# Search for Electroweak Single-Top Quark Production with the CDF II Experiment

Matthias Bühler

## Diploma Thesis

at the Faculty for Physics
of the University of Karlsruhe

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

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

August 30, 2006

# Introduction

The CDF II experiment and the Tevatron proton-antiproton collider are parts of the Fermi National Laboratories (Fermilab). The Fermilab is located in the vicinity of Chicago, USA. Today, the Tevatron is the only collider which is able to produce the heaviest known elementary particle, the top quark.

The top quark was discovered at the Tevatron by the CDF and the DØ collaborations in 1995 [1]. So far, all the top quarks found are produced via the strong interaction as top-antitop pairs. The Standard Model of elementary particle physics also predicts single-top quark production via the electroweak interaction. This production mode has not yet been observed. The CDF and the DØ collaborations have set upper limits on the cross section for that process in Run I [2, 3] and improved those results in Run II [4, 5].

Single-top quark production is one of the major interests in Run II of the Tevatron as it offers several ways to test the Standard Model and to search for potential physics beyond the Standard Model. The measurement of the cross section of singly produced top quarks via the electroweak interaction offers the possibility to determine the Cabbibo-Kobayashi-Maskawa (CKM) matrix element $V_{tb}$ directly. The CKM matrix defines the transformation from the eigenstates of the electroweak interactions to the mass eigenstates of the quarks. $V_{tb}$ gives the strength of the coupling at the $Wtb$ vertex. The single-top quark is produced at this vertex and therefore the cross section of the single-top quark production is directly proportional to $|V_{tb}|^2$. In the Standard Model, three generations of quarks and the unitarity of the CKM matrix are predicted. This leads to $V_{tb} \approx 1$. Up to now, there is no possibility to measure $V_{tb}$ without using the assumption that there are a certain number of quark generations. Since the measurement of the cross section of single-top quark production is independent of this assumption it could verify another prediction of the Standard Model or give hints towards physics beyond the Standard Model such as a fourth generation of quarks. In addition, electroweak single-top quark production is an important background for the Higgs boson search in the mass range of 90 GeV/$c^2$ to 130 GeV/$c^2$ at the Tevatron in the $WH$ channel.

Two single-top quark production modes are dominant at the Tevatron, the $t$-channel or $W$-gluon fusion and the $s$-channel or $W^*$ process. Since it is challenging to separate the signal from the various background events we use a neural network to combine several variables into one powerful discriminant. The simulated Monte Carlo sample outputs of the neural networks are used as templates for a likelihood fit to the outputs of the neural networks of the data. In this thesis CDF II data corresponding to an integrated luminosity of 695 pb$^{-1}$ is discussed. As this analysis yields no significant evidence of electroweak single-top production it is not possible

to measure any cross sections. Consequently we determine upper limits on the cross sections of the $t$- and $s$-channel production separately and combined.

This thesis is organized as follows. After giving a brief overview of the theoretical backgrounds in the first chapter, the Tevatron and the CDF II experiment are described in the second chapter. The third chapter provides details on the used Monte Carlo and data samples, and on the event selection. Chapter four presents the event yield and the signal and background estimation. Chapter five introduces the technique used to discriminate between signal and background events. Since we use a neural network package, the general usage of this package is described, as well as the particular neural network training which is done in this analysis. In addition, the method to create the templates needed in the likelihood function is shown. Chapter six incorporates the description of the likelihood function used and explains the two different approaches to determine upper limits on the single-top quark production cross section, the combined and the separate search. The results and an outlook can be found in the last chapter.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Top Quark Production within the Standard Model

The Standard Model of Elementary Particle Physics (SM) describes the fundamental particles of matter and their interactions except gravitation. The SM uses a quantum field theory to explain the interactions between particles [6, 7, 8].

In the past, the SM has been very successful in predicting a vast majority of properties of particles and interactions in different energy ranges explored by high energy collider experiments such as DØ and CDF at the Tevatron, H1 and ZEUS at HERA in Hamburg, Germany, or the four LEP experiments in Geneva, Switzerland. The following sections will give a brief overview of the SM, the top-antitop production and especially the single-top quark production via the electroweak interaction.

## 1.1 The Standard Model of Elementary Particle Physics

There are four fundamental forces which are the source of any known dynamics in nature: the strong force, the electromagnetic force, the weak force and the gravitation (table 1.1).

| force | couples to | effect | rel. strength | range |
|-------|-----------|--------|---------------|-------|
| strong | color charge | binds quarks and gluons | $10^0$ | $10^{-15}$ m |
| electro-magnetic | electric charge | interaction between el. charged particles | $10^{-2}$ | infinite |
| weak | weak charge | radioactive decay | $10^{-5}$ | sub nuclear |
| gravitation | mass | attraction of masses | $10^{-38}$ | infinite |

Table 1.1: The four fundamental forces in nature and their most important characteristics [9].

Three out of those four forces are explained in the SM (the strong, the weak and the electromagnetic force). Gravitation is explained by Einsteins Theory of General

Relativity. The SM provides a very elegant theoretical framework using quantum fields to describe the interactions. In these quantum field theories, the fundamental forces of the SM are transmitted by gauge bosons. The gauge bosons are exchanged between the matter particles.

The SM predicts two kinds of particles:

- Fermions: These are particles with spin $s = \frac{1}{2}$ and they follow the Fermi statistics. All matter particles in the SM are fermions.

- Gauge Bosons which are the carriers of the forces. They are particles with spin $s = 1$ and they follow the Bose statistics.

Fermions come as quarks or leptons. They can be ordered in three generations or families. Like all elementary particles, quarks and leptons are described by quantum numbers. Each generation consists of two quarks, a charged lepton and its neutrino. The different quark-types are called flavors. The SM predicts six different quark flavors. Three of them are so called up type quarks (up, charm and top quark) and the other three are down type quarks (down, strange and bottom quark). The three different generations and some of their major characteristics are listed in table 1.2. For all the mentioned particles we also have to consider the existence of their antiparticles, the antiquarks and antileptons.

| name | category | symbol | el. charge | mass |
|:---:|:---:|:---:|:---:|---:|
| up | quark | $u$ | $\frac{2}{3}$ | $(1.5 - 3.0)$ |
| down | quark | $d$ | $-\frac{1}{3}$ | $(3.0 - 7.0)$ |
| electron | lepton | $e$ | $-1$ | $0.511$ |
| $e$-neutrino | lepton | $\nu_e$ | $0$ | $< 2 \cdot 10^{-6}$ |
| charm | quark | $c$ | $\frac{2}{3}$ | $(1.25 \pm 0.09) \cdot 10^3$ |
| strange | quark | $s$ | $-\frac{1}{3}$ | $(95 \pm 25)$ |
| muon | lepton | $\mu$ | $-1$ | $106$ |
| $\mu$-neutrino | lepton | $\nu_\mu$ | $0$ | $< 0.190$ |
| top | quark | $t$ | $\frac{2}{3}$ | $(172.5 \pm 2.3) \cdot 10^3$ |
| bottom | quark | $b$ | $-\frac{1}{3}$ | $(4.20 \pm 0.07) \cdot 10^3$ |
| tau | lepton | $\tau$ | $-1$ | $1777$ |
| $\tau$-neutrino | lepton | $\nu_\tau$ | $0$ | $< 18.2$ |

Table 1.2: Properties of the fermions (spin-$\frac{1}{2}$ particles) [10]. The electric charge is in units of the positron charge, the mass is in units of MeV/$c^2$.

The quarks carry an additional quantum number, the color charge, which comes in three different types: red, green and blue. Free colored particles are not observed in nature; quarks always form colorless bound states, which are called hadrons. There are two different kinds of hadrons:

- Mesons which consist of two quarks e.g. $\pi^+$ (*up* ('red'), *antidown* ('antired')).

- Baryons which consist of three quarks with different colors, e.g. protons (*up* ('red'), *up* ('green'), *down* ('blue')) or neutrons (*up* ('green'), *down* ('red'), *down* ('blue')).

The quarks feel all the forces described by the SM. The leptons on the other hand do not feel the strong force. This is because in the SM the fundamental forces are transmitted by gauge bosons. Only particles that carry the charge of a force can interact via this force, e.g. the gluon, the carrier of the strong force, does only couple to particles which have a color charge. Since leptons do not have a color charge they do not feel the strong force. A summary of the gauge bosons of the three forces described by the SM and some of their properties are given in table 1.3.

| name | force | symbol | el. charge | mass |
|--------|-----------------|-----------|------------|-------------------|
| gluon | strong | g | 0 | 0 |
| photon | electromagnetic | $\gamma$ | 0 | 0 |
| W boson | weak | $W^{\pm}$ | $\pm 1$ | $80.403 \pm 0.029$ |
| Z boson | weak | $Z^{0}$ | 0 | $91.188 \pm 0.002$ |

Table 1.3: Properties of the gauge bosons (spin-1 particles) [10]. The electric charge is in units of the electron charge, the mass is in units of GeV/$c^2$.

Two charged particles interact by the emission and reabsorption of a gauge boson. These processes can be described by Feynman diagrams. In the diagrams only elementary particles (quarks, leptons and bosons) are allowed. The Feynman diagrams visualize the physical process in the momentum space. There are Feynman rules used to translate the Feynman diagram into formulas which calculate the transition amplitude $\mathcal{M}$ for the corresponding physical process. With the transition amplitude $\mathcal{M}$ it is possible to derive the cross section of the physical process by integrating over all initial and final states, the phase space. As an example the Feynman diagram for electron-electron scattering via the exchange of a virtual photon is shown in figure 1.1.
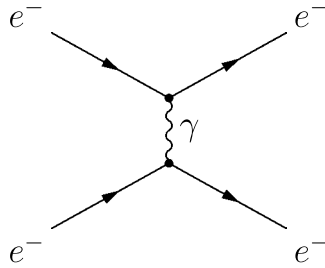


Figure 1.1: Feynman diagram for electron-electron scattering. At the left hand side are the two incoming electrons and on the right hand side are the scattered final state electrons. The interaction between the two electrons is mediated by the exchange of a photon.

As mentioned above the Feynman rules are based on elementary particles as initial and final states. At the Tevatron, the colliding particles are protons and antiprotons

which are composite particles. The virtual gluons inside the protons and antiprotons which hold the valence quarks together can split into quark-antiquark pairs, the seaquarks. This leads to the situation, that the momentum of the proton $p_p$ is shared by all three valence quarks, seaquarks and gluons. These components of the proton or antiproton are also called partons. The fraction of the momentum $x_i = \frac{p_i}{p_p}$, carried by each quark and gluon, is described by the Parton Distribution Function (PDF) $f_{i,p}(x_i, \mu^2)$. It depends on the scale $\mu$, which characterises the typical energy regime of the interaction under consideration. For top quark production it is usually set to the order of the top quark mass, $\mu = m_t$. Figure 1.2 shows the CTEQ5M1 parton distribution function for the scale $\mu^2 = (175 \text{ GeV})^2$ [11]. These PDFs have to be folded with the partonic cross sections to calculate the measurable cross section in $p\bar{p}$ collisions.
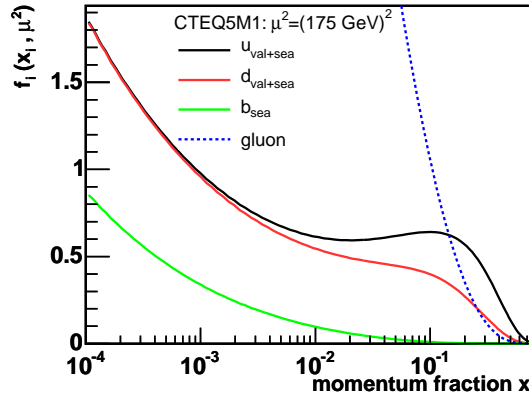


Figure 1.2: The CTEQ5M1 parton distribution function at $\mu^2 = (175 \text{ GeV})^2$ [11].

Another important aspect of the SM is the Cabibbo-Kobayashi-Maskawa mixing matrix (CKM-Matrix). This matrix is relevant for electroweak processes. There was experimental evidence found that the mass eigenstates $(s, d, b)$ of the quarks are not equivalent to the flavor eigenstates $(s', d', b')$. The transformation of mass eigenstates into flavor eigenstates is accomblished by multiplicaton with a $3 \times 3$ matrix, the CKM-matrix [12].

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

The elements $V_{q1q2}$ of the matrix have to be determined experimentally and are proportional to the coupling of two quarks $q_1$, $q_2$ to a $W$ boson. The most recent values are given in (1.2) for 90 % Confidence Level [10]. The value for $V_{tb}$ is determined by exploiting the unitarity of the CKM-matrix, not by direct measurement.

$$\begin{pmatrix} 0.9739 \text{ to } 0.9751 & 0.221 \text{ to } 0.227 & 0.0029 \text{ to } 0.0045 \\ 0.221 \text{ to } 0.227 & 0.9730 \text{ to } 0.9744 & 0.039 \text{ to } 0.044 \\ 0.0048 \text{ to } 0.014 & 0.037 \text{ to } 0.043 & 0.9990 \text{ to } 0.9992 \end{pmatrix} \tag{1.2}$$

## 1.2  Production Modes of the Top Quark in the Standard Model

The top quark belongs to the third generation of quarks and is the heaviest and most recently found one. It has been discovered by the CDF and the DØ collaborations in 1995 [1]. The best current measurement of the top quark mass is $m_{\text{top}} = 172.5 \pm 2.3$ GeV/$c^2$ [13]. Like every quark, the top quark takes part in electroweak and in strong interactions. Due to that, there are several different production modes for top quarks. The next section gives a brief overview of the pair production via the strong force. Afterwards the single-top production via the electroweak force is discussed in more detail.
A detailed review on top quark physics can be found in reference [14].

### 1.2.1  Top Quark Pair Production

In proton-antiproton collisions the dominant production mode for top quarks is the pair production mode via the strong interaction. The cross section for this production is approximately 2.5 times higher than for the electroweak production of single-top quarks. Two leading order perturbation theory Feynman graphs are given in figure 1.3.
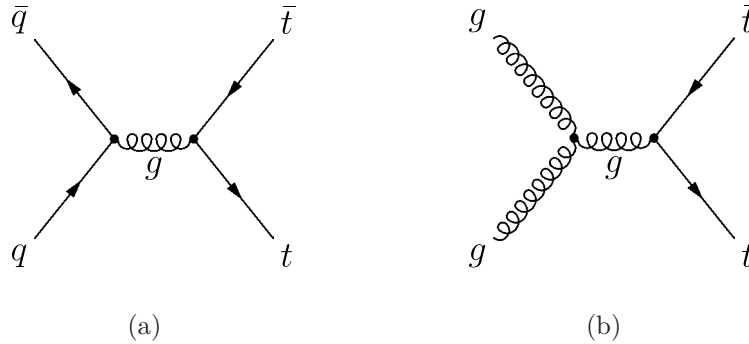


Figure 1.3: Some leading-order Feynman diagrams for top quark pair production: (a) quark-antiquark annihilation and (b) gluon fusion.

The predicted $t\bar{t}$ cross section for a top mass of 175 GeV/$c^2$ at the Tevatron is 6.7±0.9 pb [15]. 80-90% of this cross section is due to the quark-antiquark annihilation shown in figure 1.3 (a). This also includes next to leading order calculations. The uncertainty in the prediction is due to the chosen set of parton distribution function parameterizations of the proton and the uncertainty in choosing the interaction scale $Q^2$ for $t\bar{t}$ events.

## 1.2.2   Single-Top Quark Production

The production of single-top quarks via electroweak interactions is predicted by the SM but has not yet been observed. This is due to a larger amount of background events with a very similar signature as the single-top production compared to the pair production of top quarks. There are two dominating production modes for single-top quarks at the Tevatron. One is the $t$-channel or $W$-gluon fusion process and the other one is the $s$-channel or $W^*$ process. Those two processes will be outlined here in more detail. For completeness also the third process, the associated production, will be described briefly. This process can be neglected in the search at the Tevatron because its cross section is too tiny but it will contribute significantly at the Large Hadron Collider (LHC). The two main production modes are labeled by the relevant Mandelstam variables $t$ and $s$ involved in the transition matrix elements $\mathcal{M}$. The production involves the $Wtb$ vertex for all production modes. They can be distinguished by the virtuality $Q^2 = -q^2$ of the $W$ boson (q is the four momentum of the participating $W$ boson).

- $t$-channel ($q^2 = t$)
  In the $t$-channel or $W$-gluon fusion a space-like (virtual) $W$ boson ($q^2 = t < 0$) strikes a b quark inside the proton or antiproton. The name $W$-gluon fusion originates from the fact that the $b$ quark is a seaquark which is produced through a gluon splitting into a $b\bar{b}$ pair. The $t$-channel is the dominant mode at the Tevatron as well as at the LHC. The leading order (LO) and some next-to-leading order (NLO) Feynman graphs are shown in figure 1.4. The cross section for the $t$-channel production is predicted to be $1.98^{+0.28}_{-0.22}$ pb at a center of mass energy of $\sqrt{s} = 1.96$ TeV and a top mass of $m_{\text{top}} = 175$ GeV/$c^2$ [16, 17] in $p\bar{p}$ collisions.
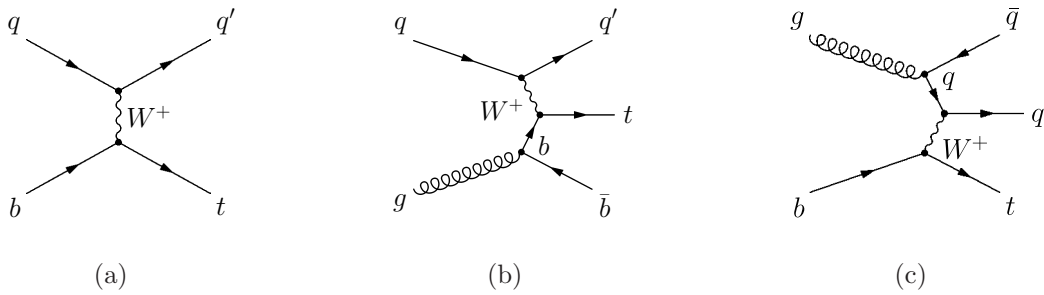


Figure 1.4: The leading order Feynman diagram and some NLO Feynman diagrams for t-channel single-top quark production: leading order (a) and NLO (b), (c).

- $s$-channel ($q^2 = s$)
  In the $s$-channel or $W^*$ production a time-like $W$ boson with $q^2 \geq m_{top} + m_b$ is produced by the fusion of two quarks. The leading order and some next-to-leading order Feynman graphs are shown in figure 1.5. The cross section

is predicted to be $0.88 \pm 0.11$ pb under the same premises as outlined for the $t$-channel [16, 17].
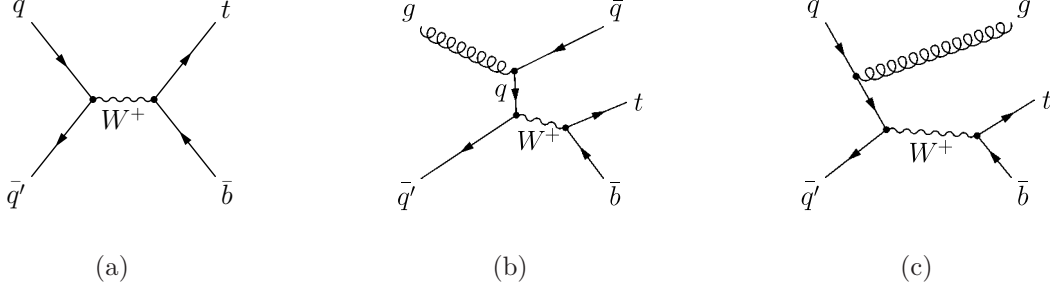


Figure 1.5: The leading order Feynman diagram and some NLO Feynman diagrams for s-channel single-top quark production: leading order (a) and NLO (b), (c).

- associated production
  As the top quark is produced in association with a real, or close to real, $W$ boson in the associated production this mode can be distinguished from the other two. The leading order Feynman graph is shown in figure 1.6. The initial state $b$ quark is a seaquark again. As mentioned above, the associated production can be neglected at the Tevatron, but will have a significant contribution at the LHC.
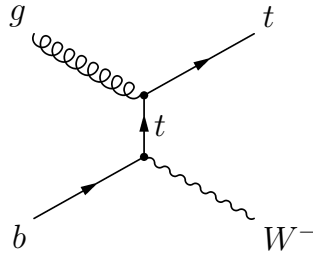


Figure 1.6: The leading order Feynman diagram for associated production.

There are several other production modes for single-top quarks via the electroweak interaction but they are all strongly suppressed by tiny CKM matrix elements as they include vertices like $Wts$ and $Wtd$.

## 1.3 Fourth Generation Quarks Scenario

The Standard Model of elementary particle physics describes all known processes in the field of high energy physics, but it is also known to have some weaknesses in answering some fundamental questions. Several extensions to the Standard Model

have been suggested to address these questions. Some of these theories are also known as theories about so called new physics. Most of them are Grand Unification Theories which have the goal to explain the unanswered questions of the Standard Model and especially to find the one point where all the gauge couplings meet.

One extension of the Standard Model which is closely related to the search for single-top quark production is the fourth generation scenario. This far, most of the determinations of the CKM matrix elements in the third row, e.g. $V_{tb}$, can only be done with the help of the unitarity constraints of the CKM matrix. The recent measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency by the CDF collaboration [18] allows to determine the ratio between $V_{td}$ and $V_{ts}$ without any unitarity constraints to:

$$|V_{td}/V_{ts}| = 0.208^{+0.001}_{-0.002}(\text{exp}) \ ^{+0.008}_{-0.006}$$

Another fraction of third row elements can be determined through top quark pair production.

$$\frac{|V_{tb}|}{\sqrt{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}} = \sqrt{\frac{BR(t \to Wb)}{\sum_{q=d,s,b} BR(t \to Wq)}} \tag{1.3}$$

$BR$ stands for branching ratio which is the ratio between the decay rate of an individual decay mode and the total decay rate. Equation 1.3 is not a direct measurement of $V_{tb}$. It can only be used to measure $V_{tb}$ if the unitarity constraint $|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2 = 1$ is fulfilled. The single-top quark production cross section on the other hand is directly proportional to $|V_{tb}|^2$. Therefore, the measurement of this cross section would yield a direct measurement of $V_{tb}$. A fourth generation of quarks would change the CKM matrix to a $4 \times 4$ matrix which would still be unitary but would of course not fulfill the unitarity constraints of the $3 \times 3$ CKM matrix.

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} & V_{uY} \\ V_{cd} & V_{cs} & V_{cb} & V_{cY} \\ V_{td} & V_{ts} & V_{tb} & V_{tY} \\ V_{Xd} & V_{Xs} & V_{Xb} & V_{XY} \end{pmatrix} \tag{1.4}$$

Since all the measurements of the third row CKM matrix elements are based on those constraints, their values would not be well known anymore in case of a fourth generation of quarks. Therefore, the direct measurement of the matrix element $V_{tb}$ would give the possibility to restrict the value of the matrix elements $V_{tY}$ and $V_{Xb}$. The results of this analysis were already used to constrain the CKM matrix element values $|V_{td}|$ and $|V_{ts}|$ [19]. This far those results are less restricting than the unitarity constraints given for a $4 \times 4$ matrix:

$$|V_{td}| \leq 0.03, \qquad |V_{ts}| \leq 0.2, \qquad |V_{tb}| \geq 0.8$$

Even though the fourth generation scenario is a possible and interesting approach to explain some aspects in high energy physics which are not understood yet, there

are several constraints additional to the ones arising from the unitarity. The fourth generation is only allowed to have very little mixing with the other generations. The quarks of the fourth generation have to be heavier than the top quark and the corresponding charged lepton and neutrino have to be heavier than the $Z$ boson [20].

# Chapter 2

# The Experiment

The CDF II experiment is located at the Tevatron collider at the Fermi National Accelerator Laboratory, abbreviated Fermilab, in Batavia, Illinois (USA). Batavia itself is located in the Western vicinity of Chicago. Fermilab is used by more than 2,500 scientists from throughout the world to carry out research in the areas of high-energy physics and astrophysics. The Tevatron is the accelerator with the highest center-of-mass energy currently in operation in the world. The next section describes the accelerator in more detail. After that the CDF II detector is explained. An aerial photo of the Fermilab area is shown in figure 2.1.



Figure 2.1: Aerial shot of the Tevatron main ring. The CDF site in this view is located at the eight o'clock position direct at the outside of the ring, the DØ site is at the twelve o'clock position. The Fermilab main building is visible in the foreground.

## 2.1 The Accelerators

The Tevatron produces collisions of protons and antiprotons. In 1996, after collecting an integrated luminosity of 106 pb$^{-1}$ of data in the so called Run I, the first

experimental evidence of the top quark was found. It was followed by the precise determination of its mass. Starting in 1996, the accelerator complex was upgraded to increase the instantaneous luminosity and the center-of-mass energy. Luminosity is an important quantity to characterize a collider. The number $n$ of produced events in a given period is given by $n = \sigma \int \mathcal{L} \, dt$. The quantity $\mathcal{L}$ is called instantaneous luminosity and $\int \mathcal{L} \, dt$ is the integrated luminosity over time. The collider characteristics and goals for Run II can be found in table 2.1.

| Run | II(goals) |
|---|---|
| Colliding bunches | $(36 \times 36)$ |
| Beam Energy [GeV] | 980 |
| Antiproton Bunches | 36 |
| $\beta^*$ [cm] | 35 |
| Bunch Length(rms) [cm] | 37 |
| Bunch Spacing [ns] | 396 |
| Interactions/Crossing | 2.3 |
| Typical Luminosity [cm$^{-2}$s$^{-1}$] | $8.6 - 16.1 \times 10^{31}$ |

Table 2.1: Operational goals for Run II of the Tevatron [21, 22].

