# Method for the Measurement of the $t\bar{t}$ Cross Section in the Lepton Plus Jets Channel Using Event Kinematics

R. Hughes, R. Marginean, E. Thomson, B. Winer
*The Ohio State University*

R. Erbacher, R. Roser
*Fermi National Accelerator Laboratory*

J. Conway
*Rutgers University*

CDF NOTE 6802

Version 1.2

January 27, 2004

**Abstract**

With 50% more data, we have updated our previous measurement of the $p\bar{p} \to t\bar{t}$ cross section in the lepton+jets channel using a kinematic analysis. Systematic errors are unchanged from our previous note, since the dominant systematic, from the data-MC agreement of the jet energy scale, will remain the same until new simulation becomes available. Kinematic quantities are used to distinguish the top signal from the backgrounds in the "lepton+jets" channel. This analysis provides a new independent technique to measure the top cross section as well as an important cross check of the traditional b-tag analysis. This note describes the most basic analysis - a kinematic fit using the $H_T$ variable. A variety of other kinematic and shape variables are also presented for consistency. Subsequent notes will make use of other energy and shape variables as well as make use of more advanced analysis techniques such as neural nets. Using this method, from a sample of data corresponding to 195 pb$^{-1}$ of integrated luminosity, the $t\bar{t}$ cross section is measured to be 4.7 $\pm$ 1.6(stat) $\pm$ 1.8(syst)pb in the W+$\geq$ 3 jet mode. In the W+$\geq$4 jet mode, the cross section is measured to be 7.8 $\pm$ 2.0(stat) $\pm$ 3.0(syst)pb.

# 1 Changes

## 1.1 Changes from v1.1 to v1.2

- Used V4.0 good run list minus run 164844, and with database problem fixed for two runs.
- Clarified some points based on questions we received.

## 1.2 Changes from v1.0 to v1.1

- Updated W+1,2,3 jet comparison plots with full data sample.
- Added cross-checks of fit result without $t\bar{t}$ and letting QCD float.
- Added plots showing fit result as a function of assumed top mass.

## 1.3 Changes since CDF note 6206

- Added 50% more data
- Used V2.0 good run list minus run 164844.
- Used good quality($\geq$ 12) ZVertexModule object closest to tight lepton as event z position instead of tight lepton $z_0$. Added a cut requiring $z_0$ of lepton to be within 5 cm of event z vertex position from ZVertexModule.
- Removed top mass as a systematic, instead quote result for a top mass of 175 GeV.

# 2 Motivation

In Run I, CDF discovered the top quark and began to measure its properties. While a coherent picture of this elusive quark is beginning to take shape, there are still many unanswered questions as well as inconsistencies in our current understanding. The $t\bar{t}$ cross section measurements in the lepton+jets and dilepton channels are quite different. More events are observed in the tagged lepton+jets W+2 jet bin than expected based on our "method 2" background calculations. Furthermore, the width of the mass peak in Run I was narrower than our Monte Carlos indicate we should expect. Therefore, it is important to measure the properties of the top quark in as many different ways as possible in order to develop a complete understanding.

This note describes the method for measuring $\sigma_{t\bar{t}}$ by performing a maximum likelihood fit using the kinematic variable $H_T$ (defined as the scalar sum of all reconstructed lepton, jet, and missing $E_T$ in the event) for events with an isolated lepton ($e$ or $\mu$), missing $E_T$, and three or more hadronic jets. Other kinematic and shape variables are presented here as well to indicate the degree of sensitivity and consistency in this measurement.

This work is the first step of our overall analysis strategy. From this foundation, we will expand this analysis to try to better separate signal from background. This will involve using additional N-jet bins, other kinematic and shape variables as well as b-tagging. Neural network techniques will also be exploited. By making use of as much information as there is available in the lepton+jets $t\bar{t}$ system, we expect to be able to measure the cross section to greater precision than what has been achieved to date.

| Time Period | Run Range | Integrated Luminosity ($pb^{-1}$) | |
|---|---|---|---|
| | | CEM and CMUP | CMX |
| 23 Mar 2002 - 12 Jan 2003 | 141544-156487 | 75.3 | 57.0 |
| 12 Jan 2003 - 27 May 2003 | 156488-163527 | 52.3 | 52.3 |
| 27 May 2003 - 9 Sep 2003 | 163528-168889 | 65.9 | 65.9 |
| Total | 141544-168889 | 193.5 | 175.3 |

Table 1: Integrated luminosity after good run requirement.

# 3 Data Sample

This analysis uses the official Top Group high $p_T$ lepton datasets [1], which are based on the 4.8.4 offline release with calorimeter-dependent reconstruction remade in 4.11.1 in order to pick up improved calibrations. Good quality requirements are enforced by V4.0 of the DQM good run list [2], with run 164844 excluded since part of this run was lost for the high $p_T$ lepton trigger stream (B). This data sample includes both the 126 $pb^{-1}$ used for the summer conferences as well as 69 $pb^{-1}$ of new data. The total sample size is 193.5 $\pm$ 11.4 $pb^{-1}$ for CEM and CMUP triggers and 175.3 $\pm$ 10.3 $pb^{-1}$ for CMX triggers. This offline luminosity has been corrected by a factor of 1.019 [3] and the uncertainty on the integrated luminosity is taken to be 5.9% [4] . The detailed numbers can be found in Table 1, which differs slightly from Table 1 in CDF note 6206 only due to the updated good run list.

# 4 Signal and Background Simulation

Official Top group 4.9.1 Monte Carlo samples and public Top Lepton+Jets group TopNtuples of these samples were used in this analysis [5]. The $t\bar{t}$ signal sample was generated with both PYTHIA and HERWIG. Due to the lack of QED radiation in the HERWIG sample, we choose to use PYTHIA to model the $t\bar{t}$ signal. The dominant background from $W$+jets was modeled with the leading order W+N parton matrix element generator ALPGEN interfaced to the parton shower HERWIG. To eliminate problems with double counting between the matrix element and the parton shower, only the W+N parton MC is used to model the W+N jet bin. For example, a plot of $H_T$ in the reconstructed W+2 jet bin only used the ALPGEN+HERWIG W+2 parton MC. Other backgrounds like Z+jets, WW+jets, WZ+jets, ZZ+jets were also modeled with ALPGEN+HERWIG.

The background from QCD multi-jets is difficult to model with MC. Therefore, we model this QCD background from the data. Two separate samples were used and compared. The first was constructed of non-isolated ($I > 0.2$) electrons and muons. This is the sample used for the central values of the top cross section. The second sample was built from conversion electrons as an alternate sample to set the systematic error. In both cases, all other cuts like lepton ID, $\not{E}_T$, number of jets, etc. were the same.

# 5 Event Selection

Top-quark events in the lepton+jets channel are characterized by a high-momentum lepton, a substantial amount of missing energy due to the leptonic $W$ decay and, in principle, 4 high energy jets due to the hadronic decay of the second $W$ and the two $b$ quarks in the event.

We use the electron and muon definitions established by the W and Z cross section groups in this analysis. We use CEM electrons and CMUP and CMX muons only. Given that all of our events must include, by definition, a high-momentum electron or muon, the data sample should only contain events that came in on CEM, CMUP or CMX triggers. However, there are some volunteers. The requirement that the lepton came in on the appropriate trigger is made.

To clean up the sample, electron events in which the primary electron is identified as a conversion are removed. Similarly, muon events containing a cosmic ray are removed. Signal events should contain only one high-$p_T$ lepton. We veto events consistent with a leptonic $Z$ decay. Although we do not yet use the data sample with plug electron triggers, we do veto events with an additional PHX electron. Since the dilepton group has not yet provided a filter with which to cut dilepton events, we have chosen to make our own. This is not quite the same as removing dilepton events and thus our data samples are not strictly orthogonal. In the future, we will implement a true dilepton cut once it becomes available.

For consistency with Run I, the same jet clustering algorithm, JETCLU with cone of 0.4, and jet selection criteria, $E_T$>15 GeV and $|\eta| < 2.0$, are used. However in this analysis, we select events after jet corrections have been made up to and including level 4 in order to flatten calorimeter response, apply the Run I raw energy scale factor and correct for multiple interactions. The MET is corrected for the muon $p_T$ in TopEventModule. We further correct the MET only for the effect of jet corrections (for all jets with level 4 corrected $E_T$>8GeV and $|\eta| < 2.5$) and for the curvature correction to the muon $p_T$.

In order to ensure the jets and lepton are reconstructed from the same interaction, we now require that the event z vertex position, as defined by the closest good quality ZVertexModule object ($\geq 12$ meaning at least 2 good COT tracks) to be within 5 cm of the tight lepton $z_0$.

To reduce the QCD background, we require the $\Delta\phi$ between the MET and the most energetic jet to be within [0.5,2.5] radians if the MET is below 30 GeV. This cut is described in more detail later in this section.

For completeness, a summary of all of our selection criteria can be found in Table 2. Further details on the event selection criteria such as cosmic ray removal and $Z$ veto can be found in [10].

## 5.1 Data events

Table 3 contains the number of observed events in the data as a function of reconstructed jet multiplicity bin. All of the selection criteria described above have been applied. The data is broken out in terms of electron+jets and muon+jets. For comparison, the number of expected $t\bar{t}$ events is also presented based on an assumed cross section of 7.0 pb.

## 5.2 $\Delta\phi$ cut

With the current relatively small non-isolated lepton data sample, it is difficult to construct appropriate templates of the discriminant variable with sufficient statistics in each bin in order to obtain a smooth shape for the QCD background. Furthermore, the non-isolated lepton sample may not model the background shapes in the isolated region perfectly. Therefore, even though QCD is not our dominant background, it is still useful to reduce its contribution in our signal region in order to minimize our sensitivity to a shape that is not known very well.

| Variable | Value |
|---|---|
| **Global event quantities** | |
| $Z$ and cosmic veto | applied |
| Missing $E_T$ | $>$ 20 GeV |
| $\Delta\phi$(MET,Lead Jet) | $>$ 0.5 and $<$2.5 if $\not{E}_T <$ 30 GeV |
| # tight leptons | 1 |
| **Jets** | |
| $E_T$ | $\geq$ 15 GeV (corrected to level 4) |
| $\|\eta\|$ | $<$ 2 |
| **Tight electrons** | |
| Fiducial and CEM | = 1 |
| $E_T$ | $\geq$ 20 GeV |
| $p_T$ | $\geq$ 10 GeV/$c$ |
| $E/p$ (if $p_T \leq$ 50 GeV/$c$) | $<$ 2.0 |
| $E_{had}/E_{EM}$ | $\leq$ 0.055 + 0.00045*E |
| $L_{shr}$ | $<$ 0.2 |
| $\|\Delta z\|$ | $<$ 3 cm |
| $\|\Delta x\|$ | $>$ -1.5 and $<$ 3.0 cm |
| $\chi^2_{Strip}$ | $<$ 10 |
| $\|\Delta z_0\|$ | $<$ 60 cm |
| $\|Z_{vtx} - z_0\|$ | $<$ 5 cm |
| # axial segments | $\geq$ 3 |
| # stereo segments | $\geq$ 3 |
| Isolation | $<$ 0.1 |
| Not a conversion | $\Delta\Lambda >$ 0.04 and $\Delta r >$ 0.2 |
| **Tight muons** | |
| Region | CMUP or CMX |
| $p_T$ | $\geq$ 20 GeV/$c$ |
| $E_{EM}$ | $<$ max(2, 2 + 0.0115*($p$-100)) |
| $E_{HAD}$ | $<$ max(6, 6 + 0.0280*($p$-100)) |
| $\|\Delta x\|_{CMU}$ | $<$ 3.0 cm |
| $\|\Delta x\|_{CMP}$ | $<$ 5.0 cm |
| $\|\Delta x\|_{CMX}$ | $<$ 6.0 cm |
| $\|\Delta z_0\|$ | $<$ 60 cm |
| $\|Z_{vtx} - z_0\|$ | $<$ 5 cm |
| # axial segments | $\geq$ 3 |
| # stereo segments | $\geq$ 3 |
| Isolation | $<$ 0.1 |

Table 2: The "standard" event selection criteria used in the top group.

| Njets | CEM | CMUP | CMX | Total | Expected $t\bar{t}$ |
|---|---|---|---|---|---|
| 0 | 99454 | 49672 | 26531 | 175657 | 0.3 |
| 1 | 9407 | 4617 | 2365 | 16389 | 4.6 |
| 2 | 1442 | 705 | 349 | 2496 | 23.9 |
| 3 | 254 | 101 | 46 | 401 | 44.8 |
| $\geq$4 | 78 | 33 | 7 | 118 | 53.0 |

Table 3: The number of events selected as a function of jet multiplicity for 193 pb$^{-1}$ of CEM and CMUP and 175 pb$^{-1}$ of CMX data. For $t\bar{t}$ expectation, a cross section of 7 pb is assumed.

In order to motivate the cut, it is useful to look at Figures 1 and 2. These two-dimensional plots show the $\Delta\phi$ between $\not{E}_T$ and leading jet versus $\not{E}_T$ for the relevant signal and background samples. Figure 1 plots the non-isolated lepton data that we will use to model the QCD background in the 1 jet bin. The top plot of this figure is electron+1 jet data and the bottom is the muon+1 jet data. In Figure 2, the same 2D distributions are shown for $t\bar{t}$ (top) and W+jets(bottom) MC events in the reconstructed W+3 jet bin. From these plots, it is clear that by eliminating events in the region OUTSIDE $0.5 < \Delta\phi < 2.5$ for the $\not{E}_T$ band between 20 and 30 GeV, we can reduce the QCD background while having a minimum impact on the signal. In the signal region (W+$\geq$ 3 jets), this cut removes 40% of the QCD background while removing only 5.3 $\pm$ 1 % of the $t\bar{t}$ signal.

During the summer 2003 conference preblessing talk, there was some concern that the $\Delta\phi$ cut was doing something strange to the data – namely it was cutting harder than the MC predicted. That is, if one used the efficiency of this cut for background and signal, that number of events observed passing the cut is out of line with what one would expect. In Table 4, we present the impact of the $\Delta\phi$ cut on electrons and muons separately as well as our ability to predict the outcome. The table shows the number of events before and after the $\Delta\phi$ cut is made for each jet multiplicity bin. The predicted number of events after the $\Delta\phi$ cut is calculated from this formula:

$$n_{\text{predicted}} = n_{\text{before}} * f^{\text{QCD}}_{\text{before}} * \epsilon^{\text{QCD}}_{\Delta\phi} + n_{\text{before}} * (1.0 - f^{\text{QCD}}_{\text{before}}) * \epsilon^{\text{W}}_{\Delta\phi}$$

where the $\epsilon^{\text{QCD}}_{\Delta\phi}$ is the ratio of the MET vs ISO predicted background before and after the $\Delta\phi$ cut is made. The $\epsilon^{\text{W}}_{\Delta\phi}$ is the taken as 93% from the ALPGEN W+jets MC sample. The predicted numbers agree with the observed numbers within the statistical precision of the data sample. Based on this study, the cut is behaving exactly as we expect.

# 6 Top kinematic quantities

Due to the large mass of the top quark, $t\bar{t}$ events have very large total energy, leading to large $H_T$ compared with the non-top background processes. The main background to $t\bar{t}$ production in the "lepton+jets" channel comes from $W$ plus multiple hadronic jets, called simply $W$+jets here. Figure 3 (upper left) shows a comparison of the shape of the distribution of $H_T$ for $t\bar{t}$ and $W$+3 parton samples in events with three or more jets. The $t\bar{t}$ signal has, on average, much more total transverse energy than the background $W$+jets signal as one would expect.

While the choice of $H_T$ can certainly be motivated by first principles, it is not necessarily the only choice for a kinematic variable. Other choices such as the sum of all the jet energies in

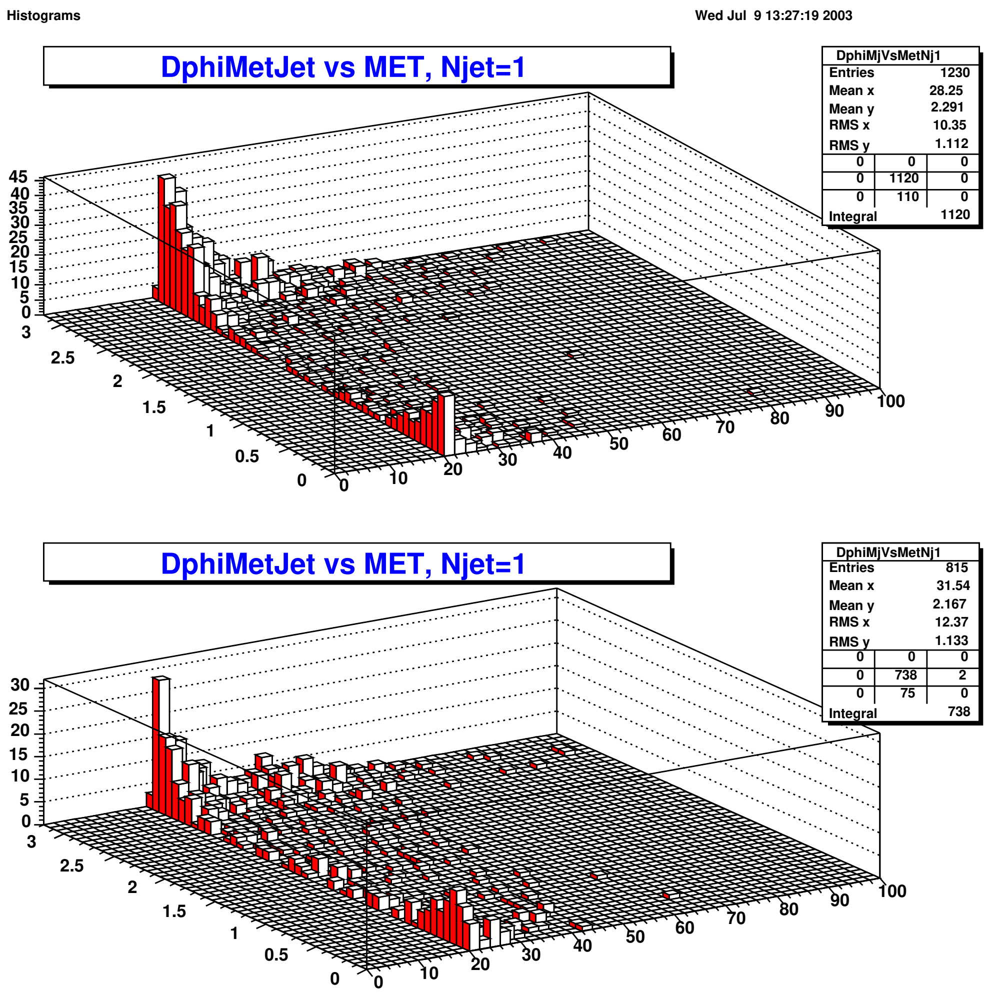



Figure 1: Distribution of $\Delta\phi$ versus $\not\!\!E_T$for non-isolated leptons in data for the one jet bin where there are reasonable statistics. The top plot is electrons and the bottom muons. This sample is a reasonable model of the QCD background. Note the pile-up of events in the $20 < \not\!\!E_T < 30$ and $0.5 > \Delta\phi > 0$ as well as $2.5 > \Delta\phi > 3.14$. Those are the background events that we are trying to remove.

