

Search for Heavy Top $t' \rightarrow Wq$ in Lepton Plus Jets Events in $\int \mathcal{L} dt = 2.3 \text{ fb}^{-1}$

John Conway, David Cox, Robin Erbacher, Andrew Ivanov,
Will Johnson, Alison Lister, Tom Schwarz
University of California, Davis

February 21, 2008

Abstract

This is an update on the search for the heavy top (t') quark pair production decaying to Wq final states using 2.3 fb^{-1} data sample of lepton+jets.

We reconstruct the mass of the t' quark and perform a 2D-fit of the observed (H_T, M_{reco}) distribution to discriminate the new physics signal from Standard Model backgrounds. We exclude Standard Model fourth-generation t' quark with mass below 315 GeV at 95%CL.

1 Introduction

The topic of interest of this study is to investigate whether the present data allow or preclude the production of hypothetical new quarks which decay to final states with a high- p_T lepton, large \cancel{E}_T , and multiple hadronic jets, having large total transverse energy H_T , and thus mimicking top quark pair event signatures in the lepton+jets decay channel.

We performed previous iterations of this analysis using 194 pb⁻¹ [1], 347 pb⁻¹ [2] and 760pb⁻¹ [3] Run II datasets. In each of the latter two, several improvements were consecutively made to the technique in order to increase the sensitivity to new physics as detailed in the respective notes. The 760pb⁻¹ analysis has been submitted to PRL [4] .

This current analysis uses 2.3 fb⁻¹ of integrated luminosity.

We refer to the hypothetical new quark as t' for brevity, although such a signature could be a standard fourth-generation up-type heavy quark in which the splitting between the t' mass and the b' mass is less than the mass of the W boson (so that the decay is predominantly to Wq), as well as any up-type quark, such as for example, arising from “beautiful mirrors” model [5] with the same quantum numbers as the top, or the heavy top predicted in Little Higgs models [6]. Further discussion of theoretical models can be found in the introduction to the PRL [4].

For the purposes of this analysis we make the following assumptions. The new quark

- is pair-produced strongly,
- has mass greater than the top quark, and
- decays promptly to Wq final states.

Due to a large number of possible interpretations: a variety of theoretical models predicting similar signatures as well as a number of free parameters within each model, in case of an excess we have arranged a priori to estimate its significance without attributing the events to a particular new physics model, but with a purpose to investigate them in more detail. Or otherwise in case of no excess observed we set a limit on the fourth-generation t' quark pair production cross section (times branching ratio $t' \rightarrow Wq$). Inevitably, however, other proposed models will have kinematic distributions and acceptances different from a generic fourth generation quark.

2 Data Samples and Monte Carlo Simulation

We make use of the official Top Group high p_T lepton datasets **bhelxx** and **bhmuxx** up to and including period 13 data (up to run 246231). These datasets are split by lepton type into CEM for electrons and CMUP and CMX for muons. The data is processed using the 6.1.1 production release. The analysis is performed in the 6.1.4 offline release.

Version 19 of the good run requirements [7] are applied with no silicon requirements. This data sample corresponds to the total integrated luminosity of 2339 pb⁻¹ for CEM and CMUP triggers and 2283 pb⁻¹ for CMX trigger [8].

For simulation of the Standard Model processes we use the official Top Group 6.1.4mc Monte Carlo samples [9]. The dominant backgrounds to $t'\bar{t}'$ are W +jets, which is modeled with ALP-GEN+HERWIG, and $t\bar{t}$ modeled with PYTHIA. We use a sample of $t\bar{t}$ events with mass of 175 GeV (ttkt75). Other backgrounds include W +heavy flavor jets, Z +jets, WW +jets, WZ +jets and single top. These backgrounds were investigate only in as much as to verify that using the W +jets MC to model our Electroweak background is a valid approximation to make.

The QCD background is modeled from the data sample of anti-electrons.

$t\bar{t}$ pair production is simulated in PYTHIA assuming 100% branching fraction to Wb . We generated several run-dependent Monte Carlo samples of $\sim 1\text{M}$ events each, in 20 GeV mass steps from 180 to 400 GeV, as well as 450 and 500 GeV samples.

3 Event Selection

We follow event selection for Winter 2008 analyses as documented in the Top Properties web-page [10].

We require one high- p_T electron or muon, large missing transverse energy \cancel{E}_T and at least four energetic jets (corrected to Level 5).

Additional cuts are applied specifically to our analysis and are motivated elsewhere in this note

- Cut in the $\Delta\Phi$ between the corrected missing energy and the lepton vs missing energy plane: $\Delta\phi \geq A_1 - 1/B_1 MET$ where $A_1 = 4.408$; $B_1 = 6.11$;
- Cut in the $\Delta\Phi$ between the corrected missing energy and the leading jet vs missing energy plane: $\Delta\phi \geq A_2 - 1/B_2 MET$ where $A_2 = 1.888$; $B_2 = 21.6$;

These cuts are designed to remove a significant fraction of the QCD background in our sample [11].

- Leading jet $E_T > 60$ GeV to further reduce the QCD backgrounds.
- $\Delta\phi(\text{MET} \rightarrow \text{muon}) < 3.05$ to remove very high p_T muons due to muon reconstruction issues.
- ΔZ between the muon and the nearest Zvtx with vertex quality $\iota=12$ and with $\iota=2$ tracks to remove events where the muon is the only track in a vertex.

Using these event selection criteria we observe a total of 1118 events, 667 events of which are CEM electron + jets, 300 CMUP and 151 CMX muon + jets.

3.1 QCD removal

The signal we are searching for is in the high H_T and high M_{rec} region relative to all SM backgrounds, QCD in particular. In order to cut out as much unwanted background as possible, to reduce any effect of mis-modeling of the QCD background, we chose to increase the leading jet E_T cut relative to that of the standard Lepton + Jets selection.

It was possible to increase the leading jet E_T cut from 20 GeV up to 60 GeV without losing more than 6% of the expected signal at 240 GeV, 1.7% at 300 GeV. This cut removes 53% of the QCD background that remained after the previously mentioned QCD cut.

3.2 High p_T muon removal

We defined out control regions as the same basic selection but requiring only 2 or 3 jets instead of at least 4. While investigating these regions we observed a few events with extremely high p_T muons. The intrinsic detector resolution is limiting our ability to reconstruct accurately the momentum of muons above a certain threshold because the curvature becomes essentially too flat. Events with muons at unmeaningful p_T will tend to have the missing E_T , MET, back to back with the muons, because the MET is corrected for the presence of the muon. This is shown in figure 1 for event in our control region (lepton + Met + 2 or 3 jets). From this plot it is clear that all muons above ~ 400 GeV have a $\Delta\phi$ between the muon and the MET above 3.05. We expect that this behavior is the same in our signal (4-jet) region. In order to remove such events from our sample, we place an additional cut on this angle for muons only.

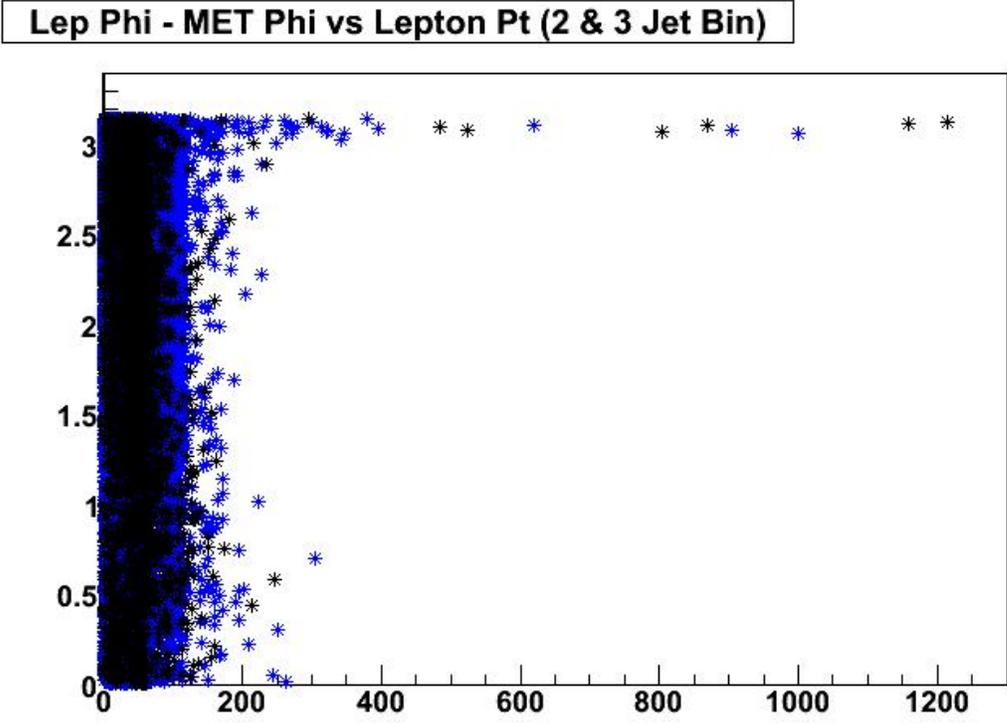


Figure 1: $\Delta\phi$ between the muon and the MET in the 2- (blue) and 3-jet (black) control regions.

4 Kinematic Variables and Mass Reconstruction

The kinematic variables used for this analysis are the same as the ones used for the previous version [3]; We perform a 2-D likelihood fit to H_T and M_{reco} , where

$$H_T = \sum_{jets} E_T + E_{T,\ell} + \cancel{E}_T, \quad (1)$$

and M_{reco} is the reconstructed mass of the top-like system as described below.

These variables serve as a good discriminators between Standard Model and new physics processes associated with production of high mass particles, such as the t'

We utilize the fact that t' decay chain in the regime we are sensitive to is identical to the one of the top quark, and attempt to reconstruct its mass similarly to as it is done in the top quark mass measurement analyses. We adopt the template method for top quark mass reconstruction [18], which is based on χ^2 -fit of kinematic properties of final top decay products.

For each event there are total $4!/2 = 12$ combinations of assigning 4 jets to partons. In addition, there are two solutions for unknown P_z neutrino momentum. The MINUIT minimization is performed for each of the 24 combinations, and then the permutation with the lowest value of χ^2 is selected. The χ^2 is given by the following expression:

$$\begin{aligned} \chi^2 = & \sum_{i=\ell,4jets} \frac{(p_T^{i,fit} - p_T^{i,meas})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(p_j^{UE,fit} - p_j^{UE,meas})^2}{\sigma_j^2} \\ & + \frac{(m_{jj} - m_W)^2}{\Gamma_W^2} + \frac{(m_{\ell\nu} - m_W)^2}{\Gamma_W^2} + \frac{(m_{bjj} - m_t)^2}{\Gamma_t^2} + \frac{(m_{b\ell\nu} - m_t)^2}{\Gamma_t^2}, \end{aligned} \quad (2)$$

where invariant masses of W decay products m_{jj} and $m_{\ell\nu}$ are constrained to the pole mass of the W boson m_W , and masses of top and anti-top (t' and \bar{t}') quarks are required to be equal. Jet, lepton and underlying event energies are allowed to float within their uncertainties, while the transverse component of neutrino momentum is calculated at each step of the fit, as follows

$$\vec{p}_T^\nu = -(\vec{p}_T^\ell + \sum \vec{p}_T^{jet} + \vec{p}_T^{UE}). \quad (3)$$

The longitudinal component p_z^ν is unconstrained parameter in the fit and initialized with the value such that $m_{\ell\nu}$ acquires W pole mass m_W .

The m_t is the free parameter initialized with $m_t = 175$ GeV, and its value in the best fit is declared to be the reconstructed mass M_{reco} of top (or t' respectively). In accordance with [18] to assure more accurate mass reconstruction the fitted jets are corrected to Level 5, and then *top-specific* jet corrections are applied, which differ for b-jets and quark jets from W 's. The top-specific corrections are derived from Monte Carlo studies. They provide better matching between parton and reconstructed jet energies.

Up to this point we follow the standard top prescription exactly. However:

1. We do not apply a χ^2 cut on reconstructed events.
2. We do not make use of b -tagging information.

Note that we do not attempt to reconstruct the mass of the t' quark as best as possible, but instead use this variable to discriminate t' signal from Standard Model backgrounds.

The χ^2 cut leads to unnecessary loss of statistics, and jeopardizes our sensitivity to possible new physics signals other than t' .

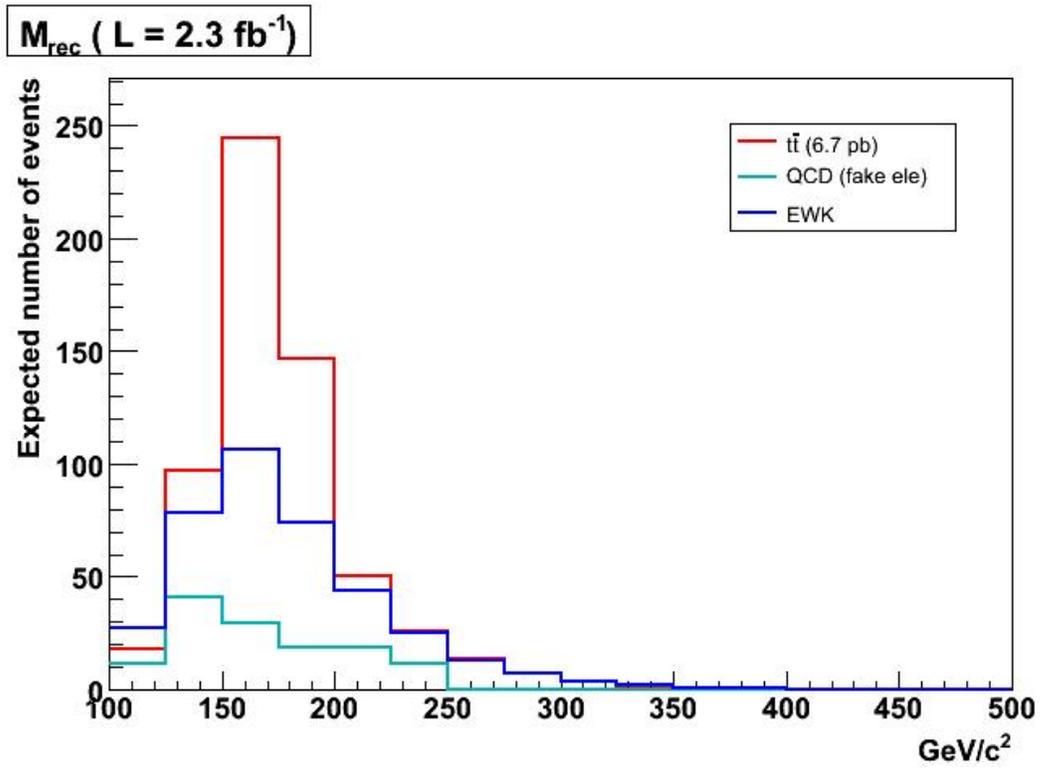


Figure 2: Standard Model M_{reco} templates.