In the beginning of Run II the Tevatron did not meet the designed goals, partially because the new Main Injector was not well understood. With more knowledge about the accelerator the instantaneous luminosity also increased. The amount of data delivered by the Tevatron and written to tape is presented as a function of time in figure 2.2.
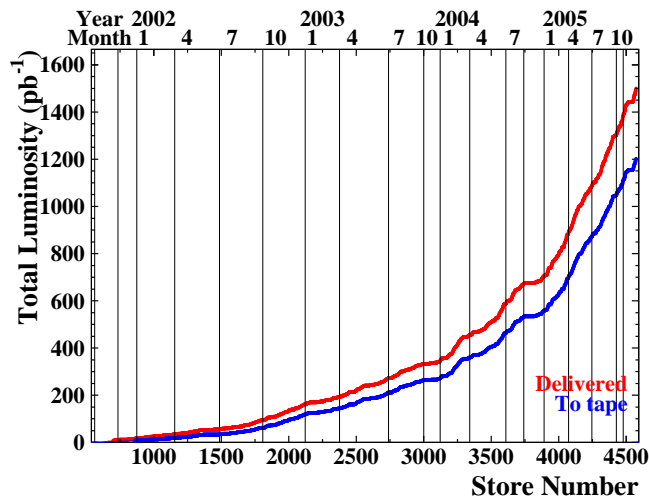


Figure 2.2: Delivered (upper curve) and recorded (lower curve) integrated luminosity since the start of Run II until the end of 2005. This thesis uses all data until September 2005.

CDF and the second Tevatron experiment DØ, were upgraded as well. Run II

started in June 2001 and is supposed to run until 2009. The baseline goal is to achieve an integrated luminosity of 4.4 fb$^{-1}$, the design integrated luminosity goal is 8.4 fb$^{-1}$. As it looks now, the design integrated luminosity will be reached. The design integrated luminosity for the end of fiscal year 2005 was 1.2 fb$^{-1}$, a goal that was attained. The Tevatron is running at a center-of-mass energy of 1.96 TeV. To reach such high energies there needs to be a chain of accelerators leading to the final accelerator, the Tevatron. A schematic overview of the accelerators is shown in figure 2.3.
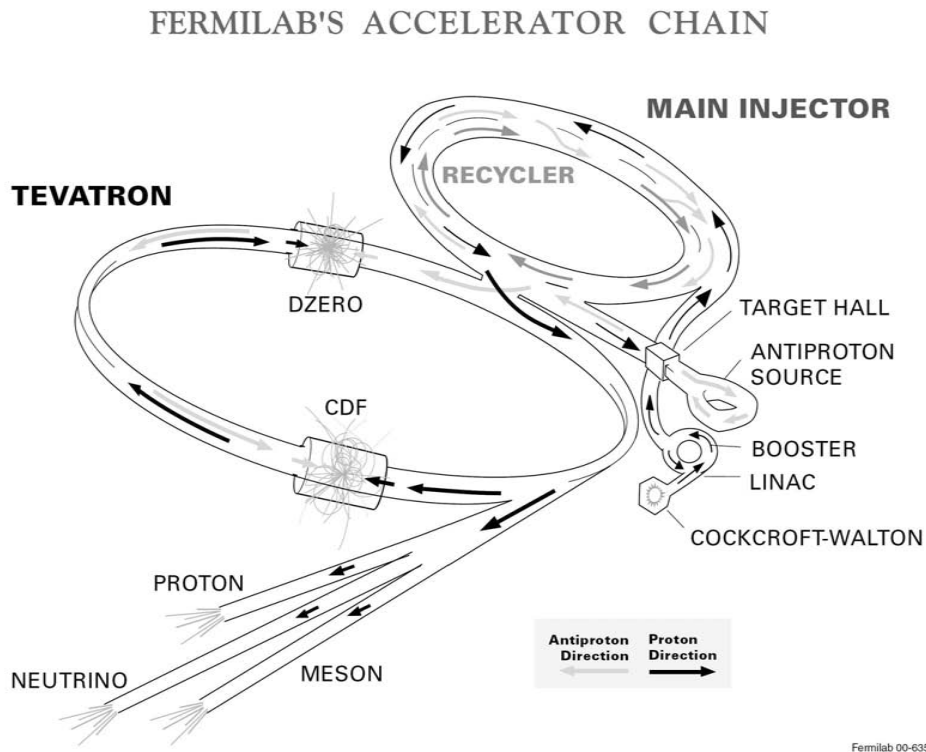


Figure 2.3: Schematic overview of the accelerator chain for Run II of the Tevatron.

The first accelerator in the chain is the Cockcroft-Walton pre-accelerator. In the Cockcroft-Walton pre-accelerator hydrogen gas is ionized to create negative ions. These negative ions are accelerated by a positive voltage to an energy of 750 keV. The pre-accelerated ions now enter a linear collider (LINAC) which is roughly 130 m long. The acceleration to 400 MeV is done by means of an oscillating electric field. The next step in the accelerator chain is the Booster. Before the ions go to the Booster they are sent through a carbon foil where the electrons are stripped off and therefore only protons enter the next accelerator. The Booster is a circular accelerator which accelerates the beam of protons within some 20.000 revolutions to an energy of 8 GeV. Like every circular accelerator the Booster also uses magnets to bend the beam onto a circular path. The protons are then sent to the Main Injector where they are accelerated from 8 GeV to 150 GeV. Some of the protons are not accelerated up to 150 GeV but are used for the antiproton production. These are also

accelerated in the Main Injector but only until they reach 120 GeV. Once they get to this energy they are sent to the antiproton source, where the protons collide with a nickel target. The collision produces a wide range of secondary particles including antiprotons. The antiprotons are collected, focused, and stored in the Accumulator ring at an energy of 8 GeV. As soon as a sufficient number of antiprotons has been produced they are sent back to the Main Injector tunnel which also incorporates the so-called Recycler. The Recycler was originally designed to collect the antiprotons coming back from the Tevatron. This plan has been abandoned. Now the purpose of the Recycler is to stack and cool fresh antiprotons before they are sent to the Main Injector for further acceleration up to 150 GeV. The electron cooling of antiprotons has been introduced in July 2005 and every shot at the Tevatron has passed the electron cooling by the end of August 2005. The final step in the chain of accelerators is the Tevatron which is an accelerator and storage ring. It has a circumference of six kilometers and accelerates the protons and antiprotons to 0.98 TeV. Protons and antiprotons circulate in opposite directions in the ring. That leads to a center-of-mass energy of 1.96 TeV at the two interaction points DO, where the DØ detector is located, and BO, where CDF is located.

## 2.2   The CDF II Detector

The CDF II experiment is a general purpose solenoid detector measuring proton antiproton collisions at the interaction point BØ. The CDF II detector covers most of the $4\pi$ solid angle around the beam spot and it also has azimuthal and forward-backward symmetry. Just like every general purpose solenoid detector the CDF II detector can track charged particles in its core, the tracking system, measure energy deposits in the calorimeters, and identify muons in the muon system. Strong magnetic fields help to identify charged particles. A more detailed description of the CDF detector can be found in its technical design report [23].

The detector was built and is maintained by scientists from 50 different institutions and eleven different countries throughout the world. The Institut für Experimentelle Kernphysik (EKP) in Karlsruhe is the only German institution in the CDF collaboration.

Figure 2.4 shows an elevation view of one half of the CDF II detector. In the following angles and directions are referred to as defined in the CDF coordinate system shown in figure 2.5. The polar angle $\theta$ is measured with respect to the proton beam axis ($z$-axis), pointing in east direction. The azimuthal angle $\varphi$ is measured from the plane of the Tevatron. Transverse and longitudinal are meant with respect to the proton beam, i.e. parallel or perpendicular to the proton beam respectively. An often-used quantity is the pseudorapidity defined by $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$.

CDF II uses the same solenoid as in Run I but the complete tracking system has been replaced. The superconducting solenoid is 4.8 meters long and has a radius of 1.5 meters. It generates a 1.4 Tesla magnetic field.

The tracking system consists of four parts: Layer 00, the SVX II (Silicon Vertex Detector), the ISL (Intermediate Silicon Layers) and the COT (Central Outer
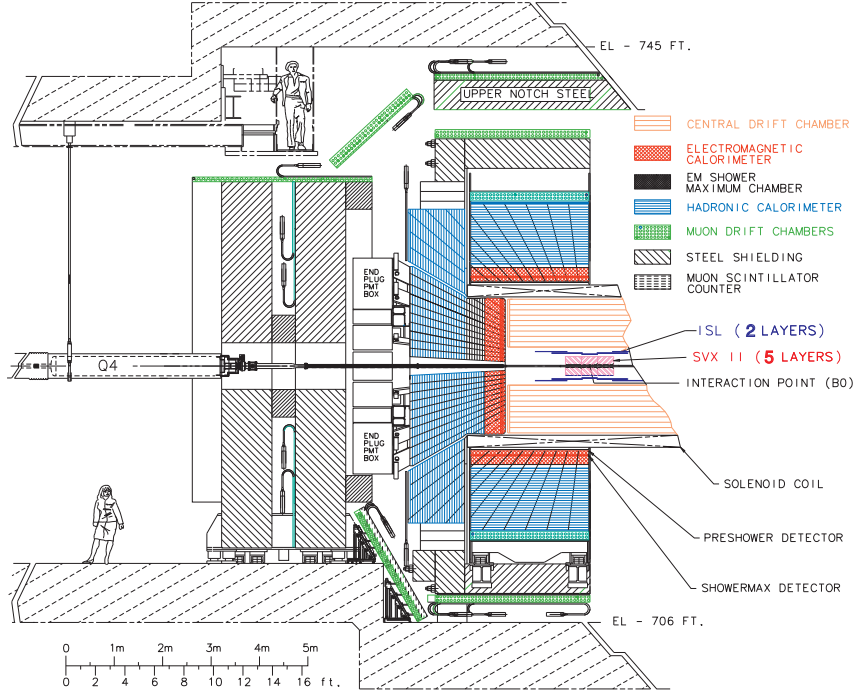
Figure 2.4: Elevation view of one half of the CDF II detector.

Tracker), which is an open drift cell chamber. Layer 00 improves the precision of track measurements and tagging efficiencies. Layer 00 is glued directly to the beam pipe therefore it is closest to the beam, with its modules placed at radii $r = 1.35$ cm and $r = 1.62$ cm of the beam pipe. It is a single-sided radiation hard silicon microstrip detector and provides a coverage of $|\eta| < 4.0$. Layer 00 is enclosed by the SVX II. The SVX II detector design is driven by high luminosity, the Tevatron short bunch spacing of 396 ns and by the physics requirement of $B$ hadron decay vertex identification within collimated high-$P_T$ jets. SVX II is comprised of three cylindrical barrels which cover $\approx 2.5\sigma$ of the interaction region providing track information to pseudorapidity $|\eta| < 2$. Five layers of double-sided silicon sensors at radii from 2.4 to 10.7 cm supply $r - \varphi$ as well as 3 $r - z$ and 2 small angle stereo measurements. The results provide good pattern recognition and 3-d vertex reconstruction with an impact parameter resolution $\sigma_\varphi < 30\mu m$ and $\sigma_{z0} < 70\mu m$ for central high momentum tracks. The impact parameter is the distance of closest approach of the track helix to the beam axis measured in the plane perpendicular to the beam. The SVX II provides coverage up to $|\eta| \approx 2$. In the region $|\eta| < 1$ the combination of the SVX II and the COT can provide full 3D tracking, but the reconstruction will need to be anchored on COT tracks. To increase the tracking volume, the three layers of the silicon detector ISL are placed between the SVX II and the COT.

The outer part of the tracking system is a drift chamber. The Central Outer Tracker (COT) is 3.1 meters long and has radii between 40 and 137 cm. It covers $|\eta| < 1.0$, and can measure particles with transverse momenta as low as 400 MeV/c. All this is done by 96 measurement layers organized into alternating axial and $\pm 2$ stereo
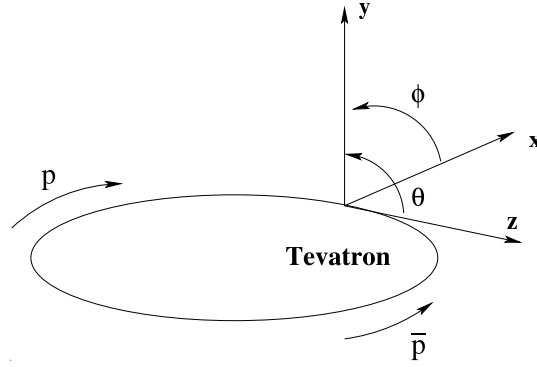
Figure 2.5: The CDF coordinate system.

superlayers. The hit position resolution is approximately $14 \mu m$ and the momentum resolution $\sigma(p_T)/p_T^2 = 0.0015 \, (\text{GeV } /c)^{-1}$. Due to the high luminosity and the short bunch spacing, the COT is designed to operate with a maximum drift time of 100 ns by reducing the maximum drift distance and by using a gas mixture (Argon-Ethane (50:50)) with a fast drift velocity.

Between the COT and the solenoid, a Time-of-Flight system (TOF) is installed mainly for particle identification. It consists of scintillator panels which provide both timing and amplitude information. The timing resolution is 100 ps. The detector covers the central region up to $|\eta| < 1.1$ and is capable of distinguishing kaons from pions by their flight time difference with at least $2\sigma$ separation up to kaon momenta of 1.6 GeV/c.

The tracking system and the solenoid are surrounded by the calorimeters. They are designed to measure the energy of all particles and jets in the detector by absorbing all the particles except muons and neutrinos. There are two main $\eta$ regions covered by the calorimeters, central and plug. Each region has an electromagnetic calorimeter (CEM, Central ElectroMagnetic calorimeter, and PEM, endPlug ElectroMagnetic calorimeter) with lead absorbers and an hadronic calorimeter (CHA/WHA, Central HAdron/end-Wall HAdron calorimeter, and PHA, end-Plug HAdron calorimeter) with iron absorbers. Those calorimeters deliver a coverage of $2\pi$ in azimuth and $|\eta| < 3.64$. Each calorimeter is segmented in azimuth and in pseudorapidity to form a projective tower geometry which points back to the nominal interaction point. The calorimeters are sampling calorimeters with scintillators as active medium.

The most outer part of the detector is the muon system. Figure 2.6 shows a schematic picture of the CDF II muon system. The additional information on the figure which is not explained is not needed here. The system consists of four separate systems of scintillators and drift tubes. The central hadron calorimeters act as absorber for the Central Muon detection System (CMU). Four layers of drift chambers outside of the central hadronic calorimeter form the CMU. It covers a range of
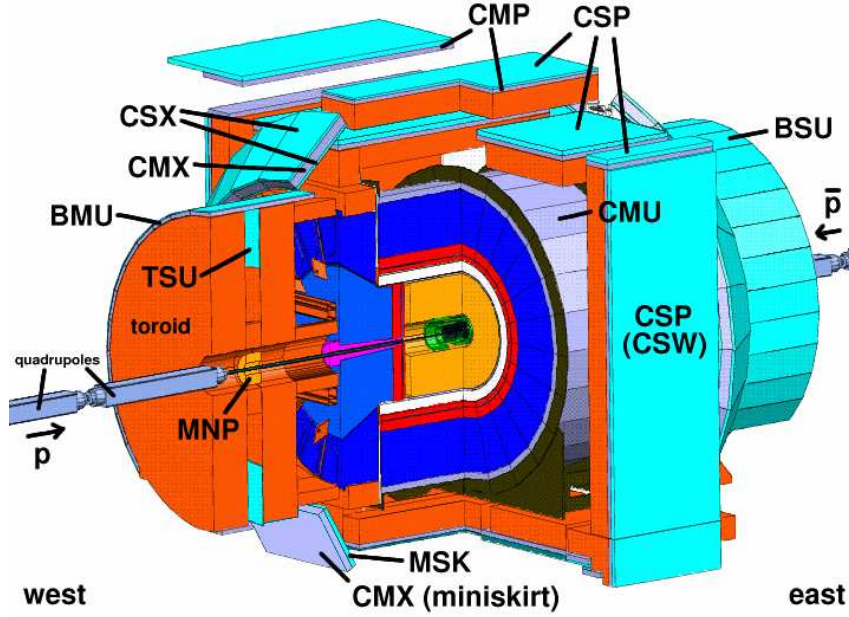
Figure 2.6: The CDF II muon system.

$|\eta| < 0.6$ and muons with a transverse momentum above 1.4 GeV/$c$ can be detected. Four additional layers of drift chambers are located behind a 0.6 m thick absorber layer of steel. These drift chambers are called Central Muon uPgrade (CMP). The CMP covers the same $\eta$ range as the CMU. To cover a bigger $\eta$ range the Central Muon eXtension (CMX) is used. It covers the pseudorapidity range $0.6 < |\eta| < 1.0$. Two of the chambers of the CMX, the miniskirts and keystones, can be seen in figure 2.7. The Intermediate Muon System (IMU) covers a range of $1.0 < |\eta| < 1.5$.

The trigger system plays an important role to efficiently extract the most interesting physics events from the large number of minimum bias events, because the collision rate is equal to the mean crossing rate of 1.7 MHz while, the tape writing speed is about 75 Hz at present. The CDF trigger is a three level system with each level providing a sufficient rate reduction for the processing of the next level.

The first two triggers are hardware triggers, the last one is a software trigger running on a Linux PC farm. Level-1 uses custom designed hardware to find physics objects based on a subset of the detector. The hardware consists of three parallel synchronous processing streams: one to identify calorimeter based objects, another one to identify muons, while the third one does tracking in the COT using the eXtremly Fast Tracker (XFT). The decision is done by simple counting these objects (e.g. one electron with 12 GeV). If an event is accepted by the Level-1 trigger, the data are moved to one of the four on-board Level-2 buffers, to average out the rate fluctuations. The typical rate of the Level-1 triggers is at present 24 kHz.

The Level-2 trigger does a limited event reconstruction using custom-designed hardware. The hardware consists of several asynchronous subsystems, e.g. the hardware
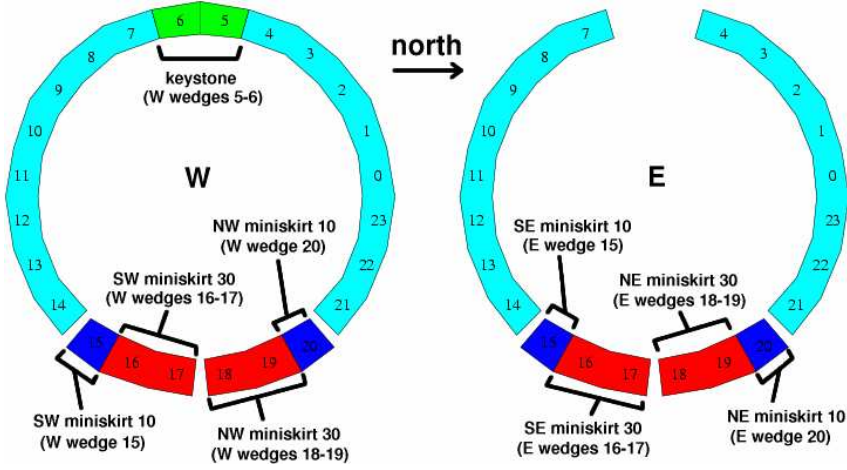
Figure 2.7: Schematic view of the CMX-wedges.

cluster finder using calorimeter information. In addition, data from the shower maximum detector can be used to improve the identification of electrons and photons. The most challenging part of the Level-2 trigger is the Silicon Vertex Tracker. The SVT allows to select tracks with large impact parameter, which opens a complete new window for physics measurements at a hadron collider. The level-2 trigger accepts 300 events per second, which are transferred to the Level-3 processor farm.

At the processor farm the events are reconstructed and filtered, using the algorithms run in the "offline" reconstruction, and are written to permanent storage with approximately 75 Hz at present. To facilitate the handling of the huge data volumes collected with the CDF II detector, events passing the Level-3 trigger are split into eight different streams.

# Chapter 3

# Event Selection

In this chapter, the Monte Carlo samples and the data samples we use in this analysis are explained. In addition, the event selection procedure which all those samples have to pass is outlined.

## 3.1 Monte Carlo Samples

Monte Carlo generators randomly generate hard parton interactions according to the probability density of phase space and the matrix element of the physics process. The resulting partons are then processed by a parton showering program to simulate gluon radiation, fragmentation and decays. The resulting particles are then handed to the detector simulation. The CDF collaboration uses the GEANT3 [24] package to simulate the tracking of particles through matter. To model the drift chamber (COT), a GARFIELD simulation is used, that was tuned to match the data [25]. The calorimeter simulation GFLASH [26] was also tuned using test beam data for electrons and high transverse momentum pions. A more detailed description of the CDF II simulation can be found elsewhere [27].

The Monte Carlo samples used for this analysis are generated by the matrix element generator MADEVENT [28], the Monte Carlo program PYTHIA [29] or a combination of ALPGEN [30] and HERWIG [31]. Simulation and production were performed using the 5.3.3 release of the CDF offline software. The reconstructed events were passed through TopFind, linked against CDF software release 5.3.3_nt, to produce TopNtuples [32]. Table 3.1 gives an overview of the Monte Carlo samples we use. All these Monte Carlo samples were simulated with run dependent settings covering the run range from run number 141544 to 203799. MADEVENT implements the correct polarization of the top quark. The theory of electroweak interactions predicts that single-top quarks are 100% polarized along the $d$ quark direction in the top quark rest frame [33, 34]. This will be outlined in more detail in chapter 5.2. The modeling of the single-top $t$-channel has a particular intricacy. If only the leading order diagram, figure 1.4 (a), is used, the $\bar{b}$ (also called the second $b$) is obtained by backward evolution with the shower Monte Carlo program PYTHIA. The kinematic distributions of the $\bar{b}$ are not well modeled by this procedure. In particular, the $p_\mathrm{T}$ distribution of the second $b$ is too soft, and the $\eta$ distribution is

| Sample | Process | Generator | N events |
|---|---|---|---|
| ttop1oNew | $t$-channel $(2 \rightarrow 2)$ | MadEvent | 400554 |
| ttop2oNew | $t$-channel $(2 \rightarrow 3)$ | MadEvent | 199506 |
| ttop1oNew+ ttop2oNew_match | $t$-channel (matched) | MadEvent | 600060 |
| ttop0oNewCat | $s$-channel | MadEvent | 577276 |
| ttopvl+ttopkl+ttopyl+ ttoptl | $t\bar{t}$ | Pythia | 2965060 |
| wtop1w+wtop2w | $WW$ | Pythia | 606877 |
| wtop1z+wtop2z | $WZ$ | Pythia | 594502 |
| ztopcz+ztopfz | $ZZ$ | Pythia | 809509 |
| ztop7i | $Z \rightarrow ee$ | Pythia | 3251489 |
| ztop0i+ztopei | $Z \rightarrow \mu\mu$ | Pythia | 1794600 |
| ztop1i+ztop4i+ztop5i | $Z \rightarrow \tau\tau$ | Pythia | 3538885 |
| ltop0b | $Wbb + 0p \ (W \rightarrow e\nu_e)$ | Alpgen+Herwig | 315072 |
| ltop1b | $Wbb + 0p \ (W \rightarrow \mu\nu_e)$ | Alpgen+Herwig | 315297 |
| ltop0c | $Wc\bar{c} + 0p \ (W \rightarrow e\nu_e)$ | Alpgen+Herwig | 271274 |
| ltop1c | $Wc\bar{c} + 0p \ (W \rightarrow \mu\nu_e)$ | Alpgen+Herwig | 310336 |
| ltop1a | $Wc + 1p \ (W \rightarrow e\nu_e)$ | Alpgen+Herwig | 263613 |
| ltop2a | $Wc + 1p \ (W \rightarrow \mu\nu_e)$ | Alpgen+Herwig | 217200 |
| ltop1n | $W + 1p \ (W \rightarrow e\nu_e)$ | Alpgen+Herwig | 280109 |
| ltop1m | $W + 1p \ (W \rightarrow \mu\nu_e)$ | Alpgen+Herwig | 232837 |
| ltop2n | $W + 2p \ (W \rightarrow e\nu_e)$ | Alpgen+Herwig | 161434 |
| ltop2m | $W + 2p \ (W \rightarrow \mu\nu_e)$ | Alpgen+Herwig | 184531 |
| ltop3n | $W + 3p \ (W \rightarrow e\nu_e)$ | Alpgen+Herwig | 178293 |
| ltop3m | $W + 3p \ (W \rightarrow \mu\nu_e)$ | Alpgen+Herwig | 173988 |

Table 3.1: Signal and background Monte Carlo samples used in the single-top analysis. In all single-top samples, the $W$ from the top quark decays to leptons, $W \rightarrow e, \mu, \tau + \nu$.

too much forward (shifted to high values of $|\eta|$). The problem can be obviated by generating two $t$-channel samples with MADEVENT, one $2 \to 2$ ($qb \to q't$) and one $2 \to 3$ ($qg \to q'tb$) sample, see figure 1.4 (b). Both samples are mixed to better reproduce the theoretically predicted distributions of the second $b$ quark. The exact mixing prescription is documented in detail in references [16, 35].

