A Measurement of the Production of Jets in Association with a W Boson in Proton-Antiproton Collisions at the Tevatron Using Data Collected with the CDF Experiment

Ben Cooper

2006

University College London

PhD thesis submitted to the University of London in fulfilment of the requirements for the award of the degree of Doctor of Philosophy.

Declaration

2

I, Ben Cooper, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

A measurement of the differential $p\bar{p} \to W + \geq n$ jets production cross-section with respect to a range of different jet kinematic variables is presented. The W is required to decay leptonically to an electron and neutrino, and the kinematics of the decay products are restricted such that the electron $P_T > 20$ GeV, electron $\eta < 1.1$, neutrino $P_T > 30$ GeV and the W transverse mass > 20 GeV/c^2. Jets are defined using the CDF JetClu cone algorithm with a radius of R = 0.4, requiring $\eta_{jet} < 2.0$ and $E_T > 15$ GeV. The jet energies are corrected individually such that they are on average equal to the energy of the hadrons within the jet, with additional correction factors being applied to account for the impact of detector resolution on the jet spectra.

The differential cross-section is measured with respect to the first, second, third and fourth jet E_T spectra, the separation in R and the invariant mass of the two leading jets. In addition, the jet E_T measurements can be integrated to form $W + \geq n$ Jets cross-section measurements for 1,2,3 and 4 jets with a range of minimum jet E_T requirements, the lowest being 15 GeV.

These cross-section results are compared to Enhanced Leading Order (ELO) W + jets theoretical predictions and predictions made using the leading order CKKW and MLM matching prescriptions. The ELO predictions are not suited to describing the absolute rate of W + jet production, but display success in reproducing certain relative rates, dependent on the choice of renormalisation scale. The comparisons with matching prescriptions indicate that these new approaches could be used successfully to improve W + Jets LO predictions.

Acknowledgements

First and foremost this thesis is dedicated to my parents, John and Carol Cooper, whose encouragement, support and interest throughout and beyond my school days has given me the confidence, skills and the opportunity to undertake and complete these studies.

A PhD in Particle Physics is a demanding but ultimately very rewarding undertaking. I have been lucky enough to have worked with some extremely talented people in this field, who have taught me a great deal in the many and varied disciplines one needs to succeed. My supervisor, Dr David Waters, despite being in huge demand at CDF, has always made time to advise and aid, and I have learned alot from him. One of the nicest and most unique aspects of academic research is the opportunity for collaboration with institutions of other countries, and it has been a great pleasure to have worked with Andrea Messina (Rome, INFN) on W + Jets over the past few years. I would also like to thank Joey Huston and Jay Dittmann for their encouragement and advice on QCD issues, Stuart Boogert, Angela Wyatt and Jon Cauchman for their exhaustive and entirely voluntary efforts in teaching me how to programme in my first year, and Pasha Murat for his help and inspiration.

Carrying out these studies at CDF also gave me the opportunity to live in Chicago for over two years, which was an experience made fantastic by a great city and some great people who I was fortunate enough to meet there. To avoid any unintentional offence caused by memory failure I won't name any names here, but I hope we can all stay in touch. I would also like to thank PPARC and the British tax payer, without which all this would not have been possible.

Contents

List of Figures

18

List of Tables

Chapter 1

Introduction

Particle physics experiments explore the interactions and properties of the fundamental particles of nature, and thus enable us to test our understanding of the universe at the smallest known distance scales. Over the last fifty or so years a great number of important discoveries have been made at such experiments which have prompted, informed and confirmed theoretical models, in particular leading to the establishment of the Standard Model of particle physics, a theoretical framework which has proved extremely successful in describing experimental observations. Currently the highest energy collider in the world is the Tevatron, located near Chicago, USA, which collides protons and antiprotons at a centre-of-mass energy of 1.96 TeV. The data used in this thesis was collected at the CDF (Collider Detector at Fermilab) experiment, one of two multipurpose detectors at the Tevatron.

The Standard Model predicts that in $p\overline{p}$ (proton-antiproton) interactions a W boson may be produced along with one or more quarks or gluons, observable as "jets" of hadrons in a detector. This thesis presents a detailed study of W boson production in association with one or more high momentum jets, where the W decays to an electron and neutrino. In particular, the differential cross-section for W production is measured as a function of several jet kinematic variables. Although for an accurate measurement it is important to understand the reconstruction of the W decay products in the CDF detector, the focus is very much on the properties of the jets that are produced in association.

The W + Jets process provides an excellent laboratory for testing the predictions of Quantum Chromodynamics (QCD), the model of the quark-quark and quark-gluon "strong" interactions. It is crucial that QCD predictions of the W + Jets process are vigorously tested, not only from the standpoint of confirming or otherwise our understanding of strong interactions, but because many of the most important and interesting processes under study at the Tevatron and at the future Large Hadron Collider (LHC) have a final state consisting of a W boson and accompanying jets. These processes occur at rates which are significantly lower than that for W

+ Jets, and thus if they are to be measured precisely it is vital that there exists a theoretical model that is able to accurately describe the large W + Jets background that will be present in the data.

In Chapter 2 we describe the Standard Model of particle physics, in particular focusing on the predictive power of QCD. In Chapter 3 we discuss the motivations for studying the W + Jets process, and describe how QCD theory is applied to the W + Jets process at hadron colliders to make predictions that can be compared with data. The experimental apparatus used to make this measurement, namely the Tevatron collider and CDF detector, is detailed in Chapter 4. In Chapter 5 we define precisely the measurement that we are making, focusing in particular on how this definition is chosen in order to produce a measurement that is independent of the CDF detector and theoretical approaches. In Chapters 6, 7 and 8 we detail the major components of the cross-section measurement, the candidate event selection, background estimation and calculation of acceptance factors respectively. In Chapter 9 we describe the procedure to extract the "unsmearing" correction factors which are necessary to account for the resolution of the jet energy measurements. In Chapter 10 we present our W + Jets differential cross-section results and in Chapter 11 we compare our results to theory, examining a range of theoretical approaches and exploring the successes and limitations of these predictions. Finally we summarise and discuss possible extensions to the analysis in Chapter 12.

Chapter 2

The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) [86, 106, 110] is a quantum field theory (QFT) [94] description of the fundamental building blocks of matter and the interactions between them. In this Chapter we introduce the fundamental particles and interactions of the SM, focusing in particular on how we calculate the rates of processes in QFT using perturbation theory, and the limitations of this approach.

2.1 The Fundamental Particles and Interactions

2.1.1 The Fermions

In the SM all matter consists of fundamental particles called *fermions*. The fermions can be subdivided into two families; *leptons* and *quarks* [70]. Table 2.1 details the properties of the six leptons (the electron, muon, tau and associated neutrinos) and the six quarks (up, down, charm, strange, beauty and top). Notice that both the quarks and the leptons can be subdivided into three doublets or *generations* according to their mass, as shown below:

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}$$

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$$

In addition each fermion has an associated *antiparticle* [65] which has opposite electric charge to its matter partner.

Fermion	Symbol	Charge (e)	Mass (MeV/c^2)	Interactions
Quarks				
up	u	+2/3	1.5-4	strong/EM/weak
down	d	-1/3	4-8	strong/EM/weak
charm	c	+2/3	1150-1350	strong/EM/weak
strange	s	-1/3	80-130	strong/EM/weak
top	t	+2/3	173500±3900	strong/EM/weak
bottom	b	-1/3	4100-4400	strong/EM/weak
Leptons				
electron	e	+1	0.511	EM/weak
neutrino	ν_e	0	0	weak
muon	μ	+1	106	EM/weak
neutrino	ν_μ	0	0	weak
tau	τ	+1	1777	EM/weak
neutrino	ν_τ	0	0	weak

Table 2.1: Properties of the 12 fermions of the standard model [70]. Not shown are the partner antiparticles, which have opposite charge. Charges refer to electric charge and are in units of electron charge.

2.1.2   Gauge Symmetries and the Bosons

In the quantum field formalism the fermions are states of a quantum field ϕ. In an analogue to classical mechanics the properties of this field, including the fermion dynamics, are completely described by the Lagrangian density. Interactions between the matter particles are introduced into the theory by imposing gauge symmetry conditions on the quantum fields [24, 37, 53]. This symmetry requires that the Lagrangian remains invariant under a Lie group of local space-time transformations of the field ϕ. The most important groups in this context are $U(n)$, the group of $n \times n$ Unitary matrices, and $SU(n)$, the group of $n \times n$ Special Unitary matrices (required to have determinant equal to one). In quantum theory these transformation matrices are built up from quantum mechanical generators which correspond to a particular observable, such as charge. An $SU(n)$ group of transformations has $n^2 - 1$ generators. The gauge invariance condition requires that the observable associated with the generator of the transformation is conserved. In order for this condition to be met, it is found that one has to introduce additional gauge field terms to the Lagrangian, one for each generator, and these gauge fields describe the propagation of fundamental spin 1 force carrying particles, the *bosons*.

Boson	Symbol	Charge (e)	Mass (MeV/c^2)	Force Mediated
photon	γ	0	0	EM
W	W^+,W^-	+1,-1	80.4	weak
Z	Z	0	91.2	weak
gluon	g	0	0	strong

Table 2.2: Properties of the bosons [70]. Charges refer to electric charge and are in units of electron charge.

The properties of the SM bosons are described in Table 2.2. Each boson couples to a certain "charge" and is thus responsible for the mediation of a certain force, or interaction, in which the total charge must always be conserved. The SM describes three of the four fundamental forces of nature; the electromagnetic, weak and strong forces. Currently there exists no quantum field formalism for gravitational interactions. However, the force of gravity is sufficiently weak that it can be ignored at energies less than the Planck scale $E < M_{Planck} \approx 10^{19} GeV$.

2.1.3   Quantum Electrodynamics (QED)

The theory of Quantum Electrodynamics (QED) [86, 106, 110] describes electromagnetic interactions between electrically charged particles. QED requires gauge invariance under the $U(1)$ group of transformations. The $U(1)$ group has one generator that is associated with the observable of electric charge, and hence one boson, the photon, which mediates the electromagnetic interaction between charged particles, in which charge must always be conserved. Since there is only one generator the $U(1)$ transformations are commutative (the group Abelian), and this means that the associated bosons cannot interact with one another i.e. the photon is electrically neutral. The coupling strength (defined more precisely in Section 2.2) in QED is proportional to the electric charge of the particle.

2.1.4   Electroweak Theory

Electroweak theory provides a unified description of both the electromagnetic and weak interactions [80, 115, 116, 125]. It demands invariance under the $SU(2) \times U(1)$ group of transformations. This requires the introduction of four gauge bosons. The $SU(2)$ group has $2^2 - 1 = 3$ bosons, which we label W^-,W^+ and W^0. We have already seen that $U(1)$ symmetry introduces one boson into the theory. In this context we label this boson B^0. In contrast to the $U(1)$ transformations, the $SU(2)$ matrices are non-commutative, making electroweak theory non-Abelian. This leads
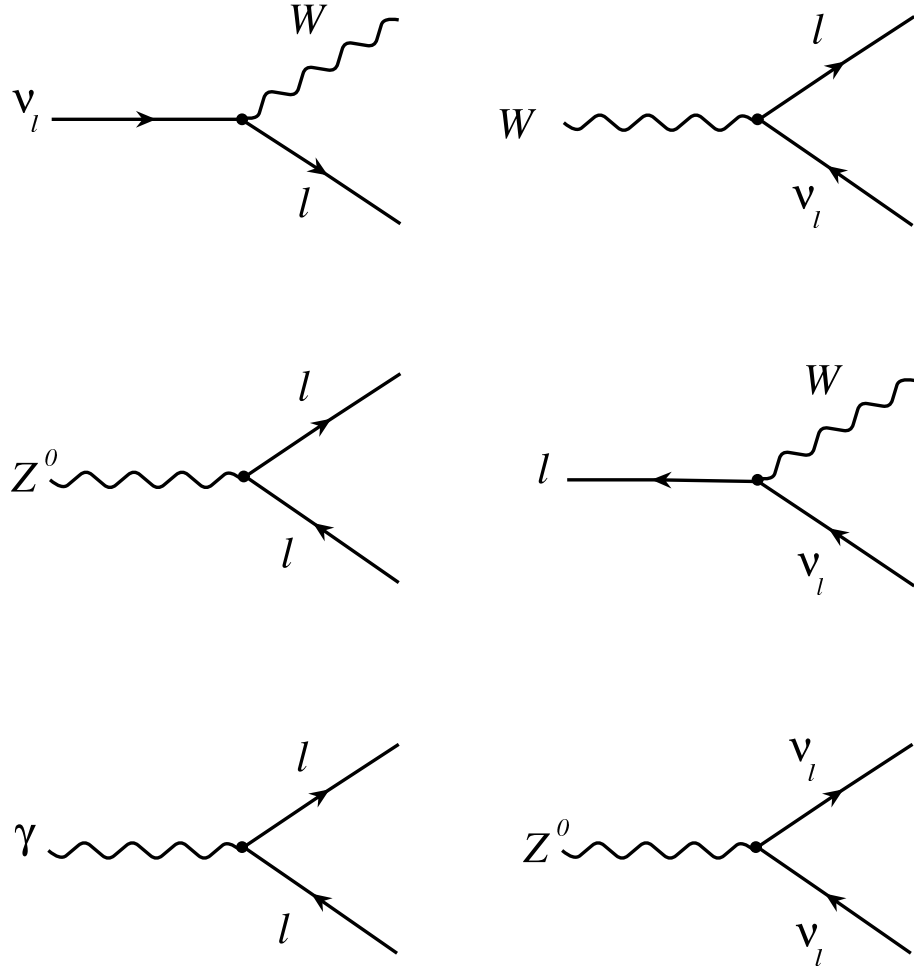
Figure 2.1: Some of the possible processes involving electroweak interaction vertices in the lepton sector. The lepton on each line must be the same type l to conserve lepton number.

to the important result that the gauge bosons of the theory are self-interacting. The W^0 and B^0 states mix quantum mechanically, producing the physical states of the Z^0 boson (associated with the weak interaction) and the photon (associated with the electromagnetic interaction).

In principle electroweak bosons can couple to any fermion, but interactions must conserve electric charge and also lepton (quantum) number. This restricts the basic interaction vertices in the lepton sector; couplings only occur between leptons within the same generation or doublet. Some examples of basic electroweak processes in the lepton sector are given in Figure 2.1. The coupling strengths g_W and g_z of the W and Z bosons respectively are independent of the particular lepton types (e, μ, τ) involved, a property known as *lepton universality.*

The theory also requires that there can be no flavour changing neutral currents. However, flavour changing charged currents are permitted, and hence the $W^{\pm}$ bosons "convert" quarks from one flavour to another. It is this feature of electroweak inter-
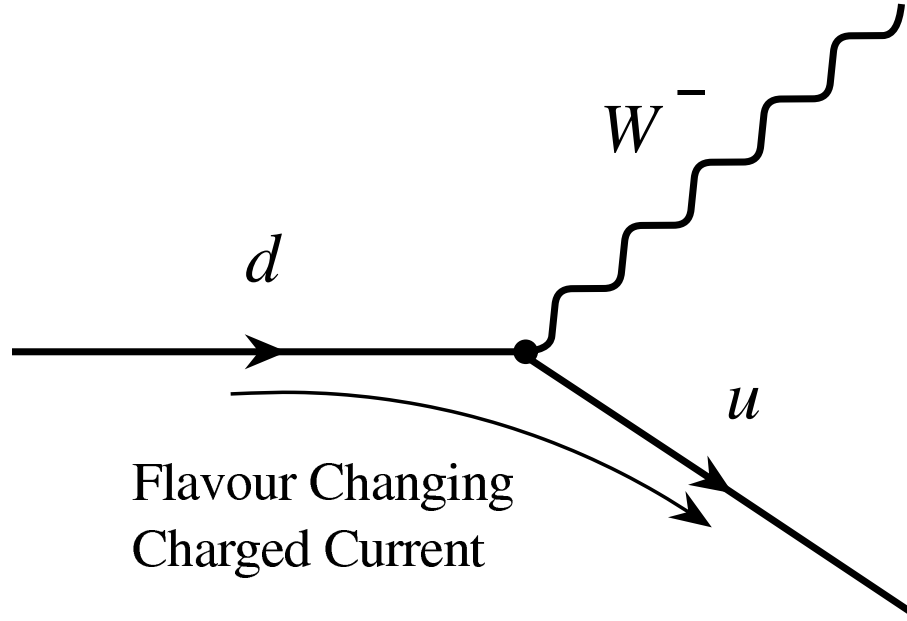
Figure 2.2: Diagram of the $d \to W^- u$ electroweak vertex.

actions which is responsible for radioactive β decay, in which a neutron is converted into a proton via a $d \to W^- u$ electroweak vertex, illustrated in Figure 2.2. For W interaction vertices involving the ud, cs and tb doublets the coupling strength is the same as for the lepton sector W vertices, g_W. However, unlike the leptonic interactions, the W can interact with quarks of differing generations, albeit with a reduced coupling strength. The coupling of the W boson to a particular flavour changing vertex is modified according to the experimentally determined coefficients of the Cabbibo Kobayashi Maskawa (CKM) matrix, shown in Equation 2.1 below [70], such that the modified coupling strength between quarks of flavours x and y is $g_W' = V_{xy}g_W$.

$$
\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \approx \begin{pmatrix} 0.964 \text{ to } 0.975 & 0.221 \text{ to } 0.227 & 0.003 \text{ to } 0.005 \\ 0.221 \text{ to } 0.227 & 0.973 \text{ to } 0.974 & 0.039 \text{ to } 0.044 \\ 0.005 \text{ to } 0.014 & 0.037 \text{ to } 0.043 & 0.999 \text{ to } 0.999 \end{pmatrix} \quad (2.1)
$$

See that for ud, cs and tb vertices the CKM coefficient is essentially unity. The coupling between quarks of different generations is said to be *Cabbibo suppressed*.

2.1.5 The Higgs Mechanism

When electroweak theory was first postulated there was one major obstacle to it's acceptance. All four bosons are predicted by the theory to be massless and this is in direct conflict with what is observed in nature. If mass terms for the parti-

cles were simply added to the Lagrangian the gauge symmetry would be broken. An alternative solution is required. In the so-called *Higgs mechanism* [90, 91] an additional scalar field, the Higgs field, is introduced into the theory, which has a potential function of a form allowing degenerate vacuum solutions with a non-zero vacuum expectation value. The vacuum states are invariant only to the $U(1)$ transformations of the electromagnetic theory, and thus the $SU(2) \times U(1)$ symmetry is "spontaneously broken". Interaction with the Higgs field allows the W and Z bosons, as well as other particles, to acquire mass, with the value of the mass dependent on the strength of the coupling. However, the mediator of this field, the spin-0 Higgs boson, has yet to be observed experimentally.

2.1.6   Quantum Chromodynamics (QCD)

The theory of Quantum Chromodynamics (QCD) [72, 84, 111] describes the strong interaction of quarks and requires invariance under $SU(3)$ transformations. In this theory a quark can be in one of three *colour* states (red, green or blue) which form the fundamental representation of the group. The group has $3^2 - 1 = 8$ generators and thus gauge invariance requires the introduction of 8 massless *gluon* fields. These gluons couple to colour charge and since the $SU(3)$ group is non-Abelian the gluons are self-interacting, carrying colour charge themselves. As will be explained this has far reaching consequences for the predictive power of QCD theory.

2.2   Calculation of Observables in QFT

For any theory to be truly successful it has to have the power to predict observables; quantities that can be measured experimentally and hence allow the theoretical model to be tested. The cross-section for a particular well-defined process is a very useful observable to calculate as it can be directly measured at particle collider experiments (see Section 4.1.2). It is the effective area over which the particles in the initial state interact to produce the final state, and is directly proportional to the rate of a process.

QED processes involve the coupling of massless photon fields to charged particles, and QCD similarly involves the coupling of massless gluon fields to coloured particles. Although we shall see that the strength of this coupling, and particularly it's dependency on the energy scale of the interaction, are crucially different in QED vs QCD, formally the procedure for the calculation of cross-sections in QED and QCD processes is exactly the same. In this section we shall outline the calculation of the cross-section for a $2 \rightarrow 2$ QED process; $e^+e^- \rightarrow \gamma^* \rightarrow \mu^+\mu^-$ scattering [86]. Figure 2.3 shows a possible Feynman diagram for this process, involving two QED coupling vertices. An analogous QCD process would be $q\bar{q} \rightarrow g^* \rightarrow q\bar{q}$, a possible
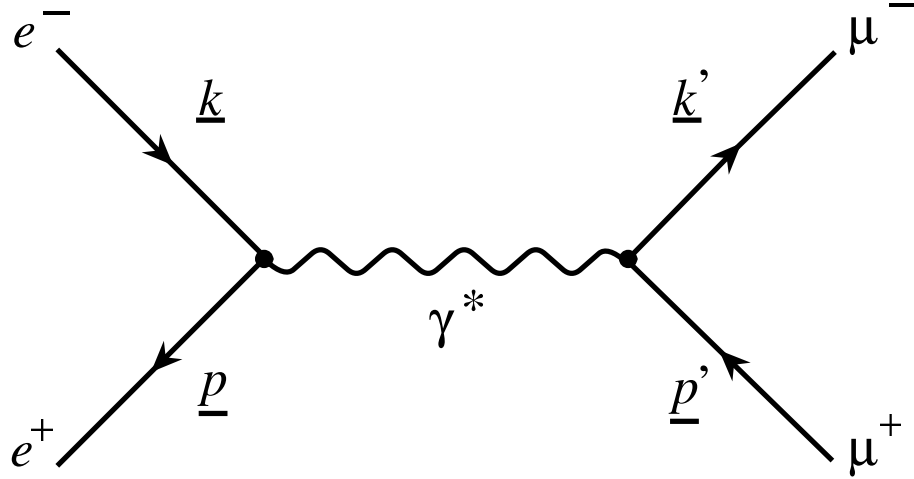
Figure 2.3: The lowest-order Feynman diagram for the process $e^+e^- \to \gamma^* \to \mu^+\mu^-$.
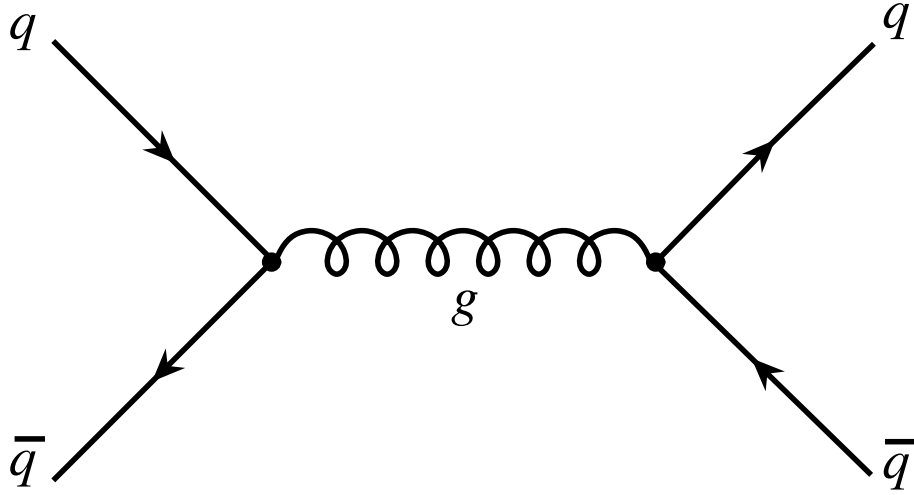
Figure 2.4: The lowest-order Feynman diagram for the process $q\bar{q} \to g^* \to q\bar{q}$.

Feynman diagram of which is shown in Figure 2.4, involving two QCD coupling vertices.

Shown on the diagram in Figure 2.3 are the four-vectors $\underline{k}$, $\underline{k}'$, $\underline{p}$ and $\underline{p}'$ of the initial and final state particles. From such a diagram one can calculate a probability amplitude $\mathcal{M}$. Each line in the Feynman diagram can be related to a propagator term in the amplitude calculation, and each interaction vertex introduces a factor $\sqrt{\alpha}$. In this case α is the QED coupling constant, for a QCD process it would be the QCD coupling constant α_S. The squared modulus of the amplitude for the diagram in Figure 2.3, $|\mathcal{M}_1|^2$, is given by[1]:

$$|\mathcal{M}_1|^2 = 32\pi^2\alpha^2\left(\frac{t^2 + u^2}{s^2}\right) \tag{2.2}$$

---

[1]Throughout this thesis we use natural units, such that $\hbar = c = 1$

where the variables s, t and u are frame invariant scalar quantities formed from the initial and final state particle four-vectors as shown below:

$$s = (\underline{k} + \underline{p})^2$$
$$t = (\underline{k} - \underline{k'})^2$$
$$u = (\underline{k} - \underline{p'})^2$$

Note that $\sqrt{s}$ is the centre-of-mass energy of the interaction. To obtain a cross-section the squared modulus of the amplitude, known as the interaction probability, must be integrated over all possible final and initial state phase space i.e. all kinematically allowed values of $\underline{k}$, $\underline{k'}$, $\underline{p}$ and $\underline{p'}$. Performing such an integration yields the following cross-section for $e^+e^- \rightarrow \gamma^* \rightarrow \mu^+\mu^-$ scattering:

$$\sigma(e^+e^- \rightarrow \mu^+\mu^-) = \frac{4\pi\alpha^2}{3s} \tag{2.3}$$

Consider this process in the centre-of-mass frame. The total energy in the initial state is $\sqrt{s}$, and the total momentum zero. If the photon propagator is to obey energy and momentum conservation it must acquire a non-zero rest mass equal to $\sqrt{s}$ via a quantum fluctuation permitted by the Heisenberg uncertainty principle. The photon is described as being *off-mass-shell* or *virtual*, and the *virtuality* of the photon Q is in this case the effective rest mass of the photon, or viewed another way, the energy scale at which the quantum fluctuations which permit the interaction occur. Via the uncertainty principle this energy scale Q can be related to a distance scale of order $1/Q$, and it is common also to talk of the interaction occurring at short or large distance scales, depending on the size of Q. Note that since in this case $Q = \sqrt{s}$, the $e^+e^- \rightarrow \gamma^* \rightarrow \mu^+\mu^-$ scattering cross-section given in Equation 2.3 depends on $1/Q^2$: as the required virtuality of the photon propagator increases the cross-section decreases. It is also proportional to α^2, a consequence of there being two interaction vertices in the Feynman diagram considered.

However, Figure 2.3 is far from the only diagram that can be drawn for the $e^+e^- \rightarrow \gamma^* \rightarrow \mu^+\mu^-$ process. Just a few of the possible higher order diagrams are shown in Figure 2.5. These diagrams involve two additional interaction vertices, and thus have an associated amplitude involving a factor α^2. The first six are referred to as loop diagrams, because they contain virtual photon loops as shown, and since there is essentially no restriction on the number of photon loops which are permitted in the theory an infinite number of such diagrams can be drawn[2]. A

---

[2]Here we are considering the $e^+e^- \rightarrow \gamma^* \rightarrow \mu^+\mu^-$ process with zero photons in the final state. A more inclusive calculation would permit one or more *real* photons in the final state, and this would be more suitable for comparison to data. Such a calculation would involve a new type of higher order diagram that does not contain a photon loop but only one additional QED vertex,
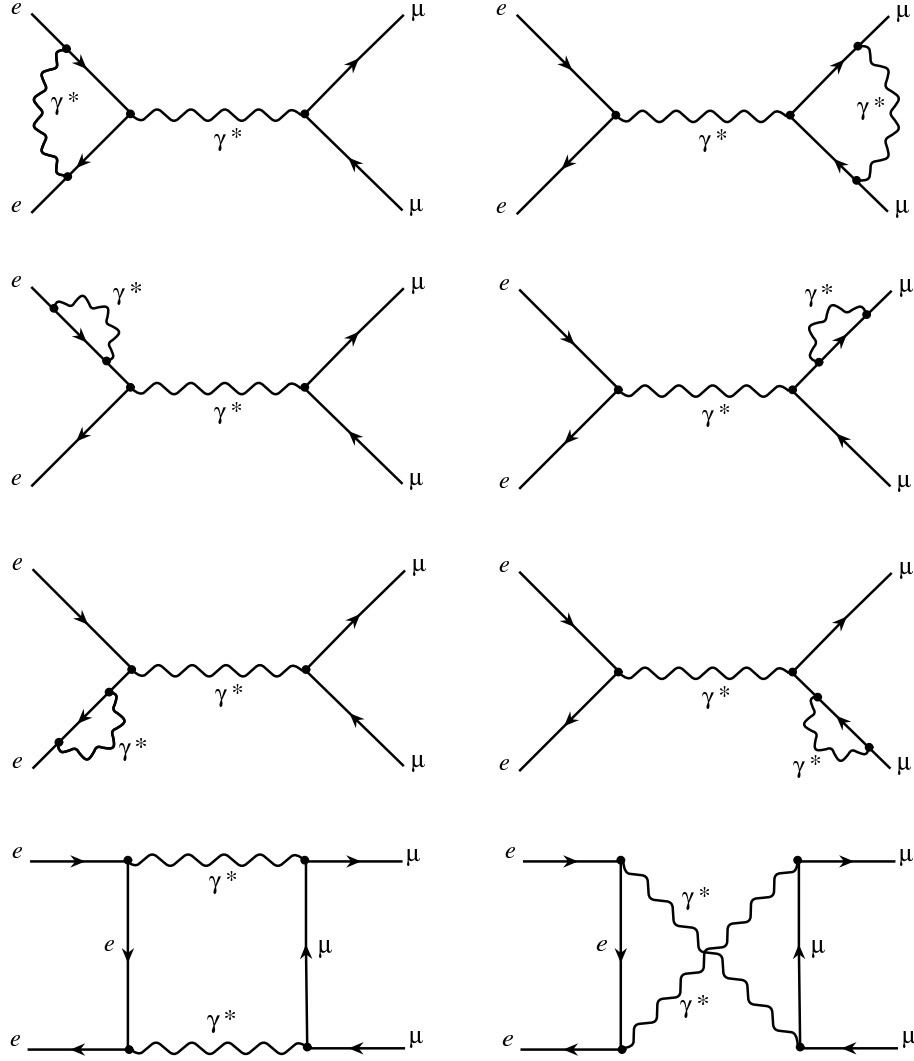
Figure 2.5: Higher order diagrams for the process $e^+e^- \to \gamma^* \to \mu^+\mu^-$.

complete evaluation of the cross-section would involve calculating the amplitude for every possible diagram, summing all these amplitudes to get the total amplitude $\mathcal{M}_{tot}$, squaring $\mathcal{M}_{tot}$ to get the interaction probability and then integrating over phase space. The total amplitude will be of the form:

$$\mathcal{M}_{tot} = A_1\alpha + A_2\alpha^2 + \text{higher orders ...} \tag{2.4}$$

where A_1 is $\mathcal{M}_1/\alpha$ and similarly A_2 is $\mathcal{M}_2/\alpha^2$, where $\mathcal{M}_2$ is the total amplitude of the diagrams in Figure 2.5. The interaction probability, directly proportional to the cross-section, is then given by:

$$|\mathcal{M}_{tot}|^2 = A_1^2\alpha^2 + 2Re(A_1 \cdot A_2)\alpha^3 + A_2^2\alpha^4 + \text{higher orders ...} \tag{2.5}$$

where a photon couples to one of the leptons in the initial or final state. These diagrams would have an associated amplitude involving a factor $\alpha^{3/2}$

This is known in QFT as a *perturbation series.* Here we have only defined three terms, but since the number of loop diagrams one can draw is infinite, the series is infinite also. Thus in practice when performing a cross-section calculation one has to terminate the perturbation series at some fixed order. Terminating the series at $\mathcal{O}(\alpha^2)$ would in this case be defined as the *leading order* (LO) cross-section calculation, and only involve evaluating the amplitude of the diagram in Figure 2.3. Terminating the series at $\mathcal{O}(\alpha^3)$ would be defined as the *next-to-leading-order* (NLO) calculation, and would involve evaluating the amplitudes of the diagrams in Figure 2.5, in particular the complex interference of these amplitudes with that of Figure 2.3. Similarly terminating the series at $\mathcal{O}(\alpha^4)$ would be the *next-to-next-to-leading-order* (NNLO) calculation, and so on[3].

Although theoretically the complete perturbation series is an exact expression for the total cross-section of a process, several issues remain if it is to be practically useful:

- Unless the coupling constant α is sufficiently small the series will not converge to a finite cross-section and termination of the series will be subject to large higher order corrections.

- When integrating the amplitudes over all phase space to obtain the cross-section we encounter a problem with loop diagrams. Within the loop the momentum of the photon propagator is unconstrained and thus the integral is divergent, known as *ultraviolet divergence.*

These issues will be addressed in the context of QED and QCD in the sections that follow.

2.3 Coupling Strength in QED

Recall that in the gauge theory of QED the couping α is proportional to the charge of the particle. The inherent uncertainties of quantum mechanics mean that the charge of a particle is not as well defined as you might expect. In QED quantum fluctuations allowed for by the Heisenberg uncertainty principle permit an electron to emit a photon with virtuality Q. This fluctuation occurs on a distance scale of order $1/Q$. The virtual photon can in turn produce a virtual e^+e^- pair, which in turn can radiate a photon, and so on, such that at sufficiently large distance scales the electron is surrounded by a virtual cloud of electron-positron pairs as illustrated in Figure 2.6. This cloud becomes polarized, with the virtual particles of opposing

---

[3]Note that our series expansion in Equation 2.5 is not actually complete to NNLO. Additionally one needs to consider two-loop diagrams which will interfere with the diagram of Figure 2.3 to yield $\mathcal{O}(\alpha^4)$ contributions.
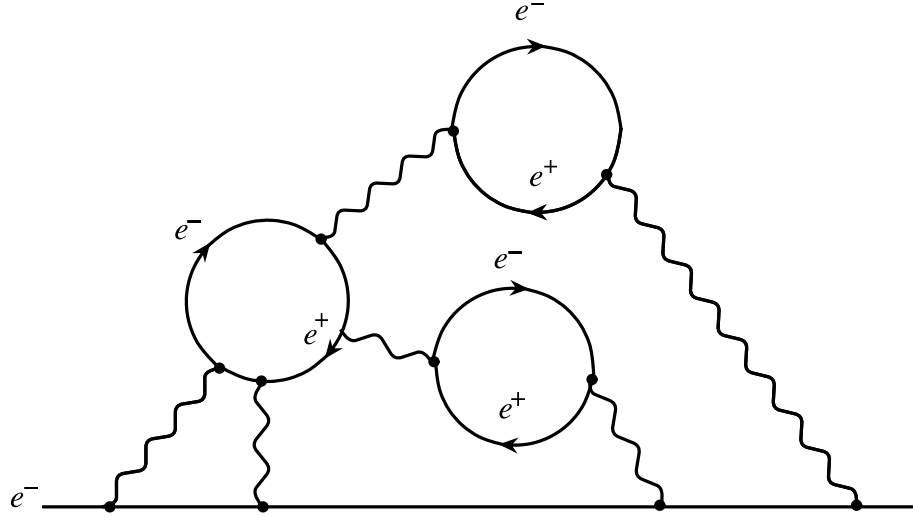
Figure 2.6: Diagram illustrating the formation of a virtual cloud of electron-positron pairs around a bare electron charge.

charge being attracted to the electron charge, and the virtual particles of like charge being repelled, as illustrated in Figure 2.7. The net effect is that the *bare* electron charge is screened and reduced. The implication of this is that the QED coupling strength α is dependent on the energy scale Q of the interaction. At large Q scales, short distance scales, less of the virtual cloud is resolvable and the coupling is larger than at small Q scales.

The theory of QED is renormalised (see Section 2.5) such that the coupling constant α is proportional to the experimentally measured charge e as follows:

$$\alpha = e^2/4\pi \qquad (2.6)$$

Effectively we have evaluated the coupling at $Q = 0$; this is the fine structure constant $\alpha^{-1} \approx 137$. As the energy scale Q decreases the coupling strength decreases asymptotically to the fine structure constant. Thus in the context of QED all energy scales currently achievable in experiments can be considered small, that is, the interactions occur on a large distance scale such that the coupling α is $<< 1$. This makes perturbation theory a very effective tool for QED predictions, as α is always small enough that a perturbation series will converge after only a few terms.

2.4 Asymptotic Freedom and Confinement in QCD

In QCD the coupling strength is proportional to the colour charge of the particle. Similarly to QED, a bare colour charge such as a quark or gluon will be surrounded by a virtual cloud of quark-antiquark pairs, illustrated in Figure 2.8, that form on a distance scale $1/Q$, where Q is the energy of the interaction. However, since gluons
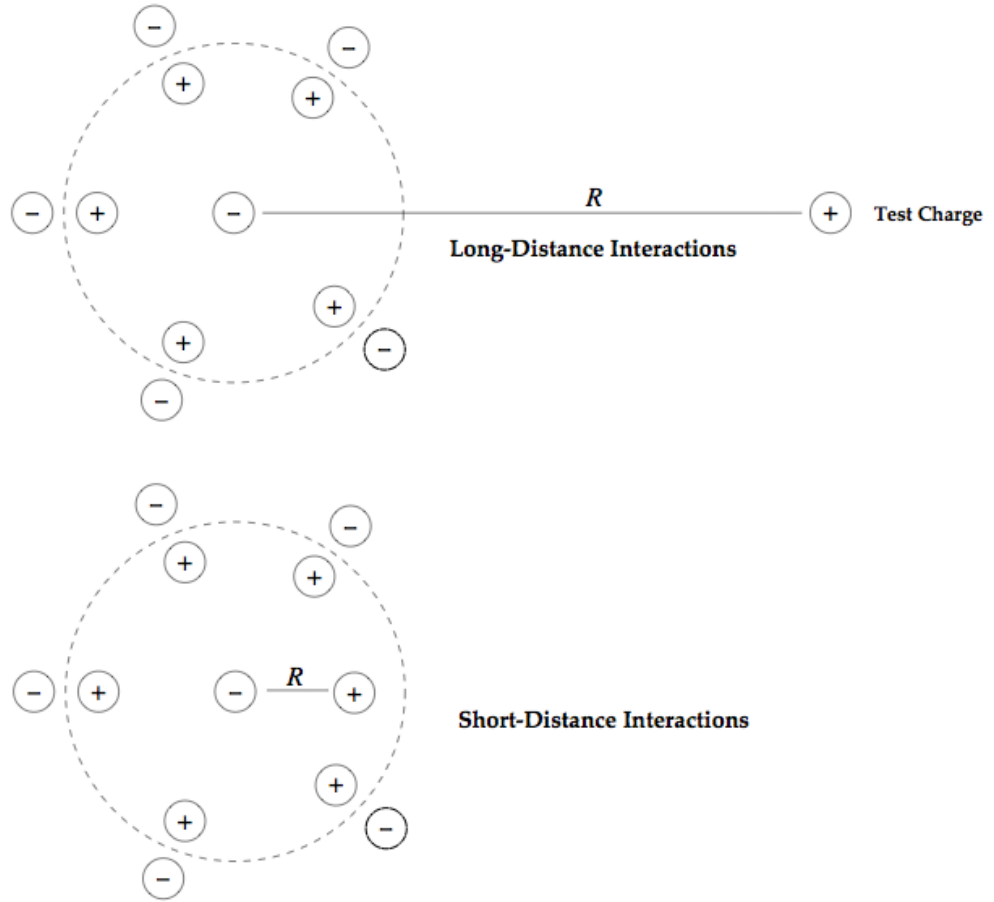
Figure 2.7: Diagram illustrating the effect of QED charge screening by virtual electron-positron pairs.

are self-interacting, the cloud will also contain gluon-gluon pairs which themselves carry colour charge, also shown in Figure 2.8. This is a crucial difference with respect to QED. It means that the net effect of the virtual cloud in QCD is not to screen and reduce the colour charge, but to increase it.

Thus in QCD as the energy scale Q decreases, the resolved colour charge, and hence the strength of the strong coupling, increases. Taken to it's limit, over an infinitely large distance scale the QCD coupling strength, and hence potential energy, would itself be infinite. The massless property of the gluon means that the strong force has infinite range, and thus it is always preferential to create lower energy colourless bound states of quarks. This is the property of *colour confinement*; that the physical states in QCD are not the coloured quarks and gluons but colourless bound states of three quarks, the hadrons (rgb or $\overline{rgb}$), or two quarks, the mesons ($r\bar{r}$,$g\bar{g}$ or $b\bar{b}$). Final-state quarks and gluons created in hard-scatter interactions at hadron colliders undergo a process called *hadronization* or *fragmentation* [72] and are observed as jets of hadrons in the detector. This is described in detail in Chapter 3.
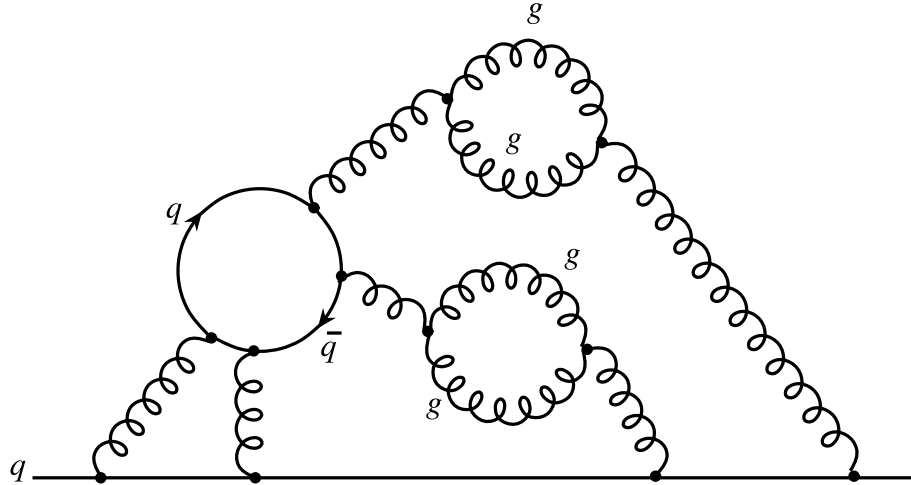
Figure 2.8: Diagram illustrating the formation of a virtual cloud of quark-antiquark pairs around a bare electron charge.

Hadronization is a *soft* QCD process, meaning that it occurs at a low Q scale where the strong coupling constant α_s is too large for perturbation theory to be possible. The perturbation series for the process will not converge to a finite result. At higher Q scales the coupling constant becomes small enough for the perturbation series to converge and perturbative predictions of *hard* processes can be made. This is known as *asymptotic freedom*. Precisely what scales we are referring to when we talk of hard and soft processes is detailed in the next section.

Figure 2.9 shows the running of α_s with scale $Q = \mu$. As Q increases the coupling decreases. One can see that at the energy scale of the Z boson mass $\alpha_s \approx 0.12$, and that this is not going to reduce significantly at higher energies. This is roughly one order of magnitude larger than the QED coupling constant, and as a consequence the termination of series in a QCD perturbative calculation is subject to more significant higher order corrections than in QED.

2.5   Renormalisation

In theory the charge screening effects of QED and QCD could be calculated using an infinite series of Feynman diagrams containing photon propagator loops like those of Figure 2.5. Unfortunately, as we have already noted, these diagrams themselves contain ultraviolet divergences. Renormalisation is the introduction of one or more dimensionless physical parameters into the theory, dependent on some unphysical *renormalisation scale*, which effectively absorbs these ultraviolet divergences and thus retrieves the ability to make perturbative calculations [37, 72, 86].

In QED the renormalisation parameter is the experimentally measured charge of the electron e. We know that this quantity must take into account the ultraviolet

divergences, as effectively it measures the electron charge at $Q = 0$. Therefore if we introduce the measured charge e into the theory as a parameter we can effectively absorb these ultraviolet divergences. This is done by defining the QED coupling constant at $Q = 0$ as the fine structure constant (Equation 2.6).

It is important to note that the choice of energy scale at which the theory of QED can be renormalised is not unique. Choosing $Q = 0$ is simply convenient since the fine structure constant can be measured very precisely at this scale. If we were to evaluate the electric charge at a higher scale $Q > 0$, we would, as we have already seen, resolve a larger electric charge and hence have a larger QED coupling constant. This implies that by introducing a renormalisation parameter (evaluated at a particular scale) into a quantum field theory one must also necessarily introduce a dependence of this parameter on a renormalisation energy scale Q. As we have already noted, the coupling strength varies with Q, the so-called *running coupling constant*. The dependence of a renormalised coupling constant on Q is determined by the *renormalisation group equation* (RGE) [86]:

$$Q^2 \frac{\delta \alpha}{\delta Q^2} = \beta(\alpha) \tag{2.7}$$

Where the β function can be determined to a certain order using perturbation theory, thus allowing for a solution $\alpha(Q^2)$ to a certain order.

In QCD the property of confinement means that we cannot renormalise the theory at $Q = 0$. Instead we have to measure the strong coupling constant in processes that occur at large Q scales, where perturbation theory can be applied. The leading order solution to the RGE (Equation 2.7) [86] gives the dependence of the strong coupling α_s on the renormalisation scale Q^2 as:

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2n_f) ln(Q^2 / \Lambda_{QCD}^2)} \tag{2.8}$$

where n_f is the number of active fermion flavours and Λ_{QCD} is a scale parameter that must be determined from experimental measurement of α_s at known Q scale. Such measurements yield $\Lambda_{QCD} = 0.2 \pm 0.1 \text{GeV}$. Physically Λ_{QCD} may be interpreted as the energy scale at which perturbation theory breaks down, as $\alpha_s \to 1$. Perturbative evaluation of QCD processes can only be made if $Q >> \Lambda_{QCD}$. Thus in some sense Λ_{QCD} is the boundary between *soft* and *hard* QCD processes. Figure 2.9 shows the running of α_s with scale $Q = \mu$. The points are experimental measurements of α_s at particular renormalisation scales, and the curve is a NNLO evaluation of the running of α_s using the average Λ_{QCD} determination of these results [70].
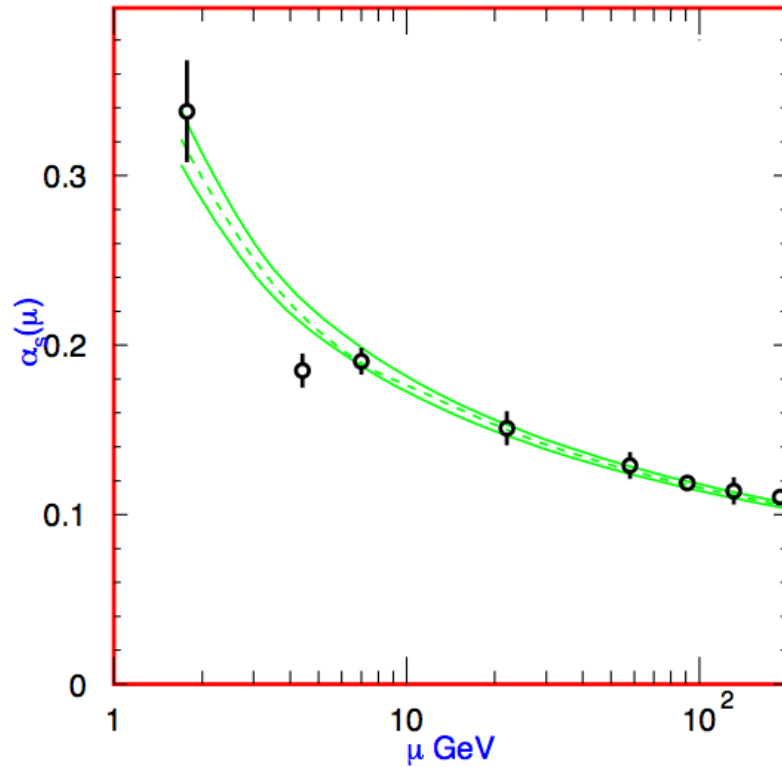
Figure 2.9: Graph showing the running of the strong coupling constant with renormalisation scale $Q = \mu$. The points show the experimental determination of α_s at a particular Q^2 scale, and the curve is a NNLO evaluation of the running of α_s using the average Λ_{QCD} determination of these results. Taken from [70].

2.6   Unanswered Questions

The Standard Model is a hugely successful description of the fundamental particles that make up the Universe and the interactions between them. However, a number of important experimental observations are not explained by the SM. Additionally, at the present time not all of the components of the SM have been verified experimentally.

Currently there exists no direct experimental evidence for the final fundamental particle predicted by the SM, the Higgs boson. However, precise measurement of SM parameters constrain the allowed Higgs mass in the SM framework. For example, precision measurement of the Top quark and W boson masses allow one to set limits on the Higgs mass, as shown in Figure 2.10. Electroweak measurements such as these have constrained the Higgs mass to be lower than about 175 GeV/c^2, and direct searches at the LEP experiments have constrained the Higgs mass to be greater than 114 GeV/c^2. Failure to discover the Higgs boson within this mass range would imply the existence of physics outside the Standard Model framework. A key goal of the CDF Run II experiment is to further constrain the predicted Higgs mass with increased precision measurements of the W boson and Top quark masses.
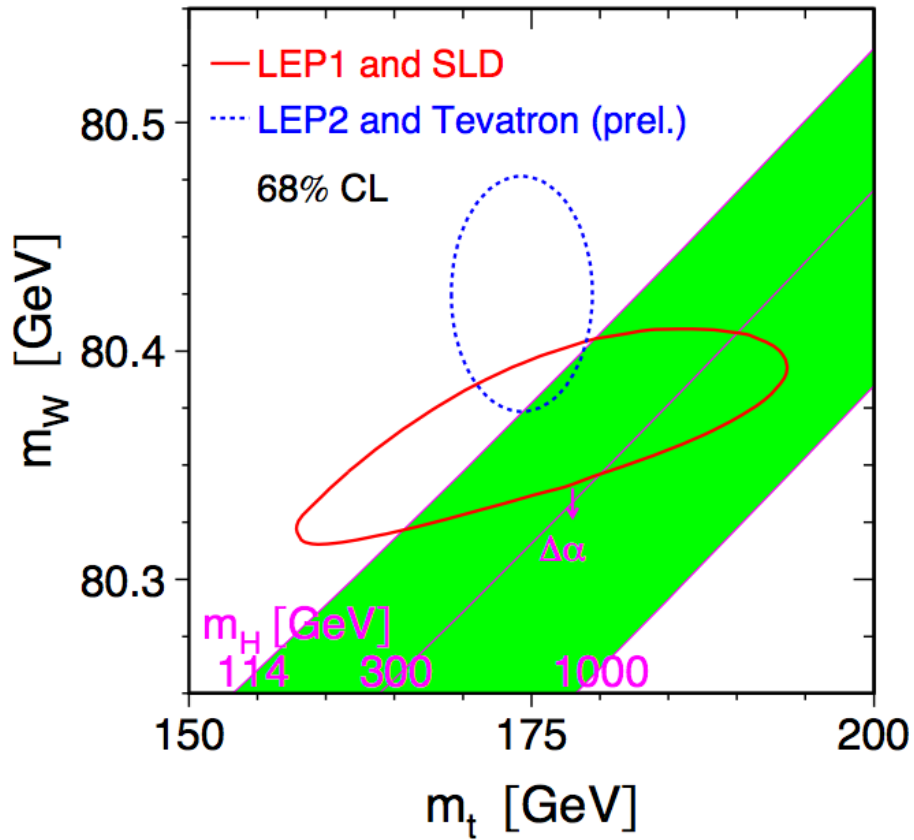
Figure 2.10: Figure showing the 68% confidence level results of the Top quark and W mass measurements along with Higgs mass contours allowed by the Standard Model. The indirect measurements from LEP I and SLD (solid contour) are compared to preliminary direct measurements from LEP II and Tevatron CDF and D0 experiments (dashed contour). Taken from [26].

Unfortunately the cross-section for Higgs production at the Tevatron is predicted to be rather small and sufficient data for a discovery is unlikely to be collected before the more powerful Large Hadron Collider (LHC) experiment begins data taking in 2008.

Discovery of the Higgs boson would undoubtedly be another huge triumph for the Standard Model, but it would not complete the story. Ever since the existence of the Higgs boson has been postulated it has been known that loops of virtual SM particles will introduce quantum corrections to the Higgs mass that could be as large as the Planck scale $M_{Planck} \approx 10^{19} GeV$. How then do such contributions cancel so completely such that the Higgs mass is of the order 100 GeV? This is known as the *hierarchy problem*. Solutions to this problem require the introduction of new physics beyond the SM framework presented in this chapter. The most popular extension to the SM is the theory of *supersymmetry* [109], in which each SM particle has a supersymmetric partner with opposite spin statistics allowing precise cancellation. Others include the recently proposed "Little Higgs" models [32], which

introduce new particles which achieve the cancellation with same sign spin statistics contributions, and theories of extra dimensions [64, 113].

Closely linked to the hierarchy problem is how the fundamental force of gravity fits into the SM. In particular why is the Planck scale so much larger than the electroweak scale, or put another way, why is gravity so weak? In fact the SM does not predict the coupling strengths of any of the fundamental forces, just as it does not predict the masses of the fundamental particles, nor the existence of three generations of fermions. These are all parameters of the model that have to be established through experimental observation. Is there some theory underlying the SM within which these parameters can be predicted?

There are also recent experimental observations which challenge the SM picture of the Universe. A huge wealth of evidence, from the rotation of spiral galaxies to the cosmic microwave background measurements of the WMAP satellite, point to a Universe dominated by cold dark matter, that is, non-luminous, non-baryonic matter. Additionally observation of neutrino flavour change [78, 79] imply crucially that neutrinos must have mass, in direct contradiction to the massless SM prediction.

Chapter 3

The W + Jets Process at Hadron Colliders

3.1 Motivations for Study

Since its discovery at CERN's SPS p$\overline{\text{p}}$ collider in 1983 [33] the W boson has been the subject of many detailed studies. In the 1980's the UA1 [25] and UA2 [121] experiments made the first measurements of W properties such as the inclusive W production cross-section $\sigma(p\overline{p} \to W) \times B(W \to l\upsilon)$ at $\sqrt{\text{s}} = 630$ GeV, the W mass and W width [25, 121]. In the 1990's the LEP experiments [1, 23, 62] made improved mass and width measurements [117], and the first measurements of the W decay branching ratios and W pair production cross-sections [101]. At the Tevatron both the CDF and D0 experiments have measured the inclusive W production cross-section at $\sqrt{\text{s}} = 1.8$ TeV [5, 11] and at $\sqrt{\text{s}} = 1.96$ TeV [3, 4, 17]. Measurements of the W mass and width [88] have also been made at the Tevatron, although currently these are not at the same level of precision as those at LEP. Measurements of the W inclusive cross-section at hadron-hadron colliders are made primarily to test the impact of QCD effects on production. As was discussed in Section 2.6, precision measurements of the W mass, when combined with Top mass measurements, provide important constraints on the allowed Higgs mass within the SM [26]. Knowledge of the W mass and Fermi coupling constant allow the width of the W to be precisely calculated within the SM, and thus measurements of this quantity provide a further important consistency check of the SM.

In this study we are interested in events where a W boson is produced in association with one or more final state quarks or gluons, which manifest themselves as jets of hadrons in the detector. Although such a process will inevitably involve a contribution from an electroweak interaction vertex, the emphasis is on the properties of the jets which are produced in strong interactions. The production of the W boson in the hard scatter ensures that the interaction occurs at a sufficiently high Q

scale that perturbation theory can be applied to these interactions. In addition, one can identify with relative ease events where the W decays leptonically to an electron or muon with an accompanying neutrino (although here we restrict ourselves to the electron decay channel only). The W + Jets process is thus a convenient "laboratory" for testing the predictions of QCD. The simplest observable one can think of forming is the cross-section for W production in association with at least n jets above a certain energy threshold. However, one can investigate QCD predictions in more detail if one looks at less inclusive observables. In this study we measure the dependence of the W$\rightarrow$ eυ + $\geq n$ jet cross-section on the kinematic properties of the jets in the event; their energy spectra, separation in $\eta - \phi$ space and dijet invariant mass.

It is interesting to test in detail the predictions of QCD simply from the point of view of testing our understanding of the universe, and of course there is always the potential for observing the effects of new physics if we see a significant deviation. However, there are also more practical motivations for testing these predictions. The final state of a W boson accompanied by high energy jets is identical to that of many of the most important processes that one hopes to measure or discover in the coming years at both the Tevatron and imminent LHC experiments [68,69]. In order to make precise measurements of these processes or claim a discovery, one has to be able to separate in some way the signal from this W + Jets "background". In practice this will require an accurate model of the kinematic structure and/or rate of W + Jets events, and a realistic estimate of the systematic uncertainties associated with such models can only be obtained via direct comparisons with real W + Jets data. In addition these comparisons could also help one to tune parameters of the models such that they provide an improved description of W + Jets, resulting in a correspondingly reduced systematic.

A key goal of CDF Run II is to make increasingly precise measurements of the properties of the Top quark, namely the Top mass and pair production cross-section, which help us to test the consistency of the SM and, in the case of the mass measurement, constrain the allowed Higgs mass range. Figure 3.1 illustrates the "lepton + jets" channel of t$\bar{\text{t}}$ production at the Tevatron. The Top quark is so massive that on production in the hard scatter it immediately decays with $\sim 100\%$ branching ratio to a beauty quark and W boson. In the lepton + jets channel, in which the most precise measurements of Top properties can be made, one W decays leptonically and one hadronically. Thus the final state consists of two jets resulting from b quarks, at least two light quarks jets, and a lepton and neutrino. Identification of Top events at CDF is made easier by the use of a secondary vertex-finding algorithm [18], which identifies with reasonable efficiency ($\sim 50\%$) b quark jets. However, some fraction of W + Jets events will naturally contain heavy flavour
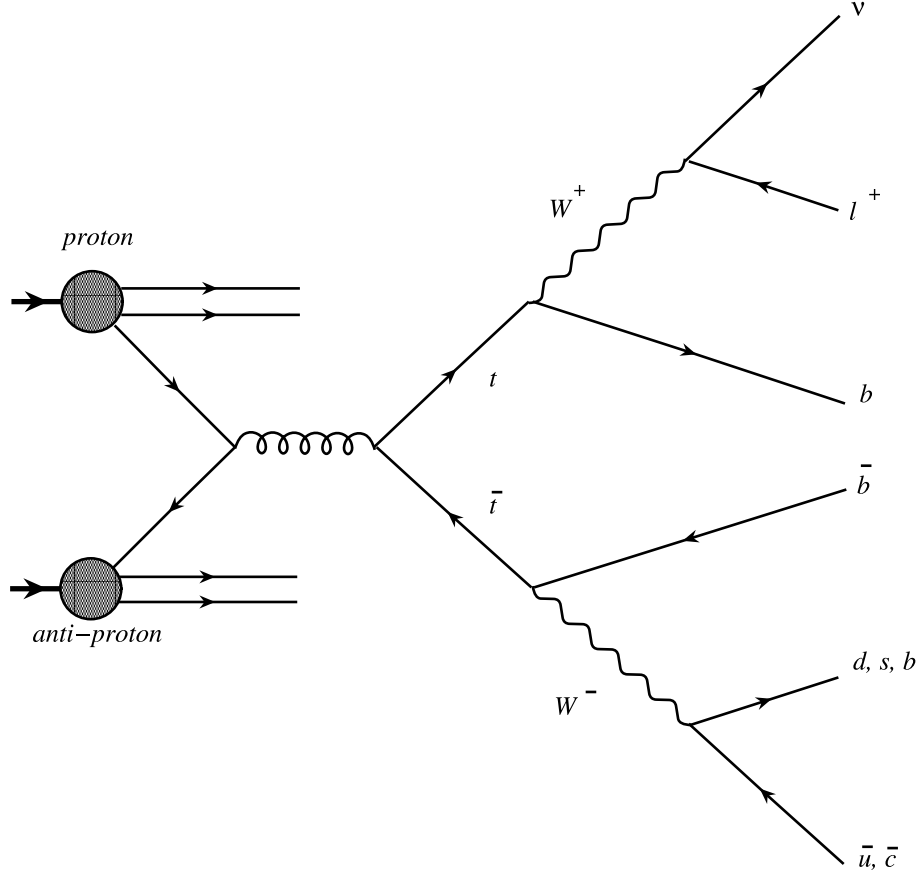
Figure 3.1: Diagram illustrating the lepton + jets channel of $t\bar{t}$ production at the Tevatron.

quarks produced directly in the hard scatter, and in addition there is the non-negligible possibility that light quark jets are misidentified as b quark jets. In the case of the Top mass the theoretical modelling of the background kinematics forms a crucial component of the measurement [15]. Similarly the W + Jets process is a crucial background in searches for single Top production at the Tevatron.

   At the Tevatron and future LHC experiments a key aim is to search for the Higgs boson and, if discovered, measure its properties as precisely as possible. There are four main SM Higgs production channels, illustrated in the diagrams of Figure 3.2. The predicted pp $\rightarrow$ H + X cross-sections of each process vary with the assumed Higgs mass as shown for both Tevatron and LHC energies in in Figure 3.3. If the Higgs is of low mass ($M_H \leq 135$ GeV) it will decay predominantly to $b\bar{b}$, but at higher mass values decay to WW and ZZ become dominant. The consensus is that Higgs channels with leptonic or photonic final states are the only ones which offer any chance of Higgs discovery because of the problem of overwhelming multijet backgrounds. At the Tevatron the most promising channel for discovery is, depending on the Higgs mass, gluon fusion where the Higgs decays to WW^*, or vector boson associated production where the Higgs decays to $b\bar{b}$ [122]. For the second channel in
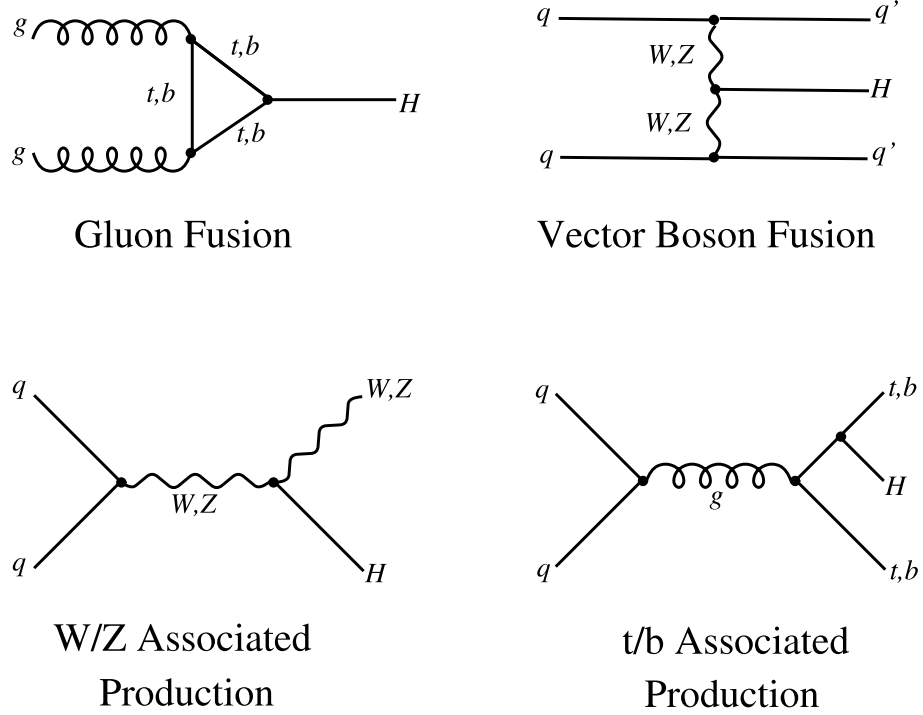
Figure 3.2: Diagrams illustrating the main SM Higgs production channels.

particular the W/Z + Jets process is a major background and thus requires under-standing in the same detail as is required for Top analyses. At the LHC the picture is not so clear cut, with the increased luminosity reach enabling some of the lower branching ratio modes, such as $H \to \gamma\gamma$, to play an important role.

There are also more indirect ways in which W + Jets studies can be beneficial to future Higgs searches. The vector boson fusion (VBF) channel for Higgs production (top right diagram of Figure 3.2) is characterized by two high energy jets in the forward direction (i.e. close to the beam direction) and, provided the Higgs decays non-hadronically, an absence of hadronic activity in the central region [36]. By studying the rate for such topologies in W + Jets events, we can get an idea of how significant W + Jets and Z + Jets backgrounds will be in VBF Higgs production. However, such studies are beyond the scope of this thesis.

3.2   Parton Distribution Functions

The theory of QCD can be used to predict the cross-section for quark-quark in-teractions so long as the energy scale of the interaction is sufficiently large that a perturbative expansion will converge. However, confinement means it is not possible to collide individual quarks together. In order to produce hard interactions between quarks one has to collide hadrons together with large centre-of-mass energies. At the Tevatron, one collides protons and antiprotons at $\sqrt{s} = 1.96$ TeV. Thus if one is to
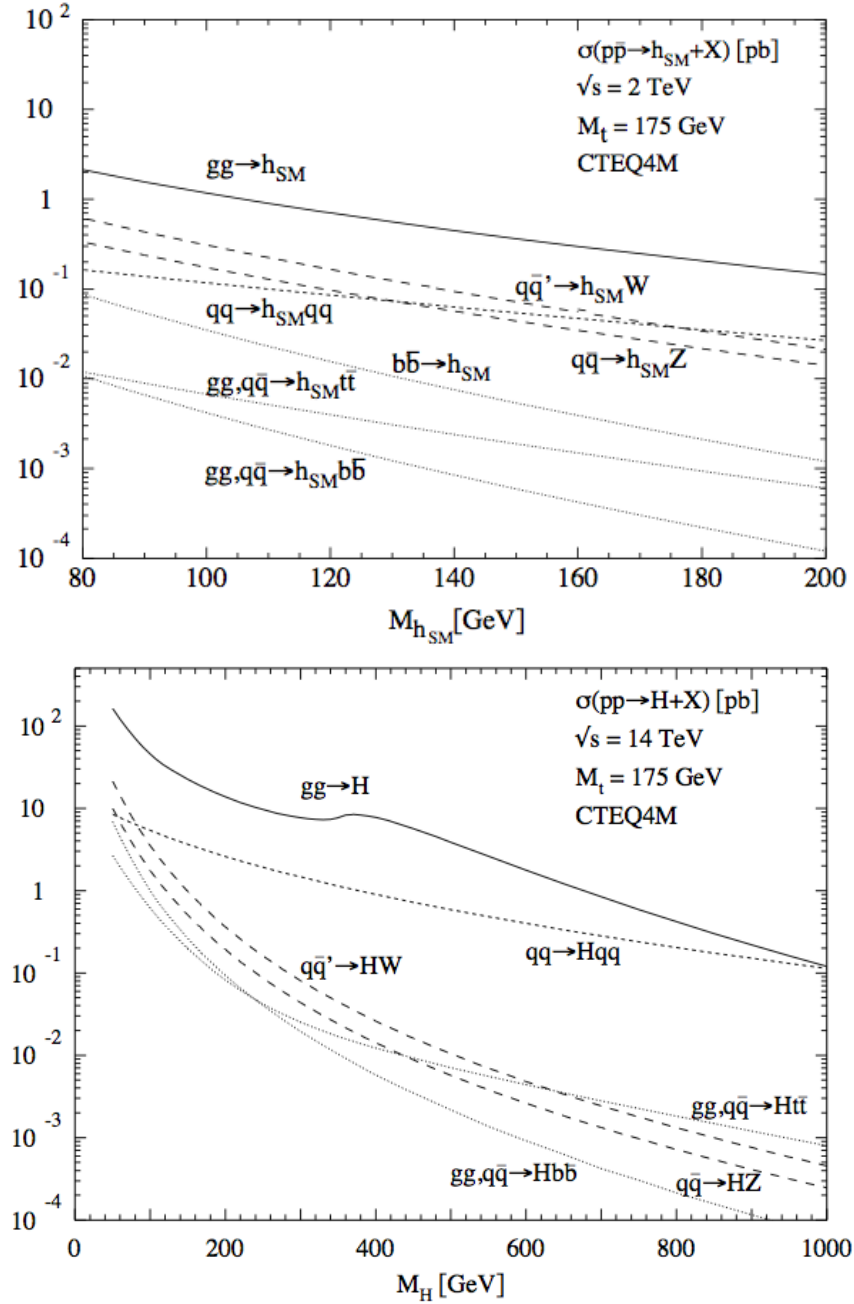
Figure 3.3: The predicted $pp \to H + X$ cross-sections of various Higgs production processes at the Tevatron (top, taken from [122]) and LHC (bottom, taken from [120]) collider energies.

compare theory to experimental data, one has to be able to predict cross-sections for proton-antiproton interactions. This involves some understanding of the complex structure of the proton.

Baryons are strongly bound states of three quarks, the proton comprising of *uud* quarks, known as the *valence* quarks. The interactions between the valence quarks occur at low energy scales of the order of the hadron mass, corresponding to timescales $t_{had} \sim \hbar/E_{had}$. Over such timescales virtual quarks and gluons will be created and destroyed via quantum fluctuations, and the proton thus consists of these virtual particles as well as the valence quarks, which are together known as *partons*. These interactions are soft and non-perturbative, and thus the partonic structure of the proton is not predictable in QCD, but has to be determined experimentally. In deep inelastic scattering (DIS) experiments [72] the proton is probed via QED interactions with a high energy electron. A schematic diagram for DIS is shown in Figure 3.4. The energy scale of the interaction Q, the exchanged four-momenta of the photon "probe", can be determined from the momenta of the scattered electron. Interactions with $Q >> E_{had}$ occur with a timescale $t_Q << t_{had}$, short enough that the probe will resolve not only the valence quarks but also the virtual quarks and gluon *sea*. As the energy of the interaction decreases the timescale increases and thus less of the virtual component of the proton is resolvable. The structure of the proton is thus dependent on the energy scale of the interaction.

In the *parton model* [72] the structure of the proton is specified by a set of *parton distribution functions* (PDFs), $f_i(x, \mu_F^2)$, that describe how the total momentum of the proton is distributed amongst the constituent partons. The effective number of partons of type i with a fraction of the proton momentum between x and $x + dx$ is given by $f_i(x, \mu_F)dx$, where μ_F is the *factorisation scale*, the energy scale at which the proton structure is resolved. PDFs are determined from fits to DIS data and are universal and process independent [112]. Although they cannot be derived from first principles, their evolution as a function of μ_F is predictable in perturbation theory, in much the same way that the evolution of α_s with Q is predictable [29, 67, 82]. Figure 3.5 shows the PDF (actually $xf_i(x, \mu_F)dx$) of the proton for $\mu_F^2 = 20 GeV^2$ and $\mu_F^2 = 10^4 GeV^2$ [105]. At lower μ_F the structure of the proton is dominated by the gluon constituents but the u and d valence structure is still important at high x. However, at higher μ_F we see that the gluons become even more dominant, as well as increased importance of c,s and b sea quarks.

3.3 Perturbative Predictions of W+Jets

Let us begin by considering the cross-section for inclusive W production at a proton-antiproton collider, $p\bar{p} \rightarrow W + X$, where X can be anything. The "simplest"
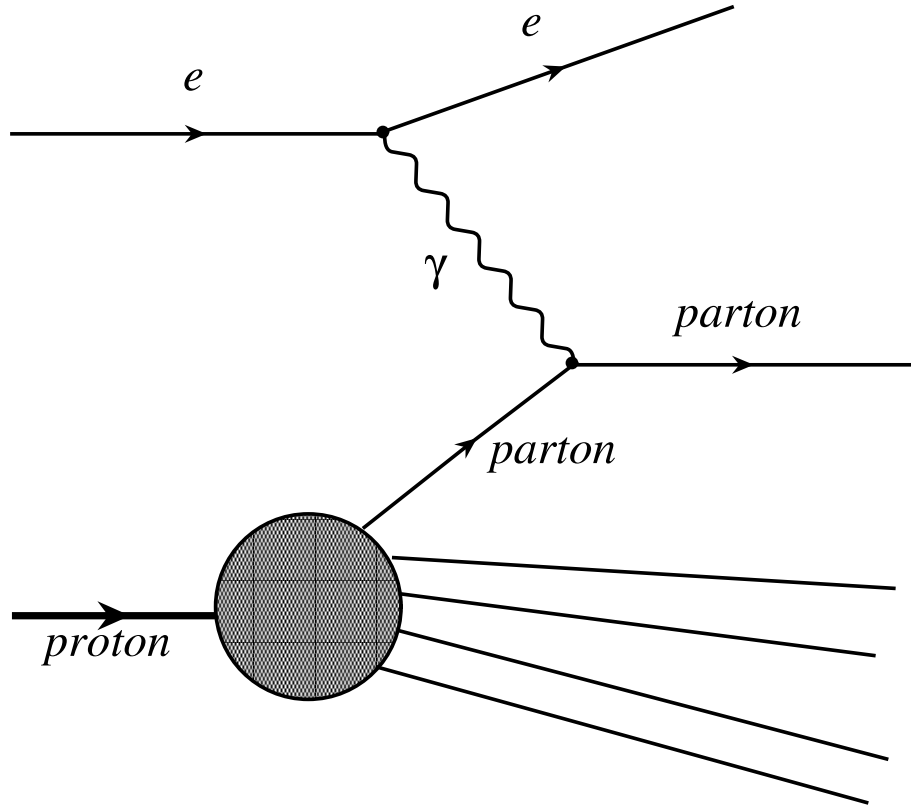
Figure 3.4: A schematic diagram of the deep inelastic scattering (DIS) process.
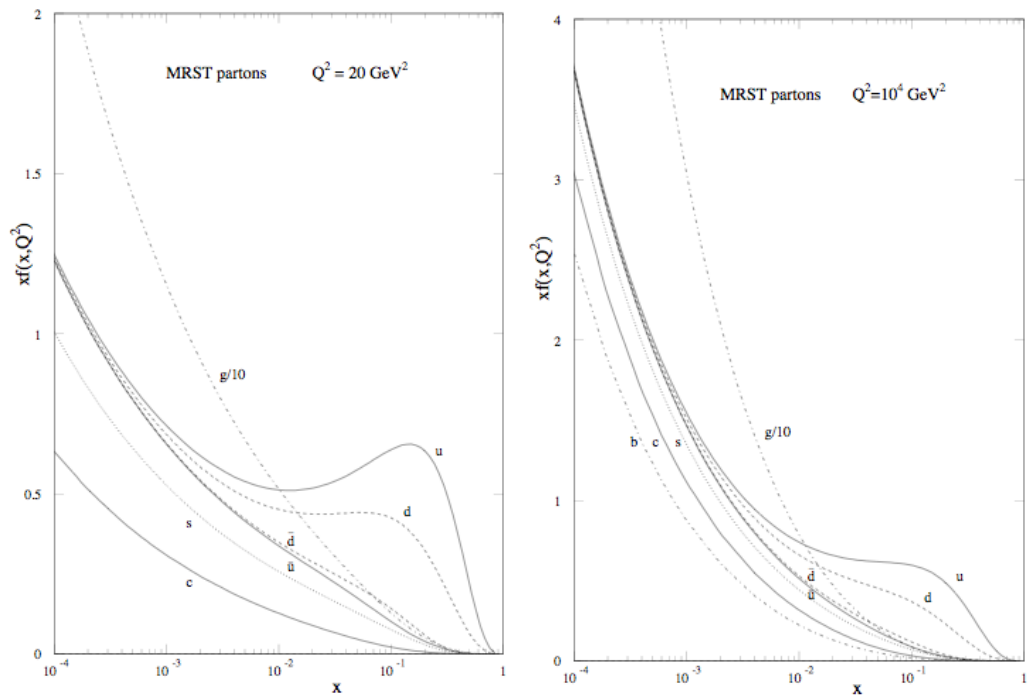
Figure 3.5: The PDFs of the proton at two different factorisation scales, $\mu_F^2 = 20 GeV^2$ and $\mu_F^2 = 10^4 GeV^2$. Taken from [105].
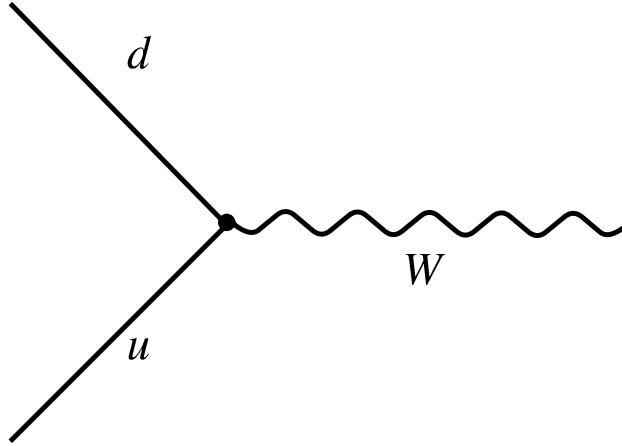
Figure 3.6: Born-level diagram for inclusive W production at a hadron collider

process by which a W boson can be produced at a hadron-hadron collider is shown in Figure 3.6. An (anti)quark from the proton "fuses" with an (anti)quark from the antiproton with sufficient energy to create a W boson. This diagram involves one electroweak vertex, but no QCD interaction vertices. It is the lowest-order (or *Born-level*) diagram in an infinite hierarchy of possible diagrams for W inclusive production with increasing numbers of QCD vertices[1]. Some of the possible higher order diagrams are shown in Figure 3.7.

The $p\bar{p} \to W + X$ cross-section can thus be expanded as a perturbation series in α_s :

$$\sigma_{p\bar{p} \to W+X} = A_0 + A_1\alpha_s + A_2\alpha_s^2 + A_3\alpha_s^3 + ... \tag{3.1}$$

Where the coefficients A_n are calculated using the Feynman diagrams that contribute at that order. Note that the creation of a real W in the interaction ensures that the energy scale Q of the process is large enough that perturbation theory can be applied. However, the exact choice of renormalisation scale at which to evaluate α_s is not clear.

3.3.1 Leading Order $p\bar{p} \to W + X$ Cross-Section A_0

The Born-level diagram is the only diagram that contributes to the leading order W inclusive cross-section A_0. The leading order $q\bar{q} \to W + X$ cross-section, calculated from the Born-level diagram using electroweak theory, is [72]:

$$\hat{\sigma}_{q\bar{q}' \to W+X} = \frac{\pi}{3}\sqrt{2}G_F M_W^2 |V_{qq'}|^2 \delta(\hat{s} - M_W^2) \tag{3.2}$$

---

[1]Higher order diagrams involving QED vertices will also contribute to the inclusive W cross-section. However, the QED coupling strength is significantly smaller than that of QCD and thus the QCD diagrams dominate.
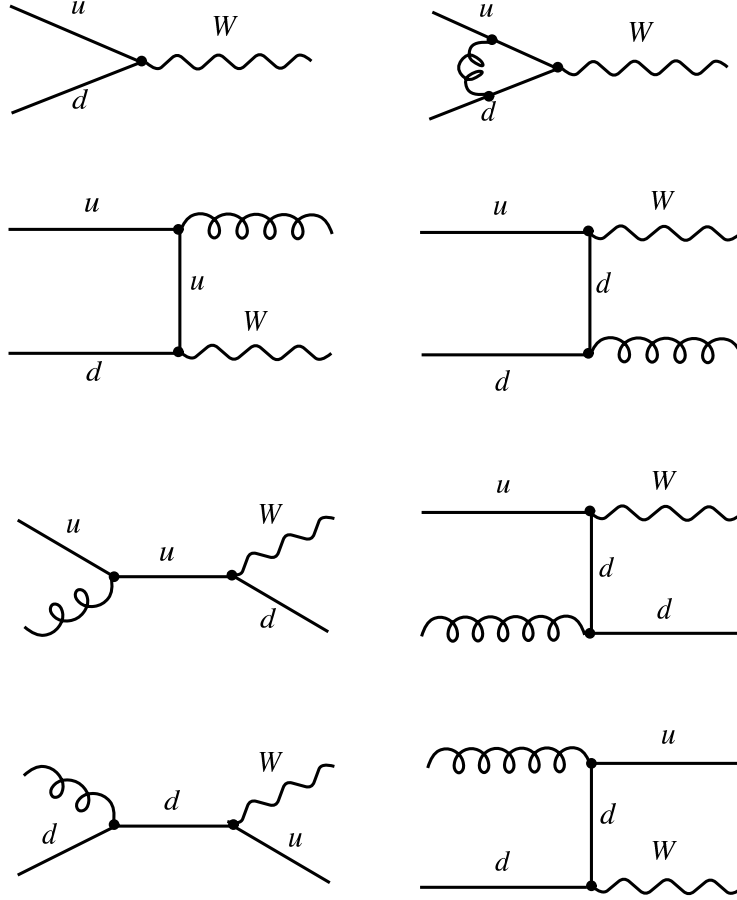
Figure 3.7: Next-to-leading-order diagrams for inclusive W production.

where G_F is the Fermi coupling constant, M_W is the mass of the W boson, $\sqrt{\hat{s}}$ is the centre-of-mass energy of the $q\bar{q}$ pair and $V_{qq'}$ is the appropriate CKM factor for the quarks involved. In this calculation one treats the W boson as a stable particle of fixed mass, which is a reasonable assumption given that the W width (Γ_W = 2.15 GeV) is much less than its mass (M_W = 80.41 GeV). The delta function is thus present to impose the resulting kinematic restriction of $\hat{s} = M_W^2$.

At a collider with a fixed centre-of-mass energy $\sqrt{s}$ in the $p\bar{p}$ frame, the fractions of the proton and antiproton momentum, (x_p and $x_{\bar{p}}$ respectively) carried by each interacting quark completely determines the centre-of-mass energy of the $q\bar{q}$ system. Assuming $\sqrt{s}$ is so large that the proton mass is negligible we can write $\hat{s} = x_p x_{\bar{p}} s$. Thus in order to calculate the leading order $p\bar{p} \rightarrow W + X$ cross-section, one must use this relationship along with the PDFs to integrate equation 3.2 over all possible x values and quark flavours:

$$\sigma_{p\bar{p}\rightarrow W} = \sum_i \int_{x=0}^{1} \left[f_i^q(x_p) f_i^{\bar{q}}(x_{\bar{p}}) + f_i^{\bar{q}}(x_p) f_i^q(x_{\bar{p}}) \right] \hat{\sigma}_{q\bar{q}'\rightarrow W} dx_p dx_{\bar{p}} \qquad (3.3)$$

Where $f_i^q(x_p)$ is the parton density for a quark of flavour i in the proton and $f_i^{\bar{q}}(x_{\bar{p}})$

is the parton density for an antiquark of the same flavour in the antiproton. The second term in the square brackets allows for the probability that the antiquark comes from the proton and the quark from the antiproton.

3.3.2 Higher Order Corrections to the $p\overline{p} \rightarrow W + X$ Cross-Section

Figure 3.7 shows the diagrams whose amplitudes must be evaluated in order to calculate the W inclusive cross-section at NLO[2] i.e. to evaluate A_0 and A_1 and terminate the series at $\mathcal{O}(\alpha_s)$. Included in this set of diagrams is the Born-level diagram, six diagrams which have one QCD vertex and will result in the addition of a parton in the final state, and a loop diagram which will not result in a parton in the final state. The $\mathcal{O}(\alpha_s)$ contribution from the diagrams which have a final state parton is known as the real contribution to the NLO calculation. The $\mathcal{O}(\alpha_s)$ contribution resulting from the interference of the loop diagram with the Born-level diagram is known as the virtual contribution.

However, problems arise when it comes to evaluating the NLO cross-section. As was noted in Section 2.2 loop diagrams contain ultraviolet divergences that are regulated by renormalising the theory of QCD. Additionally the loop diagram and one parton diagrams both contain *infrared divergences*, the result of the divergent behavior of the QCD coupling at low energy scales. Fortunately, renormalisation means that the divergence in the loop diagram exactly cancels with the divergence in the one parton diagrams to yield a finite $\mathcal{O}(\alpha_s)$ contribution to the NLO cross-section.

We can take this a step further and consider W inclusive production at NNLO. This calculation would include real contributions from tree-level diagrams with 2 final state partons, the infrared divergences of which would cancel with loop diagrams deriving from the tree-level 1 parton diagrams, providing a description up to $\mathcal{O}(\alpha_s^2)$. However, as one attempts to calculate at higher and higher orders the number of tree and loop diagrams that need to be considered increases, and arranging for the cancellation of divergences between these diagrams becomes increasingly complex. The NNLO result for the inclusive W cross-section, 2687 ± 54pb [87], represents the current limit of such calculations.

3.3.3 Leading Order $p\overline{p} \rightarrow W + n\ parton$ Cross-Sections

Thus far we have considered the W inclusive cross-section. A NLO calculation provides a description of one parton final states, and similarly a NNLO calculation

---

[2]Strictly these are not all the diagrams. One should also consider the set of charge conjugated diagrams and diagrams for the other quark flavours.
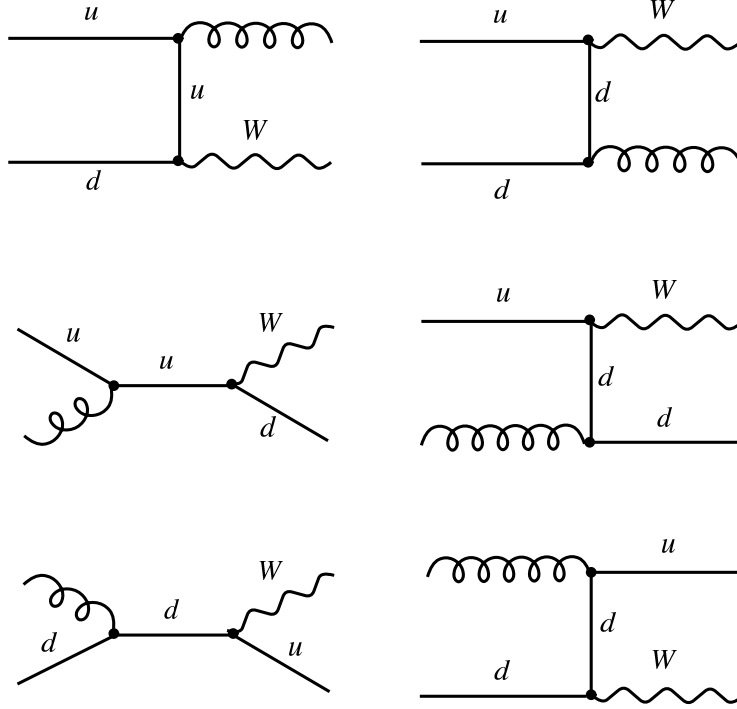
Figure 3.8: Leading-order diagrams for W + 1 parton production.

would additionally describe two parton final states.

However, we can also consider the process $p\bar{p} \to W + np$, explicitly requiring n final state partons in addition to the W. In general this can be expanded as a perturbation series:

$$\sigma_{p\bar{p}\to W+np} = B_n\alpha_s^n + B_{n+1}\alpha_s^{n+1} + B_{n+2}\alpha_s^{n+2} + ... \tag{3.4}$$

where the coefficients B_n are calculated using the diagrams that contribute at that order, and $B_n \neq A_n$.

Figure 3.8 shows the diagrams that contribute to the W + 1 parton LO cross-section calculation. Note that these are a subset of the diagrams that contribute to the NLO W inclusive calculation. The requirement of a parton in the final state means that W inclusive Born-level diagram and it's associated loop diagram do not contribute. Consequently the infrared divergences of the 1 parton diagrams cannot be canceled, and thus to avoid such divergences it is necessary to restrict the phase space of the calculation with a minimum final state parton P_T cut, $P_T^{\min}$.

Generally a W + n parton LO calculation provides only a partial evaluation of the cross-section at $\mathcal{O}(\alpha_s^n)$, and thus only a partial description of the W + n parton final state. A W + n parton NLO calculation will provide a description of W + n parton and W + n+1 parton final states. Currently, NLO calculations can be made for W + 1 parton and W + 2 parton processes [49], but any more partons in the final state and the calculation becomes too complex. However, nature is not restricted

to considering each W + n parton process in isolation. How, therefore, do we fill in the gaps of the restricted order calculations such that they can be compared to data?

3.4   Parton Showering Models

Parton showering is a way of providing a description of the $n + 1, n + 2...n + x$ parton final states to a LO W + n parton ME calculation. In other words, it approximates higher order corrections to the process which in an ideal world would be evaluated in the ME approach with a NLO or NNLO calculation that accounts fully for all the interference and helicity effects.

In a parton showering model [29] the initial and final state partons that participate in the hard interaction are permitted to radiate or "branch" in a $1 \rightarrow 2$ process allowed by QCD i.e. $q \rightarrow qg, g \rightarrow gg$ or $g \rightarrow q\bar{q}$. The daughter partons can themselves branch, and thus a tree-like multiparton configuration, or parton shower, can develop. Practically there is no difference between initial and final state radiation, both can result in additional partons in the final state and thus shape the event kinematics, and from an experimental stand point it would be impossible to differentiate between the two. However, the evolution of the parton shower does differ in each case.

