Entdeckung elektroschwacher Produktion einzelner Top-Quarks mit dem CDF II Experiment

Jan Lück

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

genehmigte

Dissertation

von

Dipl. Phys. Jan Lück
aus Bretten

Tag der mündlichen Prüfung: 24.07.2009

Referent: Prof. Dr. Th. Müller, Institut für Experimentelle Kernphysik,
Universität Karlsruhe (TH)

Korreferent: Prof. Dr. W. Wagner, Fachbereich C - Experimentelle Teilchenphysik,
Bergische Universität Wuppertal

meinem Vater

Zusammenfassung

Das wohl bekannteste Beispiel einer Materialuntersuchung ist die Lichtmikroskopie, die mittels sichtbaren, an einem zu untersuchenden Objekt gestreuten Lichts zu einer über eine Lupe vergrößerten Abbildung führt. Das erreichbare Auflösungsvermögen ist dabei aber grundsätzlich sowohl durch die geringe Energie der Lichtteilchen des sichtbaren Lichts als auch durch die Leistungsfähigkeit des menschlichen Auges beschränkt. Verwendet man jedoch höher energetische Teilchen, wie beispielsweise schnelle Elektronen bei der Elektronenmikroskopie, und gleichzeitig empfindliche elektronische Nachweisgeräte, so sind deutlich höhere Auflösungen möglich. Man spricht bei diesen hochauflösenden Studien des abzubildenden Objektes auch von Streuexperimenten. Entsprechend diesem Ansatz verbessern die Experimente der Hochenergiephysik seit Jahrzehnten sukzessive die erzielbaren Auflösungen solcher Streuversuche, indem sie sowohl höhere Teilchenenergien als auch immer empfindlichere Nachweismethoden verwenden. Die dadurch erreichte enorme Steigerung erlaubt somit die Substruktur der Materie sichtbar zu machen, um sie im Detail zu untersuchen. Dabei wurden anfangs Untersuchungen von Atomen und ihren Kernen möglich, später ließen sich die Konstituenten der Atomkerne, die Quarks, nachweisen und studieren.

Die höchsten Energien werden von Beschleunigern erreicht, die einen mittels elektrischer Felder beschleunigten Teilchenstrahl durch Magnete auf eine Kreisbahn zwingen. In solchen ringförmigen Beschleunigern werden die Teilchen auf eine Energie beschleunigt, die bei Protonen und Antiprotonen durch die erreichbare Stärke der Ablenkmagnete limitiert ist. Entweder werden die Strahlen dann extrahiert und auf extern angebrachte Materialien geschickt (Target), oder es werden gegenläufige Strahlen zur Kollision gebracht, um maximale Energien im Schwerpunkt zu erzeugen. Mittels hochempfindlicher Teilchendetektoren von der Größe eines mehrstöckigen Hauses, die jeweils rund um einen Treffpunkt beider Teilchenstrahlen angeordnet sind, lassen sich die elektronischen Abbildungen der Streukollision aufzeichnen. Hiermit können anschließend Art und Weise der Kollisionen rekonstruiert und somit die Struktur und Eigenschaften der Materie mit immer höherer Detailgenauigkeit untersucht werden.
Bei allen bis heute durchgeführten Experimenten zeigt sich, dass Quarks keine weiteren beobachtbaren Substrukturen erkennen lassen und es sich bei Quarks daher um elementare Bausteine der Materie handelt.

Der seit über zwei Jahrzehnten höchstenergetische Ringbeschleuniger, das Tevatron, befindet sich am Fermi National Accelerator Laboratory (Fermilab) in der Nähe von Chicago, USA. Hier werden Proton-Antiproton-Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 1.96$ TeV von den beiden Teilchendetektoren CDF und DØ aufgezeichnet. Mit der Entdeckung des Top-Quarks im Jahre 1995 wurde am Tevatron das bislang schwerste bekannte Elementarteilchen nachgewiesen [1, 2]. Es reihte sich nahtlos neben dem Bottom-Quark in die 1973 vorhergesagte dritte Generation von Quarks [3] (Physik-Nobelpreis Makoto Kobayashi und Toshihide Maskawa, 2008) der als Standardmodell der Teilchenphysik bezeichneten theoretischen Beschreibung aller bekannten Elementarteilchen und deren Wechselwirkungen ein.
Die dritte Generation von Leptonen, die mit der Entdeckung des τ-Leptons [4] (Physik-Nobelpreis Martin L. Perl, 1995) bereits 1975 begründet wurde und inzwischen durch die Entdeckung des τ-Neutrinos 2001 [5] komplettiert wurde, deutete auch experimentell schon vor der Entdeckung des Bottom-Quarks 1977 [6] eine dritte Quarkgeneration an. Das Standardmodell traf zahlreiche Vorhersagen, die nachträglich experimentell verifiziert werden konnten und hat sich daher im Allgemeinen als äußerst erfolgreiches theoretisches Konzept bewährt, das in der Lage ist, drei der vier als fundamental angenommenen Kräfte zu beschreiben. Außer der Gravitation, die durch die bislang nicht einbeziehbare Allgemeine Relativitätstheorie [7] (Physik-Nobelpreis Albert Einstein, 1921) beschrieben wird, umfasst es alle übrigen bekannten Wechselwirkungen der elektromagnetischen, starken und schwachen Kraft, wobei die elektromagnetische und die schwache zur elektroschwachen Kraft vereinheitlicht werden konnte [8–10] (Physik-Nobelpreis Sheldon Glashow, Abdus Salam und Steven Weinberg, 1979).

Nachdem in den 80er Jahren des letzten Jahrhunderts die vorhergesagten W und Z Bosonen der schwachen Wechselwirkung am Vorgänger des Tevatrons, dem Super-Proton-Antiproton-Synchrotron (Sp$\bar{\text{p}}$S) der Europäischen Organisation für Kernforschung (CERN) in der Nähe von Genf, Schweiz, entdeckt wurden [11–14] (Physik-Nobelpreis Carlo Rubbia und Simon van der Meer, 1984), stand die intensive Suche nach dem vom Standardmodell vorhergesagten Top-Quark an. Da zum damaligen Zeitpunkt keine konkrete Vorhersage der Top-Quark-Masse vorlag und man daher nicht wusste, dass erst die deutlich höhere Energie des Tevatrons eine Produktion von Top-Quarks ermöglichen würde, wurde auch am Sp$\bar{\text{p}}$S-Speicherring nach Top-Quark-Ereignissen gesucht [15]. Tatsächlich wurden von der UA1-Kollaboration sechs Ereigniskandidaten für die Produktion von Top-Quarks mit einer vergleichsweisen geringen Masse von ca. 40 GeV/c^2 gefunden und als erster Hinweis für die Existenz von Top-Quarks interpretiert [16]. Nachträglich stellte sich dies jedoch durch genauere Überprüfungen und höhere Datenstatistik als Fehleinschätzung heraus.

Hinweise auf eine deutlich höhere Top-Quark-Masse lieferten in den folgenden Jahren unter anderem Präzisionsmessungen am Elektron-Positron-Speicherring LEP am CERN, die schließlich mit der Entdeckung des Top-Quarks 1995 am Tevatron [1, 2] eindrucksvoll bestätigt werden konnten [17]. Die während der ersten Datennahmeperiode des Tevatrons bei einer Energie von $\sqrt{s} = 1.8$ TeV gesammelten Kollisionsereignisse zeigten einen signifikanten Überschuss von Top-Quark-Kandidaten auf. Die

mit verschiedenen Methoden tatsächlich in den beiden Teilchendetektoren CDF bzw. DØ gezählten 49 bzw. 17 Ereignisse übertrafen die abgeschätzte Anzahl von etwa 22 bzw. 4 ausschließlich durch andersartige Prozesse hervorgerufenen Hintergrundereignisse bei weitem. Dabei war zudem die gemessene Ereignisrate in guter Übereinstimmung mit der erwarteten Rate des Produktionsmechanismus von Top-Quark-Paaren durch die starke Kraft. Somit ließ sich die Existenz des Top-Quarks zweifelsfrei wissenschaftlich belegen und eröffnete am Tevatron eine neue Ära der experimentellen Vermessung seiner Eigenschaften. Die damalige, vom CDF-Experiment gemessene Top-Quark-Masse von $m_t = 176 \pm 13$ GeV/c^2 war in sehr guter Übereinstimmung mit den Vorhersagen indirekter Messungen, und sie ist auch heute noch sehr gut verträglich mit dem aktuellen, von beiden Experimenten am Tevatron bestimmten Wert von $m_t = 173,1 \pm 1,3$ GeV/c^2 [18].
Wegen seiner großen Masse spielt das Top-Quark eine besondere Rolle im Standardmodell. Fast analog zur damaligen Vorhersage der Top-Quark-Masse ermöglicht dessen präzise Bestimmung nun eine Abschätzung der Masse des Higgs-Bosons, des letzten noch nicht entdeckten Teilchens des Standardmodells. Somit liefert die Top-Quark-Physik wertvolle Hinweise zur Suche nach dem Higgs-Boson. Eine Konsequenz der hohen Masse des Top-Quarks ist seine extrem kurze Lebensdauer von etwa $0.5 \cdot 10^{-24}$ s. Im Gegensatz zu allen anderen fünf bekannten Quarksorten geht das Top-Quark somit keinen gebundenen Zustand ein und bietet demnach die einzigartige Möglichkeit, die Eigenschaften eines frei zerfallenden Quarks zu studieren.

Ein Ziel der seit 2002 laufende zweite Periode der Datennahme am Tevatron ist unter anderem die präzise Vermessung der Eigenschaften des Top-Quarks. Schon 1986 wurde vorhergesagt, dass das Top-Quark auch über einen zweiten Produktionsmechanismus auftreten sollte, der durch die elektroschwache Kraft vermittelt wird [19]. Hierbei wird das Top-Quark einzeln statt in Paaren erzeugt. Dieser Produktionsprozess konnte bislang noch nicht nachgewiesen werden, da dessen Erzeugungsrate, ausgedrückt durch den Wirkungsquerschnitt, am Tevatron weniger als die Hälfte der Rate der Paarerzeugung entspricht. Der hauptsächliche Grund ist jedoch die deutlich höhere Rate an Untergrundereignissen, die sich nur schwer von den gesuchten Signalen der einzeln erzeugten Top-Quarks unterscheiden lassen. Am Tevatron sind zwei Kanäle für Einzel-Top-Quark-Produktion dominierend: der s-Kanal mit einem theoretisch berechneten Wirkungsquerschnitt von $0,88 \pm 0,11$ pb, und der t-Kanal-Prozess mit $1,98^{+0,28}_{-0,22}$ pb, jeweils bei einer angenommenen Top-Quark-Masse von 175 GeV/c^2. Der dritte mögliche Kanal, die so genannte assoziierte Wt-Produktion, hat am Tevatron einen vernachlässigbar kleinen Wirkungsquerschnitt von $0,26 \pm 0,06$ pb.

Der Wirkungsquerschnitt der Produktion einzelner Top-Quarks ist dabei direkt proportional zum Quadrat des $|V_{tb}|$-Elements der Cabbibo-Kobayashi-Maskawa-Matrix. Letztgenannte beschreibt die elektroschwache Mischung der Quarksorten untereinander und ist eine Konsequenz aus der schon oben erwähnten Vorhersage der dritten Generation von Quarks. Das $|V_{tb}|$-Element, ausschlaggebend für die elektroschwache Umwandlung von Top- in Bottom-Quarks, ist dabei neben der Top-Quark-Masse der einzige nicht theoretisch vorhersagbare freie Parameter des Standardmodells im

Top-Quark-Sektor. Die bisherigen indirekten Messungen von $|V_{tb}|$ basieren darauf, dass genau drei Quarkfamilien existieren, woraus sich ein Wert von $|V_{tb}| \approx 1$ ableiten lässt. Eine signifikante Abweichung nach unten könnte auf eine weitere Generation von Quarks hindeuten, andererseits könnte eine entsprechende Abweichung nach oben auf zusätzliche Top-Quark-Produktionsmechanismen jenseits des Standardmodells, etwa flavour-ändernde neutrale Ströme, hinweisen [20]. Der Versuch des experimentellen Nachweises der elektroschwachen Top-Quark-Produktion und die anschließend mögliche modellunabhängige Bestimmung von $|V_{tb}|$ sind somit äußerst interessante Themen auf dem Gebiet der Teilchenphysik.

Das Institut für Experimentelle Kernphysik an der Universität Karlsruhe arbeitet nun schon seit zehn Jahren innerhalb der CDF-Kollaboration an der Suche nach elektroschwacher Top-Quark-Produktion. Die sukzessive Verbesserung der Suchmethoden, anfangs noch mittels einfacher Zählexperimenten [21, 22] analog zu den verwendeten Methoden bei der Top-Quark-Entdeckung, standen dabei im Vordergrund. Die auch im Rahmen dieser Arbeit betriebene Weiterentwicklung ausgefeilter statistischer Methoden [23,24] führte schließlich zu dem Einsatz neuronaler Netzwerke [25–30]. Bei dieser Analysemethode werden Software-Algorithmen mittels simulierter Kollisionsereignisse auf die Mustererkennung von Signalereignissen optimiert. Das hierbei benutzte NeuroBayes®-Netz [31, 32] lernt mittels einer großen Anzahl von Signal- und Untergrundereignissen, neue, noch nicht bekannte Ereignisse in eine der beiden Klassen einzuordnen. Dabei ist der Ausgabewert des neuronalen Netzes selbst ein Maß für die Trefferwahrscheinlichkeit der Klassifizierung. Hat man ein solch trainiertes Netzwerk ausgiebig mit simulierten als auch mit gemessenen Daten überprüft, ermöglicht dessen Anwendung die leistungsstarke Klassifizierung von aufgezeichneten Kollisionsereignissen. So kann der Anteil von Signalereignissen und die Signifikanz des gemessenen Signals bestimmt werden.
Der Einsatz neuronaler Netzwerke ermöglichte der CDF-Kollaboration kurz nach der DØ -Kollaboration [33, 34], erste Hinweise auf die elektroschwache Produktion einzelner Top-Quarks zu finden [35]. Gleichzeitig wurden die am Institut entwickelten Methoden auch in den Suchen nach dem Higgs-Boson [36], als auch nach Top-Quark-Produktion durch flavour-ändernde neutrale Ströme [37–39] eingesetzt.

In der hier vorgestellten Arbeit wird eine kombinierte Suche nach s- und t-Kanal-Produktion von einzelnen Top-Quarks durchgeführt, in welcher neuronale Netze darauf trainiert werden, beide Kanäle als Signal zu erkennen. Hierbei wird die Annahme gemacht, dass das Verhältnis der s- und t-Kanal-Wirkungsquerschnitte der Standardmodellvorhersage entspricht. Insgesamt werden vier Netze in unterschiedlichen Ereigniskategorien trainiert, die den entsprechenden möglichen Signaturen eines im Detektor zerfallenden einzelnen Top-Quarks entsprechen. Aus den Ausgabewerten dieser 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. Der untersuchte Datensatz entspricht dabei einer integrierten Luminosität von $3,2\ \mathrm{fb}^{-1}$.

Die erwartete Sensitivität und die gemessene Signalsignifikanz der Analysemethode wird mit Hilfe von Hypothesentests bestimmt. Hierbei ermittelt man in Simulationen, in wie vielen hypothetischen Fällen die Analyse durch statistische Fluktuationen mindestens solch einen Wert erreicht, wie er tatsächlich im aufgezeichneten Datensatz gemessen wurde. Dabei wird die Simulation der Analyse ohne Berücksichtigung des zu erwarteten Signalanteils durchgeführt, es werden also ausschließlich die erwarteten Untergrundprozesse miteinbezogen. Für die Bestimmung der Sensitivität verwendet man statt des in den Daten gemessenen Wertes den Mittelwert von Simulationen, die unter zusätzlicher Berücksichtigung des zu erwarteten Signalanteils durchgeführt wurden. Unterschreitet die so ermittelte Wahrscheinlichkeit, die gemessenen Daten ausschließlich durch Untergrundprozesse erklären zu können, einen Wert von $1,35 \times 10^{-3}$ bzw. $0,287 \times 10^{-6}$, so spricht man von einem 3σ-Hinweis bzw. einer 5σ-Beobachtung des gesuchten Prozesses. Letzteres entspricht bei einem bislang unbeobachteten Prozess einer Entdeckung.
Es ergibt sich in der hier vorgestellten Analyse eine Sensitivität entsprechend einer erwarteten Signifikanz von $5,2\sigma$. Dies stellt, zusammen mit einer weiteren CDF-Analyse, die höchste je erreichte Sensitivität einer Analyse zur Suche nach einzelnen Top-Quarks dar. Unter der Annahme von Standardmodellbedingungen liegt das 5σ-Entdeckungspotential dieser Analysen so hoch, dass man mit einer Wahrscheinlichkeit von knapp über 50% erwarten kann, die erste Entdeckung elektroschwacher Produktion einzelner Top-Quarks zu machen. Die in dieser hier vorgestellten Analyse in Kollisionsereignissen gemessene Signalsignifikanz liegt bei $3,5\sigma$. Somit bevorzugen die gemessenen Daten deutlich die Hypothese des Standardmodells mit entsprechender Produktion einzelner Top-Quarks. Jedoch ist das Signal nicht ausgeprägt genug, um von einer wissenschaftlich zweifelsfreien 5σ-Entdeckung zu sprechen, es kann lediglich von einem starken Hinweis die Rede sein. Der gemessene kombinierte Wirkungsquerschnitt, unter der Annahme einer Top-Quark-Masse von von 175 GeV/c^2, beträgt $\sigma_{\text{kombiniert}} = 1,8 \pm 0,6$ pb, was mit der theoretischen Standardmodellvorhersage von $\sigma^{\text{theo}}_{\text{kombiniert}} = 2,9 \pm 0,4$ pb verträglich ist.

Zusätzlich wird in dieser Arbeit eine separate Messung des s- und t-Kanal-Wirkungsquerschnittes durchgeführt, indem in die Maximum-Likelihood-Methode beide Signalparameter unabhängig eingehen und simultan bestimmt werden. Neuronale Netze sind hierfür darauf optimiert worden, Signalereignisse sowohl vom Untergrund als auch untereinander unterscheiden zu können. Die separate Messung ergibt Wirkungsquerschnitte von $\sigma_t = 0,7^{+0,5}_{-0,5}$ pb für den t-Kanal und $\sigma_s = 2,0^{+0,7}_{-0,6}$ pb für den s-Kanal. Das Ergebnis entspricht einer Fluktuation nach unten im Falle des t-Kanals und einer nach oben für den s-Kanal, jeweils in einer Größenordnung von etwa 2 Standardabweichungen.

Abbildung 1 zeigt die erwarteten und gemessenen Signifikanzen der letzten vier auch im Rahmen dieser Arbeit von der Karlsruher CDF-Arbeitsgruppe durchgeführten Analysen als Interpolation der integrierten Luminosität. Die Analyse im Jahr 2006, basierend auf einem Datensatz von $1,0$ fb^{-1}, erwartete im Mittel eine Signalsignifikanz von 2.6σ, konnte jedoch kein Signal etablieren. Nachträglich ließ sich dies hauptsächlich auf unvollständige Simulationen von einem der wichtigsten Unter-

grundprozesse zurückführen. Anhand der mit der Wurzelfunktion der Luminosität extrapolierten erwarteten Signifikanzen lässt sich der Fortschritt in der Weiterentwicklung der Analysemethoden direkt erkennen, beide Analysen aus 2008 liegen jeweils oberhalb ihres jeweiligen Vorgängers. Dass die letzte Analyse mit $3,2\ \mathrm{fb}^{-1}$ wiederum unterhalb ihres Vorgängers liegt, lässt sich mit dem Einfluss systematischer Unsicherheiten erklären. Die Annahme, das sich der Verlauf der erwarteten Signifikanz durch eine Wurzelfunktion beschreiben lässt, stimmt höchstens bei Vernachlässigung störender systematischer Effekte, da diese eben meist nicht mit der Wurzelfunktion skalieren.

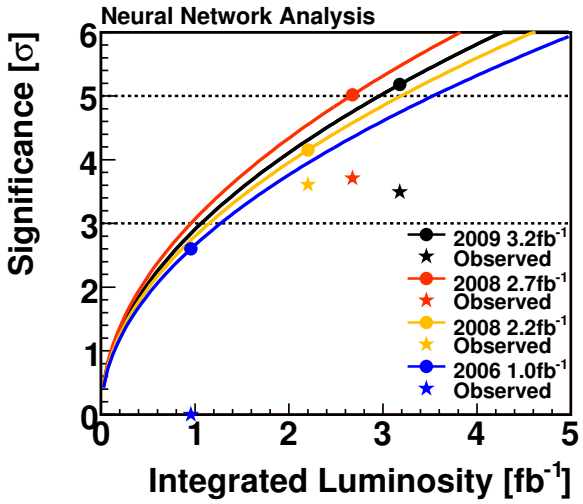


Abbildung 1: Extrapolation der erwarteten Signifikanzen der verschiedenen von der Karlsruher CDF-Arbeitsgruppe durchgeführten Analysen mit Neuronalen Netzwerken und der dazugehörenden in den Daten gemessenen Signalsignifikanzen

Innerhalb der CDF-Kollaboration wurden insgesamt sechs Analysen zur Suche nach kombinierter Produktion einzelner Top-Quarks durchgeführt. Alle Analysen, darunter auch die hier vorgestellte, arbeiten dabei mit dem selben Datensatz, abgesehen von einer Analyse, die für eine leicht anderen Signalsignatur optimiert ist und somit einen unabhängigen Datensatz verwendet, was zu einer Erhöhung der gesamten Signalakzeptanz von etwa 30% führt. Da die ersten fünf Analysen etwa in der Größenordnung von 70% untereinander korreliert sind, wird erwartet, durch eine Kombination eine höhere Sensitivität zu erreichen. Mittels einer auf neuronalen Netzen basierten Super-Diskriminanten werden die einzelnen fünf Analysen miteinander kombiniert. Schließlich wird in einer finalen Messung simultan aus der Super-Diskriminanten und der verbleibenden sechsten unabhängigen Analyse mittels der Maximum-Likelihood-Methode das Signal extrahiert, um das offizielle Ergebnis der CDF-Kollaboration zu erhalten.
Für die erwartete Signifikanz ergibt sich ein Wert von $> 5,9\sigma$. Das gemessene Signal entspricht einer Signifikanz von 5,0 Standardabweichungen, demzufolge die elektroschwache Produktion einzelner Top-Quarks wissenschaftlich schlüssig nachgewiesen ist. Wiederum unter der Annahme einer Top-Quark-Masse von von 175 GeV/c^2 beträgt der gemessene kombinierte Wirkungsquerschnitt $\sigma_{\mathrm{kombiniert}} = 2,3^{+0,6}_{-0,5}$ pb. Dies ist mit der Standardmodellvorhersage verträglich. Mit diesem gemessenen Wert

kann, unter der Annahme von $V_{\mathrm{tb}} \gg V_{\mathrm{ts}}, V_{\mathrm{td}}$, eine Abschätzung des CKM-Matrix-Elements $|V_{tb}|$ vorgenommen werden: $|V_{tb}| = 0,91^{+0,11}_{-0,11}(\mathrm{stat+syst}) \pm 0,07(\mathrm{theo})$. Die Genauigkeit der Messung ist somit noch nicht ausreichend, um eine Aussage darüber zu treffen, ob $|V_{tb}|$ signifikant von eins abweicht.

Eine zeitgleich von der DØ -Kollaboration durchgeführte Kombination dreier Analysen, basierend auf einem Datensatz mit einer integrierten Luminosität von $2,3\ \mathrm{fb}^{-1}$, erreicht ebenfalls eine Signalsignifikanz von 5,0 Standardabweichungen, wobei die erwartete Sensitivität lediglich $4,5\sigma$ beträgt. Der gemessene kombinierte Wirkungsquerschnitt, unter der Annahme einer Top-Quark-Masse von von 170 GeV/c^2, beträgt $\sigma_{\mathrm{kombiniert}} = 3,9 \pm 0,9$ pb, verträglich mit der Standardmodellvorhersage von $\sigma^{\mathrm{theo}}_{\mathrm{kombiniert}} = 3,5 \pm 0,2$ pb. Die DØ -Kollaboration extrahiert aus ihrer Messung einen Wert für $|V_{tb}|$ von $|V_{tb}| = 1,07 \pm 0,12(\mathrm{theo})$.

Jüngst hat die CMS-Kollaboration am zukünftigen Large-Hadron-Collider (LHC) die aktuellen Aussichten auf das zu erwartende Wiederentdeckungspotentials des am LHC dominierenden t-Kanal-Prozesses bei einer Schwerpunktsenergie von $\sqrt{s} =$ 10 TeV mittels Simulationen untersucht [40]. Es wird erwartet, dass eine 5σ-Wiederentdeckung mit einer gesammelten CMS-Datenmenge entsprechend einer integrierten Luminosität von etwa $0.7\ \mathrm{fb}^{-1}$ möglich sein sollte.

Vierzehn Jahre nach seiner Entdeckung wurde nun das Top-Quark ein zweites Mal, jedoch diesmal über die elektroschwache statt der starken Produktion, entdeckt, wiederum zeitgleich von den CDF- und DØ-Kollaborationen am Tevatron [41, 42]. Die Zukunft der präzisen Vermessung der Wirkungsquerschnitte der Produktionsmechanismen einzelner Top-Quarks und des CKM-Matrix-Elements $|V_{tb}|$ liegt jedoch am LHC [43].

Observation of Electroweak Single Top-Quark Production with the CDF II Experiment

Jan Lück

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

genehmigte

Dissertation

von

Dipl. Phys. Jan Lück
aus Bretten

Tag der mündlichen Prüfung: 24.07.2009

Referent: Prof. Dr. Th. Müller, Institut für Experimentelle Kernphysik,
Universität Karlsruhe (TH)

Korreferent: Prof. Dr. W. Wagner, Fachbereich C - Experimentelle Teilchenphysik,
Bergische Universität Wuppertal

Introduction

The standard model of elementary particle physics (SM) predicts, besides the top-quark pair production via the strong interaction, also the electroweak production of single top-quarks [19]. Up to now, the Fermilab Tevatron proton-antiproton-collider is the only place to produce and study top quarks emerging from hadron-hadron-collisions. Top quarks were directly observed in 1995 during the Tevatron Run I at a center-of-mass energy of $\sqrt{s} = 1.8$ TeV simultaneously by the CDF and DØ Collaborations via the strong production of top-quark pairs [1, 2]. Run II of the Tevatron data taking period started 2001 at $\sqrt{s} = 1.96$ TeV after a five year upgrade of the Tevatron accelerator complex and of both experiments. One main component of its physics program is the determination of the properties of the top quark including its electroweak production. Even though Run II is still ongoing, the study of the top quark is already a successful endeavor, confirmed by dozens of publications from both Tevatron experiments. A comprehensive review of top-quark physics can be found in reference [44].
The reasons for searching for single top-quark production are compelling. As the electroweak top-quark production proceeds via a Wtb vertex, it provides the unique opportunity of the direct measurement of the CKM matrix element $|V_{tb}|$, which is expected to be $|V_{tb}| \approx 1$ in the SM. Significant deviations from unity could be an indication of a fourth quark generation, a production mode via flavor-changing neutral currents, and other new phenomena [20], respectively.

There are two dominating electroweak top-quark production modes at the Fermilab Tevatron: the t-channel exchange of a virtual W boson striking a b quark and the s-channel production of a timelike W boson via the fusion of two quarks. In proton-antiproton-collisions the third electroweak production mode, the associated Wt production of an on-shell W boson in conjunction with a top quark has a comparatively negligible small predicted cross section. Therefore, the vast majority of the CDF and DØ single top-quark analyses search for the combined s- and t-channel signal, with the production ratio to be given by the SM.
In Tevatron Run I, several limits on the single top-quark production cross section were set by CDF [45, 46] and DØ [47, 48], whereas in Run II, even stronger limits followed by both collaborations [24, 49, 50]. Furthermore, limits on the non-SM production of single top-quarks via flavor-changing neutral currents could be obtained [38, 51, 52]. The electroweak production of single top-quarks has not yet been observed up to the time of this thesis, although the DØ and shortly thereafter the CDF Collaborations found first evidence [33–35].

The experimental challenge of the search for single top-quark production is the tiny expected signal beneath a large and imprecisely known amount of background processes. The relative fraction of background events is at the order of about ten times higher compared to the top-quark pair production. Consequently, the expected signal amounts to about 5% of the full candidate event sample whose background contribution is only known to a level at the order of 20%. Furthermore, the signal events themselves are expected to be not as distinct from the background as the top-quark pair production since there is only one heavy object present in the event. Thus, experimental methods like simple counting experiments are not sufficiently sensitive and the development of more sophisticated analysis techniques is required to distinguish such small signals from alike and inaccurately known background processes.

Neural networks comply with those requirements. They can be used to distinguish between signal and background processes by combining the information contained in several variables into a powerful discriminant, while each variable has a rather low separation capability. The application of those neural network discriminants to collision data provide a method for the extraction of the signal fraction and its significance.

This thesis presents a neural network search for combined as well as separate s- and t-channel single top-quark production with the CDF II experiment at the Tevatron using 3.2 fb^{-1} of collision data. It is the twelfth thesis dealing with single top-quark production performed within the CDF Collaboration, whereas three have been done in Run I [53–55] and eight in Run II [23, 25, 28, 39, 56–59].

Contents

Introduction **i**

1 Introduction to Top-Quark Production **1**

1.1 Standard Model of Elementary Particle Physics 1

1.1.1 Fermions of the Standard Model 2

1.1.2 Gauge Bosons and their Interactions 2

1.1.3 CKM Matrix and Electroweak Symmetry Breaking 5

1.2 Top-Quark Production within the Standard Model 7

1.2.1 Top-Quark Pair-Production via Strong Interaction 7

1.2.2 Electroweak Single Top-Quark Production 8

2 Experimental Apparatus **13**

2.1 Tevatron Accelerator Complex 15

2.1.1 Proton Source and Preacceleration 16

2.1.2 Main Injector 16

2.1.3 Antiproton Source and Accumulation 17

2.1.4 Recycler 17

2.1.5 Tevatron Collider 18

2.1.6 Tevatron Performance 19

2.2 CDF II Experiment 22

2.2.1 Tracking System 24

2.2.2 Calorimetry 26

2.2.3 Muon System 27

2.2.4 Luminosity Counters 27

2.2.5 Data Acquisition System 28

3 Event Simulation and Reconstruction 31

3.1 Monte Carlo Event Simulation 31

3.1.1 Matrix Element Event Generators 32

3.1.2 Showering and Hadronization Event Generators 33

3.1.3 Detector Simulation 34

3.2 Event Reconstruction 35

3.2.1 Charged Particle Tracking 35

3.2.2 Primary Vertex Reconstruction 36

3.2.3 Charged Lepton Reconstruction 36

3.2.4 Jet Reconstruction 36

3.2.5 Neutrino Reconstruction 38

3.2.6 Secondary Vertex Reconstruction 39

4 Event Modeling 41

4.1 Simulated Signal Event Modeling 41

4.1.1 s-channel Single Top-Quark Production 42

4.1.2 t-channel Single Top-Quark Production 43

4.2 Background Event Modeling 46

4.2.1 Simulation Based Background Event Modeling 46

4.2.2 Data Based Background Event Modeling 48

5 Candidate Event Selection and Yield Estimate 51

5.1 Candidate Event Selection 51

5.1.1 Trigger Requirements 51

5.1.2 Jet Identification and Missing Transverse Energy Requirement 55

5.1.3 Lepton Candidate Identification 56

5.1.4 Event Vetoes 58

5.2 Candidate Event Yield Estimate 61

5.2.1 Simulation Based Event Yield Estimate 61

5.2.2 QCD-Multijet Event Yield Estimate 63

5.2.3 W+Heavy Flavor Event Yield Estimate 64

5.2.4 W+Light Flavor Event Yield Estimate 64

6 Candidate Event Classification 67

6.1 Technique of Neural Network Event Classifier 67

6.1.1 Neural Network Event Classifier 67

6.1.2 Training of Neural Network Event Classifier 69

6.1.3 Training Event Sample Composition 70

6.2 Definition of Input Variables 72

6.2.1 KIT Flavor Separator 73

6.2.2 Top-Quark Reconstruction 74

6.2.3 Input Variables 74

6.3 Training Result and Template Construction 85

6.3.1 Templates for Combined Single Top-Quark Search 85

6.3.2 Templates for Separate s- and t-channel Search 87

7 Analysis 93

7.1 Statistical Method 94

7.1.1 Binned Likelihood Function 94

7.1.2 Hypothesis Testing 97

7.1.3 Cross-Section Measurement Method 99

7.2 Optimization of Event Classification Technique 100

7.3 Validation of Event Classification 103

7.3.1 Coverage of Classification Technique 103

7.3.2 Control-Region Validation of Classifier 107

7.4 Systematic Model Uncertainty 109

7.4.1 Systematic Normalization Rate Uncertainty 110

7.4.2 Systematic Template Shape Uncertainty 116

7.5 Results 122

7.5.1 Significance of Combined Single Top-Quark Search 122

7.5.2 Combined s- and t-channel Single Top-Quark Production Cross-Section Measurement 126

7.5.3 Variables in the High-Output Region of the Combined Search 127

7.5.4 Separate s- and t-channel Single Top-Quark Production Cross-Section Measurement 128

8 First Observation of Single Top-Quark Production 131

8.1 Lepton+Jets Analyses 132

8.1.1 Neural Network Analysis 132

8.1.2 Matrix Element Analysis 133

8.1.3 Likelihood Function Analysis 134

8.1.4 Boosted Decision Tree Analysis 135

8.1.5 Likelihood Function s-channel Analysis 136

8.2 Combination 137

8.2.1 Lepton+Jets Super Discriminant Analysis 137

8.2.2 $\not\!\!E_{\mathrm{T}}$+Jets Analysis 138

8.3 Results of the Combination 139

8.3.1 Application to Observed Candidate Events 139

8.3.2 Observed Single Top-Quark Candidates 144

Summary 147

Chapter 1

Introduction to Top-Quark Production

1.1 Standard Model of Elementary Particle Physics

At present there are four fundamental forces considered to be the generators of any known interaction in nature: the strong nuclear force, the weak nuclear force, the electromagnetic force, and gravitation (table 1.1).

The strong, weak, and electromagnetic forces have been successfully combined

Force	Couples with	Effect	Rel. Strength	Range
strong	color charge	binds quarks and gluons	10^{0}	10^{-15} m
electro-magnetic	electric charge	interaction between el. charged particles	10^{-2}	infinite
weak	weak charge	radioactive decay	10^{-5}	sub-nuclear
gravitation	mass	attraction of masses	10^{-38}	infinite

Table 1.1: The four fundamental forces in nature and some of their characteristics

in a complete theory, the standard model of elementary particles (SM). Gravitation is not yet included in this concept and is described by the Theory of General Relativity [7]. The SM provides a very elegant theoretical framework based on a quantum field theory that explains the nuclear and electromagnetic interactions between particles as resulting from the introduction of local symmetries. It has been very successful in predicting a variety of properties of particles and their interactions, which have been and are still tested by a large number of experiments. Up to now, no strong evidence has been found that falsifies the SM, even though it is known that it has several weaknesses. For example, recent results indicate that a very large amount of matter and energy, called dark matter and dark energy, in the universe is not explained at all [60, 61].

To describe the state of elementary particles, a series of quantum numbers is introduced by the SM, e.g. electric charge, color charge, and weak isospin. Particles are divided into two groups by their spin: the constituents of matter named fermions, half integer spin particles which follow Fermi statistics, and the carriers of the forces named gauge bosons, which have integer spin and follow Bose statistics.

1.1.1 Fermions of the Standard Model

The fermions form the building blocks of all known matter. They consist of quarks and leptons and can be ordered in three generations as shown in table 1.2. Only particles of the first generation form matter, the second and third generation fermions can solely be observed in high energy interactions since they subsequently decay into first generation particles. Each generation involves two types of quarks (called weak partners), a lepton and its corresponding neutrino.
Quarks participate in the strong, weak, and electromagnetic interaction. The different quark-types are called flavors. To date there are six different quark flavors known, up (u) and down (d), charm (c) and strange (s), and top (t) and bottom (b), carrying fractional electric charge of $\frac{2}{3}$ and $-\frac{1}{3}$, respectively. They cannot be observed as free particles, all but the top quark are confined in bound states with integer charge named hadrons, which are classified into baryons (three quark states) and mesons (quark-antiquark states). The top quark immediately decays into a lower generation type quark and therefore cannot form a baryon. Due to the Pauli exclusion principle, it is necessary that quarks carry an additional quantum number, the color charge, which comes in three different types: red, green, and blue.
Leptons interact only by electromagnetic and weak forces and thus don't carry a color quantum number associated with the strong interaction. The electric charged leptons, namely electrons, muons, and taus, have integer charge, their corresponding neutrinos are neutral, hence don't participate in the electromagnetic interaction. Originally, the SM assumes neutrinos to be massless. Due to several measurements, e.g. see references [62–65], it is necessary to extend the standard model in such a way that neutrinos have non-zero masses. The neutrino masses cited in table 1.2 are limits obtained from direct measurements. Cosmological constraints force the sum of all neutrino masses to be in the order of a few eV [66], leading to stricter limits on the masses of ν_μ and ν_τ.
For every fermion one also has to consider the existence of its antiparticle, thus antiquarks and antileptons.

1.1.2 Gauge Bosons and their Interactions

The standard model is a quantum field theory that describes the strong, weak, and electromagnetic force by gauge theories, which are based on the idea that symmetry transformations can be performed both locally and globally. In a gauge theory, the Lagrangian, a mathematical description of a physical system, is invariant under local gauge transformations, i.e. local change of variables. The interactions between

Flavor	Fermion	Symbol	El. Charge [e]	Mass [MeV/c^2]
up	quark	u	$\frac{2}{3}$	$(1.5 - 3.3)$
down	quark	d	$-\frac{1}{3}$	$(3.5 - 6.0)$
electron	lepton	e	-1	0.511
e-neutrino	lepton	ν_e	0	$< 2 \cdot 10^{-6}$
charm	quark	c	$\frac{2}{3}$	$(1.27^{+0.07}_{-0.11}) \cdot 10^3$
strange	quark	s	$-\frac{1}{3}$	104^{+26}_{-34}
muon	lepton	μ	-1	105.658
μ-neutrino	lepton	ν_μ	0	< 0.190
top	quark	t	$\frac{2}{3}$	$(171.2 \pm 2.1) \cdot 10^3$
bottom	quark	b	$-\frac{1}{3}$	$(4.20^{+0.17}_{-0.07}) \cdot 10^3$
tau	lepton	τ	-1	1776.84 ± 0.17
τ-neutrino	lepton	ν_τ	0	< 18.2

Table 1.2: The three generations of fermions and some of their properties [66]

Boson	Force	El. Charge [e]	Mass [GeV/c^2]
gluon g	strong	0	0
photon γ	electromagnetic	0	0
$W^\pm$	weak	± 1	80.425 ± 0.038
Z^0		0	91.188 ± 0.002

Table 1.3: The gauge bosons, their related forces and some of their properties [66]

particles result from the introduction of local symmetries into the Lagrangian. The SM requires a Lagrangian that is invariant under local phase transformations. To guarantee the invariance, additional 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 1.3. In general, gauge bosons are described as massless particles. To obtain massive particles as observed in nature, the introduction of a mass term into the Lagrangian is necessary. It turns out, that such a term is not gauge invariant under local gauge transformations. To solve this problem without abandoning the very successful concept of local gauge invariance, the Higgs mechanism has been introduced, see chapter 1.1.3. According to the Noether theorem [67], each symmetry induces a conserved current corresponding to charge conservation. Therefore, only particles carrying the charge of a certain force can interact via this specific force.

The electromagnetic force is transmitted by 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.

Strong interactions via gluon exchange are described by Quantum Chromo Dynamics (QCD) [68–70]. The massless gluon does not only couple to color charge but carries color itself, leading to gluon self-interactions. This leads to a decrease of the strength of strong interactions at short distances, a feature termed asymptotic freedom, and that the strength increases at large distances, called confinement, which constrains quarks to immediate creation of colorless bound hadrons, referred to as hadronization. Since gluons must carry some charge and some anticharge, all possible color combinations would lead to nine gluons. The color singlet is physically not significant, reducing the number of transmitting gluons to eight.

Weak interactions are mediated via massive gauge bosons, namely the electrically charged $W^{\pm}$ bosons, also referred to as W bosons in the following, and the electrically neutral Z^0 boson, also referred to as Z boson, inducing limited range. The W boson only couples to the left-handed (right-handed) component of the fermion (antifermion) wave function. In general, massive particles have both left- and right-handed components, their wave function ψ consists of a left-handed part ψ_L and right-handed part ψ_R, $\psi = \psi_L + \psi_R$. 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 through charged currents by $W^{\pm}$ boson exchange.

Interactions between elementary particles can be described by Feynman diagrams. Exemplarily, the diagram for electron-electron scattering via the exchange of a virtual photon is shown in figure 1.1. The Feynman diagram gives a visualization of the

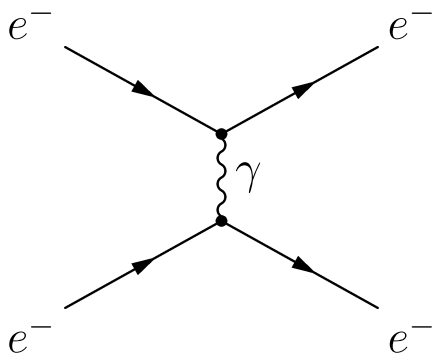


Figure 1.1: A Feynman diagram of electron-electron scattering. At the left-hand side are the two incoming initial state electrons. They interact via a virtual photon (γ) and the scattered final state electrons phase out to the right. The coupling of the electrons to the photon depends on the vertices, shown in the diagram by the dots.

physical process in the momentum space at parton level. Only elementary particles, quarks, leptons, and bosons called partons in this context, are allowed to interact with each other. The Feynman rules provide a prescription of how to translate the diagram into a formula to calculate the transition amplitude $\mathcal{M}$. From this the cross section of the process, a measure of the probability of the interaction occurring, can be derived by integrating over all initial and final states, the phase space.

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. 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. This leads to the situation, that the momentum of the proton $\mathcal{P}_p$ is shared by all three valence quarks, sea-quarks and gluons. The fraction of the momentum $\mathcal{P}_i = x_i \cdot \mathcal{P}_p$, carried by each quark and gluon, is described by the parton distribution function (PDF) $f_{i,p}(x_i, \mu^2)$. It depends on the scale μ, describing the typical energy scale of the considered interaction, which for top-quark production is usually set to the order of the top-quark mass, $\mu = m_t$. Figure 1.2 shows exemplarily the CTEQ5L [71] parton distribution function for the scale $\mu^2 = (175\ \text{GeV})^2$. These PDFs have to be folded with the partonic cross sections to calculate the measurable cross section in $p\bar{p}$ collisions.

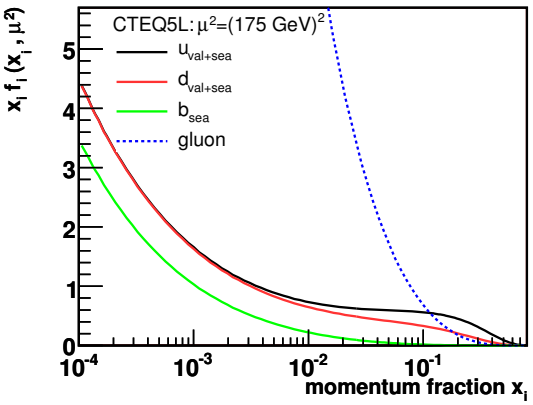


Figure 1.2: The CTEQ5L parton distribution function at $\mu^2 = (175\ \text{GeV})^2$ [71].

1.1.3 CKM Matrix and Electroweak Symmetry Breaking

In weak interactions, flavor change is possible by W boson exchange. Here, the flavor eigenstates of quarks differ from the mass eigenstates. This was experimentally found and is implemented in the theory by flavor-mixing. The transformation from one base into the other is described by the Cabibbo-Kobayashi-Maskawa matrix (CKM matrix) (1.1) [3,72]. 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'):

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

The single elements of this matrix are parameters of the SM which cannot be predicted by theoretical calculations and thus have to be determined experimentally. $V_{q_1 q_2}$ is proportional to the coupling strength of two quarks q_1, q_2 to a W boson. The values [66] for the different elements $V_{q_1 q_2}$ are given by

$$V_{\mathrm{CKM}} = \begin{pmatrix} 0.97419 \pm 0.00022 & 0.2257 \pm 0.0010 & 0.00359 \pm 0.00016 \\ 0.2256 \pm 0.0010 & 0.97334 \pm 0.00023 & 0.0415^{+0.0010}_{-0.0011} \\ 0.00874^{+0.00026}_{-0.00037} & 0.0407 \pm 0.0010 & 0.999133^{+0.000044}_{-0.000043} \end{pmatrix} \tag{1.2}$$

In the late 1960s the weak interaction was unified with the electromagnetic interaction in the electroweak theory by Glashow, Salam, and Weinberg [8–10]. Although these two forces appear very different at low energies, the theory models them as two different aspects of the same 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 [73–76].
By incorporating a scalar quantum field and choosing a favorable gauge, the mechanism introduces a mass term for the bosons of the weak interaction, while keeping the local gauge invariance of the SM Lagrangian. As a result, the Higgs mechanism within the SM predicts two massive, electrically charged $W^{\pm}$ bosons, the massive uncharged Z^0 boson, and the massless photon. There is one remaining degree of freedom of this field which leads to an observable particle, the Higgs boson H, a scalar particle yet undiscovered.
Through an introduced coupling, called Yukawa coupling, of the Higgs field to the fermion field, one finds a mass term for charged leptons being locally gauge invariant. Since the Yukawa coupling is a free parameter, the Yukawa mechanism gives no prediction of the masses of the charged leptons. The couplings of the leptons to the Higgs field are very small and proportional to the lepton masses, so first $H\tau\tau$ leptonic Yukawa coupling might be experimentally observable at future experiments. The origin of non-zero neutrino masses is not yet finally understood.
Similar to the leptonic Yukawa couplings, mass terms of quarks can be found by the coupling of quark fields to the Higgs field. Considering the weak eigenstates of quarks described by the CKM matrix, the Yukawa mechanism is also able to explain the fermion masses in the quark sector. Analog to the leptonic sector, the couplings are proportional to the quark masses, hence the coupling of the H boson to the top quark is dominating. And again, the Yukawa couplings as well as the components of the CKM matrix are free parameters of the SM, making predictions of the fermion masses impossible.

1.2 Top-Quark Production within the Standard Model

The top quark, the heaviest fundamental particle known so far, was discovered in 1995 by the CDF and the DØ Collaborations at the Tevatron [1, 2]. It belongs to the third fermion generation and participates in both strong and electroweak interactions. The top quark is extremely heavy, with a mass about 40 times higher than that of the next heaviest quark, its weak partner of the third generation, the b quark. The recent preliminary Tevatron combination of CDF and DØ results yields a top-quark mass of $m_t = 173.1 \pm 1.3$ GeV/c^2 [18]. Although the top quark decays via the weak force, its predicted lifetime of $\tau \approx 0.5 \cdot 10^{-24}$ s is an order of magnitude smaller than the time scale for hadronization. Thus, no top-flavored hadrons can form and the top quark provides the unique opportunity to study a bare quark which passes all its properties including spin information to its decay products. It decays almost exclusively into a b quark and a W boson ($|V_{tb}| \approx 1$). Decays into quarks of the first and second generation are strongly suppressed by small CKM matrix elements.

1.2.1 Top-Quark Pair-Production via Strong Interaction

In proton-antiproton ($p\bar{p}$) collisions at Fermilab's Tevatron, the dominant top-quark production mode is $t\bar{t}$ pair production via the strong interaction. Even though $t\bar{t}$ pairs can be produced via the electroweak interaction through an exchange of a Z boson or photon, this contribution is negligible compared to the strong QCD cross-section.
At a symmetric $p\bar{p}$ collider, the square of the center-of-mass energy of the interacting partons is defined by $\hat{s} = (x_1\mathcal{P}_p + x_2\mathcal{P}_{\bar{p}})^2 = 4x_1x_2\mathcal{P}_p^2 = x_1x_2s$, where s is the square of the center-of-mass energy of the colliding hadrons and $\mathcal{P}$ the 4-momentum of the participating hadrons. At energies above the kinematic threshold of $t\bar{t}$ production $\hat{s}_{min} = (\mathcal{P}_t + \mathcal{P}_{\bar{t}})^2 = 4m_t^2$ at the Tevatron ($\sqrt{s} = 1.96$ TeV), the momentum fractions of the partons are given by $x_1x_2 \geq \frac{\hat{s}_{min}}{s} = \frac{4m_t^2}{(1.96\ \text{TeV})^2} \approx 0.032 \approx (0.18)^2$. In this region of x_i and above, the incoming partons are mostly valence quarks, as visible in the PDFs (figure 1.2). Therefore, about 85% of the cross section is due to quark-antiquark annihilation (figure 1.3 (a)), only about 15% due to gluon fusion (figure 1.3 (b), (c), (d)).

The corresponding quark-antiquark annihilation cross-section calculation is based on the factorization theorem, i.e. the PDFs $f_{i,a}(x_i, \mu^2)$ of the incoming hadrons $ab = p\bar{p}$ have to be folded with the cross section $\hat{\sigma}_{ij}$ of the initial partons $ij = q\bar{q}$:

$$\sigma(ab \to t\bar{t}) = \sum_{i,j} \int \mathrm{d}x_i \mathrm{d}x_j f_{i,a}(x_i, \mu^2) f_{j,b}(x_j, \mu^2) \cdot \hat{\sigma}_{ij}(ij \to t\bar{t}) \qquad (1.3)$$

The parton-parton cross-section $\hat{\sigma}_{ij}$ can be calculated as a perturbation series in the strong QCD running coupling constant $\alpha_s(\mu^2)$. The differential cross section for

leading-order (LO) quark-antiquark annihilation, which contributes with α_s^2 to the perturbation series, is given by:

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\hat{t}}(q\bar{q} \to t\bar{t}) = \frac{4\pi\alpha_s^2}{9\hat{s}^4} \cdot \left[(m_t^2 - \hat{t})^2 + (m_t^2 - \hat{u})^2 + 2m_t^2\hat{s}\right], \tag{1.4}$$

where $\hat{s} = (\mathcal{P}_q + \mathcal{P}_{\bar{q}})^2$, $\hat{t} = (\mathcal{P}_q - \mathcal{P}_t)^2$ and $\hat{u} = (\mathcal{P}_q + \mathcal{P}_{\bar{t}})^2$ are the Lorentz-invariant Mandelstam variables of the process with $\mathcal{P}_i$ being the corresponding momentum 4-vector of the quark i. Figure 1.3 shows the leading order Feynman diagrams for $t\bar{t}$ production.

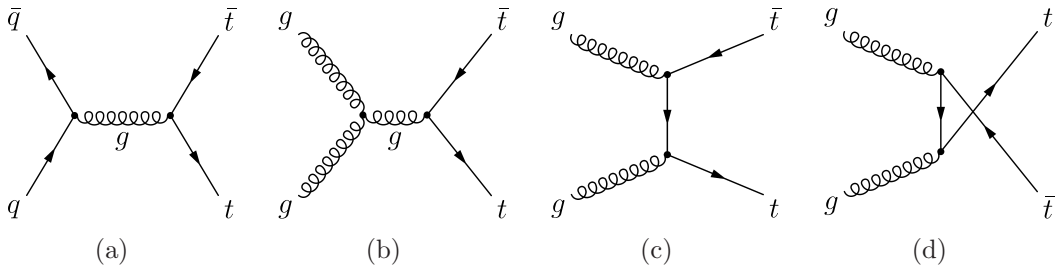


Figure 1.3: Leading-order Feynman diagrams of top-quark pair production: quark-antiquark annihilation (a) and gluon fusion (b), (c), (d).

Calculations at next-to-leading order (NLO), including initial- and final-state gluon bremsstrahlung, gluon splitting, and virtual additions to the LO processes, contribute with α_s^3 to the perturbation series. At the Tevatron, the corrections to the cross section are dominated by initial-state gluon radiation. Further calculations of these soft radiative corrections at higher orders lead to an overall enhancement compared to NLO. The current approximate next-to-next-to-leading-order (NNLO) prediction of the $t\bar{t}$ production cross-section at the Tevatron is $6.9^{+0.5}_{-0.6}$ pb [77], for a top-quark mass of $m_t = 175$ GeV/c^2.

1.2.2 Electroweak Single Top-Quark Production

Besides the strong production of top-quark pairs, the production of single top-quarks via electroweak interaction is predicted by the SM as well. Two electroweak production modes are dominating at the Fermilab Tevatron: the t-channel process (figure 1.4 (a)) and the s-channel process (figure 1.4 (b)). In $p\bar{p}$ collisions the third electroweak production mode, the Wt production (figure 1.4 (c), (d)) has by comparison a negligible small predicted cross section.
Since electroweak top-quark production proceeds via a Wtb vertex, it provides the unique opportunity of the direct measurement of the CKM matrix element $|V_{tb}|$. All three production modes are distinguished by the virtuality $Q^2 = -q^2$, where q is the 4-momentum of the participating W boson. The two dominating processes are labeled by the Mandelstam variable involved in the transition matrix element.

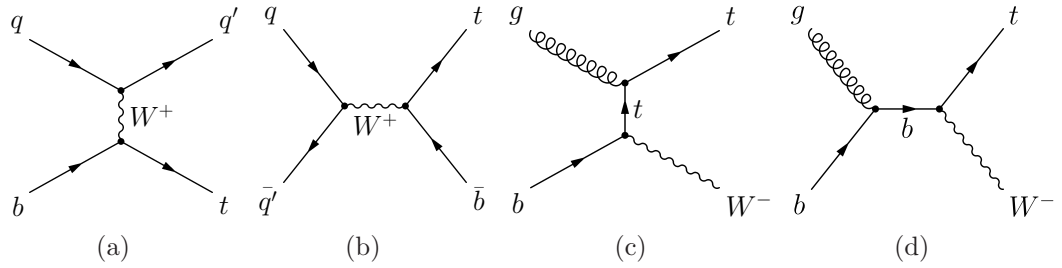


Figure 1.4: Leading-order Feynman diagrams of single top-quark production modes: t-channel (a), s-channel (b), and Wt production (c), (d).

t-channel Production Mode

In t-channel production a virtual spacelike W boson ($q^2 = \hat{t} < 0$) strikes a b quark inside the proton or antiproton. The predicted NLO cross-section at the Tevatron is $\sigma_{t-\text{channel}} = 1.98^{+0.28}_{-0.22}$ pb [78, 79], assuming $m_t = 175$ GeV/c^2. The overall uncertainty includes the choice of 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 uncertainty. The most important NLO correction to the t-channel leading-order process shown in figure 1.4 (a) is the 2→3 process, which is known as W-gluon fusion, where an initial gluon splits into a $b\bar{b}$ pair (figure 1.5 (a)).

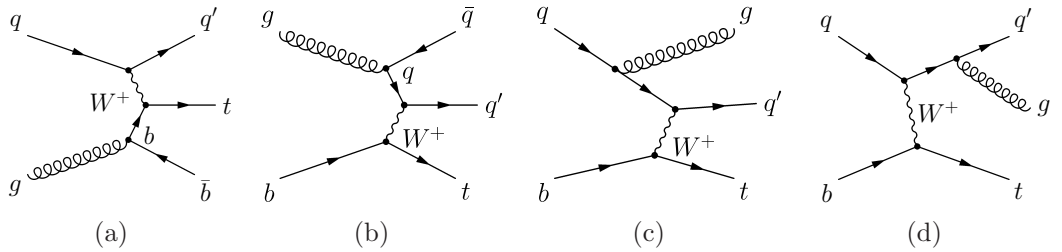


Figure 1.5: Some NLO Feynman diagrams of t-channel single top-quark production: W-gluon fusion (a), initial-state gluon splitting (b), initial-state gluon radiation (c) and final-state gluon radiation (d).

