

SUCHE NACH ELEKTROSCHWACHER PRODUKTION
EINZELNER TOPQUARKS MIT DEM
CDF II EXPERIMENT

SVENJA RICHTER

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

genehmigte

Dissertation

von

Dipl. Phys. Svenja Richter
aus Karlsruhe

Tag der mündlichen Prüfung: 02.11.2007

Referent: Prof. Dr. M. Feindt, Institut für Experimentelle Kernphysik

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

Zusammenfassung

Die Welt verstehen — Dieses Ziel treibt die Menschheit seit Anbeginn des bewussten Denkens an. Von besonderem Interesse ist dabei schon immer die Beschaffenheit der Materie gewesen, von welcher wir heutzutage eine komplexe Vorstellung haben: Atome bestehen aus Elektronen und Nukleonen, wobei selbst Nukleonen nicht elementar, sondern aus Quarks und Gluonen zusammengesetzt sind. Das derzeit gültige theoretische Modell, das diese elementaren Komponenten der Materie sowie deren Wechselwirkungen beschreibt, ist das Standardmodell der Teilchenphysik.

Die Substruktur der Materie kann nur in Streuexperimenten sichtbar gemacht werden, wobei diese Art von Experimenten in der Hochenergiephysik an Teilchenbeschleunigern durchgeführt wird. Der zur Zeit höchstenergetischste Beschleuniger, das Tevatron, befindet sich in der Nähe von Chicago auf dem Gelände des Fermi National Accelerator Laboratory (Fermilab), wo Proton–Antiproton–Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 1.96$ TeV von den zwei Vielzweckdetektoren DØ und CDF II aufgezeichnet werden.

Das schwerste Quark des Standardmodells, das Topquark, wurde im Jahr 1995 am Tevatron entdeckt [2, 3] und kann bis zur Inbetriebnahme des Large Hadron Colliders in Genf an keinem anderen Beschleuniger erzeugt werden. Die Produktion von Topquarks geschieht hauptsächlich in Paaren mittels der starken Wechselwirkung. Darüber hinaus wird vom Standardmodell eine elektroschwache Produktion einzelner Topquarks über einen Wtb -Vertex vorhergesagt. Diese Art der Topquarkerzeugung ist von großem Interesse, da sie eine Vielzahl an Möglichkeiten bietet, das Standardmodell zu testen und sensitiv auf Anzeichen neuer Physik jenseits der bestätigten Theorie ist. Die Produktionswahrscheinlichkeit einzelner Topquarks ist direkt proportional zum Quadrat des $|V_{tb}|$ -Elements der Cabbibo–Kobayashi–Maskawa-Matrix, wodurch eine direkte Bestimmung dieser elementaren Größe möglich wird. Die bisherigen indirekten Messungen von $|V_{tb}|$ basieren darauf, dass vom Standardmodell genau drei Quarkfamilien vorhergesagt werden, woraus sich durch die Unitarität der CKM-Matrix ein Wert von $|V_{tb}| \approx 1$ ergibt. Im Prinzip ist es jedoch möglich, dass $|V_{tb}|$ signifikant kleiner als eins ist, wenn man mehr als drei Quarkgenerationen annimmt. Darüber hinaus ist das Verständnis der Produktion einzelner Topquarks wichtig für Analysen zur Suche nach dem vom Standardmodell vorhergesagten Higgsboson im WH -Kanal, in welchem das Higgsboson zusammen mit einem W -Boson erzeugt wird. Bei einer Masse im Bereich zwischen $90 \text{ GeV}/c^2$ und $130 \text{ GeV}/c^2$ zerfällt das Higgsboson hauptsächlich in b -Quark-Paare, wodurch die

Ereignissignatur der von Einzeltopquarkereignissen gleicht.

Zur Produktion einzelner Topquarks tragen am Tevatron im Wesentlichen zwei Prozesse bei, die anhand der Virtualität $Q^2 = -q^2$ des ausgetauschten W -Bosons mit dem Viererimpuls q unterschieden werden. Diese beiden Prozesse erhalten ihren Namen durch die Mandelstamvariablen, mit denen die Übergangsmatrixelemente beschrieben werden können. Im ersten Prozess wird ein einzelnes Topquark durch den Austausch eines virtuellen W -Bosons zwischen einem leichten Quark und einem b -Quark aus dem Quarksee der kollidierenden Protonen und Antiprotonen erzeugt. Das ausgetauschte W -Boson ist raumartig ($q^2 = t < 0$), so dass man von einer Produktion im t -Kanal spricht. Im Gegensatz dazu annihilieren bei der Produktion im s -Kanal ein Quark und ein Antiquark zu einem virtuellen zeitartigen W -Boson ($q^2 = s \geq (m_{\text{top}} + m_b)^2$), welches ein einzelnes Topquark und ein b -Quark erzeugt. Die durch störungstheoretische Rechnungen in nächsthöherer Ordnung vorhergesagten Wirkungsquerschnitte betragen $\sigma_t = 1.98^{+0.28}_{-0.22}$ pb für den t -Kanal und $\sigma_s = 0.88^{+0.12}_{-0.11}$ pb für den s -Kanal.

Um Einzeltopquarkereignisse zu selektieren, wird die Tatsache ausgenutzt, dass das Topquark mit einem Verzweigungsverhältnis von annähernd 100% in ein W -Boson und ein b -Quark zerfällt. Da in dieser Analyse lediglich leptonische Zerfälle des W -Bosons betrachtet werden, erwartet man eine Ereignissignatur von einem isolierten Lepton, vom Neutrino hervorgerufene fehlende Transversalenergie und zwei Jets, von denen mindestens einer als b -Quark-Jet identifiziert sein muss. Für die Identifikation von b -Quark-Jets wird die lange Lebensdauer von b -Hadronen ausgenutzt, die verursacht dass die Spuren der Zerfallsprodukte des b -Hadrons nicht durch den primären Wechselwirkungspunkt führen. Unter Verwendung dieser Spuren, die einen großen Impaktparameter haben, ist es möglich, den Zerfallsvertex des b -Hadrons zu bestimmen. Aus diesem Grund werden Jets, in denen solch ein Sekundärvertex rekonstruiert werden kann, standardmäßig als b -Quark-Jets angesehen.

Da aber auch c -Hadronen eine lange Lebensdauer haben und durch fehlrekonstruierte Spuren Sekundärvertizes in leichten Jets gebildet werden können, enthält ein wesentlicher Teil der Jets, in denen ein Sekundärvertex rekonstruiert wurde, keine b -Quarks. Im Fall der Suche nach Einzeltopquarkproduktion trifft das auf 48% der Ereignisse zu, die im selektierten CDF-Datensatz von 955 pb^{-1} erwartet werden. Ein wesentliches Ziel dieser Arbeit ist die Unterdrückung dieses Untergrundes, was durch die Entwicklung des ersten neuronalen Netzes zur Identifikation von b -Quark-Jets an einem Hadronbeschleuniger realisiert wurde. Das Konzept dieses neuronalen Netzes basiert auf der Erkennung von Sekundärvertizes, die vom Zerfall eines b -Hadrons stammen. Dies wurde dadurch erreicht, dass bei der Optimierung des Netzes neben der langen Lebensdauer von b -Hadronen auch ihre große Masse, ihre hohe Zerfallsmultiplizität und die Wahrscheinlichkeit für semileptonische Zerfälle berücksichtigt werden. Da sich dieses neuronale Netz als eines der wichtigsten Kriterien zur Unterdrückung von Untergrundereignissen erwiesen hat, findet es in allen Analysen zur Suche nach Einzeltopquarkproduktion bei CDF Anwendung. Eine weitere interessante Möglichkeit der Verwendung dieses Identifikationsnetzwerks liegt in der Bestimmung der Flavorzusammensetzung einer gegebenen Menge von Ereignissen.

Im Augenblick wird das neuronale Netz bei der Abschätzung der erwarteten Ereignisse als Gegenprobe zur Standardmethode verwendet. Diese Art der Anwendung wird in Zukunft durch die mit wachsender Statistik abnehmenden Unsicherheiten noch mehr an Bedeutung gewinnen. Darüber hinaus ist erwähnenswert, dass das b -Quark-Identifikationsnetz ein universelles Werkzeug ist. Obwohl es im Rahmen der Suche nach elektroschwacher Produktion einzelner Topquarks entwickelt wurde, hängt die Verteilung des Ausgabewerts nur vom Flavorinhalt eines hochenergetischen Jets ab und nicht vom physikalischen Produktionsprozess. Daher kann das Netz prinzipiell in allen Topquark-Analysen, Suchen nach dem vorhergesagten Higgsboson und weiteren Analysen von Ereignissen mit hohen Transversalimpulsen eingesetzt werden, wenn dort die Identifikation von b -Quark-Jets vonnöten ist.

Auch in der Suche nach der Erzeugung einzelner Topquarks werden neuronale Netze verwendet, um die beiden Produktionskanäle von den verschiedenartigen Untergründen zu unterscheiden. Hierbei werden zwei unterschiedliche Ansätze verfolgt. Zum Einen wird eine kombinierte Suche nach t - und s -Kanal durchgeführt, in welcher ein neuronales Netz darauf trainiert wird, beide Kanäle als Signal zu erkennen. Hierbei wird die Annahme gemacht, dass das Verhältnis der Wirkungsquerschnitte so ist, wie es vom Standardmodell vorhergesagt wird. Zum Anderen werden in der separaten Suche die Wirkungsquerschnitte der beiden Kanäle unabhängig voneinander simultan bestimmt, indem zwei unterschiedliche Netze trainiert werden: eins zur Identifikation von t -Kanal-Ereignissen und eins zur Identifikation von s -Kanal-Ereignissen, wobei der jeweils andere Kanal als Untergrund angesehen wird. Aus den Ausgabewerten dieser drei neuronalen Netze werden Verteilungen erstellt, die unter Berücksichtigung der erwarteten Anzahl von Untergrundereignissen und der systematischen Unsicherheiten durch eine Maximum-Likelihood-Methode an die Verteilung gemessener Daten angepasst werden.

Die erwarteten Sensitivitäten der beiden unterschiedlichen Ansätze werden mit Hilfe von Hypothesentests bestimmt. In der kombinierten Suche ergibt sich eine erwartete Signifikanz von 2.6σ , im Fall der separaten Suche 2.7σ , was die höchste *A-priori*-Sensitivität aller Suchen nach Produktion einzelner Topquarks am Tevatron ist. Die erwartete Obergrenze auf den Wirkungsquerschnitt bei 95% Konfidenz ist in der kombinierten Suche 5.7 pb, während für die separate Suche Obergrenzen von 3.8 pb für den t -Kanal und 2.9 pb für den s -Kanal erwartet werden.

Durch Maximierung der Wahrscheinlichkeitsfunktion bei der Anpassung der Musterverteilungen an die gemessenen Daten ergibt sich in der kombinierten Suche eine Rate von null Einzeltopquarkereignissen. Die daraus resultierende Obergrenze auf den Wirkungsquerschnitt ist 2.6 pb bei 95% Konfidenz. In der separaten Suche ergeben sich Wirkungsquerschnitte von $\sigma_t = 0.2^{+1.1}_{-0.2}$ (stat. + syst.) pb für den t -Kanal und $\sigma_s = 0.7^{+1.5}_{-0.7}$ (stat. + syst.) pb für den s -Kanal, woraus Obergrenzen von 2.6 pb beziehungsweise 3.7 pb folgen. Zusammenfassend lässt sich sagen, dass es sowohl in der kombinierten als auch in der separaten Suche nach Einzeltopquarkproduktion nicht möglich ist, ein Signal nachzuweisen.

Parallel zu dieser Arbeit wurden zwei weitere Analysen zur kombinierten Suche nach elektroschwacher Produktion einzelner Topquarks bei CDF durchgeführt. Hiervon

war es in einer Analyse möglich, einen Wirkungsquerschnitt von $2.7^{+1.5}_{-1.3}$ pb bei einer Signifikanz von 2.3σ zu messen [130]. In der zweiten Analyse wurde ein Ergebnis ähnlich dem der in dieser Arbeit beschriebenen kombinierten Suche gefunden, woraus eine Obergrenze von 2.7 pb bei 95% Konfidenz resultiert [131]. Detaillierte Studien, die unter Einbeziehung der Korrelationen zwischen den unterschiedlichen Analysen durchgeführt wurden, haben ergeben, dass die Ergebnisse zu etwa 1% konsistent sind, was durch statistische Fluktuationen erklärt werden kann. Zusätzlich wurden auch in der DØ –Kollaboration Suchen nach Einzeltopquarkproduktion durchgeführt, wobei in jeder dieser Analysen eine Evidenz für ein kombiniertes Signal von t - und s -Kanal gefunden wurde. Jedoch entsprechen die Messergebnisse fast dem Zweifachen des vorhergesagten Wirkungsquerschnitts und liefern einen gemeinsamen Wert von 4.7 ± 1.3 pb mit einer Signifikanz von 3.6σ [133], was allerdings mit 95% Konfidenz von der kombinierten Suche, die hier vorgestellt wird, ausgeschlossen werden kann. Sowohl für die in der DØ –Kollaboration durchgeführten Analyse als auch für die hier beschriebene kombinierte Suche ist die Wahrscheinlichkeit, mit der man unter Annahme der Standardmodellvorhersage das jeweilige Ergebnis erhält, nur wenige Prozent und damit die Übereinstimmung der Resultate mit dem theoretisch erwarteten Wert in beiden Fällen etwa gleich groß. Die Interpretation der unterschiedlichen Ergebnisse ist, dass die Abweichungen wahrscheinlich durch statistische Fluktuationen zustande kommen. Die Suche nach elektroschwacher Produktion einzelner Topquarks ist daher immer noch nicht abgeschlossen, aber möglicherweise wird diese spannende Fragestellung in der nächsten Analyserunde gelöst werden können, wenn ein größerer Datensatz zur Verfügung steht.

SEARCH FOR ELECTROWEAK SINGLE TOP–QUARK PRODUCTION WITH THE CDF II EXPERIMENT

SVENJA RICHTER

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

genehmigte

Dissertation

von

Dipl. Phys. Svenja Richter
aus Karlsruhe

Tag der mündlichen Prüfung: 02.11.2007

Referent: Prof. Dr. M. Feindt, Institut für Experimentelle Kernphysik

Korreferent: Priv.–Doz. Dr. W. Wagner, Institut für Experimentelle Kernphysik

Introduction

Understanding the world — This aim drives human kind since the beginning of conscious thinking. Especially the nature of matter has been of major interest. Nowadays, we have a complex image of the constitution of matter. Atoms consist of electrons and nucleons. But even nucleons are not elementary. Their basic constituents are called quarks [1]. Physicists developed a model describing the elementary components of matter as well as the forces between them: the standard model of elementary particle physics.

The substructure of matter is only visible in scattering experiments. In high energy physics, these experiments are done at particle accelerators. The world's highest-energetic collider, the Tevatron, is hosted by the Fermi National Accelerator Laboratory (FNAL), also called Fermilab, in the vicinity of Chicago. The proton–antiproton collisions with a center-of-mass energy of $\sqrt{s} = 1.96$ TeV are recorded by two multipurpose detectors, namely DØ and CDF II.

The heaviest of the standard-model quarks, the top quark, was discovered at the Tevatron by the DØ and CDF collaborations in 1995 [2, 3]. Still, Tevatron is the only facility able to produce top quarks. They are predominantly produced in top–antitop quark pairs via the strong interaction. The standard model also allows production of single top-quarks through the electroweak interaction. Single top-quark production is one of the major interests of the Tevatron experiments, since it offers several ways to probe the standard model and to provide indications of physics beyond the standard model. The electroweak production mode involving a Wtb vertex features the opportunity of a direct measurement of the $|V_{tb}|$ element of the quark-mixing matrix (named Cabbibo–Kobayashi–Maskawa or CKM matrix), since the production cross section is proportional to $|V_{tb}|^2$. Indirect measurements of $|V_{tb}|$ have to assume that the number of quark generations is three as implemented within the standard model. With this assumption, the unitarity of the CKM matrix leads to $|V_{tb}| \approx 1$. If there are more than three quark families, $|V_{tb}|$ could be significantly smaller than one. Moreover, single top-quark production is an important background to the search of a potential Higgs boson in the mass range of 90 GeV/ c^2 to 130 GeV/ c^2 in the WH channel.

At the beginning of this analysis, no evidence for single top-quark production has yet been observed. In Tevatron Run I, several limits on the single top-quark production cross section were set by DØ [4, 5] and CDF [6, 7] which were outperformed in Run II [8, 9]. For the current round of analyses using a set of collision data

corresponding to an integrated luminosity of $\approx 1 \text{ fb}^{-1}$, a sensitivity between 2σ and 3σ is expected.

At the Tevatron, two processes contribute significantly to the production of single top-quarks: (1) production via t -channel, also called W -gluon fusion, and (2) production via s -channel, known as well as W^* process. To distinguish the two single top-quark production channels from the various backgrounds, we use neural networks to combine many quantities into one discriminating variable. The output distributions of simulated events are used to create templates which are fitted to the output distribution of observed events to determine the cross section. We pursue two different approaches: The first is a combined search of the production modes via t - and s -channel with the assumption that the ratio of the cross sections is as predicted by the standard model. The second is a separate search in which the cross sections of t - and s -channel production are determined separately and simultaneously.

One important improvement to previous searches for single top-quark production, which was developed in the context of this thesis, is the extension of the standard method of b -quark jet identification. The identification of b -quark jets is crucial for top-quark physics, since the top quark decays with a branching fraction of almost 100% into a W boson and a b quark. The standard identification method of reconstructing a secondary vertex mainly utilizes the long lifetime of b hadrons. To further reject secondary vertices due to decaying c hadrons or false reconstruction, the identification of b jets is extended by exploiting many characteristics of b hadrons using a neural network.

Contents

Introduction	i
1 Top Quarks within the Standard Model	1
1.1 The Standard Model	1
1.2 Top Quarks	5
1.2.1 Production of Top–Quark Pairs via Strong Interaction	6
1.2.2 Production of Single Top–Quarks via Electroweak Interaction	6
1.3 Is $V_{tb} \simeq 1$?	10
2 The Experimental Design	13
2.1 Accelerating and Colliding the Particles	14
2.1.1 Luminosity	16
2.2 The Collider Detector at Fermilab	18
2.2.1 Tracking System	19
2.2.2 Time-of-Flight	20
2.2.3 Calorimetry	21
2.2.4 Muon System	21
2.3 The CDF Trigger System	22
3 Neural Networks	25
3.1 Functionality of NeuroBayes [®]	26
3.1.1 The Training Process	27
3.1.2 Preprocessing of the Variables	27
3.1.3 Automatic Variable Selection	28
3.1.4 Training Result	28

4	Event Simulation and Reconstruction	31
4.1	Monte Carlo Event Generation	31
4.1.1	Showering and Hadronization Event Generators	32
4.1.2	Tree-Level Matrix Element Generators	33
4.2	Detector Simulation	34
4.3	Reconstruction Prerequisites	34
4.3.1	Track Reconstruction	34
4.3.2	Primary Vertex and Beam Line Reconstruction	37
4.3.3	Jet Reconstruction	38
4.3.4	Secondary Vertex Reconstruction	43
5	Event Selection and Background Estimate	49
5.1	Measured Data Samples	49
5.2	Requirements for Candidate Events	50
5.3	Data Modeling	53
5.3.1	Signal MC	55
5.3.2	Modeling of backgrounds	57
5.4	Event Yield and Background Estimate	58
5.4.1	Expected Event Yield of Single-Top-Quark Events	58
5.4.2	Background Estimate	59
6	A Neural-Network b-Tagger	63
6.1	Discriminating Variables	64
6.1.1	Input Variables	68
6.2	Result of the Training	68
6.3	Comparison of Observed and Simulated Events	71
6.3.1	Used Samples	71
6.3.2	Modeling of Input Variables	73
6.3.3	Output Distributions of b -Quark Jets	73
6.3.4	Output Distributions of Mistagged Jets	81
6.4	Application in Analyses	89
6.4.1	Implementation in the Search for Single Top Quarks	89
6.4.2	Estimation of Flavor Fractions	91

7	Search for Single Top–Quark Production	93
7.1	Neural–Network Input–Variables	93
7.1.1	Top–Quark Reconstruction	94
7.1.2	Kinematic Fitter	95
7.1.3	Utilized Variables	95
7.2	Training of the Neural Networks	106
7.2.1	Training Samples	106
7.2.2	Relevance of Input Variables	108
7.2.3	Output Distributions and Templates	111
7.3	Systematic Uncertainties	116
7.4	Likelihood Function	119
7.5	Ensemble Tests and Expected Sensitivity	121
7.6	Application to Observed Events	126
7.6.1	Comparison to Expectation	126
7.6.2	Fit Results	127
	Summary and Outlook	133
A	Baseline Cuts for Lepton Identification	135
A.1	Electron Identification	135
A.1.1	Central Electrons	135
A.1.2	Forward Electrons	136
A.2	Muon Identification	137
A.3	Selection of Heavy–Flavor Enhanced Events	140
B	Used MC Samples	143
C	More Discriminating b–Tagger Variables	147
C.1	Discriminating Variables	147
C.2	Discriminating Variables in Data and MC	155
D	Distributions Concerning the Single–Top Networks	165
D.1	Data–MC Comparison of Input Variables	165
D.2	Correlations between Input Variables	170
D.3	Systematic Uncertainties	172

Chapter 1

Top Quarks within the Standard Model

1.1 The Standard Model of Elementary Particle Physics

The standard model of elementary particle physics is a quantum field theory that describes the fundamental particles making up all matter as well as the fundamental forces between them, namely strong, weak, and electromagnetic force. Gravitation is not included in this model but described by general relativity. The elementary particles comprise both fermions, particles with spin $s = \frac{1}{2}$, and bosons, particles with spin $s = 1$.

In the standard model, all matter is composed of fermions. They obey the Pauli exclusion principle which states that fermions cannot share the same quantum state at the same time. To describe the state of elementary particles, quantum numbers are used, e.g. the electromagnetic charge or the spin projection. The standard model incorporates twelve different types of fermions arranged in three families, shown in table 1.1. Hereof only particles of the first generation form all ordinary matter we encounter in our natural environment; the higher-generation fermions can solely be observed in high-energy interactions, since they decay into first-generation particles.

There are two different kinds of fermions: quarks and leptons. The standard model contains six quarks which cannot be observed as free particles but only in bound states named hadrons. Those bound states can either consist of three quarks (baryons) or quark-antiquark pairs (mesons). Due to the Pauli exclusion principle, it is necessary that quarks carry an additional quantum number [10, 11], so-called color. Leptons, i.e. electron, muon, tau, and the corresponding neutrinos, do not carry color charge. Each elementary particle has a corresponding antiparticle carrying opposite quantum numbers.

In the standard model, neutrinos are originally assumed to be massless. Due to recent observations [12, 13, 14, 15], it is necessary to extend the standard model

name	category	symbol	el. charge [e]	mass [MeV/c^2]
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
muon	lepton	μ	-1	105.66
muon neutrino	lepton	ν_μ	0	≤ 0.190
charm quark	quark	c	$+\frac{2}{3}$	$(1.25 \pm 0.09) \cdot 10^3$
strange quark	quark	s	$-\frac{1}{3}$	95 ± 25
tau	lepton	τ	-1	1777.0
tau neutrino	lepton	ν_τ	0	≤ 18.2
top quark	quark	t	$+\frac{2}{3}$	$(174.2 \pm 3.3) \cdot 10^3$
bottom quark	quark	b	$-\frac{1}{3}$	$(4.20 \pm 0.09) \cdot 10^3$

Table 1.1: The elementary fermionic particles (spin $s = \frac{1}{2}$) with their electric charge in units of the electron charge and their mass [16]. The top-quark mass is based on published results [17].

in such a way that neutrinos have non-zero masses. The neutrino masses cited in table 1.1 are limits obtained from direct measurements. Cosmological constraints force the sum of all neutrino masses to be in the order of a few eV [16], leading to stricter limits on the masses of ν_μ and ν_τ .

The three standard model forces, electromagnetic, strong, and weak force, are mathematically described by gauge theories. Gauge theories are based on the idea that symmetry transformations can be performed locally as well as globally. In general, a physical system is described by Lagrangians. In a gauge theory, those Lagrangians are invariant under gauge transformations, i.e. local change of variables. To guarantee the invariance of the Lagrangian, additional fields, the gauge fields, must be introduced to compensate for the local change of variables. In quantum field theory, the excitations of the gauge fields represent particles transmitting the forces, namely the gauge bosons shown in table 1.2. According to the Noether theorem [18], each symmetry meets a conserved current, eventually corresponding to charge conservation. This leads to the fact that only particles carrying the charge of a certain force can interact via this specific force.

The gauge boson transmitting the electromagnetic force is the photon which couples to the electric charge. Electromagnetic interactions are theoretically described by Quantum Electro Dynamics (QED), predicting the photon as a massless, chargeless boson to preserve the invariance under redefinition of the electrostatic potential. The fact that the photon is massless leads to an infinite range of the electromagnetic force. The experimental limits on the photon's mass and electric charge are displayed in table 1.2.

Strong interactions via gluon exchange are described by Quantum Chromo Dynamics (QCD) [19, 20, 21]. The massless gluon does not only couple to color charge but carries color itself, leading to gluon self-interactions. This explains that the strength

of strong interactions decreases at short distances, a feature termed asymptotic freedom, and that it increases at large distances, the so-called confinement, which constrains quarks to immediate creation of colorless bound hadrons. Since gluons must carry some charge and some anticharge, all possible color combinations would lead to nine gluons. The color singlet is not physically significant, reducing the number of transmitting gluons to eight.

To understand the weak interaction, it is crucial to introduce the idea of chirality. Something is chiral if it cannot be superimposed on its mirror image. The term chirality must not be confused with the notion helicity. Helicity is the projection of a particle's spin to its direction of motion, i.e. an observable which is only Lorentz invariant for massless particles. A massive particle has both left- and right-handed chiral components or in other words, its wave function ψ consists of a left-handed chiral part ψ_L and right-handed chiral part ψ_R , $\psi = \psi_L + \psi_R$. This means for example that a massive particle with positive helicity has a left-handed component. For a massless (ultrarelativistic) particle, chirality equals (approximates) helicity, since there is no frame of reference in which its helicity would change.

Weak interactions transmit via massive bosons, namely the electrically neutral Z boson and the electrically charged W^\pm bosons, inducing limited range. The W boson only couples to the left-handed chiral component of the fermion wave function and respectively to the right-handed component in case of antifermions. The Z boson couples to both handednesses, even though with different strengths. The weak force is the only standard model force to affect neutrinos. In weak interactions, flavor change is possible by W boson exchange. Here, the quantum states of quarks differ from the mass eigenstates. The transformation from one base into the other is described by the Cabibbo–Kobayashi–Maskawa matrix (CKM matrix) (1.1) [22, 23]. By convention, the matrix is expressed as a 3×3 unitary matrix V which operates on the quark mass eigenstates (d , s , and b), resulting in the weak eigenstates (d' , s' , and b').

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

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

Table 1.2: The gauge bosons mediating the forces of the standard model. The electric charge and the mass of the bosons are shown as given by reference [16]. The limit on the photon mass is based on the magnetohydrodynamics of the solar wind. The limit on the photon's electric charge is obtained from astronomical time-of-flight methods.

The coupling strength of two quarks q_1 and q_2 to a W boson is proportional to the matrix element $V_{q_1 q_2}$ which has to be determined experimentally. The current 90% confidence limits on the magnitude of the elements, assuming unitarity of the CKM matrix, are [16]:

$$\begin{pmatrix} 0.97383^{+0.00024}_{-0.00023} & 0.2272 \pm 0.0010 & (3.96 \pm 0.09) \cdot 10^{-3} \\ 0.2271 \pm 0.0010 & 0.97296 \pm 0.00024 & (42.21^{+0.10}_{-0.80}) \cdot 10^{-3} \\ (8.14^{+0.32}_{-0.64}) \cdot 10^{-3} & (41.61^{+0.12}_{-0.78}) \cdot 10^{-3} & 0.999100^{+0.000034}_{-0.000004} \end{pmatrix} \quad (1.2)$$

The weak interaction was unified with the electromagnetic interaction in the electroweak theory by Glashow, Salam, and Weinberg [24, 25, 26]. Although these two forces appear very different at low energies, the theory models them as two different aspects of the same force. Above the unification energy, they would merge into a single electroweak force. The difference between electromagnetic and weak force appears due to the electroweak symmetry breaking. The most favored description of this symmetry breaking is the Higgs mechanism [27, 28, 29, 30] which ascribes masses to the originally massless gauge bosons, as well as quarks and leptons, by coupling to a scalar quantum field. In this symmetry breaking, the photon remains massless, while the W and Z bosons become massive. The quantum excitation of the Higgs field is the Higgs boson, the only standard model particle not yet observed.

The interaction of two charged elementary particles via the exchange of a gauge boson can be visualized using Feynman diagrams. Exemplarily, the electron–electron scattering via exchange of a virtual photon is shown in figure 1.1. Each part of a Feynman diagram can be translated into a mathematical expression to calculate the transition amplitude M . Propagation of particles and antiparticles in space–time is represented by lines; their coupling is visualized by vertices. The cross section of a process, a measure of the probability of the interaction occurring, can be derived from the transition amplitude by integrating over all initial and final states, the phase space.

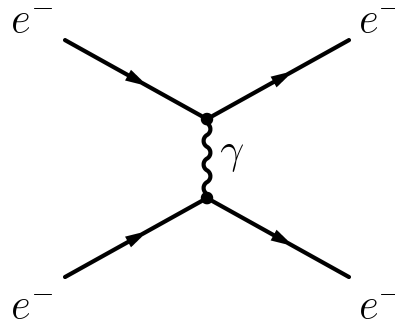


Figure 1.1: The incoming electrons are represented by the lines on the left hand side, the outgoing electrons by the lines at the right hand side. Their coupling to the photon is visualized at the vertices represented by the dots. The transmission of the force via a photon is described by its propagator.

Nevertheless, at the Tevatron, composed particles (protons and antiprotons) are collided, making the calculation of a measurable cross section more complicated: it has to be taken into account that the momentum of the proton (or antiproton) is shared among all the elementary constituent particles, also called partons. The proton consists of two u and one d quark, the antiproton of the respective antiparticles. Those constituents, called valence quarks, are bound by virtual gluons which can split into quark–antiquark pairs, the so-called sea quarks. Since the Feynman rules apply to elementary particles, the fraction of momentum carried by those partons is crucial to be known, being described by the Parton Distribution Function (PDF) $f_{i,p}(x_i, \mu^2)$. This function depends on the scale μ describing the typical energy scale of the considered interaction. For top–quark physics, this scale is often set to the top–quark mass $\mu = m_{\text{top}}$. Figure 1.2 shows the CTEQ5L PDF [31] for $\mu^2 = (175 \text{ GeV})^2$. These PDFs have to be folded with the partonic cross sections to calculate the measurable cross section in $p\bar{p}$ collisions.

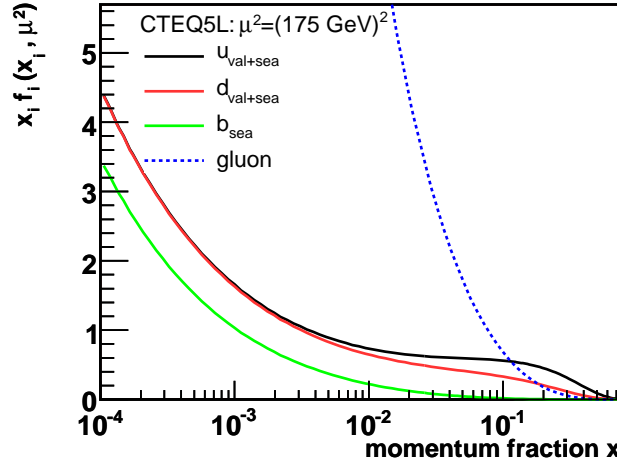


Figure 1.2: The CTEQ5L parton distribution function at $\mu^2 = (175 \text{ GeV})^2$ [31].

1.2 The Top Quark

The top quark, the heaviest of the standard model quarks, was discovered in 1995 at the Tevatron [2, 3]. Like every quark, it participates in both strong and electroweak interactions, leading to a variety of production modes described in the next sections. The main difference between the top quark and the rest of the standard model quarks lies in its large mass, nearly as heavy as a gold nucleus. The current value of the top–quark mass based on published results is $m_{\text{top}} = 174.2 \pm 3.3 \text{ GeV}/c^2$ [16, 17]; including preliminary results, the average mass is $m_{\text{top}} = 170.9 \pm 1.8 \text{ GeV}/c^2$ [32]. Although the top quark decays via the weak force, its predicted lifetime of $\tau \approx 0.5 \cdot 10^{-24} \text{ s}$ is shorter than the hadronization scale. Hence, the top quark is expected to decay before top–flavored hadrons can form. This provides the unique opportunity to study

a bare quark which passes all its properties to its decay products, in almost 100% of the cases a W boson and a b quark due the value of $|V_{tb}| \approx 1$ (see limits on CKM matrix elements (1.2)). A detailed review on top-quark physics in hadron collisions can be found in reference [33].

1.2.1 Production of Top-Quark Pairs via Strong Interaction

In hadron collisions, top quarks are predominantly produced in pairs via the strong interaction. In proton-antiproton collisions at the Tevatron, two production mechanisms are dominant: top-quark pair production via quark-antiquark annihilation (figure 1.3(a)) and via gluon fusion (figure 1.3(b)).

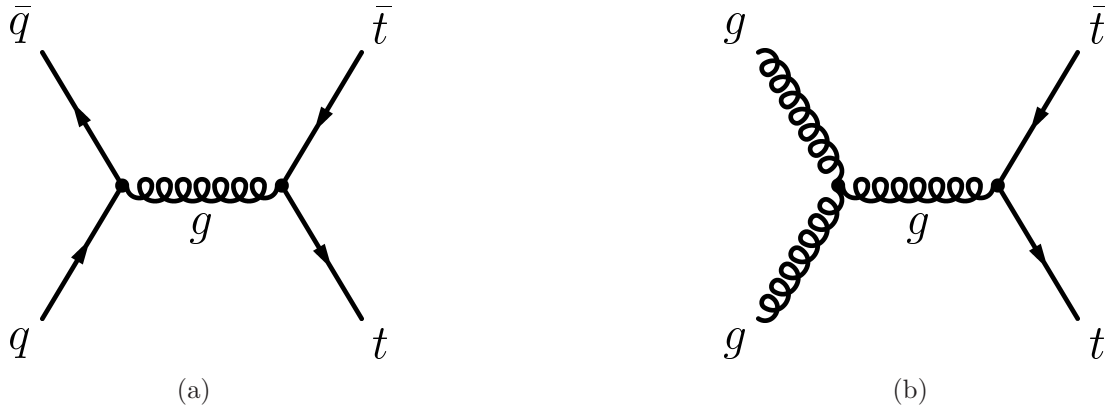


Figure 1.3: Leading-order Feynman diagrams of top-quark pair production: (a) quark-antiquark annihilation and (b) gluon fusion.

For a top-quark mass of $m_{\text{top}} = 175 \text{ GeV}/c^2$, the production cross section at the Tevatron with $\sqrt{s} = 1.96 \text{ TeV}$ is predicted to be $\sigma_{t\bar{t}} = 6.70^{+0.71}_{-0.88} \text{ pb}$ [34]. In next-to-leading order calculations, the annihilation contributes about 85%, the gluon fusion about 15% to the total cross section.

1.2.2 Production of Single Top-Quarks via Electroweak Interaction

Besides the production of top-quark pairs involving strong interactions, the production of single top-quarks via electroweak interaction is predicted by the standard model. Since electroweak top-quark production proceeds via a Wtb vertex, it provides the opportunity of a direct measurement of the CKM matrix element $|V_{tb}|$.

The three electroweak production modes in $p\bar{p}$ collisions are distinguished by the virtuality Q^2 of the participating W boson, with $Q^2 = -q^2$ where q is the four-momentum of the W boson. The two dominating production modes are named according to the relevant Mandelstam variables involved in the transition matrix

elements. In t -channel production, the produced virtual W boson is space-like ($q^2 = t < 0$), while it is time-like ($q^2 = s \geq (m_{\text{top}} + m_b)^2$) in s -channel production. The third production mode does not fit in this name scheme, since an on-shell (or almost on-shell) W boson is produced in conjunction to the top quark.

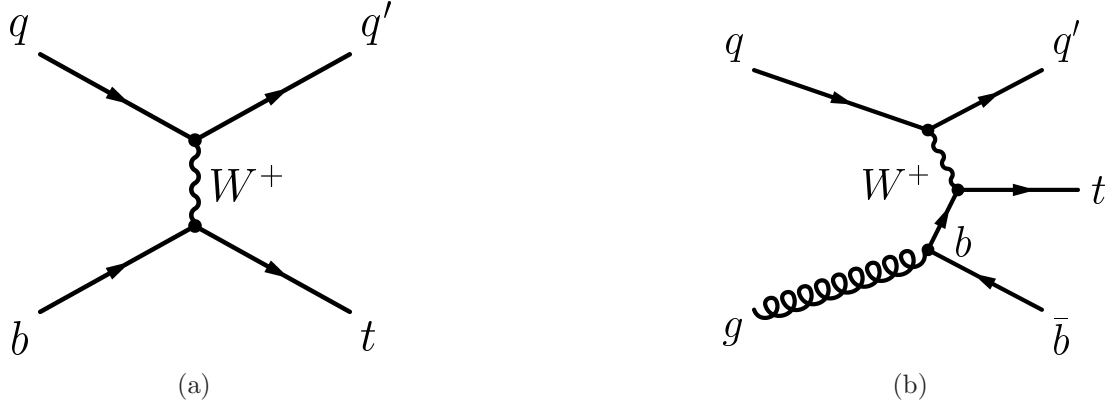


Figure 1.4: Single top-quark production in the t -channel at (a) leading order and (b) including initial state gluon splitting. For antitop-quark production, the charge conjugate processes apply.

In t -channel top-quark production, also known as W -gluon fusion, a virtual space-like W boson strikes a b sea-quark inside the proton or antiproton as shown by the Feynman graphs in figure 1.4. The leading order $2 \rightarrow 2$ process is displayed in figure 1.4(a). The most important next-to-leading order correction, illustrated in figure 1.4(b), is a $2 \rightarrow 3$ process which takes into account that the b quark originally stems from gluon splitting, leading to an additional \bar{b} quark in the event. Due to the large mass difference between top and b quark, the b quark is taken as massless in theoretical calculations, leading to a singularity if the \bar{b} -quark's direction of motion is collinear to the gluon. This calculative problem can be canceled by introducing a PDF for the b quark being part of the quark sea (as illustrated in figure 1.2), resulting in the $2 \rightarrow 2$ process in figure 1.4(a) as the leading order process. For non-zero b -quark masses, this collinear singularity is reflected in the tendency for the \bar{b} quark to be produced at high rapidity and low transverse momentum [35, 36]. The predicted cross section of W -gluon fusion at next-to-leading order at the Tevatron is $\sigma_t = 1.98^{+0.28}_{-0.22}$ pb, assuming $m_{\text{top}} = 175 \text{ GeV}/c^2$. The cross section is taken from reference [37]; the uncertainty was evaluated in reference [38] and includes uncertainties due to the factorization scale ($\pm 4\%$), the choice of PDF parameterization ($^{+11.3\%}_{-8.1\%}$), and the uncertainty in the top-quark mass ($^{+6.9\%}_{-7.5\%}$). The mass of the b quark and the error in α_s play an insignificant role in the overall uncertainty for measurable quantities.

In s -channel production, two quarks annihilate into a highly virtual W boson which produces a top (antitop) and a \bar{b} (b) quark. The leading order Feynman graph for top-quark production is displayed in figure 1.5(a). An example of next-to-leading order correction including initial state gluon splitting is shown in figure 1.5(b). This production process has the same initial and final state as the $2 \rightarrow 3$ W -gluon fusion production displayed in figure 1.4(b). In contrast to the t -channel process, where

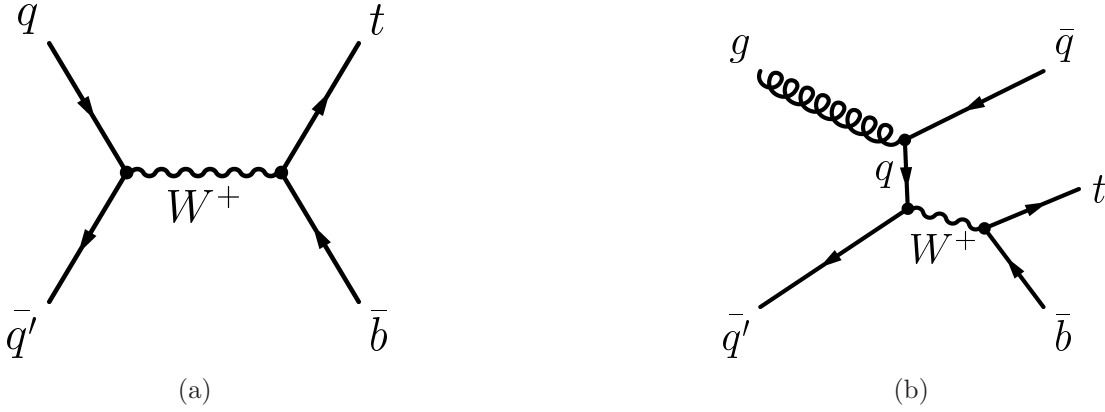


Figure 1.5: Single top-quark production in the s -channel at (a) leading order and (b) including initial state gluon splitting. For antitop-quark production, the charge conjugate processes apply.

the $t\bar{b}$ pair forms a color-octet state since it originates from a gluon, the top and the \bar{b} quark form a color-singlet state in s -channel since both quarks stem from a W boson. Therefore, both t - and s -channel are separately gauge invariant and cannot interfere. At the Tevatron, the predicted cross section of s -channel production at next-to-leading order is $\sigma_s = 0.88^{+0.12}_{-0.11}$ pb for $m_{\text{top}} = 175 \text{ GeV}/c^2$ [37, 38]. Here, the uncertainty includes $\pm 2\%$ due to the factorization scale, $^{+4.7\%}_{-3.9\%}$ due to the PDF parameterization and $^{+10.0\%}_{-11.7\%}$ due to the uncertainty in the top-quark mass.

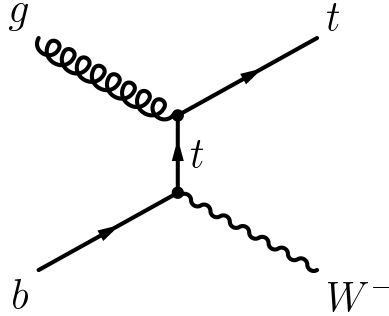


Figure 1.6: Associated single top-quark production at leading order. For antitop-quark production, the charge conjugate processes apply.

The leading order Feynman graph of associated single top-quark production, also named Wt production, is shown in figure 1.6. The predicted cross section of this production mode at the Tevatron for $m_{\text{top}} = 175 \text{ GeV}/c^2$ is $\sigma_a = 0.094^{+0.015}_{-0.012}$ pb [39], assuming $\sqrt{s} = 2.0 \text{ TeV}$. Since this cross section is negligible, the further discussion is restrained on t - and s -channel single top-quark production.

At the Large Hadron Collider (LHC) currently under construction at CERN in Geneva (Switzerland), where proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 14.0 \text{ TeV}$ will take place, the situation will be different. The cross sections of the single top-quark production mechanisms via t - and s -channel will differ for top-quark and antitop-quark production. The predicted cross section for the production

via t -channel is (156 ± 8) pb for top quarks and (91 ± 5) pb for antitop quarks. For the s -channel production mode, the cross section will be (6.6 ± 0.6) pb for top quarks and (4.1 ± 0.4) pb for antitop quarks, both comparable to the uncertainties on the t -channel production cross sections [37, 38]. The associated production will gain importance at the LHC. The cross section for W^-t and $W^+\bar{t}$ will be equal, leading to a total cross section of 62^{+17}_{-4} pb [40].

The single top-quark production t - and s -channel processes illustrated in figures 1.4 and 1.5 are dominated by contributions from u and d quarks coupling to the W boson. Contributions from s or c sea-quarks in the initial state are small, an effect of only $\approx 6\%$ for t -channel and $\approx 2\%$ for s -channel production [41]. Besides the presented Feynman graphs involving a Wtb vertex, there are several other production channels via Wtd or Wts vertices which are strongly suppressed due to tiny CKM matrix elements.

In single top-quark production via t - and s -channel, the top quark is almost exclusively created with negative helicity due to the left-handed structure of the Wtb coupling. Since top quarks are not produced as ultrarelativistic particles, the chirality eigenstates are not identical to the helicity eigenstates, leading to a small spin asymmetry $A_{\uparrow\downarrow} = \frac{N_{\uparrow} - N_{\downarrow}}{N_{\uparrow} + N_{\downarrow}}$. The spin polarization is passed to the decay products, since the top quark decays as a quasi free particle; i.e. in the top-quark rest frame, the angular distributions of the decay products are sensitive to the spin polarization of the top quark. According to reference [42], the angle $\theta_{\ell\vec{s}}$ between the charged lepton ℓ from the W decay and the top-quark spin axis \vec{s} has maximum analyzing power. The angular distribution is given by

$$\frac{dN}{d(\cos\theta_{\ell\vec{s}})} = \frac{1}{2} (1 + A_{\uparrow\downarrow} \cos\theta_{\ell\vec{s}}). \quad (1.3)$$

In reference [36], it is shown that the direction of the d -type quark provides the most effective spin axis for all single-top quark production mechanisms. The direction of the d -type quark is experimentally not known and has to be chosen by taking into account the Feynman graphs shown in figures 1.4 and 1.5.

Looking at the leading order s -channel production mechanism in figure 1.5(a), it is obvious that the d -type quark is in the initial state, i.e. stemming from either the proton or the antiproton. Thus, the beam axis provides the ideal basis for analyzing the spin polarization in s -channel events. In antitop-quark production, the initial state contains a d quark, in most cases provided by the proton. In top-quark production, respectively, the initial \bar{d} quark stems mostly from the antiproton; that is why, in the top-quark rest frame, the top-quark spin is aligned with the antiproton beam 98% of the time.

In t -channel production, the d -type quark can either be in the final state (77% of the cases [36, 43]) or in the initial state (23% of the cases). However, in both cases, the observed light-quark jet is a good basis to analyze the top-quark spin, since the final state u -type quark tends to move in the same direction as the initial state d -type quark. In the basis of the light-quark jet, the polarization of the top

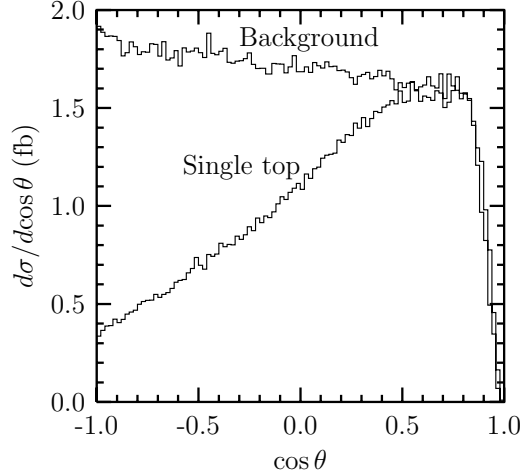


Figure 1.7: Angular distribution of the charged lepton in t -channel events at the Tevatron ($\sqrt{s} = 2$ TeV) with respect to the light-quark jet. For comparison, the distribution of the main background processes (W +jets and $t\bar{t}$) is shown as well. The plot is taken from reference [43].

quark is 95%. As illustration, figure 1.7 shows a theoretical prediction of the angular distribution (1.3) between the charged lepton and the light-quark jet in t -channel events [43]. The suppression at $\cos\theta \approx 1$ is due to a cut on the distance between lepton ℓ and jet j , $\Delta R_{\ell j} = \sqrt{(\eta_\ell - \eta_j)^2 + (\phi_\ell - \phi_j)^2} > 0.7$, which resembles the effect of jet clustering and the isolation requirement of leptons employed in experimental analyses.

1.3 Is $V_{tb} \simeq 1$?

The value of the CKM matrix element V_{tb} , related to the top-bottom charged current, is often considered to be known to a very satisfactory precision (see table 1.2 for the current boundaries). However, this range is determined using a full set of tree-level processes and relies on the unitarity of the 3×3 CKM matrix. The unitary assumption is mainly supported by three experimental facts, namely the fact that the first two rows are consistent with the unitary condition, the measurement of the B_s -mixing frequency constrains the ratio $|V_{td}/V_{ts}|$, and the measurement of $R \equiv |V_{tb}|^2/(|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2) \approx 1$ implies a strong hierarchy between V_{tb} and the other two matrix elements.

On the other hand, contrary to what has sometimes been argued, none of these experimental facts are directly constraining V_{tb} . In fact, even its determination from R comes simply from taking the square root of R , assuming the unitarity of the CKM matrix. Since no measurement of the single-top-quark production cross section yet exists, it is theoretically still possible that $V_{tb} \neq 1$.

In reference [44], two minimal extensions of the standard model are discussed allowing a value for V_{tb} considerably different from one. In the first model, the intro-

duction of a new vector-like top singlet leads to a global rescaling of V_{td} , V_{ts} , and V_{tb} leaving R unchanged. In the second model, a complete new fourth generation is added and the R measurement is used as a direct constraint.

A vector-like t' quark: The introduction of one vector-like t' quark with $Q = 2/3$ is the minimal way to implement a rescaling of all V_{ti} entries:

$$V_{ti} = V_{ti}^{(0)} \cos \theta \quad (1.4)$$

with $V_{ti}^{(0)}$ being the standard-model value. If this hypothetical iso-singlet quark has a mass around the electroweak scale, it naturally mixes with its nearest neighbor, i.e. the standard-model top quark. Since such an enlargement of the 3×3 CKM matrix does not spoil the unitarity of the first two rows of the matrix, this extension of the standard model introduces only two new parameters, namely the t - t' mixing-angle θ and the t' -quark mass $m_{t'}$. Additionally, flavor changing neutral currents are implied whose non-observation constrains θ . A theoretical bound on V_{tb} is found to be $|V_{tb}| \simeq |\cos \theta| > 0.71$, while the most stringent experimental constraint is $|\cos \theta| \geq 0.91$. A limit on the mass of the t' quark has been measured to be $m_{t'} \geq 256 \text{ GeV}/c^2$ [45].

Fourth Generation: Another possible extension of the CKM structure of the standard model is the addition of a fourth generation. In this case, the presence of a b' quark implies a unitary 4×4 matrix such that tree-level flavor changing neutral currents in hadronic Z -boson decays are forbidden. The new CKM matrix elements can be written as

$$V_{ui} = \cos \theta_w V_{ui}^{(0)} \quad (1.5)$$

$$V_{ci} = \cos \theta_v V_{ci}^{(0)} - \sin \theta_w V_{ui}^{(0)} \quad (1.6)$$

$$V_{ti} = \cos \theta_u V_{ti}^{(0)} - \sin \theta_u \sin \theta_v V_{ci}^{(0)} - \sin \theta_u \cos \theta_v \sin \theta_w V_{ui}^{(0)} \quad (1.7)$$

$$V_{t'i} = \sin \theta_u V_{ti}^{(0)} + \cos \theta_u \sin \theta_v V_{ci}^{(0)} + \cos \theta_u \cos \theta_v \sin \theta_w V_{ui}^{(0)}. \quad (1.8)$$

Theoretical considerations as well as experimental results constrain the three parameters θ_u , θ_v , and θ_w to $|\theta_w| \leq O(\lambda^2)$, $|\theta_v| \leq O(\lambda)$, and $|\cos \theta_u| \geq 0.93$ with the Wolfenstein parameter $\lambda \equiv \sin \Theta_c \simeq 0.22$. Theoretical limits on the values of V_{ti} are found to be $|V_{td}| \leq 0.03$, $|V_{ts}| \leq 0.2$, and $|V_{tb}| \geq 0.8$, respectively.

Chapter 2

The Experimental Design

The CDF II experiment is hosted by the Fermi National Accelerator Laboratory, also known as FNAL or Fermilab. Fermilab is located in about 70 km distance to Chicago in Batavia, Illinois (USA). An aerial view of the Fermilab accelerator complex is given in figure 2.1.



Figure 2.1: Aerial view of the Fermilab facility. The course of the Tevatron ring is indicated by the rear circular maintenance road.

Currently, the Tevatron is the accelerator with the highest center-of-mass energy \sqrt{s} in operation. The first proton-antiproton collisions at $\sqrt{s} = 1.80$ TeV took place in 1985. So-called Run I continued until 1996, leading to the discovery of the top quark in 1995. Starting 1996, the accelerator complex was upgraded to increase the initial luminosity and the center-of-mass energy. In 2001, Tevatron Run II started colliding particles at $\sqrt{s} = 1.96$ TeV and is supposed to continue until 2009.

Collisions at energies of this amount make high demands on the engineering of both accelerator system and detectors, being described in the following sections.

2.1 Accelerating and Colliding the Particles

To generate protons and antiprotons finally colliding with a center-of-mass energy $\sqrt{s} = 1.96$ TeV, a complex system of succeeding accelerators is realized at Fermilab. A schematic overview of the accelerator chain is given in figure 2.2.

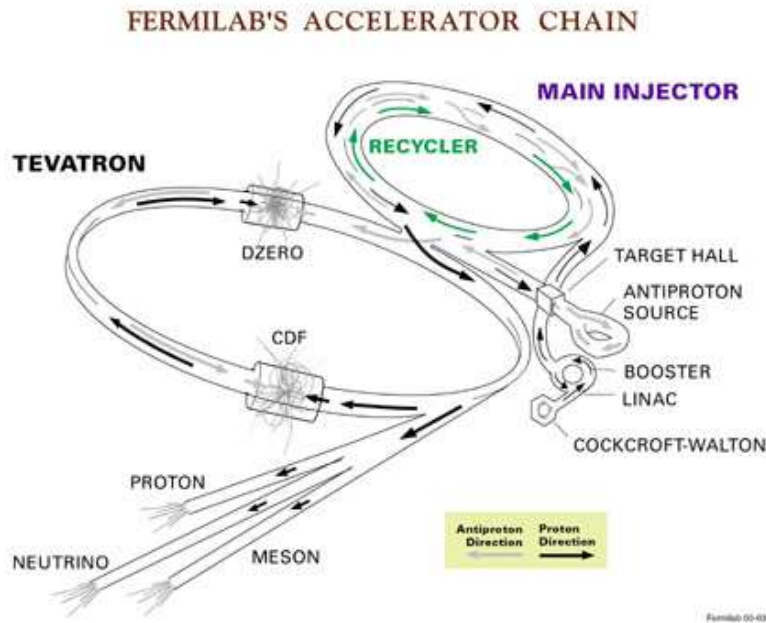


Figure 2.2: Schematic overview of the accelerator chain at Fermilab.

The first step in the accelerator system is the proton extraction using a source of negatively charged hydrogen ions. The ion source is housed in an electrically charged dome (-750 kV) which is part of the Cockcroft–Walton–type pre-accelerator. This source converts hydrogen gas to H^- ions which are accelerated to an energy of 750 keV while being guided by titanium electrodes from the charged dome to the grounded wall. Those pre-accelerated ions are passed to the LINear ACcelerator (LINAC) where they are accelerated to 400 MeV by radio frequency (RF) resonators. At the end of the LINAC, after about 130 m, the negative hydrogen ions enter the first synchrotron in the acceleration chain, the Booster, passing a carbon foil where the electrons are stripped off. In this circular accelerator with a circumference of about 475 m, the remaining protons are accelerated to 8 GeV before entering another synchrotron with seven times the circumference, the Main Injector.

The Main Injector has several operation modes, depending on the status of the acceleration process. For antiproton production, or “stacking” as it is called, one proton batch from the Booster is injected and accelerated to 120 GeV. The more

commonly used method is to inject two Booster batches and merge them together during acceleration, called “slip–stacking”. The proton beam is extracted from the Main Injector and sent to the Antiproton Source where it hits a nickel target. The high–energetic protons striking the target produce a spray of all sorts of secondary particles which are focused by a lithium lense. Using magnets acting as charge–mass spectrometer, antiprotons with energies of approximately 8 GeV are collected out of this spray and sent to the Debuncher. In this first synchrotron of the Antiproton Source with a mean radius of 90 m, the antiproton bunches are rotated by radio frequency manipulations. This is necessary because the protons striking the nickel target are bunched, leading to bunched antiprotons. Those antiprotons have a large spread in energy which is stochastically cooled [46, 47] in the Debuncher before the antiprotons are sent to the second synchrotron of the Antiproton Source, the Accumulator. In this synchrotron with a mean radius of 75 m, housed in the same tunnel as the Debuncher, the antiprotons are stored at an energy of 8 GeV and further cooled down, using both RF and stochastical cooling systems.

From the Antiproton Source, the antiprotons are sent to the Recycler which is located along the ceiling of the Main Injector. The proposed purpose of the Recycler was to recycle antiprotons from the Tevatron, cooling and storing them alongside those from the Antiproton Source. Those plans have been abandoned due to early problems in Run II. The Recycler now accepts transfers only from the Antiproton Source and cools the antiprotons further than the Accumulator is capable. In addition to stochastical cooling, electron cooling is necessary for higher intensities. Electron cooling works on the principle of momentum transfer between electrons and antiprotons of the same average velocity [48]. The antiproton beam is overlaid with a 4.3 MeV electron beam traveling about 20 m along the same path in the Recycler [49]. Coulomb scattering leads to energy transfer from the antiprotons to the co–streaming electrons until thermal equilibrium is attained. After cooling the so–called “stash”, the antiprotons are mined into nine parcels, each split into four bunches.

To prepare the final injection to the Tevatron main ring, i.e during the so–called “shot setup”, seven bunches of 8 GeV protons are transferred from the Booster to the Main Injector and accelerated to 150 GeV. After being coalesced to one bunch, the protons are shot into the Tevatron ring. This procedure is repeated 36 times. After transferring all proton bunches to the Tevatron ring, four antiproton bunches are extracted from the Recycler and accelerated to 150 GeV in the Main Injector before they are shot into the Tevatron. This procedure is repeated nine times, leading to a 36×36 bunch structure of the Tevatron beam.

The final acceleration stage is the Tevatron, a circular synchrotron with a circumference of about 6 km. It is the world’s first superconducting synchrotron and the only cryogenically cooled accelerator at Fermilab. The superconducting coils are made of niobium/titanium alloy which needs to be kept at $T \approx 4$ K. The advantage of superconductive magnets is that they do not dissipate power, since the wires do not have any electrical resistance.

The Tevatron accepts both 150 GeV protons and antiprotons from the Main Injector

and accelerates them to 980 GeV. The beam is split into three trains, each containing 12 bunches with 396 ns separation. The bunch revolution time is approximately 21 μs . Empty spaces between the trains, the abort gaps of $\approx 2.6 \mu\text{s}$, allow kicker magnets to ramp up. Those kicker magnets are used to abort the beam by guiding it into graphite beam dumps.

The Tevatron beam is brought to collision at two intersection points: B0, the location of the CDF II experiment, and D0 where the DØ detector records the collisions. Before the experiments start data taking, the beam needs some fine tuning. At the intersection points, the beam is focused to smaller transverse size, called “low β squeeze”, before collisions are initiated. To obtain a beam profile of maximum quality, the beam halo is removed by moving stainless steel collimators close to the beam, so-called “scraping”.

One important point for colliding beams is to keep the abort gaps clean. During collisions, those are gradually filled with a direct current (DC) component of the beam, since particles diffuse out of bunches and are no longer captured in RF buckets. The DC component in the abort gaps is cleaned by shooting an low-energy electron beam, called “electron lens”, into the gaps and thus excite the DC beam particles to very large amplitudes until they are lost [50].

Once the beam is stable, the experiments ramp up the high voltage of the various detector components and record data until the store (the period between Tevatron beam initialization and abortion) ends. The typical duration of a store is several hours; the longest store lasted more than 60 hours.

2.1.1 Luminosity

The instantaneous luminosity \mathcal{L} is a measure of the ability of a collider to produce collisions:

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

where n is the number of bunches, f is the revolution frequency, N_p ($N_{\bar{p}}$) is the number of protons (antiprotons) per bunch, and σ_x and σ_y represent the average transverse width of the bunches. \mathcal{L} is measured in units of $\text{cm}^{-2}\text{s}^{-1}$, typifying a particle flux. Integrating the instantaneous luminosity over time gives the so-called integrated luminosity, $\mathcal{L}_{\text{int}} = \int \mathcal{L} dt$, which is a measure of the amount of collected data. To calculate the event rate of a certain physics process, the probability of the process occurring is essential. This probability, the cross section σ , is usually given in cm^2 . In particle physics, cross sections are preferably quoted in picobarn (pb), where $1 \text{ pb} = 10^{-36} \text{ cm}^2$. For a particular cross section, the number of events N in a given amount of data can be calculated by $N = \sigma \cdot \mathcal{L}_{\text{int}}$.

The maximum luminosity is reached at the beginning of a store because the number of protons and antiprotons diminishes during collisions, leading to an exponential

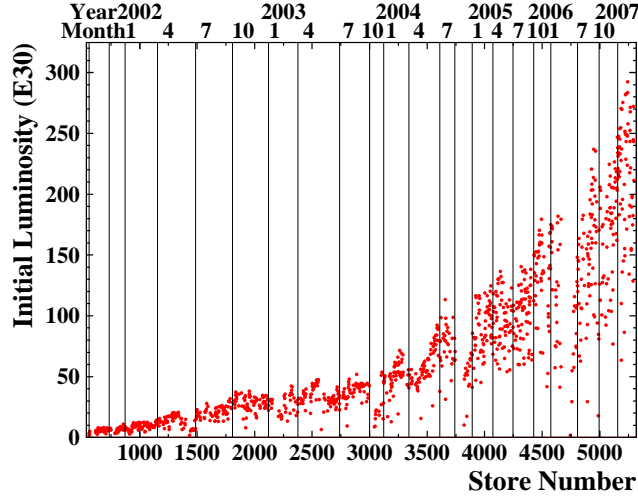


Figure 2.3: The initial luminosity in Run II as a function of time.

decrease in luminosity. Figure 2.3 shows the development of the initial luminosities over time. At the beginning of Run II, the initial luminosities were rather small but increased with improved understanding and handling of the accelerators. The design luminosity of $270 \cdot 10^{30} \text{ cm}^{-2}\text{s}^{-1}$ [51] was reached in Winter 2006/2007, the current record luminosity is $292.3 \cdot 10^{30} \text{ cm}^{-2}\text{s}^{-1}$.

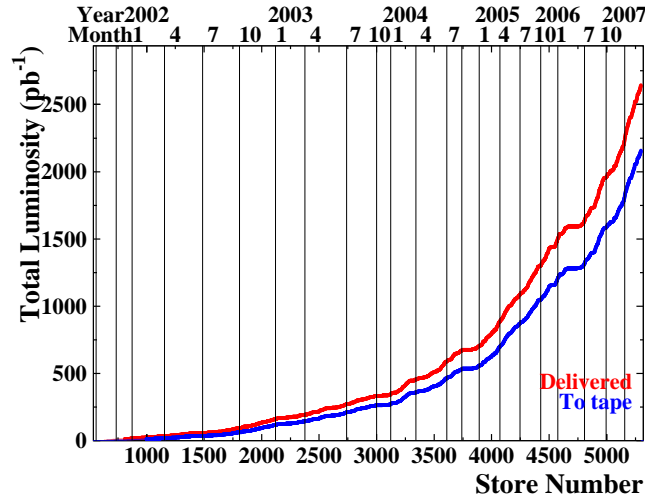


Figure 2.4: Evolution of the delivered and recorded integrated luminosity. The integrated luminosity delivered by the Tevatron is displayed in red, the integrated luminosity recorded by CDF is shown in blue.

Figure 2.4 illustrates the development of \mathcal{L}_{int} over time. Up to date (May 2007), the Tevatron delivered about 2.6 fb^{-1} whereof CDF managed to record approximately 2.1 fb^{-1} , leading to an average data taking efficiency of 82%. The Tevatron Run II goal is to deliver an integrated luminosity between 4.4 fb^{-1} and 8.5 fb^{-1} until August

2009 [51]. The data used in the analysis described in this thesis were taken from February 2002 (store 955) until February 2006 (store 4666) and correspond to an integrated luminosity of 955 pb^{-1} .

2.2 The Collider Detector at Fermilab

The Collider Detector at Fermilab (CDF), located at intersection point B0, is an azimuthally and forward-backward symmetric general purpose solenoid detector [52]. It combines precision charged particle tracking with fast projective calorimetry and fine grained muon detection to record as much information as possible.

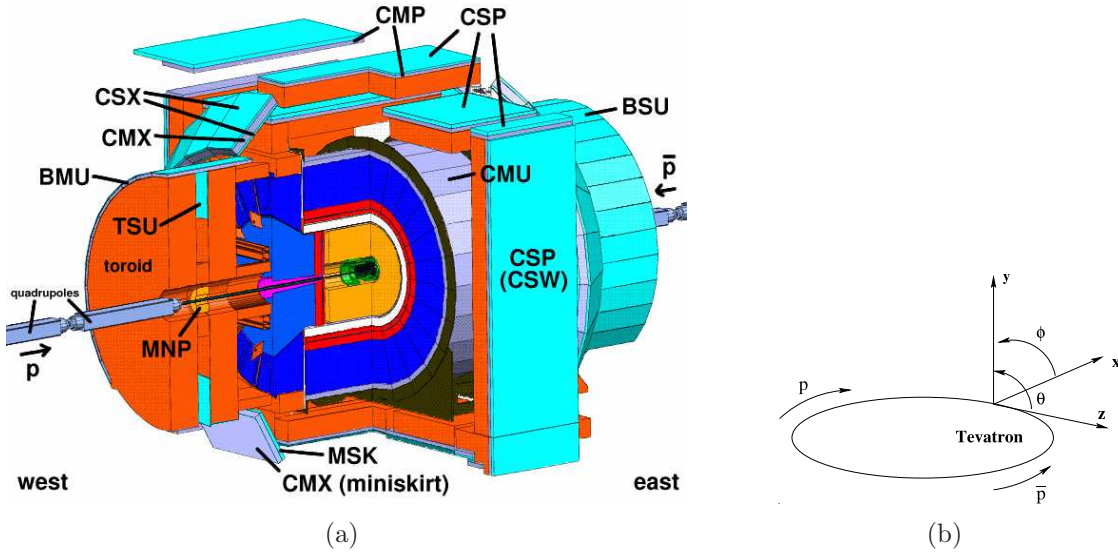


Figure 2.5: (a) Schematic view of the CDF II detector with its different components. The inner green and orange parts represent the tracking system and the red and blue ones the calorimeters. The labels refer to the different components of the muon system. (b) The CDF II coordinate system.

Figure 2.5(a) shows a cutaway view of the CDF II detector with its diverse components. The tracking systems are contained in a superconducting solenoid made of an aluminium-stabilized niobium/titanium conductor. This helium-cooled solenoid, 1.5 m in radius and 4.8 m in length, generates a 1.4 T magnetic field parallel to the beam axis. Calorimetry and muon systems are arranged outside the solenoid.

The detector is described using a cylindrical coordinate system with the z axis along the proton beam, azimuthal angle ϕ , and polar angle θ as displayed in figure 2.5(b). The azimuthal angle is defined with respect to the outgoing radial direction, the polar angle with respect to the proton beam direction. Usually the polar angle is expressed through the pseudorapidity $\eta = -\ln(\tan \frac{\theta}{2})$, motivated by the fact that production of relativistic particles is constant as a function of rapidity. Additionally, the difference in η of two relativistic particles is independent of Lorentz boosts along

the beam axis. The transverse energy and momentum of a particle are defined as $E_T = E \cdot \sin \theta$ and $p_T = p \cdot \sin \theta$, respectively.