The evolution of a parton shower depends crucially on two variables; Q, the transverse momentum scale of the branching, and z, which gives the sharing of the parent parton energy and momentum amongst the two daughters. The differential probability $d\mathcal{P}$ for a parton a to branch into daughters b and c, a $\rightarrow$ bc, is given by the Altarelli-Parisi equation [29]:

$$d\mathcal{P} = \sum_{b,c} \frac{\alpha_s(t)}{2\pi} P_{a \rightarrow bc}(z) dt dz \qquad (3.5)$$

Where $t = \ln(Q^2/\Lambda_{QCD}^2)$. The sum runs over all possible branchings, with the dependence on z of each branching given by the Altarelli-Parisi splitting kernels [29]:

$$
\begin{aligned}
P_{q \rightarrow qg}(z) &= C_F \frac{1+z^2}{1-z} \\
P_{g \rightarrow gg}(z) &= N_C \frac{(1 - z(1-z))^2}{z(1-z)} \\
P_{g \rightarrow q\bar{q}}(z) &= T_R(z^2 + (1-z)^2)
\end{aligned}
$$

$$(3.6)$$

Where $C_F = 4/3$, $N_C = 3$ and $T_R = N_f/2$ (N_f being the number of quark flavours

considered).

One can think of the shower as "evolving" in t. The hard scatter occurs at some scale Q_{max}. For the final state shower this represents the maximum allowed Q scale; the virtuality is gradually decreased along the shower, with potentially multiple $a \to bc$ branchings, until we reach some cut-off scale Q_0^2, at which stage hadronization takes over. Conversely, for initial state radiation, we begin at the hard scatter and trace the shower "back in time", that is, given a parton b we try and find the parton a that branched into b. As we "rebuild" the shower the virtuality again decreases, but the momentum fraction (of the proton) x increases. Note that the use of PDFs in the calculation of the hard scatter means that the inclusive effects of initial state radiation are already included, but parton showering enables the inclusion of initial state radiated partons in the final state.

Integrating 3.5 over all allowed z values we can define the probability that a branching occurs during a small range of t values, δt, as:

$$
\begin{aligned}
\mathcal{P}(t, t + \delta t) &= \sum_{b,c} \mathcal{I}_{a \to bc}(t) \delta t \\
\mathcal{I}_{a \to bc}(t) &= \int_{z_-(t)}^{z_+(t)} dz \frac{\alpha_s(t)}{2\pi} P_{a \to bc}(z)
\end{aligned}
\tag{3.7}
$$

Note that due to the conservation of energy and momentum the allowed z values depend on the scale t. The probability that no-branching occurs in the same interval of t is $1 - \mathcal{P}(t, t + \delta t)$. Thus the finite probability for a parton to evolve between the scales t_0 and t without any branching is given by the Sudakov form factor $\Delta(t_0, t)$ [72]:

$$
\mathcal{P}_{no-branching}(t_0, t) = \exp\left[- \int_{t_0}^{t} dt' \sum_{b,c} \mathcal{I}_{a \to bc} \right] = \Delta(t_0, t)
\tag{3.8}
$$

It then follows that the probability for a branching to occur at time t is given by:

$$
\frac{\mathcal{P}(t)}{dt} = \Delta(t_0, t) \mathcal{P}(t, t + \delta t)
\tag{3.9}
$$

These are the basic equations which enable a parton shower to evolve. As the power of α_s in equation 3.5 suggests, these showers approximate to *real* corrections at $\mathcal{O}(\alpha_s)$. In some respects, they go farther than a calculation at that order. The shower allows all partons to evolve further past the initial branching, conserving energy and momentum, and with an α_s factor that runs as a function of the Q^2 scale of the branching. They also approximate the higher-order destructive interference phenomena that prohibits large angle radiations, known as coherence effects, by requiring decreasing emission angles as the shower evolves.

It is in fact conceivable that we could abandon the matrix element approach completely and solely rely on the parton shower model for our description of the hard scatter event. However, because they do not utilise full matrix element expressions, parton showers have limited power in predicting the rate for well-separated jets. Instead they provide a good description of the soft and collinear parton emission which can have a non-negligible effect on the final state kinematics.

3.5 Hadronization Models

The theoretical tools thus far described rely on perturbation theory to predict the final state in terms of parton configurations. However, such a theory is only valid at the short distance scales of the hard scatter, where QCD interactions are relatively weak. As the final state partons evolve away from the hard interaction we re-enter the long distance, low momentum transfer regime where α_s becomes too large for perturbation theory to be practical. The colour singlet parton states are transformed by a soft QCD process known as *hadronization* (or *fragmentation*) into "jets" of colourless hadrons which satisfy the requirements of QCD colour confinement. Typically the hadrons lie close in angle to the direction of the parent parton, hence the term "jets". It is these jets which we observe in our detector, and thus to make comparisons between data and the parton level predictions outlined above we must have some model of hadronization, even if only to correct our data for it. Since we cannot predict this process from first principles such models are phenomenological. Three main schools of model exist; cluster fragmentation [75], string fragmentation [31, 35] and independent fragmentation [74]. Since the independent model is not used in this measurement we describe only the string and cluster models.

The cluster model of hadronization [75] is a relatively simple approach motivated by the colour preconfinement property [30] of parton branching. At the end of the parton shower we have some configuration of quarks, antiquarks and gluons. Due to the conservation of colour in the branching process neighboring quark-antiquarks will tend to have opposite colour, and thus can be combined into colour-neutral clusters. The gluons are split (non-perturbatively) into q$\bar{\text{q}}$ pairs before the colourless cluster is formed. These clusters will in general decay isotropically in their rest frame into two hadrons, although this depends to some extent on their mass. Very heavy clusters may first decay to lighter clusters, and very light clusters may be permitted to decay to just one hadron. Throughout this process the kinematics of the perturbative shower are preserved, such that the direction of resulting "daughter" hadrons is likely to be highly correlated with that of their "parent" partons.

The Lund String Model [31, 35] is a more complicated description of fragmen-

tation that utilises the concept of a linear quark confining potential. It is best illustrated with the example of $e^+e^- \rightarrow q\bar{q}$ production. As the quark and antiquark separate the QCD interactions between them increase in strength. In the string model the colour force lines collapse into a string like configuration between the quarks which, provided the distribution of the force lines within the string is uniform, will have a constant strong potential energy per unit length. Thus, as one might expect, the energy stored in the string will increase with increasing quark separation. When this stored energy exceeds the energy needed to create a $q\bar{q}$ pair from the vacuum, the string will fragment into two $q\bar{q}$ colour-singlet systems i.e. two new strings. These strings can go on to break-up further until we are left with $q\bar{q}$ pairs that are on-mass-shell hadrons. Although it may sound relatively simple from this description, the details of exactly how and when the strings fragment are complex. The model requires many phenomenological parameters to describe the energy-momenta spectra and flavour composition of the produced hadrons.

3.6   The Underlying Event

In the perturbative calculation of the hard-scatter process we consider the interaction between one parton in the proton and one parton in the anti-proton, so called *initiator* partons. However, in reality, the hard-scatter process is just a small component of a larger proton-antiproton interacting system. As well as the hard-scatter process, which will in general result in final state particles with large transverse momentum, there will be a whole spectrum of softer interactions which can also result in final state particles with non-negligible transverse momentum. In addition, these interactions are not isolated from one another but are colour connected, and correlations between them can be important. In this thesis we define the underlying event as those interactions which occur between the proton and antiproton which are not the hard-scatter process we are interested in. Note that this definition is not free from ambiguity since the additional interactions are not isolated from the hard-scatter process.

Modelling the underlying event is very challenging. Naively, one might think that a reasonable model would be to simply allow for multiple parton-parton $2 \rightarrow 2$ QCD interactions to occur within the same event. As well as the high Q^2 hard-scatter, there may be additional lower Q^2 scatterings. However, this has several problems:

- How do we determine the multiplicity of additional scatterings?

- Only high Q^2 interactions can be calculated perturbatively. Thus we can only calculate additional scatterings down to some minimum P_T scale.

- Correlations between the scatterings are not understood.

3.7 The ELO Monte Carlo Event Generator

Monte Carlo event generators are a way of using all the theoretical concepts discussed in Section 3.3 to produce "simulated" events of the process that can be readily compared to experimental data. In the "Enhanced Leading Order" [20] approach, a LO matrix element event generator, which handles the hard scatter via a tree-level calculation, is interfaced with a separate parton showering programme to simulate softer radiation and provide hadronization. Discussions of the applicability and limitations of the ELO approach are reserved for Section 3.9. Here we simply describe the mechanics of an ELO Monte Carlo event generator. In this analysis we make use of the ALPGEN matrix element event generator [104], interfaced with either the PYTHIA [119] or HERWIG [58] programmme to provide the parton showering. PYTHIA and HERWIG are Monte Carlo event generators capable of modelling a wide range of physics processes, but in this study we only make use of their parton showering, hadronization and underlying event models. The ELO event generation process can be divided into the following steps that are explained in detail below:

1. Set-up the parameters of the hard process.

2. Using LO matrix element calculations construct the differential cross-section distribution in the allowed phase space.

3. Use this distribution to generate an events sample.

4. Run parton showering and hadronization programmes on these events.

5. Potentially pass hadron-level events through a detector simulation.

The restrictive nature of the leading order calculation requires the setting of a number of parameters related to the hard scatter process. These include:

- The number of final state partons in the tree-level calculation i.e. LO W + n partons

- Choice of factorisation and renormalisation scales. These are generally but not necessarily set equal. The factorisation scale is the scale at which the PDF is evaluated, whilst the renormalisation scale determines the α_s coupling constant used in the matrix element evaluation.

- Selection of a PDF set

- Definition of the phase space for the calculation. At LO infrared and collinear divergences have to be avoided by a matrix element parton $P_T^{\min}$ and $\Delta R^{\min}$ [3] cut.

---

[3]The separation in $\eta \times \phi$ space of two objects a and b is given by $\Delta R = \sqrt{(\phi_a - \phi_b)^2 + (\eta_a - \eta_b)^2}$

In order to be able to generate events that kinematically reflect the results of the perturbative calculation a Monte Carlo generator must know something about the distribution of rates in parton phase space i.e. we must form differential cross-section distributions. However the number of diagrams and complexity involved in even a LO matrix element calculation means that the integration over phase space must be done numerically. We must gradually build-up a picture of the differential cross-section by sampling random points in the allowed phase space $d\phi$ and calculating the cross-section at this point $d\sigma$.

Parton-level differential cross-section distributions are clearly useful in their own right, but parton production rates are not directly observable in our detector. In order to make progress we need to generate "events" with particular final state parton configurations and kinematics, the frequency of which reflects the underlying differential cross-section. This is done by assigning each point in phase space a "weight" equal to $d\sigma$ at that point. At this stage we can think of each point in phase space as representing a possible event, and thus we have a collection of "weighted" events. The events are "unweighted" using the *acceptance-rejection* technique as follows:

1. Find the maximum event weight $d\sigma_{max}$.

2. Choose an event at random and generate a random number f uniformly in the interval (0,1).

3. If the ratio of event weight over maximum weight exceeds the random number ($d\sigma/d\sigma_{max} > f$) accept the event, otherwise reject it.

In the limit of an infinite number of "unweighted" events the resulting distribution in phase space would be exactly that of the differential cross-section. With limited statistics the "unweighted" events reflect what might be observed in a trial experiment.

In the ELO approach these events are then passed to a parton showering programme which performs initial and final state parton showering as described in Section 3.4. In doing this the kinematics, flavour and colour information of each event is passed to the PS such that it can be used in the shower evolution. The parton shower models used by PYTHIA and HERWIG are very similar, differing only in the exact choice of evolution parameter and in their treatment of QCD coherence effects. Once the parton shower has been terminated at the fragmentation scale the resulting partons are hadronized as described in Section 3.5. PYTHIA uses the string fragmentation model, whereas HERWIG uses the simpler cluster model. Both programmes also make some attempt to model the underlying event. HERWIG uses a model that was parameterized using $\sqrt{s} = 630$ GeV $p\bar{p}$ collider data collected at

UA5, whereas PYTHIA uses a model tuned on CDF Run I data at $\sqrt{s} = 1.8$ TeV, the so-called PYTHIA "Tune A" [76].

3.8 Jet Clustering Algorithms and the Jet Definition

A theoretical prediction is useless unless it can be tested by experimental observation. At first sight, this statement seems to sink the perturbative calculation. The predictions made here, in terms of parton configurations, are fundamentally unobservable and only via a little understood process do they become observable. How then can we take what we observe, the jets of hadrons, and relate it back to that which we can actually predict, the hard scattering partons, given that we do not understand in any detail how partons become jets? This is the role of jet clustering algorithms.

Lets begin with an ultra naive (and incorrect) statement: it should be possible to associate all final state hadrons to a parent final state parton. If we could make this association perfectly, with a perfect jet clustering algorithm, then conservation of energy and momentum dictates that it would simply be a case of combining the hadron 4-vectors to reconstruct the parent parton. In this picture the details of the hadronization process do not matter, we can bypass this and go directly to the parton level to compare to our perturbative prediction.

There are several problems with this oversimplified picture. Firstly, what do we mean by a parent parton? A parton resulting from the hard scatter? But what about subsequent parton showering? In reality there is no hard boundary between parton showering and the onset of hadronization. This means that even at parton level, after any parton showering but before hadronization, we need a jet clustering algorithm, a concept of a jet. One of the properties of an ideal jet algorithm would be that it reconstructs identical jets at both the parton level and the hadron level (after hadronization). Then we would again have an algorithm that is insensitive to hadronization.

However, even the most modern calorimeters cannot hope to measure the energy of individual hadrons. Each tower in the CDF central calorimeter is 15 degrees in azimuth by about 0.11 in pseudorapidity (see Chapter 4), and thus many hadrons may deposit their energies in a single tower. Thus when it comes to forming jets out of real data, we are reduced to treating these towers as the particles that make up the jets. If we are to make detector independent measurements, then clearly the calorimeter segmentation should not impact the jets that are reconstructed: we should reconstruct identical jets at the parton, hadron *and* detector level. In reality, not only is the calorimeter segmented, but the energy measurement of the towers is

imperfect. However, this energy can be corrected to on average reflect the energy of the hadrons, as is discussed later in Section 6.2.2.

In general, as well as being insensitive to non-perturbative physics like fragmentation, we want our jet definition to be robust to higher order corrections which are not fully accounted for in our theory. In our discussion of perturbative predictions we have stressed that LO calculations contain infrared and collinear divergences that are only cancelled at higher orders, and thus it is necessary to impose a parton $P_T^{\min}$ cut to avoid them. The hole left in phase space by such a cut can be filled in by parton showering, but this is an approximation as opposed to a complete description. Thus, as well as the jet clustering algorithm, a complete jet definition should include some minimum cut on the jet P_T such that we remain in the kinematic regime where perturbative predictions are valid. Even with such a cut imposed, a jet algorithm can still be sensitive to infrared and/or collinear radiation manifestly observable in the final state jets that are reconstructed. Such a situation is to be avoided, as we are then in effect sensitive to the expected limitations of our theory.

The ideal properties of a jet algorithm can be summarised as follows:

- Insensitive to non-perturbative effects: reconstructs identical jets at parton and hadron level.

- Insensitive to higher order effects: Infrared and collinear *safe.*

- Detector independent.

- A more practical consideration: clustering can be easily implemented and performed quickly.

We shall now go on to describe each of the three main jet algorithms in turn: the cone, midpoint and K_T algorithms.

3.8.1   Cone Algorithms

Cone algorithms form jets by associating together particles (be they partons, hadrons or towers) whose trajectories lie within a circle of a specific radius R in $\eta \times \phi$ space. Defining some initial cone centre or *axis*, the scalar E_T -weighted centroid of the cone is calculated by considering all the particles that lie within the cones radius:

$$\eta^C = \frac{\sum_{i \subset C} E_T^i \eta^i}{E_T^C}, \; \phi^C = \frac{\sum_{i \subset C} E_T^i \phi^i}{E_T^C} \tag{3.10}$$

where the sum is over all particles within the cone C and the scalar sum E_T of the cone is given by:

$$E_T^C = \sum_{i \subset C} E_T^i \tag{3.11}$$

This centroid then becomes the axis for a new cone. We assess which particles lie within this new cone and recompute the cones centroid, using this to define a new cone and so on. This process terminates when the new centroid is no different from the cones existing axis. We have then identified a "stable" cone, or jet. Several choices of recombination scheme exist, but that which is used in the CDF JetClu cone algorithm is now described [9]. The particles are considered massless, such that the jet 4-vector $(E^J, P_x^J, P_y^J, P_z^J)$ is given by:

$$
\begin{aligned}
P_{x,y,z}^J &= \sum_{i \subset C} P_{x,y,z}^i \\
E^J &= \sum_{i \subset C} E^i
\end{aligned}
\tag{3.12}
$$

where particle quantities $P_{x,y,z}^i$ are defined

$$
\begin{aligned}
P_x^i &= E_T^i \cdot cos(\phi^i) \\
P_y^i &= E_T^i \cdot sin(\phi^i) \\
P_z^i &= E^i \cdot cos(\theta^i) \\
E_T^i &= E^i \cdot sin(\theta^i)
\end{aligned}
\tag{3.13}
$$

In the case of tower clustering the energy E^i of the particle would be the total energy measured in the tower, and the particle direction (ϕ^i, θ^i) that defined by a unit vector pointing from the detector centre to the tower centre.

However, when it comes to practically implementing such an algorithm, we find that it is an incomplete description of a jet clustering procedure. Firstly, in the above description we have neglected to describe how the initial cone centre is chosen. In principle one could examine every tower as a starting point for the algorithm. In practice this requires far too lengthy computation times, and the concept of the "seed" had to be introduced. At CDF, only towers with energies in excess of 1 GeV are considered as initial cone centres. Secondly, there is nothing in the above prescription that prevents two stable cones overlapping. In this case, what happens to the particles which lie in the overlap region? Particles clearly should not be assigned to more than one jet, as this would be to double-count their energy. This problem is overcome by introducing a merging fraction f, which in the CDF algorithm case is set at 75%. If the shared energy is greater than 75% of the energy of the lower energy cone the cones are merged into one large jet. If it is smaller, the particles in the shared region are distributed between the jets according to which axis is closest in $\eta \times \phi$ space.

Unfortunately these issues lead to the cone jet algorithm being sensitive to higher
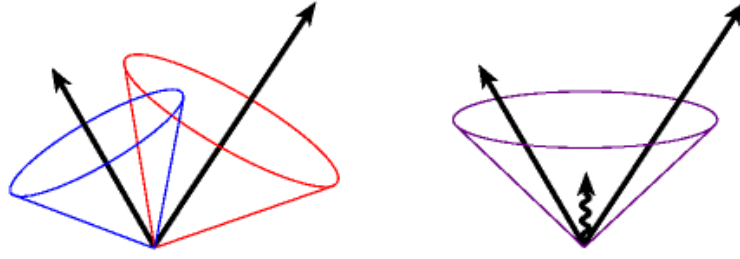
Figure 3.9: Diagram illustrating the sensitivity of cone algorithms to infrared radiation. Taken from [46].

order corrections. Figure 3.9 shows a situation were two high energy particles "seed" two different jets. If additionally there is some soft radiation present in between the two jets a merging into one jet can result, which would otherwise not have happened. This illustrates the algorithms potential infrared sensitivity. Figure 3.10 shows a case where the use of a seeds leads to problems with collinear sensitivity. In the diagram on the right a single particle seeds a single jet. However, if there is some collinear radiation, as in the left hand diagram, the energy of that particle is instead divided among several particles/towers. It could be that none of the particles now pass the seed threshold, and thus no jet is reconstructed at all. The situations in Figures 3.9 and 3.10 are obviously rather extreme, but they serve to illustrate why theorists in general do not approve of cone jet algorithms.
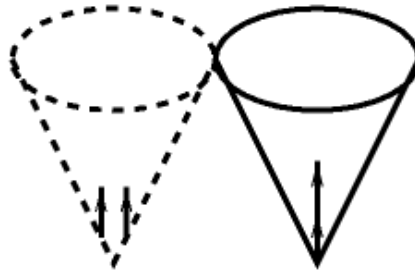
Figure 3.10: Diagram illustrating the sensitivity of cone algorithms to collinear radiation. Taken from [46].

3.8.2 The Midpoint Cone Algorithm

The midpoint cone algorithm is very similar to the seeded cone algorithm, but contains the following important modifications which are intended to minimise the problems of seeded cone algorithms described above:

- After all the stable cones have been identified additional seeds are placed at the "midpoint" positions. These positions are defined by the centroids of all

possible cone multiplets e.g. all possible pairs $P_i^J + P_j^J$, all possible triplets $P_i^J + P_j^J + P_k^J$ etc. The seeds at the midpoints are used to define new initial search cones, thus reducing the sensitivity of the algorithm to soft radiation.

- The CDF MidPoint algorithm [42] uses 4-vectors throughout the clustering and recombination process, as opposed to the scalar E_T and pseudorapidity approach of JetClu. Whether or not a tower is clustered into a cone is determined by the separation in "true" rapidity-phi space, $y \times \phi$. The final jet 4-vector is the true combination of the 4-vectors of all the particles/towers that were clustered:

$$(E^J, P_x^J, P_y^J, P_z^J) = \sum_{i \subset C}(E^i, P_x^i, P_y^i, P_z^i) \tag{3.14}$$

In the limit of massless particles/towers, the recombination approaches of Jet-Clu and MidPoint should be equivalent.

- The JetClu algorithm contains the feature known as "ratcheting": a jet cone must always contain the original seed tower that initiated it, even if the seed tower does not lie within the final stable cone. This can result in non-circular jet topologies, and is clearly a feature very dependent on seeds and thus not modelled by perturbation theory. The CDF MidPoint algorithm does not contain "ratcheting".

3.8.3 The K_T Algorithm

K_T algorithms [51] are a rather different class of jet clustering algorithm, relying not on the containment of particles within a cone, but on the relative transverse momentum between particles to determine whether or not they should be clustered together. A flowchart of the algorithm is shown in Figure 3.11. Initially we label each tower/particle as a "precluster", which we assume to be massless. For each precluster we define:

$$d_i = E_i^2 sin^2\theta_i = P_{T,i}^2 \tag{3.15}$$

and for each pair (i, j) of preclusters define:

$$d_{i,j} = \min(P_{T,i}^2, P_{T,j}^2)\frac{\Delta R_{i,j}^2}{D^2} \tag{3.16}$$

where $\min(P_{T,i}^2, P_{T,j}^2)$ means take the smallest P_T^2 from (i, j) and $\Delta R_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$. In the small angle limit, $\Delta R_{i,j} << 1$, and with the parameter $D \approx 1$, $d_{i,j}$ reduces to the minimal relative transverse momentum squared, K_T^2, between preclusters i and j. We now find the smallest value of all the d_i and $d_{i,j}$ and label
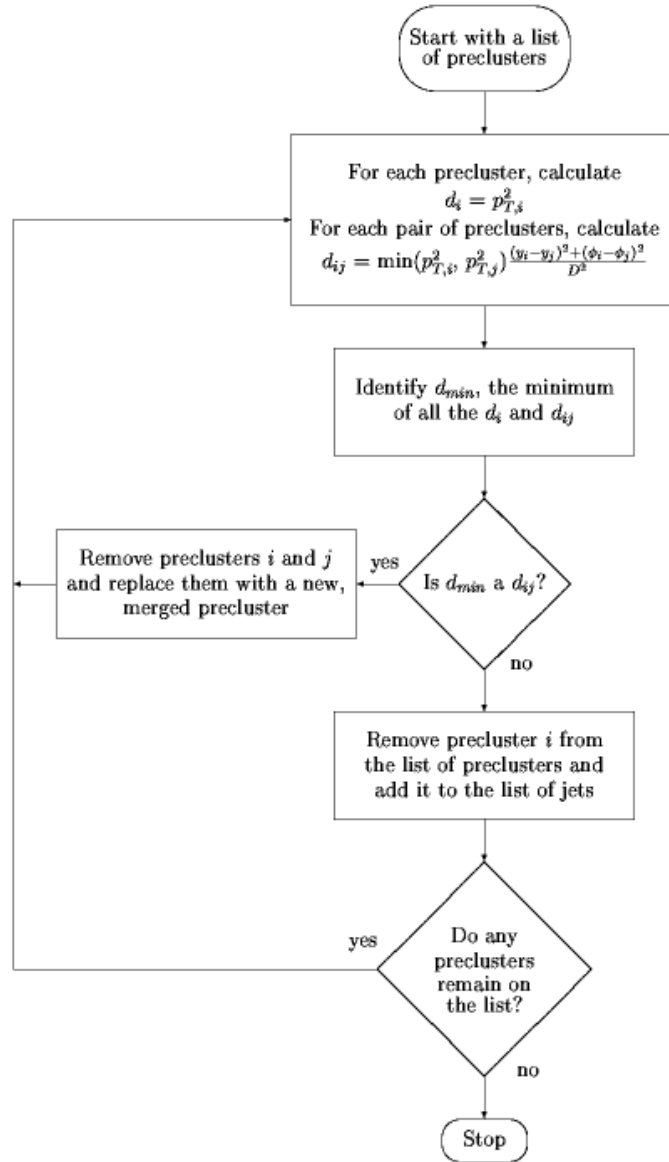
Figure 3.11: Flow diagram illustrating the workings of the inclusive K_T algorithm. Taken from [46].

this d_{min}. If d_{min} is a $d_{i,j}$ then the preclusters i and j are merged into one. If d_{min} is a d_i then the precluster i is removed from the list of preclusters and added to the list of jets. The quantities $d_{i,j}$ and d_i are then recalculated with the new list of preclusters, a new d_{min} found and so on. The algorithm terminates only when there are no preclusters remaining.

The algorithm described above is that which is in use at CDF, the "inclusive" K_T algorithm. It is so-called because every particle/tower in the event is associated to a unique jet. The parameter D plays a similar role to the cone size R in cone algorithms in that it controls to some extent the "size" of the resulting jets.

There also exists a slightly different implementation known as the "exclusive" K_T algorithm. This differs as follows:

- If d_{min} is a d_i then the precluster is considered a "beam-jet". It is removed from the list of preclusters but does not become a final state jet. In contrast to the inclusive algorithm, not all particles become associated with a final state jet in the exclusive algorithm.

- The merging process terminates when d_{min} is greater than a predefined K_T resolution cutoff parameter d_{cut}.

In the exclusive algorithm d_{cut} plays a similar role to D in controlling the size of the jets. If d_{cut} is small then the process will terminate early with a large number of final state jets.

More extensive use of the K_T algorithm has long been advocated by theorists. Its use of the K_T variable in the jet clustering process means that is by construction both infrared and collinear safe. There is also no overlapping of jets and hence no need for unphysical merging/splitting procedures. However, one of the problems in the practical implementation of the K_T algorithm at hadron colliders is that the jets have no fixed size. This makes data-driven prescriptions for the on average correcting of jet energies for the underlying event and multiple interactions difficult.

3.9 ELO and the Double-Counting Problem

The so called "Enhanced Leading Order" approach (ELO), where LO matrix element generation is combined with a parton showering model, is by now a well established method for making W+Jets predictions. It has been demonstrated in CDF Run I [20] and more recently in Run II [107] that a W + n parton ELO Monte Carlo event generator is able to predict with considerable success the W+ $\geq$ n jet cross-section, as well as the shapes of jet kinematic variables in $\geq$ n jet bins, such as jet E_T , jet-jet ΔR and invariant mass distributions. This should not be surprising. The W+ $\geq$ n jet cross-section is dominated by events with exactly n jets. Provided we have chosen a "hard" jet definition (i.e. selected jets from those clustered with a minimum E_T requirement), the rate of production and kinematics of these n jet events should be well predicted by an n parton LO matrix element calculation, given a sensible choice of P_T^{min} . The > n jet events are a sub-leading contribution to the $\geq$ n jet cross-section, and the parton showering model of these events is reasonable.

However, this is not to say that the lack of true higher order corrections in an ELO prediction goes completely unnoticed. Figure 3.12 is taken from the CDF Run I W + Jets study [20], and shows the fraction of events with $\geq$ 1 jets that have exactly 1 jet, $N_{1j}/N_{\geq 1j}$, as a function of lead jet E_T . A W + 1 parton ELO prediction is compared with the Run I W data. As the lead jet E_T increases the number of jets in the event increases also, thus reducing the fraction. The deficit of higher multiplicity
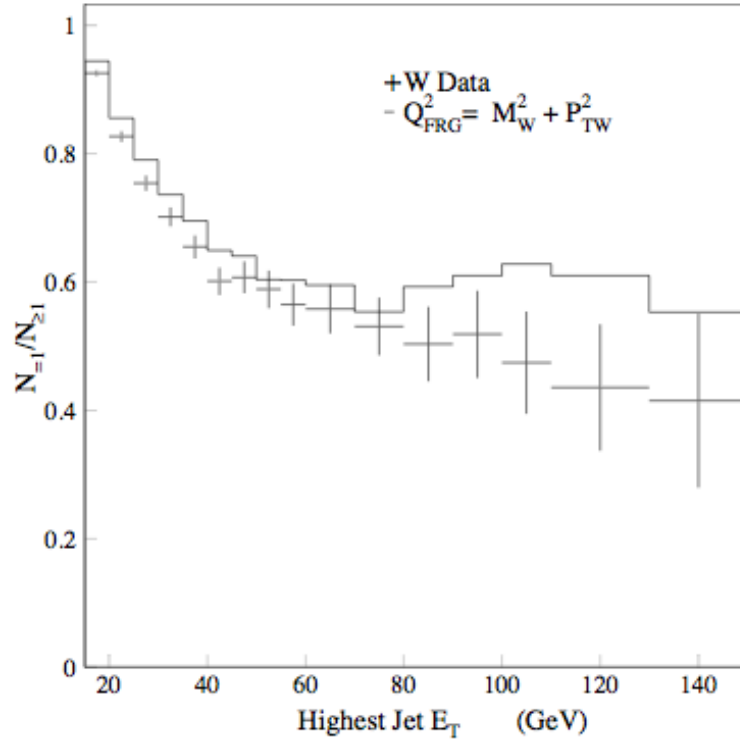
Figure 3.12: Fraction of W events with ≥ 1 jets that have exactly 1 jet, $N_{1j}/N_{\geq 1j}$, as a function of lead jet E_T . The solid line is a VECBOS + HERWIG ELO W + 1 parton sample and the points Run I W + Jets data. Taken from the Run I W + Jets study [20].

events in the ELO theory curve at high jet E_T is indicative of the limitations of ELO. In this region parton showering fails to predict accurately the production of additional hard jets, which should be evaluated using a ME calculation. This same limitation is also evident in jet E_T spectrum comparisons.

One proposed way of addressing such limitations is to combine ELO W + 0,1,2,3 and 4 parton samples into one, each sample contributing a number of events equivalent to its relative cross-section. The production of additional hard jets in the n parton sample is then in effect handled by the leading order matrix element calculation of the n + 1 parton sample. Parton showering is still required to provide higher order approximations, with the PS of the n - 1 parton sample filling in the region of phase space "below" the P_T^{min} generation cut of the n parton sample. In this way we populate all phase space, and thus dependence on the P_T^{min} generation cut should be greatly reduced. However, there is a problem with naively combining the samples in this way. In the standard ELO approach there is no explicit restriction on the "hardness" of the parton shower; it is kinematically possible that within the shower development a gluon can be radiated with P_T (relative to the mother parton) exceeding P_T^{min} . Not only is the parton shower filling in the phase space below P_T^{min} ,

but it is infringing on the phase space already covered by the ME calculation of the n + 1 parton sample. This phase space "overlap" results in an effective double counting of events in the combined sample.

The double counting problem arises essentially because there is a lack of communication between the ME and PS elements of the ELO approach. "ME-PS matching schemes" are an attempt to introduce communication and thus avoid phase space overlap, allowing the samples to be combined into a complete W + Jets sample. Two schemes exist currently; that proposed by Michelangelo Mangano, which we shall refer to as the "MLM Matching Scheme" [103], and that proposed by Catani, Kuhn, Kraus and Webber, commonly referred to as the "CKKW Matching Scheme" [52] [99].

3.9.1   CKKW Matching Scheme

The CKKW matching scheme makes extensive use of the K_T clustering algorithm (see Section 3.8.3). First weighted events are generated using an n parton LO matrix element calculation, with some minimum K_T separation of the final state partons, $d_{ij} > d_0$, to avoid infrared/collinear divergences. There then follows a PS-like reweighting of the event relative to the original ME weighting. By repeatedly clustering using the K_T algorithm it is possible to determine the minimum K_T resolution cutoff variable d_N above which we reconstruct N partons, such that $d_1 > d_2 > d_3... > d_n > d_0$. One can think of these as representing a resolution or energy scale at each vertex. This information is then used to reweight the event in two ways:

1. Reweight by $\alpha_s(d_1)\alpha_s(d_2)\alpha_s(d_3)...\alpha_s(d_n)/\alpha_s(d_0)^n$.

2. Compute the Sudakov form factor (Equation 3.8) on each line $\Delta(d_i, d_j)$ and reweight by these factors also.

These recalculated weights are then utilised in the usual unweighting procedure, and each event that passes the unweighting undergoes parton showering. The starting scale for the evolution of the shower along each line is given by the resolution variable d_i of that line. If however any radiation is produced with K_T that exceeds d_0 then this event is vetoed and does not enter the final n parton sample. In this way we prevent the parton shower from infringing on the phase space $d_{ij} > d_0$ which is already covered by the matrix element. The reweighting of the matrix element additionally ensures that the dependence on the choice of the d_0 scale largely cancels out.

3.9.2   MLM Matching Scheme

Unlike CKKW, the MLM scheme can be applied to the ELO samples after the normal event generation procedure described in Section 3.7. To use it requires a description of the jets in the event at the hadron level i.e. after parton showering and hadronization, as well as the 4-vectors of the ME hard partons i.e. the final state partons from the unweighted ME event before parton showering. The hadron level jets are defined by some jet clustering algorithm with a minimum jet E_T cut. The matching scheme works by establishing whether or not there is a correspondence between the ME partons and the hadron level jets. Take an event from an n parton ELO sample within which we have reconstructed N hadron jets. In the "exclusive" MLM scheme the event is only kept if $N = n$, otherwise it is vetoed and does not enter the final generated event sample. The idea here is that the n parton matrix element calculation covers the N jet phase space only, and does not infringe on the $N + 1$ or $N - 1$ phase space. In the "inclusive" MLM scheme the event is only kept if $N \geq n$. If no $n + 1$ parton sample exists the "inclusive" scheme should be used with the n parton sample, such that we allow $N > n$ jet final states.

Hence in the MLM scheme the jet clustering is used to regulate the phase space covered by each n parton calculation and prevent overlap. However, within reason, the final "combined" sample should be independent of the particular parameters of the jet clustering algorithm used in the matching e.g. the jet E_T cut, cone size.

Chapter 4

The Experiment

The data that is used in this analysis was collected using the CDF detector, one of two multipurpose detectors that are used to examine the high energy collisions of proton-antiproton beams produced by the Tevatron collider at Fermilab, Illinois.

Fermilab has been the site of several important discoveries that have helped to confirm the standard model of particle physics. In 1977 the E288 experiment observed the Upsilon $b\bar{b}$ meson [89], providing the first evidence of the bottom quark's existence, and in 1995 the CDF and D0 Tevatron experiments completed the quark sector of the standard model with the first observation of the top quark [8].

The top quark discovery came at the end of the "Run I" period of data taking at the Tevatron, which ran from 1985 to 1995. Between 1995 and 2000 major upgrades were made to both the Tevatron and the CDF and D0 experiments, increasing performance and physics reach, and in 2001 the "Run II" data taking period began.

4.1   The Tevatron

With a centre-of-mass energy of $\sqrt{s} =$1.96 TeV, the Tevatron is currently the world's highest energy particle collider, and the only operational hadron-hadron collider. In this section we describe how the proton and antiproton beams are produced and collided, and explain some of the measures of Tevatron performance.

4.1.1   Producing Proton-Antiproton Collisions

The process of producing proton and antiproton beams of the required energies and densities for collision requires five accelerator stages [108]: the Cockroft-Walton, the Linac, the Booster, the Main Injector and the final Tevatron accelerator stage. These make up the Fermilab accelerator chain, a schematic diagram of which is shown in Figure 4.1.

The production of protons begins by accelerating H^- ions to an energy of 750
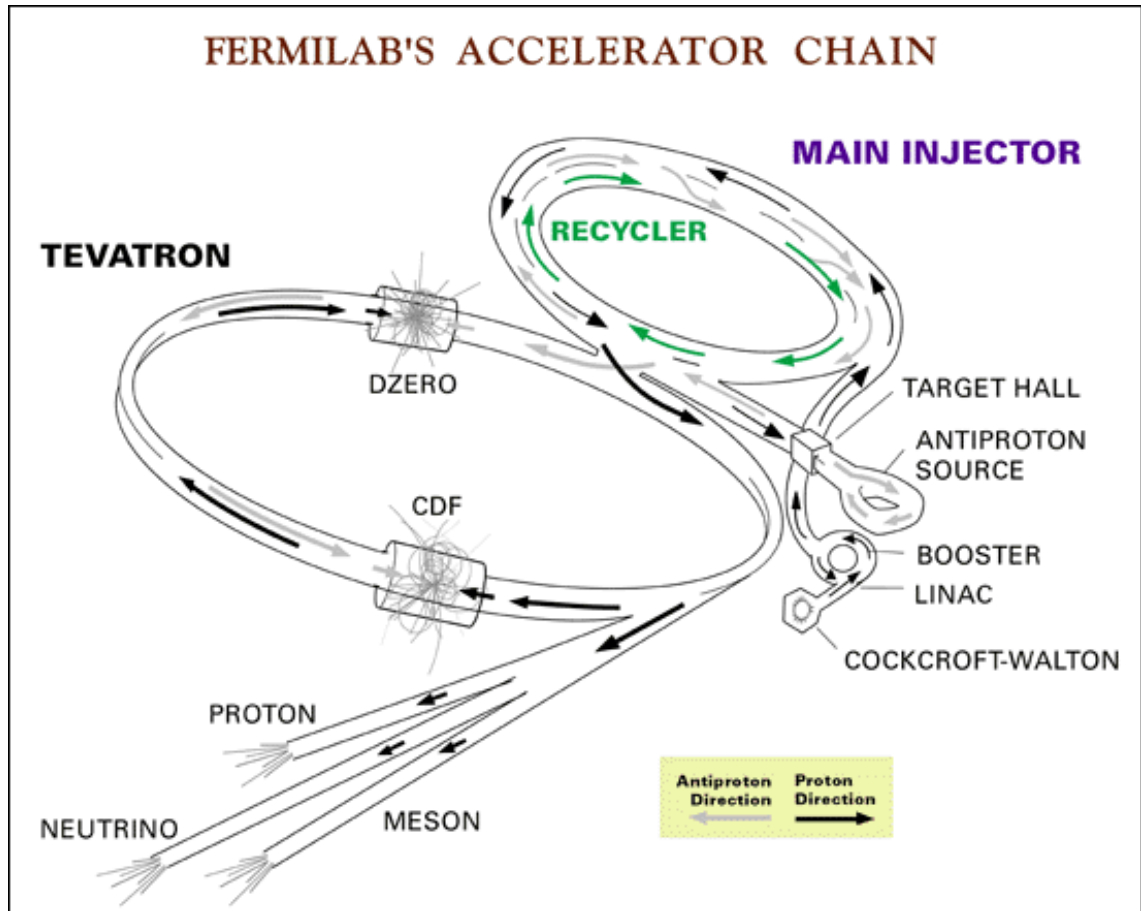
Figure 4.1: Schematic diagram of the Fermilab accelerator chain.

KeV in the Cockroft-Walton accelerator. This beam of ions is then injected into the Linac, a 150m long linear accelerator which uses radio frequency (RF) cavities to incrementally increase the beam energy to 400 MeV. At the end of the Linac the ion beam is passed through a graphite foil. This strips the H^- ions of their two electrons, leaving a pure proton beam which is then passed into the Booster, a 474m circumference synchrotron. Here RF cavities gradually accelerate the protons to an energy of 8 GeV, at which point they are collected into bunches and transferred to the Main Injector.

The Main Injector is a 3km circumference synchrotron that accelerates the proton bunches to an energy of 150 GeV before insertion into the Tevatron itself. However, at 120 GeV some of the proton bunches are extracted from the Main Injector and directed onto a nickel target. The resulting collisions produce many particles, most importantly antiprotons, which can be isolated and selected using lithium lenses and a magnetic field, as shown in Figure 4.2. These antiprotons are then passed into a series of two storage rings prior to Tevatron insertion. First is the Debuncher, a triangular storage ring of mean radius 90m, which reduces the longitudinal momentum spread of the beam. The antiprotons then pass into the Accumulator, housed in the
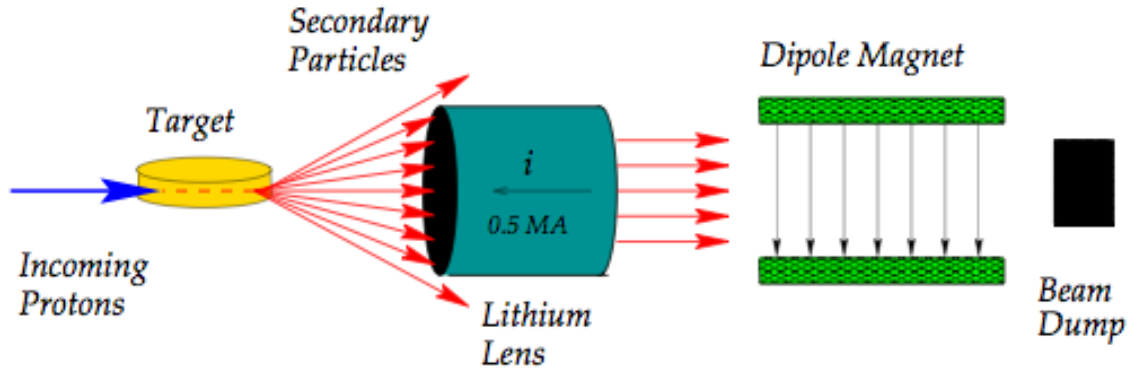
Figure 4.2: Schematic diagram illustrating the production and selection of antiprotons.

same enclosure as the Debuncher, which stores antiprotons in bunches until there are enough collected for injection into the Tevatron. Both the Debuncher and the Accumulator also use a feedback process called *stochastic cooling* [123] to gradually reduce the *emittance* (area in phase space) of the beam. Once 10^{12} antiprotons have been stored in the Accumulator they are inserted into the Main Injector at 8 GeV, where they are accelerated to 150 GeV and injected, along with the protons, into the Tevatron.

The Tevatron is a 1km radius proton-antiproton colliding beam synchrotron that uses superconducting magnets to hold the protons and antiprotons in intertwined helical orbits, the antiprotons traveling anticlockwise around the ring, and the protons clockwise. Once there are 36 bunches each of protons and antiprotons (a "store") circulating in the Tevatron the energies of each beam are ramped up simultaneously to 980 GeV and then brought into collision at the interaction points B0 and D0. As long as there are no problems a single store can last for several days, with the beams continually orbiting in the machine and colliding every 396 ns, a bunch crossing rate of 2.5MHz.

4.1.2 Tevatron Performance

A key measure of the performance of a collider is the instantaneous luminosity, $\mathcal{L}_{\text{inst}}$. Related to the flux of the colliding beams, it essentially determines the number of proton-antiproton interactions that will occur per unit time, and is given by the formula below [110]:

$$\mathcal{L}_{inst} = \frac{fN_P N_{\overline{P}} N_B}{4\pi\sigma_{xy}} \tag{4.1}$$
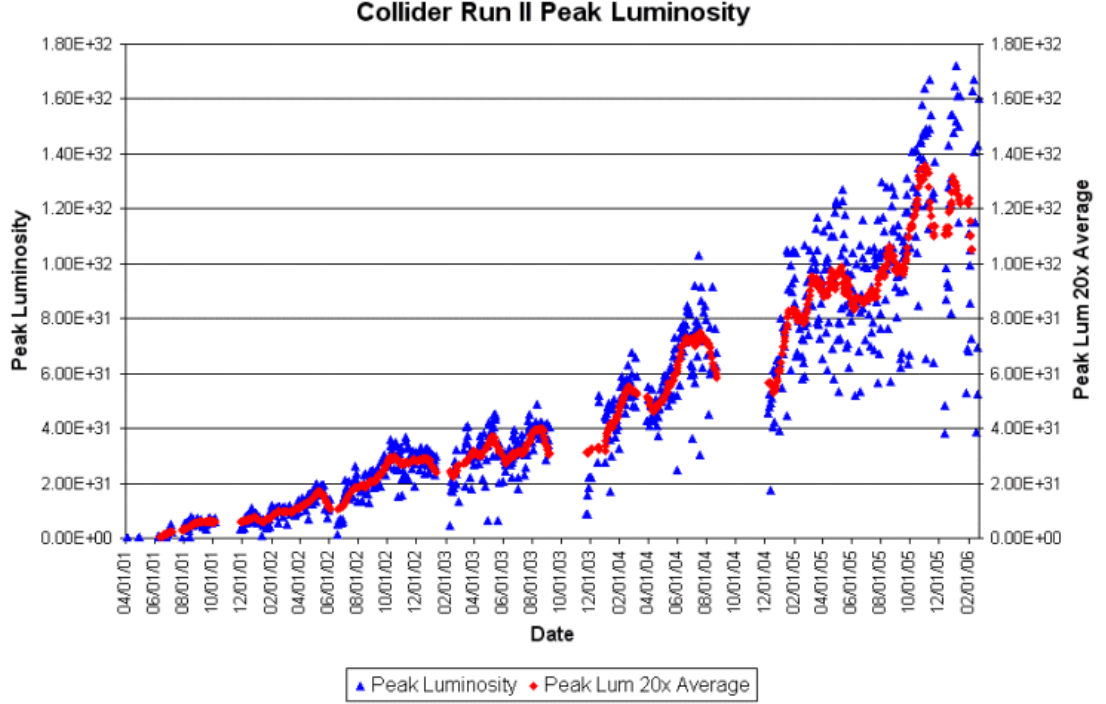
Figure 4.3: The Tevatron peak instantaneous luminosity at the beginning of each store as a function of time (blue triangles). Also shown is the peak luminosity averaged over the last twenty stores (red diamonds).

Where f is the orbital frequency of the synchrotron, N_P is the number of protons and $N_{\overline{P}}$ the number of antiprotons per bunch, N_B is the number of bunches and σ_{xy} is the RMS width of the Gaussian transverse beam profile. The conventional units are $cm^{-2}s^{-1}$. Figure 4.3 shows the instantaneous luminosity of the Tevatron as a function of time since Run II began. Recently the peak instantaneous luminosity has approached the planned value for Run II of $2 \times 10^{32} cm^{-2}s^{-1}$.

For a process of a specific cross-section, σ, the number of events one expects to observe in a given time interval dt is given by:

$$N = \sigma \cdot \int_t^{t+\delta t} \mathcal{L}_{inst} dt = \sigma \mathcal{L} \tag{4.2}$$

Where $\mathcal{L}$ is the integrated luminosity, the instantaneous luminosity integrated over time. The integrated luminosity is often used as a measure of the amount of data collected by an experiment, expressed in units of $barns^{-1}$.

To date the Tevatron has delivered over $1.5 fb^{-1}$ of data, as shown in Figure 4.4. In this analysis we use data collected between August 2002 and August 2004, corresponding to an integrated luminosity of $320 pb^{-1}$.
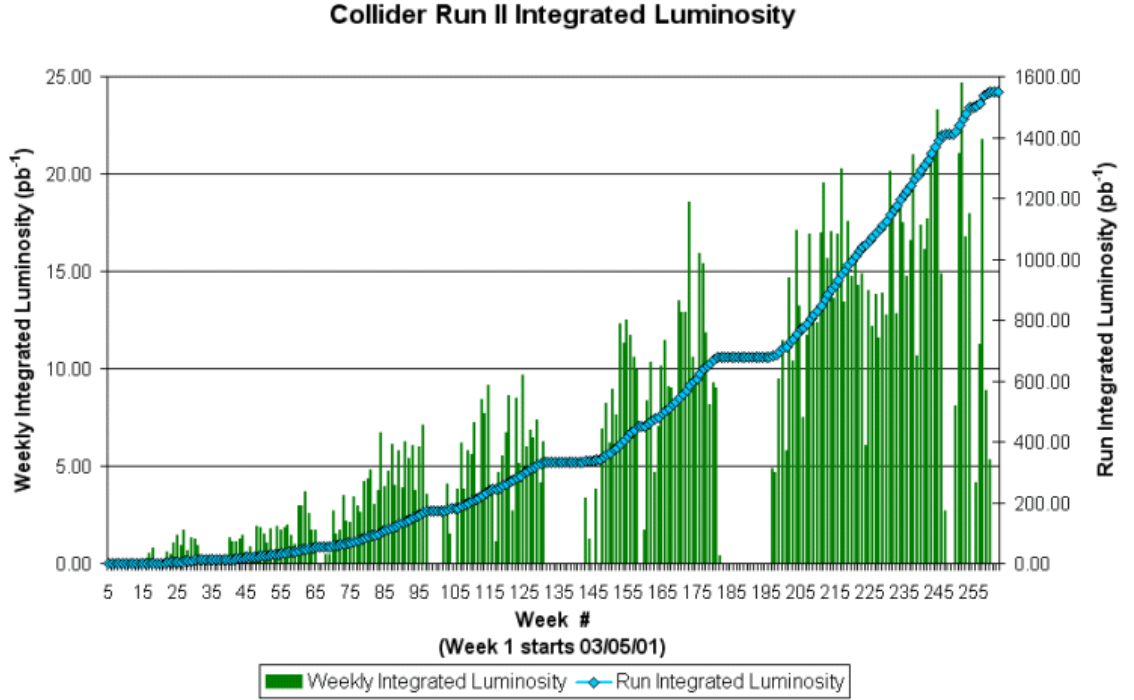
Figure 4.4: Graph showing the accumulation of integrated luminosity delivered by the Tevatron over time. The blue curve is the total integrated luminosity and the green bars show the integrated luminosity for that week number. Note that there are several periods where no data was taken. These correspond to "shutdown" periods necessary to service and maintain the detector.

4.2   The Collider Detector at Fermilab

The CDF II detector [100] is located at one of the six nominal p$\overline{\text{p}}$ interaction regions of the Tevatron. It is a multipurpose detector, designed to study a wide range of processes occurring in p$\overline{\text{p}}$ collisions. It's various components together allow for the identification and characterization of high energy electrons, muons, photons and jets, as well as more specialised measurements such as the identification of B mesons. Figure 4.5 shows an isometric view of the CDF detector, from which one can observe the cylindrical symmetry of the central barrel region about the beam direction, and the two end-caps (or "plugs") which extend the instrumentation to almost hermetic coverage of the interaction region. It measures approximately 27m from end-to-end, is about 10m high and weighs over 5000 tons.

Figure 4.6 is a schematic side-on view of the CDF detector showing all the major detector elements which are discussed in detail in the paragraphs that follow.

CDF Coordinate Systems

The right-handed Cartesian coordinate system of CDF has the z-axis defined by the beam direction, with the x-axis lieing on the plane defined by the Tevatron
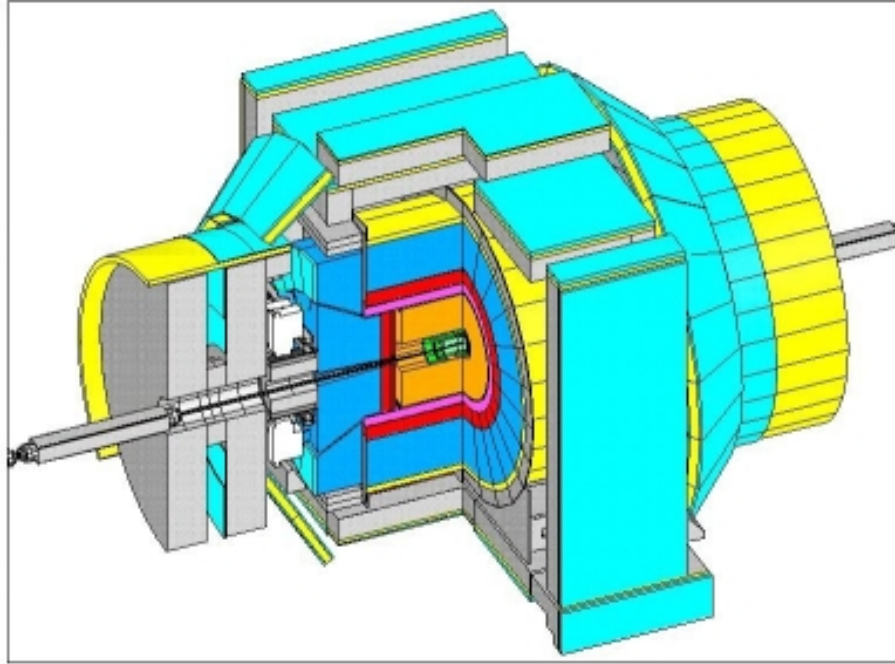
Figure 4.5: Isometric view of the CDF detector. Note the cylindrical symmetry of the detector about the beam pipe.
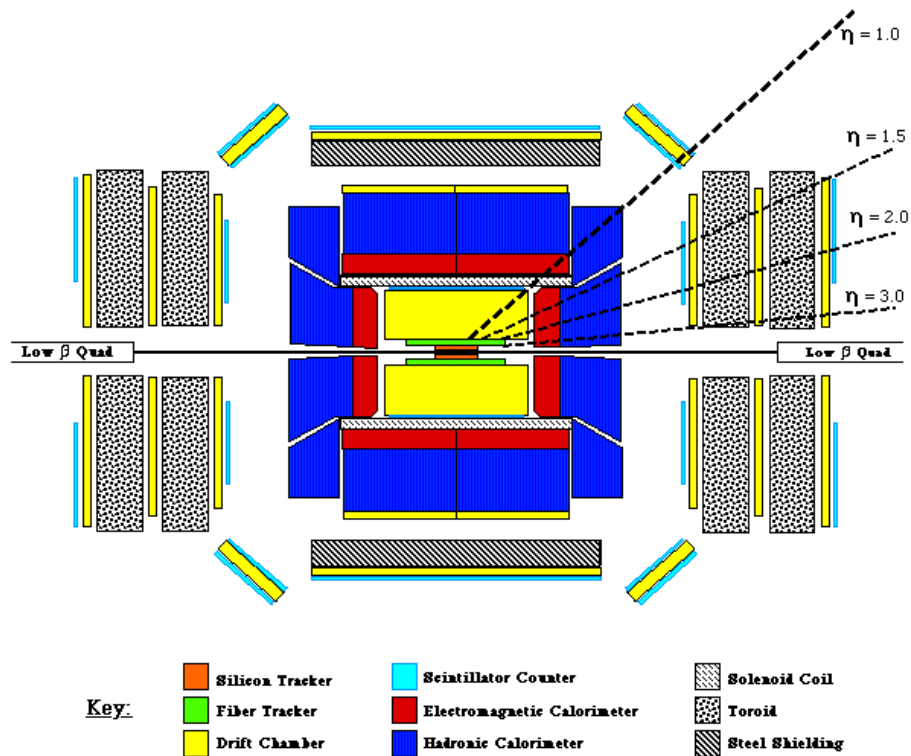
Figure 4.6: Side on view of the CDF II detector, showing all the main components.

ring. Cylindrical polar coordinates are often used also, where the polar angle θ is measured from the z-axis and the azimuthal angle ϕ is measured from the x-axis, in the $x - y$ plane. The coordinates r, ϕ and θ are then defined as:
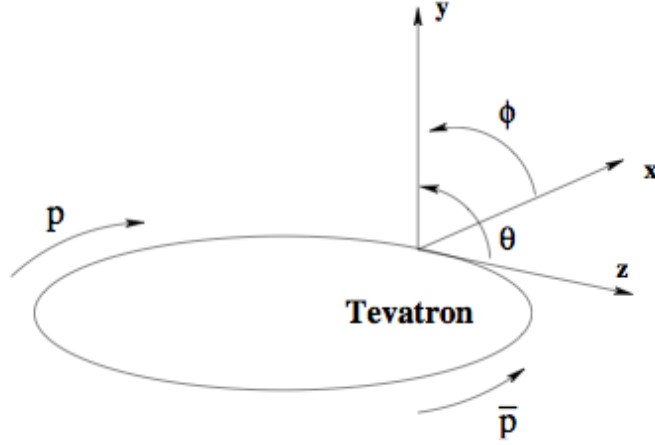
Figure 4.7: The CDF coordinate system.

$$r = \sqrt{x^2 + y^2}$$
$$\phi = \tan^{-1}\left(\frac{y}{x}\right)$$
$$\theta = \tan^{-1}\left(\frac{r}{z}\right)$$

Figure 4.7 illustrates the CDF coordinate system.

As opposed to θ the variable *pseudorapidity* η is often used:

$$\eta = -\log\left(\tan\left(\frac{\theta}{2}\right)\right) \tag{4.3}$$

The partonic nature of hadrons means that even though the proton-antiproton beam have nominally the same energy, the partons that interact can have very different initial momenta in the z-direction. Pseudorapidity is convenient because it is relatively insensitive to boosts along the z-axis.

4.2.1 The Central Tracking System

The main purpose of the tracking system [100] is to reconstruct the trajectories of charged particles coming from the interaction vertex. The tracking system is contained within a superconducting solenoid, 5m in length and 3.2m in diameter, which produces a uniform 1.41 Tesla magnetic field parallel to the beamline. Thus charged particles produced in this region experience a Lorentz force acting in the $x - y$ plane and follow a helical trajectory. Precise knowledge of the curvature of the helix in the $x - y$ plane allows one to determine the particles transverse momentum P_T, that is, the component of the total momentum in the $x - y$ plane:
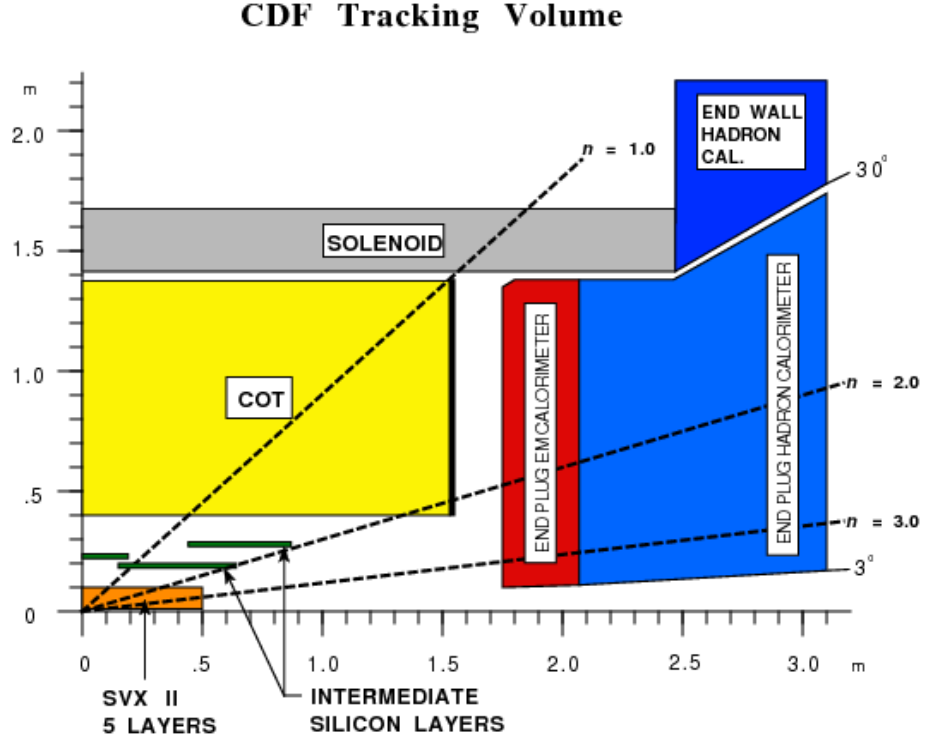
Figure 4.8: Schematic diagram showing the pseudorapidity coverage of the CDF detector elements and plug calorimeters.

$$p_T = B \cdot |q| \cdot e \cdot \rho \qquad (4.4)$$

where q is the particle charge, e is the electron charge, B is the magnetic field and ρ is radius of curvature in the $x - y$ plane. In addition the direction of curvature relative to the magnetic field indicates the particles charge.

The tracking system comprises of four subsystems, which listed in order of their distance from the beampipe are: Layer 00 (**L00**) [92], the Silicon Vertex Detector (**SVX**) [118], the Intermediate Silicon Layers (**ISL**) [21] and the Central Outer Tracker (**COT**) [22]. Figure 4.8 shows a schematic side-on view of the tracking system, and from this one can see the η and radial coverage of the various systems.

The Central Outer Tracker

The Central Outer Tracker (COT) [22] is a cylindrical open-cell drift chamber that occupies the radial region 40 to 138cm and measures 310cm along the z-axis, allowing for the reconstruction of particle trajectories within the central region $|\eta| < 1.0$.

The basic unit of the COT is the *cell*, an $r - \phi$ view of which is shown in Figure 4.9. Each cell runs the length of the COT and comprises of *potential* and *sensor* wires. The potential wires establish a high voltage electric field within the cell unit. The entire COT volume is filled with a mixture of argon, ethane and
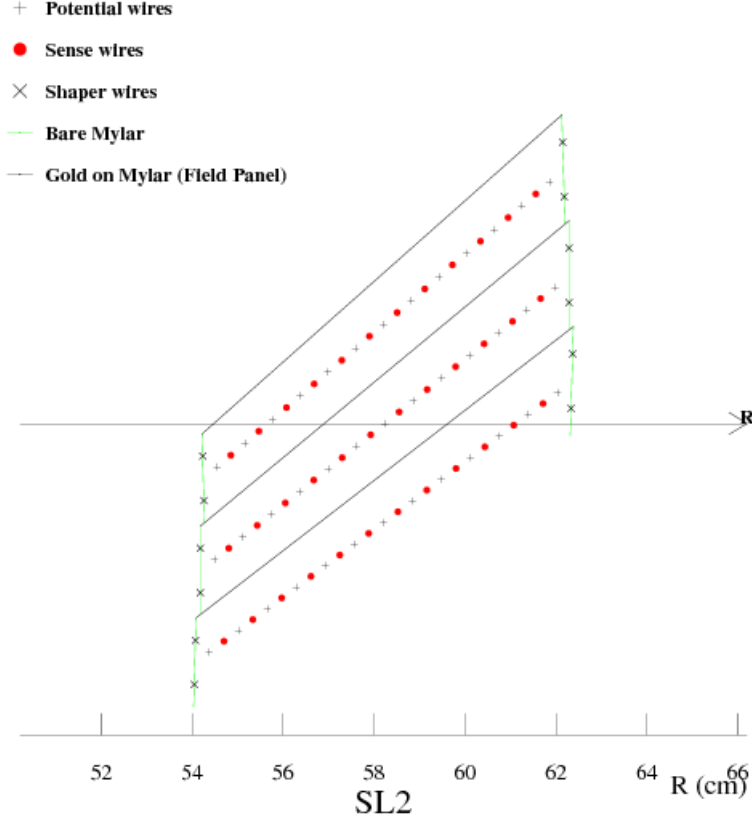
Figure 4.9: End-on $r - \phi$ diagram showing three cells of axial superlayer SL2.

isopropyl alcohol gases in the ratio 500:500:173. When a charged particle passes through the cell this gas becomes ionised, effectively leaving behind a footprint of the particles path. Under the influence of the electric field the ions drift towards the sense wires, and as they approach the wires the field strength increases and more ions are produced in a cascade, amplifying the charge deposited on the sense wire. The timing and charge deposition information of this "hit" is digitized by the COT front-end electronics.

The cells in the COT are arranged into eight *superlayers* (labeled **SL1 - SL8**) as shown in Figure 4.10. Four of the superlayers have their wires arranged parallel to the z-axis (the *axial* layers) and four have their wires offset by $3°$ from the z-axis (the *stereo* layers). Precise knowledge of the timing of the hits, geometry of the cells, the electric field, the drift time and the transmission properties of the sensor wires allow for the reconstruction of 3-dimensional particle tracks. The axial superlayers provide track measurements in the $r - \phi$ plane, and the stereo layers allow for measurements in the $r - z$ plane. The hit position resolution in the COT is around 140μm, resulting in a track momentum resolution of $\sigma_{P_T}/P_T = 0.15\%$.
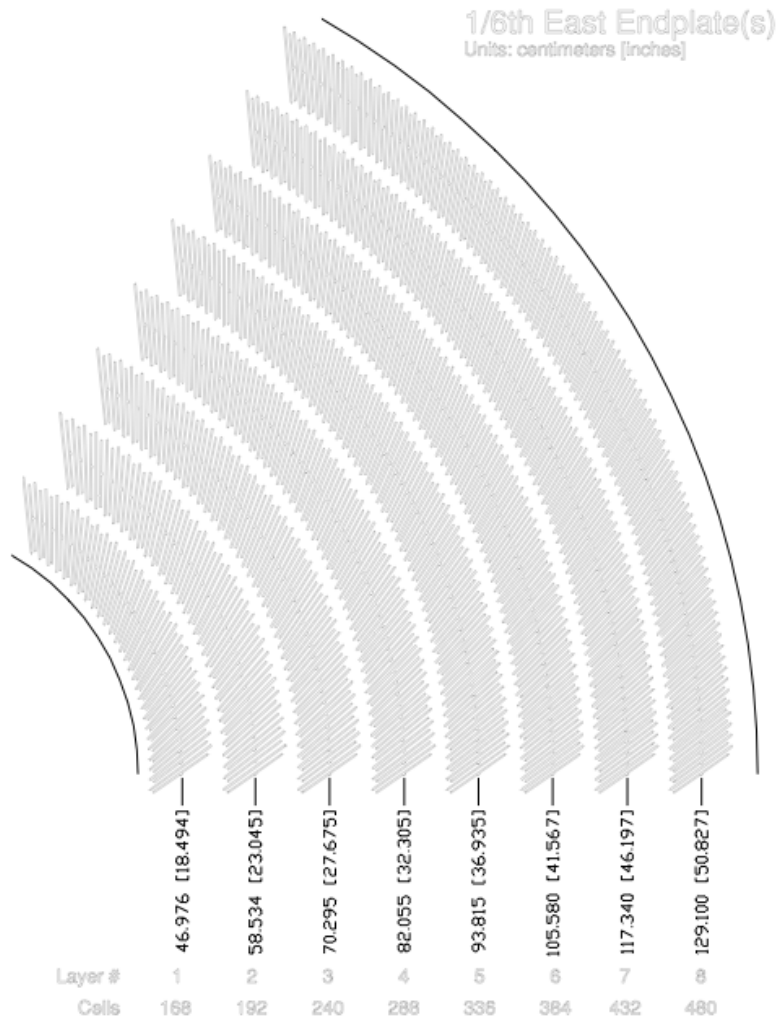
Figure 4.10: End-on $r - \phi$ diagram showing the COT superlayers.

The Silicon Detectors

The silicon detectors (L00 [92], SVX [118] and ISL [21]) are located at small r, close in to the interaction region, and their main use is to provide very high spatial resolution measurements of particle trajectories in this region. As well as helping to improve the tracking resolution this means that one can very accurately reconstruct the interaction vertex and in particular secondary vertices produced by the decay of B mesons. When combined the information from the SVXII and ISL detectors gives an impact parameter d_0 resolution of 40μm and a Z_0 resolution of 70μm[1]. Also, the silicon coverage extends out to $|\eta| < 2.0$ providing the possibility to reconstruct tracks beyond the $|\eta| < 1.0$ region covered by the COT. An end-on $r - \phi$ schematic of the silicon detectors is shown in Figure 4.11.

Layer 00 (L00) is a single-sided radiation-hard silicon microstrip detector that

---

[1]The impact parameter of a track is the transverse distance of closest approach of the track to the primary interaction vertex and Z_0 the point of intersection of the track with the $z - axis$.
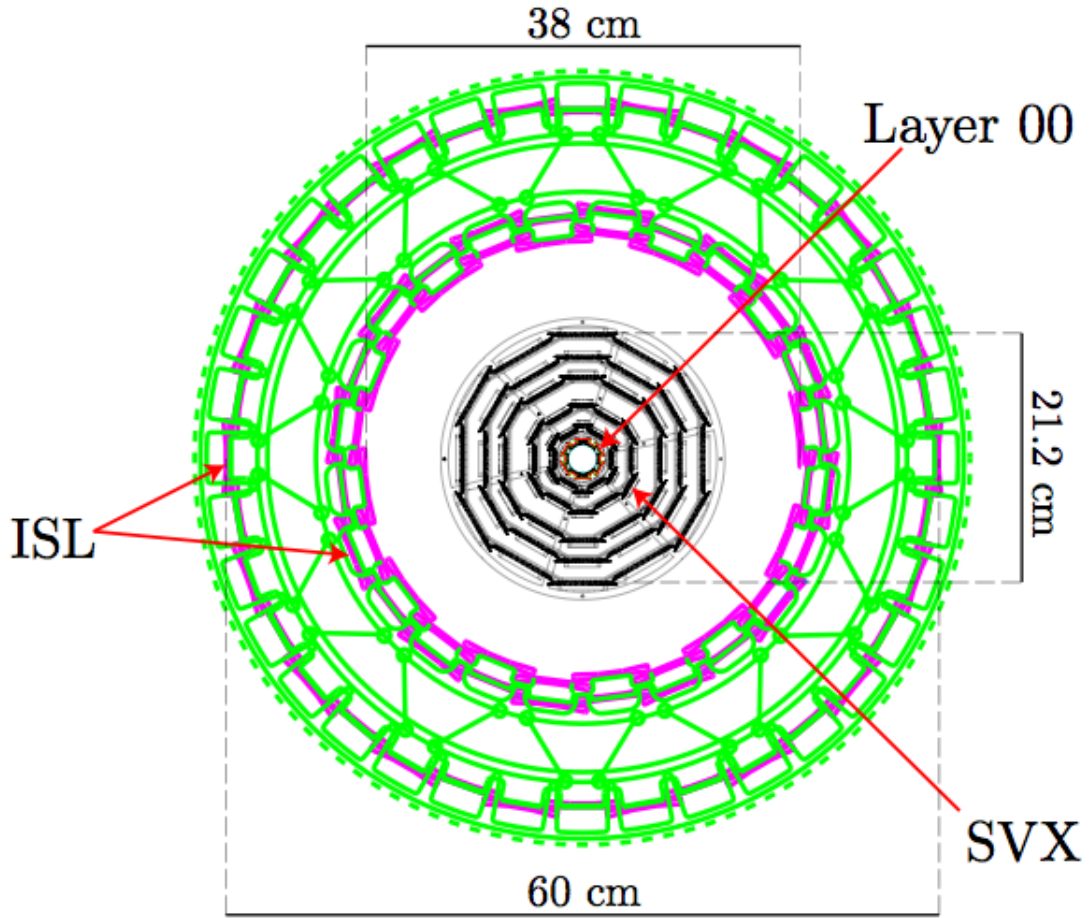
Figure 4.11: End-on $r - \phi$ view of the CDF silicon detectors.

immediately surrounds the beampipe with an inner radius of 1.15cm and an outer radius of 2.1cm. It's primary purpose is to enhance track impact parameter resolution. The Silicon Vertex Detector (SVX) comprises of five layers of double-sided silicon microstrip detectors, covering the radial region from 2.5 - 10.6cm. Like the COT it's design allows for the reconstruction of tracks in three-dimensions, but with a substantially better hit position resolution of 20μm. The Intermediate Silicon Layers (ISL) are comprised of overlapping layers of double-sided silicon microstrip detectors, positioned between at radii between 19 and 30cm.

4.2.2 The Calorimetry System

The purpose of the CDF calorimeter system [73, 97] is to measure the energy and direction of charged and neutral particles. There are two distinct types of calorimeter at CDF, electromagnetic and hadronic, both of which are sampling, that is, consisting of alternating layers of dense, absorbent material and active scintillator material. These layers are segmented into projective "towers", which point back towards the centre of the detector ($z = 0$). The Central Electromagnetic (CEM) [39] and Central
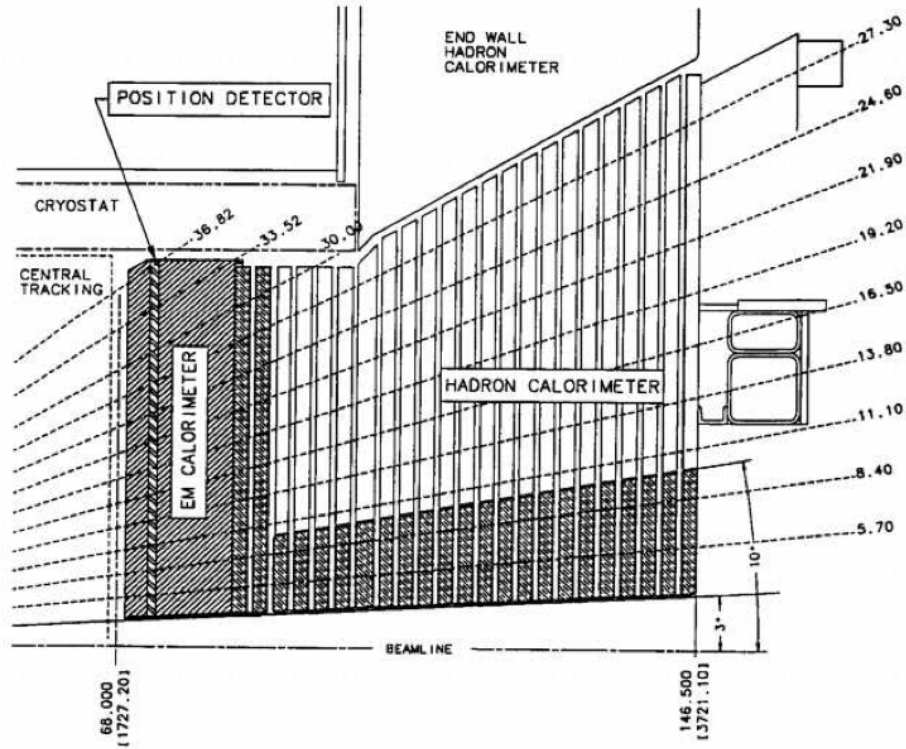
Figure 4.12: Schematic of the forward PEM, PHA and WHA calorimeter systems.

Hadronic (CHA) [41] calorimeters cover the pseudorapidity region $|\eta| < 1.1$ and have complete coverage in ϕ. Immediately outside of the solenoid is the CEM, and immediately behind this is the CHA. The CEM and CHA share the same tower geometry; they are split into two halves about $\eta = 0$, each half organized into 24 wedges in ϕ, with each wedge containing 10 projective towers of width $\Delta\eta = 0.11$. The Plug Electromagnetic (PEM) [27] and Plug Hadronic (PHA) [63] calorimeters cover the region $1.1 < |\eta| < 3.64$. Additionally the Wall Hadronic Calorimeter (WHA) [41] covers the gap between the central and plug calorimeters. The location and geometry of the forward calorimeter systems is shown in Figure 4.12. Table 4.1 shows the vital parameters of the five calorimeter subsystems; their pseudorapidity coverage, thickness in radiation lengths $\chi_0{}^2$ or nuclear interaction lengths $\lambda_0{}^3$ and energy resolution.

The electromagnetic calorimeters are designed to absorb and measure the energy of electrons and photons. When one of these particles is incident on a tower an electromagnetic shower is initiated and propagated by the lead absorber layers, and the energy of the shower sampled by the polystyrene scintillator layers. A high energy particle incident on the scintillator will ionise the material, and the resulting

---

[2]The radiation length χ_0 of a material is the distance that, on average, a high-energy electron will loose all but 1/e of it's energy via bremsstrahlung radiation.

[3]The nuclear interaction λ_0 length of a material is defined as the mean free path of a particle before undergoing an inelastic nuclear interaction.

System	$\mid\eta\mid$ Coverage	Thickness	Energy Resolution
CEM	0.0 - 1.1	$19\chi_0$	$13.5\%/\sqrt{E_T}$
PEM	1.1 - 3.6	$21\chi_0$	$(14.4\%/\sqrt{E}) + 0.7\%$
CHA	0.0 - 0.9	$4.5\lambda_0$	$50\%/\sqrt{E}$
WHA	0.7 - 1.3	$4.5\lambda_0$	$75\%/\sqrt{E}$
PHA	1.2 - 3.6	$7\lambda_0$	$80\%/\sqrt{E}$

Table 4.1: Parameters of the CDF calorimeter subsystems. CEM and PEM energy resolutions are determined using an electron test beam, and CHA, PHA and WHA energy resolutions determined with a pion test beam.

atomic transitions result in the energy of ionisation being converted into visible light. This light is then transmitted to photomultiplier tubes (PMTs) located at the tops of the towers via wavelength shifting fibers and acrylic light guides. Integrating the charge collected by the PMTs gives a measure of the energy deposited in the calorimeter. Figure 4.13 is a schematic of a CEM wedge.

Hadrons are too heavy to initiate electromagnetic showers, but when they reach the hadronic calorimeter they hit the denser steel absorber layers and initiate hadronic showers via strong interactions with the steel nuclei, the energy of which is sampled by scintillator layers.

Examining the details of an electromagnetic shower can reveal useful information on the particle that produced it. The CEM contains two wire proportional chambers for this purpose: the Central Electromagnetic Strip (CES) chamber and the Central Pre-Radiator (CPR) (shown on Figure 4.13). The CES is a 2-dimensional wire strip chamber located 5.9 radiation lengths into the CEM, and allows for the measurement of the transverse profile of the shower where it is expected to be at it's maximum lateral extent. This aids in particle identification as follows:

- The profile yields a measurement of exactly where the particle was incident on the tower. This can be matched to COT tracks, thus improving electron vs photon discrimination.

- Showers initiated by pions ($\pi^0 \to \gamma\gamma$) and prompt photons will have a different transverse shower profile.

- Hadronic showers can also occur in the CEM. However, the pulse height can be determined using the CES, and is different for electromagnetic and hadronic showers.

Similarly the Plug Electromagnetic Strip (PES) chamber provides shower profile measurements in the PEM. The CPR is a proportional chamber which sits between
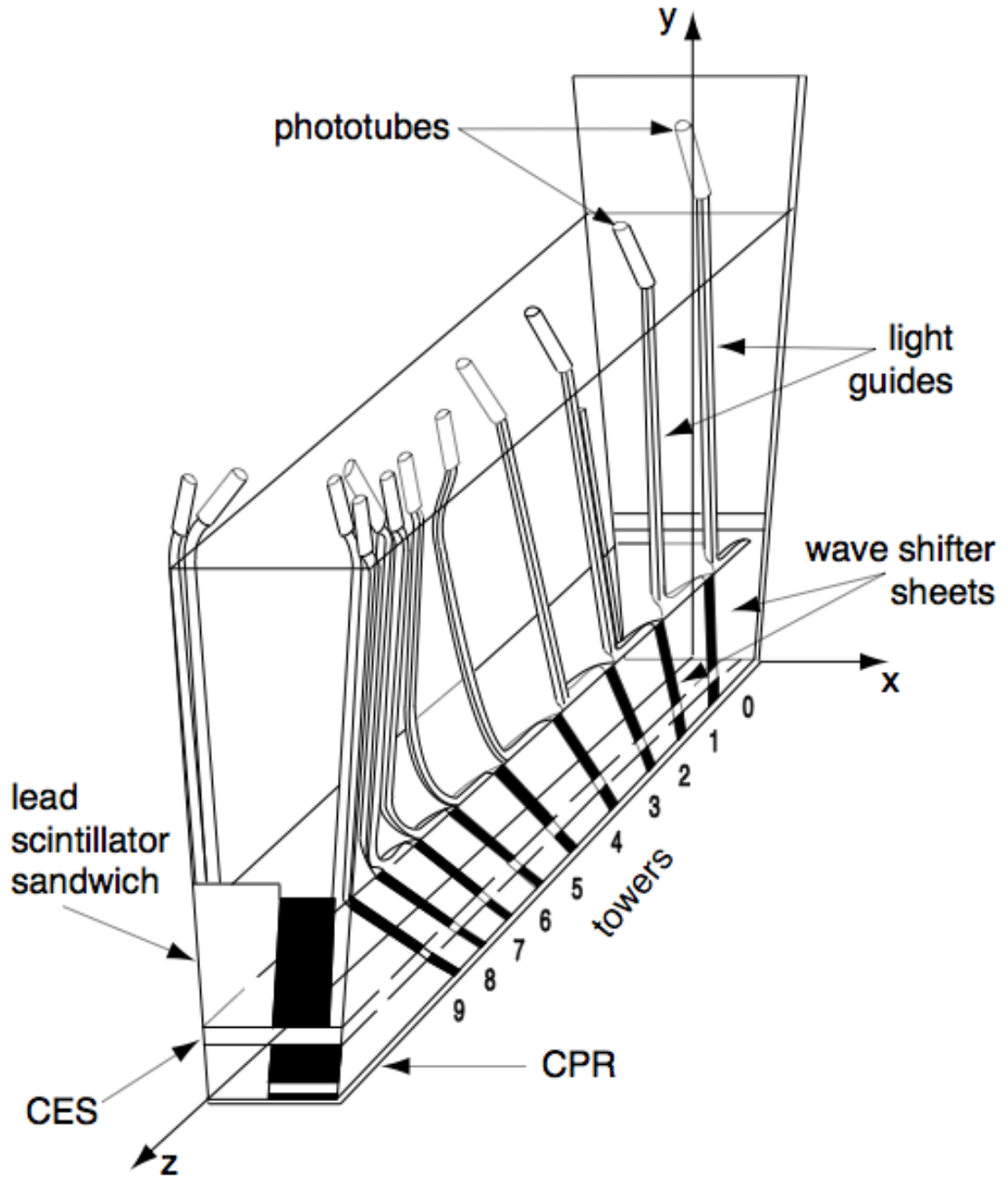
Figure 4.13: Schematic of a wedge in the CDF Central Electromagnetic (CEM) calorimeter.

the first lead layer and the solenoid in the CEM. It provides a measurement of the initial stages of the electromagnetic shower and aids in particle identification in much the same way as the CES, particularly in pion-photon discrimination.

4.2.3 The Muon System

The power P emitted by an electrically charged particle via bremsstrahlung radiation as it traverses a material layer is dependent on the mass m of the particle as [47]:

$$P \propto \frac{1}{m^6} \tag{4.5}$$

Hence the muon, some 207 times more massive than the electron, penetrates the calorimetry system without initiating any electromagnetic showers and only losing energy relatively slowly through ionisation.

Muon detection and measurement is achieved via the Central Muon Detector (CMU), the Central Muon Upgrade (CMP) and the Central Muon Extension (CMX) subsystems [34]. These are the outermost detector subsystems, with $\eta - \phi$ coverage shown in Figure 4.14. They consists of systems of small drift chambers (drift tubes) and scintillation tiles connected to PMTs. Muons will leave an ionisation path in the drift chambers and produce a scintillation signal, allowing for a coarse reconstruction of the muon trajectory. The CMU is separated from the rest of the detector by a layer of steel shielding. This should absorb any particles other than muons which may penetrate the calorimeter, such as charged pions. Thus a charged track in the COT that matches deposits in the CMU is a good muon candidate.

4.2.4 The Trigger and Data Acquisition (TDAQ) System

The detector components described above provide a wealth of information that can potentially be used by physicists to reconstruct in great detail the p$\bar{\text{p}}$ interactions that occur within the detector. However, in order to do this the detector information has to be read out and stored such that it can be analysed "offline" at a later date. This is the role of the CDF Trigger and Data Acquisition (TDAQ) system [54, 100].

At an instantaneous luminosity of $1 \times 10^{32} \text{cm}^{-2}\text{s}^{-1}$ almost every bunch crossing will contain at least one p$\bar{\text{p}}$ interaction (see Figure 6.17). Thus combined with a bunch crossing rate of 2.5MHz, we expect around 2-3 million p$\bar{\text{p}}$ interactions per second. However, the vast majority of p$\bar{\text{p}}$ interactions are soft collisions that are of limited physics interest. The interesting processes, such as the creation of a W boson in the interaction, have much smaller cross-sections and thus occur much more rarely. Thus, purely from a physics analysis point of view it is desirable to substantially filter the events, rejecting the soft interactions. Of course this could be done offline, if it wasn't for the severe technological constraints imposed on recording the detector data associated with 3 million events per second. The permanent storage media used by CDF is magnetic tape, and this can only be written to at a maximum rate of 75Hz, if you are to record the complete set of detector data that is required for offline analysis. Thus an online event rejection factor of about 10^5 is necessary, and it is the role of the trigger system to implement this rejection, whilst still ensuring that interesting physics events are recorded with high efficiency.

Practically the trigger system works by implementing a set of event selection
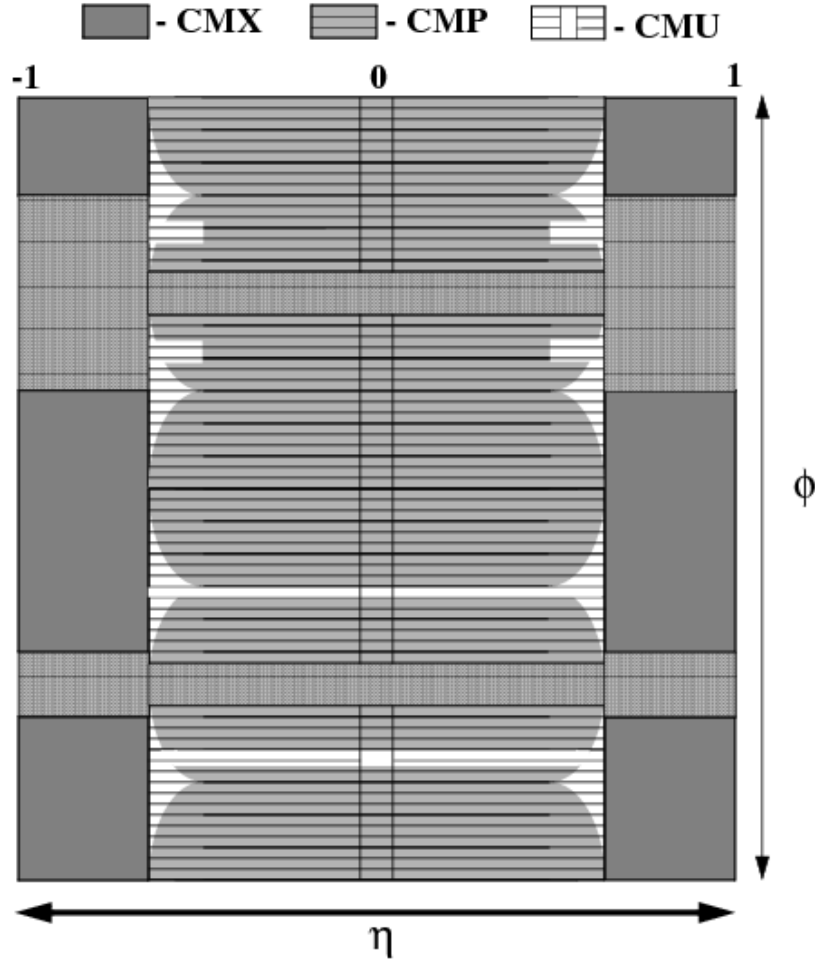
Figure 4.14: Schematic showing the $\eta - \phi$ coverage of the muon detector subsystems.

criteria, using hardware or software based algorithms to analyse detector information "on the fly". The time taken for the trigger system to process the information from an event and make a decision to reject or keep is non-negligible when one is dealing with such rapid bunch crossings. A further design concern of the trigger system is to minimize the "deadtime". This occurs when the trigger system is still processing the previous event when the next beam crossing takes place, and results in the loss of potentially interesting events.

The CDF Run II trigger system is designed to operate with zero deadtime. In order to achieve this goal without throwing away too many interesting events it is necessary to have a three tier system; Level 1, Level 2 and Level 3 (L1,L2 and L3 respectively). Ideally, to make the best informed decision about whether or not to keep an event we would like to reconstruct each event with all available detector data, and essentially this is done at Level 3. However, it takes the L3 processors
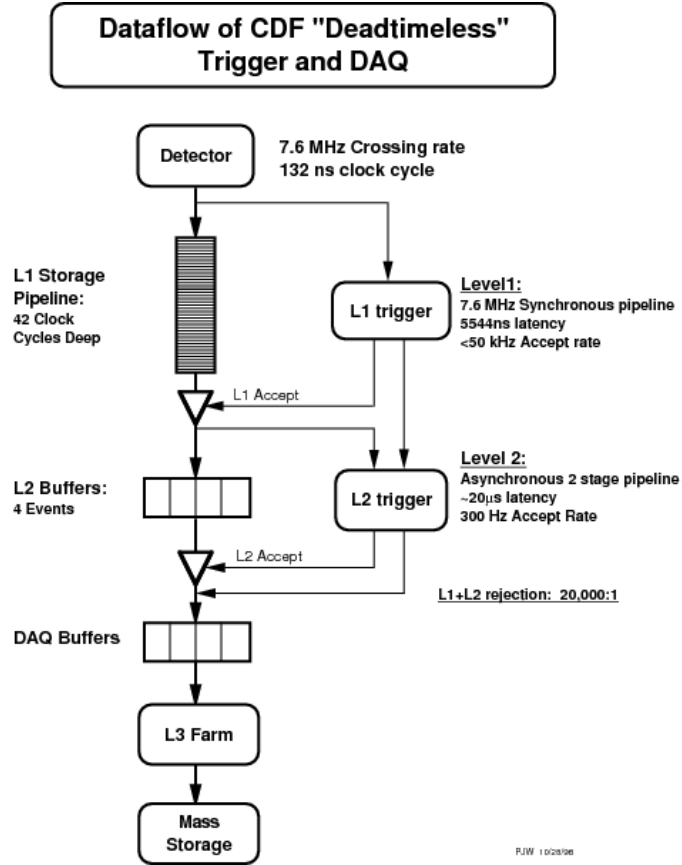
Figure 4.15: Block diagram illustrating the flow of data through the CDF Run II TDAQ system. Taken from [100]

some 20ms to do this, far longer than the bunch crossing time of 396ns. Therefore, three trigger stages of increasing complexity are used, at each stage excluding events so as to reduce the input rate to the next stage, allowing for more complex, accurate and efficient reconstruction. Figure 4.15 is a block diagram illustrating the flow of data through the TDAQ system.