If the b quark is considered massless in the computation of the W-gluon fusion matrix element, the gluon splits into a real $b\bar{b}$ pair with the final state $\bar{b}$ quark (in the following called 2^{nd} b quark in order to distinguish from the b quark coming from the decay of the top quark) being collinear with the incoming gluon. Given that the internal b quark is on-shell, its propagator is infinite and the Feynman diagram becomes singular. As in reality the b quark is not massless, its mass m_b regulates the collinear singularity, which is described by terms of $\ln[(Q^2 + m_t^2)/m_b^2]$, where Q^2 is the virtuality of the W boson. The W-gluon fusion cross-section contains these logarithmic terms of order $\ln^n[(Q^2 + m_t^2)/m_b^2]/n!$ at every order n of the perturbative expansion in the strong coupling due to the collinear emission of gluons from the

internal b quark propagator. This leads to the fact that, since the logarithms are large, the perturbation series does not converge quickly.
However, by introducing a b quark distribution function, one can sum up all collinear logarithms and bypass the convergence difficulty. Now, the b quark appears in the initial state, leading to simplified calculations since they are based on the 2→2 diagram (figure 1.4 (a)). Due to this method the cross section of this process is of order $\ln[(Q^2+m_t^2)/m_b^2]$, the diagram in figure 1.5 (a) also contains these terms, even though they are already summed into the b quark distribution function. So, to avoid double counting, one needs to remove these terms. This subtraction method, carried out in reference [80], achieves the same results within the errors as the phase space slicing method [78], upon which the ZTOP software [79] is based. This analytic form of the NLO cross-section calculation is fully differential, therefore experimental cuts can be implemented.
Very recently, NLO calculations [81] based on the 2→3 process (figure 1.5 (a)) were performed, keeping a finite b mass. Within theoretical uncertainties, the production cross-section calculation, developed in the framework of the MCFM software [82], is consistent with NLO calculations based on the 2→2 process.

s-channel Production Mode

In s-channel production (figure 1.4 (b)), also called W^* production, a timelike W boson ($q^2 = \hat{s} \geq (m_t + m_b)^2$) is produced by the fusion of two quarks. At the Tevatron, the predicted cross section at NLO is $\sigma_{s-\text{channel}} = 0.88^{+0.12}_{-0.11}$ pb [78, 79] for $m_t = 175$ GeV/c^2. 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.
The NLO contributions to the s-channel, shown in figure 1.6, will mostly lead to additional soft light quarks, since the probability of the gluon to split into a heavy quark-antiquark pair is quite low.

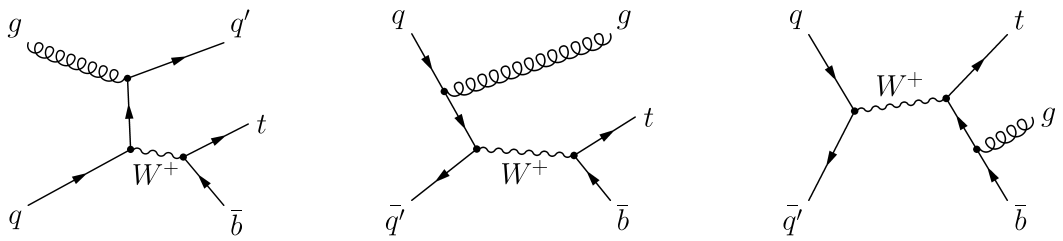


Figure 1.6: Some NLO Feynman diagrams of s-channel single top-quark production: initial-state gluon splitting (a), initial-state gluon radiation (b) and final-state gluon radiation (c).

Wt Production Mode

In the associated production an on-shell (or close to on-shell) W boson ($q^2 = m_W^2$) is produced in conjunction with the top quark (figure 1.4 (c), (d)). The predicted NNNLO cross-section at next-to-leading logarithmic (NLL) accurancy of this production mode at the Tevatron for $m_t = 175$ GeV/c^2 is $\sigma_{Wt} = 0.26 \pm 0.06$ pb [83]. Since this cross section is negligible, the further discussion is restrained on the t- and s-channel single top-quark production modes.

General Properties of Electroweak Single Top-Quark Production

NLO corrections to the s-channel diagrams (figure 1.6 (a)) can lead to the the same initial and final state as the W-gluon fusion diagram (figure 1.5 (a)). In contrast to the t-channel process, where 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 production, since both quarks stem from a W boson. Therefore, interference between s- and t-channel diagrams vanish and their separation is maintained at NLO. Higher order corrections lead to nonzero interference terms. Recent results show, that those terms are heavily color-suppressed and won't prevent a meaningful separation of both channels [81].

The s- and t-channel processes 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 about 7% for t-channel and about 2% for s-channel production [83, 84]. There are also other electroweak production modes than those involving the Wtb vertex. But the channels involving a Wtd or Wts vertex are strongly suppressed, since the involved CKM matrix elements are small. The contribution of such processes to the total single top-quark cross-section is about 1% and thus negligible at the Tevatron. The same circumstance is valid for the top-quark decay. Nearly 100% decay into a b quark and a W boson. Decays into d or s quarks are strongly CKM suppressed.

In s- and t-channel production the top quarks are produced 100% polarized along the direction of the down-type quark (q' in figure 1.4 (a), (b)) in the top-quark rest frame [84–86]. This is due to the fact that the W boson couples only to left-handed fermions. This polarization can be measured since top quarks decay before the typical time scale for hadronization is reached. Hadronization describes the transition from partons to hadrons, and in this stage information such as polarization is typically lost. Consequently, the decay products of the top quark, especially the W boson, carry the polarization information. A good variable to investigate the top-quark polarization experimentally is the angle $\Theta_{l,q}$ between the lepton coming from the W boson decay, and the light quark jet axis in the top-quark rest frame.

As already mentioned, electroweak single top-quark production provides the unique opportunity to measure the value of the CKM matrix element $|V_{tb}|$. The single top-quark production cross-section is directly proportional to $|V_{tb}|^2$, hence its measurement yields a direct extraction of the size of $|V_{tb}|$. As seen in equation 1.2, $|V_{tb}|$ is determined experimentally to be very close to unity, under the SM assumption of a unitarian 3×3 CKM matrix. By only assuming that the branching ratio (BR) for a top quark decaying into a W boson and a b quark is 100%, $\mathrm{BR}(t \to Wb) = 1$, which is equivalent to $|V_{tb}| \gg |V_{td}|$, $|V_{ts}|$, the experimental extraction of the matrix element is feasible without assuming neither the 3×3 structure of the CKM matrix nor its unitarity.

Theoretical new physics models beyond the SM exist which predict a significant deviation of $|V_{tb}|$ from unity [87]. One minimal extension of the SM is the introduction of a t' quark with a mass around the electroweak scale and an electric charge of $\frac{2}{3}$. Through a mixing with the SM top quark, this specific model implements a rescaling of the third row to conserve its unitarity and wouldn't interfere with the first and second row of the CKM matrix.
The scenario of a whole fourth generation of fermions is another interesting extension of the SM allowing for a value of $|V_{tb}|$ considerably deviant from unity. Contrary to general accepted opinion, an unitarian 4×4 quark mixing matrix could provide sizeable mixing between the SM quarks and the fourth generation, assuming no correlations to a lepton mixing matrix [88]. In this model $|V_{td}|$ and $|V_{ts}|$ could differ by a factor of 3 from the SM values as well as $|V_{tb}|$ could be as low as 0.75. The non-observation of those effects in other processes could be explained by huge cancellations due to virtual corrections of heavy b' and t' quarks. The most stringent bounds on the mixing with the fourth generation is expected to be obtained from direct measurements of the third row CKM matrix elements, $|V_{td}|$, $|V_{ts}|$, and $|V_{tb}|$. Preliminary lower limits from the CDF Collaboration on the mass of a b' and a t' quark have been measured to be $m_{b'} \geq 325\ \mathrm{GeV}/c^2$ [89, 90] and $m_{t'} \geq 311\ \mathrm{GeV}/c^2$ [91, 92] at the 95% C.L., respectively.

Chapter 2

Experimental Apparatus

The Collider Detector at Fermilab (CDF) II experiment is located at the Tevatron collider at the Fermi National Accelerator Laboratory (Fermilab), Batavia, Illinois (USA) in the western vicinity of Chicago. The Tevatron, completed in 1983, is the accelerator with the highest center-of-mass energy ($\sqrt{s} = 1.96$ TeV) currently in operation. Here, protons and antiprotons circulate in opposite directions in a ring with a circumference of about 6.3 km (figure 2.1) and are brought to collision inside two general purpose detectors, the CDF II and DØ experiments.

Figure 2.1: Aerial photograph of the Tevatron collider complex: The circular ring is the maintenance road of the Tevatron, the Tevatron collider itself sits in an underground tunnel about 8 m beneath an earthwork next to the road. The oval ring to the left depicts the perimeter of the Main Injector tunnel, which also houses the Recycler. The CDF site in this view is located at the twelve o'clock position direct at the outside of the Tevatron, the DØ site is at the four o'clock position. Wilson Hall, Fermilab's main building, is visible at the eleven o'clock position at the Tevatron.

The Tevatron is not the only high energy hadron collider built today, there are several predecessors and already some successors. The first hadron collider at the energy frontier was realized by the Intersecting Storage Rings (ISR) at the European Organization for Nuclear Research (CERN), Geneva (Switzerland). The proton-proton collider with a diameter of about 300 m initiated first collisions in 1971, running until 1984 with a collision energy up to $\sqrt{s} = 63$ GeV. The ISR also became the first proton-antiproton collider in 1981.
In the same year, also at CERN, the Super Proton Antiproton Synchrotron (Sp$\bar{\text{p}}$S) was commissioned. With a circumference of about 6.9 km it delivered center-of-mass energies of $\sqrt{s} = 546$ GeV and $\sqrt{s} = 630$ GeV in collider runs from 1981 to 1983 and from 1983 to 1984, respectively.
In the meantime, the United States started construction of the Intersecting Storage Accelerator (ISABELLE) at Brookhaven National Laboratory (BNL), Upton, New York (USA) in 1978. Delays due to tests with failing superconducting magnets lead to the cancellation of the accelerator project in 1983. ISABELLE was planned to be a proton-proton collider at $\sqrt{s} = 400$ GeV with circumference of about 3.8 km.
In 1991, eight years after the completion of the Tevatron, the US started the construction of a giant, 87.1 km-long oval collider, the Superconducting Super Collider (SSC) at Waxahachie, Texas (USA). It was designed to accelerate protons to a head-on collision energy of $\sqrt{s} = 40$ TeV. Several reasons lead to the cancellation by the US Congress in 1993, including, amongst others, exploding cost estimates due to problems with the superconducting magnets, the end of the necessity to prove the supremacy of US science with the collapse of the Soviet Union, and the financial competition with the International Space Station ISS. At that time, already 23 km of the SSC tunnel were bored and almost 2 billion dollars were spent.
The only high energy lepton-hadron collider was built at Deutsches Elektronen Synchrotron (German Electron Synchrotron, DESY) in Hamburg (Germany). The Hadron-Elektron-Ring-Anlage (Hadron-Electron Ring Accelerator, HERA) was an asymmetric lepton-proton collider with a circumference of about 6.3 km. From 1992 to 2007, an electron or positron beam of up to 27.6 GeV was collided with a proton beam of up to 920 GeV, leading to collision energies of up to $\sqrt{s} = 319$ GeV.
The Relativistic Heavy Ion Collider (RHIC) at BNL is, besides the Tevatron, the only hadron collider still in operation. RHIC was completed in 2000 in the 3.8 km-long ring tunnel of the canceled ISABELLE collider. Its unique capability is to accelerate protons and heavy ions like gold nuclei and bring them to head-on collision at an energy of $\sqrt{s} = 500$ GeV (protons) and $\sqrt{s} = 200$ GeV (gold nuclei), respectively.
At present, the completion of the Large Hadron Collider (LHC) at CERN is in its final stage. By the end of 2009, proton-proton collisions are scheduled inside the re-used 26.7 km-long circular tunnel of the Large Electron-Positron Collider (LEP, 1989-2000) at a initial center-of-mass energy of 7 and 10 TeV, respectively. The proton-proton design energy of the LHC is $\sqrt{s} = 14$ TeV, whereas it is not yet clear whether the superconducting magnets can be steadily operated at that high energy. Further operation modes will allow for heavy ion, mainly lead nuclei collisions with an energy up to $\sqrt{s} = 1150$ TeV.

2.1 Tevatron Accelerator Complex

The first Tevatron collisions at a center-of-mass energy of $\sqrt{s} = 1.8$ TeV were initiated in 1985 . The collision data delivered from 1988 to 1996 in the so called Run I phase amount to an integrated luminosity of about 130 pb^{-1}, leading, amongst other interesting results, to the discovery of the top quark in 1995. Starting in 1996, the Tevatron accelerator complex was upgraded to increase the center-of-mass energy and the instantaneous luminosity, a measure for the collision rate. Main parts of the CDF and DØ experiments were upgraded as well. The Run II phase with collisions at $\sqrt{s} = 1.96$ TeV started at the end of 2001 and is scheduled at least until 2010. The Tevatron derives its name from its original design beam energy of 1 TeV and its accelerator type, a synchrotron. Even though the Run I and the actual Run II beam energy is 900 and 980 GeV respectively, the Tevatron has attained a beam energy as high as 1.012 TeV during accelerator studies. To reach such high beam energies, a complex chain system of succeeding accelerators gradually increasing the energy of protons and antiprotons is realized at Fermilab (figure 2.2).

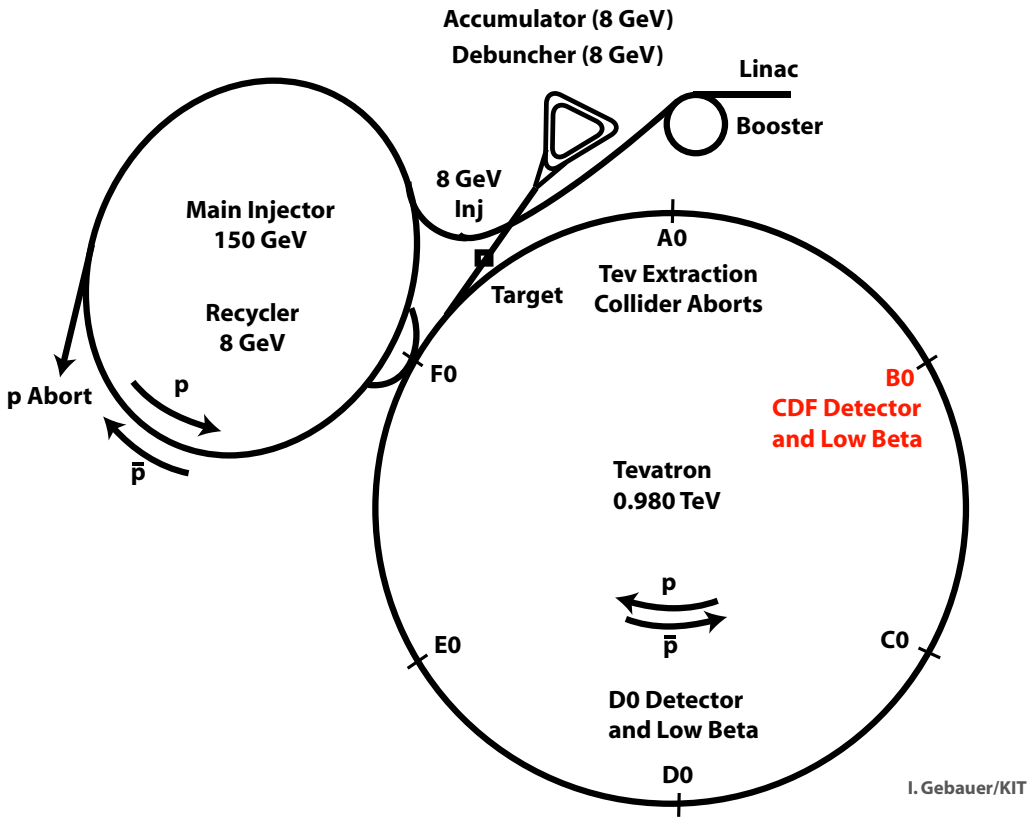


Figure 2.2: Sketch of the Tevatron accelerator complex

2.1.1 Proton Source and Preacceleration

The first step in the accelerator chain is the Cockroft-Walton electrostatic preaccelerator and its Proton Source, which houses a hydrogen ion source in an electrically charged enclosure (−750 kV). A cavity inside the source ionizes very pure hydrogen gas into electrons and H^+ ions through the generation of a dense plasma. The positive ions strike a cathode surface made of vapor coated cesium, which has a low work function and thus looses electrons easily. Occasionally, the H^+ ions convert into H^- ions by absorbing two electrons from the cathode surface and are subsequently accelerated to an energy of 750 keV by the difference of the electrostatic potential between the charged enclosure and the grounded wall. Those preaccelerated H^- ions are passed through a transfer line of focusing magnets to the Linear Accelerator (Linac).
The Linac consists of two stages of radio frequency (RF) resonators, which produce an oscillating electromagnetic field inside drift tube and side-coupled cavities respectively, with an overall length of about 130 m. The first low energy section consists of five drift tube cavities made of copper, which accelerates the negative hydrogen ions to an energy of about 116 MeV within 75 m. The second high energy Linac section consists of seven RF stations, which feed their signal into a series of side coupled cavity modules. Each module is made of alternating accelerating and coupling cell cavities, whose design is optimized to allow for maximum concentrated accelerating fields. The individual drift tubes with their intermediate gaps and the alternating cell cavities are constructed such, that the ions are only exposed to the accelerating part of the alternating-current RF field, and shielded by the drift tube and coupling cells from the decelerating part, respectively. During the linear RF acceleration to an energy of 400 MeV the ions are collected into bunches, leading to a H^- beam with a characteristic RF bunch structure. At the end of the Linac the negative hydrogen ions are sent along a transfer line to the first synchrotron in the acceleration chain, the Booster.
The Booster is a circular proton accelerator with a diameter of about 150 m, working at the same duty cycle as the Linac, which is 15 Hz. Magnets bend the circulating protons and the injected H^- bunches from the Linac through carbon foils, where the electrons are stripped off. A similar set of magnets guide the resulting protons back into the ring while removing any leftover H^- ions. The Booster is capable of storing several bunches of protons at once in its ring to increase the beam intensity. After about 33 ms and some 10.000 revolutions of acceleration by multiple RF cavities located around the ring, the proton beam leaves the Booster with an energy of 8 GeV.

2.1.2 Main Injector

Protons are then transferred to the Main Injector, a large oval synchrotron with a mean diameter of about 1 km. The Main Injector has dozens of accelerating and focusing conventional water-cooled dipole and quadrupole electromagnets, allowing for an acceleration cycle time to full energy in about 2 s. It is operated in several

functional modes, the two most important ones being the consecutive acceleration of protons to 120 GeV every cycle time to be sent to the Antiproton Source and the acceleration of protons and antiprotons to 150 GeV for injection into the Tevatron, respectively. In the latter mode, seven (six) proton (antiproton) bunches are coalesced together into one RF bunch for the injection into the Tevatron.

2.1.3 Antiproton Source and Accumulation

To produce the antiprotons, the Main Injector receives two proton bunches from the Booster, which are merged together during the acceleration to 120 GeV. The resulting bunch is extracted and sent to the Antiproton Source where it hits a stack of nickel targets and copper cooling disks. A lithium lens directly behind the target focuses the spray of all sorts of secondary particles, whereas a subsequent charge-mass spectrometer made of magnets collects antiprotons with energies of 8 GeV. The production efficiency is about $15 \cdot 10^{-6}$ antiprotons per proton hitting the Antiproton Source. These antiprotons are directed through a transfer line into the Debuncher.
The Debuncher is a rounded triangular-shaped synchrotron with a mean diameter of about 180 m. The Debuncher's primary purpose is to efficiently reduce the high momentum spread of the antiprotons, using a RF manipulation called bunch rotation. Thereby, the antiprotons loose their RF bunch structure resulting in a roughly uniform momentum. Additionally, the Debuncher maintains a stable antiproton beam at a constant energy of 8 GeV. Beam stabilization is achieved through beam-cooling systems, the so-called stochastic cooling [93, 94]. The antiprotons remain in the Debuncher for about 2 s until the next bunch of protons is sent to the Antiproton Source, at which point the antiprotons are sent to the Accumulator.
The Accumulator is also a synchrotron, which is housed in the same tunnel as the Debuncher. It is a medium-term antiproton storage ring, where the antiprotons are stacked at an energy of 8 GeV and further cooled down, using both RF and stochastical cooling systems. Extraction from the Accumulator requires the antiprotons to be collected into RF bunches again, which is done by adiabatic RF activation. When the antiproton stack in the Accumulator reaches its maximum optimal capacity, the bunches are then transferred back to the Main Injector and injected into the Recycler.

2.1.4 Recycler

The Recycler is a ring of permanent magnets along the ceiling of the Main Injector tunnel. It is a storage ring providing a long-term stash for antiprotons. Its proposed purpose was to recycle antiprotons left over from the previous Tevatron store and merging them together with those from the Antiproton Source. Those plans have been abandoned in the early Run II phase due to the tremendous effort needed to eventually achieve not enough benefit. The Recycler now accepts transfers only from the Antiproton Source and reduce the antiproton's momentum spread further than

the Accumulator is capable to. 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 [95]. The antiproton beam is overlaid with a 4.3 MeV electron beam traveling about 20 m along the same path in the Recycler [96]. Transverse momentum from the antiprotons is passed to the much lighter electrons by Coulomb scattering, causing the antiprotons to loose momentum spread, resulting in more compact and brighter bunches. Making use of some electrically powered energy correction magnets, the Recycler keeps the antiprotons at 8 GeV until the Tevatron is ready for injection.

2.1.5 Tevatron Collider

The Tevatron is the largest of the Fermilab accelerators, a circular synchrotron with a diameter of about 2 km, corresponding to a circumference of about 6.3 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 a temperature of about 4 K cooled by liquid helium. Therefore, heavily increased magnetic fields up to 4.2 T are achieved without the dissipation of electrical power through excessive resistive heating from high currents. The Tevatron collider is primarily a storage ring, circulating protons and antiprotons in opposite directions in the same beampipe. 774 dipole magnets, 240 quadrupole magnets, and multiple correction magnets keep each beam inside the beampipe, while both horizontally and vertically oriented electrostatic separators at strategic locations around the Tevatron ring allow the formation of a helical beam orbit, called helix. The protons and antiprotons spiral around each other as they revolve in opposite directions, preventing the beams from unintentional collisions and hence proton and antiproton losses. Figure 2.3 shows a photograph of the Tevatron tunnel.

Figure 2.3: Photograph of the Tevatron tunnel

The injection of the beams starts with transferring 36 consecutive bunches of 150 GeV protons from the Main Injector into the Tevatron. Afterwards, four antiproton bunches at a time are extracted from the Recycler, accelerated to 150 GeV in the Main Injector, and subsequently sent into the Tevatron. This process is repeated 9 times to give a total of 36 antiproton bunches. Both beams are simultaneously accelerated to an energy of 980 GeV within about 90 s by eight accelerating RF cavities, leading to a final bunch revolution time of approximately 21 μs. Already during filling, both the proton and the antiproton beam is split into three bunch trains, each containing 12 bunches with intermediate separation gaps of 0.396 μs. The abort gaps, empty spaces between the individual bunch trains with a size of about 2.6 μs, allow very fast kicker magnets to ramp up. For safe beam removal at the end of a Tevatron store or in case of a Tevatron malfunction, those kicker magnets dump the beams by guiding them into the Collider Aborts made of graphite.
To obtain colliding beams of maximum quality, stainless steal collimators are moved close to the beams for about 10 min, scraping away the beam halos. Additionally, clean abort gaps need to be maintained to have a safe and proper environment for collisions. The abort gaps are gradually filled with a direct current (DC) component of the beams through diffusion of particles out of the RF bunches. 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.
The Tevatron proton and antiproton beams are brought to collision at two intersection points: B0, the location of the CDF II experiment, and D0 where the DØ detector is located. To initiate the collisions at the center of the detectors, both beams of 980 GeV each are focused to a minimal transverse size, a procedure called low β^* squeeze which takes about 2 min. At the same time, the injection helix is changed to the collision mode, whereby special electrostatic separators next to the interaction points prevent an overlap of the beams. Upon completion of the squeeze, the polarity of the separators are switched, leading to symmetric proton-antiproton head-on collisions with a center-of-mass energy of $\sqrt{s} = 1.96$ TeV.

2.1.6 Tevatron Performance

The performance of the Tevatron is characterized by the instantaneous luminosity $\mathcal{L}$, a measure of the rate of the proton-antiproton collision, given by:

$$\mathcal{L} = n \cdot f \cdot \frac{N_p N_{\bar{p}}}{2\pi(\sigma_p + \sigma_{\bar{p}})} \cdot F\left(\frac{\sigma_l}{\beta^*}\right), \tag{2.1}$$

where n is the number of bunches, f is the revolution frequency, and N_p and $N_{\bar{p}}$ are the average number of protons and antiprotons in each bunch. The denominator contains σ_p and $\sigma_{\bar{p}}$, which are the spacial widths of the proton and antiproton bunches at the interaction point. F is a form factor depending, amongst others, upon the bunch length σ_l and the value of the β^* function at the interaction point. $\mathcal{L}$ is measured in units of $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, typifying a particle flux.

At the end of Run I, the typical initial instantaneous luminosity at the beginning of each store was $\mathcal{L} = 25 \cdot 10^{30}$ cm^{-2}s^{-1} [97]. At the beginning of Run II, $\mathcal{L}$ did not meet the baseline goal and did not even reach the Run I value, partially because the new Main Injector was not well understood. As knowledge about the accelerators grew, the situation improved. Figure 2.4 shows the Run II initial instantaneous luminosity at the beginning of each store as a function of the consecutive store numbers.

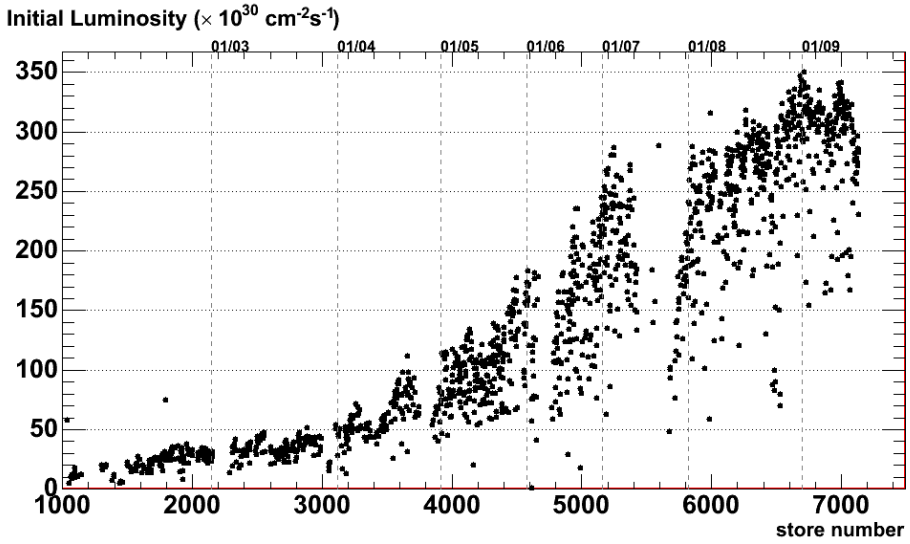


Figure 2.4: Initial instantaneous luminosity since the start of Run II

By increasing the number of bunches n and the revolution frequency f through the reduction of the bunch spacing from 3.5 μs to 0.396 μs, by enlarging the number of particles N_p and $N_{\bar{p}}$ per bunch, and by further progress, the Tevatron is currently achieving an instantaneous luminosity more than a factor 10 better compared to Run I. Recent accelerator upgrades during the Run II operation lead to initial luminosities up to $350 \cdot 10^{30}$ cm^{-2}s^{-1}, exceeding the design goal of $270 \cdot 10^{30}$ cm^{-2}s^{-1} [97] in Winter 2006/2007. A main reason is the increase of the total numbers of antiprotons in the beam. Amongst others, this was made possible by the high stacking rate of the antiproton production, which changed from $6 \cdot 10^{10}$ $\bar{p}/h$ up to $30 \cdot 10^{10}$ $\bar{p}/h$. Such high rates became accessible by stashing the antiprotons in the Recycler, allowing for optimal stacking conditions in the Accumulator. The antiproton beam quality was simultaneously improved by reducing $\sigma_{\bar{p}}$ through electron cooling in the Recycler. Additionally, the β^* function has been decreased from 0.35 m down to 0.28 m since 2005, which immediately increased the instantaneous luminosity.

Integrating the instantaneous luminosity over time gives the integrated luminosity, $\mathcal{L}_{\text{int}} = \int \mathcal{L} \mathrm{d}t$, which corresponds to 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 given in cm^2 is preferably quoted in picobarn (pb), where 1 pb = 10^{-36} cm^2. For a particular process with a given

cross section, the number of events N in a given amount of collected data can be calculated by $N = \sigma \cdot \mathcal{L}_{\text{int}}$. Figure 2.5 shows the Run II integrated luminosity of each store over the consecutive store numbers.

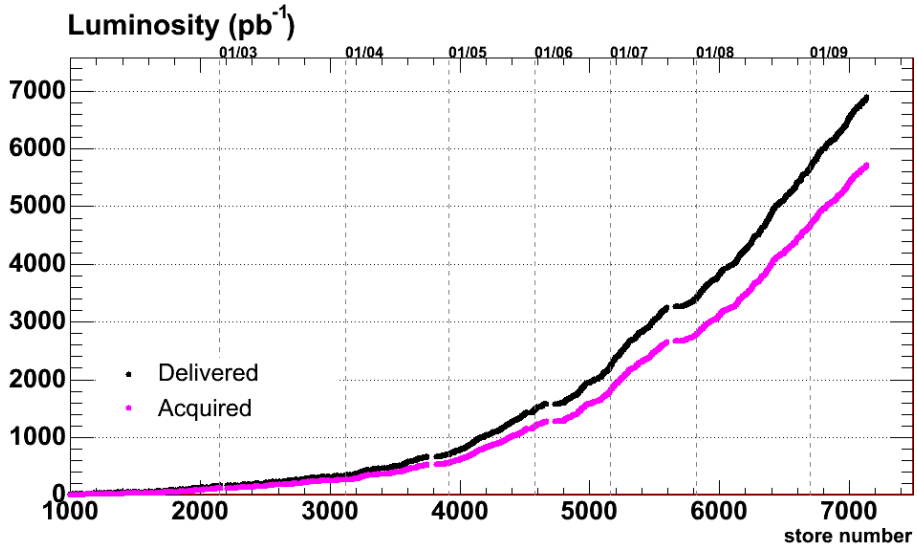


Figure 2.5: Delivered (upper curve) and recorded to tape (lower curve) integrated luminosity at the CDF interaction point since the start of Run II

The Fermilab Accelerator Division further optimized several procedures to maximize the valuable integrated luminosity per week delivered to both experiments. Since the number of protons and antiprotons diminishes during ongoing collisions, the luminosity decreases exponentially during a store. By modifying the helix tune of both beams, unfavorable antiproton losses could be significantly decreased, leading to a reduction of the exponential decline of the mid-store instantaneous luminosity. Additionally, the shortage of the store duration, as well as the minimization of dead time between two consecutive stores maximized the integrated luminosity per week up to 70 pb^{-1}.
Up to the summer shutdown started in June 2009, the Tevatron delivered about 6.9 fb^{-1} whereof CDF managed to record approximately 5.7 fb^{-1}, leading to an average data taking efficiency of about 83%. The Tevatron Run II baseline goal of delivering an integrated luminosity of 4.4 fb^{-1} until August 2009 has been exceeded, whereas the design goal of 8.5 fb^{-1} is unobtainable [97]. The collision data used in the analysis described in this thesis were taken from February 2002 (store 955) to August 2008 (store 6378) and correspond to an integrated luminosity of 3.2 fb^{-1}. It is planned to continue the Tevatron collider operation at least until the end of the fiscal year in September 2010.

2.2 CDF II Experiment

The Collider Detector at Fermilab (CDF), located at the Tevatron intersection point B0, is an azimuthally and forward-backward symmetric general purpose solenoid detector [98]. It combines precision charged particle tracking with fast projective calorimetry and fine grained muon detection to record as much information as possible. Some components of CDF and the data acquisition system have been upgraded for the Tevatron Run II to deal with the increased instantaneous luminosity. Therefore, the upgraded experiment is referred to as CDF II experiment. Shortly before the start of Run II in 2001, CDF was moved from the assembly hall, where the upgrades were performed, to the collision hall, where it is installed into the Tevatron beam pipe around the interaction point. Figure 2.6 shows a photograph of the upgraded CDF II experiment on its way to the collision hall. The assembly hall and the collision hall, where CDF is permanently located during the whole Run II phase, are about 8 m below the surface. The central part of the detector weighs about 5000 tons, arranged in a cubical volume with overall dimensions of about 12 m $\times$ 12 m $\times$ 12 m. It is operated and maintained by the CDF Collaboration, an international association of several hundreds of physicists from more than 50 universities and research laboratories. CDF began initial studies in 1985 and started collecting data in 1988, hence it is the longest lasting physics experiment ever.

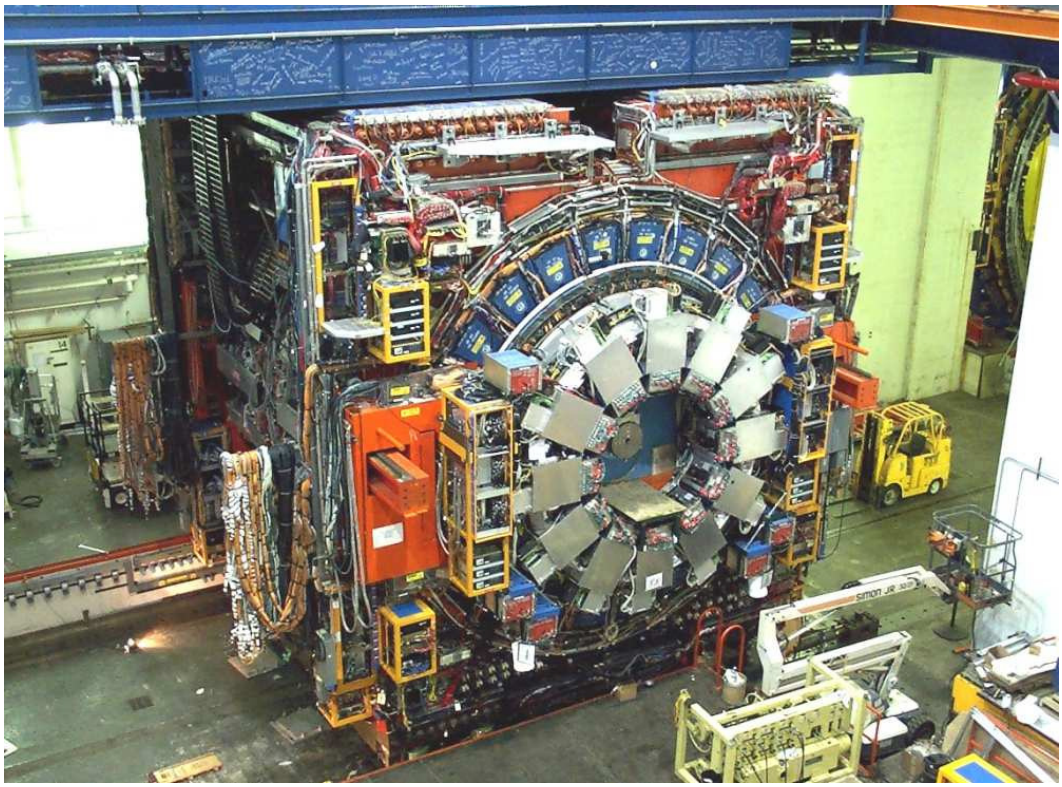


Figure 2.6: The CDF II experiment in the assembly hall on its way to its permanent position in the collision hall, which is partly visible at the outer right side of the photograph

The CDF II experiment is composed of several sub-detectors to identify and measure the emerging particles of interesting collision events. Figure 2.7 shows an isometric view of the CDF II experiment and its main sub-detectors. With Tevatron collisions occurring at each bunch crossing every 0.396 μs, the rate is far too high to allow for the storage of all data. Since the overwhelming dominant fraction of the events are ordinary and unaspiring interactions, a trigger system filters out the interesting ones.

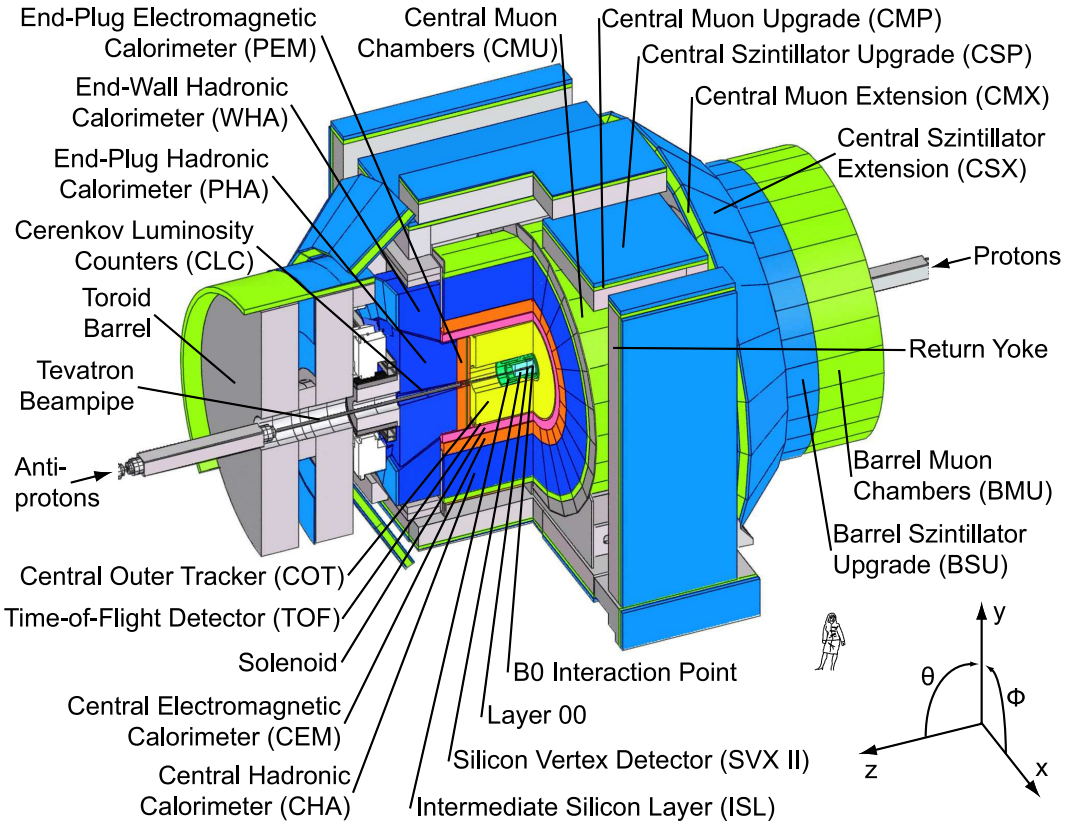


Figure 2.7: Isometric view of the main components of the CDF II experiment and the cylindrical coordinate system

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 2.7. 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 $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$, 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 $E_{\mathrm{T}} = E \cdot \sin\theta$ and $p_{\mathrm{T}} = p \cdot \sin\theta$, respectively.

2.2.1 Tracking System

The CDF tracking system, installed close to the beam pipe, is surrounded by a superconducting solenoid [99] made of an aluminum-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, bending the tracks of charged particles to helices. Hence, precise measurement of particle momenta is possible. The tracking system, schematically displayed in figure 2.8 (a), consists of two main parts: a barrel shaped silicon microstrip system and an open-cell drift chamber surrounding the silicon system. The silicon detector comprises three sub-detectors, namely the so-called Layer 00 [100], the Silicon Vertex Detector (SVX II) [101], and the Intermediate Silicon Layers (ISL) [102], all shown in a frontal view in figure 2.8 (b).

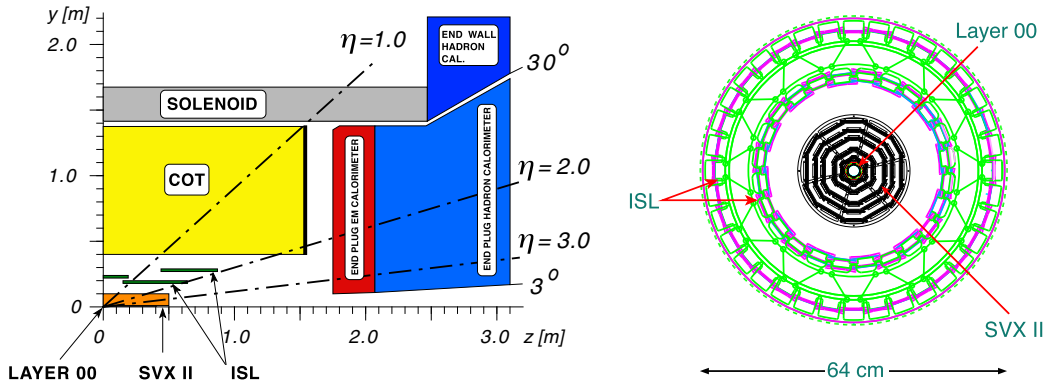


Figure 2.8: Elevation view of the CDF tracking system and the end-plug calorimetry (a) and frontal view of the silicon microstrip system (b)

Layer 00 is a radiation hard, single-sided silicon microstrip detector, directly glued on the Tevatron 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 $|\eta| \leq 4.0$. 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 are rotated by 90° 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 remaining two layers named small angle stereo (SAS) layers, are twisted by 1.2° with respect to the axially aligned strips.

The 1.9 m long ISL consists of double-sided SAS sensors. 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 is about 40 μm including 30 μm contribution from the beam width. The z_0 resolution is about 70 μm. Figure 2.9 shows a photograph of the silicon tracker system.

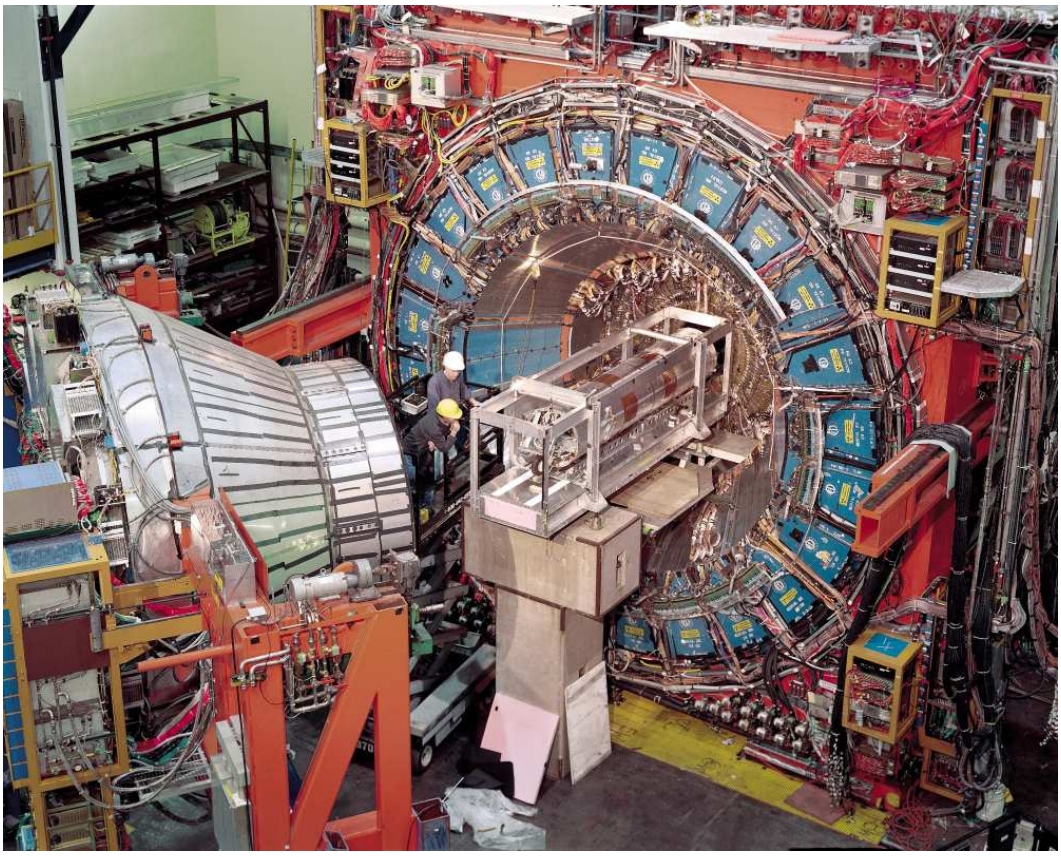


Figure 2.9: The silicon microstrip detectors before the installation into the heart of CDF

The silicon microstrip system is surrounded by the Central Outer Tracker (COT) [103], 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 with $|\eta| \leq 1$. The drift chamber provides 96 measurement layers, organized into eight alternating axial and $\pm 2^\circ$ stereo angle superlayers. The hit position resolution is 140 μm, the momentum resolution $\sigma(p_\mathrm{T})/p_\mathrm{T}^2 = 0.0015$ (GeV/c)$^{-1}$. Furthermore, the COT provides dE/dx information for the tracks.
A Time-of-Flight Detector (TOF) [104], based on plastic scintillators and fine-mesh photomultiplier tubes (PMT), is installed in a few centimeters clearance just outside the COT. The TOF resolution is $\approx$ 100 ps and it provides at least two standard deviation separation between $K^\pm$ and $\pi^\pm$ for momenta $p < 1.6$ GeV/c.

2.2.2 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) [105] and Central Hadron (CHA) [106] Calorimeters, covering the η range up to $|\eta| \leq 1.1$ and $|\eta| \leq 0.9$, respectively, and the End-Wall Hadron Calorimeter (WHA) [106] and the End-Plug Electromagnetic (PEM) and Hadron (PHA) Calorimeters [107, 108], 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 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 λ_I	$50\%/\sqrt{E} \oplus 3\%$
WHA	$0.9 \leq \|\eta\| \leq 1.3$	4.7 λ_I	$75\%/\sqrt{E} \oplus 4\%$
PHA	$1.3 \leq \|\eta\| \leq 3.64$	6.8 λ_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 and hadronic interaction lengths λ_I, respectively. The depths and energy resolutions are taken from reference [109].

The CDF electromagnetic calorimetry includes some additional smaller sub-detectors to further improve particle identification. The Central Pre-Radiate (CPR) and End-Plug Pre-Radiate (PPR) Chambers are the innermost layers of scintillator tiles of the CEM and PEM calorimeter, respectively. Their pre-shower measurements enhance the discrimination between electrons and pions and improve the photon identification. On average, an electromagnetic shower reaches its maximum about six radiation lengths into the calorimeter. The Central Electromagnetic Shower Maximum (CES) as well as the End-Plug Electromagnetic Shower Maximum (PES) Detectors are inserted at this point into the CEM and PEM systems. They are more finely sectioned than the surrounding layers to allow for better position resolution of the calorimeter clusters. CES is a series of strip and wire chambers, while PES is a series of scintillator strips. Both systems improve the cluster position resolution and make it easier to match clusters with tracks from the tracking detectors.

2.2.3 Muon System

The muon system [110] is located outside the calorimetry to ensure that all particles except for muons and neutrinos are already absorbed. Four systems of proportional chambers, filled with a 50 : 50 ratio of argon and ethane (bubbled through ethanol), and scintillators are used for detection of muons over the region $|\eta| \leq 1.5$. The Central Muon Chambers (CMU) [111], located around the outside of the CHA central barrel, consist 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) and Central Scintillator Upgrade (CSP) [112] are located outside the magnet return yoke, which forms a box made of 0.6 m of absorbing steel around the central detector. The CMP and CSP consist of a set of four layers of drift chambers and covering scintillators, measuring the momentum and timing of muons with $p_T \geq 2.0$ GeV/c in the same η range as the CMU.
The Central Muon Extension (CMX) and Central Scintillator Extension (CSX) [112] are conical sections of drift tubes and scintillators located at each end of the central detector and extend the pseudorapidity coverage of the muon system from 0.6 to 1.0.
The fourth muon system, the Barrel Muon Chambers (BMU) and Barrel Scintillator Upgrade (BSU) provides momentum and timing measurement by drift chambers and scintillators. Both sub-detectors are installed on top of the forward toroid barrels (not powered in Run II) and cover 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 muon systems

2.2.4 Luminosity Counters

Measuring the instantaneous luminosity of the Tevatron beam collisions at the B0 interaction point is performed by the Cerenkov Luminosity Counters (CLC) [113]. They are mounted surrounding the beamline in the end-plug at a pseudorapidity of $3.7 \leq |\eta| \leq 4.7$. The CLC is made of 48 gaseous Cerenkov light detectors filled with isobutane, designed to detect the burst of Cerenkov radiation resulting from charged particles flying through the medium. The radiated Cerenkov light is collected and sent to PMT to be read out.

2.2.5 Data Acquisition System

Due to the bunch structure of the Tevatron beams with separation gaps of 0.396 μs, the bunch crossing rate is approximately 2.5 MHz. Taking into account the train structure with their intermediate abort gaps of 2.6 μs, 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 Data Acquisition System (DAQ) [114], a three-level trigger system illustrated in figure 2.10. The first two trigger levels are realized by special-purpose hardware, whereas the third one is implemented by software running on a Linux processor farm.

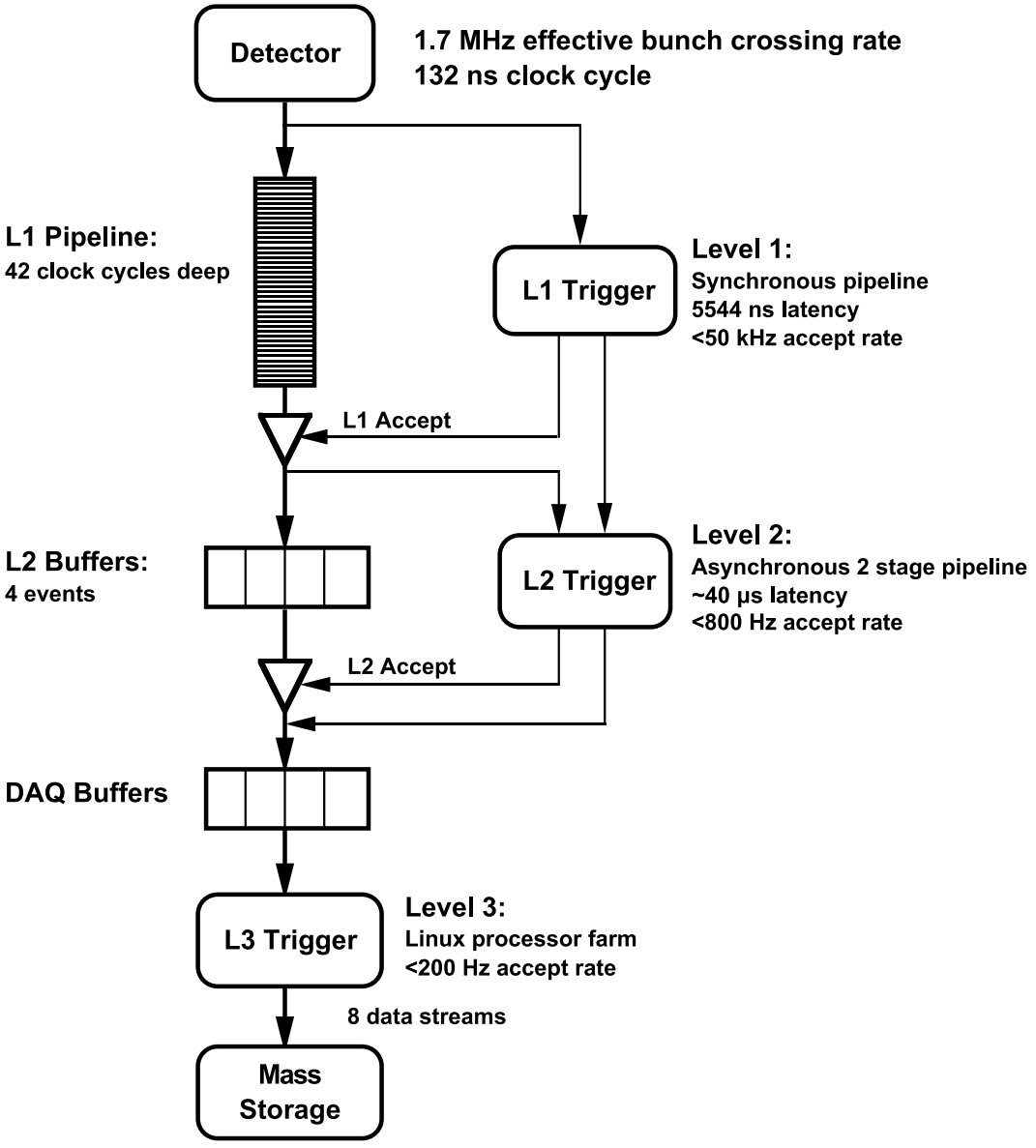


Figure 2.10: Functional block diagram of the CDF data flow

The first level (L1 Trigger) finds physics objects based on information from a small set of sub-detectors. Three parallel systems examine each event synchronous with the clock cycle. Calorimeter trigger boards find calorimeter-based objects, while muon trigger cards identify muons. Additionally, the Extremely Fast Trigger (XFT) [115] 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 within the fixed latency time of about 5 μs whether an event is passed to the second trigger level. The typical L1 accept rate is below 30 kHz.
The L2 Trigger performs minimal event reconstruction within about 40 μs using custom-designed hardware consisting of several asynchronous subsystems. Besides calorimeter, track, and muon based streams, L2 incorporates additional information from CES and SVX II. The Silicon Vertex Trigger (SVT) [116, 117] allows selection of tracks with a large impact parameter which is most important for B hadron reconstruction. Because of the large number of silicon readout channels, the readout chips hold the data of each event on an onboard circular buffer. This information is only forwarded to the SVT when an event passes the L1 Trigger. L2 currently accepts up to 800 events per second which are passed to the third trigger level.
The L3 Trigger [118] is a processor-based filtering mechanism implemented in software which has access to the whole event record, drawing its conclusions based on the full reconstructed event topology. Accepted events are written to permanent storage media with up to 200 Hz, to be further processed at a later date with offline reconstruction algorithms for data analysis. To facilitate the handling of the huge collected data volumes, the events passing all three trigger levels are split into eight different data streams by the Consumer Server/Logger (CSL). The decision to which stream an event belongs to depends on the fired triggers an event has passed; e.g. all events passing any of the high-p_{T} lepton triggers end up in Stream B. Additionally, the CSL copies a fraction of events for immediate data quality online monitoring.

Chapter 3

Event Simulation and Reconstruction

The understanding of efficiencies, acceptances and kinematic properties of collision events is essential for data analyses and thus requires deep knowledge of the physics processes and the detector response. To evaluate measured data, it is therefore useful to simulate all physics processes expected to contribute to the respective data sample. This is realized by Monte Carlo (MC) event generators which randomly produce collision data 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 objects are subject to the same event reconstruction algorithms, which allow for a direct comparison between simulated processes and observed data.

3.1 Monte Carlo Event Simulation

The hard interaction of the incoming beams results in the production of up to hundreds of outgoing particles. Unfortunately, a full theoretical quantum-mechanical treatment is unfeasible due to two main 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. Including finite higher-order corrections through the exact computation of a given number of emissions is performed by matrix element event generators. The common approach to estimate the effects due to emissions at all orders in perturbation theory is done by the parton showering technique. This procedure is 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.

The combination of matrix element generators and showering programs has become essential for analyses of multi-jet event configurations. However, combining both methods may cause severe problems. 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.

3.1.1 Matrix Element Event 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, which compute tree-level matrix elements for a fixed number of partons in the final state. These software programs generally do not include any form of hadronization, thus the final states consist of leptons and 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 jets and hard quarks and gluons.

Alpgen

Alpgen [119] 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. Alpgen allows for the matching of matrix-element hard partons and shower-generated jets, following the so-called MLM prescription.

MadEvent

MadEvent [120] is a multi-purpose tree-level generator which is powered by the matrix element generator MadGraph [121]. Given a standard model process, MadGraph automatically generates the amplitudes for all relevant sub-processes 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 showering MC programs. Recent versions of MadGraph/MadEvent [122] releases provide several approaches of jet-parton matching, including MLM, CKKW [119, 123, 124], and further methods [125].

3.1.2 Showering and Hadronization Event Generators

Showering and hadronization event 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 sub-process, which takes place at high momentum transfer scale Q^2, where the strong coupling constant α_s is small, thus the scattering processes can be calculated in perturbation theory. Higher order effects are added 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 and hadronized into color-singlet hadrons and resultant resonances are decayed. Since the hadronization of quarks and gluons, which describes the formation of jets, takes place at low Q^2 and large α_s perturbation theory cannot be applied. The phenomenologic models, used to describe hadronization in the absence of any firm theoretical understanding are different for distinct Monte Carlo generators. Finally, the underlying event structure and pile-up effects are generated by inclusion of beam remnants and interactions from additional partons in the initial hadrons and between other hadrons in the colliding beams, respectively.

Pythia

The main emphasis of the Pythia [126] event generator is on multi-particle production in collisions between elementary particles. 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 [127, 128] is implemented. In this model the color field between final state quarks and antiquarks is illustrated by color flux tubes or strings. If the potential energy in such a tube is high enough, it can be transformed in new quark-antiquark pair. Thus, colorless hadrons can

be built. PYTHIA additionally provides several sets of parameters, called tunes, which are optimized to model several aspects of collision data like underlying event structure or the contribution from pile-up collisions.

HERWIG

HERWIG [129] 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 $2 \to n$ scattering processes for both the standard model and its supersymmetric extension. Particular emphasis lies thereby on the detailed simulation of QCD parton showers. Other special features of HERWIG are QCD jet evolution with soft gluon interference considered via angular ordering. The cluster model used by HERWIG exploits the preconfinement property of perturbative QCD [130] to form color-neutral clusters which decay into colorless hadrons.

3.1.3 Detector Simulation

A well known description of the response of CDF II detector to the final state particles is crucial for data analyses. The modeling of the detector response is based on a detailed simulation using the GEANT3 package [131]. 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 package [132], a general drift chamber simulation program. To speed up the simulation, the charged particle ionization and drift properties in the COT are parametrized and tuned to data. The calorimeter simulation based on the shower development package GFLASH [133, 134] 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 [135].