2.2.1 Tracking System

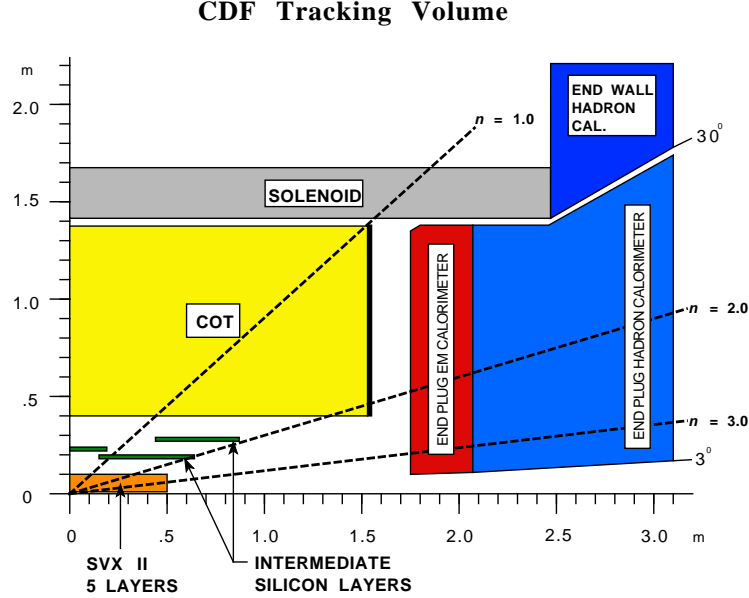


Figure 2.6: Longitudinal view of the CDF II tracking system and the plug calorimetry.

The CDF tracking system, installed close to the beam pipe, is surrounded by a 1.4 T magnetic field bending the tracks of charged particles to helices. Hence, precise measurement of particle momenta is possible. The tracking system, schematically displayed in figure 2.6, consists of two main parts: a barrel shaped silicon microstrip system and an open-cell drift chamber [53] that surrounds the silicon system. The silicon detector comprises three subdetectors, namely the Silicon Vertex detector (SVX II) [54], the Intermediate Silicon Layers (ISL) [55], and so-called Layer 00 [56].

Layer 00 is a radiation hard, single-sided silicon microstrip detector, directly attached to the beam pipe. There are 12 sensors along the beam line for a total length of 94 cm. To prevent gaps, the silicon sensors are arranged in two overlapping sub-layers at radii $r = 1.35$ cm and $r = 1.62$ cm, covering $|\eta| \leq 4.0$. The inner modules are smaller than the outer ones as shown in the schematic overview in figure 2.7(a). Due to the position nearby the beam pipe and hence close to the primary interaction point, Layer 00 provides important contributions to high-quality track reconstruction. SVX II is built in three cylindrical barrels whereof each supports five layers of double-sided silicon strip detectors with radii $2.5 < r < 10.7$ cm. The silicon vertex detector with a total length of 96 cm covers the region with $|\eta| < 2$. The sensors of all SVX II layers feature strips parallel to the z axis on one side, providing detailed measurement of the coordinates in the r - ϕ plane. The strips on the backside of the sensors in layers 0, 1, and 3 (see schematic view in figure 2.7(b)) are rotated by 90°

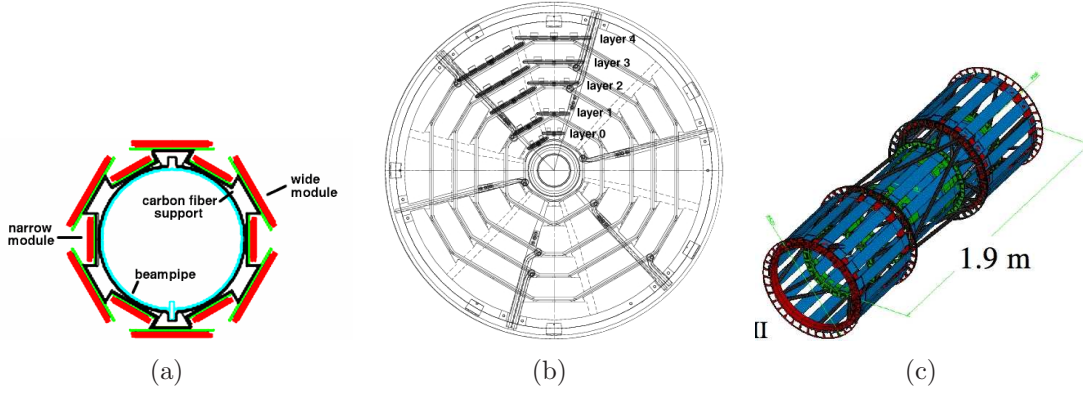


Figure 2.7: Schematic overview of the components of the silicon microstrip detector: (a) Layer 00, (b) SVX II, and (c) Intermediate Silicon Layers.

with respect to the axially aligned strips and hence combine the r - ϕ measurement with precise determination of the z coordinate. To obtain unique three-dimensional information, the strips on the backside of the sensors in the other two layers (layer 2 and layer 4), named small angle stereo (SAS) layers, are twisted by 1.2° with respect to the axially aligned strips. The geometry of the 1.9 m long ISL, consisting of double-sided SAS sensors, is schematically shown in figure 2.7(c). In the central region ($|\eta| < 1.0$), a single ISL layer is placed at a radius of 22 cm. In the plug region, $1.0 \leq |\eta| \leq 2.0$, two layers of silicon are placed at radii of 20 cm and 29 cm. This entire tracking system provides precise three-dimensional track reconstruction. The impact parameter (the closest distance of approach of the particle trajectory to the beam line) resolution of the combination of SVX II and ISL is $40 \mu\text{m}$ including $30 \mu\text{m}$ contribution from the beam width. The z_0 resolution of SVX II and ISL is $70 \mu\text{m}$.

The silicon microstrip system is surrounded by the central outer tracker (COT), a 3.1 m long cylindrical drift chamber filled with a 50 : 50 mixture of argon and ethane gas. The active volume of the COT covers the radial range from 43.4 cm to 132.3 cm and the η range of $|\eta| \leq 1$. The drift chamber provides 96 measurement layers, organized into eight alternating axial and $\pm 2^\circ$ stereo angle superlayers. The hit position resolution is approximately $140 \mu\text{m}$, the momentum resolution $\sigma(p_T)/p_T^2 = 0.0015 (\text{GeV}/c)^{-1}$. Furthermore, the COT provides dE/dx information for the tracks.

2.2.2 Time-of-Flight

A Time-of-Flight (TOF) detector, based on plastic scintillators and fine-mesh photomultipliers, is installed in a few centimeters clearance just outside the COT. The TOF resolution is $\approx 100 \text{ ps}$ and it provides at least two standard deviation separation between K^\pm and π^\pm for momenta $p < 1.6 \text{ GeV}/c$.

2.2.3 Calorimetry

Outside the solenoid, scintillator-based calorimetry allows the measurement of the particle and jet energies by fully absorbing all particles except muons and neutrinos. The calorimeters are segmented into projective towers, each one covering a small range in pseudorapidity and azimuth. The system consists of five units whose coverages, depths, and energy resolutions are summarized in table 2.1: the Central ElectroMagnetic (CEM) [57] and Central HAdron (CHA) [58] calorimeters, covering the η range up to $|\eta| \leq 1.1$ (or $|\eta| \leq 0.9$, respectively), the end-Wall HAdron calorimeter (WHA) [58] and the end-Plug ElectroMagnetic (PEM) and HAdron (PHA) calorimeters [59], which complete a coverage of 2π in azimuth up to $|\eta| \leq 3.64$.

The CEM, directly located outside the solenoid, is a sampling calorimeter consisting of 31 layers of 5 mm thick polystyrene scintillator as the active medium, interspersed with 3.2 mm thick lead, with a depth of 18 radiation lengths. The energy resolution in the CEM is $13.5\%/\sqrt{E} \oplus 1.5\%$. The CHA is 32 layers deep, using acrylic scintillator as active medium, interspersed with 2.5 cm steel absorber, corresponding to 4.7 interaction lengths. Its energy resolution is $50\%/\sqrt{E} \oplus 3\%$ for charged pions that do not interact in the CEM.

Calorimeter	η range	Depth	Energy Resolution
CEM	$ \eta \leq 1.1$	$18 X_0$	$13.5\%/\sqrt{E} \oplus 1.5\%$
PEM	$1.1 \leq \eta \leq 3.64$	$23.2 X_0$	$16\%/\sqrt{E} \oplus 1\%$
CHA	$ \eta \leq 0.9$	$4.7 \lambda_I$	$50\%/\sqrt{E} \oplus 3\%$
WHA	$0.9 \leq \eta \leq 1.3$	$4.7 \lambda_I$	$75\%/\sqrt{E} \oplus 4\%$
PHA	$1.3 \leq \eta \leq 3.64$	$6.8 \lambda_I$	$80\%/\sqrt{E} \oplus 5\%$

Table 2.1: Overview of the η range, depth and energy resolution of the different calorimetry systems. The depth is quoted in radiation lengths X_0 or hadronic interaction lengths λ_I , respectively. The depths and energy resolutions are taken from reference [60].

2.2.4 Muon System

The muon system [61] is located outside the calorimetry to ensure that all particles except for muons and neutrinos are already absorbed. Four systems of scintillators and proportional chambers, filled with a 50 : 50 ratio of argon and ethane (bubbled through ethanol), are used for detection of muons over the region $|\eta| \leq 1.5$. The Central MUon Detection (CMU) system [62], located around the outside of the CHA, consists of four layers of planar drift chambers which cover $|\eta| \leq 0.6$ and is able to detect muons with a transverse momentum $p_T \geq 1.4$ GeV/ c . The Central Muon uPgrade (CMP) [63], forming a box around the central detector, consists of a second set of four layers of drift chambers covered by scintillators (CSP). The CMP is located outside the magnet return yoke, which constitutes another 0.6 m of absorbing steel. It detects tracks of muons with $p_T \geq 2.0$ GeV/ c in the same

η range as the CMU. The Central Muon eXtension (CMX) [63] consists of conical sections of drift tubes and scintillators (CSX) located at each end of the central detector and extends the pseudorapidity coverage of the muon system from 0.6 to 1.0. The fourth muon system, the Barrel MUon detection system (BMU) consisting of drift chambers and scintillators (BSU), is installed on top of the forward toroids (not powered in Run II) and covers the pseudorapidity interval $1.0 \leq |\eta| \leq 1.5$.

Table 2.2 summarizes the main properties of the different muon chambers.

	CMU	CMP	CMX	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]	6.35×2.68	2.54×15.24	2.54×15.24	2.54×8.38
chamber length [cm]	226	640	183	363
drift time [ns]	800	1500	1600	800
min. p_T [GeV/ c]	1.4	2.0	1.4	1.4 - 2.0

Table 2.2: Design parameters of the CDF II muon detectors.

2.3 The CDF Trigger System

Due to the bunch structure of the Tevatron beam, the bunch crossing rate is approximately 2.5 MHz (the original plan to upgrade the Tevatron to a crossing rate of 7.6 MHz was not realized because of beam stability issues). Taking into account the train structure, this is reduced to an effective crossing rate of 1.7 MHz. Since it is impossible to record each collision, it is necessary to draw decisions whether a specific event is worth to be recorded on an event-by-event basis. This is achieved by the CDF three-level trigger system [64], illustrated in figure 2.8. The first two trigger levels are realized by special-purpose hardware, whereas the third one is implemented by software running on a Linux computer farm.

The first level (L1) finds physics objects based on a subset of the detector. Three parallel systems examine each event: calorimeter trigger boards find calorimeter-based objects, muon trigger cards identify muons and the eXtremely Fast Trigger (XFT) reconstructs tracks in the COT and matches those tracks to energy depositions in calorimeter towers or hits in the muon chambers. Information from all three systems is used independently to determine whether an event is passed to the second trigger level (L2). The typical L1 accept rate up to date is 25 kHz.

The L2 trigger performs minimal event reconstruction using custom-designed hardware consisting of several asynchronous subsystems. Besides calorimeter, track, and muon based streams, L2 incorporates information from the CEntral Shower maximum detector (CES) and SVX II. The CES, a strip chamber which is placed in the CEM at a depth corresponding to the average maximum of an electromagnetic shower, provides additional information about photons and electrons. The Silicon Vertex Trigger (SVT) [65] allows selection of tracks with a large impact parameter

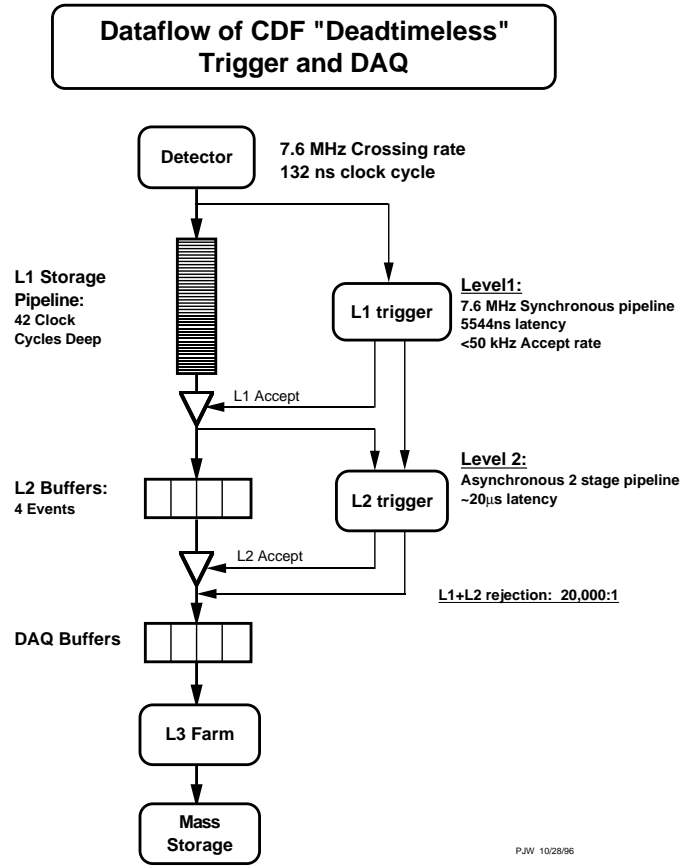


Figure 2.8: Functional block diagram of the CDF II data flow, quoting the initial design rates.

which is most important for b -Meson physics. L2 currently accepts about 600 events per second which are passed to the third trigger level (L3).

The L3 trigger [66] is a processor-based filtering mechanism which has access to the full event record, drawing its conclusions based on the event topology. Accepted events are written to permanent storage media with approximately 100 Hz. To facilitate the handling of the huge collected data volumes, the events passing all three trigger levels are split into eight different data streams. The decision to which stream an event belongs depends on the triggers an event has passed; e.g. all events passing any of the high- p_T lepton triggers end up in “stream B”.

Chapter 3

Neural Networks

Neural Networks are an attempt to emulate the functionality of the human brain with neurons connected among themselves. One successful approach are feed-forward networks with nodes arranged in separate layers, illustrated in figure 3.1. In this approach, each node has only connections from nodes of the preceding and to nodes of the succeeding layer; there is no connection among nodes of the same layer. I.e. information is only “fed forward” from one layer to the next until the output layer is reached.

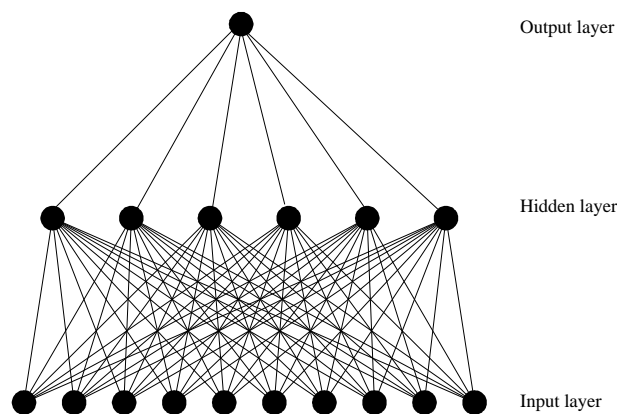


Figure 3.1: General geometry of a three-layer neural network.

Basic functionalities of feed-forward networks as well as specific properties of the NeuroBayes[®] [67] neural network package utilized in this thesis are described in the following sections.

3.1 Functionality of NeuroBayes[®]

NeuroBayes[®] combines a three-layer feed-forward neural network with a complex robust preprocessing. This preprocessing is performed before the input variables are fed to the neural network. The neural network uses Bayesian regularization techniques for the training process. The network infrastructure consists of one input node for each input variable plus one bias node, an arbitrary number of hidden nodes, and one output node which gives a continuous output in the interval $[-1, 1]$.

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

$$a_j(\mathbf{x}) = \sum_i \omega_{ij} x_i + \mu_{0,j} \quad (3.1)$$

and passed to the transfer function which gives the output of the node. The bias $\mu_{0,j}$ implements the threshold of node j . The output of each node is determined by a transformed sigmoid function

$$S(\mathbf{x}) = \frac{2}{1 + e^{-a(\mathbf{x})}} - 1 \quad (3.2)$$

which gives an output of -1 for background and $+1$ for signal. As can be seen in figure 3.2, the sigmoid function is only sensitive to a relatively small range around zero. By this transformation, the interval $[-\infty, +\infty]$ is mapped to the interval $[-1, +1]$. For very large ($x \rightarrow \infty$) or very small ($x \rightarrow -\infty$) values, a saturation effect is reached. The bias mentioned above shifts the mean of the sum of the weighted input data distribution $\sum_i \omega_{ij} x_i$ to the linear part of the sigmoid function.

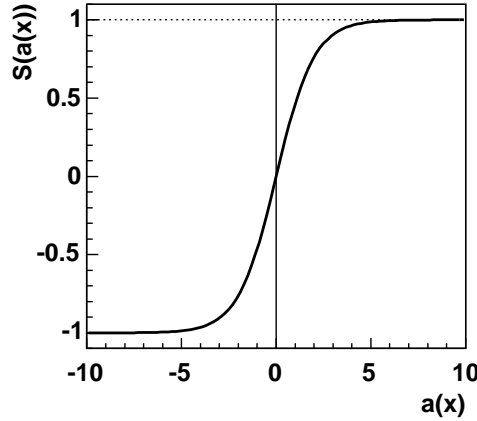


Figure 3.2: The transformed sigmoid activation function $S(a(x))$ as given by equation 3.2.

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

$$o = S\left(\sum_{j=0}^M \omega_j \cdot S\left(\sum_{i=0}^d \omega_{ij} x_i + \mu_{0,j}\right)\right) \quad (3.3)$$

where d is the number of input nodes and M the number of hidden nodes. ω_{ij} denotes the weights from the input to the hidden layer, ω_j the weights from the hidden layer to the output node. $\mu_{0,j}$ is the weight that connects the bias node with the hidden nodes.

3.1.1 The Training Process

The training of a neural network is done by minimizing the deviation between the true output and the one calculated by using the actual weights. The error function minimized in this neural network is the entropy error function

$$E_D = \sum_i \log\left(\frac{1}{2} \cdot (1 + T_i) \cdot o_i + \epsilon\right), \quad (3.4)$$

where the target value T_i is a binary number to classify event i as signal or background, o_i represents the output as given by equation 3.3. ϵ is a small regularization constant which is introduced in order to avoid numerical problems at the beginning of the training. This constant is reduced in each training iteration and is zero after just a few iterations.

The aim of the training of a neural network is to find the minimum in the multidimensional structure of the error function which may exhibit many peaks and valleys. As this task can be difficult to solve, the training process is done by the combined method of backpropagation and gradient descent, i.e. the change of each weight $\Delta\omega_{ij}$ is adjusted proportional to the current gradient of the error function $\Delta\omega_{ij} = -\eta \frac{\partial E_D}{\partial \omega_{ij}}$. The step width η is adapted individually for each weight during the training. Since the target value is not known for hidden nodes, the error induced by the current weights has to be propagated backwards from the output node by applying the chain rule for partial derivatives.

The neural network is trained with regularization techniques to improve generalization performance and to avoid overtraining. During the training process, the weights are systematically reduced in addition to the variation calculated by the gradient descent procedure. Thus, only recurring structures are intensified while the influence of statistical fluctuations is reduced by so-called weight decay. Connections (and even nodes) that have become completely insignificant are pruned away. This reduces the number of free parameters and hence improves the signal-to-noise ratio by removing the cause of the noise, leading to an improved generalization ability. For details of the above mentioned features see references [67, 68].

3.1.2 Preprocessing of the Variables

To find the optimal starting point for minimizing the error function, the input variables are preprocessed. This preprocessing is done in a completely automatic way. Equalizing the input variables and scaling them to be distributed between -1

and 1 before passing the variables to the neural network reduces the influence of extreme outliers. Those flattened distributions are then converted into Gaussian distributions, centered at zero with standard deviation one. At the beginning of the training, this avoids saturation of the nodes due to the above mentioned shape of the activation function (see figure 3.2) and assures that also the inputs to the next layers are distributed with mean zero and width one. To decorrelate the preprocessed input variables, at first, their covariance matrix is calculated. Diagonalizing the covariance matrix using Jacobi rotations [69] and dividing the rotated input vectors by the square root of the corresponding eigenvalue transforms the covariance matrix into a unit matrix.

The above mentioned transformation to a Gaussian distribution may be altered by individual variable preprocessing like fitting a spline curve to the flattened distribution. In addition, discrete variables can be treated as members of classes. The preprocessing of those kinds of variables can also deal with a certain order of values, e.g. the number of tracks in a jet. The preprocessing is also able to deal with variables that are only given for a subset of events by assigning the missing values to a δ function.

3.1.3 Automatic Variable Selection

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

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

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

After the preprocessing process, it is possible to cut on the significance of the variables to incorporate only those that include relevant information that is not already incorporated by other variables. The number of discarded variables is determined by scanning the sorted list, starting with the least relevant one, until the first quantity has a significance larger than the required minimum value.

3.1.4 Training Result

As already mentioned above, the network output of signal events piles up at +1, while background events accumulate at outputs around -1. This is illustrated in

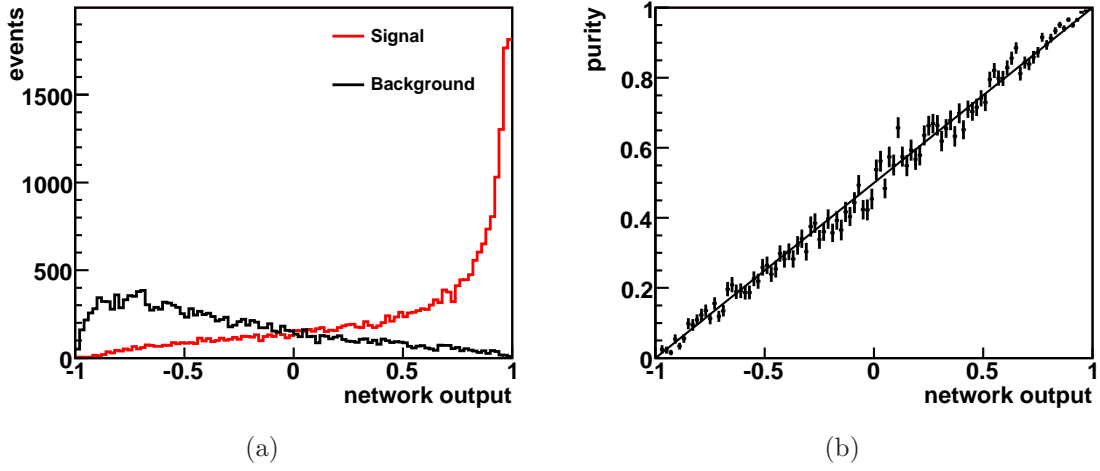


Figure 3.3: Illustration of the training result. Figure (a) shows the output distributions for signal and background events. Figure (b) shows the signal purity of each output bin and illustrates the expected linear dependence.

figure 3.3(a). After minimizing the entropy error function (3.4), the output, rescaled to the interval $[0, 1]$, can be interpreted as Bayesian *a posteriori* probability, if the *a priori* probability is correct, i.e. if a realistic mixture of signal and background has been chosen. Hence, the quality of the training can be checked by plotting the signal purity for each output bin, as illustrated in figure 3.3(b). If the network is well trained, all the points should lie on the diagonal.

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

Chapter 4

Event Simulation and Reconstruction

To evaluate measured data, it is useful to simulate physics processes expected to contribute to the respective data sample. This is usually realized by Monte Carlo generators which randomly produce hard parton interactions according to the probability density of phase space and the matrix element of a given process. To emulate the detector response, the resulting particles are passed to a simulation of the CDF II detector system.

Both measured and simulated events are subject to the same reconstruction algorithms which aim to retrace the physical process starting from measurements in the detector, allowing direct comparison between simulated processes and observed data.

4.1 Monte Carlo Event Generation

The hard interaction of two incoming beams results in the production of up to thousands of outgoing particles. So far, no evidence has been found that contradicts the belief that this process is described by the standard model of strong and electroweak phenomena. Unfortunately, full quantum-mechanical treatment is infeasible due to two reasons: first of all, the number of particles involved gives rise to a tremendous number of interfering contributions that grows factorially with the number of particles. Furthermore, perturbation theory is not able to account for the transition of partons to hadrons. This failure of perturbation theory necessitates other strategies to obtain a detailed description of the production of multiple particles, like the utilization of Monte Carlo event generators.

Any theoretical model describing an elementary process starts from the knowledge of its cross section and must both contain a way to compute or to estimate the effects of higher-order perturbation theory and a way to describe hadronization effects. One way to include higher-order corrections is to exactly compute the

result of a given number of emissions, which forms the core of the tree-level matrix element generators briefly introduced in section 4.1.2. Another way is to estimate the dominant effects due to emissions at all orders in perturbation theory, based on the observation that the dominant effects in certain regions of the phase space have almost trivial dynamics, so extra emissions can be described recursively. One approach in this context is the parton shower technique which is, by construction, the core of the event generators described in section 4.1.1.

4.1.1 Showering and Hadronization Event Generators

Showering and hadronization generators are general purpose tools able to simulate a wide variety of initial and final states, making use of phenomenological models to describe the parton-hadron transition. They begin with a leading order hard subprocess, adding higher order effects by allowing the partons to split into $q\bar{q}$ or gg pairs and by taking gluon emission from quarks into account. The resultant partons are then grouped together or “hadronized” into color-singlet hadrons and resonances are decayed. Finally, the underlying event structure of the event is generated by inclusion of beam remnants, interactions from other partons in the initial hadrons, and collisions between other hadrons in the colliding beams.

Some examples of shower and hadronization generators, applied in this analysis, are shortly described in the next sections.

Pythia

The main emphasis of the PYTHIA [70] event generator is on multi-particle production in collisions between elementary particles, in particular hard interactions involving e^+ , e^- , p , or \bar{p} . PYTHIA contains theory and models for a number of physics aspects, including hard and soft interactions, parton distributions, initial and final-state parton showers, multiple interactions, fragmentation and decay. The program is largely based on original research, but also borrows many formulae and other knowledge from the literature. For the treatment of the hadronization process, the Lund string model [71, 72] is implemented. In this model, color flux tubes are stretched between final-state quarks and antiquarks. The potential energy stored in these strings can be converted into new quark-antiquark pairs which then build colorless hadrons.

Herwig

HERWIG [73] is a general purpose event generator for the simulation of lepton-lepton, lepton-hadron, and hadron-hadron collisions. The program includes a large range of hard scattering processes together with initial and final state radiation using the angular-ordered parton shower, hadronization and hadron decays, and underlying event simulation. HERWIG is particularly sophisticated in its treatment of

the subsequent decay of unstable resonances, including full spin correlations for most processes. The program contains a large library of hard $2 \rightarrow n$ scattering processes for both the standard model and its supersymmetric extension. Particular emphasis lies thereby on the detailed simulation of QCD parton showers. The main difference between HERWIG and PYTHIA lies in the different modeling of the hadronization process. The cluster model used by HERWIG exploits the preconfinement property of perturbative QCD [74] to form color-neutral clusters which decay into colorless hadrons.

4.1.2 Tree-Level Matrix Element Generators

Effects of higher-order corrections in perturbation theory can be taken into account by exact computation of the result of a given (and usually small) number of emissions. This can be realized by considering only those diagrams corresponding to the emission of real particles. Basically, the number of emissions coincides with the perturbative order in α_s . This approach forms the core of the parton-level generators described in this section, which compute tree-level matrix elements for a fixed number of partons in the final state.

These programs generally do not include any form of hadronization, thus the final states consist of bare quarks and gluons. The kinematics of all hard objects in the event are explicitly represented and it is simply assumed that there is a one to one correspondence between hard partons and jets.

However, this assumption may cause problems when interfacing these codes to showering and hadronization programs such as HERWIG or PYTHIA, a step which is necessary in order to obtain a physically sensible description of the production process. In fact, a kinematic configuration with n final-state partons can be obtained by starting from $n - m$ partons generated by the tree-level matrix element generator with the extra m partons provided by the shower. This implies that, although the latter partons are generally softer than or collinear to the former, there is always a non-zero probability that the same n -jet configuration is generated starting from different $(n - m)$ -parton configurations. Hence, to avoid double-counting of certain parts of the phase space, this necessitates a matching of the diverse parton configurations generated by the matrix element generator.

The combination of tree-level matrix element generators and showering programs is essential for analyses based on multi-jet configurations where the standard showering codes are basically unable to describe the kinematics correctly.

Alpgen

ALPGEN [75] is designed for the generation of standard model processes in hadronic collisions with emphasis on final states with large jet multiplicities. It is based on the exact leading order evaluation of partonic matrix elements with the inclusion of b -quark and top-quark masses (c -quark masses are implemented in some cases,

where necessary) and top-quark and gauge-boson decays with helicity correlations. The code generates events in both a weighted and unweighted mode. Weighted generation allows for high-statistics parton-level studies, while unweighted events can be produced in an independent run through shower evolution and hadronization programs, such as HERWIG and PYTHIA.

MadEvent

MADEVENT [76] is a multi-purpose tree-level generator which is powered by the matrix element generator MADGRAPH [77]. Given a standard model process (at any collider, e.g. e^+e^- , ep , pp , $p\bar{p}$), MADGRAPH automatically generates the amplitudes for all relevant subprocesses and produces the mappings for the integration over the phase space. This process-dependent information is passed to MADEVENT and a stand-alone code is produced that allows the user to calculate cross sections and to obtain unweighted events. Once the events have been generated, they may be passed to a shower Monte Carlo program, such as HERWIG or PYTHIA, where partons are perturbatively evolved through the emission of QCD radiation and eventually turned into physical states by hadronization.

4.2 Detector Simulation

It is crucial to not only model the final state particles correctly, but also to describe the response of the CDF II detector to those particles. The modeling of the detector response is based on a detailed simulation using the GEANT3 package [78]. The charge deposition in the silicon layers is calculated using a simple geometrical model based on the path length of the ionizing particle. The drift model used in the COT simulation is based on the GARFIELD [79, 80] package, a general drift chamber simulation program whose parameters were scaled to describe the data. The calorimeter simulation GFLASH [81] was also tuned using test-beam data for electrons and high- p_T pions.

A detailed description of the CDF II detector simulation can be found in reference [82].

4.3 Reconstruction Prerequisites

4.3.1 Track Reconstruction

Detection and tracking of charged particles is an essential part of event analyses at CDF. Since the tracking detectors only measure distinct positions of particles, the reconstruction of tracks is the task of combining all point measurements (hits) along the trajectory to a particle track.

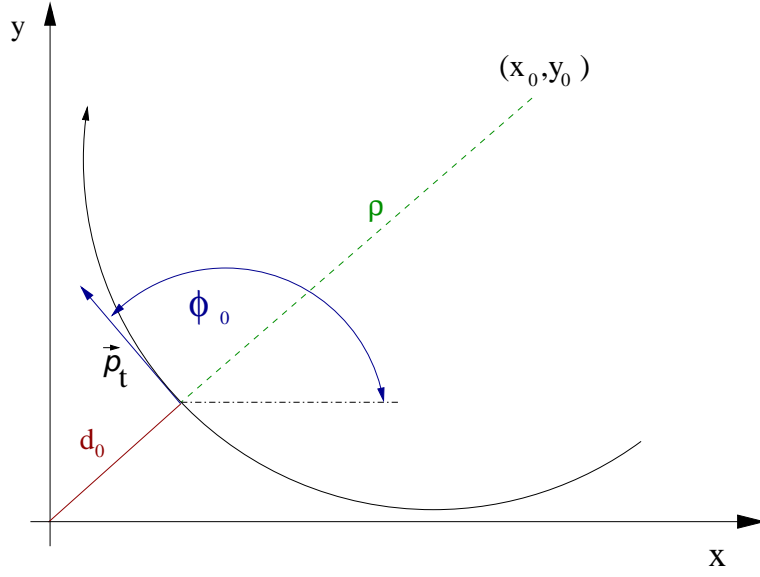


Figure 4.1: Parameterization of a charged particle track in the r - ϕ plane. The parameters are defined with respect to the perigee.

Charged particles moving in a uniform magnetic field, as inside the CDF tracking system, have a helicoidal trajectory. The curvature of the helix depends on the momentum and charge of the particle, the helix' direction points back to its origin. At CDF, such helices are described using five parameters, defined with respect to the point of minimum approach to the origin, the perigee, as illustrated in figure 4.1:

- $\cot \theta$, the cotangent of the polar angle at the perigee, is a measure of the helix pitch and the z component of the momentum. It corresponds to the pseudorapidity η .
- C is the half-curvature of the helix circle in the r - ϕ plane and has the same sign as the charge Q of the particle: $C = \text{sign}(Q)/2\rho$, where ρ is the radius of the circle centered at the perigee (x_0, y_0) . C quantifies the momentum of the particle in the r - ϕ plane, the transverse momentum p_T .
- z_0 is the z coordinate at the perigee.
- d_0 , the signed impact parameter, is the distance between the helix and the origin at the perigee: $d_0 = \text{sign}(Q) \cdot \left(\sqrt{x_0^2 + y_0^2} - \rho \right)$.
- ϕ_0 is the azimuthal direction of the helix at the perigee

The described parameterization uses the origin $(0,0)$ as reference. Due to a shifted beam line position, the coordinate system of the parameterization has to be adapted to the primary interaction point whose reconstruction is shortly described in section 4.3.2.

Since the CDF II tracking system is built of two dissimilar subdetectors, the COT and the silicon system, whose properties are quite different, the pattern recognition strategies vary for each detector. The fit methods in both systems are based on χ^2 minimization.

Tracking in the COT

The active volume of the COT has larger radii than the silicon tracker, which leads to a lower track density and more isolated tracks, resulting in less combinatorics. Thus, the track reconstruction in the COT is purer and faster as in the silicon system. The tracking algorithm can only reconstruct tracks of particles passing the entire volume of the COT. Due to the detector geometry, those tracks are limited to the range of $|\eta| \leq 1$ and $p_T > 0.5 \text{ GeV}/c$ ¹. The track reconstruction is performed in two steps: first, r - ϕ tracks are found using the four axial superlayers; second, stereo segments and hits are attached to these tracks. The axial track reconstruction is performed by two different algorithms run in parallel. The first one, adopted from Run I [83], reconstructs and links segments in the superlayers to find the trajectory. The second algorithm operates at the level of single hits, starting with a single segment in the outermost superlayer to reconstruct reference trajectories. For each hit in a search window, a new trajectory is calculated; the distance of this trajectory to the reference is filled into a histogram which is searched for track candidates. Due to the large activity, the beam line is firstly used as constraint, being released once the majority of tracks has been found. This histogram tracking algorithm is used with small variations at the third trigger level L3. The stereo track reconstruction consists of two algorithms run in series. The first one, stereo segment linking, matches stereo segments to existing axial tracks, while the second one uses z vertex seeds produced using stereo-segment-linked tracks and scans $\cot\theta$ for the best stereo hit usage. Further descriptions of tracking in the COT can be found in references [84, 85]

Tracking in the Silicon Detector

The standard reconstruction consists of three major phases: Tracks from the COT are extrapolated into the silicon detectors, tracks are reconstructed only from silicon measurements, and tracks reconstructed in the silicon are extrapolated into the drift chamber.

In the *outside-in* (OI) approach, tracks from the COT are extrapolated into the silicon detectors by adding hits to the track and recalculating the track parameters. The decision on hits belonging to the track is drawn by two different algorithms run in series. The first [86] is an extended version of the algorithm used in Run I which uses a generic progressive fitter. The second [87] utilizes a fitter (Kalman Fitter) based on the principle of a filter method proposed by R.E. Kalman [88]. To reduce CPU time and to avoid duplicating COT tracks in the OI track list, the second strategy ignores hits which are already used by the first one.

¹With smaller p_T , the particle does not reach the COT due to the curvature of the track.

The *silicon stand-alone* (SiSA) tracking [87] basically works on the same principle as the OI tracking using the Kalman Fitter. To reduce combinatorics, hits already belonging to any other track are not considered. Thus, the main issue of the SiSA tracking is to find tracks in the forward region up to $|\eta| \leq 2$ which is not covered by the COT.

In the last phase of silicon reconstruction, silicon-only tracks are extrapolated into the COT. This *inside-out* (IO) [89, 90] strategy recovers COT information of tracks from particles which do not traverse the entire drift chamber, since those cannot be found by the COT tracking. This confirmation of silicon-only tracks by COT measurements also decreases the fake rate. The geometrical acceptance of the COT in forward direction and the efficiency of the SiSA tracking result in a coverage of about $1.1 \leq |\eta| \leq 1.7$ for IO tracks.

Phoenix Tracking

For the identification of electrons in the forward region, as described in section 5.2, a special algorithm is used which resembles the OI tracking strategy. In this forward tracking algorithm, an energy cluster in the PEM (instead of a COT track) and the primary vertex are used to construct seed tracks. For each seed, two hypotheses about the charge of the particle are considered by computing the curvature for both an electron and a positron corresponding to the deposited energy. The extrapolation of those seed helices into the silicon detector works similarly to the OI tracking algorithm.

4.3.2 Primary Vertex and Beam Line Reconstruction

Accurate determination of the primary interaction point is essential for any high precision analysis, but it is especially important when selecting on the lifetime of a particle. In many applications, the position of the beam line can be used to estimate the primary vertex position which is defined as the origin of all tracks with zero or small impact parameters, so-called prompt tracks. The precision of this method is limited by the size of the beam width, but proved to be sufficient for most applications in b -quark physics where the beam line position is determined by exploiting the correlation between the impact parameter d_0 and the azimuthal direction ϕ_0 of a track.

For events with high decay multiplicity (e.g. top-quark pair production), the primary vertex can be found with better precision on an event-by-event basis. Primary vertices are reconstructed by fitting prompt tracks fulfilling certain quality requirements to a common vertex. Tracks contributing a large χ^2 to the fitted vertex are iteratively removed if the χ^2 exceeds a given threshold. The iteration stops either if no track fails the χ^2 cut or the number of tracks associated to the vertex falls below a minimum quantity.

At CDF, different primary vertex reconstruction algorithms are implemented. The

distribution of primary vertices reconstructed with the `VxPrim` [91] algorithm is used to derive the position of the beam line [92]. Here, profile plots of the vertex positions in x respectively y over z are used to fit the beam line. The position of the primary interaction point (e.g. necessary to determine secondary decay vertices) is evaluated using the `PrimeVtx` [93] package which utilizes another vertex fitter function than `VxPrim`. `PrimeVtx` is applied to tracks around a seed vertex which is commonly given by the beam line in the transverse plane. The seed's z component is given by the highest- p_T z vertex from `ZVertexColl` which is based on associating reconstructed tracks to pre-tracking vertices. If a certain number of tracks passing minimal quality criteria is consistent with a vertex, the z position of the vertex is calculated by the error weighted mean of the z_0 parameters of the tracks.

Pre-track primary vertices, i.e. vertices which are reconstructed without tracks, are built from hits in the SVX II and in the ISL using a histogram method [94]. The algorithm requires two 3D hits within 9° in azimuth and tries to find a third confirmation hit within a certain window. If no third track is found, the beam spot is used instead to form a triplet. Assuming the hits belong to a track, the corresponding z_0 is calculated and filled into histograms. From those histograms, the z positions of the vertices are determined by selecting the maxima.

4.3.3 Jet Reconstruction

In hadron-hadron collider experiments like CDF II, a quark (or antiquark) in its final form manifests itself as one or more calorimeter jets, which appear as energy deposits shared among several detector calorimeter towers. The development from scattered partons to detector jets is schematically illustrated in figure 4.2.

At CDF, the jet energy is calculated from the energy deposited in the calorimeter towers using different types of jet clustering algorithms. In this analysis, jets are clustered using a cone algorithm with a fixed cone size in which the center of the jet is defined as $(\eta_{\text{jet}}, \phi_{\text{jet}})$ and the size of the cone as $R = \sqrt{(\eta_{\text{tower}} - \eta_{\text{jet}})^2 + (\phi_{\text{tower}} - \phi_{\text{jet}})^2} \leq 0.4$.

The clustering algorithm groups calorimeter towers with $E_{T_i} > 1$ GeV into jets. Here, $E_{T_i} = E_i \cdot \sin \theta$ is the transverse energy deposited in tower i with respect to the primary vertex z position, E_i is the sum of energies measured in the electromagnetic and hadronic compartments of that tower. The algorithm begins with creating a list of seed towers, sorted by decreasing E_{T_i} . For each seed tower, adjacent towers within a radius of size R with respect to its position are used to build clusters. Once an initial list of towers is obtained, the transverse energy and the position of the cluster are calculated. This procedure is repeated iteratively, a new list of towers around the new center is determined. The jet E_T and direction are recalculated until the list of towers assigned to the cluster is stable, that is, when the geometrical center of the tower corresponds to the cluster centroid. Jets are merged if they overlap by more than 50%; otherwise, each tower in the overlap region is assigned to

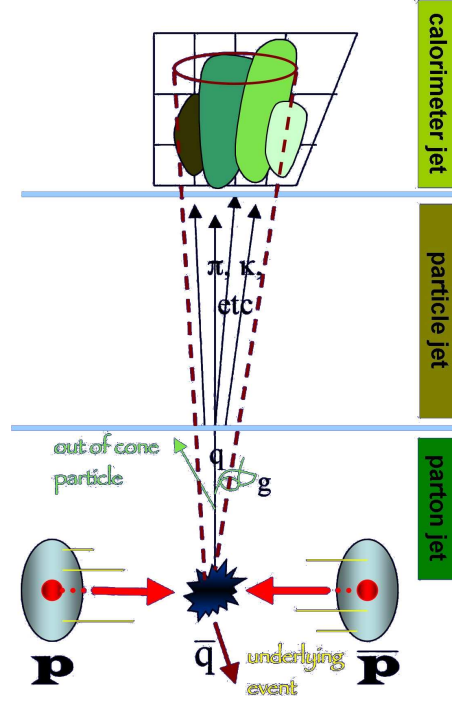


Figure 4.2: Schematic overview of jet reconstruction.

the nearest jet. The final jet energy is computed from the final list of towers N_{tow} :

$$E_{\text{jet}} = \sum_{i=0}^{N_{\text{tow}}} E_i \quad (4.1)$$

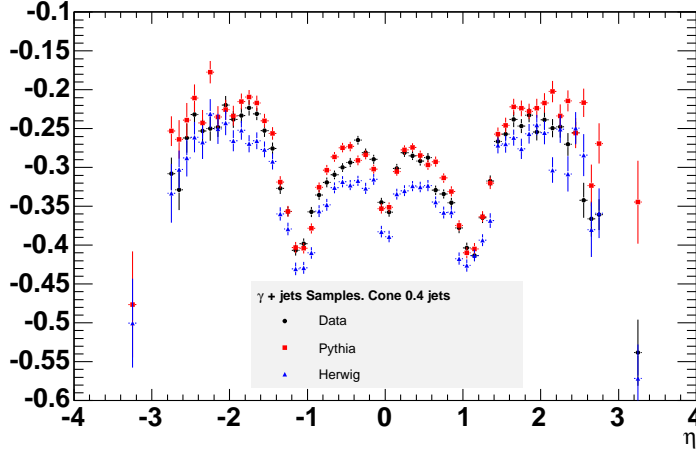
Jet–Energy Correction

At CDF, the clustered energy of a detector jet has to be corrected for several effects that can distort the measured jet energy, such as response of the calorimeter to different particles, non-linear response of the calorimeter to the particle energies, uninstrumented regions of the detector, spectator interactions, and energy radiated outside the jet cone. Since the corrections are divided into discrete levels to accommodate different effects, a subset of these corrections can be applied, depending on the analysis.

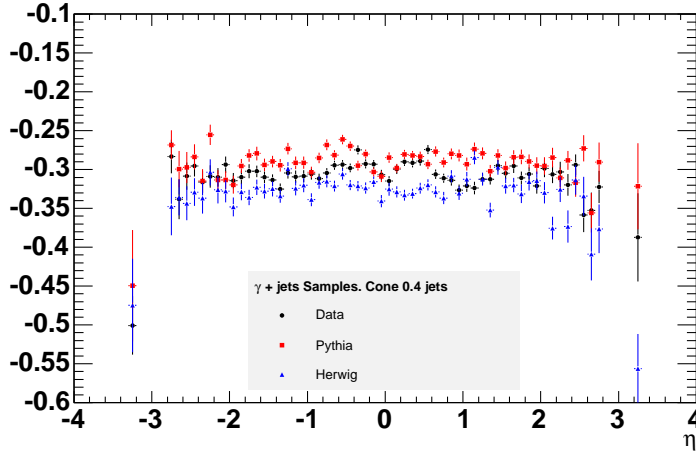
It is possible to scale the measured energy of a jet back to the energy of the final-state particle-level jet; additionally, there are corrections to associate the measured jet energy to the parent parton energy, so that direct comparison to the theory can be made. Detailed information about the jet–energy corrections can be found in reference [60].

Level 1: η –dependent Corrections Due to the geometry of the CDF calorimeter, its response is not uniform in pseudorapidity as illustrated in figure 4.3(a). The

dependencies on η arise from the separation of calorimeter components at $\eta = 0$ where the two halves of the central calorimeter join and at $\eta \approx 1.1$ where the plug and the central calorimeter abut. The different responses of the plug and the central calorimeter also cause a dependence on η .



(a)



(b)

Figure 4.3: Calorimeter response of photon-jet events from data and MC samples as a function of pseudorapidity (a) before and (b) after the correction of η -dependence.

The η -dependent corrections are determined based on the assumption that the two jets in dijet events should be balanced in p_T in absence of hard QCD radiation. Since the central region of the calorimeter is better understood, the forward region response is intended to be scaled to the central region. In addition, a region far away from the cracks has to be selected as reference. Hence, a “trigger jet” with $0.2 < |\eta| < 0.6$ is chosen, the other jet is defined as “probe jet”. Scaling the p_T of the probe jet to balance the trigger jet defines the η -dependent correction.

To check the correction, photon-jet events are utilized, since the photon should be balanced by a jet in these kind of events. After the application of the correction, the calorimeter response as a function of η is flat as displayed in figure 4.3(b).

Level 2: not applied in Run II

Level 3: not applied in Run II

Level 4: Multiple $p\bar{p}$ Interactions At the current luminosities, more than one $p\bar{p}$ interaction occurs in the same bunch crossing. These additional $p\bar{p}$ interactions increase the energy of the jets from the hard scatter if their final-state hadrons accidentally overlap with the jets. This extra energy therefore needs to be subtracted from the jet energy.

To estimate the number of interactions in a bunch crossing, utilizing the number of reconstructed vertices, N_{vtx} , is the best approach. Using minimum-bias data triggered with hits in gaseous Cherenkov-light detectors (CLC) in the very forward direction $3.7 < |\eta| < 4.7$, the transverse energy in a cone around a random seed tower in the central region $0.2 < |\eta| < 0.6$ is measured in dependence on N_{vtx} . The slope of the straight line fitted to this distribution, displayed in figure 4.4, gives the extra transverse energy per interaction as a function of N_{vtx} .

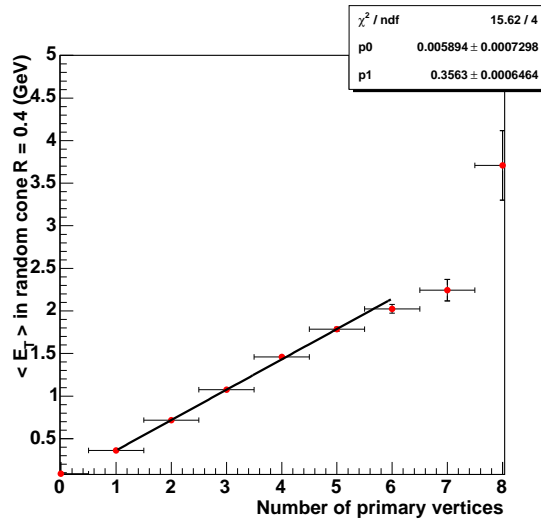


Figure 4.4: The dependence of the jet E_T on the number of $p\bar{p}$ interactions, i.e. the number of primary vertices N_{vtx} . The slope of the fitted straight line gives the extra transverse energy per interaction as a function of N_{vtx} .

Level 5: Absolute Jet Energy Scale The absolute correction aims to transform the jet energy measured in the calorimeter into the energy corresponding to

the underlying particle jet. After correcting for any non-linearity and energy loss in the uninstrumented regions of each calorimeter, the energy scale of a jet is independent of the CDF detectors. Since the calorimeter simulation has been optimized to reproduce the measured single particle response, it is possible to rely on the simulation to derive corrections over a large range of jet transverse momenta. Hence, the accuracy of this method depends on how well jets are modeled by the simulation. In particular, it depends on the multiplicity and p_T spectrum of the particles inside a jet and on the calorimeter response to an individual particle.

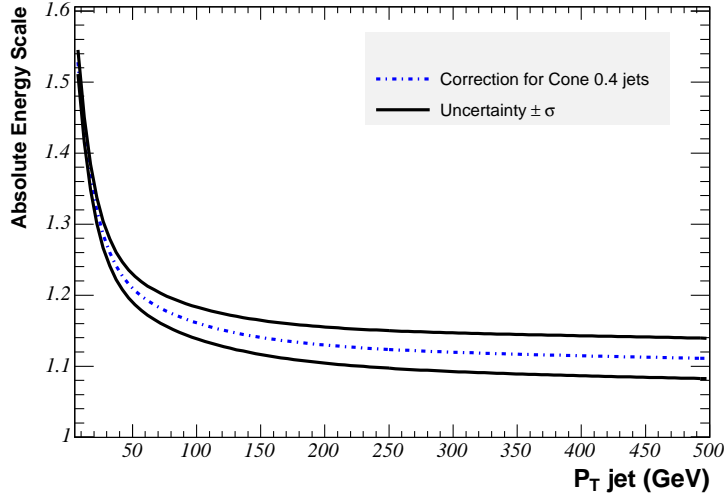


Figure 4.5: Absolute jet energy corrections as a function of jet p_T .

The correction displayed in figure 4.5 is obtained by mapping the total p_T of the particle jet to the p_T of the calorimeter jet, where the jets are required to be in the central region $0.2 < |\eta| < 0.6$. The particle jet consists of particles within a cone of $R = 0.4$ and must be within $\Delta R < 0.1$ of the calorimeter jet.

Level 6: Underlying Event & Level 7: Out-of-Cone It is often desirable to reconstruct the energy of the original parton rather than the energy of the jet, e.g. for the measurement of the top-quark mass or the search for the Higgs boson, where parton energies are used to compute the invariant mass of the decaying products.

The reconstruction of the parton energy from the particle jet is subject to several difficulties. A fraction of the parton energy can be lost from the jet cone due to final-state gluon-radiation (FSR) at large angles with respect to the parent parton or due to particles exiting the cone either in the fragmentation process or due to low p_T particles bending in the magnetic field. This energy is called “out-of-cone” energy. On the other hand, the particle jet can also have contributions not related to the actual mother parton of the hard interaction of interest defining the jet, such as particles from initial state gluon radiation (ISR) or particles from spectator partons

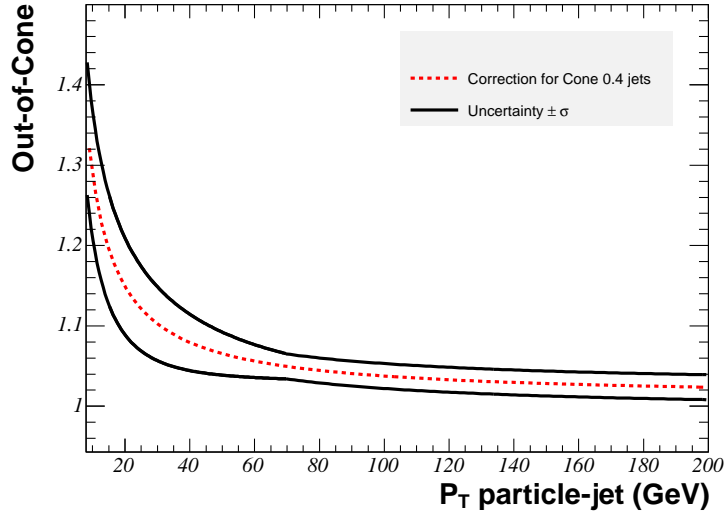


Figure 4.6: Out-of-cone jet energy corrections as function of the p_T of the particle jet.

with color connection to the other partons of the proton (“beam-beam-remnant”). These two contributions are called “underlying event”.

Final state radiation and hadronization effects are correlated with the primary jet direction and the jet energy and are expected to decrease with increasing distance from the jet core. The underlying event is thought to be uncorrelated with the direction of the outgoing parton and thus independent of the distance from the jet in η - ϕ space and almost independent of the jet energy.

The corrections for the out-of-cone energy and the underlying event are derived simultaneously using MC samples which have been tuned to describe the underlying event in data. Since any η -dependence of the out-of-cone corrections is already taken into account by the first level corrections, jets with $0.2 < |\eta| < 0.6$ are used to obtain the corrections. Those corrections are solely determined from simulation at particle generator level and hence independent from the CDF detector.

The corrections are obtained from simulated dijet samples using particle jets which match a primary parton within $\Delta R < 0.4$. The out-of-cone correction is shown in figure 4.6. For particles with $p_T = 20$ GeV/ c , the correction is about +18%. The underlying event’s transverse energy is estimated to be about 0.4 GeV. Hence, for jets with a cone of $R = 0.4$, the reconstruction of the parton energy from the particle jet is dominated by the correction for out-of-cone losses.

4.3.4 Secondary Vertex Reconstruction

The identification of b -quark jets is essential for top-quark physics, since the top quark decays in almost 100% of the cases into a W boson and a b quark. The b quark

originating from the top–quark decay hadronizes almost immediately to form a jet of particles; included are a b meson (for example B^0, B^\pm, B_s^0) or a b baryon (e.g. Λ_B). The b hadron usually carries most of the momentum of the original b –quark and has a relatively long lifetime of about 1.6 ps. Given their long lifetime and large boost, b hadrons created in this way travel a macroscopic distance away from the primary interaction point in the laboratory rest frame before decaying into several charged and neutral particles. Reconstruction of charged particle tracks enables us to look for the trajectories of the decay products that have a large impact parameter and hence are inconsistent with originating from the initial interaction point. Several of those tracks with large impact parameters, called displaced tracks, can be determined to originate from a common location and can be used to construct a secondary vertex, as illustrated in figure 4.7.

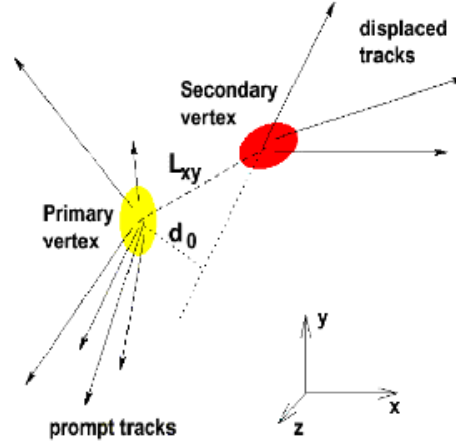


Figure 4.7: Schematic view of displaced tracks with impact parameter d_0 forming a secondary vertex. L_{xy} is defined as the projection of the two–dimensional decay length on the jet axis.

The CDF secondary vertex reconstruction algorithm, SecVtx [95], runs on an per–jet basis within each event. SecVtx starts by considering silicon tracks within each jet ($\Delta R < 0.4$); the tracks must be seeded (OI) or confirmed (IO) by a track in the COT. To be considered for SecVtx, the tracks within the jet are demanded to have $p_T > 0.5$ GeV/ c , d_0 significance $S_{d_0} \equiv |d_0/\sigma_{d_0}| > 2.0$ with respect to the primary vertex, and a minimum number of hits in the silicon detectors. The hit requirements are a function of the detector geometries and the track reconstruction quality. Tracks are further demanded to not exceed a maximum d_0 requirement (0.15 cm) in order to protect against poorly reconstructed tracks as well as tracks from long–lived light–flavor hadrons or nuclear interactions in the detector material.

The selected tracks are then ordered in p_T and a secondary vertex is sought among these tracks. The construction of a two–track seed vertex is first attempted among the qualifying tracks. If a seed vertex is found, the remaining tracks are considered for vertexing with the seed tracks. Each additional track is considered singly; after attaching all qualifying tracks to the vertex, the vertex χ^2 is recalculated and tracks

iteratively pruned from the vertex if they contribute too greatly to the overall χ^2 .

If the pruned vertex retains three or more tracks, this vertex is then subject to a final round of quality cuts. Vertices from material and nuclear interactions are removed by restraining the absolute value of L_{xy} which is defined as the 2D decay length of the fitted vertex with respect to the primary vertex, projected on the jet axis; all vertices with a radius greater than 2.5 cm with respect to the center of the SVX II are vetoed. In addition, vertices are excluded if their invariant mass is consistent with the masses of K_S or Λ , two prominent long-lived light-flavor hadrons. Finally, the vertex is demanded to have $S_{L_{xy}} > 7.5$, where $S_{L_{xy}}$ is the significance of L_{xy} , defined as $S_{L_{xy}} \equiv |L_{xy}/\sigma_{L_{xy}}|$. If the vertex satisfies all of the above criteria, a secondary vertex is defined to be found.

If no candidate vertex is found in the first pass (pass 1), a second attempt (pass 2) to reconstruct a vertex is made. Efficiency is gained by only requiring two or more tracks satisfying more stringent track quality requirements. A summary of all track and vertex requirements can be found in reference [96].

If a secondary vertex is found, the jet is said to be “ b tagged”. If the dot product of the 2D displacement vector from the primary vertex to the secondary vertex and the jet’s momentum vector is positive (i.e. the vertex is in the same hemisphere of the detector as the jet), the tag is called “positive”. If the secondary vertex and jet momentum have a negative dot product (the vertex is in the opposite hemisphere of the detector actually behind the jet), the tag is called “negative”. Such vertices are due to finite tracking resolution of the detector and are predominantly not related to heavy-flavor decays.

The SecVtx algorithm exists as a loose, a tight, and an ultra-tight version. The above described criteria are valid for the tight version which is used by default. The loose algorithm features an increased tag efficiency, while the ultra-tight version increases the purity.

Tag Efficiency

The efficiency of the SecVtx algorithm is defined as the fraction of b -quark jets fiducial to the COT and calorimetry that possess a positive SecVtx b tag. Measuring the efficiency of a b -tagging algorithm in MC events is straightforward, since one has the complete knowledge of the produced particles and thus it is unpretentious to identify the fiducial jets that come from b -quark production and the fraction which are tagged. Unfortunately, the efficiency in MC jets is not accurate; reliable modeling of b tagging requires precise understanding of the charge deposition in the silicon detectors, accurate simulation of the tracking, and realistic b hadron production and decay models. Since none of these effects are perfectly modeled in the simulation, it is imperative to measure the b -tag efficiency in the data.

The challenge in measuring the tag efficiency in data events is that the nature of individual jets is not explicitly known. The tag efficiency measurement in data relies upon constructing a pure sample of b -quark jets within the large dijet sample. Two

methods currently in use at CDF utilize high- p_T leptons matched to jets to identify jet pairs consistent with heavy-flavor production.

The first technique requires a high energetic electron to be embedded in a jet. This so-called “electron jet” is paired with a back-to-back jet, the “away jet”, which is demanded to possess a positive SecVtx tag. This jet pair (one jet containing a secondary vertex, the other having evidence for a high- p_T semileptonic decay) is consistent with coming from heavy-flavor production. Electron jets being consistent with coming from a photon conversion, i.e. electrons that have a conversion partner, are not further considered, since they are not compliant with semileptonic hadrons decays. The tag efficiency is then basically determined by the rate that the electron jet is tagged in addition to the away jet, taking into account the heavy-flavor fraction. One method first measures the heavy-flavor fraction of the electron jets in the untagged sample by identifying those electron jets that also contain a D meson [97], while another method solves simultaneously for the efficiency and heavy-flavor fraction [98]. The tagging efficiency obtained by this method in data is $\epsilon_{\text{data}} \approx 34\%$ integrated over the complete jet- E_T range.

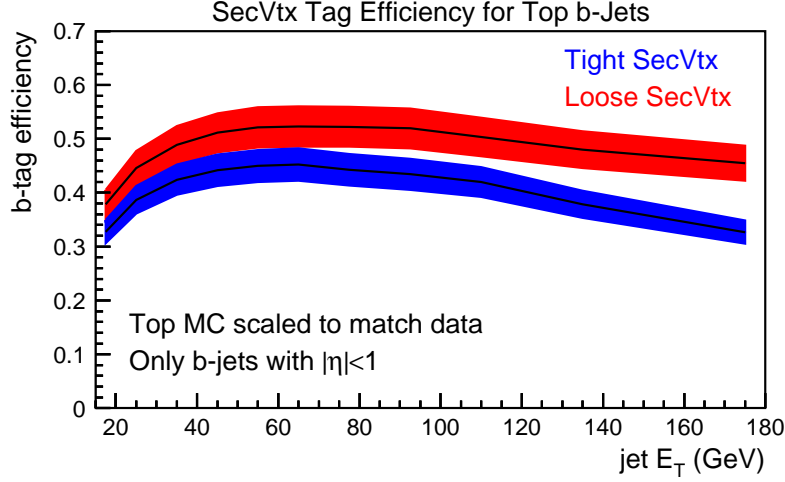
The tag efficiency is used most often when assessing signal acceptance which is typically done in MC samples. As discussed above, MC b -quark jets are not guaranteed to perfectly match data b -quark jets. So it is necessary to construct a data-to-MC scale factor for the tag efficiency, which encapsulates the discrepancies between b -quark jet tagging in the simulation and data. In an appropriate dijet MC sample matching the conditions used above to select data events, the tag efficiency was measured to be $\epsilon_{\text{MC}} \approx 39\%$. The scale factor, which is basically independent of the jet E_T , is calculated to be $SF = 0.89 \pm 0.07$.

A second technique uses high- p_T muons instead of electrons, but works on the same principle. This method leads to a scale factor of $SF = 0.92 \pm 0.06$. A combination of both methods leads to $SF = 0.91 \pm 0.06$. The corrected tagging efficiency for b jets resulting from decaying top quarks can be seen in figure 4.8.

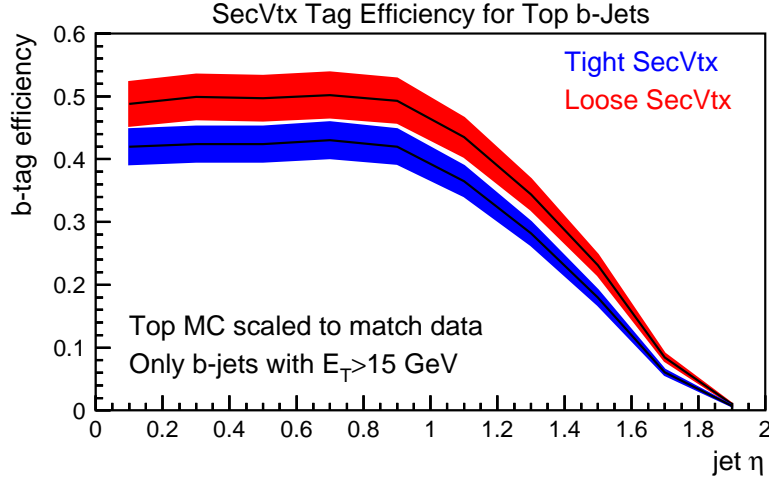
Since the determination of a scale factor using high energetic muons was not yet available when this analysis was performed, a scale factor of $SF = 0.89 \pm 0.07$ is used.

False Positive Tagging Rate

False positive tags, or mistags, in SecVtx come from the spurious identification of a secondary vertex in a jet which doesn’t contain a heavy-flavor quark. Jets from light-quark production should be consistent with zero lifetime. However, tracks within a light-quark jet can still have large impact parameters and hence satisfy the secondary vertex requirements. Sources of such spurious large impact parameters are limited detector resolution, long-lived light-particle decays (Λ , K_S), and material interactions. Since mistags due to limited detector resolution are expected to be symmetric in L_{xy} , the ensemble of negatively tagged jets ($L_{xy} < 0$) is a good estimate of the light-flavor jet contribution to the positive tag sample.



(a)



(b)

Figure 4.8: SecVtx tag efficiency for b jets in top-quark decay scaled to data as a function of (a) jet E_T and (b) jet η for two different operating points. In this analysis, the tight version is utilized.

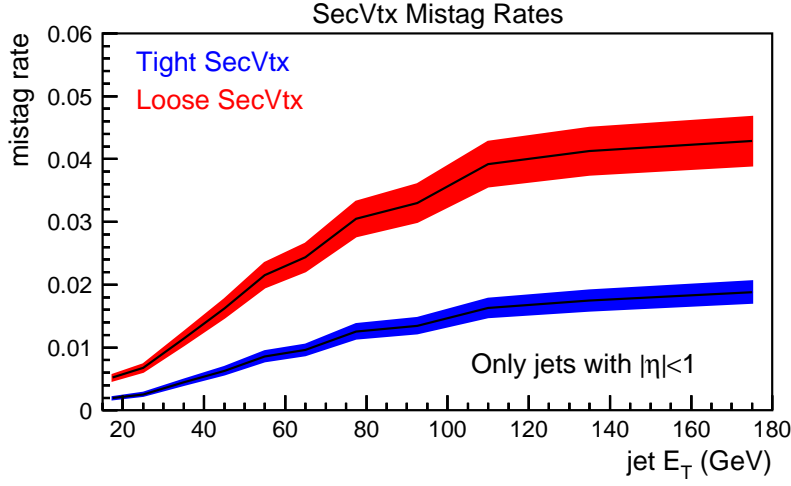
At CDF, events collected by a set of triggers requiring a minimum amount of calorimeter energy, generic jet samples, are used for the calibration of the mistag rate. The probability for a given jet to be a mistag is determined from the probability that the jet is a negative tag, parameterized in five variables: E_T , η , and ϕ of the jet, track multiplicity and the sum of the transverse energies of the taggable jets in the event. A jet is considered as taggable, if it fulfills the standards to be potentially b tagged, namely $E_T > 10$ GeV, $|\eta| < 2.4$, and at least two reconstructed tracks.

This per-jet probability can be used to estimate the expected contribution from events with mistagged jets to a given data sample.

