CDF/THESIS/ELECTROWEAK/PUBLIC/6658
IEKP-KA/2003-20

Resultate zur Erzeugung und zum Nachweis von W-Bosonen mit dem CDF-Experiment in Proton-Antiproton-Kollisionen bei 1.96 TeV Schwerpunktsenergie

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

genehmigte

Dissertation

von

Dipl. Phys. Hartmut Stadie
aus Kiel

Tag der mündlichen Prüfung: 25.07.2003

Referent: Prof. Dr. Th. Müller, Institut für Experimentelle Kernphysik

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

Zusammenfassung

Das CDF-Experiment befindet sich am Tevatron, einem Proton-Antiproton-Kollider in der Nähe von Chicago. Erste Ereignisse wurden schon 1985 bei einer Schwerpunktsenergie von 1.8 Teraelektronenvolt genommen. In den folgenden elf Jahren wurde ein Datensatz gesammelt, der einer integrierten Luminosität von 130 pb^{-1} entspricht. Mit diesen Daten konnte 1995 das Topquark entdeckt werden.

Nach einer längeren Unterbrechung begann im Sommer 2001 eine neue Datennahmenphase, Run II genannt. In der Pause wurden die Schwerpunktsenergie des Tevatron auf 1.96 TeV erhöht und die Luminosität deutlich gesteigert. Auch der *Collider Detector at Fermilab*, CDF, wurde in dieser Phase verbessert und ausgebaut. Neue Spurdetektoren und ein neues Vorwärtskalorimeter erweitern die Akzeptanz für Teilchen in Vorwärtsrichtung. Zudem wurden die Ausleseelektronik und das Triggersystem an die höheren Ereignisraten angepasst. Ziel ist es, in den ersten Jahren des Betriebs Daten, die einer integrierten Luminosität von 2 fb^{-1} entsprechen, zu nehmen.

Bis zur Inbetriebnahme des *Large Hadron Colliders* am CERN ist das Tevatron der einzige Beschleuniger, mit dem Topquarks erzeugt werden können. Das Studium der Eigenschaften dieses schwersten Quarks wird eine Hauptaufgabe des CDF-Experimentes sein. Der wichtigste Kanal zum Nachweis von Topquarks besteht aus einem leptonisch zerfallenen W-Boson und Teilchenjets. Genaue Messungen der Topquark- und W-Bosonmassen erlauben eine präzisere Vorhersage der Masse des Higgsbosons, das noch unentdeckt ist.

Die effiziente Erkennung von W-Boson-Ereignissen ist eine wichtige Voraussetzung für Analysen zur W-Boson- und Topquarkphysik. Das Ziel meiner Arbeit ist, Methoden zur Erkennung von W-Boson-Ereignissen zu erarbeiten und zu vergleichen.

Nach einer kurzen Einführung in die Theorie der schwachen Wechselwirkung und einem Überblick über die verschiedenen Prozesse der W-Boson-Erzeugung am Tevatron im ersten Kapitel, wird im folgenden Kapitel das CDF-Experiment beschrieben. Das dritte Kapitel gibt einen Einblick in die wichtigsten Schritte bei der Rekonstruktion der Ereignisse. Zudem wird der für diese Arbeit verwendete Datensatz vorgestellt. Er umfasst alle brauchbaren Daten mit Vertexdetektorinformation, die bis Januar 2003 genommen wurden. Der Datensatz entspricht einer integrierten Luminosität von 55.5 ± 3.3 pb^{-1}.

Bei der Suche nach W-Bosonen habe ich mich auf den Zerfall $W \to e\nu$ beschränkt. Mit dem CDF kann man diese Ereignisse an Hand der Signatur des Elektrons im Detektor und fehlender Energie orthogonal zum Strahl (transversal) nachweisen, da das Neutrino den Detektor, ohne wechselzuwirken, verlässt. Wichtige Voraussetzung ist also eine gute Methode zum Elektronennachweis.

Dies ist das Thema des vierten Kapitels. Aus den Daten wurden ein Datensatz mit einem sehr hohen Anteil an wirklichen Elektronen (Signal) und ein Datensatz mit fehlerhaft erkannten Elektronen (Untergrund) ausgewählt. Mit Hilfe dieser Datensätze konnten die Unterschiede zwischen Signal und Untergrund untersucht werden. Zudem ermöglichen sie, die Effizienz und Reinheit verschiedenen Elektronenidentifikationsmethoden zu ermitteln. Den CDF-Standardalgorithmus, der Schnitte auf bestimmte Elektronvariablen vornimmt, haben ich auf diese Weise analysiert.

Eine weitere Möglichkeit ist das Trainieren eines künstlichen Neuronalen Netzes zur Elektronenidentifikation auf diesen Datensätzen. Um die Durchführbarkeit dieses Ansatzes zu überprüfen, wurden zwei Identifikationsverfahren mit dem Neuronalen Netz mit den Standardalgorithmus verglichen. Ein lockerer Schnitt auf die Netzausgabe selektiert einen Datensatz mit besserer Effizienz und Reinheit als das Standardverfahren. Ein harter Schnitt reduziert den Untergrund gewaltig.

Als weiteren Test dieses Identifikationsverfahren habe ich den W-Boson-Wirkungsquerschnitt für jedes Verfahren einzeln bestimmt. Ein wichtiger Unterschied zur bisherigen Analyse war eine andere Abschätzung des Untergrundanteils und ein lockerer Schnitt auf die fehlende Transversalenergie, die dem Neutrinotransversalimpuls entspricht. Für die Abschätzung des Untergrundanteils wurde ein Datensatz selektiert, der Elektronkandidaten enthält, die sehr wahrscheinlich keine echten Elektronen sind. Die Verteilung der fehlenden Transversalenergie für die Ereignisse in diesem Datensatz wurden zusammen mit der Verteilung des mit Monte Carlo-Methoden generierten und simulierten W-Boson-Signals an die Daten angepasst und so der Untergrundanteil bestimmt. Auf diese Weise ist nicht nur der Anteil, sondern auch die Form des Untergrunds bekannt.

Bei einem erwarteten Wirkungsquerschnitt von 2.73 ± 0.10 nb ergaben die Analyse mit den Standard-CDF-Schnitten $2.74 \pm 0.02_{\text{stat}} \pm 0.12_{\text{syst}} \pm 0.16_{\text{lum}}$ nb und die Analyse mit einem vergleichbaren Schnitt auf die Ausgabe des Neuronalen Netzes $2.76 \pm 0.01_{\text{stat}} \pm 0.12_{\text{syst}} \pm 0.16_{\text{lum}}$ nb als Wirkungsquerschnitt für die Erzeugung von W-Bosonen, die in Elektron und Neutrino zerfallen. Beide Ergebnisse stimmen sehr gut miteinander und der theoretischen Vorhersage überein. Die Analyse mit dem Neuronalen Netz ergibt etwas geringere Fehler. Dies ist eine Folge der besseren Trennung von Signal und Untergrund mit Hilfe des Netzes.

Diese Analyse wurde mit einem Vierzigstel der in den nächsten Jahren erwarteten Datenmenge durchgeführt. Eine Verringerung der systematischen Unsicherheiten, die jetzt schon bei dieser Messung überwiegen, ist also von höchster Bedeutung. Die in dieser Ar-

beit vorgestelle Methode der Elektronenidentifikation erlaubt es, die Reinheit des Analysedatensatzes zu erhöhen und somit die Unsicherheiten durch Untergrund zu verringern. Zudem kann diese Identifikationsmethode in vielen Analysen, wie zum Beispiel in der Suche nach elektroschwacher Topquarkproduktion oder der Suche nach Higgsbosonen, angewandt werden, um Untergründe zu unterdrücken. Mehr Daten erlauben auch die Selektion reinerer Datensätze für das Trainieren und die Effizienzbestimmung dieser Netze. Ferner können Neuronale Netze auch zum Nachweis anderer Teilchen, die mit dem CDF schwieriger zu identifizieren sind, benutzt werden.

Preface

The CDF experiment is located at the Tevatron proton-antiproton collider near Chicago. First events, running the collider at a center-of-mass energy of 1.8 TeV, were recorded in 1985. In the following eleven years, a dataset corresponding to an integrated luminosity of 130 pb^{-1} has been collected. These data allowed the discovery of the top quark in 1995.

After a long shutdown, a new data taking period, the Run II, started in summer 2001. In this break, the Tevatron and its detectors have undergone major upgrades. The Tevatron now reaches a center-of-mass energy of 1.96 TeV in proton-antiproton collisions and has a much higher luminosity. The changes at the *Collider Detector at Fermilab*, CDF, include a completely new tracking system, a new plug calorimeter, and a reworked data acquisition system to handle higher event rates. The study of the properties of the top quark is still one of the major tasks for this experiment. The best channel to detect top quarks involves a leptonically decaying W boson and hadronic jets. Besides the measurement of the top quark mass, a precision measurement of the W boson mass is needed to constrain the mass of the, as yet unseen, Higgs boson.

Therefore, an efficient detection of W boson events is crucial for many interesting measurements. Furthermore, a good understanding of the backgrounds is necessary for precision measurements. In this study, I investigate new methods to detect W boson events and estimate the size of the background.

After a brief introduction to the theory of electroweak interactions in the first chapter, I describe the CDF experiment in the second chapter. The third chapter summarizes the main steps of the event reconstruction and characterizes the data sets used in this study. I have limited myself to the decay $W \to e\nu$, which required a preselected sample containing electron candidates. This study is based on data collected with the CDF up to January 2003. This corresponds to an integrated luminosity of 55.5 pb^{-1}. I had to reprocess the data to apply the latest calibration constants and use improved versions of the reconstruction code.

A good electron identification is needed to separate the W boson events from strong interaction background which has a much higher cross-section. In the fourth chapter, I study the efficiency and purity of the electron identification used at CDF using signal and background samples obtained from data. With these samples, the variables used in the

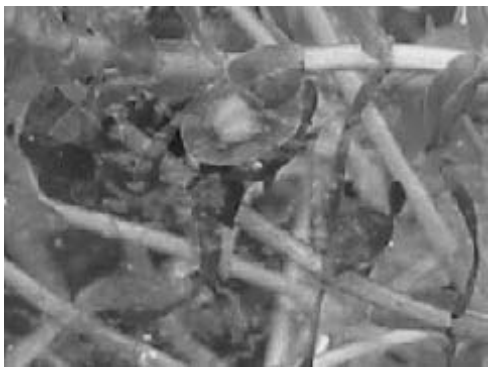
standard identification algorithm are analyzed. Furthermore, I have trained an Artificial Neural Net on these samples to identify electrons. Two different cuts on the net output are compared with the standard algorithm.

To check the consistency of the standard electron identification algorithm with the two alternative methods derived in chapter four, I have measured the W boson cross-section in three separate analyses. To do so, the acceptance and efficiency of my event selection have been determined and the systematic uncertainties have been estimated. Furthermore, I present an alternative method to estimate the contributions from events through the strong interaction. This is covered in chapter five. After the presentation of these cross-section measurements, I conclude this thesis.

Contents

List of Figures

List of Tables



Chapter 1

The W Boson in the Standard Model

In the Standard Model, the W boson mediates the weak interaction of two particles. A detailed description of the electroweak theory can be found in any text book of particle physics, e.g. [1, 2]. Instead of giving such a description, the first section will focus on the history of the weak interaction. I will try to sketch its evolution from the first observation of a weak interaction to the formulation of the Electroweak Standard Model. The later sections will summarize our current knowledge about the W boson and its production at a hadron collider.

1.1 The History of the Weak Interaction

The first observed weak interaction was the nuclear β-decay. Experiments showed the emission of an electron from a heavy nucleus. Unlike in nuclear γ-emissions, the energy of the electron was found to be continuous. For a two-body decay this observation contradicts energy and momentum conservation. The solution was the postulation of a low mass, neutral particle that is emitted with the electron by Wolfgang Pauli in 1930. As the electron carries spin $\frac{\hbar}{2}$, spin $\frac{\hbar}{2}$ was also assigned to this particle to conserve angular momentum. In 1933, Enrico Fermi applied the formalism of quantum electrodynamics on the β-decay [3, 4]. In analogy to γ-emission of an excited nucleus, Fermi replaced the photon by an electron-neutrino [1] pair. Figure 1.1 illustrates this for the β-decay process: $n \to pe\bar{\nu}_e$. Fermi's original proposal for the Lagrangian of the interaction had the form:

$$\mathcal{L} = G(\bar{\Psi}_p \gamma_4 \gamma_\mu \Psi_n)(\bar{\Psi}_e \gamma_4 \gamma_\mu \gamma_5 \Psi_\nu), \tag{1.1}$$

where Ψ_p,Ψ_n,Ψ_e, and Ψ_ν denote the wave functions of the four particles. Since the neutrino and the electron have a low mass compared to the kinetic energies in β-decay, the theory

[1] In Rome where Fermi worked at that time the postulated low mass, neutral particle was called *n*eutrino. With the success of Fermi's theory this name was adopted worldwide.

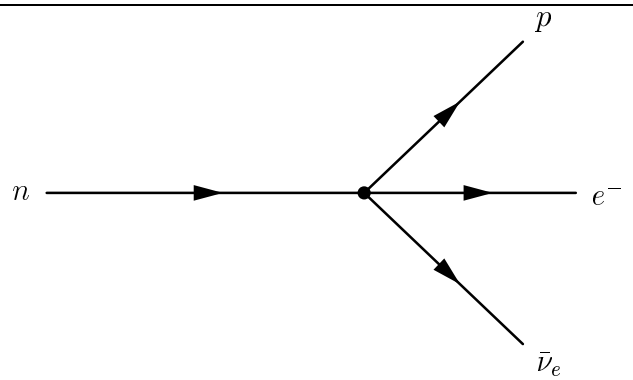


Figure 1.1: Four-fermion interaction diagram for neutron β-decay.

must be formulated relativistically. The wave functions are solutions of the free *Dirac* equation $(\not{p} - m)\Psi = 0$, that is four component spinors. The γ_5 matrix in the second term has no physical implication and can be omitted. Equation 1.1 does not allow the spin of a decaying nucleus to change. This is in contradiction to the β-decay of members of the Thorium family and in 1936 George Gamow and Edward Teller modified the Lagrangian to allow the spin change [5].

The most general, Lorentz-invariant form of the Lagrangian is:

$$\mathcal{L} = \sum_i C_i(\bar{\Psi}_p \hat{O}_i \Psi_n)(\bar{\Psi}_e \hat{O}_i \Psi_\nu), \tag{1.2}$$

where Ψ_p,Ψ_n,Ψ_e, and Ψ_ν denote once again the wave functions of the four particles. The quantities O_i are appropriate operators which characterize the decay and which are weighted by the constants C_i. $(\bar{\Psi}_p \hat{O}_i \Psi_n)(\bar{\Psi}_e \hat{O}_i \Psi_\nu)$ must be a Lorentz scalar. This reduces the possibilities for the O_i to 16 matrices: 1, γ^μ, $\sigma_{\mu\nu}$, $\gamma^\mu\gamma_5$, and γ_5. In other words the current $(\bar{\Psi}\hat{O}_i\Psi)$ can transform like a scalar (S), a vector (V), a tensor (T), an axial vector (A), or a pseudo scalar (P). Fermi had chosen the vector current in his ansatz.

Since in nuclear β-decay protons and neutrons move non-relativistically, the matrix elements can be simplified for the hadronic current. The two cases are the *Fermi* transitions, where the scalar and vector currents simplify to $\Phi_P^\dagger \Phi_n$, and the *Gamow-Teller* transitions, where the tensor and the axial vector currents simplify to $\Phi_P^\dagger \sigma \Phi_n$. Measurements of the lifetimes of several nuclei showed that the coupling strength of Fermi and Gamow-Teller transitions are about equal in magnitude. Therefore, the Lagrangian must be a combination of a scalar or vector coupling and a tensor or axial vector coupling of equal strength.

In 1948, Oscar Klein discovered that the decay of a muon and the β-decay can be described by the same four-fermion interaction. This lead to the hypothesis of the *Universal Fermi Interaction*, but the true structure of the weak interaction was not yet discovered.

There was another unresolved mystery at that point. Two mesons had been discovered with closely identical mass and lifetime, whereas their decay products had different parity. The parity operator is defined as: $\hat{P}\Psi(\vec{x}) = \Psi(-\vec{x})$. Mesons with angular momentum l have parity $P = -1^{l+1}$. A vector has negative parity ($\hat{P}V = -V$), while an axial vector has positive parity ($\hat{P}A = A$). As parity was assumed to be conserved in weak interactions, as it is in the electromagnetic and the strong interactions, these mesons had to be different particles, called θ and τ mesons at that time [2]. Hence, the problem concerning these particles was called the $\tau - \theta$ *puzzle*. The breakthrough came in 1956 when Tsung-Dao Lee and Chen-Ning Yang pointed out that there is no experimental proof for parity conservation in weak decays and suggested possible experiments to test parity conservation [6].

A short time after that, Wu, Ambler, Hayward, Hoppes, and Hudson proved that parity conservation is indeed violated in the β-decay of atomic nuclei [7]. In their now famous experiment, they analyzed the β-decay of Co^{60} nuclei that were polarized in an magnetic field. It was found that the emission of electrons is more favored in the direction opposite to that of the nuclear spin. In other words, the electrons have predominantly negative helicity. The helicity operator is defined as $\hat{\Lambda} = \vec{\sigma}\vec{p}/|\vec{p}|$, that is, the spin is projected on the direction of the momentum vector. As during a reflection the angular momentum vector of the spin keeps its direction, but the momentum vector of the electron reverses its direction, parity is violated in this experiment.

In another now famous experiment, Goldhaber, Grodzins, and Sunyar studied the decay of Eu^{152} into Sm^{152} [8]. In this decay, the nucleus captures an electron and emits a neutrino. The Samarium atom is then in an exited state and subsequently emits a photon to pass over to the ground state. Goldhaber, Grodzins, and Sunyar observed that these photons are always right circularly polarized. Due to momentum conversation, the neutrino must be emitted in the opposite direction than the photon. To conserve angular momentum the emitted neutrino then must always have negative helicity.

The discovery that predominantly electrons and neutrinos with negative helicity interact weakly, lead to the conclusion that the Lagrangian, e.g. for the muon decay, has the form [9, 10]:

$$\mathcal{L} = \frac{G_F}{\sqrt{2}}(\bar{\Psi}_{\nu_\mu}\gamma_\mu(1-\gamma_5)\Psi_\mu)(\bar{\Psi}_e\gamma_\mu(1-\gamma_5)\Psi_{\nu_e}) \tag{1.3}$$

Figure 1.2 illustrates this process. The operator $\frac{1}{2}(1-\gamma_5)$ projects out states with negative chirality. Using the helicity operator as suggested by experiment leads to a theory that is not Lorentz-invariant. Nevertheless, for massless particles the helicity projection operator equals the chirality projection operator. As a result, this form implies that due to their low mass only neutrinos with negative helicity interact weakly in charged current reactions.

[2] In todays nomenclature, the puzzle arose from the decays: $K^+ \to \pi^+\pi^0$ (P = +1) and $K^+ \to \pi^+\pi^+\pi^-$ (P = -1).

The structure of the interaction is a vector minus axial vector $(V - A)$ coupling. As a vector is parity odd and an axial vector is parity even, this coupling violates parity maximally.

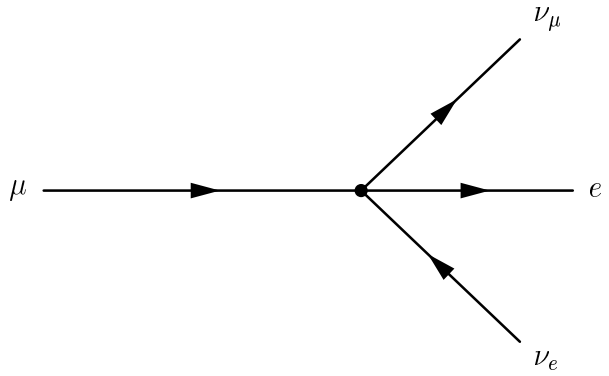


Figure 1.2: Four-fermion interaction diagram for muon decay.

As the interaction is point-like, the cross-section corresponding to diagram 1.2 diverges for large transferred momenta. The more momentum is transferred, the larger the phase space for the electron and neutrino. Thus, Fermi's theory of the weak interaction is only valid for low center-of-mass energy reactions. To avoid this divergence, the constant coupling G_F was replaced by an energy dependent term $\frac{g_W^2}{q^2+m_W^2}$, where q is the transferred momentum. For low center-of-mass energies this term can be approximated by $\frac{g_W^2}{m_W^2} \propto G_F$. So this new theory becomes the proven Fermi Theory for low energies. In this new ansatz, the interaction is not point-like, but mediated by the exchange of an "intermediate" boson with mass m_W, the W boson. The diagram of the muon decay mediated by a W boson is shown in Figure 1.3. To mediate the interaction of charged currents, these bosons have to carry charge, so there are W^+ and W^- particles.

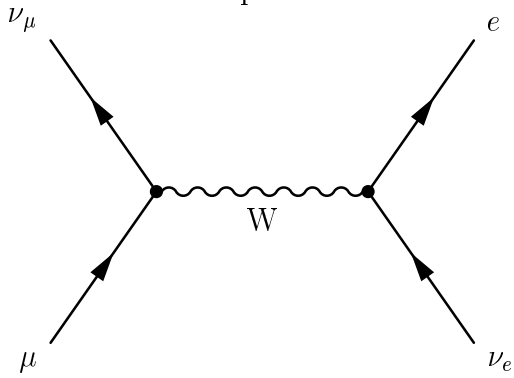


Figure 1.3: Diagram for muon decay mediated by a W boson.

Unfortunately, for a vector or axial vector coupling new divergencies arise, that is,

the theory is not renormalizable, if the intermediate boson is massive. A solution to this problem was introduced by Englert, Brout, and Higgs [11, 12]. The trick is to start with a massless W boson, but couple it to a new spin-0 field ϕ, the Higgs field, by means of an interaction term $g^2\phi^2 W_\mu$. If the scalar field ϕ assumes a constant value $\phi_0 = M_W/g$, the term for the interaction of the W boson with the Higgs field $g^2\phi_0^2 W_\mu$ plays the same role as the mass term $M_W^2 W_\mu$. This way of assigning mass to particles is called the "Higgs mechanism".

The theory of the weak interaction has now a similar form as the theory of the electromagnetic interaction. In 1961, Glashow wrote a first simple model for an electroweak gauge theory [13], a united theory for the electromagnetic and weak interactions. This approach was finalized by Salam and Weinberg in 1967 [14, 15, 16] and forms the core of the Standard Model of Particle Physics.

This electroweak theory is based on the invariance of the Lagrangian under local gauge transformations generated by the $SU(2)_L \times U(1)_Y$ group. In analogy to the spin that is based on the same $SU(2)$ group, the quantity that is conserved due to this $SU(2)$ symmetry is called *isospin*. The "L" in $SU(2)_L$ denotes that only the left-handed leptons form isospin doublets. With the discovery of the τ lepton in 1975 by Perl [17], there are three known doublets for the leptonic sector:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L , \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L , \text{ and } \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L .$$

This theory assumes that a right-handed neutrino does not exist, so the right-handed fermion fields of charged leptons appear in isospin singlets: e_R^-, μ_R^-, and τ_R^-. The $U(1)$ symmetry is essential to incorporate the electromagnetic interaction. The conserved quantum number Y is called weak hypercharge. It is defined as:

$$Q = I_3 + \frac{Y}{2} , \tag{1.4}$$

where I_3 is the third component of the weak isospin and Q is the electric charge. As the ν_e is the upper component in the doublet, it has $I_3 = \frac{1}{2}$. As it is neutral, its weak hypercharge has to be $Y = -1$.

The massless gauge fields due to the $SU(2)_L \times U(1)_Y$ symmetry are an isospin triplet W_μ^i and a singlet B_μ. The two charged fields in the triplet with $(I_3 = \pm 1)$ correspond to the $W^\pm$ bosons and only couple to the doublets containing the left-handed fermion fields. The neutral field mixes with the field B. The physical fields, the photon field A and the Z field, are obtained by rotating (W^3, B) by the angle θ_W, the weak mixing angle. Consequently, the vector bosons of the weak interaction are the $W^\pm$ and the Z^0 bosons. The electromagnetic coupling α, the Fermi coupling G_F, and the weak mixing angle θ_W

can be used to predict the boson masses:

$$m_W = \frac{A}{sin\theta_W} \text{ and } m_Z = \frac{A}{sin\theta_W cos\theta_W}, \text{where } A = \sqrt{\frac{\pi\alpha}{\sqrt{2}G_F}} \tag{1.5}$$

I will now focus on the evolution of the Glashow-Salam-Weinberg model for hadrons. In 1964, GELL-MANN and ZWEIG postulated that baryons and mesons are made up of quarks [18]. In their ansatz, they predicted the existence of three different *flavors* of quarks, called *up, down,* and *strange*. The quark model received special support from the discovery of further mesons [19, 20] that can be explained with a new quark flavor, the charm quark. As there are three families of leptons, one would expect two more quark flavors. Only three years later the fifth quark flavor, the bottom quark, was discovered [21]. Due to its extremely large mass, it took a long time to find the sixth quark flavor: the top quark was discovered in 1995 at the *Tevatron* by the CDF and D0 experiments [22, 23].

Hence, there are as well three weak isospin doublets for the quark sector:

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \text{ and } \begin{pmatrix} t \\ b' \end{pmatrix}_L$$

and the corresponding singlets: u_R, d'_R, c_R, s'_R, t_R, and b'_R. The up-type quarks carry charge $\frac{2}{3}$ and the down-type quarks charge $-\frac{1}{3}$. There is no reason that the mass eigenstates of the quarks are eigenstates of the electroweak interaction. By convention, the d', s', and b' states are a superposition of the mass eigenstates d, s, and b. NICOLA CABIBBO compared the decay rates for $K^+ \to \mu^+\nu$ and $\pi^+ \to \mu^+\nu$ to extract the angle of the rotation that translates the mass eigenstates of the first two quark families into weak interaction eigenstates, the *Cabibbo angle* [24]. In 1973 KOBAYASHI and MASKAWA extended this ansatz to three families of quarks [25]. The *Cabibbo-Kobayashi-Maskawa mixing matrix*, CKM matrix, defines the transformation from the mass eigenstates of the quarks to the eigenstates of the electroweak interaction. By convention, it is expressed by a 3×3 unitary matrix V operating on the charge $-1/3$ quarks:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.6}$$

The lowest order diagram for the β-decay using the quark picture is shown in Figure 1.4.

The $W^\pm$ and Z^0 bosons predicted by the Standard Electroweak Model were found with the $Sp\bar{p}S$ collider at CERN in 1983 [26]. In the following twenty years, the Standard Electroweak Model has well been tested with high precision data from experiments at

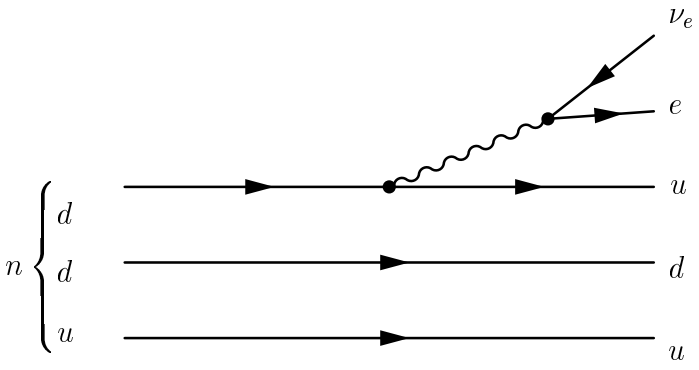


Figure 1.4: Interaction diagram for neutron β-decay.

the *Stanford Linear Collider*, the *Large Electron-Position Storage Ring* (LEP), and the *Tevatron*, a proton-antiproton collider. However, the Higgs boson, the particle that is responsible for breaking the electroweak symmetry in the Standard Model by making the W and Z bosons massive, has not yet been discovered. The search for the Higgs boson or an alternative mechanism for generating the particle masses is one of the major goals in high energy physics. Besides the direct search, the study of the W boson, e.g. measuring its mass, allows to constrain the mass of the Higgs boson and exclude some of the alternative models. This will be shown in the next section.

1.2 Properties of the W Boson

1.2.1 The Mass

Since the discovery of the W boson in 1983 and the first mass measurements by UA1 and UA2, the W mass measurements have significantly been improved by the CDF and $D0$ experiments at the Tevatron and by the four experiments at LEP. The world average of these measurements is $m_W = (80.423 \pm 0.039)$ GeV/c^2 [27]. Figure 1.5 summarizes the latest measurements from the different experiments, with these results the average mass is $m_W = (80.426 \pm 0.0034)$ GeV/c^2 [28].

In lowest order, the mass of the W boson and Z boson are given by equation 1.5. However, radiative corrections have to be applied to these results. Within the Standard Model, the good understanding of these corrections and the precision of the mass measurements allow good predictions of the Higgs boson mass. Figure 1.6 shows one diagram for contributions to the W boson mass that depend on the top quark mass and one for contributions that depend on the Higgs boson mass. As the Higgs boson has not yet been discovered, measurements of the top and the W boson mass put an important constraint on possible mass values for the Higgs boson in the Standard Model, see figure 1.7.

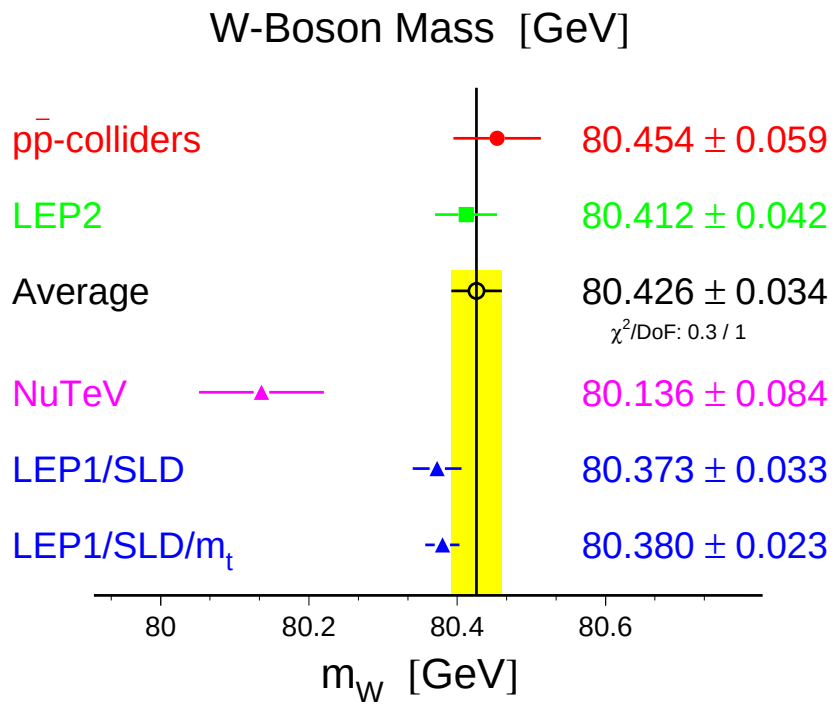


Figure 1.5: World average of W boson mass measurements [28].

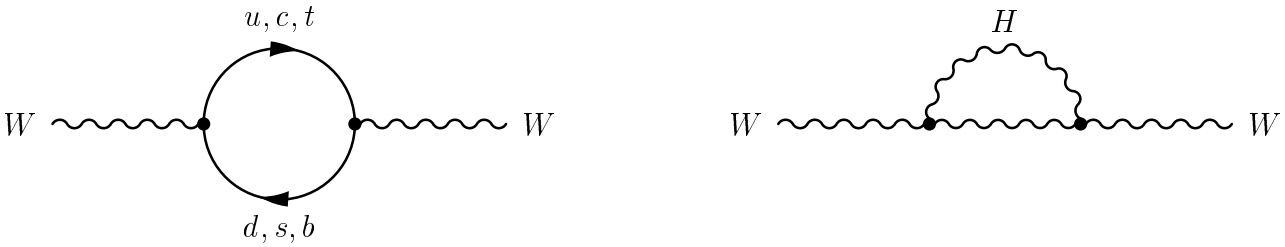


Figure 1.6: Higher order contributions to the W propagator.

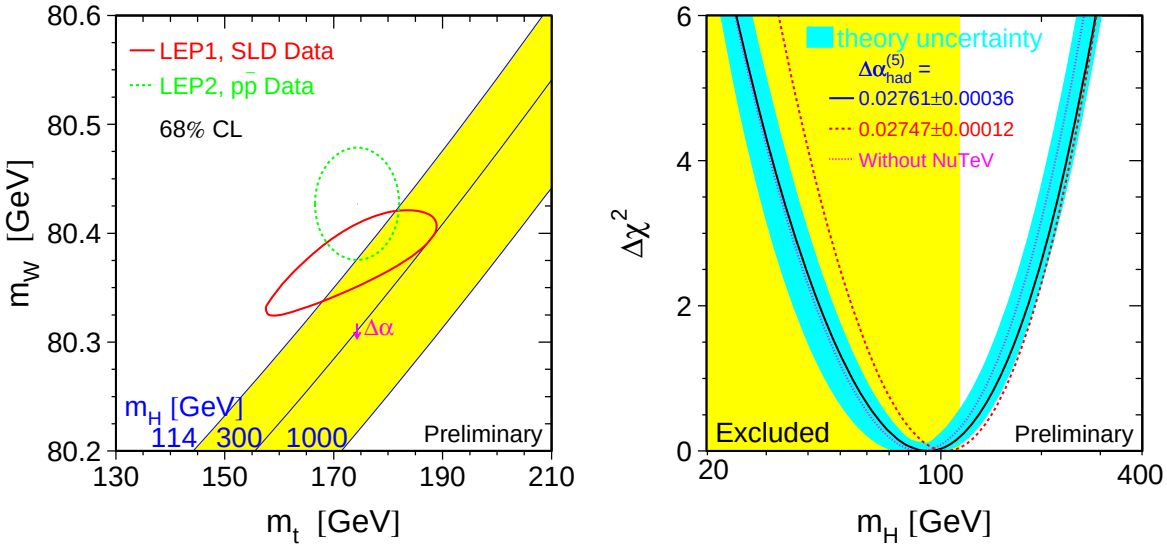


Figure 1.7: Constraints on the Higgs mass as a function of the Higgs mass and the top quark mass [28].

1.2.2 The Width

The W boson decays in roughly one out of three cases in lepton pairs and in two out of three cases in hadrons. As the top quark is heavier than the W boson, the W boson does not decay into any quark-antiquark pair containing a top quark. So the possible final states are:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L$$

The partial widths for the decays into any of the three lepton pairs are equal. Due to the fact that there are three different color charges, the decay widths into quark-antiquark pairs are enhanced by a factor of three. Figure 1.8 shows the diagrams for the decay into quark-antiquark pairs and into lepton pairs. The branching ratios for the different quark-antiquark combinations depend on the Cabibbo-Kobayashi-Maskawa matrix.

In this study, I limited myself to W bosons decaying into an electron and neutrino pair. The W width for this decay channel is approximately:

$$\Gamma(W^- \to e\bar{\nu}_e) \approx \frac{G_F}{6\pi\sqrt{2}} M_W^3 \approx 225 \text{ MeV}$$

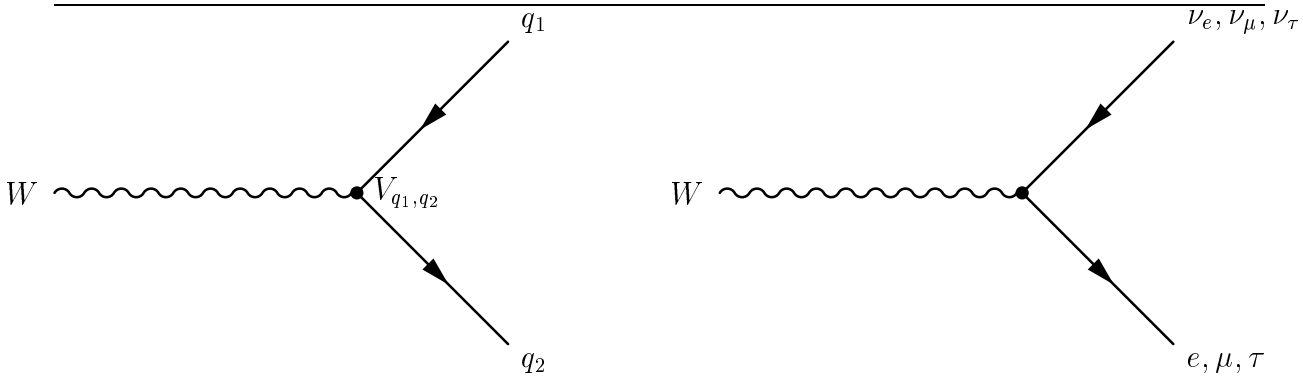


Figure 1.8: Diagrams for W boson decay. q_1 and q_2 denote quark mass eigenstates.

Its branching ratio is 10.72 ± 0.16 % [27]. Due to the $V-A$ coupling in the W decay, the direction of the electron depends on the polarization of the W boson. In the case of a W^+ boson with positive helicity, the positron is predominantly emitted in the direction of the W^+ boson. In contrast, in the case of negative helicity, the positron predominantly travels backwards with respect to the direction of the W^+ boson. For transversely polarized W^- bosons with helicity $h = \pm 1$, the decay distributions of the electron in the W^- rest frame are

$$\frac{d\Gamma_\pm}{d\cos\hat{\theta}} \sim (1 \mp \cos\hat{\theta})^2,$$

where $\hat{\theta}$ is the angle of the electron with respect to the longitudinal axis. For longitudinally polarized W^- bosons (helicity $h = 0$), the electron decay distribution is

$$\frac{d\Gamma_L}{d\cos\hat{\theta}} \sim 1 - \cos^2\hat{\theta}.$$

The world average of the measurements of the full width is $\Gamma_W = 2.118 \pm 0.042$ GeV [27]. This includes indirect measurements of the width. Figure 1.9 summarizes the latest width measurements.

1.3 The Production of W Bosons at the Tevatron

In this study, I analyzed W boson events produced at the Tevatron. The Tevatron is a proton-antiproton collider with a center-of-mass energy of 1.96 TeV. The W bosons are produced through interactions of quarks and gluons inside the proton and antiproton. Thus, a good knowledge of the structure of the proton is necessary to predict the production rates. Furthermore, the momenta of the initial state partons affect the kinematics of the produced W boson. Figure 1.10 shows the *CTEQ5L* parton distribution function (PDF) [29] for $u, d, \bar{d}$ quarks and gluons inside a proton, as a function of the fraction x

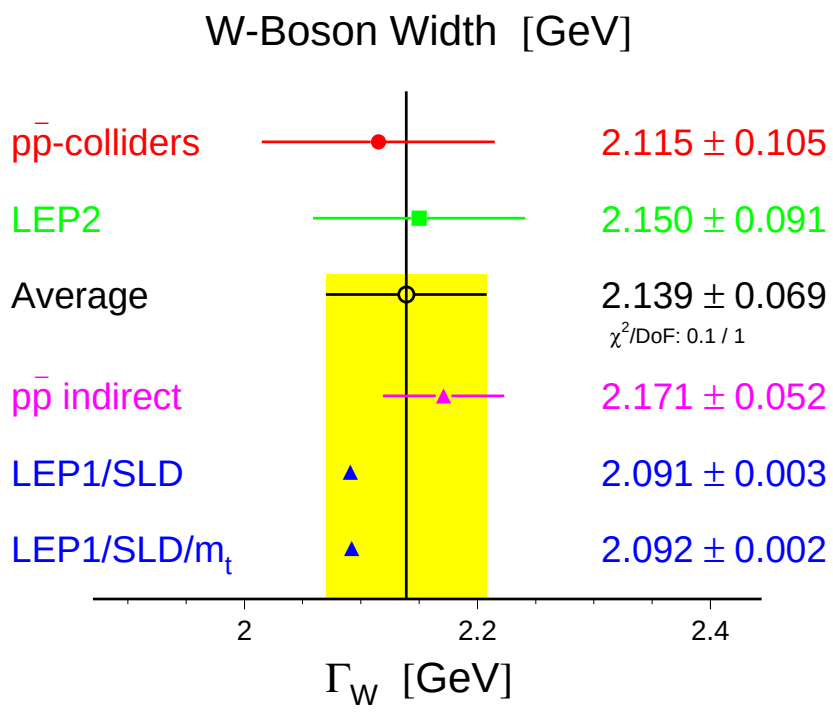


Figure 1.9: World average of W boson width measurements [28].

of the hadron momentum carried by the parton. One can see, that the u quark carries a slightly larger momentum fraction than the d quark, this is due to the fact that two of the three valence quarks in the proton are up quarks, and only one is a down quark. Therefore, it is more likely for an up quark in the proton to be a valence quark than for a down quark. This results in a forward-backward asymmetry for W bosons from direct production, i.e. the W^+ boson has more likely a boost in the u quark direction. Another source of W bosons is the decay of top quarks. Table 1.1 lists the cross-sections for the most important Standard Model processes of W boson production at the Tevatron.

1.3.1 Direct Production

The main source of W bosons at a hadron collider like the Tevatron is direct production through quark-antiquark annihilation. A tree level diagram for the production of W bosons is shown in figure 1.11. A quark and an antiquark annihilate to produce a W boson. Bosons produced in this Drell-Yan process have no transverse momentum[3] and are almost fully polarized along the antiproton direction due to the $V - A$ coupling. Hence,

[3] Throughout this thesis *longitudinal* means parallel to the proton beam and *transverse* means normal to the proton beam.

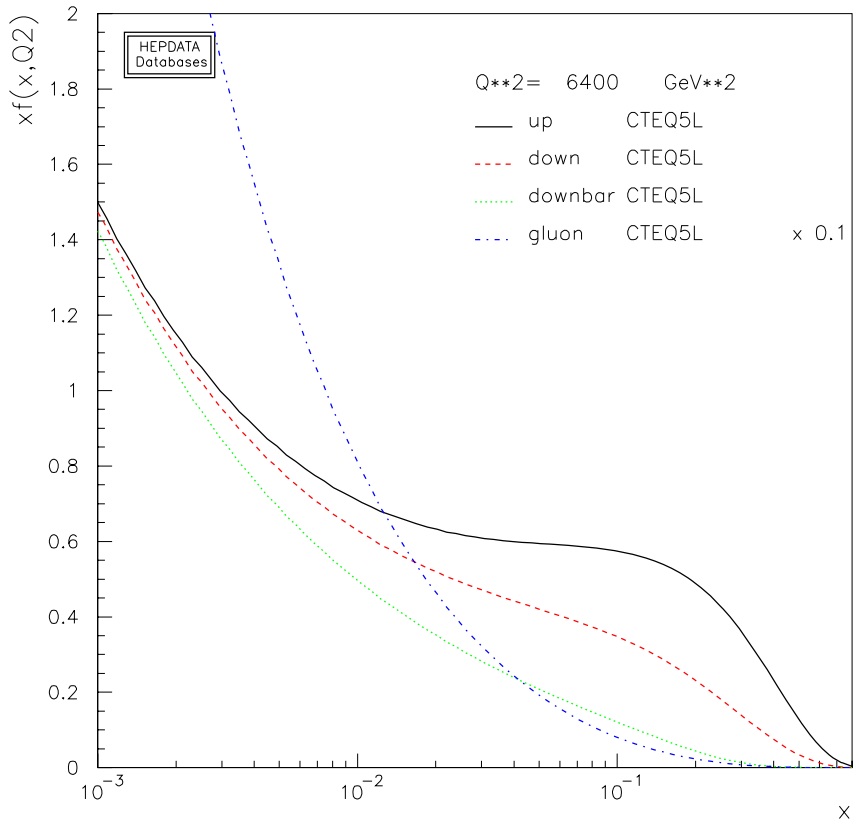


Figure 1.10: CTEQ5L parton distribution function [30].

process	cross-section [pb]
$p\bar{p} \to W + X \to e\nu_e + X$	2731 [31, 32]
$WW \to e\nu_e W$	2.7 [33]
$WZ \to e\nu_e Z$	0.2 [33]
$W\gamma \to e\nu_e\gamma(E_{T,\gamma} > 5$ GeV,$\Delta R_{e,\gamma} > 0.2, \lvert\eta_\gamma\rvert < 1.0)$	43.4 [34]
WH	< 0.2 [35]
$t\bar{t}$	7.56 [36]
single top, t-channel	1.98 [37]
single top, s-channel	0.88 [37]

Table 1.1: Cross-sections for W production at the Tevatron with 1.96 TeV center-of-mass energy.