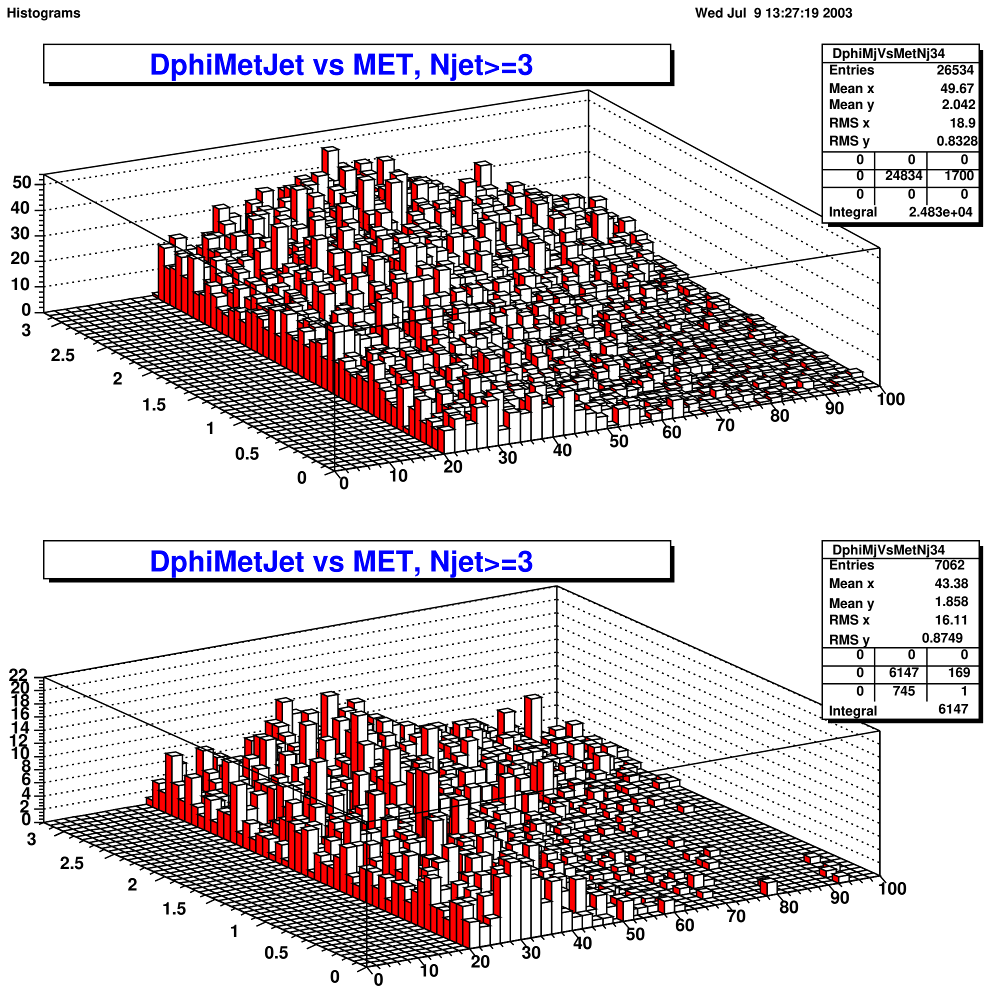



Figure 2: Distribution of $\Delta\phi$ versus $\not\!\!E_T$ in $t\bar{t}$ MC (above) and ALPGEN W+3 parton MC (below) in the $\geq$3 jet bin (signal region). Here the cut $0.5 < \Delta\phi < 2.5$ has very little impact on either $t\bar{t}$ or the dominant W+jet background.

| Njets | Before $\Delta\phi$ Cut | | After $\Delta\phi$ Cut | | |
|---|---|---|---|---|---|
| | Events | Bkg % | Events | Bkg % | Predicted |
| electrons | | | | | |
| 1 | 10797 | $9.9 \pm 0.3$ | 9501 | $4.3 \pm 0.2$ | 9504.15 |
| 2 | 1755 | $14.6 \pm 0.7$ | 1459 | $6.5 \pm 0.5$ | 1508.81 |
| 3 | 305 | $17.9 \pm 2.0$ | 254 | $10.9 \pm 1.6$ | 265.973 |
| $\geq$4 | 87 | $14.9 \pm 3.7$ | 79 | $6.2 \pm 2.1$ | 74.1893 |
| muons | | | | | |
| 1 | 7815 | $5.4 \pm 0.2$ | 7050 | $3.0 \pm 0.2$ | 7106.86 |
| 2 | 1198 | $4.6 \pm 0.3$ | 1067 | $2.0 \pm 0.2$ | 1087.43 |
| 3 | 183 | $6.0 \pm 1.0$ | 152 | $4.0 \pm 0.9$ | 167.227 |
| $\geq$4 | 44 | $6.6 \pm 3.1$ | 40 | $3.5 \pm 2.0$ | 39.7763 |

Table 4: A test of the $\Delta\phi$ cut on the lepton+jets as a function of Njet.

the event, the sum of the two leading jet $E_T$, $E_{T12} \equiv E_T(j1) + E_T(j2)$ and the minimum dijet invariant mass $m_{jj}^{min}$ $\equiv$min($m_{jj}$) could also prove useful. CDF Note 6078 looked at a number of such discriminating variables in Run I. The final set of variables we use in this analysis is based on that study as well as our own investigations. Figures 3, 4, and 5 show the degree of separation between $t\bar{t}$ and $W$+3 parton for several variables for the W+$\geq$ 3 jet signal region.

In order to decide which variable to use as the discriminant in this cross section analysis, a single input neural network was used to discriminate $t\bar{t}$ from the dominant background of W+jets. A wide range of variables were studied. Table 5 lists some of the most promising – some are based on $E_T$ and the rest are based on shape. The numbers in the table represent the percentage of correct classifications that are possible with the particular variable in question. For example, using the $E_T$ of the third leading jet, it is possible to correctly classify 74.4% of a sample of W+jets and $t\bar{t}$ while the sphericity variable can only obtain the proper classification 60% of the time.

Based on the results in Table 5, there are several $E_T$ based variables that have very similar discriminating power. This analysis could have used $E_T$ of jet 3, $H_T$, $E_{T2} + E_{T3}$, and $E_T$ of jet 2. All four of these variables have very similar separation power. In the future, the final choice of variables will be motivated based on the combined systematic and statistical uncertainty as estimated from pseudo experiments. For now, we will choose the historical favorite from Run I: the kinematic variable $H_T$.

# 7 Cross section extraction

In this analysis, the $t\bar{t}$ cross section comes from a binned likelihood fit in MINUIT of the observed distribution of $H_t$ to a combination of:

- Signal: This is modeled using PYTHIA $t\bar{t}$, using a top mass value of 175 $GeV/c^2$.
- W+Jets background: When fitting the W+$\geq$ 3 jets sample, the W+3 parton ALPGEN sample is used. When fitting the W+$\geq$ 4 jets sample, the W+4 parton ALPGEN sample is used.
- non-W Background: This is modeled using the non-isolated lepton sample.

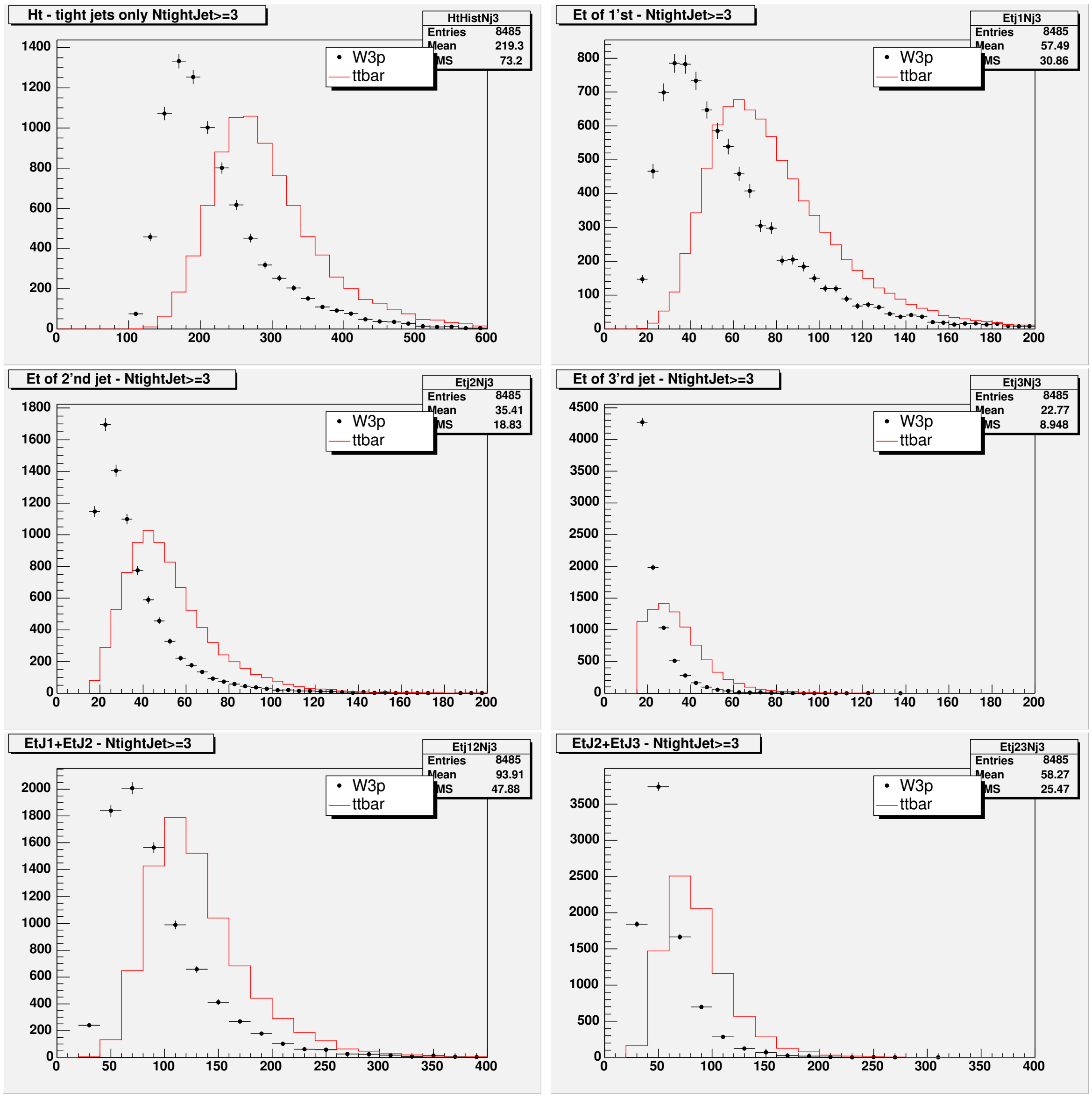



Figure 3: Comparison plots of kinematic variables using ALPGEN W+3 parton (black points) and $t\bar{t}$ (red curve) for a variety of kinematic variables in the reconstructed W+$\geq$ 3 jet bin. The curves are normalized to equal area.

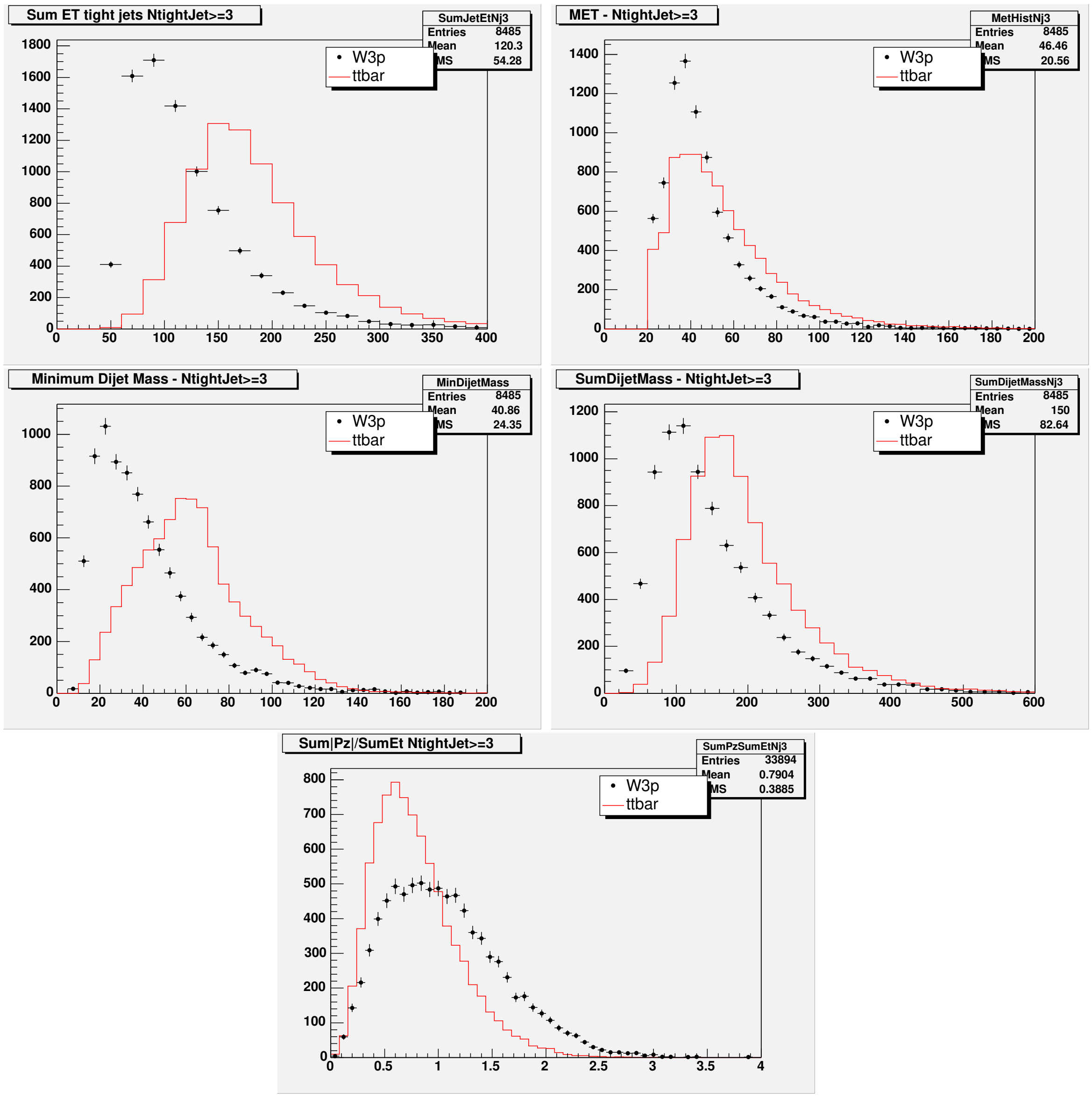



Figure 4: Comparison plots of kinematic variables using ALPGEN W+3 parton (black points) and $t\bar{t}$ (red curve) for a variety of kinematic variables in the reconstructed W+$\geq$ 3 jet bin. The curves are normalized to equal area.

