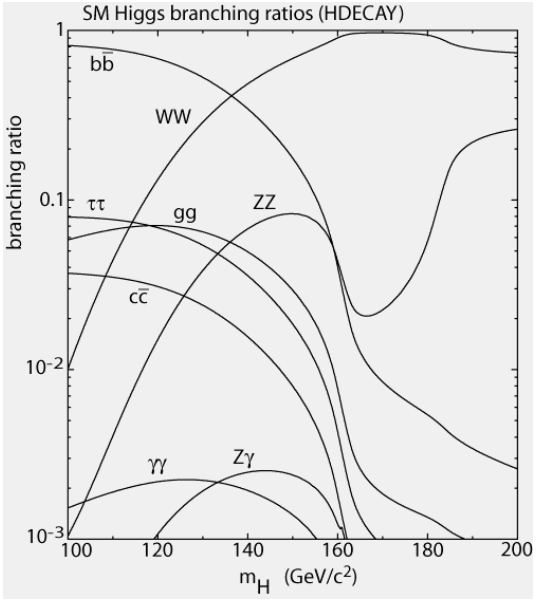

Figure 1.6: branching ratios of the Higgs boson decay

of the gluon-gluon fusion ranges from 1.0 pb to 0.1 pb for a Higgs mass between 100 GeV and 200 GeV.
The associated Higgs production (see figure 1.4) has a cross section between 0.3 pb and 0.03 pb for the WH process in the mentioned mass range. The cross section for ZH is about two times smaller, since there are two W bosons (W^+ and W^-) and only one Z boson contributing to the respective process. In these channels two quarks from the proton and antiproton form a weak gauge boson, which radiates a Higgs boson.
The decay modes of the Higgs boson and their branching ratios are shown in figure 1.6 [10]. For a $m_h < 140$ GeV the most important decay is $H \to b\bar{b}$. The coupling to the top quark and the weak gauge bosons W and Z is much larger than the coupling to the b quark, but the decay into $t\bar{t}$ respectively WW or ZZ is highly suppressed for m_h much smaller than twice the top quark mass or twice the mass of the gauge bosons. There are only some contributions from a Higgs decaying into light fermions, especially $c\bar{c}$ and $\tau\bar{\tau}$.
Similar to the gluon-gluon fusion in the production, the Higgs boson may decay into gluon or photon pairs as well as into Z and a photon via a fermion or a W loop.
For higher masses, the Higgs boson can decay into a pair of weak gauge bosons. Especially if $m_h > 2m_{W,Z}$, this decay becomes dominant. If $m_h < 2m_{W,Z}$, at least one W or Z boson has to be produced off-shell, resulting in a strong decrease of the decay into W and Z pairs for a lower Higgs mass.
The process discussed in this analysis is $WH \to l\nu b\bar{b}$. In this channel the Higgs boson is produced via associated production with a W boson. The Higgs is assumed to decay into a b quark pair, the W boson into a charged lepton and a neutrino.

This process is the most promising channel at the Tevatron, since the leptonically decaying W causes a specific signature in the detector, which allows to trigger the candidate events. The cross section of $gg \to H \to b\bar{b}$ is a few times larger than the one of $WH \to l\nu b\bar{b}$, but for this process we expect a huge backgound originating from QCD $b\bar{b}$ production.

Chapter 2

The Experimental Site

The experiment delivering the data for our analysis is the CDF II detector at the Fermi National Accelerator Laboratory (FNAL or Fermilab). CDF II and DØ are the two detection facilities analysing the collisions of protons and antiprotons, each accelerated by the Tevatron to 980 GeV, so the center-of-mass energy $\sqrt{s}$ in the collisions is 1.96 TeV.
Fermilab is located approximately 60 km west of Chicago in Batavia, Illinois (USA). Until the start of the Large Hadron Collider (LHC), planed for 2008, the Tevatron will be the accelerator with the highest center-of-mass energy. The Tevatron started to operate in 1983, the first proton antiproton collisions were observed in 1985 at CDF. One of the main discoveries at the Tevatron was the observation of the top-quark in 1995 at CDF and DØ . From 1996 to 2001 the accelerator system and the detectors were upgraded to increase the center-of-mass energy from 1.8 TeV to 1.96 TeV and to accomplish a higher brunch crossing rate. The operation time of the Tevatron since 2001 is called Run II. For this analysis the data accumulated between April 2002 and summer 2007 is used.

2.1 The Accelerator System

To accelerate protons and antiprotons to the final energy of 980 GeV a system of several pre-accelerators is used (see figure 2.2).
The first part of the accelerator chain is a Cockcroft-Walton pre-accelerator, wherein negative hydrogen ions (H^-) produced in a magnetron are brought to an energy of 750 keV using electrostatic lenses for the acceleration. These ions are injected into a linear accelerator (LINAC). The LINAC raises the energy of the hydrogen ions to 400 MeV in a series of radio frequency cavity resonators. Before entering the Booster, the ions are sent through a carbon foil to strip the electrons. The Booster is a synchrotron with a radius of 75 m. It started to operate in 1970 with initially 16 radio frequency cavities, which were increased to 19 cavities in 2004. The advantage of a circular accelerator over a LINAC is the possibility to accelerate particles with the same cavity several times, but one needs deflecting magnets to keep the beam on a circular trajectory. The Booster accelerates the protons to 8 GeV.
The protons from the Booster are sent to the Main Injector, another synchrotron

Figure 2.1: The area of Fermilab. The ring on the left is the Main Injector, the ring on the right is the Tevatron with a circumference of about 6 km.

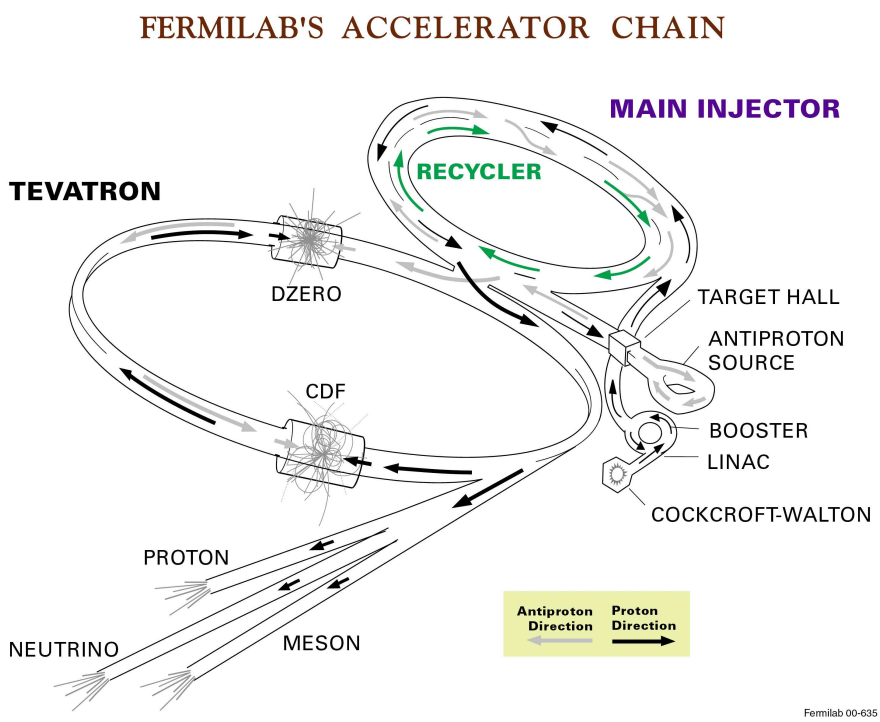

Figure 2.2: Schematic view of the accelerator chain.

with a circumference of about 1 km. The Main Injector can accelerate protons coming from the Booster as well as antiprotons from the antiproton source. It has two operation modes: the acceleration of protons and antiprotons to 150 GeV for the injection into the Tevatron and the acceleration of protons to 120 GeV for antiproton production. The protons used for the antiproton production are sent to a nickel target. In proton-nucleus interactions many secondary particles including antiprotons are produced. The produced particles are focussed by a lithium lens and magnets select particles with negative charge and an energy around 8 GeV. The only stable particles selected by this procedure are antiprotons. The antiprotons are stored in the Accumulator, a triangle-shaped storage ring. As soon as enough antiprotons have been produced, they are injected into the Recycler. The Recycler is a storage ring built in the same tunnel as the Main Injector. The Recycler was originally built to recover antiprotons coming back from the Tevatron ring to store them for further runs, but due to technical issues it does not operate in this function. The Recycler is now only used to store antiprotons coming from the Accumulator to increase the total number of available antiprotons. Another task of the Recycler is the electron cooling of the antiprotons. By inserting an electron beam with an energy of 4.34 MeV parallel to the antiproton beam on one of the straight sections of the Recycler, it is possible to reduce the divergence and energy spread of the antiproton beam. The interactions of electrons and antiprotons lead to a thermal equilibrium. If the injected electrons are nearly mono-energetic (low temperature), the temperature of the electron beam increases by Coulomb scattering between electrons and antiprotons, whereas the temperature of the antiprotons is lowered. This means, that the antiproton beam is focussed in the energy spread. The electron cooling system operates since 2005 and is essential to increases the number of stored particles and to lower the emittance in the Tevatron which directly correspond to the taken data in the detectors (see figure 2.3).
From the Recycler the antiprotons are extracted into the Main Injector, where they as well as the protons are accelerated to 150 GeV. The last component of the accelerator chain is the Tevatron. Here the protons and antiprotons coming from the Main Injector are then accelerated to the final energy of 980 GeV. The Tevatron is a synchrotron with superconducting magnets and a radius of 1 km. The protons and antiprotons running in opposite directions are brought to collisions at exactly two points, called B0 and D0, where the two detectors CDF and DØ are located. The particles are usually stored for 16 hours in the Tevatron. This period of collisions and data taking is called store. It takes approximately 10 to 16 hours to produce enough antiprotons to fill the Tevatron again and to run the next store. Since the acceleration by radio frequency cavities allows no continuous beam, the protons and antiprotons are accumulated in bunches.
An import quantity of a collider is the luminosity. The luminosity is the ratio of the event rate and the cross section of a particular physical process.

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

The cross section σ depends only on the energy of the colliding particles. All other collider effects are combined in the luminosity:

$$\mathcal{L} := \frac{N_p N_{\bar{p}} B f}{4\pi\sigma_x\sigma_y}. \tag{2.2}$$

Here N_p and $N_{\bar{p}}$ are the numbers of protons and antiprotons in a bunch respectively, B is the number of bunches, σ_x and σ_y denote the Gaussian width of the beam profile, f is the revolution frequency. For the Tevatron the number of bunches is typically 36 for each type of accelerated particle, the revolution frequency is about 48 kHz. Another useful quantity is the integrated luminosity. It is defined as:

$$\mathcal{L}_{\text{int}} := \int \mathcal{L} dt. \tag{2.3}$$

To obtain the overall number of events N of a certain physical process with cross section σ, one has to multiply the integrated luminosity with the cross section:

$$N = \sigma\mathcal{L}_{\text{int}}. \tag{2.4}$$

The luminosity is usually measured in $\text{cm}^{-2}\text{s}^{-1}$, the integrated luminosity in inverse picobarn (pb^{-1}). The unit barn is defined as 1b:= 10^{-28}m^{-2}. For the Tevatron the planed peak luminosity lies in the range 1.6 to $2.7 \cdot 10^{32}$ $\text{cm}^{-2}\text{s}^{-1}$, which has been reached. For this analysis data corresponding to an integrated luminosity of 1.9 pb^{-1} are used, which has been delivered between 2001 and 2007 (see figure 2.3).

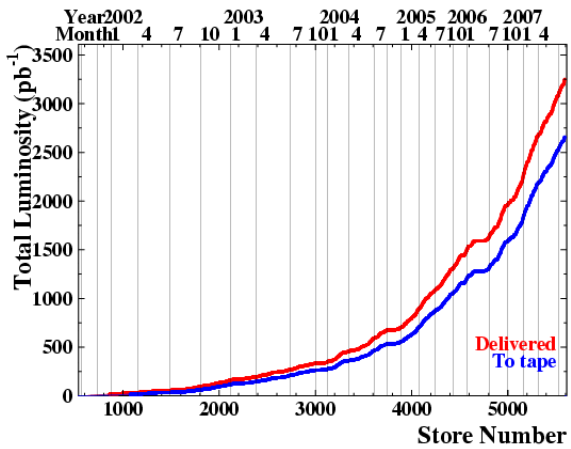

Figure 2.3: The integrated luminosity delivered by the Tevatron (upper curve) and recorded by CDF (lower curve).

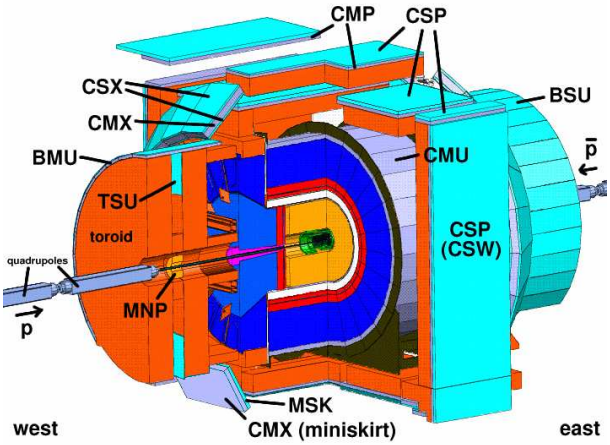

Figure 2.4: Schematic view of the CDF II detector

2.2 The CDF II Experiment

The Collider Detector at Fermilab (CDF, called CDF II for the Run II period) is a multi-purpose solenoid particle detector situated at the collision point B0 of the Tevatron. CDF is built azimuthally and forward-backward symmetrically and combines the measurement of charged particle tracks, energy deposition in the calorimeters as well as the identification of muons to deliver as much information as possible to analyze proton antiproton collisions. Figure 2.4 shows a cut-away view of the detector with its different components. The detector consists of the tracking system, contained in a superconducting solenoid, calorimetry and muon systems, which are localized outside the solenoid. The solenoid is 1.5 m in radius and 4.8 m in length and generates a 1.4 T magnetic field parallel to the beam axis. A more precise description of the CDF II can be found in [15]. Between the COT and the solenoid a Time-Of-Flight (TOF, [16]) system consisting of scintillator panels is placed. It allows the identification of kaons and pions by their flight-time difference, which is particulary essential to identify the decay products of B mesons.
The coordinate system used to describe events in the detector is shown in figure 2.5. The positive z-direction is defined as the flight direction of the protons. The azimuthal angle ϕ is measured from the Tevatron plane, the polar angle θ is measured with respect to the z-axis. For many applications it is useful to use the pseudo-rapidity η instead of the polar angle:

$$\eta := -\ln \tan \frac{\theta}{2}. \tag{2.5}$$

The pseudo-rapidity is equal to the real rapidity for massless particles. Rapidity differences are Lorentz-invariant quantities. Also for massive particles measureable in the detector, like electrons, muons, pions, and kaons, η approximates the rapidity quite well.

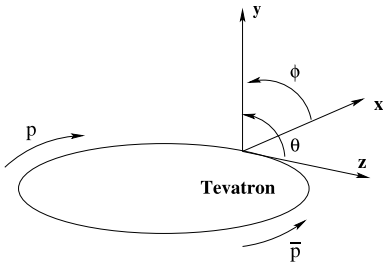

Figure 2.5: The coordinate system of CDF

2.2.1 Tracking System

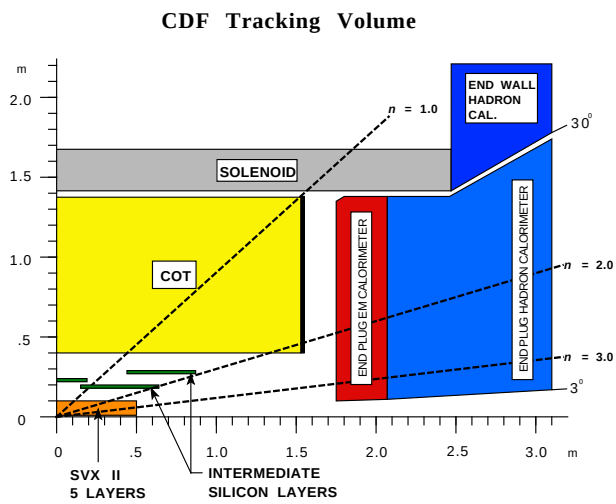

Figure 2.6: Longitudinal view of the CDF II tracking system and the plug calorimetry. Not shown is the Time-Of-Flight system located between the COT and the solenoid.

The tracking system is used to reconstruct tracks of charged particles. These measurements are not only important for the determination of interaction vertices, but they are also suitable to determine the charge and the momentum of particles, since the tracking system is placed in a magnetic field, wherefore charged particles are forced on a curved track, whose curvature is reciprocally proportional to the particle momentum.
The tracking system of CDF II consists of four parts, shown in figure 2.6. The three inner subsystems are the silicon microstrip detectors Layer00 [17], the Silicon

Vertex Detector (SVX II, [18]) and the Intermediate Silicon Layers (ISL, [19]). In total these three detectors consist of eight silicon layers in a barrel geometry ranging from a radius of $r = 1.5$ cm to $r = 28$ cm. They are surrounded by the Central Outer Tracker (COT, [20]), a gas drift chamber.
The closest silicon layer to the beam pipe is the Layer00. It is a radiation-hard, single sided microstrip detector, whereas the layers of the remaining silicon detectors are double sided sensors. The next five layers ranging from $r = 2.4$ cm to $r = 10.7$ cm form the SVX II. Three of the five layers combine an r-ϕ measurement on one side of the layers with an r-z measurement on the other side. Two layers of the SVX II supply an r-ϕ measurement with a stereo angle of 1.2°. The stereo measurements allow a precise three dimensional track resolution. The ISL consist of two layers for the plug region ($1.0 \leq |\eta| \leq 2.0$) placed at $r = 20$ cm and $r = 28$ cm and one layer in the central region ($|\eta| \leq 1.0$).
The silicon system is surrounded by the COT. The COT is a cylindrical open-cell drift chamber with an active region length of 3.1 m. The inner radius is 43.4 cm, the outer one 132.2 cm. It provides a coverage for $|\eta| \leq 1.0$. The COT contains 30,240 sense wires running between two end plates. Approximately half of the wires run parallel to the z-axis and half are small angle (2°) stereo, which provides additional z information. The wires are arranged in eight concentric superlayers each consisting out of twelve sublayers. The drift chamber is filled with a mixture of about 50% argon and 50% ethane. The filled gases allow a high drift velocity of ionised particles with a maximum drift time of 100 ns, which has to be less than the bunch crossing time of 132 ns (which would be reached for 108x108 bunches, the Tevatron currently operates only with 36x36 bunches). The COT allows the measurement of particle transverse momenta with a precision of $\sigma(p_T)/p_T^2 = 0.0015$ (c/GeV). The combined use of ISL, SVX II, and COT provides a precision of $\sigma(p_T)/p_T^2 = 0.0007$ (c/GeV).

2.2.2 Calorimetry

The calorimeters are detector components specified to measure the energy of particles. Especially the energy of uncharged particles like photons or neutral hadrons has to be measured solely with the calorimeters, since neutral particles produce no signals in the tracking system.
The calorimetry of CDF II is placed around the tracking system and the solenoid. Overall there are five calorimeters: the Central ElectroMagnetic calorimeter (CEM, [21]), the Central HAdron calorimeter (CHA, [22]), the end-Wall HAdron (WHA) calorimeter, the end-Plug ElectroMagnetic (PEM) and the end-Plug HAdron calorimeters (PHA, [23]). The calorimetry covers the full azimuthal range as well as a pseudo-rapidity range of $|\eta| \leq 3.64$. The two central calorimeters cover a pseudo-rapidity range of $|\eta| \leq 1.1$.
The CEM is localized directly outside the solenoid. The active scintillator material is polystyrene. There are 23 layers of scintillator with a thickness of 4.5 mm interspersed with 4 mm thick lead sheets as absorption material. For the read-out of the scintillator signals multi-anode photomultipliers are used. The elements of the CEM calorimeter are arranged pointing towards the interaction region, which can

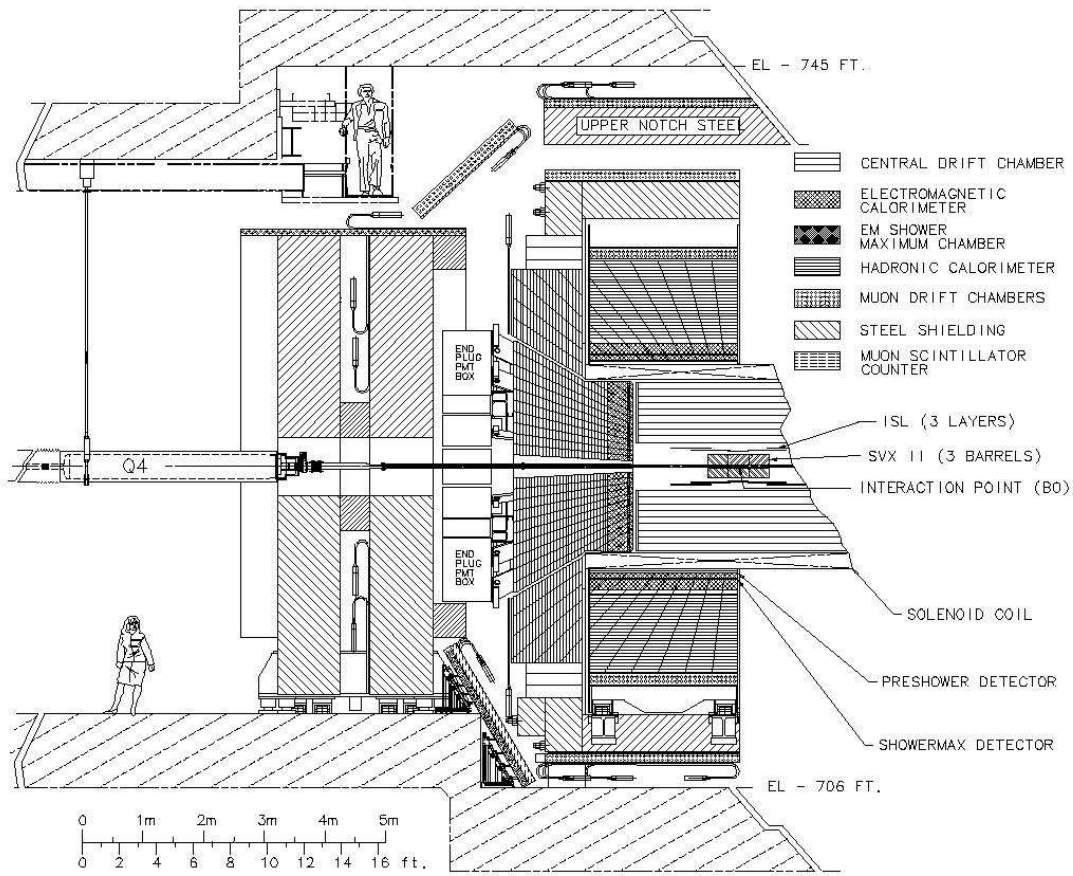

Figure 2.7: Elevation view of one half of the CDF II detector

be seen in figure 2.7.
The CHA calorimeter is built out of 23 iron and scintillator layers. The 5 cm thick iron absorbers are interspaced with 6 mm scintillators. The calorimeters in the plug region are constructed in a similar way. The electromagnetic and hadronic calorimeters have a uniform pattern of matched projective towers. The towers cover regions of 7.5° to 15° in ϕ and between 0.1 (central) and 0.6 (outer) in η. The thickness of the CEM correponds to $19X_0$, the thickness of the PEM to $21X_0$, where X_0 is the radiation length of an electron. For the central hadronic calorimeter the thickness is equivalent to 4.5λ and 7λ for the plug calorimeter. Here λ is the hadronic decay length. The energy resolution of the calorimeters is $13.5\%/\sqrt{E/\mathrm{GeV}}\oplus 2\%$ for the CEM and $14.4\%/\sqrt{E/\mathrm{GeV}}\oplus 2\%$ for the PEM calorimeters. The hadronic calorimeters have an energy resolution of about $75\%/\sqrt{E/\mathrm{GeV}}\oplus 3\%$.

2.2.3 Muon System

Nearly all particles except muons and neutrinos are absorbed by the calorimeters, the magnet return yoke, or the additional steel walls surrounding the detector. The remaining muons are detected by the muon system [24], which is located beyond the calorimeters. The muon system consists of four subsystems covering a pseudo-rapidity range of $|\eta| \leq 2.0$. The arrangement of the muon chambers is illustrated in figure 2.4.
The central region ($|\eta| \leq 0.6$) of the detector is covered by the Central MUon detection (CMU) and the Central Muon uPgrade (CMP). The CMU consists of planar drift chambers arranged in four layers and can detect muons with a transverse momentum larger than 1.4 GeV/c. Behind the 60 cm steel wall of the return yoke the four drift chambers of CMP cover the same region in η. It forms a box enclosing the central detector and detects muons with $p_T \geq 2.0$ GeV/c. The Central Muon eXtension (CMX) covers the range from $|\eta| \geq 0.6$ to $|\eta| \leq 1.0$. CMX consists of conical sections of drift tubes and scintillation counters. The last part of the muon system is the Intermediate MUon detection (IMU) system. It covers a region of $1.0 \leq |\eta| \leq 1.5$ and has replaced the forward muon system used in Run I. For Run II also several new chambers have been added to the existing muon system to increase the coverage in the ϕ-η plane.

2.3 Trigger System

The Tevatron has been designed to operate with up to 108 proton and 108 antiproton bunches, which corresponds to a crossing rate of 7.6 MHz, respectively one collision every 132 ns. For the operation of the Tevatron with 36x36 bunches the crossing frequency is only 2.5 MHz, but also this smaller rate leads to huge amount of information produced by the detector. The 36 bunches are arranged in three trains with twelve bunches each. During the train pass the crossing rate is 2.5 MHz, followed by a gap of 2.6 μs, so the averaged crossing rate is reduced to 1.7 MHz.
The crossing rate of 1.7 MHz is still to large to be written to tape, which can only be done with a rate of approximately 150 Hz in maximum. Therefore a decision system (trigger) is required to reduce the number of selected events by rejecting events of less interest. The CDF II trigger system [25] consists of three levels, schematically shown in figure 2.8.
The Level-1 and Level-2 triggers are hardware based triggers, whereas the Level-3 trigger is a software trigger running on a PC farm. On the first trigger level preliminary information of a subset of the detector is used to reduce the data flow. There are three systems running parallel on the Level-1 trigger stage: the calorimeter trigger to find energy depositions in the calorimeters, the muon triggers and the tracking system, which provides a fast track reconstruction in the COT and compares the found tracks with matchings in the calorimeters or the muon chambers. All three informations are combined to decide whether an event is rejected or passed to the Level-2 trigger. The information of up to 42 collisions is stored in a storage pipeline. In this pipeline crossing events are buffered until the trigger makes a decision. The

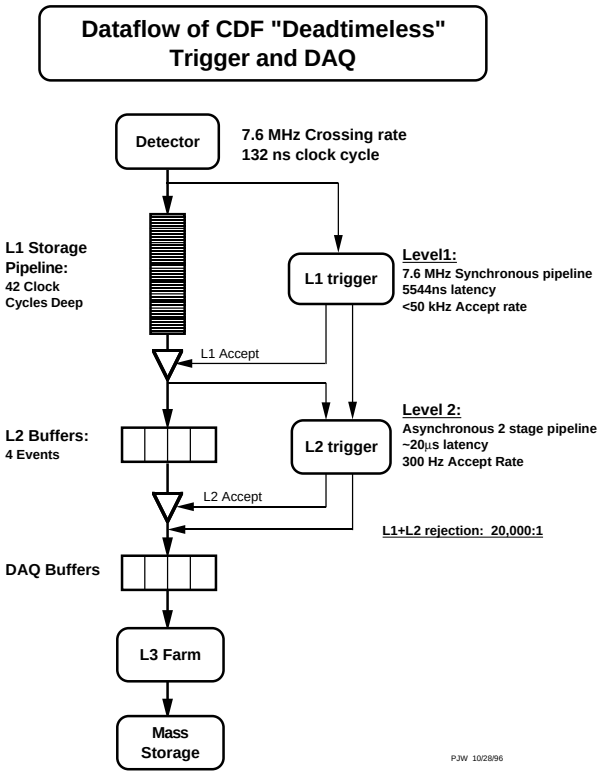

Figure 2.8: Schematic view of the CDF II trigger system. The trigger system has originally been designed to handle a 7.6 MHZ crossing rate (currently the crossing rate is reduced to 1.7 MHz).

maximum accept rate after the Level-1 trigger is of the order of 50 kHz. The buffer pipeline for the Level-2 trigger allows to store four events.
The Level-2 trigger accomplishes a limited event reconstruction by combining electron and photon identification in the shower maximum strip chamber, which is a part of the CEM, as well as tracking information from the SVX II. The Silicon Vertex Tracker (SVT, [27]) provides the determination of tracks with a large impact parameter. The Level-2 trigger also allows a first jet reconstruction by cluster finding. The event rate is reduced to about 800 Hz to 500 Hz.
The last trigger stage is the Level-3 trigger, which provides a full read out of the event. The pass decision is made out of the particle content and the event topology. About 150 events per second passing the last trigger are written to permanent storage. There exist different so-called trigger paths, which combine different selections in the three level stages.