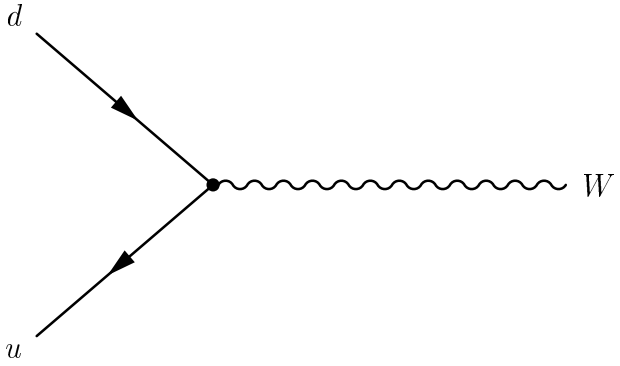


Figure 1.11: Diagram of Drell-Yan W production.

the electrons from the W decay have an angular distribution according to Equation 1.2.2. However, a gluon emitted from one of the initial partons generates the transverse momentum of the W boson. The same is true for a process, in which a gluon splits into a quark-antiquark pair and one of these quarks interacts with a quark from another hadron to produce a W boson. In short, a quark is emitted instead of a gluon. Figure 1.12 shows diagrams of these types.

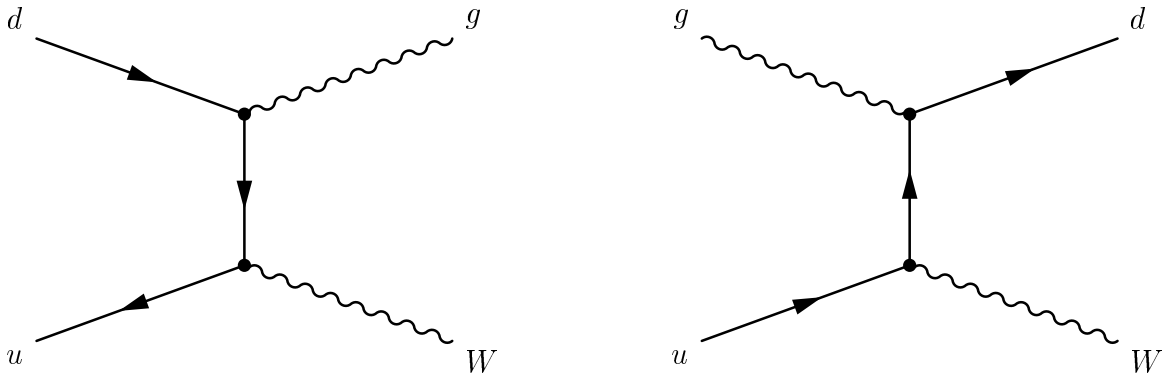


Figure 1.12: Example diagrams for W boson production with a quark or gluon in the final state.

The number of emitted gluons or quarks can be infinitely large, making the calculation of the cross-section for these processes very difficult. These calculations are based on Quantum Chromodynamics (QCD), the theory of the strong interactions. There are different formalisms to resum the divergent terms. The measurement of the transverse momentum spectrum of W bosons is a good test for the resummation formalisms used to predict this spectrum and the QCD theory itself. Figure 1.13 shows a predicted transverse momentum spectrum of W bosons at the Tevatron. This spectrum has been calculated with the *Collins-Soper-Sterman* resummation formalism [38].

The angular distribution in the $W \to e\nu_e$ decay is predicted by next-to-leading order

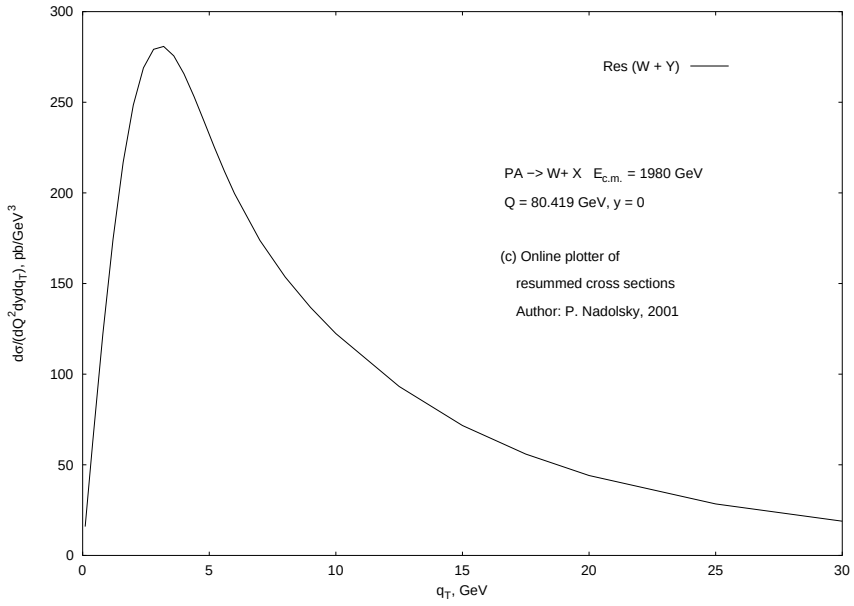


Figure 1.13: Theoretical prediction of the W transverse momentum spectrum [38].

QCD to

$$1 \pm \alpha_1 cos\hat{\theta} + \alpha_2 cos^2\hat{\theta},$$

where $\hat{\theta}$ is the polar-angle of the decay lepton in the Collins-Soper frame [39], that is the angle between the lepton momentum and the beam axis in the W boson rest frame. The parameters α_1 and α_2 are functions of the W boson transverse momentum.

The theoretical prediction for the cross-section of the process $p\bar{p} \to WX \to e\nu_e X$ at the Tevatron with a center-of-mass energy of 1.96 TeV is 2.731 nb [31, 32]. The uncertainty of the calculation is dominated by uncertainties in the parton distribution functions and is estimated to be between $3 - 5\%$ [40].

WW and WZ/γ

In the Standard Model, W bosons may be accompanied by another electroweak gauge boson, so the final states are: WW,WZ, and $W\gamma$. The cross-sections of these processes are very sensitive to the structure of the electroweak interaction. Their values for the Tevatron are given in table 1.1.

Higgs-Strahlung

A virtual W boson might as well radiate a Higgs boson to become real. The cross-section depends on the mass of the Higgs boson. Searches at LEP excluded masses of the Higgs boson below 114.4 GeV [41]. This allows to set an upper bound on the WH production cross-section of approximately 0.2 pb [35].

1.3.2 Top Quark Decay

Another source for W bosons at the Tevatron is the decay of top quarks. Due to its large mass of $m_t = 174.3 \pm 5.1$ GeV [27], the weak decay of top quarks is mediated by real W bosons. The top quark decays into a bottom quark with a ratio of nearly 100%, the decays into a down or strange quark are heavily suppressed by the CKM matrix elements.

Top Quark Pair Production

The dominant source of top quarks at the Tevatron is their pair production via the strong interaction. Theoretical predictions of the cross-section are $7.56^{+0.10}_{-0.55}$ pb [36] and $6.70^{+0.71}_{-0.88}$ pb [42]. The first prediction was evaluated for a center-of-mass energy of 2.0 TeV, instead of 1.96 TeV. The errors given for the second value are the maximal changes from using different parton distribution functions. Figure 1.14 shows two main diagrams of top pair production at the Tevatron. The dominant production mode is through quark-antiquark annihilation (left diagram). Its contribution to the total cross-section is around 90%. The initial state partons have to carry a large fraction x of the hadron momentum to exceed the energy threshold of top pair production. As can be seen in figure 1.10, the probability for large values of x is higher for quarks than for gluons. Two W bosons from the top decays are in the final state of the top pair production process.

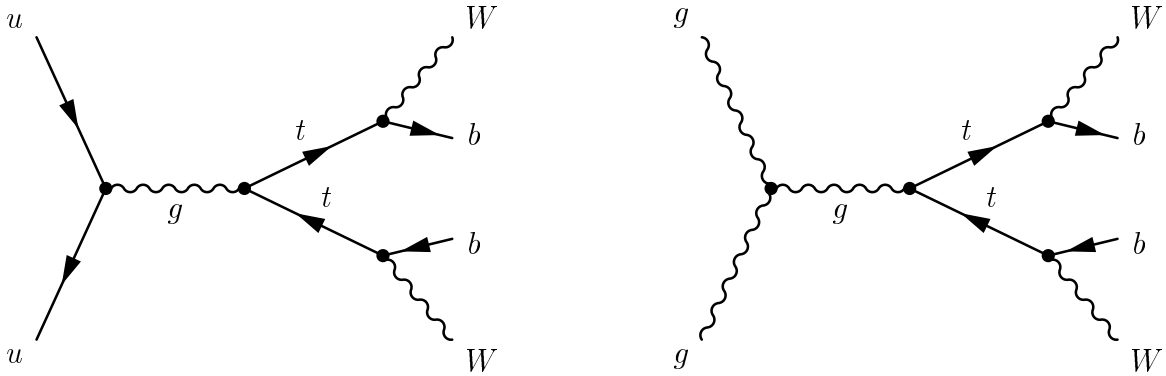


Figure 1.14: Diagrams of top pair production at the Tevatron.

Single Top Quark Production

So far, only the top quark pair production has been discovered in experiment. Nevertheless, single top quarks can as well be produced by electroweak interactions in conjunction with a bottom quark. In this case, a virtual W boson interacts with a bottom quark producing a top quark in the final state, see Figure 1.15. In Run I, an upper limit on the electroweak top quark production has been set by the CDF and D0 experiments [43, 44].

The s-channel process has the least theoretical uncertainties. This process is very similar to the direct production of on-shell W bosons. However, to decay into a top quark and a bottom quark pair, the W has to be off-shell. The large virtuality of the W boson heavily reduces the cross-section. It is expected to be 0.88 pb [37] at the Tevatron. In the the W-gluon fusion / t-channel process, a gluon splits into a $b\bar{b}$ pair. The bottom quark interacts with a virtual W boson emitted from another quark and becomes a top quark. This process has larger theoretical uncertainties in the gluon PDF, but a larger cross-section as well. Its predicted value is 1.98 pb [37].

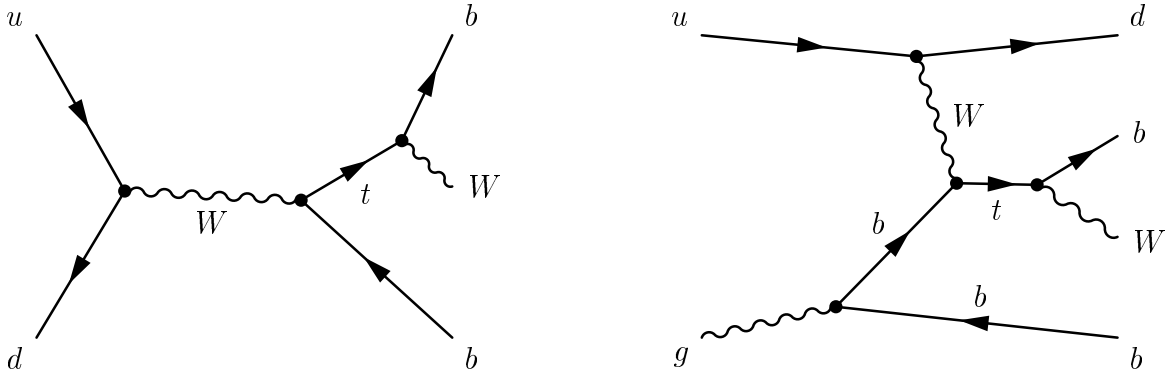


Figure 1.15: Diagrams of electroweak top quark production. The main contributions at the Tevatron arise from the s-channel process (left) and the W-gluon fusion process (right).

Chapter 2

The Experiment

The *Fermi National Accelerator Laboratory*, abbreviated Fermilab, is located approximately 40 miles west of Chicago in the State of Illinois. More than 2,500 scientists from throughout the United States of America and around the world use Fermilab's facilities to carry out research in high-energy physics. An aerial shot of the Fermilab is shown in Figure 2.1.

Figure 2.1: Aerial shot of the Fermilab. The grey circle in the back is the inner maintenance road of the Tevatron, the one in the front indicates the outer maintenance road of the main injector and the recycler.

2.1 The Accelerators

To reach the center-of-mass energies achieved at Fermilab, a system of accelerators is needed. The last stage of the acceleration takes place in the *Tevatron*, a collider with a circumference of about six kilometers. Here, the protons and antiprotons reach an energy of nearly 1 TeV. During Run II, the two beams collide with a center-of-mass energy of 1.96 TeV.

Another quantity characterizing a collider is *luminosity*. Luminosity $\mathcal{L}$ is the product of incident beam flux with the mean target density. The event rate for a particular type of event with the cross-section σ is given by the product $\mathcal{L} \cdot \sigma$. The typical luminosity for Run Ib was $\mathcal{L} = 1.6 \times 10^{31}$ cm^{-2}s^{-1}. During Run Ia and Ib from 1992 to 1996, the *Collider Detector at Fermilab*, CDF, collected data corresponding to an *integrated luminosity* of approximately 130 pb^{-1}.

For Run II, the accelerators were upgraded. The *main ring* was replaced by the *main injector* and a new antiproton storage ring, the *recycler*, was built. These two improvements should allow to increase the instantaneous luminosity up to $\mathcal{L} = 20 \times 10^{31}$cm^{-2}s^{-1}.The plan for the first phase of Run II, called Run IIa, is to deliver an integrated luminosity of 2 fb^{-1} [45]. Table 2.1 summarizes the collider characteristics in Run Ib and the goals for Run II.

Run	Ib(1993-1995) (6×6)	IIa(goals) (36×36)	IIb(goals) (140×121)
Energy [GeV]	900	998	998
Antiproton Bunches	6	36	121
β^* [cm]	35	35	35
Bunch Length(rms) [cm]	60	37	37
Bunch Spacing [ns]	~ 3500	396	132
Interactions / Crossing	2.5	2.3	1.3
Typical Luminosity [cm^{-2}s^{-1}]	1.6×10^{31}	8.6×10^{31}	16.1×10^{31}

Table 2.1: Operational performance of the Tevatron in Run I and goals for Run II [45].

Run IIa started in June 2001. Unfortunately, the instantaneous luminosities achieved by the Tevatron have not yet met the design goals, but are steadily increasing. This can be seen in figure 2.2 that presents the instantaneous luminosities of every store of Run IIa. Figure 2.3 displays the increase of the integrated luminosity delivered by the Tevatron since the start of Run II.

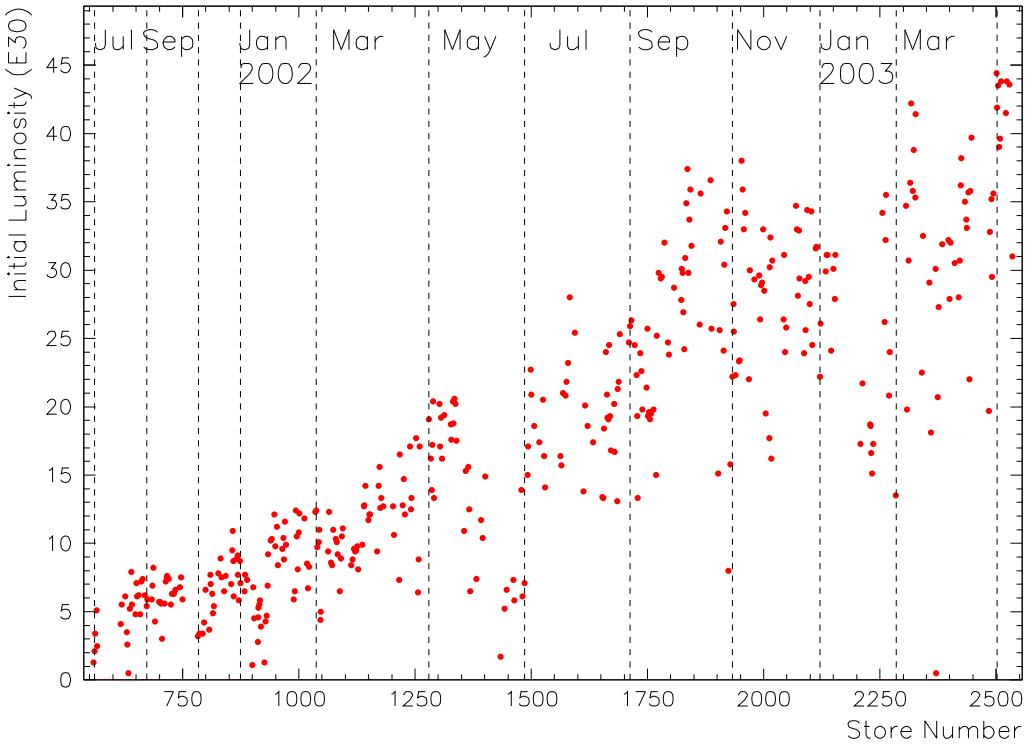


Figure 2.2: Initial luminosity per store in $cm^{-2}s^{-1}$.

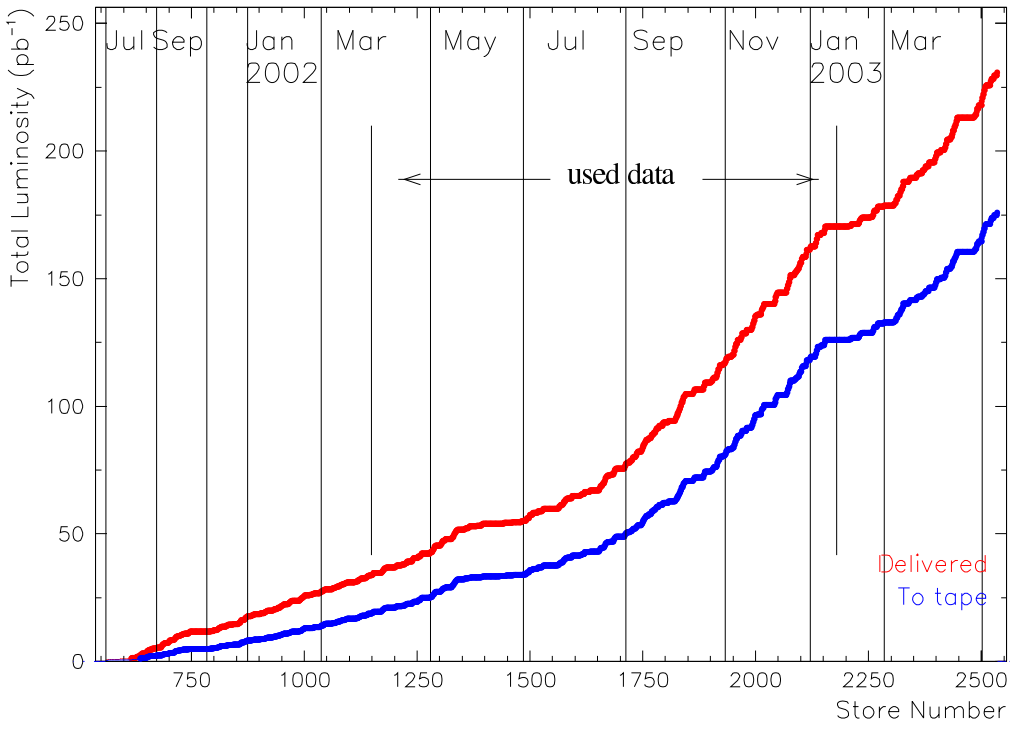


Figure 2.3: Delivered (upper curve) and recorded (lower curve) integrated luminosity since the start of Run II. The data taken from March 2002 to January 2003 was used for the analysis presented in this thesis.

2.2 The Collider Detector at Fermilab in Run II

There are two multipurpose detectors located at the Tevatron: CDF and D0. The top quark was found in 1995 with these two detectors [46, 23] during Run I of the Tevatron .

CDF is a general purpose solenoidal detector. It combines charged particle tracking with calorimetry and muon detection. The detector has both azimuthal and forward-backward symmetry. The CDF is built and maintained by a collaboration of more than 50 institutions in eleven countries. The only German institute in the collaboration is the *Institut für Experimentelle Kernphysik* in Karlsruhe.

To deal with the higher luminosities of the Tevatron in Run II, some parts of the Run I detector and most of the data acquisition system have been replaced, However, there have as well been improvements to extend the coverage and capabilities of the existing subdetectors. A more detailed description of the CDF II detector can be found in its technical design report [48]. The Run I detector is described in detail elsewhere [49, 50].

Figure 2.4 shows an elevation view of one half of the CDF II detector. In the CDF coordinate system, the polar angle θ in cylindrical coordinates is measured from the proton beam axis (z-axis) and the azimuthal angle φ from the plane of the Tevatron. Throughout this thesis, *longitudinal* means parallel to the proton beam and *transverse* means perpendicular to the proton beam. The pseudorapidity is defined by $\eta = -\ln(\tan\frac{\theta}{2})$.

2.2.1 The Tracking System

CDF II uses the same solenoid as in Run I, but the complete tracking system of the Run I detector has been replaced. The superconducting solenoid is of length 4.8 m and radius 1.5 m and generates a 1.4 T magnetic field. As can be seen in Figure 2.5, the tracking system consists of silicon detectors near the interaction region and a drift chamber. Between the drift chamber and the solenoid, a time-of-flight detector has later been added to the design to improve the particle identification capabilities of CDF [51].

The silicon vertex detector system [52] consists of eight layers arranged in cylinders spanning radii from 1.35 cm to 28 cm and lengths from 90 cm to nearly two meters for a total of six square meters of silicon and 722.000 readout channels.

Layer 00, the innermost layer, utilizes radiation tolerant silicon and low-mass readout cables between the sensors and readout electronics. Layer 00 was later added to the design of the vertex detector system to enhance its resolution and longevity [51]. To replace this layer relatively easily, it is supported by the beam pipe.

The new silicon tracker, *SVX II*, consists of five double sided layers at radii from 2.4 to 10.7 cm. The SVX II is 96 cm long and covers the pseudorapidity interval $|\eta| < 2$. The layers are assembled in three cylindrical barrels with beryllium "bulkheads"at each

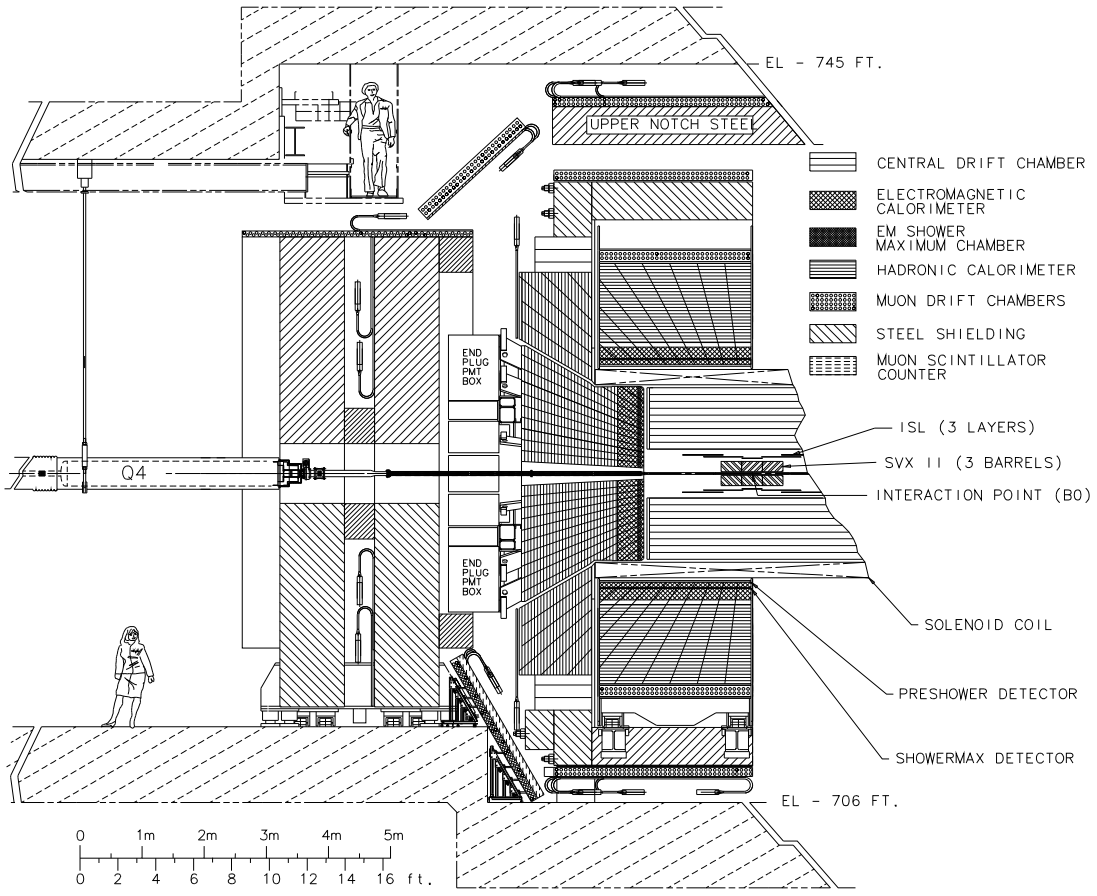


Figure 2.4: Elevation view of one half of the CDF II detector.

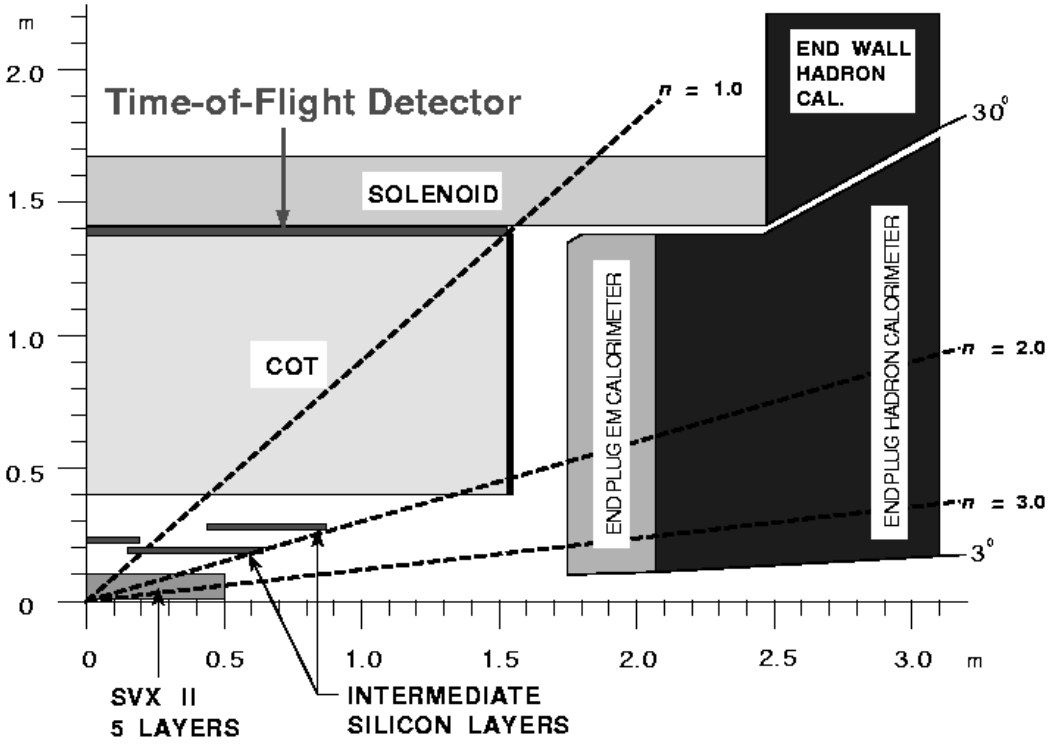


Figure 2.5: Longitudinal view of the CDF II tracking system.

end for support and cooling of the modules. The modules consist of two readout units of silicon with a hybrid at each end mounted directly atop the silicon to avoid gaps. The design improves upon the previous SVX′ [53] by enlarging coverage and elimination of gaps.

With an outer radius for the SVX II of 10.7 cm, additional tracking information is needed to robustly link to tracks found in the drift chamber. This is achieved by an additional silicon layer at a radius of 22 cm covering the interval $|\eta| < 1$ and two layers in the forward and backward direction. These *Intermediate Silicon Layers* (ISL) allow stand-alone silicon tracking over the whole region of $|\eta| < 2$. The Institut für Experimentelle Kernphysik in Karlsruhe participated in the construction of the ISL [54].

The impact parameter resolution of the silicon system is expected to be better than the resolution of the SVX′. The resolution in Run I has been $\sigma_D(p_T) = (13 + 40/p_T)\ \mu$m [55] , where p_T is the transverse momentum of the particle in units of GeV/c. The impact parameter D is the distance of closest approach of the track helix to the beam axis measured in the plane perpendicular to the beam. Figure 2.6 shows the positive effect of Layer 00 on the expected impact parameter resolution.

The drift chamber used during Run I, the *Central Tracking Chamber* (CTC), would suffer from severe occupancy problems at $\mathcal{L} \geq 1 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$. Hence, it has been replaced by the *Central Outer Tracker* [57], COT, that uses smaller drift cells and a faster gas to reduce drift times. The COT is a cylindrical open-cell drift chamber with inner and outer radii of 44 and 132 cm. It is designed to find charged particles in the region $|\eta| \leq 1$ with transverse momenta as low as 400 MeV/c. The COT is segmented into four axial and four stereo super-layers. Each super-layer contains 12 sense wires alternated with 13 potential wires which shape the field within the cell, yielding a total of 96 measurement. The COT uses Argon-Ethane (50:50) as the drift gas, during running with a 396 ns beam crossing time. This will give a maximum drift time of 180 ns. The resolution of the COT is expected to be comparable to the resolution of the CTC. Using the silicon detectors and the COT, the overall momentum resolution for charged particles is $\delta p_T/P_T^2 < 0.1\%$ GeV/c.

Between the COT and the solenoid, a *Time-of-Flight system* (TOF) is installed mainly for particle identification. It consists of scintillator panels which provide both timing and amplitude information. The timing resolution is 100 ps. The detector covers the central region out to $|\eta| < 1.1$ and will be capable of identifying kaons from pions by their flight time difference with at least 2σ separation up to kaon momenta of 1.6 GeV/c. Figure 2.7 illustrates its performance using data.

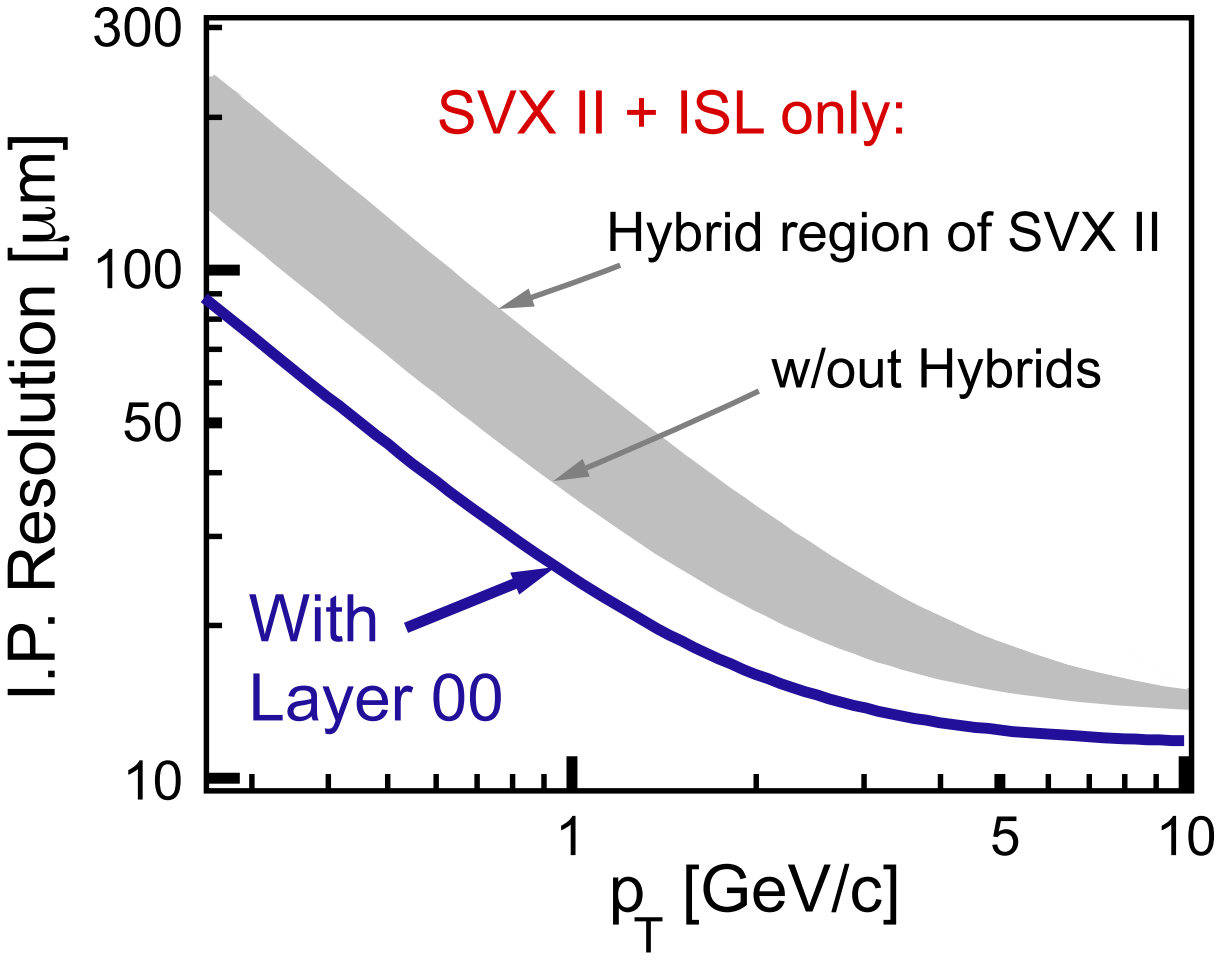


Figure 2.6: Effect of Layer 00 on the impact parameter resolution as a function of transverse momentum. The gray region shows the smearing due to mounting the readout electronics on the SVX II and ISL ladders [56].

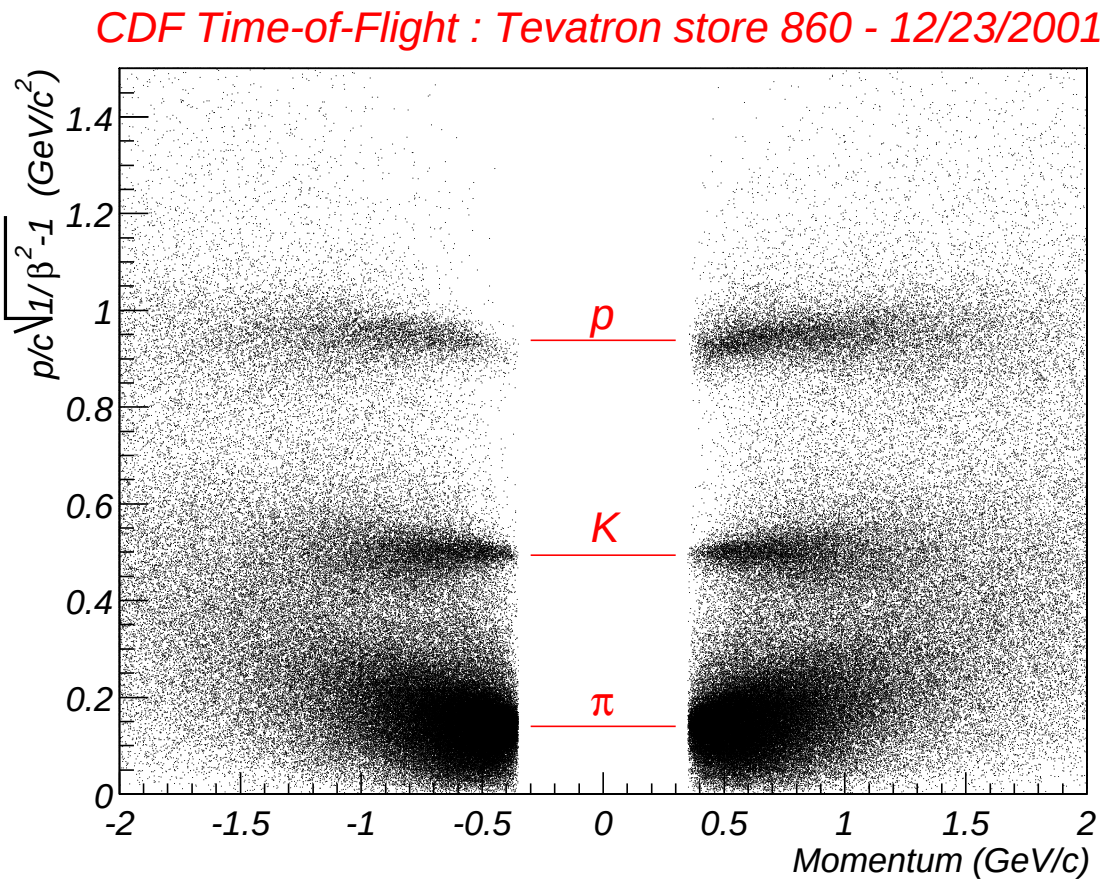


Figure 2.7: Performance of the time of flight system (TOF).

2.2.2 The Calorimeter Systems

The solenoid and tracking volume of CDF II is surrounded by calorimeters which cover 2π in azimuth and range in pseudorapidity from $\eta = -3.64$ to $\eta = 3.64$. They are segmented in azimuth and in pseudorapidity to form a projective tower geometry which points back to the nominal interaction point. As one can see in Figure 2.4, there are two main η regions of calorimeters, central and plug. Each region has an electromagnetic calorimeter (CEM and PEM) with lead absorbers and an hadronic calorimeter (CHA/WHA and PHA) with iron absorbers. Both electromagnetic calorimeters have pre-shower and stereo shower maximum detectors to improve their spatial resolution. Table 2.2 summarizes the properties of the calorimeters in Run II.

System	η Range	Thickness	Energy Resolution
CEM	$\lvert\eta\rvert < 1.1$	19 X_0, 1 λ	$13.7\%/\sqrt{E_T} \oplus 2\%$
PEM	$1.1 < \lvert\eta\rvert < 3.64$	21 X_0, 1 λ	$16\%/\sqrt{E} \oplus 1\%$
CHA	$\lvert\eta\rvert < 0.9$	4.5 λ	$50\%/\sqrt{E_T} \oplus 2\%$
WHA	$0.7 < \lvert\eta\rvert < 1.3$	4.5 λ	$75\%/\sqrt{E} \oplus 4\%$
PHA	$1.2 < \lvert\eta\rvert < 3.64$	7 λ	$80\%/\sqrt{E} \oplus 5\%$

Table 2.2: Summary of CDF calorimeter properties in Run II. The energy resolutions for the electromagnetic calorimeters are for incident electrons and photons, and for the hadronic calorimeters for incident isolated pions. The $\oplus$ signifies that the constant term is added in quadrature. The transverse energy E_T and the energy E are given in GeV.

For Run II, the existing scintillator-based calorimeters in the central region continue to perform well. However, the gas calorimeters in the region $|\eta| > 1$ were incompatible with the crossing rates for Run II, and have been replaced with a new scintillating tile calorimeter. Its electromagnetic section [58] has a energy resolution of approximately $16\%/\sqrt{E}$ with a 1% constant term. The overall segmentation of the calorimeters is shown in Table 2.3.

2.2.3 The Muon Systems

Four systems of scintillators and drift tubes are used to detect muons with the CDF. The central calorimeters act as a hadron absorber for the *Central Muon Detection System* (CMU). In Run I, the CMU consisted of four layers of drift chambers located outside the central hadronic calorimeter. It covered 84% of the solid angle for the pseudorapidity interval $|\eta| < 0.6$ and could be reached by muons with transverse momenta greater than

System	$\|\eta\|$ Range	$\Delta\phi$	$\Delta\eta$
CEM/CHA/WHA	0.0 - 1.1(1.2h)	15°	~ 0.1
PEM/PHA	1.1(1.2h) - 1.8	7.5°	~ 0.1
PEM/PHA	1.8 - 2.1	7.5°	~ 0.16
PEM/PHA	2.1 - 3.64	15°	$0.2 - 0.6$

Table 2.3: CDF II calorimeter segmentation.

1.4 GeV/c. In 1992, the system was upgraded by adding 0.6 m of steel behind the CMU and additional four layers of drift chambers behind the steel. This new system has been called *Central Muon Upgrade* (CMP). For $|\eta| < 0.6$ the CMP covered 63% of the solid angle while both systems overlapped in 53% of the solid angle. In addition, the pseudorapidity range of $0.6 < |\eta| < 1.0$ was covered by the *Central Muon Extension* (CMX) to 71% of the solid angle.

The changes for Run II in the muon systems represent incremental improvements. New chambers have been added to the CMP and CMX systems to close gaps in the azimuthal coverage and the shielding is improved. The forward muon system has been replaced with the *Intermediate Muon System*, IMU, covering a range from $1.0 \leq |\eta| \leq 1.5$. Table 2.4 gives an overview of the different muon systems in Run II.

	CMU	CMP/CSP	CMX/CSX	IMU
coverage	$\|\eta\| < 0.6$	$\|\eta\| < 0.6$	$0.6 < \|\eta\| < 1.0$	$1.0 < \|\eta\| < 1.5$
drift tubes (Run I)	2304	864	1536	none
drift tubes (Run II)	2304	1076	2208	1728
counters (Run I)		128	256	none
counters (Run II)		269	324	864
min p_T	1.4 GeV/c	2.2 GeV/c	1.4 GeV/c	$1.4 - 2.0$ GeV/c

Table 2.4: Design parameters of the CDF II muon detectors. The first row is the pseudorapidity coverage of the different detector systems. The total number of drift tubes used in Run I and II is listed in the second and third rows. The fourth and fifth rows list the total number of counters used in Run I and II. The last row states the minimal transverse momentum of a detectable muon.

2.3 Data Acquisition and Monitoring

2.3.1 Electronics and Triggering

The CDF electronics systems have substantially been altered to handle Run II accelerator conditions. The increased instantaneous luminosity requires a similar increase in data transfer rates and the reduced separation between accelerator bunches a new architecture for the readout system. Figure 2.8 shows the flow chart of the new trigger system and data acquisition (DAQ).

Due to the higher raw collision rate, the trigger in Run II must have a larger rejection factor while maintaining high efficiency for the broad range of physics topics. Figure 2.9 shows the block diagram of the trigger system for Run II. The lowest level trigger, "Level 1", uses output from the muon detectors for muon triggers and from all the calorimeters for electron and jet triggers. An addition to the previous run is the reconstruction of tracks using COT information already in "Level 1". This is done by the *XFT*, an improved version of the *central fast tracker* (CFT) [59]. A typical rate of the "Level 1" triggers is at present 6 kHz.

The second level trigger, "Level 2" uses the calorimeter trigger information with greater sophistication by running a cluster finder. In addition, data from the shower maximum detector (CES) can be used to improve the identification of electrons and photons. The most challenging addition for "Level 2" is the *Silicon Vertex Tracker* (SVT). The SVT provides the ability to trigger not only on electrons, muons, and jets, as during Run I, but to select events with tracks which have large impact parameters. This opens a complete new window for physics measurements at a hadron collider, like charm physics with high statistics and the study of hadronic B decays. The "Level 2" output rate is approximately 300 Hz.

The third level of triggering, the "Level 3", uses the algorithms run in the "offline" reconstruction. The software is run on a PC farm. The output rate of the "Level 3" trigger is approximately 75 Hz at present. The accepted events are then transferred to the *Feynman Computing Center* via network and stored on tape. To facilitate the handling of the huge data volumes collected with the CDF, the data coming from "Level 3" is currently split into eight different streams. The triggers an event has passed decide to which streams this event belongs, e.g. all events passing any of the highly energetic lepton triggers end up in "stream B".

