First Year Transfer Report

Ben Cooper

June 2003

High Energy Physics Group
Department of Physics and Astronomy

Abstract

This report describes a preliminary investigation into $p\overline{p} \rightarrow \mathrm{W} + \geq \mathrm{n}$ jet events at the CDF Run II detector. Outlined is the physics motivation for this study, particuarly in the wider context of CDF Run II as a whole, and the results that have been achieved thus far. These include a largely successful reproduction of the equivalent analysis in Run I, and detailed comparisons with leading order Monte Carlo event simulations.

1 The CDF II Experiment

1.1 Run II Detector Upgrade

The Collider Detector at Fermilab (CDF) is a general purpose experiment for the study of $p\overline{p}$ collisions at the Fermilab Tevatron Collider. The first collisions were produced and detected in October 1985 and after ten years of data acquisition, which resulted most notably in the discovery of the top quark [1], a major upgrade to both detector and collider was undertaken [2]. The collider luminosity and centre of mass energy were to ultimately be increased to $2\times10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and 2 TeV respectively, and to cope with the increased bunch crossing rates, radiation exposure and physics demands the detector had to be upgraded. These included the central tracking chamber, the silicon vertex detector and the forward and plug calorimeters. In addition the η coverage of the muon chambers was extended.

The intial luminosity goal for Run II was $5\times10^{31}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, with an integrated luminosity target of $2\mathrm{fb}^{-1}$ being collected within two to three years, and $15\mathrm{fb}^{-1}$ prior to LHC turn on [3]. Run IIa began nominally in July 2001 at a centre of mass energy of 1.96 TeV , but reliable data acquisition was not acheived until March 2002. Since then the average instantaneous luminosity has been $2.3\times10^{31}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, and only $146\mathrm{pb}^{-1}$ of 'good' data (i.e. detector components functioning) has been collected. Deficiencies in the $\overline{p}$ stacking rate, p beam brightness and unforeseen problems with the recycler are among the many reasons for the shortfall, and although the introduction of electron cooling in the recycler could bring improvements, the initial luminosity goals aleady look far from achievable [4]. Recent projections put the total Run II accumulated luminosity at 6.5 - $11\mathrm{fb}^{-1}$, limiting Higgs sensitivity somewhat, but still allowing for an extensive programme of precision physics measurements.

1.2 Run II Physics Goals

The dramatically increased statistics of Run II over Run I, as well as improvements to the detector itself, should enable all Standard Model measurements that were made in Run I to be repeated with far greater precision. In addition the increased centre of mass energy when combined with the higher luminosity opens up the possibility of the discovery of new

physics beyond the Standard Model, such as supersymmetric particles, large extra dimensions or unexpected high mass resonances.

The discovery of the top quark in CDF Run I was crucial in confirming the Standard Model, and full characterization of the properties of the top quark is an obvious starting point for Run II. Not only is the top mass one of the key electroweak parameters of the Standard Model, but being the only elementary fermion that is predicted to strongly couple to the electroweak symmetry breaking mechanism, measurement of the top mass also indirectly constrains that of the Higgs within the Standard Model framework [5].

Precise measurement of the other electroweak parameters should also be possible in Run II. With $2fb^{-1}$ of data some 4.3 million W boson events and 600,000 Z boson events should be identifiable [2]. Using techniques already developed in Run I the W mass, Z mass and W width will all be measured to record precision and diboson production cross section measurements, previously unmeasureable in Run I, will probe EWK couplings and also be sensitive to new physics.

Unlike the electroweak sector, QCD is a component of the Standard Model which has not yet been subjected to precise testing. Being a high statistics, high energy $p\overline{p}$ collider and detector system, CDF Run II provides an important testing ground for the latest generation of perturbative QCD predictions. A wide range of QCD events can be triggered on and studied at Q^2 scales never before tested in an experiment, such as mulitjet events and direct photon, W and Z boson production in association with a number of jets. In studying these processes, the ability of pertrubative QCD to not only predict cross-sections but also a wide range of event shapes can be put to the test.

Of course, the ultimate goal of CDF Run II is the unambiguous discovery of the Higgs boson before the LHC is active. Direct Higgs production via gluon fusion is predicted to be the dominant production mechanism at the Tevatron, but this channel suffers from extreme backgrounds. For $m_h \leq 130$ GeV/c^2 the most likely process to result in discovery is associated production V + h, where V = W, Z and h$\rightarrow b\overline{b}$. For a heavier Higgs, the dominant decay mode is no longer $b\overline{b}$, but WW, and so different search strategies have to be employed. It has been predicted that by combining all the data from both CDF and D0 in all available channels, the Tevatron could produce a 5 σ discovery of the Higgs with $30fb^{-1}$ of data [6]. However, since that level of data accumulation is looking less likely, a more realistic goal is perhaps to exclude the Higgs mass up to 130 GeV/c^2 at the 95% confidence level with $6fb^{-1}$ [7].

2 Direct W + Jets Processes at CDF II

2.1 Process Description

At the Tevatron Run II protons and antiprotons are collided at a centre-of-mass energy of 1.96TeV. Since $\sqrt{s} \gg M_W$ direct real W production is not limited to the high-x region of the

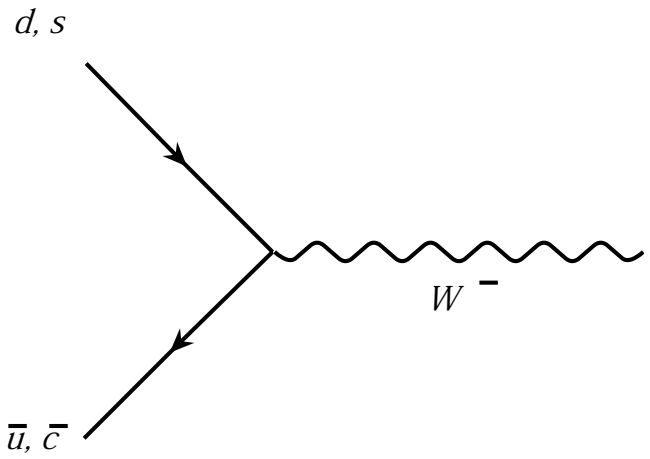

Figure 1: Feynman diagram for direct W^- production

parton density functions, and the inclusive cross-section is large; 29.77nb according to recent NNLO calculation [8]. A leading order Feymann diagram for such a process is shown in Figure 1. Note that although contributions from top and bottom quarks are allowed by the CKM matrix, the probability of this is effectively zero due to the parton density functions.

What singles this process out as direct is that the boson has been produced by parton fusion in the initial hard scatter, $q\overline{q} \rightarrow$ W. QCD predicts that one or more partons can also be produced in the hard scatter in association with the W boson, and leading order Feynmann diagrams for direct W + 1 parton processes are shown in Figure 2. The first of these diagrams shows the intial state quark in W production coupling to a gluon via the strong interaction, and thus a high P_T gluon is "seen" in the final state. The second diagram shows a gluon in the initial state which splits into a quark pair, one of which is involved in W prodution, and the other is "seen" in the final state.

One of the fundamental predictions of QCD is that we can only ever directly observe the colour singlet states of hadrons and mesons, that is bound states of two or three quarks. Thus after production in the hard scatter, final state partons immediately undergo a process of hadronization which results in a large number of hadrons with momenta distributed around that of the parent parton, a so called jet. It is in this form that the partons are "seen" or detected in the final state. Exactly how a jet is defined is a discussed in sections 2.3.2 and 3.2.1.

Of the decay channels available to the W boson, W$\rightarrow$ q$\overline{\text{q}}$ (branching fraction 67.8$\pm$1.0%) dominates. However, searches in this channel would be virtually impossible due to the huge number of QCD multijet events produced at the Tevatron. The W$\rightarrow$ eυ channel (branching fraction 10.9$\pm$0.4%) allows for much easier identification of the W boson via the presence of a high P_T electron and significant missing transverse energy. Due to the high W production cross-section this channel alone can provide sufficient statistics that the more complex muon and tau channels need not be considered.

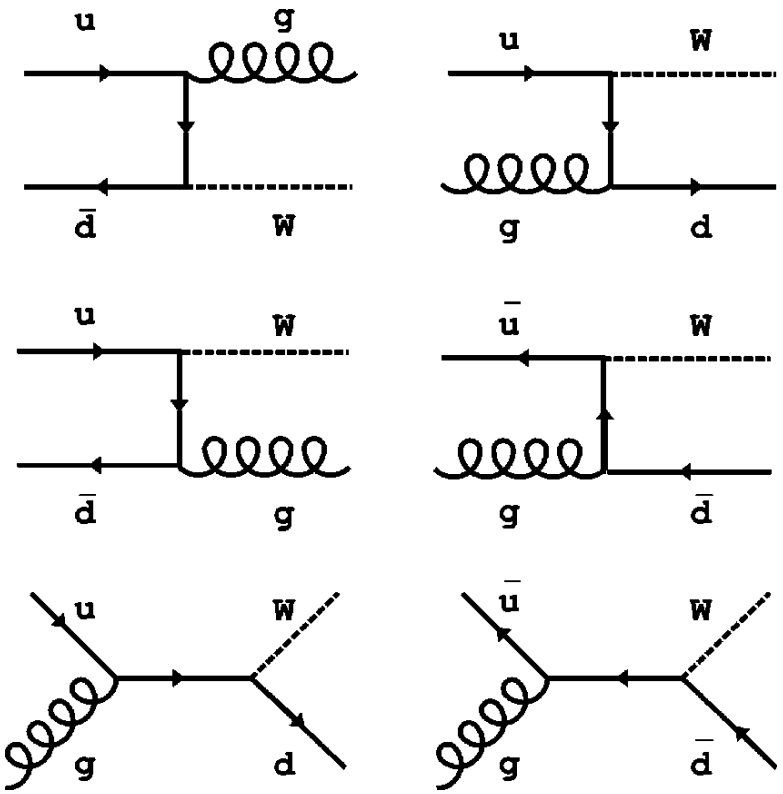

Figure 2: Feynman diagrams for some of the leading order processes that produce a W boson with an associated jet.

2.2 Physics Motivation

Why should we choose to make a detailed study of this process in particular? Firstly, it is a major background to two of the most important Run II processes, top quark production and light Higgs production in association with a W boson (see section 1.2). Figures 3 and 4 show the Feynman diagrams for the particular channels of these processes within which direct W + jets production makes its unwanted contribution. Both result in a high P_T lepton and several high P_T jets in the final state. In order to be able to accurately estimate both the cross-section and reconstructed event properties (such as mass) for these processes it is crucial to have a good understanding of the background contribution to the sample [1] [9]. Of course in practice this understanding is implemented via the latest Monte Carlo tools, but to have faith in such predictions you need to test them rigorously by making detailed comparisons with real data. This should involve not only testing the ability of the Monte Carlo to predict the correct number of events, but also to accurately reproduce various event shapes, such as jet E_T and angular separation distributions. Such comparisons are interesting and important in their own right, especially for the W + jets process, as they test the very latest QCD predictions and hence our current understanding of this challenging aspect of the standard model, another key goal for CDF Run II.

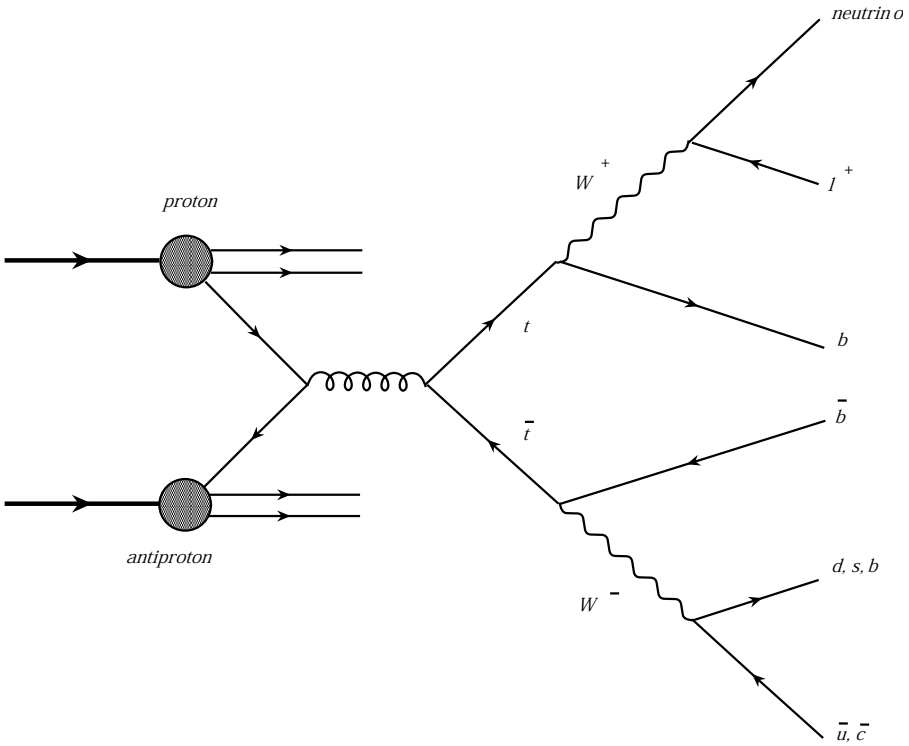

Figure 3: Feynman diagram for the lepton + jets channel of top production at CDF

2.3 Implementing Theoretical Predictions for W + Jets

2.3.1 Monte Carlo Generators

ALPGEN and VECBOS are both generators which can be used to produce W + n parton events at leading order [14]. From the full (infinite) set of possible diagrams, they make an approximate calculation by using only a subset of these diagrams with the lowest order of α_S, the so called leading order diagrams. These are evaluated at a predetermined "renormalisation" energy scale, Q^2, which effectively determines α_S. There is no correct choice of renormalisation scale as such, but it should reflect the energy scale of the hard scatter. Another parameter which also has to be predetermined is the factorisation scale, that used to evaluate the proton/anti-proton parton distribution functions (PDFs). This is generally set equal to the renormalisation scale.