Chapter 3

Background and Signal Modeling

It is impossible to create a pure W-Higgs sample, since there exist a couple of different physical interactions in proton antiproton collisions, which lead to the same signature in the detector as the W-Higgs signal. For this reason we have to find adequate methods to describe these backgrounds. In this chapter we will give an overview of the different backgrounds and how they are modelled.

3.1 The Backgrounds of the W-Higgs Production

The background to the W-Higgs process consists of several different processes. The dominant background is the W+heavy flavor production. This background is composed of $W + b\bar{b}$, $W + c\bar{c}$ and $W + c$ processes. The leading order Feynman diagrams are shown in 3.1 and 3.2. The produced quarks in these processes cannot be seen isolated in the detector. Due to the confinement of the strong interaction quarks and gluons produced at parton level cannot be observed as free particles, since the separation of colored objects leads to an increase of the gluon field strength between these objects. The increasing field strength allows the production of new quark-antiquark pairs in the field, which form colorless hadrons. This process is called hadronization. The created hadrons may again decay into several secondary particles, so at the end a shower of particles reaches the calorimeters. These particle showers are called jets.
Another background is the production of top quarks. The top quark is mainly produced in pairs of t and $\bar{t}$ via the strong interaction, but we also expect a smaller background coming from electroweak single-top production. The dominant Feynman diagram for the $t\bar{t}$ production is shown in figure 3.3, a small contribution also stems from gluon-fusion $gg \rightarrow g \rightarrow t\bar{t}$. We distinguish between three kinds of $t\bar{t}$ processes, depending on the decay mode of the two W bosons. If both W bosons decay leptonically, the decay mode belongs to the dilepton category. This category contributes to the background of the Higgs production, if one lepton is not reconstructed. Other decay modes of the two W bosons are the hadronic decay, if both gauge bosons decay into hadrons, and the semileptonic decay, if one W goes to hadrons and one W decays into a charged lepton plus a neutrino. From the full-hadronic $t\bar{t}$ channel events with a lepton coming out of a b quark decay or a

misidentified jet in addition with loose jets may fake the signature of the W-Higgs signal.
There are also two kinds of single-top production processes, called t- and s-channel (see figures 3.4 and 3.5). The names s- and t-channel refer to the Mandelstam variables which indicate the four-momentum squared of the W boson in the leading order Feynman diagrams. A further background is the diboson production, where WW, WZ and ZZ productions are included (see figures 3.6, 3.7, 3.8). The production of a single Z boson together with heavy quark jets is also a small background of the W-Higgs process.
Events with a W or Z boson together with light flavor jets, i.e. jets which do not contain any heavy quark like c or b quarks, are called mistag events. In these events a light flavor jet is mistakenly tagged as a b jet. The Feynman diagram of the dominant W+light jets production is similar to the ones of $W + b\bar{b}$ and $W + c\bar{c}$ but with light quarks, i.e u, d, or s quarks, instead of the heavy quarks. A further background process of is the $Z \to \tau\tau$, shown in figure 3.10, where a τ might be misidentified as a b quark.
Another background process consists of events, which do not contain any electroweak gauge boson. In these events there should not appear a high-energetic charged, isolated lepton, but in some cases a hadronic jet is misidentified as a charged lepton and the mismeasurement of transverse energy may fake a neutrino signature. Also the decay products of heavy quarks in a jet can contain neutrinos and isolated charged leptons with large momenta. Events of this kind are called QCD or non-W events. One example for such a process is shown in figure 3.11, where one of the light jets might be misidentified as a lepton, the second one might get lost.
The cross sections of most of these background processes are much larger than the cross section of the Higgs process, which is of the order of 0.1 pb (see section 1.7). All background processes and their predicted cross sections are listed in table 3.1. The given cross sections correspond to the sum of the cross sections of individual subprocesses, for instance the cross section of $W + b\bar{b}$ includes the processes $W + b\bar{b} + n \cdot p$ with different numbers n of additional partons p, radiated in initial or final state. Also the three decay modes $W \to e\nu_e$, $W \to \mu\nu_\mu$ and $W \to \tau\nu_\tau$ are combined. The QCD background is not listed in this table, since it consists of many different processes with very large cross sections of the order of up to 10^6 pb, but these backgrounds contribute to the signal-like events only due to mismeasurements in the detector and not due to the same final state.

3.2 Monte Carlo Event Generators

To perform a description of signal and background, proton antiproton collisions are simulated with computer programs. Since scatterings of elementary particles are quantum mechanical processes, in which the production rate, the lifetime, the momentum, and many other quantities of a particle are stochastically distributed, the basis of every simulation program is a random number generator. Computer programs based on random number generators are called Monte Carlo programs.
There exist several Monte Carlo event generators, which can be adjusted to gener-

process	cross section σ in pb
single-top s-ch.	0.88
single-top t-ch.	1.98
$t\bar{t}$	6.7
W+light	6200
$W + b\bar{b}$	12.5
$W + c\bar{c}$	22.3
$W + c$	21.1
Z+light	551
$Z + b\bar{b}$	2.0
$Z + c\bar{c}$	4.3
WW	7.7
WZ	2.3
ZZ	2.56

Table 3.1: Cross sections of the background processes.

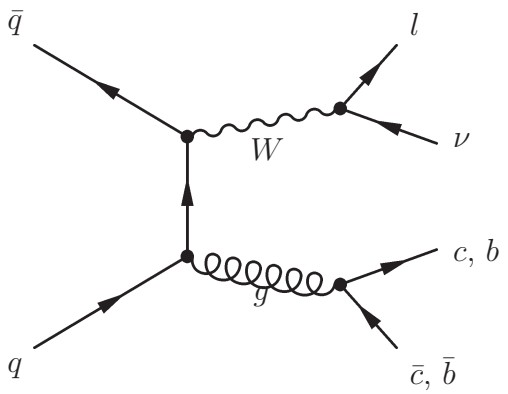

Figure 3.1: $Wc\bar{c}$, $Wb\bar{b}$

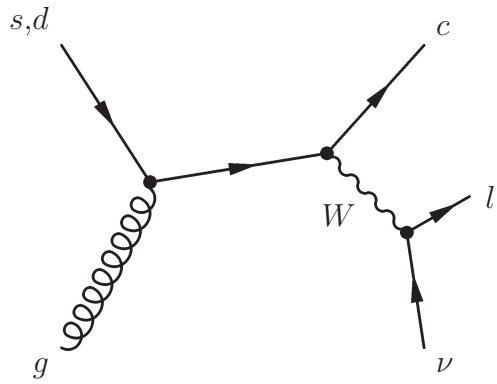

Figure 3.2: Wc

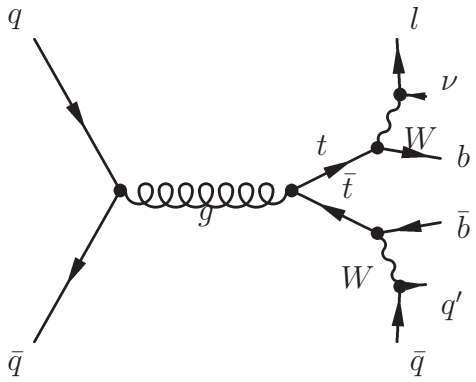

Figure 3.3: $t\bar{t}$

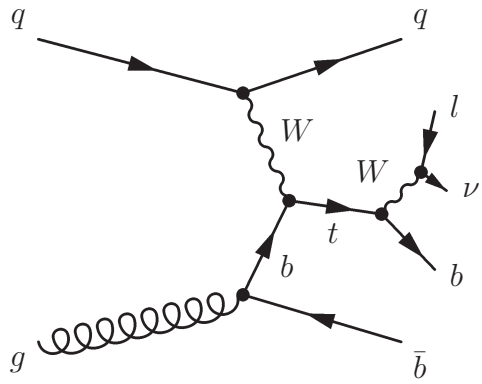

Figure 3.4: single-top t-channel

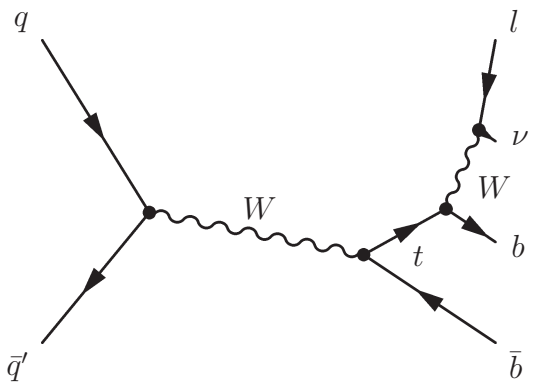

Figure 3.5: single-top s-channel

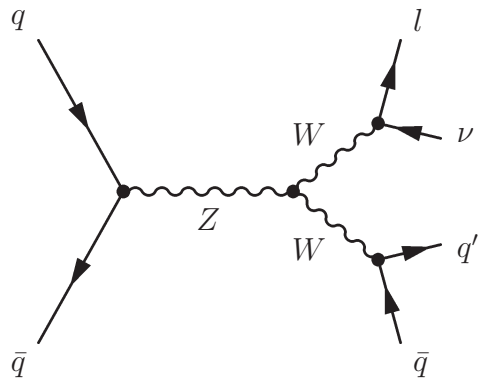

Figure 3.6: WW

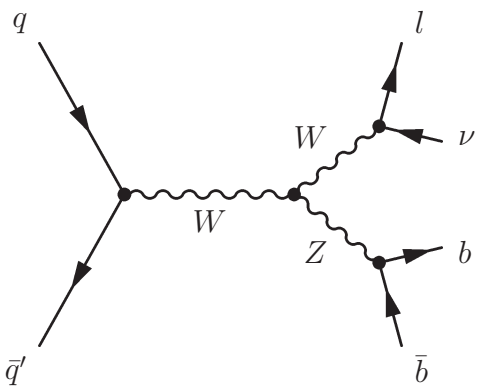

Figure 3.7: WZ

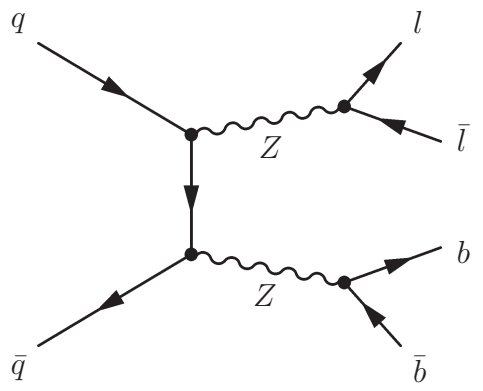

Figure 3.8: ZZ

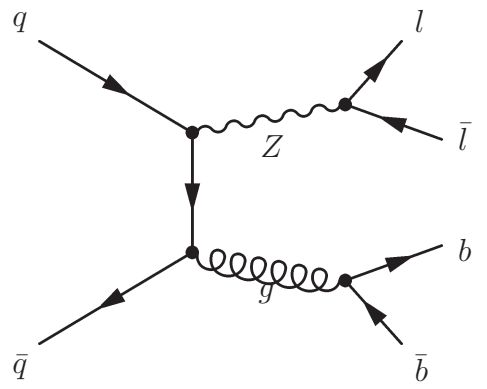

Figure 3.9: Z+jets

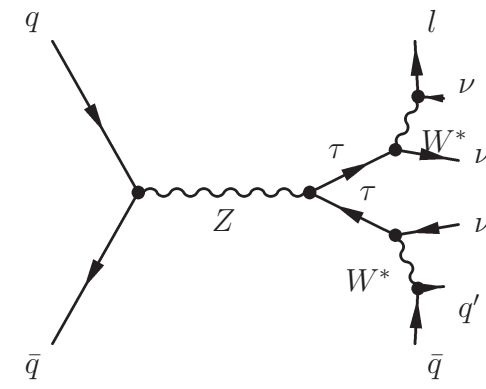

Figure 3.10: $Z \to \tau\tau$

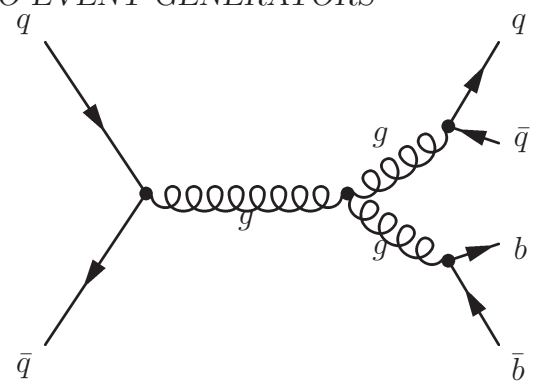

Figure 3.11: QCD background

ate events of a certain process as well as to produce events of different theoretical models. The quantities of the particles in such a generated event follow the theoretical predicted probability density of the transition matrix element of a certain elementary process. Usually in collisions of high energetic particles many interactions occur. Besides the production of primary particles in the hard scattering process, secondary particles may be radiated. Also the decay of instable particles and the hadronization of colored objects is simulated by the Monte Carlo generator. For collisions of protons and antiprotons, which are composed particles consisting of quarks and gluons, also the substructure of these particles has to be modelled. The parton distribution function $f(x, Q^2)$ (PDF) describes the probability to find a certain parton with a momentum fraction x of the full proton momentum and which transfers a momentum Q^2 to the hard scattering process.
To model the backgrounds for this analysis different Monte Carlo event generators are used and are shortly described in this section. One main difference of these generators is for instance the modeling of the hadronization. The hadronization process takes place at a low Q^2 scale of the strong interaction, where due to the large coupling constant α_s perturbation theory becomes inaccesable. Different effective theories are implemented. The different event generators describe the hadronization and formation of jets.

Alpgen
The Monte Carlo event generator Alpgen [29] has been developed to describe only hadronic collisions like protons and antiprotons at the Tevatron. It provides a leading order perturbation theory calculation of the matrix elements of QCD and electroweak hard scattering processes. The included processes in Alpgen are Standard Model processes including W and Z bosons decaying into leptons or quarks, but also pure QCD jet production is implemented. One feature of Alpgen is the handeling of the color flow and the spin correlation, which is taken into account for the calculation of the scattering amplitude up to the hadronization process. Alpgen strictly divides the calculation of the hard scattering process and the underlying event, which is calculated with a secondary program like [31]. Therefore an event, where in the hard process the number of quarks from outgoing jets has been fixed,

may also include a larger number of jets stemming from the underlying process.

Herwig
HERWIG [30] is an event generator, which includes the full simulation of several initial particle interactions like lepton-lepton, lepton-hadron or hadron-hadron collisions, with particular emphasis on the detailed simulation of QCD parton showers. It provides the initial- and final-state QCD jet evolution, includes the color coherence of partons in all subprocesses, and takes into account the azimuthal correlation within and between jets due to gluon interference and polarization. Furthermore, HERWIG uses a cluster model for jet hadronization based on non-perturbative gluon splitting. A similar model is implemented to generate soft and underlying hadronic events. Besides all Standard Model interactions HERWIG also provides the simulation of many supersymmetric processes.

Pythia
PYTHIA [31] provides a similar functionality like the HERWIG program. It can be used to simulate the high-energetic collisions of electrons, positrons, protons, and antiprotons. PYTHIA combines many theoretically based and effective models to describe parton distributions, initial- and final-state QCD radiation, fragmentation, and multiple interactions. The main difference between PYTHIA and HERWIG is the hadronization model. HERWIG uses a cluster algorithm, whereas PYTHIA uses the Lund string model. In the Lund model [32], the color field between two color-connected partons is approximated by a massless relativistic string, which can break by the creation of $q\bar{q}$ pairs to produce hadrons.

3.3 Monte Carlo and Data Based Models

To model the backgrounds of the W-Higgs process we use different Monte Carlo event generators and showering programs. The $t\bar{t}$ and the diboson processes are simulated using the PYTHIA program, which is also used to produce the signal W-Higgs Monte Carlo. For W+heavy flavor and Z+jets ALPGEN and PYTHIA are used. The single top sample is produced using MADEVENT [33] and PYTHIA for the showering. HERWIG is only used to model some samples, to determine systematic uncertainties.
In a second step all Monte Carlo based background models are submitted to the CDF detector simulation [34], which translates the simulated particles of an event into observable signals in the detector. The detector simulation turns the events produced with a Monte Carlo simulation into a comparable format to the data events. The used software package is GEANT3 [35], which allows to create a model of the whole detector and to simulate the interactions of particles and detector material. The reconstruction of the signals in the drift chamber (COT) is reproduced by a GARFIELD [36] simulation. The showering of particles inside the electromagnetic and hadronic calorimeters are simulated by GFLASH [37], which allows a parametrization of the calorimeter response, since this program runs about 100 times faster then the

usual GEANT3 program.
The only background process, which is not modelled by a Monte Carlo program, is the non-W model. The non-W events contributing to the W-Higgs data set are all events of a signature, which is untypical for most QCD events. Hence we would have to produce a huge number of Monte Carlo generated events, from which only a tiny fraction fakes the signature of the signal. Therefore, we use a non-W model based on real data events.

3.3.1 Heavy Flavor Overlap Removal

The modeling of the backgrounds with Monte Carlo event generators is based on the use of random number generators. To produce events of a certain process only the primary process is fixed. In the showering for instance it is possible to create different event signatures due to the randomly generated particles, since the event generator allows every physically allowed process to occur. This effect may lead to the appearance of heavy quarks (c or b) in a light flavor Monte Carlo sample. Since it is important for the b tagging and for the determination of efficiencies to have clearly separated light and heavy flavor samples, it is necessary to reorder the Monte Carlo events in these samples. There are two different methods to separate the light and heavy flavor events, the P_T and the jet based heavy flavor overlap removal.

P_T Based Heavy Flavor Overlap Removal
The P_T based algorithm looks for heavy quarks in the list of particles produced by the Monte Carlo generator. All events containing a heavy quark with a transverse momentum P_T greater than 8 GeV and a pseudo rapidity of $|\eta| \leq 3$ are marked. If there is no such so called tight c or b quark findable in the generator list, the algorithm searches in addition for pairs of c and $\bar{c}$ respectively b and $\bar{b}$ with $\Delta R := \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.4$. If the sum of the four-vectors of these two quarks has a transverse momentum greater than 8 GeV and $|\eta| \leq 3$, the event is also marked.
From the light flavor samples all events with such heavy quarks or quark pairs are removed. In the $Wb\bar{b}$ sample all events are kept. For the samples of the $Wc\bar{c}$ and the Wc processes all events containing a b quark or a $b\bar{b}$ pair are assigned to be events of the $Wb\bar{b}$ process. Additionally, all events with $c\bar{c}$ in the Wc samples are reordered to be events of the $Wc\bar{c}$ sample.

Jet Based Heavy Flavor Overlap Removal
For the light flavor samples the jet based algorithm keeps all events without any heavy quark in the particle list. If there are heavy quarks in the event, they are assigned to be c or b like, except if there are $c\bar{c}$ or $b\bar{b}$ pairs produced in the parton shower, which lie inside the cone of a reconstructed jet, that means that both quarks have to have $\Delta R \leq 0.4$ with respect to the jet axis. These events are not assigned to be heavy quark events. Otherwise all events from the heavy quark samples are removed, if there is a $c\bar{c}$ or $b\bar{b}$ pair produced on matrix element level and being located inside a jet with $\Delta R \leq 0.4$. In this way the jet based overlap removal exploits the relative strength of matrix element and parton shower approaches. It is

assumed, that the matrix element does not describe the distribution of close angles correctly due to infrared divergencies, whereas such signatures are better described by the parton showering.
The result of both algorithms is very similar. For our analysis the jet based heavy flavor overlap removal is used.

Chapter 4

Selection of Events

To enrich the number of signal events in the analyzed data sample, several selection criteria starting from trigger selections to more specialized methods like b-tagging or the QCD veto are applied in the offline reconstruction. The expected decay products

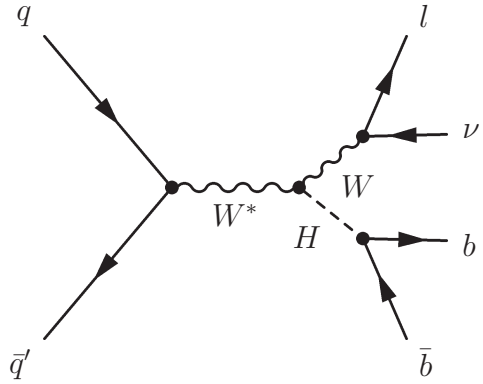

Figure 4.1: Feynman diagram of the W-Higgs production

of the associated W-Higgs process are two hard jets originating from the Higgs decay into two b quarks as well as a single charged lepton and a neutrino stemming from the W decay. We do not consider hadronically decaying W bosons in this analysis. The following selection criteria have been optimized for the search for electroweak single-top quark production [38], but since the final state of the single-top processes is similar to the one of the W-Higgs, it is possible to adopt these requirements.

4.1 Lepton Selection

All events are triggered via a high-p_T electron or a muon. Events consisting a τ originating from the W decay are not separately triggered. The applied trigger criteria are as follows:

- For central electrons the trigger requires a COT track with $p_T > 9$ GeV and an energy cluster in the CEM calorimeter with $E_T > 18$ GeV.
- Forward electrons have to have E_T >20 GeV in the PEM with $E_{\mathrm{HAD}}/E_{\mathrm{EM}} <$

0.075, where E_{HAD} is the energy deposited in the hadronic calorimeter and E_{EM} the energy measured in the electromagnetic calorimeter.

- The muon triggers require a COT track with high p_T as well as either a simultaneous detection in CMU and in CMP (called CMUP muons) or a detection in the CMX calorimeter.

In our offline reconstruction we try to improve the purity of the trigger lepton selection. In case of electrons we require, that a reconstructed track with $p_T > 10$ GeV has to match a CEM cluster with $E_T > 20$ GeV; also the electron has to fulfil $E_{HAD}/E_{EM} < 0.055 + 0.00045 \cdot E$, to avoid hadronic jets to fake an electron, and $E/p < 2.0$ for track momenta smaller 50 GeV. For all electron events we apply a veto to reject electrons stemming from photon conversion into e^+, e^-: If we find a second high p_T-track which has a common vertex with the track of the triggered electron but an opposite track curvature, the event is rejected. For the plug electrons in addition at least three r-ϕ hits in the silicon and no z or stereo signal are required. Such a track signature is called Phoenix track (PHX, [39]). For the muons we require $p_T > 20$ GeV and we demand, that the energy deposited in the calorimeters along the track has to be consistent with a minimum ionizing particle. We also require additional criteria to reject cosmic muons or muons originating from the decay of long lived particles. For these events a search for hits in the muon chambers and simultaneously for backward-moving tracks in the COT is accomplished. Such a track appears, if a charged particle flys from outside into the detector [40].
In the final lepton selection we require an isolation cut, i.e. in a cone of $\Delta R < 0.4$ the deposited E_T not assigned to the lepton has to be less than 10% of the lepton E_T. We also reject events with more than one tight lepton or an additional loose lepton to remove events from processes like $Z \to l^+l^-$ or $t\bar{t}$. A loose lepton is defined as a lepton detected only in one of the muon systems CMU, CMP or BMU or a lepton which passes all selection criteria except the isolation cut.

4.2 Jet Selection

Jets consisting of a shower of particles originate from the production of high energetic quarks or gluons. The jets in this analysis are defined by using the JETCLU algorithm [41]. JETCLU is a cone algorithm, which searches for clusters of deposited energy in the calorimeters. The cone algorithm searches for such clusters placed inside a cone with radius $\Delta R \leq 0.4$ in the η-ϕ plane. If there is a lepton, which falls into the cone of the jet, the energy of this lepton is neglected for the determination of the jet energy. From the measured jet energy it is possible to reconstruct the momentum of the initial particles approximately. The energy of the jets is corrected in several steps, called level 1 to level 7, in order to get a measure of the momenta of the initial partons:
Level 1, Relative Energy Corrections: corrects the η dependence of the calorimeter response.
Level 2, unused
Level 3, unused

Level 4, Multiple Interaction Corrections: Energies originating from different $p\bar{p}$ collisions in the same bunch increase the measured jet energy. For a luminosity of 10^{32} cm^{-2}s^{-1} the averaged number of collisions is three. The level 4 factor corrects this mismeasurement.
Level 5, Absolute Energy Corrections: Corrects the calorimeter energy for nonlinearities and uninstrumented parts of the detector. The result of the level 5 correction is assumed to be the energy on particle level after the fragmentation process.
Level 6, Underlying Energy Corrections: Subtracts the energy stemming from the underlying event. Underlying events can occur in the same proton antiproton collision originating from an interaction of other partons of the proton.
Level 7, Out of Cone Corrections: Corrects the particle energy to parton energy. This correction adds the energy of particles produced during the fragmentation, that are not placed inside the cone of the clustering.
For the selection of jets we use level 5 corrected energies, later for the reconstruction of the event we correct to level 7. More details of the jet reconstruction can be found in reference [42].
In the W-Higgs channel we expect to have two b quarks forming jets. Therefore we accept only events with exactly two so called tight jets. A tight jet must have corrected $E_T > 20$ GeV and detector $|\eta| < 2.8$. The jet is called loose jet, if it has $E_T > 12$ GeV and detector $|\eta| < 2.8$.

4.2.1 b Tagging

The decay of the analyzed Higgs decay contains two b quarks. The b quark has - as part of a b hadron - a large lifetime of averaged 1.5 ps. A high energetic b quark as it originates from the Higgs decay is able to travel a measurable distance of a few mm to cm as part of a hadron after the hadronization process before it decays. The b hadron decays into several particles, whose tracks can be measured to originate from a point displaced from the main interaction vertex. The used SecVtx algorithm [43] searches for such tracks forming a secondary vertex inside a jet cone. The main quantities to classify a track, are the impact parameter d_0, which is defined as the smallest distance of the track with respect to the primary vertex of the reconstructed jet, and the distance between the secondary and the primary vertex in the η-ϕ plane, L_{xy}. They are illustrated in figure 4.2.

The SecVtx algorithm only takes tracks into account, which are assigned to be "good". A "good" track is defined as a displaced track with a minimum number of hits in the silicon detectors and an impact parameter $d_0 \leq 0.3$ cm. All events with at least two "good" tracks are called "taggable", since two is the minimum number of tracks needed to reconstruct a secondary vertex. If an event is assigned to be taggable, the SecVtx algorithm attempts to find a common vertex of the displaced tracks. In the case of a found secondary vertex, the distance L_{xy} and their significance $\frac{L_{xy}}{\sigma_{xy}}$ is determined. To label a jet as b tagged, the significance $\frac{L_{xy}}{\sigma_{xy}}$ has to be larger than 7.5.
For our analysis we require at least one jet to be tagged as a b jet. Since the efficiency

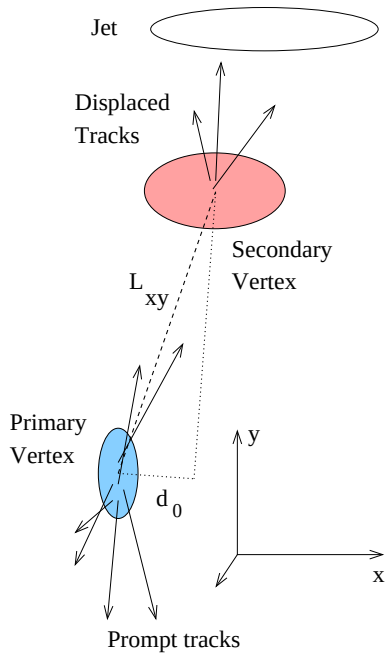

Figure 4.2: Illustration of an event with a secondary vertex.

of the SecVtx algorithm to label a jet which consists a real b hadron is less than about 50%, the requirement of two b tags would lead to a too large reduction of the signal. There are also jets assigned to be b jets, which do not contain any b quark. The first possibility is the existence of c hadrons, whose lifetimes are smaller than the lifetime of the b quark, but still large enough to create a secondary vertex. Another contribution to the tagged events comes from light flavor jets with mismeasured tracks. The mistag rate is of the order of less than 1%, but due to the large cross section of light flavor production (see table 3.1) these events play an important role in the background estimation. The mistag rate as well as the b tagging efficiency depends on the jet momentum and on the pseudo-rapidity of the jet.

4.3 Missing Transverse Energy

Due to the effect that the neutrino interacts only weakly, it cannot be detected directly in the detector. We assume that an undetected neutrino leads to an imbalance of the E_T distribution. Therefore we define the missing transverse energy:

$$\vec{\not{E}}_T := -\sum_i E_T^i \hat{n}_i \tag{4.1}$$

where i is the index of the calorimeter tower with $|\eta| < 3.6$ and $\hat{n}_i$ is the unit vector in the x-y-plane pointing towards the calorimeter with index i. E_T^i is the corrected

transverse energy measured in the ith calorimeter cell. To reject events without a neutrino we demand at least $\not{E}_T = |\vec{\not{E}}_T| = 25$ GeV.

