Entwicklung und Anwendung einer Softwareumgebung für die Neuronale-Netzwerk-Analyse zur Messung einzelner Topquarks mit dem CDF-II-Experiment

Irja Schall

Diplomarbeit

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

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

an der Fakultät für Physik der Universität Karlsruhe (TH)

26. Februar 2008

Zusammenfassung

Ziel der Teilchenphysik ist das Verständnis der fundamentalen Bausteine der Materie, sowie deren Wechselwirkungen. So erkannte man in der zweiten Hälfte des 20. Jahrhunderts nicht nur die Substruktur von Atomen, sondern es wurden auch neue Teilchen entdeckt, welche in der uns umgebenden Materie nicht vorkommen. Die elementaren Teilchen sowie deren Wechselwirkungen werden durch das Standardmodell der Teilchenphysik beschrieben.

Das letzte Elementarteilchen, das entdeckt wurde, ist das Topquark. Aufgrund seiner hohen Masse, die ungefähr der eines Goldatomes entspricht, konnte es erst 1995 am Tevatron nachgewiesen werden. Das Tevatron ist ein Hochenergie-Teilchenbeschleuniger, der zur Zeit mit der weltweit höchsten Schwerpunktsenergie von 1.96 TeV operiert. Die vorliegende Arbeit untersucht Daten des CDF II Experimentes, bei dem es sich um eines der beiden Großexperimente am Tevatron handelt. Die integrierte Luminosität des Datensatz beträgt 2.2 fb^{-1}.

Topquarks werden vorwiegend paarweise über die starke Wechselwirkung erzeugt. Vom Standardmodell wird aber auch die seltenere Erzeugung einzelner Topquarks über die elektroschwache Wechselwirkung vorhergesagt. Dieser Produktionsmechanismus ist von besonderem Interesse, da er auf verschiedene Arten eine Überprüfung des Standardmodells ermöglicht. Da die Topquarkmasse von der gleichen Größenordnung wie die Skala der Symmetriebrechung der elektroschwachen Wechselwirkung ist, handelt es sich bei der elektroschwachen Erzeugung einzelner Topquarks um einen Prozeß, der Aufschluß über Abweichungen vom Standardmodell bieten kann. Einzelne Topquarks werden über einen Wtb Vertex erzeugt, so daß eine direkte Messung des Cabbibo-Kobayashi-Maskawa (CKM) Matrixelementes V_{tb} möglich wird. Eine Abweichung von dem aufgrund der postulierten Unitarität der CKM Matrix erwarteten Wert von eins, kann zum Beispiel durch die Existenz einer weiteren Quarkfamilie verursacht werden.
Am Tevatron werden einzelne Topquarks vorwiegend durch zwei Prozesse erzeugt. Für den s-Kanal, bei dem zwei Quarks zu einem W-Boson fusionieren, das anschließend in ein Topquark zerfällt, wird ein Wirkungsquerschnitt von $\sigma_s = 0.88^{+0.12}_{-0.11}$ pb vorhergesagt. Für den t-Kanal, bei dem zwei Quarks aneinander über den Austausch eines W-Bosons streuen, bei dessen Zerfall wiederum ein Topquark entsteht, beträgt der erwartete Wirkungsquerschnitt $\sigma_t = 1.98^{+0.28}_{-0.22}$ pb.

In nahezu 100% aller Fälle zerfällt das Topquark in ein W-Boson und ein b-Quark.

Es werden nur leptonische W-Boson Zerfälle betrachtet. Die Signatur elektroschwach erzeugter Topquarks besteht also aus einem Lepton, aus fehlender Transversalenergie, hervorgerufen durch das nicht nachzuweisende Neutrino, und einem b-Quark-Jet. Der b-Quark-Jet wird dabei mithilfe der langen Lebensdauer von b-Hadronen identifiziert, welche versetzt zum primären Wechselwirkungspunkt an einem Sekundärvertex zerfallen.

Aufgrund des niedrigen Wirkungsquerschnittes und der Signatur der Einzel-Topquarkereignisse, die vielen anderen Prozessen gleicht, wird ein hoher Untergrund erwartet. Um die Einzel-Topquarkereignisse von den Untergrundereignissen zu trennen, werden neuronale Netze verwendet. Bei dem Training der neuronalen Netze wurde nicht nur zwischen den verschiedenen Einzel-Topquarkprozessen unterschieden, sondern es wurde auch nach Anzahl der Jets, sowie der Anzahl der ein b-Quark enthaltenden Jets differenziert. So wurde ein s-Kanal Netzwerk mit Ereignissen, die über zwei Jets verfügen, von denen beide als b-Quark-Jet identifiziert wurden, trainiert. Für den t-Kanal wurden drei Netzwerke trainiert. Eines mit Ereignissen mit zwei Jets, von denen einer als b-Quark-Jet erkannt wurde. Ein anderes mit Ereignissen, in denen drei Jets vorkommen, wobei einer aus einem b-Quarkzerfall herrühren muss. Das dritte wurde mit Ereignissen trainiert, die über drei Jets verfügen, von denen zwei aus einem b-Quarkzerfall stammen. Der jeweilige andere Einzel-Topquarkprozeß wird beim Training dieser Netzwerke nicht als Untergrund betrachtet. In dieser Arbeit wird die kombinierte Suche nach Einzel-Topquarkerzeugung beschrieben. Dabei werden zwar die einzelnen Netzwerke auf einen Kanal optimiert, im weiteren werden aber die beiden Einzel-Topquarkprozesse zusammengefaßt, weil anstatt der Messung der getrennten Wirkungsquerschnitte, der Nachweis einzeln erzeugter Topquarks bei dieser Analyse im Vordergrund steht.
Um die Analysemethode zu überprüfen wurden desweitern zwei Kontrollnetzwerke trainiert. Eines dient der Messung des Anteils an $W b\bar{b}$-Produktion in Ereignissen mit genau einem Jet. Das andere, auf Erkennung von $t\bar{t}$ Ereignissen optimierte Netzwerk, mißt den $t\bar{t}$ Wirkungsquerschnitt. Dieser stimmt mit anderen Messungen überein.

Um den Aufwand, der mit einer so großen Anzahl an Netzwerken verbunden ist, zu verringern, wurde eine Softwareumgebung geschrieben, welches die Anzahl der Schritte, die für das Training jedes einzelnen Netzwerks nötig sind, reduziert und vereinfacht. Durch die Softwareumgebung wird sowohl eine höhere Effizienz, als auch eine gesteigerte Robustheit und eine bessere Portierbarkeit der Software erzielt.

Die Musterverteilungen, die man aus der Anwendung der neuronalen Netzwerke auf die Datensätze der simulierten Signal- und Untergrundprozesse erhält, fließen bei der kombinierten Suche nach Einzel-Topquarkerzeugung gemeinsam in eine Likelihood-Funktion ein. Unter Berücksichtigung der erwarteten Anzahl an Untergrundereignissen und der systematischen Unsicherheiten, wird der negative Logarithmus der Likelihood-Funktion minimiert, so daß die Musterverteilungen an die Daten angepaßt werden.
Um die erwartete Sensitivität der Messung zu bestimmen werden Ensembletests

durchgeführt. Der zu erwartende Fehler auf den Wirkungsquerschnitt beträgt 26.3%, was 0.75pb entspricht. Als erwartete Signifikanz der Messung erhält man 4.4σ.

Der Wirkungsquerschnitt der elektroschwachen Erzeugung einzelner Topquarks in den gemessenen Daten ergibt sich zu $2.0^{+0.9}_{-0.8}$pb. Dies entspricht $70.5^{+30.1}_{-28.4}$% der Vorhersage, es handelt sich also im Rahmen der Fehler um einen mit dem Standardmodell in gutem Einklang stehenden Wert. Die Signifikanz der Messung beträgt 3.2σ. Im Gegensatz zu vorangegangenen Messungen mit neuronalen Netzen bei CDF II, konnte somit in der vorliegenden Analyse, eine Evidenz für die elektroschwache Erzeugung einzelner Topquarks gefunden werden. Dieses Ergebnis wird sowohl von den anderen aktuellen Analysen zur Messung einzelner Topquarks bei CDF II, als auch, auf einem geringeren Datensatz, von dem zweiten Tevatron Experiment DØ bestätigt.

In Bild 1 ist der kombinierte Output der vier neuronalen Netzwerke gezeigt. Die vorhergesagten Verteilungen sind dabei auf die Fitergebnisse normiert. Aus den angegebenen χ^2- und KS-Test-Werten entnimmt man, daß diese Verteilungen gut mit den gemessenen Daten übereinstimmen.
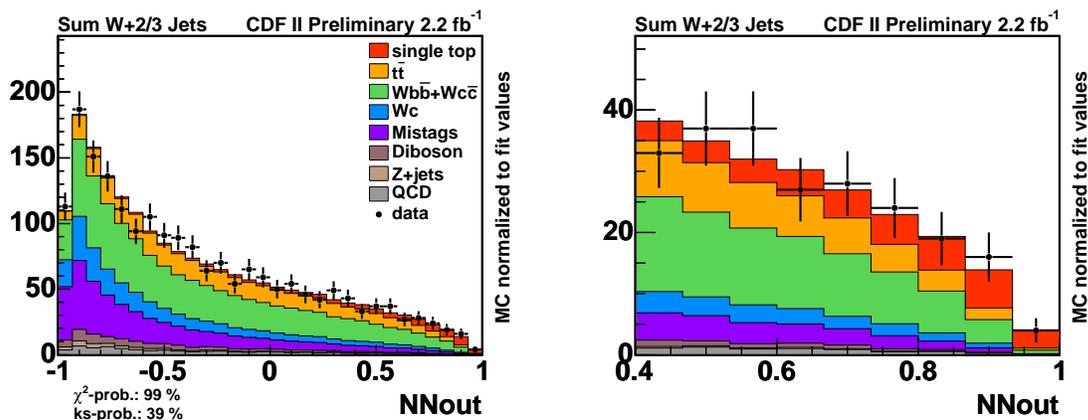
(a) Kombinierte Suche   (b) Kombinierte Suche, Signalregion

Abbildung 1: Zusammengefaßter Output der neuronalen Netzwerke in der kombinierten Suche. Die Output-Verteilungen der vier neuronalen Netzwerke sind aufaddiert und auf die Fitergebnisse normiert. Bild (a) zeigt den ganzen Bereich des Ouputs, Bild (b) die Signalregion.
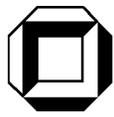
Development and Application of a Software Framework for the Single Top Quark Analysis using neural networks with the CDF II Experiment

Irja Schall

Diploma Thesis

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

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

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

February 26, 2008

Introduction

The goal of particle physics is to understand the fundamental structure of matter. In the last century the constituents of atoms have been deciphered and new particles, naturally not occurring around us, have been discovered. Elementary particles as well as their interactions are described by the standard model of particle physics.

The last elementary particle that has been discovered is the top quark. Due to its high mass, that approaches the one of a gold atom, it eluded discovery until 1995, when it was observed by the two Tevatron experiments CDF II and DØ [1, 2]. The Tevatron is hosted by the Fermi National Accelerator Laboratory (Fermilab) and with 1.96 TeV it is up to date the accelerator with the highest center of mass energy.

The predominant production mode of top quarks is the production of top-antitop quark pairs via the strong interaction. Top quarks can also be produced singly via the electroweak interaction. This production mode is of particular interest, for it provides a sensitive probe for several standard model parameters. Since the top quark mass is of the order of the electroweak symmetry breaking scale, it is a promising place to look for deviations from the standard model. Single top quarks are produced from a Wtb vertex, the measurement of single top quark production thus allows the determination of the Cabbibo-Kobayashi-Maskawa (CKM) matrix element V_{tb}. Assuming the unitarity of the CKM matrix, V_{tb} is expected to equal 1. A differing value could be caused by a fourth quark generation or other exotic phenomena.

At the Tevatron mainly two processes contribute to the electroweak production of single top quarks. In an s-channel event the fusion of two quarks produces a W boson, which decays into a top and a b quark. The theoretical cross section for this channel is $\sigma_s = 0.88^{+0.12}_{-0.11}$ pb. In a t-channel event, two quarks interact through the exchange of a W boson that again yields a top quark. The predicted cross section for this channel is $\sigma_t = 1.98^{+0.28}_{-0.22}$ pb.

The top quark almost exclusively decays into a W boson and a b quark. Only leptonically decaying W bosons were considered. The signature of single top quark production consists thus of a lepton, a neutrino that manifests itself in missing transverse energy and a b quark. The b quark jet is identified by the long lifetime of b-hadrons, that travel a small distance from the primary interaction point before decaying at a secondary vertex.

Due to its small cross section and its signature, single top quark production is dominated by a large background. To separate single top quark from background

events, four neural networks are trained. A network to distinguish s-channel events is trained on events with two jets and two b tags. Three networks are optimized to recognize t-channel events. One is trained on events with two jets and one b tag, another on events with three jets and one b tag, the last one is trained on events with three jets and two b tags.

In the combined search for single top quark production the observation of single top quark events has a higher priority than the measurement of the two separate channels. Although the different networks are optimized to recognize single top quark events of a given channel, the other channel is not considered as a background in the training. After the training they are therefore comprised in a single process, weighted with the fraction predicted by the standard model.

To validate the analysis' methods, two control networks are trained. One measures the fraction of $W b\bar{b}$ events with one jet. The other was trained on $t\bar{t}$ events with three jets and one b tag. It measures the $t\bar{t}$ cross section.

To reduce the effort related with the training of a large number of networks, a neural network analysis framework was written. Its purpose is to reduce the number of steps necessary for each network as well as to simplify them. It accomplishes an increased efficiency, robustness and portability of the software.

Contents

Chapter 1

Top Quarks within the Standard Model

1.1 The Standard Model of Elementary Particle Physics

With the discovery of the top quark in 1995 at the Tevatron [1, 2] the last elementary matter particle of the standard model of elementary particle physics has been found.

The standard model describes the interactions between all elementary particles that can be observed with the means of todays high energy particle accelerators, as well as these particles themselves.

The interactions covered by the standard model are the electromagnetic, the weak and the strong interaction. Gravitation is not included. The standard model particles comprise two main groups:

- elementary matter particles with spin 1/2, fermions

- the mediators of the fundamental forces, bosons, with spin 1

All fermions interact weakly. Every fermion has an antiparticle that is identical to its partner except that it has opposite quantum numbers. It can be distinguished between two classes of fermions: leptons and quarks. The leptons consist of the electron, the muon, the tau and their respective neutrinos. The leptons can be ordered into three generations as is shown in table 1.1 that lists the mass and the charge of the leptons.

Like the leptons, the quarks interact via the electromagnetic and the weak interaction, but also via the strong interaction. Quarks exist only in bound states (hadrons) of two's (quark-antiquark pairs called mesons) or three's (baryons). Since it is forbidden by the Pauli principle that several fermions are in the same quantum state,

name	symbol	el. charge $[e]$	mass $[\mathrm{MeV}/c^2]$
electron	e	-1	0.51
electron neutrino	ν_e	0	$\leq 2 \cdot 10^{-6}$
muon	μ	-1	105.66
muon neutrino	ν_μ	0	≤ 0.190
tau	τ	-1	1777.0
tau neutrino	ν_τ	0	≤ 18.2

Table 1.1: Mass in units of MeV/c^2 [3] and electrical charge in units of the electron charge of the fermions.

an additional quantum number, a color charge [4, 5], was attributed to the quarks. There are six quarks (plus six antiquarks): the up, the down, the charm, the strange, the top and the bottom quark. As for the leptons there are three quark generations. In table 1.2 the mass and the electrical charge of the quarks, sorted by their generation, are given.

name	symbol	el. charge $[e]$	mass $[\mathrm{MeV}/c^2]$
up quark	u	$+\frac{2}{3}$	$1.5 - 3.0$
down quark	d	$-\frac{1}{3}$	$3 - 7$
charm quark	c	$+\frac{2}{3}$	$(1.25 \pm 0.09) \cdot 10^3$
strange quark	s	$-\frac{1}{3}$	95 ± 25
top quark	t	$+\frac{2}{3}$	$(172.5 \pm 2.7) \cdot 10^3$
bottom quark	b	$-\frac{1}{3}$	$(4.20 \pm 0.09) \cdot 10^3$

Table 1.2: Mass in units of MeV/c^2 [3] and electrical charge in units of the electron.

Elementary particles, their production and the interaction between them are described with quantum field theories. Quantum field theories make use of perturbation theory to calculate transition amplitudes and cross sections. Different orders in such a perturbation series are graphically represented by Feynman diagrams such as the one in figure 1.1. Particles are drawn as lines. The interaction between particles is visualized by a vertex.

Electromagnetic Interaction

The quantum field theory of the electromagnetic interaction is called quantum electrodynamics (QED). The electromagnetic force between electrically charged particles is mediated by the photon. It is massless and bears no charge. Charged particles interact through the exchange of a photon. They can also absorb or radiate photons. The fundamental vertex diagram, that characterizes the electromagnetic interaction
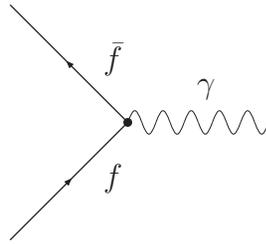
is shown in figure 1.1

Figure 1.1: QED vertex diagram: f can be any fermion, γ designates the photon, the time flow in this sort of diagram goes to the right.

Weak Interaction

The $W^{\pm}$ and the Z bosons mediate the weak interaction. They are massive and the $W^{\pm}$ bosons are electrically charged so that they interact via the electromagnetic interaction as well. Due to the high masses of the gauge bosons, $m_W = 80.403 \pm 0.029$ and $m_Z = 91.1876 \pm 0.0021$ [3], the weak interaction is of short range.

One particularity of the weak interaction is, that it doesn't couple in the same way to the left- and righthanded parts of the wavefunction of a particle: The W^- boson couples only to lefthanded particles (particles whose spins point in the opposite direction of their motion), whereas the Z boson couples to both left- and righthanded particles, although not with the same strength.

The neutral weak interaction, mediated by the Z boson, has the structure shown in figure 1.2
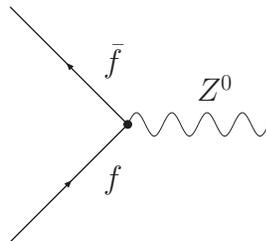

Figure 1.2: Vertex diagram of the neutral weak interaction: f can be any fermion, Z^0 is the mediator of the neutral weak interaction.

The charged weak interactions distinguishes slightly between leptons and quarks. The fundamental leptonic vertex is given in figure 1.3(a):
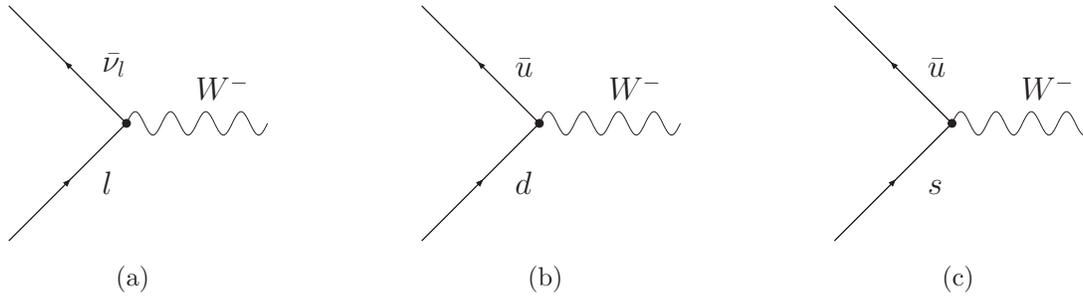
Figure 1.3: Vertex diagrams of the charged weak interaction.
Figure (a): the vertex diagram of the charged leptonic weak interaction: l designates any of the leptons electron, muon or tau, ν_l is its corresponding neutrino and W is the mediator of the charged weak interaction. The charge conjugated process is implied.
Figure (b): Vertex diagram of the charged hadronic weak interaction examplarily shown for a u-d pair. Instead of the u-d pair there could also be a c-s or a t-b pair.
Figure (c): Vertex diagram of the charged hadronic weak interaction with a cross-generational coupling.

The leptons form SU(2) doubletts under the weak interaction:

$$\left(\begin{array}{c} e \\ \nu_e \end{array} \right)_L , \left(\begin{array}{c} \mu \\ \nu_\mu \end{array} \right)_L , \left(\begin{array}{c} \tau \\ \nu_\tau \end{array} \right)_L \qquad (1.1)$$

where L stands for the lefthanded part of the wavefunctions.

Similar to the leptonic vertex a there is also one involving quarks, shown in figure 1.3(b).
Furthermore a cross-generational coupling can occur (although not as frequently). An example is shown in the diagram in figure 1.3(c).
The hadronic coupling is the same has the leptonic one, if, instead of the quark states

$$\left(\begin{array}{c} u \\ d \end{array} \right)_L , \left(\begin{array}{c} c \\ s \end{array} \right)_L , \left(\begin{array}{c} t \\ b \end{array} \right)_L \qquad (1.2)$$

one uses the Cabbibo rotated states:

$$\left(\begin{array}{c} u \\ d' \end{array} \right)_L , \left(\begin{array}{c} c \\ s' \end{array} \right)_L , \left(\begin{array}{c} t \\ b' \end{array} \right)_L \qquad (1.3)$$

that are related to the physical quarks through the Cabbibo-Kobayashi-Maskawa matrix (CKM matrix) [6, 7]

$$\left(\begin{array}{c} d' \\ s' \\ b' \end{array} \right) = \left(\begin{array}{ccc} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{array} \right) \cdot \left(\begin{array}{c} d \\ s \\ b \end{array} \right) \qquad (1.4)$$

The coupling of the top to the down quark, for instance, is then given by V_{td}.

The current values of the CKM matrix are (90% confidence level) [3]:

$$
\begin{pmatrix}
0.97383^{+0.00024}_{-0.00023} & 0.2272 \pm 0.0010 & (3.96 \pm 0.09) \cdot 10^{-3} \\
0.2271 \pm 0.0010 & 0.97296 \pm 0.00024 & (42.21^{+0.10}_{-0.80}) \cdot 10^{-3} \\
(8.14^{+0.32}_{-0.64}) \cdot 10^{-3} & (41.61^{+0.12}_{-0.78}) \cdot 10^{-3} & 0.999100^{+0.000034}_{-0.000004}
\end{pmatrix}
\tag{1.5}
$$

These values have been obtained by direct measurement as well as by using the unitarity of the CKM matrix.

Unified Electro-Weak Theory

In 1961 Glashow proposed that the electromagnetic and the weak interaction be merely two different manifestations of a single, more fundamental force – the electroweak interaction [8]. The difficulty of the large difference in the strength of the weak and of the electromagnetic coupling, which in fact reduces to the difference in mass of the photon and the W and Z bosons had to be overcome. In 1967 Weinberg and Salam showed how to achieve unification of the electromagnetic and weak interaction by spontaneous symmetry breaking [9, 10] through the Higgs mechanism [11, 12, 13, 14]. Hereby an additional field, the Higgs field, is introduced. Below a certain energy the ground state of the vacuum does not reflect the symmetry underlying the unified electroweak theory anymore, the symmetry is broken, forcing the boson fields of the electroweak interaction to mix, thus introducing mass terms for the $W^{\pm}$ and Z bosons, whereas the photon remains massless. A mixed coupling of the photon to the W and Z boson, as well as a possible coupling among the W and Z boson are shown in figure 1.4. Associated with the Higgs field is the Higgs boson, whose discovery still eludes.

Figure 1.4: Some diagrams showing the couplings of electroweak gauge bosons. Figure (a) shows the coupling among weak gauge bosons, figure (b) the coupling of the photon to the W and Z bosons.

Strong Interaction

The strong interaction is mediated by the gluons. It acts only on color charged parti-
cles: quarks and gluons themselves. The quantum field theory describing the strong
interaction is therefore known as quantum chromodynamics (QCD) [15, 16, 17].
There are three colors. The quarks carry only one color whereas the gluons are a
superposition of two or three color states. There are eight such gluons. When a
quark couples to a gluon, the quark color changes and the difference is carried of by
the gluon. A typical vertex diagram of the strong interaction is shown in figure 1.5.
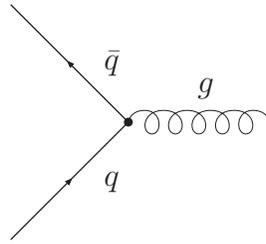

Figure 1.5: Vertex diagram of the strong interaction: a quark q and its antiparticle $\bar{q}$ interact via
a gluon g.

At short distances the coupling of the strong interaction is very small (quarks inside
a proton can move around more or less freely). This phenomenon is known as
asymptotic freedom. At large distances however the strong interaction becomes
more pronounced. As two quarks are pulled apart, so much energy is required that
between them a new pair of quarks is created, so that only color neutral particles
will emerge. When quarks are produced in a collision experiment, so called jets,
consisting of a multitude of hadrons will form. Quarks are thus always confined in
hadrons that themselves no longer carry color charges. The strong interaction is
therefore of short range.

Since gluons carry color themselves there can be gluon self-interaction, see figure 1.6.

Parton Distribution Function   In experiments where hadrons are collided, it is
important to know, what momentum fraction the constituents of the hadron carry.
For protons and antiprotons, for example, these constituents are of course the up and
the down quark. However these so called valence quarks are not the only particles
that can be found inside a hadron. Since the valence quarks interact permanently
via gluon exchange, the gluons themselves are part of the hadron. Furthermore the
gluons can split into quark pairs, the so called sea quarks. The momentum fraction
carried by all these constituents are described by a parton distribution function
(PDF). This function depends on the energy of the interaction in question. For top

(a) (b)

Figure 1.6: Vertex diagram of the strong interaction: Gluons carry color themselves and therefore couple to each other.

quark physics this scale is often set to the top quark mass: $\mu = m_{top}$. In figure 1.7 the CTEQ5L parton distribution functions [18] for $\mu^2 = (175\,GeV)^2$ is shown.
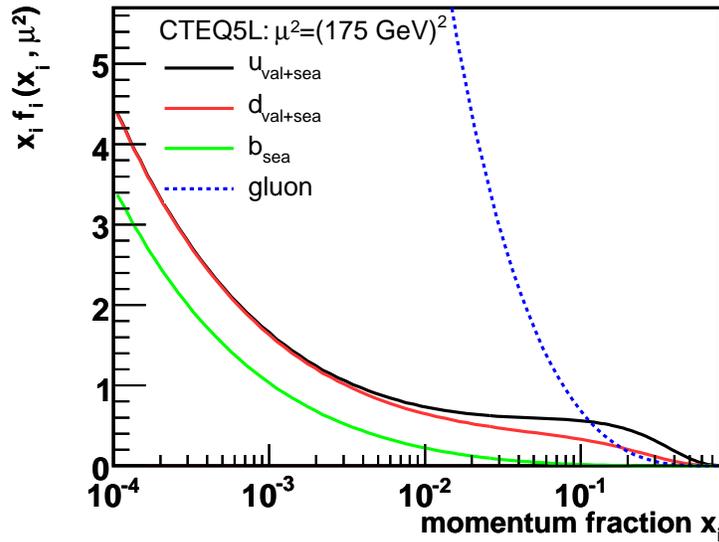

Figure 1.7: The CTEQ5L parton distribution function at $\mu^2 = (175\text{ GeV})^2$ [18]

1.2 Top Quark Production and Decay Modes

Since quarks interact both via the electroweak and the strong interaction, top quarks can be produced either by one or the other. The top quark decays via the weak interaction. Due to its short lifetime ($\approx 0.5 \cdot 10^{-24} s$), which is an order of magnitude smaller than the time scale for hadronization, the top quark decays before top-flavored hadrons can form. This has the useful consequence that the spin information carried by the top quark is still visible in its decay products, thus giving physicists the opportunity to study a bare quark. The top quark decays almost exclusively into

a b quark and a W boson; the decays into other quarks being strongly suppressed by small CKM matrix elements.

The dominant process for top quark production at the Tevatron is top quark pair production via the strong interaction. At next to leading order the $q\bar{q}$ annihilation Feynman diagram 1.8(a) contributes 85%, the gluon fusion diagram 1.8(b) 15% for $t\bar{t}$ production.

For a top quark mass of $m_{\text{top}} = 175$ GeV/c^2, the production cross section at the Tevatron with $\sqrt{s} = 1.96$ TeV is predicted to be $\sigma_{t\bar{t}} = 6.70^{+0.71}_{-0.88}$ pb [19].

(a)   (b)

Figure 1.8: Leading-order Feynman diagrams of top quark pair production: (a) quark-antiquark annihilation and (b) gluon fusion.

Additionally top quarks can be produced singly from a Wtb vertex via the electroweak interaction. Within the standard model three processes resulting in a single top quark in the final state can be distinguished. They differ in the virtuality of the produced W boson $Q^2 = -q^2$, where q is the 4-momentum of the W boson:

s-channel W^* Production   In case of the W^* production a W boson is produced by the annihilation of two quarks q and $\bar{q'}$; one comes from the proton, the other from the antiproton (see Feynman diagram 1.9(a)). The center of mass energy of this process is given by the Mandelstam variable $s = (p_q + p_{q'})^2$. It follows that the squared 4-momentum q^2 of the W boson is positive: $q^2 = s \geq (m_{\text{top}} + m_b)^2$. The W boson is thus space like. This production mode is named **s-channel** according to the relevant Mandelstam variable. In figures 1.9(b) and 1.9(c) some next to leading order corrections of this process are shown. They include initial state gluon splitting.

At the Tevatron, the predicted cross section of s-channel production at next to leading order is $\sigma_s = 0.88^{+0.12}_{-0.11}$ pb for $m_{\text{top}} = 175$ GeV/c^2 [20, 21].

In figure 1.9 only diagrams involving a u quark in the initial state are shown examplarily. The coupling to the (u,d) pair is by far the dominant process. Diagrams with a s or c quark in the initial state contribute only $\approx 2\%$ in the s-channel[22].

t-channel W-Gluon Fusion Production   In the W-gluon fusion production a b sea quark from the proton or antiproton scatters off a light quark in the other baryon
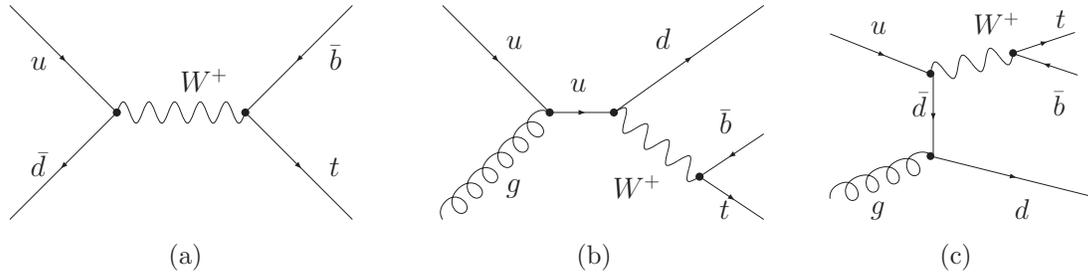
Figure 1.9: Single top s-channel. Figure (a) shows the leading order diagram, figure (b) and figure (c) are next to leading order diagrams

by exchanging a W boson, see Feynman diagram in figure 1.10(a). The squared 4-momentum q^2 of the W boson is negative: $q^2 = t = (p_{\rm b}+p_{\rm top})^2 < 0$. The W boson is therefore virtual or time like. The Mandelstam variable t is a characteristic variable for a scattering process since it is proportional to the scattering angle in the center of mass frame. This production mode is therefore called **t-channel**.
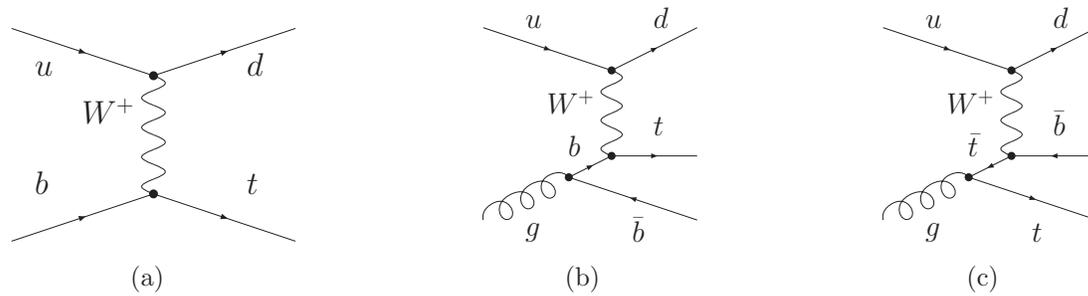

Figure 1.10: Some W gluon fusion diagrams. Figure (a) shows the leading order t-channel diagram, figure (b) is a next to leading order t-channel diagram and figure (c) represents the crossed (or u-channel) next to leading order diagram

At the Tevatron the b quark must stem from a gluon splitting as shown in the diagram in figure 1.10(b) and 1.10(c). In the theoretical calculations of the cross section, the b quark is taken to be massless. This leads however to a singularity if the initial state gluon is collinear with the final state $\bar{b}$ quark. In consequence the perturbation series describing the total cross section will converge rather slowly. To avoid this problem, a parton distribution function is introduced for the b quark. As a result the leading order process will then be 1.10(a). The predicted cross section of the W gluon fusion process at next to leading order at the Tevatron is given by $\sigma_t = 1.98^{+0.28}_{-0.22}$ pb, where $m_{\rm top} = 175$ GeV$/c^2$ was assumed. The cross section is taken from reference [20], the uncertainty from reference [21]; it includes uncertainties due to the factorization scale ($\pm 4\%$), the choice of PDF parameterization ($^{+11.3\%}_{-8.1\%}$), and the uncertainty in the top quark mass ($^{-6.9\%}_{+7.5\%}$).

