Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

IEKP-KA/2008-17

B-Jet und C-Jet Identifikation mit Neuronalen Netzen sowie Kombination von multivariaten Analysen zur Einzel-Top-Quark Produktion

Manuel Renz

Diplomarbeit

an der Fakultät für Physik
der Universität Karlsruhe (TH)

Referent: Priv.-Doz. Dr. Wolfgang Wagner
Korreferent: Prof. Dr. Thomas Müller

Institut für Experimentelle Kernphysik

Juni 2008

Zusammenfassung

Bis zum Ende des 19. Jahrhunderts hielt man Atome für die kleinsten Bausteine der Materie. Mit der Erkenntnis, dass Atome aus negativ geladenen Elektronen bestehen, die einen positiv geladenen Kern umgeben, änderte sich dieses Bild grundlegend. In den folgenden Jahrzehnten führten zahlreiche Entdeckungen, wie beispielsweise die Beobachtung von Protonen und Neutronen oder der Nachweis von Quarks, zu einem tieferen Verständnis der Prozesse, die dem Aufbau der uns bekannten Materie zugrunde liegen. Heute werden die fundamentalen Teilchen und ihre Wechselwirkungen mit dem, in der Zeit um 1970 entwickelten Standardmodell der Teilchenphysik beschrieben. Die Suche nach dem Higgs-Boson, dem letzten Teilchen, welches vom Standardmodell vorhergesagt wird, aber bisher noch nicht entdeckt werden konnte, sowie die Überprüfung weiterer, vom Standardmodell abgeleitete Vorhersagen, sind Gegenstand der aktuellen Forschung. Zu diesem Zweck werden Teilchenbeschleuniger gebaut, in denen Elektronen mit Positronen, Protonen mit Antiprotonen oder Protonen mit Protonen zur Kollision gebracht werden. Momentan erreicht das Tevatron, ein Proton-Antiproton Beschleuniger in der Nähe von Chicago (USA), mit 1.96 TeV weltweit die höchste Schwerpunktsenergie in Teilchenkollisionen. Mit dem diesjährigen Start des Large Hadron Colliders (LHC) am Europäischen Zentrum für Teilchenphysik (CERN) in Genf können sogar Schwerpunktsenergien von 14 TeV erreicht werden.
Das schwerste der bisher bekannten Quarks, das Top-Quark, wurde 1995 von der CDF- und der DØ-Kollaboration am Tevatron Beschleuniger entdeckt. Neueste Messungen ergeben eine Top-Quark-Masse von 172.6 ± 1.4 GeV/c^2 [3], was ungefähr der Masse eines Goldatoms entspricht. Am Tevatron werden Top-Quarks hauptsächlich über die starke Wechselwirkung in so genannten "Top-Quark-Paaren" erzeugt. Hierbei entstehen gleichzeitig ein Top- und ein Antitop-Quark. Theoretische Berechnungen sagen für diesen Prozess einen Wirkungsquerschnitt von $\sigma_{\mathrm{t\bar{t}}} = 6.70^{+0.71}_{-0.88}$ pb vorher [11].
Zusätzlich gibt es im Standardmodell die Möglichkeit, einzelne Top-Quarks über die schwache Wechselwirkung zu erzeugen. Dabei unterscheidet man drei verschiedene Produktionsprozesse, die sich vor allem in der Virtualität $Q^2 = -q^2$ des ausgetauschten W-Bosons unterscheiden, wobei q hier den W-Boson Viererimpuls bezeichnet. Zum einen gibt es die t-Kanal und die s-Kanal Produktion, welche ihren Namen von der, zur Beschreibung des auftretenden Übergangsmatrixelements benutzten Mandelstamvariable erhalten. Zum anderen wird die assoziierte Produktion vorhergesagt, in der das Top-Quark zusammen mit einem reellen W-Boson erzeugt wird. Am Tevatron Beschleuniger ist der Wirkungsquerschnitt für die assoziierte Produktion vernachlässigbar

klein. Die Wirkungsquerschnitte der s-Kanal bzw. t-Kanal Produktion sind mit vorhergesagten $\sigma = (0.88 \pm 0.11)$ pb bzw. $\sigma = 1.98^{+0.28}_{-0.22}$ pb zusammen fast halb so groß, wie der Wirkungsquerschnitt für $t\bar{t}$ Produktion.
Eine präzise Messung des Wirkungsquerschnitts für die Produktion einzelner Top-Quarks, würde einen interessanten Test des Standardmodells ermöglichen. Jeder der oben beschriebenen Prozesse zur Produktion einzelner Top-Quarks verfügt über einen Wtb Produktionsvertex. Der Wirkungsquerschnitt dieser Prozesse ist deshalb proportional zum Quadrat des $|V_{\mathrm{tb}}|$-Matrixelements der Cabibbo-Kobayashi-Maskawa-Matrix (CKM-Matrix), was eine direkte Messung von $|V_{\mathrm{tb}}|$ ermöglicht. Bisher kann $|V_{\mathrm{tb}}|$ nur indirekt, unter Einbeziehung der Unitaritätsbedingung für die CKM-Matrix, zu $|V_{\mathrm{tb}}| \approx 1$ bestimmt werden. Für Modelle, welche beispielsweise eine weitere Quarkgeneration favorisieren, wäre $|V_{\mathrm{tb}}|$ aber deutlich von 1 verschieden.
Die große Herausforderung an eine Analyse, die sich die Messung des Wirkungsquerschnitts für die Produktion einzelner Top-Quarks zum Ziel gesetzt hat, ist die Reduktion des immensen Untergrundes. Eine bewährte Methode hierfür ist die Identifikation so genannter b-Quark Jets: Im Standardmodell zerfallen Top-Quarks nach ca. $0.4 \cdot 10^{-24}$ Sekunden in ein W-Boson und ein Bottom-Quark. Aufgrund von Hadronisationsprozessen bildet das Bottom-Quark einen Jet von Hadronen aus, welcher üblicherweise als b-Quark Jet bezeichnet wird. Da in vielen Untergründen zur Produktion einzelner Top-Quarks keine b-Hadronen bzw. b-Quark Jets auftreten, kann durch eine effiziente Identifikation dieser Jets eine signifikante Reduktion des Untergrundes erreicht werden. Der für CDF-Analysen standardmäßig eingesetzte Algorithmus zur b-Quark Jet Identifikation nutzt die extrem lange b-Hadron Lebensdauer von ungefähr 1.6 ps aus. Hierzu sucht man in Jets nach den Spuren, die ihren Ursprung nicht am Primärvertex haben. Aus diesen Spuren versucht man schließlich einen so genannten Sekundärvertex zu rekonstruieren. Diese Sekundärvertizes stammen mit hoher Wahrscheinlichkeit vom Zerfall eines b-Hadrons und ermöglichen somit die Identifikation von b-Quark Jets. Leider ist auf diese Weise keine perfekte Identifikation möglich, da zum einen die Lebensdauer von c-Hadronen ebenfalls von Null verschieden ist, was zur Rekonstruktion sekundärer Vertizes in c-Quark Jets führen kann. Zum anderen können, durch die begrenzte Auflösung des Detektors, Spuren unzureichend rekonstruiert werden, was durchaus zur Fehlidentifikation von Jets aus leichten Quarks als b-Quark Jets führen kann.
In den Analysen zur Messung des Wirkungsquerschnitts der Produktion einzelner Top-Quarks werden nur Ereignisse selektiert, in denen das W-Boson aus dem Top-Quark Zerfall leptonisch zerfällt. Aufgrund der speziellen Ereigniskinematik werden weiterhin zwei oder drei hochenergetische Jets gefordert, von denen mindestens einer als b-Quark Jet identifiziert worden sein muss. Schnitte, wie zum Beispiel auf die fehlende Transversalenergie in einem Ereignis, führen zu einer weiteren Unterdrückung von Untergründen. Nach der Selektion der Einzel-Top-Quark Kandidaten beträgt das Verhältnis zwischen Signal- und Untergrundereignissen immer noch ungefähr 1/15, was den Einsatz multivariater Analysemethoden unumgänglich macht. Des Weiteren bestehen trotz der Forderung nach mindestens einem als b-Quark Jet identifizierten

Jet 48% des Untergrundes aus Ereignissen, in denen kein b-Hadron vorhanden ist. Dies ist hauptsächlich darauf zurückzuführen, dass die b-Quark Jet Identifikation nicht perfekt ist. Da der oben angesprochene Algorithmus nicht alle Eigenschaften, mit denen b-Quark Jets von c-Quark Jets und leichten Jets unterschieden werden können, ausnutzt kann die Identifikation noch verbessert werden.
Um eine verbesserte Trennung von Ereignissen mit und ohne b-Hadronen zu erreichen wurde der so genannte KIT-Flavor-Separator entwickelt, dessen aktuelle Version im ersten Teil der vorliegenden Arbeit vorgestellt wird. Es handelt sich hierbei um ein Neuronales Netz, welches neben der langen Lebensdauer weitere b-Hadron Eigenschaften, wie ihre Masse von ungefähr 5 GeV/c^2, ihre hohe Zerfallsmultiplizität und die Wahrscheinlichkeit semileptonisch zu zerfallen, ausnutzt, um b-Quark Jets von c-Jets und Jets aus leichten Quarks zu unterscheiden. Insgesamt wurden 54 mögliche Eingabevariablen für das Neuronal Netz betrachtet, von denen schließlich 31 zum Trainieren des Netzwerks benutzt werden. Vier dieser Variablen werden in der aktuellen Flavor-Separator-Version zum ersten Mal benutzt. Die Ausgabe des Neuronalen Netzes ist kontinuierlich zwischen -1 und 1 verteilt, wobei sich b-Quark Jets bei 1 und c-Jets bzw. leichte Jets in der Nähe von -1 anhäufen. Es ist nun möglich durch einen Schnitt auf diese Ausgabeverteilung eine sehr reine b-Quark Jet Datensatz zu erhalten. Für die Messung des Wirkungsquerschnitts der Produktion einzelner Top-Quarks hat es sich jedoch als vorteilhaft erwiesen, die kontinuierliche Ausgabe als Eingabevariable weiterer multivariater Analysen zu benutzen. Da die Ausgabeverteilung des Neuronalen Netzes nicht auf den der Produktion des b-Quark Jets zugrunde liegenden physikalischen Prozess sensitiv ist, ist der KIT-Flavor-Separator universell in Analysen von Ereignissen mit hohem Transversalimpuls, die auf die Identifikation von b-Quark Jets angewiesen sind, einsetzbar.
Da das Neuronale Netz mit simulierten Ereignissen (auch Monte-Carlo Ereignisse genannt) trainiert wurde, muss überprüft werden, ob die Simulation gemessene Ereignisse richtig beschreibt. Dies ist notwendig, da zum Beispiel die Fehlmodellierung einer der Eingabevariablen zu einer Ausgabe des Neuronalen Netzes führen kann, die die Realität nicht mehr korrekt beschreibt. Für diese Überprüfung werden zunächst die Verteilungen der Eingabevariablen von simulierten und gemessenen Ereignissen verglichen. Hierauf werden auch die Ausgabeverteilungen des Neuronalen Netzes für b-Quark Jets und leichte Jets mit gemessenen Daten verglichen. Kleine Unterschiede in der simulierten und gemessenen Verteilung leichter Jets werden mit einer Korrekturfunktion behoben.
Neben den bereits erwähnten Einsatzgebieten kann der KIT-Flavor-Separator auch zur Bestimmung der Flavorzusammensetzung eines gegebenen Datensatzes eingesetzt werden. Aktuell wird dies zur Korrektur der Untergrundabschätzung für die Einzel-Top-Quark Analyse eingesetzt. Hierzu werden die Ausgabeverteilungen für b-, c- und leichte Jets an die Datenverteilung eines Seitenbandes angepasst und Korrekturfaktoren bestimmt, mit denen die Vorhersage für die jeweiligen Prozesse korrigiert wird. In der vorliegenden Arbeit wird eine Anpassung an eine Datenmenge von 2.2 fb^{-1} gezeigt. Gerade für die zuletzt genannte Anwendung wäre es wünschenswert, eine bessere Trennung der Ausgabe von c-Jets und leichten Jets zu erreichen. Zu diesem Zweck wurde

ein weiteres Neuronales Netz, das c-Netz, trainiert, welches eine hervorragende Separierung dieser beiden Jetflavor ermöglicht. In der c-Netz − Flavor-Separator Ebene häufen sich die drei Jetflavor in räumlich voneinander getrennten Bereichen an. Durch eine zweidimensionale Anpassung an die Datenverteilung des Seitenbandes können die Fehler auf die obige Anpassung um die Hälfte reduziert werden. Da für die Validierung der c-Netz Ausgabeverteilung ein möglichst reiner c-Quark Jet Datensatz selektiert werden muss, und dies leider äußerst schwierig zu realisieren ist, konnte das c-Netz bisher noch nicht in einer "offiziellen" Analyse eingesetzt werden.
Der zweite Teil der vorliegenden Arbeit beschäftigt sich mit der Kombination dreier multivariater CDF Analysen zur Messung des Wirkungsquerschnitts der Produktion einzelner Top-Quarks in einer Datenmenge von 2.2 fb^{-1}. Zum einen wird in dieser Studie gezeigt, inwiefern Neuronale Netze zur Kombination von hochkorrelierten Einzelanalysen geeignet sind, zum anderen wird eine Messung des Wirkungsquerschnitts der Produktion einzelner Top-Quarks durchgeführt und hieraus ein Wert für $|V_{\mathrm{tb}}|$ bestimmt.
Drei unterschiedliche CDF Analysen zur Suche nach einzelnen Top-Quarks wurden in der Kombination verwendet: Eine Methode beruht auf einer Likelihood Funktion, eine weitere ist eine so genannte Matrixelement-Methode. Die dritte Analyse benutzt ein Neuronales Netz zur Unterscheidung von Signal- und Untergrundprozessen. Alle drei Analysen greifen auf dieselbe Kandidatenselektion zurück, berücksichtigen die gleichen Quellen systematischer Unsicherheiten und benutzen dieselbe statistische Interpretation zur Auswertung der Messdaten. Dies vereinfacht eine Kombination der Analysen deutlich.
Die in dieser Arbeit vorgestellte Methode benutzt nun die Diskriminanten der drei Analysen als Eingabe zu einem Neuronalen Netz, welches eine weitere Trennung von Signal und Untergrund erreicht. Insgesamt werden vier Netze trainiert, da in vier verschiedenen Kategorien von Ereignissen nach einzelnen Top-Quarks gesucht wird. Dies sind Ereignisse mit zwei bzw. drei Jets, von denen einer bzw. zwei als b-Quark Jet identifiziert wurden. Die Ausgabeverteilungen dieser vier Netze werden mit Hilfe einer Likelihood Funktion gleichzeitig an die Datenverteilung angepasst. In dieser Anpassungsrechung wird der Wirkungsquerschnitt für die Produktion einzelner Top-Quarks zu $\sigma_{\mathrm{st}} = 2.2^{+0.8}_{-0.7}$ pb bestimmt. Der zentrale Wert liegt damit unter dem vom Standardmodell vorausgesagten, ist innerhalb der angegeben Unsicherheiten aber mit diesem kompatibel. Die Sensitivität der Kombination, die mit Hilfe von Ensembletests bestimmt wurde, liegt mit 4.8σ etwas über der beobachteten Signifikanz von 3.9σ, die schließlich in Daten gefunden wurde. Insgesamt kann somit eine Verbesserung von 8% im Vergleich zu den einzelnen Analysen erzielt werden. Aufgrund der schon angesprochenen Proportionalität zwischen dem gemessenen Wirkungsquerschnitt und $|V_{\mathrm{tb}}|^2$ kann $|V_{\mathrm{tb}}| = 0.88^{+0.14}_{-0.12}(\mathrm{exp.}) \pm 0.07(\mathrm{theo.})$ bestimmt werden. Wie erwartet sind die Unsicherheiten auf diesen Wert zu groß, um eine abschließende Aussage über die Gültigkeit der Unitarität der CKM-Matrix treffen zu können.

Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

IEKP-KA/2008-17

B-Jet and C-Jet identification with neural networks as well as combination of multivariate analyses for the search for single top-quark production

Manuel Renz

Diploma Thesis

at the faculty for physics of
the university of Karlsruhe (TH)

Referee: Priv.-Doz. Dr. Wolfgang Wagner
Co-Referee: Prof. Dr. Thomas Müller

Institut für Experimentelle Kernphysik

June 2008

Introduction

Until the end of the 19th century, atoms have been considered to be the smallest components natural matter is built of. With the knowledge that atoms are built up of electrons surrounding a nucleus this view has rapidly changed. In the following decades many discoveries, like the observation of protons and neutrons or the detection of quarks, have led to a deeper understanding of the physical processes which are central to the formation of matter.
Today the fundamental particles and the forces between them are described by the Standard Model (SM) of particle physics which was developed in the 1970s. The search for the Higgs boson, the last Standard Model particle which has not been discovered so far, and the check of various predictions made by the Standard Model are subjects of the current research. For this purpose large particle accelerators, colliding e^+e^-, $p\overline{p}$ or pp-pairs, are built. At the moment, the proton-antiproton collider Tevatron at Chicago is the collider with the world's highest center of mass energy (1.96 TeV). When the Large Hadron Collider at CERN begins to operate this year, center of mass energies up to 14 TeV can be reached.
In 1995 the heaviest known quark, the top quark, was discovered by the CDF and the DØ collaborations at the Tevatron collider. With a mass of 172.6 ± 1.4 GeV/c^2 the top quark is almost as heavy as a gold nucleus. At the Tevatron, top quarks are generally produced in pairs via the strong interaction. Nevertheless the production of single top-quarks via the weak interaction is also predicted by the Standard Model, but its cross section is roughly 2.5 times smaller than the $t\overline{t}$ one. A precise measurement of the single-top cross section would enable a interesting test of the SM, since it provides a direct measurement of the Cabbibo-Kobayashi-Maskawa matrix element $|V_{\mathrm{tb}}|$. The SM predicts $|V_{\mathrm{tb}}| \approx 1$, however models including a fourth quark generation predict $|V_{\mathrm{tb}}|$ to be significantly smaller.
The main challenge for a single-top analysis at CDF is the reduction of the tremendous backgrounds. In the Standard Model, top quarks decay with a probability of nearly 100% into a W boson and a bottom quark. Due to hadronization processes b quarks form so-called b jets. Since roughly 99% of the backgrounds to single-top production do not contain any b quark, identifying b jets would significantly reduce the large background. CDF's standard b jet identification algorithm, also called b-tagging algorithm, tries to reconstruct secondary vertices in jets by taking advantage of the long b-hadron lifetime (≈ 1.6 ps). Unfortunately, a large amount of events containing c-quark jets or light-quark jets have false reconstructed secondary vertices and thus are identified

as b jets. This is due to the non-zero lifetime of c hadrons and due to the spurious reconstruction of secondary vertices within light-quark jets. The KIT Flavor Separator is a neural network which further detaches tagged b jets from tagged non-b events with high efficiency. In addition to the b-hadron lifetime, it exploits further b-hadron properties like their mass of nearly 5 GeV/c^2, their high decay multiplicity and their probability to decay semileptonically. In this diploma thesis, we show the improved KIT Flavor Separator and its application in the search for single top-quarks.
For some applications, additionally a good separation between c-quark and light-quark jets is useful. Thus, we report the training of an additional neural network which is able to distinguish between tagged c-quark and tagged light-quark jets. Furthermore an example for its application is given.
Currently there are three different multivariate single-top analyses at CDF, namely a Likelihood Function method, a Matrix Element method and a Neural Network analysis. These analyses use the same single-top candidate selection, the same signal and background estimate and the same treatment of systematic uncertainties. Hence, the discriminants of the different analyses can be combined to one super-discriminant. In this thesis we present a combination method which uses a neural network to distinguish between single-top and background events. In order to measure the single top-quark production cross-section, the output distributions for simulated single-top and background events are fitted to the distribution observed from CDF data. Furthermore the statistical sensitivity of the combination is quoted and a value for $|V_{\mathrm{tb}}|$ is calculated.

Contents

1 Theoretical Background **1**
1.1 Main ideas of the standard model . . . 1
1.2 The top quark . . . 5
1.2.1 Top quark pair production at hadron colliders . . . 5
1.2.2 Single-top quark production . . . 6

2 Experimental Setup **10**
2.1 The Accelerator Chain . . . 11
2.1.1 Luminosity . . . 12
2.2 The CDF II Detector . . . 14
2.2.1 Tracking System . . . 14
2.2.2 Calorimetry . . . 16
2.2.3 Muon Chambers . . . 17
2.3 The Trigger System . . . 17

3 Experimental techniques **20**
3.1 Neural networks . . . 20
3.1.1 The NeuroBayes® package . . . 21
3.2 Monte Carlo generators and detector simulation . . . 25
3.2.1 Monte Carlo event generation . . . 25
3.2.2 Detector simulation . . . 26
3.3 Track and jet reconstruction . . . 27
3.3.1 Reconstruction of tracks . . . 27
3.3.2 Jet clustering algorithm . . . 28
3.3.3 Jet energy correction . . . 29
3.3.4 Reconstruction of secondary vertices . . . 29

4 The KIT Flavor Separator **34**
4.1 Training of the neural network . . . 35
4.1.1 Event selection and used MC-Samples . . . 35
4.1.2 Discriminating variables . . . 35
4.1.3 Network training . . . 39
4.2 Comparison of MC-Events and measured data . . . 44
4.2.1 Utilized samples . . . 44

4.2.2 Comparison of input distributions 47
4.2.3 Validation of the network output for b-quark jets 49
4.2.4 Correction of the mistag output 52
4.3 Application in analyses 58
4.3.1 Estimation of heavy flavor contents 58
4.4 A neural network to identify c-quark jets 60

5 The single-top analyses 64
5.1 Event selection and utilized data samples 64
5.1.1 Data samples 65
5.1.2 Requests for single-top candidates 65
5.1.3 Backgrounds to single-top production 68
5.2 Signal and background estimate 69
5.2.1 Expected single top-quark events 71
5.2.2 Estimation of expected background events 71
5.3 Multivariate analyses for the single-top search 72
5.3.1 The Likelihood Function method 72
5.3.2 The Matrix Element method 75
5.3.3 The Neural Network method 77

6 Combination of single-top analyses 80
6.1 Statistical interpretation 80
6.1.1 Measurement of the single-top production cross section 80
6.1.2 Ensemble tests and sensitivities 83
6.2 Training of the neural networks 84
6.3 Systematic uncertainties 90
6.4 Results of the individual analyses 101
6.5 Combination results 103
6.5.1 Dertermination of the sensitivity 103
6.5.2 Application to observed events 103
6.5.3 $|V_{\mathrm{tb}}|$ measurement 106

7 Summary and Outlook 108

A 110
A.1 $|\eta|$ dependences for pass 1 vertices 110
A.2 $|\eta|$ dependences for pass 2 vertices 112

Chapter 1

Theoretical Background

Since the 1970s, the standard model of elementary particle physics has been very successful in describing the properties of elementary particles and their interactions, which are the electromagnetic, the strong and the weak force. Gravitation, the fundamental force we know from every days life, is not included but explained by the theory of general relativity [1]. In the following we give a short overview of the main ideas of the standard model (SM) and then concentrate on the physics of the top quark, especially the elektroweak production of single top quarks which is one of the main aspects of this thesis.

1.1 Main ideas of the standard model

The standard model is a quantum field theory that is consistent with quantum mechanics and special relativity. Quantum numbers, like the electromagnetic charge or the spin, are used to describe the quantum state of any particle.
According to the SM, all matter is made up of twelve fundamental *fermions*, shown in table 1.1. These particles carry spin $s = \frac{1}{2}$ and therefore obey the Pauli exclusion principle, which states that particles with half integer spin cannot occupy the same quantum state, i.e. not all quantum numbers of two fermions can have the same values at the same time. As can be seen in table 1.1, the fermions are arranged in three generations. Particles of the first generation form most of the natural matter we know, higher-generation particles, however, can only be observed in high energy interactions since most of them decay very quickly to particles of the first generation.

Furthermore, the twelve fermions are divided into six quarks and six leptons. The electron and its heavy "brothers" muon and tau belong to the leptons and carry electric elementary charge. Each of these particles has an neutral partner, the neutrino. Neutrinos are also leptons and occur in three different types: $\nu_{\mathrm{e}}, \nu_{\mu}, \nu_{\tau}$. The standard model predicts the neutrinos to be massless but measurements during the last decade showed that there are oscillations between the different neutrino types [4]. However, such oscillations are only possible if neutrino masses are not exactly zero and so the SM has to be extended for massive neutrinos. The neutrino masses listed in table 1.1

name	fermion-type	symbol	el. charge[e]	mass [MeV/c^2]
1.generation:				
electron	lepton	e	-1	0.51
electron neutrino	lepton	ν_e	0	$\leq 2 \cdot 10^{-6}$
up quark	quark	u	$+\frac{2}{3}$	$1.5 - 3.0$
down quark	quark	d	$-\frac{1}{3}$	$3 - 7$
2.generation:				
muon	lepton	μ	-1	105.66
muon neutrino	lepton	ν_μ	0	≤ 0.19
charm quark	quark	c	$+\frac{2}{3}$	$(1.25 \pm 0.09) \cdot 10^3$
strange quark	quark	s	$-\frac{1}{3}$	95 ± 25
3.generation:				
tau	lepton	τ	-1	1776.9
tau neutrino	lepton	ν_τ	0	≤ 18.2
top quark	quark	t	$+\frac{2}{3}$	$(172.6 \pm 1.4) \cdot 10^3$
bottom quark	quark	b	$-\frac{1}{3}$	$(4.20 \pm 0.07) \cdot 10^3$

Table 1.1: The standard model fermions with their electric charge, given in units of the electron charge, and their mass [2]. The top quark mass is based on recent Tevatron results published in [3].

are limits given by direct measurements but cosmological models predict the sum of all neutrino masses to be in the order of 0.1 eV.
Quarks carry fractional electric charges, of $+\frac{2}{3}|e|$ or $-\frac{1}{3}|e|$ and appear in the following six different types (or flavour): up, down, charm, strange, top and bottom. Owing to confinement, a process that will be discussed later, quarks cannot be observed as free particles but only in bounded quark-antiquark (meson) or three-quark (baryon) states. Because of the existence of baryons, consisting of three identical quarks, and the Pauli exclusion principle, quarks have to carry a further quantum number called color which appears in three different types: red, green and blue. Antiquarks carry anticolor (anti-red, anti-green and anti-blue). All observable particles are colorless (or white), i.e. in baryons each of the three colors has to be carried by one of the quarks and the colors of quarks in mesons are for example red and anti-red.
An antiparticle corresponds to each of the above mentioned fermions. These were first predicted by Dirac in 1931 and with the discovery of the positron, the antiparticle of the electron, in 1932 a first experimental validation could be made. All the mentioned particles are predicted by the standard model nevertheless it makes no statement on their masses. Hence, the fermion masses enter the SM as free parameters.
In the standard model the interactions between the fundamental fermions are described by gauge theories. These theories predict massless, so-called gauge-bosons which mediate the three already mentioned forces described by the SM. The gauge-bosons are particles with spin quantum number $s = 1$. Electromagnetic forces are mediated by exchange of massless and chargeless photons. For example, they are responsible for the

name	symbol	force	el. charge[e]	mass [MeV/c^2]
photon	γ	electromagnetic	$\leq 5 \cdot 10^{-30}$	$\leq 6 \cdot 10^{-23}$
gluon	g	strong	0	0
W boson	$W^{\pm}$	weak	± 1	80.403 ± 0.029
Z boson	Z^0	weak	0	91.188 ± 0.002

Table 1.2: Properties of the standard model gauge-bosons. The electric charge and the masses are taken from reference [2].

bound states of electrons with nuclei and the intermolecular forces in liquids and solids. Theoretically, the electromagnetic interaction is described by the theory of Quantum Electro Dynamics (QED). The strong force is described by Quantum Chromo Dynamics (QCD) which predicts massless gluons, coupling to the color charge and carrying color themselves. Gluon self-interactions are therefore possible and thereby the different strength of the strong force, which is small for short distances (asymptotic freedom) and increases for large distances (confinement), can be explained. The strong interaction is responsible for the binding of quarks to baryons or mesons and also for the formation of nuclei out of protons and neutrons. An example for the weak interaction is the β-decay of a radioactive nucleus. The mediators of the weak force are the $W^{\pm}$ and the Z^0 bosons, which have been found to be very massive (roughly 100 times the proton mass). A mechanism to "create" boson masses will be explained later in the text. The properties of the mentioned gauge-boson are summarized in table 1.2.

In the following we give a more detailed description of phenomena concerning the weak interaction. First of all, W bosons only couple to the left-handed (right-handed) part of the wave function of a particle (antiparticle). A wave function ψ is called left-handed (LH) if $\gamma_5\psi = -\psi$, it is called right-handed (RH) if $\gamma_5\psi = \psi$. Wave functions of massive particles, are superpositions of LH and RH parts. The weak interaction is furthermore the only force which acts on neutrinos and also the only possibility for changing a quark's flavor is the exchange of a W boson. Transitions between quark flavors of different generations are possible, because the weak eigenstates of the quarks (d', s', b') are not the same as the mass eigenstates (d, s, b). Transformations between the two bases are described by a unitary 3×3 matrix, the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix:

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

The matrix elements $V_{q_\mathrm{i} q_\mathrm{j}}$ represent the strength of the coupling between the quarks q_i, q_j and the W-boson. The following values have been measured for the elements of

the CKM-matrix [2]:

$$V_{CKM} = \begin{pmatrix} 0.97383^{+0.00024}_{-0.00023} & 0.2272 \pm 0.001 & (3.96 \pm 0.09) \cdot 10^{-3} \\ 0.2271 \pm 0.001 & 0.97296 \pm 0.00024 & (42.21^{+0.1}_{-0.8}) \cdot 10^{-3} \\ (8.14^{+0.32}_{-0.64}) \cdot 10^{-3} & (41.61^{+0.12}_{-0.78}) \cdot 10^{-3} & 0.999100^{+0.000034}_{-0.000004} \end{pmatrix} \tag{1.2}$$

In the late 1960s Weinberg, Salam and Glashow described how it should be possible to treat the electromagnetic and the weak interaction as different aspects of a single electroweak force [5, 6, 7]. This unification appears for high energies, for low energies the symmetry is broken which explains the different behaviour of the electromagnetic and the weak interaction. Up to now, the most favoured mechanism for breaking the elektroweak symmetry is the so-called Higgs mechanism [8, 9], which explains also the masses of the weak gauge-bosons. Higgs introduced another scalar field called Higgs field whose quantum is the Higgs boson, the only standard model particle that hasen't been discovered so far. All fundamental particles, including the Higgs boson itself, get their masses by interacting with this Higgs field.

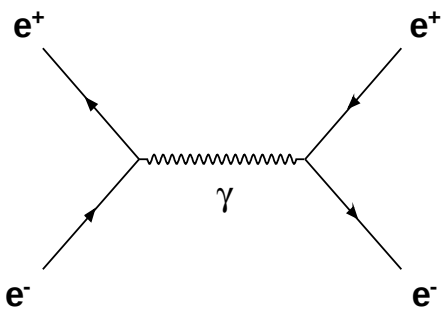


Figure 1.1: Feynman diagram of electron-positron scattering.

Feynman graphs are useful symbolic notations of physical processes. Figure 1.1 shows the annihilation of a electron and positron to a virtual photon and the following recreation of an electron positron pair. Due to convention the x-axes represents the time coordinate and the y-axes the spacial coordinate. Straight lines represent the space-time propagation of fermions while bosons are illustrated by wavy lines. Vertices stand for interaction points and contain the coupling nature and strength. By using the so-called Feynman rules, all parts of Feynman diagrams can be translated into mathematical expressions which allow the calculation of the matrix element of the corresponding interaction. The cross section, a measure for the probability of the process, can be calculated by integrating the matrix element over all possible initial and final states.

1.2 The top quark

The top quark belongs to the third fermion generation and is the heaviest fundamental particle discovered so far. Recent measurements of the CDF- and the DØ-collaboration found its mass to be $m_{\mathrm{top}} = 172.6 \pm 1.4$ GeV/c^2 [3]. With a predicted lifetime of $\tau \approx 0.4 \cdot 10^{-24}$ s, the top quark decays before any hadronization process starts. It is therefore the only quark which passes all its properties to its decay products. Since $|V_{\mathrm{tb}}| \approx 1$ (see eq. 1.1, 1.2), these decay products are almost always a W boson and a bottom quark. Below we will give a short overview of top quark production via the strong and the electroweak force. A more detailed review of top quark physics in hadron collisions can be found in [10].

1.2.1 Top quark pair production at hadron colliders

In this subsection we will focus on $\mathrm{t\bar{t}}$ production by the strong interaction, since the cross sections for electroweak top-pair productions are completely negligible for hadron collisions. For the calculation of the $\mathrm{t\bar{t}}$ cross section a factorization ansatz is made:

$$\sigma(AB \to t\bar{t}) = \sum_{i,j} \int f_{i,A}(x_i, \mu^2) f_{j,B}(x_j, \mu^2) \cdot \hat{\sigma}_{ij}(ij \to t\bar{t}; \hat{s}, \mu^2)\, dx_i dx_j \ . \tag{1.3}$$

A and B are hadrons consisting of nearly-free quarks and gluons (partons). The fraction of a hadron's longitudinal momentum a specific parton $i(j)$ is carrying, is represented by $x_{\mathrm{i}} = p_{\mathrm{i}}/p_{\mathrm{H}}$. $\hat{s}$ is the square of the center-of-mass energy of the colliding partons. The so-called renormalization scale μ describes the typical energy scale of the considered interaction and is often set to $\mu = m_{\mathrm{top}}$. $\hat{\sigma}_{\mathrm{ij}}$ represents the hard parton-parton cross section which can be calculated in perturbative QCD. The Feynman diagrams of the two dominant $\mathrm{t\bar{t}}$ production processes at a proton-antiproton collider, namely quark-antiquark annihilation and gluon-gluon fusion, are shown in figure 1.2.

For the Tevatron, a proton-antiproton collider with a center-of-mass energy of 1.96 TeV, next-to-leading order (NLO) calculations predict the fraction of the annihilation process to be 85 percent of the whole cross section. In eq. 1.3 the parton-parton cross section is folded with $f(x, \mu^2)$ which represents the parton distribution functions (PDF) for the hadrons A and B. For example a proton consists of two u and one d valence quark which are bound by the exchange of gluons. Gluons can split into quark-antiquark pairs, called see-quarks. Each of these proton constituents could be the interacting parton described in eq. 1.3. A PDF now gives the probability of finding the parton i with a longitudinal momentum fraction x_i within a proton. Figure 1.3 shows the CTEQ5L proton PDF parametrization for $\mu^2 = (175\ \mathrm{GeV})^2$. For a top quark mass of $m_{\mathrm{top}} = 175$ GeV/c^2, the $\mathrm{t\bar{t}}$ production cross section is predicted to be $\sigma_{\mathrm{t\bar{t}}} = 6.70^{+0.71}_{-0.88}$ pb [11].

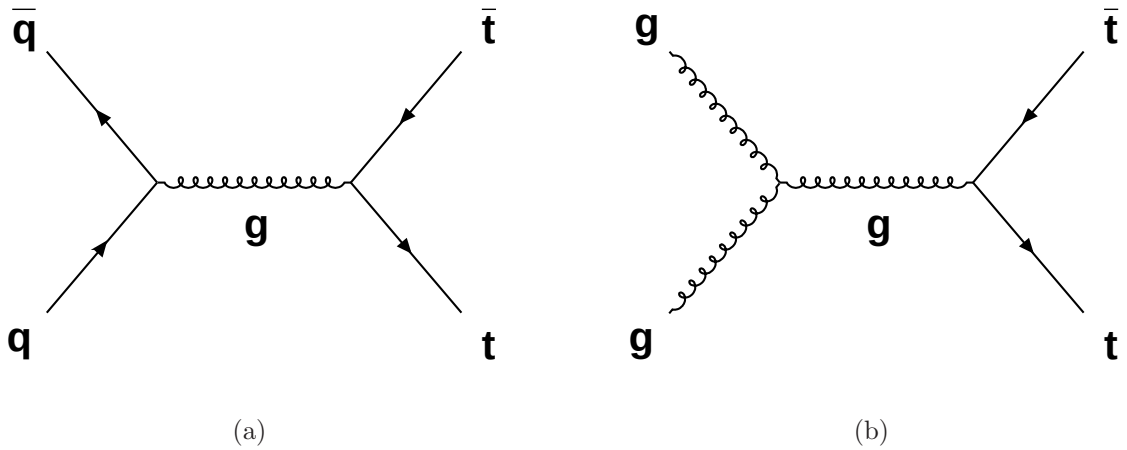


Figure 1.2: Leading order diagramms for top-pair production at the Tevatron: (a) quark-antiquark annihilation, (b) gluon-gluon fusion.

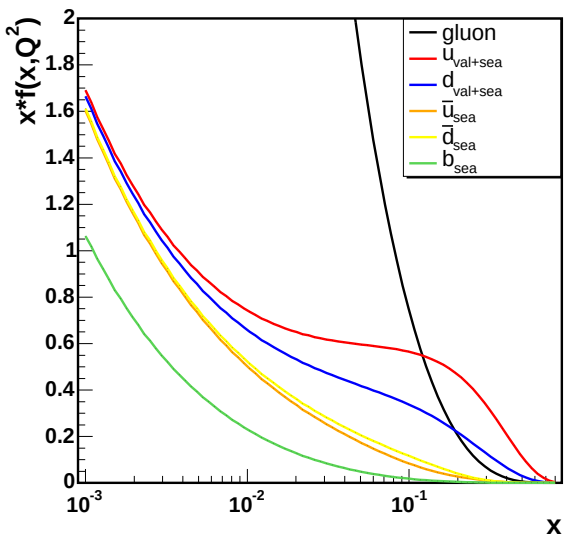


Figure 1.3: The CTEQ5L proton PDF parametrization.

1.2.2 Single-top quark production

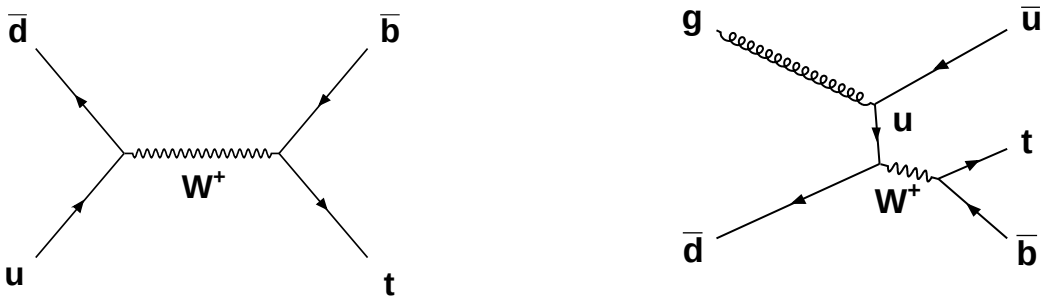
In the standard model, top quarks can also be produced singly by the electroweak interaction. There are three different production modes: the t-channel, the s-channel and the associated production. The first and the second get their names from the Mandelstam variable which gives the virtuality Q^2 and are the dominant production

channels at the Tevatron. Furthermore the different processes can be distinguished by the virtuality Q^2 of the participating W boson ($Q^2 = -q^2$, where q is the four momentum of the W). In the following we discuss the three production modes a bit more detailed:

- t-channel($q^2 = \hat{t}$)
  This channel is also named the W-gluon fusion. A virtual, space-like ($q^2 < 0$) W boson hits a bottom see-quark in a proton or an antiproton. The see-quark stems from a gluon splitting in a $b\bar{b}$ pair, thus the name W-gluon fusion. With a cross section of $\sigma = 1.98^{+0.28}_{-0.22}$ pb at a top quark mass of 175 GeV/c^2 [12, 13], the t-channel is predicted to be the dominant production mode at the Tevatron. The Feynman diagrams of the leading order (LO) and the most important NLO process with an inital state gluon splitting are shown in figure 1.4.

Figure 1.4: The single-top t-channel production mode for (a) leading order, and (b) for NLO with an initial gluon splitting.

- s-channel($q^2 = \hat{t}$)
  In the s-channel production mode a timelike W boson with $q^2 \geq (m_{\text{top}} + m_{\text{b}})^2$ is produced by the fusion of two quarks. Assuming again a top quark mass of 175 GeV/c^2 the s-channel cross section is $\sigma = (0.88 \pm 0.11)$ pb at the Tevatron collider [12, 13]. Figure 1.5 shows Feynman diagrams for the LO process as well as an example for an NLO process including initial state gluon splitting.



(a) (b)

Figure 1.5: The single-top s-channel production mode at (a) leading order, and (b) an example for NLO with an initial gluon splitting.

- associated production
  In this process, the top quark is produced in association with a real W boson ($q^2 \approx m_W^2$) whereas the initial bottom quark is a see-quark within a proton or an antiproton. At the Tevatron, the cross section for the associated production is very small, $\sigma = 0.094^{+0.015}_{-0.012}$ pb [14], and therefore negligible. For the proton-proton collisions at a center-of-mass energy of 14 TeV at the Large Hadron Collider (LHC), which is located near Geneva, Switzerland, this cross section will increase and even exceed the one of the s-channel. The leading order Feynman diagram of the associated production is shown in figure 1.6.

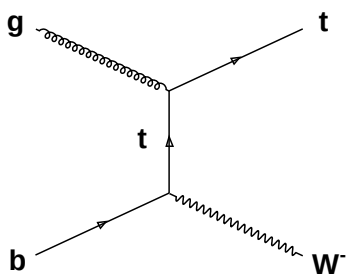


Figure 1.6: Leading order process for associated single-top production

Since all of the above mentioned production channels have a Wtb vertex, the corresponding cross sections are all proportional to the square of the CKM matrix element V_{tb}. Hence, measuring the single-top cross section allows for direct $|V_{\mathrm{tb}}|$. The value of V_{tb} is very interesting for the validation of the standard model and the exclusion of other models. In the standard model the CKM matrix is unitary, which leads, for example,

to the following unitary constraint: $|V_{\mathrm{td}}|^2 + |V_{\mathrm{ts}}|^2 + |V_{\mathrm{tb}}|^2 = 1$. This constraint enables an indirect measurement of $|V_{\mathrm{tb}}|$: $0.9990 < |V_{\mathrm{tb}}| < 0.9992$ at 90% confidence level (CL). The unitary assumption is mainly based on three experimental facts: First the measurements of V_{ub} and V_{cb} in B-meson decays shows, that the hierarchy of the elements belonging to the first two rows of the CKM matrix is in excellent agreement with the unitary condition. Furthermore the measurements of the $B_s^0 - \overline{B}_s^0$ oscillation frequency constraints the ratio of V_{td} and V_{ts} to $0.20 < |V_{\mathrm{td}}/V_{\mathrm{ts}}| < 0.22$ [15, 16]. The unitarity hypothesis predicts this ratio to be of order $\sin(\theta_{\mathrm{c}}) \simeq 0.22$ which is in very good agreement with the experimental results. Finally, the ratio $R \equiv |V_{\mathrm{tb}}|/\sqrt{|V_{\mathrm{td}}|^2 + |V_{\mathrm{ts}}|^2 + |V_{\mathrm{tb}}|^2}$ is measured to be larger than 0.6 at 95% CL [17, 18]. From the definition of R it is clear that $R \simeq 1$ implies a strong hierarchy between V_{tb} and the other two matrix elements, as expected in the unitary case.
If there is for example a fourth quark generation the 4×4 CKM matrix would be unitary too, but the 3×3 unitary constraint won't be fulfilled anymore. Theoretical calculations including different experimental results found the following limits for the "four-generation" CKM matrix elements V_{ti} : $|V_{\mathrm{td}}| \lesssim 0.03$, $|V_{\mathrm{ts}}| \lesssim 0.2$, $|V_{\mathrm{tb}}| \gtrsim 0.8$ [19]. Hence, a precise measurement of the single-top cross section is necessary to be able to distinguish between three and four quark generations. As it is shown in section 5.3 this will be hard to realize at the Tevatron. At the LHC, the uncertainty of V_{tb} from t-channel cross section measurements in 10 fb^{-1} is estimated to be 5% [19].

Chapter 2

Experimental Setup

Figure 2.1: Aerial view of the Fermilab area with the Main Injector in the foreground and the Tevatron ring in the back. The CDF-experiment is located at the ten o'clock position of the Tevatron.

The CDF II experiment is located at the Fermi National Accelerator Laboratory (Fermilab or FNAL) in Batavia, Illinois. About 60 km away from Chicago, the Fermilab is one of the centers for High-Energy-Physics in the United States of America. An overview of the 28 km^2 area can be seen in figure 2.1. Since the first high-energy particle beam at Fermilab in 1972, a lot of successful experiments have been conducted, which led for example to the discovery of the bottom quark in 1977. Measurements of the two experiments CDF and DZero (DØ) at Fermilab's proton-antiproton collider Tevatron, yielded the first observation of the top quark in 1995. In the following, we will give a short description of the Tevatron ring and the main components of the CDF

II detector.