4.4 Further Selection Cuts

The selection of events is improved by some other criteria, to reject events with an impure signature or events taken with a suboptimal detector performance. Hence the following cuts are applied:

The Good Run List
Every store of the Tevatron is usually divided into several runs for the CDF data recording. For some runs the performance of the detector is suboptimal; for instance the function of a subdetector is not available. To avoid the use of data taken during such a run, the quality of every run is checked on- and offline and for each subsystem of the detector separately. All runs useable for an analysis are registered in a good run list [44]. Every data event recorded during a run not listed on a good run list is rejected.

z Vertex Cut
The z vertex cut requires that the primary z vertex of the events is placed in a region within ± 60 cm of the center of the detector.

Lepton z Vertex Cut
If the track of the tight lepton originates from a point z_0, which differs more than 5.0 cm from the primary vertex in z direction, the event is rejected to ensure, that the lepton comes from the same interaction.

Z Boson Veto
To reduce the background coming from events with a Z boson, we define a veto for Z-like events. An event is rejected, if it is possible to combine the tight lepton with a jet, a loose jet or a loose lepton to form an invariant mass in the range of the Z mass. The Z mass range is defined as 76 GeV to 106 GeV.

4.5 QCD Veto

After the explained selection, there is still a remaining fraction of events coming from events without any W boson in the hard scattering process. Most of these events passing the cuts explained above are QCD events in which a mismeasured jet fakes a lepton or the lepton originates from heavy quark decays.
To model these QCD originating events, we had a look at two different data based models. The first model uses data, which have been selected with a jet trigger, wherefore this sample is called jet-electron sample. We require events with a jet fulfilling $E_T > 20$ GeV and $0.05 < E_{HAD}/E_{EM} < 0.2$ as well as at least four

reconstructed tracks. We hope to reject events with real electrons or muons and to keep only events consisting of QCD processes using this trigger. Since we need QCD events faking an electron or muon, we take this data sample and assign one jet per event to be a misidentified lepton. The charge and the flavor, i.e electron or muon, of this faked lepton is chosen randomly.
An alternative model for the QCD background, which is called anti-electron sample, is also taken from data. For this model we require events with an electron, which fails two of five lepton selection criteria, but not the isolation cut. The taken events are used twice, once to simulate electron triggered QCD events, once to simulate muonic events. In the second case the electron is labelled as a muon. The anti-electron sample is not yet available for the forward electrons detected in the PEM calorimeter.
To analyze these QCD background events, we made studies on a QCD enriched sample. For this sample we removed the requirement of the b tag. Hence we expect a sample, which is dominated by W+light jets events and by non-W production. The data events selected without the b tag are compared with the Monte Carlo generator based backgrounds, i.e. all backgrounds except the non-W model. The Monte Carlo samples are combined according to their predicted cross sections and their selection efficiencies. The overall amount of Monte Carlo events is scaled to the observed data. Later on we will use a more precise combination of the background Monte Carlo samples in the tagged sample, but this will be done, if all final selection cuts are defined, and it will only change the fractions in-between light and heavy quark samples.
Up to now we do not include any model for the QCD events. Therefore we expect to see a discrepancy between the Monte Carlo simulated sample and the experimental data. We had a look at several kinematic variables like angles between the four measured objects, lepton, jets and $\vec{\not{E}}_T$ as well as momenta, masses etc., separately for the four lepton selections CEM, PHX, CMUP and CMX. In most of these variables all W+jets samples, light and heavy as well as samples with and without additional partons, show a similar behaviour. The most critical variables, where we find a large difference between data and Monte Carlo events, are the missing transverse energy $\not{E}_T$, the $\not{E}_T$-significance MET_{sig}, the transverse mass of the reconstructed W boson, and, depending on the lepton type, the azimuthal angles between $\vec{\not{E}}_T$ and the jets. Here MET_{sig} is defined as:

$$MET_{sig} := \frac{\not{E}_T}{\sqrt{\sum_{jets} C_{JES}^2 \cos^2(\Delta\phi_{jet,MET}) + \cos^2(\Delta\phi_{vtx,corr})}} \tag{4.2}$$

Here C_{JES} is the level 5 jet energy correction factor; $\Delta\phi_{vtx,corr}$ is the azimuthal angle between corrected and uncorrected missing transverse energy.
The transverse W mass is calculated via:

$$M_{T,W} = \sqrt{2p_T^{lep}\not{E}_T - \vec{p}_T^{lep} \cdot \vec{\not{E}}_T}, \tag{4.3}$$

where p_T^{lep} is the transverse momentum of the charged lepton.
MET_{sig} is similar to the missing transverse energy, but it takes also into account

the angular isolation of the direction of $\vec{\not{E}}_T$ from jets. So it is obvious, that this quantity guarantees a good separation of mismeasured $\not{E}_T$ like in QCD events and $\not{E}_T$ originating from a neutrino. Also from the transverse mass of the reconstructed W boson we can discriminate QCD events, since these events do not include any W boson, whereas for events including a W boson, $M_{T,W}$ are accumulated around the W mass at 80.4 GeV. The distributions of the data and the Monte Carlo sample for these critical variables are presented in figures 4.3 and 4.4. Here also the shape of the jet-electron QCD model is shown and it can be seen, that especially for the angular distributions it does not model the difference in data and Monte Carlo correctly. The anti-electron sample has a similar shape in these variables.

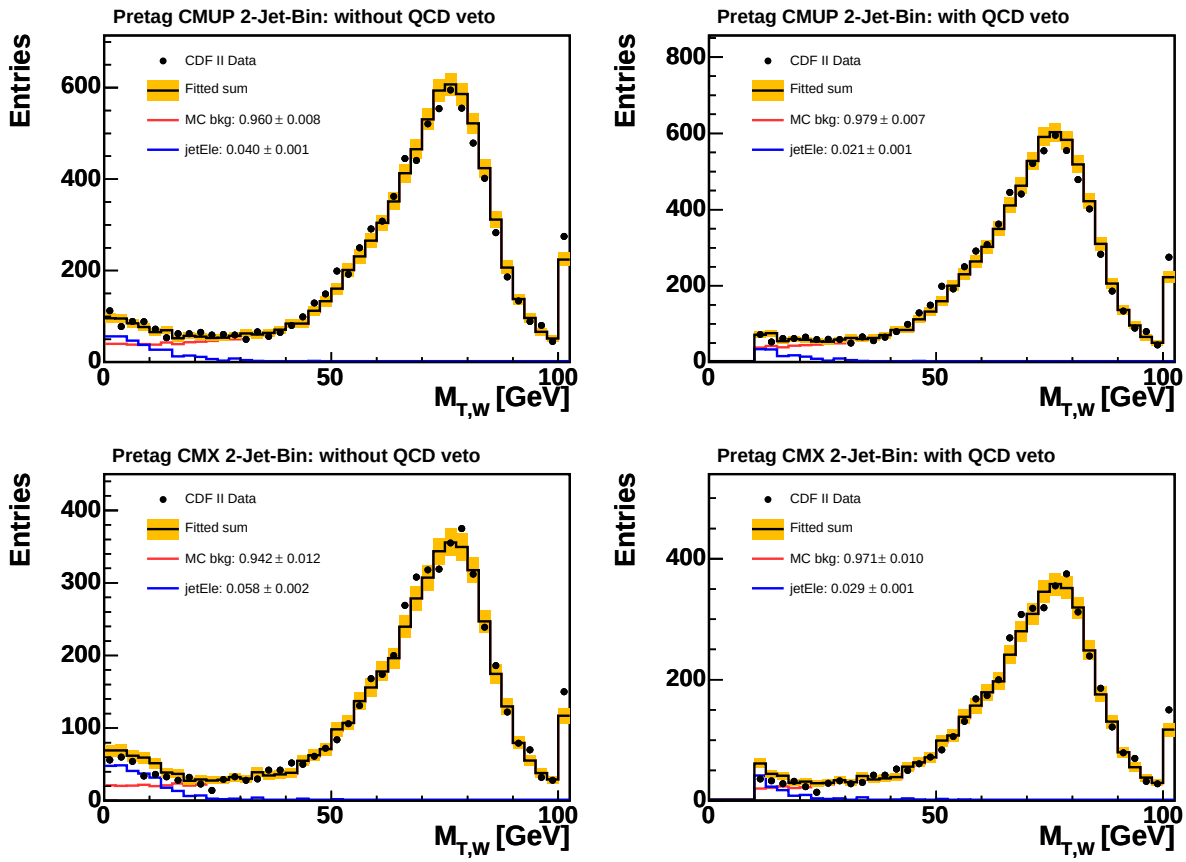

Figure 4.3: Distribution of $M_{T,W}$ for the muonic events

All critical variables are plotted against each other in two-dimensional distributions, to isolate areas with a high fraction of QCD events.
In figure 4.5 for instance the two dimensional distribution of MET_{sig} and $M_{T,W}$ for the CEM calorimeter for events with 2 jets as well as the distribution of $\not{E}_T$ against $\Delta\phi$ between lepton and jet for CMUP in the one-jet bin are shown. From these two-dimensional distributions we try to find cuts, which can reduce the amount of QCD events. Areas in these two-dimensional plots, where much more data events occur than the Monte Carlo sample describes, are assumed to be dominated by

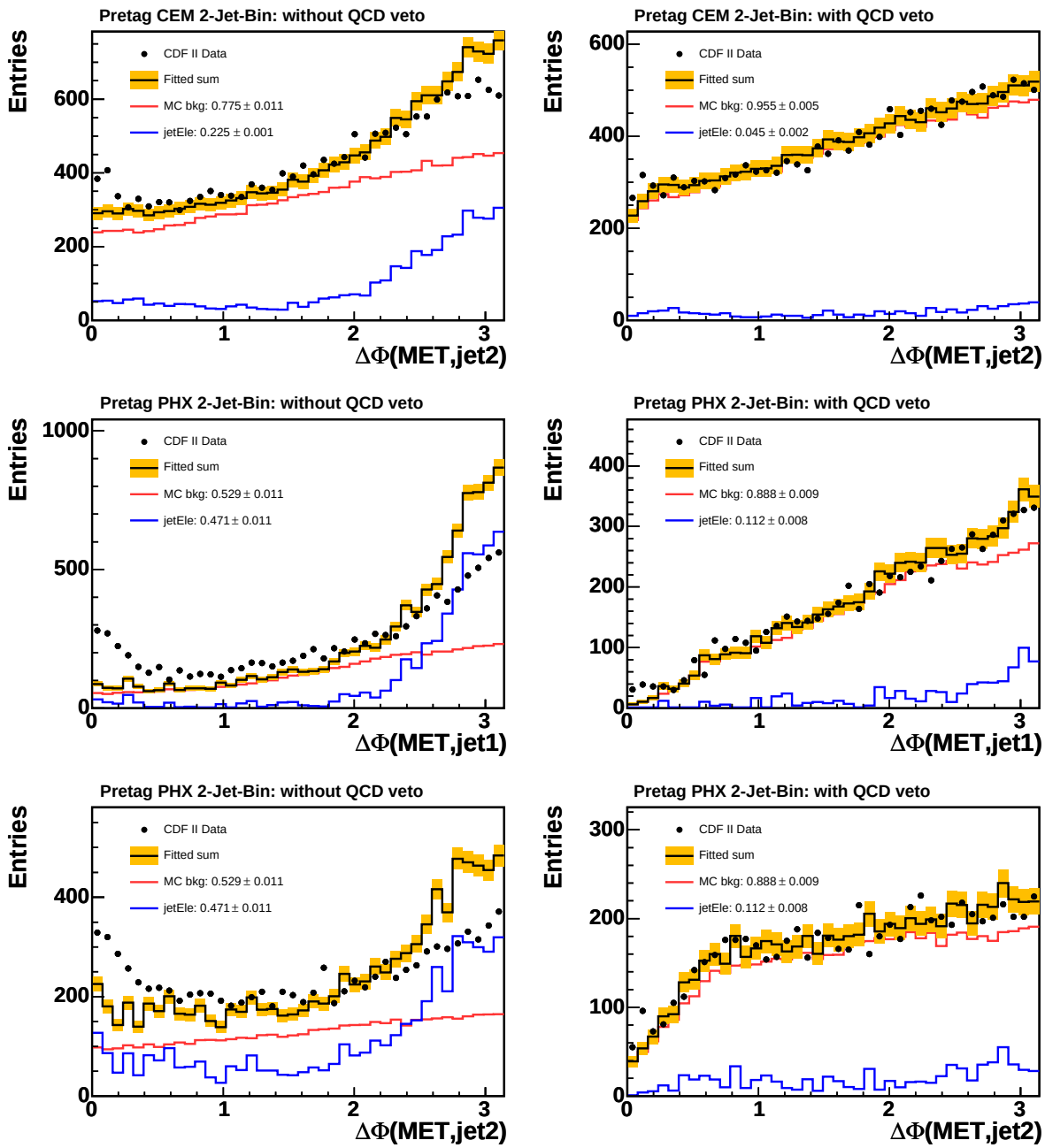

Figure 4.4: Distributions of $\Delta\Phi(\vec{\not{E}}_T, \text{jet})$ in the 2-tag bin. For the PHX events the $\Delta\Phi$ distributions are not well modelled by the jet-electron sample, wherefore QCD veto cuts are applied to these variables.

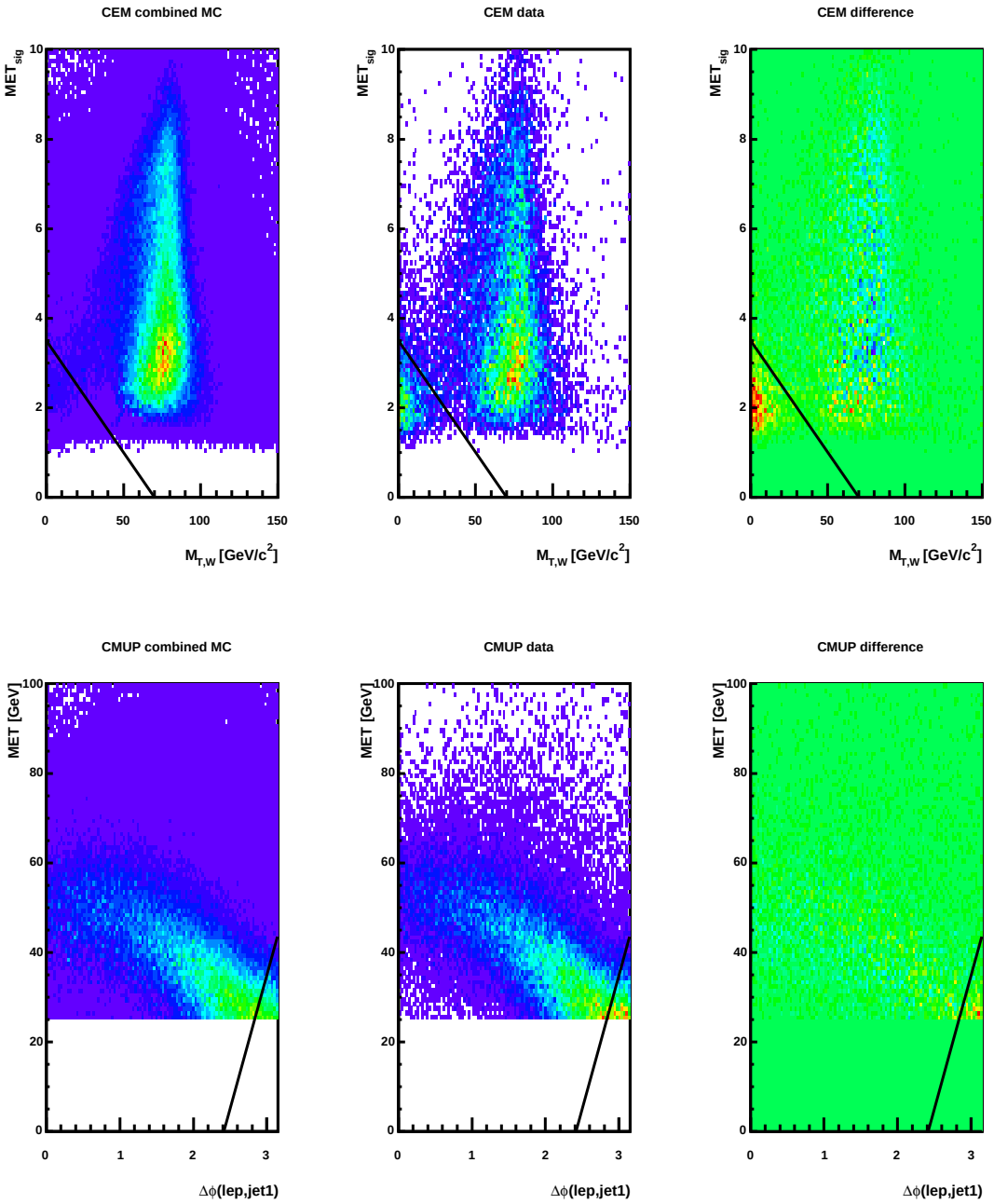

Figure 4.5: Data and Monte Carlo comparison in CEM and CMUP for the QCD veto: The plot on the left hand side shows the distribution for the W+Jets Monte Carlo, the central plot shows the data distribution, in the third plot the difference of the first two plots is presented. The red area indicates a region, where we find an excess of data events which is not modelled by the Monte Carlo, and is therefore assumed to be QCD background.

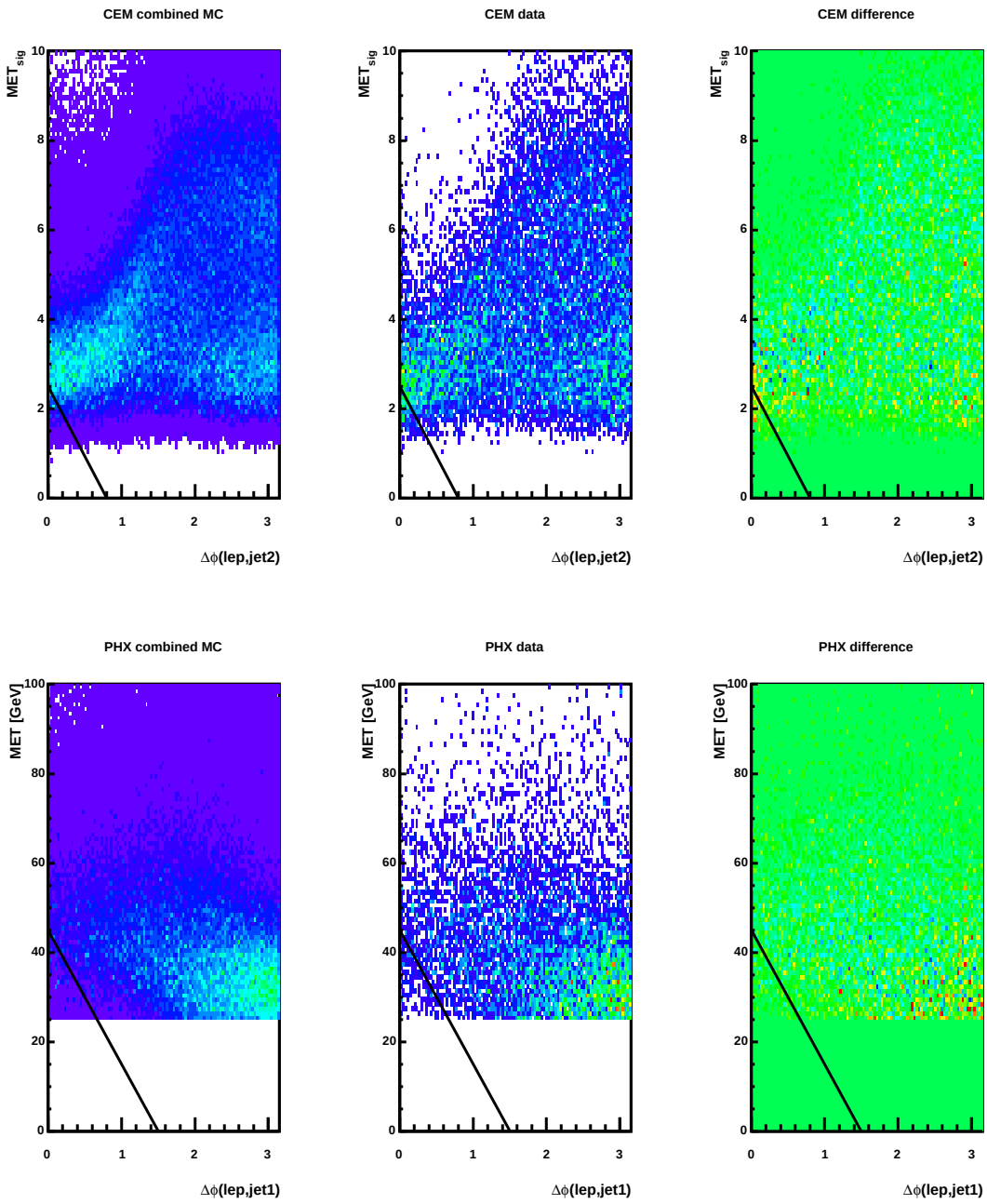

Figure 4.6: Secondary cuts to CEM and PHX events after applying the CEM cut shown in figure 4.5 respectively demanding $MET_{sig} > 2.0$ for the PHX events.

QCD background. If we find such an area with a significant excess of data events, we try to apply a veto to cut these events away. The QCD veto studies have been performed for the explained selection requiring one, two or three tight jets. From this study we decide to use the following cuts:

- For CMX events we require the transverse mass of the reconstructed W boson ($M_{T,W}$) to be greater than 10 GeV.
- For CMUP we reject all events with $M_{T,W} < 10$ GeV and in addition for the one-jet bin all events with $\not\!\!E_T < 60$ GeV$\cdot|\Delta\phi| - 145$ GeV, where $\Delta\phi$ is the angle in the x-y-plane between the lepton and the jet.
- Events, which pass the electron triggers, must have $M_{T,W} > 20$ GeV to be selected.
- For CEM events we require additionally $MET_{sig} > -0.05\ \mathrm{GeV}^{-1} \cdot M_{T,W} + 3.5$ and for the two-jet and the three-jet sample $MET_{sig} > 2.5 - 2.5 \cdot |\Delta\phi|/0.8$ with $\Delta\phi$ the azimuthal angle between $\vec{\not\!\!E}_T$ and the second leading jet (jet with second highest uncorrected E_T). For the one-jet bin we require $MET_{sig} > -7.6 + 3.2 \cdot |\Delta\phi|/0.8$, now with $\Delta\phi$ the azimuthal angle between charged lepton and the jet.
- PHX events must have $MET_{sig} > 2.0$ and for the first two leading jets $\not\!\!E_T > 45$ GeV-30 GeV$\cdot|\Delta\phi|$, where $\Delta\phi$ is the azimuthal angle between $\vec{\not\!\!E}_T$ and the respective jet. If we only have one selected jet, only this requirement is applied, $\Delta\phi$ being the angle between $\vec{\not\!\!E}_T$ and the solely selected jet.

The cuts of the QCD veto are illustrated in figures 4.5 and 4.6. The amount of lost signal events in the respective subdetectors ranges from 18.6% for PHX, where we apply the hardest QCD cuts to remain with a describable sample, to 7.5% for CMUP. All cut efficiencies for the QCD veto and for different processes are summarized in table 4.1. The QCD veto has been optimized to reduce the signal fractions as little as possible, while remaining with a describable QCD background.

subdetector	single-top		W-Higgs	W+jets	non-W
	s-channel	t-channel			
CEM	94.0%	95.6%	92.3%	95.6%	9%
PHX	86.4%	86.7%	81.4%	85.9%	11%
CMUP	97.2%	98.3%	96.4%	98.0%	37%
CMX	97.3%	97.9%	96.5%	98.1%	41%

Table 4.1: Efficiencies of the QCD veto. For the calculation of the non-W efficiencies the jet-electron sample is used.

Although a large number of QCD events are cut away by the QCD veto, there is still a remaining contribution to the data sample, which originates from such non-W events. As mentioned above, there exist different methods to model the QCD

background using data based samples.
We have no satisfying theoretical prediction of the QCD event rate, since many different processes contribute to the QCD background, most of them only due to mismeasurements. Therefore, we determine the rate of non-W events based on the observed data. The rate of the non-W background is estimated by fitting the missing transverse energy distribution. The selection of events is performed as explained above, now including the QCD veto, but without the cut on the missing transverse energy at 25 GeV and the requirement of a b tag. In this sample the fraction of QCD events is highly enriched. We use this selection to prove the quality especially of the different QCD models as well as the modeling of the W+light flavor samples. In this selection, we can also be sure, that the contribution stemming from signal or top-quark events can be neglected.
Due to the different shapes of the non-W and the other backgrounds in the missing transverse energy distribution, it is possible to fit the non-W model and the combined remaining Monte Carlo samples to the data to determine the fraction of QCD events in the observed data. Here the PHX Monte Carlo samples have to be weighted with the trigger turn-on curve [47]. The acceptance of the three trigger stages of the PHX trigger depends on the uncorrected missing transverse energy $\not\!\!E_T^{\mathrm{raw}}$ and the transverse momentum of the electron E_T^{ele}. For low values the trigger has a reduced acceptance, which is not considered in the Monte Carlo samples. Therefore all Monte Carlo generated events are weighted with the following factor:

$$\frac{1}{1+e^{-b_1(\not\!\!E_T^{\mathrm{raw}}-a_1)}} \cdot \frac{1}{1+e^{-b_2(E_T^{\mathrm{ele}}-a_2)}} \cdot \frac{1}{1+e^{-b_3(E_T^{\mathrm{ele}}-a_3)}} \tag{4.4}$$

Here every factor models the turn-on curve of one of the three trigger levels. The parametrization factors are taken from [47] and are listed in table 4.2.

	Level 1	Level 2	Level 3
a_i	14.8 GeV	21.46 GeV	20.69 GeV
b_i	0.37 GeV^{-1}	0.42 GeV^{-1}	0.78 GeV^{-1}

Table 4.2: Parameters for the PHX turn-on curve

A likelihood fit to the $\not\!\!E_T$ distribution is performed. From this fit we can conclude the QCD fraction in the final selection by integrating the fitted distributions of the non-W model and the Monte Carlo samples above 25 GeV and determing the ratios of the integrals. The fits are performed separately for the different subdetectors CEM, PHX, CMUP and CMX. The fit range includes the $\not\!\!E_T$ distribution from 0 to 100 GeV and one overflow bin for all events with $\not\!\!E_T$ >100 GeV. Due to the turn-on curve for the PHX electrons the fit range starts at 15 GeV in this subsample. In figure 4.7 and table 4.3 the results of the fits for the sample with two selected jets are shown using the jet-electron sample as QCD model. In addition, the fits are also performed without including the QCD veto. As can be seen, it is impossible to describe the $\not\!\!E_T$ distribution without the QCD veto in all subdetectors under the assumption, that the jet-electron sample models the QCD background correctly.