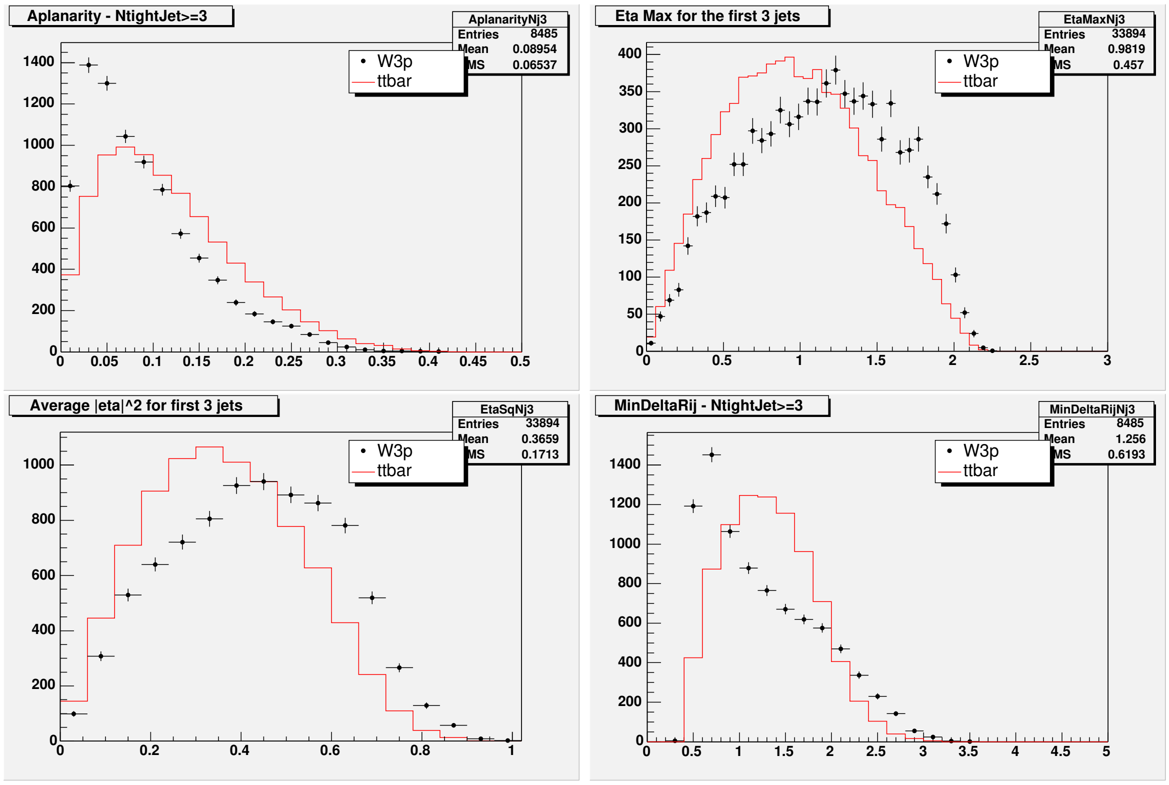



Figure 5: Comparison plots of shape variables using ALPGEN W+3 parton (black points) and $t\bar{t}$ (red curve) for a variety of kinematic variables in the reconstructed W+$\geq$ 3 jet bin. The curves are normalized to equal area.

| | **Herwig $t\bar{t}$** **ALPGEN W+jets** | **Pythia $t\bar{t}$** **ALPGEN W+jets** |
|---|---|---|
| Et jet3 | 74.4% ± 0.5% | 73.7% ± 0.5% |
| $H_T$ | 73.9% ± 0.5% | 72.5% ± 0.5% |
| Et(2+3) | 73.9% ± 0.5% | 73.4% ± 0.5% |
| Et jet2 | 73.7% ± 0.5% | 71.8% ± 0.5% |
| Min Dijet Mass | 72.3% ± 0.6% | 71.0% ± 0.6% |
| Et (1+2) | 71.0% ± 0.6% | 71.0% ± 0.6% |
| Event Inv. Mass | 70.8% ± 0.6% | 70.1% ± 0.6% |
| Et jet4 | 69.4% ± 0.6% | 69.5% ± 0.6% |
| Et jet1 | 67.7% ± 0.6% | 67.1% ± 0.6% |
| Sum Dijet Mass | 66.6% ± 0.6% | 64.9% ± 0.6% |
| Sum Ez/SumEt | 63.3% ± 0.6% | 64.0% ± 0.6% |
| Aplanarity | 59.2% ± 0.6% | 60.2% ± 0.6% |
| Min Jet Sep. | 60.3% ± 0.6% | 60.6% ± 0.5% |
| Sphericity | 60.0% ± 0.6% | 59.6% ± 0.6% |
| $\not\!\!E_T$ | 57.7% ± 0.6% | 55.7% ± 0.6% |

Table 5: Using a single input neural net to discriminate $t\bar{t}$ from W+jets, the table gives the percentage of correct classifications that are possible with a given kinematic or shape variable. The errors are statistical only.

The non-W background is actually constrained to the value expected using the MET versus Isolation technique, as described in the next section. Other background sources ($WW$, $WZ$, $W \to \tau\nu$, $Z \to ee/\mu\mu/\tau\tau$) are estimated from Monte Carlo to account for on order ~10% of the total observed events in our sample. In figure 6, the $H_T$ variable for the backgrounds listed above(red) are compared to the ALPGEN W+3 parton MC (points). While the agreement between the two are not perfect, they are very close. As a result, we could have just used the W+jets background shape. However, we choose to add these backgrounds to the W+jets shape according to the normalisation given in Table 6 for W+≥3 jets and in Table 7 for W+≥4 jets.

We actually have two separate fit methods, which are described in the following subsections. The first is based on fitting the $t\bar{t}$ cross section directly as a fit parameter, while the second fits for the signal fraction. We have checked that both methods return the same results.

## 7.1 Cross section fit method

The cross section fit method is based on a likelihood $\mathcal{L}$ formed from the product of the Poisson probabilities of observing $n_{ij}$ events in bin $ij$ of number of reconstructed jets and $H_T$, given that $y_{ij}$ such events are predicted:

$$\mathcal{L} = \prod_{ij} \frac{y_{ij}^{n_{ij}} e^{-y_{ij}}}{n_{ij}!} \quad . \tag{1}$$

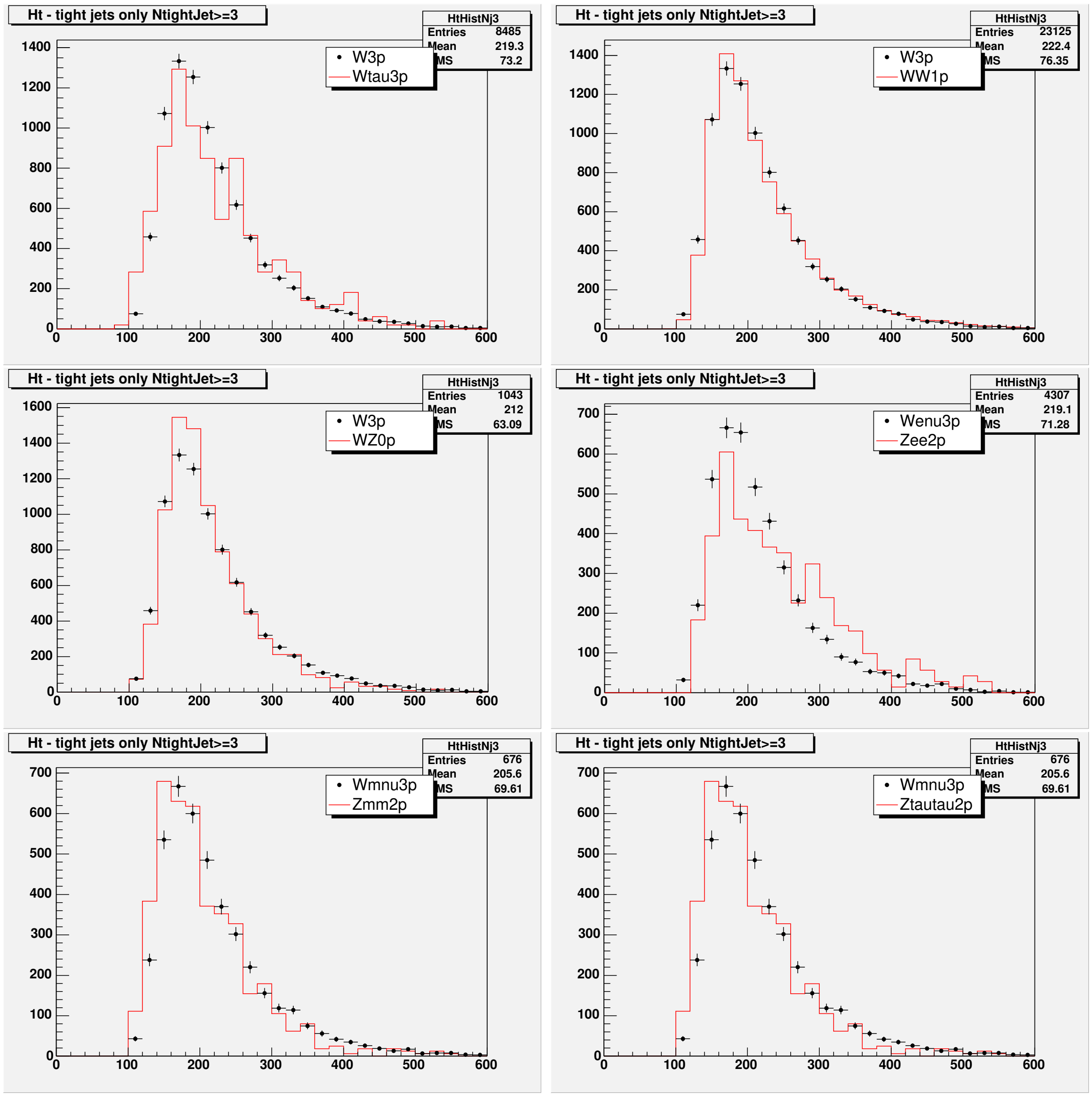



Figure 6: $H_T$ shape comparison between ALPGEN W+3 parton MC (points) and 6 other backgrounds(red) listed in the text.

| MC Sample | MC $\sigma$ | $\ell_{type}$ | $N_{gen}$ | $N_{Acc}$ | $F^{acc}_{data/mc}$ | $L_{data}$ | Contribution |
|---|---|---|---|---|---|---|---|
| W($e\nu$)+3p | 89.9±0.6 | CEM | 268682 | 4309 | 0.908 ± 0.017 | 195.3 | 255.7 |
| W($\mu\nu$)+3p | 89.9±0.6 | CMUP | 293994 | 2716 | 0.735 ± 0.027 | 195.3 | 119.2 |
| | | CMX | | 1209 | 0.925 ± 0.024 | 177.1 | 60.6 |
| W($e\nu$)+bb+1p | 1.11±0.01 | CEM | 236005 | 2805 | 0.908 ± 0.017 | 195.3 | 2.3 |
| W($mu\nu$)+bb+1p | 1.11±0.01 | CMUP | 171610 | 1112 | 0.735 ± 0.027 | 195.3 | 1.0 |
| | | CMX | | 509 | 0.925 ± 0.024 | 177.1 | 0.5 |
| W($\tau\nu$)+3p | 89.9±0.6 | CEM | 261917 | 220 | 0.908 ± 0.017 | 195.3 | 13.4 |
| | | CMUP | | 128 | 0.735 ± 0.027 | 195.3 | 6.3 |
| | | CMX | | 60 | 0.925 ± 0.024 | 177.1 | 3.4 |
| WZ+0p | 2.37±0.01 | CEM | 191011 | 555 | 0.908 ± 0.017 | 195.3 | 1.2 |
| | | CMUP | | 320 | 0.735 ± 0.027 | 195.3 | 0.6 |
| | | CMX | | 140 | 0.925 ± 0.024 | 177.1 | 0.3 |
| WW+1p | 4.38±0.00 | CEM | 926616 | 13267 | 0.908 ± 0.017 | 195.3 | 11.1 |
| | | CMUP | | 7828 | 0.735 ± 0.027 | 195.3 | 5.3 |
| | | CMX | | 3221 | 0.925 ± 0.024 | 177.1 | 2.5 |
| Z($ee$)+2p | 23.3±0.1 | CEM | 287734 | 316 | 0.908 ± 0.017 | 195.3 | 4.5 |
| Z($\mu\mu$)+2p | 23.3±0.1 | CMUP | 292446 | 507 | 0.735 ± 0.027 | 195.3 | 5.8 |
| | | CMX | | 240 | 0.925 ± 0.024 | 177.1 | 3.1 |
| Z($\tau\tau$)+2p | 23.3±0.1 | CEM | 296823 | 281 | 0.908 ± 0.017 | 195.3 | 3.9 |
| | | CMUP | | 171 | 0.735 ± 0.027 | 195.3 | 1.9 |
| | | CMX | | 72 | 0.925 ± 0.024 | 177.1 | 0.9 |
| $t\bar{t}$ | 7 | CEM | 384875 | 17800 | 0.908 ± 0.017 | 195.3 | 57.4 |
| | | CMUP | | 10906 | 0.735 ± 0.027 | 195.3 | 28.5 |
| | | CMX | | 4025 | 0.925 ± 0.024 | 177.1 | 12.0 |
| single top | 3 | CEM | 440000 | 2526 | 0.908 ± 0.017 | 195.3 | 3.1 |
| | | CMUP | | 1485 | 0.735 ± 0.027 | 195.3 | 1.5 |
| | | CMX | | 614 | 0.925 ± 0.024 | 177.1 | 0.7 |

Table 6: Background information for >=3 Jets.

This likelihood is maximized by minimizing (utilizing the MINUIT package) the negative log likelihood:

$$-\ln \mathcal{L} = \sum_{ij} n_{ij} \ln y_{ij} - y_{ij} \tag{2}$$

where we have dropped the term involving $\ln n_{ij}!$ since it simply results in an additive offset.

The number of events predicted in each bin, $y_{ij}$ is simply the sum over all sources $k$, each of which is the product of the integrated luminosity $L$, cross section $\sigma_k$, and acceptance for that source:

$$y_{ij} = \sum_{k} L f_{MC} \sigma_k \frac{N^{acc}_{ijk}}{N^{gen}_{k}} \quad . \tag{3}$$

Here the superscripts "acc" and "gen" refer to accepted and generated Monte Carlo events, respectively, for source $k$. The integrated luminosity $L$ and data to Monte Carlo scale factor

| MC Sample | MC $\sigma$ | $\ell_{type}$ | $N_{gen}$ | $N_{Acc}$ | $F^{acc}_{data/mc}$ | $L_{data}$ | Contribution |
|---|---|---|---|---|---|---|---|
| W($e\nu$)+4p | 32.5±0.5 | CEM | 292802 | 2182 | 0.908 ± 0.017 | 195.3 | 42.9 |
| W($\mu\nu$)+4p | 32.5±0.5 | CMUP | 297500 | 1403 | 0.735 ± 0.027 | 195.3 | 22.0 |
| | | CMX | | 638 | 0.925 ± 0.024 | 177.1 | 11.4 |
| W($e\nu$)+bb+1p | 1.11±0.01 | CEM | 236005 | 401 | 0.908 ± 0.017 | 195.3 | 0.3 |
| W($mu\nu$)+bb+1p | 1.11±0.01 | CMUP | 171610 | 168 | 0.735 ± 0.027 | 195.3 | 0.2 |
| | | CMX | | 80 | 0.925 ± 0.024 | 177.1 | 0.1 |
| W($\tau\nu$)+3p | 89.9±0.6 | CEM | 261917 | 38 | 0.908 ± 0.017 | 195.3 | 2.3 |
| | | CMUP | | 26 | 0.735 ± 0.027 | 195.3 | 1.3 |
| | | CMX | | 10 | 0.925 ± 0.024 | 177.1 | 0.6 |
| WZ+0p | 2.37±0.01 | CEM | 191011 | 73 | 0.908 ± 0.017 | 195.3 | 0.2 |
| | | CMUP | | 48 | 0.735 ± 0.027 | 195.3 | 0.1 |
| | | CMX | | 29 | 0.925 ± 0.024 | 177.1 | 0.1 |
| WW+1p | 4.38±0.00 | CEM | 926616 | 2589 | 0.908 ± 0.017 | 195.3 | 2.2 |
| | | CMUP | | 1481 | 0.735 ± 0.027 | 195.3 | 1.0 |
| | | CMX | | 588 | 0.925 ± 0.024 | 177.1 | 0.5 |
| Z($ee$)+2p | 23.3±0.1 | CEM | 287734 | 67 | 0.908 ± 0.017 | 195.3 | 1.0 |
| Z($\mu\mu$)+2p | 23.3±0.1 | CMUP | 292446 | 75 | 0.735 ± 0.027 | 195.3 | 0.9 |
| | | CMX | | 39 | 0.925 ± 0.024 | 177.1 | 0.5 |
| Z($\tau\tau$)+2p | 23.3±0.1 | CEM | 296823 | 72 | 0.908 ± 0.017 | 195.3 | 1.0 |
| | | CMUP | | 34 | 0.735 ± 0.027 | 195.3 | 0.4 |
| | | CMX | | 15 | 0.925 ± 0.024 | 177.1 | 0.2 |
| $t\bar{t}$ | 7 | CEM | 384875 | 9597 | 0.908 ± 0.017 | 195.3 | 31.0 |
| | | CMUP | | 5983 | 0.735 ± 0.027 | 195.3 | 15.6 |
| | | CMX | | 2176 | 0.925 ± 0.024 | 177.1 | 8.1 |
| single top | 3 | CEM | 440000 | 461 | 0.908 ± 0.017 | 195.3 | 0.6 |
| | | CMUP | | 292 | 0.735 ± 0.027 | 195.3 | 0.3 |
| | | CMX | | 113 | 0.925 ± 0.024 | 177.1 | 0.1 |

Table 7: Background information for >=4 Jets.

$f_{MC}$ multiply all sources equally. The sources include $t\bar{t}$ signal, $W + n$ partons from ALPGEN, and other Monte-Carlo-estimated sources, as well as those estimated from data. (The latter are incorporated by direct addition rather than by a Monte Carlo acceptance, clearly.)

The cross sections $\sigma_k$ are the free parameters of the fit, and vary within gaussian constraints determined as described above. The constraints appear as additive terms in $-\ln\mathcal{L}$ as

$$-\ln\mathcal{L} = \sum_{ij} n_{ij}\ln y_{ij} - y_{ij} + \sum_k \frac{(\sigma_k - \tilde{\sigma}_k)^2}{2s_k^2} \quad . \tag{4}$$

The $s_k$ here are the gaussian widths of the constraints. Of course, not all sources have such constraints; this equation illustrates the form of those that do. The integrated luminosity $L$ times MC scale factor $f_{MC}$ can also have a constraint term reflecting its uncertainty.

The MINUIT algorithms MINOS and MIGRAD find the minimum of $-\ln\mathcal{L}$ by varying the $\sigma_k$, and report the values at the minima and their correlation matrix. The final uncertainty in the $t\bar{t}$ cross section is simply taken from those reported by the result of MINOS.