## 3.2  Data Samples

Our analyses will focus on the so called lepton+jets channel. Therefore, we use the data stream including all inclusive high-$p_T$ lepton samples. Data is re-processed offline before being used by the different physics groups. This re-processing was done with `CDFSOFT2` version 5.3.3 for the so-called 0d data set and with `CDFSOFT2` version 6.1.2 for the 0h samples. The calibrations used online are checked and corrected, silicon alignment is corrected, tracks are re-fitted, cluster energies are checked, leptons are found, and "tagging" algorithms for secondary vertices were run. After that the data were stripped into the datasets bhel0d, bhel0h, bpel0d, bpel0h, bhmu0d and bhmu0h. The data sets are subdivided into the different sets based on the lepton that has triggered the event and according to the run range the event came from. The central electron data (bhel0d/h) have to pass level 3 trigger `ELECTRON_CENTRAL_18`, the forward electron data (bpel0d/h) have to pass level 3 trigger `MET_PEM` and the muon data (bhmu0d/h) have to pass level 3 trigger `MUON_CMUP18` or `MUON_CMX18`. Those triggers are all high $p_T$ lepton triggers which demand that in case of a electron the transverse energy is $E_T > 18$ GeV and in case of a muon the transverse momentum is $p_T > 18$ GeV/$c$. Other requirements of these trigger paths are a track in the COT and a matching signal in the corresponding calorimeter component for the electrons. The 0d samples include runs from run number 138425 up to 186598 which were taken from April 2002 till August 2004. The 0h samples include the runs from 190697 up to 203799 and were taken in the period from December 2004 till September 2005. For this analysis the good run list version 10a was used. The list is created by the lepton+jets group [36]. It contains all runs, in which all detector components needed for this analysis were functional and included in the data taking process. Additionally, run 164844 was removed due to problems with the *Consumer Server Logger*. The largest samples (CEM and CMUP) correspond to an integrated luminosity of 695.5 pb$^{-1}$.

## 3.3  Event Selection Procedure

In order to separate single-top quark events from various background events it is necessary to reconstruct the events with very high precision. Within the Standard Model, the top quark decays to almost 100% to a $W$ boson and a bottom quark. To suppress Quantum ChromoDynamic (QCD) multijet events the selection focuses on the leptonic decay modes of the $W$ boson. We only consider the decay channels $W \to e\nu_e$ and $W \to \mu\nu_\mu$, because high purity $\tau$ lepton identification has low efficiency at CDF. This means that the characteristic signatures we are looking for in the data are

the leptonic decay products of the $W$ boson originating from the top quark decay, i.e. we are looking for an isolated lepton and missing transverse energy. A lepton is isolated if the energy deposition in the electromagnetic calorimeter in a cone of radius $\Delta R$=0.4 around the cluster, excluding the cluster itself, is determined and the ratio of this energy to the cluster energy is less than 0.1. In addition to the isolated lepton we expect hadronic jets, from which at least one should be identified as a $b$ jet. There are several vetos and requirements to restrict the background even more. The selection mainly follows the standard cuts of the CDF top physics group for lepton+jets analyses [37]. The following sections will show the main requirements used throughout the analysis and point out differences to the standard cuts in lepton+jets analyses. A typical cut flow can be seen in table 3.2 for the matched t-channel sample ttop1oNew+ttop2oNew_match.

### 3.3.1 Lepton Cuts

Since one isolated high $p_T$ lepton originating from the $W$ decay is one of the main characteristics for single-top production, this criteria is the starting point of our cuts. The first thing that needs to be done is to decide whether a lepton candidate fulfills the requirements we need for our event. Leptons are divided into different categories. The ones we need are tight leptons and loose leptons. A lepton is called tight if it belongs to one of the four main lepton categories (CEM, PHX, CMUP and CMX). The electrons are required to have a transverse energy $E_T > 20$ GeV. The CEM electrons cover the $\eta$-range up to $|\eta| < 1.1$ and the PHX electrons the range $1.2 < |\eta| < 2.0$. In case of muons a transverse momentum of $p_T > 20$ GeV/$c$ is required. Muon candidates are accepted in the $\eta$-range of $0.0 < |\eta| < 1.0$. In particular the CMUP muons have $\eta$ values between $0.0 < |\eta| < 0.6$ and the CMX muons $0.6 < |\eta| < 1.0$, respectively. Loose leptons are leptons which fulfill less stringent ID requirements. The lepton cuts also include requirements on the shower profile $L_{\mathrm{shr}}$, the ratio of the energy between the hadronic calorimeter to the electromagnetic calorimeter $E_{\mathrm{had}}/E_{\mathrm{em}}$, the electromagnetic energy divided by the momentum of the maximum $p_T$ matched track E/P, and several other. They can all be found in detail in [37] for the different lepton types.

There is another requirement especially concerning the leptons in each event that has to be fulfilled. An event is only accepted if there is exactly one isolated tight lepton. That means no additional leptons are allowed, neither another tight lepton nor a loose lepton.

### 3.3.2 Event Cuts

In addition to the cuts only dealing with the leptons there are requirements the whole event has to fulfill. The major ones are listed here.

- The $Z$ boson veto is used to reduce the background coming from $Z$+jet production. Therefore, events are rejected in which a combination of the identified tight lepton and a second object found in the event forms an invariant mass, that falls in the window of

$$76 \text{ GeV}/c^2 < M_{ll} < 106 \text{ GeV}/c^2$$

around the $Z$ boson mass.

- The hadronization of a final state quark creates a jet of hadrons. Hadronization describes the transition from colored partons to color neutral hadrons. These particles then form a particle jet. Since the energy of hadrons is measured in the hadronic calorimeters, the momentum of the initial quark can be reconstructed by combining the energy measurements in the calorimeter towers that belong to the jet. In CDF the JetClu algorithm [38] is used to reconstruct jets. Because the reconstruction is distorted by several effects a set of jet corrections has been derived. In our analysis all the jets are level 4 corrected. This means that they are corrected for relative $\eta$-dependence and for multiple interactions. More information on the jet corrections can be found elsewhere [39]. The jets have been reclustered after the removal of the towers assigned to the selected isolated lepton from the leptonic $W$ decay and after a correction of the measured transverse energy in the towers for the location of the primary vertex.
  In $t$-channel single-top events the light flavor jet is emitted in forward direction. Therefore, the $\eta$ range is defined wider than in the standard CDF analyses. The jets have to have $E_T \geq 15$ GeV and $|\eta_{\text{detector}}| \leq 2.8$. The pseudorapidity $\eta_{detector}$ is the pseudorapidity coordinate of the calorimeter tower with respect to the origin of the coordinate system, the center of the detector. Through the extension of the $\eta$ range the acceptance for t-channel events increases about 30%, whereas the background acceptances do not increase much.
  In addition to the jets with $E_T \geq 15$ GeV which are called 'tight' jets there is also a category of jets with $E_T \geq 8$ GeV and $E_T \leq 15$ GeV which are called 'loose' jets. In this analysis we require exactly two 'tight' jets.

- The neutrino from the $W$ decay cannot be detected in any of the parts of the detector. Therefore, there has to be some missing transverse energy ($\not{E}_T$) in the event. Since a neutrino associated with the high momentum lepton from the $W$ boson decay is likely to carry a lot of transverse energy, the missing transverse energy has to fulfill corrected $\not{E}_T \geq 20$ GeV.
  $\not{E}_T$ is calculated as the vector sum of the energy in each calorimeter tower multiplied by the azimuthal direction of the tower. If the identified lepton of such an event is a muon, $\not{E}_T$ is corrected by the subtraction of the muon energy deposit in the calorimeter and by addition of the transverse muon momentum to the vector sum. If we would not apply this correction missing transverse energy would be artificially created. Since corrected jet energies are used for the jet counting, one needs to correct $\not{E}_T$ to be consistent.

- The QCD veto is applied to reduce the QCD background, i.e. events where no real $W$ boson is created. Therefore, those events are referred to as non-$W$ events. We want to reject events with low missing transverse energy and where

the leading jet and the $\not{E}_T$ vector are aligned. In these events $\not{E}_T$ is created by jet mismeasurements. Consequently, we accept all the events with $\not{E}_T > 30$ GeV. Out of the events with $\not{E}_T < 30$ GeV we accept those which have an opening angle $\Delta\phi$ between the $\not{E}_T$ vector and the transverse momentum vector of the leading jet between $0.5 < \Delta\phi < 2.5$. The leading jet is chosen after the jet corrections (level 4) are applied.

- At least one of the jets in the event must be identified to be likely to contain a $B$ hadron. Consequently, at least one $SecVtx$ tagged jet is required [40]. The $b$-tagging relies on the reconstruction of displaced secondary vertices with a silicon micro-strip detector. An event which contains a $B$ hadron is likely to have a displaced secondary vertex, since the bottom hadron carries most of the momentum of the original quark due to its relatively large mass. Consequently, the hadron is boosted and, due to its mean lifetime of approximately 1.5 ps, it travels a sizable distance before it decays.

| Cut | 0 jet | 1 jet | 2 jet | 3 jet | 4 jet | $\geq$ 5 jets | all |
|---|---|---|---|---|---|---|---|
| Total | 11514 | 135394 | 287304 | 136399 | 25299 | 4150 | 600060 |
| Good Run | 10270 | 120666 | 256023 | 121352 | 22473 | 3689 | 534473 |
| OBSV < 60.0 | 10043 | 117878 | 249895 | 117374 | 21735 | 3567 | 520492 |
| CMX Good Total | 9948 | 117253 | 248886 | 118075 | 21893 | 3600 | 519655 |
| $\geq$ 1 Tight Std. lepton | 3600 | 40171 | 71688 | 14446 | 2254 | 294 | 132453 |
| LooseLepton cut | 3598 | 40154 | 71675 | 14446 | 2254 | 294 | 132421 |
| CEM electrons | | | | | | | |
| Tight Di-Lepton Veto | 1566 | 18507 | 34122 | 6852 | 1075 | 133 | 62255 |
| Z Vertex Cut | 1566 | 18504 | 34114 | 6846 | 1074 | 132 | 62236 |
| Z veto | 1555 | 18302 | 33533 | 6740 | 1056 | 129 | 61315 |
| Missing Et | 1466 | 16376 | 30279 | 6056 | 955 | 118 | 55250 |
| QCD Veto | 1466 | 15641 | 28729 | 5744 | 910 | 107 | 52597 |
| taggable | 0 | 11484 | 25675 | 5453 | 890 | 105 | 43607 |
| b tag $\geq$ 1 | 0 | 5522 | 13061 | 3168 | 541 | 80 | 22372 |
| PHX electrons | | | | | | | |
| Tight Di-Lepton Veto | 699 | 6452 | 9725 | 1871 | 279 | 39 | 19065 |
| Z Vertex Cut | 699 | 6452 | 9725 | 1871 | 279 | 39 | 19065 |
| Z veto | 692 | 6401 | 9599 | 1833 | 277 | 39 | 18841 |
| Missing Et | 647 | 5589 | 8367 | 1591 | 244 | 30 | 16468 |
| QCD Veto | 647 | 5237 | 7747 | 1497 | 218 | 29 | 15375 |
| taggable | 0 | 3956 | 7168 | 1453 | 215 | 28 | 12820 |
| b tag $\geq$ 1 | 0 | 1855 | 3582 | 847 | 146 | 13 | 6443 |
| CMUP muons | | | | | | | |
| Tight Di-Lepton Veto | 870 | 10383 | 19368 | 3967 | 620 | 74 | 35282 |
| Z Vertex Cut | 870 | 10383 | 19366 | 3966 | 620 | 74 | 35279 |
| Z veto | 869 | 10345 | 19217 | 3930 | 615 | 73 | 35049 |
| Missing Et | 813 | 9230 | 17130 | 3494 | 546 | 65 | 31278 |
| QCD Veto | 813 | 8798 | 16208 | 3284 | 508 | 61 | 29672 |
| taggable | 0 | 6470 | 14468 | 3107 | 498 | 61 | 24604 |
| b tag $\geq$ 1 | 0 | 3059 | 7215 | 1784 | 313 | 38 | 12409 |
| CMX muons | | | | | | | |
| Tight Di-Lepton Veto | 435 | 4619 | 8171 | 1690 | 262 | 47 | 15224 |
| Z Vertex Cut | 435 | 4619 | 8171 | 1690 | 262 | 47 | 15224 |
| Z veto | 434 | 4594 | 8128 | 1683 | 259 | 47 | 15145 |
| Missing Et | 416 | 4079 | 7212 | 1497 | 225 | 44 | 13473 |
| QCD Veto | 416 | 3873 | 6805 | 1394 | 209 | 40 | 12737 |
| taggable | 0 | 2841 | 6106 | 1345 | 206 | 40 | 10538 |
| b tag $\geq$ 1 | 0 | 1370 | 3116 | 779 | 128 | 18 | 5411 |
| All | | | | | | | |
| Tight Di-Lepton Veto | 3570 | 39961 | 71386 | 14380 | 2236 | 293 | 131826 |
| Z Vertex Cut | 3570 | 39958 | 71376 | 14373 | 2235 | 292 | 131804 |
| Z veto | 3550 | 39642 | 70477 | 14186 | 2207 | 288 | 130350 |
| Missing Et | 3342 | 35274 | 62988 | 12638 | 1970 | 257 | 116469 |
| QCD Veto | 3342 | 33549 | 59489 | 11919 | 1845 | 237 | 110381 |
| taggable | 0 | 24751 | 53417 | 11358 | 1809 | 234 | 91569 |
| b tag $\geq$ 1 | 0 | 11806 | 26974 | 6578 | 1128 | 149 | 46635 |

Table 3.2: Cut flow for the *t*-channel sample ttop1oNew+ttop1oNew_match.

# Chapter 4

# Signal and Background Estimate and Event Yield

One of the most crucial parts in many high energy physics analyses is the estimation of the number of predicted events for the amount of data that is analysed. This theoretical prediction has to be consistent within errors with the actually observed number of events in the data. Otherwise this is a strong hint that the physical processes are not well understood and that the theoretical presumptions need some more careful crosschecks.

The determination of the event yield is not only done for the so-called two jet bin which is the jet bin we focus on in our analysis but also for the one and the three jet bin. This gives us the possibility to cross check our results. The event yield is determined by a combination of two different techniques. For some of the physical processes the number of expected events can be calculated with the help of Monte Carlo samples and theoretical cross sections. The other processes can only be estimated through data. The first approach is only suitable for physical processes that are theoretically very well understood. We use the Monte Carlo based technique to determine the number of predicted events for the signal as well as for $t\bar{t}$- and diboson backgrounds. Those processes are well described through Monte Carlo samples and their cross sections are known for proton antiproton collisions at the center of mass energy of the Tevatron. The second section includes all the backgrounds that are not as well understood as $t\bar{t}$ and diboson or are not even real physical processes but mismeasurements in the detector, namely mistags and non-$W$ background. Therefore, the estimation of the number of expected events has to be done differently. Those backgrounds are estimated based on data. This is why they are referred to as data based backgrounds.

# 4.1   Monte Carlo Based Estimates

## 4.1.1   Signal

In this section, the calculation of the event detection efficiency for single-top quark production is explained. As the same method is also used in section 4.1.2 and 4.1.3 for the so called Monte Carlo based background estimate some aspects of the technique are explained from a general point of view in this section.

To derive the event detection efficiencies for the physical processes mentioned above we make use of a Monte Carlo generator technique. We can use this technique because the processes of single-top production are theoretically well described and already implemented in most of the common Monte Carlo generators. The final aim of every single-top quark analysis should be to calculate the single-top quark production cross section $\sigma_{\text{st}}$ for the reasons outlined in the Introduction. The calculation is based on the following formula:

$$\sigma_{\text{st}} = \frac{N_{\text{signal}}}{(\epsilon_{\text{evt}} \cdot \int \mathcal{L}\, dt)} \tag{4.1}$$

Here $N_{\text{signal}}$ is the number of observed signal events, which we obtain for example from a maximum likelihood fit. $\epsilon_{\text{evt}}$ is the event detection efficiency which is the average probability of a single-top event to be detected, i.e. to be found in the sample of selected candidates. $\int \mathcal{L}\, dt$ is the integrated luminosity, which we will abbreviate as $\mathcal{L}_{\text{int}}$ in the paragraphs below. Equation (4.1) is used to calculate the number of expected signal events. For this purpose it takes the form:

$$N_{\text{signal}}^{\text{predict}} = \sigma_{\text{st}}^{\text{theo}} \cdot \epsilon_{\text{evt}} \cdot \mathcal{L}_{\text{int}} \tag{4.2}$$

The purpose of this section is to describe the calculation of $\epsilon_{\text{evt}}$, the event detection efficiency. This is done using the Monte Carlo samples listed in table 3.1. $\epsilon_{\text{evt}}$ can be decomposed into 4 factors:

$$\epsilon_{\text{evt}} = \epsilon_{\text{evt}}^{\text{MC}} \cdot \epsilon_{\text{BR}} \cdot \epsilon_{\text{corr}} \cdot \epsilon_{\text{trig}} \tag{4.3}$$

Here $\epsilon_{\text{evt}}^{\text{MC}}$ is the event detection efficiency as we obtain it from our samples of simulated events. In some of these samples, the $W$ boson was only allowed to decay into leptons: $W \rightarrow e/\mu/\tau + \nu$. This has to be taken into account by applying the factor $\epsilon_{\text{BR}} = 0.324$ [10]. $\epsilon_{\text{corr}}$ is a correction factor which takes into account the difference between simulated and data events and thereby gives a measure of how well the Monte Carlo simulation models the detector. $\epsilon_{\text{trig}}$ is the trigger efficiency. The correction factor is again composed out of several parts:

$$\epsilon_{\text{corr}} = \frac{\epsilon_{\text{z0}}^{\text{data}}}{\epsilon_{\text{z0}}^{\text{MC}}} \cdot \frac{\epsilon_{\text{ID/Reco}}^{\text{data}}}{\epsilon_{\text{ID/Reco}}^{\text{MC}}} \cdot \frac{\epsilon_{\text{tag}}^{\text{data}}}{\epsilon_{\text{tag}}^{\text{MC}}} \tag{4.4}$$

Since trigger and lepton id efficiencies vary for different subdetectors (we use CEM, PHX, CMUP and CMX), we have different $\epsilon_{\text{evt}}$ in the four cases: $\epsilon_{\text{evt}}^{\text{CEM}}$, $\epsilon_{\text{evt}}^{\text{PHX}}$, $\epsilon_{\text{evt}}^{\text{CMUP}}$ and $\epsilon_{\text{evt}}^{\text{CMX}}$. We use the following values for the trigger efficiencies and reconstruction and identification scale factors which are derived from data:

| Efficiencies | CEM | PHX | CMUP | CMX |
|---|---|---|---|---|
| Trigger | 0.9691±0.0042 | 0.8861±0.0025 | 0.915± 0.005 | 0.9276± 0.0034 |
| ID/Reco. | 0.981± 0.005 | 0.942 ± 0.013 | 0.8789± 0.0044 | 0.9593 ± 0.004 |

Table 4.1: Electron and muon trigger efficiencies and ID efficiency scale (correction) factors.

- The electron and muon trigger and identification efficiencies $\epsilon_{\text{trig}}$ and $\epsilon_{\text{ID/Reco}}^{\text{data}}/\epsilon_{\text{ID/Reco}}^{\text{MC}}$ are listed in table 4.1 and can also be found in references [41, 42, 43]. We apply numbers averaged over the entire run range.

- $\epsilon_{z0}^{\text{data}} = 0.955 \pm 0.003$ is the $z$ vertex cut efficiency in data [44].

- The $b$ tagging efficiency differs between data and Monte Carlo. Therefore, we need to correct our acceptance calculation which is based on Monte Carlo events. $SF_{\text{tag}} = \epsilon_{\text{tag-jet}}^{\text{data}}/\epsilon_{\text{tag-jet}}^{\text{MC}} = 0.89 \pm 0.07$ is the correction factor for the $b$ tagging efficiency [45]. This correction factor is valid per tagged $b$ jet. If a Monte Carlo sample contained only events with one and only one $b$ jet per event, the factor would be applicable globally. However, since there are also double-tagged or even triple-tagged events, a global correction factor has to be determined. It is possible to bin the distribution of the number of $b$ tags in the Monte Carlo samples and then use the per $b$ tag jet scale factor given above to calculate the event correction factor given per jet bin. If the fractions of 1-tag, 2-tag, and 3-tag events in a given jet bin are $f_1$, $f_2$, and $f_3$ , then for that jet bin the correction factors are calculated as shown below.

$$K_1 = SF_{tag} + 2 \cdot \frac{f_2}{f_1} SF_{tag}(1 - SF_{tag}) + 3 \cdot \frac{f_3}{f_1} SF_{tag}(1 - SF_{tag})^2 \qquad (4.5)$$

$$K_2 = SF_{tag}^2 + 3 \cdot \frac{f_3}{f_2} SF_{tag}^2 (1 - SF_{tag}) \qquad (4.6)$$

$$K_{12} = f_1 SF_{tag} + f_2(2SF_{tag} - SF_{tag}^2) + f_3(3SF_{tag} - 3SF_{tag}^2 + SF_{tag}^3) \qquad (4.7)$$

The results on the correction factor $K = \epsilon_{\text{tag}}^{\text{data}}/\epsilon_{\text{tag}}^{\text{MC}}$ obtained with this method are given in table 4.2 for 3 cases: (1) 1 or 2 $b$ tags ($K_{12}$), (2) exactly 1 $b$ tag ($K_1$), (3) exactly 2 $b$ tags ($K_2$).

| B-tag Efficiencies | | | | |
|---|---|---|---|---|
| Process | Sample | $K_{12}$ | $K_1$ | $K_2$ |
| t-chan 1 Jet bin | ttop1oNew+ttop2oNew_match | 0.89±0.07 | 0.89±0.07 | – |
| t-chan 2 Jet bin | ttop1oNew+ttop2oNew_match | 0.89±0.07 | 0.89±0.07 | 0.79±0.12 |
| t-chan 3 Jet bin | ttop1oNew+ttop2oNew_match | 0.91±0.06 | 0.93±0.05 | 0.79±0.13 |
| s-chan 1 Jet bin | ttop0oNewCat | 0.89±0.07 | 0.89±0.07 | – |
| s-chan 2 Jet bin | ttop0oNewCat | 0.92±0.06 | 0.96±0.03 | 0.79±0.12 |
| s-chan 3 Jet bin | ttop0oNewCat | 0.92±0.05 | 0.96±0.03 | 0.79±0.13 |

Table 4.2: Correction factor for the $b$ tagging efficiency of the various single-top samples.

We determine the event detection efficiency based on Monte Carlo events. We apply all selection and identification cuts to the simulated data. Table 4.3 summarizes the calculation for the $t$-channel. The number of remaining events in the 1-jet, 2-jet and 3-jet bin after the complete event selection are given in the line named `b tag` $\geq$ `1`. The number of expected events in every subdetector for a luminosity of $695.5\,\mathrm{pb}^{-1}$ can be found in the line named $N_{\mathrm{evt}}$. We use the number of events passing the good run criteria as denominator (line `Good Run`).

To calculate $\epsilon_{\mathrm{evt}}$ we need to derive $\epsilon_{\mathrm{corr}}$ first. We have to do that for each sample separately, since the $b$ tagging efficiency per event depends on the sample. We calculate the errors on $\epsilon_{\mathrm{corr}}$ by adding the relative errors on $\epsilon_{\mathrm{z0}}^{\mathrm{data}}/\epsilon_{\mathrm{z0}}^{\mathrm{MC}}$, $\epsilon_{\mathrm{leptonid}}^{\mathrm{data}}/\epsilon_{\mathrm{leptonid}}^{\mathrm{MC}}$, $\epsilon_{\mathrm{reco}}^{\mathrm{data}}/\epsilon_{\mathrm{reco}}^{\mathrm{MC}}$ and $\epsilon_{\mathrm{tag}}^{\mathrm{data}}/\epsilon_{\mathrm{tag}}^{\mathrm{MC}}$. Having calculated $\epsilon_{\mathrm{corr}}$ we compute $\epsilon_{\mathrm{evt}}$ based on equation (4.3). Using the values for $\epsilon_{\mathrm{evt}}$ we calculate the number of expected events according to equation (4.2). The result is presented in table 4.3 for the $t$-channel. The Monte Carlo event detection efficiency is given separately for the different subdetectors. The last row in each subdetector section gives the number of expected events in the corresponding subdetector for a luminosity of $695.5\,\mathrm{pb}^{-1}$. $\epsilon_{\mathrm{evt}}$ refers to the total event detection efficiency including the corresponding scale factors. The last section of the table summarizes the number of expected events. The numbers given refer to an integrated luminosity of $695.5\,\mathrm{pb}^{-1}$. The number of expected events for the $t$-channel and the $s$-channel in the different subdetectors are also summarized in tables 4.6 - 4.10.