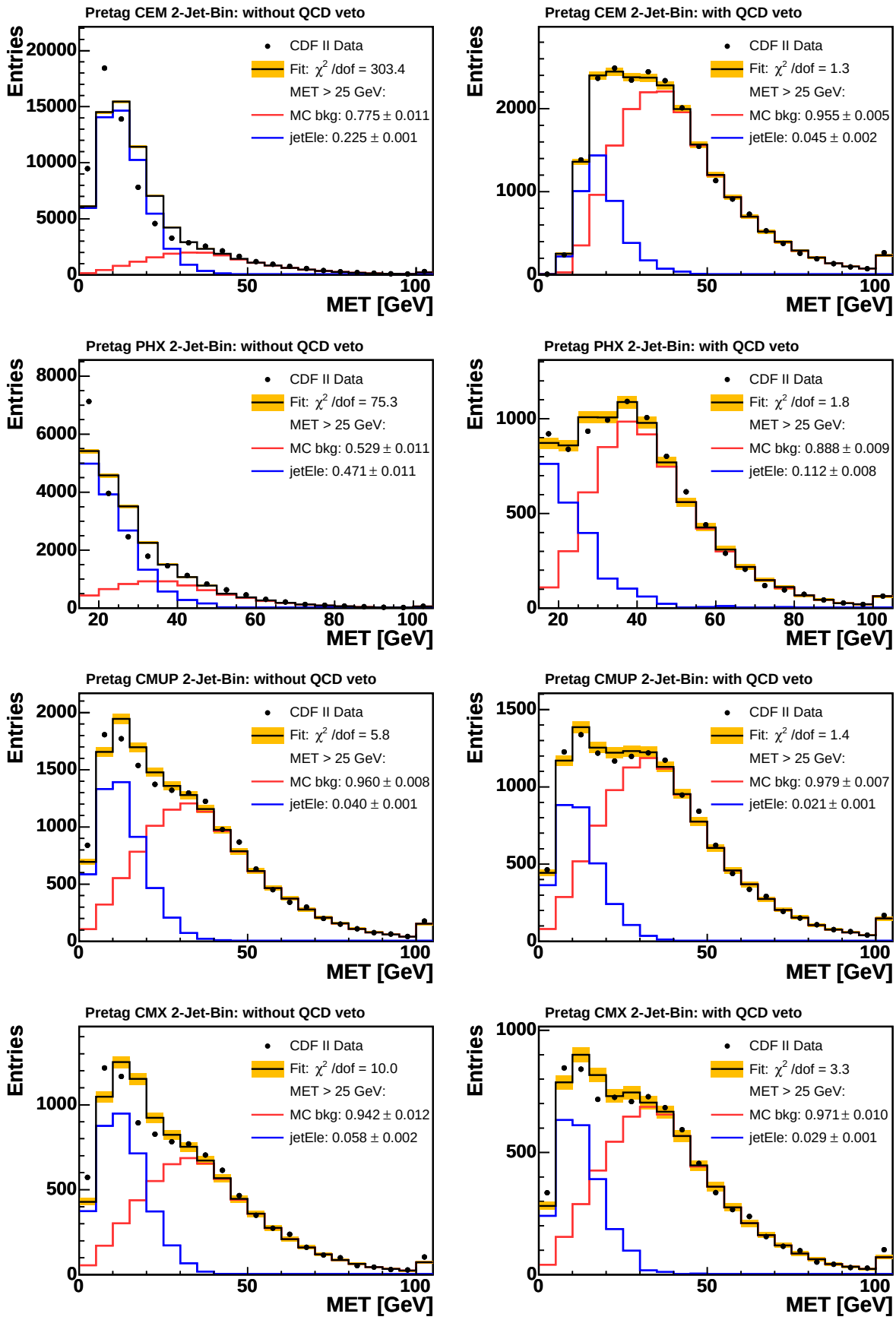

Figure 4.7: $\not\!\!E_T$ fits without (left) and with QCD veto (right) in the pretag sample with two tight jets

From the bad agreement between data and model it is also unfeasible to determine the fraction of QCD events in the selected data.

	non-W fraction	MC background
CEM	4.5 ± 0.2 %	95.5 ± 0.5 %
PHX	11.2 ± 0.8 %	88.8 ± 0.9 %
CMUP	2.1 ± 0.1 %	97.9 ± 0.7 %
CMX	2.9 ± 0.1 %	97.1 ± 1.0 %

Table 4.3: Results of the fit to the $\not\!\!E_T$ distribution applying the QCD veto.

For the anti-electron sample the results of these fits look quite similar, but since it is not available for the plug electrons and has a lower statistic, we decided to model the QCD background using the jet-electron sample.
As a cross-check of the QCD veto we compare the data with the predicted distributions of many kinematic variables. The composition of non-W and further backgrounds are accomplished according to the prediction of the $\not\!\!E_T$ fits. Especially in variables, which show a different shape distributions for the non-W and the Monte Carlo generated backgrounds, we see an improvement of the description. Some of these variables, which are also used in the QCD veto to discriminate the non-W background, are shown in figures 4.3 and 4.4.

4.6 Signal Expectation

To estimate the number of Higgs events we have to combine theoretical predictions as well as detector effects and efficiencies. The number of expected events $\hat{\nu}$ for a process is generally determined by:

$$\hat{\nu} = \sigma \epsilon_{\text{evt}} \mathcal{L}_{\text{int}}, \tag{4.5}$$

where σ is the theoretically predicted cross section of the process, ϵ_{evt} is the event detection efficiency and $\mathcal{L}_{\text{int}}$ is the integrated luminosity. For this analysis we use an amount of data corresponding to 1.9 fb^{-1}.
For the W-Higgs process σ in equation 4.5 is the product of the Higgs production cross section and the branching ratios for $H \to b\bar{b}$ and $W \to l\nu$, where l indicates electron, muon or tau. The branching ratio for a leptonic decay of the W boson is $BR(W \to l\nu) = 0.3257 \pm 0.0028$ [8]. The values for the Higgs cross section multiplied with the branching ratio for $H \to b\bar{b}$ can be found in table 4.4. Cross sections have been calculated up to next-to-next-to leading order (NNLO) [45].
The event detection efficiency is calculated as follows:

$$\epsilon_{\text{evt}} = \epsilon_{\text{evt}}^{\text{MC}} \epsilon_{\text{corr}} \epsilon_{\text{trigger}}, \tag{4.6}$$

where $\epsilon_{\text{evt}}^{\text{MC}}$ is the event detection efficiency obtained from the Monte Carlo sample and $\epsilon_{\text{trigger}}$ is the trigger efficiency. Trigger and Monte Carlo based efficiencies are

mass	$\sigma \cdot BR(H \to bb)$
110 GeV	0.162 pb
115 GeV	0.131 pb
120 GeV	0.103 pb
130 GeV	0.063 pb

Table 4.4: Expected cross section for the Higgs process

calculated separately for each used subdetector CEM, PHX, CMUP and CMX as well as for the periods of data acquisition. For the determination of these factors see references [46], [47], [48] and [49]. We take numbers averaged over all periods and weighted by the respective luminosity. The averaged efficiency factors are listed in table 4.5.

	CEM	PHX	CMUP	CMX
trigger efficiency	0.965±0.002	0.946±0.001	0.913±0.002	0.900±0.002
ID/reco efficiency	0.979±0.001	0.937±0.002	0.930±0.002	0.979±0.003

Table 4.5: Expected efficiencies for trigger and lepton ID/reconstruction

ϵ_{corr} is the efficiency correction factor, which consists of the following terms:

$$\epsilon_{\text{corr}} = \frac{\epsilon_{\text{z}_0}^{\text{data}}}{\epsilon_{\text{z}_0}^{\text{MC}}} \frac{\epsilon_{\text{leptonIdReco}}^{\text{data}}}{\epsilon_{\text{leptonIdReco}}^{\text{MC}}} \frac{\epsilon_{\text{tag}}^{\text{data}}}{\epsilon_{\text{tag}}^{\text{MC}}} \tag{4.7}$$

This correction factor accounts for differences in the cut efficiencies in data and Monte Carlo samples. ϵ_{z_0} is the efficiency of the z vertex cut. For the data the z vertex efficiency is $\epsilon_{\text{z}_0}^{\text{data}} = 0.958 \pm 0.002$ [51]. The Monte Carlo z vertex efficiency is determined from the CDF detector simulation.
The scale factor $\frac{\epsilon_{\text{leptonIdReco}}^{\text{data}}}{\epsilon_{\text{leptonIdReco}}^{\text{MC}}}$ incorporates the differences of identification and reconstruction of leptons. The values for the subdetectors can be found in [50] and [49]. Again we average the numbers for the lepton identification and reconstruction over the different periods of data collection. The last factor in equation 4.7 considers different b tagging efficiencies for data and Monte Carlo. The b tagging efficiency for one single jet is given by [52]:

$$SF_{\text{tag}} := \frac{\epsilon_{\text{tag-jet}}^{\text{data}}}{\epsilon_{\text{tag-jet}}^{\text{MC}}} = 0.95 \pm 0.01(\text{stat.}) \pm 0.04(\text{syst.}) \tag{4.8}$$

If there is one and only one tagged jet per event, this factor would be the factor, which has to be applied in equation 4.7. Due to the possibility of finding double-tagged events (we only select events with two jets, so there are no triple-tagged events) the correction factor for the b tagging efficiency reads slightly different. If f_1 and f_2 are the fractions of 1-tag and 2-tag events in the Monte Carlo signal sample,

the corresponding correction factors K_1 and K_2 are given by:

$$K_1 = SF_{\text{tag}} + 2\frac{f_2}{f_1}SF_{\text{tag}}(1 - SF_{\text{tag}}) \tag{4.9}$$

$$K_2 = SF_{\text{tag}}^2 \tag{4.10}$$

If we want to estimate the number of single-tagged events, we have to apply $\frac{\epsilon_{\text{tag}}^{\text{data}}}{\epsilon_{\text{tag}}^{\text{MC}}} = K_1$ in equation 4.7, for the estimation in the double-tagged bin we have to set $\frac{\epsilon_{\text{tag}}^{\text{data}}}{\epsilon_{\text{tag}}^{\text{MC}}} = K_2$. The fractions f_1 and f_2 are determined from the Monte Carlo sample and calculated separately for each subdetector. For K_1 we find values of about 0.98 for the four analyzed Higgs mass Monte Carlo samples, K_2 has always the same value of 0.90, since it is independent of the Monte Carlo prediction. The expected

mass	$\hat{\nu}$	$\hat{\nu}_{1\text{Tag}}$	$\hat{\nu}_{2\text{Tags}}$
110 GeV	5.28	4.13	1.14
115 GeV	4.43	3.47	0.97
120 GeV	3.54	2.75	0.80
130 GeV	2.26	1.75	0.51

Table 4.6: Expected number $\hat{\nu}$ of Higgs events

number of events $\hat{\nu}$ of this estimation are listed in table 4.6. The total observed number of events for $\mathcal{L}_{\text{int}} = 1.9$ fb^{-1} is 1355, 1282 events are observed with two jets and one b tag, 73 events are observed demanding two b tags.

4.7 Background Estimation

After the selection cuts having been defined, a precise estimation, especially for the b tagged sample, of the background composition is needed. The expected number of background events are obtained by using the Method 2 prediction [53]. Method 2 is a technique to estimate the fraction of W+jets events in the pretag and the tagged sample. The estimation for the diboson, $t\bar{t}$, and the single-top processes is similar to the method for the expectation of Higgs events explained above. The rate of the QCD background is estimated by fitting the non-W model and the remaining Monte Carlo samples to the $\not{E}_T$ distribution. These missing transverse energy fits are applied in the tagged and in the pretag sample as well as for the four subdetectors. From these results the number of W+jets events in the pretag N_{pre}^W and in the tagged sample N_{tag} are estimated as follows:

$$N_{\text{pre}}^W = N_{\text{pre}}^{\text{data}}(1 - F_{\text{nonW}}) - N_{\text{pre}}^{\text{diboson}} - N_{\text{pre}}^{t\bar{t}} - N_{\text{pre}}^{\text{singletop}} \tag{4.11}$$

$$N_{\text{tag}}^W = N_{\text{pre}}^W \left(\sum_i \epsilon_i^{\text{hf}} \cdot K \cdot F_i^{\text{hf}} + \epsilon^{\text{lf}}(1 - \sum_i K \cdot F_i^{\text{hf}}) \right). \tag{4.12}$$

In these equations $N_{\text{pre}}^{\text{data}}$ is the number of observed events in the pretag sample, $N_{\text{pre}}^{\text{diboson}}$, $N_{pre}^{t\bar{t}}$, and $N_{\text{pre}}^{\text{singletop}}$ are the predicted numbers of the respective processes, F_{nonW} is the fraction of QCD events obtained from the $\not\!\!E_T$ fit. In equation 4.12 the sum runs over all heavy flavor samples. ϵ_i^{hf} are the tagging efficiencies of these samples, ϵ^{lf} is the tagging efficiency of the light flavor events. F_i^{hf} denotes the fraction of the heavy flavor sample i in the tagged sample. The fractions of the different heavy flavor samples are estimated from leading order Monte Carlos. The overall fraction of light and heavy flavor events is predicted from fits to the output of the neural network flavor separator (see also section 5.5). This leads to a correction factor K. Every heavy flavor fraction is multiplied with K, so that $N_{\text{pre}}^{W} \cdot \epsilon^{\text{lf}}(1 - \sum_i K \cdot F_i^{\text{hf}})$ becomes the number of tagged light quark events. The correction factor for our background estimation is $K = 1.4 \pm 0.4$.
The results of the Method 2 prediction are presented in table 4.7. We collect samples belonging to the same or to similar physical processes, for example all samples including at least one c quark and one W boson but several numbers of additional partons are collected in the template $Wc/Wc\bar{c}$. The errors include uncertainties to the theoretically predicted cross section, the luminosity, the b tagging efficiency, the K factor, and the lepton identification/reconstruction.

process	$\hat{\nu}$	$\hat{\nu}_{1\text{Tag}}$	$\hat{\nu}_{2\text{Tags}}$	Δ
single-top s-ch.	30.2	23.4	6.8	14.2%
single-top t-ch.	46.6	45.4	1.3	14.6%
$Wb\bar{b}$	406.3	359.2	47.0	35.1%
mistags	274.6	273.5	1.1	12.8%
$Wc/Wc\bar{c}$	411.5	406.0	5.4	35.7%
QCD	56.2	55.2	1.1	40.0%
WW,WZ,ZZ	51.3	48.5	3.0	7.7%
Z+jets	17.4	16.6	0.7	14.5%
$t\bar{t}$	104.1	85.1	19.0	20.9%
total prediction	1398.1	1312.8	85.4	21.0%

Table 4.7: Expected number of events $\hat{\nu}$ and the relative errors Δ of each background process

4.8 Event Reconstruction

The measured signals in the detector are not all directly convertible into physical variables. From the applied cuts we expect to see two jets, the track and the energy depositions in the calorimeter of one tight lepton and missing transverse energy, but the relevant variables we are interested in are the four-vectors and quantum numbers of the b quarks, the lepton and the neutrino. Having reconstructed these values it is possible to calculate the properties of the Higgs and the W boson, respectively the

corresponding properties of the background processes.
The charged lepton can be reconstructed very easily. The momentum and energy are obtained from calorimeter and tracking information, the charge is determined by the curvature of the track.
The neutrino is the next object to be reconstructed. Since the neutrino is not visible in the detector we have to combine the information obtained from the lepton and missing transverse energy ($\vec{\not{E}}_T$). The vector $\vec{\not{E}}_T$ is assigned to be the transverse components $p_{x,\nu}$ and $p_{y,\nu}$ of the neutrino four-momentum p_ν. The only unknown property of the neutrino is the polar angle, or respectively the z component of the neutrino momentum.
We assume, that the lepton and the neutrino originate from a W decay. Therefore, the sum of the four-vectors of neutrino p_ν and charged lepton p_l has to be equal to the four-momentum of the W boson p_W. Furthermore, we can assume, that the square of the W four-vector corresponds to the W mass $M_W = 80.4$ GeV. From this argumentation we find the following relation:

$$(p_l + p_\nu)^2 = p_W^2 = M_W^2. \tag{4.13}$$

Equation 4.13 is quadratic in $p_{z,\nu}$, the z component of p_ν, and therefore in general we expect to find two (in some cases complex) solutions. If we find two real solutions for $p_{z,\nu}$, we choose the solution with the smaller absolute value. In Monte Carlo studies for single-top and $t\bar{t}$ it has been shown, that the smaller $|p_{z,\nu}|$ corresponds to the correct value of the neutrino momentum in more than 70% of all events. The two solutions for $p_{z,\nu}$ are given by:

$$p_{z,\nu}^{\pm} = \frac{\mu p_{z,l}}{E_l^2 - p_{z,l}^2} \pm \sqrt{\frac{(\mu p_{z,l})^2}{(E_l^2 - p_{z,l}^2)^2} - \frac{E_l^2 p_{T,\nu}^2 - \mu^2}{E_l^2 - p_{z,l}^2}} \tag{4.14}$$

with $\mu = \frac{M_W^2}{2} + \cos(\Delta\phi) p_{T,l} p_{T,\nu}$ and $\Delta\phi$ is the azimuthal angle difference between charged lepton and missing transverse energy.
Due to mismeasured $\vec{\not{E}}_T$ the solutions for $p_{z,\nu}$ can become complex in some cases. This corresponds to the case where the measured transverse mass of the W is greater than M_W. If we take only the real part of the solution, equation 4.13 is no longer fulfilled. Since we assume a mismeasurement of $\vec{\not{E}}_T$, we do not longer take $\not{E}_{T,x}$ and $\not{E}_{T,y}$ to be the components of the neutrino momentum. If we fix the transverse mass $M_{T,W}$ to be equal to M_W (which corresponds to a vanishing of the square root in equation 4.14), we find a quadratic dependence of $p_{x,\nu}$ on $p_{y,\nu}$, i.e $p_{y,\nu} = p_{y,\nu}(p_{x,\nu})$ with in general two different and sometimes complex solutions. In a second step we assume $\vec{\not{E}}_T$ not to be mismeasured completely wrong. Therefore, we try to find a corrected solution for $p_{x,\nu}$ and $p_{y,\nu}$, which is as close as possible to the original missing transverse energy vector. We define the difference

$$\delta(p_{x,\nu}) := \sqrt{\left(p_{x,\nu} - \not{E}_{T,x}\right)^2 + \left(p_{y,\nu} - \not{E}_{T,y}\right)^2}. \tag{4.15}$$

δ is being minimized with respect to $p_{x,\nu}$, in which we only allow $p_{x,\nu}$ to be in a range where $p_{y,\nu}$ does not become complex. Since there are in general two solutions

for $p_{y,\nu}$, we may find two different values for δ. In this case we take the solution for $p_{x,\nu}$ and $p_{y,\nu}$, which is closest to the measured $\vec{\not{E}}_T$, i.e. the smallest δ.
To reconstruct the Higgs boson it is essential to reconstruct the two b quarks, since the quarks are assigned to be the direct decay products of the Higgs. The measured four-momenta (i.e. energy and direction) of the jets are multiplied with the level 7 correction factor as explained above to gain the four-momenta of the quarks. The properties of the Higgs boson are reconstructed by adding the quark vectors.

Chapter 5

Neural Networks

Neural networks are advanced tools of data analysis. They have their origin in the research of artificial intelligence, but today they are applied in a huge range of disciplines. Neural networks try to copy the functions of the human brain to describe complex highly non-linear problems. The general principle of using a neural network is divided into two steps: Learning of historical or simulated data and applying the learned information to further problems.
In our analysis neural networks are applied to separate background and Higgs signal processes. They are also used to improve the identification and separation of b, c, and light quark jets [54]. A detailed description of the used neural-network program-package NeuroBayes® can be found in [56].

5.1 Basics of Neural Networks

Similar to the structure of the brain a neural network is built out of neurons connected with dendrites. Instead of the human brain, which consists of about 10^{10} neurons, the neural network is a very simplified mathematical model of the biological network. A neural network used to run on a computer consists of neurons (or nodes), which receive several weighted input informations which are combined to a common output. There exist different possibilities to calculate the output of a node. One of the easiest models is the McCulloch-Pitts model. For this model the output o_i of a node with index i at the discrete time-step t is given by:

$$o_i(t) = \Theta(\sum_j w_{ij} o_j(t-1) - \mu_i). \tag{5.1}$$

Here the sum runs over all nodes of the neural network. w_{ij} is the weight of the connection between node i and node j. The output of every node at the time $t-1$ is $o_j(t-1)$. μ_i is a threshold for the node i. Since $\Theta(x)$ is the Heavyside function, which is defined as $+1$ for positive arguments x and 0 for negative x, the output of every node in this model can only be 1 or 0. The output of all nodes is updated in discrete steps (from $t-1$ to t) according to equation 5.1.
In more advanced neural networks the Heavyside function in equation 5.1 is replaced by a continuous function. For many networks sigmoid functions are used. A sigmoid

function is a strictly increasing function which maps $(-\infty, \infty)$ to $[-1, 1]$. A popular choice for a sigmoid function is the hyperbolic tangens. An also often used sigmoid function reads as follows:

$$S(x) := \frac{2}{1 + e^{-x}} - 1. \tag{5.2}$$

Important for the training of neural networks is the network topology, i.e. the number of nodes and their connections. A usual topology consists of two layers: a first layer of nodes which receive external input information and a second layer, which delivers the final output of the network. For a network, which shall deliver only a binary classification, only one output node is needed, for more complex applications it is possible to built network topologies with several output nodes. There is often a certain number of intermediate layers, called hidden layers. The connections are arranged in such a way, that every node of one layer is connected with every node of the following layer, but not with the nodes of the same layer. Every connection of the neural network gets an initial weight w_{ij}. In figure 5.1 a typical network topology with one hidden layer is shown. A network with one output node is usually used for classification problems.

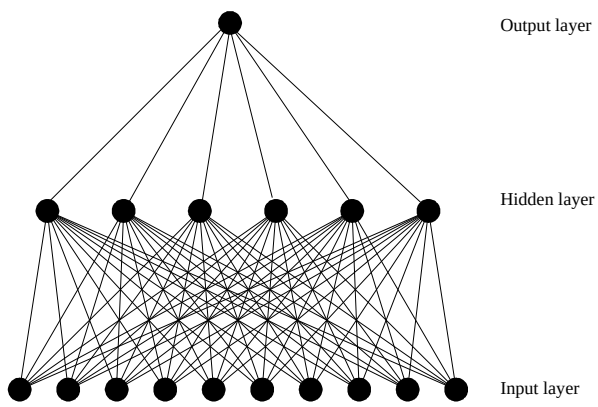

Figure 5.1: Network with one hidden layer and one output node

To each layer of such a multi-layer network an additional single bias node, always having the same output value -1, may be added. This is an alternative way to control the thresholds of the nodes in the following layers by changing the weights of the connections between the bias node and the following nodes. For multi-layer networks, which are networks without circular connections, it is not longer necessary to take the output values of the nodes time dependent, since the output of each layer depends only on the values of the previous nodes. For example the output of node k in the final layer of a network with one hidden layer is given by:

$$o_k = S(\sum_j w_{jk} S(\sum_i w_{ij} x_i)), \tag{5.3}$$

where the first sum runs over all nodes in the second layer j, the second sum runs over all nodes in the first layer i, and x_i is the external input given to the first layer node i.

5.2 Training the Network

The aim of the network training is, that the output of the network for a certain given set of input variables comes as close as possible to an unknown target, which can be a false-true decision or a continuous variable. In the training process the network receives input samples, for which the target is specified. In every training iteration the weights of the connections between the network nodes w_{ij} are changed in order to change the network output to come closer to the target. To measure the quality of the network output for a set of weights, the quadratic error function $E([w])$ is defined:

$$E([w]) := \chi^2 = \sum_k \frac{1}{2} \sum_n \left(T_{kn} - o_{kn}([w])\right)^2, \tag{5.4}$$

where T_{kn} is the target value in the output node k for the training event n. $o_{kn}[w]$ denotes the output given by the network for a certain set of weights $[w]$. In the mathematical view the training of a neural network becomes equivalent to the minimization of the quadratic error function χ^2 with respect to all variables w. An optimal training would lead to $E([w]) = 0$, which means, that the network gives the correct output for every training event.
Usually it is impossible to find the minimum of the error function analytically, since a minimization of a function in a high dimensional space (the number of dimensions is equal to the number of weights in the network) is required. From the numerical point of view it is often more practically to minimize the entropy loss function $-E_D[w]$ instead of the quadratic error function χ^2. The entropy loss function is defined as:

$$E_D([w]) := \sum_k \sum_n \log\left(\frac{1}{2}\left(1 + T_{kn} o_{kn}([w]) + \epsilon\right)\right), \tag{5.5}$$

where ϵ is a small numerical regularization factor, which decreases in every training iteration and will be zero after a few iterations. For the entropy loss function target and output of the network are assumed to be in a range of $[-1, 1]$. The advantage of this function is the divergence which occurs, if target and output are maximally discrepant, which means for instance $o_{kn} = +1$ and $T_{kn} = -1$ or vice versa.
A useful algorithm to minimize a function is the method of gradient descent. At the beginning of the training the weights are usually initialised with random numbers and the error function for these weights is calculated. In every iteration step the weights are updated proportional to the gradient of the error function with respect to the respective weight. The change Δw_{ij} of every weight w_{ij} is calculated via:

$$\Delta w_{ij} = -\eta \frac{\partial E([w])}{\partial w_{ij}}. \tag{5.6}$$

Here η is a parameter, which regularises the step width of the minimum search. If η is chosen too large, the algorithm might pass over the searched minimum, if η is chosen too small, it will take too long until the algorithm converges. For multilayer networks it is advisable to combine the method of gradient descent with a backpropagation algorithm. In a first step the overall error of the output layer is calculated. In a second step also the errors of the previous layers are calculated backwards beginning from the output layer, since the target values are only known for the last nodes. This procedure allows an individual adaption of the weight change Δw_{ij}, by considering the impact of each weight in the hidden nodes.

5.3 Methods for an Efficient Training

Many methods have been developed to improve the sensitivity of neural networks and to avoid overtraining. Especially the preprocessing of the input variables has a huge impact on the training result. It has turned out, that an improved preprocessing of the variables might have a larger effect than the network training itself.

5.3.1 Preprocessing

The NeuroBayes® package used for our analysis provides a completely automated preprocessing of all input variables. To ensure a stable training process all variables are scaled to a codomain of $[-1, 1]$. These scaled variables are flattened into Gaussian distributions with a mean of 0 and a standard derivation of 1 by binning the distribution and scaling the width of each bin. By this transformation all variables lie in the central region of the sigmoid function 5.2, which avoids a saturation process occuring for very large or small values given to equation 5.2. Hence the training process is much more stable and converges faster.
In NeuroBayes® there exists the possibility to handle every input variable individually. Instead of transforming a variable into a Gaussian it is possible to apply a spline fit to the distribution of this variable. In this case not the real value of the variable but the result of the spline fit is treated as input for the network. This procedure avoids the overestimation of statistical fluctuations. Another preprocessing can be chosen for discrete variables, which can be handled as different classes. Also variables only defined on a subset of the training sample can be used. For the subset, on which these variables are undefined, flags are set not to use these inputs for the training.
The input variables given to the network might be correlated towards each other, i.e. two variables include similar information. For this reason not the original but the decorrelated variables are given to the network. To decorrelate the variables, the covariance matrix has to be calculated. The covariance matrix is a $N \times N$ matrix, if N is the number of input variables. The decorrelated variables can be found by diagonalizing the covariance matrix, what means, that the eigenvectors of the covariance matrix divided by the square-root of the corresponding eigenvalues are the decorrelated variables we are looking for. An algorithm to diagonalize a large matrix numerically can be performed using Jacobi rotations [58].

5.3.2 Automatic Selection of Relevant Variables

NeuroBayes® provides the possibility to reject variables with a low correlation to the target automatically. Although a small information might get lost by removing variables, the dimension of the minimization problem is reduced, what can also lead to an improved training result. The significance of each input variable is calculated in an iterative way. In a first step always one input variable is removed and the overall correlation of the remaining $N-1$ variables to the target is calculated. The variable with the smallest loss of the correlation is treated as the least important variable. For the remaining $N-1$ variables this process is repeated. At the end of this procedure all variables are listed according to their relevance. Now it is possible to reject all variables from the list, beginning with the least significant one, until the significance of one variable is found to be larger than a previously defined minimal value.

5.3.3 Optimization during the Training

The training process can be optimized in different ways. Since it is a minimization process in a high dimensional space, it is advantageous to reduce the number of dimensions. This can be accomplished by pruning weights away. If a connection becomes nearly insignificant during the training, the weight of this connection is set to zero. In this way it is also possible to prune complete nodes away, if all connections between this node with an other node are weighted to zero.
Another optimization applied during the training process is the adaption of the step width η in the gradient descent in equation 5.6. The optimal step width is related to the curvature of the error function. For a large curvature the step width has to be smaller than for a low curvature in order to find the minimum as fast as possible. Since the error function depends on several weights, the curvature is parametrized by the Hessian matrix. In this case the optimal step width is determined by the largest eigenvalue of the Hessian matrix. In figure 5.2 the optimization of the step width for a one-dimensional problem is illustrated.

5.4 The Network Output

A trained NeuroBayes® network produces an output, which can be interpreted as a Bayesian a posteriori probability. For a network with for instance one output node, that has been trained to distinguish two classes A and B of events, the output is a number in the interval $[-1, 1]$. Events of class A should pile up at -1, whereas events of class B should produce an output close to $+1$ (or vice versa) to have an optimal separation of both classes, if the network is well trained. Scaling the output linearly to $[0, 1]$, we receive a probability for the event to be out of class B respectively to be not out of class A.
A quality check of the network training is performed by determing the dependence

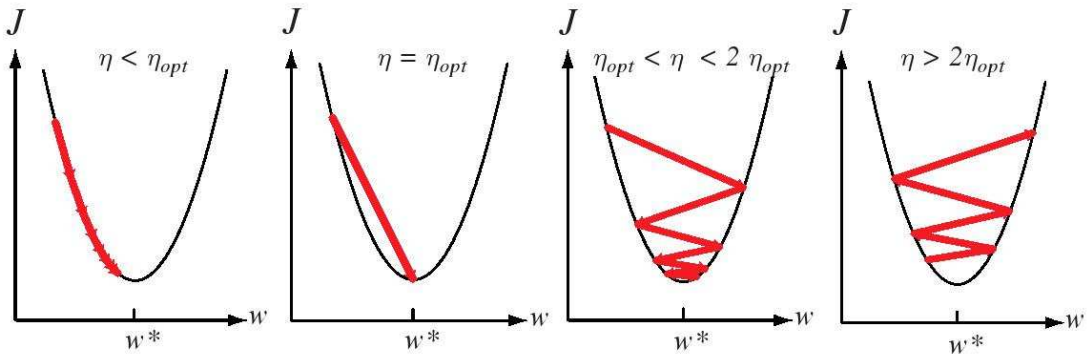

Figure 5.2: Gradient descent for a one-dimensional minimization problem. The optimal step width depends on the curvature of the function $J(w)$, which has to be minimized. If the step width η is too small, the minimum w^* will be reached, but it takes unnecessary many steps. For a too large η the algorithm escapes from the minimum [57].

of the purity on the network output. The purity P for an output o is defined as:

$$P(o) = \frac{N_B}{N_A + N_B}, \tag{5.7}$$

where N_A and N_B are the numbers of events of class A respectively B, for which the network output lies in an interval $[o, o + \delta o]$. For a well trained network the purity P should be proportional to the output o.