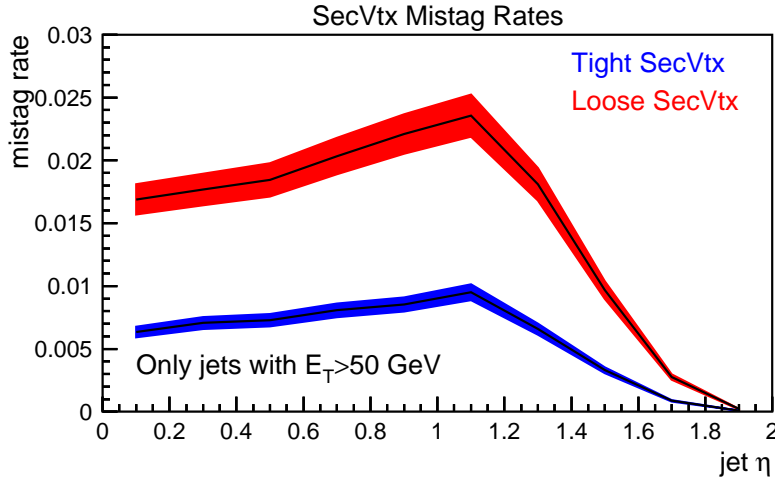
However, since not all mistags are from resolution effects alone, simply assuming

that all mistagged jets are symmetric around the origin will lead to an underestimate in the rate of false positive tags. Contributions from long-lived particles and material interactions are at strictly positive values of L_{xy} and thus introduce a mistag asymmetry. The asymmetry factor γ was evaluated by a fit of the pseudo-lifetime $c\tau = L_{xy} \times \frac{M_{\text{vtx}}}{p_{T\text{vtx}}}$ to be $\gamma \approx 1.36$, but depends on the sum of the transverse energies of all jets in the event.

With this parameterization and asymmetry, the mistag rate in a given sample can be determined as shown in figure 4.9.



(a)



(b)

Figure 4.9: SecVtx mistag rate in data as a function of (a) jet E_T and (b) jet η for two different operating points. In this analysis, the tight version is utilized.

Chapter 5

Event Selection and Background Estimate

As outlined in section 1.2.2, single-top-quark events feature a signature of jets plus the decay products of a W boson. Studies have shown that about 60% of single-top-quark events possess exactly two jets [99]. For events produced via t -channel, this is true because the \bar{b} quark produced in the $2 \rightarrow 3$ process, illustrated in figure 1.4(b), has a tendency to be produced in forward direction. In this analysis, only leptonic decays of the W boson into $e\nu_e$ or $\mu\nu_\mu$ are targeted in order to reduce background events from multi-jet production via the strong interaction.

For multivariate methods, it is of major importance to understand and to model the selected events correctly. Hence, not only the modeling of single top-quark production is subject to thorough evaluation, but also backgrounds leading to the same event topology have to be studied diligently.

5.1 Measured Data Samples

In this analysis, data taken from February 2002 until February 2006 are used. Since the focus lies on leptonically decaying W bosons, data passing the high- p_T lepton triggers are used. The central electron trigger ELECTRON_CENTRAL_18 requires a COT track with $p_T > 9$ GeV/ c matched to an energy cluster in the CEM with $E_T > 18$ GeV. The shower profile of this cluster has to be consistent with the expectation obtained by measurements with test-beam electrons. Forward electron candidates have to pass the MET_PEM trigger requiring an energy deposition of at least 20 GeV in the PEM. The ratio of hadronic-to-electromagnetic energy $E_{\text{HAD}}/E_{\text{EM}}$ has to be less than 0.075. The muon triggers MUON_CMUP18 and MUON_CMX18 require a COT track with $p_T > 18$ GeV/ c matched to a track segment in the muon chambers.

The data are stripped into different datasets according to the trigger they passed: the `bhel` datasets contain central electrons, the `bhmu` datasets incorporate muons,

and the **bpe1** datasets comprise forward electrons.

Before being used in analyses, it is crucial to reprocess the data offline. During reprocessing, online calibrations are checked and readjusted, the alignment of the silicon detector is corrected, tracks are refit, cluster energies are checked, and leptons are identified. In addition, the jet clustering and the secondary vertex fit are performed. The data taken from February 2002 until August 2004 were reprocessed with CDFSOF2 [100] version 5.3.3 and stripped into the **0d** datasets, while data taken from December 2004 until February 2006 were reprocessed with CDFSOF2 version 6.1.2 and stripped into the **0h** and **0i** datasets. All nine datasets used in this analysis are presented in table 5.1.

Sample	Run Range	Dates	Events	\mathcal{L} [pb ⁻¹]
bhe10d	138425 — 186598	02/04/2002 — 08/22/2004	1255715	333
bhmu0d	138425 — 186598	02/04/2002 — 08/22/2004	552401	333/320
bpe10d	138425 — 186598	02/04/2002 — 08/22/2004	358693	333
bhe10h	190697 — 203799	12/07/2004 — 09/04/2005	1176549	363
bhmu0h	190697 — 203799	12/07/2004 — 09/04/2005	574704	363
bpe10h	190697 — 203799	12/07/2004 — 09/04/2005	326672	363
bhe10i	203819 — 212133	09/05/2005 — 02/22/2006	730697	259
bhmu0i	203819 — 212133	09/05/2005 — 02/22/2006	358639	259
bpe10i	203819 — 212133	09/05/2005 — 02/22/2006	225722	259

Table 5.1: Used data samples with run range, date of data taking, number of events after stripping of the samples, and integrated luminosity. The luminosity values for **bhmu0d** correspond to CMUP/CMX.

Each store includes several periods of continuous data taking, so-called runs, since from time to time, data acquisition stops due to hardware or software failures. It is of vital importance to revise every single run to verify that every detector component was both functional and active. Every run marked as “good”, i.e. all necessary system components were in operation, can be used in analyses and is incorporated into the so-called “goodrun” list [101]. In this analysis, version 13 [102] of the goodrun list was used, requiring the whole tracking system, calorimetry, and muon chambers to be operating. This leads to a total integrated luminosity of $\mathcal{L} = (955 \pm 75) \text{ pb}^{-1}$.

5.2 Requirements for Candidate Events

Lepton Identification

After offline reconstruction, the lepton candidates have to pass further cuts in order to improve the purity. For central electrons, a reconstructed track with $p_T > 9 \text{ GeV}/c$ has to match a cluster in the CEM with $E_T > 20 \text{ GeV}$. Furthermore, $E_{\text{HAD}}/E_{\text{EM}} < 0.055 + 0.00045 \cdot E$ is required; the ratio of cluster energy

to track momentum E/p has to be smaller than 2.0 for track momenta ≤ 50 GeV/ c . Electron candidates in forward direction are defined by a cluster in the PEM with $E_T > 20$ GeV and $E_{\text{HAD}}/E_{\text{EM}} < 0.05$. The cluster position and the primary vertex are combined to form a search trajectory in the silicon tracker and seed the pattern recognition of the tracking algorithm. For the electron candidate to pass the selection, the found track has to fulfill certain quality criteria. Electron events are rejected, if an additional high- p_T track is found that forms a common vertex with the track of the electron candidate and has a curvature of opposite sign. These events are likely to stem from the conversion of a photon.

Muons are identified by requiring a COT track with $p_T > 20$ GeV/ c that extrapolates to a track segment in a muon chamber. Signal muons have to be detected in the CMU and CMP simultaneously or in the CMX. In order to minimize background contaminations further requirements are imposed. The energy depositions in the electromagnetic and hadronic calorimeters have to correspond to the expectation regarding minimum ionizing particles. To reject cosmic muons or muons from in-flight decays of long-lived particles (such as K_S , K_L , or Λ), the impact parameter d_0 of the track must be small. Cosmic muons are further rejected through their characteristic track timing and topology.

Furthermore, exactly one isolated lepton candidate is required, whereby a candidate is considered isolated if the E_T not assigned to the lepton in a cone of $R = 0.4$ centered around the lepton is less than 10% of the lepton E_T or p_T , respectively. This lepton is called tight lepton.

A detailed description of all lepton requirements can be found in appendix A of this thesis as well as in references [103, 104, 105].

Dilepton Veto

To ensure that there is exactly one tight lepton, events are rejected which have either an additional tight lepton or a loose lepton. Loose leptons are leptons which pass all cuts except the isolation cut, or are identified in the CMP, CMU or BMU solely.

Jet Reconstruction and Selection

In this analysis, jets are reconstructed with a cone of $R = 0.4$ without taking into account calorimeter towers which are associated to any tight isolated electron. The jet energy is corrected up to level 4, i.e for the η -dependence of the calorimeter response and for multiple $p\bar{p}$ interactions. Candidate jets, required to have detector $|\eta| < 2.8$, must have corrected $E_T > 15$ GeV to be called tight jets, whereas loose jets must have corrected E_T between 8 GeV and 15 GeV. Detector η 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. Only events with exactly two tight jets are accepted, whereby at least one of the jets must be tagged as a b -quark jet by requiring a displaced secondary vertex within the jet.

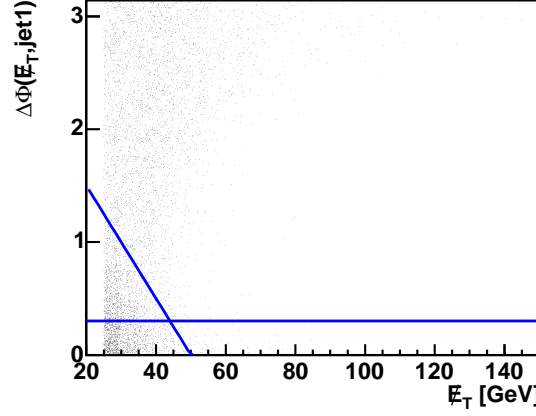


Figure 5.1: The rejection of QCD background events in the forward calorimeter is illustrated. Exemplarily, the cuts applied on the angle between missing transverse energy and the leading jet are demonstrated: $|\Delta\phi| > 0.3$ and $|\Delta\phi| > 2.5 - E_T/(20 \text{ GeV})$.

Missing Transverse Energy

The missing E_T (\vec{E}_T) is defined by

$$\vec{E}_T = - \sum_i E_T^i \hat{n}_i, \quad (5.1)$$

where i denotes the calorimeter tower number with $|\eta| < 3.6$, \hat{n}_i is a unit vector perpendicular to the beam axis which points at the i^{th} calorimeter tower. Additionally, $E_T = |\vec{E}_T|$ is defined. Because this calculation is based on calorimeter towers, \vec{E}_T has to be adjusted for the effect of the jet corrections for all tight and loose jets.

Since muons pass the calorimeters without showering, i.e. as minimum ionizing particle, a correction is applied by adding all transverse momenta of the traversing muons to the sum and by removing the average ionization energy. The corrected E_T is required to be greater than 25 GeV.

Rejection of QCD multi-jet background

To further suppress events in which no real W boson is produced, additional cuts are applied [106]. The cuts are based on the assumption that these events do not produce E_T by nature but due to lost or mismeasured jets. Therefore, one would expect small E_T and small values of the angle $\Delta\phi$ between \vec{E}_T and a jet. For central electrons, the requirements are $|\Delta\phi| > 1.9 - E_T/(20 \text{ GeV})$ for the angle between \vec{E}_T and the jet with the higher E_T and $|\Delta\phi| > 1.8 - E_T/(25 \text{ GeV})$ for the angle between \vec{E}_T and the second jet, respectively. Electrons detected in the plug calorimeter must have $|\Delta\phi| > 0.3$ and $|\Delta\phi| > 2.5 - E_T/(20 \text{ GeV})$ for the angle between \vec{E}_T and the jet with the higher E_T and $|\Delta\phi| > 2.2 - E_T/(20 \text{ GeV})$ for the angle between \vec{E}_T and

the second jet. As illustration of the QCD background rejection, the cuts applied on the angle between \vec{E}_T and the leading jet are exemplarily demonstrated in figure 5.1.

Z-Boson Veto

To remove Z -boson events, events are rejected in which the charged lepton can be paired with any more loosely defined jet or lepton to form an invariant mass consistent with the Z peak, defined as the range from $76 \text{ GeV}/c^2$ to $106 \text{ GeV}/c^2$.

5.3 Data Modeling

Using multivariate methods, it is crucial to model the observed data correctly. Therefore, the complete spectrum of expected processes contributing to the $W + 2$ jets bin has to be modeled. Most of the processes are described using Monte Carlo simulation, while some background processes are derived from data. A complete list of Monte Carlo samples used in this analysis can be found in table B.1.

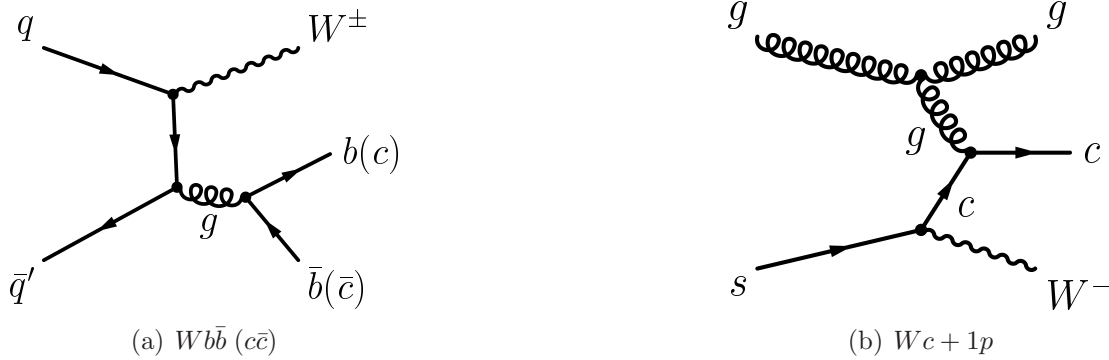


Figure 5.2: Examples of Feynman graphs showing W +heavy flavor background production: (a) illustrates production of a W boson in association to a gluon splitting into a heavy-quark pair. (b) shows the production of a W boson in addition to a c quark and a gluon as extra parton.

The background to single top-quark production in the b -tagged lepton+jets sample is dominated by W -boson production in association with heavy-flavor quarks, illustrated in figure 5.2: $Wb\bar{b}$, $Wc\bar{c}$, and Wc production, called W +heavy flavor in the following. Additional background sources are top-quark pair production being illustrated in figure 1.3, diboson production (including WW , WZ , and ZZ) whereof an example is given in figure 5.3(a), as well as Z -boson production in association with quarks. The latter, exemplarily shown in figure 5.3(b), is also called Z +jets production and contains leptonically decaying Z bosons. W -boson, diboson, and Z -boson events with a light-flavor jet mistakenly tagged as heavy flavor are named mistags¹. A substantial background arises from QCD-induced multijet events, e.g. direct $b\bar{b}$

¹The label “mistag” has to be interpreted in the respective context, since it can denominate both a mistagged jet and an event with a mistagged jet.

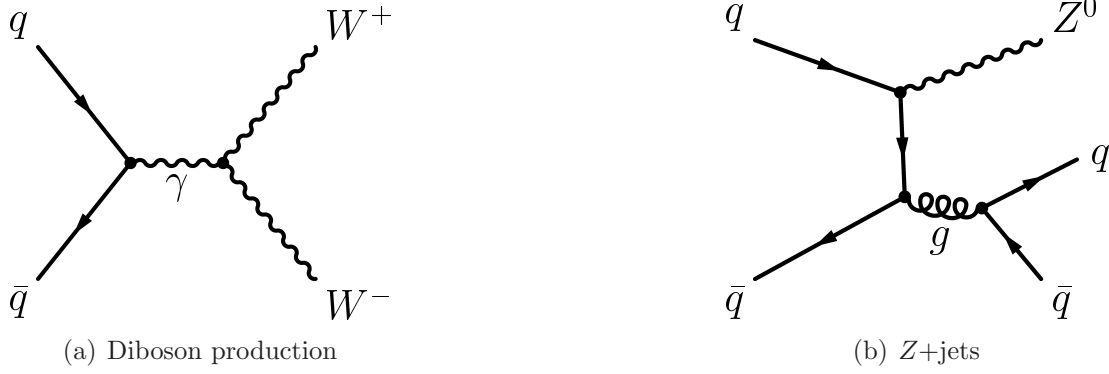


Figure 5.3: Examples of (a) diboson and (b) Z +jets background production. Variations of (a) can produce ZZ and WZ events. For Z +jets production, exemplarily illustrated in (b), only leptonic Z -Boson decays are considered.

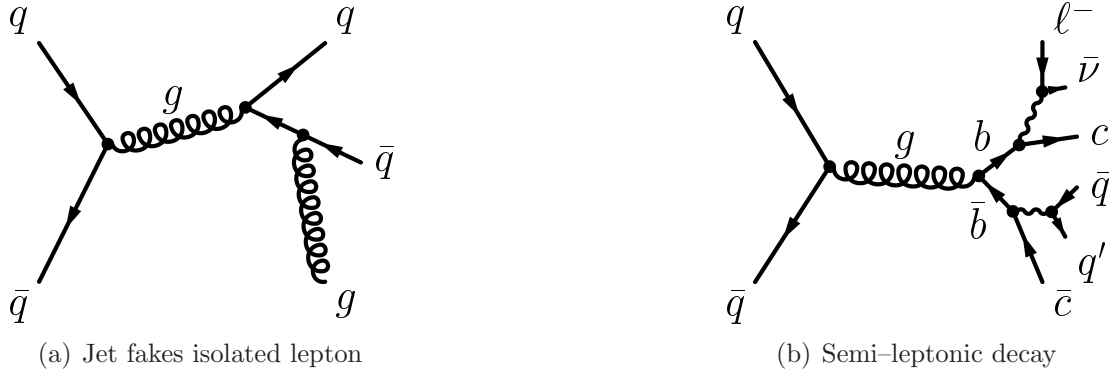


Figure 5.4: Examples of Feynman graphs showing production of QCD backgrounds. In events like (a), a jet can be misidentified as isolated lepton to fake the signature of single-top-quark events. In $b\bar{b}$ events, a semi-leptonically decaying b hadron can lead to the same signature as single-top-quark events if the lepton is spuriously identified as isolated lepton, as illustrated in (b).

production with additional gluon radiation as illustrated in figure 5.4. Those events mimic the signature of W -boson production if a jet (or a lepton from a semileptonic decay) is erroneously identified as an isolated lepton and transverse energy is mis-measured, leading to artificial \cancel{E}_T . Since no on-shell W boson is produced, those events are called non- W events in the following.

5.3.1 Signal MC

Several authors pointed out [107, 38] that the leading order contribution to single top-quark production via t -channel, as modeled in leading order parton shower Monte Carlo programs, does not adequately represent the measured final states.

The leading order process is a $2 \rightarrow 2$ process with a b quark in the initial state as given by figure 1.4(a): $b + u \rightarrow d + t$ or $b + \bar{d} \rightarrow \bar{u} + t$. For antitop-quark production, the charge conjugate processes are implied. As already pointed out in section 1.2.2, the b quark being part of the quark sea is described by a b -quark PDF for the calculation. Since the b quark originally stems from a gluon splitting into a $b\bar{b}$ pair, a \bar{b} quark has to be present in the event. Leading order parton shower programs create this \bar{b} quark through backward evolution following the DGLAP scheme [108, 109, 110]. Thereby, only the soft regime of the transverse momentum distribution of the \bar{b} quark is modeled well, while the high- p_T tail is not estimated adequately. Also, the pseudorapidity spectrum expands too far into the forward region.

One can improve the modeling of single top-quark production via t -channel by producing two samples of simulated events with matrix element generators and applying a parton shower Monte Carlo program to the final-state partons. In this analysis, the matrix element generator MADEVENT, interfaced to the CTEQ5L [31] parameterization of the parton distribution functions (PDF), was used to produce simulated events. Parton showering and hadronization were performed using PYTHIA.

The first sample is the $2 \rightarrow 2$ process $b + q \rightarrow q' + t$ given by the Feynman graph in figure 1.4(a); the second process is a $2 \rightarrow 3$ process with a gluon in the initial state, $g + q \rightarrow q' + t + \bar{b}$, which is shown in figure 1.4(b). In the second process, the \bar{b} quark, called 2nd b quark in the following, is produced directly in the hard scattering described by the matrix element. This sample describes the most important next-to-leading order (NLO) contribution to t -channel production and is therefore suitable to describe the high- p_T tail of the p_T distribution of the 2nd b quark.

However, the two samples, the $2 \rightarrow 2$ process and the $2 \rightarrow 3$ process, have to be matched to form one unified sample of simulated events. This has been realized by a procedure of adjusting the ratio between the $2 \rightarrow 2$ and $2 \rightarrow 3$ processes in such a way that the rate of events with a detectable 2nd b -quark jet, that is $p_T > 15 \text{ GeV}/c$ and $|\eta| < 2.8$, matches the prediction made by ZTOP [38], a program which operates at NLO in perturbation theory. The matched t -channel sample is created in such a manner that the p_T distribution of 2nd b -quarks in the matched t -channel sample consists of $2 \rightarrow 2$ events for transverse momenta below a certain cutoff and of $2 \rightarrow 3$ events for transverse momenta above the cutoff. This cutoff and hence the ratio between the two processes are varied until the rate of detectable 2nd b -quark jets matches the prediction [111]. The final ratio is found to be $R = 1.3$, the cutoff is derived to be $K_T = 9 \text{ GeV}/c$, as displayed in figure 5.5. As a result, all detectable 2nd b quarks ($p_T > 15 \text{ GeV}/c$) in the matched t -channel sample are provided by the $2 \rightarrow 3$ process. Looking at figure 5.5(b) which shows the p_T distribution of the 2nd b quark on a logarithmic scale, it is apparent that the 2nd- b -quark p_T spectrum in the matched t -channel sample is much harder than the one provided by the $2 \rightarrow 2$

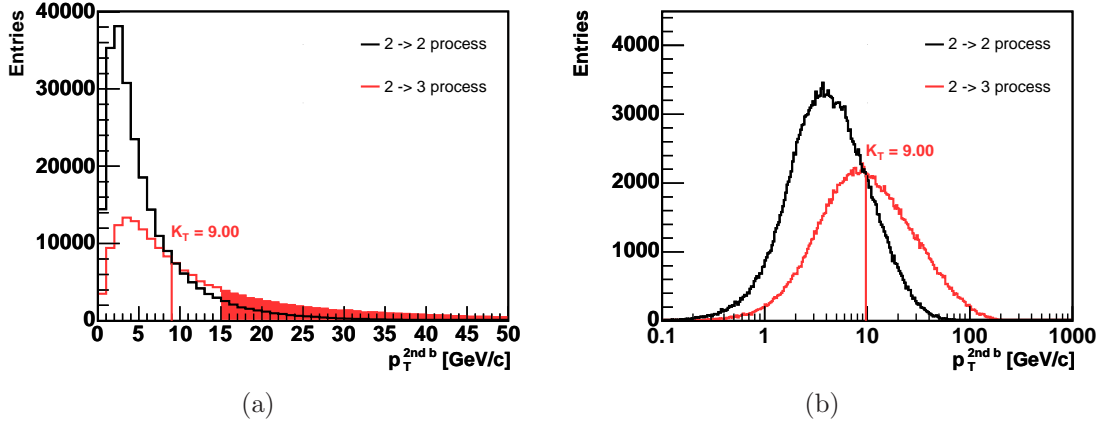


Figure 5.5: Matching of single-top-quark events produced by the $2 \rightarrow 2$ and the $2 \rightarrow 3$ t -channel processes. The p_T distributions of the 2nd b quark in the event are shown, (a) on a linear scale and (b) on a logarithmic scale. The ratio of $2 \rightarrow 2$ to $2 \rightarrow 3$ events is adjusted such that the rate of 2nd b quarks with $p_T > 15$ GeV/ c and $|\eta| < 2.8$ matches the NLO prediction. The fraction of these events is illustrated by the shaded area in (a).

process alone. The tail of the distribution extends beyond 100 GeV/ c , while it would otherwise become distinct around 50 GeV/ c .

Evaluation of Single-Top Monte Carlo Samples

It is important to evaluate the quality of the modeling of single top-quark events in a quantitative fashion. For that purpose, the kinematic distributions of the primary partons extracted from the MADEVENT samples are compared to NLO differential cross sections calculated with ZTOP. In general, very good agreement is found. In particular, it is important to notice that the pseudorapidity distribution of the 2nd b quark is modeled very well even though it was not used for the matching of the two t -channel samples.

The remaining difference between the simulation and the ZTOP NLO calculation can be obtained by reweighting the simulated events in such a way that they match the calculation and by investigating the influence on the acceptance. The weights are derived from comparing six kinematic distributions, namely p_T and η of both the top quark and the two highest E_T jets not originating from the top-quark decay. The correlations between those variables are considered by determination of the covariance matrix using simulated events. In case of single top-quark production via t -channel, one distinguishes between b -quark jets and light-quark jets to account for the specific final state. The first step to estimate a systematic uncertainty is to determine the acceptance of simulated events by counting events passing the selection cuts. The second step is to sum up the weights of accepted events to estimate the influence of the difference between calculation and simulation on the acceptance. In total, a discrepancy of $-1.8\% \pm 0.9\%$ (MC stat.) for t -channel events and $-0.3\% \pm 0.7\%$ (MC stat.) for s -channel events is found, respectively. This indicates

that the estimate of the acceptance using simulated events is slightly higher than the NLO prediction for t -channel events, while excellent agreement for s -channel events is found.

The general conclusion from our studies is that the MADEVENT Monte Carlo events give a very good representation of the single top-quark production process. For t -channel, the influence of NLO corrections is sufficiently considered by taking into account the main NLO effect. The remaining differences are covered by assigning a systematic uncertainty of $\pm 2\%$ or $\pm 1\%$ on the acceptance for t -channel or s -channel events, respectively.

More details on the t -channel matching procedure and the comparison to ZTOP can be found in reference [111].

5.3.2 Modeling of backgrounds

For the modeling of $t\bar{t}$, diboson, and Z +jets production, simulated events generated with PYTHIA are used, while the W +heavy flavor background was simulated using a combination of ALPGEN and HERWIG.

The non- W background is modeled using two different approaches [112]. The central electron and muon models are obtained from central electron trigger data. The events are required to pass all kinematic electron cuts but to fail two of the five non-kinematic cuts². Even though their kinematic properties resemble those of W -like events, those events are non- W -enriched, since the non-kinematic criteria serve primarily to filter out QCD-induced multijet events. For the forward electron sample, such a model is not yet available. For this reason, an additional non- W model is introduced, based on the idea that for a non- W event to pass the selection criteria, a jet has to resemble an electron. Hence, events from jet trigger data are required to have a jet with $E_T > 20$ GeV, $0.05 < E_{\text{HAD}}/E_{\text{EM}} < 0.2$, and at least four reconstructed tracks. The latter makes it unlikely that the event contains a real electron. Once a jet is identified as a fake electron, its charge is assigned randomly, and it is further considered as a tight forward electron.

The event candidates of both approaches have to pass all but the lepton selection criteria and the b -tag requirement to contribute to the corresponding non- W model. Since demanding a tagged jet would cause too low statistics, the b tag in the event has to be faked using taggable³ jets. If, in a given event, only one jet is taggable, this one is considered to be the tagged jet. If there are two taggable jets in the event, the event is duplicated and the first jet is assigned as tagged for one event, and the second jet as tagged for the other event, i.e. both hypothesis are utilized. To avoid a bias towards events with two taggable jets, events with only one taggable jet are duplicated as well.

Due to the usage of the neural-network b -tagger described in section 6, it is ad-

² $Q \cdot \Delta x, |\Delta z|, E_{\text{HAD}}/E_{\text{EM}}, L_{\text{shr}}, \chi_{\text{strip}}^2$

³ $E_T > 10$ GeV, $|\eta| < 2.4$, and $N_{\text{trk}} \geq 2$.

ditionally necessary to assign a hypothesis of what kind of quark flavor the jet is, b , c , or light-quark flavor. Therefore, all events are tripled by taking into account each hypothesis. Each event is weighted by the corresponding flavor fraction. This fraction is estimated by applying the neural-network b -tagger to the $\cancel{E}_T < 15$ GeV sideband of the observed data. In this sideband sample, a flavor composition of 45% b -quark jets, 40% c -quark jets, and 15% light-quark jets is found.

In order to describe events with mistagged light-quark jets, W +light flavor events simulated with ALPGEN and showered with HERWIG are used. Due to the very small fraction of tagged events in this sample, a large amount of this kind of events would be needed in simulation. Thus, the pretag sample is utilized where taggable jets are assigned to be tagged [113]. If an event has two taggable jets, both hypotheses are used for the particular event. Each hypothesis of each event is weighted by the mistag probability of the jet considered as tagged. This probability is defined by the negative tag rate and the correction factor for the mistag asymmetry, both introduced in section 4.3.4.

For both non- W and mistag model, the output of the neural network b tagger is randomly assigned to the jet attributed as tagged. For this purpose template output distributions obtained from jets of simulated events corresponding to the respective flavor are utilized.

5.4 Event Yield and Background Estimate

5.4.1 Expected Event Yield of Single-Top-Quark Events

The number of expected events is given by

$$\hat{\nu} = \sigma \cdot \varepsilon_{\text{evt}} \cdot \mathcal{L}_{\text{int}} \quad (5.2)$$

where σ is the theoretically predicted cross section of the respective process, ε_{evt} is the event detection efficiency, and \mathcal{L}_{int} is the integrated luminosity. The predicted cross sections for single top-quark production via t -channel and s -channel are quoted in section 1.2.2.

The event detection efficiency is estimated by applying the selection cuts to the samples of simulated events. In addition, one has to account for differences between the simulation and the real experimental setup. Since the trigger simulation is not used in this analysis, the event detection efficiency obtained from the Monte Carlo simulation, ε_{mc} is reduced by the trigger efficiency $\varepsilon_{\text{trig}}$. Differences in the identification efficiencies of charged leptons and b quark jets between data and simulation are accounted for by a correction factor, $\varepsilon_{\text{corr}}$ [99]. Here, the scale factor correcting the b tag efficiency, $SF = 0.89 \pm 0.07$ (see section 4.3.4), is valid per tagged heavy-flavor jet. For c -quark jets, the uncertainty on the scale factor SF is doubled. Since the simulation is too optimistic concerning mistagged light jets, the number of expected mistags per sample is estimated using the mistag matrix introduced in section 4.3.4.

The samples of simulated events are produced in such a way that the W boson emerging from the top-quark decay is only allowed to decay into lepton pairs, that is $e\nu_e$, $\mu\nu_\mu$, and $\tau\nu_\tau$. The value of ε_{mc} is therefore multiplied by the branching fraction of W bosons into leptons, $\varepsilon_{\text{BR}} = 0.324$. In total, the event detection efficiency is given by

$$\varepsilon_{\text{evt}} = \varepsilon_{\text{mc}} \cdot \varepsilon_{\text{BR}} \cdot \varepsilon_{\text{corr}} \cdot \varepsilon_{\text{trig}} \quad (5.3)$$

Including all trigger and identification efficiencies, $\varepsilon_{\text{evt}}(t\text{-channel}) = (1.2 \pm 0.1)\%$ and $\varepsilon_{\text{evt}}(s\text{-channel}) = (1.8 \pm 0.1)\%$ are found. The expected number of single-top-quark events is 22.4 ± 1.8 for production via t -channel and 15.4 ± 1.0 for s -channel, respectively; the quoted errors include only the uncertainty on ε_{evt} .

5.4.2 Background Estimate

To estimate the number of expected background events, two different approaches are realized. The expectations of theoretically understood processes are determined using simulated events [99]. Processes with larger theoretical uncertainties are treated by a data-based method [114] utilizing so-called “pretag” data, i.e. the data sample with all cuts applied except the b -tag requirement. This pretag sample is corrected for the expected number of events estimated by the Monte-Carlo-based method. Afterwards, the number of pretag non- W events is evaluated and subtracted to enable the determination of the expected pretag W +heavy flavor events which are scaled by the tagging efficiency to result in the expected number of events in the b -tagged sample. Using the remaining number of pretag events, the expected number of mistags is estimated. In the final step, the number of tagged non- W events is determined. In the following sections, this procedure is outlined more precisely.

Monte-Carlo-Based Backgrounds

Contributions of $t\bar{t}$, diboson, and Z +jets production to the b -tagged lepton+jets sample are derived from samples of simulated events. The corresponding event detection efficiencies ε_{evt} are calculated as described in section 5.4.1 in case of single top-quark production. Using (5.2), the number of expected events of a given process is estimated by multiplying the event detection efficiency by the integrated luminosity and by the appropriate theoretically predicted cross section. Therefore, it is essential that the given physical process is theoretically well understood, i.e. the kinematics are well described through simulated events and the cross section is well known.

Events from top-quark pair production are expected to be present in the selected data sample due to limited detector acceptance. In principle, $t\bar{t}$ events can be distinguished by the decay mode of both W bosons coming from the top-quark and antitop-quark decays. If both W bosons decay leptonically the event belongs to the dilepton category. If one of those leptons is not detected, the event can possibly pass the event selection cuts. In case of one W boson decaying hadronically, several

lost or misidentified jets can induce such a non-dilepton event being accepted by the single-top-quark event-selection. Even though a veto on dilepton events is applied, the rejection is not perfect. To compensate for a higher dilepton veto efficiency in simulated events induced by deviation in lepton identification and reconstruction efficiencies, the dilepton ε_{evt} is corrected using measured dilepton events from Z -boson decays. The normalization of $t\bar{t}$ production is based on the theoretical NLO cross section predicted by Bonciani *et al.* [115, 34]. The differences to the cross sections calculated by Berger *et al.* [116] and Kidonakis *et al.* [117] are incorporated, as well as an additional systematic uncertainty due to a variation of the top-quark mass, leading to a cross section of $\sigma_{t\bar{t}} = (6.70 \pm 1.32)$ pb.

To obtain the expected number of diboson events in our selected data sample, we use the theoretical cross sections predicted for a center of mass energy of $\sqrt{s} = 1.80$ TeV and $\sqrt{s} = 2.00$ TeV [118], respectively. The rescaling to $\sqrt{s} = 1.96$ TeV is realized by taking the mean of a linear and quadratic interpolation, leading to $\sigma_{WW} = (13.30 \pm 0.40)$ pb, $\sigma_{WZ} = (3.96 \pm 0.12)$ pb, and $\sigma_{ZZ} = (1.57 \pm 0.05)$ pb.

The expected number of Z +heavy flavor events is estimated in an analogous way; events of Z -boson production in association with light-flavor quarks and coincident decay of $Z \rightarrow e^+e^-$ or $Z \rightarrow \mu^+\mu^-$ contribute to the background category of mistags and are hence estimated differently. The normalization for Z +jets production is determined using $\sigma_Z \cdot \text{BR}(Z \rightarrow l^+l^-) = (336.0 \pm 8.0)$ pb, where $\text{BR}(Z \rightarrow l^+l^-)$ is the branching ratio of events with leptonically decaying Z bosons.

Estimation of the Pretag Non- W Event Rate

The estimation of the expected number of non- W events makes use of the pretag sideband of the lepton trigger data. To obtain the fraction of non- W events in the pretag sample, the selection requirement of \cancel{E}_T is additionally omitted. The \cancel{E}_T distribution of a W +jets template obtained from ALPGEN simulated events and a non- W template is fitted to the \cancel{E}_T spectrum of the measured pretag data.

For the muon and forward electron trigger data, the non- W template is built out of corresponding non-isolated sideband lepton data. By inverting the lepton calorimeter isolation requirement, the on-shell W -boson content is intensely reduced.

For the central electron sample, the non- W template is obtained from electron trigger data by requiring those events to pass all kinematic selection criteria but to fail two of five lepton identification criteria as described in section 5.3.2.

The estimate of the tagged non- W event yield is performed in a similar way after the rates of all other background processes are specified. This is mandatory to establish a \cancel{E}_T distribution according to the full background prediction.

W+Heavy Flavor Contributions

W +heavy flavor production contributes significantly to the b -tagged lepton+jets sample. Several Monte Carlo generators are capable of performing matrix element calculations for W +jets, but merely performing calculations at leading-order in perturbation theory. As a result, the overall normalization of these calculations has a large theoretical uncertainty, even though the relative contributions of the important diagrams for $Wb\bar{b}$, $Wc\bar{c}$, and Wc production are well-defined. For this reason, the number of pretag W +heavy flavor events, N_{HF} , is estimated by multiplying the number of W +jets events in the pretag data sample, $N_{W+\text{jets}}$, by the relative W +heavy flavor fractions, f_{HF} . The heavy flavor fractions were obtained from samples of ALPGEN simulated events using corrections derived from observed multijet events [119]. Hence, the number of pretag W +heavy flavor events is given by

$$N_{\text{HF}} = f_{\text{HF}} \cdot N_{W+\text{jets}} = f_{\text{HF}} \cdot (N \cdot (1 - f_{\text{non-}W}) - N_{\text{HFsim}}) \quad (5.4)$$

where N is the number of observed events in the pretag data set, $f_{\text{non-}W}$ is the fraction of non- W events in the pretag sample, and N_{HFsim} is the expected number of pretag heavy-flavor events derived by the described technique using simulated events, i.e. $t\bar{t}$, diboson, and Z +jets events. The number of tagged W +heavy flavor events $\hat{\nu}_{\text{HF}}$ is predicted by applying tagging efficiencies ε_{tag} , derived from simulated events of $Wb\bar{b}$, $Wc\bar{c}$, and Wc production.

$$\hat{\nu}_{\text{HF}} = N_{\text{HF}} \cdot \varepsilon_{\text{tag}} \quad (5.5)$$

Rates of Events with Mistagged Jets

The number of expected mistagged events is estimated by making use of the mistag matrix described in section 4.3.4. The per-jet mistag rate is used to estimate the number of negatively tagged events, $N_{\text{tag-}}$, in the observed pretag sample. To correct for the mistag asymmetry, the correction factor γ has to be applied. Additionally, to obtain the number of expected mistagged events $\hat{\nu}_{\text{mistags}}$, one has to further correct for non- W and heavy-flavor contributions to the pretag sideband:

$$\hat{\nu}_{\text{mistags}} = N_{\text{tag-}} \cdot \gamma \cdot \frac{N \cdot (1 - f_{\text{non-}W}) - N_{\text{HFsim}} - N_{\text{HF}}}{N} \quad (5.6)$$

Estimation of the b -Tagged Non- W Event Rate

To estimate the number of expected tagged non- W events, likelihood template fits are performed to the \cancel{E}_{T} spectrum of the tagged data sample. As a tagged W +jets template, rather than a sample of simulated W +jets events, the \cancel{E}_{T} distribution of the full background prediction is used with proper normalization and all background processes included except the non- W fraction. To obtain the non- W template in the tagged sample, the pretag non- W template is weighted by an \cancel{E}_{T} -dependent tagging transfer function extracted from central electron trigger data.

The predicted numbers of background events, as well as the number of expected single-top-quark events, are given in table 5.2.

Process	Number of Events
$t\bar{t}$ dilepton	19.9 ± 4.6
$t\bar{t}$ non-dilepton	38.5 ± 8.9
total $t\bar{t}$	58.4 ± 13.5
$Wb\bar{b}$	170.9 ± 50.7
$Wc\bar{c}$	63.5 ± 19.9
Wc	68.6 ± 19.0
total W +heavy flavor	303.0 ± 89.6
Mistags	136.1 ± 19.7
Non- W	26.2 ± 15.9
WW	5.5 ± 1.0
WZ	8.0 ± 0.8
ZZ	0.3 ± 0.1
total Diboson	13.7 ± 1.9
$Z \rightarrow ee$	1.2 ± 0.5
$Z \rightarrow \mu\mu$	5.3 ± 1.6
$Z \rightarrow \tau\tau$	5.5 ± 2.4
total Z +jets	11.9 ± 4.4
total background	549.3 ± 95.2
t -channel	22.4 ± 3.6
s -channel	15.4 ± 2.2
total single-top	37.8 ± 5.9
total prediction	587.1 ± 96.6
observation	644

Table 5.2: Summary of predicted numbers of signal and background events in the selected data sample with all systematic uncertainties included (see section 7.3).

Chapter 6

A Neural–Network b –Tagger

In the selection of single–top–quark events, the identification of b –quark jets is realized by requiring a reconstructed secondary vertex which mainly exploits the long lifetime of b hadrons ($\tau \approx 1.6$ ps). Due to the non–zero lifetime of c hadrons and spurious reconstruction of secondary vertices in light–flavored jets, 48% of the expected backgrounds to the search of single top–quark production do not contain any b quarks, as demonstrated in table 5.2. Selected events without any b quarks can be further rejected by extending the vertex requirement by the inclusion of other characteristics of b hadrons. Appropriate properties are their large mass of about $5 \text{ GeV}/c^2$, the high decay multiplicity, and the probability of semileptonic decays.

In this thesis, a neural network was developed (making use of the NeuroBayes[®] package described in section 3.1), which was trained with a variety of variables exploiting those quantities in order to identify secondary vertices related to the decay of b hadrons [120]. This approach has the advantage that all given information and all correlations are optimally combined to one continuous variable. For the training, tagged jets of simulated single–top–quark events and the main backgrounds, viz production of W +jets and $t\bar{t}$, are used in a mixture according to the expectation given by table 5.2. As signal, all processes containing b quarks are used, namely production of single top quarks, $t\bar{t}$, and $Wb\bar{b}$, while processes containing charm and light flavors, that is $Wc\bar{c}$, Wc , and mistags, act as background. Out of those processes, tagged jets are used for the training of the neural network, taking into account the knowledge about a particle’s origin which is a given information for simulated events. This is particularly important for the selection of mistagged light–quark jets. Since those have to be extracted from W +jets events comprising all quark flavors, only tagged jets in events without any heavy–flavored quarks or hadrons are utilized as mistags for the training. In case of the b –quark (c –quark) samples, only tagged jets are used with at least one track coming from the decay of a b hadron (c hadron).

For the training of the neural–network b –tagger, the single–top–quark event selection described in section 5.2 is omitted, since those requirements reduce the number of available jets in simulated events. Nevertheless, only events with exactly one identified tight lepton are considered. This is necessary to prevent confusing the unidentified lepton from the W –boson decay with a light–flavored jet.

6.1 Discriminating Variables

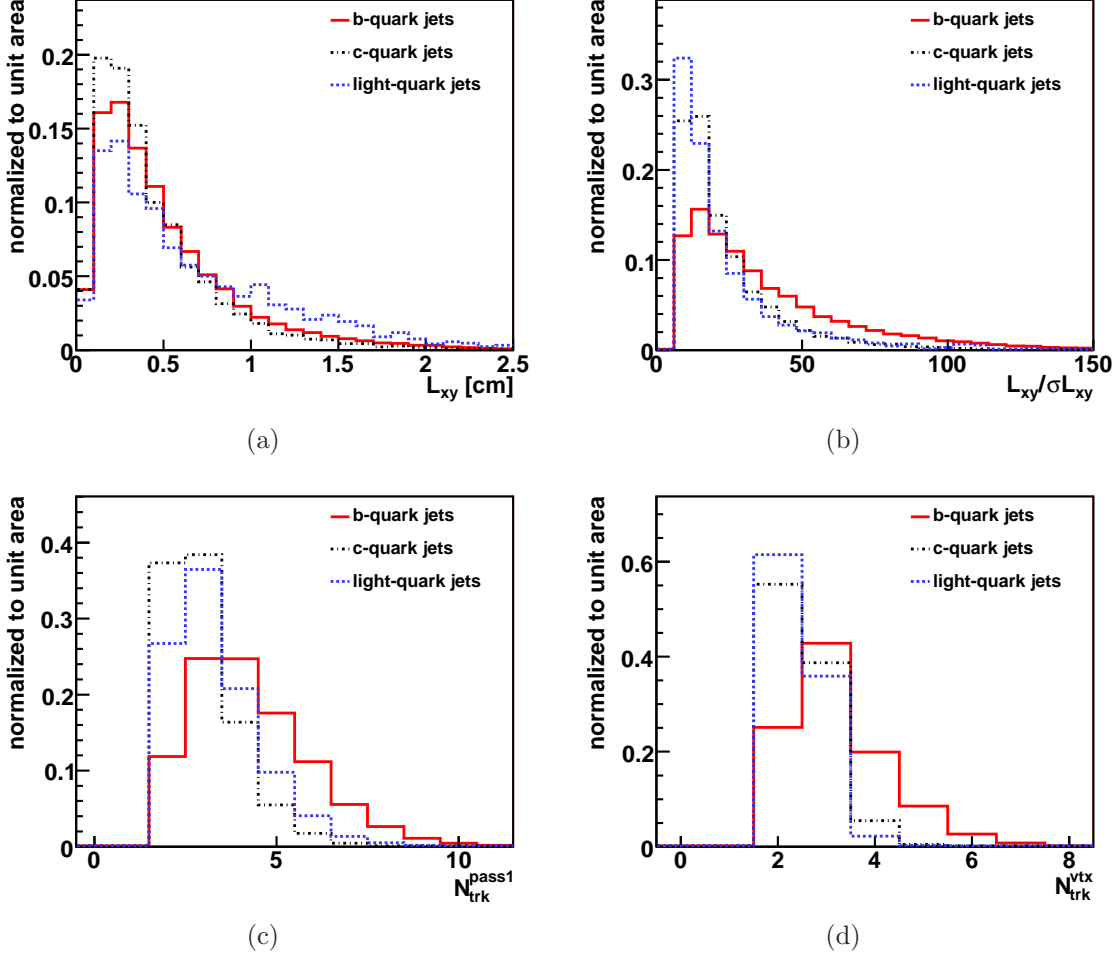


Figure 6.1: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: (a) the 2-dimensional decay length of the secondary vertex, (b) the significance of the 2-dimensional decay length, (c) the number of displaced tracks fulfilling the quality requirements for pass 1 of the vertex fit, and (d) the number of tracks fitted to the secondary vertex.

A variety of variables is suitable to exploit lifetime, mass, and decay multiplicity of b hadrons. Many of them are related to the reconstructed secondary vertex, some are reflected by the properties of the tracks in the tagged jet. To account for the probability of semileptonic b -hadron decays, the number and momenta of leptons in the jet are useful quantities.

For instance, the decay length of the reconstructed secondary vertex is related to the lifetime of the decaying particle. Even though figure 6.1(a) illustrates that spuriously reconstructed vertices can have a large transverse distance L_{xy} to the primary interaction point, the significance of the transverse decay length ($L_{xy}/\sigma_{L_{xy}}$) is small, as clarified by figure 6.1(b): as expected, vertices induced by the decay of a b hadron are more significantly dislodged from the primary interaction point than

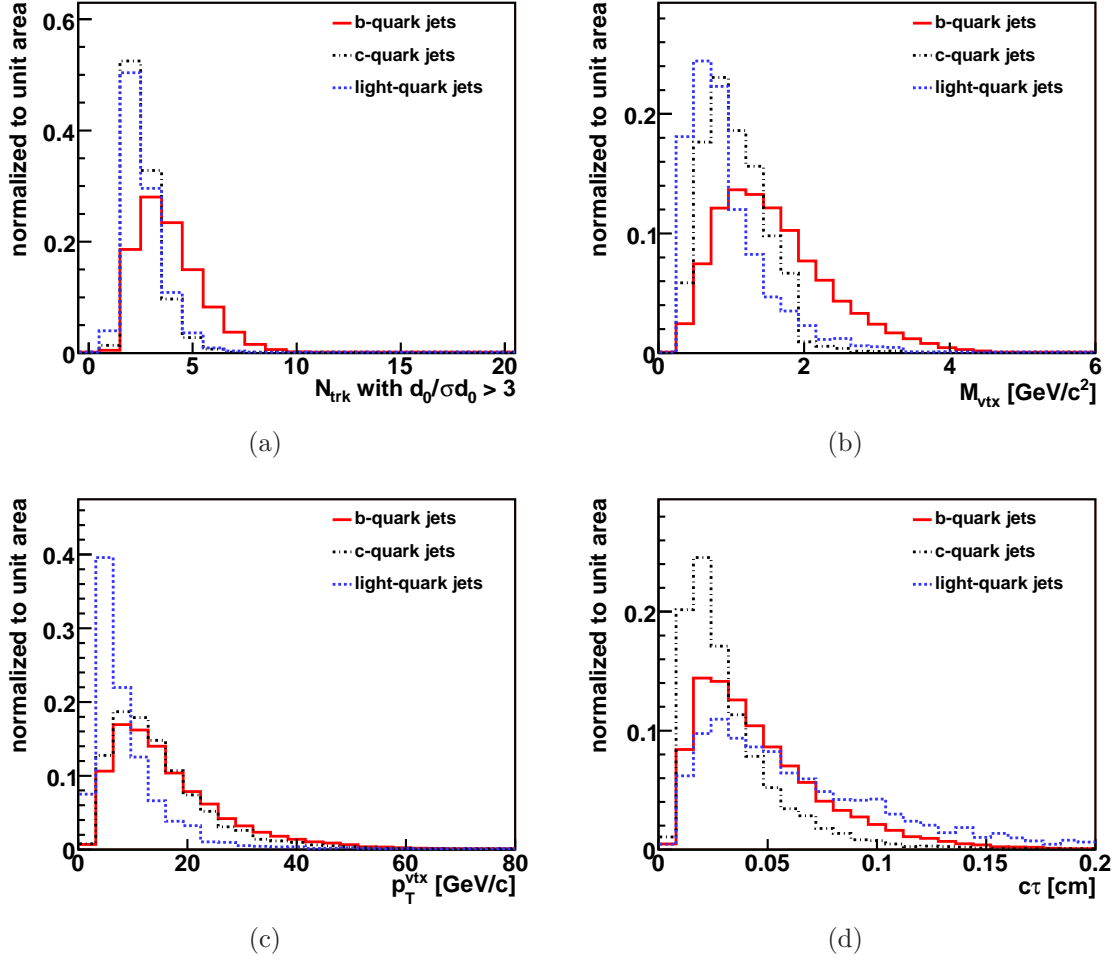


Figure 6.2: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: (a) the number of tracks in the tagged jet with an impact parameter significance larger than 3σ , (b) the invariant mass of the vertex, (c) the transverse momentum at the secondary vertex, and (d) the pseudo-lifetime of the secondary vertex.

vertices originating from decaying c hadrons or falsely reconstructed vertices.

The high decay multiplicity of b hadrons is for example reflected by the number of tracks N_{trk} per jet, fulfilling minimum-quality requirements; another sensitive quantity is the number of displaced tracks $N_{\text{trk}}^{\text{pass1}}$ qualified for the use in pass 1 of the vertex fit described in section 4.3.4, as displayed in figure 6.1(c). Figure 6.1(d) illustrates that the number of tracks associated to the vertex, $N_{\text{trk}}^{\text{vtx}}$, contains information about the origin of the vertex. Additionally, many tracks with a certain significance of the impact parameter, d_0/σ_{d_0} , are a sign of the presence of a long-lived particle in the jet. This is exemplarily shown in figure 6.2(a) for the number of tracks with $d_0/\sigma_{d_0} > 3$.

The large b -hadron mass leads to higher invariant vertex masses M_{vtx} if the vertex stems from the decay of a b -hadron as demonstrated in figure 6.2(b). Since the combination of tracks is arbitrary in case of spurious vertices, the invariant mass is

usually small. In case of c -hadron decays, the mass distribution shows a cutoff at around $2 \text{ GeV}/c^2$, induced by the c -quark mass.

Using the transverse momentum of the tracks associated to the secondary vertex, p_T^{vtx} , displayed in figure 6.2(c), one can compute the vertex pseudo-lifetime $c\tau = L_{xy} \cdot M_{\text{vtx}}/p_T^{\text{vtx}}$, depicted in figure 6.2(d), mirroring the fact that falsely reconstructed vertices can have large distances to the primary interaction point, as mentioned above.

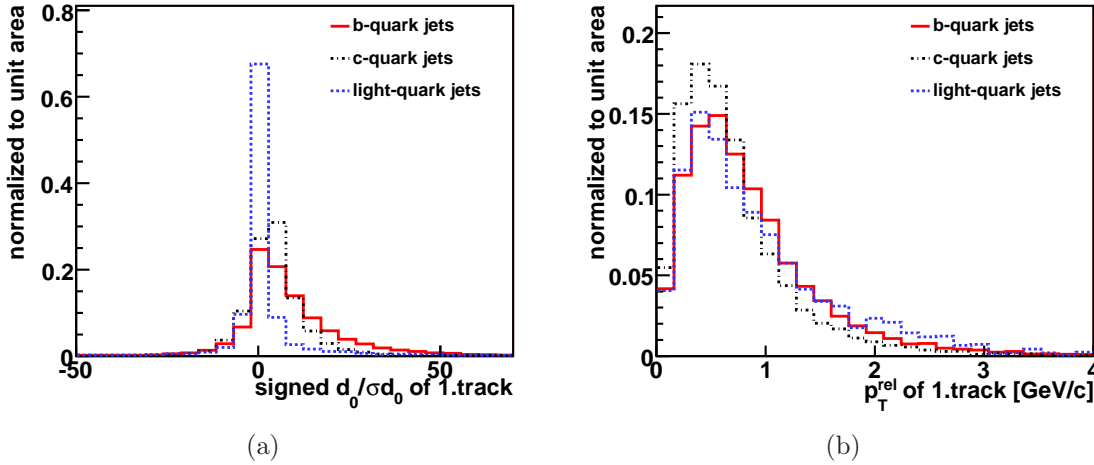


Figure 6.3: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: (a) the lifetime–signed impact parameter significance and (b) the transverse momentum p_T^{rel} with respect to the jet axis of the most energetic track in the tagged jet.

In addition to vertex properties, attributes of the tracks in the jet are suitable to discriminate jets containing a b hadron. For track information, the three most energetic tracks in each tagged jet are considered, named “1.track”, “2.track”, and “3.track” in the following. Both their impact parameters d_0 and their lifetime–signed impact parameters, $(\epsilon^{zik} \cdot p_{\text{jet}}^i \cdot p_{\text{track}}^k)/(|\epsilon^{zik} \cdot p_{\text{jet}}^i \cdot p_{\text{track}}^k|) \cdot d_0$, have been examined to account for tracks not originating from the primary interaction point; the corresponding significances are taken into account as well. Since the impact parameters of the tracks mirror the lifetime of the mother particle, the displacement from the primary interaction point is larger in case of tracks originating from the decay of a b hadron than for other tracks. As illustration, the significance of the lifetime–signed impact parameter of the most energetic track in the jet, i.e. the 1.track, is shown in figure 6.3(a). Other quantities reflecting the decay kinematics of b hadrons are the momenta p of the tracks in the jet, as well as their transverse momenta p_T^{rel} and rapidities y^{rel} with respect to the jet axis. Exemplarily, figure 6.3(b) depicts p_T^{rel} of the 1.track.

As already described in section 4.3.4, the secondary–vertex fitting procedure consists of two passes. The algorithm initially tries to form a vertex consisting of at least three tracks. If this first pass fails, a second attempt is made using tracks fulfilling more stringent criteria. In this second pass, at least two tracks must be attached

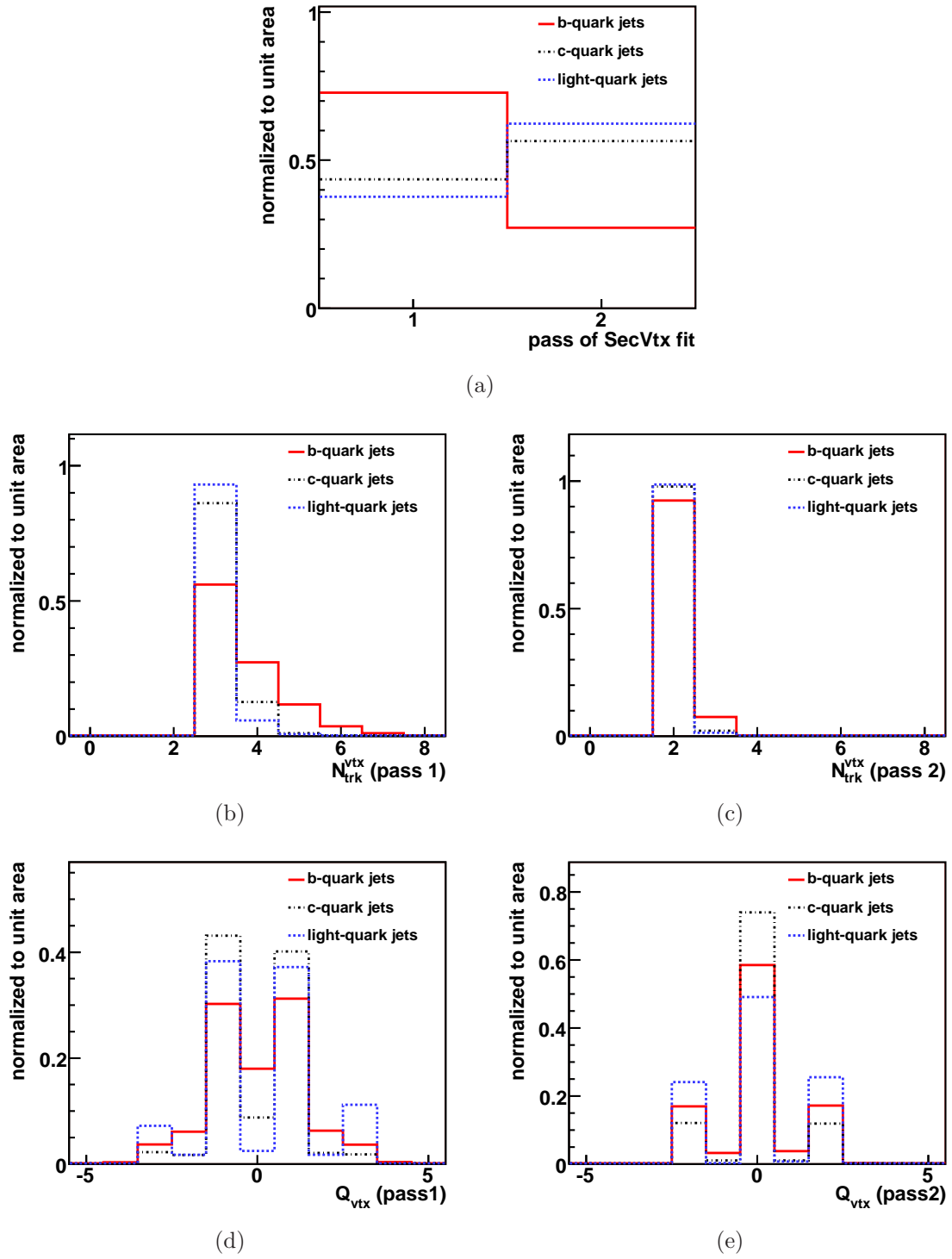


Figure 6.4: Some variables strongly depend on the pass of the secondary vertex fit whose distribution is displayed in (a): Examples are the number of tracks associated to the vertex ($N_{\text{trk}}^{\text{vtx}}$) for (b) pass-1 vertices and (c) pass-2 vertices and the charge at the secondary vertex for (d) pass-1 vertices and (e) pass-2 vertices.

to the vertex. Figure 6.4(a) demonstrates that, due to the decay multiplicity, more than 70% of the reconstructed b –decay vertices are already built in pass 1, whereas more than half of the reconstructed c –decay vertices or falsely reconstructed vertices can only be fitted in the second pass. Due to construction, $N_{\text{trk}}^{\text{vtx}}$ differs for the two passes: while b –hadron–decay vertices can contain up to six tracks if they have been reconstructed in pass 1 (see figure 6.4(b)), they mostly contain two or three tracks if they have been fitted in pass 2 (see figure 6.4(c)). The track multiplicity is reflected by the charge at the secondary vertex Q_{vtx} , as displayed in figures 6.4(d) and 6.4(e): to obtain an odd value for Q_{vtx} , an odd number of tracks is necessary, while an even number of tracks results in an even value for Q_{vtx} .

6.1.1 Input Variables

As already explained, the distributions of many variables depend on the pass of the vertex fit. Even though the network is able to learn those correlations by itself, the training procedure was facilitated by explicitly incorporating those connections. This is realized by splitting up the relevant variables into two separate ones, whereof one is only filled in case of pass 1, the other one for pass 2. In case of pass 2, the value of the pass–1 variable is assigned to a δ function (see section 3.1.2) and vice versa.

In the first approach, 50 variables were investigated whose distributions can be found in appendix C.1 if not already presented before. Variables containing only few additional information are discarded by the automatic variable selection being performed during the preprocessing as described in section 3.1.3. In case of the neural–network b –tagger, the selection was restrained to variables contributing a significance of at least 3σ . The complete list of the 25 variables eventually utilized to train the neural–network b –tagger (ordered by significance) can be found in table 6.1; the list of discarded variables is given in table 6.2.

6.2 Result of the Training

The output distributions of b –quark jets, c –quark jets, and light–quark jets with a reconstructed secondary vertex are shown in figure 6.5(a). The separation between b –hadron–decay vertices and vertices not related to the decay of a b hadron is very good. Additionally, the neural network is able to distinguish between c –decay vertices and falsely reconstructed ones. Since c hadrons have a non–zero lifetime, the network classifies tagged c –quark jets as less background–like than mistags.

Figure 6.6 demonstrates that the pass of the vertex fit is reflected by the output distributions. Since pass–1 vertices have to consist of at least three displaced tracks and pass–2 vertices are only built if the first pass already failed, one can consider pass–1 vertices as being of higher quality. Due to their high decay multiplicity, b –quark jets tend to have a reconstructed vertex fitted in pass 1, while pass–2 vertices

Rank	Variable	Relative Significance (in σ)
1	N_{trk} with $d_0/\sigma d_0 > 3$	80.5
2	signed $d_0/\sigma d_0$ of 1.track	42.3
3	signed $d_0/\sigma d_0$ of 2.track	27.1
4	signed $d_0/\sigma d_0$ of 3.track	22.7
5	M_{vtx} (pass 2)	21.1
6	M_{vtx} (pass 1)	20.6
7	pass of SecVtx fit	19.7
8	$p_{\text{T}}^{\text{rel}}$ of muon	14.5
9	vertex $c\tau$	13.3
10	$L_{xy}/\sigma_{L_{xy}}$	2.7
11	$L_{xyz}/\sigma_{L_{xyz}}$	10.4
12	$p_{\text{T}}^{\text{vtx}}$ (pass 1)	9.3
13	jet E_{T}	10.1
14	$N_{\text{trk}}^{\text{vtx}}$ (pass 1)	9.7
15	y^{rel} of 1.track	6.5
16	N_{trk}	6.1
17	L_{xy}	6.0
18	$p_{\text{T}}^{\text{rel}}$ of electron	5.8
19	N_{trk} with $d_0/\sigma d_0 > 1$	5.2
20	$N_{\text{trk}}^{\text{vtx}}$ (pass 2)	4.0
21	χ^2 per DOF of SecVtx fit (pass 2)	4.2
22	χ^2 per DOF of SecVtx fit (pass 1)	3.9
23	p of 1.track	3.6
24	$p_{\text{T}}^{\text{rel}}$ of 3.track	3.3
25	jet η	3.2

Table 6.1: Set of discriminating variables used to train the neural-network b -tagger, sorted by relevance. The quoted relative significances are determined as described in section 3.1.3, i.e. calculated in terms of reduced matrices. These variables passed the requirement of the significance being larger than 3σ . The variables which did not survive the cut on the significance are listed in table 6.2.

are only reconstructed in b -quark jets if some of the tracks of the decay products are missing. The fact that pass-2 vertices are of minor quality is distinctly recognized by the neural-network b -tagger. Figure 6.6(a) illustrates that pass-1 vertices in b -quark jets are clearly identified as signal-like, while figure 6.6(b) indicates that pass-2 vertices have a broader output distribution not prominently peaking in the signal region. In contrast, tagged c -quark and light-quark jets are identified as more background-like if the vertex was reconstructed in pass 2, whereas pass-1 vertices in these kinds of events feature a flatter distribution.

The performance of the neural network in the training sample is illustrated in figure 6.5(b) showing the signal purity $N_{\text{sig}}^{\text{cut}}/(N_{\text{sig}}^{\text{cut}} + N_{\text{bkg}}^{\text{cut}})$ in dependence of the signal efficiency $N_{\text{sig}}^{\text{cut}}/N_{\text{sig}}$, where N_{sig} is the initial number of b -quark jets in the training sample and $N_{\text{sig}}^{\text{cut}}$ ($N_{\text{bkg}}^{\text{cut}}$) the number of b -quark jets (non- b -quark jets) passing a

Rank	Variable	Relative Significance (in σ)
26	y^{rel} of 2.track	2.7
27	d_0 of 1.track	2.4
28	signed d_0 of 3.track	2.3
29	d_0 of 3.track	2.4
30	y^{rel} of 3.track	2.1
31	signed d_0 of 2.track	1.7
32	d_0 of 2.track	2.0
33	signed d_0 of 1.track	1.7
34	p_T^{rel} of 1.track	1.8
35	p of 3.track	1.4
36	electron E_T	1.3
37	# electrons per jet	1.2
38	N_{trk} with $d_0/\sigma d_0 > 2$	0.3
39	$N_{\text{trk}}^{\text{pass1}}$	1.1
40	$d_0/\sigma d_0$ of 1.track	1.1
41	p_T^{vtx} (pass 2)	1.0
42	p of 2.track	0.9
43	$d_0/\sigma d_0$ of 3.track	0.8
44	p_T^{rel} of 2.track	0.2
45	muon p_T	0.3
46	$d_0/\sigma d_0$ of 2.track	0.2
47	Q_{vtx} (pass 2)	0.1
48	# muons per jet	0.1
49	L_{xyz}	0.1
50	Q_{vtx} (pass 1)	0.0

Table 6.2: Set of discarded variables sorted by relevance. The quoted relative significances are determined as described in section 3.1.3, i.e. calculated in terms of reduced matrices. These variables did not survive the cut on the significance $> 3\sigma$.

certain cut on the neural network output. A signal efficiency equal to one corresponds to simply omitting the cut on the neural network output. Hence, the signal purity at efficiency equal to one is in accordance to the signal ratio in the training sample of about 50%. The signal purity can be improved by cutting on the network output, which provides the possibility to obtain a purity of more than 95% at the expense of losing about 50% of the signal events. As further illustration, table 6.3 shows the efficiencies for an arbitrary cut on the output at -0.3 , demonstrating that 90% of the tagged b –quark jets are kept, while 44% of the tagged c –quark jets and 73% of the mistags are rejected.

Besides using the network in a cut–based manner, it is possible to exploit the different continuous output distributions without losing events in addition to the secondary–vertex tagging efficiency, e.g. in a multivariate analysis. This is realized in the search for single top–quark production as described in section 6.4.1.