Again only diagrams with a u quark in the initial state are shown in figure 1.9. The contributions of diagrams with c or s quarks in the initial state are only ≈ 6 % for

the t-channel [22].

Associated Production   In the associated production an on-shell W boson ($q^2 = m^2_W$) is produced in association with a top quark as shown in figure 1.11.
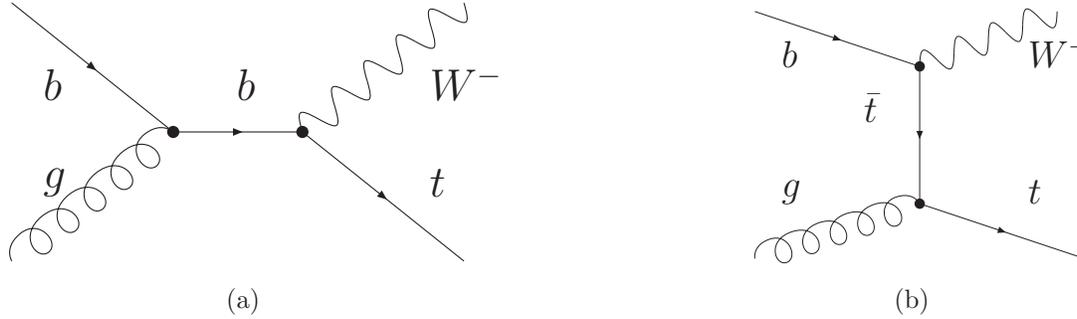

Figure 1.11: Some diagrams contributing to the associated production of single top quarks

At the Tevatron this process is neglegible ($\sigma_a = 0.094^{+0.015}_{-0.012}$ pb [23], assuming $\sqrt{s} = 2.0$ TeV and $m_{\text{top}} = 175$ GeV/c^2) and will therefore be ignored in the rest of this analysis. At the LHC, this process will however gain in importance [24].

The interest   in the measurement of the single top quark production mode lies in the different possibilities it provides to probe the standard model. It is the only process that makes it currently possible to do a direct measurement of the Cabbibo-Kobayashi-Maskawa matrix element V_{tb}. Indeed its determination has so far always relied on the unitarity of the CKM matrix. A measurement of $V_{tb} \not\simeq 1$ would require an extension of the standard model, as it now exists.

A detailed overview   of top quark physics can be found in reference [25].

Chapter 2

The Experiment

In order to discover and study elementary particles, increasingly complicated particle accelerators and detectors have been build in the second half of the 20th century, thus leading to the emergence of the Standard Model of particle physics in the first place.

At the Fermi National Accelerator Laboratory (Fermilab), located in Illinois (USA), protons and antiprotons are accelerated to an energy of 1.96 TeV and brought to collision at the two experimental sites B0 and D0. Until the launch of the LHC at the CERN, the Tevatron remains the accelerator with the highest center of mass energy. In figure 2.2 an aerial view of the Fermilab is given.

Figure 2.1: Aerial view of the Fermilab.

2.1   The Accelerator Chain

A complicated process of producing and accelerating protons and antiprotons takes place before they can finally be collided. A schematic overview of the Tevatron accelerator system is shown in figure 2.2.
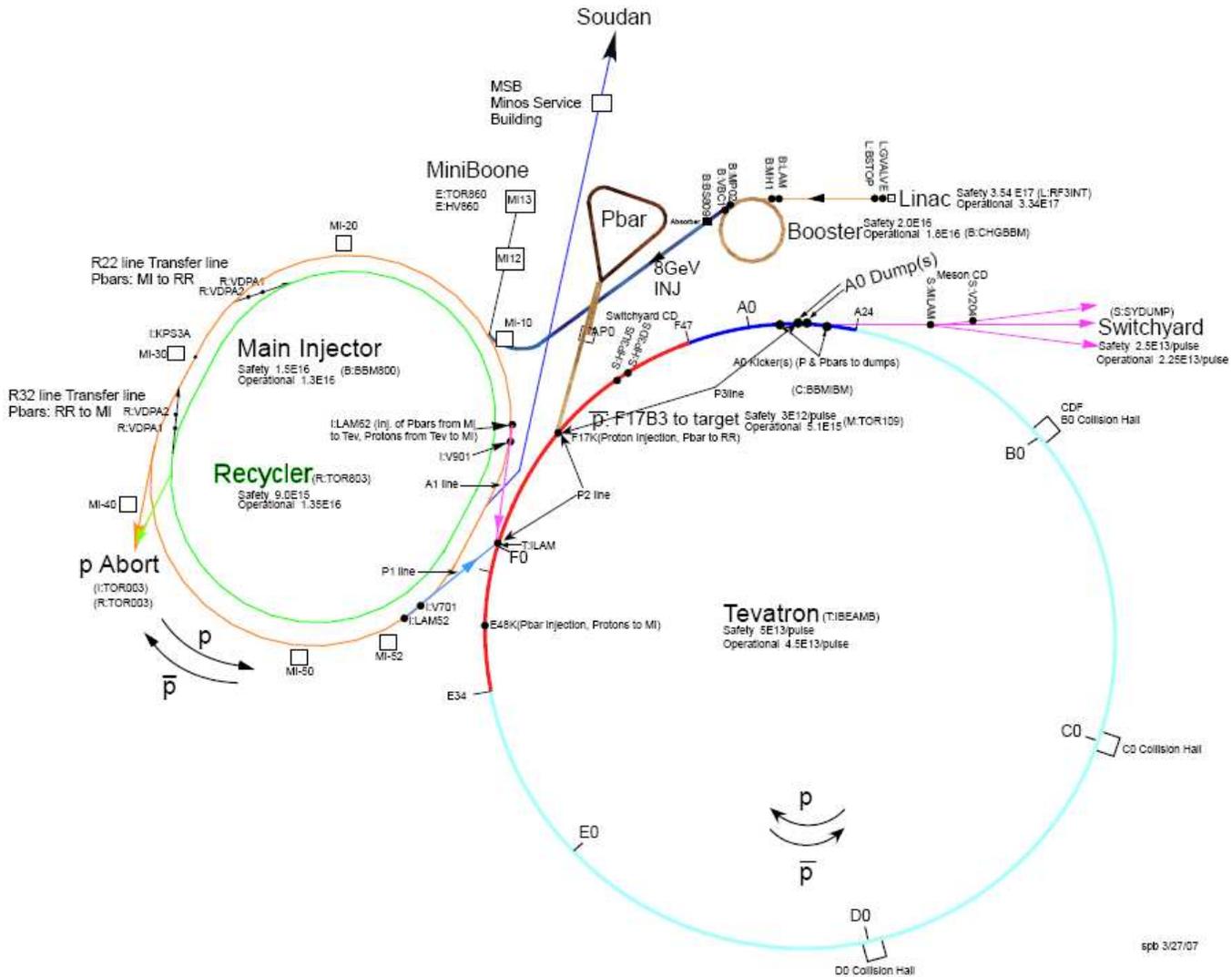

Figure 2.2: Accelerator Overview.

The first part in the accelerator chain is the Cockroft-Walton preaccelerator.

Cockroft-Walton

At the Fermilab protons are obtained from hydrogen, that in a first step is ionized to H$^-$ ions in one of the two Cockroft-Walton style electrostatic preaccelerators. A

30ft^3 gas bottle provides extremely pure hydrogen and is sufficient for six months of source operation. The hydrogen gas is ionized to H$^-$ ions in a direct-extraction magnetron type, that is part of the Cockroft-Walton. Since the dome, in which the ion source is housed, is kept at a negative potential of -750kV, the negative hydrogen ions are accelerated when leaving the source and will have gained an energy of 750 keV when they reach the grounded preaccelerator wall. A transport line guides the ions to the next stage of the accelerator chain, the Linear Accelerator (Linac).

Linac

The Linac is composed of two different types of accelerating components: drift tubes and side-coupled cavities. Five drift tubes make up the first stage of the Linac. A drift tube is a copper tube positioned inside a Linac RF cavity that protects the particle beam from the sinusoidally varying accelerating field. When the beam travels out of the drift tube and into the accelerating gap the beam sees a positive gradient and is accelerated. The drift tube Linac accelerates the beam to 116 MeV. The length of the drift tube Linac is 75 m. Seven side-coupled cavities constitute the second stage of the Linac. A side-couple cavity gives a particle an energy boost for each cavity it goes through. Each cell acts as an accelerating cavity that is coupled to another cell. Such a Fermi cavity module consists of 16 accelerating cells and 15 coupling cells. A particle is accelerated by applying small, but increasingly larger, pulses of RF energy. The side coupled Linac has a length of 64 meters and accelerates the beam to 400 MeV. In order to increase the energy of the H$^-$ ions further, they are sent to the Booster.

Booster

By entering the Booster the H$^-$ ions pass through a carbon foil, that removes the electrons, leaving only the protons. The Booster is a synchrotron with a diameter of 150 m, which the protons circulate about 20,000 times experiencing repeatedly the accelerating electrical fields until they have gained an energy of 8 GeV. The protons are then send to the Main Injector.

Main Injector

The last stage in the acceleration of the protons before they are injected into the Tevatron is the Main Injector. It is a circular ring seven times the circumference of the Booster and slightly more than half the circumference of the Tevatron. It is visible in the foreground of figure 2.1. The most interesting part of the Main Injector is how many different jobs, or "modes of operations", it has: It supplies protons for antiproton production, it accelerates the protons and antiprotons for the Collider

mode, the fixed target experiments and for the neutrino experiment NuMi.

When the Main Injector is used to provide beam to the antiproton production target, it receives either a single batch of 8 GeV protons from the Booster or else two Booster batches that are merged together during the acceleration to 120 GeV. The protons are then send to the antiproton source.

When used to inject into the Tevatron main ring, the Main Injector accepts seven bunches of 8 GeV protons from the Booster and accelerates them to 150 GeV. The bunches from the Booster get coalesced, that is, they are pushed together to form one narrow, high intensity bunch. The Main Injector injects the coalesced bunch into the Tevatron in a single turn and continues to do so until it has transferred 36 coalesced proton bunches in the Tevatron.

To inject antiprotons, the Main Injector extracts four antiproton bunches from the Recycler (that is used to store antiprotons) and accelerates them to 150 GeV. It then sends four coalesced bunches of antiprotons to the Tevatron. The antiprotons travel counterclockwise in the Tevatron, in the opposite direction of the protons. This process is repeated until there are 36 coalesced bunches of antiprotons and protons circulating in the Tevatron at 150 GeV.

Antiproton Source

For the antiproton production protons from the Main Injector are send to the Antiproton Source. The Antiproton Source itself is made up of three parts:

- The nickel target is bombarded with the protons yielding a spray of secondary particles, among them 8 GeV antiprotons, that are filtered out by magnets that act as a charge-mass spectrometer.

- The triangular shaped Debuncher Ring captures the antiprotons coming off of the target and reduces their momentum spread through RF bunch rotation and adiabatic debunching [26, 27].

- The Accumulator is the storage ring for the antiprotons. It accumulates the antiprotons by momentum stacking successive pulses of antiprotons from the Debuncher. In the Accumulator the antiprotons are further cooled down. The Accumulator occupies the same triangular shaped tunnel as the Debuncher.

The entire 120 GeV antiproton production cycle takes about 2 seconds. To make as many antiprotons as possible, the target is continuously bombarded at the fastest possible rate. The antiprotons are then send to the Recycler.

Recycler

The Recycler is located directly above the Main Injector ring. It stores antiprotons from the Antiproton Source and cools them down further than is possible in the Accumulator. In addition to stochastical cooling, electron cooling is necessary for higher intensities. Electron cooling works on the principle of momentum transfer between electrons and antiprotons of the same average velocity [28]. The antiproton beam is overlaid with a 4.3 MeV electron beam travelling about 20 m along the same path in the Recycler [29]. Coulomb scattering leads to energy transfer from the antiprotons to the co-streaming electrons until thermal equilibrium is attained. After cooling the so-called "stash", the antiprotons are mined into nine parcels, each split into four bunches.
From the Recycler the antiprotons are again let into the Main Injector that transfers them into the Tevatron.

Tevatron

The Tevatron is the largest of the Fermilab accelerators, with a circumference of approximately 4 miles. The Tevatron ring can be seen on the right site in picture 2.1. The Tevatron is a superconducting alternating gradient synchrotron. It ramps the coalesced bunches of protons and antiprotons to an energy of 980 GeV. The beam is split into three trains, each containing 12 bunches with 396 ns separation. The bunch revolution time is approximately 21 μs. Once the final energy has been reached, the beam is brought to intersection at the two experimental sites B0 and D0 where the particle detectors CDFII and DØ begin taking data; the Tevatron then becomes a storage ring and a "store" has been established.

The performance of the Tevatron is characterized by the luminosity.

$$\mathcal{L} = n \cdot f \cdot \frac{N_p N_{\bar{p}}}{4\pi\sigma_x\sigma_y}, \qquad (2.1)$$

where n is the number of bunches, f is the revolution frequency, N_p ($N_{\bar{p}}$) is the number of protons (antiprotons) in each bunch, and σ_x and σ_y represent the average transverse width of the bunches. $\mathcal{L}$ is measured in $\text{cm}^{-2}\text{s}^{-1}$ or $\text{b}^{-1}\text{s}^{-1}$ where $1\text{b} = 10^{-24}\text{cm}^2$.
The luminosity is a measure of particle interaction, specifically the chance that a proton will collide with an antiproton. To achieve high luminosity the two beams of protons and antiprotons are focused and brought together by Low Beta Magnets at each of the two experiments, DØ and CDFII. Figure 2.3 shows how the initial luminosity increased over time with the improved understanding and handling of the accelerators.

The integrated luminosity is the integral of the luminosity with respect to time and is a measure of the amount of collected data. The integrated luminosity delivered by the Tevatron and recorded by the CDFII detector are shown in figure 2.4
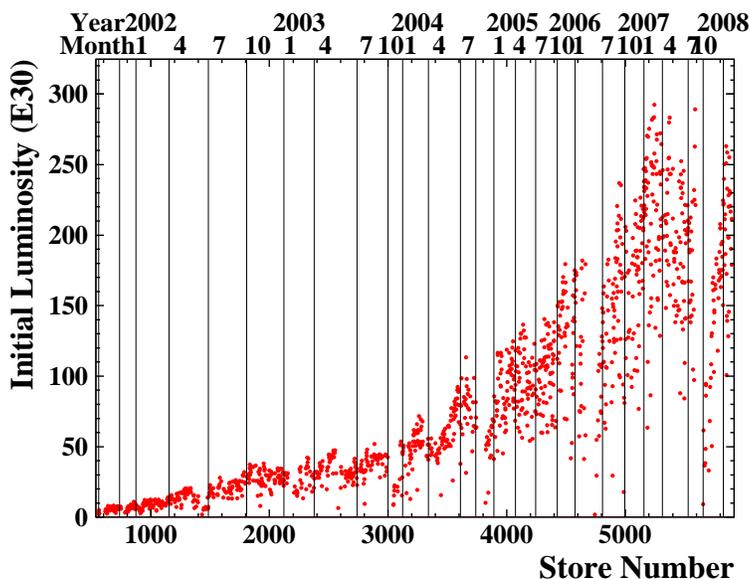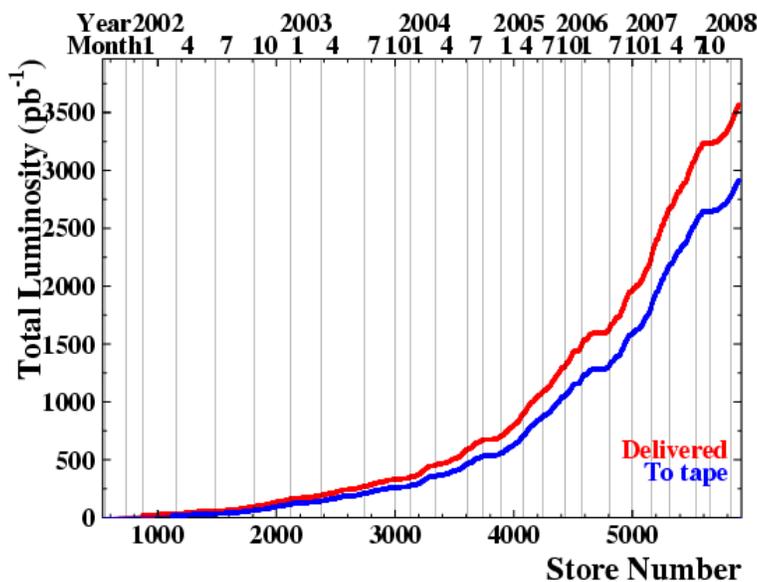
Figure 2.3: Initial luminosity in Run II.

Figure 2.4: Luminosity delivered by the Tevatron and luminosity written to tape by CDFII.

2.2   The Collider Detector at Fermilab

In order to identify the particles produced at the collision point B0 the Collider Detector at Fermilab (CDFII) does a tracking of the particles as they part from the collision point and spread into the detector, it measures their energy deposits in the calorimeters and performs a precise muon detection [30]. A solenoid generates a

1.4 T magnetic field in which the tracks of the charged particles bend so that their momentum and their electrical charge can be inferred.
In figure 2.5 a picture of the CDFII detector is shown. Figure 2.7 gives a schematic view of the detector.
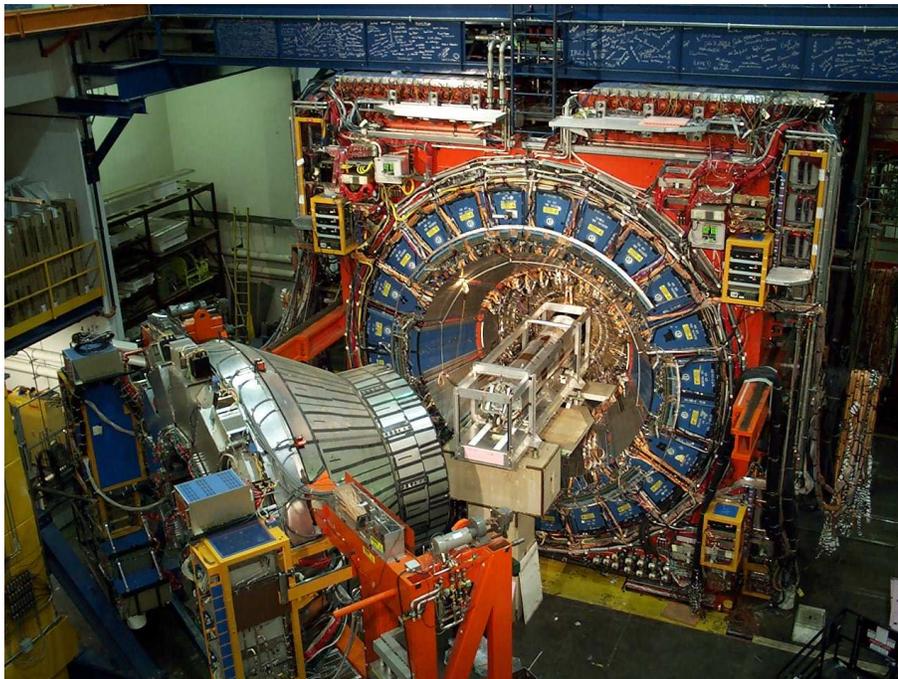

Figure 2.5: Picture of the CDFII detector, during installation of the Silicon detector

The coordinate system for the CDFII detector is shown in figure 2.6(b). The z-axis points in the direction of the protons. Instead of the polar angle θ the pseudorapidity $\eta = -ln(tan(\frac{\theta}{2}))$ is usually used. The pseudorapidity is a handy variable to approximate the rapidity of relativistic particles if their mass and momentum are not known. The pseudorapidity is used, since statistical particle distributions are flat in the rapidity for many particle physics production models. Additionally the rapidity differences are a Lorentz invariant quantity for boosts along the z-axis. A difference is also made between detector η, that gives the pseudorapidity with respect to the origin of the coordinate system, located in the center of the detector and the kinematic η, where the pseudorapidity is given with respect to the actual interaction point. A frequent variable is transverse energy and the transverse momentum of a particle; they are defined as $E_{\mathrm{T}} = E \cdot sin\theta$ and $p_{\mathrm{T}} = p \cdot sin\theta$, respectively.

In the following the different components of the detector will be described, starting with those parts close to the collision point and then following the produced particles to the outside regions of the detector. In figure 2.7(a) the inner part of the CDF detector is shown. Located around the beam pipe, closest to the interaction point is the tracking system, consisting of an silicon microstrip system and of an opencell wire drift chamber [31] surrounding the silicon detector.
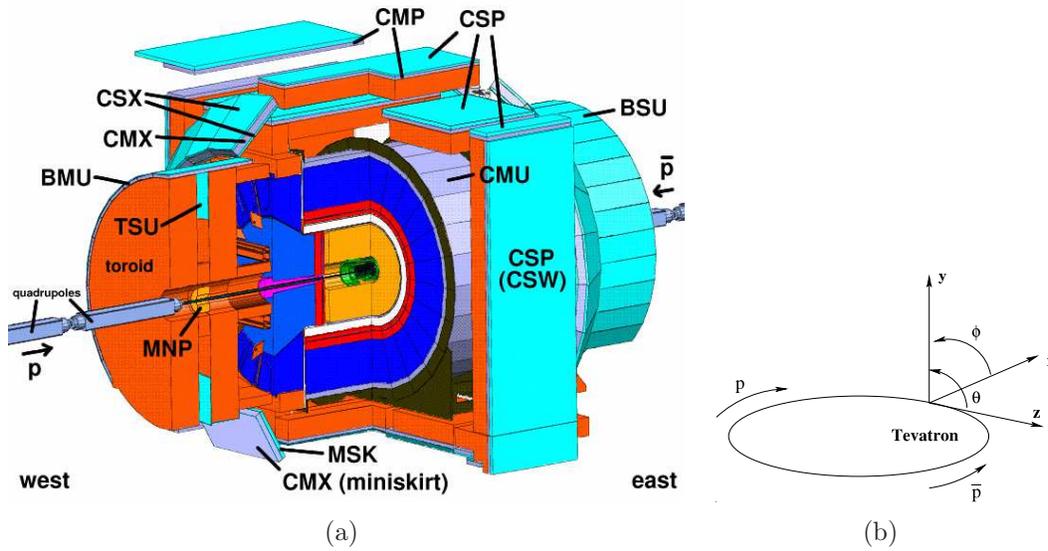
Figure 2.6: Figure (a) is a schematic view of the CDFII detector, depicting its different components. Figure (b) shows the coordinate system of the CDFII detector

Tracking System

The silicon microstrip system is made up of eight radiation-hard layers. The innermost is called Layer 00 [32], it is situated immediately outside the beampipe, at a radius of approximately r = 1.6 cm and covers the eta range $|\eta| \leq 4.0$. It is a single-sided silicon microstrip detector. The remaining seven layers are double-sided silicon microstrip detectors.

The five layers around L00 (layers 0 to 4) comprise the SVX II (Silicon Vertex) detector [33]. It extends from r = 2.1 cm to r = 17.3 cm and covers $|\eta| \leq 2.0$. The purpose of these five layers is the high-precision tracking and secondary vertex detection at inner radii for b hadron identification.

The two outer layers comprise the ISL (Intermediate Silicon Layer) system [34]. The central layer is at r = 22 cm and forward/backward layers are at r = 20 cm and r = 28 cm. The central layer covers $|\eta| < 1.0$ and the forward/backward layers cover $1.0 < |\eta| < 2.0$. In the central region the ISL is designed to provide enhanced linking of tracks between the SVX II and the surrounding drift chamber COT (Central Outer Tracker). In the forward region, where the COT coverage is incomplete, it provides improved silicon-only tracking capabilities.

The impact parameter (the closest distance of approach of the particle trajectory to the beam line) resolution of the combination of the SVX II and the ISL is 40 μm including 30 μm contribution from the beam width. The z_0 resolution of the SVX II and the ISL is 70 μm.

The Central Outer Tracker constitutes the second part of the tracking system. It expands from r = 40 cm to r = 137 cm and covers the central η range $|\eta| \leq 1.0$. It is an open-cell drift chamber with argon-ethane gas in a 50/50 mixture and it performs a general-purpose tracking in the central regions of the detector.

(a)

Figure 2.7: Cut away view of the CDF inner detector. The tracking system is shown in green and orange

The hit position resolution is approximately 140 μm, the momentum resolution $\sigma(p_\mathrm{T})/p_\mathrm{T}^2 = 0.0015\,(\mathrm{GeV}/c)^{-1}$.

Solenoid

The superconducting solenoid provides the magnetic field for tracking in the central detector. It is made of an aluminium-stabilized niobium titanium conductor and is usually used with a current of 4650 A and a 1.4 Tesla field parallel to the beam axis. The solenoid is contained within a cryostat where it is cooled by liquid helium.

Calorimeters

Placed outside the solenoid are the electromagnetic and hadronic sampling calorimeters. Only muons and neutrinos can penetrate them. As all other particles are absorbed by them, their energy deposits are measured.

There are three calorimeters covering the central region of the detector.The Central Electromagnetic Calorimeter (CEM)[35] measures the energy of electromagnetic showers in the range $|\eta| < 1.1$. Its energy resolution is $13.5\%/\sqrt{E} \oplus 1.5\%$. It is a lead/scintillator sampling detector. The energy measurements of hadronic showers is done by the Central Hadronic Calorimeter (CHA) [36] which goes up to $|\eta| < 0.9$ and has an energy resolution of $50\%/\sqrt{E} \oplus 3\%$. The range of CHA is extended by the Endwall Hadronic Calorimeter (WHA) [36] which covers $0.8 < |\eta| < 1.2$ and has the energy resolution $75\%/\sqrt{E} \oplus 4\%$. Both of them are iron/scintillator devices. In addition, for improved particle identification, the Central Shower Maximum de-

tector (CES) combined with a Central Preradiate Chamber (CPR) is integrated into CEM. In the forward region the Plug Electromagnetic Calorimeter (PEM) [37] performs the energy measurements of the electromagnetic showers in the η range $1.1 < |\eta| < 3.6$ with the energy resolution $16\%/\sqrt{E} \oplus 1\%$, whereas the Plug Hadronic Calorimeter measures the hadronic energy deposits for $1.2 < |\eta| < 3.6$ and has the energy resolution $80\%/\sqrt{E} \oplus 5\%$ The Plug Electromagnetic Shower Maximum detector (PES) and the Plug Pre-Radiate detector (PPR) are responsible shower maximum measurements.

Muon Chambers

The muon chambers [38] reside beyond the calorimetry to ensure only muons will penetrate there. The geometric and engineering problems to cover the full η region lead to 4 separate but very similar muon systems consisting of scintillators and proportional chambers. In the central region ($|\eta| < 0.6$) the muons are detected by the Central Muon Chambers (CMU) [39]. The Central Muon Upgrade (CMP) [40] is located behind CMU and the magnet return yoke. It covers the same η range as CMU and serves as a confirmation of CMU tracks. Being behind more material than CMU, CMP hits have a higher signal-to-background ratio and increase the trigger efficiency of the CMU/CMP combination. The central muon coverage is extended by the Central Muon Extension (CMX) [40] chamber that covers the η region $0.6 < |\eta| < 1.0$. In the forward region muon detection is done by the Barrel Muon Chambers (BMU), they cover $1.0 < |\eta| < 1.5$.

Trigger System

Due to the high collision rate at the Tevatron (1.7 MHz resulting from 396 ns long 36x36 bunches) it is not possible to record each event (the tape writing rate is only 100 Hz). CDF therefore uses a 3-level trigger system [41] to filter out the interesting events that can then be recorded. The data is rejected in steps, so that at each level a more sophisticated choice can be performed.
An overview of the trigger system is given in figure 2.8.

The first level trigger is a synchronous hardware trigger. It can process one event every 132 ns, an L1 decision always occurs 5 μs after a beam collision. It combines the information of calorimeter triggers that identify calorimeter based objects, of muon triggers which identify muons and of the Extremely Fast Tracker (XFT) which matches the tracks in COT to energy deposits in the calorimeters. It is then decided which events will be passed on to the next stage of the trigger. The L1 rate can be up to 35kHz.
The Level 2 trigger is a combination of hardware and software trigger. Its average processing time is 30 μs. It performs a limited event reconstruction using a custom-designed hardware consisting of several asynchronous subsystems. L2 uses
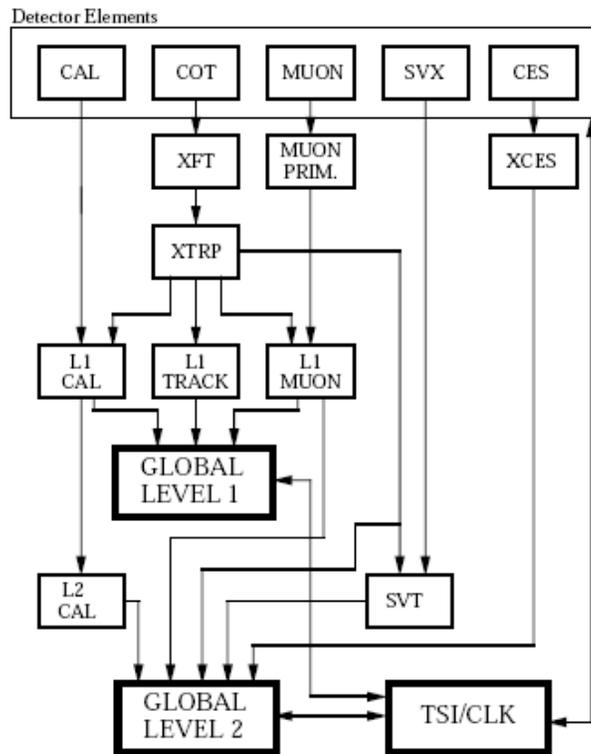
Figure 2.8: Run II trigger system

for instance the shower maximum triggers, the Silicon Vertex Tracker (SVT) that permits the selection of tracks with a large impact parameter, which is of interest in b physics. The L2 rate is typically 600 Hz.

Level 3 [42] is a pure software trigger; it reconstructs complete events, that are written to permanent storage with approximately 100 Hz.