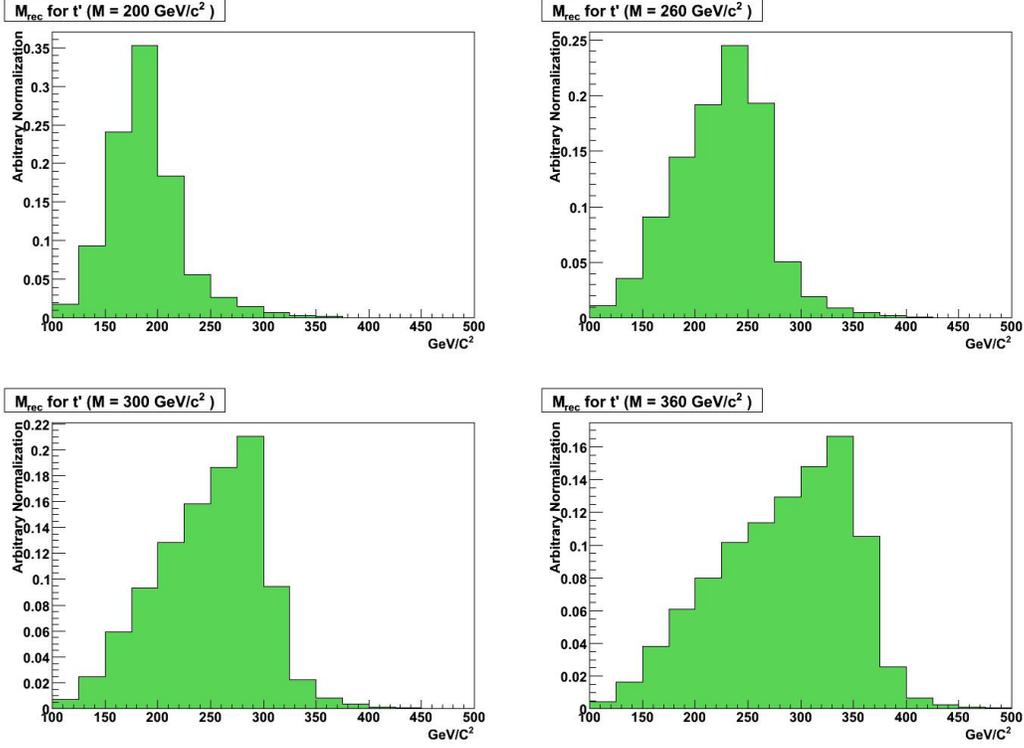


Figure 3: t' M_{reco} templates for masses of 200, 260, 300 and 360 GeV/c^2 .

Templates of M_{reco} for main Standard Model backgrounds (W +jets, QCD and $t\bar{t}$) and t' for various masses are shown in Figures 2 and 3.

Templates of H_T for main Standard Model backgrounds (W +jets, QCD and $t\bar{t}$) and t' for various masses are shown in Figures 4 and 5.

5 Analysis Method

We perform a binned likelihood fit in H_T and M_{reco} to extract the t' signal and/or set an upper limit on its production rate. The likelihood is defined as the product of the Poisson probabilities for observing n_i events in 2-d bin i of (H_T, M_{reco}) :

$$\mathcal{L}(\sigma_{t'}|n_i) = \prod_i P(n_i|\mu_i) \quad . \quad (4)$$

The expected number of events in each bin, μ_i , is given by the sum over all sources indexed by j , which we further subdivide into separate e +jets and μ +jets sub-sources:

$$\mu_i = \sum_j L_j \sigma_j \epsilon_{ij} \quad . \quad (5)$$

Here the L_j are the integrated luminosities, the σ_j are the cross sections, and the ϵ_{ij} are the efficiencies per bin of (H_T, M_{reco}) .

We calculate the likelihood as a function of the t' cross section, and use Bayes' Theorem to convert it into a posterior density in $\sigma_{t'}$. We can then use this posterior density to set an upper limit on (or if we get lucky, measure) the production rate of t' .

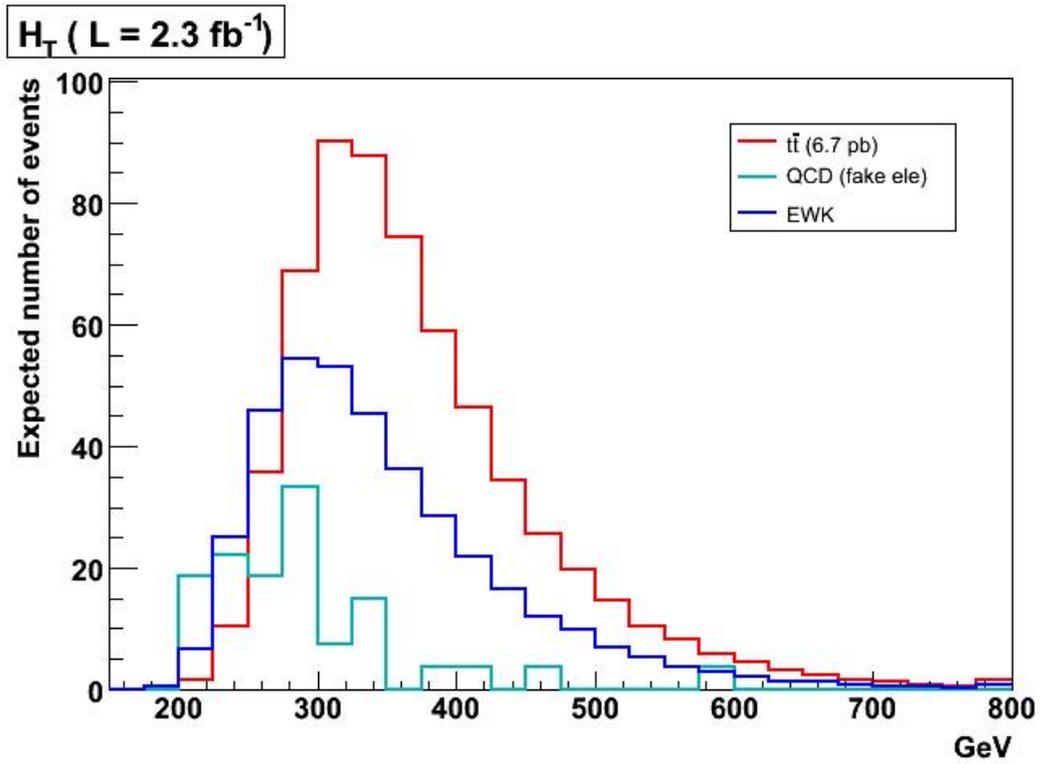


Figure 4: Standard Model H_T templates.

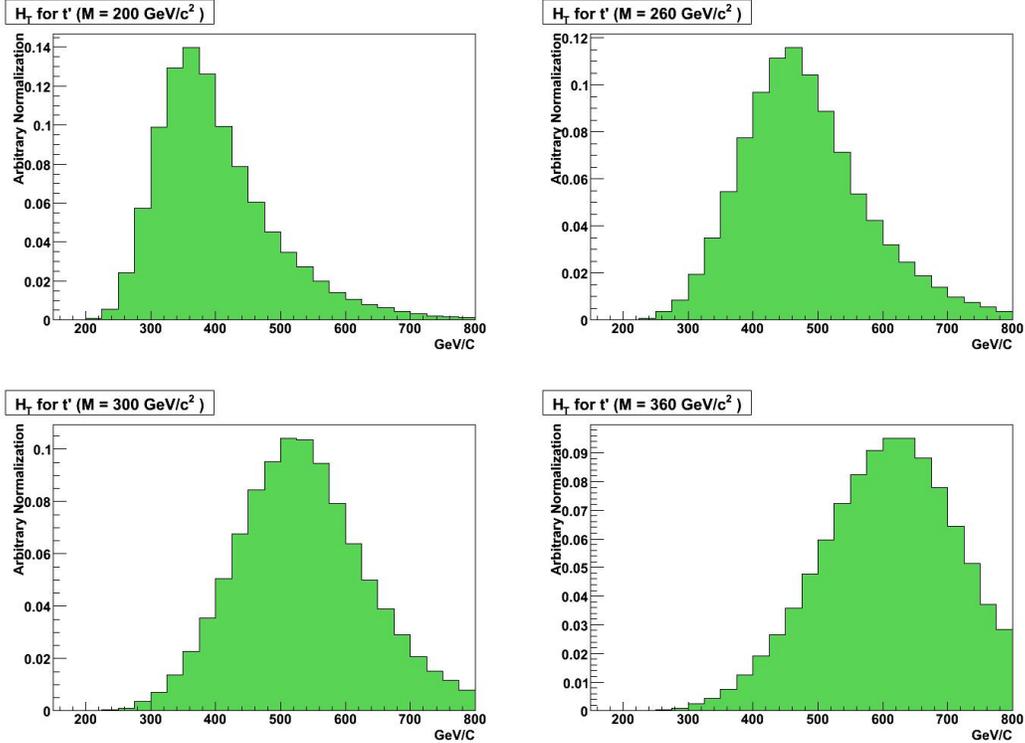


Figure 5: t' H_T templates for masses of 200, 260, 300 and 360 GeV/c^2 .

The production rate for W +jets is a free parameter in the fit. Other parameters, such as the $t\bar{t}$ production cross section, lepton ID, data/MC scale factors, integrated luminosity are related to systematic errors and treated in the likelihood as nuisance parameters constrained within their expected (normal) distributions. We adopt the profiling method [1] for dealing with these parameters, i.e. the likelihood is maximized with respect to the nuisance parameters. The other (marginalization) technique, where the likelihood is integrated over all possible values of the nuisance parameters, is more CPU-intensive, but a cross check is performed and it gave consistent results.

Taking this into account the likelihood takes the following expression:

$$\mathcal{L}(\sigma_{t'}|n_i) = \prod_{i,k} P(n_i|\mu_i) \times G(\nu_k|\tilde{\nu}_k, \sigma_{\nu_k}) \quad , \quad (6)$$

where ν_k are the nuisance parameters, such as $\sigma_{t\bar{t}}$, L_j and etc. $\tilde{\nu}_k$ are their central nominal values and σ_{ν_k} are their uncertainties.

6 Monte Carlo Scaling

The Monte Carlo samples used to model W +jets background as well as for $t\bar{t}$ and t' were generated using run dependent settings with a luminosity profile corresponding to the data period up to $1.1fb-1$ (periods 0-8). For all analyses using further run periods, the Joint Physics group recommends using these MC samples and re-weighting them according to the full luminosity. The re-weighting method used for this analysis is similar to the ones used for the FCNC analysis [17] but extended to include

Lepton Type	Scale Factor	Trigger Efficiency	Integrated Luminosity pb ⁻¹
CEM	0.9802 ± 0.0049	0.9699 ± 0.0046	2339
CMUP	0.9272 ± 0.0062	0.9131 ± 0.0047	2339
CMX	0.9781 ± 0.0075	0.9343 ± 0.0050	2283

Table 1: Lepton ID / reconstruction scale factors for each lepton type. Also shown are the trigger efficiencies and integrated luminosities for each lepton type.

period 13 data. This method considers separately the contribution from period 0 where not all detector components, such as the CMX mini-skirt and keystones, had performance problems. This contribution is kept unchanged but the contributions from period 1-8 are scaled to map onto those from periods 1-13. The effect of not scaling period 0 was estimated to be negligible on the templates for the backgrounds obtained from MC.

We accounted for the effect of this scaling on the average Scale Factor described in 6.1.

6.1 Lepton ID / Reconstruction Data/MC Scale Factors

The lepton ID / Reconstruction Scale Factors, SF, are evaluated for electrons [21] and muons [?] using the standard method suggested by the Joint Physics group, and using the Joint Physics Scale Factor package. The Scale Factors represent the factor by which the MC simulation needs to be scaled in order to match the data.

The average SF is computed separately for each lepton type: CEM, CMUP and CMX. Special care for CMX is needed to account for the fact that the mini-skirt and key-stones were not used in data taking for period 0.

Table 1 shows the ID/Reconstruction scale factors obtained for each lepton type along with their expected uncertainties. The luminosities used in this table are obtained from the official top group Luminosity accounting up to period 13 [8] which is based on the Joint Physics [7].

7 Triggers and Trigger Efficiencies

The triggers used for this analysis are shown in table 2.

The trigger efficiencies for the triggers used in this analysis are calculated according to the method suggested by the Joint Physics group. These Trigger efficiencies are shown, with their errors, in table 1.

Trigger Name	Run Range
ELECTRON_CENTRAL_18	138425-246231
MUON_CMUP18	138425-246231
MUON_CMX18_L2	138425-200272
MUON_CMX18_L2_PT5	138425-226194
MUON_CMX18_L2_PT5_LUMI_200	200272-226194
MUON_CMX18	_JET10 226195-246231
MUON_CMX18	_JET10_LUMI_270 226195-246231
MUON_CMX18	_JET10_DPS 226195-246231

Table 2: Triggers used for this analysis, with run ranges.

8 Data Validation

8.1 Control Region: Lepton + MET + 3 Jets

This section contains many kinematic distributions comparing the data to the sum of all expected backgrounds. In this control region we do not expect a significant contribution from t' . The various background normalizations are obtained as follows:

- The QCD normalization is obtained from a fit to the MET before the MET cut at 20 GeV is applied. The low MET region is where we expect most QCD and hence fitting this region will give us the best estimate of the QCD fraction after all cuts are applied.
- The $t\bar{t}$ normalization is obtained assuming the SM cross-section of 6.7 pb with the acceptances given by the MC.
- The EWK background normalization is obtained by subtracting from the total amount of data the amount of QCD and of $t\bar{t}$.

The plots shown here are

- Missing E_T distributions, figures 6 and 7
- Lepton p_T distributions, figures 8 and 9
- Jet E_T distributions, figures 10, 11, 12, 13, 14 and 15
- H_T distributions, figure 16 and 17
- $\Delta\phi$ between the MET and the leading jet, figure 18
- $\Delta\phi$ between the MET and the lepton, figure 19

These plots show us that the agreement between data and the predictions is satisfactory in the 3-jet control region. In particular the variables that are used in the kinematic fitting are well described. There are no anomalous events in the tails of any of the distributions. Of particular concern were the tails in the lepton p_T spectrum which now appear to agree with data very well.