The collection of criteria which specify the acceptance requirements at each level of the triggering process is known as a trigger path. The system allows many different trigger paths to be in operation at any one time, and thus a wide range of different types of physics events have a chance to be recorded each bunch crossing.

Level 1 Trigger

The Level 1 trigger consists of custom-built electronics mounted directly onto the detector. These are organised into three parallel synchronous processing streams which feed inputs of the single Global L1 Decision unit, as illustrated in Figure 4.16. Each stream constructs primitive physics objects from raw detector data; one stream finds calorimeter objects, one muon objects and one COT track objects (using the eXtremely Fast Tracker XFT). The Global L1 Decision unit must then decide whether

RUN II TRIGGER SYSTEM
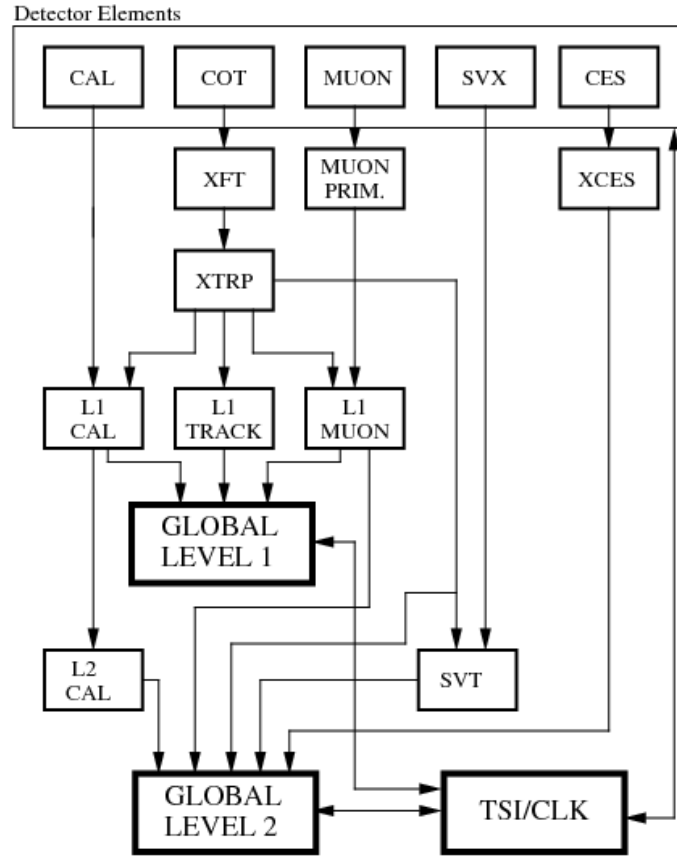
Figure 4.16: Block diagram illustrating the operation of the Level 1 and Level 2 trigger systems. Taken from [100]

the L1 criteria specified by the trigger table have been satisfied, essentially involving a simple counting of the physics objects. This decision process takes 5.544μs. Given that the bunch spacing is 396ns, how can this trigger be deadtimeless? The answer is that within the DAQ electronics of each L1 detector component there is a 42 "bucket" data pipeline. For each $p\bar{p}$ bunch crossing event data enters the pipeline, and is moved along one bucket every 132ns. It will thus take an event 5.544μs to reach the end of the pipeline, giving L1 the time needed to make a decision. The maximum acceptance rate for the L1 trigger is 20kHz, although during the period of data taking for this analysis it was more typically 12kHz.

Level 2 Trigger

The Level 2 trigger is a custom-hardware based system which processes events that have received a L1 accept in a time ordered fashion, and is illustrated in Figure 4.16.

Again in order to produce zero deadtime L2 has buffering for four events. These buffers differ from the pipeline in L1 in that data is resident in the buffer until a decision is made and cannot be lost. If a L1 accept does occur whilst all four of the L2 buffers are occupied then deadtime is incurred. However, the L2 latency is designed to be 20μs and thus with a 20kHz L1 accept rate deadtime should be minimal.

As in L1, each detector subsystem used by L2 has an associated L2 hardware subsystem which forms physics objects from the detector information. All of the information used in the L1 decision is available at L2 and in addition L2 makes use of information from the CES and SVX detector subsystems. The L2 subsystems perform more complex reconstructions than in L1 in order to increase the event rejection. A good example of this is the Level 2 Cluster Finder (L2CAL). In the L1CAL system only individual calorimeter tower energies are considered, and the L1 Global Decision unit makes a decision based on the number of towers with energies above a certain threshold. However, the energy of a hadronic jet is in general spread over several towers. Therefore the L1 thresholds must be set much lower than the jet energy to provide an efficient jet trigger, but this results in rates that are too high for readout into L3. To provide a reduction in the jet trigger rates at L2 the L2CAL subsystem performs a simple clustering of calorimeter towers, and then L2 triggers on the E_T of this cluster. The L2 trigger decision is made by the Global L2 Decision unit, which collates the information from each of the subsystems in order to determine whether the L2 trigger requirements have been satisfied. The L2 trigger is designed to work with a maximum acceptance rate of 300Hz.

Level 3 Trigger and DAQ

The Level 3 trigger system is closely connected to the Data Acquisition System (DAQ), and the flow of data through this final stage is shown in Figure 4.17. The L3 trigger system is a processing farm of 136 dual-CPU processing nodes which reconstruct the event using all available detector data and significantly more complex algorithms than previous stages, allowing for detailed particle identification and event topology criteria to be tested. Results from the lower trigger levels are used to drive the algorithms. For example, if the L2 trigger indicates the presence of an electron candidate only, only those portions of the algorithms relevant to verifying it are invoked. Should the event pass the L3 trigger requirements the Data-Logging subsystem delivers the event to mass storage (i.e. magnetic tape) and also to online monitoring processes which verify that the detector, trigger and DAQ systems are functioning correctly. The maximum accept rate for Level 3 is 75Hz, the rate at which events can be written to magnetic tape.

The DAQ is responsible for collecting data fragments from the front-end elec-
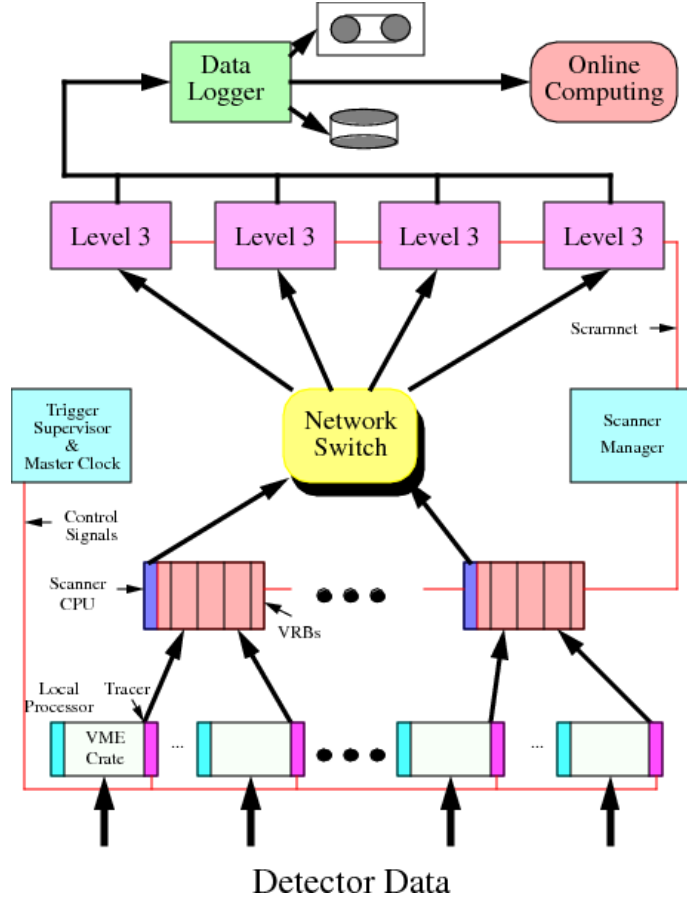
Figure 4.17: Block diagram illustrating the flow of data through the DAQ and Level 3 Trigger systems.

tronics systems for events satisfying the L2 trigger and sending them up to the L3 trigger system. This happens via "scanner" CPUs, which readout the full event data from the front-end electronics and send it to the network switch. The network switch then organises distribution of the event fragments among the L3 processors.

4.2.5 Luminosity Measurement at CDF

Equation 4.2 tells us that in order to determine from the number of events observed the cross-section for a particular process one needs an accurate measurement of the luminosity. Equation 4.1 specifies the instantaneous luminosity in terms of beam parameters, and although it is useful in terms of understanding how the luminosity can be increased, unfortunately these parameters cannot be determined accurately enough to give a precise luminosity measurement. Instead, we have to effectively work backwards from Equation 4.2, calculating the luminosity by counting the number of events for a process that has a well-known and large cross-section. At the Tevatron inelastic $p\bar{p}$ interactions are the best candidate for this job, having a cross-section of $\sigma_{p\bar{p}} = 60.7 \pm 2.4$mb at $\sqrt{s} = 1.96$TeV [96] (determined from E710,

E811 and CDF Run I experimental data). The luminosity is calculated using the following [59]:

$$\mathcal{L} = \frac{\mu \cdot f_{BC}}{\sigma_{p\bar{p}}} \tag{4.6}$$

where f_{BC} is the bunch-crossing frequency and μ is the average number of inelastic $p\bar{p}$ interactions per bunch crossing.

At CDF μ is determined using the Cerenkov Luminosity Counter (CLC) [16]. This detector consists of two CLC modules, one in each of the forward and backward end-plug calorimeters, covering a pseudorapidity range of $3.7 < |\eta| < 4.7$. Each module contains 48 conical Cerenkov counters which point towards the interaction region. The counters are filled with isobutane gas at atmospheric pressure. At the end furthest from the interaction region are light collectors and PMTs in order to collect the ultra violet Cerenkov light that is emitted by high momentum particles when they pass through the gas-filled counter. Prompt particles produced in inelastic $p\bar{p}$ collisions will pass through the entire length of the counter and thus produce a large (~ 100 photoelectrons) PMT signal. Conversely, background processes such as interactions of the beam-halo will be incident at larger angles to the counter axis and hence produce a smaller signal. By setting a suitable threshold one can effectively eliminate these backgrounds, and the number of "hits" in the CLC can be used to determine the average number of inelastic $p\bar{p}$ interactions per bunch crossing as follows:

$$\mu_\alpha = \frac{\overline{N}_\alpha}{\overline{N}_\alpha^1} \tag{4.7}$$

where, for a certain threshold requirement α, $\overline{N}_\alpha^1$ is the average number of CLC hits for a single $p\bar{p}$ interaction (determined from low luminosity running) and $\overline{N}_\alpha$ is the average number of CLC hits per bunch crossing. Combining with Equation 4.6 gives [59]:

$$\mathcal{L} = \frac{\overline{N}_\alpha \cdot f_{BC}}{\overline{N}_\alpha^1 \cdot \xi_\alpha \cdot \sigma_{p\bar{p}}} \tag{4.8}$$

where ξ_α is the CLC acceptance, the probability to detect a single $p\bar{p}$ interaction, that has to be calculated using a CDF tuned simulation of minimum bias events.

The uncertainty on the CLC luminosity measurement is 6% [95], dominated by the error on the CLC acceptance and inelastic $p\bar{p}$ cross-section determination.

Chapter 5

The Measurement Definition

5.1 Previous Vector Boson + Jets Measurements

The first measurements made of jets produced in association with vector bosons at hadron colliders were made by the UA1 and UA2 experiments at the CERN $p\overline{p}$ collider in the 1980's. These measurements [98] [28] were of the ratio of the experimentally observed production rate of W events produced in association with one jet to that with no jets. From such a ratio a measurement of the strong coupling α_s at the scale $Q^2 = M_W^2$ was extracted, assuming a certain renormalisation scheme and PDF set.

The first Tevatron vector boson plus jets measurement was made by D0 in 1995 [2] and was essentially a repeat of the UA1/UA2 strong coupling determination measurements. In 1997 CDF Run I published the first measurements of W + Jets [12] [20] and Z + Jets [10] inclusive cross-sections. The inclusive production cross-sections for $p\overline{p} \rightarrow W+ \geq n$ Jets and $p\overline{p} \rightarrow Z+ \geq n$ Jets, for jets exceeding a transverse energy threshold of 15 GeV, were measured using 108pb^{-1} of data. Statistics allowed for n up to 5 (4 in Z case). In addition, the distribution of W/Z + Jet events with respect to jet transverse energies, separation of jets in $\eta - \phi$ space and jet-jet invariant masses were determined. Although attempts were made to correct these distributions for backgrounds, they fell short of a true differential cross-section determination. The cross-section measurements were made at the parton level, that is, the energies of the jets were corrected such that on average they equalled the energy of the parent final state parton. The inclusive cross-sections and kinematic distributions were compared to the detector-simulated predictions of ELO generators.

In 1998 CDF Run I published a measurement of the $\sigma(p\overline{p} \rightarrow W+ \geq 1 \text{ Jet})/\sigma(p\overline{p} \rightarrow W)$ cross-section ratio for jet transverse energy thresholds ranging from 15 to 95 GeV [13], using 108pb^{-1} of data, and made comparisons to NLO predictions. Other than the measurement presented in this thesis, this represents the only measurement

of a differential W + Jets production cross-section with respect to jet kinematic variables.

So far at CDF Run II an update of the Run I inclusive W + Jets cross-section and kinematic distributions measurement has been made using 127pb^{-1} of Run II data [107]. These measurements were again made at the parton level, with results entirely consistent with those observed at the lower Run I energies.

As well as direct W/Z + Jets cross-sections measurements other measurements have been made at the Tevatron which provide similar tests of QCD predictions at vector boson mass scales. Both CDF and D0 have made measurements of the differential W or Z boson production cross-section with respect to the boson transverse momentum. Vector bosons directly produced in the hard scatter acquire transverse momentum by recoiling against final state quarks or gluons, at leading order via the processes of Figure 3.8. At low boson P_T one is sensitive to non-perturbative physics of soft and collinear gluon radiation, but at higher P_T the production rate should be described by perturbative QCD predictions. CDF measured the Z boson differential transverse momentum cross-section using 110pb^{-1} [19] and D0 measured both W and Z boson differential transverse momentum cross-section using a similar amount of data [6] [7].

5.2 Our Measurement Definition

If a measurement is to be truly useful in terms of theoretical comparisons it must be independent of the detector and technique used to make the measurement. The inclusive $p\bar{p} \rightarrow W+ \geq n$ Jets cross-section measurements of the previous studies described above [20] [107] have been made in this way. However, in order to test QCD predicitions in greater detail one would like also to be able to compare theoretical predictions of jet kinematic distributions, such as the transverse energy spectra of the jets, to the data. In previous studies (with the exception of [13]) the only option for such comparisons has been to run the Monte Carlo W + Jets event samples through the CDF detector simulation and W event selection criteria such that they can be compared to (background corrected) distributions of W + Jet candidate events in the data. Unfortunately theorists do not generally have access to the CDF detector simulation, let alone the W selection code used, and thus such comparisons are only ever a one-off. Ultimately what is desired in order to explore the kinematic predictions of perturbative theory is a differential cross-section measurement with respect to these variables which fully takes into account all detector effects.

In addition to detector independence, we wish to ensure that this measurement is as independent as possible from any theoretical models used. We cannot have the measurement already biased by the very theoretical models we will ultimately be

comparing to. However, we shall see that in order to make the measurement, and in particular to correct for certain detector effects, it is necessary to use detector simulated W + Jets Monte Carlo samples. The dependence of such corrections on the theory can be assessed by varying the parameters of the Monte Carlo. Where dependence is observed there should be an associated systematic error which reflects this.

Measurement of the W + n jet cross-section involves the identification of such events in our data via the decay products of the W boson. W bosons can decay both leptonically and hadronically, but the hadronic channel is virtually indistinguishable from the more prolific multijet QCD production. Due to the relative ease of their identification we select only W$\rightarrow$eυ "candidate" events from the data, and hence our measurement is of the cross-section $\sigma_{p\bar{p}\rightarrow W} \times BR(e\nu)$, which we shall refer to as the W$\rightarrow$eυ cross-section. An inclusive W$\rightarrow$eυ cross-section can be defined in the following way:

$$\sigma = \frac{C - B}{\mathcal{A} \cdot \mathcal{L}} \tag{5.1}$$

Where C is the total number of W$\rightarrow$eυ candidate events observed in the data and B the estimated number of background events in this sample. This background comprises of non-signal events that nevertheless reproduce the W$\rightarrow$eυ signature and thus make it into your candidate sample. The numerator, $C - B$, is then the actual number of W$\rightarrow$eυ signal events you believe have been observed. The acceptance factor $\mathcal{A}$ accounts for the "loss" of real W$\rightarrow$eυ events due to the W selection criteria you necessarily impose on the data to form the candidate sample. Finally the integrated luminosity of the data sample, $\mathcal{L}$, is necessary to scale the W$\rightarrow$eυ events counted to a cross-section.

A differential cross-section measurement can be made by binning the W$\rightarrow$eυ sample in the particular kinematic variable of interest and making the cross-section calculation in each bin. Clearly to do this we require not only the distribution of candidate events in the variable, but also a knowledge of how the background events and acceptance factors are dependent on the variable. As it turns out, this proves to be one of the most challenging aspects of such a measurement, and is detailed in Chapters 7 and 8. In this study we make a measurement of the differential W$\rightarrow$eυ + ≥ 1 jet cross-section with respect to the E$_T$ of the highest E$_T$ jet[1], the differential W$\rightarrow$eυ + ≥ 2 jet cross-section with respect to the E$_T$ of the second highest E$_T$ jet and so on up to the differential W$\rightarrow$eυ + ≥ 4 jet cross-section, where each jet is required to pass a rapidity restriction of $|\eta| < 2.0$. These cross-sections shall be written as:

---

[1]Throughout this study we shall refer to the highest E$_T$ jet as the first jet, the second highest E$_T$ jet as the second jet and so on.

$$\frac{d\sigma_{\geq 1j}}{dE_T^{1stj}}; \ \frac{d\sigma_{\geq 2j}}{dE_T^{2ndj}}; \ \frac{d\sigma_{\geq 3j}}{dE_T^{3rdj}}; \ \frac{d\sigma_{\geq 4j}}{dE_T^{4thj}}$$

The measurement is made down to a minimum jet E_T of 15 GeV. The same information can easily be integrated to form the inclusive W$\rightarrow e\upsilon$ + $\geq$n jet cross-sections for 1,2,3 and 4 jets with a range of minimum jet E_T thresholds. We also measure the W$\rightarrow e\upsilon$ + $\geq$ 2 jet differential cross-section with respect to the ΔR separation and invariant mass of the first and second jets, written as:

$$\frac{d\sigma_{\geq 2j}}{d\Delta R_{jj}}; \ \frac{d\sigma_{\geq 2j}}{dM_{jj}};$$

where the two jets are required to pass a minimum E_T threshold of 15 GeV and the $|\eta| < 2.0$ rapidity restriction. The dijet invariant mass, M_{jj}, is defined as:

$$M_{jj}^2 = (E_1 + E_2)^2 - (\overline{p}_1 + \overline{p}_2)^2 \qquad (5.2)$$

Where the 4-vector of each jet, $(E, \overline{p})$, is determined using the JetClu recombination scheme given in Equation 3.12.

It is important to note that in making this measurement we are interested not in the properties of the W boson, but in the kinematics of the jets that are produced in association. The presence of the W boson, detected by the signatures of its high P_T electron and neutrino decay products, ensures that the hard scattering process has occurred at a scale that allows perturbative QCD predictions to be made. We measure the differential cross-section with respect to jet kinematic variables in order to test these QCD predictions. Thus this is really a QCD measurement, with the W simply acting as a clean and easily identifiable "trigger" for high Q^2 QCD events. With this philosophy in mind, we have made a further departure from previous W + Jets measurements in redefining what we mean by the W$\rightarrow e\upsilon$ + $\geq n$ jet cross-section. Instead of this relating to the rate of production of all W bosons that decay to electrons in association with n jets, we modify the definition such that it is a measure of the rate of W + n jet production where the W decay is restricted to the following region of phase space:

$$
\begin{aligned}
\upsilon \ P_T \ &> \ 30 \text{ GeV} \\
e \ P_T \ &> \ 20 \text{ GeV} \\
e \ \eta \ &< \ 1.1 \\
W \ M_T \ &> \ 20 \text{ GeV} \qquad (5.3)
\end{aligned}
$$

where M_T is the W boson transverse mass[2]. In doing so we do not at all compromise the measurements usefulness. The decay of the W boson is a well understood electroweak phenomenon, the W mass and width being known with great precision [70], and thus in comparing perturbative W + Jets predictions to our results theorists can with confidence impose the same W decay restrictions in the theoretical predictions. In Chapter 8 we explain that modifying the cross-section definition in this way amounts to a redefinition of the acceptance factor $\mathcal{A}$ which reduces the significance of theoretical model dependence in our measurement.

Another important aspect of the cross-section measurement is the exact jet definition used. In line with previous W + Jets measurements we use the CDF cone jet algorithm JetClu to cluster the towers into jets and define the calorimeter-level jet energies. Clearly to avoid detector dependence in the measurement one has to correct the calorimeter-level jet energy such that it is at least on average equal the total energy of all the hadrons contained within the jet [45]. Such corrections are derived for a generic jet using dijet QCD Monte Carlo event samples that have been passed through the CDF detector simulation, comparing the energy of a jet clustered at the hadron-level with the energy of the same jet reconstructed in the calorimeter (see Section 6.2.2 for details). After this jets are said to have been corrected to the hadron-level. However, previous W + Jets measurements [20] [107] took this one step further by attempting to correct the hadron-level energies to the matrix element parton level. That is, they attempt to account for the energy of the parent parton that is "lost" from the jet cone due to the parton showering and hadronization processes. This so-called "out-of-cone" correction [45] is done in a similar way to the absolute correction, using Monte Carlo event samples to compare the energy of a jet at the hadron level to the energy of the parent parton that it originated from. However, this correction is clearly entirely dependent on the particular parton showering and hadronization models used by the Monte Carlo sample. In our measurement we correct the jets to the hadron-level only, a correction that hopefully removes detector dependence in a model independent fashion. This also affords extra flexibility for comparisons to our results in the future when more advanced fragmentation models may be available.

However, if we are to make a truly detector independent hadron-level measurement there is one further effect of the detector that must be taken into account; the resolution of the calorimeter energy measurements. The absolute corrections correct the energy of jets to be *on average* equal to the hadronic energy. However, even if a "perfect" absolute correction was applied to the data, the detector resolu-

---

[2]Since the CDF detector is not hermetic, only the transverse component of the neutrino momentum can be reconstructed. This means that only the transverse mass of the W boson, $M_T = \sqrt{2E_T^e E_T^\nu (1 - cos\Delta\phi_{e\nu})}$, can be reconstructed from the electron and neutrino decay products.

tion can result in a measured jet E_T spectra that is different from the true hadron spectrum. Thus it is necessary to apply an additional bin-by-bin "unsmearing" correction to the measured cross-sections. This is extracted by comparing hadron-level and detector-level cross-sections as a function of jet E_T, ΔR_{jj} and M_{jj} in detector simulated W$\rightarrow$ eυ Monte Carlo samples, detailed in Chapter 9.

Previous W + Jets analyses [20] [107] also made other jet corrections in their attempts to return to the energy of the parent parton. These corrected for sources of jet energy that were not associated with the hard-scatter process under study:

1. Energy arising from multiple p$\overline{\text{p}}$ interactions in the same bunch crossing.

2. Energy associated with the underlying event.

At the Tevatron the high instantaneous luminosities involved mean that we expect more than one interaction per bunch crossing on average. These extra soft interactions may contribute additional hadrons to a jet, and thus additional energy. The amount of additional energy that on average results from multiple interactions is clearly dependent on Tevatron parameters such as the instantaneous luminosity i.e. it is another detector dependence. It thus makes sense for us to make an on average correction to the jet energy to account for this affect, and this will be described in Section 6.2.3.

Additional hadrons can also result from the interactions of spectator partons in the hard-scatter p$\overline{\text{p}}$ collision, the so-called "underlying event" described in Section 3.6. In contrast to the multiple interaction phenomena, the underlying event is a "physics effect"; something that theorists can model which is not detector dependent. Previous W + Jets analyses used data-driven methods [45] to correct the energy of their jets for the average contribution expected from the underlying event. However, whether or not one should do this depends very much on the theoretical comparisons that are to be made. If the theoretical model to which you compare has an underlying event model, then it may be desirable to leave the energy uncorrected in order to test the success of this model. The results presented in this study are not corrected for the underlying event.

Chapter 6

W→ eυ + n Jet Candidate Event Selection

The first stage in making a measurement of the W→ eυ + n jet differential cross-section is to identify W→ eυ candidate events from the multitude of p$\overline{\text{p}}$ interactions that have been logged. To do this we look for the distinctive two-part signature of the W boson decay; namely a high P_T electron accompanied by a high P_T neutrino. Whereas electrons can be directly detected via tracking and calorimetry measurements, the weakly-interacting neutrino does not leave deposits of any kind in the CDF detector, and thus its presence must be inferred via the constraint of conserved transverse momentum. Section 6.1 describes in detail the process of selecting a W→ eυ event sample.

Once such a sample has been identified our attention turns to the hadronic activity. For each event we examine the jets present and categorise according to the number of jets, the hadron-level jet E_T and the ΔR and invariant mass of the first and second jets. To do so requires a particular jet definition, along with a prescription to correct the E_T of the jets to the hadron-level. The formation of the candidate W→ eυ + n jet E_T distributions is described in full in Section 6.2.

6.1 Forming the W→ eυ Event Sample

6.1.1 Event Vertex Requirements

The hard-interaction that produces the W and associated jets can occur anywhere the proton and antiproton bunches overlap. The x, y, z position of the interaction is known as the event vertex. The average beam width is 30μm in the (x, y) plane [126], but the size of the interaction region in z is much larger. In this analysis we require precise knowledge of the z position of the event vertex such that the kinematic quantities such as the jet and electron E_T are calculated accurately.
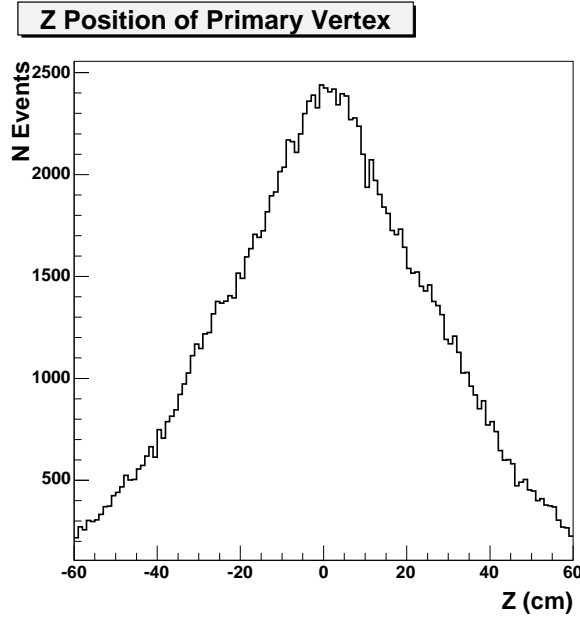
Figure 6.1: Distribution in z of the primary event vertex for W selected events.

Identifying this vertex can be a problem for some analyses, requiring a detailed vertex reconstruction algorithm [56]. However, once we have identified a W candidate through all the requirements that are detailed below, we have a very good handle on the event vertex position. Since the W travels negligible distance before decay the z position of the interaction is given by intersection of the well-defined high P_T electron track and the beamline. The distribution in z of the event vertex for the W selected events is given in Figure 6.1. We make a cut of 60 cm to keep the interaction within the fiducial volume of the detector and to maintain the projective geometry of the calorimeter towers.

6.1.2 High P_T Electron Selection

High P_T electrons leave a distinctive signature in the detector; a straight, well defined and high P_T track in the COT and large, localised and mainly electromagnetic energy deposits in the calorimeter. More specialised measurements can also be made using sub-detector systems such as the CES to enable better electron identification. Our goal is to achieve the highest possible efficiency in identifying electrons whilst minimising the false identification of other "objects", such as photons or highly electromagnetic jets, which can fake electron-like signatures. In order to do this we first use an algorithm to search for and define electromagnetic cluster objects, essentially localised collections of calorimeter towers which have had substantial EM energy deposits, then construct, from specific detector measurements relating to a cluster, a series of high P_T electron identification variables on which we make selec-

tion "cuts". If a cluster passes all the cuts, then it has been successfully identified as a high P_T electron candidate, and is referred to as a "tight electron".

Electromagnetic Clusters and Electron Candidates

An electron incident on the calorimeter will deposit energy via an electromagnetic shower. This shower is unlikely to be contained within one calorimeter tower, and thus if we are to measure the total energy deposited by an electron we need some way of associating multiple towers that contain the energy from one shower process [124]. This is done by searching the calorimeter for seed towers which have a minimum tower transverse energy[1] of 3 GeV. The "shoulder towers" (those adjacent in η) of each seed are then examined, and if they have at least 100 MeV of energy, they are merged with the seed to form a cluster. The maximum cluster size is thus three towers adjacent in η. The total energy and centroid of this cluster is defined using the electromagnetic energy component only. To prevent the sharing of towers between clusters, the seeds are ordered in decreasing transverse energy and the clustering performed on each in turn, with towers that are clustered being removed from the seed tower list.

An EM cluster is then retained as an "electron candidate" if the following conditions are met:

- The total tower E_T exceeds 5 GeV

- The ratio of the total hadronic to electromagnetic energy of the cluster is less than 0.125.

- A good quality COT track can be matched to the cluster, ruling out the possibility that the cluster has been produced by a photon. This matching is done by looking for COT tracks that when extrapolated to the CEM lie within 25cm in the $x-y$ plane and 38cm in the z-direction from the centre of the cluster seed tower. The highest P_T track which fulfills the requirements is considered as the electron track. The track quality requirements are given below.

The COT track that is matched is then also extrapolated to the CES to find a CES shower match for the electron candidate.

Electron Identification Variables and Cuts

In the following paragraphs we describe the variables and the cuts made on these variables in order to identify a high P_T electron:

---

[1]Tower transverse energy $E_T = E \times \sin\theta_D$ where θ_D is the angle of the tower measured from the detector centre.

- **Central**

  The EM cluster should lie within the bounds of the CEM, that is, within the detector pseudorapidity range $\eta_D < 1.1$. This keeps the electron within a region of the calorimeter where the response of the calorimeter is best understood.

- **Fiducial**

  The electron should be incident on a well-instrumented region of the CEM in order for the cluster energy measurements to be reliable. For this the electrons position is taken as that given by the CES shower. The electron should be at least 3cm from the edge of a tower, such that the shower is largely contained within the active region. It should not be in the region where the two halves of the central calorimeter are joined ($|\eta_D| < 0.05$), nor the uninstrumented "chimney" region ($0.77 < \eta_D < 1.0, 75° < \phi < 90°$) where the cryogenic system connects to the solenoid magnet. Additionally the outer half of the last CEM tower is excluded ($1.05 < \eta_D < 1.10$) as it is prone to leakage into the hadronic calorimeter.

- **$E_T > 20$ GeV**

  The total energy E of the electron candidate is defined as the total electromagnetic energy of the towers within the EM cluster, as explained above. This energy is corrected for differences in tower-to-tower gain [55], variations in response within single towers ("face corrections") [127] and changes in the CEM gain over time. The transverse energy E_T is given by $E_T = E \times \sin\theta$, where θ is the angle of the COT track associated with the EM cluster, and is required to be greater than 20 GeV.

- **Had/EM**

  This variable is used primarily to distinguish electrons from jets. The vast majority of the electrons energy should be deposited via an electromagnetic cascade in the EM calorimeter towers, although some small leakage into the hadronic calorimeter is expected. Conversely a jet will most likely deposit the majority of its energy in the hadronic calorimeter. Thus the energy in the hadronic towers corresponding to the electron cluster is examined, and the ratio of hadronic to electromagnetic cluster energy not allowed to exceed the level given by the following formula:

$$\text{Had/Em} < 0.055 + 0.00045 \times E$$

Where the cut is a function of the electron energy E because leakage of electron energy into the hadronic calorimeter increases as a function of electron energy.

- **Isolation < 0.1**

  Again this variable is designed to distinguish electrons from jets. An electron should produce a localised energy deposit that is well separated from other energy in the calorimeter. The electron "isolation" variable is constructed in the following way:

  $$\text{Isolation} = \frac{E_T^{0.4} - E_T^{\text{cluster}}}{E_T^{\text{cluster}}}$$

  Where $E_T^{0.4}$ is the total energy (hadronic and electromagnetic) in a $\eta \times \phi$ cone of radius R $\leq$ 0.4 about the cluster centroid (including the cluster energy itself) and E_T^{cluster} is the cluster energy (electromagnetic only). The fraction of energy within the 0.4 cone that is not cluster energy must not exceed 10%.

- **P_T > 10 GeV**

  The transverse momentum measurement of the COT track associated to the electron should exceed 10 GeV.

- **Track Quality Cuts**

  The COT track associated to the electron should have been reconstructed from at least 3 axial and 2 stereo superlayers with at least 5 hits per superlayer.

- **Energy-Momentum Ratio E/P $\leq$ 2.0**

  A relativistic electron has negligible mass and thus should have an energy-momentum ratio very close to unity. In practice the electron may emit bremsstrahlung radiation as it passes through detector material, and this can distort the energy-momentum picture, since the emitted photons will likely end up in the clustered calorimeter energy, but the momentum as measured by the COT track is reduced. We require that the energy-momentum ratio be less than 2.0, keeping most of the bremsstrahlung electrons without introducing too many fakes. However, as electron momentum increases the COT momentum resolution gets progressively worse whilst the calorimeter energy resolution is improving, and the E/P measurement becomes inaccurate. Thus we remove the E/P cut if the electron track P_T exceeds 50 GeV.

- **Strip Chamber Profile $\chi^2_{\text{str}} < 10$**

  A transverse profile of the electromagnetic shower at shower maximum is provided by the CES detector. The pulse height in z of the shower profile of the

electron candidate can be compared to that of test beam electrons. If the χ^2 of this comparison is greater than 10 the shower profile is not considered electron-like enough and the candidate rejected.

- **CES-COT Track Matching**

  As described above the electron candidate has associated to it a COT track and a CES shower. The CES shower position is required to be in good agreement with the location of the track when extrapolated to the CES detector. The constraints on their separation in the $r - \phi$ plane, Δx, and z direction Δz are given below:

$$-3.0 cm < Q_e \cdot |\Delta x| < 1.5 cm$$
$$|\Delta z| < 3.0 cm$$

  Where Q_e is the charge on the electron (as given by the direction of the COT track curvature). The asymmetry in the Δx requirement is due to the bending of the electron track in the magnetic field which may take it away from the shower centroid that includes bremsstrahlung photons.

- **Lateral Energy Sharing** $L_{shr} < 0.2$

  The electromagnetic shower produced by an incident electron will generally not be contained in a single tower, and some energy will be deposited in the shoulder towers. From test beam data we can parameterise how much lateral energy sharing we expect from an electron, and construct the following L_{shr} variable as a discriminant:

$$L_{shr} = 0.14 \cdot \sum_i \frac{E_i^{adj} - E_i^{exp}}{\sqrt{(0.14)^2 \cdot E + (\Delta E_i^{exp})^2}} \tag{6.1}$$

  Where the sum is over the two (or one) shoulder towers of the EM cluster, E_i^{adj} is the energy of the shoulder tower, E_i^{exp} is the expected energy in that shoulder tower from the test beam parameterisation, ΔE_i^{exp} is the uncertainty on that expected energy and $0.14 \cdot \sqrt{E}$ is the uncertainty on the cluster energy. Essentially the L_{shr} variable tells us how significant is the disagreement between the test beam lateral energy sharing and that of the cluster, and we require this to be less than 0.2.

Figures 6.2 and 6.3 show the distribution of each of the electron ID variables. These are plotted for electromagnetic clusters that are required to pass all the electron ID requirements with the exception of the variable being plotted, which remains untested. Thus these distributions show the shape of the ID variables in high quality
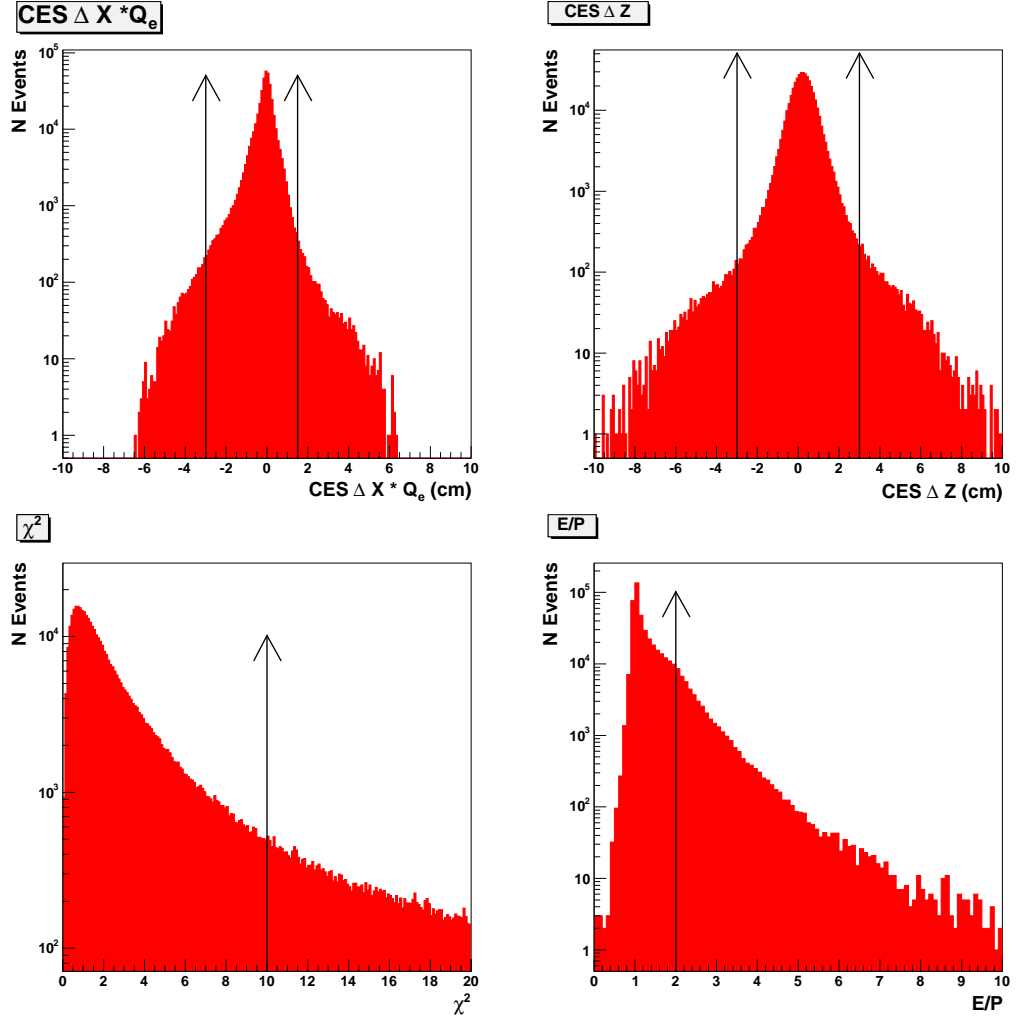
Figure 6.2: The distribution of the CES-COT track matching $\Delta X * Q_e$ and ΔZ, χ^2_{str} and E/P electron ID variables, for electromagnetic clusters which are required pass all the electron ID requirements with the exception of the variable being plotted, which remains untested. The position of the tight electron selection cut is marked with an arrow.

electromagnetic objects. Note that the peculiar pedestal shape in the Had/Em and L_{shr} distributions is due to the trigger requirements that are already imposed on these electromagnetic clusters, as will be detailed in Section 6.1.5.

6.1.3   High P_T Neutrino Selection

Since neutrinos do not interact at all with the detector their presence has to be inferred from the constraints of energy-momentum conservation. In a p$\bar{\text{p}}$ interaction we can use the constraint of momentum conservation in the plane transverse to the beamline because we know that the initial-state protons have nominally zero transverse momentum on interaction. However, we cannot use the same constraint on the total momentum component in the z direction, because we do not know
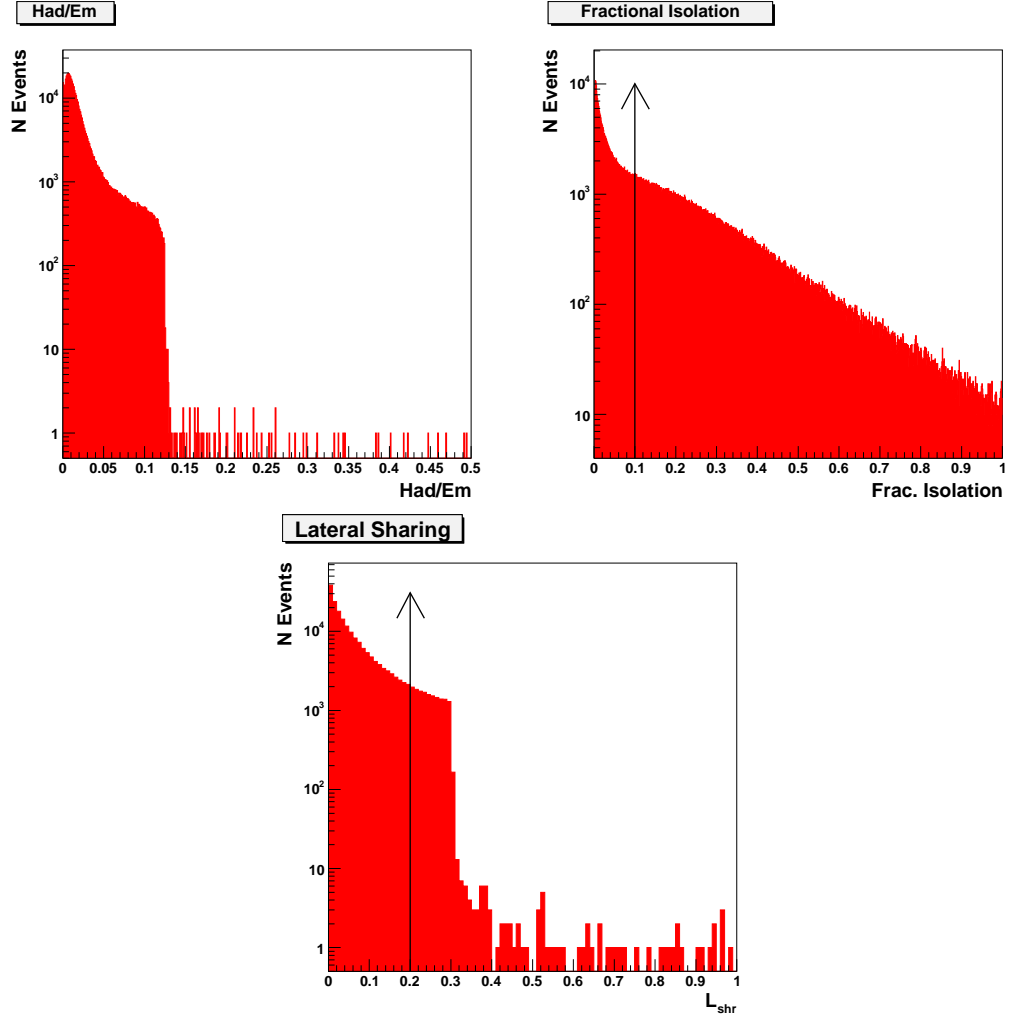
Figure 6.3: The distribution of the Had/Em, fractional isolation and L_{shr} electron ID variables, for electromagnetic clusters which are required to pass all the electron ID requirements with the exception of the variable being plotted, which remains untested. The position of the tight electron selection cut is marked with an arrow (not possible for Had/Em since this is a sliding cut dependent on the electron energy).

the beam axis component of the initial-state partons. If we vectorially sum the transverse energy measured by each tower in the calorimeter, making use of both the central and forward calorimeters to give us the most hermetic coverage possible, we will obtain a vector in the x-y plane, $\hat{\mathbf{E}}_{\mathbf{T}}^{\mathbf{sum}}$, which indicates the magnitude and direction of the transverse momentum imbalance of the event:

$$\hat{\mathbf{E}}_{\mathbf{T}}^{\mathbf{sum}} = \sum_i \hat{\mathbf{E}}_{\mathbf{T},i}^{\mathbf{tower}} \tag{6.2}$$

Where the sum is over all calorimeter towers i and $\hat{\mathbf{E}}_{\mathbf{T},i}^{\mathbf{tower}}$ is calculated using the event primary vertex. If there is no hard-scatter neutrino present one would hope that the magnitude of this vector, $|\hat{\mathbf{E}}_{\mathbf{T}}^{\mathbf{sum}}|$, would be zero to within the limits of the

calorimeter energy resolution. If a neutrino is present, then the zero total transverse momentum constraint means that the neutrino transverse momentum vector is given by $\hat{\nu}_{\mathbf{T}} = -\hat{\mathbf{E}}_{\mathbf{T}}^{\mathbf{sum}}$, as demonstrated in the vectorial equation below:

$$
\begin{aligned}
\hat{\nu}_{\mathbf{T}} + \hat{\mathbf{E}}_{\mathbf{T}}^{\mathbf{sum}} &= 0 \\
\therefore\ \hat{\nu}_{\mathbf{T}} &= -\hat{\mathbf{E}}_{\mathbf{T}}^{\mathbf{sum}} \\
\text{or}\ \ \hat{\not{\mathbf{E}}}_{\mathbf{T}} &= -\hat{\mathbf{E}}_{\mathbf{T}}^{\mathbf{sum}}
\end{aligned}
\tag{6.3}
$$

The vector $\hat{\nu}_{\mathbf{T}}$ is known as the "missing E_T vector" and is more commonly written as $\hat{\not{\mathbf{E}}}_{\mathbf{T}}$, with its magnitude being the event "missing E_T", written as $\not{E}_T$. In an ideal world $\hat{\not{\mathbf{E}}}_{\mathbf{T}}$ would perfectly describe the neutrino's P_T and direction in the transverse plane. Unfortunately, the presence of a neutrino in the event is not the only possible source of missing transverse energy. In practice, in order to get close to the true neutrino P_T, the missing E_T has to be corrected for the presence of weakly interacting muons and calorimeter energy mismeasurements. These corrections are described below.

A cut of 30 GeV is made on the corrected $\not{E}_T$ in order to select W$\rightarrow$ eν events with a reasonable purity. Figure 6.4(a) shows the missing E_T distribution of tight electron events before this missing E_T cut is made. One can clearly see the W$\rightarrow$ eν signal peaked at around 40 GeV, but also the presence of a large lower $\not{E}_T$ peak which is caused by other sources of tight electron candidate events that are investigated in detail in Chapter 7. In Figure 6.4(b) the increase in the purity of the candidate sample following the missing E_T cut is demonstrated; the tight electron E_T spectra only becomes convincingly W-like (i.e. peaked at $\sim$ 40 GeV) after the $\not{E}_T$ cut is made.

Missing E_T Corrections

In this study we perform two corrections to the "raw" missing E_T vector calculated using equations 6.2 and 6.3:

1. Correcting for the presence of muons.

2. Correcting for the mismeasurement of jet energy.

Muons, which in this analysis will largely result from the fragmentation and subsequent hadronic decays of heavy quark jets, interact only weakly with the calorimeter, and thus their calorimeter measured energy is just a fraction of the actual energy of the muon. Clearly we have to account for the energy of the muon that has been "missed", otherwise we shall mistakenly attribute this energy to the neutrino. The
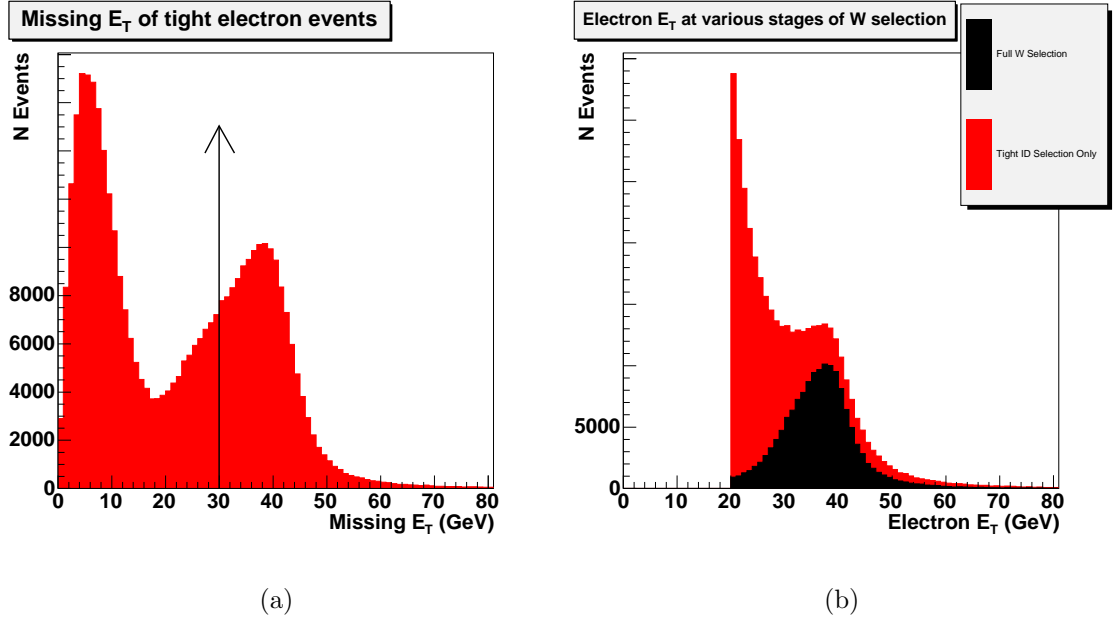
Figure 6.4: Distributions demonstrating the impact of the missing E_T cut on the $W \to e\upsilon$ candidate sample. The left-hand plot shows the missing E_T distribution of tight electron events before any $\not{E}_T$ cut is made (with an arrow indicating the position of the $\not{E}_T$ cut), and the right-hand plot shows the electron E_T distribution of tight electron events before and after the missing E_T cut.

first stage in such a correction is to identify events which contain high P_T muons. This proceeds in a similar way to the high P_T electron identification. Muon candidates are identified by searching for "stubs" in the muon chambers that can be matched to a COT track. A series of quality requirements is then made on the muon candidate, such as the track P_T and quality, how well the track matches to the stub and how well isolated the muon is from significant energy deposits in the calorimeter [85]. If a muon candidate passes all these requirements then it is used to correct the $\not{E}_T$. Naively one might think that we simply add the muon momentum as measured by the COT to the missing E_T vector in order to correct for it. However, doing so would result in double-counting, as some small fraction of the muons energy is deposited in the calorimeter. The equations below give the change in the x and y components of the missing E_T vector in order to account for the muon:

$$\Delta \not{E}_x^\mu = P_x^\mu \left(1 - \frac{E_{CAL}^\mu}{|P^\mu|}\right)$$

$$\Delta \not{E}_y^\mu = P_y^\mu \left(1 - \frac{E_{CAL}^\mu}{|P^\mu|}\right) \tag{6.4}$$

Where P_x^μ is the x component of the muon momentum and E_{CAL}^μ is the total energy deposited by the muon in the calorimeter. Since we only identify 3384 high P_T muons in 403009 events containing a high P_T electron, the muon corrections to the missing E_T are largely insignificant.

A far more important correction to the missing E_T is that for the mismeasurement of jet energies in the event. In Chapter 5 we explained the need for corrections to the raw calorimeter measured jet energies such that they better reflect the energy of the hadrons inside the jet. However, the tower sum E_T vector $\hat{\mathbf{E}}_{\mathbf{T}}^{\mathbf{sum}}$ is calculated using the raw calorimeter-level measurements, and thus this should be somehow modified by the jet corrections to better describe the true transverse energy imbalance of the final state particles in the event. The resulting correction to the x and y component of the missing E_T vector is given by:

$$\Delta\ \not{E}_x^{jet} = \sum_j E_{x,j}^{jet}(K_j - 1)$$

$$\Delta\ \not{E}_y^{jet} = \sum_j E_{y,j}^{jet}(K_j - 1) \tag{6.5}$$

Where the sum is over all jets selected using the criteria described in Section 6.2, $E_{x,j}^{jet}$ is the x component of the uncorrected calorimeter-level energy of the *jth* jet and K_j is the hadron-level correction factor applied to the *jth* jet. The final missing E_T corrected for both muons and jets is thus given by:

$$\not{E}_T = \sqrt{(\not{E}_x^{raw} - \Delta\ \not{E}_x^\mu)^2 + (\not{E}_y^{raw} - \Delta\ \not{E}_y^\mu)^2} \tag{6.6}$$

Where $\not{E}_x^{raw}$ and $\not{E}_y^{raw}$ are the x and y components of the raw uncorrected missing E_T .

One might reasonably ask that if we correct the $\not{E}_T$ for the mismeasurement of the energy that is clustered into jets, why do we not attempt to similarly correct for the energy that isn't clustered into jets, often referred to the "unclustered energy correction". The reason, as will later be explained in Section 6.2.2, is that as the energy decreases the response of the calorimeter becomes increasingly non-linear and hard to determine. The Run I W + Jets analyses [13] [60] [66] used P_T balancing in Z→ ee events to determine a constant factor that could be used to make an average correction for the response of the calorimeter to the unclustered energy. However, in events which contain a hard jet this unclustered energy is a small component of the total transverse energy flow of the event, and making such an average correction does very little to improve the missing E_T resolution.

Figure 6.5 shows the corrected missing E_T resolution for three different simulated W + Jets Monte Carlo samples. The corrected missing E_T is on average within 2%
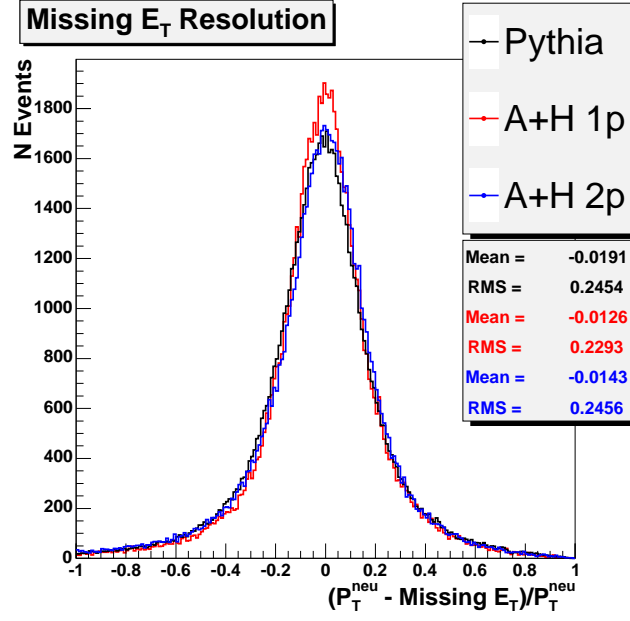
Figure 6.5: Missing E_T resolution from PYTHIA and ALPGEN+HERWIG Monte Carlo samples. The missing E_T is corrected for jets and muons, and is within 2% of the neutrino P_T on average.

of the true neutrino P_T . This average offset, along with the effect of the resolution, is accounted for by the acceptance factor, as is explained in Chapter 8.

6.1.4 Additional Selection Criteria

Once we have made the high P_T electron and neutrino selection described above we should already have a very reasonable W$\rightarrow$eυ candidate sample. We make the following additional requirements of the events in order to ultimately improve the accuracy of the cross-section measurement.

Z$\rightarrow$ ee Rejection

It is possible that our W$\rightarrow$eυ sample could contain a small fraction of Z$\rightarrow$ ee events. Such events also contain a high P_T electron in the final state, and although the vast majority should be removed by the $\not{E}_T$ cut, severe energy mismeasurements can mean that some Z$\rightarrow$ ee events with anomalously high $\not{E}_T$ can enter. To cut down this background we apply a Z$\rightarrow$ ee veto algorithm. The algorithm looks for a second "object" which with our tight electron forms an invariant mass in the [76,106] GeV Z mass window and additionally satisfies the following criteria:

- It is an opposite sign (track required) EM cluster which satisfies some loose electron cuts: $E_T > 10$ GeV, Had/Em < 0.12 and isolation < 0.15.
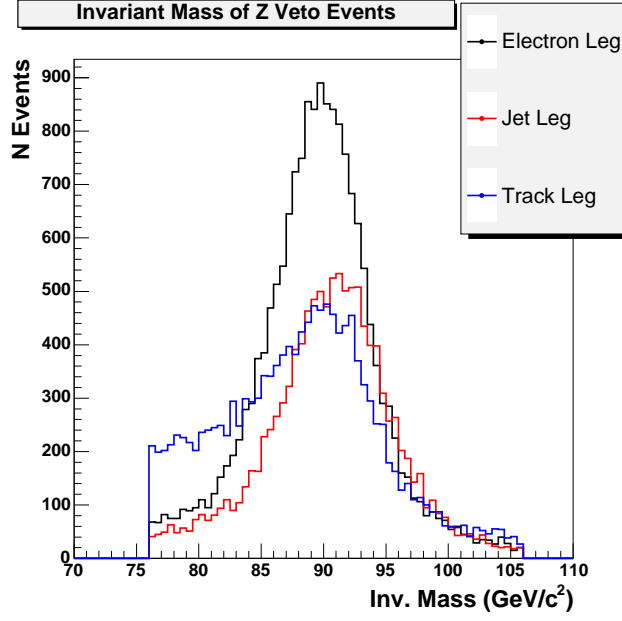
Figure 6.6: Invariant mass of the Z candidate electron, track or jet with the tight electron for events which are rejected by the Z rejection algorithm. The histograms are not normalised and so their relative areas correspond to the relative frequency with which each type of Z rejection event is identified.

- It is a jet with uncorrected $E_T > 15$ GeV, $|\eta| < 2.0$, EMF > 0.95 and number of associated tracks < 3.

- It is an opposite sign track with $P_T > 10$ GeV, absolute track isolation < 4.0 GeV and $|Z_0 - Z_{vtx}| < 10$ cm.

If such a second object is found the event is disregarded and doesn't enter the W$\rightarrow$eυ candidate sample. Figure 6.6 shows the invariant mass of the Z candidate electron, track or jet with the tight electron for events which are rejected by the Z$\rightarrow$ee veto algorithm.

It is of course possible that this algorithm removes a fraction of true W$\rightarrow$eυ events from the candidate sample, but from Monte Carlo studies this fraction is found to be less than 1% (see Chapter 8).

Conversion Veto

High energy photons that convert into electron-positron pairs are another source of high P_T electrons which can thus fake a W$\rightarrow$eυ signal. We perform an explicit search for such events in our W$\rightarrow$eυ sample by examining all other COT tracks in the event. If a track is found that is opposite charge to the electron track and which appears to originate from a common decay vertex in x-y-z then the event is flagged as a conversion and does not enter our W$\rightarrow$eυ sample. However, a W
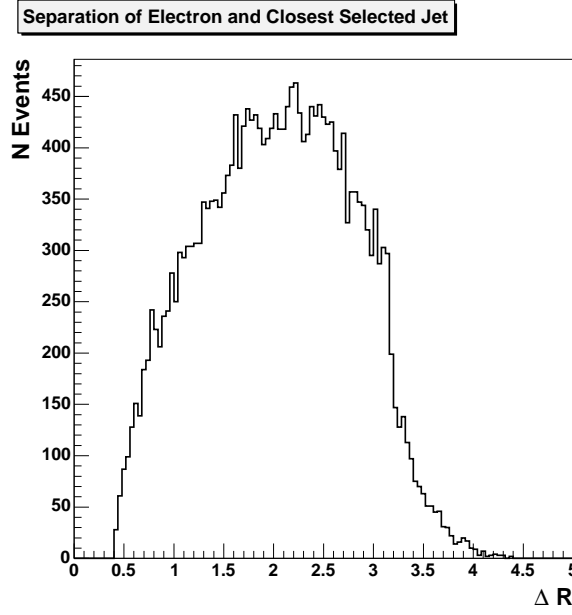
Figure 6.7: Separation in R between the tight electron and nearest selected jet before the $\Delta R_{ej} \geq 0.52$ cut is made.

electron can emit *bremmstrahlung* radiation which can convert, and the resulting conversion tracks can appear to form a perfectly good conversion candidate with the W electron. To try and avoid the incorrect vetoing of such "trident" events, we test that each conversion partner track to the W electron does not have a conversion partner track of its own.

Jet-Electron ΔR Cut

For all jets selected using the criteria described in Section 6.2 we require that the separation in R between the tight electron and the jet, ΔR_{ej} , is greater than 0.52, or 1.3 times the jet cone radius of 0.4.

$$\Delta R_{ej} = \sqrt{\Delta \phi_{ej}^2 + \Delta \eta_{ej}^2} \geq 0.52 \qquad (6.7)$$

Where the electron ϕ, η is determined using the electron track. Figure 6.7 shows the ΔR_{ej} distribution between the tight electron and the nearest selected jet before the $\Delta R_{ej} \geq 0.52$ cut is made. This cut removes the events in the $0.4 < \Delta R_{ej} < 0.52$ region; 176 out of 22535 ≥ 1 jet events or 0.8%.

Why do we make such a requirement? The isolation, Had/Em and E/P electron identification variables can all potentially be affected by the presence of a nearby jet. The higher the jet multiplicity of the event, the greater the chance that a jet will be close enough to an electron to affect these variables. Thus the efficiency of these cuts can be dependent to some extent on the jet multiplicity. As will be

explained in Chapter 8, we use W + Jet Monte Carlo samples to determine the efficiency of these electron identification cuts. A ΔR_{ej} cut can be used to minimise the dependence of these efficiencies on the jet multiplicity distribution of the Monte Carlo sample, thus hopefully minimising any Monte Carlo model dependence.

W Transverse Mass Cut

The W transverse mass of the W$\rightarrow$eυ event can be reconstructed from the tight electron and missing E_T vector as follows:

$$M_T = \sqrt{2E_T^e \ \not{E}_T(1 - \cos\Delta\phi_{e\nu})} \tag{6.8}$$

Where E_T^e is the tight electron transverse energy and $\Delta\phi_{e\nu}$ is the separation in ϕ of the electron and missing E_T transverse vectors. In Chapter 7 we demonstrate that in the interests of minimising the QCD background and the associated systematic it is desirable to impose a W transverse mass cut of 20 GeV on the W$\rightarrow$eυ event sample.

6.1.5   The Dataset

To test every event that has been logged in CDF Run II against the W$\rightarrow$eυ selection criteria would be an extremely time consuming and inefficient enterprise. Instead, we restrict our search to those events which passed the online high P_T central electron trigger described below. We also require that for each of these events the detector was in a satisfactory state to provide the measurements we need, via the implementation of the "good run" requirement (detailed below). The final dataset within which we search for W$\rightarrow$eυ events contains 6.4 million events, and corresponds to an integrated luminosity of 320pb^{-1}.

High P_T Electron Trigger

As was explained in Chapter 4 a three-level trigger system is used to select hard p$\overline{\text{p}}$ interactions from the generally less interesting softer collisions. There are many different requirements that can be made at each of the levels in order to select different kinds of event, and a particular combination of requirements is known as a trigger path. The `ELECTRON_CENTRAL_18` trigger path [93] is designed to select events containing high P_T electrons. It has the following requirements at each level:

- **Level 1:** At least one central calorimeter tower with $E_T > 8$ GeV, Had/EM ≤ 0.125 and a good quality COT track (found using the XFT[2]) with $P_T \geq$

---

[2]The eXtremely Fast Tracker (XFT) is a special track reconstruction algorithm that can reconstruct tracks in the time demanded by Levels 1 and 2.

8.34 GeV/c pointing to the tower.

- **Level 2:** Additionally a central EM cluster with electromagnetic component $E_T \geq 16$ GeV, Had/EM ≤ 0.125 and an XFT track of $P_T \geq 8.34$ GeV/c matched to the cluster.

- **Level 3:** A central EM cluster with $E_T > 18$ GeV, Had/EM ≤ 0.125, $L_{shr} < 0.4$ and a COT track reconstructed using the full algorithm with $P_T > 9$ GeV and matched using the CES to the cluster.

Good Run Requirements

At CDF each event can be associated to a particular "run", a time period of continuous data taking where the detector is known to be in a particular state. For each run that we use in this study we require that all the detector subsystems (with the exception of the silicon tracking system) were in good condition i.e. detector and electronics components are functioning as intended, in terms of gain, readout etc. The silicon is a detector subsystem that is often operated with zero gain due to potentially damaging beam conditions, and thus requiring "good" silicon results in a much restricted dataset. Although useful in terms of improving tracking and vertex resolution, silicon is not crucial to this analysis, and thus it is not a good run requirement.

6.1.6   The W$\rightarrow$eυ Sample

The results of applying the W$\rightarrow$eυ selection criteria to the dataset are shown in Table 6.1. Figure 6.8 shows the W transverse mass (reconstructed from the tight electron and missing E_T) for the W$\rightarrow$eυ event sample and a simulated PYTHIA W + Jets Monte Carlo sample. Good agreement with the Monte Carlo indicates we have a W sample of high purity, although one can see the clear presence of backgrounds in the tails of the data distribution.

6.2   Forming the W$\rightarrow$eυ $+$ n Jet E_T Distributions

Once we have formed our W$\rightarrow$eυ event sample the events are categorised according to the number and transverse energy of their jets, thus forming the candidate jet E_T distributions.

6.2.1   Jet Definition and Selection

The jet definition used in this analysis can be summarised as follows:

	N Events	Fraction
Electron Cuts		
Trigger	6383282	1.0000
Central ($|\eta_{\mathrm{D}}^{\mathrm{CEM}}| \leq 1.1$)	6383182	1.0000
Beam Constrained $|Z_0| < 60$cm	6156908	0.9645
Fiducial Region	5803827	0.9092
$E_T > 20$ GeV	4173640	0.6538
$P_T > 10$ GeV	3772762	0.5910
Had/EM $< 0.055 + 0.00045 \times E$	2878500	0.4509
E/P < 2.0 or $P_T > 50$ GeV	2008889	0.3147
-3.0cm $< Q_e \times \Delta X < 1.5$cm	1900208	0.2977
$|\Delta Z| < 3$cm	1823237	0.2856
Strip $\chi^2 < 10$	1448195	0.2269
$L_{\mathrm{shr}} < 0.2$	1314019	0.2059
COT Segments (min 5 hits 3 Ax, 2 St)	1310637	0.2053
Fractional Isolation < 0.1	733239	0.1149
Conversion removal	403545	0.0632
$|\Delta R_{\mathrm{ej}}| > 0.52$	403009	0.0631
W Selection Cuts		
Missing $E_T > 30$GeV	149313	0.0234
Veto Two Tight Electrons	149262	0.0234
$Z(\rightarrow e^+e^-)$ Rejection	147798	0.0232
W Transverse Mass > 20GeV	147008	0.0231
W Candidates	147008	0.0231

Table 6.1: Number of events after each W candidate selection cut is applied.
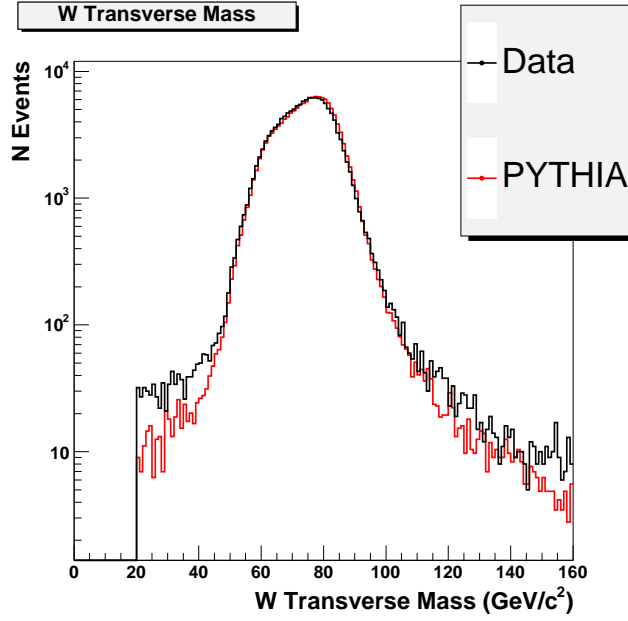
Figure 6.8: W Transverse Mass reconstructed using the tight electron and missing E_T. Data compared with simulated PYTHIA W→ eυ Monte Carlo.

1. Jets are clustered from the calorimeter using the JetClu algorithm described in Section 3.8.1.

2. The energies of these jets are then corrected to the hadron-level and also to account for multiple interactions.

3. Finally the following requirements are made of the corrected E_T and detector location of the jets:

    - $E_T \geq 15$ GeV

    - $|\eta| < 2.0$.

    At low energies the response of the calorimeter is very hard to parameterise, thus the corrected E_T cut ensures that the jets are of sufficient energy that the hadron-level corrections can be reliably applied. The η cut ensures that the jets remain in the central region of the calorimeter where the response is best understood.

Unfortunately in this analysis there is a further complication in the jet selection. The energy of the tight electron is not removed from the calorimeter before jet clustering is preformed. Thus it is always the case that one of the jets that results from clustering is the electron. Clearly we do not want this "electron-jet" to be counted as an analysis jet and included in the cross-section. Thus the absolute separation in R between the tight electron and each jet in the event is determined,

Jet Multiplicity	N Events
≥ 0	147008
≥ 1	22535
≥ 2	4402
≥ 3	885
≥ 4	208

Table 6.2: Inclusive jet multiplicity distribution.

and the highest E_T jet which has $\Delta R < 0.4$ is considered to be the "electron jet" and removed from consideration. Of the 147008 W candidate events, only 3 fail to find an "electron jet" in this way. Figure 6.9 shows the absolute separation in R between the "electron jet" and the tight electron for all W candidate events.
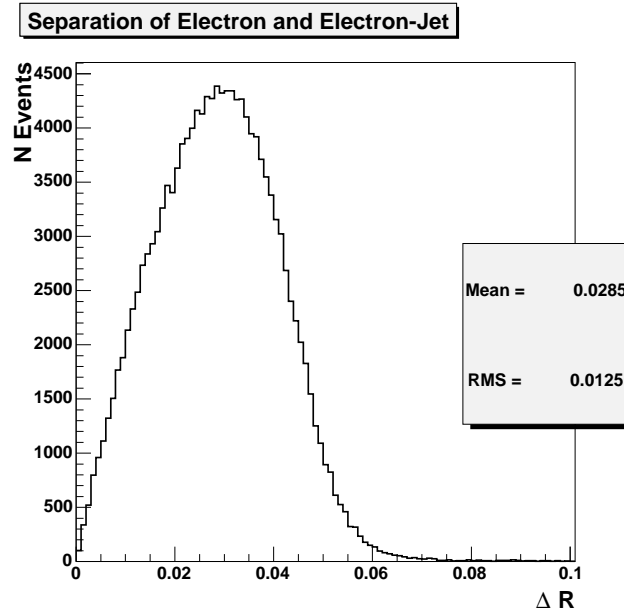
Figure 6.9: Absolute separation in R between the jet which is identified as resulting from the electron energy deposits and the tight electron for all W$\rightarrow$ ev candidate events.

Table 6.2 and Figure 6.10 describe the jet multiplicity distribution of the W$\rightarrow$ ev event sample i.e. the number of events with n jets. Note that although we have events in the five, six and seven jet bins, there are not enough events to make a reasonable differential cross-section measurement here. Figure 6.11 and 6.12 show the candidate W$\rightarrow$ ev + n jet E_T , ΔR_{jj} and M_{jj} distributions that are used to form the differential cross-sections.

In the sections that follow we will describe how the energy of the jets is corrected.
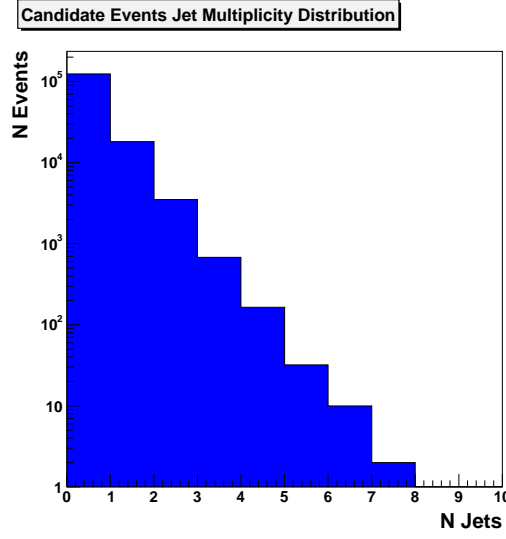
Figure 6.10: Exclusive jet multiplicity distribution of W candidate events.

6.2.2   Correcting Jet Energy to the Hadron-Level

In Chapter 5 we discussed the need to correct the energies of the observed jets back to the hadron-level in order to produce a CDF independent differential cross-section measurement. This correction is made in two distinct steps:

1. A "relative" correction factor is applied to make the response of the calorimeter uniform in η and equal to the response of the central region $0.2 < \eta < 0.6$.

2. An "absolute" correction factor is then applied to on average correct the calorimeter-level energy in the central region to the energy of the hadrons within the jet cone.

Relative Corrections

The relative correction factor is found by examining the P_T balance in data dijet events [45] i.e. events containing two JetClu reconstructed jets of a minimum energy, requiring that the energy of any additional jets is small. The method works on the principle that, due to the constraint of conserved transverse momentum in $p\bar{p}$ collisions, the P_T of the two jets should be equal. One jet, the "trigger" jet, is required to be in the central region $0.2 < \eta < 0.6$, whilst the other jet, the "probe", can be anywhere. If both jets are within the central region, the "probe" and "trigger" are assigned randomly. Figure 6.13 shows the ratio of the probe jet P_T to the trigger jet P_T ($\beta = P_T^{probe}/P_T^{trigger}$) as a function of the probe jet η. In other words, under the assumption that the P_T of the probe and trigger jets should balance, it shows the correction factor, $1/\beta$, required to make the response across the calorimeter equal to the response in the $0.2 < \eta < 0.6$ region.

(a) Lead Jet E$_T$ for ≥ 1 jet

(b) Second Jet E$_T$ for ≥ 2 jet

(c) Third Jet E$_T$ for ≥ 3 jet

(d) Fourth Jet E$_T$ for ≥ 4 jet
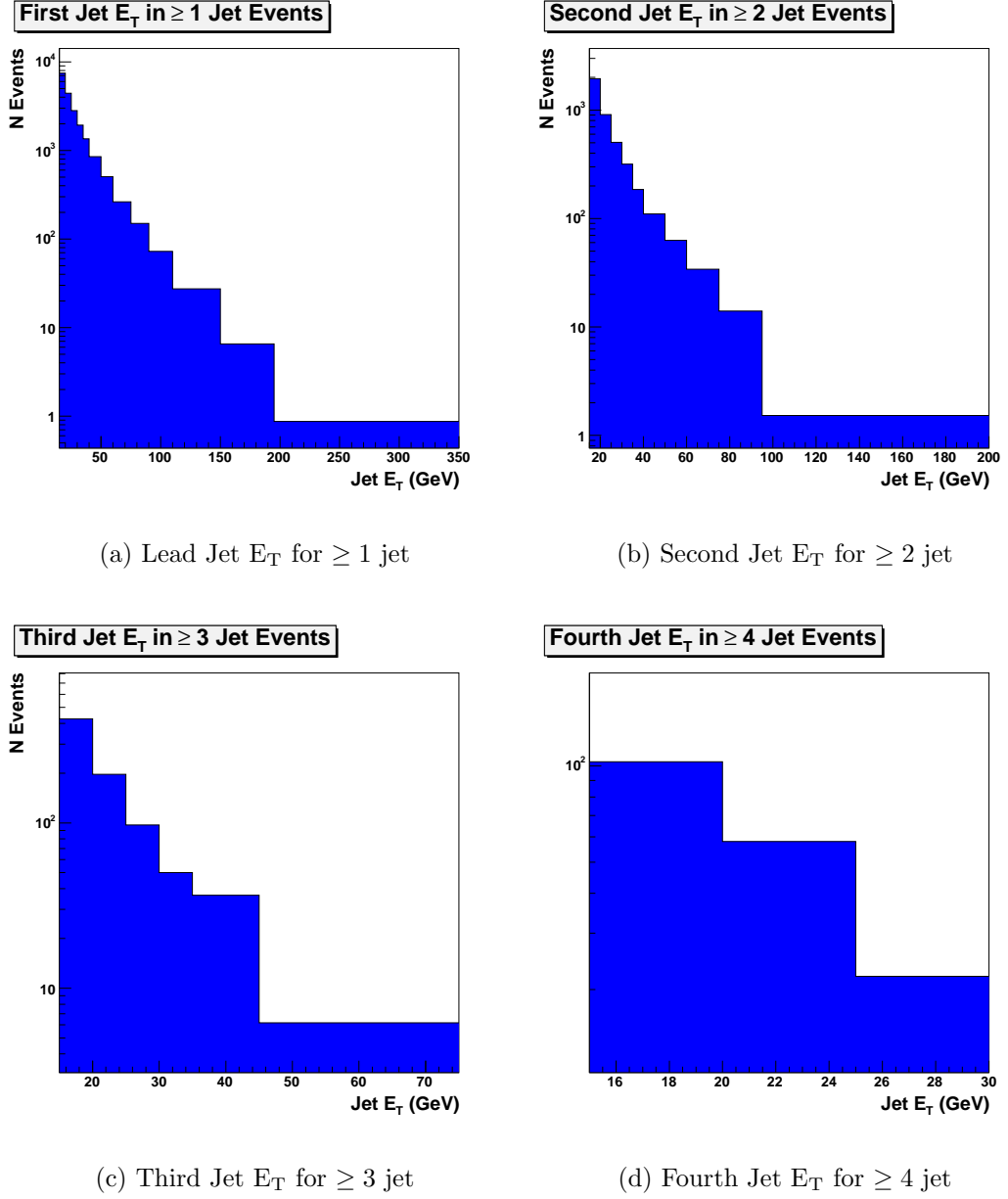
Figure 6.11: Candidate W$\rightarrow$ eυ + jet E$_T$ distributions.

The response of the calorimeter will also change with time, due mainly to the ageing of the phototubes. In Figure 6.13 the relative correction factor is shown for seven different time periods, and one can see that there are subtle changes in the P$_T$ balance over time.

It is important to note that the relative correction factors obtained using detector simulated PYTHIA or HERWIG dijet Monte Carlo samples are not quite the same as those derived from the data. This is largely because the detector simulation does not perfectly reproduce the η dependence of the calorimeter response.

(a) First-Second Jet ΔR_{jj}

(b) First-Second Jet M_{jj}
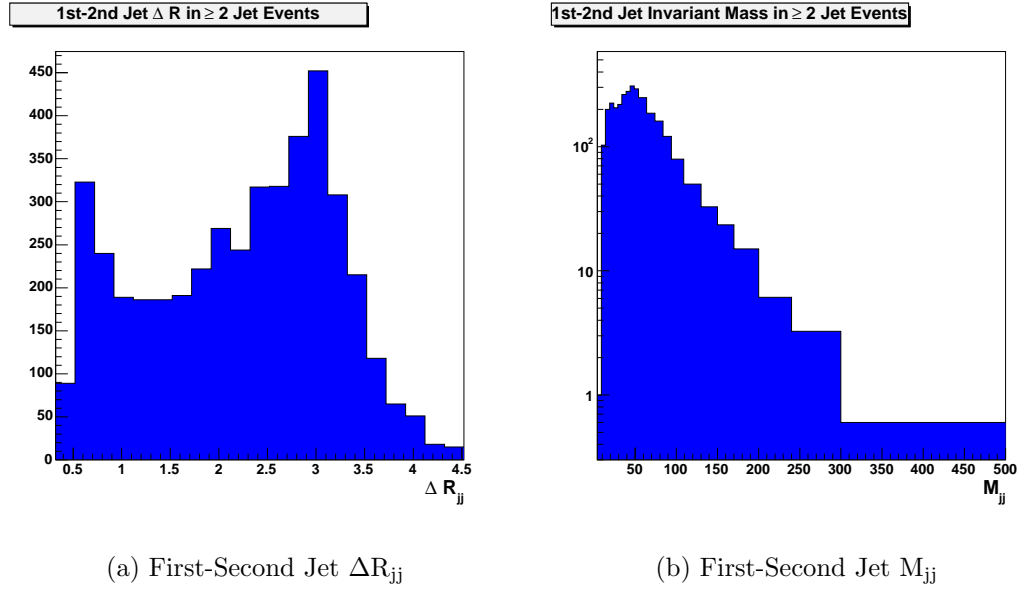
Figure 6.12: Candidate ΔR_{jj} and M_{jj} distributions.

Absolute Corrections

The response of the central CDF calorimeter to an individual hadron depends largely on the type of particle and its momentum. Electromagnetic particles, such as π^0 (which decay to photons), electrons and photons deposit the majority of their energy in the CEM, and here the observed energy is very close to the actual energy of the particle. However, other hadrons will deposit their energy largely in the CHA, and unfortunately the response of the CHA is non-linear as a function of the particle momentum, as demonstrated in Figure 6.14. In addition to this there are effects introduced by the magnetic field. Particles with momenta below 400 MeV will be confined inside the COT by the field and thus never reach the calorimeter. Particles with marginally higher energies may have their trajectory so significantly altered that they end up outside the calorimeter-level jet cone. The hadron-level energy of a particular jet is thus a complicated dependence on the make-up of the particles in the jet and the distribution of the momenta of these particles.

A considerable amount of effort has been made to tune the CDF detector simulation to accurately reproduce the response of the CDF calorimeter to the particles within jets, as well as to simulate the effects of the magnetic field and interactions in the material between the event vertex and the calorimeter. This enables us to use PYTHIA dijet Monte Carlo samples that have been detector simulated to derive a jet E_T dependent correction factor that takes us, on average, from the calorimeter-level to the hadron-level jet energy. The extraction of the correction factor is relatively simple. For each event in the Monte Carlo sample the jets are clustered using Jet-Clu at both the calorimeter-level and the hadron-level. Where possible each central
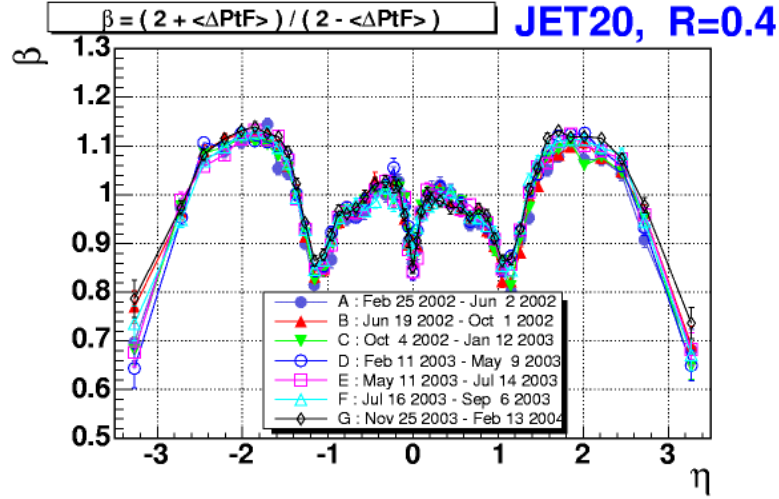
Figure 6.13: The correction factor required to achieve P_T balancing in dijet events as a function of detector η. Taken from [44].

$(0.2 < \eta < 0.6)$ jet at calorimeter-level is then associated, via a maximum separation in $\eta - \phi$ space, to a jet at the hadron-level, and the calorimeter-level energies of these matched jets plotted against the hadron-level energies. From such a distribution we can parameterize a so-called "Cal2Had transfer function", which provides a mapping between calorimeter-level and hadron-level jet energies. At CDF, the transfer function currently in use involves some ten parameters and is dependent on the calorimeter-level energy of the jet (η dependence being handled by the relative correction). It is shown in Figure 6.15 for a cone size of 0.4. The maximum, or average, value of the transfer function can be used to define an "absolute" correction factor as a function of calorimeter-level E_T, shown in Figure 6.16. This correction factor will take the calorimeter-level energy of the jet and correct it to the average equivalent hadron-level energy.

6.2.3 Correcting Jet Energy for Multiple Interactions

Figure 6.17 shows the probability of n interactions per bunch crossing for different Tevatron instantaneous luminosities. One can see that at an instantaneous luminosity of $1 \times 10^{32} \mathrm{cm}^{-2}\mathrm{s}^{-1}$ the most probable number of interactions is two. In our $W \rightarrow e\upsilon$ event sample, one of the interactions will always be the hard-scatter interaction which fired the high P_T electron trigger. It is extremely unlikely that any additional interactions will be hard-scatter events, far more probable that they will be soft $p\bar{p}$ interactions. However, a soft interaction will still deposit some energy in the calorimeter, and, as was discussed in Chapter 5, it is sensible to correct jets for this energy in order to produce a Tevatron independent measurement.

The energy deposition of the additional soft interactions should be well modelled
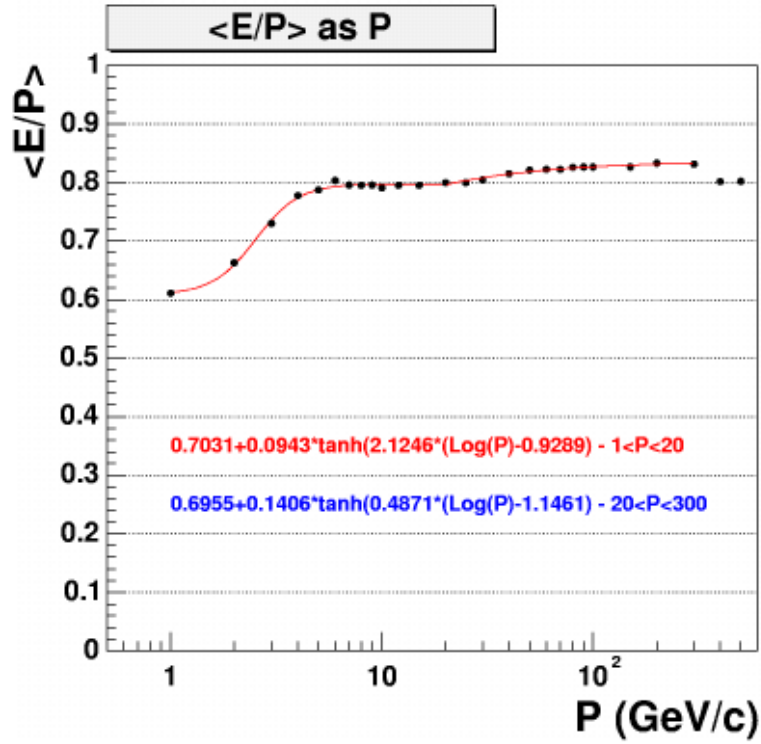
Figure 6.14: Results from the CDF detector simulation for the response of the central calorimeter to single hadrons as a function of particle momentum. Here response is defined as the ratio of the simulated calorimeter energy to the particle's momentum. Taken from [43].

by data events collected using the minimum bias trigger. This trigger essentially accepts all $p\bar{p}$ crossings with the exception of those where zero interactions occurred. Using such a sample we can measure the transverse energy in a cone of radius 0.4 randomly located in the central region ($0.2 < \eta < 0.6$) as a function of the number of reconstructed primary vertices in the event, shown in Figure 6.18. This can then be used to parameterise the amount of extra interaction energy we expect to overlap with a cone 0.4 jet by the number of additional primary vertices in the event [45], which is of course related to the instantaneous luminosity. Figure 6.19 shows the number of primary vertices reconstructed per event in our W candidate event sample. Just over one third of the candidate events (51975/147008) have more than one primary vertex reconstructed. Jets in these events will receive a correction of the order 0.5 GeV.