2.1 The Accelerator Chain

Currently, the Tevatron is the accelerator with the highest center-of-mass energy in the world. A proton and an antiproton beam, each with an energy of 980 GeV, collide in two interaction regions with a center-of-mass energy of 1.96 TeV. As illustrated in figure 2.2 several different pre-accelerators are needed to reach such high beam energies.
The proton beam begins life as a pulsed, 25 keV negative hydrogen ion beam from a

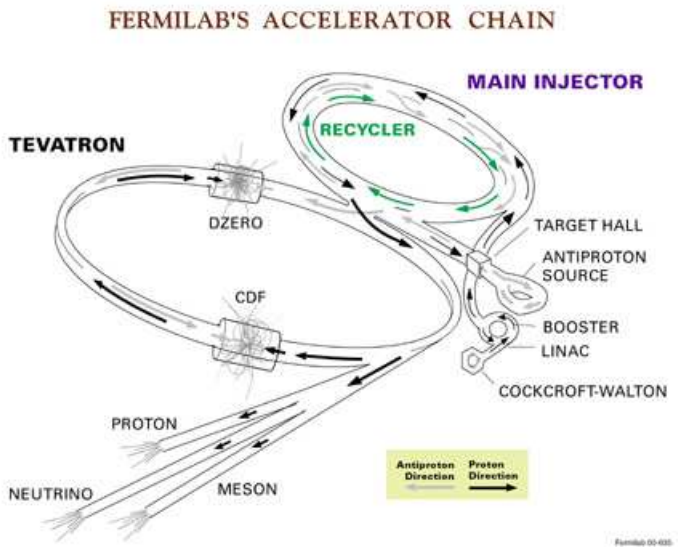


Figure 2.2: Overview of Fermilab's Accelerator Chain

magnetron surface-plasma source [20]. This type of source consists of an oval-shaped cathode surrounded by an anode. The approximately 1 mm wide gap between the two electrodes is filled with hydrogen gas which turns into a dense plasma by the impact of electric and magnetic fields. Positive hydrogen ions then hit the cathode, absorbe two electrons and are reflected. Caesium vapor coating the cathode's surface and lowering its work function, increases the probability for the mentioned process. After formation, some of the H^- ions are extracted through the anode aperture and are accelerated through an extraction plate.
The whole source is located in a metallic enclosure kept at -750 keV by a Cockcroft-Walton generator. While flying from the source to ground potential, the extracted H^- ions are accelerated to 750 keV and are then sent to the Linac, an 400 MeV linear accelerator. Here, the ions are accelerated in cavities by alternately oscillating electric fields. At the end of the Linac is a carbon foil which removes the electrons from the ions, leaving only the protons.
In the third step, the resulting proton beam enters the so-called Booster, a proton

synchrotron with a diameter of roughly 150 m. Unfortunately, the momentum spread of the Linac-proton beam is about 0.3% which possibly leads to a radius difference of nearly 8 mm in the Booster [20]. Thus a radio frequency (RF) cavity which decelerates too fast and accelerates too slow particles, is placed between the two accelerator stages. After spinning around 20,000 times in the Booster, protons have reached an energy of 8 GeV and enter the Main Injector. This second synchrotron has a radius of 524 meters and increases the beam energy from 8 GeV to 150 GeV. Futhermore, the Main Injector provides a 120 GeV proton beam for the Antiproton Source.

In the Antiproton Source the high energy proton beam hits a stack of nickel target and copper cooling disks whereby many secondary particles, especially antiprotons, are produced [21]. A cylindric lithium lense, which focusses the secondary particles along trajectories parallel to the cylinder axes, is placed directly behind the target. Negatively charged particles are then selected by a dipole magnet and sent to the Debuncher which has the shape of a "rounded triangle". The purpose of the Debuncher is to accept pulses of antiprotons from the source and to reduce their momentum spread through RF bunch rotation and adiabatic debunching from 4% to 0.2% [20]. Additionally, the phase space is reduced by stochastic cooling [22].
The beam then enters the Accumulator which is placed in the same tunnel as the Debuncher. Its task is to store antiprotons for several hours while they are cooled with different techniques like momentum or betatron cooling [21]. Thereby the physical size and energy spread of the particle beam is reduced without any accompanying beam loss. When enough antiprotons have been collected, they are sent to the Recycler which is installed in the same tunnel as the Main Ring. Here they are further cooled by stochastic and electron cooling processes and finally stored until they are shot into the Tevatron.

The Tevatron is the last part of Fermilab's accelerator chain. On their way around the 6.3 km long ring, protons and antiprotons, which are flying in opposite directions, are accelerated to 980 GeV. This is currently the highest particle energy ever reached with a collider. All magnets, that are guiding particles on their track, are superconducting coils cooled with liquid helium to nearly 4 K. The acceleration itself happens in RF-cavities. Therefore the particle beams have to be bunched. Finally, 36 bunches of protons and antiprotons, consisting of billions of particles, circulate with a revolution time of 20 μs.
The two beams are then forced to cross each other at two interaction regions where the two detectors, CDF and DØ, are ready to detect the products of the proton-antiproton collisions.

2.1.1 Luminosity

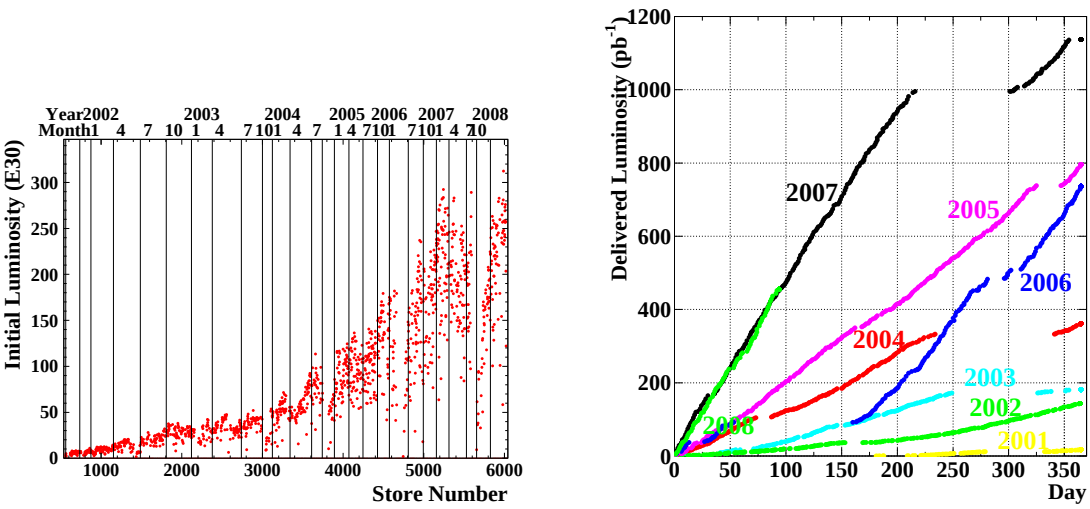
An imporant characteristic of an accelerator is its luminosity. The luminosity is given by

$$\mathfrak{L} = j \cdot \frac{v}{C} \cdot \frac{N_a \cdot N_b}{4\pi\sigma_x\sigma_y} \ , \tag{2.1}$$

where j is the number of bunches, v the velocity, C the circumference of the accelerator, N_a (N_b) the number of particles of type a (b) per bunch and σ_x (σ_y) the average bunch width. It is therefore a measure for the ability of an accelerator to produce collisions and is often given in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Frequently, the overall performance of storage rings is also represented by their integrated luminosity $\int \mathfrak{L}\, dt$, which is a measure for the amount of data collected. Hence, the number of events of a specific physiscs process can easily be calculated by multiplying its cross section, often given in units of b (1 b = 1 barn = 10^{-24} cm^2), with the integrated luminosity:

$$N_\mathrm{proc} = \sigma_\mathrm{proc} \cdot \int \mathfrak{L}\, dt \ . \tag{2.2}$$

Since the number of protons and antiprotons per bunch decreases with the number of collisions the maximum luminosity is reached at the beginning of each run (initial luminosity). In figure 2.3(a) the initial luminosity is shown as a function of the store number, where a store denotes the period between Tevatron beam initialization and abortion. The luminosity delivered by the Tevatron per calender year is illustrated in figure 2.3(b). During winter 2006/2007 the Tevatron achieved its design luminosity of $270 \cdot 10^{30}$ cm^{-2}s^{-1} . On March, the 17th 2008 the current record initial luminosity of $315,9 \cdot 10^{30}$ cm^{-2}s^{-1} was reached. Figure 2.4 compares the integrated luminosity



(a) (b)

Figure 2.3: (a): Initial Tevatron luminosities given as a function of time. (b): Luminiosity delivered by the Tevatron per year.

delivered by the Tevatron with the integrated luminosity which the CDF detector recorded. As can be seen, the Tevatron yielded approximately 3.8 fb^{-1} till April 2008 whereas CDF managed to write about 3.1 fb^{-1} to tape. The measurement of the luminosity at CDF is performed with low mass Cherenkov detectors [23].

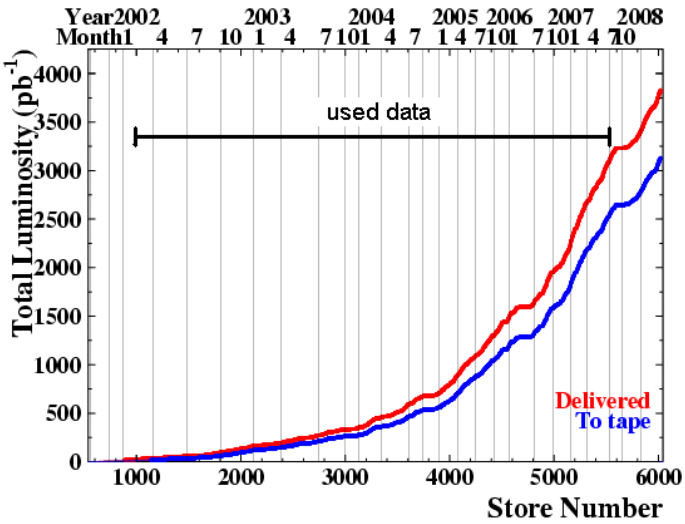


Figure 2.4: Integrated luminosity delivered by the Tevatron (red) and the luminosity recorded by CDF (blue) till April 2008. The black line represents the time period in which the data used for the single-top analyses have been taken.

2.2 The CDF II Detector

The Collider Detector at Fermilab (CDF) is one of the two detectors installed at Tevatron. It is an azimuthally and forward-backward symmetric general purpose solenoidal detector which combines precision charged particle tracking with fast projective calorimetry and fine grained muon detection [24].

Figure 2.5(a) shows a cutaway view of the CDF II detector. The tracking systems are enveloped by a superconducting solenoid, 1.5 m in radius and 4.8 m in length, which produces a 1.4 T magnetic field along the beam axes. All calorimetry and muon systems are located outside this solenoid. As can be seen in figure2.5(b) , the detector is described by a right-handed spherical coordinate system, where the z axis points along the proton beam. The azimuthal angle ϕ is measured from the Tevatron plain, and the polar angle θ is measured from the proton direction. One usually replaces the polar angle by the pseudorapidity η, which is defined as $\eta = -\ln(\tan\frac{\theta}{2})$. This is reasonable, because the production of relativistic particles is a constant as a function of rapidity.

2.2.1 Tracking System

The CDF Tracking System, which is responsible for a high momentum resolution, consists of two main parts: a barrel-shaped silicon system and an open cell drift chamber (COT) surrounding the first. Moreover the silicon system consists of three subsystems, namely the Layer00 [25], the Silicon VerteX detector (SVXII) [26] and the Intermediate Silicon Layers (ISL) [27]. The different components are illustrated in figure 2.6.

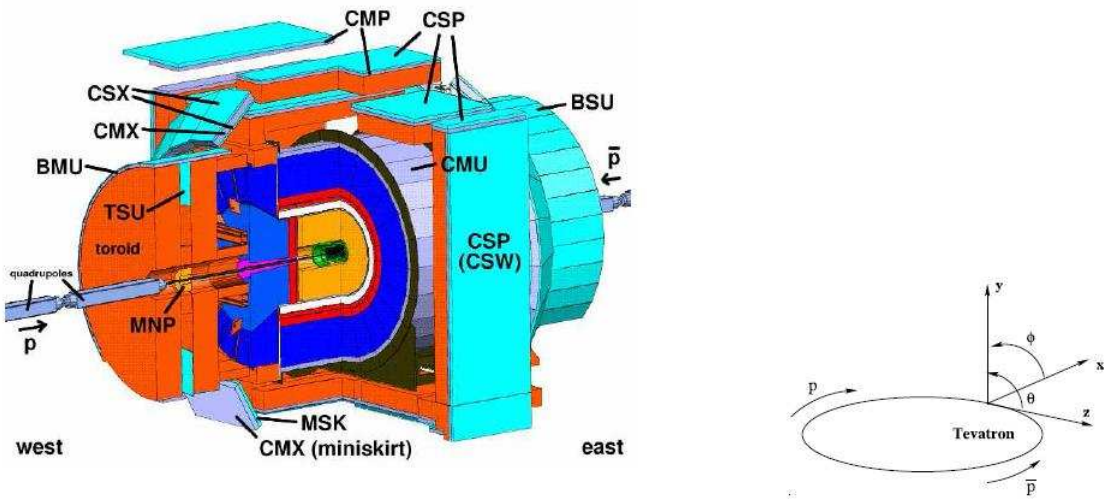


Figure 2.5: (a) Cutaway view of the CDFII detector with its different components. The inner green and orange parts belong to the tracking system, the red and blue ones to the calorimeters. The labeled objects are part of the muon system. (b) The CDFII coordinate system.

Layer00, added to the detector design in 1999, is a single-sided layer of silicon microstrips, which provides a full azimuthal coverage. With their location at radii of 1.35 cm and 1.65 cm the two sensors are very close to the beam pipe and responsible for a high impact parameter[1] resolution i.e. for high-quality track reconstruction. SVXII consists of five layers double-sided silicon wafers covering the region $|\eta| < 2$. On the front side, the strips are placed parallel to the z-axis, hence one has a high resolution in the $r - \phi-$plane. In contrast to this, the strips on the backside of three sensors are rotated by an angle of 90° with respect to beam-parallel strips and allow the measurement of a particle's z coordinate with high precision. The strips on the backside of the other two layers of sensors are only rotated by an angle of 1.2°. In the central region ($|\eta| < 1$), a single ISL layer is placed at a radius of 22 cm and in the forward-backward region, $1 \leq |\eta| \leq 2$, two further ISL layers are located at radii of 20 cm and 28 cm. SVXII and ISL together are a single functional system which provides stand-alone silicon tracking over the full $|\eta| \leq 2$ region. The whole CDF silicon system reaches an impact parameter resolution of 40μm.
In the central $|\eta| \leq 1$ region the silicon system is surrounded by the COT. This drift chamber is roughly 3 m long and covers radii between 44 cm and 132 cm. It is segmented in 8 super-layers alternating stereo and axial, with a stereo angle of $\pm 2°$. Each super-layer contains 12 sense wires alternating with 13 potential wires which provide the field shaping within the cell, yielding a total of 96 measurement layers. The COT is filled with a mixture of Argone-Ethane-CF_4(50:35:15) and the drift field is about 2.4 keV/cm which limits the maximum drift time to less than 100 ns. It provides track

[1]impact parameter: closest distance of a particle track to the primary interaction point

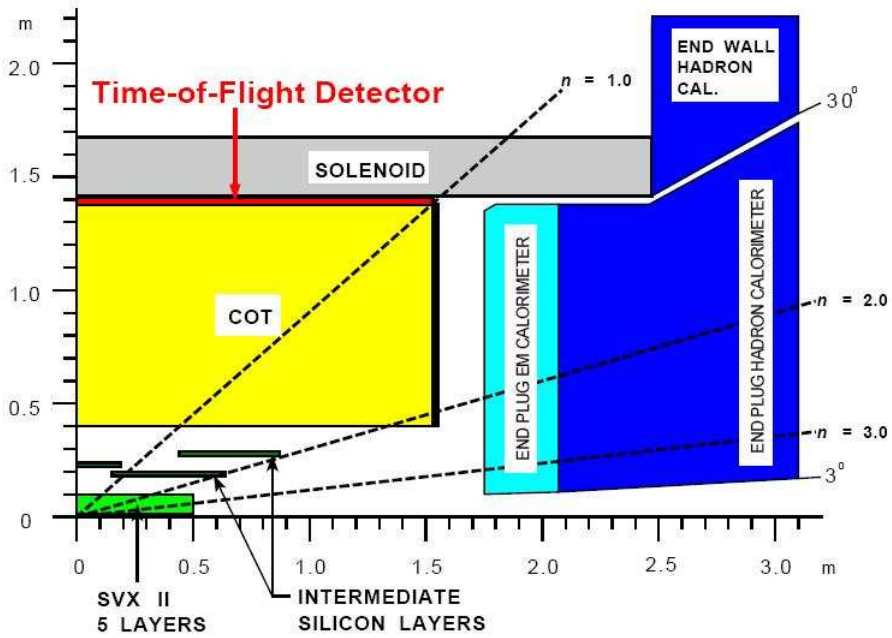


Figure 2.6: Longitudinal view of the CDF II tracking system and the plug calorimetry

momentum measurements with resolution $\frac{\sigma_{p_T}}{p_T} = 0.0015\ \mathrm{GeV}^{-1} \times p_T$.
In 2001 a Time-Of-Flight (TOF) detector has been installed just outside the COT. The motivation for this was to enhance the particle identification capability provided by the central drift chamber. With a time-of-flight resolution of nearly 100 ps, TOF is capable of identifying charged kaons from pions by their flight time differences with at least two standard deviation separation up to kaon momenta of 1.6 GeV/c.

2.2.2 Calorimetry

Outside the solenoid, scintillator-based calorimetry allows the measurement of particle and jet energies. Except of muons and neutrinos all particles are fully absorbed in different electromagnetic or hadronic calorimeters. The whole system consists of five parts: The Central ElectroMagnetic (CEM) calorimeter [28], the Central HAdron (CHA) [29] calorimeter, the end-Plug ElectroMagnetic (PEM) calorimeter, the end-Plug HAdron (PHA) calorimeter and the end-Wall HAdron (WHA) calorimeter. All of these calorimeters are segmented into towers pointing towards the center of the detector.
The electromagnetic calorimeters use layers of lead alternating with polystyrene scintillators (central region $0 \leq |\eta| \leq 1.1$) or plastic scintillators (forward-backward region $1.1 \leq |\eta| \leq 3.64$) as active detector medium. CEM has an average energy resolution, $\sigma(E)/\sqrt{E[\mathrm{GeV}]}$, of $13.5\%/\sqrt{E}$ as well as a position resolution of ± 2 mm at 50 GeV and PEM possesses a energy resolution of $16\%/\sqrt{E} \oplus 1\%$, where $\oplus$ represents the sum

	CMU	CMP/CSP	CMX/CSX	BMU
coverage	$\|\eta\| \leq 0.6$	$\|\eta\| \leq 0.6$	$0.6 \leq \|\eta\| \leq 1.0$	$1.0 \leq \|\eta\| \leq 1.5$
chamber area[cm^2]	2.68×6.35	2.5×15	2.5×15	2.5×8.4
chamber length[cm]	226	640	180	363
max. drift time[ns]	800	1400	1400	800
min. p_T[GeV/c]	1.4	2.2	1.4	$1.4 - 2.0$

Table 2.1: Design Parameters of the CDF II muon system

in quadrature, as well as a position resolution between 5 mm for 50 GeV and 1 cm at 5 GeV.
The central and the end-Wall hadronic calorimeters consist of steel absorbers and acrylic scintillators in contrast to PHA which uses iron absorbers. CHA and WHA have energy resolutions of $50\%/\sqrt{E} \oplus 3\%$ and $75\%/\sqrt{E} \oplus 4\%$ respectively, the end-Plug hadronic calorimeter however $80\%/\sqrt{E} \oplus 5\%$.

2.2.3 Muon Chambers

The key to many precision measurements and new discoveries at the high-energy frontier is the efficient detection of muons over a large fiducial and kinematic range. CDF's muon system is divided into four subdetectors. The Central MUon (CMU) detector [30] encloses the outside of the central hadron calorimeter at a radial distance of 3470 mm from the beam axis and covers the $|\eta| \leq 0.6$ range. It consists of 144 modules with 16 rectangle cells in each module. Every cell is further divided in four layers of drift chambers which are filled with a 50%/50% argon-ethane gas. The Central Muon uPgrade (CMP) consists of another set of muon chambers behind additional 60 cm of steel in the region $55° \leq \theta \leq 90°$ which form a box around the central detector. Additionally a layer of 216 scintillation counters (the CSP) covers the outside surface of the chambers. The Central Muon Extension, which consists of conical sections of drift tubes (CMX) and scintillation counters (CSX), is located at each end of the central detector and ranges in polar angle from 42° to 55°. At 55° it slightly overlaps the coverage provided by the CMU and extends its pseudorapidity coverage up to 1.0. The last part of the system is the Barrel MUon detection system (BMU), which covers the $1.0 \leq \eta \leq 1.5$ region. The heart of this detector is a barrel of CMP-like chambers and CSP-like scintillation counters. Table 2.1 lists some design parameters of the CDF II muon system and figure 2.5(a) shows its different parts. .

2.3 The Trigger System

As mentioned in section 2.1 the Tevatron beam is splitted in bunches. The average time between two particle bunches is 396 ns. Additionally the 36 bunches are subdivided into three trains, each containing 12 bunches, with an "empty space time" of roughly

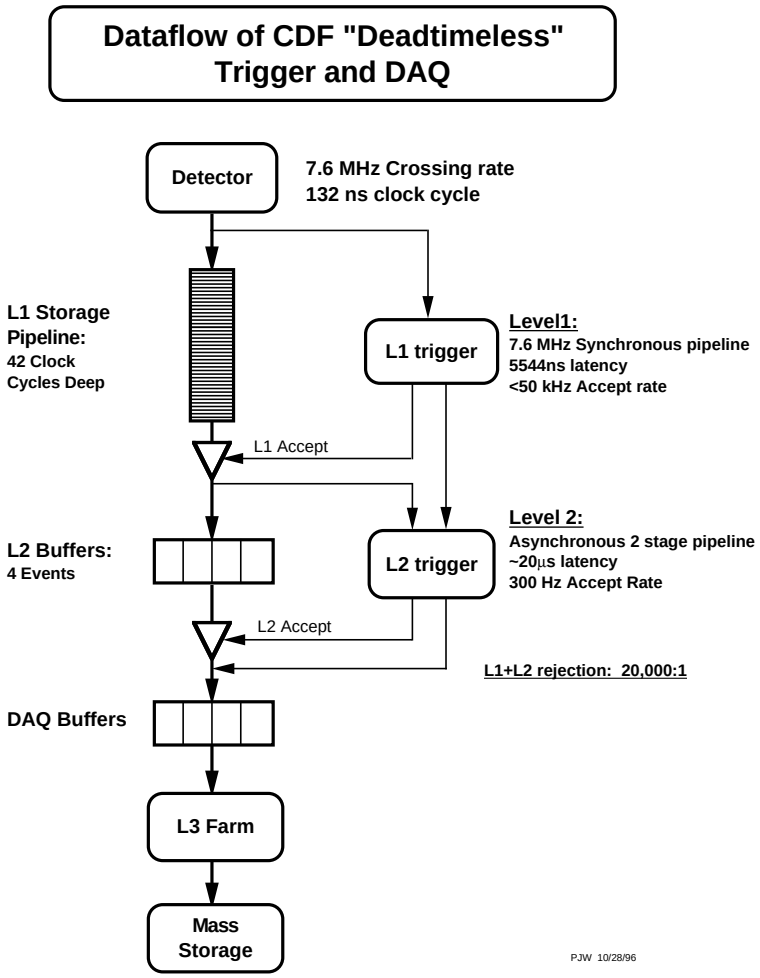


Figure 2.7: Overview of the CDF II data flow which shows the initial design rates. The expected bunch crossing rate of 7.6 MHz could not be reached because of beam stability criteria.

2.6 μs between two trains. Therefore the effective bunch crossing rate is approximately 1.7 MHz. Since the tape writing speed is limited to an event rate of 100 Hz one has to decide for every event, whether it is one of the rare interesting or one of the numerous background events. This decision is made by a three-level trigger system [31], which is illustrated in figure 2.7. Level-1 (L1) is a purely hardware trigger, Level-2 (L2) is a combination of hardware and software triggers and Level-3 (L3) is a software trigger, running on a Linux computer farm.

L1 uses custom designed hardware to find physics objects based on a subset of the detector information and makes a decision based on simple counting these objects. The following three systems are used to evaluate the events: The calorimeter trigger boards are searching for electrons, photons and jets by applying thresholds to individual calorimeter trigger towers. They are also senitive to an events's total transverse energy and missing transverse energy by applying thresholds after summing over all

towers. Muon trigger cards use hits in the muon chambers to identify single and dimuon objects. Finally, the eXtremely Fast Tracker (XFT) reconstructs tracks in the COT and matches these to energy depositions in calorimeter towers or hits in muon chambers. In conclusion, L1 reduces the event rate by a factor of 42.5 to 40 kHz.
The second trigger system performs minimum event reconstruction utilizing a custom-designed hardware, which consist of several asynchronous subsystems, for the pre-processing, while the L2 decision is taken on a high performance commodity PC. In addition to the informations from the calorimeters, muon chambers and the COT, L2 processes data from the CEntral Shower maximum detector (CES) and SVXII. CES is placed in the CEM at a depth which corresponds to the average maximum of an electromagnetic shower and improves the identification of electrons and photons. The Silicon Vertex Trigger (SVT) allows to trigger for tracks with large impact parameters which is important for processes involving hadronic decays of b-quarks. Jet reconstruction, which is provided by the Level-2 cluster finder, is also a part of the L2 trigger that accepts in average 600 events per second.
Events passing L2 are sent to the Level-3 computer farm. The main goal of this last trigger is the decrease of the 600 Hz event rate after L2 to about 100 events per second, which can be written to permanent storage. L3 has access about the full event record and its decisions are based on the particle content and the event topology.

Chapter 3

Experimental techniques

3.1 Neural networks

Neural networks(NNs) are simplified models of the neural processing of our brain. During a training process they are able to learn how to perform a special task, e.g. the determination whether a given candidate belongs to a class A or a class B (classification). Once trained, NNs can be applied to unknown datasets. Unlike cut-based analyses, neural networks are able to account for the information and correlations of various variables. In the following we will concentrate on feed-forward NNs.

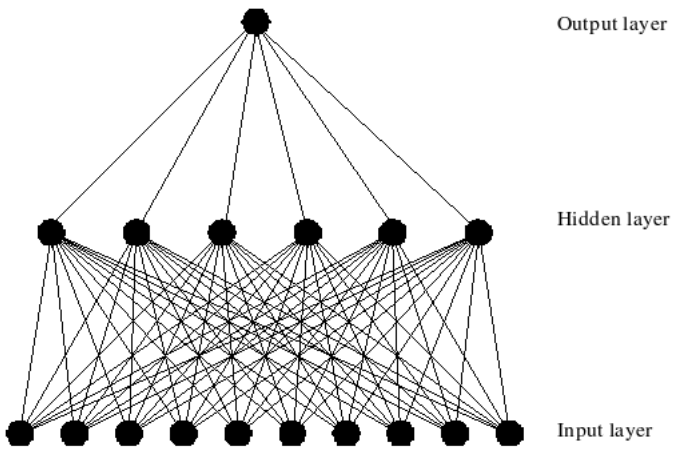


Figure 3.1: Three-layer feed-forward Neural Network with one neuron in the output layer.

A feed-forward net is made up of different layers. The first one is the input-, the last one the output-layer. Layers between those two are called hidden-layers. Furthermore, every layer consists of a specific number of neurons (or simply nodes). There is no connection between nodes of the same layer only the nodes of two adjoining layers

can be connected. All of these connections have specific weights which represent the strength of the contacts. During the network training, the weights are changed and connections can actually disappear (pruning). Figure 3.1 illustrates a feed-forward Neural Network with one node in the output layer.

3.1.1 The NeuroBayes® package

The NeuroBayes neural network package [32, 33] is a highly sophisticated tool to perform multivariate analysis of correlated data. It combines a three-layered feed-forward neural network with an automated preprocessing of input variables and uses Bayesian regularization techniques for the network training.

Preprocessing of input variables

In order to find an optimal starting point for the network training, input variables are automatically preprocessed before being fed to the neural network. In a first step variables are flattened, i.e. they are transformed such that the resulting variables are distributed uniformly between 0 and 1. Then they are scaled to lie between -1 and 1. Such transformations reduce the influence of extreme outliers. The flattend distributions are then converted into Gaussian distributions, centered at zero with standard deviation one. Therefore also the output of the hidden-layer is distributed with mean zero and width one. That provides optimal conditions for a fast initial learning and avoids neuron saturation. Optionally the above mentioned transformation to a Gaussian distribution can be replaced by an individual preprocessing for different variables, like fitting a spline curve to the flattened distributions. The individual preprocessing is also able to deal with variables that are only given for a subset of events by assigning the missing values to a δ-function. Further, it is useful to decorrelate the input variables before starting the network training. Hence, NeuroBayes calculates the covariance matrix of the preprocessed variables and diagonalizes it using iterative Jacobian rotations [34]. Afterwards the rotated input vectors are divided by the square root of the corresponding eigenvalues. Finally the covariance matrix of the transformed variables has become a unit matrix.
After preprocessing NeuroBayes ranks the input variables according to the significance of their correlation to the target. To achieve this, the correlation matrix and the total correlation to the target for the whole set of variables are calculated. In the following, one variable is removed from the original set and the total correlation to the target is recalculated. This process is done with each input variable. Finally, the one with the least loss of information, i.e. the least significant one, is removed from the set and gets the last place in the significance ranking list. The procedure is then repeated for $N-1$ variables and so on until only one variable, reasonably the most significant one, remains. For this ranking the significance is calculated by multiplying the loss of correlation to the target of the individual variable with $\sqrt{n}$, where n is the number of events in the training sample.
Based on the significance ranking of the input variables, one can fix a significance limit

for the variables which will be used to train the network. Hence, one ensures that only variables with relevant information, that is not already given by other variables, enter the network training. For this purpose, NeuroBayes starts at the end of the ranking list and looks for the first variable with a significance higher than the cut value. This variable and all the higher ranked once are then finally used to train the network.

The "independent correlation" option

If one uses individual preprocessing for the variables and fits spline curves to the flattened distributions, there is the possibility of using the "independent correlation" option, as it is named in the following. In this case, only the part of a variable's correlation to target which is independent of the correlation to the target of another variable is incorporated.
In general the flattened distributions are divided in bins. Every bin is represented by its signal probability, i.e. the number of signal events in the bin devided by the overall number of entries in the bin. A spline curve is then fit to this binned signal probability distribution and the resulting signal probabilities for each bin are assigned to every event which corresponds to the bin. Now, suppose we want to have the variable $V_2(x_2)$ independent of the correlation to the target of the variable $V_1(x_1)$ and look to an event which lies for V_1 in the bin i with a signal probability of $p_i(x_1)$ and for V_2 in the bin j with a signal probability of $p_j(x_2)$. Instead of using $p_j(x_2)$ as the signal probability for the given event one makes the following ansatz:

$$p_j(x_2) = p_i(x_1) \cdot f_2(x_2) \tag{3.1}$$

Here $f_2(x_2)$ is a variation function that represents the excess of information given by V_2 in comparison to V_1. $f_2(x_2)$ is calculated for every event of bin j and the same procedure is done for all other bins. Eventually, the spline curves are fitted to the binned $f_2(x_2)$ distribution and the signal probabilities are assigned to the events in the same way as mentioned above.

The network structure

As already mentioned, NeuroBayes is a three-layer feed-forward neural network. Its input-layer consists of one node for each input variable, labeled by an index i, and one bias node. There is further an arbitrary number of nodes in the hidden layer (index j) and for classification trainings discussed in this thesis, there is only one output node which gives a continuous output in the interval $[-1, 1]$. Connections between nodes of different layers are represented by the weights ω_{ij}, between the input- and the hidden-layer, and ω_j, between the hidden- and the output-layer. Weights between the bias node, which implements the threshold for the nodes j, and hidden nodes are given by ν_{0j}. The input of a node j in the hidden layer is calculated via the biased weighted sum,

$$h_j(\mathbf{x}) = \sum_i \omega_{ij} x_i + \nu_{oj} \ , \tag{3.2}$$

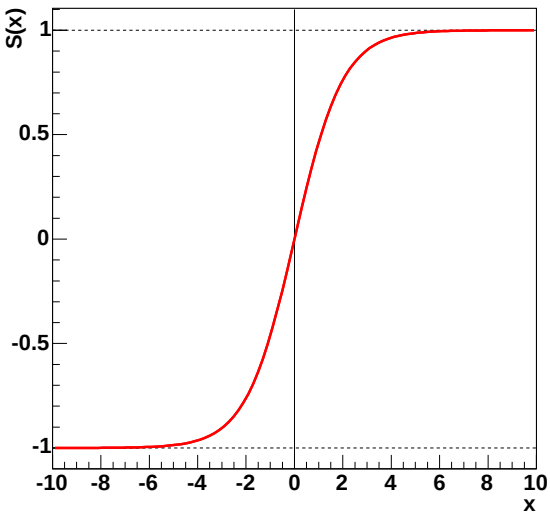


Figure 3.2: The sigmoid function in $[-10, 10]$

where the x_i are the input data. A sigmoid transfer function,

$$S(h(x)) = \frac{2}{1 + e^{-h(x)}} - 1 \ , \tag{3.3}$$

then gives the output of the particular node. This function maps the interval $]-\infty, \infty[$ onto $]-1, 1[$. As can be seen in figure 3.2, the sigmoid function is only sensitive to a small region around zero. For large $|x|$ the sigmoid function is flat whereby saturation effects can appear. Hence, the task of the bias weight is to shift the weighted sum of input values, $\sum_i \omega_{ij} x_i$, to the linear part of the transfer function. Finally, the network output is calculated by

$$o = S(\sum_j^H \omega_j \cdot S(\sum_i^V \omega_{ij} x_i + \nu_{oj})) \ , \tag{3.4}$$

where H is the number of hidden nodes and V is the number of input variables.

The Training Process

Historical or simulated data, for which the expected network output is known, is used for the network training. The training itself is done by minimizing the entropy error function

$$E_{\mathrm{D}} = \sum_i^N log(\frac{1}{2}(1 + T_i \cdot o_i + \epsilon)) \ , \tag{3.5}$$

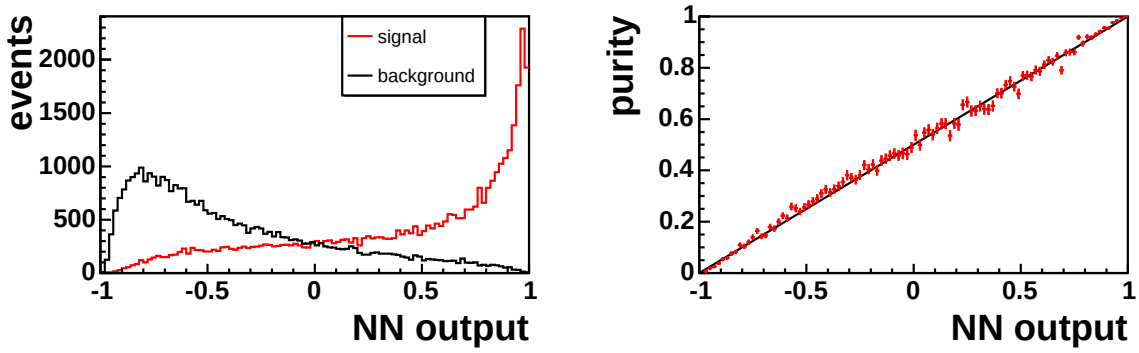
where N is the number of training events, T_i is a binary number which denotes the event as signal ($T_i = 1$) or background ($T_i = -1$) and o_i is the calculated network

output for the specific event, see eq. 3.4. A small constant ϵ was inserted to avoid numerical problems at the beginning of the training (e.g. total misclassification $T_i = 1$ and $o_i = -1$ or vice versa would lead to an infinitely large E_{D}). After just a few training iterations ϵ is equal to zero.
The main challenge during the network training is to find a set of weights for which the multidimensional structure of the error function has a minimum close to the global one. For this purpose NeuroBayes uses a combined method of backpropagation and gradient decent. Here the change of the weights $\delta\omega_{ij}$ is proportional to the current gradient of the error function and the step size η_{ij}:

$$\delta\omega_{ij} = -\eta_{ij}\frac{\partial E(\omega)}{\partial \omega_{ij}} \tag{3.6}$$

The step size is calculated individually for each weight, see [32] for details. Due to the already mentioned pruning, unsignificant weights or even nodes can be neglected. This reduces the number of free parameters, i.e. the dimension of the error function structure. Another minimization is based on the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method.

Training Results



(a) (b)

Figure 3.3: Results of the network training: (a) network output for signal(red) and background(black) (b) signal purity in each output bin.

After a successful training, the distribution of network outputs looks similar to the one illustrated in figure 3.3(a). Signal events cumulate at outputs next to $+1$ and background events gather at -1. If the network was trained with a realistic mixture of signal and background events and if the signal purity in each output bin lies on the diagonal(see figure 3.3 (b)), the output, scaled to $[0, 1]$, can be interpreted as a Bayesian *a posteriori* probability. Again, see [32] for further details.

3.2 Monte Carlo generators and detector simulation

In order to get a description of the different physics processes happening in hadron collisions, the complete physical knowledge about those is implemented in Monte Carlo (MC) event generators. These computer programs use statistical methods to randomly produce hard parton interactions according to the probability density of phase space and the matrix element of a given process. The behaviour of the various particles originating from the hard process, from particle decays or radiation processes in the detector is simulated by the CDFII detector simulation.

3.2.1 Monte Carlo event generation

Besides the hard parton interactions, which take place at high momentum transfer scales Q^2, initial and final state radiation, decays of unstable particles and hadronization processes have to be simulated. At high Q^2 the strong coupling constant α_s is small, thus the scattering processes can be calculated in perturbation theory. Since the hadronization of quarks and gluons, which describes the formation of jets, takes place at low Q^2 and large α_s perturbation theory cannot be applied. The phenomenologic models, used to describe hadronization in the absence of any firm theoretical understanding are different for distinct Monte Carlo generators. In the following we give a brief introduction to the MC generators used to simulate the different processes for the KIT Flavor Separator training (see chapter 4).

Pythia

The Pythia [35] program is a standard tool for the generation of high-energy collisions, comprising a coherent set of physics models for the evolution from a few-body hard process to a complex multihadronic final state. It is able to simulate the collisions of pp, $p\overline{p}$, e^+e^- and $\mu^+\mu^-$. Models for initial- and final-state parton showers, multiple parton-parton interactions, beam remnants and particle decays are also included. The hadronization process is based solely on the Lund string fragmentation framework. In this model the color field between final state quarks and antiquarks is illustrated by color flux tubes (or strings). If the potential energy in such a tube is high enough, it can be transformed in new quark-antiquark pair. Thus colorless hadrons can be built.

ALPGEN

ALPGEN [36] is designed for the simulation of multiparton hard processes in hadronic collisions like the proton-(anti)proton collisions at the Tevatron or the LHC. It performs the calculation of the exact matrix element in leading order perturbation theory for QCD and electroweak hard scattering processes. ALPGEN focusses on multijet final states originating directly from hard QCD radiative processes or the decay of massive

particles, such as for example W and Z bosons. Besides the evaluation of matrix elements, ALPGEN further provides a complete parton-level simulation which offers the possibility to carry out the shower evolution and hadronization of the partonic final states. In order to use these results in practical analyses of experimental data, the calculations need to be completed with the treatment of higher-order corrections leading to the development of partonic cascades, and with the subsequent transformation of the partons into observable hadrons. For this task MC programs like Pythia can be used.

MadEvent

The multipurpose event generator MadEvent [37] is based on the calculation of arbitrary tree level helicity amplitudes with the MadGraph [38] program. Given a process, MadGraph automatically identifies all the relevant subprocesses, generates both the amplitudes and the mappings needed for an efficient integration over the phase space, and passes them to MadEvent. MadEvent then produces a stand-alone code that allows to calculate cross sections and to obtain unweighted events. Once the events have been generated they can be passed to any shower Monte Carlo program (such as Pythia) where partons are perturbatively evolved through the emission of QCD radiation, and eventually turned into physical states (hadronization).

3.2.2 Detector simulation

In order to get a useful simulation of the different physics processes it is necessary to model not only the particle final states correctly but also to get an idea how they will interact with the CDF detector. For this purpose it is crucial to simulate the behaviour of the various particles in the different detector subsystems. The CDF detector simulation framework (cdfSim) is integrated into an AC++ [39] application used to process events at the CDF experiment. GEANT3 [40] performs the tracking of particles through matter.
The main purpose of simulating the silicon detector is to describe the charge deposition of traversing tracks on the silicon strips well enough to evaluate silicon tracking performance. This charge deposition is again calculated by the GEANT3 package. CDF's Central Outer Tracker is modeled by the GARFIELD [41] program, which is a general drift chamber simulation package. The most challenging part in the muon simulation is the description of the complicated geometry of the muon systems. Nevertheless a good agreement between data and simulation has been found [42]. The simulation of the CDF calorimeters is based on the GFLASH [43] parameterization package interfaced with the GEANT3 simulation. Further details of the detector simulation are given in [42].

3.3 Track and jet reconstruction

A crucial prerequisite for the KIT Flavor Separator, as for any other b jet identification algorithm, is the reconstruction of tracks and jets. In the following, we therefore give an introduction to the tracking in the different CDF subdetectors, the jet cluster algorithm, the jet energy corrections and the standard CDF high-p_{T} b-tagging algorithm.

3.3.1 Reconstruction of tracks

In a homogeneous magnetic field, a charged particle travels on a helix whose axis is parallel to the applied magnetic field. The shape of this helix depends on the particle charge and momentum. Particles flying on their helices are detected from the silicon detector and the COT (i.e. the particles produce so-called "hits" in the tracking systems). The challenge for tracking algorithms is the reconstruction of particle tracks from these hits.
In a first step, tracks in the COT are reconstructed. The COT is placed at larger radii than the silicon detector, hence the density of tracks is lower and so the tracks are more isolated which leads to a much easier reconstruction than in the silicon detector. Currently two different reconstruction algorithms for tracks in the COT are in use. The first one, adopted from Run I, reconstructs and links segments in the super-layers to find the trajectory [44]. The second, faster approach [45] uses axial segments found in the outermost super-layers as seeds. For every seed the algorithm calculates a "reference circle" which is tangent to the segment and also includes the beam line. Furthermore for each hit within a large window around the circle, a new trajectory, which goes through the middle of the segment and includes the hit and the beam spot, is calculated. The distances of these trajectories to the reference circle are filled in a curvature histogram. Since real tracks lie along a single "hit circle", they appear as a peak in this histogram. For each track that has been found the algorithm performs a 3D fit using all combinations of stereo segments in the outermost super-layers that could be consistent with the given axial track. The newly estimated 3D trajectory is then used to determine which of the segments in the remaining stereo layers are consistent with belonging to the track.
The track reconstruction in the silicon detector consists of three major phases. In the first, tracks found in the COT are extrapolated into the silicon detector ("outside-in" [46]). Axial and stereo silicon hits are added using a progressive fit. The second phase of silicon reconstruction consists of stand-alone pattern recognition in the silicon detectors [47]. Since it is possible to have as many as $50,000$ seeds from a few tens of tracks it is necessary to reduce the combinatorics. For example hits already added to outside-in tracks are excluded. Hence, the silicon stand-alone algorithm essentially looks for tracks in the forward-backward region, $1 \leq |\eta| \leq 2$, not covered from the COT. The final phase is called "inside-out" tracking [48]. Tracks found with the silicon stand-alone algorithm are extrapolated to the COT to identify tracks which have no hit in the outermost super-layers and thus could not be found by the COT algorithms.

3.3.2 Jet clustering algorithm

Due to hadronization processes, final state quarks create jets of colorless hadrons which are detected in the calorimeters. For these jets different reconstruction algorithms are possible. In this analysis only jets clustered with a cone algorithm [49] are used. An example for the evolution from the $p\overline{p}$-collision via the final state quarks to calorimeter jets is illustrated in figure 3.4.

For the clustering only calorimeter towers with $E_{\mathrm{T},i} > 1$ GeV are taken into account.

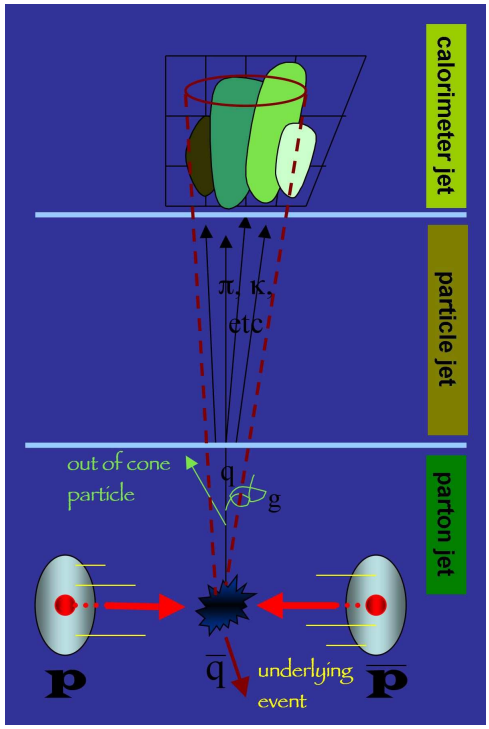


Figure 3.4: Schematic overview of jet evolution and reconstruction

Here $E_{\mathrm{T},i} = E_i \cdot \sin(\theta_i)$ is the transverse energy of a tower with respect to the z-position of the $p\overline{p}$ interaction, and the energy E_i is the sum of the energies measured in the electromagnetic and hadronic compartments of that tower. The jet cluster algorithm starts with the definition of "seed towers" in order of decreasing $E_{\mathrm{T},i}$. All towers within a radius $R = \sqrt{(\eta^{\mathrm{tower}} - \eta^{\mathrm{jet}})^2 + (\phi^{\mathrm{tower}} - \phi^{\mathrm{jet}})} = 0.4$ with respect to its position (η^{jet},ϕ^{jet}) are used to build a "cluster" around each seed tower. In a further step the jet energy and the current center of the cluster is calculated. The above procedure is then repeated (e.g. adding all towers within R around the new center and calculate the new jet energy and the new center) until the jet location is stable. If two jets overlap by more than 50% they are merged together. The final jet energy is calculated by $E_{\mathrm{jet}} = \sum_{i=0}^{N_{\mathrm{tower}}} E_i$, where N_{tower} is the number of calorimeter towers added to the final cluster.