Such LO predictions are expected to perform well when the final state partons are well separated. However, the exclusion of higher order diagrams in the matrix element evaluation means that they cannot predict the existence of additional final state partons that have resulted from initial or final state gluon radiation. Next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) calculations include a larger set of the higher order diagrams and thus can to some extent make these predictions, but these are only just being implemented in the latest gen-

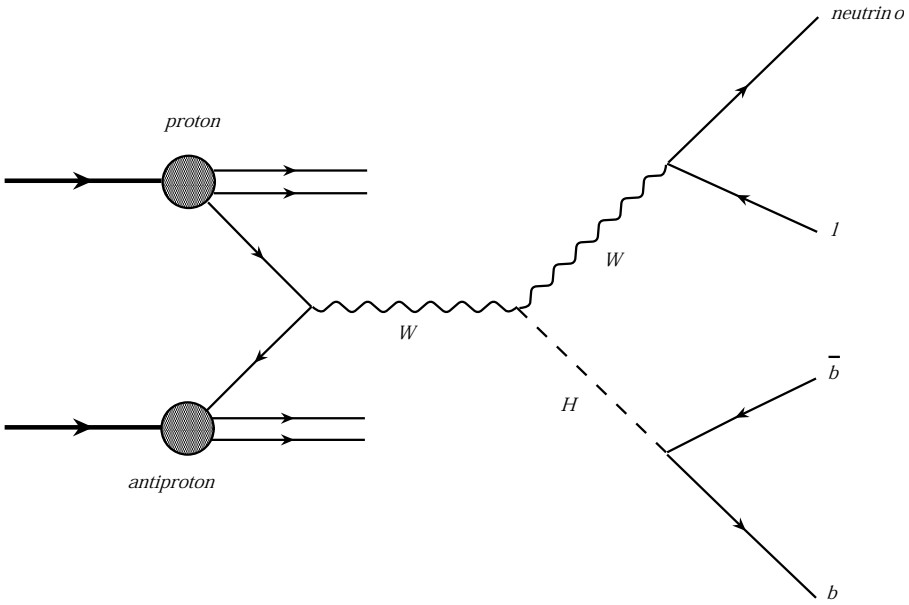

Figure 4: Feynman diagram for Higgs production in association with a W boson that decays leptonically

eration of Mone Carlo generators, and are pushing the practical boundaries of matrix element computation. The current "solution" to dealing with higher order corrections is to interface a leading order matrix element calculation of the hard scatter process with a parton showering simulation, such as HERWIG [15]. HERWIG provides a color-coherent shower evolution of the final state using a model well motivated by theory but which doesn't include matrix element calculations. One sets a "fragmentation" scale which effectively determines how soft the radiated gluon distribution will be, and thus can have a significant impact on the number of final state partons. The combination of a tree-level W + n parton generator such as ALPGEN with a parton shower program such as HERWIG is commonly referred to as an Enhanced Leading Order (ELO) approach.

Thus far we have been discussing W + n parton event generation. If we are to directly compare real data with Monte Carlo generated data, we clearly need some method to evolve these final state partons into hadronic jets which can be passed through a detector simulation. Unfortunately, hadronization is a long distance, low energy QCD process which simply cannot be calculated peturbatively. In practice HERWIG is used to implement phenomenological models of hadronization that have been tuned to describe data.

2.3.2 Defining a Jet: The Jet Algorithm

The final state of a W + n parton process will consist of a large number of high energy hadrons, the vast majority of which will have resulted from the hadronization of a final state parton.

These hadrons will manifest themselves as localised energy deposits in the calorimeter towers of the detector, real or simulated. The purpose of the jet clustering algorithm is to "cluster" these deposits into jets which accurately reflect the energy and momentum of the parent partons.

However, as the current focus of this and other QCD studies is to make meaningful comparisons between perturbative QCD predictions and real data, one does not want a clustering procedure which is sensitive to the details of the showering and hadronization processes that are not explicitly calculated. This translates into the desire for a jet algorithm that is infared and collinear safe. The cone algorithm Jet Clu is that which was favoured in CDF Run I, but it's use of seeding to reduce computation time means it does not meet these requirements [16]. In addition, a further problem of any cone algorithm is how to treat energy sharing between overlapping cones. Generally it is accepted that an additional parameter which dictates the limit in ΔR[1] jet separation before two jets are merged (R_{Sep}) has to be incorporated into the cross-section definition [17], something that is clearly undesirable when one is trying to compare with perturbative predictions. It is clearly highly desirable to move to a more theoretically accceptable algorithm as soon as possible, such as the K_T clustering algorithm [18].

Once the jets have been formed by the clustering procedure a few additional requirements are necessary to ensure a sensible jet definition. All definitions include some minimum jet E_T requirement to restrict our studies to jets from partons produced in the hard scatter. Depending on the nature of the study, the energy of the jets may also be subject to corrections intended to bring the measured energy closer to that of the parent parton, as discussed in section 3.2.1.

2.3.3 The Double Counting Problem of ELO Simulations

Unfortunately there is an inherent problem associated with the ELO solution of simply adding showers to a leading order calculation, known as "double counting". Naively one might assume that data and Monte Carlo could be compared in exclusive jet multiplicity bins. The W + 1 jet sample would be produced using the W + 1 parton Monte Carlo, the W + 2 jet produced using the W + 2 parton Monte Carlo and so on. However, this ignores the overlapping of phase space between an enhanced leading order W + 1 parton calculation and a leading order W + 2 parton calculation. For example, in which bin should a 2 jet event produced by showering in the ELO W + 1 parton generation be placed? Such an event is already accounted for to some extent in the W + 2 parton leading order calculation, and thus to place it in the 2 jet bin involves "double counting".

To overcome this problem data and Monte Carlo are compared in inclusive multiplicity bins such that W + $\geq$ n jet events are produced using the ELO W + n parton Monte Carlo. However, this is far from a satisfactory solution, as it effectively introduces phase space deficiencies that are evident in certain distributions (see Section 3.4.2). An ELO W + 1 parton Monte Carlo cannot be expected to predict the rate of 3 or 4 jet events well, for example.

[1] $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2})$

2.4 Run I Studies

The W + Jets process was studied in detail using $108pb^{-1}$ of data taken during Run Ia and Run Ib [10] [11] [12] [13]. The main features of the analysis were:

- Detailed calculation of the inclusive cross-sections for $p\overline{p} \rightarrow$ W$\rightarrow$ eυ + $\geq$ n jet events.
- Comparison of these with VECBOS + HERWIG ELO simulations for two choices of renormalisation/factorisation scale: $\langle P_T \rangle^2$ and $(M_W^2 + P_{TW}^2)$.
- Examination of the $\sigma_n/\sigma_{(n-1)}$ ratio.
- Examination of jet kinematic distributions for various n and comparison of these to the same Monte Carlo: Jet E_T, combined invariant mass and separation in R.

The Run I study concluded that the ELO QCD predictions reproduced the main qualitative featues of the data for cross sections and jet kinematics. However, it was also clear that some of the distributions would benefit from true higher order corrections. In particular several kinematic distributions highlighted the failure of ELO QCD to produce the observed number of multijet events in the higher jet E_T regions, attributed to the absence of higher order QCD production diagrams. The large variation in cross section predictions with the renormalization scale also indicated that higher order corrections to the LO cross sections are substantial. The lower scale agreed better in magnitude, whilst the higher scale agreed better with the slope of cross-section versus the number of jets. At this level of calculation which renormalization scale does one choose to believe?

3 Preliminary Study into Direct W($\rightarrow e\upsilon$) + Jets Production at CDF Run II

3.1 Overview

The 89.9pb^{-1} of data collected at CDF between March 2001 - December 2002 using the high P_T ELECTRON_CENTRAL_18 trigger has been used to produce a preliminary study of W + Jets events in Run II. Using the same selection criteria as employed in Run I 44836 W + Jets events have been identified. The dominant contribution to this sample will be direct W production in association with n jets, but contributions from background processes such as top quark production and QCD multijet events will also be present. These have not yet been accounted for. Event numbers and jet kinematic distributions for various jet multiplicities have been produced and compared with those from the equivalent Run I analysis. In addition, comparisons have also been made with samples produced using ELO Alpgen + HERWIG and Pythia Monte Carlo Simulations.

Electron Cut	Requirement
E_T	$\geq$ 20 GeV
P_T	$\geq$ 13 GeV
$\mid \eta \mid$	≤ 1.1
Fiducial Volume	Yes
E/P	≥ 0.5 and ≤ 2
Had/Em	$\leq 0.055 + .0045$*E
Frac. Isolation	≤ 0.1
Z Vertex	$\mid z \mid \leq 60$ cm
L_{shr}	≤ 0.2
$\mid \Delta x \mid$	≤ 1.5 cm
$\mid \Delta z \mid$	≤ 3.0 cm
χ^2_{str}	≤ 10.0

Table 1: Table summarising the tight electron quality cuts. The E_T kinematic cut has also been included for completeness.

3.2 Forming the W($\rightarrow e\upsilon$) + Jets Sample

Since we wish first to compare the Run II data to that of the equivalent Run I analysis the method used to form the W($\rightarrow e\upsilon$) + jets sample has been as far as possible a replica of that used in Run I, discussed in [10]. The procedure can be summarised as follows:

1. Select only those events which have a high P_T electron that passes stringent ID cuts.
2. Subject these to a missing transverse energy cut to enrich the sample in W boson events.
3. Reject events consistent with Z boson production.
4. Classify the event according to the number of jets present.

The initial ELECTRON_CENTRAL_18 sample requires at least one electromagnetic cluster with E_T greater than 18 GeV, as can be clearly seen in Figure 5. However, these clusters are only electrons in the loosest sense. In order to increase the electron purity of the sample a variety of electron quality cuts are applied which are summarised in Table 1. In particular the Had/Em and fractional isolation cuts are very important in reducing the background from electron-like jets. Additionally geometric cuts, such as the fiducial and pseudorapidity requirements, are present to limit the electrons to regions of the calorimeter which are relatively well understood. These remain the same as in Run I, but since the detector has changed may need revision. Once the quality cuts have established the presence of one or more tight electrons in the event, the $E_T > 20$ GeV requirement selects high transverse energy electrons. Such a sample will comprise not only of W$\rightarrow e\upsilon$ decay events, but Z $\rightarrow e^+e^-$ decay events also. Figure 6 shows

Electron Cut	Requirement
E_T	$\geq$ 20 GeV
Had/Em	$\leq$ 0.055 + .0045*E
Frac. Isolation	$\leq$ 0.1
M_{ee}	76 GeV/c^2 $\leq M_{ee} \leq$ 106 GeV/c^2

Table 2: Table summarising cuts applied to a second electron to test for consistency with Z $\rightarrow e^+e^-$ production. M_{ee} is the invariant mass of the already identified tight electron and the second electron. Note that there is no pseudorapidity restriction on this second electron.

the missing transverse energy ($\not\!\!E_T$) distribution for such events. The presence of neutrinos from W$\rightarrow$ eυ is clearly evident in the high $\not\!\!E_T$ peak. To isolate this W component one requires that the corrected $\not\!\!E_T$ exceeds 30 GeV. Details of the $\not\!\!E_T$ correction can be found in [10].

This $\not\!\!E_T$ cut will remove the vast majority of Z $\rightarrow e^+e^-$ events from the sample. However, to increase the purity further we demand that each event contain exactly one high E_T tight electron, and in addition subject the other electromagnetic clusters in the event to a looser set of electron cuts designed to identify the second electron from Z decay. These cuts are outlined in Table 2. Any event with a second cluster that passes these cuts is rejected.

One is left with a sample of W candidate events, the purity of which is well illustrated by a plot of the reconstructed W transverse mass, Figure 7. The next stage is to classify each event according to how many jets it contains. The exact defintion of a jet is somewhat arbitrary, and was discussed in section 2.3.2. In this analysis, the Run I jet definition is largely adopted with a few differences. Jets are formed by clustering calorimeter energy deposits with the Run I seeded cone algorithm JetClu. In Run I, the energy of the event's "true" electron was first removed from the calorimeter before any clustering took place. However, due to the nature of the ntuples used in this analysis this is not currently possible, and thus one of the jets in each event is identified as the electron via matching in R^2 and removed from the jet list. The ramaining jets have their energies fully corrected (see section 3.2.1), and are finally defined as an associated jet if they pass the kinematic cuts $\mid \eta \mid < 2.4$ and $E_T > 15$ GeV.