Chapter 3

Event Selection and Classification

Among the myriads of particle interactions recorded by the CDFII detector, only those matching the single top signature are used for this analysis.

As already alluded in section 1.2 single top quark events are characterized by a signature of jets plus the decay products of a W boson in the final state. In the single top quark analysis only leptonically decaying W bosons are considered, in order to reduce QCD background. The single top signal signature is:

- two or three jets

- one isolated lepton

- missing transverse energy

In figure 3.1 a t-channel event with its decay products is shown exemplarily.
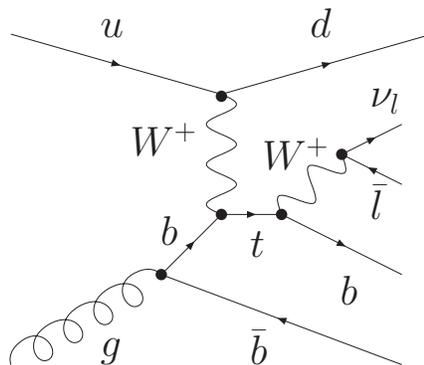
Figure 3.1: t-channel single top event with the typical signal signature

3.1   Event Selection

The data used in this analysis was recorded in the time from february 2002 until
may 2007. The recorded data is split into several samples, according to the date
when they were taken and which trigger they passed. **bhel** samples comprise events
with central electrons, **bhmu** contains events with central muons and **bpel** events
comprising forward electrons. In table 3.1 a list of the used data samples is shown.
During a store it happens that the process of taking data is interrupted due to

Period	Sample	Run Range	Dates	$\mathcal{L}$ [pb^{-1}]
0	bhel0d	138425 — 186598	02/04/2002 — 08/22/2004	331.47
0	bhmu0d	138425 — 186598	02/04/2002 — 08/22/2004	331.47/318.11
0	bpel0d	138425 — 186598	02/04/2002 — 08/22/2004	331.47
1–4	bhel0h	190697 — 203799	12/07/2004 — 09/04/2005	362.94
1–4	bhmu0h	190697 — 203799	12/07/2004 — 09/04/2005	362.94/359.50
1–4	bpel0h	190697 — 203799	12/07/2004 — 09/04/2005	362.94
5–7	bhel0i	203819 — 212133	09/05/2005 — 02/22/2006	258.37
5–7	bhmu0i	203819 — 212133	09/05/2005 — 02/22/2006	258.37/258.37
5–7	bpel0i	203819 — 212133	09/05/2005 — 02/22/2006	258.37
8	bhel0i	217990 — 222426	06/09/2006 — 09/01/2006	166.29
8	bhmu0i	217990 — 222426	06/09/2006 — 09/01/2006	166.29/166.29
8	bpel0i	217990 — 222426	06/09/2006 — 09/01/2006	166.29
9	bhel0i	222529 — 228596	09/01/2006 — 11/22/2006	156.76
9	bhmu0i	222529 — 228596	09/01/2006 — 11/22/2006	156.76/152.78
9	bpel0i	222529 — 228596	09/01/2006 — 11/22/2006	156.76
10	bhel0i	228644 — 233111	11/24/2006 — 01/30/2007	243.19
10	bhmu0i	228644 — 233111	11/24/2006 — 01/30/2007	243.19/243.49
10	bpel0i	228644 — 233111	11/24/2006 — 01/30/2007	243.19
11	bhel0j	233133 — 237795	01/31/2007 — 03/30/2007	234.99
11	bhmu0j	233133 — 237795	01/31/2007 — 03/30/2007	234.99/229.98
11	bpel0j	233133 — 237795	01/31/2007 — 03/30/2007	234.99
12	bhel0j	237845 — 241664	04/01/2007 — 05/13/2007	162.01
12	bhmu0j	237845 — 241664	04/01/2007 — 05/13/2007	162.01/155.25
12	bpel0j	237845 — 241664	04/01/2007 — 05/13/2007	162.01
13	bhel0j	241665 — 246231	05/17/2007 — 08/04/2007	280.86
13	bhmu0j	241665 — 246231	05/17/2007 — 08/04/2007	280.86/268.35
13	bpel0j	241665 — 246231	05/17/2007 — 08/04/2007	280.86

Table 3.1: Used data samples for the different periods with run number, run range, date of data
taking, and integrated luminosity. The luminosity values for `bhmu0` correspond to CMUP/CMX.

software or hardware failures. If this happens a new run number is given. The
functionality of all detector components is recorded along with the run number. In
this analysis only runs being part of the "good run list" are taken. They include
the runs where the whole tracking system, the calorimetry and the muon chambers

were fully operational. The total integrated luminosity taken during those runs is $\mathcal{L} = (2.2 \pm 0.1)\ \text{fb}^{-1}$.

Before it can be used in physics analyses, the data is reprocessed offline with the CDF software CDFSOFT2. It checks and readjusts online calibrations, corrects the alignment of the silicon detector, refits tracks, checks cluster energies, identifies leptons, performs the jet clustering and the secondary vertex fit.

It is of course of great importance that the particle identification in these data samples is reliable. To improve the purity of the data samples several cuts are imposed on the reconstructed particles or jets.

Lepton identification

Electrons   must fulfill the following requirements:

- in the central detector region electrons must exhibit a reconstructed track with $p_\text{T} > 10$ GeV/c that matches a cluster in the CEM with $E_\text{T} > 20$ GeV. To reduce events where jets fake an electron signal it is required that $E_\text{HAD}/E_\text{EM} < 0.055 + 0.00045 \cdot E$ and that the ratio of cluster energy to track momentum E/p be smaller than 2.0 for track momenta ≤ 50 GeV/c.

- in the forward region electrons are defined by a cluster in the PEM with $E_\text{T} > 20$ GeV and $E_\text{HAD}/E_\text{EM} < 0.05$. The combined information of the positions of the cluster and the primary vertex is used to search for the required track in the silicon tracker.

To reject events that arise from the conversion of a photon, electron events are furthermore required not to have a vertex with an additional high-p_T track, that has a curvature with opposite sign than the track of the electron candidate.

Muons   are identified by a COT track with $p_\text{T} > 20$ GeV/c that can be linked to a segment in a muon chamber, whereby the muon must hit CMX or both CMU and CMP. Furthermore the energy depositions in the electromagnetic and hadronic calorimeters have to correspond to those of minimum ionizing particles. Cosmic muons and muons from in-flight decays of long lived particles are excluded by demanding a small impact parameter of the track. Cosmic muons are furthermore recognized and rejected by their typical tracks going from outer space to the inside of the detector.

An isolated lepton   is required in this analysis, to further reduce the possibility of a jet being interpreted as a lepton and to exclude leptons form decays of heavy flavor quarks. Hereby the energy deposits in a cone of $R = 0.4$ around the track of the lepton candidate, that are not attributed to the lepton, must not exceed 10%

of the E_T or p_T, respectively of the lepton. Such an isolated lepton is also called a tight lepton.

A dilepton veto ensures that there is really only one lepton in the event, by rejecting those events, that have either an additional tight lepton or a loose lepton. Loose leptons are leptons which pass all cuts except the isolation cut, or are identified in the CMP, CMU or BMU solely.

A detailed description of all lepton requirements can be found in references [43, 44, 45].

Jet Reconstruction and Selection

Jets stem from the production of high-energetic quarks or gluons. They are reconstructed with an algorithm that takes all clusters of deposited energy in the calorimeters that are within a cone of $R = 0.4$. Calorimeter towers that are associated with an isolated lepton are not taken into account. Several corrections are applied to the energy of the jets. Jet energy corrections up to level 5 are used in this analysis; they effectuate an η-dependent correction that makes the calorimeter response to jet energies uniform in η (level 1), energy stemming from multiple $p\bar{p}$ interactions are subtracted (level 4), any non-linearity and energy losses in the un-instrumented regions of the central calorimeter are corrected and the jet energy measured in the calorimeter is corrected up to particle level (level 5), that is to the energy of the underlying particle jet. More detailed informations about the jet-energy corrections can be found in reference [46].
Tight Jets have corrected $E_T > 20$ GeV and detector $|\eta| < 2.8$. Loose jets have corrected E_T between 12 GeV and 20 GeV.
Only events with exactly two or three tight jets are accepted.

b-Tagging

Since the top quark produced in a single top quark event decays almost exclusively into a b quark, only events likely to feature a b quark are selected. b quarks can be identified by their long lifetime ($\approx$ 1.5ps), that allows them to travel a small distance in the form of a b-hadron before they decay at a secondary vertex, that is displaced from the primary interaction point, for an illustration see figure 3.2.
The secondary vertex reconstruction algorithm SecVtx [47] searches for silicon tracks within the $\Delta R < 0.4$ cone of a jet. These tracks must fulfill certain criteria, such as to have $p_T > 0.5$ GeV/c or not to exceed an impact parameter of 0.15 cm to the primary interaction point. SecVtx then tries to reconstruct a secondary vertex with these tracks. In order to be attributed a b tag, the secondary vertex must pass further quality cuts, e.g. the invariant mass of the vertex must not correspond to the masses of long lived light flavor hadrons, such as the K_S or Λ and the vertex is demanded to have $S_{L_{xy}} > 7.5$, where $S_{L_{xy}}$ is the significance of the distance of the
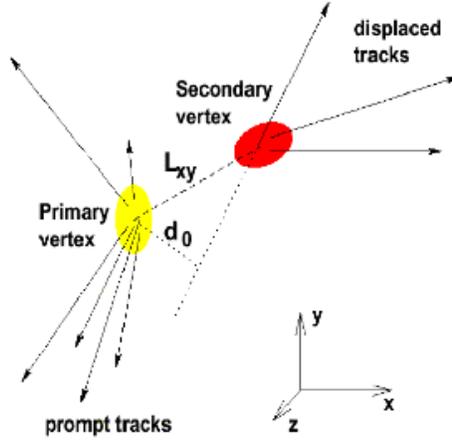
Figure 3.2: Schematic view of an event with a secondary vertex. The secondary vertex has an impact parameter d_0 to the primary interaction point. L_{xy} is the distance of the secondary to the primary vertex in the η-ϕ plane.

secondary to the primary vertex in the η-ϕ plane L_{xy}, defined as $S_{L_{xy}} \equiv |L_{xy}/\sigma_{L_{xy}}|$. A summary of all track and vertex requirements can be found in reference [48].

Missing Transverse Energy

Since neutrinos are subject only to the weak interaction, which renders their detection very difficult, they cannot be measured directly in the detector. Their transverse energy can however be inferred indirectly, from the energy deposits of the other particles in the event and from the fact that the transverse energy distribution of an interaction must be symmetric around the beampipe. The missing E_{T} ($\vec{\not{E}}_{\mathrm{T}}$) is thus defined by

$$\vec{\not{E}}_{\mathrm{T}} = -\sum_i E_{\mathrm{T}}^i \hat{n}_i, \tag{3.1}$$

where i denotes the calorimeter tower number with $|\eta| < 3.6$, $\hat{n}_i$ is a unit vector perpendicular to the beam axis which points at the ith calorimeter tower. To compensate mismeasurements in the calorimeter, jet corrections are incorporated int the calculation of $\vec{\not{E}}_{\mathrm{T}}$. Since muons pass the calorimeters without showering, i.e. as minimum ionizing particle, a correction is applied by adding all transverse momenta of the traversing muons to the sum and by removing the average ionization energy. To ensure that there really is a neutrino in the event, the corrected $\not{E}_{\mathrm{T}}$ is required to be greater than 25 GeV.

Rejection of QCD Multi-Jet Background

There still remain events that contain in fact not a real W boson, but are really QCD events, where jets are lost or mismeasured leading to apparent missing E_T, where jets fake a lepton or where the lepton stems from the decay of heavy quarks. For these events one can expect small $\not{E}_\mathrm{T}$, small $\not{E}_\mathrm{T}$ significance $\not{E}_\mathrm{T sig}$, a small transverse W boson mass $M_\mathrm{T}(W)$, and small values of the angle $\Delta\phi_{\vec{\not{E}}_\mathrm{T},\mathrm{jet}}$ between $\vec{\not{E}}_\mathrm{T}$ and a jet are expected. It is therefore possible to impose cuts on these variables to exclude more of these QCD events. An example of such a cut is shown in figure 3.3.
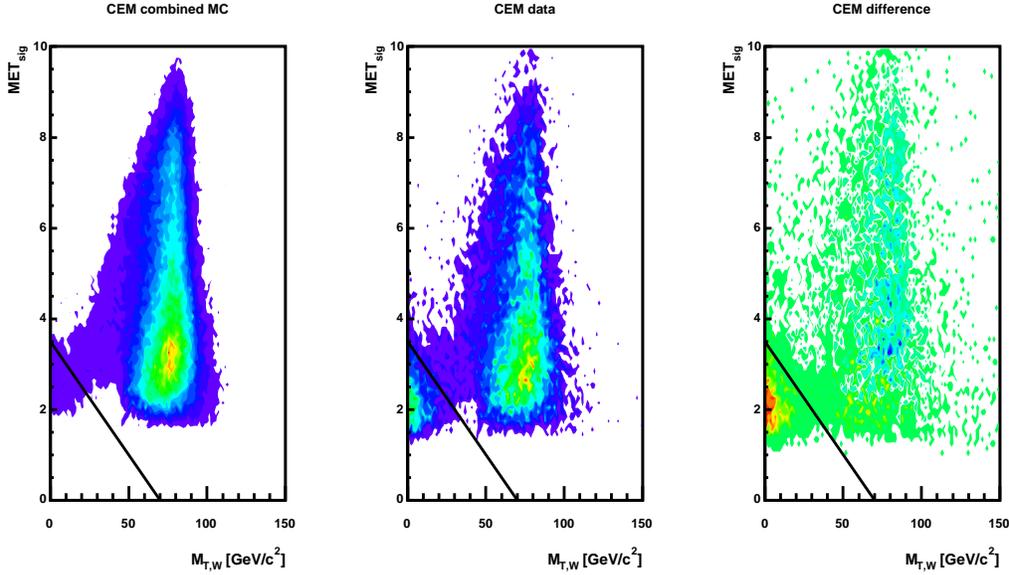
Figure 3.3: Illustration of the CEM QCD veto in the 2-jet bin. The distribution of $\not{E}_\mathrm{T sig}$ versus $M_\mathrm{T}(W)$ is shown in the pretag sample. On the left hand side, the distribution of the W+jets sample is shown. The plot in the middle shows the distribution of pretag data. On the right hand side, the difference between the other two distributions is shown. The line represents the cut $\not{E}_\mathrm{T sig} > -0.05\,\mathrm{GeV}\cdot c^2 \cdot M_\mathrm{T}(W) + 3.5\,\mathrm{GeV}$. This cut rejects events which are not modeled by the W+jets sample.

The $\not{E}_\mathrm{T}$ significance is defined as

$$\not{E}_\mathrm{T sig} = \frac{\not{E}_\mathrm{T}}{\sqrt{\sum_\mathrm{jets} C_\mathrm{JES}^2 \cos^2(\Delta\phi_{\vec{\not{E}}_\mathrm{T},\mathrm{jet}}) + \cos^2(\Delta\phi_{\vec{\not{E}}_\mathrm{T}^\mathrm{uncorr},\vec{\not{E}}_\mathrm{T}^\mathrm{corr}})}} \tag{3.2}$$

with the level 5 jet-correction factor C_JES and the azimuthal angle $\Delta\phi_{\vec{\not{E}}_\mathrm{T}^\mathrm{uncorr},\vec{\not{E}}_\mathrm{T}^\mathrm{corr}}$ between uncorrected and corrected $\vec{\not{E}}_\mathrm{T}$. The transverse W boson mass is given by

$$M_\mathrm{T}(W) = \sqrt{2p_\mathrm{T}^\ell \not{E}_\mathrm{T} - \vec{p_\mathrm{T}}^{\,\ell} \cdot \vec{\not{E}}_\mathrm{T}} \tag{3.3}$$

with $p_\mathrm{T}{}^\ell$ being the transverse momentum of the charged lepton.

Events passing the electron trigger must have $M_\mathrm{T}(W) > 20$ GeV. For central electrons, it is additionally required that $\displaystyle{\not}E_\mathrm{Tsig} > -0.05$ GeV $\cdot\, c^2 \cdot M_\mathrm{T}(W) + 3.5$ GeV. In the 2- and 3-jet bin, $\displaystyle{\not}E_\mathrm{Tsig} > 2.5$ GeV $- 2.5|\Delta\phi_{\vec{\not}E_\mathrm{T},\,\mathrm{jet2}}|/0.8 \cdot$ GeV must be fulfilled, with the azimuthal angle $\Delta\phi_{\vec{\not}E_\mathrm{T},\,\mathrm{jet2}}$ between $\vec{\not}E_\mathrm{T}$ and the second leading jet. In the 1-jet bin, $\displaystyle{\not}E_\mathrm{Tsig} > -7.6$ GeV $+ 3.2|\Delta\phi_{\ell,\,\mathrm{jet}}|/0.8 \cdot$ GeV must be satisfied, with $\Delta\phi_{\ell,\,\mathrm{jet}}$ being the azimuthal angle between the charged lepton and the jet.

Z Boson Veto

To exclude events whose decay products really stem from the production of a Z boson, events are rejected if it is possible to combine the tight lepton with a tight jet, a loose jet or a loose lepton to form an invariant mass in the proximity of the Z boson mass ranging from 76 GeV$/c^2$ to 106 GeV$/c^2$.

z Vertex Cut

The primary interaction point is required to lie within $\pm$ 60 cm of the center of the detector.

3.2 Signal and Background Simulation

Even after the careful reconstruction and the application of the cuts described in the previous section, there remain events that have the same signature as single top quark events, but must be attributed to background processes. In order to measure the single top cross section, it is therefore crucial that both the signal and the background processeses are modeled correctly. Most of these processes are described with Monte Carlo (MC) methods, a couple of background processes are derived from data.

Matrix element generators are used to simulate hard particle interaction. ALPGEN [49] is a generator for processes occurring in hadron collisions. MADEVENT is a multi-purpose tree-level generator [49]. Simulated events produced with a matrix element generator still have to be showered and hadronized. The showering adds higher order effects by allowing the partons to split into $q\bar{q}$ or gg pairs and by taking gluon emission from quarks into account. Since these events still contain bare quarks and gluons in the final state, they must then undergo a hadronization that groups them together in color neutral hadrons. The generators PYTHIA [50] and HERWIG [51] were used for this purpose in this analysis.

MC events have furthermore to be subjected to the detector simulation, in order to

render them comparable with data. The detector simulation models the response of the CDFII detector to the events produced with the MC. The simulated particles are turned into observable signals in the detector. For this purpose the GEANT3 package [52] is taken. The reconstruction of signals in the drift chamber (COT) is done with the GARFIELD [53, 54] software. The showering of particles inside the calorimeters is simulated with GFLASH [55].

3.2.1   Signal MC

The Feynman diagrams in figure 1.10 and 1.9 represent contributions for different orders in perturbation theory. In case of the single top quark t-channel production it is not sufficient to take only the leading order process, the so-called $2 \rightarrow 2$ process, but the next to leading order process, the $2 \rightarrow 3$ process, must be included as well. Events of the leading order process can be calculated with a matrix element generator. By applying a parton showering on these events next to leading order processes are partially modelled, e.g. by creating the $\bar{b}$ quark through backward evolution of the DGLAP [56, 57, 58] equation. However the events produced in this manner do not adequately describe the expected distributions of the $\bar{b}$ quark. Only the soft part of the transverse energy distribution of the $\bar{b}$ quark is well modeled, whereas the high p_T tail lacks a correct description. The pseudorapidity expands to far in the forward region.

For this reason the leading and next to leading order processes are generated separately. For the matrix element generation the MC generator MADEVENT was taken, using the CTEQ5L parameterization of the parton distribution function. The showering and hadronization is done with PYTHIA.

The two produced samples have then to be taking together to form one unified single top t-channel sample. This is done by the so called matching procedure. Hereby the fraction of $2 \rightarrow 2$ and $2 \rightarrow 3$ processes is varied until the number of events with a detectable 2nd b quark ($p_T > 20 \, \text{GeV}/c$ and $|\eta| < 2.8$) corresponds to the one predicted by ZTOP (a program to calculate NLO s-channel and t-channel single top quark production distributions). Below the point, where the p_T of the 2nd b quark is equal for the $2 \rightarrow 2$ and $2 \rightarrow 3$ processes, only the $2 \rightarrow 2$ are taken into the matched sample, above only the $2 \rightarrow 3$ are taken. It is found that the ratio of the $2 \rightarrow 2$ and the $2 \rightarrow 3$ process is $R = 2.1$, the value for the p_T cutoff is 20 GeV/c. It follows that the detectable 2nd b quarks are all described by the $2 \rightarrow 3$ process. The p_T distribution for both samples is shown in figure 3.4, a more detailed description of the matching process can be found in reference [59].

3.2.2   Background Processes

Several background processes survive the cuts described in section 3.1, either because they match the single top signature or because of mismeasurements. They have to be modeled properly as well. The matrix elements of the processes W + heavy flavor and Z + heavy flavor were generated with ALPGEN. For the other processes
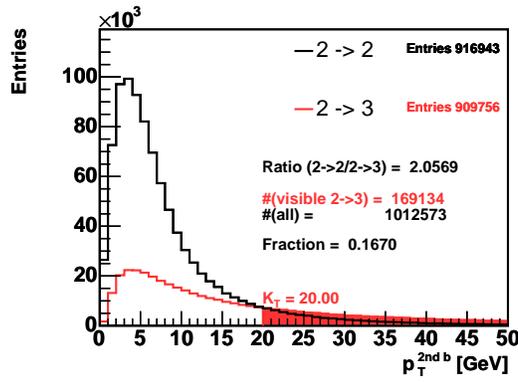
Figure 3.4: Matching of single top quark events produced by the $2 \to 2$ and the $2 \to 3$ t-channel processes. The p_T distributions of the 2nd b quark in the event are shown. The ratio of $2 \to 2$ to $2 \to 3$ events is adjusted such that the rate of 2nd b quarks with $p_T > 20 \, \text{GeV}/c$ and $|\eta| < 2.8$ matches the NLO prediction. The fraction of these events is illustrated by the shaded area.

PYTHIA was taken. To be consistent all processes were showered and hadronized with PYTHIA. The following background processes occur:

W + heavy flavor

The W + heavy flavor background consists of the processes $Wb\bar{b}$, $Wc\bar{c}$ and Wc. Feynman diagrams of these processes are shown exemplarily in figure 3.5. In these processes a leptonically decaying W boson is produced along with the heavy quark flavors b and c.

(a) $Wc\bar{c}$, $Wb\bar{b}$

(b) Wc

Figure 3.5: Feynman diagrams contributing to the W + heavy flavor processes.

W + heavy flavor represents the major background in the two jet bin.

The W + heavy flavor events consist of events with matched matrix-element hard-partons and shower-generated jets, following the so called MLM matching [60, 61]: At first, parton level configurations are generated for a given hard parton multiplicity. The partons are constrained by $p_T > p_T^{\text{min}}$ and $\Delta R_{\text{jj}} > R^{\text{min}}$ cuts. The

jet showering is then performed using the default PYTHIA algorithms. Before the hadronization the showered events have to be processed with a cone jet algorithm, defined by $E_{\mathrm{T}}^{\min}$ and R_{jet}, that clusters the partons. The partons can then be matched to the jets. For each parton the jet with the smallest $\Delta R_{\mathrm{j,parton}}$ to the parton is selected. If $\Delta R_{\mathrm{j,parton}} < \Delta R_{\mathrm{jet}}$ the parton is matched. Furthermore a jet can only be matched to a single parton. For the inclusive sample, the event is kept if all partons are matched, otherwise it is discarded. For the exclusive sample it is moreover necessary that the number of jets equals the parton multiplicity.

Top Quark Pair Production

The strong production of a top-antitop pair is a background, particularly important in the three jet bin. There are three different configurations of possible final states of this process, see Feynman diagrams figure 3.6.

(a) leptonic decay (b) semileptonic decay (c) hadronic decay
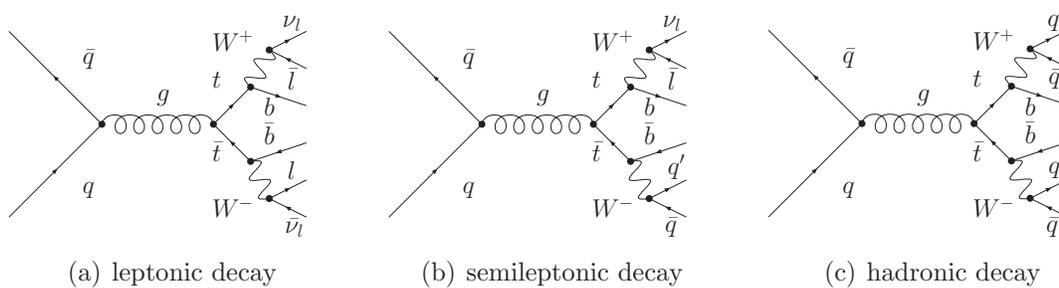
Figure 3.6: Possible $t\bar{t}$ decays.

An event where both W bosons decay leptonically (dilepton event) can match the signal signature if one of the leptons is lost or else is taken for a jet. In the case of a semi-leptonic decay one W boson decays into a lepton and a neutrino. If a number of the produced jets is lost or qualified as loose, $t\bar{t}$ events can be observed in the two or more likely the three jet bin. The all-hadronic case hardly occurs.

Diboson

In diboson events two massive vector bosons are produced; as in the case of $t\bar{t}$ production they can decay into leptons or hadrons. Different Feynman diagrams contributing to these processes are shown exemplarily in figure 3.8.
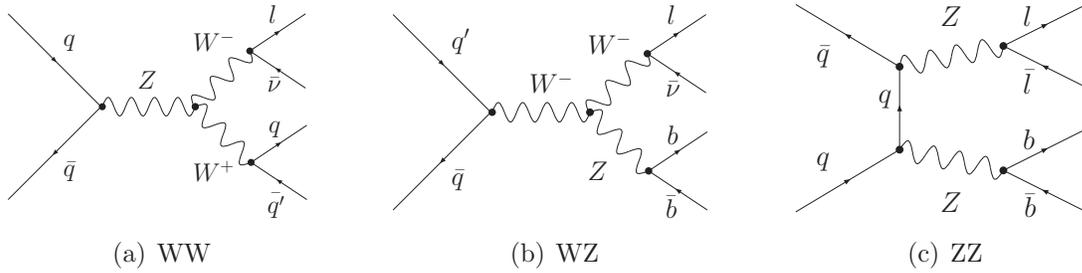
Figure 3.7: Processes involving two weak gauge bosons

Z + jets

The Z + jets background designates those processes, where a Z boson is produced along with a jet. A couple of Feynman diagrams contributing to these processes are shown in figure 3.8.

Figure 3.8: Feynman diagrams contributing to the production of Z+jets

Mistags

Events containing a W or Z boson in association with light flavored jets (u, d, s) can fake a single top event, if one of these jets is falsely assigned a b-tag (called mistag). The events are similar to the W + heavy flavor ones, but with a light instead of a heavy flavored quark. Only a small fraction of light jets are mistagged. A typical mistag rate is 0.5%. Therefore it is quite challenging to produce a sufficient sample of mistagged W+light events. A lot of MC events would therefore have to be generated. Instead the pretag sample is used. In the pretag sample, taggable jets are assigned to be tagged. A taggable jet has $E_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.4$, and more than two tracks ($N_{\mathrm{trk}} \geq 2$). If an event has more than one taggable jet, the probability that a specific jet is selected as the tagged one is given by the ratios of the mistag probabilities. This probability is defined by the negative tag rate and the correction factor for the mistag asymmetry. Additionally, each event is weighted by the mistag probability of the jet considered as tagged [62].

QCD

QCD events do not contain any on-shell W boson, but can appear to have the same signature. They are therefore also referred to as non-W events. Typical Feynman diagrams are shown in figure 3.9. Similar to the $t\bar{t}$ background these events can match the single top signal signature if a jet fakes a lepton.

<div align="center">(a)   (b)</div>

Figure 3.9: Possible QCD processes: in figure 3.9(a) shows a $b\bar{b}$ event where one b decays leptonically, the other hadronically, in figure 3.9(b) a semileptonically decaying gluon pair is produced

Two different models exist to describe these events. Both are taken from measured data.

For the central region of the detector, events with electrons are taken from central electron trigger data to simulate the non-W events. These electrons are required to fail two of the five non-kinematic selection criteria, but must pass the isolation cut. Those events have the kinematic properties of W boson events, but are not non-W enriched, since the kinematic cuts are applied to weed out QCD multijet events. This non-W model is used to simulate electron triggered QCD events as well as muon events.

For the forward region no such anti-electron non-W model exists.

Additionally, jet trigger data is used. Since QCD events will only pass the single top cuts if a jet is taken for an electron, the events are required to have a jet with $E_T > 20$ GeV and $0.05 < E_{HAD}/E_{EM} < 0.2$. In order to avoid an electron enriched sample, these jets must have at least four reconstructed tracks. This jet-electron sample is taken and one jet per event is identified as a fake electron. This data is doubled to simulate electron and muon QCD events. Since requiring a tagged jet in non-W events would result in a very low statistic, only a taggable jet is requested. If there is only one taggable jet in the event, it is considered as the tagged jet. In events with more than one taggable jet, one of them is chosen randomly to be the tagged jet. In order to be able to use the neural network flavor separator (described in section 5.1.1), it is additionally necessary to assign a hypothesis of what kind of quark flavor the jet is, b, c, or light quark flavor. The probability that a specific quark flavor is assigned to a certain jet is given by the expected flavor composition of the non-W background. This composition is estimated by applying the neural

network flavor separator to the $\not{E}_\mathrm{T} < 15$ GeV sideband of the observed data. In this sideband sample, a flavor composition of 45% b quark jets, 40% c quark jets, and 15% light quark jets is found [63].

Heavy Flavor Overlap Removal During the showering of MC events it can happen that in an originally light flavored sample, heavy quarks will appear. These events have to be moved, since it is important (for the neural network flavor separator for example) to have pure samples. In this analysis a jet based heavy flavor overlap removal is applied: Heavy quark pairs $b\bar{b}$ and $c\bar{c}$ are moved to a heavy flavor sample if they lie within $\Delta R < 0.4$ of the cone of a jet, since in that regime the showering gives a better description of the observation; otherwise the events are completely removed.

3.3 Event Yield

For a process that is theoretically well understood the number of expected events can be derived from the predicted cross section:

$$\hat{\nu} = \sigma \cdot \varepsilon_\mathrm{evt} \cdot \mathcal{L}_\mathrm{int} \tag{3.4}$$

where ε_evt is the event detection efficiency, and $\mathcal{L}_\mathrm{int}$ is the integrated luminosity. The event detection efficiency is given by:

$$\varepsilon_\mathrm{evt} = \varepsilon_\mathrm{mc} \cdot \varepsilon_\mathrm{BR} \cdot \varepsilon_\mathrm{corr} \cdot \varepsilon_\mathrm{trig} \tag{3.5}$$

The event detection efficiency is estimated by applying the cuts described in section 3.1 to the samples of simulated events. This aspect is comprised in the factor ε_mc. Sometimes only a subprocess is considered in this analysis, e.g. in single top production the W boson is required to decay leptonically. ε_mc is therefore multiplied with the appropriate branching ratio ε_BR. Differences in the identification efficiencies of charged leptons and b quark jets between data and simulation are accounted for by a correction factor, $\varepsilon_\mathrm{corr}$. $\varepsilon_\mathrm{trig}$ designates the trigger efficiency. Processes with larger theoretical uncertainties are estimated from pretag data. For this analysis the estimated proportions of signal and background processes were obtained with the so called "Method 2 for you" [64, 65]. They can be inferred from table 3.2.