3.2 Event Reconstruction

Once detector data is obtained, either from Tevatron collisions or simulated MC events, the raw data needs to be converted to reconstructed physics objects. This is done in two steps: first, information from the sub-detectors is combined to form high-level detector objects, e.g. tracks in the tracking system, energy clusters in the calorimeters. Afterwards, all objects passing certain quality requirements are associated with corresponding candidates of physical objects. The resulting electron, muon, jet, or neutrino candidates can then be further identified at analysis level.

3.2.1 Charged Particle Tracking

Detection and tracking of charged particles is an essential part of collision data 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, called hits, along the trajectory to a particle track. 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. Each helix 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 below. Since the CDF II tracking system is built of two dissimilar sub-detectors, the COT and the silicon system, the pattern recognition strategies vary for each sub-detector. The fit methods in both systems are based on χ^2 minimization.
In a first step, tracks in the COT are reconstructed. The active volume of the COT has larger radii than the silicon tracking system, 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 is only able to reconstruct tracks of charged 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 $p_{\mathrm{T}} > 0.5$ GeV/c. Two different reconstruction algorithms for tracks in the COT are in use. The first one, adopted from Run I, reconstructs and links segments in the super-layers to find the trajectory. The second, faster approach [136] operates at the level of single hits. It starts with a single segment in the outermost superlayer to reconstruct reference trajectories. Due to the large activity, the beam line is firstly used as constraint, being released once the majority of tracks has been found.
The silicon track reconstruction consists of three major approaches. In the first method, called outside-in tracking, tracks found in the COT are extrapolated into the silicon detector [137]. Axial and stereo silicon hits are subsequently added to the track using a progressive fit. The silicon stand-alone tracking method [137] consists of stand-alone pattern recognition in the silicon sub-detectors. To reduce combinatorics, hits already belonging to any other track are not considered. Thus, the main

purpose is to find tracks in the forward region up to $|\eta| \leq 2$ which is not covered by the COT. The third method is called inside-out tracking [138]. Tracks found with the silicon stand-alone algorithm are extrapolated to the COT to identify tracks from particles which did not traverse the entire COT and thus could not be found by the COT tracking algorithms. This results in a significant decrease of fake rates for tracks up to about $|\eta| \leq 1.7$.
For the identification of electrons in the forward region, a special algorithm, called Phoenix (PHX) tracking, is used. This forward tracking algorithm is analog to the outside-in tracking. 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 sub-detector works similarly to the outside-in tracking algorithm.

3.2.2 Primary Vertex Reconstruction

Accurate determination of the primary interaction point of the hard scattering is essential for any high precision analysis, especially when selecting on the lifetime of a particle. Because the interaction region of the particle beams has a substantial volume, the knowledge of the position of the primary vertex affects the measured kinematic properties of the event objects. 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.

3.2.3 Charged Lepton Reconstruction

Depending on their type, the charged leptons are reconstructed from different high-level detector objects. Electron candidates in the central and forward region of the CDF II detector are built of an isolated track matched to an electromagnetic calorimeter cluster in CEM and PEM, respectively. A muon candidate requires an isolated track matched to a track segment in the muon chambers.

3.2.4 Jet Reconstruction

Emerging quarks manifest themselves through fragmentation and hadronization as calorimeter jets, which appear as calorimeter energy deposits from the wide variety of particles that can develop in the shower. The jet energy is calculated from the energy deposited in the electromagnetic and hadron calorimeter towers using a jet clustering algorithm. The cone algorithm clusters the jet 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

is $R = \sqrt{(\eta_{\text{tower}} - \eta_{\text{jet}})^2 + (\phi_{\text{tower}} - \phi_{\text{jet}})^2} \leq 0.4$. The clustering algorithm groups calorimeter towers with $E_{\text{T}i} > 1$ GeV which are not already associated with charged lepton candidates. Here, $E_{\text{T}i} = E_i \cdot \sin\theta$ 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 $E_{\text{T}i}$. 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 the nearest jet. The final jet energy is computed from the final list of towers N_{tow}:

$$E_{\text{jet}} = \sum_{i=0}^{N_{\text{tow}}} E_i \tag{3.1}$$

Jet Energy Correction

Accurate determination of the energy of jets is one of the most difficult measurements. The response of the calorimeter, uninstrumented regions, and further aspects have to be accounted for by several jet energy corrections [109]. Since the corrections are divided into discrete levels to accommodate different effects, a subset of these corrections can be applied, depending on the analysis. By using all levels of corrections, it is possible to estimate the energy of a parton from the energy of its reconstructed jet.

η-dependent Corrections (referred to as Level 1) Due to the geometry of the CDF calorimeter, its response is not uniform in pseudorapidity. This results from the difference in the calorimeter clustering performance between the central and forward region, and from inefficiencies due to cracks between sections of the calorimeter.

Multiple $p\bar{p}$ Interactions (Level 4) At higher luminosities, more than one $p\bar{p}$ interaction occurs at the same bunch crossing. The average number of interactions per bunch crossing varies from one up to eight. These additional interactions cause extra energy to be deposited in the calorimeter, which therefore needs to be subtracted from the jet energy.

Absolute Jet Energy Scale (Level 5) 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 further energy

losses in the uninstrumented regions of each calorimeter, the energy scale of a jet is independent of the CDF II detector and thus comparable with other high-energy experiments.

Underlying Event (Level 6) It is often desirable to determine the energy of the original parton rather than the energy of the corresponding jet, especially for the reconstruction of the invariant mass of a decayed massive particles. Particles from initial-state gluon radiation (ISR) and particles from spectator partons with color connection to the beam remnant lead to energy contributions from the underlying event not related to the original parton.

Out-of-Cone (Level 7) The out-of-cone correction adjusts the particle-level energy for leakage of radiation outside the clustering cone up to $R = 1.3$ due to final-state gluon radiation (FSR) at large angles with respect to the parent parton.

Splash-Out (Level 8) On average, 0.5 GeV of leakage energy is estimated to be outside of the $R = 1.3$ cone and thus the parton energy has to corrected for this splash-out contribution.

3.2.5 Neutrino Reconstruction

Neutrinos cannot be detected by the CDF II experiment, but energetic neutrinos carry away energy from the collision event. Therefore, an escaping neutrino will manifest itself as missing energy in the overall three-dimensional energetic balance of the event. Since the longitudinal component of the energy of the colliding initial-state partons is not known, while the initial transverse component is assumed to be zero, the transverse energetic balance is used for the neutrino reconstruction. The missing transverse energy vector $\vec{\not\!\!E}_{\mathrm{T}}$ is defined by

$$\vec{\not\!\!E}_{\mathrm{T}} = -\sum_i E_{\mathrm{T}}^i \hat{n}_i, \tag{3.2}$$

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. After $\vec{\not\!\!E}_{\mathrm{T}}$ is calculated, it needs to be corrected for the position of the primary vertex of the event, which will cause an adjustment in the direction of the vector associated with each tower. Finally, missing transverse energy $\not\!\!E_{\mathrm{T}} \equiv |\vec{\not\!\!E}_{\mathrm{T}}|$ is defined.

3.2.6 Secondary Vertex Reconstruction

By exploiting special features of the mechanism of jet formation, a method is provided to classify jets into two separate categories of their underlying quark content, namely light and heavy quark-flavor jets. Jets originating from top and bottom quarks can be identified by a observable secondary vertex inside its cone. The top quark decays instantaneously after production into a W boson and a b quark, the latter itself, regardless of its production mode, hadronizes almost immediately. The resulting B hadron, either a b meson (for example $B^0, B^{\pm}, B_s^0$) or a b baryon (e.g. Λ_B), usually carries most of the large transverse momentum of the original b quark. Since B hadrons only decay through weak interactions, they have a considerable long lifetime of about 1.6 ps. In conjunction with a commonly large relativistic boost, this causes B hadrons to travel an average distance L_{xy} of a couple of millimeters transversely before they decay. Therefore, B hadron decays produce secondary vertices displaced from the primary interaction point. By reconstructing charged tracks with large impact parameters d_0, i.e. displaced tracks with a low probability of coming from the primary vertex, possible secondary vertices can be found as illustrated in figure 3.1.

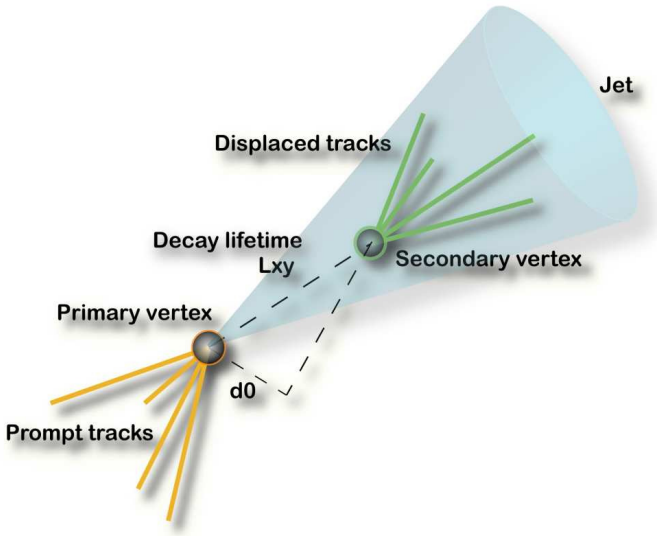


Figure 3.1: Schematic view of displaced tracks forming a secondary vertex

A procedure like the CDF secondary vertex reconstruction algorithm SecVtx [139] is often also referred to as b tagging. SecVtx runs on an per-jet basis within each event and starts with the selection of tracks associated to a jet, which have large impact parameters d_0 with respect to the primary vertex. The selected tracks have to pass certain quality requirements 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. SecVtx uses two passes of a vertex finder approach, the first one with at least three tracks emerging from a joint vertex, the second with only two tracks yet satisfying more stringent track quality requirements.

If a secondary vertex is found, the jet is said to be b-tagged. If the dot product of L_{xy} and the jet axis is positive, i.e. the jet and the reconstructed secondary vertex are in the same detector hemisphere, the jet is called positively tagged. Otherwise the jet is called negatively tagged. Since secondary vertices corresponding to the decay of heavy-flavor hadrons are expected to have large positive L_{xy}, vertices with a negative tag cannot be consistent with heavy-flavor decays and are due to the finite tracking resolution of the CDF tracking system.

Chapter 4

Event Modeling

To perform a search for a yet undiscovered signal process such as electroweak single top-quark production, accurate models predicting the range of expected data are needed for both the signal being tested and for all the expected SM background processes. As the goal of this analysis is to discover single top-quark production if it is present and exclude it if it is absent, the sensitivity of the analysis and the false discovery rate must be evaluated, along with the significance of a signal, if observed. Using a multivariate method, it is crucial to model the observed data correctly in the whole relevant kinematic phase space. Therefore, the complete spectrum of expected physics processes as well as instrumental contributions have to be modeled carefully. Most of the processes are described using Monte Carlo simulation, while some background processes are derived from data.

4.1 Simulated Signal Event Modeling

The CDF Run I single top-quark analyses used PYTHIA to generate signal s- and t-channel MC events. Several authors have pointed out that the LO contribution to single top-quark t-channel production as modeled in parton shower Monte Carlo programs like PYTHIA does not adequately represent the expected distributions of observable jets [79, 140]. To bypass this problem, the CDF Run II single top-quark analyses chose MADEVENT as MC generator [26], which brings along two advantages:
First, it provides the opportunity to generate two different t-channel samples which can subsequently be matched together to better model theoretical expectations.
Second, MADEVENT fully incorporates the spin of the top quark in contrast to the PYTHIA generator. As already discussed, one interesting feature of electroweak top-quark production is that the top quark in its rest frame is almost 100% polarized along the direction of the down-type quark. It is important to include this feature in the Monte Carlo description since it can be used to discriminate single top-quark events from background.

Because MadEvent is designed to produce events at parton level, one needs a parton showering software to generate all expected final-state particles. For this purpose the MadEvent output is passed to Pythia, where a strong-angular-ordered showering is done through emission of QCD radiation.

4.1.1 *s*-channel Single Top-Quark Production

According to reference [79], NLO distributions for *s*-channel single top-quark production are the same as the LO ones multiplied only by a constant known correction, called k-factor. In the same reference, the Ztop software package is introduced. Ztop provides NLO *s*-channel and *t*-channel single top-quark production distributions within the geometrical acceptance of a given detector, which is very appropriate to verify event samples of MC generators. The Ztop software calculates, that at least one jet (apart from the 1^{st} b quark jet coming from the top-quark decay) should be visible within the detectors acceptance in about 90% of all *s*-channel events. In the majority of events this p_T-leading jet is a b-quark jet, which in the following is referred to as 2^{nd} b quark, contrary to the 1^{st} b quark from the top decay. In MadEvent, it is represented by a 2^{nd} b quark at parton level, see figure 4.1 (a). Only in about 10% Ztop expects a light quark jet, which could come from hard initial or final-state gluon radiation and initial-state gluon splitting, respectively. Since the used *s*-channel MadEvent sample is LO and does therefore not include matrix elements with real gluon corrections, which are at the end modeled by the Pythia showering, their contributions do not appear in parton level distributions.

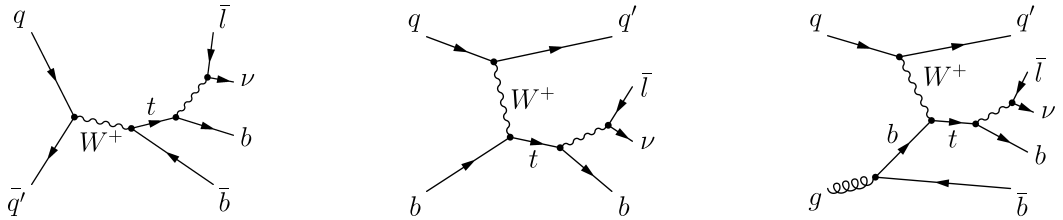


Figure 4.1: Some Feynman diagrams of single top-quark production modes: *s*-channel (a), *t*-channel 2→2 (b), and *t*-channel 2→3 (c).

Figure 4.2 compares the differential NLO p_{T} and $Q_{\mathrm{l}} \cdot \eta$ cross-section distributions with the simulated LO MadEvent *s*-channel signal events. In both the distributions of the leading jets and the top quark, good agreement can be observed. The MadEvent p_{T} spectrum of the leading jet as well as of the top quark seem to be slightly harder, even though the difference is only a few percent.

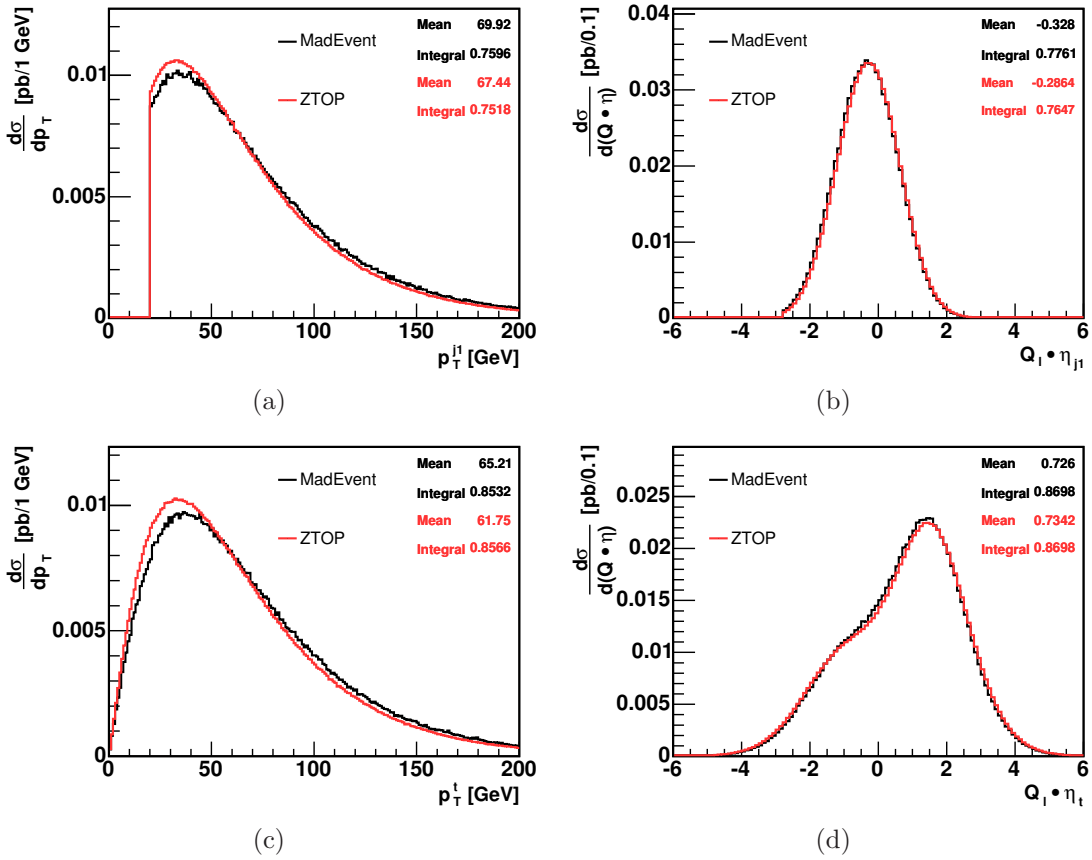


Figure 4.2: Both the s-channel rate and the shape of the differential NLO p_T (a) and $Q_l \cdot \eta$ (b) cross-section distributions of the leading jet are in good agreement. This also applies to both the distributions of the top quark (c) and (d), whose rates of the ZTOP and MADEVENT distributions are normalized to the total theoretical cross section.

4.1.2 t-channel Single Top-Quark Production

The t-channel LO process is a $2 \rightarrow 2$ process with a b quark in the initial state: $b + u \rightarrow d + t$ or $b + \bar{d} \rightarrow \bar{u} + t$, see figure 4.1 (b). For MC event simulation, a b-quark parton distribution function is used, since the b quark stems originally from a virtual gluon splitting into a $b\bar{b}$ sea-quark pair. Since flavor is conserved in the strong interaction, a $\bar{b}$ quark, called 2^{nd} b quark, has to be present in the event as well.

It is known that PYTHIA generates too soft and too far forward distributed 2^{nd} b quarks. The reason for that is that PYTHIA starts with the LO 2→2 diagram, i.e. with a b-quark PDF and then creates the initial state through backward evolution using the DGLAP scheme [141–143]. Using this method, only the soft region of the transverse momentum of the 2^{nd} b quark is well modeled, while the hard region is underestimated.

With MADEVENT, this shortcoming can be avoided by performing a matching [26, 140] of both the LO 2→2 and the most important NLO 2→3 matrix element, see

Figure 4.1 (c). The 2→2 process with a b-quark PDF is expected to cover its soft p_T range, while the 2→3 process with an initial-state gluon splitting into a $b\bar{b}$ pair will model the hard p_T range of the b quark. Both processes differ in the number of final-state partons. The 2→2 matrix element includes the light quark and the decay products of the top quark, namely a lepton, a neutrino and the so-called 1^{st} b quark in the final state. The 2→3 matrix element includes the same final-state partons plus an additional $\bar{b}$ quark, the already mentioned 2^{nd} b quark.

t-channel Simulation Matching Procedure

A joint event sample is created by matching the p_{T} spectrum of the 2^{nd} b quark to the differential cross section predicted by the ZTOP program. The matched t-channel sample is created in such a manner that the p_{T} distribution of 2^{nd} b quarks in the matched t-channel sample consists of $2 \to 2$ events for transverse momenta below a certain cutoff K_{T} and of $2 \to 3$ events for transverse momenta above K_{T}. The ratio R between the two processes are varied until the rate of events with a detectable 2^{nd} b quark jet, that is $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ and $|\eta| < 2.8$, matches 16.7%, as predicted by ZTOP. The final ratio is found to be $R = 2.1$ and the cutoff is derived to be $K_{\mathrm{T}} = 20\,\mathrm{GeV}/c$. As a result, all detectable second b quarks, $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$, of the joint t-channel sample are simulated using the $2 \to 3$ sample. Figure 4.3 illustrates the matching procedure.

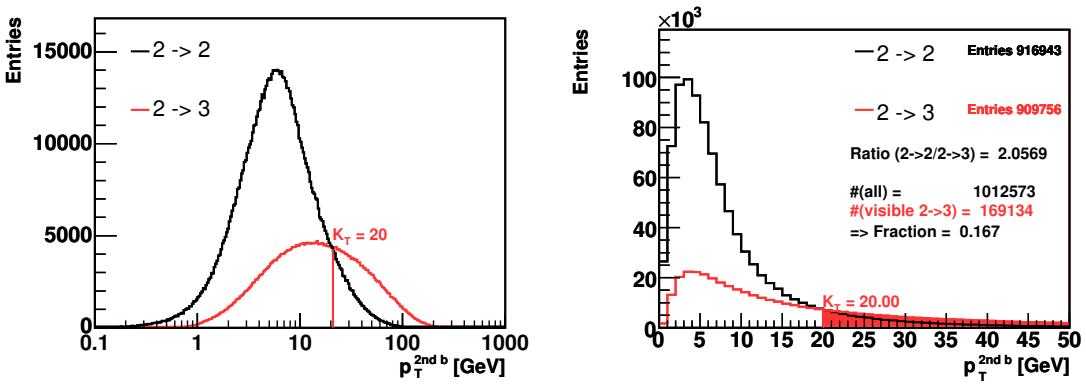


Figure 4.3: Matching of t-channel events of the $2 \to 2$ and the $2 \to 3$ process. The p_{T} distributions of the 2^{nd} b quark are shown, (a) on a logarithmic scale, and (b) on a linear scale. The ratio of $2 \to 2$ to $2 \to 3$ events is adjusted such that the rate of 2^{nd} b quarks with $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ and $|\eta| < 2.8$ matches the NLO prediction of 16.7%. The fraction of these events is illustrated in (b) by the shaded area.

t-channel Simulation Matching Validation

To evaluate the quality of the matching procedure in a qualitative fashion, the kinematic distributions of the primary partons obtained from the matched MadEvent single top-quark samples were checked against NLO differential cross sections calculated with Ztop. Figure 4.4 compares the matching outcome with the differential NLO p_{T} and $Q_{\mathrm{l}} \cdot \eta$ cross-section distributions of the 2$^{\mathrm{nd}}$ b quark and the top quark, respectively. In general, very good agreement is found. Beside the agreement of the rate of the 2$^{\mathrm{nd}}$ b quark distributions, which was the aim of the matching procedure, both the falling p_{T} spectrum and the slightly asymmetric shape of the $Q_{\mathrm{l}} \cdot \eta$ distribution are well modeled by the matched MadEvent sample. In particular, the pseudorapidity distribution of the 2$^{\mathrm{nd}}$ b quark is modeled very well even though it was not used for the matching of the two t-channel samples.

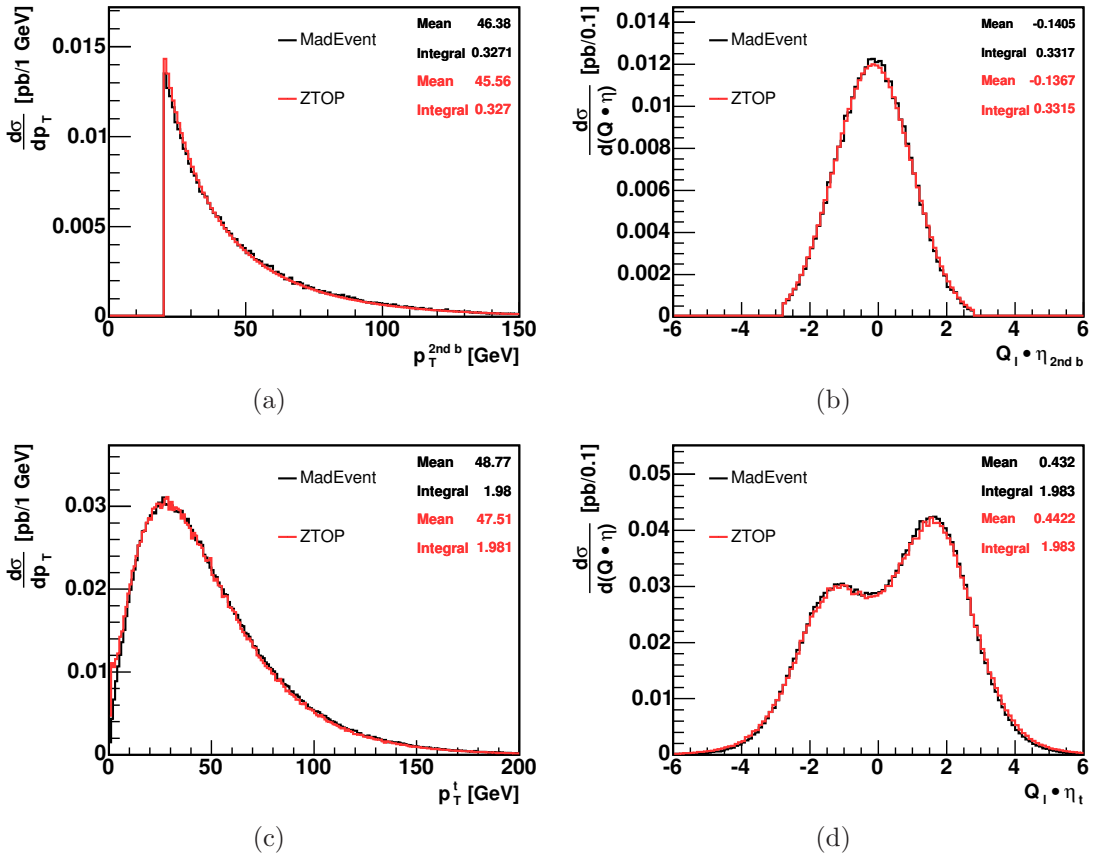


Figure 4.4: Both the t-channel rate and the shape of the differential NLO p_{T} (a) and $Q_{\mathrm{l}} \cdot \eta$ (b) cross-section distributions of the 2$^{\mathrm{nd}}$ b quark are well modeled. This also applies to both the distributions of the top quark (c) and (d), whose rates of the Ztop and MadEvent distributions are normalized to the total theoretical cross section.

4.2 Background Event Modeling

Background processes to single top-quark production in the considered W+jets sample mimic the signal event signature of a single high-energetic isolated charged lepton and large missing transverse energy from the undetected neutrino, both originating from the leptonically decaying W boson, and several additional jets, at least one from heavy-flavor quark production. Since the CDF secondary vertex reconstruction algorithm SecVtx is used to enrich the sample of candidate events with heavy-flavor jets, light-flavor events are heavily suppressed and are only expected to enter through faking a secondary vertex.
The candidate event sample, in the following referred to as lepton+jets sample, is dominated by W-boson production, namely $Wb\bar{b}$, $Wc\bar{c}$, and Wc production, called W+heavy flavor. Since c quarks also have a significant long lifetime, charm production can lead to the reconstruction of a real secondary vertex. In contrast, the W+light flavor process $Wq\bar{q}$ also makes a significant large contribution due to its high production rate which compensates for the low probability of faking secondary vertices.
Additional background sources belong to the following categories: top-quark pair production $t\bar{t}$, two different processes of QCD-multijet events, Z+jets production modes, and diboson production WW, WZ, and ZZ.

4.2.1 Simulation Based Background Event Modeling

For the modeling of $t\bar{t}$ and diboson production, simulated events generated with Pythia are used. Events of $t\bar{t}$ production contribute to the lepton+jets sample if one W boson decays leptonically and the decay products of the other one are lost due to detector acceptance, see figure 4.5 (a).
The diboson production modes WW and WZ contribute as well since their final states include heavy-flavor jets, a neutrino, and a lepton, see figure 4.5 (b) and (c). ZZ events mimic a lepton+jets signature only if one of the two leptons is lost, which coincidently fakes a neutrino through the resulting missing transverse lepton energy, see figure 4.5 (d).

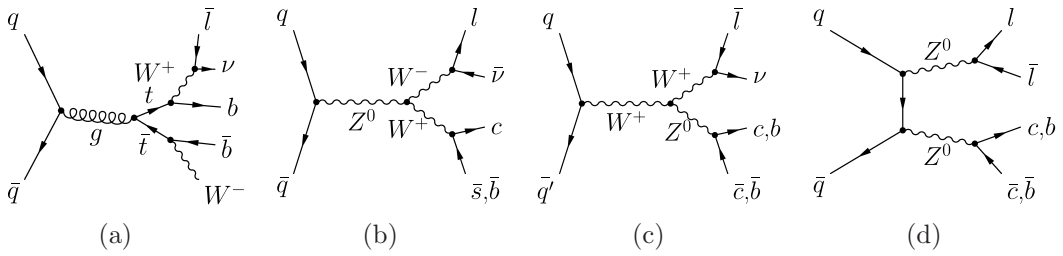


Figure 4.5: Feynman diagrams of $t\bar{t}$ production (a) and the diboson production modes WW (b), WZ (c), and ZZ (d)

W+heavy flavor and Z+jets production are simulated using a combination of ALPGEN matrix element generation and PYTHIA showering. $Wb\bar{b}$ and $Wc\bar{c}$ events have a similar final state as the signal, see figure 4.6 (a). Wc events, as shown in figure 4.6 (b), feature the lepton+jets characteristics by additional jet production through higher order effects which is modeled by showering procedures.
Z+jets events stem mainly from $Zb\bar{b}$ and $Zc\bar{c}$ production, and from $Z \to \tau\bar{\tau}$, shown in figure 4.6 (c) and (d), respectively. Analog to ZZ production, the first two Z+jets modes only contribute if one of the two leptons is lost.

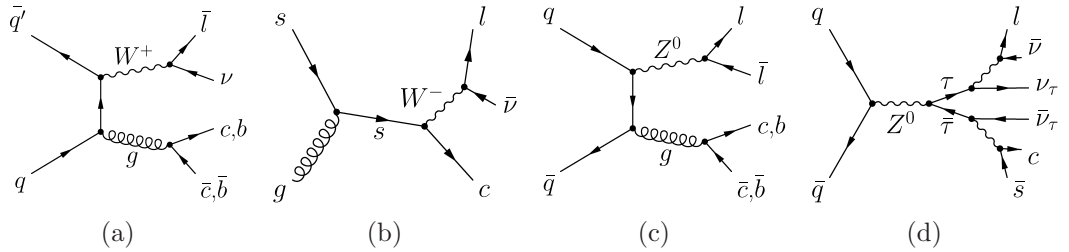


Figure 4.6: Feynman diagrams of W+heavy flavor production $Wb\bar{b}$ and $Wc\bar{c}$ (a), as well as Wc production (b); Z+jets production modes $Zb\bar{b}$ and $Zc\bar{c}$ (c), and $Z \to \tau\bar{\tau}$ (d)

In order to describe events with light-quark jets with falsely reconstructed SECVTX secondary vertices, $Wq\bar{q}$ events, see figure 4.7 (a), simulated with ALPGEN and showered with PYTHIA are used. Due to the very small fraction of events with jets containing reconstructed SECVTX tags in this sample, a huge amount of $Wq\bar{q}$ events would be needed in simulation to reach an acceptable modeling. Thus, the requirements of the SECVTX algorithm are loosened for the $Wq\bar{q}$ modeling and all jets containing at least two tracks with a certain quality, called taggable jets, are assigned to be SECVTX tagged [58].

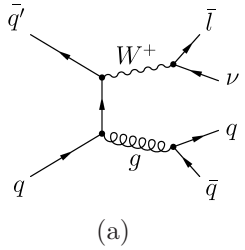


Figure 4.7: Feynman diagram of W+light flavor production $Wq\bar{q}$ (a)

If an event has two taggable jets, both hypotheses are taken into account for the particular event. Each hypothesis is weighted by a tagging probability of the jet considered as tagged. This probability of being falsely b-tagged, also referred to as being mistagged, is defined by a negative tag rate and a correction factor for the mistag asymmetry. The mistag asymmetry is introduced by contributions from long-lived light-flavor particles and material interactions. The negative tag rate is extracted from observed generic jet samples and is parameterized in six variables,

namely the E_{T} and η of the jet, the number of tracks in the jet, the scalar sum of transverse energy of the jets, the number of reconstructed primary vertices, and the z coordinate of the primary vertex associated with the hard interaction.
For the W+jet and Z+jet events produced by ALPGEN and showered by PYTHIA, it is important to take into account that the same n-jet configuration can be generated starting from different $(n - m)$-parton configurations, where the additional m partons are provided by the PYTHIA shower. 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. This is done following the MLM prescription [119].
Since PYTHIA showering can also produce c-quark jets and b-quark jets in events which were generated as light and charm processes, it is important for a proper modeling to promote those original light and charm events to the corresponding $Wc\bar{c}$ and $Wb\bar{b}$ category, respectively [36].

4.2.2 Data Based Background Event Modeling

A substantial background contribution arises from QCD-induced multijet events. Those events mimic the signal signature, if either in direct $b\bar{b}$ production, see figure 4.8 (a), a lepton from a semileptonic b decay or a jet from strong gluon production, see figure 4.8 (b), is erroneously identified as an isolated lepton. Thereby, SECVTX tags arise from heavy-flavor jets or from falsely reconstructed secondary vertices in light-flavor jets. Simultaneous energy mismeasurements can lead to large artificial missing transverse energy, which additionally features the characteristics of lepton+jets events. Since no on-shell W boson is produced, those events are also sometimes referred to as non-W events.

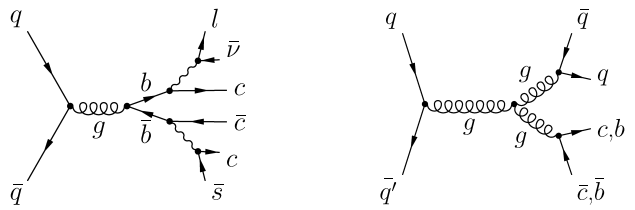


Figure 4.8: Feynman diagrams of two QCD-multijet production modes: direct $b\bar{b}$ production (a) and strong gluon production (b)

While the probability of both the fake of an isolated lepton and the production of large artificial missing transverse energy at once is very small, the huge cross section of QCD-multijet events makes this background non-trivial. Furthermore, because of the resulting rareness of QCD events with such characteristics, it is not feasible to simulate these events.

The QCD background is modeled using two different approaches based on observed data [57,58]. The first one, called anti-electron model, is obtained from central electron trigger data. The events are required to pass all kinematic electron cuts but to fail two of the five electron identification requirements. Even though their kinematic properties resemble those of W-like events, those events are QCD-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, a second QCD model, called jet-electron model, is introduced. It is based on the idea that for a QCD event to enter the lepton+jets sample, a jet has to resemble an electron. Hence, events from generic jet trigger data are required to have a high-energetic jet with a large fraction of energy from the electromagnetic calorimeter.
The anti-electron and jet-electron models do a remarkable job of describing the kinematic properties of the QCD-multijet sample even for muon events, and so they are also used for the muon sample. To assign the events to their appropriate muon sub-detectors, their pseudorapidities are used as criteria. Once a jet is identified as a fake lepton, its charge is assigned randomly, and it is further considered as a real lepton.
Since demanding a SecVtx tagged jet would cause too low statistics, the b tag in the QCD model events has to be faked using taggable jets. If, in a given QCD-multijet event, only one jet is taggable, this one is considered to be the tagged jet. If there are more taggable jets in the event, one of those jets is randomly assigned to be the tagged one. In doing so, each taggable jet has the same probability to be selected.
It is additionally necessary to assign a hypothesis of what underlying quark flavor this newly SecVtx tagged jet is, namely b, c, or light quark flavor. This is done using the flavor composition of the QCD multijet enriched events measured in low missing transverse energy sideband data [28].

Chapter 5

Candidate Event Selection and Yield Estimate

5.1 Candidate Event Selection

The common CDF event selection for single top-quark candidates exploits the kinematic features of the s- and t-channel signal final state, which contains a W boson and the 1^{st} b quark from the top-quark decay, the 2^{nd} b quark, and possibly additional light quarks. To reduce multijet background events, the W originating from the top quark is demanded to decay leptonically into an electron e or muon μ and its corresponding anti-neutrino $\bar{\nu}_{e/\mu}$. The non-detectable neutrino is required to manifest itself as large missing transverse energy in the detector's energy balance. The leptonic W-boson decay into a tau τ and a tau-neutrino $\bar{\nu}_\tau$ is not included because of CDF's low τ identification efficiency. However, W bosons which decay into $\tau\bar{\nu}_\tau$ where the τ further decays into e/μ $\bar{\nu}_{e/\mu}$ enter the event selection.
As a result of the event kinematics of t-channel single top-quark events, the 2^{nd} b quark as well as the scattered incoming light quark is mainly produced in forward direction. This is taken into account by expanding the jet acceptance into more forward regions compared to CDF's standard lepton+jets selection. Finally, candidate events are required to feature two or three jets, where at least one is tagged by the SecVtx algorithm.

5.1.1 Trigger Requirements

To collect collision data in the W+jets sample, it is sensible to trigger on its distinct features. The most prominent features in leptonically decaying W bosons are the appearance of an energetic charged electron or muon and the large missing transverse energy $\not\!\!E_{\mathrm{T}}$, respectively. Thus, high-p_{T} lepton triggers and $\not\!\!E_{\mathrm{T}}$ triggers are used. Each trigger is a composite of the three levels in the CDF trigger system, with stricter quality requirements imposed at each level.

Central electrons have to pass the ELECTRON_CENTRAL_18 trigger that requires a COT track with $p_{\mathrm{T}} > 9$ GeV/c matching an energy deposition in the CEM with $E_{\mathrm{T}} > 18$ GeV. The shower profile of this cluster has to be consistent with the expectation obtained by measurements with test-beam electrons.
Forward electron (PHX) candidates have to pass the MET_PEM trigger, which does not trigger solely on clusters in the PEM calorimeter. This is because of higher background energy depositions from elastic collisions and beam remnants in the forward region. Additionally, due to its central geometry, COT tracking is not available for PHX electron candidates. Instead, this trigger relies on the presence of large missing transverse energy, as well as requiring an energy deposition of at least 20 GeV in the PEM. The lack of tracking requirements for this trigger makes it less pure, thus requiring additional quality selection cuts to purify the sample.
The central muon trigger MUON_CMUP18 requires a track in the COT with $p_{\mathrm{T}} > 18$ GeV/c matched to track segments in both central muon chambers CMU and CMP simultaneously.
Forward muon candidates collected by the MUON_CMX18 trigger have a COT track matched to hits in the CMX muon chambers. Compared to the coincidence of CMUP muon hits in both sub-detectors, CMX candidates are reconstructed from hits in only one sub-detector, which furthermore is much less shielded. Even by requiring a timing signal from CSX scintillators consistent with particles coming from Tevatron collisions, this trigger leads to much higher background rates of non-muon entries.
CDF's method of triggering muons lead to fewer fake events compared to electrons, but at the expense of a lower muon acceptance. Thus, many muon events missed by the muon triggers are potentially recoverable at a later stage during offline reconstruction if the events are collected by a non-muon trigger [57]. Since the most prominent feature of W+jets events besides the high-p_{T} leptons are large missing transverse energy and several jets, it is feasible to trigger on this signature. The MET35_&_TWO_JETS trigger requires $\not\!\!E_{\mathrm{T}} > 35$ GeV and two jets with uncorrected $E_{\mathrm{T}} > 10$ GeV. As instantaneous luminosity increased at the Tevatron, it was necessary to modify this trigger. The resulting MET35_&_CJET_&_JET trigger additionally requires that one of the jets is central, $|\eta_{\mathrm{det}}(\mathrm{jet})| < 1.1$, where $\eta_{\mathrm{det}}(\mathrm{jet})$ is calculated with respect to the origin of the coordinate system at the center of the detector. As a result, the overall trigger rate was reduced without loosing much signal acceptance.
CDF collision data taken from period 0, started in February 2002, until period 19, ended in August 2008, is used to search for single top-quark production. Each Tevatron store is composed of one or more consecutive CDF data acquisition time segments, called runs, which are, from time to time, interrupted by dead time due to detector component failures. By using the so-called good-run list version 25, requiring the whole tracking system, calorimetry, and muon chambers to be in operation, only collision data runs with highest quality are accepted. Run 141544 recorded during store 1120 on March, 23^{rd}, 2002 includes the first good quality data considered. The newest collision data was acquired in run 266513 during store 6378 on August, 24^{th}, 2008, more than 5 years and 5000 Tevatron stores later.

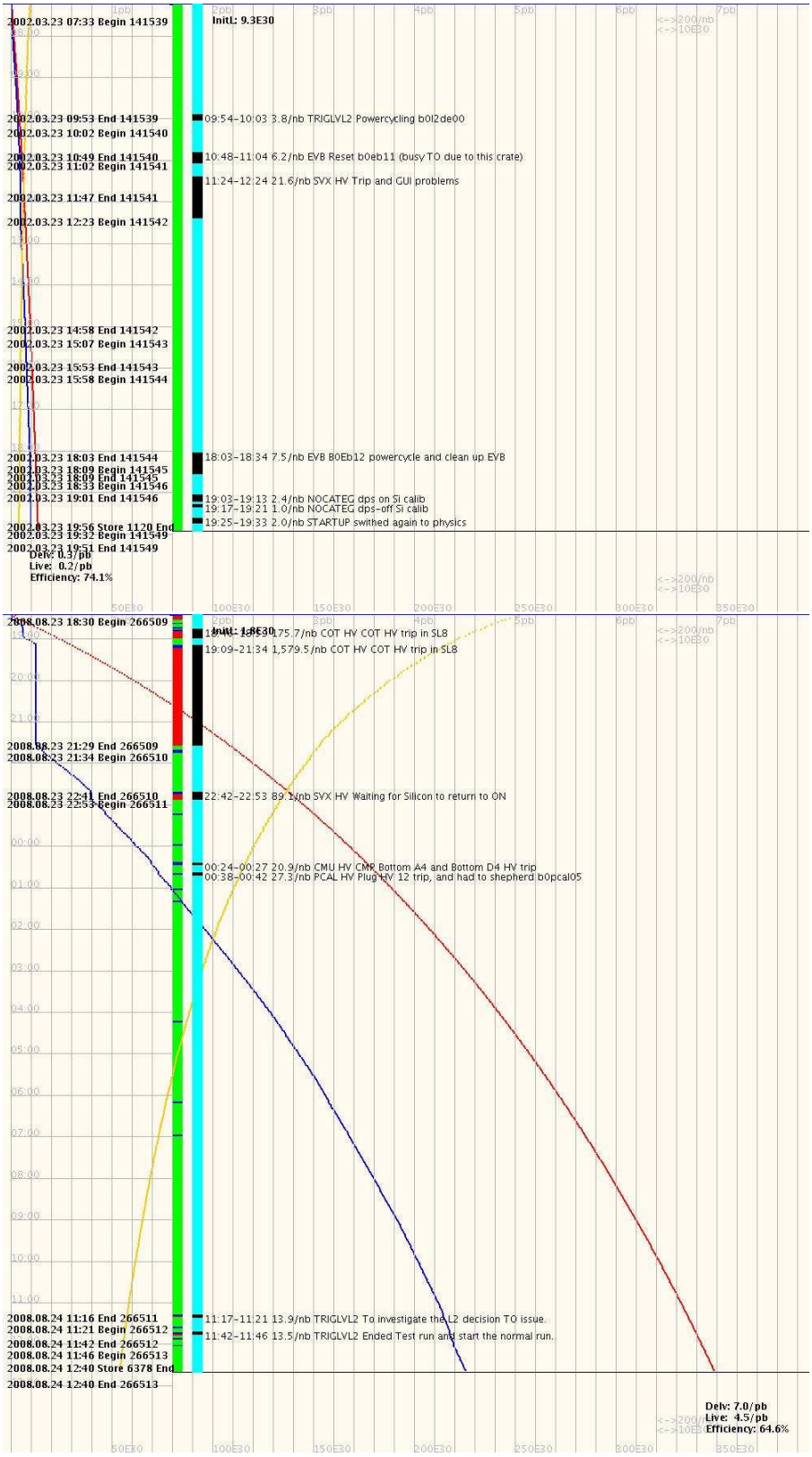


Figure 5.1: First and last Tevatron stores, which are used in the search for single top-quark production, and their development over time from top to bottom: decreasing instantaneous luminosity (yellow curve), delivered integrated luminosity (red curve), integrated luminosity recorded by CDF (blue curve). The first figure shows store 1120, which started on March, 23^{rd}, 2002. Below, store 6378, ended on August, 24^{th}, 2008, is displayed.

Figure 5.1 shows the progress of CDF's data acquisition during those two Tevatron stores. The yellow curve corresponds to Tevatron's exponentially declining instantaneous luminosity with its axis at the bottom. Store 1120 had a initial luminosity of about $10 \cdot 10^{30}$ cm^{-2}s^{-1} compared to about $250 \cdot 10^{30}$ cm^{-2}s^{-1} of store 6378. The red and blue curves are the integrated luminosities delivered by the Tevatron and recorded by CDF, respectively. Their corresponding axis is located at the top of each figure. Both curves separate over time due to unavoidable trigger dead time and detector failures. The latter causes are also visible in the labeled interruptions of the vertical green and blue bars representing different aspects of CDF's data acquisition system.

The largest samples of collected CEM, PHX, and CMUP trigger data amounts to an integrated luminosity of 3.2 fb^{-1} each. Since $\not\!\!E_{\mathrm{T}}$ trigger event rates would exceed the overall trigger accept rates at high luminosities, only an known fraction of those events are accepted, leading to a smaller $\not\!\!E_{\mathrm{T}}$ trigger data sample of 3.0 fb^{-1}. The latter applies also for CMX trigger events, even though with a higher fraction. Additionally, the CMX trigger was not available at the beginning of Run II, which further reduces the collected data to finally 3.1 fb^{-1}. Figure 5.2 shows the time period of collected Tevatron collision data used in the search for single top-quark production.

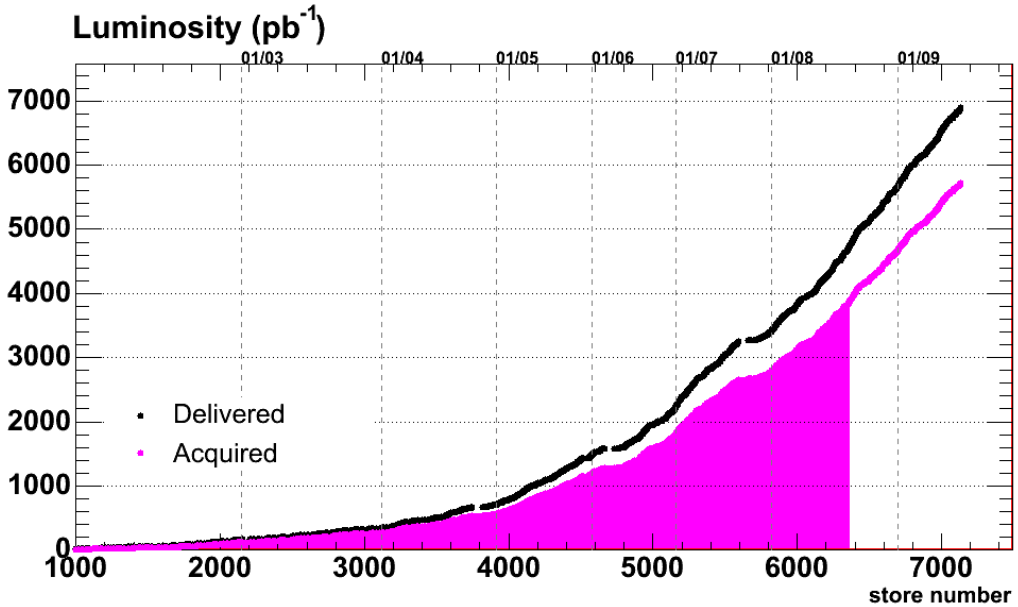


Figure 5.2: Delivered (upper curve) and recorded to tape (lower curve) integrated luminosity at the CDF since the start of Run II, the filled area corresponds to the collision data analyzed.

5.1.2 Jet Identification and Missing Transverse Energy Requirement

Hadronic jet candidates are reconstructed with a cone of $R = 0.4$ without taking into account energy deposits associated to any reconstructed charged lepton. Since single top-quark signal events feature forward jets, the jet acceptance is enlarged up to $|\eta_{\text{det}}(\text{jet})| < 2.8$, while CDF's standard lepton+jets selection is limited to $|\eta_{\text{det}}(\text{jet})| < 2.0$. All jet candidates have to fulfill $E_{\text{T}} > 20$ GeV. Thereby, the jet energies are corrected up to Level 5, which transforms them into an absolute scale. At least two jets with corrected $E_{\text{T}} > 25$ GeV and a distance of $R = 1.0$ in-between are required for $\not\!\!E_{\text{T}}$ trigger events, and, additionally, one of the reconstructed jets must be central, $|\eta_{\text{det}}(\text{jet})| < 0.9$. Figure 5.3 shows the jet multiplicity of s- and t-channel signal MC events after trigger requirements and jet identification.

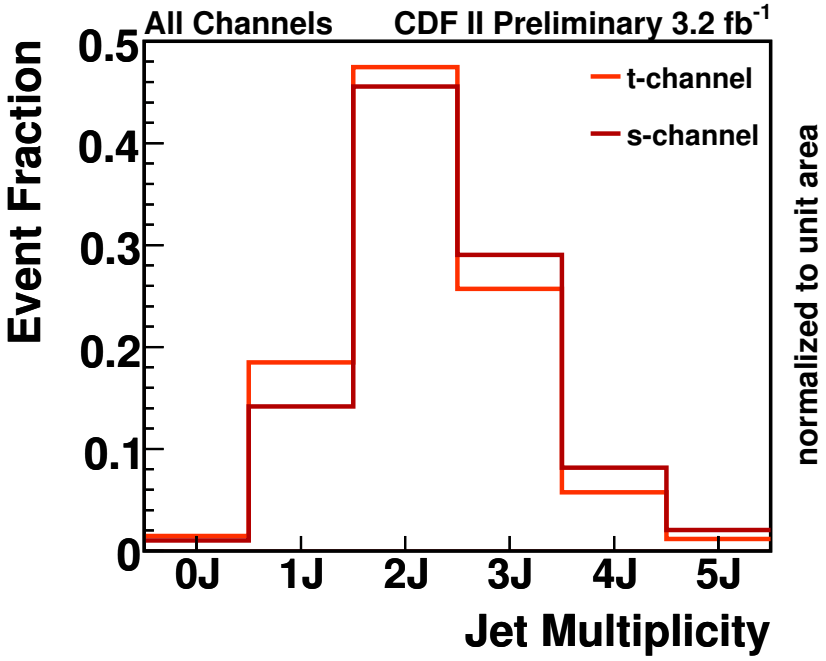


Figure 5.3: Simulated jet multiplicity for s- and t-channel signal events after all trigger and jet identification requirements

Since the vast majority of signal events feature two or three jets, such a number of jets is required for signal candidate events. To reduce the amount of background candidates, in particular events from light-flavor processes, at least one of these jets has to be tagged as a b-quark jet by using displaced secondary vertex SecVtx information.

Further reduction of background processes is achieved by selecting on $\not\!\!E_{\text{T}}$. Because its calculation is based on calorimeter towers, $\vec{\not\!\!E}_{\text{T}}$ has to be adjusted for the effect of the jet corrections. This additionally includes jet energy corrections for loose jets, which are defined to have E_{T} between 12 GeV and 20 GeV. Since muons pass

the calorimeters without showering as minimum ionizing particles, corrections are applied by adding all transverse momenta of the traversing muons to the vector sum. Subsequently, the corresponding average ionization energies are removed. The resulting corrected $\not\!\!E_{\mathrm{T}}$ is required to be greater than 25 GeV.
Since the $\not\!\!E_{\mathrm{T}}$ and MET_PEM triggers use uncorrected $\not\!\!E_{\mathrm{T}}$ at trigger levels, the corresponding data samples are biased with respect to fully corrected $\not\!\!E_{\mathrm{T}}$ used for candidate event selection. To compensate, all simulated signal and background MC events are weighted by a $\not\!\!E_{\mathrm{T}}$ turn-on function to sculpt their kinematics to match the data. This function is determined from W-boson-enriched collision data. Analogously, the same applies to the electron transverse energy in MET_PEM events. Those events are triggered, amongst others, on electron E_{T} from the PEM, which has non-negligible energy corrections compared to CEM. Hence, an additional forward electron E_{T} turn-on function is applied to MET_PEM trigger MC events to simulate the trigger bias in collision data.

5.1.3 Lepton Candidate Identification

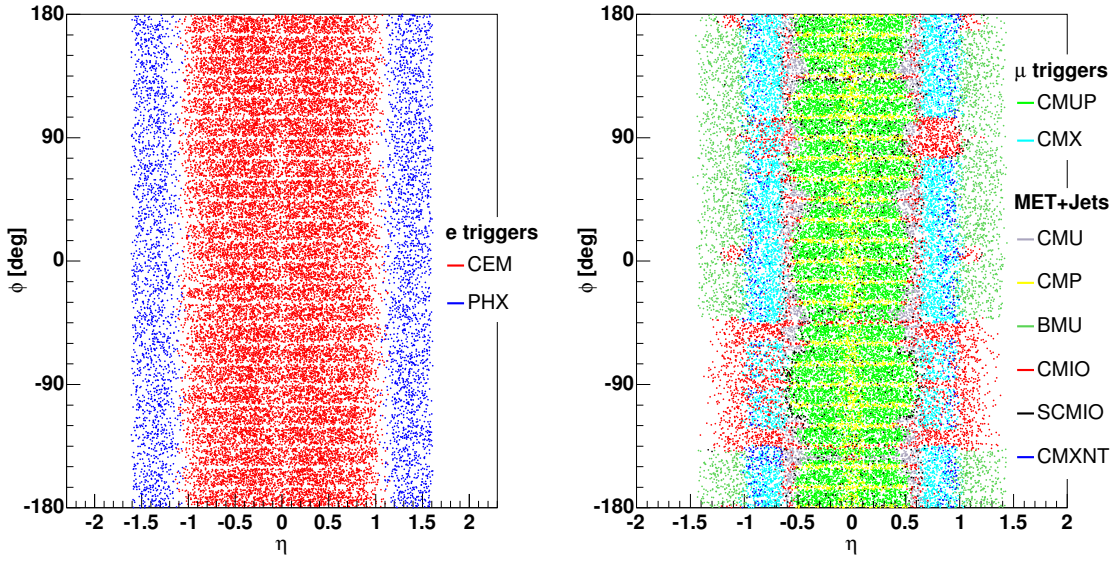
Charged lepton candidates from the offline reconstruction have to pass further identification criteria in order to improve their purity. The lepton candidates are divided in several groups of lepton types depending on the fired trigger. Muons and electrons from events of the high-p_{T} lepton triggered coverage, called TLC events, are of four different types, namely CEM, PHX, CMUP, and CMX. Muon events collected by the $\not\!\!E_{\mathrm{T}}$ trigger are called extended muon coverage (EMC) events.

CEM electrons have a pseudorapidity range of $|\eta| \leq 1.1$. Their energy deposition in the CEM has to be larger than 20 GeV, their reconstructed track $p_{\mathrm{T}} \geq 10$ GeV/c. In addition, the ratio of the energy in the hadronic calorimeter and the energy in the electromagnetic calorimeter ($E_{\mathrm{HAD}}/E_{\mathrm{EM}}$) for the cluster matching the CEM candidate, has to be smaller than $(0.055 + (0.00045 \times E))$.
The PEM calorimeter detects electrons in the forward region, whereas only electrons with $1.1 \leq |\eta| \leq 1.6$ are taken into account. Furthermore, $E_{\mathrm{T}} \geq 20$ GeV, $E_{\mathrm{HAD}}/E_{\mathrm{EM}} \leq 0.05$, and a track with at least three silicon hits pointing from the calorimeter cluster to the primary vertex, are required for PHX electrons.
In general, electron candidates are rejected, if an additional high-p_{T} track is found forming a common vertex with the track of the candidate and has a curvature of opposite sign, which makes them likely to stem from photon conversion.

CMUP and CMX muons are identified by requiring a COT track with $p_{\mathrm{T}} > 20$ GeV/c that extrapolates to tracks in the corresponding muon chambers. Candidates have to be detected in the CMU and CMP chambers simultaneously, or in the CMX, covering a pseudorapidity region of $|\eta| \leq 0.6$ and $0.6 \leq |\eta| \leq 1.0$, respectively.
EMC muon candidates collected by the $\not\!\!E_{\mathrm{T}}$ trigger consists of six orthogonal types, namely CMU, CMP, BMU, CMIO, SCMIO, and CMX_NT. Each of them have to pass slightly different quality criteria due to their geometry. CMU only and CMP only candidates have tracks in the CMU and CMP chambers, but not in the corre-

sponding overlapping CMP and CMU chambers, respectively. This can be caused by slightly different muon chamber coverages and configurations. BMU muons have entries in the forward, $1.0 \leq |\eta| \leq 1.5$, barrel muon chambers. CMIO and SCMIO muons are isolated tracks matched to calorimeter clusters that do not point towards a muon detector or that point towards a non-fiducial part of a muon detector, respectively. These tracks are required to be matched to a COT track and a low-energy calorimeter cluster consistent with the criteria of minimum ionizing particles. CMX_NT muon candidates leave tracks in the most forward region of the CMX detector, which is not triggerable due to the limited COT geometry. They have the same quality requirements as the CMX muon sample.
For all muon types, further quality requirements are made. The energy depositions in the electromagnetic and hadronic calorimeters have to correspond to the expectations regarding minimum ionizing particles. To reject cosmic muons or muons from in-flight decays of long-lived particles such as K_S, K_L, and Λ, the impact parameter d_0 of the track must be small. Cosmic muons are further rejected through their characteristic track timing and topology.