In this analysis we also correct for the rare cases where an additional $p\bar{p}$ interaction contributes an entire extra jet with E_T exceeding 15 GeV. This so-called "jet promotion" background correction will be described later in Section 7.5.
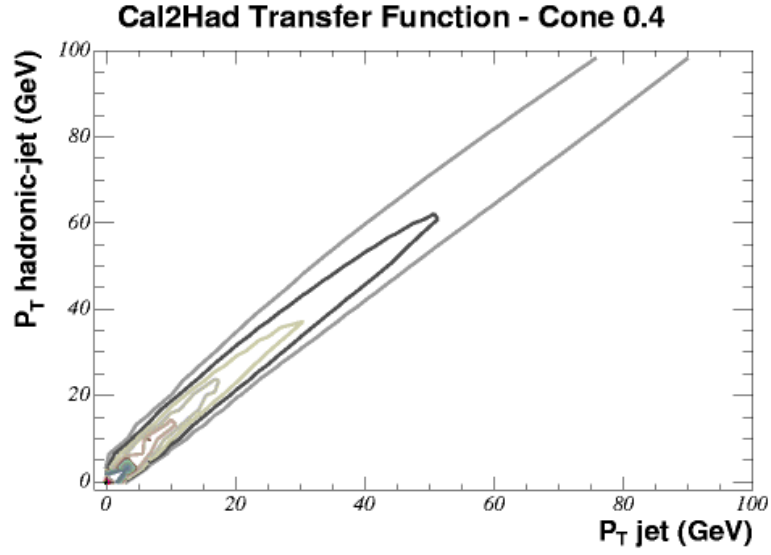
Figure 6.15: The Cal2Had transfer function for a cone size of 0.4, which provides a mapping between calorimeter-level (x-axis) and hadron-level (y-axis) jet transverse energies. Note that at each calorimeter-level energy there are a range of possible hadron-level energies, and vice versa. Taken from [43].

6.2.4   Jet Energy Correction Systematics

The relative, absolute and multiple interaction corrections described above all have associated systematic uncertainties. Figure 6.20 shows the contribution to the total jet energy scale systematic from each of the correction factors as a function of the jet E_T . Note that since we only correct our jet energy using the relative, absolute and multiple interaction corrections, our total systematic is these corrections only added in quadrature. This produces a systematic uncertainty on the jet energy scale from between 2 and 4% depending on the jet P_T , dominated by the absolute correction systematic.

The uncertainty in the data-derived relative correction arises from the definition of the dijet data sample used. As with any data sample it is formed using certain selection cuts, and a strong dependence of the correction factor on the exact cuts used could indicate, for example, a changing background content that would weaken the assumption that this is a pure dijet sample. The event selection cuts are varied and the resulting change in the relative corrections taken as a systematic. In addition, the relative corrections are applied to alternative data and PYTHIA Monte Carlo dijet samples, and any deviation from perfect jet balancing is taken as a systematic. The total systematic is a function of jet η and P_T . In the region $\eta < 2.0$ it is always less than 1.5%, but at $\eta > 2.0$ it increases sharply to as much as 7.5% at low P_T [45].

The uncertainty in the MC-derived absolute corrections arises from two sources:
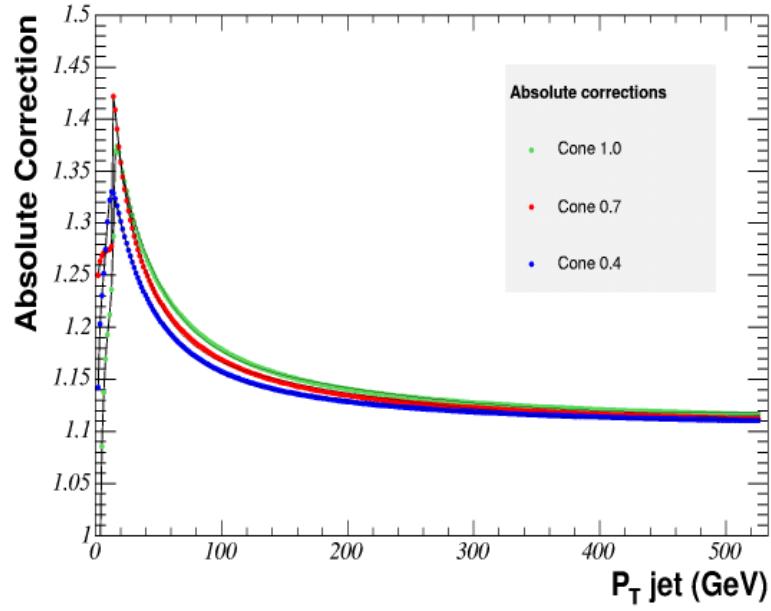
Figure 6.16: The absolute correction factor as a function of calorimeter-level jet transverse energy for 3 different cone sizes. Taken from [43].

1. The uncertainty in the simulated calorimeter response to single particles, both electromagnetic and hadronic.

2. The uncertainty on the particle make-up of the jets, both in terms of the type of particles and their momentum spectra, in the Monte Carlo i.e. the fragmentation model.

The first source is evaluated by comparing the E/P of single particles measured in data and simulated Monte Carlo for a range of momenta. For hadronic particles the uncertainty on the calorimeter energy response is 2.5-4%, depending on the particle P_T . For electromagnetic particles it is 1.7%, independent of P_T . However, since on average 70% of the jet energy is hadronic, these numbers translate to an uncertainty on the jet energy of 1.8-2.5% from hadronic and 0.5% from electromagenetic particles [45]. The uncertainty due to the fragmentation model is found by comparing the calorimeter response to jets in data and Monte Carlo simulation, via dijet balancing. We know that the response of the calorimeter to electromagnetic particles is unity, and have a formulation for the response to single hadronic particles which we know works on average. Thus any difference in the average calorimeter response to jets in data and HERWIG/PYTHIA Monte Carlo is due to differences in the fragmentation. This uncertainty is determined to be 1% independent of jet P_T [45]. Figure 6.21 shows the absolute correction systematic as a function of jet P_T .

The uncertainty on the multiple interaction correction is found by testing the assumption that the n vertex parameterisation does not depend on the topology of
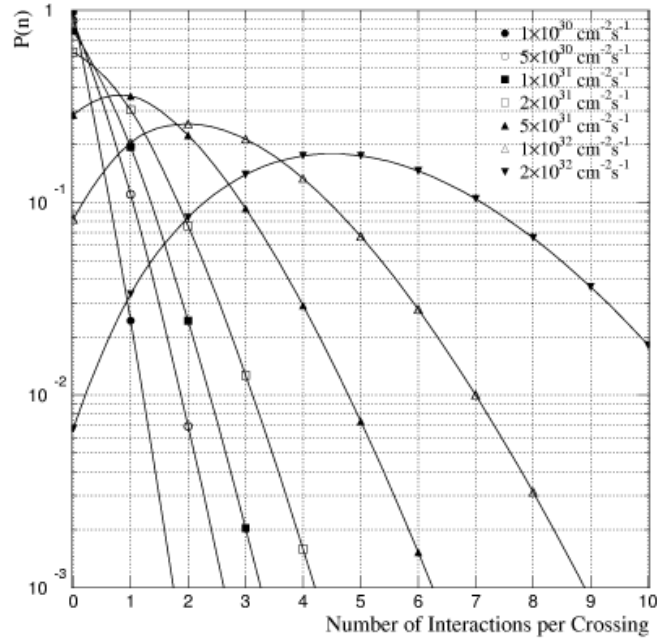
Figure 6.17: The probability of n interactions per bunch crossing for different Tevatron instantaneous luminosities. Taken from [40].

the hard interaction event or the instantaneous luminosity e.g. certain topologies or high occupancy events could lead to more fake vertices or a different vertex finding efficiency than is observed in minimum bias. The multiple interaction measurement (Figure 6.18) is repeated for a variety of different event samples and as a function of instantaneous luminosity, and a change in the slope of the n vertex parameterisation at the level of 15% is observed.
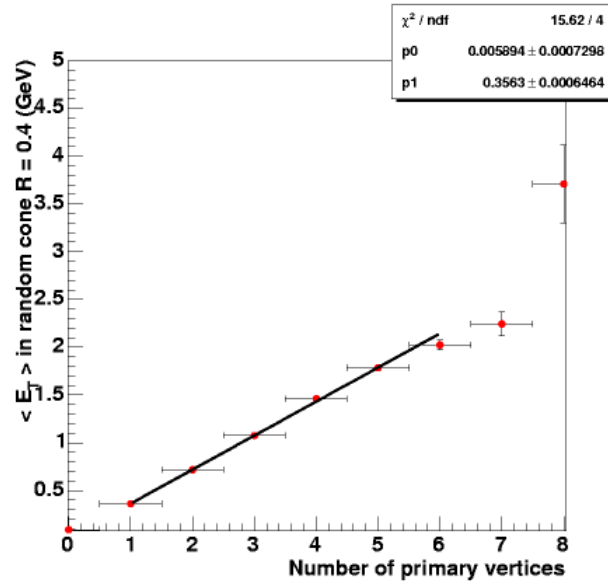
Figure 6.18: Average energy in a random, central 0.4 cone in minimum bias events. This gives the multiple interaction correction factor as a function of the number of extra vertices. Taken from [57].
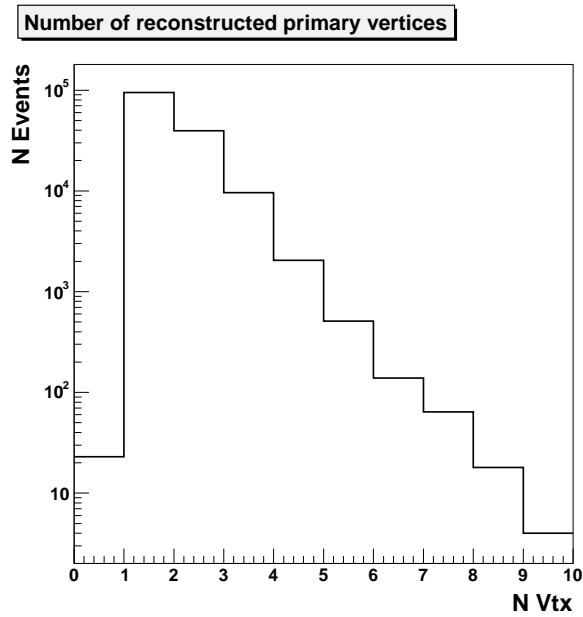
Figure 6.19: Number of primary vertices reconstructed per event in our W candidate event sample.
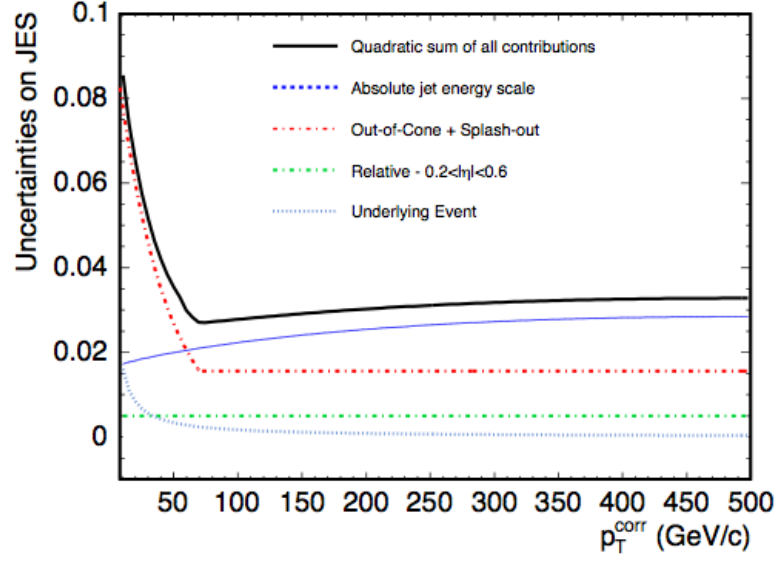
Figure 6.20: Jet energy scale systematic for each level of correction. Since we only correct our jet energy using the relative, absolute and multiple interaction corrections, our total systematic is these corrections only added in quadrature. Taken from [45].
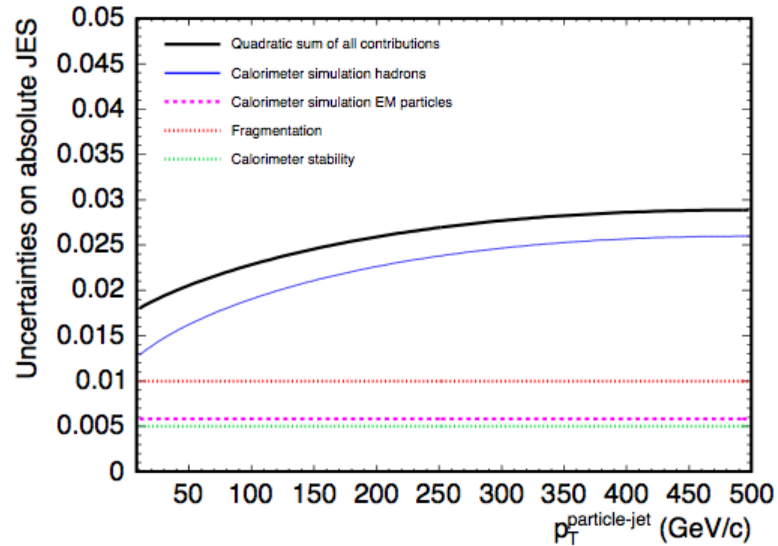
Figure 6.21: Absolute correction systematic as a function of jet P_T . The dominant contribution is from the uncertainty on the simulated calorimeter response to hadronic particles. Taken from [45].

Chapter 7

Backgrounds to W→ eυ + n Jet Events

The W→ eυ candidate event sample is formed by selecting events which contain a good quality high P_T electron along with large missing E_T . However, several other standard model processes can produce this signature, by "faking" one or both of the electron/neutrino and/or actually containing a real high P_T electron/neutrino in their final-state. In order to make an accurate measurement of the W + Jets cross-section an estimate of the number of background events in the candidate sample must be made, and in the case of a differential cross-section there must be some estimate of the background "contamination" in each bin. The backgrounds which are considered in this analysis are:

- **QCD multijet production**. It has been estimated that approximately 1 in every 10000 jets can "fake" an electron [83] i.e. pass all the electron ID selection criteria. If in the same event there is also a significant mismeasurement of jet energy, producing large missing E_T , then we have a QCD multijet event that has produced a W-like signature. Although the probability of this is extremely small, the cross-section for multijet production is much greater than W production, and thus this background can be significant.

- **t$\bar{\text{t}}$ production**. Figure 3.1 shows a diagram for the leptonic mode of t$\bar{\text{t}}$ production, where one of the W bosons decays leptonically, and one hadronically. This produces a final-state which is indistinguishable from our W→ eυ + Jets signal process, and which will tend to populate the 3 and 4 jet bins. Additionally the dileptonic mode, where both W bosons decay leptonically, may also contribute, although this has a smaller cross-section and will populate the lower jet multiplicity bins where the direct W production rate is higher.

- **Z→ ee production** Although the Z inclusive cross-section is 10.69 times smaller than the W and we attempt to explicitly remove Z events from our sample

with both a cut on missing E_T and the Z removal algorithm, it is still possible that some Z events make it into our sample.

- **W→ $\tau\upsilon$ production** Assuming lepton universality, a W produced in a $p\bar{p}$ collision is equally likely to decay to a τ as an electron. The τ has a significant branching fraction to electrons (18%), and thus W($\to \tau\upsilon$) production can produce a three neutrino one electron final state which is largely indistinguishable from the signal.

- **WW production** The inclusive diboson cross-section is around 2000 times smaller than for W production. However, approximately 7% of the time, one W boson will decay to an electron and neutrino and the other decay hadronically, contributing background events in the ≥ 2 jet bin where W production is suppressed by α_s^2 .

In addition we also consider what is referred to as the "jet promotion" background. Unlike the above sources this does not contribute to the total number of W→ eυ events in our sample, instead it promotes events between different jet multiplicity bins. It is described separately in Section 7.5.

7.1 Modelling the Backgrounds

7.1.1 Monte Carlo Background Models

The Z→ ee , W→ $\tau\upsilon$, WW and t$\bar{\text{t}}$ backgrounds can be safely modelled by simulated Monte Carlo samples of these processes. Table 7.1 details the Monte Carlo samples used. For the Z→ ee and W→ $\tau\upsilon$ backgrounds the 1 parton sample is used in the ≥ 1 jet bin, the 2 parton sample in the ≥ 2 jet bin, and so on. The WW and t$\bar{\text{t}}$ Monte Carlo samples include all possible decay modes of the W and top quark respectively. A simple estimate of the Monte Carlo modelled backgrounds could be made by passing the MC samples through the W selection criteria, and then use the Monte Carlo predicted cross-sections to normalise the number of events that pass in each jet E_T bin to the luminosity of our dataset. However, such reliance on leading order Monte Carlo cross-section predictions is not desirable.

7.1.2 Antielectron QCD Background Model

The rate and kinematic properties of the QCD background depend on the very complex details of a jet faking an electron signature, and is not something that one should rely on the detector simulation to accurately describe. Instead we select "antielectron" events from our high P_T electron dataset that we believe are overwhelmingly

MC Sample	Q^2 Scale	Parton Generation Cut	N_{gen}
W$\to$ eυ + Jets			
A+H 1 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	193K
A+H 2 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	180K
A+H 3 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	251K
A+H 4 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	219K
Z$\to$ ee + Jets			
A+H 1 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	449K
A+H 2 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	393K
A+H 3 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	299K
A+H 4 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	296K
W$\to$ $\tau\upsilon$ + Jets			
A+H 1 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	277K
A+H 2 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	272K
A+H 3 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	245K
A+H 4 parton	$M_W^2 + sumP_T^2(p)$	minP$_T$ 8 GeV, ΔR 0.2	243K
$t\bar{t}$			
PYTHIA	M_t^2	None	1.2M
WW			
PYTHIA	$(2M_W)^2$	None	420K

Table 7.1: Description of the Monte Carlo samples used to model backgrounds and signal processes. A+H = ALPGEN+HERWIG ELO sample.

QCD and representative of multijet events that can fake the W signature. Much like a Monte Carlo sample, we can then use this "antielectron" event sample to model the QCD background in any kinematic distribution of our choosing.

One can divide the electron selection criteria (Table 6.1) into two categories, as shown in Table 7.2. One group contains variables designed to discriminate electrons from fakes, which we will refer to as "ID variables". The other is meant to enrich the candidate sample in high E$_T$ electrons to ensure the presence of a W production event, we will refer to this group as "kinematic variables". The antielectron sample is selected from the analysis dataset by requiring an electron cluster which passes all the kinematic variables but fails at least two identification variables, using exactly the same cuts as in the signal analysis (Table 6.1). Any event which contains a tight electron as well as an antielectron is discarded and does not enter our antielectron sample. Table 7.3 shows the statistics of the antielectron sample in each jet multi-

kinematic	E_T	P_T	central	fiducial	Z_0	E/P	Isolation
ID	Had/Em	χ^2	L_{shr}	CES ΔX	CES ΔZ		

Table 7.2: Table showing the division of electron selection variables into "kinematic" and "identification" categories. Our antielectron sample is formed by requiring all kinematic variables to pass, but at least two ID variables to fail.

Jet Bin	N_{anti}	$N_{\mathrm{anti}}|_{\not{E}_T \geq 30\mathrm{GeV}}$
≥ 0 jets	23493	1068
≥ 1 jets	12710	449
≥ 2 jets	2831	149
≥ 3 jets	618	38
≥ 4 jets	131	16

Table 7.3: Table showing the statistics of the antielectron sample.

plicity bin, and Table 7.4 shows the sample composition in terms of the fraction of events in the sample which fails each of the ID cuts.

The important assumption made in forming the sample in this way is that the kinematic properties of multijet events that fake electrons are not dependent on the ID variables. In other words, by reversing only the ID cuts, we do not introduce any kinematic bias into our QCD sample. This is crucial because, as we shall see in Section 7.2, we rely on the QCD kinematics being well modelled by our antielectron sample in order to extract the differential background. By requiring at least two ID cuts to fail we hopefully select an antielectron sample which is QCD dominated, whilst maintaining reasonable statistics in each jet multiplicity bin.

7.2 The Background Extraction Method

Figure 6.4(a) in Chapter 6 shows the missing E_T distribution of the W$\rightarrow$ eυ candidate event sample in the ≥ 0 jet bin, with the 30 GeV missing E_T cut is removed. One can clearly see the W$\rightarrow$ eυ signal, peaking at around 40 GeV, but also a large peak at low missing E_T, produced by QCD events that fake the electron signature. Less obvious contributions will be made by the W$\rightarrow \tau\upsilon$, Z$\rightarrow$ ee , WW and t$\bar{\mathrm{t}}$ processes. To within the statistics available the antielectron sample describes the $\not{E}_T$ shape of the QCD background, and the Monte Carlo samples can be passed through the W selection criteria (minus the $\not{E}_T$ cut) to give us $\not{E}_T$ shapes for the other background sources and signal. We can thus make a binned maximum likelihood fit of the background and signal shapes to the data in each inclusive jet multiplicity sample, and the area above 30 GeV of each background curve will give us the normalisation with

ID Cut	Pass %	Fail %
Strip χ^2	20	80
CES Δx	73	27
CES Δz	72	28
L_{shr}	61	39
Had/EM	35	65

Table 7.4: Table showing the composition of the antielectron sample in terms of the fraction of events in the sample which fails each of the ID cuts. These numbers are for the entire antielectron sample, but the composition does not change significantly with increasing jet multiplicity.

respect to the candidates.

Note that the method described here allows one to extract the background composition of each W + n jet multiplicity sample. In principle to extract the differential background one could use the same method in each bin of the jet E_T, ΔR_{jj} or M_{jj} distribution. However, this would be very time consuming and statistics limited in some bins, so instead the inclusive background result is used to normalise the kinematic shapes of the background models, as described in Section 7.3. Cross-checks have been performed that demonstrate the equivalence of these two approaches.

7.2.1   Fitting to the Missing E_T Distribution

The number of observed candidate events in the ith bin of the W + $\geq n$ jet data $E\!\!\!/_T$ distribution is given by:

$$N_n^i(\text{data}) = N_n^i(\text{W} \to \text{e}\upsilon) + N_n^i(\text{QCD}) + N_n^i(\text{W} \to \tau\upsilon) + N_n^i(\text{Z} \to \text{ee}) + N_n^i(\text{t}\bar{\text{t}}) + N_n^i(\text{WW})$$

(7.1)

Where $N_n^i(\text{W} \to \text{e}\upsilon)$ is the number of signal events in the ith bin, $N_n^i(\text{QCD})$ the number of QCD events and so on. Naively one might think that there are six normalisation factors to be found in the fit, one for each contribution in Equation 7.1. However, in fact the fit has only two free parameters, which we shall label K_Q and K_W. Regardless of the fit to the data, the relative normalisations of the W$\to$ eυ, Z$\to$ ee and W$\to$ $\tau\upsilon$ contributions are determined by the well-known standard model relationships between their cross-sections. These distributions can thus be combined to form an "electroweak template", the normalisation of which is the first free parameter of the fit, K_W. The following relationships are used to scale the Z$\to$ ee and W$\to$ $\tau\upsilon$ missing E_T distributions relative to W$\to$ eυ to form the electroweak template:

$$N_n^i(W \rightarrow \tau\upsilon)' \;\; = \;\; N_n^i(W \rightarrow \tau\upsilon) \cdot \frac{G_n(W \rightarrow e\upsilon)}{G_n(W \rightarrow \tau\upsilon)} \tag{7.2}$$

$$N_n^i(Z \rightarrow ee)' \;\; = \;\; N_n^i(Z \rightarrow ee) \cdot \frac{G_n(W \rightarrow e\upsilon)}{G_n(Z \rightarrow ee)} \cdot \frac{1}{R_{WZ}} \tag{7.3}$$

Where $N_n^i(W \rightarrow \tau\upsilon)$ is the number of unscaled W$\rightarrow \tau\upsilon$ MC events in the ith bin after W $+ n$ jet selection, $G_n(W \rightarrow e\upsilon)$ is the number of generated signal MC events with n jets, $G_n(W \rightarrow \tau\upsilon)$ is the number of generated W$\rightarrow \tau\upsilon$ MC events with n jets and R_{WZ} is the ratio of W to Z inclusive cross-sections. Note that we assume lepton universality and that R_{WZ} is a constant with respect to jet multiplicity. The second free parameter of the fit is the normalisation of the antielectron $\not{E}_\mathrm{T}$ distribution, or "QCD template", K_Q . The normalisation of the t$\bar{\mathrm{t}}$ and WW contributions are fixed by the cross-section of these processes and the luminosity of the dataset, and are thus not free parameters of the fit. We use the most recent CDF measurement of the t$\bar{\mathrm{t}}$ cross-section, 7.1 ± 1.0pb [14], and the NLO theoretical prediction for the WW production cross-section at $\sqrt{\mathrm{s}}$ =1.96 TeV, 12.4 ± 0.8pb [48]. One advantage of this approach is that, with the exception of the t$\bar{\mathrm{t}}$ and WW contributions, we do not directly use any previous cross-section measurements to determine the background normalisations, only σ_W/σ_Z in which many systematic errors cancel. Equation 7.1 can thus be rewritten as:

$$N_n^i(data) \;\; = \;\; K_{W_n} \cdot [N_n^i(W \rightarrow e\upsilon) + A_Z N_n^i(Z \rightarrow ee) + A_\tau N_n^i(W \rightarrow \tau\upsilon)]$$
$$+ K_{Q_n} \cdot N_n^i(QCD) + A_{\mathrm{t\bar{t}}} N_n^i(t\bar{t}) + A_{\mathrm{WW}} N_n^i(WW) \tag{7.4}$$

Where A_Z, A_τ, $A_{\mathrm{t\bar{t}}}$ and A_{WW} are the fixed normalisation factors discussed above and $N_n^i([\mathrm{process}])$ is the number of events observed in the ith bin of the Monte Carlo after W $+ n$ jet selection. Note that each jet multiplicity bin will likely have different K_W and K_Q factors determined from the fit. The Monte Carlo samples used in each multiplicity bin were described in Section 7.1.1.

To perform the fit we use the `ROOT` TFractionFitter class [77]. This takes as input the electroweak and QCD templates along with the data $\not{E}_\mathrm{T}$ distribution and fixed normalisation top and WW distributions. Assuming Poisson counting statistics, it then makes a binned maximum likelihood fit [38] in which the only free parameters are the fractional contributions of the QCD and electroweak templates to the data, F_Q and F_W respectively, which can easily be related to the factors K_Q and K_W . However, instead of the "theoretical" prediction in each bin being fixed, Poisson statistics are also assumed in the template predictions resulting in additional contributions to the likelihood in each bin. One can think of this as the fitter having additional freedom to vary the shape of the electroweak and QCD

templates within the statistics of each bin. In this way the statistical uncertainties in both the MC and antielectron distributions are accounted for in the fit and more importantly the error on the fit parameters.

7.2.2   Electroweak Corrections to the Antielectron QCD Template

We have thus far assumed that the only contribution to the antielectron sample is from QCD multijet events. However, it is possible that the antielectron sample includes not only jets faking electrons but also real electron events from W$\rightarrow$eυ , Z$\rightarrow$ee , W$\rightarrow$$\tau\upsilon$, WW and top production which for some reason fail two or more of the electron identification cuts. In order to account for this we run our electroweak and top Monte Carlo samples through the antielectron sample selection and obtain an antielectron $\not{E}_T$ distribution for each. However, before these can be used to produce a corrected QCD template they must be normalised *vis a vis* the data antielectron distribution. The top and WW normalisation is fixed by the assumed cross-sections and data luminosity as already described. For the remaining contributions we can as before linearly combine to form an "electroweak template" using the known cross-section relationships. The normalisation factor for this template distribution will simply be the same K_W factor as is used to normalise the signal distributions. Here the relative contribution of the electroweak template will be much less significant than in the signal samples because the large inefficiency of the antielectron selection means the initial normalisation of the histograms is much smaller. The QCD background template is then formed by subtracting the correctly normalised electroweak and top antielectron distributions from the data antielectron distribution. In practice this requires an iterative procedure because we only know K_W once the fit has been performed. In the first iteration the fit is performed with K_W set to zero *i.e.* no electroweak corrections to the QCD template applied. The next iteration uses the K_W from the first, thus applying an electroweak correction, and so on until the K_W returned by the fit is smaller than the previous iteration result by less than $\sigma_{K_W}/10$.

The results of these corrections are shown in Figure 7.1 for each jet multiplicity bin. The black curve shows the original antielectron sample distribution, and the red curve the distribution after subtracting the correctly normalised electroweak contributions shown. One can see that the most significant correction occurs in the ≥ 0 jet bin above 25 GeV. The electroweak distributions peak around this value and their impact is enhanced in the ≥ 0 jet bin because of the dijet nature of QCD production: high E_T QCD fake electrons are more likely to be produced in association with another high E_T jet. The electroweak corrections in the other jet multiplicity bins are insignificant.

Jet Bin	K_W	σ_{K_W}/K_W	K_Q	σ_{K_Q}/K_Q	χ^2/DOF	χ^2 Probability
≥ 0	2.84439	0.44 %	2.88044	0.87 %	4.97684	0.00000
≥ 1	1.70081	0.97 %	2.74368	1.15 %	2.22816	0.00000
≥ 2	0.55455	1.83 %	2.73934	2.44 %	1.56752	0.00773
≥ 3	0.11400	4.07 %	2.45626	5.35 %	1.25971	0.20827
≥ 4	0.03871	12.69 %	2.58447	12.32 %	0.61973	0.87967

Table 7.5: Table showing the K_W and K_Q factors that result from the background maximum likelihood fit procedure in each jet multiplicity bin. Also shown is the χ^2 results of the fit.

7.2.3 Fit Results

Table 7.5 shows the K_W and K_Q factors found using the maximum likelihood fit procedure in each jet multiplicity bin, along with the χ^2 results of the fit. Figures 7.2 and 7.3 show the "fit normalised" background and signal missing E_T distributions and their combination alongside the distribution of the data candidates. One can see in all jet multiplicity samples that the combined signal plus background distributions describe the $\not{E}_T$ shape seen in the data very well. As the jet multiplicity increases the errors on the fit parameters also increase. This reflects the reduction in statistics of the data and QCD template distributions as we increase jet multiplicity, as can be seen in Figure 7.2. The error bars grow, giving the fitter increased flexibility to make a "good" fit at the expense of uncertainty in this fit. The very high statistics of the ≥ 0 jet and ≥ 1 jet samples reveal small shape differences between the template distributions and the data, resulting in the χ^2 probability of the fit reducing to zero. However, we are simply using the fitter as a tool to calculate the best normalisation factors for our background estimation and a reasonable statistical error on these factors. We do not pretend that we can reproduce the data across the full $\not{E}_T$ range in such high statistics samples. The uncertainty on the shape of our QCD and electroweak templates is a systematic of the method which we attempt to address in Section 7.4.

One can further test the shape and normalisation of these backgrounds by comparing to W candidates in other W-related kinematic variables. Figures 7.4 and 7.5 show the background plus signal combination compared to data in the tight electron E_T and reconstructed W transverse mass distributions. Here the normalisations of the backgrounds and signal as determined by the missing E_T fitting procedure are used, and the missing E_T cut is again not made such that we can better examine the modelling of the QCD background. One can see that the agreement between the combined (red) distribution and the observed data (black) distribution is reasonable in both variables. That is a powerful cross-check of our fitting procedure and background modelling.

One can see from Figure 7.5 that a W transverse mass cut of $WM_T > 20$ GeV has been made, as was described in Section 6.1.4. This cut was motivated by the distribution shown in Figure 7.6(b). This shows a comparison of the background plus signal WM_T distribution with the data in the ≥ 0 jet bin, where the normalisations have been determined by the missing E_T fit procedure without a WM_T cut applied. The results of this fit are shown in Figure 7.6(a). Although the agreement in the missing E_T distribution is still reasonable, Figure 7.6(b) shows that the description of the QCD background in the region $WM_T < 20$ GeV is problematic. Since we do not gain much signal acceptance in this region we decided to exclude it in the interests of a better controlled QCD background estimation.

7.2.4   Inclusive Background Results

The missing E_T fit procedure described above is performed separately for each inclusive jet multiplicity bin, and each inclusive background estimated by integrating the appropriate curve above $\not{E}_T > 30$ GeV. The inclusive background results are given in Table A.1.

QCD is the dominant inclusive background component in the ≥ 1 jet and ≥ 2 jet bins, but is exceeded in the ≥ 3 jet and ≥ 4 jet bins by the $t\bar{t}$ background. This is understandable since the leptonic mode of the $t\bar{t}$ process (see Figure 3.1) will tend to result in three or four final state jets, and thus the $t\bar{t}$ acceptance is concentrated in these bins. Conversely the inclusive $W \rightarrow \tau\upsilon$ background remains constant at around 1.7%, because, as expected, the acceptance relative to the signal of $W \rightarrow \tau\upsilon$ events remains the same as we add more jets. The $Z \rightarrow ee$ background is insignificant, which is unsurprising since we implement a $Z \rightarrow ee$ veto algorithm in our W selection (see Section 6.1.4).

7.3   Differential Background Results

In order to make a differential cross-section measurement we require knowledge of the backgrounds in each bin of the jet E_T , ΔR_{jj} and M_{jj} candidate distributions. The inclusive background results in each jet multiplicity bin are used to normalise the jet E_T , ΔR_{jj} and M_{jj} shapes of each background sample relative to the candidates. In Figure 7.7 is shown the total background in each bin of the jet E_T , ΔR_{jj} and M_{jj} distributions, as well as the individual contribution from each background source, with the normalisations determined in this way. Figure 7.8 shows the background in each bin as a fraction of the number of candidate events in that bin. Tables A.5 - A.10 summarise the differential background results in each of the jet E_T , ΔR_{jj} and M_{jj} distributions.

In the first jet and second jet E_T distributions QCD is the dominant background component at low jet E_T , but at higher jet E_T the $t\bar{t}$ background becomes equally important. The jets in $t\bar{t}$ events tend to be hard, resulting either from the decay of a W boson or top quark. One can see in Figures 7.7(a) and 7.7(b) that the $t\bar{t}$ first and second jet E_T distributions are peaked at around 60 GeV, whilst the QCD background is exponentially decaying. The $t\bar{t}$ background has a significant impact on the shape and normalisation of the background in the tails of the first and second jet E_T distributions. In the third and fourth jet E_T distributions, the $t\bar{t}$ background dominates both the shape and normalisation of the total background. It is largely the impact of the $t\bar{t}$ background which results in the background fraction increasing with higher jet E_T across all jet multiplicity bins (with the exception of the fourth jet), from 10-20% at low E_T up to 60-80% at high E_T .

The structure of the Top and WW backgrounds is evident in the M_{jj} differential background distribution. The WW background is peaked at 80 GeV, as one might expect for a pair of jets resulting from W decay, and the $t\bar{t}$ background peaks at around 90 GeV, or $M_t/2$. Both make important contributions to the shape of the total background, which is otherwise QCD dominated. The ΔR_{jj} differential background is a similar picture, the flat QCD background dominating, with the WW and Top backgrounds peaked at a back-to-back dijet configuration ($\Delta R_{jj} = \pi$).

In all the distributions shown in Figure 7.8 is reported the promotion background fraction. Discussion of this background component is postponed until Section 7.5.

7.4 Background Systematics

The following sources of systematic error are considered in the estimation of the background as a function of jet E_T , ΔR_{jj} and M_{jj} :

- The definition of the antielectron sample which we use to model the QCD background. How good a model of the $\not{E}_T$ and jet E_T , ΔR_{jj} and M_{jj} distributions of QCD fake electron events is this?

- The models of the Z$\rightarrow$ ee , W$\rightarrow \tau \upsilon$, $t\bar{t}$ backgrounds and W$\rightarrow$ eυ signal $\not{E}_T$ distribution. How dependent are the background predictions on the details of these Monte Carlo samples?

- Disagreement between the missing E_T scale and resolution in data and detector simulated Monte Carlo.

- The error on the $t\bar{t}$ and WW cross-sections that are used to normalise these background contributions.

Jet Bin	N_{anti}	$N_{\text{anti}}|_{\not{E}_T \geq 30\text{GeV}}$
≥ 0 jets	16509	648
≥ 1 jets	8406	294
≥ 2 jets	1721	93
≥ 3 jets	386	24
≥ 4 jets	85	11

Table 7.6: Table showing the statistics of the redefined antielectron sample where the ID cuts are varied by 20% away from the signal region.

- The error on the $\sigma_{\text{W}}/\sigma_{\text{Z}}$ cross-section ratio that is used to normalise the $Z \to$ ee $\not{E}_T$ distribution relative to that of the signal.

- The error on the least likelihood fit that results from the statistical error on the antielectron, Monte Carlo and data $\not{E}_T$ distributions.

- The statistical error on the antielectron and MC background sample jet E_T shapes.

The error resulting from the uncertainty in the antielectron QCD model is found by varying the antielectron sample definition. The sample is defined by selecting events which fail two or more of the ID cuts listed in Table 7.2, where the exact cuts used are those which define the tight electron in our W candidate selection, listed in Table 6.1. If in our antielectron definition we vary these cuts, we will vary the composition of the antielectron sample and test the assumption that the ID variables are uncorrelated with the kinematic properties of the sample. The ID cuts are each varied by 20% "away" from the signal region and a new antielectron sample selected from the analysis dataset by requiring that at least two of the redefined ID variables fail. Tables 7.6 and 7.7 show the statistics and composition of this redefined antielectron sample respectively. Note that the statistics are reduced because the cuts required to fail are effectively tighter in this definition. The full missing E_T procedure is then repeated with this new antielectron sample to obtain a new bin-by-bin background estimate for each background source. Figure 7.9 compares the differential QCD background obtained using the central antielectron definition and this new definition. One can see that the variation is small and within the statistical error on the background. This variation is propagated as a systematic on the QCD background estimation. The change in the other backgrounds is negligible.

In Chapter 5 we discussed the importance of making a differential cross-section measurement which is as independent as possible from any theoretical models used. In the above background extraction method we make use of ALPGEN+HERWIG Monte Carlo samples to model the $W \to \tau\upsilon$ and $Z \to$ ee backgrounds and, most significantly, the $W \to$ eυ signal $\not{E}_T$ distributions, using the n parton sample in the

ID Cut	Pass %	Fail %
Strip χ^2	16	84
CES Δx	69	31
CES Δz	68	32
L_{shr}	64	36
Had/EM	31	69

Table 7.7: Table showing the composition of the redefined antielectron sample where the ID cuts are varied by 20% away from the signal region.

$\geq n$ jet bin. Thus, through our background estimation, we are potentially building the "Enhanced Leading Order" approach into our differential cross-section measurement. In order to test the dependence of our background estimation on the ELO model, we repeated the full background calculation using "skewed" Monte Carlo samples, that is, the $n + 1$ parton sample in the $\geq n$ jet bin. Table A.2 gives the inclusive background results when the Monte Carlo samples used are skewed in this way, and we observe some variation in the Z$\rightarrow$ ee and W$\rightarrow \tau\upsilon$ backgrounds.

In Section 6.2.4 we discussed the uncertainty in the absolute jet energy corrections arising from disagreement in the jet energy response between detector simulated Monte Carlo and data. Given this discrepancy, which is at the level of a few percent, it seems likely that there could be a similar level of disagreement in the missing E_T scale between data and simulated Monte Carlo. To assess the possible impact of this on the background estimates we repeat the full background calculation with a $\pm 5\%$ systematic variation in the corrected missing E_T for the Monte Carlo signal and background models, keeping the candidates and antielectron sample unchanged. Tables A.3 and A.4 show the inclusive background estimates with this missing E_T shift. The variation in backgrounds is within the statistical error on the template shapes.

In the normalisation of the backgrounds we use the theoretical values for the ratio of the W to Z cross-section, R_{WZ}, the measured $t\bar{t}$ cross-section and the NLO WW cross-section, as described above. The quoted errors on the $t\bar{t}$ and WW cross-sections are propagated as a systematic on the background by varying the cross-sections (and hence normalisations) by their uncertainty and repeating the full background calculation. However, R_{WZ} controls only the normalisation of the Z$\rightarrow$ ee background relative to the signal, and since the Z$\rightarrow$ ee contribution is very small this can safely be ignored as a potential systematic on the total background. The dependence of the background on the assumed top mass has also been investigated. The $t\bar{t}$ Monte Carlo sample we use assumes a top mass of 175 GeV/c^2. The latest combined Top mass measurement from CDF and D0 is $M_T = 172.7 \pm 2.9$ GeV/c^2. Using Top MC samples of mass 170 and 180 GeV/c^2 has

negligible impact on the top background within the top cross-section error.

In Table 7.3 we report the statistics available in the antielectron sample. One can see that in the signal region ($\not{E}_T > 30$ GeV) the statistics are rather limited, and this translates into a large statistical uncertainty in the QCD background bin-by-bin. Much smaller statistical errors on the Monte Carlo modelled backgrounds are also present. Additionally, these limited statistics translate into an error on the normalisation parameters, K_Q and K_W, obtained from the maximum likelihood fit of the background models to the candidate the missing E_T distribution, as was shown in Table 7.5. However, the errors on the fit parameters are well within the statistical errors on the total background, and can thus be safely ignored.

In Tables A.11-A.16 are reported the statistical and systematic errors on each background component, where the systematic contribution is the sum in quadrature of the effect of varying the antielectron definition, "skewing" the Monte Carlo samples and the top and WW cross-section uncertainties.

7.5 Jet Promotion Background

In Section 6.2.3 we discussed the need for an average correction to the jet energy to account for the energy deposited in the calorimeter by additional $p\bar{p}$ interactions in the same bunch crossing. This correction was parameterised by the number of additional vertices in the W candidate event i.e. the number of vertices exceeding one. Additional interactions are very likely to be soft and thus result in an isotropic distribution of energy. However, there is a small probability that, due to some inhomogeneity in the energy deposition or a semi-hard interaction, the energy is localised enough to be reconstructed as whole new jets in the event. The impact on the analysis of such occurrences will be the "promotion" of W candidate events from a lower to a higher jet multiplicity bin. Thus this is not a true background in the sense that it doesn't contribute non-signal events to the candidate sample.

Additional reconstructed vertices in W candidate events are likely to be correlated with the presence of an additional interaction, and thus additional jets. As was described in Section 6.2.3, additional interactions should be well modelled by data collected using a minimum bias (MB) trigger. Using a MB event sample we can estimate the probability of n extra jets per additional vertex, and thus estimate the size of the promotion background from knowledge of the W candidate vertex distribution.

The MB sample used contains 20,586,700 events with 18,650,748 vertices. The jet multiplicity for each vertex multiplicity sample is reported in Table 7.8. These events have been selected from the same runs as are used to form the W candidate sample; in this way the different run conditions as well as the changes in the instantaneous

N_{vtx}	1		2		3		inc	
$N_{vtx} \times$Evts	$1 \times 13{,}306{,}100$		$2 \times 2{,}133{,}810$		3×285014		$18{,}650{,}748$	
nj	evt	prob	evt	prob	evt	prob	evt	prob
1	58677	4.41×10^{-3}	21881	5.12×10^{-3}	5071	5.93×10^{-3}	88361	4.73×10^{-3}
2	10521	7.90×10^{-4}	3669	8.59×10^{-4}	849	9.92×10^{-4}	15474	8.29×10^{-4}
3	959	7.25×10^{-5}	361	8.46×10^{-5}	91	1.06×10^{-4}	1459	7.82×10^{-5}
4	166	1.24×10^{-5}	52	1.22×10^{-5}	16	1.86×10^{-5}	245	1.31×10^{-5}

Table 7.8: Different vertex multiplicity minimum bias samples are used to measure the probability to have n jets generated by a single interaction vertex. The second row reports the number of vertices in each jet multiplicity sample (i.e. the vertex number of each sample $\times$ the number of events in that sample), the jet exclusive yield and the relative probability as given by Equation 7.5. The probabilities for the 1, 2 and 3 vertex samples are given separately, while the last column is for the entire MB sample.

luminosity are taken into account.

Using the MB sample the probability for n jets to be produced per vertex is the ratio of the number of events containing n jets to the number of total vertices contained in the same event sample. Using only events with exclusively one vertex the probability to have n extra jets per vertex is given by $P(nj|1\ vtx) = N_{nj,1vtx}/N_{1vtx}$, where N_{1vtx} is the number of events with one vertex, and $N_{nj,1vtx}$ is the number of events with n jets and one vertex. For events with an arbitrary number of vertices i the previous expression can be generalised to:

$$P(nj|1\ vtx) = \frac{N_{nj,ivtx}}{i \times N_{ivtx}} \qquad (7.5)$$

The underlying assumption here is that all the jets in a single MB event come from a single active vertex. In other words, the probability of having two active vertices in the same bunch crossing is negligible. This assumption is confirmed by the results given in Table 7.8. Here samples of different vertex multiplicity are used to calculate the probability to produce n jets per vertex. The first three columns detail MB samples with exactly 1, 2 and 3 vertices respectively; the last column reports the results obtained using all the MB events. From this table it appears that the probability to generate n jets per vertex is fairly independent of the vertex multiplicity, within 20-30%. For the calculation of the promotion background we use the results of the last column, the inclusive sample, in order to have the largest possible statistics.

The probability per W candidate event to have an extra n jets (P_n) is thus the probability of n jets per vertex scaled by the average number of extra vertices per

W event:

$$P_n = \frac{Nextra\ vtx(W)}{N_{evt}(W)} \times P(nj|1\ vtx); \qquad (7.6)$$

Where $\frac{Nextra\ vtx(W)}{N_{evt}(W)} = 0.35$. Examining the results in Table 7.8, one can see that the probability of extra jets is rather small. However, these probabilities have to be weighted by the jet multiplicity spectrum found in the W candidate events. For example, if 1% of the events migrate from the ≥ 0 jet to ≥ 1 jet jet bin, this represents about 5% of the ≥ 1 jet events, because the ≥ 1 jet sample is roughly 5 times smaller than the ≥ 0 jet sample. The number of observed W candidate events with n jets, C_n, can be related to the actual number of W $+ n$ jet events (before promotion effects), W_n, by the following system of linear equations:

$$\overline{C} = P \times \overline{W}$$

$$
\begin{pmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_{\geq 4} \end{pmatrix}
=
\begin{pmatrix}
1 - \Sigma_{i=1}^{\infty} P_i & 0 & 0 & 0 & 0 \\
P_1 & 1 - \Sigma_{i=1}^{\infty} P_i & 0 & 0 & 0 \\
P_2 & P_1 & 1 - \Sigma_{i=1}^{\infty} P_i & 0 & 0 \\
P_3 & P_2 & P_1 & 1 - \Sigma_{i=1}^{\infty} P_i & 0 \\
P_4 & P_3 & P_2 & P_1 & 1
\end{pmatrix}
\begin{pmatrix} W_0 \\ W_1 \\ W_2 \\ W_3 \\ W_{\geq 4} \end{pmatrix}
$$

where $\Sigma_{i=1}^{\infty} P_i$ represents the total probability that one W $+ n$ jet event is promoted from the nth bin. We assume that the probability for an additional interaction to result in 4 jets is negligible, truncating the sum at $n = 4$. The promotion background p in each exclusive jet multiplicity bin is then given by:

$$\overline{p} = (\overline{C} - \overline{W})$$

where

$$\overline{W} = (P^{-1} \times \overline{C})$$

Note that we have to calculate the promotion background in exclusive jet multiplicity bins because the P_n probabilities are for exclusive number of jets n. The promotion background as a fraction of the number of W candidates in each inclusive jet multiplicity bin is reported in Table 7.9.

To obtain the differential promotion background as a function of first, second, third and fourth jet E_T we use the leading jet E_T spectrum from the minimum bias sample normalized to the appropriate background fraction in Table 7.9. Here we make the assumption that the extra promotion jets are always softer than the jets produced in the hard-scatter event in association with the W boson. In the

njet $\geq$	1	2	3	4
bkgd$_{frac}$ %	2.4	3.6	3.8	3.3

Table 7.9: Fraction of W candidate events that are the result of promotion background as a function of inclusive jet multiplicity.

ΔR_{jj} and M_{jj} distributions we make the assumption that the extra promotion jet is totally uncorrelated with the jets produced in the hard scatter i.e. the distribution in ΔR_{jj} of the promotion background is flat. This implies that, since the inclusive promotion background in the ≥ 2 jet bin is only 3.6%, the promotion background can be safely ignored in the ΔR_{jj} and M_{jj} distributions.

The results of the promotion background calculation are shown in the last two columns of Tables A.5-A.8. Here the error on the promotion is the sum in quadrature of the statistical error (from the minimum bias sample statistics) with a flat (in jet E_T) 30% systematic to account for the dependence on the vertex multiplicity used.

The other process which can potentially promote W + n jet events to W + (n+1) jet is $W + \gamma$. The photon produced in association with the W deposits energy in the EM calorimeter, and thus could be reconstructed as an analysis jet when it falls in the region $|\eta| \leq 2.0$ and has energy is greater than 15 GeV. We passed a simulated $W\gamma$ Monte Carlo event sample of known cross-section through our event selection criteria to estimate the cross-section for a photon to be reconstructed as a 15 GeV jet. The ratio of this to the W$\rightarrow$ eυ inclusive cross-section then gives the probability, P_γ, that a W$\rightarrow$ eυ + zero jet event will be promoted into the one jet bin. We found that this probability was 8.6×10^{-4}, within the uncertainty on the multiple interaction promotion effect described above. Therefore we do not give further consideration to this effect.

(a) ≥ 0 jet

(b) ≥ 1 jet

(c) ≥ 2 jet

(d) ≥ 3 jet

(e) ≥ 4 jet
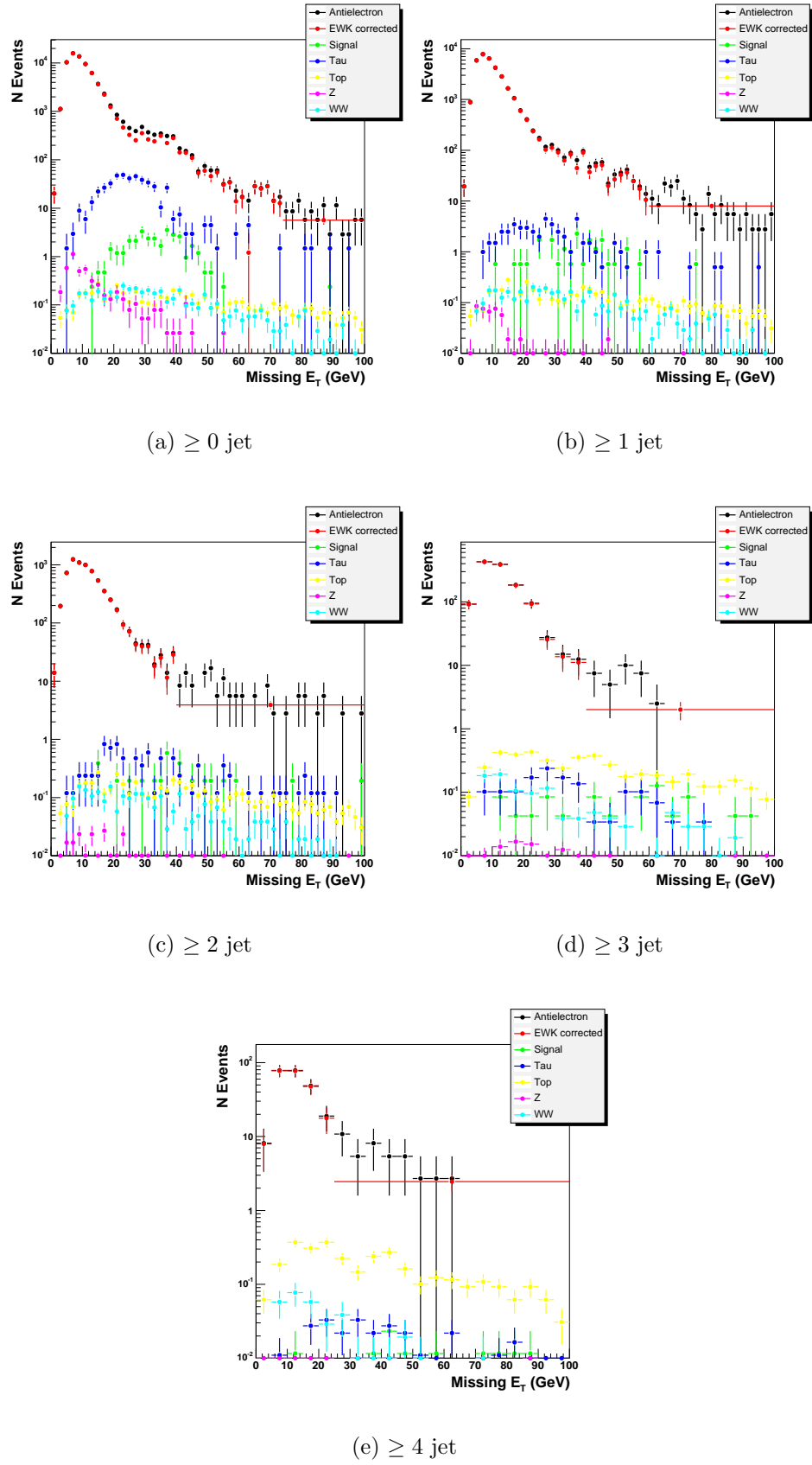
Figure 7.1: For each jet multiplicity bin is shown the antielectron missing E_T distribution before (black) and after (red) correcting for potential contamination from electroweak and top processes. Also shown are the correctly normalised electroweak and top missing E_T distributions used to make the corrections.

(a) ≥ 0 jet

(b) ≥ 1 jet

(c) ≥ 2 jet

(d) ≥ 3 jet

(e) ≥ 4 jet
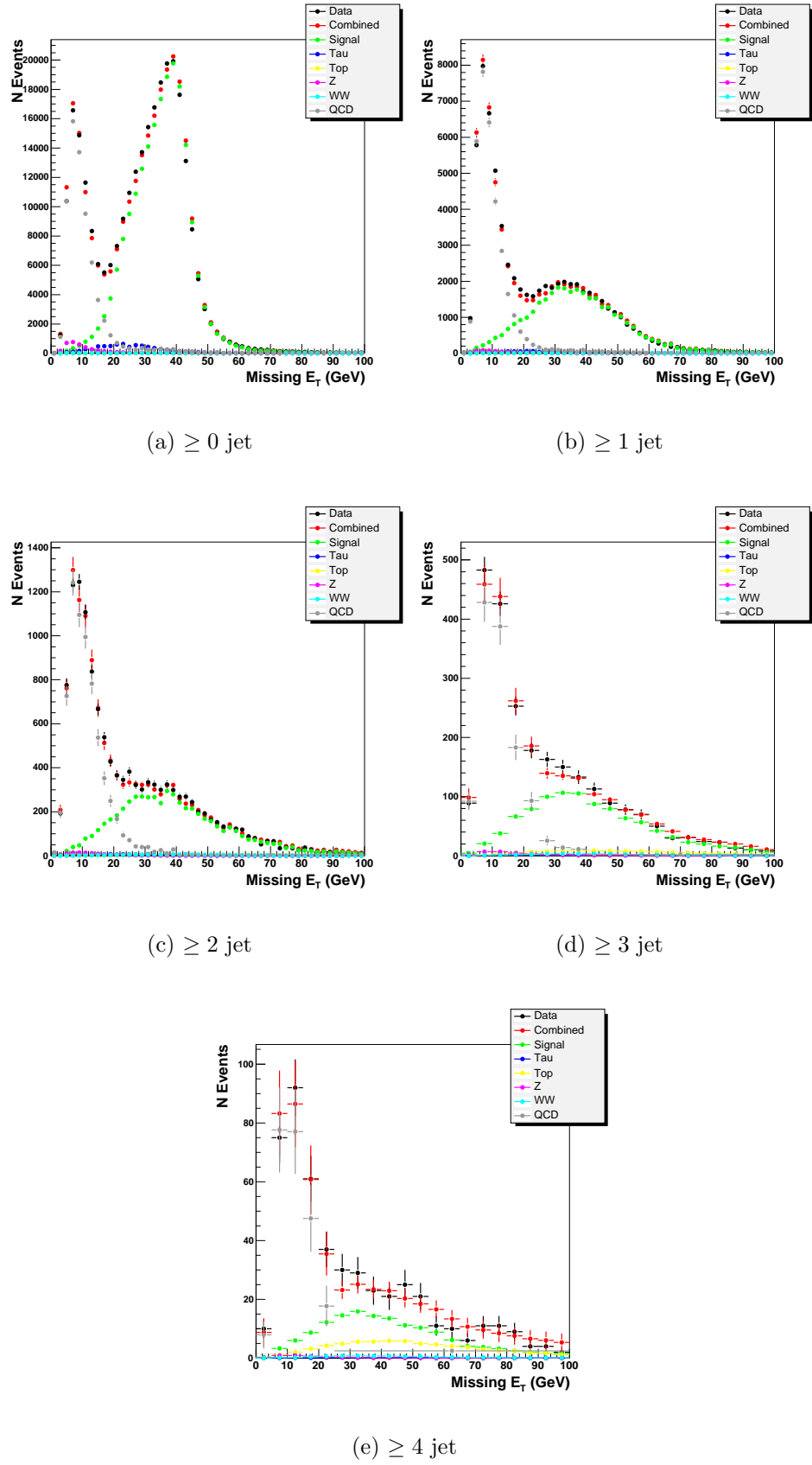
Figure 7.2: For each jet multiplicity bin is shown the results of fitting the background and signal missing E_T distributions to that of the candidates. The candidate data distribution is shown in black, and the combined signal and background distribution in red. For log scale plots see Figure 7.3.

(a) ≥ 0 jet

(b) ≥ 1 jet

(c) ≥ 2 jet

(d) ≥ 3 jet

(e) ≥ 4 jet
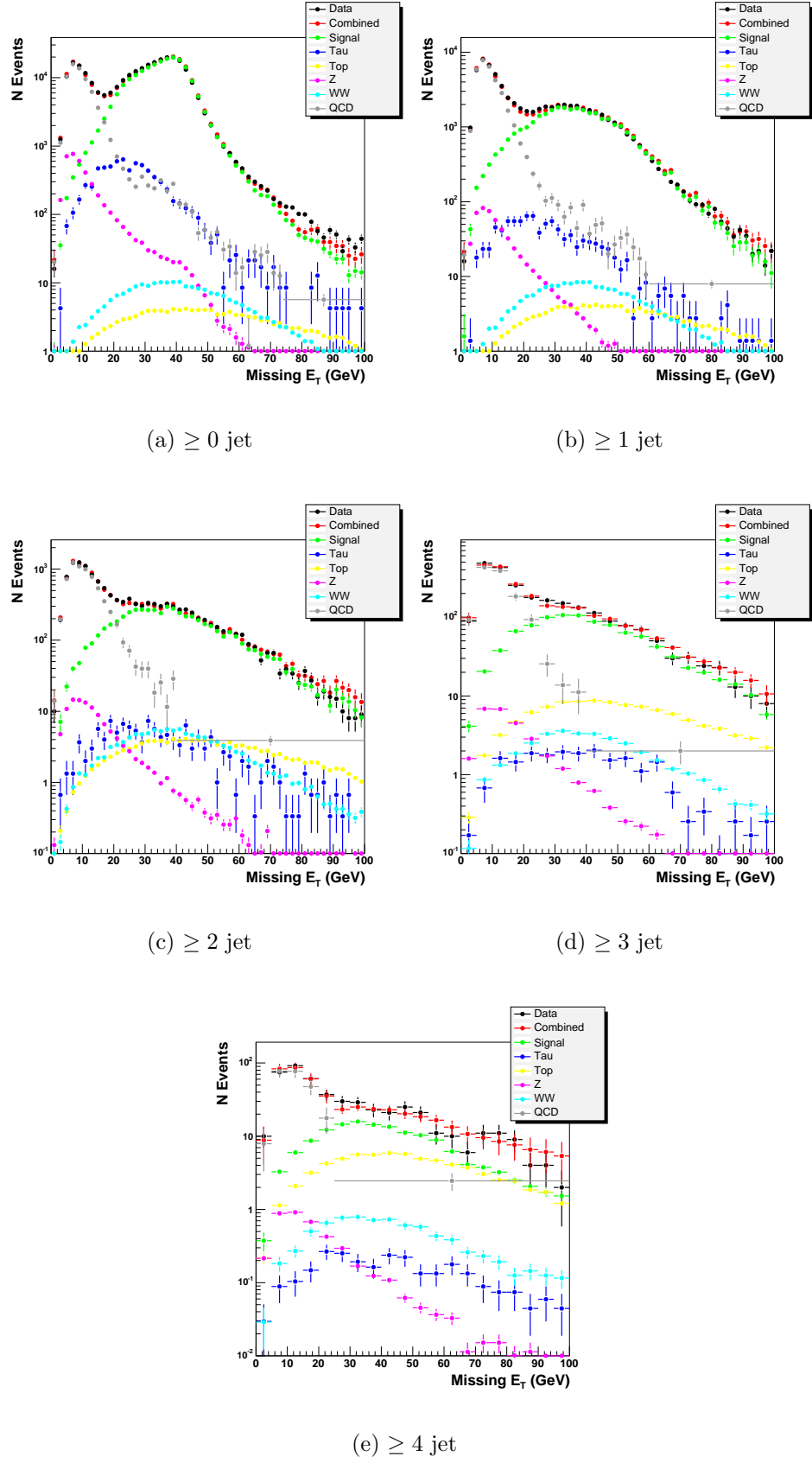
Figure 7.3: Identical to Figure 7.2 except on a log scale.

(a) ≥ 0 jet

(b) ≥ 1 jet

(c) ≥ 2 jet

(d) ≥ 3 jet

(e) ≥ 4 jet
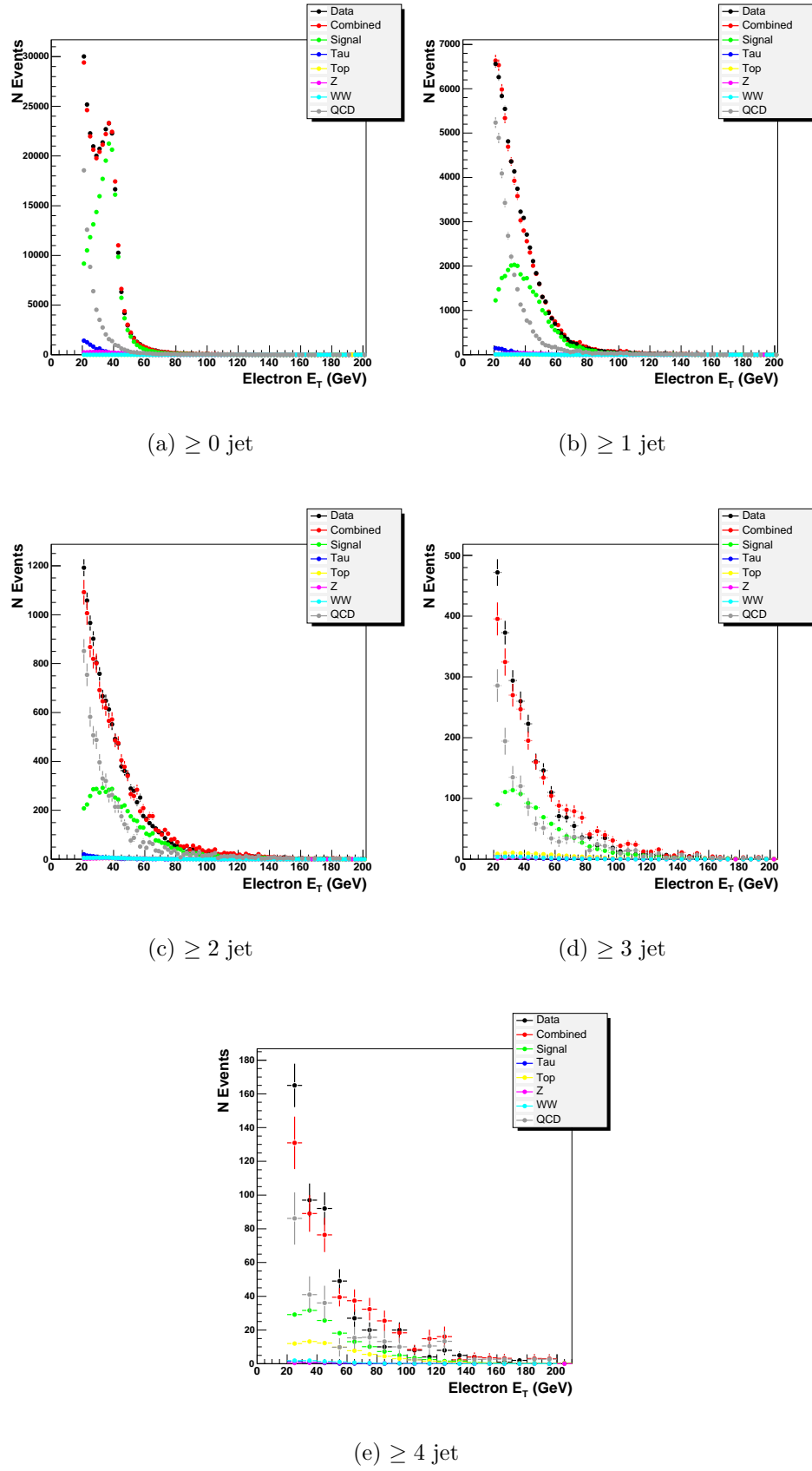
Figure 7.4: Signal plus background compared with W candidate events in the distribution of tight electron E_T . The normalisations are taken from the missing E_T fit procedure.

(a) ≥ 0 jet

(b) ≥ 1 jet

(c) ≥ 2 jet

(d) ≥ 3 jet

(e) ≥ 4 jet
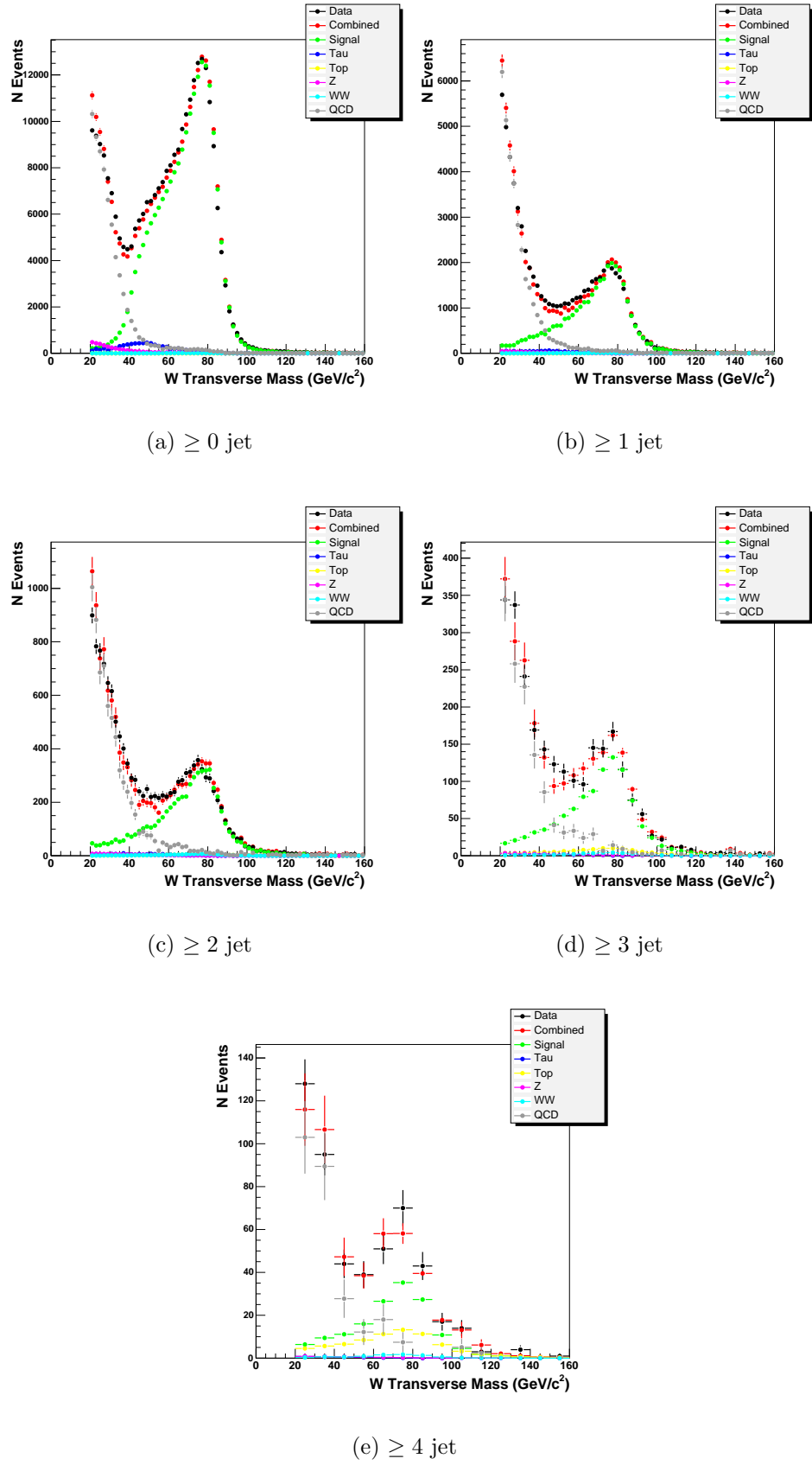
Figure 7.5: Signal plus background compared with W candidate events in the distribution of W transverse mass. The normalisations are taken from the missing E_T fit procedure.

(a) ≥ 0 jet   (b) ≥ 0 jet
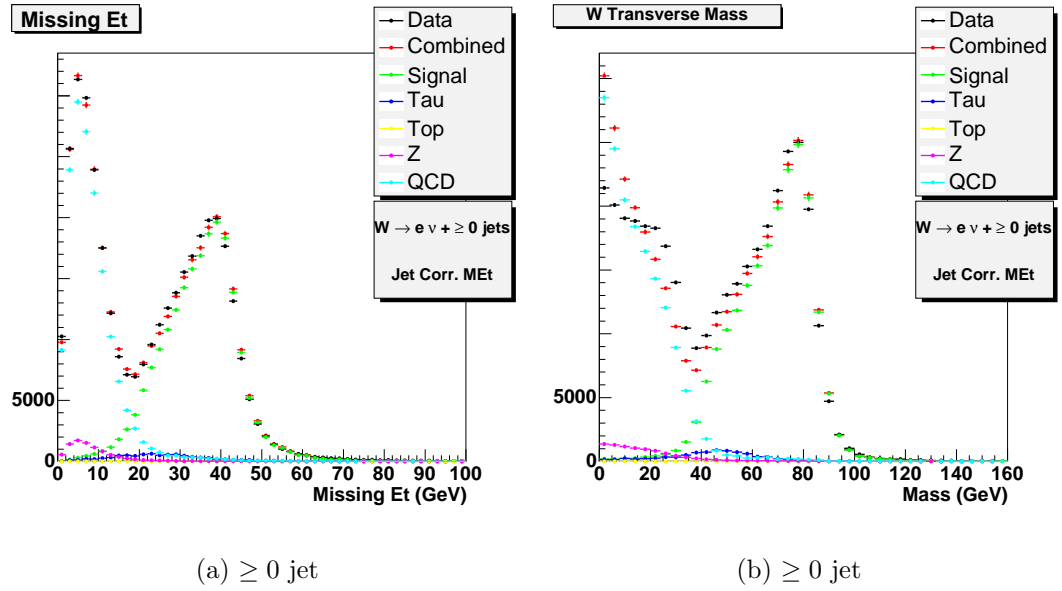
Figure 7.6: The results of removing the W transverse mass cut in the ≥ 0 jet bin. The missing E_T fit is shown on the left (note the enhanced QCD peak) and on the right the resulting picture in the W transverse mass distribution. Note the disagreement between the combined distribution (red) and the W candidates (black) below 20 GeV which motivated the inclusion of a W transverse mass cut.

(a) Lead jet E_T in ≥ 1 jet

(b) Second jet E_T in ≥ 2 jet

(c) Third jet E_T in ≥ 3 jet

(d) Fourth jet E_T in ≥ 4 jet

(e) ΔR_{jj} in ≥ 2 jet

(f) M_{jj} in ≥ 2 jet
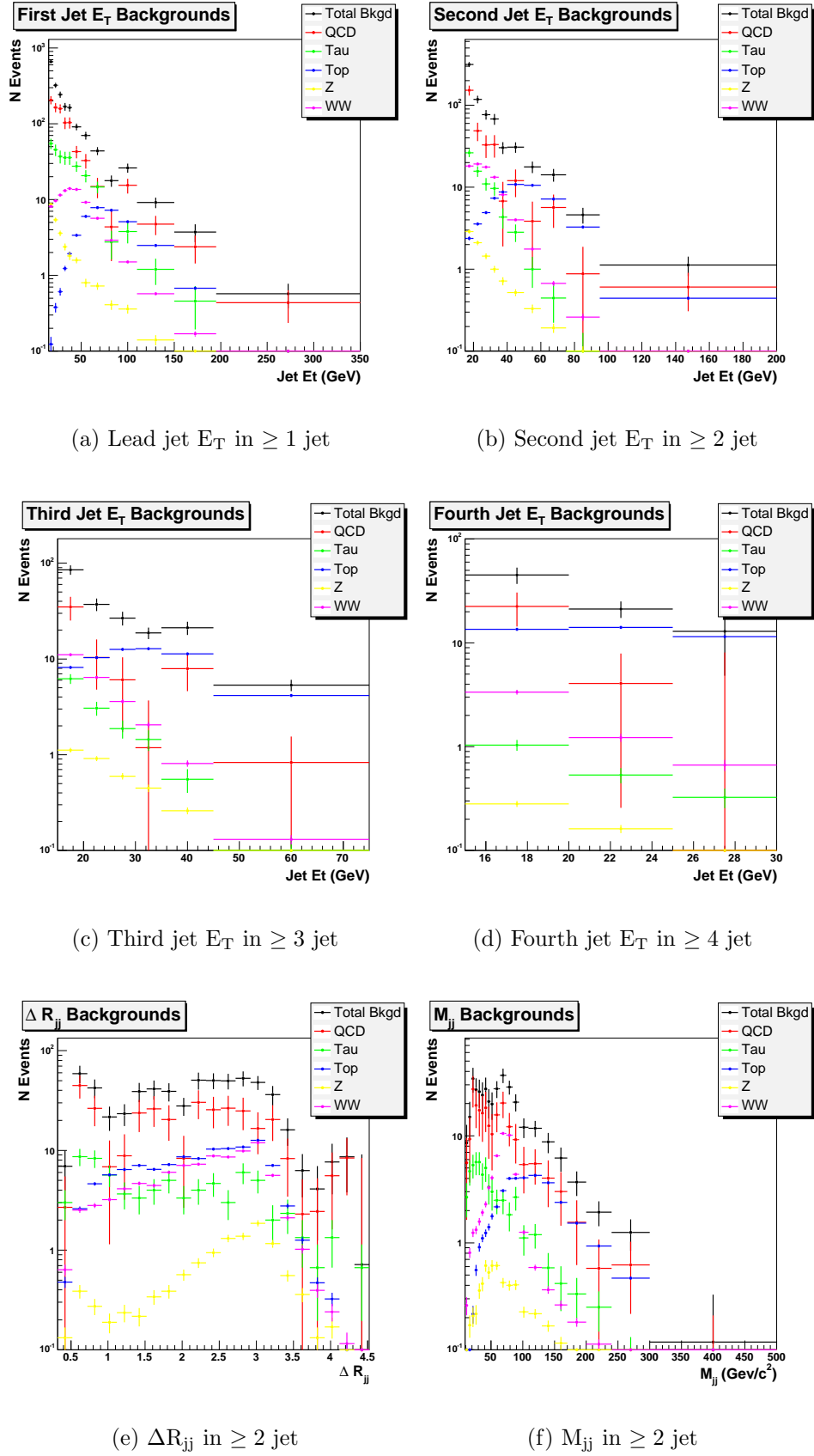
Figure 7.7: Composition of the total background as a function of jet E_T. The correctly normalised contribution of each background source is shown, along with their total.
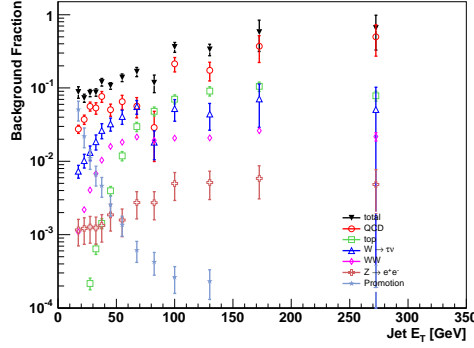
(a) Lead jet E_T in ≥ 1 jet
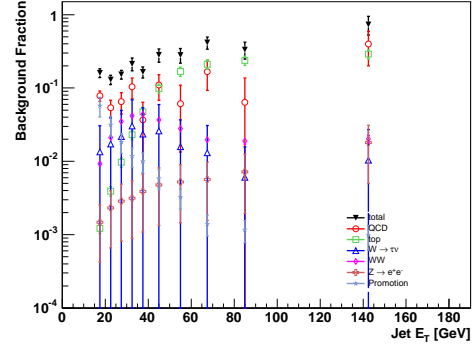
(b) Second jet E_T in ≥ 2 jet
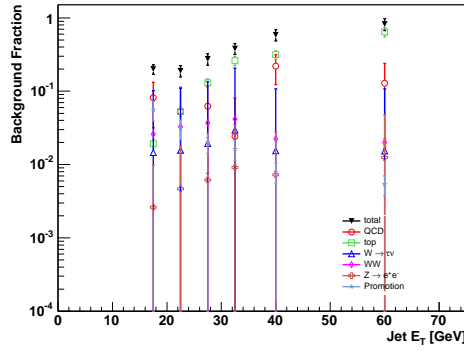
(c) Third jet E_T in ≥ 3 jet
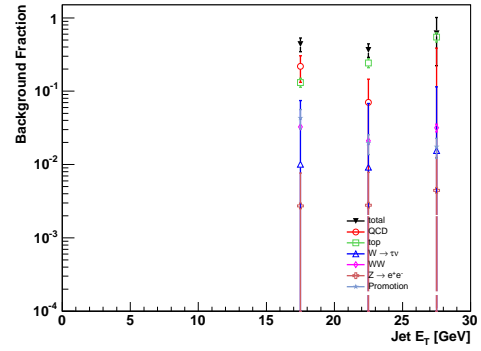
(d) Fourth jet E_T in ≥ 4 jet
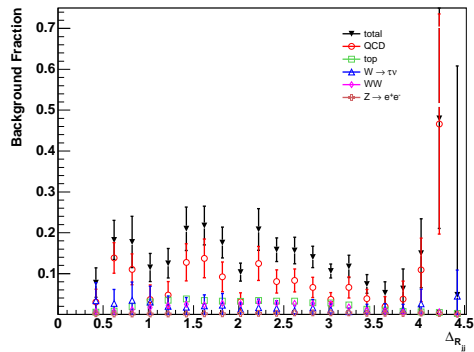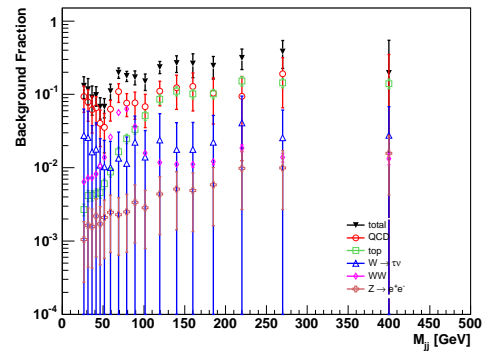
(e) ΔR_{jj} in ≥ 2 jet

(f) M_{jj} in ≥ 2 jet

Figure 7.8: Background as a fraction of the number of candidate events in each bin. Shown is the total background fraction as well as that from each individual background source.
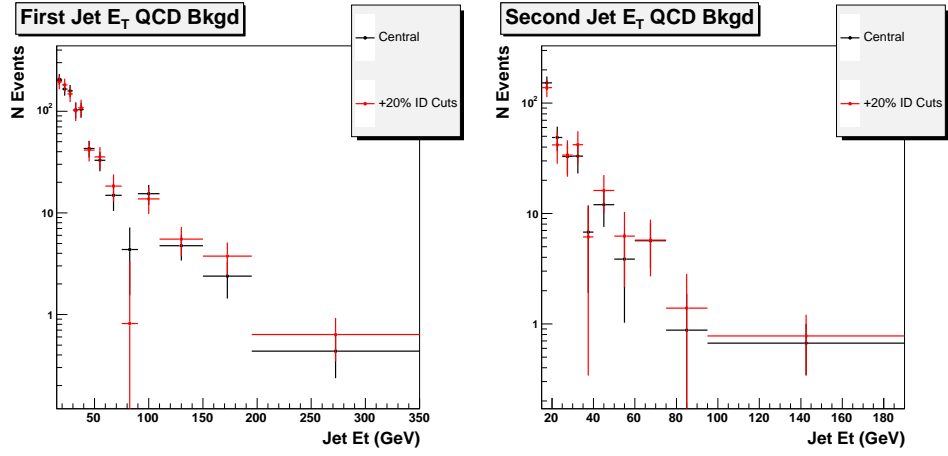
(a) Lead jet E_T in ≥ 1 jet

(b) Second jet E_T in ≥ 2 jet
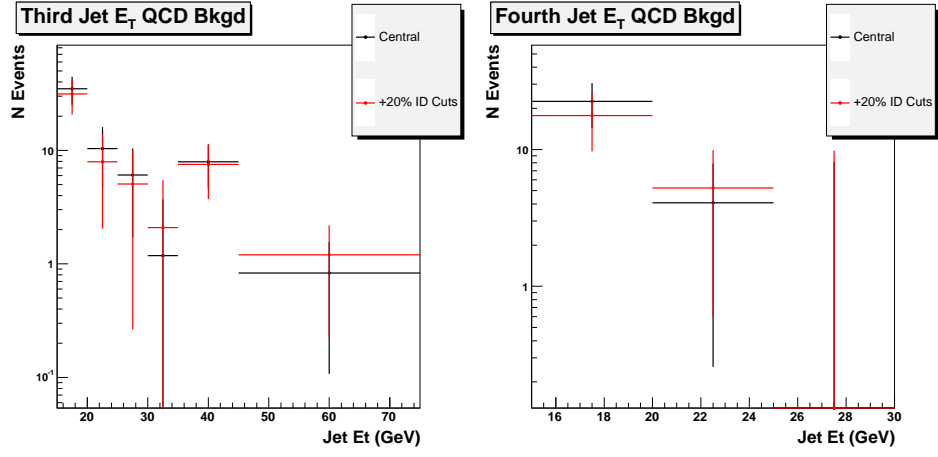
(c) Third jet E_T in ≥ 3 jet

(d) Fourth jet E_T in ≥ 4 jet
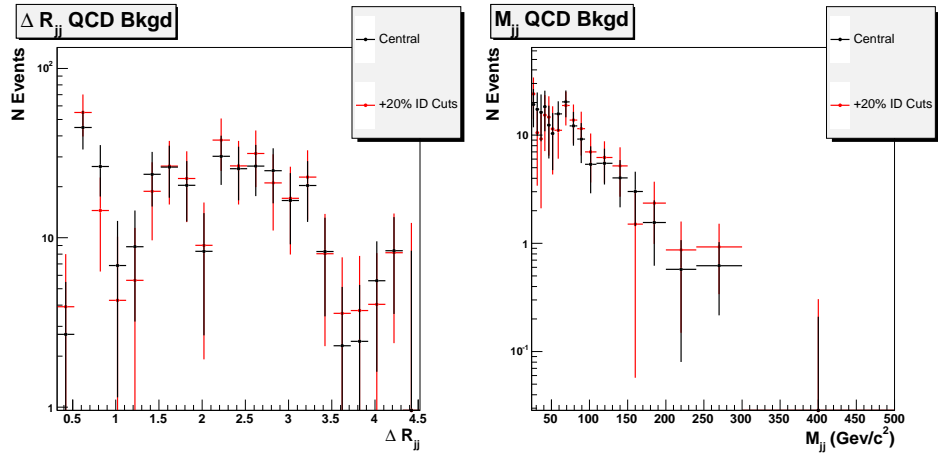
(e) ΔR_{jj} in ≥ 2 jet

(f) M_{jj} in ≥ 2 jet

Figure 7.9: Comparison of the differential QCD background estimate obtained using the standard antielectron sample (black) and a redefined antielectron sample (red) where the ID cuts are varied by 20% away from the signal region.

Chapter 8

Acceptance and Efficiency Corrections

The W selection criteria described in Section 6.1 are necessarily imposed on the dataset to select W→ ev events from the myriad of processes that will be present. We have already seen that the W candidate event sample contains not only W events but also background events, and that these have to be corrected for. In addition, by imposing geometric and kinematic restrictions on the electron and neutrino, and requiring the electron to pass stringent quality requirements, we "lose" W→ ev events that are actually present in the dataset. To illustrate this phenomenon, Figure 8.1 shows the $\not{E}_T$ distribution of simulated W→ ev Monte Carlo, where all the tight selection criteria have been imposed with the exception of the 30 GeV missing E_T cut. One can see that a fraction of the W→ ev events have a missing E_T less than 30 GeV. This is quite expected given the physics of W production at hadron colliders. A W→ ev decay at rest in the transverse frame would produce a neutrino with $E_T = M_W sin\theta/2$ and an electron with $E_T = M_W sin\theta/2$, where θ is the angle of the decay products to the z-axis. The presence of initial state radiation and/or recoiling jets produced in the hard scatter means that the W boson is rarely if ever decaying at rest in the transverse frame, and subsequently this E_T distribution of the decay products is "smeared". Additional smearing also results from the resolution of the detector measurements. If we are to measure W→ ev cross-sections, we have to correct our event yield for the W→ ev events that are "lost" in this way.

Generally the term "acceptance" is taken to mean a correction factor applied to the event yield to account for the "loss" of signal events due to geometric and kinematic cuts made in the event selection, whereas "efficiency" is a factor that accounts for the loss of signal events due to the requirements of certain lepton identification criteria and is to first order independent of the particular signal process under study. As will be discussed shortly, in this analysis we define a total acceptance factor $\mathcal{A}$ which accounts for all the W event selection criteria; all the geometric,
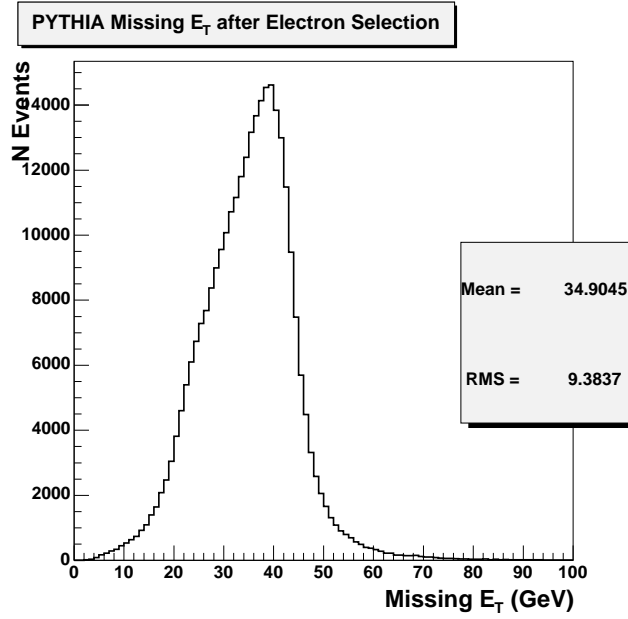
Figure 8.1: Missing E_T distribution from a simulated PYTHIA W$\to$eυ Monte Carlo sample where all the tight selection criteria have been imposed with the exception of the 30 GeV missing E_T cut.

kinematic and electron identification cuts. It is this which is used in Equation 5.1 to calculate the cross-section. However, in this chapter we shall also break down the total acceptance $\mathcal{A}$ into its various acceptance and efficiency components to examine individually the impact of each of the selection cuts. This is done purely to investigate the validity of the acceptance correction procedure.

Essentially an acceptance factor is the ratio of signal events "accepted" by the cuts to the total number of signal events that were passed through the selection criteria. Since we don't know how many signal events are in the initial dataset, the only option for determining acceptance factors is to use detector simulated signal Monte Carlo samples, as detailed in Section 8.1 below. In our case we will be using ELO ALPGEN + PYTHIA W$\to$eυ + n parton simulated Monte Carlo samples for this purpose, described in Table 8.1. In terms of making a measurement that is independent of theoretical models used this is potentially a problem, and in Section 8.3 we investigate the dependence of the acceptance factors on the details of the ELO approach used.

In using W$\to$eυ Monte Carlo to calculate the total acceptance factor $\mathcal{A}$ we are implicitly relying on simulated W$\to$eυ Monte Carlo to reproduce the efficiency of the electron identification criteria. This is a departure from previous W + Jets analyses [20] [107], where this factor has been estimated using a Z$\to$ee data sample, as will be described in Section 8.4. However, since ours is a differential cross-section measurement, we need to know the acceptance and efficiency factors in each bin of

MC Sample	σ pb	Q^2 Scale	Parton Generation Cut	N_{gen}
W$\to$ eυ + n parton				
A+P 1 parton	712	$M_W^2 + sumP_T^2(p)$	minP$_T$ 10 GeV, ΔR 0.4	305K
A+P 2 parton	249.5	$M_W^2 + sumP_T^2(p)$	minP$_T$ 10 GeV, ΔR 0.4	239K
A+P 3 parton	85.16	$M_W^2 + sumP_T^2(p)$	minP$_T$ 10 GeV, ΔR 0.4	224K
A+P 4 parton	28.94	$M_W^2 + sumP_T^2(p)$	minP$_T$ 10 GeV, ΔR 0.4	67K

Table 8.1: Description of the Monte Carlo samples used in the differential acceptance calculation. A+P = ALPGEN + PYTHIA ELO sample.

the jet E$_T$, ΔR$_{jj}$ and M$_{jj}$ distributions. There simply are not enough Z$\to$ ee data statistics available to do this. In Section 8.4 we attempt to validate this approach within the available statistics by comparing the efficiency factors determined from W$\to$ eυ Monte Carlo and Z$\to$ ee data.