Process	Number of Events			
	2 jets		3 jets	
	1 tag	2 tags	1 tag	2 tags
$t\bar{t}$ dilepton	30.6 ± 4.3	8.4 ± 1.4	24.1 ± 3.4	8.1 ± 1.3
$t\bar{t}$ non-dilepton	64.2 ± 9.0	12.7 ± 2.1	180.0 ± 25.1	52.2 ± 8.6
total $t\bar{t}$	94.8 ± 13.3	21.1 ± 3.5	204.1 ± 28.5	60.3 ± 9.9
$Wb\bar{b}$	376.2 ± 113.4	49.7 ± 15.5	106.7 ± 32.2	17.6 ± 5.5
$Wc\bar{c}/Wc$	361.4 ± 111.4	4.8 ± 1.6	92.7 ± 28.5	2.4 ± 0.8
total W+heavy flavor	737.6 ± 224.8	54.5 ± 17.1	199.4 ± 60.7	20.0 ± 6.3
Mistags	308.3 ± 51.1	1.2 ± 0.4	88.6 ± 14.8	0.9 ± 0.3
Non-W	55.8 ± 22.3	1.5 ± 0.6	21.3 ± 8.5	0.2 ± 0.1
WW	37.2 ± 4.1	0.2 ± 0.0	12.4 ± 1.4	0.2 ± 0.0
WZ	14.8 ± 1.1	2.9 ± 0.3	4.1 ± 0.3	0.9 ± 0.1
ZZ	0.4 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	0.0 ± 0.0
total Diboson	52.4 ± 5.2	3.2 ± 0.3	16.7 ± 1.7	1.1 ± 0.1
total Z+jets	19.1 ± 2.8	0.9 ± 0.1	7.1 ± 1.0	0.5 ± 0.1
total background	1268.0 ± 319.5	82.4 ± 22.0	537.2 ± 115.2	83.0 ± 16.8
t-channel	50.6 ± 7.4	1.4 ± 0.2	13.1 ± 1.9	2.1 ± 0.3
s-channel	26.3 ± 3.7	7.6 ± 1.2	8.2 ± 1.2	2.7 ± 0.4
total single top	76.9 ± 11.1	9.0 ± 1.4	21.3 ± 3.1	4.8 ± 0.7
total prediction	1345.0 ± 231.9	91.3 ± 17.6	558.7 ± 68.8	87.8 ± 11.6
observation	1312	82	491	95

Table 3.2: Summary of predicted numbers of signal and background events in the selected data sample. All systematic uncertainties are included (see section 6.3).

Chapter 4

Neural Net Analysis Framework

To measure the single top quark cross section, a neural net analysis is conducted. A framework was written to facilitate the procedure of the training, as well as the steps following the training. First an overview of the different parts of the single top quark analysis is given, then the neural network framework is described more closely.

4.1 Structure of the Single Top Quark Analysis

The single top quark analysis is performed on the so called `TopNtuples` [66]. `TopNtuples` are ROOT based trees, containing high-level objects, that provide the complete information on one event. The `TopNtuples` comprise data events and simulated events of signal and background processes.

In figure 4.1 the path of these events in the single top quark analysis is shown. In a first step the `Preselection` is applied on them. The `Preselection` performs all cuts described in section 3.1. The `Preselection` thus leaves only the event information that is useful for the single top quark analysis and writes them to new, in their size dramatically reduced ROOT files, the so called `MiniTrees`. A further program is run on the `MiniTrees`, the `SingleTopReconstruction`. In particular, it effectuates the reconstruction of the top quark. Since there are several possible hypotheses leading to the top quark reconstruction, see section 5.4, a container class, the `SingleTopInterpretation`, stores these different hypotheses, so that all of them can later be accessed in the analysis. The resulting `RecoTrees` are then in a convenient format for the analysis.

`Training Samples` for the training of neural networks can now be made. They are composed of signal and background events in any desired ratio, whereas the different background processes are in the expected proportions, given by the "Method 2" [64, 65].

The actual `Training` of the networks is done with the neural net package Neuro-Bayes® [67, 68].
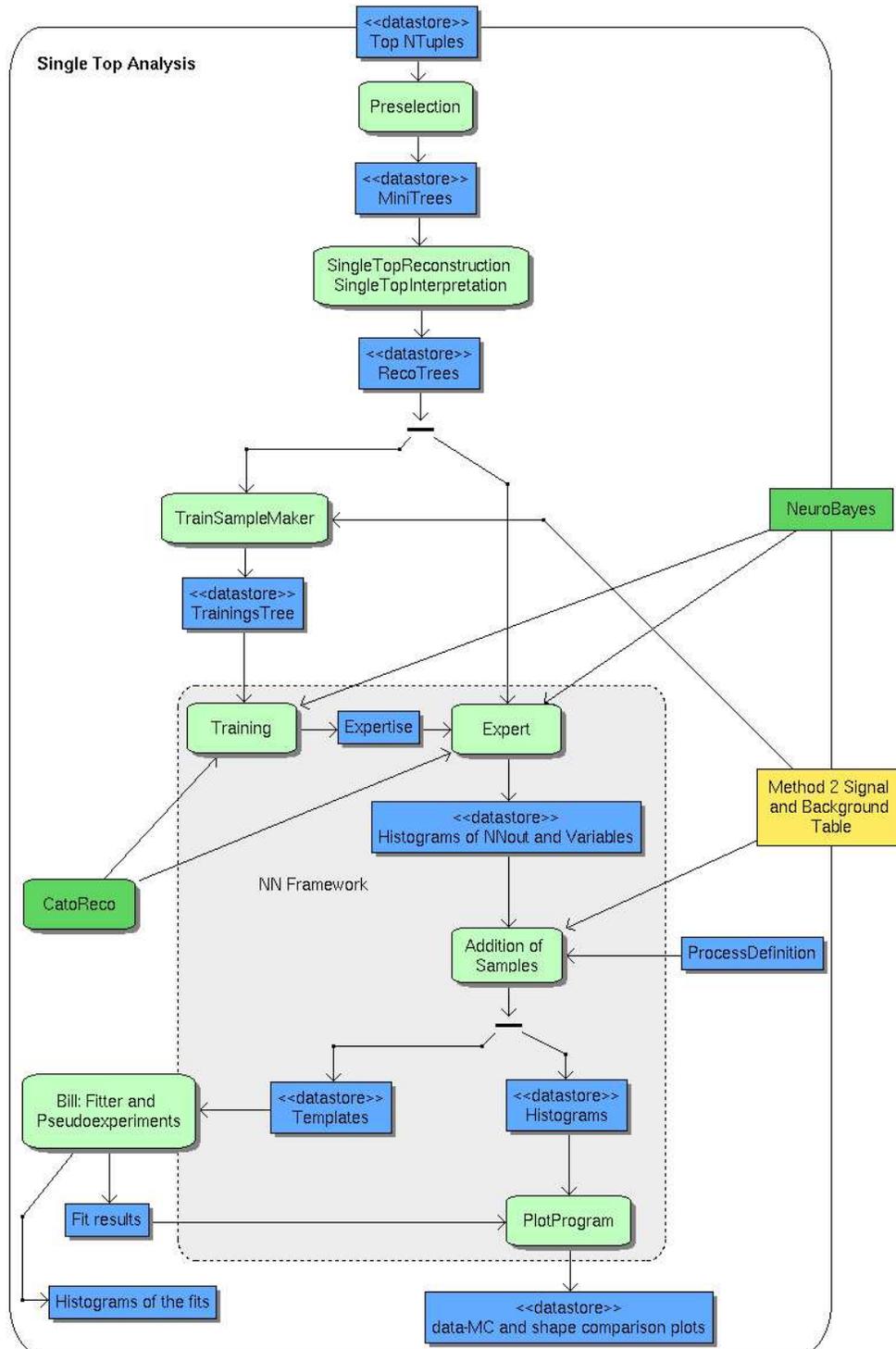
Figure 4.1: UML activity diagram of the single top quark analysis : Overview of the different parts of the single top quark analysis . The green boxes stand for actions, usually a program that is run. The blue boxes designate data stores, where incoming data (information) is stored permanently

The calculation of complex variables, needed for the `Training`, can be done with the `CatoReco` program, that provides functions for more frequently used variables. The result of the training is an `Expertise`, which will be able to classify an event as background- or signal-like. The `Expert` is then run over the `RecoTrees` and ROOT files containing histograms with the neural network output, as well as histograms of all variables used in the training are obtained.
The histograms of a main processes are then added up and normalized to the expected number of events, yielding templates for fitting and histograms ready to be plotted.

4.2 Single Top Neural Network Framework

Due to the large number of neural networks that were planned to train, and that have indeed been trained in this years iteration of the analysis, a neural network framework was created. Its different components are shown in the shaded area in figure 4.1. The idea was to reduce the number of steps necessary for the training of each network, as well as to simplify them, by providing a framework.
A framework achieves:

- an easier handling, since functions are provided for all the main tasks

- an increased efficiency - duplication of code is avoided by providing the framework, it follows that only a minimal number of settings has to be specified

- an improved robustness - reduction of individual code also reduces the number of errors that can be made

- a higher portability - functions written by one person in the group are made available to others through the framework

4.2.1 Overview of the Framework

An overview of the neural network framework is given in figure 4.2. The different classes provided by the framework are displayed.
At the heart of the framework is the `NNframe` class, which will be described in more detail in the next section. It provides the functions necessary for the training and for running the neural network expert. It makes use of the classes `Variable`, `Subdetector` and `Category` which respectively provide the different attributes of a variable, a subdetector and allow to distinguish easily between different samples. For each network, a new `Network` class, that inherits from `NNframe` must be written, containing merely the informations specific for this network. There can be any number of `Network` classes. The training and running the expert can now easily be effectuated with the help of this `Network` class. When the expert has been

run, the resulting files are added up to form samples of main processes or samples comprising several processes with the help of the `AddHistograms` class. The `Plot` class is separated from the rest of the classes. Its purpose is to draw the distributions of processes, obtained with `AddHistograms` in different ways, e.g. normalized to unit area to compare their shapes or normalized to the predicted events in order to compare them to the measured one.
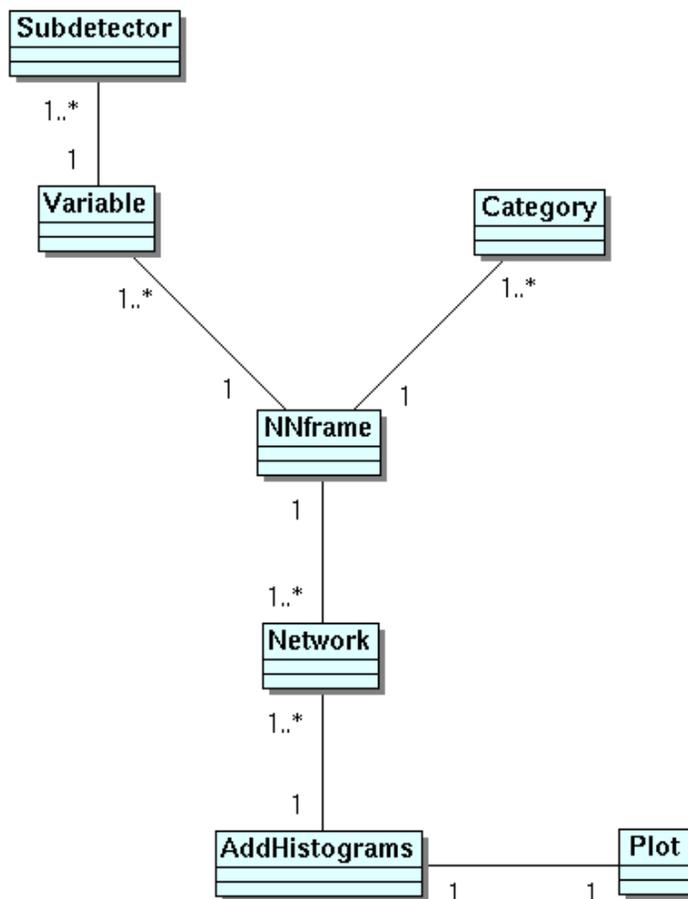
Figure 4.2: UML class diagram of the neural network framework. The blue rectangles are each representing a class. The boxes inside the rectangles are reserved for data members and member functions of the class. They are not specified in this overview diagram. The lines stand for connections among the classes. The numbers designate the multiplicities, for example there can be any number of `Network` classes, but there is only one `NNframe` class.

4.2.2   Infrastructure for the Training of Neural Networks

A more detailed view of the classes, needed for the training of a neural net and for running the expert on the samples is given in figure 4.3. The most important data members and functions are shown exemplarily.
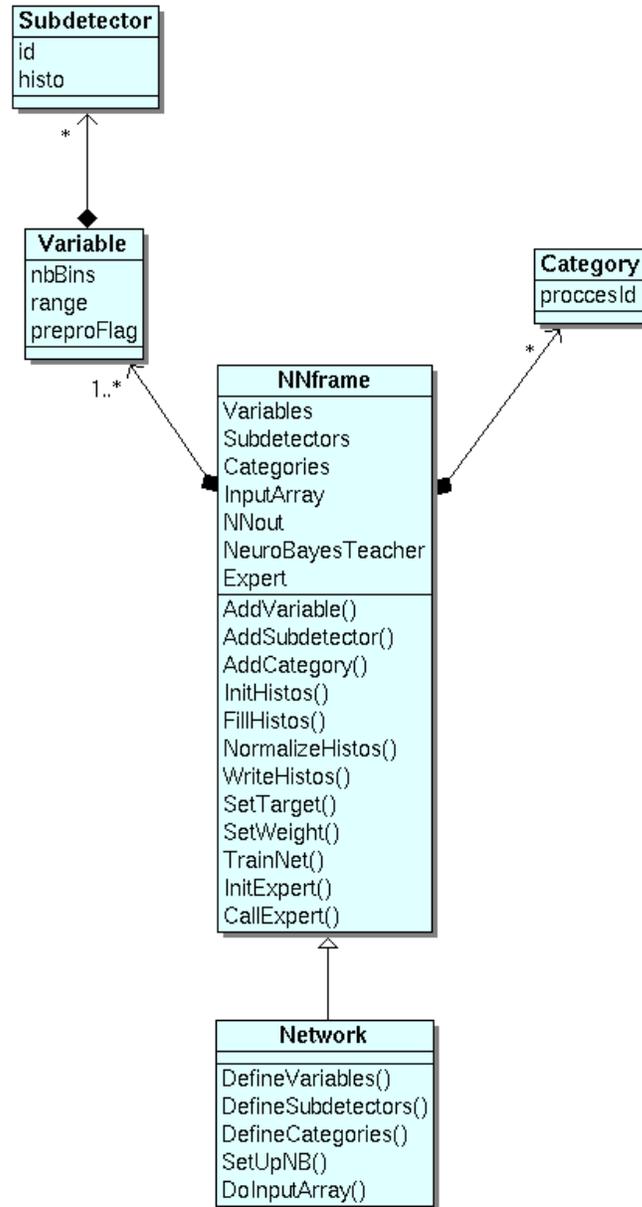
Figure 4.3: Detailed UML class diagram of the part of the neural network framework that is used for the training and for the expert. Data members and member functions of the classes are specified. The line with the diamond defines the relationship of composition for two objects, for example the `NNframe` class can possess any number of instances of the `Variable` class. The line with the open triangle designates a generalization: `Network` inherits from `NNframe`.

Among the data members of the `NNframe` class are container for the `Variables` and the `Subdetectors`. A variable is defined with the help of the `Variable` class. It

provides data members defining histogram specifications such as the binning and the range; another data member contains the preprocessing flag of the variable in the training. A variable and all its parameters can be set with the `AddVariable()` function. A variable furthermore owns instances of the `Subdetector` class. This class contains on its part a histogram, since histograms of the variables must be filled for all subdetectors separately. The `AddSubdetector()` function allows to add the desired subdetectors. The `NNframe` functions `InitHistos()`, `FillHistos()` and `WriteHistos()` are responsible for the creation of the histograms, their correct filling and writing them to ROOT files.

The `Category` class can be used for aliasing the process id's, so that a process (consisting of different samples) can easily be referred to in the training or while running the neural network expert. The declaration of the process id's is done with the `AddCategory()` function.

For every network to be trained, a `Network` class inheriting from `NNframe` is created. It contains the declaration and definition of variables (`DefineVariables()` and `DoInputArray()`), of subdetectors (`DefineSubdetectors()`), and the network topology (`SetUpNB()`). The `Network` class is a container for all the settings that are identical for the training and for the expert.

The interface to NeuroBayes® is done in `NNframe` with an instance of `NeuroBayes-Teacher` for the training and with an instance of `Expert` for running the expert. The `InputArray` must contain the values of the variables in the right order. It is one advantage of the framework, that the `InputArray` does not have to be filled by hand, since it is needed for both the training and the expert, and a mistake can easily be made. Taking the `Network` class for the training and the expert ensures that the `InputArray` is always filled correctly. For the training the function `SetTarget(processname)` sets the target to 1 for a signal event of the trainingstree and otherwise to zero, `SetWeight()` is used for training with weights. `TrainNet()` does the actual training, including the preprocessing. The function `InitExpert()` initializes the expert and `CallExpert()` runs the Expert, that calculates the neural network output for every event. During expertizing, the samples are normalized to their expected cross section with the `NormalizeHistos()` function.

4.2.3   Network Classes used in the Single Top Search

The neural networks trained for the single top quark analysis are shown in figure 4.4. There are five `Network` classes. Two of them are control networks, the others are used to discriminate the single top quark signals from the background. For the pretag networks, network classes have been created as well. Furthermore the neural network framework is used in the search for anomalous single top quark production.

4.2.4   Addition of Samples

After the expert has been run, the samples are still split into subprocesses. The `AddHistos` class permits the addition of samples to processes with the help of a
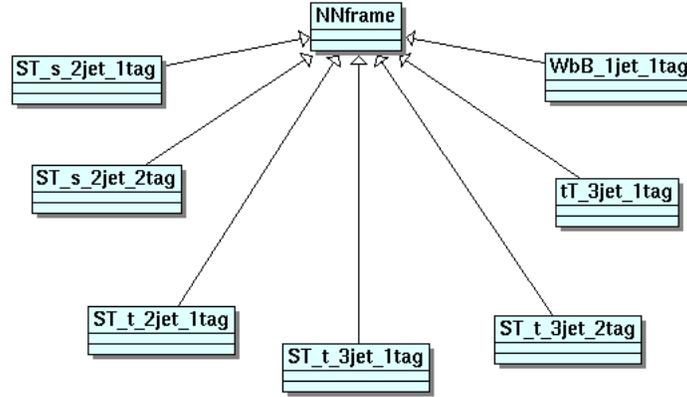
Figure 4.4: UML class diagram of the network classes used in the single top analysis. There are two control networks ($Wb\bar{b}$ in the $W+1$ jet sample with 1 b tag and $t\bar{t}$ in the $W+3$ jets sample also with 1 b tag), and five single top quark networks (t- and s-channel in the $W+2$ jets sample with 1 b tag, s-channel in the $W+2$ jets sample with 2 b tags and two t-channel networks in the $W+3$ jets sample, with 1 or 2 b tags).

configuration file. The configuration file specifies which subprocesses shall be combined to form which process.

Figure 4.5 shows the `AddHistograms` class along with its structures `Process`, `Detector` and `Subprocess`.

First the `VariableNames` of the histograms that shall be added are read from the `Network` class in question, in order to avoid repeating the definition of variables. The same holds for the `DetectorNames`.

The `BuildTemplates()` function is then called to do the addition of the samples that result from running the Expert. It will do the following:

The `ProcessNames` are read from the configuration file with the `ReadProcess()` function. Any number of subprocesses can be defined to form a process. The composition of subprocesses can be different for the subdetectors. This is for example the case for the QCD background model, where the process for the PHX subdetector is composed only of samples of the jet-electron model, while the other subdetectors also include the anti-electron samples, see section 3.2.2.

This is modelled by attributing the `Subprocesses` to the `Detector` structure. The `Detectors` belong on their part to the `Process` structure.

It is during the addition of samples, that the processes are normalized to their expected number of events, which is given by the "Method 2" tables [64, 65]. The `GetBackgroundEstimate()` function reads these tables and writes the factors that are different for every process and different in the subdetectors to the `ScaleFactor` data member of the `Detector` structure. The `AddUpSubProcesses()` function then adds up the histograms of the subprocesses and scales the resulting histograms of the processes to the expected number of events with the help of this `ScaleFactor`. Since they are now scaled correctly, the histograms of the subdetectors can be added
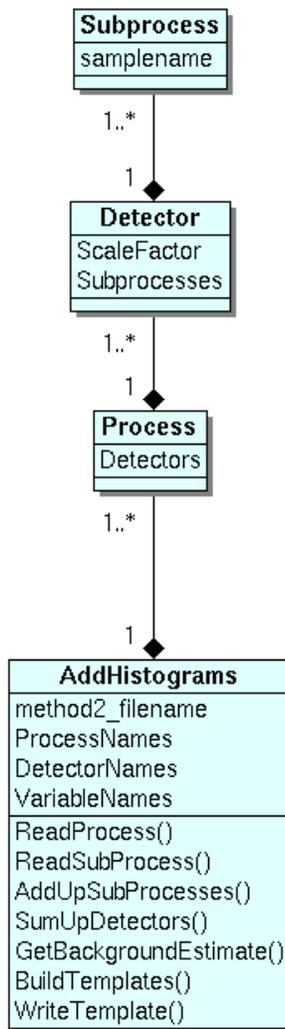
Figure 4.5: Detailed UML class diagram of the part of the neural network framework that is used for the addition and normalization of samples

up with the `SumUpDetectors()` function. `WriteTemplate()` writes the histograms to two different files, one contains the templates that are required to perform the fits, the other is needed to draw the distributions of processes for different comparisons.

Chapter 5

Analysis Tools and Methods

In the search for single top quark production several analysis tools and methods are used. The separation of signal and background events is accomplished with neural networks. The actual measurement of the fraction of single top quark events in the data is done by minimizing a negative log-likelihood function. An important method is the choice of the top quark reconstruction.

5.1 NeuroBayes®

Neural networks are a multivariate method for statistical analyses. They can extract patterns and detect trends that are too complex to be noticed by other techniques. Neural networks therefore represent powerful tools in the domain of classification and prediction.
The neural network package used for this analysis is NeuroBayes® [67]. NeuroBayes® combines a three layer feed forward neural network with a complex and robust pre-processing.

Training

To train a neural net, discriminating variables must be fed into it. An ideal variable would look completely different for signal and background. During the training the neural network learns characterizing differences of the signal and the background sample, and is then able to distinguish between unknown signal and background events on that basis.
The geometry of a three layer neural network can be seen in figure 5.1. In the case of a feed forward network information in the network is passed only in one direction, from the input to the output layer.
The first (the input) layer consists of one node for each variable that is used in
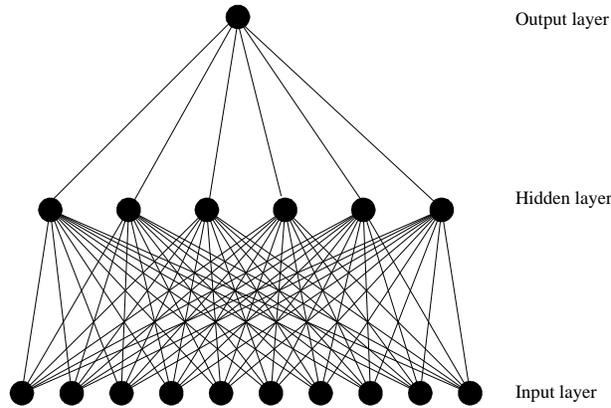
Figure 5.1: General geometry of a three layer neural network.

the training, and of one node, the bias node. The second (the hidden) layer can comprise arbitrarily many nodes. There is one output node in the third layer, that returns a value between -1 and 1, according to whether it attributes an event to the background or to the signal. The nodes imitate the function of neurons in a nervous system such as the brain. To obtain the output of each node j in the hidden layer, a biased weighted sum of the values of the input layer x_i is calculated

$$a_j(\mathbf{x}) = \sum_i \omega_{ij} x_i + \mu_{0,j} \tag{5.1}$$

where $\mu_{0,j}$ is the threshold that has to be exceeded in order for the node to be activated. The weighted sum is passed to a transformed sigmoid function, which yields the output of each node

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

As can be seen in figure 5.2, the sigmoid function is only sensitive to a relatively small range around zero. By this transformation, the interval $[-\infty, +\infty]$ is mapped to the interval $[-1, +1]$. For very large ($x \to \infty$) or very small ($x \to -\infty$) values, a saturation effect is reached. The bias $\mu_{0,j}$ shifts the mean of the sum of the weighted input data distribution $\sum_i \omega_{ij} x_i$ to the linear part of the sigmoid function.

The output of the neural network (of the output node) is calculated in the same way as the outputs of the nodes of the hidden layer:

$$o = S(\sum_{j=0}^{M} \omega_j \cdot S(\sum_{i=0}^{d} \omega_{ij} x_i + \mu_{0,j})) \tag{5.3}$$

where d is the number of input nodes and M the number of hidden nodes. ω_{ij} denotes the weights from the input to the hidden layer, ω_j the weights from the
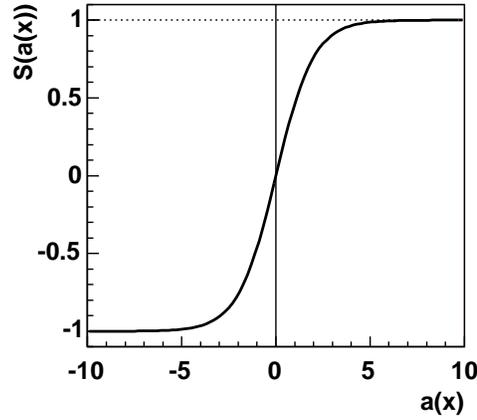
Figure 5.2: The transformed sigmoid activation function S(a(x)) as given by equation 5.2.

hidden layer to the output node. $\mu_{0,j}$ is the weight that connects the bias node with the hidden nodes.

In order to train a neural network a training pattern, consisting of historical or simulated data, where the classification is known, is needed. The deviation between the true output and the one calculated by the neural network is reduced by minimizing the error function, which in the case of NeuroBayes® is the entropy function

$$E_D = \sum_i \log(\frac{1}{2} \cdot (1 + T_i) \cdot o_i + \epsilon), \tag{5.4}$$

where the target value T_i is a binary number to classify event i as signal or background, o_i represents the output as given by equation 5.3. ϵ is a small regularization constant which is introduced in order to avoid numerical problems at the beginning of the training. This constant is reduced in each training iteration and is zero after just a few iterations.

The error functions depends strongly on the combination of weights so that its surface has a complex shape in a multidimensional parameter space. In order to find its minimum the weights are adjusted using the method of steepest descent. The change of each weight $\Delta\omega_{ij}$ is thereby proportional to the current gradient of the error function

$$\Delta\omega_{ij} = -\eta \frac{\partial E_D}{\partial \omega_{ij}} \tag{5.5}$$

where η is the step width (or learning rate), that determines how fast the weights are changed. Since the target value is not known for hidden nodes, the error induced by the current weights has to be propagated backwards from the output node.

The neural network is trained with regularization techniques to improve generalization performance and to avoid overtraining. During the training process, the weights are systematically reduced in addition to the variation calculated by the gradient descent procedure. Thus, only recurring structures are intensified while the influence of statistical fluctuations is reduced by so-called weight decay. Connections

(and even nodes) that have become completely insignificant are pruned away. This reduces the number of free parameters and hence improves the signal-to-noise ratio by removing the cause of the noise, leading to an improved generalization ability. For details of the above mentioned features see references [67, 68].

Preprocessing

In order to find the minimum in the error function the preprocessing, that takes place before the actual training is performed, plays an important role. A *global* preprocessing is done for all variables: Equalizing the input variables and scaling them to be distributed between -1 and 1 before passing the variables to the neural network reduces the influence of extreme outliers. Those flattened distributions are then converted into Gaussian distributions, centered at zero with standard deviation one. At the beginning of the training, this avoids saturation of the nodes due to the above mentioned shape of the activation function (see figure 5.2) and assures that also the inputs to the next layers are distributed with mean zero and width one. To decorrelate the preprocessed input variables, at first, their covariance matrix is calculated. Diagonalizing the covariance matrix using Jacobi rotations [69] and dividing the rotated input vectors by the square root of the corresponding eigenvalue transforms the covariance matrix into a unit matrix.

An *individual* preprocessing of variables can also be performed. For instance the above mentioned transformation to a Gaussian distribution may be altered by fitting a spline curve to the flattened distribution. In addition, discrete variables can be divided into ordered or unordered classes. The preprocessing is also able to deal with variables that are only given for a subset of events by assigning the missing values to a δ function.

Automatic Variable Selection

The significances of the training variables are determined automatically during the preprocessing in NeuroBayes®.

The correlation matrix of all preprocessed input variables is calculated including the correlation of all variables to the target. One variable after the other is omitted to determine the loss of total correlation to the target caused by its removal. The variable with the smallest loss of correlation is discarded leading to an $(n-1)$-dimensional correlation matrix. The same procedure is repeated with the reduced correlation matrix to find the least important of the $(n-1)$ remaining variables.

The significance of each variable is calculated by dividing the loss of correlation induced by its removal at the relevant point of the successive procedure by the square root of the sample size, i.e those significances are relative numbers in terms of the reduced correlation matrices.

After the preprocessing process, it is possible to cut on the significance of the variables to incorporate only those that include relevant information that is not already
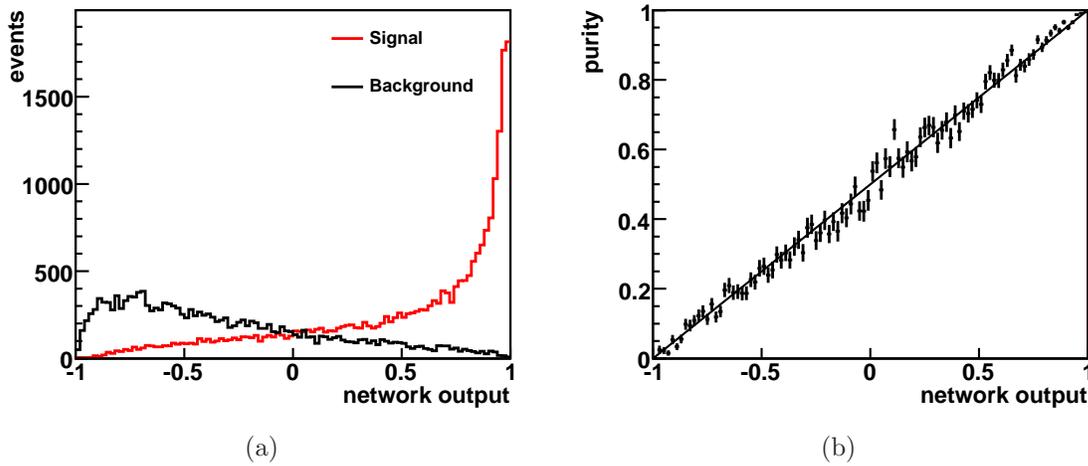
Figure 5.3: Illustration of the training result. Figure (a) Output distribution for signal and background events Figure (b) Signal purity versus network output

incorporated by other variables. The number of discarded variables is determined by scanning the sorted list, starting with the least relevant one, until the first quantity has a significance larger than the required minimum value.

Performance

There are two variables that are used to estimate the performance of a neural network: the purity and the efficiency. The purity is the number of selected signal events compared to all selected events (for a given cut x_{cut} on the output of the neural network):

$$P(NNout) = \frac{N_s(NNout > x_{cut})}{N_s(NNout > x_{cut}) + N_b(NNout > x_{cut})} \tag{5.6}$$

The efficiency is given by the ratio of selected signal events to all signal events of the sample:

$$\epsilon(NNout) = \frac{N_s(NNout > x_{cut})}{N_s} \tag{5.7}$$

Hereby N_s is the number of signal events and N_b the number of background events. As already mentioned above, the network output of signal events piles up at $+1$, while background events accumulate at outputs around -1. This is illustrated in figure 5.3(a). If the network is trained optimally, the network output can be scaled to the interval $[0,1]$ and be interpreted as a Bayesian a posteriori probability, if the a priori probability is correct, i.e. if a realistic mixture of signal and background has been chosen. Furthermore if the global minimum of the error function has been found, the purity is a linear function of the network output as shown in figure 5.3(b).