Finally, only isolated lepton candidates are considered, whereby a candidate is defined isolated if the E_{T} and p_{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} and p_{T}, respectively. Figure 5.4 illustrates the different lepton types dependent on the geometry of their corresponding coverage in η and ϕ.



(a) (b)

Figure 5.4: The geometrical coverage of all considered electron (left-hand side) and muon (right-hand side) candidates

5.1.4 Event Vetoes

Further event selection requirements are imposed to suppress certain sources of background events, mostly to improve the background modeling.

Dilepton Veto

To ensure that there is exactly one lepton per event, events are rejected which include additional leptons, even if the definition is more loose. Loose leptons have to pass all lepton quality criteria except the isolation requirement.

z Vertex Cut

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

Z Boson Veto

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

QCD Veto

Additional requirements are introduced to further suppress QCD-multijet events in which no real W boson appears [36]. CEM and PHX events must have $M_{\mathrm{T}}\ (W) > 20$ GeV/c^2, while muon-candidate events are required to have $M_{\mathrm{T}}\ (W) > 10$ GeV/c^2. The transverse mass of the W boson is given by

$$M_{\mathrm{T}}\ (W) = \sqrt{2 \cdot p_{\mathrm{T}}\ (\mathrm{lep}) \cdot \not\!\!E_{\mathrm{T}} - \vec{p}_{\mathrm{T}}\ (\mathrm{lep}) \cdot \vec{\not\!\!E}_{\mathrm{T}}}\ , \tag{5.1}$$

where p_{T} (lep) is the transverse momentum of the charged lepton.
CEM events are additionally required to be above the triangular cut $\not\!\!E_{\mathrm{T,sig}} > -0.05\ c^2/\mathrm{GeV} \cdot M_{\mathrm{T}}\ (W) + 3.5$, which is motivated in figure 5.5. The significance of the missing transverse energy $\not\!\!E_{\mathrm{T,sig}}$ is thereby defined as

$$\not\!\!E_{\mathrm{T,sig}} \equiv \frac{\not\!\!E_{\mathrm{T}}}{\sqrt{\sum_{\mathrm{jet}} C_{\mathrm{JES}}^2 \cdot \cos^2\left(\Delta\phi\ \left(\mathrm{jet}, \not\!\!E_{\mathrm{T}}\right)\right) + \cos^2\left(\Delta\phi\ \left(\not\!\!E_{\mathrm{T}}^{\mathrm{raw}}, \not\!\!E_{\mathrm{T}}^{\mathrm{cor.}}\right)\right)}}\ , \tag{5.2}$$

where C_{JES} is the jet energy correction factor and $\Delta\phi\ \left(\not\!\!E_{\mathrm{T}}^{\mathrm{raw}}, \not\!\!E_{\mathrm{T}}^{\mathrm{cor.}}\right)$ is the azimuthal angle between uncorrected and corrected missing transverse energy.

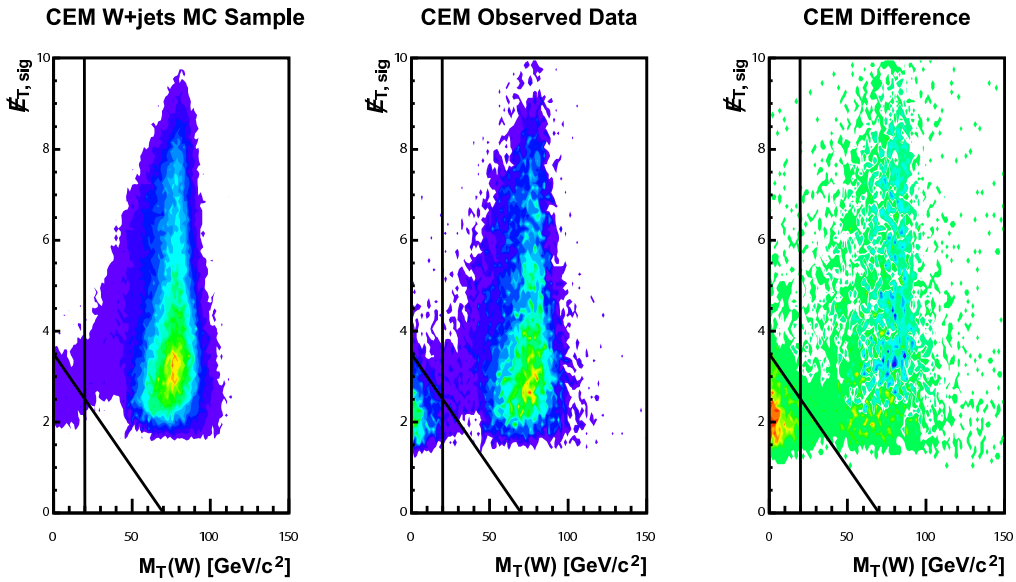


Figure 5.5: Illustration of the QCD veto for CEM events. The distributions of $\not\!\!E_{\mathrm{T,sig}}$ versus M_{T} (W) is shown for several samples before applying SecVtx tags. On the left-hand side, the predicted distribution of the W+jets sample is shown. The middle plot shows the distribution of collision data. On the right-hand side, the difference between prediction and observation is shown. The lines represents the triangular and M_{T} (W) cuts.

Furthermore, CEM events have to fulfill $\not\!\!E_{\mathrm{T,sig}} > 2.5 - 2.5 \cdot \left|\Delta\phi\ \left(\vec{\not\!\!E}_{\mathrm{T}}, \mathrm{jet1}\right)\right| /0.8$ where $\Delta\phi\ \left(\vec{\not\!\!E}_{\mathrm{T}}, \mathrm{jet1}\right)$ is the azimuthal angle between $\vec{\not\!\!E}_{\mathrm{T}}$ and the jet with the second largest uncorrected E_{T}.
PHX events must fulfill $\not\!\!E_{\mathrm{T,sig}} > 2.0$ and, for the first two leading jets, $\not\!\!E_{\mathrm{T}} > 45$ GeV-30 GeV$\cdot \left|\Delta\phi\ \left(\vec{\not\!\!E}_{\mathrm{T}}, jet\right)\right|$, where $\Delta\phi\ \left(\vec{\not\!\!E}_{\mathrm{T}}, jet\right)$ is the azimuthal angle between $\vec{\not\!\!E}_{\mathrm{T}}$ and the corresponding jet.
All muon candidate events are only required to have M_{T} $(W) > 10$ GeV$/c^2$, except for the SCMIO candidates, which requires M_{T} $(W) > 20$ GeV$/c^2$ because of greater observed QCD contamination.

Figure 5.6 shows the jet and b-tag multiplicity of simulated s- and t-channel single top-quark events after passing all selection requirements. It is obvious, that the channel with two jets and one b tag is by far the most populated one for both signal processes.

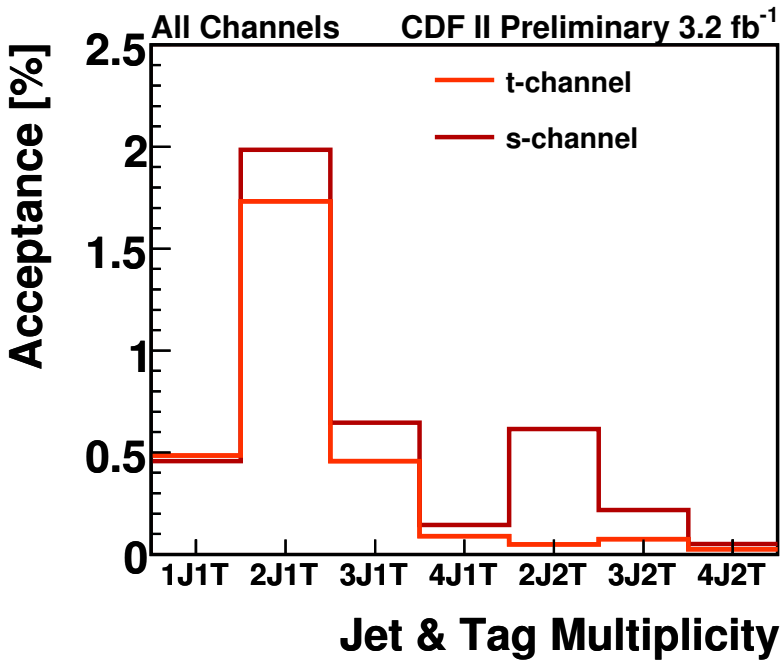


Figure 5.6: Simulated jet and b-tag multiplicity for s- and t-channel single top-quark events after passing all selection requirements not including trigger and identification efficiencies

5.2 Candidate Event Yield Estimate

The expected number of events under SM assumptions for the signal and all contributing background processes are derived from the CDF's so-called Method II. It is a procedure to calculate the normalization of processes in the SecVtx tagged lepton+jets sample. Method II assumes that all processes contributing significantly to the lepton+jets sample are known. In case of the search for single top-quark production analyses, these are single top-quark, top-quark pair, and diboson production, W+jets and Z+jets production, and QCD-multijet events.
To derive the corresponding expectations, both simulated MC events as well as observed sideband data and measured quantities are incorporated. First, the expectations of theoretically well understood processes are determined using simulated events. Processes with large theoretical uncertainties are treated by a data-based method utilizing the pretag sideband data, which is corrected for the expected number of pretag events estimated with simulated MC events. The number of pretag QCD-multijet events is evaluated and also subsequently subtracted to enable the determination of the expected pretag W+jets fraction. To finally obtain the number of expected W+jets events in the b-tagged sample, one has to differentiate between W+heavy flavor (W+HF) and W+light flavor (W+LF) processes even in the pretag sample. This is done using heavy-flavor fractions derived from the corresponding simulated MC event samples. The latter is also used to extract b-tagging rates, which, applied to each W+HF component, results in the estimate of the rate of W+HF events in the b-tagged sample.
However, since the MC simulation does not properly predict the heavy-flavor fractions, an additional correction factor, called K-factor, obtained in the one jet sideband data, is required to match the predictions to collision data. Using the remaining number of pretag events, the expected number of W+LF events is estimated. In the final step, the number of b-tagged QCD-multijet events is determined.
The Method II procedure is performed individually in eight orthogonal lepton+jets channels considered. Those are events with two or three jets, each with one or at least two b tags divided in TLC and EMC event candidates, respectively.

5.2.1 Simulation Based Event Yield Estimate

Samples of simulated events are used to derive the expectation of both signal processes, single top-quark s- and t-channel production, as well as the background contributions of $t\bar{t}$, diboson, and Z+jets production to the b-tagged lepton+jets sample. 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. For a certain process, the number of expected events is given by

$$\hat{\nu} = \sigma \cdot \varepsilon_{\text{evt}} \cdot \mathcal{L}_{\text{int}} \ , \tag{5.3}$$

where σ is the theoretically predicted cross section of the respective process, ε_{evt}

is the event detection efficiency, and $\mathcal{L}_{\text{int}}$ is the integrated luminosity. The event detection efficiency is estimated by performing the event selection on 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, $\varepsilon_{\text{corr}}$. The samples of simulated events are produced such, 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

$$\varepsilon_{\text{evt}} = \varepsilon_{\text{mc}} \cdot \varepsilon_{\text{BR}} \cdot \varepsilon_{\text{corr}} \cdot \varepsilon_{\text{trig}} \ . \tag{5.4}$$

Signal Single Top-Quark Production

For the signal, $\varepsilon_{\text{evt}}(t\text{-channel}) = (1.8 \pm 0.3)\%$ and $\varepsilon_{\text{evt}}(s\text{-channel}) = (2.8 \pm 0.4)\%$ are found, including all trigger and identification efficiencies . The predicted cross sections for s- and t-channel single top-quark production are $\sigma_{s-\text{channel}} = 0.88^{+0.12}_{-0.11}$ pb and $\sigma_{t-\text{channel}} = 1.98^{+0.28}_{-0.22}$ pb, respectively, assuming $m_t = 175$ GeV/c^2 [78, 79]. In 3.2 fb^{-1} of Tevatron collision data, the expected total number of signal t-channel events is 113.8 ± 16.9, while 77.3 ± 11.2 s-channel events are predicted. All quoted uncertainties include the systematic uncertainties on the theoretical cross-section calculation, the trigger efficiency, the lepton ID efficiency, the b-tagging correction factor, and the luminosity.

Top-Quark Pair Production

For top-quark pair production, PYTHIA MC events are used to determine ε_{evt}. $t\bar{t}$ 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. Events from top-quark pair production are expected to be present in the selected data sample due to limited detector acceptance. Either one lepton in dilepton events is lost or several jets are lost or misidentified in non-dilepton events. To compensate for a higher dilepton veto efficiency in simulated events due to differences in lepton identification and reconstruction efficiencies, the dilepton ε_{evt} are corrected using measured dilepton events from Z-boson decays. The $t\bar{t}$ production normalization is based on theoretically predicted NLO cross-section [144, 145]. The differences to other cross-section predictions [146, 147], as well as the uncertainty due to a variation of the top-quark mass is incorporated as additional systematic uncertainty to the cross section. This leads to a cross section of $\sigma_{t\bar{t}} = (6.70 \pm 0.83)$ pb.

Diboson Production

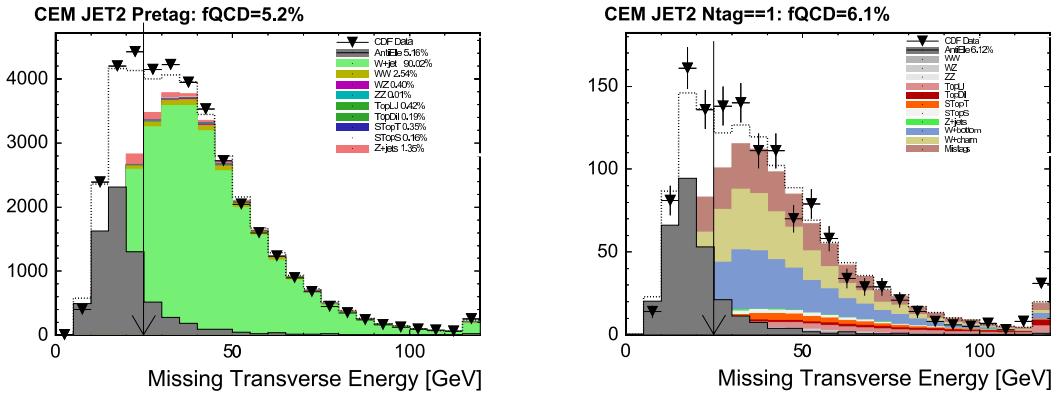
To obtain the expected number of diboson events in the selected data sample, the theoretical NLO cross-section predicted for a center-of-mass energy of $\sqrt{s} = 2.00$ TeV [148] are rescaled to $\sqrt{s} = 1.96$ TeV. This leads to $\sigma_{WW} = (12.40 \pm 0.25)$ pb, $\sigma_{WZ} = (3.96 \pm 0.06)$ pb, and $\sigma_{ZZ} = (1.58 \pm 0.02)$ pb. As for $t\bar{t}$ production, simulated events from PYTHIA are used to determine ε_{evt}.

Z+jets Events

In an analogous manner using simulated ALPGEN events, the expectation for Z+jets events are derived, except that CDF's measured inclusive Z+jets cross-section of $\sigma_{Z+\text{jets}} = (787.4 \pm 50.0)$ pb [149] is used instead of a theoretical one.

5.2.2 QCD-Multijet Event Yield Estimate

The estimation of the expected number of QCD-multijet events makes use of the pretag sideband data. Here all selection criteria but the secondary vertex tag requirement are applied to the collision data. To obtain the fraction of QCD-multijet events in the pretag sample, the selection requirement of $\not\!\!E_{\text{T}}$ is additionally omitted. The $\not\!\!E_{\text{T}}$ distribution of all contributing pretag processes and the corresponding anti-lepton QCD-model template are fitted to the observed $\not\!\!E_{\text{T}}$ spectrum of the pretag data, as shown in figure 5.7 (a) for CEM events with two jets. The estimate of the SECVTX tagged QCD event yield is performed in a similar way after the rates of all other background processes are specified, e.g. see figure 5.7 (b). This is mandatory to establish the $\not\!\!E_{\text{T}}$ distribution according to the full background prediction.



(a) (b)

Figure 5.7: Fits to the $\not\!\!E_{\text{T}}$ distribution to obtain the QCD-multijet fraction in the pretag sample (a) and b-tagged sample (b). The plots exemplarily show CEM candidate events with two jets.

5.2.3 W+Heavy Flavor Event Yield Estimate

The number of expected W+jets events in the pretag sample, $\hat{\nu}^{\text{pre}}_{W+\text{jets}}$, can be calculated by subtracting the number of QCD-multijet $\left(\hat{\nu}^{\text{pre}} \cdot \left(1 - F^{\text{pre}}_{\text{QCD}}\right)\right)$, electroweak diboson $(\hat{\nu}^{\text{pre}}_{\text{ewk}})$ and top-quark $\left(\hat{\nu}^{\text{pre}}_{\text{top}}\right)$ events in the pretag sample from the overall number of observed pretag events $\hat{\nu}^{\text{pre}}$ by

$$\hat{\nu}^{\text{pre}}_{W+\text{jets}} = \hat{\nu}^{\text{pre}} \cdot \left(1 - F^{\text{pre}}_{\text{QCD}}\right) - \hat{\nu}^{\text{pre}}_{\text{ewk}} - \hat{\nu}^{\text{pre}}_{\text{top}} \, . \tag{5.5}$$

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

$$\hat{\nu}_{W+\text{HF}} = \hat{\nu}^{\text{pre}}_{W+\text{HF}} \cdot \epsilon_{\text{tag}} = (\hat{\nu}^{\text{pre}} \cdot (1 - F_{\text{QCD}}) - \hat{\nu}_{\text{ewk}} - \hat{\nu}_{\text{top}}) \cdot f_{\text{HF}} \cdot K_{\text{HF}} \cdot \epsilon_{\text{tag}} \, , \tag{5.6}$$

where the tagging efficiency ϵ_{tag} is at the order of up to 40% and derived from Monte Carlo simulations. It is different for $Wb\bar{b}$, $Wc\bar{c}$, and Wc processes, leading to a different amount of expected events for these heavy-flavor background processes. The fraction f_{HF} of W+HF events in the pretagged sample is also derived from simulations. The heavy-flavor scale factor K_{HF} adjusts the predictions such, that the latter match the observation from collision data in the one jet sideband. This control sample is W+jets dominated and has negligible signal contamination and high statistics, making it ideal to measure the heavy-flavor content. This is done by fitting templates of all contributing processes to the one jet data distribution of a flavor discriminating variable. This variable, the KIT flavor separator, is able to distinguish between lepton+jet events with light-quark, c-quark, and b-quark jets [28, 30]. It exploits several intrinsic properties of the SecVtx reconstructed secondary vertex, which are fed into a neural network. Its output distribution allows for the separation of the underlying quark flavor. Since CDF measures the Wc cross-section consistent with Alpgen predictions, K_{HF} is set to unity for this process [150]. For $Wb\bar{b}$ and $Wc\bar{c}$, a common factor of $K_{\text{HF}} = 1.4 \pm 0.4$, in its conservative uncertainty consistent with studies in higher jet-multiplicity samples, is derived.

5.2.4 W+Light Flavor Event Yield Estimate

The number of expected W+LF events is estimated by a parameterization of the rate of negative secondary vertex tags (mistags) in generic jet trigger data [139]. A mistag is defined as a jet which does not result from the fragmentation of a heavy-flavor quark, yet has a reconstructed secondary vertex. Mistags are caused mostly by random overlap of tracks which are displaced from the primary vertex due to tracking errors, although there are other contributions from long-living particles and nuclear interactions with the detector material as well. A jet is said to be negatively tagged if the transverse decay-length significance $\Delta L_{2D}/\sigma_{2D}$ is smaller than -7.5, contrary to $\Delta L_{2D}/\sigma_{2D} > 7.5$ for positively tagged jets. Since the secondary vertex finder algorithm is symmetric in its treatment of L_{2D}, the tracking-related

mistags should occur at the same rate for $L_{2D} > 0$ and $L_{2D} < 0$. Therefore, a good estimate of the positive mistag rate due to resolution effects can be obtained from the negative tag rate.
The rate of negative tags for all jets is measured in an inclusive sample of jet trigger data. The per-jet tag rate is parameterized in six kinematic variables: the E_{T} and η of the jet, the number of tracks in the jet, the scalar sum of transverse energy of the jets, the number of reconstructed primary vertices, and the z coordinate of the primary vertex associated with the hard interaction. This rate is used to estimate the number of negatively tagged events $\hat{\nu}^{\mathrm{pre}}_{\mathrm{tag}-}$ in the observed pretag sample. Since the negative tag rate does not fully reflect the positive mistags due to long-living particles or detector material interactions, a correction factor for the mistag asymmetry γ, determined using fits to measured pseudo-$c\tau$ spectra of tagged vertices [139], has to be applied. Additionally, to obtain the number of expected mistagged W+LF events $\hat{\nu}_{W+\mathrm{LF}}$, one has to further correct for QCD and heavy-flavor contributions to the pretag sideband:

$$\hat{\nu}_{W+\mathrm{LF}} = \hat{\nu}^{\mathrm{pre}}_{\mathrm{tag}-} \cdot \gamma \cdot \frac{\hat{\nu}^{\mathrm{pre}}_{W+\mathrm{jets}} - \hat{\nu}^{\mathrm{pre}}_{W+\mathrm{HF}}}{\hat{\nu}^{\mathrm{pre}}} \tag{5.7}$$

Finally, to estimate the number of expected b-tagged QCD-multijet events, template fits are performed to the $\not\!\!E_{\mathrm{T}}$ spectrum of the secondary vertex tagged data sample. The $\not\!\!E_{\mathrm{T}}$ distribution of the full background prediction, with proper normalization and all background processes included except the QCD fraction, is used. To obtain the QCD template in the b-tagged sample the pretag QCD template is weighted by a b-tagging transfer function depending on $\not\!\!E_{\mathrm{T}}$, which is extracted from CEM trigger data.

The results of the candidate event yield estimate are summarized in table 5.1, all quoted uncertainties include the systematic uncertainties on the theoretical cross-section calculations, the trigger efficiency, the lepton ID efficiency, the b-tagging correction factor, the positive mistag rate, the heavy-flavor scale factor, the QCD-multijet estimate, and the luminosity, where appropriate.

Figure 5.8 compares the final result of the candidate event yield estimate in terms of jet multiplicity with the number of observed collision events. Even though events with one or four jets are not considered for the search of single top-quark production, the outcome in those data sidebands is shown. The uncertainty corresponds to the overall sum of all contributing processes derived in the Method II procedure. The signal expectation is much smaller than the overall uncertainty on all background processes, making a simple counting experiment impossible.

Process	Number of Predicted Events in 3.2 fb^{-1}	
	W + 2 jets	W + 3 jets
t-channel Single Top	87.6 ± 13.0	26.2 ± 3.9
s-channel Single Top	58.1 ± 8.4	19.2 ± 2.8
$t\bar{t}$	204.1 ± 29.6	482.0 ± 69.8
$Wb\bar{b}$	656.9 ± 198.0	201.3 ± 60.8
$Wc\bar{c}$	292.2 ± 90.1	98.1 ± 30.2
Wc	250.4 ± 77.2	52.1 ± 16.0
$Wq\bar{q}$	501.3 ± 69.6	151.9 ± 21.4
WW	58.5 ± 6.6	21.2 ± 2.4
WZ	28.9 ± 2.4	8.5 ± 0.7
ZZ	0.9 ± 0.1	0.4 ± 0.0
Z+jets	36.5 ± 5.6	15.6 ± 2.4
QCD Multijets	89.6 ± 35.8	35.1 ± 14.0
Total Single Top	145.7 ± 21.4	45.4 ± 6.7
Total Prediction	2265.0 ± 375.4	1111.5 ± 129.5
Observed in Data	2229	1086

Table 5.1: Summary of predicted numbers of signal and background events in the selected W+jets data sample

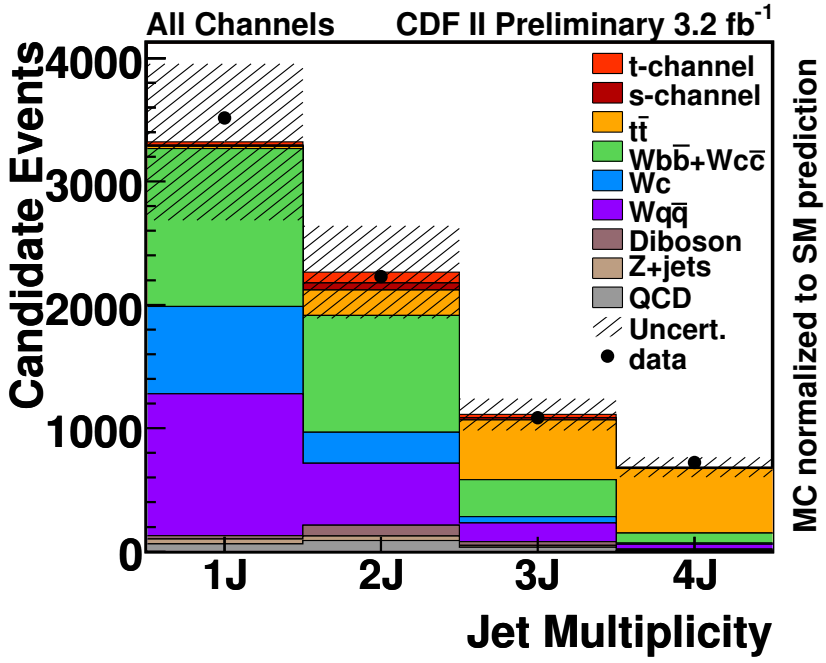


Figure 5.8: Jet multiplicity of the candidate event yield estimate compared to the number of observed events in collision data

Chapter 6

Candidate Event Classification

Since the number of expected single top-quark signal events are not only smaller than the predicted background events, but even smaller than the uncertainty on the latter, a conventional counting experiment analysis is not feasible. A different approach is to exploit kinematic differences of signal and background events by a pattern recognition algorithm to discriminate between signal and background processes. Thereby, the aim is to find a region of phase space where the signal fraction is highly enhanced such that the background uncertainty (even if it remains at the same relative level) does not shadow the signal anymore. In this thesis, artificial neural networks constructed with and trained by the NeuroBayes® package, are utilized [31, 32].

6.1 Technique of Neural Network Event Classifier

The output of a well trained and validated neural network (NN) can be used to create template distributions of the considered signal and background processes. Those templates are to be fitted to the output distribution of the observed events to extract the single top-quark signal significance and the corresponding production cross-section.

6.1.1 Neural Network Event Classifier

NeuroBayes® combines a three-layer feed-forward neural network with a complex robust preprocessing into a powerful event classifier. The 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

$$a_j(\mathbf{x}) = \sum_i \omega_{ij} x_i + \mu_{0,j} \tag{6.1}$$

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

$$S(\mathbf{x}) = \frac{2}{1 + e^{-a(\mathbf{x})}} - 1 \tag{6.2}$$

which gives an output of -1 for background and $+1$ for signal. As can be seen in figure 6.1, the sigmoid function is only sensitive to a relatively small range around zero. By this transformation, the interval $[-\infty, +\infty]$ is mapped to the interval $[-1, +1]$. For very large ($x \to \infty$) or very small ($x \to -\infty$) values, a saturation effect is reached. The bias 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.

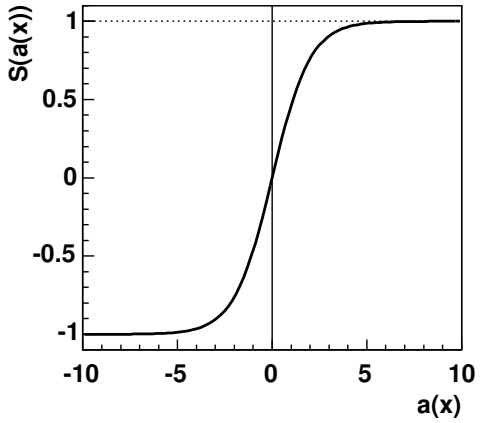


Figure 6.1: The transformed sigmoid activation function S(a(x)) as given by equation 6.2.

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

$$o = S\left(\sum_{j=0}^{M} \omega_j \cdot S\left(\sum_{i=0}^{d} \omega_{ij} x_i + \mu_{0,j}\right)\right) \tag{6.3}$$

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.

6.1.2 Training of Neural Network Event Classifier

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 the implemented neural networks is the entropy error function

$$E_D = \sum_i \log\left(\frac{1}{2}\cdot(1 + T_i \cdot o_i + \epsilon)\right), \tag{6.4}$$

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 (6.3). 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 back-propagation 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 the search of electroweak single top-quark production, six NN classifiers are trained. Three of them are optimized for events with two jets and one b tag, one for two-jet two-tag events, and a fifth and sixth one in the three-jet bin, for both tagging categories. The sample of events with two jets and one tag is the largest subsample and dominates the search for single top-quark production. Four of the six trained neural networks, one in each jet and tag category, are applied for the search for combined s- and t-channel signal events. The combined search is based on the assumption that the ratio of both signal process rates will occur as predicted by the SM. For the separate search, the remaining two networks, each trained for s- and t-channel, respectively, are simultaneously applied in the dominant two-jet one-tag sample. In conjunction with the neural networks in the other categories, this allows for a distinct separation of the two single top-quark production modes.

For validation purposes, four additional control neural networks are trained. Two of them are trained for t-channel single top-quark classification in the kinematically similar sideband region of SecVtx untagged events with two and three jets, respectively. Another neural network is trained in the control sample of events with only one b-tagged jet, but this time to classify $Wb\bar{b}$ events. Above mentioned sideband regions are expected to be dominated by W+LF and W+HF production, respectively. On the contrary, the control region with three jets and with one b tag of the remaining fourth validation classifier trained for $t\bar{t}$ production is expected to be populated equally by W+jets and $t\bar{t}$ events.
Furthermore, the analysis of the search for single top-quarks produced via flavor-changing neutral currents (FCNC) also used NeuroBayes® neural networks [38, 39]. They were trained to separate SM events from SM suppressed FCNC events in the candidate event signal sample with one b-tagged jet as well as in its high statistic control sample of events with one untagged jet. All twelve trained neural network classifiers are listed in table 6.1.

1 Jet		2 Jets			3 Jets		
0 Tag	1 Tag	0 Tag	1 Tag	2 Tags	0 Tag	1 Tag	2 Tags
FCNC single top-quark validation	FCNC single top-quark search	t-channel combined signal validation	**t-channel combined signal search**	**s-channel combined signal search**	t-channel combined signal validation	**t-channel combined signal search**	**t-channel combined signal search**
	$Wb\bar{b}$ signal validation		**t-channel separate signal search**			$t\bar{t}$ signal validation	
			s-channel separate signal search				

Table 6.1: Overview of all trained NN classifier for the combined and separate single top-quark signal extractions (**bold**) and for validation purposes, respectively

6.1.3 Training Event Sample Composition

For the training of the 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 due to the small number of predicted signal events, a different composition of the training samples has to be chosen.

All training samples used are composed in such a way that the respective signal process contributes 50% to the total number of events. The relative fractions of all considered background processes are given by the respective number of expected events derived in chapter 5, except for QCD events, whose event properties and kinematics are hard to model. Therefore, it's more conservative, to not train against the available QCD model.
For the separate search one further needs rather good separation between s- and t-channel single top-quark events. Therefore, the training of those networks are special is the way, that the s-channel single top-quark sample is used as background sample for the training of the t-channel network and vice versa. Tables 6.2 and 6.3 list the training sample composition for the combined search networks and separate search networks, respectively.

Category	2 jets 1 tag	2 jets 2 tags	3 jets 1 tag	3 jets 2 tags
t-channel	50.0%	—	50.0%	50.0%
s-channel	—	50.0%	—	—
$t\bar{t}$	5.1%	15.9%	22.0%	37.0%
$Wb\bar{b}$	13.7%	27.2%	7.9%	11.2%
$Wc\bar{c}/Wc$	14.0%	3.2%	7.8%	1.8%
$Wq\bar{q}$	14.0%	—	10.0%	—
Diboson	2.4%	2.2%	1.6%	—
Z+jets	0.8%	0.6%	0.7%	—

Table 6.2: Composition of the training samples used to train the neural networks for the combined search.

Category	2 jets 1 tag		2 jets 2 tags	3 jets 1 tag	3 jets 2 tags
t-channel	50.0%	9.6%	—	50.0%	50.0%
s-channel	5.1%	50.0%	50.0%	—	—
$t\bar{t}$	4.6%	4.1%	15.9%	22.0%	37.0%
$Wb\bar{b}$	12.3%	11.1%	27.2%	7.9%	11.2%
$Wc\bar{c}/Wc$	12.6%	12.3%	3.2%	7.8%	1.8%
$Wq\bar{q}$	12.5%	11.3%	—	10.0%	—
Diboson	2.2%	1.9%	2.2%	1.6%	—
Z+jets	0.7%	0.7%	0.6%	0.7%	—

Table 6.3: Composition of the training samples used to train the neural networks for the separate search.

6.2 Definition of Input 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 6.1) and assures that also the inputs to the next layers are distributed with mean zero and width one.
To decorrelate the preprocessed input variables, their covariance matrix is calculated. Diagonalizing the covariance matrix and dividing the rotated input vectors by the square root of the corresponding eigenvalue transforms the covariance matrix into a unit matrix. The transformation to a Gaussian distribution may be altered by individual variable preprocessing like fitting a spline curve to the flattened distribution to handle statistical fluctuations. 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.
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, 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.

For all neural network trainings, three categories of input variables were considered: some are directly measured in the detector, others are reconstructed out of measured values, and a few are calculated by advanced algorithms. Finally, more then 50 input variables were investigated; hereof only those variables were used which contribute to a significant amount to the discriminating power of the neural network. This reduces the number of training variables to 11 to 18 depending on the jet and tag category. The number of hidden nodes used for the network is found to be arbitrary in terms of discriminating power and is chosen to be 15 for all considered networks.

6.2.1 KIT Flavor Separator

The KIT flavor separator is a neural network based tool to improve the purity of the b-quark identification in high-p_{T} jets [28,30]. It gives an additional handle to reduce the large background components where no real b-quarks are contained, namely $Wq\bar{q}$ and c quark backgrounds. Both of them amount to about 50% in the dominant two jet data sample even after imposing the requirement that one jet is SecVtx tagged. The SecVtx algorithm merely provides a binary decision whether a jet contains a reconstructable secondary vertex. Furthermore, this quantity is based only on a few requirements applied to track information resulting from the long B-hadron lifetime. The KIT flavor separator is a neural network trained on simulated SecVtx tagged jet samples. It exploits many different informations as inputs, which are combined by the network. Besides the lifetime based informations, like the impact parameter d_0, and the decay length significance $\frac{L_{xy}}{\sigma_{xy}}$, the KIT flavor separator makes use of further variables like the reconstructed SecVtx vertex mass and its decay multiplicity. Finally, the KIT flavor separator output can be treated as a measure for the probability of a true b quark being present within the SecVtx tagged jet. The output of the KIT flavor separator is shown in figure 6.2. For jets containing a b quark, the output of the network accumulates at $+1$, whereas jets without any heavy quark produce an output close to -1. It is also possible to distinguish jets with c but no b quarks, their output distribution lies in-between the b respectively light jets.

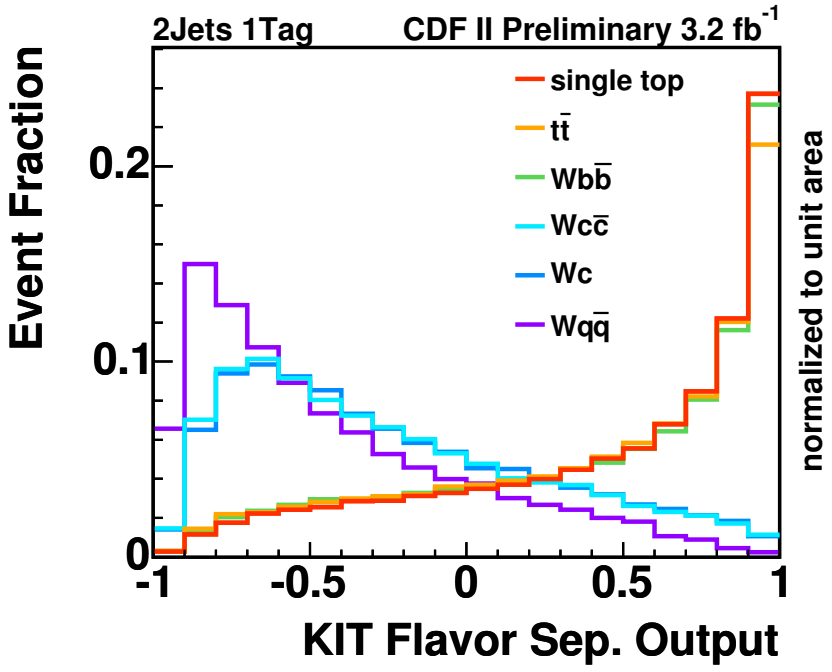


Figure 6.2: The output of the KIT flavor separator can distinguish between processes with real b-quark content (single top, $t\bar{t}$, and $Wb\bar{b}$), c-quark content ($Wc\bar{c}$ and Wc), and light-quark content ($Wq\bar{q}$).

6.2.2 Top-Quark Reconstruction

For some of the variables, the reconstruction of the top quark is necessary. The top-quark four-momentum is built out of the reconstructed W boson and the b-tagged jet. The W boson is reconstructed from the measured tight lepton and the reconstructed neutrino. The transverse momentum of the neutrino is derived from $\vec{\not{E}}_{\mathrm{T}}$. The z component of the neutrino momentum is obtained using a quadratic constraint of the W-boson decay kinematics with the W-boson mass constrained to 80.4 GeV/c^2 [66]. There are two in general complex solutions for the z component. In case the solutions are real, the solution with the smaller absolute value is chosen. In case of complex solutions, a kinematic fit allowing for varying the x and y components of $\vec{\not{E}}_{\mathrm{T}}$ is performed to find a solution as close as possible to $\vec{\not{E}}_{\mathrm{T}}$ [39].
If there is more than one jet with a reconstructed secondary vertex, the tagged jet with the largest product of the charge Q_ℓ of the tight lepton and the jet pseudo-rapidity, $Q_\ell \cdot \eta$, is assigned to belong to the top-quark decay. At the top-quark reconstruction level, additional jet corrections are applied. The calorimeter energy is corrected to the hadron level, the energy of the underlying event is subtracted, and the energy radiated outside the jet cone is added.

6.2.3 Input Variables

In the following, the input variables of all NN classifiers applied in the search for single top-quark production are listed in descending order of their significance.

Input Variables for the Two-Jet One-Tag Category

14 input variables are used to distinguish t-channel signal from background events for the combined search as well as the separate search for single top-quark production:

- M (lνb), the reconstructed top-quark mass is built out of the charged lepton, the reconstructed neutrino, and, in case of the one-tag category, the b-tagged jet. As illustrated in figure 6.3(a), the invariant mass between these three objects peaks around 170 GeV/c^2 for top-quark processes, while all remaining background processes accumulate at much smaller values.

- KIT Flavor Sep. Output, the neural network output of the KIT flavor separator of the b-tagged jet is able to discriminate processes with underlying B hadrons from charm processes and from light-quark flavor processes with mis-reconstructed secondary vertices, see figure 6.3(d).

- M (jet1,jet2), the invariant mass of the two jets reflects the characteristics of the underlying object, if both jets originate from the same one. As visible in figure 6.3(f), the diboson processes peak at the value of the W and Z boson, as expected from two jets coming from an hadronically decaying boson. For

W+jets events, the invariant mass is much lower in agreement with two jets stemming from a gluon. Since both jets are expected to come from different objects in single top-quark and $t\bar{t}$ events, the corresponding distribution is much harder.

- Q (lep) $\cdot\, \eta$ (l-jet), the product of the charge of the lepton and the pseudorapidity of the light-quark jet is very distinct for single top-quark events [23,24]. As can be seen in figure 6.3(i), this variable is very asymmetric for signal events, while it is central for all background processes. This asymmetry, which only occurs for t-channel signal events, is induced by the parton distribution function of the proton. A top-quark is mostly produced by an initial state u-quark stemming from the proton, since the latter consists of two u and only one d valence quark. The initial $\bar{d}$ quark originating from the antiproton makes a smaller contribution to the production of top quarks. Since the light-quark jet in t-channel events stems from the W-boson exchange of the initial valence quark, while the initial b quark is a sea quark carrying a significantly smaller fraction of the proton momentum, the light-quark jet has a strong tendency of propagating in the direction of the valence quark. Hence, most of the light-quark jets in top-quark production propagate in proton direction, which is equivalent to positive pseudorapidity. Since the top-quark decays into b quark and a W^+ boson, which subsequently further decays into a neutrino and a positively charged lepton, the charge of the lepton Q (lep) distinguishes between top- and antitop-quark production. Thus, t-channel top-quark events will accumulate in the positive forward regime of the product Q (lep) $\cdot\, \eta$ (l-jet). The charge conjugate case applies for antitop-quark production.

- $\mathrm{M_T}$ (lνb), the transverse mass of the reconstructed top quark, see figure 6.4(a), is obtained without any information of the uncertain z components of the considered objects:

$$\begin{aligned} \mathrm{M_T}\ (\mathrm{l}\nu\mathrm{b}) = \Big(& \left(p_\mathrm{T}\,(\mathrm{l}) + p_\mathrm{T}\,(\nu) + p_\mathrm{T}\,(\mathrm{b})\right)^2 \\ & - \left(p_x\,(\mathrm{l}) + p_x\,(\nu) + p_x\,(\mathrm{b})\right)^2 - \left(p_y\,(\mathrm{l}) + p_y\,(\nu) + p_y\,(\mathrm{b})\right)^2 \Big)^{\frac{1}{2}} \end{aligned} \tag{6.5}$$

- $\cos\Theta\ (\mathrm{lep,l\text{-}jet})_{\mathrm{l}\nu\mathrm{b\ r.f.}}$, the cosine of the angle in the top-quark rest frame between the charged lepton and the light-quark jet is sensitive to the t-channel spin polarization of the top-quark, as already discussed in chapter 1.2.2. This variable separates the signal from almost all background processes, as shown in figure 6.4(d).

- $\mathrm{E_T}$ (l-jet), the transverse energy of the light-quark jet, which is much harder for top-quark processes, is illustrated in figure 6.4(g).

- $\cos\Theta\ (\mathrm{lep}_{\mathrm{l}\nu\ \mathrm{r.f.}}, W)$, the cosine of the polar angle between the charged lepton in the W-boson rest frame and the direction of the W boson is one of the few variables which separate Wc production from the remaining W+jets processes, as seen in figure 6.4(j).

- η (lν), the pseudorapidity of the reconstructed W boson is depicted in figure 6.5(a).
- M_T (lν), the transverse mass of the reconstructed W boson, see figure 6.5(d), is already introduced in chapter 5.1.4.
- η (jet1,jet2), the scalar sum of the pseudorapidities of the two jets is shown in figure 6.5(g).
- P_T (lep), the transverse momentum of the charged lepton is rather soft for QCD-multijet events and Z+jets events, as seen in figure 6.5(j).
- H_T, the scalar sum of transverse energies of the two jets, the transverse momentum of the charged lepton, and the transverse energy of the reconstructed neutrino distinguishes between $t\bar{t}$ production and all remaining background processes, whereas the signal is somewhere in-between, see figure 6.6(a).
- $\cos\Theta^*$ (lep,W), the cosine of the polar angle between the charged lepton in the W-boson rest frame and the reconstructed W boson in the top-quark rest frame, shown in figure 6.6(d), is a variable sensitive to the helicity of the W boson in top-quark events.

17 input variables are used to distinguish s-channel signal from background events for the separate search for single top-quark production:

- KIT Flavor Sep. Output, the neural network output of the KIT flavor separator of the selected b-quark jet from the top-quark decay.
- M (lνb), the reconstructed top-quark mass.
- H_T, the scalar sum of transverse energies of the two most energetic jets, the transverse momentum of the charged lepton, and the transverse energy of the reconstructed neutrino.
- $\cos\Theta$ $(\text{jet1}, \text{jet2})_{\text{l}\nu\text{b r.f.}}$, the cosine of the angle in the top-quark rest frame between the two jets.
- $\cos\Theta$ $(\text{lep}_{\text{l}\nu\text{ r.f.}}, W)$, the cosine of the polar angle between the charged lepton in the W-boson rest frame and the direction of the W boson.
- E_T (l-jet), the transverse energy of the light-quark jet.
- M_T (lν), the transverse mass of the reconstructed W boson.
- M_T (lνb), the transverse mass of the reconstructed top quark.
- η (lν), the pseudorapidity of the reconstructed W boson.
- M (jet1,jet2), the invariant mass of the two most energetic jets.

- P_T (lep), the transverse momentum of the charged lepton.
- $\cos\Theta$ (lν, lνb), the cosine of the angle between the W boson and the direction of the top quark.
- $\cos\Theta$ (lν, lνb)$_{\text{l}\nu\text{b r.f.}}$, the cosine of the angle between the W boson in the top-quark rest frame and the direction of the top quark.
- Q (lep) $\cdot$ η (l-jet), the product of the charge of the lepton and the pseudorapidity of the light-quark jet.
- $\cos\Theta^*$ (lep, lν), the cosine of the angle between the charged lepton in the W-boson rest frame and the direction of the W boson in the top-quark rest frame.
- $\cos\Theta$ (lep, beam), the cosine of the angle between the charged lepton and the direction of the beam pipe.
- η (l), the pseudorapidity of the charged lepton.

Input Variables for the Two-Jet Two-Tag Category

The two-jet two-tag neural network is trained with 11 input variables to distinguish s-channel signal from background events for the combined search for single top-quark production:

- M (lνjet1jet2), the invariant mass of the system composed by the reconstructed W boson and the two most energetic jets.
- M_T (lν), the transverse mass of the reconstructed W boson.
- M_T (lνb), the transverse mass of the reconstructed top quark.
- $\cos\Theta$ (jet1, jet2)$_{\text{l}\nu\text{b r.f.}}$, the cosine of the angle in the top-quark rest frame between the two most energetic jets.
- M (lνb), the reconstructed top-quark mass.
- KIT flavor sep. Output (jet1+jet2), the sum of the neural-network output of the KIT flavor separator of both the two most energetic jets.
- η (lν), the pseudorapidity of the reconstructed W boson.
- $\not{E}_T$, the transverse energy of the reconstructed neutrino.
- M (jet1,jet2), the invariant mass of the two most energetic jets.
- E_T (b-jet), the transverse energy of the b-quark jet from the top-quark decay.
- η (l), the pseudorapidity of the charged lepton.

Input Variables for the Three-Jet One-Tag Category

The three-jet one-tag neural network is trained with 18 input variables to distinguish t-channel signal from background events for the combined search for single top-quark production:

- Q (lep) $\cdot$ η (l-jet), the product of the charge of the lepton and the pseudorapidity of the most energetic light-quark jet.
- KIT flavor sep. Output, the neural-network output of the KIT flavor separator of the selected b-quark jet from the top-quark decay.
- $\mathrm{H_T}$, the scalar sum of transverse energies of the three most energetic jets, the transverse momentum of the charged lepton, and the transverse energy of the reconstructed neutrino.
- M (jet1,jet3), the invariant mass of the most and third most energetic jet.
- M (lνb), the reconstructed top-quark mass.
- $\mathrm{P_T}$ (lνbjet1jet2), the transverse momentum of the system composed by the reconstructed W boson, the b-quark jet from the top-quark decay, and the remaining two most energetic jets.
- M (jet2,jet3), the invariant mass of the second and third most energetic jet.
- $\cos\Theta$ $(\mathrm{lep,l\text{-}jet})_{\mathrm{l\nu b\ r.f.}}$, the cosine of the angle in the top-quark rest frame between the charged lepton and the most energetic light-quark jet.
- $\Delta\eta$ (jet1, jet2), the difference in pseudorapidity between the two most energetic jets.
- η (jet1, jet2, jet3), the scalar sum of the pseudorapidities of the three most energetic jets.
- $\not\!\!E_{\mathrm{T,sig}}$, the significance of the transverse energy of the reconstructed neutrino, is introduced in chapter 5.1.4.
- M (jet1,jet2,jet3), the invariant mass of the three most energetic jets.
- $\mathrm{E_T}$ (jet2,jet3), the transverse energy of the second and third most energetic jet.
- $\mathrm{E_T}$ (b-jet), the transverse energy of the b-quark jet from the top-quark decay.
- $\mathrm{M_T}$ (lνb), the transverse mass of the reconstructed top quark.
- $\Delta\eta$ (lνb, l-jet), the difference in pseudorapidity between the reconstructed top quark and the most energetic light-quark jet.
- $\mathrm{E_T}$ (jet1,jet3), the transverse energy of the first and third most energetic jet.
- $\mathrm{E_T}$ (jet1,jet2), the transverse energy of the two most energetic jets.

Input Variables for the Three-Jet Two-Tag Category

The three-jet two-tag neural network is trained with 15 input variables to distinguish t-channel signal from background events for the combined search:

- Q (lep) $\cdot$ η (l-jet), the product of the charge of the lepton and the pseudorapidity of the light-quark jet.
- $\mathrm{P_T}$ (lνbjet1jet2), the transverse momentum of the system composed by the reconstructed top quark and the remaining two most energetic jets.
- M (jet1,jet2), the invariant mass of the two most energetic jets.
- $\cos\Theta$ (lep,l-jet)$_{\mathrm{l}\nu\mathrm{b\ r.f.}}$, the cosine of the angle in the top-quark rest frame between the charged lepton and the light-quark jet.
- $\mathrm{E_T}$ (jet1,jet2), the transverse energy of the two most energetic jets.
- M (jet1,jet3), the invariant mass of the most and third most energetic jet.
- $\Delta\eta$ (jet2, jet3), the difference in pseudorapidity between the second and third most energetic jet.
- $\mathrm{E_T}$ (b-jet2), the transverse energy of the jet not from the top-quark decay.
- $\mathrm{M_T}$ (lνb), the transverse mass of the reconstructed top quark.
- $\Delta\eta$ (jet1, jet2), the difference in pseudorapidity between the two leading jets.
- $\mathrm{E_T}$ (jet3), the transverse energy of the third most energetic jet.
- Centrality, the centrality of the event defined as the sum of the transverse energy of the two most energetic jets divided by $\sqrt{\hat{s}}$.
- $\sqrt{\hat{s}}$, the energy in the s-channel center-of-mass system.
- $\cos\Theta$ (jet1, jet2)$_{\mathrm{l}\nu\mathrm{b\ r.f.}}$, the cosine of the angle in the top-quark rest frame between the two most energetic jets.

Since the neural networks are 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 the signal and background models described in chapter 4. 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 candidate event yield estimate in chapter 5. A reasonable agreement between the modeling and the observation in the signal regions as well as in background dominated sidebands is found. Figures 6.3-6.6 show the distributions of the network input variable for TLC events with two jets and one b tag. The corresponding distribution of EMC events, of the s-channel network in the 2-jet bin with 2 b tags, the t-channel network in the 3-jet bin with 1 and 2 b tags show comparable good agreement.

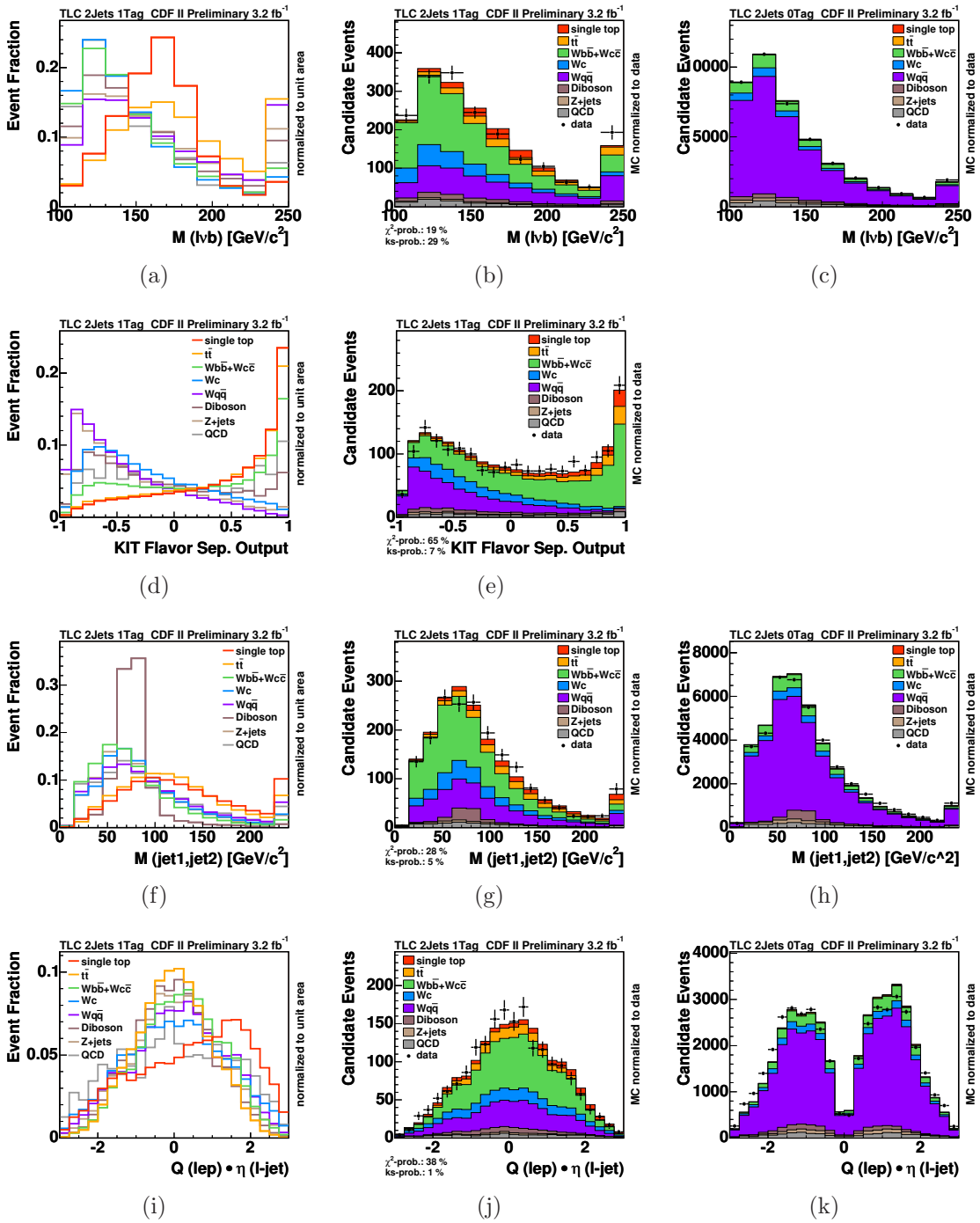


Figure 6.3: Shape comparison (first column) and MC modeling validation (second column) in the two-jet one-tag signal region, and MC modeling validation (third column) in the untagged two-jet sideband of the discriminating input-variables for TLC events.