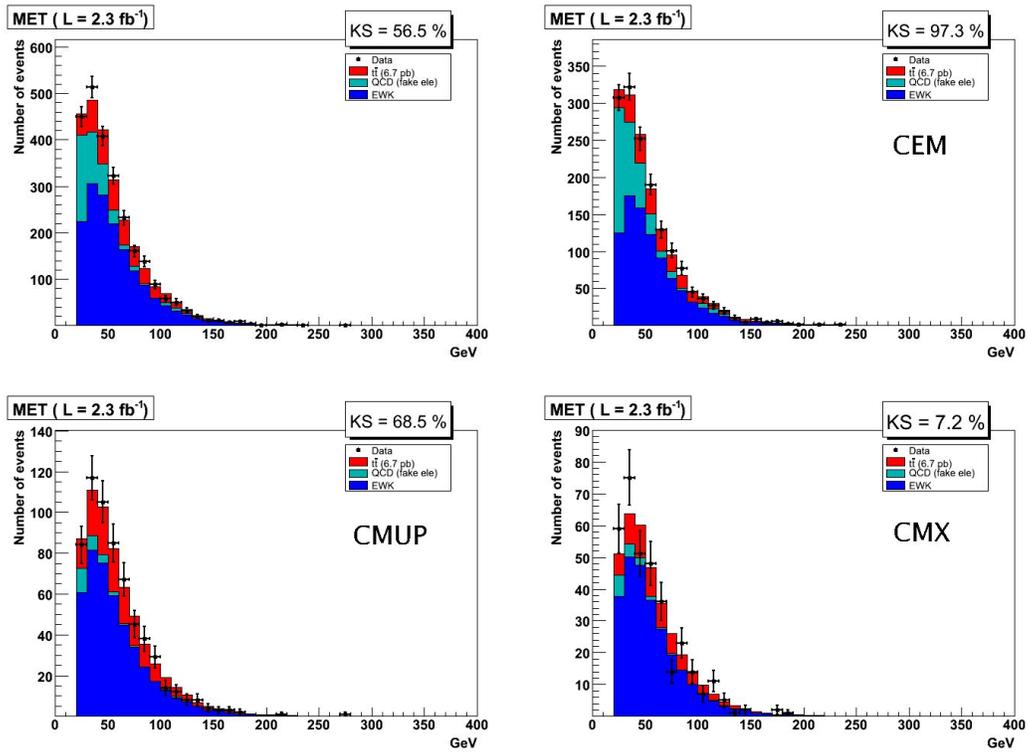


Figure 6: Missing E_T in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

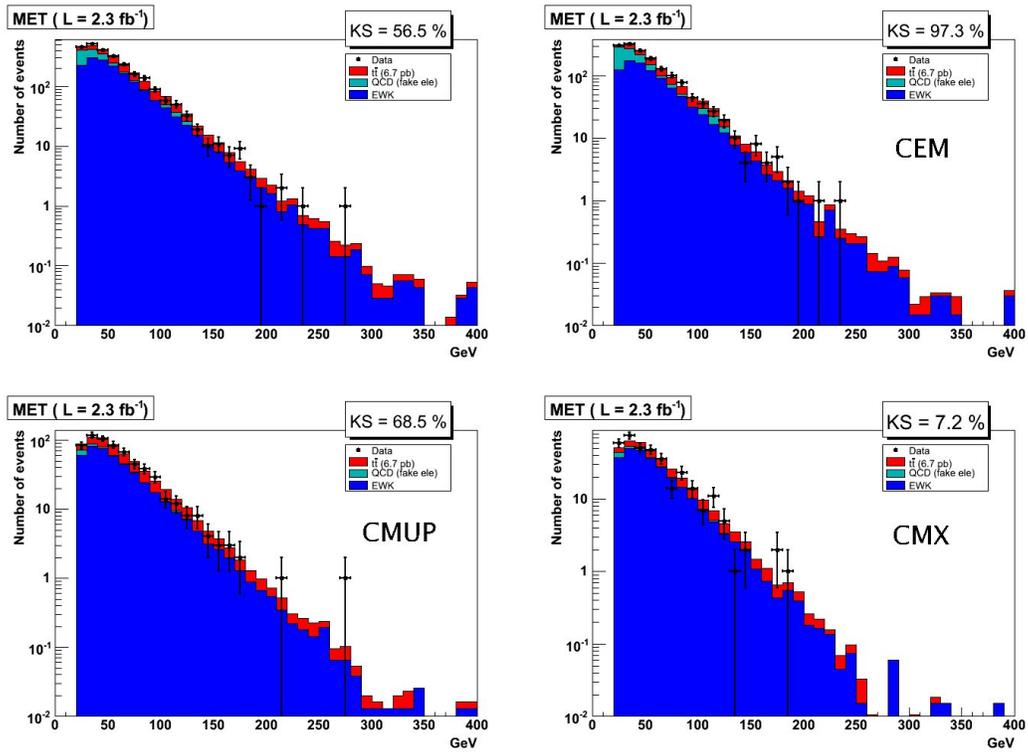


Figure 7: Log scale distributions of the missing E_T in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

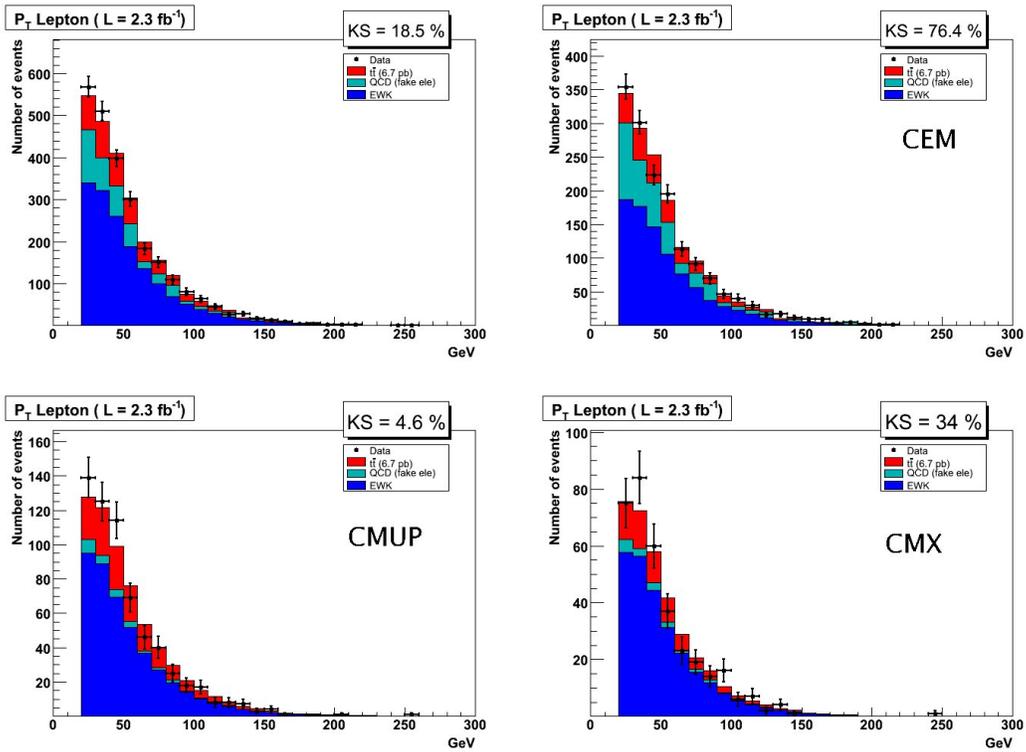


Figure 8: Lepton p_T in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Middle Left: CMUP; Bottom Right: CMX; Bottom left: CMUP in log-scale

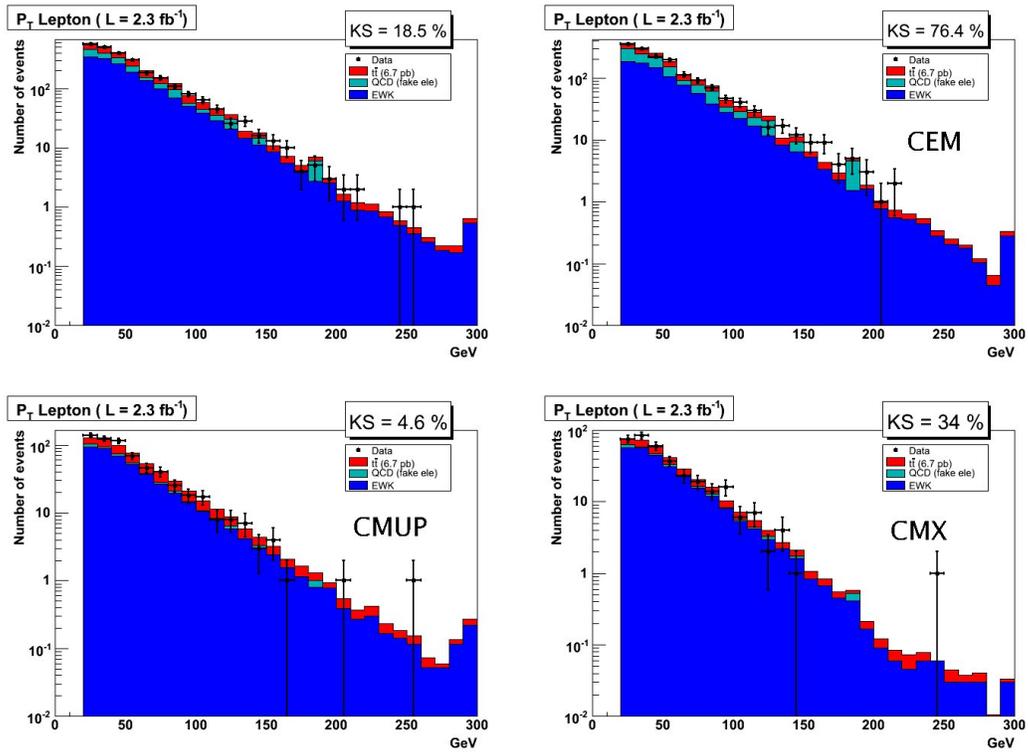


Figure 9: Log scale distributions of the lepton p_T in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom Left: CMUP; Bottom Right: CMX

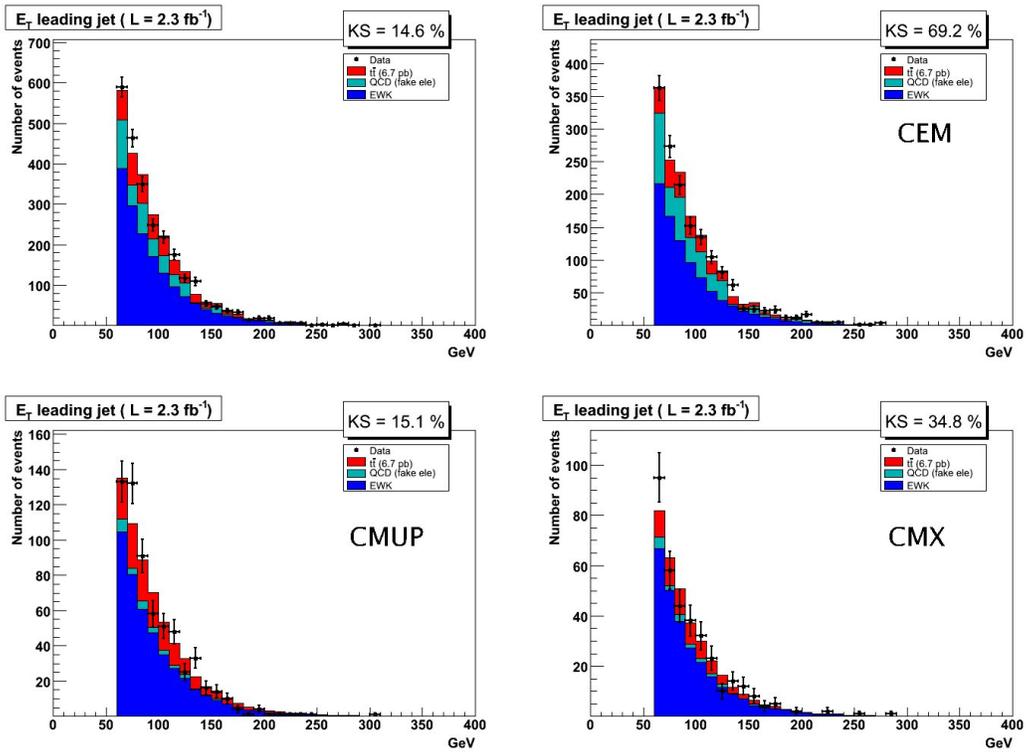


Figure 10: E_T of the leading jet in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom Left: CMUP; Bottom Right: CMX

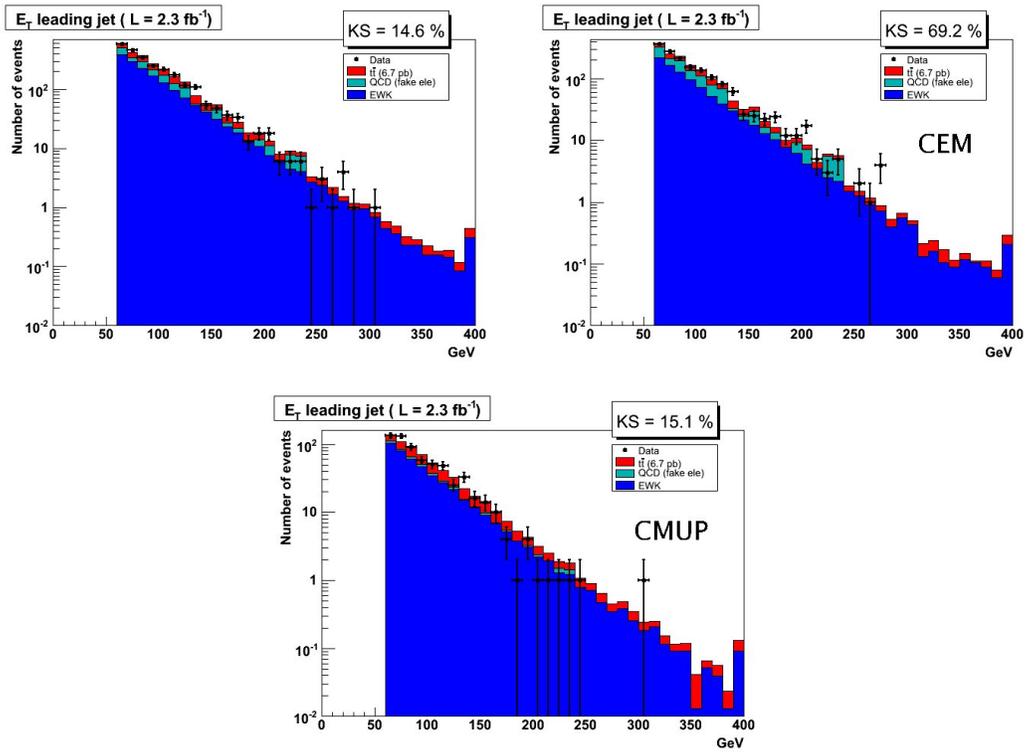


Figure 11: Log scale distributions of the E_T of the leading jet in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom Left: CMUP; Bottom Right: CMX