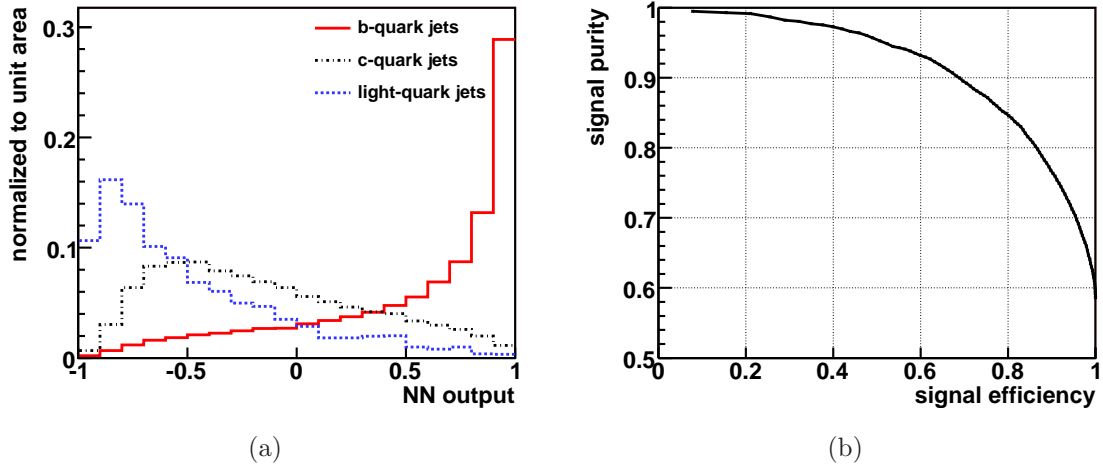


Figure 6.5: (a) The output of the neural-network b -tagger for b -quark jets, c -quark jets, and light-quark jets with a reconstructed secondary vertex and (b) the signal purity in the training sample in dependence of the signal efficiency.

flavor	$N_{\text{events}}^{\text{output} > -0.3} / N_{\text{events}}$
b -quark jets	90%
c -quark jets	56%
light-quark jets	27%

Table 6.3: To demonstrate the performance of the neural-network b -tagger, the efficiency of a cut at output > -0.3 is shown for b -quark jets, c -quark jets, and mistags. The numbers were obtained by using the distributions in figure 6.5(a)

6.3 Comparison of Observed and Simulated Events

As the neural-network b -tagger is trained using simulated events, it is essential to verify that the simulated events describe the data correctly. A decisive point are the shapes of the input-variable distributions whose simulated distributions are compared to data in section 6.3.2. Furthermore, it is important to scrutinize the output distributions of both b signal and non- b background, as described in sections 6.3.3 and 6.3.4.

6.3.1 Used Samples

In principle, one would like to select jets in observed events corresponding to those utilized for the training of the neural network. In case of a neural-network b -tagger, this is not possible, since the flavor of a jet is crucial to be known. Hence, it is necessary to exploit special kinds of data to obtain samples which are enhanced in specific quark flavors.

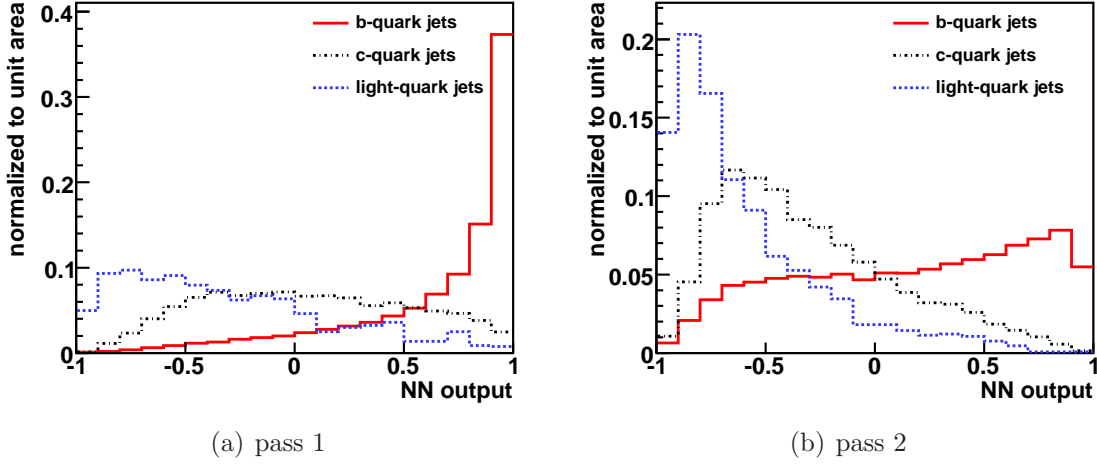


Figure 6.6: The output of the neural-network b -tagger for b -quark jets, c -quark jets, and light-quark jets with a secondary vertex reconstructed in (a) pass 1 and (b) pass 2.

Heavy-Flavor Enhanced Samples

Inclusive Electron Data For the comparison of b -quark jets in data and simulation, b -quark pair production and semileptonic b -hadron decays are exploited to select heavy-flavor enriched samples. Hence, data triggered with the inclusive electron trigger, `blpc0d`, are utilized. This trigger requires a COT track with $p_T > 8$ GeV/ c to match an energy cluster with $E_T > 8$ GeV. The shower profile of this cluster has to be consistent with the expectation obtained by test-beam measurements. The data sample has been stripped for an 8 GeV electron and two 10 GeV jets.

MC To model the inclusive electron data sample, generic $2 \rightarrow 2$ scattering QCD events (generated with HERWIG) with a minimum outgoing transverse momentum of 15 GeV/ c (20 GeV/ c) are used where all flavors are created. A filter requiring a 7 GeV/ c (8 GeV/ c) electron in the generator parton list discards most light-flavor events. In these simulated samples, every event contains a lepton originating from semileptonic decays whereas in data, the required lepton is possibly produced by material interactions.

To emulate the trigger efficiency of measured data, simulated events are accepted or rejected according to a parameterization of the 8 GeV L1 lepton trigger [121].

$$freq\left(\left(\sqrt{E_T} - \sqrt{E_{T0}}\right)/(2\sigma_E)\right) \times freq\left(\left(p_T - p_{T0}\right)/(2\sigma_p p_T p_{T0})\right) \times \left(1 - e^{(2.0 - E_T)/(2\sigma_{E2})}\right),$$

with $E_{T0} = 7.93$, $\sigma_E = 0.096$, $\sigma_{E2} = 1.10$, $p_{T0} = 7.56$, $\sigma_p = 0.0139$ (base unit is GeV), and $freq(x)$ the normal frequency function

$$freq(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

Mistag Samples

Jet–Trigger Data For the comparison of mistagged jets in data and simulation, events selected by generic jet triggers are utilized. Those triggers require a jet, reconstructed with a cone of $R = 0.7$, having a transverse energy of $E_T > 20$ GeV (gjt10d) or $E_T > 50$ GeV (gjt20d), respectively. In those events, negatively tagged jets are utilized.

MC As a model of the above described data, dijet events generated with PYTHIA and HERWIG, having a minimum p_T of 18 GeV/ c and 40 GeV/ c , respectively, are used.

6.3.2 Modeling of Input Variables

As already mentioned, the distributions of the input variables in data and simulation have to be compared. For this purpose, the inclusive lepton samples described before are utilized. Exemplarily, some comparisons are shown in figures 6.7, 6.8, and 6.9; all other comparisons can be found in appendix C.2.

Overall, good agreement between data and simulation is found. Some minor differences are observed using the inclusive electron control samples, caused by the fact that the electron in data can be induced by material interactions. This fact can be compensated by demanding minimum–quality requirements for the electron track. E.g. in case of the comparison of the transverse momentum at the secondary vertex, illustrated in figure 6.8(c), the electron track is required to be attached to the secondary vertex. Nevertheless, this difference between simulation and data is specific to the stripped inclusive lepton samples and not expected in genuine b –quark jets.

6.3.3 Output Distributions of b –Quark Jets

To quantify differences between the output distributions of observed and simulated b –signal events, both b –quark pair production and semileptonic b –hadron decays are taken into account to select b –quark enhanced events. This is realized by requiring an energetic electron within a jet and an additional jet opposite in ϕ .

The electron must have $E_T > 9.0$ GeV and a matched track with $E_T > 8.0$ GeV/ c . Furthermore, the ratio between hadronic and electromagnetic energy, $E_{\text{HAD}}/E_{\text{EM}}$, is required to be less than 0.05; the ratio of clustered energy to track momentum, E/p , has to be $0.5 < E/p < 2.0$. Additionally, the shower profile of the cluster has to be consistent with the expectation obtained by measurements with test–beam electrons. If more than one electron is present, the one with the highest E_T is selected. Events are rejected if the electron is identified as coming from the conversion of a photon. Details on the electron selection can be found in appendix A.3.

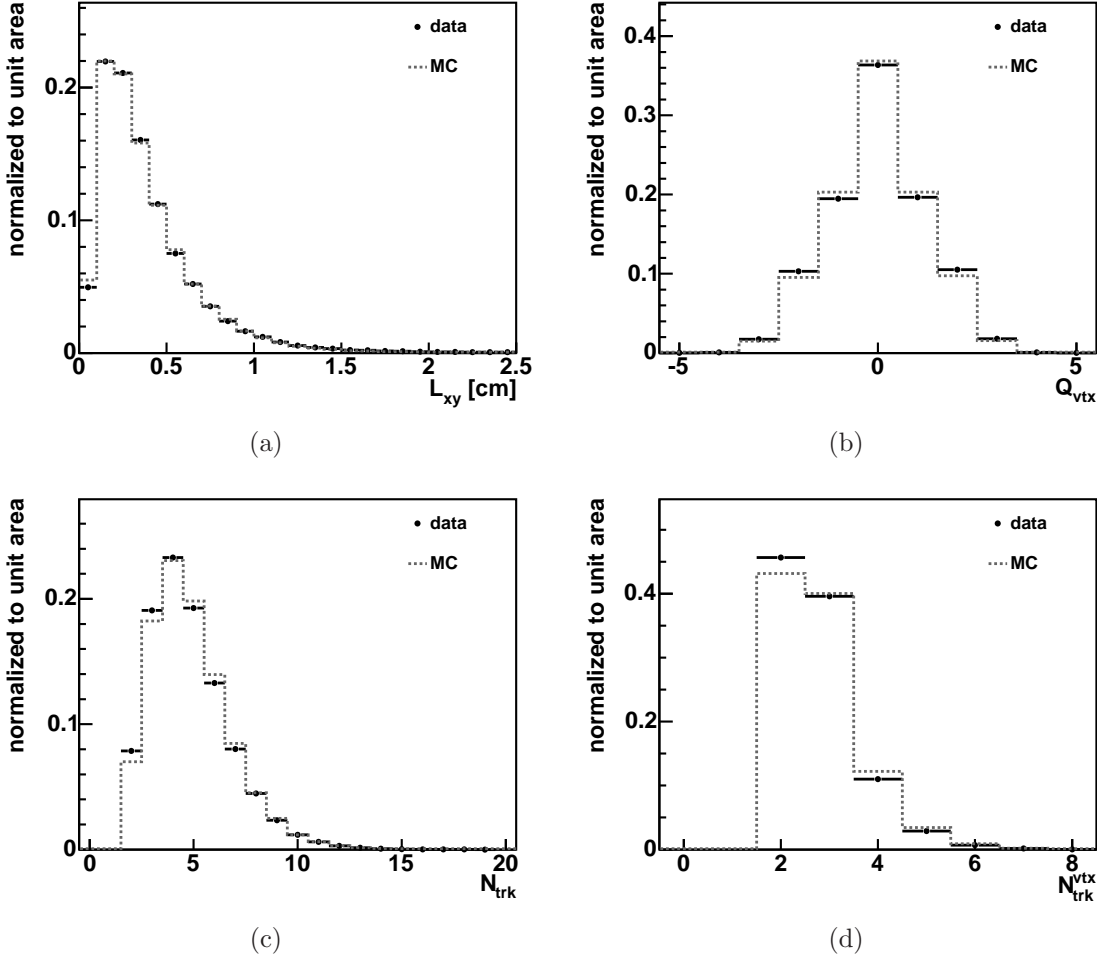


Figure 6.7: Comparison of discriminating variables in data and simulation: (a) the 2-dimensional decay length of the secondary vertex, (b) the charge at the secondary vertex, (c) the number of tracks per jet, and (d) the number of tracks fitted to the secondary vertex.

The selected electron must be within $\Delta R = 0.4$ of a 15 GeV jet, called “electron jet”. If more than one jet is present, the closest one is selected.

To account for the fact that b quarks are mostly produced in pairs, an additional 15 GeV jet with $\Delta\phi > 2$ radians relative to the electron jet and $\eta < 1.5$ is required, the so-called “away jet”. If there is more than one candidate, the one with the smallest $|\Delta\phi - \pi|$ is chosen.

To evaluate the differences between the output distributions of b -quark jets in data and simulation, the away jet is considered, since utilizing the electron jet would introduce a strong bias due to the presence of the lepton.

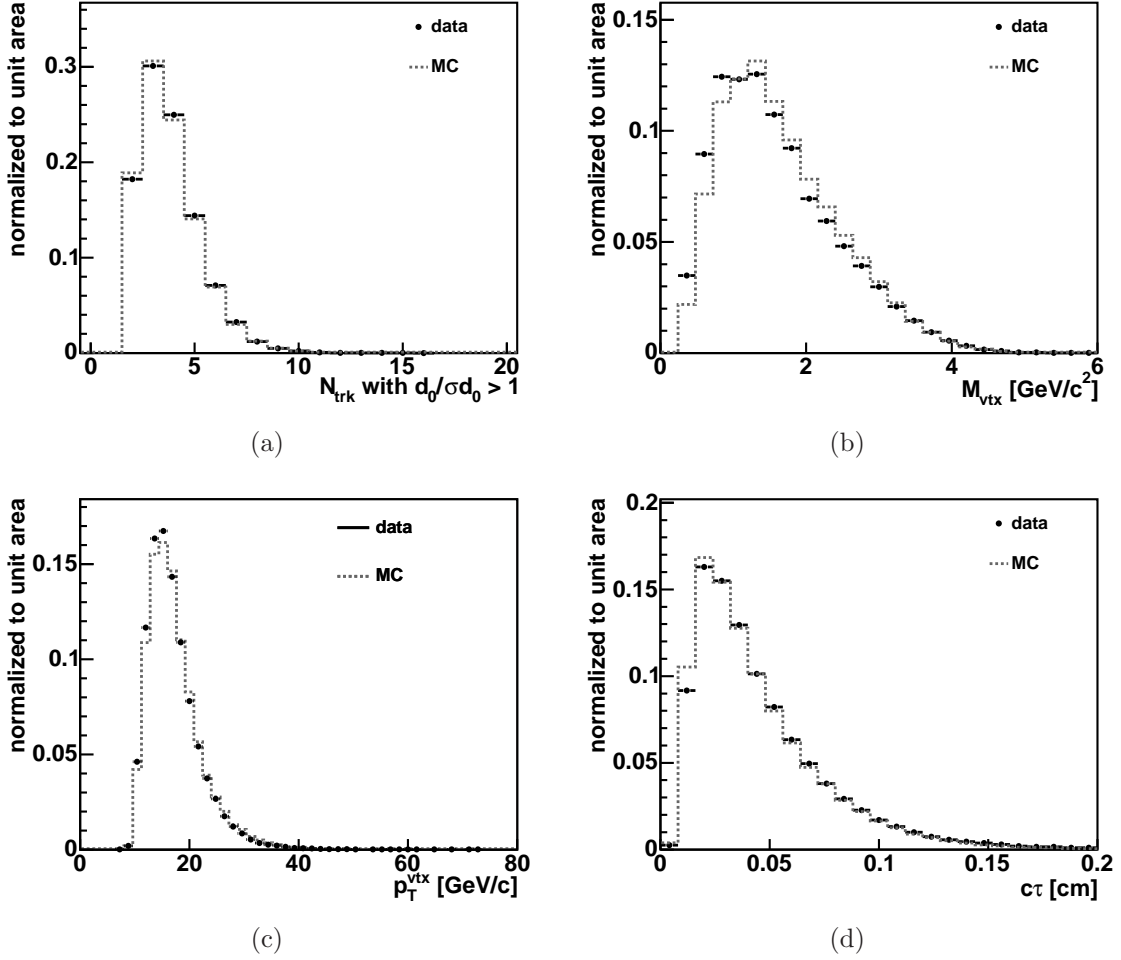


Figure 6.8: Comparison of discriminating variables in data and simulation: (a) the number of tracks in the tagged jet with an impact parameter significance larger than 1σ , (b) the invariant mass of the vertex, (c) the transverse momentum at the secondary vertex if the electron track is used for the fit, and (d) the pseudo-lifetime of the secondary vertex.

Comparability of Flavor Compositions

Even though the selected events are enriched in heavy flavors, the contributions of the different flavors are expected to vary between data and simulation. Hence, a first step in evaluating the differences between data and simulation is to understand the flavor composition of the considered jets.

In simulated events, the contributions of the different flavors can be determined by looking for heavy-flavor quarks within a cone of $R = 0.4$ around the jet axis. In data, this task is more difficult to perform. One established method to estimate the flavor composition of tagged jets is to make use of the distribution of the mass at the secondary vertex, M_{vtx} , being very distinct for the differently flavored jets. For this estimation, template distributions of M_{vtx} , calculated at secondary vertices in jets of different flavors, are fitted to the distribution observed in data. Those template

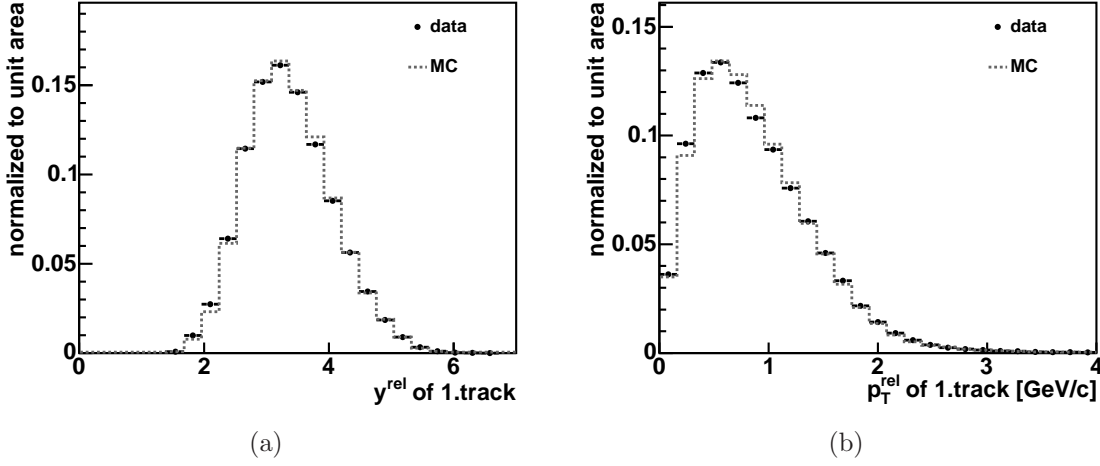


Figure 6.9: Comparison of discriminating variables in data and simulation: (a) the rapidity y^{rel} and (b) the transverse momentum p_T^{rel} with respect to the jet axis of the most energetic track in the tagged jet.

distributions, displayed in figure 6.10(a), are obtained from away jets in simulated events. The composition of tagged away jets is estimated for two cases: (1) both electron and away jet are required to be tagged and (2) only the away jet is required to be tagged (the electron jet is not necessarily tagged). The fitted distributions of both scenarios are shown in figures 6.10(b) and 6.10(c), describing the observed distributions well.

(a) away jet tagged, electron jet tagged

	b -quark fraction [%]	c -quark fraction [%]	light-quark fraction [%]
MC	95.0	3.2	1.8
data	87.2 ± 3.0	9.9 ± 3.8	2.9 ± 2.0

(b) away jet tagged

	b -quark fraction [%]	c -quark fraction [%]	light-quark fraction [%]
MC	88.9	8.8	2.3
data	72.9 ± 1.9	17.2 ± 2.8	9.9 ± 1.9

Table 6.4: Flavor composition of tagged away jets in data and simulation. The flavor composition in data was estimated by fitting templates of M_{vtx} to the distribution in data, as illustrated in figure 6.10. The flavor compositions are given for (a) both electron and away jet required to be tagged and (b) only the away jet required to be tagged.

The composition of away jets, estimated in case of observed data and determined for simulated events, is given in table 6.4. As expected, the contribution of b -quark jets to the sample of tagged away jets increases if the electron jet is additionally required to be tagged. Nevertheless, the b -quark fraction measured in the simulation is larger than estimated for away jets in data. If only the away jet is required to be tagged,

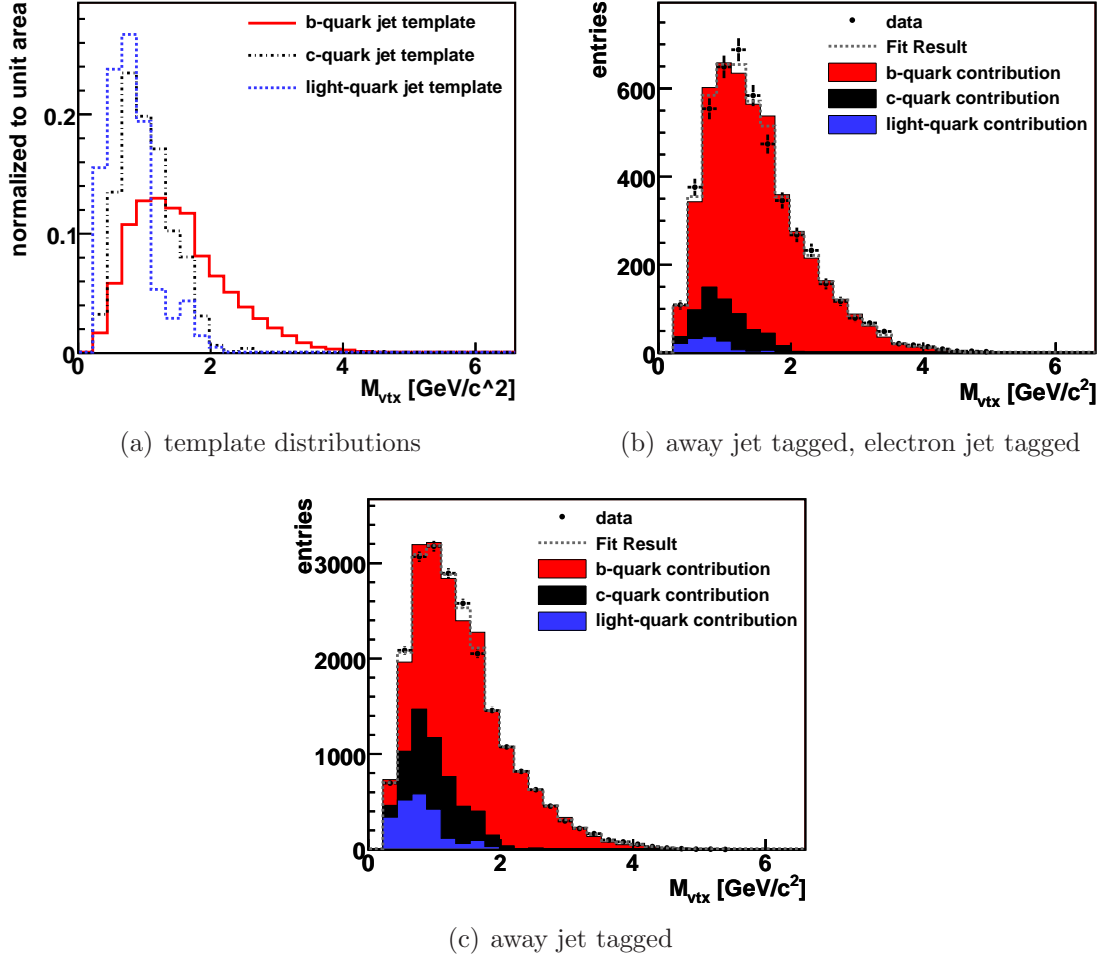


Figure 6.10: Fit of M_{vtx} templates to estimate the flavor composition of the away jet in data. Figure (a) shows the templates obtained from tagged simulated away jets which are fitted to the M_{vtx} distribution of away jets in data. Figure (b) shows the fit result if both electron and away jet are required to be tagged. Additionally, the contributions of the different templates are shown as given by table 6.4. The fit result is not equal to the weighted sum of the templates, since the fit takes statistical uncertainties of the templates into account. The result of the fit if only the away jet is required to be tagged is shown in figure (c).

the flavor composition in the simulation is measured to be 88.9%, while the estimate in data is only 72.9%, i.e. the b -quark fraction in the simulation is 22% larger than estimated in the data. If the electron jet is additionally required to be tagged, the b -quark fraction in simulated away jets increases to 95%, whereas the estimate in data is 87.2%. Since the composition seems to agree at a level of 9%, events with both jets tagged are expected to be more qualified to evaluate the differences between the output distributions of b -quark jets in data and simulation.

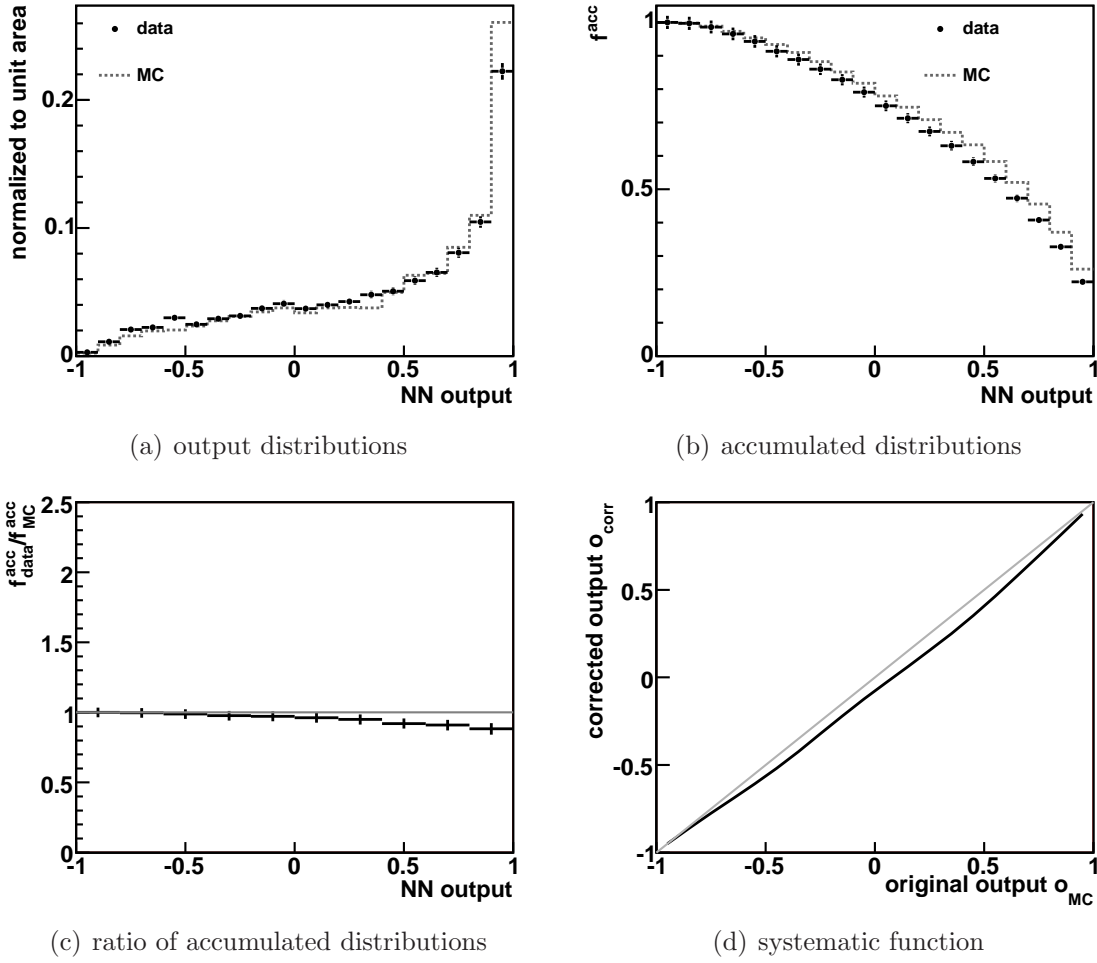


Figure 6.11: The estimation of systematic uncertainties due to differences between data and simulation is illustrated. For all distributions, tagged away jets in events with tagged electron jets are utilized. The output distributions of away jets in data and MC are shown in figure (a), figure (b) shows the accumulated output distributions, figure (c) shows the ratio of the accumulated distribution in data divided by the accumulated distribution in the simulation, and (d) shows the function quantifying the systematic uncertainties due to the observed differences between b -quark jets in data and simulation.

Evaluating the Systematic Uncertainties

Figure 6.11(a) shows the output distributions of tagged away jets in observed and simulated events with tagged electron jets. As expected from the estimated flavor composition, the output distribution is more b -quark-like in simulated events than in data. Nevertheless, adjusting the flavor composition of simulated jets to the estimated composition of jets in observed events would introduce a large systematic uncertainty, since the fit of M_{vtx} templates only provides a rough estimate.

To quantify the differences between data and simulation, the output distributions are integrated from right to left, i.e. they are accumulated, having the advantage that the accumulated distributions f^{acc} are monotonic decreasing as displayed in

figure 6.11(b). The value of the accumulated distribution at each output value o , $f^{\text{acc}}(o)$, is equivalent to the ratio of events with outputs larger than o divided by the total number of events: $f^{\text{acc}}(o) = N_{\text{events}}(\text{output} \geq o)/N_{\text{events}}$. If the output distributions were completely alike, the ratio of the accumulated distribution in data $f_{\text{data}}^{\text{acc}}(o)$ divided by the accumulated distribution in the simulation $f_{\text{MC}}^{\text{acc}}(o)$, displayed in figure 6.11(c), should be equal to one. If the neural-network b -tagger is implemented by cutting on the output, this ratio can be used to correct the cut efficiency in the simulation in such a way that it fits the one in data.

The use of the b tagger as a continuous variable, like it is realized in the search for single top-quark production, necessitates a method allowing to correct the output event-by-event for differences between data and simulation. For this purpose, the output o_{MC} of simulated events is shifted to the value o_{corr} where the accumulated distribution in data has the same value as the accumulated distribution in the simulation at that point: $f_{\text{data}}^{\text{acc}}(o_{\text{corr}}) \equiv f_{\text{MC}}^{\text{acc}}(o_{\text{MC}})$. The dependence of the corrected output o_{corr} on the output o_{MC} is demonstrated in figure 6.11(d). The application

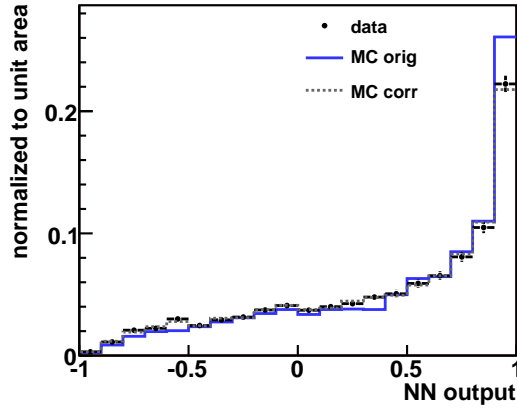


Figure 6.12: The impact of the systematic correction of b -quark jets is demonstrated by comparing the output distribution of tagged away jets in observed events to the one in simulation before and after the application of the systematic correction function.

of this function to simulated b -quark jets leads to a less signal-like distribution, as illustrated in figure 6.12 showing the output distribution of tagged away jets in observed events (with the electron jet additionally tagged) compared to simulated distributions before and after the correction. This closure test of the systematic correction demonstrates that the corrected output describes the observed distribution well.

Since the observed deviations of the simulated distribution from the one in data, illustrated in figures 6.11(c) and 6.11(d), are small and expected to be partly induced by different flavor compositions in simulation and data, the displayed function is not used as a default correction of the output of simulated b -quark jets but rather to estimate a systematic uncertainty, as explained in section 6.4.1.

Comparison of the Output Distributions of the Electron Jet

As already mentioned, the electron jet has quite specific properties and cannot be used to evaluate systematic uncertainties. Nevertheless, examining its output distribution in data and simulation is an interesting cross check.

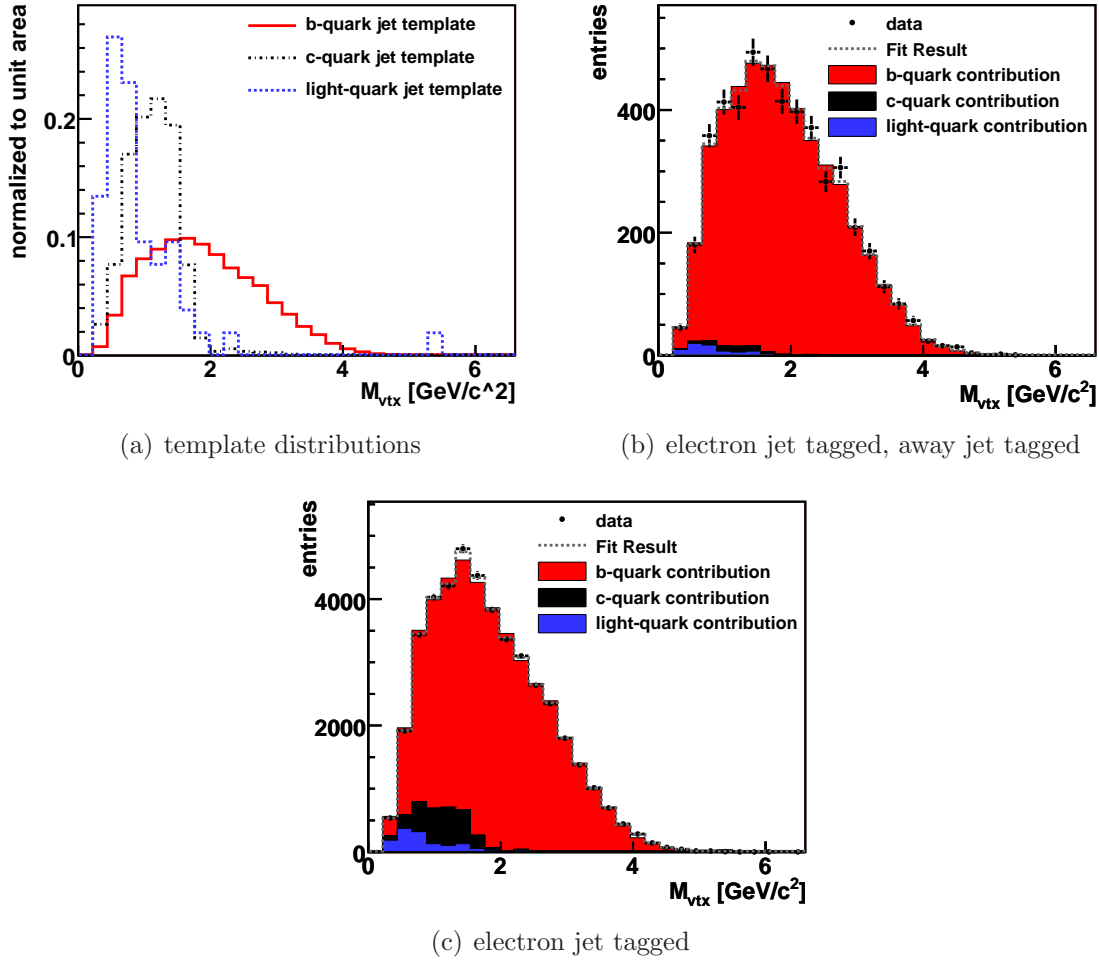


Figure 6.13: Fit of M_{vtx} templates to estimate the flavor composition of the electron jet in data. Figure (a) shows the templates obtained from tagged simulated electron jets which are fitted to the M_{vtx} distribution of electron jets in data. Figure (b) shows the fit result if both electron and away jet are required to be tagged. Additionally, the contributions of the different templates are shown as given by table 6.5. The fit result is not equal to the weighted sum of the templates, since the fit takes statistical uncertainties of the templates into account. The result of the fit if only the electron jet is required to be tagged is shown in figure (c).

As described before in case of the away jet, the flavor composition of the electron jet is determined by fitting template distributions of M_{vtx} to the distribution observed in data. The template distributions obtained from electron jets in simulated events are shown in figure 6.13(a), the results of the fits to the observed distributions are illustrated in figure 6.13(b) for the case that both jets are tagged and in figure 6.13(c) for the case that only the electron jet is required to be tagged.

(a) electron jet tagged, away jet tagged			
	b -quark fraction [%]	c -quark fraction [%]	light-quark fraction [%]
MC	98.0	1.9	0.04
data	97.6 ± 2.1	0.9 ± 1.7	1.5 ± 1.0

(b) electron jet tagged			
	b -quark fraction [%]	c -quark fraction [%]	light-quark fraction [%]
MC	92.7	7.2	0.1
data	90.7 ± 1.1	6.2 ± 1.2	3.1 ± 1.1

Table 6.5: Flavor composition of tagged electron jets in data and simulation. The flavor composition in data was estimated by fitting templates of M_{vtx} to the distribution in data, as illustrated in figure 6.13. The flavor compositions are given for (a) both electron and away jet required to be tagged and (b) only the electron jet required to be tagged.

The estimated flavor composition of the electron jet in data is shown in table 6.5 in addition to the composition determined for simulated events. Due to the presence of the electron, the b -quark fraction is larger, while the deviations of the composition in the simulation from the expectation in data are smaller than in case of the away jet. If both jets are required to be tagged, the b -quark fraction in the simulation is almost identical to the one estimated in data. Therefore, the deviations in the output distributions are expected to be smaller than in case of the away jet. This is demonstrated in figure 6.14. Comparing the accumulated distributions shown in figure 6.14(b) or the ratio of the accumulated distributions displayed in figure 6.14(c), it can be seen that the deviations of the output distribution of simulated electron jets from the observed one are negligible.

6.3.4 Output Distributions of Mistagged Jets

The comparison of the output distributions of light-quark jets in data and simulation is a difficult challenge, since the neural-network b -tagger is applied to events already selected by the secondary-vertex b -tagger. Nevertheless, events with falsely reconstructed secondary vertices can be selected by making use of negatively tagged jets, i.e. secondary vertices with a negative decay length as introduced in section 4.3.4. Under the assumption that negatively tagged vertices are a valid description of positively mistagged vertices, those events can be utilized for the comparison of mistagged jets in data and simulation. One caveat is that the decay length is negative per definition. This is taken into account by using its absolute value.

The simulation of mistagged jets is a challenging task, since it requires adequate simulation of detector effects. Hence, a first important check is to quantify if mistags are modeled consistently in simulation. Therefore, in simulated dijet events, all-flavor negatively tagged jets are compared to negatively tagged light-flavored jets and to positively tagged light-flavored jets. As shown in figure 6.15(a), the output

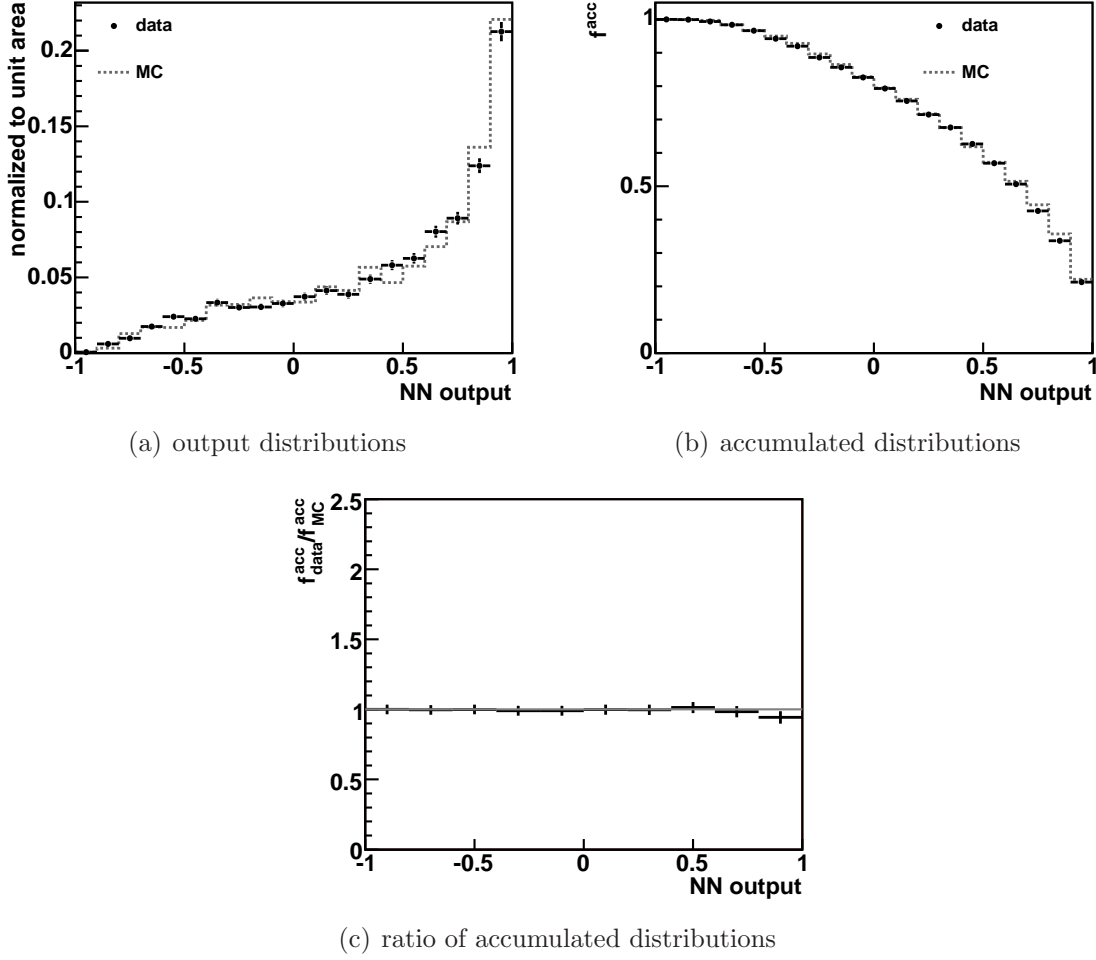


Figure 6.14: The differences in the output distributions of tagged electron jets in data and simulation are evaluated. For all distributions, the away jet is additionally required to be tagged. The output distributions of electron jets in data and MC are shown in figure (a), figure (b) shows the accumulated output distributions, and figure (c) shows the ratio of the accumulated distribution in data divided by the accumulated distribution in the simulation.

distributions of the various kinds of mistagged jets look consistently.

The comparison between negatively tagged jets in simulated and observed jet events, presented in figure 6.15(b), reveals that the output distribution of mistags in data appears less background-like than indicated by simulated mistags. Hence, it is necessary to correct the output distribution of simulated mistags in such a way that it describes the observation. As described in case of the estimation of systematic uncertainties due to the modeling of b -quark jets (see section 6.3.3), the output distributions are accumulated to quantify the differences between mistags in data and simulation, being depicted in figure 6.15(c). As expected, the ratio of the accumulated data distribution divided by the accumulated MC distribution, shown in figure 6.15(d), exhibits significant deviations from one, confirming the conclusion that the output distribution of simulated mistagged jets needs to be corrected.

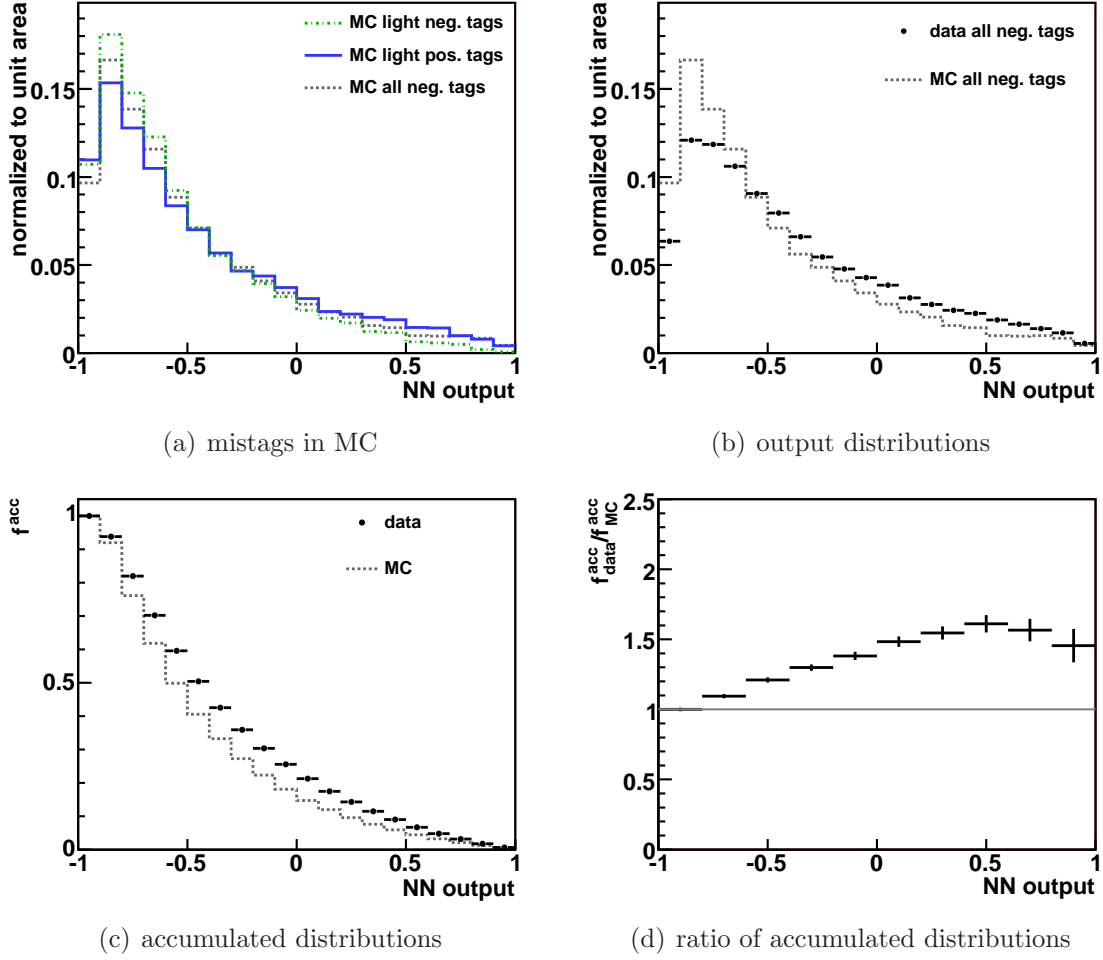


Figure 6.15: Comparison of mistag output distributions: in figure (a), the output distributions of negatively tagged jets are compared to negatively and positively tagged light-quark jets in simulated events, figure (b) shows the output distributions of negatively tagged jets in data and simulation, figure (c) depicts the accumulated output distributions of negatively tagged jets in data and simulation, and figure (d) displays the ratio of the accumulated data distribution divided by the simulated one.

Since the deviation of simulated mistags from observed ones differs for pass-1 and pass-2 vertices, as demonstrated in figure 6.16, the corrections are derived separately for the two passes of the vertex fit.

Parameterization of the Correction Function

As noted in section 4.3.4, the mistag probability depends on several parameters, namely E_T , η , and ϕ of the jet, track multiplicity N_{trk} and the sum of the transverse energies of the taggable jets in the event ($\sum E_T$). Hence, it is prospective that the differences between the output distributions of mistagged jets in data and simulation also depend on those quantities, which is investigated in the following. Figures 6.17

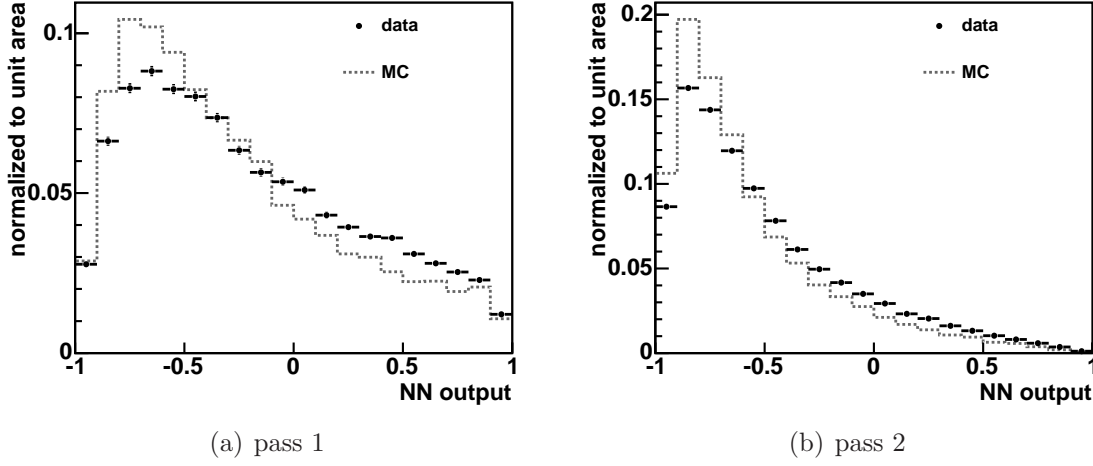


Figure 6.16: The output distributions of negatively tagged (a) pass-1 and (b) pass-2 vertices in data and simulation are shown.

and 6.18 show the ratio of the accumulated output distribution in data divided by the one in simulation for different bins of the above mentioned variables for pass-1 and pass-2 vertices, respectively. Here, the dependences are treated completely inclusive, i.e. possible correlations between the different bins are ignored. Within the given statistics, pass-1 mistags seem only to depend on jet E_T , while pass-2 mistags show dependences on jet E_T , $\sum E_T$, and N_{trk} . In both cases, no angular dependences can be observed.

The function correcting for the differences observed between the output distributions of mistagged jets in data and simulation is exclusively parameterized dependent on the above mentioned quantities, as demonstrated in figure 6.19. As described in section 6.3.3, the corrected output o_{corr} is given by the accumulated distributions: $f_{\text{data}}^{\text{acc}}(o_{\text{corr}}) \equiv f_{\text{MC}}^{\text{acc}}(o_{\text{MC}})$. It can be seen that the corrections are negligible for some bins, e.g. $E_T < 30$ GeV in case of pass-1 vertices and both $\sum E_T$ bins for $E_T \geq 30$ GeV and $N_{\text{trk}} < 6$ in case of pass-2 vertices. Other bins show larger deviations of the simulated output from the observed one, like $E_T \geq 30$ GeV for pass-1 vertices and both $\sum E_T$ bins for $E_T \geq 30$ GeV and $N_{\text{trk}} \geq 6$ in case of pass-2 vertices.

Due to construction, the application of the correction function to the output of simulated mistags leads to a less background-like distribution. In figure 6.20, the output distribution of mistags in data is compared to the distribution of simulated mistags before and after correction for the observed differences between data and simulation. It can be seen that the corrected distribution describes the data well for both pass-1 (see figure 6.20(a)) and pass-2 (see figure 6.20(b)) vertices.

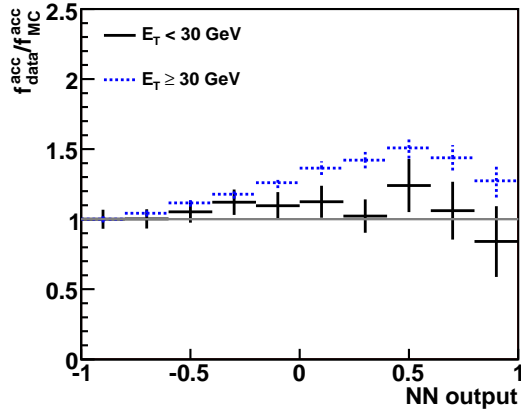
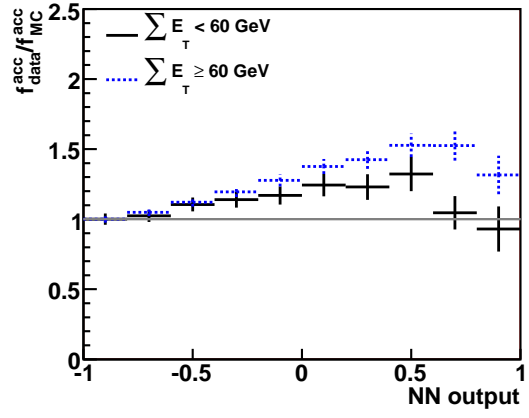
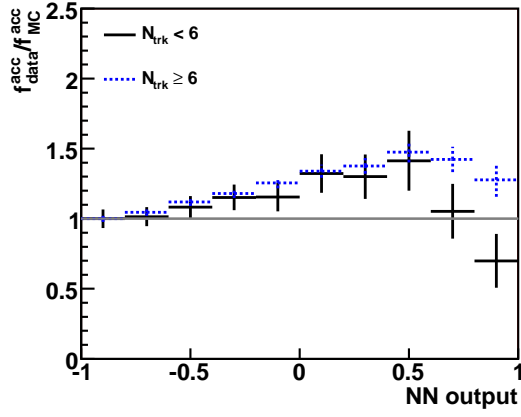
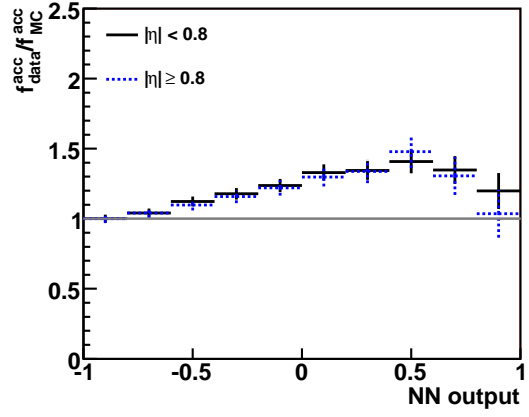
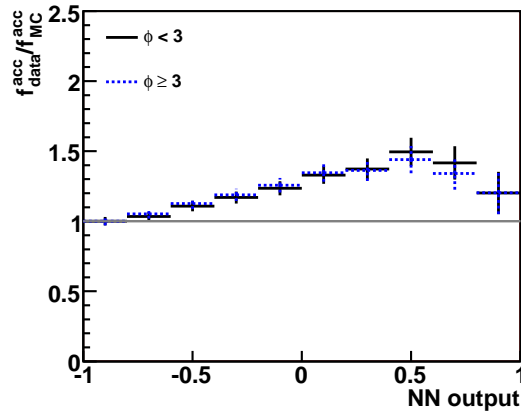
(a) pass 1 in dependence of E_T (b) pass 1 in dependence of $\sum E_T$ (c) pass 1 in dependence of N_{trk} (d) pass 1 in dependence of η (e) pass 1 in dependence of Φ

Figure 6.17: The ratio of the accumulated data distribution divided by the accumulated MC distribution for pass-1 vertices is shown in dependence of (a) jet E_T , (b) the sum of the transverse energies of the taggable jets, (c) the number of tracks per jet, (d) jet η , and (e) jet ϕ

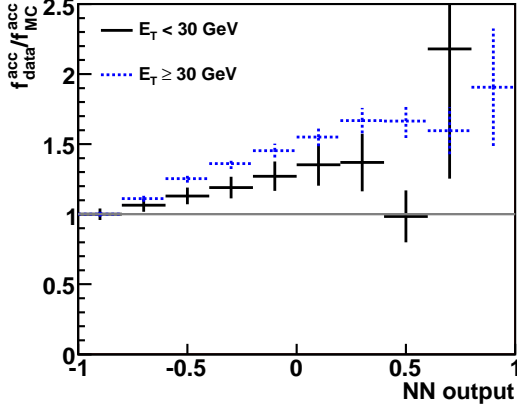
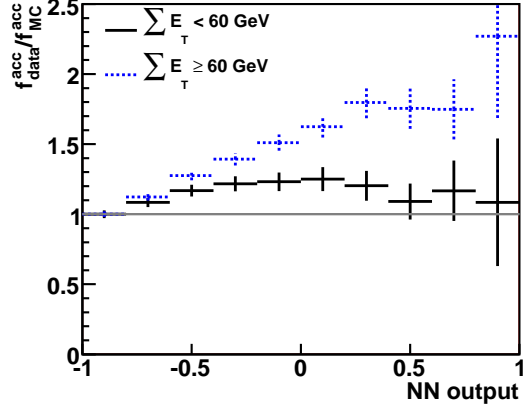
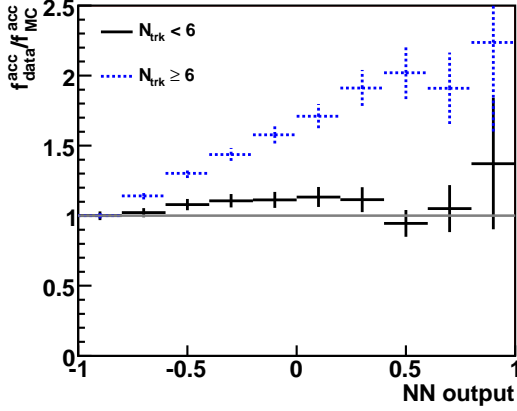
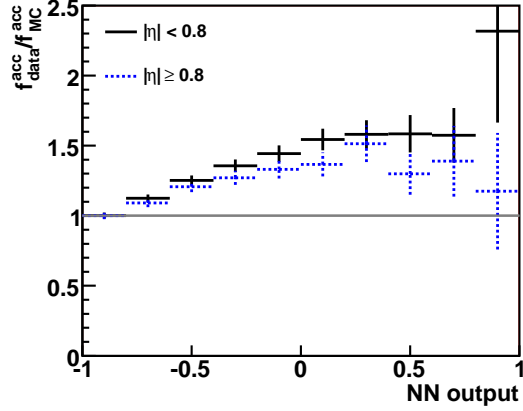
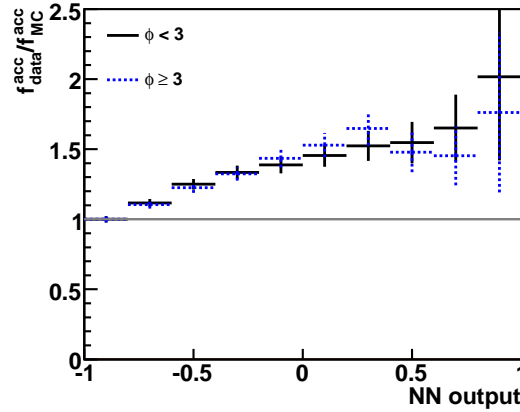
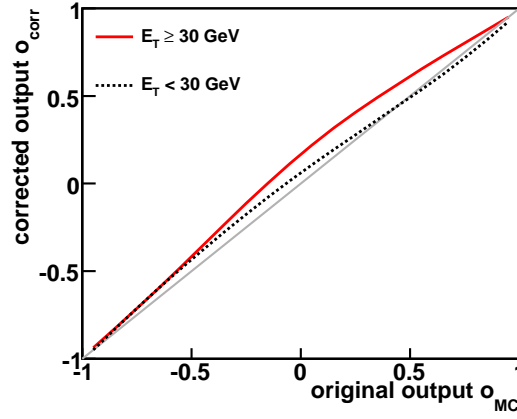
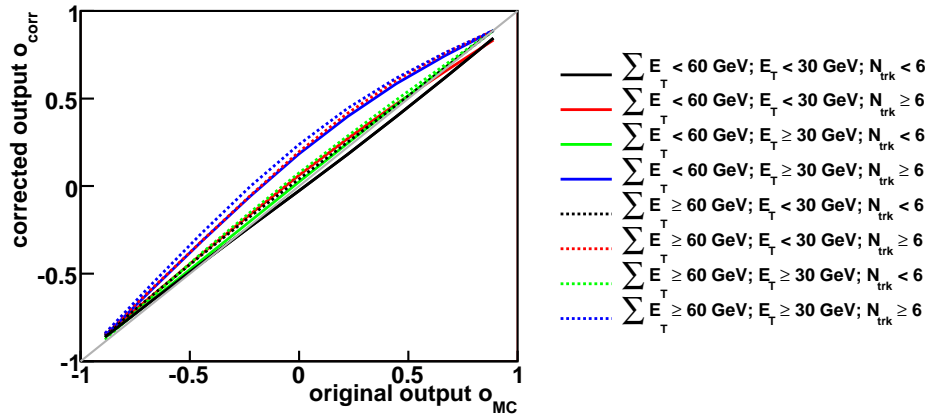
(a) pass 2 in dependence of E_T (b) pass 2 in dependence of $\sum E_T$ (c) pass 2 in dependence of N_{trk} (d) pass 2 in dependence of η (e) pass 2 in dependence of Φ

Figure 6.18: The ratio of the accumulated data distribution divided by the accumulated MC distribution for pass-2 vertices is shown in dependence of (a) jet E_T , (b) the sum of the transverse energies of the taggable jets, (c) the number of tracks per jet, (d) jet η , and (e) jet ϕ



(a) pass 1



(b) pass 2

Figure 6.19: Functions correcting for the observed differences between mistags in data and simulation for (a) pass-1 and (a) pass-2 vertices. The functions are parameterized according to the dependences observed in figures 6.17 and 6.18.

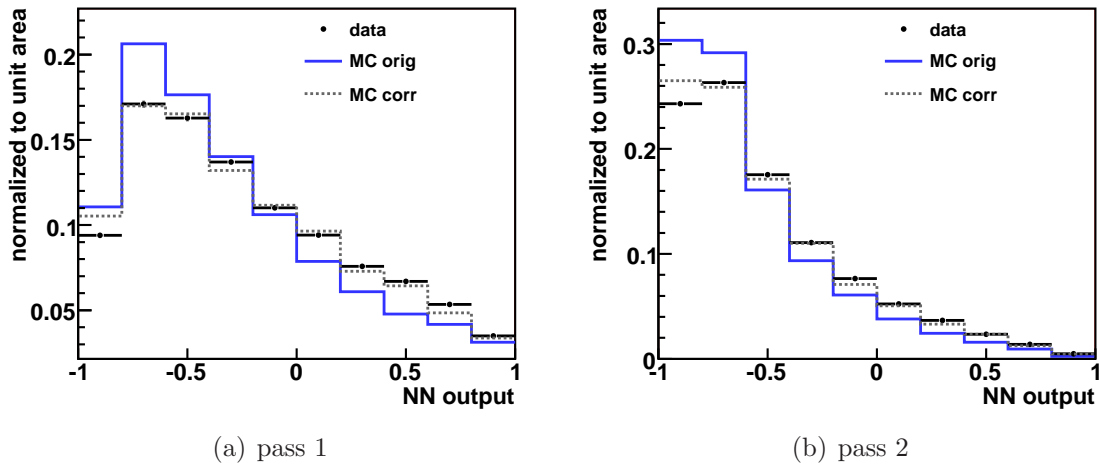


Figure 6.20: The mistag correction is demonstrated by comparing the observed mistag output distribution to the one in simulation before and after the application of the correction function for (a) pass-1 and (b) pass-2 vertices.

6.4 Application in Analyses

The b -tagger described above is the first neural-network b -tagger developed at a hadron collider. At CDF, it has successfully been implemented into all searches for single top-quark production, where it has proven to be an important quantity to reject events not containing any b quarks. As demonstrated in section 7, the b -tagger output is one of the most important variables to discriminate single-top-quark events.

Even more, the output of the b -tagger can not only be used as a discriminating variable, but also to estimate the flavor composition of a given data sample. This way of application has a large potential to improve the estimation of the expected event yield, gaining importance with increasing amount of collected data.

Furthermore, the neural-network b -tagger is a universal tool which can be utilized in all high- p_T analyses at CDF, requiring identification of b -quark jets, like the search for the Higgs boson predicted by the standard model.

6.4.1 Implementation in the Search for Single Top Quarks

As mentioned above, the output of the neural-network b -tagger can be used as a continuous variable by exploiting the different shapes of the output distributions of the various physics processes, as done in the search for single top-quark production. This way of application is illustrated in figure 6.21 for simulated events passing the single-top-quark event selection described in section 5.2. Exemplarily, the distri-

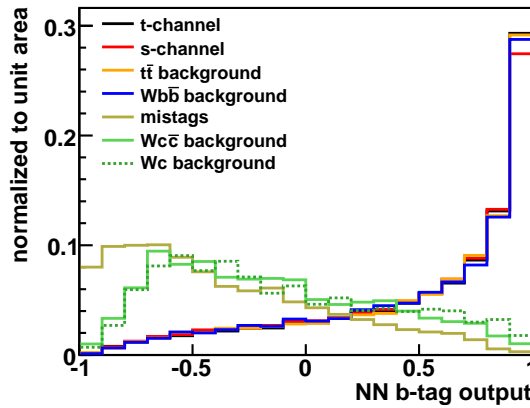


Figure 6.21: The b -tagger output distributions of simulated events passing the event selection are shown for production of single top quarks, $t\bar{t}$, $Wb\bar{b}$, Wc , and $Wc\bar{c}$ in addition to the distribution of the mistag model described in section 5.3.2. The output of the tagged jet with the largest E_T in events passing the single-top-quark event selection is shown.

butions of single-top-quark events are compared to production of $t\bar{t}$ and W +heavy flavor, in addition to the mistag model described in section 5.3.2. The output of

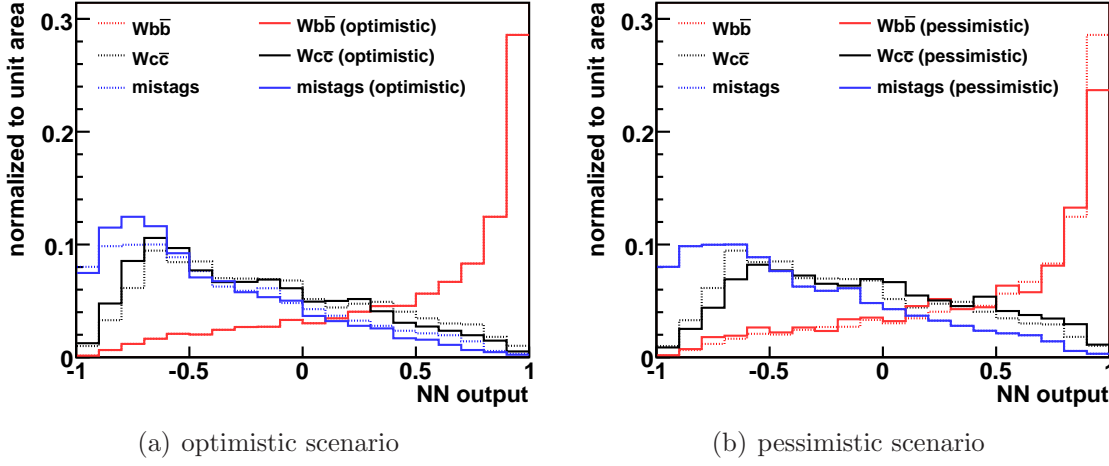


Figure 6.22: A systematic shape uncertainty for the output of the neural-network b -tagger can be estimated by utilizing the correction function illustrated in figures 6.11(d) and 6.19. This is realized by defining two scenarios: (a) an optimistic and (b) a pessimistic one. In the optimistic case, the separation between b -quark jets on one hand and c -quark and mistagged jets on the other hand is increased, while it is decreased in the pessimistic scenario. The dashed lines show the default output distributions, while the solid lines illustrate the output distributions utilized in the corresponding systematic scenario.

the light-flavored jets was corrected for the observed differences between data and simulation using the correction functions illustrated in figure 6.19.

Besides the expected separation between b -quark jets on one hand and c -quark and light-quark jets on the other hand, it can be seen that the output distributions of jets of the same quark flavor look similar, no matter of the physical production mode. This leads to the conclusion that the neural-network b -tagger can be used as a generic tool in high- p_T analyses where flavor separation is needed, like in searches for Higgs bosons in the mass range below $140 \text{ GeV}/c^2$ where it dominantly decays into $b\bar{b}$.

Estimating a Systematic Uncertainty

The shapes of the b -tagger output distributions of simulated events are afflicted with a systematic uncertainty. This uncertainty can be estimated by making use of the above derived correction functions for b -quark jets (figure 6.11(d)) and mistags (figure 6.19), realized by defining two scenarios, an optimistic and a pessimistic one. In the optimistic case, illustrated in figure 6.22(a), the distributions are altered in such a way that the discrimination between b -quark jets on one hand and c -quark and mistagged jets on the other hand is increased. This is achieved by not correcting the output of mistagged light-quark jets and by utilizing the systematic correction function (derived for b -quark jets) to shift the output of c -quark jets to the more background-like region. In the pessimistic scenario, demonstrated in figure 6.22(b), the discrimination between b -quark and non- b -quark jets is decreased by applying

the systematic b -quark function to b -quark jets and the mistag correction functions to c -quark jets. The distribution of simulated mistags is corrected as usual.