For a detailed discussion of the output interpretation, see reference [67].
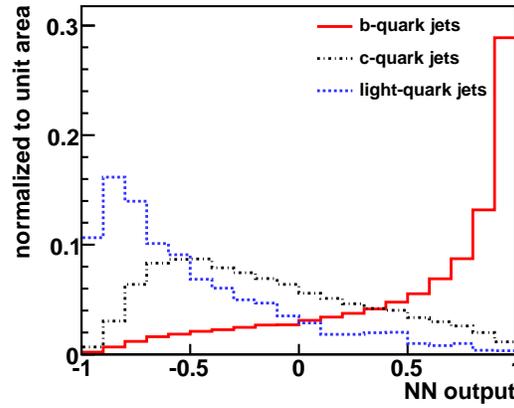
(a)

Figure 5.4: The output of the neural network flavor separator for b quark jets, c quark jets, and light-quark jets with a reconstructed secondary vertex.

5.1.1   Neural Network Flavor Separator

An application of a neural network and important tool in this analysis is the neural network flavor separator [70]. Its purpose is to give the probability for a jet to have a certain flavor. Contrary to the SecVtx tag this probability is not merely a binary number but a continuous one. The information given by the flavor separator is based on the discriminating power of a neural network that has been trained on SecVtx tagged jet samples produced by MC generators. Among the input variables used for the training of the flavor separator are the invariant mass and the decay length of the secondary vertex. The separation of b quark jets from those jets not containing a b quark is shown in figure 5.4. It can also be seen that there is a discrimination between c quark jets and light quark jets.

5.2   Likelihood Function

A binned likelihood fit is done to estimate the number of single top quark events. The used likelihood function has the following structure:

$$
L(\beta_1, \dots, \beta_C; \delta_1, \dots, \delta_S) = \prod_{t=1}^{N_t} \left(\prod_{k=1}^{N_{\mathrm{bins},t}} \frac{e^{-\mu_{k,t}} \cdot \mu_k^{n_{k,t}}}{n_{k,t}!} \right) \cdot \prod_{j=A}^{C} G(\beta_j, 1.0, \Delta_j) \cdot \prod_{i=1}^{S} G(\delta_i, 0.0, 1.0) \,.
$$

(5.8)

The likelihood function comprises Poisson terms for the individual bins of the histograms, Gaussian constraints on the background and Gaussian constraints on the systematic uncertainties.

The first part of the likelihood function consists of Poisson terms for each bin k of each template histogram t in the fit. The inner product multiplies all bins in one

histogram; $n_{k,t}$ is the number of observed events in bin k; $N_{\text{bins},t}$ is the number of bins in the histogram. The outer product loops over the N_t template histograms that are simultaneously fitted. The expectation value $\mu_{k,t}$ of events in a bin k of template histogram t is:

$$\mu_{k,t} = \sum_{j=1}^{C} \beta_j \hat{\nu}_{j,t} \alpha_{jk,t} \tag{5.9}$$

where $\hat{\nu}_{j,t}$ is the predicted expectation value for the number of events of a process j in template histogram t; $\alpha_{jk,t}$ is the fraction of events of the histogram for process j in bin k. Furthermore systematic uncertainties, modifying $\mu_{k,t}$, are included:

$$\mu_{k,t} = \sum_{j=1}^{C} \beta_j \cdot \hat{\nu}_j \cdot \left\{ \prod_{i=1}^{S} (1 + \delta_i \cdot \epsilon_{ji,t}) \right\} \cdot \alpha_{jk,t} \cdot \left\{ 1 + \sum_{i=1}^{S} (\delta_i \cdot \kappa_{jik,t}) \right\} \;, \tag{5.10}$$

δ_i hereby denotes the strength, given in units of one standard deviation, of one of the S systematic effects listed in 6.3. It constitutes an additional fit parameter. The rate uncertainties caused by the systematic effects are designated $\epsilon_{ji,t}$. The shape uncertainties are reflected by relative uncertainties in the bin content of bin k, given by the factor $\kappa_{jik,t}$.

The values of $\kappa_{jik,t}$ are calculated from the systematically shifted normalized template histograms $\alpha_{jik,t}^{+}$ and $\alpha_{jik,t}^{-}$

$$\kappa_{jik,t} = \frac{\alpha_{jik,t}^{+} - \alpha_{jik,t}^{-}}{2\,\alpha_{jk,t}} \;. \tag{5.11}$$

By construction the $\kappa_{jik,t}$ satisfy the normalization condition

$$\sum_{k=1}^{N_{\text{bins},t}} \delta_i \cdot \kappa_{jik,t} = 0 \;. \tag{5.12}$$

The second term imposes Gaussian constraints on the rates of the considered background processes:

$$G(\beta_j, 1.0, \Delta_j) = \frac{1}{\sqrt{2\,\pi\,\Delta_j^2}} \cdot \exp\left(\frac{-(\beta_j - 1.0)^2}{2\,\Delta_j^2} \right) \tag{5.13}$$

The β_j are free parameters of the fit, they are given by the expected value divided by the prediction: $\beta_j = \nu_j / \hat{\nu}_j$. Δ_j designates the expected relative error on the expected number of events. C is the number of processes going into the fit.

The third term contains a Gaussian factor for each systematic uncertainty:

$$G(\delta_i, 0.0, 1.0) = \frac{1}{\sqrt{2\,\pi}} \cdot \exp\left(\frac{-\delta_i^2}{2} \right) \;. \tag{5.14}$$

The single top quark rate is measured by fitting the parameters of the likelihood function β_j and δ_j to the observed data. The negative logarithm of the likelihood function is hereby minimized, with respect to these parameters.

The likelihood function can be calculated as a function of β_1 by minimizing the negative log-likelihood with respect to all other free parameters, yielding a one dimensional function, the reduced likelihood $L_{red}(\beta_1)$.

5.3 Pseudo Experiments and Expected Sensitivity

To determine the sensitivity of the likelihood fit, pseudo experiments are done. For every process j in the fit, the number of events N_j in the pseudo experiment is determined by drawing a random number from a Poisson distribution with mean $\hat{\nu}_j$. N_j numbers are then drawn from the neural network distribution of the templates and filled into a histogram that is taken as neural network output for the pseudo experiment. One pseudo experiment always includes single top quark events at the rate predicted by the standard model, another is done without any single top quark events. The likelihood fit is then performed with the random neural network output of the pseudo experiment.

A hypothesis test is done to estimate the significance of the measured results. Hereby the null hypothesis H_0 supposes that the single top quark cross section is zero ($\beta_1 = 0$), H_1 assumes a cross section as predicted by the standard model ($\beta_1 = 1$). The pseudo experiments yield two distributions, for the so called Q-value, which is defined as:

$$Q = -2\left(\ln L_{\mathrm{red}}(\beta_1 = 1) - \ln L_{\mathrm{red}}(\beta_1 = 0)\right), \tag{5.15}$$

where $L_{\mathrm{red}}(\beta_1 = 1)$ is the value of the reduced likelihood function at the standard model prediction and $L_{\mathrm{red}}(\beta_1 = 0)$ is the value of the reduced likelihood function for a single top quark cross section of zero.

To obtain a measure of how well the observed value Q_0 fits the hypotheses H_0, the p-value is computed. It is given by:

$$p(Q_0) = \frac{1}{I_q} \cdot \int_{-\infty}^{Q_0} q_0(Q')\, dQ' , \tag{5.16}$$

where q_0 is the distribution of Q-values for the null hypothesis H_0 and

$$I_q = \int_{-\infty}^{+\infty} q_0(Q')\, dQ' . \tag{5.17}$$

The expected p-value is defined by $\hat{p} = p(\tilde{Q}_1)$, where $\tilde{Q}_1$ is the median of the Q-value distribution q_1 for the hypothesis H_1. If H_1 is true, it is expected that $p < \hat{p}$ with a probability of 50%.

5.4 Top Quark Reconstruction

Some variables used for the training of the neural networks necessitate a reconstruction of the top quark from the final state particles in the event. The top quark decays into a W boson and a b quark. The event selection has been chosen in a way that only leptonically decaying W bosons are considered. It is thus assumed that the lepton and the missing transverse energy $\not{E}_T$ in the event originate from the W boson decay. In a first step this information is used to reconstruct the z component of the 4-momentum of the neutrino. The following relation exists between the 4-momentum of the neutrino p_ν, the 4-momentum of the lepton p_l and the W boson mass m_W^2:

$$(p_l + p_\nu)^2 = (p_W)^2 = m_W^2 = 80.4\,\text{GeV}/c^2 \tag{5.18}$$

This quadratic equation can be solved for the z-component of the 4-momentum of the neutrino. There are two solutions:

$$p_z^\nu = \frac{\kappa p_z^\ell}{(E_\ell)^2 - (p_z^\ell)^2} \pm \frac{1}{2\left((E_\ell)^2 - (p_z^\ell)^2\right)} \tag{5.19}$$
$$\cdot \sqrt{(2\kappa p_z^\ell)^2 - 4\left((E_\ell)^2(p_T^\nu)^2 - \kappa^2\right) \cdot \left((E_\ell)^2 - (p_z^\ell)^2\right)}$$
$$\text{with } \kappa = \frac{1}{2}(m_W^2 - m_\ell^2) + \cos(\phi_\ell - \phi_\nu) \cdot p_T^\ell p_T^\nu. \tag{5.20}$$

The masses used in this equation are given in section 1.1. If both solutions are real, the one with smallest absolute value is taken, since neutrinos produced in top quark decays are rather central. It can also happen, that the p_z^ν solutions become complex if $\not{E}_T$ has been mismeasured. If one would take only the real part of the solution, equation 5.18 would no longer be fulfilled. In reference [71] it is shown how a physically reasonable solution can be found nevertheless by assuming a mismeasurement of $\not{E}_T$ and doing a minimization that corrects $\not{E}_T$ and leads a real p_z^ν solution.

The energy of the neutrino can then be calculated with $(E_\nu)^2 = (\not{E}_T)^2 + (p_z^\nu)^2$ and the 4-momentum of the W boson is reconstructed to $p_W = p_l + p_\nu$.

A choice has to be made concerning the identification of one of the jets in the event as the jet originating from the b quark of the top quark decay. The event selection always requires that at least one jet has been attributed a b tag. For the reconstruction of t-channel events with exactly one tagged jet, this jet is attributed to the b quark emanating from the top quark. In the t-channel MC sample this is the best possible choice for 75% of the events with two jets and one b tag. For t-channel events with two tags the tagged jet with the largest product of its pseudorapidity times the lepton charge $Q_\ell \cdot \eta_{lj}$ is taken. For s-channel events the jet in the event with the highest $Q_\ell \cdot \eta_{lj}$ is chosen as being the jet coming from the top quark decay [72]. In the s-channel MC sample this is the best possible choice for 60% of the events with two jets and one b tag and for 40% of the events with two jets and two b tags. The mass of the b-quark jet is set to $m_b = 5$ GeV/c^2, its energy is calculated by $(E_b)^2 = (m_b)^2 + (\vec{p})^2$. The top quark 4-vector can then be reconstructed: $p_t = p_b + p_W$.

Chapter 6

Measurement of Single Top Quark Production

For the single top quark analysis many neural networks have been trained. Several of them are control networks such as the one to measure $Wb\bar{b}$ or $t\bar{t}$. In the search for single top quark production four different networks have been trained. As is the case for all physical measurements statistical as well as systematic uncertainties have to be determined and taken into account.

6.1   Control Networks

The purpose of the control networks is to measure a known background process for events with a given number of jets, to demonstrate that it yields the expected event yield, as a cross check for the single top quark networks. Two such control networks have been trained. One measures the number of $Wb\bar{b}$ events with exactly one jet, the other one, the $t\bar{t}$ background for events with three jets and one b tag.

6.1.1 $Wb\bar{b}$ Network

The $Wb\bar{b}$ network is trained on events with one jet. The signal to background ratio of the events used for the training is 50%; the background processes are in the proportions predicted by the "Method 2" [64, 65]. The composition of the training sample is shown in table 6.1.

The variables used for the training can be found in table 6.2. They are ranked by their significance after the preprocessing.

To compare the shapes of the distributions of the different processes for the three most important variables, the distributions are normalized to unit area, as shown on the left-hand side in figure 6.1. To compare measured and predicted distributions

Process	t-channel
Wbb	50.0%
t-channel	1.0%
s-channel	0.0%
$t\bar{t}$	0.5%
$Wc\bar{c}$	10.0%
Wc	13.5%
mistags	23.2%
Diboson	0.5%
Z+jets	1.3%

Table 6.1: Composition of the training sample used to train the $Wb\bar{b}$ neural network in the $W + 1$ jet sample. The training sample contains 105986 events.

Rank	Variable	Relative Significance (in σ)
1	flavor separator output	185.0
2	E_{T} (jet1)	31.8
3	η (loose jet1)	26.9
4	p_{T} ($\ell\nu b$)	14.2
5	$M_{\ell\nu b}$	7.8
6	# loose jets	6.1
7	$\cos\Theta(\ell, \text{beam})$	5.8
8	$\cos\Theta(W, t)$	2.7
9	$\cos\Theta(W, t)_{\mathrm{lab}}$	4.0
10	η_ℓ	4.5
11	$\Delta\phi(\ell, \vec{E}_{\mathrm{T}})$	1.9
12	$M_{\mathrm{T}}(W)$	3.4
13	$\Delta R(\ell, b)$	3.3
14	H_{T}	3.1

Table 6.2: Set of discriminating variables used for the training of the $Wb\bar{b}$ neural network in the $W + 1$ jet sample.

of these variables, the background distributions are normalized to the prediction, as shown on the right-hand side in figure 6.1.

By far the most important variable is the neural network flavor separator. As can be seen in figure 6.1(a) it separates the $Wb\bar{b}$ signal from those processes that do not contain a real b quark, i.e. mistags, $Wc\bar{c}$ and Wc. Since these are the most important background processes, the flavor separator is the variable with the best discriminating power.

The next variable, the transverse energy of the jet E_{T} (jet1), shown in figure 6.1(c), divides $Wb\bar{b}$, from those events where a real top quark was produced, that is $t\bar{t}$ and single top quark production. The energy of the jet is higher for these events, since a larger center of mass energy is required to produce one or two top quarks in the

first place.

The third most important variable is the pseudorapidity of the most energetic loose jet η (loose jet1), shown in figure 6.1(e). It is more central for $Wb\bar{b}$ then for the other processes. The distribution for the t-channel is very different, since the tagged jet in these events can either come from the top quark or the light quark jet, which is produced along the direction of the proton, i.e. in forward direction.

The values of the χ^2 test and the KS test can be inferred from figures 6.1(b), 6.1(d) and 6.1(f), they give the probability that the data describes the predicted distributions. They are mostly greater than 1%.

The template distributions of the neural network output, which are fitted to the observed distribution are shown in figure 6.2(a). There is a clear separation between $Wb\bar{b}$ and the most important background processes $Wc\bar{c}$, mistags and Wc. Processes containing a top quark also look very signal-like, but since the expected number of events in these processes is very small, their contribution is less important. In figure 6.2(b) the expected distributions normalized to the prediction as well as the observed distribution are shown.

The expected uncertainty on the measurement of the $Wb\bar{b}$ cross section is 18.2%, as can be inferred from figure 6.2(c). The results of the fit are given in table 6.3. The fit yields 24% more $Wb\bar{b}$ events than predicted by the "Method 2" [64, 65].

The distributions normalized to the fit result, shown in figure 6.2(d), are thus in good correspondence with the observed distribution.

process	fitted rate / predicted rate
Wbb	1.24 ± 0.09
Mistags	1.00 ± 0.11
$Wc\bar{c}$	0.99 ± 0.35
Wc	0.85 ± 0.14
$t\bar{t}$	1.01 ± 0.21
t-channel	1.00 ± 0.14
s-channel	1.00 ± 0.15
Diboson	1.00 ± 0.10
Z+jets	1.00 ± 0.16
non-W	0.99 ± 0.40

Table 6.3: The fitted rates of the considered processes in the fit of the output of the $Wb\bar{b}$ network.
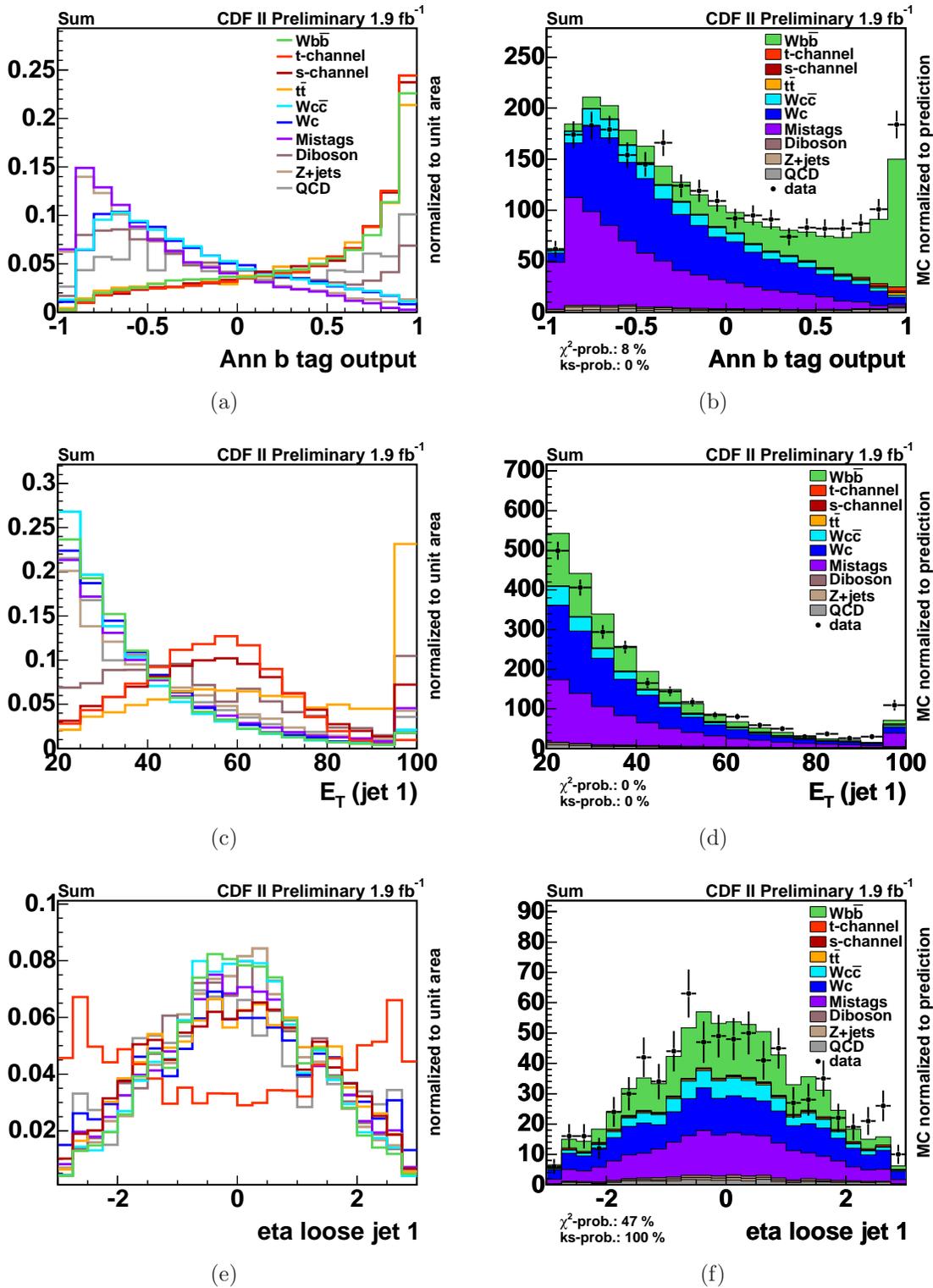
Figure 6.1: The three most important variables in the training of the neural network used to discriminate $Wb\bar{b}$ events with one jet and one b tag are shown:
(a) and (b) the output of the neural network flavor separator,
(c) and (d) the transverse energy of the jet,
(e) and (f) the pseudorapidity of the most energetic loose jet.
The predicted distribution in figures (a), (c) and (e) are scaled to unit area, the ones in figures (b), (d) and (f) are normalized to the number of predicted events.
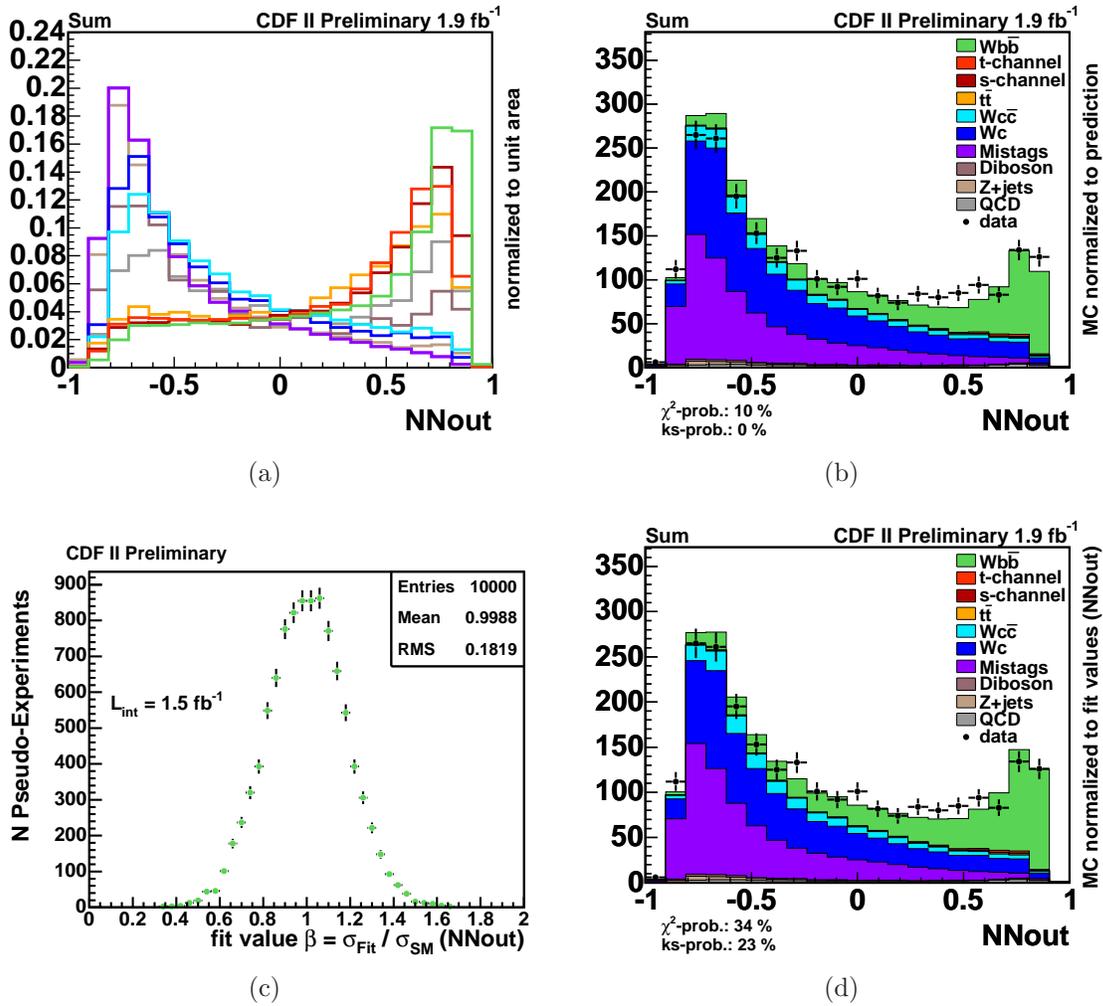
Figure 6.2: Neural network output of the $Wb\bar{b}$ network and illustration of the fit result. (a) depicts the templates which are fitted to the observed distribution. (b) shows the observed and the expected distribution normalized to the prediction. (c) shows the ratio of the fitted cross section divided by the expectation obtained by pseudo experiments and illustrates the expected uncertainty of the measurement. (d) shows the distributions normalized to the fit result.

6.1.2 $t\bar{t}$ Network

As a further cross check $t\bar{t}$ was measured in the $W + 3$ jets sample. Hereby exactly one tagged jet was required. The training sample for the $t\bar{t}$ network is again composed of 50% signal events, while the background proportions correspond to the "Method 2" [64, 65] predictions. The exact composition of the training sample is given in table 6.4.

Process	t-channel
$t\bar{t}$	50.0%
t-channel	2.2%
s-channel	1.4%
$Wb\bar{b}$	13.5%
$Wc\bar{c}$	7.5%
Wc	6.0%
mistags	16.2%
Diboson	2.0%
Z+jets	1.2%

Table 6.4: Composition of the training sample used to train the $t\bar{t}$ neural network in the $W + 3$ jets sample, with one b tag. The number of events in the training sample is 47214.

The variables used in the training, sorted by their significance after the preprocessing, are shown in table 6.5.

A comparison of the shapes of the predicted distributions can be inferred from the left-hand side of figure 6.3, where the predicted distributions normalized to unit area are shown for the three most important variables. On the right-hand side the distributions of predicted events normalized to the number of expected events, as well as the observed distributions are shown.

The most relevant variable is the sum of the transverse energies H_T of all particles produced in the event; it consists of the missing transverse energy, the transverse energy of the lepton, of the tight jets and also of the loose jets, see figure 6.3(a). Only a very high center of mass energy can lead to the production of two top quarks, resulting in a large H_T. In case of the production of a single top quark much less energy is required. For the other processes, that do not contain a top quark in the event, the H_T distributions are shifted even further to lower energies. H_T is thus a variable that distinguishes $t\bar{t}$ from all other processes.

The second most important variable is the neural network flavor separator. It separates $t\bar{t}$ from those processes where no b quark was produced, see figure 6.3(c).

The third variable is the aplanarity. The definition of the aplanarity is based on the normalized momentum tensor

$$M^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |p_i|^2} \tag{6.1}$$

Rank	Variable	Relative Significance (in σ)
1	H_{T}	107.0
2	flavor separator output	65.8
	for b quark from top quark	
3	aplanarity	31.8
4	M_{j2j3}	29.7
5	$p_{\mathrm{T}}\,(\ell\nu bjj)$	26.2
6	M_{j1j3}	20.2
7	$\Delta\eta(j1,j3)$	18.1
8	$\Delta\eta(j1,j2)$	10.2
9	M_{j1j2}	15.3
10	$M_{\mathrm{T}}\,(W)$	11.0
11	$E_{\mathrm{T}}\,(\text{jet } 3)$	10.7
12	$\not{E}_{\mathrm{T}}$	10.3
13	# loose jets	11.5
14	$M_{\ell\nu b}$	10.3
15	η_{ℓ}	8.2
16	$\Delta\eta(j2,j3)$	7.0

Table 6.5: Set of discriminating variables used for the training of the $t\bar{t}$ neural network in the W + 3 jets sample with one tagged jet.

where the Greek indices are spatial components and $\sum_i$ denotes the sum over all particles in the event. The eigenvalues of (6.1) satisfy the normalization constraint $\lambda_1 + \lambda_2 + \lambda_3 = 1$ with $\lambda_1 \geq \lambda_2 \geq \lambda_3$. The aplanarity distinguishes spherical from planar and linear event topologies, it is given by

$$A = \frac{3}{2}\lambda_3. \tag{6.2}$$

Since $t\bar{t}$ events have a more spherical event topology than the other processes, they have a higher aplanarity.

The χ^2 and KS probabilities in figures 6.3(b), 6.3(d) and 6.3(f) show how well the measured data corresponds to the predicted distributions.

In figure 6.4(a) the neural network output for the $t\bar{t}$ network is shown. $t\bar{t}$ is well separated from the major background processes. Figure 6.4(b) shows the observed distribution and the predicted distribution that is scaled to the number of predicted events. The template distributions in figure 6.4(a) are fitted to the observed distribution. The fit results are shown in table 6.6. The expected uncertainty on the measurement of the $t\bar{t}$ cross section is 10.8% as illustrated in figure 6.4(c). In figure 6.4(d) the different templates are scaled to the fit results in table 6.6. The fit results can thus be compared with the observed distributions. They are in good correspondence as can be concluded from the χ^2 and the KS probabilities in figure 6.4(d), that are greater than 1%. The fitted $t\bar{t}$ cross section 7.5 $\pm$ 0.8 pb is in good agreement with the CDFII average of 7.3 $\pm$ 0.5 (stat) $\pm$ 0.6 (syst) $\pm$ 0.4 (lumi) pb [73].

process	fitted rate / predicted rate
$t\bar{t}$	1.12 ± 0.12
$Wb\bar{b}$	0.65 ± 0.27
Mistags	0.96 ± 0.13
$Wc\bar{c}$	0.79 ± 0.33
Wc	0.81 ± 0.32
t-channel	0.99 ± 0.14
s-channel	1.00 ± 0.14
Diboson	1.00 ± 0.10
Z+jets	1.00 ± 0.14
non-W	0.89 ± 0.39

Table 6.6: The fitted rates of the considered processes in the fit of the output of the $t\bar{t}$ network.
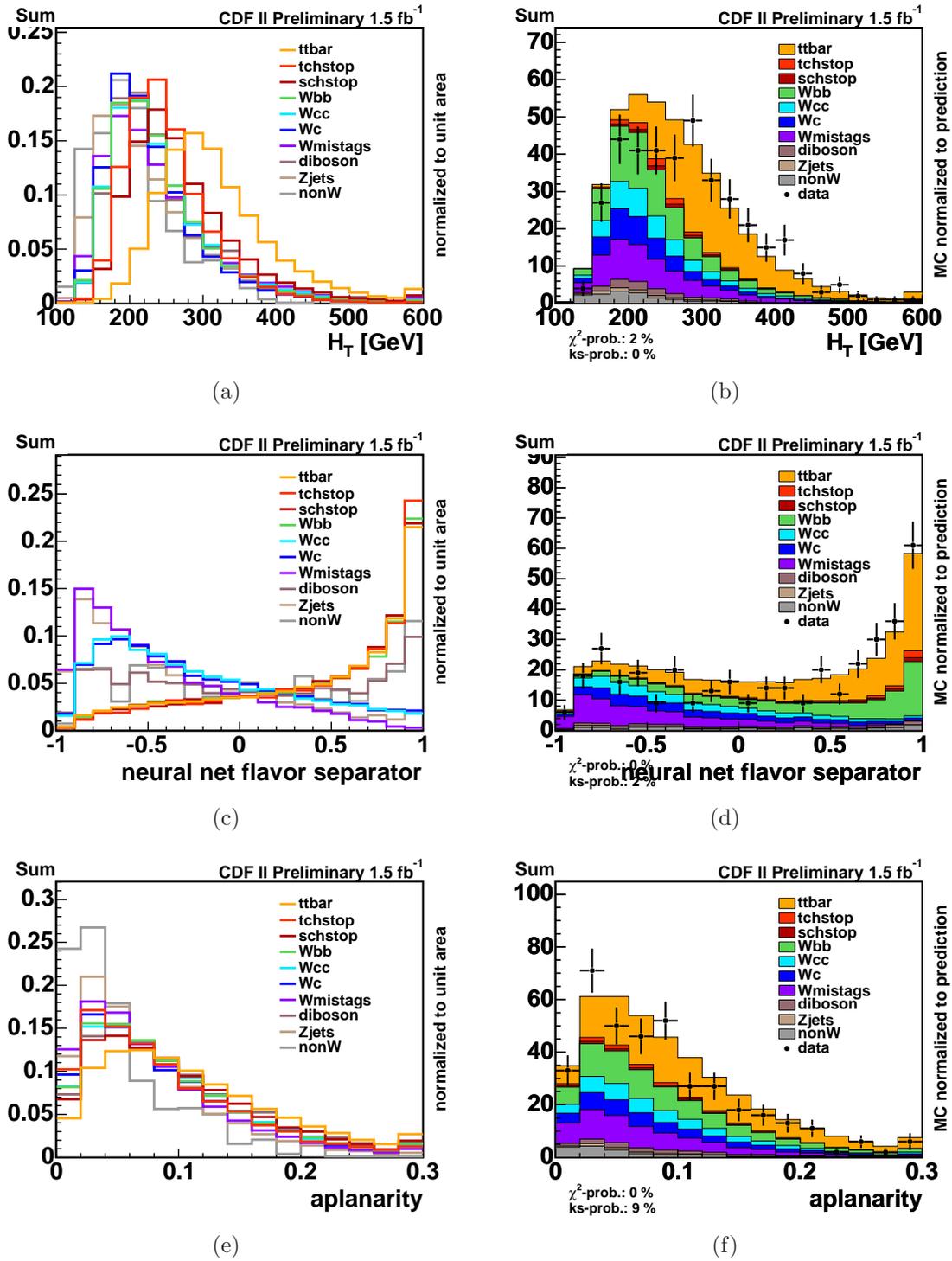
Figure 6.3: The three most important variables in the training of the neural network used to discriminate $t\bar{t}$ events with three jets and one b tag are shown:
(a) and (b) the transverse energy H_T of all particles produced in the event,
(c) and (d) the output of the neural network flavor separator,
(e) and (f) the aplanarity of the event.
The distribtions in figures (a), (c) and (e) are scaled to unit area, the distributions of expected events in figures (b), (d) and (f) are normalized to the number of predicted events.