Throughout this chapter we will use the following nomenclature to make the acceptance formulae less cluttered:

$$A = \frac{N^{num}}{N^{den}} = \frac{N\left[X^{num}\right]}{N\left[X^{den}\right]} \tag{8.1}$$

where X^{num} represents all the cuts in the numerator of the acceptance calculation, X^{den} represents all the cuts in the denominator, and N^{num} and N^{den} are the number of events which pass the numerator and denominator cuts respectively. In calculating the acceptance using detector simulated Monte Carlo samples, we make cuts on detector reconstructed quantities, such as missing E$_T$ or fiduciality, but also cuts on "truth" or "generator" level quantities, such as the electron P$_T$ as generated by the Monte Carlo. We distinguish cuts made at the generator and detector simulated level by the following notation:

Reco(X) - Cuts X made at detector level

Gen(X) - Cuts X made at generator level

8.1   Our Acceptance Definition

In this analysis we use the ELO ALPGEN + PYTHIA W$\to$ eυ Monte Carlo samples detailed in Table 8.1 to calculate a total acceptance factor $\mathcal{A}_i$ in each bin i of each jet E$_T$, ΔR$_{jj}$ and M$_{jj}$ distribution, which is then used in the cross-section calculation

in that bin. Using the nomenclature explained above the calculation of the factor $\mathcal{A}_i$ can be expressed in the following formula:

$$\mathcal{A}_i \;=\; \frac{N_i\left[\mathrm{Reco}(E_T; P_T;\, \not\!E_T; WM_T; \eta^D; \mathrm{fid}; \Delta R_{ej}; ID; ZVeto)\right]}{N_i\left[\mathrm{Gen}(P_T^e; \eta^e; P_T^v; WM_T)\right]} \qquad (8.2)$$

For an event to be counted in the numerator it must pass ALL of the W candidate selection criteria described in Section 6.1 (including electron ID criteria) at the detector reconstructed level. For an event to enter the denominator the generator-level W electron is required to have $P_T > 20$ GeV and $\eta < 1.1$, the generator-level W neutrino $P_T > 30$ GeV, and the generator-level W boson a transverse mass of > 20 GeV. The events in both the numerator and denominator must pass the jet requirements of the bin i. Note that, since the numerator is not a subset of the denominator, this definition allows acceptance factors to be greater than one. Exactly how we arrive at this definition for the total acceptance $\mathcal{A}$ is described in the paragraphs that follow.

In previous W + Jets analyses [20] [107] detector simulated W$\rightarrow$ev Monte Carlo samples have been used to calculate the acceptance of the geometric and kinematic requirements using the following formula:

$$A = \frac{N\left[\mathrm{Reco}(E_T; P_T;\, \not\!E_T; WM_T; \eta^D; \mathrm{fid})\right]}{N_{\mathrm{Gen}}} \qquad (8.3)$$

Where the numerator is the total number of events that pass the geometric and kinematic cuts at the detector level and N_{Gen} is the total number of events in the MC sample (the number of MC events generated). In this study we calculate both the acceptance *and* efficiency using W Monte Carlo and hence require all the W selection criteria in Table 6.1 to be present in the numerator:

$$A = \frac{N\left[\mathrm{Reco}(E_T; P_T;\, \not\!E_T; WM_T; \eta^D; \mathrm{fid}; \Delta R_{ej}; ID; ZVeto)\right]}{N_{\mathrm{Gen}}} \qquad (8.4)$$

At this stage we have an acceptance definition which would correct the candidate event yield to that of the full W$\rightarrow$ev production cross-section. However, in this analysis we do not attempt to make such a measurement. As was discussed in Chapter 5, we kinematically restrict the decay of the W in our cross-section definition to reflect the kinematic cuts made in the event selection. Practically this redefinition of the cross-section translates as a redefinition of the acceptance. In comparing Equation 8.2 with Equation 8.4, we see that in our definition of the total acceptance $\mathcal{A}$ the denominator is now no longer every event generated in the Monte Carlo,
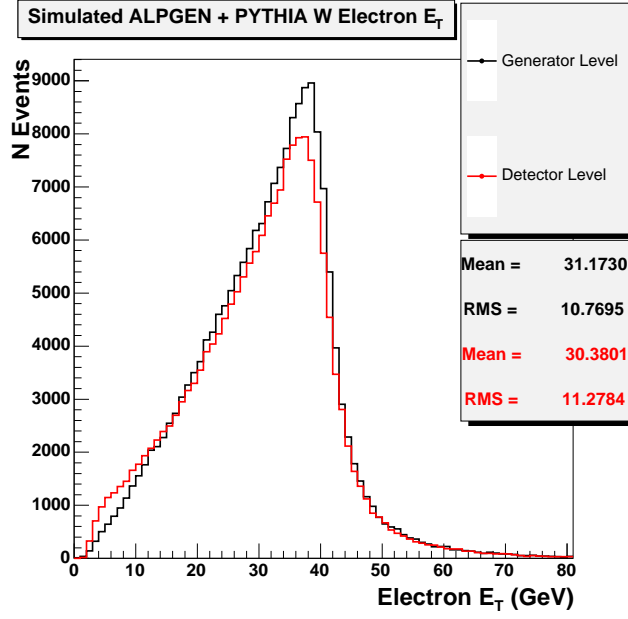
Figure 8.2: Generator-level and Detector-level W electron E_T distributions, where the W electron at generator-level has been matched in R (requiring $\Delta R < 0.005$) to the detector-level EM cluster.

but the number of events that at generator level pass the kinematic and geometric restrictions on the cross-section.

The reason for redefining the acceptance in this way is illustrated by Figure 8.2. This shows the generator-level W electron E_T distribution (black) and for the same electron the detector-measured E_T (red). One can see the differences in these distributions, caused by the detector resolution. In previous analyses the acceptance of the electron E_T cut would be defined using the detector-level (red) distribution as the ratio of the area above 20 GeV to the total area. Thus the acceptance will be very much dictated by the shape of the underlying generator-level distribution i.e. the W + Jet physics. In our new definition the acceptance is given by the ratio of the area of the detector-level distribution above 20 GeV to the area of the generator-level distribution above 20 GeV. This should only be dependent on the resolution of the detector measurements and the local shape of the generator-level distribution around 20 GeV i.e. much less dependent on the underlying W + Jets theoretical model.

8.1.1   Factorising the Acceptance

In order to understand how the various W selection criteria contribute to the acceptance factor it is useful to factorise $\mathcal{A}$, given by Equation 8.2, into a number of different component factors, each of which is related to the acceptance of a particular cut or collection of cuts:

$$\mathcal{A} = A_{geo} \cdot A_{kin} \cdot A_{je\Delta R} \cdot \xi_{ID} \cdot A_{Zrej} \tag{8.5}$$

Note that these component acceptances are constructed such that their product is equal to the total acceptance $\mathcal{A}$. However, it is important to note that these component acceptance factors are not used in the cross-section calculation; we only use the total acceptance factor $\mathcal{A}$ as defined in Equation 8.2.

The acceptance of the geometric and kinematic W selection requirements is defined as follows:

$$A_{geo} = \frac{N\left[Reco(\eta^D; fid) \ \& \ Gen(P_T^e; P_T^v; WM_T)\right]}{N\left[Gen(P_T^e; \eta^e; P_T^v; WM_T)\right]} \tag{8.6}$$

$$A_{kin} = \frac{N\left[Reco(E_T; P_T; \not{E}_T; WM_T; \eta^D; fid)\right]}{N\left[Reco(\eta^D; fid) \ \& \ Gen(P_T^e; P_T^v; WM_T)\right]} \tag{8.7}$$

Note that an implicit requirement in the numerator of A_{geo} is that there is an electromagnetic cluster in the event against which we can test the η^D and fiducial requirements. This will not always be the case, one can imagine several instances where a W electron may not be reconstructed as an electromagnetic cluster:

1. The W electron is incident on one of the cracks in the calorimeter, thus not depositing sufficient electromagnetic energy to be reconstructed as an electron cluster object.

2. The W electron overlaps with a high P_T jet produced in the hard scatter and is "obliterated", again the reconstruction code is unable to form an electron cluster object.

Thus there is an "electron reconstruction" efficiency built into the acceptance $\mathcal{A}$ definition.

The efficiency of the jet-electron ΔR cut (see Section 6.1.4) is defined relative to A_{kin} as follows:

$$A_{je\Delta R} = \frac{N\left[X_{kin}^{num} \ \& \ Reco(\Delta R_{ej})\right]}{N_{kin}^{num}} \tag{8.8}$$

Where X_{kin}^{num} represents all the cuts in the numerator of A_{kin}, and N_{kin}^{num} is the number of events which pass the cuts in the numerator of A_{kin}.

Similarly the efficiency of the electron identification cuts (including the conversion veto) is defined relative to $A_{je\Delta R}$ as follows:

$$\xi_{\mathrm{ID}} = \frac{N\left[X^{\mathrm{num}}_{\mathrm{je}\Delta R} \ \& \ \mathrm{Reco}(\mathrm{fIso};\mathrm{Had/EM};\chi^2;L_{\mathrm{shr}};\Delta X;\Delta Z;\mathrm{COTseg};\mathrm{conv})\right]}{N^{\mathrm{num}}_{\mathrm{je}\Delta R}} \quad (8.9)$$

and the efficiency of the Z veto (see Section 6.1.4) defined relative to ξ_{ID} as:

$$\xi_{\mathrm{Zrej}} = \frac{N\left[X^{\mathrm{num}}_{\mathrm{ID}} \ \& \ \mathrm{Reco}(\mathrm{ZVeto})\right]}{N^{\mathrm{num}}_{\mathrm{ID}}} \quad (8.10)$$

8.1.2 Subtleties in the Acceptance Definition

There are several subtle issues which require careful consideration when calculating the acceptance from $W \to e\upsilon$ Monte Carlo:

- The counting of jets in the numerator and denominator.

- The definition of the W electron on which we make generator-level denominator cuts.

- The definition of the detector-level electron on which we make the reconstructed level cuts.

As was noted in Section 6.2.1, the electron is not removed from the calorimeter before jet clustering and is thus always present in the jet collection unless explicitly removed. In the analysis we remove this electron from the jet collection by matching the selected tight electron to a jet in R. However, in the denominator of the acceptance factor $\mathcal{A}$ (equation 8.2) we are making cuts at the generator-level level, no tight electron at the detector-level has yet been defined. In order to correctly bin the denominator by the number and E_{T} of detector-level jets the electron must be removed, and this is done by matching the generator-level electron in R to a jet. Since generator and detector-level electrons can differ in R there could be a small discrepancy between this procedure and that used in the analysis.

There exists in the event record[1] of the $W \to e\upsilon$ Monte Carlo samples W daughter electrons and neutrinos at various stages of the event evolution e.g. before and after final state gluon and/or QED radiation. The exact generator-level electron/neutrino that is used for the denominator cuts is closely linked to the cross-section definition. If for example we take the post-QED FSR electron, theorists will have to implement a QED FSR model in order to compare to our result. If we take the pre-QED FSR electron, we will have a QED FSR model built into our acceptance. The same

---

[1]The Monte Carlo event record contains descriptions of all the generator-level particles (the type of particle, P_{T}, mass etc.) at each stage of the event evolution, from the initial $p\bar{p}$ interaction to the final state.

applies to gluon radiation, which also effects the neutrino definition. However, QED FSR is a well understood and predicted phenomenon, and thus having a QED FSR model built into the acceptance is not a problem. Gluon radiation on the other hand is a QCD phenomenon, and one of the main motivations behind making the measurement is to enable theorists to compare models of gluon radiation to data. Thus we do not want any model of gluon radiation to be built into our result. We define our generator-level electron and neutrino as those before QED FSR radiation but after final state gluon radiation.

In the $W \to e\upsilon$ candidate event selection every electromagnetic object in the event is subjected to the tight electron criteria. The same can of course be done in the computation of the acceptance factor $\mathcal{A}$, giving every reconstructed electromagnetic object in the event the chance to pass the numerator cuts of equation 8.2, and this is the correct way to compute this factor. However, if we are to factorise $\mathcal{A}$ into different component acceptance factors the picture becomes more complicated. The numerator and denominator requirements of the A_{geo} and A_{kin} definitions involve making detector level geometric and kinematic cuts before any tight identification cuts have been made. Thus jets can quite conceivably be present as an electromagnetic object and pass these cuts, leading to geometric and kinematic acceptances which are inconsistent with the total $\mathcal{A}$ factor. Furthermore, these jets that pass the geometric and kinematic requirements will very likely fail the electron ID cuts, diluting the ξ_{ID} factor. In order to prevent such jet contamination it is thus necessary to require that the detector-level electromagnetic object matches the generator-level W electron with $\Delta R < 0.4$. However, in making such a matching requirement we are neglecting from consideration a class of $W \to e\upsilon$ event where an electromagnetic object is produced that passes all the identification criteria but which is not the W decay electron, for example a jet with anomalously large electromagnetic calorimeter deposits. Figure 8.3(a) shows the lead jet E_T distributions of $W \to e\upsilon$ Monte Carlo events which pass all the selection criteria, one requiring a match to the generator-level W electron and the other not. The ratio of these distributions, shown in Figure 8.3(b), indicates that this effect is at the level of at most 1%, and thus negligible.

8.2 Acceptance Results

Figure 8.4 shows the differential acceptance results for each of the jet E_T, ΔR_{jj} and M_{jj} distributions, which are also summarised in Tables B.1 -B.6. These results were obtained using the ELO ALPGEN + PYTHIA detector simulated $W \to e\upsilon$ Monte Carlo samples described in Table 8.1. As with the estimation of the background, the n parton sample is used in the $\geq n$ jet bin.
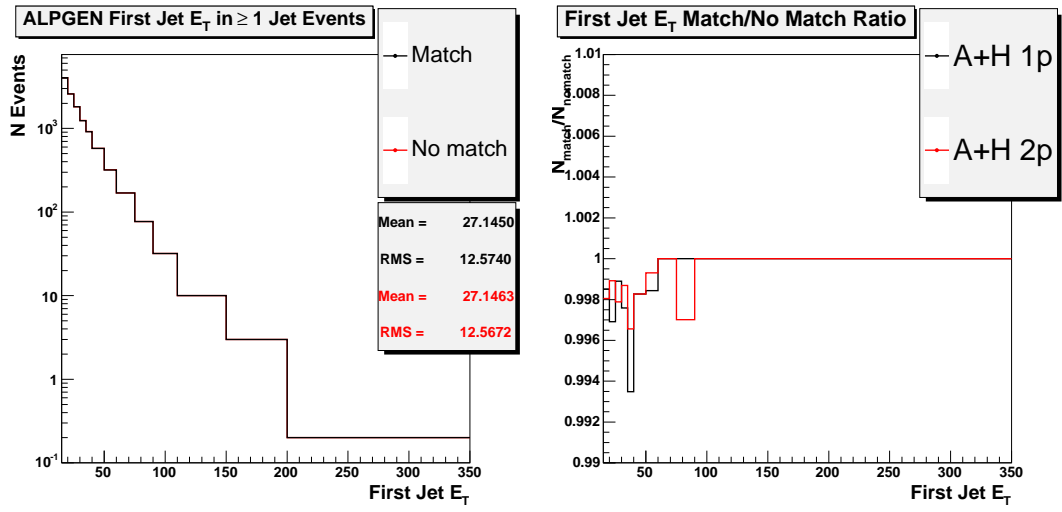
(a) Lead jet E_T distribution of W selected PYTHIA W$\rightarrow$eυ events; the black curve is where the reconstructed electron has been required to match the generator-level electron, in the red curve no such match is required.

(b) Ratio of black and red curves opposite. Also shown for ALPGEN+HERWIG 1 and 2 parton samples.

Figure 8.3: Studying the impact on the acceptance calculation of matching the generator-level electron to the reconstructed electron.

One can see that both the total and component acceptance factors are more or less constant as a function of jet E_T, ΔR_{jj} and M_{jj}. The largest structure in the total acceptance is observed in the M_{jj} distribution at low jet-jet invariant mass, at the level of 5% (absolute) i.e. the structure of the differential cross-section will thus be affected by at most 5%. As will be shown in Section 10.3, this is within the systematics of the cross-section measurement. In other words, the total acceptance effectively has no dependence on jet E_T, ΔR_{jj} or M_{jj}.

This result is not entirely unexpected. We are measuring the acceptance of selection cuts which are concerned with W production and decay physics, detector resolutions and the identification of electrons in the detector. To first order these should be unrelated to the hadronic activity in the event. With the acceptance definition used in previous analyses, Equation 8.3, one might expect a greater correlation with jet properties in the geometric and kinematic acceptance, since the P_T and η spectra of the electron and neutrino are related to the W P_T distribution.

8.3 Theoretical Model Dependence

To determine the differential acceptance factors we use the simulated ELO ALPGEN + PYTHIA W$\rightarrow$eυ Monte Carlo samples detailed in Table 8.1, in particular the
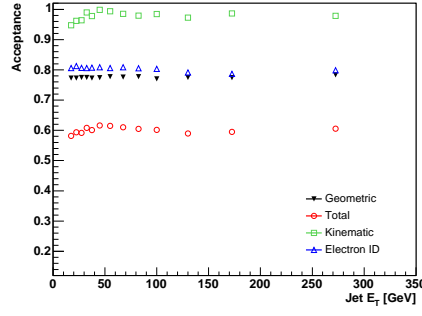
(a) Lead jet E_T in ≥ 1 jet
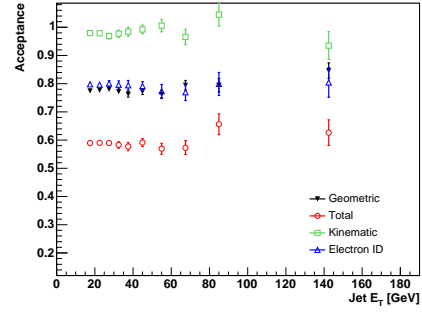
(b) Second jet E_T in ≥ 2 jet
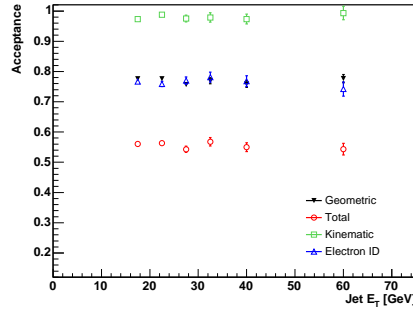
(c) Third jet E_T in ≥ 3 jet
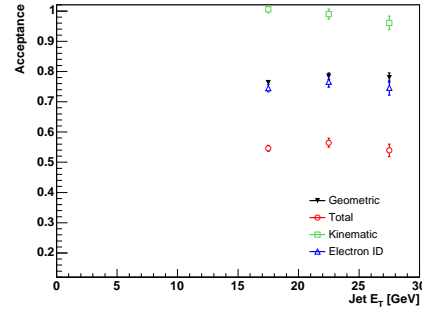
(d) Fourth jet E_T in ≥ 4 jet
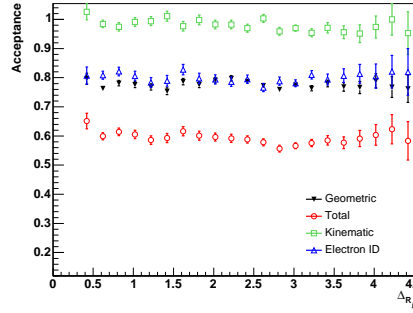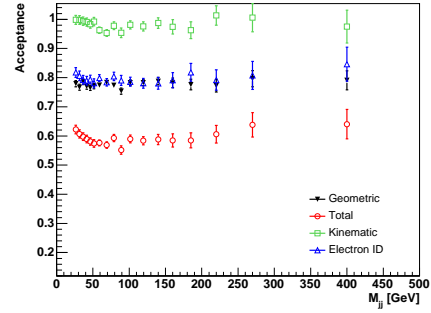
(e) ΔR_{jj} in ≥ 2 jet

(f) M_{jj} in ≥ 2 jet

Figure 8.4: Acceptance results for each of the jet E_T , ΔR_{jj} and M_{jj} distributions. Shown is the total acceptance $\mathcal{A}$, and the geometric (A_{geo}), kinematic (A_{kin}) and electron ID (ξ_{ID}) factors.

n parton sample in the $\geq n$ jet bin. In Section 8.1 it was explained that our new acceptance definition should limit the dependence of the acceptance result on the particular theoretical model used. However, some theoretical dependence may still remain, and to estimate this we examine the differential acceptance factors obtained in the jet E_T distributions when we use the $n+1$ and $n-1$ parton samples in the

$\geq n$ jet bin. Broadly speaking, using the $n-1$ parton sample means that the nth jet is likely generated by the parton shower, whereas using the $n+1$ parton sample will result in events with more jets than the n parton sample, and hence a harder W P_T distribution. Therefore, by varying the number of partons in this way, we probe the sensitivity of the acceptance to the limitations of the ELO approach, that is, it's ability to predict W $+ n$ jet cross-sections and the parton shower description of additional jets. This is important, since it is limitations of this kind which we wish to evaluate when we make theoretical comparisons to our measured cross-sections.

Figures 8.5 and 8.6 show the A_{geo} and A_{kin} factors as a function of the first, second, third and fourth jet E_T for different n parton multiplicity ALPGEN + PYTHIA MC samples detailed in Table 8.1. One can see that the geometric acceptance is completely unaffected by changes in the parton multiplicity. At low first and second jet E_T the kinematic acceptance displays some dependence on the parton multiplicity, showing variation on the scale of 5%.

Figures 8.7 and 8.8 show the $A_{je\Delta R}$ and ξ_{Zrej} factors as a function of jet E_T. These factors are constant as a function of jet E_T and are effectively independent of the parton multiplicity. Theoretical model dependence in the ID efficiency estimation is examined in the next section.

8.4 Validating Monte Carlo ID Efficiency Against Z Data

The Had/EM, E/P, ΔX, ΔZ, χ^2, L_{shr}, COT track quality and fractional isolation cuts detailed in Section 6.1.2 are specific detector measurements designed to discriminate real electrons from fakes. In estimating the efficiency of these cuts from W$\rightarrow$ eυ Monte Carlo we rely heavily on the CDF detector simulation, and to first order the efficiency should be process independent.

Given the complex nature of many of these measurements it is perhaps unsurprising to find that the detector simulation is somewhat limited in it's ability to reproduce them. Figures 8.9 and 8.10 compare the distribution of the ID variables in the W candidate events to those of simulated Z$\rightarrow$ ee Monte Carlo events. In each distribution all of the tight electron selection requirements have been imposed with the exception of the cut under consideration. One can see that in most variables the simulated Monte Carlo reproduces the data shape reasonably well, but there are clear differences in the Had/EM, fractional isolation and CES ΔX and ΔZ distributions. However, these differences largely occur well within the cuts used to identify the electron (e.g. the CES ΔZ cut is $|\Delta Z| < 3.0$). In other words, after all the other cuts have been made the differences between data and Monte Carlo in a particular variable do not appear significant for tight electron selection.
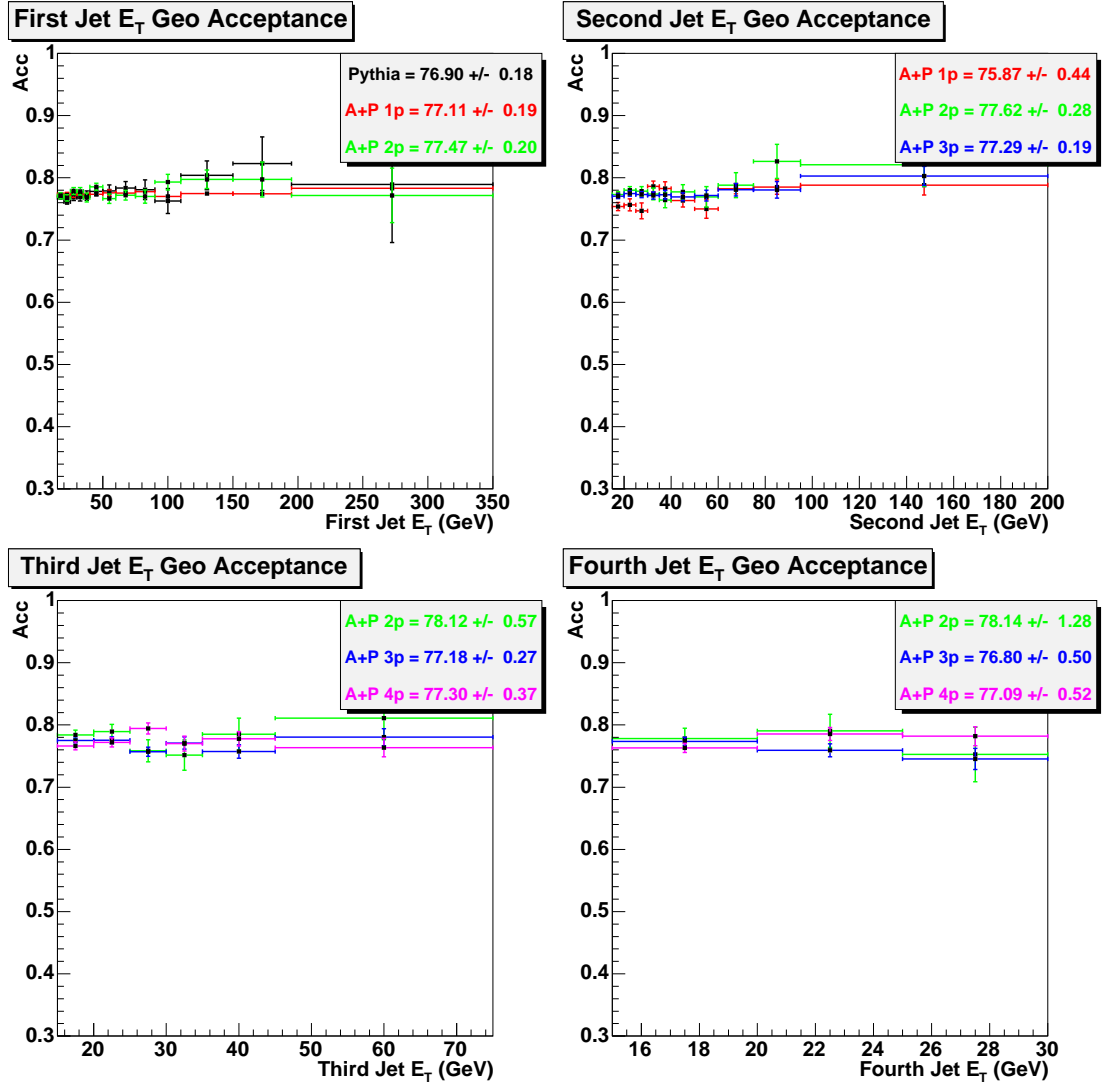
Figure 8.5: A_{geo} as a function of jet E_T for the four inclusive jet multiplicities considered. The results for different ALPGEN + PYTHIA (A+P) n parton samples are shown. The inclusive efficiency of each sample is shown in the legend.

A more quantitative approach to estimating the potential error in using simulated Monte Carlo to estimate ID efficiencies can be made by directly comparing to ID efficiency measurements made using a Z→ ee data sample. In the following Section 8.4.1 we describe how the ID efficiency is measured using Z→ ee data events, and then go on to make the comparison with W→ ev Monte Carlo in Section 8.4.2.

8.4.1   Electron ID Efficiency from Z→ ee Data

A Z→ ee data sample is selected from the same high P_T electron dataset as used in the signal analysis. This is done by requiring at least two "loose" electrons which pass the following cuts that are a subset of the W selection requirements:

- Central ($|\eta_D^{CEM}| \leq 1.1$)
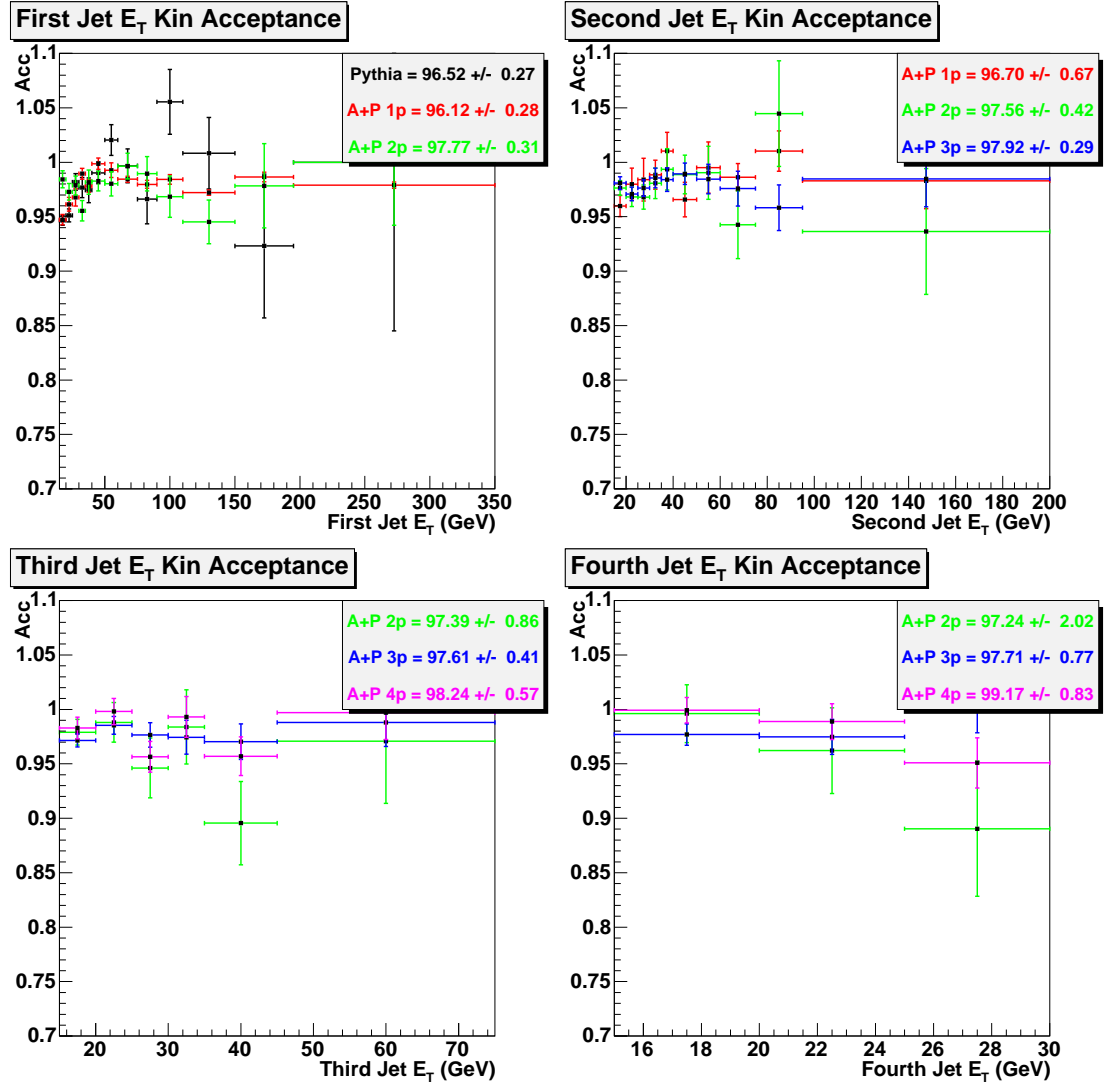
Figure 8.6: A_{kin} as a function of jet E_T for the four inclusive jet multiplicities considered. The results for different ALPGEN + PYTHIA (A+P) n parton samples are shown.

- Beam Constrained $|Z_0| < 60$cm

- Fiducial Region

- $E_T > 20$ GeV

- $P_T > 10$ GeV

In addition the electrons must together form an invariant mass within the Z mass window of {76,106} GeV, have opposite charge and at least one of the two electrons in the event must pass all of the baseline high P_T electron selection criteria listed in Table 6.1. This is the "tight" electron, and the other "loose" leg can be thought of as the "probe" electron. The method hinges on the assumption that all the events in this sample are Z→ ee events, and thus that the "probe" electron is a Z decay
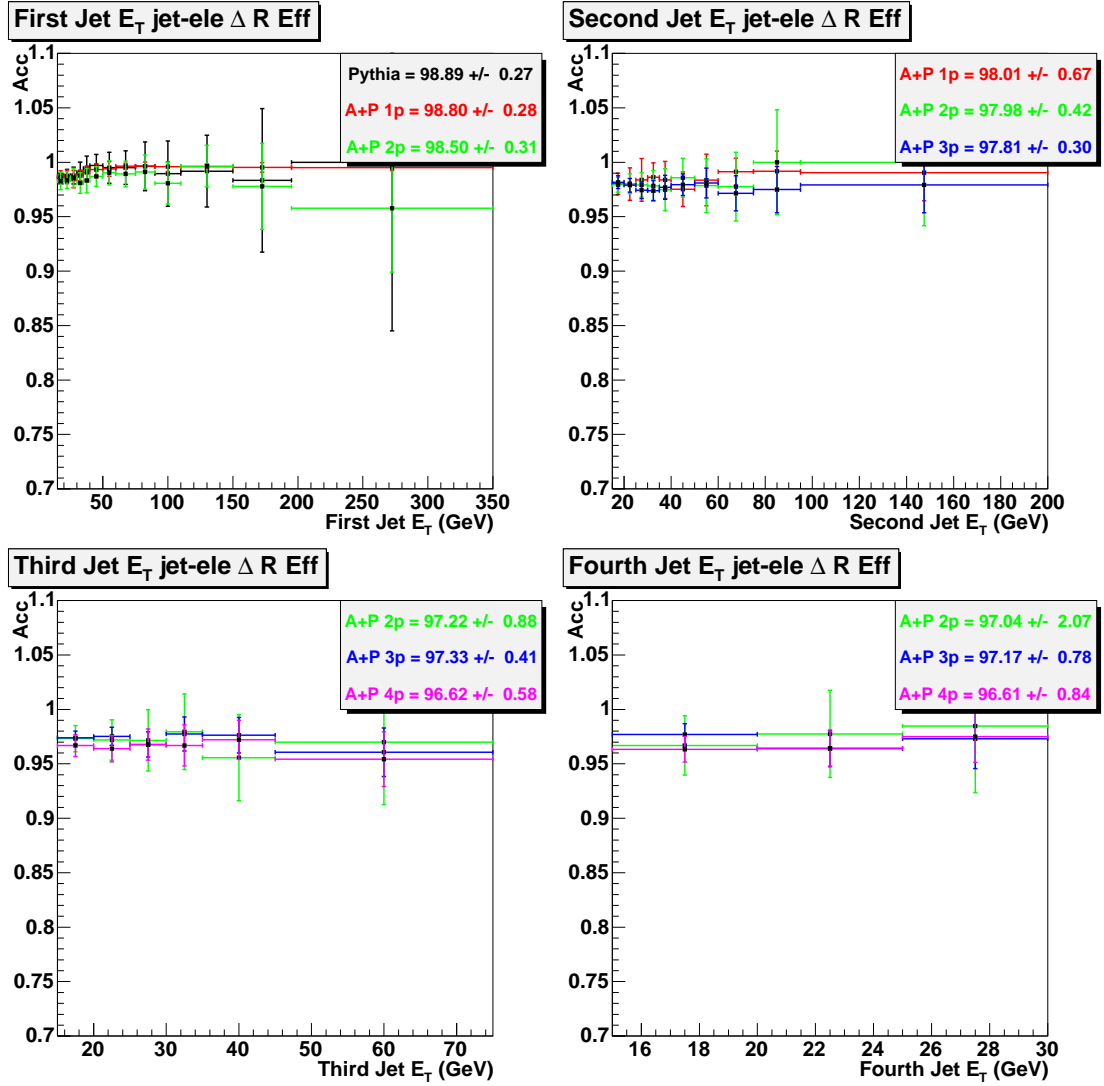
Figure 8.7: $A_{je\Delta R}$ as a function of jet E_T for the four inclusive jet multiplicities considered. The results for different ALPGEN + PYTHIA (A+P) n parton samples are shown.

electron, even though it has not yet been identified as an electron by any tight requirements. If this assumption holds we can use the "probe" electron to test the efficiency of the ID cuts. Figure 8.11 shows the invariant mass of the electrons in our Z data sample, along with that of a simulated PYTHIA Z→ ee Monte Carlo sample if we apply the same selection. The good agreement between the data and Z→ ee Monte Carlo indicates that we have a Z sample of reasonable purity.

Figure 8.12 is a tree diagram to aid explanation of how we use the Z data sample to calculate the ID efficiency. Imagine we begin with a pure sample of Z→ ee events. We select one of the two electrons at random and test it against the tight ID requirements, it can either pass or fail. We then make the same test on the second leg. However, the constraints of the sample described above mean that if the first leg has already failed, the second leg can only pass, otherwise the event
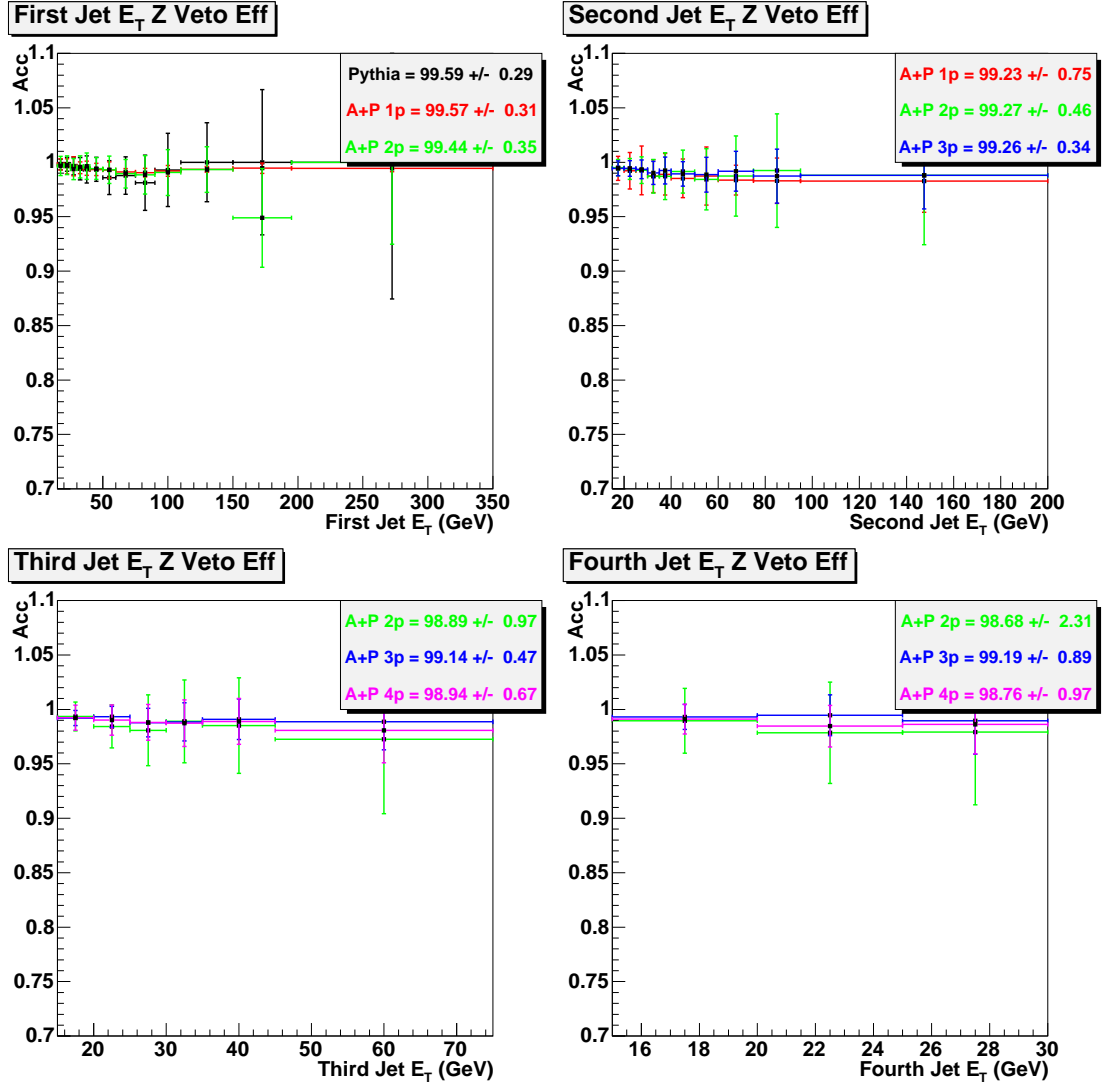
Figure 8.8: ξ_{Zrej} as a function of jet E_T for the four inclusive jet multiplicities considered. The results for different ALPGEN + PYTHIA (A+P) n parton samples are shown.

would not be in our Z sample. We thus have two classes of events in our Z sample:

- N_{TT} is the number of events containing two tight electrons.

- N_{TF} is the number of events with one tight electron and a second "loose" electron which doesn't pass the tight requirements.

The total number of events in our sample is thus $N_{\mathrm{TP}} = N_{\mathrm{TT}} + N_{\mathrm{TF}}$. The total number of Z events, N_Z, is unknown. Using these quantities the ID efficiency can be derived as follows. Label each electron e_a or e_b. The probability to observe two tight electrons is given by:

$$P_{\mathrm{TT}} = \xi_{\mathrm{ID},e_a} \times \xi_{\mathrm{ID},e_b} = \xi_{\mathrm{ID}}^2 \qquad (8.11)$$

and the probability to observe one tight electron and one which fails the electron ID cuts is given by:

$$P_{TF} = [\xi_{ID,e_a} \times (1 - \xi_{ID,e_b})] + [\xi_{ID,e_b} \times (1 - \xi_{ID,e_a})] = 2\xi_{ID}(1 - \xi_{ID}) \qquad (8.12)$$

Therefore:

$$N_{TT} = \xi_{ID}^2 N_Z \qquad (8.13)$$

$$N_{TF} = 2\xi_{ID}(1 - \xi_{ID})N_Z \qquad (8.14)$$

Eliminating N_Z in the above equations we get for ξ_{ID}:

$$\xi_{ID} = \frac{2N_{TT}}{N_{TF} + 2N_{TT}} \qquad (8.15)$$

In this scheme one electron leg must always pass the full ID requirements as this is the "tight" control leg. However, the "probe" electron can be required to pass or fail any combination of the tight selection cuts in order to probe the efficiencies of individual cuts and their correlations.

This procedure is complicated somewhat by the potential breakdown in the key assumption; that all events in the selected Z sample are actually Z$\to$ ee events. There will be some contamination from W$\to$ eυ + jet events, which can mean that the "probe" electron is not actually an electron at all. One can attempt to account for this background by looking at the number of same-sign events that are present in the sample. To first order the number of W+Jets events in the opposite-sign sample will be equal to the number of same-sign events, assuming that a jet from a W+Jet event is equally likely to be opposite or same-sign. N_{TT} and N_{TF} would then be corrected as follows:

$$N_{TT}^{corr} = N_{TT}^{OS} - N_{TT}^{SS} \qquad (8.16)$$

$$N_{TF}^{corr} = N_{TF}^{OS} - N_{TF}^{SS} \qquad (8.17)$$

Where N_{TT}^{OS} is the number of tight-tight opposite-sign events, and N_{TT}^{SS} is the number of tight-tight same-sign events. There are two corrections to this approximation. Firstly, it has been determined that the same-sign fraction in W + 1 jet fakes of zero-jet Z$\to$ ee events is actually 40% [81]. We assume that this fraction remains constant with increasing jet multiplicity. Secondly, a pure Z$\to$ ee sample will contain a small fraction of same-sign events due to the presence of tridents. This can be

accounted for by correctly normalising the number of same-sign events observed in Z→ ee Monte Carlo. Taking these effects into account we get:

$$N_{TT}^{corr} = N_{TT}^{OS} - \frac{0.6}{0.4} \times \left[N_{TT}^{SS} - MC_{TT}^{SS} \cdot \frac{N_{TT}^{total}}{MC_{TT}^{total}} \right] \tag{8.18}$$

$$N_{TF}^{corr} = N_{TF}^{OS} - \frac{0.6}{0.4} \times \left[N_{TF}^{SS} - MC_{TF}^{SS} \cdot \frac{N_{TF}^{total}}{MC_{TF}^{total}} \right] \tag{8.19}$$

Where MC_{TT}^{SS} is the number of same-sign tight-tight events from Z→ ee Monte Carlo, MC_{TT}^{total} is the total number of tight-tight events ($MC_{TT}^{total} = MC_{TT}^{OS} + MC_{TT}^{SS}$) and N_{TT}^{total} is the total number of tight-tight events in the Z data sample.

In order to form a differential ID efficiency we simply perform the calculation above in each bin of the differential distribution. In the case where $N_{TT}^{SS} = 0$ we do not make any trident correction i.e. $N_{TT}^{corr} = N_{TT}^{OS}$.

8.4.2 Comparison of Monte Carlo and Data ξ_{ID}

The method described above is used to compute the ID efficiency using simulated Z→ ee Monte Carlo, the only difference being that we don't make the same-sign background correction. The ID efficiency in W→ eυ Monte Carlo is found using Equation 8.9, as was described in Section 8.1.1.

Inclusive ID Efficiency Results

Table B.7 compares the ID efficiency obtained using the Z data sample and a simulated PYTHIA Z→ ee Monte Carlo sample in each inclusive jet multiplicity bin. When counting jets in this case we remove both the tight and loose electrons from the list of jets.

One can see that the scale factor, the ratio of data to MC ID efficiencies, is consistent with unity for all jet multiplicity bins with the exception of the ≥ 2 jet bin. This is demonstrated graphically in Figure 8.13(a). The reason for the discrepancy in the ≥ 2 jet bin is thought to be a source of background that is not sufficiently corrected for by the same-sign corrections described above. Evidence for this hypothesis is presented in Figure 8.14. This shows the ID efficiency as a function of the mass window width for data and Z→ ee Monte Carlo. Reducing the window width should reduce the background present in the data. We observe that narrowing the window has little impact on either the data or MC efficiency in the ≥ 0 jet and ≥ 1 jet bins, but in the ≥ 2 jet bin there is a clear trend of the data efficiency increasing to be in agreement with the MC as the window is reduced. Figure 8.13(b) shows the scale factor as a function of jet multiplicity for a reduced mass window of {88,94} GeV.

Differential ID Efficiency Results

Due to the limited statistics of the Z$\rightarrow$ ee data sample we can only make meaningful comparisons of the differential ID efficiency in the ≥ 1 jet and ≥ 2 jet bins. Here we use the mass window {88,94} GeV in order to exclude backgrounds.

Figure 8.15 shows the differential ID efficiency in the first and second jet E_T distributions, obtained using the Z$\rightarrow$ ee data sample and Alpgen + PYTHIA n parton W$\rightarrow$ eυ Monte Carlo samples. One can see that, within the statistics available, all the Monte Carlo samples do a reasonable job of reproducing the data differential ξ_{ID} , at least to within 5% (absolute). Good agreement is observed between the $n-1$ and n parton samples, but the $n+1$ parton sample appears to be systematically shifted by order -5% (relative). This is thought to be due to the increased jet activity in the $n+1$ parton sample which, despite the presence of a jet-electron ΔR cut, has a small impact on the ID efficiency.

8.5 Acceptance Systematics

We consider three possible sources of systematic error on the total acceptance $\mathcal{A}$:

- The uncertainty on using detector simulated Monte Carlo to estimate an electron ID efficiency.

- Theoretical model dependence of the acceptance result.

- The statistics of the W$\rightarrow$ eυ Monte Carlo sample used in each jet E_T ,ΔR$_{jj}$ and M$_{jj}$ bin.

In Section 8.4.2 we demonstrate that the Data/Theory scale factor for the electron ID efficiency is consistent with unity when comparing in inclusive jet multiplicity bins. On examining the differential ID efficiency, we observe no evidence to suggest that the scale factor is strongly dependent on jet properties. However, we do see potential theoretical model dependance of the ID efficiency, at the level of 5%. A similar level of model dependance is observed in the kinematic acceptance factor.

In light of these studies we propagate an absolute error on the acceptance of 5% as a conservative estimate of the impact of the scale factor and model dependance systematics described above. This will translate directly into a 5% global error on the differential cross-section. In addition the statistical error on the acceptance resulting from the Monte Carlo statistics is propagated into the differential cross-section, which will vary from bin to bin. In Section 10.3 we discuss the impact of the acceptance systematics on the cross-section.

8.6   Global Efficiency Factors

There are two further sources of inefficiency which we thus far have not considered:

- The dataset is formed using the high P_T electron trigger path described in Section 6.1.5. The efficiency of this trigger to accept high P_T electrons can be measured by using a W candidate sample selected from a looser trigger path and observing for what fraction of these events the high P_T electron trigger fired. The trigger efficiency is found to be $0.9620^{+0.0063}_{-0.0066}$ [93], with the small inefficiency arising from the XFT track reconstruction.

- As part of the W event selection criteria we require that the position of the reconstructed primary vertex be $|Z_0| < 60$cm. However, the full luminous region in which $p\bar{p}$ collisions can occur extends beyond this, and the luminosity as measured by the CLC relates to the full luminous region. Studies using minimum bias samples have estimated the full extent of the region in which $p\bar{p}$ collisions occur, and thus the efficiency of this vertex requirement. It is found to be 0.955 ± 0.003 [114].

These factors are applied uniformly across our differential distributions, and are thus assumed to be independent of the properties of the jets in our events. The uncertainties associated with these factors are negligible compared to the 5% systmatic already applied to the acceptance factor, and can thus be safely ignored.

(a) Had/EM

(b) E/P

(c) CES ΔZ

(d) $Q_e \cdot$ CESΔX
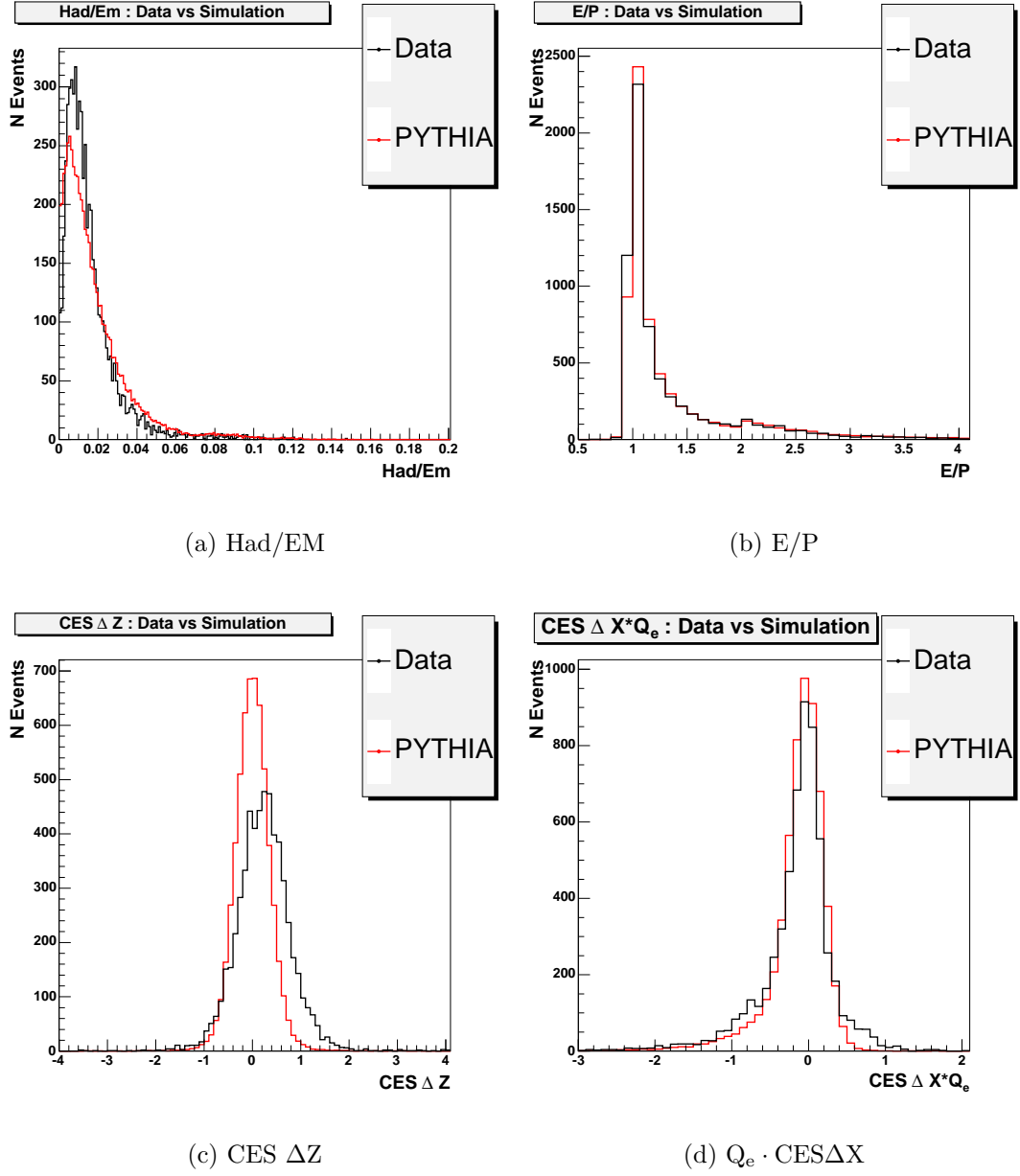
Figure 8.9: Comparison of electron identification variables in data and detector simulated Z$\rightarrow$ ee Monte Carlo. In each distribution all of the tight electron selection requirements have been imposed with the exception of the cut under consideration, and the Monte Carlo is normalised to the data.
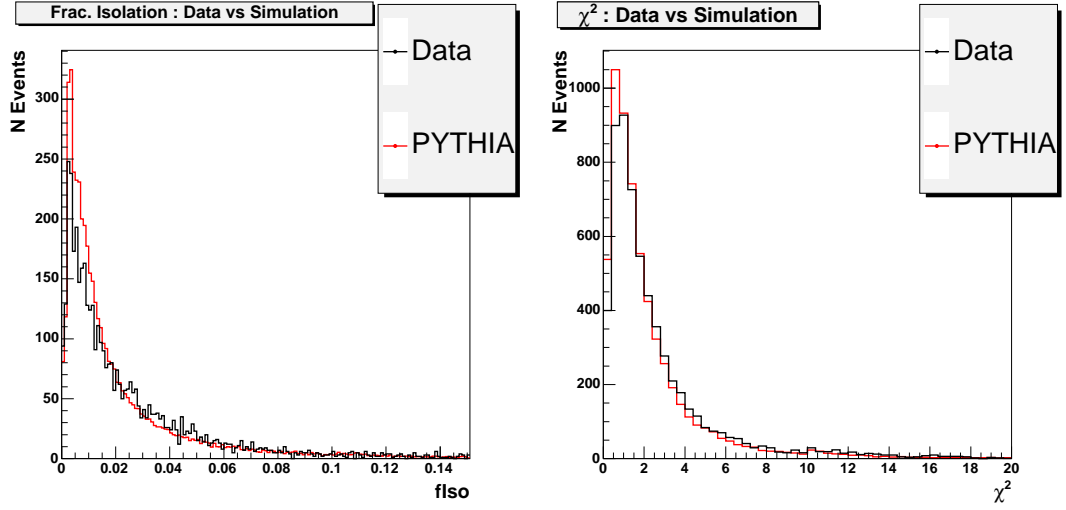
(a) Fractional Isolation

(b) χ^2
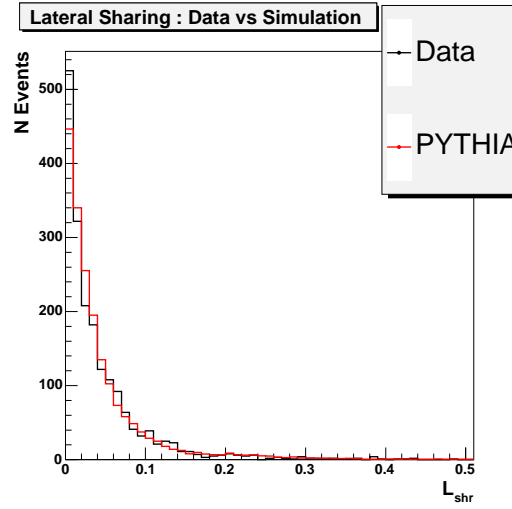
(c) L_{shr}

Figure 8.10: Comparison of electron identification variables in data and detector simulated Z$\rightarrow$ ee Monte Carlo. In each distribution all of the tight electron selection requirements have been imposed with the exception of the cut under consideration, and the Monte Carlo is normalised to the data.
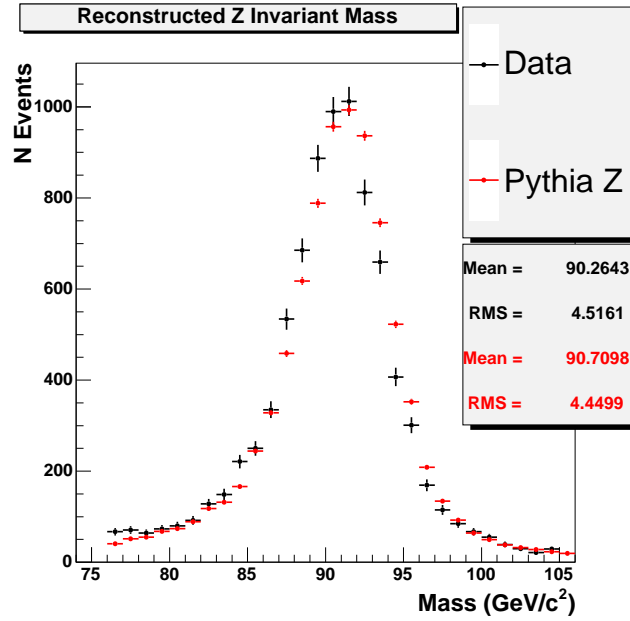
Figure 8.11: Invariant Mass of "tight" and "probe" electrons in our Z data sample and a PYTHIA simulated Z$\rightarrow$ ee Monte Carlo sample.
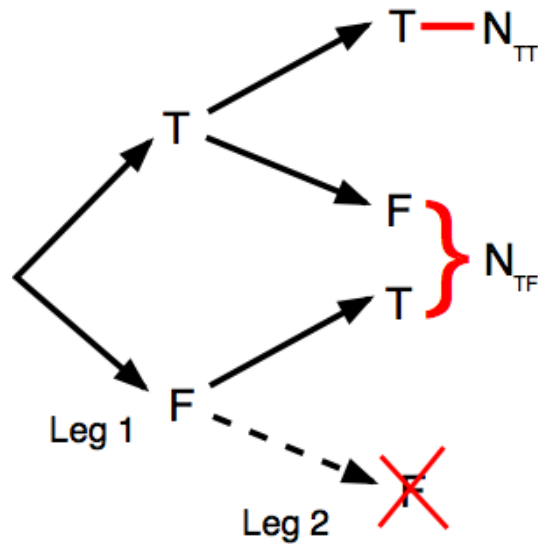
Figure 8.12: Tree diagram to illustrate the calculation of electron ID efficiency from the Z$\rightarrow$ ee data sample. The requirement that our sample contains at least one tight electron means that if the first leg fails the second leg can only pass otherwise the event does not enter our Z sample.

(a) Mass window (76,106)   (b) Mass window (88,94)
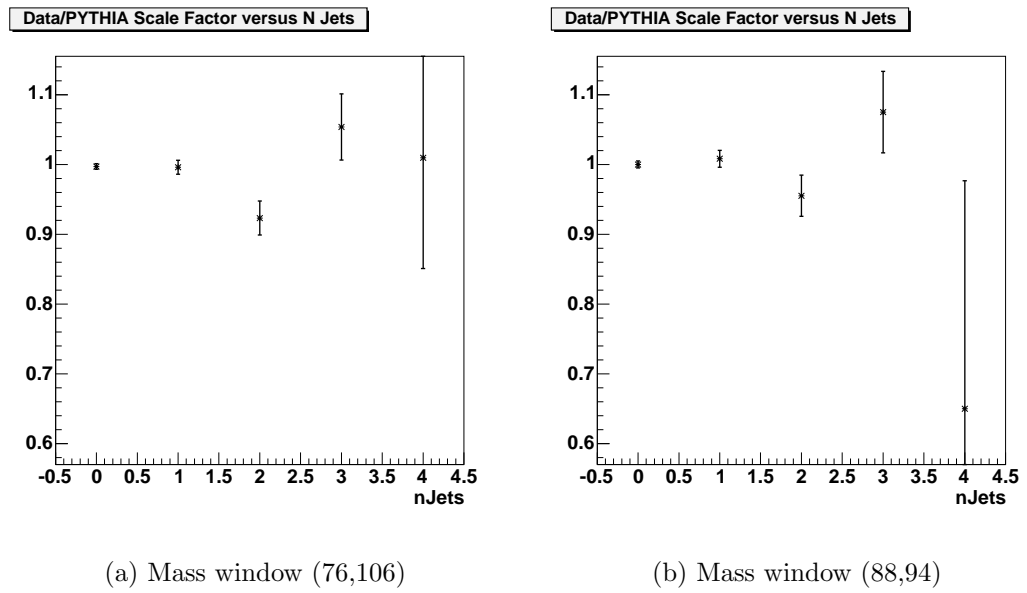
Figure 8.13: Data/PYTHIA Scale Factor as a function of the number of jets.

(a) ≥ 0 jet

(b) ≥ 1 jet

(c) ≥ 2 jet
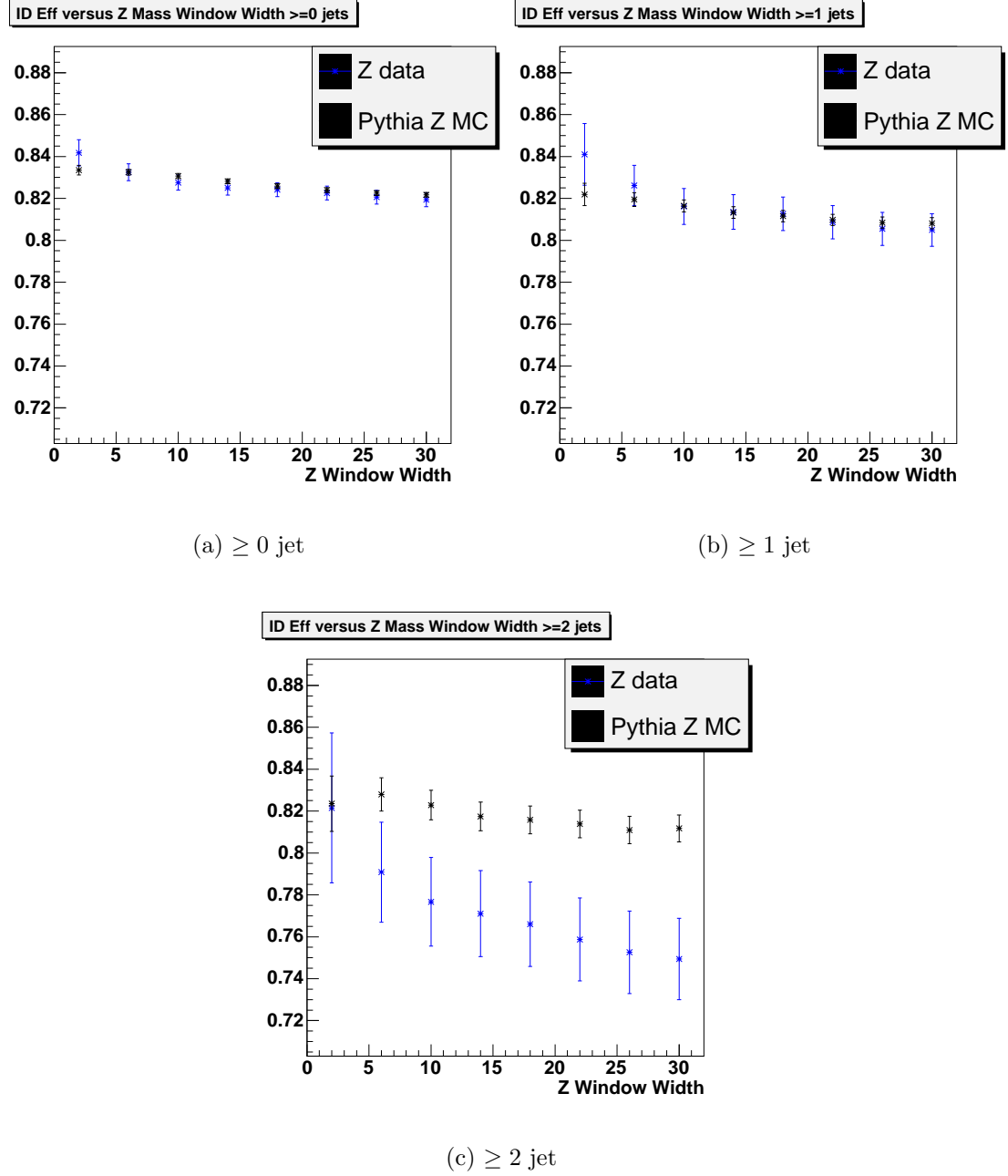
Figure 8.14: Data and Z→ ee Monte Carlo ID Efficiency as a function of the Z mass window for ≥ 0 jet ,≥ 1 jet and ≥ 2 jet events. The mass window is always centred on 91GeV, thus a window width of 30 GeV corresponds to {76,106}.

(a) ≥ 1 jet   (b) ≥ 2 jet
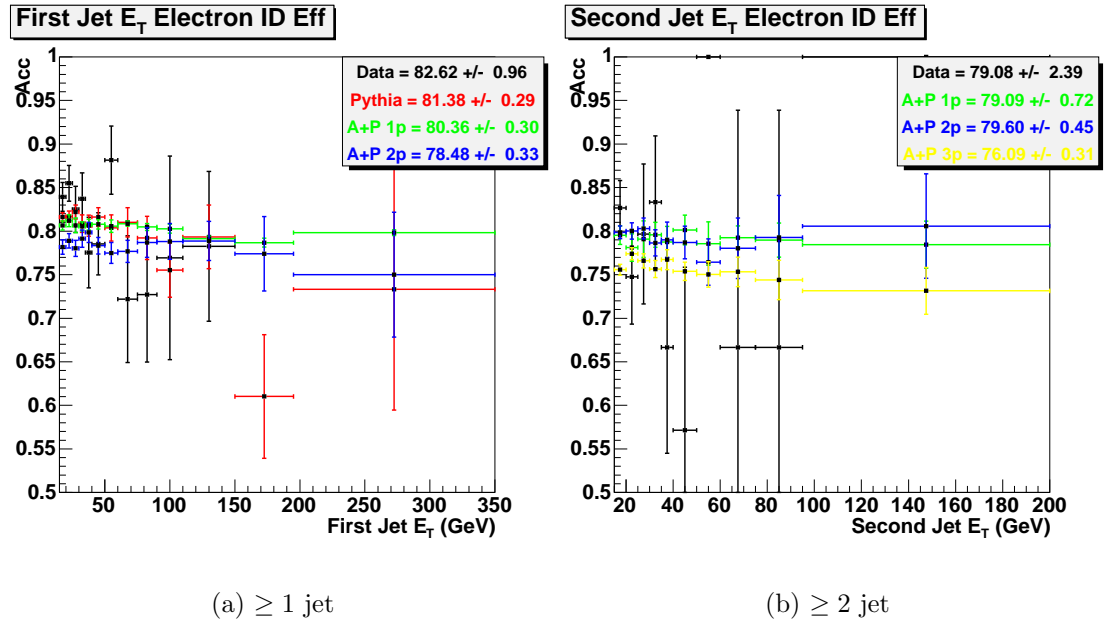
Figure 8.15: Comparison of the differential ID efficiency obtained with the Z→ ee data sample and Alpgen + PYTHIA simulated Monte Carlo samples in the first and second jet E_T diistributions.

Chapter 9

The Unsmearing Correction

9.1 The Need for Unsmearing

In Section 6.2 we described how we apply an absolute correction to our jets that should correct the calorimeter measured energy such that it is on average equal to the energy of the hadrons contained within the jet. These corrections are determined from detector simulated PYTHIA dijet Monte Carlo samples by examining the correlation in E_T between hadron-level and calorimeter-level jets (see Section 6.2.2 for details) and should be to first order process independent.

In a detector simulated Monte Carlo sample we can cluster jets using the same JetClu algorithm at both the hadron-level and the calorimeter-level. Figure 9.1(a) is a plot of the raw (uncorrected) calorimeter-level jet energy against the hadron-level jet energy for a detector simulated $W \rightarrow e\upsilon + 1$ parton ALPGEN+PYTHIA sample. On the left (Figure 9.1) is the two-dimensional scatter plot, and on the right (Figure 9.1(b)) is a Y-axis profile plot of the same distribution, where a linear fit has been made to the profile points. For each bin in X the profile plot plots the mean of the distribution in Y, with an error equal to the error on the mean. As one might expect, the linear fit has a slope which is substantially different from unity, indicating the systematic offset of the calorimeter energy scale.

Figure 9.2 shows the same distributions, but in this case the calorimeter-level jet energy is corrected using the relative and absolute corrections as in the analysis. One can see that, as expected, the linear fit to the mean of the corrected jet energy has a slope close to unity. This confirms that the absolute corrections are performing as we expect. However, what we can also notice from this distribution is that there is a substantial degree of smearing as we go from the hadron to calorimeter-level. Figure 9.3 shows the calorimeter resolution (CAL-HAD/HAD) for a range of slices in hadron jet E_T , where the calorimeter energy is corrected (Figure 9.3(a)) and where it is left uncorrected (Figure 9.3(b)). The effect of the absolute corrections is to centre these distributions close to zero, but they have little impact on the

(a) Scatter plot

(b) Y-axis profile plot with linear fit.
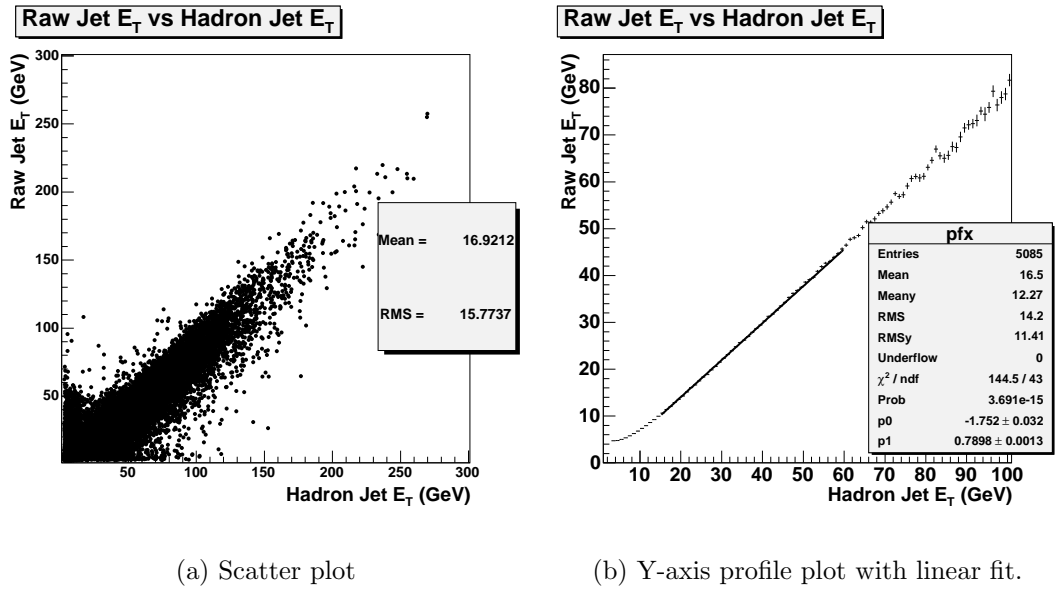
Figure 9.1: A 2-D histogram of the raw (uncorrected) calorimeter-level jet energy plotted against the hadron-level jet energy for a detector simulated $W \rightarrow e\upsilon + 1$ parton ALPGEN+PYTHIA sample.

(a) Scatter plot

(b) Y-axis profile plot with linear fit.
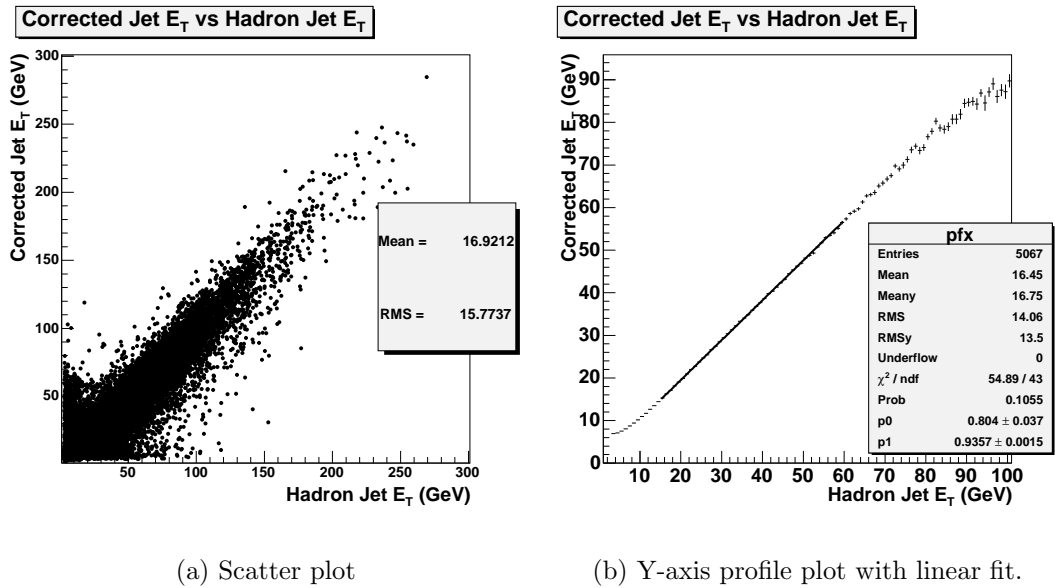
Figure 9.2: A 2-D histogram of the corrected calorimeter-level jet energy plotted against the hadron-level jet energy for a detector simulated $W \rightarrow e\upsilon + 1$ parton ALPGEN+PYTHIA sample.

resolution. The resolution ranges from 15-25%, decreasing with increasing jet E_T .

In order to illustrate how the resolution of the calorimeter can affect the measurement we have used a $W \rightarrow e\upsilon$ Monte Carlo sample and implemented a simple model of a detector; in this model the calorimeter-level jet energy is the hadron-level
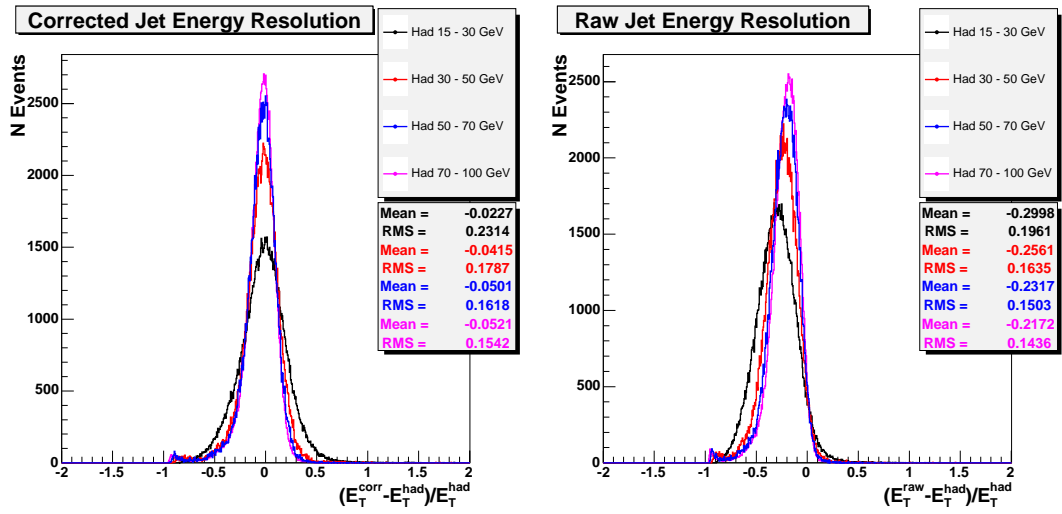
(a) Corrected Jet Energy Resolution (CAL-HAD)/HAD.

(b) Raw (uncorrected) Jet Energy Resolution (CAL-HAD)/HAD.

Figure 9.3: Jet energy resolution for a range of slices in hadron-level jet E_T .

energy smeared by a Gaussian with the following properties:

$$
\begin{aligned}
\mu &= \text{HAD}_{E_T} \\
\sigma &= 2 \times \left[(0.1 \cdot \text{HAD}_{E_T}) + 1 \right]
\end{aligned}
\tag{9.1}
$$

This smearing function is an approximation of the calorimeter resolution, with an additional factor of two to exaggerate the smearing effect for this demonstration. Figure 9.4(a) shows the overall resolution (CAL-HAD/HAD for every jet) produced by this smearing, and Figure 9.4(b) shows the CAL vs HAD correlation obtained. These distributions show that we have a sample which behaves as if "perfect" absolute corrections have been applied at the calorimeter-level, leaving the effects of the resolution only. Figure 9.4(c) shows the hadron-level lead jet E_T distribution alongside the smeared distribution. One can clearly see the impact of the smearing on the spectrum. The smearing causes a migration of events between jet E_T bins, and thus the measured spectrum is no longer that of the underlying hadron jets. Thus, even if the absolute corrections are perfect the spectrum measured in the calorimeter will not be that of the underlying hadrons, and you will not have made a truly hadron-level measurement. However, in this case, where we have knowledge of the spectrum at the calorimeter-level and the hadron-level, it is clear how we can take account of the smearing effect; simply take the ratio of the hadron and calorimeter distributions in each bin and apply this "unsmearing factor" to the calorimeter distribution:
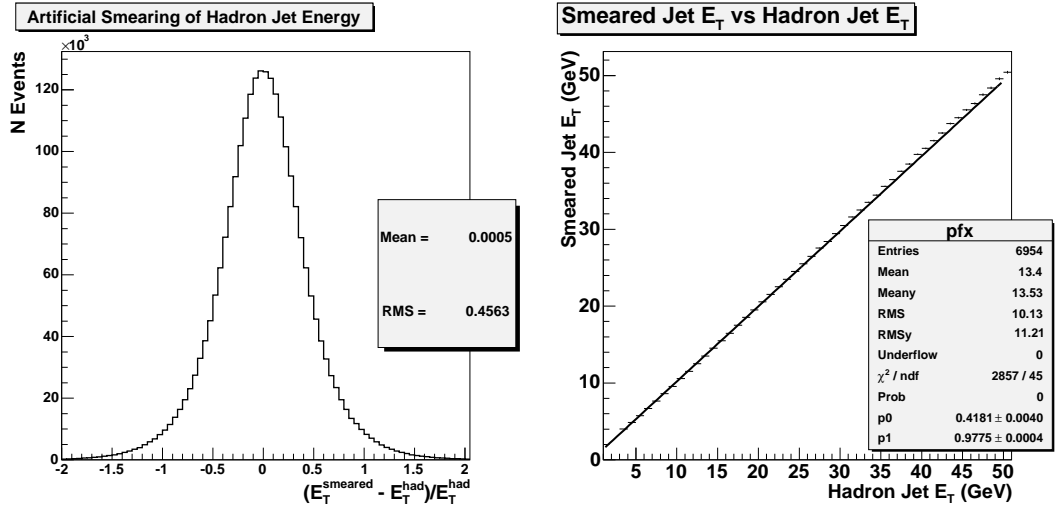
(a) Gaussian Smearing.

(b) SMEAR-HAD correlation profile plot.
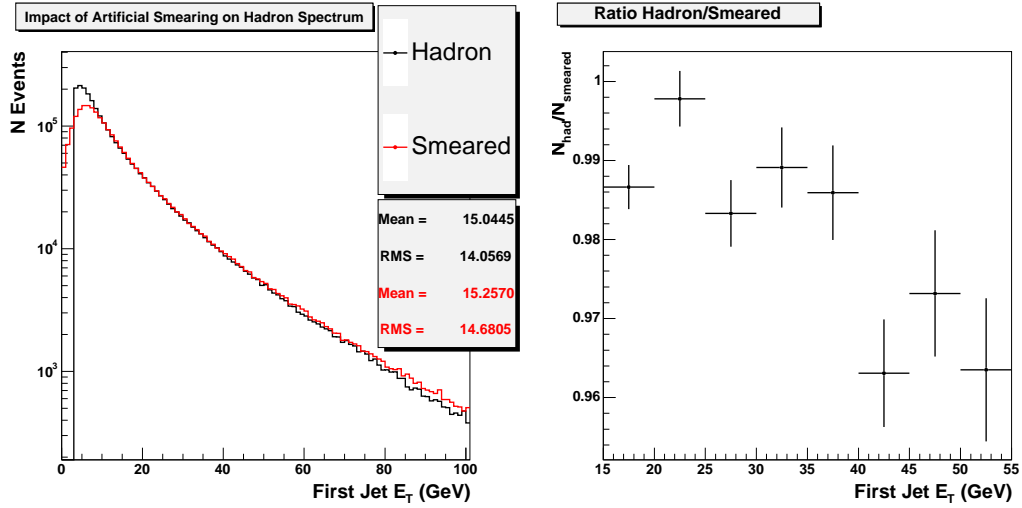
(c) Impact of smearing on spectra.

(d) Ratio Hadron/Smeared.

Figure 9.4: Distributions to investigate the impact of calorimeter resolution, where we have implemented an artificial Gaussian smearing of hadron-level jet energies.

$$\mathcal{S}_i = \frac{N_i^{had}}{N_i^{cal}} \tag{9.2}$$

where $\mathcal{S}_i$ is the unsmearing factor in the ith bin and N_i^{cal} and N_i^{had} are the number of events in the ith bin of the calorimeter-level and hadron-level distributions respectively. In general the unsmearing factor will depend on the size of the detector resolution as a function of jet E_T, and the shape of the underlying hadron distribution. The unsmearing factors for this simple model are shown in Figure 9.4(d).

(a) Lead jet E_T in ≥ 1 jet

(b) Second jet E_T in ≥ 2 jet

(c) Third jet E_T in ≥ 3 jet

(d) Fourth jet E_T in ≥ 4 jet
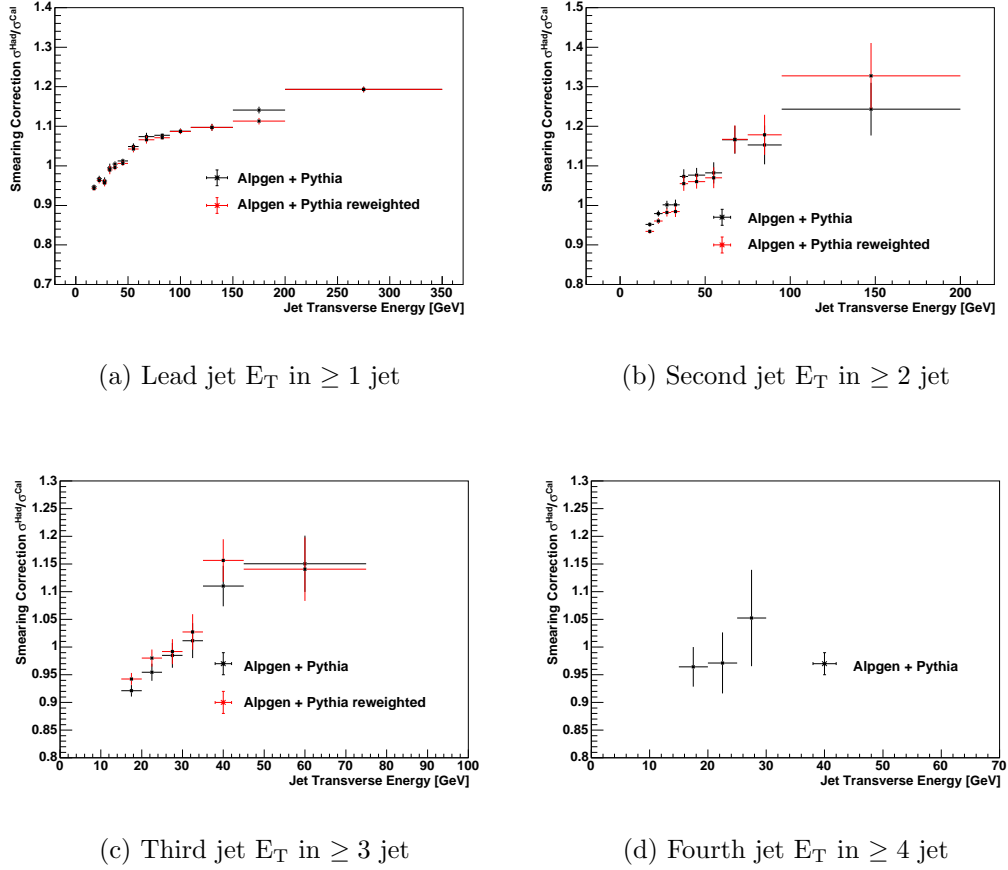
Figure 9.5: Unsmearing factors, shown before and after reweighting Monte Carlo on data cross-section.

9.2 Unsmearing the Data

A detector simulated $W \rightarrow e\upsilon + n$ parton Monte Carlo sample can be used to correct the measured cross-section for the resolution effects of the CDF calorimeter in exactly the same way. Figure 9.5 shows the unsmearing factors for the first,second,third and fourth jet E_T distributions using simulated ALPGEN+PYTHIA $W \rightarrow e\upsilon + n$ parton Monte Carlo samples, where the n parton sample is used in the $\geq n$ jet bin. Here the calorimeter-level jet E_T is corrected using absolute and relative corrections, and the $\eta < 2.0$ selection requirement is applied to both the calorimeter and hadron-level jets.

However, it has already been noted that the unsmearing correction will depend on the shape of the underlying hadron distribution. How do we know that the $W \rightarrow e\upsilon$ Monte Carlo samples describe this accurately? Put another way, how can we ensure that we are not introducing theoretical model dependence with this correction? The answer is to reweight the calorimeter-level jet E_T distribution in the Monte Carlo on the data measured cross-section. Since there is a one-to-one cor-

respondence between the calorimeter and hadron jets, reweighting the calorimeter-level distribution requires that the hadron-level distribution changes also, as determined by the 2-D CAL-HAD correlation plot of Figure 9.2(a). Figure 9.5 shows the unsmearing factors before any reweighting is done (black) and after several iterations of the reweighting procedure (red). One iteration of reweighting proceeds as follows:

1. Correct the data measured cross-section by the initial unsmearing factors of the Monte Carlo (the black curve in Figure 9.5).

2. Take the ratio of the unsmearing corrected data and the hadron-level Monte Carlo distributions to find "reweighting factors".

3. Use these to reweight the hadron-level Monte Carlo distribution, using the 2-D CAL-HAD correlation to correspondingly reweight the calorimeter-level Monte Carlo distribution.

4. One can now define new unsmearing factors using the reweighted Monte Carlo calorimeter-level and hadron-level distributions. These can be used in the next iteration of reweighting, or taken as the final unsmearing factors.

Figure 9.6 shows the ratio of the unsmearing corrected data cross-section and the hadron-level Monte Carlo cross-section. The black curve is the initial ratio before any reweighting is done (i.e. the reweighting factors of the first iteration), and the red curve after two reweighting iterations. One can see that the initial agreement between the data and Monte Carlo shapes is already very reasonable within the data-dominated statistical uncertainties, and that reweighting can improve this marginally. It was found that more than two iterations of the reweighting procedure made very little difference to the agreement. One can see from Figure 9.5 that the impact of reweighting on the unsmearing factors is insignificant. In the case of the fourth jet E_T distribution no reweighting is performed since one cannot reasonably define a shape with only three data points. However, one can see that (within the data-dominated statistical errors) the initial agreement is already reasonable, and given the insignificance of the impact of reweighting on the unsmearing factors in the other distributions we believe that successful reweighting would have negligible impact on the fourth E_T unsmearing factors.