### 7.2 Signal fraction fit method

The signal fraction fit method is very similar to the above, except the free parameters of the fit are the fractions, rather than the cross-section. It is however, an independent piece of code!

## 8 QCD Background estimate

The sample that we are using to model the background from non-W is the non-isolated lepton sample. This sample presumably includes fake leptons, as well as $b\bar{b}$ in which one of the b's decays semi-leptonically to an electron or muon. This sample requires that the isolation of the lepton be $> 0.2$, but otherwise, all of the other cuts, both on the lepton as well as the MET and Jet cuts, are identical. Note however, that:

- MET is corrected for non-isolated muons.
- Any jet within 0.1 in $\Delta R$ of a non-isolated lepton is removed, both to avoid double-counting energy as a lepton and as a jet and since isolated leptons do not have jets in such close proximity.

With the $\not{E}_T$ vs Isolation technique that has been used since the Run I "evidence" paper, we can determine the expected QCD background fraction. Table 8 contains the results of this fit on the isolated-lepton sample with the $\Delta\phi$ cut applied. We see that in the signal region, we expect approximately 7.0 $\pm$ 1.7% of the background to be attributed to QCD. The electron+jets sample has substantially more (factor of 3) QCD background than that in the muon sample. This makes sense since there are substantial numbers of conversions in the electron sample that are not present in the muon sample.

The 7% background in the $W+\geq$ 3 jet bin is however an over estimate of the total QCD background. We need to correct for "W-like" processes that contaminate our sample. In order to correct for this, we determine the fraction of W-like events which populate each of the MET vs ISO regions. Events which we considered include $W \to \ell\nu$+3 partons, $WW$+1 parton, $Z \to \ell\ell$+2partons, where $\ell = e, \mu, \tau$. We determine the corrected QCD background to be 6.3 $\pm$ 1.6%. The relative contributions to this background is shown in the last row of Table 8. These numbers are as in the summer conference note. Since the background levels did not change significantly with more data, we still use these values in the fits to the full dataset.

## 9 Expected Sensitivities

In order to determine the expected precision of the $H_T$ variable as well as the others presented above on the $t\bar{t}$ cross section, a series of pseudo-experiments were performed. We used the acceptance numbers based on our cuts, 125 pb$^{-1}$ of data, and an assumed $t\bar{t}$ cross section of 7 pb. The number of QCD background events was determined from MET vs ISO in the non-isolated lepton sample. The $\Delta\phi$ cut was imposed. (The process of determining the size of the QCD sample will be explained in detail in a later section on backgrounds.) The number of W+jet events is derived from the number of observed data events minus the number of events expected from QCD and $t\bar{t}$ production. Each of the three numbers (QCD, $t\bar{t}$, and W+jets) correspond to the mean of a Poisson. The $H_T$ is used as our primary discriminating variable

| W+N jets | Ele Bckg | Muo Bckg | Total Bckg |
|---|---|---|---|
| **1** | 3.8 ± 0.2% | 2.9 ± 0.2% | 3.4 ± 0.3% |
| **2** | 6.1 ± 0.5% | 2.0 ± 0.2% | 4.3 ± 0.5% |
| **≥3** | 8.9 ± 1.5% | 4.0 ± 0.9% | 7.0 ± 1.7% |
| **≥3 corrected** | 7.7 ± 1.3% | 3.1 ± 0.9% | 6.3 ± 1.6% |

Table 8: QCD Background fraction as a function of Njet. Numbers were determined using the MET vs ISO technique. The isolated lepton sample from the summer conferences was used and the $\Delta\phi$ cut was applied. The *corrected* row means that the QCD background calculation was corrected for W-like event contamination in our isolated lepton sample. The value of 6.3 ± 1.6 % will be the QCD background fraction used in the fits.

but a number of other variables are also checked. Each pseudo-experiment samples the $H_T$ distribution from $t\bar{t}$MC, ALPGEN W+3 parton MC, and the non-isolated lepton data sample for QCD.

The results of these pseudo-experiments can be found in figure 7. The top plot in this figure contains the expected sensitivity for a variety of kinematic and shape variables in the W+≥3 jet bin. The bottom plot in the same figure is for the W+≥4 jet bins. The data point for each variable indicates the mean returned uncertainty from fitting the pseudo experiments. The error bars represent the one sigma range of possible returned uncertainties from the fit. Just to belabor the point and be perfectly clear, the smaller the median value and the smaller the error bars the better. From the plots, one can see, as in the neural net studies above, that $H_T$ is not the only choice for a kinematic fit. Others like $E_T$ of the 3rd jet and sum $E_T$ of all the jets in the event do equally well. For the $H_T$ variable in the $\geq$ 3 jet bin, we expect a fractional uncertainty of 25% in the $\geq$ 3 jet bin and 35% in the ≥4 jet bin.

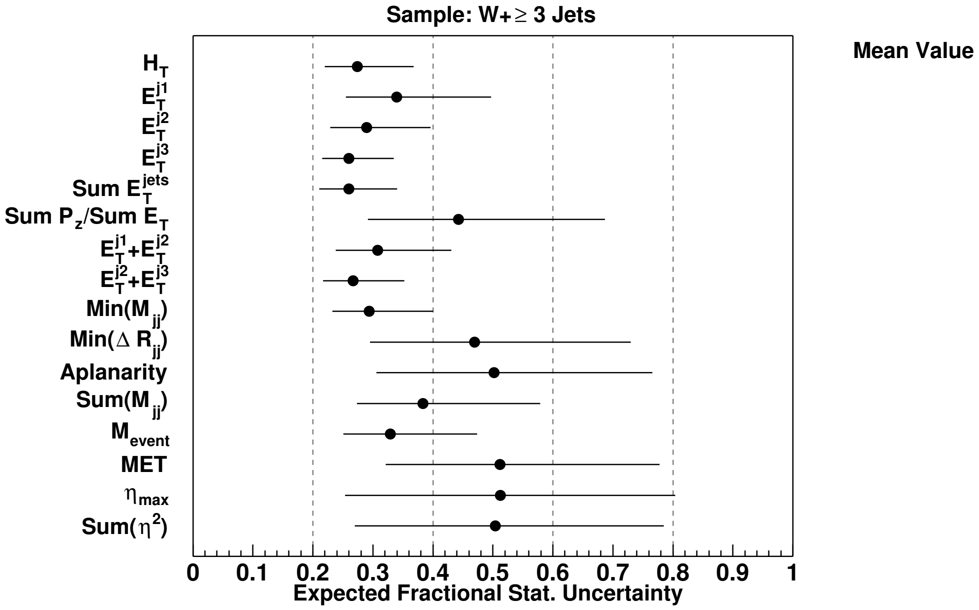
With 195 pb$^{-1}$, we have 50% more data and we expect the uncertainties to scale down correspondingly by $\sqrt{1.5}$, that is to a fractional uncertainty of 20% in the ≥3 jet bin and 29% in the ≥4 jet bin.

# 10 MC background shape validation

Since this is a "shape" analysis and not a counting experiment, it is imperative that the MC's model the distributions properly. As stated before, the dominant background to $t\bar{t}$ is W+jets which we model with ALPGEN W+N parton MC. To validate the MC, we need to look in a region that is top-depleted.

The top plot in figure 8 shows the observed W+jet data (red) along with a prediction for the amount of $t\bar{t}$ (blue) one might expect as a function of jet multiplicity. The theoretical cross section used for the blue curve ($t\bar{t}$) is 7 pb. The lower plot in the same figure is the ratio of the red and blue curves. From the ratio, one can easily see that very little top is expected in the W+1 and W+2 jet bins and so these are good choices to test the background MC shapes with data. Furthermore, even in the exclusive W+3 jet bin, the amount of top expected is only about 12% of the total and thus can also be used as a cross check. Where we expect the bulk of the signal ($\geq$ 4 jet bin), the top to W ratio is approximately 1:1.

In the W+1 jet sample, we perform a two-component fit using shapes from ALPGEN W+1 parton MC and non-isolated leptons. In the W+2 jet bin, the ALPGEN sample is the more



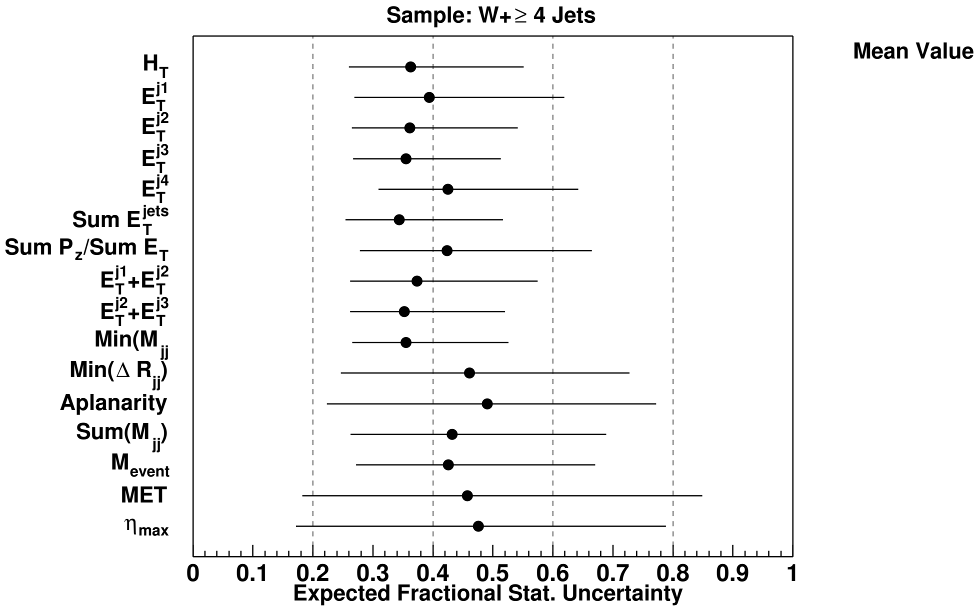




Figure 7: Expected fractional statistical uncertainty as determined by pseudo-experiments based on the summer conference data sample size. The points represent the mean returned fractional uncertainty from fitting the pseudo-experiments and the error bar represents the 1 sigma range of possible returned uncertainties from the fit. The $\Delta\phi$ cut was imposed. The inputs to the PE's were $t\bar{t}$ MC, ALPGEN W+3,4 parton MC and QCD as drawn from the non-isolated lepton sample. The top plot is for $\geq$3 jet bin and the bottom is the $\geq$4 jet bin.

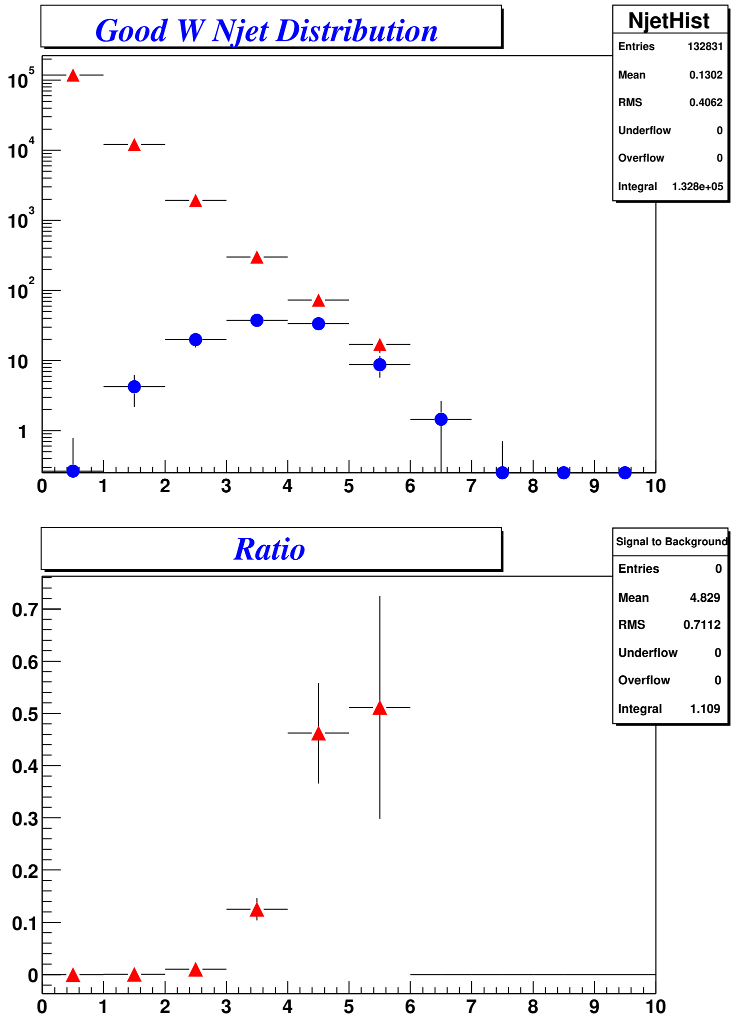



Figure 8: W+njet distributions and the comparison to $t\bar{t}$ MC.

appropriate W+2 parton MC. In the W+3 jet bin, there is a non negligible amount of $t\bar{t}$ of order 12% as described above. Therefore, we also add a third component from PYTHIA $t\bar{t}$ MC to the ALPGEN W+3 parton MC and QCD shapes. When $t\bar{t}$ is introduced in this 3 jet bin, we do not fix its amount but rather allow it to float.

Rather than just state that the resulting fit is poor or good by eye, we tried to quantify the difference by using the Kolmogorov-Smirnov statistic. We first performed a binned KS test but people were unhappy with this approach. It is misleading (skewed) at large KS values. Instead, we follow a suggestion from Igor Volobouev. Precisely, we fit the data to the MC and determine the Kolmogorov distance D_d as the maximum distance between the integral distributions of the fitted MC shape and the data. We then generate pseudo-experiments according to this MC fit. We determine the Kolmogorov distance D_MC between each random sample and the MC parent. We keep track of whether the D_MC is larger or smaller than D_d for the 1000 pseudo-experiments performed. The probability that the data agree with the MC is equal to the number of times that D_d< D_MC.

Table 9 lists the results of the Kolmogorov distance measurement described above in the W+1, 2 and 3 jet regions. Figures 9, 10, 11 show the distributions for lepton $E_T$, $\not\!\!E_T$, leading jet $E_T$ and $H_T$. Based on both the plots and the table, we see that the modeling of the variables is adequate in the 3 jet region but not in the 1 or 2 jet regions. Surprisingly, the least well modeled variable is the lepton $E_T$.

The suspicion is not that the jet energy scale is bad or that the higher statistics in the 1 and 2 jet bin make discrepancies more visible, but that the feed-down from the W+2 parton and W+3 parton MC significantly alters the lepton $E_T$ distribution from the W+1 parton MC in the W+1jet region [6]. Recently, MC has become available that allows one to simply add up the contributions from different W+n parton samples to obtain a complete W+n jet spectrum [7]. Note that a matching scheme from Michelangelo Mangano also does this but we haven't tried that yet. We have this new MC for $W \to e^+\nu$ events only, so we have simply compared our $W \to e^+\nu$ data to this MC in Figures 12 and 13 for the 1 and 2 jet regions. Table 10 lists the results of the Kolmogorov distance measurement described above in the W+1 and 2 jet regions. The agreement is now adequate for the lepton $E_T$ and $H_T$ but still bad for $\not\!\!E_T$ and leading jet $E_T$- perhaps future inclusion of small background contributions from $W \to \tau\nu$, Z and WW, WZ, ZZ will improve the agreement?

| Variable | **Probability of Fit (%)** | | |
|---|---|---|---|
| | W+1jet | W+2jet | W+3jet |
| $H_T$ | 0.0 | 0.9 | 42 |
| **lepton** $E_T$ | 0.0 | 11 | 93 |
| $\not\!\!E_T$ | 0.0 | 16 | 81 |
| **Leading Jet** $E_T$ | 0.0 | 5 | 70 |
| **Sum Jet** $E_T$ | - | 0.8 | 78 |

Table 9: Probability of W+1,2,3 jet data fits as determined using the Kolmogorov distance described in the text.