## 4.1.2   $t\bar{t}$ Background

As mentioned in section 1.2.1 the top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV is dominated by top-pair production via the strong interaction. Quark-antiquark annihilation is the dominating sub-process, contributing about 85% of the cross section. Next-to-leading order (NLO) corrections to the cross section are available since the late 1980's [46, 47]. More recent calculations try to improve the predictions by resumming leading and next-to-leading order logarithmic terms in the cross section. Table 4.4 shows the predictions of three different groups for $m_{\mathrm{top}} = 175$ GeV/$c^2$. The results of Berger and Contopanagos (BECO) are scaled down from their predictions for $\sqrt{s} = 2.00$ TeV. In our analysis we use the prediction by Bonciani et al. (BCMN) [49, 50] to calculate the number of expected $t\bar{t}$ events. Two reasons motivate that decision:

1. BCMN work with the most recent set of PDFs.

2. The error assigned by BCMN includes systematic uncertainties due to the choice of the PDF.

To take into account different predictions by BECO and Kidonakis we assign half the difference between BCMN and BECO as additional systematic uncertainty $\Delta_2 = 0.23$ pb and add it in quadrature to the error assigned by BCMN. Since we use a Gaussian constraint on the background in our analysis we also need to symmetrize the errors. We do that by taking the average between the negative and positive errors. Additionally, we need to add the uncertainty in $\sigma_{t\bar{t}}$ due to the top mass

| | $t$-channel | | |
|---|---|---|---|
| Cut | 1 jet | 2 jet | 3 jet |
| Total | 135394 | 287304 | 136399 |
| Good Run | 120666 | 256023 | 121352 |
| OBSV< 60.0 | 117878 | 249895 | 117374 |
| $K_1$ | 0.89±0.07 | 0.89±0.07 | 0.93±0.05 |
| $K_2$ | 0.89±0.07 | 0.79±0.12 | 0.79±0.13 |
| $K_{12}$ | 0.89±0.07 | 0.89±0.07 | 0.91±0.06 |

CEM electrons

| b tag $\geq 1$ | 5522 | 13061 | 3168 |
|---|---|---|---|
| $\epsilon_{\text{evt}}^{MC}$ [%] | 1.033 | 2.444 | 0.593 |
| $\epsilon_{\text{evt}}$ [%] | 0.2685±0.0262 | 0.6296±0.0606 | 0.1551±0.0128 |
| $N_{evt}$ | 3.70±0.36 | 8.67±0.83 | 2.14±0.18 |

PHX electrons

| b tag $\geq 1$ | 1855 | 3582 | 847 |
|---|---|---|---|
| $\epsilon_{\text{evt}}^{MC}$ [%] | 0.347 | 0.670 | 0.158 |
| $\epsilon_{\text{evt}}$ [%] | 0.079±0.0089 | 0.1516±0.0168 | 0.0364±0.0035 |
| $N_{evt}$ | 1.09±0.12 | 2.09±0.23 | 0.50±0.05 |

CMUP Muons

| b tag $\geq 1$ | 3059 | 7215 | 1784 |
|---|---|---|---|
| $\epsilon_{\text{evt}}^{MC}$ [%] | 0.572 | 1.350 | 0.334 |
| $\epsilon_{\text{evt}}$ [%] | 0.1258±0.0145 | 0.2942±0.0335 | 0.0739±0.0072 |
| $N_{evt}$ | 1.73±0.20 | 4.05±0.46 | 1.02±0.10 |

CMX Muons

| b tag $\geq 1$ | 1370 | 3116 | 779 |
|---|---|---|---|
| $\epsilon_{\text{evt}}^{MC}$ [%] | 0.256 | 0.583 | 0.146 |
| $\epsilon_{\text{evt}}$ [%] | 0.0624±0.0065 | 0.1406±0.0144 | 0.0357±0.0032 |
| $N_{evt}$ | 0.84±0.09 | 1.90±0.20 | 0.48±0.04 |

**Number of expected events**

| $N(695.50\text{pb}^{-1})$ | 7.36±0.77 | 16.71±1.72 | 4.14±0.37 |
|---|---|---|---|

Table 4.3: $t$-channel event yield estimate for the single-top event selection.

| | $m_{top}$ | $\sigma_{t\bar{t}}$ | Ref. |
|---|---|---|---|
| Berger and Conto. | 175 GeV/$c^2$ | $7.15^{+0.09}_{-0.55}$ pb | [48] |
| Bonciani et al. | 175 GeV/$c^2$ | $6.70^{+0.71}_{-0.88}$ pb | [49, 50] |
| Kidonakis | 175 GeV/$c^2$ | $6.77 \pm 0.42$ pb | [15] |
| Bonciani et al. | 170 GeV/$c^2$ | $7.83^{+0.86}_{-1.04}$ pb | [49, 50] |
| Bonciani et al. | 180 GeV/$c^2$ | $5.75^{+0.59}_{-0.75}$ pb | [49, 50] |

Table 4.4: Cross section predictions by three different groups of theorists for $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV.

uncertainty $\Delta m_{\text{top}} = 5$ GeV/$c^2$. We take the average difference between the cross section for 170/175 GeV/$c^2$ and 180/175 GeV/$c^2$, which is $\Delta_3 = 1.04$ pb. Adding all three uncertainties in quadrature we get:

$$\Delta \sigma_{t\bar{t}} = \sqrt{0.795^2 + 0.23^2 + 1.04^2} \text{ pb} = 1.32 \text{ pb} \qquad (4.8)$$

Thus, we use $\sigma_{t\bar{t}} = (6.70 \pm 1.32)$ pb. Including the scaled uncertainty due to the other

| | |
|---|---|
| 170 GeV/$c^2$ | $(7.83 \pm 1.54)$ pb |
| 175 GeV/$c^2$ | $(6.70 \pm 1.32)$ pb |
| 180 GeV/$c^2$ | $(5.75 \pm 1.13)$ pb |

Table 4.5: Cross section predictions used in our analysis to predict the number of $t\bar{t}$ background events.

theoretical predictions (also for $m_{\text{top}} = 170, 180$ GeV/$c^2$) we get the numbers given in table 4.5. The event detection efficiency $\epsilon_{\text{event}}$ for $t\bar{t}$ events is calculated using the PYTHIA Monte Carlo generator. We use the $t\bar{t}$ Monte Carlo samples listed in table 3.1. To determine the number of expected $t\bar{t}$ events we use the method described in section 4.1.1. The results are summarized in tables 4.6 - 4.10.

### 4.1.3   Diboson Background

To predict the number of diboson events in our selected data sample we use the theoretical cross sections predicted by Campbell and Ellis [51]:

| $\sqrt{s}$ | WW | WZ | ZZ |
|---|---|---|---|
| 2.00 TeV | 13.5 pb | 4.02 pb | 1.60 pb |
| 1.96 TeV | 13.30 pb | 3.96 pb | 1.57 pb |

Campbell and Ellis give a relative error on the cross sections of 3%. Their numbers are given for $\sqrt{s} = 2.00$ TeV. We rescale those numbers to 1.96 TeV. Therefor, the mean of a linear and a quadratic interpolation is taken. The number of expected diboson events is calculated in the same way as for single-top and $t\bar{t}$ production. We also determine the number of expected events for the processes $Z \to \tau\tau$ and $Z \to \mu\mu$. For these processes, the cross section is determined using the Z peak for

the invariant mass region of the lepton pair and and a K-factor of 1.4. This leads to $\sigma_Z \cdot \mathrm{BR}(Z \to \ell^+\ell^-) = (336.0 \pm 8.0)\,\mathrm{pb}$. The uncertainty due to the luminosity is not included, since it will be considered when calculating the number of expected events. We use the Monte Carlo samples listed in table 3.1. The results are shown in tables 4.6 - 4.10.

| Process | 1 jet | 2 jets | 3 jets |
|---|---|---|---|
| | for 695.5 pb$^{-1}$ of data in CEM | | |
| $t$-channel | $3.70 \pm 0.36$ | $8.67 \pm 0.83$ | $2.14 \pm 0.18$ |
| $s$-channel | $1.75 \pm 0.17$ | $6.04 \pm 0.44$ | $1.25 \pm 0.09$ |
| $t\bar{t}$ | $2.53 \pm 0.25$ | $20.29 \pm 1.61$ | $46.32 \pm 3.46$ |
| $WW$ | $0.77 \pm 0.07$ | $1.86 \pm 0.18$ | $0.51 \pm 0.05$ |
| $WZ$ | $1.25 \pm 0.12$ | $2.85 \pm 0.24$ | $0.65 \pm 0.05$ |
| $ZZ$ | $0.01 \pm 0.00$ | $0.06 \pm 0.00$ | $0.05 \pm 0.00$ |
| $Z \to \tau\tau$ | $0.93 \pm 0.09$ | $1.60 \pm 0.16$ | $0.69 \pm 0.07$ |
| $Z \to \mu\mu$ | $0.00 \pm 0.00$ | $0.00 \pm 0.00$ | $0.00 \pm 0.00$ |
| Total MC bkg. | $5.49 \pm 0.30$ | $26.66 \pm 1.65$ | $48.22 \pm 3.46$ |

Table 4.6: Summary of number of expected events for $t$-channel and $s$-channel single-top, $t\bar{t}$ and diboson production in the CEM detector.

| Process | 1 jet | 2 jets | 3 jets |
|---|---|---|---|
| | for 695.5 pb$^{-1}$ of data in PHX | | |
| $t$-channel | $1.09 \pm 0.12$ | $2.09 \pm 0.23$ | $0.50 \pm 0.05$ |
| $s$-channel | $0.42 \pm 0.05$ | $1.33 \pm 0.11$ | $0.26 \pm 0.02$ |
| $t\bar{t}$ | $0.52 \pm 0.06$ | $4.70 \pm 0.43$ | $10.31 \pm 0.89$ |
| $WW$ | $0.27 \pm 0.03$ | $0.73 \pm 0.08$ | $0.19 \pm 0.02$ |
| $WZ$ | $0.64 \pm 0.07$ | $1.39 \pm 0.13$ | $0.34 \pm 0.03$ |
| $ZZ$ | $0.00 \pm 0.00$ | $0.03 \pm 0.00$ | $0.01 \pm 0.00$ |
| $Z \to \tau\tau$ | $1.02 \pm 0.11$ | $0.00 \pm 0.00$ | $0.20 \pm 0.02$ |
| $Z \to \mu\mu$ | $0.00 \pm 0.00$ | $0.00 \pm 0.00$ | $0.00 \pm 0.00$ |
| Total MC bkg. | $2.45 \pm 0.15$ | $6.85 \pm 0.46$ | $11.05 \pm 0.89$ |

Table 4.7: Summary of number of expected events for $t$-channel and $s$-channel single-top, $t\bar{t}$ and diboson production in the PHX detector.

| Process | 1 jet | 2 jets | 3 jets |
|---|---|---|---|
| | for 695.5 pb$^{-1}$ of data in CMUP | | |
| $t$-channel | $1.73 \pm 0.2$ | $4.05 \pm 0.46$ | $1.02 \pm 0.10$ |
| $s$-channel | $0.83 \pm 0.10$ | $2.87 \pm 0.25$ | $0.62 \pm 0.05$ |
| $t\bar{t}$ | $1.27 \pm 0.15$ | $10.78 \pm 1.01$ | $24.67 \pm 2.18$ |
| $WW$ | $0.42 \pm 0.05$ | $0.83 \pm 0.10$ | $0.26 \pm 0.03$ |
| $WZ$ | $0.52 \pm 0.06$ | $1.25 \pm 0.12$ | $0.32 \pm 0.03$ |
| $ZZ$ | $0.03 \pm 0.00$ | $0.09 \pm 0.01$ | $0.02 \pm 0.00$ |
| $Z \to \tau\tau$ | $0.59 \pm 0.07$ | $0.78 \pm 0.09$ | $0.58 \pm 0.07$ |
| $Z \to \mu\mu$ | $5.92 \pm 0.68$ | $3.33 \pm 0.37$ | $0.25 \pm 0.03$ |
| Total MC bkg. | $8.75 \pm 0.70$ | $17.06 \pm 1.09$ | $26.15 \pm 2.18$ |

Table 4.8: Summary of number of expected events for $t$-channel and $s$-channel single-top, $t\bar{t}$ and diboson production in the CMUP detector.

| Process | 1 jet | 2 jets | 3 jets |
|---|---|---|---|
| | for 695.5 pb$^{-1}$ of data in CMX | | |
| $t$-channel | $0.84 \pm 0.20$ | $1.90 \pm 0.20$ | $0.48 \pm 0.04$ |
| $s$-channel | $0.39 \pm 0.04$ | $1.29 \pm 0.10$ | $0.27 \pm 0.02$ |
| $t\bar{t}$ | $0.54 \pm 0.06$ | $4.58 \pm 0.39$ | $10.37 \pm 0.83$ |
| $WW$ | $0.23 \pm 0.02$ | $0.39 \pm 0.04$ | $0.13 \pm 0.01$ |
| $WZ$ | $0.37 \pm 0.04$ | $0.61 \pm 0.05$ | $0.19 \pm 0.02$ |
| $ZZ$ | $0.00 \pm 0.00$ | $0.04 \pm 0.00$ | $0.01 \pm 0.00$ |
| $Z \to \tau\tau$ | $0.43 \pm 0.04$ | $0.21 \pm 0.02$ | $0.21 \pm 0.02$ |
| $Z \to \mu\mu$ | $3.91 \pm 0.41$ | $1.11 \pm 0.11$ | $0.28 \pm 0.03$ |
| Total MC bkg. | $5.48 \pm 0.42$ | $6.94 \pm 0.42$ | $1.19 \pm 0.84$ |

Table 4.9: Summary of number of expected events for $t$-channel and $s$-channel single-top, $t\bar{t}$ and diboson production in the CMX detector.

| Process | 1 jet | 2 jets | 3 jets |
|---|---|---|---|
| | for $695.5\,\mathrm{pb}^{-1}$ of data | | |
| $t$-channel | $7.36 \pm 0.77$ | $16.71 \pm 1.72$ | $4.14 \pm 0.37$ |
| $s$-channel | $3.40 \pm 0.36$ | $11.52 \pm 0.91$ | $2.41 \pm 0.19$ |
| $t\bar{t}$ | $4.85 \pm 0.51$ | $40.34 \pm 3.45$ | $91.66 \pm 7.37$ |
| $WW$ | $1.68 \pm 0.18$ | $3.81 \pm 0.40$ | $1.10 \pm 0.12$ |
| $WZ$ | $2.77 \pm 0.29$ | $6.09 \pm 0.55$ | $1.50 \pm 0.13$ |
| $ZZ$ | $0.05 \pm 0.01$ | $0.21 \pm 0.02$ | $0.10 \pm 0.01$ |
| $Z \to \tau\tau$ | $2.96 \pm 0.32$ | $2.59 \pm 0.27$ | $1.68 \pm 0.18$ |
| $Z \to \mu\mu$ | $9.84 \pm 1.09$ | $4.44 \pm 0.48$ | $0.53 \pm 0.06$ |
| Total MC bkg. | $22.15 \pm 1.29$ | $57.48 \pm 3.56$ | $96.57 \pm 7.37$ |

Table 4.10: Summary of number of expected events for $t$-channel and $s$-channel single-top, $t\bar{t}$ and diboson production.

## 4.2    Data Based Background Estimates

We estimate the $W$+jets background and the non-$W$ background with the standard method used in the CDF top physics group, the so-called *method 2* [52]. The first step is the estimation of the non-$W$ fraction in the pretag sample. The sample is called pretag because the events have to pass all the criteria outlined in section 3.3 except for the requirement of a $b$ tagged jet.

### 4.2.1    Non-W

Since the cross section for jet production via QCD processes is much larger than the cross section for the electroweak production of $W$ bosons we still have a substantial rate of multi-jet events passing our selection cuts. In these events, a jet is mis-identified as an electron or muon. This is the so-called non-$W$ background. To reduce this background, the QCD veto was introduced.

We use the standard method to estimate the rate of non-$W$ events: the $\not{E}_T$-versus-Isolation method, which we call here the four sector method. The basic assumption of this method is that the variables $\not{E}_T$ and isolation are uncorrelated for multi-jet events. Four sectors of the $\not{E}_T$ versus isolation plane are used for the analysis. They are shown and defined in figure 4.1.

1. sector A: iso $\geq 0.2$ and $\not{E}_T \leq 15$ GeV

2. sector B: iso $\leq 0.1$ and $\not{E}_T \leq 15$ GeV

3. sector C: iso $\geq 0.2$ and $\not{E}_T \geq 20$ GeV

4. sector D: iso $\leq 0.1$ and $\not{E}_T \geq 20$ GeV
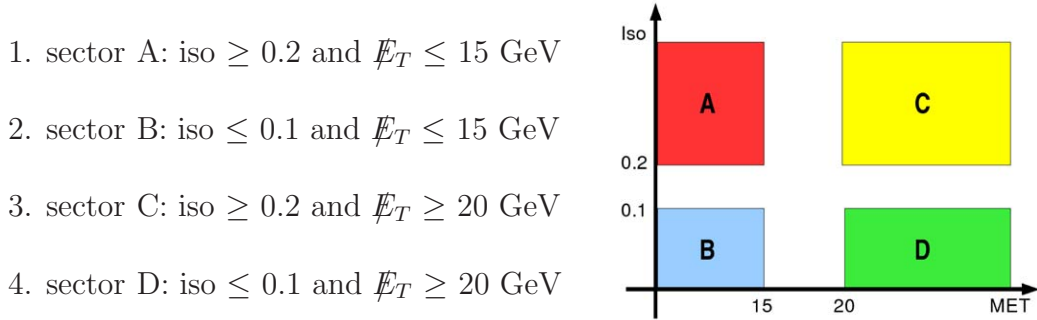


Figure 4.1: Definition of the four sectors used in the four sector method.

We apply our selection criteria to the data samples, omitting the cuts on isolation of the lepton, the $\not{E}_T$ and the $b$ tag. We count the number of events in the four sectors. Sector D is the signal region, sectors A, B and C are sideband regions. The analysis is performed separately for four different lepton categories, i.e. CEM, PHX, CMUP, and CMX.

Events with non-isolated leptons need a special treatment compared to the standard selection of events with isolated leptons:

1. For non-isolated electrons, we remove jets from the jet list that match the electron candidate within $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$. This is relevant for the jet counting. Otherwise the same physical object would be counted twice, once

as an electron and secondly as a jet. These matched jets are also not used to correct the missing transverse energy.

2. For non-isolated muons we apply a correction to the $\not{E}_T$. We subtract the muon $p_T$ vectorially, but add the muon calorimeter energy vectorially back in.

The raw event rates in the sideband regions are corrected for the signal content, i.e. the number of $W$+jets events in the sidebands. To predict the distribution of $W$+jets events in the sectors we use $W$+jets Monte Carlo events. The Monte Carlo samples we use are: ltop1n, ltop2n, ltop3n, ltop1m, ltop2m, and ltop3m listed in table 3.1.

The non-$W$ fraction $f_{\text{nonW}}$ in the signal sector D is then predicted using the corrected event counts in the sideband regions.

$$f_{\text{nonW}} = \frac{N_B^{\text{corr}} \cdot N_C^{\text{corr}}}{N_A^{\text{corr}} \cdot N_D} \tag{4.9}$$

The corrected event count in the background regions is obtained from:

$$N_i^{\text{corr}} = N_i - M_i \frac{S_D}{M_D} \tag{4.10}$$

where $S_D$ is the number of signal events in sector D, which is given by $S_D = N_D \cdot (1 - f_{\text{nonW}})$. $M_i$ is the number of Monte Carlo events in region $i$. Solving this equation we find an ambiguity. The unphysical solution (negative number of events) is dismissed. The results for the non-$W$ fractions are given in table 4.11. The

|      | 1 jet | 2 jets | 3 jets |
|------|-------|--------|--------|
| CEM  | 0.021 | 0.039  | 0.065  |
| PHX  | 0.16  | 0.26   | 0.26   |
| CMUP | 0.016 | 0.010  | 0.014  |
| CMX  | 0.018 | 0.014  | 0.014  |

Table 4.11: Predicted non-$W$ fractions for the different lepton categories broken up into 3 jets. The relative uncertainty on these fractions is 25% for CEM, CMUP and CMX. For PHX we use an uncertainty of 30%. The given non-$W$ fractions include the cut on the $\Delta\phi$ between the axis of leading jet in the transverse plane and the $\not{E}_T$ vector (QCD cut).

systematic uncertainty for CEM, CMUP and CMX is taken to be 25%, the same number as used in $t\bar{t}$ analyses (e.g. reference [52]). We tried to use the 4-sector method also for the PHX events taken with the MET_PEM_v trigger. But it turns out that the standard method is not applicable here because there are too few events predicted. One reason could be a potential trigger bias, since the $\not{E}_T$ distribution is distorted by the trigger. That is why we use events taken with the `PLUG_ELECTRON_20` trigger. The resulting numbers are given in table 4.11. We assign a systematic uncertainty of 30% to these pretag non-$W$ fractions for the PHX samples.

To predict the number of non-$W$ events in the $b$ tagged sample we multiply the

|           | 1 jet | 2 jets | 3 jets |
|-----------|-------|--------|--------|
| Electrons | 0.026 | 0.029  | 0.029  |
| Muons     | 0.030 | 0.045  | 0.029  |

Table 4.12: $b$ tag fractions.

|       | 1 jet           | 2 jets          | 3 jets        |
|-------|-----------------|-----------------|---------------|
| CEM   | $20.0 \pm 5.0$  | $7.2 \pm 1.8$   | $2.0 \pm 0.5$ |
| PHX   | $88.1 \pm 26.4$ | $110 \pm 38$    | $6.9 \pm 2.1$ |
| CMUP  | $8.5 \pm 2.1$   | $1.3 \pm 0.3$   | $0.3 \pm 0.1$ |
| CMX   | $5.6 \pm 1.4$   | $1.0 \pm 0.3$   | $0.2 \pm 0.1$ |
| Total | $122.1 \pm 34.9$ | $119.5 \pm 40.4$ | $9.3 \pm 2.7$ |

Table 4.13: Number of predicted non-$W$ events with $b$ tag.

number of pretag events by the tagging rates $\epsilon_{\text{tag}}$, as quoted in table 4.12, and the non-$W$ fraction in table 4.11. The resulting numbers of $b$ tagged non-$W$ events are listed in table 4.13.

Since the background is still underestimated in the PHX $W + 2$ jets sample after $b$ tagging, the described standard method to estimate the non-$W$ fraction does not work for the PHX sample. Therefore, we fit the distribution of $\Delta\phi$ between the second leading jet and the $\not{E}_T$ vector to determine the number of non-$W$ events we expect in the PHX region. The distribution can be seen in figure 4.2. This distribution has a significant peak at low values where the $\not{E}_T$ points in direction of the jet. To model the shape of non-$W$ events which we need for the fit of the distribution we use a sample of generic jet events, where one jet with a high fraction of energy deposited in the electromagnetic part of the calorimeter is used to mock an electron candidate. Using the fit result we find $110 \pm 38$ non-$W$ events in the PHX sample which is also listed in table 4.13. This corresponds to a non-$W$ fraction of about 50%.

## 4.2.2   Mistags

One substantial background arises from falsely reconstructed secondary vertices which lead to a false $b$ tag for $W +$ light flavor jets events. We estimate the number of mistagged events by applying the mistag matrix provided by the CDF high-$p_{\text{T}}$ $b$ tagging group [53]. This matrix gives the probability to mistag a light jet as a heavy jet. The result of the mistag estimate is shown in table 4.14. The uncertainties of the prediction are calculated taking the correlations between jet bins into account. In addition to the statistical uncertainties quoted in table 4.14 there is an additional systematic uncertainty of 8% on the number of predicted negative tags. $\alpha \cdot \beta$ is a scale factor.
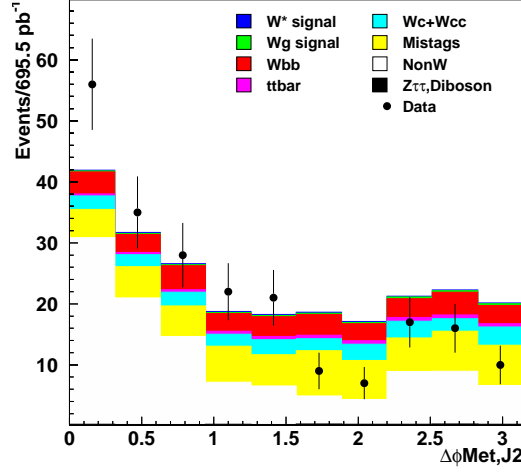
Figure 4.2: Distribution of the difference in azimuth, $\Delta\phi$, between the 2nd leading jet and the $\not{E}_T$ Vector. This distribution is used to extract the non-$W$ fraction in the PHX sample.

### 4.2.3   Heavy Flavor

The heavy flavor backgrounds are estimated by applying heavy flavor fractions $f_{HF}$ measured in generic jet data to the pretag $W$+jets samples. The numbers can be found in reference [54]. The number of $W$ + heavy flavor events $N_{HF}$ is estimated according to

$$N_{HF} = N_{pretag} \cdot (1 - f_{nonW}) \cdot f_{HF} \cdot \epsilon_{tag} \ . \tag{4.11}$$

The number of events in the pretag sample $N_{pretag}$ is given in table 4.15. Using equation 4.11 we get the result for the expected heavy flavor backgrounds as given in table 4.15.