(a)

(b)

(c)

(d)
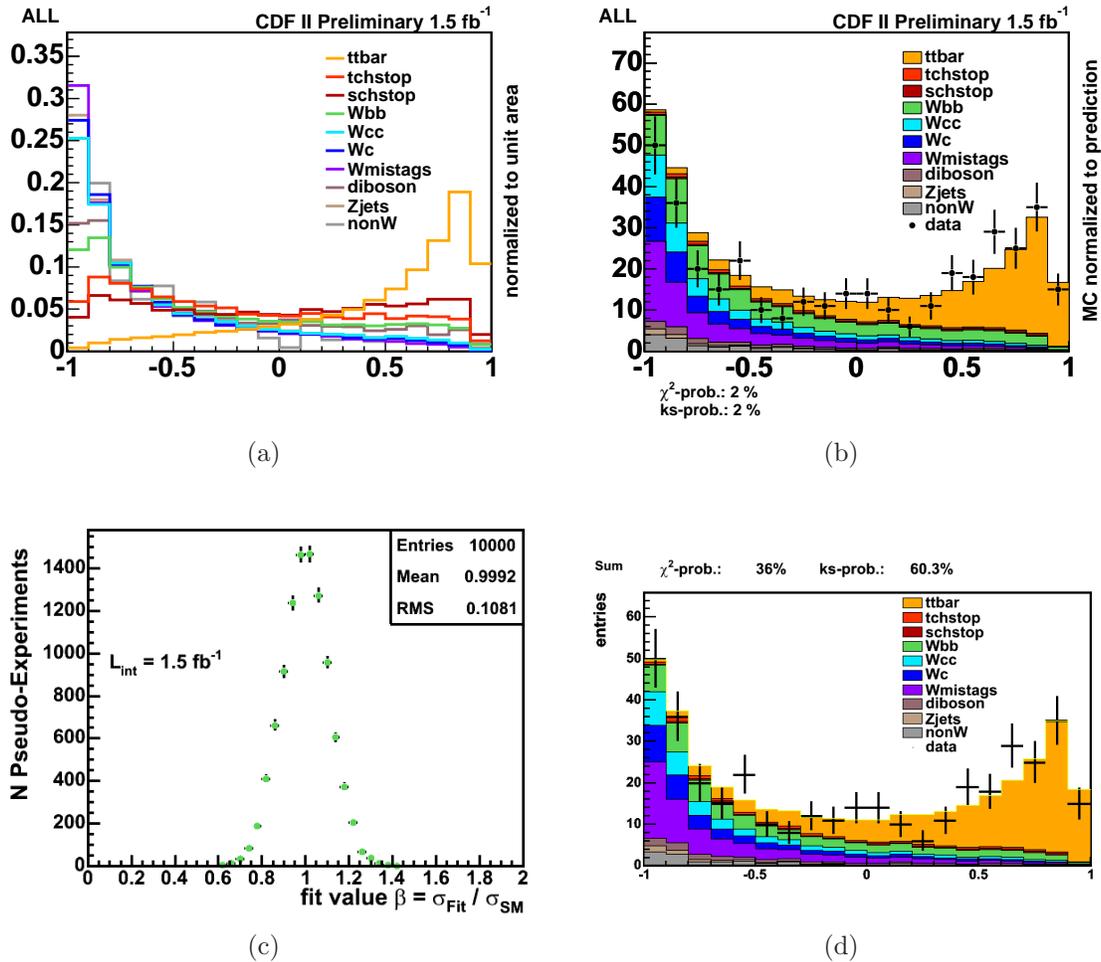
Figure 6.4: Neural network output for $t\bar{t}$ and illustration of the fit results. (a) depicts the templates which are fitted to the observed distribution. (b) shows the observed and the expected distribution normalized to the prediction. (c) shows the ratio of the fitted cross section divided by the expectation obtained by pseudo experiments and illustrates the expected uncertainty of the measurement. (d) shows the distributions normalized to the fit result.

6.2   Single Top Quark Networks

For the combined search, where t- and s-channel single top production are added up in one process with the ratio predicted by the standard model, four neural networks have been trained. One is trained on s-channel events with two jets and two b tags. The others are trained on t-channel events with two jets and one b tag or three jets with one, respectively two b tags.

The signal to background ratio for all networks is again 50%. The background processes are in the proportions given by "Method 2" [64, 65]. The composition of the training samples is shown in table 6.7.

Process	t-channel 2 jets; 1 b tag	s-channel 2 jets; 2 b tags	t-channel 3 jets; 1 b tag	t-channel 3 jets; 2 b tags
t-channel	50.0%	0.0%	50%	50%
s-channel	0.0%	50.0%	0.0%	0.0%
$t\bar{t}$	5.1%	15.9%	22.0%	37.0%
$Wb\bar{b}$	13.7%	27.2%	7.9%	11.2%
$Wc\bar{c}$	6.7%	2.0%	4.3%	0.2%
Wc	7.3%	1.2%	3.5%	1.6%
mistags	14.0%	0.0%	10.0%	0.0%
Diboson	2.4%	2.2%	1.6%	0.0%
Z+jets	0.8%	0.6%	0.7%	0.0%
# events	77296	13501	19899	3346

Table 6.7: Composition of the training samples used to train the neural networks to discriminate single top quark events. The contributions of the respective single top quark events is 50% of the complete training sample. The relative fractions to the background category are given by table 3.2. The absolute fractions are determined by the requirement that the contributions of all background processes sum up to 50%. The number of events used for every training is specified.

t-channel, 2 jets, 1 b-tag   The most relevant input variables for the training of the t-channel network on events with two jets and 1 b tag are listed in table 6.8.

In figure 6.5 the expected distributions of the input variables of the processes normalized to unit area, as well as the expected distributions normalized to data and the observed distribution are shown.

The most important variable for the training is the reconstructed top quark mass, $M_{l\nu b}$, that is the invariant mass of the lepton, the neutrino and the b quark jet. As expected there is a peak at the top quark mass for single top quark and $t\bar{t}$ production, while it is much smaller for processes that do not involve a top quark, see figure 6.5(a). Furthermore since two b quarks are produced in $t\bar{t}$ events, resulting in an ambiguity for the choice of the tagged jet, the $t\bar{t}$ distribution tends to higher reconstructed top quark masses and has a broader shape than the single top quark process.

Rank	Variable	Relative Significance (in σ)
1	$M_{\ell\nu b}$	114
2	flavor separator output for b quark from top quark	86
3	M_{j1j2}	74
4	$Q_\ell \cdot \eta_{lj}$	50
5	$M_T(\ell\nu b)$	29
6	$\cos\Theta(\ell, lj)$	27
7	E_T (light jet)	19
8	$\cos\Theta(\ell, W)$	17
9	η_W	16
10	$M_T(W)$	10
11	$\sum_{j1j2}(\eta_{jets})$	7
12	$p_{T\ell}$	5
13	H_T	5
14	$\cos\Theta(\ell, W)_{lab}$	5

Table 6.8: Set of discriminating variables used for the training of the t-channel neural network on events with two jets and one b tag.

The second most important variable is the output of the neural network flavor separator. As for the $Wb\bar{b}$ and the $t\bar{t}$ network, it is a good discriminator between processes featuring a b quark and those without, thus separating single top quark events from processes not containing a b quark, as can be seen in figure 6.5(c).
The third best discriminating variable is the dijet mass M_{j1j2}, figure 6.5(e). It is again a variable that is different for processes where a top quark has been produced. The χ^2 and KS probabilities in figures 6.5(b), 6.5(d) and 6.5(f) are greater then 1% and thus show that the data is in good agreement with the predicted distributions.

Variables 4 to 6 in the ranking are shown in figure 6.6.
The fourth variable is the product of the charge of the lepton times the pseudorapidity of the light quark jet (the jet that is not associated with the top quark) $Q_\ell \cdot \eta_{lj}$. It is a very asymmetric distribution for the t-channel, which results in an asymmetric distribution for the combined single top process as can be seen in figure 6.6(a). Meanwhile the other processes are centered around zero. The characteristic shape for the t-channel arises from the structure of the proton. The virtual W boson in t-channel events is produced either from an u quark in the proton or from a $\bar{d}$ quark in the antiproton. The latter case is however rarer since there are two up quarks, but only one down quark in the proton. Since the u quark is a valence quark the ensuing light quark jet carries a large momentum fraction of the proton and is therefore emitted in the direction of the proton. One therefore expects that single top quarks are produced in proton direction in roughly 2/3 of the cases, in 1/3 of the cases in the direction of the antiproton. The charge conjugated case of antitop quark production, where the light quark jet goes into the direction of the

antiproton, is compensated by multiplying with the charge of the lepton in the final state.

The next variable is the transverse mass of the reconstructed top quark $M_{\mathrm{T}}(\ell\nu b)$, defined by the equation:

$$
\begin{aligned}
M_{\mathrm{T}}(\ell\nu b) \;=\; \Bigg(\;\; & (P_{\mathrm{T}}(l) + P_{\mathrm{T}}(\not{E}_{\mathrm{T}}) + P_{\mathrm{T}}(b))^2 \\
& -(P_{\mathrm{x}}(l) + P_{\mathrm{x}}(\not{E}_{\mathrm{T}}) + P_{\mathrm{x}}(b))^2 - (P_{\mathrm{y}}(l) + P_{\mathrm{y}}(\not{E}_{\mathrm{T}}) + P_{\mathrm{y}}(b))^2 \Bigg)^{\frac{1}{2}}
\end{aligned}
\tag{6.3}
$$

The discrimination between processes involving top quarks (peaking at high transverse masses) and the background processes is particularly pronounced for QCD and Z + jets events; they have a peak at low transverse masses, see figure 6.6(c).

Another discriminating variable for the t-channel is the cosine of the angle between the lepton and the spin axis of the top quark [74] (which corresponds to the direction of the light quark jet in this channel). The polarization of the top quark can be inferred from this angular distribution, since it passes on its spin information to its decay products (as already mentioned in section 1.2). Due to the left-handed structure of the electroweak interaction the top quark is produced almost exclusively with negative helicity, leading to the typical $\cos\theta(l, lj)$ distribution for t-channel events, shown in figure 6.6(e).

Again the χ^2 and KS probabilities in figures 6.6(b), 6.6(d) and 6.6(f) are greater than 1%, so that a good agreement of the data to the prediction is found.

In figure 6.7 variables 7 to 9 are shown. That top quark production necessitates a high center of mass energy is also visible in the transverse energy of the light quark jet E_{T} (light jet), shown in figure 6.7(a).

The angle between the lepton and the W boson is particularly efficient to separate single top quark events from Z + jets and Wc (figure 6.7(c)).

The pseudorapidity η_W of the W boson originating from a top quark decay is more central than for the other decays, as can be seen in figure 6.7(e).

The χ^2 and KS probabilities in figures 6.7(b), 6.7(d) and 6.7(f) are greater than 1% and thus show the agreement between observed data and predicted events.

The last variable is the transverse mass of the W boson, shown in figure 6.8. A large discrimination can be seen between single top quark events on the one hand and QCD, Z + jets and $t\bar{t}$ events on the other hand (figure 6.8(a)). The predicted distribution in figure 6.8(b) is consistent with data as can be inferred from the χ^2 and KS probabilities, that are greater than 1%.

In figure 6.9 the neural network output is shown. As can be seen in figure 6.9(a) there is a very good separation between single top events and background processes. Figure 6.9(b) shows the neural network output of the predicted distributions normalized to the number of expected events and the observed distribution.
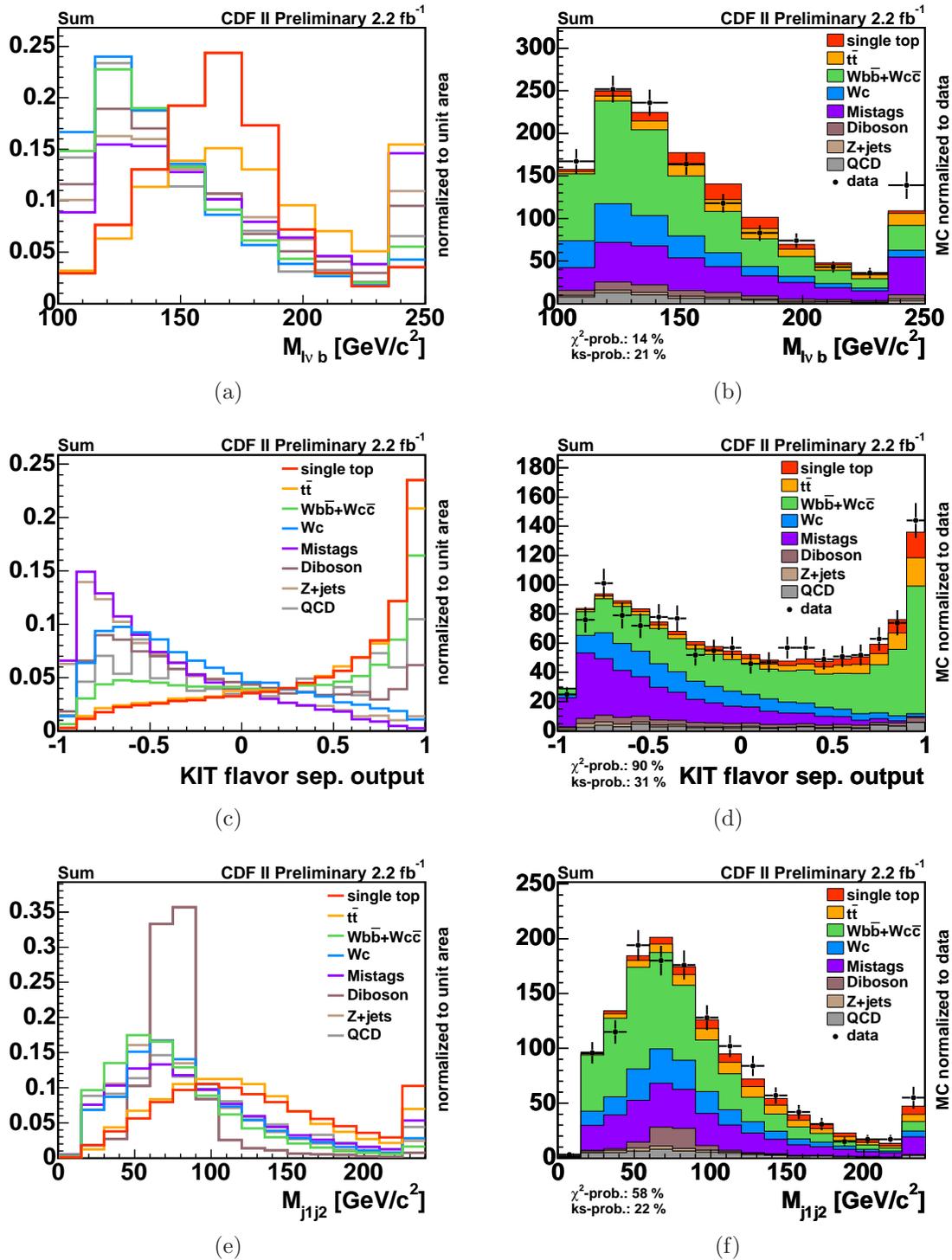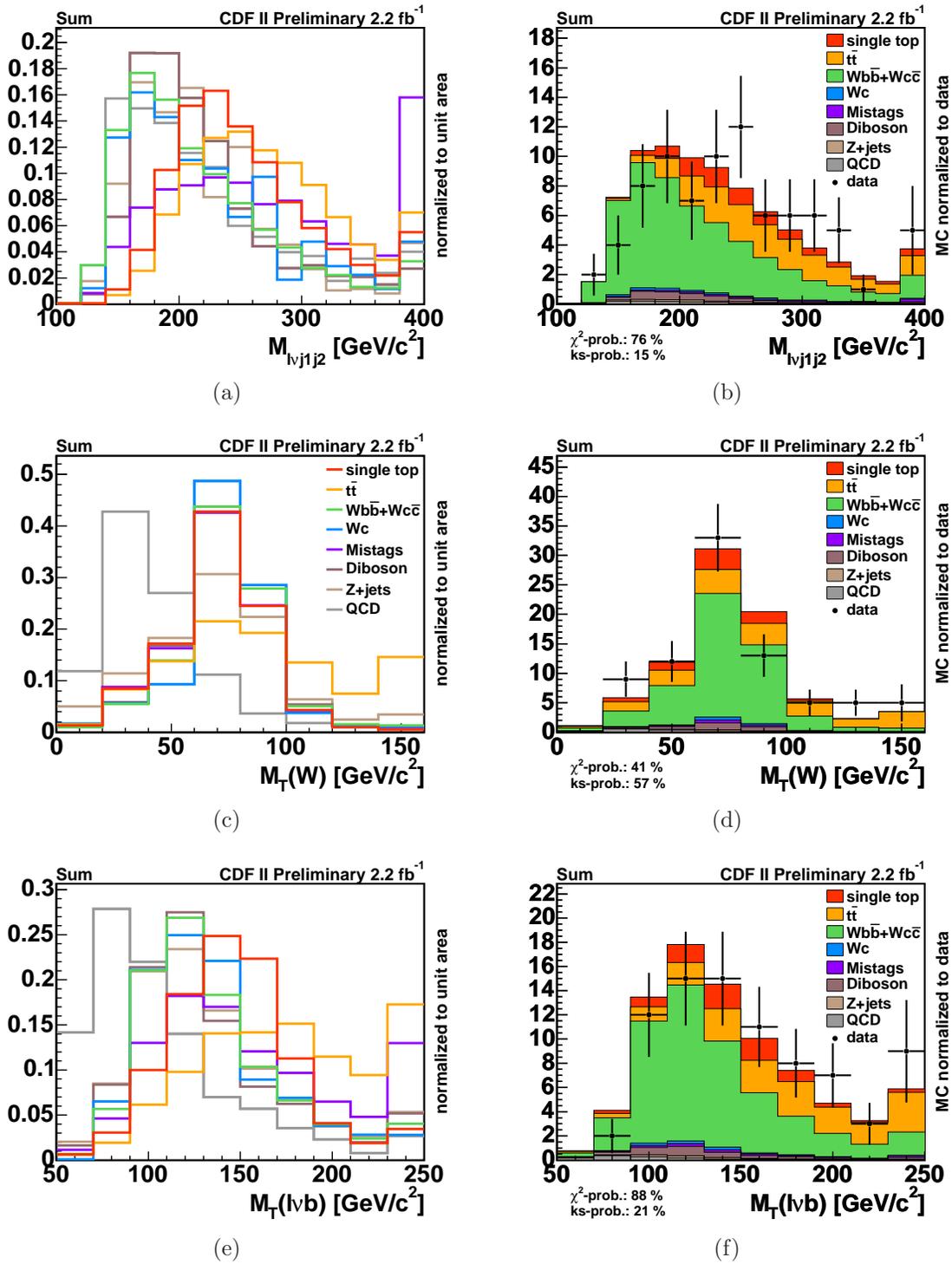
Figure 6.5: The three most important variables in the training of the neural network used to discriminate t-channel events with two jets and one b tag are shown:
(a) and (b) the invariant mass of the lepton, the neutrino and the b quark jet $M_{l\nu b}$,
(c) and (d) the output of the neural network flavor separator,
(e) and (f) the dijet mass M_{j1j2}.
The predicted distributions in figures (a), (c) and (e) are scaled to unit area, the ones in (b), (d) and (f) are normalized to data.
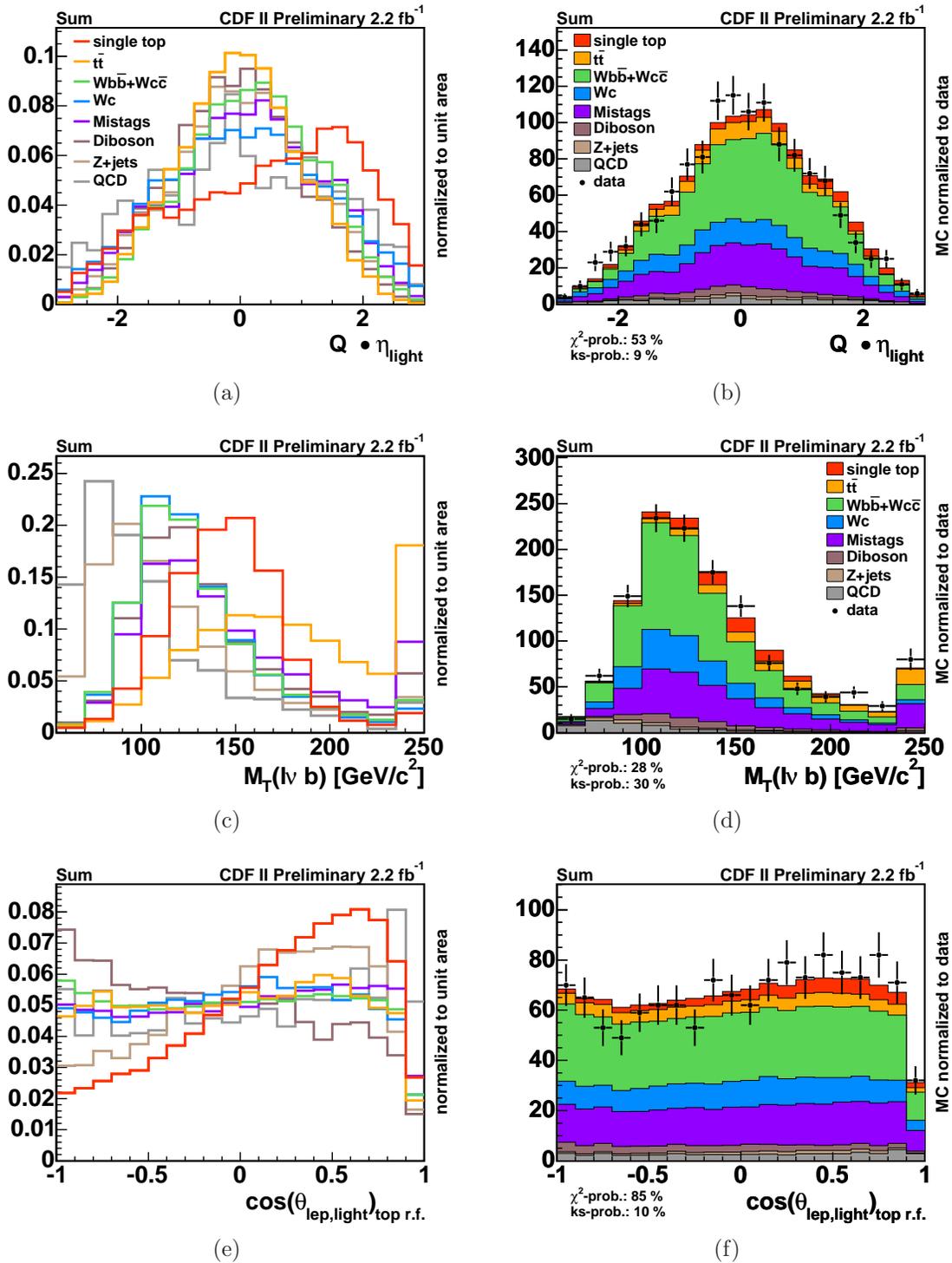
Figure 6.6: Variables 4 to 6 in the training of the neural network used to discriminate t-channel events with two jets and one b tag are shown:

(a) and (b) the product of the lepton charge times the pseudorapidity of the light quark jet $Q_\ell \cdot \eta_{lj}$,

(c) and (d) the transverse mass of the neutrino, the lepton and the b quark jet $M_T (\ell\nu b)$,

(e) and (f) the cosine of the angle between the lepton and the light quark jet $\cos\theta(l, lj)$.

The predicted distributions in figures(a), (c) and (e) are scaled to unit area, the ones in figures (b), (d) and (f) are normalized to data.

(a)

(b)

(c)

(d)

(e)

(f)
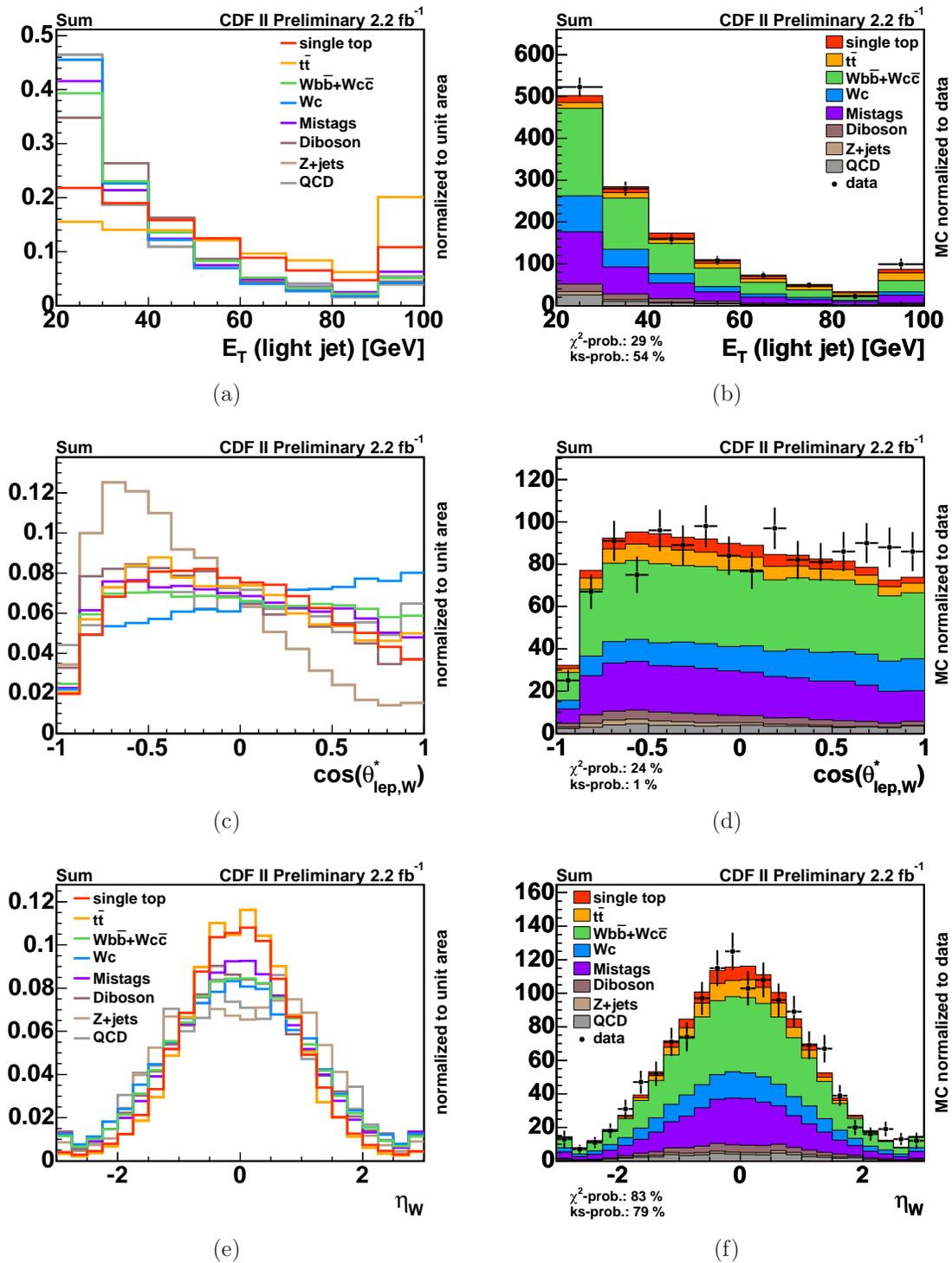
Figure 6.7: Variables 7 to 9 in the training of the neural network used to discriminate t-channel events with two jets and one b tag are shown:
(a) and (b) the transverse energy of the light quark jet E_T (light jet),
(c) and (d) the cosine of the angle between the lepton and the W boson $\cos\theta(l, W)$,
(e) and (f) the pseudorapidity of the W boson η_W.
The predicted distributions in figures (a), (c) and (e) are scaled to unit area, the ones in figures (b), (d) and (f) are normalized to data.
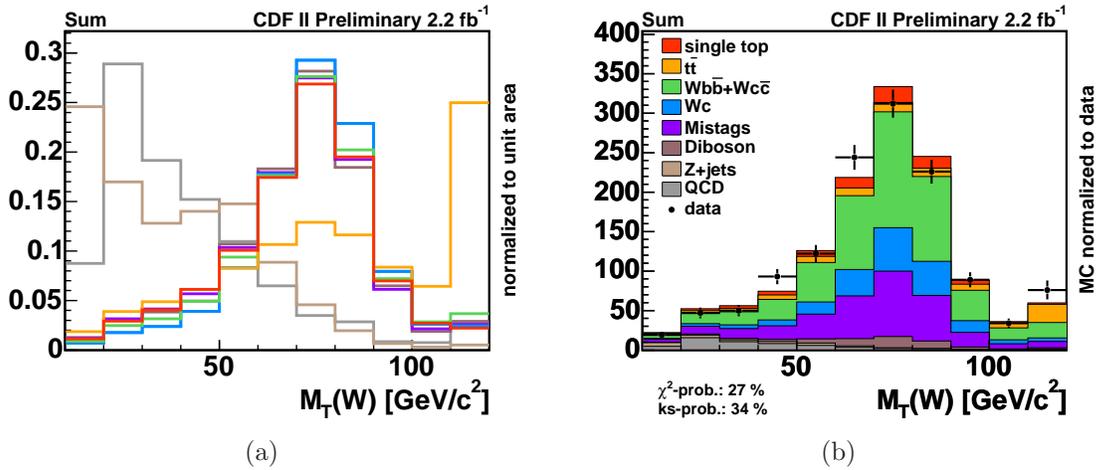
Figure 6.8: Variable 10 in the training of the neural network used to discriminate t-channel events with two jets and one b tag is shown.
(a) and (b) the transverse energy of the W boson $M_T(W)$.
The distributions in (a) are scaled to unit area, the predicted distributions in (b) are normalized to data.
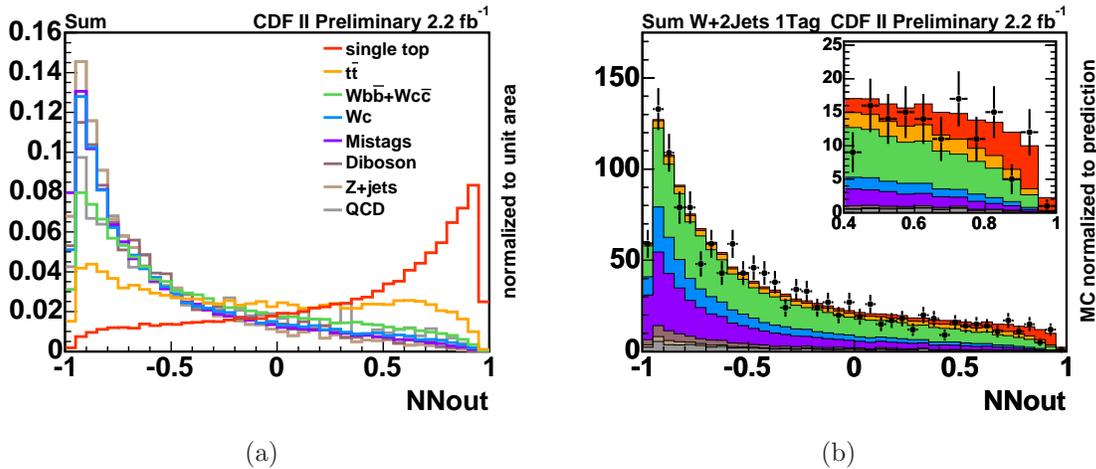
Figure 6.9: Neural network output for the t-channel network trained on events with two jets and one b tag. (a) depicts the templates which are fitted to the observed distribution. (b) shows the expected distribution normalized to prediction as well as the observed distribution.

s-channel, 2 jets, 2 b-tags   For the training of the s-channel network on events with two jets and two b tags the variables in table 6.9 were used.