3.2.1 Jet Energy Corrections

The Run II corrections are summarised in Table 5, and explained in full in [19]. Some of these corrections are simply to implement the hadronic energy scale of the calorimeter which varies across the detector in η, and should clearly always be implemented. However, the out-of-cone, underlying event and muiltiple interaction corrections are intended to correct from the energy measured in the jet cone back to the energy of the parent parton. Thus, the extent to which these corrections are applied depends on exactly how one chooses to define a jet.

[2] R = $\sqrt{\eta^2 + \phi^2}$

No. Jets	N_{RI}	L_{RII}/L_{RI}	σ_{RII}/σ_{RI}	$N_{RI} \times L_{RII}/L_{RI} \times \sigma_{RII}/\sigma_{RI}$	N_{RII}	% Diff.
≥ 0	51431	0.83	1.10	47092	44836	-5%
1	8548	0.83	1.14	8112	7153	-12%
2	2016	0.83	1.18	1980	1873	-5%
3	454	0.83	1.24	468	441	-6%
4	105	0.83	1.19	104	123	+15%

Table 3: Table showing the deficiency in the number of events in Run II (N_{RII}) when compared to the number one would expect on scaling the Run I results (N_{RI} Scaled)

The out-of-cone correction attempts to account for energy which falls outside of the jet's cone but which has actually originated from the same parent parton, increasing the jet energy in doing so. The multiple interaction and underlying event corrections reduce the jet energy according to the hypothesis that part of this energy on average results from additional $p\overline{p}$ interactions within the same bunch crossing and soft particles produced in the long range QCD interactions between the p and $\overline{p}$ remnants.

3.3 The Monte Carlo Samples

Table 4 details each of the Monte Carlo data samples that were used. A full decription can be found at [20]. Note that PYTHIA was used for the zero parton generation as opposed to ALPGEN + HERWIG due to a bug in the ALPGEN + HERWIG 0 parton samples. Hereafter the samples will be referred to by the names given in Table 4.

3.4 Results So Far

3.4.1 Comparisons with Run I

In order to establish that the event selection and reconstruction procedure is reliable one should first compare the number of W + Jet events that have been obtained in Run II (N_{RII}) to those of Run I (N_{RI}). To make a meaningful comparison one should obviously scale the Run I numbers by the Run II/Run I luminosity ratio (L_{RII}/L_{RI}), and also take account of the increase in cross-section due to the increased centre-of-mass energy in Run II (σ_{RII}/σ_{RI}). Table 3 compares Run II event numbers to Run I event numbers scaled in this fashion. Note that these are not background corrected in either case. This comparison shows a significant deficiency in the number of events observed in Run II.

In Figure 8 a comparison of the shape of the jet multiplicity distribution in Run I and Run II is made. One can see that the level of agreement is good in the lower multiplicity bins, but

begins to fail in the region of higher jet multiplicity, with Run II displaying a larger fraction of high multiplicity events.

Figure 9 compares the jet transverse energy distribution of Run I and Run II across a range of jet multiplicities. Exactly how and where the distributions differ can be more clearly seen in Figure 10. Although there is a high level of agreement in the lower energy bins, the Run II results show a greater number of events in the high energy bins across all jet multiplicities, indicating a "harder" jet energy distribution overall. The use of greater accumulated data in the future will help to confirm whether this is indeed the case. One possibile explanation is that the larger centre-of-mass energy used in Run II has resulted in more energetic jets.

It is possible that the Run I to Run II discrepancies seen in Figures 8 and 10 are related. Figure 12 shows a positive correlation between the E_T of the highest E_T jet in the event and the jet multiplicity of that event. Thus potentially the larger fraction of high multiplicity events in Run II is due to the harder jet distribution.

Figure 11 further highlights the relationship between the jet energy and jet multiplicity of an event. As the energy of the highest E_T jet in the event increases the fraction of one jet events decreases. The apparent levelling of the curve beyond $\sim$80 GeV is reasonably consistent with the shape of Figure 12. One would hope that there would not be significant disagreement between Run I and Run II in this plot, as the relationship between the energy of the highest E_T jet in an event and its jet multiplicity is something that should be invariant, as long as the same jet definition has been used. With the exception of one bin, both curves agree within statistical uncertainties.

Finally we compare the angular correlations of jets using two variables: the dijet invariant mass (M_{jj}) and dijet angular separation in R (ΔR_{jj}). These are shown in Figures 13 and 14 respectively. The Run I and Run II ΔR_{jj} distributions agree reasonably well, and one would hope they would given that we have used the same jet definition in Run II. However, although the M_{jj} distributions agree well in the low mass region, they are markedly different in the high mass region. The value of M_{jj} depends on the angular separation and energy of the jets. Given the agreement seen in the ΔR_{jj} distributions, one might attribute the high M_{jj} discrepancy to the harder jet E_T distribution observed in Run II.

3.4.2 Comparisons with Monte Carlo Samples

As was explained in section 2.3.3, the "double counting" problem means that ELO W + n parton generated distributions are best compared to inclusive W + $\geq$ n jet data. Figure 15 is the equivalent plot to Figure 9 for data and Monte Carlo (MC) comparisons. Here the W + $\geq$ 1 jet distribution is compared to to the A+H 1 Parton data, the W + $\geq$ 2 jet distribution compared to the A+H 2 Parton data and so on. Again, the level of agreement is best examined in Figure 16. In all cases the theoretical predictions show a deficit in the high and low jet E_T regions. The A+H 2 Parton MC predicts the shape of the W + $\geq$ 2 jet data well in the intermediate 30 to 90 GeV E_T range, whereas the A+H 1 Parton MC does not do so well. One

has to view these plots with care, as the fact that the MC data has been normalised to the Run II data means that the deficits in the theory in the high and low E_T regions will cause an offset from zero in the intermediate regions. However, this is not sufficient to explain the level of disagreement in the intermediate region for the A+H 1 Parton MC. It is clearly more than an offset and in fact somewhat worse than was observed in Run I [11].

The deficit seen in the theoretical predictions at high jet E_T can be explained in terms of the need for true higher order corrections in these regions, and this is well highlighted by Figure 17. This shows the equivalent plot to Figure 11, but with the predictions of the various Monte Carlo generated samples compared to Run II. The PYTHIA and A+H 1 Parton MC samples predict the one jet event fraction reasonably well at low E_T, but display a deficiency in > 1 jet events at high E_T. The Run II events in this region are predominantly multijet, as demonstrated in Figure 12, but these Monte Carlos do not contain the full set of higher order diagrams necessary to predict the correct multijet rate. It is interesting to note that the A+H 1 Parton predictions seem to fail at around 100 GeV in both the jet E_T and one jet fraction distributions. As one would expect the A+H 2 and 3 Parton MC samples do not well describe the one jet fraction distribution as they do not evaluate one parton leading order diagrams.

Figures 18, 19, 20 and 21 compare the Run II M_{jj} and ΔR_{jj} distributions to those produced by the various Monte Carlo samples for W + $\geq$ 2 jet and W + $\geq$ 3 jet events. One would expect the W + $\geq$ 2 jet distribution to be reproduced best by the A+H 2 Parton MC sample, and similarly the W + $\geq$ 3 jet distribution to be reproduced best by the A+H 3 Parton MC sample. In the M_{jj} distributions it is hard to ascertain whether this is the case, although the A+H 2 Parton sample certainly predicts the shape of the W + $\geq$ 2 jet data better than the A+H 1 Parton sample. Interestingly, the A+H 3 Parton sample reproduces the W + $\geq$ 2 jet distribution with arguably the same precision as the A+H 2 Parton sample. This is surprising given that the 3 parton sample does not evaluate any two parton leading order diagrams.

In the ΔR_{jj} distributions the level of agreement between data and Monte Carlo can be seen more clearly. The A+H 2 Parton MC sample reproduces the W + $\geq$ 2 jet distribution particuarly well, markedly better than the 1 parton sample. This distribution displays the difference between the A+H 2 Parton and A+H 3 Parton predictions much more clearly than the equivalent M_{jj} distribution, but although the 2 parton sample's reproduction of the data is better, the 3 parton sample again does surprisingly well. The situation is not so clear cut in the W + $\geq$ 3 jet distributions. It is arguable whether the A+H 3 Parton sample predicts the distribution any better than the 2 or even 1 parton samples.

3.5 Preliminary Conclusions

- There is a deficiency in the number of events observed in Run II when compared to scaled Run I results, despite a close replication of the Run I selection procedure. The reasons for this are as yet unclear.
- The jet energy distribution in Run II appears to be harder when compared to Run I,

potentially due to the increase in $\sqrt{s}$.

- The shape of the jet multiplicity distribution seems to have changed from Run I to Run II, with Run II displaying a greater proportion of high jet multiplicity events. It is possible that this is connnected to the harder jet energy distribution, and hence in turn to the increased centre-of-mass energy.
- The Run I and Run II ΔR_{jj} distributions agree well, which is evidence of a good replication of the Run I jet definition. Disagreements in the M_{jj} distributions can be attributed to the difference in jet energy spectrum.
- The same limitations of ELO QCD predictions are evident in the Run II results as were found in the Run I analysis. Figures 15 and 17 highlight the need for true higher order corrections that are not provided by HERWIG.
- ELO Limitations are not so evident in the ΔR_{jj} and M_{jj} Monte Carlo distributions, indicating that these angular variables are not so sensitive to higher order corrections.
- The ΔR_{jj} and M_{jj} Monte Carlo distributions show that W + $\geq$ n jet data is best reproduced at ELO by W + n parton Monte Carlo data, although some curious results have been observed.

4 Future Plan

Short Term:

- Attempt to explain the discrepancy between Run I and Run II event numbers by comparing efficiency estimates.
- Invesitigate the hypothesis that the harder jet energy spectrum of Run II is due to the increase in centre-of-mass energy via examination of Monte Carlo samples of varying sqrts.
- Fully correct the sample to account for contamination from background processes.

Long Term:

- Compare data to NLO QCD predictions using Monte Carlo generators such as MCFM. Will these reduce the sensitivity to higher order corrections?
- Investigate in detail the effect of varying renormalisation/factorisation and fragmentation scales within these Monte Carlos.

- Examine other event shapes which could also display the limitaions of ELO and the need for higher order corrections, such as thrust.
- Move to a more theoretically sound jet algorithm such as the K_T clustering or Midpoint algorithms. These could make a considerable difference to NLO comparisons.
- The K_T algorithm introduces new challenges relating to the difficulty in applying current jet correction procedures to K_T defined jets. These will have to be addressed if the algorithm is to be implemented successfully.
- Do the full inclusive cross-section measurement for the W + Jets process using the increased statistics of Run II.

Sample Name	Generator	Num. Partons	PDF	$\sqrt{s}$ (TeV)	Q^2_{REN}	Num. Events
PYTHIA	PYTHIA 6.203	0	CTEQ5L	1.96	M^2_W	64000
A+H 1 Parton	ALPGEN 1.1 + HERWIG 6.4	1	CTEQ5L	1.96	$M^2_W + \sum P^2_T$	45000
A+H 2 Parton	ALPGEN 1.1 + HERWIG 6.4	2	CTEQ5L	1.96	$M^2_W + \sum P^2_T$	40000
A+H 3 Parton	ALPGEN 1.1 + HERWIG 6.4	3	CTEQ5L	1.96	$M^2_W + \sum P^2_T$	32500

Table 4: Table detailing the Monte Carlo Samples used in this study. Q^2_{REN} refers to the renormalisation scale. By default the factorisation scale is always set equal to the renormalisation scale.

Correction	Affect on Jet Energy	Description
Relative Energy	Increase	Corrects for the η variation in hadronic energy scale.
Time Dependance	Increase	Accounts for the degrading of calorimeter components over time.
Raw Energy Scale	Increase	Attempts to correct back to the Run I raw energy scale.
Multiple Interactions	Decrease	Account for energy from additonal $p\overline{p}$ interactions.
Absolute Energy	Increase	Uniformly implements the hadronic energy scale of the calorimeter.
Underlying Energy	Decrease	Account for energy produced by the underlying event.
Out of Cone	Increase	Account for energy radiated outside the cone.

Table 5: Table briefly explaining the various jet energy corrections used in this study.