### 4.2.4   Event Yield Summary

If we add up all the numbers of expected events for the different processes we estimated in the sections before, we get our total event yield. It is shown for the three jet bins and divided into the different processes in table 4.16. The last line shows the number of events we find in data. Our prediction matches the data in all three jet bins well. Table 4.17 shows the event yield observed in data split up into lepton categories as well as our prediction. Since our analysis focuses on the $W + 2$ jets sample the column 2 `jets` is of most interest. For our analysis we expect $674.1 \pm 96.1$ events and we find 689 events in data. This is a good agreement between prediction and observation.

| | 1 jet | 2 jets | 3 jets |
|---|---|---|---|
| | Predicted negative tags | | |
| CEM | $115.3 \pm 6.9$ | $53.5 \pm 1.4$ | $19.4 \pm 0.3$ |
| PHX | $73.4 \pm 3.0$ | $42.5 \pm 0.8$ | $15.81 \pm 0.3$ |
| CMUP | $52.5 \pm 3.0$ | $23.3 \pm 0.8$ | $7.3 \pm 0.1$ |
| CMX | $30.8 \pm 2.1$ | $14.2 \pm 0.5$ | $3.9 \pm 0.1$ |
| Total | $272.0 \pm 10.9$ | $133.5 \pm 2.2$ | $46.4 \pm 0.5$ |
| | Observed negative tags | | |
| CEM | 122 | 53 | 12 |
| PHX | 62 | 40 | 16 |
| CMUP | 48 | 20 | 5 |
| CMX | 39 | 14 | 1 |
| Total | 271 | 127 | 34 |
| | Correction factor $\alpha \cdot \beta = 1.37 \pm 0.15$ | | |
| | Predicted positive mistags | | |
| Total | $351.3 \pm 52.5$ | $164.3 \pm 29.6$ | $56.0 \pm 7.6$ |

Table 4.14: Number of predicted and observed negative tagged events broken up into the different lepton categories. The numbers given in this table are before subtracting the non-$W$ fraction. The number of predicted positive mistags, however, does include the correction for the non-$W$ fraction.

| | 1 jet | 2 jets | 3 jets |
|---|---|---|---|
| $N_{\mathrm{pretag}}$ | $81080 \pm 1550$ | $13934 \pm 550$ | $2394 \pm 146$ |
| $Wb\bar{b}$ | $270.0 \pm 81.2$ | $170.7 \pm 49.2$ | $48.5 \pm 12.5$ |
| $Wc\bar{c}$ | $94.7 \pm 23.7$ | $64.5 \pm 17.3$ | $20.7 \pm 5.6$ |
| $Wc$ | $268.5 \pm 56.4$ | $69.4 \pm 15.3$ | $12.2 \pm 2.6$ |
| Total HF | $633.2 \pm 161.3$ | $304.6 \pm 81.9$ | $81.4 \pm 20.7$ |

Table 4.15: Number of predicted heavy flavor events.

|  | 1 jet | 2 jets | 3 jets |
|---|---|---|---|
| $W b \bar{b}$ | $270.0 \pm 81.2$ | $170.7 \pm 49.2$ | $48.5 \pm 12.5$ |
| $W c \bar{c}$ | $94.7 \pm 23.7$ | $64.5 \pm 17.3$ | $20.7 \pm 5.6$ |
| $W c$ | $268.5 \pm 56.4$ | $69.4 \pm 15.3$ | $12.2 \pm 2.6$ |
| Total HF | $633.2 \pm 161.3$ | $304.6 \pm 81.9$ | $81.4 \pm 20.7$ |
| Non-$W$ | $122.1 \pm 34.9$ | $119.5 \pm 40.4$ | $9.3 \pm 2.7$ |
| Mistags | $351.3 \pm 52.5$ | $164.3 \pm 29.6$ | $56.0 \pm 7.6$ |
| Total data based background | $1106.6 \pm 169.7$ | $588.4 \pm 96.0$ | $146.7 \pm 22.0$ |
| $t \bar{t}$ | $4.85 \pm 0.51$ | $40.34 \pm 3.45$ | $91.66 \pm 7.37$ |
| $WW$ | $1.68 \pm 0.18$ | $3.81 \pm 0.40$ | $1.10 \pm 0.12$ |
| $WZ$ | $2.77 \pm 0.29$ | $6.09 \pm 0.55$ | $1.50 \pm 0.13$ |
| $ZZ$ | $0.05 \pm 0.01$ | $0.21 \pm 0.02$ | $0.10 \pm 0.01$ |
| $Z \to \tau\tau$ | $2.96 \pm 0.32$ | $2.59 \pm 0.27$ | $1.68 \pm 0.18$ |
| $Z \to \mu\mu$ | $9.84 \pm 1.09$ | $4.44 \pm 0.48$ | $0.53 \pm 0.06$ |
| Total MC bkg. | $22.2 \pm 1.3$ | $57.5 \pm 3.6$ | $96.6 \pm 7.4$ |
| $t$-channel | $7.36 \pm 0.77$ | $16.71 \pm 1.72$ | $4.14 \pm 0.37$ |
| $s$-channel | $3.40 \pm 0.36$ | $11.52 \pm 0.91$ | $2.41 \pm 0.19$ |
| Total Single-Top | $10.8 \pm 1.1$ | $28.2 \pm 2.6$ | $6.5 \pm 0.6$ |
| Total Prediction | $1139.4 \pm 161.3$ | $674.1 \pm 96.1$ | $249.8 \pm 23.2$ |
| Observation | 1141 | 689 | 280 |

Table 4.16: Expected and observed event yield split up into different processes.

| Predicted Number of Events | | | |
|---|---|---|---|
|  | 1 jet | 2 jets | 3 jets |
| CEM | $465.1 \pm 99.1$ | $250.7 \pm 47.0$ | $112.7 \pm 13.1$ |
| PHX | $315.4 \pm 74.7$ | $236.8 \pm 52.4$ | $57.6 \pm 10.7$ |
| CMUP | $226.4 \pm 47.2$ | $120.6 \pm 21.9$ | $53.5 \pm 5.8$ |
| CMX | $132.5 \pm 27.8$ | $65.8 \pm 12.6$ | $26.0 \pm 3.1$ |
| Total | $1139.4 \pm 161.3$ | $674.1 \pm 96.1$ | $249.8 \pm 23.2$ |
| Observed Number of Events | | | |
|  | 1 jet | 2 jets | 3 jets |
| CEM | 440 | 259 | 127 |
| PHX | 329 | 221 | 85 |
| CMUP | 222 | 137 | 47 |
| CMX | 150 | 72 | 21 |
| Total | 1141 | 689 | 280 |

Table 4.17: Predicted and observed event yield split up into lepton categories.

# Chapter 5

# Neural Network Discriminants

In this chapter, the used technique to distinguish between signal and background is introduced. We use neural networks to combine several discriminants into one more powerful discriminant. First the NeuroBayes® package [55], the neural network we use, is introduced. After that, the training of the networks for our analysis is explained and finally the creation of the templates out of the neural network outputs is outlined in detail.

## 5.1 NeuroBayes

The NeuroBayes® package used in this analysis is provided by the company Phi-T. It combines a three-layer feed forward neural network as seen in figure 5.1 (a) with a complex robust preprocessing of the input variables. Bayesian regularization techniques are used for the training process. There is one input node for each input variable plus one bias node. The number of nodes in the hidden layer can be freely chosen by the user. There is one output node which gives a continuous output in the interval [-1,1].

The nodes of two consecutive layers are connected with variable weights. For each node $j$, a biased weighted sum of the values of the previous layer $x_i$ is calculated

$$a_j(\mathbf{x}) = \sum_i \omega_i x_i + \mu_{0,j} \qquad (5.1)$$

and passed to the transfer function which gives the output of the node. The bias $\mu_{0,j}$, which is calculated for each input, implements the thresholds of the several nodes: if the input to a node is larger than its threshold, the node will send an input to the next layer. The output of each node is determined by a transformed sigmoid function

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

which gives an output of -1 for background and +1 for signal.

As can be seen in figure 5.1 (b), the sigmoid function is only sensitive to a relatively small range about 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 mentioned above shifts the mean of the input data distribution to the linear part of the sigmoid function.
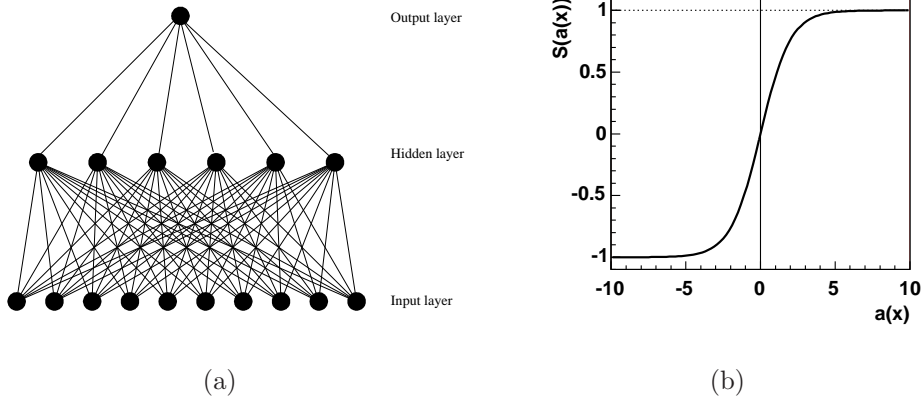


(a)            (b)

Figure 5.1: (a) The geometry of a 3 layer neural network and (b) the sigmoid function given by formula (5.2).

The output of the neural network for the output node is calculated by

$$o = S(\sum_{j=0}^{M} \omega_j^{(2)} \cdot S(\sum_{i=0}^{d} \omega_{ji}^{(1)} x_i + \mu_{0,j})) \tag{5.3}$$

where $d$ is the number of input and $M$ the number of hidden nodes. $\omega_{ji}^{(1)}$ denotes the weights from the input to the hidden layer, $\omega_j^{(2)}$ the weights from the hidden layer to the output node. $\mu_{0,j}$ is the weight that connects the bias node with the hidden nodes.

## 5.1.1 The Training Process

The training of the neural network is done by minimizing the deviation between the true output and the one calculated by using the actual weights. The error function minimized in this neural network is the entropy error function, which is essentially given by the sum of the logarithms of the output values. The aim of the training of the neural network is to find the minimum in the multidimensional structure of the error function with local maxima and minima. As this task can be difficult to solve, the training process is done by the combined method of gradient descent and backpropagation. The neural network is trained with regularization techniques to improve generalization performance and to avoid overtraining. During the training process, weights (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. This leads again to an improved generalization ability. Figure 5.2 (b) displays a control plot of the training

process. It shows an error in arbitrary units that is minimized during the training process. In this case the error is minimized after only few training iterations. For more information on the the above mentioned features of NeuroBayes® see [55].
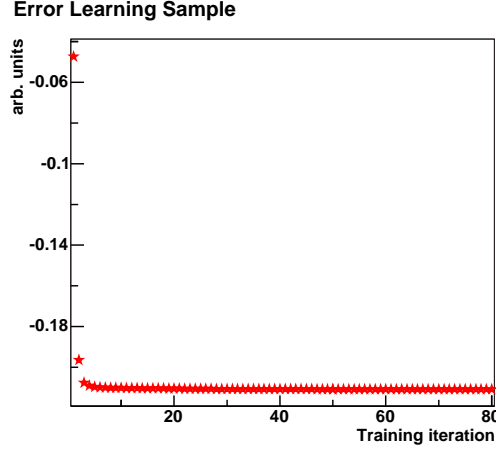


Figure 5.2: Control plot for the neural network training: an error in arbitrary units that needs to be minimized against the number of training iterations.

## 5.1.2  Preprocessing of the Variables

To find the optimal starting point for minimizing the error function, the input variables are preprocessed. This preprocessing is done in a completely automatic way. To reduce the influence of extreme outliers, the input distributions are equalized to lie between -1 and 1. Those flattened distributions are then converted into a Gaussian distribution, centered at zero with standard deviation 1. This avoids saturation of the nodes due to the above mentioned shape of the activation function (see figure 5.1 (b)). After this transformation, the input variables are linearly decorrelated by diagonalization and rotation of the covariance matrix into a unit matrix.

The above mentioned transformation to a Gaussian distribution may be altered by an individual variable preprocessing like fitting the flattened distribution with splines if this is considered to be sensible. In addition, discrete variables can be treated as members of classes. The preprocessing of those kinds of variables can also deal with a certain order of values if this is important, e.g. the number of jets in an event. The preprocessing is also able to deal with variables that are not given for every event by assigning the missing values to a $\delta$-function.

Of course, those preprocessed input variables do not have any physical meaning anymore, unlike the original ones.

## 5.1.3  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 the 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. At the end of the preprocessing, we get a list of variables ordered according to their correlation to the target.

The significance of each variable is calculated by dividing the loss of correlation induced by its removal at the relevant point in the procedure described above 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, it is possible to cut on the significance of the variables. Only those variables are taken into account that include relevant information that is not already incorporated by other variables.

## 5.2 Training of the Neural Networks

### 5.2.1 Training Samples

Since we want to measure a combined single-top cross section $\sigma_{t+s}$ as well as separate cross sections for the $t$-channel $\sigma_t$ and $s$-channel $\sigma_s$, we need three different trained neural networks in this analysis. Consequently, we train separate neural networks for $t$-channel, $s$-channel and $t$- and $s$-channel combined for the 2 jet bin. Therefore, three separate training samples are needed. Each used training sample includes roughly 17000 events. The samples contain a ratio of signal to background of 35% to 65%. In the training of an $s$-channel network, the $t$-channel events are treated as background and vice versa, although this is negligible in our case. Due to the fact that there is no non-$W$ Monte Carlo we add the non-$W$ fraction to 80% to the $Wb\bar{b}$ fraction and to 20% to the $Wc\bar{c}$ fraction. This estimation is based on a fit performed to the output of the neural network $b$ tagger [56] in a control sample enriched in $b$ tagged QCD multi-jet events. The configurations of the training samples are shown in table 5.1. The distribution of the 65% background to the different background processes is the same for all three networks. The difference is that in the training sample for the $t$-channel 1% of the 17000 events in the sample are $s$-channel events which are treated as background, whereas in the $t+s$-channel training sample 14% of the 17000 events are $s$-channel events, 15% are $t$-channel ($2 \rightarrow 2$) and 6% ($2 \rightarrow 3$) events which are all treated as signal. The same can be seen for the $s$-channel training sample where 34% are $s$-channel events which are marked as signal in the training process and 1% are $t$-channel events which are treated as background.

### 5.2.2 Input Variables for the Neural Networks

A neural network is a tool to combine information from different sources concerning one target into one powerful discriminant considering correlations between the dif-

|  | $t$-channel | $s$-channel | $t$- and $s$-channel |
|---|---|---|---|
| $t$-channel 2— >2 | 23.0% | 0.5% | 15.0% |
| $t$-channel 2— >3 | 11.0% | 0.5% | 6.0% |
| $s$-channel | 1.0% | 34.0% | 14.0% |
| $t\bar{t}$ | 4.6% | 4.6% | 4.6% |
| $WW$ | 0.4% | 0.4% | 0.4% |
| $WZ$ | 0.7% | 0.7% | 0.7% |
| $ZZ$ | 0.0% | 0.0% | 0.0% |
| $Z—>\tau\tau$ | 0.3% | 0.3% | 0.3% |
| $Wb\bar{b}$ 0 parton (e,$\nu$) | 11.8% | 11.8% | 11.8% |
| $Wb\bar{b}$ 0 parton ($\mu$,$\nu$) | 11.8% | 11.8% | 11.8% |
| $Wc\bar{c}$ 0 parton (e,$\nu$) | 4.2% | 4.2% | 4.2% |
| $Wc\bar{c}$ 0 parton ($\mu$,$\nu$) | 4.2% | 4.2% | 4.2% |
| $Wc$ 1 parton (e,$\nu$) | 4.4% | 4.4% | 4.4% |
| $Wc$ 1 parton ($\mu$,$\nu$) | 3.6% | 3.6% | 3.6% |
| mistags | 18.9% | 18.9% | 18.9% |

Table 5.1: Configuration of the training samples for the 2 jet bin (35% signal, 65% background).

ferent sources. In our case the input for the neural network are variables measured in the detector or reconstructed out of the measured values. A perfect variable would have a completely different shape for signal events, $t$-channel and $s$-channel, than for background events. Unfortunately, there is no single "golden" variable and most variables have only a small discrimination power.

As mentioned in section 5.1.2, NeuroBayes$^{\circledR}$ is able to select the most relevant variables automatically. We studied approximately 40 variables and pruned away all which had a significance of less than three sigma. After that all three networks have 14 input variables. Table 5.2 shows the 14 variables ranked according to the output of NeuroBayes$^{\circledR}$. The significance $\sigma$ is also shown in the list. It is possible that a variable has a better rankt than another variable even though its significance is smaller. This can happen because of the way in which the two values are calculated by NeuroBayes$^{\circledR}$ as explained in section 5.1.3.

The input variables are measured or derived values for every event. Since we focus on the so called $W+2$ jets channel, each event we select has exactly one tight lepton and missing transverse energy both originating from the $W$ decay as well as two tight jets. The input variables can be reconstructed from these values, e.g. the invariant mass of the lepton, the missing transverse enregy and the $b$ jet, $M_{l\nu b}$, or they are already directly measured in the detector, e.g. the transverse energy of the leading jet $E_T(j_1)$. The third category of input variables are values which are calculated by algorithms out of the measured values, e.g. the neural network $b$-tagger ANN B-Jets1 [56]. In the following the 14 input variables are described and a Monte Carlo study for four different processes is done. The $t$-channel is shown in black, the $s$-channel in red, the $t\bar{t}$ events in yellow and the $Wb\bar{b}$ in dashed blue. All the

| Rank | Variable | correlation | significance $[\sigma]$ |
|------|----------|-------------|------------------------|
| 1. | ANN B-Jets1 | 33.39 | 44.52 |
| 2. | $M_{l\nu b}$ | 27.20 | 36.27 |
| 3. | $Q \cdot \eta$ | 18.17 | 24.23 |
| 4. | $E_T(j_2)$ | 16.11 | 21.48 |
| 5. | $cos(\Theta_{l,q})$ | 12.25 | 16.33 |
| 6. | $E_T(j_1)$ | 8.51 | 11.35 |
| 7. | $\log_{10}(\Delta_{34})$ | 6.97 | 9.30 |
| 8. | $\sum_i \eta(j_i)$ | 6.35 | 8.47 |
| 9. | $P_T(lep)$ | 6.53 | 8.70 |
| 10. | $\eta_W$ | 5.48 | 7.31 |
| 11. | $\chi_3^2$ | 4.31 | 5.75 |
| 12. | $M_{j_1,j_2}$ | 3.76 | 5.02 |
| 13. | $\sum \eta(j_T, j_l)$ | 2.49 | 3.32 |
| 14. | $\Delta\chi^2$ | 2.49 | 3.33 |

Table 5.2: Input Variables of the $t+s$-channel neural network ordered according to their correlation to the target. The significance of the variable is also shown.

curves are normalized for comparison reasons. Those four processes are chosen since we want to see how the two signal processes look like for the input variables, the $t\bar{t}$ background is the only background with a real top in it and the $Wb\bar{b}$ background resembles all the other backgrounds not including a real top. This Monte Carlo study should display how different the input variables look like for the different processes.

- The first two input variables described are the so called dijet mass $M_{j_1,j_2}$ and the reconstructed top mass $M_{l\nu b}$. Both variables are reconstructed masses which are calculated using the measured values from the detector. The dijet mass is the invariant mass of the two jets in the event. The two Lorentz vectors of the jets are added vectorially and the mass of the resulting vector is the invariant dijet mass of the event, figure 5.3 (a). The two jets in the $Wb\bar{b}$ sample originate from gluon splitting. Therefore, the sum of their Lorentz vectors gives a lower value for the mass than in the cases of the other three processes where the jets do not originate from the same object. This explains why the curve of the $Wb\bar{b}$ sample peaks for lower values. The reconstructed mass of the top quark is calculated using the fact that the top decays in nearly 100% of the cases into a $W$ boson and a $b$ jet. Therefore, the reconstructed top mass is the sum of the reconstructed $W$ boson and the identified $b$ jet. If there is more than one $b$ jet in the event, the jet with the larger value of $Q \cdot \eta$ is chosen. The $W$ boson is reconstructed from the measured tight lepton and the missing transverse energy in the event. Since only the missing transverse energy is known, there is an ambiguity for the $p_z$ of the neutrino. We choose the solution with the smaller value for $|p_z|$. The reconstruction also explains

the different shapes of the three different processes that include real tops. First of all, the $Wb\bar{b}$ background peaks at lower values since there is no top in the events which can be seen in figure 5.3 (b). The other three processes do all peak around the top mass since they all include real tops. The fact that the three processes still have different shapes can be explained if we take a look at the reconstruction. For $t$-channel events there should only be one $b$ tagged jet and therefore the only ambiguity arises from the $W$ boson reconstruction. In case of the $s$-channel it is very likely that the wrong jet is assigned as the $b$ jet, since there are two $b$ jets produced. In the case that there are two $b$ tagged jets an additional ambiguity occurs. Only in about 55% of the cases the assignment of the $b$ jet is correct. Therefore, the peak of the $s$-channel is not as sharp as the peak of the $t$-channel. The $t\bar{t}$ peak is even less sharp than the $s$-channel peak since there is one real top and one real antitop in the event which both produce a $b$ jet once they decay. This results in even more combinatorial options.
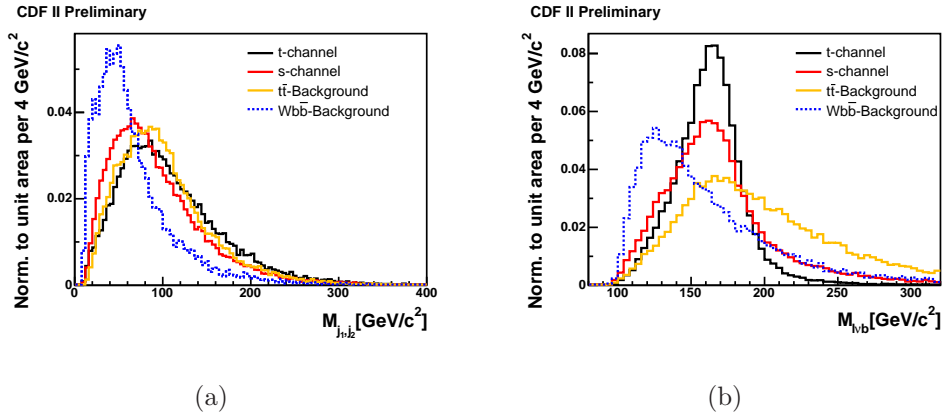


(a)                  (b)

Figure 5.3: (a) The invariant mass of the two leading jets and (b) the reconstructed top mass.

- The $Q \cdot \eta$ variable is used as an input variable for the neural network since it is a very strong discriminant for the $t$-channel. This can be seen in figure 5.4 (a) where the $t$-channel peaks in the forward direction, whereas the other distributions peak in the central region. The variable $Q \cdot \eta$ is defined as the product of the charge of the lepton and the pseudorapidity $\eta$ of the light quark jet (the jet which is not assigned as the $b$ jet from the top decay). The asymmetry of the $t$-channel is due to the parton distribution function of the proton. A proton consists of two up quarks, one down quark, gluons and seaquarks. Since in the $t$-channel production mode the initial state partons are a valence quark of the proton and a $b$ quark from gluon splitting, the top (antitop) quark and the light quark (antiquark) are produced propagating in the proton direction. There are two times more top quarks than antitop quarks produced if the initial state parton comes from the proton. The lepton in an antitop event

has always a negative charge. The multiplication of the pseudorapidity of the light quark with the charge of the lepton leads to the characteristic peak in the positive regime of the $Q \cdot \eta$ distribution for $t$-channel events. The argument is made for the protons, but it is the same for antiprotons with the difference that the quarks (antiquarks) propagate in antiproton direction and two times more antiprotons are produced. The charge of the lepton switches the peak from the negative to the positive range.

The transverse momentum of the lepton in the event is another input variable. Its Monte Carlo study can be seen in figure 5.4 (b). The lepton cut at 20 GeV/$c$ can be seen easily.



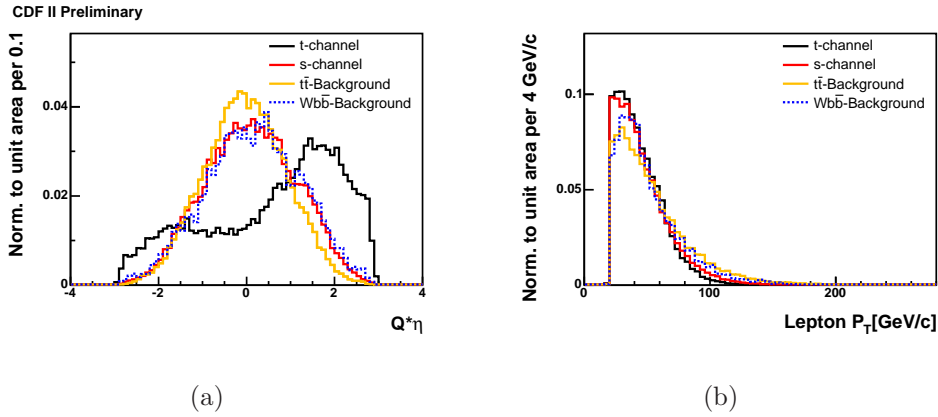(a)                                              (b)

Figure 5.4: (a) The charge of the lepton times pseudorapidity of the leading light jet and (b) the transverse momentum of the lepton.