3.3.3 Jet energy correction

From the measured jet energy the transverse momentum of the inital partons can be estimated. For this purpose it is necessary to correct the jet energy for instrumental effects and for radiation and fragmentation effects. This correction is done in different levels, which for example allows the application of a subset only.

- **Level 1: "η-dependent" corrections**
  Level 1 corrects for the nonuniform response in η of the CDF calorimeters, which arises from the separation of calorimeter components at $\eta = 0$ where the two halves of the central calorimeter join and at $|\eta|$=1.1 where the plug and central calorimeters join.

- **Level 2: unused**

- **Level 3: unused**

- **Level 4: "Multiple interactions"** This correction subtracts the contributions to the cluster energy of different $p\overline{p}$ interactions during the same bunch crossing (pile-up).

- **Level 5: "Absolute jet energy scale"** Level 5 corrects the jet energy measured in the calorimeter for any non-linearity and energy loss in the uninstrumented regions of each calorimeter. The jet energy measured is corrected to $\sum p_T$ of the particles within the cone of same size around the parton direction which matched the jet direction with $R < 0.4$.

- **Level 6: "Underlying event"** The underlying event is defined as the energy associated with the spectator partons in a hard collision event. Level 6 subtracts this energy from the particle-level jet energy.

- **Level 7: "Out-of-cone" and Level 8: "Splash-out"** The "out-of-cone" corrections adjusts the particle-level energy for leakage of radiation outside the clustering cone used for jet definition, taking the "jet energy" back to "parent parton energy". The level 7 systematic uncertainties are evaluated by looking at energy leakage from the jet clustering cone up to $R = 1.3$. The uncertainty for the leakage outside $R = 1.3$ is referred to as level 8 "splash-out" uncertainty.

3.3.4 Reconstruction of secondary vertices

A crucial prerequisite for any top quark analysis is the identification of b quark jets with high efficiency. Since nearly 100% of the top quarks decay into a W boson and a bottom quark, identifying those b jets would significantly reduce the background. Common algorithms used to identify b jets are predicated on the reconstruction of secondary vertices and take advantage of the long b-hadron lifetime which is about 1.6 ps. The considerable lifetime in conjunction with a large relativistic boost causes

b hadrons to fly on average several millimeters before decaying. Therefore, b-hadron decays produce secondary vertices displaced from the primary interaction point. By reconstructing charged tracks with large impact parameters, i.e. tracks with a low probability of coming from the primary interaction point, possible displaced vertices can be found. Figure 3.5 illustrates the mentioned event topology. The secondary vertex reconstruction algorithm used by CDF, SecVtx [50], is described below. We focus on so-called *tight* SecVtx vertices, because only those are used in the following chapters.

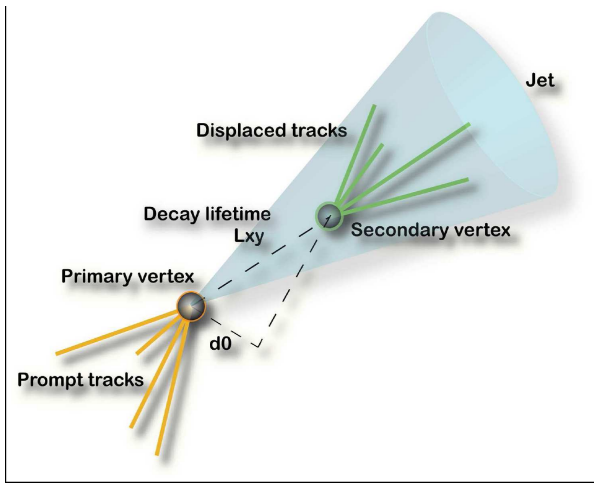


Figure 3.5: Event topology expected for b-quark jets.

SecVtx starts with the selection of tracks associated to a jet with $R < 0.4$, which have large impact parameters d_0. The impact parameters are calculated with respect to the primary $p\overline{p}$ vertex determined for each event separately. In a first step all tracks with $p_T > 0.5$ GeV/c and an impact parameter significance $S_{d_0} = |\frac{d_0}{\sigma_{d_0}}| > 2.0$ are selected. Furthermore tracks with too large d_0 (> 0.15 cm) values are discarded, because they are likely to originate from nuclear interactions in the detector material or the decay of long-lived particles. The tracks are than ordered by their S_{d_0} and their p_T. Starting with the high ranked onse, tracks are paired to form a seed vertex. Around such a seed vertex the algorithm searches for tracks with an impact parameter significance smaller than three with respect to the seed vertex. With these additional tracks a vertex fit is performed. If there is a track which contributes to the vertex fit χ^2 more than 50, the track is discarded and the vertex is refitted until all tracks contribute less than 50 to χ^2. The vertex fit is successful if the number of remaining tracks is ≥ 3 and if at least one of the tracks, used for the secondary vertex fit, has $p_T \geq 1$ GeV/c. The requirements explained above have to be fulfilled for *pass 1* vertices. If the reconstruction of a pass 1 vertex fails, the algorithm tries to find a *pass 2* vertex. For pass 2 vertices tracks are required to have $p_T > 1.0$ GeV and $S_{d_0} > 3.5$. In contrast to the pass 1 procedure, for pass 2 no seeding is performed, hence all tracks are fitted

to a secondary vertex. Again, tracks which contribute too much to the vertex χ^2 are discarded. The vertex fit is successful if at least two tracks remain. Before a jet is said to be "tagged", all secondary vertices have to pass a last round of quality criteria. For each vertex the two dimensional decay length with respect to the primary vertex (L_{xy}) is calculated. The decay length significance, $S_{L_{\mathrm{xy}}} = |\frac{L_{\mathrm{xy}}}{\sigma_{L_{\mathrm{xy}}}}|$ is demanded to be greater than nine. A complete list of all requirements for secondary vertices is given in [55]. Jets for which a good secondary vertex is found are said to be b-tagged. If the dot product of the 2D decay length and the jet axis is positive, i.e. the jet and the reconstructed secondary vertex lie in the same detector hemisphere, the jet is called positively tagged. Otherwise the jet is called negatively tagged. Since secondary vertices corresponding to the decay of heavy flavor hadrons are expected to have large positive L_{xy}, vertices with a negative tag cannot be consistent with heavy flavor decays and are tue to the finite tracking resolution of the CDF II detector. Figure 3.6 illustrates a positive and a negative SecVtx tag. Two further SecVtx vertex versions are available at CDF, the *loose* definition leads to a higher tagging efficiency but less purity. The *ultra-tight* vertex definition increases the tagging purity but decreases the efficiency.

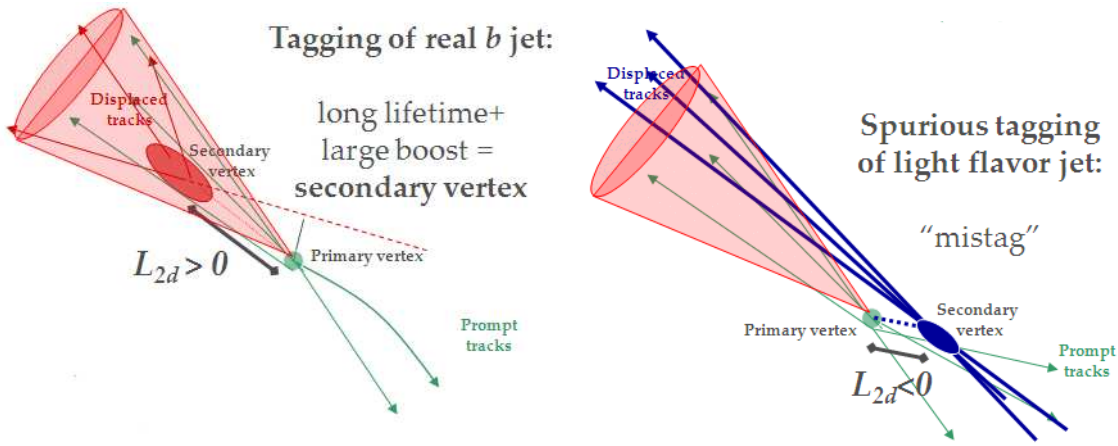


Figure 3.6: On the left hand side a positive SecVtx tag is illustrated. An example for a negative SecVtx tag is shown on the right hand side.

Tagging efficiency

The tagging efficiency of the SecVtx algorithm is defined as the ratio of jets coming from bottom quark production that have a positive tag among all tagged jets. This efficiency is very easy to calculate from Monte Carlo simulations because one has the complete knowledge of the particles within a jet, and thus it is straightforward to identify jets that come from b-quark production and the fraction of tagged ones. Unfortunately, the b-tagging in Monte Carlo simulations has non negligible weaknesses: The charge deposition in the silicon detector is not understood precisely, there is no perfect simulation of tracking and several detector effects, like noise for example. Therefore,

the b-tagging efficiency has to be measured in data. Here the main challenge is to separate a pure b-jet sample from a large amount of dijet events. Two methods for this measurement exist at CDF. Both are utilizing high p_{T} leptons matched to jets to identify jet pairs consistent with heavy flavor decays.
The first method looks for high-p_{T} muons within a jet, the so-called "muon jet". This jet has to be paired with a back-to-back jet, called "away jet", which is required to have a positive SecVtx tag. Such jet pairs are consistent with coming from heavy flavor quark production. Furthermore, the p_{T} of the muon relative to the jet axis ($p_{\mathrm{T}}^{\mathrm{rel}}$) is a powerful discriminator of b jets from jets of charm or light flavor. Fitting the $p_{\mathrm{T}}^{\mathrm{rel}}$ spectra of b and non-b jets derived from Monte Carlo to the tagged and untagged data sample, the fraction of tagged b-jets (i.e. the tag efficiency) can be determined. In 1.2 fb^{-1} of CDF data the muon method finds a b-tag efficiency of $0.388 \pm 0.006(\mathrm{stat}) \pm 0.006(\mathrm{syst})$ in data [51]. The tag efficiency is often used when assessing signal acceptance, which is typically done in signal Monte Carlo samples. Hence it is necessary to construct a data-to-Monte Carlo scale factor for the tagging efficiency, which accounts for the differences between b-tagging in data and Monte Carlo. The muon method calculates this scale factor to $SF_{\mathrm{tag}}^{\mathrm{muo}} = \frac{\epsilon_{\mathrm{tag}}^{\mathrm{data}}}{\epsilon_{\mathrm{tag}}^{\mathrm{MC}}} = 0.932 \pm 0.016(\mathrm{stat}) \pm 0.015(\mathrm{syst})$ [51].
The second method roughly works the same way as discussed above but uses electrons instead of muons. A detailed description of this method can be found in [52]. The result of the electron method for the scale factor is $SF_{\mathrm{tag}}^{\mathrm{ele}} = 0.977 \pm 0.018(\mathrm{stat}) \pm 0.030(\mathrm{syst})$, which is in good agreement with the results of the muon method using the same integrated luminosity. Reference [53] explains how the two scale factors can be combined to $SF_{\mathrm{tag}} = 0.95 \pm 0.01(\mathrm{stat}) \pm 0.04(\mathrm{syst})$. For the single-top analyses discussed in chapter 5 and 6 CDF data corresponding to 2.2 fb^{-1} is used. Since the scale factors doesn't change significantly for the new data the ones for 1.2 fb^{-1} are used [54].

False positive rate

Typically the performance of b-tagging algorithms is also characterized by the rate of light flavor jets which have been mistakenly tagged (mistags). Due to the limited detector resolution, material interaction and decays of long-lived particles (e.g. Λ, K_{S}), tracks in light quark jets can still have large impact parameters and therefore fulfill secondary vertex requirements.
Mistags resulting from the limited detector resolution are expected to be symmetric in L_{xy}. Hence the number of jets with negative tags is a good estimate for the contribution of mistags to the positive tag sample. At CDF an a priori prediction of the mistag rate is calculated from the inclusive jet samples, which are collected by a set of simple triggers gathering events with a certain minimum amount of calorimeter energy. The per jet probability to be a mistag is determined from the probability of getting a negative tag, parametrized in six variables (mistag matrix) [56]:

- Jet E_{T}, the uncorrected transverse jet energy
- N_{trk}, the number of tracks per jet (track multiplicity)

- $|\eta_{\text{jet}}|$, the absolute value of a jet's pseudorapidity
- n_{PV}, the number of primary vertices in an event
- Z_{PV}, the z-position of the primary vertex
- $\sum E_{\text{T}}$, the sum of the transverse energy of all taggable[1] jets

From the mistag probabilities of the whole jet sample, the contribution of mistagged positive jets to a given data sample, due to the limited detector resolution, can be predicted.
However, not all mistags stem from limited detector resolutions alone. Since secondary vertices in light quark jets which are reconstructed from tracks coming from material interactions or long-lived particle decays are strictly at positive L_{xy}, the assumption of a symmetric L_{xy} leads to an underestimation of the false positive tag rate. Hence a correction factor $\alpha\beta$ is applied [57]. Here α is given by the number of light positive tags divided by the number of negative tagged jets in data. The numerator of this fraction is derived from a fit to the vertex mass M_{vtx}. The factor β includes the number of jets in the pretag sample divided by the number of number of light jets in the same sample. Finally the positive mistag rate is given by

$$\frac{N^{+}_{light}}{N^{pre}_{light}} = \alpha\beta \cdot R^{-}_{mistag} \ , \tag{3.7}$$

where R^{-}_{mistag} is the mistag rate predicted from negative tags which is calculated as explained above. Results for the scale factors α and β in bins of E_{T} are given in [57].

[1] A jet is said to be taggable, if $E_T > 10$ GeV, $|\eta| < 2.4$ and at least two tracks are reconstructed within the jet.

Chapter 4

The KIT Flavor Separator

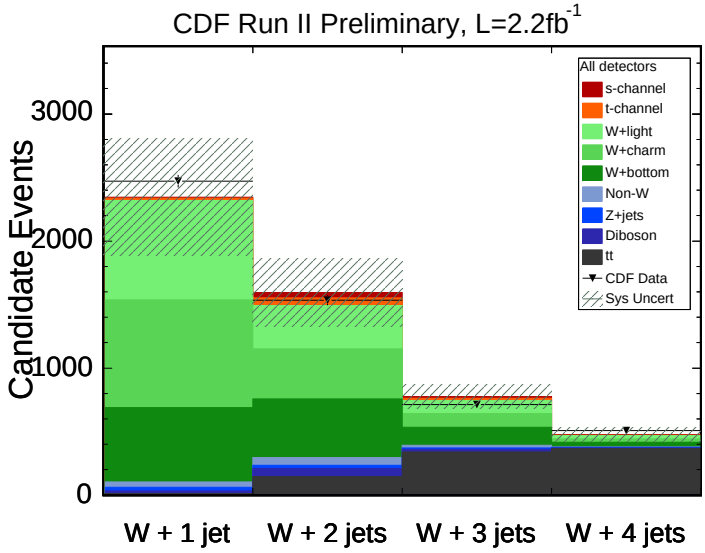


Figure 4.1: Jet multiplicity distribution for signal and background processes. The predicted number of events in each W+jet bin is compared with the number of observed events. Uncertainties on the observation are statistical; the hatch marks represent systematic errors in the background estimate. The single-top analyses are performed using W+2jet and W+3jet events.

After the single-top candidate selection, which is discussed in detail in chapter 5, the expected signal and background composition for events with two and three jets is shown in figure 4.1. In spite of the requirement of at least one jet identified as a b jet, still 50% of the selected single-top candidates do not contain any bottom quark. The main contributions to this background stem from W+charm and W+light quark events. Jets containing a charm quark can be misidentified as a b jet due to the non-zero lifetime of C hadrons which can therefore produce a secondary vertex. An inaccurate track reconstruction leads to a spurious reconstruction of secondary vertices in light

flavored jets. Since the SecVtx-tag method is mainly based on the long b-hadron lifetime of $\tau \approx 1.6$ ps, non-b backgrounds can be further suppressed by taking other b-hadron characteristics into account. Suitable properties are the large b-hadron mass of roughly 5 GeV/c^2, their high decay multiplicity and their probability for decaying semileptonically.
In this chapter we present an improved version of the KIT Flavor Separator [58], a neural network which is able to distinguish between tagged b jets and tagged c/light jets. For this purpose 54 input variables have been investigated and then delivered to the NeuroBayes® package. After preprocessing, 31 variables which have passed a 3.5σ significance cut are used for the network training. The Flavor Separator output is continuously distributed between -1 and 1. Most of the c/light jets have outputs next to -1 and b jets accumulate at outputs next to 1. Furthermore, we show how the output distributions for the different processes can be validated or, if necessary, corrected. In the last section we will give some examples of use for our Flavor Separator.

4.1 Training of the neural network

4.1.1 Event selection and used MC-Samples

Since the requirements of the single-top candidate selection drastically reduce the number of available events, we only demand that an event has to have exactly one tight (see lepton identification in subsection 5.1.2) electron or muon. As explained in section 3.3, jets within an event are reconstructed with a jet-cone cluster-algorithm ($R < 0.4$). Furthermore, each jet used for the network training has to have a positive SecVtx tag and a corrected transverse energy of $E_{\mathrm{T}} > 20$ GeV. Jet energies are corrected up to hadron level.
For the modelling of b signal we use single-top s-channel, t-channel, $\mathrm{t\bar{t}}$, and $Wb\bar{b}$ simulated events. Only jets in which at least one track can be identified as produced by a particle of the b-Hadron decay, using the Monte Carlo truth, are taken into account. The W+light jet events are derived from $W+n$-parton ($W+np$) samples by excluding events which contain a b or c hadron. $Wc\bar{c}$ and Wc simulated events are used to describe the background containing c-quarks. From these events we use only jets which contain a track produced by a particle of a c-hadron decay. In all W boson samples the W boson is forced to decay leptonically, i.e. $W \to e\nu_e$, $W \to \mu\nu_\mu$ or $W \to \tau\nu_\tau$. Table 4.1 gives on overview of the used Monte Carlo samples and the particular MC generator.

4.1.2 Discriminating variables

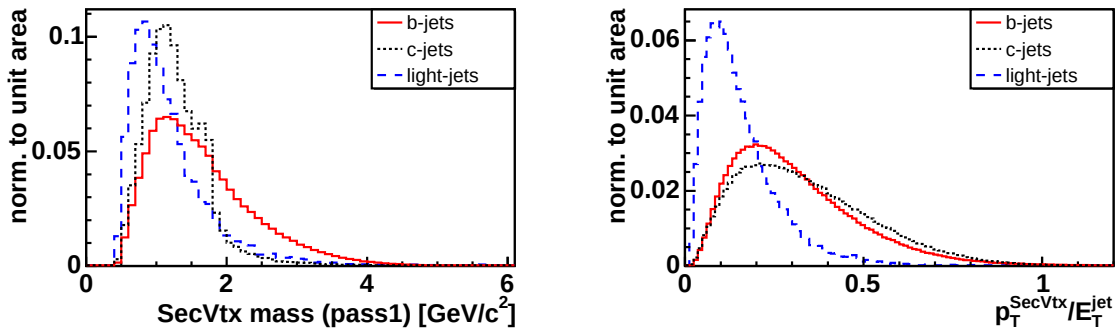
Lots of the variables used to discriminate b jets from c and light jets are direct results of the secondary vertex fit. Since the fit procedure is very different for pass 1 and pass 2 vertices the shapes of these variables are expected to look different too. Hence, we decided to split up pass-dependent variables into two separate ones (one for pass 1

process	Generator	datasets
single top t-channel	MadEvent + PYTHIA	stopm0
single top s-channel	MadEvent + PYTHIA	stop00
$t\bar{t}$	PYTHIA	ttop75
$Wbb + np$	AlpgenV2 + PYTHIA	btop0w, btop1w, btop2w btop5w, btop6w, btop7w dtop0w, dtop1w, dtop2w
$Wc\bar{c}$ +np	AlpgenV2 + PYTHIA	ctop0w, ctop1w, ctop2w ctop5w, ctop6w, ctop7w etopw0, etopw1, etopw2
$Wc + np$	AlpgenV2 + PYTHIA	stopw0, stopw1, stopw2, stopw3 stopw5, stopw6, stopw7, stopw8 stopwa, stopwb, stopwc, stopwd
Mistags	AlpgenV2 + PYTHIA	ptop0w, ptop1w, ptop2w, ptop3w ptop4w, ptop5w, ptop6w, ptop7w ptop8w, ptop9w, utop0w, utop1w utop2w, utop3w, utop4w

Table 4.1: Monte Carlo samples used for the training of the neural network with the utilized generator and the CDF internal dataset name.

and one for pass 2). In case of a pass 2 vertex, the corresponding pass 1 variables are assigned to a δ-function (see 3.1.1) and vice versa. The pass of a given vertex itself is used as an input variable, because for roughly 70% of the b jets a secondary vertex is already found with the pass 1 algorithm. However, more than 50% of the vertices in c and *light* quark jets can only be fitted in the pass 2 approach.
Below we illustrate the distributions of four variables, each exploiting one of the main b-hadron characteristics mentioned above. Moreover the distributions of four "new" variables, that haven't been used in previous versions of our Flavor Separator, are shown. The distributions of all other variables can be found in [59] or in documentations of previous tagger versions [58, 60]. Nevertheless all used variables are explained in detail.

The invariant mass of the tracks fitted to the secondary vertex is shown in figure 4.2(a). As can be seen the secondary vertex mass is larger for jets containing a b hadron. The cutoff between 1.7 and 1.8 GeV/c^2 in the distribution for c-jets originates from the c-quark mass. $p_{\mathrm{T}}^{\mathrm{SecVtx}}$ is the transverse momentum of all the tracks used for the secondary vertex reconstruction. One new variable accounts for the ratio of the transverse momentum at the secondary vertex and the transverse energy of the jet, $p_{\mathrm{T}}^{\mathrm{SecVtx}}/E_{\mathrm{T}}^{\mathrm{jet}}$. In figure 4.2(b) one can see that light quark jets accumulate at small $p_{\mathrm{T}}^{\mathrm{SecVtx}}/E_{\mathrm{T}}^{\mathrm{jet}}$ values.
The two and three dimensional decay length (L_{xy} and L_{xyz}) are variables which incorporate the long b-hadron lifetime. False reconstructed vertices in light quark jets are

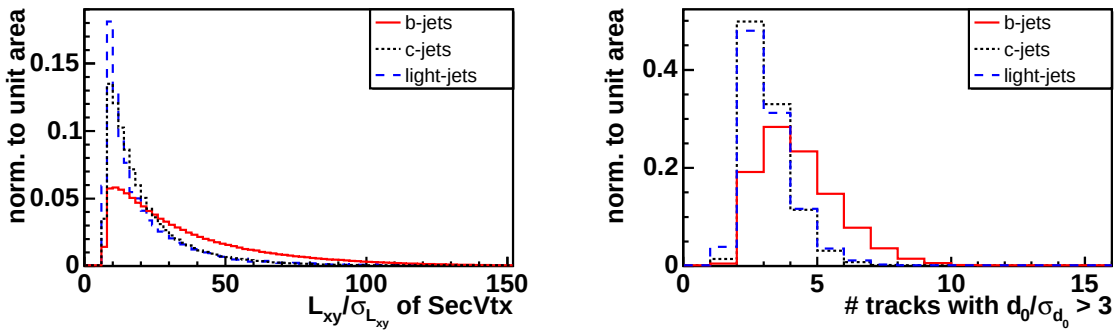


(a) (b)

Figure 4.2: Variables used for the discrimination of b jets from non-b jets: (a) the invariant mass of the tracks fitted to a pass 1 secondary vertex (b) $p_{\mathrm{T}}^{\mathrm{SecVtx}}$ devided by the transverse energy of the jet, $E_{\mathrm{T}}^{\mathrm{jet}}$.

likely to have large distances from the primary vertex but are expected to have smaller decay length significances ($L_{\mathrm{xy}}/\sigma_{L_{\mathrm{xy}}}$ and $L_{\mathrm{xyz}}/\sigma_{L_{\mathrm{xyz}}}$) than vertices stemming from b- or c-hadron decays. The $L_{\mathrm{xy}}/\sigma_{L_{\mathrm{xy}}}$ distribution is illustrated in figure 4.3(a). Using the two dimensional decay length, the secondary vertex mass and the transverse momentum at the secondary vertex, the pseudo-lifetime of a vertex, $c\tau = L_{\mathrm{xy}} \cdot M_{\mathrm{SecVtx}}/p_{\mathrm{T}}^{\mathrm{SecVtx}}$, can be calculated. Spuriously reconstructed secondary vertices can have large $c\tau$ values. The number of "good SecVtx" tracks N_{trk} within a jet represents the decay multiplicity of b hadrons. Here "good"' means that the tracks have to fulfill some minimal quality requirements [55]. Also the number of tracks used in the vertex fit $N_{\mathrm{trk}}^{\mathrm{vtx}}$ which is large for vertices stemming from b-hadron decays, contains relevant information. Another sensitive quantity is the number of tracks which are suitable for the pass 1 reconstruction algorithm explained in section 3.3. Figure 4.3(b) shows the distribution of the number of tracks with an impact parameter significance d_0/σ_{d_0} larger than three. As it is illustrated secondary vertices in b jets consist of a large number of tracks with $d_0/\sigma_{d_0} > 3$. The absolute value of the charge $|Q_{\mathrm{vtx}}|$ at the secondary vertex is a further input variable. A measure for the quality of the vertex fit is the χ^2/dof value. For false reconstructed vertices in light quark jets, the distribution of χ^2/dof is shifted to larger values.

Besides the variables related to the secondary vertex, jet and track based variables are also taken into account. In case of jets the transverse energy $E_{\mathrm{T}}^{\mathrm{jet}}$ and the jet's pseudorapidity η^{jet} are included. For the purpose of including track-based information, the good SecVtx tracks within a tagged jet are sorted according to their momentum. The three tracks with the largest momentum are labeled as the 1st, 2nd and 3rd track. All the following track-based variables are related to each of these tracks. Variables which incorporate the long b-hadron lifetime are the impact parameter of a track d_0 and its



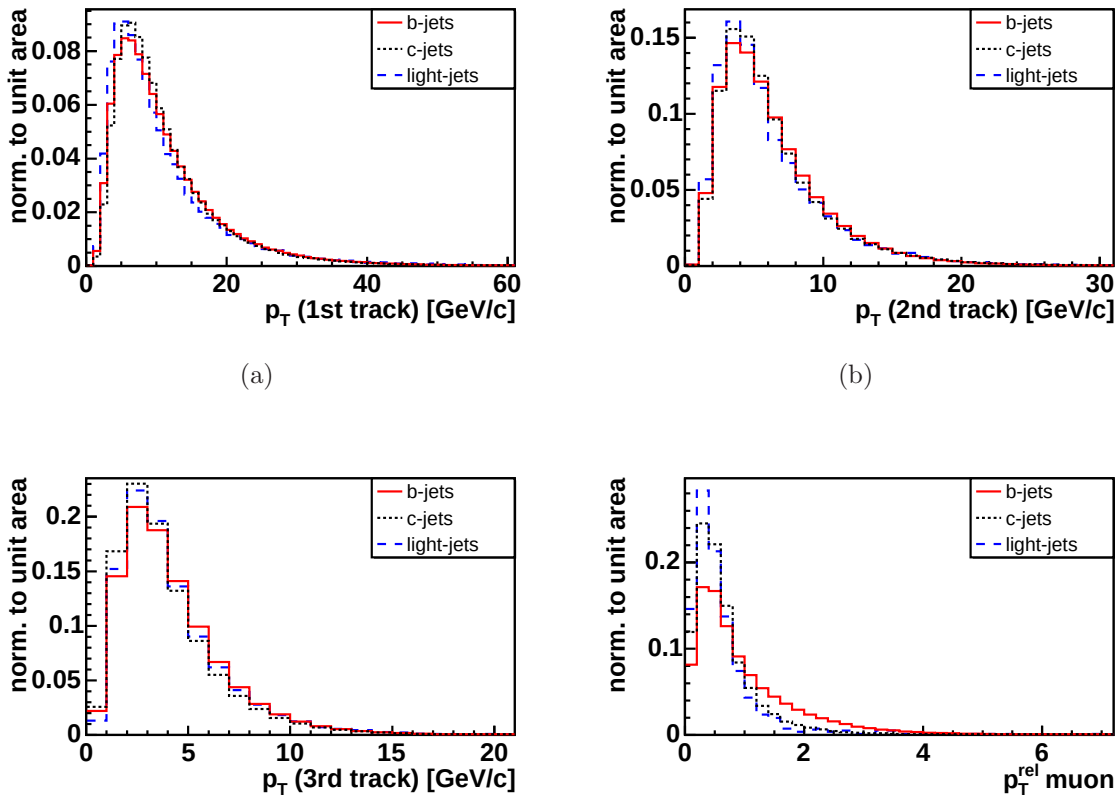
(a) (b)

Figure 4.3: Variables used for the discrimination of b jets from non-b jets: (a) the significance of the two dimensional decay length derived from the secondary vertex fit (b) the number of tracks in the jet with an impact parameter significance larger than three.

significance. Tracks coming from a real secondary vertex are likely to have large impact parameters and high values for d_0/σ_{d_0}. This is also true for the lifetime-signed impact parameter, which is defined as

$$\frac{\epsilon^{zik} \cdot p^i_{jet} \cdot p^k_{track}}{|\epsilon^{zik} \cdot p^i_{jet} \cdot p^k_{track}|} \cdot d_0 \tag{4.1}$$

and the corresponding significance. Quantities which exploit the b-hadron decay kinematics are the momentum p of the tracks, their transverse momentum with respect to the jet axis $p^{\mathrm{rel}}_{\mathrm{T,track}}$ and their rapidities with respect to the jet axis $y^{\mathrm{rel}}_{\mathrm{track}}$. In the current version of the Flavor Separator, additionally the transverse momentum of the tracks $p^{\mathrm{track}}_{\mathrm{T}}$ is used. Figures 4.4(a) - 4.4(c) show the $p^{\mathrm{track}}_{\mathrm{T}}$ distributions for the 1st, 2nd and 3rd track.
The number of leptons within a jet, their energy and momenta are also taken into account, because b hadrons have a large probability for decaying leptonically. As an example figure 4.4(d) illustrates the transverse momentum of the muon with respect to the jet axis $p^{\mathrm{rel}}_{\mathrm{T,muo}}$. Furthermore the number of muons and electrons in a jet, their transverse momentum $p^{\mathrm{muo}}_{\mathrm{T}}$ and $p^{\mathrm{ele}}_{\mathrm{T}}$, the transverse energy of the electron and its p_{T} with respect to the jet axis are used to separate b jets from c and light jets.
Since most of the above mentioned variables show only small differences between b and non-b processes and are additionally highly correlated, a cut-based analyses won't succeed. We therefore use a neural network to combine all available information to one discriminant.



(a) (b) (c) (d)

Figure 4.4: Variables used for the discrimination of b jets from non-b jets: (a)-(b) p_T distribution of the 1st, 2nd and 3rd track (d) transverse momentum of the muon with respect to the jet axis.

4.1.3 Network training

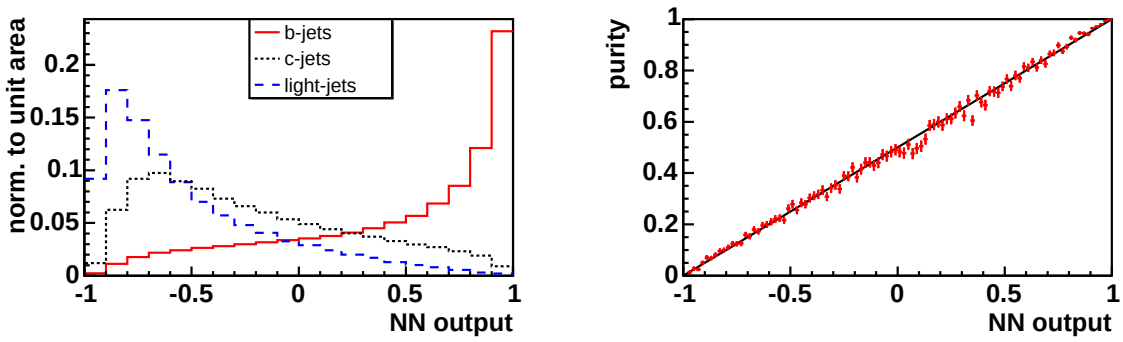
For the network training we use a realistic mixture of b-like signal and non-b background processes. The signal template consists of single-top s-channel, single-top t-channel, $t\bar{t}$ and $Wb\bar{b}$ events. Together the signal processes make up 53% of all training events. The remaining 47% background events are composed of $Wc\bar{c}$, Wc and W+light samples. The different subprocesses are mixed with fractions close to the ones predicted from the single-top background estimate (see section 5.2). Table 4.2 summarizes the different processes and the particular fractions.

As explained above, 54 variables have been investigated for the separation of b jets from c/light jets. These variables are fed into the NeuroBayes package. After preprocessing the variables are ranked according to their relative significance (see 3.1.1 for details). We decided to perform a cut on the significance at 3.5σ, i.e. only variables with a relative significance larger than 3.5σ are used for the network training. Table 4.3 lists the 31 variables which survived the significance cut sorted by their relevance.

b signal	Fraction of the complete sample
t-channel	2.9%
s-channel	2.0%
$t\bar{t}$	19.0%
$Wb\bar{b}$	29.1%
Σ b signal	53.0%
non–b background	Fraction of the complete sample
$Wc\bar{c}$	11.0%
Wc	11.1%
Mistags	24.9%
Σ non–b background	47%

Table 4.2: Mixing ratios of the training samples used for the neural network training.

The residual variables are shown in table 4.4 also sorted according to their significance. The neural network has the following structure: 32 input nodes, ten hidden nodes and one output node. After the network training the output for b-quark jets, c-quark jets and light-quark jets look like illustrated in figure 4.5(a). A good separation of the b-like signal and the non-b backgrounds can be found. Vertices stemming from C-hadron decays are classified to be less background-like than the spurious reconstructed vertices in light jets, which is mainly due to the non-zero C-hadron lifetime leading to "real" secondary vertices. Since the network is trained with a realistic mixture of signal and background events and, as shown in figure 4.5(b), the purity in each output bin lies on the diagonal, the network output (transformed to [0,1]) can be directly interpreted as a Bayesian *a posterior* probability for a jet to be a b-jet.



(a) (b)

Figure 4.5: (a): Output of the Flavor Separator for b, c and light jets. (b): Purity of the network output in each output bin

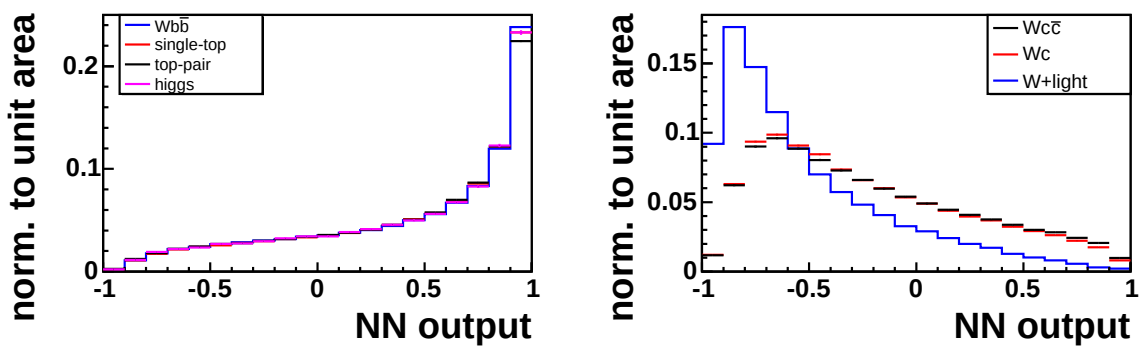
Rank	Variable	rel. Significance (in σ)
1	number of tracks with d_0 significance > 3	116.1
2	signed d_0 significance of the 1st track	69.2
3	signed d_0 of the 2nd track	46.3
4	signed d_0 of the 3rd track	37.3
5	pass of SecVtx fit	28.1
6	invariant mass at secondary Vertex (pass 1)	27.5
7	p_T of the muon with respect to jet axis	25.7
8	invariant mass at secondary vertex (pass 2)	21.4
9	number of good tracks in jet	20.6
10	pseudolifetime at secondary vertex	18.5
11	number of used tracks in fit (pass 1)	14.9
12	E_{T} of the jet	13.5
13	η with respect to jet axis of the 2nd track	12.5
14	momentum of the 3rd track	12.3
15	significance of the 3D decay length	11.5
16	momentum of the 1st track	10.5
17	p_{T} of the 1st track with respect to the jet axis	9.1
18	E_{T} of the electron	8.8
19	significance of the 2D decay length	7.7
20	χ^2 per degree of freedom of SecVtx fit (pass 1)	7.2
21	number of used tracks in fit (pass 2)	6.8
22	d_0 of the 3rd track	5.4
23	d_0 significance of the 3rd track	5.8
24	p_{T} at secondary vertex (pass 1)	6.3
25	p_{T} at secondary vertex over E_T of jet	6.0
26	d_0 of the 2nd track	5.4
27	momentum of the 2nd track	5.2
28	d_0 significance of the 2nd track	5.1
29	number of muons in jet	5.1
30	p_{T} of the muon	4.3
31	charge at secondary vertex (pass 2)	3.9

Table 4.3: Set of discriminating variables used to train the Flavor Separator, sorted by relevance. The quoted relative significances are determined as described in section 3.1.1. These variables passed the requirement of the significance being larger than 3.5σ.

Rank	Variable	rel. Significance (in σ)
32	η with respect to jet axis of the 3rd track	3.4
33	d_0 significance of the 1st track	3.1
34	2D decay length	3.1
35	signed d_0 of the 1st track	2.7
36	η with respect to jet axis of the 1st track	2.3
37	p_{T} of the 1st track	2.0
38	p_{T} of the 2nd track	2.5
39	3D decay length	2.4
40	χ^2 per degree of freedom of SecVtx fit (pass 2)	1.5
41	number of tracks with d_0 significance > 2	1.5
42	number of pass 1 tracks in jet	1.4
43	η of the jet	1.3
44	number of electrons in the jet	1.2
45	p_{T} at secondary vertex (pass 2)	1.2
46	number of tracks with d_0 significance > 1	1.2
47	signed d_0 significance of the 2nd track	1.1
48	charge at secondary vertex (pass 1)	1.0
49	p_{T} of the electron with respect to the jet axis	0.7
50	p_{T} of the 2nd track with respect to the jet axis	0.7
51	p_{T} of the 3rd track with respect to the jet axis	0.5
52	signed d_0 significance of the 3rd track	0.4
53	p_{T} of the 3rd track	0.3
54	d_0 of the 1st track	0.2

Table 4.4: Set of discarded variables sorted by relevance. The quoted relative significances are determined as described in section 3.1.1. These variables did not survive the cut on the significance $> 3.5\sigma$.

Figure 4.6(a) shows the KIT Flavor Separator output for different processes containing b jets. As can be seen the output is completely independent of the origin of a b jet. Hence, the Flavor Separator can for example also be used in a Higgs analysis ($WH \rightarrow b\bar{b}$). The output distributions for the different background processes are illustrated in figure 4.6(b). Again the output for $Wc\bar{c}$ and Wc events look very similar. In the figures 4.9(a) - 4.9(d) the neural network output for signal and background processes is split for pass 1 and pass 2 vertices. Because of the high b-hadron decay multiplicity, the vertices from b-hadron decays are mostly reconstructed by the pass 1 algorithm. This quantity has been learned by the neural network and thus b jets with reconstructed pass 2 vertices look more background-like. For the same reason the neural network output distribution for for c-quark jets, with pass 1 vertices, looks more signal-like.
Further studies comparing the Flavor Separator output for different numbers of primary vertices (Nzvtx) in an event, are performed. As can be seen in figure 4.7 the outputs for b, c and light jets look similar up to the number of five primary vertices per event. Moreover, studies show that the Flavor Separator output does not depend on the type of the lepton in which the W boson decays (see figure 4.8).



(a) (b)

Figure 4.6: (a) : Flavor Separator output for different processes which contain b jets. (b): Flavor Separator output for $Wc\bar{c}$, Wc and W+light jets.

Just to get an idea of how well the Flavor Separator works, we apply a cut on its output at -0.4. After this cut still 90% of the b jets remain, however the number of c jets is reduced by a factor of 1.75 and actually only 31% of all light-quark jets pass this cut. Cutting for example at 0.9 leads to an almost pure b-sample whereas roughly 1/4 of all b jets remain. Like for every cut, one has to compromise between the sample purity and the number of remaining events. It turned out that in practice it is better to use the continuous Flavor Separator output as input to further multivariate analyses (see sections 4.3 and 5.3).

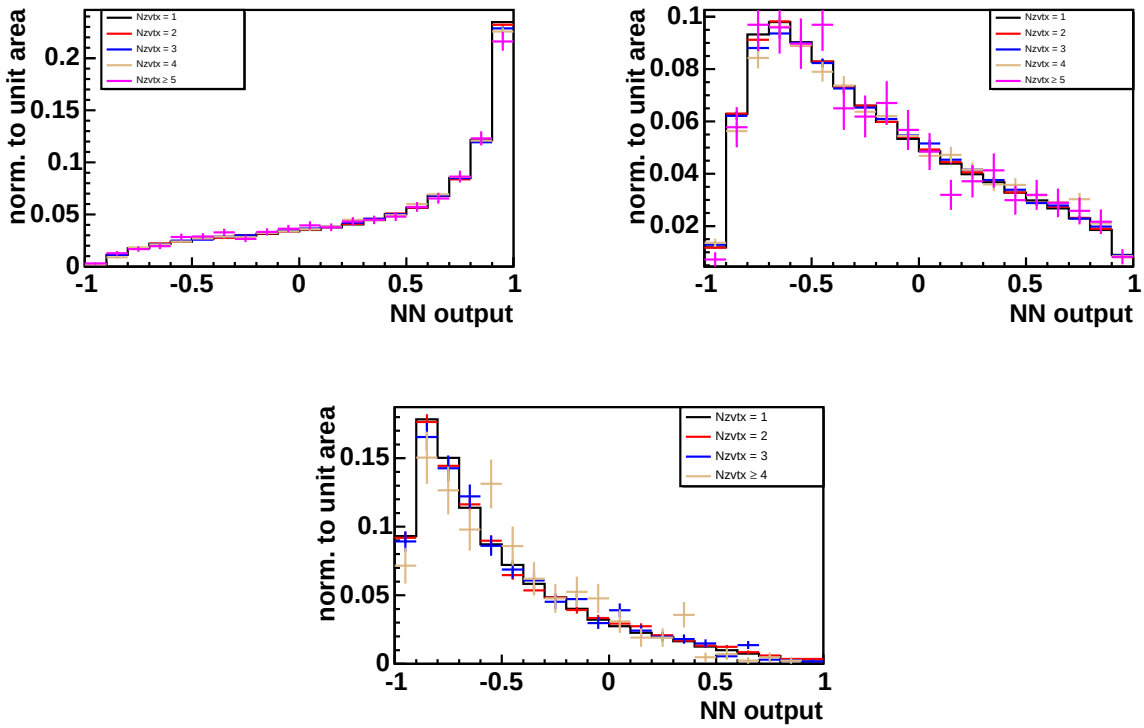


Figure 4.7: Comparison of the Flavor Separator output for different numbers of primary vertices. On the upper left hand side the distributions for b-jets are shown. The plot on the upper right hand side shows the comparison for c-jets. Beneath the distributions of light-jets are illustrated.

4.2 Comparison of MC-Events and measured data

As mentioned above, the neural network is trained with simulated events. Hence, one has to ensure that the simulation describes observed events well, i.e. avoid any mismodelling of input variables or network outputs. For this purpose three steps are necessary: First, one has to verify that the shapes of the input variables look the same in observed and simulated events. In the following step the Flavor Separator output of a b jet enhanced data sample is compared with the output of simulated events. Finally, the same comparison is done for light-quark jets.