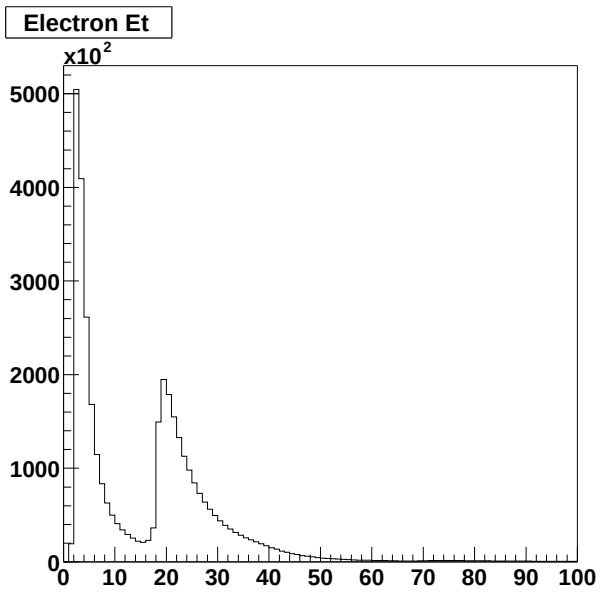

Figure 5: Et distribution for every electron in the ELECTRON_CENTRAL_18 sample

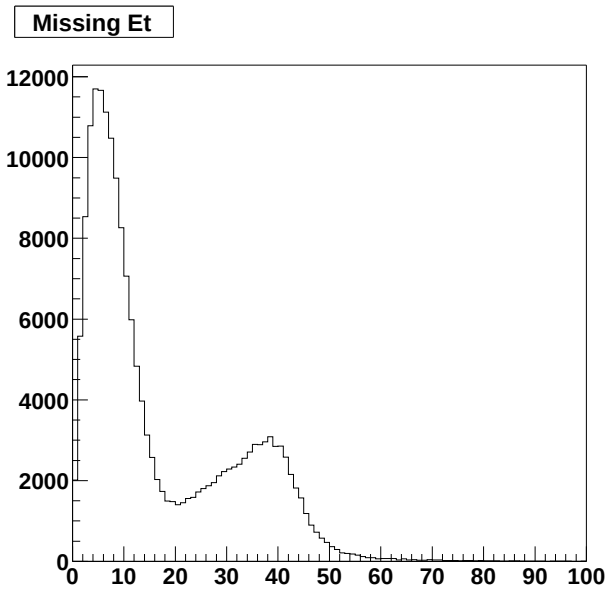

Figure 6: $\not\!\!E_{\mathrm{T}}$ distribution for events which contain one or more tight electrons

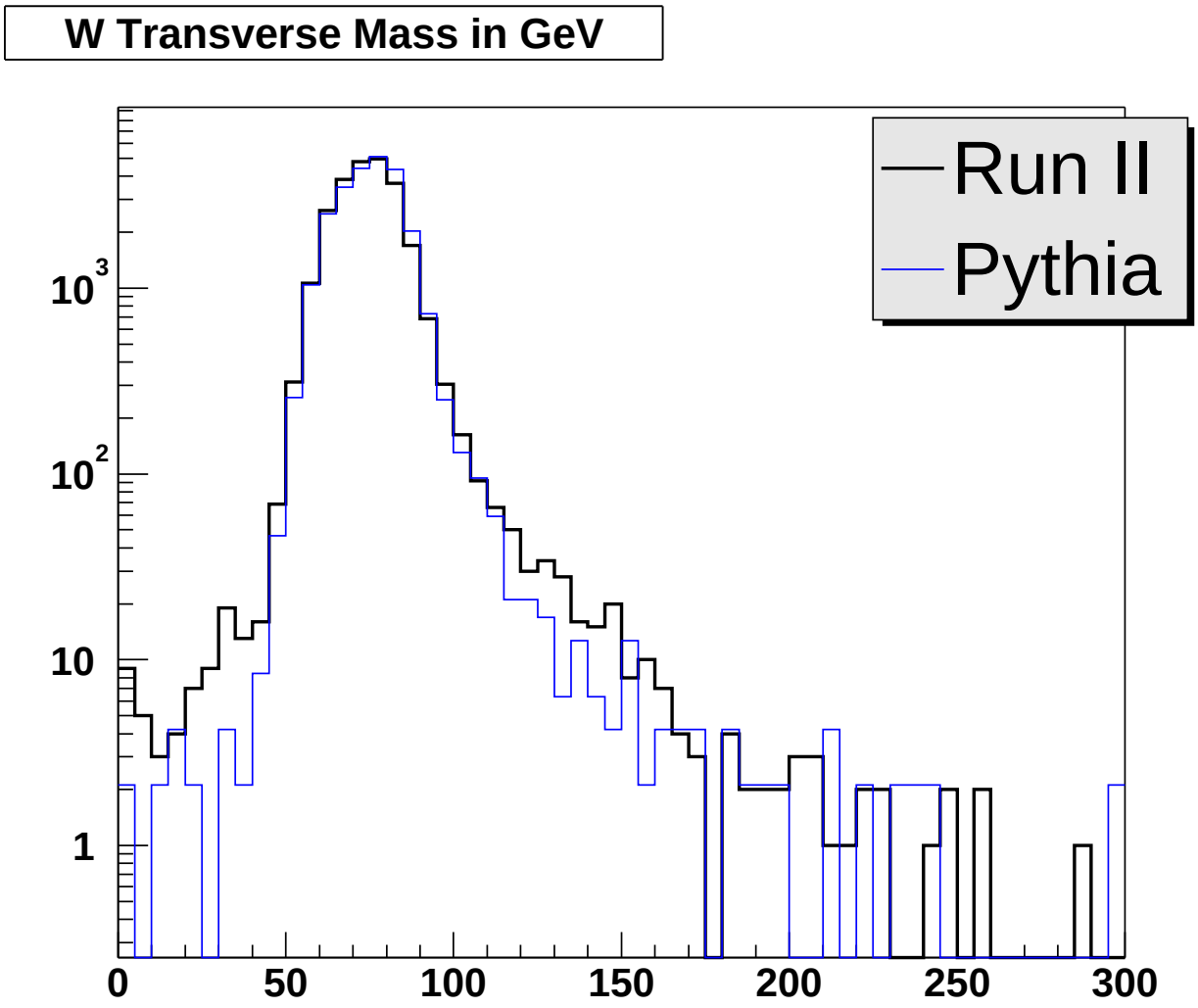

Figure 7: Plot of reconstructed W Transverse Mass illustrating the purity of the W sample. PYTHIA generated Monte Carlo data is also shown for comparison.

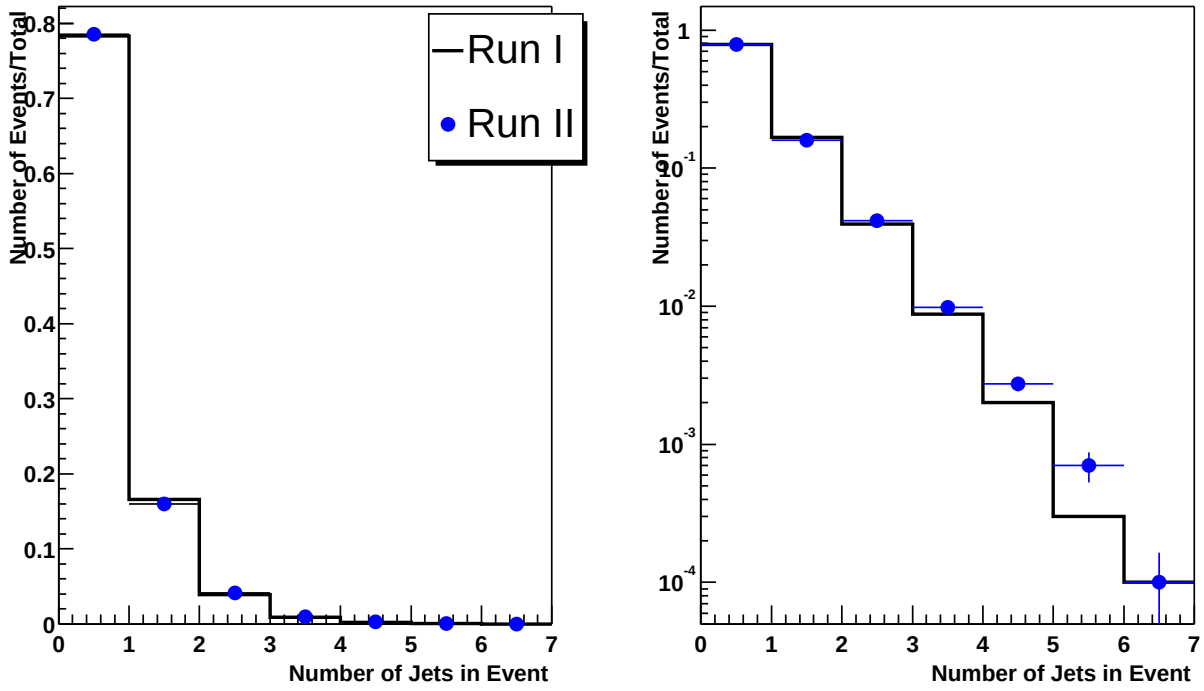

Figure 8: Comparison of Jet Multiplicity Distribution between Run I and Run II, normalised to the total number of events. The plot on the right-hand side differs only in the logarithmic scale. Errors shown are statistical only.

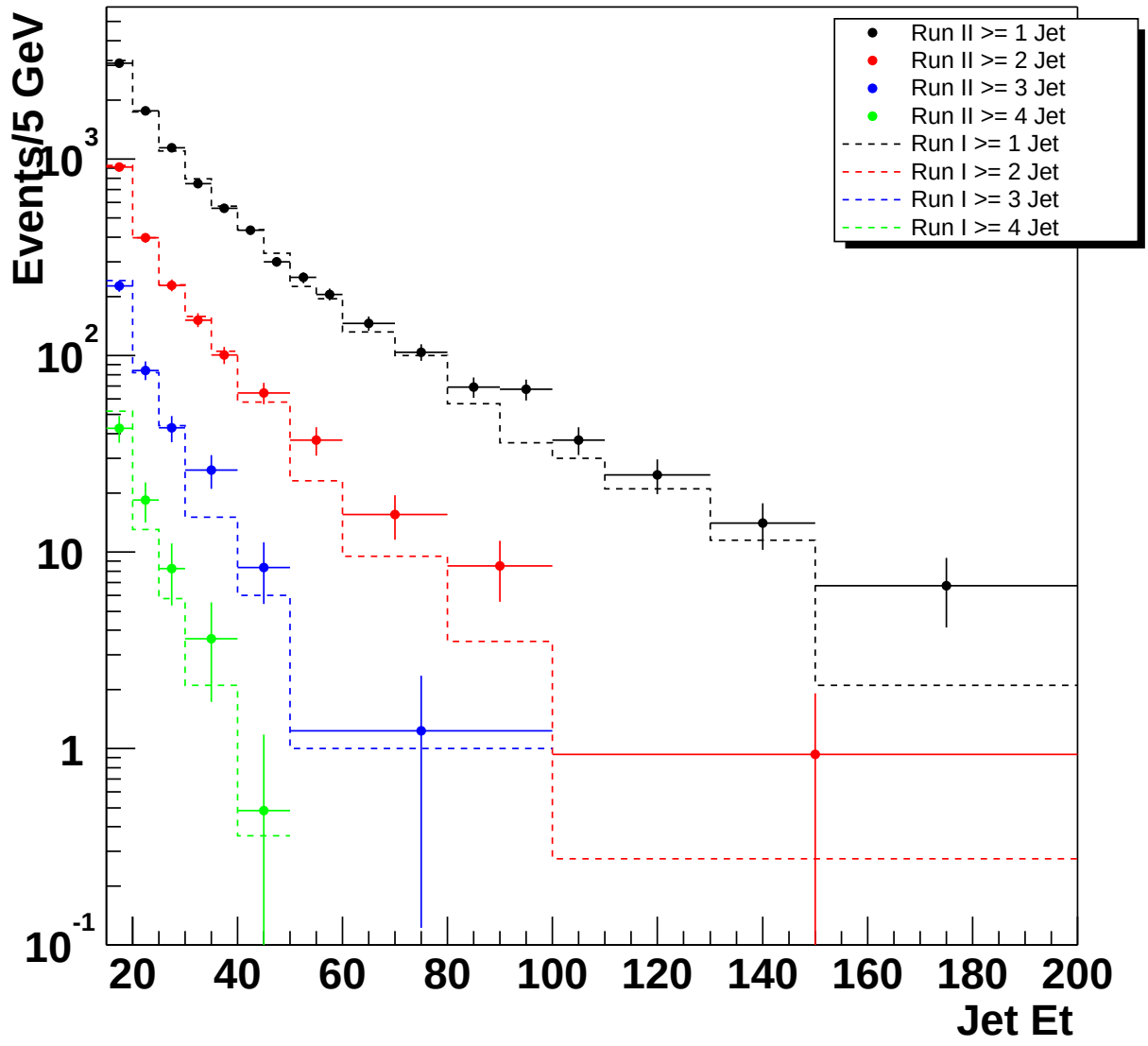

Figure 9: Comparison of the Jet E_T distribution between Run I and Run II for various inclusive jet multiplicities. Shown is the highest E_T jet in W + $\geq$ 1 jet events, the second highest E_T jet in W + $\geq$ 2 jet events, the third highest E_T jet in W + $\geq$ 3 jet events and the fourth highest E_T jet in W + $\geq$ 4 jet events. Run II is scaled to Run I. Errors shown are statistical only.