The influence of the application of this procedure to the result of the search for single top-quark production is illustrated in section 7.3.

6.4.2 Estimation of Flavor Fractions

Due to the different shapes of the output distributions of b -quark, c -quark, and light-quark jets, the neural-network b -tagger can be utilized to estimate the flavor composition of a given data sample by fitting template output distributions, obtained from differently flavored simulated jets, to the output distribution of the selected data events.

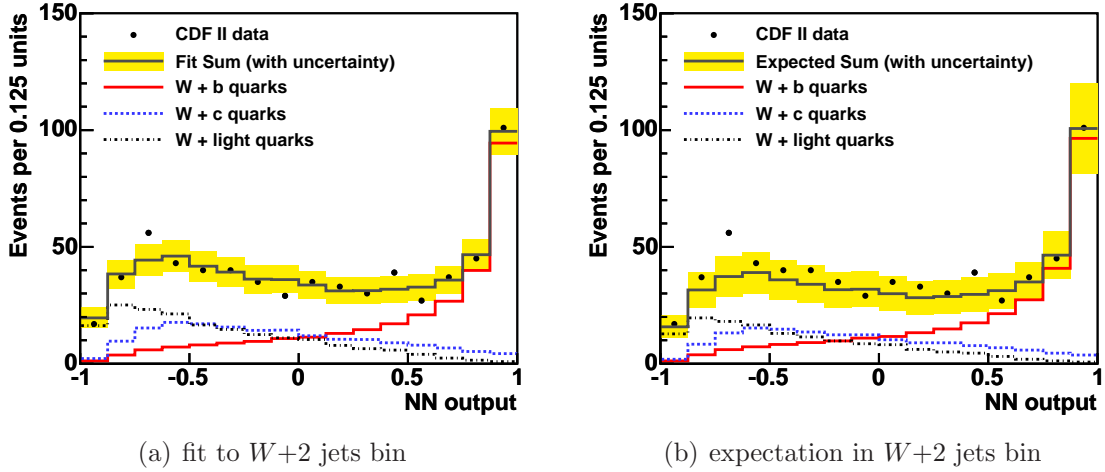


Figure 6.23: The use of the neural-network b -tagger to estimate the flavor composition is illustrated exemplarily for the observed events passing the single-top-quark event selection. Figure (a) shows the result of a template fit to the data distribution including the statistical uncertainty (yellow band). The templates are scaled to the fitted values given in table 6.6. As a comparison, figure (b) illustrates the respective distributions scaled to the expected event yield, derived in section 5.4. The accordant contributions of the different flavors are also given in table 6.6. The uncertainty on the prediction includes all systematics as explained in section 7.3.

In the search for single top-quark production [122], such a fit has been performed to check the estimated event yield derived in section 5.4. The result of the template fit to the b -tagger output distribution of the observed events passing the single-top-quark event selection is illustrated in figure 6.23(a). The fitted contributions of the different flavors are quoted in table 6.6. For comparison, the expected contributions of each flavor (based on the event yield for the different physics processes presented in table 5.2) are given in table 6.6 as well. In figure 6.23(b), the respective expected contributions to the total number of expected events are compared to the distribution observed in data. Both estimated output distributions describe the one in data well.

flavor	b –tagger fit	event yield
b	292.8 ± 26.3	299.0 ± 56.8
c	171.6 ± 53.8	148.1 ± 39.4
light	179.5 ± 42.5	140.0 ± 19.8
total	644	587.1 ± 96.6

Table 6.6: Expected contributions of the different flavors to the observed data events passing the single–top–quark event selection. The numbers obtained by a template fit of the output distribution of the neural–network b –tagger (see figure 6.6) are given as well as the predicted contributions according to the event yield (described in section 5.4). The uncertainties on the fit result are statistical only, while those on the prediction include systematics as explained in section 7.3.

Within the given uncertainties, the fit reproduces the expected flavor contributions, even though the fitted central values of the c –quark and light–quark contributions are slightly higher than the ones obtained from the event yield estimation described in section 5.4. This indicates that the observed excess is mainly due to charm and mistagged events. However, since the uncertainties on the fitted flavor contributions are still large (but are expected to decrease with increasing statistics), the fitted numbers can currently only be used as a cross check. Nevertheless, this way of utilizing the neural–network b –tagger has a large potential to improve the estimation of the expected number of events in future analyses.

Chapter 7

Search for Single Top–Quark Production

As outlined in section 1.2.2, the standard model predicts two dominant electroweak production modes of single top quarks at the Tevatron: t -channel and s -channel single-top-quark production, both investigated in this analysis. In this search for single top-quark production, only leptonic W -boson decays in events with exactly two tight jets are targeted.

Since a small signal must be extracted from a huge amount of background events, multivariate methods are an appropriate approach, exploiting a large number of characteristic quantities. In this analysis, neural networks are used to discriminate single-top-quark events by employing the NeuroBayes[®] package, described in section 3.1. To measure the cross section of single top-quark production, signal and background templates of the neural network output are created from simulated events, which are fitted to the output distribution of observed events.

For the search of single-top-quark events, two different approaches are taken. One is to assume the ratio of the cross sections of both production channels to be as predicted by the standard model (see section 1.2.2). In this approach, one neural network is trained to identify both t - and s -channel events, the so-called combined network. The other is to determine the cross sections of both channels independently and simultaneously, named separate search. Here, two different neural networks are trained, one to identify only t -channel events and one to identify only s -channel events. For the t -channel neural network, s -channel events are treated as background and vice versa.

7.1 Neural–Network Input–Variables

For all three 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 neural-network

b -tagger described in chapter 6 and the kinematic fitter introduced in section 7.1.2. Furthermore, some quantities need the knowledge about the four-vector of the top quark whose reconstruction is described in section 7.1.1.

7.1.1 Top-Quark Reconstruction

For some of the variables, the reconstruction of the top quark is necessary. At this, the top-quark four-momentum is built out of the reconstructed W boson and a defined b -quark jet.

The first step in top-quark reconstruction is the selection of the b -quark jet. If there is only one tagged jet, this jet is utilized to reconstruct the top quark. If there is more than one jet with a reconstructed secondary vertex, the tagged jet with the largest product of the charge of the tight lepton (Q_ℓ) and the jet pseudorapidity, $Q_\ell \cdot \eta$, is assigned to belong to the top-quark decay [123]. At the top-quark reconstruction level, the momentum p of the selected b -quark jet is corrected up to level 7, i.e. parton level. The mass of the b -quark jet is set to $m_b = 5 \text{ GeV}/c^2$, its energy calculated by $(E_b)^2 = (m_b)^2 + (\vec{p})^2$.

The second step is the reconstruction of the four-momentum of the W boson, being built from the four-momenta of the measured tight lepton and the reconstructed neutrino. In good approximation, the transverse neutrino momentum can be derived from the missing transverse energy \cancel{E}_T . To be consistent with the jet energy correction applied to the b -quark jet, jets corrected up to level 6 are utilized to calculate \cancel{E}_T . Out-of-cone, i.e. level 7, corrections are omitted to avoid double counting. The z component of the neutrino momentum is obtained by using a quadratic constraint of the W -boson decay-kinematics, assuming a leptonically decaying W boson and a massless neutrino: $(\mathbf{p}_W)^2 = (\mathbf{p}_e + \mathbf{p}_\nu)^2 = m_W^2$. Solving for p_z^ν results in the following expression:

$$p_z^\nu = \frac{\kappa p_z^\ell}{(E_\ell)^2 - (p_z^\ell)^2} \pm \frac{1}{2((E_\ell)^2 - (p_z^\ell)^2)} \quad (7.1)$$

$$\cdot \sqrt{(2\kappa p_z^\ell)^2 - 4((E_\ell)^2(p_T^\nu)^2 - \kappa^2) \cdot ((E_\ell)^2 - (p_z^\ell)^2)}$$

$$\text{with } \kappa = \frac{1}{2}(m_W^2 - m_\ell^2) + \cos(\phi_\ell - \phi_\nu) \cdot p_T^\ell p_T^\nu. \quad (7.2)$$

In this formula, the masses are used as given by tables 1.1 and 1.2: $m_W = 80.4 \text{ GeV}/c^2$, $m_e = 0.511 \text{ MeV}/c^2$, and $m_\mu = 0.106 \text{ GeV}/c^2$. Out of the two solutions of equation (7.1), the one with the smallest $|p_z^\nu|$ is chosen, since neutrinos produced in top-quark decays are rather central. If the solution of the equation is complex, which happens in about 30% of the cases, the real part of the solution is utilized. Finally, the neutrino energy is calculated by $(E_\nu)^2 = (\cancel{E}_T)^2 + (p_z^\nu)^2$.

Using the four-momentum of the W boson, $\mathbf{p}_W = \mathbf{p}_e + \mathbf{p}_\nu$, the four-momentum of the top quark can be calculated by $\mathbf{p}_{\text{top}} = \mathbf{p}_b + \mathbf{p}_W$.

7.1.2 Kinematic Fitter

To evaluate the ambiguities in the reconstruction of the top quark due to selection of the b -quark jet and the neutrino p_z solution, a kinematic fitter [124] is utilized. This fitter varies p_b , p_T^ν , and ϕ_ν to minimize a χ^2 function for all possible combinations of selecting the b -quark jet and neutrino p_z :

$$\chi^2 = \frac{(p_b - p_b^{\text{obs}})^2}{\sigma_{p_b}^2} + \frac{(p_T^\nu - \cancel{p}_T)^2}{\sigma_{\cancel{p}_T}^2} + \frac{(\phi_\nu - \phi_{\cancel{E}_T}^{\text{obs}})^2}{\sigma_{\phi_\nu}^2} + \frac{(M_{\ell\nu b} - m_t)^2}{\sigma_{m_t}^2} + Y \cdot (\text{Im}(p_z))^2. \quad (7.3)$$

In this context, p_b^{obs} is the measured momentum of the jet considered as the b -quark jet from the top-quark decay, \cancel{p}_T is the transverse momentum corresponding to \cancel{E}_T , $\phi_{\cancel{E}_T}^{\text{obs}}$ is the measured azimuthal angle of \cancel{E}_T , and $M_{\ell\nu b}$ is the invariant mass of the considered system of b -quark jet, charged lepton, and neutrino. The parameters $\sigma_{p_b} = 12.9 \text{ GeV}/c$, $\sigma_{\cancel{p}_T} = 9.0 \text{ GeV}/c$, and $\sigma_{\phi_\nu} = 0.2$ were derived by studying simulated events. $\sigma_{m_t} = 0.5 \text{ GeV}/c^2$ constrains $M_{\ell\nu b}$ to $m_t = 175 \text{ GeV}$ in this fit. Additionally, a penalty term $Y \cdot (\text{Im}(p_z))^2$ with $Y = 3.0$ [125] pushes the fit away from unphysical neutrino solutions.

Since there are two possibilities to assign a b -quark jet and two possible solutions for p_z^ν , the kinematic fit results in four different χ^2 values which are used for the training of the neural networks.

7.1.3 Utilized Variables

As already mentioned, a variety of variables is suitable to discriminate single-top-quark events. For the training of the neural networks, 26 variables were initially investigated, whereof a total number of only 23 variables is utilized, since variables with a significance smaller than 3σ are rejected by the automatic variable selection explained in section 3.1.3. Those used variables are depicted in figures 7.1 to 7.6, comparing the distributions of single-top-quark events to those of $t\bar{t}$ events and, exemplarily for W +jets production, $Wb\bar{b}$ events.

Some of the input variables are directly measured in the detector. Figures 7.1(a) and 7.1(b) show the pseudorapidities of the jets η_{j1} and η_{j2} , where $j1$ is the jet with the higher E_T and $j2$ the jet with the smaller E_T . These variables discriminate t -channel single-top-quark events, because both jets are less central than in case of s -channel single-top-quark events and production of $t\bar{t}$ or W +jets. Even though this tendency is only small in case of the leading jet $j1$, the second-leading jet $j2$ has a broad η distribution with a trend towards the forward direction. These features are also reflected in the distributions of $\eta_{j1} + \eta_{j2}$ and $\eta_{j1} - \eta_{j2}$, displayed in figure 7.1(c) and figure 7.1(d), respectively. In case of $Wb\bar{b}$ events, it can be seen that both jets have a tendency to have a small difference in η . Further directly measured variables are the transverse energies of the two jets, illustrated in figures 7.2(a) and 7.2(b). Comparing the spectra of jets generated by different physics processes, it can be seen that both jets in $Wb\bar{b}$ events have softer E_T spectra than those in $t\bar{t}$ or

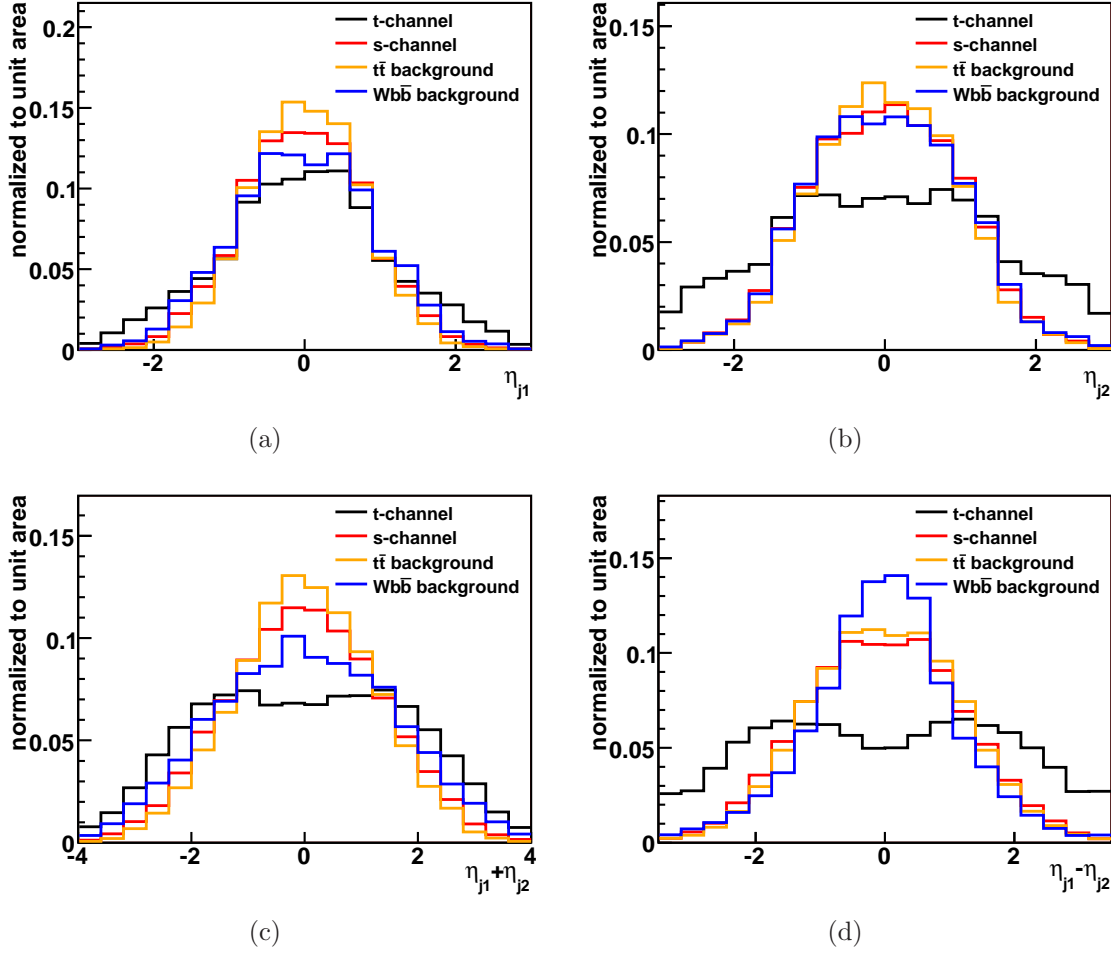


Figure 7.1: Input variables for the training of the neural networks used to discriminate single-top-quark events: (a) the pseudorapidity of the leading jet j_1 , (b) the pseudorapidity of the second-leading jet j_2 , (c) the sum of the pseudorapidities of the two jets $\eta_{j1} + \eta_{j2}$, and (d) the difference between the pseudorapidities of the two jets $\eta_{j1} - \eta_{j2}$. t - and s -channel events are compared with $t\bar{t}$ events and, exemplarily for W +jets, $Wb\bar{b}$ events.

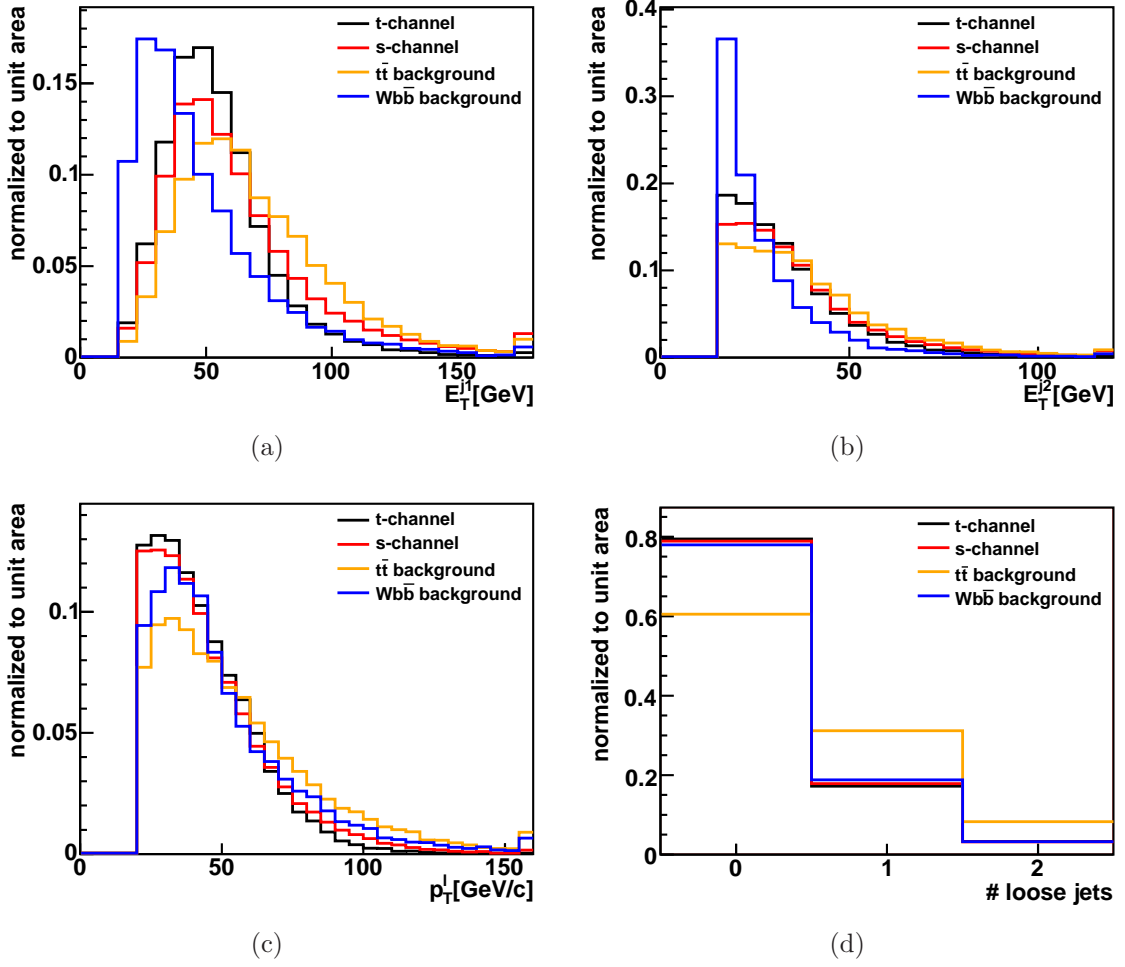


Figure 7.2: Input variables for the training of the neural networks used to discriminate single-top-quark events: (a) the transverse energy of the leading jet $j1$ (is only shown for completeness but, due to the automatic variable selection described in section 3.1.3, is not used for the training), (b) the transverse energy of the second-leading jet $j2$, (c) the transverse momentum of the tight lepton, and (d) the number of loose jets. t - and s -channel events are compared with $t\bar{t}$ events and, exemplarily for W +jets, $Wb\bar{b}$ events.

single-top-quark events whereas $t\bar{t}$ events feature the hardest jet E_T spectra. It is important to note that the distribution of the transverse energy of the leading jet j_1 is only shown for the sake of completeness, since it is rejected by the automatic variable selection.

In addition to the transverse jet energies, the transverse momentum of the tight lepton p_T^ℓ provides information about the generating physics process: as demonstrated in figure 7.2(c), leptons in single-top-quark events have softer p_T spectra than those in $Wb\bar{b}$ events, while leptons in $t\bar{t}$ events have a slightly harder spectrum.

An important variable being a measure of the total activity per event is the number of loose jets, depicted in figure 7.2(d). As expected, about 80% of the single-top-quark and $Wb\bar{b}$ events do not contain any loose jets, while about 40% of the $t\bar{t}$ events contain at least one loose jet.

Besides those directly measured variables, some reconstructed quantities are considered. One example is the invariant mass of the two jets, $M_{j_1j_2}$, illustrated in figure 7.3(a). This variable reflects the fact that both jets in $Wb\bar{b}$ events stem from the same object, namely a gluon splitting into a $b\bar{b}$ pair, while they originate from different objects in single-top-quark and $t\bar{t}$ events, leading to larger invariant masses.

Figure 7.3(b) depicts the mass of the top quark reconstructed out of the measured tight lepton, the reconstructed neutrino, and a selected b -quark jet as described in section 7.1.1. It is obvious that the invariant mass between these three objects is smaller in $Wb\bar{b}$ events, since those kinds of events do not contain any top quarks, while the reconstructed invariant mass peaks at the expected value in case of single-top-quark and $t\bar{t}$ events. Nevertheless, the distribution of the reconstructed top-quark mass in s -channel events is broader than the one in t -channel events, induced by larger ambiguities due to the selection of the b -quark jet. This equivocality, caused by the fact that there are two b quarks in s -channel events, has an even larger impact in $t\bar{t}$ events having a larger jet multiplicity.

An important variable to discriminate t -channel events is the product of the charge of the tight lepton Q_ℓ and the pseudorapidity of the light-quark jet $Q_\ell \cdot \eta_{lj}$ [8]. In this connection, the light-quark jet is the jet which has not been assigned to the decay of the top quark. As can be seen in figure 7.3(c), this variable is very asymmetric for t -channel events, while it is rather balanced around zero for all other processes. This asymmetry, induced by the parton distribution function of the proton, can be explained by the Feynman graphs shown in figure 1.4. The light-quark jet is produced via emission of a W boson from a u -type or d -type valence quark, while the second participating initial-state parton, a b quark stemming from gluon splitting, is a sea quark carrying a significantly smaller fraction of the proton momentum (see figure 1.2). As already mentioned in section 1.2.2, this leads to the fact that the light-quark jet has a strong tendency of propagating in the direction of the u -type or d -type valence quark. Since the proton consists of two u quarks and only one d quark, a top quark is mostly produced by an initial-state u -quark stemming from the proton, while the \bar{d} originating from the antiproton makes a smaller contribution to the production of top quarks. Hence, most of the top quarks propagate in proton

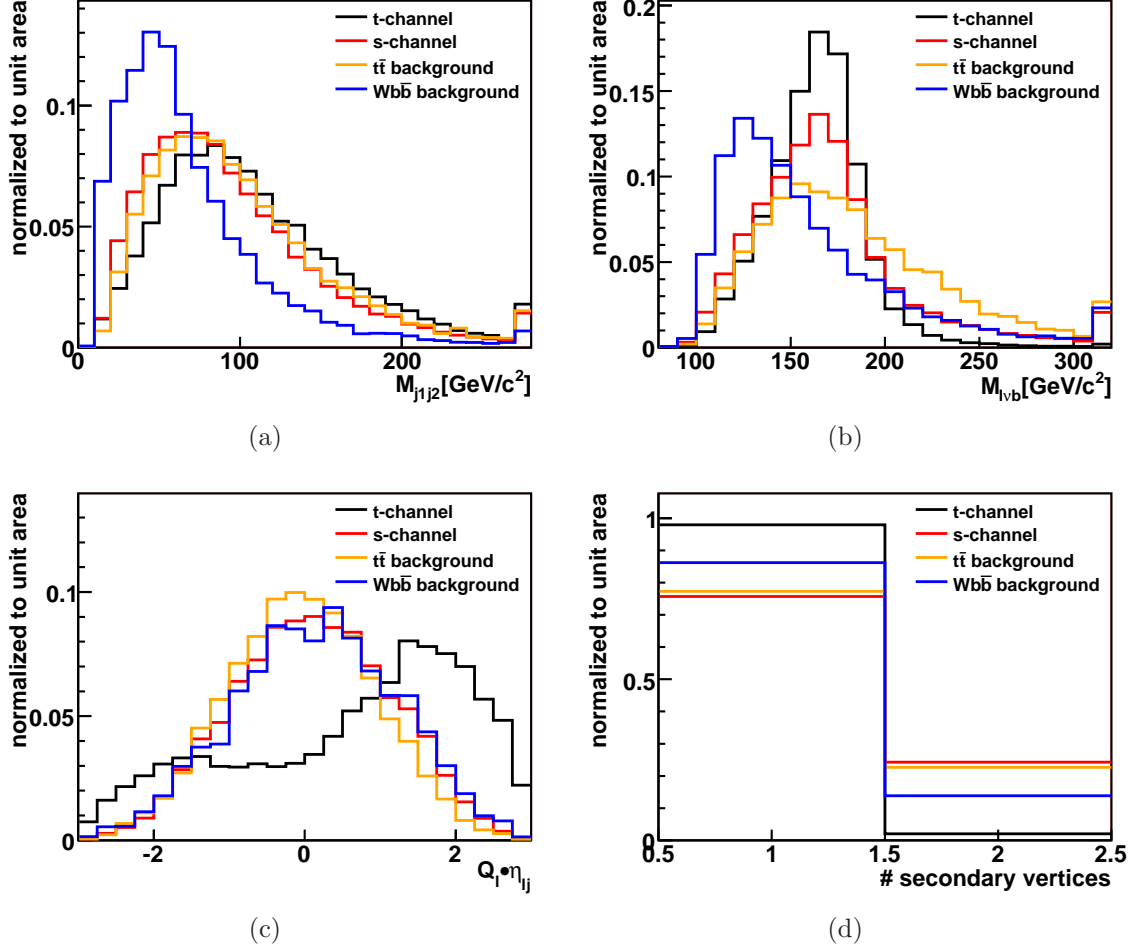


Figure 7.3: Input variables for the training of the neural networks used to discriminate single-top-quark events: (a) the invariant mass of the two jets, (b) the invariant mass of the reconstructed top quark (see section 7.1.1), (c) the product of the charge of the tight lepton Q_ℓ and the pseudorapidity of the jet not assigned to the decay of the top quark, and (d) the number of reconstructed secondary vertices. t - and s -channel events are compared with $t\bar{t}$ events and, exemplarily for W +jets, $Wb\bar{b}$ events.

direction, being reflected in the larger peak in the positive regime of the $Q_\ell \cdot \eta_{lj}$ distribution. For antitop-quark production, the charge conjugate case applies which is compensated by the multiplication of the pseudorapidity by the charge of the lepton.

Since some of the considered processes contain more than one b quark, it is important to account for the number of reconstructed secondary vertices. As illustrated in figure 7.3(d), most t -channel events have only one jet with a reconstructed secondary vertex, while both jets are tagged in about 20% of s -channel and $t\bar{t}$ events. In $Wb\bar{b}$ events, the fraction of double-tagged events is slightly smaller than in case of s -channel and $t\bar{t}$ events, even though they contain two b quarks as well. This can be attributed to the fact that there is a certain probability for both b quarks to be embedded in the same jet.

Another set of variables is motivated by reference [126] where studies have shown that the difference between the rapidity of the top quark and the rapidity of the light-quark jet is correlated to the sum of those two quantities. The difference of the rapidities is related to the scattering angle in the parton-parton center-of-mass system, while the sum reflects the boost of the parton-parton center-of-mass system with respect to the laboratory system. Due to this reason, $\eta_{\text{top}} + \eta_{lj}$ and $\eta_{\text{top}} - \eta_{lj}$ are utilized in this analysis. As demonstrated in figures 7.4(a) and 7.4(b), those variables do not provide much separation power when considered separately. Nevertheless, the correlations between those two variables are quite specific for t -channel events compared to s -channel, $t\bar{t}$, and $Wb\bar{b}$ as illustrated in figures 7.4(c) to 7.4(f), raising the discrimination power of those input variables.

To account for the total amount of energy in each event, the scalar sum of the transverse energy of tight lepton, the transverse energies of all tight and loose jets, and the missing transverse energy, $H_T = \sum_{\text{jets}} E_T + p_T^\ell + \cancel{E}_T$, is suitable. Figure 7.5(a) demonstrates that $t\bar{t}$ events have a tendency towards larger values of H_T , while the distribution of $Wb\bar{b}$ is slightly shifted to smaller values compared to single-top-quark events.

The physics process which created the W boson can be ascertained by the pseudorapidity of the reconstructed W boson. Looking at figure 7.5(b), it can be seen that W bosons originating from decaying top quarks in $t\bar{t}$ events are more central than those in single-top-quark events. Furthermore, W bosons in $Wb\bar{b}$ events have a broader η distribution slightly expanding into the forward region.

As already mentioned in section 1.2.2, the top quark is almost exclusively produced with negative helicity. In t -channel events, this spin polarization can be observed by examining the cosine of the angle between the light-quark jet and the lepton in the top-quark rest frame, as demonstrated in figure 7.5(c). In s -channel events, the best basis to analyze the polarization of the top-quark (antitop-quark) spin is the cosine of the angle between the lepton and the positive (negative) beam axis in the top-quark rest frame. Due to ambiguities in reconstruction of the top-quark rest frame, the polarization of s -channel single-top-quarks is smeared out in the distribution of this variable, as illustrated in figure 7.5(d). Nevertheless, this variable provides

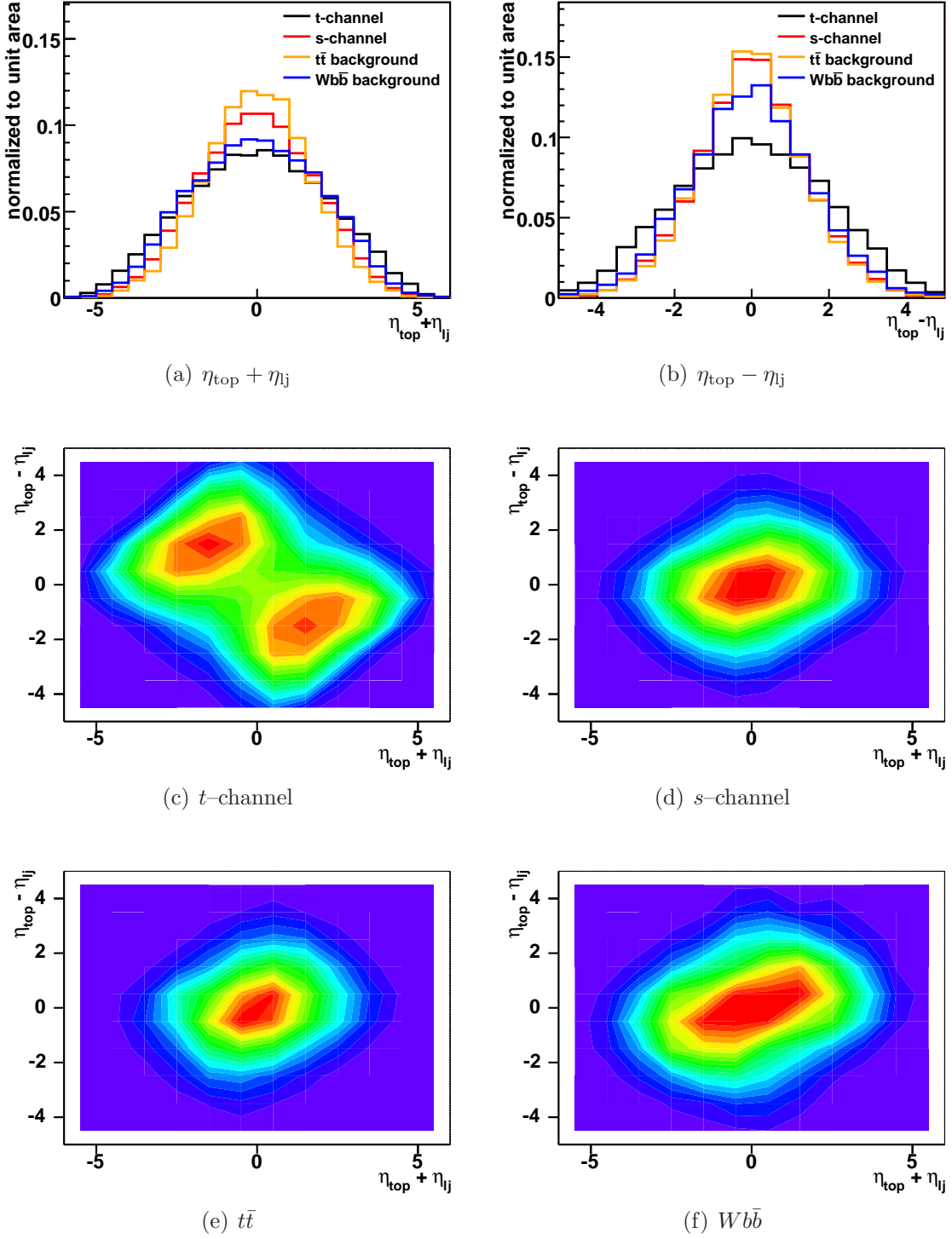


Figure 7.4: Input variables for the training of the neural networks used to discriminate single-top-quark events: (a) the sum of the pseudorapidities of the top quark and the light-quark jet (the jet not assigned to the top-quark decay) $\eta_{\text{top}} + \eta_{\text{lj}}$, (b) the difference between the pseudorapidity of the top quark and the pseudorapidity of the light-quark jet $\eta_{\text{top}} - \eta_{\text{lj}}$, and the dependence of $\eta_{\text{top}} - \eta_{\text{lj}}$ on $\eta_{\text{top}} + \eta_{\text{lj}}$ for (c) t -channel events, (d) s -channel events, (e) $t\bar{t}$ events, and (f) $Wb\bar{b}$ events. The choice of these variables is motivated by reference [126].

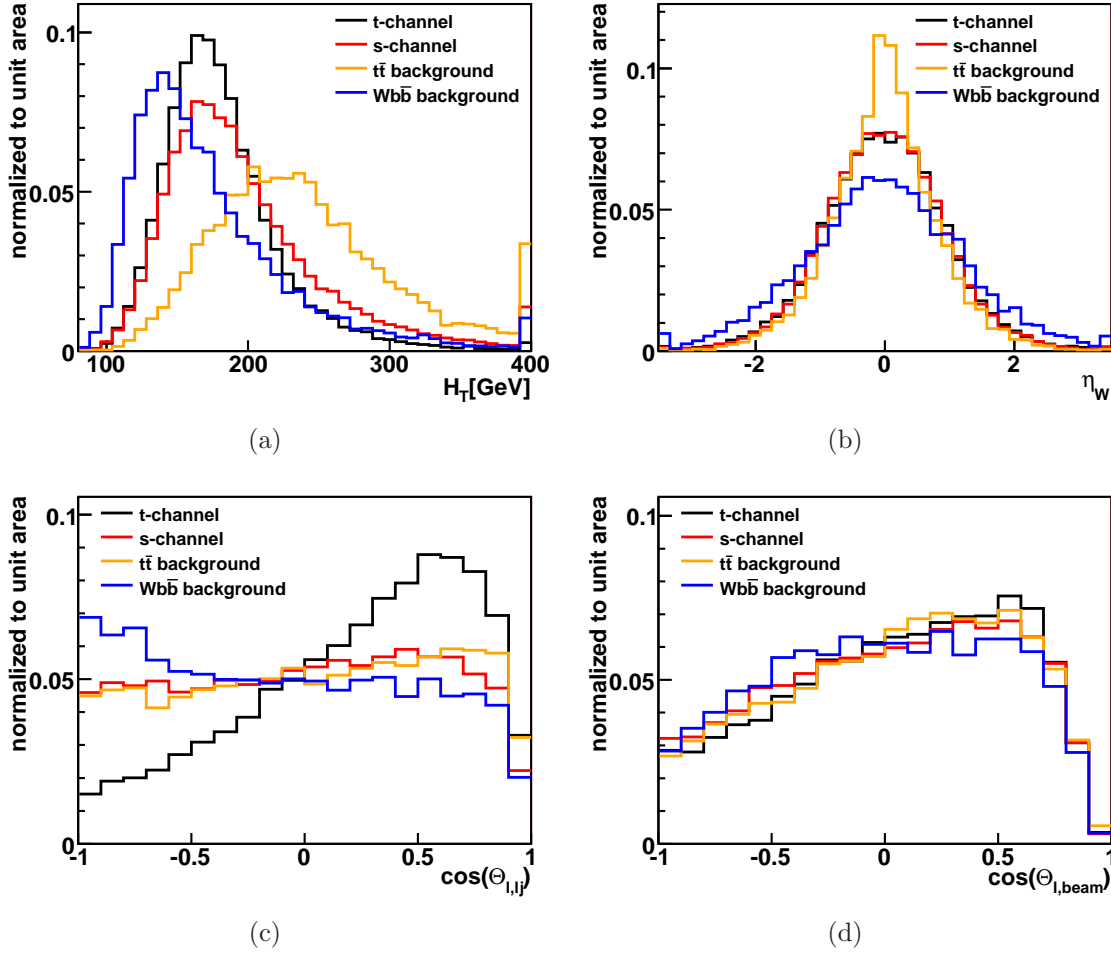


Figure 7.5: Input variables for the training of the neural networks used to discriminate single-top-quark events: (a) $H_T = \sum_{\text{jets}} E_T + p_T^\ell + \cancel{E}_T$ where $\sum_{\text{jets}} E_T$ is the sum of the transverse energies of the tight and loose jets, (b) the pseudorapidity of the reconstructed W boson (see section 7.1.1), (c) the cosine of the angle in the top-quark rest frame between the tight lepton and the jet not assigned to the top-quark decay, and (d) the cosine of the angle between the tight lepton and the beam axis in the top-quark rest frame. t - and s -channel events are compared with $t\bar{t}$ events and, exemplarily for W +jets, $Wb\bar{b}$ events.

enough information to be incorporated in the training of the neural network.

Some of the input variables are calculated by advanced algorithms. One example is the output of the neural-network b -tagger described in section 6, trained to identify secondary vertices related to the decay of a b hadron. As already shown in figure 6.21, the output value of this b tagger is continuously distributed between -1 and 1 . In this distribution, events containing b hadrons (like t -channel, s -channel, $t\bar{t}$, and $Wb\bar{b}$) pile up at values around 1 , while events without any b hadrons accumulate at outputs around -1 . Additionally, a small separation between c -quark and light-quark jets can be observed. For the training of the neural networks used to discriminate single-top-quark events, the neural-network b -tagger output of the most energetic tagged jet is taken.

Another set of variables is calculated by the kinematic fitter introduced in section 7.1.2. As a measure of the top-quark reconstruction potential, all four χ^2 values calculated using equation (7.3) are considered. In this context, the assignment of the b -quark jet to the top-quark decay and the selection of p_z^ν is the following:

	b -quark jet	p_z^ν solution
χ_1^2	$j1$	larger p_z^ν solution
χ_2^2	$j2$	larger p_z^ν solution
χ_3^2	$j1$	smaller p_z^ν solution
χ_4^2	$j2$	smaller p_z^ν solution

It can be seen that all possible combinations of the objects in $Wb\bar{b}$ events lead to higher χ^2 values than in case of events containing a real top quark. In addition to the four χ^2 values, the difference between the best χ_{best}^2 and the second-best $\chi_{2\text{ndbest}}^2$ value is used as illustrated in figure 7.6. In this distribution, the peak at -10 represents the underflow bin, i.e. those events with only one solution for p_z^ν , leading to no difference between χ_{best}^2 and $\chi_{2\text{ndbest}}^2$.

Comparison between Observation and Simulation

Since the neural network is trained with simulated events, it is crucial to check if the input variables are modeled correctly. Hence it is necessary to compare the shape of each input variable in observed events with the shape obtained by modeled data embracing the signal and background models described in section 5.3. For this comparison, each modeled process is scaled in such a way that it contributes as many events to the compound model as predicted by the estimated event yield quoted in table 5.2.

Exemplarily, figure 7.7 shows comparisons of the distributions of the output of the neural-network b -tagger, the reconstructed top-quark mass $M_{\ell\nu b}$, the invariant mass of the two jets M_{j1j2} , and the product of the lepton charge times the pseudorapidity of the jet not assigned to the decay of the top quark $Q_\ell \cdot \eta_{lj}$. The comparison of all other utilized variables can be found both in appendix D.1 and in reference [127]. In general, good agreement between the distributions in observed events and those

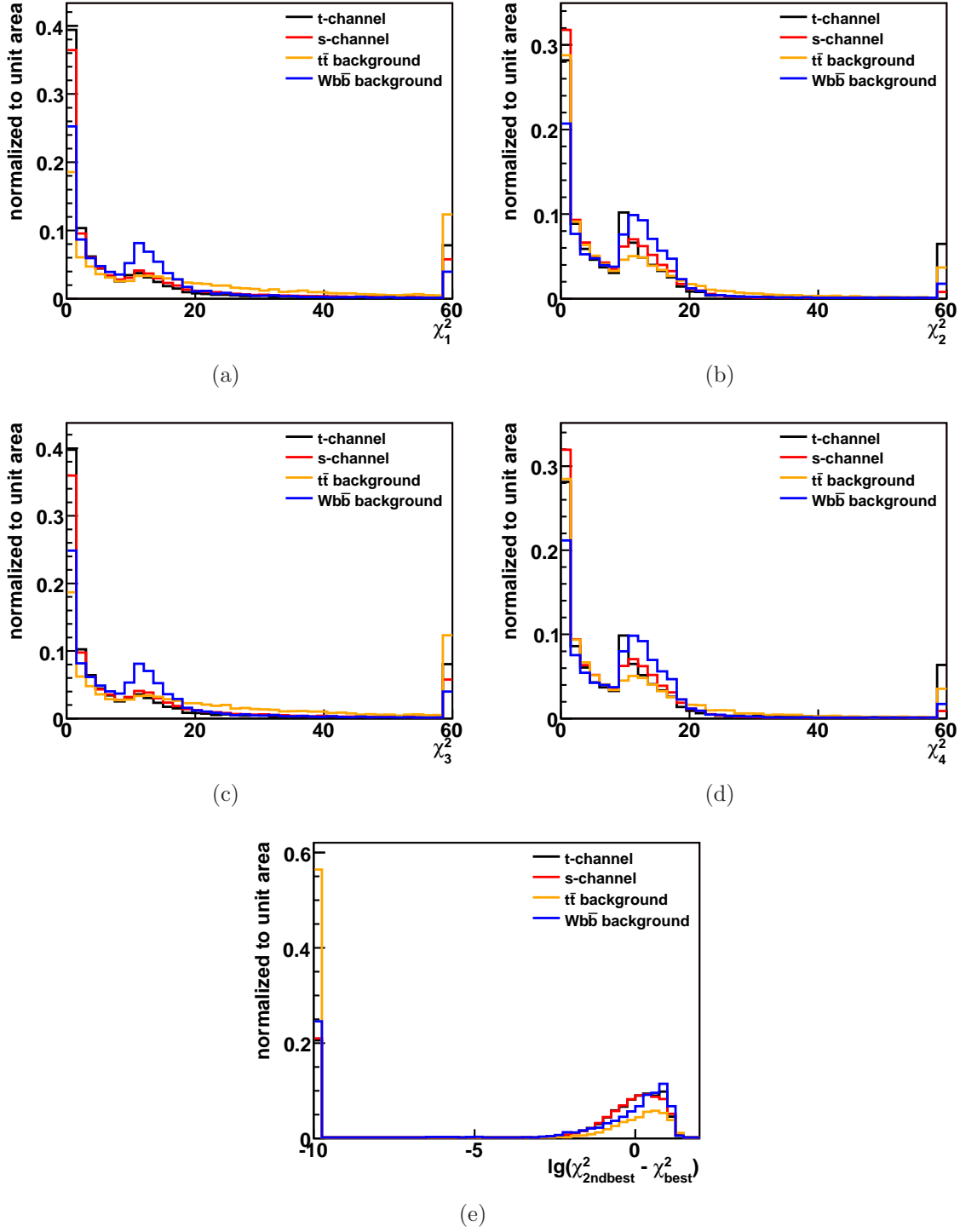


Figure 7.6: Input variables for the training of the neural networks used to discriminate single-top-quark events, calculated by the kinematic fitter introduced in section 7.1.2: the χ^2 values of the four possible combinations of the objects to reconstruct the top quark: (a) χ_1^2 , (b) χ_2^2 , (c) χ_3^2 , (d) χ_4^2 , and (e) the logarithm of the difference between the χ^2 values of the second-best combination and the best combination $\chi_{2ndbest}^2 - \chi_{best}^2$. t - and s -channel events are compared with $t\bar{t}$ events and, exemplarily for W +jets, $Wb\bar{b}$ events.

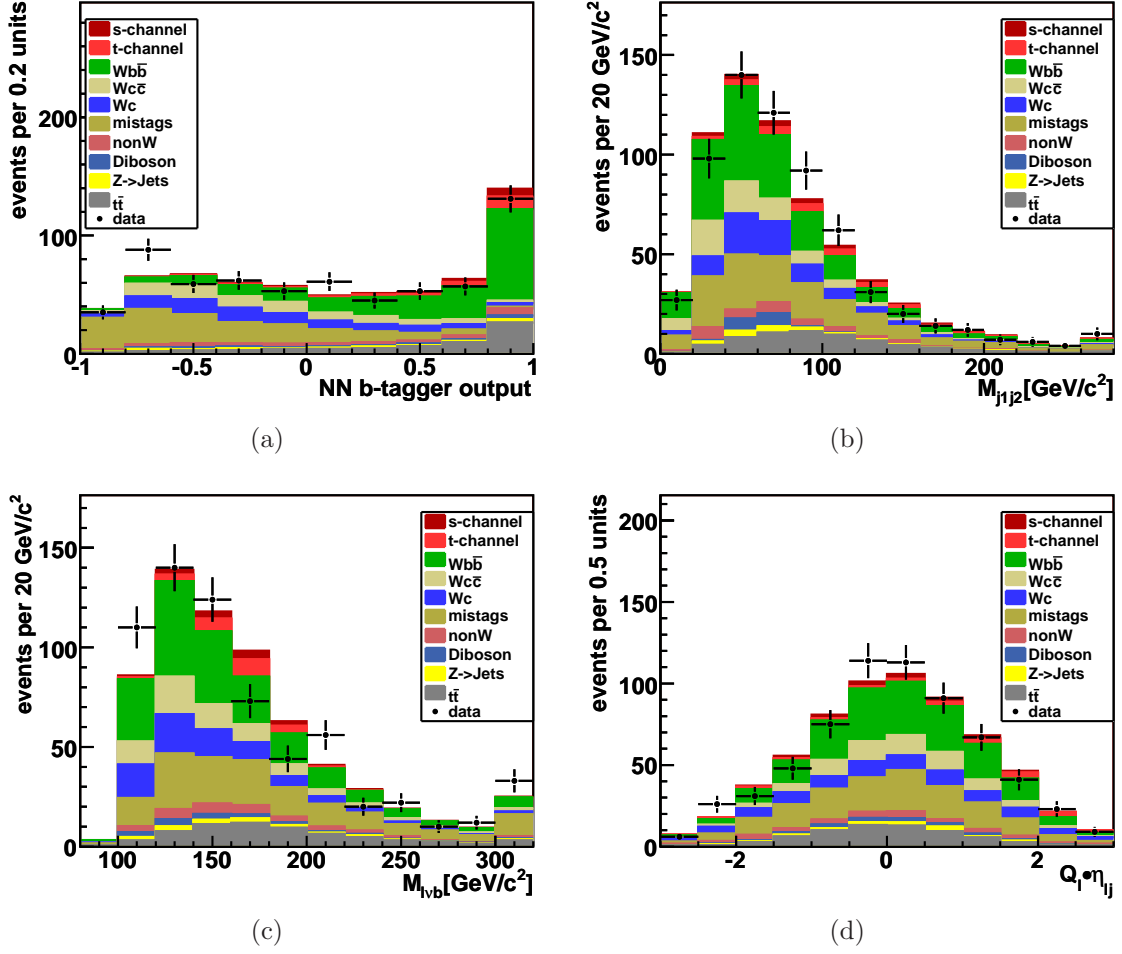


Figure 7.7: Data-MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events: (a) the output of the neural-network b -tagger of the most energetic tagged jet, (a) the invariant mass of the two jets, (b) the invariant mass of the reconstructed top quark (see section 7.1.1), and (c) the product of the charge of the tight lepton Q_ℓ and the pseudorapidity of the jet not assigned to the decay of the top quark. The modeled distributions are scaled to the number of observed events.

obtained by the composite model is achieved. This leads to the conclusion that the signal and background models described before are suitable to train the neural networks.

7.2 Training of the Neural Networks

For the training of the three networks, it is necessary to arrange training samples consisting of the relevant modeled physics processes, each with reasonable statistics. A natural approach would be to select a mixture corresponding to the estimated composition of the observed events. Since this is not practicable in case of the search for single top–quark production, a different composition of the training samples has to be chosen as will be discussed in section 7.2.1.

Before the training, the input variables are preprocessed as discussed in section 3.1.2 and the respective correlations to the target are determined, providing an ordered list of input variables being presented in section 7.2.2. In this connection, only variables with a significance larger than 3σ are utilized for the training. An illustration of the correlations between the considered variables can be found in appendix D.2.

Each training results in a continuous output distributed between -1 and 1 . At this, the output distribution of signal events piles up at 1 , while background events accumulate at -1 . In section 7.2.3, the output distributions of the different physics processes relevant for the search for single top–quark production are presented. Exemplarily, the distributions of the combined network are discussed.

7.2.1 Training Samples

For the training of the three neural networks, a realistic mixture of all expected processes would be a desirable approach. However, since the number of expected single-top–quark events is quite small compared to the number of expected background events, this concept is not realizable.

The samples used to train the three neural networks are composed in such a way that both single-top–quark production–processes add up to 35%. For the training of the combined neural network, the respective contributions of both single-top–quark production–modes are given by the cross sections predicted by the standard model. In the separate search, s –channel events are treated as background for the training of the t –channel neural network and vice versa. In both trainings, the respective background single-top–quark production–channel contributes 1% to the total fraction of single-top–quark events.

The individual contributions of all other background processes are alike for both approaches, separate and combined search, and add up to 65%. The fractions of all considered processes relative to this total contribution are given by the respective number of expected events quoted in table 5.2.

Process	combined search	t -channel	s -channel
t -channel $2 \rightarrow 2$	15.0%	23.0%	0.5%
t -channel $2 \rightarrow 3$	6.0%	11.0%	0.5%
s -channel	14.0%	1.0%	34.0%
$t\bar{t}$		7.3%	
WW		0.7%	
WZ		1.0%	
ZZ		0.0%	
Z +jets, $Z \rightarrow ee$		0.1%	
Z +jets, $Z \rightarrow \mu\mu$		0.2%	
Z +jets, $Z \rightarrow \tau\tau$		1.2%	
$Wb\bar{b}$, $W \rightarrow e\nu$		12.6%	
$Wb\bar{b}$, $W \rightarrow \mu\nu$		8.2%	
$Wb\bar{b}$, $W \rightarrow \tau\nu$		0.4%	
$Wc\bar{c}$, $W \rightarrow e\nu$		5.7%	
$Wc\bar{c}$, $W \rightarrow \mu\nu$		3.2%	
$Wc\bar{c}$, $W \rightarrow \tau\nu$		0.2%	
Wc , $W \rightarrow e\nu$		4.0%	
Wc , $W \rightarrow \mu\nu$		3.2%	
Wc , $W \rightarrow \tau\nu$		0.3%	
mistags, $W \rightarrow e\nu$		10.2%	
mistags, $W \rightarrow \mu\nu$		6.0%	
mistags, $W \rightarrow \tau\nu$		0.7%	

Table 7.1: Composition of the training samples used to train the neural networks to discriminate single-top-quark events. The contributions of single-top-quark events add up to 35% of the complete training sample. The contributions of the background processes are the same for both approaches, separate and combined search. The relative fractions to the background category are given by table 5.2. The absolute fractions are determined by the requirement that the contributions of all background processes sum up to 65%. The contributions of the different W -boson decay-modes in case of W +jet events have been obtained from the acceptance of simulated events.

Rank	Variable	Relative Significance (in σ)
1	NN b -tag output	42.7
2	$M_{\ell\nu b}$	35.3
3	M_{j1j2}	28.0
4	$Q_\ell \cdot \eta_{lj}$	18.8
5	# loose jets	14.8
6	H_T	10.4
7	$\cos \Theta(\ell, lj)$	11.3
8	η_W	9.3
9	$\lg(\chi_{2\text{ndbest}}^2 - \chi_{\text{best}}^2)$	8.9
10	p_T^ℓ	7.6
11	# secondary vertices	7.0
12	χ_3^2	5.5
13	χ_2^2	5.2
14	$\eta_{j1} + \eta_{j2}$	4.8
15	E_T^{j2}	4.0
16	$\eta_{\text{top}} - \eta_{lj}$	2.9
17	$\eta_{j1} - \eta_{j2}$	3.7
18	η_{j1}	3.0
19	$M_{\ell\nu bb}$	2.7
20	$\eta_{\text{top}} + \eta_{lj}$	2.1
21	χ_1^2	2.1
22	$\cos \Theta(\ell, \text{beam})$	1.8
23	χ_4^2	1.9
24	η_{j2}	1.7
25	E_T^{j1}	0.9
26	$E_T^{j1} + E_T^{j2}$	0.3

Table 7.2: Set of discriminating variables investigated for the training of the combined neural network. The quoted relative significances are determined as described in section 3.1.3, i.e. calculated in terms of reduced matrices. Only variables with a significance $> 3\sigma$ are used for the training.

A list of all used processes, modeled as described in section 5.3, as well as their contributions to the training samples is given in table 7.1.

7.2.2 Relevance of Input Variables

As already mentioned, 26 variables have initially been investigated for the training of the neural networks. Making use of the automatic variable selection, only variables are used for the training which have a significance of more than 3σ , calculated as described in section 3.1.3. Since the relevance of the variables varies for the three trainings, each neural network utilizes a different set of input variables passing the cut on the significance, quoted in tables 7.2, 7.3, and 7.4.

Rank	Variable	Relative Significance (in σ)
1	$M_{\ell\nu b}$	50.1
2	$Q_\ell \cdot \eta_j$	41.1
3	NN b -tag output	29.9
4	$\cos \Theta(\ell, l_j)$	21.2
5	$M_{j_1 j_2}$	13.1
6	# loose jets	10.9
7	H_T	11.3
8	η_W	9.0
9	# secondary vertices	8.1
10	$\eta_{j1} + \eta_{j2}$	6.8
11	$\lg(\chi_{2\text{ndbest}}^2 - \chi_{\text{best}}^2)$	6.0
12	p_T^ℓ	4.9
13	χ_3^2	4.2
14	E_T^{j2}	4.1
15	$\eta_{\text{top}} - \eta_j$	3.1
16	$\eta_{j1} - \eta_{j2}$	4.3
17	$\cos \Theta(\ell, \text{beam})$	3.3
18	χ_2^2	3.0
19	η_{j2}	3.1
20	χ_1^2	2.1
21	$\eta_{\text{top}} + \eta_j$	2.0
22	$M_{\ell\nu bb}$	1.7
23	χ_4^2	1.2
24	E_T^{j1}	0.5
25	$E_T^{j1} + E_T^{j2}$	1.1
26	η_{j1}	0.2

Table 7.3: Set of discriminating variables investigated for the training of the t -channel neural network. The quoted relative significances are determined as described in section 3.1.3, i.e. calculated in terms of reduced matrices. Only variables with a significance $> 3\sigma$ are used for the training.

Rank	Variable	Relative Significance (in σ)
1	NN b -tag output	41.1
2	# secondary vertices	27.4
3	M_{j1j2}	21.8
4	$M_{\ell\nu b}$	20.0
5	# loose jets	14.5
6	η_W	11.4
7	χ_2^2	10.6
8	p_T^ℓ	10.3
9	χ_3^2	9.3
10	E_T^{j2}	7.9
11	$\eta_{j1} - \eta_{j2}$	6.6
12	$\lg(\chi_{2\text{ndbest}}^2 - \chi_{\text{best}}^2)$	5.9
13	H_T	5.2
14	$\eta_{\text{top}} - \eta_{\text{j}}$	4.3
15	χ_4^2	3.9
16	χ_1^2	4.0
17	$\eta_{j1} + \eta_{j2}$	2.2
18	$\eta_{\text{top}} + \eta_{\text{j}}$	3.7
19	$\cos \Theta(\ell, l_j)$	3.6
20	$M_{\ell\nu bb}$	1.6
21	$E_T^{j1} + E_T^{j2}$	1.3
22	$Q_\ell \cdot \eta_{lj}$	1.2
23	η_{j2}	1.5
24	$\cos \Theta(\ell, \text{beam})$	1.4
25	η_{j1}	0.4
26	E_T^{j1}	0.1

Table 7.4: Set of discriminating variables investigated for the training of the s -channel neural network. The quoted relative significances are determined as described in section 3.1.3, i.e. calculated in terms of reduced matrices. Only variables with a significance $> 3\sigma$ are used for the training.

For all three trainings, the output of the neural-network b -tagger, the reconstructed top-quark mass $M_{\ell\nu b}$, the invariant mass of the two jets $M_{j_1 j_2}$, and the number of loose jets are some of the most important variables. Additionally, $Q_\ell \cdot \eta_{lj}$ and $\cos \Theta(\ell, lj)$ are of major importance in both networks considering t -channel events as signal, i.e. the combined and the t -channel network, while they are negligible in case of the s -channel neural network. In contrast, the number of secondary vertices is a significant quantity to identify s -channel events.

Three out of the 26 variables are not utilized in any of the neural networks, since they do not provide significant additional information. These variables are the transverse energy of the most energetic jet E_T^{j1} , the sum of the transverse energies of both jets $E_T^{j1} + E_T^{j2}$, and the invariant mass of lepton, missing transverse energy and the two jets $M_{l\nu bb}$ reconstructed under the assumption of s -channel events (two b -quark jets). Altogether, 23 variables are used for the training of all three neural networks whereof 18 variables are employed by the combined network, while both t - and s -channel networks utilize different subsets consisting of 19 variables.

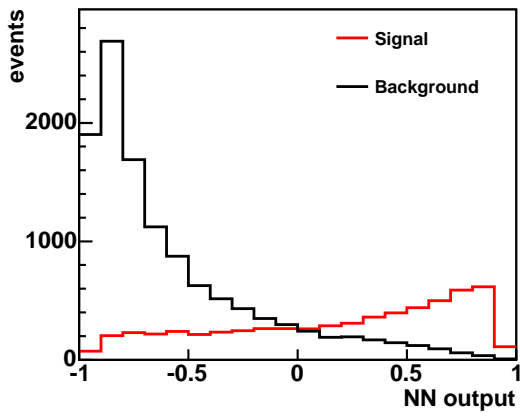
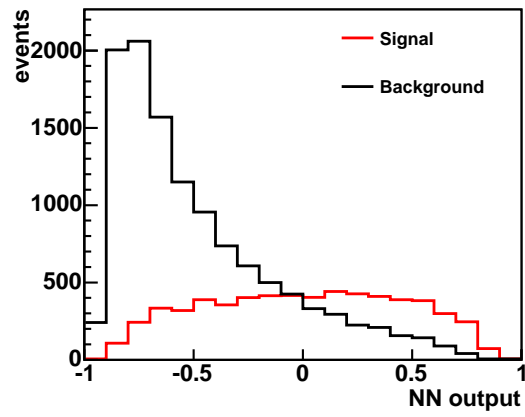
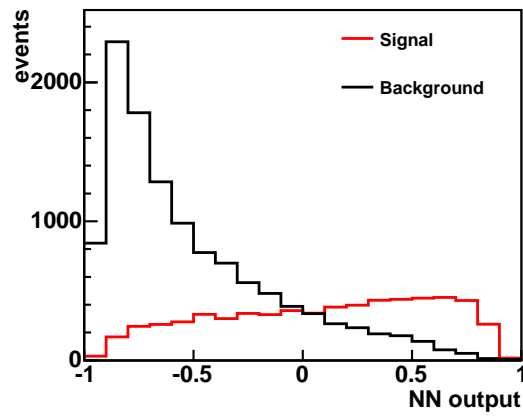
7.2.3 Output Distributions and Templates

The training of a neural network results in one output variable continuously distributed between -1 and 1 . Figure 7.8 illustrates the result of the training of the three networks utilized in the search for single top-quark production. Figure 7.8(a) depicts the output of the t -channel neural network applied to the mixed sample used for the training, featuring an efficient separation between t -channel single-top-quark events and the background processes, induced by both excellent background rejection and reasonable signal identification. Nevertheless, it can be seen that there are obviously no signal events being identified with absolute doubtlessness.

Furthermore, figure 7.8(b) illustrates that the s -channel neural network provides less definite identification of signal events, compared to the t -channel neural network. The reason for this rather flat, indetermined signal output distribution around zero is that there are no variables featuring a powerful discrimination of s -channel events, as opposed to t -channel events.

The output of the neural network supposed to identify both t - and s -channel events as signal reflects the above mentioned features as shown in figure 7.8(c). The output distribution of single-top-quark events looks more signal-like than in case of the s -channel neural network but has a slightly less prominent peak in the signal region compared to the t -channel neural network.

The output of the different neural networks is used to create templates which are to be fitted to the output distribution of observed events. In general, it would be possible to create a template for each of the relevant processes. Nevertheless, comparing the different processes, it can be seen that some of the output distributions feature very similar shapes. As illustrated in figure 7.9 exemplarily for the combined search, three non-top templates are created: $Wb\bar{b}$ and WZ form the b -like template as illustrated in figure 7.9(a). The c -like template consists of $Wc\bar{c}$, Wc , WW , and mistags

(a) t -channel network(b) s -channel network

(c) combined network

Figure 7.8: The neural network outputs of the (a) t -channel neural network, (b) s -channel neural network, and (c) combined neural network in the respective training samples.

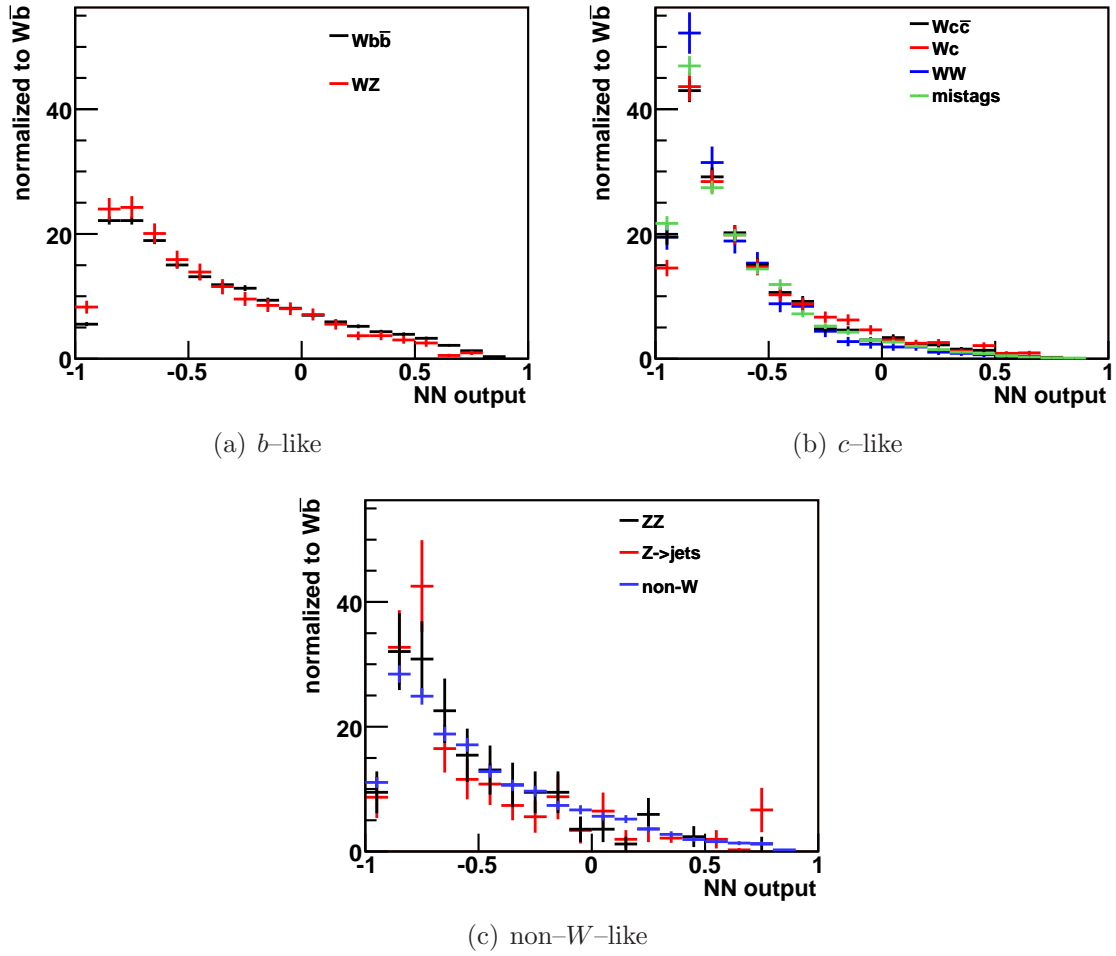


Figure 7.9: The construction of the different templates is illustrated exemplarily for the combined neural network. Due to similarity of the output distributions, three non-top templates are created: (a) b -like template, (b) c -like template, and (c) non- W -like template. For comparison, all distributions are scaled to the same area.

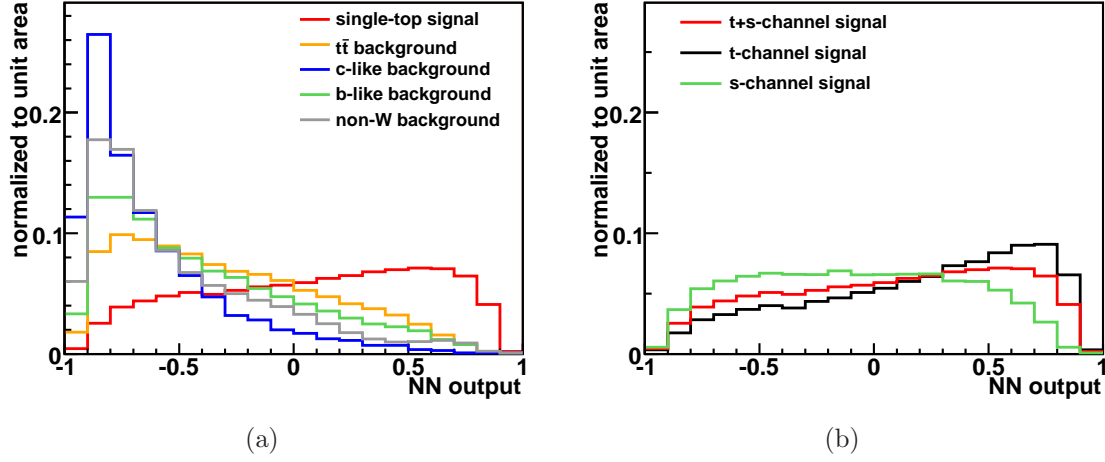


Figure 7.10: The templates utilized in the combined search are illustrated: (a) shows the different template distribution and (b) illustrates the combination of the t - and s -channel distributions into the single-top template. The templates are composed as illustrated in figure 7.9.

as shown in figure 7.9(b). In case of the third template, the non- W background is combined with Z +jets and ZZ as depicted in figure 7.9(c). Constructing the templates, the contributions of the individual processes are weighted according to the number of expected events. Thus, together with the $t\bar{t}$ template, a total number of four background templates is utilized in both combined and separate search.