4.2.1 Utilized samples

In general one wants to compare the events, which were used for the network training, with measured events. In our case this is not feasible, because the flavor content of a jet cannot be determined easily in observed events. Thus, only data samples in which special quark flavors are enriched can be taken. For signal-like events we use dijet-data with an identified electron or muon. Mistag events are obtained from data collected

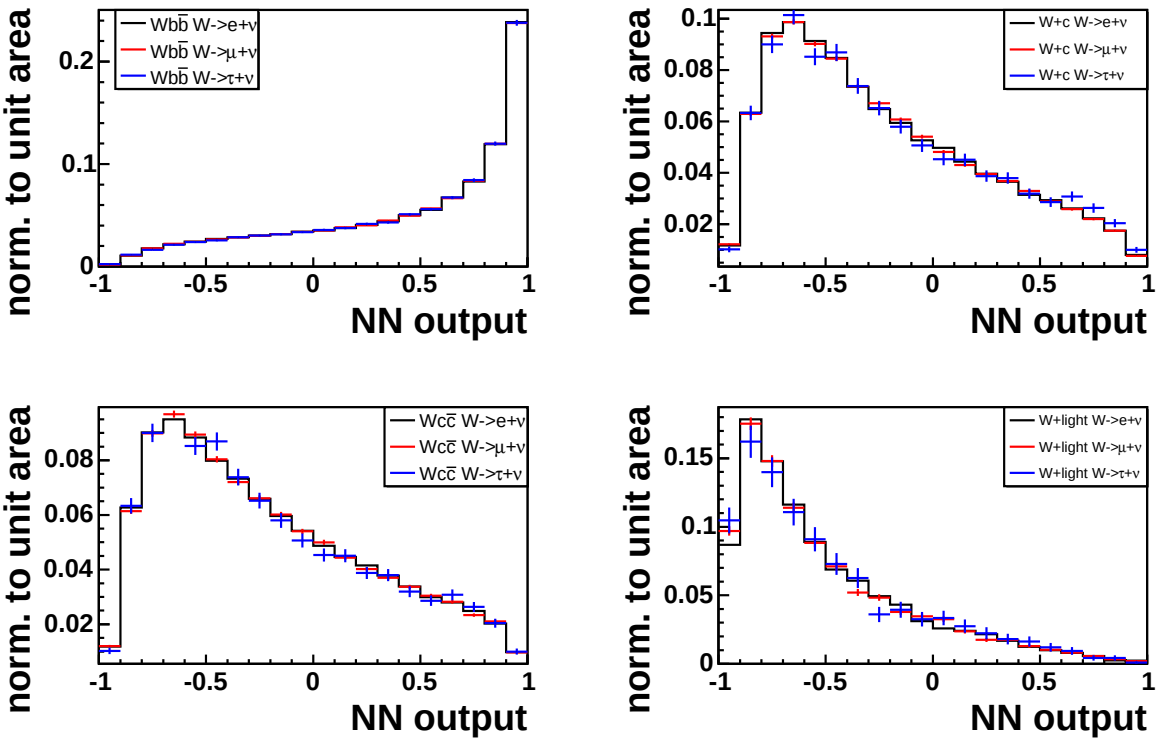


Figure 4.8: Comparison of the Flavor Separator output distributions for the different types of leptons the W boson decays into. On the upper left hand side the distributions for b-jets are shown. The plot on the upper right hand side shows the comparison for Wc processes. On the lower left hand side the comparison is done for jets stemming from $Wc\bar{c}$ processes. The fourth plot shows the output of light-quark jets.

with different jet triggers, which require different minimum jet energies.

Heavy-Flavor Samples

- **Inclusive** 8 **GeV electron data** We use data collected with the inclusive electron trigger (*blpc0d, blpc0h, blpc0i* and *blpcmi* datasets), which requires a COT track with $p_{\mathrm{T}} > 8$ GeV/c matching an energy cluster with $E_{\mathrm{T}} > 9$ GeV. The samples have been stripped for an 8 GeV electron and two 10 GeV jets.

- **Inclusive** 8 **GeV muon data** We use data collected with the MUON_CMUP8 trigger (*bmcl0d, bmcl0h, bmcl0i* and *bmclmi* datasets). The samples have been stripped for an 8 GeV muon and two 10 GeV jets.

- **Monte Carlo** Pythia has been used to generate some generic $2 \rightarrow 2$ scattering QCD events with a minimum outgoing transverse momentum of 20 GeV where all flavors are generated (btopla sample). A filter requiring a 7 GeV electron or

a 8 GeV muon in the generator parton list discards most light flavor events. In this sample, unlike data, every event contains a lepton that was not created by material interactions. The events are accepted or rejected according to a trigger parameterization of the 8 GeV L2 lepton trigger, which describes the trigger's efficiency to identify an electron, based on its fully reconstructed E_{T}, p_{T} and, in part, calorimeter isolation. The parameterization that has been chosen is the one that corresponds to a moderate isolation cut ($\mathrm{Isol}_{0.4} < 4$) [61], although the shape of the function does not vary strongly with isolation.

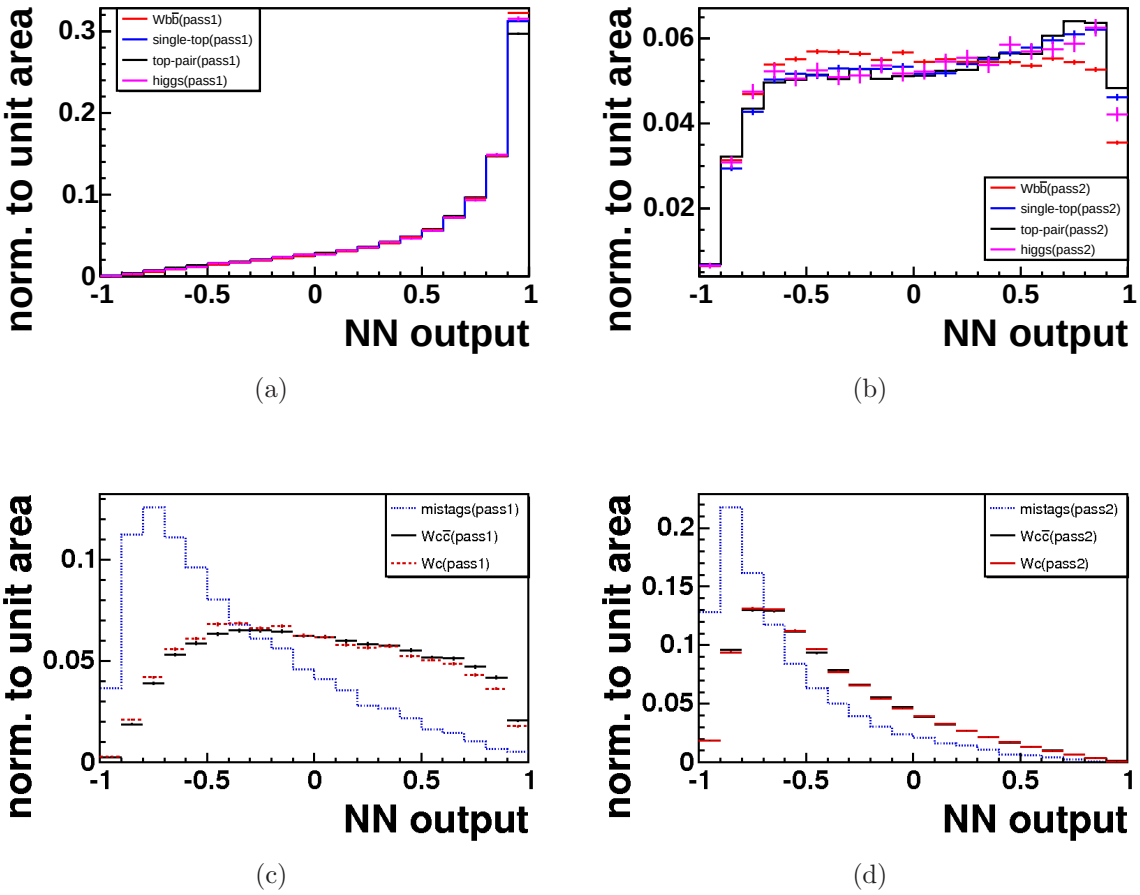


Figure 4.9: (a) : Flavor Separator output for different processes which contain b jets with pass 1 secondary vertices. (b): Flavor Separator output for different processes which contain b jets with pass 2 secondary vertices. (c)Flavor Separator output for $Wc\bar{c}$, Wc and W+light jets with pass 1 secondary vertices. (d)Flavor Separator output for $Wc\bar{c}$, Wc and W+light jets with pass 2 secondary vertices.

Mistag Samples

- **Jet-trigger data** The dijet data used for the comparison of mistagged jets is selected by generic jet triggers, which require a jet, reconstructed with a cone of $R = 0.7$, having a transverse energy of $E_{\mathrm{T}} > 20$ GeV (gjt1XX), $E_{\mathrm{T}} > 50$ GeV (gjt2XX), $E_{\mathrm{T}} > 70$ GeV (gjt3XX) or $E_{\mathrm{T}} > 100$ GeV respectively.

- **Monte Carlo** Dijet events simulated with Pythia are used to model the data. The jets are required to have at least $p_{\mathrm{T}} > 18$ GeV/c (btopqb), $p_{\mathrm{T}} > 40$ GeV/c (btoprb), $p_{\mathrm{T}} > 60$ GeV/c (btopsb) or $p_{\mathrm{T}} > 90$ GeV/c (btoptb) respectively. Since the jet trigger is not simulated for these events the shapes for the corrected E_{T}^{jet} are different in data and Monte Carlo. However, if one cuts on $E_{\mathrm{T}}^{\mathrm{jet}} > 50$ GeV (gjt2XX, btoprb), $E_{\mathrm{T}}^{\mathrm{jet}} > 70$ GeV (gjt3XX, btopsb) and $E_{\mathrm{T}}^{\mathrm{jet}} > 100$ GeV (gjt3XX, btopsb) the shapes look more comparable. The two different scenarios are shown in figure 4.10.

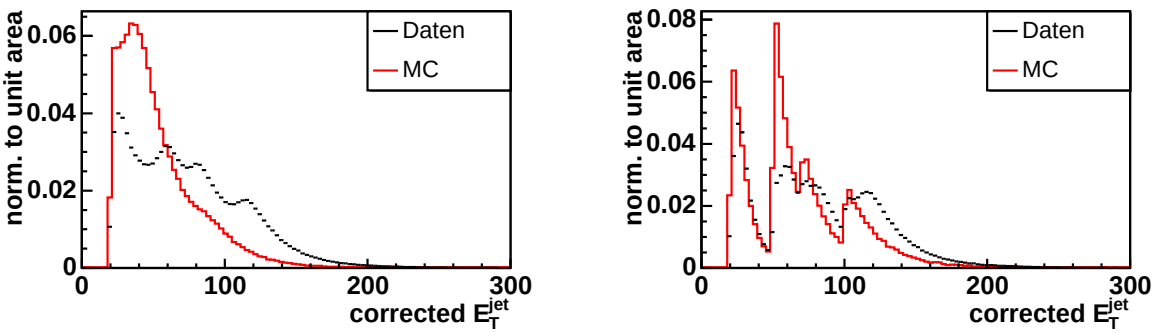


Figure 4.10: $E_{\mathrm{T}}^{\mathrm{jet}}$ distributions for the mistags data and respective Monte Carlo samples. On the left hand side no cuts are applied. On the right hand side the cuts explained in the text are applied which leads to more comparable shapes.

4.2.2 Comparison of input distributions

As already explained, the shapes of the input variables in data and simulation have to be compared, because the mismodelling of an input variable could lead to corrupt network trainings. For the comparison we use the inclusive lepton data and the corresponding MC-samples introduced above. Each shape is checked twice, once for electron and once for muon data. In the following we show the distributions for the four variables additionally included in this version of the Flavor Separator (figures 4.11 and 4.12). All other distributions are shown in [59].

Most input variable shapes show a good agreement in data and simulation. The only significant deviations can be seen for the distribution of the transverse momentum at

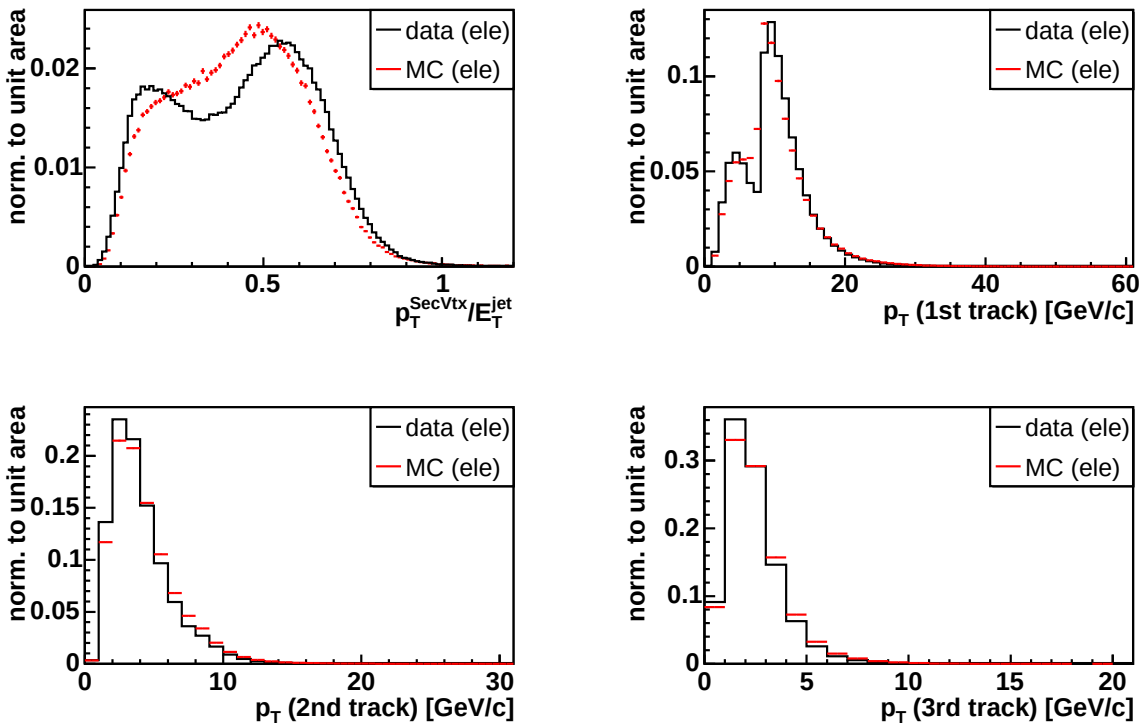


Figure 4.11: Comparison of input distributions with inclusive electron data. The four new variables, added to the Flavor Separator training since the last version, are shown. All other distributions can be found in [59].

the secondary vertex, for the shape of p of the track with the highest momentum in the jet and for the ratio of p_T at the secondary vertex and the E_T of the jet. All these distributions show a bump at low momenta which can't be found in simulation. There is evidence to suggest, that those bumps stem from the track of the high-p_T lepton, which, in data, fails more often to be a "good SecVtx" track than in simulation. In [60] it has been demonstrated, that if one requires the trigger electron (i.e. the electron fulfilling the cuts described in the previous section) to be in the same jet as the considered track and, in addition, that the track fulfills the appropriate requirements to be a SecVtx track (i.e. $p_T > 0.5$ GeV), the distributions of p of the track with the highest momentum in the jet, fit perfectly for measured and MC events. In addition, the track of the lepton seems to be used less often for the secondary vertex fit in data, than in simulation. This has been proven by requiring the lepton track to be used in the fit and compare the transverse momentum at the secondary vertex for data and MC. With these requirements, those distributions fit quite well [60]. Hence the deviations are specific for the used control sample and not expected in genuine b-quark jets. In summary one can say that all input variables are modeled well and thus it is reasonable

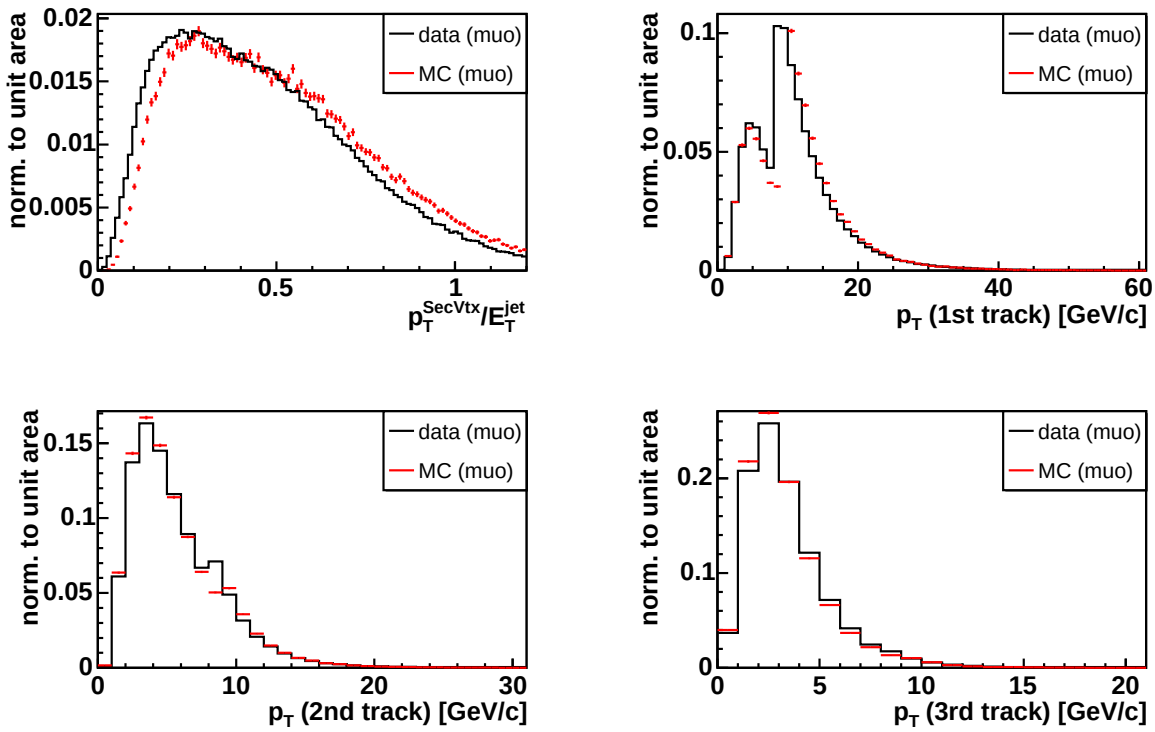


Figure 4.12: Comparison of input distributions with inclusive muon data. The four new variables, added to the Flavor Separator training since the last version, are shown. All other distributions can be found in [59].

to train the neural network with simulated events.

4.2.3 Validation of the network output for b-quark jets

For a validation of the network output for b-quark jets, first a heavy flavor enhanced sample has to be selected. In a second step, the composition of the selected sample has to be determined. Finally, the output distributions of observed and simulated events are compared.

Preparation of the samples

To enhance heavy-flavor quarks within the inclusive electron sample, a method similar to the one used in the determination of the SecVtx scale factor (see 3.3.4) is incorporated. An energetic electron, passing the same cuts as performed in [52], is required within a jet, because it is often a signal for a heavy-flavor quark decay. Since these quarks are dominantly produced in pairs, a further jet, opposite in ϕ, called the "away jet" is also required. This additional demand increases the purity of the heavy-flavor

sample. In contrast to the calculation of the SecVtx scale factor we focus on the away jet for the following reason: The network is very sensitive to the presence of any lepton. This is not only due to the use of lepton specific variables but also due to the use of track based variables such as the momentum and the rapidity of the 1st track with respect to the jet axis. Thus, a comparison including the electron jet would lead to a strong bias.

Composition of the away jet sample

The contribution of different quark flavors to the heavy-flavor enhanced sample is expected to differ for measured and simulated events. However, for a comparison it is necessary to have roughly the same contributions. Therefore, the understanding of the flavor composition of the away jet sample is mandatory. In simulated events the flavor of a jet is determined by using Monte Carlo true values to look for a b or c quark with $p_{\mathrm{T}} > 5$ GeV/c within a cone of $R = 0.4$ around the jet axis. If such a heavy quark is found, the jet is assigned as a b or c jet, respectively. In the other case the jet is labeled as light.
To estimate the heavy-flavor content in data a method, which is well established at CDF, is used. This method takes advantage of the distribution of the invariant mass at the secondary vertex. The flavor content can be estimated by fitting the secondary vertex mass distributions of b jets, c jets and light jets derived from Monte Carlo events to the distribution found in data. For this purpose two possibilities are imaginable. On one hand side events for which only the away jet has a reconstructed secondary vertex can be used. On the other hand side both jets can be required to have a SecVtx-tag. Studies [59, 58] showed, that demanding a SecVtx tag for both jets leads to a better suppression of c-quark and light-quark events and increases the comparability of data and simulated events. In figure 4.13 the mass templates derived from Monte Carlo events and the fit to the data distribution are illustrated. The fit results are summarized in table 4.5. Within the uncertainties, a good agreement of the heavy-flavor composition in data and simulation can be found.

sample	b fraction	c fraction	light fraction
MC	94.9%	2.8%	2.3%
data	98.3% ± 2.2%	0.0% ± 3.4%	1.6 ± 1.5%

Table 4.5: Fractions of b, c and light away jets in Monte Carlo events and data: The composition in the simulation was obtained by using Monte Carlo true values. The fractions in data are estimated by fitting secondary vertex mass distributions derived from simulation to the data. For these numbers, both the electron and the away jet are required to be tagged by SecVtx.

Comparison of output distributions

In figure 4.14 the Flavor Separator outputs for events with a tagged electron jet and a tagged away jet in data and Monte Carlo events are compared. As can be seen,

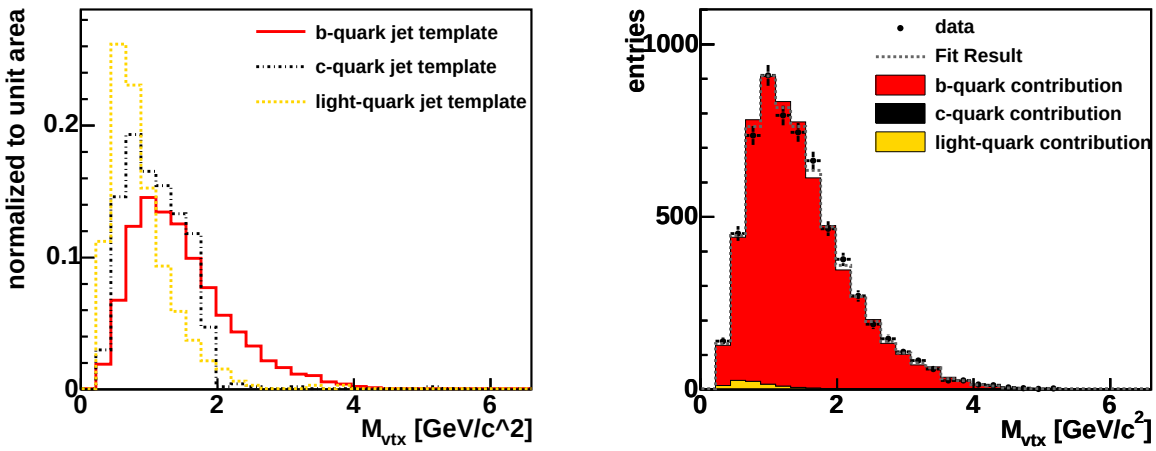


Figure 4.13: On the left hand side one can see the mass templates used for the fit. On the right hand side, the secondary vertex mass distributions normalized to their fit values and the measured distribution are illustrated. Both, the electron and the away jet are required to have a SecVtx tag.

no significant deviations between the two distributions can be found. In contrast to previous Flavor Separator versions, there is no need for the determination of a correction function this time. The differences observed in the previous version of the Flavor Separator are most likely due to different compositions of the away jet sample. Looking to table 4.5 and [58] one can see that the contribution of c quarks and light quarks to the heavy-flavor content were much higher last time, which led to a more background-like data distribution.

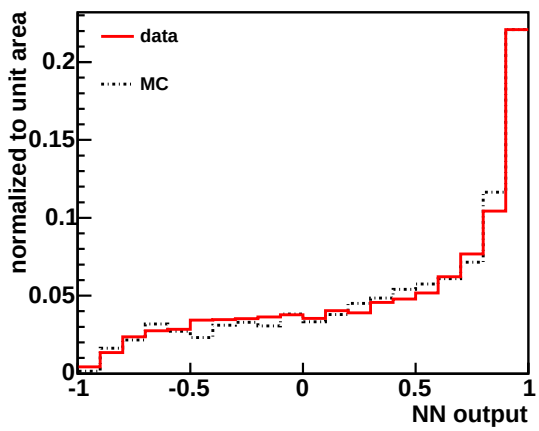


Figure 4.14: Comparison of the Flavor Separator output for observed and simulated b-quark jets.

4.2.4 Correction of the mistag output

Since the neural network is trained with jets which are already tagged by the SecVtx algorithm, it is difficult to select light-quark jets. A model, similar to the one used for the calculation of the mistag matrix (see subsection 3.3.4), assumes that jets with a negative SecVtx-tag, i.e. with a negative decay length L_{xy}, can describe positive mistagged jets. However, the jets used for the network training have all positive tags, and thus per definition a positive decay length. This caveat is taken into account by using the absolute value of L_{xy} when applying the network to negatively tagged jets. The simulation of mistagged jets is complicated, because it requires a good knowledge of the detector effects leading to a spurious reconstruction of a secondary vertex. Thus, one has to verify that light-quark jets are simulated consistently. In figure 4.15 the Flavor Separator output of simulated events with a negative tag is compared with the output distributions of negative and positive tagged light-quark jets. As can be seen the distributions look similar which leads to the conclusion that negatively tagged jets are a suitable model for positively mistagged jets.

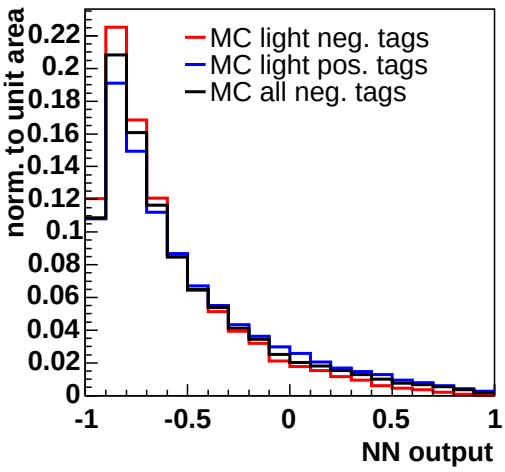


Figure 4.15: The output distribution of negative tagged jets, is compared with the output distributions of positive tagged light jets and negative tagged light jets.

If there is a deviation between the Flavor Separator output in measured and simulated events, one wants to determine a correction function which allows an event-by-event correction of the differences. For the determination of a correction function, so-called "accumulated" distributions (f^{acc}) are used, which have the advantage of decreasing monotonically. These distributions are calculated by integrating the output from the left to the right, i.e. at an output value i the value of the accumulated distribution is the integral of the output distribution from i to 1 divided by the integral from -1 to 1. If the output distributions of measured and simulated events are completely alike, the ratio of $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ for each output value should be equal to one.

The comparison of simulated and observed negatively tagged jets is shown in figure 4.16(a). One can see that the output distribution of observed events is more signal-like than the distribution for simulated events. As already mentioned before, these distributions are accumulated to quantify the difference in the outputs. The ratio of the accumulated distributions of data and Monte Carlo events, $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$, is illustrated in figure 4.16(b). Since the output distributions are quite different for jets with pass 1 and pass 2 vertices (figures 4.16(c) and 4.16(d)), these two jet classes are treated separately in the following. Due to the significant deviations in the output of observed and simulated events the calculation of a correction function is necessary.

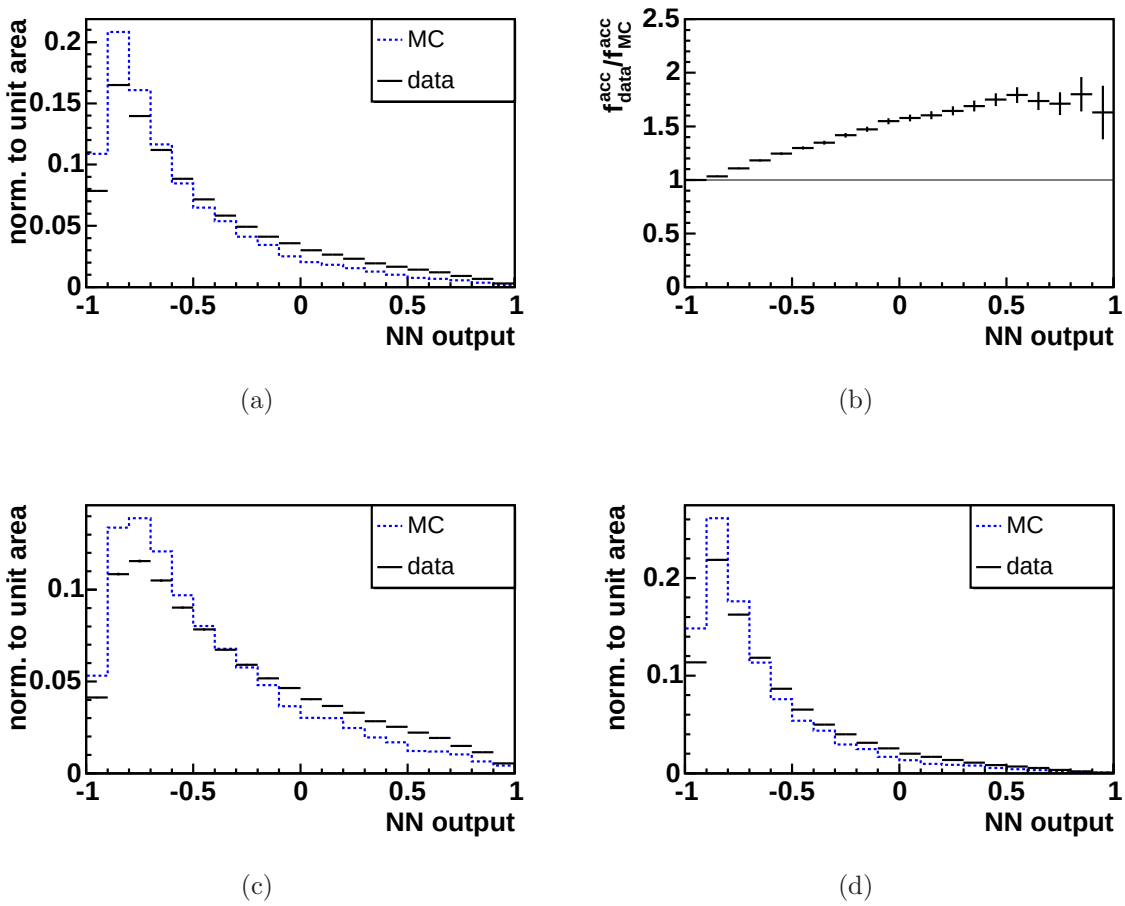


Figure 4.16: (a): Comparison of the output distributions of observed and simulated mistag events. (b): Ratio of the accumulated data and Monte Carlo distribution (c): Comparison of the output distributions of observed and simulated mistag events for jets with pass 1 vertices. (d): Comparison of the output distributions of observed and simulated mistag events for jets with pass 2 vertices.

Parametrization of the correction function

As described in subsection 3.3.4 the probability of a given jet to be mistagged can be parametrized with six variables: E_{T} of the jet, number of good tracks in the jet N_{trk}, $|\eta|$ of the jet, number of z-vertices in an event, z-positon of the primary vertex z_{PV} and sum of the transverse energy of all taggable jets, $\sum E_{\mathrm{T}}$. Therefore, it is reasonable that the deviations found for the Flavor Separator output of mistagged jets between data and simulation also rely on these variables. The dependencies are investigated by dividing each variable into two bins (e.g. jets with $E_{\mathrm{T}} < 60$ GeV and $E_{\mathrm{T}} \geq 60$ GeV). The variables cannot be divided in more than two bins, because the number of simulated events is limited and thus the number of events in the different bins would be too small to see non-statistical dependencies. For each bin the $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ ratio is calculated. If the $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ distribution looks alike for the two bins, there is no need to correct for this variable. If significant deviations for the two bins can be found, the variable is taken into account for the correction function. The $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ distributions of the six variables for jets with a pass 1 vertex are shown in figure 4.17. Deviations can be seen for jets with $E_{\mathrm{T}} < 60$ GeV and $E_{\mathrm{T}} \geq 60$ GeV, events with $\sum E_{\mathrm{T}} < 90$ GeV and $\sum E_{\mathrm{T}} \geq 90$ GeV, jets with $N_{\mathrm{trk}} < 8$ and $N_{\mathrm{trk}} \geq 8$ and jets with $|\eta| < 0.8$ and $|\eta| \geq 0.8$. For mistagged jets with pass 2 vertices the same dependencies can be found, as it is illustrated in figure 4.18. Furthermore the dependency of $|\eta|$ is found to mainly rely on a specific subset of events. In the case of pass 1 vertices, these events have $E_{\mathrm{T}} \geq 60$ GeV, $N_{\mathrm{trk}} \geq 8$ and $\sum E_{\mathrm{T}} \geq 90$ GeV. For jets with a pass 2 vertex, mainly events fulfilling $E_{\mathrm{T}} \geq 60$ GeV, $N_{\mathrm{trk}} < 8$ and $\sum E_{\mathrm{T}} < 90$ GeV depend on $|\eta|$. The $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ distributions of these two subsets are shown in figure 4.19. $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ distributions of $|\eta|$ for all other possible combinations of E_{T}, N_{trk} and $\sum E_{\mathrm{T}}$ are shown in appendix A.

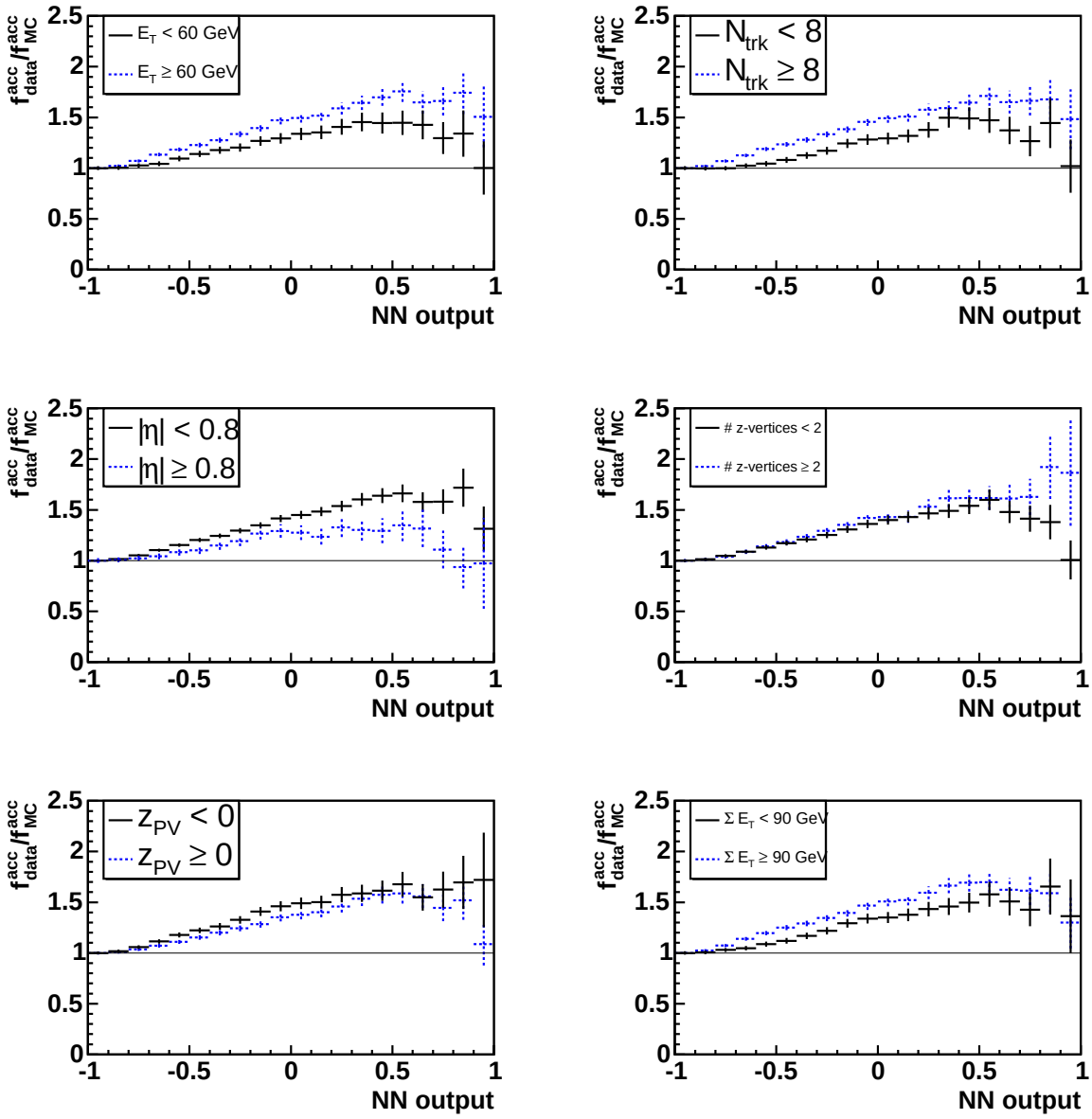


Figure 4.17: $f^{acc}_{data}/f^{acc}_{MC}$ dependencies on the six variables mentioned in the text. Only jets with a pass 1 vertex are taken into account.

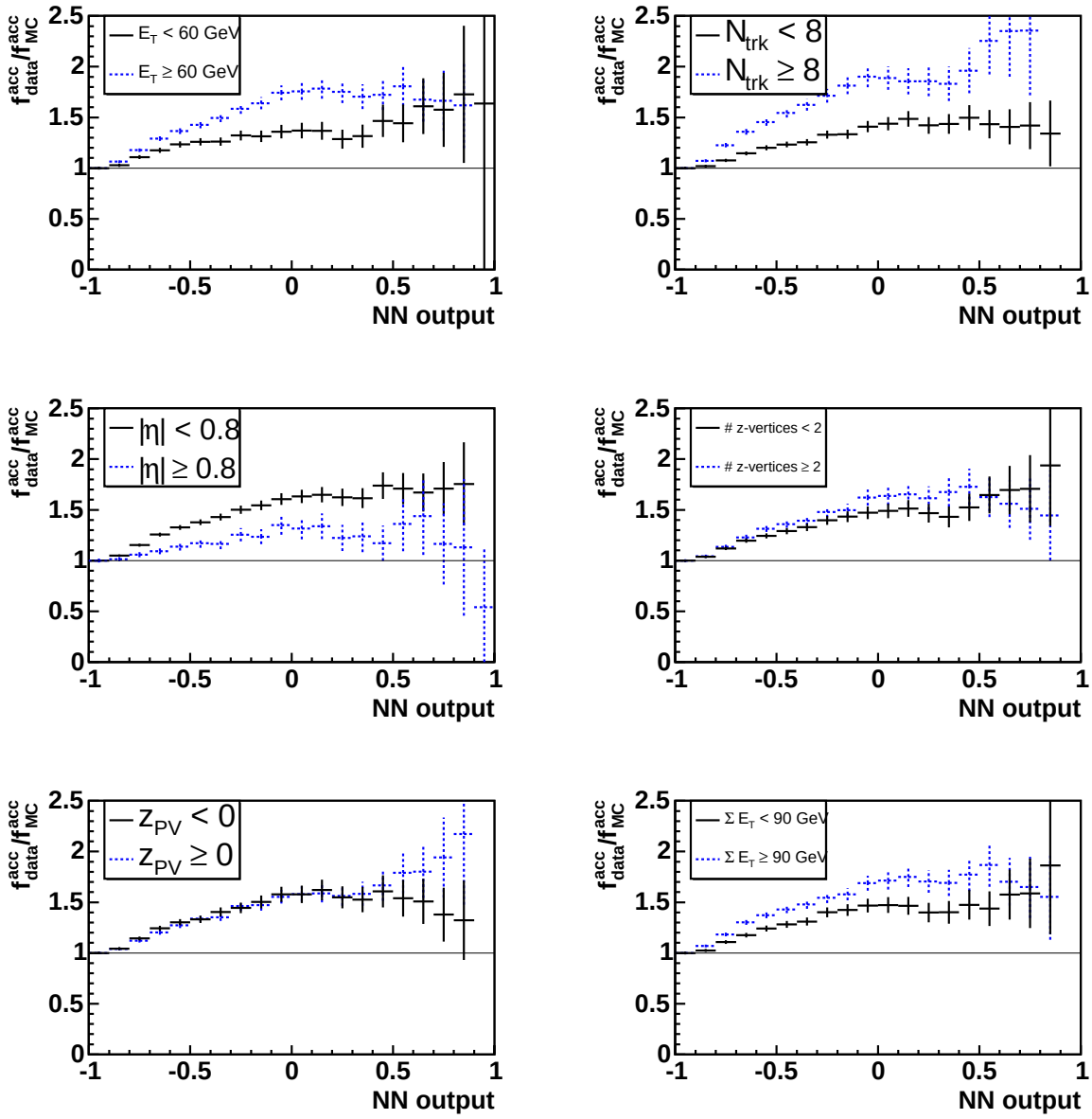


Figure 4.18: $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ dependencies on the six variables mentioned in the text. Only jets with a pass 2 vertex are taken into account.

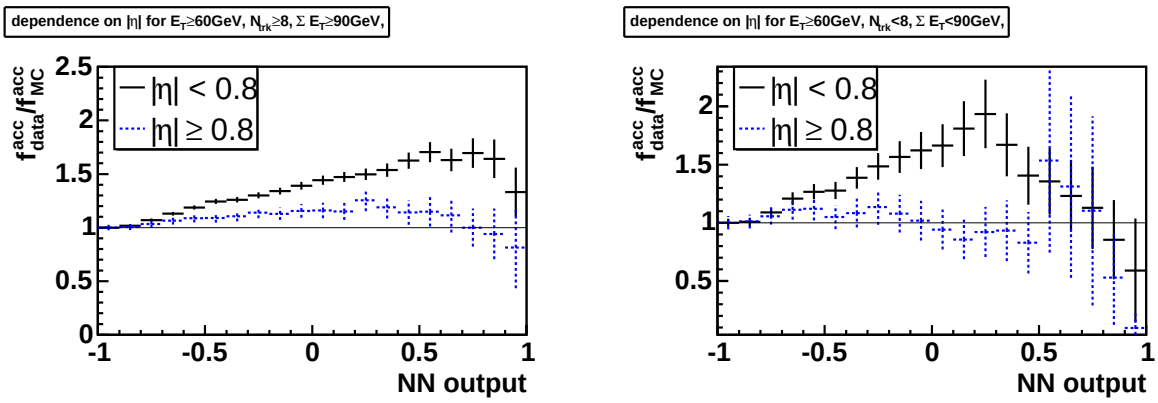


Figure 4.19: On the left hand side the dependency on eta of jets with pass 1 vertices fulfilling $E_{\mathrm{T}} \geq 60$ GeV, $N_{\mathrm{trk}} \geq 8$ and $\sum E_{\mathrm{T}} \geq 90$ GeV is shown. On the right hand side the dependency on eta of jets with pass 2 vertices fulfilling $E_{\mathrm{T}} \geq 60$ GeV, $N_{\mathrm{trk}} < 8$ and $\sum E_{\mathrm{T}} < 90$ is illustrated.

The functions used to correct the differences in the output of mistagged jets in data and simulation are parameterized in the three variables E_{T}, N_{trk} and $\sum E_{\mathrm{T}}$. For the particular subsets of events mentioned above also $|\eta|$ is taken into account. Overall nine different correction functions have to be calculated for jets with pass 1 and pass 2 vertices respectively. For this purpose the output of each simulated jet o_{orig} is replaced by the output value o_{corr} where the accumulated data distribution has the same value as the accumulated distribution of the simulated jets, i.e. $f^{\mathrm{acc}}_{\mathrm{data}}(o_{\mathrm{corr}}) \stackrel{!}{=} f^{\mathrm{acc}}_{\mathrm{MC}}(o_{\mathrm{orig}})$. In figure 4.20 the correction functions for pass 1 and pass 2 vertices are illustrated.

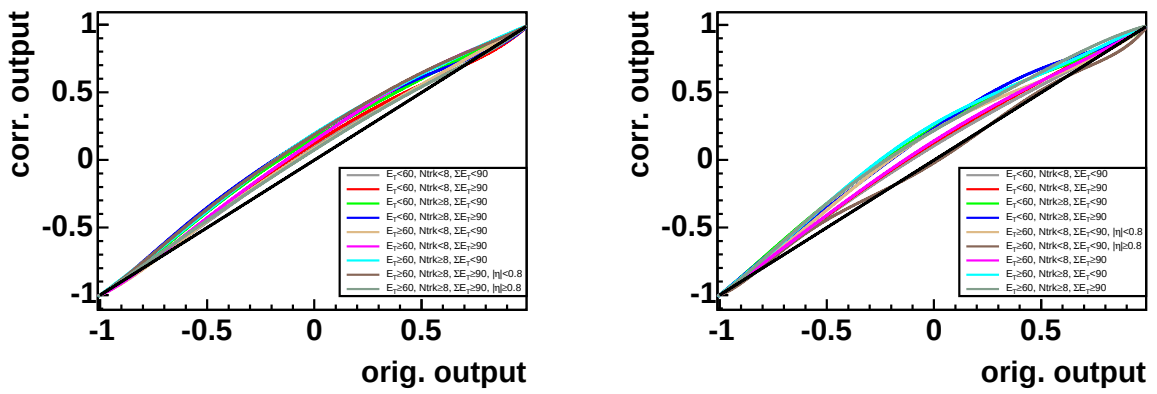


Figure 4.20: On the left hand side the functions used to correct the output of simulated jets with a pass 1 vertex are illustrated. The correction functions for jets with pass 2 vertices are shown on the right hand side.