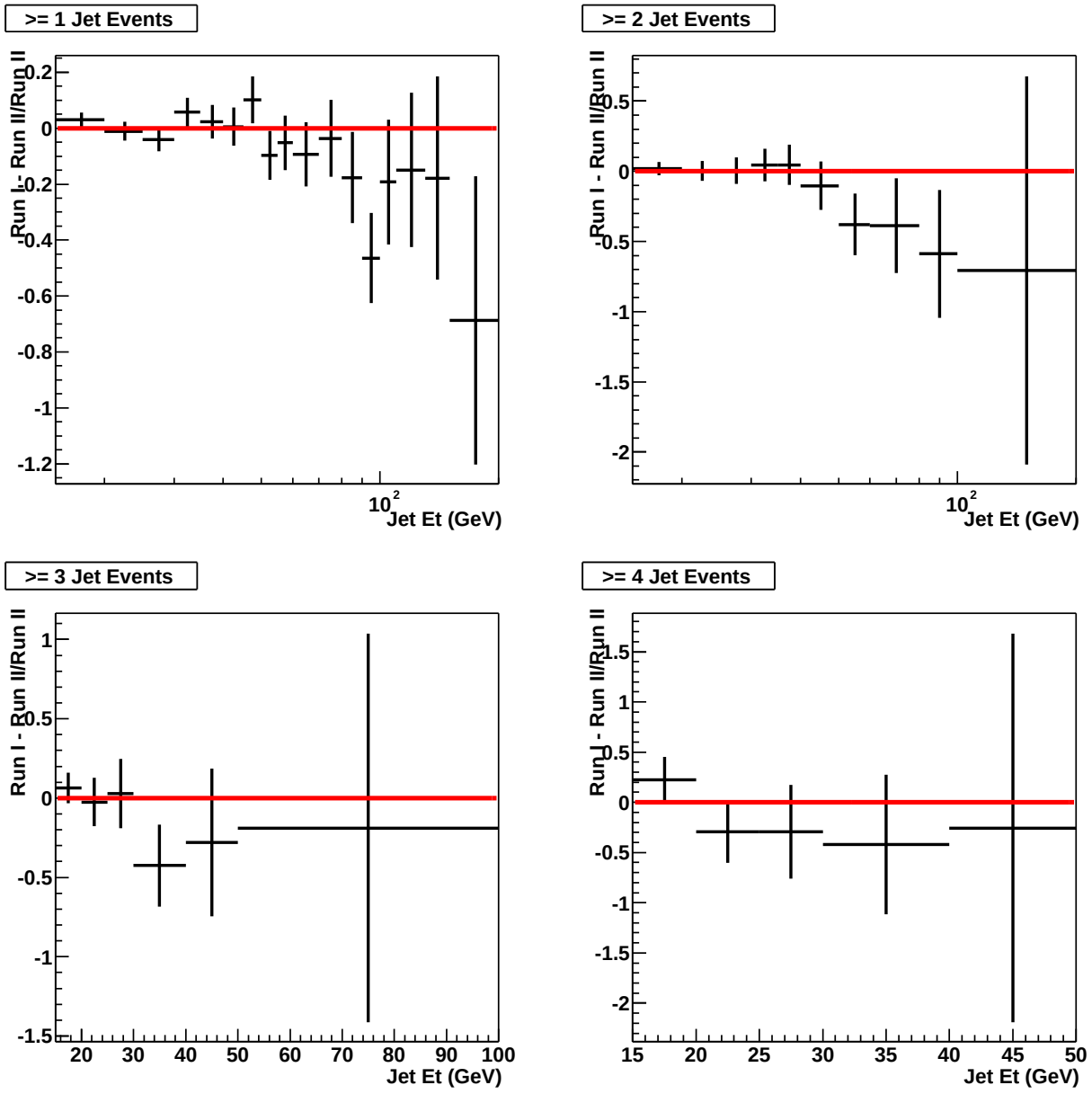

Figure 10: These graphs highlight the difference in the Jet E_T distribution between Run I and Run II, and relate directly to the four curves of Figure 9. Note the negative values at high jet E_T, indicating that the Run II jets are somewhat harder. Errors shown are statistical only.

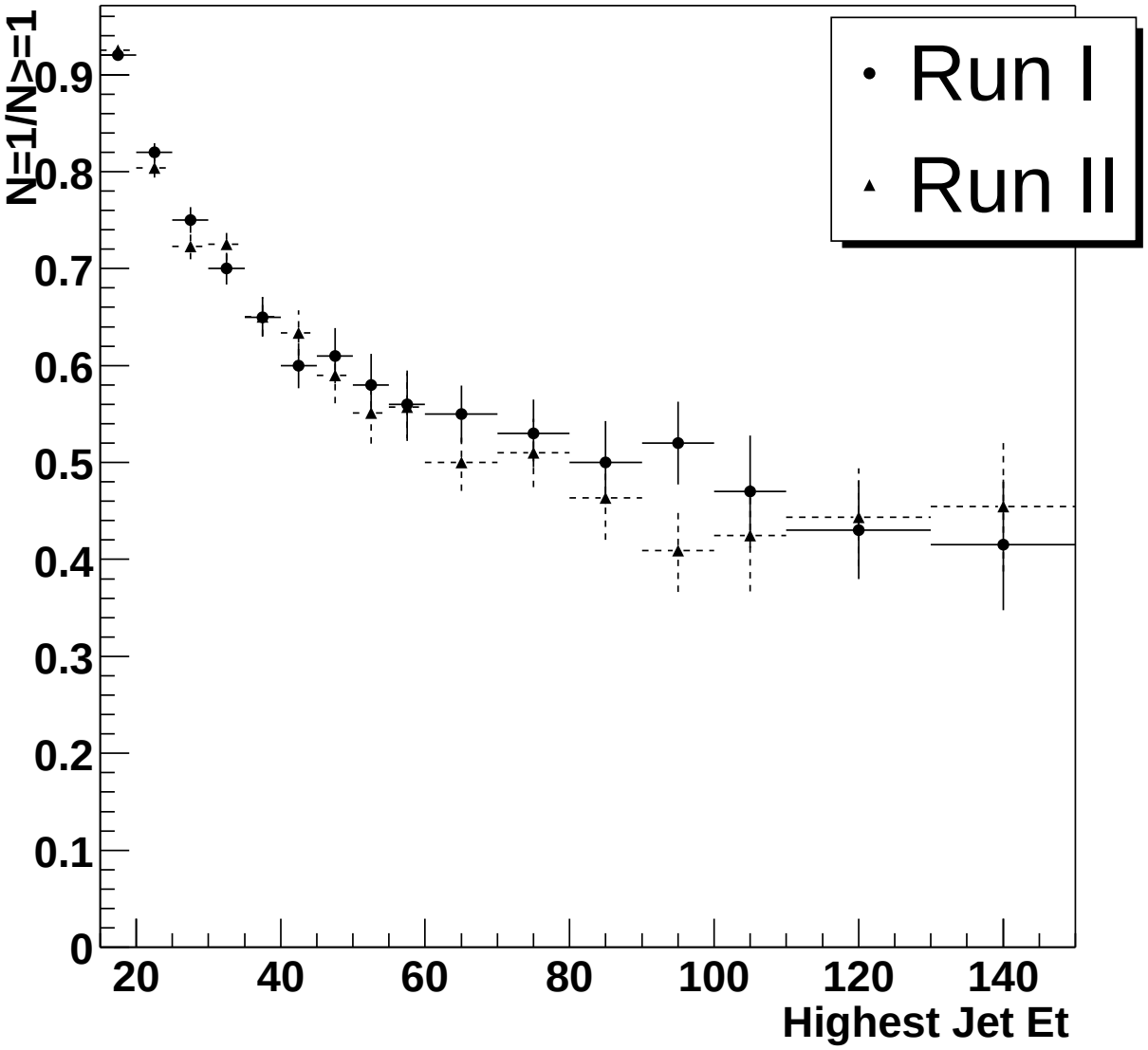

Figure 11: This plot shows the fraction of 1 jet events in ≥ 1 jet events as a function of the E_T of the highest E_T jet in the event. Both Run I and Run II data is displayed for comparison. Errors shown are statistical only.

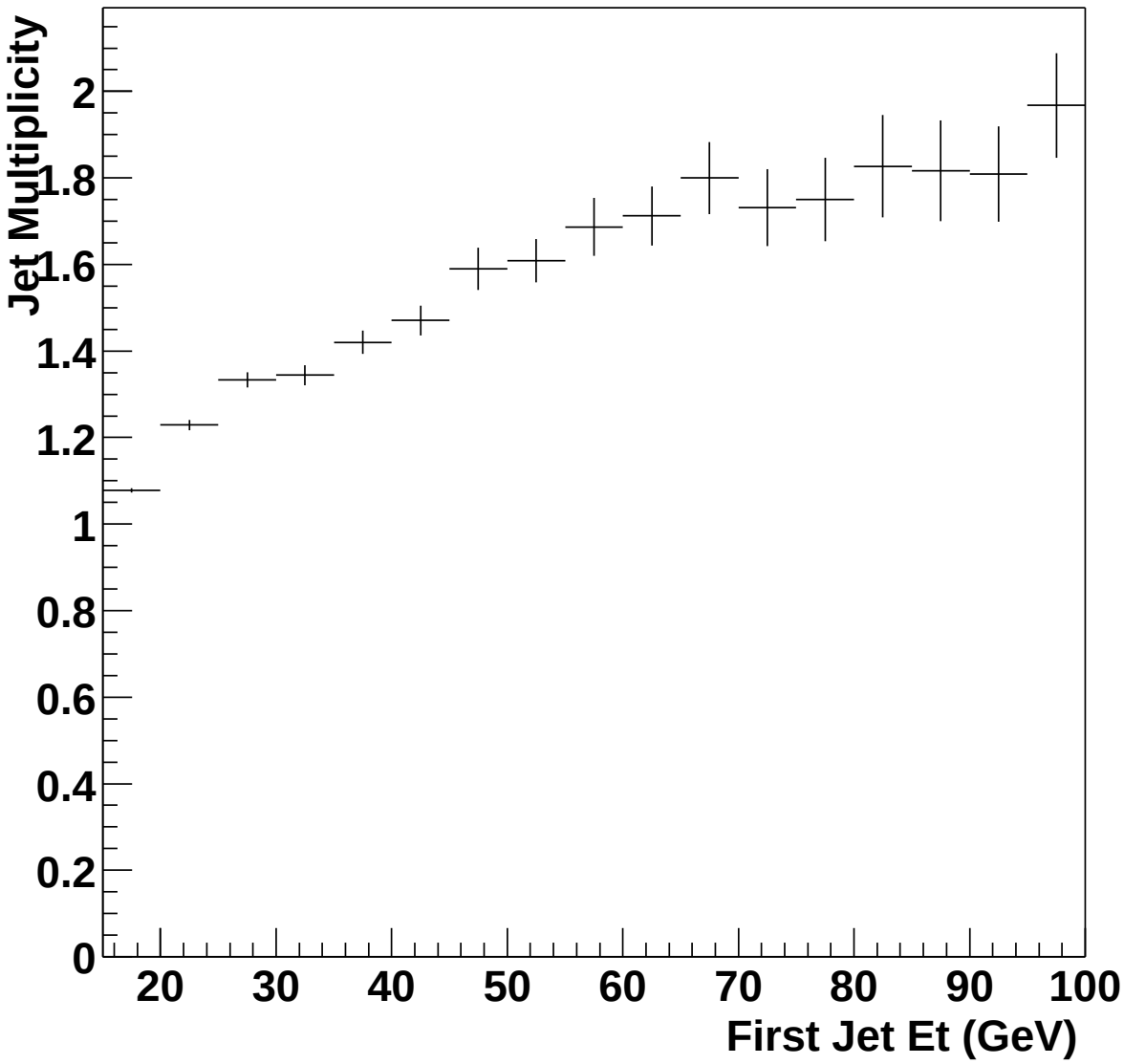

Figure 12: This is a profile plot showing the average jet multiplicity of events versus the E_T of the highest E_T jet in the event. One can clearly see a direct correlation between the two.