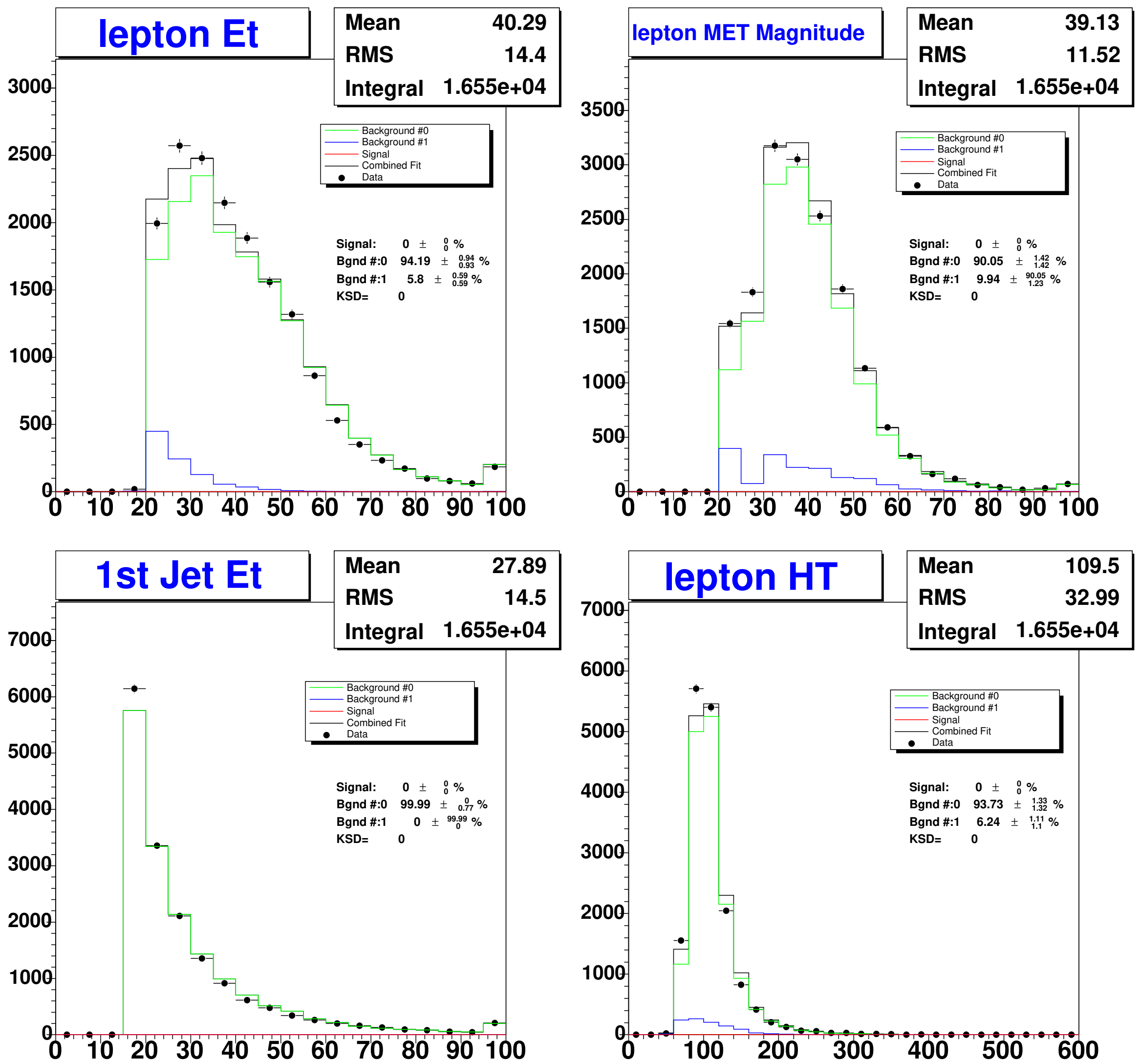



Figure 9: Monte Carlo to Data comparison of the lepton $E_T$, MET, leading jet $E_T$ and $H_T$ variables in the W+1 jet bin. The combined fit (black curve) of W+1 parton MC (green curve) and non-isolated leptons for the QCD Fake background (blue curve) is compared to the data (points).

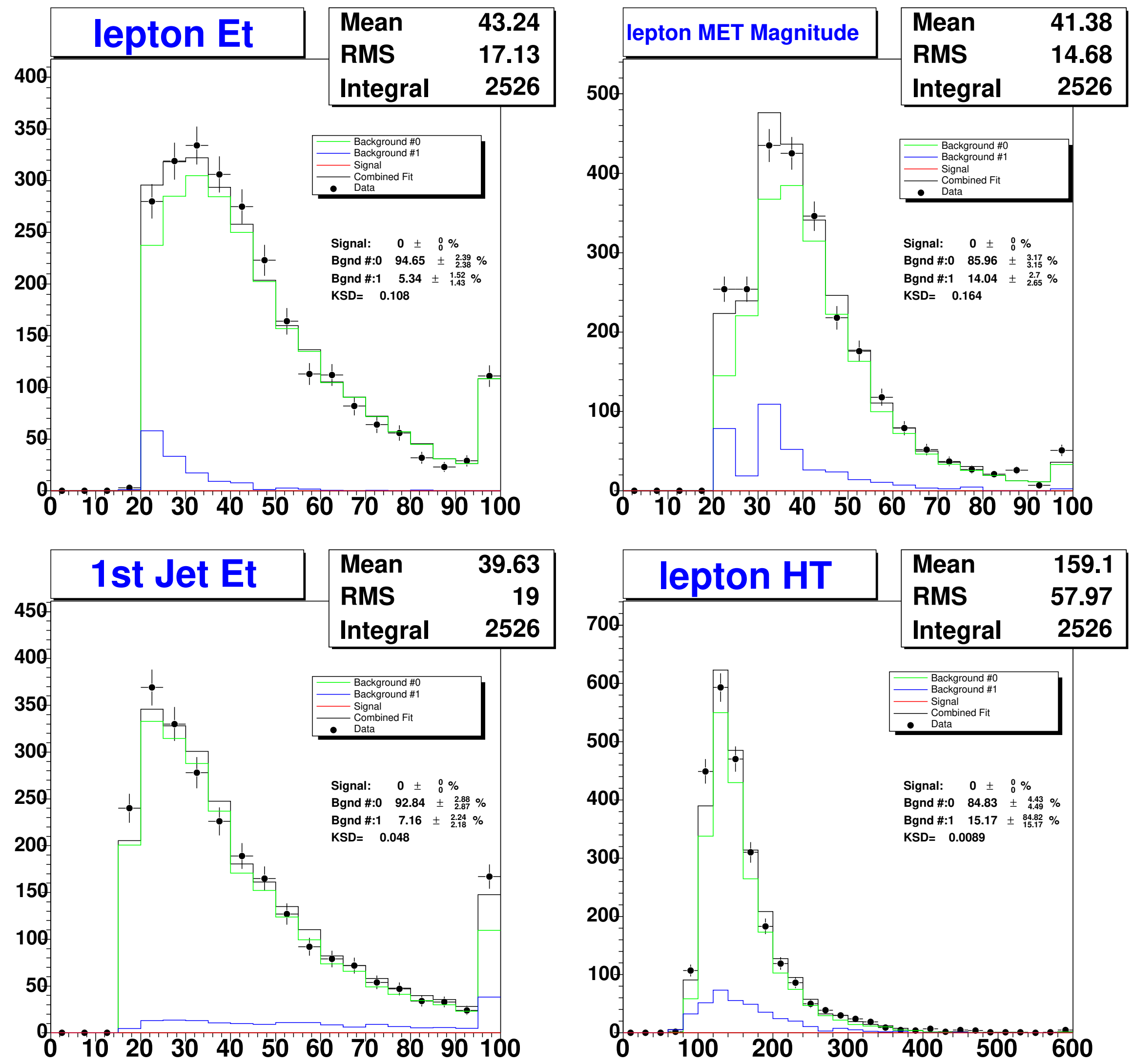



Figure 10: Monte Carlo to Data comparison of the lepton $E_T$, MET, leading jet $E_T$ and $H_T$ variables in the W+2 jet bin. The combined fit (black curve) of W+2 parton MC (green curve) and non-isolated leptons for the QCD Fake background (blue curve) is compared to the data (points).

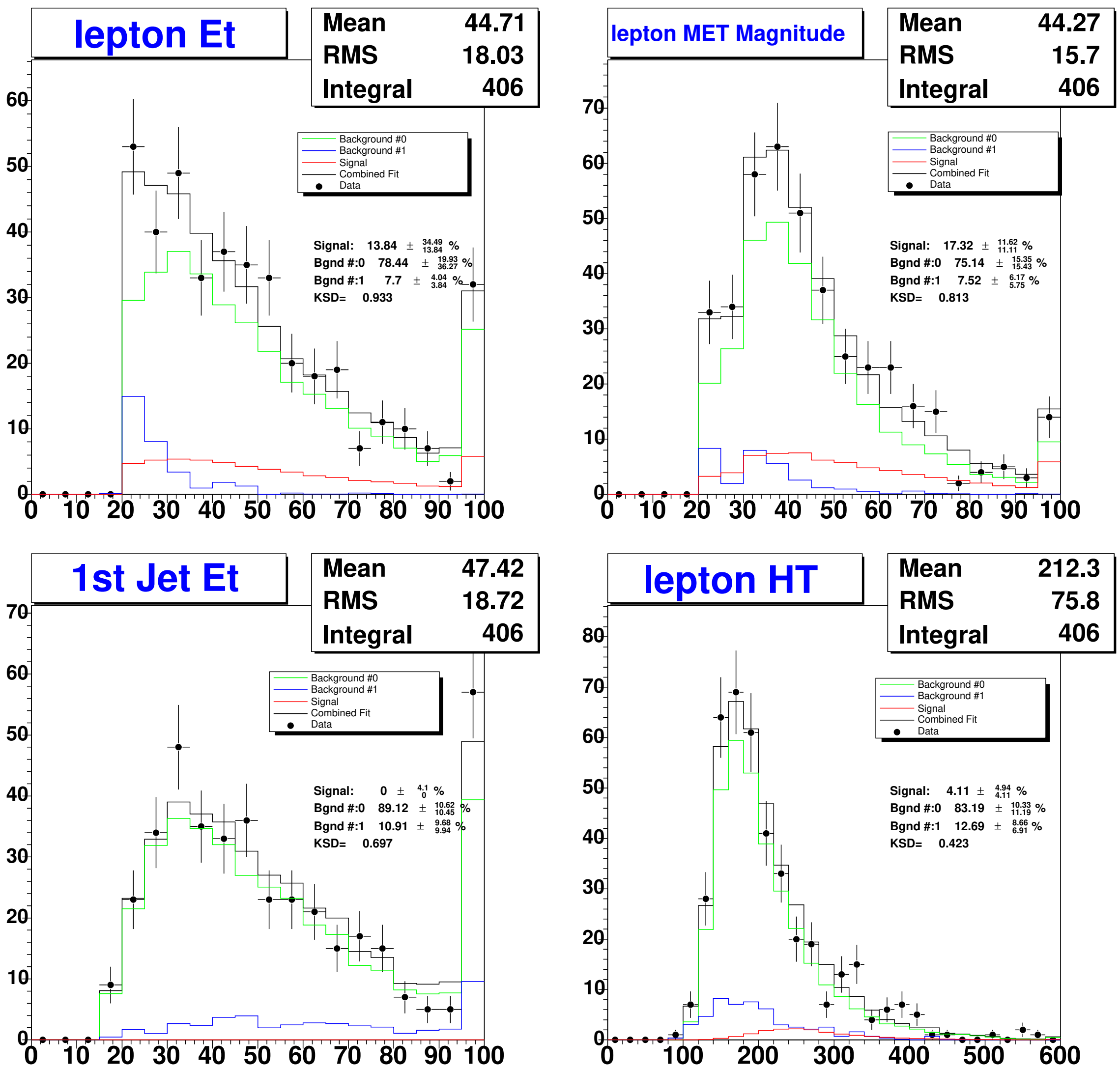



Figure 11: Monte Carlo to Data comparison of the lepton $E_T$, MET, leading jet $E_T$ and $H_T$ variables in the W+3 jet bin. The combined fit (black curve) of W+3 parton MC (green curve), $t\bar{t}$(red curve) and non-isolated leptons for the QCD Fake background (blue curve) is compared to the data (points).

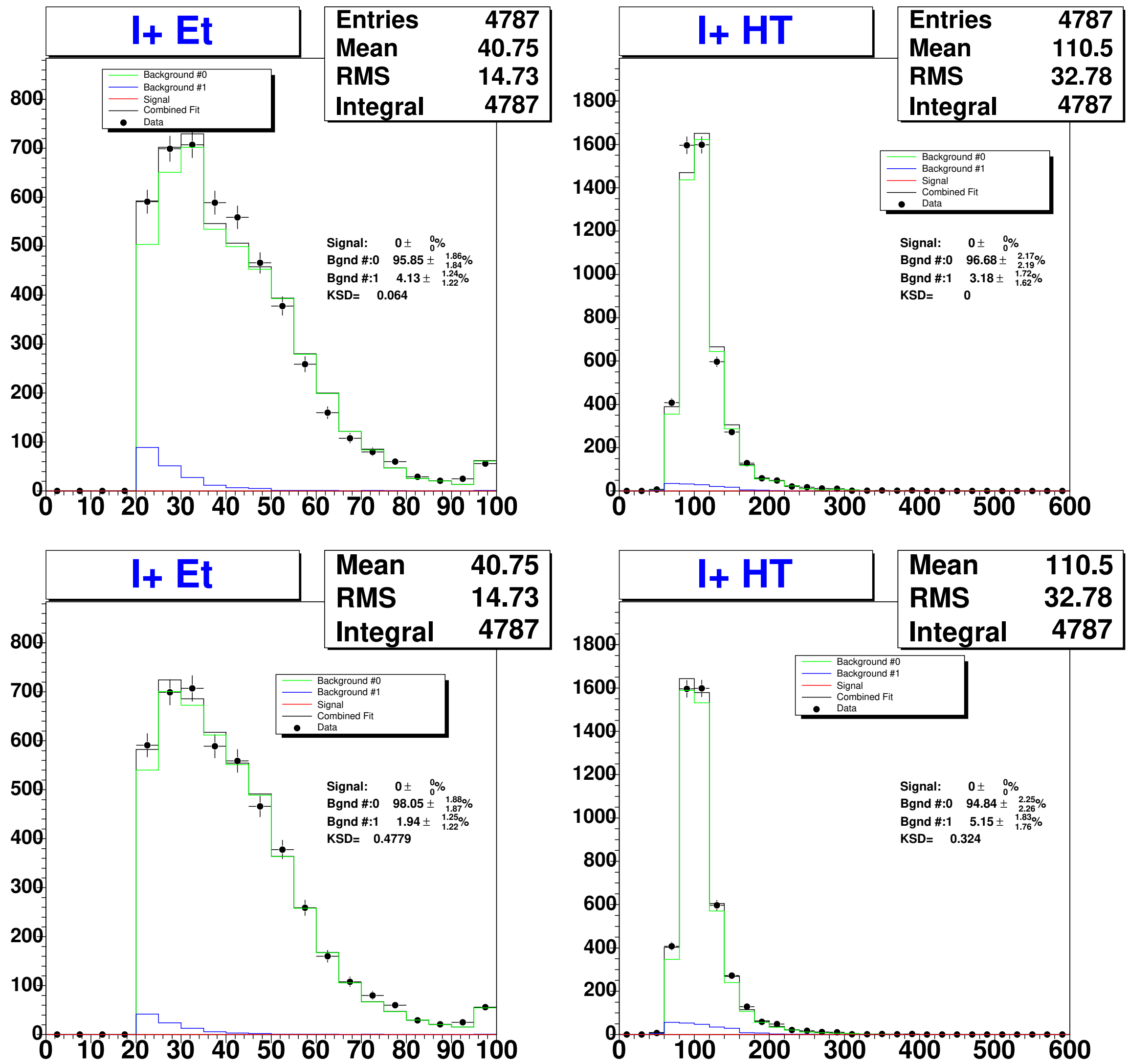



Figure 12: Monte Carlo to Data comparison of the lepton $E_T$ and $H_T$ variables in the $W^+ \to e^+\nu$+1 jet bin to W+1p MC (top) and $\Sigma_{n=0}^{4}$W+np MC (bottom). The combined fit (black curve) of W+jet MC (green curve) and non-isolated leptons for the QCD Fake background (blue curve) is compared to the data (points).

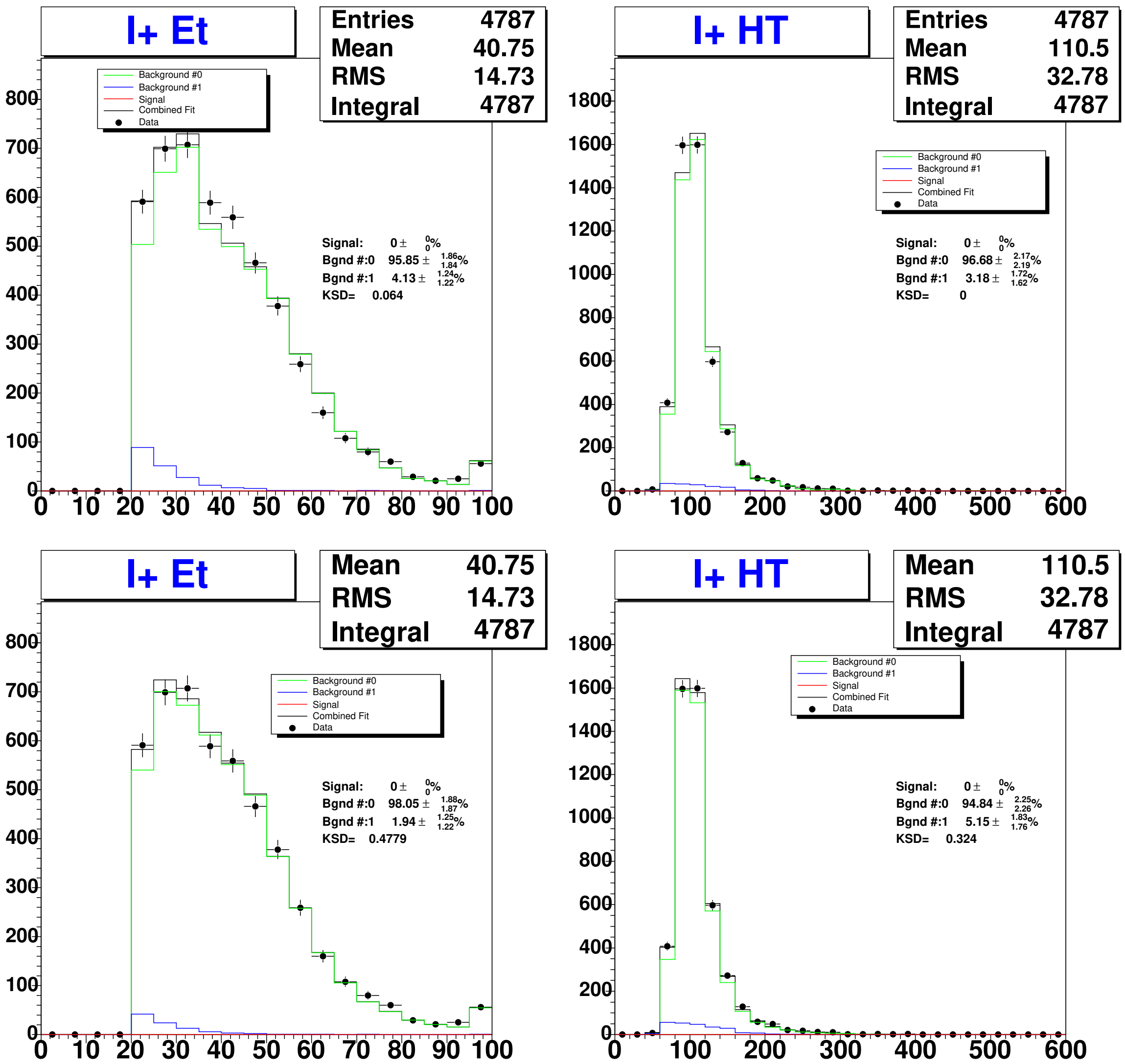



Figure 13: Monte Carlo to Data comparison of the lepton $E_T$ and $H_T$ variables in the $W^+ \rightarrow e^+\nu$+2 jet bin to W+2p MC (top) and $\Sigma_{n=0}^{4}$W+np MC (bottom). The combined fit (black curve) of W+jet MC (green curve) and non-isolated leptons for the QCD Fake background (blue curve) is compared to the data (points).

| Variable | Probability of Fit (%) | | | |
|---|---|---|---|---|
| | W+1jet | | W+2jet | |
| | $\Sigma_{n=0}^{4}$W+np | W+1p MC | $\Sigma_{n=0}^{4}$W+np | W+2p MC |
| $H_T$ | 39 | 0 | 62 | 3 |
| **lepton** $E_T$ | 48 | 6 | 53 | 23 |
| $\not{E}_T$ | 0 | 3 | 9 | 4 |
| **Leading Jet** $E_T$ | 0 | 0 | 18 | 28 |
| **Sum Jet** $E_T$ | 0 | 0 | 36 | 7 |

Table 10: Probability of $W^+ \rightarrow e^+\nu$+1,2 jet data fits to W+np MC and to $\Sigma_{n=0}^{4}$W+np MC as determined using the Kolmogorov distance described in the text.