Rank	Variable	Relative Significance (in σ)
1	$M_{\ell\nu j1j2}$	34
2	$M_{\mathrm{T}}(W)$	24
3	$M_{\mathrm{T}}(\ell\nu b)$	18
4	$\cos\Theta(j,j)$	18
5	$M_{\ell\nu b}$	11
6	$\sum_{\mathrm{jets}}$(flavor separator output)	8
7	η_W	7
8	$\not{E}_{\mathrm{T}}$(level 5)	6
9	M_{j1j2}	6
10	$E_{\mathrm{T}}(b)$	5
11	η_ℓ	3

Table 6.9: Set of discriminating variables used for the training of the s-channel neural network on events with two jets and two b tags.

In figure 6.10 the distributions of the processes for the best three variables are displayed. The first variable in the ranking is the invariant mass of the two jets, the lepton and the neutrino, $M_{l\nu j1j2}$. It corresponds to the center of mass energy $\sqrt{\hat{s}}$ for an s-channel event. As shown in figure 6.10(a) it distinguishes single top quark production from all other processes.
The next variable is the transverse mass of the W boson $M_{\mathrm{T}}(W)$. It is in particular a very well discriminator to separate single top quark events from $t\bar{t}$ production and QCD events, as can be inferred from figure 6.10(c).
The third most important variable is the transverse mass of the reconstructed top quark $M_{\mathrm{T}}(\ell\nu b)$, as shown in figure 6.10(e).
The χ^2 and KS probabilities in figures 6.10(b), 6.10(d) and 6.10(f), show that the predicted distributions correspond to the measured ones.

The neural network output is shown in figure 6.11. Single top quark events are separated from the background processes as can be seen in figure 6.11(a). In figure 6.11(b) shows the predicted distributions normalized to the number of expected events and the observed distribution.

Figure 6.10: The three most important variables in the training of the neural network used to discriminate *s*-channel events with two jets and two *b* tags are shown:
(a) and (b) the invariant mass of the two jets, the neutrino and the lepton $M_{l\nu j1j2}$
(c) and (d) the transverse mass of the W boson $M_T(W)$,
(e) and (f) the transverse mass of the lepton, the neutrino and the *b*-jet $M_T(\ell\nu b)$.
The expected distributions in figures (a), (c) and (e) are scaled to unit area, the ones in figures (b), (d) and (f) are normalized to data.
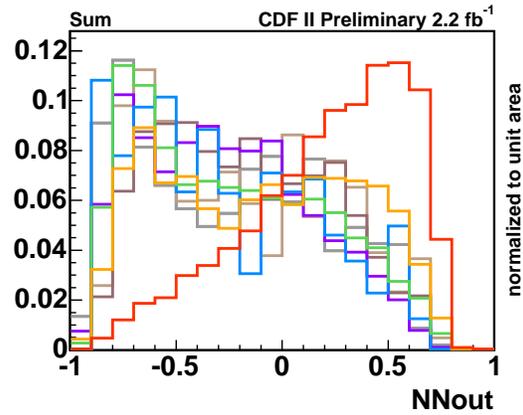
(a)
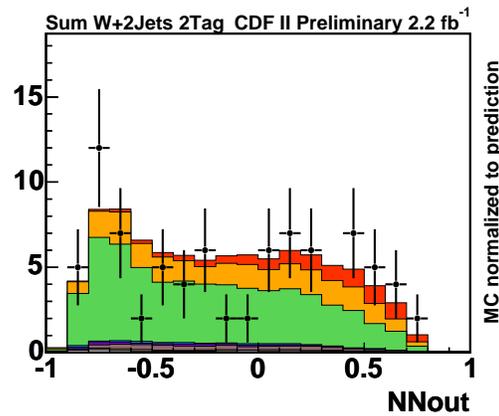
(b)

Figure 6.11: Neural network output for the *s*-channel network trained on events with two jets and two *b* tags. (a) depicts the templates which are fitted to the observed distribution. (b) shows the observed and the expected distribution normalized to the prediction.

t-channel, 3 jets, 1 b-tag The ranking of the variables used for the training of the t-channel network on events with three jets and one b tag are given in table 6.10.

Rank	Variable	Relative Significance (in σ)
1	$Q_\ell \cdot \eta_{lj}$	38
2	flavor separator output	35
	for the b quark from the top quark	
3	H_{T}	25
4	M_{j1j3}	23
5	$M_{\ell\nu b}$	19
6	$p_{\mathrm{T}}(\ell\nu bjj)$	17
7	M_{j2j3}	13
8	$\cos\Theta(\ell, lj)$	12
9	$\Delta\eta(j1, j2)$	11
10	$\sum_{\mathrm{j1j2j3}}(\eta_{jets})$	7
11	$\not{E}_{\mathrm{Tsig}}$	6
12	M_{j1j2j3}	7
13	$E_{\mathrm{T}}(j2, j3)$	7
14	$E_{\mathrm{T}}(b)$	6
15	$M_{\mathrm{T}}(\ell\nu b)$	6
16	$\Delta\eta(t, lj)$	5
17	$E_{\mathrm{T}}(j1, j3)$	4
18	$E_{\mathrm{T}}(j1, j2)$	5

Table 6.10: Set of discriminating variables used for the training of the t-channel neural network on events with three jets and one b tag.

The distributions of the three best variables are shown in figure 6.12.
The most important variable is the product of the charge of the lepton times the pseudorapidity of the light flavored jet $Q_\ell \cdot \eta_{lj}$, see figure 6.12(a). In second place is the neural network flavor separator as shown in figure 6.12(c). The third best variable is the tight sum of transverse energies H_{T}. It comprises the transverse energies of all particles in the events except for the loose jets. The χ^2 and KS probabilities in figures 6.12(b), 6.12(f) and 6.12(d) are all greater than 1%, there is thus a good correspondence between data and predicted distributions.

Figure 6.13 shows the neural network output. In figure 6.13(a) the separation between the single top quark processes and the background processes can be seen. The predicted distributions in figure 6.13(b) are normalized to prediction, the observed distribution is shown as well.
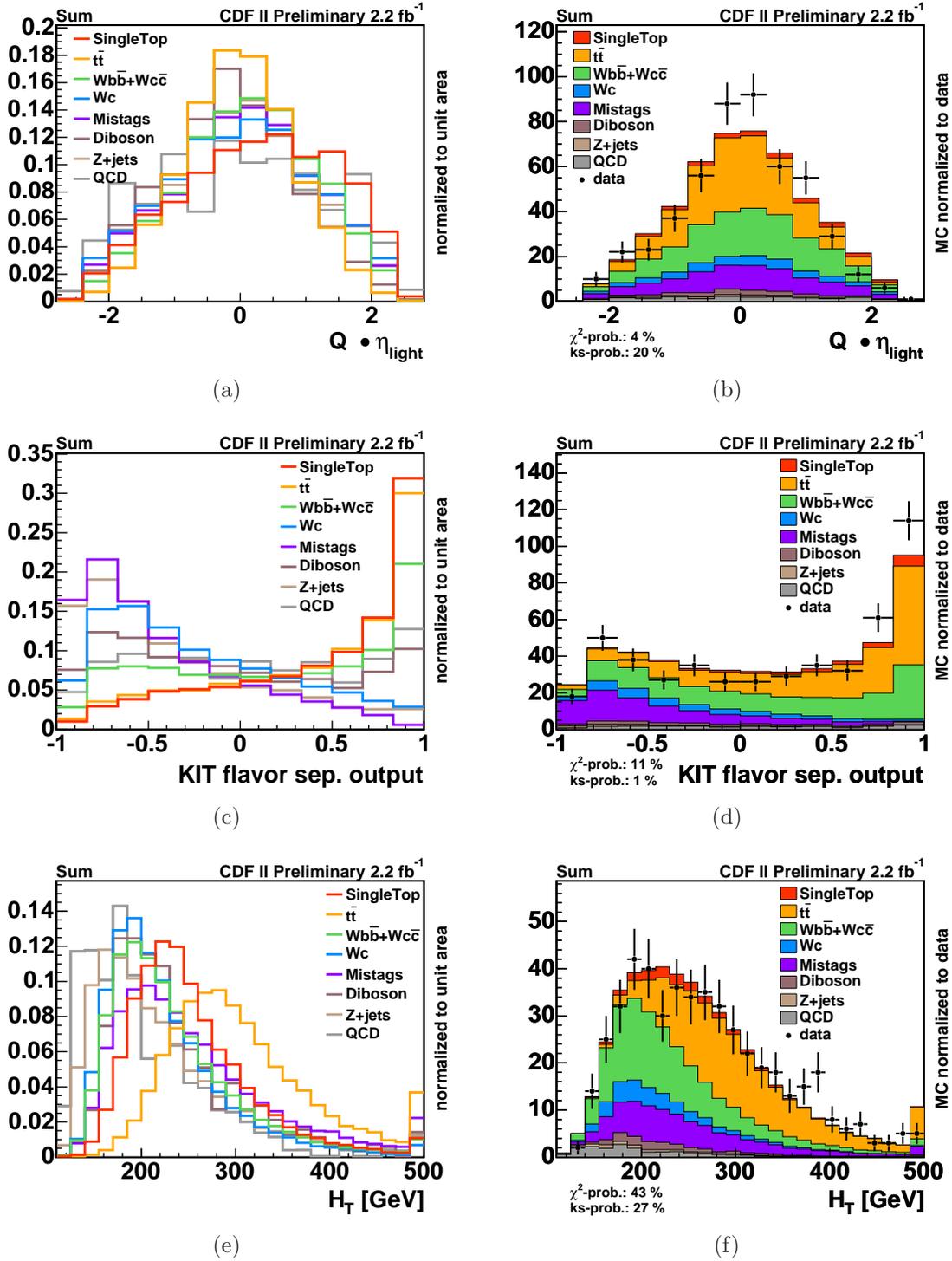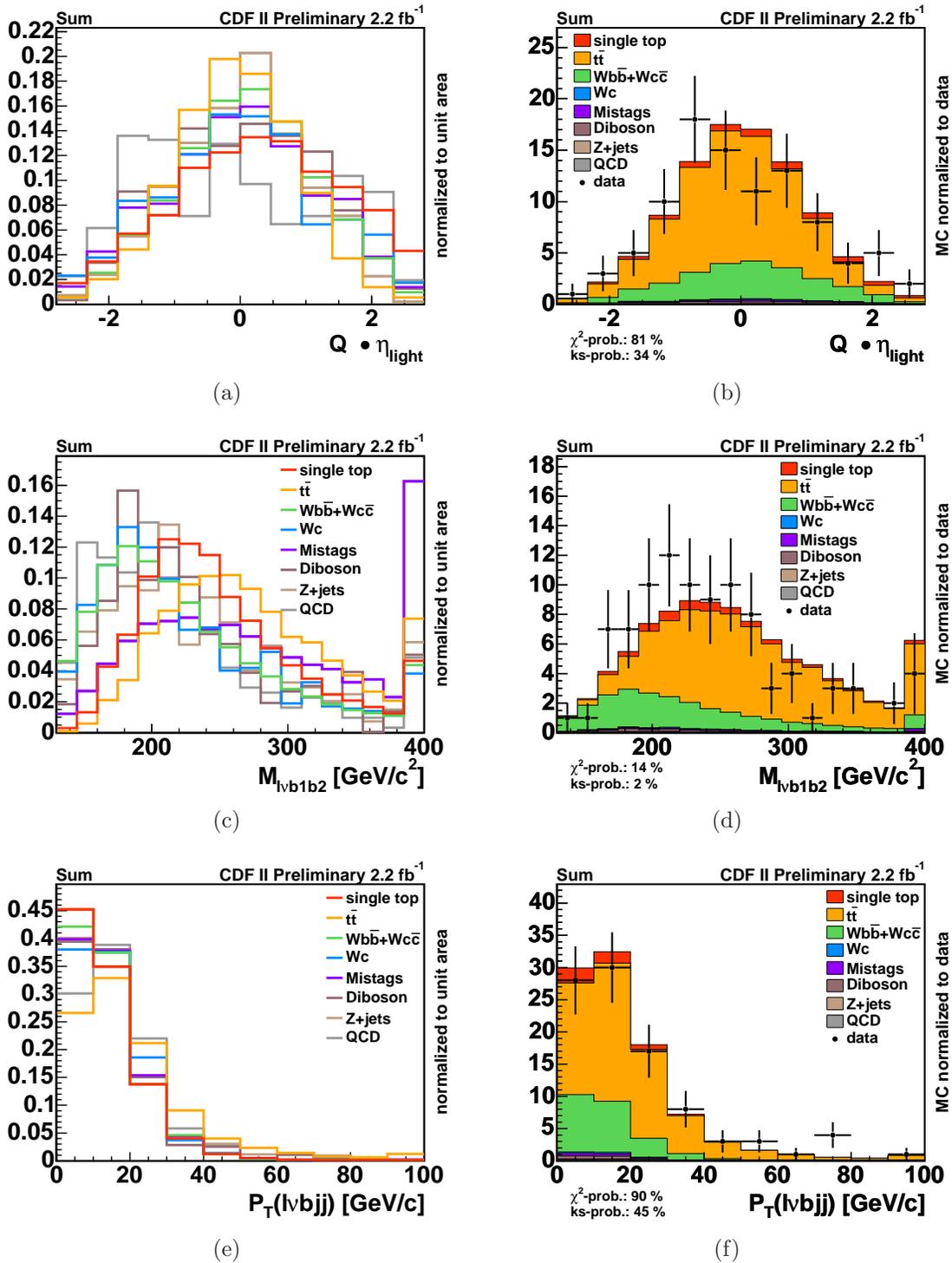
Figure 6.12: The three most important variables in the training of the neural network used to discriminate t-channel events with three jets and one b tag are shown:
(a) and (b) the product of the lepton charge times the pseudorapidity of the light quark jet $Q_\ell \cdot \eta_{lj}$
(c) and (d) the output of the neural network flavor separator.
(e) and (f) the tight transverse energy H_T of all particles produced in the event (with the exception of the loose jets),
The distributions in (a), (e) and (c) are scaled to unit area, the predicted distributions in (b), (f) and (d) are normalized to data.
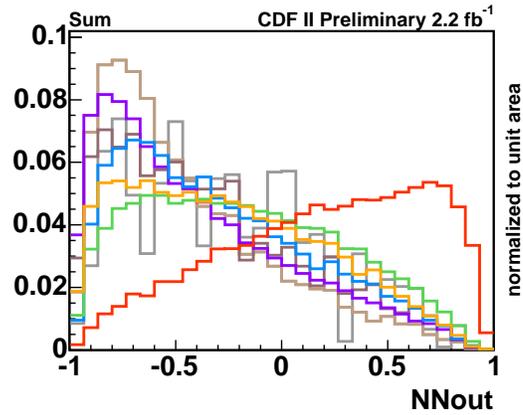
(a)
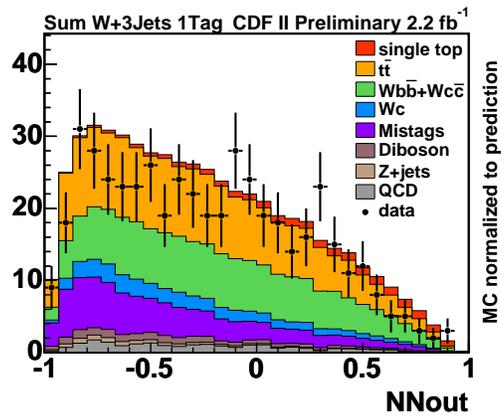
(b)

Figure 6.13: Neural network output for the *t*-channel network trained on events with two jets and one *b* tag. (a) depicts the templates which are fitted to the observed distribution. (b) shows the observed and the expected distribution normalized to the prediction.

t-channel, 3 jets, 2 b-tags   For the t-channel network trained on events with three jets and two b tags the variables given in table 6.11 were taken.

Rank	Variable	Relative Significance (in σ)
1	$Q_\ell \cdot \eta_{lj}$	25
2	$M_{\ell\nu bb}$	11
3	$p_{\mathrm{T}}(\ell\nu bjj)$	9
4	M_{j1j2}	6
5	$\cos\Theta(\ell, lj)$	5
6	$\sum_{\mathrm{j1j2}}(E_{\mathrm{T,jets}})$	5
7	M_{j1j3}	4
8	$\Delta\eta(j2, j3)$	4
9	$E_{\mathrm{T,2nd-b}}$	4
10	$M_{\mathrm{T}}(\ell\nu b)$	3
11	$\Delta\eta(j1, j2)$	3
12	$E_{\mathrm{T}}(\mathrm{jet3})$	2
13	centrality $\sum_{\mathrm{j1j2}}(E_{\mathrm{T,jets}})/\hat{s}$	2
14	$\hat{s}$	3
15	$\cos\Theta(j, j)$	3

Table 6.11: Set of discriminating variables used for the training of the t-channel neural network on events with three jets and two b tags.

The three most important variables are shown in figure 6.14.

The first variable in the ranking is again the product of the charge of the lepton times the pseudorapidity of the light flavored jet $Q_\ell \cdot \eta_{lj}$, shown in figure 6.14(a). The second best variable is the invariant mass of the lepton, the neutrino, and the two tagged jets $M_{l\nu b\bar{b}}$. In third place is the transverse momentum of the reconstructed top quark and the two remaining jets $p_{\mathrm{T}}(\ell\nu bjj)$. The χ^2 and KS probabilities in figures 6.14(b), 6.14(d) and 6.14(f) give probabilities greater than 1% for the predicted distribution to describe the observed distributions.

Figure 6.15 shows the neural network output. Single top quark events are well separated from the main background process $t\bar{t}$, while the discrimination to the other processes is not very pronounced, as can be seen in figure 6.15(a). Figure 6.15(b) shows the expected distribution normalized to prediction as well as the observed distribution.

Figure 6.14: The three most important variables in the training of the neural network used to discriminate t-channel events with three jets and two b tags are shown:
(a) and (b) the product of the lepton charge times the pseudorapidity of the light quark jet $Q_\ell \cdot \eta_{lj}$
(c) and (d) the invariant mass of the lepton, the neutrino, and the two tagged jets $M_{l\nu b\bar{b}}$,
(e) and (f) the transverse momentum of the reconstructed top quark and the two remaining jets $p_T(\ell\nu bjj)$.
The predicted distributions in figures (a), (c) and (e) are scaled to unit area, the ones in figures (b), (d) and (f) are normalized to the data.
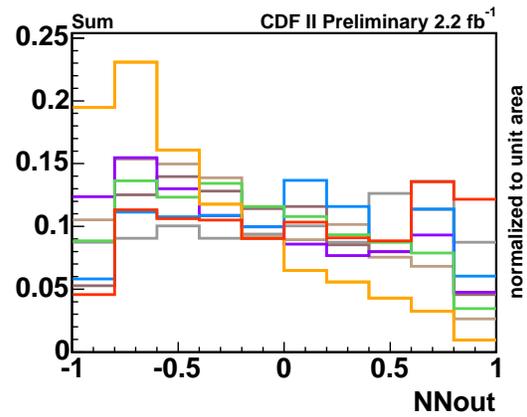
(a)
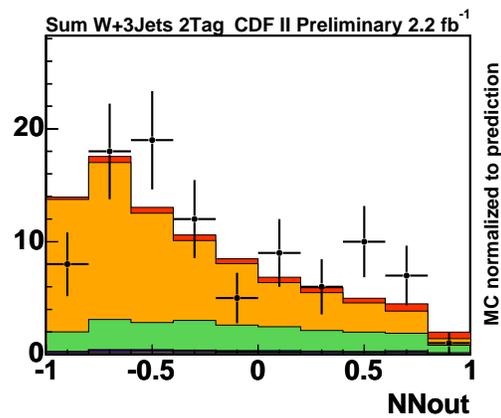
(b)

Figure 6.15: Neural network output for the *t*-channel network trained on events with two jets and one *b* tag. (a) depicts the templates which are fitted to the observed distribution. (b) shows the observed and the expected distribution normalized to the prediction.

6.3 Systematic Uncertainties

Systematic uncertainties in the measured results arise from detector effects as well as from the modeling of physics processes. They can influence the rate of the predicted signal and background events and the shape of the template distributions used in the fit.

The following sources for systematic uncertainties are taken into account:

- the uncertainty on the jet energy scale

- the uncertainty in modeling initial-state gluon radiation (ISR)

- the uncertainty in modeling final-state gluon radiation (FSR)

- the choice of the parameterization of the parton distribution functions (PDF) used for the event simulation

- the choice of the Monte Carlo event generator

- comparison to NLO calculations

- the uncertainty in the event detection efficiency

- the uncertainty in modeling the output of the neural network flavor separator

- the uncertainty in the factorization and renormalization scale for the simulation of W+heavy flavor processes

- the modeling of instrumental backgrounds, that is mistag events and non-W events

- the uncertainty in the luminosity determination

- the uncertainty in the modeling of the distributions of $\Delta R_{j_1,j_2}$, η_{j2} for events with two jets and η_{j3} for events with three jets

The impact of these sources of uncertainties is evaluated by altering the modeling of the corresponding processes or effects within their uncertainties or by assigning a plausible alternative model. As a result, relative changes of the event rates and shifted template distributions are obtained.

The effect of the uncertainty on the jet energy scale is quantified by varying the scale within positive and negative deviations forming an envelop of possible values. [46]. The corresponding alternative template distributions are calculated for all signal and background processes and are shown exemplarily for single top quark and $t\bar{t}$ production in figure 6.16.

The influence of initial-state and final-state gluon radiation is estimated by producing samples of simulated events for which the simulation was altered to produce

either less or more gluon radiation compared to the standard setting [75]. Specifically, two parameters controlling the parton shower in the PYTHIA program are varied: Λ_{QCD} and the scale factor K to the transverse momentum scale of the showering. The different settings are derived from studies of ISR in Drell-Yan events. Using these specific ISR and FSR samples of simulated events, alternative template shapes are produced for single top quark and $t\bar{t}$ events.

The impact of the uncertainties on the PDF parameterization are studied by reweighting single top quark and $t\bar{t}$ events with weights associated with the 20 pairs of CTEQ6M eigenvectors. The rate uncertainty on the signal model is determined by a comparison to differential cross sections computed with the ZTOP program. The modeling of $t\bar{t}$ events is studied by using simulated events produced with the MC@NLO program and showered by HERWIG as an alternative generator. The factorization and renormalization scale is varied in the simulation to derive an additional set of altered template histograms for W+heavy flavor events. The default W+jets Monte Carlo samples are generated with a dynamic scale $\mu^2 = Q^2$.

The uncertainty in the event detection efficiency ϵ_{evt} includes the uncertainties on the trigger efficiency, on the lepton identification efficiency, and on the b tagging efficiency which is the dominating factor. Since no cut is applied on the output of the neural network flavor separator, the uncertainty associated with this quantity does not imply a rate uncertainty, but only a shape uncertainty on the template distributions. Systematic effects are studied by utilizing the correction function derived for the mistags. Therefore we consider two scenarios. The pessimistic one, in which we apply the correction function on the c-like templates, that they get more signal like. And the optimistic one, in which we use the uncorrected mistag shape, that the mistag template gets more background like. The flavor composition of the non-W sample is varied: the default model assumes a composition of 45% b quark jets, 40% c quark jets, and 15% light-quark jets, whereas the alternative model uses a composition of 60:30:10 [63], respectively. To evaluate the systematic effect on the shapes of the distributions caused by the modeling of mistagged light-quark jet events, an alternative model is utilized to create template distributions. This is realized by replacing the default mistag model based on simulated events as described in section 3.2.2 by a description on the basis of measured W+jets events before b tagging. The uncertainty on the modeling of the distributions of ΔR and η_{j2} in the pretag sample is determined by reweighting the distributions.

Tables 6.12-6.16 summarize the relative rate uncertainties on the event prediction of the various Monte Carlo samples.

Source	t-channel	s-channel	single-top	$t\bar{t}$
ISR less/more	2.8/-0.2 %	0.3/6.7 %	1.9/2.1 %	-2.6/-7.1 %
FSR less/more	4.2/-1.3 %	5.9/0.4 %	4.8/-0.7 %	-5.1/-2.6 %
PDF	3.4/-3.4 %	2.2/-2.2 %	3.0/-3.0 %	1.8/-1.8 %
MC	2.0/-2.0 %	1.0/-1.0 %	1.7/-1.7 %	-2.7/2.7 %
ϵ_{evt}	4.2/-4.2 %	2.3/-2.3 %	3.6/-3.6 %	2.9/-2.9 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Cross section	12.6/-12.6 %	12.4/-12.4 %	12.6/-12.6 %	12.4/-12.4 %
M_{top} 170/180	1.3/-0.8 %	2.4/-1.7 %	1.7/-1.1 %	-3.1/1.4 %
	Diboson	Z+jets		
ϵ_{evt}	7.6/-7.6 %	8.3/-8.3 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Cross section	1.9/-1.9 %	10.8/-10.8 %		

Table 6.12: Systematic rate uncertainties for events with two jets and one b tag.

Source	t-channel	s-channel	single-top	$t\bar{t}$
ISR less/more	-4.9/-6.9 %	1.3/9.2 %	0.4/6.7 %	0.5/-9.5 %
FSR less/more	3.9/-6.6 %	8.1/2.2 %	7.5/0.8 %	-8.1/-1.8 %
PDF	2.0/-2.0 %	2.0/-2.0 %	2.0/-2.0 %	1.7/-1.7 %
MC	2.0/-2.0 %	1.0/-1.0 %	1.2/-1.2 %	4.6/-4.6 %
ϵ_{evt}	10.0/-10.0 %	8.7/-8.7 %	8.9/-8.9 %	9.0/-9.0 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Cross section	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
M_{top} 170/180	-4.7/-4.1 %	2.1/0.1 %	1.0/-0.5 %	0.4/3.0 %
	Diboson	Z+jets	Mistags	
ϵ_{evt}	9.8/-9.8 %	10.6/-10.6 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Double tag			23.4/-23.4%	
Cross section	1.9/-1.9 %	10.8/-10.8 %		

Table 6.13: Systematic rate uncertainties for events with two jets and two b tags.
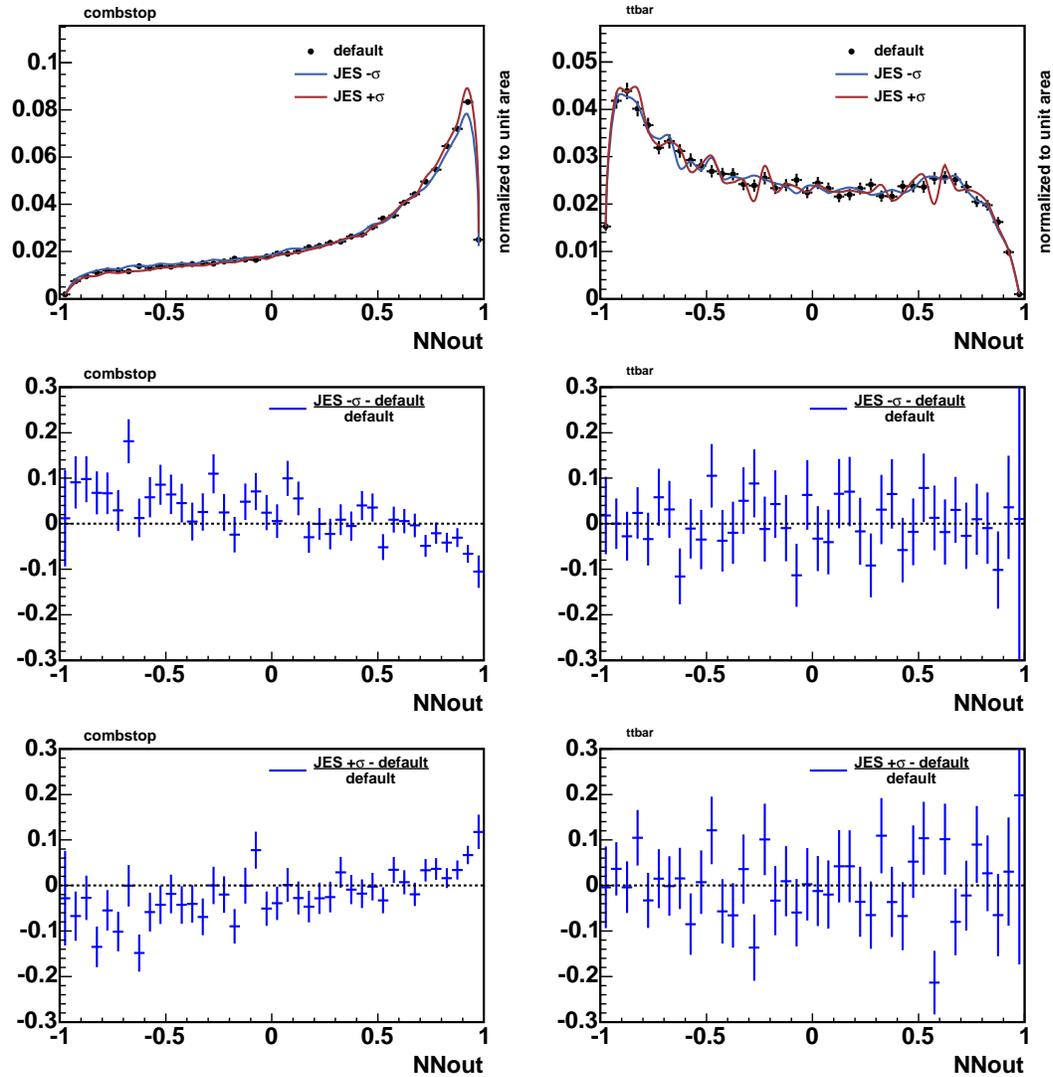
Figure 6.16: Shape systematics due the uncertainty on the jet energy correction for single top and $t\bar{t}$ events. In the upper row the default distribution is shown in comparison to the shifted distributions. In the two lower rows the relative difference between the shifted distribution and the default is plotted.