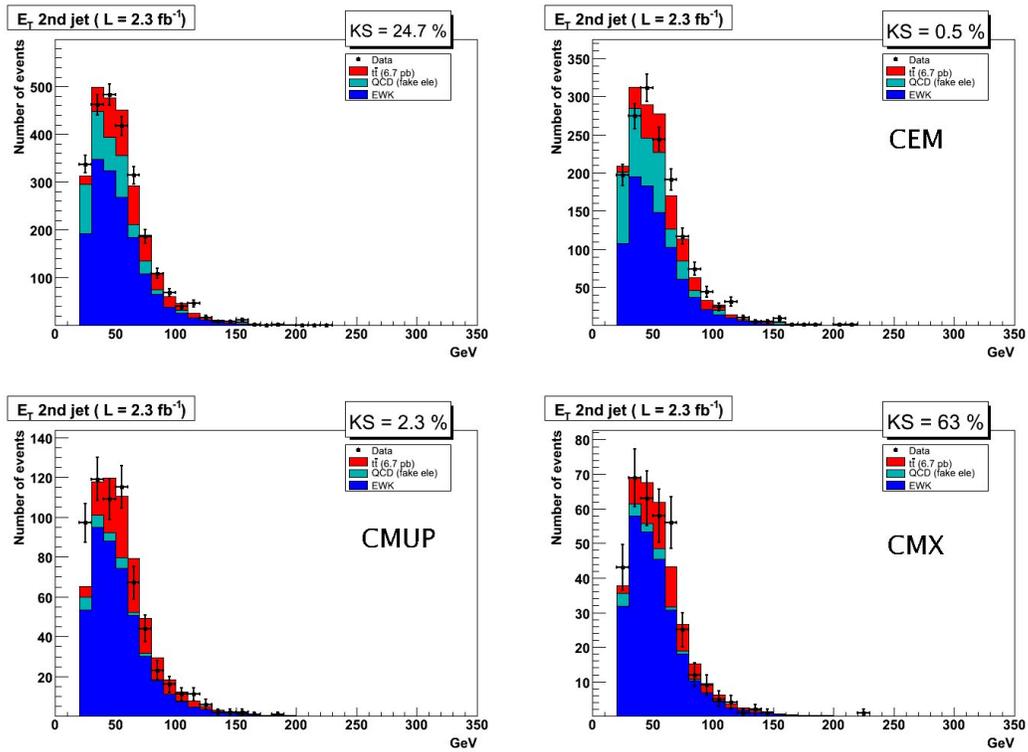


Figure 12: E_T of the 2nd jet in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

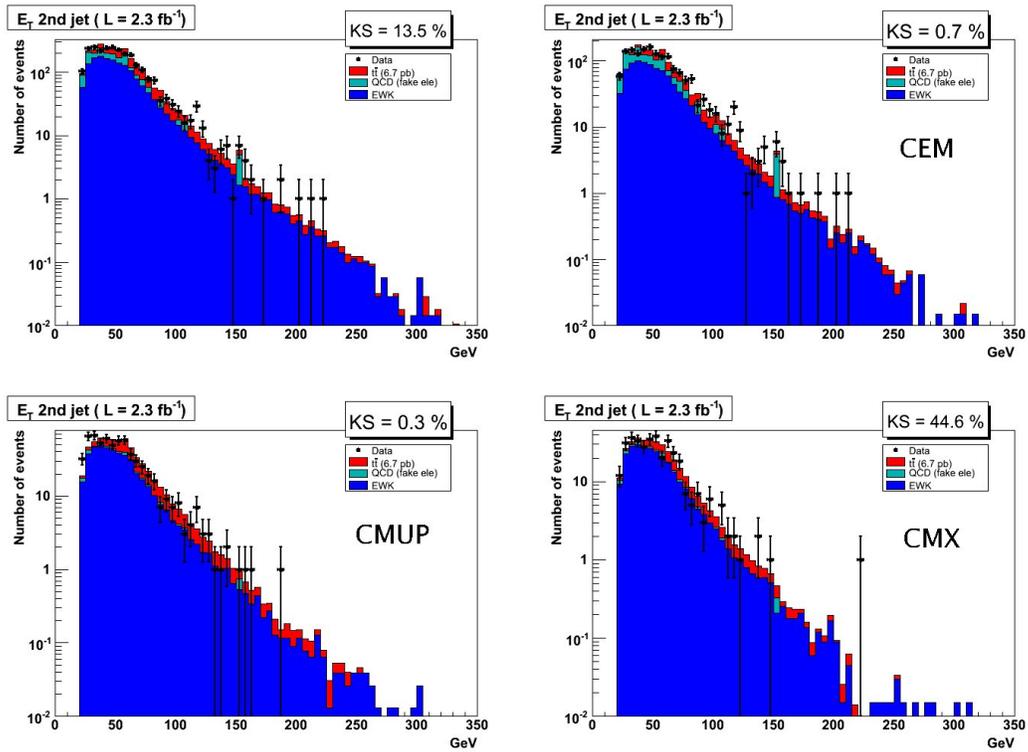


Figure 13: Log scale distributions of the E_T of the 2nd jet in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

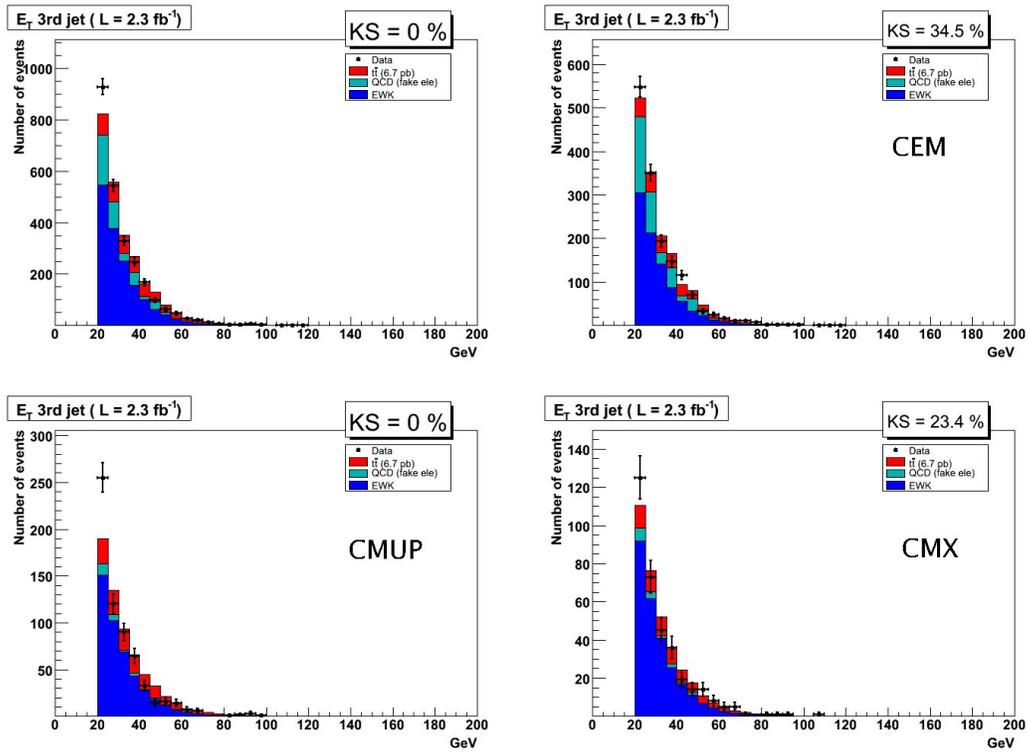


Figure 14: E_T of the 3rd jet in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

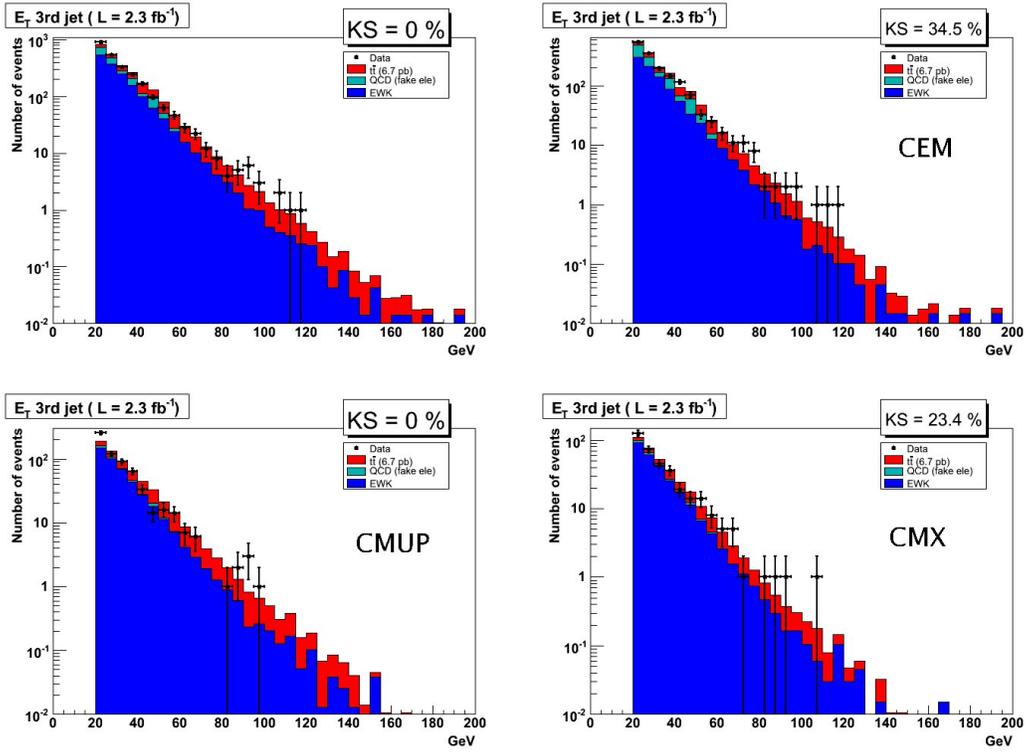


Figure 15: Log scale distribution of the E_T of the 3rd jet in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

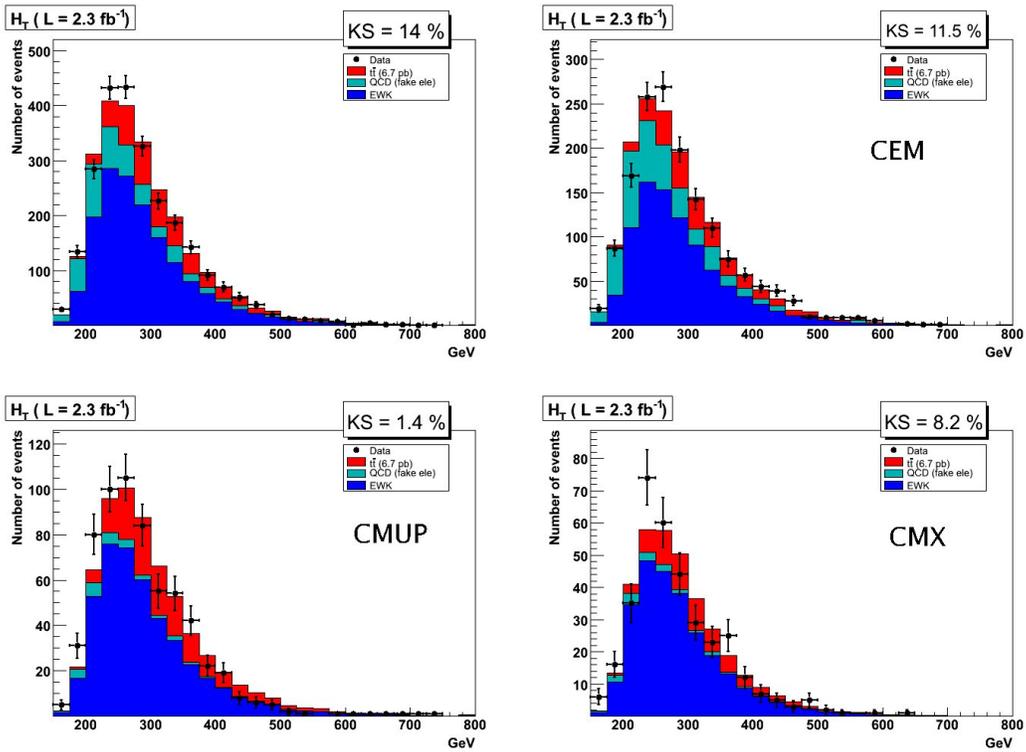


Figure 16: H_T of the event in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

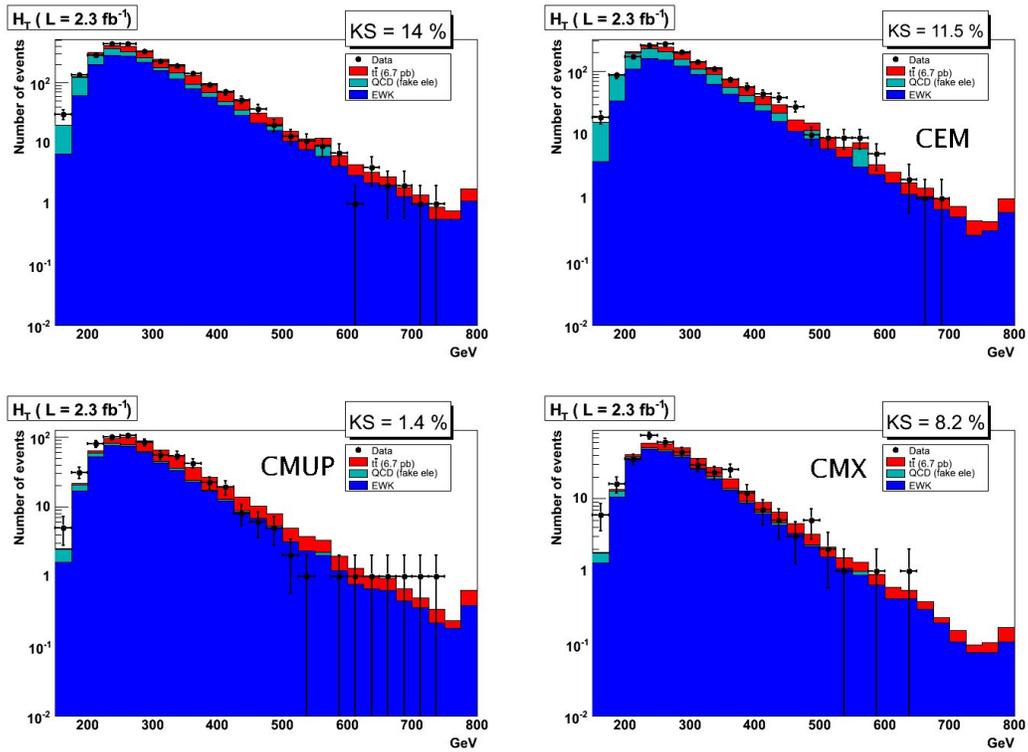


Figure 17: Log scale distributions of the H_T of the event in the 3-jet control region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