### 10.1 Conclusions of MC and Data Comparisons

In summary, our MC background shapes model our data in the W+1, 2 and 3 jet bins quite well across a wide range of kinematic and shape variables. Accounting for the feed-down from the W+2p and W+3p samples significantly improves the agreement in the W+1jet region and likewise in the W+2jet region for the lepton $E_T$and $H_T$. However, the W+3 parton sample along with the non-isolated lepton sample agree with the W+3jet data sufficiently well, within our current level of statistical precision, that they can be used to determine a signal fraction in the $\geq 3$ and $\geq 4$ jet bins.

## 11 Results for $t\bar{t}$ fit fraction

Based on pseudo-experiments, we expect our best measurement of the signal to come from the W+$\geq$3 jet bin. To check for consistency, we also fit the W+$\geq$4 jet bin which has a better signal/noise ratio but fewer statistics. For the QCD background, we have chosen the non-isolated lepton sample. From the background section, the amount of QCD in our fits is set to 6.3%. The W+jets background is taken from ALPGEN W+3 and W+4 parton MC(for the two separate signal regions). PYTHIA $t\bar{t}$is used to model the signal region. For our final answer, we will use $H_T$ but we will investigate the behavior of the fit fraction for the other kinematic variables as well.

Figures 14 and 16 show the resulting $t\bar{t}$ fit fraction in the W+$\geq$3 and W+$\geq$4 jet bins respectively for four separate kinematic and shape variables. The $\Delta\phi$ cut is used and the QCD shape is modeled using non-isolated leptons. A summary of the fit fraction results in both signal bins can be found in figures 15 and 17. In all of these plots, it is clear that the data prefers a substantial fraction of $t\bar{t}$, in both the W+$\geq$3 and W+$\geq$4 samples. In addition, there is general agreement within each sample for the approximate fraction. The W+$\geq$3 sample prefers about 15% $t\bar{t}$, while the W+$\geq$4 sample prefers about 50%. Using the $H_T$ variable, the extracted $t\bar{t}$ fit fraction in the W+$\geq$3 jet bin is 0.126 $\pm$ 0.042(stat). In the W+$\geq$4 jet bin, the resulting fit fraction is 0.504 $\pm$ 0.130(stat).

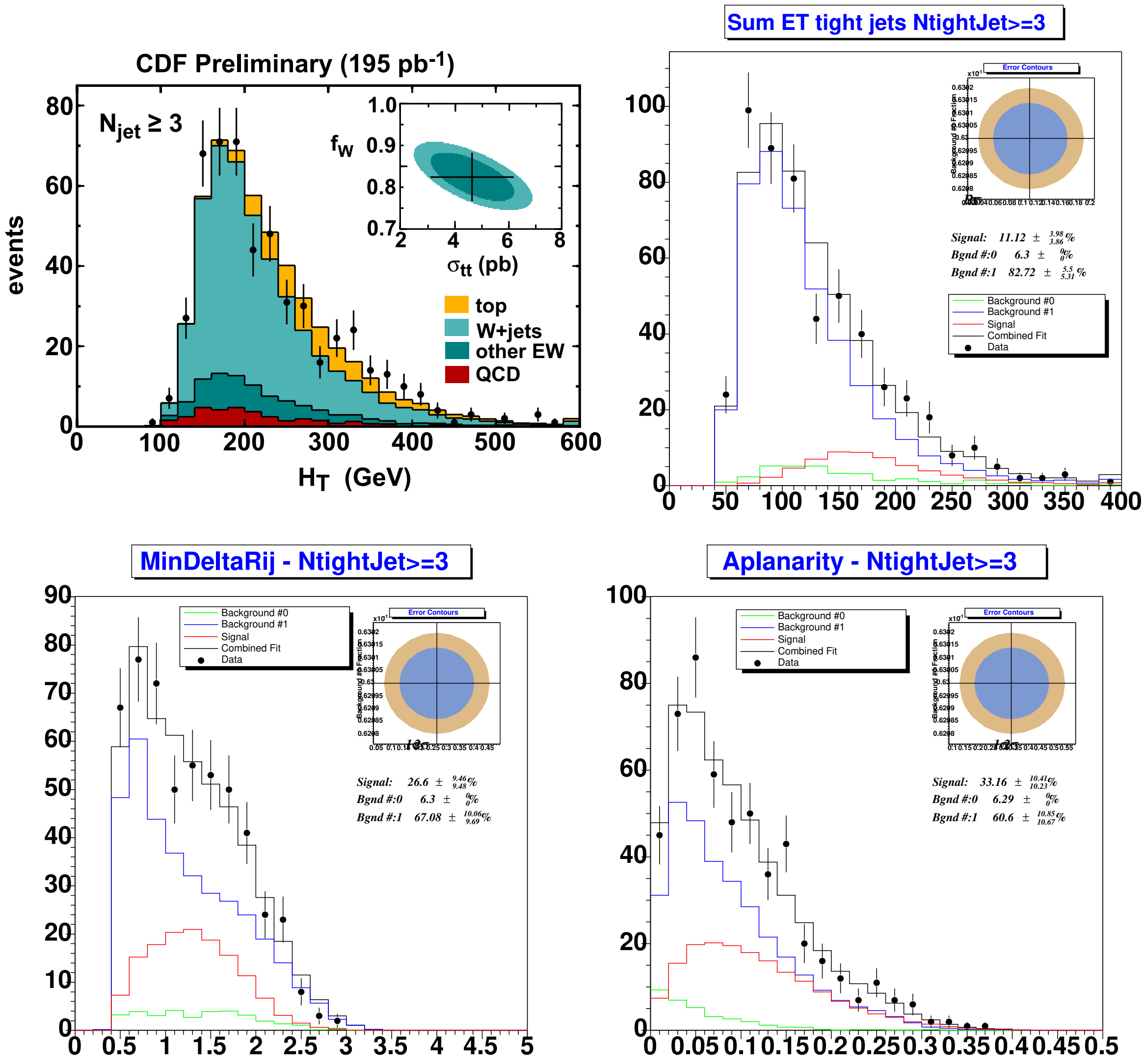



Figure 14: Fits of the data to W+jets and QCD background templates for a range of kinematic and shape variables in the W+$\geq$ 3 jet bin

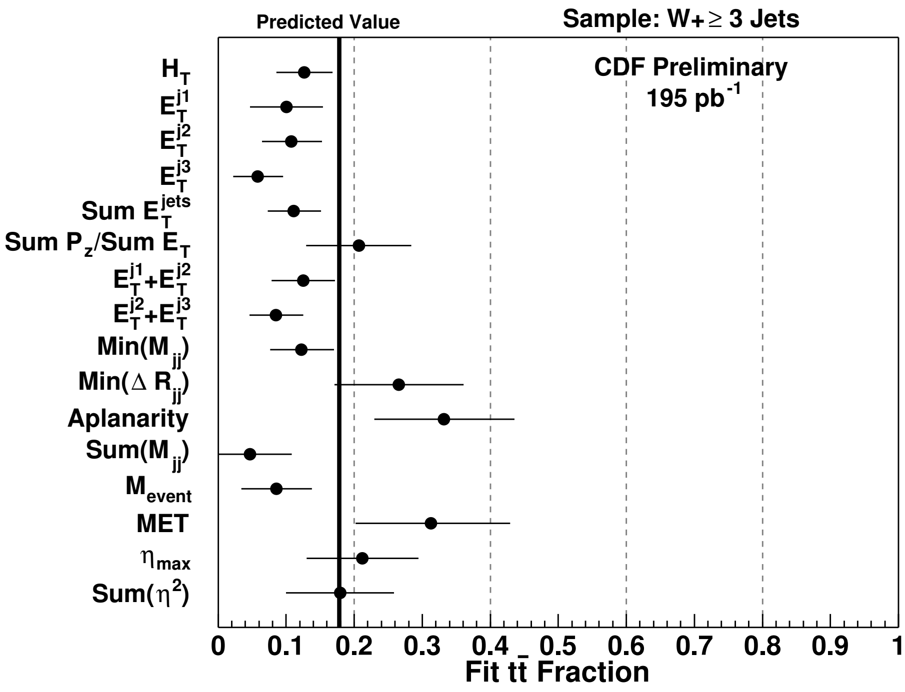



Figure 15: Summary plot of the fit fraction of $t\bar{t}$ results for a variety of kinematic and shape variables in the W+$\geq$3 jet bin. Non-isolated lepton sample was used as the QCD model and the $\Delta\phi$ cut was applied.

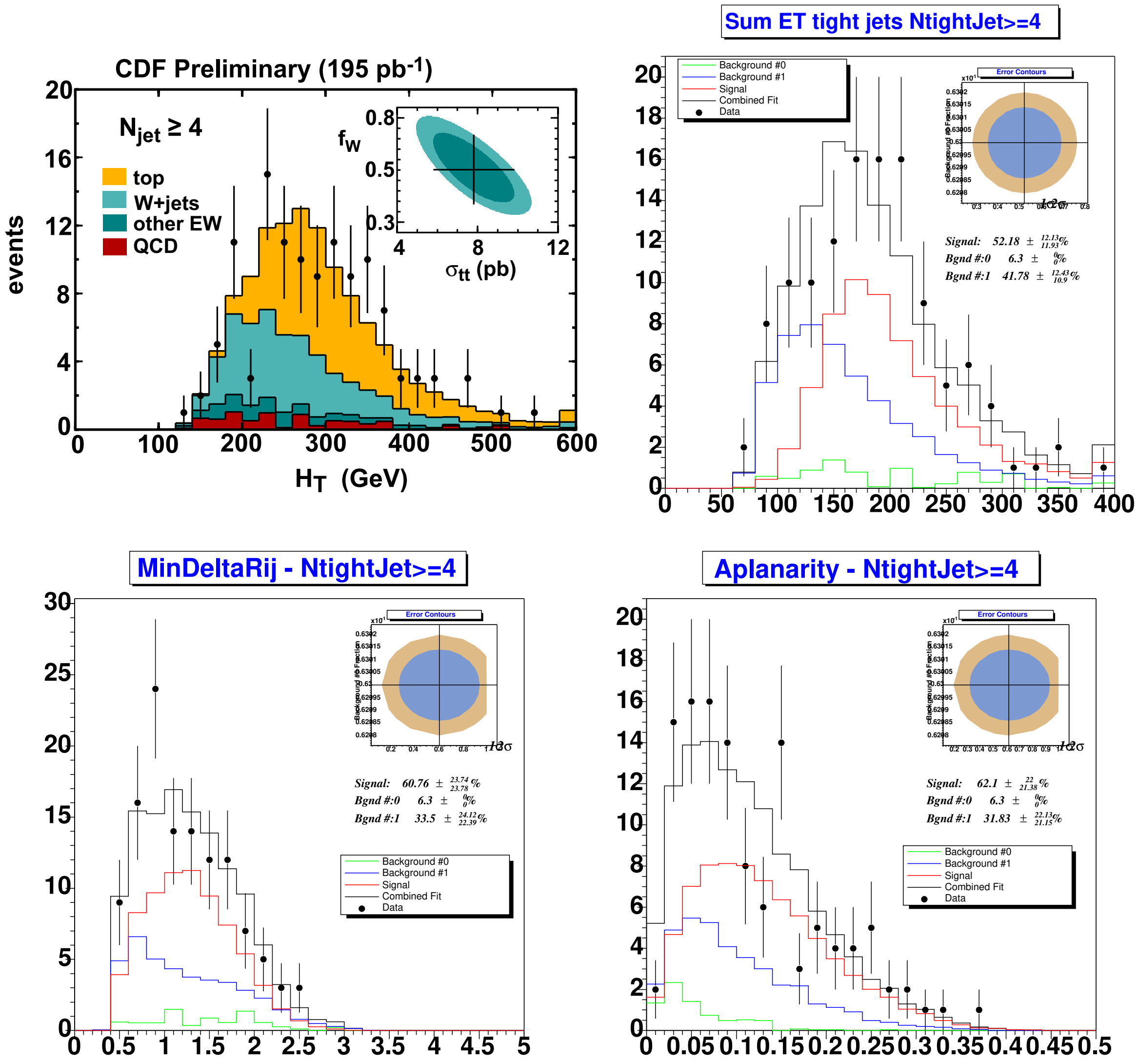



Figure 16: Fits of the data to W+jets and QCD background templates for a range of kinematic and shape variables in the W+$\geq$ 4 jet bin.

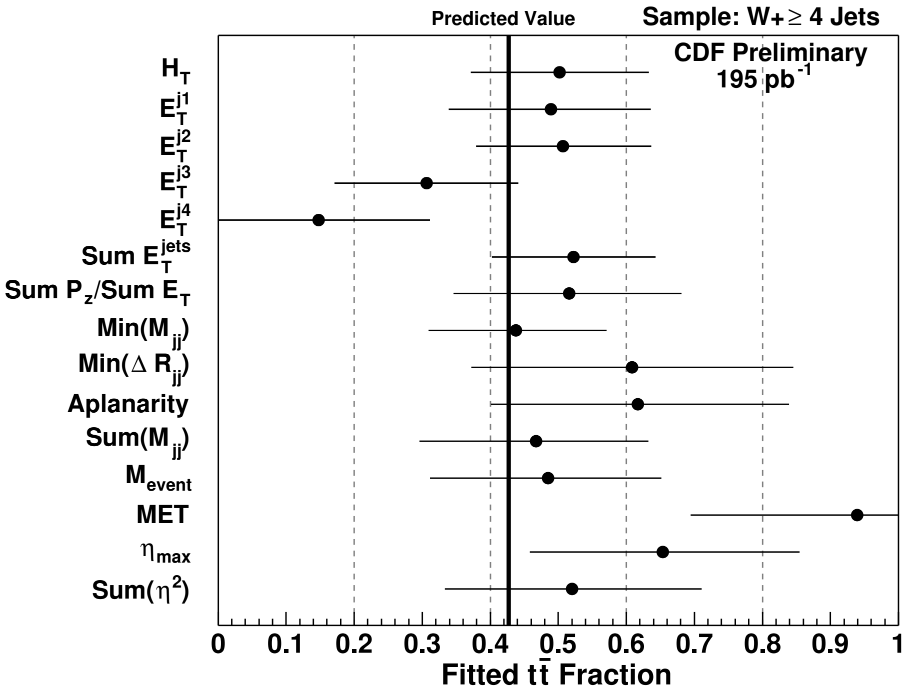



Figure 17: Summary plot of the fit fraction of $t\bar{t}$ results for a variety of kinematic and shape variables in the W+$\geq$4 jet bin. Non-isolated lepton sample was used as the QCD model and the $\Delta\phi$ cut was applied.

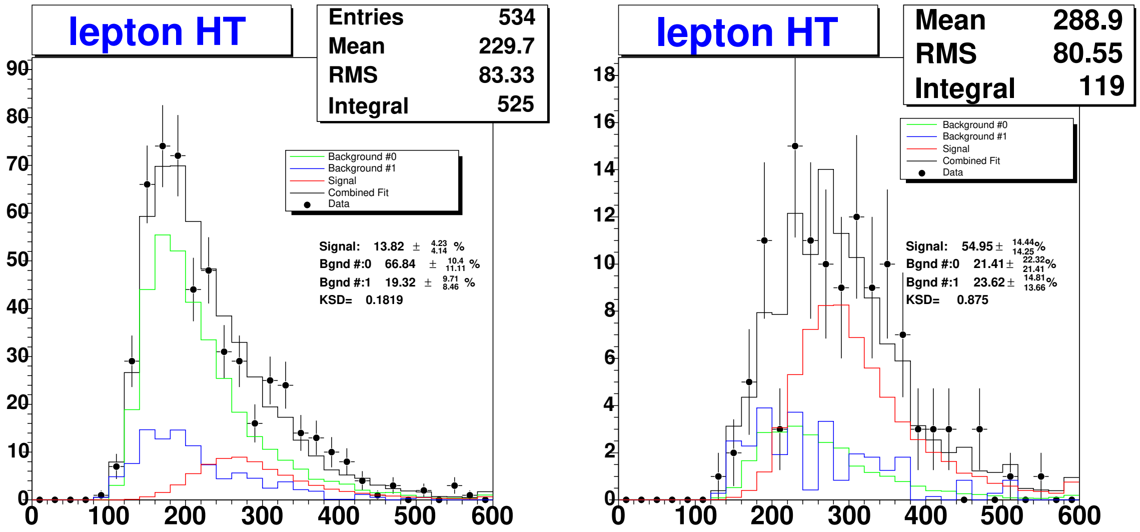



Figure 18: $H_T$ fits allowing the QCD background fraction to float in the W+$\geq$ 3(left) and $\geq$ 4(right) jet bins respectively. No small backgrounds were included in the W+jets MC shape.