5.5 Application: The Neural Network Flavor Separator

The flavor separator [54], [55] is an application of a neural network to improve the purity of the b quark identification in high-p_T jets. The SecVtx algorithm explained in section 4.2.1 provides only a binary decision whether a jet contains a b quark or not. Furthermore, this binary quantity based only on a few cuts applied on the track information resulting from the long b hadron lifetime. The flavor separator is a neural network trained on SecVtx tagged jet samples taken from Monte Carlo generators. It combines many different informations as input variables, which are combined by the network to an output, that can be treated as a probability for b jet identification. Besides the lifetime based informations, like the impact parameter d_0, and the decay length significance $\frac{L_{xy}}{\sigma_{xy}}$, the flavor separator makes use of further variables, like for instance the vertex mass and the decay multiplicity.

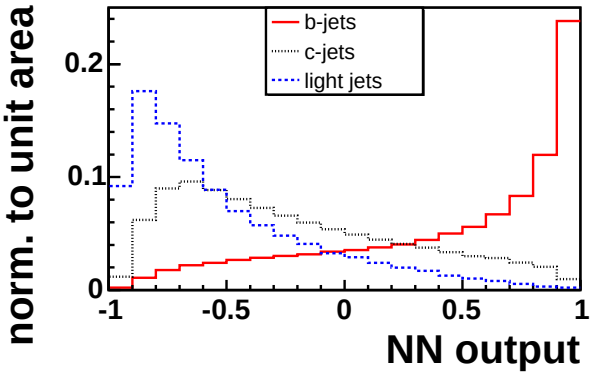

Figure 5.3: The output of the neural network flavor separator.

The output of the flavor separator is shown in figure 5.3. For jets containing a b quark the output of the network accumulates at $+1$, whereas jets without any heavy quark produces an output close to -1. It is also possible to distinguish jets with c but no b quarks. Their output distribution lies between the outputs for b jets respectively light jets.

Chapter 6

Measurement of upper Limits on the Higgs Cross Section

The aim of this analysis is to find an upper limit of the cross section on the W-Higgs process. The measurement of this upper limit is performed in the following steps: We train neural networks with Monte Carlo generated signal events and apply the trained network to the experimental data. In a likelihood fit the output template of the data is compared with the templates of the background and signal models. From the fitted ratio of signal and background events we calculate limits on the Higgs cross section.

6.1 Neural Network Analysis

As explained in the previous chapter, a neural network is a powerful tool for classification problems. In our analysis neural networks are used to distinguish background and signal events. The output of the networks will give a probability for every event to be Higgs- or background-like.
At first the networks have to be tought with information of signal and background events. Therefore, a training sample has to be composed, which consists of Higgs as well as of background events. In this analysis the training samples are filled half-and-half with simulated background and Monte Carlo generated signal events. The composition of the background events corresponds to the expected composition from Method 2 (see section 4.7). The result of the network training depends also on the size of the training sample. In our case the number of training events is limited by the W-Higgs Monte Carlo samples, which contain about 300,000 events for each assumed Higgs boson mass. From these generated events about 15,000 are selected by our event selection, demanding exactly one SecVtx b tag, about 5,000 events have two b tags. We train two networks for each mass sample, one network is trained on the sample with one b tag, another one on the sample with two b tags. The analysis is performed assuming Higgs masses of 110, 115, 120 and 130 GeV. The separation of the trainings for the different masses is performed due to the different distributions of some input variables, for instance the mass of the reconstructed Higgs boson. The separation into one and two tag trainings is used due to the different background

composition.
The information delivered to the network is made up of several physical variables. These variables are chosen in such a manner, that every variable shows a slightly different distribution for the Higgs process and at least one background process. Although every of these variables has only a small discrimination power, the combination of these inputs to the neural network leads to good separation of the Higgs-signal from the background. For the training the following input variables are used:

1. $M_{b\bar{b}}$, the invariant mass of the reconstructed Higgs boson. For the W-Higgs process we expect all events to lie around the assumed mass of the Higgs boson. In events including top quarks the invariant mass of the two b quarks has a maximum at higher values, since the b quarks originate from heavy massive t quark decays, whereas for W+jets the distribution of this variable has a lower maximum than for Higgs events.

2. neuro-Tag 1, the output of the neural network flavor separator for the jet with highest E_T. This variable is taken only into account if the jet with highest E_T has a secondary vertex b tag. As explained in section 5.5 the flavor separator distinguishes events with jets originating from b quark production from events with c or light quarks. For the Higgs the flavor separator can only divide mistag and $W + c/W + c\bar{c}$ events from the signal, $W + b\bar{b}$ and top quark production looks similar to the Higgs events in this variable.

3. neuro-Tag 2, the output of the neural network flavor separator for the jet with lowest E_T. This variable is taken only into account if the jet with lowest E_T has a secondary vertex b tag.

4. $\sum_{\text{jets}} E_{T,\text{jet}}$, the scalar sum of all tight and loose jets. This and the following E_T and P_T variables as well as the jet masses show larger values for top-like events and smaller values for W+jets and diboson backgrounds, where the jets do not originate from the decay of massive objects like the top quark or the Higgs boson.

5. $E_{T,b}$, the transverse energy of the reconstructed b quark four vector with the highest neural network b tagger output.

6. $P_{T,b}$, the transverse momentum of the reconstructed b quark four vector with the highest neural network b tagger output.

7. M_{min}, the invariant mass of the jet with lowest energy (correction level 7).

8. M_{max}, the invariant mass of the jet with highest energy (correction level 7).

9. $P_{T,min}$, the transverse momentum of the jet with lowest energy.

10. $P_{T,W}$, the transverse momentum of the reconstructed W boson.

11. $Q \cdot \eta_{lep}$, the product of the charge and the pseudo rapidity of the lepton. Multiplying the pseudo rapidity with the lepton charge symmetrizes this distribution. For the Higgs signal we expect to find a symmetric distribution for this variable, whereas for some backgrounds the charge and the direction of the lepton might be correlated.

12. η_{W^*}, the pseudo rapidity of the W boson before Higgs radiation. This variable discriminates $t\bar{t}$ and Z+jets events from the other samples.

13. $\cos(\theta^*)$, where θ^* is the azimuthal angle between the Higgs boson momentum and the proton beam axis measured in the W^* rest frame. For the Higgs signal process the spin state of the initial particles has to be $(s, m_s) = (\frac{1}{2}, -\frac{1}{2})$ for the quark and $(\frac{1}{2}, \frac{1}{2})$ for the anti-quark, since only these states are interacting weakly. From the addition theorem for angular momenta and the fact, that the Higgs is a spin-0 particle, the only allowed spin states for the W and the Higgs are $(1, 0)$ and $(0, 0)$. From figure 6.1 it can be seen, that the preferred direction of the transversally polarized W is orthogonal to the direction of the initial partons. There is no such spin correlation for most of the backgrounds, wherefore θ^* shows a flat distribution, respectively in the $\cos(\theta^*)$ distribution the backgrounds cumulate at ± 1, whereas the Higgs signal is centred around 0.

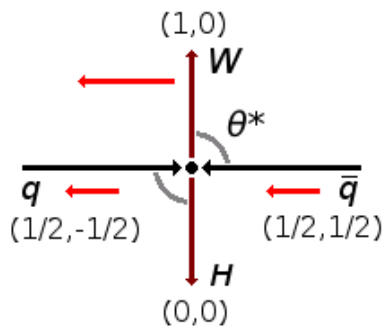

Figure 6.1: Spin correlation in the Higgs production

14. ϕ-topology, which is calculated as follows (here $\Delta\phi$ is the azimuthal angle between the reconstructed neutrino vector and the closest jet four vector):

    - ϕ-topology $= \Delta\phi$ if in the x-y-plane the ordering of the two jets, the lepton and the neutrino is jet-jet-lepton-neutrino or jet-jet-neutrino-lepton.
    - ϕ-topology $= 2\pi - \Delta\phi$ if in the x-y-plane the ordering is jet-lepton-jet-neutrino or jet-neutrino-jet-lepton.

    This variable separates slightly $W + c/W + c\bar{c}$ and especially QCD events, where the lepton and the neutrino four vectors are almost close together.

15. $\eta_{2^{nd}b}$, the pseudo rapidity of the b quark four vector of the jet with lowest neural network b tagger output (if both jets are b tagged) or respectively of

the jet without b tag. This variable especially discriminates the single-top t-channel.

16. $\Delta\phi_{H,W}$, the azimuthal angle between the Higgs and the W boson. For the Higgs process $\Delta\phi_{H,W}$ should be always close to π, since we expect the Higss and the W to fly in opposite directions in the x-y-plane.

17. $\Delta\phi_{MET,lep}$, the azimuthal angle between the missing transverse energy direction and the lepton vector. For QCD events $\vec{\not{E}}_T$ and the $P_{T,lep}$ are often close together, whereas for W+jets events a large angle between these two vectors is preferred.

18. M_{W^*}, the invariant mass of the reconstructed W boson before Higgs radiation. For Higgs events we expect $M_{W^*} \gtrsim m_h + m_W \approx 200$ GeV.

19. $\not{E}_T \sin(\Delta\phi_{min})$, where $\Delta\phi_{min}$ is the angle between $\vec{\not{E}}_T$ and the closest object in the x-y-plane. This quantity is highly correlated to $\not{E}_T$, but it also takes the isolation of the neutrino into account and shows a slightly more significant discrimination than the $\not{E}_T$ only.

20. $\Delta R_{b\bar{b}}$ the distance in the η-ϕ-plane of the two reconstructed quark vectors. For top-like events the two b quark jets originate from the decay of different particles, whereas for instance in W+jets and also in the W-Higgs signal the two jets stem from the decay of the same particle. Therefore, for top-like events the averaged distance between the two jets is slightly larger than for the other events.

21. $M_{l\nu b}$, which corresponds to the invariant top mass if we have a single top or top-antitop event.

22. $P_{T,H}$, the transverse momentum of the reconstructed Higgs boson. $P_{T,H}$ provides a similar information like the P_T variables of the jets.

23. H_T, the scalar sum of the P_T of all jets, lepton and neutrino. This variable shows larger values for $t\bar{t}$ events, whereas W+jet and diboson events have a softer distribution. The Higgs and single-top events lie in between these backgrounds.

24. $\log(A)$, where A is the aplanarity which is given by $A := \frac{3}{2}Q_1$. Here Q_1 is the smallest eigenvalue of the sphericity tensor S_{ij}. The components of the sphericity tensor are given by:

$$S_{ij} := \frac{\sum_{k=1}^{N} p_i^k p_j^k}{\sum_{k=1}^{N} |p^k|^2} \tag{6.1}$$

where i and j denote the coordinates x, y and z, and the sum goes over all N objects of the event, i.e. in our case two jets, charged lepton and neutrino. The sphericity tensor describes an ellipsoid with the aplanarity as the smallest

elongation in the direction of one of the three principle axes. Therefore, the aplanarity indicates the flatness of the momentum distribution of an event. The discrimination power of the aplanarity is not very large, but it is a variable, which is not highly correlated to the other ones.

Altogether we considered about 50 further variables, but we only took variables into account, which are described well by the background model in the b tagged as well as in the pretag sample and which provide a significant discrimination between signal and background. For the final selection of variables we use the automatic variable selection of the NeuroBayes® package (see section 5.3.2). The chosen topology for all networks contains, besides the nodes for the selected variables, one bias node in the first layer. The second layer is filled with twelve nodes. We also trained networks with a different number of secondary nodes, but the training result was very stable under this variation. In the third and final layer we have one node, since we want the network only to deliver a probability for every event to be Higgs- or background-like.
All used input variables have been checked in the tagged and in the pretag sample separately for the lepton categories CEM, PHX, CMUP and CMX. They all show good agreement with the observed data. In the pretag sample we apply all explained selection cuts except the requirement of a secondary vertex b tag.
For the tagged sample we apply an ordinary χ^2 test and calculate the corresponding probability. Due to the higher statistic in the pretag sample the ordinary χ^2 probability is almost zero. For this reason we define a $\tilde{\chi}^2$ probability which is based on pseudo experiments and also incorporates jet energy scale uncertainties.
At first the ordinary χ^2_{data} of the data Monte Carlo comparison is calculated. In the next step we make N pseudo experiments with randomly built distributions following the original Monte Carlo samples. To produce an systematic shape uncertainty we use the jet energy scale $\pm 1\sigma$ errors, which are the largest systematic uncertainties to the W+jets process, of each Monte Carlo sample to smear every single distribution. We also include the rate uncertainties of each physical process to vary the fraction of every subsample. The smeared samples are put together and normalized to the data. For every pseudo experiment we calculate the χ^2 for the comparison of the smeared pseudo experiment with the original Monte Carlo distribution. We count all pseudo experiments, which have $\chi^2 > \chi^2_{data}$. The number of these pseudo experiments is called M. The $\tilde{\chi}^2$ probability is then given by $\frac{M}{N}$. A variable is accepted to be used in our analysis if the $\tilde{\chi}^2$ probability is greater than 0.01.
On the following pages some of the most important input variables of the neural network are shown (figures 6.2 to 6.6). The upper left plot shows the data and Monte Carlo distribution of the variable after applying the secondary vertex b tag. In the upper right histogram the shapes of the different background processes are compared with the expected shape of the W-Higgs process. The further four plots show the data compared to the pretag Monte Carlo expectation for each subdetector separately. For further considered variables (see figures 6.7 to 6.25) only the comparison between data and background model in the tagged sample and the shape comparison with the Higgs model is shown.

For the modeling of the input variables - and also later on of the neural network output - the number of selected events after requiring the b tag in the mistag Monte Carlo sample as well as in the QCD background model is not large enough to deliver smooth distributions. The QCD background is based on data events, wherefore an enlargement of this sample is not accomplishable. From the mistag Monte Carlo only a small fraction of produced events is selected. To produce enough mistag events with a real b tag, much more CPU time would be necessary.
Therefore, for the mistag and QCD backgrounds we only demand at least one jet to be taggable. In this case also the output of the neural network flavor separator has to be chosen randomly. For the mistag events we draw a random number following the output for light jets (see figure 5.3). For the non-W background in a first step the flavor of an individual event is assigned to be b like (45% of all cases), c like (40%) or light flavor like (15%) [59]. Depending on the chosen flavor, the output of the flavor separator for this jet is drawn randomly from the respective output template (see figure 5.3).

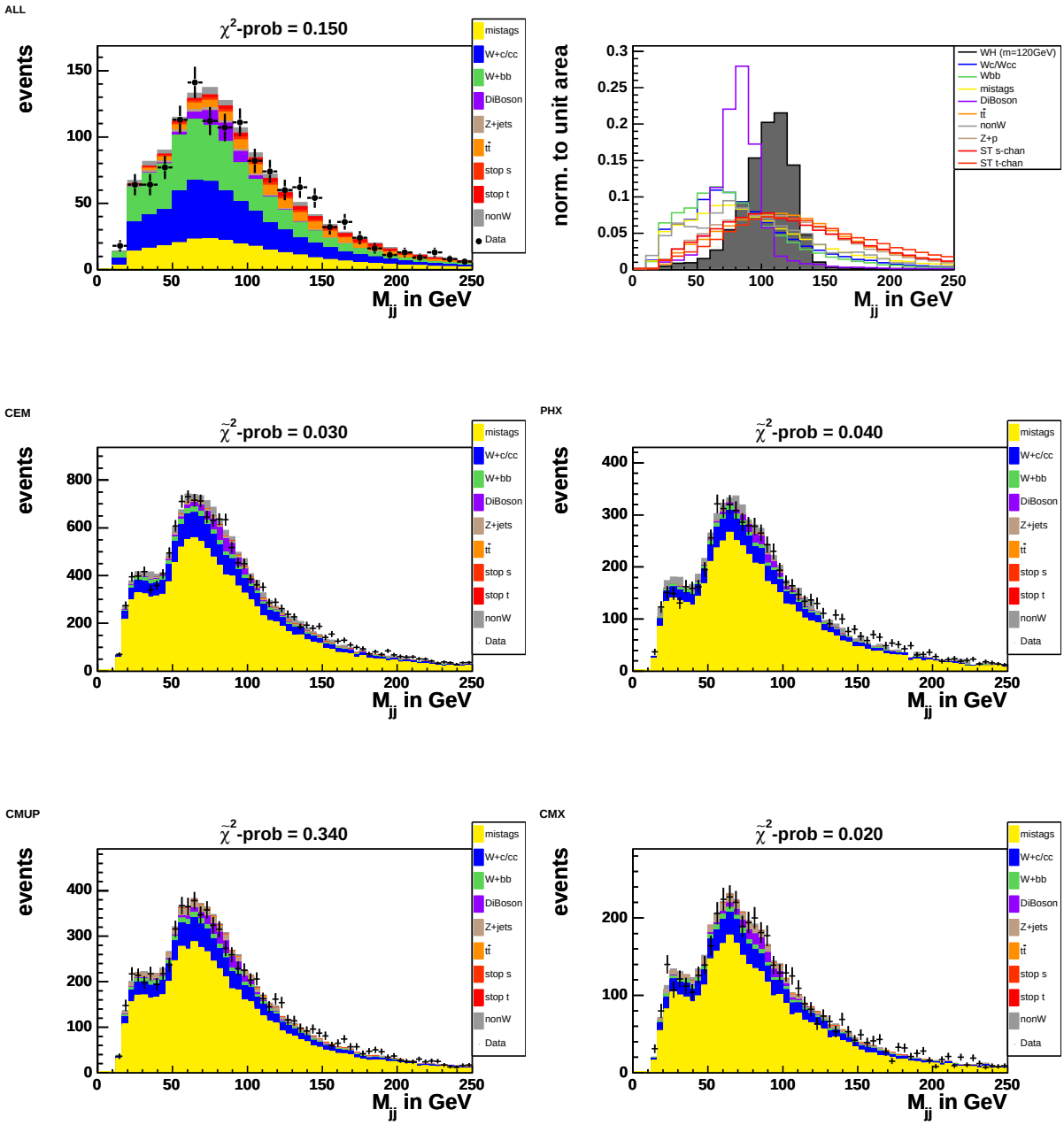

Figure 6.2: The invariant mass of the two quarks, for the W-Higgs process this variable corresponds to the Higgs mass