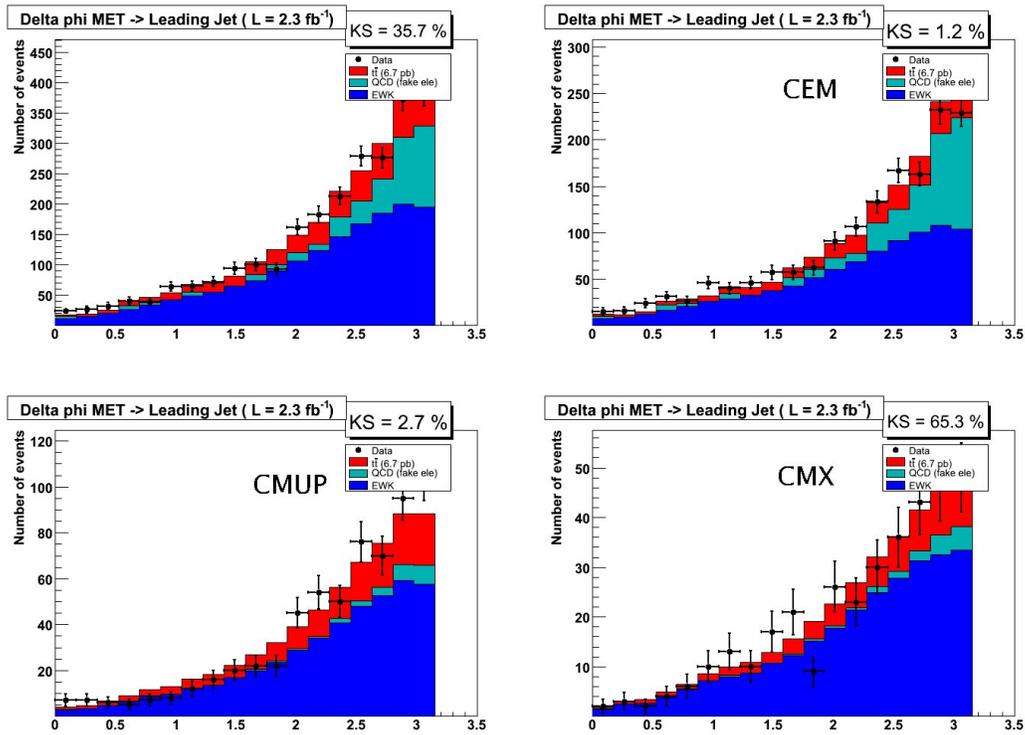


Figure 18: $\Delta\phi$ between the MET and the leading jet; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX

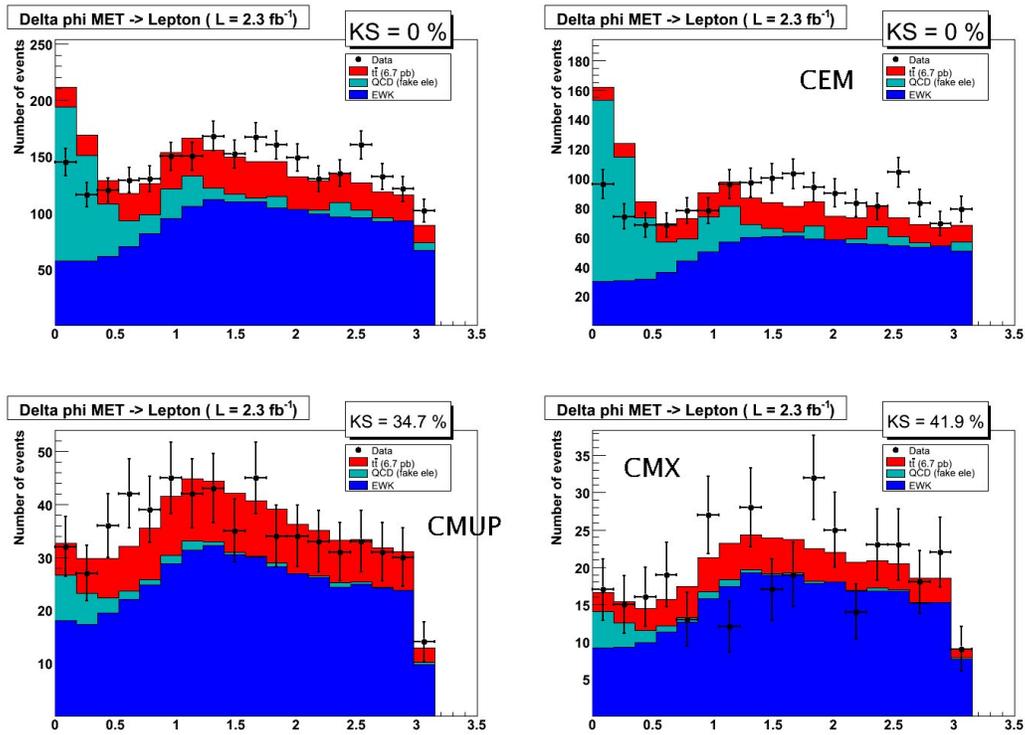


Figure 19: $\Delta\phi$ between the MET and the lepton; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

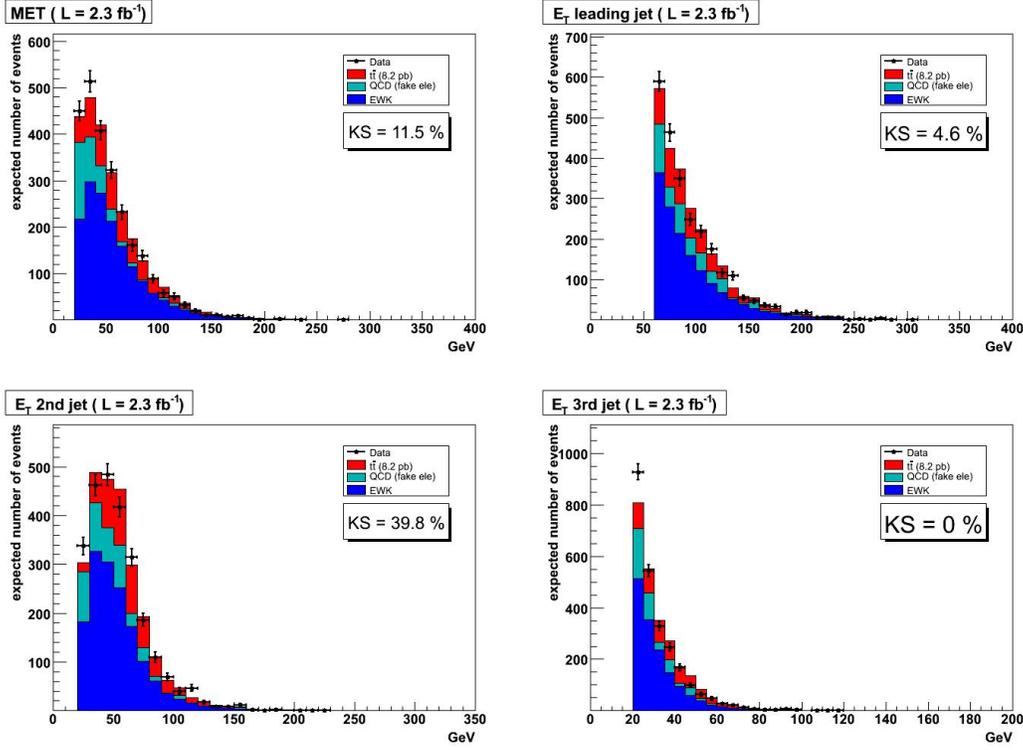


Figure 20: Kinematics for all leptons in the 3-jet control region assuming $t\bar{t}$ cross-section of 8.2 pb; Top Left: MET; Top Right: Leading jet E_T ; Bottom left: E_T of 2nd jet; Bottom Right: E_T of 3rd jet

8.2 Dependence on the top cross-section

These plots are similar to the ones in the previous section but assuming a top mass of 175 GeV but a $t\bar{t}$ cross-section of 8.2 pb. The templates for top therefore don't change, only the normalization. As we state that the QCD normalization fit is dominated by the low MET tails, we keep the normalization for QCD the same. Thus the increase in expected number of $t\bar{t}$ events is compensated by a decrease in the number of EWK backgrounds.

Figure 20 shows the MET and Jet E_T distributions for the 3 jets in the event. Figure 21 shows the lepton p_T distributions on both a linear and a log-scale. Figure 22 shows the H_T distribution.

Although the KS of the distributions are somewhat less than for the lower $t\bar{t}$ cross-section, the agreement between data and predictions is satisfactory.

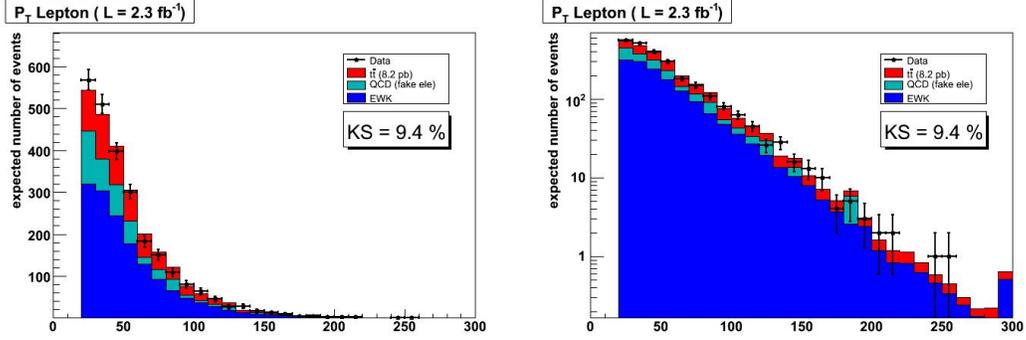


Figure 21: Lepton p_T in the 3-jet control region assuming $t\bar{t}$ cross-section of 8.2 pb; Left: inclusive leptons; Right: log-scale plot

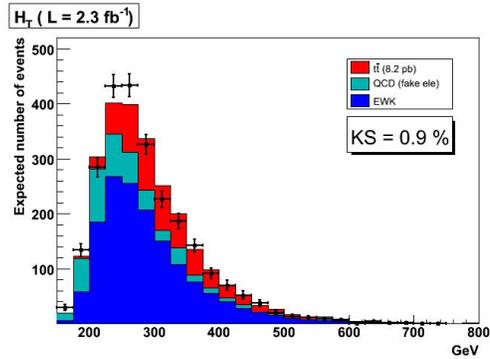


Figure 22: H_T of the event in the 3-jet control region; Top Left: inclusive leptons.

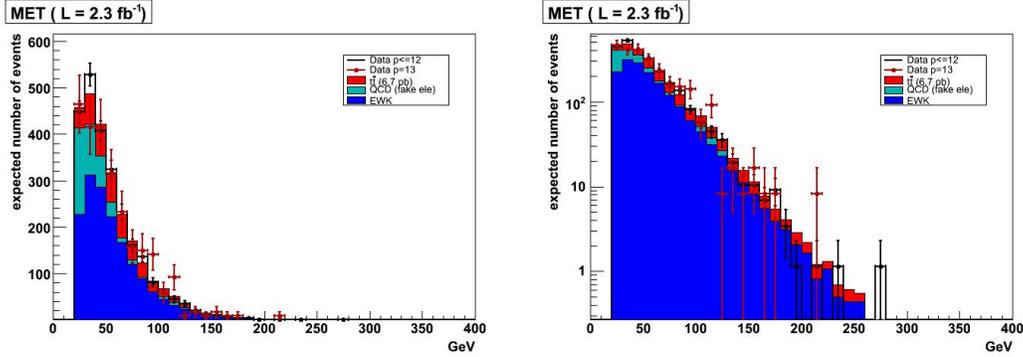


Figure 23: Missing E_T distribution in the 3-jet control region comparing the p13 data with the p12 and earlier data; Left: MET; Right: MET on log-scale

8.3 Period 13 Data Validation in 3-jet Control Region

As we are one of the first groups to use the Period 13 data in the pretag sample, a few validation plots seem in order.

The plots compare the data up to and including period 12 with the period 13 data. The distributions are both scaled up to the total number of events in the whole run period. The errors are scaled in proportion to the scaling.

Figure 23 shows the MET distribution in log and linear scales. Figure 24 shows the lepton p_T distribution for all leptons as well as separated out by lepton type. Figure 25 shows the E_T distributions of the 3 jets as well as the event H_T .

These plots show no significant deviations between period 13 data and the previous data.

8.4 W+lf as Model for all EWK Backgrounds

We assume that the templates for all EWK backgrounds can be described by the W+lf ALPGEN MC for W+n parton where n varies from 0 to 4 partons.

This cross-check is done in the signal region ($i=4$ jets).

The plots in figure 26 show the comparison between the W+lf templates and the templates where all other backgrounds are considered, each weighted by its expected cross-section, except for W+lf which is left to float up to fill in the remainder of the data after all other backgrounds have been taken into account. Other backgrounds are: W+heavy flavor, Z+jets, WW+jets, WZ+jets, ZZ+jets, single top.

9 Signal Region Plots

This section contains the data vs predictions plots in the signal box once we opened the box. The plots shown are

- H_T of the event, figures 27 and 28
- M_{rec} of the event, figures 29 and 30
- p_T of the lepton, figures 31 and 31
- E_T of the leading jet, figures 33 and 33

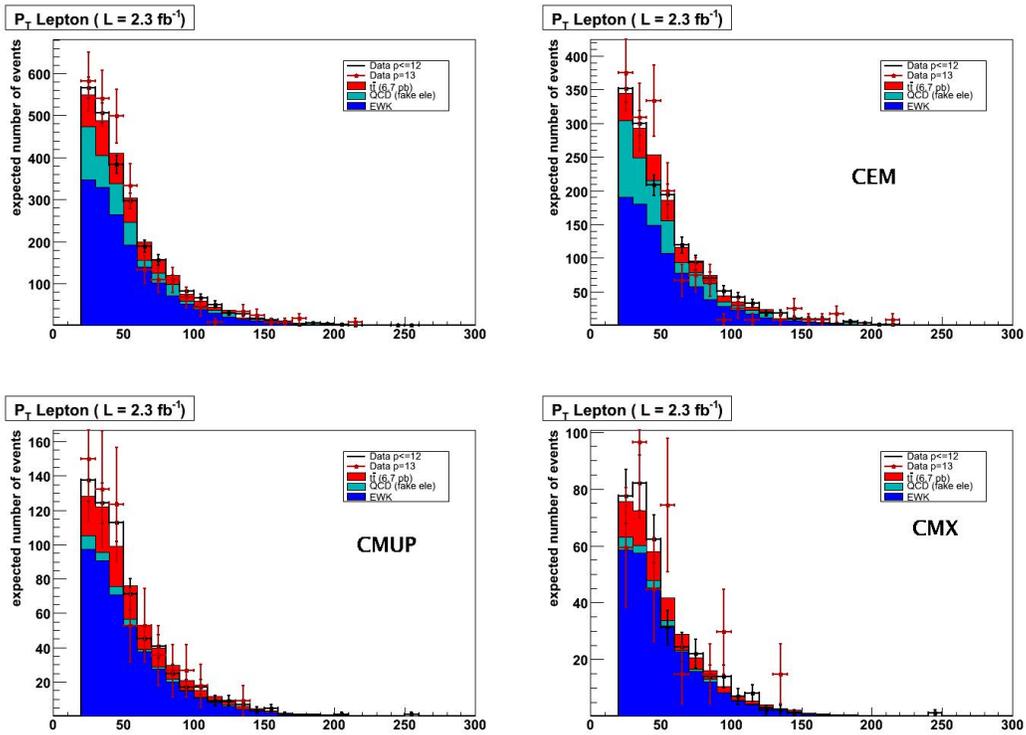


Figure 24: Leptons p_T distributions in the 3-jet control region comparing the p13 data with the p12 and earlier data; Top Left: All leptons; Top Right:CEM; Bottom left: CMUP ; Bottom Right: CMX