- The logarithm of the $\Delta_{34}$ value which is calculated by the $K_T$ jet clustering algorithm [57]. $\Delta_{34}$ is the cut off parameter at which the algorithm merges four jets of the event into three jets. It is a good variable to discriminate $t\bar{t}$ events. The logarithm of the cut off parameter, $\log_{10}(\Delta_{34})$, can be seen in figure 5.5 (a). The neural network $b$ tagger output is a very strong variable to discriminate backgrounds which do not include a real $b$. The neural network b-tagger is described in more detail in [56]. Since all the four Monte Carlo samples we look at include a real $b$ jet and therefore peak at high values we also show the distribution for mistags in green (dashed). It peaks at low values and thus is easily separated from the other processes which is shown in figure 5.5 (b).

- The input variable $cos(\Theta_{l,q})$ is again a very strong variable to separate the t-channel from the other processes as shown in figure 5.6 (a). This is because the top quarks are produced 100 % polarized along the direction of the $d$-type quark in the top rest frame in case of the $t$-channel as well as the $s$-channel. This effect is due to the P violation in the electroweak force. Because the $W$ boson only couples to left handed fermions and because of the short lifetime of
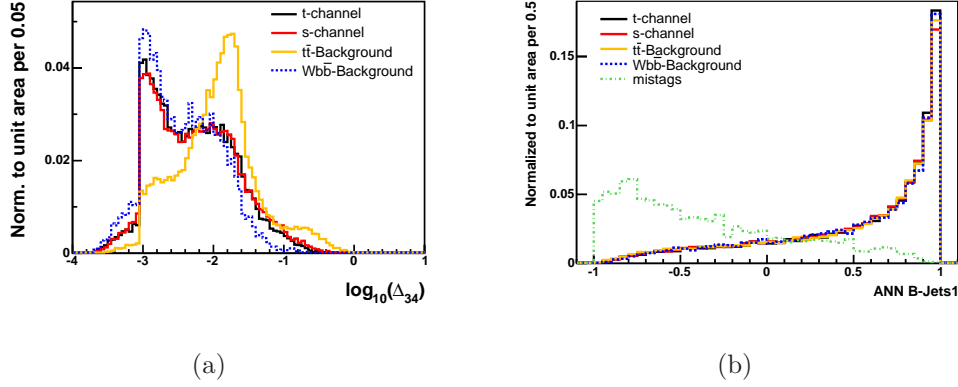
(a)                                                    (b)

Figure 5.5: (a) The logarithm of $\Delta_{34}$, calculated by the $K_T$ jet clustering algorithm and (b) the neural network output of the $b$ tagger for the first b tagged jet.

the top quark (no hadronization) the properties of the top quark are transfered to its decay products. This polarization can experimentally be seen through the angle $\Theta_{q,l}$. This is the angle between the lepton from the top decay and the light quark jet axis in the top rest frame. In case of the $s$-channel, the light quark jet axis is the beam axis. We also investigated the variable $cos(\Theta_{l,beam})$ but this variables does not have as much power to distinguish between signal and background as $cos(\Theta_{q,l})$.

The pseudorapidity of the $W$ boson is another input variable. The $W$ boson is reconstructed the same way as described above. The difference of the shapes of the different processes can be seen in figure 5.6 (b). The shape of the $t\bar{t}$ sample is different because the top and the antitop are produced more central and therefore the $W$ boson is more central for this process as well. The shapes of the two signal samples look very much alike and the shape of the $Wb\bar{b}$ sample is much broader since it contains a real $W$ boson not originating from the top quark decay.

- The transverse energies of the two jets, $E_T(j_1)$ and $E_T(j_2)$, are measured in the hadronic calorimeters. They are both direct measurements from the detector. Their discriminants can be seen in figure 5.7. The different shapes can be explained in a similar way as for the shapes of the reconstructed dijet mass. The $Wb\bar{b}$ sample peaks at low values since both jets originate from the same gluon. Those jets lie in low energy ranges. For the $t\bar{t}$ sample on the other hand the possibility to have a high energetic jet is significantly higher and therefore the shape is broader.

- The sum of the pseudorapidities of the two jets $\sum_i \eta(j_i)$ is an input variable that is calculated from the two reconstructed jets of the event. The relevant Monte Carlo distributions can be seen in figure 5.8 (a). The shape of the $t$-channel sample can be explained through the fact that the light quark is

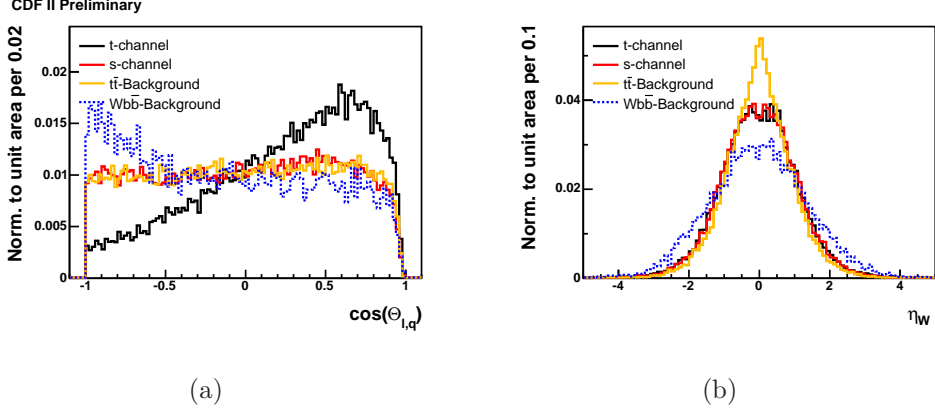(a)                                            (b)

Figure 5.6: (a) The cosine of the angle between the lepton and the light jet in the top rest frame and (b) the pseudorapidity of the $W$ boson.

produced in forward direction. Therefore, the distribution of the pseudorapidity is wider. The sum of the pseudorapidity of the reconstructed top vector and the pseudorapidity of the light quark jet $\sum \eta(j_T, j_l)$ on the other hand is again influenced by the the fact that the top and antitop quarks are produced more central for the $t\bar{t}$ events. Figure 5.8 (b) shows the shape differences.

- The last two input variables are based on a Kinematic Fitter developed at the University of Illinois [58]. This fitter uses kinematic variables to reconstruct the objects in the event through a fit routine. The fits do all have $\chi^2$ values to give a measure how well the fit matches the data. Since there are two possible assignments for the $b$ jet and two possible assignments for $p_z$, there are four different fits and we get four $\chi^2$ values. The $\chi^2$ is calculated according to equation 5.4.

$$\chi^2 = \frac{(P_b - P_b^{\text{obs}})^2}{\sigma_{P_b}^2} + \frac{(\eta_b - \eta_b^{\text{obs}})^2}{\sigma_{\eta_b}^2} + \frac{(\Phi_b - \Phi_b^{\text{obs}})^2}{\sigma_{\Phi_b}^2} + \tag{5.4}$$
$$\frac{(\not{E}_T - \not{E}_T{}^{\text{obs}})^2}{\sigma_{\not{E}_T}^2} + \frac{(\not{\Phi}_\nu - \not{\Phi}_\nu{}^{\text{obs}})^2}{\sigma_{\not{\Phi}_\nu}^2} + \frac{(M_{l\nu b} - M_t^{\text{obs}})^2}{\sigma_{M_t}^2} + Y(Im(P_z))^2$$

The variables defining the $\chi^2$ can all be calculated from the Lorentz vector of the $b$ jet and the missing transverse energy in the event. $P_b$ is the energy, $\eta_b$ the pseudorapidity, and $\Phi$ the angle in the x-y plane of the $b$ jet. $\not{E}_T$ is the missing transverse energy in the event, $\not{\Phi}$ is the angle in the x-y plane of the missing energy, and $M_{l\nu b}$ is the reconstructed top mass. We set the top mass $M_t$ to 175 GeV/$c^2$. A penalty term proportional to the square imaginary part of the neutrino's $p_z$ solution with $Y = 5$ pushes the kinematic fit away from unphysical neutrino solutions.
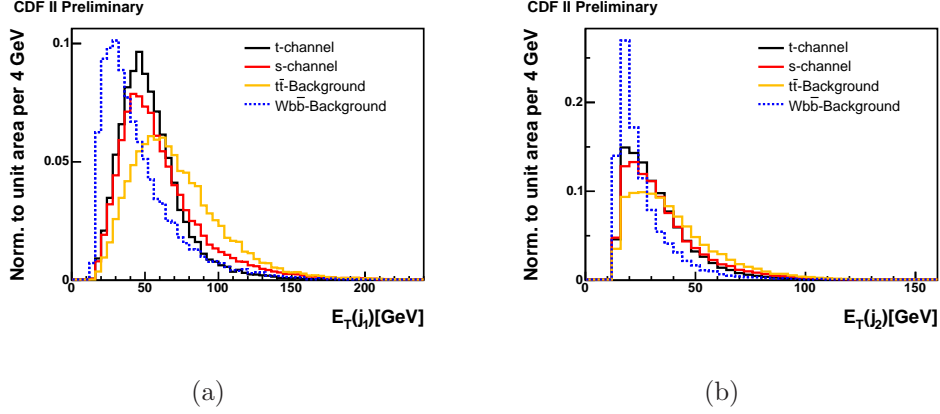
(a)                                                             (b)

Figure 5.7: (a) The transverse energy of the leading jet $E_T(j_1)$ and (b) the transverse energy of the second jet $E_T(j_2)$.



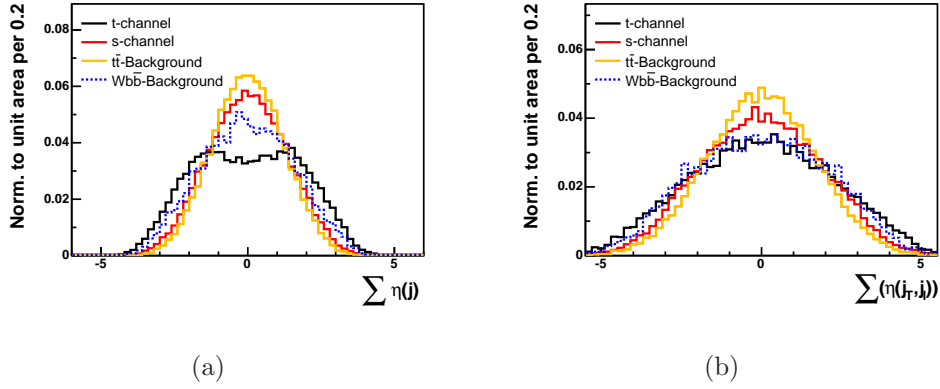(a)                                                             (b)

Figure 5.8: (a) The sum of the pseudorapidities of the jets in the event and (b) the sum of the pseudorapidity of the reconstructed top quark and the light jet.

We examined all four $\chi^2$ plus some combinations of the four values. Based on the variable ranking given by NeuroBayes$^\circledR$ we decided to take the third $\chi^2$ which can be seen in figure 5.9 (a) as well as the logarithm of the difference between the first and the second $\chi^2$ calculated by the Kinematic Fitter $\Delta(\chi^2_{1,2})$ shown in figure 5.9 (b).

The neural network package provides a feature that shows the training success of the neural network. In figure 5.10 (a), the neural network output for signal and background can be found. This graphic is just for illustration purpose, since it depends on the ratio of signal to background in the training sample. The graphic just proves that the neural network is able to distinguish between background and signal events. In figure 5.10 (b), the purity in each bin versus the neural network output is shown. For a well trained network the points should lie on the diagonal.
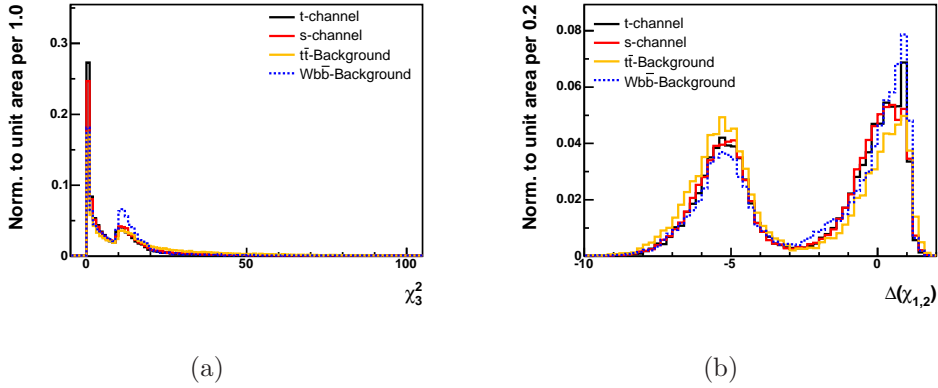
(a)                                        (b)

Figure 5.9: (a) The $\chi^2$ of the Kinematic Fitter for top reconstruction and (b) the difference between the first and the second $\chi^2$ of the Kinematic Fitter for top reconstruction.

The presented graphics are examples taken from the training of the neural network for the $t$-channel 2 jet bin. The other networks have different training outputs but are generally comparable and are for that reason not shown here.
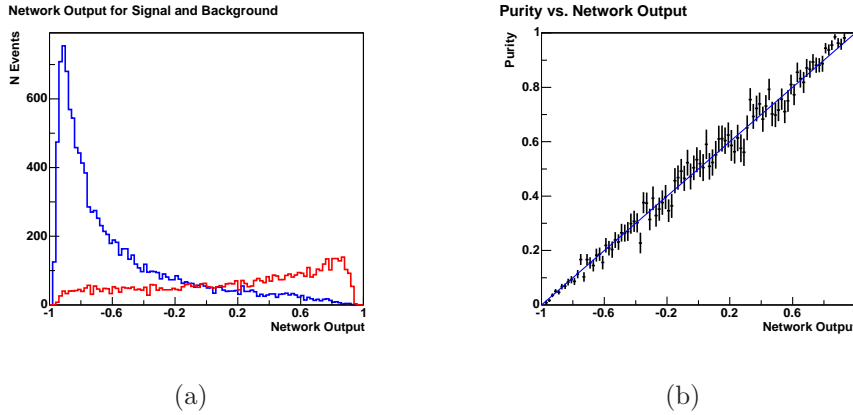


(a)                                        (b)

Figure 5.10: (a) Neural network output for signal and background of the $t$-channel network and (b) purity versus neural network output for the $t$-channel network.

## 5.2.3   Validation of Input Variables

To make sure that the training is done in the right way it is not only important to check that the purity versus neural network output plot in figure 5.10 (b) looks good and that the error shown in figure 5.2 reached its minimum but it is also important that the Monte Carlo matches the data. Therefore, we need to do data Monte Carlo comparisons for the input variables. For some of the input variables the comparison is shown in figure 5.11 and 5.12. The calculation of the variables is done in the same

way as described in section 5.2.2. The Monte Carlo sample is put together according to our event prediction in table 4.16 from the Monte Carlo samples listed in table 3.1. In the plots the data is shown with black dots, the Monte Carlo estimate is shown in red and the yellow band is the statistical error on the Monte Carlo prediction, i.e. only statistical uncertainties are included in these plots, no systematic uncertainties. The statistical uncertainty is the square-root of the number of entries in a certain bin.

The data Monte Carlo comparison shows good agreement within the uncertainties for the input variables. Therefore it is sensible to assume that the presumptions we made are correct.
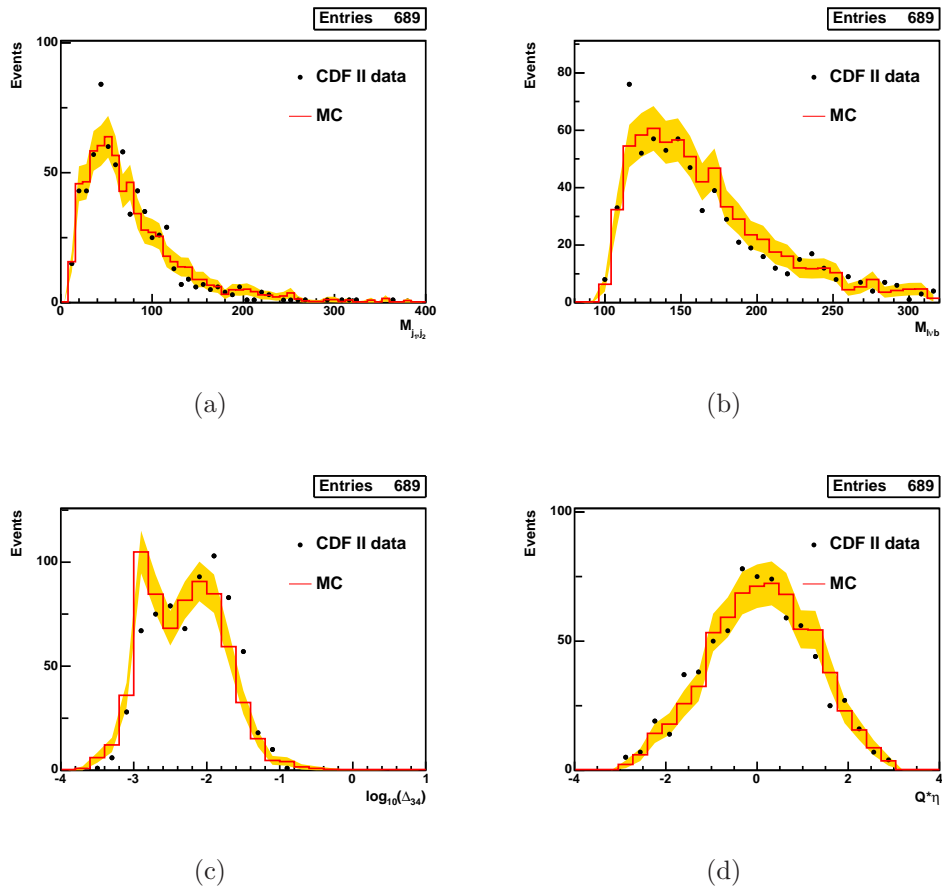


Figure 5.11: Some of the input variables for the neural network, comparing the data and the prediction: (a) the invariant mass of the two leading jets, (b) the reconstructed top mass, (c) the logarithm of $\Delta_{34}$, calculated by the $K_T$ jet clustering algorithm [57], and (d) the lepton charge times pseudorapidity of the light jet. The Monte Carlo is normalized to the data.

(a)

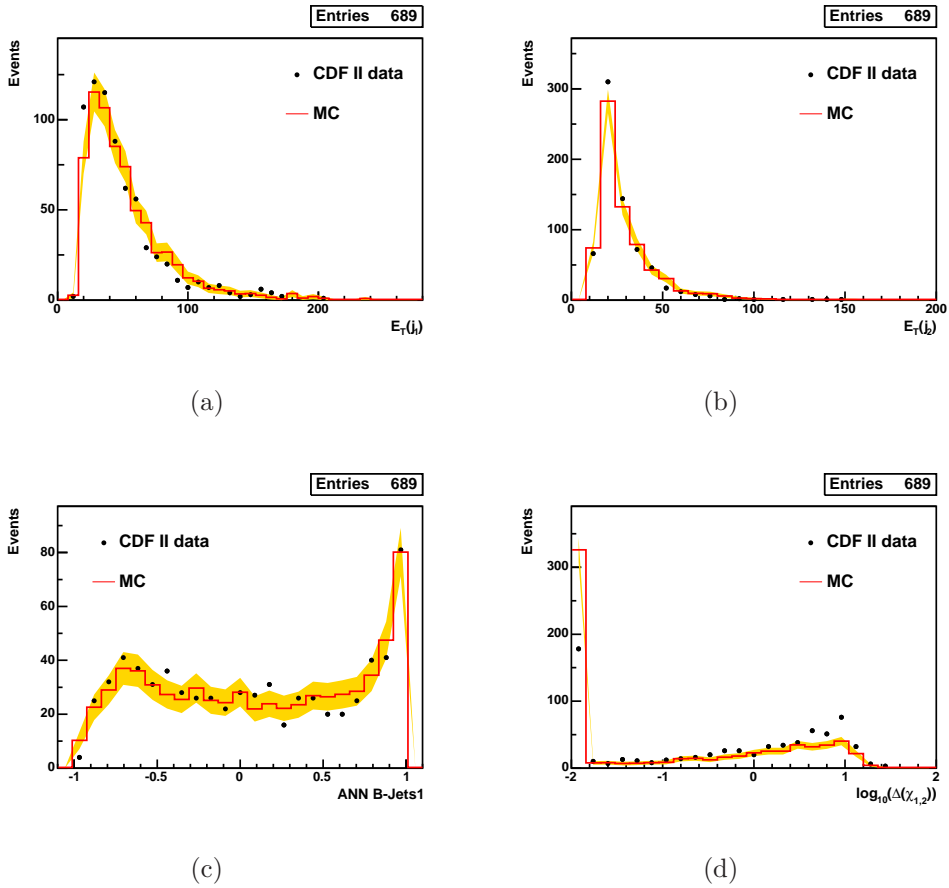(b)

(c)

(d)

Figure 5.12: Some of the input variables for the neural network, comparing the data and the prediction: (a) the transverse energy of the leading jet, (b) the transverse energy of the second jet, (c) the neural network output of the $b$ tagger for the first b tagged jet, and (d) the $\chi^2$ of the Kinematic Fitter for top reconstruction. The Monte Carlo is normalized to the data.

## 5.3 Neural Network Output Templates

Since our final aim in this analysis is to measure single-top content in data, we use the neural network output as discriminant between signal and background. We use a binned likelihood function to search for single-top in data or to calculate upper limits on the cross section of single-top production in case we find no evidenz for single-top. Therefore we use shape templates from Monte Carlo samples for different processes in the likelihood fit. In addition we want to divide the output of the neural network into a signal and a background region.

Therefore, we determine the predicted cut efficiencies which display the ratio of each process that would survive a cut on a certain neural network output. Figure 5.13 (a) shows four examples of the cut efficiencies against the neural network output. The $t$-channel is shown in black, the $s$-channel is shown in red, $Wb\bar{b}$ dashed in blue and $t\bar{t}$ in yellow. This illustrates the use of this cut since the shapes of the curves show that there are certain cuts where we loose significantly more on background than we loose on signal. These cut efficiencies and the number of expected events lead to an estimate for the ratio of signal $S$ to square-root of background $\sqrt{B}$ for the different neural network outputs. This is a good indicator to choose the cutting point for our neural network. The graph for $\sigma = \frac{S}{\sqrt{B(1+0.2)}}$ can be found in figure 5.13 (b). It includes a systematic error of 20% on the background which makes the $\frac{S}{\sqrt{B}}$ smaller to be conservative.
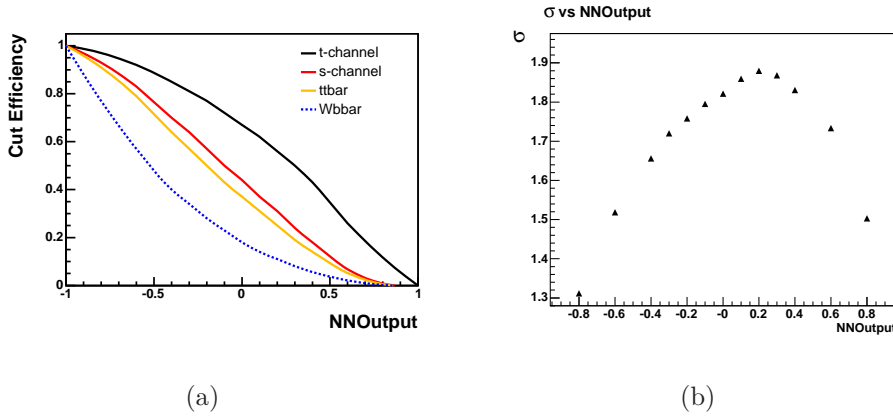


Figure 5.13: (a) Cut efficiency against the neural network output for different samples, (b) $\sigma$ versus neural network output.

The sigma distribution has its maximum around NNOutput = 0.2; therefore, we choose to cut at this value. The expected $\sigma$ is roughly 1.9. The output of the neural network is divided into two regions: The background region from -1.0 to 0.2 and the signal region from 0.2 to 1.0.

It would be possible to create templates for every physical process, but the more templates are handed to the likelihood fit the more difficult it gets, since the fit has more free parametersa and is not able to distinguish between the different templates.

Looking at the different network outputs for the different processes in figure 5.14, we realize that it is possible to combine the different processes into four templates because the shape of the neural network output is very similar. The four templates are: Single-top ($t$- and $s$-channel), $t\bar{t}$, charm-like and bottom-like. The charm-like template contains $Wc\bar{c}$, $Wc$, $WW$ and mistags. The bottom-like template is $Wb\bar{b}$, $WZ$, $ZZ$, $Z->\mu\mu$ and $Z->\tau\tau$.



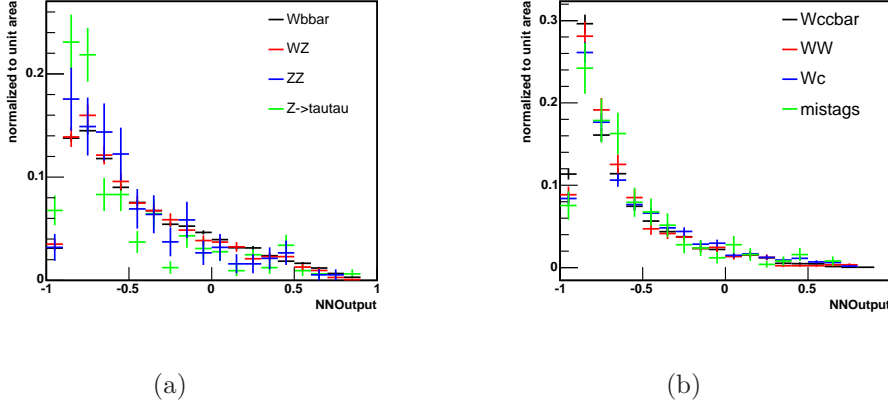(a)                                                      (b)

Figure 5.14: (a) Neural network output for the different physical processes plotted for the bottom like processes and (b) the charm like processes.