2.3.2 Online Monitoring of the Data Taking

A complex multi-purpose detector, like the CDF, consists of many different detector systems. To take data with high efficiency and high quality, it is necessary to quickly

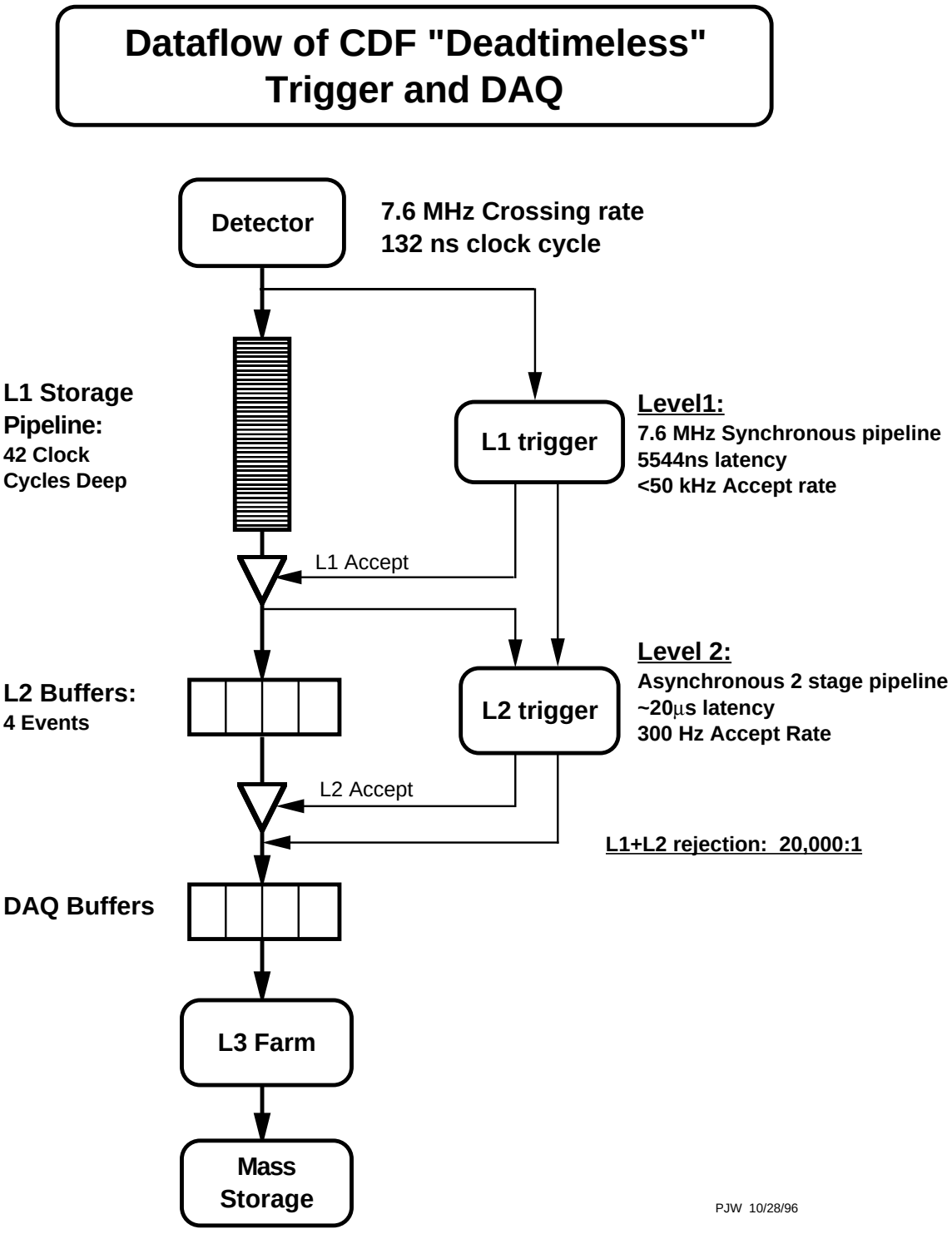
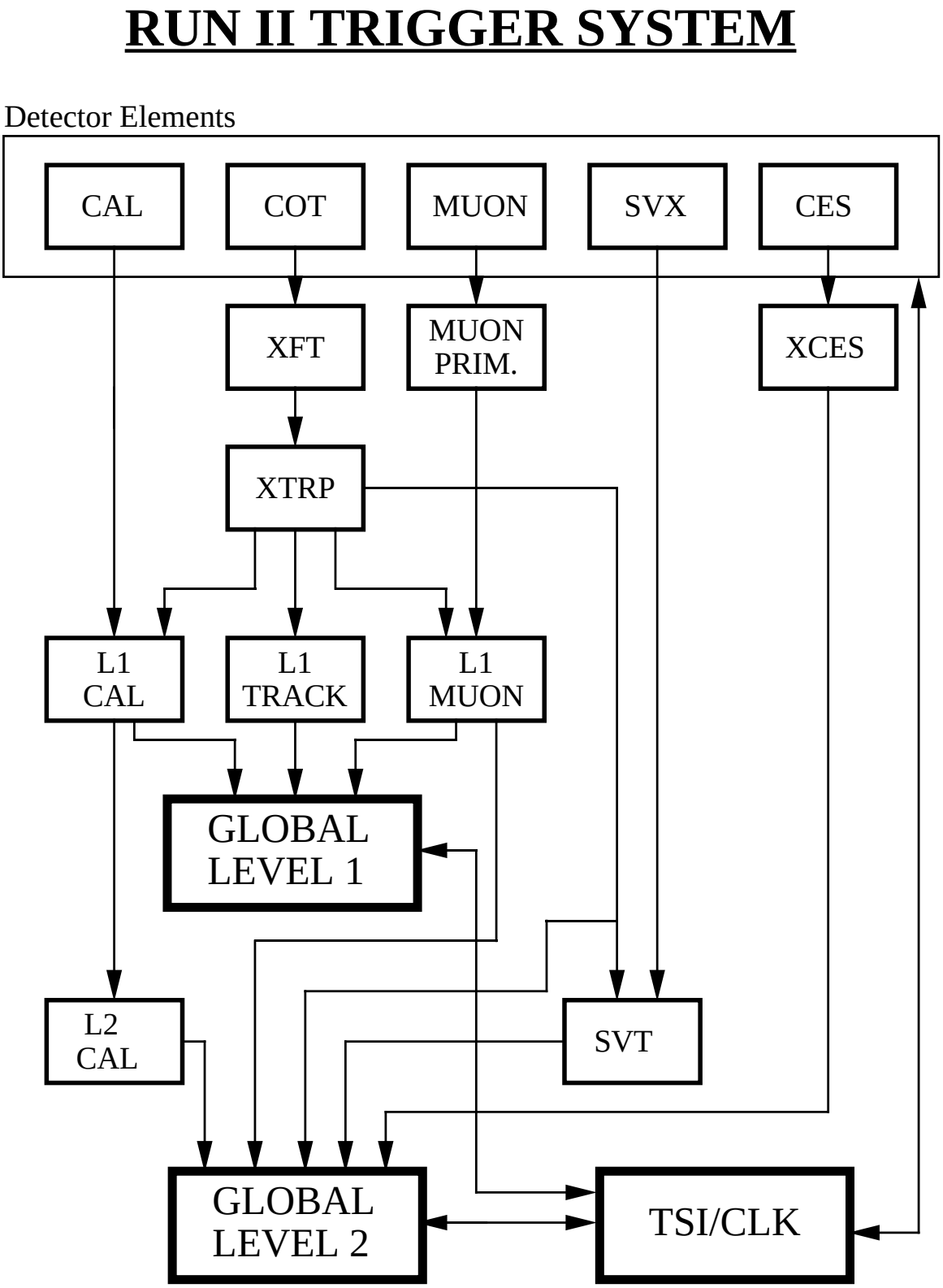


Figure 2.8: Functional block diagram of the CDF II data flow.



PJW 9/23/96

Figure 2.9: Block diagram of the CDF trigger system for Run II

spot problems with one of these subdetectors. This can be achieved by monitoring the data during data taking. During Run I, this was done by programs that were directly connected to the DAQ. At CDF, all processes receiving data from the DAQ are called *consumers*. This term is now jargon in Run II for all processes monitoring the detector online [60, 61, 62].

Requirements for the Online-Monitoring

These consumer monitors have to fulfill the following requirements:

1. The consumer monitor should present the data in a way that allows to easily spot problems with a subdetector. In most cases, histograms are filled with the data, e.g. for a calorimeter a histogram of the occupancy of each channel allows to spot dead or hot regions. Furthermore, even plots only useful for experts should be routinely filled, to allow the experts to identify and solve a problem.

2. The monitoring processes should not lower the data taking rate. On one hand, this means that copies of the physics events are sent to the consumer, on the other hand, that the consumer process itself cannot affect the DAQ, e.g. an abnormal termination of a consumer process does not crash the data logger.

3. It should be easy to add or remove the monitoring of certain detector components to have maximum flexibility.

4. The consumer processes should be very stable to allow to collect enough statistics to even spot small problems.

5. Each consumer should offer the same user interface for viewing its results to lower the learning curve for the shift crew. In Run I, each monitor had its own graphical user interface.

The last point caused a lot of maintenance problems with the consumer code during Run I. To avoid this problem in Run II, a group was formed to create a uniform framework for all processes monitoring the detector with event data online.

Design of the Consumer Framework

As the consumer framework group consists of relatively few people from Fermilab, Waseda University, and the Universität Karlsruhe, the basic idea has been to heavily use existing packages to implement the framework. The standard software framework of CDF II, *AC++*, [63] has been used to read the data. Data analysis, graphical user interface, and

socket connections have been programmed using *ROOT*. Its functionality is very well summarized on its official web page[64]:

"The ROOT system provides a set of OO frameworks with all the functionality needed to handle and analyze large amounts of data in a very efficient way. Having the data defined as a set of objects, specialized storage methods are used to get direct access to the separate attributes of the selected objects, without having to touch the bulk of the data. Included are histograming methods in 1, 2 and 3 dimensions, curve fitting, function evaluation, minimization, graphics and visualization classes to allow the easy setup of an analysis system that can query and process the data interactively or in batch mode."

We designed a system to match the requirements listed above. All components are written in C++ using ROOT. The basic framework is independent of the CDF software environment and can be reused by other experiments. Figure 2.10 visualizes the interaction of the components of the consumer framework and figure 2.11 shows the class diagram of the consumer framework [65].

The Consumer-Server/Logger

The item 2 in the requirements list, monitoring the data without impeding data taking, was fulfilled by the *Consumer-Server/Logger (CSL)* [66]. The main task of the CSL is to log the data coming from the "Level 3" farm and transfer it to the Feynman computing center for storage on tape. However, a subset of the events is copied and can be sent via network to the monitoring programs. A special AC++ input module has been written that allows to connect to the CSL to get events during data taking. For the monitoring of some detector components, only events that have passed certain triggers can be used. The CSL allows to specify the triggers that an event has to have passed to be sent to the consumer. The overall rate of events sent to consumer monitors is about 10 Hz.

The Consumers

The monitoring programs get the data from the CSL and use this to fill histograms and tables. All consumer monitoring programs contain a module that inherits from the *ConsumerFrameworkModule*. The ConsumerFrameworkModule starts the server and administers a list of all ROOT objects used for the monitoring. After a certain number of events has been processed, the histograms, canvases, and all other ROOT objects are sent to the server via a socket connection. At the end of each run, the monitoring results are saved to a file. These files are archived on tape for later inspection and studies of the detector performance.

The actual monitoring is done by objects inheriting from *BaseMonitor2*. The ConsumerFrameworkModule contains a list of all these objects and calls their *addEvent* func-

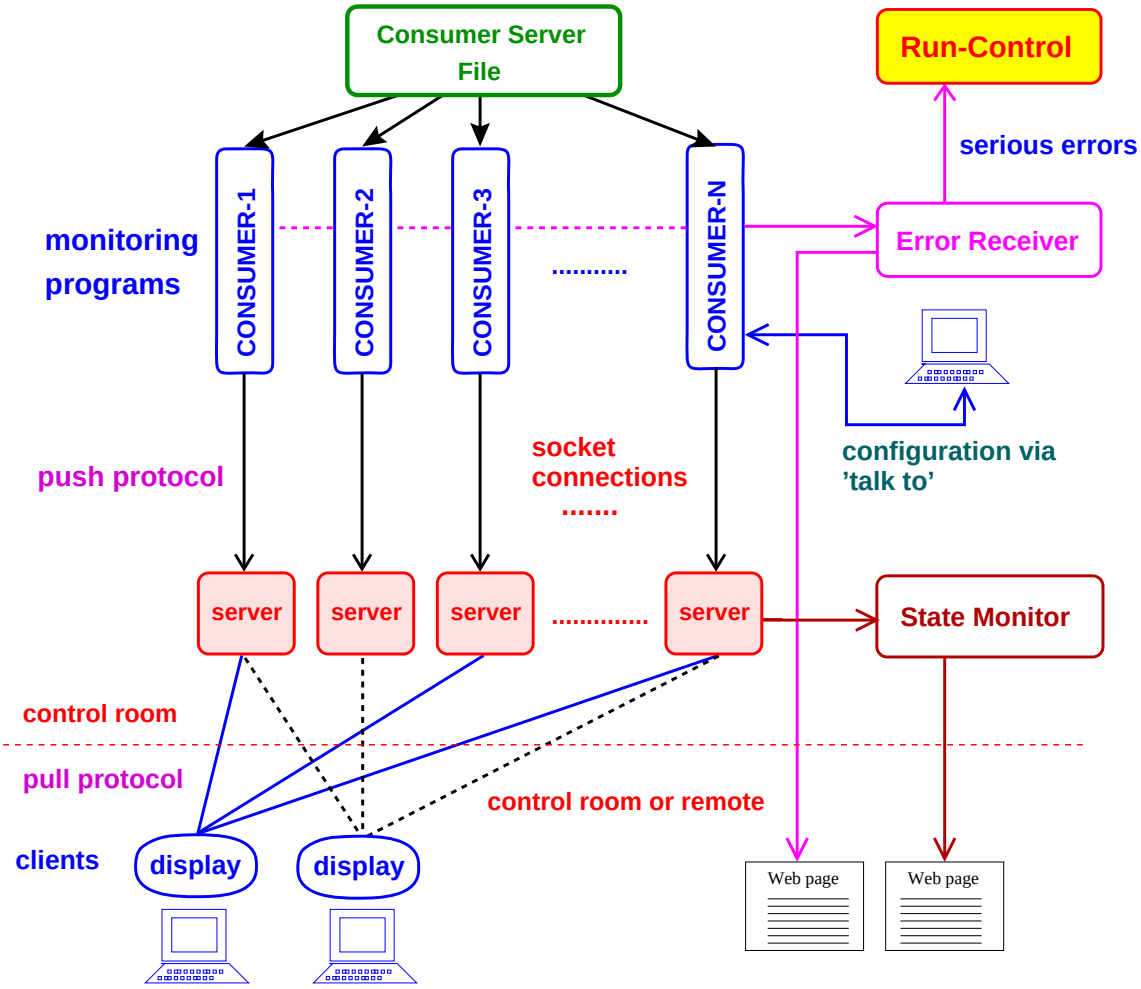


Figure 2.10: Overall design of the consumer framework.

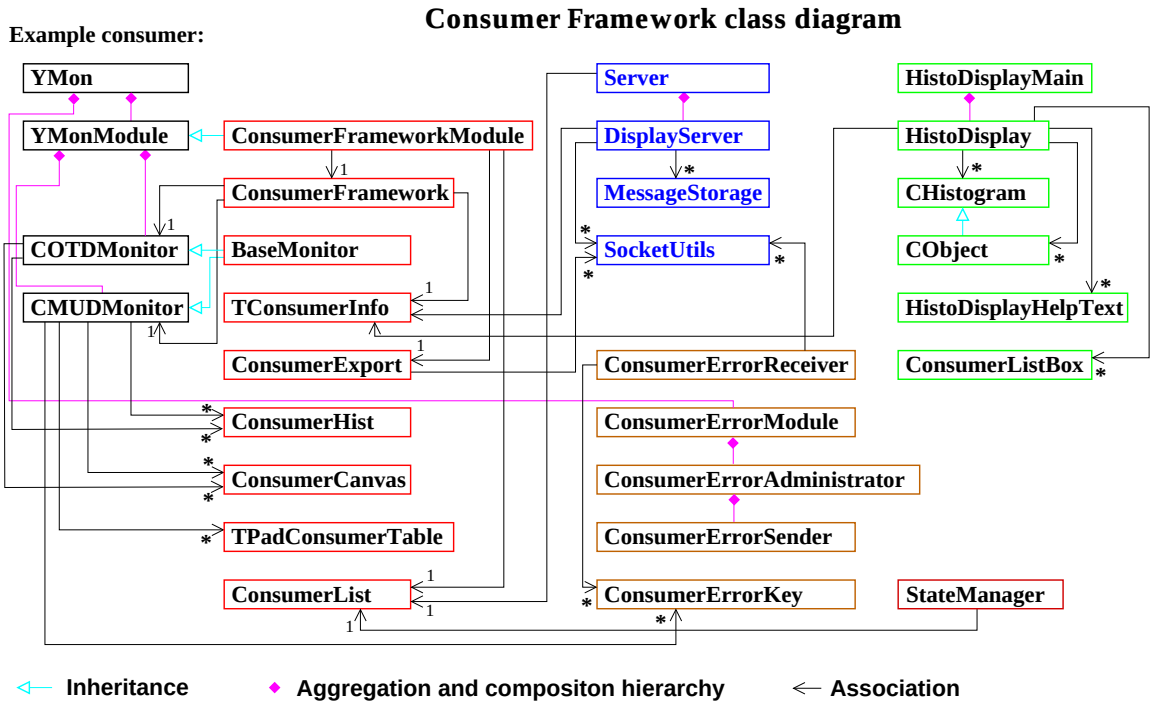


Figure 2.11: Class diagram of the consumer framework [65].

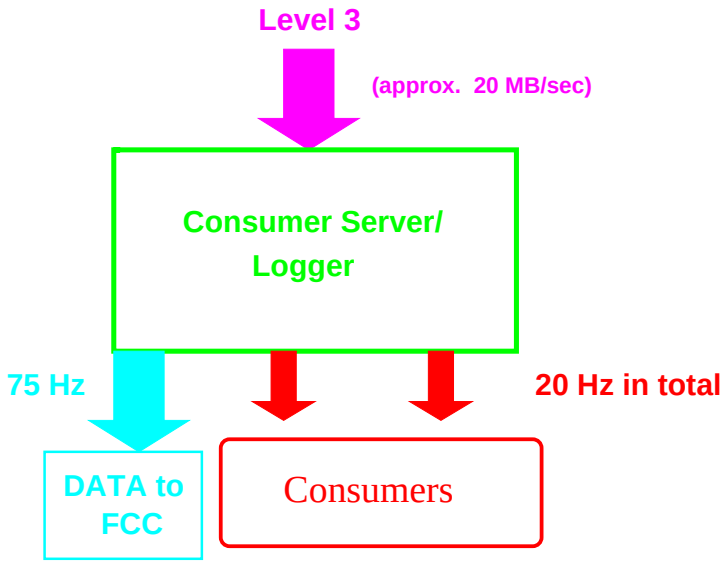


Figure 2.12: Diagram of the data flow through the Consumer-Server/Logger.

tion every event. In this function, the event data is used to evaluate the detector performance, e.g. fill histograms. The histograms and text are usually put on a ROOT canvas. To achieve a unified look-and-feel, a special canvas class has been implemented, that displays the consumer name, detector name, run and event number, and date and time in the uppermost part of the canvas. An example can be seen in the lower right window on figure 2.14.

This special canvas class is one example of classes implemented for the framework, that facilitate the writing of consumer monitors. Other classes take care of starting the server, sending the monitor output, and all other framework related tasks. The authors of the monitoring code only have to write the part that analyzes the data. This framework and the support by the framework group were crucial for allowing detector experts to write the monitoring code, as many of them were not very familiar with the CDF offline framework and C++. On the other hand, only the experts for a detector can implement all the plots needed to identify and solve problems with this detector (requirement 1).

Examples for consumer monitors are:

- *YMon* monitoring the rates and distributions of each detector sub-system looking for bad channels, cards, voltage supplies, etc.
- *TrigMon* monitoring the various trigger quantities in each trigger bank looking for hot/failed channels, loose cables, trigger bits fired, bunch counter mismatches, photomultiplier spikes, etc.
- *XMon* monitoring the trigger rates
- *Stage0* fitting the drift constants for the gas in the drift chamber (COT)
- *LumMon* monitoring the luminosity measurements and the performance of the *Cherenkov Luminosity Counters* (CLC)
- *ObjectMon* monitoring the "Level 3" reconstructed objects, e.g. jets, electrons, muons, and tracks
- *SVXMon* and *SiliMon* monitoring the performance of the silicon detectors
- *BeamMon* fitting the position of the beam line for each run
- *DAQMon* monitoring the readout performances of almost all of the different crates, the different readout times, and the event sizes

The Display Server

The server process runs on the same machine as the monitoring program and is started by this program at the beginning. The server receives the monitoring output from the consumer and updates its internal list of ROOT objects accordingly. The server reacts on requests by display clients and sends the requested objects to the display programs. This connection is as well implemented with sockets. The special server processes guarantee that the monitoring is not slowed down by many requests from displays. Furthermore, a broken connection might crash the server process. In this case, the server is restarted by the monitoring program. If the server had been integrated in the monitoring process, this process would have to be restarted in violation with requirement 4. Due to this design, the number of display clients is only limited by network traffic and bandwidth and the stability of the consumer monitor is greatly enhanced.

The State Manager

After receiving the first data from the monitoring program, the server sends a status message to a central process, the *state manager*. This status message contains the name of the consumer, the number of the analyzed run, the number of events processed by this consumer and the location of the server. The state manager collects this information from every running server and updates a web page accordingly. A screen-shot of this web page is displayed in figure 2.13. Every two minutes the server sends a new status message. Clients can use this page to find the right server to connect to.

The Consumer-Display

The display program offers a graphical user interface (GUI) to view the monitor output. It is based on ROOT. The user can connect to a server or open a file containing the monitor results. In case of a server connection, a list of all available objects is send from the server to the consumer-display on request by the display client. All available objects are displayed in a tree structure and the user can start to browse the results. Figure 2.14 shows a screen-shot of a typical display session. The display automatically requests newer versions of the displayed objects and updates them. A special feature is the so called "slide show". In this mode, the display cycles through all objects (canvases) in a special directory showing a new plot every ten seconds. Thus, the person on shift, the *consumer operator*, can easily inspect all important plots. Figure 2.15 gives an impression of the work place of the consumer operator in the CDF control room.

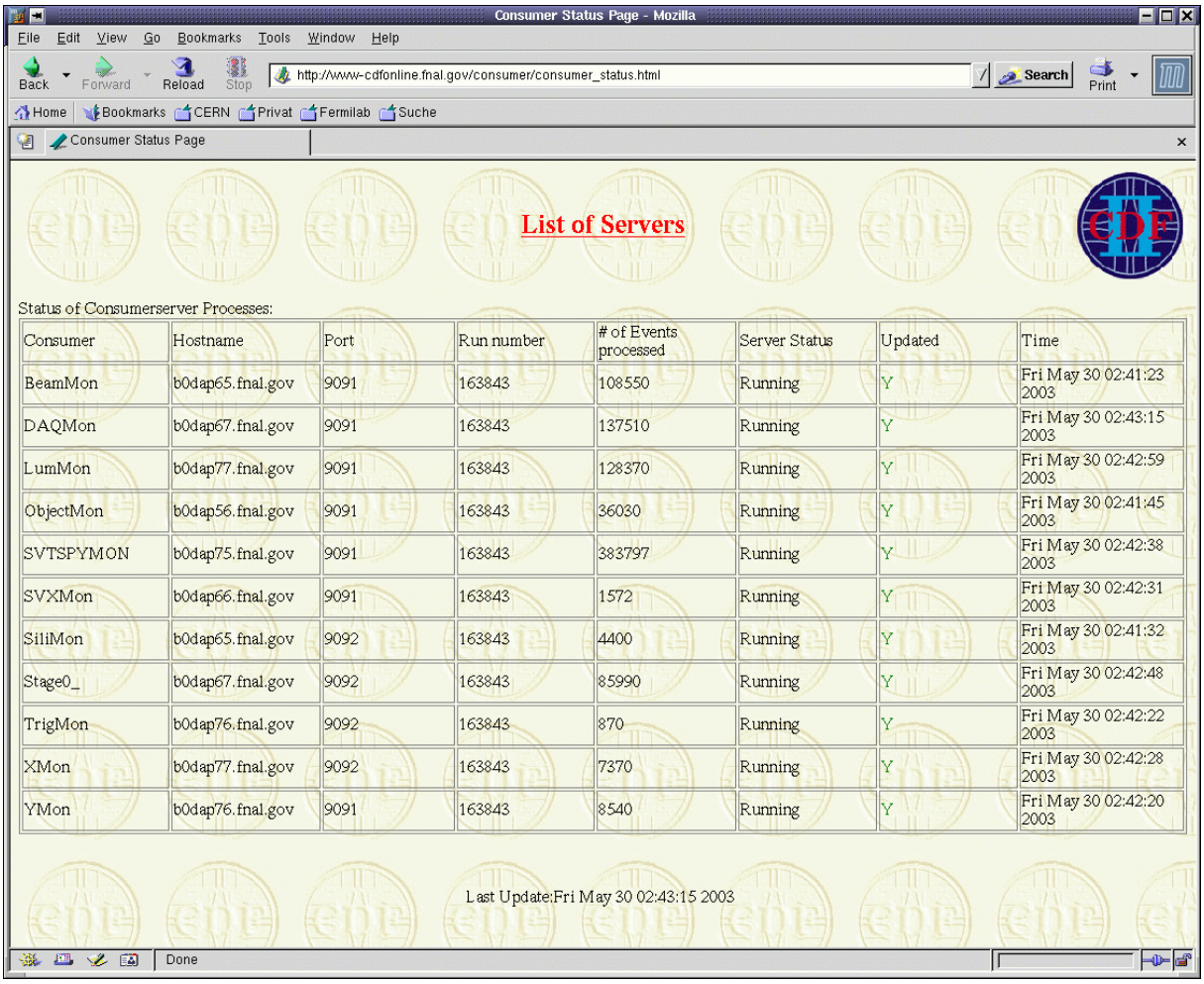


Figure 2.13: Screen-shot of the consumer status page.

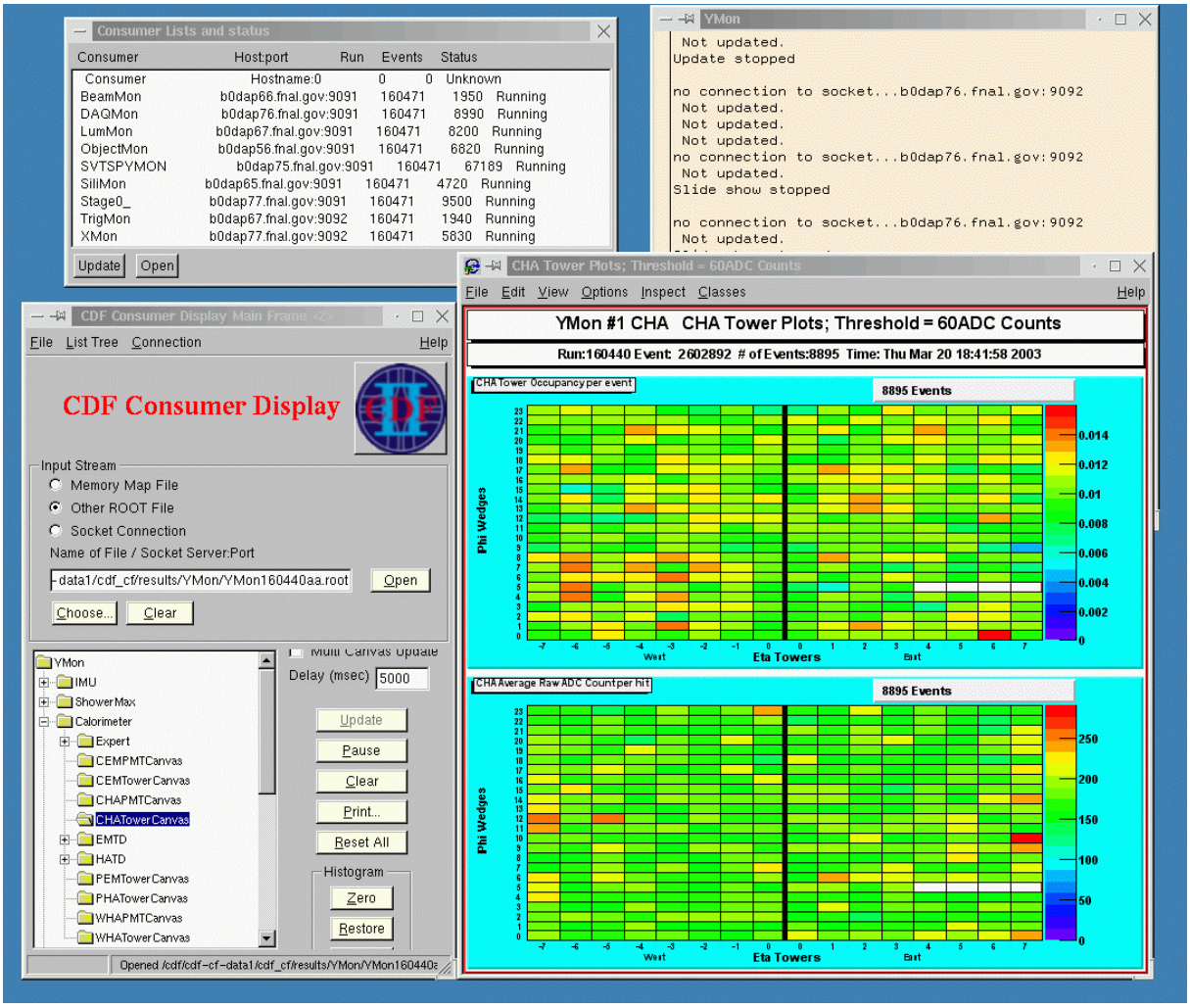


Figure 2.14: Screen-shot of the display program. The lower left window shows the display GUI , the lower right window is a typical canvas created by a consumer.

Figure 2.15: Picture of the CDF control-room. One can see the monitors used by the consumer operator to watch the detector performance during data taking.

Error Receiver

To facilitate the spotting of errors or problems with one detector, some consumers automatically check the histograms. This is for example done by comparing the actual histogram with a reference plot from a former run or by fitting a function describing the expected distribution to the data. The consumer framework offers two ways to communicate a found error.

- The histogram showing an unexpected distribution can be put on a special canvas with a red background and a text describing the problem. After this object has been sent to the server, the server forwards this *warning canvas* to all display clients requesting objects. The display will pop up an extra window showing this canvas to alert the consumer operator.

- The other way uses the error logging mechanism offered by the CDF offline software. This allows to log errors and specify their severity. Usual destinations for logged errors are output in the terminal or a log file. The monitoring programs contain a module that adds a special destination to the error logger. This destination sends all errors passing a specified severity level via network to a central error receiver. This error receiver updates a web page showing the error messages. Explicitely registered error messages are sent to the main control GUI, *run-control GUI*, to warn the person responsible for data taking, the *DAQ ace*. Run-control can automatically react on certain messages, e.g. hold the run or reset a DAQ board.

  This is, for example, used to reduce the number of failures in the silicon system. The consumer monitor for the silicon system, *SVXMon*, sends a message to run-control, as described before, if a problem with a read-out chip is found, and run-control then resets this chip. This improved greatly the data taking efficiency with the silicon detectors.

Consumer Operation

Since the start of Run II, the online monitoring is an integral part of the CDF data taking. It allows to spot problems quickly and react to them. Some part of this is done automatically. The entire detector performance is efficiently monitored from event data. The framework has met all goals listed in section 2.3.2. The consumer framework and the monitoring code have reached a very high level of stability. Figure 2.16 shows many of the people of the consumer framework group.

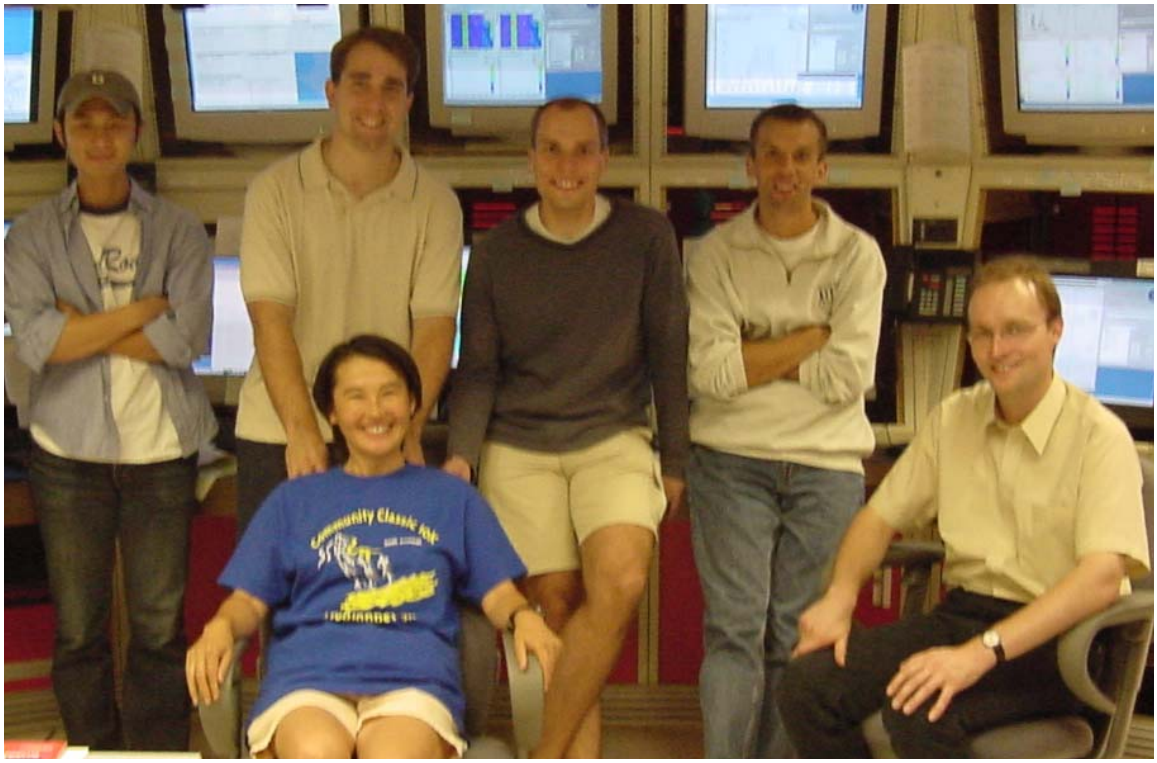


Figure 2.16: Picture of the consumer framework group. In the back from left to right are Koji Ikado, Greg Veramendi, the author, and Hans Wenzel, sitting in the front are Kaori Maeshima and Wolfgang Wagner.

Chapter 3

Event Reconstruction and Data Samples

The events in the different raw data streams are reconstructed on a PC farm [67]. The events with their reconstructed jets, tracks, muons, electrons, and other high level objects are then written into different data sets based on the triggers passed by the event. All events passing the high-E_T electron triggers are in the *BHEL08* data set for example. The *"08"* indicates the production version and corresponds to production with the *4.8.4* version of the CDF software. This data set was used in this analysis. The following sections will briefly describe the main reconstruction steps.

3.1 Track Reconstruction

Using information from the tracking detectors, particle trajectories can be reconstructed. Inside the solenoid, charged particles travel on a helix with its axis parallel to the magnetic field. At CDF the following five parameters are used to describe this helix [68]:

- $\cot\theta$: the cotangent of the polar angle at minimum approach to the origin
- C : the half curvature (same sign as the charge of the particle)
- z_0 : the z position at the point of minimum approach to the origin of the helix
- D : the signed impact parameter; the distance between the helix and the origin at minimum approach
- φ_0 : direction of the track at the point of minimum approach

3.1.1 Tracking in the Central Outer Tracker

In a first step, tracks in the *Central Outer Tracker* (COT) are reconstructed. The drift chamber is the tracking detector with the largest distance from the beam axis. That is, its occupancy is lower and the tracks are more isolated, making the reconstruction easier for this detector than for the silicon detectors.

There are two algorithms to find tracks in the COT. One algorithm is based on the code used in Run I to reconstruct tracks in the CTC [69]. In this approach, segments are reconstructed in the super-layers. These segments are then linked together to reconstruct the trajectory.

The other algorithm [70] uses one segment in the outer super-layers and the expected beam line to construct a reference track. The distances of the hits in the other super-layers from this reference are filled into a histogram. This histogram is used to determine the track parameters. Due to this procedure, the tracks are already beam constrained which improves the momentum resolution. However, the exact position of the beam line is not known when the reconstruction is done and the tracks reconstructed by this algorithm have a bias towards the assumed beam position used in the construction of the reference tracks. In an early production pass, the origin has been used as the beam position, although the beam had a few millimeter offset in reality. Figure 3.1 shows the resulting big difference in the impact parameter for tracks from the unbiased segment linking algorithm and from the biased histogram linking algorithm.

3.1.2 Silicon Tracking

There are two different approaches to reconstruct tracks in the silicon system: *outside-in* and *stand-alone* tracking. The *outside-in* tracking algorithm propagates a track found in the COT into the silicon system and tries to add hits to the track. After a hit has been added, the track parameters are recalculated using this additional information. In the CDF software, there are two implementations of this algorithm. One is based on the Run I code and uses a progressive fitter [71]. The other uses a Kalman fitter, which is the optimal fitter for this task. This fitter and the algorithms based on it have been developed at the Institut für Experimentelle Kernphysik in Karlsruhe [72].

The *stand-alone* tracking algorithm is as well based on this Kalman fitter. The COT does not cover the forward and backward regions. Thus, only the information of the silicon detectors can be used to find tracks up to $|\eta| < 2.0$. This is the task of the *stand-alone* algorithm. To reduce combinatorics, the algorithm uses only hits not used by the two *outside-in* strategies. The position of the beam line is needed for the construction of the track candidates causing a small bias towards the assumed beam position.

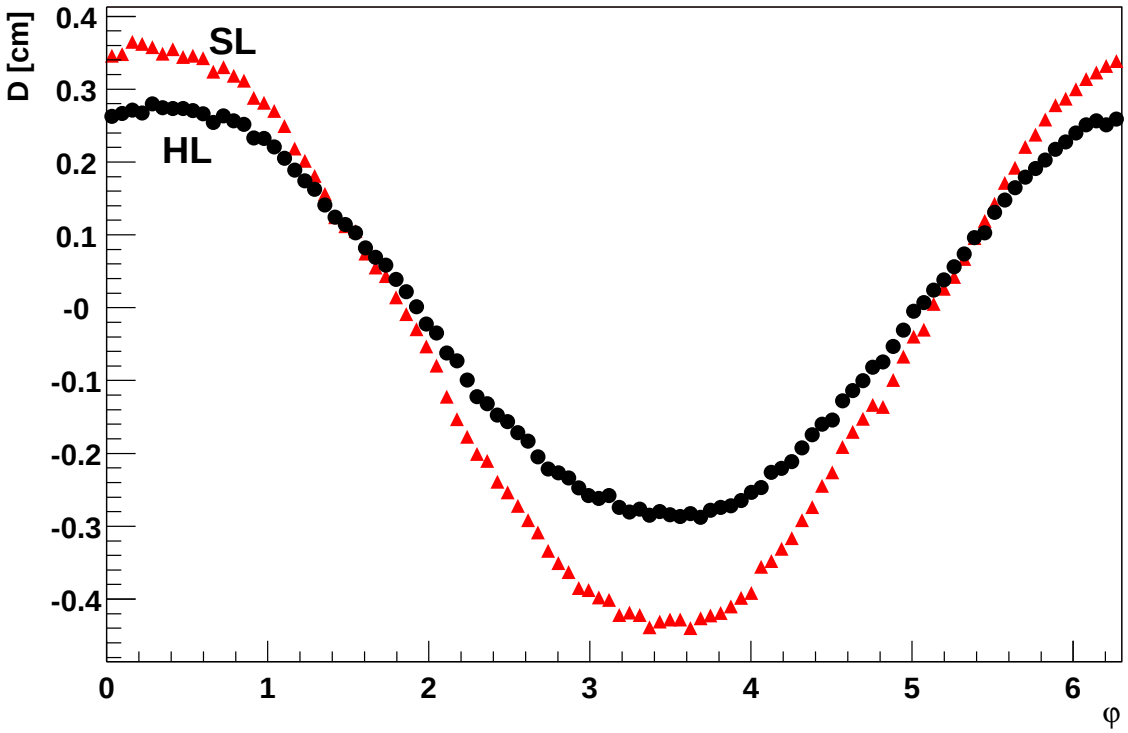


Figure 3.1: The average impact parameter D as a function of φ_0 for all tracks found with the segment linking algorithm (triangles) and the histogram linking algorithm (circles).

3.2 Primary Vertex Reconstruction

Many analyses like life time measurements and analyses needing a b-tag require a good knowledge of the primary vertex position for each event. In most cases in Run I, the position of the beam line was used to estimate the primary vertex position in x and y, if the z coordinate was known. This method proved to be sufficient for most applications in b-physics. Vxprim [73] was used to find the primary vertex with a better precision than the beam width for events with high multiplicity (e.g. $t\bar{t}$). To achieve this goal, Vxprim fitted the primary vertex using reconstructed tracks. In Run I, one method to fit the beam line on a run by run basis [74] used the primary vertices found by Vxprim. In addition, in Run II primary vertex reconstruction will be an useful tool to distinguish different interactions in a single bunch crossing.

The Vxprim algorithm and the underlying vertexing functions [75] have been ported to C++ [76]. An AC++ module that calls the Vxprim algorithm is run in production. The results of this module are used to determine the beam line positions [77]. The code for the module resides in the VertexMods package [78]. The VertexFitter class that houses the vertexing algorithms can be found in the VertexAlg package [79]. The Vxprim module stores found primary vertices in a VertexColl object [80].

3.2.1 The Algorithm

The algorithm is the same as in Run I [73]. The paragraph titled "Finding the Primary Vertex" describes the fitting function `vxgtpr`. This function can be used outside the Vxprim module by using the VertexFitter class directly.

The Track Selection in the Vxprim Module

The Vxprim module reads in CdfTrack [81] objects and applies cuts on the tracks. The Vxprim module allows to cut on the transverse momentum of the tracks and on their impact parameter [68]. At least two stereo and two axial super layers with at least six hits each have to be assigned to a COT track to accept this track. Silicon tracks reconstructed by an outside-in algorithm are required to have at least four $r-\varphi$ hits, for tracks from the stand-alone algorithm four $r-\varphi$ and two stereo hits are required.

Finding the Primary Vertex

The main task of the algorithm is to find the tracks that originate from the primary vertex and remove all other tracks from the vertex fit. Using a wrong track in the fit results in getting a primary vertex position and covariance matrix that is not compatible with the

right vertex position. To remove the tracks that do not originate from the primary vertex, the fitting procedure is iterative. It starts with fitting a vertex with all tracks that passed the track selection cuts. Then it loops over all the tracks and subtracts one track a time from the fit and calculates the χ^2 of this track with respect to the fitted vertex. If the highest χ^2 value for any of the tracks exceeds a specified value (*maxtrackchi2*), this track is removed from the track sample. Then all remaining tracks are used to fit a vertex and this pruning procedure is repeated. If all tracks pass the χ^2 cut, the tracks go through the same procedure again doing a vertex fit with steering of the track parameters this time. This pruning of the track collection stops, if a specified minimum number of tracks is left or all tracks pass the χ^2 cut. A last vertex fit is done with the remaining tracks to find the primary vertex position. In the version of vxprim currently running in production, the second pruning loop using a vertex fit with steering of the track parameters is omitted.

The Vertex Selection in the Vxprim Module

In the Vxprim module, the pruning is allowed to go down to *minn* - 1 tracks. If at least *minn* tracks are used in the final fit, none of the tracks has a χ^2 with respect to the vertex that is greater than *maxtrackchi2*. In this case, a Vertex object with the results of the fit is created and appended to the event record as part of a VertexColl. There is the possibility to rerun this algorithm using all tracks that have been dropped in the pruning to search for additional primary vertices. All tracks with a z_0 parameter in a specified window around the z coordinate of the previously found vertex are removed from the sample beforehand to reduce combinatorics.

3.2.2 Vxprim using Beam Data from Run II

The goal of Vxprim is to find a primary vertex with a resolution below the beam width for events with a high track multiplicity and not to find a primary vertex for every event. Nevertheless, we have investigated how often Vxprim finds a primary vertex. The percentage of events for which at least one vertex has been found in an event during production is shown in table 3.1 for the 4.8.4 production version.

3.3 The Beam Line

The transverse profile of the luminous region inside the CDF can be described by Gaussian functions in x and y. The means of these Gaussians define the beam position. The width of these Gaussians varies along the length of the interaction region due to the focusing of

data file	cdfsoft version	Vxprim using		
		def. tracks	COT tracks	Si tracks
br0234be.002cphys	4.8.4a_08	94.22%	94.69%	79.99%
br0234be.009ephys	4.8.4a_08	94.00%	94.97%	78.61%

Table 3.1: The percentage of events for which at least one vertex has been found. The second column gives the name of the production version that has been used. The third column shows the results for the Vxprim module using the default track collection. The fourth column shows the results for the Vxprim module using COT tracks and the fifth shows the results for Vxprim using silicon tracks.

the beams. This can be expressed by the following equation [82, 83]:

$$\sigma(z) = \sqrt{\epsilon(\beta^* + \frac{(z - z_{\min})^2}{\beta^*})}, \tag{3.1}$$

where ϵ is the transverse emittance, β^* the amplitude function at the interaction point, and $z_{\min}$ the z position of the minimum of the amplitude function. Using the Run IIa design parameters for β^* of 35 cm and ϵ[1] of 1.40×10^{-9} m [45] the minimal transverse width is approximately 22 μm. Thus the knowledge of the beam position in x and y gives a good estimate of the primary vertex position of an event. Unfortunately, the beams are not centered in the detector and not parallel to the detector z-axis. The dependence of the x and y position of the beam on z can be described by a straight line:

$$\bar{x} = x_{slope} z + x_0; \bar{y} = y_{slope} z + y_0 \tag{3.2}$$

The longitudinal luminosity distribution can be approximated as [82, 84]:

$$\frac{dL}{dz} \propto \frac{exp(-(z - z_0)^2/2\sigma_z^2)}{1 + \frac{(z - z_{\min})^2}{\beta^{*2}}}, \tag{3.3}$$

where z_0 is the point at which the centers of the proton and anti-proton bunches collide, the cogging point, $z_{\min}$ is the z position of the minimum of the amplitude function, β^* is the value of the amplitude function at the minimum and σ_z is the longitudinal bunch length. Using the expected values for Run IIa, this equation can be approximated by a Gaussian with a width of 30 cm.