After applying the correction functions to the output of simulated events, their Flavor Separator output matches the observed distribution very well. A comparison between the original, the corrected and the observed outputs is shown in figure 4.21. The difference between the original and the corrected output is used as systematic uncertainty on the Flavor Separator output.

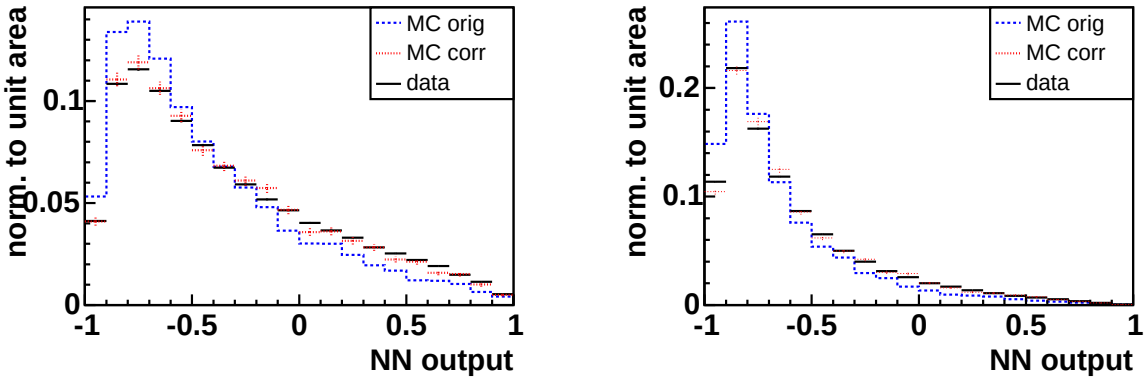


Figure 4.21: On the left hand side one can see the comparison of the original, the corrected and the observed Flavor Separator output of mistagged jets with a pass 1 vertex. The same comparison for jets with a pass 2 vertex is shown on the right hand side.

4.3 Application in analyses

The KIT Flavor Separator was primarily developed for the search of single top-quarks. It is integrated in all CDF single-top analyses and is very appropriate for the identification of events containing b quarks. Its application in the different analyses is discussed more detailed in section 5.3.
Furthermore, the continuous Flavor Separator output can be used for the estimation of the heavy flavor content of a given data sample. Below we show how the fraction of b and c jets can be estimated in a sideband (events with one jet) of the single-top analyses.
Because of the already mentioned universality of the Flavor Separator output it can be used in every analysis which has to distinguish between b and non-b events. For example, the Flavor Separator is used in a search for Higgs bosons [62].

4.3.1 Estimation of heavy flavor contents

Since the output distributions of b jets, c jets and light jets have different shapes, the fraction of each flavor can be estimated. Therefore one fits the output templates from

simulated events of each flavor to the output distribution of observed events.
In the single-top analyses, the predicted heavy flavor fractions for events with two jets are corrected with a correction factor K which is derived from a fit to the output distributions of events with one leptonic W and one jet passing the single-top candidate selection (see chapter 5). This factor incorporates the differences between the fitted and the predicted heavy flavor fractions.
In the following we show the results of a fit to 2.2 fb^{-1} of data. In this fit, the fractions of $Wb\bar{b}$, $Wc\bar{c}$+Wc and W+light processes are all allowed to float freely. In figure 4.22(a) the fitted templates and the fit result is shown. The yellow band represents the uncertainty on the fit result. For comparison we perform also a fit to the Flavor Separator output distribution of events with two jets whereof one jet has a SecVtx-tag. The results of this fit is illustrated in figure 4.22(b). In table 4.6 the fit results are compared with the predicted fractions. Within the given uncertainties the fit results are well comparable with the predicted numbers. Nevertheless the central values for c-jets and light-jets are somewhat different from the expected central values. The predicted

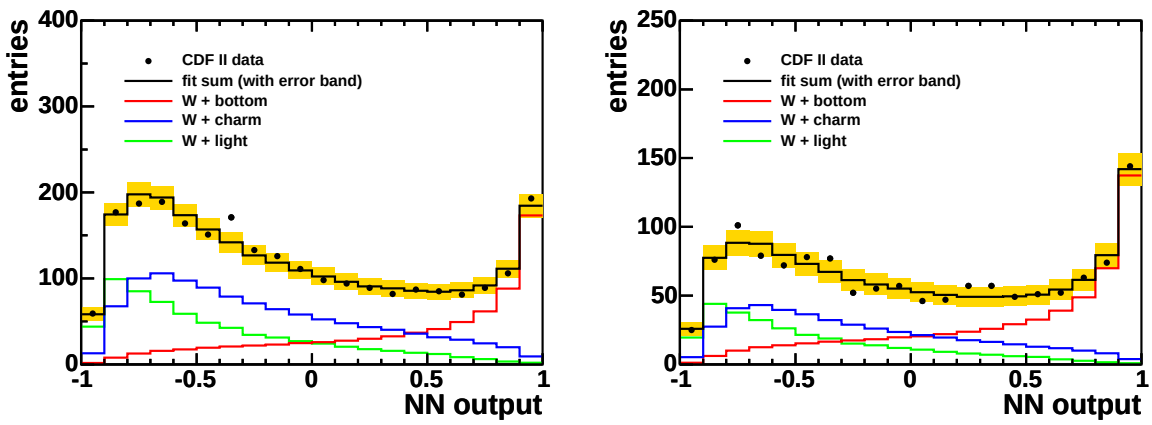


Figure 4.22: Usage of the Flavor Separator output for the estimation of the heavy flavor content in a given data set. In (a) a template fit to the output distribution of events with one jet is shown. In (b) the fit is performed on the output of events with two jets. In both cases one jet has to have a SecVtx tag.

fractions for b and c jets quoted in table 4.6 are already scaled with a correction factor $K = 1.4 \pm 0.4$. This factor has been derived from a fit to the Flavor Separator output distribution in 1.5 fb^{-}1 of data. In contrast to our fit, the fraction of light-quark jets has been fixed to its prediction in this fit [63].

	W+1 jet		W+2 jet	
flavor	fitted fraction [%]	pred. fraction [%]	fitted fraction [%]	pred. fraction [%]
b	29.5 ± 1.8	29.0 ± 7.6	44.1 ± 2.8	45.4 ± 8.7
c	43.6 ± 4.6	37.4 ± 11.3	33.4 ± 6.2	31.3 ± 8.3
light	26.9 ± 3.8	33.6 ± 5.7	22.5 ± 4.9	23.3 ± 3.8

Table 4.6: The fitted fractions are compared with the predicted numbers. Within the given uncertainties the fit results are well comparable with the predicted numbers.

4.4 A neural network to identify c-quark jets

Especially for the estimation of the heavy flavor content, a better separation of the output for c jets and light jets would increase the sensitivity of the fit. For this purpose we train a further neural network using the same input variables and the same Monte Carlo samples as for the Flavor Separator training. In this training, c jets are assigned to signal, light-jets to background. After preprocessing, the 26 variables shown in table 4.7 survive a 3σ cut on the relative significance and are therefore used for the network training. In figure 4.23 one can see the result of the network training. A very good separation between signal and background can be found.

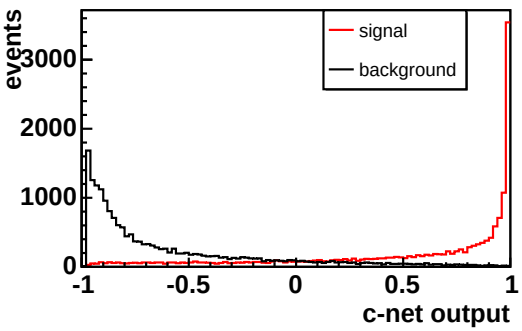


Figure 4.23: Result of the c-net training. A very good separation of c-jets and light-jets can be found.

This, so-called c-net, together with the Flavor Separator enables to identify b jets, c jets and light jets in a two dimensional plane. Figure 4.24 shows the distributions of c-net output versus Flavor Separator output for the three jet flavors. As can be seen b jets accumulate in the upper right corner whereas light jets have outputs next to $(-1, -1)$. Moreover c jets are spread along the upper border and have some small contributions to the region in the lower left corner.

Rank	Variable	Significance (in σ)
1	p_{T} at secondary vertex over E_{T} of jet	113.9
2	number of good tracks in jet	38.2
3	pseudolifetime at secondary vertex	31.0
4	significance of the 3D decay length	25.1
5	signed d_0 of the 2nd track	18.4
6	χ^2 per degree of freedom of SecVtx fit (pass1)	17.9
7	χ^2 per degree of freedom of SecVtx fit (pass2)	16.5
8	η of the jet	15.7
9	signed d_0 of the 3rd track	14.7
10	E_{T} of the jet	13.8
11	signed d_0 of the 1st track	14.3
12	charge at secondary vertex (pass2)	12.0
13	p_{T} of the 3rd track	12.3
14	charge at secondary vertex (pass1)	10.0
15	p_{T} of the 1st track	7.8
16	η with respect to the jet axis of the 2nd track	6.3
17	d_0 of the 1st track	5.6
18	number of muons in the jet	5.5
19	number of used tracks in fit (pass1)	5.3
20	p_{T} at secondary vertex (pass2)	3.8
21	p_{T} at secondary vertex (pass1)	5.1
22	d_0 significance of the 2nd track	2.9
23	d_0 of the 2nd track	5.0
24	p_{T} with respect to the jet axis of the 1st track	4.0
25	signed d_0 significance of the 1st track	4.1
26	number of used tracks in fit (pass2)	3.2

Table 4.7: Set of discriminating variables used for the training of the c-net sorted by relevance. The quoted relative significances are determined as described in section 3.1.1. These variables pass a cut on the significance $> 3.0\sigma$.

The two dimensional output templates can be used for 2D fits of the heavy flavor fractions. For this purpose we again allow the fraction of b, c and light jets to float freely in the fit of the 2D output distribution of events with exactly one jet and, for comparison, events with two jets. In both cases one jet has to be tagged by the SecVtx algorithm. Figure 4.25 illustrates the fit results of the two fits, projected on the c-net and Flavor Separator axes respectively. The yellow band again represents the uncertainties on the fitted distributions. Table 4.8 lists the fitted and predicted fractions. Within the given uncertainties the fit results are well comparable with the predicted numbers. Also the fractions obtained from the Flavor Separator output fit in the last section match quite well the numbers found in the 2D fit. The advantage of the 2D fit is based on the uncertainties of the c jet and light jet fraction, which are

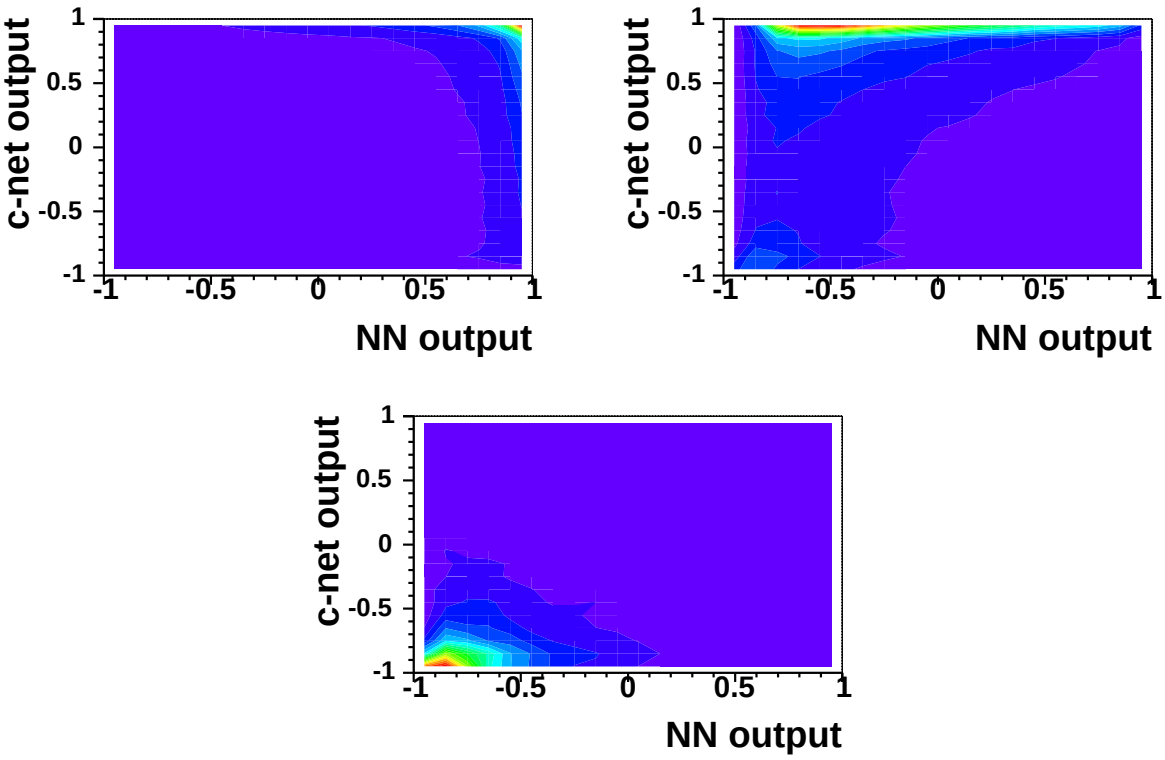


Figure 4.24: Two dimensional output distributions of b, c, and *light* jets. The c-net output is plotted on the ordinate, the Flavor Separator output on the abscissa. Red regions represent a high point density, however blue regions represent low point densities.

roughly halved.

For the validation of the c-net output for c jets one needs a sample of observed events which is enriched with jets originating from c-quarks. Unfortunately this is very hard to realize and hasen't been done so far. Therefore the c-net output cannot be validated and hence the c-net hasen't been used in any analyses so far. For the 2D fit explained above the output for c jets and also for light jets is uncorrected. Nevertheless the c-net together with the Flavor Separator offers a further probability to estimate the heavy flavor composition of a given data set.

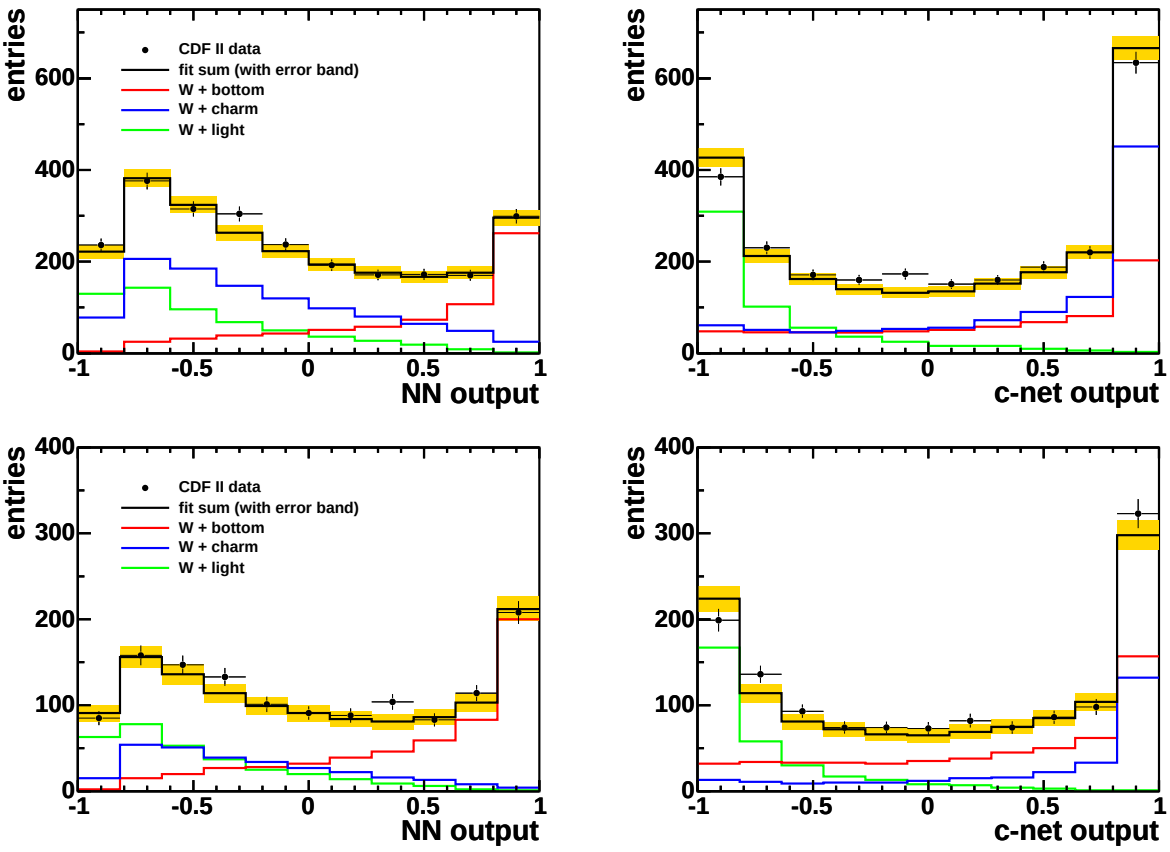


Figure 4.25: The figures in the upper row show the templates and the fit result for events with exactly one jet, projected on the Flavor Separator (left hand side) and the c-net axes (right hand side). In the lower row the same plots are illustrated for the fit to the output distribution of events with two jets. In both cases one jet has to have a SecVtx tag.

	W+1 jet		W+2 jet	
flavor	fitted fraction [%]	pred. fraction [%]	fitted fraction [%]	pred. fraction [%]
b	30.1 ± 1.7	29.0 ± 7.6	46.2 ± 2.7	45.4 ± 8.7
c	44.6 ± 2.1	37.4 ± 11.3	26.2 ± 2.5	31.3 ± 8.3
light	25.4 ± 1.5	33.6 ± 5.7	27.6 ± 2.0	23.3 ± 3.8

Table 4.8: The fractions obtained from the 2D fit are compared with the predicted numbers. Within the given uncertainties the fit results are well comparable with the predicted numbers and with the fractions resulting from the Flavor Separator fit only.

Chapter 5

The single-top analyses

In this chapter we give a brief overview over the three CDF single-top analyses which have been performed so far. First we report which data samples have been used and how single-top candidates have been selected. In the second step the background estimation for the selected candidates is explained. Finally we discuss the three multivariate analyses, namely the Likelihood Function Method (LF), the Matrix Element Method (ME) and the Neural Network Method (NN), in more detail.

5.1 Event selection and utilized data samples

As can be seen in the figures 1.4 and 1.5 the single-top final state contains a real W boson, one or two bottom quarks, and possibly additional light quark jets. As a result of the event kinematics of a t-chanel single-top event, the second b quark in t-channel events is primarily produced in forward direction. Thus, in most events only two jets, the b and the light quark jet or the b and the second b jet in single-top s-channel events respectively, can be observed. The acceptance of the candidate selection can be increased by also including events with three jets where at least one jet is tagged by SecVtx. To reduce multijet backgrounds, the W boson stemming from the top quark decay is required to decay leptonically into electrons or muons. Leptonic W-boson decays in $\tau\nu_\tau$ are not included because of CDF's low τ identification efficiency. However, W boson which decay into $\tau\nu_\tau$ where the τ decays further into e/μ $\nu_{e/\mu}$ enter the event selection with a small fraction. Summarised, selected events are required to have:

- two or three jets, whereas at least one jet is identified as a b-jet with a displaced secondary vertex
- exactly one isolated lepton (see lepton identification in subsection 5.1.2)
- large missing transverse energy from the non-detectable neutrino.

5.1.1 Data samples

Since the W boson is required to decay leptonically, the analyses described below use data which passes the high-p_{T} lepton trigger. Central electrons have to pass the Level-3 ELECTRON_CENTRAL_18 trigger that requires a COT track with $p_{\mathrm{T}} > 9$ GeV/c matching an energy deposition $E_{\mathrm{T}} > 18$ GeV in the CEM. The MET_PEM trigger, which forward electrons have to pass, demands an energy deposition of 20 GeV in the PEM. Muons have to satisfy the requirements of the L3 MUON_CMUP_18 or MUON_CMX_18 trigger (COT track with $p_{\mathrm{T}} > 18$ GeV/c matched to a track segment in the corresponding muon chamber).
The data samples get their names from the trigger they have passed, i.e. *bhel* data are events fulfilling the central electron trigger, *bpel* data samples contain forward electrons and muon data is integrated in *bhmu* samples. Furthermore the data is reprocessed offline before being used in any physics analyses. During the processing with CDFSOFT2 version 5.3.3 (*0d* samples) or CDFSOFT2 version 6.1.2 (*0h, 0i* and *0j* data sets) the online calibrations are checked and improved, the misalignment of the silicon detector is corrected, tracks are refitted, energy clusters are checked, leptons are identified and jet clustering as well as secondary vertex algorithms are performed. Because of hardware or software problems, some detector components may be inoperative during accelerator stores. The runs, where all the detector elements necessary for the single-top analyses were operating, are listed in the so-called "goodrun" list [64]. The mentioned analyses have used the single-top goodrun list version 19 and data which have been taken from February 2002 till August 2007. The largest data samples (CEM and CMUP) correspond to an integrated luminosity of 2.2 fb^{-1}.

5.1.2 Requests for single-top candidates

Lepton identification

Leptons are divided in two groups: isolated and non-isolated leptons. A lepton is called isolated, if the energy in a cone with $R = 0.4$ around the reconstructed lepton, excluding the lepton itself, is smaller than ten percent of the lepton energy.
Depending on the subdetector in which a lepton is detected, four lepton-types, namely CEM, PHX, CMUP, CMX, are defined. Central electrons are detected in the CEM and their pseudorapidity range is therefore $|\eta| \leq 1.1$. The energy deposition in the calorimeter tower has to be larger than 20 GeV and the reconstructed track $p_{\mathrm{T}} \geq 10$ GeV/c. In addition, the ratio of the energy in the hadronic calorimeter and the energy in the electromagnetic calorimeter ($E_{\mathrm{HAD}}/E_{\mathrm{EM}}$) for the cluster matching the electron candidate, has to be smaller than $(0.055+(0.00045\times E))$. The PEM calorimeter detects electrons in the forward-backward region whereas only electrons with $1.1 \leq |\eta| \leq 1.6$ are taken into account. Furthermore, $E_{\mathrm{T}} \geq 20$ GeV, $E_{\mathrm{HAD}}/E_{\mathrm{EM}} \leq 0.05$ and a track with at least three $r-\phi$ hits in the silicon which points from the calorimeter cluster to the primary vertex (Phoenix track), are required for forward electrons.
Events, containing an electron track and a further track which possesses positron char-

acteristics, are rejected if the two tracks form a common vertex, because these elcetrons are likely to originate from photon conversion processes. More details about the electron selection can be found in reference [65].
Muons have to have a COT track with $p_{\mathrm{T}} > 20$ GeV/c which can be extrapolated to a muon chamber segment and a pseudorapidity $|\eta| < 1.0$. Again, further requirements can be found in [66]. When identifying muons additionlly a further problem has to be solved. Cosmic muons or muons coming from decays of long lived particles have to be rejected. One possibility for this is to ask for impact parameters $d_0 \leq 0.2$ cm [66]. Additionally the "unwanted" muons can be identified through their characteristic track topology.
Furthermore, one has to differentiate between tight leptons, which are isolated and fulfill all the requirements mentioned above, and loose leptons, which are non-isolated and fulfill less quality requirements. An event is identified as a single-top candidate if exactly one tight lepton is found, i.e. there is neither another tight lepton nor a further loose lepton allowed (dilepton-veto).

Jet selection

Hadronic jets are reconstructed as explained in section 3.3. For this analysis the jet energy is corrected up to hadron level. A jet candidate is required to have $|\eta_{\mathrm{det}}| < 2.8$, where η_{det} is defined as the pseudorapidity of the jet, calculated with respect to the origin of the coordinate system which is located in the center of the detector. Jets are called "tight", if they fulfill $E_{\mathrm{T}} > 20$ GeV, however "loose" jets have an E_{T} between 12 GeV and 20 GeV. Two or three tight jets are required to be present in a single-top candidate event. A large fraction of the backgrounds is removed by demanding at least one of these jets to be tagged as a b-quark jet by using displaced secondary vertex information.

Missing transverse energy

Since the neutrino, originating from the W-boson decay, cannot be detected with any part of the CDFII detector, there has to appear a lot of missing transverse energy ($\not\!\!E_{\mathrm{T}}$). The $\not\!\!E_{\mathrm{T}}$ is defined as the magnitude of $\vec{\not\!\!E}_{\mathrm{T}} = -\sum_i E_i \hat{n}_i$, where $\hat{n}_i$ is a unit vector parallel to the transverse component of the vector pointing at the ith calorimeter tower, and E_i is the energy therein. Missing transverse energy determined this way has to be corrected because of two reasons: On one hand the jet energies are corrected for the energy scale, thus this corrections have to be taken into account also for $\not\!\!E_{\mathrm{T}}$. On the other hand, muons pass the calorimeters without showering. Therefore the transverse energy of the muons has to be added to the sum above and the small average muon ionization energy has to be removed. The corrected $\not\!\!E_{\mathrm{T}}$ is required to be greater than 25 GeV for single-top candidate events.

Z-boson veto

The Z-boson veto is used to reduce the background due to Z+jets production. If the tight lepton together with another jet or lepton can be formed to an object with the invariant mass of the Z boson, the event is rejected. For this purpose the Z-boson mass is assumed to lie between 76 GeV/c^2 and 106 GeV/c^2.

QCD veto

Events in which no real W boson appears (often called QCD-multijet events), have to be further suppressed by additional cuts. These events can pass the requirements explained above, because of jets which are misidentified as leptons or because of leptons originating from semileptonic heavy quark decays. Two data based samples are used to model QCD-multijet events: The jet-electron and the anti-electron sample [62]. Monte Carlo samples of all expected backgrounds except for the QCD-multijet model are fitted to W+jets data, on which no secondary vertex algortihm is applied (pretag sample). By looking at different kinematic variables, the differences between measured and predicted distributions, for the four lepton types (CEM, PHX, CMUP and CMX) separately, are identified as QCD-multijet background. The variables, showing the largest discrepancies, are the missing transverse energy $\not\!\!E_{\mathrm{T}}$, the $\not\!\!E_{\mathrm{T}}$-significance $\not\!\!E_{\mathrm{T,sig}}$, the transverse mass of the reconstructed W boson $M_{\mathrm{T,W}}$ and, depending on the lepton type, the azimuthal angle between $\vec{\not\!\!E}_{\mathrm{T}}$ and one jet. $\not\!\!E_{\mathrm{T,sig}}$ is defined as

$$\not\!\!E_{\mathrm{T,sig}} \equiv \frac{\not\!\!E_{\mathrm{T}}}{\sqrt{\sum_{\mathrm{jets}} C_{\mathrm{JES}}^2 \cdot \cos^2(\Delta\phi_{\mathrm{jet}}, \not\!\!E_{\mathrm{T}}) + \cos^2(\Delta\phi_{\mathrm{raw,corr}})}} \ , \tag{5.1}$$

where C_{JES} is the jet energy correction factor and $\Delta\phi_{\mathrm{raw,corr}}$ is the azimuthal angle between corrected and uncorrected missing transverse energy. The transverse mass of the W boson is given by

$$M_{\mathrm{T,W}} = \sqrt{2p_{\mathrm{T,lep}}\not\!\!E_{\mathrm{T}} - \vec{p}_{\mathrm{T,lep}}\vec{\not\!\!E}_{\mathrm{T}}} \ , \tag{5.2}$$

where $p_{\mathrm{T,lep}}$ is the transverse momentum of the charged lepton. Exemplarily, figure 5.1 shows the two dimensional distribution of $\not\!\!E_{\mathrm{T,sig}}$ and $M_{\mathrm{T,W}}$ for events with a central electron and two jets. The black line in the bottom left corner illustrates the applied "triangle" cut. A more detailed description of the QCD veto can be found in [62].
The following requirements are applied to single-top event candidates:

- Events passing the electron triggers must have $M_{\mathrm{T,W}} > 20$ GeV/c^2
- Central electrons are required to have $\not\!\!E_{\mathrm{T,sig}} > -0.05\ c^2/\mathrm{GeV}\cdot M_{\mathrm{T,W}} + 3.5$ and events with two or three jets have to fulfill $\not\!\!E_{\mathrm{T,sig}} > 2.5 - 2.5\cdot|\Delta\phi|/0.8$ where $\Delta\phi$ is the azimuthal angle between $\vec{\not\!\!E}_{\mathrm{T}}$ and the jet with the second largest uncorrected E_{T} (second leading jet).

- Forward electrons must have $\not\!\!E_{\mathrm{T,sig}} > 2.0$ and, for the first two leading jets, $\not\!\!E_{\mathrm{T}} > 45$ GeV-30 GeV$\cdot|\Delta\phi|$, where $\Delta\phi$ is the azimuthal angle between $\vec{\not\!\!E}_{\mathrm{T}}$ and the corresponding jet, has to be fulfilled.

- The transverse W-boson mass of central muon events has to exceed 10 GeV. Furthermore events with a CMUP muon and one jet are rejected, if $\not\!\!E_{\mathrm{T}} < 60$ GeV$\cdot|\Delta\phi| - 145$ GeV. Here, $\Delta\phi$ is the angle between the tight lepton and the jet in the x-$y-$plane.

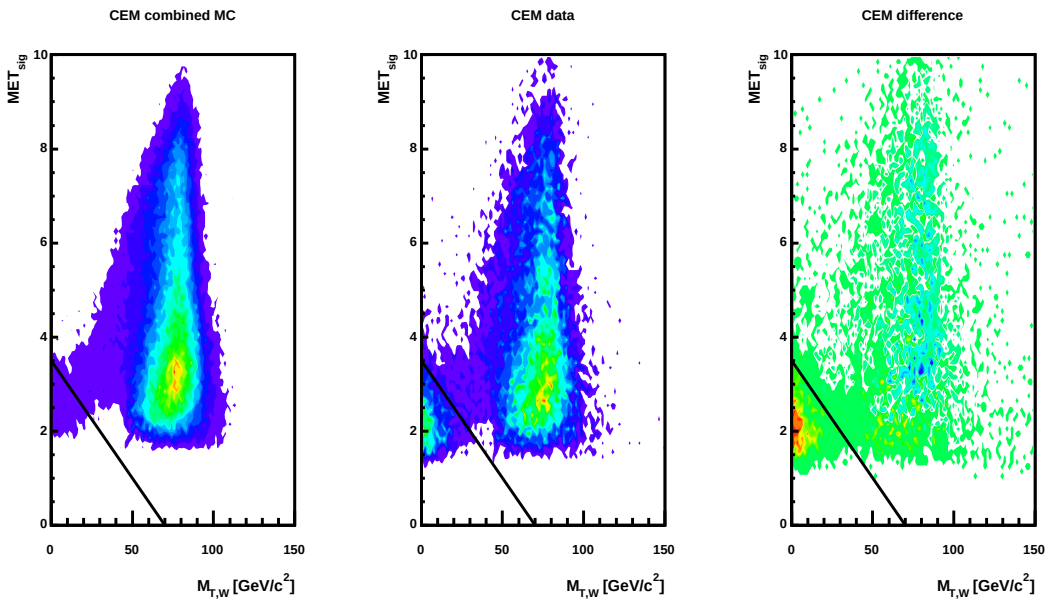
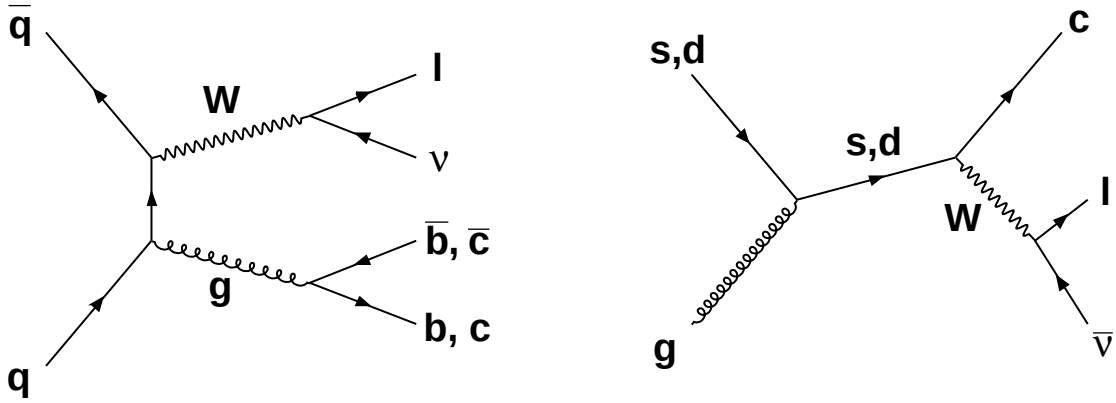


Figure 5.1: Illustration of the QCD veto in the central electron sample with 2 jets. The distribution of $\not\!\!E_{\mathrm{T,sig}}$ versus $M_{\mathrm{T,W}}$ is shown. On the left hand side, the predicted distribution of the W+jets sample is shown. The plot in the middle shows the distribution for events without applying a SecVtx-tag. On the right hand side, the difference between prediction and observation is shown. The line represents the cut $\not\!\!E_{\mathrm{T,sig}} > -0.05$ GeV$-1 \cdot M_{\mathrm{T,W}} + 3.5$.

5.1.3 Backgrounds to single-top production

After the above mentioned event selection cuts, the ratio between single-top and background events is still less than S/B $\approx 1/15$. The background is dominated by W-boson production in association with heavy-flavor quarks. Figures 5.2(a) and 5.2(b) show examplary Feynman diagrams for $Wb\bar{b}$, $Wc\bar{c}$ and Wc production. Further backgrounds are events from top-pair production (see 1.2.1) and the diboson processes WW, WZ

and ZZ. An example for diboson production is illustrated in figure 5.3(a). The production of a Z boson in addition with jets, in the following called Z+jets, is shown in figure 5.3(b). For the decay of the Z boson only leptonic processes are considered. Each event with a W or Z boson and a light quark jet, which is falsely identified as a b-jet, contributes to the background called mistags. Furthermore the QCD-multijet processes, discussed for the QCD veto, have still a substantial contribution to the background. A QCD process in which a jet can be misidentified as an isolated lepton is illustrated in figure 5.3(c). Events like the one shown in figure 5.3(d) fake a single-top event structure due to semi-leptonic bottom quark decays. We won't discuss the Monte Carlo samples used to model the different backgrounds in this thesis. A detailed description of these can be found in [73, 74].



(a) (b)

Figure 5.2: Exemplary Feynman diagrams of background processes to single top-quark production: (a) $Wb\bar{b}, Wc\bar{c}$, (b) Wc,

5.2 Signal and background estimate

The expected number of events for the different signal and background processes are derived from the so-called Method II, which is a procedure to calculate the normalization of processes in the SecVtx tagged lepton plus jets data set. Method II assumes that all processes, contributing to the lepton plus jets sample, are known. In case of the single-top analyses these are, $t\bar{t}$, single-top, electroweak, QCD and W+jets processes. Below, we give a short description of the Method II procedure. Details can be found in [67].

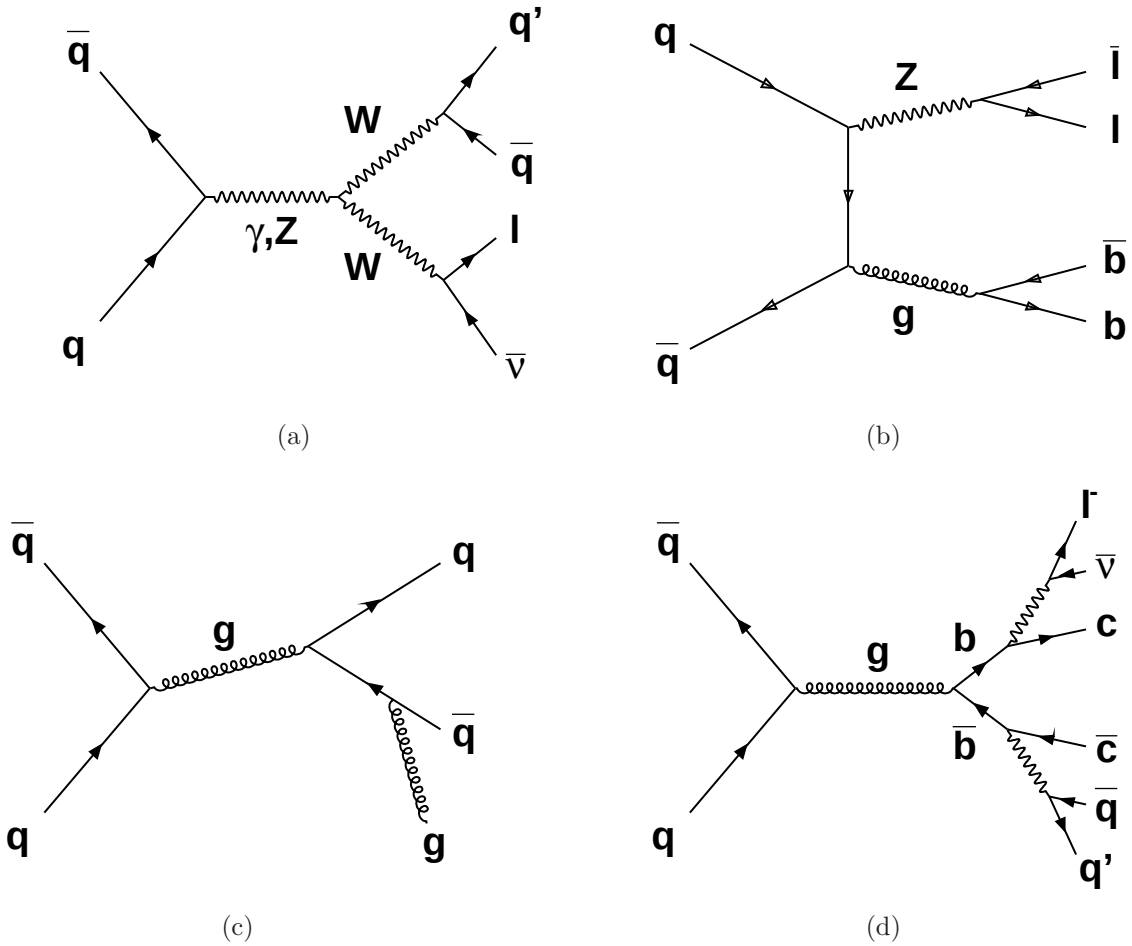


Figure 5.3: Exemplary feynman diagrams of background processes to single top-quark production: (a) WW, as an example for the diboson background (b) Z+jets, (c) QCD-multijet background, where a gluon jet can be misidentified as an isolated lepton, (d) QCD-multijet background from semileptonic heavy quark decays.

5.2.1 Expected single top-quark events

In general the number of expected signal events is calculated by

$$N_{\text{signal}} = \sigma \cdot \epsilon_{\text{evt}} \cdot \int \mathfrak{L}\, dt \ , \tag{5.3}$$

where σ is the single-top production cross section given in section 1.2.2, ϵ_{evt} is the event detection efficiency and $\int \mathfrak{L}\, dt$ the integrated luminosity. The event detection efficiency cannot simply be calculated by the ratio of the number of selected events and the number of generated events from Monte Carlo samples, because trigger, b-tagging and lepton identification efficiencies are different for meaured and simulated samples. Hence, ϵ_{evt} has to be calculated via

$$\epsilon_{\text{evt}} = \epsilon_{\text{evt}}^{\text{MC}} \cdot \epsilon_{\text{BR}} \cdot \frac{\epsilon_{z0}^{\text{data}}}{\epsilon_{z0}^{\text{MC}}} \cdot \frac{\epsilon_{\text{leptonidreco}}^{\text{data}}}{\epsilon_{\text{leptonidreco}}^{\text{MC}}} \cdot \frac{\epsilon_{\text{tag}}^{\text{data}}}{\epsilon_{\text{tag}}^{\text{MC}}} \cdot \epsilon_{trigger} \ . \tag{5.4}$$

Here $\epsilon_{\text{evt}}^{\text{MC}}$ is the event detection efficiency derived from Monte Carlo. The branching ratio for the leptonic decay of the W bosons is taken into account by $\epsilon_{\text{BR}} = 0.3257 \pm 0.0028$ [2]. $\frac{\epsilon_{z0}^{\text{data}}}{\epsilon_{z0}^{\text{MC}}}$ incorporates the z-vertex cut efficiency, for data $\epsilon_{z0}^{\text{data}} = 0.958 \pm 0.002$ as given in [68], for Monte Carlo $\epsilon_{z0}^{\text{MC}}$ is taken from simulated events. The differences in the identification and reconstruction of leptons for each trigger are included by the $\frac{\epsilon_{\text{leptonidreco}}^{\text{data}}}{\epsilon_{\text{leptonidreco}}^{\text{MC}}}$ factor. These factors are determined by averaging over the numbers for the different run ranges shown in [66, 69]. Different b-tagging efficiencies are introduced via $\text{SF}_{\text{tag}} = \frac{\epsilon_{\text{tag}}^{\text{data}}}{\epsilon_{\text{tag}}^{\text{MC}}}$ which is different for events with one tag, two tags and three tags respectively. The value for single tag events is given in subsection 3.3.4 and a method for calculating numbers of double and tripple tag events is explained in [65]. Finally the last factor in eq. 5.4 incorporates the trigger efficiencies which are different for each trigger [70, 71, 72, 66]. Again a value averaged over different run ranges is used. The estimated number of single-top events in 2.2 fb^{-1} is listed in table 5.1.

5.2.2 Estimation of expected background events

The estimation of expected background events is done in three steps. First the number of electroweak (WW, WZ, ZZ, $Z + jets$) and $t\bar{t}$ events is calculated in the same way as the number of signal events in the previous subsection, since their cross section can be calculated theoretically. Furthermore, the contribution of the QCD-multijet background to the sample of events, to which no secondary vertex algorithm has been applied, is determined by fitting the $\not\!\!E_{\text{T}}$ distibution of a QCD-multijet template and a W+jets Monte Carlo sample to data. In the tagged sample the whole expected background (excluding QCD-multijets) and the QCD-multijet model are fitted to the $\not\!\!E_{\text{T}}$ distribution, to evaluate the QCD-multijet background fraction. In both scenarios the number of QCD-multijet events is simply calculated by

$$N_{\text{QCD}} = F_{\text{QCD}} \cdot N_{\text{evt}} \ . \tag{5.5}$$

The number of expected W+jets events in the pretag sample, $N^{\text{pre}}_{W+\text{jets}}$, can be calculated by substracting the number of QCD-multijet ($N_{\text{pre}} \cdot (1 - F^{\text{pre}}_{\text{QCD}})$), electroweak ($N^{\text{pre}}_{\text{ewk}}$) and top quark ($N^{\text{pre}}_{\text{top}}$) events in the pretag sample from the number of pretag events:

$$N^{\text{pre}}_{W+\text{jets}} = N_{\text{pre}} \cdot (1 - F^{\text{pre}}_{\text{QCD}}) - N^{\text{pre}}_{\text{ewk}} - N^{\text{pre}}_{\text{top}} \ . \tag{5.6}$$

To obtain the number of expected W+jets events in the tagged sample one has to differentiate between W+heavy flavor (W+hf) and W+light (W+lf) flavor processes. The contribution of W+hf events is calculated via

$$N^{\text{tag}}_{W+\text{hf}} = (N_{\text{pre}}(1 - F_{\text{QCD}}) - N_{\text{ewk}} - N_{\text{singletop}} - N_{t\bar{t}}) \cdot f_{\text{HF}} \cdot K \cdot \epsilon_{\text{tag}} \ , \tag{5.7}$$

where K is the scale factor mentioned in section 4.3 and f_{HF} is the fraction of W+hf events in the pretagged sample which is derived from Monte Carlo simulations. The tagging efficiency ϵ_{tag} is different for $Wb\bar{b}$, $Wc\bar{c}$ and Wc processes leading to a different amount of events for these three backgrounds. Using again the model for light quark jets explained in subsection 3.3.4, the number of $W + \text{lf}$ events can be calculated by

$$N^{\text{tag}}_{W+\text{lf}} = \frac{N^{+}_{\text{light}}}{N_{\text{pre}}} \cdot (N^{\text{pre}} - N^{\text{pre}}_{t\bar{t}} - N^{\text{pre}}_{\text{QCD}} - N^{\text{pre}}_{W+\text{hf}} - N^{\text{pre}}_{\text{ewk}} - N^{\text{pre}}_{\text{singletop}}) \ , \tag{5.8}$$

where N^{+}_{light} is the number of light jets with a positive tag. The results of the background estimate for 2.2 fb^{-1} are shown in table 5.1.

5.3 Multivariate analyses for the single-top search

In this section we give a short introduction to the different single-top analyses. The main ideas of the analyses technique are explained and the discriminants which are finally used to separate single top-quarks from the background are shown. In a first step all analyses measure the combined single top-quark production cross section, i.e. the cross section of s-channel and t-channel together. Nevertheless also the cross sections of the different channels can be measured separately. The treatment of systematic uncertainties and the statistic interpretation are the same for all analyses. Since they are also the same in our combination method, they are discussed in chapter 6. The results of the different analyses are also given in chapter 6 together with the results of the combination.