The number of expected events for the complete region can be found in table 5.3. It also contains the cut efficiencies for NNOutput = 0.2. The numbers of the expected events in the signal region is derived by the multiplication of the cut efficiency with the total number of expected events. Consequently, the number of expected events in the background region is calculated by subtracting the number of expected events in the signal region from the total number of expected events. Figure 5.15 shows the signal region of the neural network.

The expectation based on table 5.3 can be seen as the stacked histogram of the different predicted processes. The data is shown as black marks with the statistical uncertainty shown as black error bars. We expect $65.1 \pm 9.3$ events in the signal region, while we observe 59 in data. The distribution shows no obvious evidence for single-top production.

Now the templates are created. The $t\bar{t}$ template is created by taking the output of the neural network of the $t\bar{t}$ sample. As mentioned before, the single-top template, the bottom like and the charm like templates consist of different processes. Every template has to be put together in the right proportion. Therefore, we take the resulting templates for every process and weight it according to the number of expected events for this process. Finally we have to create the three needed templates by adding up the right templates. That means for the single-top template the $t$-channel and $s$channel template, for the charm like template we need to add the $Wc\bar{c}$, $Wc$, $WW$ and mistags templates, and for the bottom-like template we need
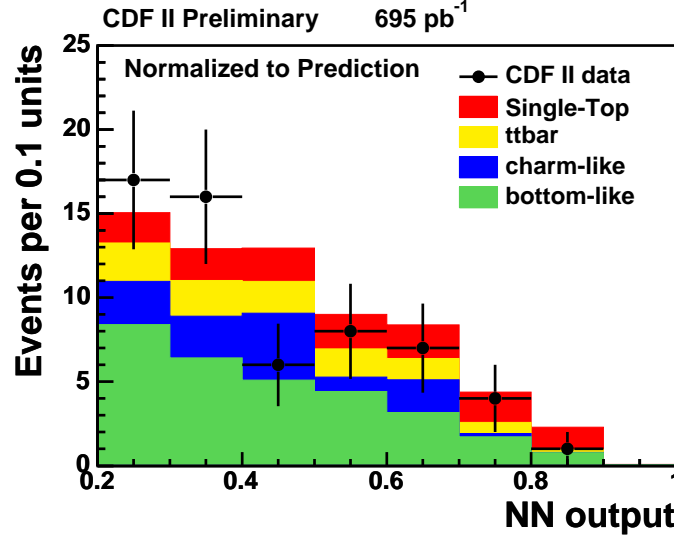
Figure 5.15: Data distribution of the neural network output in the signal region. The expectation is displayed as a stacked histogram and the data as black marks.

to add the $W b\bar{b}$, $WZ$, $ZZ$, $Z->\mu\mu$ and $Z->\tau\tau$ templates. Since the templates of every process were weighted according to the number of expected events for the corresponding process it can happen that the final templates is dominated by one process, e.g. the bottom like template is strongly dominated by the $W b\bar{b}$ sample. The resulting four different templates can be seen in figure 5.16 (a) for the combined neural network. The templates are normalized to unit area for better comparison. The single-top template is shown in black and even though the shape is fairly flat it can be seen that it is the template which is the most signal-like. The templates for the charm-like background in green and the bottom-like background in blue peak for low neural network outputs, which classifies them clearly as background. Since $t\bar{t}$ includes a real top, the template for this process in red is also flat. But in comparison to the single-top template it has its peak in the background region. Figure 5.16 (b) shows the neural network output for the two single-top production modes we consider for the combined neural network, $t$-channel in red and $s$-channel in green. The $t$-channel is better recognized as signal than the $s$-channel. In black, the combined single-top template can be seen. They are also normalized to unit area.

In figure 5.17, the five templates for the separate search are shown. They are created in the same way as the templates for the combined search except for the fact that they are two dimensional with the output of the $t$-channel neural network on the x-axis and the output of the $s$-channel neural network on the y-axis. Figure 5.17 (a) shows that $t$-channel events are detected by the $t$-channel network as signal as well as by the $s$-channel network. The $s$-channel events on the other hand are not as well detected as the $t$-channel events which can be seen in figure 5.17 (b). Figure 5.17 (c) shows the $t\bar{t}$ template which looks a bit like the $s$-channel template, which approves the combined templates where $s$-channel and $t\bar{t}$ also looked a bit alike.

| Process | bkg region | signal region | total | cut eff. |
|---|---|---|---|---|
| total | 609.0 | 65.1 | 674.1 | — |
| $t$-channel | 7.35 | 9.36 | 16.71 | 0.56 |
| $s$-channel | 7.95 | 3.57 | 11.52 | 0.31 |
| single-top | 15.30 | 12.93 | 28.23 | — |
| $t\bar{t}$ | 30.26 | 10.09 | 40.34 | 0.25 |
| $WZ$ | 5.53 | 0.56 | 6.09 | 0.092 |
| $ZZ$ | 0.19 | 0.02 | 0.21 | 0.074 |
| $Z \rightarrow \mu\mu$ | 4.03 | 0.41 | 4.44 | 0.092 |
| $Z \rightarrow \tau\tau$ | 2.35 | 0.24 | 2.59 | 0.092 |
| $Wb\bar{b}$ | 232.68 | 28.76 | 261.44 | 0.11 |
| bottom-like | 244.79 | 29.98 | 274.8 | — |
| $Wc\bar{c}$ | 90.46 | 2.70 | 93.16 | 0.029 |
| $Wc$ | 66.07 | 3.33 | 69.40 | 0.048 |
| $WW$ | 3.71 | 0.1 | 3.81 | 0.026 |
| mistags | 158.39 | 5.91 | 164.3 | 0.036 |
| charm-like | 318.62 | 12.05 | 330.67 | — |

Table 5.3: Number of expected events in the signal, background, and total region and the cut efficiency for a cut at NNOutput = 0.2.

The difference between them is again that the $t\bar{t}$ background template peaks a bit towards lower neural network output values. The charm-like template (figure 5.17 (d)) and the bottom-like template (figure 5.17 (e)) both peak in the lower left corner which indicates that they are recognized as background by both neural networks.
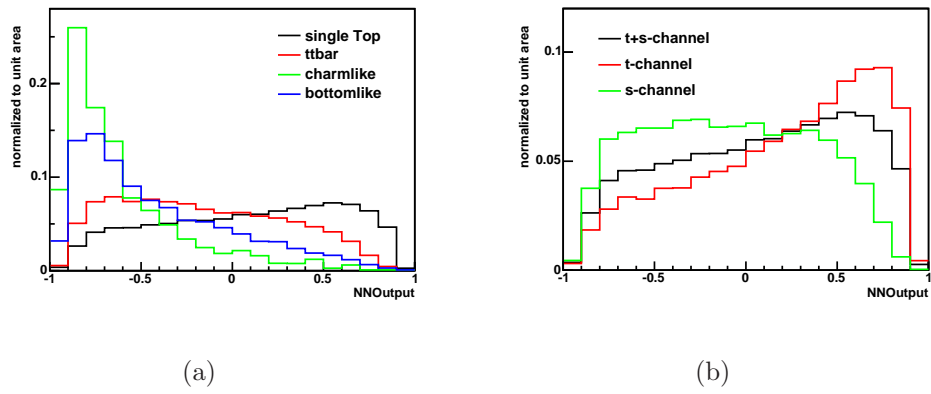
Figure 5.16: (a) The four templates: single-top, $t\bar{t}$, charm like and bottom like; (b) the two different single-top processes and the single-top template.
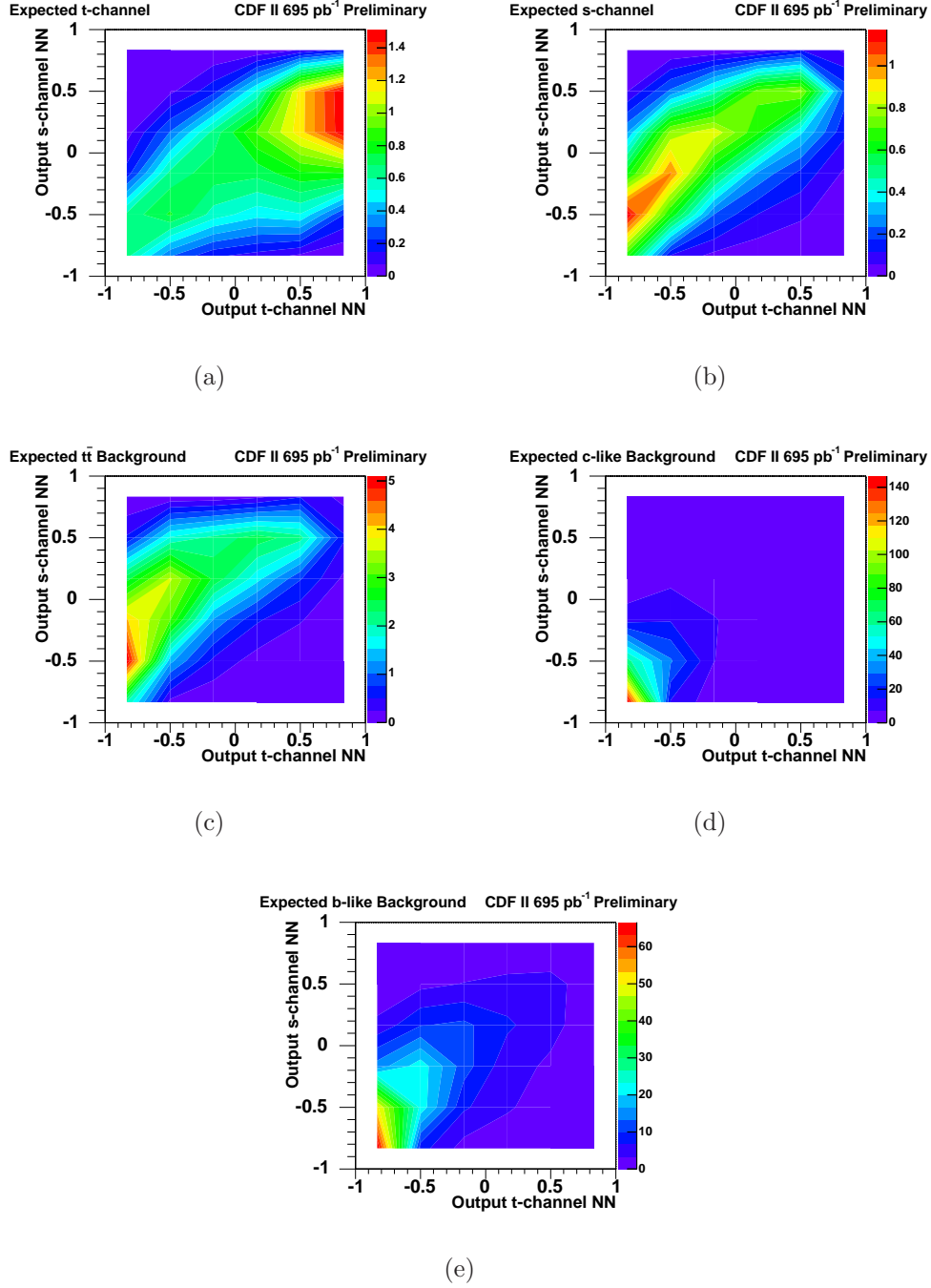
(a)

(b)

(c)

(d)

(e)

Figure 5.17: Two signal 2D templates for (a) $t$-channel and (b) $s$-channel and the three background 2D templates of (c) $t\bar{t}$, (d) charm-like and (e) b-like backgrounds.

# Chapter 6

# Single-Top Search

In this chapter, the analysis techniques used in this analysis are presented as well as the results for the different searches. The likelihood function which is used to measure the single-top contents of the CDF data samples and to calculate upper limits on the single-top production cross sections is introduced in the first part of the chapter. After that, the results of the combined search with neural networks are presented followed by the results of the simultaneously measured cross section in the separate search.

## 6.1 Likelihood Fit

In our analysis, a binned likelihood function, based on reference [59], is used. Binned likelihood functions use the contents $n_i$ of a set of bins $i$ derived from a distribution as observables. The parameter aimed to measure is the production cross section $\sigma$. For convenience reasons, the parameter $\beta = \sigma/\sigma^{SM}$ is used, which is the cross section normalized to its Standard Model prediction.
The binned likelihood function $L(\beta)$ is defined by

$$L(\beta) = \prod_{i=1}^{B} P(n_i, \mu_i) \tag{6.1}$$

The statistical content $n_i$ of a bin $i$ is described by the Poisson distribution

$$P(n_i, \mu_i) = \frac{\mu_i^{n_i}}{n_i!} e^{-\mu_i} \tag{6.2}$$

where $\mu_i$ denotes the expectation value of the distribution in this bin $i$. This expectation value $\mu_i$ is the sum of all signal and background process expectation values $\mu_{ji}$ contributing to the specific bin $i$. Each $\mu_{ji}$ is the product of the normalized production cross section $\beta_j$, the integrated luminosity $\mathcal{L}_{\mathsf{int}}$, the event detection efficiency of process $j$, $\nu_j$, times the normalized content of bin $i$ of the template histogram for process $j$, $\alpha_{ji}$. The product of $\alpha_{ji}$ times $\nu_j$ is sometimes called the bin specific acceptance function $\nu_{ji}$. Thus, the Poisson mean for a bin $i$ in the presence of $d$

signal and background processes is

$$\mu_i = \sum_{j=1}^{d} \mu_{ji} = \sum_{j=1}^{d} \nu_{ji} \cdot \beta_j \cdot \mathcal{L}_{\text{int}} = \sum_{j=1}^{d} \nu_j \cdot \alpha_{ji} \cdot \beta_j \cdot \mathcal{L}_{\text{int}} \qquad (6.3)$$

The likelihood consisting of $B$ bins within this scenario is then defined by

$$L(\beta) = \prod_{i=1}^{B} \frac{\mu_i^{n_i}}{n_i!} e^{-\mu_i} \qquad (6.4)$$

We use the same likelihood function as used for the single-top search described in reference [4]. In the following, the likelihood fit is explained considering the combined search. Therefore, there are only four classes of processes. In the case of the separate search, five processes are considered. In the combined case the classes are: the combined single-top production modes, $t\bar{t}$ background, charm-like background and bottom-like background.

All the formulas do not include any systematic uncertainties this far. If we include a certain number $S$ of systematic uncertainties the expected mean in bin $i$ changes to:

$$
\begin{aligned}
\mu_i \;=\; & \sum_{j=1}^{4} \beta_j \cdot \nu_j \cdot \mathcal{L}_{\text{int}} \cdot \left\{ \prod_{m=1}^{S} (1 + \delta_m \cdot \epsilon_{jm}) \right\} \cdot \alpha_{ji} \\
& \cdot \left\{ \prod_{l=1}^{S} (1 + |\delta_l| \cdot (\kappa_{jli+} H(\delta_l) + \kappa_{jli-} H(-\delta_l))) \right\}
\end{aligned}
\qquad (6.5)
$$

Some of the variables used in equation 6.5 were already explained before but they are mentioned here for consistency reasons again. The following variable definitions are used:

1. We consider four processes: $j =$ index over the processes
   $j = 1$ : single-top, $j = 2$ : $t\bar{t}$, $j = 3$ : bottom-like backgrounds, $j = 4$ : charm-like backgrounds.

2. The event detection efficiency of process $j$ is named $\nu_j$.

3. The normalized content of bin $i$ of the template histogram for process $j$ is $\alpha_{ji}$.

4. We consider six effects which cause systematic uncertainties in acceptance and in the shape of the template histograms: jet energy scale, Parton Distribution Functions (PDF), Initial and Final State Radiation (ISR, FSR), signal generator, and neural net $b$ tag shapes. The sources of systematic uncertainties are indexed with $m$. The relative acceptance uncertainties due to these sources are named $\epsilon_{ji}$. The relative uncertainties in the bin content of bin $i$ of the template histograms are called $\kappa_{jli}$. Most shape uncertainties are asymmetric. Thus, we have to introduce positive and negative uncertainties: $\kappa_{jli+}$ and $\kappa_{jli-}$.

5. The variation in strength of a systematic effect $m$ is measured with the variable $\delta_m$.

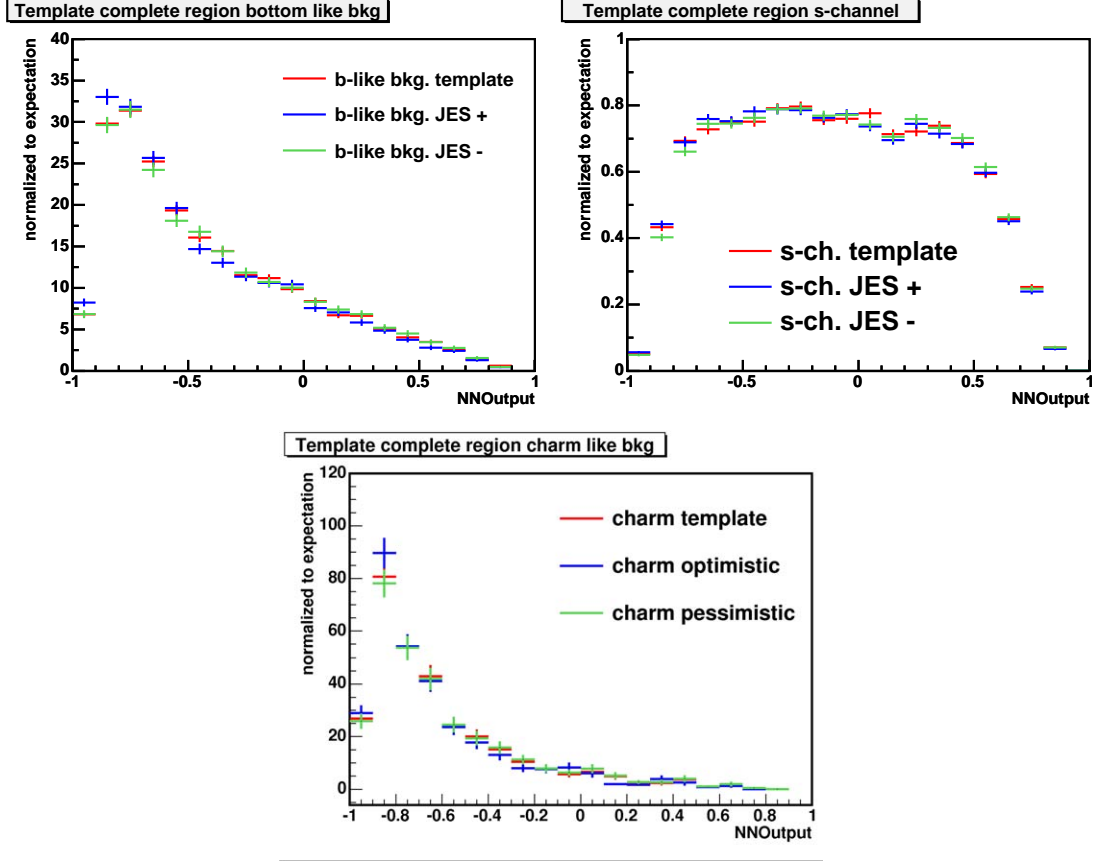Figure 6.1 illustrates how shape uncertainties are treated. We use shifted histograms



Figure 6.1: Shifted histograms to estimate the shape uncertainty due to the uncertainty in the jet energy scale for the $b$-like background and the $s$-channel, and due to the uncertainty in the modeling of the neural network $b$ tagger.

for the different uncertainties. As an example the shifted histograms due to the jet energy scale for the bottom-like background and the $s$-channel are shown. Another example is the shifted histogram of the charm-like background due to the uncertainty in the modeling of the neural network $b$ tagger. We estimate the individual contribution of a particular source of systematics by investigating the shift in the central value. These results are given in table 6.1. We add an additional systematic uncertainty for the modeling of the non-$W$ contribution. We vary the assignment of the non-$W$ to the two different background templates we use. The default is bottom-like:charm-like = 80:20. We vary between bottom-like:charm-like = 100:0 and bottom-like:charm-like = 50:50. This gives us large modeling uncertainties of about 30%, mainly because the non-$W$ background is so large in the PHX lepton category. It sticks out that the total systematic uncertainties for the $s$-channel and the combined search are dominated by the non-$W$ modeling, whereas the total systematic uncertainty of the $t$-channel search is mainly dominated by the jet energy

| Syst. Source | $t$-channel | $s$-channel | $t$- and $s$-channel combined |
|---|---|---|---|
| ISR | ±2% | ±1% | ±2.7% |
| FSR | ±3% | ±1% | ±2.1% |
| Jet Energy | +7% / -5% | ±4% | +5% / -1% |
| PDF | ±3% | ±1.5% | ±1.8% |
| NN b tagger | ±1% | ±1% | ±6% |
| non-$W$ modeling | ±2% | +170% / -90% | +26% / -31% |
| Total | +8.7% / -7.2% | +170.1% / -90.1% | +27.4% / -31.8% |

Table 6.1: Estimate of the systematic uncertainty on the central value of the likelihood fits to the neural network outputs. The $t$- and $s$-channel values refer to the separate search, the number of $t$- and $s$-channel combined to the combined search.

scale uncertainty.

The inclusion of systematic uncertainties is described in more detail in reference [59]. To obtain the probability distribution for the SM normalized signal cross sections $\beta_{t+s-ch}$, we integrate out all nuisance parameters from the likelihood function:

$$L^*(\beta_{t+s-ch}) = \int_0^{+\infty} \int_0^{+\infty} \int_0^{+\infty} L(\beta_1, ..., \beta_4) \, d\beta_2 d\beta_3 d\beta_4 \qquad (6.6)$$

This reduced or marginalized likelihood only depends on the normalized combined single-top cross section. This technique is also referred to as marginalization. The integration is implemented as a Monte Carlo integration. For each integration point random numbers are generated for the nuisance parameters $\beta_k$ according to Gaussian distributions $G(\beta_k, 1.0, \Delta_k)$. With other words in case of 10000 pseudo-experiments we first throw 10000 random numbers, for every background process that we consider, based on a Gaussian distribution. In addition, all random numbers we need for the systematic uncertainties are thrown as well and stored together with the other numbers as tuples. They are used in the same combination in every integration step. This reduces the needed time to perform the pseudo-experiments significantly. The width of the Gaussian is given by the uncertainty of the background estimation. Based on reference [60] we choose $\Delta_3 = 19\%$ for the $t\bar{t}$ background, $\Delta_4 = 13\%$ for the charm-like background, and $\Delta_5 = 23\%$ for the bottom-like background. The marginalized likelihood is finally maximized with respect to the data set.

## 6.2 Combined Search

We use a network trained with $t$-channel and $s$-channel events as signal to search for single-top events. We assume the ratio between the two single-top processes to be as predicted by the Standard Model. Now the maximum likelihood fit introduced in section 6.1 is applied to the network output. Using pseudo-experiments, we determine whether it is better to fit the entire neural network output histogram

or only the signal region. Pseudo-experiments are also known as ensemble tests. They are used to test that the analysis is not biased and to determine the predicted sensitivity as well as the expected result of the analysis. For the combined analysis, we do 10000 pseudo-experiments. We start by throwing random numbers of events for every process $N_j$ from a Poisson distributions. The Poisson distributions are well defined since we know the number of expected events for every process which is used as the expectation value of the Poisson distribution. Based on that we create the distribution of the discriminants by throwing $N_j$ random numbers for every process from the corresponding templates. This already creates a set of 10000 pseudo data sets. At this point, the usual analysis technique starts. The pseudo data sets are analysed like real data and the fitted results of the pseudo-experiments are put into a histogram and the Root Mean Square (RMS) of this distribution is used as a measure for the sensitivity, i.e. the expected statistical uncertainty of a potential measurement. From these studies we conclude that a fit to the entire distribution returns the best sensitivity since the RMS for the fit of the signal region is bigger than the RMS for the fit of the total region.