Figure 9.7(a) shows the unsmearing correction for the first-second jet invariant mass distribution. One can see that the unsmearing correction on the first three bins at low invariant mass is very large. This may not be particularly surprising. The non-overlapping cones restriction of JetClu jet clustering, combined with the $E_T > 15\text{GeV}$ requirement, means that the first-second jet invariant mass distribution has a very sudden "turn on" at low M_{jj} (see the distribution of candidates in

(a) Lead jet E_T in ≥ 1 jet

(b) Second jet E_T in ≥ 2 jet

(c) Third jet E_T in ≥ 3 jet

(d) Fourth jet E_T in ≥ 4 jet
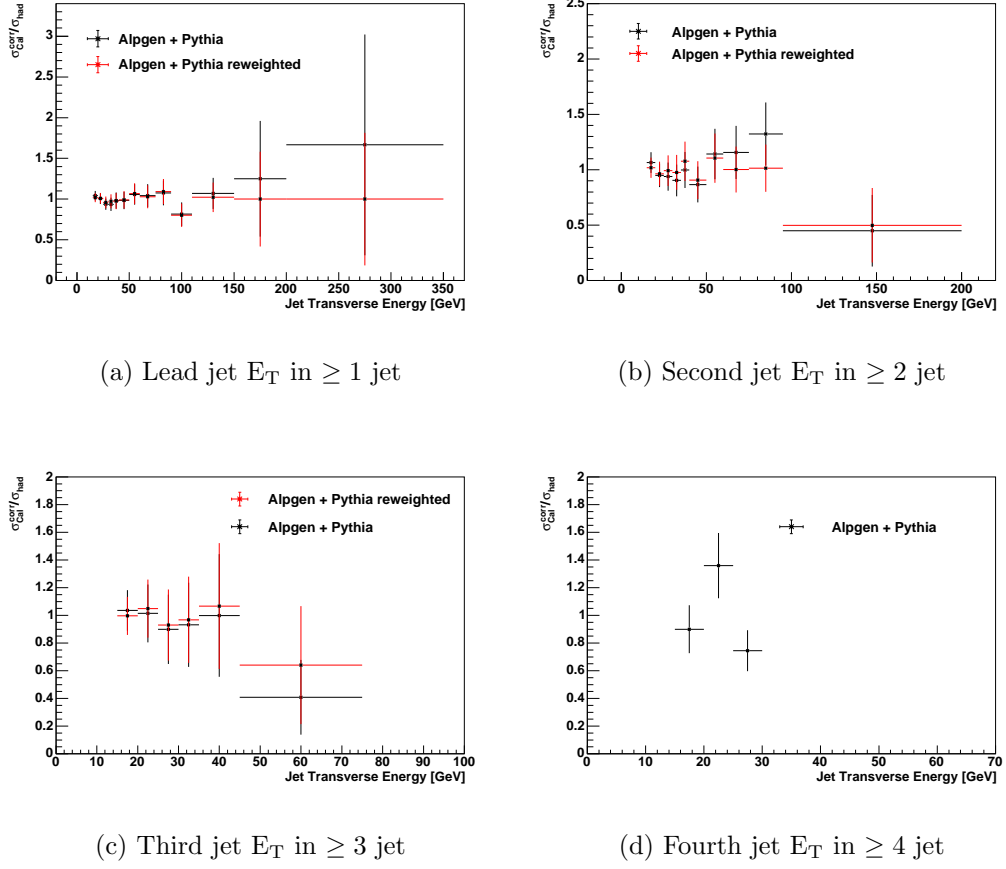
Figure 9.6: The ratio of the unsmearing corrected data cross-section and the hadron-level Monte Carlo cross-section after zero iterations (black) and several iterations (red) of the reweighting procedure.

Figure 6.12(b)), and hence the impact of the invariant mass resolution can be dramatic in this region. However, Figure 9.7(b) shows that in this crucial region where the unsmearing correction is largest we cannot demonstrate that the Monte Carlo is providing a reasonable description of the shape of the data. For this reason the invariant mass differential cross-section is cut off below $24\mathrm{GeV/c^2}$. Above this cut the invariant mass unsmearing correction is under control.

The impact of the calorimeter energy resolution on the first-second ΔR_{jj} distribution is limited to the migration of jets across the $E_T > 15\mathrm{GeV}$ cut, which both jets are required to pass in order to contribute to the distribution. This unsmearing factor is found to be 0.942, and this is applied to each bin in the ΔR_{jj} differential cross-section.

(a) Unsmearing Factor   (b) Data/Theory
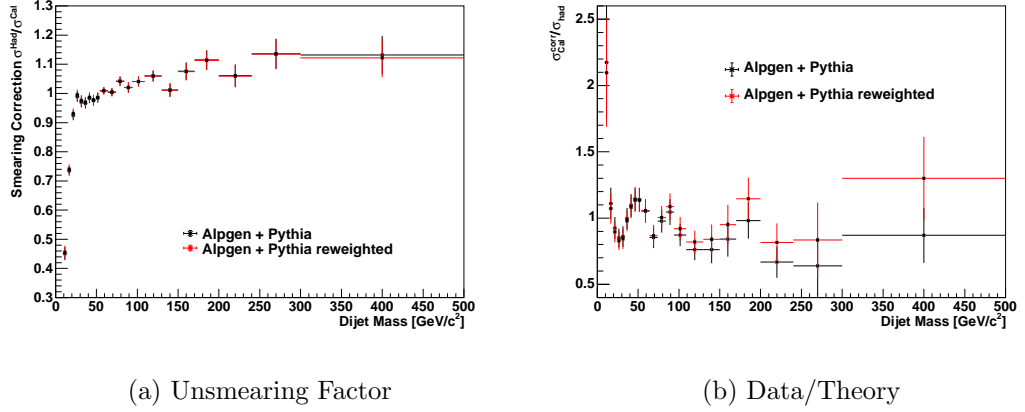
Figure 9.7: On the left the unsmearing factor and on the right the ratio of the unsmearing corrected data cross-section and the hadron-level Monte Carlo cross-section for the first-second jet invariant mass distribution.

9.3   Unsmearing Systematics

The absolute corrections have a systematic error of 2-3% due to uncertainties in the detector simulation of hadrons and the Monte Carlo fragmentation model (see Section 6.2.4). These same uncertainties are likely to impact the Monte Carlo description of the detector resolution, on which we are totally reliant in determining the unsmearing factors.

In order to assess the possible impact of an uncertainty in the detector resolution we artificially increased the resolution of the detector by smearing the calorimeter-level jet energies using a Gaussian with the following properties:

$$\mu \ = \mathrm{CAL}_{\mathrm{E_T}} \tag{9.3}$$
$$\sigma \ = \mathcal{K} \times \left[(0.1 \cdot \mathrm{CAL}_{\mathrm{E_T}}) + 1 \right]$$

Where $\mathcal{K} = 0.5$ will increase the resolution by $\sim 10\%$ and $\mathcal{K} = 1.12$ will increase the resolution by $\sim 50\%$, as shown in Figure 9.8. The impact on the first jet $\mathrm{E_T}$ unsmearing factor is shown in Figure 9.9. The only impact is in the high $\mathrm{E_T}$ region, where the unsmearing factor changes by around 5% (absolute) for an additional 10% on the resolution, and more dramatically for an additional 50%. However, as will be shown in Chapter 10, in the high $\mathrm{E_T}$ region the relative statistical error on the cross-section from the counting of candidate events is already at the 20-50% level, and thus this uncertainty on the unsmearing factor is not significant.

The only systematic uncertainty that is propagated into the cross-section measurement for the unsmearing factor is the error due to the Monte Carlo statistics, shown in Figure 9.5.

(a) $\mathcal{K} = 0$

(b) $\mathcal{K} = 0.5$

(c) $\mathcal{K} = 1.12$
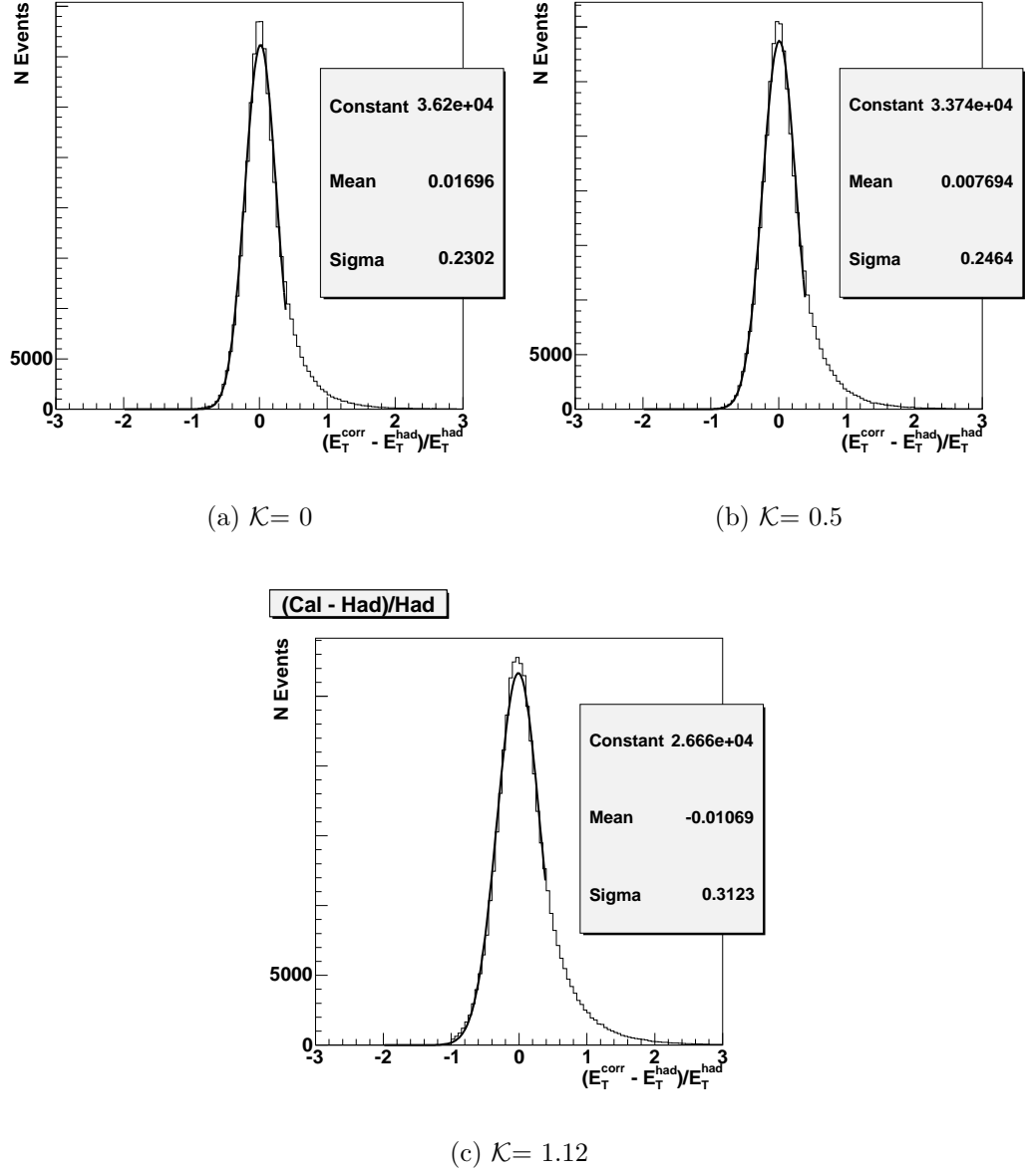
Figure 9.8: Artificially worsening the calorimeter resolution by smearing the calorimeter energy with a Gaussian $\mu = \mathrm{CAL}_{\mathrm{E_T}}$ and $\sigma = \mathcal{K} \times \left[(0.1 \cdot \mathrm{CAL}_{\mathrm{E_T}}) + 1\right]$.
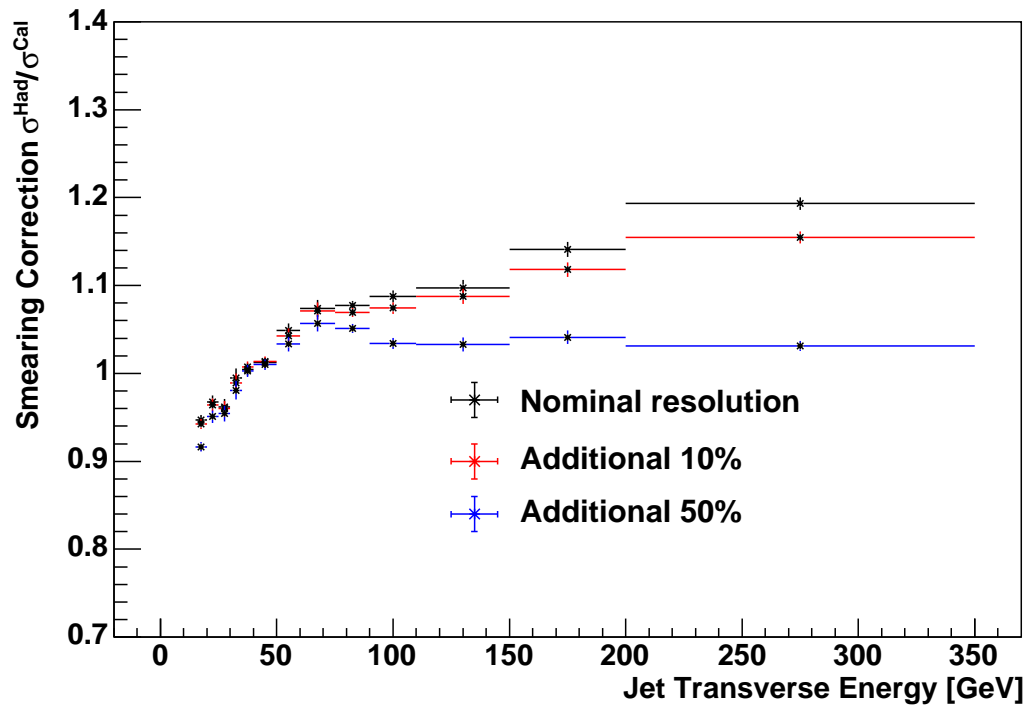
Figure 9.9: Unsmearing factor in the first jet E_T distribution with additional artificial smearing of calorimeter energy.

Chapter 10

The Cross-Section Results

In Chapters 6,7, 8 and 9 we described how the differential candidate, background, acceptance and unsmearing quantities are determined. In this chapter we describe how this information is used to assemble the following W$\rightarrow$ eυ + $\geq n$ jet differential cross-sections:

$$\frac{d\sigma_{\geq 1j}}{dE_T^{1stj}}; \quad \frac{d\sigma_{\geq 2j}}{dE_T^{2ndj}}; \quad \frac{d\sigma_{\geq 3j}}{dE_T^{3rdj}}; \quad \frac{d\sigma_{\geq 4j}}{dE_T^{4thj}}; \quad \frac{d\sigma_{\geq 2j}}{d\Delta R_{jj}}; \quad \frac{d\sigma_{\geq 2j}}{dM_{jj}};$$

We also show how this information can be "integrated" to form inclusive W + $\geq n$ jet cross-section measurements with a range of minimum jet E_T thresholds.

10.1 Inclusive $p\overline{p} \rightarrow W(e\upsilon) + X$ Cross-Section

It has been shown that the overall normalisation of our differential cross-section results will be dependent on several factors that do not vary significantly with jet activity; the electron ID, trigger and vertex effiencies and luminosity. Comparing our inclusive $p\overline{p} \rightarrow W(e\upsilon) + X$ cross-section result with an independent measurement of the same quantity is a powerful cross-check of our W$\rightarrow$ eυ event selection and estimation of these global factors.

In Table 6.1 the total number of inclusive W candidate events is reported, 147008 events. From Table A.1 one can calculate the total background in the inclusive sample, giving 5371 events. In order to calculate the inclusive W cross-section one must compute the full acceptance factor without imposing the restrictions on the W decay kinematics i.e. that given by Equation 8.3. Using a PYTHIA W$\rightarrow$ eυ sample this is found to be 0.177. Combining these numbers in the standard cross-section formula (Equation 5.1) together with a luminosity of 320pb^{-1}and the trigger and vertex efficiencies detailed in Section 8.6 gives an inclusive cross-section of:

$$\sigma(\text{W} \rightarrow e\upsilon + \text{X}) = 2722 \pm 7(\text{stat.}) \pm 136(\text{sys.}) \pm 163(\text{lum.}) \text{ pb}$$

Where the systematic error is domninated by the 5% uncertainty on the acceptance factor; other systematic sources are either not relevant (jet energy scale) or negligible in the inclusive sample (background). This is in excellent agreement with the independent CDF measurement of $\sigma(\text{W} \rightarrow \text{e}\upsilon + \text{X}) = 2780 \pm 14(\text{stat.}) \pm 60(\text{sys.}) \pm 167(\text{lum.})\text{pb}$ [17], which used 72pb^{-1} of data, and the NNLO prediction of $2687 \pm 54\text{pb}$ [87].

10.2 Differential Cross-Section Results

The differential cross-section with respect to a variable x, $\mathrm{d}\sigma/\mathrm{d}x$, is calculated by evaluating the following equation in each bin i (with bin width Δx_i) of the variable x:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x}\bigg|_i = \frac{C_i - B_i}{\mathcal{A}_i \cdot \mathcal{L}} \times \mathcal{S}_i \times \frac{1}{\Delta x_i} \tag{10.1}$$

Where C_i is the total number of W$\rightarrow$ eυ candidate events observed, B_i the estimated number of background events, $\mathcal{A}_i$ the acceptance and $\mathcal{S}_i$ the unsmearing factor in bin i and $\mathcal{L}$ is the integrated luminosity of the dataset. Tables C.1 to C.6 show the candidate, total background, acceptance and unsmearing numbers in each bin of each differential variable, along with the calculated cross-section in that bin, using an integrated luminosity of 320pb^{-1}. Note that in Tables C.1 to C.6 the quoted cross-section for a particular bin i is the inclusive cross-section in that bin, and to get the corresponding differential cross-section one simply divides by the quoted bin width. Figure 10.1 shows the first, second, third and fourth jet differential cross-sections with respect to jet E_T , and Figures 10.2 and 10.3 the differential cross-section with respect to first-second jet ΔR_jj and invariant mass. Shown also are the ELO ALPGEN + PYTHIA predictions, where the theory is normalised to the measured inclusive cross-section in each case. Comparisons with theory are discussed in detail in the next chapter.

The maximum E_T limit of our differential measurements with respect to jet E_T is determined by the requirement that the cross-section is positive and not consistent with zero (i.e. $\sigma - \Delta\sigma > 0$).

10.3 The Error on the Cross-Section

As with any measurement there are two basic types of error that we can consider: statistical and systematic. Our measurement of the number of observed candidate events C_i in each bin i is a sample from a Poisson distribution, and as such has an associated statistical error $\sigma_C = \sqrt{C_i}$ (assuming the mean of the distribution $\mu = C_i$). In column five of Tables C.1 to C.6 is reported the absolute size of the
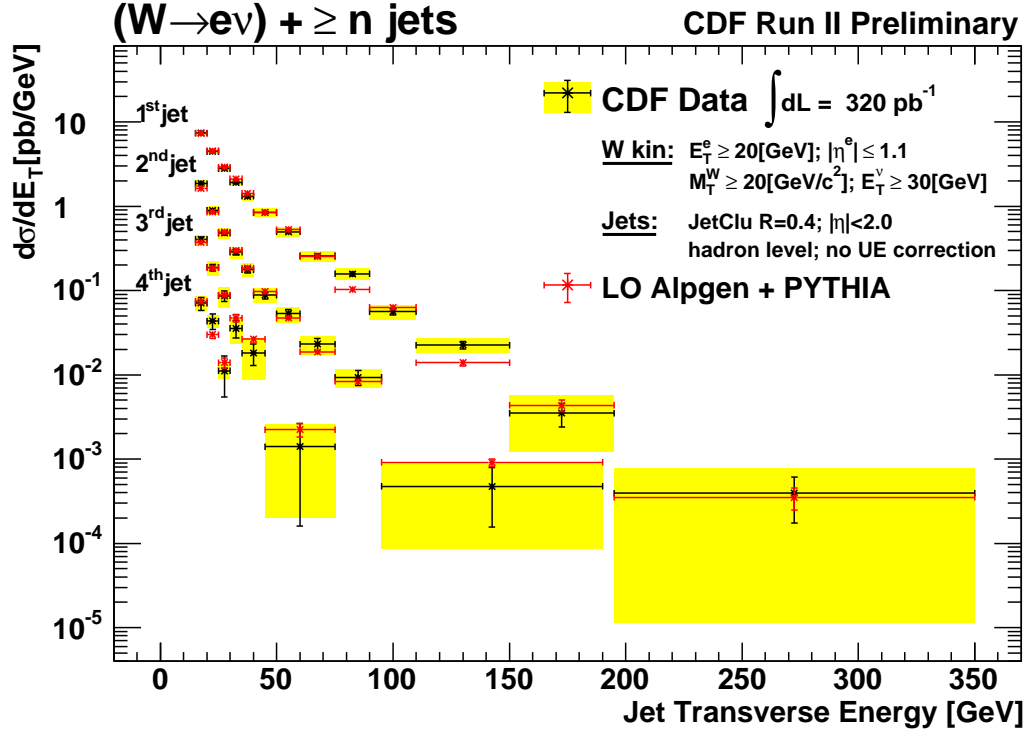
Figure 10.1: The measured first, second, third and fourth jet differential cross-sections with respect to jet E_T . The black error bars indicate the statistical error and the yellow band the systematic. Also shown is the ELO ALPGEN + PYTHIA prediction, normalised to the measured inclusive cross-section.

candidate statistical error in each bin. We define the statistical error on the cross-section as the uncertainty resulting from this counting of candidate events. It is evaluated by varying the number of candidate events by $\pm\sigma_C$ and repeating the cross-section calculation.

The systematic error on the differential cross-section results from the following uncertainties:

- The uncertainty on the background estimate, estimated in Chapter 7.

- The uncertainty on the acceptance estimate, estimated in Chapter 8.

- The uncertainty on the unsmearing procedure, estimated in Chapter 9.

- The uncertainty resulting from the 2-4% error on the jet energy scale, explained in Section 6.2.4.

With the exception of the jet energy scale, each of these uncertainties is propagated into the cross-section by individually varying the associated quantity by plus/minus the error, keeping all other quantities at their central values, and recalculating the cross-section. The systematic is then the difference between the central and "shifted"
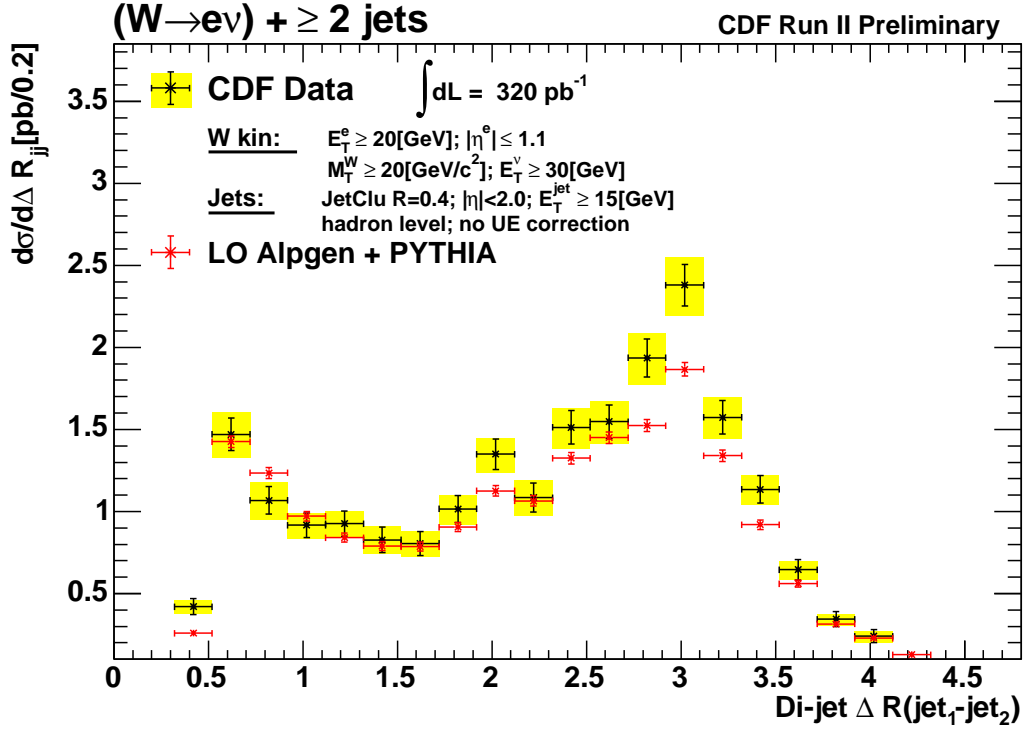
Figure 10.2: The measured differential cross-sections with respect to first-second jet ΔR_{jj}. The black error bars indicate the statistical error and the yellow band the systematic. Also shown is the ELO ALPGEN + PYTHIA prediction, normalised to the measured inclusive cross-section.

cross-sections. The sum in quadrature of each systematic contribution gives the total systematic error on the cross-section.

Evaluation of the jet energy scale systematic is not as straightforward. In principle a variation in the jet energy scale can affect each of the candidate, background and acceptance components in a particular bin. It will result in the migration of events between inclusive jet multiplicity samples and the migration of events between bins in variables that are dependent on the jet energy i.e. the first, second, third and fourth jet E_T and M_{jj} distributions. In the ΔR_{jj} distribution the only effect will be the migration of events between the ≥ 1 jet and ≥ 2 jet multiplicity bins. Both the candidate and background components can be affected in this way, but the acceptance is largely insensitive to a redefinition of the jet energy scale because it is not an explicit function of jet E_T (see Equation 8.2). The most obvious way to propagate the uncertainty on the jet energy scale is to reevaluate all the components of the measurement with a plus/minus variation of the energy scale and observe the change in the cross-section. However, in all of our differential variables we have some bins with limited statistics, and the affect of varying the jet energy scale on such bins is subject to statistical fluctuations. For example, if many of the events in a particular bin are near the bin boundary the migration of events
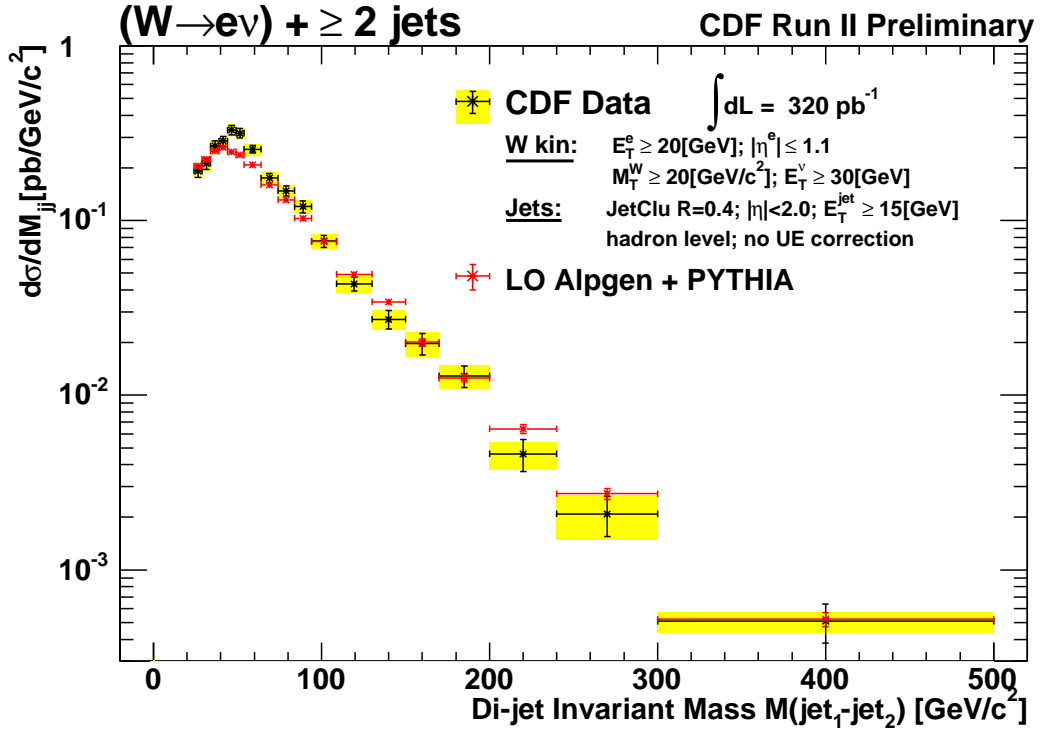
Figure 10.3: The measured differential cross-sections with respect to first-second jet invariant mass. The black error bars indicate the statistical error and the yellow band the systematic. Also shown is the ELO ALPGEN + PYTHIA prediction, normalised to the measured inclusive cross-section.

will be large compared to a bin where the events happen to be more centrally distributed. To overcome this problem we estimate the jet scale systematic by using a $W \rightarrow e\upsilon$ Monte Carlo event sample. We verified that, to within the statistics available, the proportional change in the number of candidate and background events in any particular bin when the jet energy scale is varied is the same. This means that the net change in the cross-section is reduced to the change in the signal events $(C - B)$, which, since the acceptance is very flat as a function of jet variables, are distributed as the differential cross-section we have measured. We can use Monte Carlo events to simulate the effect of varying the jet scale on signal events. First the distribution of the Monte Carlo in the variable in question is reweighted such that it agrees with that of the differential cross-section. The jet energy scale in the Monte Carlo events is then varied by plus/minus the jet energy scale uncertainty and the relative change in each bin gives the cross-section systematic.

In Tables C.7 to C.12 are detailed the errors on the differential cross-section measurements. Shown is the absolute systematic error from each of the jet energy scale, background, acceptance and unsmearing components, the total systematic error when these contributions are added in quadrature, and the absolute statistical error on the cross-section resulting from the counting of candidates. Figure 10.4

shows the relative error from each systematic component and the total relative systematic and statistical errors on the differential cross-section measurements. One can see that at low E_T the dominant systematic is that arising from the jet energy scale uncertainty, but that at high E_T the uncertainty in the background subtraction, largely arising from the limited QCD background template statistics in this region, dominates.
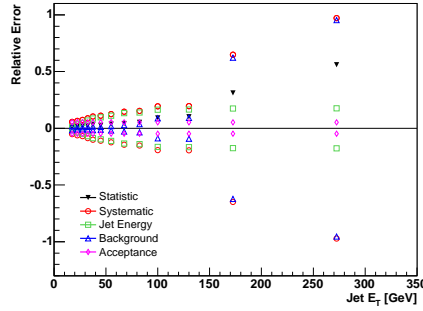
(a) Lead jet E_T in ≥ 1 jet
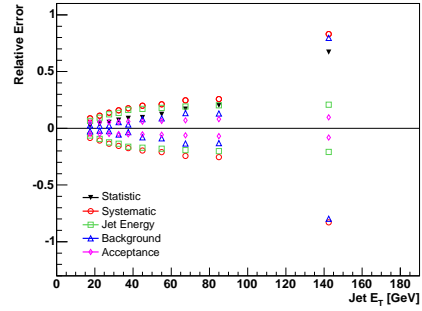
(b) Second jet E_T in ≥ 2 jet

(c) Third jet E_T in ≥ 3 jet

(d) Fourth jet E_T in ≥ 4 jet

(e) ΔR_{jj} in ≥ 2 jet
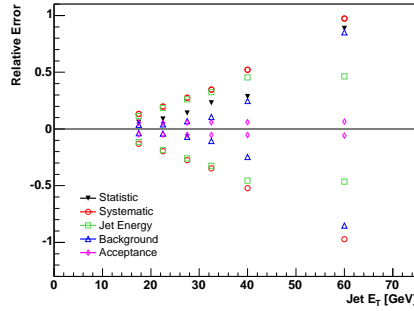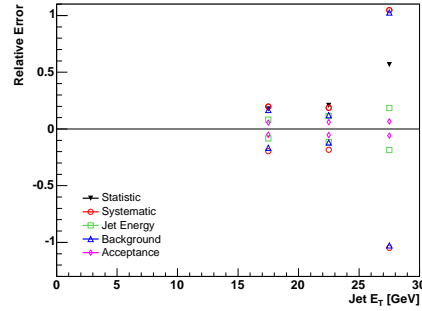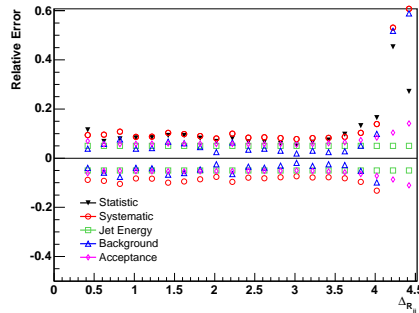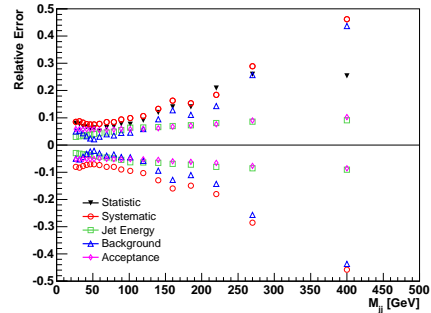
(f) M_{jj} in ≥ 2 jet

Figure 10.4: The relative error from each systematic component and the total relative systematic and statistical errors on the cross-section for each differential measurement.

The uncertainty resulting from the 6% error on the measurement of the integrated luminosity, explained in Section 4.2.5, is not included in the differential cross-section systematics shown here. All of the systematic contributions discussed above contribute to an uncertainty in the shape of the differential cross-section and are to first order uncorrelated bin-to-bin. The integrated luminosity error represents a 6% uncertainty in the overall normalisation of the differential cross-section, and to include it as a systematic on each bin would give a misleading impression of the uncertainty in the shape of the differential measurement. Similarly the relatively small uncertainties on the trigger efficiency and Z vertex acceptance (see Chapter 8) are not included.

10.4 Integrated Cross-Section Results

Previous W + Jets studies [20] [107] have measured the inclusive W $+ \geq n$ jet cross-section (for $n = 1 - 4$), where jets are required to pass a minimum $E_T > 15$GeV threshold. The differential candidate, acceptance and background quantities with respect to first, second, third and fourth jet E_T allow us to calculate these inclusive cross-sections not only for a minimum jet E_T threshold of 15 GeV, but for a whole range of thresholds from 15 GeV upwards. The integrated cross-section above a minimum threshold of E_T^{min} is given by:

$$\int_{E_T^{min}}^{\infty} \frac{d\sigma}{dE_T} dE_T = \frac{\sum_{i=E_T^{min}}^{N_{bins}} (C_i - B_i)}{\mathcal{A}_{\geq E_T^{min}} \cdot \mathcal{L}} \times \mathcal{S}_{\geq E_T^{min}} \qquad (10.2)$$

where C_i and B_i are the number of candidate and background events in bin i respectively, and the sum is over all bins above the E_T^{min} threshold. The integrated acceptance and unsmearing quantities cannot be determined by summing the differential bins in this way because they are multiplicative factors. They must be recalculated directly by considering all events which pass the minimum E_T threshold. In the case of the acceptance, $\mathcal{A}_{\geq E_T^{min}}$ is given by evaluating Equation 8.2 using events with jet $E_T > E_T^{min}$. The integrated unsmearing factor $\mathcal{S}_{\geq E_T^{min}}$ is determined by evaluating the calorimeter to hadron ratio, Equation 9.2, using all bins above the E_T^{min} threshold.

Tables C.13 to C.16 show the integrated candidate, background, acceptance and unsmearing results for a range of E_T^{min} thresholds for each of the first, second, third and fourth jet E_T distributions, and the inclusive W$\rightarrow$ eυ $+\geq n$ jet cross-sections that result. Note that the integrated cross-section for the largest E_T^{min} threshold is necessarily identical to the differential cross-section in the highest E_T bin. The errors on the integrated cross-sections are evaluated in exactly the same way as for the differential measurement. Tables C.17 to C.20 show the absolute errors on

the integrated cross-sections, and Figure 10.5 the relative statistical and systematic errors.

(a) W→ eυ + ≥ 1 jet

(b) W→ eυ + ≥ 2 jet

(c) W→ eυ + ≥ 3 jet

(d) W→ eυ + ≥ 4 jet
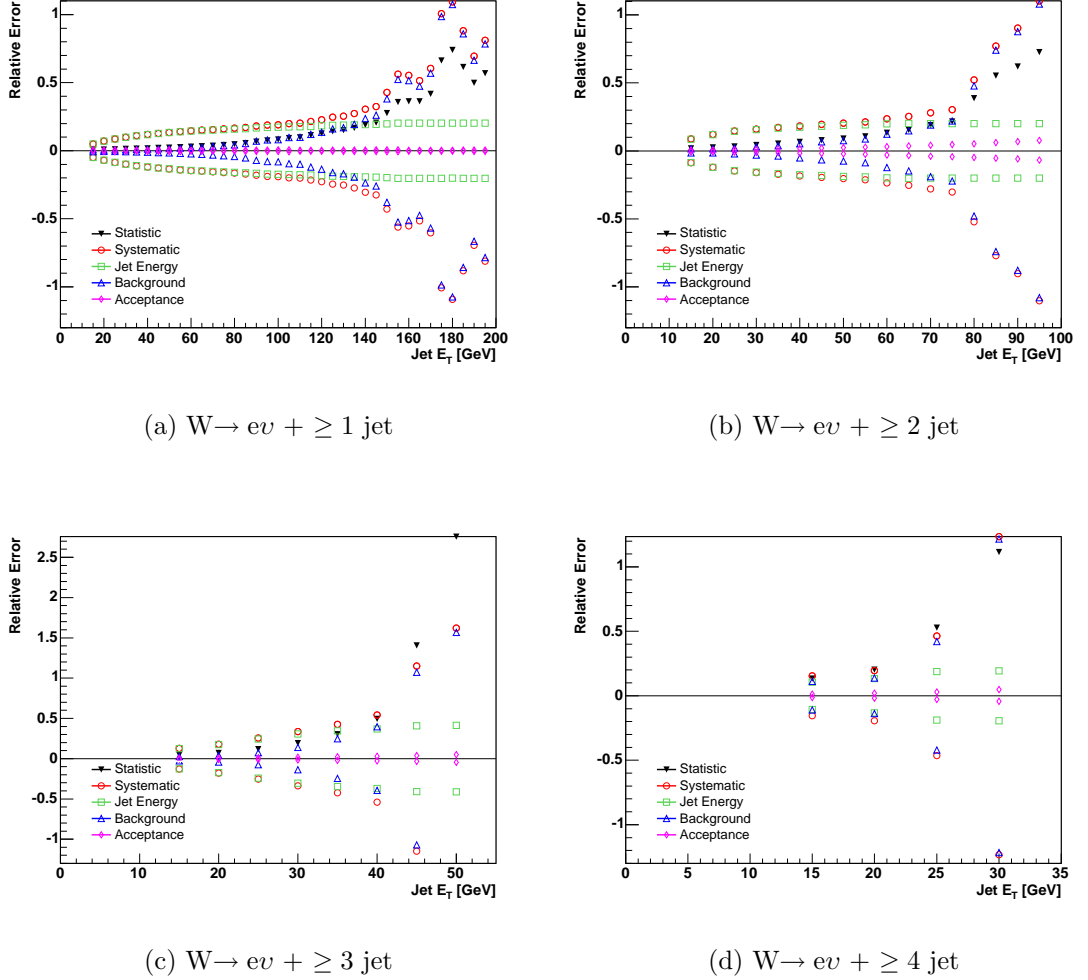
Figure 10.5: The relative statistical and systematic errors on each of the W→ eυ + $\geq$ n jet integrated cross-section measurements. The x-axis is the minimum jet E_T threshold.

In Figure 10.6 the dependence of the inclusive W→ eυ + ≥ 1 jet , ≥ 2 jet , ≥ 3 jet and ≥ 4 jet cross-section on the minimum jet E_T threshold is shown. Figure 10.7 shows this information in a different form, allowing for comparison of the change of each inclusive W→ eυ + $\geq$ n jet cross-section with increasing jet E_T threshold. Here the W→ eυ + ≥ 0 jet cross-section shown is calculated as explained in Section 10.1, but using the inclusive acceptance factor for the restricted W decay kinematics for consistency with the $\geq n$ jet results (i.e. that given by Equation 8.2). This is determined, using a PYTHIA W→ eυ event sample, to be 0.604. This gives, for the inclusive W cross section with W decay kinematics restricted, $\sigma'(W \rightarrow e\upsilon + X)$:

$$\sigma'(\mathrm{W} \to \mathrm{e}\upsilon + \mathrm{X}) = 798 \pm 2.2(\mathrm{stat.}) \pm 40(\mathrm{sys.}) \pm 48(\mathrm{lum.})\ \mathrm{pb}$$
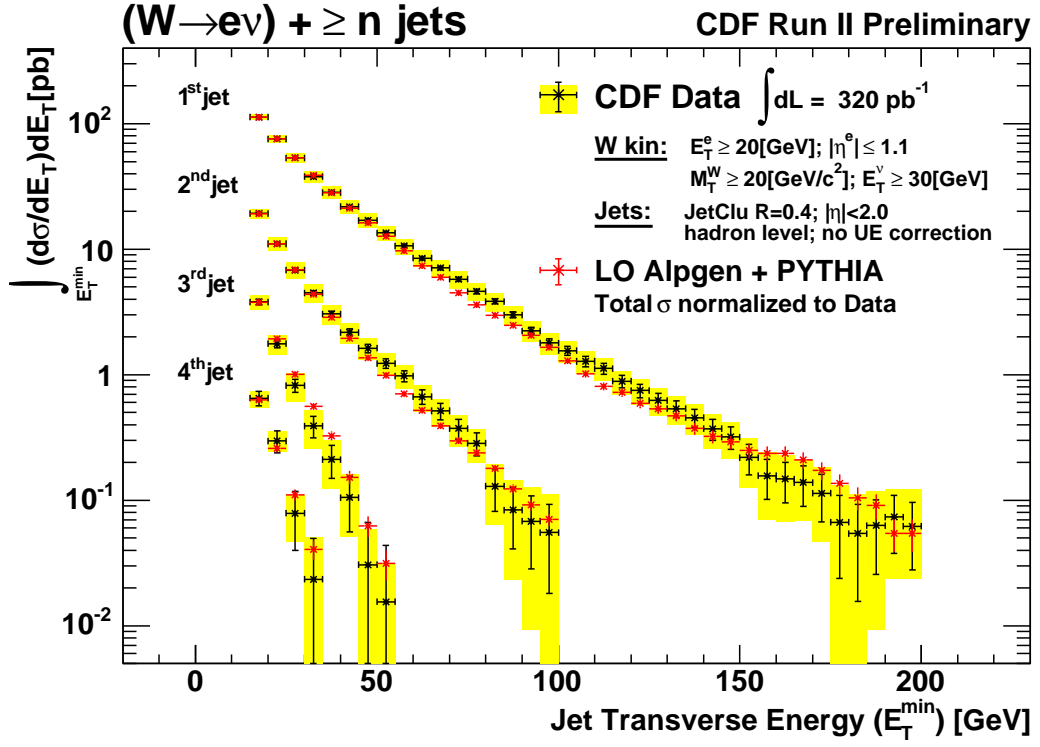
Figure 10.6: The measured first, second, third and fourth jet integrated cross-section results. The black error bars indicate the statistical error and the yellow band the systematic. Also shown is the ELO ALPGEN + PYTHIA prediction, normalised to the measured inclusive cross-section.
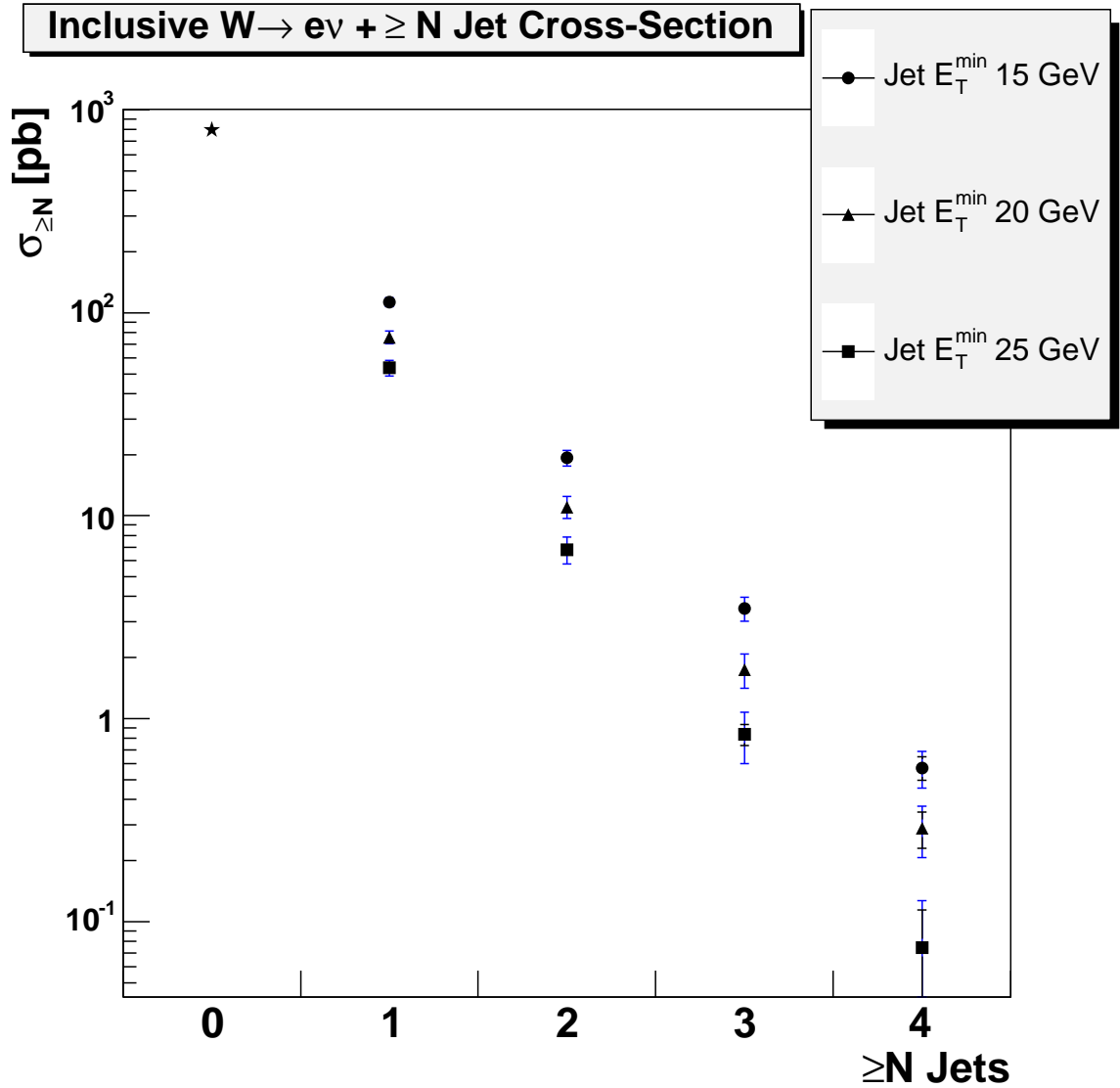
Figure 10.7: Inclusive W$\rightarrow$ eυ $+ \geq n$ jet cross-section for various jet E_T thresholds. The inner bars indicate statistical errors and the outer the total statistical plus systematic. Note that a different symbol is used for the inclusive ≥ 0 jet cross-section, since this does not depend on the jet E_T threshold.

Chapter 11

Comparison to Theory

In this chapter we make comparisons between our measured cross-sections and the predictions of leading order W + Jet Monte Carlo event generators.

Previous studies have observed considerable discrepancies between measured W$\to$eυ + $\geq$ n jet cross-sections and the predictions of n parton leading order generators. The Run I W + Jets study [20], which compared the leading order predictions of VECBOS to parton-level W$\to$eυ + $\geq n$ jet measured cross-sections, observed discrepancy at the level of 20-50%, depending on the choice of renormalisation scale. The leading order inclusive W$\to$eυ cross-section prediction of ALPGEN is 2036pb, some 30% lower than the most recent CDF measurement of 2775 ± 175pb [17], and the NNLO calculation of 2687 ± 54pb [87]. These discrepancies are due to the absence of contributions from higher order diagrams in the leading order matrix element calculation. However, whilst leading order calculations may have difficulty reproducing the absolute rates of jet production, they may perform better in their description of certain relative rates which are less sensitive to higher order corrections.

We make the following comparisons between data and Monte Carlo:

- Comparisons of the inclusive W$\to$eυ + $\geq$ n jet cross-sections $\sigma_{\geq n}$, for $n = 0, 1, 2, 3, 4$ and various jet E_T thresholds.

- Comparisons of the differential cross-section distributions in jet E_T , ΔR_{jj} and M_{jj} variables, where the Monte Carlo is normalised to the inclusive data cross-section in the variable of interest.

Two approaches to leading order W + Jets predictions are considered; the Enhanced Leading Order approach and the ME-PS matching approach, both of which are detailed in Chapter 3.

In the ELO approach an n parton leading order matrix element calculation is interfaced with a parton showering programme to enable description of $\geq n$ jet final states. We examine the dependence of the ELO predictions on the number of matrix

element partons, the parton level generation cuts, choice of renormalisation scale and the parton showering programme used. ALPGEN is used for the matrix element calculation and either HERWIG or PYTHIA used for parton showering.

The performance of two different ME-PS matching schemes are examined: the CKKW scheme (Section 3.9.1) and the MLM scheme (Section 3.9.2). By introducing communication between the matrix element calculation and the parton showering components of the ELO generation these prescriptions claim to allow the combination of different n parton samples without introducing double counting. By including contributions from $n-1, n-2$ and $n+1, n+2$ parton leading order matrix elements this combined sample should provide a far more complete description of W + Jets final states than a single n parton ELO sample, capable of describing the kinematics in any $\geq n$ jet bin and potentially improving the inclusive W$\to$ eυ + $\geq n$ jet cross-section predictions.

11.1 Generating the Theoretical Predictions

In order to make a fair data-theory comparison it is vital to ensure that the cross-section definition used in generating the theoretical predictions is fully compatible with that used in making the measurement. Since the cross-section measurement has been made at the hadron-level, this can be achieved without the need for a complex simulation of the CDF detector's response to particles. In practice any theorist could follow the steps below to produce their own predictions that could be compared to our measurement.

11.1.1 Enhanced Leading Order

The ELO generation process produces an inclusive cross-section prediction for the W$\to$ eυ + n parton process considered, where the matrix element partons are required to pass certain generation cuts. For each ALPGEN + HERWIG/PYTHIA ELO sample we use the same generation cuts: parton $P_T > 5$ GeV and $\Delta R_{pp} > 0.35$. In order to make the hadron-level cross-section predictions suitable for comparisons with our measurement we have to impose our W and jet selection criteria on the event sample, as follows :

1. Impose the restrictions on the W decay products that are required by our cross-section definition; $P_T^{ele} > 20$GeV, $|\eta^{ele}| < 1.1$, $P_T^{neu} > 30$GeV and $W_{M_T} > 20$GeV/c^2. Veto any events which don't pass these requirements.

2. Cluster the final state hadrons into jets using the JetClu cone algorithm, and require that the jets pass the $|\eta^j| < 2.0$ cut.

3. Using the hadron-level clustered jets bin the remaining events in the observable of interest, using the same binning as our data.

4. The absolute cross-section prediction in a particular bin i is given by:

$$\sigma_i = \frac{N_i}{N_{gen}} \times \sigma_{inc} \tag{11.1}$$

where N_i is the number of events in bin i, N_{gen} is the total number of Monte Carlo events generated and σ_{inc} is the inclusive W$\rightarrow$ eυ + n parton cross-section returned by the generator.

In addition, to test the dependence of the hadron-level cross-section predictions on the parton generation cuts we can identify the matrix element partons in the event record and impose selection requirements tighter than the original $P_T > 5$ GeV and $\Delta R_{pp} > 0.35$ cuts.

11.1.2   CKKW and MLM Matching Prescriptions

The CKKW predictions are generated using MADGRAPH [102] for the matrix element calculation and PYTHIA for the parton showering. The MLM predictions are generated using ALPGEN for the matrix element calculation and HERWIG for the parton showering. In both prescriptions one generates five independent samples which differ in the number of final state partons, varying from 0 to 4 partons. For each observable we then build up the combined 0p + 1p + 2p + 3p + 4p distribution, and it is this which is compared to the data. Although the matching prescriptions used are different in each case, both MLM and CKKW combined distributions are made identically as follows:

1. As with the standard ELO samples, each of the n parton matched samples is clustered at the hadron level using the JetClu algorithm.

2. The events are then subjected to the W decay and jet $|\eta| < 2.0$ restrictions of the cross-section definition and binned in the observable of interest, using the same binning as the data. Hence we have a distribution for each n parton sample, containing N_{np} events.

3. Each n parton distribution then makes the following absolute cross-section contribution to the combined distribution:

$$\sigma_{np}^{abs} = \frac{\sigma_{np}}{N_{np}^{gen}} \times N_{np} \tag{11.2}$$

where σ_{np} is the generated cross-section and N_{np}^{gen} is the total number of events generated in the n parton sample.

Figure 11.1 shows how the different n parton samples combine to form the absolute cross-section prediction in the first jet E_T , second jet E_T and ΔR_{jj} and M_{jj} distributions in the ALPGEN + HERWIG MLM matched samples. Figure 11.2 shows the same information for the MADGRAPH + PYTHIA CKKW matched samples. One can see that the manner in which the combined cross-sections are formed differ markedly between the two matching prescriptions. For example, in the CKKW prescription the 0 parton sample makes a far more significant contribution in these distributions than in the MLM prescription. Exactly how much of a contribution is provided by each sample is determined by the particular matching "cuts" used i.e. the K_T generation parameter in CKKW, and the minimum jet E_T matching cut and cone size in MLM. However, the combined predictions of each should be independent of these matching details, and this is examined in the CKKW case in Sections 11.2.2 and 11.3.2.

11.2   Inclusive W→ ev + $\geq n$ jet Cross-Section

11.2.1   ELO Comparisons

In this section we compare the inclusive W→ ev + $\geq n$ jet cross-section predictions of ALPGEN + PYTHIA and ALPGEN + HERWIG ELO samples to the data.

In Figure 11.3 the n parton ALPGEN + PYTHIA sample is used in the $\geq n$ jet bin, and the comparison made for two different minimum jet E_T definitions using different matrix element parton P_T generation cuts. Using a minimum jet E_T of 15 GeV the cross-section prediction is 10-30% larger for a 5 GeV matrix element parton P_T cut than a 10 GeV cut (see Figure 11.3(c)). This means that the parton phase space between $P_T = 5$ and 10 GeV is contributing a significant additional cross-section above jet E_T 15 GeV. The matrix element parton P_T cut is necessarily imposed to exclude regions of phase space where the leading order cross-section is divergent. Exactly where this cut is placed should not, within threshold effects, alter significantly the jet cross-section. However, even using a minimum jet E_T of 30 GeV we observe dependence on the matrix element parton cuts at the level of 5-20% (see Figure 11.3(d)). Figure 11.4 demonstrates that in the W→ ev + ≥ 1 jet and ≥ 2 jet case the dependence does not disappear until the jet E_T threshold is some 60 GeV. The suspected reason for this dependence is as follows. In the ELO approach the probability of some additional hard final state parton being produced in initial state radiation is independent of the parton P_T cut imposed in the matrix element calculation. As this P_T cut is lowered we approach the cross-section divergence and the inclusive cross-section increases, increasing the opportunity for hard ISR.

Given the dependence of the ELO cross-sections on the generation cuts it is difficult to draw meaningful comparisons between the predicted and observed rates in

Figure 11.3; the ELO predictions are simply not well defined. However, we can also examine the ability of the ELO approach to predict the relative rate of additional final state jets. Figure 11.5 shows the prediction of the ALPGEN + PYTHIA 1 parton sample, where the theory is normalised to the measured cross-section in the ≥ 1 jet bin. This comparison reveals that the shape of the predicted jet multiplicity spectrum is largely insensitive to the parton P_T generation cut applied; the generation cut impacts the absolute normalisation of the $W \to e\upsilon + \geq n$ jet cross-section predictions only. It is also evident that the 1 parton sample describes the rate of 2, 3 and 4 or more jets relative to the ≥ 1 jet rate in the data very well. This represents something of a success for the PYTHIA parton showering approach; it is the parton shower which is responsible for the generation of 2 or more jet events in the 1 parton sample. However, in Figure 11.6 we examine the dependence of the shape of the predicted jet multiplicity spectrum on the choice of renormalisation scale in ALPGEN + PYTHIA 1 parton and 2 parton samples. One observes significant dependence at the level of 50% in the 1 parton sample, and somewhat less in the 2 parton sample. This is not surprising given the dependence of the jet E_T spectrum on the renormalisation scale, discussed in Section 11.3.

In Figure 11.7 we compare the $W \to e\upsilon + \geq n$ jet cross-section predictions of ALPGEN + PYTHIA, with and without the PYTHIA underlying event model, and ALPGEN + HERWIG n parton samples, where identical generation cuts (parton $P_T > 5 \text{GeV}$) and renormalisation scale choices ($Q^2 = M_W^2 + \sum_{P_{T,j}^2}$) are used in each case, and the n parton sample is used in the $\geq n$ jet bin. The HERWIG predictions contain significantly fewer final state jets than that of PYTHIA, both with and without the underlying event model, such that the cross-section prediction of HERWIG is between 15% and 50% lower than PYTHIA, regardless of the jet energy threshold. This can only be ascribed to the differences in the parton showering models between PYTHIA and HERWIG. The absence of the underlying event model in PYTHIA reduces the predicted cross-section by some 10% using a jet threshold of 15 GeV, and less for a threshold of 30 GeV, consistent with the underlying event making a smaller contribution at higher jet energies.

11.2.2 Matched Sample Comparisons

In this section we compare the inclusive $W \to e\upsilon + \geq n$ jet cross-section predictions of MLM and CKKW ME-PS matched samples to the data.

Figure 11.8 shows the MADGRAPH + PYTHIA CKKW predictions alongside the data, for two different jet E_T thresholds ($E_T > 15$ and $E_T > 30$ GeV) and three different CKKW K_T generation cuts ($K_T > 10, K_T > 15$ and $K_T > 20$ GeV). The CKKW $W \to e\upsilon + \geq n$ jet cross-section predictions are generally within 20% of the measured values. However, Figure 11.9 shows that there is substantial dependence

of the CKKW predictions on the K_T generation cut used. This is larger in the higher jet multiplicity samples, ranging from 5% - 50% for a jet E_T threshold of 15 GeV, but somewhat smaller for a threshold of 30 GeV. Qualitatively speaking, the larger the K_T generation cut, the larger the fraction of phase space which is covered by the parton shower. However, dependence at this level is surprising; one of the major reasons for introducing ME-PS matching prescriptions is to reduce the dependence of predictions on the generation cuts.

Figure 11.10 shows the ALPGEN + HERWIG MLM predictions alongside the data, again for two different jet E_T thresholds ($E_T > 15$ and $E_T > 30$ GeV). With a jet E_T threshold of 15 GeV the predictions are some 60% smaller than the measured values, the discrepancy decreasing somewhat with the higher jet E_T threshold of 30 GeV. This is consistent with the comparisons between HERWIG and PYTHIA ELO results in the previous section; HERWIG producing fewer jets than PYTHIA.

However, the Data/Theory plots in Figures 11.8 and 11.10 indicate that whilst the absolute normalisation of the MLM predictions is worse than the CKKW, the slope, or relative change in the cross-section with additional jets, is better reproduced by the MLM predictions. In Figures 11.11 and 11.12 the CKKW and MLM predictions are normalised to either the measured ≥ 0 jet or ≥ 1 jet cross-section. These show that whilst CKKW well reproduces the measured $\sigma_{\geq 1j}/\sigma_{\geq 0j}$, it fails to predict as well the relative rate for two, three and four jet production in ≥ 1 jet events. The MLM sample, on the other hand, well reproduces the relative rate of additional jet production up to four jets in the full ≥ 0 jet sample (although with the lower jet E_T threshold of 15 GeV it has difficulty predicting the measured $\sigma_{\geq 1j}/\sigma_{\geq 0j}$ ratio).

11.3   Differential Cross-Sections

For the comparisons to the differential cross-section measurements all the Monte Carlo predictions shown are normalised to the relevant inclusive cross-section measurement. For example, in the first jet E_T distribution this is the inclusive $W \rightarrow e\upsilon +$ ≥ 1 jet cross-section for $E_T^{min} > 15$ GeV, in the second jet E_T , ΔR_{jj} and M_{jj} distribution this is the inclusive $W \rightarrow e\upsilon + \geq 2$ jet cross-section, and so on. This is done such that we can examine the Monte Carlo description of the kinematics independent of the discrepancies in the absolute cross-section predictions shown in the previous section.

11.3.1   ELO comparisons

Jet E_T Distributions

Figure 11.13 shows the measured first, second, third and fourth jet E_T differential cross-section distributions, along with the ELO ALPGEN + PYTHIA predictions for different parton P_T generation cuts, where the n parton sample is used in the $\geq n$ jet bin. The dependence of the shape of the differential cross-section distribution on the generation cuts is small.

Figures 11.14 and 11.15 show the Data/Theory ratio for first, second, third and fourth jet E_T distributions, where the data is compared to ALPGEN + PYTHIA predictions with a parton P_T generation cut of 5 GeV. The n parton sample is used in the $\geq n$ jet bin, but with varying choices of renormalisation scale; $Q^2 = M_W^2$ and $Q^2 = <P_{T,j}^2>$. For the comparatively large, fixed scale choice of $Q^2 = M_W^2$ the predicted shapes of the jet E_T spectra are somewhat harder than the data. However, using a lower, dynamic scale choice of the average parton P_T , $Q^2 = <P_{T,j}^2>$, the agreement with the data in the shape of the jet E_T distributions is improved considerably, particularly in the low E_T region. It is difficult to draw conclusions from the high E_T region, since the errors on the data are very large.

Figure 11.16 show the Data/Theory ratio for first, second, third and fourth jet E_T distributions respectively using ALPGEN + HERWIG predictions with the renormalisation scale $Q^2 = <P_{T,j}^2>$. There is very little discernible difference between these predictions and those of the ALPGEN + PYTHIA samples with the same choice of scale, indicating that the kinematic shape of the predictions is not particularly sensitive to the parton shower model.

The comparisons in Figures 11.14 and 11.15 are made using n parton sample which we expect to reproduce the data best. Figure 11.17 compares the predictions of the 1 parton and 3 parton ALPGEN + PYTHIA samples to the data in the second E_T differential cross-section. Since these Monte Carlo samples are generated with a parton P_T^{min} cut of 5 GeV, the $n-1$ parton prediction of a $\geq n$ jet distribution will come entirely from the parton showering model used. The 1 parton ELO prediction of the second jet E_T distribution is reasonable at low jet E_T , but it begins to fail in the high E_T region when compared with the 2 parton ELO prediction. This is indicative of a limitation of parton showering models; the high E_T region requires additional contributions from the n parton matrix element calculation to describe it. Conversely the shape of the 3 parton prediction is too hard and it fails to describe the shape in the low E_T region.

ΔR_{jj} and M_{jj} Distributions

Figure 11.18 shows the measured first-second jet ΔR_{jj} and M_{jj} differential cross-section distributions, along with the ELO ALPGEN + PYTHIA 2 parton predictions using different parton P_T generation cuts, where the theory is normalised to the ≥ 2 jet cross-section. These again support the conclusion that the shape of the event jet kinematic variables do not depend significantly on the generation cuts.

Figure 11.19 compares the predicted shape of the ALPGEN + PYTHIA 2 parton sample to the data in the M_{jj} and ΔR_{jj} distributions, for two different renormalisation scales. Interestingly, the shape of the ΔR_{jj} distribution is better reproduced by the fixed $Q^2 = M_W^2$ scale. Figure 11.20 shows the data over theory ratio for the ALPGEN + PYTHIA 2 parton predictions in the M_{jj} distribution. Neither scale well describes the position or shape of the the low mass peak in the M_{jj} distribution, a product of the 15 GeV cut on the jet E_T and the implicit spatial separation of the non-overlapping cone algorithm.

11.3.2 Matched Sample Comparisons

Jet E_T Distributions

Figure 11.21 shows the CKKW prediction for three different K_T generation cuts compared with the measured differential cross-section in each of the first, second, third and fourth jet E_T distributions. The shape of the predictions does not depend on the generation cuts.

Figures 11.22 and 11.23 show the Data/Theory ratio for the first to fourth jet E_T distributions, where the data is compared CKKW and MLM matched samples. The single MADGRAPH + PYTHIA CKKW combined sample describes the shape of each of the jet E_T spectra very well, although there is a consistent excess at high P_T . However, the predictions of the ALPGEN + HERWIG MLM matched sample are consistently harder than observed in the data.

ΔR_{jj} and M_{jj} Distributions

Figure 11.24 shows the measured ΔR_{jj} and M_{jj} differential cross-sections with the CKKW and MLM predictions, where the theory has been normalised to the measured ≥ 2 jet cross-section. Both the CKKW and MLM predictions describe the shape of the ΔR_{jj} distribution reasonably well, the MLM sample perhaps performing slightly better in terms of the relative sizes of the back-to-back and collinear peaks. However, the CKKW sample comes considerably closer to describing the low mass peak in the M_{jj} distribution than any other theoretical prediction tested. This distribution is a complex convolution of the first and second jet E_T and ΔR_{jj} distributions, all of which are well described by the CKKW sample.

11.4   Conclusions

These comparisons have demonstrated that, in terms of predicting absolute rates of jet production in association with a W boson, the ELO approach is not robust enough to give a definitive answer; the rate being heavily dependent on the details of the matrix element parton generation cuts. Where the ELO approach does better is in predicting relative rates of jet production; the rate for $> n$ jet production in $\geq n$ jet events, for example, and the shape of the differential cross-sections. These are largely insensitive to the parton generation cuts, and compare well with the relative rates observed in the data. One could conclude that to predict the absolute rates of QCD processes one needs the sophistication of a higher order calculation, in which higher order diagrams and their interference effects are fully accounted for. However, these relative rates, the kinematic "structure" of the events, are not so sensitive to higher order corrections, and can be well described by the ELO approach (with the notable exception of the complex dijet invariant mass distribution). To do this successfully there is clearly a need for tuning the ELO approach on data such as that presented here. The "shape" predictions are dependent on the matrix element parton multiplicity and the choice of renormalisation scale. Whilst choice of the former is obvious, the latter is considerably more complicated, with different scales being better suited to different distributions.

These studies also indicate the potential power of the ME-PS matching approach. The combined CKKW sample describes with considerable success the shape of the jet kinematic distributions across the full jet multiplicity range explored, including the problematic invariant mass distribution, something which cannot be achieved using a single n parton ELO sample. Although the combined MLM prediction does not reproduce these kinematic shapes as well, it provides an excellent description of the relative rate of jet production in W events, which is somewhat worse in CKKW. Taken in isolation these observations seem to indicate that the combination of the different n parton samples in the resepective ME-PS matching schemes successfully avoids the issue of double-counting, but the fact that both features cannot be observed in a single scheme implies that there may still be work to do on these approaches. In particular, the CKKW predictions for the absolute rates of W $+ \geq n$ jet production display considerable dependence on the choice of K_T generation cut, a somewhat disappointing observation, given that a major motivation for implementing ME-PS matching schemes is to remove this dependence.

(a) First Jet E_T

(b) Second Jet E_T

(c) ΔR_{jj}
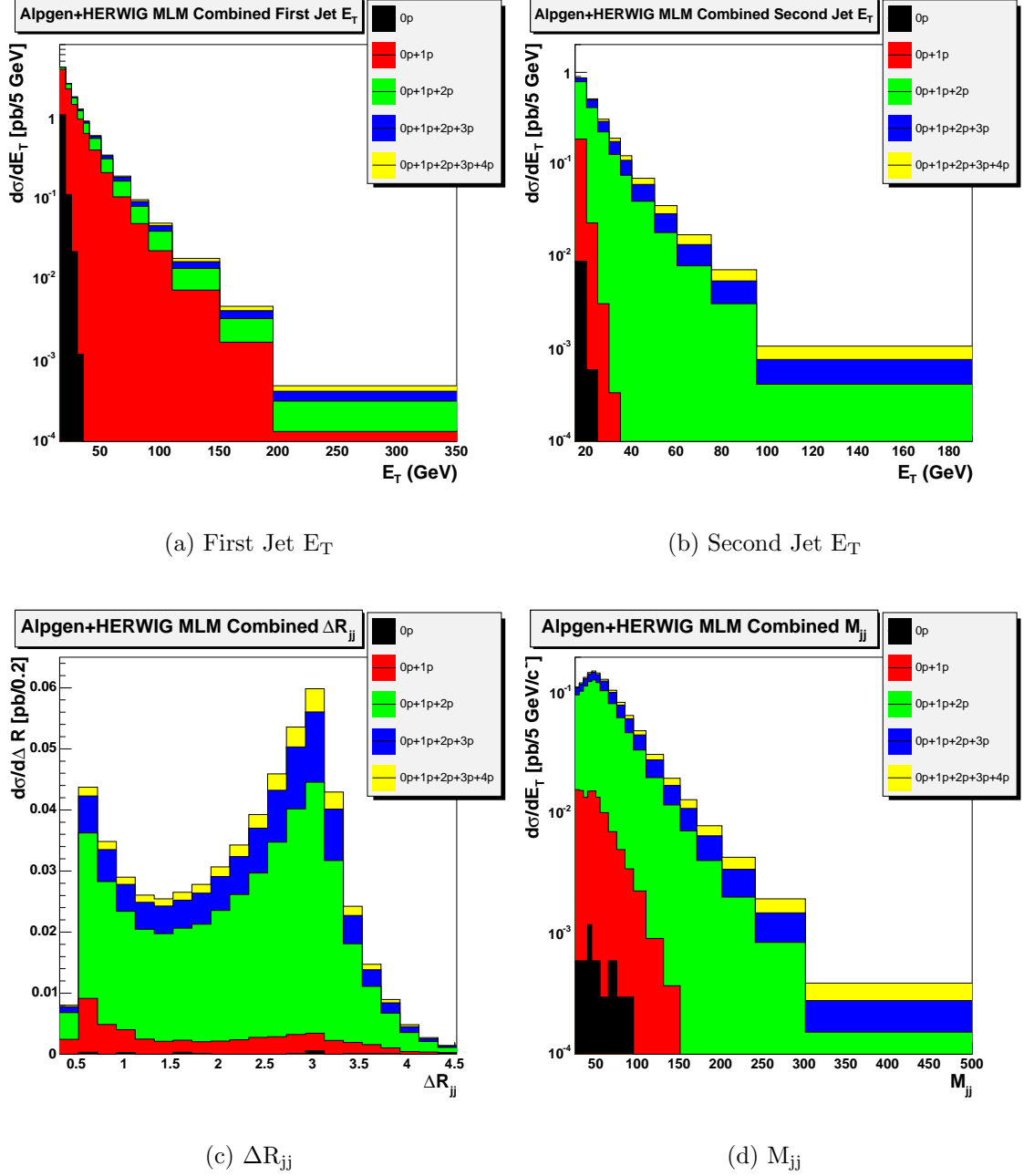
(d) M_{jj}

Figure 11.1: Plot showing the composition of the MLM matched ALPGEN + HER-WIG combined absolute cross-section prediction in the first jet E_T , second jet E_T , ΔR_{jj} and M_{jj} distributions. The black area shows the contribution from the 0 parton sample only, the red the combined contribution from the 0 parton and 1 parton samples, and so on.

(a) First Jet E_T

(b) Second Jet E_T

(c) ΔR_{jj}
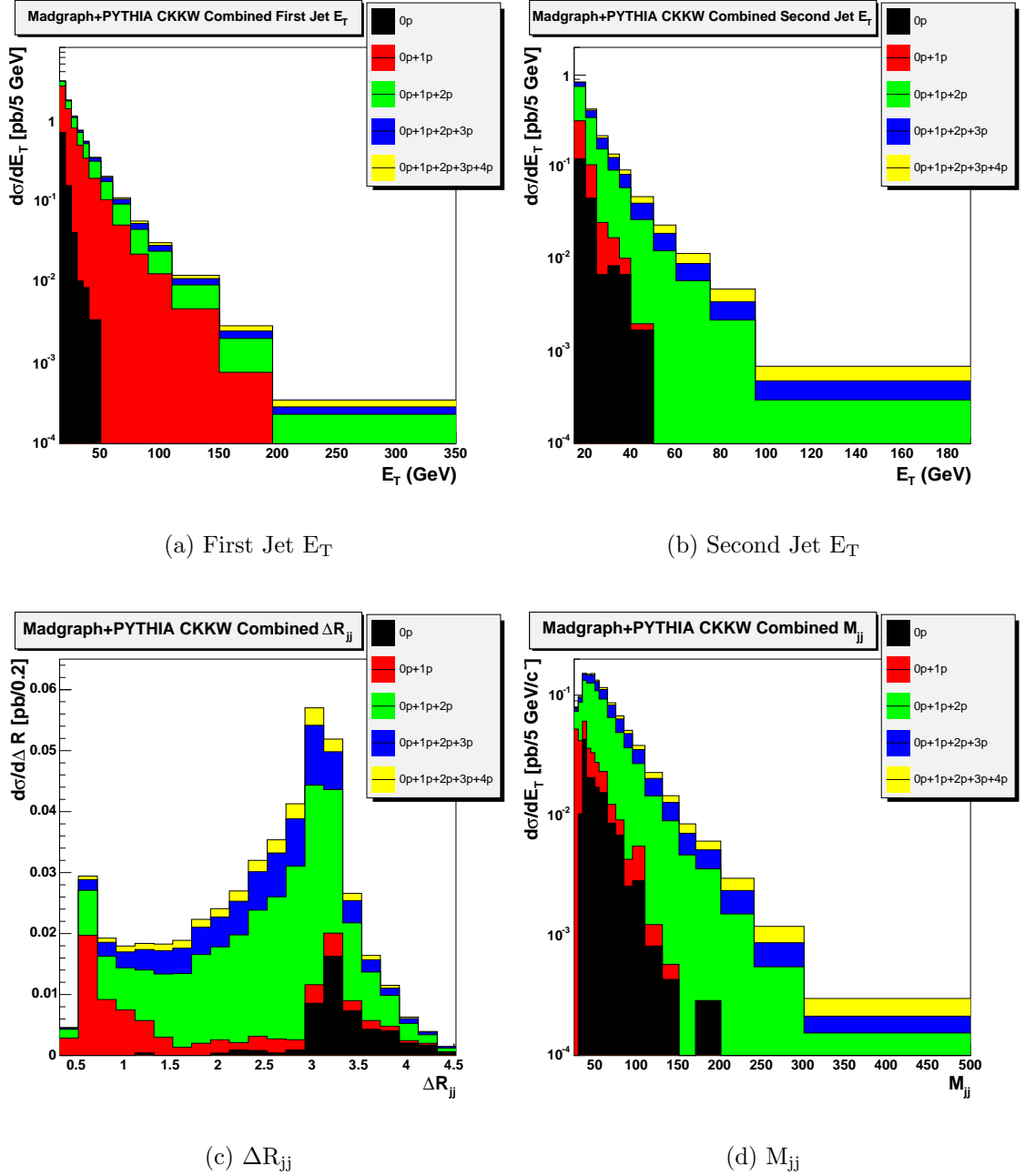
(d) M_{jj}

Figure 11.2: Plot showing the composition of the CKKW matched MADGRAPH + PYTHIA combined absolute cross-section prediction in the first jet E_T, second jet E_T, ΔR_{jj} and M_{jj} distributions.

(a) $E_T^{jet} \geq 15$ GeV

(b) $E_T^{jet} \geq 30$ GeV

(c) $E_T^{jet} \geq 15$ GeV

(d) $E_T^{jet} \geq 30$ GeV
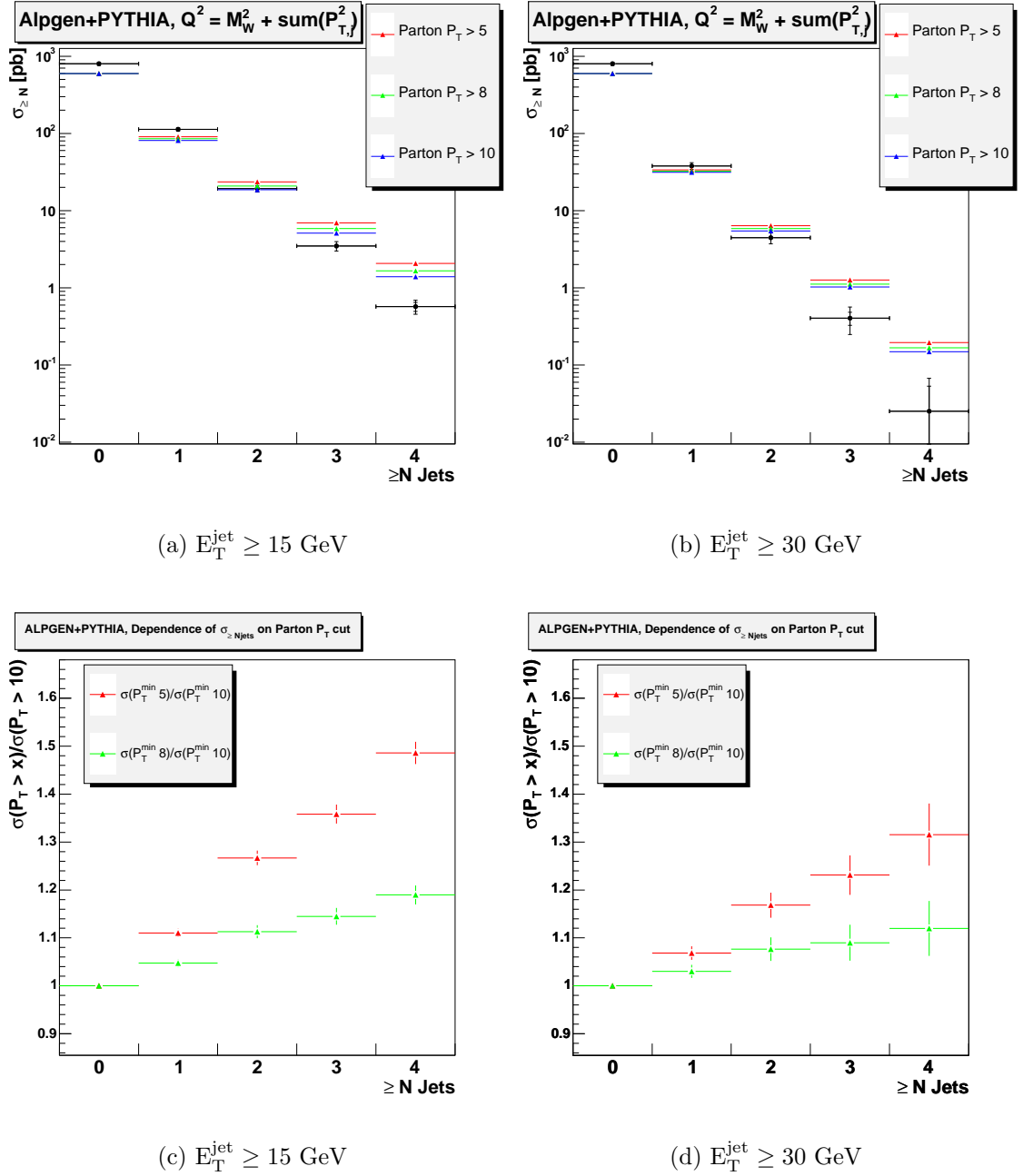
Figure 11.3: The top plots show the $W \to e\upsilon + \geq n$ jet cross-section prediction of ALPGEN + PYTHIA n parton samples compared to our measured data, where the n parton sample is used in the $\geq n$ jet bin. The two plots are for two different minimum jet E_T cuts, and in each the predictions for different parton P_T generation cuts are shown, with the measured data in black. For the data the inner error bars show the statistical and outer the total statistical plus systematic uncertainty. The bottom plots show the dependence of the cross-section prediction on the parton P_T generation cut.

(a) $W \rightarrow e\upsilon + \geq 1$ jet
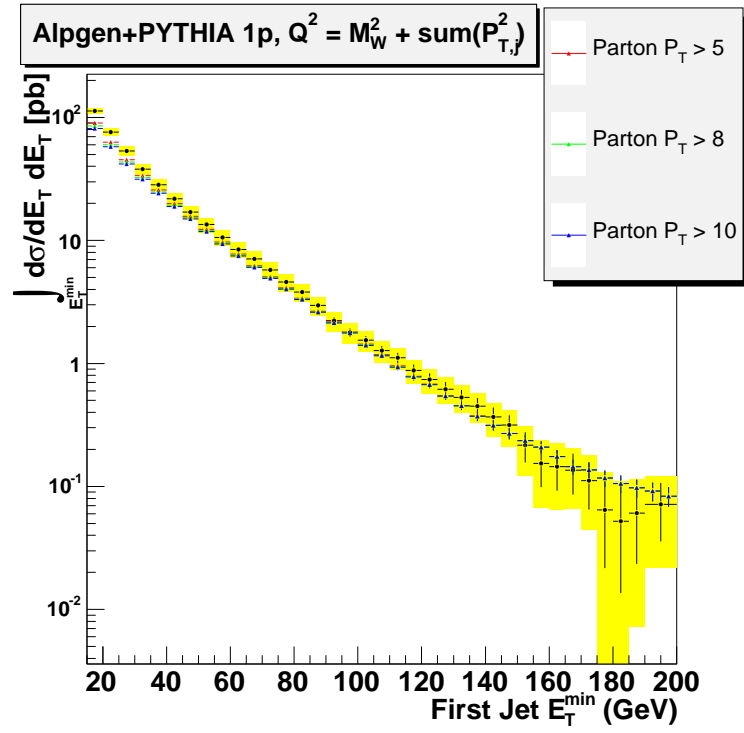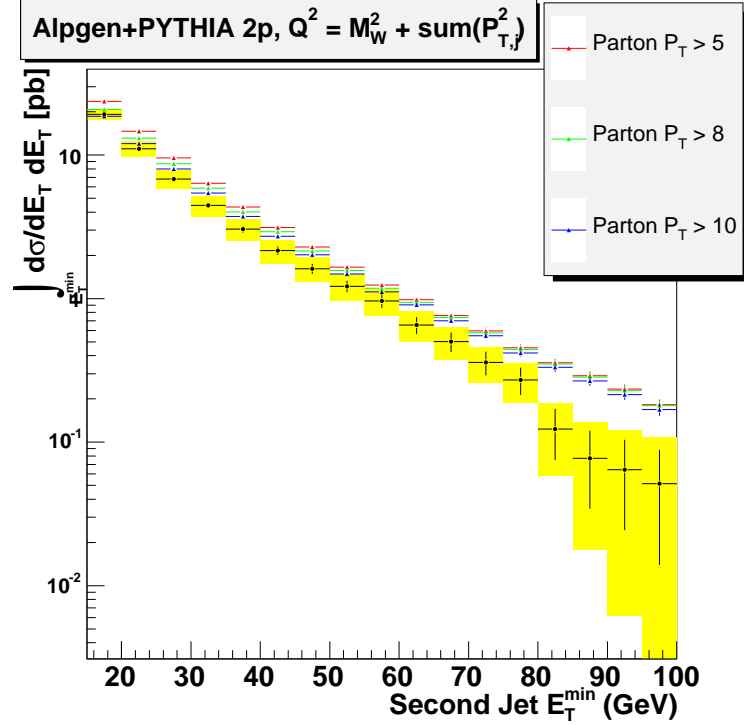
(b) $W \rightarrow e\upsilon + \geq 2$ jet

Figure 11.4: The $W \rightarrow e\upsilon + \geq 1$ jet inclusive cross-section prediction of an ALPGEN + PYTHIA 1 parton sample as a function of minimum jet E_T , compared to the data, and similarly for $W \rightarrow e\upsilon + \geq 2$ jet . In each the predictions for different parton P_T generation cuts are shown.
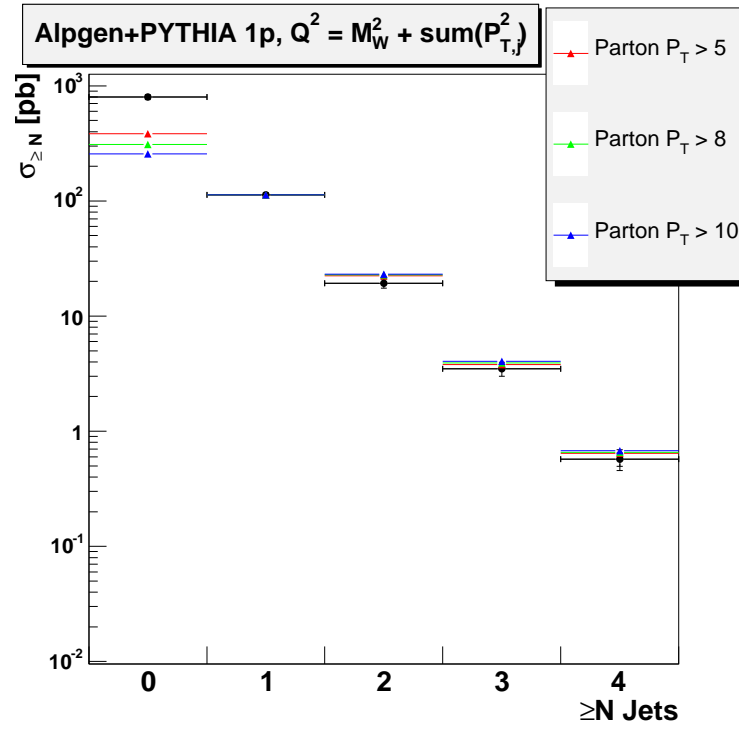
(a) $E_T^{jet} \geq 15$ GeV
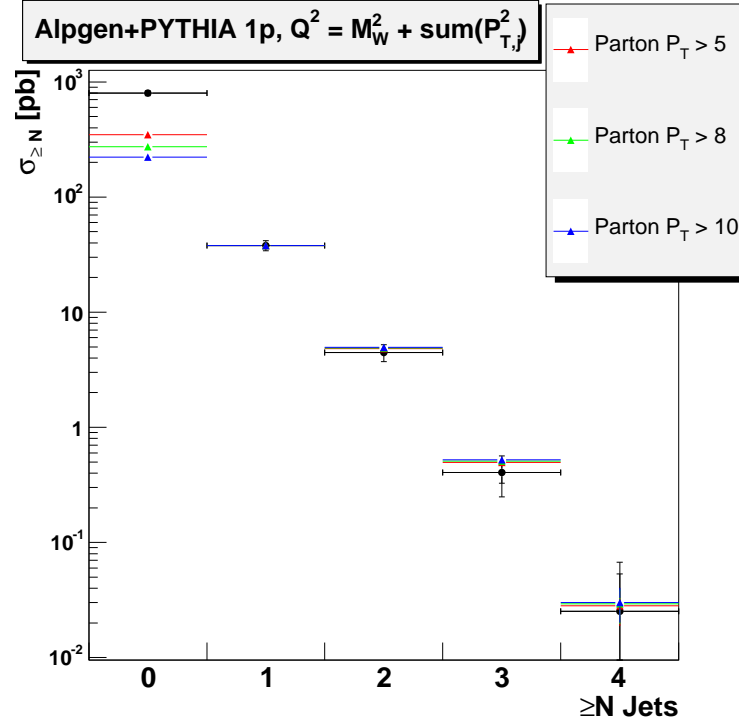
(b) $E_T^{jet} \geq 30$ GeV

Figure 11.5: The W$\rightarrow e\upsilon + \geq n$ jet cross-section prediction of the ALPGEN + PYTHIA 1 parton sample compared to our measured data, where the theory is normalised to the data ≥ 1 jet bin. The two plots are for two different minimum jet E_T cuts, and in each the predictions for different parton P_T generation cuts are shown.
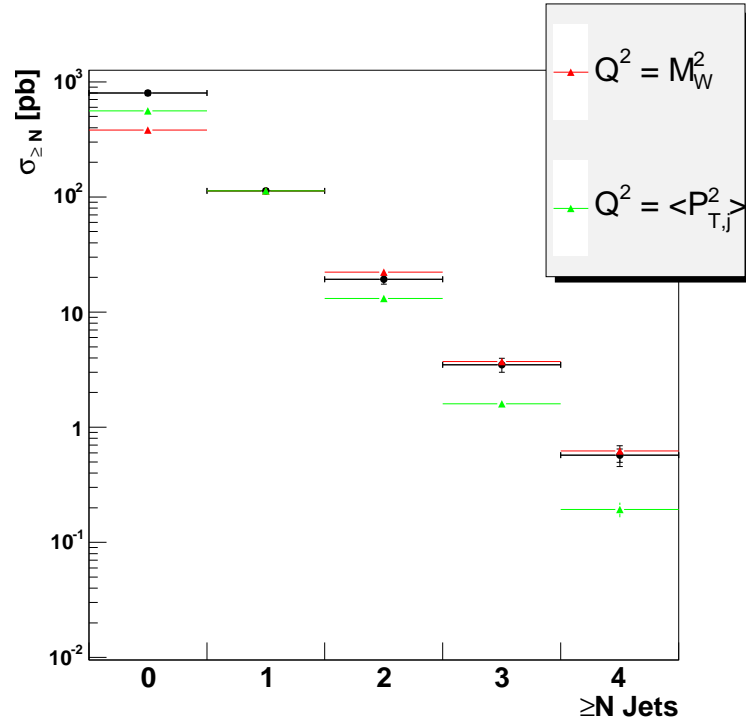
(a) ALPGEN + PYTHIA 1p
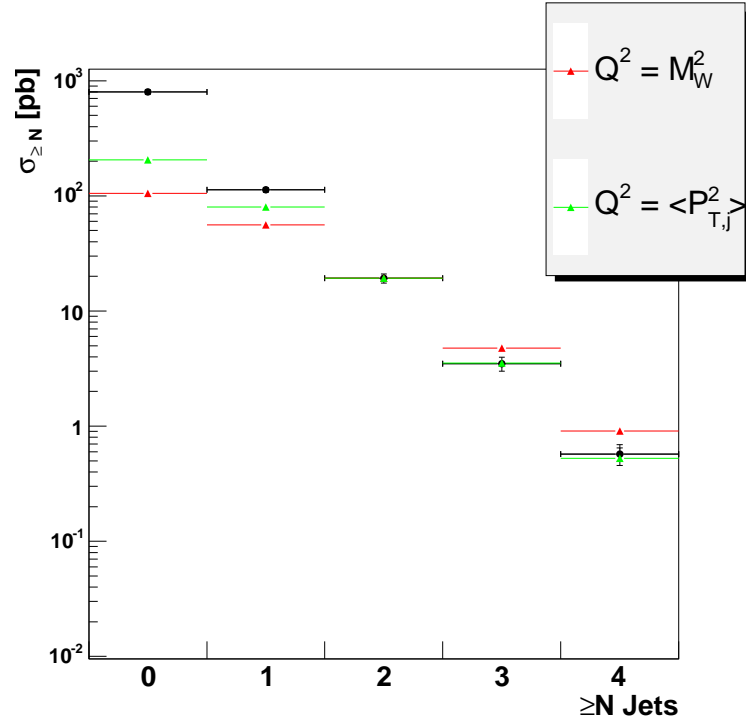
(b) ALPGEN + PYTHIA 2p

Figure 11.6: The W→ eυ + $\geq$ n jet cross-section predictions of ALPGEN + PYTHIA 1 parton and 2 parton samples compared to our measured data, where the theory is normalised to the data ≥ 1 jet and ≥ 2 jet bins respectively. The predictions for two different renormalisation scale choices are shown.