With the CDF, the beam position can be measured using tracks found with the *Central Outer Tracker* (COT) as well as with tracks found with the silicon tracking detectors

[1] using $\epsilon = \frac{\epsilon_{N,95\%}}{6\pi(\beta\gamma)_{rel}}$

(L00, SVX II, and ISL) [85]. These two measurements allow to check the global alignment of the two tracking systems and there are different tables in the database for the different measurements. The beam positions obtained with the COT are stored in the CotBeamPosition tables and the beam positions found using silicon tracks in the SvxBeamPosition tables respectively. The beam positions are calculated and stored on a run by run basis. To distinguish between the beam positions obtained for a run using different reconstruction passes, the database tables contain a so called *history code*. A list with all used codes and their meaning can be found on the web pages of the alignment group [77].

3.3.1 The Algorithms

We apply the same two algorithms that have been used in Run I [74]. The first algorithm is based on the reconstruction of the primary vertex for each event using tracks. This is the default algorithm for fitting the beam line. The second algorithm is based directly on tracks using the correlation of the D and the φ_0 parameter of a track. The results of this algorithm are stored in the database as well to allow to cross-check the default beam line. To distinguish between the beam line results from different algorithms, the database tables contain an *algorithm code*. This integer value is 20 for the vertex algorithm results and 10 for the results from the D-φ_0 fits.

Using the Primary Vertex of the Event

The primary vertex is reconstructed using tracks with the *Vxprim* module, see section 3.2. During production, two clones of the *Vxprim* module are running.

One module produces a vertex collection named *cot_vertices*. In the vertex fit, the track parameters derived from measurements in the Central Outer Tracker (COT tracks) are used. The track requirements for the COT tracks are:

- $p_T > 0.5$ GeV
- $|d_0| < 3.0$ cm
- "standard" COT track
- found by SL tracking algorithm

The other module produces a vertex collection named *svx_vertices*. Only tracks that have been reconstructed using measurements in the silicon detectors are used in the vertex fit. The track requirements for these (silicon) tracks are:

- $p_T > 0.5$ GeV

- $|d_0| < 3.0$ cm
- outside-in tracks:
  - "standard"COT track
  - at least four phi silicon hits
- at least five phi and two stereo silicon hits for stand-alone tracks

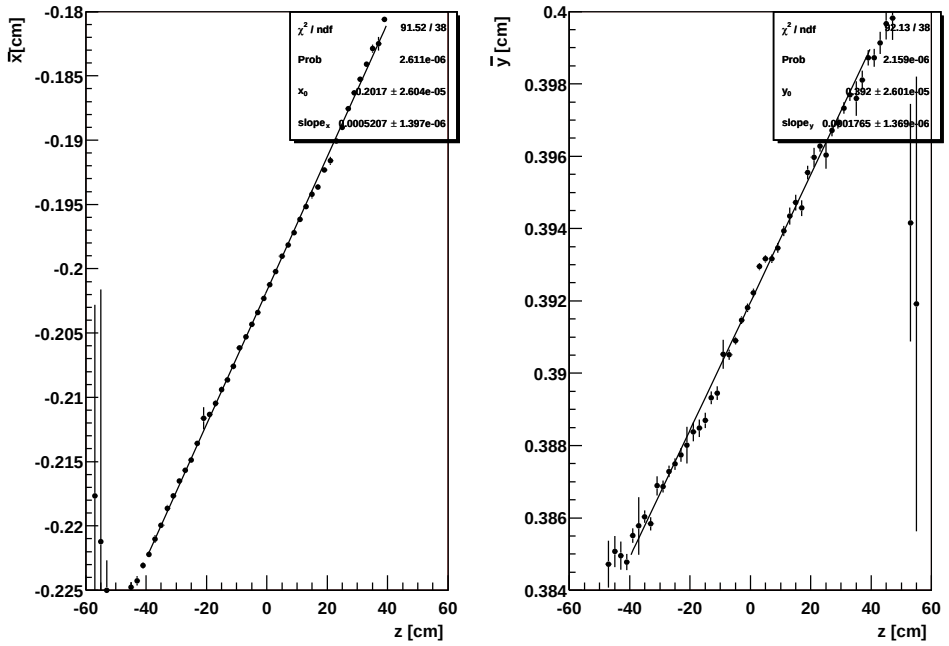


Figure 3.2: Profile plots of the vxprim vertices reconstructed with silicon tracks for run 155130.

From the distributions of the primary vertices one can directly derive the beam position and other beam parameters. Profile plots of the vertex positions in x respectively y over z are used to fit the beam line. Figure 3.2 shows the profile plots for run 155130.

During production, a special module, named *FitBeamModule*, fills the separate profile plots with the primary vertices reconstructed using COT stand-alone tracks and silicon tracks. After production, a script combines the plots for the different streams, makes line fits through the data points of the plots, and fills the data base with the results of these beam line fits.

Using the D-φ_0 Correlation

As can be seen in Figure 3.3, the impact parameter D of a track coming from a primary vertex at (x_V, y_V) can be parameterized to first order using the track parameterization of CDF, see section 3.1:

$$D(x_V, y_V, \varphi_0) = -x_V \sin\varphi_0 + y_V \cos\varphi_0 \tag{3.4}$$

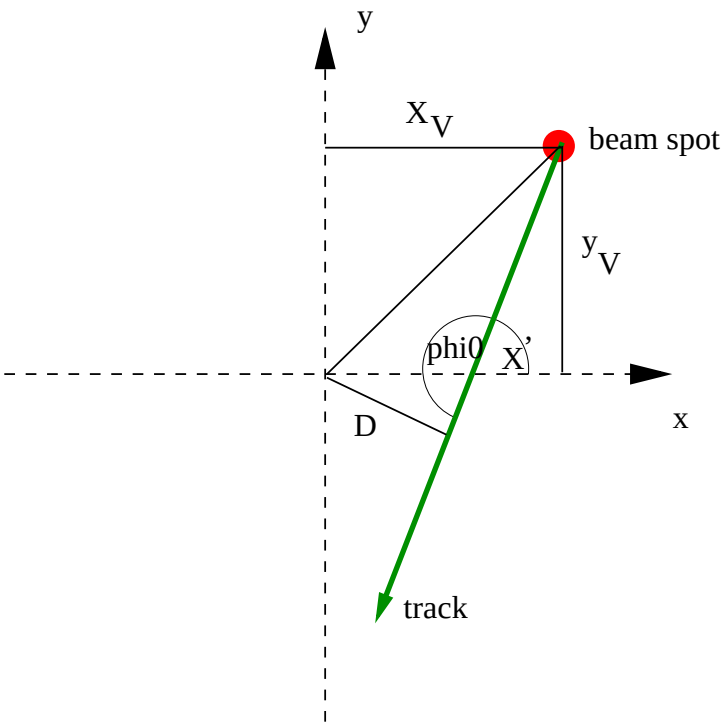


Figure 3.3: Figure to illustrate the correlation between the impact parameter D and the angle φ_0 of a track.

Using the z_0 parameter of the track and the beam line parameters $\vec{x} = (x_0, y_0, x_{\text{slope}}, y_{\text{slope}})$, x_V and y_V can be written as:

$$x_V = x_0 + z_0 x_{\text{slope}} \tag{3.5}$$

$$y_V = y_0 + z_0 y_{\text{slope}} \tag{3.6}$$

Inserting 3.5 and 3.6 in 3.4 gives:

$$D(z_0, \varphi_0) = -x_0 \sin\varphi_0 + y_0 \cos\varphi_0 - z_0 \cdot x_{\text{slope}} \sin\varphi_0 + z_0 \cdot y_{\text{slope}} \cos\varphi_0 \tag{3.7}$$

With $\vec{g} = (-\sin\varphi_0, \cos\varphi_0, -z_0 \sin\varphi_0, z_0 \cos\varphi_0)$ this can be written as:

$$D(z_0, \varphi_0) = \vec{x} \cdot \vec{g} \tag{3.8}$$

For n tracks, the χ^2 of the difference between the measured and the calculated impact parameter using the beam position is:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{D_i - \vec{x} \cdot \vec{g}}{\sigma_i} \right)^2 \text{ , with } \sigma_i^2 = \sigma_{D_i}^2 + 2\sigma_{\text{beam}}^2 \tag{3.9}$$

By minimizing this χ^2, one can obtain the beam position as:

$$\vec{x} = V\vec{b} \text{ , with } V^{-1} = \sum_{i=1}^{N} \frac{\vec{g}_i \vec{g}_i^T}{\sigma_i^2} \text{ and } \vec{b} = \sum_{i=1}^{N} \frac{D_i \vec{g}_i}{\sigma_i^2} \tag{3.10}$$

All COT tracks that pass the following cuts are used in a first fit:

- $p_T > 1.0$ GeV
- $|d_0| < 3.0$ cm
- "standard" COT track
- found by SL tracking algorithm

The track requirements for the silicon tracks are:

- $p_T > 1.0$ GeV
- $|d_0| < 3.0$ cm
- outside-in tracks:
  - "standard" COT track
  - at least four phi silicon hits
- at least five phi and two stereo silicon hits for stand-alone tracks

After the fit, any track that has a large impact parameter with respect to the calculated beam line result is cut away and the beam position is fitted again using the remaining tracks. This procedure is iterated. Between each iteration, the impact parameter cut is tightened, until less than 60% of the starting tracks survive. The values of $\vec{b}$ and V are stored in a histogram. As for the vertices, the histograms from all the data files of a run are combined and the fit result and its covariance matrix are calculated using formula 3.10. The result is put into the database.

3.3.2 Stability of the Beam Position

The *FitBeamModule* has recently been modified to create new versions of the histograms used in the beam line fits for every 500 000 "level 2" triggers. This allows to monitor the beam position during a single run. For a test, we have looked at stream G data that had been produced with offline version 4.8.4. The silicon tracks have been refitted using the *TrackRefitModule* from offline version 4.9.1. We have investigated 51 long runs. Figure 3.4 shows the maximum difference between the beam position parameters within

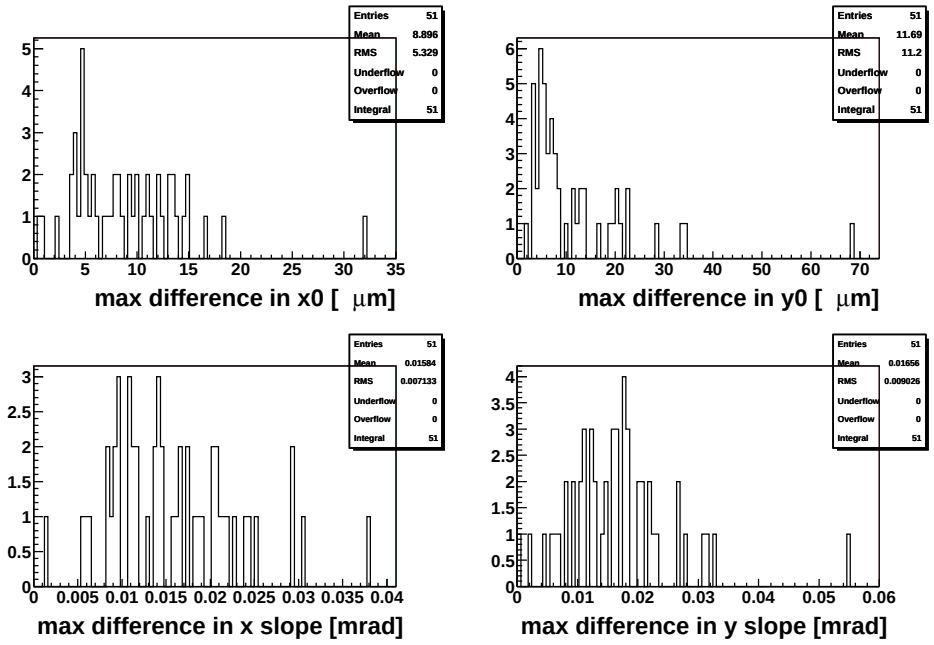


Figure 3.4: Maximum difference in the parameters for the beam position per division in the analyzed and rather long runs.

one run for the 51 investigated runs. The movement of the beam is in the order of 5 μm for most of the analyzed runs. Nevertheless, at least for two runs, there are differences in the order of a few times the expected width of the beam. The Figures 3.5 and 3.6 show the beam positions obtained for different divisions of the two runs with the largest movements: run 148824 and run 150395. To take this movement into account, the beam position is fitted separately for each 500 000 "level 2" trigger division and stored in the data base. This started with the history code 19 entries.

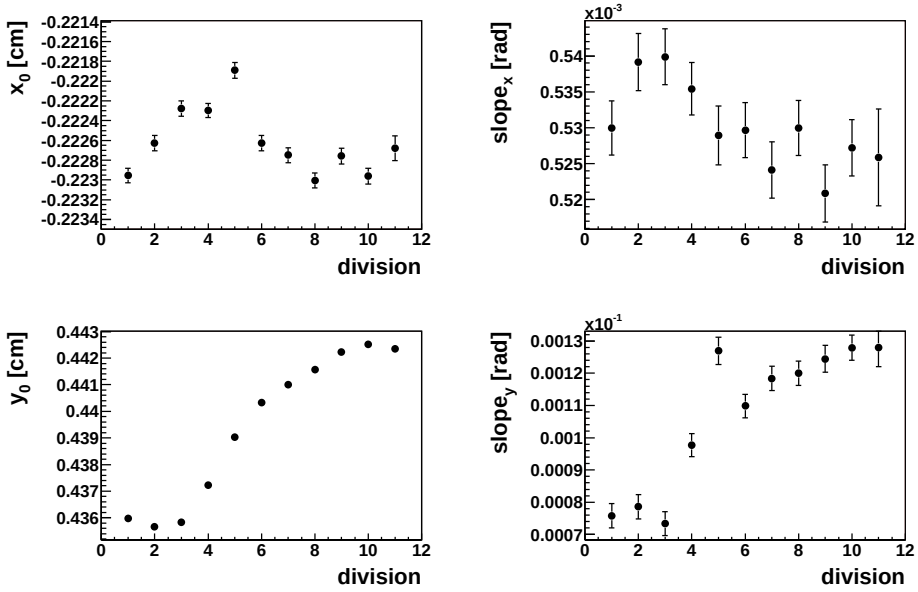


Figure 3.5: Beam position parameters for different divisions of run 148824.

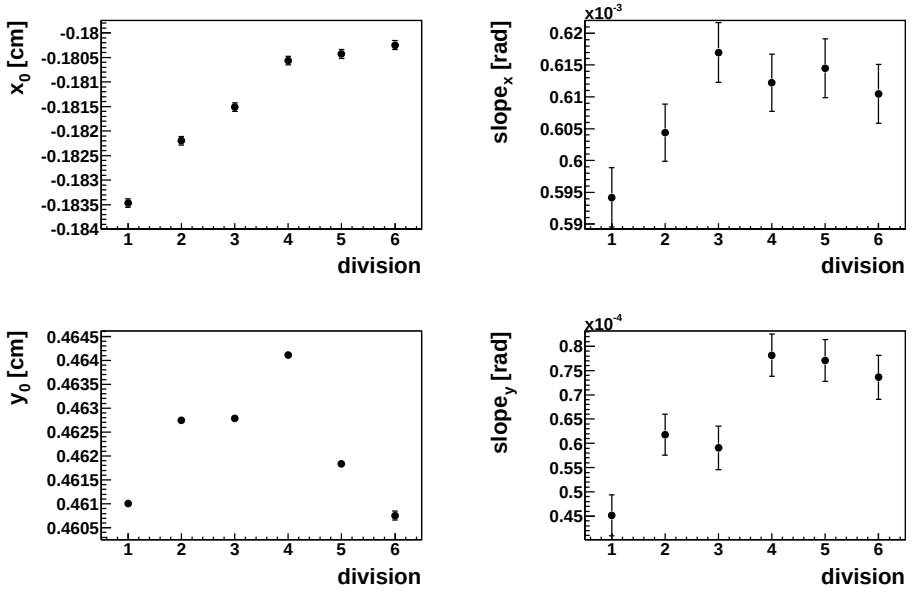


Figure 3.6: Beam position parameters for different divisions of run 150395.

3.3.3 Size of the Beam Spot

The database table for the *SVX* beam position contains fields for the beam spot size. To get the beam width, the x and y coordinates of the primary vertices are plotted for six certain z ranges. The length of these ranges is 1 cm to limit the effect of the slope of the beam line. Only vertices with an invariant mass above 50 GeV are used for the plots. Vxprim should reconstruct these vertices very well. In a first approach to obtain the mean beam width, Gaussians are fitted to the distributions. The weighted mean of the six values for the Gaussian widths is calculated. After subtracting the mean vertex fit resolution in quadrature, the obtained value is filled into the database as the mean beam width. This procedure is comparable to what was done in [86] and used in the database entries with history code 17 and 19.

This procedure can be improved by correcting for the beam movements. The histograms containing the x and y vertex positions for six different z ranges are created for every division of 500 000 "level 2" trigger as it is done for the profile plots. For each division, we correct for the fitted beam position and sum up the histograms for all divisions. Gaussians are then fitted to the resulting histograms. We correct these widths for the mean resolution of the vertex fits. After the fit, we know the width of the beam for six different values of z. Figure 3.7 shows the result for run 144674. As the width of the beam should depend on z according to equation 3.1, we have tried to fit this function to the distribution. More data points are needed to really constrain the problem. This has been done in [87]. The minimal beam width obtained in this study is $\sigma_x = (25.7 \pm 1.9)\ \mu$m and $\sigma_y = (25.8 \pm 1.7)\ \mu$m.

3.4 Jet Reconstruction

The hadronization of a final state quark creates a jet of hadrons. The energy of the hadrons is measured in the calorimeters. The momentum of the initial quark can be reconstructed by combining the energy measurements in the calorimeter towers that belong to the jet.

3.4.1 Algorithms

Three different algorithms to reconstruct jets are implemented in the CDF software [88].

JetClu

This algorithm has been the standard algorithm in Run I. Thus, its systematics are very well understood. Firstly, this algorithm selects a seed tower. Then it draws a cone around

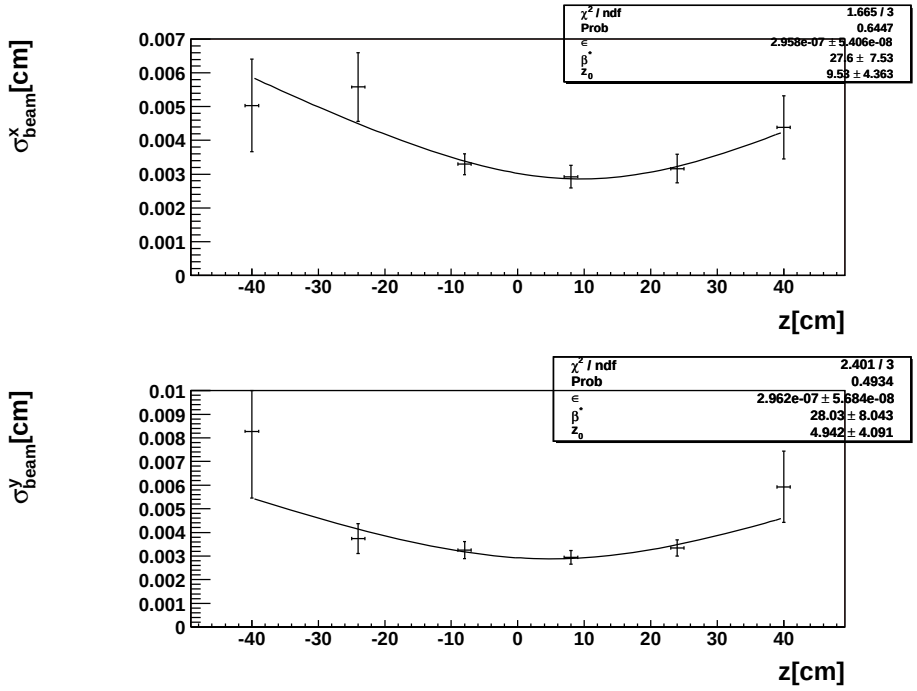


Figure 3.7: Beam width over z for run 144674.

this tower with a fixed radius in the η-φ plane. All calorimeter towers inside this cone are combined to form the jet. The axis of this jet is used as the new direction of the cone axis in the next iteration of this algorithm. If the jet axis stays stable, the reconstruction of this jet is finished. Seed towers are all calorimeter towers with a measured energy above a certain threshold.

Although this algorithm works very well in the dense environment of hadron-hadron collisions, it has two problems when it is applied to partons to make theoretical predictions:

- A single parton with energy above threshold will serve as a seed. But, if the momentum of this parton is distributed to two partons with each having half the energy, both might fail the threshold cut. This is called the collinear problem.

- If there are two high energetic partons that have a distance in the η-φ plane that exceeds the cone size, two jets will be formed by the algorithm. A gluon emitted by one of the partons might move the jet axis in a way that now both partons and the gluon form just one jet. This is called the infrared problem.

Midpoint Cone Algorithm

The midpoint cone algorithm is based on the simple cone algorithm described in the previous section. It was formulated, to get rid of the problems of its predecessor by adding the midpoints between two towers as further seeds.

K_T-Clustering Algorithm

The K_T-clustering algorithm calculates the distances between all energy depositions based on the formula $d_{ij} = \min(p_{T,i}^2, p_{T,j}^2)\frac{\Delta\mathcal{R}_{ij}^2}{D^2}$, where $\mathcal{R}_{ij}$ is the distance between two momentum vectors in the $\eta - \varphi$ plane and D is a parameter of the jet algorithm. The two momentum vectors with the smallest distance are combined and this new vector is used instead of the two original ones. All distances are then recalculated and the procedure is iterated as long as there are distances below a certain threshold value.

This algorithm works very well on the partonic level and for experiments located at an electron-positron or electron-proton collider with a rather clean environment. Right now, the results of this algorithm for the experiments at the Tevatron are not well understood and disagree with the other two algorithms for low energetic jets.

3.4.2 Jet Energy Corrections

Not all the particles created during the fragmentation of a hadron are measured in the calorimeter, nor do the jet algorithms combine all towers correctly. Furthermore, the calorimeter response is different in pseudorapidity and changes with time. To compensate for these effects, the energy of the reconstructed jet needs to be corrected to match the energy of the initial parton. Only then, one gets the right momenta and masses for objects decaying into quarks.

A group within CDF has derived a first set of jet corrections for Run II [89]. Right now seven different corrections are applied to a jet. The following corrections are performed:

- η-dependent relative corrections
- time-dependent corrections
- raw energy scale corrections for Run IIa to agree with Run Ib
- multiple interactions corrections
- absolute energy corrections (from calorimeter energy to particle energy)
- underlying event corrections

- out-of-cone corrections

The single corrections are performed sequentially. Applying "Level 4" corrections means that the first four corrections are applied. Figure 3.8 illustrates the impact of the full jet corrections on the reconstructed Z boson mass for Monte Carlo $Z \to qq$ events.

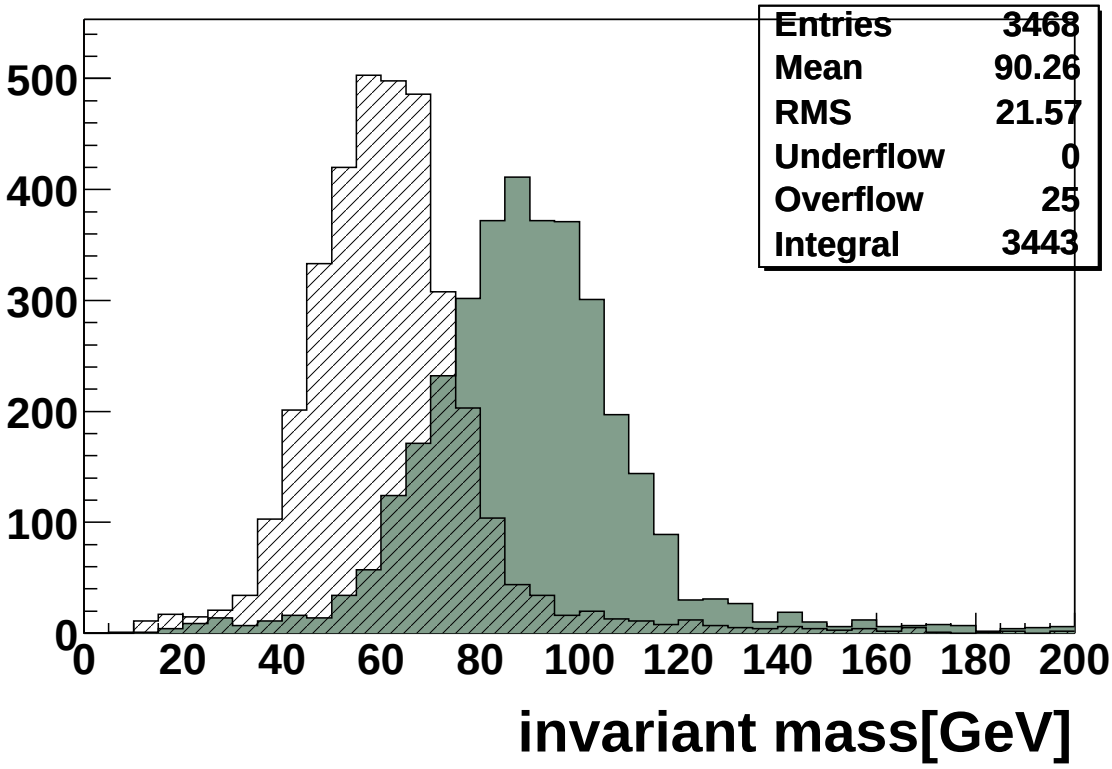


Figure 3.8: Reconstructed Z boson mass in Monte Carlo $Z \to qq$ events. The hatched histogram shows the result for uncorrected jets. The grey histogram shows the invariant mass of corrected jets.

3.5 The Identification of Bottom Jets

In many physics analyses, it is crucial to know the flavor of a jet to extract a signal. It is possible to discriminate jets originated by a bottom quark from jets originated by lighter quarks or gluons. Due to the relatively large mass of the bottom quark, the bottom hadron carries most of the momentum of the original quark. Thus, the hadron is boosted and, due to its lifetime of approximately 1.5 ps, it travels a sizeable distance before its decay.

3.5.1 The Jet Probability

One way to tag a jet is to look for tracks near the jet axis that do not originate from the primary vertex. For every track, one can calculate the likelihood that the particle belonging to this track has been produced at the primary vertex. Using all tracks associated with this jet, a probability is calculated following formula 3.11 whether this jet has been formed at the primary vertex.

$$P_N = \Pi \cdot \sum_{j=0}^{N-1} (-\log \Pi)^j / j!, \text{ with } \Pi = \prod_{i=1}^{N} p_i \tag{3.11}$$

3.5.2 The *SecVtx* Algorithm

The second algorithm in use at the CDF is called *SecVtx*. This algorithm searches for a secondary vertex directly. In a first step, tracks that are well reconstructed and satisfy a specific impact parameter significance are chosen. Then it attempts to reconstruct a secondary vertex with at least three tracks using this selected tracks. If this fails, tighter cuts are applied on the tracks and the attempt is made to find a two-track secondary vertex. If a secondary vertex has been found, the two-dimensional displacement vector from this vertex with respect to the primary vertex is calculated. A jet is tagged, if the length of this vector exceeds a specific significance. Details about the algorithm and the specific cuts used can be found in [90]. The efficiency to tag a fiducial b-jet with transverse energy above 10 GeV is $(19.6 \pm 1.1)\%$ averaged over E_T [91]. First results on the dependence of the tagging efficiency on the transverse jet energy are shown in figure 3.9 These numbers were obtained from data using a double-tag method.

3.6 Data Samples

In this section, I will briefly characterize the data samples used in the analysis and describe the reprocessing. To take advantage of newer calibrations and correct for bugs in the reconstruction code used originally, I have remade most of the high level objects. I have used CDF software version *4.9.1hpt1* for this task. The calorimeter reconstruction was redone with the latest calorimeter constants. The jets, electromagnetic objects, muons, and calculated missing transverse energy were dropped on input and remade. The tracks were refitted to take correctly the alignment into account. Consequently, new primary vertices were reconstructed and the latest version of the b-tagging algorithm was used.

To facilitate the further data analysis, the events were written out in an extended ntuple. I have used *Stntuple* with version *239* for this task. In fact, Stntuple is a *micro-DST* and offers a new framework to easily access the data and analyze it using ROOT.

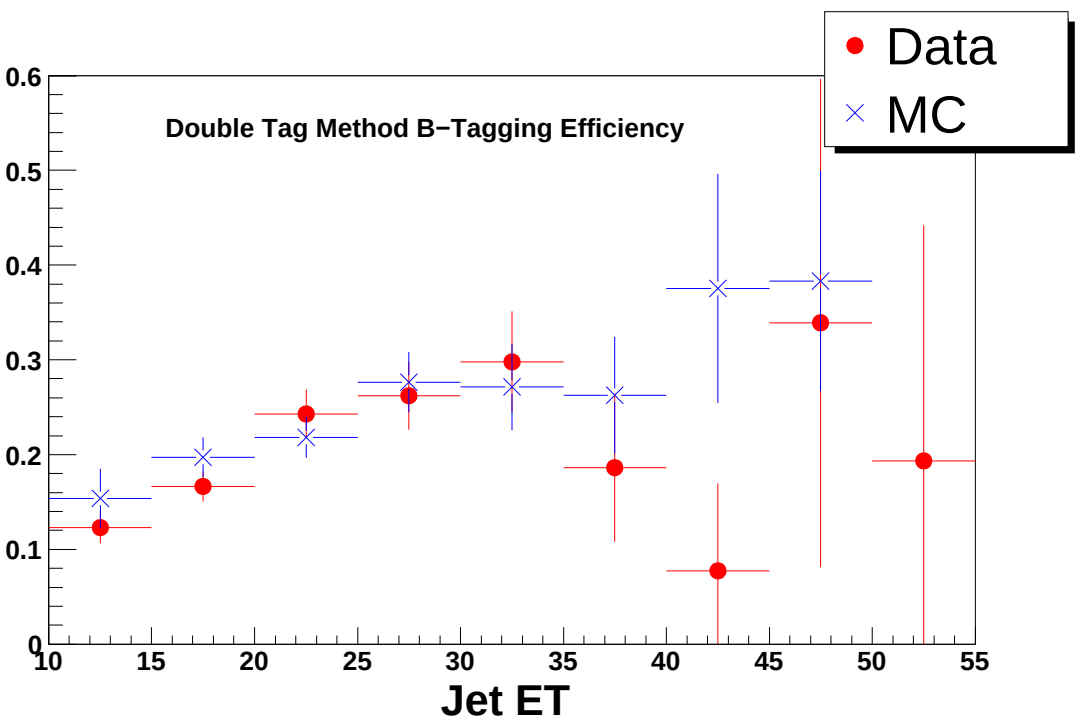


Figure 3.9: Efficiency to tag a b-jet as function of the transverse jet energy[91].

Developing code is much faster within this framework than in the CDF offline framework. Furthermore, the events can be read in much faster. All high level objects stored in the *Stntuple* contain the same information as their counterparts in the standard CDF event data model (EDM).

3.6.1 WTOP0E Sample

To determine the acceptance of the CDF detector and of my selection cuts, an official Monte Carlo sample for the process $W \to e\nu$ was used. Its name was *WTOP0E*. The events were generated with Pythia version 6.203 [92]. The W bosons were forced to decay into electron-neutrino pairs. In this analysis, I used 160.000 events from this sample. I have determined the acceptance and efficiency of the preselection on this Monte Carlo sample in section 5.3.

3.6.2 Collider Data Sample: *BTOP0G*

The data sample used in this analysis is called *BTOP0G*. It is based on the high-energetic electron sample *BHEL08*. However, a few cuts on the electron candidates were applied to clean up the sample. These cuts are discussed in detail in chapter 4.

I used the good runs list of the CDF Lepton+Jets Working Group as of 04/01/03. The *DHInputModule* writes a log file that lists all the run sections in the processed data. This information can be used to obtain the integrated luminosity of the processed data. However, when only certain (good) runs are selected by the input module, still all run sections are written into the log file. I have written a little script that removes these superfluous run sections from the log file and prints out the integrated luminosity of all selected runs. Table 3.2 lists the integrated luminosities for all processed file sets. The analyzed data corresponds to an integrated luminosity of (55.5 ± 3.3) pb^{-1} with an error on the luminosity of 6% [93]. I used all good runs up to January 2003. The data was copied to two computer clusters at Karlsruhe to be reprocessed and to fill the ntuple. The events written to the Stntuple files were required to have a corrected missing transverse energy above 15 GeV.

file set	location	luminosity [pb^{-1}]	luminosity (selected runs) [pb^{-1}]
GI0741.0	EKP	2.347	1.119
GI0741.1	EKP	2.822	1.154
GI0741.2	EKP	2.406	0.389
GI0741.3	EKP	2.274	0.636
GI0741.4	EKP	2.078	1.082
GI0741.5	EKP	1.964	1.707
GI0744.0	EKP	2.697	1.307
GI0744.1	EKP	2.023	0.000
GI0744.2	EKP	2.029	0.525
GI0744.3	EKP	2.393	0.599
GI0744.4	EKP	2.432	1.297
GI0744.5	EKP	0.915	0.532
GI0747.0	EKP	2.227	0.000
GI0747.1	EKP	2.325	0.000
GI0747.2	EKP	2.249	0.693
GI0747.3	EKP	2.314	1.107
GI0747.4	EKP	4.097	0.000
GI0747.5	EKP	2.672	0.817
GI0749.0	EKP	2.041	1.478
GI0749.1	EKP	2.027	1.838
GI0749.2	EKP	2.177	1.263
GI0749.3	EKP	1.989	1.836
GI0749.4	EKP	1.939	1.448
GI0749.5	EKP	1.950	0.391
GI0750.0	EKP	1.915	1.471
GI0750.2	EKP	1.254	0.808
GI0776.0	EKP	1.963	1.883
GI0776.1	EKP	1.942	1.942
GI0776.2	EKP	1.899	1.790
GI0776.3	EKP	2.087	1.872
GI0776.4	EKP	1.369	0.962

Table 3.2: Luminosity of the used file sets located at the EKP. The first column is the file set name. The third column is the total luminosity of all runs in the file set. The fourth is the luminosity of the good runs in the file set.

file set	location	luminosity[pb^{-1}]	luminosity (selected runs)[pb^{-1}]
last page	EKP	66.816	31.946
GI0943.0	FZK	2.403	0.992
GI0943.1	FZK	2.712	0.903
GI0943.2	FZK	1.920	1.128
GI0945.0	FZK	2.585	0.978
GI0945.1	FZK	2.593	1.056
GI0945.2	FZK	2.366	0.355
GI0949.0	FZK	2.855	0.763
GI0949.1	FZK	1.245	1.245
GI1239.0	FZK	2.430	1.471
GI1239.1	FZK	2.073	1.524
GI1239.2	FZK	2.092	1.783
GI1239.3	FZK	2.085	2.085
GI1239.4	FZK	2.067	2.001
GI1241.0	FZK	2.148	2.089
GI1241.1	FZK	2.112	1.729
GI1241.2	FZK	2.120	2.017
GI1275.0	FZK	2.156	1.193
GI1275.1	FZK	0.226	0.226
all at FZK	FZK	38.188	23.538
all	-	105.0	55.5

Table 3.3: Luminosity of the used file sets located at the FZK. The first column is the file set name. The third column is the total luminosity of all runs in the file set. The fourth is the luminosity of the good runs in the file set.

Chapter 4

Electron Identification

I limited myself to analyze data with an electron in the central region of the detector. The electrons are highly energetic, that is they have a transverse momentum of at least 20 GeV. The major source of energy loss for these electrons is photon emission in the interaction with the nuclei in the traversed medium. The energy loss by this bremsstrahlung for highly energetic electrons can be described by:

$$-\frac{dE}{dx} = 4\alpha N_A \frac{Z}{A} r_e^2 E \ln \frac{183}{Z^{1/3}}, \tag{4.1}$$

where Z and A are the atomic number and atomic weight of the medium, and r_E is the classical electron radius. This equation can be simplified by defining the radiation length X_0 for a given medium to:

$$-\frac{dE}{dx} = \frac{E}{X_0} \tag{4.2}$$

As can be seen in equation 4.2, the energy loss through bremsstrahlung is proportional to the energy. As the loss through ionization is proportional to the logarithm of the energy, the bremsstrahlung clearly dominates for electrons with an energy of tens of GeV .

Electrons from collisions can be identified with the CDF by a track in the central drift chamber (COT). They deposit their energy in the central calorimeters. The electrons emit photons in the lead absorbers of the central electromagnetic calorimeter (CEM). These photons on the other hand produce electron-positron pairs again leading to an electromagnetic cascade. The shape and the position of this electromagnetic shower is measured by the central shower maximum detector (CES). The COT track points to the shower position and the track momentum matches the energy measured in the calorimeter. Another criterion to identify electrons is the shape of the shower. A hadronic shower is much broader and longer. Electrons deposit most of their energy in the electromagnetic section of the calorimeter and only a small fraction in the hadronic one. Thus, a comparison of the energies measured in the two sections allows to distinguish electrons from

pions. Furthermore, a hadronic shower continues into the adjacent calorimeter towers, while electromagnetic showers are much narrower.

A main source for misidentified electrons is the coincidence of a π^0 and a $\pi^\pm$ in the same tower. The π^0 causes an electromagnetic shower in the calorimeter, while the $\pi^\pm$ leads to a reconstructed track that matches the electromagnetic cluster. These mesons are produced through strong interactions, thus having a much higher cross-section at an hadron collider than processes that lead to isolated electrons. However, the pions are seldomly isolated, but are part of a jet of hadrons.

4.1 Selection of a Signal and Background Sample

To analyze the efficiency and purity of the different electron cuts, a sample was created that mainly consists of real electrons and one that mainly consists of fake electrons. The selection criteria for these samples were similar to the ones used in [94]. The criteria were applied on the *BTOP0G* data set. This dataset is described in subsection 3.6.2. The preselection cuts applied on the electron candidates for this data set are discussed in section 4.4.

4.1.1 The Signal Sample

There are two main processes at a hadron collider to get events containing highly energetic electrons: W boson production and Z boson production. As the cross-section for $Z \rightarrow e^+e^-$ is ten times lower than the cross-section for $W \rightarrow e\nu$, it was chosen to select W boson events to get a large electron sample. In a first step, events in the central high-E_T electron sample were selected that have a missing transverse energy above 20 GeV. Exactly one electron candidate with a transverse energy above 15 GeV was required in the event. Events with any further jet with a tranverse energy above 5 GeV were cut away. The jet veto reduces the number of QCD events in this sample, as hadrons mimicking an electron have to be balanced by another jet. Furthermore, the missing transverse energy is in this case mainly measured by the CEM, which has a much better resolution than the hadronic calorimeter, compare table 2.2. Still, the sample will not be 100% pure. A look at the jet multiplicities obtained for W boson events and QCD background in appendix B might suggest that the impurity is in the order of 1%. Figure 4.1 displays one event in the signal sample.

4.1.2 The Background Sample

To get a sample of fake electrons, events were selected that contain two jets in the central region. The cuts on the jets were: $E_T > 15$ GeV and $|\eta| < 1.1$. The cosine of the angle

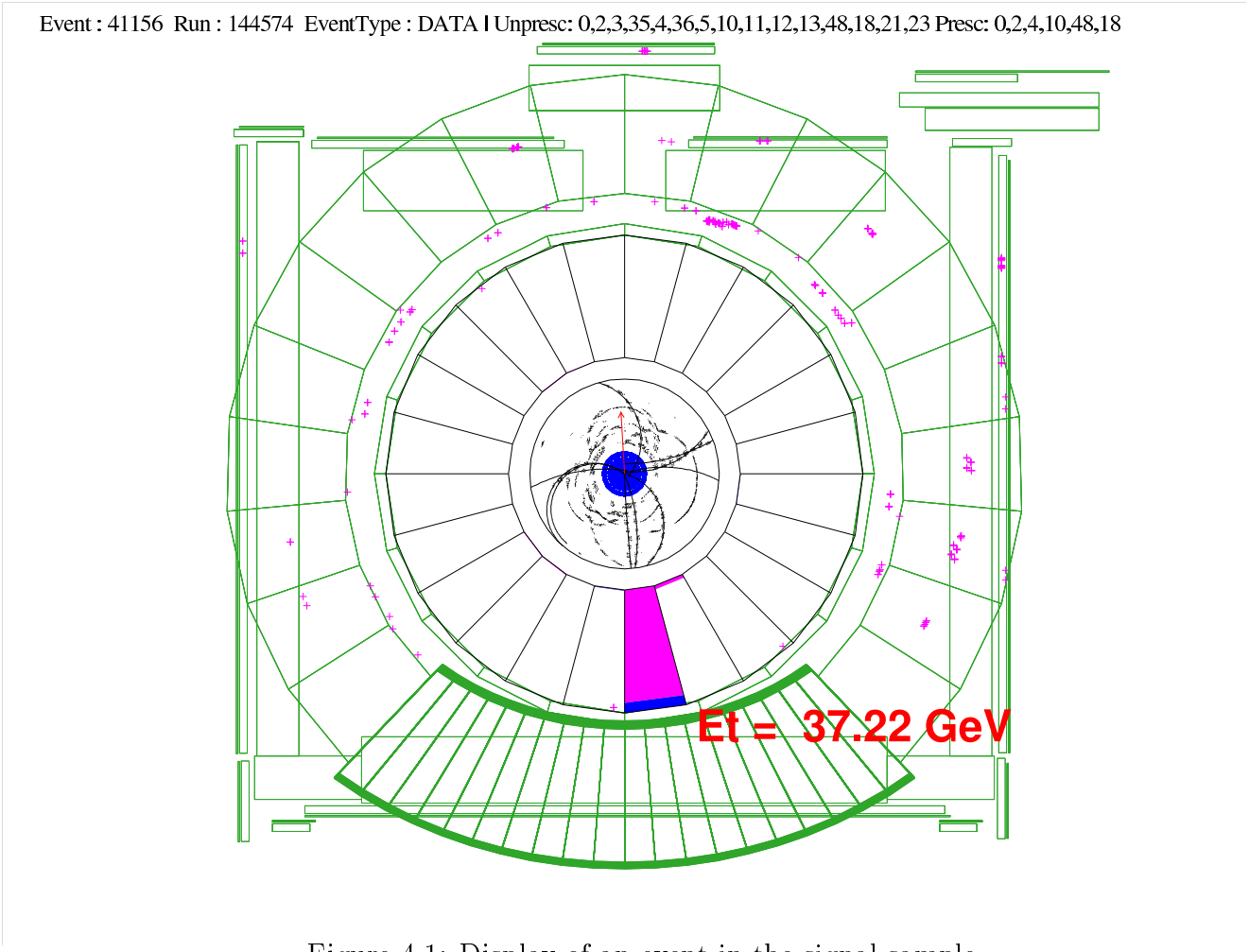


Figure 4.1: Display of an event in the signal sample.

of the jets in the transverse plane had to be above 0.95. In other words, the jets had to be back-to-back. Events with more than two jets have been vetoed. Moreover, the missing transverse energy was required to be below 5.0 GeV. The main source for such events is quark-antiquark or gluon pair production via the strong interaction. By applying this selection cuts on the central high-E_T electron sample, events were picked in that one of the two jets fakes an electron. The only relevant process to get two real electrons that would pass the selection cuts is Z boson production with the Z boson decaying into two electrons. To exclude this process, events were cut away that contain two electron candidates. This sample is called the background sample for the electron identification. Figure 4.2 displays one event in the background sample.