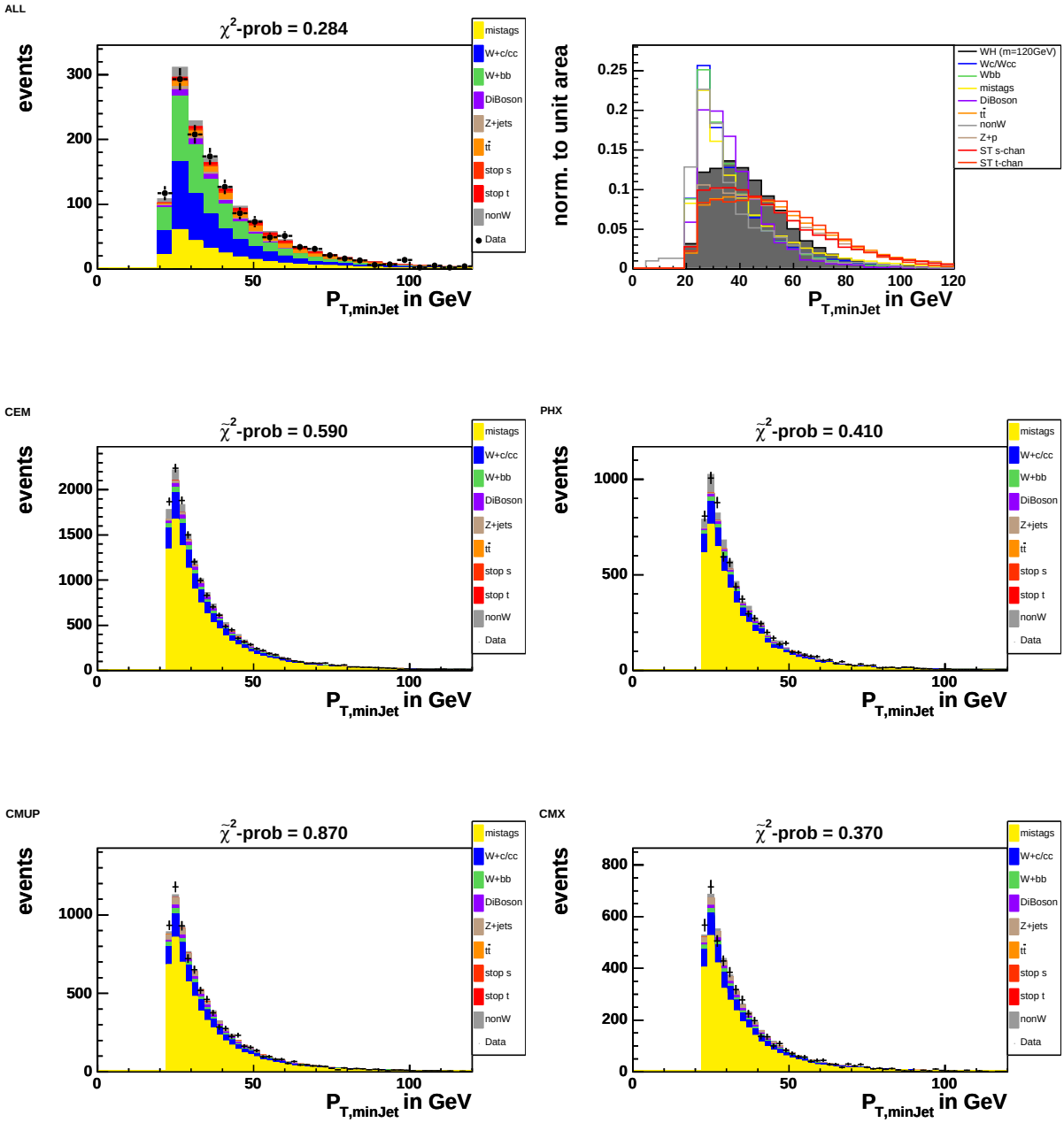

Figure 6.3: The transverse momentum of the jet with lowest energy

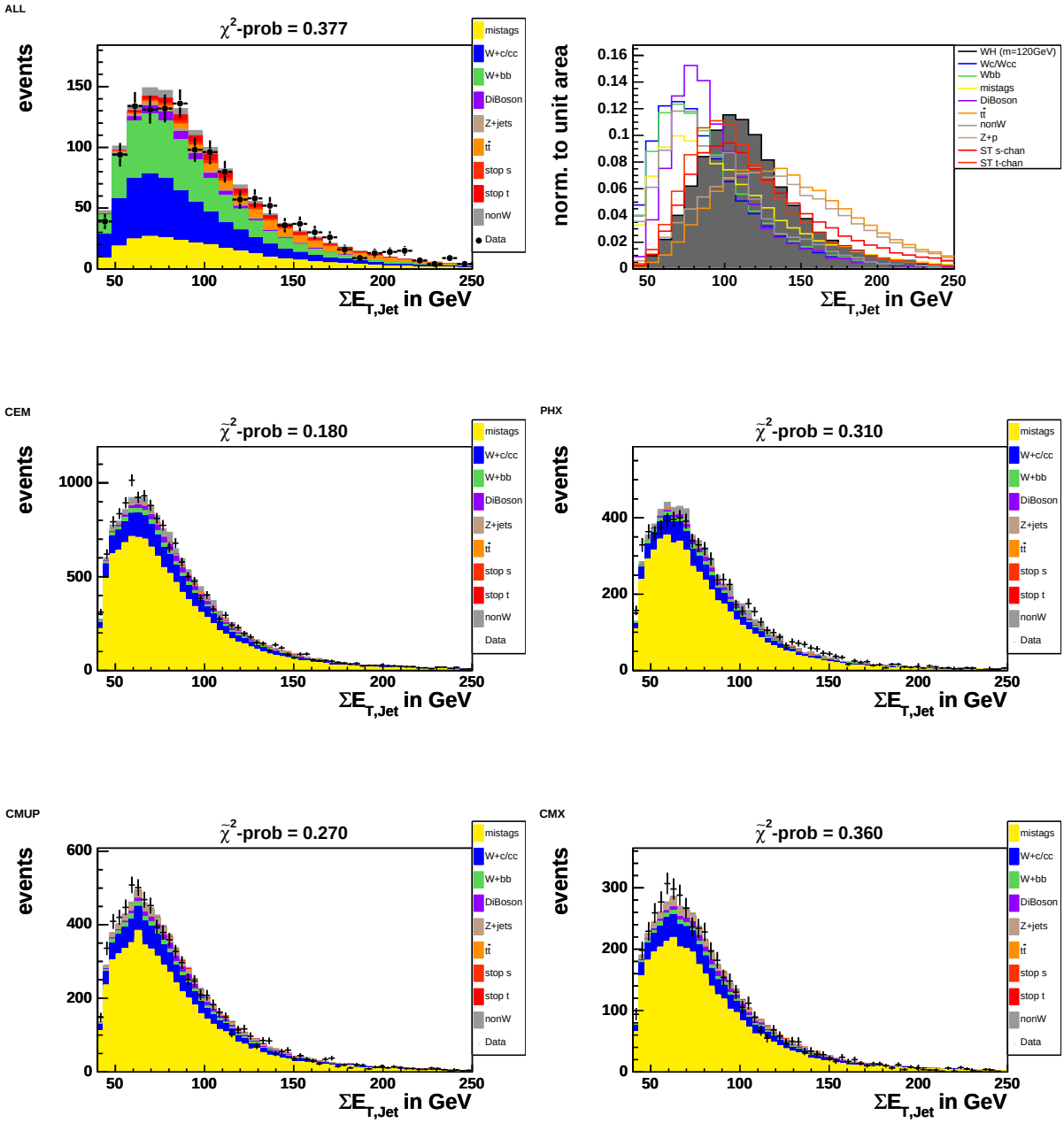

Figure 6.4: The scalar sum of all jets

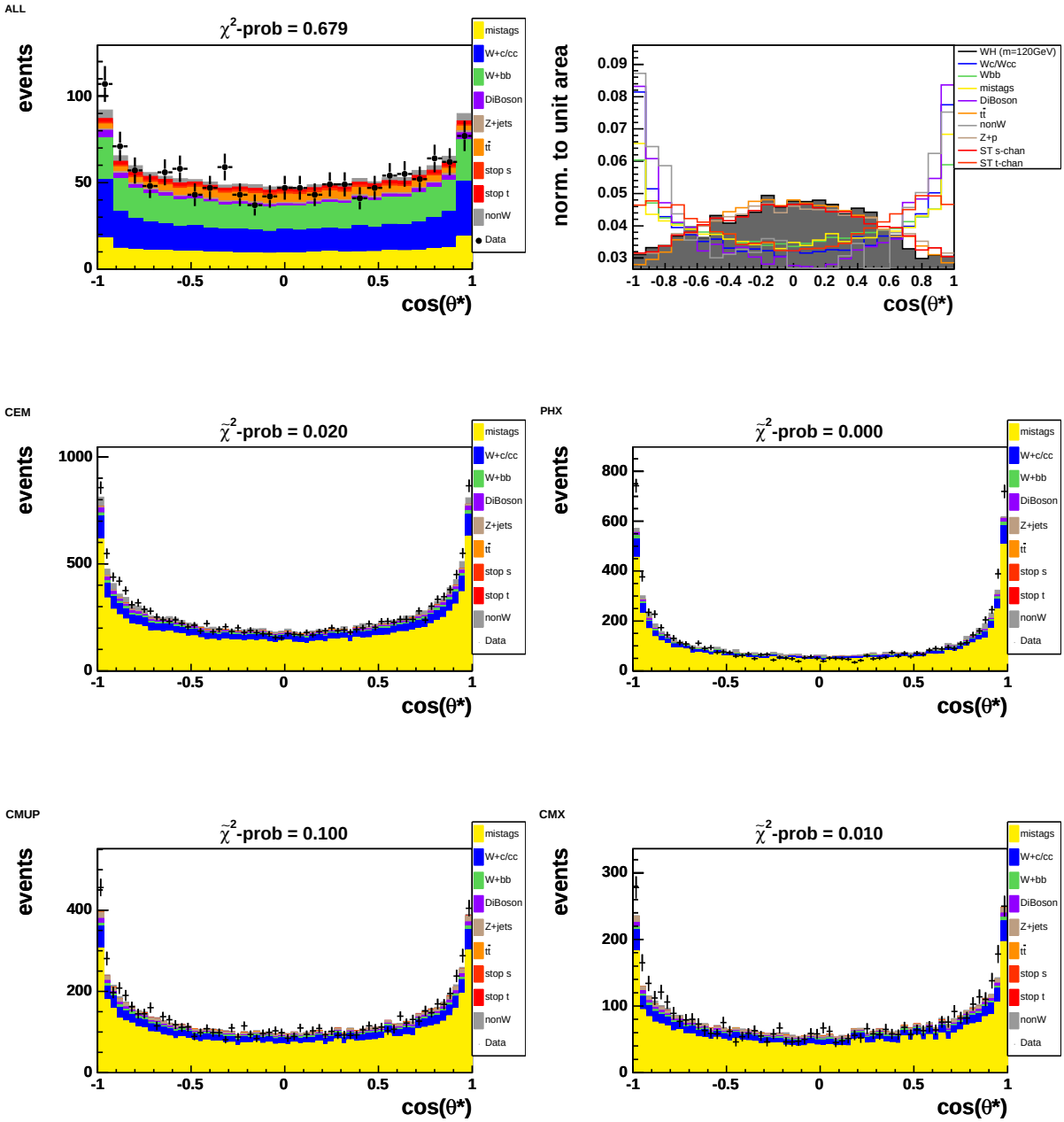

Figure 6.5: The azimuthal angle between Higgs and z-axis in the W^* rest frame

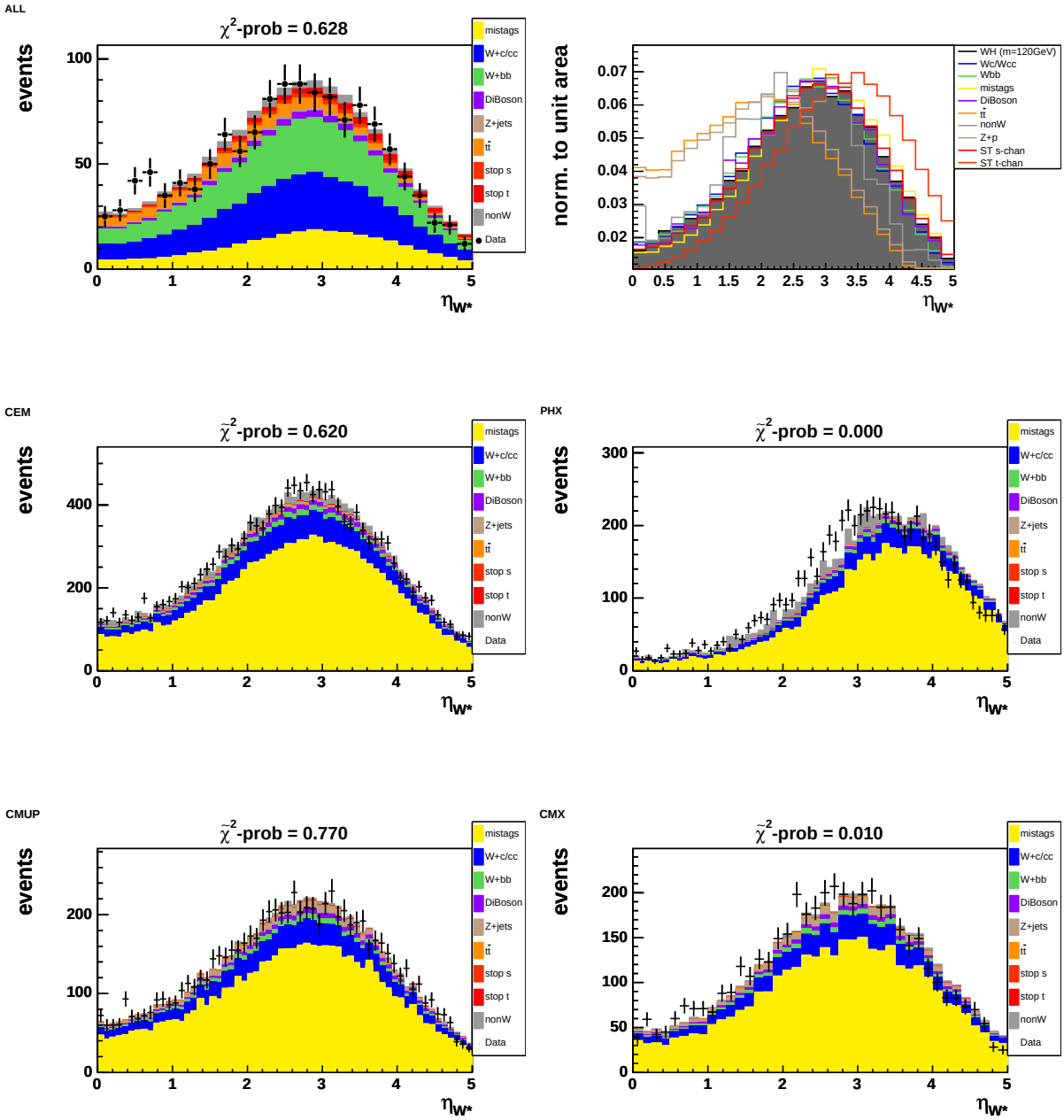

Figure 6.6: The pseudo rapidity of the reconstructed W^*

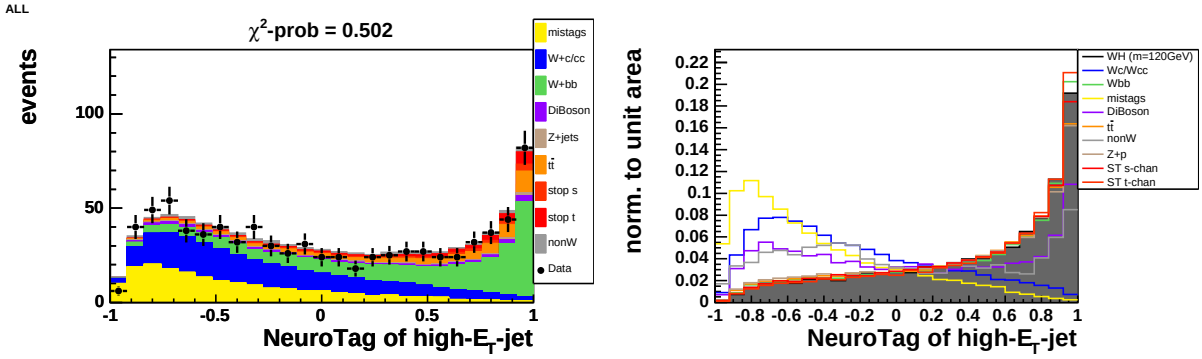

Figure 6.7: Output of the neural network flavor separator for the jet with highest E_T

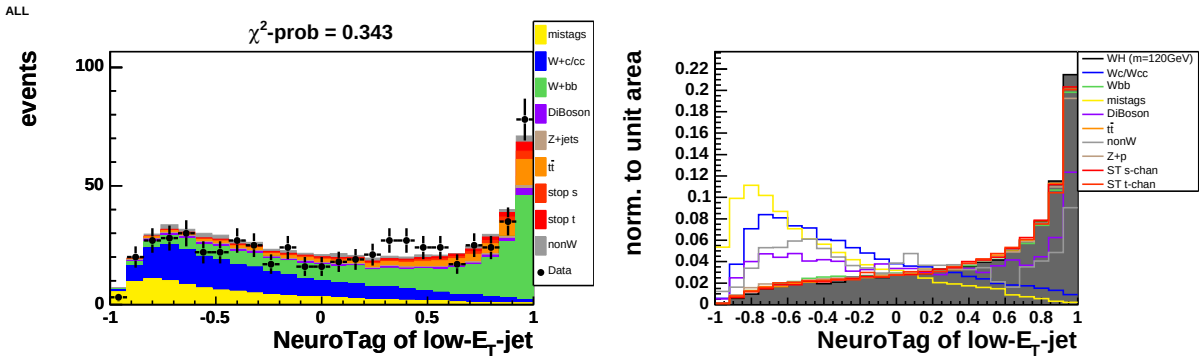

Figure 6.8: Output of the neural network flavor separator for the jet with lowest E_T

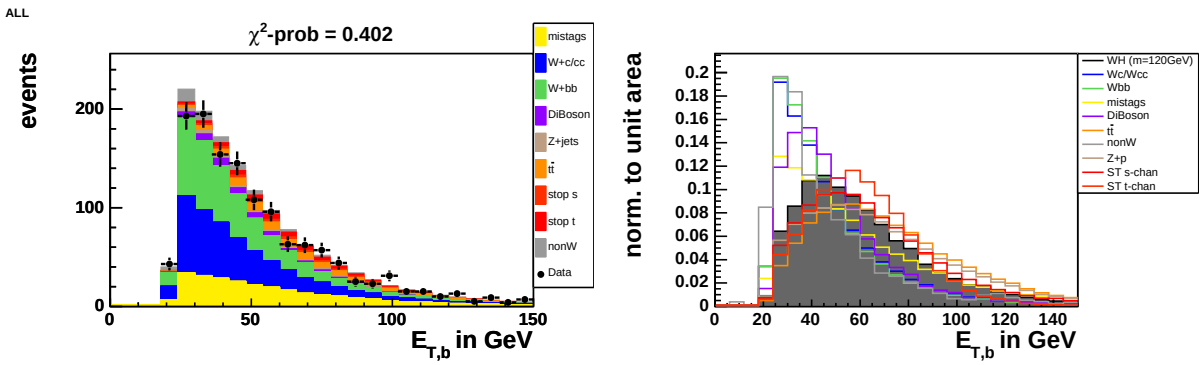

Figure 6.9: The transverse energy of the b tagged jet

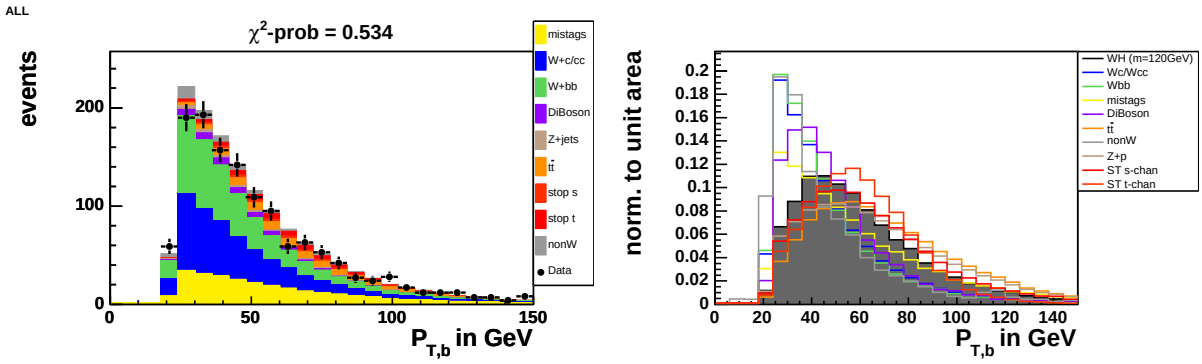

Figure 6.10: The transverse momentum of the b tagged jet

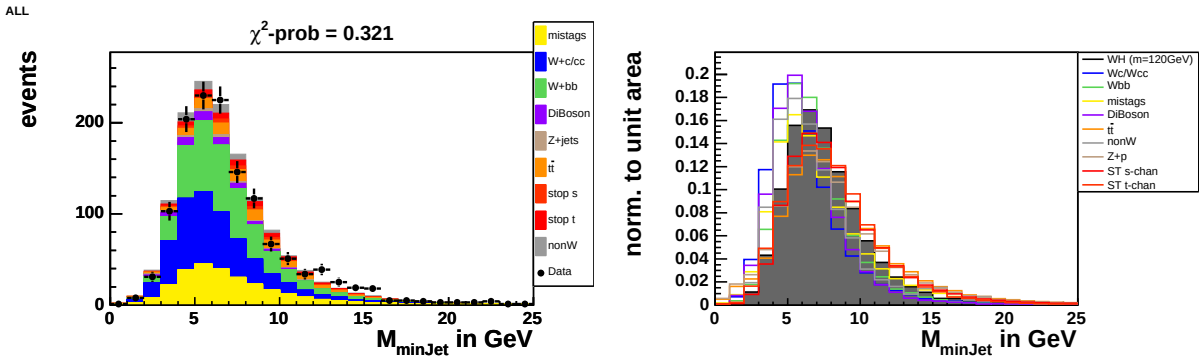

Figure 6.11: The mass of the jet with lowest energy

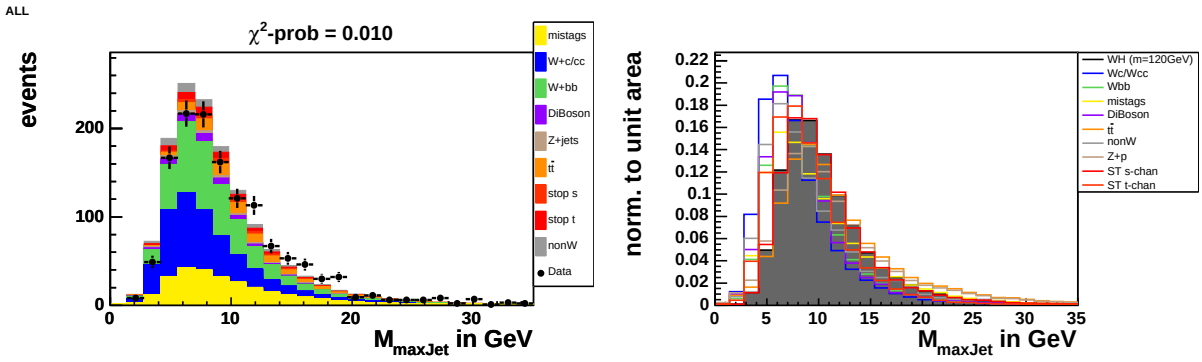

Figure 6.12: The mass of the jet with highest energy

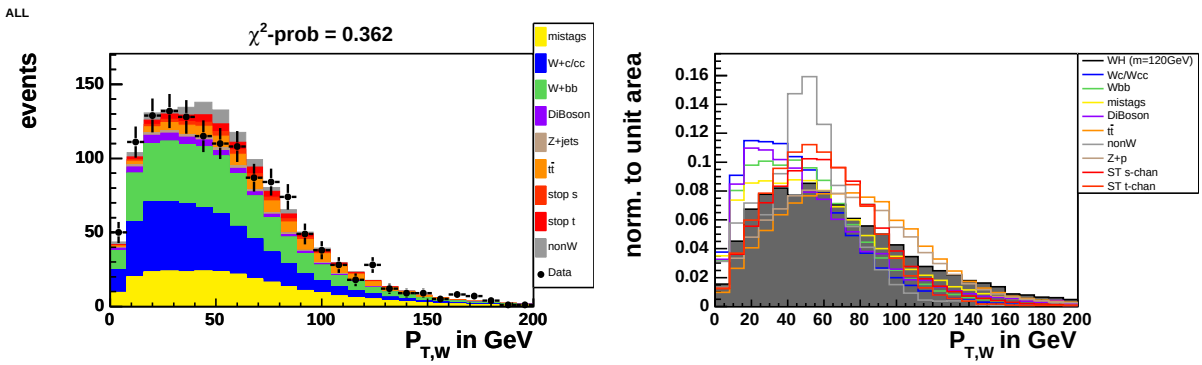

Figure 6.13: The transverse momentum of the reconstructed W boson

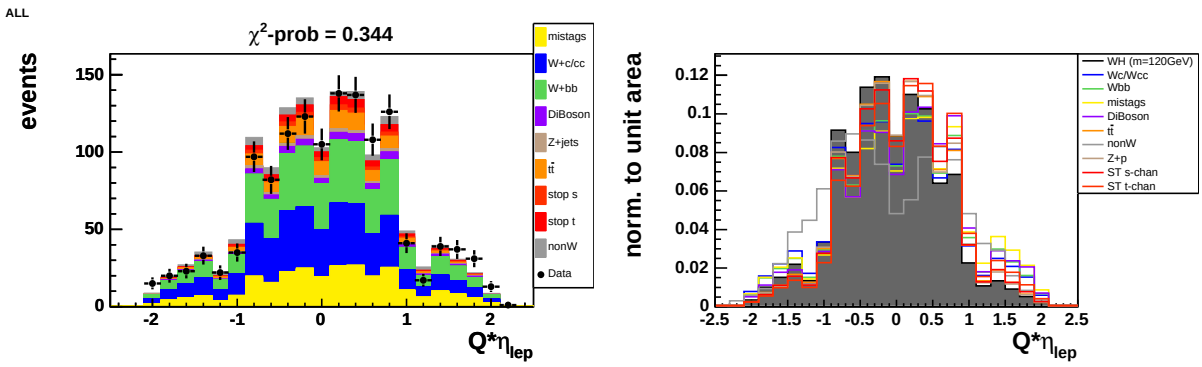

Figure 6.14: Product of charge and pseudo rapidity of the lepton

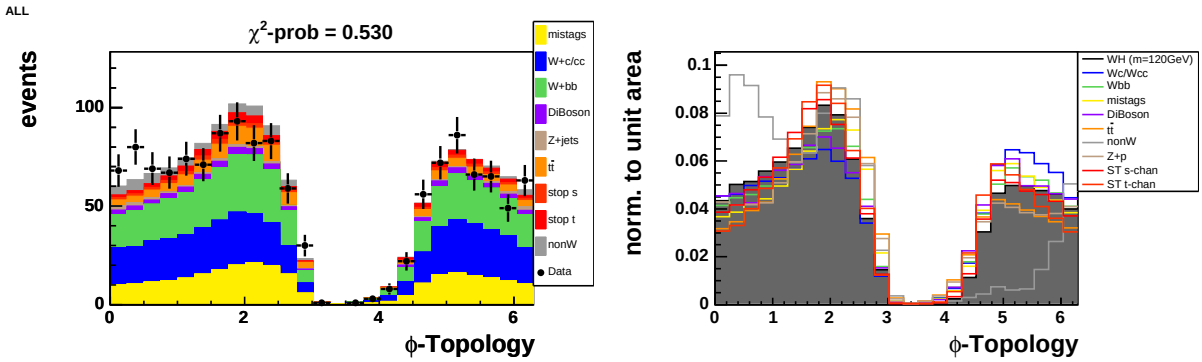

Figure 6.15: Topology of the event in the x-y-plane

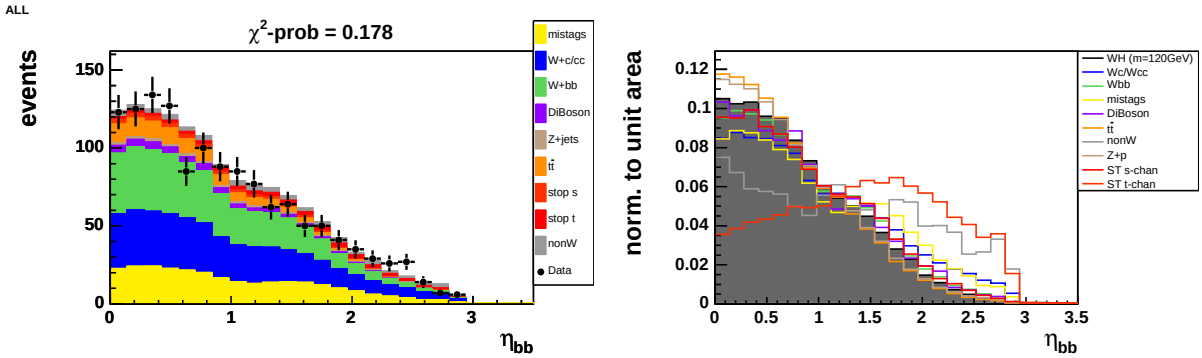

Figure 6.16: The pseudo rapidity of the 2nd b quark

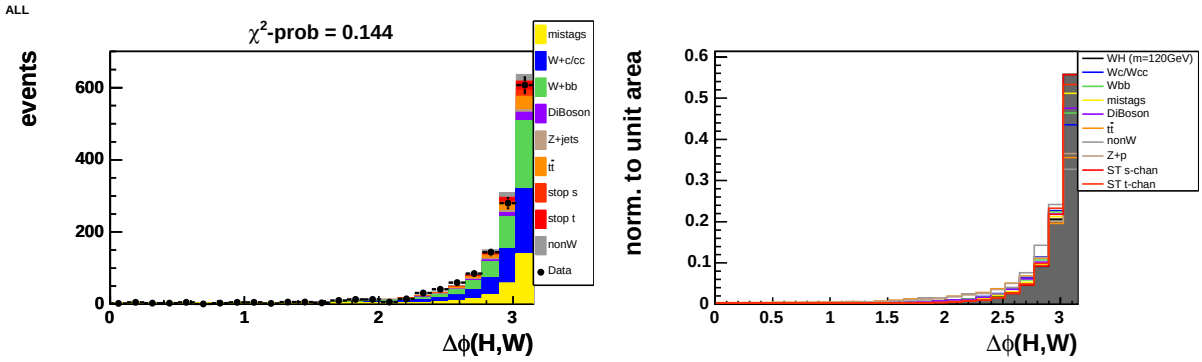

Figure 6.17: The azimuthal angle between the Higgs and the W boson

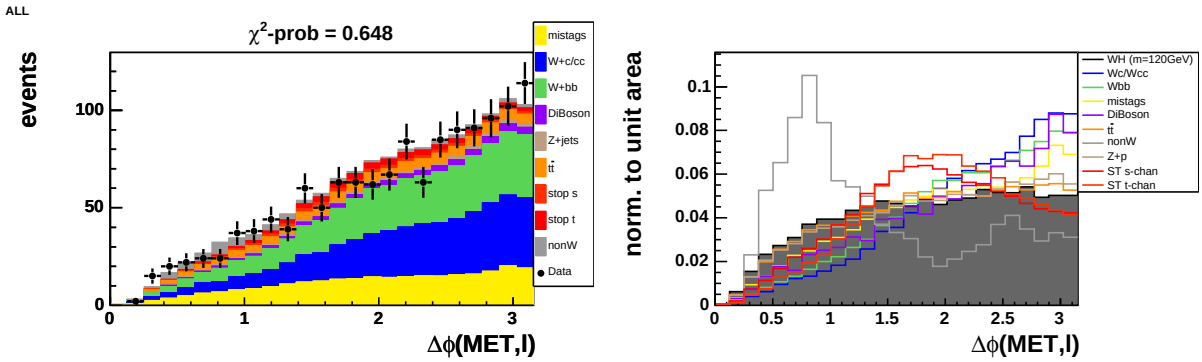

Figure 6.18: The azimuthal angle between missing transverse energy and lepton

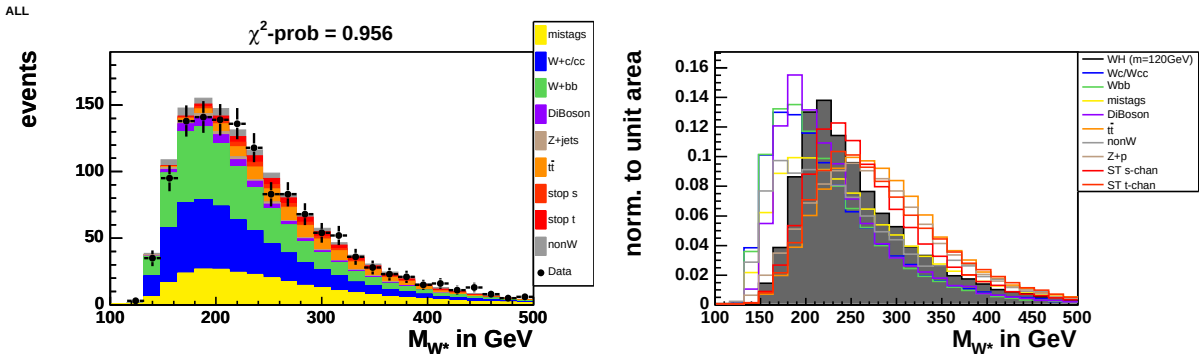

Figure 6.19: Mass of the W^*

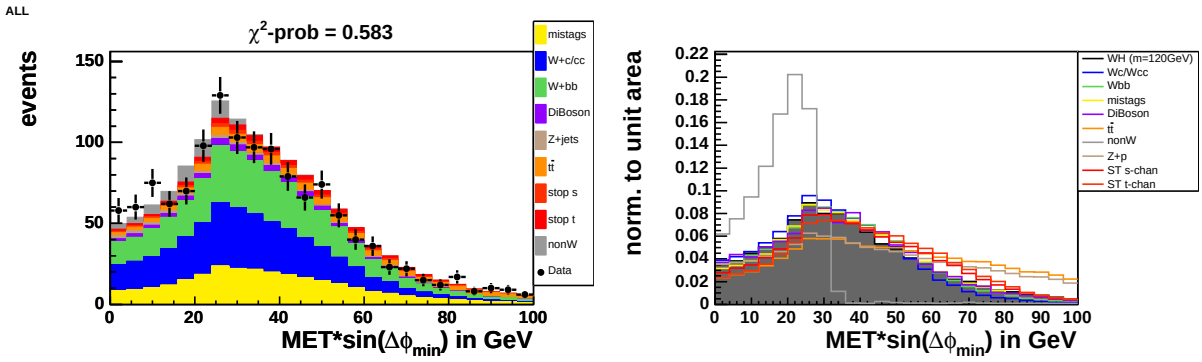

Figure 6.20: Product of $\not\!\!E_T$ and the angle between $\vec{\not\!\!E}_T$ and the closest object in the x-y-plane