5.3.1 The Likelihood Function method

This analysis combines different variables with a likelihood function in order to build a discriminant which can be used to discover a single top-quark signal. Two different likelihood function discriminants have been developed, one for events with two jets as well as one for events with three jets. In both cases the discriminant is optimized for finding single-top t-channel events [73]. As input to the discriminants basic quantities,

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

Table 5.1: Summary of predicted numbers of signal and background events in the selected data sample (2.2 fb^{-1}).

like the E_{T} of a jet, as well as quantities of reconstructed objects, like the top quark mass, are used. Therefore a method that resolves reconstruction ambiguities has been developed, the so-called kinematic solver.

The Kinematic Solver

The χ^2, resulting from a reconstruction method called the kinematic solver, is one of the input variables to the likelihood function. Here, the masses of the reconstructed W boson from the top quark decay ($M_{l\nu}$) and the mass of the top quark ($M_{l\nu b}$) are constrained to 80.4 GeV and 175 GeV respectively. Additionally the lepton momentum vector, the p_{T} of the top quark and the direction and invariant mass of the b-jet, coming from the top quark decay, are constrained to their measured values. The output from the kinematic solver is a set of neutrino and b-jet four-vector solutions, which are used to reconstruct kinematic variables. Also provided is the χ^2 output mentioned above, which indicates how far from the measured values of the b-jet energy and the $\not{E}_{\mathrm{T}}$ the solver found its solutions.

Likelihood function technique

In a first step each variable i is filled in a histogram with n_i bins. The number of entries in each bins is denoted by f_{ijk} whereas j is the bin number and k is the event class.

There are five possible event classes: The signal has $k = 1$ and the backgrounds $Wb\bar{b}$, $t\bar{t}$, $Wc\bar{c}/c$ and mistags have event classes of k=2,3,4 and 5. Furthermore the histograms are normalized, i.e. $\sum_{j=1}^{n_i} f_{ijk} = 1$ for all i and all k. The likelihood function for an event is computed by evaluating in which bin j_i the event falls in the distribution of variable i, and computing

$$p_{ik} = \frac{f_{ij_ik}}{\sum_{m=1}^{5} f_{ij_ik}} \ , \tag{5.9}$$

which is used to calculate the likelihood function

$$L_k(\{x_i\}) = \frac{\prod_{i=1}^{n_{var}} p_{ik}}{\sum_{m=1}^{5} \prod_{i=1}^{n_{var}} p_{ik}} \ . \tag{5.10}$$

t-channel likelihood function, 2 and 3 jet bin

The likelihood function ($L_{\text{tchan2jet}}$) for events with two jets uses the single-top t-channel process as signal. The outcome of $L_{\text{tchan2jet}}$ for events with two tagged jets is set to zero because only 1% of the t-channel events are expected to have two tags. In the following the seven input variables used for the likelihood are listed:

- H_{T}: the scalar sum of the transverse energies of the two jets, the lepton, and the missing transverse energy, $H_{\text{T}} = \sum_{\text{jets}} E_{\text{T}} + p_{\text{T}}^{l} + \not{E}_{\text{T}}$, where $\sum_{\text{jets}} E_{\text{T}}$ is the sum of the transverse energies of the jets.
- $Q_l \cdot \eta_{lj}$: lepton charge times pseudorapidity of the untagged jet
- χ_t^2: the χ^2 value of the kinematic solver
- $\cos\Theta(l, lj)$: cosine of the angle between the lepton and the untagged jet in the top quark rest-frame
- M_{j1j2}: the invariant mass of the two jets.
- $\log(\text{ME}_{\text{t-chan}})$: the logarithm of the MADGRAPH matrix element computed using the constrained four-vectors of the bottom quark, the lepton and the neutrino. This matrix element is a weighted average of $ub \to dt$ matrix element (weight 2/3) and the same matrix element with reversed lepton sign (weight 1/3), in order to simulate the $d\bar{b} \to u\bar{t}$.
- the KIT Flavor Separator output

$L_{\text{tchan3jet}}$, the likelihood function for events with three jets, is computed with ten different input variables:

- $M_{l\nu b}$: the reconstructed top quark mass
- the KIT Flavor Separator output
- the number of jets with a SecVtx tag

- $Q_l \cdot \eta_{lj}$
- the smallest ΔR of any two jets
- M_{jj} of the two jets which are assigned not to stem from the top quark decay
- $\cos\Theta(l, lj)$
- E_{T} of the jet with the lowest E_{T}
- η of the reconstructed W boson
- E_{T} of the jet which has been chosen to be the b-quark jet from the top quark decay

Figure 5.4 shows the output distributions for the two likelihood functions on a logarithmic scale.

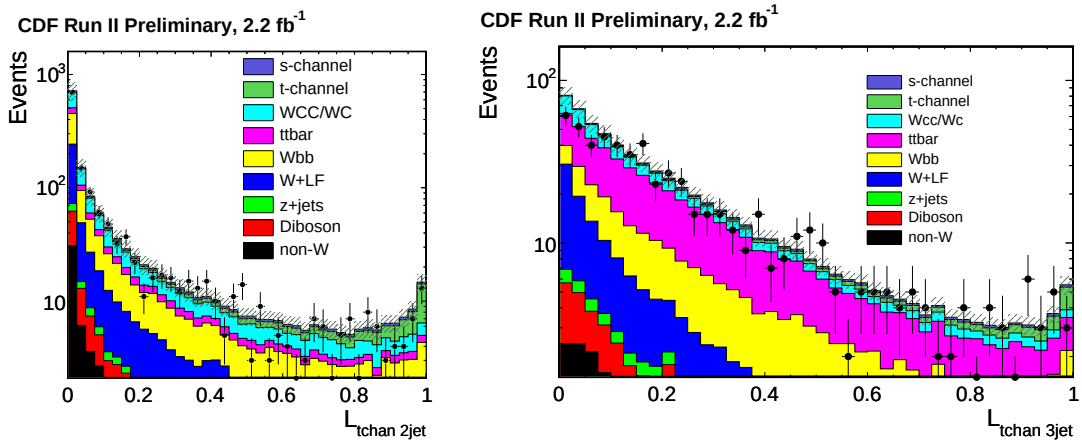


Figure 5.4: Output of the two likelihood discriminants, on the left hand side the one for events with two jets, on the right hand side the one for events with three jets. The number of events is shown on a logarithmic scale and the different processes are normalized to the standard model prediction.

5.3.2 The Matrix Element method

The main idea of this analysis is to calculate from a given set of variables (the 4-vectors of the lepton and the jets) the event probability densities that these variables results from a given underlying physical process (signal or background). Therefore the fully differential cross section of all kinematic variables has to be calculated on an event-by-event basis for the signal and the background hypothesis [75]. In general a differential cross section is given by

$$d\sigma = \frac{(2\pi)^4 |M|^2}{4\sqrt{(q_1 \cdot q_2)^2 - m_{q1}^2 \cdot m_{q2}^2}} d\Phi_n(q_1 + q_2; p_1, ..., p_n) \ , \tag{5.11}$$

where $|M|$ is the matrix element, $m_{q1(2)}$ the masses and $q_{1(2)}$ the momenta of the incident particles. The n-body phase space is represented by

$$d\Phi_n(q_1 + q_2; p_1, ..., p_n) = \delta^4(q_1 + q_2 - \sum_{i=1}^{n} p_i) \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \ . \tag{5.12}$$

Since not all momenta of the inital and final states can be measured, e.g. p_z of the neutrino, the event probability is not simply $P_{\text{evt}} \sim \frac{d\sigma}{\sigma}$ (σ is the total cross section). Rather the differential cross section has to be folded with the parton distribution functions $f(x_{1(2)})$ and integrated over all particle momenta which can't be measured. The mapping between the particle variables (y) and the measured variables (x) is established with the transfer function, $W(y, x)$. All this leads to the following equation for the event probability density:

$$P(x) = \frac{1}{\sigma} \int 2\pi^4 |M|^2 \frac{f(x_1)}{|E_{q1}|} \frac{f(x_2)}{|E_{q2}|} W(y, x) d\Phi_n dE_{q1} dE_{q2} \tag{5.13}$$

Event probabilities are calculated for single-top s- and t-channel processes as well as for $Wb\bar{b}$, $Wc\bar{c}$, Wc+jets, $W + gg$ and $t\bar{t}$ backgrounds. The probability densities for the two single-top processes are added to form a single probability density. These event probabilities and the output of the KIT Flavor Separator (b), scaled to $0 \leq b \leq 1$ are used to construct a discriminant variable for each event:

$$EPD = \frac{b \cdot P_{single-top}}{b \cdot P_{single-top} + b \cdot (P_{Wb\bar{b}} + P_{t\bar{t}}) + (1 - b) \cdot (P_{Wc\bar{c}} + P_{Wcj} + P_{Wgg})} \tag{5.14}$$

Figure 5.5 shows the EPD distribution for events with two and three jets together.

Matrix elements and transfer functions

The matrix elements, $|M|^2$ are calculated at leading order perturbation theory with the HELAS (HELicity Amplitude Subroutines for Feynman Diagram Evaluations) package [76]. Subroutines for the different processes are derived from MadGraph.
The transfer function, $W(y, x)$, represents the probability of measuring the set of observable variables (x) corresponding to the set of produced variables (y). Hence, the set (y) incorporates all final state particle momenta at the particle level, whereas the set (x) incorporates the measured momenta (of the corresponding objects) with the CDF detector. For well measured momenta, like the one of the lepton, the transfer function is taken as a δ-function, i.e. the measured momenta are used in the calculation of the differential cross section. If the detector resolution is not good enough, $W(y, x)$ is a Gaussian function. For particle momenta that cannot be measured at all, the transfer function is unity. A detailed description of the calculation of transfer functions is given in [75].

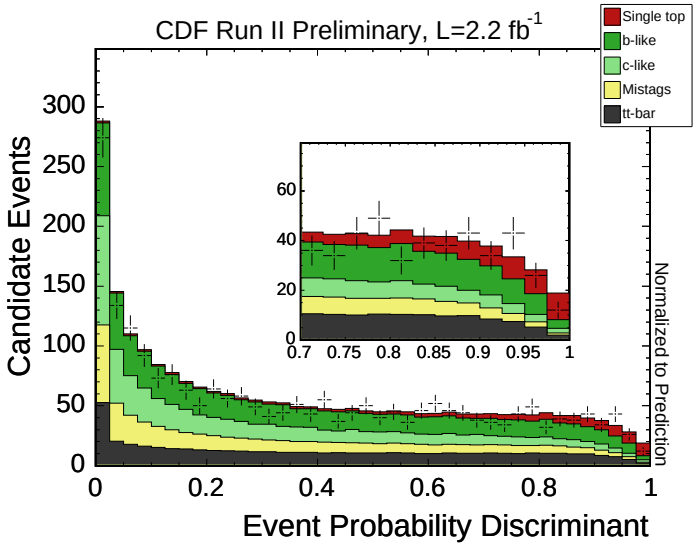


Figure 5.5: EPD output distributions for signal and background events, normalized to the predicted number of events for 2.2 fb^{-1}

Non triggered muons

The Matrix Element method is currently the only single-top analysis that includes events with leptons of a further lepton type, called loose muons. All analyses relying on the identification of muons in the final state are limited by the acceptance of the two high-p_{T} muon triggers. Also using muons of the well defined loose muon categories would increase the sensitivities of those analyses, but unfortunately there are no high efficiency triggers for the selection of this events. The loose muon events used in this analysis are selected with the missing transverse energy trigger by requiring additionally two jets with $E_{\mathrm{T}} > 25$ GeV. One of these jets has to be a central jet, i.e $|\eta| < 0.9$ [77]. These trigger requirements are reasonable, because the $\not\!\!E_{\mathrm{T}}$ at trigger-level is uncorrected. Therefore the minimum ionizing muons can enhance the $\not\!\!E_{\mathrm{T}}$ of an event over the true $\not\!\!E_{\mathrm{T}}$ due to the neutrino escaping the detection.

5.3.3 The Neural Network method

The neural network analysis uses the Neurobayes® package, introduced in section 3.1.1, to combine the information of many kinematic or event shape variables to a powerful discriminant. Overall four different networks have been trained, one for events with two or three jets and one or two tags, respectively [74]. All networks have been trained with 50% background and 50% signal events whereas the background is composed like it is predicted. The signal samples have been built only of t-channel events except of the neural network for two jets and two tags, which is trained only with s-channel events.

For all neural network trainings mentioned above, three categories of input variables are used: some are directly measured in the detector, others are reconstructed out of measured properties, and a few are calculated by advanced algorithms like the KIT Flavor Separator. For this purpose the top quark four-momentum is built out of the reconstructed W boson and a jet defined as b-jet. For the networks trained with t-channel events the tagged jet is used as the b-jet from the top quark decay. For the network trained with s-channel events, the jet with the largest product of the charge of the lepton and the jet pseudorapidity, $Q \cdot \eta$, is assigned as the b-jet from the top quark decay.

Table 5.2 lists the five most significant input variables of the individual networks. In the following a more detailed describtion of these variables is given:

rank	2jets-1tag	2jets-2tags	3jets-1tag	3jets-2tags
1	$M_{l\nu b}$	$M_{l\nu j1j2}$	$Q_l \cdot \eta_{lj}$	$Q_l \cdot \eta_{lj}$
2	KIT Flav. Sep.	$M_{T,W}$	KIT Flav. sep.	$M_{l\nu bb}$
3	M_{j1j2}	$M_{T,l\nu b}$	H_T	$p_{T,l\nu bjj}$
4	$Q_l \cdot \eta_{lj}$	$\cos\Theta(j,j)$	M_{j1j3}	M_{j1j2}
5	$M_{T,l\nu b}$	$M_{l\nu b}$	$M_{l\nu b}$	$\cos\Theta(l,lj)$

Table 5.2: The five most significant input variables of the four different neural networks. The abbreviations are explained in the text.

- $M_{l\nu b}$: mass of the reconstructed top quark
- KIT Flav. Sep.: output of the KIT Flavor Separator for the b-quark jet from the top quark decay
- M_{j1j2}: invariant mass of the two most energetic jets
- $Q_l \cdot \eta_{lj}$: lepton charge times pseudorapidity of the untagged jet
- $M_{T,l\nu b}$: transverse mass of the reconstructed top quark
- $M_{l\nu j1j2}$: invariant mass of the lepton-neutrino-jet-jet system
- $M_{T,W}$: transverse mass of the reconstructed W boson
- $\cos\Theta(j,j)$: cosine of the angle between the two jets in the top rest-frame
- H_T: the scalar sum of the transverse energies of the two jets, the lepton, and the missing transverse energy, $H_T = \sum_{jets} E_{\mathrm{T}} + p^l_{\mathrm{T}} + \not{E}_{\mathrm{T}}$, where $\sum_{\mathrm{jets}} E_{\mathrm{T}}$ is the sum of the transverse energies of the jets.
- M_{j1j3}: invariant mass of the system first and third leading jet

- $M_{l\nu bb}$: invariant mass of the reconstructed top quark and second bottom quark, which corresponds to $\sqrt{\hat{s}}$ of s-channel single top-quark events.
- $p_{T,l\nu bjj}$: the transverse momentum of the combination of the reconstructed top quark and all additional jets
- $\cos\Theta(l, lj)$: cosine of the angle between the lepton and the untagged jet in the top quark rest-frame

The output distributions of the four different networks are illustrated in figure 5.6.

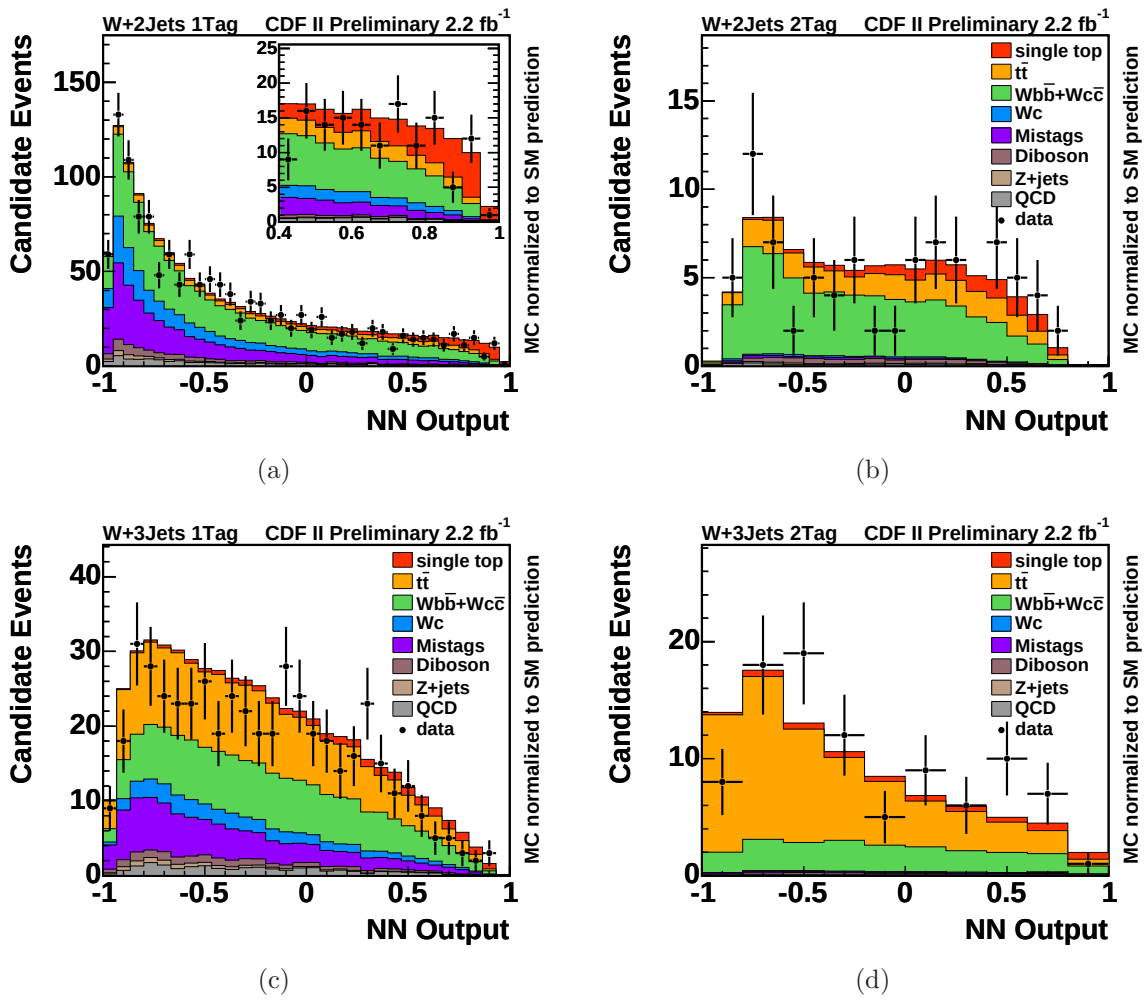


Figure 5.6: Output distributions of the four different networks: (a) 2jets-1tag network, (b) 2jets-2tags network, (c) 3jets-1tag network, (d) 3jets-2tags network. The different processes are all normalized to the standard model predictions.

Chapter 6

Combination of single-top analyses

In this chapter we report our studies concerning the combination of the three individual single-top analyses, explained at the end of the previous chapter, using the NeuroBayes® package. Since the three analyses are not 100% correlated, i.e. the analyses use different information for the identification of the single top-quark signal, a combination of the individual discriminants can improve the separation of signal and background processes.
In the following we first explain the statistic interpretation, since we optimize training parameters of the neural network training based on the sensitivity. We then explain how the neural networks used for the combination are trained.
Finally, we give an overview over the systematic uncertainties, the results of the individual analyses and the results of our combination.

6.1 Statistical interpretation

The statistical interpretation explained in the following is the same for all individual analyses and for our combination method. Two different methods are used for the measurement of the single top-quark production cross section and the determination of the sensitivity and observed significance of the analysis, respectively.

6.1.1 Measurement of the single-top production cross section

For the template fit of the discriminant output distributions to the observed distribution a likelihood function is applied. It is built up of different terms which we explain step by step below.
The first term is a Poisson distribution for the content of each bin in a fitted histogram, yielding to

$$L = \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \ , \tag{6.1}$$

where B is the number of bins, k is the bin index, μ_k is the mean value of the predicted number of events in bin k and n_k is the number of observed events in bin k. Furthermore the mean value μ_k depends on the predicted number of events of each individual physics process,

$$\mu_k = \sum_{j=1}^{A} \mu_{jk} = \sum_{j=1}^{A} \beta_j \cdot \hat{\nu}_j \cdot \alpha_{jk} \ , \tag{6.2}$$

where j is the index of a physical process, A the number of physical processes and μ_{jk} is the expectation value of the physical process j in the bin k. μ_{jk} itself depends on β_j which represents the ratio of the measured number of events and the predicted number of events for process j. β_j can also be interpreted as the ratio of the measured and the predicted cross section for process j and is one of the parameters we fit for. Moreover $\hat{\nu}_j$ is the number of expected events for process j in the given data set and α_{jk} is the relative fraction of events of process j in the bin k. Hence the α_{jk} fulfill the normalization condition $\sum_{k=1}^{B} \alpha_{jk} = 1$.
The second term in our likelihood function implements Gaussian constraints to the background rates. This is reasonable, because the knowledge about the background fractions from the background estimate is thus included. Also it is not practicable to perform a free fit of all background parameters, because there is not enough discrimination power among the various background processes. Hence the likelihood function with the Gaussian constraints is given by

$$L = \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \cdot \prod_{j=2}^{A} G(\beta_j, \Delta_j) = \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \cdot \prod_{j=2}^{A} \frac{1}{\sqrt{2\pi\Delta_j^2}} e^{-\frac{(\beta_j-1)^2}{2\Delta_j^2}} \ , \tag{6.3}$$

where the Δ_j are the relative uncertainties on the predicted number of background events of the processes j. The values of the Δ_j used in the single-top analysis are listed in table 6.1. The product in the second term starts at $j = 2$, because the signal is per definition labeled with $j = 1$.

process	Δ
$t\bar{t}$	12.4 %
$Wb\bar{b}$	30.0 %
Wcharm	30.0 %
mistags	16.6 %
Z+jets	10.8 %
diboson	1.9 %
QCD	40.0 %

Table 6.1: Relative standard deviations used for the gaussian constraints of the background rates

The last term in our likelihood function incorporates the influence of systematic

uncertainties. Three different types of systematic uncertainties are included: Uncertainties which affect the rates of a particular signal or background process (rate uncertainty), uncertainties which influence the shape of the templates used for the fit (shape uncertainty) and bin by bin uncertainties which are due to the limited number of simulated events and which are only included for the calculation of the sensitivity (see the following subsection). The systematic effects, which we account for in the single-top analyses, are discussed in detail in section 6.3. These uncertainties enter the likelihood function by shifting the expected mean of every process j in every bin k. Therefore the formula for the mean values μ_{jk} changes to

$$\begin{aligned}\mu_{jk} = \beta_j \cdot \hat{\nu}_j \cdot \left\{ \prod_{i=1}^{S} (1 + |\delta_i| \cdot (\epsilon_{ji+} H(\delta_i) + \epsilon_{ji-} H(-\delta_i))) \right\} \\ \cdot \alpha_{jk} \cdot \left\{ \sum_{l=1}^{S'} (1 + |\delta_l| \cdot (\kappa^{+}_{jlk} H(\delta_l) + \kappa^{-}_{jlk} H(-\delta_l))) \right\} ,\end{aligned} \tag{6.4}$$

where the $\delta_{i,l}$ represent the strength of an uncertainty i, the ϵ_{ij} are the relative rate uncertainties, the $\kappa^{\pm}_{jlk}$ are the relative shape uncertainties, S is the number of systematic rate uncertainties and S' is the number of systematic shape uncertainties. If a systematic effect causes both, rate and shape uncertainties, δ_i and δ_l have the same value. $H(x)$ denotes the Heavyside step function. The $\kappa^{\pm}_{jlk}$ are derived from normalized, systematic shifted histograms (α^{+}_{jlk} and α^{-}_{jlk}) which take different scenarios of a systematic uncertainty into account, e.g. the changes of the discriminant shape if one assumes more or less of a systematic effect. They are calculated via

$$\kappa^{\pm}_{jlk} = \frac{\alpha^{\pm}_{jlk} - \alpha_{jk}}{\alpha_{jk}} , \tag{6.5}$$

and fulfill $\sum_{k=1}^{B} \alpha_{jk} \cdot \kappa^{\pm}_{jlk} = 0$. Some uncertainties cause only one systematic shifted histogram, e.g. if there is only one alternative model. For these uncertainties only κ^{+}_{jlk} or κ^{-}_{jlk} is incorporated.
The strength of a systematic uncertainty $\delta_{i,l}$ is again constraint with a Gaussian distribution, centered at zero with a standard deviation of one. Hence the final likelihood function is given through

$$\begin{aligned} L &= \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \cdot \prod_{j=2}^{A} G(\beta_j, \Delta_j) \cdot \prod_{i=0}^{S} G(\delta_{i,l}; 0, 1) \\ &= \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \cdot \prod_{j=2}^{A} G(\beta_j, \Delta_j) \cdot \prod_{i=0}^{S} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\delta_{i,l})^2}{2}} . \end{aligned} \tag{6.6}$$

The measurement of the single-top cross section is done by calculating the maximum of reduced likelihood function $L_{\mathrm{red}}(\beta_1)$ which only depends on β_1. There are two main

possibilities for obtaining the reduced likelihood function. The method of "marginalization" is used for the measurement of β_1 in observed events. Here, the original likelihood function is integrated over all "unwanted" parameters (also called nuisance parameters), i.e. β_2-β_A and δ_1-δ_S. After that, a scan over β_1 is performed to determine the maximum of $L_{\mathrm{red}}(\beta_1)$.
A second method, called "profiling of the likelihood", calculates the minimum of the negative log-likelihood function, with respect to all nuisance parameters, for a fixed β_1 value, using the MINUIT [78] package, which results in a one dimensional distribution $L_{\mathrm{red}}(\beta_1)$. Profiling is used for the determination of the sensitivity and the observed significance which is explained in the next subsection.

6.1.2 Ensemble tests and sensitivities

For the calculation of the sensitivity and the observed significance a likelihood function, slightly different from the one explained above, is used. Here, the systematic uncertainties are not included directly in the likelihood function but taken into account in another way which we explain below. Thus the likelihood function is given by the equations 6.2 and 6.3.
In order to compute the sensitivity and the observed significance of an analysis ensemble test are performed. An ensemble test consists of a set of simulated, so-called pseudo experiments. For each of these pseudo experiments first the number of events of a particular process N_j is drawn from a Poisson distribution of mean $\hat{\nu}'_j$. Systematic rate uncertainties are incorporated via

$$\hat{\nu}'_j = \hat{\nu}_j \cdot \left\{ \prod_{i=1}^{S} (1 + |\delta_i| \cdot (\epsilon_{ji+} H(\delta_i) + \epsilon_{ji-} H(-\delta_i))) \right\} , \tag{6.7}$$

where the strength of the systematic uncertainties, δ_i, are drawn from gaussian distributions which are centered at zero with standard deviation one. After that, N_j random numbers are drawn from the template distributions α'_{jk} of the j physical processes. Here, the systematic shape uncertainties are included by

$$\alpha'_{jk} = \alpha_{jk} \cdot \left\{ \sum_{l=1}^{S'} (1 + |\delta_l| \cdot (\kappa^+_{jlk} H(\delta_l) + \kappa^-_{jlk} H(-\delta_l))) \right\} , \tag{6.8}$$

where the δ_l are again drawn from Gaussian distributions which are centered at zero with standard deviation one. If the α^+_{jik} and α^-_{jik} histograms mentioned above represent maximum deviations from the default values, these Gaussian distributions are truncated between -1 and 1. Additionally the uncertainty due to the limited number of simulated events is incorporated by fluctuating the number of entries in each bin k of the α_{jk} templates with a Gaussian distribution, centered at the original value with a standard deviation that represents the statistical error on the number of entries.
The N_j random numbers are filled into j histograms which represent the discriminant

values for a process j. The single-top cross section of each pseudo experiment is calculated by profiling the likelihood function as explained in the previous subsection.
Two different ensemble tests are performed. One including single top-quark events and one without any single top-quark event. For this purpose β_1 is fixed at 1, i.e. the standard model single top-quark production cross section or, 0 respectively. For the calculation of the sensitivity of an analysis two different hypotheses are tested. The hypotheses H_0 assumes that no single top-quark production exists, while the hypothesis H_1 assumes single top-quark production with the cross section predicted by the standard model. To be able to claim an observation of single top-quark production the hypothesis H_0 has to be rejected at the 99.999% C.L.. The hypotheses test is based on the Q-value which is defined by

$$Q = -2 \ln \frac{L_{\mathrm{red}}(\beta = 1)}{L_{\mathrm{red}}(\beta = 0)} \ . \tag{6.9}$$

The Q-value distributions for the different hypothesis are called q_0 and q_1, respectively. With these distributions the probability for H_0 to be true can be quoted. For this purpose the so-called p-value is defined via

$$p(Q_m) = \frac{1}{A_q} \cdot \int_{-\infty}^{Q_m} q_0(Q')\, dQ' \ , \tag{6.10}$$

where $A_{\mathrm{q}} = \int_{-\infty}^{\infty} q_0(Q')\, dQ'$ and Q_{m} is the measured Q-value, in case of observed events, or the median of the H_1 Q-value distribution, in case of pseudo experiments. Hence, the meaning of the expected p-value ($\hat{p}$) is the following: Under the assumption that H_1 is true, one expects to observe $p < \hat{p}$ with a probability of 50%.

6.2 Training of the neural networks

Four different neural networks are trained for the combination of the three single top-quark analyses, one network for events with two jets and one or two SecVtx tags, respectively and one network for events with three jets and one or two SecVtx tags, respectively. For this purpose we use the outputs of the individual discriminants as input variables for our neural networks. Typically, the samples used for the network trainings are composed of the relevant physics processes, which are mixed with ratios predicted from the standard model. Since the number of expected single top-quark events is much smaller than the expected number of all background events, this is not feasible, because the neural network would not be able to separate signal from background events. Hence, training samples with different mixtures have to be used. For each network, trainings with different training sample compositions are performed. The composition which leads to the best discrimination of signal and background events is selected utilizing ensemble tests. For this, pseudo experiments are drawn from eight different network output distributions which represent the following physics processes: Single top-quark production, $Wb\bar{b}$, $Wc\bar{c}$+Wc (Wcharm), mistags, WW+WZ+WW

(diboson), Z+jets, QCD-multijets and $t\bar{t}$ production. The output templates for WW, WZ and WW are added together according to their predicted ratios. The selection of the most discriminating network itself, is based on the so-called "o-value" which is defined by

$$o = \frac{|\overline{q}_1(Q) - \overline{q}_0(Q)|}{\sqrt{\Delta q_1^2(Q) + \Delta q_0^2(Q)}} \ , \tag{6.11}$$

where $\overline{q}_{0,1}(Q)$ is the mean of the Q-value distribution of the $H_{0,1}$ hypothesis and $\Delta q_{0,1}$ represents the "Root Mean Square" (RMS) of the two Q-value distributions. While for the calculation of a p-value at least 50 million pseudo experiments are necessary the o-value is already numerically stable for roughly 100000 pseudo experiments yielding a tremendous decrease in the needed computing time. Finally, the four networks with the highest o-values are used for the combination. This is reasonable, because the larger the separation of the Q-value distributions for the two hypothesis $H_{0,1}$ the higher is the o-value. Since a large separation of the Q-value distributions yields in a small p-value, i.e. a large sensitivity, one can in summary say that the higher the o-value, the higher the sensitivity.
In subsection 5.3.2 it is mentioned, that the Matrix Element analysis uses additionally non-triggered muon events. These events are not included in our combination.

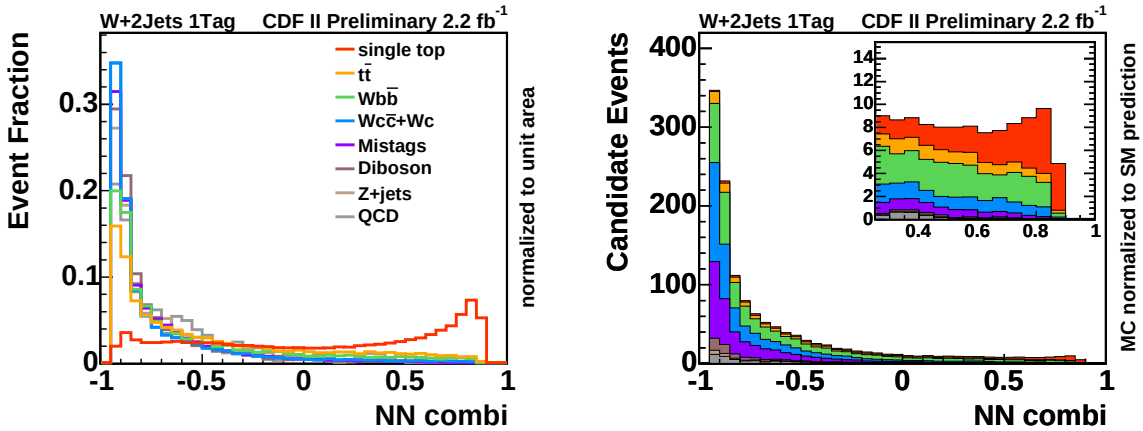
The network using two jets and one tagged jet

In a first step a neural network is trained using the outputs of the three individual analyses as input variables. After preprocessing the following correlations between the analyses are found: The correlation between the Likelihood Function method (LF) and the Matrix Element method (ME) is 81.3%. Furthermore the correlation is 89.0% between the LF method and the Neural Network analysis (NN) and 82.4% between the ME analysis and the NN method. From these input variables, the Neural Network method shows the largest correlation to the target. Because of the high correlations between the three analyses we decided to use the LF and the ME output a second time, i.e. as two further input variables, but this time preprocessed with the "independent correlation option (ICO)" (see subsection 3.1.1) with respect to the NN method output. Thus, these two variables dominantly incorporate only the excess of information given by the LF and ME method in comparison to the NN analysis.
Furthermore neural networks for eighteen different compositions of the training sample are trained. First the signal-to-background ratio is varied from 50:50 over 40:60 to 30:70. For this the background processes are mixed according to their predicted ratios. For each signal-to-background ratio six different mixtures of the signal sample are performed. The following s-channel single-top to t-channel single-top ratios are realized: 10:90, 20:80, 30:70, 34:66 (predicted by the standard model), 40:60 and 50:50. For the network trained with a signal-to-background ratio of 30 : 70 and a s-channel single-top to t-channel single-top ratio of 10 : 90 the largest o-value has been found. Table 6.2 lists the input variables ranked according to their relative significance and the correlation to the target of each input variable for this neural network. After preprocessing

a 3σ cut on the relative significance has been performed (illustrated by the horizontal line in table 6.2). Only two variables have passed this cut and have been finally used for the network training. The network output distributions for the different physics processes are illustrated in figure 6.1(a). Figure 6.1(b) shows the output distributions normalized to the standard model prediction.

rank	input variable	correlation to target [%]	relative significance in σ
1	NN	63.2	198.3
2	ME	59.1	38.8
3	ME_{ICO}	45.5	2.3
4	LF	58.2	1.4
5	LF_{ICO}	46.2	1.7

Table 6.2: Ranking of the input variables for the network using two jets and one SecVtx tag. The horizontal line represents the 3σ cut on the relative significance.



(a) (b)

Figure 6.1: (a): Shapes of the output distributions of the network using two jets and one SecVtx tag for the different processes. (b): Output distributions normalized to the number of events which are predicted by the standard model for the different processes. Within the plot, the output region NN combi > 0.3 is zoomed.

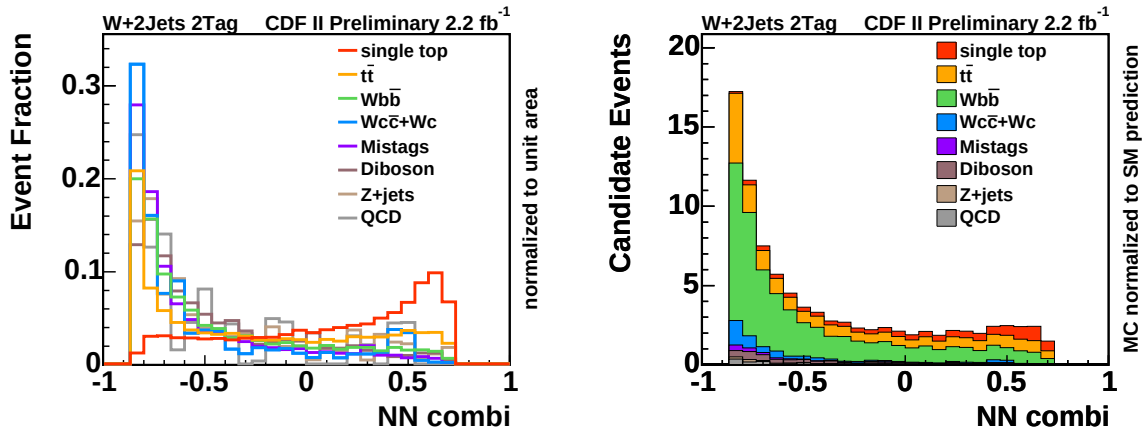
The network using two jets and two tags

Since there is no Likelihood function discriminant for events with two jets and two SecVtx tags, only the ME and the NN outputs are incorporated for the training. The correlation of these two analyses is 75.3%, the Matrix Element method output is showing the highest correlation to the target. Hence the NN output, preprocessed

with the ICO is included as a further input variable.
For six different compositions of the training sample a neural network is trained. Again the three signal-to-background ratios mentioned above are realized and the background processes are mixed according to their predicted ratios. The signal template consists purely of s-channel single-top events or is mixed of s-channel single-top and t-channel single-top events according to the standard model ratio. A maximum o-value has been found for the network which is trained with a signal-to-background ratio of 40 : 60 and with a signal template that consists purely of s-channel single-top events. Table 6.3 gives an overview of the significance and the correlation to the target of the three variables. A 3σ cut on the relative significance after the preprocessing reduces the number of input variables. Hence, the final network has been trained only with two input variables. Figure 6.2(a) illustrates the output distributions for the different physics processes. In figure 6.2(b) these distributions are normalized to the standard model prediction.

rank	input variable	correlation to target [%]	relative significance in σ
1	ME	50.6	72.1
2	NN	45.3	15.6
3	NN_{ICO}	41.8	1.7

Table 6.3: Ranking of the input variables for the network using two jets and one SecVtx tag. The horizontal line represents the 3σ cut on the relative significance.



(a) (b)

Figure 6.2: (a): Shapes of the output distributions of the network using two jets and one SecVtx tag for the different processes. (b): Output distributions normalized to the number of events which are predicted by the standard model for the different processes.

The network using three jets and one tag

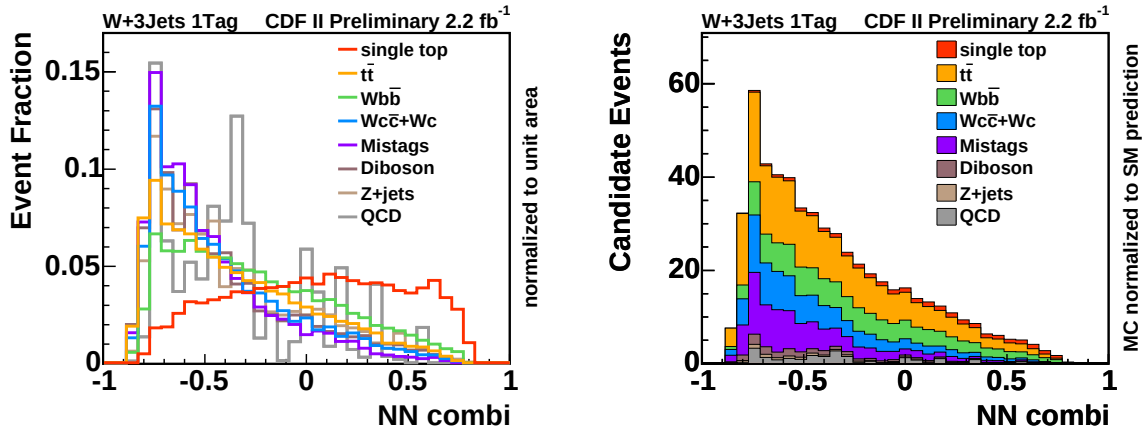
The input variables to this network are similar to the ones used for the training of the network using two jets and one SecVtx tag, i.e. additionally to the outputs of the three individual analyses the LF and ME output, preprocessed with the ICO with respect to the NN output, are included. For events with three jets and one SecVtx tag the analyses are less correlated than for the two cases already discussed above. The correlation between the Likelihood Function analysis and the Matrix Element technique is 48.6%, while the correlation between the LF analysis and the NN method is 83.2% and the correlation between the ME technique and the NN method is 51.0%. Fifteen different compositions of the training sample are realized: The three already mentioned signal-to-background ratios and different mixtures of the signal sample (s-channel single-top to t-channel single-top = 10:90, 20:80, 30:70, 38:62 (SM ratio) and 50:50). For each composition a neural network has been trained. The neural network for which the largest o-value has been found is trained with a signal-to-background ratio of 40 : 60 and with a s-channel single-top to t-channel single-top ratio of 50 : 50. In table 6.4 the input variables are ranked according to their relative significance. Four of the five input variables passed to NeuroBayes survive a 3σ cut on the significance after the preprocessing. The network output distributions for the different physics processes are illustrated in figure 6.3(a). Figure 6.3(b) shows the output distributions normalized to the standard model prediction.

rank	input variable	correlation to target [%]	relative significance in σ
1	NN	40.9	92.1
2	ME	33.1	32.1
3	LF	37.7	11.0
4	ME_{ICO}	30.2	5.6
5	LF_{ICO}	27.3	2.4

Table 6.4: Ranking of the input variables for the network using three jets and one SecVtx tag. The horizontal line represents the 3σ cut on the relative significance.

The network using three jets and two tags

In contrast to events with two jets and two SecVtx tags, a Likelihood Function discriminant exists for events with three jets and two SecVtx-tags. Hence, the outputs of all three single-top analyses can be used as input to the neural network. The ME analysis shows the largest correlation to the target. Therefore, the NN and LF outputs, preprocessed with the ICO with respect to the ME output, are used as two further input variables. After preprocessing, the following correlations between the different analyses are found: LF method - ME analysis 40.4%, LF method - NN method 63.2% and ME analysis - NN method 44.5%.
For six different compositions of the training sample a neural network is trained. Again



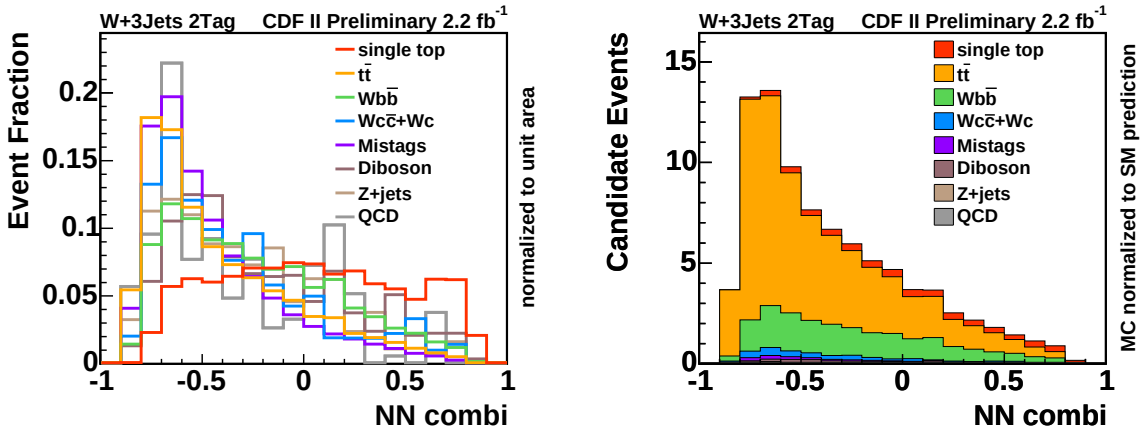
(a) (b)

Figure 6.3: (a): Shapes of the output distributions of the network using three jets and one SecVtx tag for the different processes. (b): Output distributions normalized to the number of events which are predicted by the standard model for the different processes.

three different signal-to-background ratios, 50:50, 40:60 and 30:70, are realized. For each signal-to-background ratio two different compositions of the signal sample are tested: s-channel single-top and t-channel single-top events mixed according to the ratios predicted by the standard model (56:44) and an s-channel single-top to t-channel single-top ratio of 70:30. The network reaching the largest o-value has been trained with a signal-to-background ratio of 40 : 60 and with a standard model mixture of the signal sample. For this network, table 6.5 gives an overview of the significance and the correlation to the target of the five input variables. A 3σ cut on the relative significance after the preprocessing diminishes the number of input variables. Figure 6.4(a) illustrates the output distributions for the different physics processes. In figure 6.4(b) these distributions are normalized to the standard model prediction.

rank	input variable	correlation to target [%]	relative significance in σ
1	ME	38.2	37.5
2	NN	35.3	20.1
3	LF_{ICO}	30.1	7.2
4	NN_{ICO}	34.6	1.6
5	LF	31.4	1.0

Table 6.5: Ranking of the input variables for the network using three jets and two SecVtx tag. The horizontal line represents the 3σ cut on the relative significance.