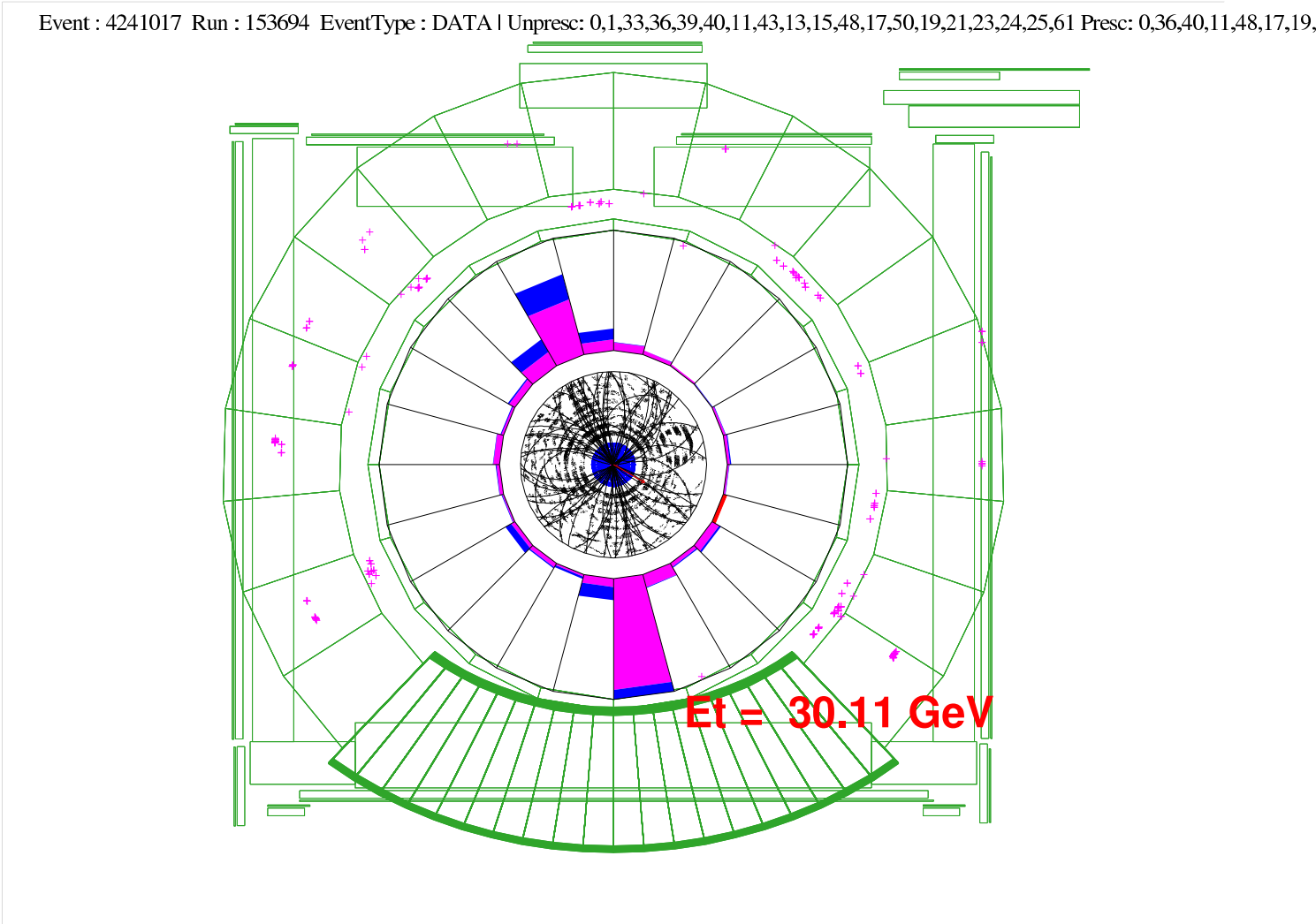


Figure 4.2: Display of an event in the background sample.

4.2 Variables to Identify Electrons

I use nearly the same variables for the electron identification as used in [95]. The following variables are used to identify electrons:

- E:
  The energy deposited by the electron in the electromagnetic and hadronic central calorimeters. An electron cluster is made from a seed electromagnetic tower and at most one more shoulder tower, passing some well defined requirements. Hence, a cluster could have two towers in pseudorapidity ($\Delta\eta \approx 0.2$) and one more tower in azimuth ($\Delta\varphi \approx 0.13$ rad) at maximum.

- E_T:
  The transverse energy deposited by the electron in the electromagnetic and hadronic central calorimeters. It is calculated by multiplying the electromagnetic cluster energy by $\sin\theta$, where θ is the polar angle of the best matching COT track.

- E_{had}/E_{em}:
  The ratio of the hadronic (CHA + WHA) calorimeter energy to the electromagnetic (CEM) calorimeter energy for the electron cluster.

- Isolation:
  The ratio of the energy inside a cone of radius $R = \sqrt{\Delta\eta^2 + \Delta\varphi^2} = 0.4$ around the electron cluster to the energy of the electron cluster. The energy of the electron cluster is subtracted from the numerator and the energy in the cone is corrected for calorimeter leakage (see [96, 97]).

- L_{shr}:
  The lateral shower profile for the electron. This variable compares the energy in the CEM towers adjacent to the seed towers with data from test beam electrons [98].

- χ^2_{wire}:
  The χ^2 comparison of the CES shower profile in the $r-\varphi$ view with the same profile extracted from test beam electrons.

- χ^2_{strip}:
  The χ^2 comparison of the CES shower profile in the $r-z$ view with the same profile extracted from test beam electrons.

- P_T:
  The transverse momentum of the track that is associated with the electron. Only tracks that are reconstructed in the COT are used. These tracks are constrained to originate from the beam line to improve the momentum resolution.

- $Q \cdot \Delta x$:
  The distance in the $r - \varphi$ plane between the extrapolated COT track and the best matching CES cluster times the charge of the track.

- Δz:
  The distance in the $r - z$ plane between the extrapolated COT track and the best matching CES cluster.

- E/P:
  The ratio of the electromagnetic transverse cluster energy to the transverse track momentum. The track is beam constrained.

4.3 Corrections

4.3.1 CEM Energy Corrections

Various corrections to the CEM response have already been applied in the offline code. To check for further corrections, we calculated the average E/P for electrons from $W \to e\nu$ data using E/P values between 0.9 and 1.1. After the offline corrections of the CEM energy, there was a dependence of the average E/P on the local x-position of the cluster. A correction factor has been derived in [99] to flatten the distribution. Figure 4.3 shows the effect of this correction.

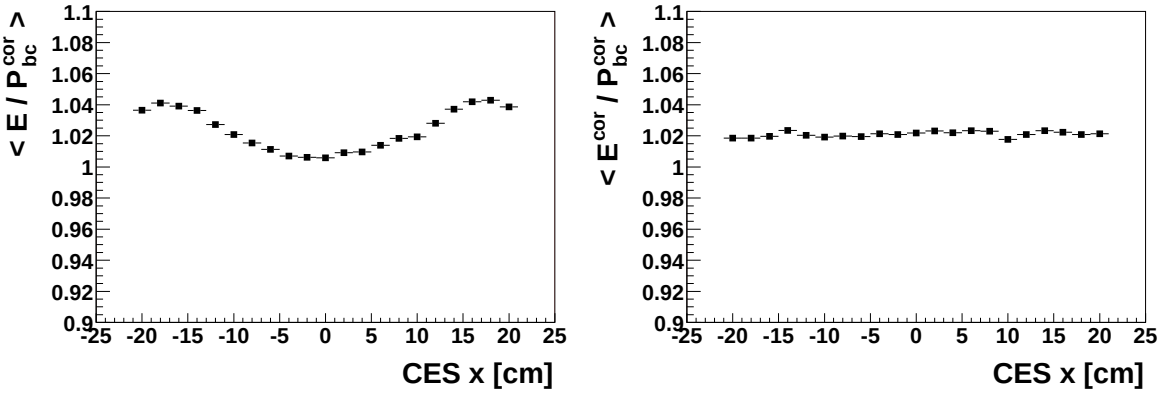


Figure 4.3: Average E/P as a function of CES-x before (left) and after (right) the CEM energy correction. The average E/P is calculated in the range between 0.9 and 1.1. Electrons from $W \to e\nu$ data are used.

4.3.2 Curvature Correction for Beam Constrained COT Tracks

Figure 4.4 shows the E/P distribution for electrons and positrons. One can easily identify

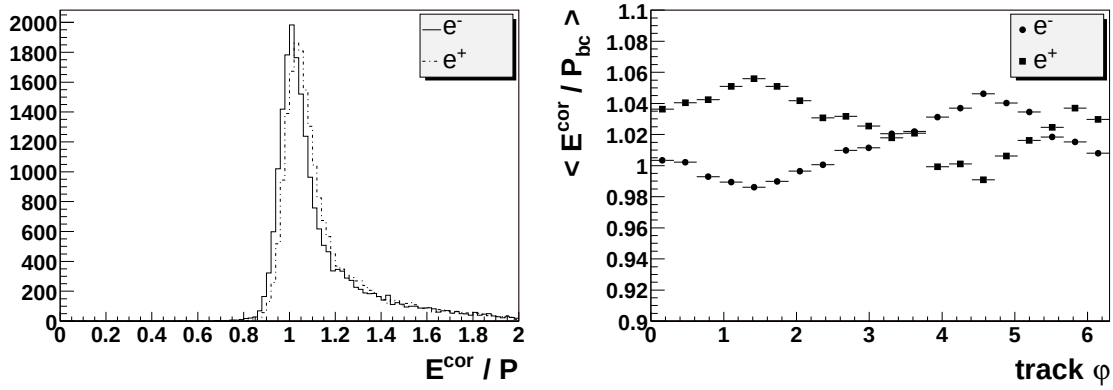


Figure 4.4: The E/P distributions for electrons and positrons (left) and the average E/P as a function of the track φ (right). The average E/P is calculated in the range between 0.9 and 1.1. Electrons from $W \to e\nu$ data are used.

a dependence on the charge and the angle φ_0 of the track. There is a systematic bias in momentum measurements in data. This bias was removed by recalculating the signed curvature of the beam constrained track as described in [99]. The effect of this correction is shown in Figure 4.5

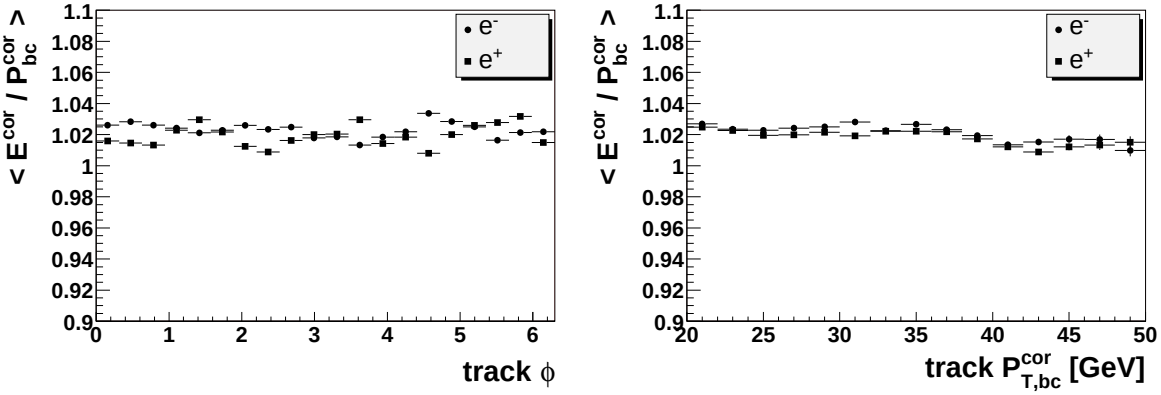


Figure 4.5: The average E/P as a function of the track φ_0 (left) and as a function of the track transverse momentum (right) after the curvature recalculation. The average E/P is calculated in the range between 0.9 and 1.1. Electrons from $W \to e\nu$ data are used.

4.4 Preselection Cuts

In this section, the cuts applied to the electron candidates beforehand are described.

4.4.1 Cuts for the *BTOP0G* Data Set

Events in the *BTOP0G* data set had to have an electron candidate passing certain cuts. There were two different sets of cuts for highly and very highly energetic electrons. Events passing one of these sets of cuts were written to the data set. Table 4.1 lists the cuts of the different sets. These cuts were applied on the uncorrected variables. The electron

Variable	highly energetic electron cuts	very highly energetic electron cuts
E_T	> 18 GeV	> 70 GeV
E_{had}/E_{em}	< 0.125	-
P_T	> 9.0 GeV	> 15 GeV
E/P	< 4.0	-
L_{shr}	< 0.3	-
$\|\Delta x\|$	< 0.3 cm	< 0.3 cm
$\|\Delta z\|$	< 0.5 cm	< 0.5 cm

Table 4.1: Cuts on electron candidates in the *BTOP0G* data set.

track was not beam constrained.

4.4.2 Further Preselection Cuts

There were some further cuts which the electron candidates had to pass in this analysis. Only central electrons were used. Their transverse energy had to be above 20 GeV, the beam constrained transverse momentum of the track had to exceed 10 GeV. The z_0 parameter of the track had to be within 60 cm. The electrons had to be found in a region with good coverage by the CEM, i.e. they had to be fiducial. Furthermore, the cuts for highly energetic electrons were applied on all candidates to get uniform background distributions in the W boson analysis presented in the next chapter. In detail, these cuts were $E_{had}/E_{em} < 0.125$, $E/P < 4.0$, and $L_{shr} < 0.3$.

4.5 Distributions of the Electron Identification Variables for Signal and Background

The figures 4.6 and 4.7 show the distribution of the electron variables for the signal and background sample. One can see in the L_{shr} distribution for example, that some cuts were already applied on the electron candidates, when creating the *BTOP0G* dataset.

The isolation offers the best discrimination power. Hadrons faking an electron candidate are often part of a jet. Therefore, the surrounding calorimeter cells are hit as well by other hadrons, while the electrons themselves cause no activity around the electromagnetic cluster.

The ratio of the hadronic to the electromagnetic cluster energy is as well very different for signal and background electron candidates. The same is true for the ratio of the measured calorimeter energy to the momentum of the associated track. For real electrons, this ratio should be one. However, the electron might radiate a photon on its way through the tracking systems (bremsstrahlung). The photon is often detected in the same calorimeter tower as the electron, thus the measured calorimeter energy is not sensitive to the photon radiation, but the measured track momentum is lowered by the radiation. This effect causes a tail in the distribution for the signal electron candidates for E/P values above one. In a perfect detector, the measured calorimeter energy should always exceed the measured track momentum for the background. However, parts of the hadronic shower in the calorimeter are undetectable. This effect and measurement errors allow to get a ratio of one even for the background.

The lateral shower profile has a narrow peak around zero for the electrons candidates from the signal sample. For the background, one gets a much broader distribution, due to the differences in the shape of an electromagnetic and a hadronic shower. The same is true for the variables comparing the shower shape with test beam electrons: χ^2_{wire} and χ^2_{strip}. Furthermore, the track from real electrons matches the shower position, thus $Q \cdot \Delta x$ and Δz are close to zero for signal.

4.6 Tight CDF Cuts to Identify High-E_T Electrons

One option to identify electrons is to use the cuts applied in the CDF analyses for the winter conferences 2003 [95]. Table 4.2 lists the cuts used to identify an electron. The efficiencies reported in table 4.2 were derived from $Z \to e^+e^-$ events. Events with two electron candidates were selected. One of the candidates had to pass tight cuts without applying any cuts on the other leg. The candidates belonging to this other leg were used to evaluate the cuts. The overall efficiency for the CDF tight cuts from data is $(85.5 \pm 0.9)\%$.

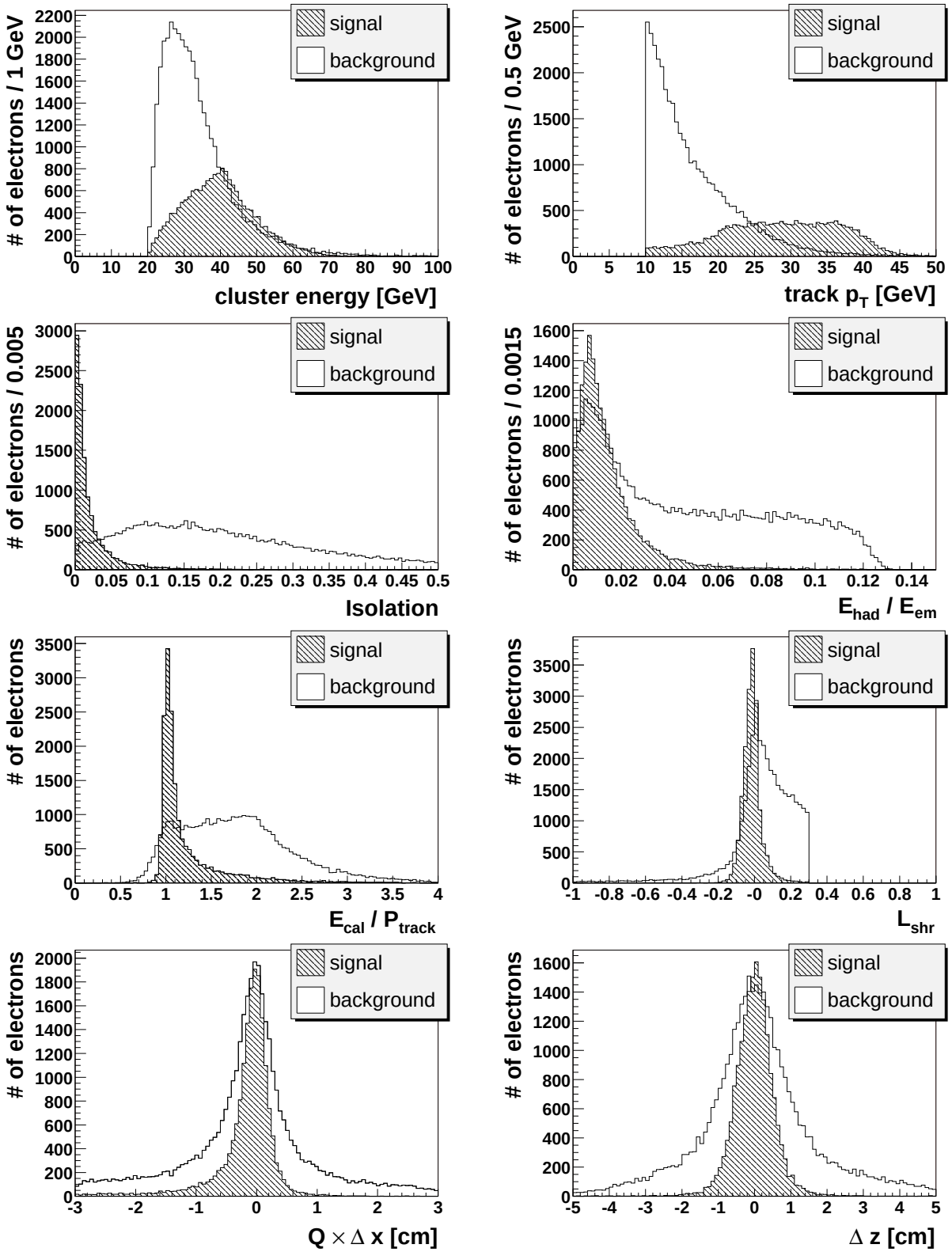


Figure 4.6: Comparison of the distributions of important electron variables for the signal sample (shaded area) and the background sample (white area). The total number of entries in the background histograms was normalized to the number of entries in the signal histograms.

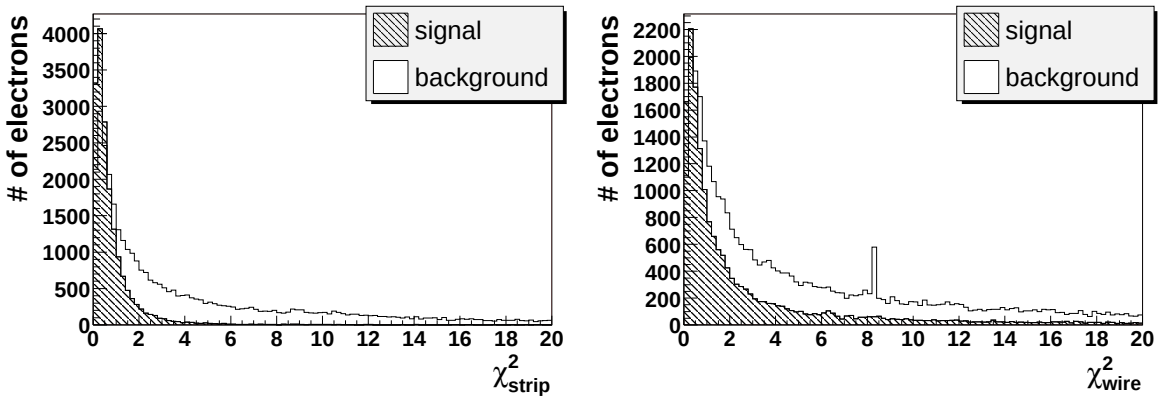


Figure 4.7: Comparison of the distributions of important electron variables for the signal sample (shaded area) and the background sample (white area). The total number of entries in the background histograms was normalized to the number of entries in the signal histograms.

Variable	Cut	efficiency(CDF-6262)
Isolation	< 0.1	$(97.1 \pm 1.0)\%$
E_{had}/E_{em}	$< 0.055 + 0.00045 \cdot E$	$(99.9 \pm 0.4)\%$
E/P	< 2.0 or $E_T > 50$ GeV	$(94.0 \pm 0.6)\%$
L_{shr}	< 0.2	$(97.7 \pm 0.5)\%$
$Q \cdot \Delta x$	> -3.0 cm and < 1.5 cm	$(99.1 \pm 0.3)\%$
$\lvert\Delta z\rvert$	< 3.0 cm	$(99.4 \pm 0.2)\%$
χ^2_{strip}	< 10	$(97.5 \pm 0.4)\%$
Axial SL	≥ 3	$(99.6 \pm 0.1)\%$
Stereo SL	≥ 3	$(97.8 \pm 0.3)\%$
all cuts		$(85.5 \pm 0.9)\%$

Table 4.2: Efficiency of the cuts used to identify tight electrons [95]. Listed are the efficiencies of each cut alone. The last row states the efficiency for all cuts combined.

To check my method of getting a signal sample, the CDF tight cuts were applied on the candidates in this sample and their efficiencies were determined. The results are listed in table 4.3.

Variable	Cut	signal	
		passed	efficiency
no cuts	-	17683	-
Isolation	< 0.1	17223	$(97.40 \pm 0.12)\%$
E_{had}/E_{em}	$< 0.055 + 0.00045 \cdot E$	17484	$(98.87 \pm 0.08)\%$
E/P	< 2.0 or $E_T > 50$ GeV	16566	$(93.68 \pm 0.18)\%$
L_{shr}	< 0.2	17565	$(99.33 \pm 0.06)\%$
$Q \cdot \Delta x$	> -3.0 cm and < 1.5 cm	17641	$(99.76 \pm 0.04)\%$
$\lvert\Delta z\rvert$	< 3.0 cm	17613	$(99.60 \pm 0.05)\%$
χ^2_{strip}	< 10	17501	$(98.97 \pm 0.08)\%$
Axial SL	≥ 3	17647	$(99.80 \pm 0.03)\%$
Stereo SL	≥ 3	17484	$(98.87 \pm 0.08)\%$
all cuts		15527	$(87.81 \pm 0.25)\%$

Table 4.3: Electron identification efficiency for signal events

The overall efficiency on my signal sample is $(87.81 \pm 0.25)\%$. This efficiency is a little bit higher than the quoted efficiency of the CDF tight cuts from $Z \to e^+e^-$ events. However, for my analysis tighter preselection cuts were applied, which results in a higher efficiency on my signal sample. To take any correlation between my signal sample and the W boson cross-section analysis performed in the next chapter and the impurity of the signal sample into account, I assume a 1% systematic error on the efficiency. Thus, the efficiency for the CDF tight cuts is $(87.81 \pm 1.03)\%$ in this analysis.

The background sample can be used to study the purity of the tight cuts. Out of the 37859 electron candidates in the sample, 3901 candidates passed the cuts. This corresponds to an efficiency of $(10.38 \pm 0.16)\%$ on the background sample. However, without the preselection cuts applied on the electron candidates during the creation of the *BTOP0G* dataset used for the background sample the efficiency on background would be much lower.

Variable	Cut	background	
		passed	efficiency
no cuts	-	37859	-
Isolation	< 0.1	9336	$(24.84 \pm 0.22)\%$
E_{had}/E_{em}	$< 0.055 + 0.00045 \cdot E$	26576	$(70.70 \pm 0.23)\%$
E/P	< 2.0 or $E_T > 50$ GeV	25991	$(69.15 \pm 0.24)\%$
L_{shr}	< 0.2	31549	$(83.93 \pm 0.19)\%$
$Q \cdot \Delta x$	> -3.0 cm and < 1.5 cm	35515	$(94.48 \pm 0.12)\%$
$\lvert\Delta z\rvert$	< 3.0 cm	34643	$(92.16 \pm 0.14)\%$
χ^2_{strip}	< 10	29239	$(77.79 \pm 0.21)\%$
Axial SL	≥ 3	37301	$(99.23 \pm 0.04)\%$
Stereo SL	≥ 3	34869	$(92.76 \pm 0.13)\%$
all cuts		3901	$(10.38 \pm 0.16)\%$

Table 4.4: Electron identification efficiency for background events

4.7 Very loose Electron Cuts

Although the efficiency of the tight cuts is much higher on signal than on background, there are still channels in which the cross-sections of background processes are so much higher than the signal cross-section, that a sizeable number of misidentified electrons is selected. To correct for this impurity, it is necessary to have a set of cuts that selects namely fake electrons. To achieve this, very loose cuts were defined based on the figures 4.6 and 4.7. These cuts have a very high efficiency for signal candidates, thus candidates that fail these cuts are very likely background.

The cuts and their efficiency on the signal sample are listed in table 4.5. 98.0% of the candidates in the signal sample pass these cuts.

4.8 Identification of High-E_T Electrons using an Artificial Neural Net

Artificial Neural Nets have proven themselves successful in discriminating signal from background in high-energy physics experiments. Especially the LEP experiments have used them extensively, as the clean signatures of electron-positron collisions and their

Variable	Cut	passed	efficiency
no cuts	-	17683	-
Isolation	< 0.2	17554	$(99.27 \pm 0.06)\%$
E_{had}/E_{em}	$< 0.08 + 0.00045 \cdot E$	17615	$(99.62 \pm 0.05)\%$
E/P	< 3.5 or $E_T > 50$ GeV	17639	$(99.75 \pm 0.04)\%$
L_{shr}	< 0.2	17565	$(99.33 \pm 0.06)\%$
all cuts		17336	$(98.04 \pm 0.10)\%$

Table 4.5: Electron identification efficiency for signal events using very loose cuts

good theoretical understanding allowed to train and study the nets on simulated events. However, at a hadron collider the situation is more difficult. It is not yet possible to simulate proton-antiproton collisions with the same accuracy. Furthermore, the main background in physics analyses at a hadron collider arises from QCD events, like two jet production, that fake the signal. The large cross-section of these processes makes it very difficult to generate enough background events. A possible solution of this problem might be to use data to train the neural net. This approach is described in this section for the electron identification.

To obtain the efficiency and the purity of the tight electron identification cuts, a signal and a background sample were selected from data, as described in section 4.1. In the signal sample, 17683 electron candidates passed the basic cuts. 37849 electron candidates passed these cuts in the background sample. I tried then to use these samples to train an Artifical Neural Net. As implementation of the neural net the *Neurobayes* package [100] was used, that is being developed at Karlsruhe and has successfully been used by the *Delphi* group at Karlsruhe in physics analyses.

As mentioned before, the purpose of the net is to discriminate real electrons from QCD fakes. Great care is needed in selecting the input variables of the net. To illustrate this, the electron transverse energy is plotted for the background and signal sample in figure 4.8. There is a clear difference in the spectrum. However, using this difference in the net would discriminate W boson events from QCD events instead of electrons from fakes. Therefore, I limited myself to variables that are less dependent on the production process. The input variables were: E_{had}/E_{em}, isolation, L_{shr}, χ^2_{strip}, $Q \cdot \Delta x$, Δz, and E/P.

Before the net is trained, the correlations of the input variables are determined by the *Neurobayes* program. A new set of orthogonal variables is then used internally to train the net. Table 4.6 shows the covariance matrix of the input variables in percentage. All signal and background electrons were used to determine the linear correlation. The background events were weighted to have the same sum of weights as the number of signal electrons.

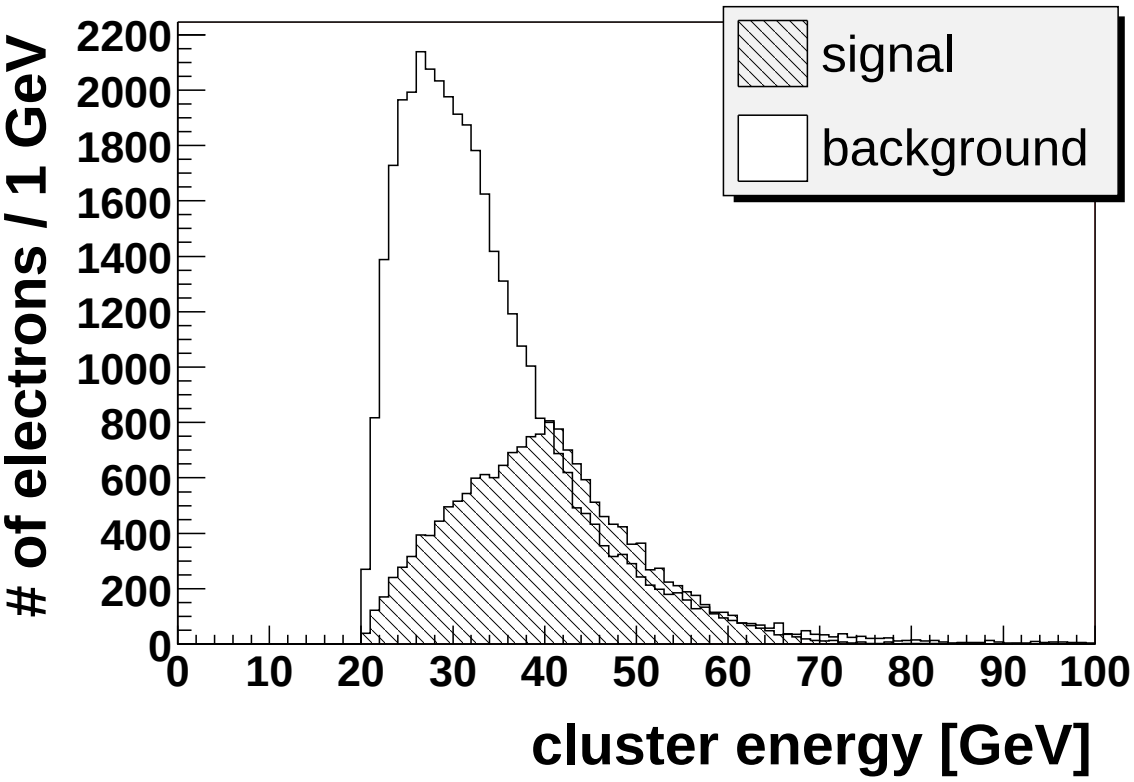


Figure 4.8: Energy of the signal and background electron candidates.

	target	E_{had}/E_{em}	iso	L_{shr}	χ^2_{strip}	$Q \cdot \Delta x$	Δz	E/P
target	100.0	-45.4	-74.0	-28.3	-49.3	-3.9	-2.0	-43.5
E_{had}/E_{em}	-45.4	100.0	41.7	13.9	35.4	2.0	2.2	17.4
iso	-74.0	41.7	100.0	24.3	42.0	2.8	1.2	33.1
L_{shr}	-28.3	13.9	24.3	100.0	9.7	2.2	0.0	7.2
χ^2_{strip}	-49.3	35.4	42.0	9.7	100.0	2.6	8.1	27.9
$Q \cdot \Delta x$	-3.9	2.0	2.8	2.2	2.6	100.0	-0.4	-13.3
Δz	-2.0	2.2	1.2	0.0	8.1	-0.4	100.0	1.4
E/P	-43.5	17.4	33.1	7.2	27.9	-13.3	1.4	100.0

Table 4.6: Covariance matrix of the input variables in percentage.

As can be seen in figures 4.6 and 4.7, the isolation has the largest correlation with the target, a value indicating whether an electron candidate belongs to the signal sample (target = 1) or the background sample (target = -1). E_{had}/E_{em}, L_{shr}, and E/P are as well quite different for signal and background. However, all four variables are correlated, thus reducing their discrimination power. One advantage of an Artificial Neural Net is that it takes these correlations into account.

70% of the data were used to train the net and the remaining events formed a test sample. The training was iterated 50 000 times. Figure 4.9 shows the net output for all electron candidates in the signal and background sample after the training. One can see a clear distinction between the electron candidates from the two samples. Table 4.7

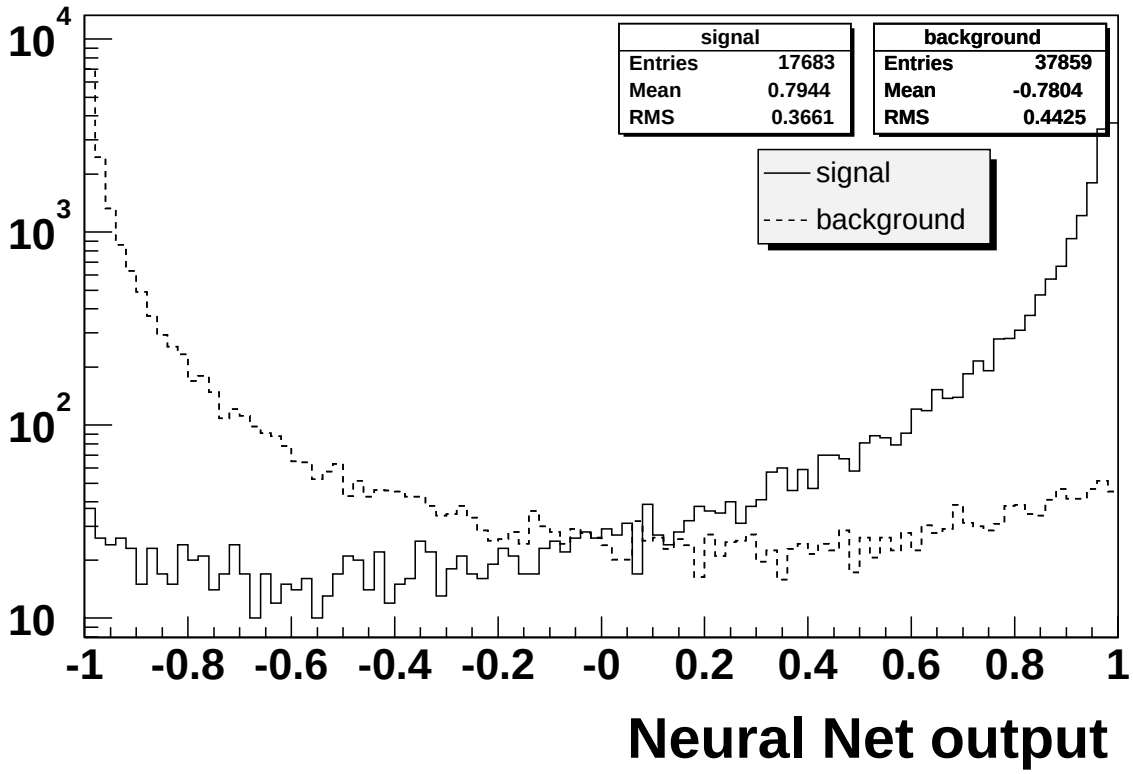


Figure 4.9: Net output for the signal (solid line) and background (dashed line) electrons after the training.

quantifies this by listing the efficiencies on background and signal for different cuts on the net output. Accepting electron candidates with a net output greater than 0.5 has an efficiency of 88.2% on the signal sample, while only 4.7% of the electron candidates from the background sample pass this cut. This should be compared with the efficiency of the tight cuts. The numbers were 87.8% on the signal sample and 10.4% on the background

sample. Thus, this cut on the net output selects only half the number of background events as the tight cuts while being as efficient for the signal. However, these numbers were obtained from the samples used for the training. As for the obtained efficiency of the CDF tight cuts, one should add a 1% systematic error due to impurities of the signal sample to the statistical error.

cut	signal efficiency	background efficiency
> 0.95	$(40.1 \pm 0.4)\%$	$(0.5 \pm 0.1)\%$
> 0.9	$(57.2 \pm 0.4)\%$	$(1.0 \pm 0.1)\%$
> 0.8	$(74.2 \pm 0.3)\%$	$(2.2 \pm 0.1)\%$
> 0.7	$(82.4 \pm 0.3)\%$	$(3.3 \pm 0.1)\%$
> 0.5	$(88.2 \pm 0.2)\%$	$(4.7 \pm 0.2)\%$
> 0	$(94.4 \pm 0.2)\%$	$(8.0 \pm 0.2)\%$
> -0.5	$(97.2 \pm 0.1)\%$	$(12.8 \pm 0.3)\%$
> -0.7	$(98.1 \pm 0.1)\%$	$(16.7 \pm 0.3)\%$
> -0.9	$(99.2 \pm 0.1)\%$	$(30.7 \pm 0.3)\%$
> -0.95	$(99.5 \pm 0.1)\%$	$(39.2 \pm 0.4)\%$

Table 4.7: Efficiency on the signal and background sample for different cuts on the net output.

To check the reported efficiencies, the missing transverse energy spectrum is plotted for all events in the *BTOP0G* dataset with an electron candidate accepted by a given identification method. Only events with a missing transverse energy above 15 GeV were used. Figure 4.10 shows the spectrum for the tight cuts and two different cuts on the net output. A typical spectrum falls off starting from the cut value of 15 GeV. Around 20 GeV there is a local minimum. The number of events increases sharply for missing transverse energies above 25 GeV and reaches its maximum at 40 GeV. This is the Jacobian peak for the neutrinos from W boson decay. Events in this region contain mostly real electrons while the number of events with a low missing transverse energy gives a measure for the purity of the cut. Electrons from events in the low missing transverse energy region were most likely misidentified candidates.

A comparison of the curves for the CDF tight cuts and the cut on the net output at 0.0 reveals that the latter selects less events in the low missing transverse energy region, but more events in the Jacobian peak region. Therefore, this cut on the net output has a better purity and efficiency than the tight cuts. Another advantage of cutting on the neural net output is that one can easily vary the cut, e.g. to select a very pure electron

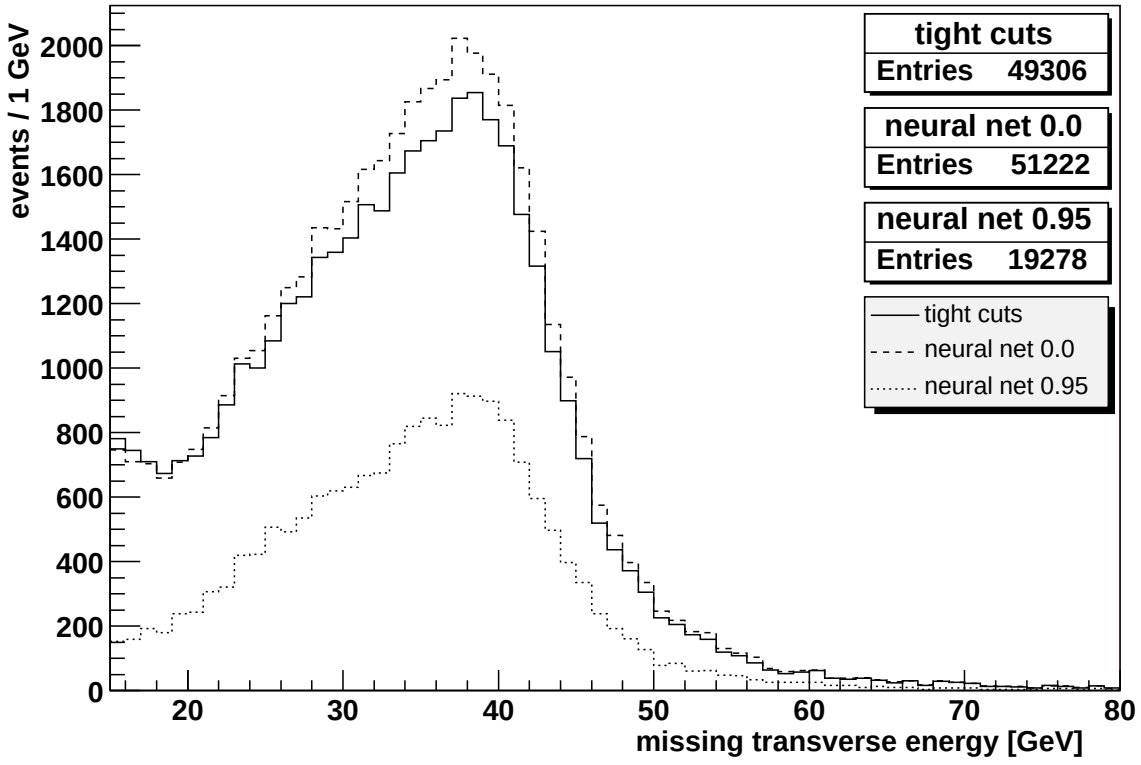


Figure 4.10: Missing transverse energy spectrum for events selected with the tight electron cuts (solid line) and events selected with the Artificial Neural Net (dashed line).

sample. This is demonstrated with the curve for the cut on the neural net output of 0.95 in figure 4.10. The number of events in the region below 25 GeV is reduced by a factor of six with this cut, while the number of events in the peak region is only by a factor of two lower. Hence, the fraction of misidentified electrons is heavily reduced using this cut. Another possibility is to select only events with a very low net output of their electron candidates. This sample would contain mostly QCD events and would allow to study the background of a physics measurement.

The trained neural net combines the seven main variables for the electron identification to one variable: the net output. As proven in the last paragraph, this new variable has a better discrimination power than the set of cuts used so far at CDF. Further improvements might be achieved by applying loose cuts on the input variables, e.g. by cutting away candidates with a very high isolation fraction. This would remove evident background before the training and improve the sensitivity of the net on this variable for the remaining events. As can be seen in figure 4.6, the isolation variable for example does not exceed 0.15 for the signal, but can have values above 0.4 for the background.

A concern about Artificial Neural Nets is that they might be more sensitive to systematic changes in the input variables. For the cuts, systematics, e.g. a change in the CEM energy scale, can be studied by varying the cuts. For a neural net, this can be done by varying the input variables, in this case by scaling E/P, and study the change in the efficiency of the net. The net does not need to be trained again for this.

If the isolation of the electrons changes due to higher luminosity and consequently more interactions per bunch-crossing, one has to study this effect with the new data and eventually retrain the net, but the same is true for cuts. There one has to reevaluate the efficiency as well and maybe adjust the cuts.

Chapter 5

W Boson Cross-Section

In this chapter, I will describe my measurement of the cross-section of W boson production in proton-antiproton collisions with a center-of-mass energy of 1.96 TeV. I limited myself to the decay $W \to e\nu$.

After the electron has been identified, the electron four-momentum is constructed with the total energy measured in the calorimeter and the direction of the track associated with the electron. At a hadron collider, only the x and y momentum of the neutrino can be derived from the momentum balance of all measured particles. This is obtained by summing up all calorimeter towers and adding the momenta of muons that penetrate the calorimeter. The missing momentum is attributed to the neutrino. Particles in the very forward direction leave the detector undetected and the z-component of the neutrino momentum remains unknown.

Therefore, the W boson can only partially be reconstructed using the missing transverse energy and the electron four-momentum. The z momentum of the W boson stays unknown. However, the transverse momentum and the transverse mass of the W boson can be determined. The transverse mass m_T is defined by:

$$m_T^2 = (p_T^e + p_T^\nu)^2 - (p_x^e + p_x^\nu)^2 - (p_y^e + p_y^\nu)^2, \tag{5.1}$$

where p^e denotes the electron momentum and p^ν the neutrino momentum.

5.1 Data Sample and Event Selection

The data sample consisted of events in the *BTOP0G* data set with one identified electron, please see section 3.6.2. The preselection cuts mentioned in section 4.4 were applied to the data. I further removed events that contain at least one muon with a transverse momentum above 20 GeV and a stub reconstructed in the muon chambers. A veto

on events with a second electron candidate passing the *BTOP0G* preselection cuts was imposed to avoid Z boson events, and other sources of dilepton events. Additionally, for all analyses the missing transverse energy in the event was required to be above 15 GeV in the preselection. The missing transverse energy was corrected for muons and recalculated for the primary vertex of each event.

The data was analyzed for three different electron identification methods. To compare with the previous results [101], the same electron identification cuts were applied in the first analysis. Furthermore, I used two different cuts on the output of the Artifical Neural Net trained to identify electrons, as described in chapter 4. In one selection, the net output for an electron candidate had to be above 0.0, in the other selection, it had to exceed 0.95. For the final analysis, the cut on the missing transverse energy was raised to 20 GeV. This is the main difference to the analysis described in [101]. There the threshold has been 25 GeV.

Figures 5.1 and 5.2 compare the distributions for the missing transverse energy and the electron transverse energy for the different electron selections. The looser cut on the Artificial Neural Net offers a slightly better efficiency for electrons in the peak region around 40 GeV than the tight cuts.

5.2 Signal Monte Carlo Sample

As the Monte Carlo sample for the signal 160.000, $W \to e\nu$ events from the *WTOP0E* data set were used. More details about the sample can be found in section 3.6.1. Unfortunately, the used generator, Pythia, does not take gluon nor quark emissions fully into account. This results in a lower transverse momentum of the W boson in Monte Carlo compared to data and a lower jet multiplicity.

To correct for this effect, we studied the energy recoiling of the W boson U, that is $\vec{U} = -\vec{\not\!\!E}_T - \vec{E}_T^e$. Here $\vec{\not\!\!E}_T$ denotes the missing transverse energy and $\vec{E}_T^e$ the electron transverse energy. As in [101], the calculated recoil energy for Monte Carlo was compared with data, see figure 5.3. One can see that the recoil energy for Monte Carlo events tends to be lower than for data. To obtain better agreement, $\vec{U}$ had to be scaled up by 5% for Monte Carlo [101]. The missing transverse energy is then recalculated as $\vec{\not\!\!E}_T = -\vec{E}_T^e - \vec{U}$. Figure 5.4 shows the improved agreement after the correction. In the further analysis, this corrected missing transverse energy was used for Monte Carlo.