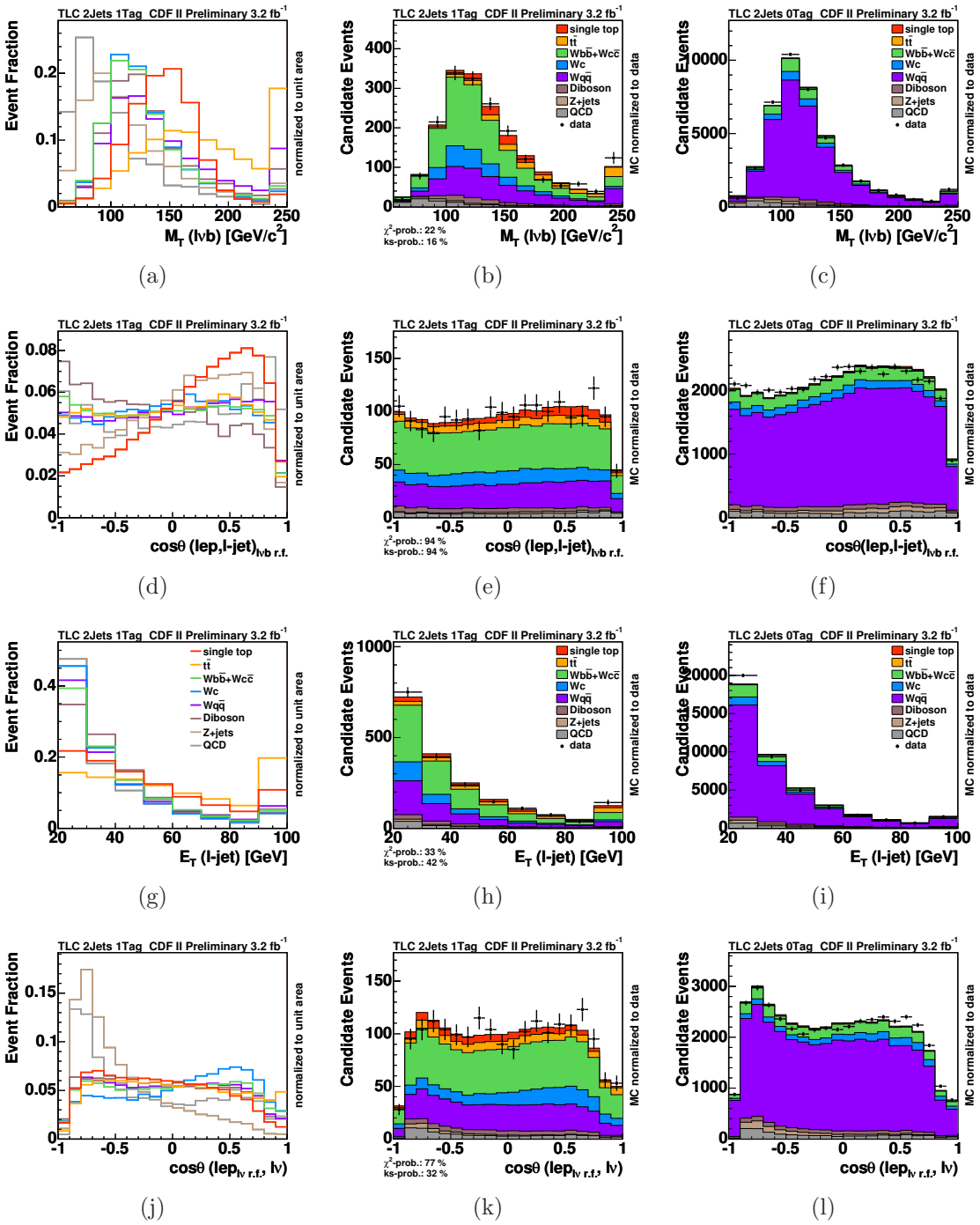


Figure 6.4: Shape comparison (first column) and MC modeling validation (second column) in the two-jet one-tag signal region, and MC modeling validation (third column) in the untagged two-jet sideband of the discriminating input-variables for TLC events.

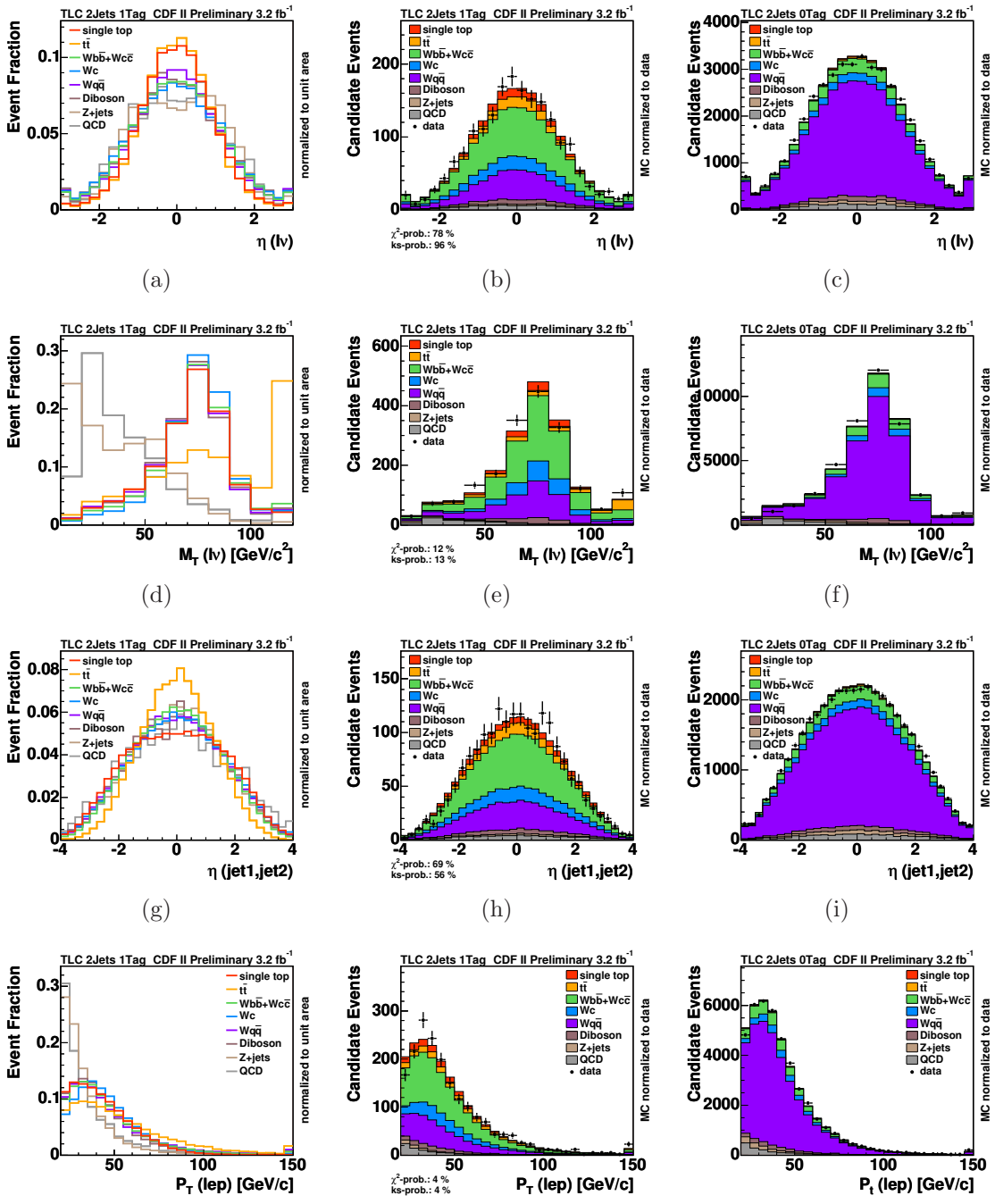


Figure 6.5: Shape comparison (first column) and MC modeling validation (second column) in the two-jet one-tag signal region, and MC modeling validation (third column) in the untagged two-jet sideband of the discriminating input-variables for TLC events.

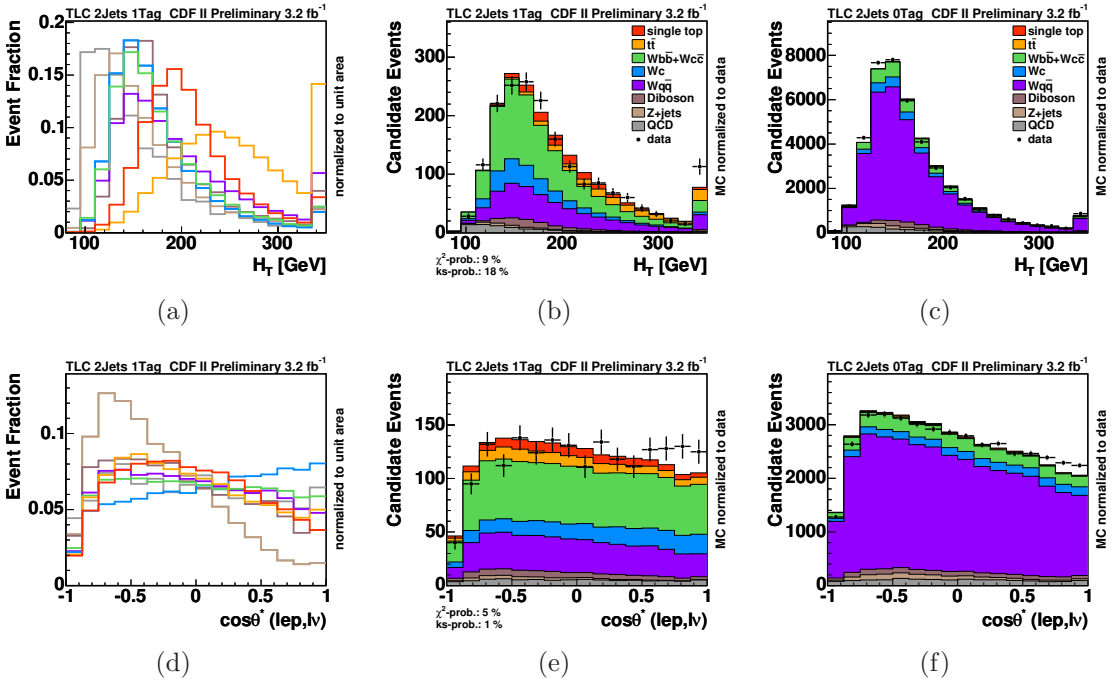


Figure 6.6: Shape comparison (first column) and MC modeling validation (second column) in the two-jet one-tag signal region, and MC modeling validation (third column) in the untagged two-jet sideband of the discriminating input-variables for TLC events.

Additionally, since neural networks use the correlation between input variables to distinguish signal from background events, the modeling of the corresponding correlations are checked. This is done using the correlation variable

$$\kappa_{ij} = \frac{x_i - \bar{x}_i}{\sigma_{x_i}} \cdot \frac{x_j - \bar{x}_j}{\sigma_{x_j}} \tag{6.6}$$

between two variables x_i and x_j, where $\bar{x}_i$ is the event ensemble mean value of variable x_i and σ_{x_i} is the corresponding width. At an early stage of the analysis, κ_{ij}, inspired by the corresponding correlation coefficient, are checked by comparing the modeling with collision data.
It turned out that this check has an additional handle to pinpoint potential mismodeling, leading to the omission of several input variables. Figure 6.7 shows the distributions of the correlation variable between the four most significant input variables of the NN trained for t-channel events with two jets and one b tag.

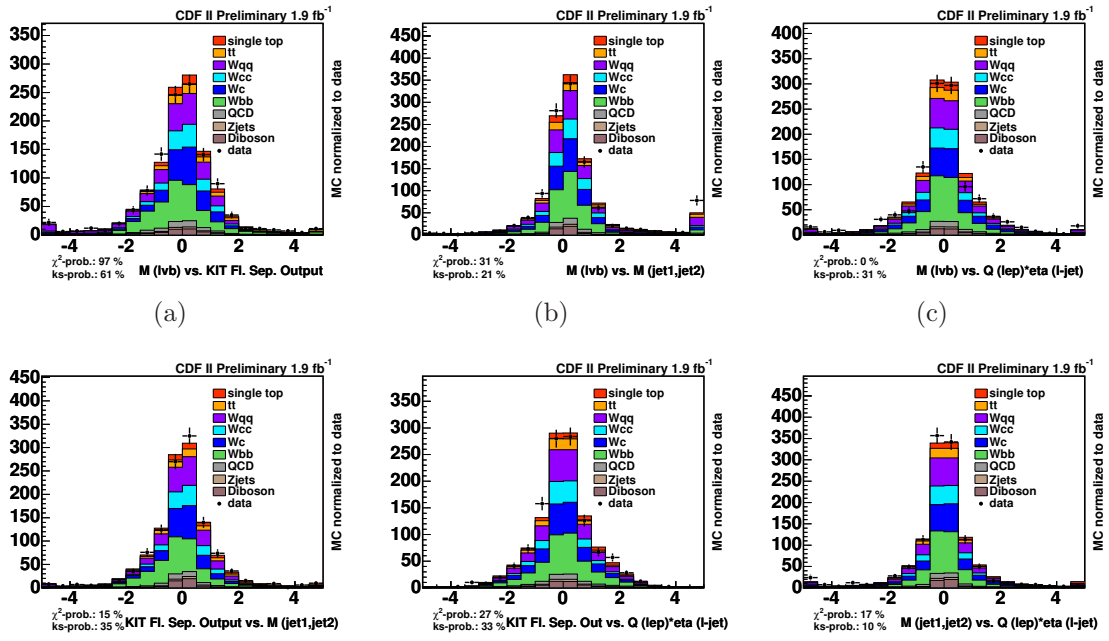


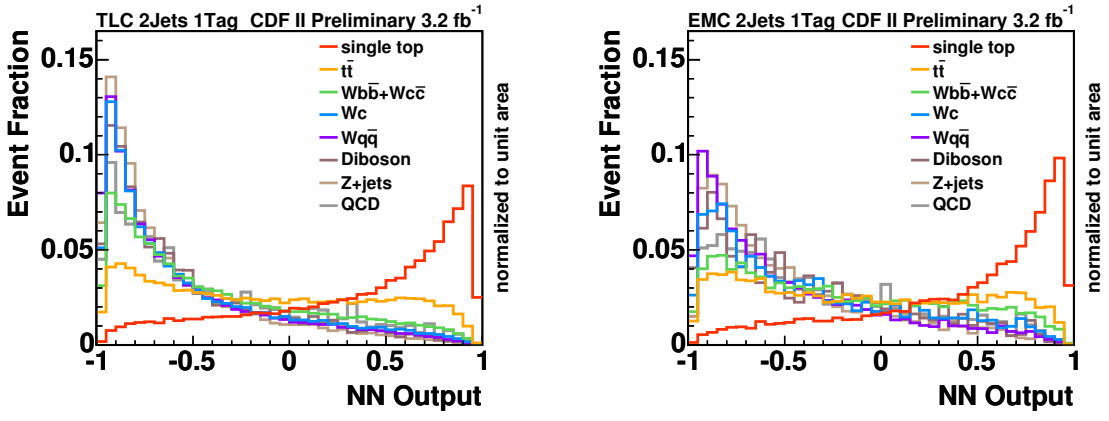
Figure 6.7: Correlation distributions between the first four most significant input variables for the t-channel neural network in the two jet one tag category

6.3 Training Result and Template Construction

The training of a NeuroBayes® neural network for classification results in one output variable continuously distributed between -1, the background like region, and the signal region at 1. The output of the various samples of signal and background models are used to create templates of each process, which finally are to be fitted to the output distribution of the observed events.
To construct the templates, it is relevant to combine background processes whose output distributions look very similar and are hence difficult to distinguish. On this account, some of the processes are merged into one template with a ratio given by the background estimation. An exception is made for the $Wb\bar{b}$ and $Wc\bar{c}$ process. Both are paired together, since their individual rate is expected to be highly correlated due to their similar underlying Feynman diagrams. Finally, seven background templates remain: $t\bar{t}$, $Wb\bar{b} + Wc\bar{c}$, Wc, $Wq\bar{q}$, Diboson, Z+jets, and QCD.
Each of the four jet and tag categories is divided into two separate channels, the Triggered Lepton Coverage (TLC) containing triggered electrons and muons, and the remaining muons from the Extended Muon Coverage (EMC) accepted through the $\not\!\!E_{\mathrm{T}}$ trigger. This separation accommodate the different signal-over-background ratio of the different samples.

6.3.1 Templates for Combined Single Top-Quark Search

The output distributions of both s- and t-channel events are combined into one signal distribution, where the ratio between the two processes is as predicted by the SM. Figures 6.8-6.9 depict the shape of the fit templates made of the neural network outputs of all signal and background events in all eight channels considered in the search for combined s- and t-channel single top-quark production.



(a) (b)

Figure 6.8: Signal and background templates in the two-jet one-tag signal region for TLC (a) and EMC (b) events for the combined single top-quark search

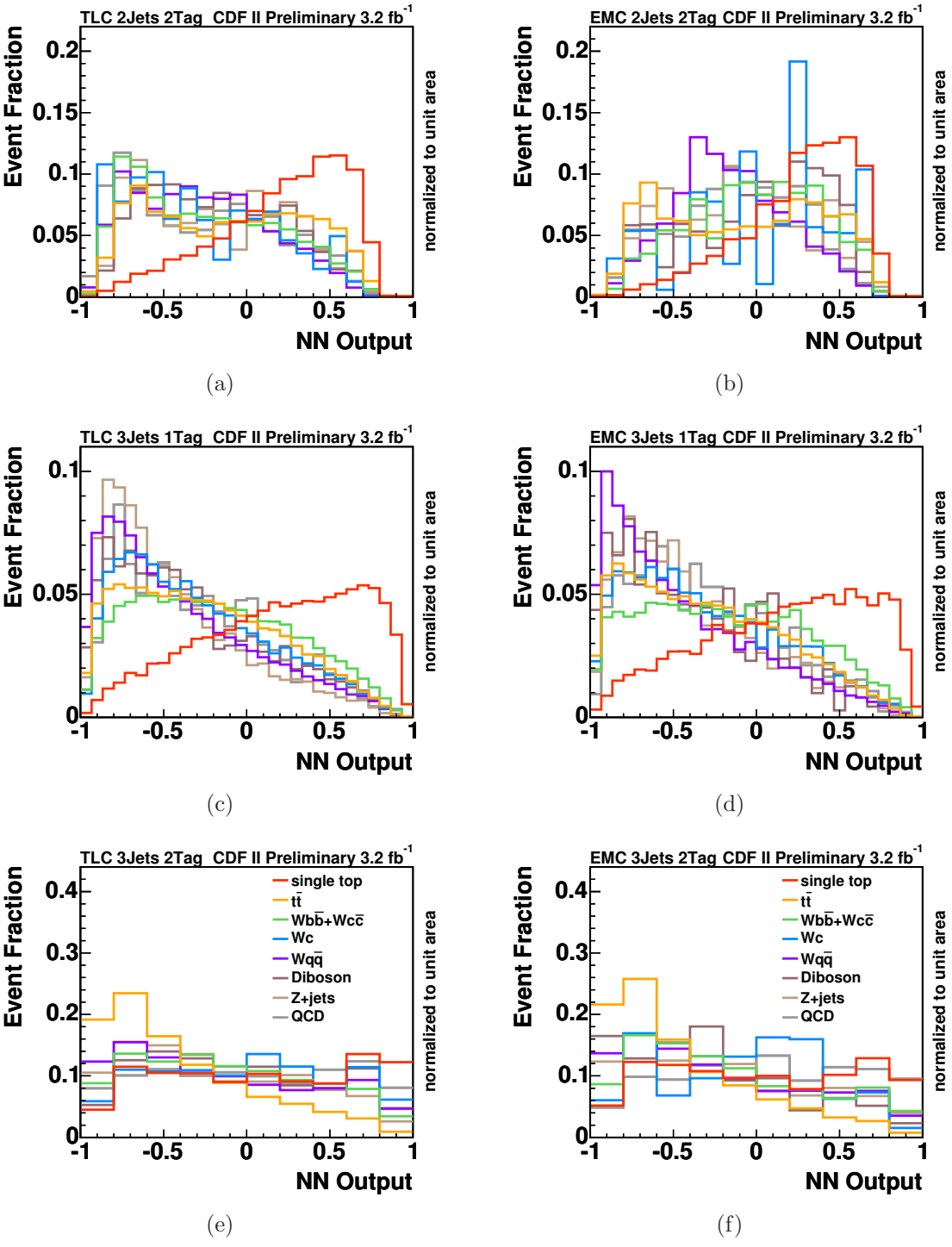


Figure 6.9: Signal and background templates in the two-jet two-tag signal region for TLC (a) and EMC (b) events, in the three-jet one-tag signal region for TLC (c) and EMC (d) events, and in the three-jet two-tag signal region for TLC (e) and EMC (f) events for the combined single top-quark search

6.3.2 Templates for Separate s- and t-channel Search

For the separate search, the combined search network in the two jet one tag category is replaced by two networks optimized for s- and t-channel classification, respectively. Those two neural networks build a 2D discriminant, which improves the performance to separate the two single top-quark production modes. The 2D templates of the separate search in the 2-jet bin with 1 b tag are illustrated in figures 6.10-6.13, showing the output of the s-channel neural network versus the output of the t-channel neural network. For the fit to the observed data the 2D templates get unwinded bin by bin from the top left to the bottom right to obtain 1D distributions. The final templates of the separate search in the 2-jet bin with 1 and 2 b tags and 3-jet bin with 1 and 2 b tags, respectively, are illustrated in figure 6.14 and 6.15.

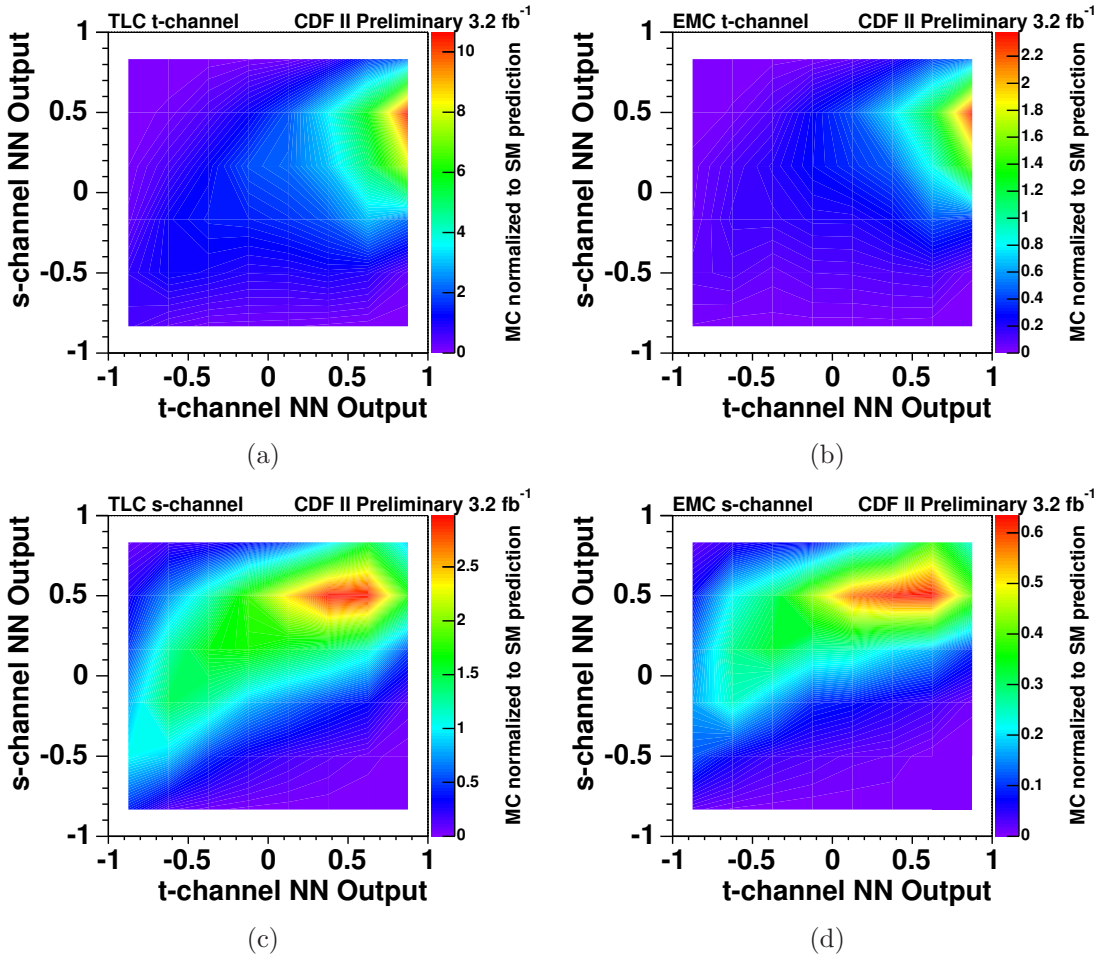


Figure 6.10: The outputs of the s-channel vs. t-channel networks are shown for TLC (left) and EMC (right) t-channel (top) and s-channel (bottom) events.

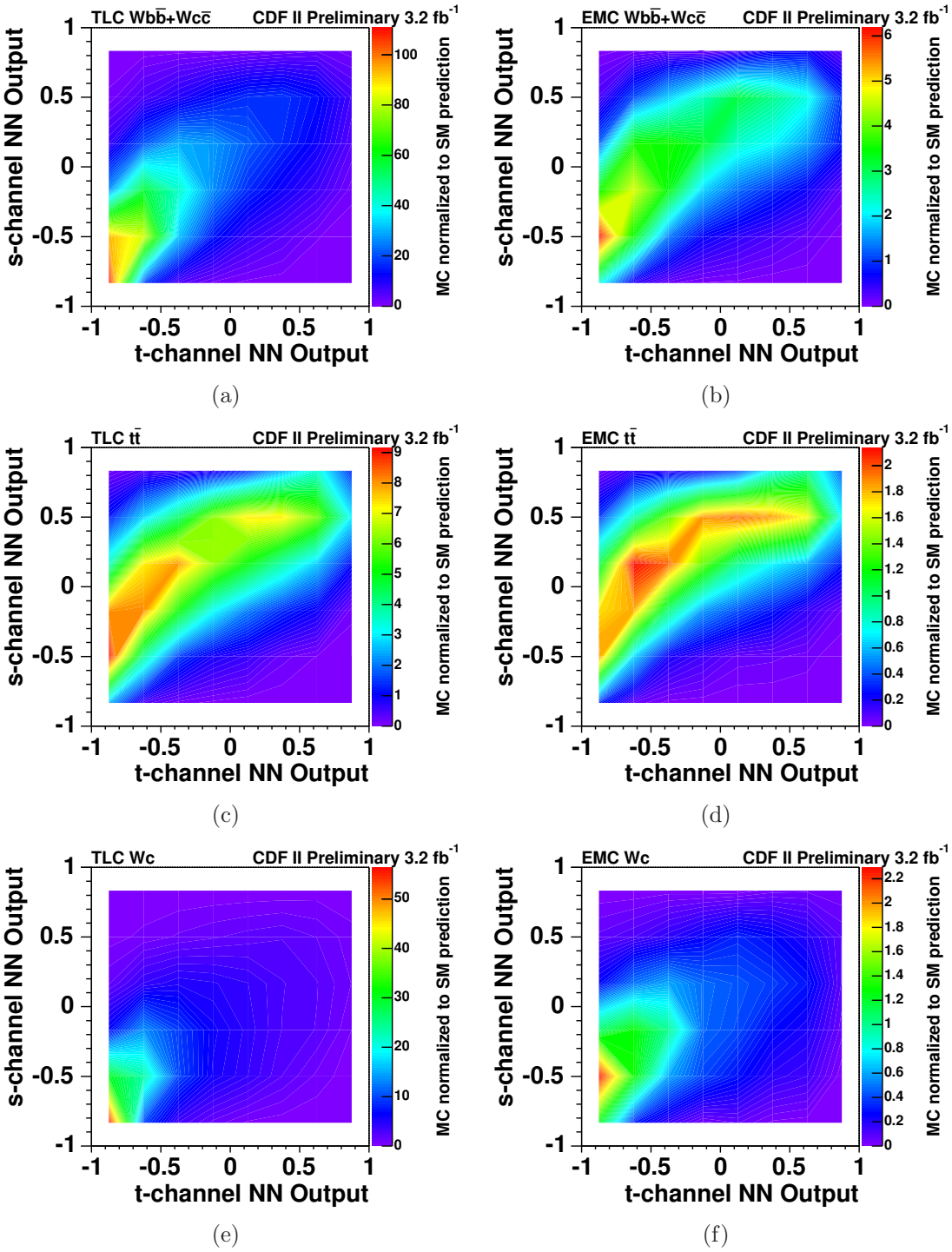


Figure 6.11: The outputs of the s-channel vs. t-channel networks are shown for TLC (left) and EMC (right) $Wb\bar{b} + Wc\bar{c}$ (top), $t\bar{t}$ (middle), and Wc (bottom) events.

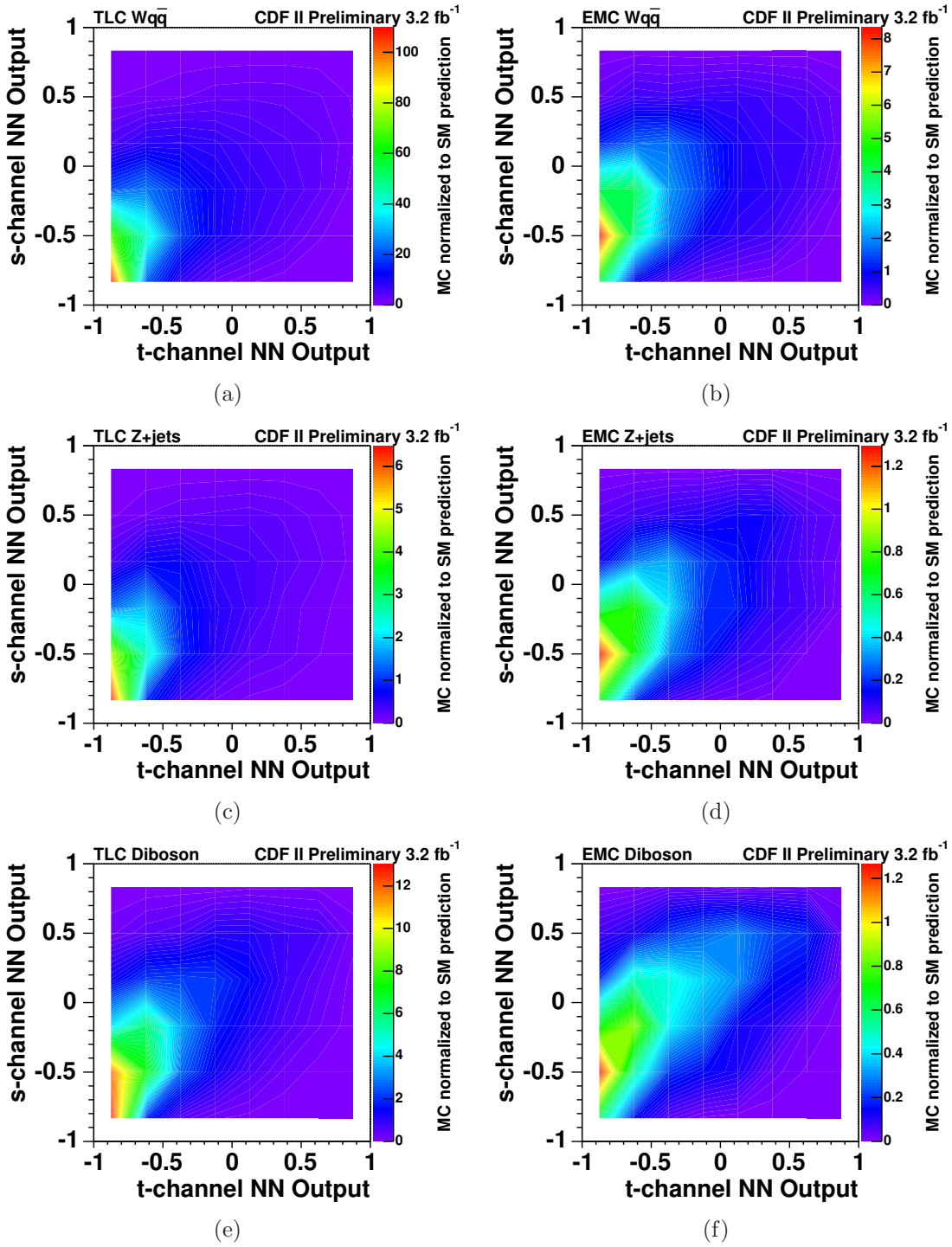
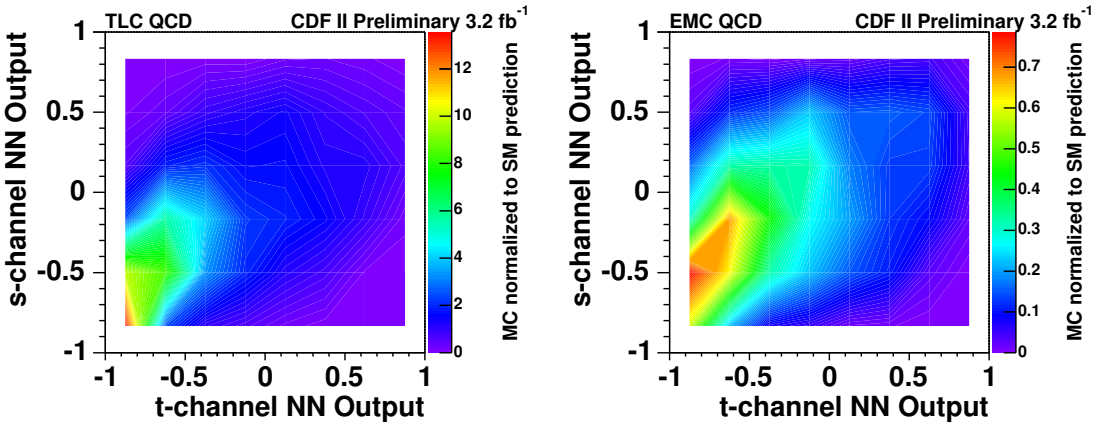


Figure 6.12: The outputs of the s-channel vs. t-channel networks are shown for TLC (left) and EMC (right) $Wq\bar{q}$ (top), Z+jets (middle), and Diboson (bottom) events.



(a) (b)

Figure 6.13: The outputs of the s-channel vs. t-channel networks are shown for TLC (left) and EMC (right) QCD-multijet events.

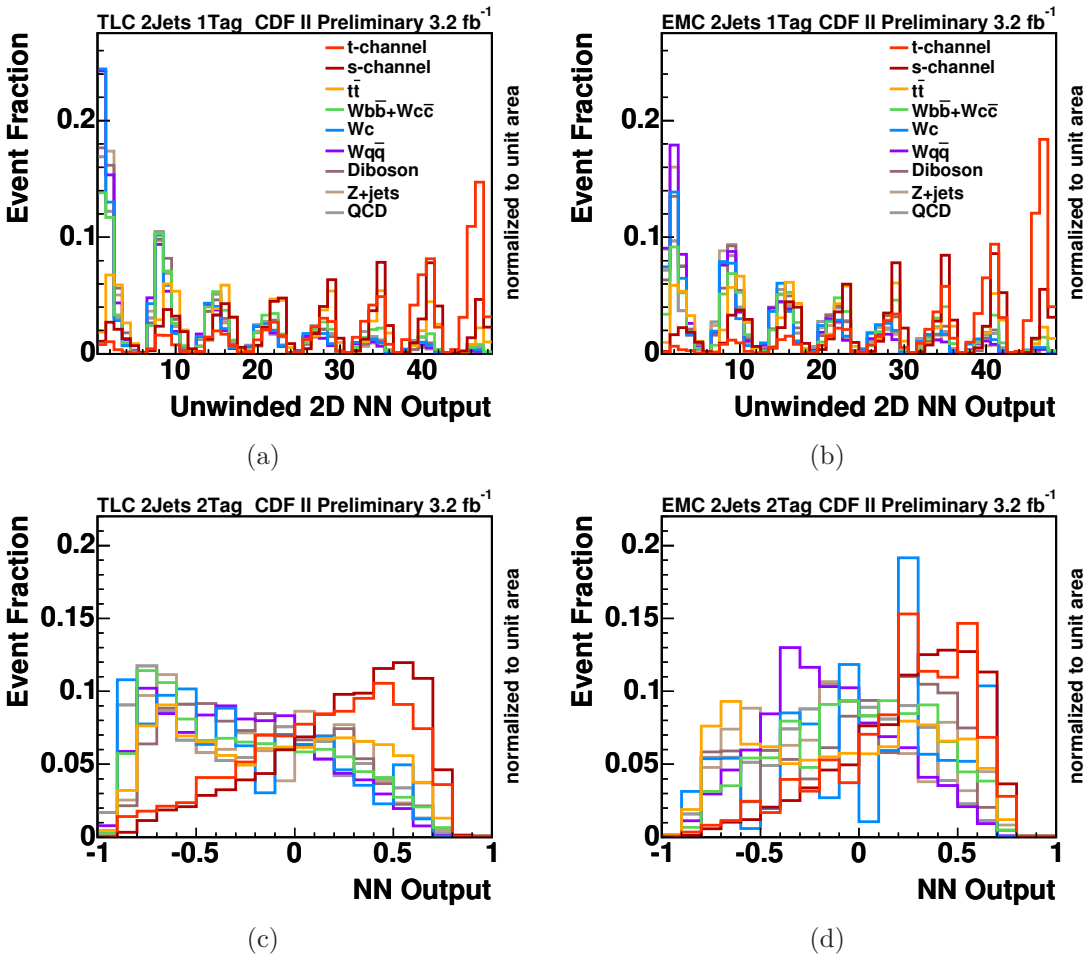


Figure 6.14: The unwinded 2D NN output of the s- and t-channel neural networks in the two-jet one-tag category (top) and the NN output of the s-channel neural network in the two-jet two-tag category (bottom) for TLC (left) and EMC (right) events

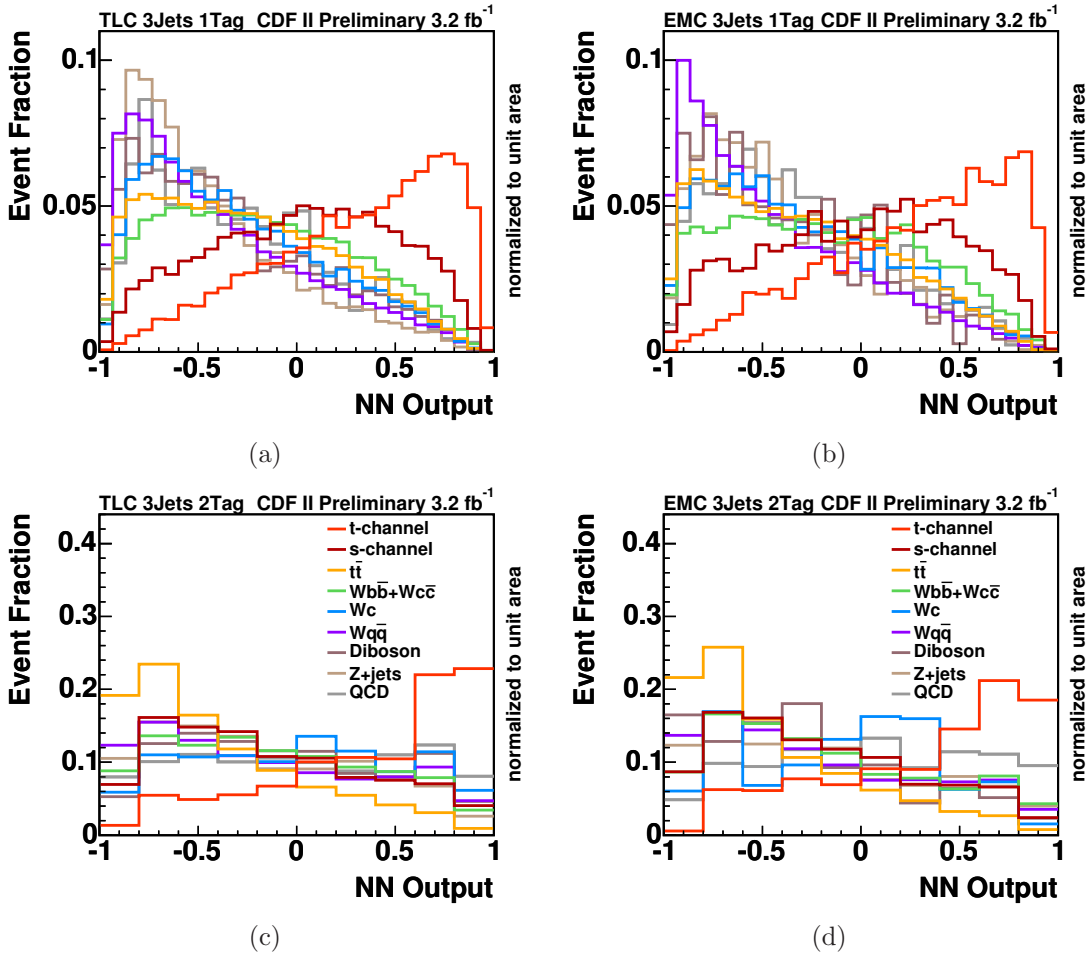


Figure 6.15: The NN output of the t-channel neural network in the three-jet one-tag category (top) and of the t-channel neural network in the three-jet two-tag category (bottom) for TLC (left) and EMC (right) events

Chapter 7

Analysis

The goal of this analysis is to discover combined s- and t-channel single top-quark production if present and exclude it if absent, using sound and robust statistical methods. Therefore, the expected significance of the analysis must be evaluated, assuming expected SM signal and background features. In case of a present signal, the determination of its observed significance is of vital importance whether to claim an observation or not.
Furthermore, the extraction of both the combined as well as the separate s- and t-channel signal production cross-sections are performed, assuming a top-quark mass of $m_t = 175\,\mathrm{GeV}/c^2$. If no signal can be established, expected and observed upper limits on the production cross-sections are determined.

The neural network classifiers of the combined search for both s- and t-channel single top-quark production are optimized for a maximal expected signal significance, while the classifiers of the separate search are focused to achieve a minimal expected uncertainty on the measurement of the s- as well as the t-channel cross-sections. Both optimizations assume production rates and classifier shapes for signal as well as background processes as predicted by the latest available theoretical SM calculations and up-to-date models. Such highly optimized discriminants need to be carefully checked for sources of mismodeling, which is done in several sidebands of collision data.
The most time consuming part is the identification of sources for systematic uncertainties, the evaluation of alternative approaches with full coverage of the expected systematic effects, and their implementation into the analysis.

Finally, the neural network output templates in all channels are fitted simultaneously to the observed output distributions to determine the observed significance as well as to possibly measure the single top-quark production cross-sections, if a signal can be established.

7.1 Statistical Method

The determination of the expected and observed significance as well as the measurement of the production cross-section is performed with a binned likelihood function. To compute the significance of a potentially observed signal, a hypothesis test is carried out to determine whether the candidate events of the collision data are preferring the signal-plus-background hypothesis or the background-only hypothesis. It is based on a test statistic represented by a likelihood ratio, which is build upon the entries in each bin of the discriminating classifier. For each of the two considered hypotheses, ensemble tests of pseudoexperiments are computed to obtain the corresponding distributions of the test statistic. To obtain the observed value of the test statistic, the likelihood ratio is determined on collision data.
Finally, the cross section for single top-quark production is extracted by fitting the templates of the signal and various background processes to the discriminating classifier distribution of the observed collision data using a binned likelihood function.

7.1.1 Binned Likelihood Function

Both, combined and separate search, apply the same binned likelihood function, built up of different terms. The first term is a Poisson distribution for the content of each bin in the considered classifier distribution, yielding to

$$L = \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \ , \tag{7.1}$$

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

$$\mu_k = \sum_{j=1}^{A} \mu_{jk} = \sum_{j=1}^{A} \beta_j \cdot \hat{\nu}_j \cdot \alpha_{jk} \ , \tag{7.2}$$

where j is the index of a physical process, A the number of physical processes and μ_{jk} is the expectation value of the physical process j in bin k. μ_{jk} itself depends on β_j which represents the ratio of the measured number of events and the estimated number of events for process j. β_j can also be interpreted as the ratio of the measured and the predicted cross section of process j. Since negative production rates are unphysical, β_j is not allowed to fluctuate negative. Moreover, $\hat{\nu}_j$ is the total number of expected events for process j in the given data set and α_{jk} is the relative fraction of events of process j in the bin k. Hence, α_{jk}, also referred to as the shape of process j, fulfills the normalization condition $\sum_{k=1}^{B} \alpha_{jk} = 1$.

An additional term in the likelihood function implements Gaussian constraints to the rates of the various background processes. It is reasonable to include the a priori knowledge of the uncertainties of the rates of the processes from the candidate event yield estimate into the likelihood function. Thus, the application of the maximum likelihood technique is more robust due to the avoidance of free floating rates of the background processes. Hence, the likelihood function with the Gaussian constraints is given by

$$L = \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \cdot \prod_{j=2}^{A} \frac{1}{\sqrt{2\pi\Delta_j^2}} e^{-\frac{(\beta_j-1)^2}{2\Delta_j^2}} = \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \cdot \prod_{j=2}^{A} G(\beta_j, \Delta_j) \ , \quad (7.3)$$

where the Δ_j are the relative uncertainties on the estimated number of background events of the processes j. For the combined search, the product in the second term starts at $j = 2$, since the rate of the combined signal process, labeled with $j = 1$, is the parameter to be measured. In case of the separate search, since both the rates of s- and t-channel signal are included separately as $j = 1$ and $j = 2$, the second term for the background rates starts at $j = 3$.

Furthermore, two different types of systematic uncertainties are incorporated by adding a last term to the likelihood function. Uncertainties which affect the rate of a particular signal or background process are called rate uncertainties, and uncertainties which change the template shape of a process are referred to as shape uncertainties. These systematic uncertainties enter the likelihood function by shifting the expected mean of every process j in every bin k. Therefore, the mean value μ_{jk} changes to

$$\begin{aligned} \mu_{jk} = \beta_j \cdot \hat{\nu}_j \cdot \left\{ \prod_{i=1}^{S} (1 + |\delta_i| \cdot (\epsilon_{ji+} H(\delta_i) + \epsilon_{ji-} H(-\delta_i))) \right\} \\ \cdot \alpha_{jk} \cdot \left\{ \sum_{l=1}^{S'} (1 + |\delta_l| \cdot (\kappa_{jlk}^{+} H(\delta_l) + \kappa_{jlk}^{-} H(-\delta_l))) \right\} \ , \end{aligned} \quad (7.4)$$

where the δ_i represent the strength of an uncertainty i, the ϵ_{ij} are the relative rate uncertainties, the $\kappa_{jlk}^{\pm}$ are the relative shape uncertainties, S and S' are the number of systematic rate and shape uncertainties, respectively. If a systematic effect causes both rate and shape uncertainties, δ_i and δ_l are 100% correlated. $H(x)$ denotes the Heaviside step function. The value of $\kappa_{jlk}^{\pm}$ is derived from normalized, systematically shifted histograms (α_{jlk}^{+} and α_{jlk}^{-}) taking into account the change of the discriminant shape due to the application of one or two systematically shifted alternative models:

$$\kappa_{jlk}^{\pm} = \frac{\alpha_{jlk}^{\pm} - \alpha_{jk}}{\alpha_{jk}} \ , \text{with} \sum_{k=1}^{B} \alpha_{jk} \cdot \kappa_{jlk}^{\pm} = 0 \quad (7.5)$$

The strength of a systematic uncertainty δ_i is again constrained by a standardized Gaussian distribution. If the alternative models representing the systematic uncertainty of source i are 1σ variations of the default model, the Gaussian distribution

is effectively not restricted by requiring the strength to be consistent within at least a 5σ deviation from the nominal value, $|\delta_i| < 5$. Some systematic sources are treated by alternative models corresponding to a maximal possible deviation. Since extrapolation above such maximal boundaries results in unphysical behavior, the corresponding systematic strength is truncated to the valid region, $|\delta_i| < 1$. Hence, the final likelihood function is given through

$$\begin{aligned} L &= \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \cdot \prod_{j=2}^{A} G(\beta_j, \Delta_j) \cdot \prod_{i=0}^{S} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\delta_i)^2}{2}} \\ &= \prod_{k=1}^{B} \frac{e^{-\mu_k} \cdot (\mu_k)^{n_k}}{n_k!} \cdot \prod_{j=2}^{A} G(\beta_j, \Delta_j) \cdot \prod_{i=0}^{S} G(\delta_i; 0, 1) \ . \end{aligned} \tag{7.6}$$

Finally, for performing hypotheses tests and the cross-section measurements, the likelihood function with respect to the signal rate β_1, the reduced likelihood $L_{\mathrm{red}}(\beta_1)$, needs to be determined. Thus, one needs to get rid of all unwanted parameters β_2-β_{A} and δ_1-δ_{S}, referred to as nuisance parameters. For hypothesis testing, the method of profiling the likelihood is performed. Here, the minimum of the negative log-likelihood function with respect to all nuisance parameters is derived for each fixed β_1 value, using the MINUIT [151] package. The method of marginalization is used for the cross-section measurement. Therefor, the likelihood function is integrated over all nuisance parameters, also resulting in the reduced likelihood $L_{\mathrm{red}}(\beta_1)$.
The reason for using two different approaches to obtain the reduced likelihood function is based on a mis-behavior within the profiling procedure with MINUIT. In the past, it has been found that fitting for nuisance parameters could yield discontinuities in the first derivatives of the likelihood function with respect to the nuisance parameters. As a result, MINUIT could obtain incorrect uncertainties for functions which have such discontinuities. In particular, this is the case for shape uncertainties, for which piecewise linear interpolations are used. Figure 7.1 shows the δ_{JES} distribution of the jet energy scale (JES) uncertainty, which is affected by both rate and shape uncertainties of signal and background processes.

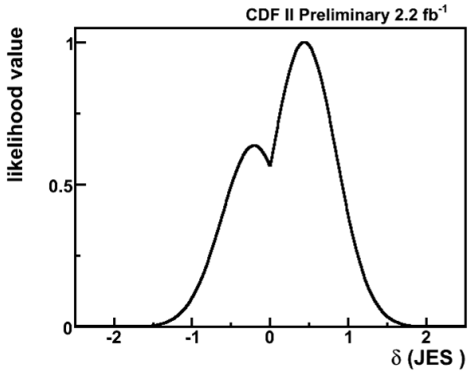


Figure 7.1: Double peak structure of the strength of the jet energy scale (JES) uncertainty δ_{JES} obtained by the profiling procedure, resulting in an incorrect estimated uncertainty.

Such mis-behavior can be avoided by using the marginalization procedure at the expense of computing time. This makes marginalization only feasible for the cross-section measurement, not for hypothesis testing.
Hence, the likelihood function for the hypothesis tests only includes uncertainties on the rates of the background processes, since the corresponding nuisance parameters don't have any discontinuities. The remaining systematic uncertainties as well as bin-by-bin uncertainties due to the limited number of simulated events are accounted for by fluctuating the values of the nuisance parameters in the generation of the sets of pseudoexperiment ensembles, referred to as prior-predictive ensembles.

7.1.2 Hypothesis Testing

In order to compute the expected and observed significance of the analysis, a modified frequentist approach is used, which extends a technique developed at LEP [152]. All but the treatment of the systematic uncertainties, which is Bayesian [153], is based on frequentist principles.
For the hypothesis testing, two sets of ensembles consisting of simulated pseudoexperiments are created. For each of the pseudoexperiments, the number of events of a particular process N_j is randomly drawn from a Poisson distribution of mean $\hat{\nu}'_j$. Systematic rate uncertainties are incorporated via

$$\hat{\nu}'_j = \hat{\nu}_j \cdot \left\{ \prod_{i=1}^{S} (1 + |\delta_i| \cdot (\epsilon_{ji+} H(\delta_i) + \epsilon_{ji-} H(-\delta_i))) \right\} , \tag{7.7}$$

where the strength of the systematic uncertainties δ_i is drawn from a standardized Gaussian distribution. After that, N_j random numbers are drawn from the template distributions α'_{jk} for each process j to obtain the shape of the pseudo-data of this process in this particular pseudoexperiment. Here, the systematic shape uncertainties are included by

$$\alpha'_{jk} = \alpha_{jk} \cdot \left\{ \sum_{l=1}^{S'} (1 + |\delta_l| \cdot (\kappa^{+}_{jlk} H(\delta_l) + \kappa^{-}_{jlk} H(-\delta_l))) \right\} , \tag{7.8}$$

where δ_l is also drawn from Gaussian distributions which are centered at zero with standard deviation one. Again, if a systematic effect causes both rate and shape uncertainties, δ_i and δ_l are 100% correlated.
Additionally, the uncertainty due to the limited amount of simulated events is incorporated by fluctuating the number of entries in each bin k of the α_{jk} templates with a Gaussian distribution, centered at the original value with a standard deviation that represents the statistical uncertainty on the number of entries. Finally, the full discriminant distribution of pseudo-data is obtained by summing up all N_j pseudo-events of all processes j.

The hypothesis test is performed by considering two hypotheses. The first one, the null hypothesis H_0, assumes that the single top-quark production cross-section is zero, setting $\beta_1 = 0$. The second one, the test hypothesis H_1, assumes that the signal production cross-section is the one predicted by the standard model, fixing $\beta_1 = 1$. The objective of the analysis is to observe single top-quark events, that means to reject the null hypothesis H_0.

The hypotheses test is based on the Q-value test statistic, a likelihood ratio which is defined by

$$Q = -2 \ln \frac{L_{\mathrm{red}}(\beta_1 = 1)}{L_{\mathrm{red}}(\beta_1 = 0)} \ . \tag{7.9}$$

This choice of the test statistic is motivated by the Neyman-Pearson lemma [154], which indicates that a likelihood ratio is the most sensitive variable for separating hypotheses.

To evaluate the expected false discovery rate of the analysis assuming the standard model and its observed false discovery rate with the given collision data, the comparison to the null hypothesis H_0 has to be performed using the Q-value distributions. The latter are referred to as q_0 and q_1 for the H_0 and H_1 hypotheses, respectively. With these distributions the probability for H_0 to be true can be quoted either for the expected significance of the SM assumption or the observed significance of the collision data. For this purpose the so-called p-value is defined via

$$p(Q) = \frac{1}{A_q} \cdot \int_{-\infty}^{Q} q_0(Q')\, dQ' \ , \tag{7.10}$$

where $A_{\mathrm{q}} = \int_{-\infty}^{\infty} q_0(Q')\, dQ'$. For the determination of the expected significance, Q_{exp} is the median expected Q-value of the distribution q_1 describing the test hypothesis H_1, which assumes SM single top-quark production. Hence, the meaning of the expected p-value $\hat{p}_{\mathrm{exp}}$ is the following: Under the assumption that H_1 is true, one expects to observe $p < \hat{p}_{\mathrm{exp}}$ with a probability of 50%.
To obtain the observed significance from collision data, the observed p-value $\hat{p}_{\mathrm{obs}}$ is determined using the Q-value Q_{obs} measured in collision data. Therefore, the reduced likelihood $L_{\mathrm{red}}(\beta_1)$ is obtained from the binned likelihood fit to collision data instead of pseudoexperiments.
By converting the median expected and observed p-values $\hat{p}_{\mathrm{exp}}$ and $\hat{p}_{\mathrm{obs}}$ into a number of standard Gaussian deviations σ using the integral of one side of a standardized Gaussian distribution, the expected and observed significances can be quoted. The criterion used to claim evidence and observation is $\hat{p} < 1.35 \cdot 10^{-3}$ and $\hat{p} < 2.87 \cdot 10^{-7}$, which means to see at least that many candidate events that the observed excess over the background processes corresponds to a background fluctuation of 3σ or 5σ, respectively.

7.1.3 Cross-Section Measurement Method

The extraction of the single top-quark production cross-section from collision data is a full Bayesian procedure. The measurement including all systematic rate and shape uncertainties is done by determining the reduced likelihood function $L_{\mathrm{red}}(\beta_1)$ for the combined search. In case of the separate search, the reduced likelihood function $L_{\mathrm{red}}(\beta_1, \beta_2)$ is a two-dimensional function of both the s- and t-channel signal cross-sections. The reduced likelihood function, obtained via marginalization, is then converted into a posterior probability density as a function of the single top-quark production cross-section. The most probable value and the corresponding uncertainties are extracted, the 1σ uncertainty is defined as the region around the most probable value that contains 68% of the total area of the probability density function and with boundaries at equal values.

7.2 Optimization of Event Classification Technique

With the expected significance as a measure of the combined search analysis' sensitivity, this figure of merit can be used to further check for optimizations to increase the discovery potential of the analysis. Over the last years of the development of the analysis, several different approaches and tunings were examined. Only modifications with reasonable improvements above 3% gain in expected significance were finally considered.

More than 50 different input variables were tested. Those tests include simple as well as rather complicated kinematic event shape and reconstruction variables. Most of them couldn't add additional discriminating power to the trained neural networks due to high correlations to already incorporated variables. A few of them had severe mis-modeling problems and had to be canceled. But several minor modifications to the calculation of some variables showed significant improvements, i.e. a different choice of jet corrections for the reconstruction of physical objects or whether to choose the b-tagged jet or the jet with maximum $Q\,(\text{lep}) \cdot \eta\,(\text{jet})$ as the b jet coming from the top-quark decay. Finally, less than 20 input variables were sufficient for the neural network training regardless of which channel. The by far biggest improvement to the analysis was achieved by the introduction of the KIT flavor separator, see chapter 6.2.1, boosting the expected significance by up to 20%.

Dividing the sample of selected candidate events into separate channels with different expected ratios of signal-over-background candidate events additionally improved the overall performance. Including candidate events with three jets didn't add much to the discrimination power at all, regardless of using this sample as an additional channel for the search of signal events or using it as an sideband to constrain the dominant top-quark pair production or using it for both simultaneously through two-dimensional discriminant classifiers. Nevertheless, the sample of three jet events is included in the final analysis to be maximally inclusive with respect to the signal acceptance. At the end, four classifiers were optimized, one per jet and b-tag multiplicity. Each of those four sub-samples was additionally subdivided by the categories of TLC and EMC lepton types, but the analysis showed no improvement by allocating special TLC and EMC optimized classifiers. Also the binning of the classifier discriminants had no influence on the outcome, mainly due to the incorporation of the statistical uncertainty specifying the limited amount of simulated MC events.

The NeuroBayes® package used to train the neural network classifiers provides several tuning options. It turned out, that the standard set of parameters seems to be already optimal. Changing the type of variable preprocessing, the error function to be minimized, the method of minimization, the learning speed of the network, the kind of classification, and the rate of weight updates during the learning process only resulted in insignificant additional gain, if at all. Even the variation of the number of hidden nodes of the network architecture from very few up to dozens of nodes didn't change the classifier's discrimination power, owing to NeuroBayes® complex and robust preprocessing.