Source	t-channel	s-channel	single-top	$t\bar{t}$
ISR less/more	-6.8/-0.0 %	2.4/-12.6 %	-3.3/-4.8 %	-0.6/-4.6 %
FSR less/more	-1.5/-3.1 %	-6.0/-4.8 %	-3.3/-3.8 %	-3.5/-2.2 %
PDF	2.7/-2.7 %	2.3/-2.3 %	2.6/-2.6 %	1.8/-1.8 %
MC	1.9/-1.9 %	1.5/-1.5 %	1.7/-1.7 %	-1.7/1.7 %
ϵ_{evt}	3.5/-3.5 %	2.3/-2.3 %	3.0/-3.0 %	2.3/-2.3 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Cross section	12.6/-12.6 %	12.4/-12.4 %	12.6/-12.6 %	12.4/-12.4 %
M_{top} 170/180	1.5/-2.8 %	6.0/-2.7 %	3.2/-2.7 %	-0.7/0.8 %
	Diboson	Z+jets		
ϵ_{evt}	7.8/-7.8 %	7.8/-7.8 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Cross section	1.9/-1.9 %	10.8/-10.8 %		

Table 6.14: Systematic rate uncertainties for events with three jets and one b tag

Source	t-channel	s-channel	single-top	$t\bar{t}$
ISR less/more	7.8/3.2 %	4.3/-11.2 %	5.8/-4.9 %	-0.5/-6.6 %
FSR less/more	15.0/1.3 %	-7.4/-5.0 %	2.4/-2.2 %	-3.4/-2.7 %
PDF	1.5/-1.5 %	2.1/-2.1 %	1.9/-1.9 %	1.7/-1.7 %
MC	1.9/-1.9 %	1.5/-1.5 %	1.7/-1.7 %	2.0/-2.0 %
ϵ_{evt}	9.1/-9.1 %	8.8/-8.8 %	8.9/-8.9 %	9.1/-9.1 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Cross section	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
M_{top} 170/180	4.2/3.0 %	1.6/-6.8 %	2.7/-2.5 %	-0.6/-1.0 %
	Diboson	Z+jets	Mistags	
ϵ_{evt}	10.8/-10.8%	11.1/-11.1 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Double tag			23.4/-23.4%	
Cross section	1.9/-1.9 %	10.8/-10.8 %		

Table 6.15: Systematic rate uncertainties for events with three jets and two b tags

The analyses are done under the assumption of a top quark mass of $M_t = 175\,\text{GeV}/c^2$. That is why the uncertainty in the top quark mass is not taken into account as a systematic uncertainty. Hence, the analyses provide rather a measurement at the specified value of the top quark mass. However, if the top quark mass is varied in the simulation by $\pm 5\,\text{GeV}/c^2$, the acceptance for single top quark events changes as shown in table 6.12-6.15.

process	2jets 1tag	2jets 2tags	3jets 1tag	3jets 2tags
t-ch	-1.1/0.6 %	4.8/-3.5 %	-10.4/10.6 %	-5.7/4.3 %
s-ch	-0.1/-0.6 %	1.2/-1.9 %	-8.3/9.4%	-7.2/7.4 %
single-top	-0.8/0.2 %	1.8/-2.2 %	-9.1/9.9 %	-6.6/6.1 %
$t\bar{t}$	9.8/-9.4 %	8.1/-7.5 %	4.6/-5.1 %	5.4/-5.2 %
$Wc\bar{c}+Wb\bar{b}$	7.0/-6.9 %	10.8/-10.6 %	8.4/-7.7 %	11.0/-12.1 %
Wc	7.0/-6.3 %	11.3/-10.3 %	8.2/-6.9 %	13.9/-15.8 %
Z+jets	-5.3/5.4 %	5.0/-5.0 %	-10.8/14.0 %	-5.0/5.0 %
Diboson	-2.7/1.7 %	-3.0/1.5 %	-12.4/11.9 %	-11.0/11.0 %

Table 6.16: Systematic rate uncertainty due to shifts of the jet energy scale by -1σ/+1σ respectively.

6.4 Measured Single Top Quark Cross Section

The expected sensitivity of the single top quark cross section of the combined single top process is computed with ensemble tests, as described in section 5.3. The root mean square (RMS) of the single top quark cross section distribution obtained by the pseudo experiments is defined as the uncertainty of the measurement. The uncertainty on the cross section for the combined single top process is determined including all systematic uncertainties. As shown in figure 6.17 the uncertainty is 26.3% of the predicted cross section or 0.75 pb.
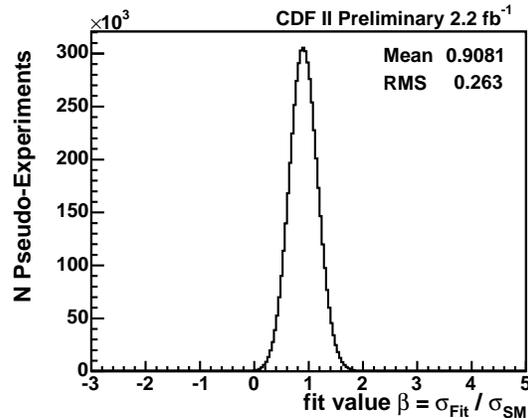
Figure 6.17: The distribution of expected measurements in the combined search is shown. The fitted cross sections are based on the ensemble test including single top quark events. The RMS value of this distribution is defined as the expected uncertainty of the measurement.

In order to estimate the significance of an observed signal a hypothesis test is conducted. The resulting Q-value distributions are shown in figure 6.18. From the Q-value the expected p-value can be computed. The p-value is defined by the median $\tilde{Q}$ of the Q-value distribution: $p = \hat{p}(\tilde{Q})$. A p-value of $\hat{p} = 0.000529\%$ is obtained. With a probability of 50% a p-value equal or smaller to this one is expected. Under the assumption that the single top quark cross section corresponds to the one predicted by the standard model, the probability that a measured excess

over the background corresponds merely to a background fluctuation is 0.000529%, or in other words, the excess over the background would correspond to a background fluctuation of 4.4 σ.
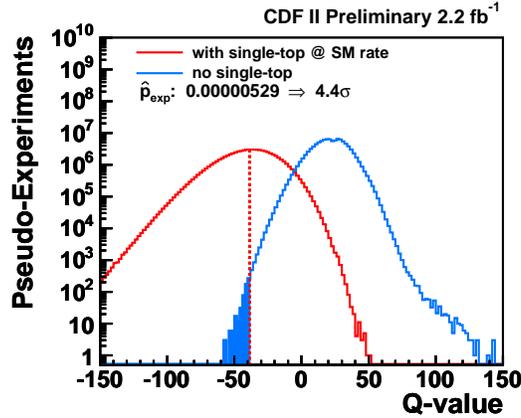
(a) combined search

Figure 6.18: Distributions of Q-values for two ensemble tests, one with single top quark events present at the expected standard model rate, one without any single top quark events.

In the combined search the output templates of the four networks, shown in figures 6.9(a), 6.11(a), 6.13(a) and 6.13(a) were fitted together. The obtained fit values for the signal and background process rates as well as the strengths of the systematic effects are given in table 6.17. For single top quark production a rate of $70.5^{+30.1}_{-28.4}\%$ of the standard model prediction is found, this corresponds to a cross section of $2.0^{+0.9}_{-0.8}$pb. It follows that the measured single top quark cross section is, under consideration of the error range, in good correspondence with the value predicted by the standard model. The observed Q-value is -13.6, which yields an observed p-value of 0.061%, see figure 6.19. The significance of the measurement is thus 3.2σ.
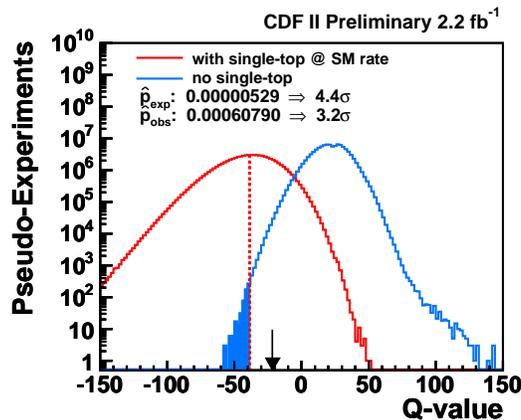
Figure 6.19: Comparison of observed Q-value to the expectation in the combined search.

Figure 6.20 displays the neural network output of the four networks used in the

process	fitted rate / predicted rate
single-top (uncorrected)	0.69 ± 0.29
$t\bar{t}$	0.91 ± 0.09
$Wb\bar{b} + Wc\bar{c}$	0.93 ± 0.13
Wc	1.07 ± 0.26
mistags	0.96 ± 0.14
diboson	1.00 ± 0.02
Z + HF jets	1.00 ± 0.11
non-W	0.89 ± 0.39

systematic source	fitted excursion / standard deviation
JES	0.44 ± 0.71
ISR	0.55 ± 0.82
FSR	0.22 ± 0.79
PDF	0.03 ± 0.96
single-top NLO	0.00 ± 0.93
$t\bar{t}$ generator	0.28 ± 0.97
luminosity	0.00 ± 0.07
acceptance	0.29 ± 0.88
double-tag rate	-0.02 ± 0.97
heavy flavor Q2	0.66 ± 0.74
mistag model	0.29 ± 0.69
non-W flavor	0.01 ± 0.93
KIT flavor separator	0.00 ± 0.23
η_{j2}	0.17 ± 0.78
η_{j3}	0.83 ± 0.76
ΔR_{j1j2}	-0.01 ± 0.97

Table 6.17: The fitted rates of the considered processes and strengths of systematic effects of the combined search are quoted. The given rate of single top quark production is the uncorrected fit result.

combined search. The distribution normalized to the fit result as well as the observed distributions are shown. They are in good agreement, as can be infered from the χ^2 and the KS probabilities.

In figure 6.21 these four neural network outputs of the combined search are comprised in a single output.
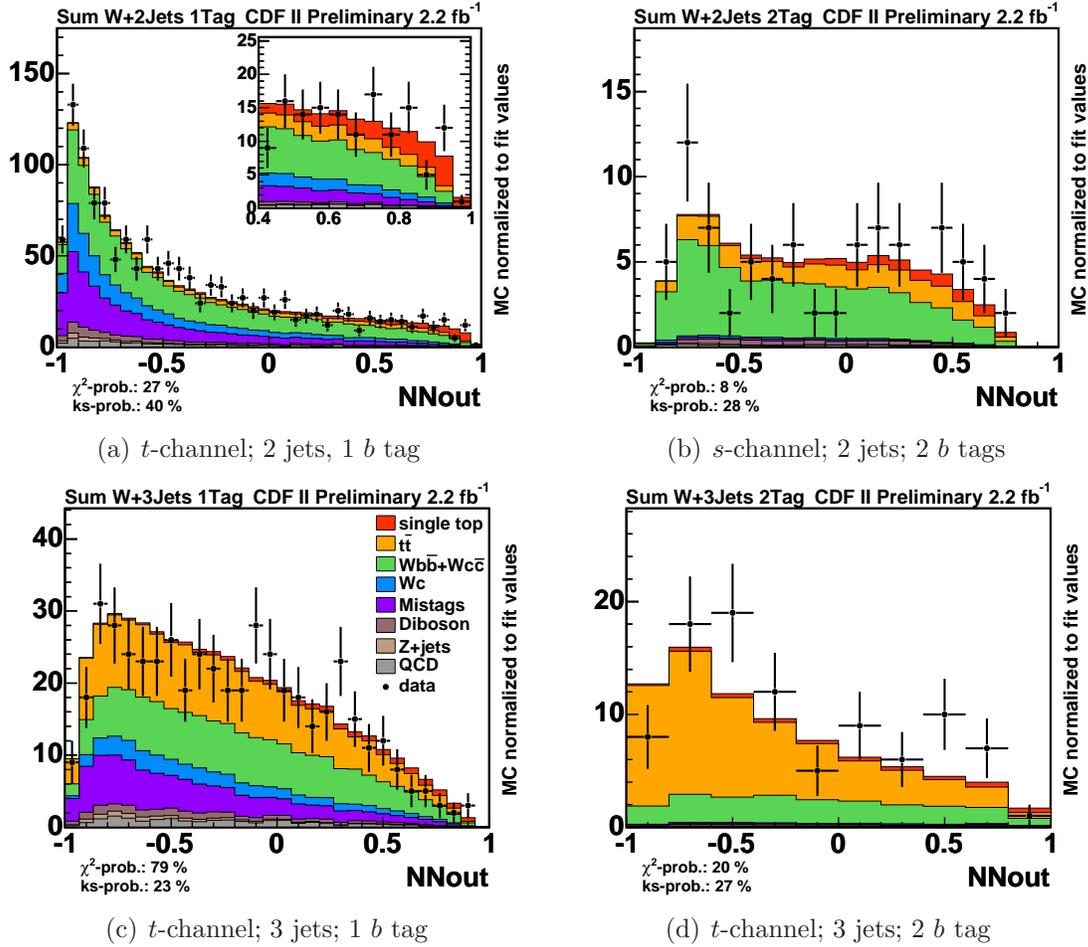
Figure 6.20: Neural network output of the four networks used in the combined search. Output of: (a) the t-channel neural network trained on events with two jets and one b tag, (b) the s-channel neural network rrained on events with two jets and two b tags, (c) the t-channel neural network trained on events with three jets and one b tag and (d) respectively two b tags. The distribtutions are normalized to the fit results
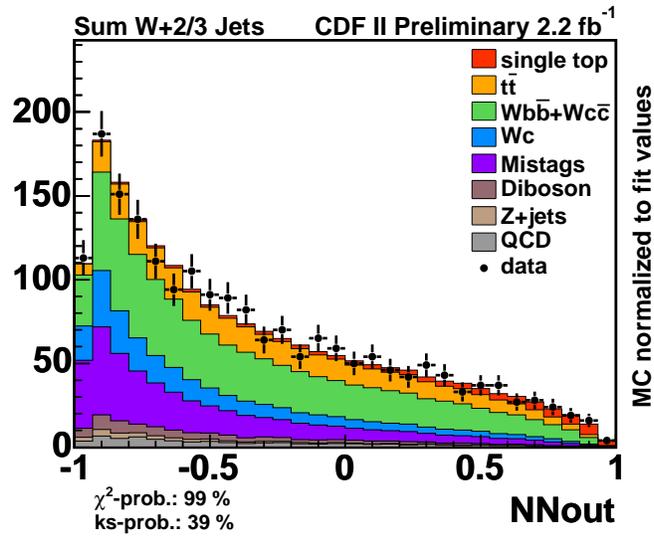
(a) combined search
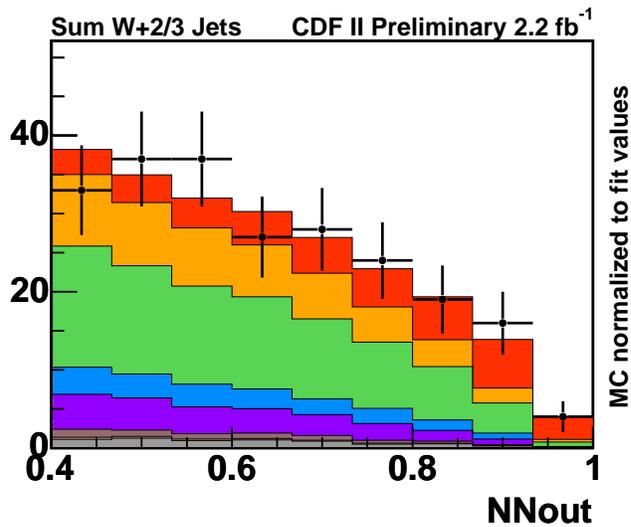
(b) combined search, signal region

Figure 6.21: Neural network output for the combined single top search. The output of the four networks used in the combined search are added up. The distributions are scaled to the fit result. In (a) the whole range of the output is shown, (b) displays only the signal region.

Summary and Outlook

In this thesis a neural network analysis to measure the single top quark cross section is presented. The examined dataset was taken by the CDF II detector at the Tevatron and comprises 2.2 fb^{-1} of integrated luminosity.
The two main processes contributing to the single top quark production are the s- and the t-channel. Their cross section is predicted to $\sigma_s = 0.88^{+0.12}_{-0.11}$ pb, respectively $\sigma_t = 1.98^{+0.28}_{-0.22}$ pb.

The analyzed data was required to correspond to the single top quark signature. Only events containing a lepton, missing transverse energy and two or three jets, of which at least one must have been identified to originate from a b quark, were taken into consideration.

To separate single top quark from background processes four neural networks were trained. One was optimized to recognize s-channel events with two jets and two b tags. The other networks were trained to identify t-channel events with two or three jets and at least one b tag. In the training of these networks the other single top quark process is not considered as a background.

As a cross check of the method two control networks were trained. The $Wb\bar{b}$ networks measures the number of $Wb\bar{b}$ events with one jet. The other network is trained with events with three jets and one b tag, and measures the $t\bar{t}$ cross section. It is found to be 7.5 $\pm$ 0.8 pb which is in good correspondence with the CDF II average of 7.3 $\pm$ 0.5 (stat) $\pm$ 0.6 (syst) $\pm$ 0.4 (lumi) pb [73].

The templates that are obtained by applying the single top quark neural networks on the samples of simulated signal and background events are fitted simultaneously to the output distribution of observed events, using a likelihood method. The negative logarithm of the likelihood function is minimized, whereby statistical as well as systematical uncertainties are taken into account.
The expected sensitivity is obtained with ensemble tests. The expected uncertainty on the cross section is estimated to be 0.75 pb, which corresponds to 26.3% of the cross section. The expected significance of the measurement is 4.4σ.
In data a single top quark cross section of $2.0^{+0.9}_{-0.8}$pb is measured. This corresponds to $70.5^{+30.1}_{-28.4}\%$ of the rate predicted by the standard model. Under consideration of the errors the measured value is in good agreement with the standard model cross section. The significance of the measurement is 3.2σ.

Evidence for the electroweak production of single top quarks has eluded for a long time. In Tevatron Run I several limits on the single top quark cross section could be set by DØ [76, 77] and CDF [78, 79]. Those limits were outperformed in RunII [80, 81]. The previous analyses with neural networks at CDF II were not able to establish a signal of single top quark production [70, 82]. First evidence for single top quark production was found by DØ [83, 84]. The probability to obtain these results under the assumption of the standard model is however at the percent level. The other two single top quark analyses at CDF II yield results comparable to the ones in this analysis [85, 86].

In conclusion it can be said that the obtained measurement of the single top quark cross section is still limited by statistics, a more precise and significant measurement will be possible with more data, either at the Tevatron or the LHC.

List of Figures

List of Tables

Bibliography

[1] F. Abe et al. Observation of top quark production in anti-p p collisions. Phys. Rev. Lett., 74:2626–2631, 1995.

[2] S. Abachi et al. Observation of the top quark. Phys. Rev. Lett., 74:2632–2637, 1995.

[3] W.-M. Yao et al. Review of particle physics. J. Phys., G33:1–1232, 2006.

[4] O. W. Greenberg. Spin and unitary spin independence in a paraquark model of baryons and mesons. Phys. Rev. Lett., 13:598–602, 1964.

[5] M. Y. Han and Y. Nambu. Three-triplet model with double SU(3) symmetry. Phys. Rev., 139:B1006–B1010, 1965.

[6] N. Cabibbo. Unitary symmetry and leptonic decays. Phys. Rev. Lett., 10:531–532, 1963.

[7] M. Kobayashi and T. Maskawa. CP violation in the renormalizable theory of weak interaction. Prog. Theor. Phys., 49:652–657, 1973.

[8] S. L. Glashow. Partial symmetries of weak interactions. Nucl. Phys., 22:579–588, 1961.

[9] A. Salam and J. C. Ward. Electromagnetic and weak interactions. Phys. Lett., 13:168–171, 1964.

[10] S. Weinberg. A model of leptons. Phys. Rev. Lett., 19:1264–1266, 1967.

[11] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. Phys. Rev. Lett., 13:321–322, 1964.

[12] P. W. Higgs. Broken symmetries and the masses of gauge bosons. Phys. Rev. Lett., 13:508–509, 1964.

[13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. Phys. Rev. Lett., 13:585–587, 1964.

[14] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. Phys. Rev., 145:1156–1163, 1966.

[15] D. J. Gross and F. Wilczek. Ultraviolet behavior of non-abelian gauge theories. Phys. Rev. Lett., 30:1343–1346, 1973.

[16] H. D. Politzer. Reliable perturbative results for strong interactions? Phys. Rev. Lett., 30:1346–1349, 1973.

[17] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the color octet gluon picture. Phys. Lett., B47:365–368, 1973.

[18] H. L. Lai et al. Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. Eur. Phys. J. C, 12:375–392, 2000.

[19] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi. The t anti-t cross-section at 1.8 TeV and 1.96 TeV: A study of the systematics due to parton densities and scale dependence. J. High Energy Phys., 04:068, 2004.

[20] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan, and S. Weinzierl. The fully differential single top quark cross section in next-to-leading order QCD. Phys. Rev., D66:054024, 2002.

[21] Z. Sullivan. Understanding single-top-quark production and jets at hadron colliders. Phys. Rev., D70:114012, 2004.

[22] A. P. Heinson, A. S. Belyaev, and E. E. Boos. Single top quarks at the Fermilab Tevatron. Phys. Rev., D56:3114–3128, 1997.

[23] T. M. P. Tait. The t W- mode of single top production. Phys. Rev., D61:034001, 2000.

[24] A. Belyaev and E. E. Boos. Single top quark tW+X production at the LHC: a closer look. Physical Review D, 63:034012, 2001.

[25] W. Wagner. Top quark physics in hadron collisions. Rept. Prog. Phys., 68:2409–2494, 2005.

[26] D. Mohl, G. Petrucci, L. Thorndahl, and S. Van Der Meer. Physics and technique of stochastic cooling. Phys. Rept., 58:73–119, 1980.

[27] S. Van Der Meer. Stochastic cooling and the accumulation of anti-protons. Rev. Mod. Phys., 57:689–697, 1985.

[28] G. I. Budker. An effective method of damping particle oscillations in proton and anti-proton storage rings. Sov. Atom. Energ., 22:438–440, 1967.

[29] S. Nagaitsev et al. Experimental demonstration of relativistic electron cooling. Phys. Rev. Lett., 96:044801, 2006.

[30] The CDF II Collaboration. The CDF II detector, 1996. Technical Design Report.

[31] A. Affolder et al. CDF central outer tracker. Nucl. Instrum. Methods A, 526:249–299, 2004.

[32] C. S. Hill. Initial experience with the CDF layer 00 silicon detector. Nucl. Instrum. Meth., A511:118–120, 2003.

[33] A. Sill. CDF Run II silicon tracking projects. Nucl. Instrum. Meth., A447:1–8, 2000.

[34] A. Affolder et al. Intermediate silicon layers detector for the CDF experiment. Nucl. Instrum. Methods A, 453:84–88, 2000.

[35] L. Balka et al. The CDF central electromagnetic calorimeter. Nucl. Instrum. Methods A, 267:272, 1988.

[36] S. Bertolucci et al. The CDF central and endwall hadron calorimeter. Nucl. Instrum. Methods A, 267:301, 1988.

[37] M. G. Albrow et al. The CDF plug upgrade electromagnetic calorimeter: Test beam results. Nucl. Instrum. Methods A, 480:524–546, 2002.

[38] C. M. Ginsburg. CDF Run 2 muon system. Eur. Phys. J., C33:s1002–s1004, 2004.

[39] G. Ascoli et al. CDF central muon detector. Nucl. Instrum. Meth., A268:33, 1988.

[40] T. Dorigo. The muon system upgrade for the CDF II experiment. Nucl. Instrum. Meth., A461:560–562, 2001.

[41] M. Albrow et al. CDF Run II triggers table and datasets plan. Internal CDF Note 4718 (unpublished), 2001.

[42] G. Gomez-Ceballos et al. Event builder and level 3 at the CDF experiment. Nucl. Instrum. Meth., A518:522–524, 2004.

[43] V. Martin. High p_T muons, recommended cuts and efficiencies for release 5.3. Internal CDF Note 7031 (unpublished), 2004.

[44] V. Martin. High-p_T muon ID cuts and efficiencies for use with 5.3.1 data and 5.3.3 MC. Internal CDF Note 7367 (unpublished), 2005.

[45] C. Hill et al. Electron identification in offline release 5.3. Internal CDF Note 7309 (unpublished), 2005.

[46] A. Bhatti et al. Determination of the jet energy scale at the Collider Detector at Fermilab. Nucl. Instrum. Methods A, 566:375–412, 2006.

[47] C. Neu. CDF b-tagging: Measuring efficiency and false positive rate, 2006. Presented at TOP 2006: International Workshop on Top Quark Physics, Coimbra, Portugal, 12-15 Jan 2006.

[48] C. Neu et al. Tight, loose and ultratight SecVtx tag rate matrix with 1.2fb^{-1}. Internal CDF Note 8519 (unpublished), 2006.

[49] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa. ALP-GEN, a generator for hard multiparton processes in hadronic collisions. J. High Energy Phys., 07:001, 2003.

[50] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. JHEP, 05:026, 2006.

[51] G. Corcella et al. Herwig 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). J. High Energy Phys., 01:010, 2001.

[52] R. Brun and Carminati C. GEANT detector description and simulation tool. CERN Program Library Writeup W5013 (unpublished), 1993.

[53] R. Veenhof. Garfield, a drift chamber simulation program. Prepared for International Conference on Programming and Mathematical Methods for Solving Physical Problems, Dubna, Russia, 14-19 Jun 1993.

[54] R. Veenhof. Garfield, recent developments. Nucl. Instrum. Meth., A419:726–730, 1998.

[55] G. Grindhammer, M. Rudowicz, and S. Peters. The fast simulation of electromagnetic and hadronic showers. Nucl. Instrum. Meth., A290:469, 1990.

[56] Y. L. Dokshitzer. Calculation of the structure functions for deep inelastic scattering and e+ e- annihilation by perturbation theory in quantum chromodynamics. (in russian). Sov. Phys. JETP, 46:641–653, 1977.

[57] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. Sov. J. Nucl. Phys., 15:438–450, 1972.

[58] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. Nucl. Phys., B126:298, 1977.

[59] J. Lueck. Kinematics of electroweak single top quark production. Master's thesis, Universitaet Karlsruhe, 2006. FERMILAB-MASTERS-2006-01.

[60] M. L. Mangano et al. Matching parton showers and matrix elements. hep-ph/0602031, 2006.

[61] M. L. Mangano. Alpgen, a generator for hard multiparton processes in hadronic collisions. JHEP 0307 (2003) 001, 2006.

[62] S. Budd et al. Mistag model used for single-top summer 2006 analyses. Internal CDF Note 8490 (unpublished), 2006.

[63] S. Elbracht et al. Fitting the flavor composition of lepton trigger data in bins of MET and lepton isolation. Internal CDF Note 9131 (unpublished), 2007.

[64] Dante Amidei et al. Calibration of heavy-flavor production in W+1 jet data. Internal CDF Note 9187 (unpublished), 2008.

[65] J. Adelman et al. Method II for you. Internal CDF Note 9185 (unpublished), 2008.

[66] Evelyn J. Thomson et al. TopNtuple: a public analysis module for the Top group for CDF Run II. Internal CDF Note 6737 (unpublished), 2003.

[67] M. Feindt. A neural Bayesian estimator for conditional probability densities. http://www.citebase.org/abstract?id=oai:arXiv.org:physics/0402093, 2004.

[68] M. Feindt and U. Kerzel. The NeuroBayes neural network package. Nucl. Instrum. Meth., A559:190–194, 2006.

[69] V. Blobel and E. Lohrmann. Statistische und numerische Methoden der Datenanalyse. Teubner Stuttgart Leibzig, 1998.

[70] S. Richter. Suche nach elektorschwacher Produktion einzelner Topquarks mit dem CDF-Experiments in Run II. PhD thesis, Universitaet Karlsruhe, 2007. FERMILAB-THESIS-2007-35.

[71] Th. Peiffer. Search for the Higgs boson in the WH channel with the CDF II experiment. Master's thesis, Universitaet Karlsruhe, 2006. unpublished.

[72] T. Walter. Suche nach elektroschwacher Erzeugung einzelner Top-Quarks mit dem CDF-Experiment in Proton-Antiproton-Kollisionen bei 1.96 TeV Schwerpunktsenergie. PhD thesis, Universitaet Karlsruhe, 2005. FERMILAB-THESIS-2005-28.

[73] The CDF Collaboration. Combination of cdf top quark pair production cross section measurements with up to 760 pb$^-$1. Internal CDF Note 8148 (unpublished), 2006.

[74] M. Jezabek. Top quark physics. Nucl. Phys. Proc. Suppl., 37B:197, 1994.

[75] A. Abulencia et al. Top quark mass measurement using the template method in the lepton + jets channel at CDF II. Phys. Rev. D, 73:032003, 2006.

[76] V. M. Abazov et al. Search for single top quark production at DØ using neural networks. Phys. Lett., B517:282–294, 2001.

[77] B. Abbott et al. Search for electroweak production of single top quarks in $p\bar{p}$ collisions. Phys. Rev., D63:031101, 2001.

[78] D. Acosta et al. Search for single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. Phys. Rev., D65:091102, 2002.

[79] D. Acosta et al. Optimized search for single top quark production at the Fermilab Tevatron. Phys. Rev., D69:052003, 2004.

[80] D. Acosta et al. Search for electroweak single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. <u>Phys. Rev.</u>, D71:012005, 2005.

[81] V. M. Abazov et al. Search for single top quark production in p anti-p collisions at s**(1/2) = 1.96 TeV. <u>Phys. Lett.</u>, B622:265–276, 2005.

[82] M. Buehler et al. Single-top search with neural networks for summer 2006. Internal CDF Note 8335 (unpublished), 2007.

[83] V. M. Abazov et al. Evidence for production of single top quarks and first direct measurement of |V(tb)|. <u>Phys. Rev. Lett.</u>, 98:181802, 2007.

[84] D0 collaboration. Updated combination results from three single top quark cross section measurements using the blue method. Public D0 note 5396, 2007.

[85] P. Dong et al. Search for single top quark production using the matrix element technique. Internal CDF Note 9117 (unpublished), 2008.

[86] S. Budd et al. Likelihood function single top analysis 1.9 fb-1. Internal CDF Note 9173 (unpublished), 2008.

Danksagung

An erster Stelle möchte ich Herrn Priv.-Doz. Dr. Wolfgang Wagner sowie Herrn Prof. Dr. Thomas Müller dafür danken, daß sie es mir ermöglicht haben meine Diplomarbeit in ihrer Arbeitsgruppe zu schreiben. Es war mit Sicherheit der Teil meines Studiums, der mich am meisten beeinflußt hat und an den ich mich immer gerne zurückerinnern werde. Eine besonders schöne und lehrreiche Erfahrung war dabei der Besuch des Fermilabs in den USA, für den ich hier noch einmal danken möchte.

Besonderer Dank gebührt Herrn Dr. Dominic Hirschbühl, von dessen Erfahrung und physikalischem, sowie algorithmischen Verständnis ich enorm profitieren konnte. Außerdem bedanke ich mich bei ihm für das zeitaufwendige Korrekturlesen meiner Arbeit, zu deren Abschluß er mit vielen Anregungen und Verbesserungsvorschlägen beigetragen hat.

Vielen Dank an Jan Lück, der in jeder Hinsicht ein ausgezeichneter Kollege ist und mit dem ich gerne zusammengearbeitet habe.

Großer Dank gilt der gesamten Topgruppe. Insbesondere danke ich Dr. Svenja Richter für die vielen Fragen, die sie mir beantwortet hat, sowie meinen weiteren Zimmergenossen Adonis Papaikonomou und Jasmin Kiefer.

Ich bedanke mich für die angenehme Arbeitsatmosphäre am EKP und für die zahlreichen Hilfestellung, die man jeder Zeit und in jedem Zimmer des achten und neunten Stockwerks erhält. Hierbei seien insbesondere das Adminteam sowie das Sekretariat erwähnt.

Zu guter letzt danke ich meiner Familie für ihre finanzielle sowie moralische Unterstützung.

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

Irja Schall

Karlsruhe, 22. Februar 2008