## 11.1 Cross Checks

Given the limited statistics available for this analysis, it is difficult to be certain that we are modeling the QCD shape properly. Monte Carlo's offer little guidance in this fake background and extracting information from the data is tenuous. Therefore, even after all the other checks have been done on the QCD shape and agreement in the W+1, 2 and 3 jet bins, we perform fits where we allow the QCD background fraction to float to whatever value the fit prefers. The results of the fit to $H_T$ are shown in Figure 18

A second check that we can perform is to fit our signal region with and without $t\bar{t}$ and then look at the KS test, as described earlier to see how well we can do. This gives us two pieces of information. It tells us how consistent our nominal fit is with the data and it tells us whether the fit really prefers top or not. The results of the fit to $H_T$ are shown in Figure 19

The results of both cross-checks are summarised in Table 11. If QCD is allowed to float, the KS test is slightly better in the 3 jet and 4 jet bins. Letting QCD float has no appreciable impact on the $t\bar{t}$ fraction that we measure. Any fit to the data WITHOUT including $t\bar{t}$ yields a KS test of less than 1% and shows that the fit is incompatible with the data.

| $t\bar{t}$ | QCD | W+$\geq$3jet | | W+$\geq$4jet | |
|---|---|---|---|---|---|
| | | Signal Fraction (%) | KS Prob (%) | Signal Fraction (%) | KS Prob (%) |
| With | Fixed | 13.2 | 16 | 49.45 | 51 |
| With | Float | 13.82 | 18 | 54.95 | 88 |
| Without | Fixed | - | 0.3 | - | 0.0 |
| Without | Flaot | - | 0.1 | - | 0.3 |

Table 11: Signal fraction and KS probability for fits (without small backgrounds) with QCD background fraction fixed vs floating and with and without a $t\bar{t}$ component.

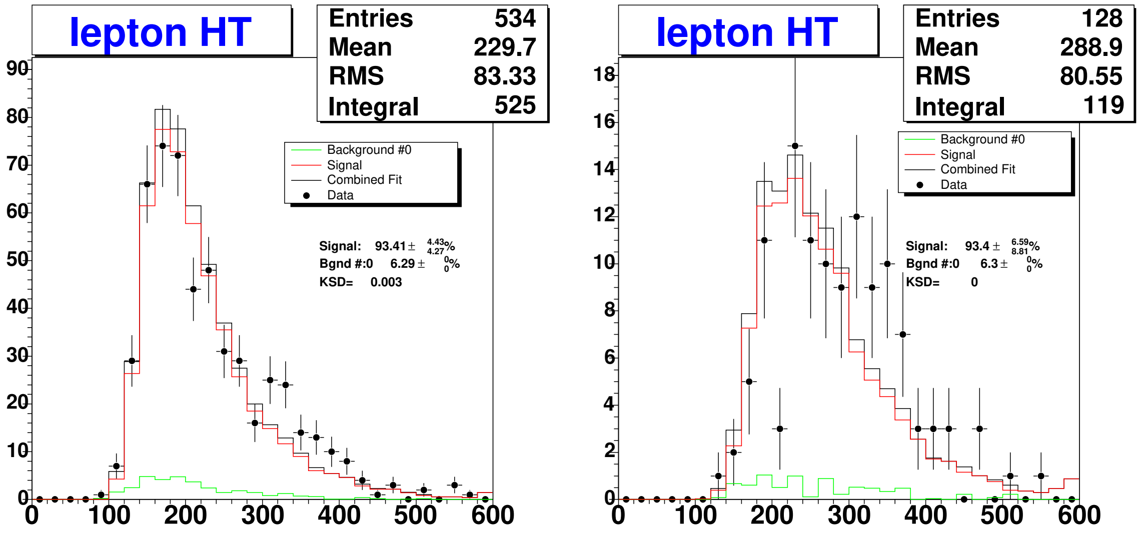



Figure 19: $H_T$ fits without any $t\bar{t}$ component and with fixed QCD background fraction in the W+$\geq$ 3(left) and $\geq$ 4(right) jet bins respectively. No small backgrounds were included in the W+jets MC shape.

# 12 Cross Section Results

The $t\bar{t}$ cross section is calculated based on the following formula.

$$\sigma(t\bar{t}) = \frac{\mathcal{F} \cdot N}{\mathcal{L} \cdot A}$$

Where $\mathcal{F}$ is the $t\bar{t}$ fit fraction, $N$ is the number of observed events in the data, $\mathcal{L}$ is the integrated luminosity, and $A$ is the acceptance.

The acceptance can be broken down into the following form:

$$\text{Acceptance} = [\varepsilon_{cmx}(t\bar{t}) \cdot \mathcal{L}_{cmx} + \varepsilon_{cmu}(t\bar{t}) \cdot \mathcal{L}_{cmu} + \varepsilon_{cem}(t\bar{t}) \cdot \mathcal{L}_{cem}] \cdot \varepsilon_{common} \cdot \varepsilon_{d\phi} \cdot \varepsilon_{jet}$$

The values for $\varepsilon_{cmx}(t\bar{t})$, $\varepsilon_{cmu}(t\bar{t})$, $\varepsilon_{cem}(t\bar{t})$, and $\varepsilon_{common}$ come from CDF Note 6574 [10] for the summer conference dataset. These numbers will be updated when the trigger and lepton identification and reconstruction efficiencies for the complete data sample are finalised at a Joint Physics Working Group meeting in early January 2004. The numbers for each of the quantities in the equations above are listed for reference in Table 12.

Putting all of the numbers together, the final results for the cross section (with statistical errors only) can be found in table 13.

## 12.1 Results as a function of assumed top mass

The result above assumed a top mass of 175GeV/c$^2$. Figure 20 shows our result for 3 different top masses of 170, 175 and 180 GeV/c$^2$. Table 14 gives the numerical results. Note these results use a HERWIG $t\bar{t}$ sample rather than the PYTHIA $t\bar{t}$ sample used for our central result.

| **Variable** | **Value** |
|---|---|
| $\mathcal{L}_{cmx}$ | $193.5 \pm 11.4$ pb$^{-1}$ |
| $\mathcal{L}_{cmu,cem}$ | $175.3 \pm 10.1$ pb$^{-1}$ |
| $\varepsilon_{d\phi}$ | $0.947 \pm 0.001$ |
| $\varepsilon_{jet}$**(n≥3)** | 1.0 |
| $\varepsilon_{jet}$**(n≥4)** | $0.554 \pm 0.002$ |
| $\varepsilon_{cem}(t\bar{t})$ | $0.0465 \pm 0.0009$ |
| $\varepsilon_{cmu}(t\bar{t})$ | $0.0232 \pm 0.0006$ |
| $\varepsilon_{cmx}(t\bar{t})$ | $0.0107 \pm 0.0003$ |
| $\varepsilon_{common}$ | $0.951 \pm 0.062$ |

Table 12: Acceptance and luminosity numbers for this cross section measurement.

| **Sample** | **Observed Evts** | $t\bar{t}$ **fraction** | $\sigma(t\bar{t})$ |
|---|---|---|---|
| **W+≥3 Jets** | 525 | $0.126 \pm 0.042$ | $4.7 \pm 1.6$ pb |
| **W+≥4 Jets** | 119 | $0.504 \pm 0.130$ | $7.8 \pm 2.0$ pb |

Table 13: The $t\bar{t}$ production cross section in the W+$\geq$ 3 and W+$\geq$4 jet bins as well as the fraction of $t\bar{t}$ events found in our data.

# 13 Systematics

Inspection of the above cross section formula shows that we will be sensitive to systematic effects from the following sources:

- The uncertainty on the luminosity $\mathcal{L}$ of 5.9%[4].
- Systematic effects on the acceptance $A$. These fall into sub-categories of those that affect the lepton ID, those that affect the efficiency for passing the MET and Jet $E_T$ cuts.
- Systematic effects on the determination of the $t\bar{t}$ fraction $\mathcal{F}$ in the fitting of the $H_T$ distribution. These are shape systematics. These include effects that the modify the $H_T$ distribution from $t\bar{t}$ and W+jets. In addition, there are systematics associated with our background model. We use non-isolated leptons as our primary non-W background model, and we check our sensitivity to this by using the conversion electron sample to model the non-W background.

| Generated Top Mass | Measured $t\bar{t}$ $\sigma$ (pb) | |
|---|---|---|
| | W+≥3jets | W+≥4jets |
| 170 | 4.88± 1.61 | 8.35± 2.09 |
| 175 | 4.62± 1.48 | 7.79± 1.88 |
| 180 | 4.30± 1.35 | 6.52± 1.66 |

Table 14: The $t\bar{t}$ production cross section in the W+$\geq$ 3 jets and W+$\geq$4 jets regions for fits to HERWIG $t\bar{t}$ MC generated at three different top masses. The error is statistical only.

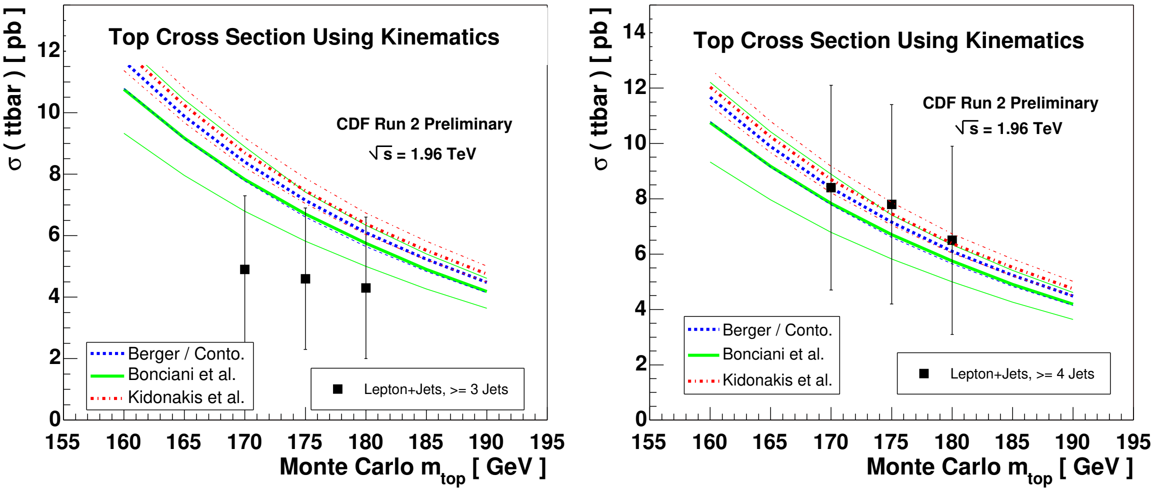



Figure 20: The $t\bar{t}$ production cross section in the W+$\geq$ 3 jets (left) and W+$\geq$4 jets (right) regions for fits to HERWIG $t\bar{t}$ MC generated at three different top masses.

Note that some effects will show up in more than one category. For example, the uncertainty in the Jet energy scale will lead to both a change in the shape of the $H_T$ distribution, as well as a change in the $t\bar{t}$ acceptance (as more or fewer jets will pass the Jet $E_T$ cut).

## 13.1 Determination of shape systematics

To determine the shape-based systematics, we use a pseudo-experiment (PE) technique. We will use the determination of the systematic due to the uncertainty of the top quark mass as an example.

### 13.1.1 Shape systematic from variation of the assumed top quark mass

In fitting our data, we assume that the top quark mass is 175 $GeV/c^2$. Our default model for fitting the $W+\geq 3$ jet sample then includes 3 samples:

- A PYTHIA $t\bar{t}$ sample, with the top quark mass set at 175 $GeV/c^2$. PYTHIA was used rather than the standard HERWIG since we had samples of PYTHIA at different top quark masses readily available. As is shown later, there is very little difference in $H_T$ shape between PYTHIA and HERWIG.
- An ALPGEN $W+3$ parton sample.
- A non-isolated sample, obtained from the data.

However, since we only know the top mass to within about 5 $GeV/c^2$, it is entirely possible that the true top quark mass is really 170 $GeV/c^2$, or 180 $GeV/c^2$. The systematic is determined using the following procedure:

1. We generate PE's using a PYTHIA 170 (or 180) $GeV/c^2$ $t\bar{t}$ sample, the standard ALPGEN $W+3$ parton sample, and the data non-isolated sample.
2. We fit each PE twice:

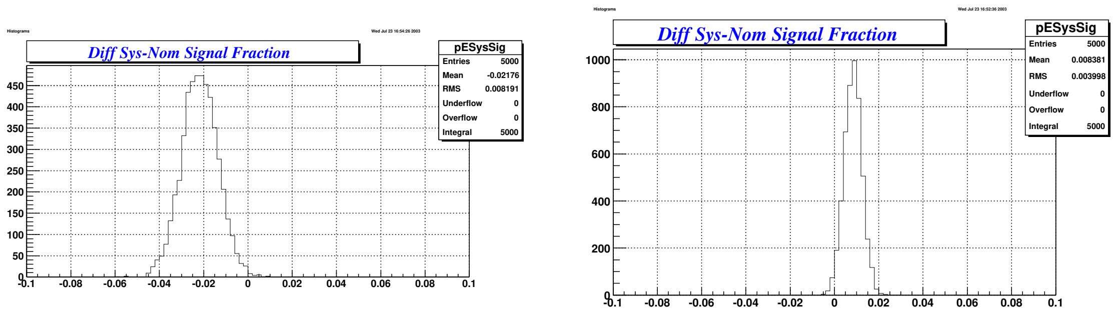



Figure 21: These plots show the difference in the fitted fraction between using the standard templates (with a top mass of $175 GeV/c^2$, and using the generated templates with a top mass of $170 GeV/c^2$ (left) and $180 GeV/c^2$ (right).

- Using the SAME model as applied to the data (that is the 175 $GeV/c^2$ $t\bar{t}$ sample, the standard ALPGEN $W + 3$ parton sample, and the data non-isolated sample).
- Using the APPROPRIATE model for this PE (that is the 170 $GeV/c^2$ $t\bar{t}$ sample, the standard ALPGEN $W + 3$ parton sample, and the data non-isolated sample).

3. We take the difference of the two fits, and plot the result for many PE's. The systematic is the mean of this distribution.

The plots of this difference are shown in Figure 21. The plot on the left shows the difference in the fitted fraction between using the standard templates (with a top mass of $175 GeV/c^2$, and using the generated templates with a top mass of $170 GeV/c^2$. The $t\bar{t}$ fraction used in the pseudoexperiments was 0.19, so the mean of -0.022 corresponds to a -11.6% systematic on the fraction. The plot on the right shows the difference in the fitted fraction between using the standard templates (with a top mass of $175 GeV/c^2$, and using the generated templates with a top mass of $180 GeV/c^2$. Using this higher top mass yields a mean difference of +0.0084, which corresponds to a -4.4% systematic on the fraction. We symmetrize this result and obtain a systematic of ±8% on the fit fraction due to varying the top quark mass.

#### 13.1.2 Shape systematic from variation of the $E_T$ scale

Our dependence on the $E_T$ scale is significant, since this will come into play in both the jet $E_T$'s, as well as in the determination of the MET. Jet $E_T$'s are corrected using Level 4, and the method used is provided by the JetUser package (Version3_4_NewSyst). To determine the systematics associated with the Jet corrections, we use a procedure outlined on the Jet Corrections web page (along with a code snippet), which can be found on:

http://www-cdf.fnal.gov/internal/physics/top/jets/jetcorr-systematics.html

In this procedure, only the Monte Carlo Jet $E_T$'s are shifted due to the systematic uncertainty on the $E_T$ scale. In our case, this means we generate what we refer to as "shifted templates" for both $t\bar{t}$ and $W+ \geq 3$ jets Monte Carlo. There will be shifted templates for both an upper systematic uncertainty shift, as well as a separate set for a lower systematic uncertainty shift. We then generate a fake dataset from PE's using the upper shifted MC, and fit both the upper shifted dataset (plus QCD shape from the data) and the default (unshifted) dataset with the nominal background templates. The difference is the systematic effect. For the upper systematic, we find a shift in the fitted fraction of +30% (meaning we systematically find 30% more $t\bar{t}$ if the $E_T$ is higher than our default). For the lower systematic, we find a shift in the fitted fraction of -26% (meaning we systematically find 26% less $t\bar{t}$ if the $E_T$ scale is lower than our default). We symmetrize this result and obtain a systematic of ±28% on the fit fraction due to varying the Jet $E_T$ scale. This systematic increases if we use the $W+\geq 4$ sample to measure the cross section. In this case, we obtain a systematic of ±33%.