(a) $E_T^{jet} \geq 15$ GeV

(b) $E_T^{jet} \geq 30$ GeV

(c) $E_T^{jet} \geq 15$ GeV

(d) $E_T^{jet} \geq 30$ GeV
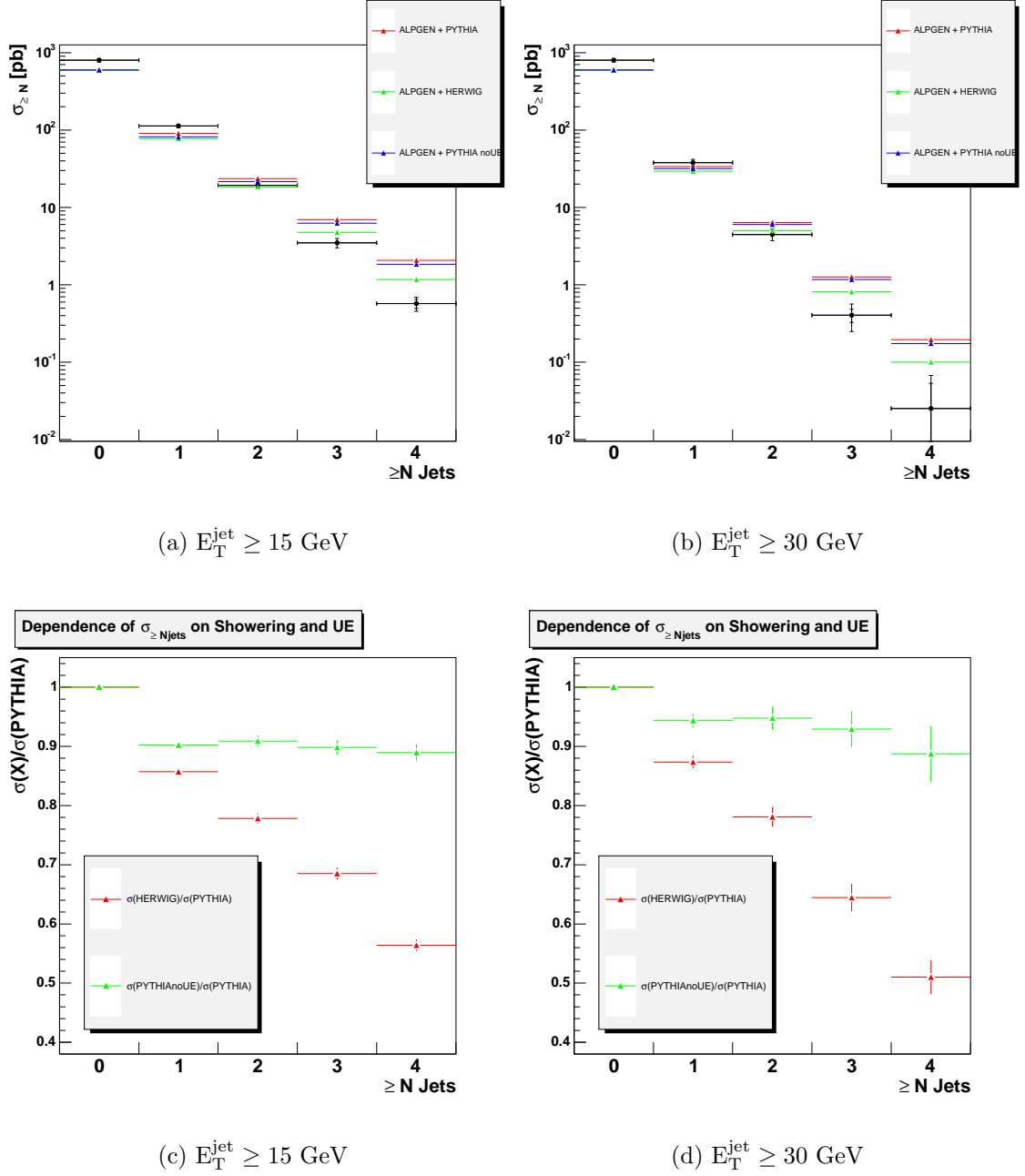
Figure 11.7: The W$\to$eυ + $\geq$ n jet cross-section predictions of ALPGEN + PYTHIA, ALPGEN + PYTHIA with underlying event disabled and ALPGEN + HERWIG n parton samples compared to our measured data, where the n parton sample is used in the $\geq n$ jet bin. The bottom plots show the ratio of the HERWIG and PYTHIA (no underlying event) predictions to those of standard PYTHIA.

(a) $E_T^{jet} \geq 15$ GeV

(b) $E_T^{jet} \geq 30$ GeV

(c) $E_T^{jet} \geq 15$ GeV

(d) $E_T^{jet} \geq 30$ GeV
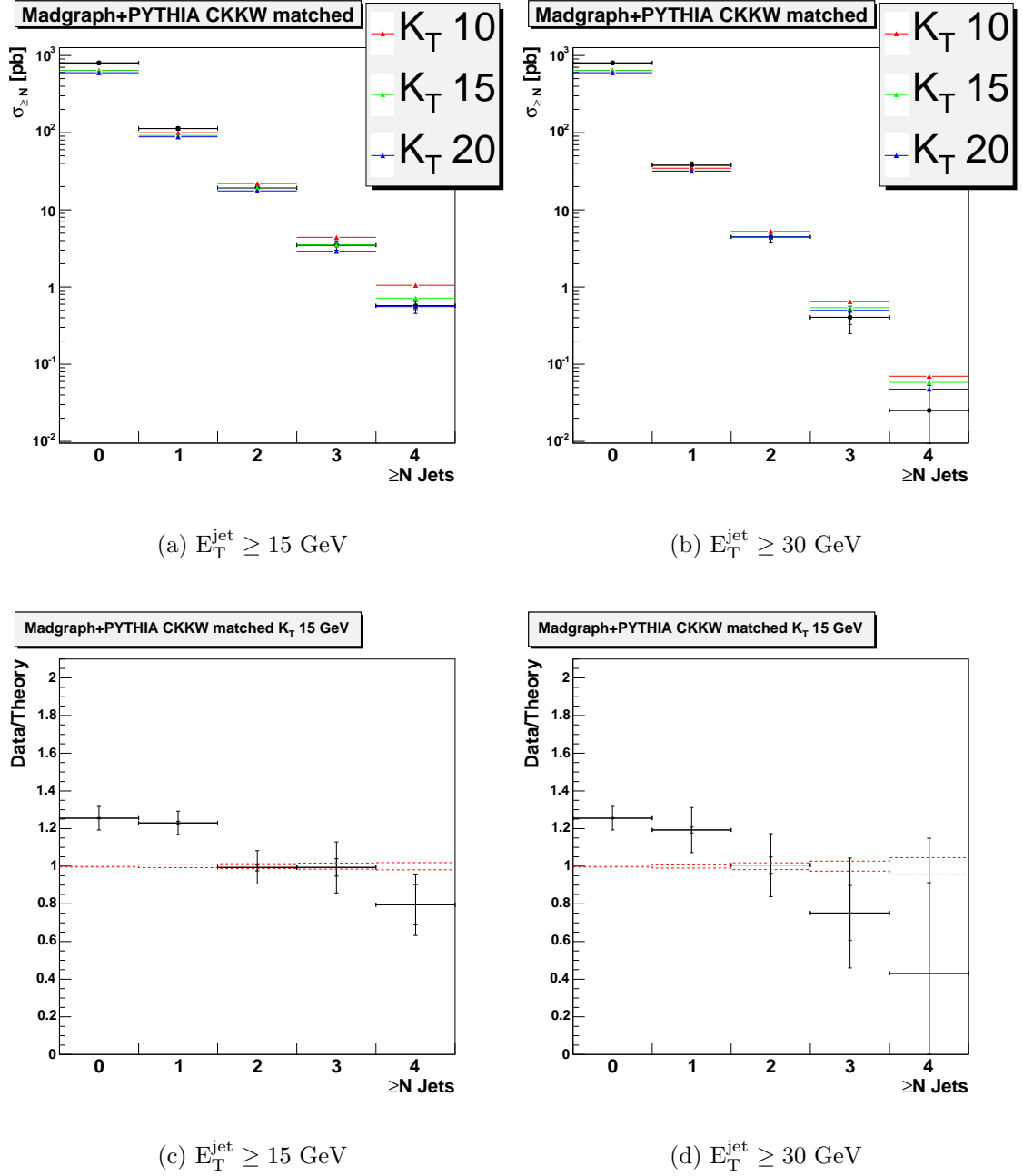
Figure 11.8: The top plots show the W$\rightarrow e\upsilon + \geq n$ jet cross-section prediction of the CKKW combined MADGRAPH + PYTHIA samples compared to our measured data. The two plots are for two different minimum jet E_T cuts, and in each the predictions for different parton K_T generation cuts are shown. The bottom plots show the Data/Theory ratio for a parton K_T generation cut of 15 GeV. The dotted red band indicates the statistical uncertainty on the Monte Carlo prediction for a ratio consistent with unity, the black error bars the data statistical (inner) and total statistical pluts systematic (outer) uncertainty.
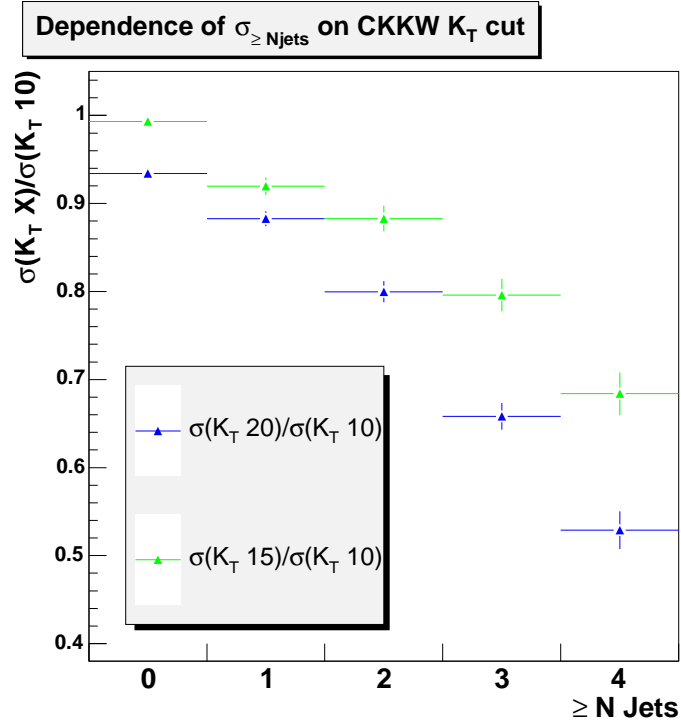
(a) $E_T^{jet} \geq 15$ GeV
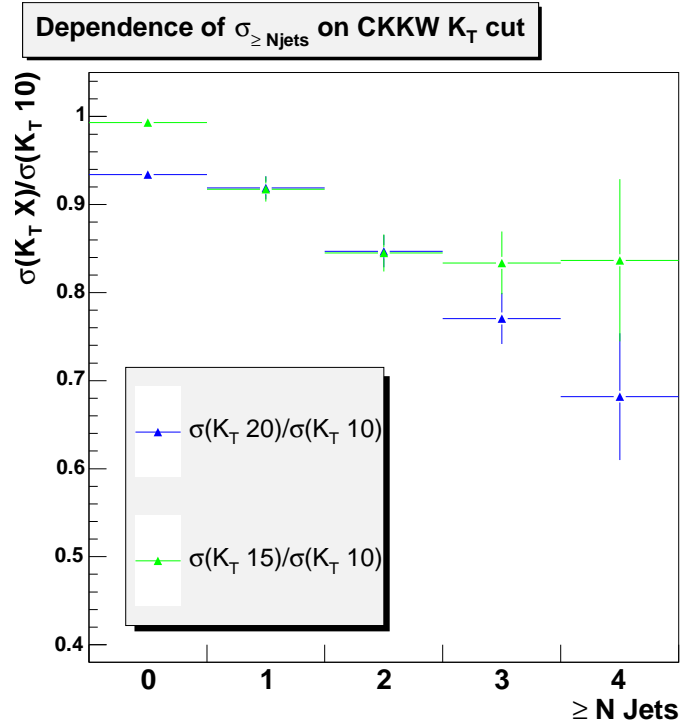
(b) $E_T^{jet} \geq 30$ GeV

Figure 11.9: The dependence of the MADGRAPH + PYTHIA CKKW cross-section prediction on the K_T generation cut.

(a) $E_T^{jet} \geq 15$ GeV

(b) $E_T^{jet} \geq 30$ GeV

(c) $E_T^{jet} \geq 15$ GeV

(d) $E_T^{jet} \geq 30$ GeV
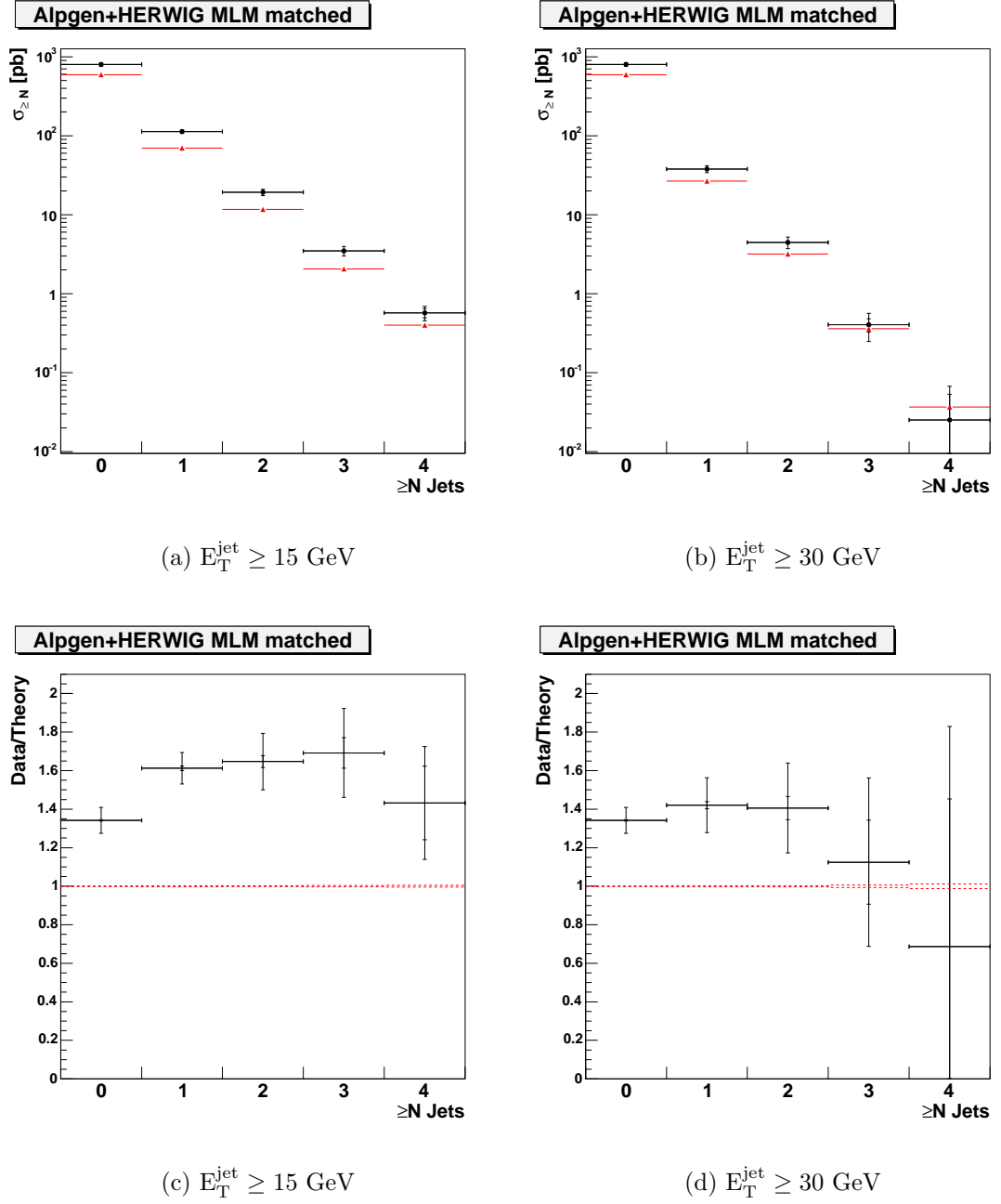
Figure 11.10: The top plots show the W$\rightarrow$ eυ + $\geq n$ jet cross-section prediction of the MLM combined ALPGEN + HERWIG samples compared to our measured data. The two plots are for two different minimum jet E_T cuts. The bottom plots show the Data/Theory ratio.

(a) $E_T^{jet} \geq 15$ GeV

(b) $E_T^{jet} \geq 30$ GeV

(c) $E_T^{jet} \geq 15$ GeV

(d) $E_T^{jet} \geq 30$ GeV
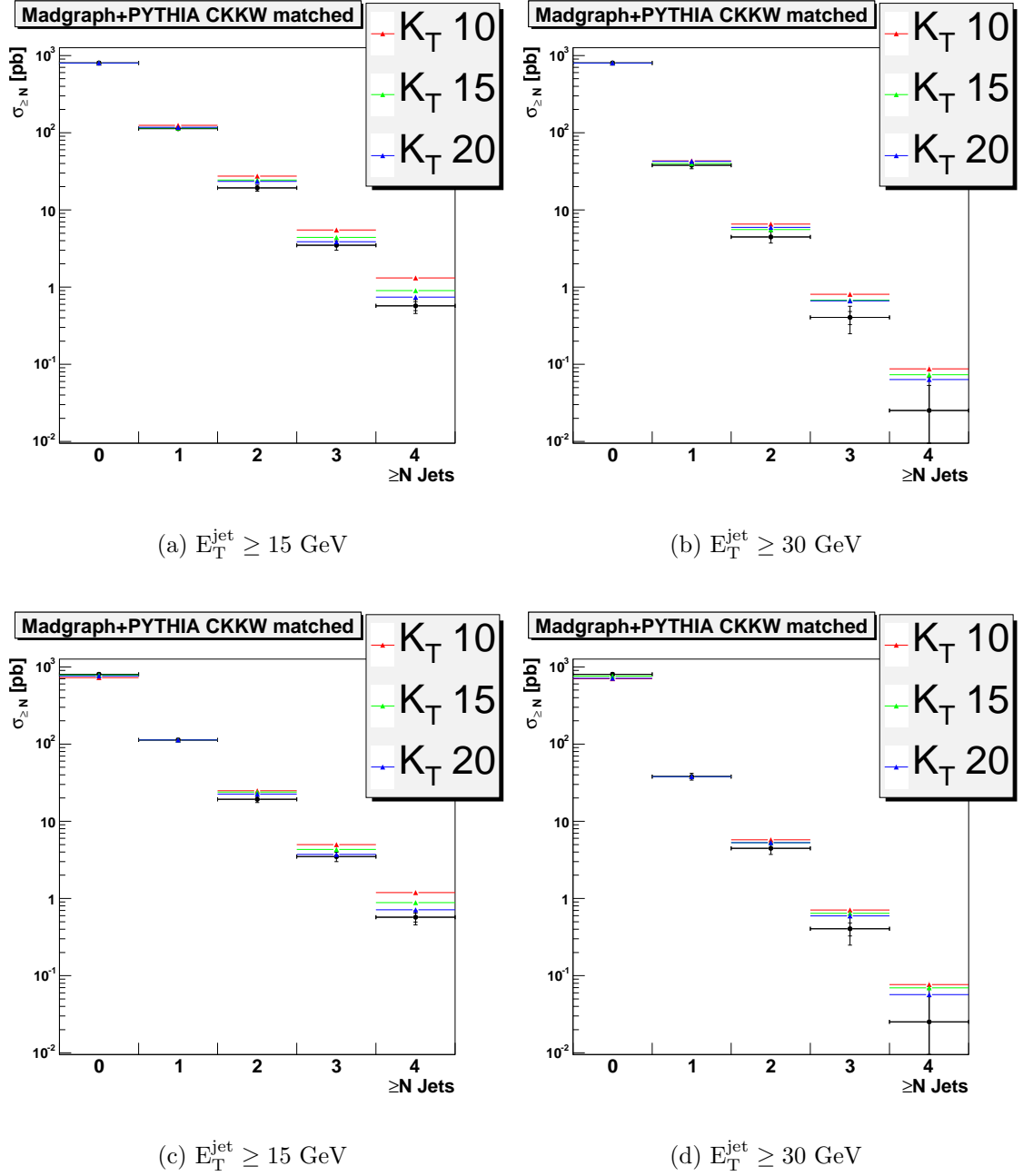
Figure 11.11: The MADGRAPH + PYTHIA CKKW W$\rightarrow$ e$\upsilon + \geq n$ jet predictions where the theory is normalised to the measured ≥ 0 jet (top plots) cross-section and ≥ 1 jet (bottom plots) cross-section. The results for two different jet E_T thresholds are shown.

(a) $E_T^{jet} \geq 15$ GeV

(b) $E_T^{jet} \geq 30$ GeV

(c) $E_T^{jet} \geq 15$ GeV

(d) $E_T^{jet} \geq 30$ GeV
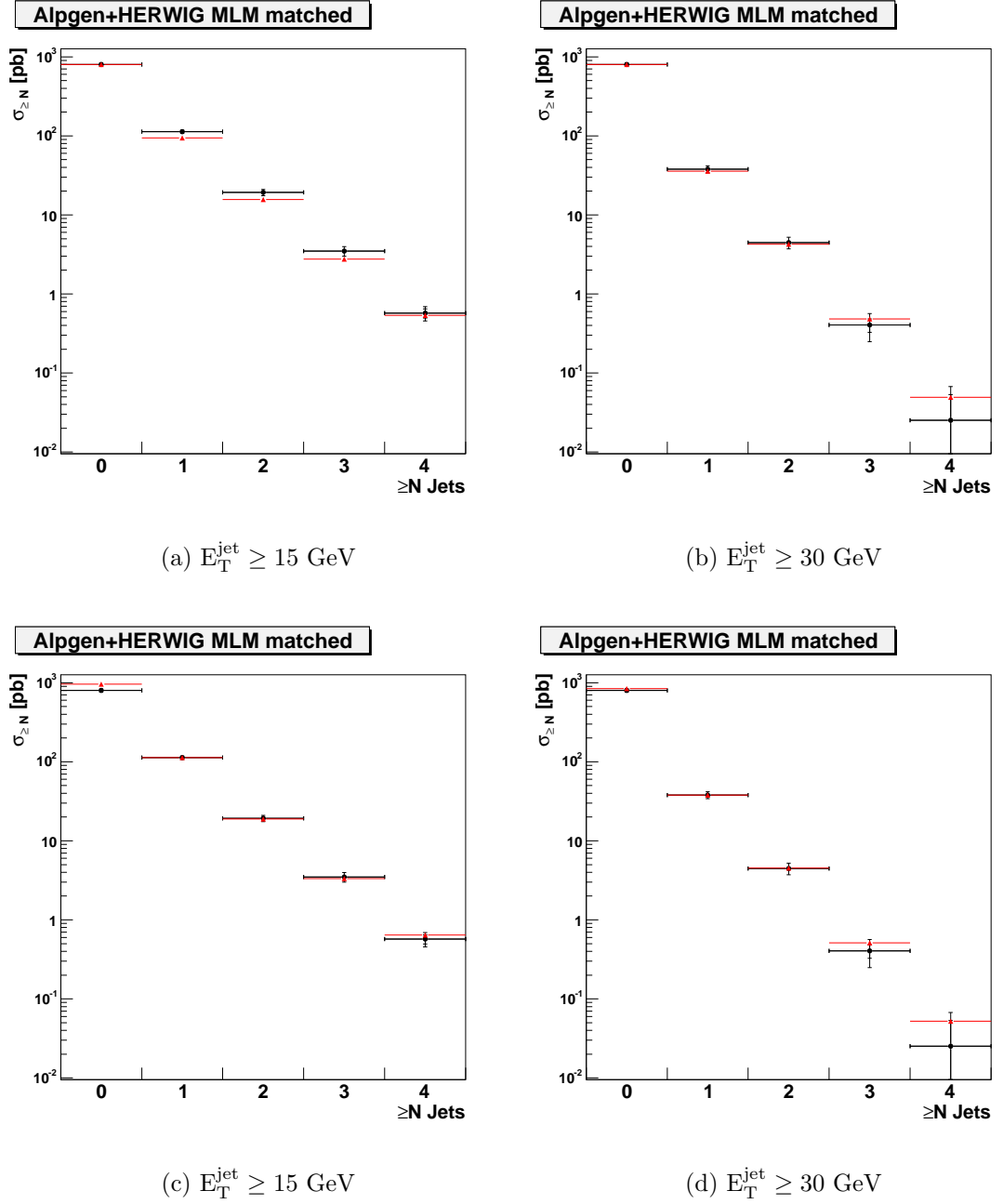
Figure 11.12: The ALPGEN + HERWIG MLM W$\to$ eυ + $\geq$ n jet predictions where the theory is normalised to the measured ≥ 0 jet (top plots) cross-section and ≥ 1 jet (bottom plots) cross-section. The results for two different jet E_T thresholds are shown.

(a) First Jet E_T

(b) Second Jet E_T

(c) Third Jet E_T
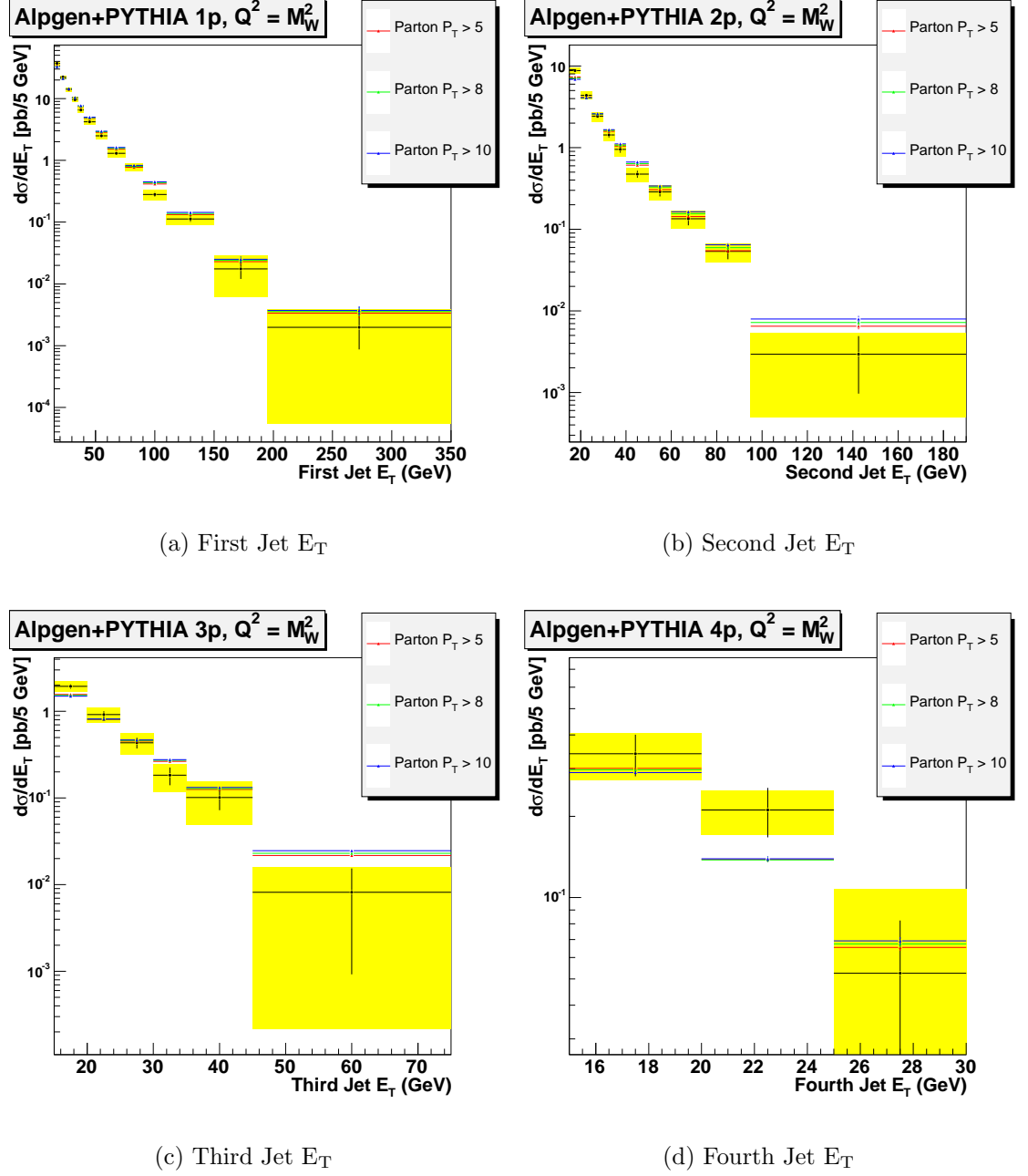
(d) Fourth Jet E_T

Figure 11.13: The first, second, third and fourth jet E_T differential cross-section distributions, along with the ELO ALPGEN + PYTHIA predictions, where the n parton sample is used in the $\geq n$ jet bin. The predictions using parton P_T generation cuts of 5, 8 and 10 GeV are shown. The black error bars represent the data statistical error, and the yellow band the total systematic error.

(a) $Q^2 = M_W^2$

(b) $Q^2 = \sum_{P_{T,j}^2}$

(c) $Q^2 = M_W^2$

(d) $Q^2 = \sum_{P_{T,j}^2}$
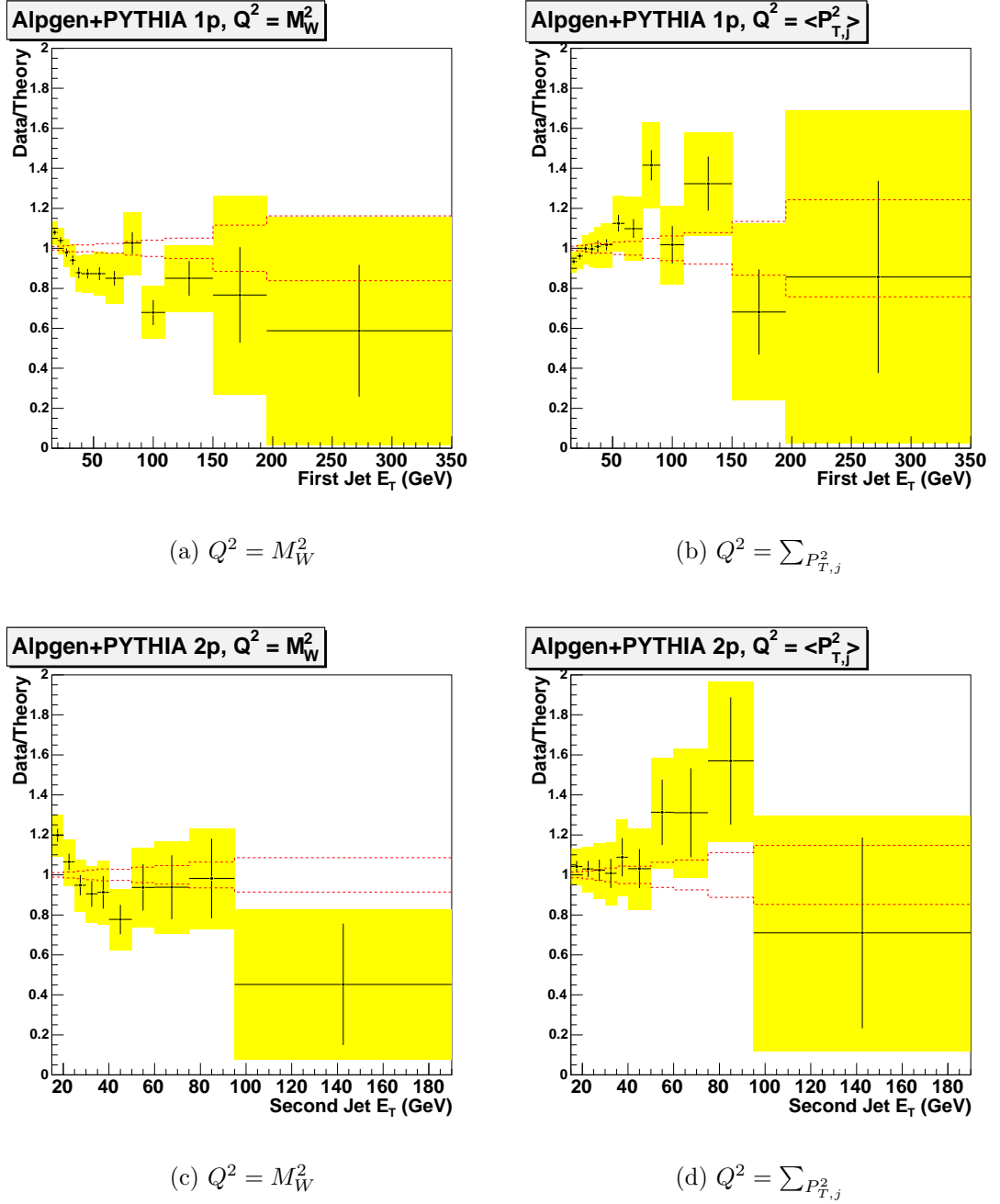
Figure 11.14: The ratio of the data first jet E_T and second jet E_T differential cross-section measurements to the predictions using ELO ALPGEN + PYTHIA 1 parton and 2 parton samples respectively with a parton P_T generation cut of 5 GeV. The prediction using two different renormalisation scales is shown. The dotted red band indicates the statistical uncertainty on the Monte Carlo prediction for a ratio consistent with unity, the black error bars the data statistical error and the yellow band the data systematic error.

(a) $Q^2 = M_W^2$

(b) $Q^2 = \sum_{P_{T,j}^2}$

(c) $Q^2 = M_W^2$

(d) $Q^2 = \sum_{P_{T,j}^2}$
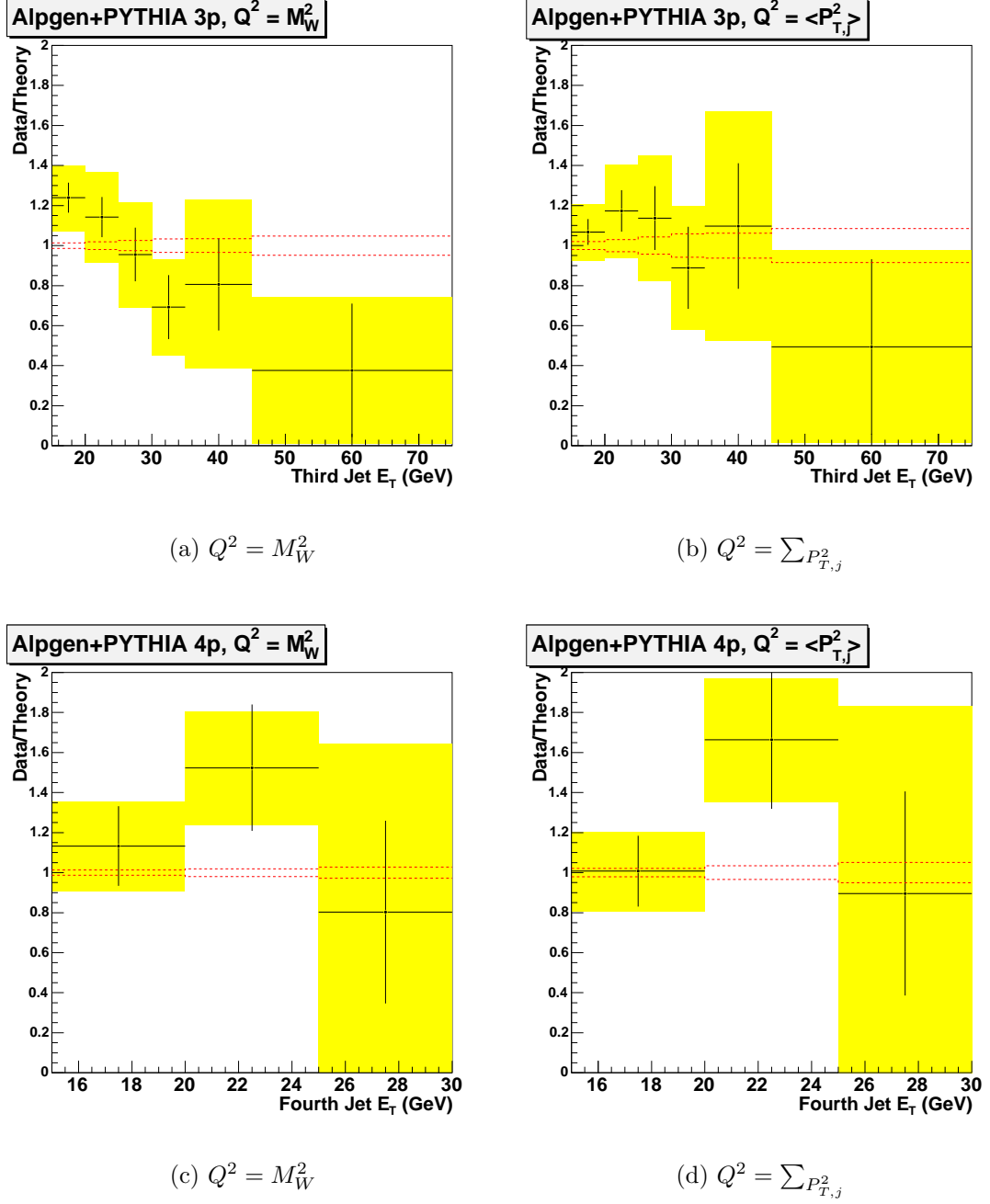
Figure 11.15: The ratio of the data third jet and fourth jet E_T differential cross-section measurements to the predictions using ELO ALPGEN + PYTHIA 3 parton and 4 parton samples respectively with a parton P_T generation cut of 5 GeV. The prediction using two different renormalisation scales is shown.
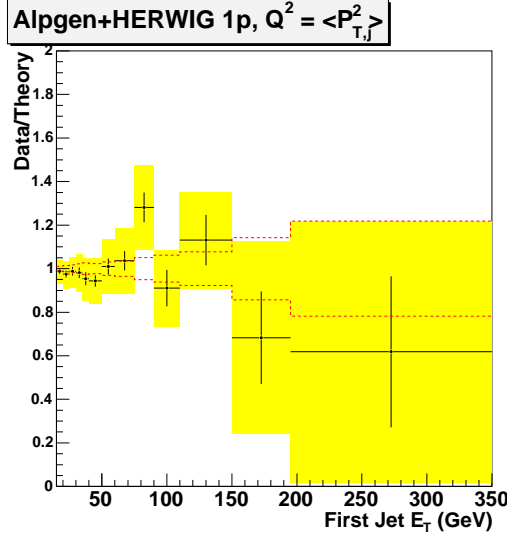
(a) First Jet E$_T$
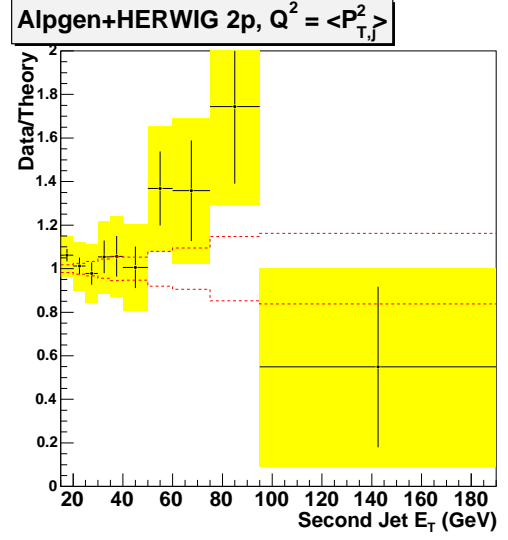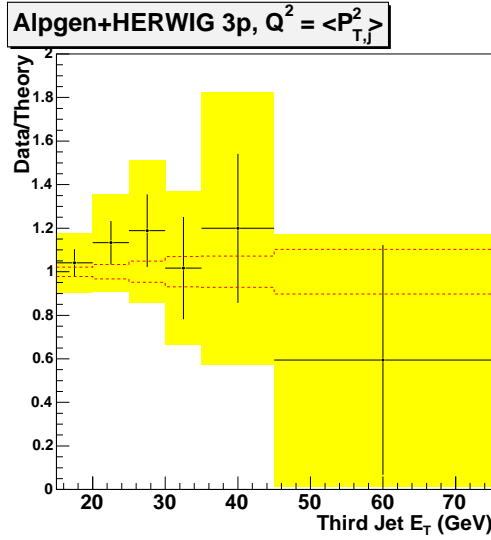
(b) Second Jet E$_T$

(c) Third Jet E$_T$
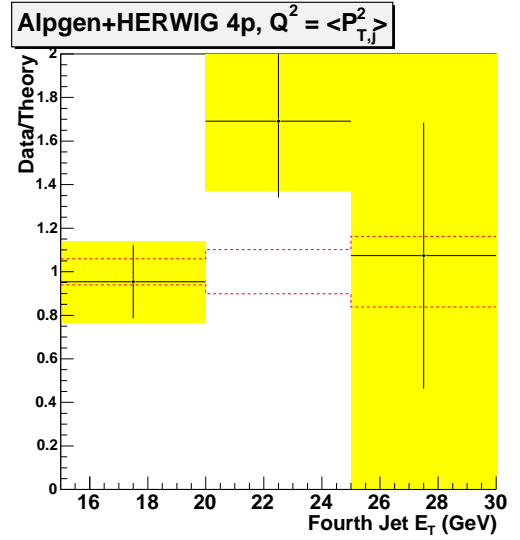
(d) Fourth Jet E$_T$

Figure 11.16: The ratio of the data first, second, third and fourth jet E$_T$ differential cross-section measurement to the ELO ALPGEN + HERWIG predictions using $Q^2 = <P_{T,j}^2>$, where the n parton sample is used in the $\geq n$ jet bin.
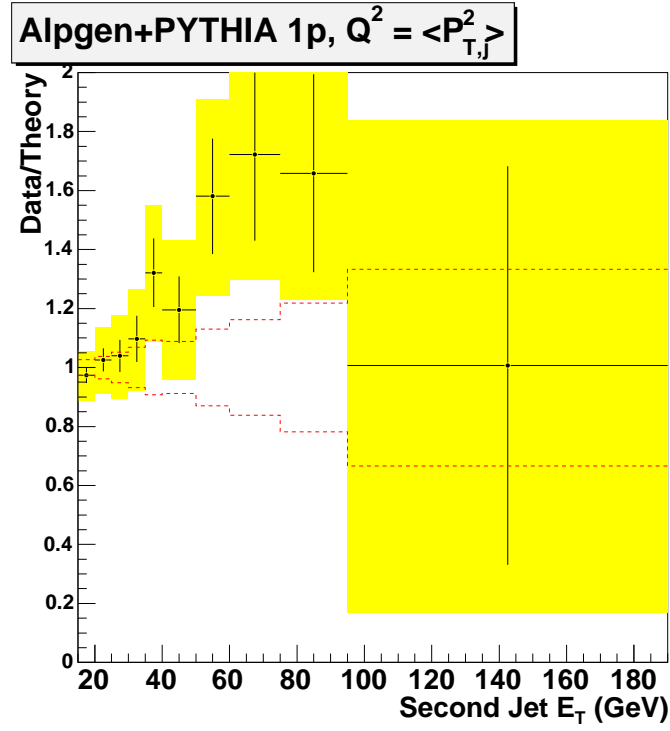
(a) W→ eν + 1p Data/Theory
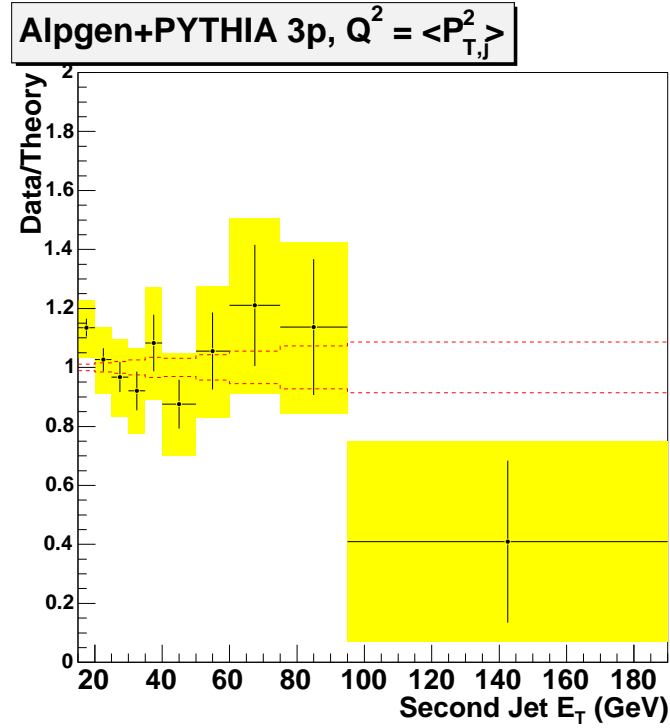
(b) W→ eν + 3p Data/Theory

Figure 11.17: Comparison of the data second jet E_T differential cross-section measurement to the prediction using ELO ALPGEN + PYTHIA 1 and 3 parton samples.
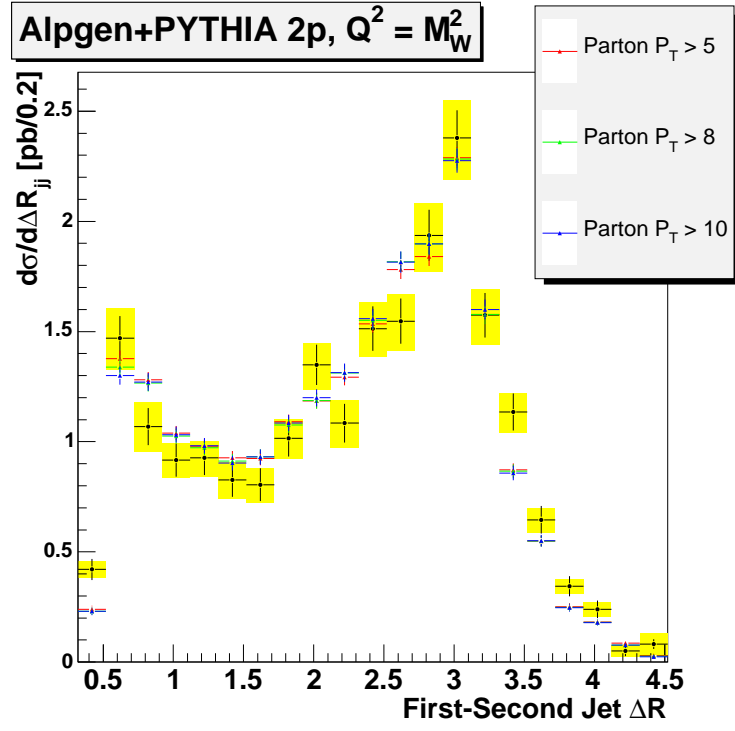
(a) ΔR_{jj}
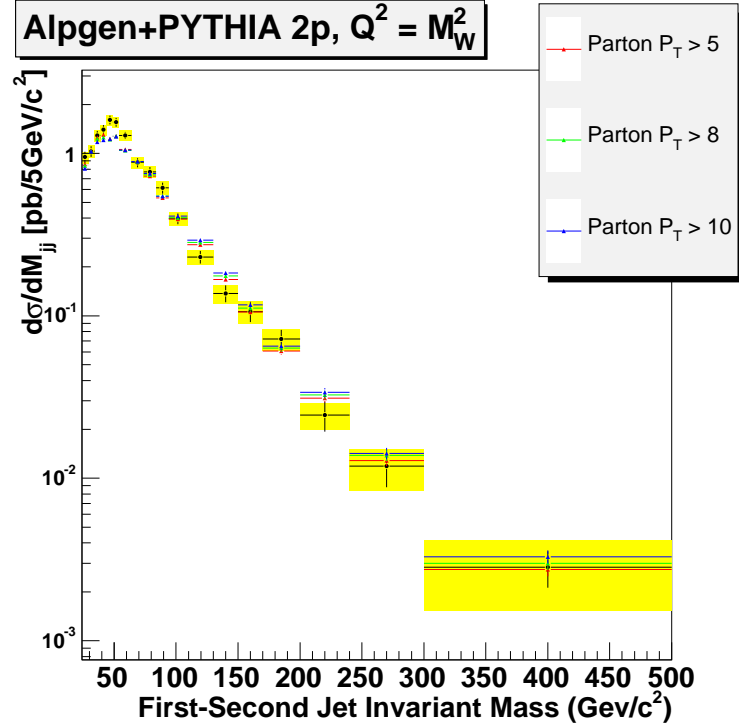
(b) M_{jj}

Figure 11.18: The ΔR_{jj} and M_{jj} differential cross-section distributions, along with the ELO ALPGEN + PYTHIA 2 parton predictions for different parton generation cuts.
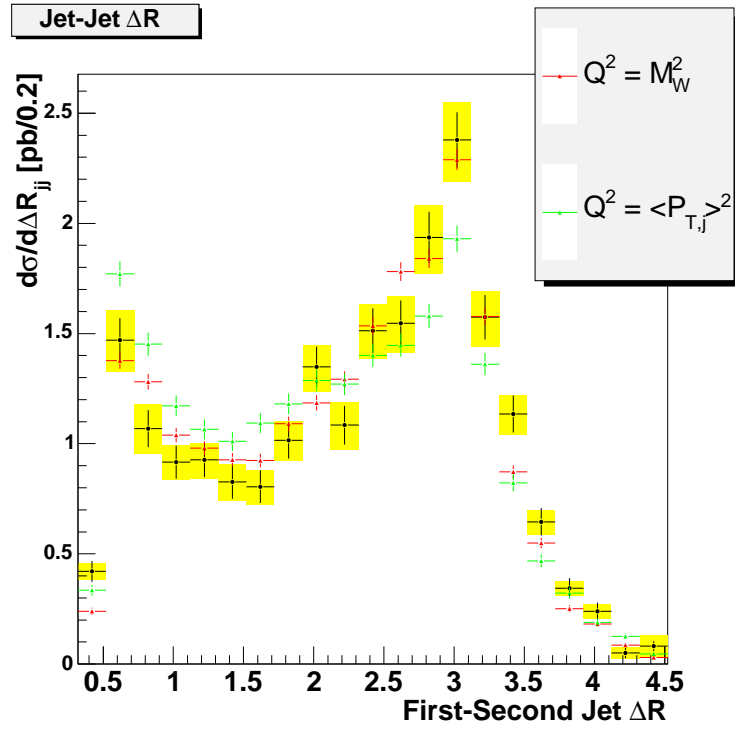
(a) ΔR_{jj}
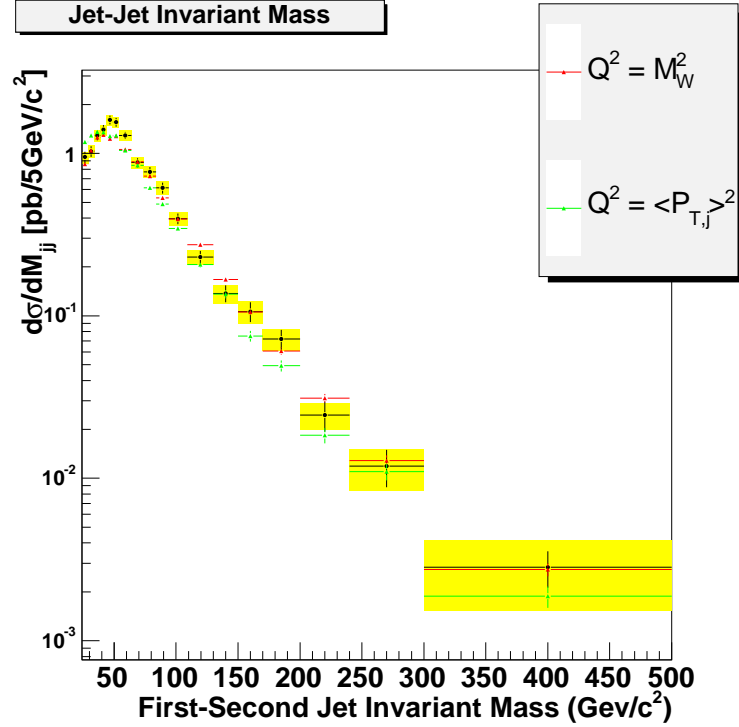
(b) M_{jj}

Figure 11.19: The ΔR_{jj} and M_{jj} differential cross-section distributions, along with the ELO ALPGEN + PYTHIA 2 parton predictions for two different renormalisation scales.
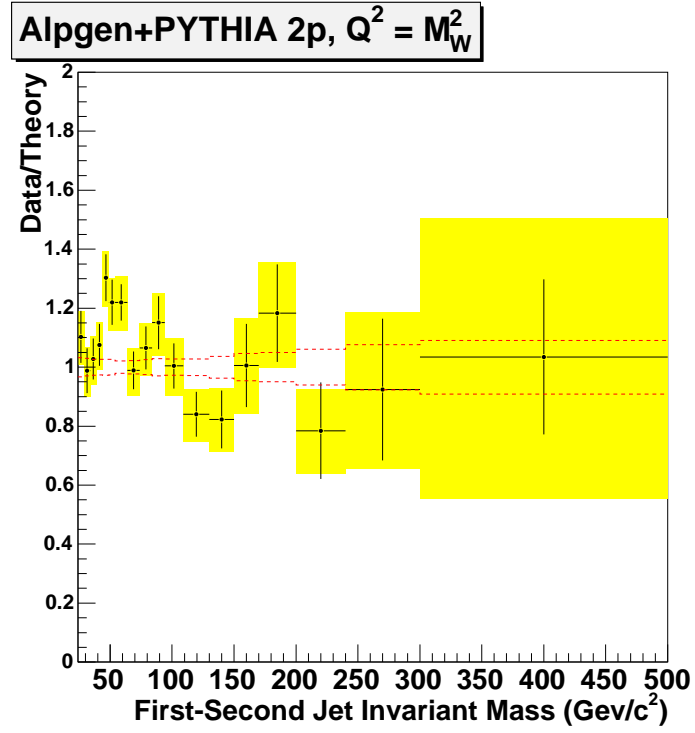
(a) $Q^2 = M_W^2$
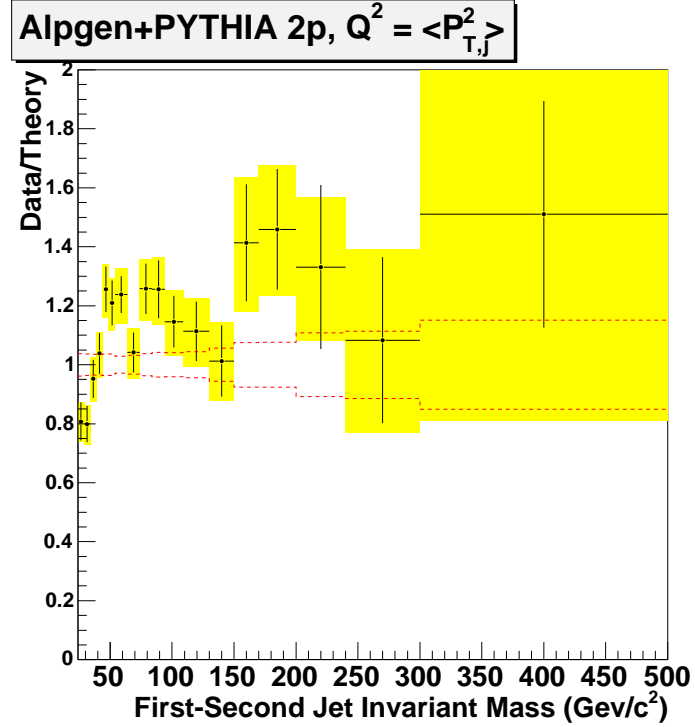
(b) $Q^2 = < P_{T,j}^2 >$

Figure 11.20: The ratio of the data M_{jj} differential cross-section measurement to the prediction using the ELO ALPGEN + PYTHIA 2 parton sample with two different renormalisation scales.

(a) First Jet E_T

(b) Second Jet E_T

(c) Third Jet E_T
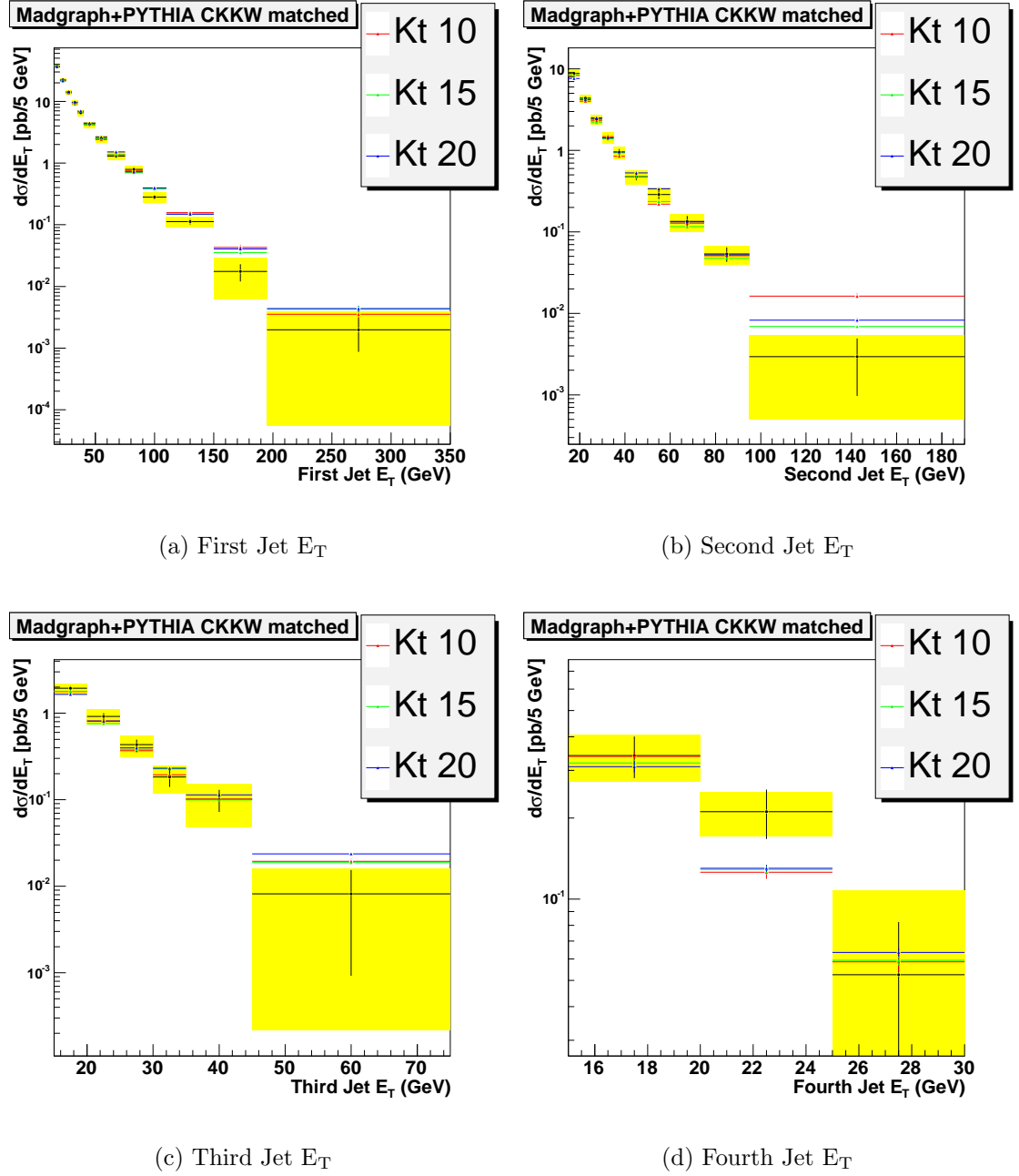
(d) Fourth Jet E_T

Figure 11.21: The first, second, third and fourth jet E_T differential cross-section distributions, along with the MADGRAPH + PYTHIA CKKW predictions, where the theory is normalised to the inclusive data cross-section in the relevant $\geq n$ jet sample. The prediction for three different K_T CKKW generation cuts is shown.
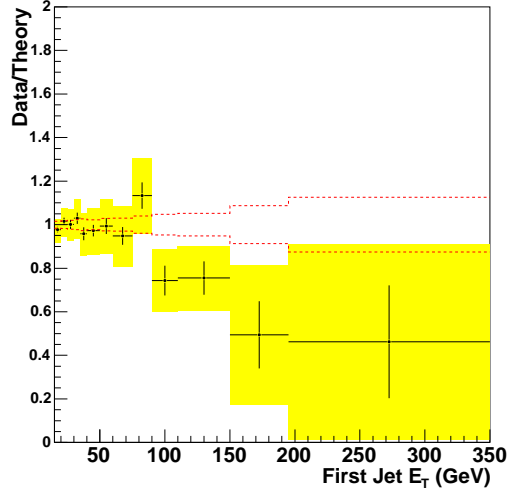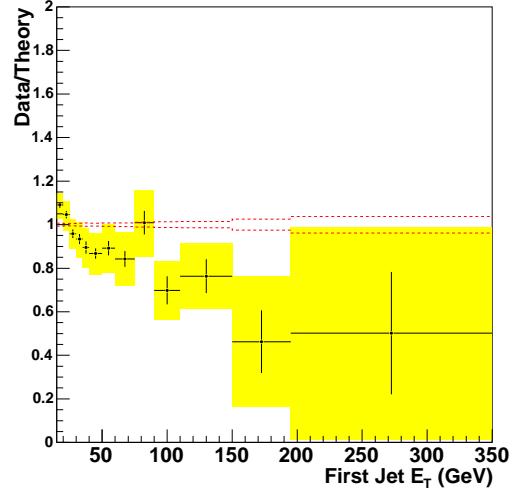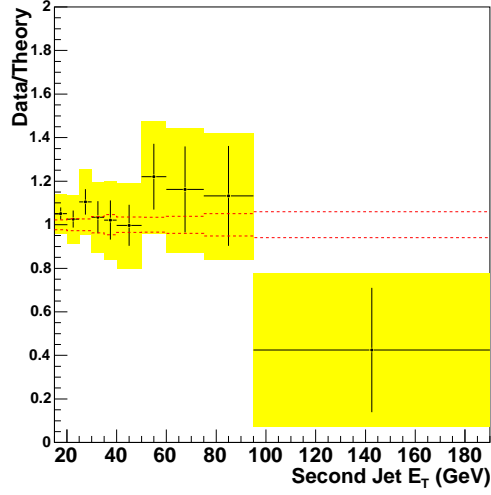
(a) MADGRAPH + PYTHIA CKKW

(b) ALPGEN + HERWIG MLM

(c) MADGRAPH + PYTHIA CKKW
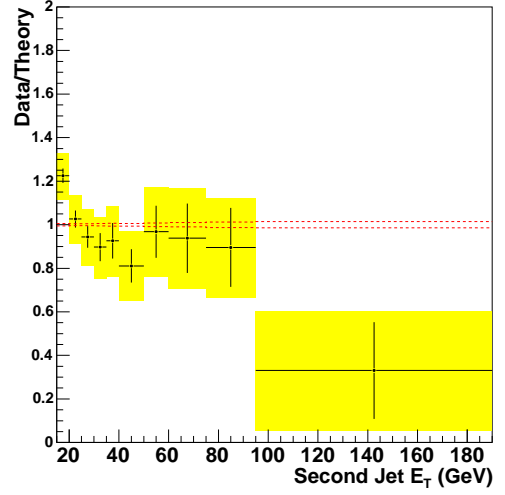
(d) ALPGEN + HERWIG MLM

Figure 11.22: The ratio of the data first and second jet E_T differential cross-section measurements to the predictions of the combined MADGRAPH + PYTHIA CKKW samples and the ALPGEN + HERWIG MLM samples.
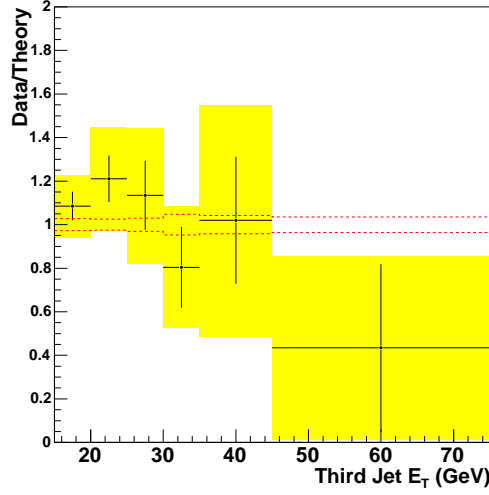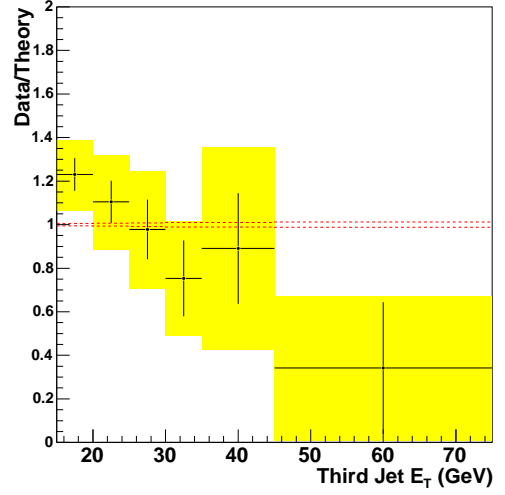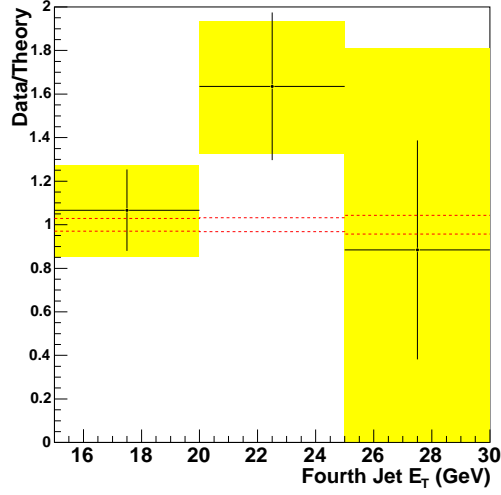
(a) MADGRAPH + PYTHIA CKKW

(b) ALPGEN + HERWIG MLM

(c) MADGRAPH + PYTHIA CKKW
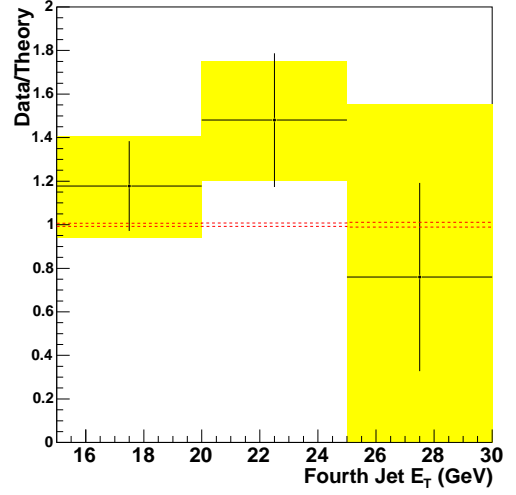
(d) ALPGEN + HERWIG MLM

Figure 11.23: The ratio of the data third jet E_T differential cross-section measurement to the predictions of the combined MADGRAPH + PYTHIA CKKW samples and the ALPGEN + HERWIG MLM samples.
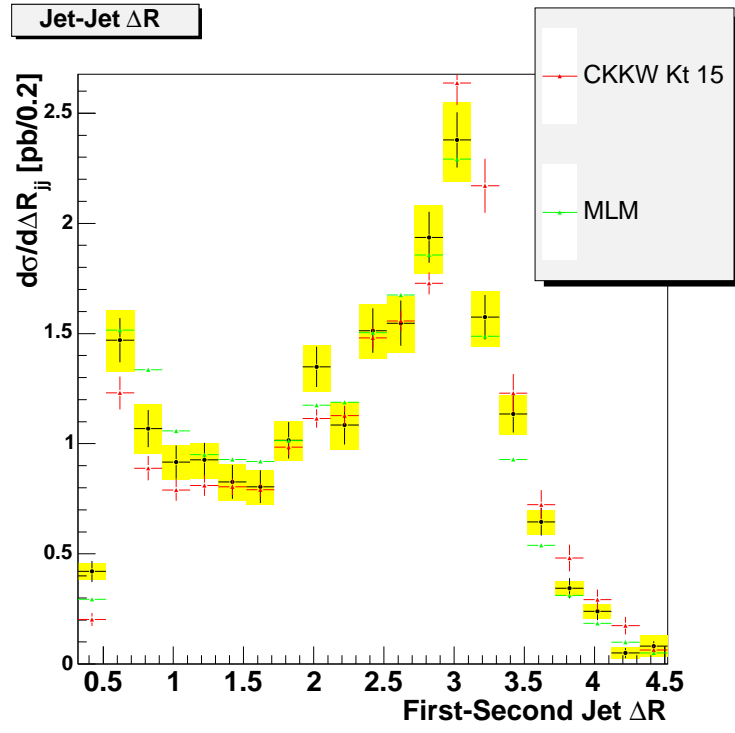
(a) ΔR_{jj}
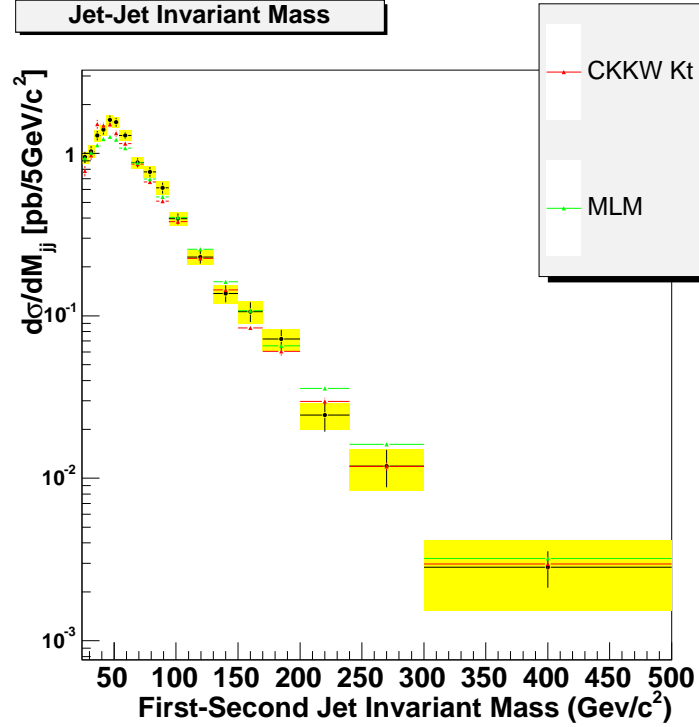
(b) M_{jj}

Figure 11.24: The ΔR_{jj} and M_{jj} differential cross-section distributions, along with the MADGRAPH + PYTHIA CKKW and ALPGEN + HERWIG MLM matched sample predictions.
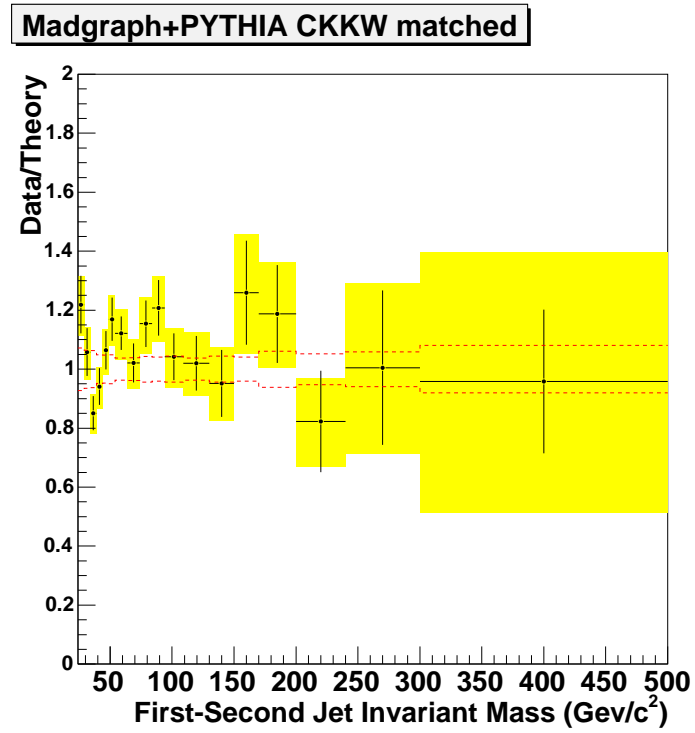
(a) MADGRAPH + PYTHIA CKKW
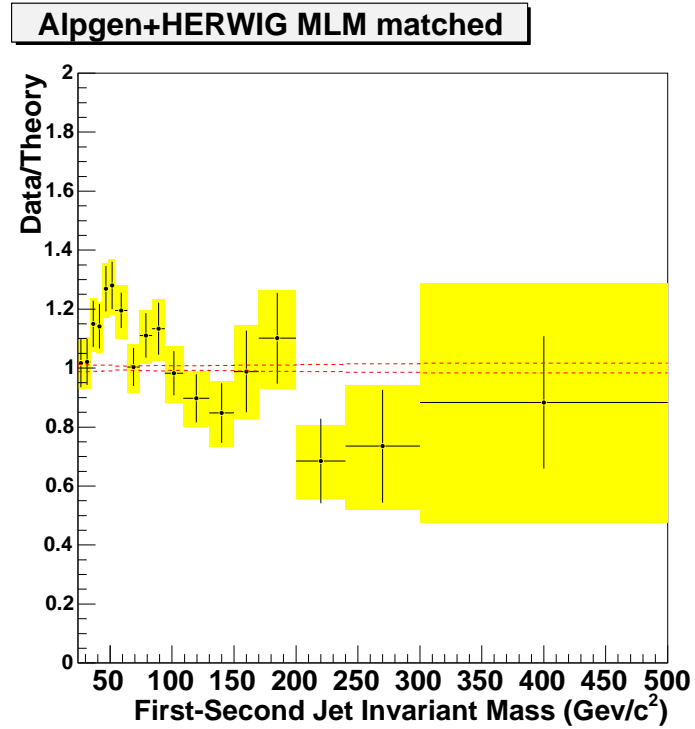
(b) ALPGEN + HERWIG MLM

Figure 11.25: The ratio of the data M_{jj} differential cross-section measurement to the predictions using the MADGRAPH + PYTHIA CKKW and ALPGEN + HERWIG MLM matched samples.

Chapter 12

Summary and Outlook

In this thesis are presented the first differential measurements of the W$\rightarrow$eυ + $\geq$ n jet production cross-section at the hadron-level with respect to jet E$_\text{T}$, ΔR$_\text{jj}$ and M$_\text{jj}$ variables, in which all known detector effects have been accounted for. Such a measurement allows detailed comparison of W + Jets data with the predictions of Monte Carlo event generators without the need for a CDF detector simulation.

This differential measurement has involved several significant departures from the way in which more inclusive W + Jets measurements have previously been made [20] [107]. The cross-section has not been corrected back to the full W decay acceptance, but to a restricted decay phase space, since this reduces the dependance of the measurement on theoretical models of the signal process. The move to a differential measurement necessitates improved kinematic modelling of the QCD multijet background, and this is achieved via the formation of an "antielectron" sample using the signal dataset, which is shown to provide an excellent description of event kinematics in several W-related variables. In order to account for the effects of calorimeter energy resolution on the measured jet spectrum we apply additional unsmearing factors in the cross-section definition which are derived in a model independent fashion using detector simulated signal Monte Carlo samples.

The comparisons of Enhanced Leading Order Monte Carlo event samples to our measurements show limitations in the ability of the ELO approach to make absolute rate predictions, but demonstrate that certain relative rates, such as the shape of the differential cross-section, can be well modelled. The comparisons to the CKKW and MLM combined ME-PS matched samples show promise in terms of delivering a single W + Jets Monte Carlo event sample capable of modelling any distribution in any jet multiplicity sample, but also indicate that more work may be needed to fine tune these approaches.

There are several possible extensions to this work. At the time of writing there is more than 1fb^{-1} of data available for study and this, coupled with the inclusion of the muon W decay channel, would greatly improve the statistical reach of the

analysis. In particular this would allow for more precise measurements at high jet E_T , which would be very useful and relevant in the coming LHC era. Additionally this $1fb^{-1}$ dataset allows for the possibility of examining $Z +$ Jets events to a similar level of detail. Such an analysis would be very complimentary to this one, being sensitive to the same QCD physics, but in a much cleaner channel with significantly reduced backgrounds. Also, as was discussed in Chapter 3, the Midpoint and K_T jet clustering algorithms are preferable to the seeded cone algorithm from the point of view of making a measurement which is not sensitive to soft QCD processes which are beyond the reach of theoretical predictions.

Appendix A

Background Results

n Jets	Total	QCD		$t\bar{t}$		$Z \to ee$		$W \to \tau \upsilon$		WW	
		N	%	N	%	N	%	N	%	N	%
≥ 0	147008	2555 ±96	1.7 %	117 ±1.0	0.1 %	213 ±4	0.1 %	2317 ±99	1.6 %	169 ±1	0.1 %
≥ 1	22535	1074 ±54	4.8 %	117 ±0.9	0.5 %	33 ±1	0.1 %	388 ±23	1.7 %	140 ±1	0.6 %
≥ 2	4402	336 ±29	7.6 %	114 ±0.9	2.6 %	11 ±0	0.3 %	77 ±5	1.7 %	92 ±1	2.1 %
≥ 3	885	72 ±13	8.2 %	95 ±0.9	10.7 %	4 ±0	0.5 %	14 ±1	1.6 %	26 ±0	2.9 %
≥ 4	208	41 ±11	20.0 %	60 ±0.7	29.0 %	1 ±0	0.3 %	2 ±0	1.0 %	6 ±0	2.8 %

Table A.1: Results of inclusive background calculation. Shown are the total number of candidate events in each jet multiplicity bin and the estimated number of background events, both as an absolute number and as a fraction of the candidates. Errors shown are those resulting from limited background template statistics only.

n Jets	Total	QCD		$t\bar{t}$		$Z \to ee$		$W \to \tau \upsilon$		WW	
		N	%	N	%	N	%	N	%	N	%
≥ 0	147002	2192 ±89	1.5 %	117 ±1.0	0.1 %	295 ±7	0.2 %	2135 ±91	1.5 %	169 ±1	0.1 %
≥ 1	22529	1026 ±55	4.6 %	117 ±0.9	0.5 %	58 ±1	0.3 %	548 ±24	2.4 %	140 ±1	0.6 %
≥ 2	4398	330 ±30	7.5 %	114 ±0.9	2.6 %	21 ±1	0.5 %	107 ±5	2.4 %	92 ±1	2.1 %
≥ 3	882	69 ±14	7.9 %	95 ±0.9	10.7 %	4 ±0	0.5 %	17 ±1	2.0 %	26 ±0	2.9 %
≥ 4	207	42 ±11	20.3 %	60 ±0.7	29.0 %	0 ±0	0.2 %	2 ±0	1.0 %	6 ±0	2.8 %

Table A.2: Results of inclusive background calculation when the Monte Carlo samples are skewed such that the $W \to e\upsilon$, $W \to \tau \upsilon$ and $Z \to ee$ $n + 1$ parton samples are used in the $\geq n$ jet bin.

n **Jets**	**Total**	**QCD**		$t\bar{t}$		$Z \to ee$		$W \to \tau\upsilon$	
		N	%	N	%	N	%	N	%
≥ 1	22529	1108 ±55	4.9 %	119 ±1.0	0.5 %	37 ±1	0.2 %	417 ±24	1.9 %
≥ 2	4398	351 ±30	8.0 %	116 ±0.9	2.6 %	12 ±0	0.3 %	77 ±5	1.8 %
≥ 3	882	76 ±14	8.6 %	97 ±0.9	11.0 %	4 ±0.1	0.5 %	14 ±1.1	1.6 %
≥ 4	207	43 ±11	20.7 %	62 ±0.7	29.7 %	1 ±0.0	0.3 %	2 ±0.2	1.0 %

Table A.3: Results of inclusive background calculation when Monte Carlo missing E_T scale is systematically shifted by +5%.

n **Jets**	**Total**	**QCD**		$t\bar{t}$		$Z \to ee$		$W \to \tau\upsilon$	
		N	%	N	%	N	%	N	%
≥ 1	22529	1054 ±53	4.7 %	114 ±0.9	0.5 %	29 ±1	0.1 %	361 ±22	1.6 %
≥ 2	4398	329 ±28	7.5 %	111 ±0.9	2.5 %	10 ±0	0.2 %	73 ±5	1.7 %
≥ 3	882	70 ±13	8.0 %	92 ±0.8	10.4 %	4 ±0.1	0.4 %	14 ±1.1	1.6 %
≥ 4	207	41 ±11	19.7 %	58 ±0.7	28.1 %	1 ±0.0	0.3 %	2 ±0.2	1.0 %

Table A.4: Results of inclusive background calculation when Monte Carlo missing E_T scale is systematically shifted by -5%.

Bin	GeV	Cand	Bkgd	QCD	Δ_{QCD}	Top	Δ_{Top}	Tau	Δ_{Tau}	WW	Δ_{WW}	Zee	Δ_{Zee}	Pro	Δ_{Pro}
1	15-20	7466	6.6e+02	206	26	0.124	0.035	55	11	8.11	0.28	8.7	3.4	3.8e+02	1.1e+02
2	20-25	4419	322	165	24	0.378	0.076	45.4	9.7	9.67	0.31	5.4	2.1	96.9	29
3	25-30	2821	243	159	22	0.61	0.11	37.1	8.5	11.46	0.33	3.6	1.4	31.3	9.4
4	30-35	1927	168	103	18	1.24	0.2	35.7	8.3	13.13	0.36	2.39	0.96	12.7	3.8
5	35-40	1353	163	104	17	1.93	0.3	35.7	8.3	14.01	0.37	1.81	0.74	6.26	1.9
6	40-50	850	91.3	43	8.1	3.37	0.49	27.5	5.5	13.59	0.26	1.59	0.63	2.18	0.66
7	50-60	504	70.1	32.8	7.2	5.98	0.86	20.6	4.5	9.19	0.21	0.8	0.33	0.69	0.21
8	60-75	263	43.9	14.9	4.6	7.8	1.1	14.7	3.2	5.65	0.13	0.73	0.29	0.162	0.051
9	75-90	150.3	17.7	4.4	2.9	7.2	1	2.7	1.2	2.911	0.097	0.41	0.17	0.0641	0.022
10	90-110	72.2	26.2	15.4	3.4	5.11	0.73	3.8	1.2	1.504	0.06	0.36	0.15	0.019	0.0075
11	110-150	27.2	9.2	4.7	1.4	2.48	0.35	1.2	0.48	0.571	0.026	0.141	0.059	0.0063	0.0028
12	150-195	6.44	3.7	2.39	0.95	0.678	0.098	0.46	0.27	0.169	0.013	0.038	0.018	0.00056	0.00059
13	195-350	0.87	0.57	0.44	0.2	0.068	0.01	0.044	0.045	0.019	0.0024	0.0042	0.0025	0.00016	0.00017

Table A.5: Results of differential background calculation for first jet E_T in the ≥ 1 jet bin. Shown are the total number of candidates (*Cand*), the total background (*Bkgd*), and then the exact contribution from each separate background source with accompanying error (statistical plus systematic quadrature sum).

Bin	GeV	Cand	Bkgd	QCD	Δ_{QCD}	Top	Δ_{Top}	Tau	Δ_{Tau}	WW	Δ_{WW}	Zee	Δ_{Zee}	Pro	Δ_{Pro}
1	15-20	1952	314	153	24	2.38	0.36	26	33	18.12	0.42	2.9	2.1	112	34
2	20-25	907	118	49	13	3.56	0.53	16	20	19.19	0.43	2.1	1.5	28.6	8.6
3	25-30	503	77	33	11	4.9	0.72	11	14	17.59	0.41	1.4	1	9.2	2.8
4	30-35	318	68	33	11	7.4	1.1	1e+01	12	13.25	0.36	1	0.72	3.8	1.1
5	35-40	184	30.5	6.8	5	8.7	1.3	4.3	5.6	8.09	0.28	0.72	0.52	1.84	0.56
6	40-50	109	30.8	12	4.6	10.8	1.5	2.8	3.6	4.01	0.14	0.52	0.37	0.644	0.19
7	50-60	63	17.7	3.9	3	10.6	1.5	1	1.3	1.764	0.092	0.33	0.24	0.204	0.062
8	60-75	34	14.2	5.7	2.5	7.2	1	0.44	0.6	0.672	0.046	0.19	0.14	0.0477	0.015
9	75-95	13.8	4.6	0.9	1	3.26	0.47	0.08	0.13	0.26	0.025	0.099	0.072	0.016	0.0054
10	95-190	1.68	1.24	0.67	0.33	0.487	0.07	0.018	0.028	0.034	0.0042	0.03	0.022	0.00165	0.00061

Table A.6: Results of differential background calculation for second jet E_T in the ≥ 2 jet bin.

Bin	GeV	Cand	Bkgd	QCD	Δ_{QCD}	Top	Δ_{Top}	Tau	Δ_{Tau}	WW	Δ_{WW}	Zee	Δ_{Zee}	Pro	Δ_{Pro}
1	15-20	426	85	35	21	8.2	1.2	6	37	11.08	0.33	1.1	3.1	24	7.2
2	20-25	196	37.2	10	12	10.3	1.5	3	18	6.43	0.25	0.9	2.5	6.09	1.8
3	25-30	97	26.7	6.1	5.3	12.6	1.8	2	11	3.61	0.19	0.6	1.7	1.96	0.59
4	30-35	49	18.7	1.2	2.8	12.8	1.8	1.4	8.7	2.05	0.14	0.4	1.2	0.8	0.24
5	35-45	36	21.1	7.9	3.5	11.3	1.6	0.6	3.3	0.81	0.062	0.26	0.72	0.287	0.086
6	45-75	6.5	5.33	0.83	0.72	4.16	0.59	0.1	0.6	0.13	0.014	0.08	0.23	0.0351	0.011

Table A.7: Results of differential background calculation for third jet E_T in the ≥ 3 jet bin.

Bin	GeV	Cand	Bkgd	QCD	Δ_{QCD}	Top	Δ_{Top}	Tau	Δ_{Tau}	WW	Δ_{WW}	Zee	Δ_{Zee}	Pro	Δ_{Pro}
1	15-20	103	45.1	22.5	8.8	13.6	1.9	1	6.6	3.35	0.18	0.28	0.51	4.46	1.3
2	20-25	58	21.3	4.1	4.4	14.1	2	0.5	3.4	1.22	0.11	0.16	0.29	1.14	0.34
3	25-30	21	12.9	0	8.1	11.5	1.6	0.3	2.1	0.665	0.08	0.09	0.17	0.367	0.11

Table A.8: Results of differential background calculation for fourth jet E_T in the ≥ 4 jet bin.