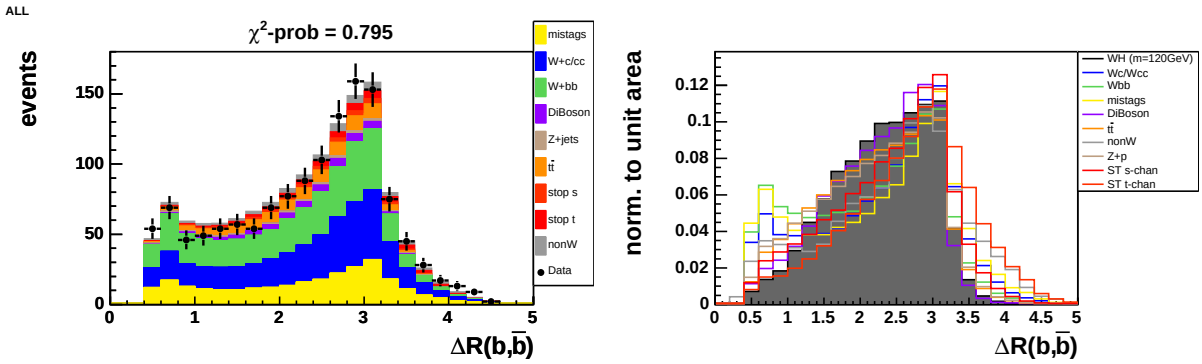

Figure 6.21: Distance in the ϕ-η-plane between the two quarks

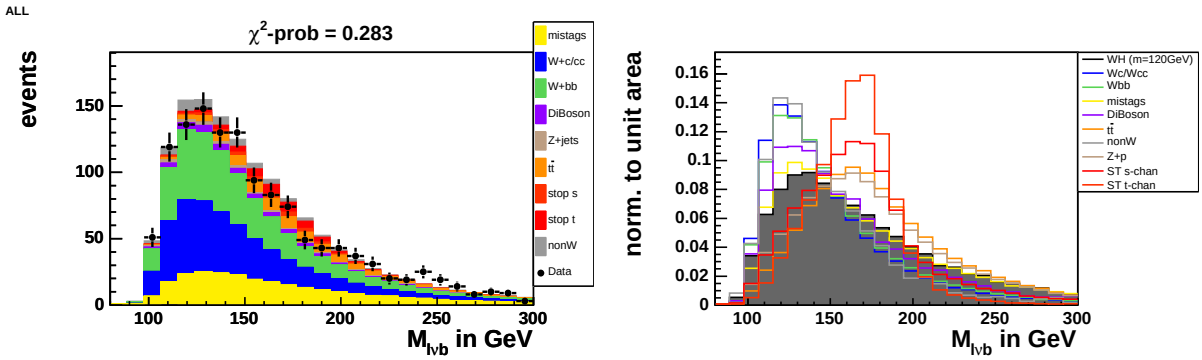

Figure 6.22: Invariant mass of the lepton-neutrino-b system

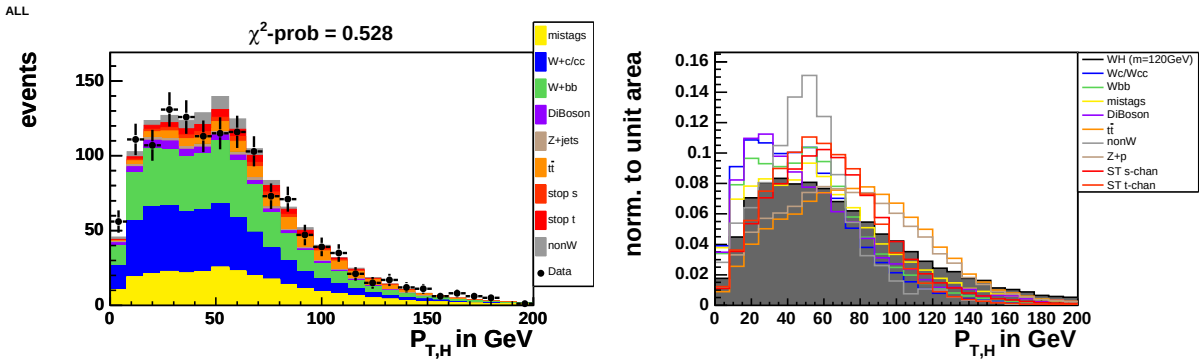

Figure 6.23: Transverse momentum of the Higgs boson

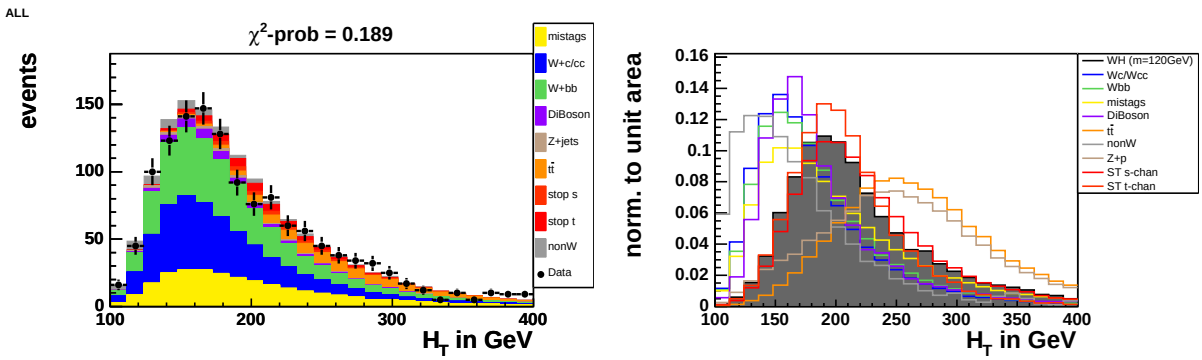

Figure 6.24: The scalar sum of the P_T of all jets, lepton, and neutrino

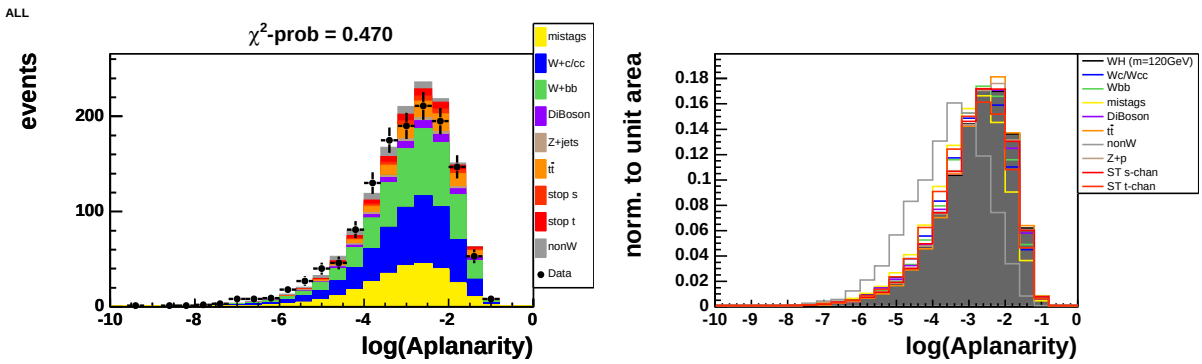

Figure 6.25: Logarithm of the aplanarity

The only variables not described by the background model are the distributions for $\cos(\theta^*)$ and η_{W^*} in the pretag sample for the PHX electrons. Here the W+jets Monte Carlo samples do not model the observed distribution correctly. Therefore, these two variables are set to a default value of -999 during the training, if an event with a PHX triggered electron passes the network. Hence for events with PHX electrons these variables are not taken into account for the network training.
The final used variables differ for each trained network. For the two networks trained to separate Higgs events with a mass of 120 GeV, the input variables selected by the automatic variable selection and their impacts on the final output are listed in tables 6.1 and 6.2.

ranking	variable	correlation to target in%	in σ
1	$M_{b\bar{b}}$	52.9	62.8
2	neuro-Tag 1	18.9	22.4
3	$P_{T,min}$	16.6	19.8
4	neuro-Tag 2	12.5	14.9
5	$\sum_{jets} E_{T,jet}$	7.6	9.0
6	$\eta_{2^{nd}b}$	6.8	8.1
7	$\Delta R_{b\bar{b}}$	4.7	5.6
8	$\cos(\theta^*)$	3.4	4.0
9	M_{max}	3.2	3.8
10	M_{min}	2.5	2.9
11	$P_{T,b}$	0.6	0.7
12	$E_{T,b}$	3.3	3.9
13	$\not{E}_T \sin(\Delta\phi_{min})$	3.3	3.9
14	$\log(A)$	2.8	3.4

Table 6.1: Input variables for the 1-tag network. The variable $cos(\theta^*)$ is not used for PHX events.

ranking	variable	correlation to target in%	in σ
1	$M_{b\bar{b}}$	61.3	41.0
2	$P_{T,min}$	7.8	5.2
3	$\eta_{2^{nd}b}$	7.5	5.0
4	$\not{E}_T \sin(\Delta\phi_{min})$	7.1	4.8
5	$\cos(\theta^*)$	6.0	4.0

Table 6.2: Input variables for the 2-tag network. The variable $cos(\theta^*)$ is not used for PHX events.

Studies have been done to improve the network output. For instance we applied different variations of the AdaBoost (adapted boost) algorithm [60], [61] to gain a better discrimination of signal and background events. Another method which has been tested is the use of a separated network trained on a sample including only

40% $t\bar{t}$, 10% single top and 50% Higgs events. The output of this network became an additional input variable of the final network. The motivation for this top-Higgs network is the fact that the $t\bar{t}$ process is one of the backgrounds, which is most difficult to separate. The gained improvement of these methods is minimal and not significant. The used NeuroBayes® network seems to extract nearly all possible information from the input variables.
The outputs of the neural networks for the Monte Carlo samples are shown in figure 6.26 for both trained networks for all four assumed Higgs masses. Since we want to use the output templates of the neural networks for a likelihood fit, we reduce the number of templates by adding templates similar in both, the neural network output shapes in both networks, as well as the fraction of one and two tag sample events. The remaining templates are composed as follows:

- single-top t-channel
- $Wb\bar{b}$
- c-like backgrounds: consists of Wc, $Wc\bar{c}$, mistags and QCD
- Diboson: WW, WZ and ZZ
- $t\bar{t}$-like: includes $t\bar{t}$, single-top s-channel and Z+jets.

From the templates for these combined channels the huge impact of the dijet mass $M_{b\bar{b}}$ can be seen. For a lower assumed Higgs boson mass the separation power between signal and diboson production becomes worse, since the maximum in the dijet mass distribution of the Higgs channel comes closer to the maximum of the diboson production, which is centred around the masses of W and Z bosons. For a higher assumed Higgs mass the separation between signal and most of the backgrounds becomes clearer. Only the top quark productions are more difficult to distinguish from the W-Higgs channel.

6.2 Likelihood Fit

From the neural network output templates we want to estimate the number of Higgs events in the observed data, respective the maximum number of Higgs events, which is still consistent with the CDF data, to gain an upper limit to the Higgs production cross section. Therefore, we apply a maximum likelihood fit to the outputs of the neural networks. The likelihood function is given by:

$$L(\beta_1, \ldots, \beta_C) = \prod_{h=1}^{H} \left(\prod_{k=1}^{B_h} \frac{e^{-\mu_{k,h}} \mu_{k,h}^{n_{k,h}}}{n_{k,h}!} \right) \prod_{j=2}^{C} G(\beta_j, 1.0, \Delta_j). \tag{6.2}$$

The fitted parameters are $\beta_j = \frac{\sigma_j}{\sigma_{SM,j}}$, the cross section σ_j of each physical process j normalized to the standard model prediction $\sigma_{SM,j}$. The number of processes taken into account for the fit is C.
The first term of this likelihood function consists a Poisson factor for each bin k

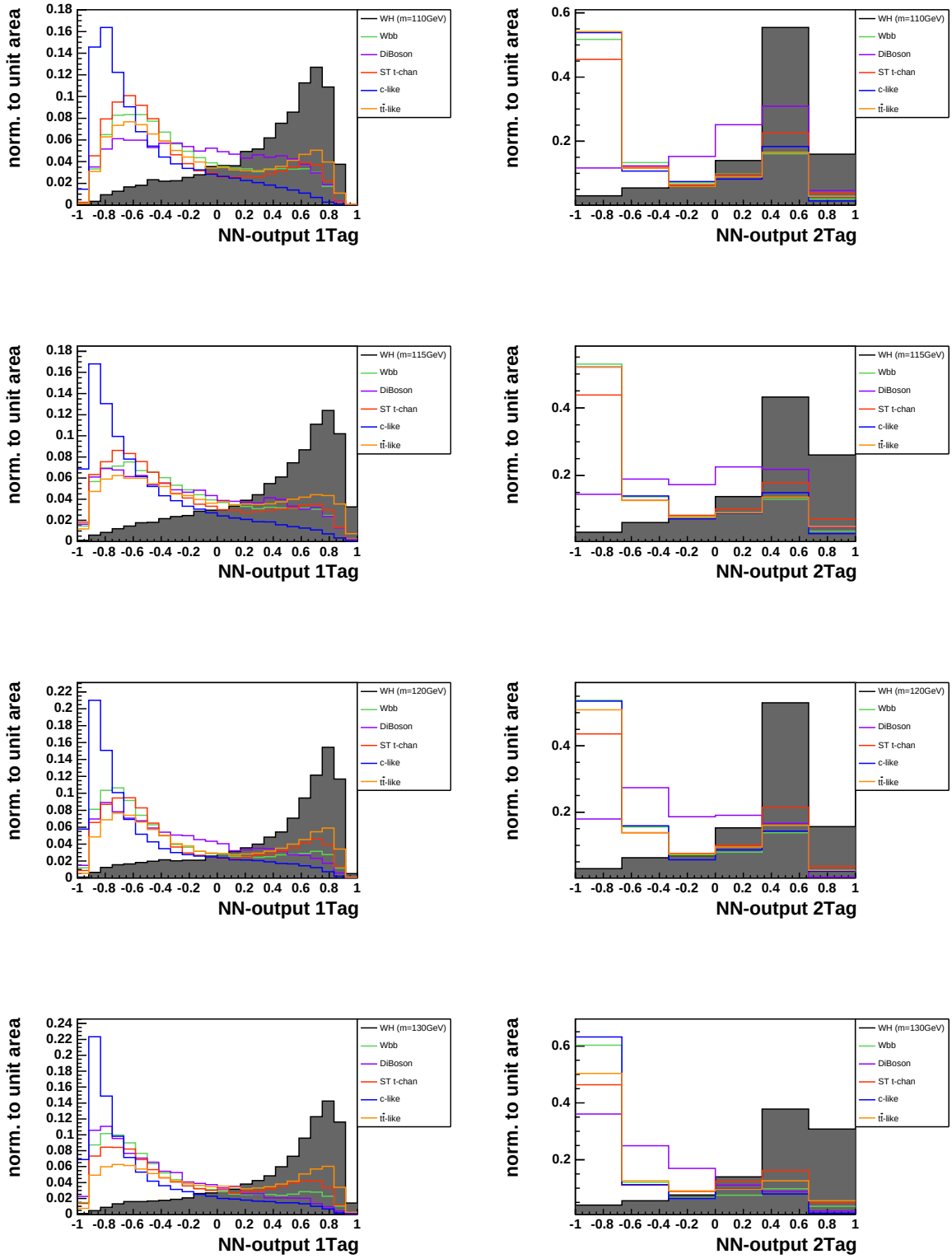

Figure 6.26: The neural network templates for the different processes used for the likelihood fit. On the left the 1-tag and on the right the 2-tag network outputs for m_h =110, 115, 120 and 130 GeV (from up to down) are shown.

in the histogram h, where $n_{k,h}$ is the number of observed data events in bin k and histogram h; B_h is the overall number of bins in histogram h, H the number of simultaneously fitted histograms (here $H = 2$ for the single and double tagged networks). The expectation value $\mu_{k,h}$ of events in the bin with index k in Histogram h is

$$\mu_{k,h} = \sum_{j=1}^{C} \beta_j \hat{\nu}_{j,h} \alpha_{jk,h}. \tag{6.3}$$

Here $\hat{\nu}_{j,h}$ is the predicted expectation value for the number of events of a certain process j in histogram h. $\alpha_{jk,h}$ denotes the fraction of events of the template histogram for process j located in bin k. Therefore, it yields:

$$\sum_{k=1}^{B_h} \alpha_{jk,h} = 1. \tag{6.4}$$

The second term in the likelihood function sets Gaussian constraints to the background rates. There is no Gaussian constraint included for the W-Higgs signal process ($j = 1$).

$$G(\beta, \mu, \Delta) = \frac{1}{\sqrt{2\pi\Delta^2}} e^{\frac{-(\beta-\mu)^2}{2\Delta^2}} \tag{6.5}$$

In this formula Δ denotes the expected relative error on the expected number of events.
For our analysis we do not handle every physical process shown in Plot 6.26 separately, because it is difficult for the likelihood fit to distinguish templates with similar shapes. Therefore, we collect all processes with similar template shapes and commit them to the likelihood fit. Finally we only fit the signal template and the five background templates listed above.

6.2.1 Systematic Uncertainties

There are several sources of systematic uncertainties, which affect the rate of expected events of a physical process as well as the shape of the neural network output. To take systematic uncertainties into account the likelihood function 6.2 has to be modified. Let there be S sources of systematic uncertainties, then the strength of the uncertainty i is denoted with δ_i, where i runs from 1 to S. δ_i is given in units of one standard derivation of the systematic effect. The δ_i become additional fit parameters of the likelihood function. Therefore, in a first step the likelihood function 6.2 is multiplied with a Gaussian factor for each systematic uncertainty. The new likelihood function is given by:

$$L(\beta_1, \ldots, \beta_C, \delta_1, \ldots, \delta_S) = \prod_{h=1}^{H} \left(\prod_{k=1}^{B_h} \frac{e^{-\mu_{k,h}} \mu_{k,h}^{n_{k,h}}}{n_{k,h}!} \right) \prod_{j=2}^{C} G(\beta_j, 1.0, \Delta_j) \prod_{i=1}^{S} G(\delta_i, 0.0, 1.0). \tag{6.6}$$

This factor only gives a Gaussian constraint on the new parameters. To incorporate the effect of the systematic uncertainties, we modify the expectation value $\mu_{k,h}$ (see eq. 6.3). The modified value is:

$$\mu_{k,h} = \sum_{j=1}^{C} \beta_j \hat{\nu}_{j,h} \left[\sum_{i=1}^{S} (1 + \delta_i \epsilon_{ji,h}) \right] \alpha_{jk,h} \left[1 + \sum_{i=1}^{S} (\delta_i \kappa_{jik,h}) \right]. \tag{6.7}$$

The first new factor in 6.7 includes the relative acceptance uncertainties $\epsilon_{ji,h}$. These systematic shift is independent of the individual bin k of histogram h; $\epsilon_{ji,h}$ only describes the overall change of the rate of process j affected by the systematic source i.
The second additional factor in 6.7 describes the shape uncertainties generated by systematic effects. Here $\kappa_{jik,h}$ indicates the relative shift induced by the uncertainty i in bin k of the template histogram h for the physics process j. The shape uncertainty is calculated from systematically shifted histograms:

$$\kappa_{jik,h} = \frac{\alpha^{+}_{jik,h} - \alpha^{-}_{jik,h}}{2\alpha_{jk,h}}. \tag{6.8}$$

In 6.8 $\alpha^{\pm}_{jik,h}$ denote the bin contents of bin k of these histograms shifted with a $+1\sigma$ and -1σ systematic effect. For some systematic effects only one histogram $\tilde{\alpha}_{jik,h}$ is calculated. In this case $\kappa_{jik,h}$ is calculated as follows:

$$\kappa_{jik,h} = \frac{\tilde{\alpha}_{jik,h} - \alpha_{jik,h}}{2\alpha_{jk,h}}. \tag{6.9}$$

The systematic shape uncertainties are constructed in the way that holds:

$$\sum_{k=1}^{B_h} \kappa_{jik,h} \cdot \alpha_{jk,h} = 0, \tag{6.10}$$

so systematic rate uncertainties are only included in $\epsilon_{ji,h}$. For our analysis several sources of systematic uncertainties are considered. The first uncertainty taken into account considers different models of the parton distribution function (PDF). The PDF used by default to generate the W-Higgs Monte Carlo sample is CTEQ5L. For every event generated with CTEQ5L and having a parton momentum fraction x and transfered momentum Q^2 the probability of finding such an event using another PDF is calculated. This probability might be larger or smaller than the standard probability using CTEQ5L. For the uncertainty shifts all events are re-weighted with respect to the standard probability. This leads to a different rate as well as a different shape of the neural network outputs. The PDFs considered as sources of systematic errors are MRST72, MRST75 (taken as a pair of $+1\sigma$ and -1σ), which considers different values of Λ_{QCD} , and the 20 pairs of eigenvectors of CTEQ6M. The neural network outputs of all PDF uncertainties are combined to one template by adding in quadrature the $\kappa_{jik,h}$ bin-wise for all uncertainties. For the rate uncertainty the event probability normalized to the standard probability is combined

into one distribution for each event and each modified PDF. The root mean square (RMS) of this distribution is defined as the averaged 1σ rate uncertainty. The PDF uncertainties are calculated for the signal sample and the $t\bar{t}$ background.
Another systematic uncertainty is applied due to the rate uncertainty of the initial and final state radiation (ISR, FSR). To consider these uncertainties, Monte Carlo events are produced changing the parameters for initial or final state radiation in Pythia with a factor of 2 respectively 0.5 for more or less radiation. We use the ISR and FSR uncertainties only as rate uncertainties on the signal, since these modified Monte Carlo samples only exist for the Higgs sample with $m_h = 120$ GeV. For the other Higgs masses we assume the rate uncertainties due to ISR and FSR to be the same as for the $m_h = 120$ GeV sample. The ISR and FSR rate uncertainties are also taken into account for single-top and $t\bar{t}$ backgrounds.
For all processes except the non-W model we take the uncertainties on the jet energy scale (JES) into account. The overall JES uncertainty is a combination of uncertainties occurring during the seven corrections (see chapter 4.2) applied to the measured jet energy. The systematic uncertainties mostly arise from the modeling of jets by the MC simulation and by the knowledge of the response of the detector to single particles. A detailed description of the JES uncertainties can be found in [42]. The JES uncertainties are considered as sources of shifts in the rate as well as in the shape of background and signal.
Further sources of systematic uncertainties are the uncertainty on the modeling of the mistags, the composition of the non-W flavors and the scale Q^2 for the $W + b\bar{b}$ samples. As a systematic shape uncertainty to the QCD background a different flavor composition of the non-W sample is assumed. For this sample we assign 60% of all jets to be b like, 30% c like and 10% to be light quark jets as recommended in [59]. The output of the neural network flavor separator is again drawn randomly. For the modeling of the mistag events we use events from Monte Carlo generators, which are only taggable. Since taggable events might not represent the real mistags in the data sample, we use an alternative way to model the mistags, which is treated as a systematic uncertainty. Therefore, all events from data without any b tag are taken. The output of the neural network flavor separator is again drawn randomly according to the light jet distribution. This sample is used as a shape uncertainty to our original mistag Monte Carlo model.
The $W + b\bar{b}$ samples are recreated with larger and smaller scale Q^2 in Alpgen. These two samples are treated as systematic shape uncertainties to the $W + b\bar{b}$ process.
The rate uncertainties affecting the signal process are listed in table 6.3 for an assumed Higgs mass of 120 GeV. For the background processes these errors are of the same order. Only the JES uncertainty of the $t\bar{t}$ process is with about 8 to 10% larger than for the other processes. The JES and PDF systematic shape uncertainties for the W-Higgs process are shown in figure 6.27. Always shown is the original distribution (black line) as well as the shifted histograms used to calculate the $\kappa_{jik,h}$ in equations 6.8 respectively 6.9. For the JES uncertainty there exist two alternative histograms for $\pm 1\sigma$ errors, the PDF uncertainty is treated with one averaged error histogram according to equation 6.9. For the PDF uncertainty the shifted

histogram is nearly identical to the default distribution, so the most important systematic shape uncertainty affecting the Higgs signal is the error to the jet energy scale. An overview of all considered shape uncertainties is listed in table 6.4.

source	1 b tag	2 b tags
JES +/-	-0.6/-0.3%	-1.4/1.0%
ISR less/more	2.6/-2.4%	5.5/-3.1%
FSR less/more	2.8/1.3%	3.6/-0.4%
PDF	±2.2%	±2.0%
luminosity	±6.0%	±6.0%

Table 6.3: Rate uncertainties on the W-Higgs process for m_h =120 GeV, separately for the samples with one and two b tags.

	JES	PDF	Q^2	non-W comp.	mistag model
W-Higgs	×	×			
single-top t-ch.	×				
$Wb\bar{b}$	×		×		
c-like	×			×	×
Diboson	×				
$t\bar{t}$-like	×	×			

Table 6.4: List of the considered shape uncertainties for the different fit templates.

6.2.2 Ensemble Tests

To determine the sensitivity of the likelihood fit we use ensemble tests based on pseudo experiments. The ensemble test is done twice:

1. including a W-Higgs contribution with the expected standard model cross section.

2. including no Higgs signal.

In each pseudo experiment we draw a random number N_j for every process j from a Poisson distribution with mean $\hat{\nu}_j$. The corresponding mean values are given in table 6.5. In the next step we draw N_j random numbers following the template distribution of each process shown in figure 6.26. These numbers are filled into a histogram which gives the resulting neural network output of one pseudo experiment. For each pseudo experiment the log-likelihood fit is applied and the cross section of W-Higgs is measured.
Figure 6.28 shows the results for the expected measurement of the Higgs cross section for both ensemble tests made with 30000 pseudo experiments each. For the

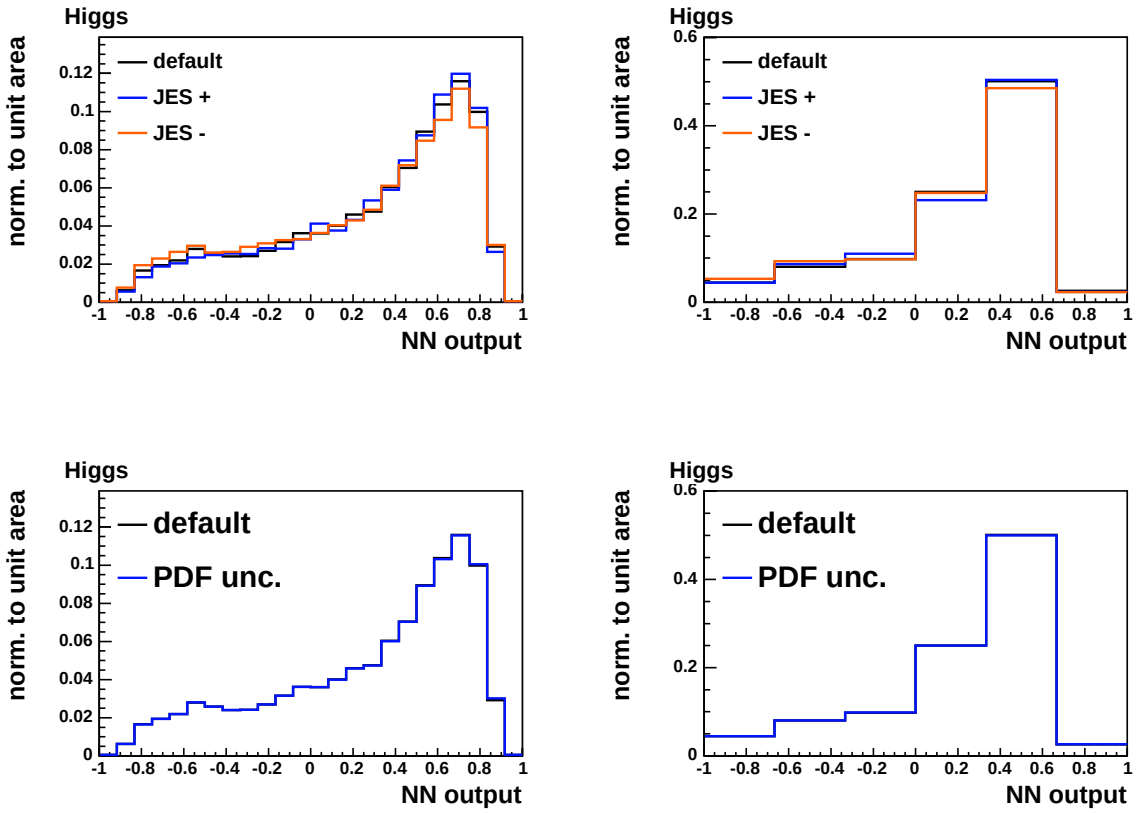

Figure 6.27: Template histograms for the systematic shape uncertainties of the W-Higgs process with m_h =120 GeV for the 1-tag (left) and 2-tag network (right)

j	process	$\hat{\nu}_{j,1}$	$\hat{\nu}_{j,2}$	Δ_j
1	W-Higgs	2.75	0.80	-
2	single-top t-ch.	45.4	1.3	15%
3	$Wb\bar{b}$	359.2	47.0	35%
4	c-like	734.7	7.6	21%
5	Diboson	48.5	3.0	8%
6	$t\bar{t}$-like	125.1	26.5	15%

Table 6.5: Expected number of events and the relative errors of the templates in the single and double tagged sample

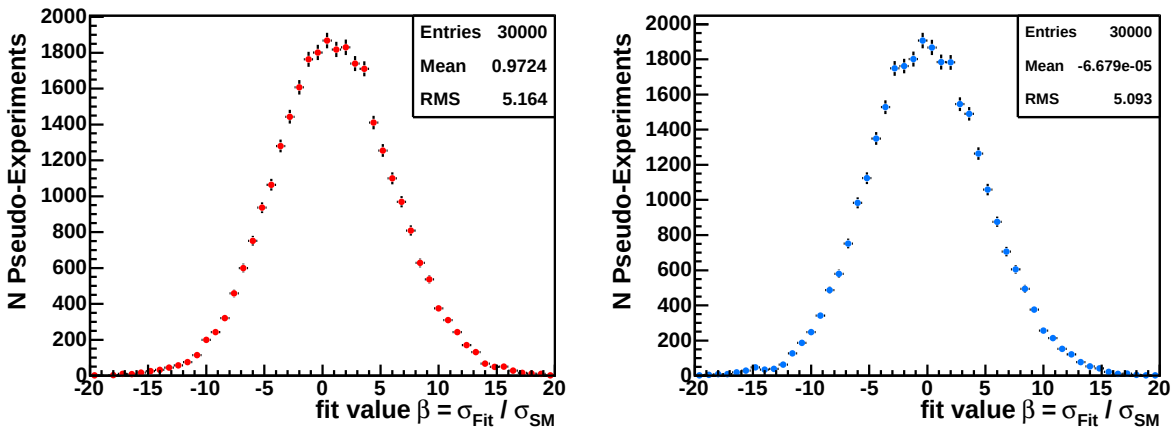

Figure 6.28: The left plot shows the expected β if we assume a Higgs with 120 GeV mass. The right plot shows the expected β under the assumption that there is no Higgs boson.

ensemble test with included Higgs we would expect to find a mean of 1. For the test without Higgs events a mean of 0 in the β distribution is expected. Due to the large width of both distributions it is nearly impossible to make a decision whether a measured β with observed data indicates the presence or absence of a Higgs signal.

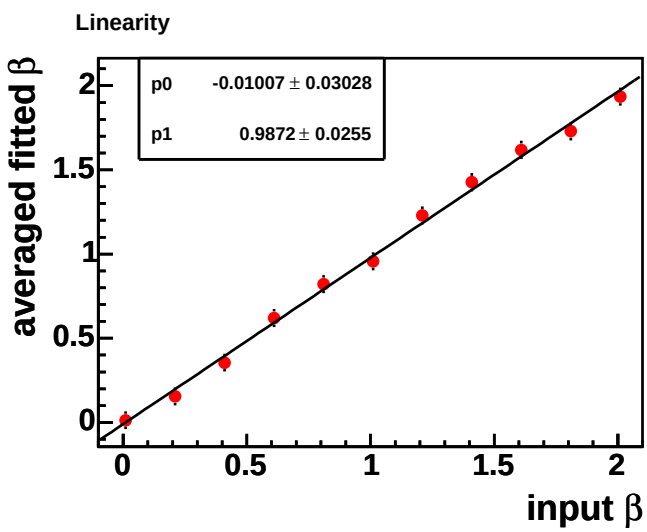

Figure 6.29: Linearity cross check: The averaged fitted $\beta = \frac{\sigma_{\text{Higgs}}}{\sigma_{SM}}$ of 10000 pseudo experiments is proportional to the input β, here shown for the fit to the $m_h = 120$ GeV network output.