Another approach to improve the analysis' performance is to optimize the composition of the sample used to train the classifier. Changing the composition of background processes from the default expected SM mixture to a sample dominated by only the most relevant background processes like top-quark pair or $Wb\bar{b}$ production seems not to increase the classifiers power to separate single top-quark events from background events. The same is noticeable when changing the fraction of signal events within the training sample composition from the default of 50% down to 20% and up to 65%. Additionally, several boosting methods were checked, no gain in expected significance was noticeable.

To achieve minimal expected uncertainties on the simultaneous measurement of the separate s- and t-channel signal cross-sections, the classifier of the dominant candidate sample with two jets and one b tag was subject to optimization studies. Several reasonable and significant improvements were found. Instead of using a classifier trained for signal t-channel separation as with the classifier of the combined single top-quark search, an additional s-channel trained network is included. This is done by combining both of them into a two-dimensional classifier. Moreover, the separation power of the s-channel network could be improved by changing the choice which jet to assign to the b jet from the top-quark decay from the default hypothesis of the b-tagged jet to the jet with maximum $Q\,(\text{lep}) \cdot \eta\,(\text{jet})$. Further improvement was achieved by not only training those s- and t-channel networks against the background mixture but also against the other single top-quark signal. Therefor, the other single top-quark signal rate in the training sample was enlarged by a factor of five. Finally, all those contributions improved the expected uncertainty on the separate measurements of the s-channel cross-section by about 14%. Figure 7.2 shows the outcome of the overall optimizations for the separate search.

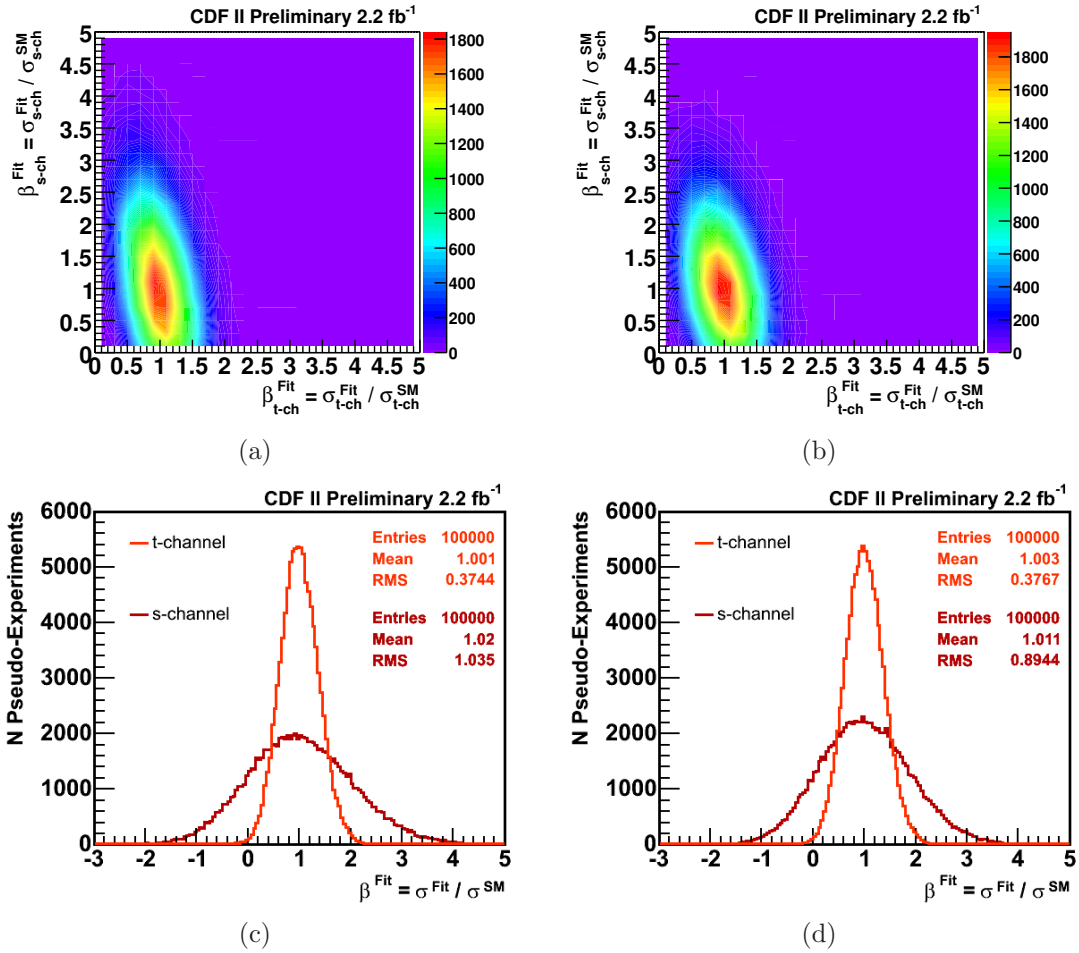


Figure 7.2: Impact of the improvements for the separate search for s- and t-channel signal cross-sections: The left-hand side shows both expected signal β distributions as a two-dimensional (a) and the corresponding one-dimensional (c) outcomes of a SM ensemble of pseudoexperiments with the classifiers trained for the combined search. The right-hand side shows the analog distributions (b) and (d) of the optimized classifiers for the separate search. The expected uncertainty on the s-channel cross-section measurement, given by the RMS value, improved by about 14%.

7.3 Validation of Event Classification

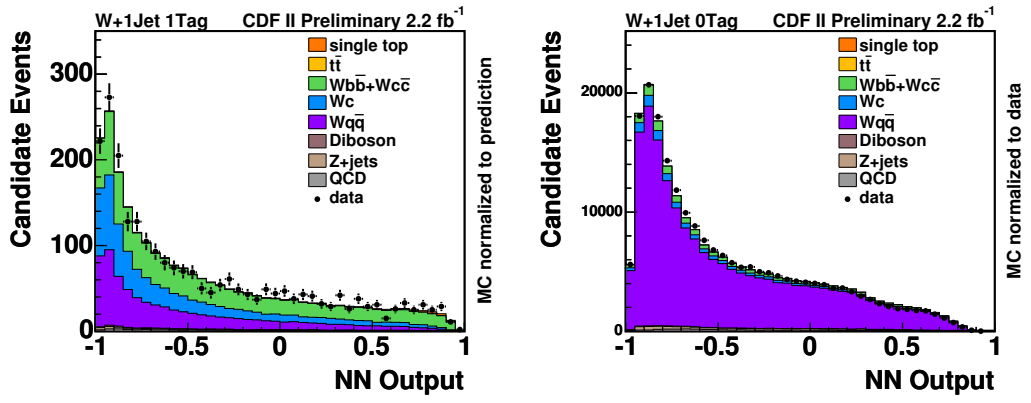
Besides the validation of the modeling of all input variables of a classifier, its output needs to be checked for mis-modeling effects. To increase the confidence in a classification technique, collision data in the signal region as well as in the sideband region can be used to verify the general usability of a technique. Even more, sideband collision data is also a testbed for the final classifiers, especially if the sideband data is expected to have the same event kinematics as the signal region.

7.3.1 Coverage of Classification Technique

To gain confidence in a rather complex procedure, it is helpful to study its characteristics in all accessible aspects. Neural network classifiers are a kind of special non-trivial calculations. To verify their robustness and validity, they should be applied in all available scopes which allow for an explicit examination.

Classifier for FCNC Single Top-Quark Production

Sideband collision data with only one jet is expected to be populated by the same background processes as the signal region with two or three jets, except for a different composition. This makes it an ideal testbed for validation studies. For the search of single top-quark production via flavor-changing neutral currents (FCNC) using 2.2 fb^{-1} at CDF [38,39], NeuroBayes® neural network classifiers were trained to separate SM events from SM suppressed FCNC events. This was done in the candidate event signal sample with one b tag as well as in its high statistics control sample of events with no b tag. Both classifiers show good agreement between the distributions of the collision data and the SM expectations, as shown in figure 7.3.



(a) (b)

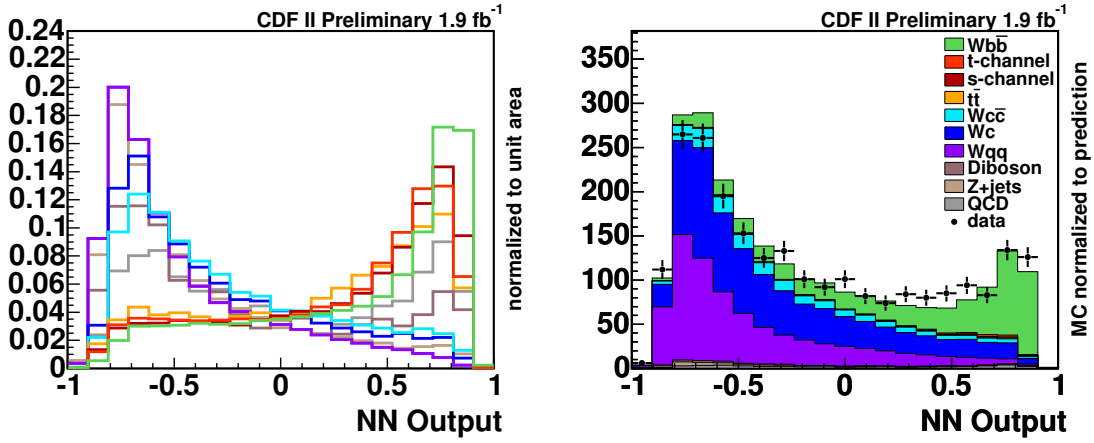
Figure 7.3: Both the signal region (a) and the control region (b) of the search for FCNC single top-quark production show good agreement of the expected and observed NN output [38,39].

Classifier for $Wb\bar{b}$ Production

In the one jet sample, an additional neural network was trained to separate $Wb\bar{b}$ events [29]. The underlying idea is to create a classifier for a known SM process and check its performance and modeling on collision data. The $Wb\bar{b}$ network is trained on events with one jet and one b tag, whereas the training and network parameters stayed at their default values.

By far the most important variable for this neural network is the KIT flavor separator. It separates the $Wb\bar{b}$ signal from those processes that do not contain a real b quark, i.e. $Wc\bar{c}$, Wc, and $Wq\bar{q}$. Since these are the most important background processes, the KIT flavor separator is the variable with the best discriminating power. The second best variable, the transverse energy of the jet, $E_{\mathrm{T}}(jet1)$, discriminates $Wb\bar{b}$ from events with a real produced top-quark, that is $t\bar{t}$ and single top-quark production. The energy of the jet is higher for these events, since a larger center-of-mass energy is required to produce one or two top quarks in the first place.

By including 12 additional input variables, the NeuroBayes® neural network classifier gains high discriminating power against the dominant contributing background processes, as can be seen in figure 7.4 (a). Figure 7.4 (b) shows the accumulation of the expected $Wb\bar{b}$ signal in the signal-like region around NN output of 1. The comparison to the observed candidate events reveals, that the collision data prefers a higher rate of signal-like events of about 24%. This is in good agreement with studies of the W+HF contribution in the one jet collision data, which resulted in the introduction of the heavy-flavor correction factor $K = 1.4 \pm 0.4$, see chapter 5.2.3.



(a) (b)

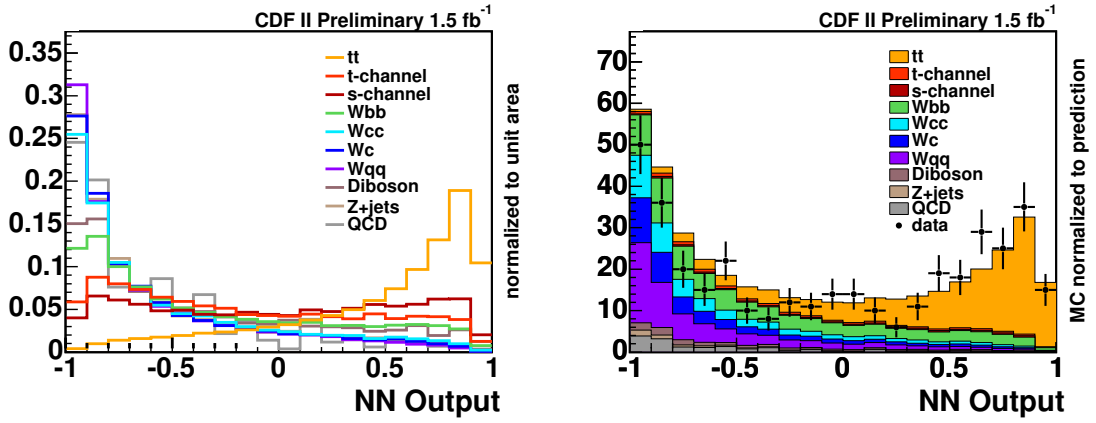
Figure 7.4: The high discriminating power of the $Wb\bar{b}$ classifier against the dominant contributing background processes is visible in the shape comparison of the NN output (a). The expected and observed NN output (b) of the $Wb\bar{b}$ classifier shows good agreement when keeping in mind the heavy-flavor correction factor $K = 1.4 \pm 0.4$, which is not included in the prediction for events with one jet.

Classifier for $t\bar{t}$ Production

As a further cross check, a $t\bar{t}$ discriminating network was trained in the $W + 3$ jets sample [29]. Hereby exactly one b-tagged jet was required. The training sample is again composed of 50% signal events, while the remaining fraction of background processes corresponds to the predictions.

The most relevant variable is the sum of the transverse energies H_{T} of all particles produced in the event, which is the missing transverse energy from the undetected neutrino, the transverse energy of the lepton, and the transverse energies of the jets. Only a very high center-of-mass energy can lead to the production of two top quarks, resulting in a large H_{T} . In case of the production of a single top-quark, much less energy is required. For the other processes that do not contain a top quark in the event, the H_{T} distributions are shifted even further to lower energies. Thus, H_{T} is a variable that distinguishes $t\bar{t}$ from all other processes. The second most important variable is again the KIT flavor separator.

Through the inclusion of 14 additional input variables, good separating power is achieved, as seen in figure 7.5 (a). The NN output of the $t\bar{t}$ classifier, see figure 7.5 (b), shows good agreement between the prediction and the candidate events. The fit to the observed candidate events extracts a $t\bar{t}$ production cross-section of $7.5 \pm 0.8\,(\mathrm{stat})$ pb, which is in good agreement with the recently obtained value of $7.2 \pm 0.4\,(\mathrm{stat.}) \pm 0.5\,(\mathrm{syst}) \pm 0.4\,(\mathrm{lumi})$ pb by a comparable CDF analysis [155].

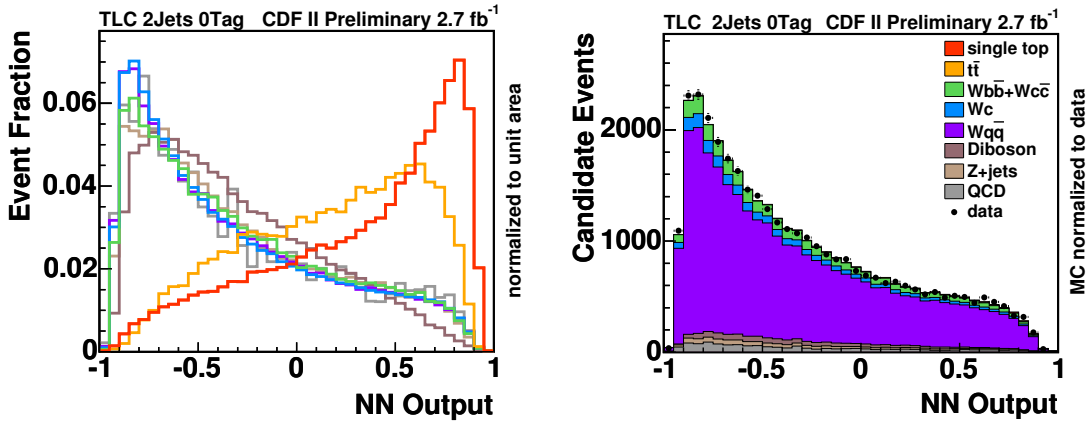


(a) (b)

Figure 7.5: The discriminating power of the $t\bar{t}$ classifier trained for events with three jets and one b tag against the contributing background processes is visible in the shape comparison of the NN output (a). The expected and observed NN output (b) of the classifier shows good agreement.

Classifier for Single Top-Quark Production in the 0-Tag Sideband

To verify NeuroBayes® neural networks trained for single top-quark production, classifier training and validation in the kinematically similar sideband region of events with two or three jets and 0 b tags is an appropriate measure. Due to its negligible signal content with respect to the dominant W+jets production, this provides a nice proving ground since no signal excess over background is expected. The training in the 0-tag sample requires to adjust the choice of the jet coming from the top-quark decay. It is changed from the tagged jet to the most central jet, since no b tag is available. Due to the same reason, the KIT flavor separator, which operates on jets identified by the SecVtx b-tagging algorithm, is also unavailable. Except for the latter, the same variables are used for the 2-jet network classifier as in the case of the t-channel network in the 2-jet bin with 1 b tag and of the s-channel network in the 2-jet bin with 2 b tags. For the 3-jet network classifier exactly the same input variables as in the case of the t-channel networks in the 3-jet bin with 1 or 2 b tags are included. Finally, the combination of 21 and 25 input variables are used to train the network classifier for t-channel single top-quark events with two or three jets, respectively, whereof the following are the most discriminating ones: the invariant mass of the two most energetic jets, the sum of the transverse energies of all particles produced in the event, the product of the charge of the lepton and the pseudorapidity of the light-quark jet, and the reconstructed top-quark mass.
Neural network classifiers are trained in both event samples with two and three jets, the validation samples are additionally split into the two lepton type categories TLC and EMC. All classifier output distributions show good agreement between the expected background processes and the observed candidate events. As an example, the achieved discrimination power and the comparison to collision data for TLC events with two jets is illustrated in figure 7.6 (a) and (b), respectively.



(a) (b)

Figure 7.6: The discriminating power of the signal classifier trained for TLC 2-jet 0-tag events is illustrated in (a). The expected and observed NN output (b) of the classifier shows good agreement in this high statistics sideband region.

7.3.2 Control-Region Validation of Classifier

In contrast to the validation of classifiers especially trained in the 0-tag sideband, the final signal region optimized classifiers can be furthermore checked with 0-tag sideband collision data. Once more, the choice of the b-tagged jet is replaced by the most central one, where appropriate. The output value of the KIT flavor separator is also unavailable, the choice of a replacement value turned out to be marginal concerning the validation outcome quality. A value of -2, -1, 0, 1, and even a uniform distributed random number between 0 and 1 doesn't change the classifier's ability to model the observed collision data distribution. Using a KIT flavor separator value of 0, the following comparisons between the expected background processes and the observed collision events show good agreement in the 2-jet (figure 7.7) and 3-jet (figure 7.8) events sample, as well as the 0-tag channels combined (figure 7.9).

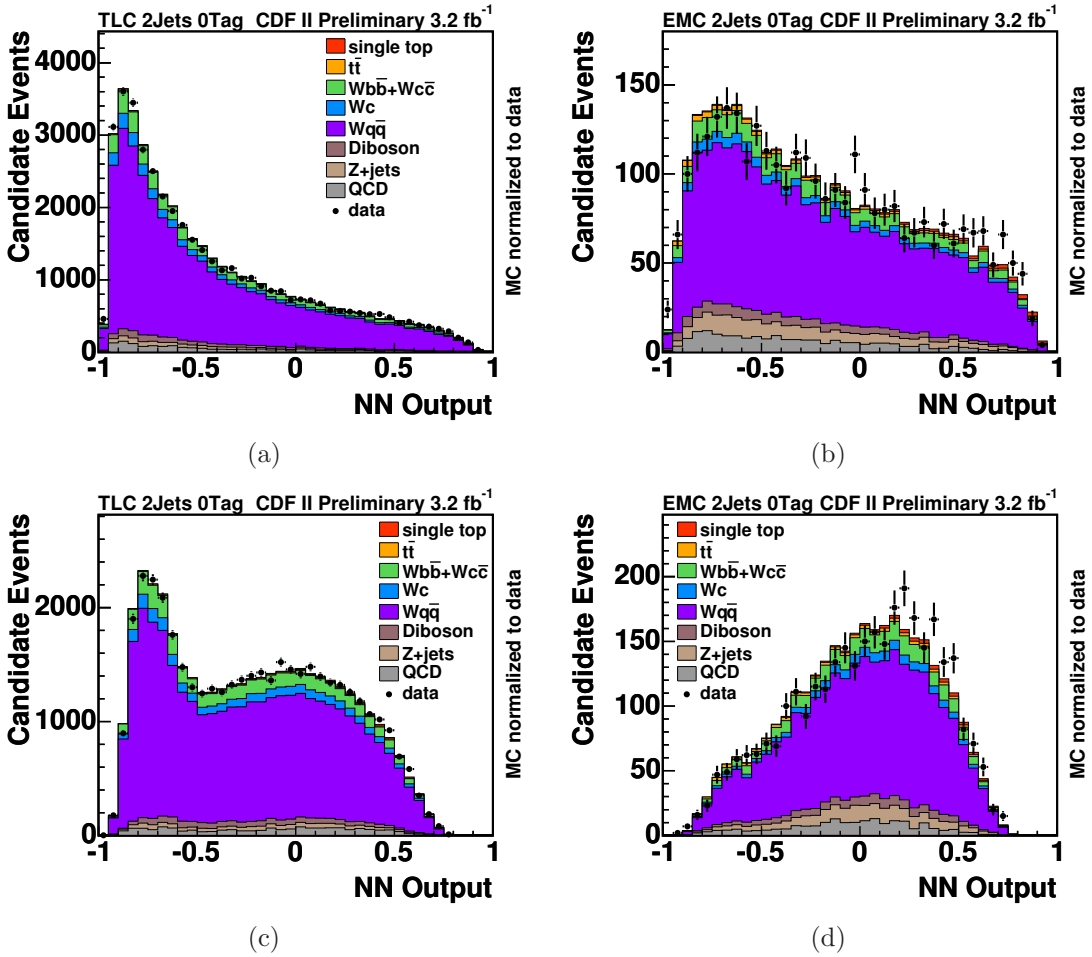


Figure 7.7: The expected and observed output of the final neural network classifier trained for t-channel 2-jet 1-tag events applied to 0-tag TLC (a) and EMC (b) events, and those trained for s-channel 2-jet 2-tag events applied to 0-tag TLC (c) and EMC (d) events shows good agreement in this high statistics sideband region.

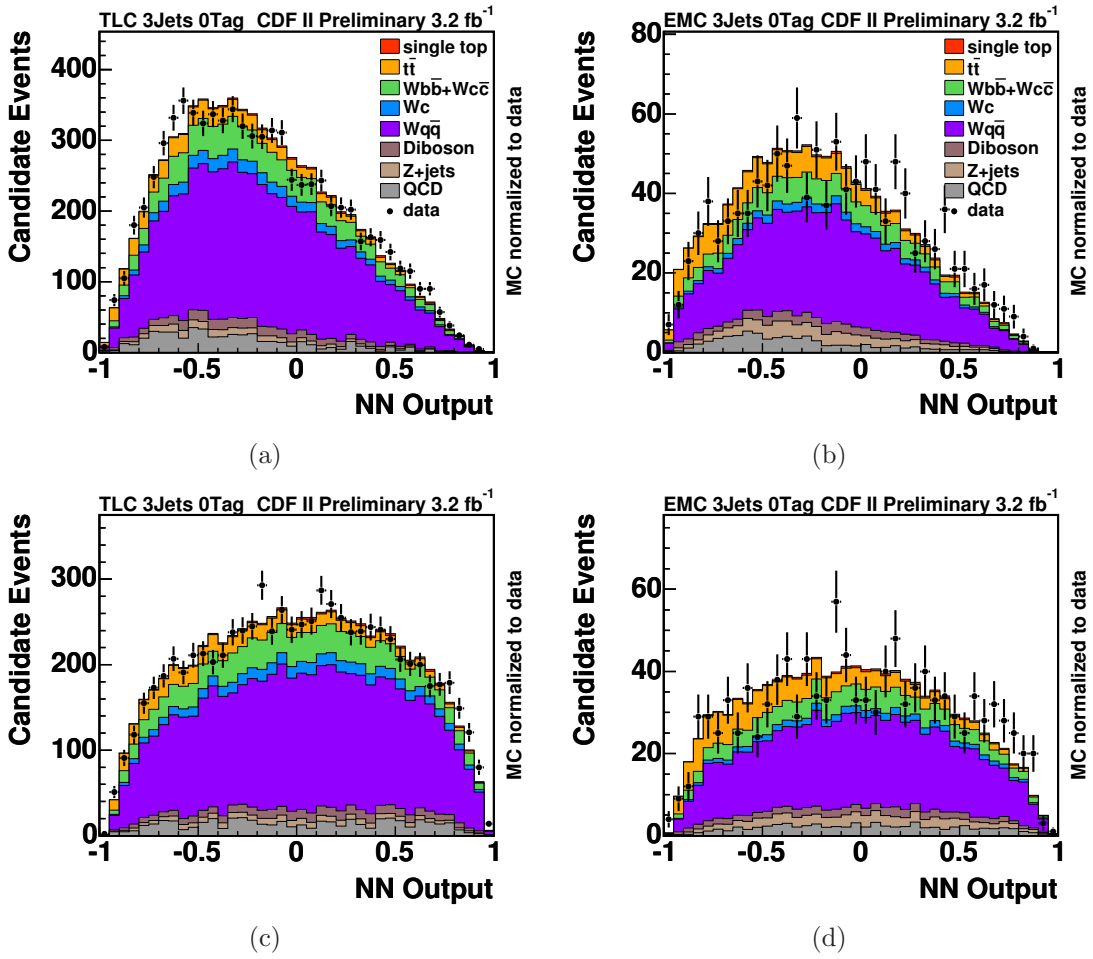


Figure 7.8: The validation of the final neural networks trained for t-channel 3-jet 1-tag events applied to 0-tag TLC (a) and EMC (b) events, and those trained for t-channel 3-jet 2-tag events applied to 0-tag TLC (c) and EMC (d) events shows good agreement.

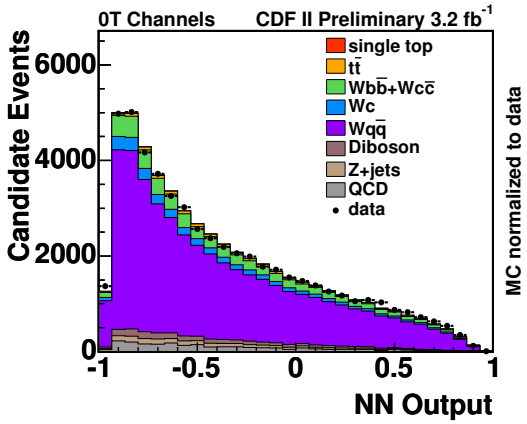


Figure 7.9: The validation of the combined final classifiers trained in the 1 tag event samples show good agreement in the 0-tag sideband region.

7.4 Systematic Model Uncertainty

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. The analyses are done under the assumption of a top-quark mass of $m_t = 175\,\mathrm{GeV}/c^2$. That is why the uncertainty in the top-quark mass is not taken into account as a systematic uncertainty when extracting the single top-quark production cross-sections. Hence, the analyses provide rather a measurement at the specified value of the top-quark mass. However, for the determination of the expected and observed significance, both rate and shape uncertainties are taken into account. The impact of the sources of uncertainties is evaluated by altering the modeling of the corresponding processes or effects within their uncertainties or by assigning a plausible alternative model. As a result, relative changes of the event rates and shifted template distributions are obtained.

The following sources of systematic uncertainties are considered. The uncertainty on the corrections for the jet energy scale (JES) is determined by varying the corrections within their $\pm 1\sigma$ uncertainties [109]. This results in rate as well as shape uncertainties for signal and all simulated background processes.
The influence of correlated initial- and final-state gluon radiation (IFSR) is estimated by producing samples of simulated events for which the simulation is altered to produce either less or more gluon radiation compared to the standard setting [156]. In case of both signal processes, the very same simulated partonic events as in the default sample are used to avoid statistical fluctuations. 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 rates and template shapes are obtained for simulated single top-quark and $t\bar{t}$ events.
The choice of the parameterization of the parton distribution functions (PDF) used for the event simulation is another source for a systematic uncertainty. Its size is estimated by reweighting single top-quark and $t\bar{t}$ events with weights associated to various alternative PDF sets, whereas the CTEQ5L PDF set was used for default event simulation [71]. The CTEQ6M PDF provides a set of 20 orthogonal pairs of eigenvalues (EV) corresponding to the uncertainties of its determination. Those are treated as independent systematic sources and their impact is included by the bin-wise quadratic sum of the weights. Moreover, the difference between the default and MRST72, as well as between the default and MRST75 are considered to be taken into account as an additional rate and shape uncertainty added up bin-by-bin in quadrature.
The impact of using LO instead of NLO Monte Carlo event generators is evaluated for single top-quark and $t\bar{t}$ events, referred to as MC uncertainty. For the signal, the uncertainty is quantified by taking the differences between the default LO Monte Carlo simulation and the theoretical NLO calculation provided by ZTOP [26]. This is done by assigning weights to simulated events, derived from a comparison of six

kinematic distributions: the p_{T} and the η of the top quark and the two highest E_{T} jets, which do not originate from the top-quark decay. The correlation between the different variables is determined from the simulated events generated by MadEvent. By applying the event selection to the reweighted Monte Carlo events, an estimate on the deviation of the acceptance in the simulation compared to the NLO prediction is found. For $t\bar{t}$ events, the systematic change of the acceptance rate between the default Pythia and the alternative MC@NLO model is considered.
The rate uncertainty in the event detection efficiency ϵ_{evt} includes the uncertainties on the trigger efficiency, on the lepton identification efficiency, and on the b-tagging efficiency which is the dominating factor. Other sources of rate-only systematics are the uncertainty on the luminosity determination and the $Wq\bar{q}$ double b-tagging efficiency, respectively.
Since no cut is applied on the output of the KIT flavor separator, the uncertainty associated with this quantity does not imply a rate uncertainty, but a shape uncertainty on the template distributions. Systematic effects are studied by utilizing the correction function derived for the $Wq\bar{q}$ events [28, 30]. Therefore two scenarios are considered. For the pessimistic one, the correction function is applied on the charm-like templates, making it more signal-like. For the optimistic one, the uncorrected $Wq\bar{q}$ shape is used, which makes the $Wq\bar{q}$ template more background-like.
The factorization and renormalization scale $\mu^2 = Q^2$ is varied by a factor of ± 2 in the simulation to derive an additional set of altered template shapes for $Wb\bar{b}$ events. Furthermore, the resulting relative bin-wise differences are also applied as a fully correlated uncertainty to the template shape of $Wc\bar{c}$ events.
A modified model describing QCD-multijet events is used to evaluate the shape uncertainty concerning its flavor composition. 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 verify the systematic effect on the shapes of the distributions caused by the modeling of mistagged light-quark jet events, an alternative $Wq\bar{q}$ model is utilized to create template distributions. This is realized by replacing the default $Wq\bar{q}$ model based on simulated events by a description on the basis of observed W+jets collision events without b-tagging requirements.
Since validation cross-checks in the 0-tagged sideband region identified not yet understood mismodeling in jets with lowest E_{T} in high η regions and in the distance ΔR between two jets in the $\eta - \phi$ plane, respectively, maximal shape systematics are assigned. Event weights are applied to the simulated 0-tagged events such, that the originally mismodeled simulation match the observed data distribution. These weights, now applied to the simulated events in the corresponding tagged samples, constitute shape uncertainties.

7.4.1 Systematic Normalization Rate Uncertainty

Each background process features an uncertainty on its production rate, Δ, given as the relative uncertainties on the estimated number of events of this process. Its

determination depends on the underlying method to estimate the process' yield. W+HF and QCD-multijet events are estimated based on the observed sideband rate of events and by a sideband-region fit to the observed missing transverse energy distribution, respectively. Those estimates are attached with large conservative uncertainties covering all imaginable effects: $\Delta(Wb\bar{b}+Wc\bar{c}) = 30\%$, $\Delta(Wc) = 30\%$, and $\Delta(\text{QCD}) = 40\%$. Well-understood processes, whose expected rates are derived from theoretical calculations, have rather small attached uncertainties. Additionally, the statistical uncertainty due to the limited amount of simulated events is incorporated as systematic uncertainty in each bin of all processes.
Tables 7.1-7.8 summarize the relative systematic rate uncertainties on the event prediction of the signal and background processes for all considered channels. During the evaluation of the likelihood function, each of those nuisance parameters is treated as 100% correlated throughout all considered channels and all affected processes.

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	7.0/-1.5 %	6.2/7.1 %	6.7/1.4 %	-7.7/-9.7 %
PDF	3.1/-3.5 %	1.7/-1.4 %	2.6/-2.8 %	1.9/-2.3 %
MC	2.0/-2.0 %	1.0/-1.0 %	1.7/-1.7 %	-2.7/2.7 %
ϵ_{evt}	4.2/-4.2 %	2.3/-2.3 %	3.6/-3.6 %	2.9/-2.9 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Δ	12.6/-12.6 %	12.4/-12.4 %	12.6/-12.6 %	12.4/-12.4 %
M_{top} 170/180	6.1/-5.3 %	9.5/-8.0 %	7.3/-6.2 %	7.8/-8.1 %
	Diboson	Z+jets	$Wq\bar{q}$	W+HF / QCD
ϵ_{evt}	7.6/-7.6 %	8.3/-8.3 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Δ	1.9/-1.9 %	10.8/-10.8 %	12.6/-12.6 %	±30% / ±40%

Table 7.1: Systematic rate uncertainties for TLC lepton events with 2 jets and 1 b tag

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	7.0/-1.5 %	6.2/7.1 %	6.7/1.6 %	-7.7/-9.7 %
PDF	3.1/-3.5 %	1.7/-1.4 %	2.6/-2.7 %	1.9/-2.3 %
MC	2.0/-2.0 %	1.0/-1.0 %	1.6/-1.6 %	2.7/-2.7 %
ϵ_{evt}	3.9/-3.9 %	1.5/-1.5 %	3.0/-3.0 %	2.6/-2.6 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Δ	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
M_{top} 170/180	6.1/-5.3 %	9.5/-8.0 %	7.3/-6.3 %	7.8/-8.1 %
	Diboson	Z+jets	$Wq\bar{q}$	W+HF / QCD
ϵ_{evt}	7.2/-7.2 %	7.6/-7.6 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Δ	1.9/-1.9 %	10.8/-10.8 %	12.6/-12.6 %	±30% / ±40%

Table 7.2: Systematic rate uncertainties for EMC lepton events with 2 Jets and 1 b tags

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	-0.9/-13.4 %	9.5/11.4 %	7.9/7.5 %	-7.5/-11.2 %
PDF	3.0/-3.3 %	1.7/-1.5 %	1.9/-1.7 %	1.9/-2.3 %
MC	2.0/-2.0 %	1.0/-1.0 %	1.2/-1.2 %	4.6/-4.6 %
ϵ_{evt}	10.0/-10.0 %	8.7/-8.7 %	8.9/-8.9 %	9.0/-9.0 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Δ	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
M_{top} 170/180	2.5/-7.3 %	9.4/-6.9 %	7.7/-6.7 %	9.9/-7.1 %
	Diboson	Z+jets	$Wq\bar{q}$	W+HF / QCD
ϵ_{evt}	9.8/-9.8 %	10.6/-10.6 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Double tag			22.0/-22.0 %	
Δ	1.9/-1.9 %	10.8/-10.8 %	12.6/-12.6 %	$\pm$30% / $\pm$40%

Table 7.3: Systematic rate uncertainties for TLC lepton events with 2 jets and 2 b tags

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	-1.0/-13.5 %	9.4/11.4 %	8.0/8.2 %	-7.4/-11.1 %
PDF	3.0/-3.3 %	1.7/-1.5 %	1.9/-1.7 %	1.9/-2.3 %
MC	2.0/-2.0 %	1.0/-1.0 %	1.1/-1.1 %	-4.6/4.6 %
ϵ_{evt}	10.0/-10.0 %	8.7/-8.7 %	8.9/-8.9 %	8.9/-8.9 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Δ	12.6/-12.6 %	12.4/-12.4 %	12.4/-12.4 %	12.4/-12.4 %
M_{top} 170/180	2.5/-7.3 %	9.4/-6.9 %	8.5/-7.0 %	9.9/-7.1 %
	Diboson	Z+jets	$Wq\bar{q}$	W+HF / QCD
ϵ_{evt}	9.3/-9.3 %	9.9/-9.9 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
double tag			22.0/-22.0 %	
Δ	3.2/-3.2 %	10.8/-10.8 %	12.6/-12.6 %	$\pm$30% / $\pm$40%

Table 7.4: Systematic rate uncertainties for EMC lepton events with 2 Jets and 2 b tags

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	-8.4/-3.1 %	-3.6/-17.4 %	-6.5/-8.6 %	-4.0/-6.7 %
PDF	3.2/-3.7 %	1.8/-1.5 %	2.7/-2.8 %	1.9/-2.3 %
MC	1.9/-1.9 %	1.5/-1.5 %	1.7/-1.7 %	-1.7/1.7 %
ϵ_{evt}	3.5/-3.5 %	2.3/-2.3 %	3.0/-3.0 %	2.3/-2.3 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Δ	12.6/-12.6 %	12.4/-12.4 %	12.6/-12.6 %	12.4/-12.4 %
M_{top} 170/180	6.2/-6.5 %	11.7/-8.6 %	6.4/-5.6 %	9.3/-8.4 %
	Diboson	Z+jets	$Wq\bar{q}$	W+HF / QCD
ϵ_{evt}	7.8/-7.8 %	7.8/-7.8 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Δ	1.9/-1.9 %	10.8/-10.8 %	12.6/-12.6 %	±30% / ±40%

Table 7.5: Systematic rate uncertainties for TLC lepton events with 3 jets and 1 b tag

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	-8.3/-3.1 %	-3.6/-17.4 %	-6.4/-9.0 %	-4.1/-6.7 %
PDF	3.2/-3.7 %	1.7/-1.5 %	2.6/-2.8 %	1.9/-2.3 %
MC	1.5/-1.5 %	1.9/-1.9 %	1.7/-1.7 %	2.7/-2.7 %
ϵ_{evt}	3.0/-3.0 %	1.5/-1.5 %	2.4/-2.4 %	2.5/-2.5 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Δ	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
M_{top} 170/180	6.2/-6.5 %	11.7/-8.6 %	8.5/-7.4 %	9.3/-8.4 %
	Diboson	Z+jets	$Wq\bar{q}$	W+HF / QCD
ϵ_{evt}	7.3/-7.3 %	14.2/-11.2 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Δ	1.9/-1.9 %	10.8/-10.8 %	12.6/-12.6 %	±30% / ±40%

Table 7.6: Systematic rate uncertainties for EMC lepton events with 3 Jets and 1 b tags

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	22.7/4.4 %	-3.1/-16.2 %	8.2/-7.2 %	-3.9/-9.3 %
PDF	3.7/-4.1 %	1.8/-1.5 %	2.6/-2.6 %	1.9/-2.3 %
MC	1.9/-1.9 %	1.5/-1.5 %	1.7/-1.7 %	2.0/-2.0 %
ϵ_{evt}	9.1/-9.1 %	8.8/-8.8 %	8.9/-8.9 %	9.1/-9.1 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Δ	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
M_{top} 170/180	7.8/-3.0 %	9.1/-11.1 %	6.9/-6.0 %	9.3/-9.5 %
	Diboson	Z+jets	$Wq\bar{q}$	W+HF / QCD
ϵ_{evt}	10.8/-10.8%	11.1/-11.1 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Double tag			22.0/-22.0%	
Δ	1.9/-1.9 %	10.8/-10.8 %	12.6/-12.6 %	±30% / ±40%

Table 7.7: Systematic rate uncertainties for TLC lepton events with 3 jets and 2 b tags

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	22.8/4.5 %	-3.1/-16.2 %	8.0/-7.3 %	-3.9/-9.3 %
PDF	3.7/-4.1 %	1.8/-1.5 %	2.6/-2.6 %	1.9/-2.3 %
MC	1.5/-1.5 %	1.9/-1.9 %	1.7/-1.7 %	-2.0/2.0 %
ϵ_{evt}	9.0/-9.0 %	8.8/-8.8 %	8.9/-8.9 %	9.1/-9.1 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Δ	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
M_{top} 170/180	7.8/-3.0 %	9.1/-11.1 %	8.5/-7.6 %	9.3/-9.5 %
	Diboson	Z+jets	$Wq\bar{q}$	W+HF / QCD
ϵ_{evt}	11.5/-11.5 %	10.8/-10.8 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
double tag			22.0/-22.0 %	
Δ	1.7/-1.7 %	10.8/-10.8 %	12.6/-12.6 %	±30% / ±40%

Table 7.8: Systematic rate uncertainties for EMC lepton events with 3 Jets and 2 b tags

The majority of systematic rate uncertainties are only determined for single top-quark and $t\bar{t}$ events, since the rates of the main background processes are estimated with observed collision data, whose large uncertainties are expected to cover all possible remaining sources. Since most of the rate uncertainties also have correlated shape uncertainties attached, it is not possible to judge the impact of a systematic source to the outcome of the analysis due to possible cancellation effects. However, one known dominating source of rate and shape uncertainty, JES, is evaluated for all processes using simulated events, as can be seen in tables 7.9 and 7.10.

process	2jets 1tag	2jets 2tags	3jets 1tag	3jets 2tags
t-ch	-1.1/0.6 %	4.8/-3.5 %	-10.4/10.6 %	-5.7/4.3 %
s-ch	-0.1/-0.6 %	1.2/-1.9 %	-8.3/9.4%	-7.2/7.4 %
single-top	-0.8/0.2 %	1.8/-2.2 %	-9.1/9.9 %	-6.6/6.1 %
$t\bar{t}$	9.8/-9.4 %	8.1/-7.5 %	4.6/-5.1 %	5.4/-5.2 %
$Wc\bar{c}$+$Wb\bar{b}$	7.0/-6.9 %	10.8/-10.6 %	8.4/-7.7 %	11.0/-12.1 %
Wc	7.0/-6.3 %	11.3/-10.3 %	8.2/-6.9 %	13.9/-15.8 %
Z+jets	-5.3/5.4 %	5.0/-5.0 %	-10.8/14.0 %	-5.9/7.2 %
Diboson	-2.7/1.7 %	-3.0/1.5 %	-12.4/11.9 %	-12.0/12.0 %

Table 7.9: Systematic JES down/up rate uncertainties for TLC lepton events

process	2jets 1tag	2jets 2tags	3jets 1tag	3jets 2tags
t-ch	-2.3/2.4 %	1.9/-1.7 %	-10.1/12.7 %	-3.9/5.8 %
s-ch	-1.7/0.4 %	-1.3/0.4 %	-9.8/8.6 %	-6.9/10.4 %
single-top	-2.1/1.7 %	-0.9/0.1 %	-10.0/11.0 %	-5.6/8.4 %
$t\bar{t}$	9.1/-8.9 %	6.1/-6.4 %	3.8/-4.2 %	4.1/-4.5 %
$Wc\bar{c}$+$Wb\bar{b}$	5.5/-5.4 %	7.0/-6.3 %	6.4/-6.4 %	10.4/-7.1 %
Wc	5.7/-4.9 %	6.9/-4.6 %	0.3/-12.7 %	10.1/-17.3 %
Z+jets	-2.8/2.8 %	-2.8/2.8 %	-11.2/14.2 %	-10.0/13.1 %
Diboson	-4.5/2.2 %	-4.5/2.3 %	-13.0/13.0 %	-13.0/13.0 %

Table 7.10: Systematic JES down/up rate uncertainties for EMC lepton events

7.4.2 Systematic Template Shape Uncertainty

For the usage in the binned likelihood function, all shape uncertainties are smoothed using a 5-bin median filter. This means, that the content n_i of bin i of the relative difference between the systematically shifted and the default distribution is given by $median(n_{i-2}, n_{i-1}, n_i, n_{i+1}, n_{i+2})$. As one can see from the formula this can be done for all bins, except for the first two and last two bins. The median filter has the advantage to remove extreme outliers and high frequency noise very efficiently, but doesn't change already smooth distributions. In figure 7.10 a comparison between the original and the smoothed relative ratio plots is shown. In the following, figures 7.11-7.19 exemplarily show all considered systematic shape uncertainties in the dominant channel of TLC events with 2 jets and 1 tag. Comparable but individually obtained shape uncertainties are also applied to all seven remaining channels. The shape uncertainty with most impact to the analysis' outcome is the ALPGEN $Wb\bar{b}$ factorization/renormalization scale Q^2 uncertainty, followed by the JES uncertainty.

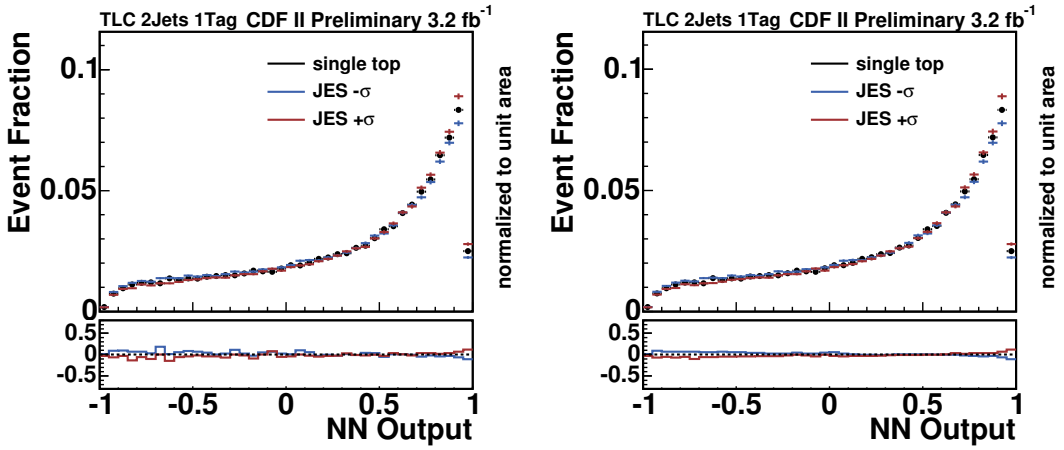


Figure 7.10: Comparison of the original (left) and smoothed (right) relative difference between the JES shifted distribution and the default distribution of single top-quark events, visible in the lower sub-figure

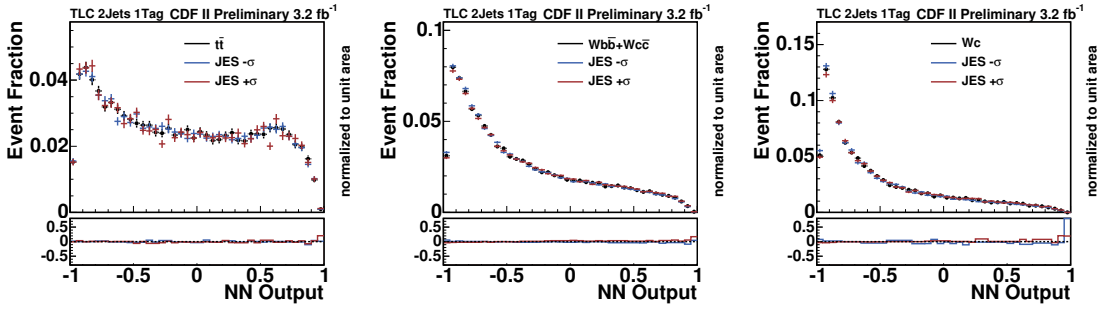


Figure 7.11: Shape systematics due the uncertainty on the jet energy scale (JES) correction for $t\bar{t}$ (left), $Wb\bar{b} + Wc\bar{c}$ (middle), and Wc (right) events

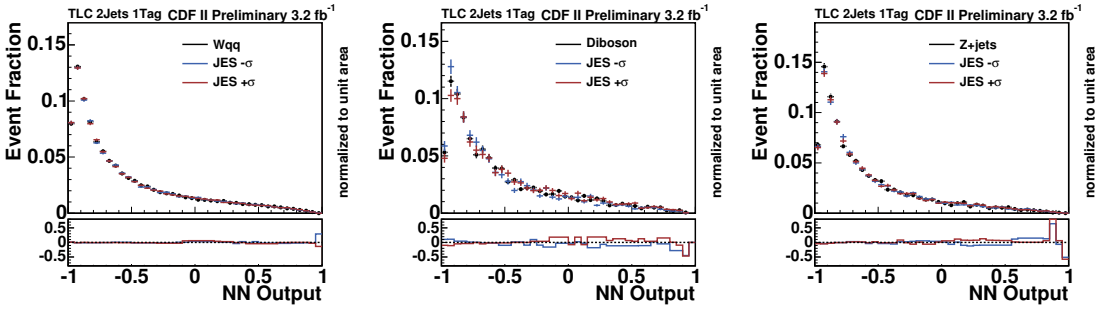


Figure 7.12: Shape systematics due the uncertainty on the jet energy scale (JES) correction for $Wq\bar{q}$ (left), Diboson (middle), and Z+jets (right) events

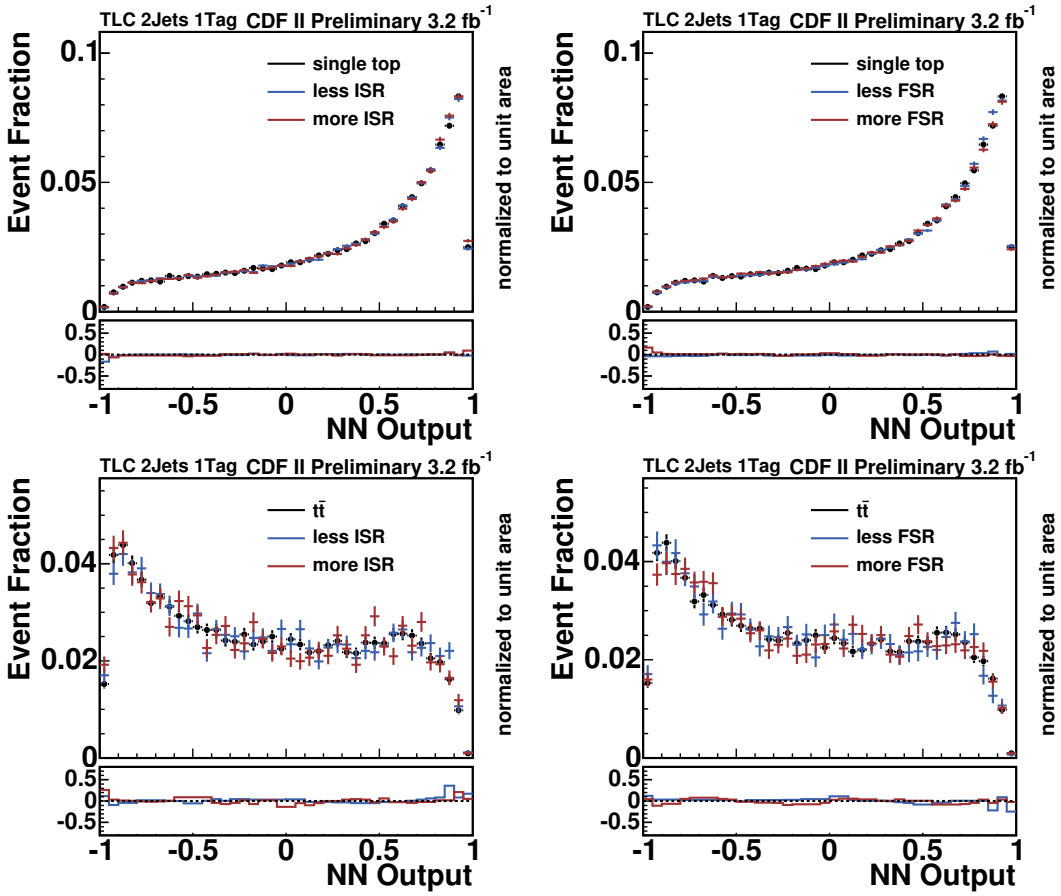


Figure 7.13: Shape systematics due the uncertainty on ISR (left) and FSR (right) for single top-quark (top) and $t\bar{t}$ (bottom) production

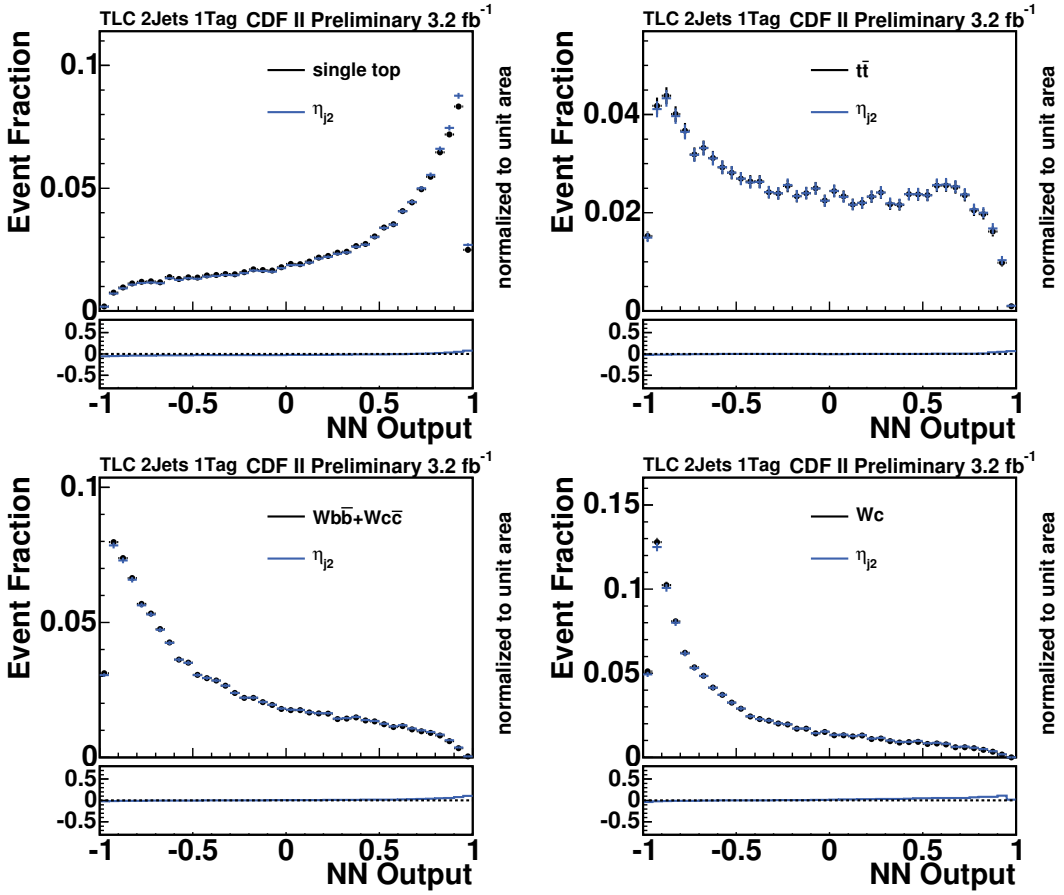


Figure 7.14: Shape systematics due the mismodeling of η of the second-leading jet for single top-quark (top left), $t\bar{t}$ (top right), $Wb\bar{b}+Wc\bar{c}$ (bottom left), and Wc (bottom right) events

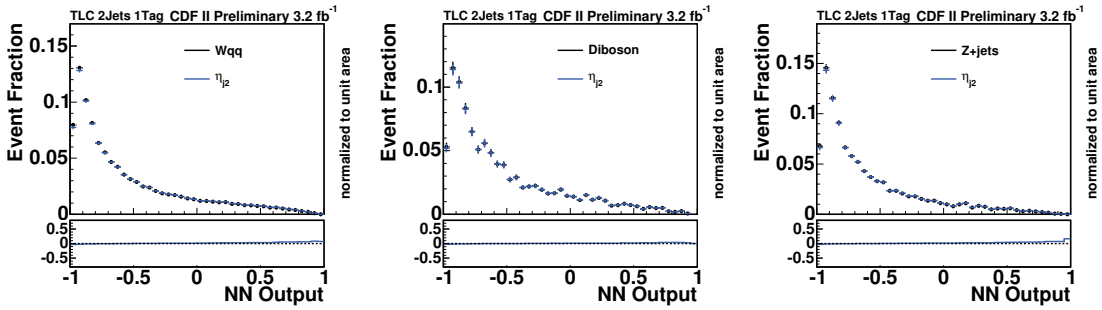


Figure 7.15: Shape systematics due the mismodeling of η of the second-leading jet for $Wq\bar{q}$ (left), Diboson (middle), and Z+jets (right) events