A comparison of the five templates used in the combined search can be found in figure 7.10. As can be seen in figure 7.10(a), the c -like template features the most background-like distribution, while the distributions are getting less background-like with increasing b -quark content: since part of the events incorporated in the non- W -like template contain b quarks, this template distribution is slightly less background-like compared to the c -like template. This trend is even more pronounced in case of the b -like template mainly consisting of $Wb\bar{b}$ events. As expected due to the presence of top quarks, the $t\bar{t}$ template features the least background-like distribution, although being distinct from the distribution of the single-top template.

In case of the separate search, the individual templates are constructed out of the same processes, except t - and s -channel which are treated separately, leading to a total of six template distributions. Those templates are illustrated in figure 7.11 showing the output of the s -channel neural network versus the output of the t -channel neural network. The two-dimensional template distribution of t -channel events is depicted in figure 7.11(a), indicating that the output distribution of the t -channel network peaks at values around one, while the output of the s -channel network has a broad symmetric distribution around zero. The situation is different for s -channel events as illustrated in figure 7.11(b): the output distribution of the t -channel network looks very background-like, while the output of the s -channel network has broad distribution not significantly peaking in the signal region as

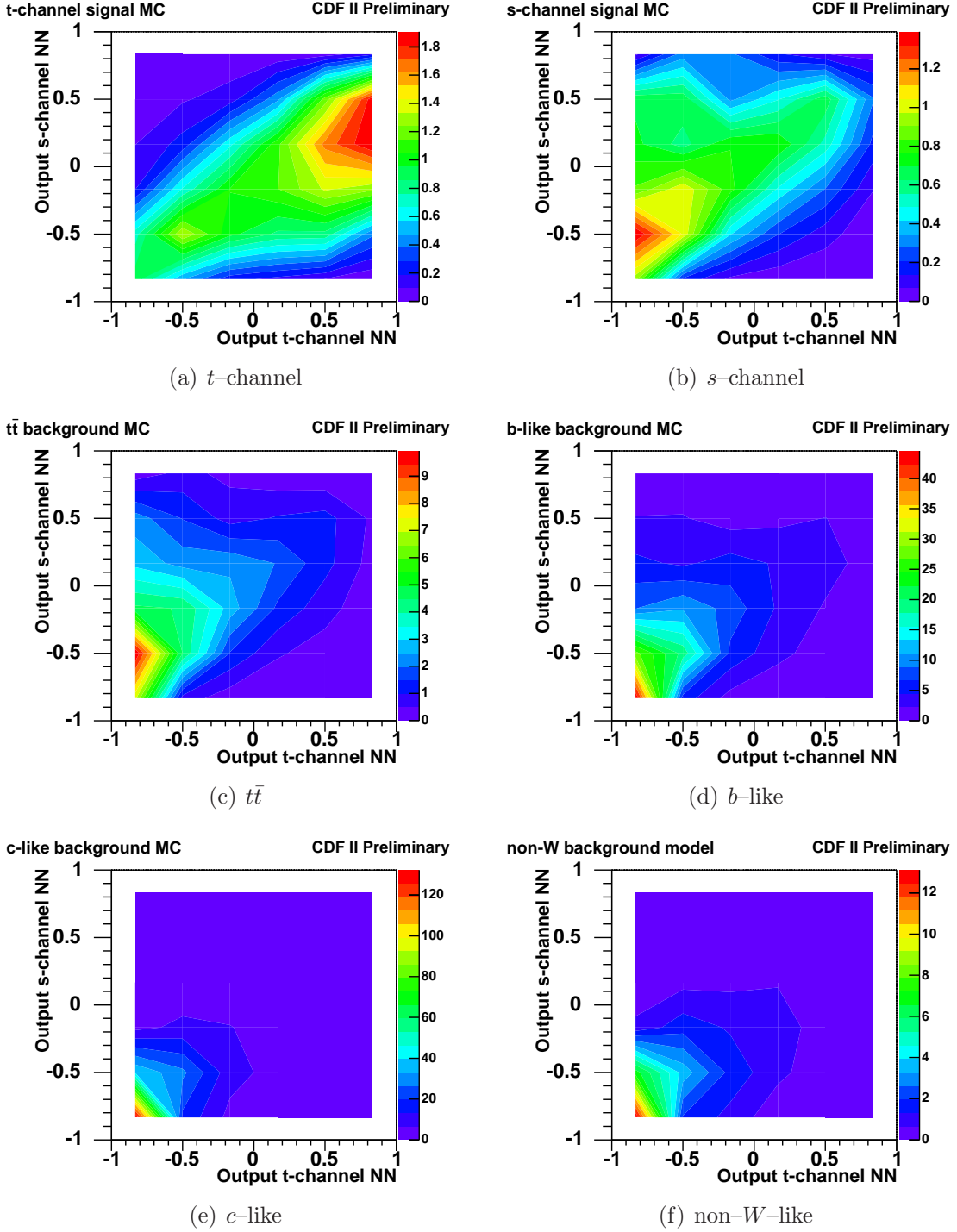


Figure 7.11: The templates utilized in the separate search are illustrated. The output of the s -channel network in dependence of the output of the t -channel network is shown for (a) t -channel events, (b) s -channel events, (c) $t\bar{t}$ events, (d) the combined b -like template, (e) the combined c -like template, and (f) the combined non- W -like template. The templates are composed as illustrated in figure 7.9.

already discussed before. Looking at figure 7.11(c), it can be seen that the t -channel neural network identifies $t\bar{t}$ events as quite background-like whereas their suppression works slightly worse in case of the s -channel network. Considering the remaining template categories illustrated in figures 7.11(d), 7.11(e), and 7.11(f), both networks classify them clearly as background. As already addressed in the discussion of the templates used in the combined search, the c -like template depicted in figure 7.11(e) has the most background-like distribution.

7.3 Systematic Uncertainties

Uncertainties in the modeling of physics processes and detector effects cause systematic uncertainties on the measurement results, affecting the rate of predicted signal and background events as well as the shape of the template histograms used in the fit to the observed data distribution. It must be noted that some effects induce only rate uncertainties, while some affect only the shape of the templates, even though most of them impact both.

The following sources of systematic uncertainties are considered: the uncertainty on the jet energy corrections, 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 in modeling the output of the neural-network b -tagger, the uncertainty in the factorization and renormalization scale for the simulation of W +heavy flavor processes, the modeling of instrumental backgrounds, that is mistag events and non- W events, and the uncertainty in the luminosity determination. The impact of these sources of uncertainties is evaluated by altering the modeling of the corresponding processes or effects within their uncertainties or by assigning a plausible alternative model. As a result, relative changes of the event rates and shifted template distributions are obtained. Rate uncertainties are only determined for single-top-quark and $t\bar{t}$ events because the rates of the main backgrounds, W +jets and non- W events, are estimated based on the observed rate of events before b tagging or by a fit to the observed missing transverse energy distribution, respectively. A more detailed discussion on the background estimation is given in section 5.4.

The effect of the uncertainty on the jet energy corrections is quantified by varying the corrections within their $\pm 1\sigma$ uncertainties [60]. The corresponding alternative template distributions are calculated for all signal and background processes. The influence of initial-state and final-state gluon radiation is estimated by producing samples of simulated events for which the simulation was altered to produce either less or more gluon radiation compared to the standard setting [128]. Specifically, two parameters controlling the parton shower in the PYTHIA program are varied: Λ_{QCD} and the scale factor K to the transverse momentum scale of the showering. The different settings are derived from studies of ISR in Drell-Yan events. Using these specific ISR and FSR samples of simulated events, alternative template shapes are

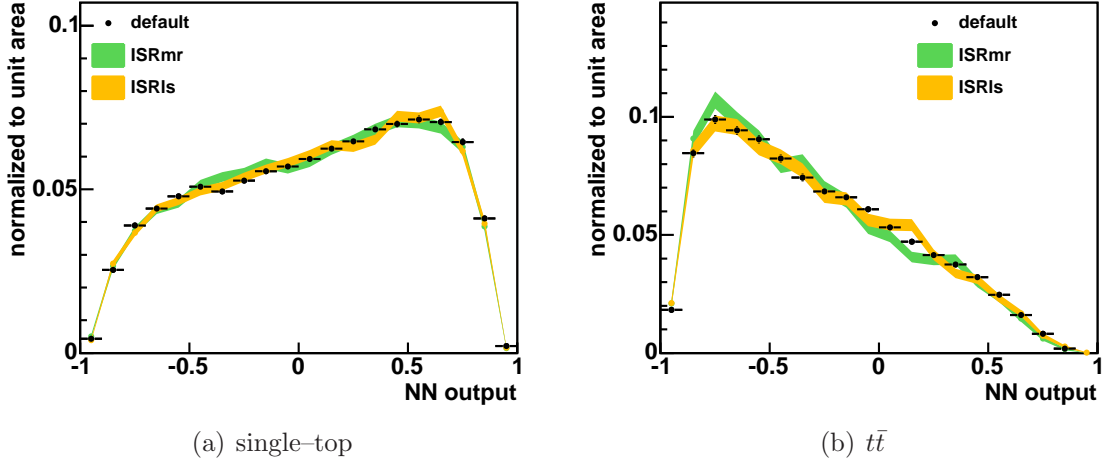


Figure 7.12: Shape variation due to the uncertainty on initial-state radiation. The template distributions for (a) single-top-quark and (b) $t\bar{t}$ events in the combined neural network analysis are shown. The default distributions are compared to the shifted distributions where the simulation was altered to produce either less or more initial-state gluon radiation compared to the standard setting [128].

produced for single-top-quark and $t\bar{t}$ events. As an example, figure 7.12 displays the specific ISR template distributions of single-top-quark and $t\bar{t}$ events in the combined neural network analysis.

The impact of the uncertainties on the PDF parameterization are studied by reweighting single-top-quark and $t\bar{t}$ events with weights associated with the 20 pairs of CTEQ6M eigenvectors. The rate uncertainty on the signal model which is based on the MADEVENT event generator is determined by a comparison to differential cross sections computed with the ZTOP program as described in section 5.3.1. The modeling of $t\bar{t}$ events is studied by using simulated events produced with the HERWIG program as an alternative generator. The shape uncertainties on the template histograms due to the modeling of W +heavy flavor events with the ALPGEN Monte Carlo program [75] are deduced by using version 2.0 of the program instead of the default version 1.3.3. The factorization and renormalization scale is varied in the simulation to derive an additional set of altered template histograms for W +heavy flavor events. The default W +jets Monte Carlo samples are generated with a dynamic scale $\mu^2 = Q^2$. To check the influence of a rather drastic systematic variation, the scale is fixed to $\mu^2 = m_W^2$.

The uncertainty in the event detection efficiency includes the uncertainties on the trigger efficiency, on the lepton identification efficiency, and on the b -tagging efficiency which is the dominating factor. Since no cut is applied on the output of the neural-network b -tagger, the uncertainty associated with this quantity does not imply a rate uncertainty, but only a shape uncertainty on the template distributions. Systematic effects are studied by utilizing the correction or systematic shift functions to define two scenarios, an optimistic one and a pessimistic one as described in section 6.4.1.

Source	t -channel	s -channel	single-top	$t\bar{t}$
jet energy corrections	2.1%	1.4%	1.8%	8.6%
initial-state radiation	2.6%	1.6%	2.2%	10.1%
final-state radiation	3.4%	1.4%	2.6%	11.4%
parton distribution functions	2.5%	2.2%	2.4%	2.4%
event generator	2.0%	1.0%	1.6%	3.0%
event detection efficiency	8.1%	6.3%	7.4%	7.1%
luminosity	6.0%	6.0%	6.0%	6.0%

Table 7.5: Relative systematic rate uncertainties for single-top-quark and $t\bar{t}$ events.

A modified model of non- W events is considered to investigate the influence of this aspect on the analysis. In comparison to the default model described in section 5.3, the alternative model uses events selected from a generic jet sample where one of the jets has a fraction of electromagnetic energy measured in the calorimeter of at least 0.8, but less than 0.95. This so-called jet-electron assumes the role of the charged lepton in the event. In addition, the flavor composition of the non- W sample is varied: the default model assumes a composition of 45% b -quark jets, 40% c -quark jets, and 15% light-quark jets, whereas the alternative model uses a composition of 60:30:10, respectively.

To evaluate the systematic effect on the shapes of the distributions caused by the modeling of mistagged light-quark jet events, an alternative model is utilized to create template distributions. This is realized by replacing the default mistag model based on simulated events as described in section 5.3 by a description on the basis of measured W +jets events before b tagging.

Table 7.5 summarizes the relative rate uncertainties on the number of expected single-top-quark and $t\bar{t}$ prediction events. The analyses are done under the assumption of a top-quark mass of $M_t = 175 \text{ GeV}/c^2$. That is why the uncertainty in the top-quark mass is not taken into account as a systematic uncertainty. Hence, the analyses provide rather a measurement at the specified value of the top-quark mass. However, if the top-quark mass is varied in the simulation by $\pm 5 \text{ GeV}/c^2$ the acceptance for single-top-quark events changes by $\pm 2.8\%$.

The influences of the sources of systematic uncertainties on the shape of the template distributions are visualized in appendix D.3, besides the examples presented in figure 7.12. For later usage in the template likelihood fit to the observed output distribution, all shape and rate uncertainties are symmetrized as illustrated in figure 7.13 exemplarily for the uncertainty on initial state radiation. The symmetrization is performed for each bin by calculating half the difference between the two altered scenarios. If there is only one systematic scenario, the full difference between the altered and default distribution is considered.

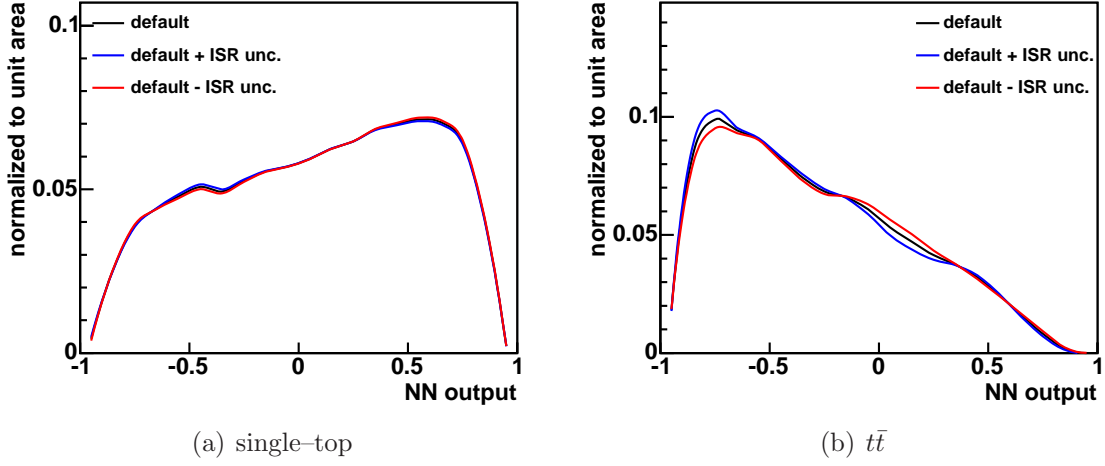


Figure 7.13: Symmetrized shape variation due to the uncertainty on initial-state radiation for (a) single-top-quark and (b) $t\bar{t}$ events in the combined neural network analysis. The default distributions are compared to the symmetrized variations. The symmetrization is performed for each bin by calculating half the difference between the two scenarios with altered initial-state radiation illustrated in figure 7.12.

7.4 Likelihood Function

Both combined and separate search use the same likelihood function for the template fit to the observed events. This likelihood function consists of Poisson terms for the individual bins of the fitted histogram, Gaussian constraints on the background rates, and Gaussian constraints on the strengths of systematic effects:

$$L(\beta_1, \dots, \beta_C; \delta_1, \dots, \delta_S) = \prod_{k=1}^B \frac{e^{-\mu_k} \cdot \mu_k^{n_k}}{n_k!} \cdot \prod_{j=A}^C G(\beta_j, 1.0, \Delta_j) \cdot \prod_{i=1}^S G(\delta_i, 0.0, 1.0) . \quad (7.4)$$

Systematic uncertainties are included as factors modifying the expectation value μ_k of events in a certain bin k , with

$$\mu_k = \sum_{j=1}^C \beta_j \cdot \hat{\nu}_j \cdot \left\{ \sum_{i=1}^S (1 + \delta_i \cdot \epsilon_{ji}) \right\} \cdot \alpha_{jk} \cdot \left\{ 1 + \sum_{i=1}^S (\delta_i \cdot \kappa_{jik}) \right\} , \quad (7.5)$$

$$G(\beta_j, 1.0, \Delta_j) = \frac{1}{\sqrt{2\pi} \Delta_j} \cdot \exp \left(\frac{-(\beta_j - 1.0)^2}{2 \Delta_j^2} \right) , \text{ and} \quad (7.6)$$

$$G(\delta_i, 0.0, 1.0) = \frac{1}{\sqrt{2\pi}} \cdot \exp \left(\frac{-\delta_i^2}{2} \right) . \quad (7.7)$$

The index j runs over the different event categories occurring in the likelihood function. In both approaches, four background categories are considered as introduced in section 7.2.3: $t\bar{t}$, b -like backgrounds, c -like backgrounds, and non- W -like events. In the combined search, the total number of event classes is $C = 5$, with $j = 1$

being the label for single-top-quark events. The labeling of background processes starts with $A = 2$. In the separate search, C is equal to 6, with $j = 1$ for t -channel single-top-quark events and $j = 2$ for s -channel single-top-quark events. The first background label is $A = 3$.

The predicted expectation values for the number of events of a certain event category are denoted $\hat{\nu}_j$, leading to $\hat{\nu}_1 = 37.8$ for the combined search and $\hat{\nu}_1 = 22.4$ and $\hat{\nu}_2 = 15.4$ for the separate search, respectively. The predicted expectation values for the background event categories as well as their relative uncertainties Δ_j are given in Table 7.6. The free parameters in the fit are given by $\beta_j = \nu_j/\hat{\nu}_j$, i.e. the

process	$\hat{\nu}_j$	Δ_j
$t\bar{t}$	58.4	19.7%
b -like	178.9	28.3%
c -like	273.7	16.1%
non- W	38.5	43.0%

Table 7.6: Predicted expectation values and their relative standard deviations used for the Gaussian background constraints in the likelihood function (7.4).

expectation values over their prediction. The normalized content of bin k of the template histogram for event category j is α_{jk} . The total number of bins is B .

In the fit, seven effects causing systematic rate uncertainties are considered as given by table 7.5. In this notation, the sources of systematic uncertainties carry the index i . The variation in strength of a systematic effect i is measured with the variable δ_i which constitutes an additional fit parameter and measures the strength of the systematic effect in units of one standard deviation. The relative rate uncertainties due to these sources are named ϵ_{ji} . As outlined in section 7.3, ten sources of uncertainties influencing the template shape are taken into account. Altogether, $S = 13$ sources of systematic uncertainties are thus considered. In the template distributions, the shape uncertainties are reflected by relative uncertainties in the bin content of bin k , being denominated as κ_{jik} . The values of κ_{jik} are calculated from the systematically shifted normalized template histograms α_{jik}^+ and α_{jik}^- according to

$$\kappa_{jik} = \frac{\alpha_{jik}^+ - \alpha_{jik}^-}{2 \alpha_{jk}} . \quad (7.8)$$

By construction the κ_{jik} satisfy the normalization condition

$$\sum_{k=1}^B \kappa_{jik} = 0 . \quad (7.9)$$

The systematically shifted template taking into account the shifts caused by all systematic effects with strengths $\{\delta_i\}$ is given by

$$\alpha'_{ji} = \alpha_{jk} \cdot \left\{ 1 + \sum_{i=1}^S \delta_i \cdot \kappa_{jik} \right\} . \quad (7.10)$$

Due to (7.9), the shifted histogram α'_{ji} is properly normalized:

$$\sum_{k=1}^B \alpha'_{ji} = 1. \quad (7.11)$$

Both normalized background rates β_j , where j runs over the background event categories, and the parameters δ_i describing the strengths of systematic excursions are constrained by Gaussian terms in the likelihood function, see (7.4): the background rates β_j are constrained within the relative uncertainties of the prediction, Δ_j , while the strengths of the systematic effects δ_i are constrained to 0.0 with a standard deviation of 1.0. The single-top-quark content is measured by fitting the parameters of the likelihood function, β_j and δ_j , to the observed data. This is achieved by minimizing the negative logarithm of the likelihood function (7.4) with respect to these parameters using the program MINUIT [129]. In doing so, the normalized expectation values of single-top-quark events, β_1 for the combined search and β_1 and β_2 for the separate search, are allowed to assume only values greater or equal than zero to avoid unphysical results.

Using this technique, one can compute the likelihood function of the combined search as a function of β_1 only by minimizing the negative log-likelihood at a fixed value of β_1 with respect to all other variables (often called nuisance parameters). This method, often named “profiling the likelihood function”, results in a one-dimensional function, the reduced likelihood $\mathcal{L}_{\text{red}}(\beta_1)$.

7.5 Ensemble Tests and Expected Sensitivity

To compute the sensitivity of both combined and separate search, ensemble tests are used. In this context, an ensemble test consists of a set of pseudo experiments. For each pseudo experiment, first the number of events N_j of each event category is determined by drawing a random number from a Poisson distribution of a mean $\hat{\nu}_j$. As a result, the pseudo experiment features a total number of $\sum N_j$ events.

In a second step, N_j random numbers are drawn from the template distributions of the neural network output for all considered event categories displayed figure 7.10(a) for the combined search and figure 7.11 for the separate search, respectively. Those random numbers are filled in a histogram which constitutes the neural network output distribution of a particular pseudo experiment. For each approach, combined and separate search, two ensemble tests are performed: one with single-top-quark events included at the predicted standard-model rate and one without any single-top-quark events. For each pseudo experiment, the single-top-quark cross section is determined using the same method as used for observed events. Based on the ensemble test with single-top-quark events included, the root mean square (RMS) of the resulting single-top-quark cross-section distribution is defined as the expected uncertainty of the measurement. For the combined neural network, a value of 1.3 pb is found including all systematic uncertainties, which is 45% of the predicted cross

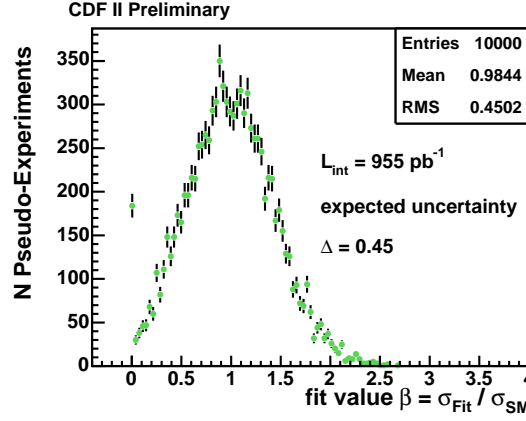


Figure 7.14: The distribution of expected measurements in the combined search is shown. The fitted cross sections are based on the ensemble test with single-top-quark events included. The RMS value of this distribution is defined as the expected uncertainty of the measurement.

section as illustrated in figure 7.14. For the separate neural network search, the expected uncertainties are 1.1 pb for t -channel figure 7.15(a)) and 1.0 pb for s -channel (see figure 7.15(b)) , corresponding to 56% or 112%, respectively.

The effect of the individual sources of systematic uncertainties are estimated by evaluating the distributions of the fitted δ_i parameters measuring the strength of a systematic excursion in units of one standard deviation. The RMS values of these distributions, summarized in table 7.7 for both combined and separate search including rate and shape uncertainties, are defined as a relative measure of the size of the systematic uncertainties. In this context, it is important to note that some sources like mistag model or Q^2 only effect specific processes.

To compute the significance of a potentially observed signal, a hypothesis test is performed, considering two hypotheses. The first one, the null hypothesis H_0 , assumes that the single-top-quark cross section is zero ($\beta_1 = 0$ for the combined search; $\beta_1 = \beta_2 = 0$ for the separate search). The second one, H_1 , assumes that the single-top-quark production cross-section is the one predicted by the standard model ($\beta_1 = 1$ for the combined search; $\beta_1 = \beta_2 = 1$ for the separate search). The objective of both analyses is to observe single-top-quark events, that means to reject the null hypothesis H_0 . The hypothesis test for the combined neural network search is based on the Q -value,

$$Q = -2 (\ln L_{\text{red}}(\beta_1 = 1) - \ln L_{\text{red}}(\beta_1 = 0)) , \quad (7.12)$$

where $L_{\text{red}}(\beta_1 = 1)$ is the value of the reduced likelihood function at the standard-model prediction and $L_{\text{red}}(\beta_1 = 0)$ is the value of the reduced likelihood function for a single-top-quark cross section of zero. The Q -value for the separate search is given in an analogous way. Using the two ensemble tests, the respective Q -value distributions are determined. The resulting distributions are shown in figure 7.16(a) for the combined search and in figure 7.16(b) for the separate search, respectively. In order to quantify the probability for H_0 to be correct, the p -value is defined.

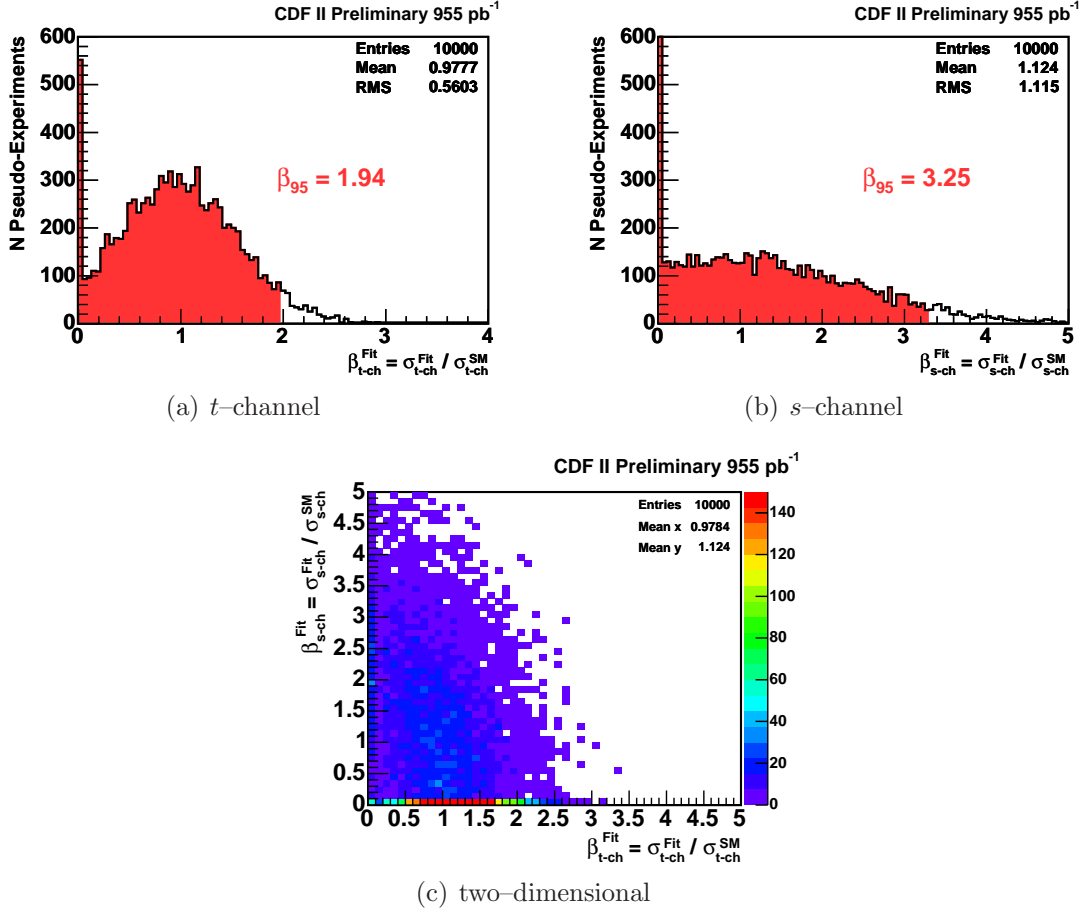


Figure 7.15: The distributions of expected measurements in the separate search are shown for (a) *t*-channel and (b) *s*-channel. The fitted cross sections are based on the ensemble test with single-top-quark events included. The RMS values of these distributions are defined as the expected uncertainties of the measurement. The red-colored areas correspond to 95% of the integral under the particular distributions, representing the expected upper limits. The two-dimensional distribution of the fitted cross sections is displayed in (c).

Source of systematic effect	RMS of δ_i distribution	
	combined search	separate search
mistag shape	0.41	0.46
ALPGEN scale	0.38	0.44
jet energy corrections	0.33	0.34
ALPGEN version	0.28	0.45
neural–network b –tagger	0.23	0.28
non– W model	0.13	0.19
initial–state radiation	0.09	0.14
final–state radiation	0.07	0.14
non– W flavor composition	0.05	0.09
event detection efficiency	0.04	0.04
signal generator	0.018	0.02
parton distribution functions	0.02	0.07

Table 7.7: RMS values of the δ_i distributions obtained from the ensemble test with single–top–quark events included. These values give a relative estimate of the size of the different systematic effects.

Assuming that the value Q_0 is observed in a particular experiment, the p –value is given by

$$p(Q_0) = \frac{1}{I_q} \cdot \int_{-\infty}^{Q_0} q_0(Q') dQ' , \quad (7.13)$$

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

$$I_q = \int_{-\infty}^{+\infty} q_0(Q') dQ' . \quad (7.14)$$

To quantify the sensitivity of the analysis, the expected p –value $\hat{p} = p(\tilde{Q}_1)$ is defined, where \tilde{Q}_1 is the median of the Q –value distribution q_1 for the hypothesis H_1 . The meaning of \hat{p} is the following: Under the assumption that H_1 is correct one expects to observe $p < \hat{p}$ with a probability of 50%. For the combined search, $\hat{p} = 0.5\%$ is found, while it is $\hat{p} = 0.4\%$ for the separate search. Both values include all systematic uncertainties. The found \hat{p} –values can be interpreted as follows: assuming the predicted single–top–quark production cross–section, the expectation is, with a probability of 50%, to see at least that many single–top–quark events that the observed excess over the background corresponds to a background fluctuation of 2.6σ in case of the combined search and 2.7σ in case of the separate search, respectively.

The expected upper limit of the combined search is defined by the median of the distributions of upper limits obtained from pseudo experiments based on the ensemble with single–top–quarks included. According to figure 7.17, the expected upper limit for the combined search is 5.7 pb at 95% confidence level. The expected upper limits for the separate search are determined in a different way, namely from the distributions of the fitted cross sections displayed in figure 7.15(a) and figure 7.15(b),

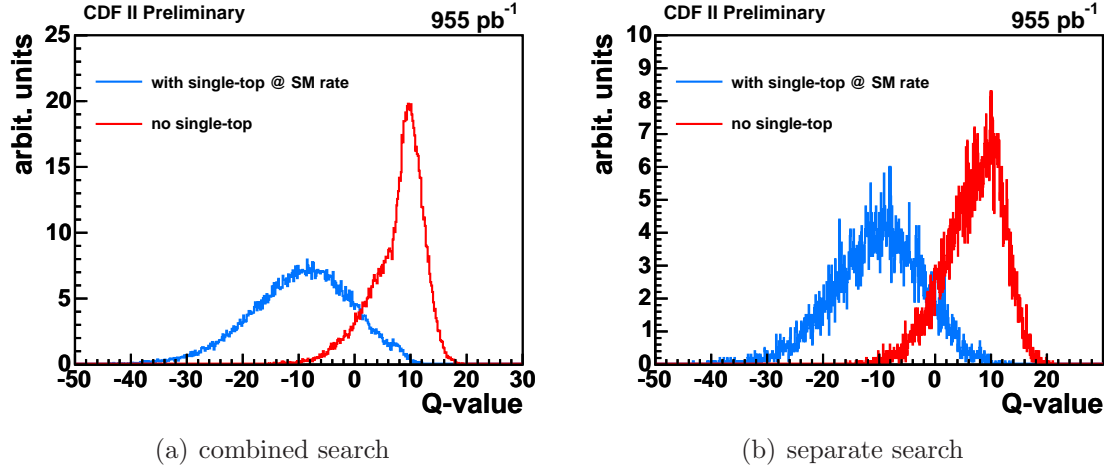


Figure 7.16: Distributions of Q -values for two ensemble tests, one with single-top-quark events present at the expected standard-model rate, one without any single-top-quark events for (a) combined and (b) separate search.

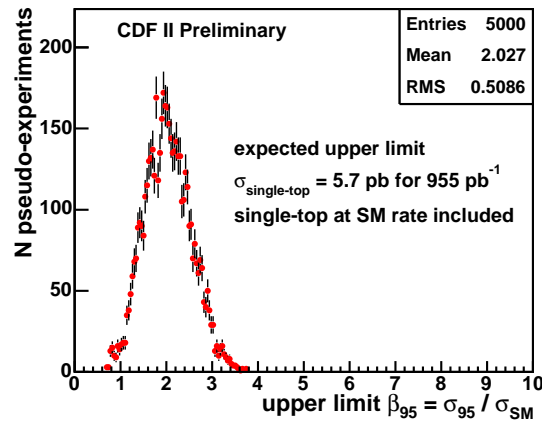


Figure 7.17: The distribution of the upper limits obtained from pseudo experiments based on the ensemble test with single-top-quarks included is shown for the combined search. The median of this distribution is defined as the expected upper limit of the combined analysis.

respectively. In this approach, the limits are derived from the area covering 95% of the integral under the particular distribution, leading to an expected upper limit of 3.8 pb for t -channel and 2.9 pb for s -channel, respectively. To illustrate the result of the pseudo experiments performed in the separate search, the two-dimensional distribution of the fitted cross sections is displayed in figure 7.15(c).

7.6 Application to Observed Events

After the expected sensitivities and the expected upper limits of both approaches have been determined, the neural networks are applied to observed events. At first, the output distributions of observed events are compared to the expected distributions. Finally, the templates are fitted to the observed distributions to determine the single-top-quark cross sections.

7.6.1 Comparison to Expectation

Due to the shapes of the template distributions, the bins with the highest output values are the most interesting ones: here, the ratio of signal to background events should be the largest. Hence, the discussion of the comparison between the distribution of observed events and the expected distribution is restrained to the so-called signal region.

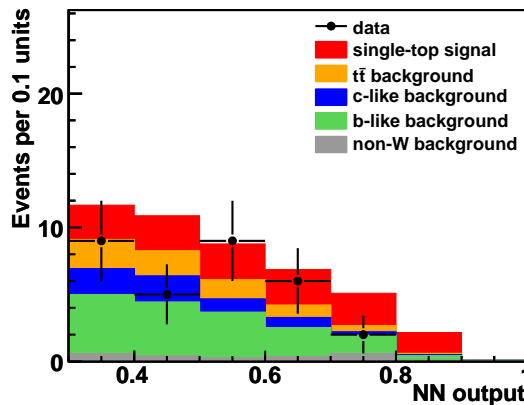


Figure 7.18: The output of observed events is compared to the expectation in the high-output region of the combined analysis.

The output distribution of the combined neural network in the bins with the highest output values is depicted in figure 7.18, illustrating that less events are observed than expected. In the third to last bin, 5.1 events are expected including 2.4 single-top-quark events, while only 2 events are observed. In the last two bins, no events are observed, while the expectation is 2.1 events including 1.6 single-top-quark events in the second to last bin and 0.1 single-top-quark events in the last bin, respectively.

Hence, the observation in the three highest output bins is completely compatible with the background expectation only.

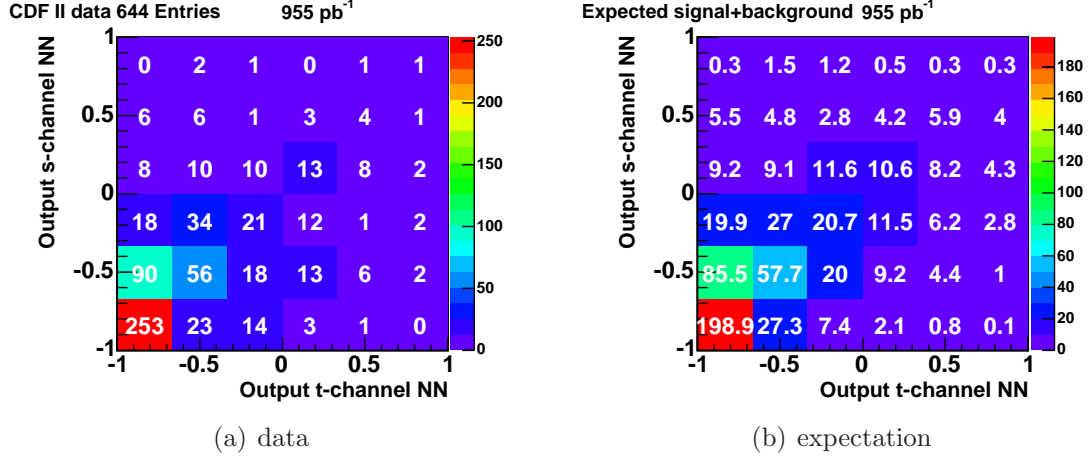


Figure 7.19: The two-dimensional output distributions of the separate search are shown for (a) observed events and (b) expectation. The number quoted for each bin denotes the number of entries in the particular bin.

The two-dimensional output distribution of observed events in the combined search is depicted in figure 7.19(a), while the expectation is illustrated in figure 7.19(b). Comparing the six bins with the highest t -channel outputs, it can be seen that 8 events are observed, while the expectation is 12.5 events. In the bins with the second-highest t -channel output, the observation is 21 events with an expectation of 25.8. In the six bins with the highest output in the s -channel network, 5 events are observed, while 4.1 events are expected. For the bins with the second-highest s -channel network output, 21 events are observed and 27.2 events expected. In conclusion, the data appears less signal-like than predicted even though the expectation agrees with the observation in the six bins with the highest s -channel output.

7.6.2 Fit Results

The likelihood fit to the neural network output for the combined search yields a rate of zero single-top-quark events, as already indicated by figure 7.18. The values obtained for the background rates and strengths of systematic effects are quoted in table 7.8. The observed Q -value is 9.13, yielding an observed p -value of 54.6%. Figure 7.20, comparing the observed Q -value with the expectation, illustrates that the observed data are thus well compatible with being a background fluctuation. As depicted in figure 7.21, the fitted distributions describe the observed output distribution well. The upper limit on the combined single-top-quark cross section is found to be 2.6 pb at the 95% confidence level (C.L.).

For the separate search, the fit yields the cross sections $\sigma_t = 0.2^{+1.1}_{-0.2}(\text{stat.} + \text{syst.})$ pb for t -channel and $\sigma_s = 0.7^{+1.5}_{-0.7}(\text{stat.} + \text{syst.})$ pb for s -channel, respectively. The

process	(fitted rate)/(predicted rate)	fitted events
single-top	0.0 ± 0.41	0.0 ± 15.6
$t\bar{t}$	1.02 ± 0.19	59.7 ± 11.2
b -like	1.13 ± 0.19	201.3 ± 34.1
c -like	1.20 ± 0.11	327.2 ± 30.2
non- W	1.11 ± 0.42	42.0 ± 16.1
systematic source	(fitted excursion)/(standard deviation)	
mistag shape	0.47 ± 0.80	
ALPGEN scale	-0.25 ± 0.81	
jet energy corrections	0.28 ± 0.88	
ALPGEN version	0.06 ± 0.89	
neural-network b -tagger	0.40 ± 0.93	
non- W model	0.043 ± 0.97	
initial-state radiation	0.011 ± 0.96	
final-state radiation	0.036 ± 0.96	
non- W flavor composition	-0.07 ± 0.97	
event detection efficiency	0.041 ± 0.97	
signal generator	0.018 ± 0.97	
parton distribution functions	0.013 ± 0.97	

Table 7.8: The fitted rates of the considered processes and strengths of systematic effects of the combined search are quoted.

fitted rates of the considered processes and strengths of systematic effects are quoted in table 7.9. At the 95% confidence level, the resulting upper limits on the t - and s -channel cross sections are 2.6 pb and 3.7 pb, respectively. As shown in figure 7.22, the observed Q -value of 2.94, yielding a p -value of 21.9%, is more compatible with the standard-model expectation than in case of the combined search. The fit result is depicted in figure 7.23 showing the contours of the negative logarithm of the reduced likelihood function in the plane of single-top-quark s -channel versus t -channel cross sections. Negative cross section values are physically meaningless and therefore not allowed. The minimum represents the best fit values and is indicated by the black dot. The black error bars quote the 1σ uncertainties on the fitted t - and s -channel cross sections, while the yellow error bars present the 95% C.L. upper limits. The true values of both cross sections have a probability of 95% (68%) to be found in the region comprised by the yellow (black) contour. The value predicted by the standard model, within its uncertainties, is illustrated by the red rectangle.

In conclusion, both approaches cannot establish a signal of single top–quark production. The result of the combined search is fully compatible with being a background fluctuation, while the result of the separate search is slightly more compliant with the standard-model expectation, especially for the s -channel.

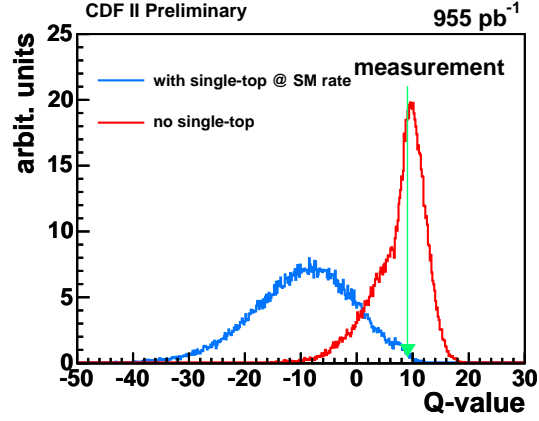


Figure 7.20: Comparison of observed Q -value to the expectation in the combined search.

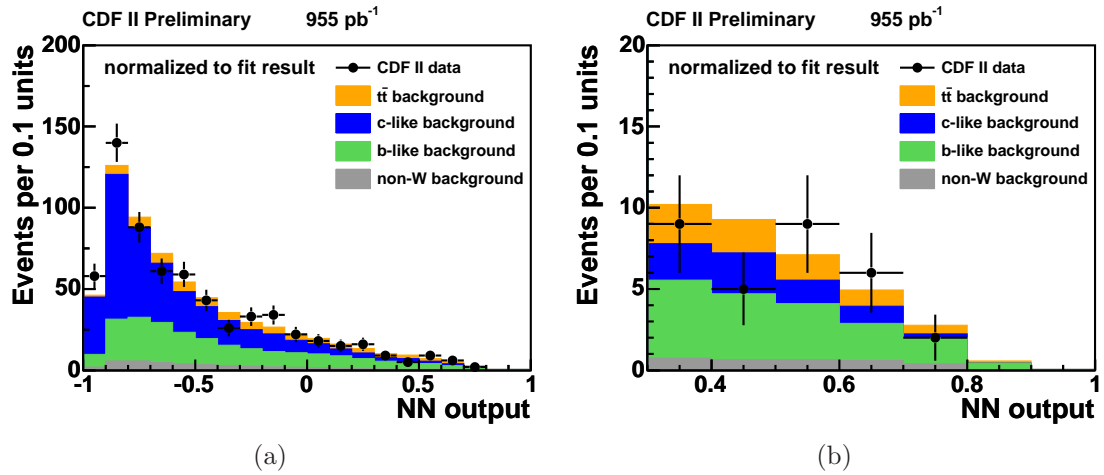


Figure 7.21: Fit result for the combined neural network search versus the observed distribution. (a) shows the entire neural network output domain, while (b) zooms in on the high neural network output region. Since the single-top-quark rate is measured to be zero, it is omitted in these histograms.

process	fitted rate / predicted rate	fitted events
t -channel	0.11 ± 0.66	2.4 ± 14.8
s -channel	0.76 ± 1.66	11.7 ± 25.6
$t\bar{t}$	0.99 ± 0.20	58.0 ± 11.4
b -like	1.12 ± 0.23	200.5 ± 41.9
c -like	1.15 ± 0.12	315.6 ± 32.6
non- W	1.20 ± 0.41	45.9 ± 15.9
systematic source	fitted excursion / standard deviation	
mistag shape	0.582 ± 0.60	
ALPGEN scale	-0.443 ± 0.58	
jet energy corrections	0.074 ± 0.87	
ALPGEN version	0.197 ± 0.54	
neural-network b -tagger	-0.399 ± 0.93	
non- W model	0.0381 ± 0.94	
initial-state radiation	0.178 ± 0.97	
final-state radiation	-0.120 ± 0.95	
non- W flavor composition	-0.080 ± 0.97	
event detection efficiency	-0.009 ± 0.97	
signal generator	-0.004 ± 0.97	
parton distribution functions	0.062 ± 0.97	

Table 7.9: The fitted rates of the considered processes and strengths of systematic effects of the separate search are quoted.

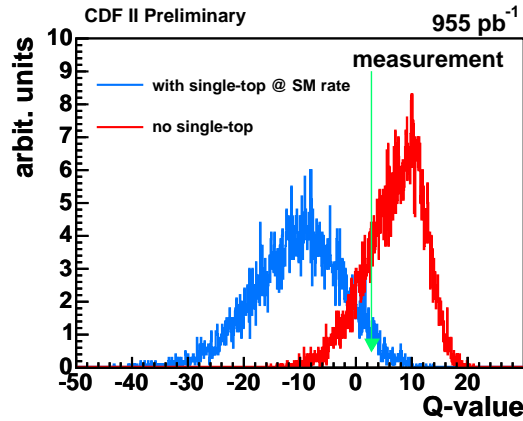


Figure 7.22: Comparison of observed Q -value to the expectation in the separate search.

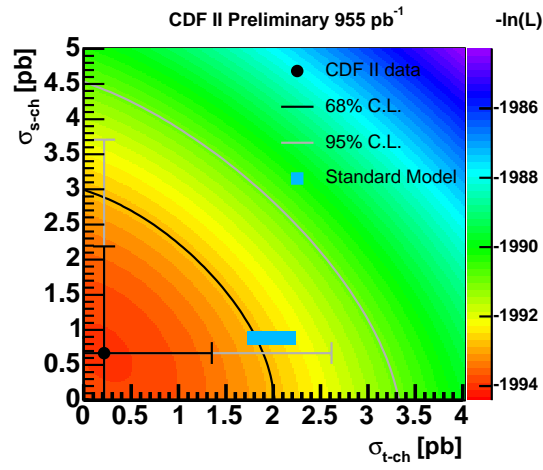


Figure 7.23: Negative log likelihood function of the separate neural network search. The one standard deviation and 95% confidence level contours are overlaid.

Summary and Outlook

In this thesis, a search for single top-quark production using neural networks has been described, utilizing a dataset corresponding to an integrated luminosity of 955 pb^{-1} . At the Tevatron, two processes contribute significantly to the production of single top-quarks: t -channel and s -channel production with predicted cross sections at next-to-leading order of $\sigma_t = 1.98^{+0.28}_{-0.22} \text{ pb}$ and $\sigma_s = 0.88^{+0.12}_{-0.11} \text{ pb}$, respectively.

In the search for single-top-quark events, only semileptonic top-quark decays are considered by requiring one isolated lepton and missing transverse energy. Additionally, exactly two jets are required whereof at least one must be tagged as a b -quark jet. It has been shown that the standard identification of b -quark jets, requiring a reconstructed secondary vertex, still leaves a significant amount of events containing no b -quarks at all. Therefore, it is important to reject secondary vertices which are not related to the decay of a b hadron by extending the identification of b -quark jets.

In this thesis, the first neural-network b -tagger at a hadron collider has been developed, being able to reject c -quark and light-quark jets with secondary vertices. The output of this tagger finds application in all searches for single top-quark production at CDF where it has proven to be of major importance to reject background events. One interesting way of utilizing the output of the neural-network b -tagger lies in the determination of the flavor composition of a given sample of observed events. Currently, this is used as a cross check to the actual estimation of the event yield but has a large potential, since the uncertainties are expected to decrease with increasing statistics. Additionally, it must be noticed that this tagger, developed in the context of searches for single top-quark production, is a universal tool which can be utilized in all kinds of high- p_T analyses requiring identification of b -quark jets, such as the search for a Higgs boson in the WH channel.

In the analysis described in this thesis, two different approaches to search for electroweak production of single top-quarks have been performed. In the first approach, the ratio of the cross sections of both production channels is assumed to be as predicted by the standard model. In this approach, one neural network is trained to identify both t - and s -channel events, the so-called combined network. In the other approach, the cross sections of both channels are determined independently and simultaneously, named separate search. Here, two different neural networks are trained, one to identify only t -channel events and one to identify only s -channel

events. For the t -channel neural network, s -channel events are treated as background and vice versa.

The output distributions of those three neural networks are utilized to create template distributions which are fitted to the output distribution of observed events using a likelihood method. In the likelihood function, Gaussian constraints on the number of predicted background events as well as on the strength of systematic uncertainties are included. Concerning the latter, both systematic rate and shape uncertainties are considered.

The expected sensitivities of the two approaches are determined using hypotheses tests, leading to an expected significance of 2.6σ in the combined search and 2.7σ in the separate search, which is the highest *a priori* sensitivity of all searches for single top-quark production currently performed at the Tevatron. The expected upper limit at 95% confidence level is 5.7 pb in case of the combined search, while it is 3.8 pb for t -channel and 2.9 pb for s -channel in case of the separate search.

The likelihood fit in the combined search yields a rate of zero single-top-quark events. The resulting upper limit at 95% confidence level is 2.6 pb. In case of the separate search, the obtained cross sections are $\sigma_t = 0.2^{+1.1}_{-0.2}(\text{stat.} + \text{syst.})$ pb for t -channel and $\sigma_s = 0.7^{+1.5}_{-0.7}(\text{stat.} + \text{syst.})$ pb for s -channel, respectively. The ensuing upper limits on the t - and s -channel cross sections are 2.6 pb and 3.7 pb, respectively. In conclusion, both approaches are not able to establish a signal of single top-quark production.

Parallel to this work, two other groups at CDF performed searches for single top-quark production, both looking for a combined signal of t - and s -channel. Hereof, only one was able to measure a combined cross section of $2.7^{+1.5}_{-1.3}$ pb with a significance of 2.3σ [130]. The other one obtained a result comparable to the combined neural-network analysis and set a limit of 2.7 pb at the 95% confidence limit [131]. Detailed studies, taking into account the correlations between the different analyses, lead to a consistency around 1%. The interpretation is that the deviations in the obtained results are due to statistical fluctuations. Additionally, several analyses looking for a combined t - and s -channel signal were performed at DØ whereof each observed evidence for single top-quark production [132]. Nevertheless, the measured cross sections are all about twice the predicted value, resulting in a combined value of 4.7 ± 1.3 pb with a significance of 3.6σ [133]. However, this value is excluded at the 95% confidence level by the combined network analysis described in this thesis. For both the DØ analysis and the combined analysis presented in this thesis, the probability to obtain the respective observed result, under the assumption that the standard-model prediction is correct, is at the per-cent level. Since the agreement with the predicted cross section is roughly comparable in both cases, the different results are most likely induced by statistical fluctuations. Therefore, the cross section of single top-quark production is still an exciting open question which can hopefully be answered in the next rounds of analyses utilizing a larger amount of data.

Appendix A

Baseline Cuts for Lepton Identification

A.1 Electron Identification

A.1.1 Central Electrons

- **Fiducial:** Flag, if the electron is within the fiducial volume of the CEM towers. The fiducial volume excludes tower 9 and the chimney, the exhaust of the cabling in tower 7. Additionally, a matched CES strip cluster and a matched CES wire cluster are required.
- **E_T :** Transverse energy deposited in the calorimeter with respect to the primary vertex z position: $E_T = E \cdot \sin \theta$.
- **$|z_0|$:** The z intersection of the track with the beam axis in the r - z plane.
- **p_T :** Transverse momentum of the associated COT track.
- **Good COT Axial Segments:** Number of track segments with at least 5 hits in the COT axial superlayers for the maximum p_T matched track.
- **Good COT Stereo Segments:** Number of track segments with at least 5 hits in the COT stereo superlayers for the maximum p_T matched track.
- **Conversion:** Flag, if the electron is identified as coming from a photon conversion. Conversions are oppositely charged electron-track pairs with similar directions and small separation in the x - y plane. Trident events, i.e. electrons radiating off a photon that converts into e^+e^- , are kept.
- **E_{HAD}/E_{EM} :** Ratio of hadronic calorimeter energy to electromagnetic calorimeter energy using all towers involved.

- **Isolation:** The energy in a cone of radius $R = 0.4$ around the seed tower divided by the total energy of the cluster. The isolation variable is corrected for calorimeter leakage, i.e. energy in adjacent towers not considered by the cluster algorithm.
- **L_{shr} :** The lateral shower profile compares the energy in CEM towers adjacent to the seed tower for data and test-beam electrons and is defined as

$$L_{\text{shr}} = 0.14 \frac{\sum_i (M_i - P_i)}{\sqrt{(0.14\sqrt{E_{\text{EM}}})^2 + \sum_i (\Delta P_i)^2}}, \quad (\text{A.1})$$

where the sums are over the towers in the electromagnetic cluster adjacent to the seed tower and in the same wedge as the seed tower. M_i is the measured energy in an adjacent tower, P_i is the predicted energy deposit in the adjacent tower, E_{EM} is the total electromagnetic energy in the cluster, and ΔP_i is an estimate of the uncertainty in P_i [134].

- **E/P :** Electromagnetic calorimeter energy E divided by the momentum P of the maximum p_T matched track.
- **$|\Delta \mathbf{z}|$:** Distance in the r - z plane between the extrapolated COT track and the best matching CES cluster.
- **$\mathbf{Q} \cdot \Delta \mathbf{x}$:** Distance in the r - ϕ plane between the extrapolated COT track and the best matching CES cluster, multiplied by the charge of the track.
- **χ^2_{strip} :** The χ^2 comparison of the CES shower profile in the r - z view with the same profile extracted from test-beam electrons.

A.1.2 Forward Electrons

- **E_T :** Transverse energy deposited in the calorimeter with respect to the primary vertex z position: $E_T = E \cdot \sin \theta$.
- **PES $|\eta|$:** Detector η of the best matching PES cluster is used as a cut variable to determine fiduciality.
- **$E_{\text{HAD}}/E_{\text{EM}}$:** Ratio of hadronic calorimeter energy to electromagnetic calorimeter energy using all towers involved.
- **Towers in χ^2 fit:** Number of towers used by the 3×3 PEM cluster fit to data from test runs.
- **PEM 3×3 χ^2 :** χ^2 value of the 3×3 PEM cluster fit to data from test runs.
- **PES cluster 5×9 ratio:** By default, PES clusters have 9 strips with the most energetic one at the center. This variable is defined as the ratio of the energy measured by the central 5 strips by the energy measured by all 9 strips of the cluster.

Central Electron Variable	Cut	TNtuple
Region	CEM	Region
Fiducial	Yes	Fiducial
E_T	≥ 20.0 GeV	Et
$ z_0 $	≤ 60.0 cm	TrkZ0
p_T	≥ 10.0 GeV/ c	TrkPt
Good COT Axial Segments	≥ 3	TrkAxSeg
Good COT Stereo Segments	≥ 2	TrkStSeg
Conversion	$\neq 1$	Conversion
$E_{\text{HAD}}/E_{\text{EM}}$	$\leq 0.055 + 0.00045 \cdot E$	Hadem
Isolation	≤ 0.1	Isol
L_{shr}	≤ 0.2	LshrTrk
E/P	≤ 2.0 unless $p_T \geq 50$ GeV/ c	EP
$ \Delta z $	≤ 3.0 cm	DeltaZ
$Q \cdot \Delta x$	≥ -3.0 cm and ≤ 1.5 cm	Charge, DeltaX
χ^2_{strip}	≤ 10.0	StripChi2

Table A.1: Baseline cuts for **central electrons** [105]. Variables given in the TNtuple column are members of the **electron** class of the **HighLevelObjects**.

- **Isolation:** The energy in a cone of radius $R = 0.4$ around the seed tower divided by the total energy of the cluster. The isolation variable is corrected for calorimeter leakage, i.e. energy in adjacent towers not considered by the cluster algorithm.
- **ΔR PEM-PES:** Distance in the η - ϕ plane ($\sqrt{\Delta\phi^2 + \Delta\eta^2}$) between the signal in the plug electromagnetic calorimeter cluster and the best matching pre-shower cluster.
- **Phoenix match:** Match between reconstructed phoenix track and calorimeter tower.
- **Number PHX Si hits:** Number of silicon hits associated to the track.
- **Number PES clusters:** Number of pre-shower clusters.
- $|z_0|$: The z intersection of the track with the beam axis in the r - z plane.

A.2 Muon Identification

The baseline cuts for **muons** are given in table A.3. The detailed description of the cut variables are:

- p_T : Transverse momentum of the corresponding track.

Plug Electron Variable	Cut	TNtuple
Region	PEM	Region
E_T	≥ 20.0 GeV	Et
PES $ \eta $	$1.2 \leq \eta \leq 2.0$	Pes2dEta
$E_{\text{HAD}}/E_{\text{EM}}$	≤ 0.05	Hadem
Towers in χ^2 fit	≥ 1	Pem3x3FitTow
PEM 3 x 3 χ^2	≤ 10.0	Pem3x3Chisq
PES U cluster 5 x 9 ratio	≥ 0.65	Pes2d5by9U
PES V cluster 5 x 9 ratio	≥ 0.65	Pes2d5by9V
Isolation	≤ 0.1	Isol
ΔR PEM-PES	≤ 3 cm	Pem3x3DetEta, Pem3x3Phi Pes2dX, Pes2dY
Phoenix match	Yes	PhxMatch
Number PHX Si hits	≥ 3	TrkSiHits
Number PES clusters	≥ 1	NumPes2d
$ z_0 $	≤ 60.0 cm	TrkZ0

Table A.2: Baseline cuts for **plug electrons** [105] using the phoenix tracking. Variables given in the TNtuple column are members of the **electron** class of the **HighLevel0bjects**.

- **E_{EM} :** Energy deposition in the electromagnetic calorimeter.
- **E_{HAD} :** Energy deposition in the hadronic calorimeter.
- **Isolation:** Total excess energy within a cone of radius $\Delta R = 0.4$ around the muon divided by the transverse momentum of the muon.
- **$|z_0|$:** The z intersection of the track with the beam axis in the r - z plane.
- **$|d_0|$:** If the track does not include silicon hits, d_0 is measured with respect to the COT beam spot. Otherwise, the SVX beam spot is taken as reference point.
- **CMU/ CMP / CMX $|\Delta \mathbf{x}|$:** Distance in r - ϕ between the stub direction and the track extrapolation to the stub.
- **Fiduciality:** A track is projected out from the tracking chambers to the plane of the muon chamber in question. The fiducial distance is then defined as the distance between the projected track position and the edge of the chamber. Two planes are used for this measurement. One is the local z direction of the chamber which is in the longitudinal direction of the chambers. The other is the local x direction of the chamber which is along the drift direction.
- **COT exit radius:** The COT exit radius check ensures that the muon passes through all 4 axial layers of the COT and could trigger the event.

Central Muon Variable	Cut	TNtuple
p_T	≥ 20.0 GeV	PtCorr
E_{EM}	$\leq 2 + \max(0, 0.0115(p - 100))$ GeV	EmEnergy
E_{HAD}	$\leq 6 + \max(0, 0.0280(p - 100))$ GeV	HadEnergy
Isolation	≤ 0.1	Isol
Good COT Axial Segments	≥ 3	TrkAxSeg
Good COT Stereo Segments	≥ 2	TrkStSeg
$ z_0 $	≤ 60.0 cm	Z0
$ d_0 $	≤ 0.2 cm if no Si hits	D0
$ d_0 $	≤ 0.02 cm if Si hits	D0
Additional requirements for specific subdetectors		
CMU $ \Delta x $	≤ 3.0 cm	CmuDx
CMP $ \Delta x $	≤ 5.0 cm	CmpDx
CMX $ \Delta x $	≤ 6.0 cm	CmxDx
Fiduciality requirements		
CMUP	CMUFidX < 0 cm, CMUFidZ < 0 cm	CmuFidX, CmuFidZ
	CMPFidX < 0 cm, CMPFidZ < -3 cm	CmpFidX, CmpFidZ
CMX	CMXFidX < 0 cm, CMXFidZ < -3 cm	CmxFidX, CmxFidZ
	COT exit radius > 140 cm	Eta, Z0

Table A.3: Baseline cuts for **CMUP** and **CMX** muons [103, 104]. Variables given in the TNtuple column are functions of the `muon` class of the `HighLevel0bjects`.

A.3 Selection of Heavy-Flavor Enhanced Events

To determine the difference between b -quark jets in data and simulation, an energetic electron within a jet and an additional jet opposite in ϕ are required.

The requirements for the electron are quoted in table A.4. Some of the quantities have already been introduced in section A.1.1, the others are explained below.

- $|z|$: The z coordinate of the track helix at a given radius
- $|z_{\text{pvtx}} - z_0|$: The distance of the z intersection of the track with the beam axis in the r - z plane (z_0) from the z position of the primary vertex (z_{pvtx}).
- $|\Delta \mathbf{x}|$: Distance in the r - ϕ plane between the extrapolated COT track and the best matching CES cluster.

If more than one candidate is present, the one with the highest E_T is taken.

Central Electron Variable	Cut	TNtuple
Fiducial	Yes	Fiducial
E_T	$\geq 9.0 \text{ GeV}$	Et
$ z $ at edges of SVX II	$< 43.5 \text{ cm}$	TrkZ0, TrkD0, Curv, Lambda
$ z_{\text{pvtx}} - z_0 $	$< 5 \text{ cm}$	fjetZV, TrkZ0
p_T	$\geq 8.0 \text{ GeV}/c$	TrkPt
Conversion	$\neq 1$	Conversion
$E_{\text{HAD}}/E_{\text{EM}}$	< 0.05	Hadem
Isolation	> 0.1	Isol
L_{shr}	≤ 0.2	LshrTrk
E/P	$0.5 < E/P < 2.0$	EP
$ \Delta z $	$\leq 5.0 \text{ cm}$	DeltaZ
$ \Delta x $	$\leq 3.0 \text{ cm}$	DeltaX
χ_{strip}^2	≤ 10.0	StripChi2

Table A.4: The baseline cuts for electrons used to exploit semileptonic b -hadron decays [97, 98]. Variables given in the **TNtuple** column are members of the **electron**, **track** and **summary** classes of the **HighLevelObjects**.

The selection of the electron jet is done the following way:

- $E_T > 15 \text{ GeV}$ (corrected up to level 4)
- $\Delta R = \sqrt{(\eta_{\text{electron}} - \eta_{\text{jet}})^2 + (\phi_{\text{electron}} - \phi_{\text{jet}})^2} < 0.4$
- If more than one candidate, the one with the smallest ΔR is taken.

The selection of the away jet is done the following way:

- $E_T > 15$ GeV (corrected up to level 4)
- $\eta < 1.5$
- $|\Delta\phi| > 2$ rad between electron and away jet
- If more than one candidate, the one with the largest $|\Delta\phi|$ is taken.

Appendix B

Used MC Samples

In the search of single top-quark production, a set of Monte Carlo simulated samples (see table B.1) is used to model the selected data sample. Single top-quark production is simulated with the matrix element generator MADEVENT and PYTHIA showering. For the modeling of $t\bar{t}$, diboson, and Z +jets production, simulated events generated with PYTHIA are used. W +heavy flavor events were simulated using a combination of ALPGEN and HERWIG. In addition, this combination forms the basis for the mistag model used in this analysis.

To evaluate the systematic uncertainties of our search for single top-quark production, a variety of MC samples is utilized (see table B.2), each containing variations due to a specific source of systematic uncertainty.

To determine the correction functions which have to be applied to the output of the neural-network b -tagger, another set of simulated processes was used (see table B.3).

Sample	Process	Generator	Events
ttop1oNew + ttop2oNew	t -channel (matched)	MADEVENT + PYTHIA	600060
ttop0oNewCat	s -channel	MADEVENT + PYTHIA	577276
ttopkl + ttopvl + ttopyl + ttoptl	$t\bar{t}$	PYTHIA	3463090
ltop0b + ltop3b + atop0t	$Wb\bar{b} + 0p$ ($W \rightarrow e/\mu/\tau + \nu$)	ALPGEN + HERWIG	814418
ltop0c + ltop3c + atop3t	$Wc\bar{c} + 0p$ ($W \rightarrow e/\mu/\tau + \nu$)	ALPGEN + HERWIG	879204
ltop1a + ltop5a	$Wc + 1p$ ($W \rightarrow e/\mu + \nu$)	ALPGEN + HERWIG	532823
ltop2n + ltop2m + atop8t	$W + 2p$ ($W \rightarrow e/\mu/\tau + \nu$)	ALPGEN + HERWIG	356454
wtop1w + wtop2w	WW	PYTHIA	606877
wtop1z + wtop2z	WZ	PYTHIA	594502
ztopcz + ztopfz	ZZ	PYTHIA	801176
ztop7i	$Z \rightarrow ee$	PYTHIA	3251489
ztop0i + ztopei	$Z \rightarrow \mu\mu$	PYTHIA	1794600
ztop1i + ztop4i + ztop5i	$Z \rightarrow \tau\tau$	PYTHIA	3538885

Table B.1: Monte Carlo samples used in the search for single top-quark production.

Sample	Comment	Events
Single-Top		
sisrls	s -channel, less ISR	294910
sisrmr	s -channel, more ISR	267491
sfsrls	s -channel, less FSR	299793
sfsrmr	s -channel, more FSR	299838
tisrls_merged	t -channel, less ISR	303293
tisrmr_merged	t -channel, more ISR	242986
tfsrls_merged	t -channel, less FSR	311474
tfsrmr_merged	t -channel, more FSR	298803
tchan170_merged	t -channel, $m_{\text{top}} = 170 \text{ GeV}/c^2$	135565
tchan180_merged	t -channel, $m_{\text{top}} = 180 \text{ GeV}/c^2$	131691
ctoph0	s -channel, $m_{\text{top}} = 170 \text{ GeV}/c^2$	100165
ctoph1	s -channel, $m_{\text{top}} = 180 \text{ GeV}/c^2$	69671
$t\bar{t}$		
ttopbr	$t\bar{t}$, less ISR	930469
ttopdr	$t\bar{t}$, more ISR	924188
ttopfr	$t\bar{t}$, less FSR	932334
ttopkr	$t\bar{t}$, more FSR	466292
W + Jets		
ltop1r + ltop3r	$Wbb + 0p, Q^2$ ($W \rightarrow e/\mu/ + \nu$)	379422
ltop5r + ltop7r	$Wc\bar{c} + 0p, Q^2$ ($W \rightarrow e/\mu/ + \nu$)	368864
ltop9r + ltopbr	$Wc + 1p, Q^2$ ($W \rightarrow e/\mu/ + \nu$)	339125
ltopdr + ltopfr	$W + 2p, Q^2$ ($W \rightarrow e/\mu/ + \nu$)	248089
mtop0x + mtop1x + mtop2x	$Wbb + 0p$ ($W \rightarrow e + \nu$) ALPGEN Version 2	1124616
mtop0y + mtop1y + mtop2y	$Wbb + 0p$ ($W \rightarrow \mu + \nu$) ALPGEN Version 2	1127743

Table B.2: Systematic Monte Carlo samples used in the search for single top-quark production.