5.3 Acceptance and Efficiency of the Preselection

I have used 160.000 $W \to e\nu$ Monte Carlo events from the *WTOP0E* data set to evaluate the acceptance and efficiency of the preselection cuts applied in my analysis.

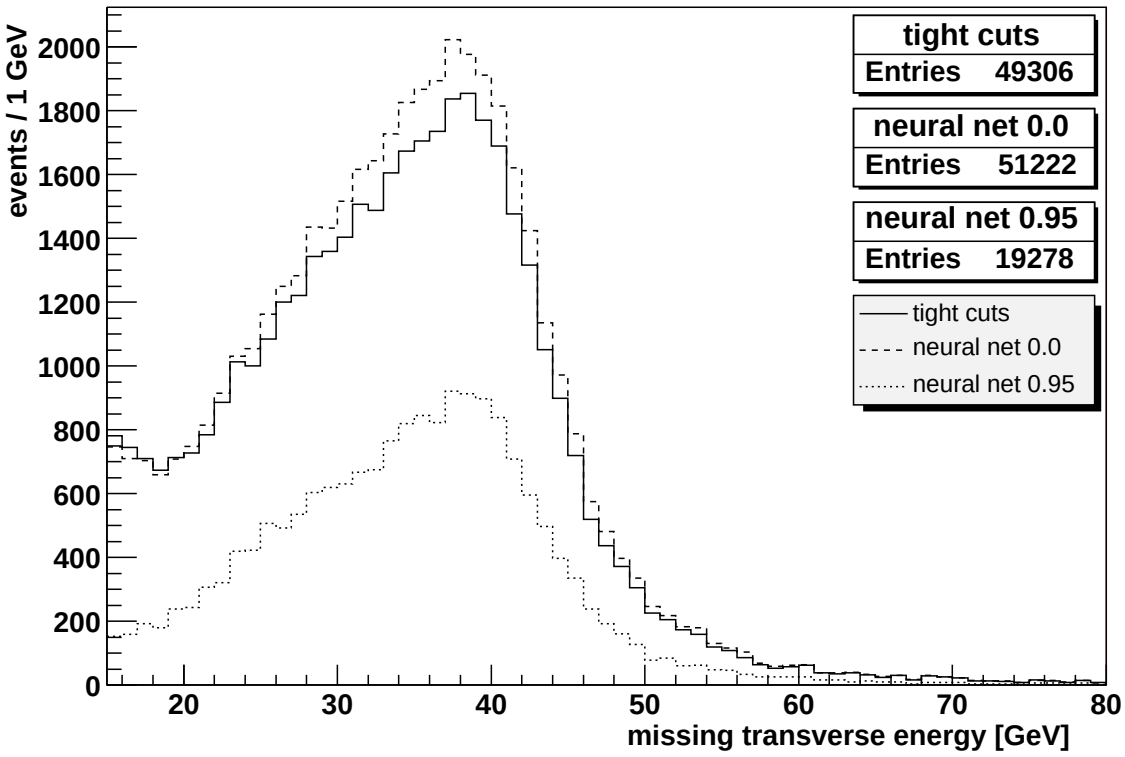


Figure 5.1: Missing transverse energy spectrum for events selected with the tight electron cuts (solid line) and events selected with two different cuts on the Artifical Neural Net output (dashed and dotted lines).

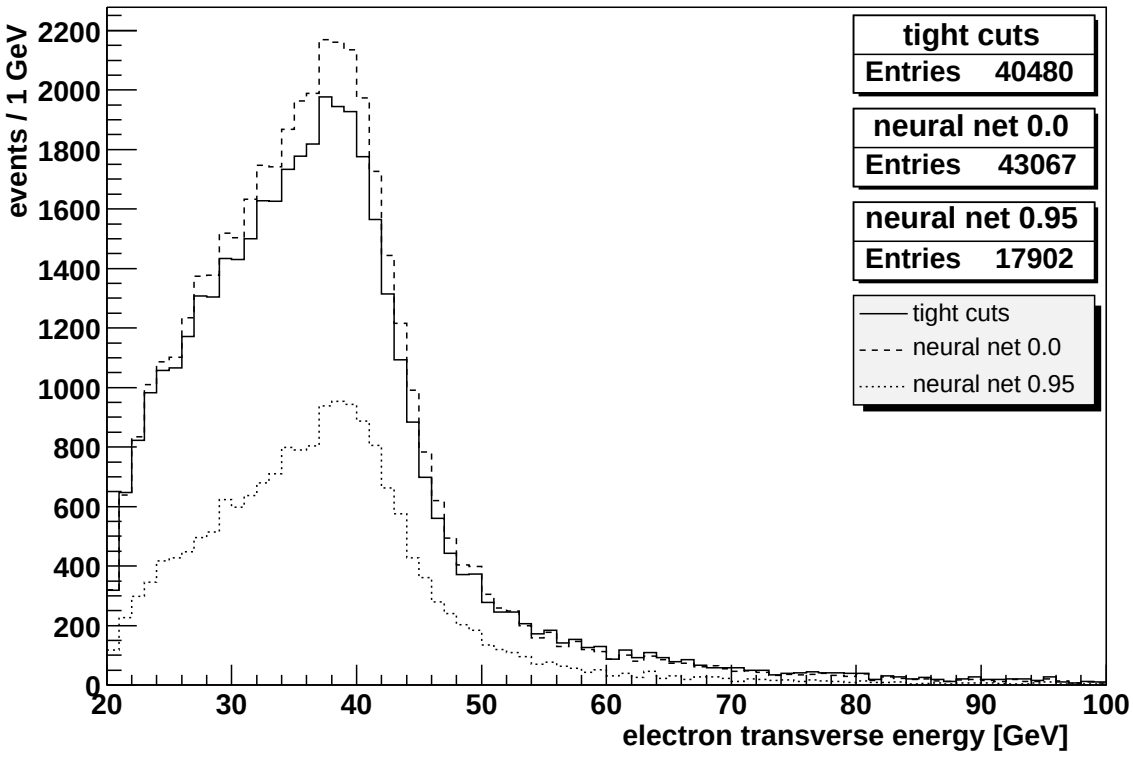


Figure 5.2: Electron transverse energy spectrum for events selected with the tight electron cuts (solid line) and events selected with two different cuts on the Artifical Neural Net output (dashed and dotted lines).

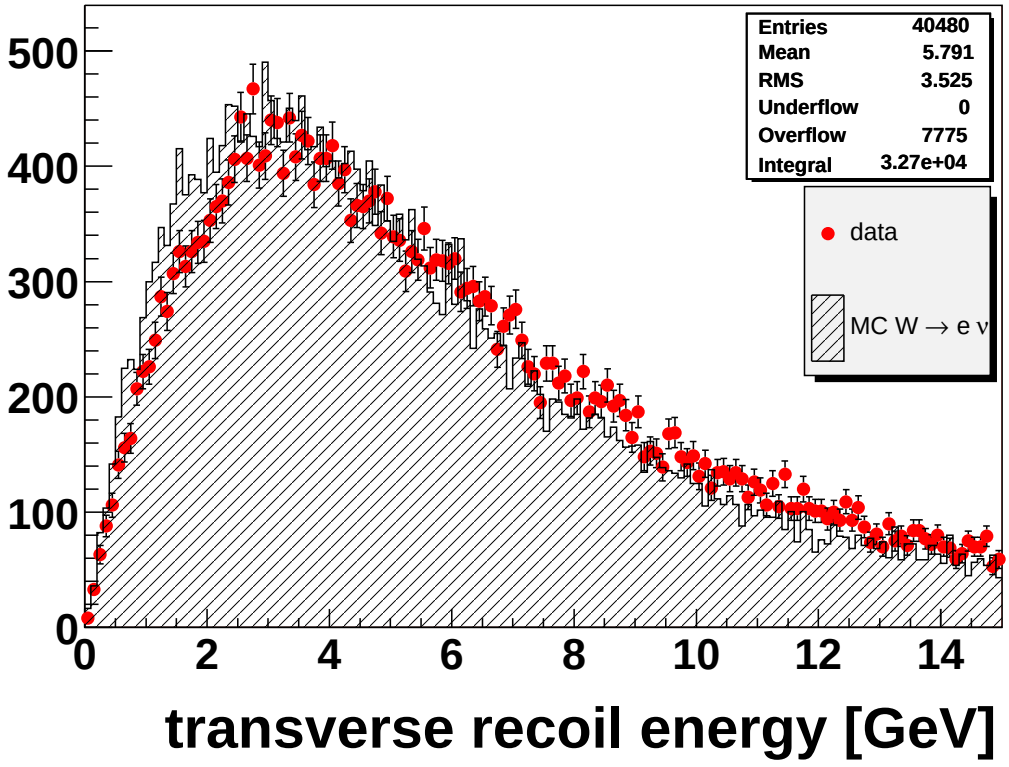


Figure 5.3: Transverse energy recoiling of the reconstructed W boson before the correction. The data points (dots) are compared with Monte Carlo $W \to e\nu$ events (shaded area).

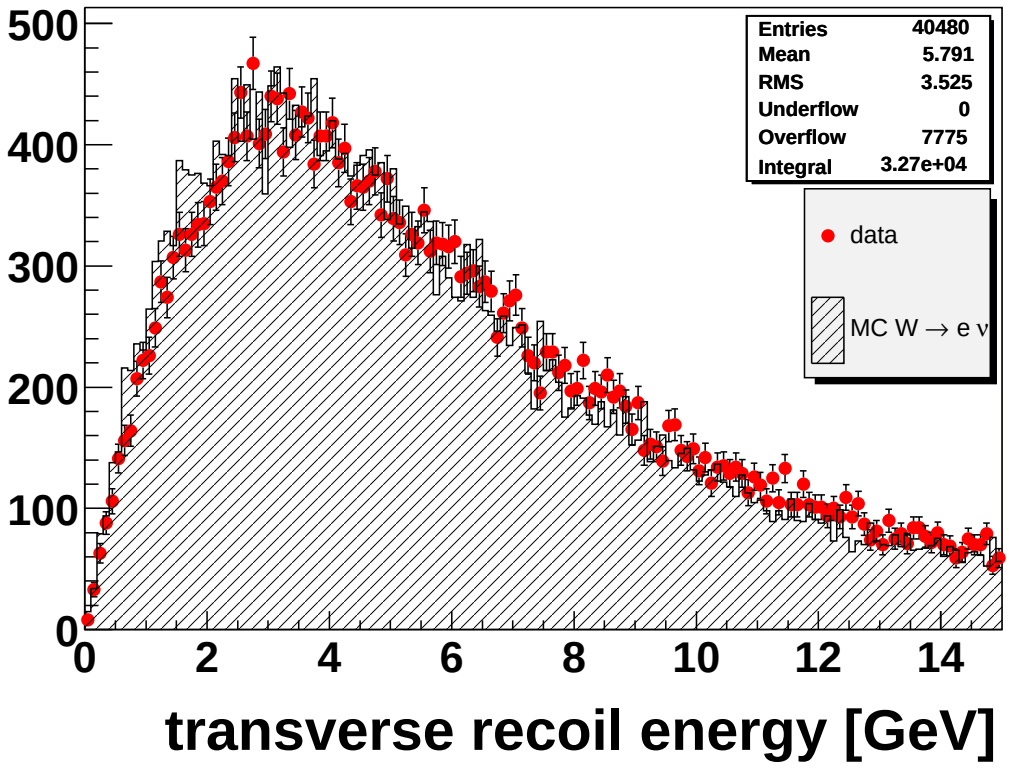


Figure 5.4: Transverse energy recoiling of the reconstructed W boson after the correction. The data points (dots) are compared with Monte Carlo $W \to e\nu$ events (shaded area).

cut	passed	acceptance and efficiency
before cuts	160000	-
$E_{T,miss} > 15$ GeV	136804	$(85.50 \pm 0.09)\%$
cuts for highly energetic electron		
$E_T > 18$ GeV	118536	86.65%
$E_{had}/E_{em} < 0.125$	118241	86.43%
$P_T > 9.0$ GeV	113741	83.14%
$E/P < 4.0$	77960	56.99%
$L_{shr} < 0.3$	56202	41.08%
$\|\Delta x\| < 0.3$	55421	40.51%
$\|\Delta z\| < 0.5$	55217	$(40.36 \pm 0.13)\%$
cuts for very highly energetic electrons		
$E_T > 70$ GeV, $P_T > 15.0$ GeV, $\|\Delta x\| < 0.3$, $\|\Delta z\| < 0.5$	339	$(0.25 \pm 0.01)\%$
combined electron cuts	55224	$(40.37 \pm 0.13)\%$
after all cuts	55224	$(34.52 \pm 0.12)\%$

Table 5.1: Cut flow table of the acceptance and efficiency of the *BTOP0G* preselection cuts on $W \to e\nu$ Monte Carlo. The acceptances and efficiencies of the two sets of cuts for electrons are calculated with respect to the 136804 events that passed the missing transverse energy cut. The last row states the acceptance and efficiency for all cuts.

At first, I evaluated the cut requiring a missing transverse energy above 15 GeV that had been applied when creating the Stntuple files. 136.804 of the 160.000 events passed this cut for the signal Monte Carlo. Thus, the efficiency of this first missing transverse energy cut is $(85.50 \pm 0.09)\%$.

5.3.1 Cuts for the *BTOP0G* Data Set

The first cut performed on the processed *BHEL08* data were the selection cuts on the electron candidates to create the used data set, *BTOP0G*. These cuts are described in section 4.4.1. By applying the same cuts on the signal Monte Carlo, the acceptances and efficiencies listed in table 5.1 were obtained. The electron tracks can only be reconstructed with high efficiency and accuracy in the central region of the detector using the drift

chamber. This explains the large reduction due to the E/P cut, as a track is needed to calculate this quantity. Furthermore, L_{shr}, Δx, and Δz are only defined for the central calorimeter. As a result, only electrons with a pseudorapidity $|\eta| < 1.1$ can pass these cuts. Events could enter the data set by passing one of the two different sets of electron cuts. Nearly all electrons from W boson decay pass the highly energetic electron cuts, thus the second path is not needed for this analysis. Its main purpose is to allow searches for very massive particles that decay into a final state with a very highly energetic electron, e.g. searches for Z' and W'.

5.3.2 Further Electron Preselection Cuts

For these analyses, further cuts on the electron candidates have been applied. The corrected electron variables were used for these cuts. To keep away from the trigger thresholds, I tightened the cuts on the transverse energy and the transverse track momentum. As will be explained in the section about the background sample, section 5.4, the highly energetic electron preselection cuts were imposed on all electron candidates. This removes events from the sample that have an electron only passing the very highly energetic electron cuts. Furthermore, the electron had to be fiducial. Besides these cuts already mentioned in section 4.4.2, events with two electron candidates passing the *BTOP0G* preselection cuts were excluded and events that contain muons with a transverse momentum above 20 GeV and a stub reconstructed in the muon chambers were removed. Table 5.2 lists the acceptances and efficiencies of these cuts on the signal Monte Carlo. It shows, that the cuts to remove events passing through the very highly energetic electron preselection, the cuts on E_{had}/E_{em}, L_{shr}, and E/P, have almost no impact on the overall acceptance and efficiency. The same is true for the cuts removing events with a second highly energetic electron or a high momentum muon.

To exclude a large fraction of QCD events, the cut on the missing transverse energy was tightened by going from 15 GeV to 20 GeV. The efficiency of this cut was evaluated on the signal Monte Carlo. Table 5.3 lists the obtained efficiencies for this cut for the different electron selections. As expected, the electron selection does not affect the efficiency of the missing transverse energy cut for signal Monte Carlo. All obtained values agree well with each other. For the cross-section calculation, the value before the electron identification was used, that is $(96.77 \pm 0.08)\%$. Using this procedure, the overall acceptance and efficiency of the selection cuts on signal Monte Carlo is $A_W = \frac{43693}{160000} = (27.31 \pm 0.11)\%$. As the same Monte Carlo sample was used as in the previous analysis [101], the same systematic error of 1.05% was assumed. This results in a final value of $A_W = (27.31 \pm 1.06)\%$.

cut	passed	acceptance and efficiency
before cuts	160000	-
passed previous cuts	55224	-
exactly one electron passing cuts 4.4.1	55208	99.97%
no muon with a stub and $p_T > 20$ GeV	55208	100.00%
further electron selection cuts		
$E_T^{cor} > 20$ GeV	53095	96.17%
$P_T^{cor} > 10.0$ GeV	51997	94.18%
$E_{had}/E_{em} < 0.125$	51996	94.18%
$L_{shr} < 0.3$	51992	94.17%
$E/P < 4.0$	51990	94.17%
fiducial	46622	84.45%
$\lvert z_0^{track} \rvert < 60.0$ cm	45151	81.78%
overall acceptance	45151	$(28.22 \pm 0.11)\%$

Table 5.2: Acceptance and efficiency of further preselection cuts on $W \to e\nu$ Monte Carlo.

selection	all events	events with $\not\!\!E_T > 20$ GeV	cut efficiency
before electron id	45156	43693	$(96.77 \pm 0.08)\%$
CDF tight cuts	39575	38265	$(96.69 \pm 0.09)\%$
ANN netout > 0.0	42244	40892	$(96.80 \pm 0.09)\%$
ANN netout > 0.95	13346	12930	$(96.88 \pm 0.15)\%$

Table 5.3: Efficiency of the cut on the missing transverse energy on the $W \to e\nu$ Monte Carlo sample for the different electron selections.

5.4 Background Estimates

There are two classes of backgrounds. Processes that have a highly energetic electron in their final state are irreducible with respect to the electron identification, e.g. weak decays. The other class consists of QCD events where hadrons fake an electron. These are of special importance in the study of W boson plus jets events.

5.4.1 Irreducible Background

The two main irreducible backgrounds are $Z \to e^+e^-$ events, where one electron has left the detector undetected and accounts for the missing transverse energy, and $W \to \tau\nu_\tau \to e\nu_\tau\nu_\tau\nu_e$. The latter is suppressed by the branching fraction of $\tau \to e\nu_\tau\nu_e$ of $(17.84 \pm 0.06)\%$ [27]. Furthermore, the electrons from this process are less energetic and fail the transverse energy cut more often than electrons directly from the W decay. In the CDF analysis [101], the contributions from these two processes have been evaluated to be $(0.92 \pm 0.04)\%$ for $Z \to e^+e^-$ events and $(2.06 \pm 0.06)\%$ in case of $W \to \tau\nu_\tau$.

The different missing transverse energy cut used in my analyses should not affect the fraction of $Z \to e^+e^-$ events. However, $W \to \tau\nu_\tau$ events tend to have a smaller missing transverse energy than $W \to e\nu_e$. To take this into account, I reevaluated the contribution from this process by determining the acceptance and efficiency for the $W \to \tau\nu_\tau$ Monte Carlo. 224000 events from the *WTOP1T* data set were used. The acceptance and efficiency was $A_W^\tau = \frac{2492}{224000} = (1.11 \pm 0.02)\%$. As the cross-sections for both processes are identical, the ratio of the acceptance and efficiency for $W \to \tau\nu_\tau$ and the sum of the acceptances and efficiencies for $W \to e\nu_e$ and $W \to \tau\nu_\tau$ corresponds to the fraction of $W \to \tau\nu_\tau$ events f_τ in the data after the subtraction of other background sources. I get $f_\tau = \frac{A_W^\tau}{A_W + A_W^\tau} = (3.91 \pm 0.16)\%$. This value is much larger than the background fraction in the previous analysis due to the looser cut on the missing transverse energy.

The combined value of the fraction of irreducible background in the data samples is: $f_{\text{bkg}}^{\text{irreducible}} = (4.83 \pm 0.20)\%$.

The acceptance and efficiency for $W \to e\nu$, A_W, is correlated with the background fraction from $Z \to e^+e^-$ and $W \to \tau\nu$ events, as its value was obtained from Monte Carlo as well. An overestimation of the acceptance and efficiency for the simulated signal will reduce the measured cross-section. An overestimation of the acceptance and efficiency for the background processes will increase the expected background fraction. This will lead to a smaller cross-section as well. Thus the absolute values of the systematic errors on these parameters need to be added, when calculating the overall systematic error.

5.4.2 Background from QCD Events

The missing transverse energy spectrum for data rises going to smaller values of missing transverse energy, see figure 5.1. Figure 5.5 shows the same distribution for the $W \to e\nu$ Monte Carlo, *WTOP0E* data sample. There the distribution steadily falls going from the peak region around 40 GeV to lower values. The excess in data is due to QCD events where a hadron fakes an electron. Thus, the low energy bins in the missing transverse energy distribution shown in figure 5.1 allow to determine the fraction of QCD events in the data sample. To model the QCD background, events from the *BTOP0G* data set were selected that were most likely QCD events, but passed the missing transverse energy cut. For the tight cut analysis, events were used for that the electron candidate failed the very loose cuts defined in section 4.7. As only 2.0% of the real electrons will enter this sample (see table 4.5), the sample should mostly consist of QCD background.

Figure 5.6 shows the transverse energy distribution of the electron candidates in this background sample. One can easily identify the effect of the two different sets of electron preselection cuts applied at the creation of the *BTOP0G* data sample. The very highly energetic electron cuts are less stringent for QCD background. As was shown in section 5.3, this second cut set is negligible for electrons from W boson decay. To remove this artefact caused by the preselection cuts for the background sample, all events in this analysis have to pass the tighter cuts of the highly energetic electron cut set.

For the analyses using Artifical Neural Nets, events were chosen with an electron candidate with a net output less -0.95 to create a background sample. The efficiency of this selection is 0.5% for real electrons and 60.8% for fakes, compare table 4.7. Figure 5.7 shows the missing transverse energy and electron transverse energy distributions for the two background samples. The fall in these distributions is typical for QCD events and confirms the selected cuts for a QCD background sample. However, the shape of the two electron transverse energy distributions is slightly different. Though there are less events with an electron transverse energy ranging from 20 to 40 GeV for the Artifical Neural Net sample, there is an excess for large transverse energies. This may be due to the fact, that electrons from W boson decay were used for the training. Thus, the signal training sample did not offer much statistics for electron candidates with a transverse energy above 50 GeV and the net has a smaller discriminative power there. The missing transverse energy spectrum in figure 5.7 shows, that the background sample selected with the loose cuts is not as pure as the sample selected using the Artificial Neural Net. The structure of the distribution looks a little bit odd and might be caused by different triggers cutting on the missing transverse energy.

To check the purity of the background sample, I looked for any forward-backward asymmetry of the electron candidates. As explained in section 1.3, at the Tevatron W^+ bosons have more likely a boost in the proton (forward) direction and W^- bosons in the backward direction. Hence, the electron has more often a negative pseudorapidity than the

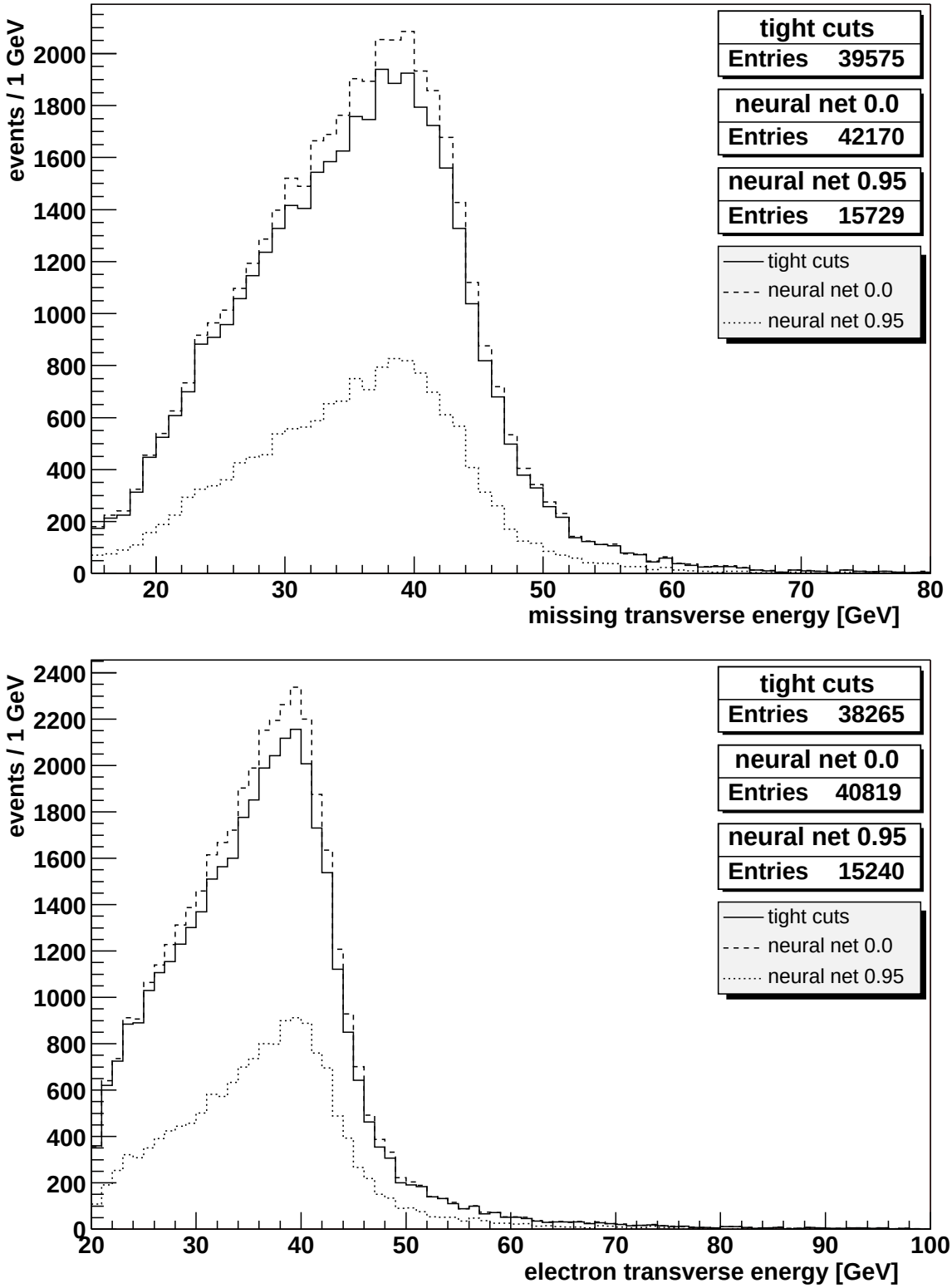


Figure 5.5: Missing transverse energy spectrum (top) and electron transverse energy spectrum (bottom) for $W \to e\nu$ Monte Carlo events selected with the tight electron cuts (solid line) with two different cuts on the Artifical Neural Net output (dashed and dotted lines).

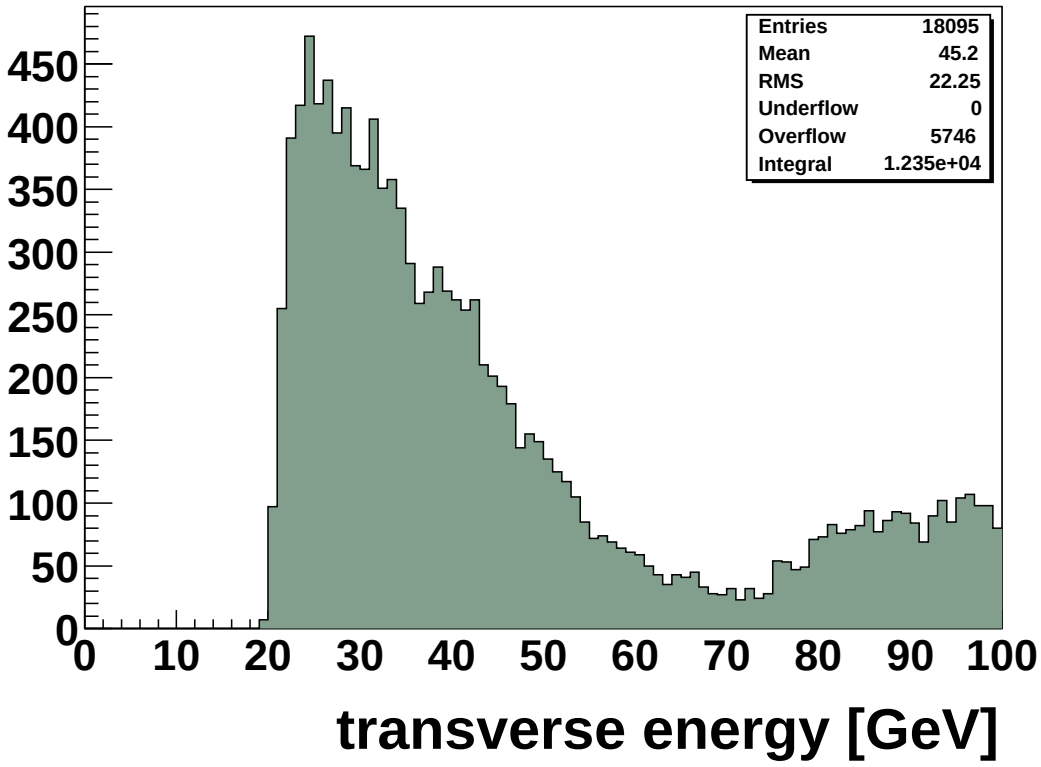


Figure 5.6: Transverse energy spectrum of electron candidates in the QCD background sample before tightening the preselection cuts.

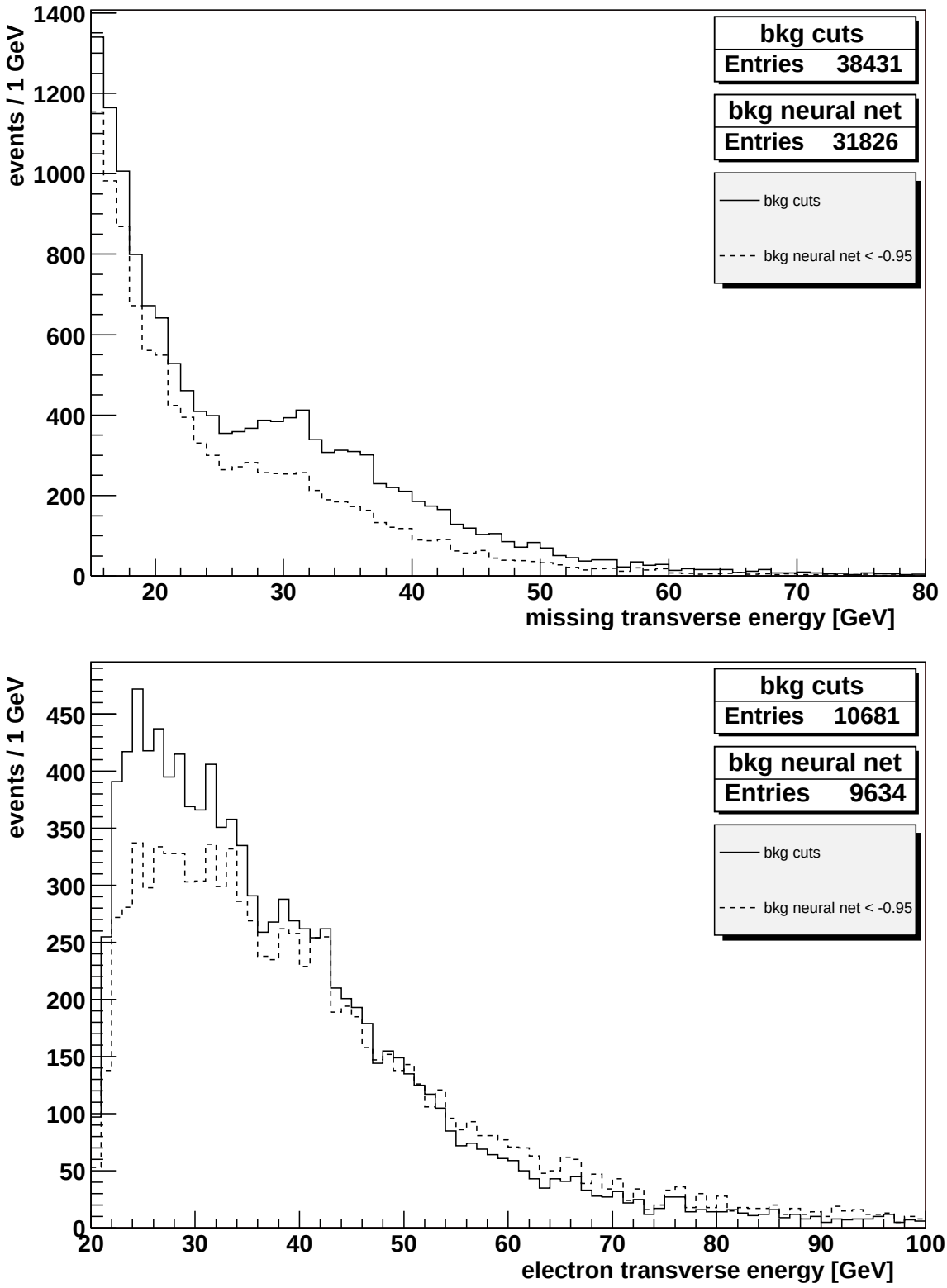


Figure 5.7: Missing transverse energy spectrum (top) and electron transverse energy spectrum (bottom) for background events selected with the very loose electron cuts (solid line) and background events selected with the Artifical Neural Net (dashed line).

positron. The ratio $\frac{N^+-N^-}{N^++N^-}$ characterizes the size of this asymmetry. As this asymmetry is sensitive to the parton distribution functions (PDFs) for the proton, one can test different PDFs by measuring the asymmetry. However, sizeable differences between the PDFs can only be found at large absolute pseudorapidities. Figure 5.8 compares the asymmetry in data selected via the tight electron identification cuts with Monte Carlo. One can clearly see the asymmetry in data. Data and Monte Carlo agree well. No asymmetry can be seen for the background sample showing that the background sample is rather free of W boson events.

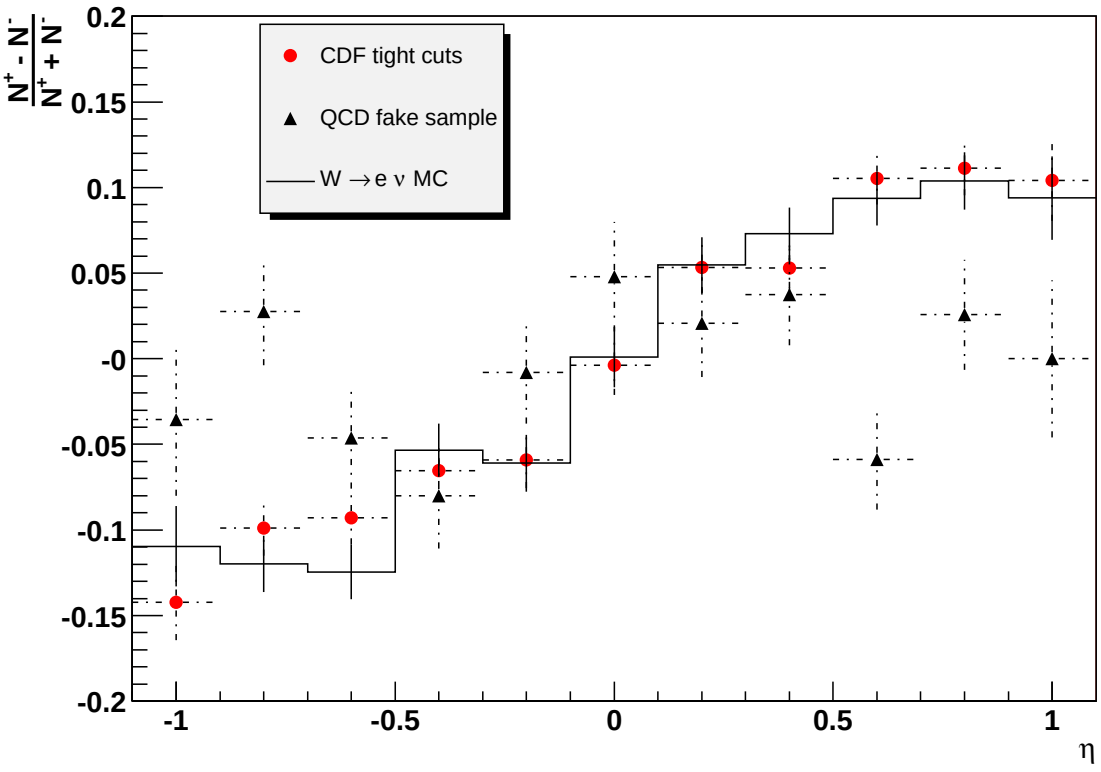


Figure 5.8: Forward-backward asymmetry of the electrons selected by the tight electron identification cuts. The data events (dots) are compared with the $W \to e\nu$ Monte Carlo (histogram) and the QCD background sample (triangles).

As the low transverse missing energy region in the data distribution is dominated by QCD background, the sum of the transverse missing energy distributions for QCD background and Monte Carlo for $W \to e\nu$ and $W \to \tau\nu$ was fitted to this distribution to determine the fraction of QCD events in the data. For the fraction of $W \to \tau\nu$ Monte Carlo events, the result from the previous subsection was used. The fit was done in the region 15 GeV $< \not\!\!E_T <$ 25 GeV using *Minuit*[102]. In the fit, I minimized the maximum

Kolmogorov distance between the summed histogram and the data histogram. After the fit, only the histogram entries with missing transverse energy above 20 GeV were used to calculate the QCD background fraction from the fit result. Figures 5.9 and 5.10 show the results for the three analyses.

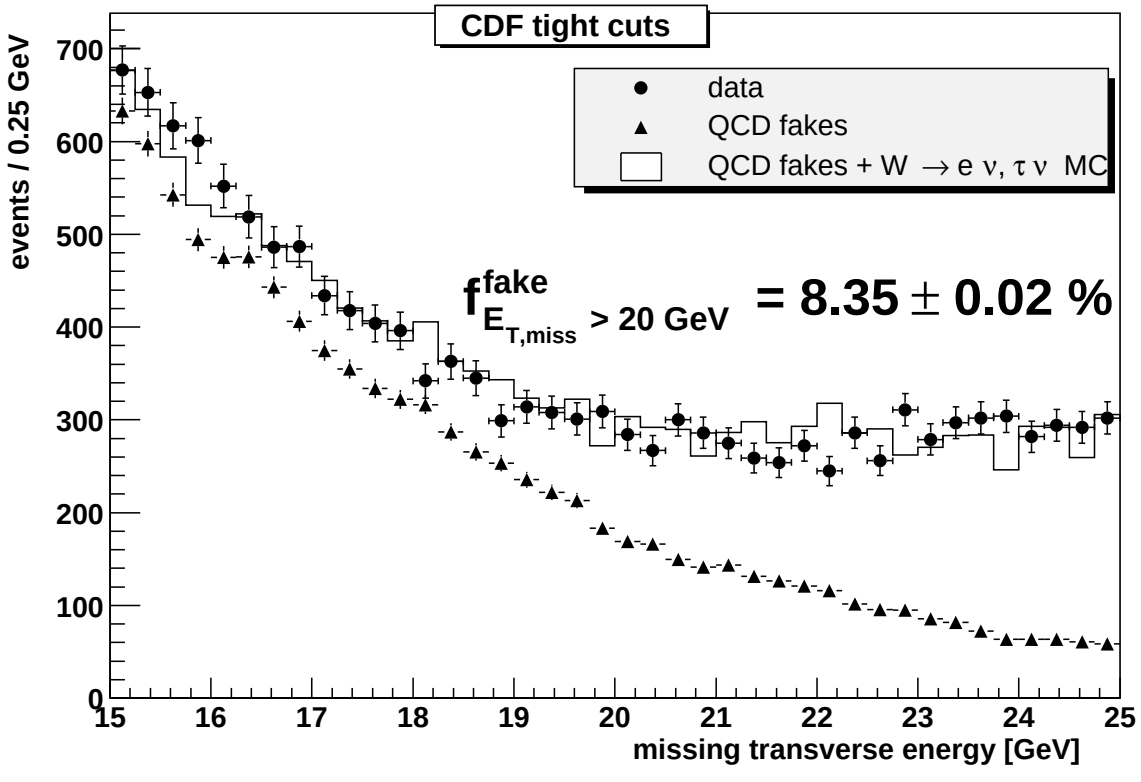


Figure 5.9: Comparison of data (dots) and Monte Carlo plus QCD background (histogram) for the tight cut analysis. The fraction of QCD background (triangles) was fitted using this distribution.

Table 5.4 lists the QCD background fractions for the three different electron selections. The statistical errors reported by the fitter were very low. However, systematic effects have to be taken into account. These may arise from different energy scales for data and Monte Carlo, bad description of the signal by Monte Carlo, contributions from other processes, or impurity of the background sample.

To estimate the size of the effects increasing the fake fraction, I assumed in a second method to determine the QCD background fraction, that all events in data with missing transverse energy between 15 and 15.5 GeV are from QCD background. This result was then extrapolated to the region with missing transverse energy above 20 GeV. In the table,

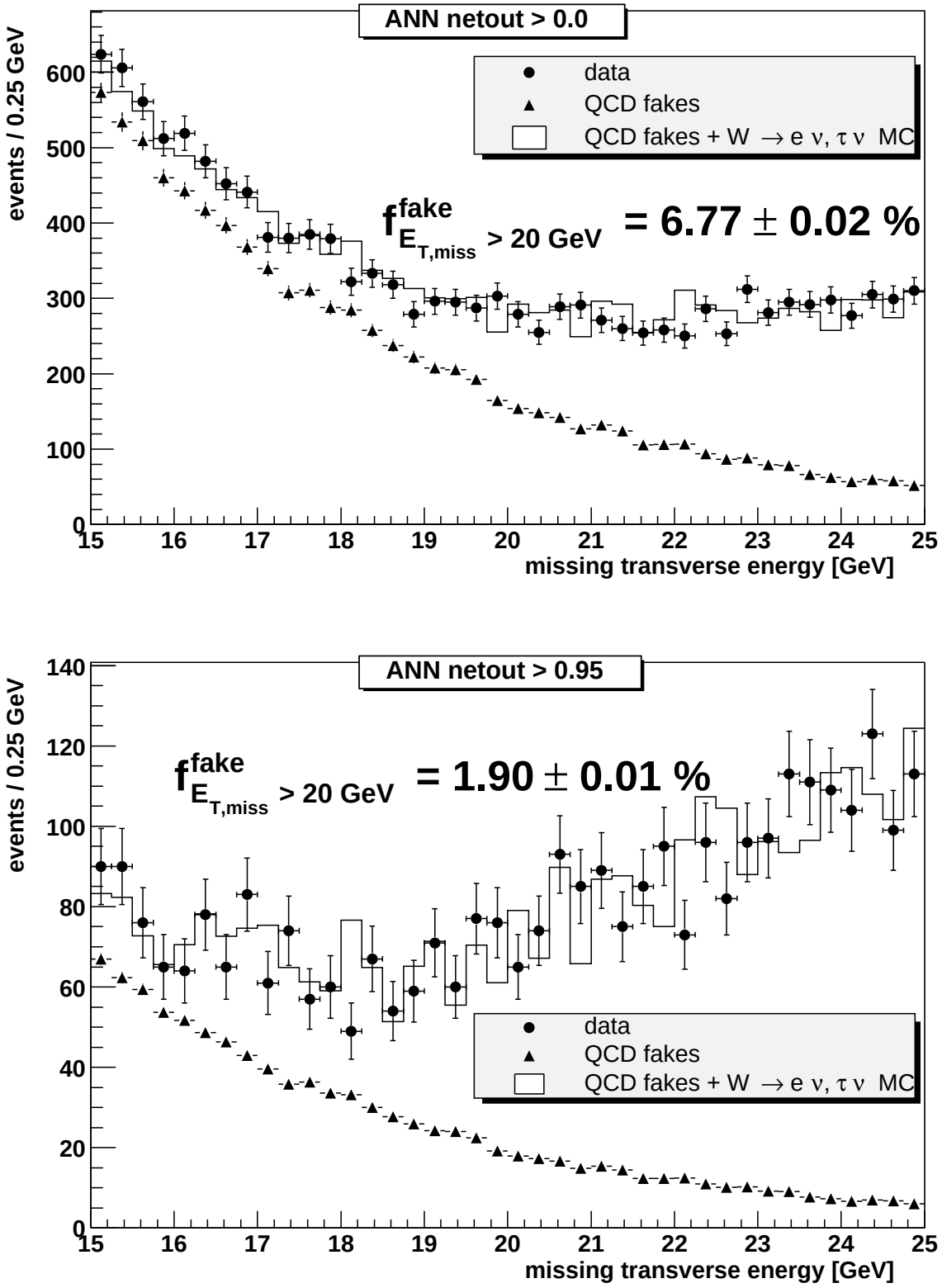


Figure 5.10: Comparison of data (dots) and Monte Carlo plus QCD background (histogram) for the ANN analyses. The fraction of QCD background (triangles) was fitted using this distribution. The upper plot is the result for the electron identification by requiring a net output above 0.0, for the lower plot the net output had to be above 0.95.

selection	all events	QCD events	QCD fraction	"simple method"
CDF tight cuts	40480	3380.1	$(8.35 \pm 0.02)\%$	$(9.02 \pm 0.30)\%$
ANN netout > 0.0	43067	2916.3	$(6.77 \pm 0.02)\%$	$(7.52 \pm 0.26)\%$
ANN netout > 0.95	17902	340.3	$(1.90 \pm 0.02)\%$	$(2.65 \pm 0.20)\%$

Table 5.4: Fitted QCD background fraction for the different electron selections.