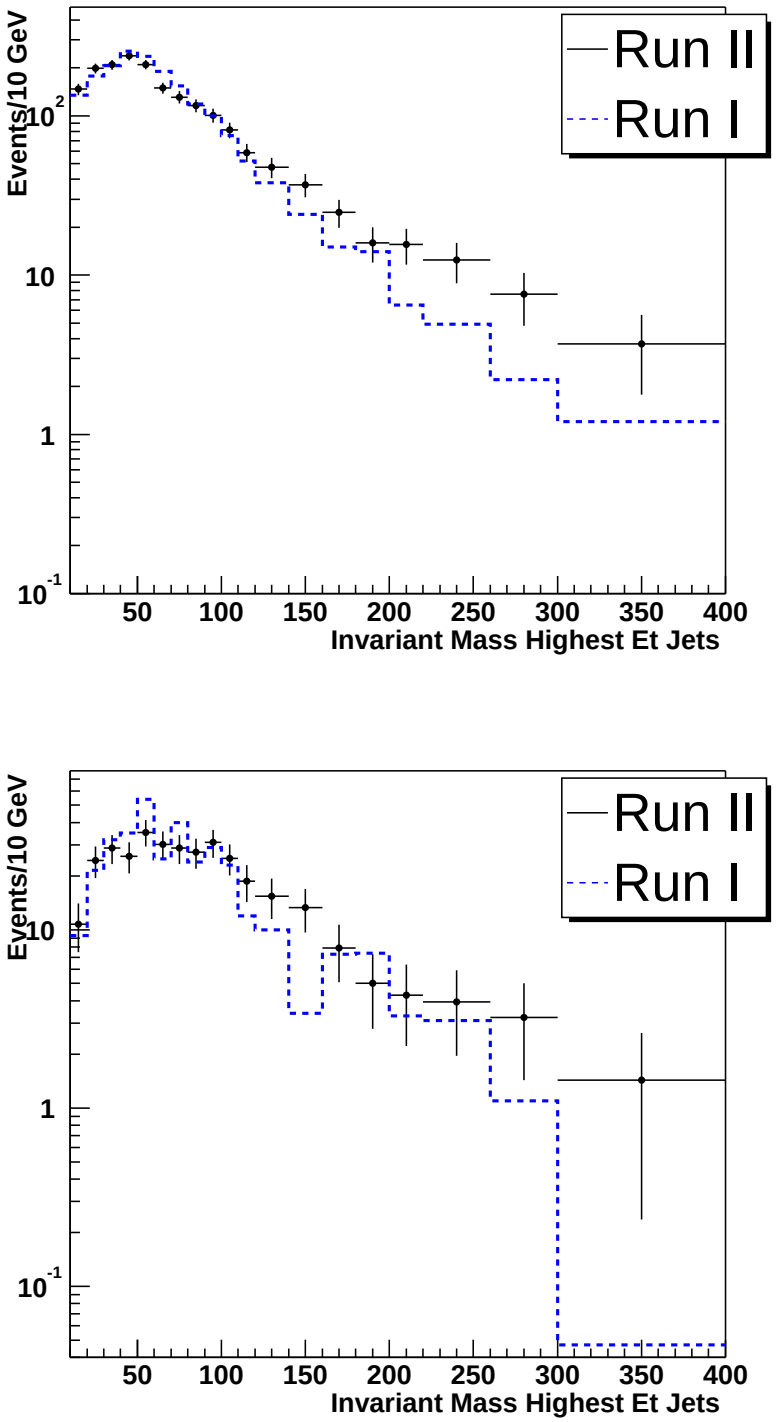

Figure 13: Plots showing M_{jj} for the two highest E_T jets in W + $\geq$ 2 jet events (top) and W + $\geq$ 3 jet events (bottom) for both Run I and Run II data. Run II scaled to Run I.

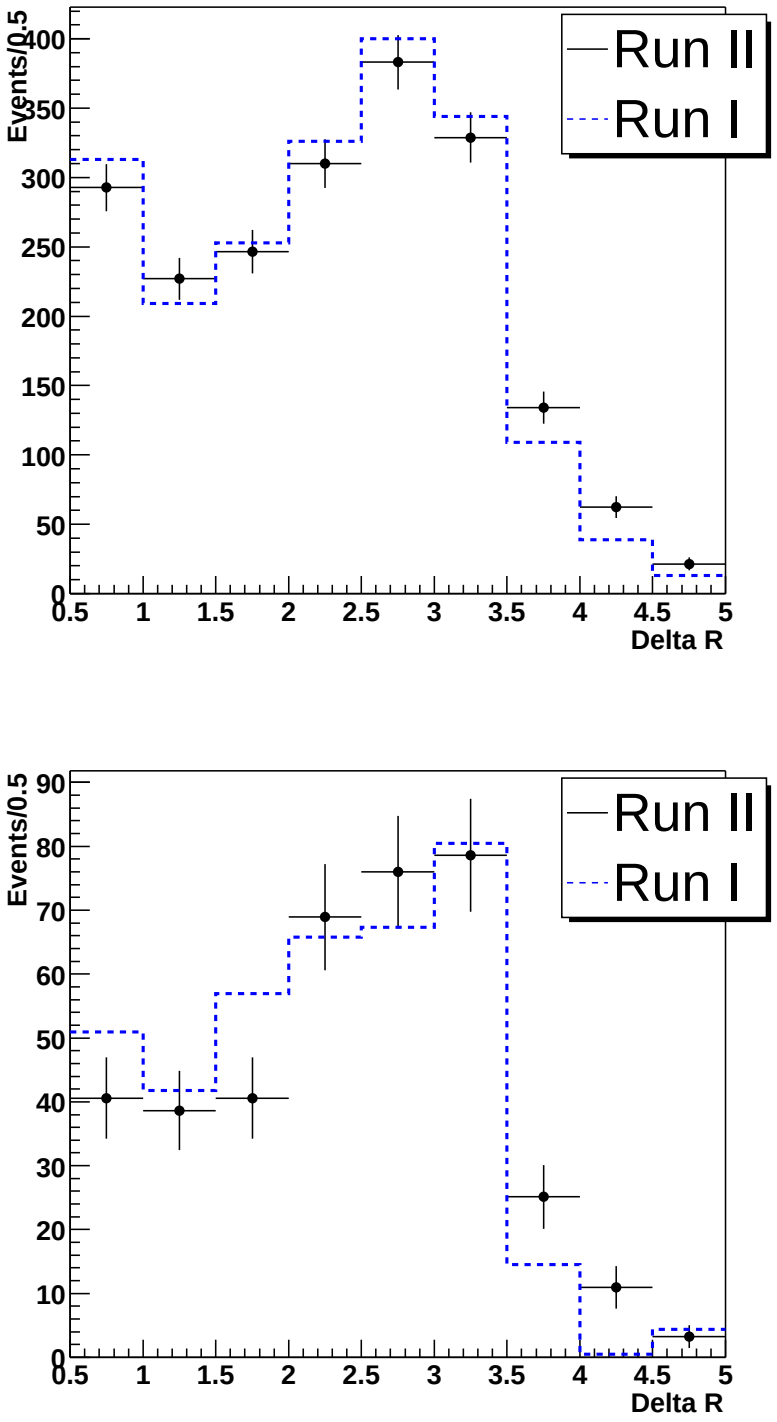

Figure 14: Plots showing ΔR_{jj} for the two highest E_T jets in W + $\geq$ 2 jet events (top) and W + $\geq$ 3 jet events (bottom) for both Run I and Run II data. Run II scaled to Run I.

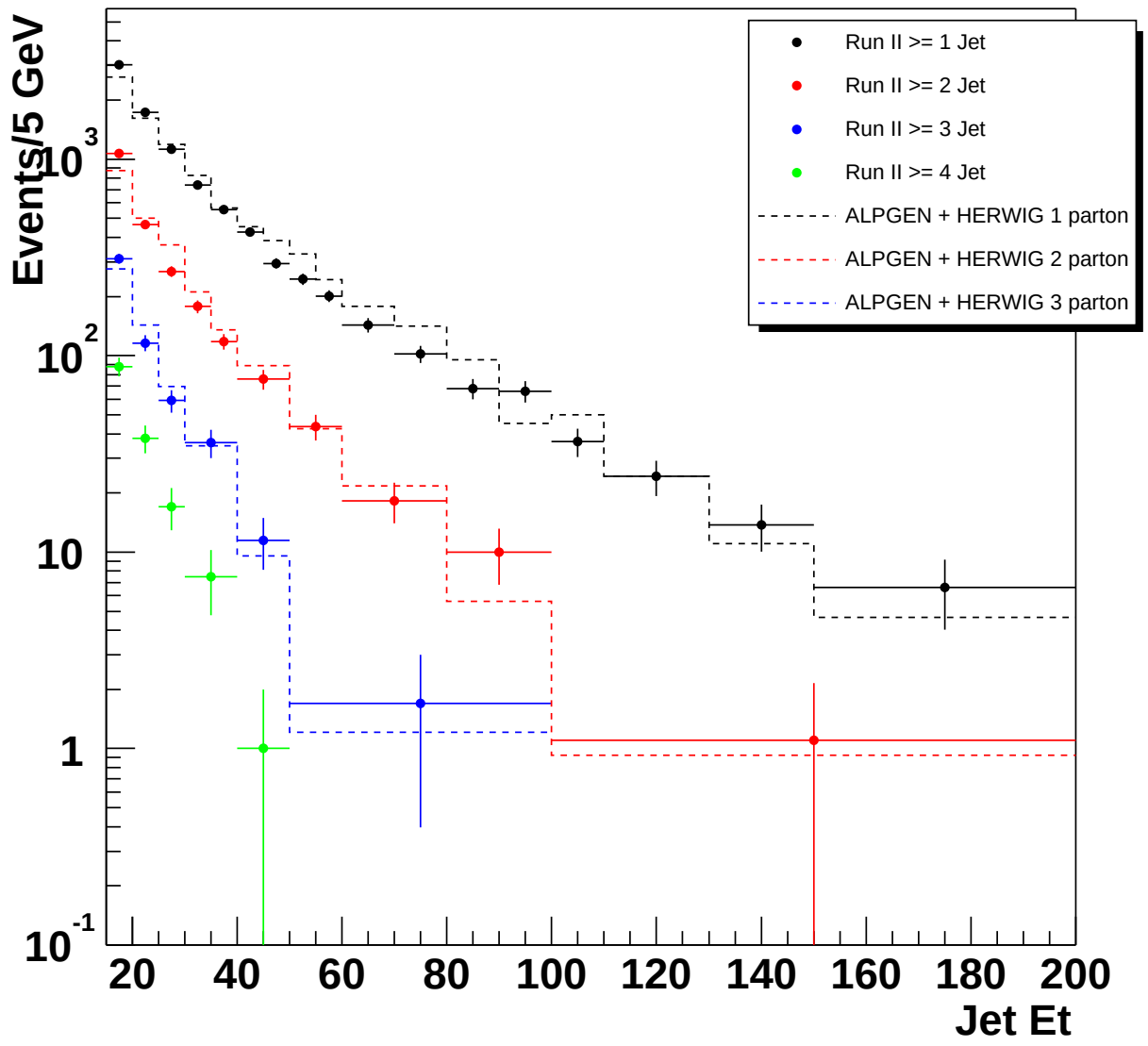

Figure 15: Comparison of the Jet E_T distribution between Run II and ALPGEN + HERWIG Monte Carlo samples for various inclusive jet multiplicities. Shown is the highest E_T jet in W + $\geq$ 1 jet events, the second highest E_T jet in W + $\geq$ 2 jet events, the third highest E_T jet in W + $\geq$ 3 jet events and the fourth highest E_T jet in W + $\geq$ 4 jet events. Monte Carlo is scaled to Run II. Errors shown are statistical only. Monte Carlo is scaled to data.