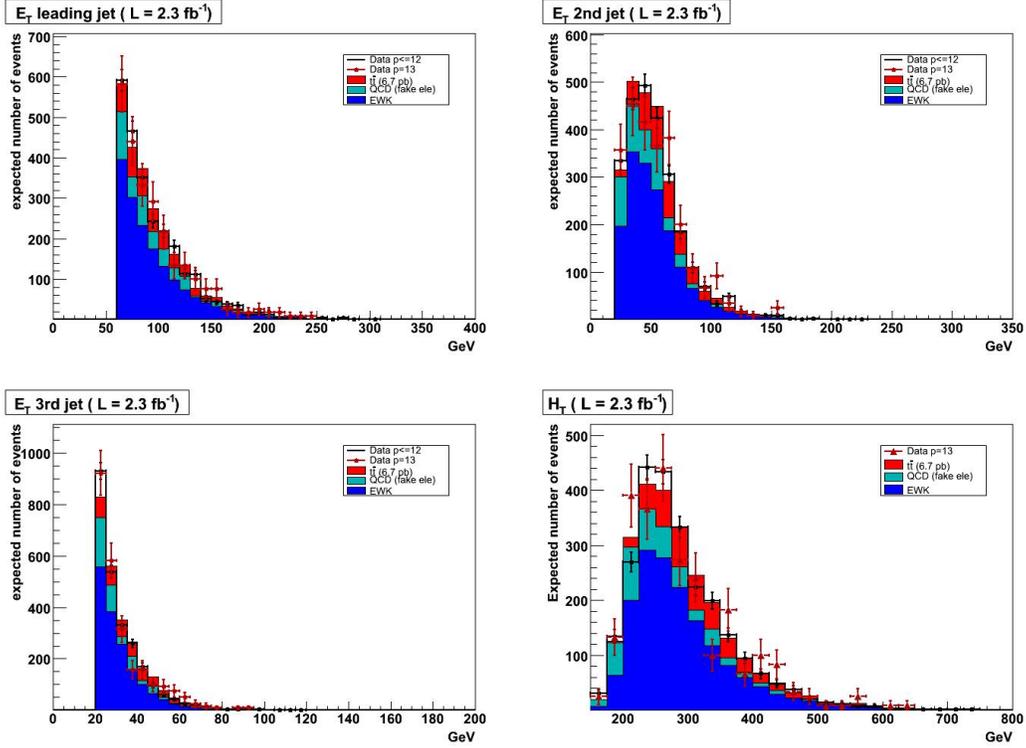


Figure 25: Jet E_T distributions and event H_T distributions in the 3-jet control region comparing the p13 data with the p12 and earlier data; Top Left: Leading jet; Top Right: 2nd jet; Bottom left: 3rd jet ; Bottom Right: H_T

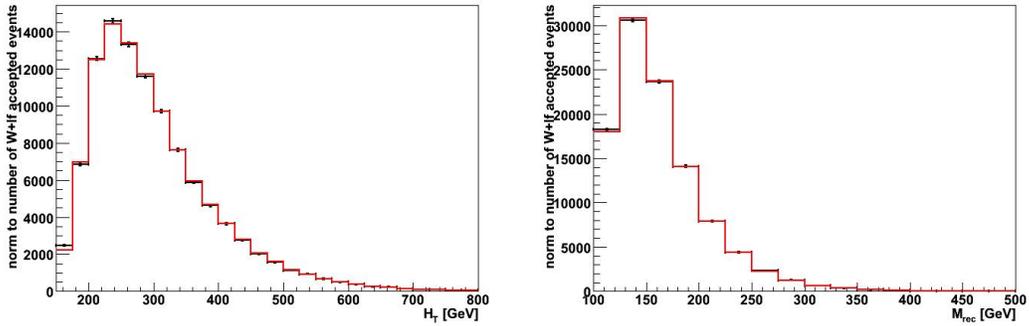


Figure 26: Templates used in the kinematic fitter. Comparing the predictions from W+lf MC only to that combining all background. Left: M_{rec} ; Right: H_T

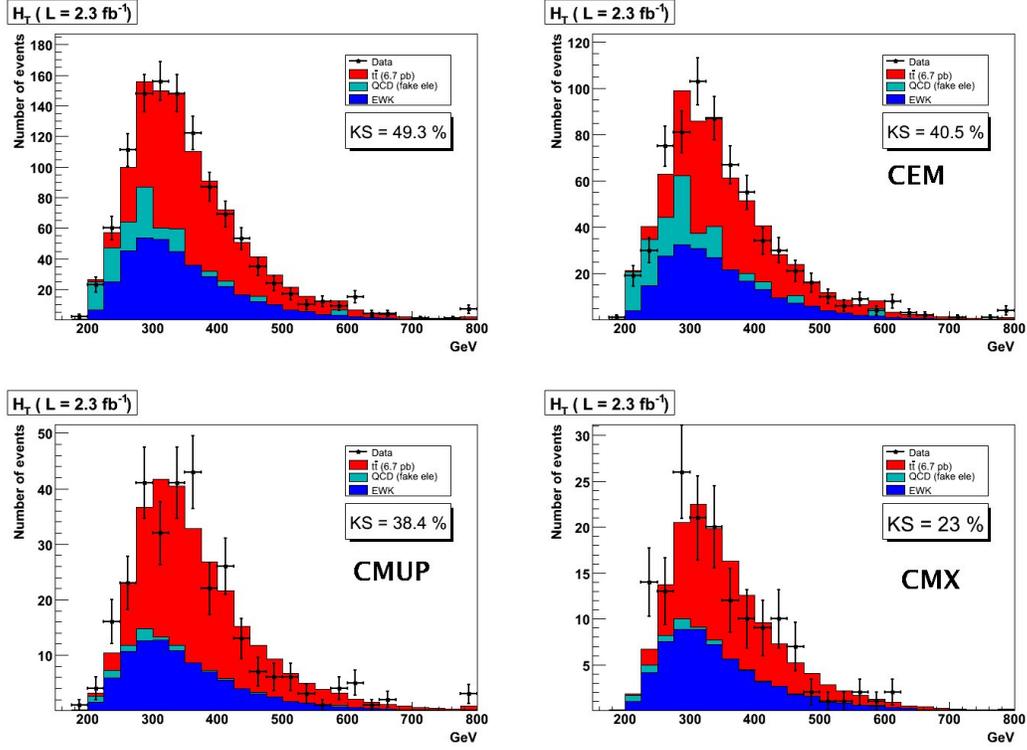


Figure 27: H_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

- E_T of the 2nd jet, figures 35 and 35
- E_T of the 3rd jet, figures 37 and 37
- E_T of the 4th jet, figures 39 and 39
- Missing E_T of the event, figures 41 and 42
- $\Delta\phi$ between the missing E_T and the leading jet, figure 43
- $\Delta\phi$ between the missing E_T and the lepton, figure 44

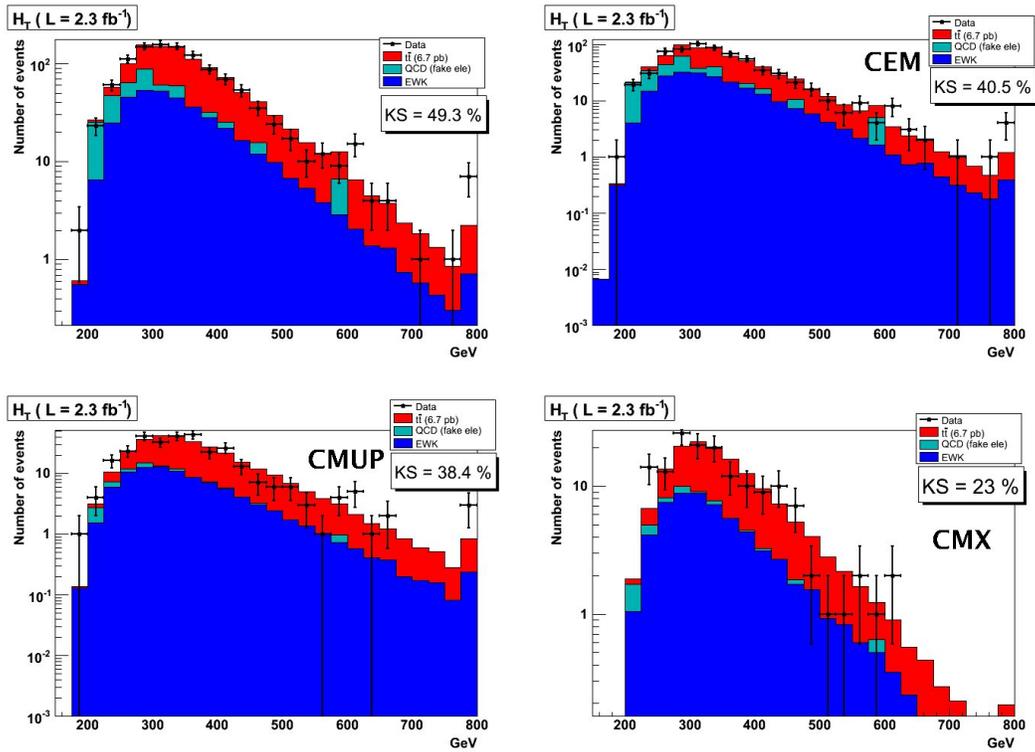


Figure 28: log-scale plots of the H_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

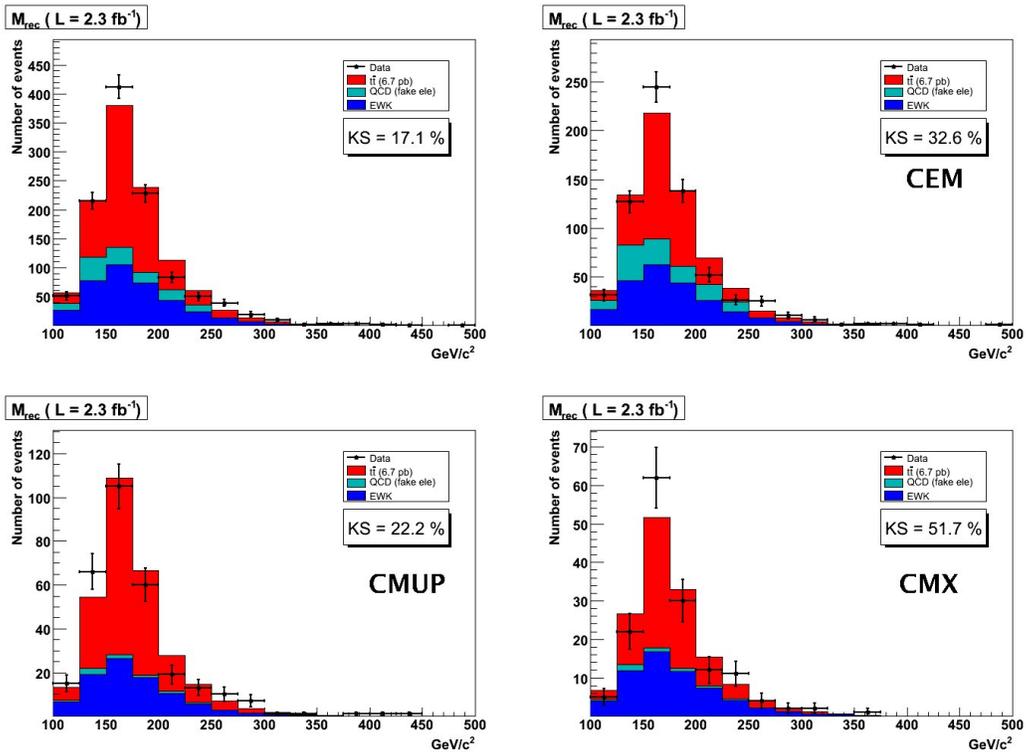


Figure 29: M_{rec} distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

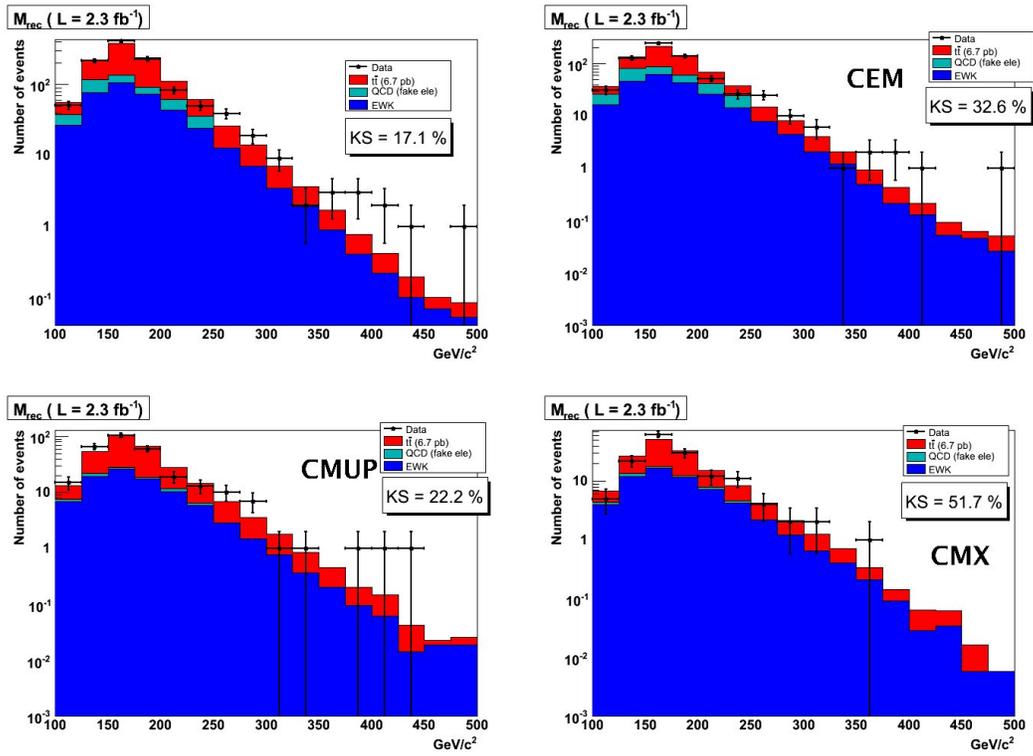


Figure 30: log-scale plots of the M_{rec} distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