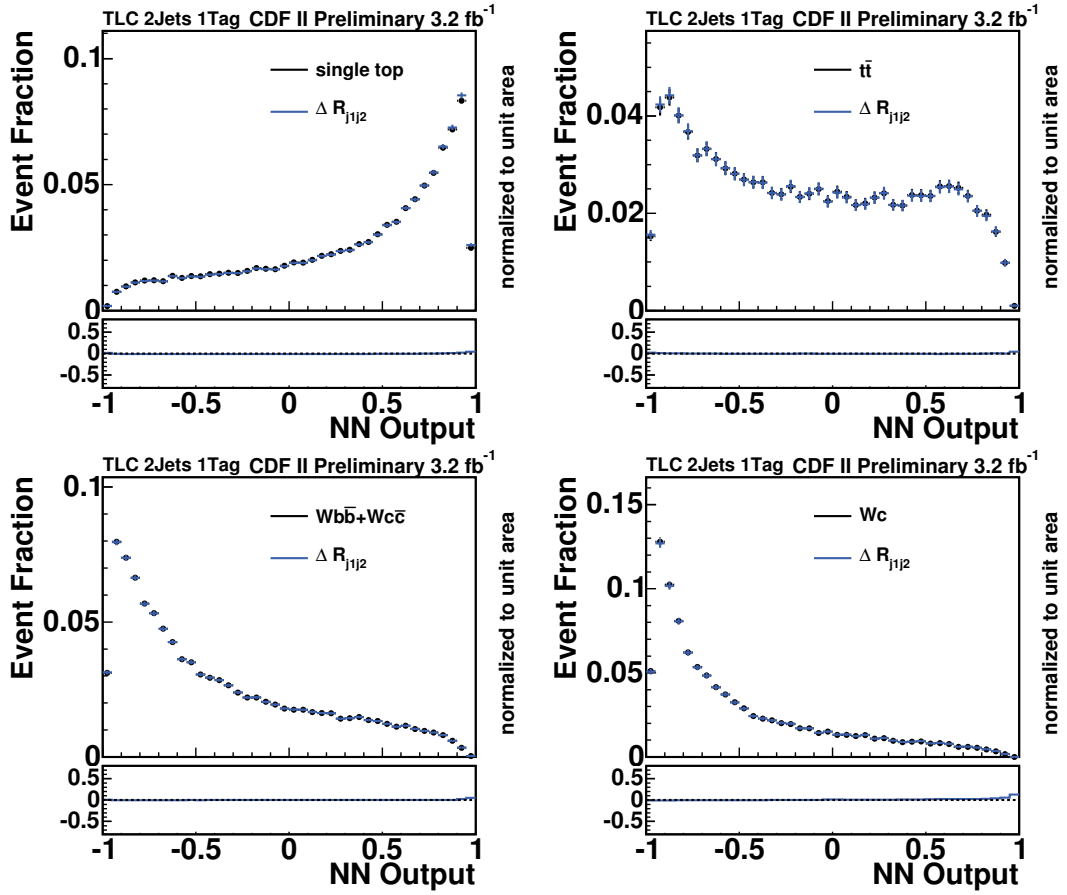


Figure 7.16: Shape systematics due the mismodeling of ΔR between both jets for single top-quark (top left), $t\bar{t}$ (top right), $Wb\bar{b}+Wc\bar{c}$ (bottom left), and Wc (bottom right) events

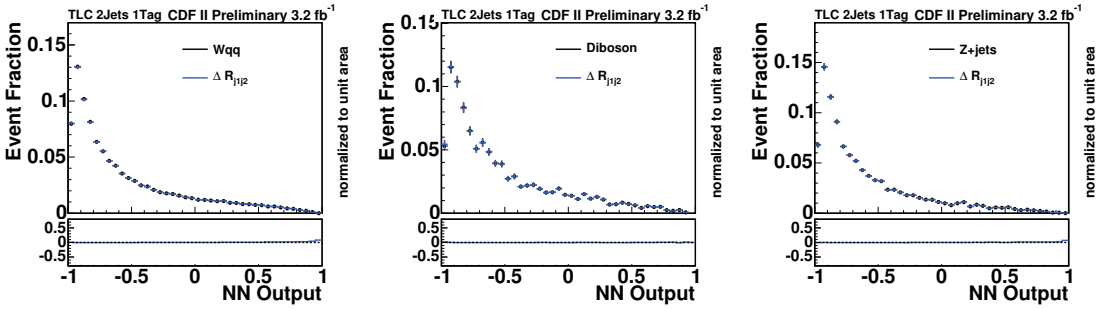


Figure 7.17: Shape systematics due the mismodeling of ΔR between both jets for $Wq\bar{q}$ (left), Diboson (middle), and Z+jets (right) events

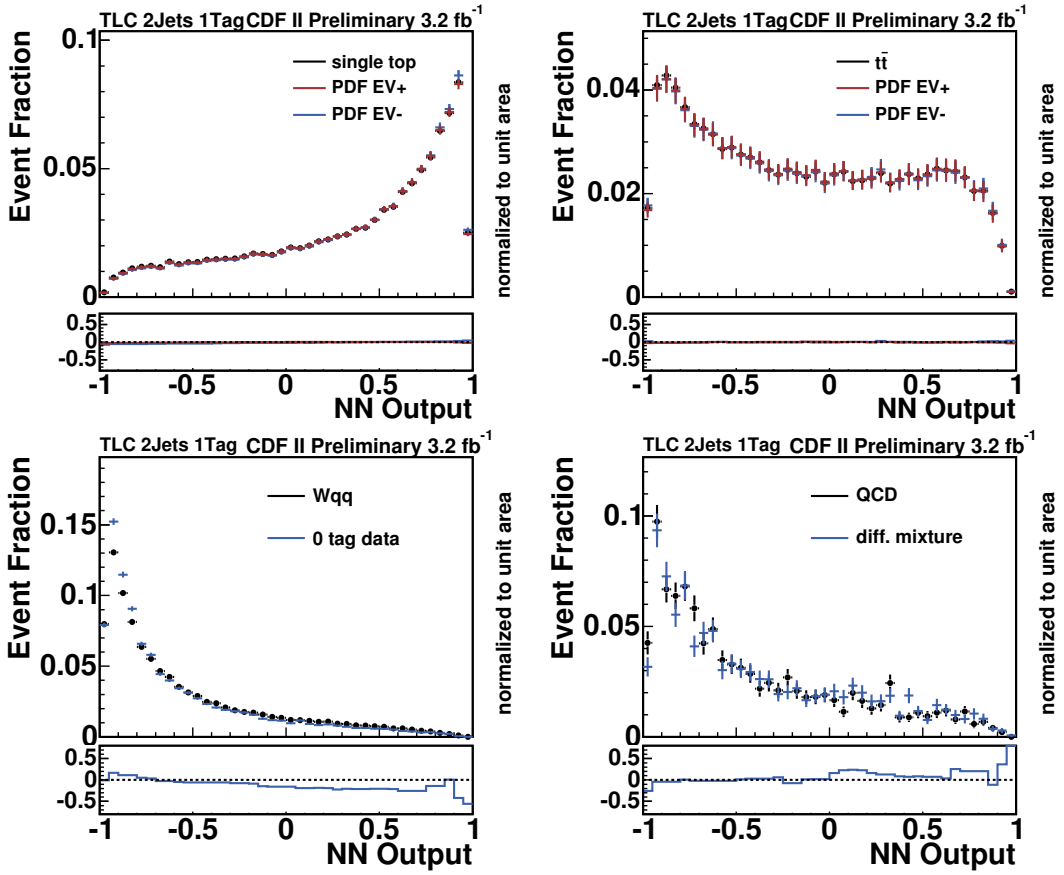


Figure 7.18: Shape systematics due the variations of the eigenvalues (EV) of the CTEQ6M PDF set for single top-quark (top left) and $t\bar{t}$ (top right) production, due to the choice of the $Wq\bar{q}$ model (bottom left), and due to the choice of the flavor composition of the QCD model (bottom right)

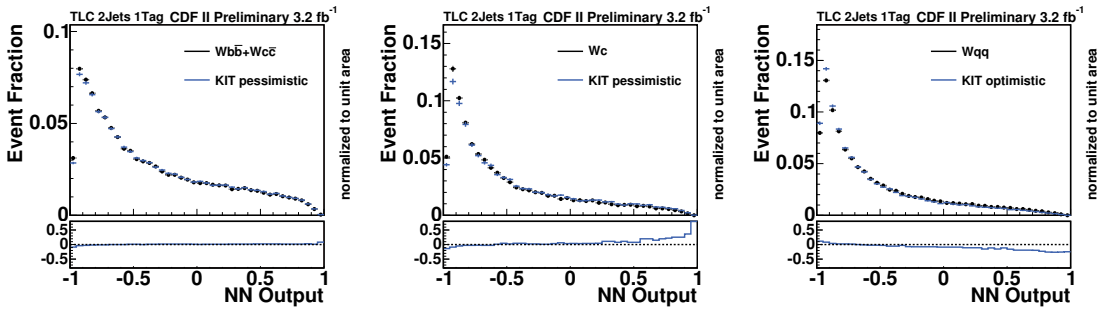


Figure 7.19: Shape systematics due the uncertainty on the KIT flavor separator $Wq\bar{q}$ corrections for a pessimistic $Wb\bar{b}+Wc\bar{c}$ (left) and Wc (middle), and an optimistic $Wq\bar{q}$ (right) scenario

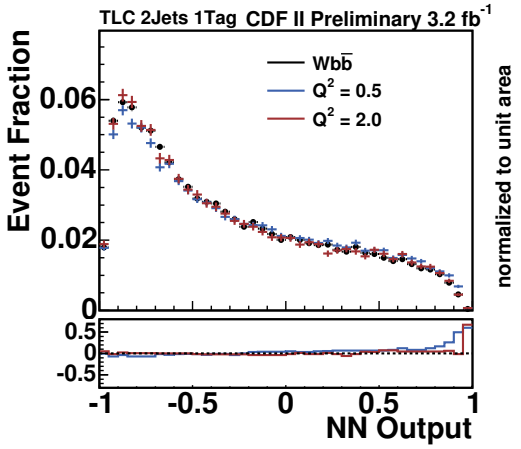


Figure 7.20: Shape systematics due the uncertainty on the ALPGEN factorization/renormalization scale Q^2 for the $Wb\bar{b}$ process. The same relative uncertainty is additionally applied to $Wc\bar{c}$ events.

7.5 Results

First, the discovery potential of the analysis performing a combined search for s- and t-channel single top-quark production is evaluated. This is done by computing the expected significance of the combined signal through hypothesis testing including all considered rate and shape uncertainties. Afterwards, the observed candidate events are compared to the expected distributions and the observed significance is determined. The corresponding combined top-quark production cross-sections is subsequently extracted.

Finally, the second analysis optimized for a simultaneous extraction of both the s- as well as the t-channel single top-quark production cross-section is conducted.

7.5.1 Significance of Combined Single Top-Quark Search

For the combined search, $\hat{p}_{\text{exp}} = 0.11 \times 10^{-6}$ is found, which corresponds to an expected significance of 5.2σ. The resulting test statistic distributions including all rate and shape systematics are shown in figure 7.21. The found value of $\hat{p}_{\text{exp}}$ can be interpreted as follows: assuming the predicted SM 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 5.2σ in case of the combined search. Thus, assuming SM circumstances, the discovery potential of the analysis is large enough that one expects to achieve first observation of single top-quark production with a probability above 50%.

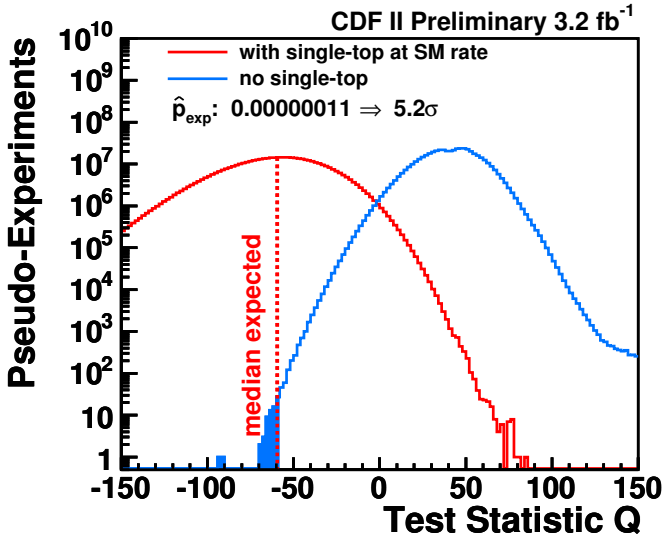


Figure 7.21: Distributions of Q-values for two ensemble tests, one with single top-quark events present at the expected SM rate, one without any single top-quark events for the combined search

After the expected sensitivity has been determined, the neural network classifiers are applied to the observed candidate events. Figures 7.22-7.23 depict the observed candidate events compared to the expected distributions. In figure 7.24 the distributions of all 8 channels are added together.

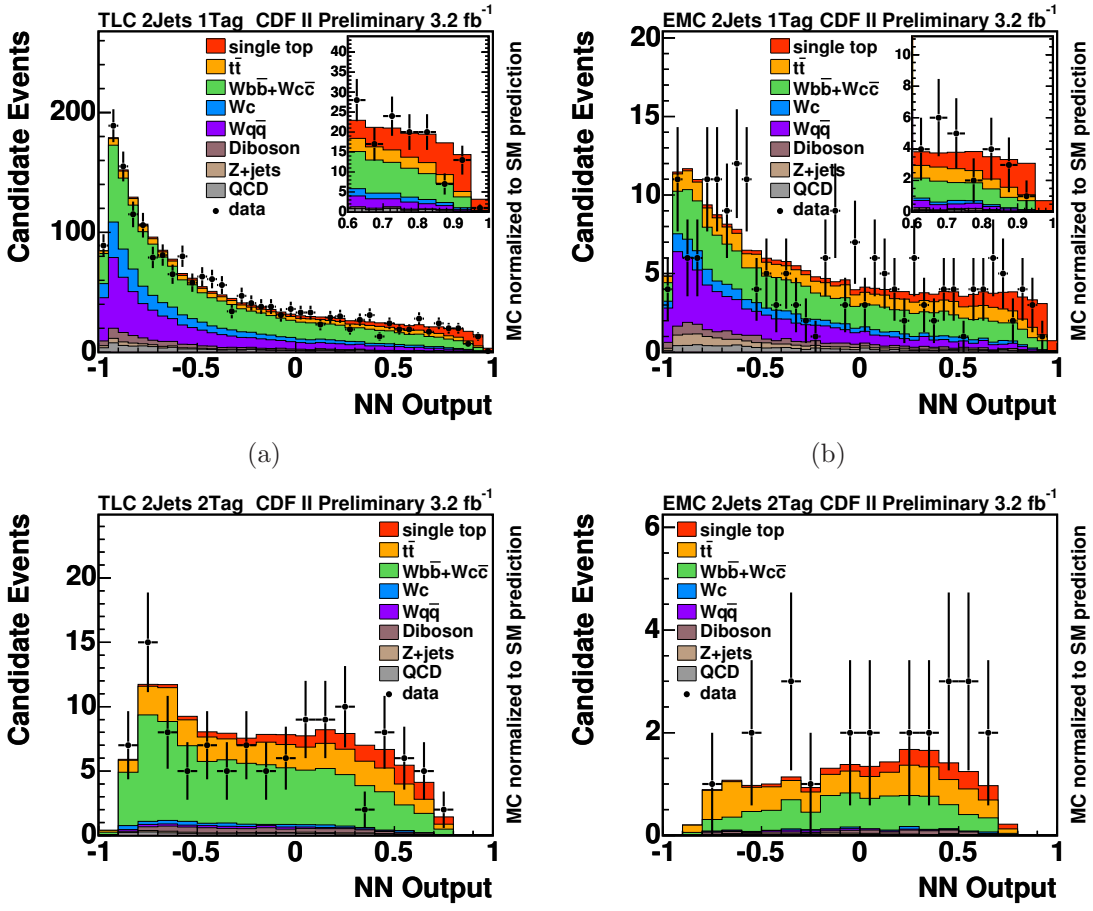


Figure 7.22: Predicted distribution in the two-jet one-tag signal region (a),(b) and in the two-jet two-tag signal region (c),(d) for TLC (left) and EMC (right) events.

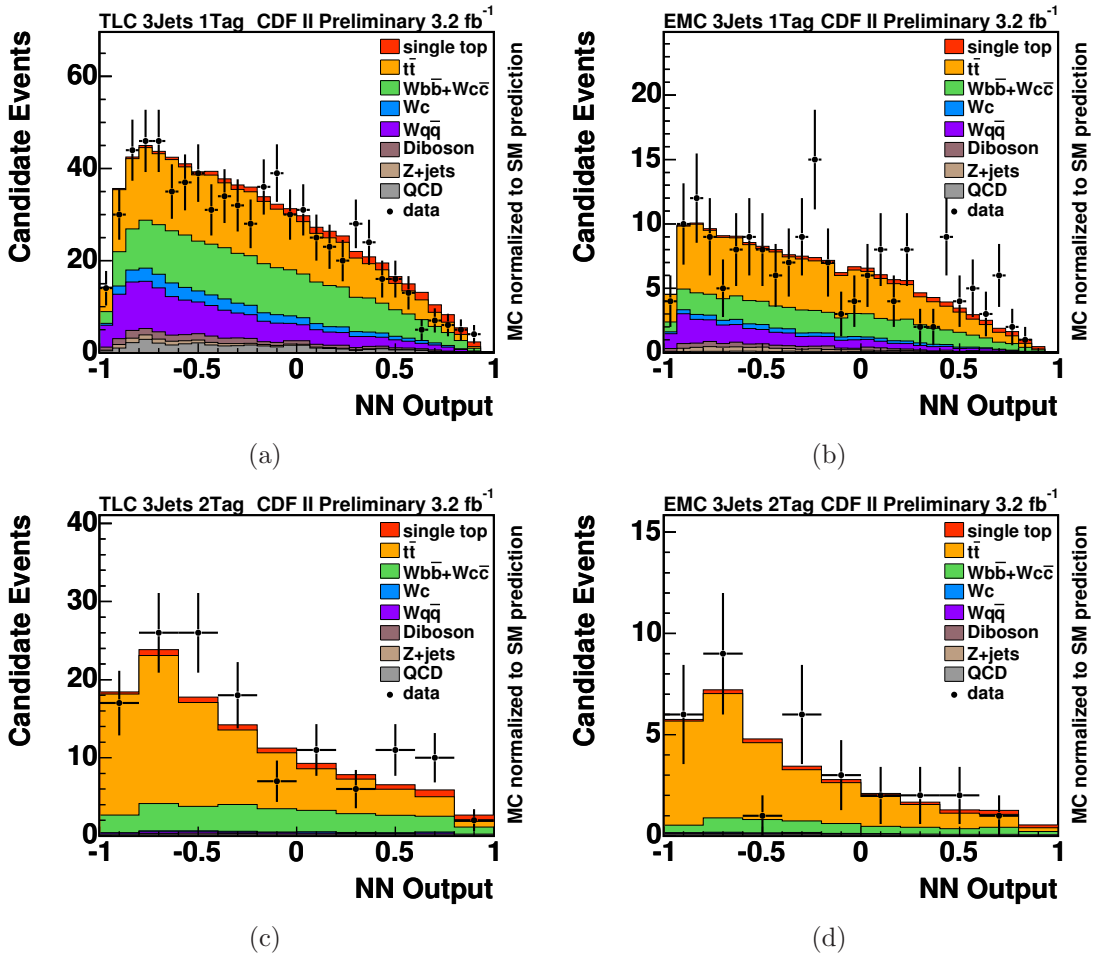


Figure 7.23: Predicted distribution in the three-jet one-tag signal region (a),(b) and in the three-jet two-tag signal region (c),(d) for TLC (left) and EMC (right) events.

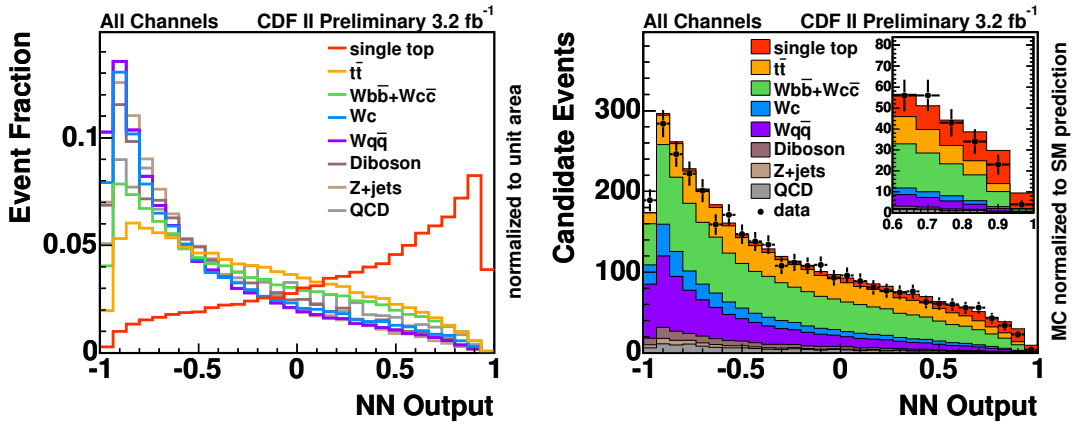


Figure 7.24: Templates for signal and background processes (left) and the predicted and observed distributions (right) of all 8 channels used in the combined search

Finally, the templates are fitted to the observed distributions to determine the observed significance of the combined signal. The simultaneous likelihood fit to the neural network outputs in all eight channel yields an observed Q-value of -27.2, resulting in an observed $\hat{p}_{\text{obs}} = 240.14 \times 10^{-6}$. This corresponds to an observed significance of 3.5σ. Figure 8.8 compares the observed Q-value with the Q-value distributions of both tested hypotheses. The collision data seems to strongly prefer the test hypothesis H_1 including SM predicted single top-quark production. But the observed candidate events don't accumulate distinct enough in the signal-like region allowing to proclaim the observation of single top-quark production, merely a strong evidence can be announced.

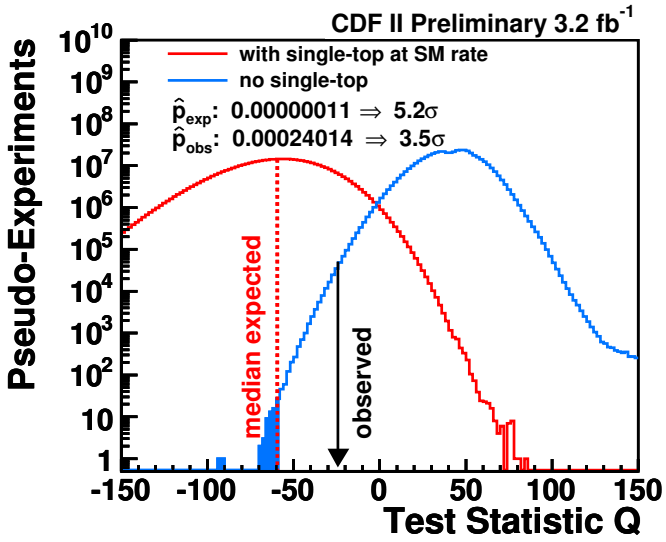
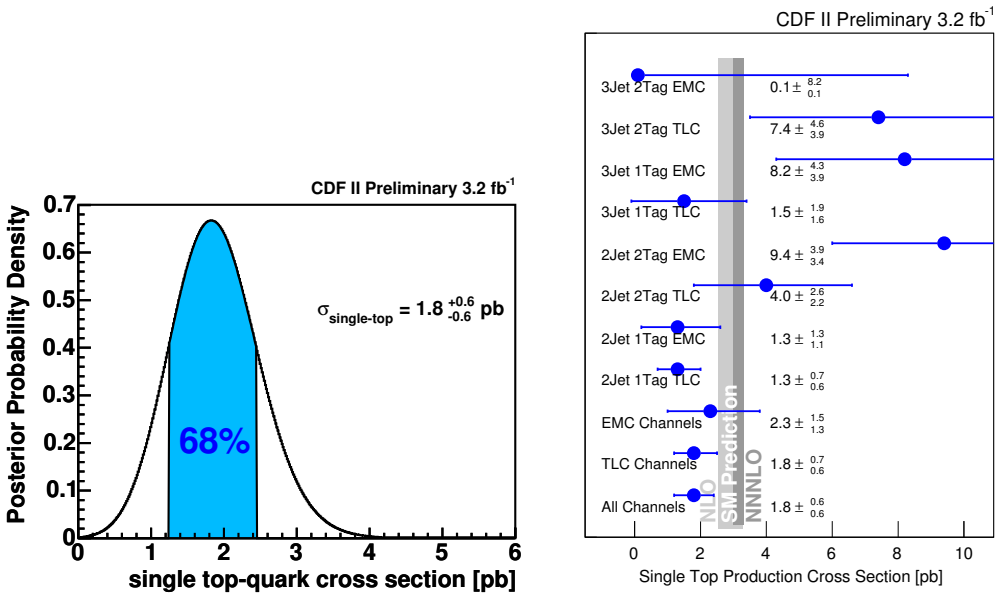


Figure 7.25: Comparison of observed Q-value to both distributions of the null hypothesis H_0 with no single top-quark production as well as the test hypothesis H_1 with signal at the SM rate.

7.5.2 Combined s- and t-channel Single Top-Quark Production Cross-Section Measurement

To extract the combined s- and t-channel single top-quark production cross-section, the reduced likelihood function is determined including all systematic rate and shape uncertainties by marginalization. It is then converted into a posterior probability density as a function of the single top-quark production cross-section and the most probable value and the corresponding uncertainties at the 68% C.L. are extracted. Figure 7.26(a) shows the extracted posterior probability density, which yields the extracted cross section for single top-quark production of $\sigma_{\text{combined}} = 1.8^{+0.6}_{-0.6}$pb. This is compatible with a downward fluctuation of about one to two standard deviations compared to the theoretical NLO prediction of the combined signal cross-section of $\sigma^{\text{theo}}_{\text{combined}} = 2.9 \pm 0.4$ pb [78, 79]. Figure 7.26(b) illustrates the signal cross-section extraction in each channel individually. It is obvious, that the three most dominating channels of TLC and EMC events with two jets and one tag and TLC events with three jets and one tag are consistent well below the expectation, respectively.



(a) (b)

Figure 7.26: Extracted posterior probability density of all channels (left-hand side) and the results for the eight different channels and the final result of the simultaneous fit in all channels compared to theoretical NLO [78, 79] and NNNLO [83] predictions (right-hand side)

7.5.3 Variables in the High-Output Region of the Combined Search

By requiring a NN output above 0.8, see inset of figure 7.27(a), in the dominating TLC and EMC candidate event sample with 2 jets and 1 b tag, about 33 single top-quark and 32 background events are expected, yielding $\frac{S}{B} = 1$, whereas 49 candidate events are observed in collision data. The corresponding high NN output distributions of the reconstructed top-quark mass, see figure 7.27(b), the KIT flavor separator, see figure 7.27(c), and the product of the lepton-charge and the pseudorapidity of the light quark, see figure 7.27(d), are shown.

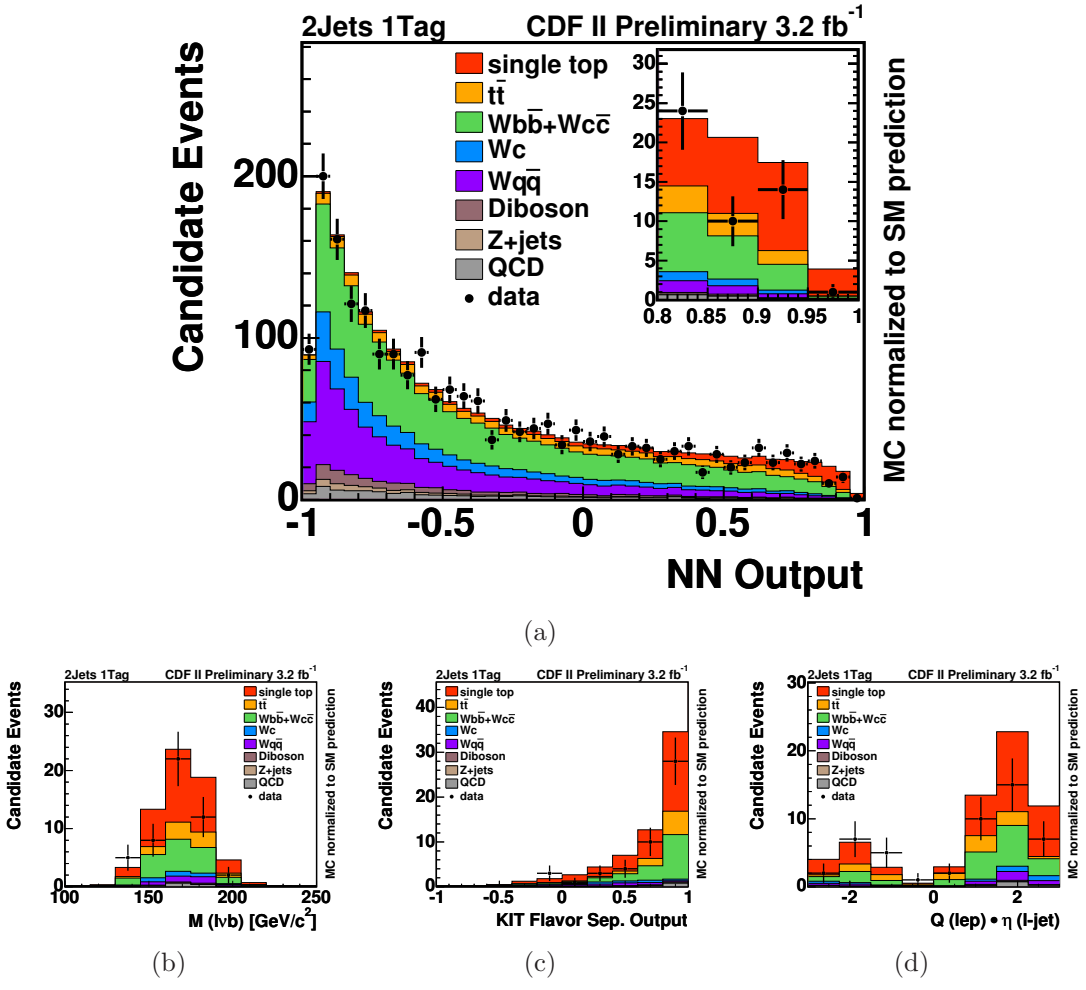


Figure 7.27: Signal-like events with a NN output above 0.8, see inset in (a), are depicted as reconstructed top-quark mass (b), as KIT flavor separator output (c), and as product of the lepton-charge and the pseudorapidity of the light quark (d).

7.5.4 Separate s- and t-channel Single Top-Quark Production Cross-Section Measurement

Contrary to the measurement of the combined single top-quark production cross-section, the separate search simultaneously extracts both the s- and t-channel production cross-section independently. This is done by a determination of both signal parameters with a two-dimensional reduced likelihood function $L_{\text{red}}(\beta_1, \beta_2)$, which is obtained via marginalization. Instead of one combined signal template, two independent signal templates are fitted together with all background templates to the candidate events in all channels. An optimized 2D discriminant, which improves the performance to separate the two single top-quark production modes in the dominant channels with 2 jets and 1 b tag, is unwinded bin-by-bin into a 1D distribution to be fitted simultaneously with all remaining channels. The expected and observed neural network output distribution of the candidate events of the separate search are shown in figure 7.28 and 7.29.

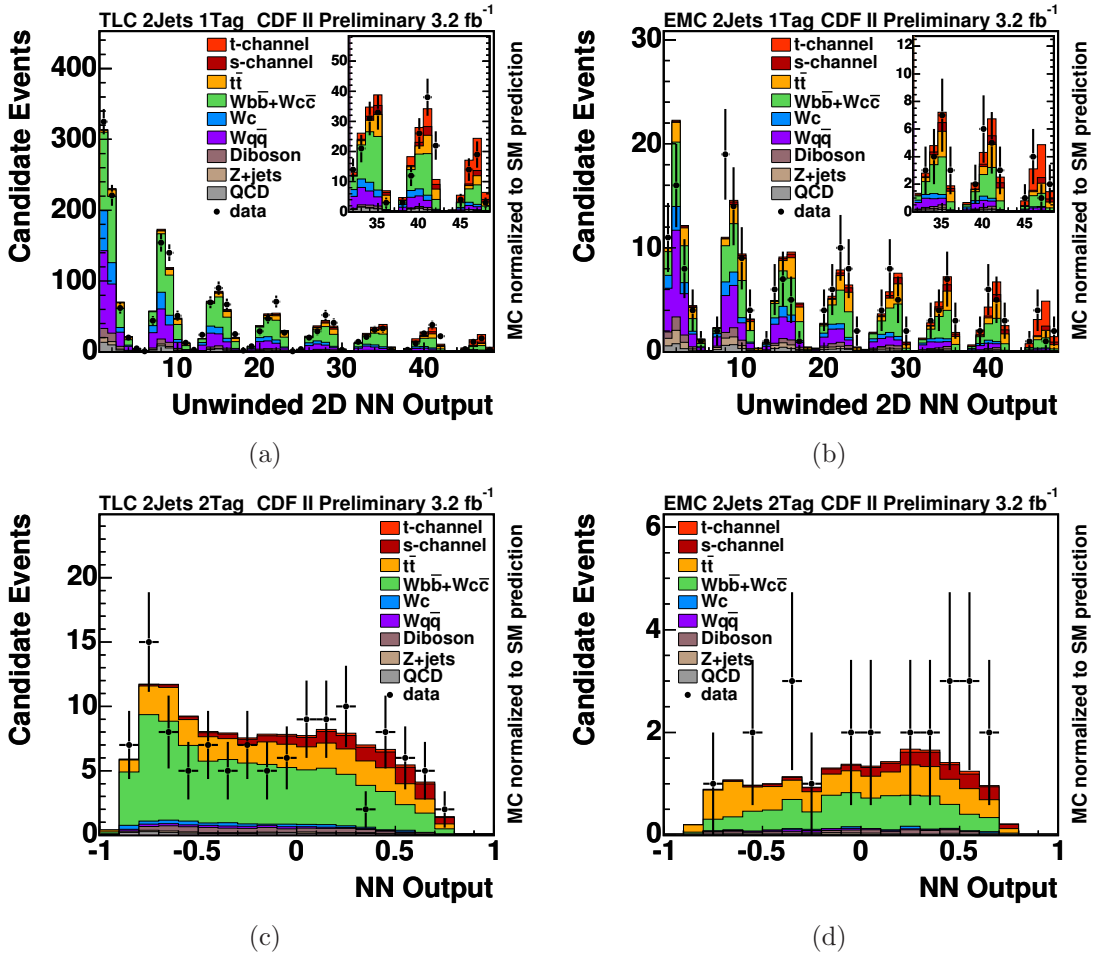


Figure 7.28: The expected and observed unwinded 2D NN output in the two-jet one-tag category (top) and the NN output of the s-channel neural network in the two-jet two-tag category (bottom) for TLC (left) and EMC (right) events

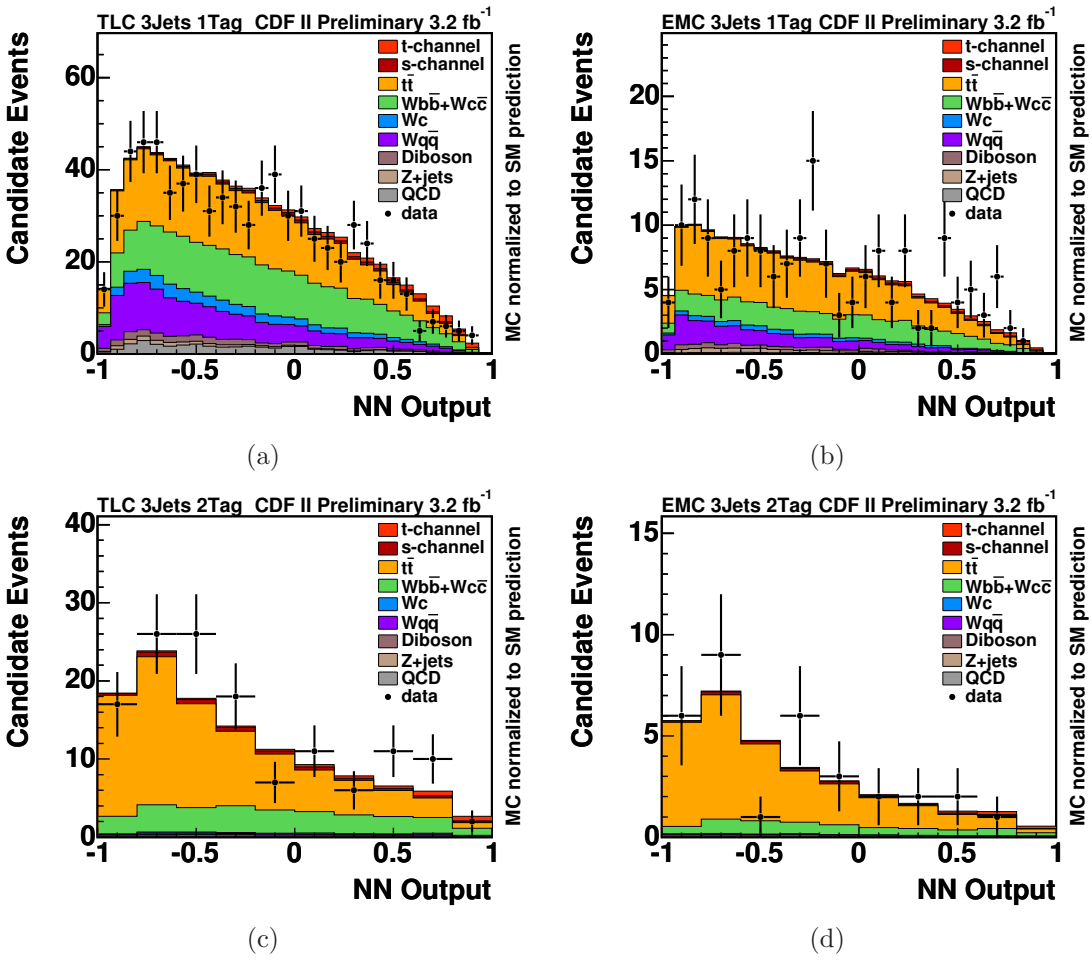


Figure 7.29: The expected and observed NN output of the t-channel neural network in the three-jet one-tag category (top) and of the t-channel neural network in the three-jet two-tag category (bottom) for TLC (left) and EMC (right) events

For the separate search, the fit yields the cross sections $\sigma_t = 0.7^{+0.5}_{-0.5}$ pb for t-channel and $\sigma_s = 2.0^{+0.7}_{-0.6}$ pb for s-channel single top-quark production, respectively. The fit result is depicted in figure 7.30 showing the difference between the logarithm of the reduced likelihood function and its minimum in the plane of single top-quark t-channel versus s-channel cross-sections. The minimum represents the most probable values and is indicated by the black dot. The error bars quote the 1σ, 2σ, and 3σ uncertainties ($\Delta ln(L)$ of 0.50, 2.0, and 4.5) on the fitted s- and t-channel production cross-sections. The true values of both cross sections have a probability of 68.3%, 95.5%, and 99.7% to be found in the region comprised by the corresponding contours ($\Delta ln(L)$ of 1.15, 3.09, and 5.92), respectively. The values predicted by two different theoretical calculations [78, 79, 83], within their uncertainties, are illustrated by the blue rectangles.
The separate result seems to be compatible with a downward fluctuation for the t-channel cross-section and with an upward fluctuation for the s-channel cross-section, respectively, both at the order of about 2σ. Since both separate signal cross-sections are anti-correlated, this is in good agreement with the extracted combined signal.

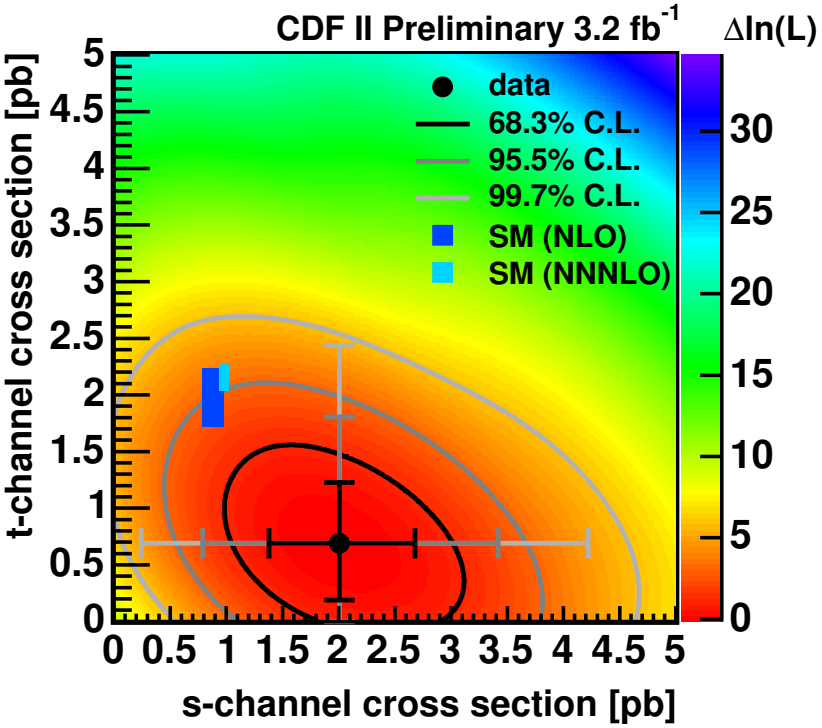


Figure 7.30: The likelihood fit estimate for the simultaneous s- and t-channel production cross-section measurement. The contours of the 1σ ($\Delta ln(L) = 1.15$), 2σ ($\Delta ln(L) = 3.09$), and 3σ ($\Delta ln(L) = 5.92$) uncertainties are valid for both production channels simultaneously. The error bars represent the 1σ ($\Delta ln(L) = 0.5$), 2σ ($\Delta ln(L) = 2.0$), and 3σ ($\Delta ln(L) = 4.5$) uncertainties of the given production channel without any assumptions on the other production channel.

Chapter 8

First Observation of Single Top-Quark Production

Within the CDF Collaboration, there are six separate analyses searching for electroweak single top-quark production. Five of them, including the NN one described in chapter 7, share the lepton+jets collision data set corresponding to an integrated luminosity of 3.2 fb^{-1}, which is selected for having a reconstructed electron or muon, missing transverse energy, and two or three jets with at least one of which is consistent with a b-quark jet. The neural network (NN) [157], matrix element (ME) [158], and boosted decision tree analyses (BDT) [159] use exactly the same events including both TLC and EMC lepton categories. The likelihood function analysis (LF) [160] uses only TLC events with two jets, one tag, and TLC events with three jets and one or two tags. Another likelihood function analysis optimized for s-channel signal (LFS) [161] considers only TLC events consistent with two jets and two tags.
Furthermore, a neural-network based analysis of collision events with missing transverse energy $\not\!\!E_{\text{T}}$ and jets (MJ) [162] adds about 30% to the signal acceptance by using an orthogonal sample of collision events corresponding to an integrated luminosity of 2.1 fb^{-1}. The MJ analysis is optimized to pick up signal candidate events with two or three jets and at least one b tag, which don't contain a reconstructed electron or muon, including events where a hadronically decaying tau is present.

Since the lepton+jets analyses are correlated among each other at the order of about 70%, additional gain in sensitivity is expected from a combination. Therefore, the lepton+jets NN, ME, BDT, LF, and LFS analyses are combined using a super discriminant technique (SD) similar to the one applied in CDF's published evidence for single top-quark production [35, 163]. The SD method uses neural networks trained to separate signal from background processes taking as inputs the discriminant outputs of the five analyses for each event.
Finally, a simultaneous fit over the two exclusive lepton+jets SD and $\not\!\!E_{\text{T}}$+jets MJ analyses will further increase the overall signal acceptance and will provide the final results of the search for combined s- and t-channel single top-quark production at CDF.

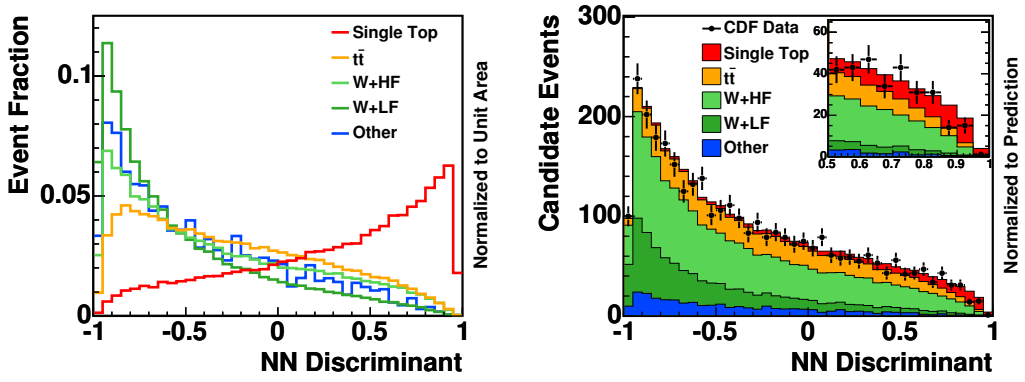
8.1 Lepton+Jets Analyses

Each of the five individual lepton+jets analyses use the same collision data sample of 3.2 fb^{-1}, the same event selection requirements, candidate event yield estimate, and systematic uncertainties. The NN, ME, and BDT analyses use all eight considered channels including both TLC and EMC lepton categories, while the LF and LFS analyses complement one another to cover all TLC only channels:

- TLC, 2-jet, 1-tag: NN, ME, BDT, LF
- TLC, 2-jet, 2-tag: NN, ME, BDT, LFS
- TLC, 3-jet, 1-tag: NN, ME, BDT, LF
- TLC, 3-jet, 2-tag: NN, ME, BDT, LF
- EMC, 2-jet, 1-tag: NN, ME, BDT
- EMC, 2-jet, 2-tag: NN, ME, BDT
- EMC, 3-jet, 1-tag: NN, ME, BDT
- EMC, 3-jet, 2-tag: NN, ME, BDT

8.1.1 Neural Network Analysis

For the sake of completeness, the NN analysis is listed here, even though a detailed description can be found in chapter 7. The NN analysis is, in common with the BDT one, the a priori most sensitive with an expected significance of 5.2σ. The NN discriminant of the sum of all eight considered channels is illustrated in figure 8.1.



(a) (b)

Figure 8.1: The power of the NN discriminant is visible in the template shape comparison (a), while in (b) the predicted distribution is compared with the observed candidate events.

8.1.2 Matrix Element Analysis

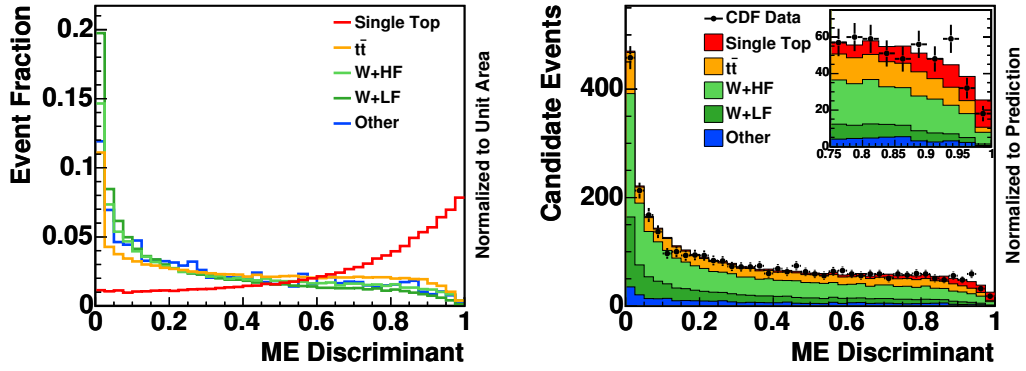
The main idea of a ME analysis is to calculate the event probability density (EPD) that a given set of four-vectors of the charged lepton and the jets resulted from a given underlying physical signal or background process [56,57,164]. The probability density is given by

$$P(x) = \frac{1}{\sigma} \int 2\pi^4 |M|^2 \frac{f(x_1)}{|E_{q1}|} \frac{f(x_2)}{|E_{q2}|} W(y,x) d\Phi_n dE_{q1} dE_{q2} \ , \tag{8.1}$$

where σ is the total cross section, $|M|$ is the matrix element, $f(x_i)$ and E_{qi} are the PDFs and energies of the incident partons, $W(y,x)$ is the transfer function between partonic and measured quantities, and $d\Phi_n$ represents the n-body phase space. Event probabilities are derived by calculating LO matrix elements for both the s- and t-channel signal, as well as for $Wb\bar{b}$, $Wc\bar{c}$, Wc, $Wq\bar{q}$ and $t\bar{t}$ background processes and by integrating over the whole phase space. The probability densities for the two signal processes are added to form a combined signal probability density. These event probabilities and the output of the KIT flavor separator (b), scaled to $0 \leq b \leq 1$ are used to construct a discriminant for each event:

$$EPD = \frac{b \cdot P_{signal}}{b \cdot P_{signal} + b \cdot (P_{Wb\bar{b}} + P_{t\bar{t}}) + (1-b) \cdot (P_{Wc\bar{c}} + P_{Wc} + P_{Wq\bar{q}})} \tag{8.2}$$

Figure 8.2 shows the resulting simulated ME discriminant template shapes and the ME discriminant applied to collision data compared to the prediction.



(a) (b)

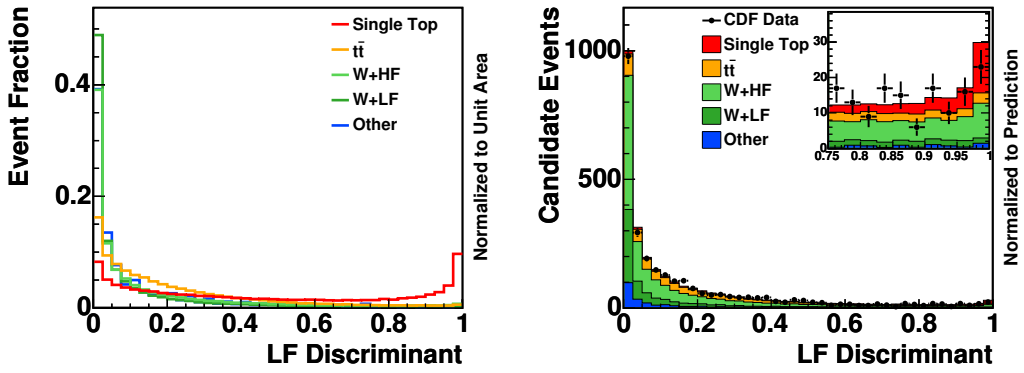
Figure 8.2: The power of the ME discriminant is visible in the template shape comparison (a), while the predicted discriminant distribution is compared with the observed candidate events (b).

8.1.3 Likelihood Function Analysis

The LF analysis combines different input variables with a LEP style projective likelihood function in order to build a discriminant [58]. In a first step each input variable i is filled in a histogram with n_i bins. The number of entries in each bin is denoted by f_{ijk}, whereas j is the bin number and k is the event process. Five possible processes are considered, the signal ($k = 1$) and the $t\bar{t}$, $Wb\bar{b}$, $Wc\bar{c}$+Wc, and $Wq\bar{q}$ background processes. Furthermore, the histograms are normalized such, that $\sum_{j=1}^{n_i} f_{ijk} = 1$ for all variables and processes. The likelihood function for a given event is computed by evaluating in which bin j_i the event falls in the distribution of variable i:

$$L_k(\{x_i\}) = \frac{\prod_{i=1}^{n_{var}} p_{ik}}{\sum_{k=1}^{5} \prod_{i=1}^{n_{var}} p_{ik}} \text{ , with } p_{ik} = \frac{f_{ij_ik}}{\sum_{l=1}^{5} f_{ij_il}} \tag{8.3}$$

The likelihood function for events with two jets and one b tag uses 7 input variables, whereas the likelihood function for events with 3 jets and one or two tags uses 10 input variables. Amongst others, the invariant mass of the reconstructed top-quark, the total scalar sum of transverse energy in the event, the charge of the lepton times the pseudorapidity of the light-quark jet, the χ^2 value of a kinematic solver used to find the most likely four-vector of the neutrino, the KIT flavor separator output, and the cosine of the angle between the lepton and the light-jet in the top-quark rest-frame are considered as input variables. Both likelihood functions are optimized for t-channel signal events. Figure 8.3 shows the LF discriminant shapes and the comparison between the LF discriminant applied to collision data and compared to the prediction of all three considered channels, respectively.

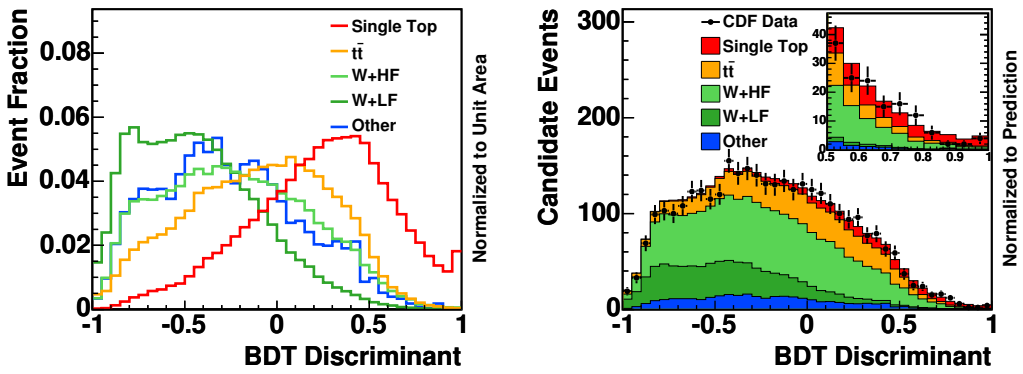


(a) (b)

Figure 8.3: The power of the LF discriminant is visible in the template shape comparison (a), while the predicted distribution is compared with the observed candidate events (b).

8.1.4 Boosted Decision Tree Analysis

The BDT discriminant uses a decision tree method that applies binary cuts iteratively to classify events, which is further improved using a boosting algorithm [165]. A decision tree is a binary tree structured classifier, where repeated accept/reject decisions are performed on a single variable at a time until some stop criterion is reached. Thus, the phase space is split up into signal like and background like regions, called leaves.
The boosting of a decision tree represents an extension to a single decision tree. Several decision trees, derived from the same training sample by reweighting, are combined to form a classifier which is given by a weighted majority vote of the individual decision trees. This boosting procedure stabilizes the response of the decision trees with respect to fluctuations in the training sample. Using boosted decision trees, many input variables are combined into a single powerful discriminant. The BDT discriminant uses over 20 input variables. Some of the most sensitive are the KIT flavor separator, the invariant mass of the reconstructed top-quark, the total scalar sum of transverse energy in the event, the charge of the lepton times the pseudorapidity of the light-quark jet, the dijet mass, and the transverse mass of the reconstructed W boson. Figure 8.4 shows the resulting simulated BDT discriminant template shapes and the BDT discriminant applied to collision data compared to the prediction.

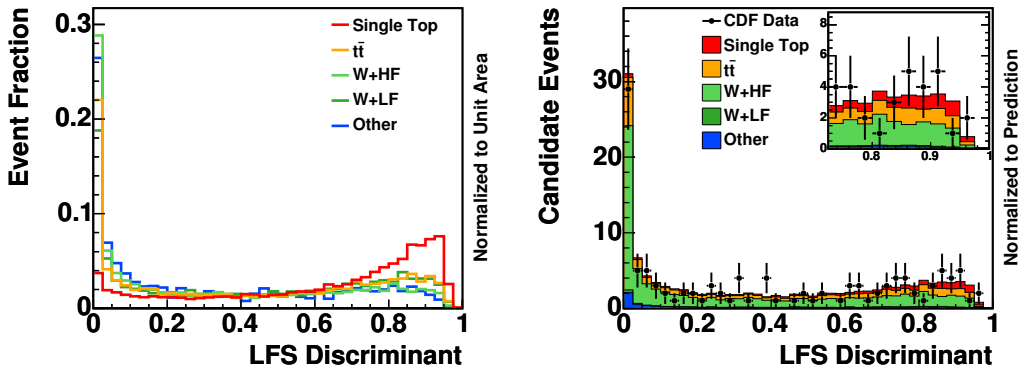


(a) (b)

Figure 8.4: The power of the BDT discriminant is visible in the template shape comparison (a), while in (b) the predicted distribution is compared with the observed candidate events.

8.1.5 Likelihood Function s-channel Analysis

The LFS analysis uses, similar to the LF analysis, a projective likelihood function to combine the separation power of several input variables [59]. Contrary to the LF analysis, the LFS discriminant is optimized to be sensitive to the s-channel signal process. Even though the LFS analysis itself is developed for further event samples, e.g. events with three jets or events from the EMC lepton category, only the subset complementary to the LF analysis is used in the lepton+jets combination. Thus, only TLC events with two b-tagged jets are used, resulting in 609 observed candidate events.
A kinematic fitter is applied to find the most likely z component of the neutrino momentum and the b jet that most likely came from the top-quark decay. In addition to the outputs of the kinematic fitter, other important inputs to the LFS discriminant are the invariant mass of both jets, the transverse momentum of the system composed of both jets, the leading jet transverse momentum, the reconstructed top-quark mass, the total scalar sum of transverse energy in the event, and the missing transverse energy. Figure 8.5 shows the resulting simulated LFS-discriminant template shapes and the LFS discriminant applied to collision data compared to the prediction.