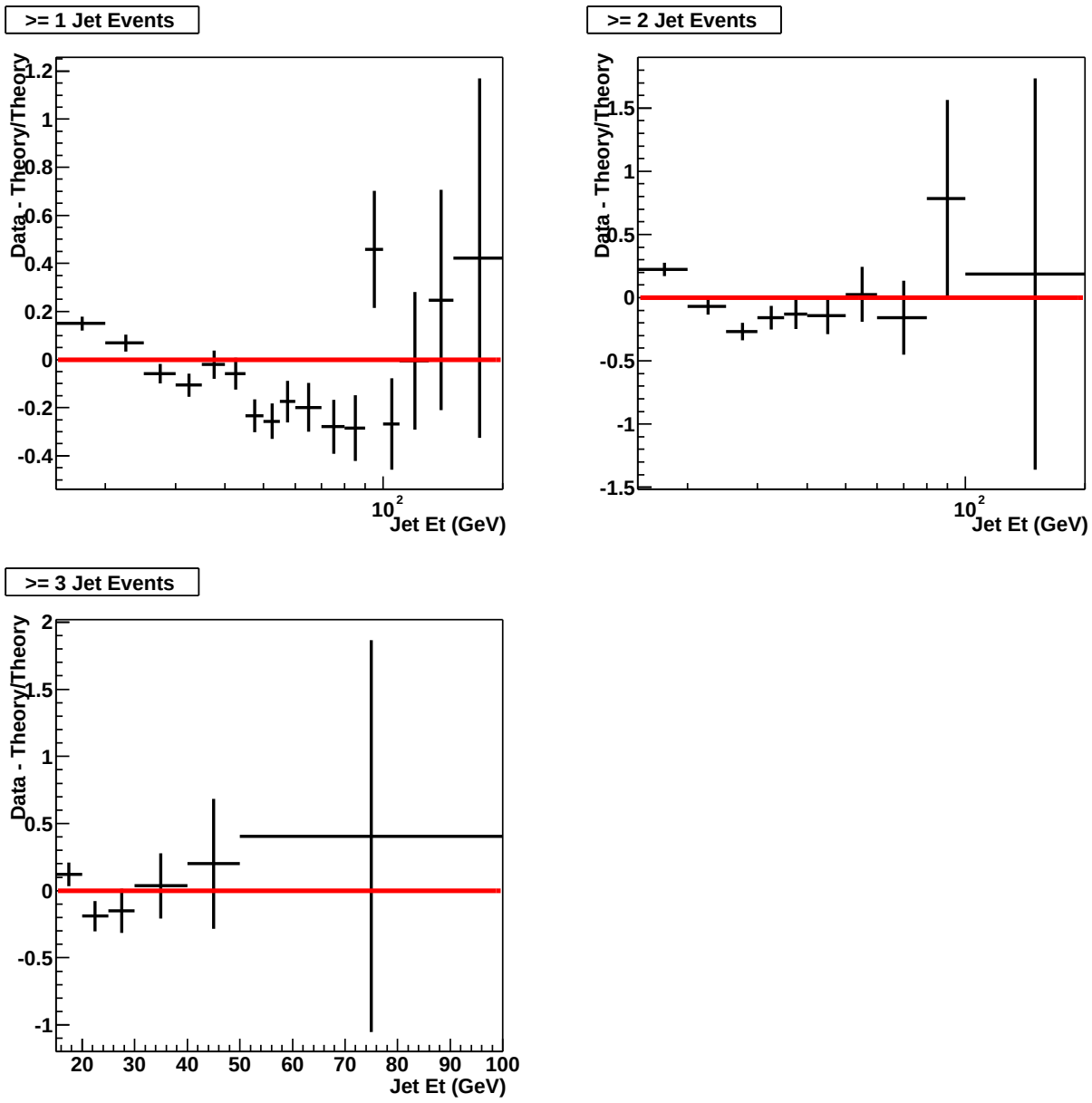

Figure 16: These graphs highlight the difference in the Jet E_T distributions between Run II and the various ALPGEN + HERWIG Monte carlo samples, and relate directly to the three curves of Figure 15 in which they are compared. Note the disagreement at high E_T which was also seen in Run I. Errors shown are statistical only.

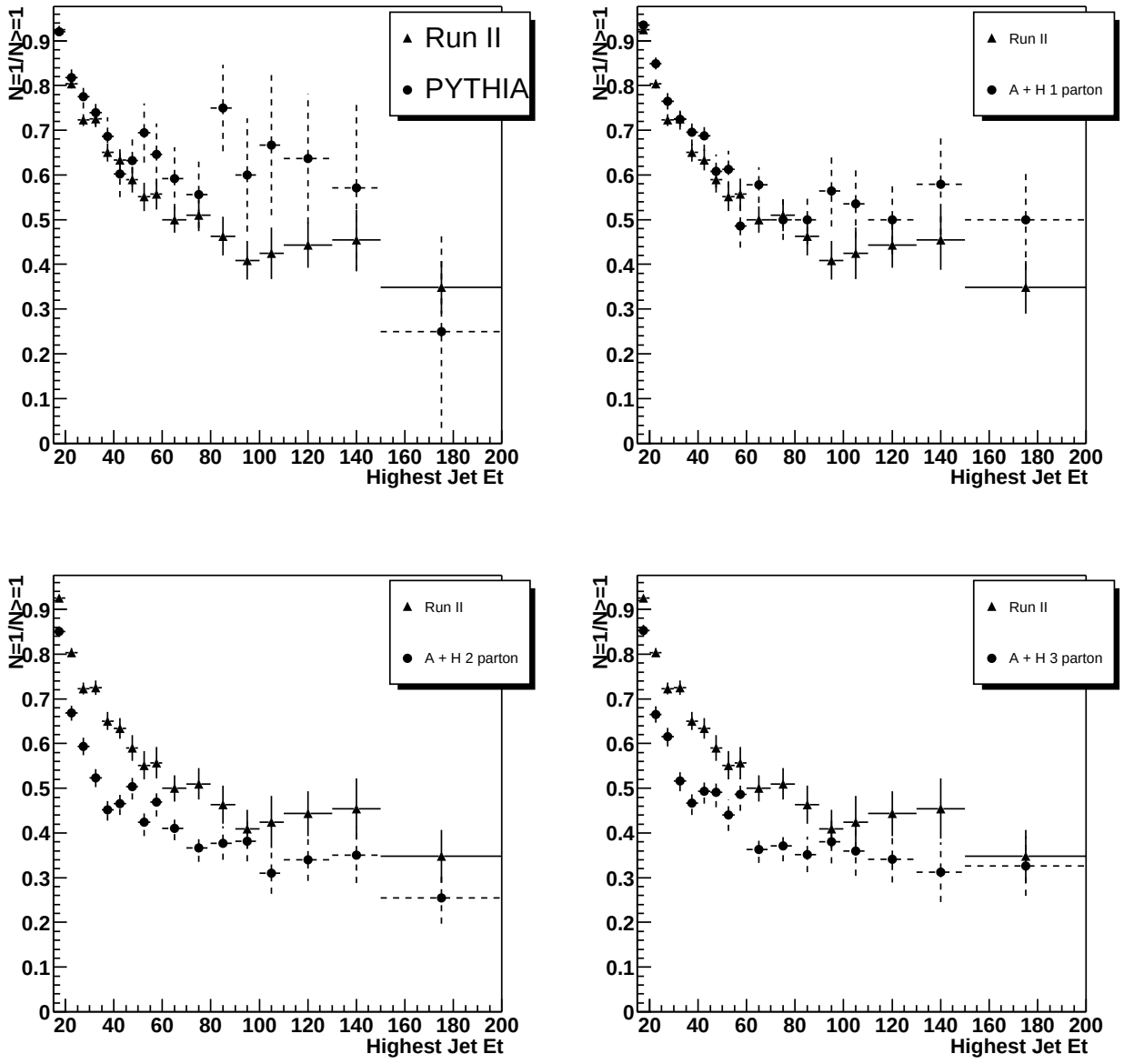

Figure 17: This plot shows the fraction of 1 jet events in $\geq$ 1 jet events as a function of the E_T of the highest E_T jet in the event. The four different Monte Carlo generated samples are compared to Run II data seperately. Errors shown are statistical only.

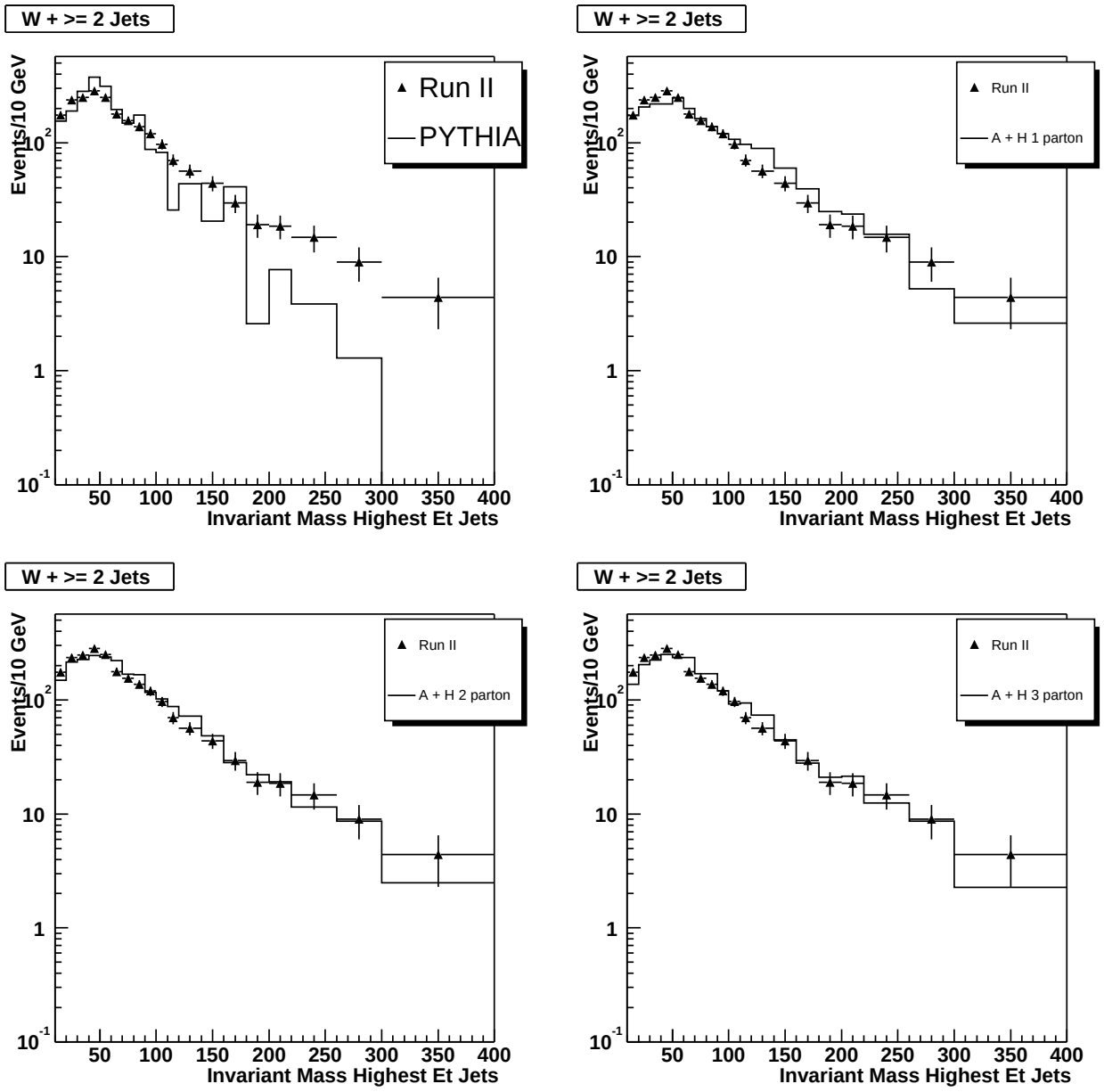

Figure 18: Plots showing M_{jj} for the two highest E_T jets in W + $\geq$ 2 jet events. The various Monte Carlo data samples are compared to Run II data in both jet multiplicity ranges. Monte Carlo is scaled to data.

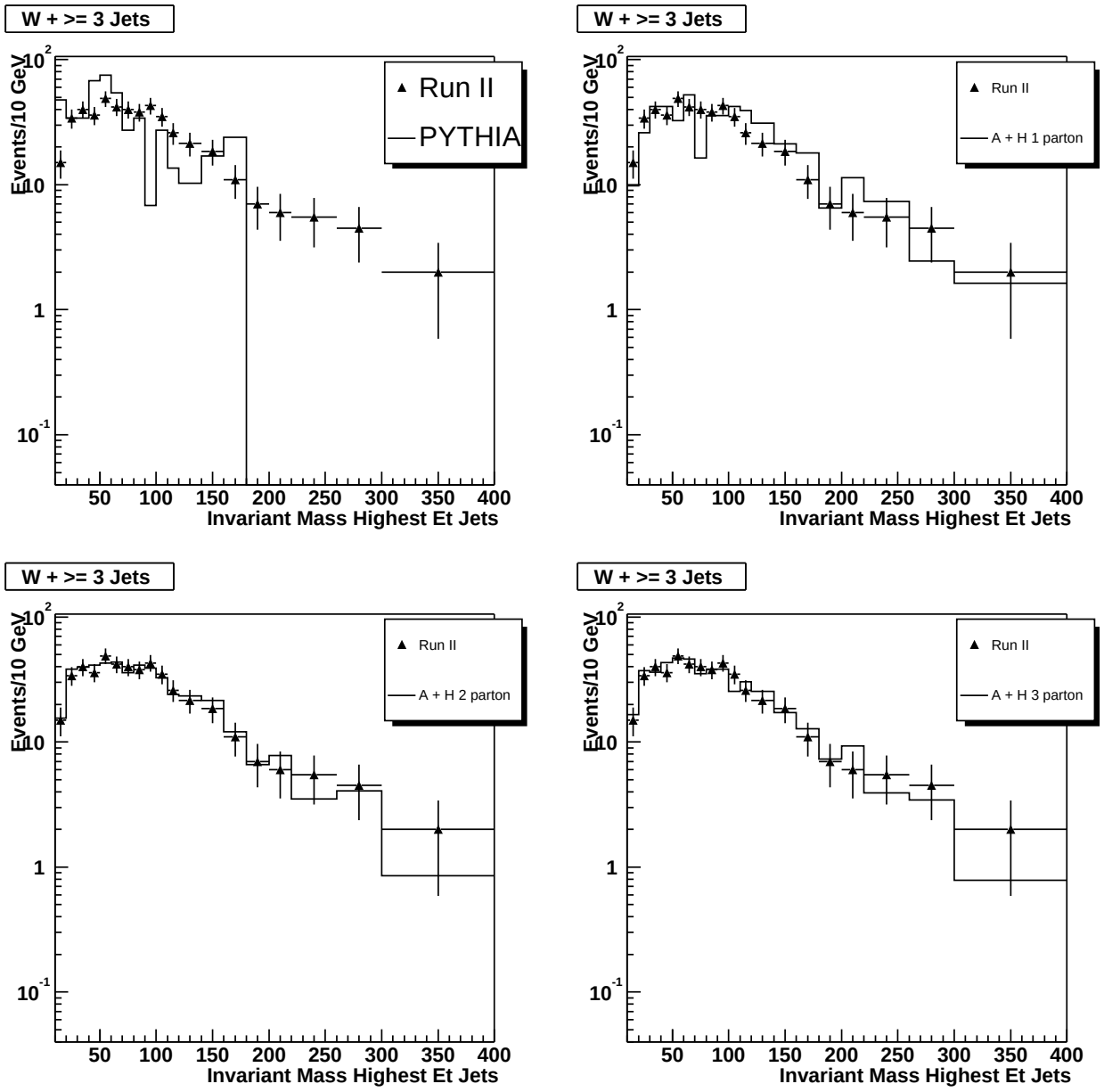

Figure 19: Plots showing M_{jj} for the two highest E_T jets in W + $\geq$ 3 jet events. The various Monte Carlo data samples are compared to Run II data in both jet multiplicity ranges. Monte Carlo is scaled to data.