(a) (b)

Figure 6.4: (a): Shapes of the output distributions of the network using three jets and two SecVtx tag for the different processes. (b): Output distributions normalized to the number of events which are predicted by the standard model for the different processes.

6.3 Systematic uncertainties

Uncertainties in the modeling of physics processes or detector effects cause systematic uncertainties on the measurement results. For the single-top analyses three different types of systematic uncertainties are incorporated: Uncertainties which affect only the rate of particular processes, uncertainties which affect the shapes of the discriminant templates for particular processes and bin by bin uncertainties which are due to the limited number of simulated events and which are only included in the calculation of the sensitivity. For sure, most of the systematic effects cause both, rate and shape uncertainties.

The following sources of systematic uncertainties are considered: the uncertainty on the jet energy scale, the uncertainty in modeling initial-state gluon radiation (ISR) and final-state gluon radiation (FSR), the choice of the parameterization of the parton distribution functions (PDF) used for the event simulation, the choice of the Monte Carlo event generator, the uncertainty in the event detection efficiency, the uncertainty of the output of the KIT Flavor Separator, the uncertainty in the factorization and renormalization scale for the simulation of W+heavy flavor processes Q^2, the modeling of instrumental backgrounds, i.e. mistag events and QCD-multijet events, the uncertainty in the luminosity determination and the uncertainty on the modeling of the distributions of $\Delta R_{\mathrm{j1,j2}}$ for events with two jets and the η distribution of the softest jet η_{j}, for events with two or three jets. The effects of these uncertainties are taken into account by varying the corresponding processes within the uncertainties or by applying an alternative model. This results in systematically shifted templates or relative shifts in the event rates. In the following we explain the systematic effects in more detail.

For each shape uncertainty the default and the shifted distributions in addition with the relative differences of the varied templates with the default ones are illustrated. In this thesis only the shape uncertainties for the network using two jets and one SecVtx tag are shown. Table 6.6 gives an overview over the processes which are affected by the different systematic uncertainties.

uncertainty	single-top	$Wb\bar{b}$	Wcharm	mistags	diboson	Z+jets	QCD	$t\bar{t}$
JES	×	×	×	×	×	×		×
ISR/FSR	×							×
PDF	×							
Q^2		×						
mistag model				×				
QCD flav. mixture							×	
Flav. Sep.			×	×				
η_{j}	×	×	×	×	×	×		×
ΔR_{j1j2}	×	×	×	×	×	×		×

Table 6.6: Processes which are affected by the different shape uncertainties

If the shifts caused by a particular systematic uncertainty represent maximum deviations, the Gaussian distribution, from which the strength δ_i is drawn in the calculation of pseudo experiments, are constraint between -1 and 1. This is done for the following systematics: JES, ISR/FSR, PDF, Q^2, KIT Flavor Separator output and the modeling of mistags, QCD-multijets, $\Delta R_{\mathrm{j1,j2}}$, and η_{j} distributions.
The effects caused by the uncertainty in the jet energy scale (JES) is incorporated by applying jet energy corrections which represent maximum deviations from the original correction [49]. The rate uncertainties caused by the jet energy scale for the different processes are listed in table 6.7. Differences in the shapes of the templates for the different processes can be seen in figures 6.5 and 6.6.

process	2jets 1tag	2jets 2tags	3jets 1tag	3jets 2tags
single-top	-0.8/0.2 %	1.8/-2.2 %	-9.1/9.9 %	-6.6/6.1 %
$t\bar{t}$	9.9/-9.4 %	8.2/-7.6 %	4.6/-5.1 %	5.4/-5.2 %
$Wb\bar{b}$	6.9/-7.6 %	10.7/-10.6 %	8.5/-8.4 %	10.6/-11.5 %
Wcharm	7.0/-6.3 %	11.3/-10.3 %	8.2/-6.9 %	13.9/-15.8 %
Z+jets	-5.3/5.4 %	5.0/-5.0 %	-10.8/14.0 %	-5.9/7.2 %
diboson	-2.7/1.7 %	-2.9/1.4 %	-12.4/11.7 %	-12.0/12.0 %

Table 6.7: Systematic jet energy scale down/up rate uncertainties

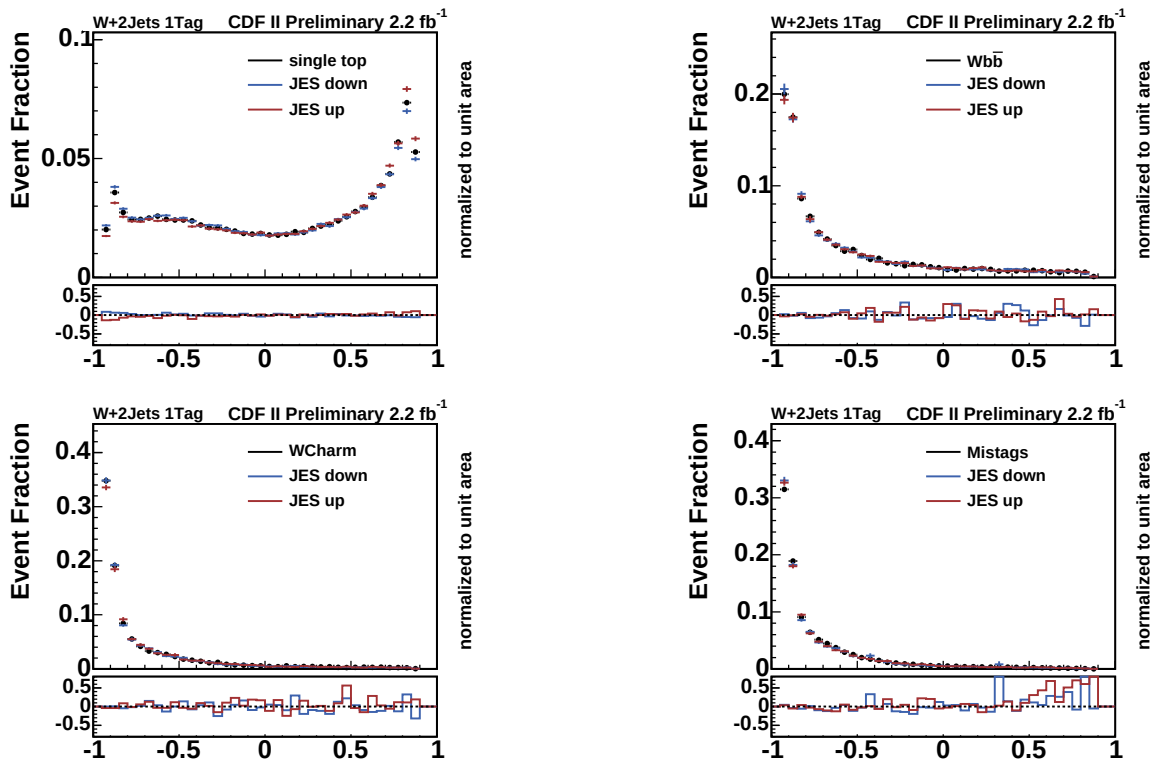


Figure 6.5: Shape systematics caused by the uncertainty on the jet energy scale. In the upper row the shifts in the single-top (left) and the $Wb\bar{b}$ (right) template are shown. In the lower left corner the uncertainties on $Wc\bar{c}$+Wc are illustrated. The shifts in the mistags distribution is displayed in the lower right corner. In addition to the particular distribution the relative deviations from the original output are shown.

The effects due to more or less initial/final state radiation are modeled by producing samples in which less or more gluon radiation in the parton shower is simulated [79]. Using these specific ISR and FSR samples of simulated events, alternative template shapes are produced for single top-quark and $t\bar{t}$ events as illustrated in figure 6.7.

To include the uncertainties resulting from the choice of a particular parton distribution function, weights obtained from 20 pairs of independent eigenvectors of the CTEQ6M parton distribution are compared to the weights of the originally used CTEQ5L PDF. The relative differences of the weights from CTEQ5L and from the eigenvectors are added quadratically considering the particular sign. PDF uncertainties affect only the shape of the single-top template, illustrated in figure 6.8(a), and the rate of the single-top and top-pair processes.

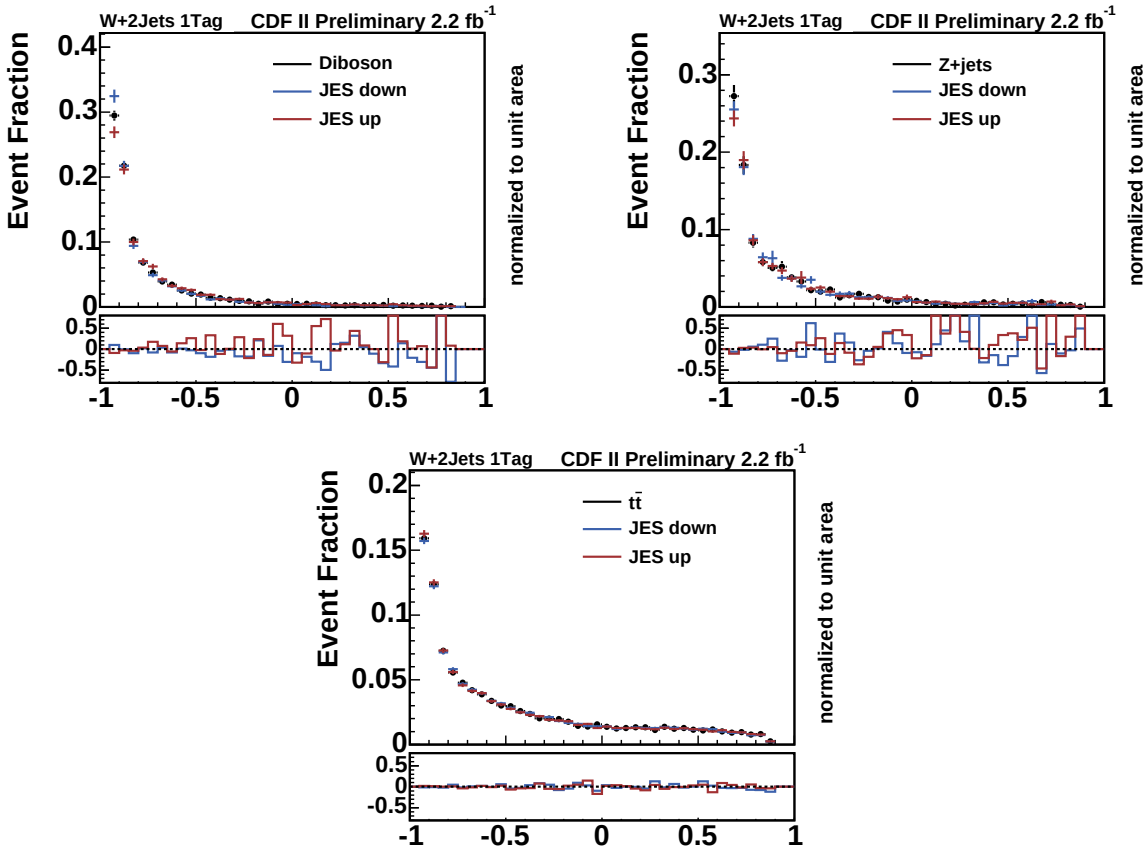
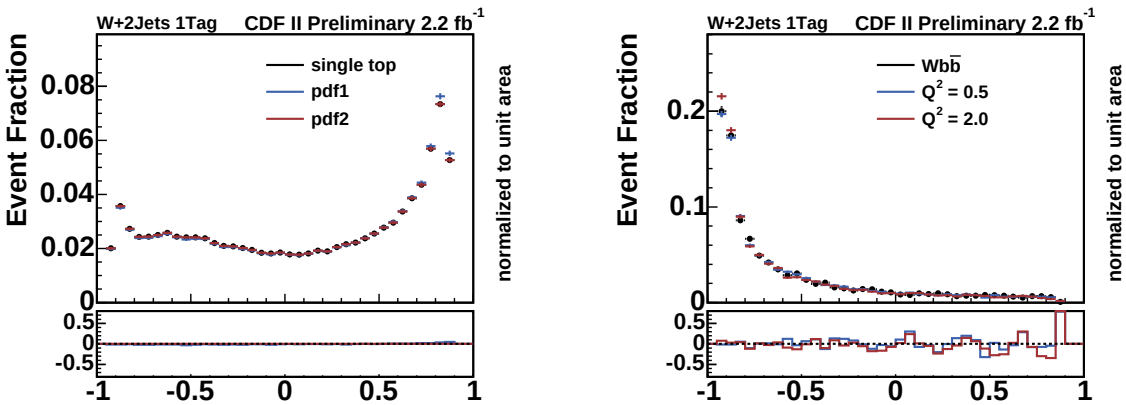


Figure 6.6: Shape systematics caused by the uncertainty on the jet energy scale. In the upper left corner the shifts in the diboson template are shown. The uncertainties on the Z+jets output are illustrated in the upper right plot. Beyond the particular distribution the relative deviations from the original output are shown.



(a) (b)

Figure 6.8: (a): Shifts in the single-top output distribution caused by different PDF scenarios. (b): Uncertainties on the $Wb\bar{b}$ template derived from different Q^2 values.

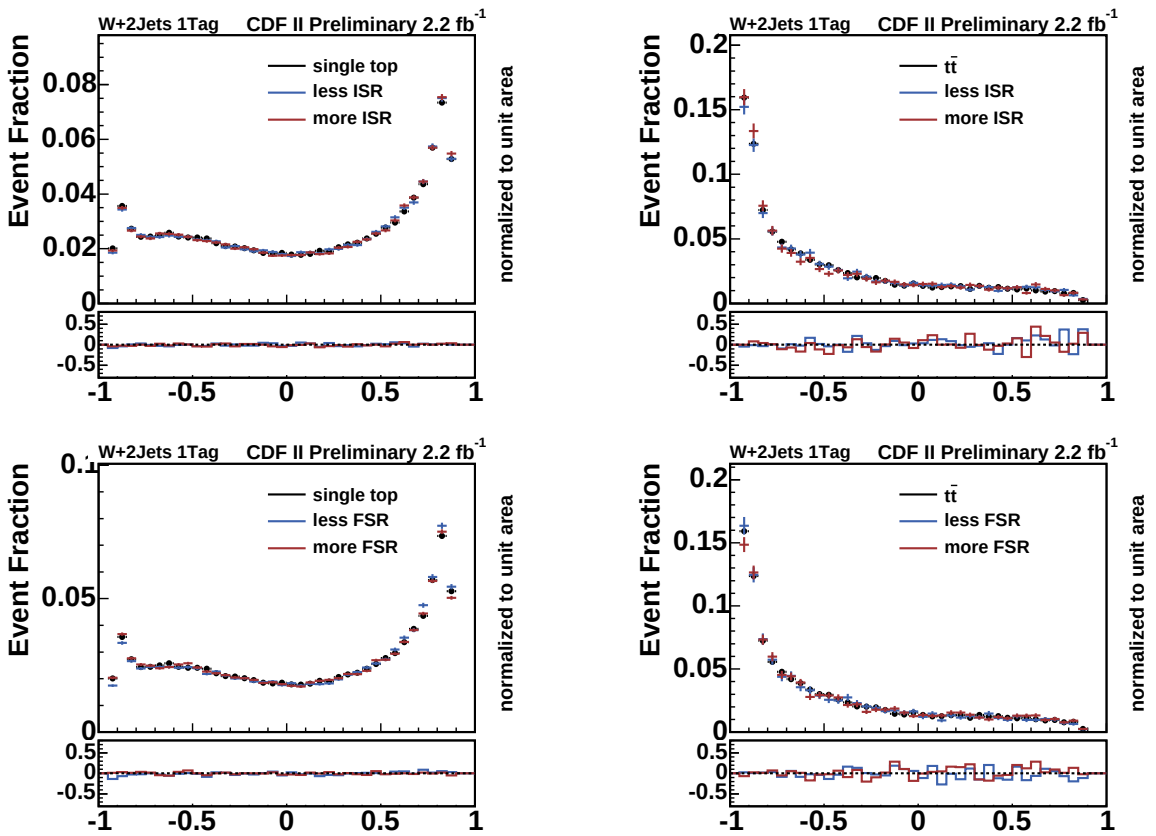


Figure 6.7: Shape systematics caused by different scenarios for initial/final state radiation. In the upper row the effects due to more or less ISR are illustrated. On the left hand side one can see the shifts in the single-top output distribution, the plot on the right hand side displays the shifts in the $t\bar{t}$ template. In the second row the effects cause by more or less FSR are shown.

To account for the uncertainties due to the choice of the renormalization and factorization scale in the production of W+heavy flavor events, two different scenarios are realized. The default W+jets Monte Carlo samples are generated with a dynamic scale $\mu^2 = Q^2$. For the two alternative scenarios this scale is halved and doubled, respectively. The influence of the variation of Q^2 on the shape of the output distribution of $Wb\bar{b}$ events is illustrated in figure 6.8(b).
Uncertainties in the rate of the single top-quark production are derived from a comparison of the distributions of different kinematic variables in NLO calculations and the MadEvent Monte Carlo sample. The uncertainty in the modelling of $t\bar{t}$ processes is determined by using events simulated with the Herwig program as an alternative generator. Furthermore, the uncertainty of the event detection efficiency ϵ_{evt} includes the uncertainties on the trigger efficiency, on the lepton identification and on the b-tagging scale factor. The correction function applied to the mistag output of the KIT Flavor Separator, is incorporated as systematic uncertainty by defining an optimistic

and a pessimistic model. For the pessimistic scenario the correction function is applied to the W+charm and the mistags output forcing both to be more signal like. For the optimistic model both output distributions are used uncorrected, which leads to more background like templates. The changes in the template shapes can be seen in figure 6.9.

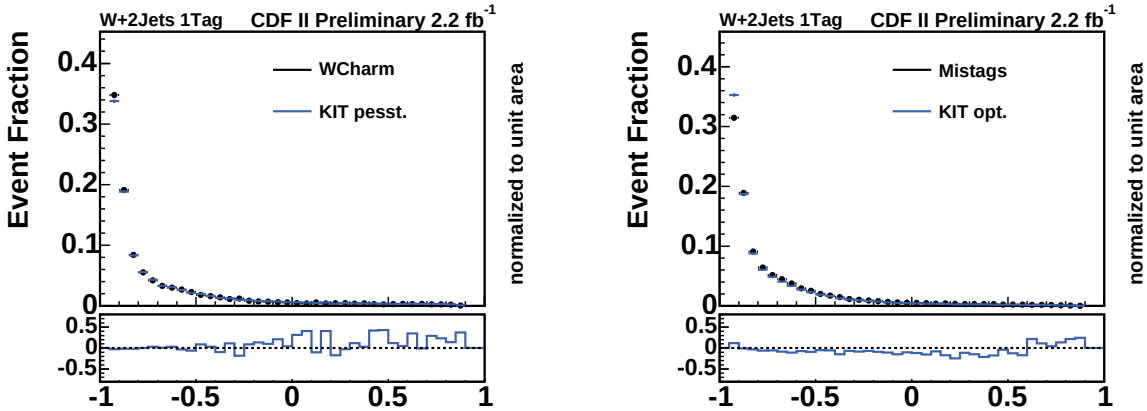
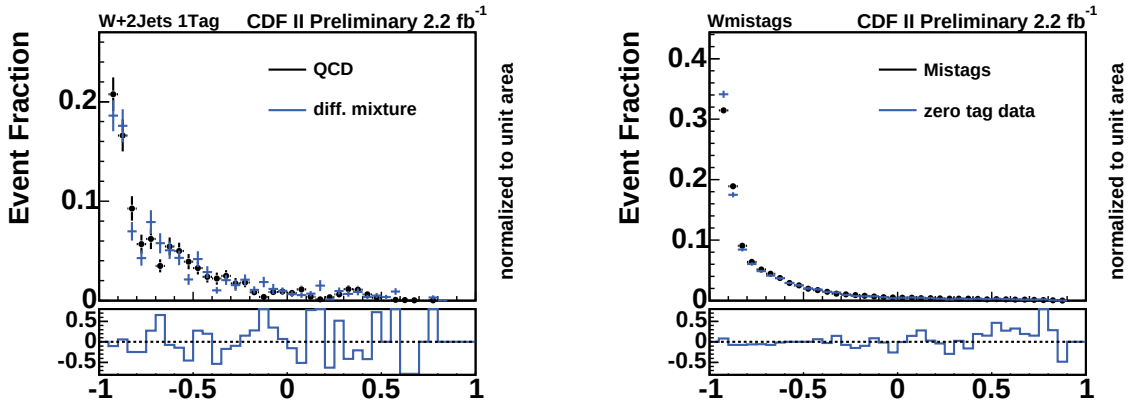


Figure 6.9: Optimistic and pessimistic model of the KIT Flavor Separator. On the left hand side the correction function is applied to the $Wc\bar{c}$+Wc template. On the right hand side the correction function is omitted for mistag events.

The influence of the heavy flavor composition of the QCD-multijet sample is derived by using samples with a different mixture. Studies with the KIT Flavor Separator showed the following ratios of heavy flavor processes: 45% b-quark jets, 40% c-quark jets and 15% light quark jets [80]. The alternative model uses an "extreme" composition of 60% b-quark jets, 30% c-quark jets and 10% light quark jets. Its effect on the QCD-multijet output is shown in figure 6.10(a).

To evaluate the systematic effect on the shape of the mistag distribution which is caused by the modeling of mistagged light-quark jet events, the default model is compared with W+jets data, for which the jets are taggable but not tagged. The differences in the template shape are illustrated in figure 6.10(b).

Moreover the possible mismodeling of ΔR between the two jets in simulated two jet events and of the pseudorapidity of the second leading jet, η_{j}, in simulated events with two or three jets is taken into account. This is done by reweighting the templates using the distributions of the particular variable in the sideband of events which have at least one taggable jet but no tagged jets. The shifts caused by these uncertainties are shown in figures 6.11, 6.12, 6.13 and 6.14.



(a) (b)

Figure 6.10: Systematic effects caused by a different heavy flavor composition of the QCD-multijet sample (a) and a different modeling of mistagged light-quark jets (b).

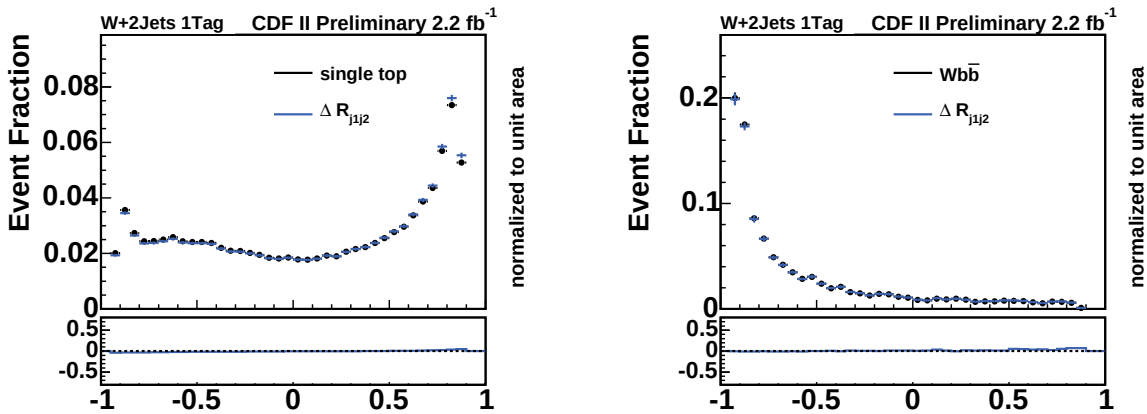


Figure 6.11: Systematic uncertainties due to the mismodeling of ΔR. On the left hand side the shift on the single-top template is illustrated. The plot on the right hand side shows the effect on the $Wb\bar{b}$ output distribution.

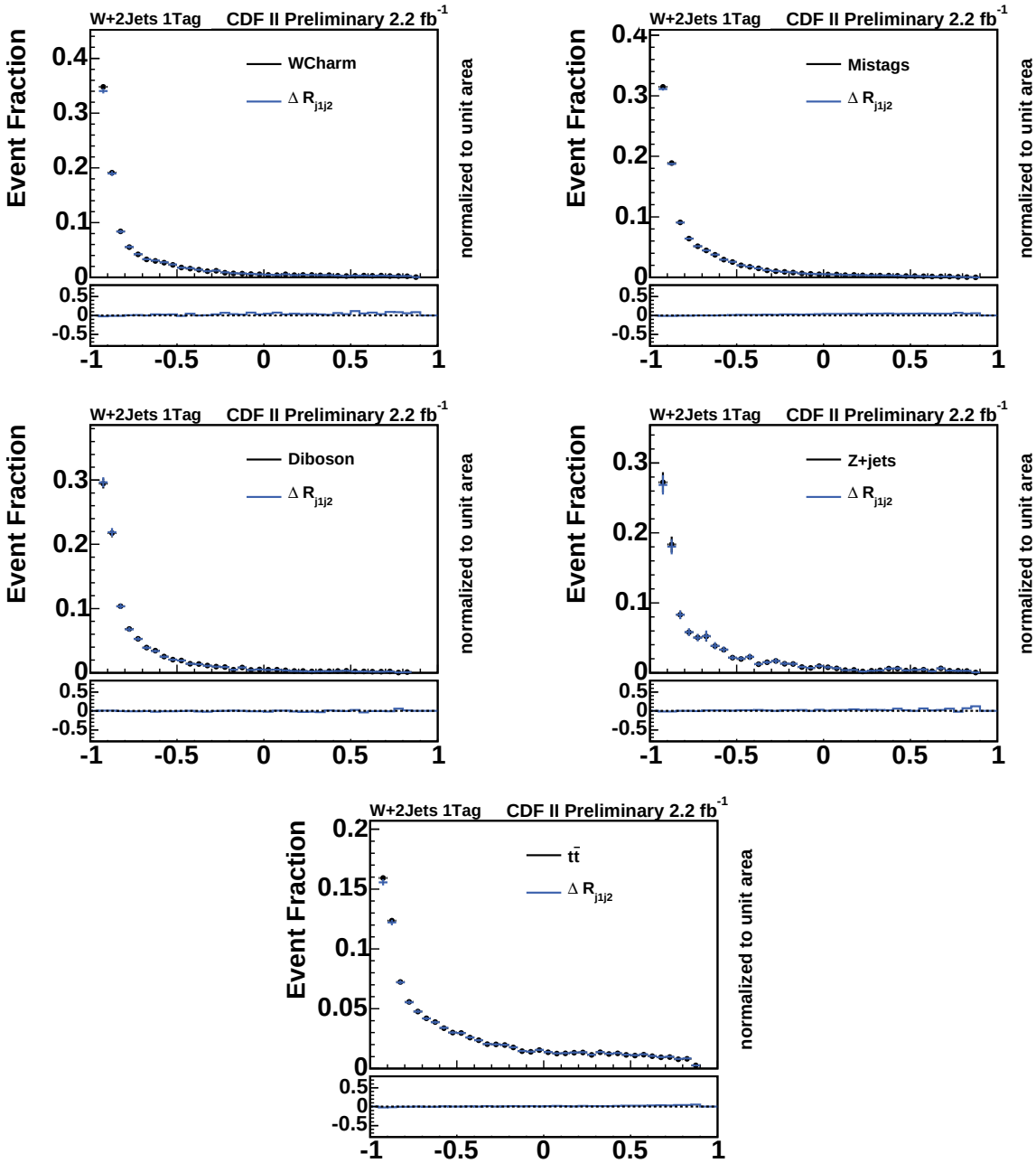


Figure 6.12: Systematic uncertainties due to the mismodeling of ΔR. In the upper left corner the effect on the Wcharm template is illustrated. The plot in the upper right corner shows the shifts on the mistags output. In the second row the uncertainties on diboson (left hand side) and Z+jets processes are displayed. The last plot shows the effect on the $t\bar{t}$ template.

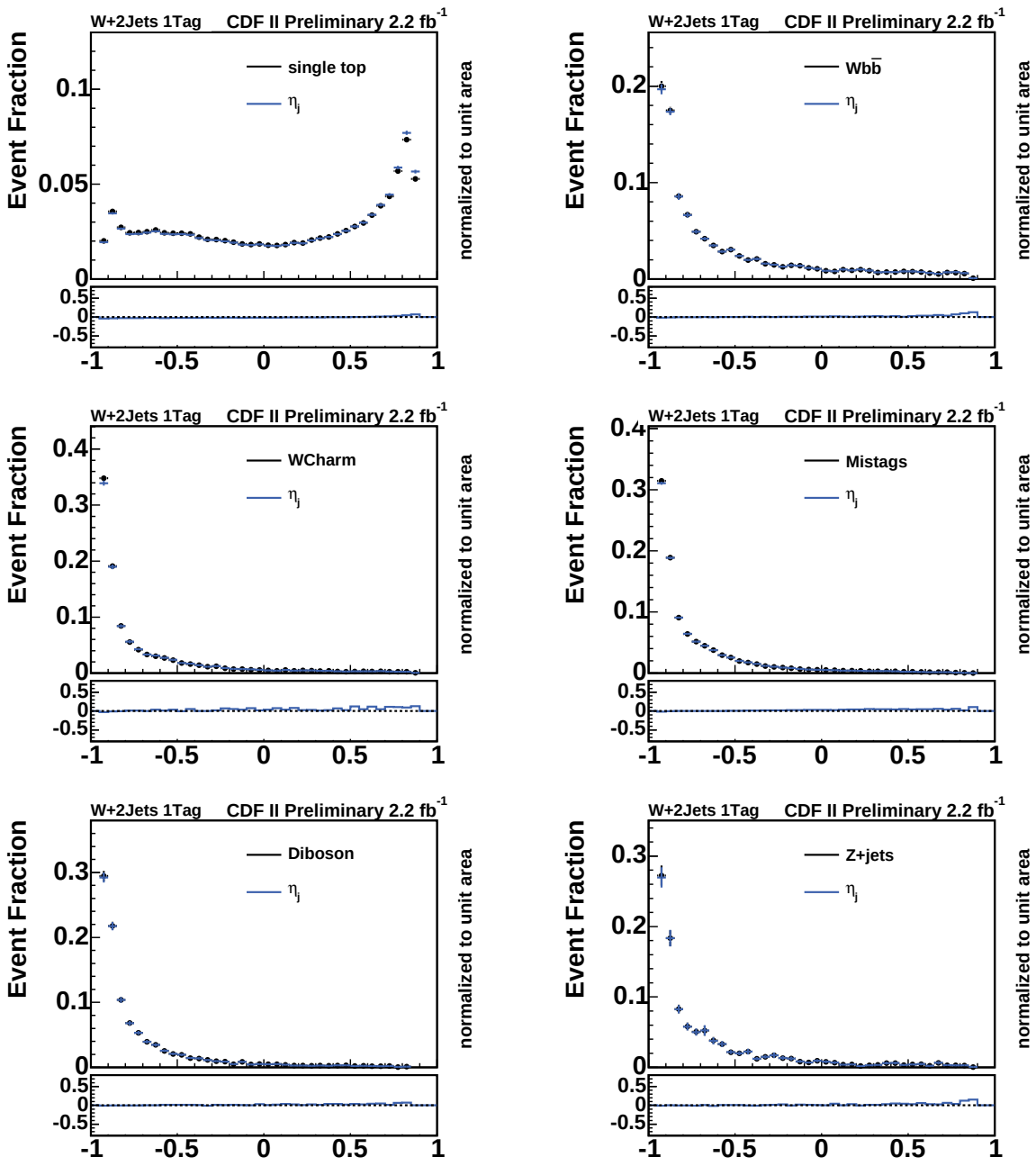


Figure 6.13: Systematic uncertainties due to the mismodeling of η_j. In the upper left corner the effect on the single-top template is illustrated. The plot in the upper right corner shows the shifts on the $Wb\bar{b}$ output. In the second row the uncertainties on Wcharm (left hand side) and mistag (right hand side) processes are displayed. In the lower left corner the impact on the diboson template is illustrated. The plot in the lower right corner shows the shifted Z+jets output.

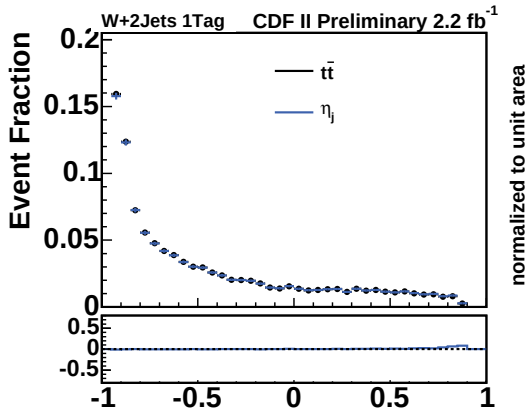


Figure 6.14: Systematic uncertainties due to the mismodeling of η_j. This plot shows the effect on the $t\bar{t}$ output distribution.

Additionally, the different rates of single top-quark and top-pair processes for different top-quark masses (170 and 180 GeV/c^2) are included for the determination of the sensitivity and the observed significance. This uncertainty is not included in the measurement of the single top-quark production cross-section. Another rate uncertainty (double tag), accounts for the differences of the uncertainties on the expected number of mistag events with one or two SecVtx tags, respectively.
The relative rate uncertainties on the different processes are listed in the tables 6.8 (events with 2jets and 1tag), 6.9 (2jets 2tags), 6.10 (3jets 1tag) and 6.11 (3jets 2tags).

Source	single-top	$t\bar{t}$	diboson	Z+jets
ISR less/more	1.9/2.1 %	-2.6/-7.1 %		
FSR less/more	4.8/-0.7 %	-5.1/-2.6 %		
PDF	3.0/-3.0 %	1.8/-1.8 %		
MC generator	1.7/-1.7 %	-2.7/2.7 %		
ϵ_{evt}	3.6/-3.6 %	2.9/-2.9 %	7.6/-7.6 %	8.3/-8.3 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
M_{top} 170/180	7.3/-6.2 %	7.8/-8.1 %		

Table 6.8: Systematic rate uncertainties for events with 2 jets and 1 SecVtx tag.

Source	single-top	$t\bar{t}$	diboson	Z+jets	mistags
ISR less/more	0.4/6.7 %	0.5/-9.5 %			
FSR less/more	7.5/0.8 %	-8.1/-1.8 %			
PDF	2.0/-2.0 %	1.7/-1.7 %			
MC generator	1.2/-1.2 %	4.6/-4.6 %			
ϵ_{evt}	8.9/-8.9 %	9.0/-9.0 %	9.8/-9.8 %	10.6/-10.6 %	
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	
M_{top} 170/180	8.3/-7.0 %	9.9/-7.1 %			
double tag					23.4/-23.4%

Table 6.9: Systematic rate uncertainties for events with 2 jets and 2 SecVtx tags.

Source	single-top	$t\bar{t}$	diboson	Z+jets
ISR less/more	-3.3/-4.8 %	-0.6/-4.5 %		
FSR less/more	-3.3/-3.8 %	-3.5/-2.2 %		
PDF	2.6/-2.6 %	1.8/-1.8 %		
MC generator	1.7/-1.7 %	-2.7/2.7 %		
ϵ_{evt}	3.0/-3.0 %	2.5/-2.5 %	7.8/-7.8 %	7.8/-7.8 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
M_{top} 170/180	8.3/-7.3 %	9.3/-8.4 %		

Table 6.10: Systematic rate uncertainties for events with 3 jets and 1 SecVtx tag.

Source	single-top	$t\bar{t}$	diboson	Z+jets	mistags
ISR less/more	5.8/-4.9 %	-0.5/-6.6 %			
FSR less/more	2.4/-2.2 %	-3.4/-2.7 %			
PDF	1.9/-1.9 %	1.7/-1.7 %			
MC generator	1.7/-1.7 %	2.0/-2.0 %			
ϵ_{evt}	8.9/-8.9 %	9.0/-9.0 %	11.0/-11.0%	11.1/-11.1 %	
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	
M_{top} 170/180	8.5/-7.6 %	9.3/-9.5 %			
double tag					23.4/-23.4%

Table 6.11: Systematic rate uncertainties for events with 3 jets and 2 SecVtx tags

6.4 Results of the individual analyses

In this section the results of the individual analysis are shown. We present the measured cross section as well as the sensitivity and the observed significance.

Likelihood Function method

As can be seen in figure 6.15 the analysis has an expected p-value of 0.0035% which corresponds to a sensitivity of 3.4σ. The observed p-value of 2.45% is in accordance with an 2.0σ excess over the standard model backgrounds. The combined s- and t-channel cross section is measured to be $\sigma_{s+t} = 1.8^{+0.9}_{-0.8}$ pb.

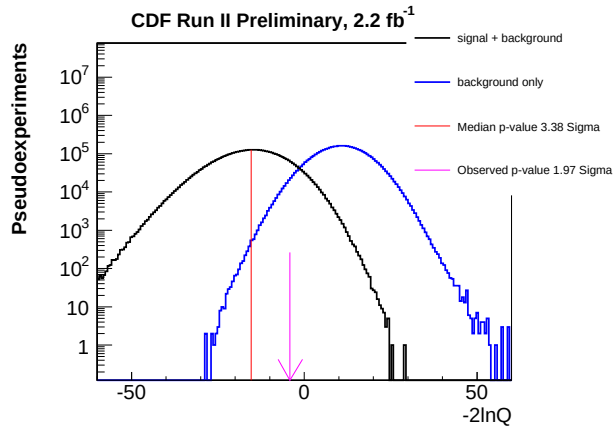


Figure 6.15: Expected and observed p-values of the Likelihood function method. The Q-value distributions for single top-quark production at the standard model rate (black curve) and without single top-quark production (blue curve) are plotted on a logarithmic scale and the observed Q-value is indicated by the arrow.

Matrix Element method

The Matrix Element method measures a combined s-channel and t-channel cross section of $2.2^{+0.8}_{-0.7}$ pb. As can be seen in figure 6.16, the probability that the observed excess originates from a background fluctuation (p-value) is 0.03% (3.4σ) and the expected p-value in pseudo experiments is 0.0003% (4.5σ).
As already mentioned before, we do not include the non-triggered muons, used in this analysis, in our combination study. Since there is no official result for the Matrix Element method without the non-triggered muons, we simulated 50 million pseudo experiments for this scenario. The obtained sensitivity for this scenario is 4.2σ.

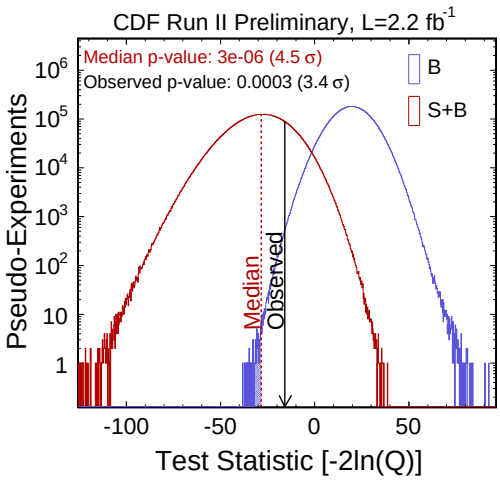


Figure 6.16: Expected and observed p-values of the Matrix Element method. The Q-value distributions for single top-quark production at the standard model rate (red curve) and without single top-quark production (blue curve) are plotted on a logarithmic scale and the observed Q-value is indicated by the arrow.

Neural Network method

The Neural Network analysis measures a combined single top-quark production cross section of $2.0^{+0.9}_{-0.8}$ pb. Figure 6.17 illustrates the expected and the observed p-values. The expected p-value of 0.00053% corresponds to an expected 4.4σ excess over the standard model backgrounds. The observed probability, that the seen excess stems from a background fluctuation is 0.063% (3.2σ).

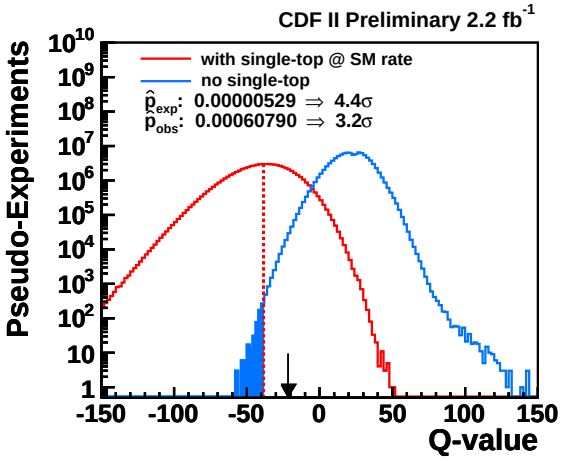


Figure 6.17: Expected and observed p-values of the Neural Network method. The Q-value distributions for single top-quark production at the standard model rate (red curve) and without single top-quark production (blue curve) are plotted on a logarithmic scale and the observed Q-value is indicated by the arrow.

6.5 Combination results

In this section, the application of the combination neural networks to measured events is shown and values for the sensitivity, the observed significance, the measured single top-quark production cross-section and for $|V_{tb}|$ are quoted.

6.5.1 Dertermination of the sensitivity

Hundred million pseudo experiments are simulated to derive the sensitivity of our neural network combination method. The output distributions of all four neural networks are included and finally fitted simultaneously. Figure 6.18 illustrates the obtained Q-value distributions which are plotted on a logarithmic scale. The Q-value distribution for single top-quark production at the standard model rate is drawn in red. Its median is represented by the dashed line. The expected p-value, i.e. the probability that the signal stems from a statistical fluctuation of the backgrounds, is 0.000097% which corresponds to an expected significance of 4.8σ.

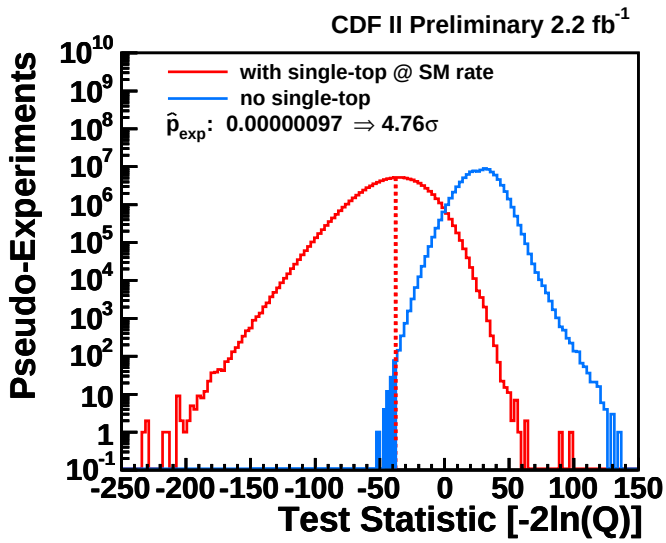


Figure 6.18: Sensitivity of the neural network combination. The Q-value distributions for single top-quark production at the standard model rate (red curve) and without single top-quark production (blue curve) are plotted on a logarithmic scale.

6.5.2 Application to observed events

Comparison with expectation

After the determination of the sensitivity the neural networks are applied to observed events. In figure 6.19 the output distributions for the background processes and the single top-quark production of the four combination networks are compared with the output distributions of observed events. For this purpose the signal and background

templates are normalized to the number of events predicted by the standard model for the particular processes (see table 5.1).