I call this method "simple method". As the contributions from signal were neglected in this method, the values are always larger than the fit results and overestimate the QCD background fraction. The difference between the two methods was used to estimate the positive systematic error on the fit result. I included the statistical error on the results from the "simple method" into this calculation:

$$\delta f_{\mathrm{bkg}}^{\mathrm{QCD}} = \sqrt{(f_{\mathrm{simple}} - f_{\mathrm{fit}})^2 + \delta f_{\mathrm{simple}}^2}\,, \tag{5.2}$$

where $\delta f_{\mathrm{bkg}}^{\mathrm{QCD}}$ is the error on the fake fraction, f_{simple} is the estimated fake fraction from the "simple method", f_{fit} is the fitted fake fraction, and $\delta f_{\mathrm{simple}}$ is the statistical error on the result from the "simple method".

Another source of uncertainty is the purity of the QCD background samples. I tried to estimate the fraction of real W boson events in these samples. In case of the background sample for the CDF tight cuts analysis, it was estimated that the cuts used to select QCD background events will select around 2% of the signal events, see table 4.5. The efficiency for the CDF tight cuts for signal was 87.8% and 40480 events were accepted for the analysis sample from the preselected sample. Neglecting the background, I could derive an upper bound on the number of signal events in the QCD background of $\frac{40480}{0.878} \cdot 0.02 = 922$. As 10681 events were in the background sample, this would estimate that the impurity is below 8.6%. As stated in table 5.4, the fit suggested that 3380 background events are in the analysis sample. This number could be lower by 8.6% due to the impurity of the QCD background sample, that corresponds to a reduction by 292 events. This procedure was repeated for the two other analyses. However, the efficiency on signal events of the method used to select the QCD background for the Artificial Neural Net analyses was 0.5%, leading to smaller errors. Table 5.5 lists the positive and negative errors on the number of QCD events in the analysis samples. The statistical error on the "simple method" result was included into both overall errors.

selection	QCD events	$+\sigma$	bkg impurity	signal in bkg	$-\sigma$
CDF tight cuts	3380.1	297.2	$\leq 8.6\%$	291.7	315.9
ANN netout > 0.0	2916.3	341.9	$\leq 2.4\%$	69.1	131.6
ANN netout > 0.95	340.3	139.0	$\leq 2.3\%$	7.9	36.7

Table 5.5: Errors ($+\sigma$,$-\sigma$) on the estimated number of QCD events in the analysis samples.

5.5 The $W \to e\nu$ Cross-Section

To determine the cross-section, the following formula was used:

$$\sigma_W = \frac{N_{\text{obs}} - N_{\text{bkg}}^{\text{QCD}} - N_{\text{bkg}}^{\text{irreducible}}}{(\int \mathcal{L}dt) A_W \epsilon_e \epsilon_t} \tag{5.3}$$

N_{obs} is the number of observed events. $N_{\text{bkg}}^{\text{QCD}}$ is the number of expected background events in the signal sample. This number was obtained from a fit to the low region in the missing transverse energy distribution. The errors on this number can arise from bad Monte Carlo description of the data, other background sources, or impurities in the selected QCD background sample, please see section 5.4.2. $N_{\text{bkg}}^{\text{irreducible}}$ is the number of background events from $W \to \tau\nu$ and $Z \to e^+e^-$. Its size was evaluated using Monte Carlo as described in section 5.4.1. $\int \mathcal{L}dt$ is the integrated luminosity of the analyzed data. The integrated luminosities for each file set are listed in section 3.6.2. A_W is the acceptance and efficiency of the preselection cuts for $W \to e\nu$ Monte Carlo events. This was obtained in section 5.3. The error reflects the uncertainties in the generation and simulation of the signal events. ϵ_e is the efficiency of the electron identification after the preselection cuts have been applied. This number was derived in chapter 4. As mentioned in this chapter, a 1% systematic error was added to the statistical error of the efficiency determination to take impurities in the sample used for the efficiency determination into account. ϵ_t is the trigger efficiency. Its value is $(96.8 \pm 0.1)\%$ [103].

Table 5.6 lists the cross-section results for the three different analyses. The previous analysis using the tight electron cuts measured a cross-section of the W boson decaying into an electron and a neutrino of $(2.67 \pm 0.02_{\text{stat}} \pm 0.13_{\text{syst}} \pm 0.27_{\text{lum}})$ nb [101]. The main difference to the previous analysis is an improvement in the luminosity estimates reducing the error on the luminosity [93]. The updated error on the luminosity is 6.0 %. Our result using the tight cuts for the electron identification is $(2.74 \pm 0.02_{\text{stat}} \pm 0.12_{\text{syst}} \pm 0.16_{\text{lum}})$ nb. This agrees well with the previous measurement and the theoretical prediction of 2.73 nb [31, 32]. Estimates of the error on this prediction range from 3 to 5 %. Thus, a better luminosity estimation and a lower systematical error are needed to test

variable	electron identification		
	CDF tight cuts	ANN netout > 0.0	ANN netout > 0.95
N_{obs}	40480 ± 202	43067 ± 208	17902 ± 134
$N_{\text{bkg}}^{\text{QCD}}$	3380^{+297}_{-316}	2916^{+342}_{-132}	340^{+139}_{-37}
$N_{\text{bkg}}^{\text{irreducible}}$	$(4.83 \pm 0.20)\% \cdot (N_{\text{obs}} - N_{\text{bkg}}^{\text{QCD}})$		
$\int \mathcal{L} dt$	(55.5 ± 3.3) pb^{-1}		
A_W	$(27.31 \pm 1.06)\%$		
ϵ_e	$(87.8 \pm 1.3)\%$	$(94.4 \pm 1.0)\%$	$(40.1 \pm 1.1)\%$
ϵ_t	$(96.8 \pm 0.1)\%$		
$\sigma_W \cdot BR(W \to e\nu)$	2741 pb	2759 pb	2841 pb
$\Delta_\sigma^{\text{stat}}$	15 pb	14 pb	22 pb
$\Delta_\sigma^{\text{sys}}$	123 pb	119 pb	142 pb
$\Delta_\sigma^{\text{lum}}$	163 pb	164 pb	169 pb

Table 5.6: Values that enter the $W \to e\nu$ cross section calculation for the different electron identifications.

these calculations. The Run I result extrapolated to the Run II center-of-mass energy is (2.72 ± 0.13) nb [104, 101].

The analysis performed by selecting electrons with a cut at 0.0 on the net output of the electron net measured a cross-section of $(2.76 \pm 0.01 \pm 0.12 \pm 0.16)$ nb. It has slightly lower statistical and systematical errors. This reflects the enhanced electron identification due to the Artificial Neural Net.

The result obtained for cutting at 0.95 on the net output is higher than the other two measured cross-sections. As the three analysis share many inputs in the cross-section determination, the results can only differ due to differences in the expected QCD background and the assumed electron identification efficiency. However, even within the systematic errors of these variables the results agree.

The errors discussed in the previous sections were used to calculate the systematic error on these measurements. The systematic errors in these measurements are of similar size than the systematic error in the previous measurement. The main contributions come from the acceptance and efficiency of my preselection, the determination of the QCD background and the electron identification efficiency. The different sources of systematic errors are listed in table 5.7. Besides A_W and $N_{\text{bkg}}^{\text{irreducible}}$, all errors were assumed to be uncorrelated.

variable	electron identification		
	CDF tight cuts	ANN netout > 0.0	ANN netout > 0.95
$N_{\text{bkg}}^{\text{QCD}}$	$^{+23}_{-22}$ pb	$^{+9}_{-23}$ pb	$^{+6}_{-23}$ pb
$N_{\text{bkg}}^{\text{irreducible}}$	± 6 pb	± 6 pb	± 6 pb
A_W	± 108 pb	± 107 pb	± 110 pb
ϵ_e	± 41 pb	± 29 pb	± 78 pb
ϵ_t	± 3 pb	± 3 pb	± 3 pb
combined error	$^{+123}_{-122}$ pb	$^{+117}_{-119}$ pb	$^{+140}_{-142}$ pb

Table 5.7: Systematic errors of the $W \to e\nu$ cross-section measurements. For the combination of the individual systematic errors I assumed all variables except A_W and $N_{\text{bkg}}^{\text{irreducible}}$ to be uncorrelated.

Chapter 6

Conclusions

In this thesis, I have studied W boson production and decay with the *Collider Detector at Fermilab*, CDF, in proton-antiproton collisions with a center-of-mass energy of 1.96 TeV. I have used the first (55.5 ± 3.3) pb^{-1} of data collected since the start of Run II in 2001, which represents all good runs with silicon tracking information up to January 2003.

I limited my study of W bosons to the decay channel $W \to e\nu$. A good electron identification is crucial to disentangle the events from this electroweak process from the large number of QCD events. I reevaluated the efficiency and purity of the standard CDF electron identification using tight cuts on a signal and background sample derived from data. Furthermore, I used these data samples to train an Artificial Neural Net for electron identification. By offering a similar efficiency, a better background rejection could be achieved with the Artificial Neural Net. This is an example that Artificial Neural Nets can successfully be used at a hadron collider, if one uses data instead of Monte Carlo for the training. The large cross-sections of processes via the strong interaction make it merely impossible to generate a sufficient amount of QCD events to simulate the background from QCD correctly.

Using the standard tight cuts and two different cuts on the net output of the Artificial Neural Net, I measured the W boson cross-section in three analyses. My result for the tight cuts analysis is $(2.74 \pm 0.02_{\text{stat}} \pm 0.12_{\text{syst}} \pm 0.16_{\text{lum}})$ nb and the result for an analysis cutting on the net output is $(2.76 \pm 0.01_{\text{stat}} \pm 0.12_{\text{syst}} \pm 0.16_{\text{lum}})$ nb. The latter has a better statistical error due to the improved electron identification of the Artificial Neural Net. These results are in good agreement with the theoretical predictions of 2.73 nb and the previous Run II measurement of $(2.67 \pm 0.02_{\text{stat}} \pm 0.13_{\text{syst}} \pm 0.27_{\text{lum}})$ nb. To estimate the amount of background from fake electrons in the data samples, I created a background sample by selecting events with an electron candidate that has a small electron probability. This sample and the signal Monte Carlo were fitted to the data in order to obtain the background fraction of each data sample. This method allowed an estimate of the shape

of the background for further analyses, e.g. measurements of the W boson mass, its transverse momentum, or the properties of the top quark.

In the coming years, the CDF experiment will continue to take data. This will largely increase the statistics for physics measurements. However, this cross-section measurement is already limited by systematics. Thus, lowering the systematic uncertainties will be of growing importance. The electron identification based on Artificial Neural Nets allows to select very clean data samples reducing systematic uncertainties due to background events. With more data, one can use $Z \to e^+e^-$ to select a signal sample for the training. This will result in a purer signal sample and improve the efficiency determination. The use of an Artificial Neural Net might be even more interesting for the identification of particles that are more difficult to detect, like τ leptons, that have gained scant attention in Run I.

Appendix A

Transverse Mass and Transverse Momentum Distribution of the Reconstructed W Bosons

I will now study the kinematics of the reconstructed W bosons. For all plots, the QCD background was subtracted from data according to its fitted contribution and the result was compared with the signal Monte Carlo generated with Pythia. This demonstrates an advantage of my method of estimating the QCD background, as this method allows to determine the shape of the background as well as its size.

A.1 Transverse Mass of the W Bosons

Figures A.1 and A.2 show the transverse mass of the W bosons for the different analyses. The expected Jacobian peak at the W boson mass value is slightly shifted for data towards lower values. This might be due to missing corrections for the transverse missing energy or a wrong calorimeter scale. Furthermore, the Jacobian peak is narrower for Monte Carlo than for data. On one hand, this might indicate a problem with the detector, its calibration, or the reconstruction code, on the other hand, this might as well reconcile a problem in the detector simulation.

There are no sizeable differences in the shape of the transverse mass distributions for the three different analyses. However, one can see the very small amount of QCD background in the analysis using the tighter cut on the net output.

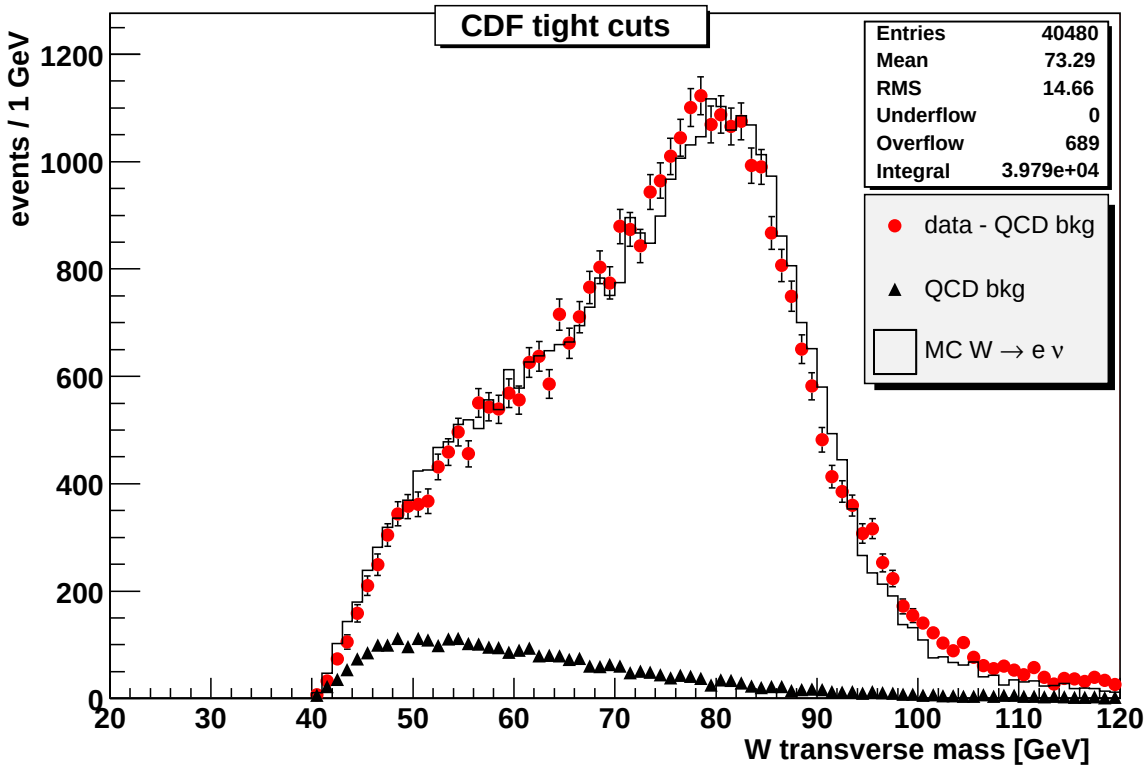


Figure A.1: Transverse mass of the reconstructed W bosons using the tight cut electron identification. The data events (dots) are corrected for contributions from QCD background (triangles) and compared with Monte Carlo $W \to e\nu$ events (histogram).

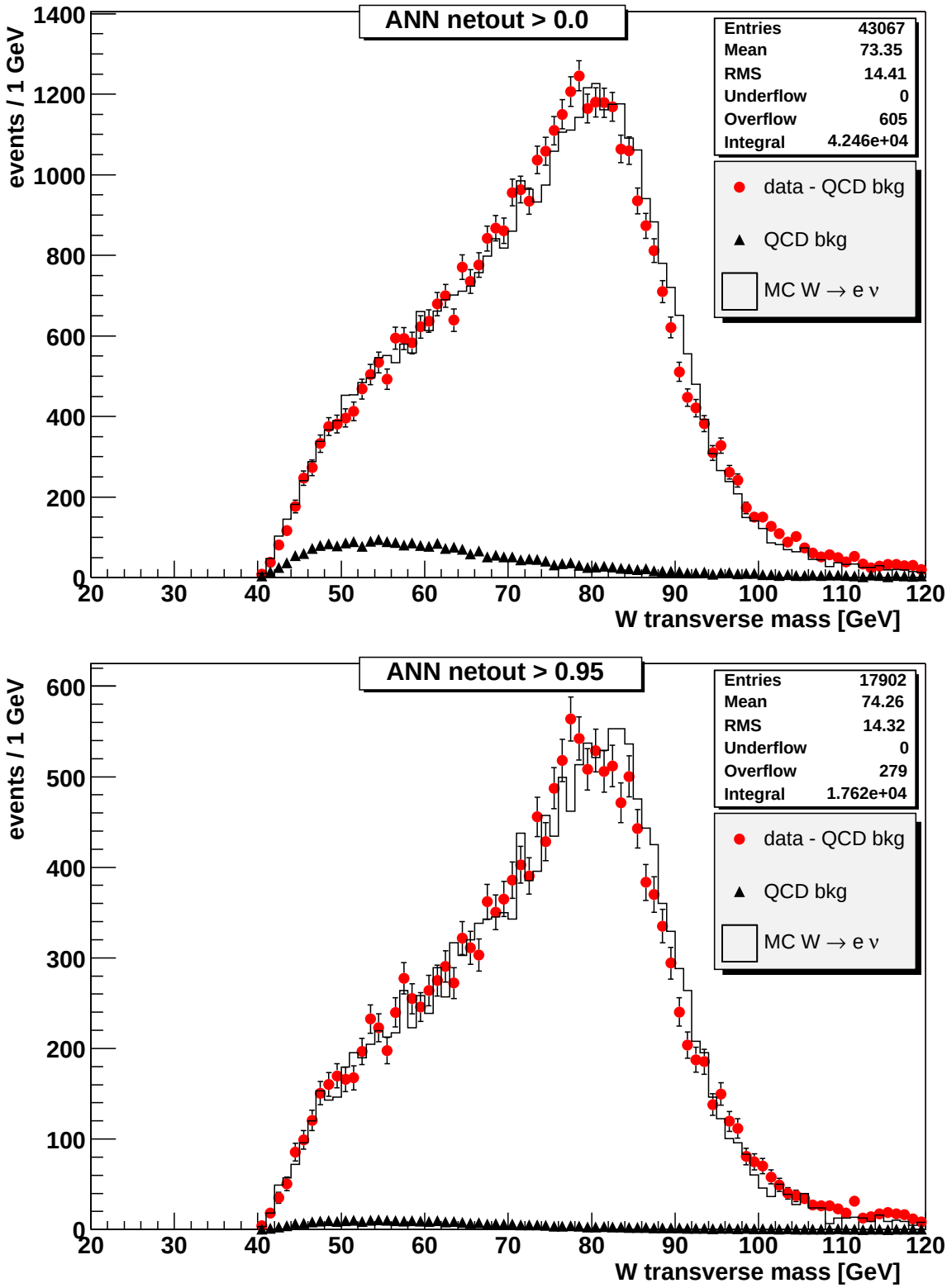


Figure A.2: Transverse mass of the reconstructed W bosons using the Artifical Neural Net for the electron identification. The data events (dots) are corrected for contributions from QCD background (triangles) and compared with Monte Carlo $W \rightarrow e\nu$ events (histogram).

A.2 Transverse Momentum of the W Bosons

The distribution of the transverse momentum of the W boson offers a good test for the different resummation methods used in theoretical calculations of the differential W boson cross-section. Figures A.3 and A.4 show the transverse momentum distribution of the reconstructed W bosons for the three different analyses. However, I have not corrected these distributions for the acceptance. Thus, a direct comparison with theory is impossible. The contributions from QCD background are nearly negligible for the third analysis, where the electron candidate was required to have a net output above 0.95. Furthermore, the peak of its distribution is shifted towards lower values than for the other two analyses.

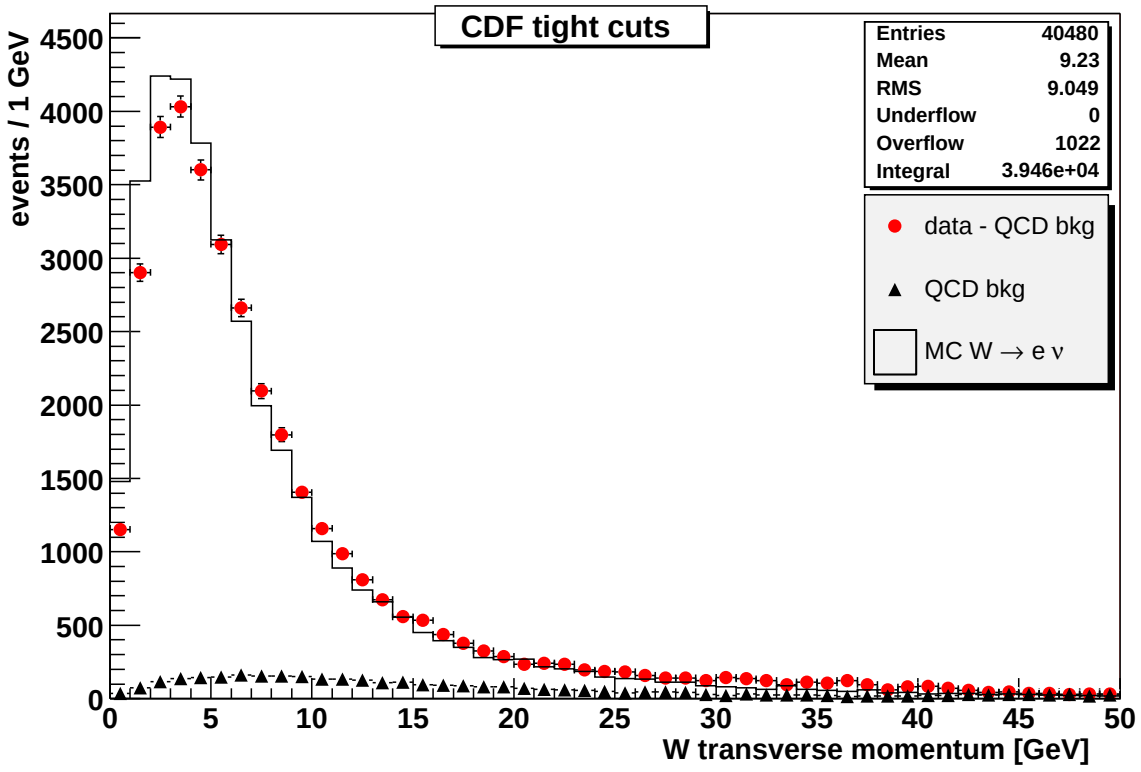


Figure A.3: Transverse momentum of the reconstructed W bosons using the tight cut electron identification. The data events (dots) are corrected for contributions from QCD background (triangles) and compared with Monte Carlo $W \to e\nu$ events (histogram) after the 5% correction, see section 5.2.

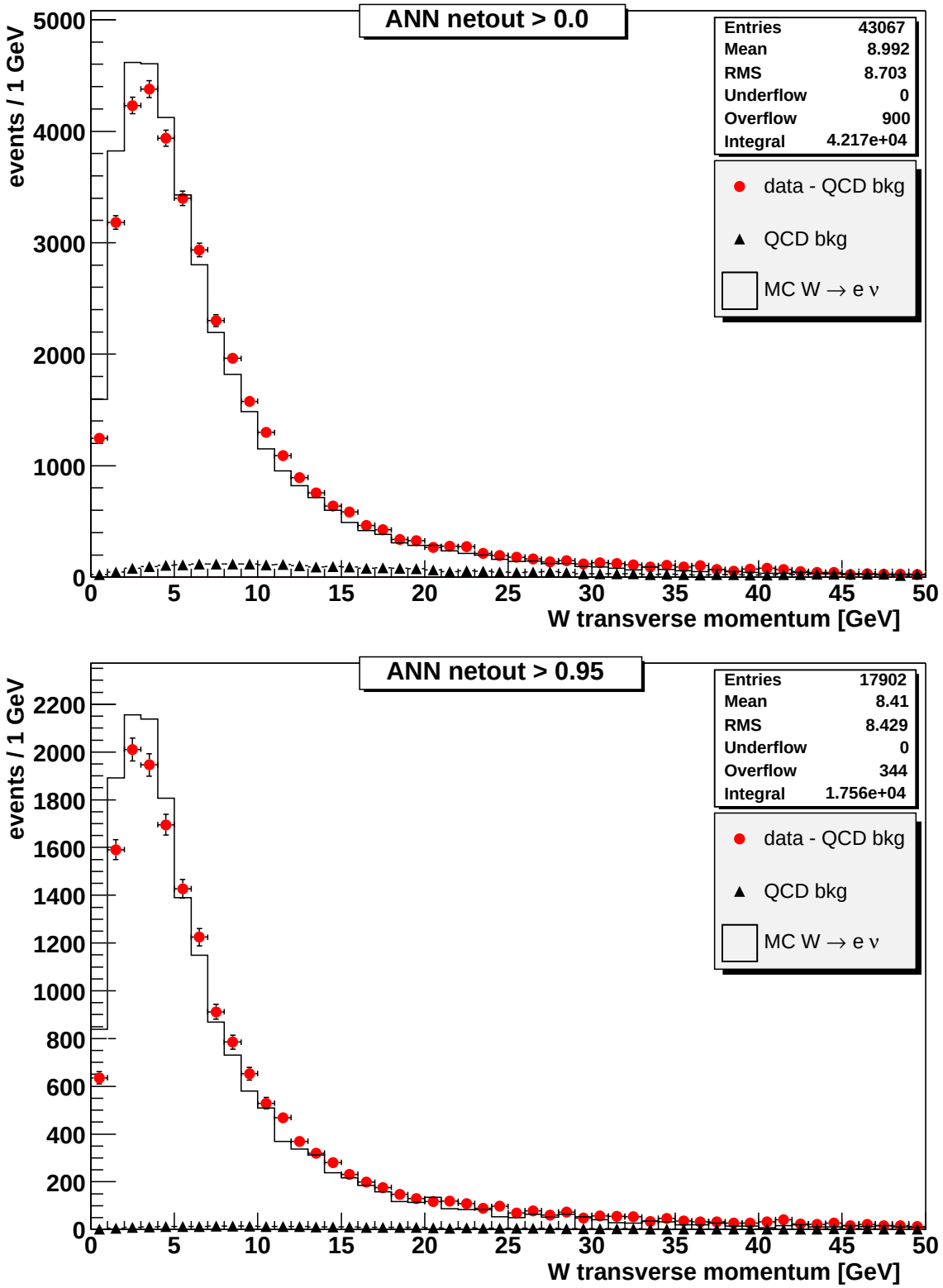


Figure A.4: Transverse momentum of the reconstructed W bosons using the Artifical Neural Net for the electron identification. The data events (dots) are corrected for contributions from QCD background (triangles) and compared with Monte Carlo $W \to e\nu$ events (histogram) after the 5% correction, see section 5.2.

Appendix B

W Boson Events with Jets

W bosons events with jets are the starting point of many physics analyses, especially in top physics. In this appendix, I will study the jet multiplicity in these events and the contributions from QCD background to these data samples. The missing transverse energy resolution significantly decreases for higher jet activity. Therefore, the fraction of QCD background events shall rise with the jet multiplicity.

The jets were reconstructed with the *JetClu* algorithm introduced in section 3.4 with a cone size of $R = 0.4$. The jet directions were recalculated for the primary vertex position of the event. All jets were corrected with the jet corrections described in section 3.4.2 up to "level 7". Finally, the following cuts were applied on the corrected transverse energy and the pseudorapidity:

- $E_T > 15$ GeV
- $|\eta| < 2.4$

These are the same cuts as used in the Run I analysis [105]. The accepted jets were ordered by their corrected energy.

In my analyses, I have not corrected the missing transverse energy for changes in the jet energies from the jet corrections. As the calorimeter response is quite different for different detector regions, the corrections should reduce the amount of QCD fakes significantly. Future analyses at CDF will use an improved missing transverse energy, thus it will be instructive to repeat this study for the corrected missing transverse energy.

For all plots presented in this appendix, QCD background was subtracted from the data. The amount of QCD fake events in data was determined in section 5.4.2 by a fit to the lower part of the missing transverse energy spectrum, where QCD background should dominate.

B.1 The Jet Multiplicity in W Boson Events identified using the CDF Tight Cuts

Firstly, I will study the jet multiplicity for W boson events identified using the tight cuts for the electron identification. Table B.1 summarizes the number of events in the data and the background sample for the different inclusive numbers of jets. The background sample consisted of events failing the very loose electron identification cuts defined in section 4.7.

Number of jets	data	QCD bkg	data - QCD bkg	$\sigma_{W\to e\nu}$
≥ 0	40480 ± 201	3380 ± 318	37100 ± 379	(2741 ± 28) pb
≥ 1	10822 ± 104	2997 ± 90	7825 ± 138	(578 ± 10) pb
≥ 2	3241 ± 57	1541 ± 34	1700 ± 66	(125.6 ± 4.9) pb
≥ 3	914 ± 30	584 ± 15	330 ± 34	(24.4 ± 2.5) pb
≥ 4	236 ± 15.4	189.2 ± 8.0	46.8 ± 17.3	(3.5 ± 1.3) pb
≥ 5	59 ± 7.7	54.7 ± 4.2	4.3 ± 8.7	(0.3 ± 0.6) pb

Table B.1: Jet multiplicity for the tight cut analysis. The second column lists the number of events in data for the different inclusive jet multiplicities. The amount of QCD events in these sub samples is listed in the third column and the difference between these two numbers in the fourth. The cross-section times branching ratio that corresponds to this number of events is shown in the last column.

I assumed that the number of data events after the QCD background subtraction for $N_{\text{jets}} = 0$ corresponds to the cross-section times branching ratio measurement in chapter 5. The cross-sections of the other inclusive jet multiplicities were calculated by simply scaling the cross-section from chapter 5 by the ratio $\frac{N^{\text{data}}_{\geq n \text{ jets}} - N^{\text{fake}}_{\geq n \text{ jets}}}{N^{\text{data}}_{\geq 0 \text{ jets}} - N^{\text{fake}}_{\geq 0 \text{ jets}}}$. As a result, the cross-sections stated in the table were calculated without taking any jet inefficiencies into account, although jets with a distance from the electron in the $\eta - \varphi$ space below 0.4 had been removed. However, the Run I analysis has shown that the acceptance and efficiency has only a small dependence on the jet multiplicity [105]. The stated errors are just the statistical error on the number of data events and the statistical and systematic error on the number of QCD events. I have omitted all other systematic errors, e.g. the errors on the acceptance or the electron identification efficiency, compare chapter 5.

When at least two jets are required, every second event in data is already due to the strong interaction. Starting with four jets, the QCD background completely dominates

the sub samples. Our cross-sections are a little bit higher than the Run I measurements, which can be attributed to the rise in the center-of-mass energy from 1.8 TeV to 1.96 TeV.

Figure B.1 compares the results listed in table B.1 with the Pythia Monte Carlo prediction. One can see the expected exponential fall of the cross-section for higher inclusive jet multiplicities for data and Monte Carlo. As expected, Pythia underestimates the jet multiplicity, as the gluons and quarks generated with the W boson tend to be less energetic than in real data.

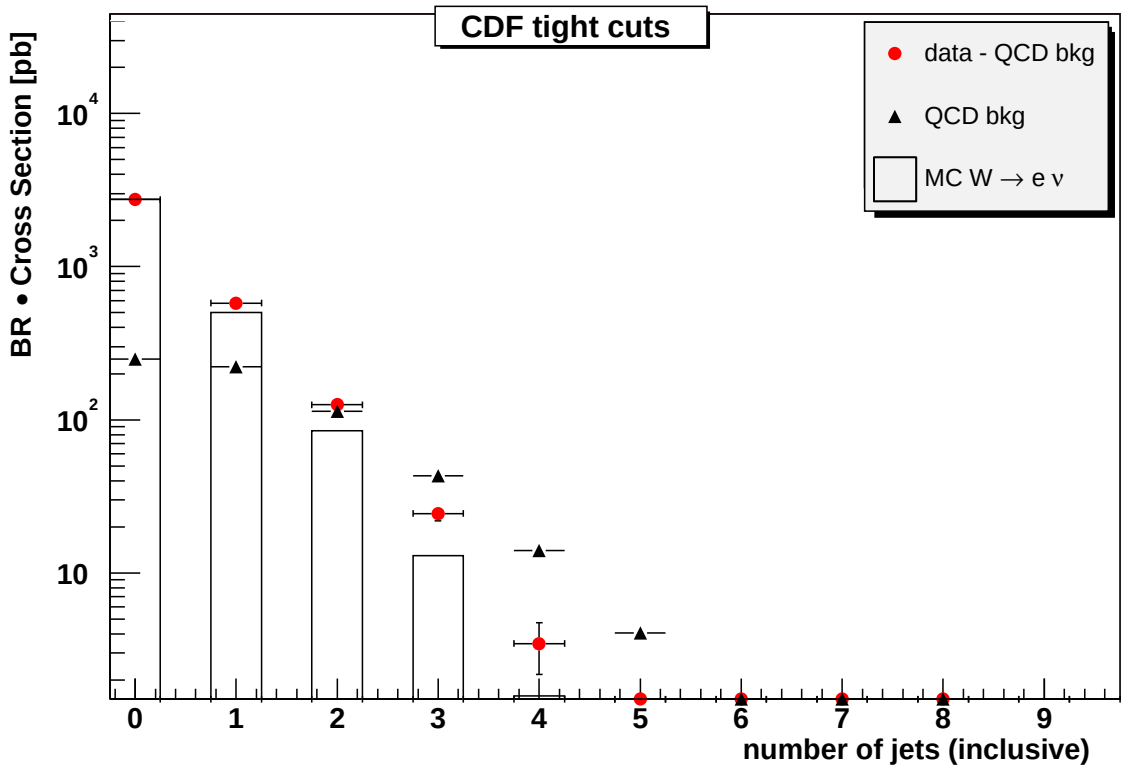


Figure B.1: $W+ \geq n$ jets cross-section for the tight cut analysis. The triangles indicate the distribution for the QCD background. The data after subtraction of the background (dots) favors higher jet multiplicities than the $W \rightarrow e\nu$ Monte Carlo generated with Pythia.

Figures B.2 and B.3 show the transverse energy and pseudorapidity distributions for the first four leading jets. Once again, one can see the large contributions from QCD fake events for higher jet multiplicities. The pseudorapidity spectrum of the leading jet from background shows two peaks at the border of the central and the plug calorimeters. The data corrected for the background looks fine in this respect. Therefore, this strange peaks

seem to be a feature of the background. The tranverse energy spectrum of the leading jet shows a steeper fall for the corrected data than for the background. The plots for the second, third and forth jet have too low statistics to identify any differences between the assumed signal and the background.

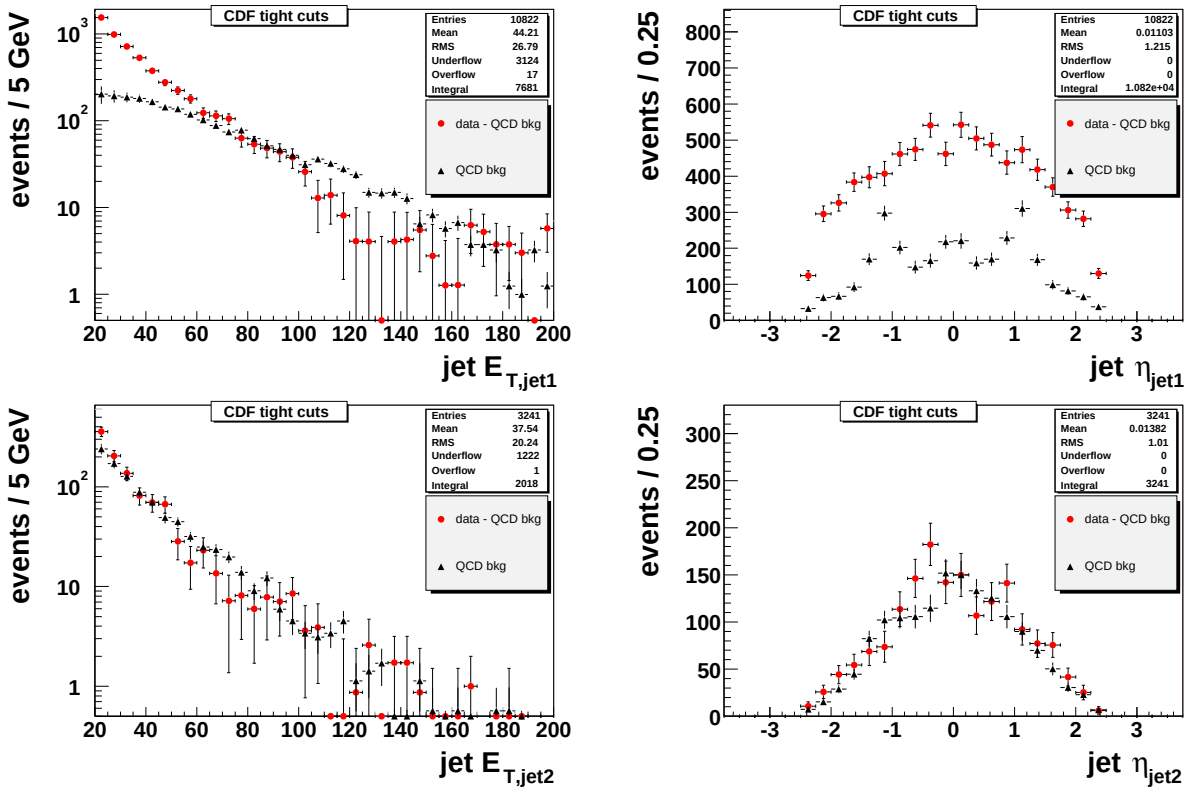


Figure B.2: Transverse energy and pseudorapidity distribution for the first (top) and second (bottom) leading jet for the tight cut analysis.

The invariant mass of the two leading jets can be seen in figure B.4. The rise at lower masses is a threshold effect due to the 15 GeV transverse energy requirement for both jets. For values above this threshold, the distribution shows the expected steep fall. The Monte Carlo was normalized to the number of two or more jets events in data after the background subtraction and agrees well with the data.

The separation ΔR of the two leading jets is shown in figure B.5. Due to the used cone algorithm, the jets have a minimal separation of $R = 0.4$. Quarks from gluon splitting tend to have a rather small opening angle. This explains the peak for small separations. The other peak is at π, meaning back-to-back jets.

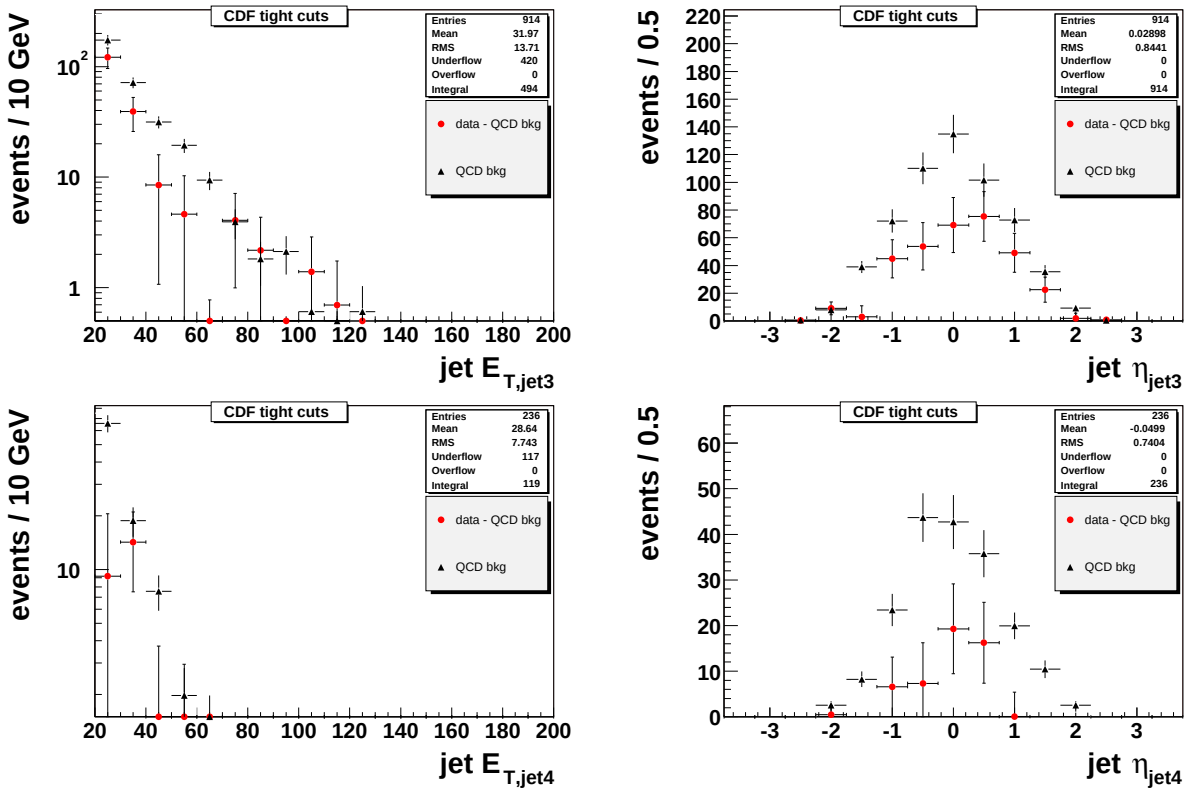


Figure B.3: Transverse energy and pseudorapidity distribution for the third (top) and forth (bottom) leading jet for the tight cut analysis.

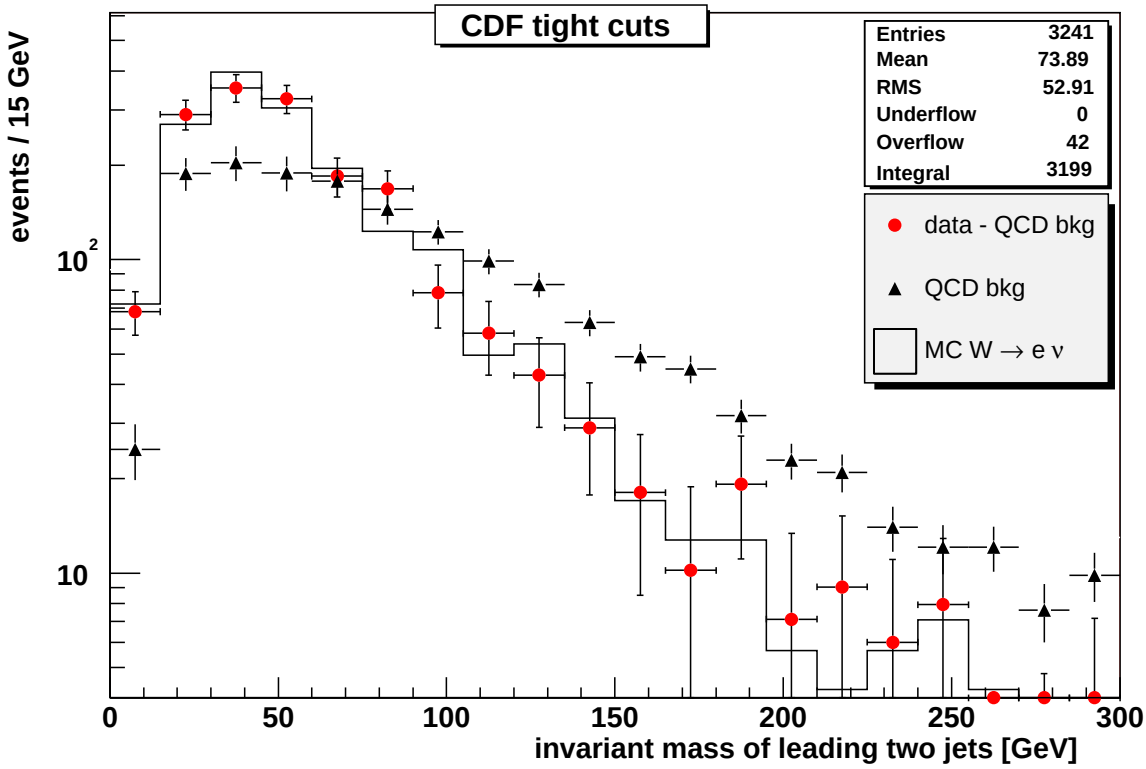


Figure B.4: Invariant mass of the leading two jets for the tight cuts analysis.