Bin	Rad	Cand	Bkgd	QCD	Δ_{QCD}	Top	Δ_{Top}	Tau	Δ_{Tau}	WW	Δ_{WW}	Zee	Δ_{zee}
1	0.32-0.52	89	6.9	2.7	2.8	0.479	0.091	3	3.9	0.636	0.078	0.13	0.1
2	0.52-0.72	323	59	45	12	2.6	0.39	9	11	2.53	0.16	0.39	0.28
3	0.72-0.92	239	42	26.4	9	4.62	0.68	8	11	2.8	0.16	0.27	0.2
4	0.92-1.12	188	21.6	6.9	6.7	5.69	0.83	5.7	7.3	3.21	0.18	0.19	0.14
5	1.12-1.32	186	23.3	8.8	6.1	6.43	0.93	3.7	4.7	4.13	0.2	0.24	0.17
6	1.32-1.52	186	39	23.7	8.5	7.1	1	3.3	4.3	4.66	0.21	0.22	0.16
7	1.52-1.72	190	41.3	26.1	9	6.46	0.94	4	5.2	4.44	0.21	0.34	0.25
8	1.72-1.92	222	39.1	20.4	8.1	7.2	1	5	6.4	6.02	0.24	0.39	0.28
9	1.92-2.12	269	27.9	8.3	5.9	8.6	1.2	3.3	4.3	7.07	0.26	0.57	0.41
10	2.12-2.32	243	51	30	10	8.3	1.2	4	5.2	7.27	0.26	0.75	0.54
11	2.32-2.52	317	50.3	25.6	9	10.3	1.5	4.7	6	8.79	0.29	0.94	0.68
12	2.52-2.72	318	50	26.5	9	10.4	1.5	3	3.9	8.59	0.29	1.31	0.94
13	2.72-2.92	376	52.9	24.9	9.4	10.8	1.6	6	7.7	9.86	0.31	1.38	0.99
14	2.92-3.12	452	48	16.6	7.6	12.6	1.8	5	6.4	11.94	0.34	1.9	1.3
15	3.12-3.32	308	36.2	20.4	8.1	7.1	1	2	2.6	5.63	0.23	1.16	0.83
16	3.32-3.52	215	16.1	8.3	4.8	2.77	0.42	2.3	3.1	2.11	0.14	0.56	0.4
17	3.52-3.72	118	6.3	2.3	2.9	1.26	0.2	1.3	1.8	1.022	0.099	0.36	0.26
18	3.72-3.92	65	4.1	2.4	2.9	0.471	0.09	0.67	0.96	0.395	0.062	0.13	0.1
19	3.92-4.12	51	7.6	5.6	4	0.324	0.068	1.3	1.8	0.241	0.048	0.17	0.13
20	4.12-4.32	18	8.6	8.4	4.8	0.093	0.03	0	0	0.116	0.033	0.038	0.033
21	4.32-4.52	15	0.7	0	8.4	0.031	0.016	0.67	0.96	0.0096	0.0096	0.009	0.012

Table A.9: Results of differential background calculation for ΔR_{ij} in the ≥ 2 jet bin.

Bin	GeV	Cand	Bkgd	QCD	Δ_{QCD}	Top	Δ_{Top}	Tau	Δ_{Tau}	WW	Δ_{WW}	Zee	Δ_{Zee}
1	24-29	206	27	19.2	7.4	0.56	0.1	5.7	7.3	1.32	0.11	0.22	0.16
2	29-34	219	26	17.3	7.9	0.91	0.15	5.7	7.3	1.58	0.12	0.36	0.26
3	34-39	263	24	16.3	8.4	1.1	0.18	4.3	5.6	1.92	0.14	0.42	0.3
4	39-44	279	27.4	18.3	7.6	1.24	0.2	5	6.4	2.29	0.15	0.61	0.44
5	44-49	308	20.9	12.4	6.4	1.41	0.22	3.3	4.3	3.28	0.18	0.53	0.38
6	49-54	292	19.8	10.4	5.7	1.77	0.28	3	3.9	4.06	0.2	0.61	0.44
7	54-64	248	27.5	15.7	4.9	2.17	0.32	2.5	3.2	6.52	0.18	0.61	0.44
8	64-74	186	36.8	20.3	5.8	3.07	0.45	2.5	3.2	10.52	0.23	0.42	0.31
9	74-84	160	28.6	12.2	4.2	4	0.58	1.8	2.4	10.16	0.22	0.4	0.29
10	84-94	120	20.7	9.2	3.7	4	0.58	2.7	3.4	4.37	0.15	0.41	0.29
11	94-109	79.3	12	5.4	2.6	4.05	0.58	1.1	1.4	1.256	0.064	0.23	0.16
12	109-130	49.8	11.8	5.5	2	4.27	0.61	1.2	1.5	0.588	0.037	0.22	0.16
13	130-150	32.8	8.8	4	2	3.64	0.52	0.58	0.77	0.364	0.03	0.17	0.12
14	150-170	23.5	6.2	3	1.6	2.39	0.34	0.42	0.56	0.263	0.025	0.116	0.084
15	170-200	15	3.7	1.56	0.97	1.53	0.22	0.33	0.44	0.182	0.017	0.088	0.064
16	200-240	6.12	1.94	0.58	0.5	0.94	0.14	0.25	0.33	0.113	0.012	0.06	0.044
17	240-300	3.25	1.25	0.62	0.41	0.468	0.068	0.08	0.12	0.045	0.006	0.032	0.023
18	300-500	0.6	0.12	0	0.21	0.084	0.013	0.017	0.024	0.008	0.0014	0.0094	0.0069

Table A.10: Results of differential background calculation for M_{jj} in the ≥ 2 jet bin.

Bin	GeV	Tot	Tot_{Syst}	Tot_{Stat}	Qcd_{Syst}	Qcd_{stat}	Tau_{Syst}	Tau_{Stat}	WW_{Syst}	WW_{Stat}	Top_{Syst}	Top_{Stat}	Zee_{Syst}	Zee_{Stat}
1	15-20	6.6e+02	1.2e+02	27	26	26	11	8.7	0.28	0.28	0.035	0.031	3.4	0.48
2	20-25	322	41	24	24	23	9.7	7.9	0.31	0.31	0.076	0.054	2.1	0.38
3	25-30	243	27	23	22	22	8.5	7.1	0.33	0.33	0.11	0.069	1.4	0.31
4	30-35	168	20	19	18	18	8.3	7	0.36	0.36	0.2	0.098	0.96	0.25
5	35-40	163	19	19	17	17	8.3	7	0.37	0.37	0.3	0.12	0.74	0.22
6	40-50	91.3	9.4	9.2	8.1	8.1	5.5	4.3	0.26	0.26	0.49	0.11	0.63	0.14
7	50-60	70.1	8.6	8.1	7.2	7.2	4.5	3.8	0.21	0.21	0.86	0.15	0.33	0.1
8	60-75	43.9	6.3	5.2	4.6	4.5	3.2	2.6	0.13	0.13	1.1	0.14	0.29	0.08
9	75-90	17.7	4.8	3	2.9	2.8	1.2	1.1	0.097	0.097	1	0.14	0.17	0.06
10	90-110	26.2	4.1	3.6	3.4	3.4	1.2	1.1	0.06	0.06	0.73	0.099	0.15	0.049
11	110-150	9.2	1.6	1.4	1.4	1.4	0.48	0.45	0.026	0.026	0.35	0.049	0.059	0.021
12	150-195	3.7	1.7	0.99	0.95	0.95	0.27	0.26	0.013	0.013	0.098	0.024	0.018	0.011
13	195-350	0.57	0.29	0.2	0.2	0.2	0.045	0.044	0.0024	0.0024	0.01	0.0041	0.0025	0.0019

Table A.11: The statistical and systematic error on each background component in each bin of the first jet E_T distribution, along with the total background and errors. The systematic is the statistical error added in quadrature with the systematics resulting from the antielectron definition, Monte Carlo model dependence and Top cross-section uncertainties. In addition the total systematic includes the jet promotion contribution.

Bin	GeV	Tot	Tot_{Syst}	Tot_{Stat}	Qcd_{Syst}	Qcd_{stat}	Tau_{Syst}	Tau_{Stat}	WW_{Syst}	WW_{Stat}	Top_{Syst}	Top_{Stat}	Zee_{Syst}	Zee_{Stat}
1	15-20	314	43	22	24	21	33	3	0.42	0.42	0.36	0.14	2.1	0.17
2	20-25	118	17	12	13	12	20	2.3	0.43	0.43	0.53	0.17	1.5	0.14
3	25-30	77	11	10	11	10	14	1.9	0.41	0.41	0.72	0.19	1	0.12
4	30-35	68	14	10	11	10	12	1.8	0.36	0.36	1.1	0.24	0.72	0.097
5	35-40	30.5	5.2	5	5	4.9	5.6	1.2	0.28	0.28	1.3	0.26	0.52	0.082
6	40-50	30.8	6.3	4.5	4.6	4.5	3.6	0.69	0.14	0.14	1.5	0.2	0.37	0.049
7	50-60	17.7	4	2.9	3	2.8	1.3	0.41	0.092	0.092	1.5	0.2	0.24	0.039
8	60-75	14.2	2.7	2.5	2.5	2.5	0.6	0.22	0.046	0.046	1	0.14	0.14	0.025
9	75-95	4.6	1.2	1	1	0.99	0.13	0.083	0.025	0.025	0.47	0.079	0.072	0.015
10	95-190	1.24	0.35	0.33	0.33	0.33	0.028	0.018	0.0042	0.0042	0.07	0.014	0.022	0.0039

Table A.12: The statistical and systematic error on each background component in each bin of the second jet E_T distribution, along with the total background and errors.

Bin	GeV	Tot	Tot_{Syst}	Tot_{Stat}	Qcd_{Syst}	Qcd_{stat}	Tau_{Syst}	Tau_{Stat}	WW_{Syst}	WW_{Stat}	Top_{Syst}	Top_{Stat}	Zee_{Syst}	Zee_{Stat}
1	15-20	85	13	9.7	21	9.7	37	0.73	0.33	0.33	1.2	0.25	3.1	0.062
2	20-25	37.2	6.6	5.6	12	5.6	18	0.51	0.25	0.25	1.5	0.28	2.5	0.056
3	25-30	26.7	4.8	4.4	5.3	4.3	11	0.4	0.19	0.19	1.8	0.31	1.7	0.046
4	30-35	18.7	3.2	2.6	2.8	2.5	8.7	0.35	0.14	0.14	1.8	0.31	1.2	0.039
5	35-45	21.1	3.7	3.3	3.5	3.3	3.3	0.15	0.062	0.062	1.6	0.21	0.72	0.021
6	45-75	5.33	0.97	0.73	0.72	0.72	0.6	0.038	0.014	0.014	0.59	0.073	0.23	0.0068

Table A.13: The statistical and systematic error on each background component in each bin of the third jet E_T distribution, along with the total background and errors.

Bin	GeV	Tot	Tot_{Syst}	Tot_{Stat}	Qcd_{Syst}	Qcd_{stat}	Tau_{Syst}	Tau_{Stat}	WW_{Syst}	WW_{Stat}	Top_{Syst}	Top_{Stat}	Zee_{Syst}	Zee_{Stat}
1	15-20	45.1	9.6	8.1	8.8	8.1	6.6	0.12	0.18	0.18	1.9	0.32	0.51	0.019
2	20-25	21.3	4.4	3.8	4.4	3.8	3.4	0.089	0.11	0.11	2	0.33	0.29	0.014
3	25-30	12.9	8.3	8.1	8.1	8.1	2.1	0.069	0.08	0.08	1.6	0.3	0.17	0.011

Table A.14: The statistical and systematic error on each background component in each bin of the fourth jet E_T distribution, along with the total background and errors.

Bin	Rad	Tot	Tot_{Syst}	Tot_{Stat}	Qcd_{Syst}	Qcd_{stat}	Tau_{Syst}	Tau_{Stat}	WW_{Syst}	WW_{Stat}	Top_{Syst}	Top_{Stat}	Zee_{Syst}	Zee_{Stat}
1	0.32-0.52	6.9	3.2	3	2.8	2.8	3.9	1	0.078	0.078	0.091	0.061	0.1	0.035
2	0.52-0.72	59	15	12	12	12	11	1.7	0.16	0.16	0.39	0.14	0.28	0.06
3	0.72-0.92	42	15	9	9	8.9	11	1.7	0.16	0.16	0.68	0.19	0.2	0.051
4	0.92-1.12	21.6	6.5	5.9	6.7	5.7	7.3	1.4	0.18	0.18	0.83	0.21	0.14	0.042
5	1.12-1.32	23.3	6.7	5.8	6.1	5.6	4.7	1.1	0.2	0.2	0.93	0.22	0.17	0.047
6	1.32-1.52	39	9.9	8.5	8.5	8.4	4.3	1.1	0.21	0.21	1	0.23	0.16	0.045
7	1.52-1.72	41.3	9	9	9	8.9	5.2	1.2	0.21	0.21	0.94	0.22	0.25	0.057
8	1.72-1.92	39.1	8.3	8.1	8.1	7.9	6.4	1.3	0.24	0.24	1	0.24	0.28	0.06
9	1.92-2.12	27.9	5.9	5.8	5.9	5.7	4.3	1.1	0.26	0.26	1.2	0.26	0.41	0.073
10	2.12-2.32	51	12	9.8	10	9.7	5.2	1.2	0.26	0.26	1.2	0.25	0.54	0.084
11	2.32-2.52	50.3	9.1	9	9	8.9	6	1.2	0.29	0.29	1.5	0.28	0.68	0.094
12	2.52-2.72	50	10	8.9	9	8.9	3.9	1	0.29	0.29	1.5	0.28	0.94	0.11
13	2.72-2.92	52.9	9.9	9	9.4	8.9	7.7	1.4	0.31	0.31	1.6	0.29	0.99	0.11
14	2.92-3.12	48	7.8	7.6	7.6	7.5	6.4	1.3	0.34	0.34	1.8	0.31	1.3	0.13
15	3.12-3.32	36.2	8.4	8	8.1	8	2.6	0.82	0.23	0.23	1	0.23	0.83	0.1
16	3.32-3.52	16.1	5	4.9	4.8	4.8	3.1	0.88	0.14	0.14	0.42	0.15	0.4	0.072
17	3.52-3.72	6.3	3.2	2.9	2.9	2.8	1.8	0.67	0.099	0.099	0.2	0.099	0.26	0.058
18	3.72-3.92	4.1	3.1	2.9	2.9	2.8	0.96	0.47	0.062	0.062	0.09	0.06	0.1	0.035
19	3.92-4.12	7.6	4.3	4	4	4	1.8	0.67	0.048	0.048	0.068	0.05	0.13	0.04
20	4.12-4.32	8.6	4.9	4.8	4.8	4.8	0	0	0.033	0.033	0.03	0.027	0.033	0.019
21	4.32-4.52	0.7	8.4	8.4	8.4	8.4	0.96	0.47	0.0096	0.0096	0.016	0.015	0.012	0.0094

Table A.15: The statistical and systematic error on each background component in each bin of the ΔR_{jj} distribution, along with the total background and errors.

Bin	GeV	Tot	Tot_{Syst}	Tot_{Stat}	Qcd_{Syst}	Qcd_{stat}	Tau_{Syst}	Tau_{Stat}	WW_{Syst}	WW_{Stat}	Top_{Syst}	Top_{Stat}	Zee_{Syst}	Zee_{Stat}
1	24-29	27	9	7.5	7.4	7.4	7.3	1.4	0.11	0.11	0.1	0.066	0.16	0.045
2	29-34	26	10	7.6	7.9	7.5	7.3	1.4	0.12	0.12	0.15	0.084	0.26	0.058
3	34-39	24	10	7.6	8.4	7.5	5.6	1.2	0.14	0.14	0.18	0.092	0.3	0.063
4	39-44	27.4	8.1	7.6	7.6	7.4	6.4	1.3	0.15	0.15	0.2	0.098	0.44	0.076
5	44-49	20.9	6.8	6.4	6.4	6.3	4.3	1.1	0.18	0.18	0.22	0.1	0.38	0.071
6	49-54	19.8	5.8	5.7	5.7	5.6	3.9	1	0.2	0.2	0.28	0.12	0.44	0.076
7	54-64	27.5	6.8	4.9	4.9	4.9	3.2	0.65	0.18	0.18	0.32	0.092	0.44	0.054
8	64-74	36.8	5.9	5.7	5.8	5.6	3.2	0.65	0.23	0.23	0.45	0.11	0.31	0.045
9	74-84	28.6	4.6	4.2	4.2	4.2	2.4	0.55	0.22	0.22	0.58	0.12	0.29	0.043
10	84-94	20.7	4.4	3.8	3.7	3.7	3.4	0.67	0.15	0.15	0.58	0.12	0.29	0.044
11	94-109	12	3	2.5	2.6	2.5	1.4	0.35	0.064	0.064	0.58	0.1	0.16	0.027
12	109-130	11.8	2.2	2	2	2	1.5	0.31	0.037	0.037	0.61	0.089	0.16	0.022
13	130-150	8.8	2.3	1.9	2	1.9	0.77	0.22	0.03	0.03	0.52	0.084	0.12	0.02
14	150-170	6.2	2.2	1.6	1.6	1.6	0.56	0.19	0.025	0.025	0.34	0.068	0.084	0.017
15	170-200	3.7	1.3	0.95	0.97	0.94	0.44	0.14	0.017	0.017	0.22	0.044	0.064	0.012
16	200-240	1.94	0.6	0.51	0.5	0.5	0.33	0.1	0.012	0.012	0.14	0.03	0.044	0.0084
17	240-300	1.25	0.51	0.41	0.41	0.4	0.12	0.048	0.006	0.006	0.068	0.017	0.023	0.005
18	300-500	0.12	0.21	0.21	0.21	0.21	0.024	0.012	0.0014	0.0014	0.013	0.004	0.0069	0.0015

Table A.16: The statistical and systematic error on each background component in each bin of the M_{jj} distribution, along with the total background and errors.

Appendix B

Acceptance Results

Bin	GeV	Acc	Δ_{Acc}	Geo	Δ_{Geo}	Kin	Δ_{Kin}	ID	Δ_{ID}
1	15-20	0.5818	0.0044	0.7731	0.0033	0.9475	0.0048	0.8058	0.0052
2	20-25	0.5934	0.0058	0.7726	0.0042	0.9617	0.0063	0.8125	0.0068
3	25-30	0.5913	0.0072	0.7751	0.0052	0.9642	0.0078	0.8065	0.0084
4	30-35	0.6078	0.0045	0.775	0.0033	0.99	0.0049	0.806	0.0052
5	35-40	0.601	0.0054	0.7727	0.0039	0.978	0.0058	0.8072	0.0062
6	40-50	0.6164	0.0047	0.7738	0.0034	0.9985	0.0051	0.8084	0.0054
7	50-60	0.6147	0.0063	0.7774	0.0045	0.9941	0.0067	0.8061	0.0072
8	60-75	0.6104	0.0026	0.7759	0.0019	0.9852	0.0028	0.8086	0.003
9	75-90	0.6047	0.0036	0.7776	0.0027	0.979	0.0037	0.8049	0.004
10	90-110	0.6013	0.0044	0.77	0.0033	0.9844	0.0045	0.803	0.005
11	110-150	0.5897	0.0025	0.7746	0.0019	0.9723	0.0024	0.7913	0.0028
12	150-195	0.595	0.0046	0.7742	0.0035	0.9867	0.0044	0.7869	0.0051
13	195-350	0.6057	0.0031	0.7833	0.0023	0.9788	0.0028	0.7983	0.0034

Table B.1: Breakdown of the differential acceptance results in the first jet E_T distribution. Shown is the total acceptance, and the component acceptance factors for the geometric, kinematic and ID cuts. Errors shown are the result of the limited Monte Carlo statistics, and do not include the flat 5% acceptance systematic.

Bin	GeV	Acc	Δ_{Acc}	Geo	Δ_{Geo}	Kin	Δ_{Kin}	ID	Δ_{ID}
1	15-20	0.5897	0.0053	0.7743	0.0038	0.9795	0.0058	0.798	0.0062
2	20-25	0.5898	0.0069	0.7769	0.005	0.9789	0.0076	0.7966	0.0081
3	25-30	0.5896	0.0088	0.7804	0.0063	0.9685	0.0095	0.801	0.01
4	30-35	0.583	0.011	0.7717	0.0082	0.977	0.012	0.797	0.013
5	35-40	0.577	0.014	0.763	0.01	0.984	0.016	0.794	0.017
6	40-50	0.591	0.014	0.7715	0.0098	0.992	0.015	0.791	0.016
7	50-60	0.569	0.02	0.762	0.014	1.006	0.022	0.774	0.023
8	60-75	0.573	0.024	0.794	0.017	0.966	0.027	0.769	0.029
9	75-95	0.656	0.037	0.794	0.024	1.045	0.04	0.799	0.041
10	95-190	0.626	0.045	0.846	0.028	0.935	0.05	0.804	0.052

Table B.2: Breakdown of the differential acceptance results in the second jet E_T distribution.

Bin	GeV	Acc	Δ_{Acc}	Geo	Δ_{Geo}	Kin	Δ_{Kin}	ID	Δ_{ID}
1	15-20	0.5603	0.0054	0.7764	0.0039	0.9729	0.006	0.7667	0.0064
2	20-25	0.5632	0.0075	0.7761	0.0054	0.9878	0.0082	0.7589	0.0089
3	25-30	0.5429	0.01	0.7573	0.0074	0.975	0.011	0.77	0.012
4	30-35	0.568	0.014	0.7697	0.01	0.979	0.016	0.781	0.017
5	35-45	0.55	0.015	0.759	0.011	0.974	0.016	0.768	0.018
6	45-75	0.544	0.019	0.776	0.014	0.993	0.022	0.742	0.023

Table B.3: Breakdown of the differential acceptance results in the third jet E_T distribution.

Bin	GeV	Acc	Δ_{Acc}	Geo	Δ_{Geo}	Kin	Δ_{Kin}	ID	Δ_{ID}
1	15-20	0.546	0.01	0.7633	0.0074	1.005	0.012	0.745	0.012
2	20-25	0.565	0.015	0.784	0.01	0.99	0.017	0.766	0.018
3	25-30	0.539	0.02	0.78	0.015	0.961	0.023	0.747	0.025

Table B.4: Breakdown of the differential acceptance results in the fourth jet E_T distribution.

Bin	Rad	Acc	Δ_{Acc}	Geo	Δ_{Geo}	Kin	Δ_{Kin}	ID	Δ_{ID}
1	0.32-0.52	0.651	0.027	0.797	0.019	1.026	0.027	0.807	0.03
2	0.52-0.72	0.599	0.012	0.7634	0.0089	0.985	0.013	0.808	0.014
3	0.72-0.92	0.614	0.013	0.7814	0.0095	0.975	0.014	0.821	0.015
4	0.92-1.12	0.606	0.014	0.776	0.011	0.991	0.015	0.806	0.016
5	1.12-1.32	0.586	0.015	0.768	0.011	0.994	0.016	0.784	0.017
6	1.32-1.52	0.593	0.016	0.753	0.011	1.011	0.017	0.789	0.018
7	1.52-1.72	0.616	0.016	0.786	0.011	0.977	0.017	0.828	0.018
8	1.72-1.92	0.601	0.015	0.777	0.011	0.998	0.017	0.798	0.018
9	1.92-2.12	0.596	0.014	0.7913	0.0096	0.983	0.015	0.794	0.016
10	2.12-2.32	0.592	0.014	0.7978	0.0095	0.982	0.015	0.786	0.016
11	2.32-2.52	0.588	0.012	0.7873	0.0087	0.97	0.014	0.795	0.014
12	2.52-2.72	0.578	0.012	0.773	0.0085	1.004	0.014	0.766	0.014
13	2.72-2.92	0.557	0.012	0.7599	0.0085	0.959	0.013	0.788	0.014
14	2.92-3.12	0.5665	0.0099	0.7749	0.0072	0.97	0.011	0.781	0.012
15	3.12-3.32	0.576	0.012	0.764	0.0092	0.954	0.014	0.81	0.015
16	3.32-3.52	0.585	0.016	0.78	0.012	0.971	0.018	0.794	0.019
17	3.52-3.72	0.577	0.02	0.769	0.015	0.956	0.022	0.807	0.024
18	3.72-3.92	0.591	0.028	0.766	0.021	0.951	0.031	0.814	0.032
19	3.92-4.12	0.603	0.035	0.79	0.024	0.974	0.038	0.805	0.042
20	4.12-4.32	0.623	0.05	0.768	0.036	1	0.054	0.821	0.056
21	4.32-4.52	0.583	0.066	0.762	0.046	0.953	0.073	0.82	0.08

Table B.5: Breakdown of the differential acceptance results in the ΔR_{jj} distribution.

Bin	GeV	Acc	Δ_{Acc}	Geo	Δ_{Geo}	Kin	Δ_{Kin}	ID	Δ_{ID}
1	24-29	0.622	0.015	0.779	0.01	0.998	0.015	0.818	0.017
2	29-34	0.607	0.014	0.767	0.01	0.998	0.015	0.806	0.016
3	34-39	0.598	0.013	0.7857	0.0092	0.994	0.014	0.792	0.015
4	39-44	0.589	0.013	0.7699	0.0093	0.991	0.014	0.79	0.015
5	44-49	0.581	0.013	0.7651	0.0092	0.983	0.014	0.794	0.015
6	49-54	0.574	0.013	0.771	0.0094	0.992	0.014	0.781	0.015
7	54-64	0.576	0.0099	0.7746	0.0071	0.963	0.011	0.799	0.012
8	64-74	0.569	0.011	0.7839	0.0078	0.953	0.012	0.784	0.013
9	74-84	0.593	0.012	0.7738	0.0089	0.978	0.013	0.805	0.014
10	84-94	0.552	0.014	0.754	0.011	0.954	0.016	0.79	0.017
11	94-109	0.59	0.013	0.783	0.0096	0.981	0.014	0.786	0.015
12	109-130	0.584	0.014	0.783	0.01	0.976	0.016	0.781	0.017
13	130-150	0.588	0.017	0.788	0.012	0.987	0.018	0.781	0.02
14	150-170	0.584	0.023	0.782	0.016	0.975	0.025	0.791	0.026
15	170-200	0.585	0.026	0.776	0.018	0.963	0.029	0.818	0.031
16	200-240	0.606	0.03	0.772	0.021	1.013	0.033	0.792	0.035
17	240-300	0.638	0.042	0.796	0.028	1.006	0.048	0.808	0.048
18	300-500	0.641	0.05	0.791	0.033	0.975	0.056	0.846	0.059

Table B.6: Breakdown of the differential acceptance results in the M_{jj} distribution.

n Jets	N_{TT}		N_Z		ξ^{raw}	ξ	Z MC ξ	Scale Factor
	OS	SS - τ	OS	SS - τ				
≥ 0	5754	13 - 6.3	8427	209 - 110.2	0.8115	0.8192 ±0.0035	0.8217 ±0.001	0.9969 ±0.004
≥ 1	1049	6 - 2.0	1611	65 - 23.8	0.7887	0.8047 ±0.0085	0.8081 ±0.003	0.9958 ±0.011
≥ 2	187	1 - 0.8	323	13 - 5.6	0.7333	0.7488 ±0.0217	0.8117 ±0.007	0.9224 ±0.027
≥ 3	45	0 - 0.0	64	2 - 1.0	0.8257	0.8368 ±0.0384	0.7938 ±0.020	1.0541 ±0.051
≥ 4	5	0 - 0.0	8	0 - 0.0	0.7692	0.7692 ±0.1296	0.7619 ±0.060	1.0096 ±0.176

Table B.7: Inclusive electron ID efficiency numbers using a Z sample defined by a mass window [76,106] GeV. OS is the number of opposite-sign events, SS the number of same-sign events and τ the estimated number of trident events. ξ^{raw} is the data ID efficiency before SS and trident correction, and ξ is the fully corrected number. Z MC ξ is the ID efficiency from PYTHIA Z$\rightarrow$ ee Monte Carlo, and the scale factor is the ratio of the data to MC efficiencies.

Appendix C

Cross-Section Results

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	15-20	36.97	0.947	0.0056	7466	86	6.6e+02	1.2e+02	0.5818	0.0044
2	20-25	22.28	0.968	0.00723	4419	66	322	41	0.5934	0.0058
3	25-30	13.99	0.962	0.00879	2821	53	243	27	0.5913	0.0072
4	30-35	9.6	0.995	0.0109	1927	44	168	20	0.6078	0.0045
5	35-40	6.55	0.991	0.0128	1353	37	163	19	0.601	0.0054
6	40-50	4.19	1.02	0.006	850	21	91.3	9.4	0.6164	0.0047
7	50-60	2.49	1.06	0.00829	504	16	70.1	8.6	0.6147	0.0063
8	60-75	1.293	1.08	0.00957	263	9.4	43.9	6.3	0.6104	0.0026
9	75-90	0.791	1.08	0.0119	150.3	7.1	17.7	4.8	0.6047	0.0036
10	90-110	0.281	1.1	0.00678	72.2	4.2	26.2	4.1	0.6013	0.0044
11	110-150	0.113	1.1	0.00825	27.2	1.8	9.2	1.6	0.5897	0.0025
12	150-195	0.0175	1.15	0.00799	6.44	0.85	3.7	1.7	0.595	0.0046
13	195-350	0.002	1.2	0.00659	0.87	0.17	0.57	0.29	0.6057	0.0031

Table C.1: Table showing the unsmearing, candidate, total background and acceptance numbers in each bin of the first jet E_T distribution and the cross-section that results. Also shown is the Poisson statistical error on the candidates, and the systematic errors on the total background and acceptance numbers (the acceptance systematic is from the statistical error on the acceptance only and does not include the flat 5% systematic on the acceptance.) Note that the quoted cross-section for a particular bin i is the inclusive cross-section in that bin, and to get the corresponding differential cross-section one simply divides by the quoted bin width

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	15-20	8.82	0.952	0.00548	1952	44	314	43	0.5897	0.0053
2	20-25	4.37	0.98	0.0077	907	30	118	17	0.5898	0.0069
3	25-30	2.41	1	0.0105	503	22	77	11	0.5896	0.0088
4	30-35	1.43	1	0.0136	318	18	68	14	0.583	0.011
5	35-40	0.952	1.07	0.0186	184	14	30.5	5.2	0.577	0.014
6	40-50	0.475	1.08	0.0179	109	7.4	30.8	6.3	0.591	0.014
7	50-60	0.287	1.08	0.0267	63	5.6	17.7	4	0.569	0.02
8	60-75	0.134	1.17	0.0352	34	3.4	14.2	2.7	0.573	0.024
9	75-95	0.054	1.15	0.0494	13.8	1.9	4.6	1.2	0.656	0.037
10	95-190	0.0029	1.24	0.0664	1.68	0.3	1.24	0.35	0.626	0.045

Table C.2: Table showing the unsmearing, candidate, total background and acceptance numbers in each bin of the second jet E_T distribution and the cross-section that results. Statistical errors only.

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	15-20	1.95	0.959	0.0107	426	21	85	13	0.5603	0.0054
2	20-25	0.919	0.977	0.0157	196	14	37.2	6.6	0.5632	0.0075
3	25-30	0.434	1.01	0.0226	97	9.8	26.7	4.8	0.5429	0.01
4	30-35	0.183	1.03	0.0318	49	7	18.7	3.2	0.568	0.014
5	35-45	0.101	1.12	0.0369	36	4.2	21.1	3.7	0.55	0.015
6	45-75	0.0082	1.16	0.0512	6.5	1	5.33	0.97	0.544	0.019

Table C.3: Table showing the unsmearing, candidate, total background and acceptance numbers in each bin of the third jet E_T distribution and the cross-section that results. Statistical errors only.

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	15-20	0.341	0.964	0.0358	103	10	45.1	9.6	0.546	0.01
2	20-25	0.211	0.971	0.0552	58	7.6	21.3	4.4	0.565	0.015
3	25-30	0.052	1.05	0.0869	21	4.6	12.9	8.3	0.539	0.02

Table C.4: Table showing the unsmearing, candidate, total background and acceptance numbers in each bin of the fourth jet E_T distribution and the cross-section that results. Statistical errors only.

Bin	Rad	$\sigma[pb]$	$Cand$	Δ_{Cand}	$Bkgd$	Δ_{Bkgd}	Acc	Δ_{Acc}
1	0.32-0.52	0.42	89	9.4	6.9	3.2	0.651	0.027
2	0.52-0.72	1.47	323	18	59	15	0.599	0.012
3	0.72-0.92	1.068	239	15	42	15	0.614	0.013
4	0.92-1.12	0.917	188	14	21.6	6.5	0.606	0.014
5	1.12-1.32	0.927	186	14	23.3	6.7	0.586	0.015
6	1.32-1.52	0.827	186	14	39	9.9	0.593	0.016
7	1.52-1.72	0.805	190	14	41.3	9	0.616	0.016
8	1.72-1.92	1.016	222	15	39.1	8.3	0.601	0.015
9	1.92-2.12	1.349	269	16	27.9	5.9	0.596	0.014
10	2.12-2.32	1.084	243	16	51	12	0.592	0.014
11	2.32-2.52	1.51	317	18	50.3	9.1	0.588	0.012
12	2.52-2.72	1.55	318	18	50	10	0.578	0.012
13	2.72-2.92	1.94	376	19	52.9	9.9	0.557	0.012
14	2.92-3.12	2.38	452	21	48	7.8	0.5665	0.0099
15	3.12-3.32	1.57	308	18	36.2	8.4	0.576	0.012
16	3.32-3.52	1.135	215	15	16.1	5	0.585	0.016
17	3.52-3.72	0.646	118	11	6.3	3.2	0.577	0.02
18	3.72-3.92	0.344	65	8.1	4.1	3.1	0.591	0.028
19	3.92-4.12	0.24	51	7.1	7.6	4.3	0.603	0.035
20	4.12-4.32	0.05	18	4.2	8.6	4.9	0.623	0.05
21	4.32-4.52	0.082	15	3.9	0.7	8.4	0.583	0.066

Table C.5: Table showing the unsmearing, candidate, total background and acceptance numbers in each bin of the ΔR_{jj} distribution and the cross-section that results.

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	24-29	0.95	0.989	0.0181	206	14	27	9	0.622	0.015
2	29-34	1.031	0.971	0.0186	219	15	26	10	0.607	0.014
3	34-39	1.289	0.966	0.0178	263	16	24	10	0.598	0.013
4	39-44	1.4	0.983	0.0165	279	17	27.4	8.1	0.589	0.013
5	44-49	1.608	0.975	0.0173	308	18	20.9	6.8	0.581	0.013
6	49-54	1.557	0.985	0.0166	292	17	19.8	5.8	0.574	0.013
7	54-64	1.291	1.01	0.0125	248	11	27.5	6.8	0.576	0.0099
8	64-74	0.878	1	0.0138	186	9.6	36.8	5.9	0.569	0.011
9	74-84	0.771	1.04	0.0161	160	8.9	28.6	4.6	0.593	0.012
10	84-94	0.613	1.02	0.0178	120	7.7	20.7	4.4	0.552	0.014
11	94-109	0.396	1.04	0.0174	79.3	5.1	12	3	0.59	0.013
12	109-130	0.23	1.06	0.0184	49.8	3.4	11.8	2.2	0.584	0.014
13	130-150	0.138	1.01	0.0229	32.8	2.9	8.8	2.3	0.588	0.017
14	150-170	0.106	1.08	0.0303	23.5	2.4	6.2	2.2	0.584	0.023
15	170-200	0.072	1.11	0.033	15	1.6	3.7	1.3	0.585	0.026
16	200-240	0.0245	1.06	0.0387	6.12	0.88	1.94	0.6	0.606	0.03
17	240-300	0.0119	1.14	0.0517	3.25	0.52	1.25	0.51	0.638	0.042
18	300-500	0.00284	1.13	0.0659	0.6	0.12	0.12	0.21	0.641	0.05

Table C.6: Table showing the unsmearing, candidate, total background and acceptance numbers in each bin of the M_{jj} distribution and the cross-section that results.

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	15-20	36.97	0.47	2.1	-1.9	0.29	-0.29	-0.64	0.64	-1.8	2	0.22	-0.22
2	20-25	22.28	0.36	1.5	-1.4	0.8	-0.8	-0.22	0.22	-1.1	1.2	0.17	-0.17
3	25-30	13.99	0.29	1	-0.97	0.66	-0.66	-0.15	0.15	-0.68	0.76	0.13	-0.13
4	30-35	9.6	0.24	0.85	-0.83	0.67	-0.67	-0.11	0.11	-0.46	0.51	0.11	-0.11
5	35-40	6.55	0.2	0.68	-0.66	0.56	-0.56	-0.11	0.11	-0.32	0.35	0.085	-0.085
6	40-50	4.19	0.11	0.47	-0.46	0.4	-0.4	-0.052	0.052	-0.2	0.22	0.025	-0.025
7	50-60	2.49	0.091	0.31	-0.31	0.28	-0.28	-0.049	0.049	-0.12	0.13	0.02	-0.02
8	60-75	1.293	0.055	0.19	-0.19	0.17	-0.17	-0.037	0.037	-0.062	0.068	0.011	-0.011
9	75-90	0.791	0.042	0.12	-0.12	0.11	-0.11	-0.028	0.028	-0.038	0.042	0.0087	-0.0087
10	90-110	0.281	0.026	0.055	-0.054	0.046	-0.046	-0.025	0.025	-0.014	0.015	0.0017	-0.0017
11	110-150	0.113	0.011	0.022	-0.022	0.019	-0.019	-0.01	0.01	-0.0054	0.0059	0.00084	-0.00084
12	150-195	0.0175	0.0055	0.011	-0.011	0.003	-0.003	-0.011	0.011	-0.00084	0.00093	0.00012	-0.00012
13	195-350	0.002	0.0011	0.0019	-0.0019	0.00035	-0.00035	-0.0019	0.0019	-9.4e-05	0.0001	1.1e-05	-1.1e-05

Table C.7: Table showing the breakdown of statistical and systematic errors on the cross-section with respect to first jet E_T. Columns $\Delta_\sigma^{JES+/-}$, $\Delta_\sigma^{Bkgd+/-}$, and $\Delta_\sigma^{Acc+/-}$ show the plus/minus absolute systematic from the uncertainty on the jet energy scale, background estimation and acceptance calculation respectively. The total systematic and statistical errors are given by columns $\Delta_\sigma^{Syst+/-}$ and Δ_σ^{Stat} respectively. Note that here the acceptance systematic includes the flat 5%.

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	15-20	8.82	0.24	0.78	-0.75	0.57	-0.57	-0.23	0.23	-0.43	0.47	0.051	-0.051
2	20-25	4.37	0.17	0.48	-0.47	0.41	-0.41	-0.093	0.093	-0.21	0.24	0.034	-0.034
3	25-30	2.41	0.13	0.33	-0.33	0.3	-0.3	-0.061	0.061	-0.12	0.13	0.025	-0.025
4	30-35	1.43	0.1	0.23	-0.22	0.19	-0.19	-0.078	0.078	-0.073	0.081	0.019	-0.019
5	35-40	0.952	0.084	0.17	-0.17	0.15	-0.15	-0.032	0.032	-0.05	0.056	0.016	-0.016
6	40-50	0.475	0.045	0.094	-0.093	0.081	-0.081	-0.038	0.038	-0.025	0.028	0.0079	-0.0079
7	50-60	0.287	0.036	0.061	-0.06	0.052	-0.052	-0.025	0.025	-0.016	0.019	0.0071	-0.0071
8	60-75	0.134	0.023	0.033	-0.033	0.026	-0.026	-0.018	0.018	-0.0083	0.0095	0.0041	-0.0041
9	75-95	0.054	0.011	0.014	-0.014	0.011	-0.011	-0.007	0.007	-0.0037	0.0043	0.0023	-0.0023
10	95-190	0.0029	0.002	0.0024	-0.0024	0.00061	-0.00061	-0.0023	0.0023	-0.00024	0.00028	0.00016	-0.00016

Table C.8: Table showing the breakdown of statistical and systematic errors on the cross-section with respect to second jet E_T .

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	15-20	1.95	0.12	0.26	-0.25	0.22	-0.22	-0.072	0.072	-0.094	0.1	0.022	-0.022
2	20-25	0.919	0.081	0.18	-0.18	0.17	-0.17	-0.038	0.038	-0.045	0.05	0.015	-0.015
3	25-30	0.434	0.061	0.12	-0.12	0.11	-0.11	-0.03	0.03	-0.022	0.024	0.0098	-0.0098
4	30-35	0.183	0.042	0.064	-0.064	0.059	-0.059	-0.019	0.019	-0.0096	0.011	0.0057	-0.0057
5	35-45	0.101	0.029	0.053	-0.053	0.046	-0.046	-0.025	0.025	-0.0054	0.0061	0.0033	-0.0033
6	45-75	0.0082	0.0072	0.0079	-0.0079	0.0038	-0.0038	-0.0069	0.0069	-0.00047	0.00053	0.00036	-0.00036

Table C.9: Table showing the breakdown of statistical and systematic errors on the cross-section with respect to third jet E_T .

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	15-20	0.341	0.06	0.068	-0.067	0.028	-0.028	-0.057	0.057	-0.017	0.019	0.013	-0.013
2	20-25	0.211	0.044	0.039	-0.039	0.024	-0.024	-0.025	0.025	-0.011	0.013	0.012	-0.012
3	25-30	0.052	0.03	0.055	-0.055	0.0097	-0.0097	-0.054	0.054	-0.0031	0.0035	0.0043	-0.0043

Table C.10: Table showing the breakdown of statistical and systematic errors on the cross-section with respect to fourth jet E_T .

Bin	Rad	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}
1	0.32-0.52	0.42	0.048	0.04	-0.037	0.021	-0.021	-0.016	0.016	-0.026	0.029
2	0.52-0.72	1.47	0.1	0.14	-0.14	0.074	-0.074	-0.086	0.086	-0.075	0.084
3	0.72-0.92	1.068	0.084	0.12	-0.11	0.053	-0.053	-0.081	0.081	-0.055	0.061
4	0.92-1.12	0.917	0.076	0.079	-0.075	0.046	-0.046	-0.036	0.036	-0.048	0.054
5	1.12-1.32	0.927	0.078	0.081	-0.077	0.046	-0.046	-0.038	0.038	-0.049	0.055
6	1.32-1.52	0.827	0.077	0.085	-0.082	0.041	-0.041	-0.056	0.056	-0.044	0.05
7	1.52-1.72	0.805	0.075	0.079	-0.076	0.04	-0.04	-0.049	0.049	-0.043	0.048
8	1.72-1.92	1.016	0.083	0.091	-0.087	0.051	-0.051	-0.046	0.046	-0.054	0.06
9	1.92-2.12	1.349	0.092	0.11	-0.1	0.067	-0.067	-0.033	0.033	-0.07	0.079
10	2.12-2.32	1.084	0.088	0.11	-0.1	0.054	-0.054	-0.07	0.07	-0.057	0.063
11	2.32-2.52	1.51	0.1	0.13	-0.12	0.076	-0.076	-0.052	0.052	-0.078	0.087
12	2.52-2.72	1.55	0.1	0.13	-0.13	0.077	-0.077	-0.059	0.059	-0.08	0.089
13	2.72-2.92	1.94	0.12	0.16	-0.15	0.097	-0.097	-0.059	0.059	-0.099	0.11
14	2.92-3.12	2.38	0.13	0.18	-0.17	0.12	-0.12	-0.046	0.046	-0.12	0.13
15	3.12-3.32	1.57	0.1	0.13	-0.12	0.079	-0.079	-0.049	0.049	-0.081	0.091
16	3.32-3.52	1.135	0.084	0.093	-0.088	0.057	-0.057	-0.028	0.028	-0.061	0.069
17	3.52-3.72	0.646	0.063	0.056	-0.053	0.032	-0.032	-0.018	0.018	-0.037	0.042
18	3.72-3.92	0.344	0.046	0.035	-0.033	0.017	-0.017	-0.018	0.018	-0.022	0.025
19	3.92-4.12	0.24	0.039	0.033	-0.032	0.012	-0.012	-0.024	0.024	-0.017	0.02
20	4.12-4.32	0.05	0.023	0.027	-0.026	0.0025	-0.0025	-0.026	0.026	-0.0043	0.0052
21	4.32-4.52	0.082	0.022	0.05	-0.049	0.0041	-0.0041	-0.048	0.048	-0.009	0.012

Table C.11: Table showing the breakdown of statistical and systematic errors on the cross-section with respect to ΔR_{jj} .

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	24-29	0.95	0.076	0.08	-0.076	0.028	-0.028	-0.047	0.047	-0.05	0.056	0.017	-0.017
2	29-34	1.031	0.079	0.09	-0.086	0.033	-0.033	-0.055	0.055	-0.054	0.06	0.02	-0.02
3	34-39	1.289	0.087	0.11	-0.1	0.044	-0.044	-0.056	0.056	-0.067	0.074	0.024	-0.024
4	39-44	1.4	0.093	0.11	-0.1	0.05	-0.05	-0.045	0.045	-0.072	0.081	0.024	-0.024
5	44-49	1.608	0.098	0.12	-0.11	0.061	-0.061	-0.038	0.038	-0.083	0.093	0.028	-0.028
6	49-54	1.557	0.098	0.12	-0.11	0.062	-0.062	-0.033	0.033	-0.081	0.09	0.026	-0.026
7	54-64	1.291	0.065	0.1	-0.095	0.055	-0.055	-0.04	0.04	-0.065	0.072	0.016	-0.016
8	64-74	0.878	0.057	0.074	-0.07	0.04	-0.04	-0.035	0.035	-0.045	0.05	0.012	-0.012
9	74-84	0.771	0.052	0.065	-0.062	0.038	-0.038	-0.027	0.027	-0.04	0.044	0.012	-0.012
10	84-94	0.613	0.048	0.057	-0.055	0.033	-0.033	-0.027	0.027	-0.033	0.036	0.011	-0.011
11	94-109	0.396	0.03	0.039	-0.038	0.025	-0.025	-0.018	0.018	-0.021	0.023	0.0066	-0.0066
12	109-130	0.23	0.021	0.024	-0.024	0.015	-0.015	-0.013	0.013	-0.012	0.014	0.004	-0.004
13	130-150	0.138	0.016	0.018	-0.018	0.009	-0.009	-0.013	0.013	-0.0076	0.0085	0.0031	-0.0031
14	150-170	0.106	0.015	0.017	-0.017	0.0074	-0.0074	-0.014	0.014	-0.0063	0.0072	0.003	-0.003
15	170-200	0.072	0.01	0.011	-0.011	0.0052	-0.0052	-0.008	0.008	-0.0045	0.0051	0.0021	-0.0021
16	200-240	0.0245	0.0051	0.0045	-0.0044	0.002	-0.002	-0.0035	0.0035	-0.0016	0.0019	0.00089	-0.00089
17	240-300	0.0119	0.0031	0.0034	-0.0034	0.001	-0.001	-0.0031	0.0031	-0.00091	0.0011	0.00054	-0.00054
18	300-500	0.00284	0.00072	0.0013	-0.0013	0.00026	-0.00026	-0.0012	0.0012	-0.00024	0.00029	0.00017	-0.00017

Table C.12: Table showing the breakdown of statistical and systematic errors on the cross-section with respect to M_{jj} .

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	≥ 15	112.86	0.98	0.00301	2.253e+04	1.5e+02	2.3e+03	1.4e+02	0.586	0.0026
2	≥ 20	75.81	0.996	0.00355	1.506e+04	1.2e+02	1639	72	0.5881	0.0032
3	≥ 25	53.5	1.01	0.00401	1.064e+04	1e+02	1316	59	0.5859	0.0038
4	≥ 30	37.93	1.03	0.00436	7821	88	1074	52	0.6086	0.0023
5	≥ 35	28.33	1.04	0.0044	5894	77	905	48	0.6088	0.0026
6	≥ 40	21.82	1.05	0.00375	4541	67	742	44	0.6112	0.003
7	≥ 45	16.97	1.06	0.00426	3570	60	650	41	0.6108	0.0034
8	≥ 50	13.46	1.07	0.00479	2841	53	559	38	0.6078	0.0038
9	≥ 55	10.59	1.08	0.00528	2263	48	484	35	0.605	0.0043
10	≥ 60	8.46	1.09	0.00576	1832	43	419	34	0.6055	0.0018
11	≥ 65	7.06	1.1	0.00617	1518	39	350	31	0.6042	0.0019
12	≥ 70	5.76	1.1	0.00639	1265	36	317	30	0.6026	0.0022
13	≥ 75	4.58	1.09	0.00629	1043	32	287	28	0.6015	0.0024
14	≥ 80	3.81	1.1	0.00566	884	30	264	26	0.599	0.0026
15	≥ 85	2.97	1.1	0.00436	726	27	244	25	0.5979	0.0029
16	≥ 90	2.23	1.11	0.0048	592	24	231	25	0.599	0.0032
17	≥ 95	1.78	1.11	0.00526	490	22	201	23	0.5991	0.0035
18	≥ 100	1.55	1.11	0.00574	420	20	170	20	0.5974	0.0038
19	≥ 105	1.27	1.11	0.00618	359	19	155	20	0.5948	0.0041
20	≥ 110	1.11	1.11	0.00664	303	17	126	17	0.5918	0.0021
21	≥ 115	0.88	1.12	0.00709	255	16	116	17	0.5928	0.0023
22	≥ 120	0.739	1.13	0.00749	217	15	101	16	0.5938	0.0025
23	≥ 125	0.618	1.13	0.00764	191	14	94	16	0.5928	0.0027
24	≥ 130	0.531	1.13	0.00756	159	13	75	14	0.5946	0.0029
25	≥ 135	0.449	1.13	0.00707	139	12	68	14	0.5953	0.0031
26	≥ 140	0.368	1.14	0.00568	121	11	63	14	0.5963	0.0034
27	≥ 145	0.315	1.16	0.00605	102	10	53	13	0.598	0.0037
28	≥ 150	0.217	1.16	0.0064	85	9.2	52	13	0.597	0.004
29	≥ 155	0.154	1.17	0.00674	71	8.4	47	12	0.5968	0.0043
30	≥ 160	0.145	1.18	0.00701	64	8	42	11	0.5992	0.0019
31	≥ 165	0.135	1.18	0.00708	56	7.5	35.4	9.7	0.6001	0.002
32	≥ 170	0.112	1.18	0.00694	50	7.1	33.1	9.6	0.6008	0.0022
33	≥ 175	0.065	1.19	0.00643	42	6.5	32.2	9.6	0.6015	0.0023
34	≥ 180	0.052	1.2	0.00518	34	5.8	26.1	8.5	0.6027	0.0025
35	≥ 185	0.061	1.2	0.0056	32	5.7	22.8	7.9	0.6045	0.0027
36	≥ 190	0.071	1.2	0.00607	29	5.4	18.2	7.2	0.6046	0.0029
37	≥ 195	0.06	1.2	0.00655	27	5.2	17.9	7.2	0.6061	0.0032

Table C.13: Table showing the integrated candidate, total background, acceptance and unsmearing numbers for a range of first jet E_T^{min} thresholds and the inclusive W→ eυ + ≥ 1 jet cross-section that results. Also shown is the statistical error on the candidates, and the systematic errors on the total background and acceptance numbers (the acceptance systematic is the Monte Carlo statistical error only and does not include the flat 5% error on the acceptance.)

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	≥ 15	19.25	0.942	0.00256	4397	66	792	53	0.5884	0.0033
2	≥ 20	11.04	0.988	0.00369	2445	49	477	31	0.5877	0.0042
3	≥ 25	6.81	1.02	0.00499	1538	39	359	26	0.5865	0.0052
4	≥ 30	4.46	1.04	0.00655	1035	32	282	23	0.5848	0.0065
5	≥ 35	3.04	1.06	0.00839	717	27	214	19	0.5858	0.008
6	≥ 40	2.16	1.09	0.0106	533	23	184	18	0.5896	0.0096
7	≥ 45	1.61	1.11	0.0129	410	20	151	17	0.591	0.011
8	≥ 50	1.22	1.11	0.0156	315	18	122	14	0.588	0.013
9	≥ 55	0.96	1.13	0.0188	249	16	101	13	0.582	0.016
10	≥ 60	0.654	1.16	0.0224	189	14	87	13	0.604	0.019
11	≥ 65	0.502	1.18	0.0263	147	12	69	11	0.61	0.022
12	≥ 70	0.359	1.2	0.0307	112	11	56	11	0.622	0.025
13	≥ 75	0.272	1.22	0.0356	87	9.3	44	9.5	0.644	0.028
14	≥ 80	0.123	1.19	0.0396	59	7.7	39.2	9.5	0.64	0.032
15	≥ 85	0.077	1.2	0.0448	45	6.7	32.9	9	0.628	0.036
16	≥ 90	0.064	1.21	0.0511	40	6.3	29.8	8.9	0.641	0.041
17	≥ 95	0.051	1.23	0.0582	32	5.7	24.2	8.4	0.625	0.045

Table C.14: Table showing the integrated candidate, total background, acceptance and unsmearing numbers for a range of second jet E_T^{min} thresholds and the inclusive W$\rightarrow$ eυ + $\geq$ 2 jet cross-section that results.

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	≥ 15	3.49	0.919	0.00485	871	30	237	17	0.558	0.0037
2	≥ 20	1.74	0.988	0.00771	445	21	152	12	0.556	0.005
3	≥ 25	0.829	1.02	0.0111	249	16	115	10	0.5503	0.0067
4	≥ 30	0.406	1.05	0.0157	152	12	87.8	8.9	0.5561	0.0089
5	≥ 35	0.226	1.09	0.0217	103	10	69.1	8.3	0.548	0.012
6	≥ 40	0.114	1.14	0.0296	62	8.1	45.7	6.4	0.544	0.015
7	≥ 45	0.03	1.17	0.0387	31	5.8	26.9	4.4	0.544	0.02
8	≥ 50	0.012	1.17	0.0495	17	4.5	15.4	2.5	0.528	0.025

Table C.15: Table showing the integrated candidate, total background, acceptance and unsmearing numbers for a range of third jet E_T^{min} thresholds and the inclusive W$\rightarrow$ eυ + $\geq$ 3 jet cross-section that results.

Bin	GeV	$\sigma[pb]$	unsmear	Δ_{uns}	Cand	Δ_{Cand}	Bkgd	Δ_{Bkgd}	Acc	Δ_{Acc}
1	≥ 15	0.572	0.889	0.0157	204	14	98	12	0.5508	0.0074
2	≥ 20	0.288	0.989	0.0273	100.8	9.9	52.1	6.7	0.556	0.011
3	≥ 25	0.074	1.02	0.0421	42.8	6.3	30.8	5	0.547	0.016
4	≥ 30	0.025	1.08	0.0649	21.8	4.4	17.8	4.8	0.558	0.025

Table C.16: Table showing the integrated candidate, total background, acceptance and unsmearing numbers for a range of fourth jet E_T^{min} thresholds and the inclusive W$\rightarrow$eυ + $\geq$ 4 jet cross-section that results.

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	≥ 15	112.86	0.84	5.6	-5.6	5.6	-5.6	-0.77	0.77	-0.49	0.5	0.35	-0.35
2	≥ 20	75.81	0.69	5.3	-5.3	5.3	-5.3	-0.41	0.41	-0.4	0.41	0.27	-0.27
3	≥ 25	53.5	0.59	4.6	-4.6	4.6	-4.6	-0.34	0.34	-0.34	0.35	0.21	-0.21
4	≥ 30	37.93	0.5	3.8	-3.8	3.8	-3.8	-0.29	0.29	-0.14	0.14	0.16	-0.16
5	≥ 35	28.33	0.44	3.1	-3.1	3.1	-3.1	-0.27	0.27	-0.12	0.12	0.12	-0.12
6	≥ 40	21.82	0.39	2.6	-2.6	2.6	-2.6	-0.25	0.25	-0.11	0.11	0.078	-0.078
7	≥ 45	16.97	0.35	2.2	-2.2	2.2	-2.2	-0.24	0.24	-0.093	0.094	0.068	-0.068
8	≥ 50	13.46	0.31	1.8	-1.8	1.8	-1.8	-0.22	0.22	-0.084	0.085	0.06	-0.06
9	≥ 55	10.59	0.28	1.5	-1.5	1.5	-1.5	-0.21	0.21	-0.074	0.075	0.052	-0.052
10	≥ 60	8.46	0.26	1.2	-1.2	1.2	-1.2	-0.2	0.2	-0.025	0.025	0.045	-0.045
11	≥ 65	7.06	0.24	1.1	-1.1	1	-1	-0.19	0.19	-0.023	0.023	0.04	-0.04
12	≥ 70	5.76	0.22	0.89	-0.89	0.87	-0.87	-0.18	0.18	-0.02	0.021	0.034	-0.034
13	≥ 75	4.58	0.2	0.73	-0.73	0.71	-0.71	-0.17	0.17	-0.018	0.018	0.026	-0.026
14	≥ 80	3.81	0.18	0.63	-0.63	0.61	-0.61	-0.16	0.16	-0.017	0.017	0.02	-0.02
15	≥ 85	2.97	0.17	0.51	-0.51	0.48	-0.48	-0.16	0.16	-0.014	0.014	0.012	-0.012
16	≥ 90	2.23	0.15	0.4	-0.4	0.37	-0.37	-0.16	0.16	-0.012	0.012	0.0097	-0.0097
17	≥ 95	1.78	0.14	0.34	-0.34	0.3	-0.3	-0.14	0.14	-0.01	0.01	0.0085	-0.0085
18	≥ 100	1.55	0.13	0.3	-0.3	0.27	-0.27	-0.13	0.13	-0.0098	0.0099	0.008	-0.008
19	≥ 105	1.27	0.12	0.25	-0.25	0.22	-0.22	-0.12	0.12	-0.0088	0.0089	0.0071	-0.0071
20	≥ 110	1.11	0.11	0.22	-0.22	0.2	-0.2	-0.11	0.11	-0.004	0.004	0.0066	-0.0066
21	≥ 115	0.88	0.1	0.19	-0.19	0.16	-0.16	-0.11	0.11	-0.0034	0.0034	0.0056	-0.0056
22	≥ 120	0.739	0.094	0.17	-0.17	0.14	-0.14	-0.1	0.1	-0.0031	0.0031	0.0049	-0.0049
23	≥ 125	0.618	0.088	0.15	-0.15	0.12	-0.12	-0.099	0.099	-0.0028	0.0028	0.0042	-0.0042
24	≥ 130	0.531	0.08	0.13	-0.13	0.1	-0.1	-0.09	0.09	-0.0026	0.0026	0.0035	-0.0035
25	≥ 135	0.449	0.075	0.12	-0.12	0.086	-0.086	-0.087	0.087	-0.0024	0.0024	0.0028	-0.0028
26	≥ 140	0.368	0.07	0.11	-0.11	0.071	-0.071	-0.087	0.087	-0.0021	0.0021	0.0018	-0.0018
27	≥ 145	0.315	0.065	0.1	-0.1	0.061	-0.061	-0.082	0.082	-0.0019	0.002	0.0016	-0.0016
28	≥ 150	0.217	0.06	0.093	-0.093	0.043	-0.043	-0.082	0.082	-0.0014	0.0015	0.0012	-0.0012
29	≥ 155	0.154	0.055	0.086	-0.086	0.031	-0.031	-0.081	0.081	-0.0011	0.0011	0.00089	-0.00089
30	≥ 160	0.145	0.053	0.08	-0.08	0.029	-0.029	-0.075	0.075	-0.00046	0.00046	0.00086	-0.00086
31	≥ 165	0.135	0.049	0.07	-0.07	0.027	-0.027	-0.064	0.064	-0.00046	0.00046	0.00081	-0.00081
32	≥ 170	0.112	0.047	0.067	-0.067	0.023	-0.023	-0.063	0.063	-0.0004	0.0004	0.00065	-0.00065
33	≥ 175	0.065	0.043	0.065	-0.065	0.013	-0.013	-0.064	0.064	-0.00025	0.00025	0.00035	-0.00035
34	≥ 180	0.052	0.039	0.057	-0.057	0.011	-0.011	-0.056	0.056	-0.00022	0.00022	0.00023	-0.00023
35	≥ 185	0.061	0.037	0.054	-0.054	0.012	-0.012	-0.052	0.052	-0.00027	0.00027	0.00028	-0.00028
36	≥ 190	0.071	0.036	0.05	-0.05	0.014	-0.014	-0.047	0.047	-0.00034	0.00035	0.00036	-0.00036
37	≥ 195	0.06	0.034	0.049	-0.049	0.012	-0.012	-0.047	0.047	-0.00031	0.00032	0.00033	-0.00033

Table C.17: Table showing the breakdown of statistical and systematic errors on the integrated $W \to e\nu + \geq 1$ jet cross-sections. Columns $\Delta_\sigma^{JES+/-}$, $\Delta_\sigma^{Bkgd+/-}$ and $\Delta_\sigma^{Acc+/-}$ show the plus/minus absolute systematic from the uncertainty on the jet energy scale, background estimation and acceptance calculation respectively. The total systematic and statistical errors are given by columns $\Delta_\sigma^{Syst+/-}$ and Δ_σ^{Stat} respectively. Note that here the acceptance systematic includes the flat 5%.

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	$\geq$ 15	19.25	0.35	1.7	-1.7	1.7	-1.7	-0.28	0.28	-0.11	0.11	0.052	-0.052
2	$\geq$ 20	11.04	0.28	1.3	-1.3	1.3	-1.3	-0.17	0.17	-0.078	0.079	0.041	-0.041
3	$\geq$ 25	6.81	0.23	1	-1	0.99	-0.99	-0.15	0.15	-0.06	0.061	0.033	-0.033
4	$\geq$ 30	4.46	0.19	0.72	-0.72	0.7	-0.7	-0.14	0.14	-0.049	0.05	0.028	-0.028
5	$\geq$ 35	3.04	0.16	0.52	-0.52	0.51	-0.51	-0.12	0.12	-0.041	0.042	0.024	-0.024
6	$\geq$ 40	2.16	0.14	0.4	-0.4	0.38	-0.38	-0.11	0.11	-0.035	0.036	0.021	-0.021
7	$\geq$ 45	1.61	0.13	0.31	-0.31	0.29	-0.29	-0.11	0.11	-0.031	0.032	0.019	-0.019
8	$\geq$ 50	1.22	0.11	0.25	-0.25	0.23	-0.23	-0.091	0.091	-0.027	0.029	0.017	-0.017
9	$\geq$ 55	0.96	0.1	0.21	-0.2	0.18	-0.18	-0.084	0.084	-0.025	0.027	0.016	-0.016
10	$\geq$ 60	0.654	0.088	0.15	-0.15	0.13	-0.13	-0.08	0.08	-0.02	0.021	0.013	-0.013
11	$\geq$ 65	0.502	0.078	0.13	-0.13	0.1	-0.1	-0.074	0.074	-0.017	0.018	0.011	-0.011
12	$\geq$ 70	0.359	0.068	0.1	-0.1	0.072	-0.072	-0.068	0.068	-0.014	0.015	0.0092	-0.0092
13	$\geq$ 75	0.272	0.059	0.082	-0.082	0.054	-0.054	-0.06	0.06	-0.011	0.013	0.0079	-0.0079
14	$\geq$ 80	0.123	0.048	0.064	-0.064	0.025	-0.025	-0.059	0.059	-0.0059	0.0065	0.0041	-0.0041
15	$\geq$ 85	0.077	0.043	0.059	-0.059	0.015	-0.015	-0.057	0.057	-0.0042	0.0047	0.0029	-0.0029
16	$\geq$ 90	0.064	0.04	0.058	-0.058	0.013	-0.013	-0.056	0.056	-0.0038	0.0043	0.0027	-0.0027
17	$\geq$ 95	0.051	0.037	0.056	-0.056	0.01	-0.01	-0.055	0.055	-0.0035	0.004	0.024	-0.024

Table C.18: Table showing the breakdown of statistical and systematic errors on the integrated $W \rightarrow e\nu + \geq 2$ jet cross-sections.

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	$\geq$ 15	3.49	0.16	0.45	-0.45	0.44	-0.44	-0.096	0.096	-0.023	0.023	0.018	-0.018
2	$\geq$ 20	1.74	0.13	0.31	-0.31	0.3	-0.3	-0.071	0.071	-0.015	0.016	0.014	-0.014
3	$\geq$ 25	0.829	0.098	0.21	-0.21	0.2	-0.2	-0.062	0.062	-0.0099	0.01	0.0091	-0.0091
4	$\geq$ 30	0.406	0.079	0.14	-0.14	0.12	-0.12	-0.056	0.056	-0.0064	0.0066	0.006	-0.006
5	$\geq$ 35	0.226	0.069	0.096	-0.096	0.078	-0.078	-0.055	0.055	-0.0047	0.0049	0.0045	-0.0045
6	$\geq$ 40	0.114	0.056	0.062	-0.062	0.042	-0.042	-0.045	0.045	-0.0031	0.0033	0.003	-0.003
7	$\geq$ 45	0.03	0.042	0.034	-0.034	0.012	-0.012	-0.032	0.032	-0.001	0.0011	0.00099	-0.00099
8	$\geq$ 50	0.012	0.033	0.019	-0.019	0.005	-0.005	-0.019	0.019	-0.00055	0.0006	0.00051	-0.00051

Table C.19: Table showing the breakdown of statistical and systematic errors on the integrated $W \rightarrow e\nu + \geq 3$ jet cross-sections.

Bin	GeV	$\sigma[pb]$	Δ_σ^{Stat}	Δ_σ^{Syst+}	Δ_σ^{Syst-}	Δ_σ^{JES+}	Δ_σ^{JES-}	Δ_σ^{Bkgd+}	Δ_σ^{Bkgd-}	Δ_σ^{Acc+}	Δ_σ^{Acc-}	Δ_σ^{Uns+}	Δ_σ^{Uns-}
1	$\geq$ 15	0.572	0.076	0.089	-0.089	0.061	-0.061	-0.063	0.063	-0.0076	0.0078	0.01	-0.01
2	$\geq$ 20	0.288	0.059	0.056	-0.056	0.038	-0.038	-0.04	0.04	-0.0055	0.0058	0.008	-0.008
3	$\geq$ 25	0.074	0.039	0.035	-0.035	0.014	-0.014	-0.031	0.031	-0.0021	0.0022	0.0031	-0.0031
4	$\geq$ 30	0.025	0.028	0.031	-0.031	0.0049	-0.0049	-0.031	0.031	-0.0011	0.0012	0.0015	-0.0015

Table C.20: Table showing the breakdown of statistical and systematic errors on the integrated $W \rightarrow e\nu + \geq 4$ jet cross-sections.

References

[1] P. A. Aarnio *et al.* [DELPHI Collaboration], "The DELPHI detector at LEP," Nucl. Instrum. Meth. A **303**, 233 (1991).

[2] S. Abachi *et al.* [D0 Collaboration], "A Study of the strong coupling constant using W + jets processes", Phys. Rev. Lett. **75**, 3226 (1995).

[3] V. M. Abazov *et al.* [D0 Collaboration], "Measurement of the Cross-Section for Inclusive W Production in the Muon Channel at $\sqrt{s} = 1.96$ TeV using the D0 Detector", D0 Public Note 4750.

[4] V. M. Abazov *et al.* [D0 Collaboration], "Measurement of the Cross-Section for W and Z Production to Electron Final States with the D0 Dectector at $\sqrt{s} = 1.96$ TeV", D0 Public Note 4403.

[5] B. Abbott *et al.* [D0 Collaboration], "Measurement of W and Z boson production cross sections", Phys. Rev. D **60**, 052003 (1999) [arXiv:hep-ex/9901040].

[6] B. Abbott *et al.* [D0 Collaboration], "Differential production cross section of Z bosons as a function of transverse momentum at s**(1/2) = 1.8-TeV", Phys. Rev. Lett. **84**, 2792 (2000) [arXiv:hep-ex/9909020].

[7] B. Abbott *et al.* [D0 Collaboration], "Differential cross section for W boson production as a function of transverse momentum in p anti-p collisions at s**(1/2) = 1.8-TeV", Phys. Lett. B **513**, 292 (2001) [arXiv:hep-ex/0010026].

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

[9] F. Abe *et al* (CDF Collaboration), "The Topology of Three Jet Events in p$\overline{p}$ collisions at $\sqrt{s} = 1.8$ TeV", Phys. Rev. **D45** (1992) 1448.

[10] F. Abe *et al.* [CDF Collaboration], "Properties of jets in Z boson events from 1.8-TeV anti-p p collisions", Phys. Rev. Lett. **77**, 448 (1996) [arXiv:hep-ex/9603003].

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

[12] F. Abe *et al.* [CDF Collaboration], "Properties of jets in W boson events from 1.8-TeV anti-p p collisions", Phys. Rev. Lett. **79**, 4760 (1997) [arXiv:hep-ex/9709016].

[13] F. Abe *et al.* [CDF Collaborations], "Measurement of the sigma(W + ¿= 1jet)/sigma(W) cross section ratio from anti-p p collisions at s**(1/2) = 1.8-TeV", Phys. Rev. Lett. **81**, 1367 (1998).

[14] A. Abulencia *et al.* [CDF Collaboration], "Combination of CDF Top Pair Production Cross-Section Measurements", CDF Public Note 7794 (2005)

[15] A. Abulencia *et al.* [CDF Collaboration], "Precision top quark mass measurement in the lepton + jets topology in p anti-p collisions at s**(1/2) = 1.96-TeV", Phys. Rev. Lett. **96**, 022004 (2006) [arXiv:hep-ex/0510049].

[16] D. Acosta *et al*, "The CDF Cherenkov Luminosity Monitor", CDF Note 5559 (2001)

[17] D. Acosta *et al.* [CDF II Collaboration], "First measurements of inclusive W and Z cross sections from Run II of the Tevatron collider", Phys. Rev. Lett. **94**, 091803 (2005) [arXiv:hep-ex/0406078].

[18] D. Acosta *et al.* [CDF Collaboration], "Measurement of the t anti-t production cross section in p anti-p collisions at s**(1/2) = 1.96-TeV using lepton + jets events with secondary vertex b-tagging", Phys. Rev. D **71**, 052003 (2005) [arXiv:hep-ex/0410041].

[19] A. A. Affolder *et al.* [CDF Collaboration], "The transverse momentum and total cross section of e+ e- pairs in the Z boson region from p anti-p collisions at s**(1/2) = 1.8-TeV", Phys. Rev. Lett. **84**, 845 (2000) [arXiv:hep-ex/0001021].

[20] T. Affolder *et al* (CDF Collaboration), "Tests of Enhanced Leading Order QCD in W Boson Plus Jets Events from 1.8 TeV $p\bar{p}$ Collisions", Phys. Rev. **D63** (2001) 072003.

[21] A. A. Affolder *et al.* [CDF Collaboration], "Intermediate silicon layers detector for the CDF experiment," Nucl. Instrum. Meth. A **453**, 84 (2000).

[22] A. A. Affolder *et al.* [CDF Collaboration], "CDF central outer tracker," Nucl. Instrum. Meth. A **526**, 249 (2004).

[23] K. Ahmet *et al.* [OPAL Collaboration], "The Opal Detector At Lep," Nucl. Instrum. Meth. A **305**, 275 (1991).

[24] I. J. R. Aitchison, A. J. G. Hey, "Gauge Theories in Particle Physics", IOP Publishing, 1989, 1.

[25] C. Albajar *et al.* [UA1 Collaboration], "Studies of intermediate vector boson production and decay in U(A)(1) at the CERN proton - anti-proton collider", Z. Phys. C **44**, 15 (1989).

[26] [ALEPH Collaboration], "A combination of preliminary electroweak measurements and constraints on the standard model", [arXiv:hep-ex/0511027].

[27] M. G. Albrow *et al.* [CDF Collaboration], "The CDF plug upgrade electromagnetic calorimeter: Test beam results", Nucl. Instrum. Meth. A **480**, 524 (2002).

[28] J. Alitti *et al.* [UA2 Collaboration], "A Determination of the strong coupling constant alpha-s from W production at the CERN p anti-p collider", Phys. Lett. B **263**, 563 (1991).

[29] G. Altarelli and G. Parisi, "Asymptotic Freedom In Parton Language," Nucl. Phys. B **126**, 298 (1977).

[30] D. Amati, G. Veneziano, "Preconfinement as a Property of Perturbative QCD", Phys. Lett. **B83** (1979) 87.

[31] B. Andersson, G. Gustafson and B. Nilsson-Almqvist, "A Model For Low P(T) Hadronic Reactions, With Generalizations To Hadron - Nucleus And Nucleus-Nucleus Collisions", Nucl. Phys. B **281**, 289 (1987).

[32] N. Arkani-Hamed, A. G. Cohen and H. Georgi, "Electroweak symmetry breaking from dimensional deconstruction," Phys. Lett. B **513**, 232 (2001) [arXiv:hep-ph/0105239].

[33] G. Arnison *et al.* [UA1 Collaboration], "Further Evidence For Charged Intermediate Vector Bosons At The Sps Collider", Phys. Lett. B **129**, 273 (1983).

[34] A. Artikov *et al.*, "Design and construction of new central and forward muon counters for CDF II", Nucl. Instrum. Meth. A **538**, 358 (2005) [arXiv:physics/0403079].

[35] X. Artru and G. Mennessier, "String Model And Multiproduction", Nucl. Phys. B **70**, 93 (1974).

[36] G. Azuelos *et al.*, "Search For The Standard Model Higgs Boson Using Vector Boson Fusion At The LHC" *Prepared for Workshop on Physics at TeV Colliders, Les Houches, France, 21 May - 1 Jun 2001*

[37] D. Bailin, A. Love, "Introduction to Gauge Field Theory", IOP Publishing, Revised Edition, 1993, 1.

[38] S. Baker, R.D. Cousins, "Clarification of the Use of Chi-Square and Likelihood Functions in Fits to Histograms", *Nuclear Instruments and Methods in Physics Research, 1983.*

[39] L. Balka *et al.* [CDF Collaboration], "The Cdf Central Electromagnetic Calorimeter", Nucl. Instrum. Meth. A **267**, 272 (1988).

[40] M. Begel, D. Edmunds, P. Laurens, R. Partridge, "D∅ Luminosity in Run II: Delivered", D∅ Note 3970 (2003)

[41] S. Bertolucci *et al.* [CDF Collaboration], "The Cdf Central And Endwall Hadron Calorimeter", Nucl. Instrum. Meth. A **267**, 301 (1988).

[42] A. Abulencia *et al.* [CDF Run II Collaboration], "Measurement of the inclusive jet cross section in p anti-p interactions at s**(1/2) = 1.96-TeV using a cone-based jet algorithm," arXiv:hep-ex/0512020.

[43] A. Bhatti, F. Canelli, "Absolute Corrections and their Systematic Uncertainties", CDF Note 7450 (2005)

[44] A. Bhatti, K. Hatakeyama, "Update on the Relative Jet Energy Corrections", CDF Note 7354 (2005)

[45] A. Bhatti *et al.*, "Determination of the jet energy scale at the Collider Detector at Fermilab", arXiv:hep-ex/0510047.

[46] G. C. Blazey *et al.*, "Run II jet physics", arXiv:hep-ex/0005012.

[47] R. N. Cahn, G. Goldhaber, "The Experimental Foundations of Particle Physics", Cambridge University Press, (1989).

[48] J. M. Campbell and R. K. Ellis, "An update on vector boson pair production at hadron colliders", Phys. Rev. D **60**, 113006 (1999) [arXiv:hep-ph/9905386].

[49] J. Campbell and R. K. Ellis, "Next-to-leading order corrections to W + 2jet and Z + 2jet production at hadron colliders", Phys. Rev. D **65**, 113007 (2002) [arXiv:hep-ph/0202176].

[50] M. Carena *et al.* [Higgs Working Group Collaboration], "Report Of The Tevatron Higgs Working Group", arXiv:hep-ph/0010338.

[51] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, "Longitudinally invariant K(t) clustering algorithms for hadron-hadron collisions", Nucl. Phys. B **406**, 187 (1993).

[52] S. Catani, F. Krauss, R. Kuhn, B.R. Webber, "QCD Matrix Elements and Parton Showers", JHEP 0111:063 (2001) hep-ph/0109231.

[53] T-P. Cheng, L-F. Li, "Gauge Theory of Elementary Particle Physics", Oxford University Press, 1984, 1.

[54] Y. S. Chung *et al.*, "The Level-3 Trigger At The Cdf Experiment At Tevatron Run Ii," IEEE Trans. Nucl. Sci. **52**, 1212 (2005).

[55] M. Coca *et al*, "A First Look At Run 2 High P_T Electrons", CDF Note 5803 (2001)

[56] B. Cooper, M. D'Onofrio, B. Heinemann, P. Murat "Z Vertex Finder", CDF Note 7227 (2004)

[57] B. Cooper, M. D'Onofrio, G. Flanagan, "Multiple Interaction Corrections", CDF Note 7365 (2004)

[58] G. Corcella *et al.*, "HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)", JHEP **0101**, 010 (2001) [arXiv:hep-ph/0011363].

[59] D. Cronin-Hennessy, A. Beretvas and P. F. Derwent [CDF Collaboration], "Luminosity monitoring and measurement at CDF," Nucl. Instrum. Meth. A **443**, 37 (2000).

[60] D. P. Cronin-Hennessy, "Tests of Perturbative QCD in W + jets events produced in s**(1/2) 1.8-TeV proton - anti-proton collisions," FERMILAB-THESIS-1997-42 http://www.slac.stanford.edu/spires/find/hep/www?r=fermilab-thesis-1997-42 SPIRES entry

[61] C.L. Cowan, Jr., F. Reines, F.B. Harrison, H.W. Kruse and A.D. McGuire, "Detection of the Free Neutrino: A Confirmation", Science 124, 103 (1956).

[62] D. Decamp *et al.* [ALEPH Collaboration], "Aleph: A Detector For Electron - Positron Annihilations At Lep," Nucl. Instrum. Meth. A **294**, 121 (1990) [Erratum-ibid. A **303**, 393 (1991)].

[63] P. De Barbaro *et al.*, "CDF plug upgrade hadron calorimeter design", UR-1360
http://www.slac.stanford.edu/spires/find/hep/www?r=ur-1360

[64] K. R. Dienes, E. Dudas and T. Gherghetta, "Extra spacetime dimensions and
unification," Phys. Lett. B **436**, 55 (1998) [arXiv:hep-ph/9803466].

[65] P. A. M. Dirac, "The Quantum Theory of the Electron", Proc. Royal. Soc.,
A117, (1928) 610.

[66] J. R. Dittmann, "Measurement of the W + ¿= 1 jet cross section in proton
antiproton collisions at s**(1/2) = 1.8-TeV," FERMILAB-THESIS-1998-07
http://www.slac.stanford.edu/spires/find/hep/www?r=fermilab-thesis-1998-
07 SPIRES entry

[67] Y. L. Dokshitzer, "Calculation Of The Structure Functions For Deep Inelas-
tic Scattering And e+e- Annihilation By Perturbation Theory In Quantum
Chromodynamics. (In Russian)", Sov. Phys. JETP **46**, 641 (1977) [Zh. Eksp.
Teor. Fiz. **73**, 1216 (1977)].

[68] V. Drollinger, "Getting ready for physics at the LHC with the CMS detector,"
arXiv:physics/0604110.

[69] I. Efthymiopoulos [ATLAS Collaboration], "Overview of the ATLAS detector
at LHC," Acta Phys. Polon. B **30**, 2309 (1999).

[70] S. Eidelman et al. (Particle Data Group), Phys. Lett. B **592**, 1 (2004) (URL:
http://pdg.lbl.gov)

[71] S. Eidelman and F. Jegerlehner, "Hadronic contributions to g-2 of the leptons
and to the effective fine structure constant alpha (M(z)**2)", Z. Phys. C **67**,
585 (1995) [arXiv:hep-ph/9502298].

[72] R.K. Ellis, W.J. Stirling and B.R. Webber, "QCD and Collider Physics", Cam-
bridge University Press, (1996).

[73] R. D. Erbacher [CDF 2 collaboration], "Calorime-
try in CDF Run 2," FERMILAB-CONF-02-251-E
http://www.slac.stanford.edu/spires/find/hep/www?r=fermilab-conf-02-
251-e SPIRES entry *Presented at 31st International Conference on High
Energy Physics (ICHEP 2002), Amsterdam, The Netherlands, 24-31 Jul 2002*

[74] R. P. Feynman, R. D. Field and G. C. Fox, "Quantum-Chromodynamic Ap-
proach For The Large-Transverse-Momentum Production Of Particles And
Jets", Phys. Rev. D **18**, 3320 (1978).

[75] R. Field, S. Wolfram "A QCD Model for e⁺e⁻ Annihilation", Nucl. Phys. **B213** (1983) 65.

[76] R. Field, "Min-Bias and the Underlying Event at the Tevatron and LHC", talk presented at the Fermilab ME/MC Tuning Workshop, Fermilab (2004).

[77] http://root.cern.ch/root/html404/TFractionFitter.html

[78] Y. Fukuda *et al.* [Super-Kamiokande Collaboration], "Evidence for oscillation of atmospheric neutrinos," Phys. Rev. Lett. **81**, 1562 (1998) [arXiv:hep-ex/9807003].

[79] S. Fukuda *et al.* [Super-Kamiokande Collaboration], "Constraints on neutrino oscillations using 1258 days of Super-Kamiokande solar neutrino data," Phys. Rev. Lett. **86**, 5656 (2001) [arXiv:hep-ex/0103033].

[80] S. L. Glasow, "Partial Symmetries of Weak Interactions", Nucl. Phys. **22**, (1961)

[81] D. Glenzinski et al, "WW Cross-Section Measurement using the Tight Lepton and Isolated Track Sample", CDF Note 6872 (2004)

[82] V. N. Gribov and L. N. Lipatov, "E+ E- Pair Annihilation And Deep Inelastic E P Scattering In Perturbation Theory", Sov. J. Nucl. Phys. **15**, 675 (1972) [Yad. Fiz. **15**, 1218 (1972)].

[83] M. Griffiths, B. Heinemann, G. Manca, "Measurement of the Fake Rate of Electrons and Muons", CDF Note 7470 (2005)

[84] D. J. Gross, F. Wilczek, "Asymptotically Free Gauge Theories", Phys. Rev. **D8**, 3633-3652, (1973).

[85] U. Grundler, A. Taffard, X. Zhang, "High-P_T Muons Recommended Cuts and Efficiencies for Winter 2006", CDF Note 7956 (2005)

[86] F. Halzen, A. D. Martin, "Quarks and Leptons: An Introductory Course in Modern Particle Physics", Wiley, 1982, 1.

[87] R. Hamberg, W.L. van Neerven, T. Matsuura, "A Complete Calculation of the Order α_s^2 Correction to the Drell-Yan K Factor", Nucl. Phys. **B359** (1991) 343-405.

[88] C. Hays [CDF and D0 Collaborations], "W boson mass and width at the Tevatron", *Prepared for 32nd International Conference on High-Energy Physics (ICHEP 04), Beijing, China, 16-22 Aug 2004*

[89] S. W. Herb *et al.*, "Observation of a dimuon resonance at 9.5-GeV in 400-GeV proton - nucleus collisions", Phys. Rev. Lett. **39**, 252 (1977).

[90] P. W. Higgs, "Brocken Symmetries, Massless Particles and Gauge Fields", Phys. Lett. **12**, (1964), 132-133.

[91] P. W. Higgs, "Brocken Symmetries and the Masses of Gauge Bosons", Phys. Rev. Lett. **13**, (1964), 508-509.

[92] C. S. Hill [On behalf of the CDF Collaboration], "Operatiaonal experience and performance of the CDFII silicon detector," Nucl. Instrum. Meth. A **530**, 1 (2004).

[93] Y. Ishizawa, J. Nielsen, "Trigger Efficiencies for High E_T Electrons", CDF Note 7401 (2005)

[94] C. Itzykson, J-B. Zuber, "Quantum Field Theory", McGraw-Hill, 3rd edition, 1987, 1.

[95] S. Jindariani *et al* "Luminosity Uncertainty for Run II up until August 2004", CDF Note 7446 (2005)

[96] S. Klimenko [CDF and D0 Collaborations], "Luminosity determination at the Tevatron", http://www.slac.stanford.edu/spires/find/hep/www?irn=5764882 *Prepared for 14th Topical Conference on Hadron Collider Physics (HCP 2002), Karlsruhe, Germany, 29 Sep - 4 Oct 2002*

[97] S. Kuhlmann *et al.*, "The CDF calorimeter upgrade for Run IIB," Nucl. Instrum. Meth. A **518**, 39 (2004) [arXiv:physics/0310155].

[98] M. Lindgren, M. Ikeda, D. Joyce, A. Kernan, J. P. Merlo, D. Smith and S. J. Wimpenny, "A Measurement of the strong coupling constant alpha-s in W boson production at the CERN proton - anti-proton collider", Phys. Rev. D **45**, 3038 (1992).

[99] L. Lonnblad, "Correcting the Colour-Dipole Cascade Model with Fixed Order Matrix Elements", JHEP 0205:046 (2002) hep-ph/0112284

[100] P. T. Lukens [CDF IIb Collaboration], "The Cdf Iib Detector: Technical Design Report," FERMILAB-TM-2198 http://www.slac.stanford.edu/spires/find/hep/www?r=fermilab-tm-2198 SPIRES entry

[101] L. Malgeri, "W W cross section and branching fraction measurements at LEP," *Prepared for 29th International Conference on High-Energy Physics (ICHEP 98), Vancouver, British Columbia, Canada, 23-29 Jul 1998.*

[102] F. Maltoni and T. Stelzer, "MadEvent: Automatic event generation with Mad-Graph", JHEP **0302**, 027 (2003) [arXiv:hep-ph/0208156].

[103] M. Mangano, "Merging Multijet Matrix Elements and Shower Evolution in Hadronic Collisions", http://mlm.home.cern.ch/mlm/talks/lund-alpgen.pdf

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

[105] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, "Parton distributions: A new global analysis", Eur. Phys. J. C **4**, 463 (1998) [arXiv:hep-ph/9803445].

[106] B. R. Martin, G. Shaw, "Particle Physics", Wiley, 1992, 1.

[107] A. Messina, S. Sarkar, "Measurement of the W$\to$ ev + $\geq$n Jet Cross-Section in p$\overline{\text{p}}$ Collisions at $\sqrt{\text{s}} = 1.96$ TeV", CDF Note 6767 (2001)

[108] R. S. Moore, "An overview of Tevatron collider run II at Fermilab," FERMILAB-CONF-02-320-E http://www.slac.stanford.edu/spires/find/hep/www?r=fermilab-conf-02-320-e SPIRES entry

[109] H. P. Nilles, "Supersymmetry, Supergravity And Particle Physics," Phys. Rept. **110**, 1 (1984).

[110] D. H. Perkins, "Introduction to High Energy Physics", Addison-Wesley, 1987, 1.

[111] H. D. Politzer, "Asymptotic Freedom: An Approach to Strong Interactions", Phys. Rept. **14**, 129-180, (1974).

[112] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky and W. K. Tung, "New generation of parton distributions with uncertainties from global QCD analysis", JHEP **0207**, 012 (2002) [arXiv:hep-ph/0201195].

[113] L. Randall and R. Sundrum, "A large mass hierarchy from a small extra dimension," Phys. Rev. Lett. **83**, 3370 (1999) [arXiv:hep-ph/9905221].

[114] W. Sakumoto, "Event $|Z_{vtx}| < 60$ Cut Acceptance for Run II", CDF Note 7935 (2005)

[115] A. Salam, J. C. Ward, "Electromagnetic and Weak Interactions", Phys. Lett. **13**, (1964).

[116] A. Salam, "Proceedings of the Eigth Nobel Symposium", ed. N. Svartholm, Amkvist and Wiksell, Stockholm, 1968, 367-377.

[117] S. Schael *et al.* [ALEPH Collaboration], "Measurement of the W boson mass and width in e+ e- collisions at LEP," Eur. Phys. J. C **47**, 309 (2006) [arXiv:hep-ex/0605011].

[118] A. Sill [CDF Collaboration], "CDF Run II silicon tracking projects," Nucl. Instrum. Meth. A **447**, 1 (2000).

[119] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands, "PYTHIA 6.3: Physics and manual", arXiv:hep-ph/0308153.

[120] M. Spira, "QCD effects in Higgs physics", Fortsch. Phys. **46**, 203 (1998) [arXiv:hep-ph/9705337].

[121] S. Stapnes [UA2 Collaboration], "New results on W and Z0 physics from the UA2 Collaboration," *Prepared for Les Rencontres de Physique de la Vallee d'Aoste: Results and Perspectives in Particle Physics, La Thuile, Italy, 3-9 Mar 1991*

[122] A. Stange, W. J. Marciano and S. Willenbrock, "Higgs Bosons At The Fermilab Tevatron," Phys. Rev. D **49**, 1354 (1994) [arXiv:hep-ph/9309294].

[123] H. Stockhorst, "Stochastic beam cooling: An introduction," *Prepared for Symposium on Threshold Meson Production in p p and p d Interaction (COSY-11), Cracow, Poland, 20-24 Jun 2001*

[124] R. Wagner, "Electron Identification for Run II: Algorithms", CDF Note 5456 (2003)

[125] S. Weinberg, "A Model of Leptons", Phys. Rev. Lett. **19**, 1264-1266 (1967).

[126] S. Wolfgang, W. Pekka, K. Sinervo "An Evaluation of the Beam Width at CDF", CDF Note 6311 (2003)

[127] K. Yasuoka, S. Mikamo, T. Kamon and A. Yamashita [CDF NW Wedge Group], "Response Maps Of The CDF Central Electromagnetic Calorimeter With Electrons", Nucl. Instrum. Meth. A **267**, 315 (1988).