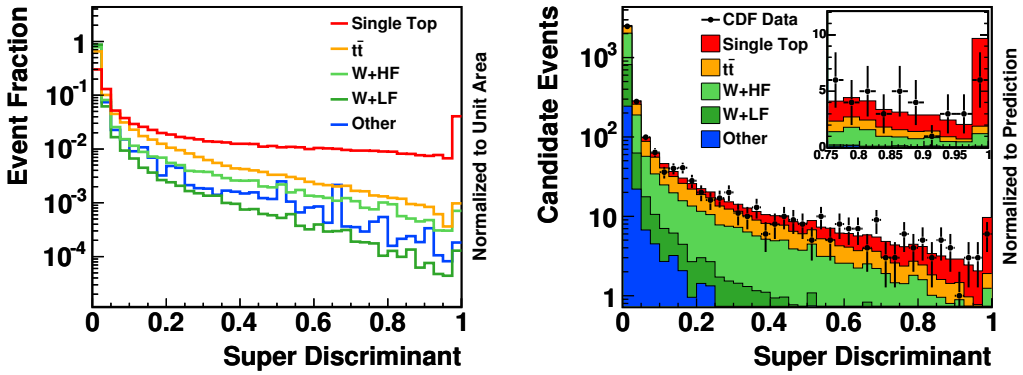
(a) (b)

Figure 8.5: The power of the LFS discriminant is visible in the template shape comparison (a), while the predicted discriminant distribution is compared with the observed candidate events (b).

8.2 Combination

8.2.1 Lepton+Jets Super Discriminant Analysis

The discriminant outputs of the individual lepton+jets analyses are combined into a single super discriminant (SD) using neural networks. The neural network weights and topology are optimized using a technique known as neuro-evolution of augmenting topologies (NEAT) [166, 167].
For each of the eight considered channels, NEAT begins from a population of neural networks generated from a seed network by randomly varying the network weights. The evolution then proceeds in generations, whereas in each generation, the following three steps are completed: first, the fitness of each neural network classifier is evaluated by calculating its performance using a figure of merit highly correlated to the expected significance. Second, neural networks with poor fitness are removed from the population. And last, the remaining classifiers are allowed to replenish the population through mutation and breeding. Possible mutations include randomly changing one or more NN weights, randomly adding a link between nodes, and randomly adding new nodes. Breeding involves blending randomly selected features from two neural networks. The population of neural networks remaining at the end of this process for one generation becomes the initial population for the next generation. At the end of the evolution, the classifiers with the highest fitness values are collected. A detailed estimation of the expected significance including all systematic uncertainties is subsequently performed in order to select the final neural network. With the SD analysis the a priori sensitivity improves by at least 13% over the best individual analysis, resulting in a median expected significance of $> 5.9\sigma$. The super discriminant of the sum of all eight considered channels is illustrated in figure 8.10.

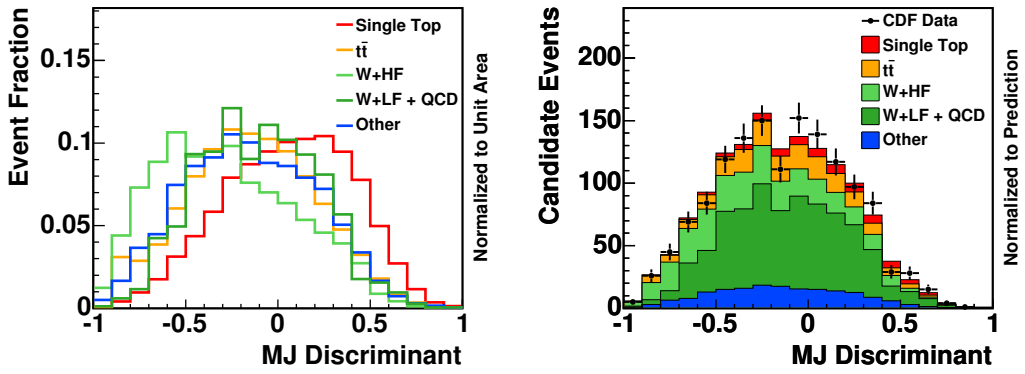


(a) (b)

Figure 8.6: The power of the super discriminant is visible in the template shape comparison (a), while in (b) the prediction is compared with the observed candidate events.

8.2.2 $\not\!\!E_T$+Jets Analysis

The MJ analysis is designed to select events with $\not\!\!E_T$+jets and to veto events selected by the lepton+jet analyses. It accepts events in which the W boson decays into taus and those in which the electron or muon fails the lepton identification criteria. Identification of b jets is done with the SecVtx b-tagging algorithm supplemented with a jet-probability algorithm based on the impact parameters of the jet [168], dividing the considered sample of candidate events into three distinct channels: events with one b tag and no impact parameter tag, events with one b tag and one impact parameter tag, and events with at least two b tags.
The main background process in the MJ analysis is QCD-multijet events in which mis-measured jet energies produce large $\not\!\!E_T$ aligned in the same direction as jets. Several powerful variables like the transverse momentum imbalance $\not\!p_T$, $\not\!\!E_T$, the angle between the latter, and the jet directions are used as inputs to a neural network (NNQCD) trained to suppress QCD-multijet events. The NNQCD output is required to pass a certain threshold, removing 77% of the QCD-background events, while keeping 91% of the signal acceptance.
The MJ analysis also applies neural networks to combine information of several input variables to discriminate single top-quark events from the remaining background events. The most important input variables are the invariant mass of the $\vec{\not\!\!E}_T$ and the second leading jet, the scalar sum of the jet energies, the $\not\!\!E_T$, and the azimuthal angle between $\vec{\not\!\!E}_T$ and the jets. Figure 8.7 shows the MJ discriminant shapes and the comparison between the MJ discriminant applied to collision data and compared to the prediction of all three considered channels, respectively.



(a) (b)

Figure 8.7: The power of the MJ discriminant is visible in the template shape comparison (a), while the predicted discriminant distribution is compared with the observed candidate events (b).

8.3 Results of the Combination

The discriminants of the SD and MJ analyses are used to extract the expected and observed signal significance and the measured signal cross-section. Therefore the templates of the simulated signal and background processes in eight lepton+jets channels and three $\not\!\!E_{\mathrm{T}}$+jets channels are fitted simultaneously, using the same statistical methods as described in chapter 7.1. As visible by eye, the super discriminant dominates by far over the MJ discriminant.
In the most signal-like bin, the SD analysis has a signal-over-background ratio of about five, while expecting about ten candidate events. This significantly reduces the sensitivity to systematic uncertainties affecting the background processes. The results of the lepton+jets and $\not\!\!E_{\mathrm{T}}$+jets candidate event yield estimates are summarized in table 8.1.

Process	lepton+jets	$\not\!\!E_{\mathrm{T}}$+jets
s-channel Signal	77.3 ± 11.2	29.6 ± 3.7
t-channel Signal	113.8 ± 16.9	34.5 ± 6.1
$t\bar{t}$	686.1 ± 99.4	184.5 ± 30.2
W+HF	1551.0 ± 472.3	304.4 ± 115.5
W+LF + QCD	777.9 ± 103.7	679.4 ± 27.9
Z+jets	52.1 ± 8.0	128.6 ± 53.7
Diboson	118.4 ± 12.2	42.1 ± 6.7
Total Prediction	3376.5 ± 504.9	1404 ± 172
Observed in Data	3315	1411

Table 8.1: Summary of predicted numbers of signal and background events in the selected lepton+jets and $\not\!\!E_{\mathrm{T}}$+jets data sample, respectively

8.3.1 Application to Observed Candidate Events

Signal Significance of Combination

By performing 400 million pseudoexperiments, a median expected p-value of $\hat{p}_{\mathrm{exp}} < 0.002 \times 10^{-6}$ is estimated, which corresponds to an expected median significance of $> 5.9\sigma$. The observed signal has an observed p-value of $\hat{p}_{\mathrm{obs}} = 0.310 \times 10^{-6}$. This corresponds to an observed signal significance of 5.0σ, thus electroweak single top-quark production is conclusively observed.
These signal significances represent an improvement of approximately 13% for the expected significance and 16% for the observed signal significance over the best single analysis, respectively. Figure 8.8 compares the observed value of the test statistic with the corresponding distributions of both tested hypotheses.

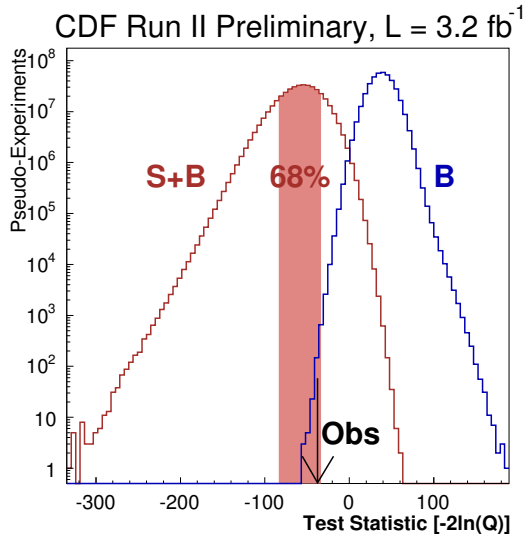


Figure 8.8: Test-statistic distributions for two ensemble tests, one with single top-quark events present at the expected SM rate (S+B), one without any single top-quark events (B) for the combination. The observed value of the test statistic is illustrated by the arrow.

Cross-Section Measurement

The most probable value of the combined s- and t-channel single top-quark production cross-sections is $\sigma_{\text{combined}} = 2.3^{+0.6}_{-0.5}$ pb, assuming a top-quark mass of 175 GeV/c^2. This is, within about one standard deviation, compatible with the theoretical NLO prediction of the combined signal cross-section of $\sigma^{\text{theo}}_{\text{combined}} = 2.9 \pm 0.4$ pb [78, 79]. The dependence on the top-quark mass is $+0.02$ pb/(GeV/c^2). Figure 8.9 shows the extracted posterior probability density.

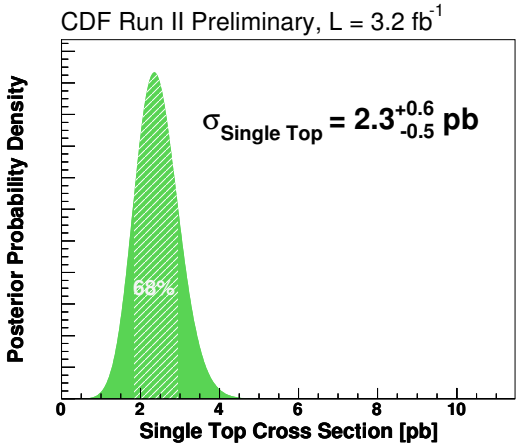
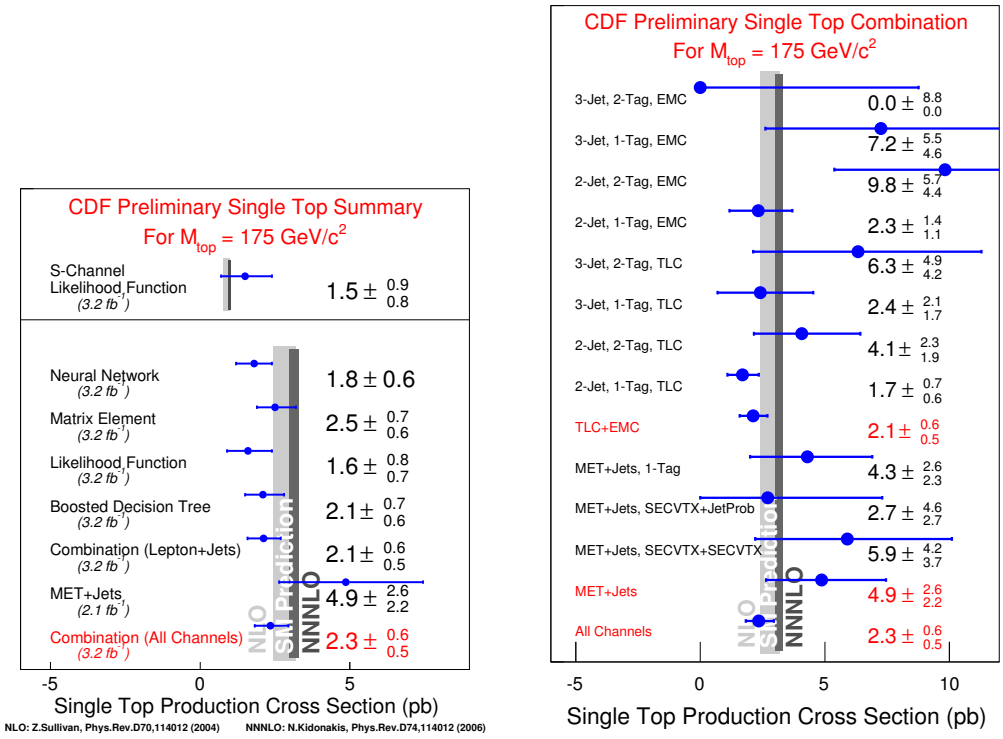


Figure 8.9: The extracted posterior probability density

Figure 8.10(a) illustrates the result for each analysis and the combination. As can be seen, the lepton+jets analyses are in good agreement with each other and consistently low compared to the theoretical combined s- and t-channel predictions, except for the LFS analysis. The latter, optimized to measure the s-channel signal, extracts a high s-channel cross section, still compatible with the theoretical prediction taking into account the quoted uncertainties. The MJ analysis measured a high combined signal cross-section, but, within uncertainties, still in agreement with the SM prediction.
The result in each of the eleven considered channels of the combination are shown in figure 8.10(b). Similar to the outcome of the NN analysis described in chapter 7.5.2, the three most precise channels, TLC and EMC events with two jets and one tag and TLC events with three jets and one tag, respectively, are below the SM expectations.



(a) (b)

Figure 8.10: The cross-section measurement results for each of the analyses and the combination (a), the results of the combination in each of the eleven channels separately, as well as the result for the lepton+jets channels, the $\not\!\!E_T$+jets (MET+jets) channels, and all channels (b)

.

Table 8.2 lists the measured cross sections and expected and observed significances for each of the analyses and the combination. Even though the MJ analysis has a rather low sensitivity, its high observed signal significance and its orthogonal candidate event sample has enough impact on the combination to interpret the excess of signal-like events over the expected background events as observation of single top-quark production with a signal significance of 5.0 standard deviations.

Analysis	Cross Section [pb]	Observed Significance [σ]	Expected Significance [σ]
LF	$1.6^{+0.8}_{-0.7}$	2.4	4.0
ME	$2.5^{+0.7}_{-0.6}$	4.3	4.9
NN	$1.8^{+0.6}_{-0.6}$	3.5	5.2
BDT	$2.1^{+0.7}_{-0.6}$	3.5	5.2
LFS	$1.5^{+0.9}_{-0.8}$	2.0	1.1
SD	$2.1^{+0.6}_{-0.5}$	4.8	> 5.9
MJ	$4.9^{+2.5}_{-2.2}$	2.1	1.4
Combined	$2.3^{+0.6}_{-0.5}$	5.0	> 5.9

Table 8.2: Results summary for the five correlated lepton+jets analyses, the SD analysis, and the MJ analysis, and the total combination. The LFS analysis measures only the s-channel production cross-section, while the other analyses measure the sum of the s- and t-channel cross-sections.

Extraction of the CKM Matrix Element $|V_{tb}|$

From the observed cross-section measurement at $m_t = 175$ GeV/c^2 including a $\pm 2.5\,\mathrm{GeV}/c^2$ uncertainty on the top-quark mass, and assuming $V_{\mathrm{tb}} \gg V_{\mathrm{ts}}, V_{\mathrm{td}}$ as well as a flat prior in $|V_{tb}|^2$ from 0 to 1, the absolute value of the CKM matrix element V_{tb} is obtained via

$$|V_{\mathrm{tb}}| = |V_{\mathrm{tb}}|^{\mathrm{theo}} \cdot \sqrt{\frac{\sigma^{\mathrm{obs}}_{\mathrm{combined}}}{\sigma^{\mathrm{theo}}_{\mathrm{combined}}}}\,, \tag{8.4}$$

where $|V_{\mathrm{tb}}|^{\mathrm{theo}}$ is the SM V_{tb} value from equation 1.2, $\sigma^{\mathrm{obs}}_{\mathrm{combined}}$ is the observed combined single top-quark production cross-section and $\sigma^{\mathrm{theo}}_{\mathrm{combined}}$ is the corresponding theoretical SM prediction [78, 79].
$|V_{tb}|$ is measured to be $|V_{tb}| = 0.91^{+0.11}_{-0.11}(\mathrm{stat} + \mathrm{syst}) \pm 0.07(\mathrm{theory})$, whereas the latter uncertainty originates from theoretical cross-section predictions. As expected, the quoted uncertainties are too large for a verification or falsification of the unitarity of the CKM-matrix. Furthermore, a lower limit is extracted with $|V_{tb}| > 0.71$ at the 95% C.L. Figure 8.11 shows the posterior probability density as a function of $|V_{tb}|^2$.

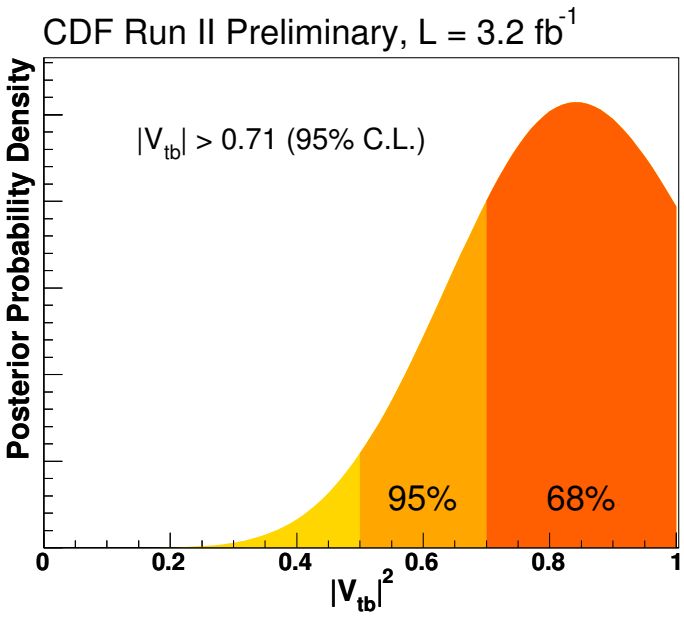


Figure 8.11: The posterior probability density as a function of $|V_{tb}|^2$

8.3.2 Observed Single Top-Quark Candidates

Figure 8.12 lists the five most t-channel single top-quark like candidate events from the most pure analysis channel of lepton+jets events with two jets and one b tag. The quantities shown in the figure include the run and event number, the lepton type, the value of the KIT flavor separator, the values of the individual discriminants, namely BDT (ranges from -1 to 1), LF (0 to 1), ME (0 to 1), NN (-1 to 1), SD (0 to 1), and the values of the three discriminating variables H_T, $M_{l\nu b}$, and $Q \times \eta$. By comparing the values of the variables with the corresponding shape distributions in figure 8.13, it is obvious that those signal like events populate almost exclusively the signal dominated regions. Moreover, it is remarkable that the average of the reconstructed top-quark mass over the five most signal like events amounts to 172.8 GeV/c^2, almost exactly the recent preliminary Tevatron result of $m_t = 173.1 \pm 1.3$ GeV/c^2 [18].

Run	Event	Lepton	KIT NN	BDT	LF	ME	NN	SD	H_T	$M_{l\nu b}$	$Q \times \eta$
148916	792764	CEM	0.94	0.76	0.94	0.97	0.94	0.99	219.0	189.7	-2.15
206282	3294678	CMUP	0.99	0.76	0.80	0.98	0.94	1.0	307.3	178.0	1.31
229936	4412760	CMX	1.00	0.95	1.00	0.99	0.97	1.0	221.0	171.1	2.03
242557	1564229	CEM	0.54	0.85	1.00	0.99	0.93	1.0	189.8	164.5	2.84
262776	4920497	CEM	0.86	0.95	1.00	1.00	0.92	1.0	191.8	160.8	2.67

Figure 8.12: The five most single top-quark like candidate events and some of their properties

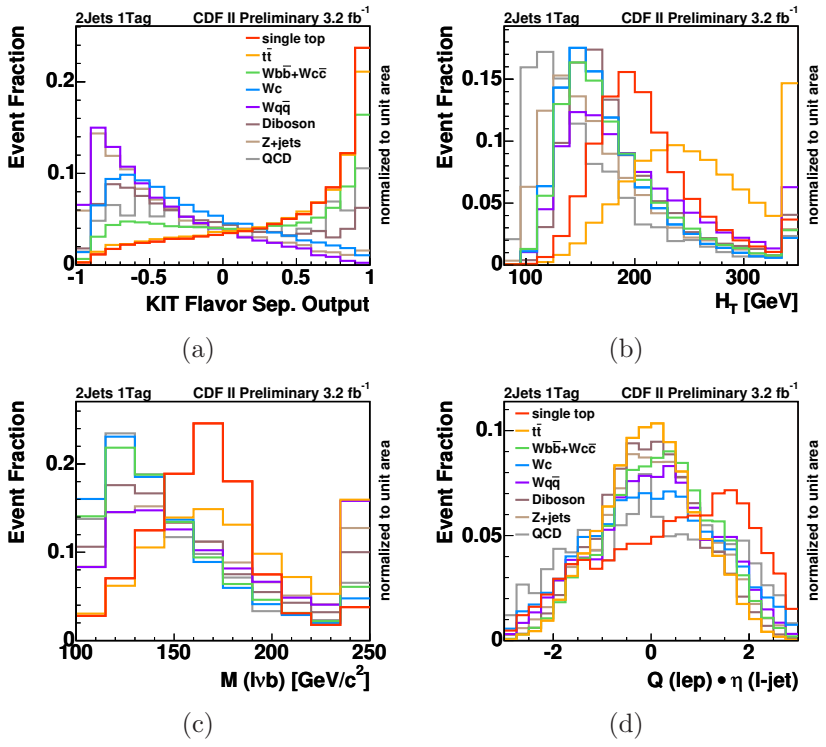


Figure 8.13: Four of the most discriminating variables for events with 2 jets and 1 b tag.

Figure 8.14 and 8.15 show the most t-channel signal like candidate event, number 4920497 from CDF run 262776, taken on March, 29th, 2008. The measured energy clusters of the CDF detector corresponding to reconstructed objects are shown in the η–ϕ plane view, figure 8.14, as well as in the 3D view, figure 8.15(a). The central electron ($p_{\mathrm{T}} = 57.2$ GeV/c) is drawn in red, the green cluster corresponds to the b-tagged jet ($E_{\mathrm{T}} = 66.7$ GeV), the forward jet ($E_{\mathrm{T}} = 31.2$ GeV), shown in blue, is not visible in the 3D view, but is at least indicated by several tracks pointing in forward direction. The undetectable neutrino causes a E_{T} imbalance amounting to $\not\!\!E_{\mathrm{T}} = 36.8$ GeV. In figure 8.15(b), the same observed event is drawn in the r–ϕ plane. The innermost part corresponds to the silicon tracking systems, surrounded by the COT drift chamber, which is heavily populated by hits of low-energetic and thus heavily curved tracks of charged particles. Outside of the COT are the electromagnetic and hadronic calorimeters, the energy depositions are illustrated in pink and blue, respectively. The central electron, indicated by the pink up-going track in both views of figure 8.15, causes the single electromagnetic energy deposit on the top, the b-tagged jet causes the electromagnetic as well as the hadronic energy deposits at the bottom. Another high-energetic pink track pointing in the same direction as the b-tagged jet is supposed to be a non-isolated electron from the semi-leptonic decay of the produced B hadron. In figure 8.15(b), the reconstructed direction of the neutrino is illustrated by the red arrow. The forward jet is not visible in this view.

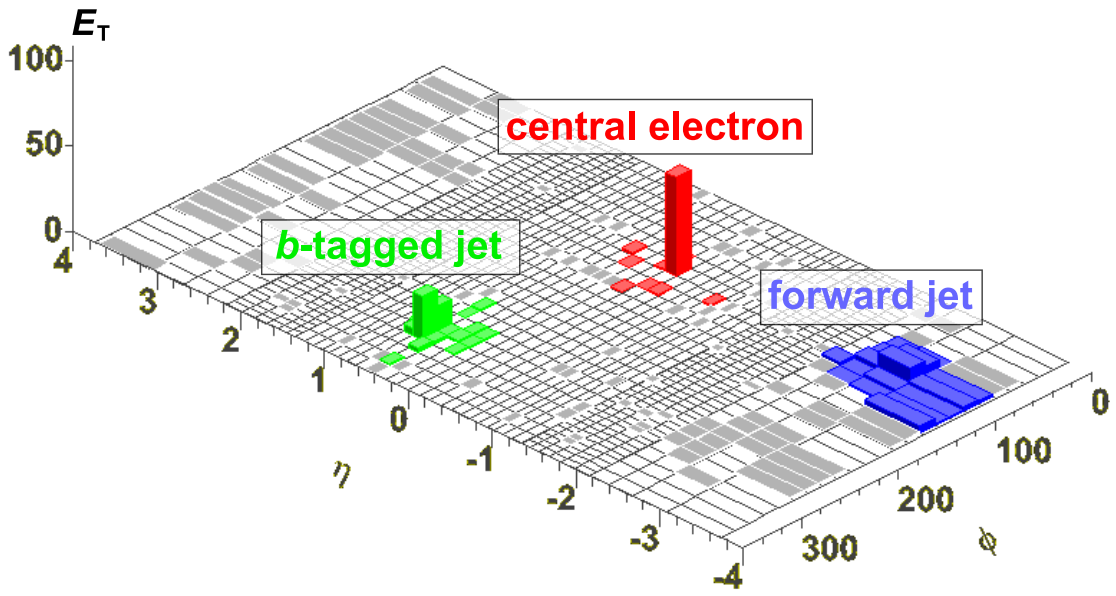
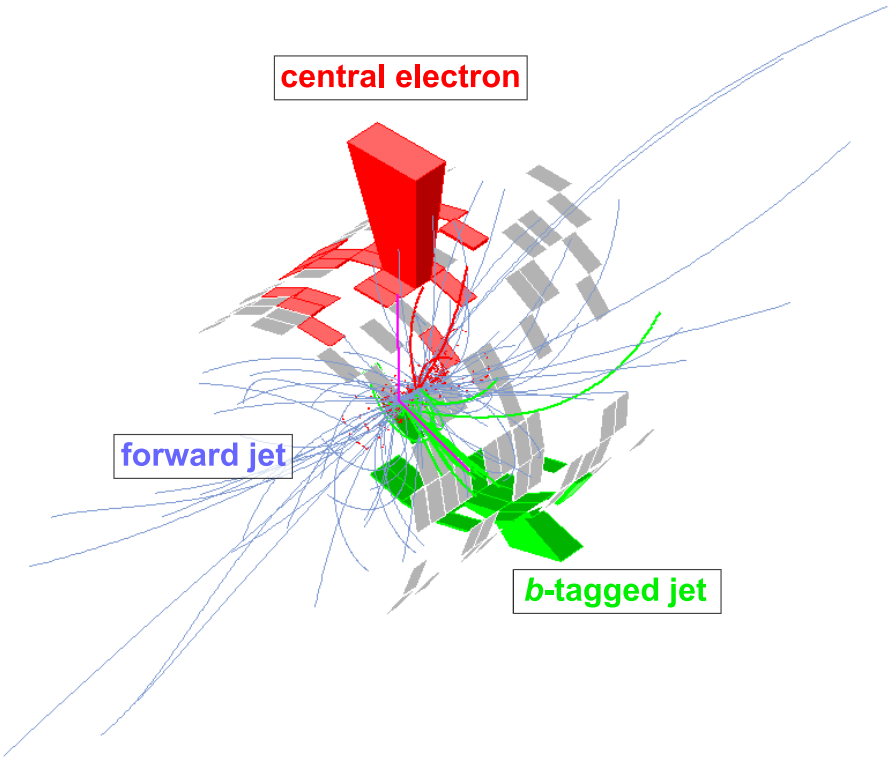
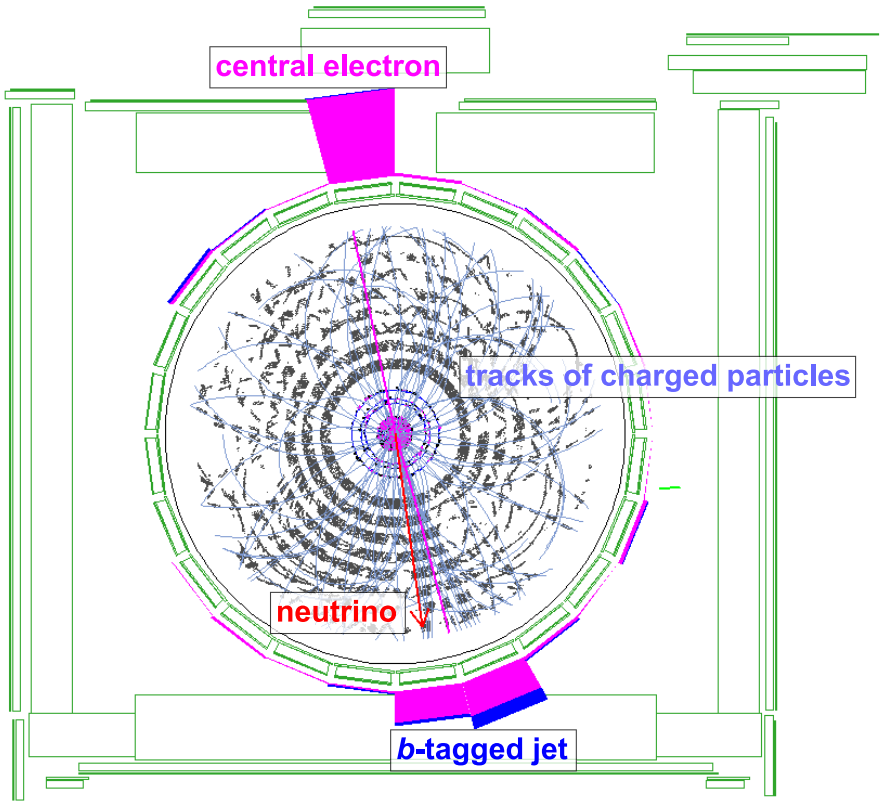


Figure 8.14: The most t-channel signal like candidate event is shown in the η–ϕ plane.



(a)



(b)

Figure 8.15: The most t-channel signal like candidate event is shown in the 3D view (a) and in the r–ϕ plane (b).

Summary

The top quark is the heaviest fundamental particle known so far and was discovered in 1995 by the CDF and the DØ Collaborations at the Tevatron [1, 2]. It was directly observed through the production of top-quark pairs via the strong interaction. Even though, the standard model of elementary particle physics (SM) also predicts electroweak single top-quark production. This thesis performs a search for single-top quark production which has not been observed up to the time of this analysis, whereas first evidence was already found by the DØ and shortly thereafter by the CDF Collaboration [33–35]. Since electroweak single top-quark production proceeds via a Wtb vertex, it provides the unique opportunity of the direct measurement of the CKM matrix element $|V_{tb}|$, which is sensitive to physics beyond the SM [20].

At the Tevatron, two single top-quark production modes are dominant. The t-channel production mode through an exchange of a virtual W boson results in a top-quark accompanied by a bottom and a light quark. The s-channel corresponds to a production of a timelike W boson via the fusion of two quarks resulting in a top-quark and a bottom quark. Since the top-quark decays almost exclusively into a bottom quark and a W boson, the experimental signature of single top-quarks is given by a W boson and a bottom quark from the top-quark decay, another bottom quark, and possibly additional light quarks. To reduce multijet background events, the W boson is demanded to decay leptonically into an electron or muon and its corresponding anti-neutrino manifesting itself as missing transverse energy.
After requiring two or three jets with at least one reconstructed secondary vertex b-tag, an isolated electron or muon, and large missing transverse energy, the expected number of signal events amounts to only about 5% of the full lepton+jets candidate event sample. Even worse, the amount of remaining background events is only known to a precision of about 20%. Thus, sophisticated analysis techniques are needed allowing for the extraction of such a small single top-quark signal from the imprecisely known amount of background processes.

In this thesis, one of the two most sensitive searches for combined electroweak s- and t-channel single top-quark production is described. With a sensitivity, in common with CDF's boosted decision tree (BDT) analysis, corresponding to an expected median significance of 5.2 standard deviations (σ) it exceeds DØ 's a priori most sensitive analysis by more than 20%. This CDF analysis is based on the application of NeuroBayes® neural networks [31, 32]. Those can be optimized to distinguish between signal and background processes by combining the information contained in

several variables each with rather low separation capability into a powerful discriminant. Such highly advanced discriminants need to be carefully checked for all sources of mismodeling in question, which is done in several sidebands of collision data. The neural network output distribution of the expected background processes and the SM prediction of the simulated single top-quark signal is compared to the observed distribution of candidate events using a Bayesian maximum likelihood technique. This includes the treatment of various systematic uncertainties affecting the rates of the predicted signal and background events and the shapes of the corresponding discriminant output distributions. Finally, the analysis method provides a sensitive and robust method for the extraction of the signal fraction and its significance.
The described analysis finds an observed signal significance of 3.5σ in a sample corresponding to 3.2 fb^{-1} of candidate events collected at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV. The combined s- and t-channel single top-quark production cross-section, assuming a top-quark mass of $m_t = 175$ GeV/c^2, is extracted to be $\sigma_{\text{combined}} = 1.8^{+0.6}_{-0.6}$pb, which is compatible with a downward fluctuation of about 1.5σ compared to the theoretical NLO prediction of the combined signal cross-section of $\sigma^{\text{theo}}_{\text{combined}} = 2.9 \pm 0.4$ pb [78, 79].
Moreover, a separate measurement of the s- and t-channel single top-quark production cross-section is performed, respectively. The additional application of a further optimized two-dimensional neural network classifier in the dominant channel with two jets and one b tag allows for the simultaneous extraction of both production cross-sections independently. The fit to the observed candidate events yields a cross sections of $\sigma_t = 0.7^{+0.5}_{-0.5}$ pb for t-channel and $\sigma_s = 2.0^{+0.7}_{-0.6}$ pb for s-channel single top-quark production, respectively, assuming a top-quark mass of $m_t = 175$ GeV/c^2. Compared to the theoretical NLO predictions of $\sigma_t^{\text{theo}} = 1.98^{+0.28}_{-0.22}$ pb and $\sigma_s^{\text{theo}} = 0.88^{+0.12}_{-0.11}$ pb [78, 79], the separate result seems to be compatible with a downward fluctuation for the t-channel cross-section and with an upward fluctuation for the s-channel cross-section, respectively, both at the order of about 2σ.

Furthermore, a combination of this analysis with four more CDF analyses in the lepton+jets candidate event sample, namely the BDT analysis, the matrix element analysis (ME), the likelihood function analysis (LF), and another likelihood function analysis optimized for s-channel signal (LFS), is performed using a super discriminant technique (SD). Thereby, the SD neural networks are trained to separate signal from background processes taking the discriminant outputs of the five analyses as inputs for each event.
Another neural-network based analysis using collision events with missing transverse energy $\not\!\!E_{\text{T}}$ and jets (MJ) adds about 30% to the signal acceptance by using an orthogonal sample of collision candidate events with no reconstructed charged lepton. Finally, a simultaneous fit over both exclusive SD and MJ analyses provides the final results of the search for combined s- and t-channel single top-quark production.
A median expected p-value of $\hat{p}_{\text{exp}} < 0.002 \times 10^{-6}$ is estimated, which corresponds to an expected median significance of $> 5.9\sigma$. The observed signal has an observed p-value of $\hat{p}_{\text{obs}} = 0.310 \times 10^{-6}$. This corresponds to an observed signal significance of 5.0σ, thus electroweak single top-quark production is conclusively observed [41].

The most probable value of the combined s- and t-channel single top-quark production cross-sections is $\sigma_{\text{combined}} = 2.3^{+0.6}_{-0.5}$ pb, assuming a top-quark mass of $m_t = 175$ GeV/c^2. This result is compatible with the standard model prediction. From the observed cross-section measurement at $m_t = 175$ GeV/c^2 and assuming $V_{\text{tb}} \gg V_{\text{ts}}, V_{\text{td}}$, the absolute value of the CKM matrix element V_{tb} is obtained to be $|V_{tb}| = 0.91^{+0.11}_{-0.11}(\text{stat} + \text{syst}) \pm 0.07(\text{theory})$, in agreement with the standard model value of $|V_{tb}| \approx 1$.

With a collision data set above 4 fb^{-1}, systematic uncertainties will start to be the limiting factor for future CDF single top-quark analyses. To measure both the combined as well as the separate production cross-sections and extract V_{tb} with significant higher precision, systematic uncertainties need to be considerably reduced. Even though, the available amount of signal candidate events will further allow the study of single top-quark properties, for instance the polarization of top-quarks only occurring through the electroweak production [84–86].

At the same time, the DØ Collaboration performed a neural network based combination of three lepton+jets analyses based on 2.3 fb^{-1} of collision data. DØ achieved an expected significance of 4.5σ and an observed signal significance of 5.0σ, thus also observed single top-quark production [42]. Under the assumption of a different top-quark mass of $m_t = 170$ GeV/c^2, DØ extracts a combined single top-quark production cross-section of $\sigma_{\text{combined}} = 3.9 \pm 0.9$ pb and the absolute value of the CKM matrix element $|V_{tb}| = 1.07 \pm 0.12(\text{theo})$.
A future combination of the results of the CDF and DØ Collaboration will provide a joint Tevatron single top-quark production cross-section measurement and the corresponding extraction of the CKM matrix element $|V_{tb}|$.

Recently, the CMS collaboration at the LHC approved an analysis evaluating the prospects for the measurement of the t-channel single top-quark production cross-section. The analysis is performed in the muon+jets sample and is based on simulated events corresponding to 0.2 fb^{-1} assuming $\sqrt{s} = 10$ TeV [40]. It is expected, that a reobservation of single-top quark production is feasible with about 0.7 fb^{-1} of CMS data.

14 years after its first observation, the top-quark has been rediscovered via the electroweak instead of the strong interaction. Again, both Tevatron collaborations CDF and DØ discovered the top-quark signal simultaneously. Though, the future of precision top-quark physics will take place at the LHC [43].

Bibliography

[1] F. Abe et al., Phys. Rev. Lett. 74, 2626–2631 (1995).

[2] S. Abachi et al., Phys. Rev. Lett. 74, 2632–2637 (1995).

[3] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. 49, 652–657 (1973).

[4] M. L. Perl et al., Phys. Rev. Lett. 35, 1489–1492 (1975).

[5] K. Kodama et al., Phys. Lett. B504, 218–224 (2001).

[6] S. W. Herb et al., Phys. Rev. Lett. 39, 252–255 (1977).

[7] A. Einstein, AdP 55, 241 (1918).

[8] S. L. Glashow, Nucl. Phys. 22, 579–588 (1961).

[9] A. Salam and J. C. Ward, Phys. Lett. 13, 168–171 (1964).

[10] S. Weinberg, Phys. Rev. Lett. 19, 1264–1266 (1967).

[11] G. Arnison et al., Phys. Lett. B122, 103–116 (1983).

[12] M. Banner et al., Phys. Lett. B122, 476–485 (1983).

[13] G. Arnison et al., Phys. Lett. B126, 398–410 (1983).

[14] P. Bagnaia et al., Phys. Lett. B129, 130–140 (1983).

[15] Th. Muller, Fortschr. Phys. 37, 339 (1989).

[16] G. Arnison et al., Phys. Lett. B147, 493 (1984).

[17] W. Hollik and Th. Muller, Phys. Bl. 53, 127 (1997).

[18] Tevatron Electroweak Working Group, FERMILAB-TM-2427-e (2009).

[19] S. S. D. Willenbrock and D. A. Dicus, Phys. Rev. D34, 155 (1986).

[20] T. M. P. Tait and C. P. Yuan, Phys. Rev. D63, 014018 (2001).

[21] H. Stadie, IEKP-KA/99-24 (1999).

[22] D. Hirschbuehl, IEKP-KA/2001-27 (2001).

[23] T. Walter, FERMILAB-THESIS-2005-28 (2005).

[24] D. Acosta et al., Phys. Rev. D71, 012005 (2005).

[25] Y. Kemp, FERMILAB-THESIS-2006-39 (2006).

[26] J. Lueck, FERMILAB-MASTERS-2006-01 (2006).

[27] M. Buehler, FERMILAB-MASTERS-2006-02 (2006).

[28] S. Richter, FERMILAB-THESIS-2007-35 (2007).

[29] I. Schall, IEKP-KA/2008-6 (2008).

[30] M. Renz, FERMILAB-MASTERS-2008-06 (2008).

[31] M. Feindt, `arXiv:physics/0402093` (2004).

[32] M. Feindt and U. Kerzel, Nucl. Instrum. Meth. A559, 190–194 (2006).

[33] V. M. Abazov et al., Phys. Rev. Lett. 98, 181802 (2007).

[34] V. M. Abazov et al., Phys. Rev. D78, 012005 (2008).

[35] T. Aaltonen et al., Phys. Rev. Lett. 101, 252001 (2008).

[36] T. Peiffer, FERMILAB-MASTERS-2008-01 (2008).

[37] A. Papaikonomou, IEKP-KA/2005-3 (2005).

[38] T. Aaltonen et al., Phys. Rev. Lett. 102, 151801 (2009).

[39] A. Papaikonomou, FERMILAB-THESIS-2009-21 (2009).

[40] J. Bauer et al., CMS Physics Analysis Summary TOP-09-005 (2009).

[41] T. Aaltonen et al., Phys. Rev. Lett. 103, 092002 (2009).

[42] V. M. Abazov et al., Phys. Rev. Lett. 103, 092001 (2009).

[43] C. E. Gerber et al., `arXiv:0705.3251` (2007).

[44] J. R. Incandela, A. Quadt, W. Wagner, and D. Wicke, `arXiv:0904.2499` (2009).

[45] D. Acosta et al., Phys. Rev. D65, 091102 (2002).

[46] D. Acosta et al., Phys. Rev. D69, 052003 (2004).

[47] B. Abbott et al., Phys. Rev. D63, 031101 (2001).

[48] V. M. Abazov et al., Phys. Lett. B517, 282–294 (2001).

[49] V. M. Abazov et al., Phys. Lett. B622, 265–276 (2005).

[50] V. M. Abazov et al., Phys. Rev. D75, 092007 (2007).

[51] V. M. Abazov et al., Phys. Rev. Lett. 99, 191802 (2007).

[52] V. M. Abazov et al., Phys. Rev. Lett. 101, 221801 (2008).

[53] T. Kikuchi, FERMILAB-THESIS-2000-26 (2000).

[54] C. I. Ciobanu, FERMILAB-THESIS-2002-42 (2002).

[55] S. K. Wolinski, FERMILAB-THESIS-2002-48 (2002).

[56] B. Stelzer, FERMILAB-THESIS-2005-79 (2005).

[57] P. J. Dong, FERMILAB-THESIS-2008-12 (2008).

[58] S. R. Budd, FERMILAB-THESIS-2008-41 (2008).

[59] K. Nakamura, FERMILAB-THESIS-2009-13 (2009).

[60] J. L. Tonry et al., Astrophys. J. 594, 1–24 (2003).

[61] J. Dunkley et al., Astrophys. J. Suppl. 180, 306–329 (2009).

[62] Y. Fukuda et al., Phys. Rev. Lett. 81, 1562–1567 (1998).

[63] S. Fukuda et al., Phys. Rev. Lett. 86, 5651–5655 (2001).

[64] Q. R. Ahmad et al., Phys. Rev. Lett. 87, 071301 (2001).

[65] Q. R. Ahmad et al., Phys. Rev. Lett. 89, 011301 (2002).

[66] C. Amsler et al., Phys. Lett. B667, 1 (2008).

[67] E. Noether, Nachr. v. d. Ges. d. Wiss. zu Goettingen 235–257 (1918).

[68] D. J. Gross and F. Wilczek, Phys. Rev. Lett. 30, 1343–1346 (1973).

[69] H. D. Politzer, Phys. Rev. Lett. 30, 1346–1349 (1973).

[70] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, Phys. Lett. B47, 365–368 (1973).

[71] H. L. Lai et al., Eur. Phys. J. C 12, 375–392 (2000).

[72] N. Cabibbo, Phys. Rev. Lett. 10, 531–532 (1963).

[73] F. Englert and R. Brout, Phys. Rev. Lett. 13, 321–322 (1964).

[74] P. W. Higgs, Phys. Rev. Lett. 13, 508–509 (1964).

[75] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Phys. Rev. Lett. 13, 585–587 (1964).

[76] P. W. Higgs, Phys. Rev. 145, 1156–1163 (1966).

[77] S. Moch and P. Uwer, `arXiv:0807.2794` (2008).

[78] B. W. Harris et al., Phys. Rev. D66, 054024 (2002).

[79] Z. Sullivan, Phys. Rev. D70, 114012 (2004).

[80] T. Stelzer, Z. Sullivan, and S. Willenbrock, Phys. Rev. D56, 5919–5927 (1997).

[81] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, Phys. Rev. Lett. 102, 182003 (2009).

[82] J. M. Campbell and R. K. Ellis, Phys. Rev. D62, 114012 (2000).

[83] N. Kidonakis, Phys. Rev. D74, 114012 (2006).

[84] A. P. Heinson, A. S. Belyaev, and E. E. Boos, Phys. Rev. D56, 3114–3128 (1997).

[85] D. O. Carlson and C. P. Yuan, Phys. Lett. B306, 386–390 (1993).

[86] G. Mahlon and S. J. Parke, Phys. Rev. D55, 7249–7254 (1997).

[87] J. Alwall et al., Eur. Phys. J. C49, 791–801 (2007).

[88] M. Bobrowski, A. Lenz, J. Riedl, and J. Rohrwild, `arXiv:0902.4883` (2009).

[89] T. Aaltonen et al., Phys. Rev. D76, 072006 (2007).

[90] M. Hickman et al., Public CDF Conference Note 9759 (2009).

[91] T. Aaltonen et al., Phys. Rev. Lett. 100, 161803 (2008).

[92] J. Conway et al., Public CDF Conference Note 9446 (2008).

[93] D. Mohl, G. Petrucci, L. Thorndahl, and S. Van Der Meer, Phys. Rept. 58, 73–119 (1980).

[94] S. Van Der Meer, Rev. Mod. Phys. 57, 689–697 (1985).

[95] G. I. Budker, Sov. Atom. Energ. 22, 438–440 (1967).

[96] S. Nagaitsev et al., Phys. Rev. Lett. 96, 044801 (2006).

[97] P. H. Garbincius, `arXiv:hep-ex/0406013` (2004).

[98] R. Blair et al., FERMILAB-PUB-96-390-E (1996).

[99] H. Minemura et al., Nucl. Instrum. Meth. A238, 18–34 (1985).

[100] C. S. Hill, Nucl. Instrum. Meth. A530, 1–6 (2004).

[101] A. Sill, Nucl. Instrum. Meth. A447, 1–8 (2000).

[102] A. Affolder et al., Nucl. Instrum. Meth. A453, 84–88 (2000).

[103] A. A. Affolder et al., Nucl. Instrum. Meth. A526, 249–299 (2004).

[104] D. Acosta et al., Nucl. Instrum. Meth. A518, 605–608 (2004).

[105] L. Balka et al., Nucl. Instrum. Meth. A267, 272 (1988).

[106] S. Bertolucci et al., Nucl. Instrum. Meth. A267, 301 (1988).

[107] R. Oishi, Nucl. Instrum. Meth. A453, 227–229 (2000).

[108] M. G. Albrow et al., Nucl. Instrum. Meth. A480, 524–546 (2002).

[109] A. Bhatti et al., Nucl. Instrum. Meth. A566, 375–412 (2006).

[110] C. M. Ginsburg, Eur. Phys. J. C33, s1002–s1004 (2004).

[111] G. Ascoli et al., Nucl. Instrum. Meth. A268, 33 (1988).

[112] T. Dorigo, Nucl. Instrum. Meth. A461, 560–562 (2001).

[113] D. Acosta et al., Nucl. Instrum. Meth. A461, 540–544 (2001).

[114] D. Amidei et al., Nucl. Instr. Meth. A269, 51 (1988).

[115] A. Abulencia et al., Nucl. Instrum. Meth. A598, 328–330 (2009).

[116] B. Ashmanskas et al., Nucl. Instrum. Meth. A518, 532–536 (2004).

[117] J. Adelman et al., Nucl. Instrum. Meth. A572, 361–364 (2007).

[118] G. Gomez-Ceballos et al., Nucl. Instrum. Meth. A518, 522–524 (2004).

[119] M. L. Mangano et al., J. High Energy Phys. 07, 001 (2003).

[120] F. Maltoni and T. Stelzer, J. High Energy Phys. 02, 027 (2003).

[121] T. Stelzer and W. F. Long, Comput. Phys. Commun. 81, 357–371 (1994).

[122] J. Alwall et al., J. High Energy Phys. 09, 028 (2007).

[123] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, J. High Energy Phys. 11, 063 (2001).

[124] F. Krauss, J. High Energy Phys. 08, 015 (2002).

[125] S. Mrenna and P. Richardson, J. High Energy Phys. 05, 040 (2004).

[126] T. Sjostrand et al., Comput. Phys. Commun. 135, 238–259 (2001).

[127] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, Phys. Rept. 97, 31 (1983).

[128] T. Sjostrand, Nucl. Phys. B248, 469 (1984).

[129] G. Corcella et al., J. High Energy Phys. 01, 010 (2001).

[130] D. Amati and G. Veneziano, Phys. Lett. B83, 87 (1979).

[131] R. Brun and C. Carminati, CERN Program Library Writeup W5013 (1993).

[132] R. Veenhof, Nucl. Instrum. Meth. A419, 726–730 (1998).

[133] G. Grindhammer, M. Rudowicz, and S. Peters, Nucl. Instrum. Meth. A290, 469 (1990).

[134] G. Grindhammer and S. Peters, `arXiv:hep-ex/0001020` (1993).

[135] E. Gerchtein and M. Paulini, `arXiv:physics/0306031` (2003).

[136] F. D. Snider, Nucl. Instrum. Meth. A566, 133–141 (2006).

[137] S. Menzemer, IEKP-KA/03-04, Internal CDF Note 5968 (2003).

[138] C. Hays et al., Nucl. Instrum. Meth. A538, 249–254 (2005).

[139] D. Acosta et al., Phys. Rev. D 71, 052003 (2005).

[140] E. E. Boos, V. E. Bunichev, L. V. Dudko, V. I. Savrin, and A. V. Sherstnev, Phys. Atom. Nucl. 69, 1317–1329 (2006).

[141] Y. L. Dokshitzer, Sov. Phys. JETP 46, 641–653 (1977).

[142] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. 15, 438–450 (1972).

[143] G. Altarelli and G. Parisi, Nucl. Phys. B126, 298 (1977).

[144] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, Nucl. Phys. B529, 424–450 (1998).

[145] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, J. High Energy Phys. 04, 068 (2004).

[146] E. L. Berger and H. Contopanagos, Phys. Rev. D57, 253–264 (1998).

[147] N. Kidonakis and R. Vogt, Phys. Rev. D 68, 114014 (2003).

[148] J. M. Campbell and R. K. Ellis, Phys. Rev. D 60, 113006 (1999).

[149] T. Aaltonen et al., Phys. Rev. Lett. 100, 102001 (2008).

[150] T. Aaltonen et al., Phys. Rev. Lett. 100, 091803 (2008).

[151] F. James and M. Roos, Comput. Phys. Commun. 10, 343–367 (1975).

[152] T. Junk, Nucl. Instrum. Meth. A434, 435–443 (1999).

[153] T. Bayes, Phil. Trans. R. Soc. Lond. 53, 269 (1763), reprinted in Biometrika 45, 293 (1958).

[154] J. Neyman and E. Pearson, Phil. Trans. R. Soc. Lond. A231, 289 (1933).

[155] R. Erbacher, A. Ivanov, and T. Schwarz, Public CDF Note 9462 (2008).

[156] A. Abulencia et al., Phys. Rev. D 73, 032003 (2006).

[157] D. Hirschbuehl et al., Public CDF Note 9716 (2009).

[158] F. Canelli et al., Public CDF Note 9711 (2009).

[159] F. Canelli et al., Public CDF Note 9701 (2009).

[160] S. Budd et al., Public CDF Note 9699 (2009).

[161] K. Nakamura et al., Public CDF Note 9712 (2009).

[162] A. Apresyan et al., Public CDF Note 9650 (2009).

[163] S. Budd et al., Public CDF Note 9251 (2008).

[164] M. F. Canelli, FERMILAB-THESIS-2003-22 (2003).

[165] A. Hocker et al., PoS ACAT, 040 (2007).

[166] K. O. Stanley and R. Miikkulainen, Evol. Comput. 10, 99–127 (2002).

[167] S. Whiteson and D. Whiteson, `arXiv:0903.0885` (2006).

[168] A. Abulencia et al., Phys. Rev. D74, 072006 (2006).

Danksagung

Acknowledgments

Mein erster Dank gilt Herrn Professor Dr. Thomas Müller dafür, dass er mich in seine Top-Arbeitsgruppe aufgenommen und mir die Durchführung dieser Doktorarbeit ermöglicht hat. Seine unentwegte Unterstützung und Rückendeckung besonders in den schwierigen Phasen der Arbeit, als auch die mir möglich gemachten Aufenthalte am Fermilab waren wichtige Motivation für diese Arbeit.

Herrn Professor Dr. Wolfgang Wagner gilt mein besonderer Dank für die geduldige Heranführung an das Thema, seine exzellente Art und Weise der Betreuung, die hervorragende Zusammenarbeit mit ihm und in der von ihm geformten Arbeitsgruppe und das mir entgegengebrachte Vertrauen, die von ihm begonnene Analyse weiterführen zu dürfen.

Ich bedanke mich bei den Mitgliedern der Top-Arbeitsgruppe für die außergewöhnliche und inspirierende Atmosphäre. Besonders herausheben möchte ich die enge Zusammenarbeit mit Frau Dr. Svenja Richter in den ersten Jahren, mit Herrn Dr. Dominic Hirschbühl in letzten Jahren und mit Herrn Dr. Antonios Papaikonomou über den gesamten Zeitraum, die zum Gelingen dieser Arbeit beigetragen hat. Bei Frau Dr. Jeannine Wagner-Kuhr bedanke ich mich für die kooperative Unterstützung und das intensive Korrekturlesen. Meinen Zimmerkolleginnen und -kollegen Matthias Bühler, Thorsten Chwalek, Irja Schall, Jasmin Gruschke und Manuel Renz danke ich zudem für das sehr angenehme und konstruktive Arbeitsklima. Für den interkollaborativen Wissensaustausch bin ich Julia Bauer und Philipp Sturm äußerst dankbar.

I wish to thank the members of the CDF Collaboration and its spokepersons, Professor Dr. Jacobo Konigsberg and Dr. Robert Roser for operating this magnificent experiment and enabling this outcome. This thesis would have been impossible without the teamwork within the CDF top group and single-top subgroup, led by Professor Dr. Takasumi Maruyama, Professor Dr. Robin Erbacher, Professor Dr. Kirsten Tollefson, Professor Dr. Florencia Canelli, Dr. Veronica Sorin, and Dr. Catalin Ciobanu, Professor Dr. Wolfgang Wagner, Professor Dr. Bernd Stelzer, Dr. Thomas Junk, Dr. Craig Group, respectively. I would like to thank Professor Dr. Rainer Wallny for his enduring support and Professor Dr. Kevin Lannon, Dr. Charles Plager, Dr. Peter Dong, Dr. Sarah Budd, Dr. Koji Nakamura, Bruno Casal Laraña, and Ji-Eun Jung for the productive collaboration.

Herrn Professor Dr. Michael Feindt und den NeuroBayes®-Experten seiner Arbeitsgruppe danke ich für die vielen hilfreichen Diskussionen und den nützlichen Tipps

und Tricks. Herrn Dr. Michael Milnik und Iris Gebauer bin ich für die zahlreichen diversen Hilfestellungen zum Dank verpflichtet. Dem EKP–Admin–Team unter der Leitung von Herrn Dr. Thomas Kuhr und Herrn Dr. Michal Kreps gebührt ein großer Dank für die Bereitstellung der Computing-Ressourcen, ohne die diese Arbeit in der vorliegenden Form nicht möglich gewesen wäre.

Bei Frau Weißmann, Frau Bräunling, Frau Schorn, Frau Gerstner und Frau Junge möchte ich mich für die sehr gute Zusammenarbeit und die organisatorische Mithilfe bedanken. Des Weiteren danke ich allen Mitgliedern des EKP für die angenehme Arbeitsatmosphäre.

Abschließend möchte ich mich bei meiner Mutter für ihren unaufhörlichen Zuspruch und ihre großzügige Unterstützung bedanken. Ebenso danke ich meiner Schwester und meinem Schwager, als auch den Eltern meiner Freundin für jegliche Hilfe und den gegebenen Rückhalt. Meiner Freundin Uli danke ich für ihre unermüdliche Kraft.

Diese Arbeit wurde unterstützt durch ein Stipendium des Ministeriums für Wissenschaft, Forschung und Kunst des Landes Baden-Württemberg im Rahmen des Strukturierten Promotionskollegs am KIT-Centrum Elementarteilchen- und Astroteilchenphysik KCETA.