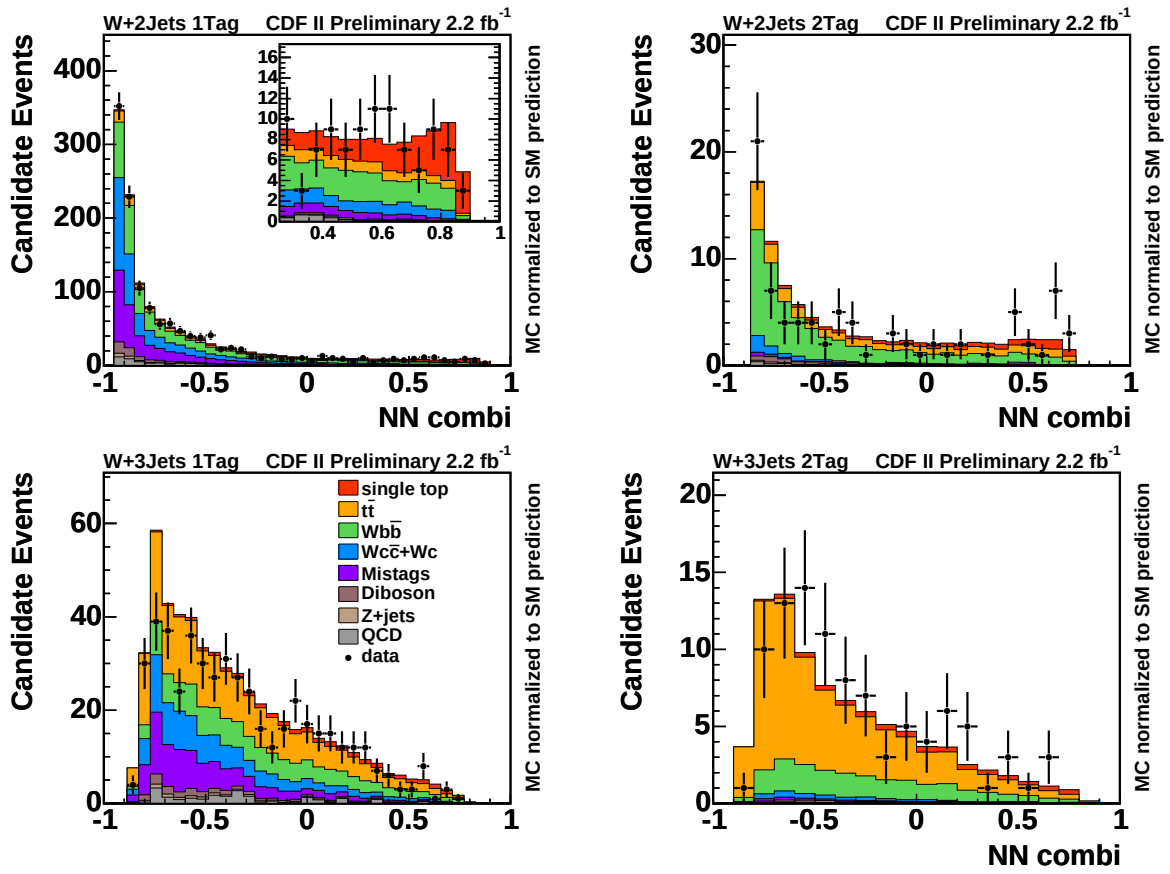
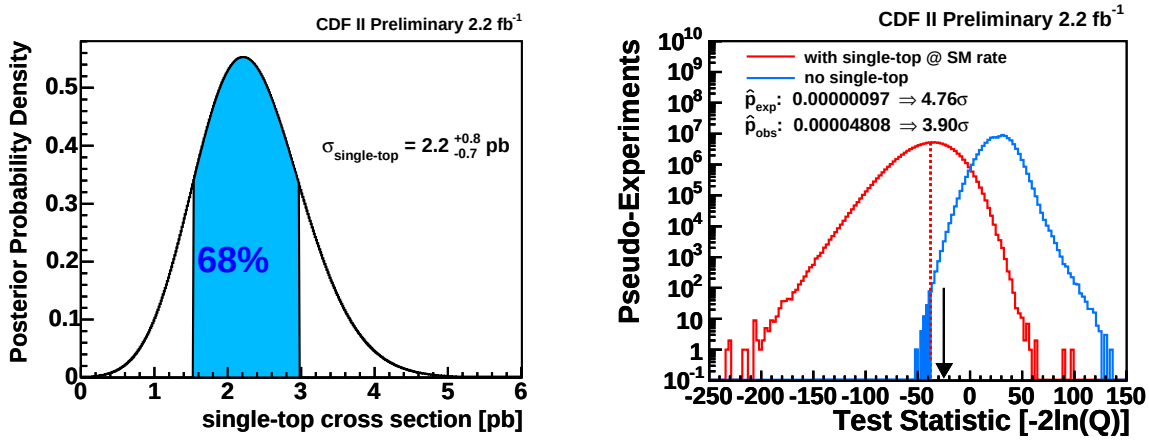


Figure 6.19: Comparison of the predicted and observed output distributions of the four combination neural networks. In the upper left corner the output of the 2jets-1tag network is shown. Within the plot, the output region NN combi > 0.3 is zoomed. The plot in the upper right corner displays the 2jets-2tags distributions. In the lower row the output distributions of the 3jets-1tag network (left hand side) and the 3jets-2tags network (right hand side) are shown.

Fit results

For the calculation of the single top-quark production cross section, the output distributions of measured events are fitted as explained in subsection 6.1.1. The combination result is obtained from a simultaneous fit to all four neural network outputs. Furthermore, a separate fit to each single network is performed.
The reduced Likelihood function is illustrated in figure 6.20(a). As can be seen, its maximum is located at $\sigma_{\text{single-top}} = 2.2^{+0.8}_{-0.7}$ pb which corresponds to $77^{+26.8}_{-23.8}\%$ of the single top-quark production cross section predicted by the standard model. Hence

the measured central value is smaller than the one predicted by the standard model, but, within the large uncertainties, both are still compatible. The observed Q-value is determined as explained in subsection 6.1.2. In figure 6.20(b) the observed Q-value, $Q_{\mathrm{m}} = -24.31$, is indicated with an arrow. The observed p-value, i.e the probability that the observed excess results from a statistical fluctuation of the background is therefore 0.004808%, which corresponds to an observed significance of 3.9σ.



(a) (b)

Figure 6.20: (a): Reduced Likelihood function obtained from marginalization. The blue band represent the $\pm 1\sigma$ uncertainties on the cross section measurement. (b): Observed sensitivity of the neural network combination. The Q-value distributions for single top quark production at the standard model rate (red curve) and without single top-quark production (blue curve) are plotted on a logarithmic scale and the observed Q-value is indicated by the arrow.

The following single top-quark production cross sections are fitted in the different jet and tag categories separately: $\sigma^{\text{2jets1tag}}_{\text{single-top}} = 1.7^{+0.8}_{-0.7}$ pb for events with two jets and one SecVtx-tag, $\sigma^{\text{2jets2tags}}_{\text{single-top}} = 4.6^{+2.5}_{-2.2}$ pb for events with two jets and two SecVtx-tags, $\sigma^{\text{3jets1tag}}_{\text{single-top}} = 2.9^{+2.7}_{-2.4}$ pb for events with three jets and one SecVtx-tag and $\sigma^{\text{3jets2tags}}_{\text{single-top}} = 8.2^{+4.4}_{-4.0}$ pb for events with three jets and two SecVtx-tags. In figure 6.21 the output distributions of the four networks are normalized to these fit results, i.e. the background processes are still normalized to the number of events predicted from the standard model only the single top-quark template is scaled with the measured cross section. The measured cross sections and their uncertainties are summarized and compared with the standard model prediction in figure 6.22.

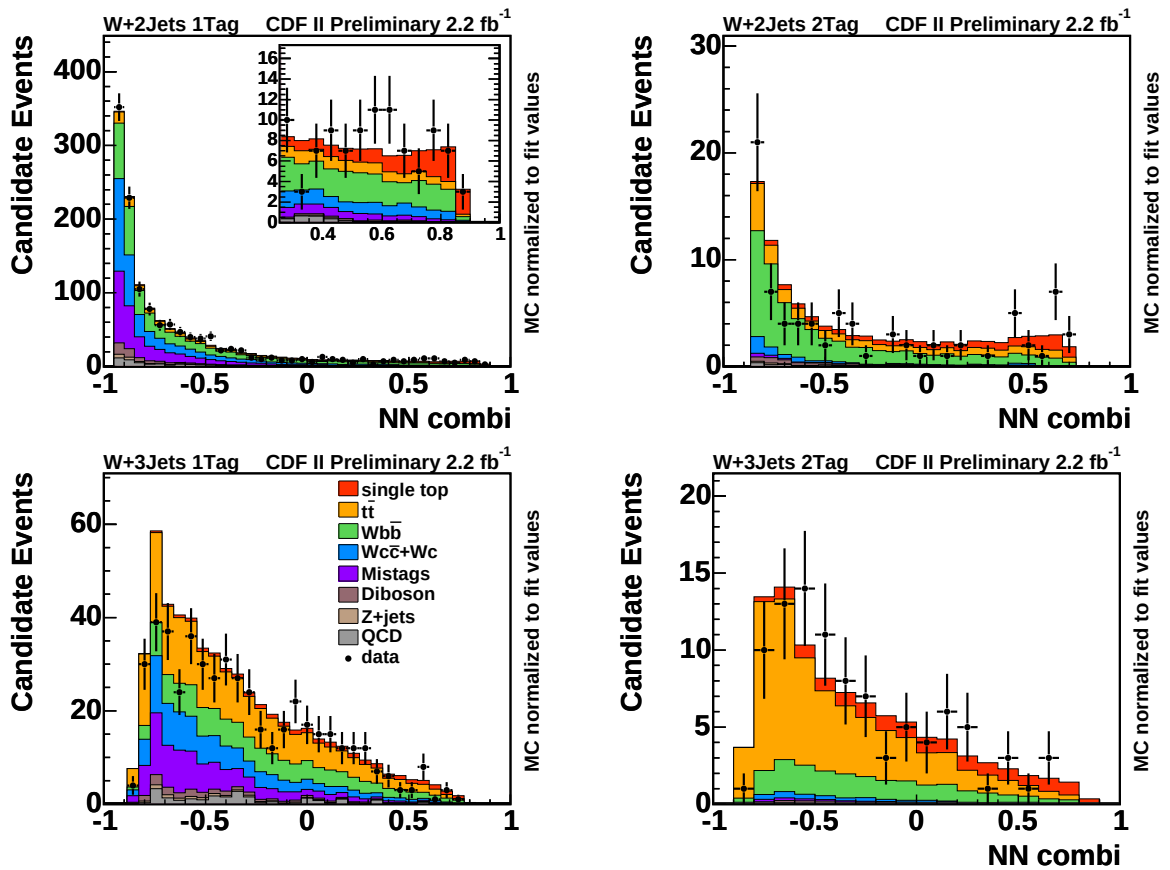


Figure 6.21: Output distributions of the four combination neural networks normalized to the fit values. The background processes are still normalized to the number of events predicted from the standard model only the single top-quark template is scaled with the measured cross section. In the upper left corner the output of the network using two jets and one tag is shown. Within the plot, the output region NN combi > 0.3 is zoomed. The plot in the upper right corner displays the outyput distributions of the network using two jets and two tags. In the lower row the output distributions of the network using three jets and one tag (left hand side) and the network using three jets and two tags (right hand side) are shown.

6.5.3 $|V_{\mathrm{tb}}|$ measurement

Due to the already mentioned proportionality of $\sigma_{\mathrm{single-top}}$ and $|V_{\mathrm{tb}}|^2$, a value for $|V_{\mathrm{tb}}|$ can be derived under the assumption that $V_{\mathrm{tb}} \gg V_{\mathrm{ts}}, V_{\mathrm{td}}$ via

$$|V_{\mathrm{tb}}|^{\mathrm{meas}} = |V_{\mathrm{tb}}|^{\mathrm{theo}} \cdot \sqrt{\frac{\sigma^{\mathrm{meas}}_{\mathrm{single-top}}}{\sigma^{\mathrm{theo}}_{\mathrm{single-top}}}} \ , \tag{6.12}$$

where $|V_{\mathrm{tb}}|^{\mathrm{theo}}$ is the V_{tb} value from equation 1.2, $\sigma^{\mathrm{meas}}_{\mathrm{single-top}}$ is the measured single top-quark production cross section and $\sigma^{\mathrm{theo}}_{\mathrm{single-top}}$ is the sum of the standard model

values for $\sigma_{\mathrm{s-channel}}$ and $\sigma_{\mathrm{t-channel}}$, which are given in chapter one. In this manner $V_{\mathrm{tb}} = 0.88^{+0.14}_{-0.12}(\mathrm{exp.}) \pm 0.07(\mathrm{theo.})$ can be found. As expected, the quoted uncertainties are too large for a verification or falsification of the unitary assumption of the CKM-matrix.

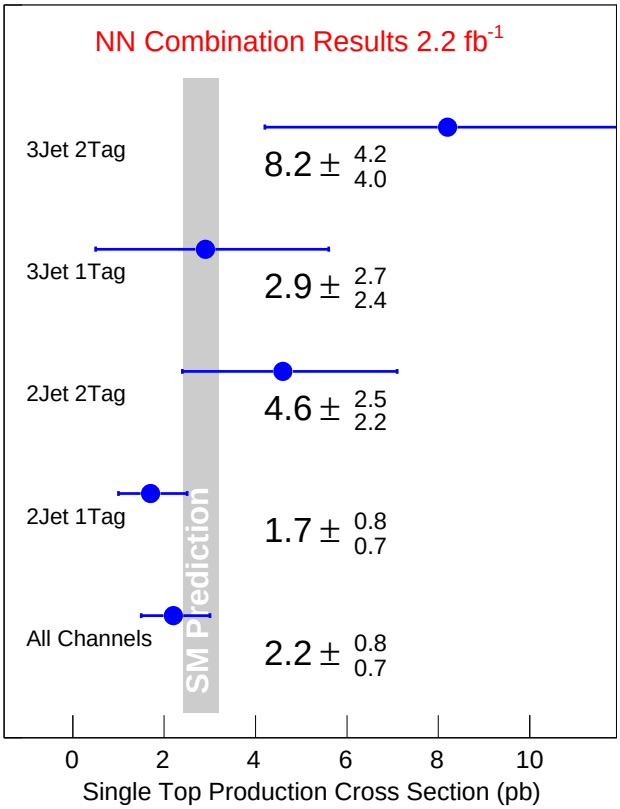


Figure 6.22: Summary of the combination cross section measurements for the different jet and tag bins.

Chapter 7

Summary and Outlook

In the first part of this diploma thesis, the current version of the KIT Flavor Separator, a neural network which is able to distinguish between tagged b-quark jets and tagged c/light-quark jets, is presented. In comparison with previous versions four new input variables are utilized and new Monte Carlo samples with a larger number of simulated events are used for the training of the neural network. It is illustrated that the output of the neural network is continuously distributed between 1 and -1, whereas b-quark jets accumulate at 1, however, c-quark jets and light-quark jets have outputs next to -1.
To ensure that the network output describes observed events correctly, the shapes of all input variables are compared in simulation and data. Thus the mismodelling of any input variable is excluded. Moreover, the b jet and light jet output distributions are compared with the output of samples of observed events, which are enhanced in the particular flavor. In contrast to previous versions, no b-jet output correction function has to be calculated, because the agreement between simulation and collision data is excellent for b-quark jets. For the light-jet output, correction functions are developed.
Different applications of the KIT Flavor Separator are mentioned. For example it provides a precious input to all three CDF single top quark analyses. Furthermore, it is shown that the KIT Flavor Separator is a universal tool, which can be used in every high-p_{T} analysis that requires the identification of b-quark jets with high efficiency. As it is pointed out, a further application is the estimation of the flavor composition of a given sample of observed events.
In addition a neural network, which is able to separate c-quark jets from light-quark jets, is trained. It is shown, that all three flavors can be separated in the c-net$-$Flavor Separator plane. As a result, the uncertainties on the estimation of the flavor composition in events with one tagged jet are cut into half.
In the second part of this diploma thesis, a method for the combination of three multivariate single-top analyses using an integrated luminosity of 2.2 fb^{-1} is presented. For this purpose the discriminants of the Likelihood Function analysis, the Matrix Element method and the Neural Network analysis are used as input variables to a neural network. Overall four different networks are trained, one for events with two or

three jets and one or two SecVtx tags, respectively. Using a binned likelihood function, the outputs of these networks are fitted to the output distribution of observed events. A single top-quark production cross section of $\sigma_{\text{single-top}} = 2.2^{+0.8}_{-0.7}$ pb is measured. Ensemble tests are performed for the calculation of the sensitivity and observed significance, which are found to be 4.8σ and 3.9σ, respectively. Hence the improvement of this combination is roughly 8% in comparison with sensitivities found by the individual analyses. Due to the proportionality of $\sigma_{\text{single-top}}$ and $|V_{\text{tb}}|^2$ and under the assumption $V_{\text{tb}} \gg V_{\text{ts}}, V_{\text{td}}$, a value for $|V_{\text{tb}}|$ is quoted: $|V_{\text{tb}}| = 0.88^{+0.14}_{-0.12}(\text{exp.}) \pm 0.07(\text{theo.})$. It can be seen, that the given uncertainties are too large for a verification or falsification of the unitarity assumption of the CKM-matrix.
Parallel to this combination a further combination method (NEAT-combination) has been developed [81]. This combination uses a neural network trained with a neuro-evolution technique, which optimizes the neural network architecture and weights through the use of genetic algorithms. In this analysis an improvement of roughly 12% could be reached. In figure 7.1 the current situation for the measurement of the single top-quark production cross section is summarized.
After collecting more data, CDF will be able to observe single top-quark production with a significance larger than 5.0σ. Nevertheless, the cross section measurement will still have large uncertainties on the level of 20%. Precise measurements on the few percent level will only be possible at the LHC. Recent studies for the CMS detector showed, that the t-channel cross section can be measured with an accuracy of 7% in 1 fb^{-1} of LHC data [82].

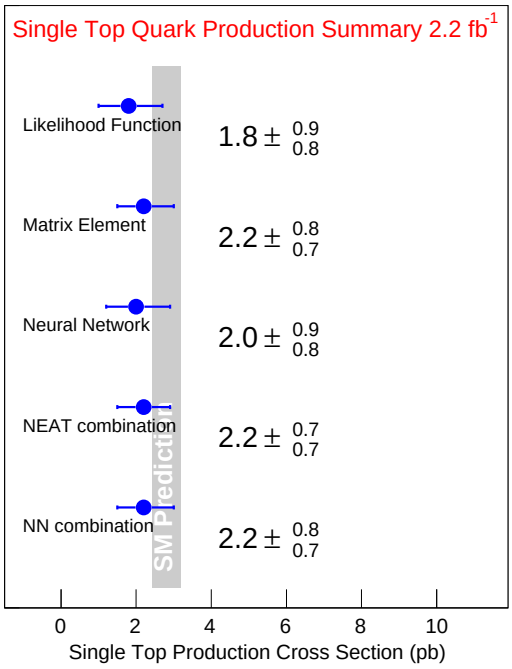


Figure 7.1: Summary of the single top production cross section measurements.

Appendix A

A.1 $|\eta|$ dependences for pass 1 vertices

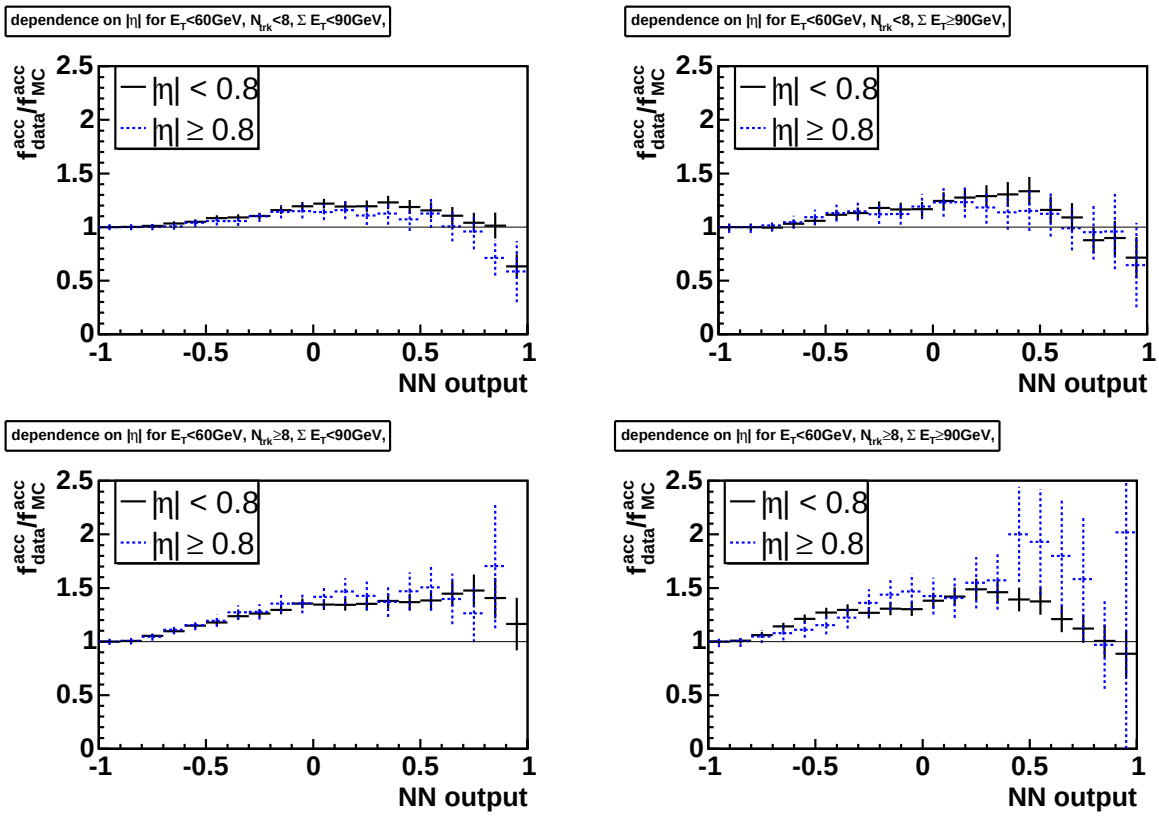


Figure A.1: $f^{acc}_{data}/f^{acc}_{MC}$ distributions of $|\eta|$ for all combinations of E_T, N_{trk} and $\sum E_T$, which are not used in the correction function pass 1 vertices.

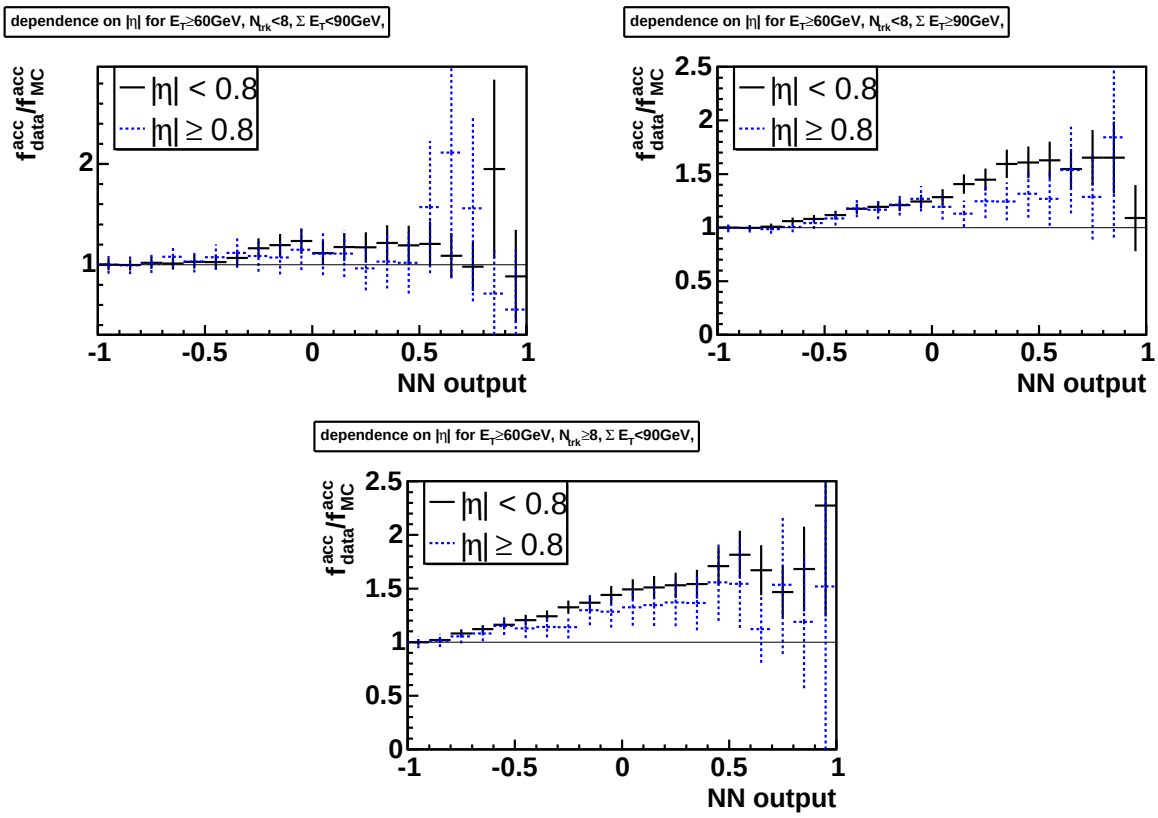


Figure A.2: $f^{acc}_{data}/f^{acc}_{MC}$ distributions of $|\eta|$ for all combinations of E_T, N_{trk} and $\sum E_T$, which are not used in the correction function pass 1 vertices.

A.2 $|\eta|$ dependences for pass 2 vertices

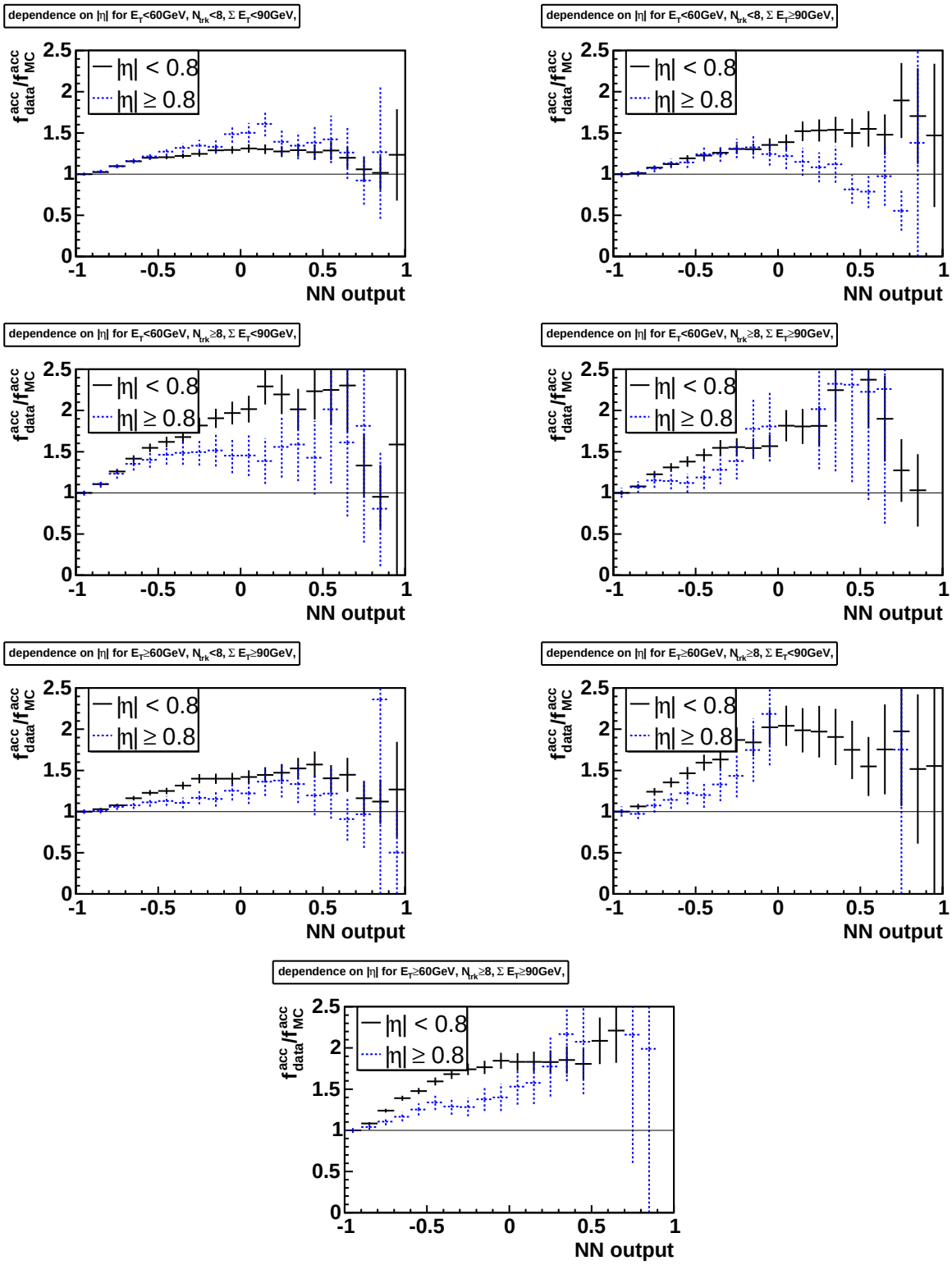


Figure A.3: $f^{acc}_{data}/f^{acc}_{MC}$ distributions of $|\eta|$ for all combinations of E_T, N_{trk} and $\sum E_T$, which are not used in the correction function for pass 2 vertices.

List of Figures

1.1 Feynman diagram 4
1.2 Top quark pair production 6
1.3 CTEQ5L PDF parametrization 6
1.4 Single-top t-channel 7
1.5 Single-top s-cahnnel 8
1.6 Associated single-top production 8

2.1 Aerial view of Fermilab 10
2.2 Accelerator Chain 11
2.3 Initial Tevatron luminosity 13
2.4 Integrated luminosity 14
2.5 Cutaway view and coordinatesystem of CDF 15
2.6 Tracking Volume 16
2.7 Trigger System 18

3.1 Feed-forward Neural Network 20
3.2 The sigmoid function 23
3.3 Results of a network training 24
3.4 Jet evolution and reconstruction 28
3.5 B-tagging event topology 30
3.6 SecVtx positiv and negative tags 31

4.1 Jet multiplicity distribution 34
4.2 Input Variables I 37
4.3 Input Variables II 38
4.4 Input Variables III 39
4.5 Flavor Separator output 40
4.6 Flavor Separator output for different processes 43
4.7 Output comparison for different numbers of primary vertices 44
4.8 Output comparison for different lepton types 45
4.9 Flavor Separator output for pass 1 and pass 2 vertices 46
4.10 $E_{\mathrm{T}}^{\mathrm{jet}}$ cuts on mistag samples 47
4.11 Comparison of input distributions with inclusive electron data 48
4.12 Comparison of input distributions with inclusive electron data 49

4.13 Compostion of away jet sample 51
4.14 Comparison of b jet output 51
4.15 Veryfication of the mistag model 52
4.16 Comparison of mistag output in data and simulation 53
4.17 $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ dependencies pass 1 55
4.18 $f^{\mathrm{acc}}_{\mathrm{data}}/f^{\mathrm{acc}}_{\mathrm{MC}}$ dependencies pass 2 56
4.19 Eta dependency 57
4.20 Correction functions 57
4.21 Corrected outputs 58
4.22 Heavy flavor composition fits 59
4.23 C-net training results 60
4.24 Two dimensional distributions for jets with different flavors 62
4.25 2D fits projected on the c-net and the Flavor Separator axes 63

5.1 Illustration of the QCD veto in the central electron sample with 2 jets . 68
5.2 Exemplary Feynman diagrams of background processes to single top-quark production 69
5.3 More backgrounds to single-top production 70
5.4 Likelihood function output distributions 75
5.5 EPD distribution 77
5.6 Neural network output distributions 79

6.1 Output for the network using two jets and one SecVtx tag 86
6.2 Output for the network using two jets and one SecVtx tag 87
6.3 Output of the network using three jets and one SecVtx tag 89
6.4 Output of the network using three jets and two SecVtx tag 90
6.5 JES shape systematics I 92
6.6 JES shape systematics II 93
6.8 PDF and Q^2 shape systematics 93
6.7 ISR/FSR shape systematics 94
6.9 KIT Flavor Separator shape systematics 95
6.10 QCD-multijet mixture and Mistag Model shape systematics 96
6.11 ΔR shape systematics I 96
6.12 97
6.13 η_{j} shape systematics I 98
6.14 η_{j} shape systematics II 99
6.15 Expected and observed p-values of the Likelihood function method . . 101
6.16 Expected and observed p-values of the Matrix Element method 102
6.17 Expected and observed p-values of the Neural Network method 102
6.18 Sensitivity of the combination 103
6.19 Comparison of predicted and observed output distributions 104
6.20 Reduced Likelihood function and observed sensitivity 105
6.21 Output distributions normalized to fit results 106
6.22 Summary of the combination cross section measurements 107

7.1 Summary of the single top production cross section measurements . . . 109

A.1 Remaining dependences on $|\eta|$ (pass1) 110
A.2 Remaining dependences on $|\eta|$ (pass1) 111
A.3 Remaining dependences on $|\eta|$ (pass2) 112

List of Tables

1.1 The fundamental fermions 2
1.2 Gauge-Boson 3

2.1 CDF Muon System 17

4.1 Used MC samples in the training 36
4.2 Ratios of the training samples 40
4.3 Set of discriminating variables used to train the KIT Flavor Separator. 41
4.4 Set of discriminating variables discarded for the training of the KIT Flavor Separator. 42
4.5 Compositon of the away jets in Monte Carlo and data 50
4.6 Heavy flavor content fit results 60
4.7 Set of discriminating variables used for the training of the c-net. 61
4.8 2D fit results 63

5.1 Expected event yield 73
5.2 Neural Network input variables 78

6.1 Relative standard deviation for Gaussian constraints 81
6.2 Ranking of the input variables for the network using two jets and one SecVtx tag 86
6.3 Ranking of the input variables for the network using two jets and two SecVtx tag 87
6.4 Ranking of the input variables for the network using three jets and one SecVtx tag 88
6.5 Ranking of the input variables for the network using three jets and two SecVtx tag 89
6.6 Processes affected by shape uncertainties 91
6.7 Jet energy scale rate uncertainties 91
6.8 Rate uncertainties for 2 jets 1 tag events 99
6.9 Rate uncertainties for 2 jets 2 tags events 100
6.10 Rate uncertainties for 3 jets 1 tag events 100
6.11 Rate uncertainties for 3 jets 2 tags events 100

Bibliography

[1] A.Einstein; Prinzipielles zur allgemeinen Relativitätstheorie; *Annalen der Physik. 55, S.241-244*, 1918.

[2] W.-M. Yao et al.; *J. Phys.* G 33, 1 (2006) and 2007 partial update for the 2008 edition.

[3] The Tevatron Elektroweak Working Group; Combination of CDF and DØ Results on the Mass of the Top Quark; *arXivx:hep-ex/0803.1683*.

[4] S. Fukuda et al.; Solar B-8 and He p neutrino measurements from 1258 days of Super-Kamiokande data; *Phys. Rev. Lett.* 86:5651-5655, 2001.

[5] S. L. Glashow; Partial symmetries of weak interactions; *Nucl. Phys.*, 22:579-588, 1961.

[6] A.Salam and J.C.Ward; Electromagnetic and weak interactions; *Phys. Lett.*, 13:168-171, 1964.

[7] S.Weinberg; A model of Leptons; *Phys. Rev. Lett.*, 19:1264-1266, 1967.

[8] P.W.Higgs; Broken symmetries and the masses of gauge bosons; *Phys.Rev.Lett.* 13:508-509, 1964.

[9] P.W.Higgs; Spontaneous symmetry breakdown without massless bosons; *Phys.Rev.* 145:1156-1163, 1966.

[10] W. Wagner; Top quark physics in hadron collisions; *Rept. Prog. Phys. 68:2409-2494*, 2005.

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

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

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

[14] T. M. P. Tait; Signals for the Electroweak Symmetry Breaking Associated with the Top Quark; *hep-ph/9907462*, 1999.

[15] DØ Collaboration, V.M. Abazov et al., *Phys. Rev. Lett.* 97, 021802 (2006).

[16] CDF Collaboration, Measurement of the $B_s^0 - \overline{B}_s^0$ Oscillation Frequency, *Phys.Rev.Lett.* 97, 062003 (2006).

[17] CDF Collaboration, D. Acosta et al., *Phys. Rev. Lett. 95*, 102002 (2005) [arXiv:hep-ex/0505091].

[18] DØ Collaboration, V.M. Abazov et al., *Phys.Lett.* B639 616-622 (2006).

[19] J.Alwall et al.; Is $V_tb \simeq 1$?; *Eur. Phys. J. C 49*, 791-801 (2007).

[20] J. Thompson; Introduction to colliding beams at Fermilab; *FERMILAB-TM-1909*.

[21] J. Morgan; The Antiproton Source Rooki book; *unpublished*, 1999.

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

[23] S.Klimenko et al.; The CDF Run II Luminosity Monitor; *internal CDF Note 4330*, 1998.

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

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

[26] A.Sill; CDF Run II silicon tracking projects; *Nucl. Instrum. Meth.* **A447**:1-8, 2000.

[27] T.Affolder et al.; Intermediate silicon layers detector for the CDF experiment; *Nucl. Instrum. Meth.* **A453**:84-88, 2000.

[28] L.Balka et al.; The CDF Central Electromagnetic Calorimeter; *Nucl. Instrum. Meth.* **A267**:272-279, 1988.

[29] S.Bertolucci et al.; The CDF Central and Endwall Hadron Calorimeter; *Nucl. Instrum. Meth.* **A267**:301-314, 1988.

[30] G.Ascoli et al.; CDF Central Muon Detector; *Nucl. Instrum. Meth.* **A268**:30-40, 1988.

[31] M.Albrow et al.; CDF Run-II Trigger Table and Datasets Plan; *internal CDF Note 4718*, 2001.

[32] M.Feindt; A neural Bayesian estimator for conditional probability densities; *arXiv:physics/0402093*.

[33] M.Feindt, U.Kerzel; The NeuroBayes neural network package; *Nucl. Instrum. Meth.* **A559**:190-194.

[34] V.Blobel, E.Lohrmann; Statistische und numerische Methoden der Datenanalyse; *Teubner, Stuttgart, Leipzig*, 1998.

[35] T.Sjöstrand et al.; PYTHIA 6.3 Physics and Manual; *hep-ph/0308153*, 2003.

[36] M.L.Mangano et al.; ALPGEN, a generator for hard multiparton processes in hadronic collisions; *arXiv:hep-ph/0206293v2*, 2002.

[37] F.Maltoni and T.Stelzer; Madevent: Automatic event generation with MadGraph; *J. High Energy Phys.*, 02:027, 2003.

[38] T.Stelzer and W.F.Long; Automatic generation of tree level helicity amplitudes; *Comput. Phys. Commun.*, 81:357-371, 1994.

[39] E.D. Frank, Bob Jacobsen, E. Sexton-Kennedy for BaBar computing group; Architecture of the Computing in BaBar reconstruction System; *Proceedings of the CHEP'97 conference*; Berlin, Germany, April 7-11 (1997).

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

[41] Rob Veenhof; GARFIELD, a drift chamber simulation program; *Version 6.22*, CERN.

[42] E.Gerchtein, M.Paulini; CDF Detector Simulation Framework and Performance; *arXiv:physics/0306031*, 2003.

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

[44] A.Mukherjee; CTC and VTX Tracking; *internal CDF Note 5490*, 1995.

[45] F.D.Snider; Tracking at CDF: Algorithms and experience from Run I and Run II; *Nucl. Instrum. and Methods* **A566**, 133-141 (2006).

[46] W.-M.Yao and K.Bloom; "Outside-in" Silicon Tracking at CDF; *internal CDF Note 5991*, 2002.

[47] B.R.Ko, S.H.Oh, C.Wang; SVXII Stand-alone Tracking; *internal CDF Note 6440*, 2003.

[48] Y.Huang, C.Hays, A.Kotwal; Inside-Out Tracking; *internal CDF Note 6707*, 2003.

[49] A.Bhatti et al.; Determination of the jet energy scale at the Collider Detector at Fermilab; *Nucl. Instrum. and Methods* **A566**, 375-412 (2006).

[50] T.Affolder et al.; Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV; *Phys. Rev. D 64 032002*, 2001.

[51] F.Garberson, J.R.Incandela, C.Neu; SecVtx B-Tag Efficiency Measurement Using Muon Transverse Momentum for 1.2 fb^{-1} Analyses; *internal CDF Note 8640*, 2007.

[52] S.Grinstein, J.Guimarães da Costa, D.Sherman; The SecVtx Scale Electron-Method Scale Factor for Winter 2007; *internal CDF Note 8625*, 2007.

[53] F.Garberson et al.; Combination of the SecVtx 1.2 fb^{-1} b-Tagging Scale Factors; *internal CDF Note 8666*, 2007.

[54] S.Grinstein, D.Sherman; SecVtx Scale Factors and Mistag Matrices for the 2007 Summer Conferences; *internal CDF Note 8910*, 2007.

[55] C.Neu et al.; SecVtx Optimizing Studies for 5.3.3. Analyses; *internal CDF Note 7578*, 2005.

[56] J.Freeman et al.; SecVtx Mistag Matrices for 2.2 fb^{-1} of Data Up to P13; *internal CDF Note 9280*, 2008.

[57] J.Freeman and S.Grinstein; SecVtx Mistag Asymmetry Corrections through 2.2 fb^{-1}; *internal CDF Note 9277*, 2008.

[58] S.Richter; PhD-Thesis: Search for Electroweak Single Top-Quark Production with the CDF II Experiment; FERMILAB-THESIS-2007-35, 2007.

[59] T.Chwalek et al.; Update of the Neural Network b Tagger for Single-Top Analyses; *internal CDF Note 8903*, 2007.

[60] M.Feindt et al.; A Neural Network b Tagger for Single-Top Analyses; *internal CDF Note 7816*, 2006.

[61] A.Holloway et al.; SecVtx Tagging Scale Factor for 5.3.3 Analysis; *internal CDF Note 7333*, 2004.

[62] T.Peiffer; Diploma-Thesis: Search for the Higgs Boson in the WH Channel with the CDF II Experiment; FERMILAB-MASTERS-2008-01, 2008.

[63] D.Amidei et al.; Calibration of Heavy-Flavor Production in W+1 Jet Data; *internal CDF Note 9187*, 2008.

[64] H. Frisch, R. St. Denis, T. Vaiciulis and W. Badgett; The good run list; *Internal CDF Note 5613* (unpublished), 2001.

[65] S.Budd et al.; Event detection efficiency for single-top events and MC based background estimate for summer 2006; *internal CDF Note 8286*, 2006.

[66] U. Grundler, L. Lovas, A. Taffard; High-P_T muons recommended cuts and efficiencies for Winter 2007; *Internal CDF Note 8618*, 2006.

[67] J.Adelman et al.; Method II For You; *Internal CDF Note 9185*, 2008.

[68] W.K.Sakumoto; Event $|Z_{vtx}| \leq 60$ cm Cut Acceptance for Run II; *Internal CDF Note 8318*, 2006.

[69] D. Hare, E. Halkiadakis, T. Spreitzer; Electron ID Efficiency and Scale Factors for Winter 2007 Analyses; *internal CDF Note 8614*, 2006.

[70] Veronique Boisvert; Trigger Efficiencies for the High E_T Central Electrons in Gen6; *internal CDF Note 7939*, 2006.

[71] Bo-Young Han, Eva Halkiadakis; MET PEM trigger efficiency for Pheonix electrons; *internal CDF Note 7940* , 2006.

[72] Bo-Young Han, Veronique Boisvert; Trigger Efficiencies for the High E_T Central Electron in the Gen6 data(run:217990 - 222426); *internal CDF Note 8629*, 2006.

[73] S.Budd et al.; Likelihood Function Analysis of Single-top Production with 2.2 fb^{-1} of Data; *Internal CDF Note 9112*, 2008.

[74] D.Hirschbühl et al.; Search for single top-quark production with neural networks using 2.2 fb^{-1}; *Internal CDF Note 9107*, 2008.

[75] F.Canelli et al.; Search for Single Top Quark Production using the Matrix Element Technique; *Internal CDF Note 9117*, 2008.

[76] E. Murayama, I. Watanabe and K. Hagiwara; HELAS: HELicity Amplitude Subroutines for Feynman Diagram Evaluations; *KEK Report 91-11*, 1992.

[77] F.Canelli et al.; Increasing Muon Acceptance with the MET Plus Jet Triggers; *internal CDF Note 9105*, 2007.

[78] F. James and M. Roos; Minuit: A system for function minimization and analysis of the parameter errors and correlations; *Comput. Phys. Commun.***10:343-367**, 1975.

[79] Y.-K. Kim, U.-K. Yang; Initial state gluon radiation studies on Drell-Yan data for top-pair production in hadron Collider; *internal CDF Note 6804*, 2003.

[80] S. Elbracht et al.; Fitting the flavor composition of lepton trigger data in bins of MET and lepton isolation; *internal CDF Note 9131*, 2007.

[81] D.Amidei et al.; Combination of Single-Top Results with 2.2 $fb^{-}1$; *internal CDF Note 9173*.

[82] P.Sturm; Diploma-Thesis: Studies for the Measurement of Single-Top-Quark-Events with the CMS-Experiment; IEKP-KA/2008-8, 2008.

Danksagung

An dieser Stelle möchte ich mich bei den Personen bedanken, die maßgeblich zum Gelingen dieser Arbeit beigetragen haben.

Zunächst möchte ich Herrn Priv.-Doz. Dr. Wolfgang Wagner danken, der das Referat und die hervorragende Betreuung dieser Arbeit übernommen hat.

Des Weiteren danke ich Herrn Prof. Dr. Thomas Müller für die Aufnahme in seine Top-Quark Arbeitsgruppe, sowie für die Übernahme des Korreferats dieser Arbeit.

Weiterhin gebührt mein großer Dank Herrn Dr. Dominic Hirschbühl für die vielen wertvollen Ratschläge, die unzähligen, hilfreichen Diskussionen sowie für das intensive Korrekturlesen dieser Arbeit.

Ich bedanke mich bei allen Mitgliedern der Top-Quark Gruppe für die gewinnbringende Zusammenarbeit und die sehr gute Stimmung innerhalb der Arbeitsgruppe. Herausheben möchte ich hier Jan Lück und Dr. Svenja Richter, die mir beide immer mit Rat und Tat zur Seite standen.

Des Weiteren danke ich den Mitgliedern der Karlsruher B-Physik Arbeitsgruppe, mit denen ich mir im vergangenen Jahr ein Büro teilen durfte, für das äußerst angenehme Arbeitsklima. Ich danke allen Mitgliedern des EKP, hier vor allem dem Admin-Team und dem Sekretariat, für die freundliche Unterstützung.

Weiterhin bedanke ich mich bei den Mitgliedern meiner Lerngruppe um Michael Dürrschnabel, Harald Flügge, Jens Müller und Thomas Peiffer, für die gute Zusammenarbeit und die lustigen Stunden die wir im Laufe unseres Physikstudiums zusammen erlebt haben.

Abschließend möchte ich mich bei meiner Schwester und meinen Eltern für ihre grenzenlose Unterstützung bedanken, ohne die ein erfolgreicher Abschluss meines Studiums und dieser Arbeit unmöglich gewesen wäre.

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

Manuel Renz

Karlsruhe, den 17. Juni 2008