Since this systematic is the primary uncertainty for this analysis, we tried to break it down and understand which piece(s) of the jet energy scale process dominates. To do this, we used the above outlined process and applied each correction (syscode) separately. [Note that for the data correction errors, the MC is still used but the jet corrections and associated shifts one would use on *data* are applied to the corrected (flattened) MC.] The $t\bar{t}$ fraction is then measured from the new PE and compared to the nominal value of 0.2042. The answer was determined separately for a shift up and a shift down. What is reported in Table 15 is just the average shift. That of course is not the entire story. We need to add in the acceptance systematic as well since events move in and out of our signal bin as a result of this energy scale change. The column labeled "Shape+Accept" is a straight addition of these effects. The bottom line by adding all of the "Shape+Accept" column in quadrature is that the overall systematic is 30.2%. The dominant component is the scale correction followed in close second with the relative correction, absolute correction and out of cone.

| **Data/MC** | **Sys Code** | **Shift** | **Shape+Accept** |
|---|---|---|---|
| Data | 1(relative) | 0.01838 | 0.09882 |
| Data | 2(cal stab) | 0.00930 | 0.04942 |
| Data | 3(scale) | 0.02810 | 0.14113 |
| MC | 1(relative) | 0.02109 | 0.12059 |
| MC | 2(cal stab) | 0.0 | 0.0 |
| MC | 3(Scale Cor) | 0.03290 | 0.18058 |
| MC | 4(mult int) | 0.0 | 0.0 |
| MC | 5(absolute) | 0.02151 | 0.12011 |
| MC | 7(out of cone) | 0.02579 | 0.04592 |
| MC | 8(splash out) | 0.00987 | 0.07256 |

Table 15: List of individual contributions to the jet energy scale systematic. The $t\bar{t}$ input fraction in the pseudo-experiments is 0.2042. The shifts quoted are all relative to this input fraction. MC Syscodes 0-4 are determined using Level 4 corrections. MC syscodes 5-8 use Level 7 corrections. In the "Shape + Acceptance" column, this takes into account the acceptance systematic as a result of the shift as a straight add to the "shift" column.

#### 13.1.3 Shape systematic from ISR and PDF's

To determine this systematic, we break it up into two pieces; one which handles the ISR shape and the second piece to determine the PDF shape systematic. The details follow.

- The model for initial state radiation (ISR) used in the generation of the top quark can also have an impact on the shape of the resulting $H_T$ distribution. To study this, we use the PYTHIA Monte Carlo, with ISR turned on (the default) and ISR turned off. Just as described above, we generate PE's using the ISR-off Monte Carlo, but fit with the default (in this case PYTHIA with ISR on) templates. The systematic is taken to be half of the result, or 0.56%.

- In order to obtain a systematic for the PDF, we followed the prescription of the top mass group. The cross section systematic for this PDF is measured using Herwig MC samples with three of the 20 fit parameters in CTEQ6M PDF varied within their uncertainties ($\pm 1\sigma$). These three parameters were chosen because their corresponding uncertainties dominate the uncertainties in the dijet cross section at high $Q^2$. To be specific, the parameters chosen were eigenvectors 10, 15 and 19. Eigenvector 15 for instance controls the amount of glue-glue in ISR. Using pseudo-experiments, we determine a cross section after varying each parameter by +1 and -1 $\sigma$. We divide the difference by two and then add them in quadrature. The results are shown in Table 16.

  The final systematic for PDF in $\geq$ 3 jets is 0.24 pb and in the $\geq$ 4 jets of 0.30 pb. In terms of top fraction, the PDF systematic for W+$\geq$3, $\geq$4 jets translates into 3.5% and 4.3% respectively.

| **Eigenvector** | **njet** | **-1** $\sigma$ | **+1** $\sigma$ | $\Delta\sigma/2$ |
|---|---|---|---|---|
| 15 | $\geq 3$ | 7.28 $\pm$ 0.02pb | 7.39 $\pm$ 0.02pb | 0.06 pb |
| 10 | $\geq 3$ | 7.36 pb | 7.15 pb | 0.11 pb |
| 19 | $\geq 3$ | 7.39 pb | 7.00 pb | 0.20 pb |
| 15 | $\geq 4$ | 6.86 $\pm$ 0.04pb | 7.04 $\pm$0.04pb | 0.09 pb |
| 10 | $\geq 4$ | 6.99 pb | 6.69 pb | 0.15 pb |
| 19 | $\geq 4$ | 6.80 pb | 6.32 pb | 0.24 pb |

Table 16: Shifts in the $t\bar{t}$cross section as determined using pseudo-experiments for $\pm 1\sigma$ shifts in the specified PDF Eigenvector parameters.

## 13.2 Determination of the $Q^2$ Shape Systematic

The choice of $Q^2$ scale in our MC's can have a significant impact on the shape of the $H_T$ distribution. To measure our sensitivity to this effect, we performed pseudo-experiments using two different ALPGEN MC samples. One sample was our nominal one which uses for a $Q^2$ value of $M_W^2$+$\Sigma p_T^2(jet)$. The second ALPGEN sample was generated using $Q^2 = M_W^2$. Pseudo-experiments were generated by selecting events from the second template and then fitting the result to the nominal template. For the W+$\geq$ 3 jet events, the difference between the two over the full range of $Q^2$ is 28%. Taking half the difference, we use 14% as the $Q^2$ systematic. For the W+$\geq$ 4 jet bin, the fraction of signal to background is much larger. Consequently, we are less sensitive to the $Q^2$ scale in the W+jets. In this bin, the systematic is 0.62%. The $H_T$ shape for the two ALPGEN samples is shown in Figure 22.

### 13.2.1 Shape systematic from variation of the top quark generator

To test the dependence of our result on the generator used for $t\bar{t}$ we examine the difference in the $H_T$ shape obtained using PYTHIA(the default) and HERWIG. Once again, the systematic

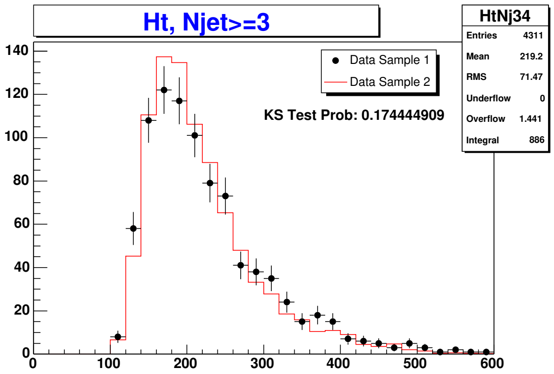



Figure 22: $H_T$ for two different ALPGEN samples, one in which $Q^2 = M_W^2$, and one in which $Q^2 = M_W^2 + \Sigma {p_T}^2(jet)$ (nominal).

is taken to be half of the result, or 0.60%.

### 13.3 Determination of acceptance systematics

Most of the systematics listed above also impact the acceptance. Unfortunately, each of these are completely correlated with the shape systematic. This can be seen by again examining the cross section formula:

$$\sigma(t\bar{t}) = \frac{\mathcal{F} \cdot N}{\mathcal{L} \cdot A}$$

For example, the if the true top mass is 170 $GeV/c^2$, the above shape systematic discussion indicates that we would determine a fit fraction $\mathcal{F}$ which is -11.6% smaller than its true value on average (since we are using an $H_T$ shape determined using a top mass of 175 $GeV/c^2$). This would result in a cross section that is -11.6% smaller than its true value. From the Monte Carlo, we find that the acceptance $A$ for events using a mass of 170 $GeV/c^2$ is approximately 5% smaller than for a mass of 175 $GeV/c^2$. Again, since we are assuming the top mass is 175 $GeV/c^2$ and don't know it is really 170 $GeV/c^2$, we use the larger acceptance and obtain a cross section which is 5% too small. These two effects go in exactly the same direction, unfortunately, and as a result, the overall effect of varying the top mass is the sum of these two effects. We have recently decided after a discussion in the lepton+jets group to not take this as a systematic but to instead quote our result as a function of top mass, as was done in Run I.

The acceptance systematics due to $E_T$ scale, ISR, and top generator are shown in tables 19 and 20, for the $W+ \geq 3$ and $W+ \geq 4$ samples, respectively.

The systematic due to lepton ID effects is also listed in these tables, and is taken from CDF Note 6574 [10]. The systematic is smaller than listed in that note, since ISR and $E_T$ scale effects were included, and we have redone those systematics for this analysis.

### 13.4 Determination of the QCD background fraction using MET vs ISO systematic

In order to estimate the systematic uncertainty of the QCD background using this MET vs ISO technique, we varied the location of the upper isolation and lower $\not\!\!E_T$ cuts. Each of these values are somewhat arbitrary. We varied Isolation boundary between 0.17 and 0.50. (0.2 is our standard) in steps of 0.03 to 0.05. We varied the MET boundary between 5 and 15 GeV in 2 GeV steps (10 is our standard). The results of these studies can be found in Tables 17 and 18. Based on the fluctuations in these numbers, we take as our background uncertainty to be 10%.

To determine a systematic, we use pseudo-experiments and shift the fraction of QCD normalization by ± 10%. In so doing, we find that the change in cross section is negligible. So, in order to quote a systematic, we will take the expected QCD fraction of 6.7%, double it and halve it in pseudo experiments, the nominal cross section goes from 7.0pb to 6.86 ± 0.02 pb (doubling it) and 7.10 ± 0.02 pb(halving it). We will take as the systematic then half the difference between these two extremes. In so doing, the systematic uncertainty on the MET vs ISO bckg calculation is 0.12pb or 1.7%.

| **MET Boundary** | **Bckg Fraction** | **Iso Boundary** | **Bckg Fraction** |
|---|---|---|---|
| 0.17 | 9.7 ± 1.6% | 5 GeV | 8.6 ± 1.7 % |
| 0.18 | 9.4 ± 1.5% | 7 GeV | 8.9 ± 1.6 % |
| 0.19 | 9.1 ± 1.5% | 9 GeV | 8.9 ± 1.5% |
| 0.20 | 8.9 ± 1.6% | 10 GeV | 8.9 ± 1.5% |
| 0.21 | 8.9 ± 1.6% | 11 GeV | 8.9 ± 1.5% |
| 0.25 | 8.0 ± 1.5% | 13 GeV | 9.6 ± 1.6% |
| 0.30 | 8.5 ± 1.7% | - | - |
| 0.40 | 8.9 ± 1.9% | - | - |
| 0.45 | 9.2 ± 2.2% | - | - |

Table 17: QCD background fractions for the electron+jet data as determined using MET vs ISO in the non-isolated lepton sample. The boundary regions are varied and the resulting fluctuations in the background fraction used to determine the QCD systematic

| **MET Boundary** | **Bckg Fraction** | **Iso Boundary** | **Bckg Fraction** |
|---|---|---|---|
| 0.17 | 4.1 ± 1.1% | 5 GeV | 4.5 ± 1.3 % |
| 0.18 | 4.1 ± 1.0% | 7 GeV | 4.0 ± 1.1 % |
| 0.19 | 4.0 ± 1.0% | 9 GeV | 4.0 ± 1.0 % |
| 0.20 | 4.0 ± 1.0% | 10 GeV | 4.0 ± 1.0% |
| 0.21 | 4.1 ± 1.0% | 11 GeV | 4.1 ± 1.0% |
| 0.25 | 3.8 ± 1.0% | 13 GeV | 4.2 ± 1.0% |
| 0.30 | 3.7 ± 0.9% | 15 GeV | 4.1 ± 0.9% |
| 0.40 | 3.3 ± 0.9% | - | - |
| 0.45 | 3.4 ± 1.0% | - | - |

Table 18: QCD background fractions for the muon+jet data as determined using MET vs ISO in the non-isolated lepton sample. The boundary regions are varied and the resulting fluctuations in the background fraction used to determine the QCD systematic

## 13.5 Determination of background model systematic

The sample that we are using to model the background from non-W is the non-isolated lepton sample. This sample requires that the isolation of the lepton be $> 0.2$, but otherwise, all of the other cuts, both on the lepton as well as the MET and Jet cuts, are identical. This sample presumably includes fake leptons, as well as $b\bar{b}$ in which one of the b's decays semi-leptonically to an electron or muon. What we don't know is how the kinematics (and composition) of this sample changes as the isolation cut is changed from $I < 0.1$ to $I > 0.2$.

One thing we can do is to see what happens if we change the composition of the sample so that it is dominantly made up of only fakes. A convenient sample for this is the conversion electron sample. This sample is obtained easily, since these events are included in the standard tight lepton sample, and are already flagged as conversions. We simply require that the electron is an identified conversion (as opposed to removing these, which is the default procedure). Conversion electrons are real electrons, but are definitely non-W in origin. In addition, since the isolation cut *is* applied, this tests some of the assumptions we make in using the non-isolated sample.

The size of this systematic also provides a very good argument for reducing the overall size of the non-W background. As should be obvious, the smaller the non-W background, the smaller the systematic associated with lack of knowledge of the non-W $H_T$ shape.

We perform pseudo-experiments for two different cut sets: without the $\Delta\phi$ cut and with the $\Delta\phi$ cut. The results are the following:

- Without the $\Delta\phi$ cut: the expected precision on the cross section is approximately 28%, and the systematic due to the uncertainty in modeling the background is 32%.
- With the $\Delta\phi$ cut: the expected precision on the cross section is approximately 28%, and the systematic due to the uncertainty in modeling the background is 16%.

So applying the $\Delta\phi$ does not change the expected precision on the cross section, but it does drop the systematic due to background modeling by a factor of 2. This is the primary motivation for making this cut, and we quote the systematic as $\pm 16\%$.

| Effect | Shape | Acceptance | Total |
|---|---|---|---|
| Et Scale | 28% | 5.1% | 30.2% |
| Generator | 0.60% | | 0.60% |
| //Top Mass | 8.0% | 5.0% | 13% |
| Lepton ID (includes PDF) | 3.5% | 5.3% | 8.8% |
| ISR | 0.56% | 0.78% | 1.3% |
| $Q^2$ | 14% | | 14% |
| QCD | 1.7% | | 1.7% |
| Luminosity | | | 5.9% |
| Background model | 16% | | 16% |
| Total | | | 38.3% |

Table 19: Relative systematic uncertainty in the fitted value of the $t\bar{t}$ cross section in the W+$\geq$3 jet sample, coming from various sources, as discussed in the text.

| Effect | Shape | Acceptance | Total |
|---|---|---|---|
| Et Scale | 20.4% | 13% | 31.5% |
| Generator | 0.60% | | 0.60% |
| //Top Mass | 8.0% | 5.0% | 13% |
| Lepton ID (includes PDF) | 4.3 | 5.3% | 9.6% |
| ISR | 0.56% | 0.78% | 1.3% |
| $Q^2$ | 0.6 % | | 0.6% |
| QCD | 3.3% | | 3.3% |
| Luminosity | | | 5.9% |
| Background model | 16% | | 16% |
| Total | | | 37.5% |

Table 20: Relative systematic uncertainty in the fitted value of the $t\bar{t}$ cross section in the W+≥4 jet sample, coming from various sources, as discussed in the text.

### 13.6 Total systematic uncertainty

The total systematic uncertainty, including effects from luminosity, acceptance and shape, and background modeling, are shown in Tables 19 and 20, for the $W+\geq 3$ and $W+\geq 4$ samples, respectively. In the cases where both a shape and acceptance systematic are listed, the total is given by the straight sum of the two. The final systematic including all effects is obtained by adding in quadrature the quantities listed in the column labeled $Total$. The systematic in the W+≥ 3 jet sample is 38.3%, while in the W+≥4 jet sample it is 37.5%.

## 14 Summary

With 50% more data, we have updated our previous measurement of the $p\bar{p} \to t\bar{t}$ cross section in the lepton+jets channel using a kinematic analysis. Systematic errors are unchanged from our previous note, since the dominant systematic, from the data-MC agreement of the jet energy scale, will remain the same until new simulation becomes available. The $H_T$ variable was chosen to separate signal from background. The background fractions were determined from a combined fit $H_T$ templates based on "W-like" backgrounds, QCD fakes and $t\bar{t}$. The $t\bar{t}$ and W background shapes were determined from MC simulations. The QCD fakes template was derived from a data sample of non-isolated leptons and then checked with conversions. Based on the 195 pb$^{-1}$ of data collected for the summer conferences, the $t\bar{t}$ production cross-section is measured to be 4.7 ± 1.6(stat) ± 1.8(syst)pb in the W+≥ 3 jet sample. In the W+≥4 jet sample, the $t\bar{t}$ production cross section is measured to be 7.8 ± 2.0(stat) ± 3.0(syst)pb. These numbers should be compared to the latest theoretical prediction of $6.7^{+0.71}_{-0.88}$ pb [11].

## References


[1] CDF note 6548 and Top group dataset page
`http://www-cdf.fnal.gov/internal/physics/top/topdata/TopData_4111.html`

[2] DQM good run list page
`http://www-cdf.fnal.gov/internal/dqm/goodrun/good.html`

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

[4] D. Acosta *et al.*, *A First Look at the CLC Luminosity Measurements*, CDF Note 6052.

[5] Information on MC samples can be found here
`http://www-cdf.fnal.gov/internal/physics/top/RunIIWjets/final4111/TopAnalysis_4111.html#MC`

[6] "CKKW use in CDF", E. Thomson MC4Run2 Workshop December 4, 2003.
`http://cepa.fnal.gov/patriot/mc4run2/MCTuning/031204.html`

[7] S. Mrenna, P. Richardson, hep-ph/0312274

[8] V. Drollinger, M. Gold *et al.*, *A Neural Network Measurement of the Top Quark Mass*, CDF note 6078.

[9] Andy Beretvas, Morris Binkley and Luc Demortier, *Binned KS Probability*, CDF Note 4373.

[10] Ken Bloom, Koji Ebina, Evelyn Thomson, Kohei Yorita, $t\bar{t}$*Event Selection and Detection Efficiency for Summer 2003 Lepton+Jets Analyses*, CDF Note 6574.

[11] R. Bonciani, S. Catani, M.L. Mangano and P. Nason, *Nucl. Phys. B 529* (1998) 424, as updated by M. Cacciari, S. Frixione, M.L. Mangano, P. Nason, G. Ridolfi, *The t-tbar cross-section at 1.8 and 1.96 TeV: a study of the systematics due to parton densities and scale dependence*, CERN-TH/2003-054