Sample	Process	Generator
<i>b</i>-like correction		
btop5a + btop6a	$2 \rightarrow 2$ scattering QCD, $p_{T\min} = 15 \text{ GeV}/c$, filtered for 7 GeV/ c electron or muon in generator parton list	HERWIG
btop7a + btop8a	$2 \rightarrow 2$ scattering QCD, $p_{T\min} = 20 \text{ GeV}/c$, filtered for 8 GeV/ c electron or 9 GeV/ c muon in generator parton list	HERWIG
mistag correction		
jqcd1f	dijet sample, $p_T > 18 \text{ GeV}/c$	PHYTIA
jqcd2f	dijet sample, $p_T > 40 \text{ GeV}/c$	PHYTIA
jqcd2h	dijet sample, $p_T > 18 \text{ GeV}/c$	HERWIG
jqcd3h	dijet sample, $p_T > 40 \text{ GeV}/c$	HERWIG

Table B.3: Monte Carlo samples used to determine the correction functions of the neural-network *b*-tagger.

Appendix C

More Discriminating b -Tagger Variables

C.1 Discriminating Variables

Figures C.1 to C.8 show the distributions of discriminating variables investigated for the neural-network b -tagger. The distributions are shown for b -quark, c -quark, and light-quark jets. More information can be found in section 6.1.1. Some further examples are shown in figures 6.1 to 6.4.

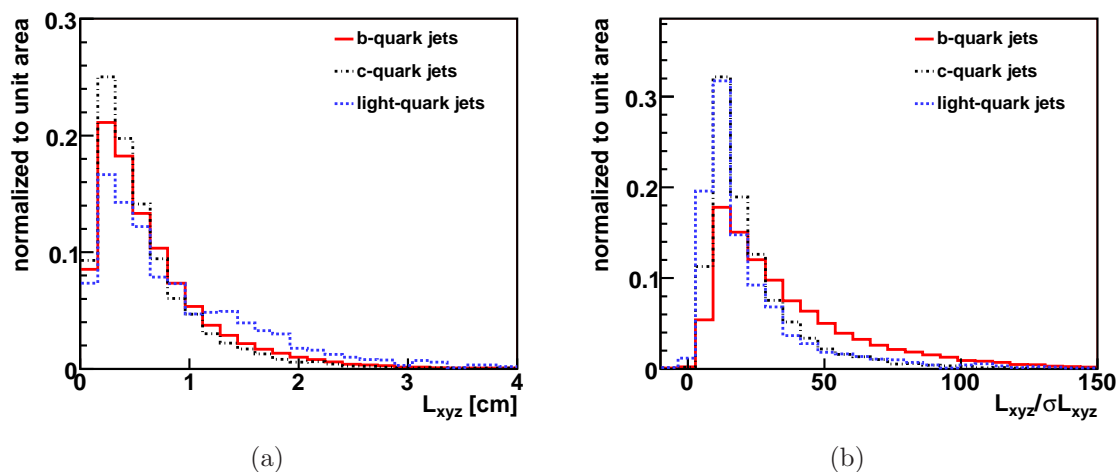


Figure C.1: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: (a) the 3-dimensional decay length of the secondary vertex, (b) the significance of the 3-dimensional decay length.

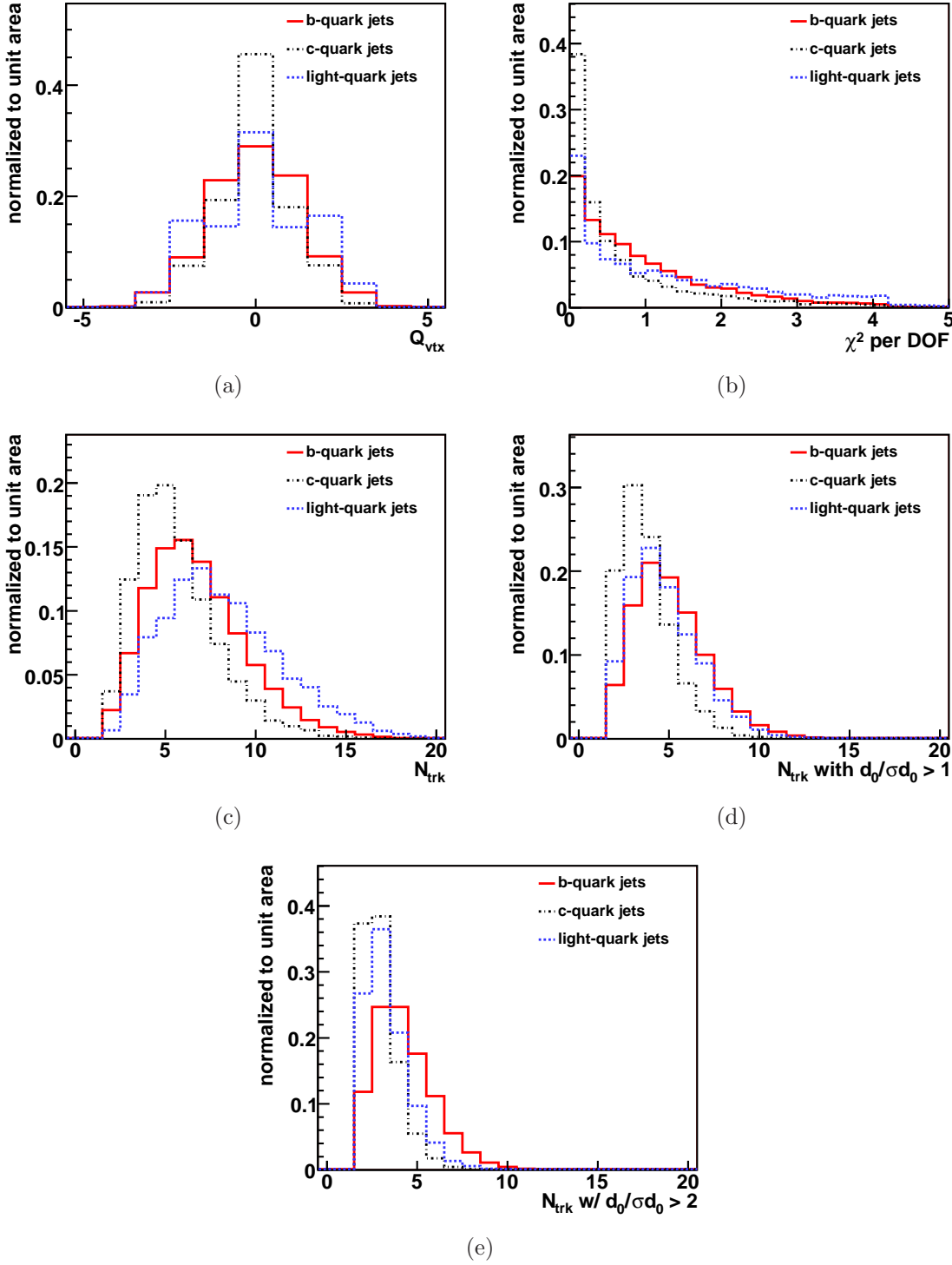


Figure C.2: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: (a) the charge at the secondary vertex, (b) the χ^2 per degree of freedom of the vertex fit, (c) the number of tracks per jet and the number of tracks in the tagged jet with an impact parameter significance larger than (d) 1σ and (e) 2σ .

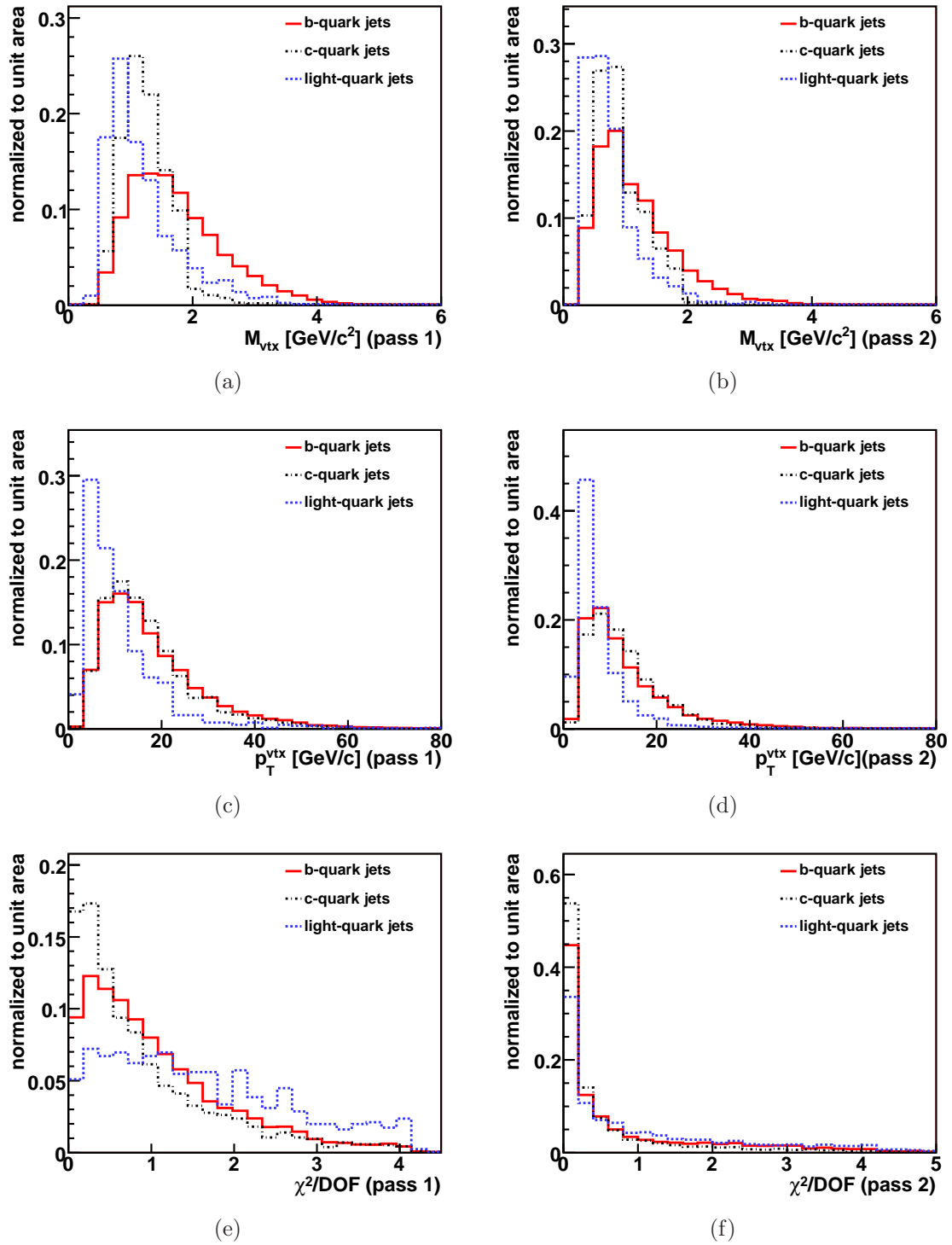


Figure C.3: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: the invariant mass of (a) pass-1 and (b) pass-2 vertices, the transverse momentum at (c) pass-1 and (d) pass-2 vertices, and the χ^2 per degree of freedom for (e) pass-1 and (f) pass-2 vertices.

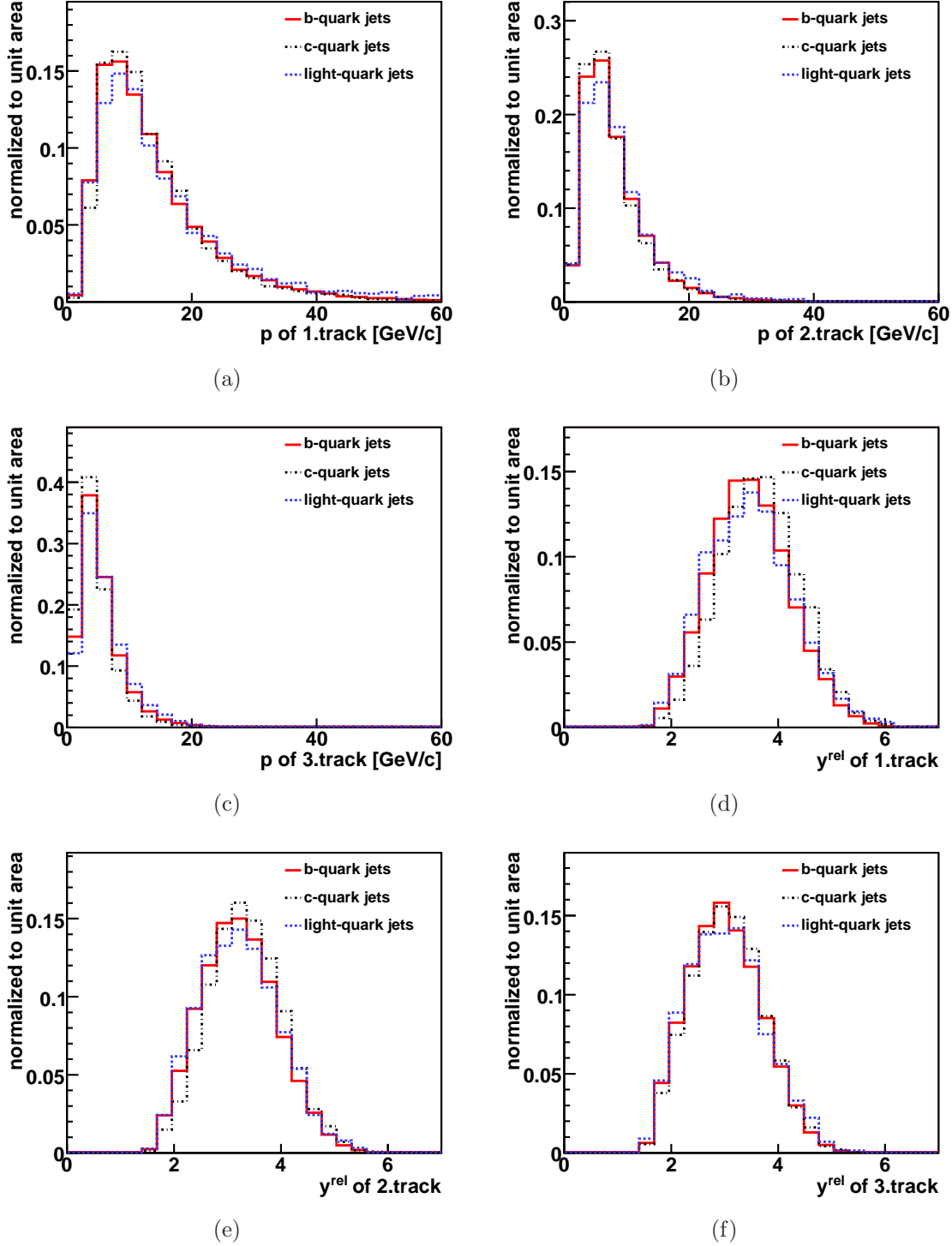


Figure C.4: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: the momentum of the (a) 1.track, (b) 2.track, and (c) 3.track, and the rapidity y^{rel} with respect to the jet axis of the (d) 1.track, (e) 2.track, and (f) 3.track.

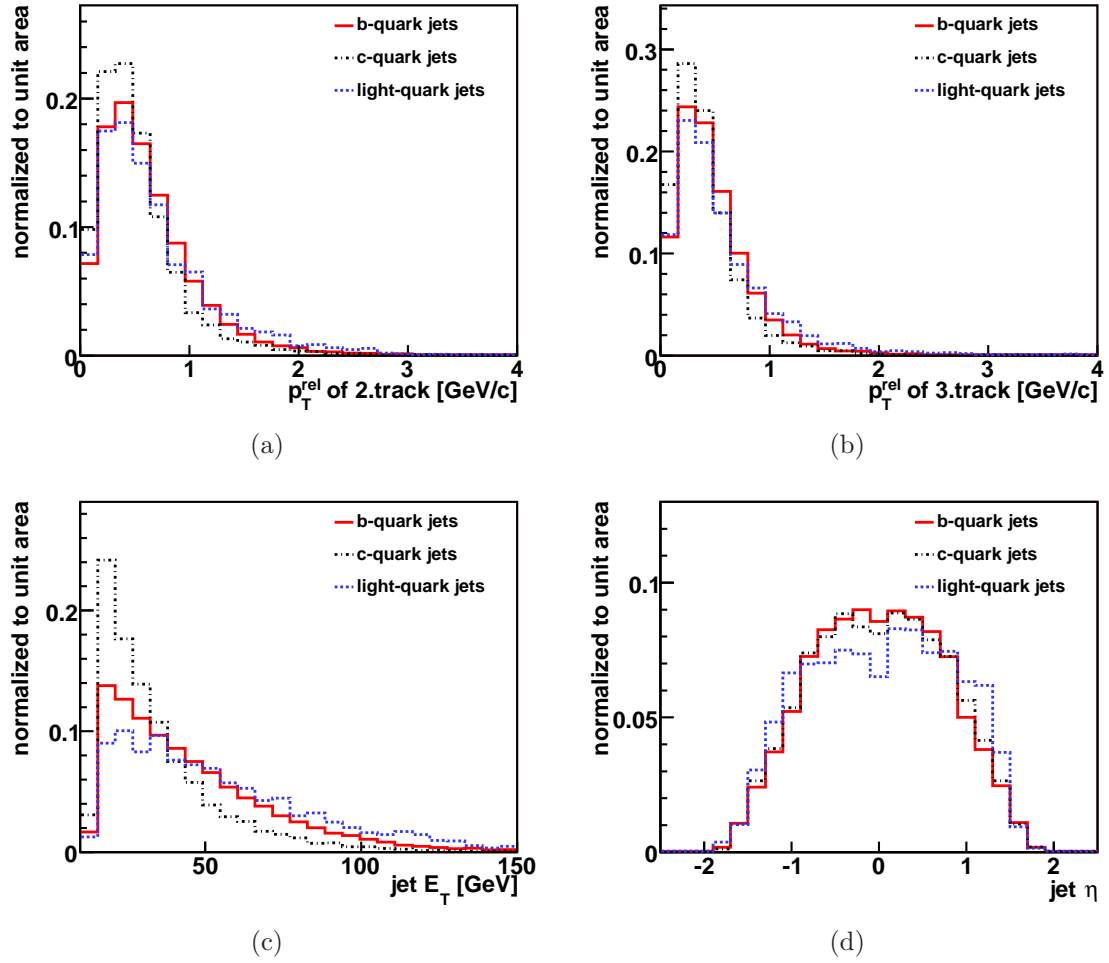


Figure C.5: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: the transverse momentum p_T^{rel} with respect to the jet axis of the (a) 2.track and (b) 3.track, (c) jet E_T , and (d) jet η .

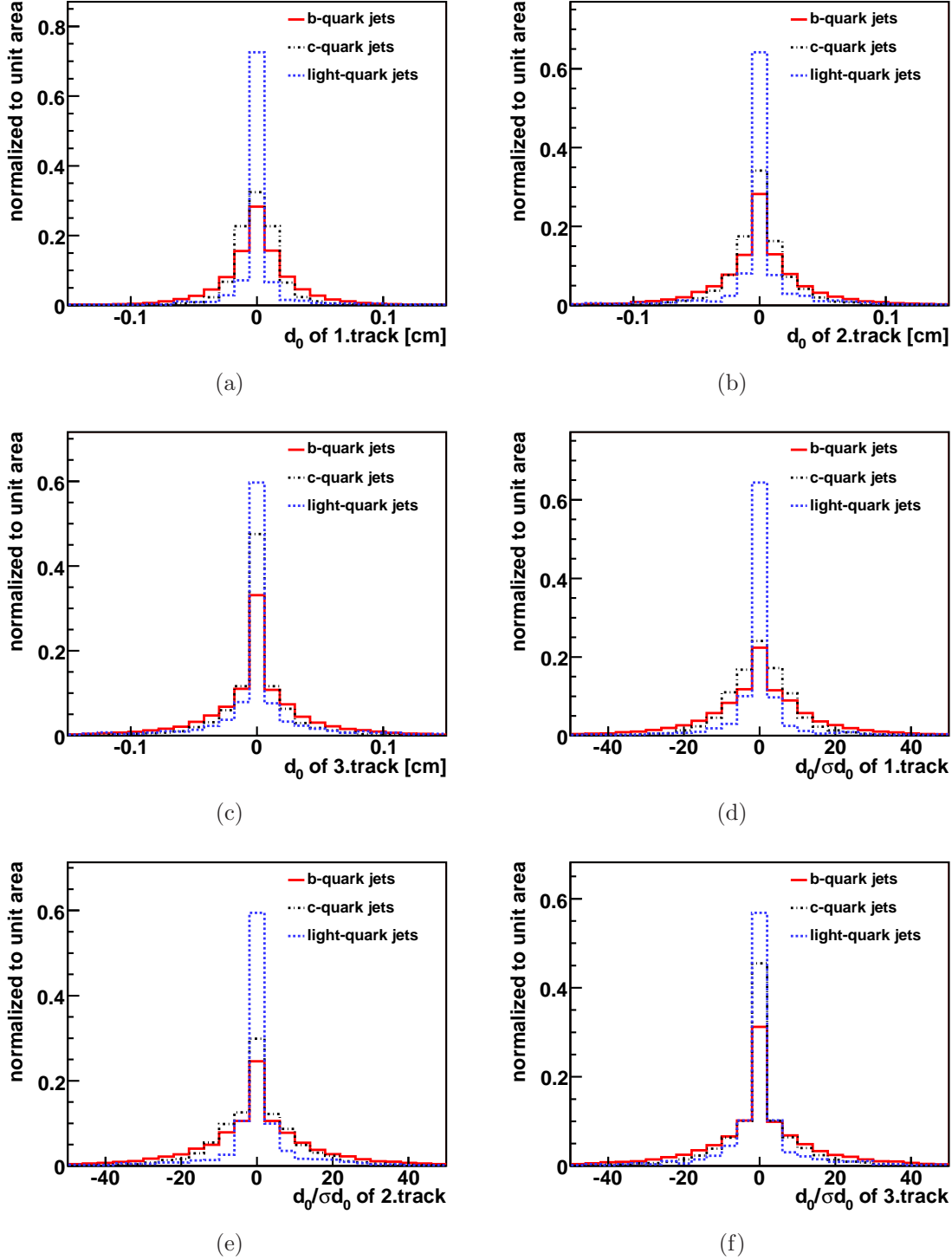


Figure C.6: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: the impact parameter d_0 of the (a) 1.track, (b) 2.track, and (c) 3.track and the significance of the impact parameter of the (d) 1.track, (e) 2.track, and (f) 3.track.

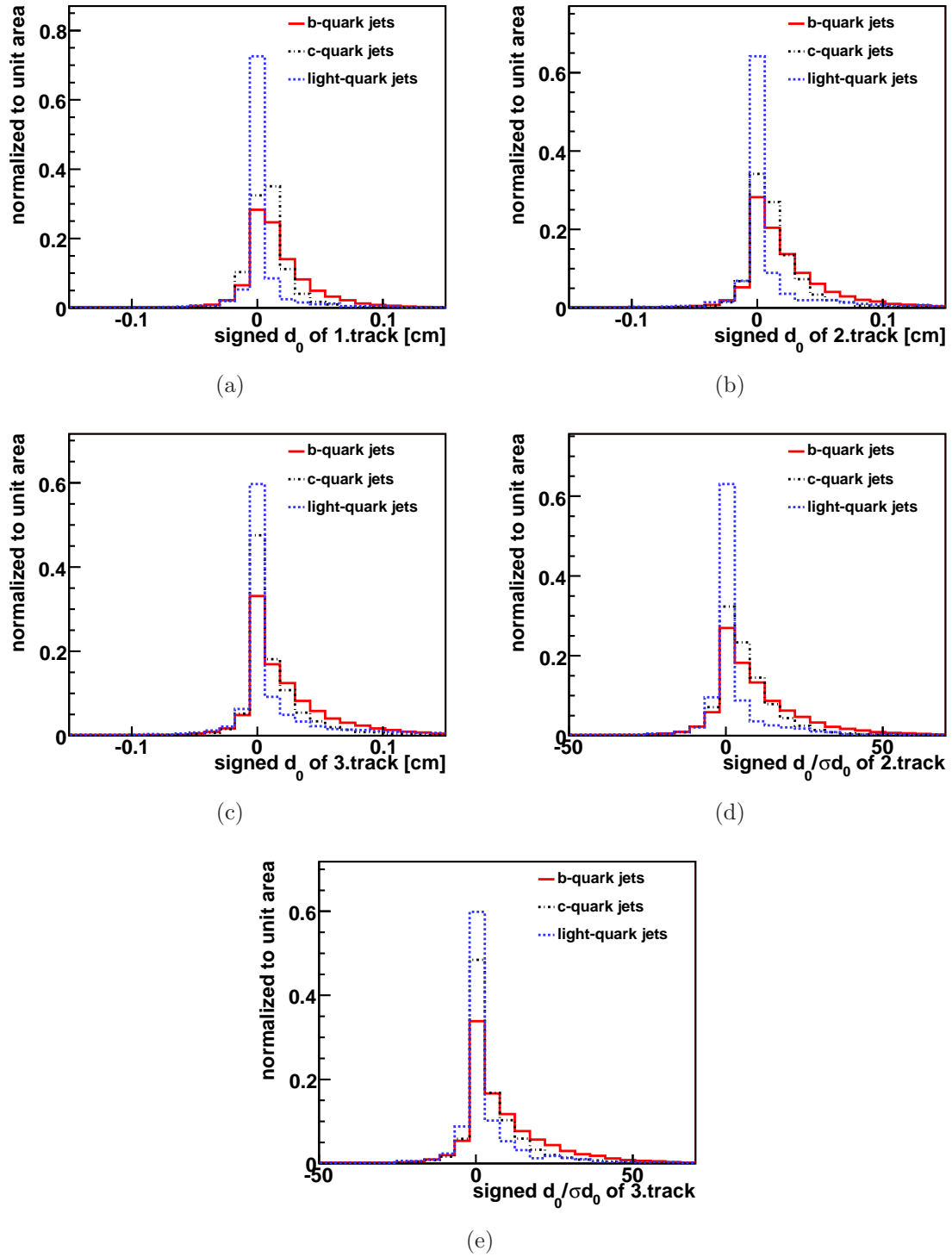


Figure C.7: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: the signed impact parameter of the (a) 1.track, (b) 2.track, and (c) 3.track and the significance of the signed impact parameter of the (d) 2.track and (e) 3.track.

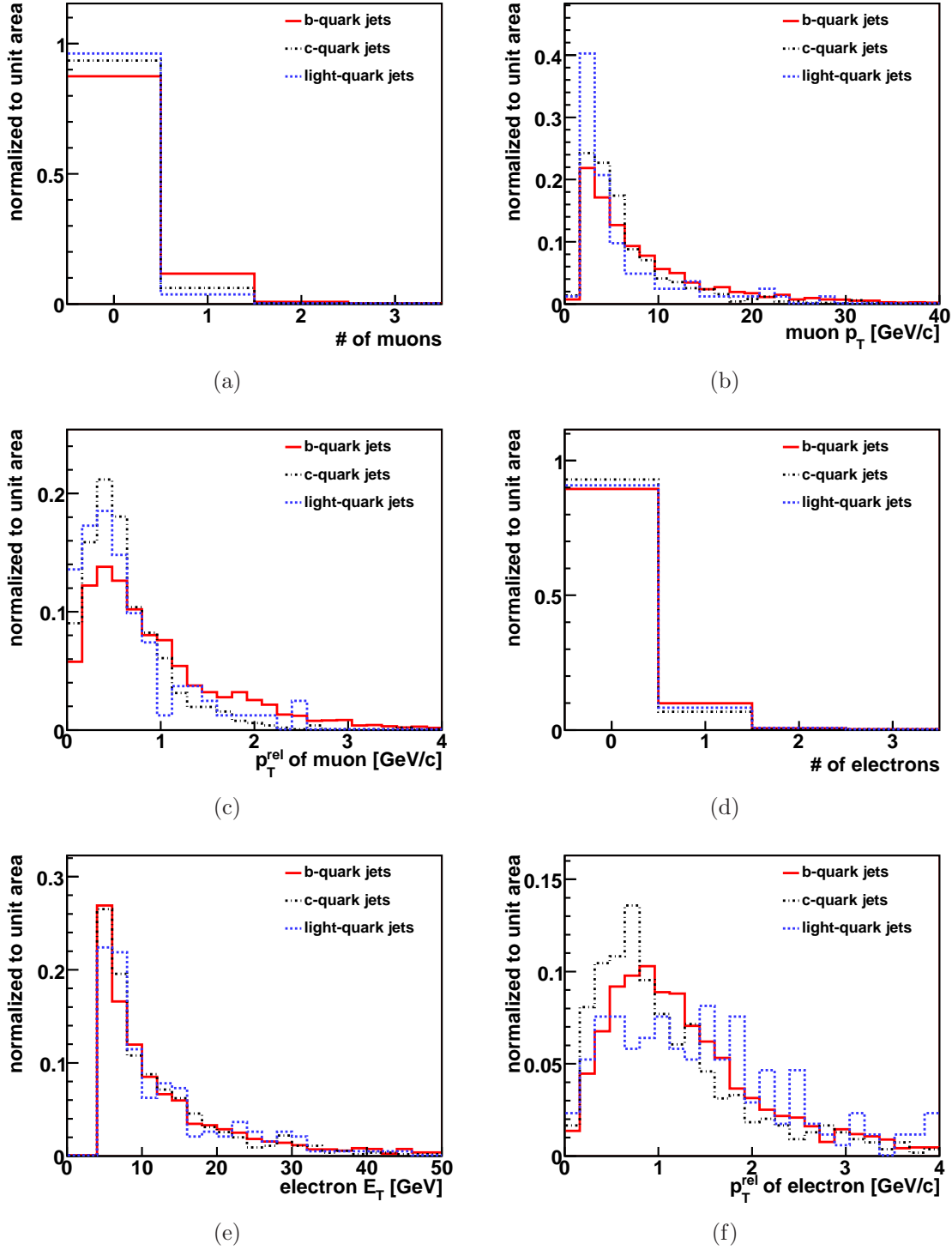


Figure C.8: Qualified variables to discriminate secondary vertices stemming from the decay of a b hadron: (a) the number of muons per jet and (b) p_T and (c) transverse momentum p_T^{rel} with respect to the jet axis of the most energetic muon in the jet, (d) the number of electrons per jet and (e) E_T and (f) p_T^{rel} of the most energetic electron in the jet.

C.2 Discriminating Variables in Data and MC

Figures C.9 to C.17 show the comparison of discriminating b -tagger variables in data and simulation. More information about the used samples can be found in section 6.3.2.

The differences observed in figures C.12(c), C.12(d), C.17(d), and C.17(e) are induced by differences between electrons in data and simulation. First, the selected electron in observed events can be created by material interactions, while this is avoided by requiring an electron at generator level in simulated events. Second, in simulation, the track of the selected electron fulfills more often some minimum quality requirements than in data. This fact additionally increases the probability of the track to be attached to the secondary vertex in simulated events. Nevertheless, this is a specific feature of the inclusive lepton samples used for this comparison and not expected in genuine b -quark jets.

Some further comparisons are shown in figures 6.7 to 6.9.

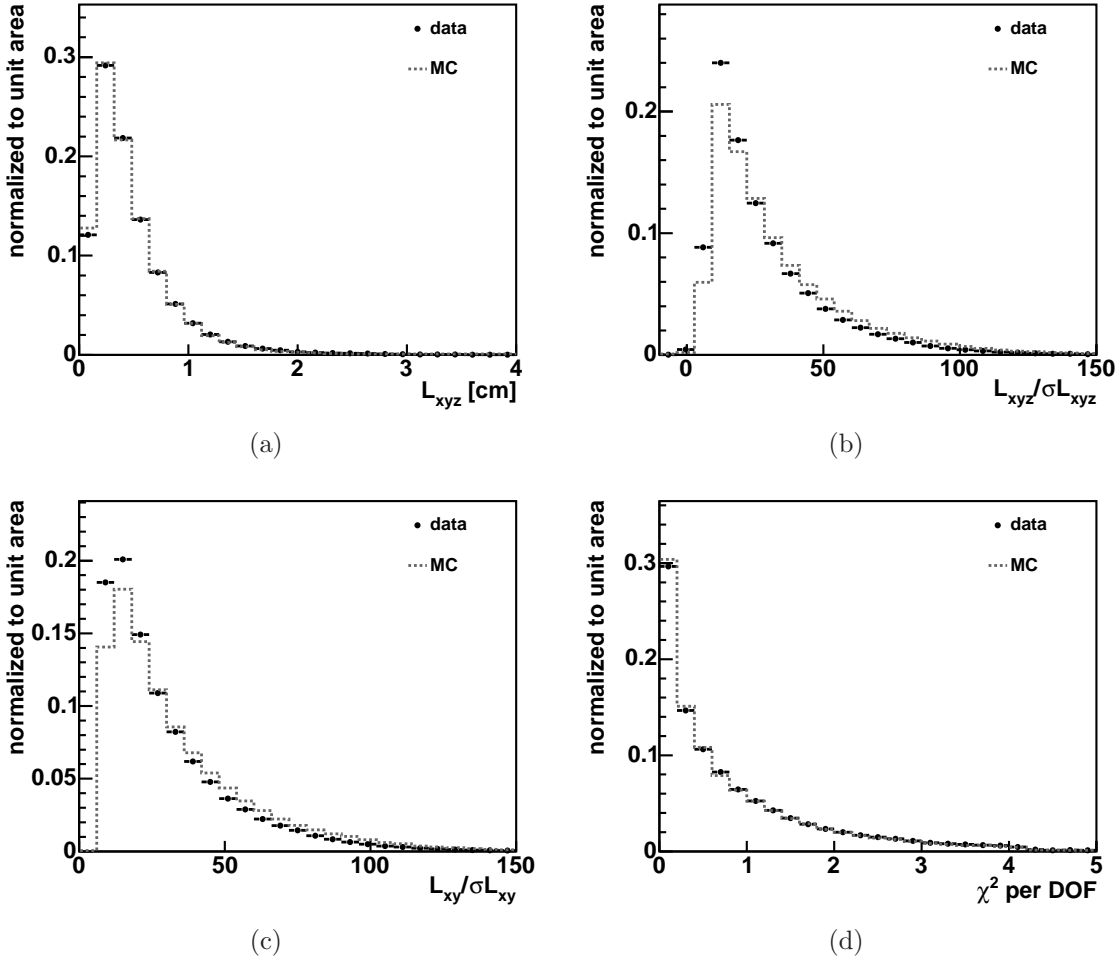


Figure C.9: Comparison of discriminating b -tagger variables in data and simulation: the (a) 3-dimensional decay length of the secondary vertex, the significance of the (b) 3-dimensional and (c) 2-dimensional decay length, and (d) the χ^2 per degree of freedom of the vertex fit,

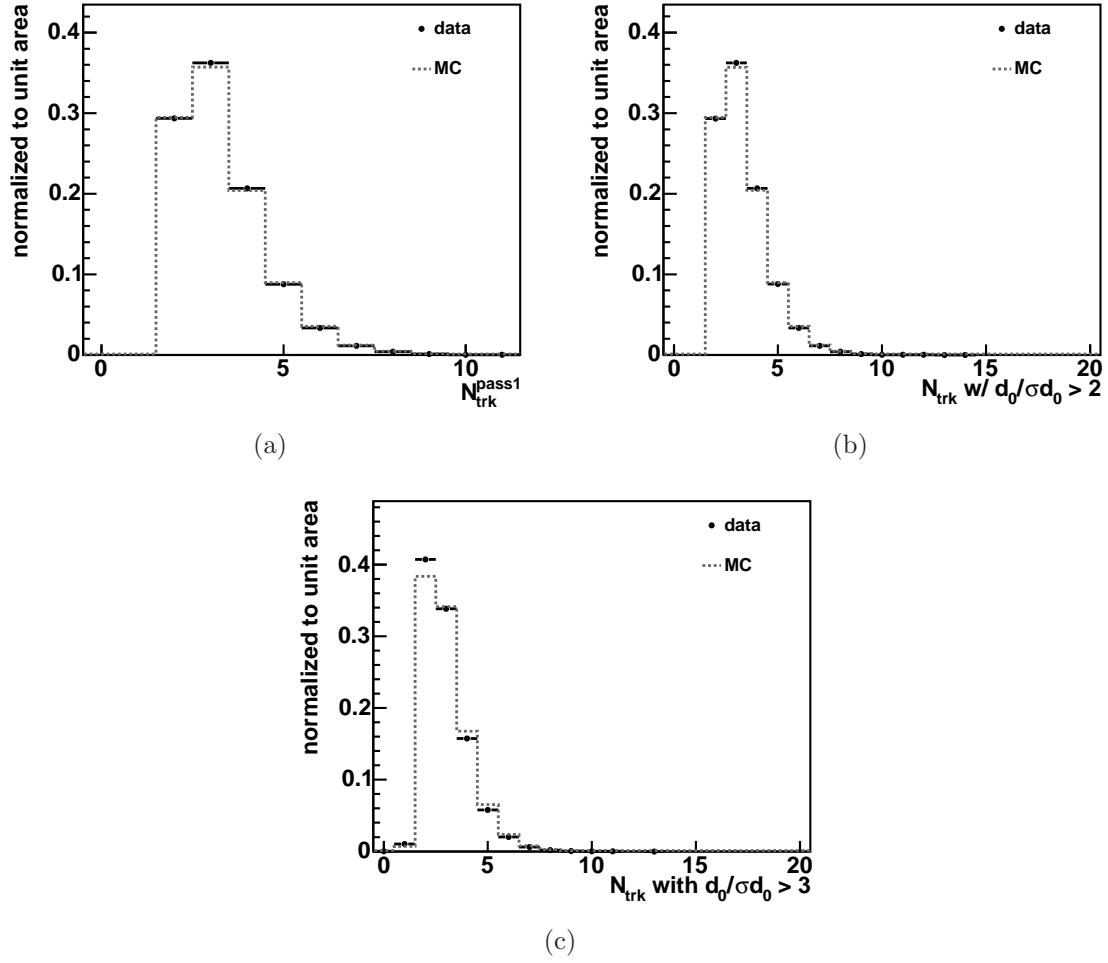


Figure C.10: Comparison of discriminating b -tagger variables in data and simulation: (a) the number of displaced tracks fulfilling the quality requirements for pass 1 of the vertex fit, the number of tracks in the tagged jet with an impact parameter significance larger than (b) 2σ and (c) 3σ .

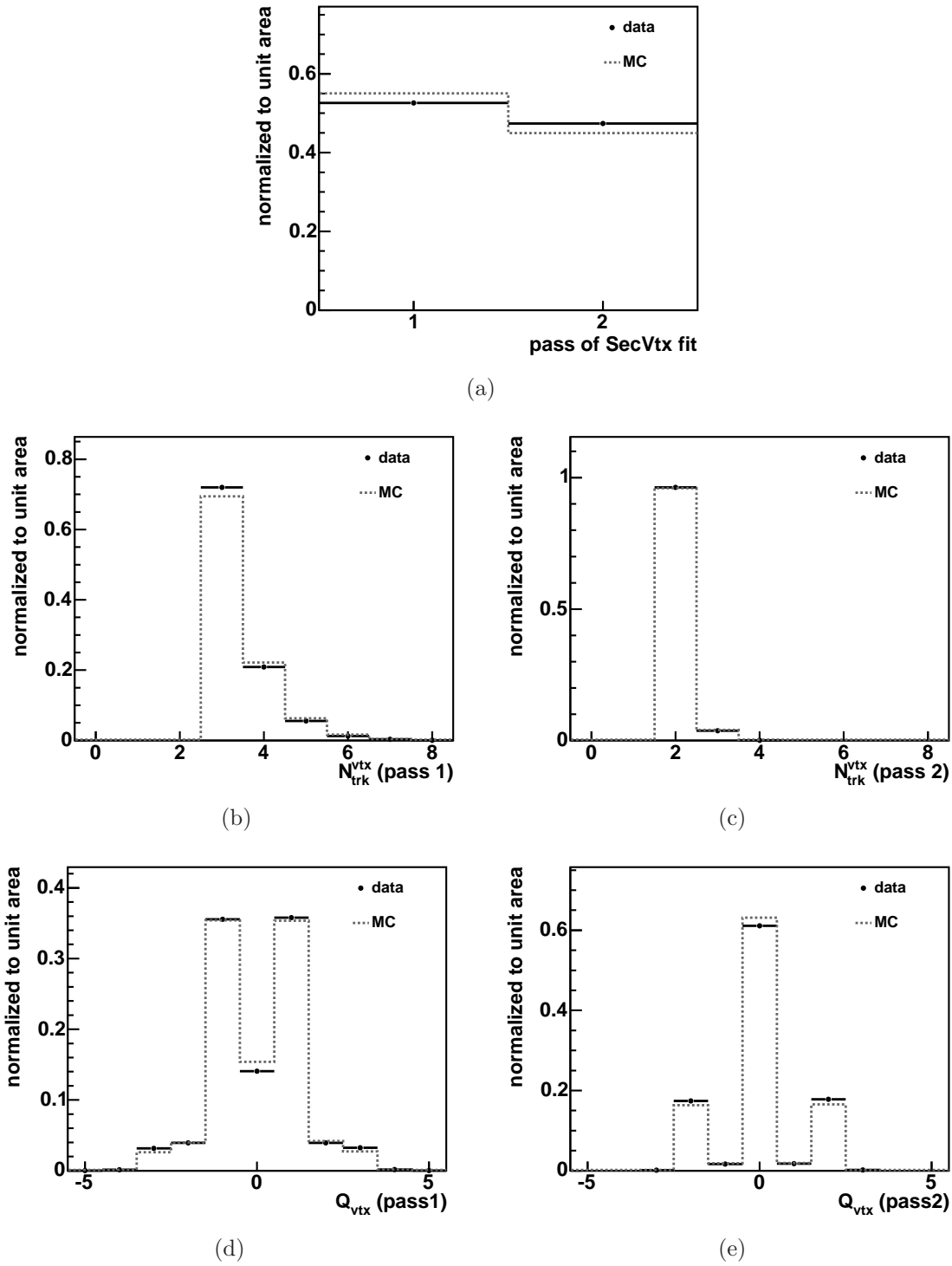


Figure C.11: Comparison of discriminating b -tagger variables in data and simulation: (a) the pass of the vertex fit, the number of tracks associated to (b) pass-1 and (c) pass-2 vertices, and the charge at (d) pass-1 and (e) pass-2 vertices.

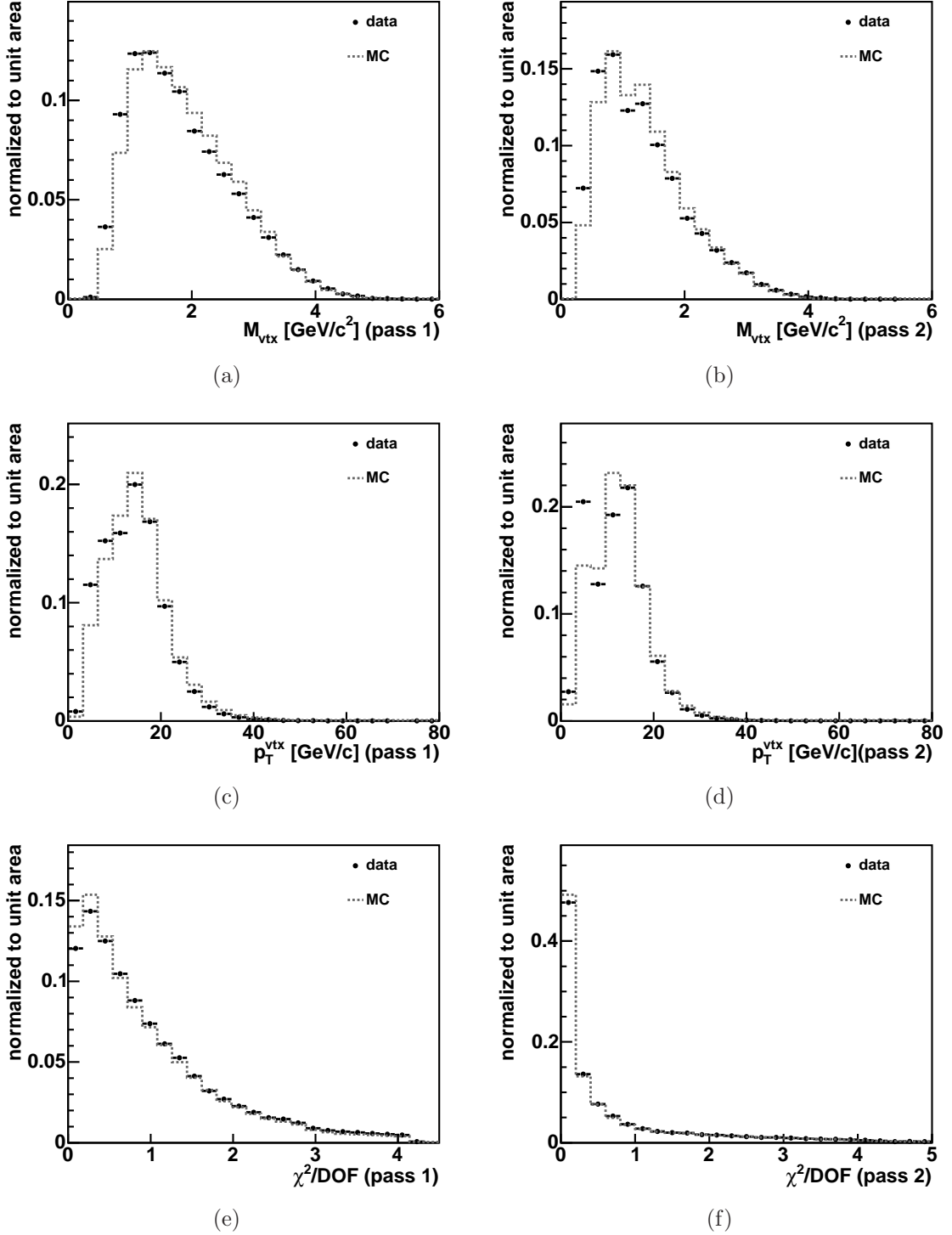


Figure C.12: Comparison of discriminating b -tagger variables in data and simulation: the invariant mass of (a) pass-1 and (b) pass-2 vertices, the transverse momentum at (c) pass-1 and (d) pass-2 vertices, and the χ^2 per degree of freedom for (e) pass-1 and (f) pass-2 vertices.

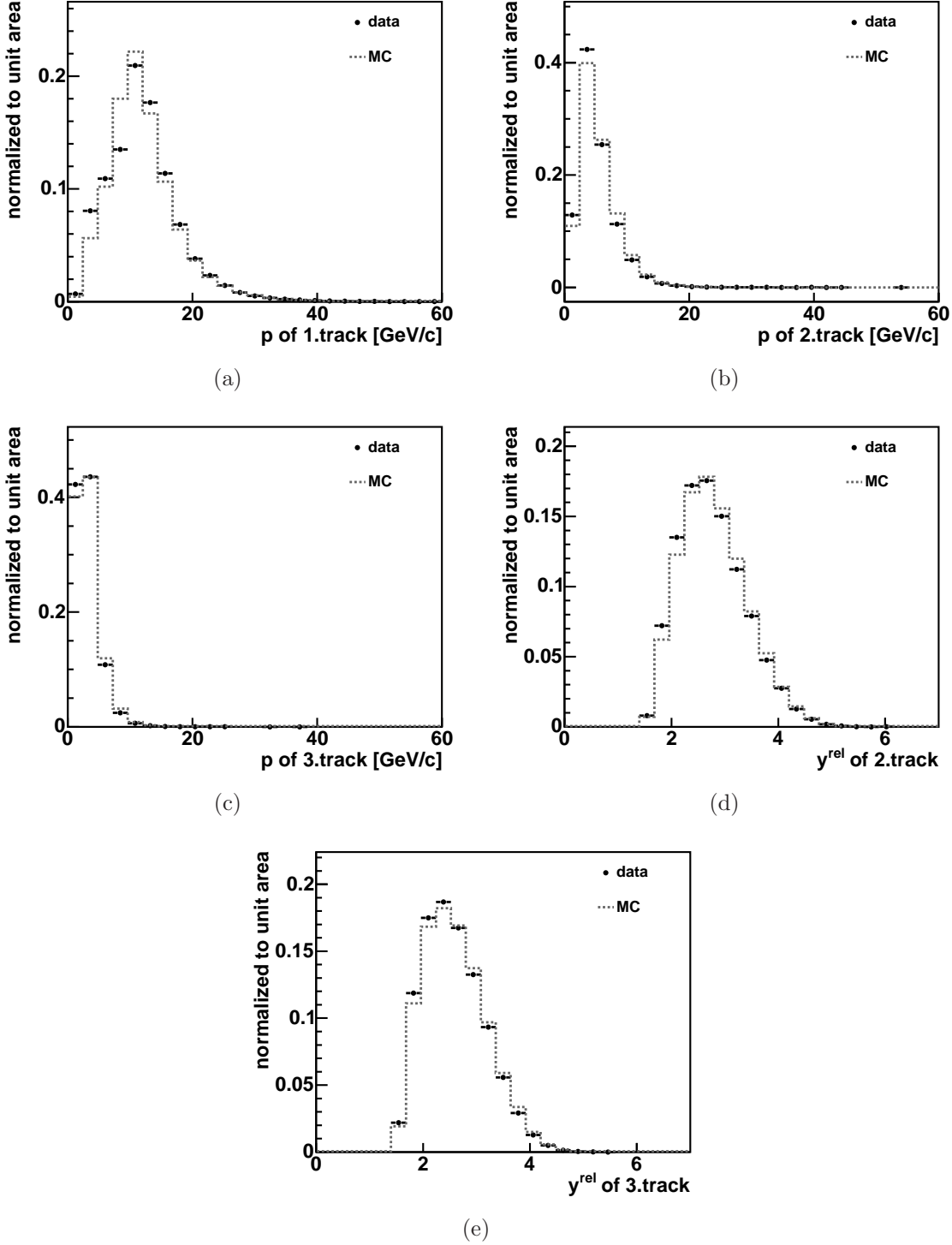


Figure C.13: Comparison of discriminating b -tagger variables in data and simulation: the momentum of the (a) 1.track, (b) 2.track, and (c) 3.track, and the rapidity y^{rel} with respect to the jet axis of the (d) 2.track and (e) 3.track.

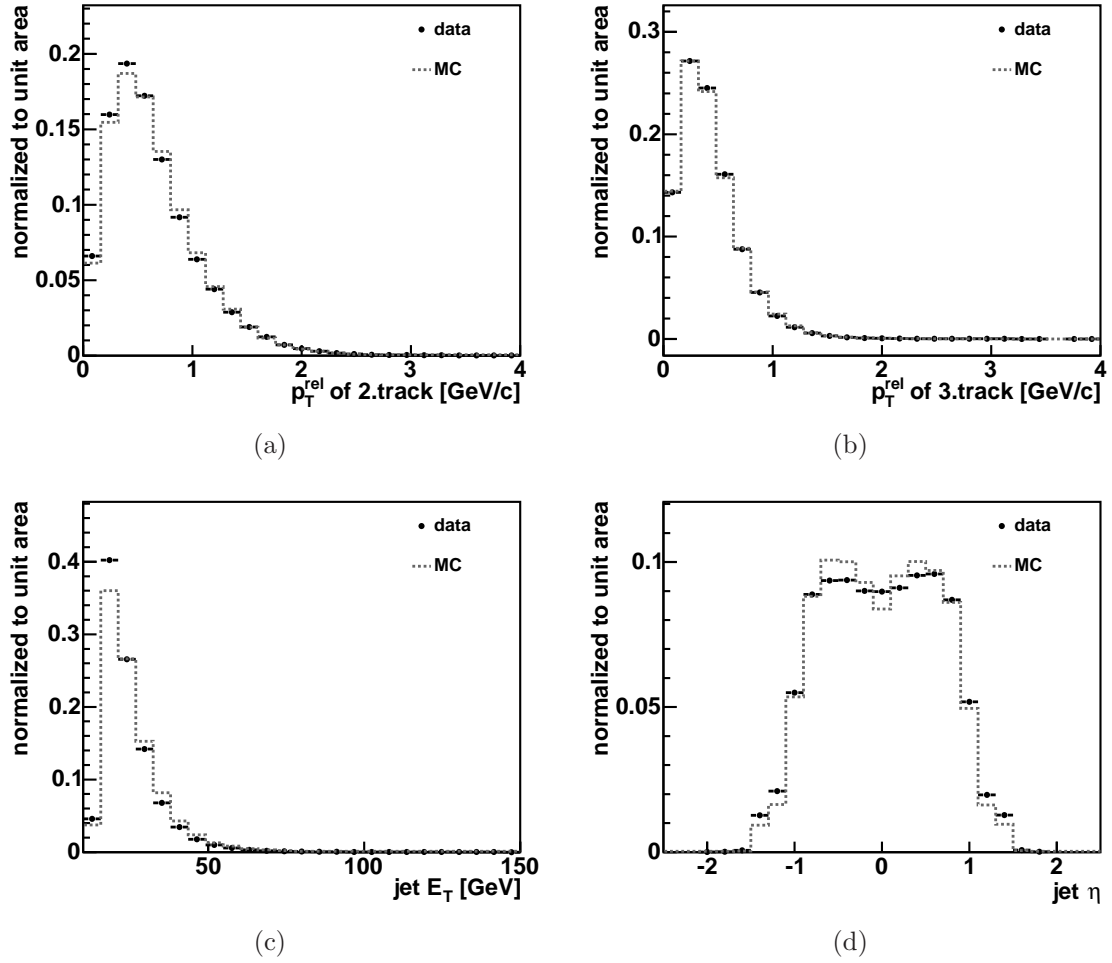
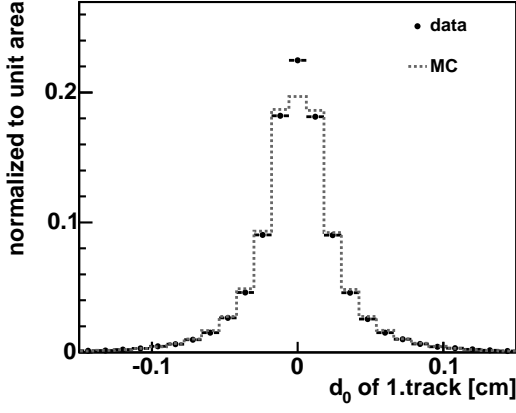
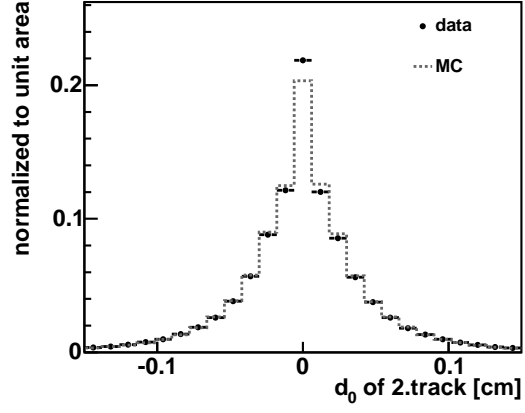


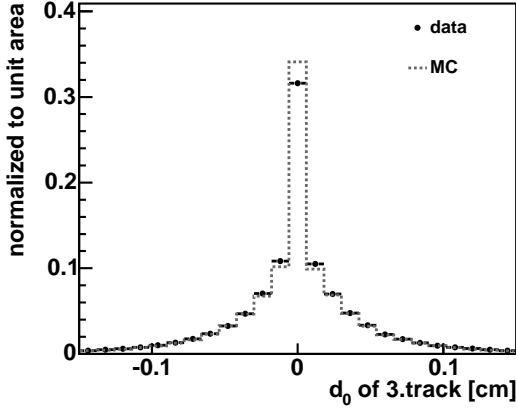
Figure C.14: Comparison of discriminating b -tagger variables in data and simulation: the transverse momentum p_T^{rel} with respect to the jet axis of the (a) 2.track and (b) 3.track, (c) jet E_T , and (d) jet η .



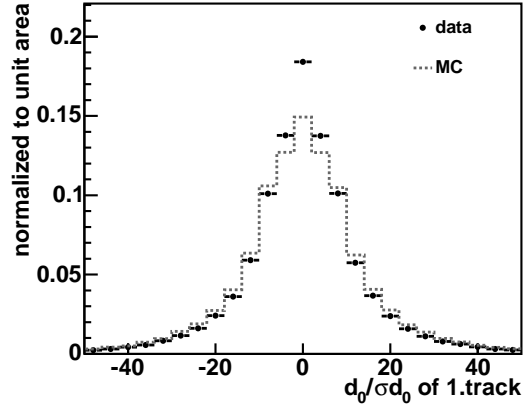
(a)



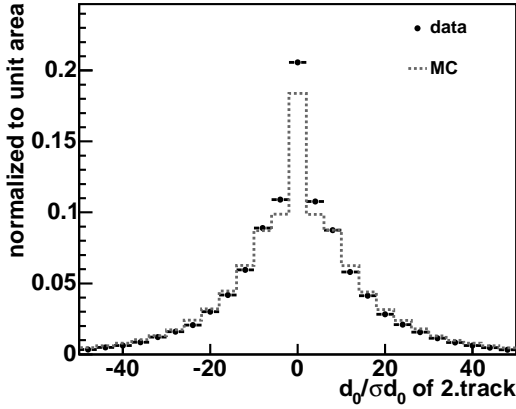
(b)



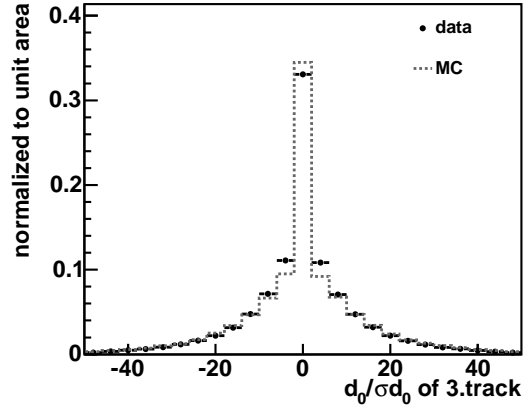
(c)



(d)



(e)



(f)

Figure C.15: Comparison of discriminating b -tagger variables in data and simulation: the impact parameter d_0 of the (a) 1.track, (b) 2.track, and (c) 3.track and the significance of the impact parameter of the (d) 1.track, (e) 2.track, and (f) 3.track.

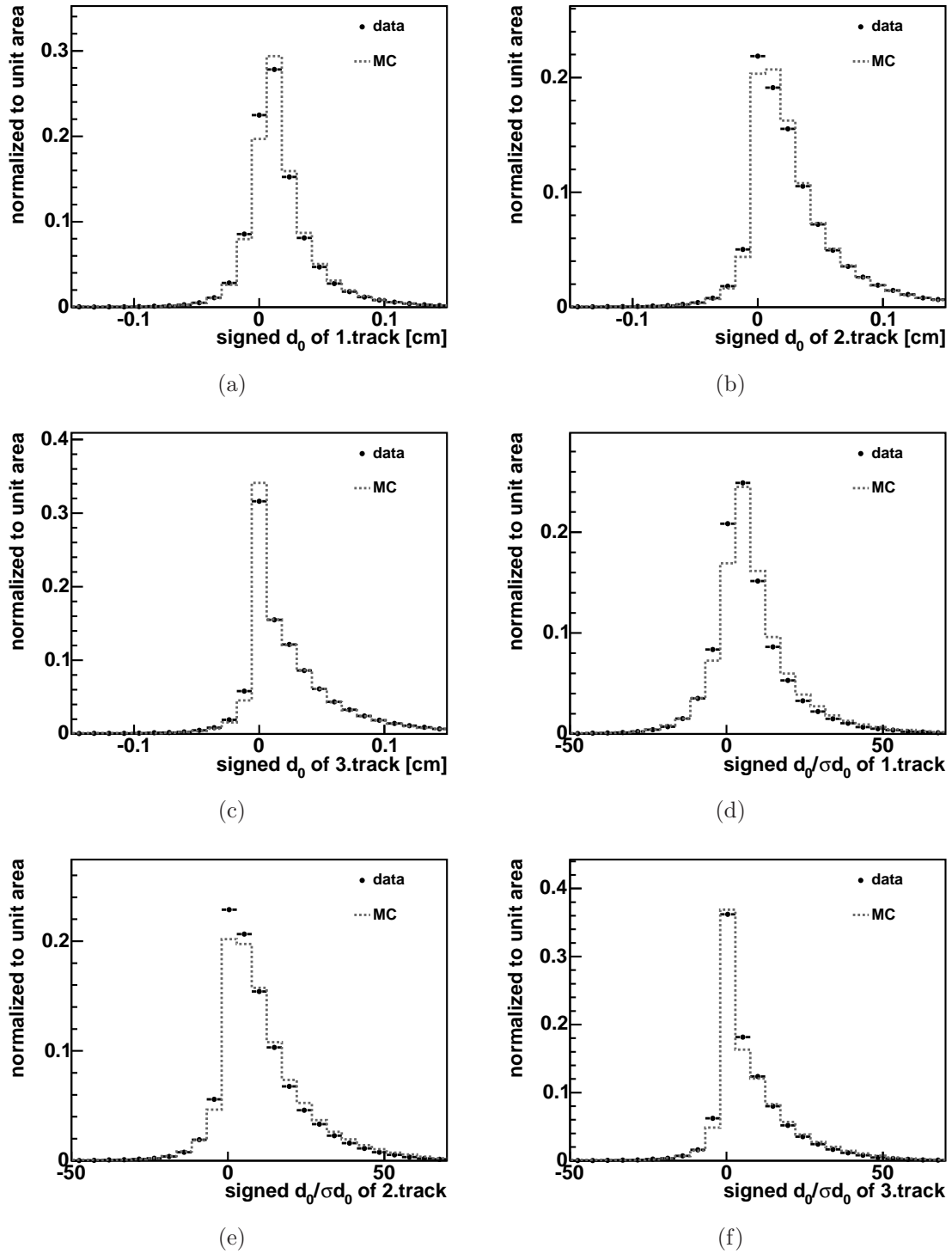
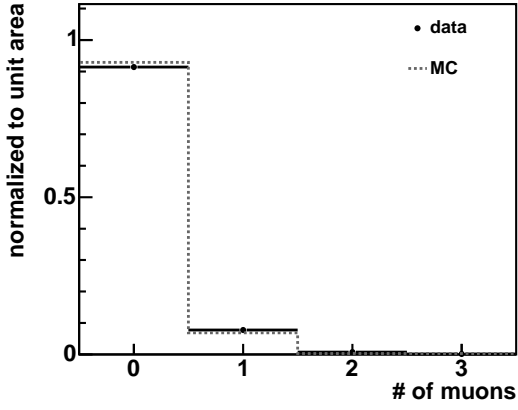
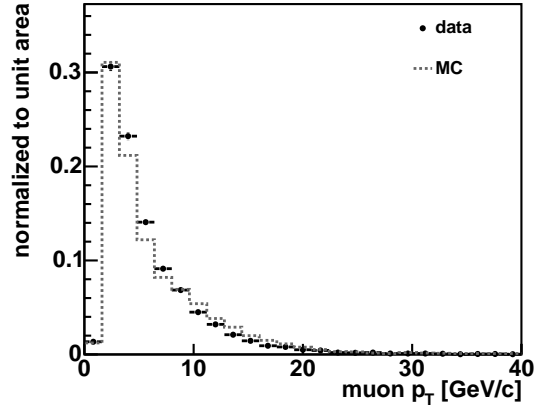


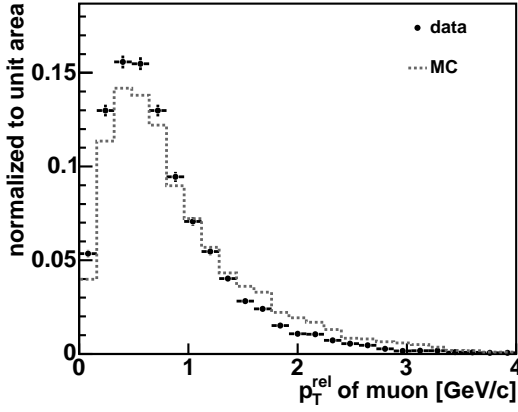
Figure C.16: Comparison of discriminating b -tagger variables in data and simulation: the signed impact parameter of the (a) 1.track, (b) 2.track, and (c) 3.track and the significance of the signed impact parameter of the (d) 1.track, (e) 2.track, and (f) 3.track.



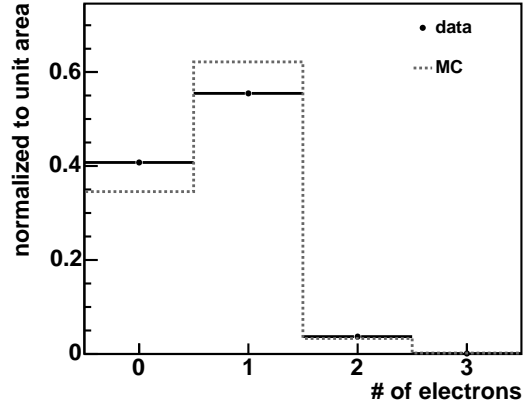
(a)



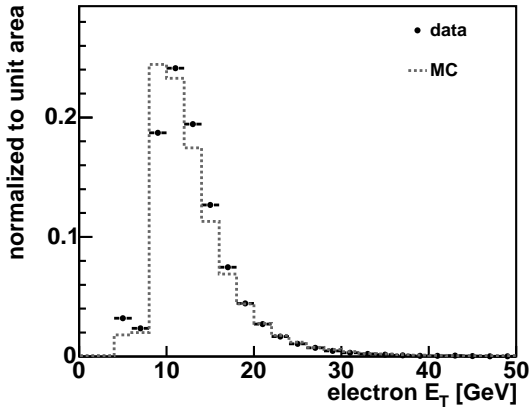
(b)



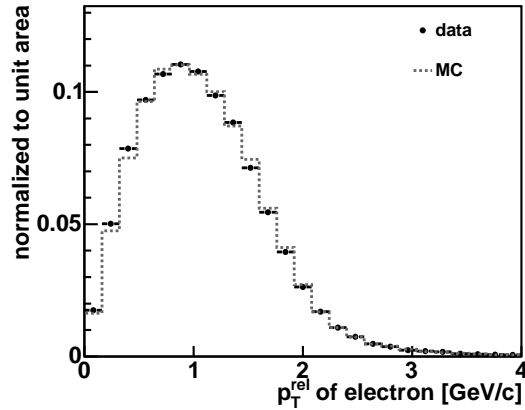
(c)



(d)



(e)



(f)

Figure C.17: Comparison of discriminating b -tagger variables in data and simulation: (a) the number of muons per jet and (b) p_T and (c) transverse momentum p_T^{rel} with respect to the jet axis of the most energetic muon in the jet, (d) the number of electrons per jet and (e) E_T and (f) p_T^{rel} of the most energetic electron in the jet.