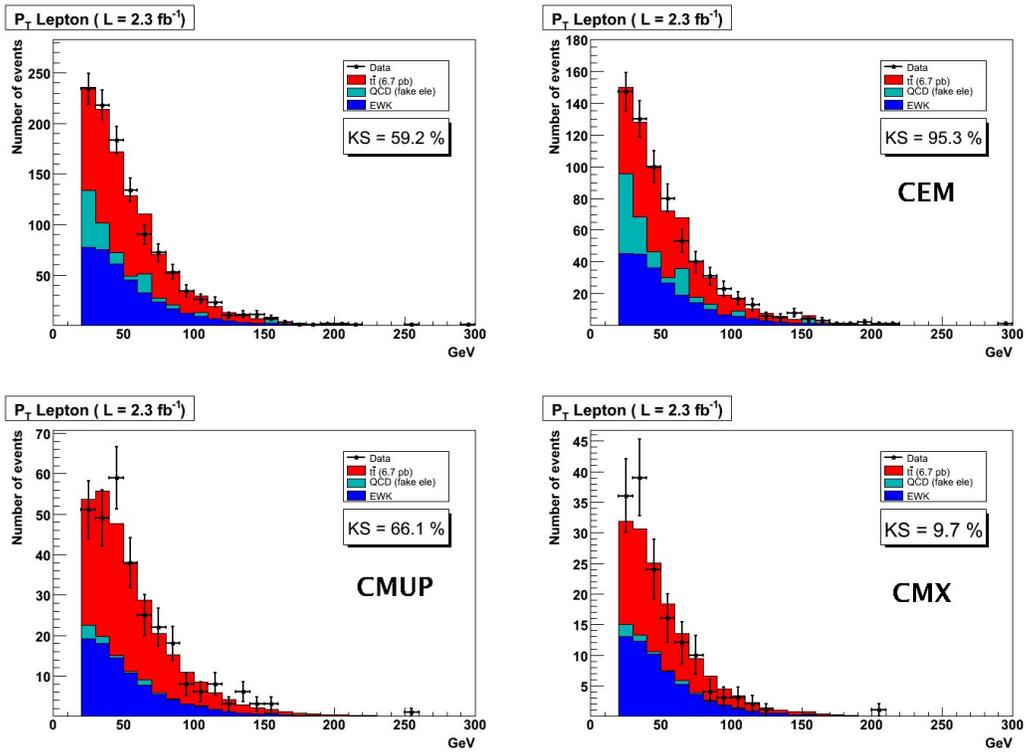


Figure 31: Lepton p_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

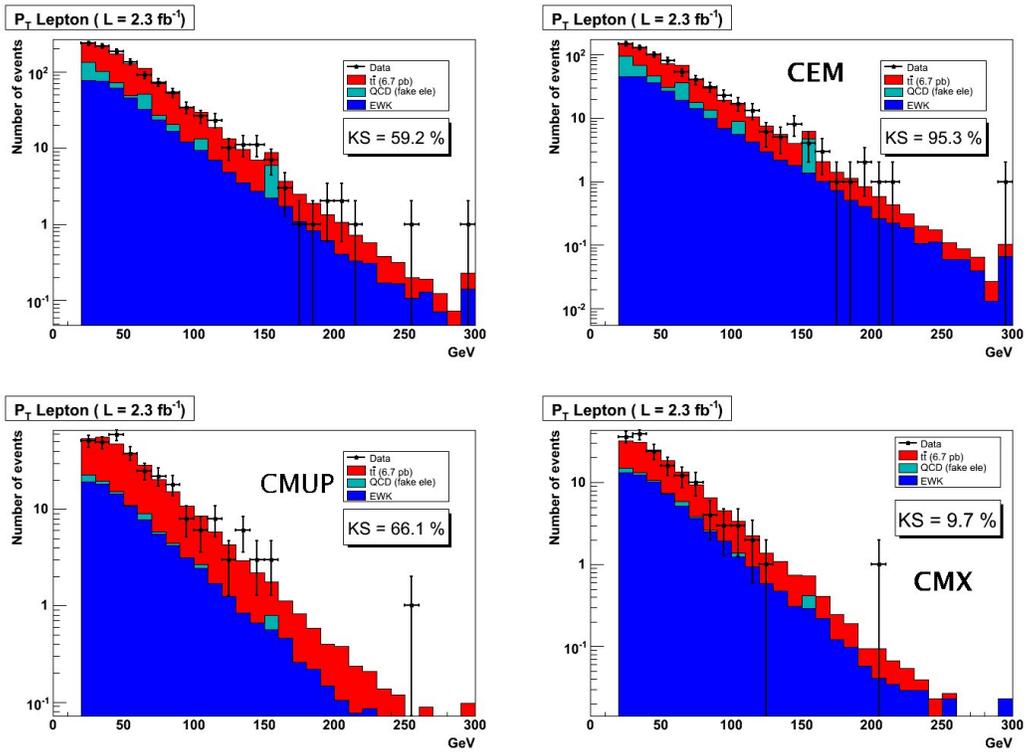


Figure 32: log-scale plots of the lepton p_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

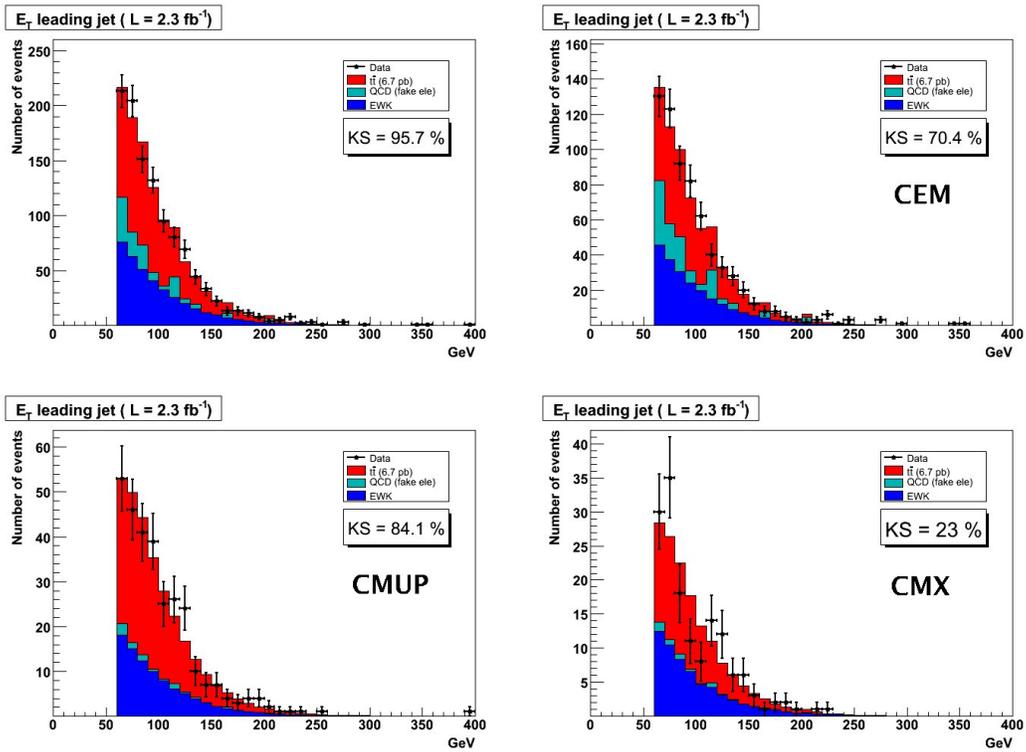


Figure 33: Leading jet E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

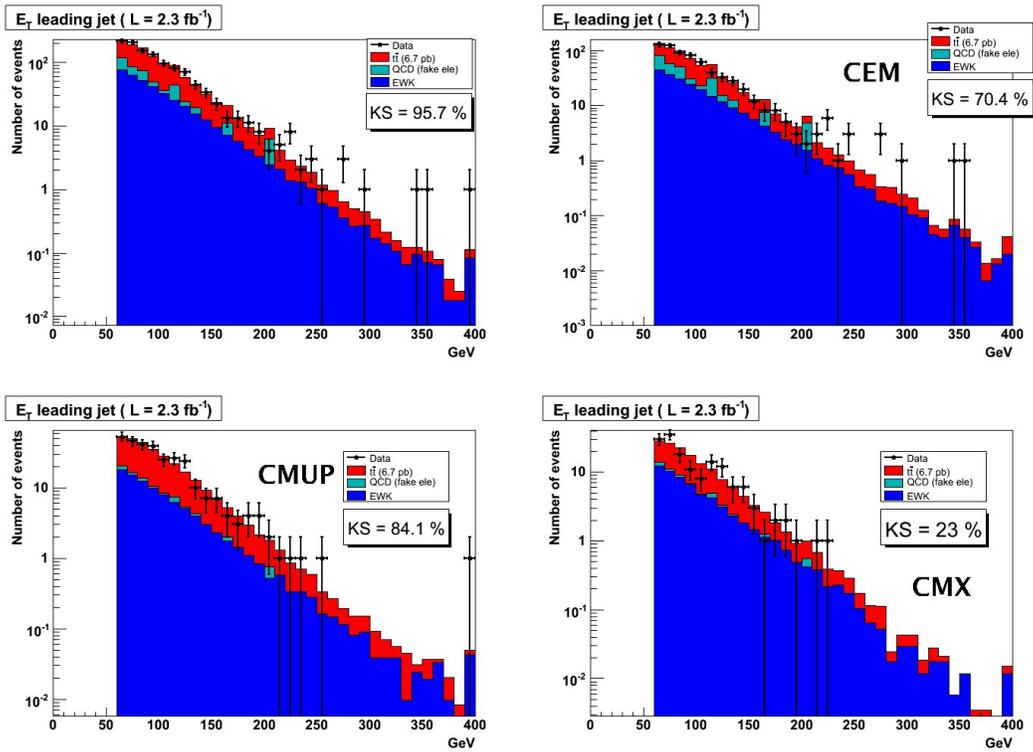


Figure 34: log-scale plots of the leading jet E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

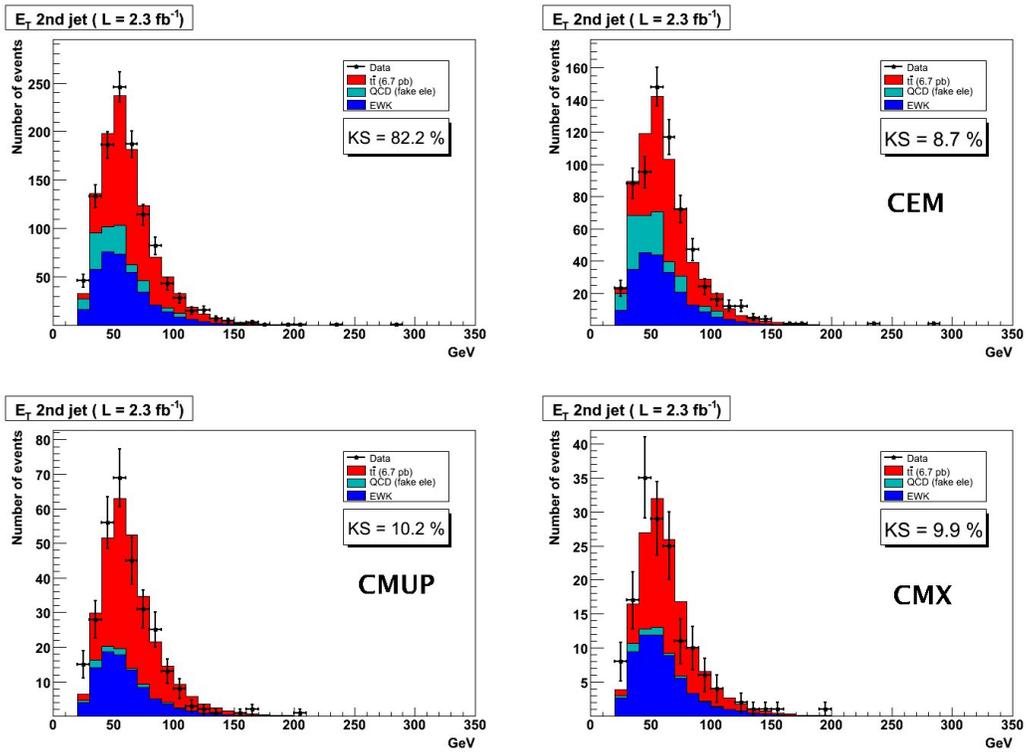


Figure 35: 2nd jet E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

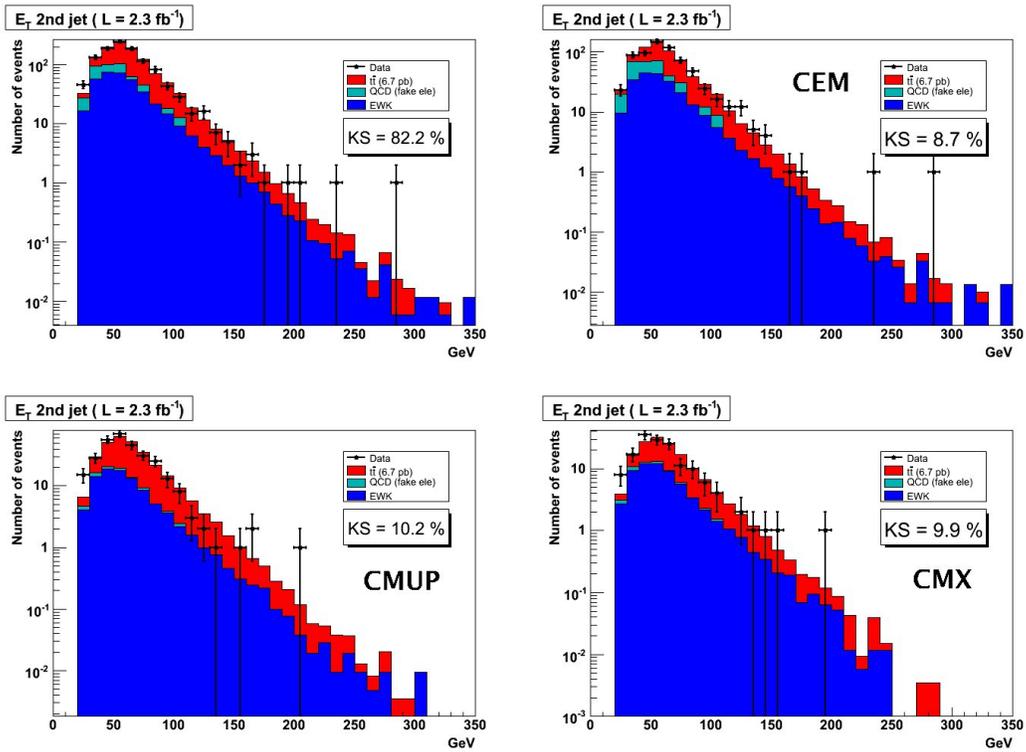


Figure 36: log-scale plots of the 2nd jet E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