As a cross check of our fit method we produce pseudo experiments including Higgs events at a cross section ranging from zero up to two times the Standard Model prediction, that means with an input β between 0 and 2. In figure 6.29 the averaged fitted β of 10000 pseudo experiments is plotted as a function of the assumed Higgs cross section. We see a linear dependence between the assumed and the fitted β.

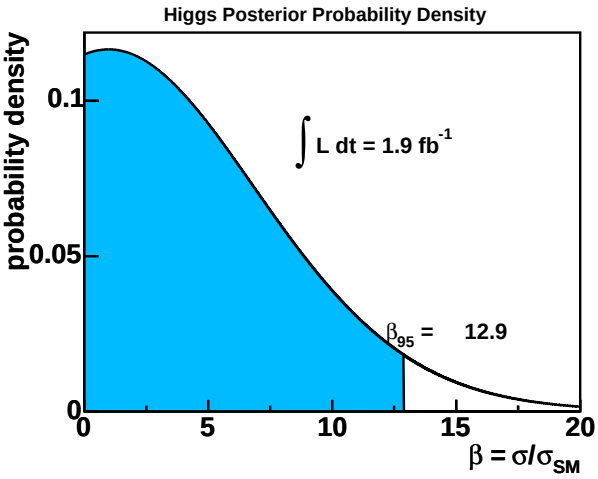

Figure 6.30: The likelihood function for one single pseudo experiment, the fitted β is 0.98, the resulting upper limit is $\beta_{95} = 12.9$.

To predict an upper limit on the W-Higgs cross section we integrate the normalized likelihood function from $\beta_{Higgs} = 0$ (we assume that the Higgs cross section cannot be negative) up to the point, where the integral reaches 0.95. This point gives us the upper limit on the Higgs cross section with 95% C.L.. The integration is performed using a profiling method. In the profiling method β_{Higgs} is increased in small steps. In every step the full likelihood fit is applied fixing β_{Higgs} and fitting the other parameters of the backgrounds and the systematics such, that the logarithm of the likelihood function is maximal in each point of the integration. Figure 6.30 illustrates this integral method for one pseudo experiment.
The integral of the likelihood function is calculated for an ensemble test consisting of 10000 pseudo experiments. This gives us an estimate on the expected upper limit for the β of the Higgs process. Again the ensemble test is made with pseudo experiments with and without included Higgs events.
In figure 6.31 the distribution of the upper limit β_{95} for 10000 pseudo experiments is shown. We define the mean of this distribution as the expected upper limit of our analysis, the uncertainty of this expectation is given by the RMS. The expected upper limit from the ensemble test made with included Higgs events is $\beta_{95} = 12.3 \pm 3.3$; if there is no Higgs in the expected mass range, we expect an upper limit of $\beta_{95} = 11.7 \pm 3.2$.

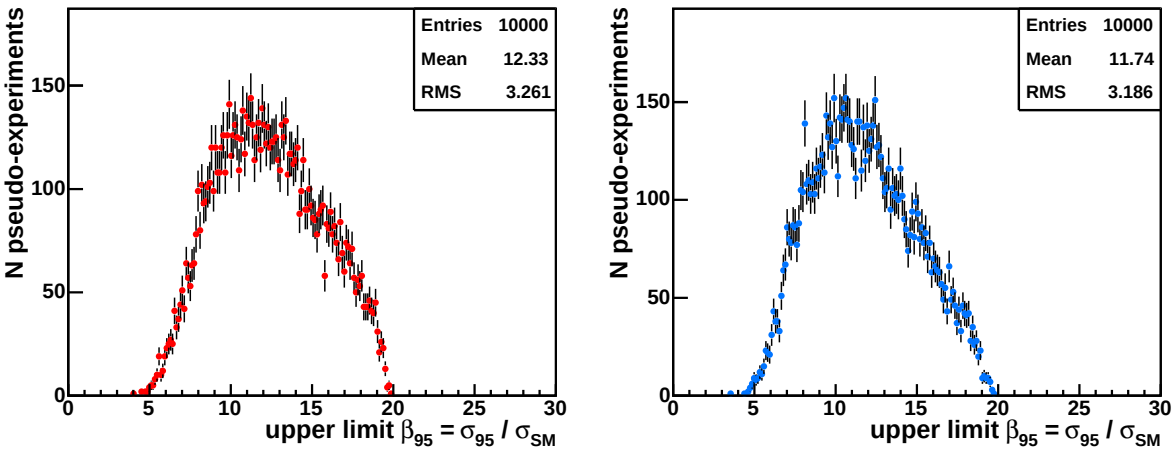

Figure 6.31: The left plot shows the expected 2σ limit if we assume a Higgs with 120 GeV mass. The plot on the right hand side shows the expected limit under the assumption that there is no Higgs boson.

6.3 Results of the Likelihood Fit

Fitting the Monte Carlo templates of the neural network output to the observed data yields $\beta = 1.8 \pm 5.2$ for the $m_h = 120$ GeV sample. This result corresponds to a cross section times branching ration for $H \to b\bar{b}$ of $\sigma = 0.19 \pm 0.54$ pb for the W-Higgs process (see also table 4.4). The fitted signal and combined background samples can be seen in figure 6.32.

M_{Higgs}	expected limit		observed	observed limit	fitted β	χ^2/dof
	with Higgs	no Higgs	limit	$\sigma \cdot BR(H \to b\bar{b})$		
110 GeV	9.9±3.1	9.0±2.8	7.8	1.3 pb	-0.6±3.7	1.18
115 GeV	11.0±3.3	10.3±3.1	11.5	1.5 pb	2.6±4.4	0.83
120 GeV	12.3±3.3	11.7±3.2	12.4	1.3 pb	1.8±5.2	0.95
130 GeV	17.3±4.8	17.4±4.8	21.2	1.3 pb	5.1±8.3	0.84

Table 6.6: Expected and observed limits to the W-Higgs process with $H \to b\bar{b}$. The upper limits are given in terms of $\beta = \frac{\sigma}{\sigma_{SM}}$ and as absolute limits to the cross section.

The analysis has been accomplished for various Higgs masses from 110 to 130 GeV. The discrimination power of the neural networks arises with the mass of the Higgs boson, since in the most important input variable, the di-jet mass, the Higgs events cumulate further away from W and Z like backgrounds. Since the cross sections (see table 4.4) decreases with increasing Higgs masses, we find the highest sensitivity for low masses. The upper limits to the cross section $\sigma \cdot BR(H \to b\bar{b})$ ranges from

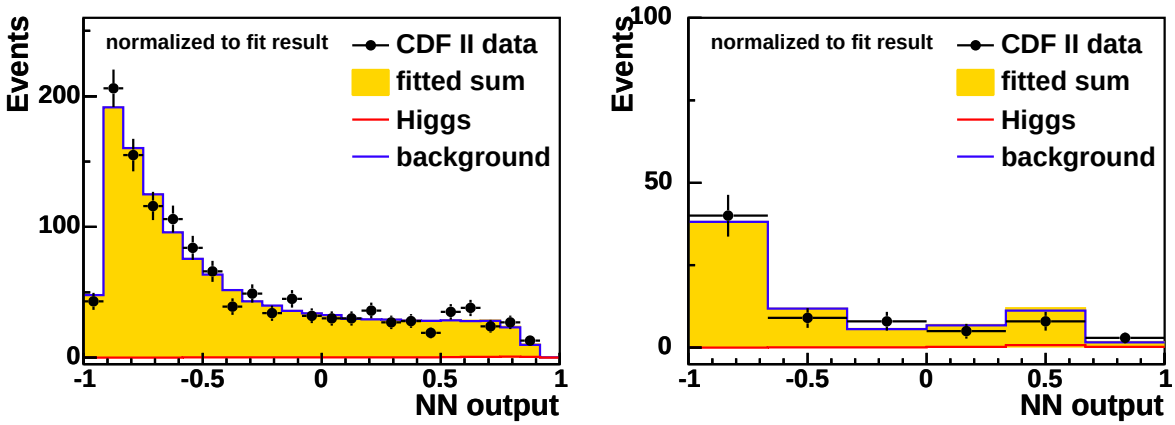

Figure 6.32: The final result of the combined fit in the single-tag (left) and double-tagged sample (right): for the W-Higgs process with $m_h = 120$ GeV we measure $\beta = \frac{\sigma}{\sigma_{SM}} = 1.8 \pm 5.2$ with an upper limit of $\beta_{95} = 12.4$. The χ^2/dof of the fit is 0.95, this corresponds to a probability of 0.54.

1.3 pb to 1.5 pb for the different masses. The results are shown in figure 6.33 and summarized in table 6.6.
All fitted limits are compatible with the expected limits for both assumptions. A significant observation of a Higgs signal is neither observed nor expected in the analyzed data sample.

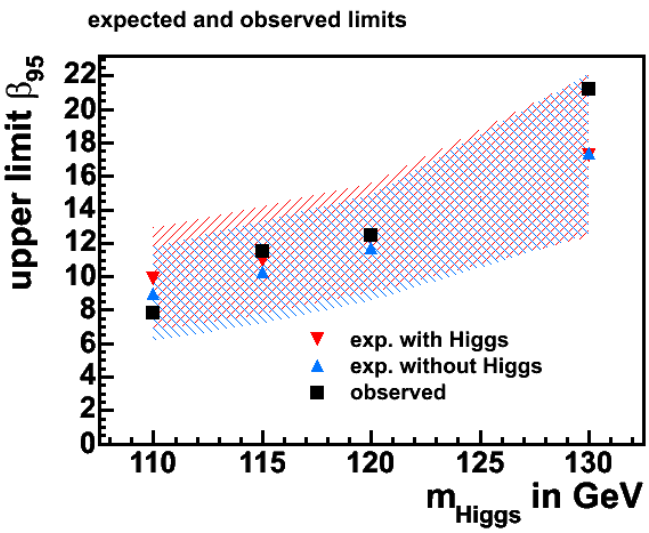

Figure 6.33: The expected and observed upper limits β_{95} for different Higgs masses. The dashed bands indicates the $\pm 1\sigma$ errors.

Chapter 7

Conclusions

7.1 Summary

In the presented analysis a search for a Standard Model like Higgs boson in the WH channel using the CDF II data corresponding to 1.9 fb^{-1} is performed. Since the cross sections of most of the background processes are much larger than the cross section of the Higgs production, we try to improve the sensitivity in several steps. In a first step a cut-based analysis is performed. Especially for the reduction of the QCD background several cuts are applied. These cuts have been developed to remain with as many as possible signal events and to reduce simultaneously the backgrounds. The cut-based part of this analysis leads to an adequate description of the remaining data set by the Monte Carlo models and especially the data based QCD model.
For a further improvement of the sensitivity we use neural networks. A separation of the selected sample into two subsamples with one and two SecVtx b tags and separate network trainings on these sub-samples is applied. The network training is also performed for different assumed Higgs boson masses between 110 and 130 GeV/c^2. For the training several input variables are used. The variables with the largest discrimination power are the reconstructed mass of the Higgs boson and the output of the neural network flavor separator, which distinguishes b jets from light quark jets. The expected sensitivity of the analysis is deduced from a likelihood fit applied to the network output for pseudo experiments. The expected upper limits to the Higgs cross section are still much larger than the theoretically predicted cross section. So the observed upper limits of 1.3 to 1.5 pb for the different assumed masses are compatible with the expected limits, but also these limits are still one order of magnitude larger than the predicted cross sections, which are of the order of 0.1 pb.
Therefore, with this analysis we are unable to decide, whether the Higgs mechanism is realized in nature or not. In the considered mass region for the Higgs boson no significant excess or deficiency is observed. A significant excess in the CDF II data in this analysis would not be a confirmation of the predicted Standard Model Higgs process, since such an excess would correspond to a process with a signature like the W-Higgs production, but with a cross section at least ten times larger than the

predicted Standard Model rate. Therefore, the non-observation of any signal excess can be interpreted as a validation of the Standard Model in the accessible region.

7.2 Results of Recent Analyses

There exist a couple of further analyses based on data recorded at the Tevatron. The results of the individual searches in different channels are combined to an overall upper limit to the Higgs cross sections. For the CDF analyses a combination of all individual Higgs searches is given in [62]. Besides searches for $WH \rightarrow l\nu b\bar{b}$, the channel considered in our analysis, further dominant channels are $ZH \rightarrow l^+l^-b\bar{b}$ and $ZH \rightarrow \nu\bar{\nu}b\bar{b}$ for the lower Higgs boson mass region and $gg \rightarrow H \rightarrow W^+W^- \rightarrow l^+l^-\nu\bar{\nu}$, which is the most significant channel for Higgs searches, if the Higgs mass exceeds twice the mass of the W boson. The combined results from these channels, which has been analysed with data corresponding to an integrated luminosity between 1.1 fb^{-1} and 1.9 fb^{-1}, are presented in figure 7.1. The results of the resent W-Higgs search, which also uses neural networks for signal and background discrimination, are compatible with the upper limits gained in our analysis.

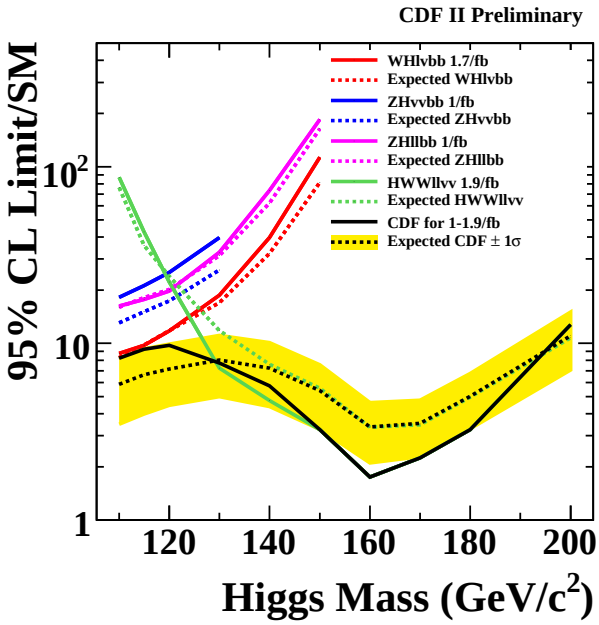

Figure 7.1: The combined upper limit as a function of Higgs boson mass between 110 and 200 GeV/c^2 for all CDF analyses.

The observed (expected) limits are factors of 9.2 (6.6) for a Higgs boson mass of 115 GeV respectively 1.8 (3.4) for 160 GeV higher than cross sections predicted by the Standard Model. Combining the results of the CDF measurements with the analyses made with the DØ experiment leads to the limits shown in figure 7.2.

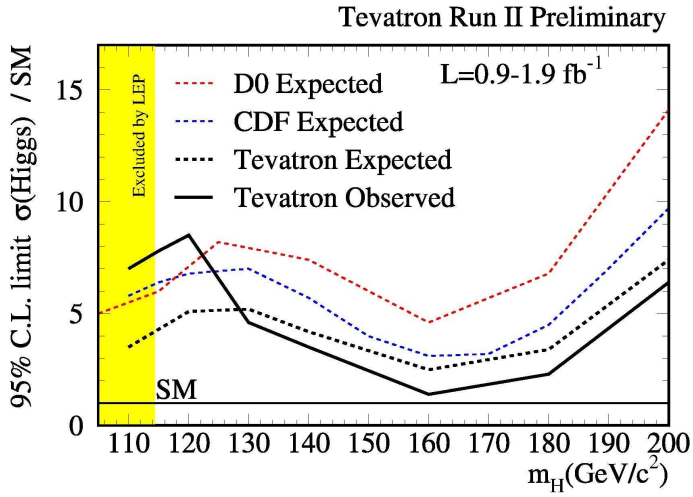

Figure 7.2: The combined upper limit from CDF and DØ as a function of the Higgs boson mass.

7.3 Outlook

The presented analysis gives upper limits to the cross section of the associated Higgs production $WH \rightarrow l\nu b\bar{b}$ for an assumed Higgs mass between 110 GeV and 130 GeV. These limits are a factor of about ten higher than the theoretical predicted cross section. The separate look at this Higgs production process will not lead to a significant observation or exclusion of the Higgs boson by increasing the analyzed data, since the Tevatron is expected to be switched off until 2009 or 2010. The conclusion from our analysis, we can give, is the confirmation of the Standard Model in the reachable parameter region. A Higgs boson with much larger couplings than predicted by the Standard Model can be excluded.
For a Higgs boson mass of about 160 GeV the combination of all Higgs searches might reach a significance, which allows the exclusion or the detection of the Higgs particle at the Tevatron, if the amount of available data is increased in the next few years. But since the measurements of the top quark mass as well as the mass of the W boson favor a lower Higgs mass, a detection of a Standard Model Higgs in this region seems to be not very likely.
If the Higgs boson has a lower mass, which is preferred from the electroweak fits, the detection becomes more difficult. The observed limits of all combined analyses of CDF and DØ lie still about more than five times above a significant detection. Although the data, which can be analyzed, will increase in the next years with a factor of perhaps two to three with respect to todays data, a significant observation seems to be nearly impossible. A final answer to the question, if the Higgs mechanism is realized, will presumably be given first by the experiments at the Large Hadron Collider at CERN.

Bibliography

[1] Peter W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. 13, 508 (1964)

[2] Peter W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. 145, 1156 (1966)

[3] Stephen Weinberg, *A Modell of Leptons*, Phys. Rev. Lett. 19, 1264 (1967)

[4] Chen-Ning Yang, Robert L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. 96, 191 (1954)

[5] Francis Halzen, Alan D. Martin, *Quarks and Leptons: Introductory Course in Modern Particle Physics*, Wiley & Sons, 1984

[6] Michael E. Peskin, Daniel V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books, 1995

[7] Ulrich Nierste, Fred-Markus Stober, *Vorlesung zur Theoretischen Teilchenphysik I-III*, Universität Karlsruhe, WS 2005/06, SS 2006, WS 2006/07

[8] W-M Yao et al., *J. Phys. G: Nucl. Part. Phys.*, 33 1

[9] The LEP Electroweak Working Group, http://lepewwg.web.cern.ch

[10] A.Djouadi, J.Kalinowski and M.Spira, *HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension*, hep-ph/9704448

[11] M. Spira, *Higgs boson production and decay at the Tevatron*, hep-ph/9810289

[12] T. Han and S. Willenbrock, *QCD correction to the* $pp \to WH$ *and* ZH *total cross sections*, Phys. Lett. B273 (1991) 167

[13] Yoshiaki Kusakabe, *Search for Higgs Boson Production in Proton-Antiproton Collisions at* $\sqrt{s} = 1.96$ *TeV*, CDF Internal Note 8714

[14] Peter H. Garbincius, *Tevatron Collider Operations and Plans*, hep-ex/0406013

[15] CDF II Collaboration, *The CDF-II detector: Technical design report*, FERMILAB-PUB-96/390-E

[16] The CDF Collaboration, *Time of Flight System for CDF*, CDF Internal Note 2573

[17] C.S. Hill et al., *Initial Experience with the CDF Layer 00 Silicon Detector*, Nucl. Instrum. Meth. A 511, 118 (2003)

[18] A. Sill et al., *CDF Run II Silicon Tracking Projects*, Nucl. Instrum. Meth. A 447,1 (2000)

[19] T. Affolder et al., *Intermediate silicon layers detector for the CDF experiment*, Nucl. Instrum. Meth. A 453, 84 (2000)

[20] A. Affolder et al., *CDF Central Outer Tracker*, Nucl. Instrum. Meth. A 526, 249 (2004)

[21] L. Balka et al., *The CDF central electromagnetic calorimeter*, Nucl. Instrum. Meth. A 267, 272 (1988)

[22] S. Bertolucci et al., *The CDF central and endwall hadron calorimeter*, Nucl. Instrum. Meth. A 267, 301 (1988)

[23] M.G. Albrow et al., *The CDF plug upgrade electromagnetic calorimeter: test beam results*, Nucl. Instrum. Meth. A 480, 524 (2002)

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

[25] The Trigger and Datasets Working Group, *Run-II Trigger Table and Datasets Plan*, CDF Internal Note 4718

[26] The Trigger Working Group, *What Data can the Trigger Cut On?*, http://www-cdf.fnal.gov/internal/upgrades/daq_trig/twg/tools/trigopts.ps

[27] B. Ashmanskas et al., *The CDF Silicon Vertex Trigger*, Nucl. Instrum. Meth. A 518, 532 (2004)

[28] G. Gomez-Ceballos et al., *Event Builder and Level 3 at the CDF experiment*, Nucl. Instrum. Meth. A 518, 522 (2004)

[29] M.L. Mangano, F. Piccinini, A.D. Polosa, M. Moretti, R. Pittau, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, hep-ph/0206293

[30] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour, B.R. Webber, *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons*, hep-ph/0011363

[31] T. Sjöstrand, P. Eden, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna, E. Norrbin, *High-Energy-Physics Event Generation with PYTHIA 6.1*, hep-ph/0010017

[32] M. Bertini, L. Lönnblad, T. Sjöstrand, *Pythia version 7-0.0 - a proof-of-concept version*, hep-ph/0006152

[33] F. Maltoni and T. Stelzer, *MadEvent: Automatic event generation with MadGraph*, JHEP 0302 (2003) 027 [arXiv:hep-ph/0208156]

[34] E. Gerchtein, M. Paulini, *CDF Detector Simulation Framework and Performance*, physics/0306031

[35] R. Brun and F. Carminati, *GEANT Detector Description and Simulation Tool*, CERN Programming Library Long Writeup W5013 (1993)

[36] Rob Veenhof, *GARFIELD, a drift chamber simulation program*, Version 6.22, CERN

[37] G. Grindhammer, M. Rudowicz, and S. Peters, *The Fast Simulation of Electromagnetic and Hadronic Showers*, Nucl. Instrum. and Methods A290, 469 (1990)

[38] D. Hirschbühl, J. Lück, Th. Müller, A. Papaikonomou, Th. Peiffer, M. Renz, S. Richter, I. Schall, J. Wagner-Kuhr, W. Wagner, *Search for single top-quark production with neural networks using 1.92 fb-1*, CDF Internal Note 9107

[39] C. Issever, B. Wagner, A. Robson, C. Mills, *Plug Electron Baseline Cuts as defined in Summer 2003 and their Efficiencies*, CDF Internal Note 6789

[40] Heather K. Gerberich, Ashutosh V. Kotwal, Chris Hays, *Cosmic Ray Tagging using COT Hit Timing*, CDF Internal Note 6089

[41] B. Flaugher, J.Mueller, *A Guide to JETCLU: The CDF Jet Cluster Algorithm*, CDF Internal Note 1814

[42] A. Bhatti et al., *Determination of the jet energy scale at the Collider Detector at Fermilab*, Nucl. Instrum. Methods A 566, 375 (2006)

[43] C.Rott, H.Bachacou, K.Hoffman, A.Dominguez, W.Yao, D.Glenzinski, D.Bortoletto, M.Shochet, *SecVtx Optimization for the 2003 Winter Conferences*, CDF Internal Note 6246

[44] W. Badgett, H. Frisch, R. St. Denis, T. Vaiciulis, *The Good Run List*, CDF Internal Note 5613

[45] Oliver Brein, Abdelhak Djouadi, Robert Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, hep-ph/0307206

[46] Veronique Boisvert, *Trigger Efficiencies for the High ET Central Electrons in Gen6*, CDF Internal Note 7939

[47] Bo-Young Han, Eva Halkiadakis, *MET PEM trigger efficiency for Pheonix electrons*, CDF Internal Note 7940

[48] Bo-Young Han, Veronique Boisvert, *Trigger Efficiencies for the High Et Central Electron in the Gen6 data(run:217990 - 222426)*, CDF Internal Note 8629

[49] Ulysses Grundler, Lubomir Lovas, Anyes Taffard, *High-Pt muons recommended cuts and efficiencies for Winter 2007*, CDF Internal Note 8618

[50] D. Hare, E. Halkiadakis, T. Spreitzer, *Electron ID Efficiency and Scale Factors for Winter 2007 Analyses* CDF Internal Note 8614

[51] W.K. Sakumoto, *Event |Zvtx| <60 Cut Acceptance for Run II (to Feb06)*, CDF Internal Note 8318

[52] Ford Garberson, Sebastian Grinstein, Joao Guimaraes da Costa, J. Incandela, Chris Neu, Daniel Sherman, *Combination of the SecVtx 1.2/fb b-Tagging Scale Factors*, CDF Internal Note 8666

[53] Franklin, Grinstein, Guimaraes da Costa, Lannon, Schwarz, Sherman, Taffard, *Method 2 Backgrounds for 1.12/fb Lepton+Jets Analyses*, CDF Internal Note 8766

[54] Michael Feindt, Thomas Müller, Svenja Richter, Wolfgang Wagner, *A Neural Network b Tagger for Single-Top Analyses*, CDF Internal Note 7816

[55] T. Chwalek, M. Feindt, D. Hirschbühl, Th. Müller, M. Renz, S. Richter, W. Wagner, *Documentation of the updated Neural Network b Tagger used for Single-Top Analyses*, CDF Internal Note 8903

[56] Michael Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, physics/0402093

[57] Richard O. Duda, Peter E. Hart, David G. Stork, *Pattern Classification*, John Wiley & Sons, Inc., 2001

[58] V. Blobel, E. Lohrmann, *Statistische und numerische Methoden der Datenanalyse*, Teubner, 1998

[59] S.Elbracht, S.Richter, W.Wagner, *Fitting the Flavor Composition of Lepton Trigger Data in Bins of MET and Lepton Isolation*, CDF Internal Note 9131

[60] Y. Bengio, H. Schwenk, *Boosting Neural Networks*, to appear in Neural Computation

[61] Y. Freund *Boosting a weak learning algorithm by majority*, Information and Computation, 121(2), 1995

[62] P. Fernandez, T. Junk, N. Krumnack, T. Masubuchi, J. Nielsen, A. Taffard, W. Yao, *Combined Upper Limit on Standard Model Higgs Boson Production at CDF for 1 fb^{-1} Data*, CDF Internal Note 8941

[63] Yves Kemp, *Studien zur Suche nach Higgsbosonen im WH-Kanal bei CDF für eine Higgsmasse von 120 GeV*, IEKP-KA/2001-26

[64] Martin Hennecke, *Search for the Higgs Boson with the CDF experiment at the Tevatron*, IEKP-KA/2005-9

[65] Dominic Hirschbühl, *Measurement of the Charge Asymmetry and the W Boson Helicity in Top-Antitop Quark Events with the CDF II Experiment*, IEKP-KA/2005-?

[66] Thorsten Chwalek, *Measurement of the W Boson Helicity in Top-Antitop Quark Events with the CDF II Experiment*, IEKP-KA/2006-14

[67] Julia Weinelt, *Measurement of the Charge Asymmetry in Top-Antitop Quark Production with the CDF II Experiment*, IEKP-KA/2006-21

[68] Svenja Richter, *Search for Electroweak Single Top-Quark Production with the CDF II Experiment*, IEKP-KA/2007-14

Danksagung

Das Zustandekommen dieser Arbeit wäre nicht möglich gewesen ohne das Zutun einiger Personen, denen ich an dieser Stelle meinen Dank aussprechen möchte.

An erster Stelle danke ich Herrn Prof. Dr. Thomas Müller dafür, dass ich meine Diplomarbeit innerhalb seiner Arbeitsgruppe erstellen konnte sowie dass er mir einen Aufenthalt am Fermilab und den Besuch des CDF Collaboration Meetings in Paris ermöglicht hat.

Desweiteren möchte ich mich bei Herrn Priv.-Doz. Dr. Wolfgang Wagner für die Übernahme des Korreferats sowie die gute Betreuung innerhalb der Top-Gruppe bedanken.

Ein besonderer Dank gebührt Frau Dr. Jeannine Wagner-Kuhr für die vielen hilfreichen Tipps, Vorschläge und Diskussionen sowie für das sehr detailierte und ausführliche Korrekturlesen dieser Arbeit.

Weiterhin danke ich allen Mitgliedern der Top-Arbeitsgruppe für die angenehme Arbeitsatmosphäre. Insbesondere seien erwähnt Dr. Dominic Hirschbühl und Jan Lück für ihre vielen oft klärenden Hinweise. Ebenfalls danke ich Thorsten Chwalek für das Korrkturlesen dieser Arbeit, Julia Weinelt und Georg Sartisohn für die gute Zusammenarbeit in userem Büro und die vielen gegenseitig gegebenen kleinen Hilfen, die oft groß erscheinende Probleme schnell beseitigt haben. Außerdem sei an dieser Stelle allen weiteren Mitgliedern des EKP gedankt, insbesondere dem Admin-Team für die freundliche Unterweisung im Umgang mit den Computing-Ressourcen.

Ich möchte mich auch bei meinen zum teil langjährigen Karlsruher WG-Mitbewohnern bedanken, mit denen man zum Glück über fast alles außer Physik reden kann.

Meinen Eltern danke ich dafür, dass sie immer und uneingeschränkt für mich da sind.

Hiermit versichere ich, die vorliegende Arbeit selbstständig verfasst und nur die angegebenen Hilfsmittel verwendet zu haben.

Thomas Peiffer

Karlsruhe, den 15. Januar 2008