The likelihood fit to data yields a best value for the cross section of

$$0.8^{+1.3}_{-0.8} \, (\text{stat.}) \, ^{+0.2}_{-0.3} \, (\text{syst.}) \, \text{pb}$$

This corresponds to $\beta = 0.27^{+0.46}_{-0.27} \, (\text{stat.})$. This value is taken from the maximum of the posterior probability density which is shown in figure 6.2. The resulting upper limit on the cross section is calculated by integrating over this distribution starting at zero until the integral covers 95% of the area under the distribution. This leads to an upper limit of $3.4 \, \text{pb}$ ($\beta_{95} = 1.17$) at the 95% confidence level.
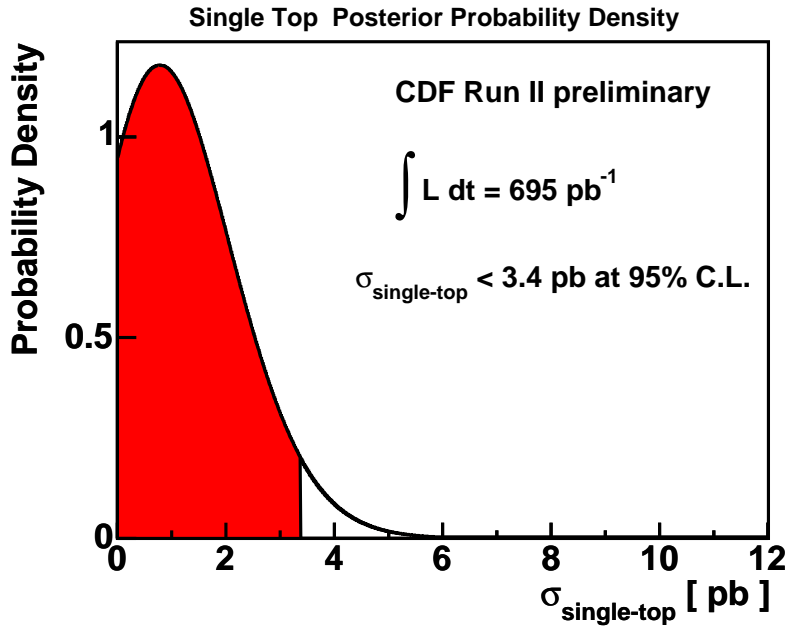


Figure 6.2: Posterior probability density for the combined search using the neural network.

The result of the likelihood fit to the data is illustrated in figure 6.3.  The data
is shown as black dots with error bars.  The fit result can be seen as the stacked
histogram added up from single-top in red, $t\bar{t}$ in yellow, charm-like backgrounds in
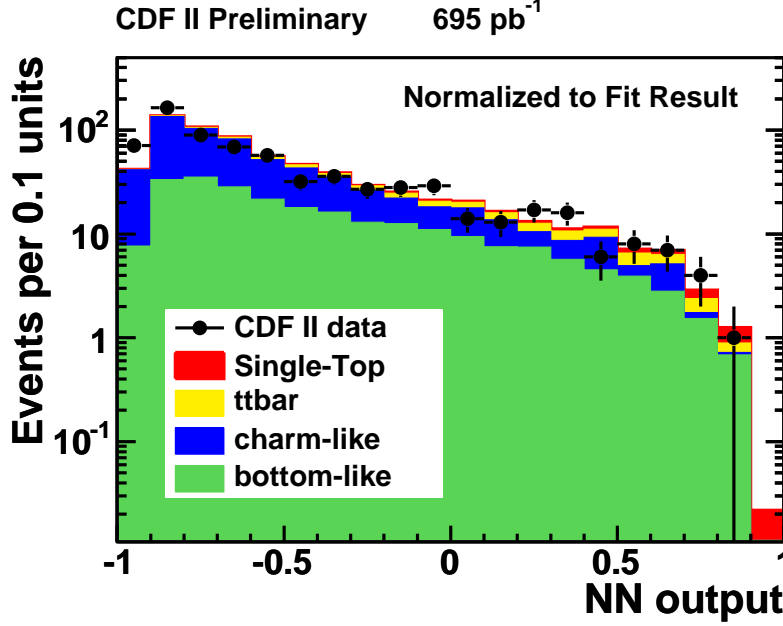blue and bottom-like backgrounds in green.



Figure 6.3: Data distribution of the neural network output in the entire output region. The fitted
components are shown as the stacked histogram.

The fit result on data has now to be compared with the expectation from pseudo-
experiments. Therefore the results of pseudo-experiments are shown in figure 6.4.
Figure 6.4 (a) shows the distribution of the upper limits obtained for the pseudo-
experiments.  Single-top events are included in these experiments at the expected
Standard Model rate.  We define the median of the distribution as the expected
upper limit, since that means that in 50% of the pseudo-experiments the expected
upper limit at the 95% C.L. was higher than the value we choose and in 50% it
was lower. With this definition we obtain 5.7 pb as the expected upper limit. The
distribution of fit values in figure 6.4 (b) shows the central values obtained from
the likelihood fit for each pseudo-experiment. The RMS of this distribution is our
expected statistical uncertainty which is 51%. That means that we can trust the
result of our likelihood fit to the data, since it is approved by the results of the
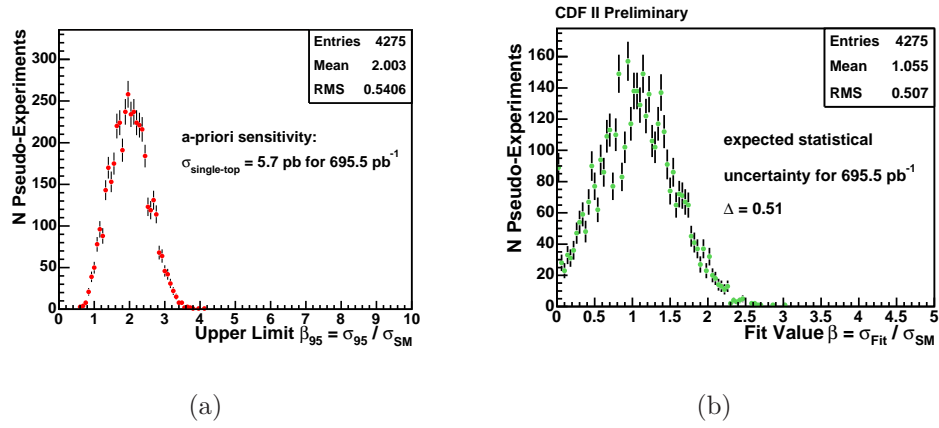pseudo-experiments within the uncertainties.

(a)                                        (b)

Figure 6.4: Results from ensemble test for the 1D neural network analysis which is used for a combined search for single-top.

## 6.3  Separate Search

We use two independent neural networks, one trained for $s$-channel and the other one for $t$-channel, which provide the opportunity to search for both channels simultaneously. The training and input variables of the networks are similar to the one described in section 5.2 except that each network is trained particularly for one channel. The creation of the templates for signal and background processes was already explained in section 5.3.

As described in section 6.2, we apply a maximum likelihood fit to the network 2D output. The only difference in the likelihood function used is the generalization for two dimensions. In order to estimate the sensitivity presuming a scenario of Standard Model single-top production, pseudo-experiments are performed in the same way as described in the last section. The resultant estimates of the likelihood fits for the Standard Model normalized $t$- and $s$-channel cross sections and the resultant 2D distribution are shown in figure 6.5. The RMS values are again taken



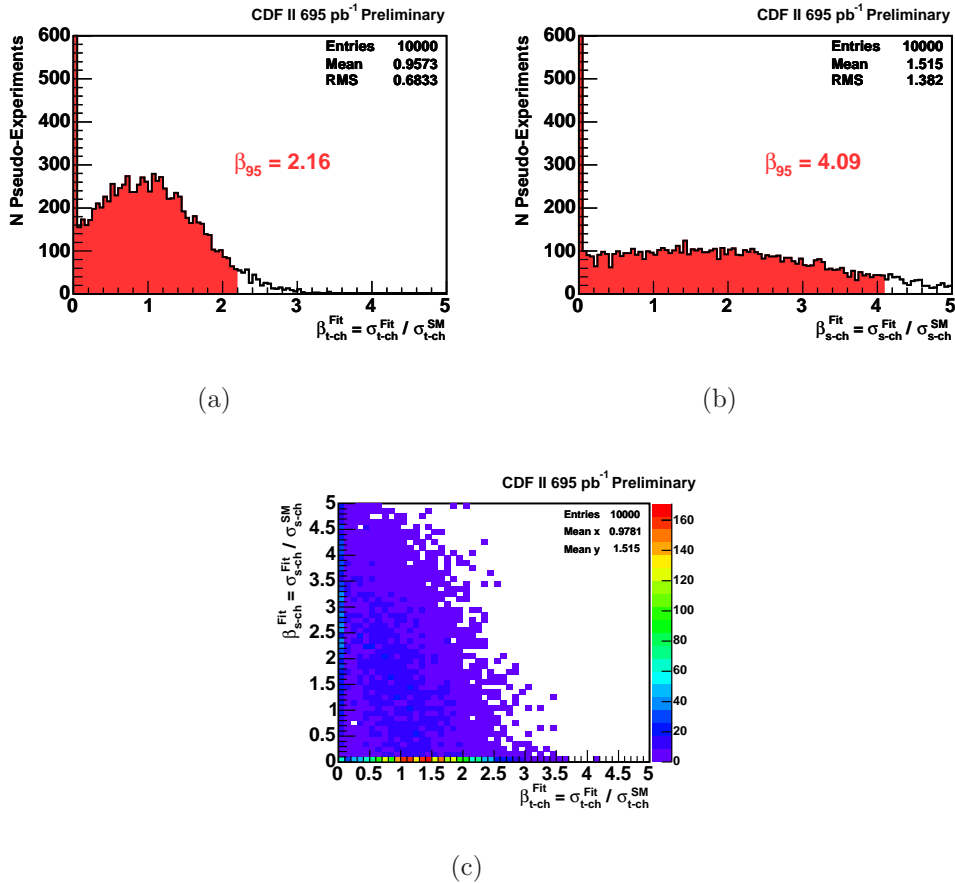(a)                                                  (b)



(c)

Figure 6.5: Pseudo-experiment distributions of (a) $t$-channel and (b) $s$-channel cross section measurement normalized to the SM prediction and (c) the 2D distribution.

to be the expected statistical uncertainties. For the $t$- and $s$-channel, the expected

statistical uncertainty is therefore 68% and 138%, respectively. Table 6.2 shows further values of the distributions. To calculate the expected upper limit we integrate

|  | $t$-channel | $s$-channel |
|---|---|---|
| 95% C.L. | 2.16 | 4.09 |
| RMS | 0.68 | 1.38 |
| # in 1st bin | 10.82% | 25.06% |

Table 6.2: Characteristical values of the pseudo-experiment distributions (in units of $\beta$). # in 1st bin shows the percentage of pseudo-experiments in which the fit finds no single-top at all.

over the distribution starting from zero until we cover 95% of the area. The value we stop at is the expected upper limit at 95% C.L. given in units of $\beta$. This translates to $\sigma_{t-ch} < 4.28$ pb at the 95% C.L. for the $t$-channel and $\sigma_{s-ch} < 3.60$ pb at the 95% C.L. for the $s$-channel. Table 6.2 also shows the percentage of pseudo-experiments in which the fit finds no single-top at all. In case of the $t$-channel network nearly 11% of the pseudo-experiments show no single-top $t$-channel fraction in the fit. For the $s$-channel network the fit result for $s$-channel is zero in 25% of the cases.

The 2D data distribution can be seen in figure 6.6 (a), the total expectation in figure 6.6 (b). The upper right corner in the two histograms is the signal region since this is the area where both the $t$-channel and $s$-channel network outputs are high. Neither a deficit nor an excess can be seen in this region. Figures 6.6 (c) and (d) show the corresponding 3D distributions for the data and the prediction on logarithmical scale. The corresponding likelihood fit to the data shown in figure 6.7 gives a cross sections for the $t$-channel of $\sigma_{t-ch} = 0.6^{+1.9}_{-0.6}(\text{stat.})^{+0.1}_{-0.1}(\text{syst.})$ pb and for the $s$-channel of $\sigma_{s-ch} = 0.3^{+2.2}_{-0.3}(\text{stat.})^{+0.5}_{-0.3}(\text{syst.})$pb. The central value is marked as the black point, the Standard Model prediction as the red point, the statistical uncertainties are given by the black one sigma line, and to get the upper limits at the 95% confidence level we take the values indicated by the yellow band on the $t$- and $s$-channel axis which are $\sigma_{t-ch} < 3.1$ pb and $\sigma_{s-ch} < 3.2$ pb at the 95% C.L., respectively.
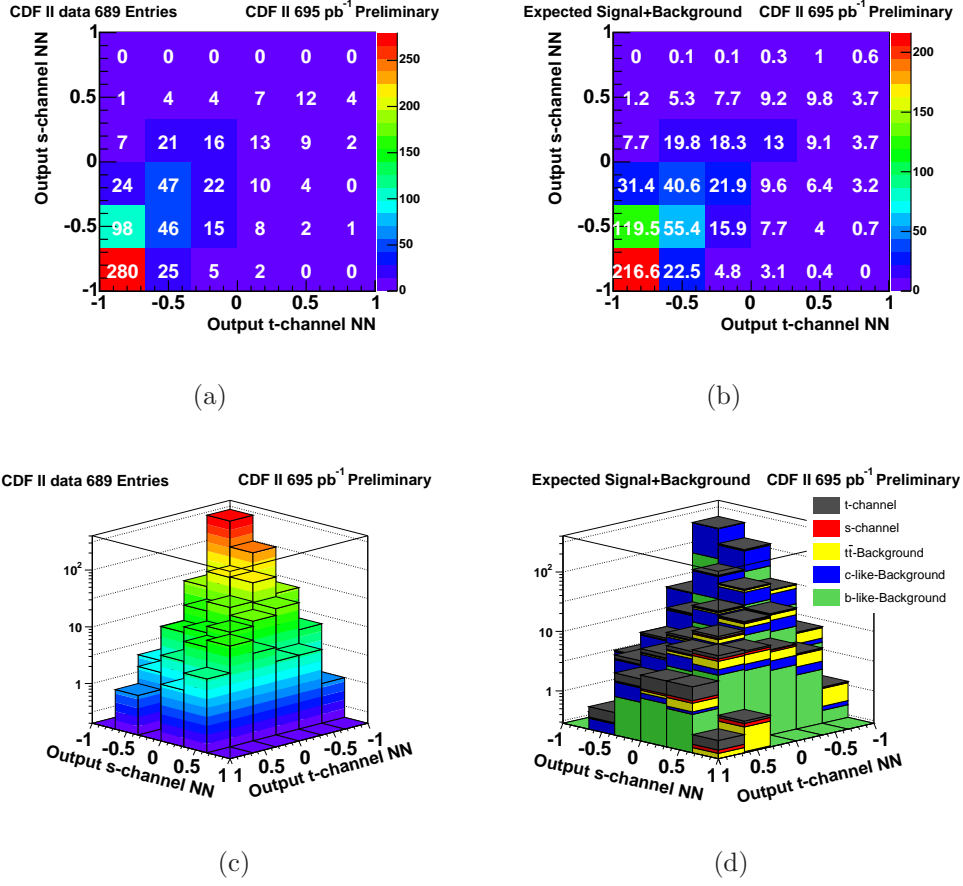
(a)



(b)



(c)



(d)

Figure 6.6: (a), (b) 2D, and (c), (d) 3D distributions of both neural network outputs of data on the left and total expectation on the right.
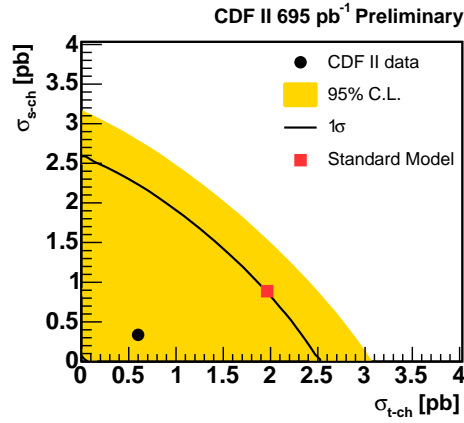


Figure 6.7: The likelihood fit estimate for $t$-channel and $s$-channel cross section measurement with the contour of the $1\sigma$ uncertainty and the contour of the upper limit at the 95% C.L.

# Chapter 7

# Conclusion and Outlook

In this thesis a data set corresponding to an integrated luminosity of 695 pb$^{-1}$ has been analysed in search for electroweak production of single-top quarks. The first step was to determine the number of expected signal and background events. It is important to make sure that the expectations match the numbers found in data well. This was the case. We used a neural network technique to separate the signal from the background events. In the process of optimizing the discrimination between signal and background we investigated the separation power of over 40 different variables. We found that 14 of those variables have a sufficient significance. They are used as input variables for the neural networks. Those neural networks were then trained with Monte Carlo events. After that we applied them to Monte Carlo and data samples. We built templates from the output of the Monte Carlo samples. Those templates were used in the likelihood fit to match the data. Since this analysis is a shape based analysis it is important that all the processes involved are well understood. This was verified by a data Monte Carlo comparisons of the input variables.

Two different approaches are taken to measure the cross sections of single-top quark production. The first one is the combined search in which only one network is used. This network is trained with $t$- and $s$-channel events as signal. The second approach is the simultaneous measurement of the $t$-channel and the $s$-channel production cross section. Here two separately trained networks are used. Since neither analysis finds any evidence for a signal we calculate upper limits on the single-top cross section using a Bayesian technique. The observed limit at the 95% Confidence Level (C.L) on the $s + t$ combined single-top cross section is 3.4 pb. For the separate search the observed limits at the 95% C.L. are 3.1 pb and 3.2 pb on the $t$-channel and $s$-channel single-top production cross section, respectively. In table 7.1, we summarize the expectations and results of our analyses. These limits are the most stringent limits for the single-top quark production determined up to date.

Based on the results of this analysis there seems to be a deficit of single-top events. But it is hard to conclude whether this deficit is caused by the nature of electroweak single-top quark production, by not well understood background processes and de-

| Technique | Expected Limit | Observed Limit |
|-----------|:--------------:|:--------------:|
| NN combined | 5.7 pb | 3.4 pb |
| NN $t$-channel | 4.3 pb | 3.1 pb |
| NN $s$-channel | 3.6 pb | 3.2 pb |

Table 7.1: Summary of expected and observed upper limits at the 95% Confidence Level. The pseudo-experiments that were used to find the expected limits included single-top events at the predicted Standard Model rate.

tector effects, or even wrong theoretical calculations. The CDF collaboration will be able to answer this question in the near future with further improved analyses and more data taken.

# Bibliography

[1] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **74**, 2626 (1995);
DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

[2] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **65**, 091102 (2002);
CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **69**, 052003 (2004).

[3] DØ Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. B **517**, 282 (2001);
DØ Collaboration, B. Abbott *et al.*, Phys. Rev. D **63**, 031101 (2001).

[4] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **71**, 012005 (2005).

[5] DØ Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. B **622**, 265 (2005).

[6] S. L. Glashow, *Partial symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579-588.

[7] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264-1266.

[8] A. Salam, *Weak and electromagnetic interactions*. Originally printed in *Svartholm: Elementary Particle Theory, Proceeding Of The Nobel Symposium Held 1968 At Lerum, Sweden*, Stockholm 1968, 367-377.

[9] Th. Müller, *Von höchsten Energien zu kleinsten Teilchen*,
Talk at the 'Physik am Samstag'-lecture series in Karlsruhe, Germany, 17. June 2006,
http://www.physik.uni-karlsruhe.de/3block1.php/Aktuelles/Physik_am_Samstag_06/PaS06_Mueller.pdf.

[10] W.-M. Yao *et al.*, J. Phys. G **33**, 1(2006)
updates for the 2006 edition available on http://pdg.lbl.gov/

[11] H. L. Lai *et al.*, Eur. Phys. J. C **12**, 375 (2000).

[12] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973);
N. Cabbibo, Phys. Rev. Lett. **10**, 531 (1963).

[13] T. E. W. Group, *Combination of CDF and DØ results on the mass of the top quark*, hep-ex/0603039, March 2006.

[14] W. Wagner, Rep. Prog. Phys. **68**, 2409 (2005).

[15] N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003);
M. Cacciari *et al.*, J. High Energy Phys **0404**, 068 (2004).

[16] Z. Sullivan, Phys. Rev. D **70**, 114012 (2004).

[17] B. W. Harris *et al.*, Phys. Rev. D **66**, 054024 (2002).

[18] CDF Collaboration, A. Abulencia *et al.*, Phys. Rev. Lett. **97**, 062003 (2006).

[19] J. Alwall *et al.*, *Is $V_{tb} \approx 1$?*, hep-ph/0607115 v2, 2006.

[20] P. Q. Hung, *Minimal SU(5) Resuscitated by Higgs Coupling Fixed Points*, hep-ph/9710297 v2, 2005.

[21] P.H. Garbincius, *Tevatron collider operations and plans*, hep-ex/0406013.

[22] Fermilab Beams Division, *Run II Handbook*, http://www-bd.fnal.gov/runII/index.html.

[23] CDF Collaboration, R. Blair *et al.*, FERMILAB-Pub-96/390-E (1996).

[24] R. Brun and F. Carminati, CERN Programming Library Long Writeup **W5013** (1993).

[25] CDF Collaboration, T. Affolder *et al.*, *CDF central outer tracker*, Nucl. Intrum. Meth. **526** (2004) 249-299.

[26] G.Grindhammer, M. Rudowicz, and S.Peters. *The fast simulation of electromagnetic and hadronic showers*, Nucl. Instrum. Meth. **A 290** (19990) 469.

[27] E. Gerchtein and M. Paulini, physics/0306031 (2003).

[28] T. Stelzer and W. F. Long, Comp. Phys. Commun. **81**, 357 (1994);
F. Maltoni and T. Stelzer, hep-ph/0208156 (2002).

[29] T. Sjöstrand *et al.*, Comp. Phys. Commun. **135**, 238 (2001).

[30] ALPGEN, a generator for hard multiparton processes in hadronic collisions, M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. Polosa, *JHEP 0307:001* (2003), hep-ph/0206293.

[31] HERWIG 6.5, G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, *JHEP 0101:010* (2001), hep-ph/0011363, hep-ph/0210213.

[32] E. Thomson *et al.*, *Top and Electro Weak Event Classification Module for CDF Run II*, CDF note 5947.

[33] D. O. Carlson and C. P. Yuan, Phys. Lett. B **306**, 386 (1993);
A. P. Heinson, A. S. Belyaev and E. E. Boos, Phys. Rev. D**56**, 3114 (1997).

[34] G. Mahlon and S. Parke, Phys. Rev. D 55:7249-7254, 1997.

[35] S. Budd *et al.*, *Validation of the MadEvent Single-Top Samples with ZTOP NLO Calculations*, CDF note 7701, Version 2.0, February 2006 (Version 1.0 from June 2005).
Jan Lück, *Kinematics of electroweak single-top quark production*, FERMILAB-MASTERS-2006-1.

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

[37] R. Erbacher *et al.*, *Preliminary Event Selection and t-tbar Signal Acceptance for the Winter 2005 Top Lepton + Jets Samp*, CDF note 7372 (2005).

[38] M. Tönnesmann, *Jet Reconstruction at CDF*, Prepared for 10th International Conference on Calorimetry in High Energy Physics (Calor 2002), Pasadena, California, 25.-30. March 2002.

[39] A. Bhatti *et al.*, *Determination of the jet energy scale at the collider detector at Fermilab*, hep-ex/0510047, October 2005.

[40] W. Yao *et al.* , *A Seed Vertex b-Tag Algorithm for Top*, CDF note 2716, 1994.
C. Neu *et al.*, *SecVtx Optimization Studies for 5.3.3 Analyses*, CDF note 7578, 2005.

[41] V. Boisvert, *Trigger Efficiencies for High $E_T$ Central Electrons in the Gen6 data*, CDF note 7939, version 2.1, December 2005.
B. Han and E. Halkiadakis, *MET PEM trigger efficiecy for Phoenix electrons*, CDF note 7940, version 2.0, December 2005.

[42] T. Spreitzer, C. Mills, and J. Incandela, *Electron Identification in Offline Release 6.1*, CDF note 7950.

[43] U. Grundler, A. Taffard, X. Zhang, *High-$P_T$ muons recommended cuts and efficiencies for Winter 2006*, CDF note 7956.

[44] W.K. Sakumoto, *Event $-Z_{vtx}| \leq 60$ cm Cut Acceptance for Run II*, CDF note 7935.

[45] F. Garberson *et al.*, *Combination of the SECVTX b-Tagging Scale Factors for Winter 2006 Analyses*, CDF note 8025, January 2006.

[46] P. Nason, S. Dawson and R.K. Ellis: Nucl. Phys. B 327, 607-633, 1988.

[47] W. Beenakker, H. Kuijf, W.L. van Neerven and J. Smith: Phys. Rev. 40, 54-82, 1989.

[48] E.L. Berger and H. Contopanagos: Phys. Rev. D 57, 253-264, 1998.

[49] R. Bonciani, S. Catani, M.L. Mangano and P. Nason: Nucl. Phys. B 529, 424, 1998.

[50] M. Cacciari, S. Fricione, M.L. Mangano, P. Nason and G. Ridolfi, hep-ph/0303085, 2003.

[51] J.M. Campbell and R.K. Ellis, *Update on vector boson pair production at hadron colliders*, Phys. REv. D 60, 113006 (1999).

[52] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **71**, 052003, hep-ex/0410041.

[53] A. Foland *et. al.*, *SecVtx Tag Rate Matrix for 5.3.3_nt*, CDF note 7326.

[54] S. Budd *et al.*, *Single-Top analysis with 0.7 fb$^{-1}$*, CDF note 8112.

[55] M. Feindt, e-Print Archive physics/0402093 (2004).

[56] M. Feindt *et al.*, *A Neural Network b Tagger for Single-Top Analyses*, CDF note 7816.

[57] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. B **406**, 187 (1993).

[58] S. Budd *et al.*, *A Multivariate Likelihood Search Search for Single Top*, CDF note 7493.

[59] C. Ciobanu *et al.*, *Likelihood Function for Single-Top Search with 162 pb$^{-1}$*, CDF note 7106.

[60] S. Budd *et al.*, *Event detection efficiency for single-top events in CDF II*, CDF note 7863.