Appendix D

Distributions Concerning the Single-Top Networks

D.1 Data-MC Comparison of Input Variables

Since the neural networks identifying single-top-quark events have been trained using simulated events, it is crucial to verify that the distributions obtained by a composition of signal and background models describe the distributions in observed events correctly. The different physics processes are modeled as described in section 5.3. Each process contributes a fraction corresponding to the respective expected number of events quoted in table 5.2. To allow a detailed comparison of the shapes of the distribution, the area of the modeled distribution is scaled to the number of observed events.

The comparisons of the neural-network b -tagger, the reconstructed top-quark mass $M_{\ell\nu b}$, the invariant mass of the two jets, and the product of the lepton charge times the pseudorapidity of the jet not assigned to the decay of the top quark $Q_\ell \cdot \eta_j$ have already been shown in figure 7.7. The comparison of the rest of the input variables introduced in section 7.1.3 can be found in figures D.1 to D.4. It can be seen that the distributions obtained by the composite model describe the ones of observed events well.

All those distributions as well as the comparisons of all distributions in all the subdetectors, comparisons of some control variables, and comparisons in control regions can be found in reference [127].

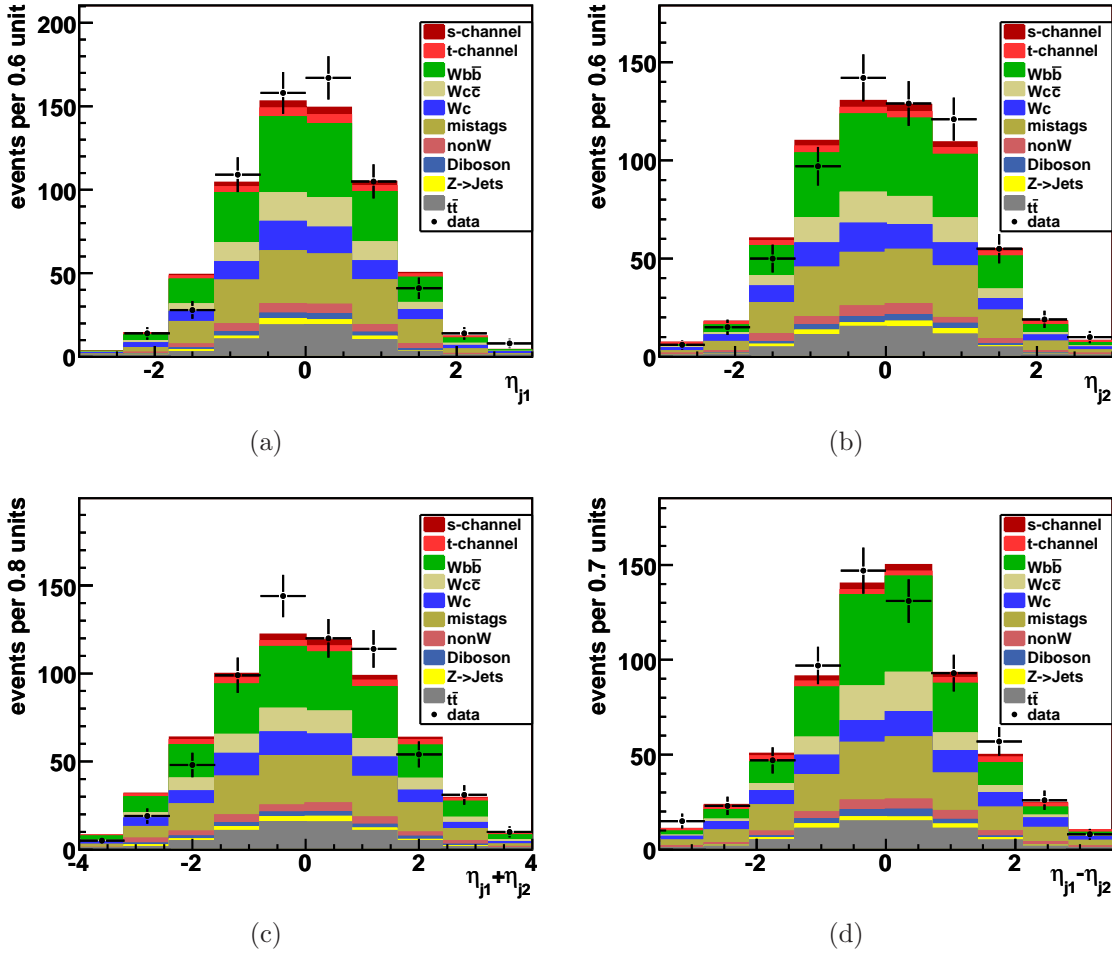


Figure D.1: Data-MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events: (a) the pseudorapidity of the leading jet $j1$, (b) the pseudorapidity of the second-leading jet $j2$, (c) the sum of the pseudorapidities of the two jets $\eta_{j1} + \eta_{j2}$, and (d) the difference between the pseudorapidities of the two jets $\eta_{j1} - \eta_{j2}$. The modeled distributions are scaled to the number of observed events.

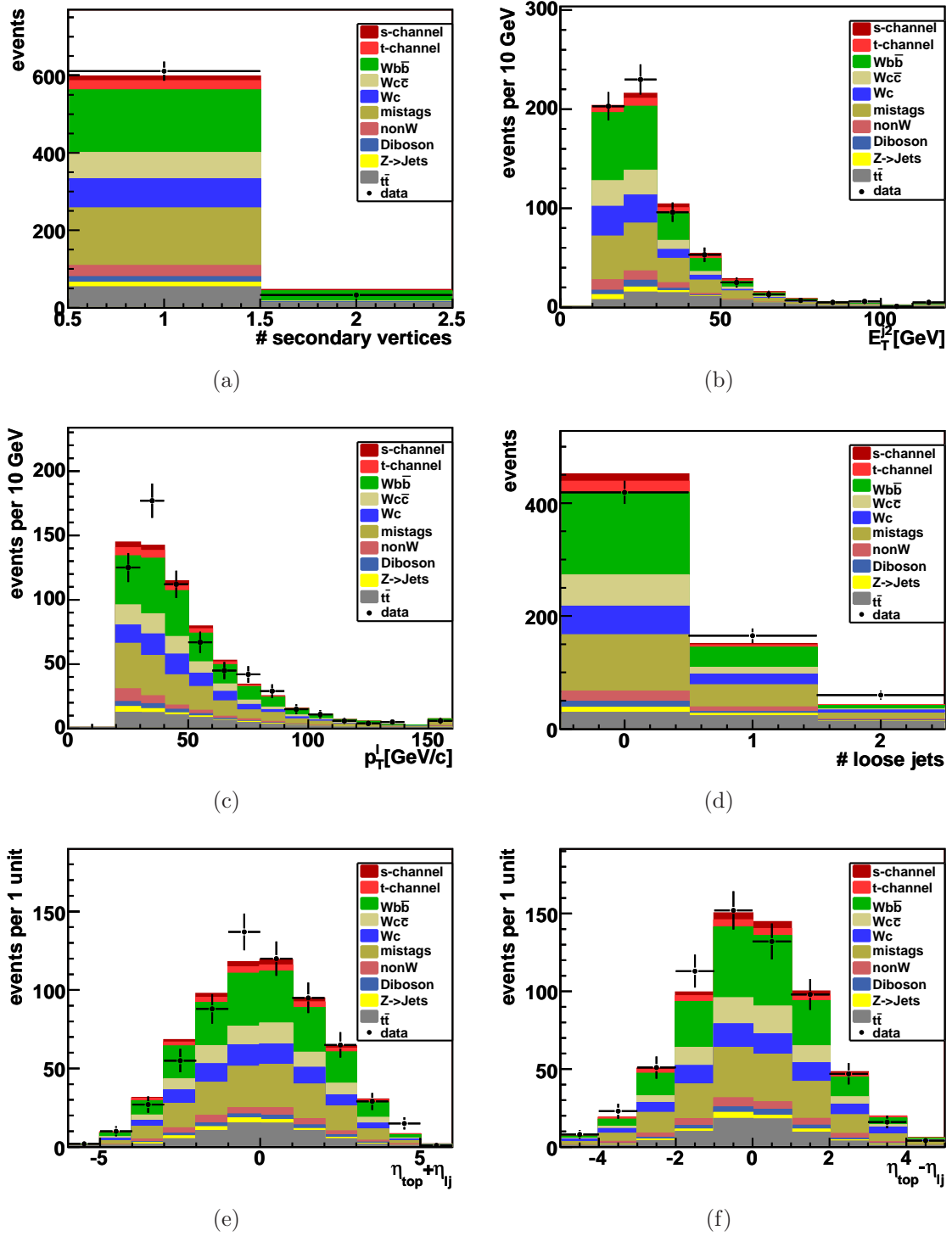


Figure D.2: Data–MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events: (a) the number of reconstructed secondary vertices, (b) the transverse energy of the second-leading jet j_2 , (c) the transverse momentum of the tight lepton, (d) the number of loose jets, (e) the sum of the pseudorapidities of the top quark and the light-quark jet $\eta_{\text{top}} + \eta_{\text{lj}}$, and (f) the difference between the pseudorapidity of the top quark and the pseudorapidity of the jet not assigned to the top-quark decay $\eta_{\text{top}} - \eta_{\text{lj}}$. The modeled distributions are scaled to the number of observed events.

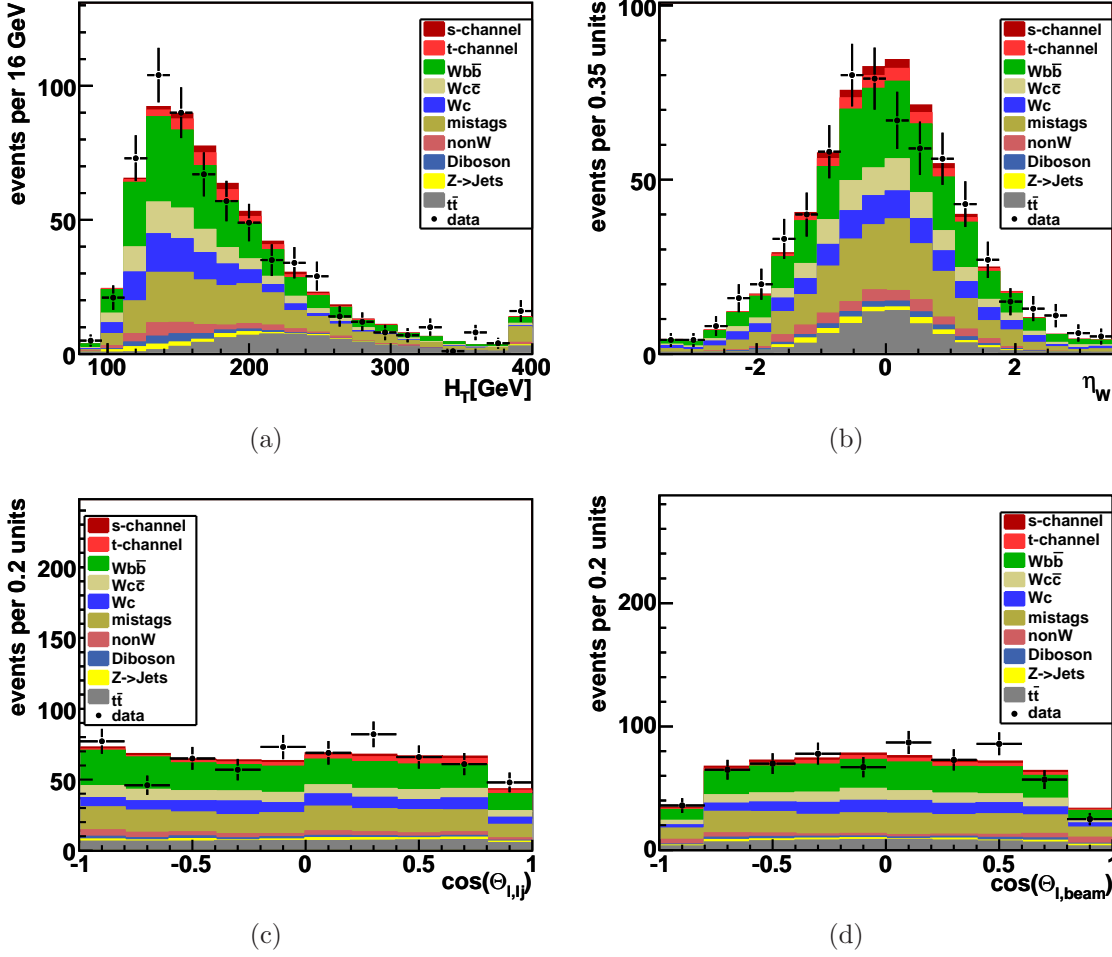


Figure D.3: Data-MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events: (a) $H_T = \sum_{jets} E_T + p_T^\ell + \cancel{E}_T$ where $\sum_{jets} E_T$ is the sum of the transverse energies of the tight and loose jets, (b) the pseudorapidity of the reconstructed W boson (see section 7.1.1), (c) the cosine of the polar angle in the top-quark rest frame between the tight lepton and the jet not assigned to the top-quark decay, and (d) the cosine of the polar angle between the tight lepton and the beam axis in the top-quark rest frame. The modeled distributions are scaled to the number of observed events.

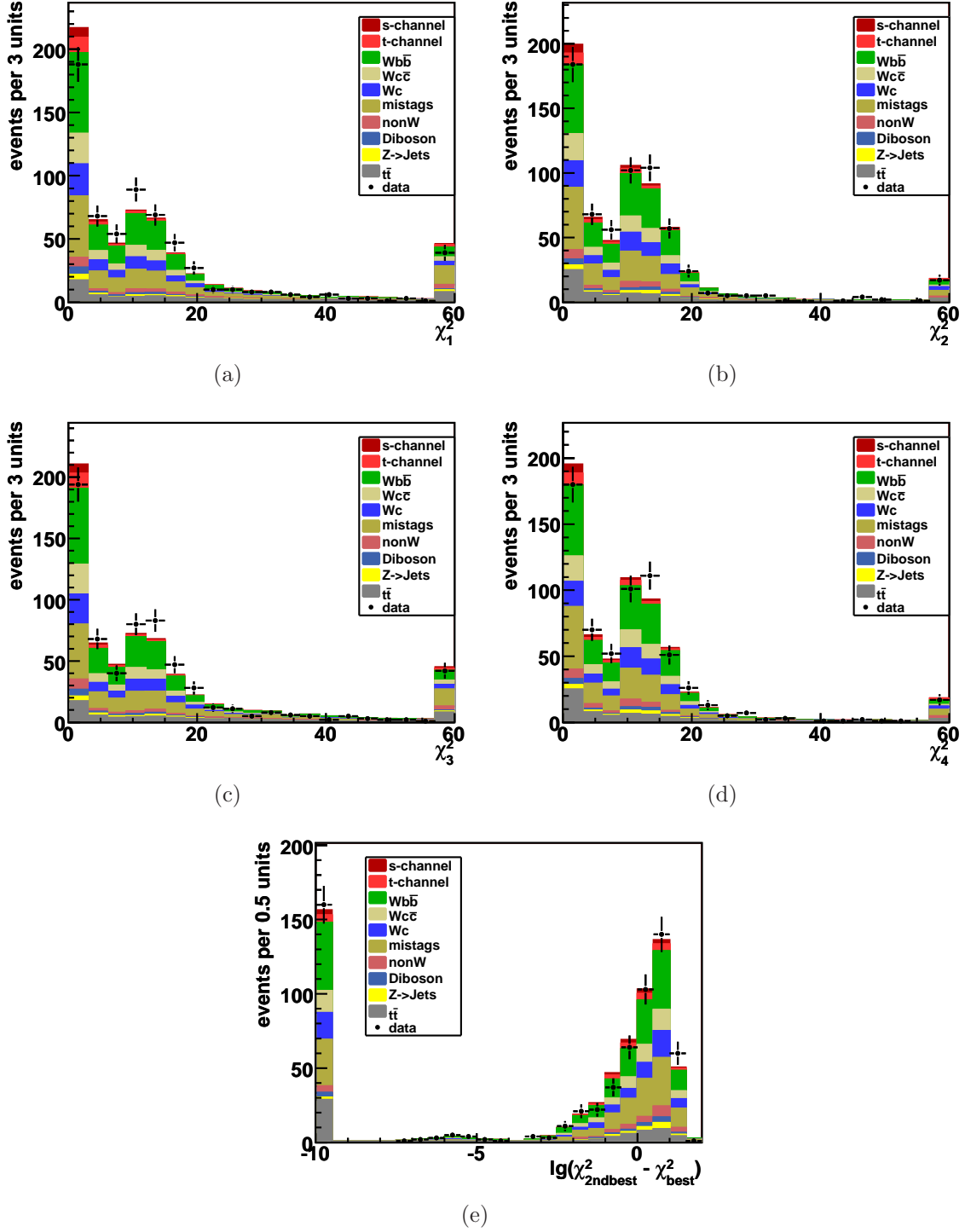
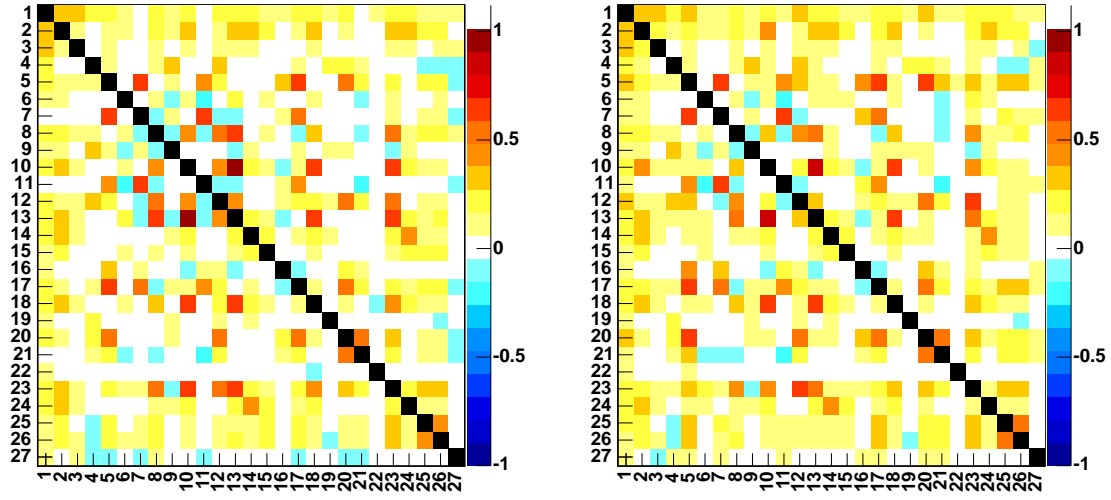


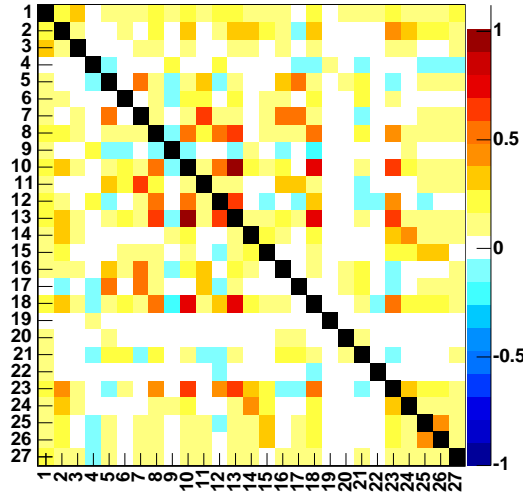
Figure D.4: Data–MC comparison of input variables for the training of the neural networks used to discriminate single–top–quark events, calculated by the kinematic fitter introduced in section 7.1.2: the χ^2 values of the four possible combinations of the objects to reconstruct the top quark: (a) χ_1^2 , (b) χ_2^2 , (c) χ_3^2 , (d) χ_4^2 , and (e) the logarithm of the difference between the χ^2 values of the second–best combination and the best combination $\chi_{2ndbest}^2 - \chi_{best}^2$. The modeled distributions are scaled to the number of observed events.

D.2 Correlations between Input Variables

During the preprocessing of the input variables, their correlations are determined as outlined in section 3.1.2. To illustrate the correlations between the variables, the correlation matrices of the three network trainings performed in the search for single top-quark production are depicted in figure D.5. The correlation matrix of the input variables for the combined network is shown in figure D.5(a), while the ones of the separate search are depicted in figures D.5(b) and D.5(c), respectively. The axes are labeled with numbers corresponding to the arbitrary order of the variables in the training process. It must be noted that label number 1 corresponds to the target value. That is why the numbering of the variables starts at 2. The key to break down those numbers into the names of the considered variables can be found in table D.5(d).



(a) combined

(b) t -channel(c) s -channel

Number	Variable	Number	Variable
2	$M_{\ell\nu b}$	15	$\lg(\chi_{2\text{ndbest}}^2 - \chi_{\text{best}}^2)$
3	NN b -tag output	16	η_{j1}
4	$\cos\Theta(\ell, l_j)$	17	η_{j2}
5	$Q_\ell \cdot \eta_{lj}$	18	H_T
6	η_W	19	$\cos\Theta(\ell, \text{beam})$
7	$\eta_{j1} + \eta_{j2}$	20	$\eta_{j1} - \eta_{j2}$
8	E_T^{j2}	21	$\eta_{\text{top}} - \eta_{lj}$
9	p_T^ℓ	22	# loose jets
10	E_T^{j1}	23	$M_{\ell\nu bb}$
11	$\eta_{\text{top}} + \eta_{lj}$	24	χ_1^2
12	M_{j1j2}	25	χ_2^2
13	$E_T^{j1} + E_T^{j2}$	26	χ_4^2
14	χ_3^2	27	# secondary vertices

(d) Meaning of Numbering

Figure D.5: The correlation matrices of the input variables are shown for the (a) combined network, (b) t -channel network, and fig:schanCorrelations s -channel network. In (d), the numbering of the axis labels is broken down into the different variables. Since label number 1 corresponds to the target, the numbering of the variables starts at 2.

D.3 Systematic Uncertainties

The uncertainties affect the rate of predicted signal and background events as well as the shape of the template histograms used in the fit to the observed data distribution. Some effects cause only rate uncertainties, some only shape uncertainties, but most of them affect both. The following sources of systematic uncertainties are considered: the uncertainty on the jet energy corrections, 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 uncertainty in modeling the output of the neural-network b -tagger, the uncertainty in the factorization and renormalization scale for the simulation of W +heavy flavor processes, the choice of the Monte Carlo event generator, the uncertainty in the event detection efficiency, the modeling of instrumental backgrounds, that is mistag events and non- W events, and the uncertainty in the luminosity determination. The impact of these sources of uncertainties is evaluated by altering the modeling of the corresponding processes or effects within their uncertainties or by assigning a plausible alternative model. As a result, relative changes of the shifted template distributions are obtained, being exemplarily illustrated by means of the template distributions utilized in the combined search. Rate uncertainties are only determined for single-top and $t\bar{t}$, quoted in table 7.5. For details of the performed variations and further information, see section 7.3.

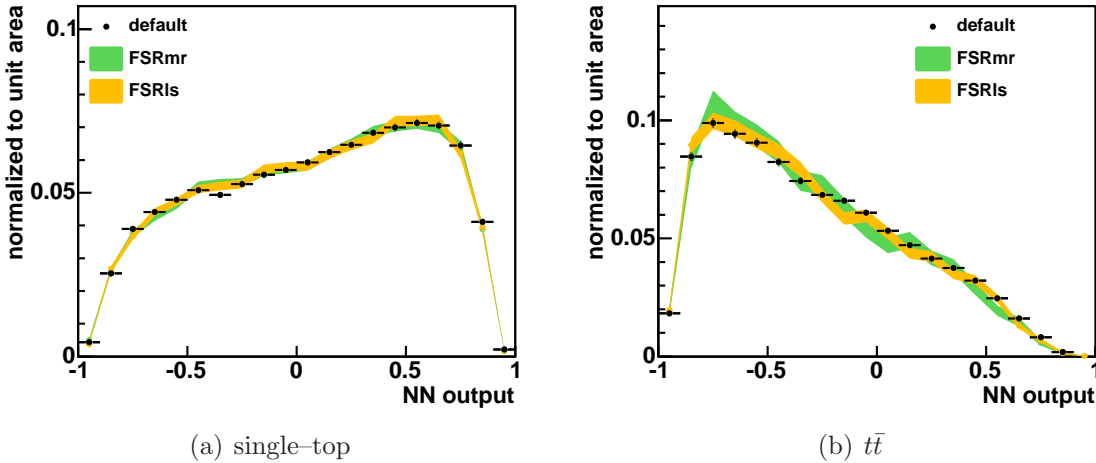


Figure D.6: Shape variation due to the uncertainty on final-state radiation. The template distributions for (a) single-top-quark and (b) $t\bar{t}$ events in the combined neural network analysis are shown. The default distributions are compared to the shifted distributions where the simulation was altered to produce either less or more final-state gluon radiation compared to the standard setting [128].

The impact of the uncertainty in the modeling of ISR on the shape of the template distributions has already been shown in figure 7.12, while the influence of the modeling of FSR is depicted in figure D.6. Systematic changes of the shape of the distributions induced by the parameterization of the PDF are obtained by combin-

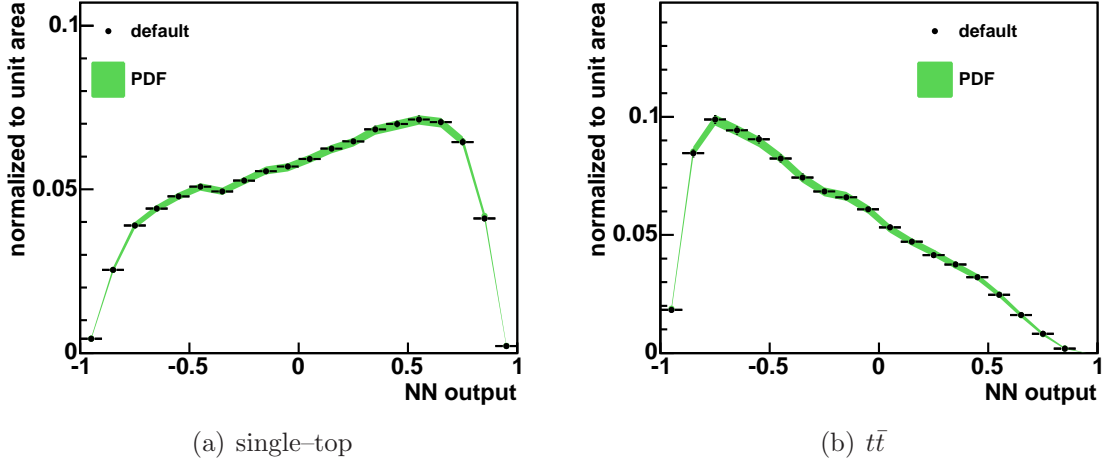


Figure D.7: Shape variation due to the choice of PDF parameterization. The template distributions for (a) single-top-quark and (b) $t\bar{t}$ events in the combined neural network analysis are shown. The default distributions are compared to the shifted distributions obtained from the 20 pairs of CTEQ6M eigenvectors [31].

ing the histograms achieved from the 20 pairs of CTEQ6M eigenvectors in such a way that the largest deviation from the standard distribution is determined bin-by-bin, leading to the variation of the output shape depicted in figure D.7. The influence of the variation of the output of the neural-network b -tagger on the shape of the template distributions is depicted in figure D.8. Figure D.9 illustrates the uncertainty induced by the selection of a different factorization and renormalization scale on the template histograms of W +heavy flavor events. The shape uncertainty due to the choice of the version of the ALPGEN Monte Carlo program is illustrated in figure D.10(a). The variations of the template shapes due to the modeling of the mistags and the non- W background models are shown in figures D.10(b) and D.11, respectively.

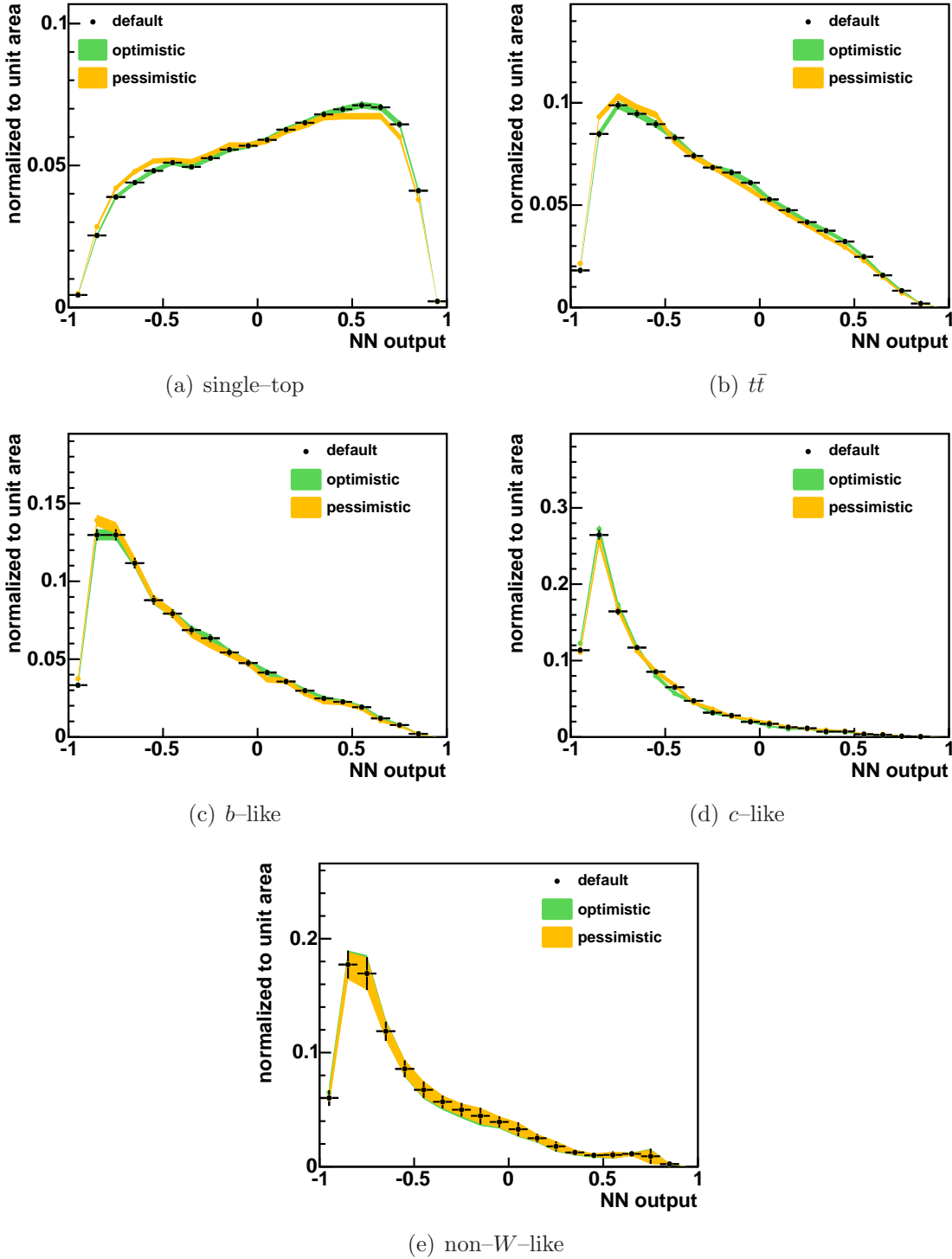


Figure D.8: Shape variation due to the uncertainty on the shape of the neural-network b -tagger. The distributions of (a) single-top template, (b) $t\bar{t}$ template, (c) b -like template, (d) c -like template, and (e) non- W -like template utilized in the combined neural network analysis are shown. The default distributions are compared to the shifted distributions obtained from the optimistic and pessimistic scenarios introduced in section 6.4.1.

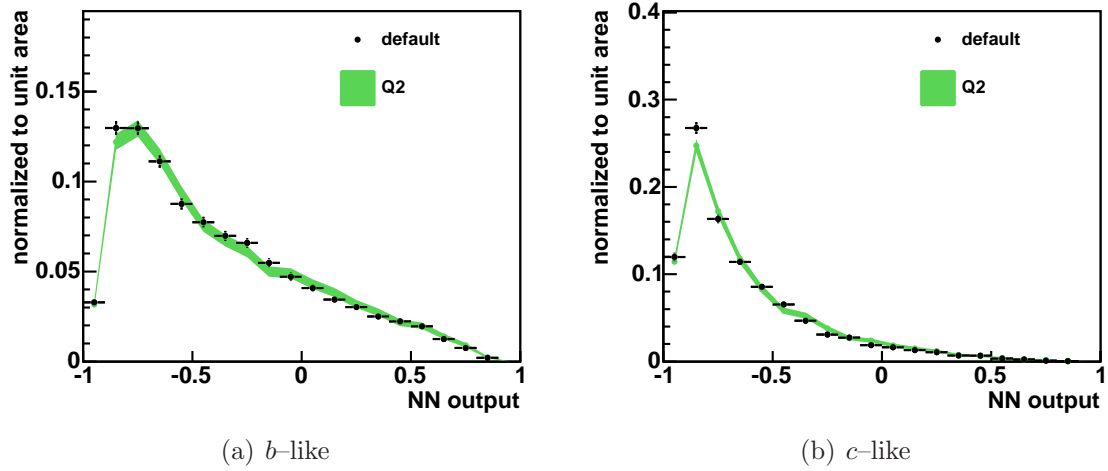


Figure D.9: Shape variation due to a different choice of the factorization and renormalization scale. The distributions of the (a) b -like and (b) c -like templates used in the combined neural network analysis are shown. The default distributions are compared to the shifted distributions obtained from events generated with a different choice of the factorization and renormalization scale.

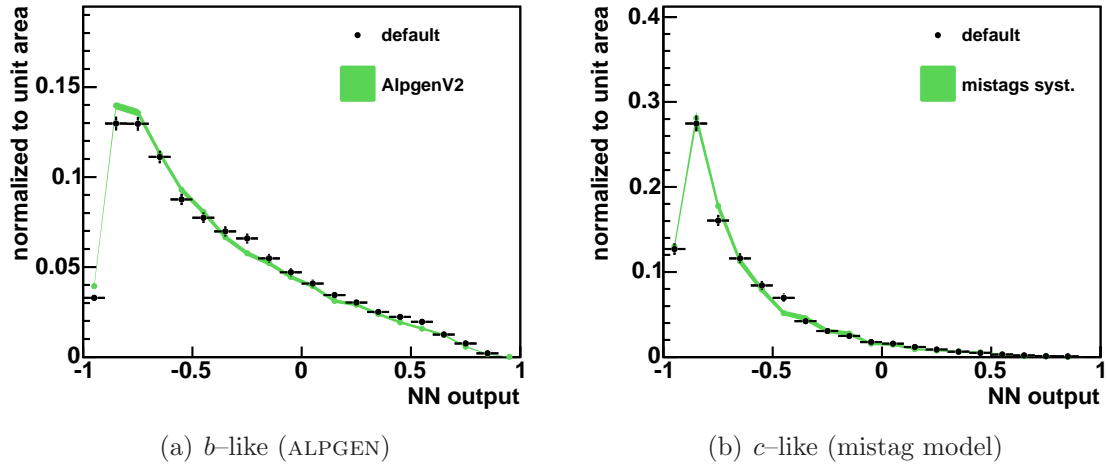


Figure D.10: Shape variation due to the choice of the utilized version of ALPGEN and the uncertainty on the mistag modeling. (a) The b -like template distributions in the combined analysis are shown. The default distribution is compared to the shifted distribution where ALPGEN version 2.0 was used instead of the default version 1.3.3. (b) The c -like template distributions in the combined neural network analysis are shown. The shape obtained by the default mistag model described in section 5.3 is compared to the shape obtained by an alternative model based on observed events.

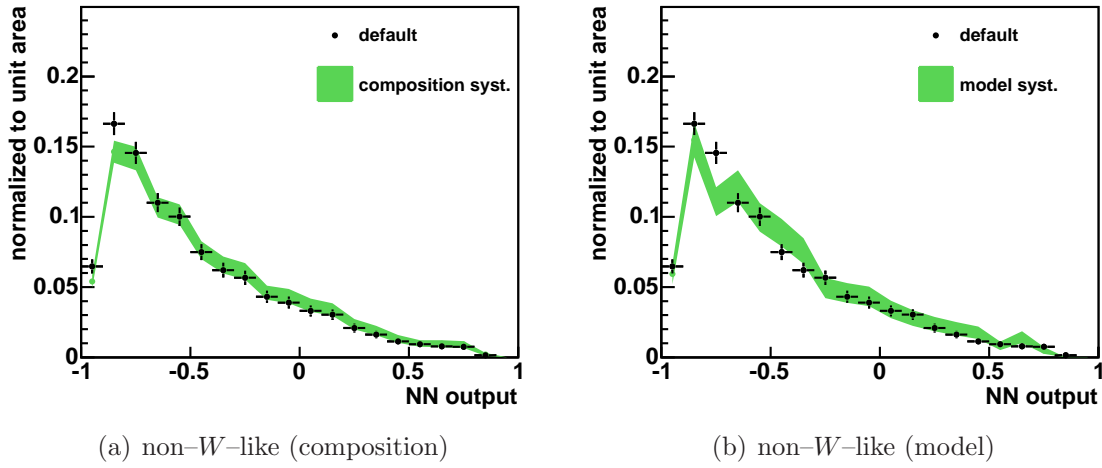


Figure D.11: Shape variation due to the uncertainty on the non- W background modeling. The default distribution of the non- W -like template is compared to distributions where (a) a different flavor composition of the non- W background is assumed and (b) a different model for the non- W background was considered. The default distribution is obtained using the non- W model described in section 5.3.

List of Figures

1.1	Feynman diagram of electron–electron scattering.	4
1.2	The CTEQ5L parton distribution function at $\mu^2 = (175 \text{ GeV})^2$	5
1.3	Leading–order Feynman diagrams of top quark pair production	6
1.4	Feynman graphs of t –channel single–top quark production	7
1.5	Feynman graphs of s –channel single–top quark production	8
1.6	Feynman graph of associated single–top quark production	8
1.7	Theoretical prediction of the top–quark polarization in t –channel events	10
2.1	Aerial view of the Fermilab facility	13
2.2	Schematic overview of the accelerator chain at Fermilab.	14
2.3	The initial luminosity in Run II as a function of time.	17
2.4	Evolution of the delivered and recorded integrated luminosity	17
2.5	Cutaway view of the CDF II detector	18
2.6	Longitudinal view of the CDF II tracking system and the plug calorime- try	19
2.7	Schematic overview of the components of the silicon microstrip detector	20
2.8	Functional block diagram of the CDF II data flow	23
3.1	General geometry of a three–layer neural network	25
3.2	The transformed sigmoid activation function $S(a(x))$	26
3.3	Illustration of Neural Network Output	29
4.1	Track Parameterization	35
4.2	Schematic overview of jet reconstruction.	39
4.3	η –dependence of calorimeter response	40
4.4	N_{vtx} –dependence of jet energy	41
4.5	Absolute jet energy corrections	42

4.6	Out-of-cone jet energy corrections	43
4.7	Schematic view of secondary vertex reconstruction	44
4.8	SecVtx tag efficiency	47
4.9	SecVtx mistag rate	48
5.1	Illustration of QCD background rejection in the forward calorimeter .	52
5.2	Examples of W +heavy flavor background production	53
5.3	Examples of diboson and Z +jets background production	54
5.4	Examples of QCD background production	54
5.5	Matching of t -channel single top-quark production	56
6.1	Discriminating variables for the identification of b -quark jets	64
6.2	Discriminating variables for the identification of b -quark jets	65
6.3	More discriminating variables for the identification of b -quark jets . .	66
6.4	Vertex variables in dependence on the pass of the SecVtx fit	67
6.5	Performance of the neural-network b -tagger	71
6.6	Output of the neural-network b -tagger in dependence on the pass of the SecVtx fit	72
6.7	Comparison of discriminating variables in data and simulation	74
6.8	Comparison of discriminating variables in data and simulation	75
6.9	Comparison of discriminating variables in data and simulation	76
6.10	Fit of M_{vtx} templates to the away jet distribution in data	77
6.11	Comparison of output distributions of away jets in data and simulation	78
6.12	Demonstration of systematic b -quark correction	79
6.13	Fit of M_{vtx} templates to the electron jet distribution in data	80
6.14	Comparison of output distributions of electron jets in data and sim- ulation	82
6.15	Output distributions of mistags in data and simulation	83
6.16	Output distributions for mistagged pass-1 and pass-2 vertices in data and simulation	84
6.17	Parameter-dependence of the difference between pass-1 mistags in data and simulation	85
6.18	Parameter-dependence of the difference between pass-2 mistags in data and simulation	86

6.19	Functions correcting for the observed differences between mistags in data and simulation	87
6.20	Demonstration of mistag correction	88
6.21	Use of b tagger in search for single top-quark production	89
6.22	Illustration of estimating a systematic shape uncertainty for the output of the neural-network b -tagger	90
6.23	Use of the b tagger to estimate the flavor composition	91
7.1	Input variables for the training of the neural networks used to discriminate single-top-quark events	96
7.2	Input variables for the training of the neural networks used to discriminate single-top-quark events	97
7.3	Input variables for the training of the neural networks used to discriminate single-top-quark events	99
7.4	Input variables for the training of the neural networks used to discriminate single-top-quark events	101
7.5	Input variables for the training of the neural networks used to discriminate single-top-quark events	102
7.6	Input variables for the training of the neural networks used to discriminate single-top-quark events	104
7.7	Data-MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events	105
7.8	Results of the trainings of the three neural networks	112
7.9	Template categories	113
7.10	Templates utilized in the combined search	114
7.11	Templates utilized in the separate search	115
7.12	Shape variation due the uncertainty on initial-state radiation	117
7.13	Symmetrized shape variation due the uncertainty on initial-state radiation	119
7.14	Expected uncertainty of the measurement of the combined search	122
7.15	The expected uncertainties of the measurements and upper limits of the separate search	123
7.16	Distributions of expected Q -values	125
7.17	Expected upper limit of the combined search	125
7.18	Expectation and observation in the high-output region of the combined analysis	126

7.19	Expectation and observation in the separate search	127
7.20	Observed Q -value of the combined search	129
7.21	Fit result of the combined search	129
7.22	Observed Q -value of the separate search	130
7.23	Fit result of the separate search	131
C.1	Discriminating variables for the identification of b -quark jets	147
C.2	Discriminating variables for the identification of b -quark jets	148
C.3	Discriminating variables for the identification of b -quark jets	149
C.4	Discriminating variables for the identification of b -quark jets	150
C.5	Discriminating variables for the identification of b -quark jets	151
C.6	Discriminating variables for the identification of b -quark jets	152
C.7	Discriminating variables for the identification of b -quark jets	153
C.8	Discriminating variables for the identification of b -quark jets	154
C.9	Comparison of discriminating b -tagger variables in data and simulation	156
C.10	Comparison of discriminating b -tagger variables in data and simulation	157
C.11	Comparison of discriminating b -tagger variables in data and simulation	158
C.12	Comparison of discriminating b -tagger variables in data and simulation	159
C.13	Comparison of discriminating b -tagger variables in data and simulation	160
C.14	Comparison of discriminating b -tagger variables in data and simulation	161
C.15	Comparison of discriminating b -tagger variables in data and simulation	162
C.16	Comparison of discriminating b -tagger variables in data and simulation	163
C.17	Comparison of discriminating b -tagger variables in data and simulation	164
D.1	Data-MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events	166
D.2	Data-MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events	167
D.3	Data-MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events	168
D.4	Data-MC comparison of input variables for the training of the neural networks used to discriminate single-top-quark events	169
D.5	Correlation matrices of input variables to the single-top-quark searches	171
D.6	Shape variation due to the uncertainty on final-state radiation	172

D.7	Shape variation due to the choice of PDF parameterization	173
D.8	Shape variation due to the uncertainty on the shape of the neural- network b -tagger	174
D.9	Shape variation due to a different choice of the factorization and renormalization scale	175
D.10	Shape variation due to the choice of the utilized version of ALPGEN and the uncertainty on the mistag modeling	175
D.11	Shape variation due to the uncertainty on the non- W background modeling	176

List of Tables

1.1	The fundamental fermionic particles	2
1.2	The gauge bosons of the standard model	3
2.1	Properties of the different calorimetry systems	21
2.2	Design parameters of the CDF II muon detectors	22
5.1	Data samples used for the search for single top quarks	50
5.2	Expected event yield	62
6.1	Set of discriminating variables used to train the neural-network b -tagger	69
6.2	Set of discriminating variables discarded for the training of the neural-network b -tagger	70
6.3	Example for the cut efficiency of the b tagger	71
6.4	Flavor composition of tagged away jets	76
6.5	Flavor composition of tagged electron jets	81
6.6	Comparison of differently estimated contributions of the various flavors to the observed data passing the single-top-quark event selection	92
7.1	Composition of the training samples for the neural networks used to discriminate single-top-quark events	107
7.2	Set of discriminating variables used to train the combined neural network	108
7.3	Set of discriminating variables used to train the t -channel neural network	109
7.4	Set of discriminating variables used to train the s -channel neural network	110
7.5	Relative systematic rate uncertainties for single-top-quark and $t\bar{t}$ events.	118

7.6	Gaussian constraints on the predicted number of events	120
7.7	Estimated strengths of systematic effects	124
7.8	Fitted rates of the considered processes and strengths of systematic effects of the combined search	128
7.9	Fitted rates of the considered processes and strengths of systematic effects of the separate search	130
A.1	Baseline cuts for central electrons	137
A.2	Baseline cuts for plug electrons using the phoenix tracking	138
A.3	Baseline cuts for CMUP and CMX muons	139
A.4	The baseline cuts for electrons used to exploit semileptonic b -hadron decays.	140
B.1	MC samples used in this analysis	144
B.2	MC samples used to estimate systematic effects	145
B.3	MC samples used to determine NN b -tagger corrections	146

Bibliography

- [1] M. Gell-Mann. A schematic model of baryons and mesons. Phys. Lett., 8:214–215, 1964.
- [2] F. Abe et al. Observation of top quark production in anti-p p collisions. Phys. Rev. Lett., 74:2626–2631, 1995.
- [3] S. Abachi et al. Observation of the top quark. Phys. Rev. Lett., 74:2632–2637, 1995.
- [4] V. M. Abazov et al. Search for single top quark production at DØ using neural networks. Phys. Lett., B517:282–294, 2001.
- [5] B. Abbott et al. Search for electroweak production of single top quarks in $p\bar{p}$ collisions. Phys. Rev., D63:031101, 2001.
- [6] D. Acosta et al. Search for single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV. Phys. Rev., D65:091102, 2002.
- [7] D. Acosta et al. Optimized search for single top quark production at the Fermilab Tevatron. Phys. Rev., D69:052003, 2004.
- [8] D. Acosta et al. Search for electroweak single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. Phys. Rev., D71:012005, 2005.
- [9] V. M. Abazov et al. Search for single top quark production in p anti-p collisions at $\sqrt{s} = 1.96$ TeV. Phys. Lett., B622:265–276, 2005.
- [10] O. W. Greenberg. Spin and unitary spin independence in a paraquark model of baryons and mesons. Phys. Rev. Lett., 13:598–602, 1964.
- [11] M. Y. Han and Y. Nambu. Three-triplet model with double SU(3) symmetry. Phys. Rev., 139:B1006–B1010, 1965.
- [12] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. Phys. Rev. Lett., 81:1562–1567, 1998.
- [13] 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.

- [14] Q. R. Ahmad et al. Measurement of the charged current interactions produced by B-8 solar neutrinos at the Sudbury Neutrino Observatory. Phys. Rev. Lett., 87:071301, 2001.
- [15] Q. R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. Phys. Rev. Lett., 89:011301, 2002.
- [16] W.-M. Yao et al. Review of particle physics. J. Phys., G33:1–1232, 2006.
- [17] Tevatron Electroweak Working Group. Combination of CDF and D0 results on the mass of the top quark, 2006.
- [18] E. Noether. Invariante Variationsprobleme. Nachr. v. d. Ges. d. Wiss. zu Goettingen, pages 235–257, 1918.
- [19] D. J. Gross and F. Wilczek. Ultraviolet behavior of non-abelian gauge theories. Phys. Rev. Lett., 30:1343–1346, 1973.
- [20] H. D. Politzer. Reliable perturbative results for strong interactions? Phys. Rev. Lett., 30:1346–1349, 1973.
- [21] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the color octet gluon picture. Phys. Lett., B47:365–368, 1973.
- [22] N. Cabibbo. Unitary symmetry and leptonic decays. Phys. Rev. Lett., 10:531–532, 1963.
- [23] M. Kobayashi and T. Maskawa. CP violation in the renormalizable theory of weak interaction. Prog. Theor. Phys., 49:652–657, 1973.
- [24] S. L. Glashow. Partial symmetries of weak interactions. Nucl. Phys., 22:579–588, 1961.
- [25] A. Salam and J. C. Ward. Electromagnetic and weak interactions. Phys. Lett., 13:168–171, 1964.
- [26] S. Weinberg. A model of leptons. Phys. Rev. Lett., 19:1264–1266, 1967.
- [27] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. Phys. Rev. Lett., 13:321–322, 1964.
- [28] P. W. Higgs. Broken symmetries and the masses of gauge bosons. Phys. Rev. Lett., 13:508–509, 1964.
- [29] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. Phys. Rev. Lett., 13:585–587, 1964.
- [30] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. Phys. Rev., 145:1156–1163, 1966.

- [31] H. L. Lai et al. Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. Eur. Phys. J. C, 12:375–392, 2000.
- [32] Tevatron Electroweak Working Group. A combination of CDF and D0 results on the mass of the top quark, 2007.
- [33] W. Wagner. Top quark physics in hadron collisions. Rept. Prog. Phys., 68:2409–2494, 2005.
- [34] 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.
- [35] S. S. D. Willenbrock and D. A. Dicus. Production of heavy quarks from W gluon fusion. Phys. Rev., D34:155, 1986.
- [36] G. Mahlon and S. J. Parke. Improved spin basis for angular correlation studies in single top quark production at the Tevatron. Phys. Rev., D55:7249–7254, 1997.
- [37] 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.
- [38] Z. Sullivan. Understanding single-top-quark production and jets at hadron colliders. Phys. Rev., D70:114012, 2004.
- [39] T. M. P. Tait. The t W^- mode of single top production. Phys. Rev., D61:034001, 2000.
- [40] A. Belyaev and E. E. Boos. Single top quark $tW+X$ production at the LHC: a closer look. Physical Review D, 63:034012, 2001.
- [41] A. P. Heinson, A. S. Belyaev, and E. E. Boos. Single top quarks at the Fermilab Tevatron. Phys. Rev., D56:3114–3128, 1997.
- [42] M. Jezabek. Top quark physics. Nucl. Phys. Proc. Suppl., 37B:197, 1994.
- [43] T. Stelzer, Z. Sullivan, and S. Willenbrock. Single top quark production at hadron colliders. Phys. Rev., D58:094021, 1998.
- [44] J. Alwall et al. Is $V(tb) = 1$? Eur. Phys. J., C49:791–801, 2007.
- [45] J. Conway et al. Search for heavy top $t' \rightarrow Wq$ in lepton plus jet events. Public CDF note 8495, 2007.
- [46] D. Mohl, G. Petrucci, L. Thorndahl, and S. Van Der Meer. Physics and technique of stochastic cooling. Phys. Rept., 58:73–119, 1980.
- [47] S. Van Der Meer. Stochastic cooling and the accumulation of anti-protons. Rev. Mod. Phys., 57:689–697, 1985.

- [48] G. I. Budker. An effective method of damping particle oscillations in proton and anti-proton storage rings. Sov. Atom. Energ., 22:438–440, 1967.
- [49] S. Nagaitsev et al. Experimental demonstration of relativistic electron cooling. Phys. Rev. Lett., 96:044801, 2006.
- [50] X. Zhang, V. Shiltsev, F. Zimmermann, and K. Bishofberger. The special applications of Tevatron electron lens in collider operation. Presented at Particle Accelerator Conference (PAC 03), Portland, Oregon, 12-16 May 2003.
- [51] P. H. Garbincius. Tevatron collider operations and plans, 2004.
- [52] The CDF II Collaboration. The CDF II detector, 1996. Technical Design Report.
- [53] A. Affolder et al. CDF central outer tracker. Nucl. Instrum. Methods A, 526:249–299, 2004.
- [54] A. Sill. CDF Run II silicon tracking projects. Nucl. Instrum. Meth., A447:1–8, 2000.
- [55] A. Affolder et al. Intermediate silicon layers detector for the CDF experiment. Nucl. Instrum. Methods A, 453:84–88, 2000.
- [56] C. S. Hill. Initial experience with the CDF layer 00 silicon detector. Nucl. Instrum. Meth., A511:118–120, 2003.
- [57] L. Balka et al. The CDF central electromagnetic calorimeter. Nucl. Instrum. Methods A, 267:272, 1988.
- [58] S. Bertolucci et al. The CDF central and endwall hadron calorimeter. Nucl. Instrum. Methods A, 267:301, 1988.
- [59] M. G. Albrow et al. The CDF plug upgrade electromagnetic calorimeter: Test beam results. Nucl. Instrum. Methods A, 480:524–546, 2002.
- [60] A. Bhatti et al. Determination of the jet energy scale at the Collider Detector at Fermilab. Nucl. Instrum. Methods A, 566:375–412, 2006.
- [61] C. M. Ginsburg. CDF Run 2 muon system. Eur. Phys. J., C33:s1002–s1004, 2004.
- [62] G. Ascoli et al. CDF central muon detector. Nucl. Instrum. Meth., A268:33, 1988.
- [63] T. Dorigo. The muon system upgrade for the CDF II experiment. Nucl. Instrum. Meth., A461:560–562, 2001.
- [64] M. Albrow et al. CDF Run II triggers table and datasets plan. Internal CDF Note 4718 (unpublished), 2001.

- [65] B. Ashmanskas et al. The CDF silicon vertex trigger. Nucl. Instrum. Meth., A518:532–536, 2004.
- [66] G. Gomez-Ceballos et al. Event builder and level 3 at the CDF experiment. Nucl. Instrum. Meth., A518:522–524, 2004.
- [67] M. Feindt. A neural Bayesian estimator for conditional probability densities. <http://www.citebase.org/abstract?id=oai:arXiv.org:physics/0402093>, 2004.
- [68] M. Feindt and U. Kerzel. The NeuroBayes neural network package. Nucl. Instrum. Meth., A559:190–194, 2006.
- [69] V. Blobel and E. Lohrmann. Statistische und numerische Methoden der Datenanalyse. Teubner Stuttgart Leibzig, 1998.
- [70] T. Sjostrand et al. High-energy-physics event generation with PYTHIA 6.1. Comput. Phys. Commun., 135:238–259, 2001.
- [71] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton fragmentation and string dynamics. Phys. Rept., 97:31, 1983.
- [72] T. Sjostrand. Jet fragmentation of nearby partons. Nucl. Phys., B248:469, 1984.
- [73] G. Corcella et al. Herwig 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). J. High Energy Phys., 01:010, 2001.
- [74] D. Amati and G. Veneziano. Preconfinement as a property of perturbative qcd. Phys. Lett., B83:87, 1979.
- [75] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. J. High Energy Phys., 07:001, 2003.
- [76] F. Maltoni and T. Stelzer. Madevent: Automatic event generation with MadGraph. J. High Energy Phys., 02:027, 2003.
- [77] T. Stelzer and W. F. Long. Automatic generation of tree level helicity amplitudes. Comput. Phys. Commun., 81:357–371, 1994.
- [78] R. Brun and Carminati C. GEANT detector description and simulation tool. CERN Program Library Writeup W5013 (unpublished), 1993.
- [79] R. Veenhof. Garfield, a drift chamber simulation program. Prepared for International Conference on Programming and Mathematical Methods for Solving Physical Problems, Dubna, Russia, 14-19 Jun 1993.
- [80] R. Veenhof. Garfield, recent developments. Nucl. Instrum. Meth., A419:726–730, 1998.

- [81] G. Grindhammer, M. Rudowicz, and S. Peters. The fast simulation of electromagnetic and hadronic showers. Nucl. Instrum. Meth., A290:469, 1990.
- [82] E. Gerchtein and M. Paulini. CDF detector simulation framework and performance, 2003. Talk given at 2003 Conference on Computing in High-Energy and Nuclear Physics.
- [83] A. Mukherjee. CTC and VTX tracking. Internal CDF Note 5490 (unpublished), 1995.
- [84] F. D. Snider. Tracking at CDF: Algorithms and experience from Run I and Run II. Nucl. Instrum. Meth., A566:133–141, 2006.
- [85] C. Hays et al. The COT pattern recognition algorithm and offline code. Internal CDF Note 6692 (unpublished), 2004.
- [86] W.-M. Yao et al. Outside-in silicon tracking at CDF. Internal CDF Note 5991 (unpublished), 2002.
- [87] S. Menzemer. Spurrekonstruktion im Silizium-Vertexdetektor des CDFII-Experiments. PhD thesis, Universitaet Karlsruhe, 2003. IEKP-KA/03-04, Internal CDF Note 5968.
- [88] R.E. Kalman. A new approach to linear filtering and prediction problems. Transactions of the ASME. Series D: Journal of Basic Engineering, 82:35–45, 1960.
- [89] Y. Huang et al. Inside-out tracking. Internal CDF Note 6707 (unpublished), 200.
- [90] C. Hays et al. Inside-out tracking at CDF. Nucl. Instrum. Meth., A538:249–254, 2005.
- [91] H. Stadie et al. Vxprim in Run II. Internal CDF Note 6047 (unpublished), 2002.
- [92] H. Stadie et al. The beam position in Run II. Internal CDF Note 6327 (unpublished), 2003.
- [93] S. Rappoccio et al. First look at SecVtx using event primary vertex finder PrimeVtx. Internal CDF Note 6417 (unpublished), 2003.
- [94] M. Feindt et al. Pre-tracking primary vertex z finder. Internal CDF Note 5988 (unpublished), 2002.
- [95] C. Neu. CDF b-tagging: Measuring efficiency and false positive rate, 2006. Presented at TOP 2006: International Workshop on Top Quark Physics, Coimbra, Portugal, 12-15 Jan 2006.
- [96] C. Neu et al. SecVtx optimization studies for 5.3.3 analyses. Internal CDF Note 7578 (unpublished), 2005.

- [97] P. Lujan et al. Efficiency of SecVtx b-tag used for 2005 winter conferences. Internal CDF Note 7343 (unpublished), 2005.
- [98] A. Holloway et al. SecVtx tagging scale factor for 5.3.3. analyses. Internal CDF Note 7445 (unpublished), 2005.
- [99] S. Budd et al. Event detection efficiency for single-top events and MC based background estimate for summer 2006. Internal CDF Note 8286 (unpublished), 2006.
- [100] E. Sexton-Kennedy and P. Murat. Commissioning the CDF offline software. physics/0306112, 2003.
- [101] H. Frisch. The good run list. Internal CDF Note 5613 (unpublished), 2001.
- [102] Good run list v13.
<http://www-cdf.fnal.gov/internal/dqm/goodrun/v13/goodv13.html>.
- [103] V. Martin. High p_T muons, recommended cuts and efficiencies for release 5.3. Internal CDF Note 7031 (unpublished), 2004.
- [104] V. Martin. High- p_T muon ID cuts and efficiencies for use with 5.3.1 data and 5.3.3 MC. Internal CDF Note 7367 (unpublished), 2005.
- [105] C. Hill et al. Electron identification in offline release 5.3. Internal CDF Note 7309 (unpublished), 2005.
- [106] P. Dong et al. Reduction of QCD background in lepton+jets events. Internal CDF Note 8295 (unpublished), 2006.
- [107] E. E. Boos, L. V. Dudko, and V. I Savrin. CMS Note 2000/065, 2000.
- [108] Y. L. Dokshitzer. Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics. (in russian). Sov. Phys. JETP, 46:641–653, 1977.
- [109] V. N. Gribov and L. N. Lipatov. Deep inelastic $e p$ scattering in perturbation theory. Sov. J. Nucl. Phys., 15:438–450, 1972.
- [110] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. Nucl. Phys., B126:298, 1977.
- [111] J. Lueck. Kinematics of electroweak single top quark production. Master’s thesis, Universitaet Karlsruhe, 2006. FERMILAB-MASTERS-2006-01.
- [112] S. Budd et al. Estimation and modeling of non-W background for single-top searches. Internal CDF Note 8489 (unpublished), 2006.
- [113] S. Budd et al. Mistag model used for single-top summer 2006 analyses. Internal CDF Note 8490 (unpublished), 2006.

- [114] S. Budd et al. Data based background estimate for summer 2006 single-top search. Internal CDF Note 8292 (unpublished), 2006.
- [115] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason. NLL resummation of the heavy-quark hadroproduction cross-section. Nucl. Phys., B529:424–450, 1998.
- [116] E. L. Berger and H. Contopanagos. Threshold resummation of the total cross section for heavy quark production in hadronic collisions. Phys. Rev., D57:253–264, 1998.
- [117] N. Kidonakis and R. Vogt. Next-to-next-to-leading order soft-gluon corrections in top quark hadroproduction. Phys. Rev. D, 68:114014, 2003.
- [118] J. M. Campbell and R. K. Ellis. An update on vector boson pair production at hadron colliders. Phys. Rev. D, 60:113006, 1999.
- [119] D. Acosta et al. Measurement of the $t\bar{t}$ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using lepton + jets events with secondary vertex b -tagging. Phys. Rev. D, 71:052003, 2005.
- [120] M. Feindt et al. A neural network b tagger for single-top analyses. Internal CDF Note 7816 (unpublished), 2006.
- [121] S. Baroiant et al. Measurement of electron trigger efficiencies for level1 and level2 8 GeV triggers. Internal CDF Note 6257 (unpublished), 2003.
- [122] M. Buehler et al. Single-top search with neural networks for summer 2006. Internal CDF Note 8335 (unpublished), 2007.
- [123] T. Walter. Suche nach elektroschwacher Erzeugung einzelner Top-Quarks mit dem CDF-Experiment in Proton-Antiproton-Kollisionen bei 1.96 TeV Schwerpunktsenergie. PhD thesis, Universitaet Karlsruhe, 2005. unpublished.
- [124] S. Budd et al. Likelihood function search for single-top production with 1 fb^{-1} . Internal CDF Note 8342 (unpublished), 2006.
- [125] S. Budd, 2007. personal correspondence.
- [126] D. Hirschbuehl. Eine Methode zur Messung von Einzel-Top-Quark Ereignissen mit dem CDF Experiment. Master’s thesis, Universitaet Karlsruhe, 2001. IEKP-KA/2001-27, additionally: Internal CDF Note 5976 ‘A study of single top quark production using event shape and rapidity observables’.
- [127] M. Buehler et al. Data - Monte Carlo comparison for single-top neural network search. Internal CDF Note 8291 (unpublished), 2006.
- [128] A. Abulencia et al. Top quark mass measurement using the template method in the lepton + jets channel at CDF II. Phys. Rev. D, 73:032003, 2006.

- [129] 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.
- [130] CDF collaboration. Search for single top quark production in 955 pb⁻¹ using the matrix element technique. Public CDF note 8588, 2006.
- [131] CDF collaboration. Multivariate likelihood search for single-top-quark production with 1 fb⁻¹. Public CDF note 8585, 2006.
- [132] V. M. Abazov et al. Evidence for production of single top quarks and first direct measurement of $|V_{tb}|$. Phys. Rev. Lett., 98:181802, 2007.
- [133] D0 collaboration. Updated combination results from three single top quark cross section measurements using the blue method. Public D0 note 5396, 2007.
- [134] R. Wagner. Electron identification for Run II: Understanding and using Lshr. Internal CDF Note 6249 (unpublished), 2003.

Danksagung

Mein allererster Dank gilt Herrn Professor Dr. Michael Feindt dafür, dass er mich als Doktorandin angenommen und mir dadurch die Durchführung dieser Arbeit ermöglicht hat. Vielen Dank auch für die stetige Unterstützung und hilfestellende Fragen, die sehr zum Gelingen dieser Arbeit beigetragen haben.

Besonderer Dank gilt Herrn Privatdozenten Dr. Wolfgang Wagner nicht nur für die Übernahme des Korreferats, sondern auch dafür, dass er mich in den letzten Jahren in meiner Arbeit sehr unterstützt hat. Die fachlichen Diskussionen mit ihm haben mich sehr viel weitergebracht.

Ich möchte mich auch bei Herrn Professor Dr. Thomas Müller für die herzliche Aufnahme in die Top-Gruppe bedanken. Ich weiß es sehr zu schätzen, dass er mich stets sehr unterstützt hat.

Ein Dankeschön auch dafür, dass ich im Rahmen meiner Doktorarbeit nicht nur die Möglichkeit hatte, ans Fermilab zu fahren und mich dort an der Datennahme beteiligen konnte, sondern auch dafür, dass ich sowohl an der Herbstschule in Maria Laach als auch an der CERN-Sommerschule in Schweden teilnehmen durfte. Das alles waren einmalige Erfahrungen für mich!

Ganz besonders möchte ich mich bei Dr. Dominic Hirschbühl, Thorsten Chwalek, Jan Lück, Julia Weinelt und Adonis Papaikonomou bedanken, die meine Arbeit vor der Abgabe zur Kontrolle durchgelesen und mich mit ihrer konstruktiven Kritik zu Verbesserungen angeregt haben.

Ein ganz besonderer Dank geht an Jan Lück für das gelebte Motto “geteiltes Leid ist halbes Leid”. Durch unsere Zusammenarbeit arteten selbst die längsten durchgearbeiteten Nächte nie komplett in Stress aus, denn selbst in den größten mitternächtlichen Krisen war immer noch ein Funke Spaß dabei.

Ich möchte mich auch bei der gesamten Top-Gruppe für all die Unterstützung und Anregungen, die ich bekommen habe, bedanken. Ganz besonderer Dank gilt Dr. Dominic Hirschbühl, von dem ich in zahlreichen Diskussionen sehr viel gelernt habe. Des Weiteren danke ich allen Mitgliedern des EKP für die angenehme Arbeitsatmosphäre. Ich hätte mir keine bessere Umgebung für diese Arbeit wünschen können. Ganz besonders gilt das für meine Zimmerkollegen, die meinen manchmal doch etwas ausdrucksstarken Arbeitsstil (meist) klaglos ertragen haben.

Ein ganz herzliches Dankeschön möchte ich auch an das EKP-Admin-Team richten,

ohne deren unermüdlichen Einsatz die Durchführung dieser Arbeit kaum möglich gewesen wäre. In diesem Zusammenhang möchte ich besonders Dr. Yves Kemp danken, der oft meine letzte Rettung in Sachen Computer war. Ich möchte diese Gelegenheit ebenfalls nutzen, mich beim gesamten Phi-T-Team und auch bei Dr. Ulrich Kerzel noch einmal für die Hilfe in technischen Fragen zu bedanken.

Auch bei Frau Weißmann, Frau Schorn und Frau Gerstner möchte ich mich dafür bedanken, dass sie mir durch ihre Unterstützung des Öfteren das Organisatorische um einiges erleichtert haben.

Ich möchte mich auch bei meinen Freunden und insbesondere bei meiner Mannschaft dafür bedanken, dass sie so oft Verständnis dafür hatten, dass ich vor lauter Arbeit in den letzten Jahren etwas weniger Zeit hatte!

Abschließend möchte ich mich bei meinen Eltern, meinen Großeltern, meiner Schwester, meinem Schwager in spe und meinem Freund bedanken. Danke, dass Ihr mir in all den Jahren fest zur Seite gestanden seid. Ohne Eure Unterstützung hätte ich das alles nicht geschafft. Danke auch dafür, dass Ihr immer wieder darauf geachtet habt, dass ich auch mal wieder den Kopf frei bekomme.