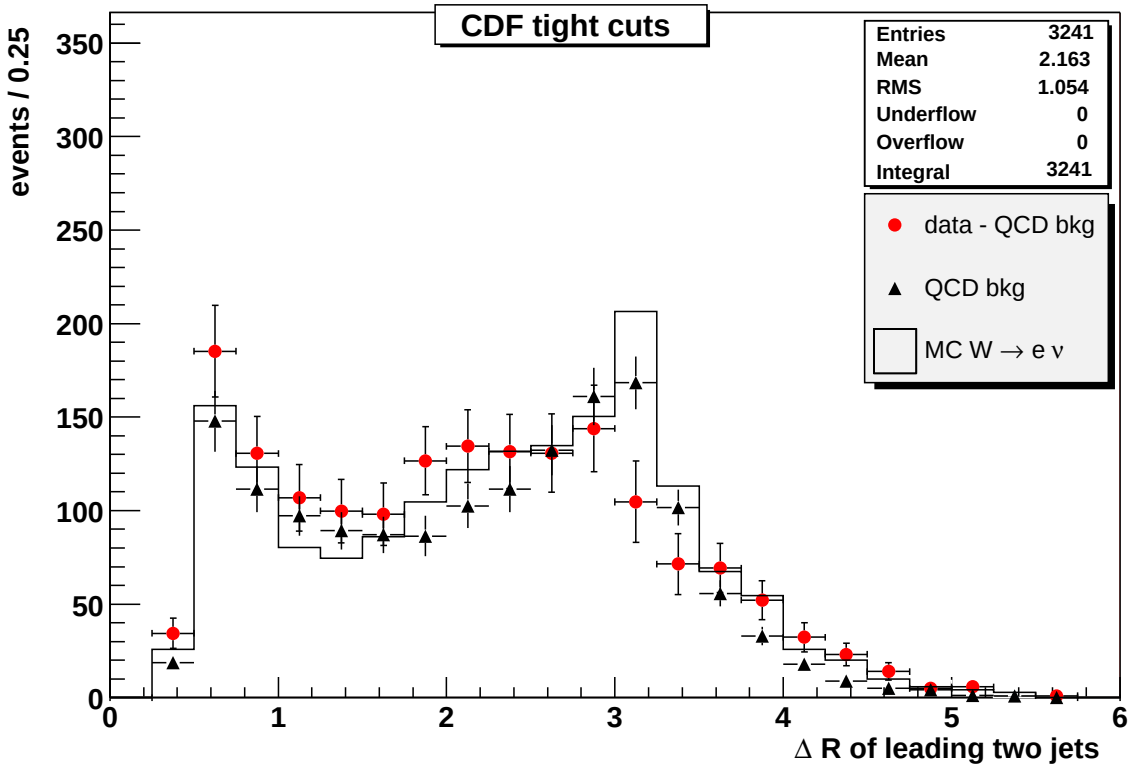


Figure B.5: Jet-jet separation of the leading two jets for the tight cuts analysis. ΔR is the distance in the η-φ plane.

B.2 The Jet Multiplicity in W Boson Events identified using the Artificial Neural Net

The electron identification using an Artificial Neural Net selected less QCD background events than the tight electron cuts used in the previous section. However, the net might be more stringent on the isolation requirement, thus removing events with a high jet multiplicity. To check for this, the jet multiplicity analysis was repeated for the two electron identifications based on an Artificial Neural Net. The background sample was selected with the net as well. For this, events were chosen that have an electron candidate with a very low net output, i.e. a very low probability of being a real electron.

Table B.2 summarizes the results for the looser electron identification. The results agree well with the tight cut analysis. The QCD background still plays an important role for higher jet multiplicities. The values listed in table B.2 were obtained by the same methods explained in the previous section. The same results can be seen in figure B.6.

Number of jets	data	QCD bkg	data - QCD bkg	$\sigma_{W\to e\nu}$
≥ 0	43067 ± 208	2916 ± 343	40151 ± 401	(2759 ± 28) pb
≥ 1	11097 ± 105	2722 ± 93	8375 ± 140	(575 ± 10) pb
≥ 2	3159 ± 56	1436 ± 33	1723 ± 65	(118.4 ± 4.5) pb
≥ 3	863 ± 29	553 ± 15	310 ± 33	(21.3 ± 2.3) pb
≥ 4	222 ± 14.9	181.0 ± 7.6	41.0 ± 16.7	(2.8 ± 1.1) pb
≥ 5	62 ± 7.9	53.6 ± 4.1	8.4 ± 8.9	(0.6 ± 0.6) pb

Table B.2: Jet multiplicity for the analysis requiring an ANN net output above 0.0. The second column lists the number of events in data for the different inclusive jet multiplicities. The amount of QCD events in these sub samples is listed in the third column and the difference between these two numbers in the fourth. The cross-section that corresponds to this number of events is shown in the last column.

For completeness, I show the invariant mass and jet-jet separation distributions for this analysis as well. The distributions in figure B.7 and B.8 are very similar to the same distributions for the tight cut analysis.

The analysis performed with a very tight cut on the output of the Artificial Neural Net shows a much smaller QCD background fraction for data. Table B.3 and figure B.9 summarize the results for this analysis.

The fraction of QCD background events is much lower than for the other two analyses. However, this tight selection does not seem to lower the jet multiplicity. In fact, the

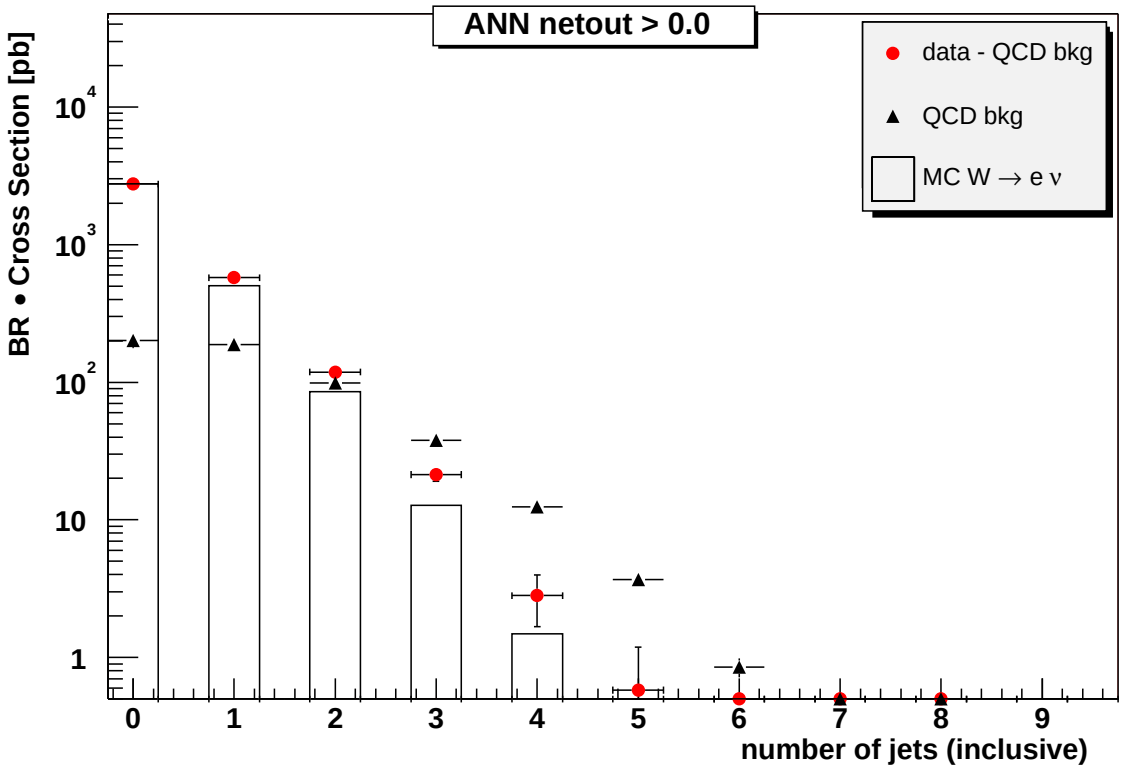


Figure B.6: $W+ \geq n$ jets cross-section for the analysis requiring an ANN net output above 0.0.

Number of jets	data	QCD bkg	data - QCD bkg	$\sigma_{W\to e\nu}$
≥ 0	17902 ± 134	340 ± 139	17562 ± 193	(2841 ± 31) pb
≥ 1	4085 ± 64	318 ± 32	3767 ± 71	(609 ± 11) pb
≥ 2	1024 ± 32	168 ± 8	856 ± 33	(138.5 ± 5.3) pb
≥ 3	245 ± 15.7	64.5 ± 2.4	180.5 ± 15.8	(29.2 ± 2.6) pb
≥ 4	61 ± 7.8	21.1 ± 1.0	39.9 ± 7.9	(6.5 ± 1.3) pb
≥ 5	9 ± 3.0	6.3 ± 0.5	2.7 ± 3.0	(0.4 ± 0.5) pb

Table B.3: Jet multiplicity for the analysis requiring an ANN net output above 0.95. The second column lists the number of events in data for the different inclusive jet multiplicities. The amount of QCD events in these sub samples is listed in the third column and the difference between these two numbers in the fourth. The cross-section that corresponds to this number of events is shown in the last column.

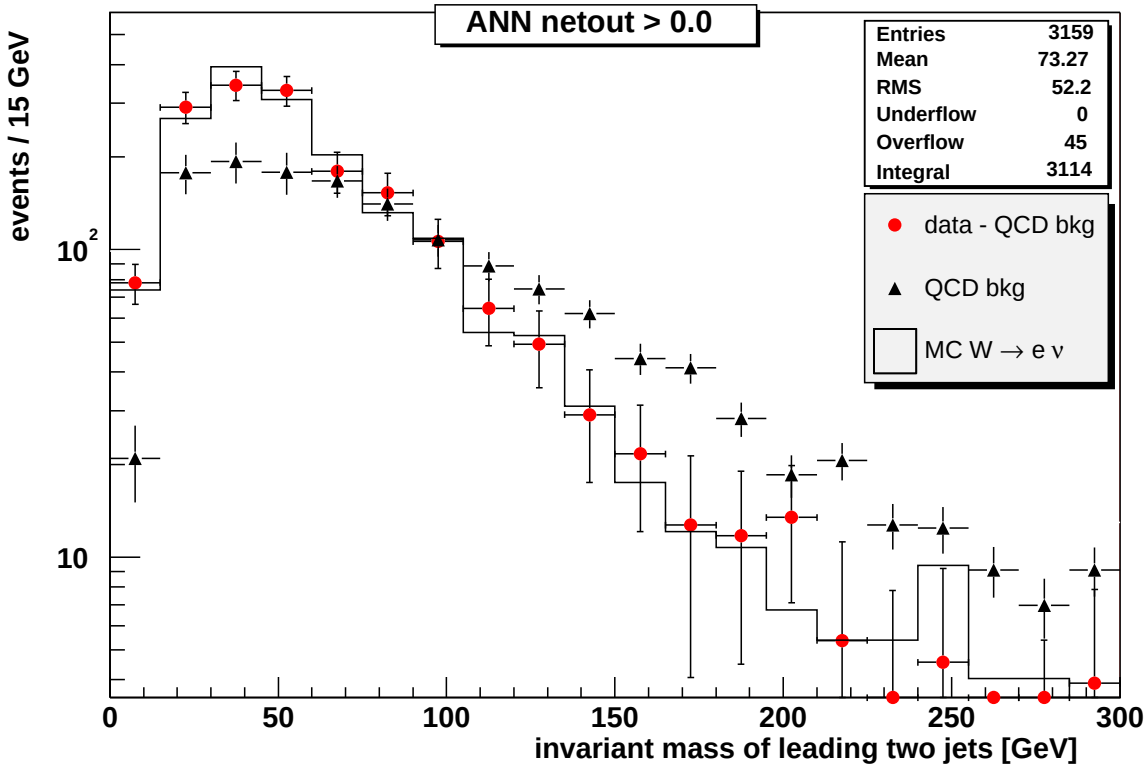


Figure B.7: Invariant mass of the leading two jets for the analysis requiring an ANN net output above 0.0.

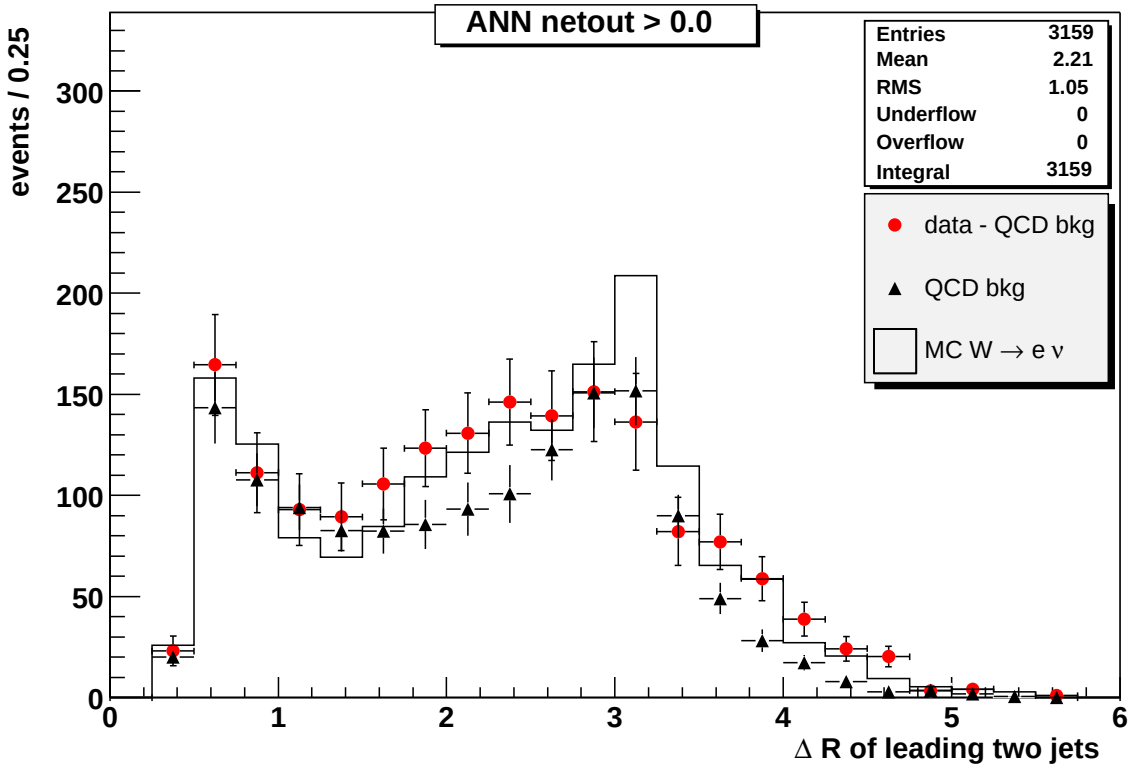


Figure B.8: Jet-jet separation of the leading two jets for the analysis requiring an ANN output above 0.0. ΔR is the distance in the η-φ plane.

cross-sections are even higher, though this is at least partly due to the higher overall cross-section measured in this analysis.

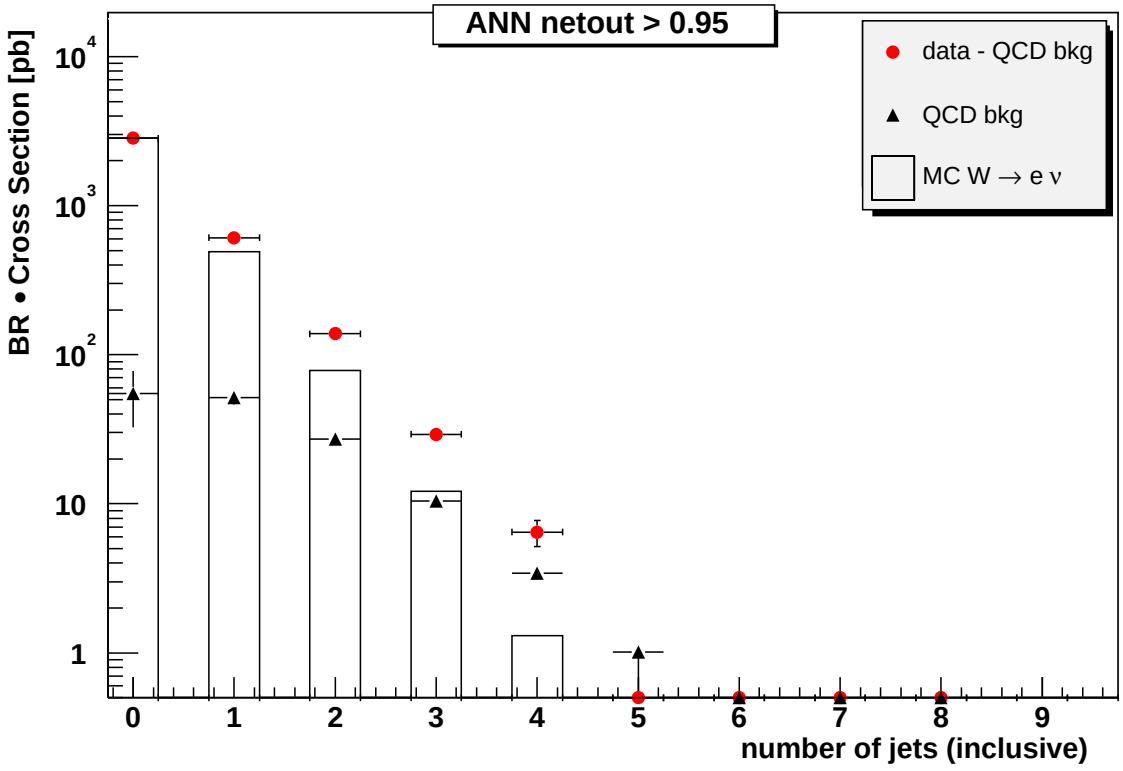


Figure B.9: $W+ \geq n$ jets cross-section for the analysis requiring an ANN net output above 0.95.

The distributions of the invariant mass and the separation of the two leading jets shown in the figures B.10 and B.11 are similar to the distributions for the other two analyses.

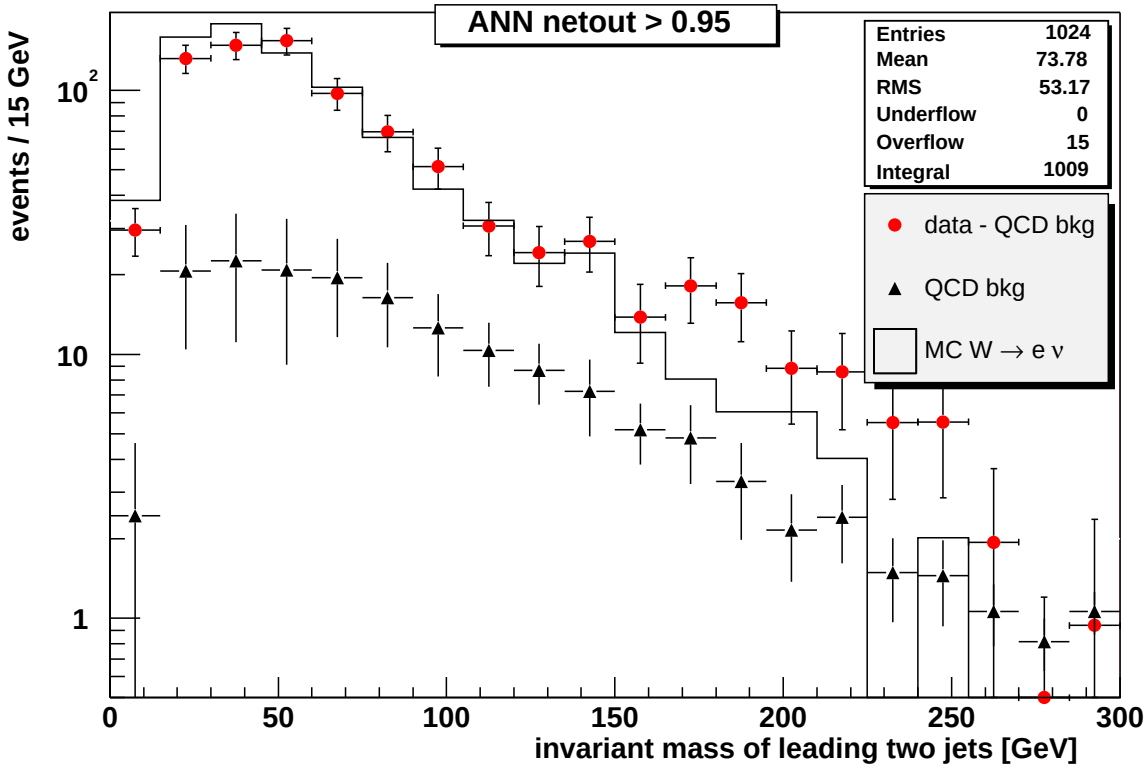


Figure B.10: Invariant mass of the leading two jets for the analysis requiring an ANN net output above 0.95.

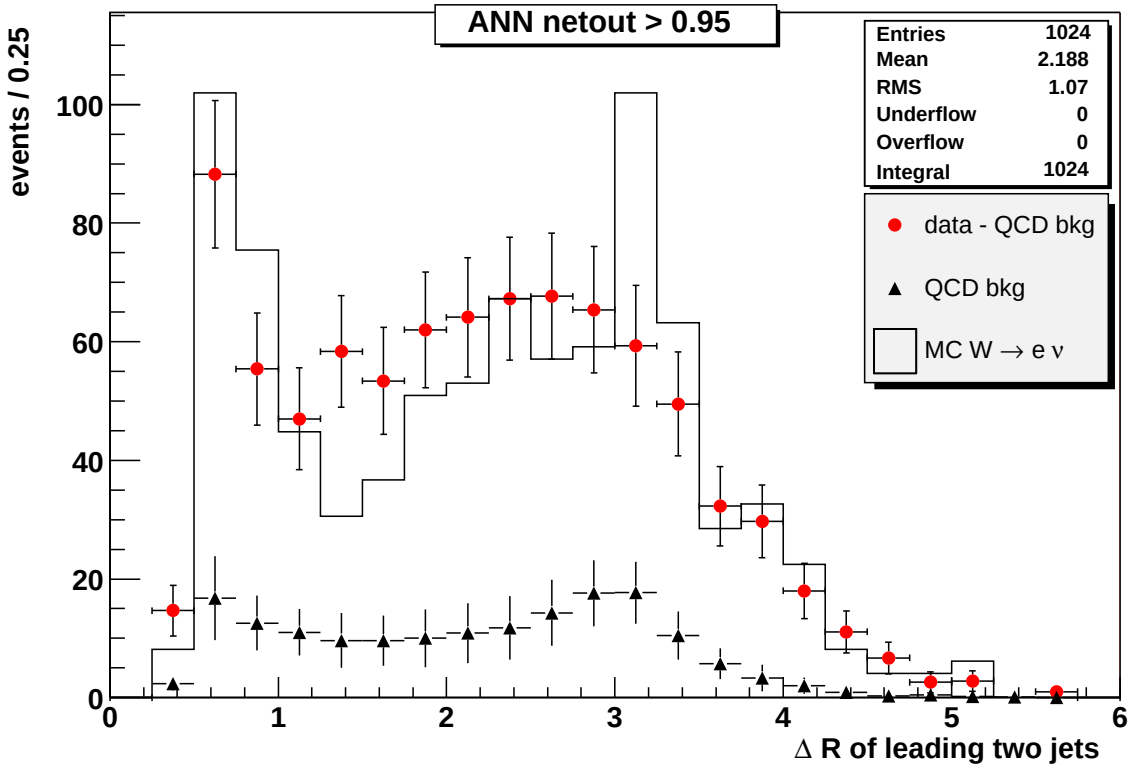


Figure B.11: Jet-jet separation of the leading two jets for the analysis requiring an ANN net output above 0.95. ΔR is the distance in the η-φ plane.

Appendix C

The Jet Multiplicity in tagged Events

I will now study the jet multiplicity in events that contain a jet that was tagged as a b-jet by the *SecVtx* algorithm, see 3.5. Tagged W boson events with four jets can be the final state of top quark pair production. Single top production can be searched for in tagged W boson events with at least two jets. Thus, the fraction of QCD background in the data and the overall cross-section for tagged W boson events is an important input in many physics analyses.

However, to measure the cross-section one has to correct the overall acceptance for the b-tagging efficiency of $(19.6 \pm 1.1)\%$ for a fiducial jet. The efficiency of tagging an event depends on the number of fiducial bottom jets, e.g. if two bottom jets are fiducial, the tag efficiency is 43.0% instead of 19.6%. Furthermore, for data one does not know, whether the second b-jet is fiducial, as long as it is not tagged. Therefore, a thorough Monte Carlo study is needed to determine the overall acceptance for each jet bin. This requires a very good simulation of the silicon detectors. As this is still work in progress, I have not tried to determine the acceptance in order to measure the cross-sections.

Still, the fraction of QCD background events is a very important input for physics analyses based on tagged W boson plus jets events. The number of data events, QCD events, and the background fraction are listed in table C.1 for the analysis using the tight cuts for the electron identification. The number for the QCD background events was obtained from the background sample selected in section 5.4.2. In contrast to the background fraction obtained for all events, see section B.1, this fraction decreases for tagged events with the number of jets. One reason might be an increase in the signal cross-section for higher jet multiplicities due to top pair production. However, the errors on the fake fractions are so large, that the results agree as well with a constant background fraction. Still, these numbers suggest that every second event in a $W + 2$ jets sample selected using the tight electron identification cuts is a QCD event.

Number of jets	data	QCD bkg	data - QCD bkg	QCD bkg fraction
≥ 1	95 ± 9.7	60.1 ± 7.1	34.9 ± 12.0	$(63 \pm 10)\%$
≥ 2	59 ± 7.7	38.9 ± 5.0	20.1 ± 9.1	$(66 \pm 12)\%$
≥ 3	32 ± 5.7	19.0 ± 3.1	13.0 ± 6.4	$(59 \pm 14)\%$
≥ 4	13 ± 3.6	6.6 ± 1.6	6.4 ± 4.0	$(51 \pm 19)\%$

Table C.1: Jet multiplicity in tagged events for the tight cut analysis.

The tagged jets from background tend to have a lower transverse energy than the jets from signal, see figure C.1. The pseudorapidity distribution of the tagged jets just reflects the acceptance of the *SecVtx* algorithm. Silicon tracks from the *stand-alone* algorithm were not used in the *SecVtx* algorithm. As the COT only covers the central region, the *SecVtx* efficiency declines when leaving the central region.

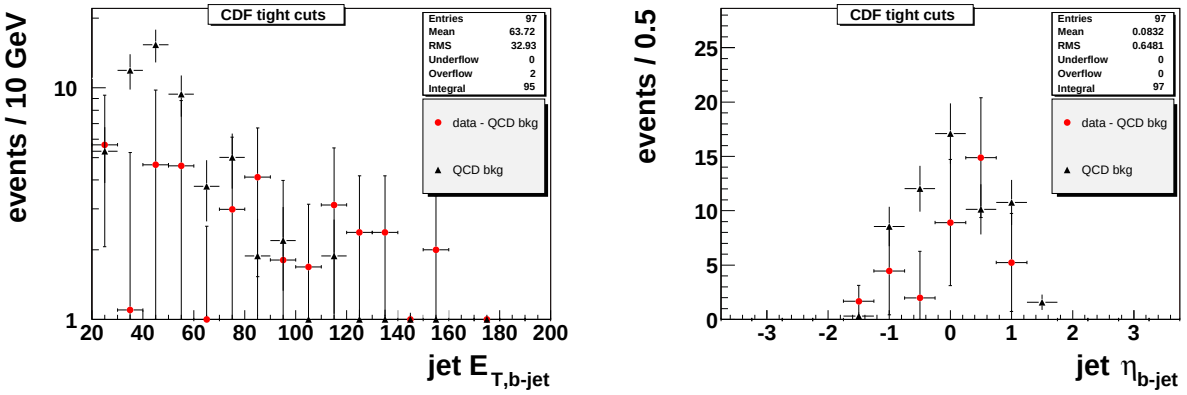


Figure C.1: Transverse energy and pseudorapidity distribution of the tagged jet for the tight cut analysis.

This study just presents a first look into the tagged W boson data sample. To measure cross-sections in this channel, a detailed Monte Carlo study is needed. However, one can see the big importance of non W background for analyses in this channel. To study the kinematical differences between signal and QCD events with hadrons faking the electron, more data is needed, but due to the large contributions from background a good understanding of the background shape is needed for top quark analyses, especially in searches for single top production.

Bibliography

[1] V. Barger and R. J. N. Phillips, "Collider Physics,". Redwood City, USA: Addison-Wesley (1987) 592 P. (Frontiers in Physics, 71).

[2] Walter Greiner and Bernd Müller, "Gauge theory of weak interactions,". Springer (2000) 402 p.

[3] E. Fermi, "Trends to a theory of beta radiation. (in italian)," *Nuovo Cim.* **11** (1934) 1–19.

[4] E. Fermi, "Versuch einer Theorie der β-Strahlen. I.," *Z. Phys.* **88** (1934) 161–177.

[5] G. Gamow and E. Teller, "Selection Rules for the β-Disintegration," *Phys. Rev.* **49** (1936) 895.

[6] T. D. Lee and C.-N. Yang, "Question of parity conservation in weak interactions," *Phys. Rev.* **104** (1956) 254–258.

[7] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental test of parity conservation in beta decay," *Phys. Rev.* **105** (1957) 1413–1414.

[8] M. Goldhaber, L. Grodzins, and A. W. Sunyar, "Helicity of neutrinos," *Phys. Rev.* **109** (1958) 1015–1017.

[9] E. C. G. Sudarshan and R. e. Marshak, "Chirality invariance and the universal fermi interaction," *Phys. Rev.* **109** (1958) 1860–1862.

[10] R. P. Feynman and M. Gell-Mann, "Theory of the Fermi interaction," *Phys. Rev.* **109** (1958) 193–198.

[11] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Phys. Rev. Lett.* **13** (1964) 321–323.

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

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

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

[15] A. Salam and J. C. Ward, "Weak and electromagnetic interactions," *Nuovo Cim.* **11** (1959) 568–577.

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

[17] M. L. Perl *et. al.*, "Evidence for anomalous lepton production in e^+e^- annihilation," *Phys. Rev. Lett.* **35** (1975) 1489–1492.

[18] M. Gell-Mann, "A schematic model of baryons and mesons," *Phys. Lett.* **8** (1964) 214–215.

[19] J. E. Augustin *et. al.*, "Discovery of a Narrow Resonance in e^+e^- Annihilation," *Phys. Rev. Lett.* **33** (1974) 1406–1408.

[20] J. J. Aubert *et. al.*, "Experimental Observation of a Heavy Particle J," *Phys. Rev. Lett.* **33** (1974) 1404–1406.

[21] S. W. Herb *et. al.*, "Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions," *Phys. Rev. Lett.* **39** (1977) 252–255.

[22] **CDF** Collaboration, F. Abe *et. al.*, "Evidence for top quark production in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev. Lett.* **73** (1994) 225–231, `hep-ex/9405005`.

[23] **D0** Collaboration, S. Abachi *et. al.*, "Observation of the top quark," *Phys. Rev. Lett.* **74** (1995) 2632–2637, `hep-ex/9503003`.

[24] N. Cabbibo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.* **10** (1963) 531–533.

[25] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.* **49** (1973) 652–657.

[26] C. Rubbia, "The physics of the proton - anti-proton collider,". Invited talk given at Int. Europhysics Conf. on High Energy Physics, Brighton, England, Jul 20-27, 1983.

[27] K. Hagiwara *et. al.*, "Review of Particle Physics," *Physical Review D* **66** (2002) 010001+.

[28] LEP Electroweak Working Group *et. al.*, "A Combination of Preliminary Electroweak Measurements and Constraints on the Standard Model,". LEPEWWG/2003-01.

[29] **CTEQ** Collaboration, H. L. Lai *et. al.*, "Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions," *Eur. Phys. J.* **C12** (2000) 375–392, `hep-ph/9903282`.

[30] The plot has been made with the on-line PDF plotter at http://www-spires.dur.ac.uk/hepdata/pdf3.html.

[31] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Parton distributions and the LHC: W and Z production," *Eur. Phys. J.* **C14** (2000) 133–145, `hep-ph/9907231`.

[32] private communication with J. Stirling according to CDF-note 6110.

[33] L. J. Dixon, Z. Kunszt, and A. Signer, "Vector boson pair production in hadronic collisions at $O(\alpha_s)$: Lepton correlations and anomalous couplings," *Phys. Rev.* **D60** (1999) 114037, `hep-ph/9907305`.

[34] D. Benjamin, H. Gerberich, A. Goshaw, H. Hayward, B. Heinemann, M. Kirby, and N. Tanimoto, "Cross Section Measurements for W+photon and Z+photon Production in Run II,". CDF Note-6366.

[35] **Higgs Working Group** Collaboration, M. Carena *et. al.*, "Report of the Tevatron Higgs working group," `hep-ph/0010338`.

[36] E. L. Berger and H. Contopanagos, "Threshold resummation of the total cross-section for heavy quark production in hadronic collisions," *Phys. Rev.* **D57** (1998) 253–264, `hep-ph/9706206`.

[37] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan, and S. Weinzierl, "The fully differential single top quark cross section in next-to-leading order QCD," *Phys. Rev.* **D66** (2002) 054024, `hep-ph/0207055`.

[38] F. Landry, R. Brock, P. M. Nadolsky, and C. P. Yuan, "Tevatron Run-1 Z boson data and Collins-Soper-Sterman resummation formalism," `hep-ph/0212159`.

[39] J. C. Collins and D. E. Soper, "Angular distribution of dileptons in high-energy hadron collisions," *Phys. Rev.* **D16** (1977) 2219.

[40] U. Baur *et. al.*, "Report of the working group on photon and weak boson production," `hep-ph/0005226`.

[41] The LEP working group for Higgs Boson Searches, "Search for the Standard Model Higgs Boson at LEP,". CERN-EP/2003-011, submitted to Phys.Lett.B.

[42] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, "The $t\bar{t}$ cross-section at 1.8 and 1.96 TeV: A study of the systematics due to parton densities and scale dependence," `hep-ph/0303085`.

[43] **CDF** Collaboration, D. Acosta *et. al.*, "Search for single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev.* **D65** (2002) 091102, `hep-ex/0110067`.

[44] **D0** Collaboration, V. M. Abazov *et. al.*, "Search for single top quark production at D0 using neural networks," *Phys. Lett.* **B517** (2001) 282–294, `hep-ex/0106059`.

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

[46] **CDF** Collaboration, F. Abe *et. al.*, "Evidence for top quark production in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev.* **D50** (1994) 2966–3026.

[47] **CDF** Collaboration, F. Abe *et. al.*, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.* **74** (1995) 2626–2631, `hep-ex/9503002`.

[48] **CDF-II** Collaboration, R. Blair *et. al.*, "The CDF-II detector: Technical design report,". FERMILAB-PUB-96-390-E.

[49] **CDF** Collaboration, F. Abe *et. al.*, "The CDF Detector: An Overview," *Nucl. Instr. Meth.* **A271** (1988) 387–403.

[50] **CDF** Collaboration, F. Abe *et. al.*, "Measurement of the W-boson mass," *Phys. Rev.* **D52** (1995) 4784–4827.

[51] **CDF-II** Collaboration, F. Abe *et. al.*, "Proposal for Enhancement of the CDF II Detector: An Inner Silicon Layer and A Time of Flight Detector,". FERMILAB-PROPOSAL-909.

[52] **CDF** Collaboration, T. K. Nelson, "The CDF-II silicon tracking system," *Nucl. Instrum. Meth.* **A485** (2002) 1–5.

[53] S. Cihangir *et. al.*, "SVX': The New CDF silicon vertex detector," *Nucl. Instrum. Meth.* **A360** (1995) 137–140.

[54] F.Hartmann and Th.Müller and M.Schilling and A.Heiss and D.Knoblauch and H.Wenzel, "Evaluation of the ISL prototype detectors,". CDF Note-4627.

[55] **CDF** Collaboration, D. Amidei *et. al.*, "The Silicon Vertex Detector of the Collider Detector at Fermilab," *Nucl. Instrum. Meth.* **A350** (1994) 73–130.

[56] **CDF** Collaboration, J. Cranshaw, "Status of CDF Run II upgrade," *Nucl. Instrum. Meth.* **A462** (2001) 170–173.

[57] **CDF** Collaboration, K. Burkett, "Design and construction of the CDF central outer tracker," *Nucl. Instrum. Meth.* **A461** (2001) 62–64.

[58] **CDF** Collaboration, M. G. Albrow *et. al.*, "The CDF plug upgrade electromagnetic calorimeter: Test beam results," *Nucl. Instrum. Meth.* **A480** (2002) 524–546.

[59] G. W. Foster, J. Freeman, C. Newman-Holmes, and J. Patrick, "A Fast Hardware Track Finder for the CDF Central Tracking Chamber," *Nucl. Instrum. Meth.* **A269** (1988) 93.

[60] **CDF** Collaboration, H. Wenzel *et. al.*, "Online Monitoring and Module Maintenance for CDF in the upcoming Fermilab Tevatron Run II," in *Computing in High Energy and Nuclear Physics (CHEP 2000)*, pp. 335–340. 2000.

[61] H. Stadie, "CDF Online Monitoring." Talk at *ROOT2001 International HENP ROOT Users Workshop*, http://www-root.fnal.gov/root2001/presentations/session5/consumer.pdf.

[62] **CDF** Collaboration, K. Maeshima *et. al.*, "Online Monitoring for the CDF Run II Experiment," in *2003 Conference for Computing in High-Energy and Nuclear Physics (CHEP 03)*. 2003.

[63] E. Sexton-Kennedy, M. Shapiro, R. Snider, and R. Kennedy, "The physical design of the CDF simulation and reconstruction software," in *Computing in High Energy and Nuclear Physics (CHEP 2000)*, pp. 161–164. 2000.

[64] R. Brun and F. Rademakers. Root system homepage http://root.cern.ch.

[65] W. Wagner, "Online Monitoring in the CDF II Experiment." Talk at *EPS HEP 2001 Conference* Budapest, http://www-ekp.uni-karlsruhe.de/~wagner/pubs/eps2001.ps.gz.

[66] M. Shimojima, B. J. Kilminster, K. S. McFarland, A. Vaiciulis, and D. J. Holmgren, "Consumer-Server/Logger System for the CDF Experiment," *IEEE Trans. Nucl. Sci.* **47** (2000) 236–239.

[67] **CDF** Collaboration, J. Antos *et. al.*, "The CDF Run 2 offline computer farms," in *CHEP'01: Computing In High-Energy and Nuclear Physics*, pp. 53–56. 2001.

[68] H. Wenzel, "Tracking in the SVX,". CDF Note-1790.

[69] A. Mukherjee, "CTC and VTX Tracking,". CDF Note-5490.

[70] P. Azzi, G. Busetto, P. Gatti, and A. Ribon, "Histogram Tracking in the COT,". CDF Note-5562.

[71] K. Bloom and W.-M. Yao, ""Outside-In"Silicon Tracking at CDF,". CDF Note-5991.

[72] S. Menzemer, *Spurrekonstruktion im Silizium-Vertexdetektor des CDF-II-Experiments*. PhD thesis, Institut für Experimentelle Kernphysik, Universität Karlsruhe, 2003.

[73] F. Bedeschi *et. al.*, "A Primary Vertex Finding Package,". CDF Note-1789.

[74] H. Wenzel, "Fitting the beam position with the SVX,". CDF Note-1924.

[75] G. Punzi and S. Dell'Agnello, "The C$VTX Vertex-Finding Package,". CDF Note-1791.

[76] H. Stadie, W. Wagner, H. Wenzel, and Th. Müller, "Vxprim in Run II,". CDF Note-6047.

[77] Run II Alignment group http://www-cdf.fnal.gov/internal/upgrades/align/alignment.html.

[78] VertexMods source in the code browser http://cdfcodebrowser.fnal.gov/CdfCode/source/VertexMods/.

[79] VertexAlg source in the code browser http://cdfcodebrowser.fnal.gov/CdfCode/source/VertexAlg/.

[80] H. Stadie and H. Wenzel, "Proposal for a Run II Vertex Object,". CDF Note-5161.

[81] J. Boudreau and R. Snider, "A User's Guide to CdfTrack and Related Classes,". CDF Note-5089.

[82] M. Martens and P. Bagley, "Luminosity Distribution During Collider Run II,". http://www-ap.fnal.gov/∼martens/luminosity/beam_size.html.

[83] H. Wenzel, "The Primary Interaction Vertex,". CDF Note-4066.

[84] B. Flaugher, "Efficiency of the 60 cm Z-vertex cut,". CDF Note-5094.

[85] H. Stadie *et. al.*, "The Beam Position in Run II,". CDF Note-6327.

[86] W. Wagner, "An estimate of the average beam width at CDF,". CDF Note-6104.

[87] S. Lai *et. al.*, "An Updated Measurement of the Beam Width at CDF,". CDF Note-6492.

[88] G. C. Blazey *et. al.*, "Run II jet physics," in *Physics at Run II: QCD and Weak Boson Physics Workshop*, pp. 47–77. 1999. `hep-ex/0005012`.

[89] J.-F. Arguin *et. al.*, "Generic Jet Energy Corrections for Run II,". CDF Note-6280.

[90] C. Rott *et. al.*, "SecVtx Optimization for the 2003 Winter Conferences,". CDF Note-6242.

[91] H. Bachacou *et. al.*, "Efficiency of SecVtx B-Tag Used for 2003 Winter Conferences,". CDF Note-6291.

[92] T. Sjostrand *et. al.*, "High-energy-physics event generation with PYTHIA 6.1," *Comput. Phys. Commun.* **135** (2001) 238–259, `hep-ph/0010017`.

[93] S. Klimenko, J. Konigsberg, and T. Liss, "Averaging of the inelastic cross-section measured by the CDF and the E811 experiments,". CDF Note-6314.

[94] I. Kleppe, "Erkennung und Analyse der Prozesse $p\bar{p} \to WX \to e\nu_e X$ und $p\bar{p} \to Z^0 X \to e^+e^- X$ mit dem CDF Experiment," Master's thesis, Institut für Experimentelle Kernphysik, Universität Karlsruhe, 2003.

[95] M. Coca and E. Halkiadakis, "Central Electron Identification Efficiencies for Winter 2003 conferences,". CDF Note-6262.

[96] P. Wilson, "Calorimeter Isolation and Lateral Shower Leakage for Photons and Electrons,". CDF Note-4170.

[97] A. Wyatt and B. Heinemann, "Correction for Leakage Energy in the Central and Plug Calorimeters in Run II,". CDF Note-6167.

[98] R. G. Wagner, "Understanding and Using Lshr,". CDF Note-6249.

[99] G. Manca and Y.-K. Kim, "Z^0 Cross Section Measurement Using Run II Electrons,". CDF Note-6281.

[100] M. Feindt, "Neurobayes - a Neural Bayesian Estimator for Conditional Probability Densities,". Report IEKP-KA/01-1, Institut für Experimentelle Kernphysik, Universität Karlsruhe.

[101] M. Coca *et. al.*, "$W \to e\nu$ Cross Section Analysis with Run II Data,". CDF Note-6300.

[102] F. James and M. Roos, "'Minuit' a System for Function Minimization and Analysis of the Parameter Errors and Correlations," *Comput. Phys. Commun.* **10** (1975) 343–367.

[103] Y.-K. Kim *et. al.*, "Trigger Efficiencies for High E_T Electrons,". CDF Note-6234.

[104] **CDF** Collaboration, F. Abe *et. al.*, "Measurement of $\sigma \cdot B(W \to e\nu)$ and $\sigma \cdot B(Z^0 \to e^+e^-)$ in $p\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV," *Phys. Rev. Lett.* **76** (1996) 3070–3075, `hep-ex/9509010`.

[105] **CDF** Collaboration, F. Abe *et. al.*, "Properties of jets in W-boson events from 1.8 TeV $\bar{p}p$ collisions," *Phys. Rev. Lett.* **79** (1997) 4760–4765, `hep-ex/9709016`.

About the Pictures at the Beginning of Each Chapter

The pictures at the beginning of each chapter show the ontogenesis of *cynops orientalis* David, 1873. Its common name is *chinese fire belly newt*. This species can be found in east-central china. The adults have a size of up to nine centimeters.

Since June 2002, I have been keeping a group of two males and four females. In February 2003, the first eggs were discovered at the plants in the aquarium. The development of these eggs (chapter one and two), the larvae (chapter three and four) and a juvenile newt (chapter five) are shown at the beginning of each chapter. An adult female fixing an egg to a plant is shown at the beginning of chapter six.

I like to thank Michael Milnik for his help with taking these pictures.

Acknowledgments

I wish to sincerely thank my supervisor Prof. Dr. Thomas Müller for offering me both guidance and freedom in my work and Prof. Dr. Michael Feindt who co-supervised this thesis.

I must thank Dr. Wolfgang Wagner for always supporting my work and for carefully reading the manuscript. For the same, I owe a debt of gratitude to Dr. Hans Wenzel who already guided my work on my diploma thesis and continued to offer me advice and support. Although I complained a lot, I enjoyed running races with him and Dr. Kaori Maeshima, when I stayed at Fermilab.

This thesis would have been impossible without the hard work of every member of the CDF collaboration on building up the experiment and keeping it running. Besides them, I like to thank the former and recent members of the Institut für Experimentelle Kernphysik in Karlsruhe, especially my office mates, for the nice and encouraging atmosphere and their support over the last few years. For her help in creating the posters and improving plots, I thank Dr. Stephanie Menzemer.

Support of this work was provided in form of a graduate scholarship by the "Graduiertenkolleg für Teilchen- und Astrophysik".

Finally, I wish to thank my family and Julia for their support.