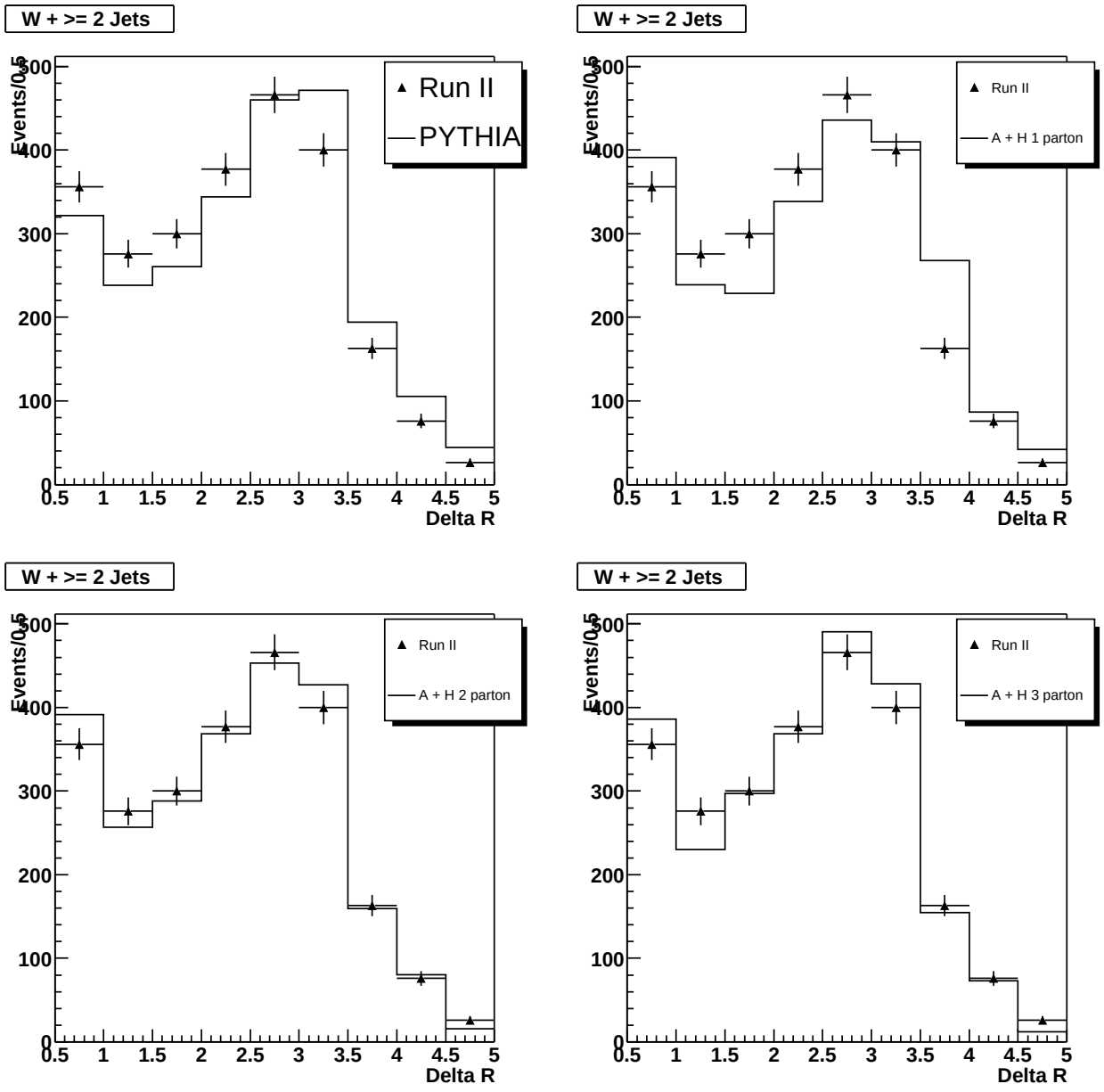

Figure 20: Plots showing ΔR_{jj} for the two highest E_T jets in W + $\geq$ 2 jet events. The various Monte Carlo data samples are compared to Run II data in both jet multiplicity ranges. Monte Carlo is scaled to data.

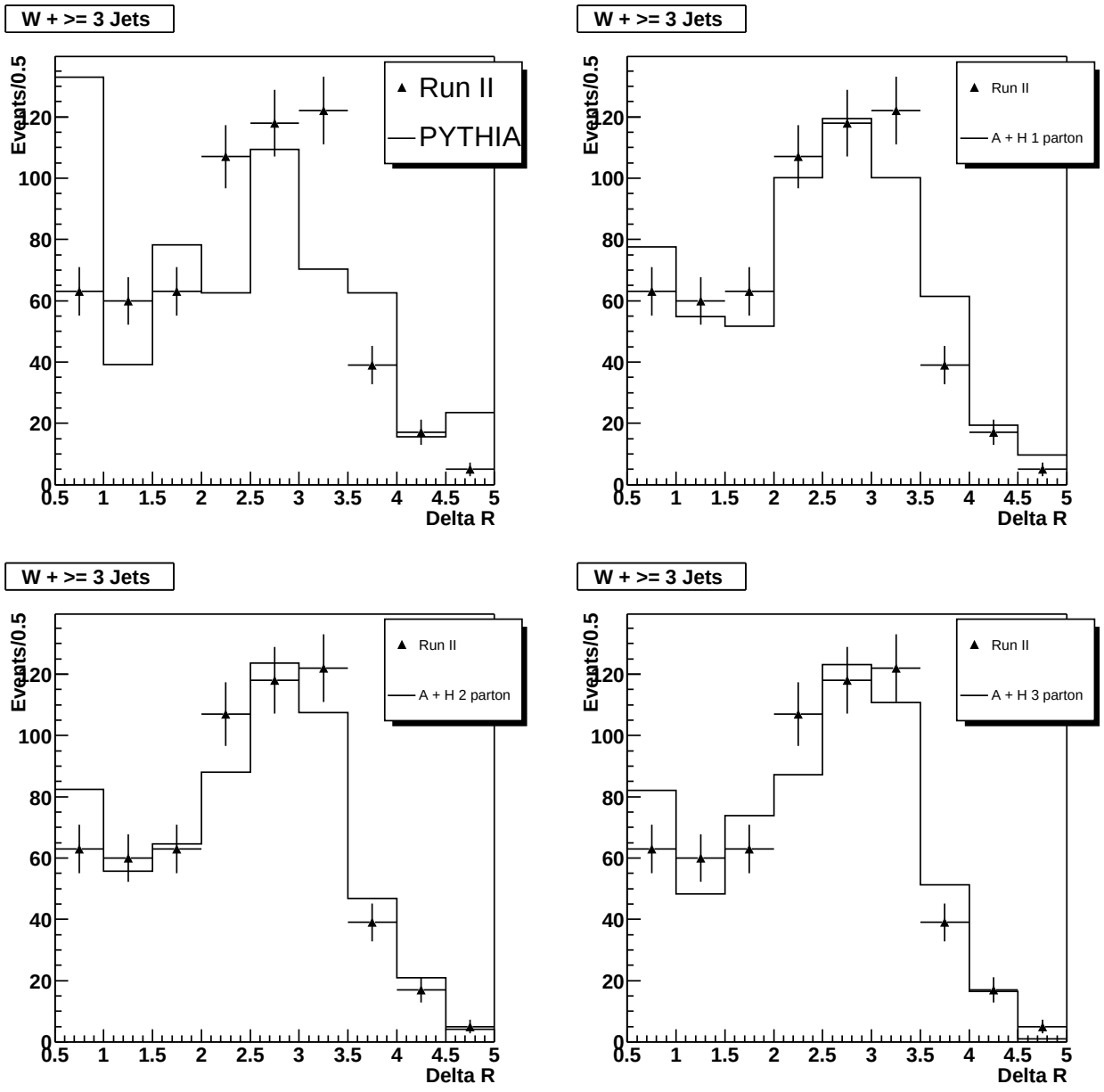

Figure 21: Plots showing ΔR_{jj} for the two highest E_T jets in W + $\geq$ 3 jet events. The various Monte Carlo data samples are compared to Run II data in both jet multiplicity ranges. Monte Carlo is scaled to data.

References

[1] F. Abe *et al.* [CDF Collaboration], "Observation of Top Quark Production in $p\overline{p}$ Collisions with the CDF Detector at Fermilab ", Phys. Rev. Lett. A **74** (1995) 2626; 2631.

[2] [CDF Collaboration], "The CDF II Detector Technical Design Report", FERMILAB-Pub-96/390-E (1996)

[3] D. McGinnis *et al.* [Fermilab Beams Division], "Run II Handbook", http://www-bd.fnal.gov/runII

[4] K. Schindl *et al.*, "DOE Review of Tevatron Luminosity Performance", http://ps-div.web.cern.ch/ps-div/Meetings/Slides/2002/20021111/FNALDOE1002.doc (2002)

[5] M. Quiros, "Bounds on the Higgs Mass in the Standard Model and Minimal Supersymmetric Standard Model", hep-ph/9411403 (1994)

[6] W. Yao, "Standard Model Higgs and Top Mass Measurements at the Tevatron", http://www.dpf99.library.ucla.edu/session1/yao0115.pdf(1999)

[7] B. Winer, "Report on Higgs Sensitivity", CDF Collaboration Meeting May 2003 http://www-cdf.fnal.gov/internal/WebTalks/0305/030523_collab.html (2003)

[8] M. Coca *et al.*, "W $\rightarrow$ $e\nu$ Cross Section Analysis with Run II Data", CDF/DOC/ELECTROWEAK/CDFR/6300 (2003)

[9] D. Amidei *et al.* [CDF Lepton + Jets Working Group], "Measurement of the $p\overline{p} \rightarrow t\overline{t}$ cross section in the l + jets SECVTX tagged sample", CDF/ANAL/TOP/CDFR/6329 (2003)

[10] D. Cronin-Hennessy, "Tests of Perturbative QCD in W + Jets Events Produced in $\sqrt{s}$=1.8 TeV $p\overline{p}$ Collisions", PhD Thesis, Duke University http://www.phy.duke.edu/research/hep/pubs/Hennessy_thesis.ps

[11] T. Affolder *et al.* [CDF Collaboration], "Tests of Enhanced Leading Order QCD in W Boson Plus Jets Events from $\sqrt{s}$=1.8 TeV $p\overline{p}$ Collisions", Physical Review D, **63**, 072003 (2001)

[12] D. Cronin-Hennessy *et al.*, "Measurement of $W^{\pm} \rightarrow e^{\pm}\upsilon$ + n Jet Cross Sections in 1.8 TeV $p\overline{p}$ Collisions", CDF/ANAL/JET/CDFR/4093 (1997)

[13] D. Cronin-Hennessy *et al.*, "Properties of Jets in W Boson Events from 1.8 TeV $p\overline{p}$ Collisions", CDF/ANAL/ELECTROWEAK/CDFR/3095 (1996)

[14] M.L. Mangano *et al.*, "Alpgen, A Generator for Hard Multiparton Processes in Hadronic Collisions", hep-ph/0206293 (2002)

[15] G. Corcella *et al.*, "Herwig 6.5: An Event Generator for Hadeon Emission Reactions with Interfering Gluons (including supersymmetric processes)", hep-ph/0011363 (2000)

[16] G.C. Blazey *et al.*, "Run II Jet Physics", CDF/PUB/JET/PUBLIC/5293 (2000)

[17] M. Seymour, "Jets in QCD", hep-ph/9506421 (1995)

[18] S. Catani, "Longitudinally-Invariant $k_\perp$-Clustering Algorithms for Hadron-Hadron Collisions", CERN-TH.6675/93 (1993)

[19] J-F. Arguin *et al.*, "Generic Jet Energy Corrections for Run II", CDF/ANAL/JET/CDFR/6280 (2003)

[20] [CDF Collaboration], "4.9.1 Monte Carlo Samples", http://www-cdf.fnal.gov/internal/physics/top/mc