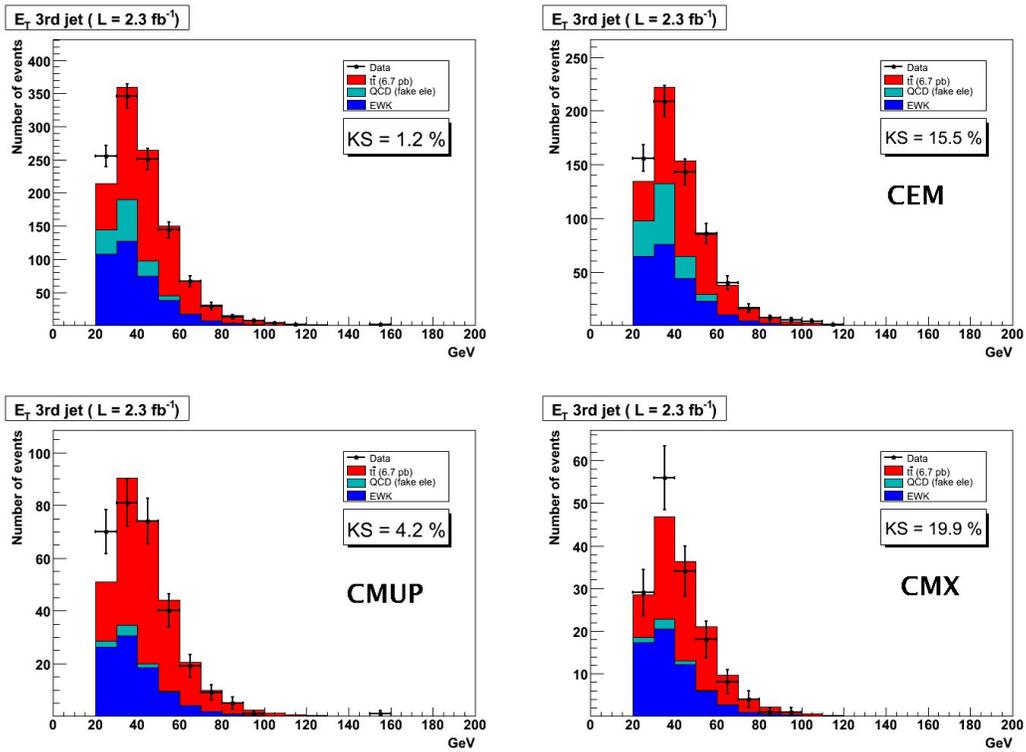


Figure 37: 3rd jet E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

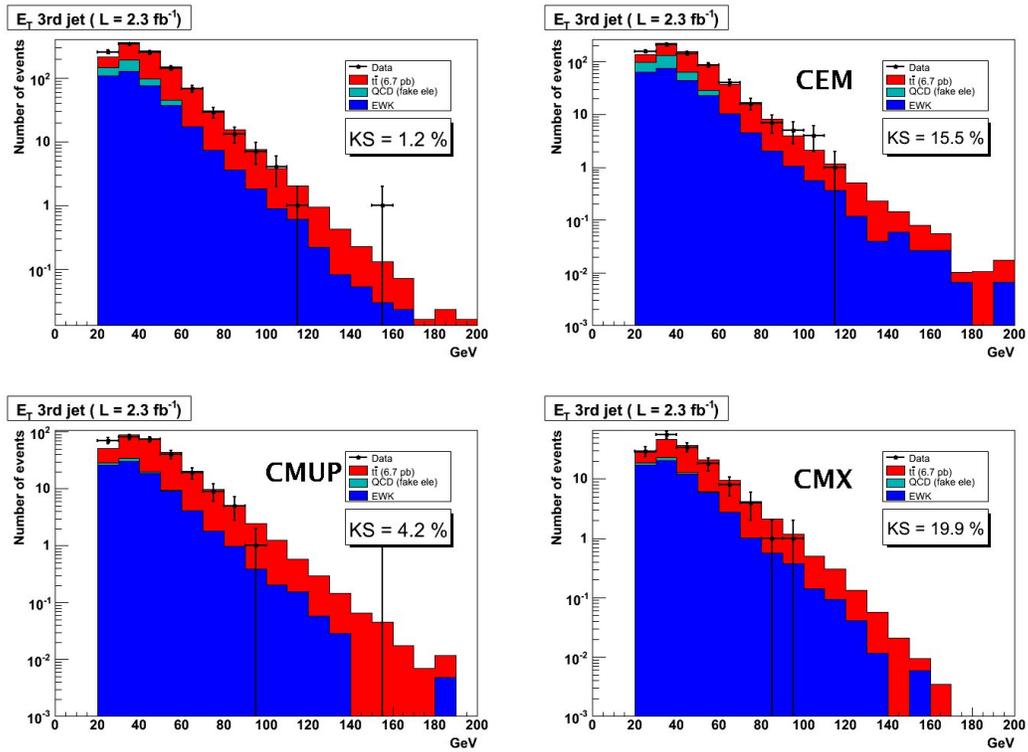


Figure 38: log-scale plots of the 3rd jet E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

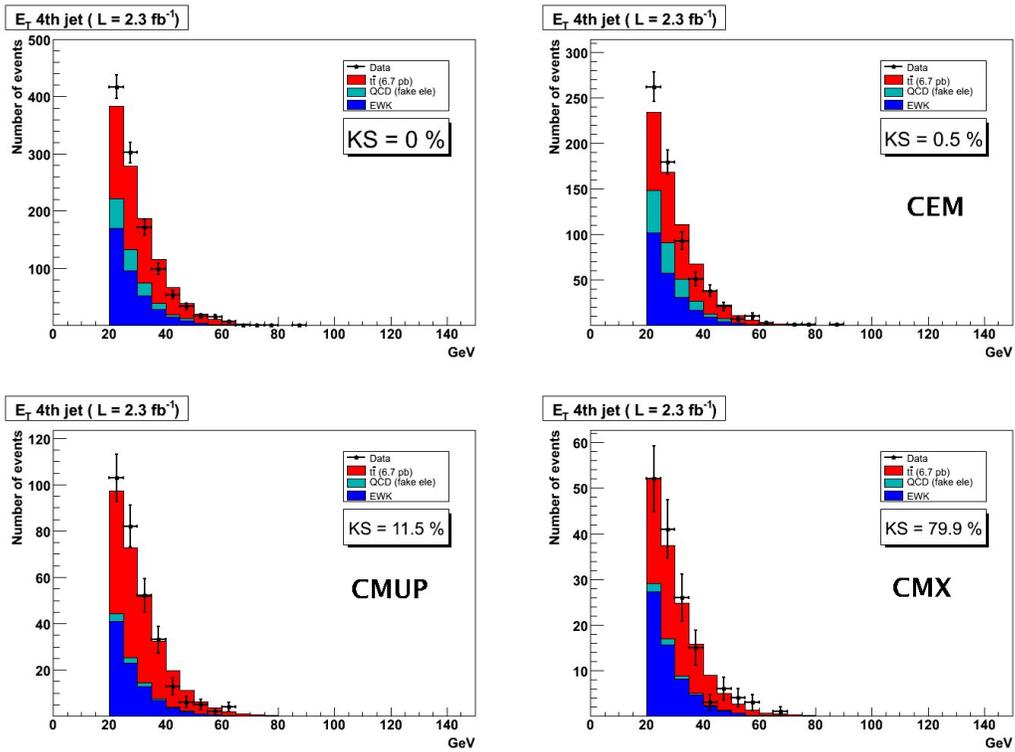


Figure 39: 4th jet E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

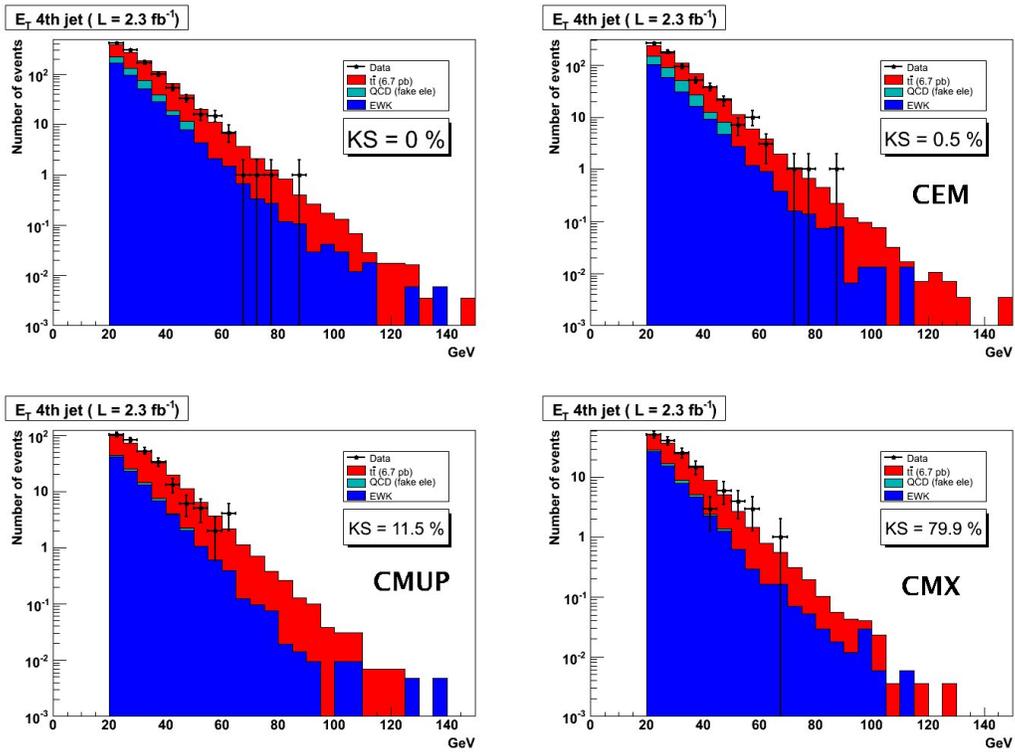


Figure 40: log-scale plots of the 4th jet E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

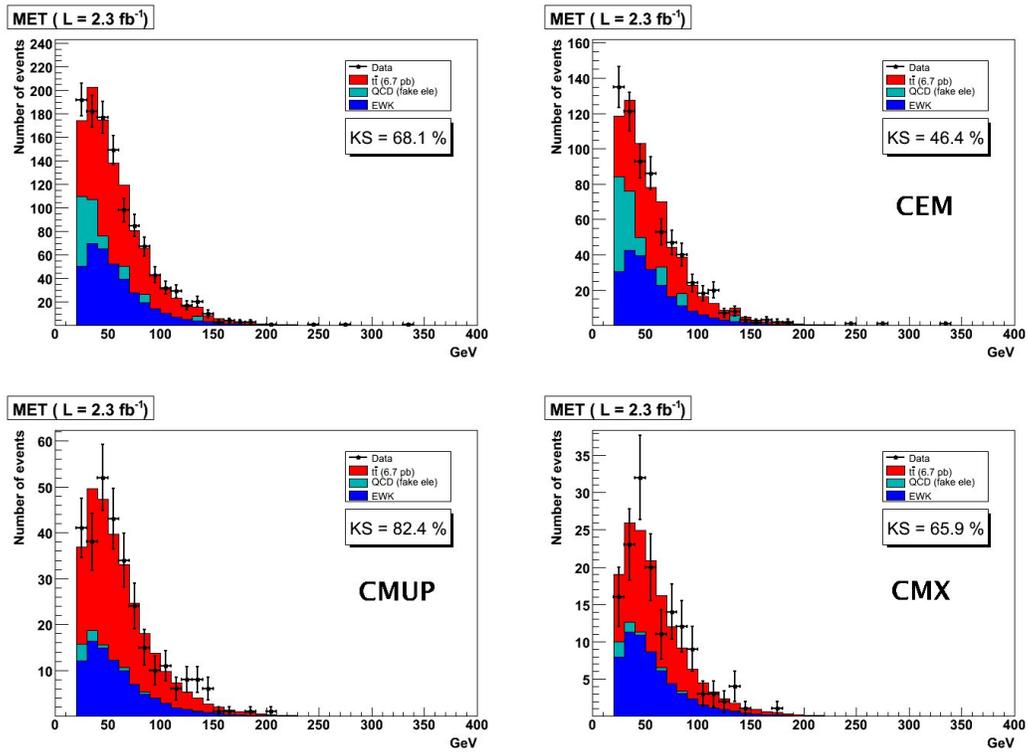


Figure 41: Missing E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

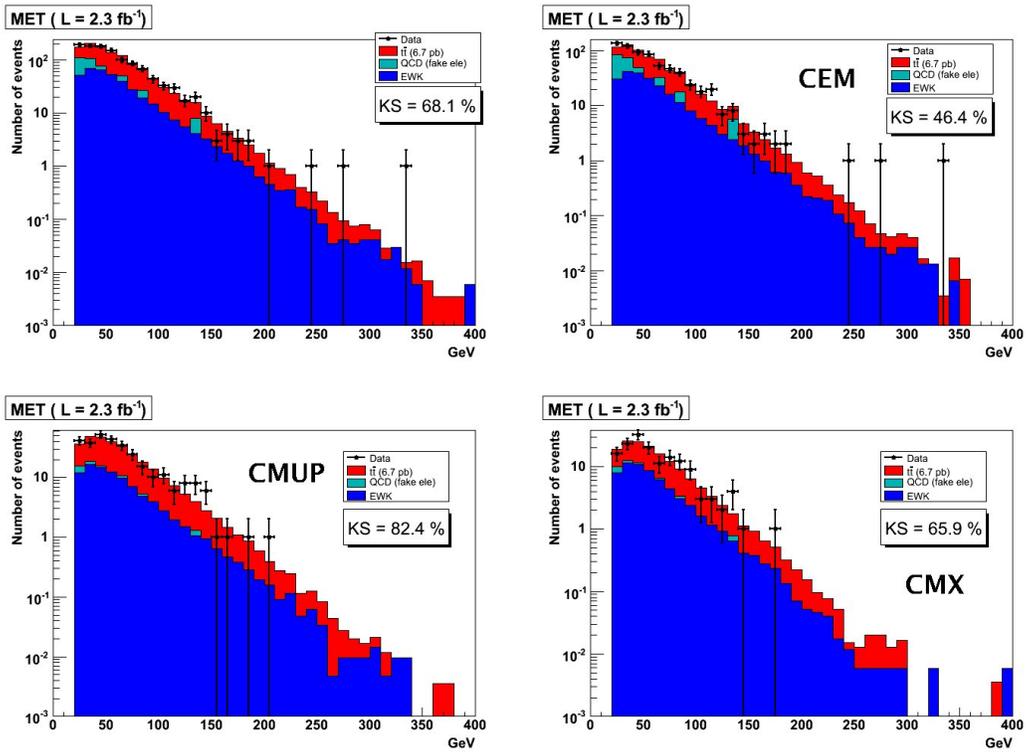


Figure 42: log-scale plots of the missing E_T distribution in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

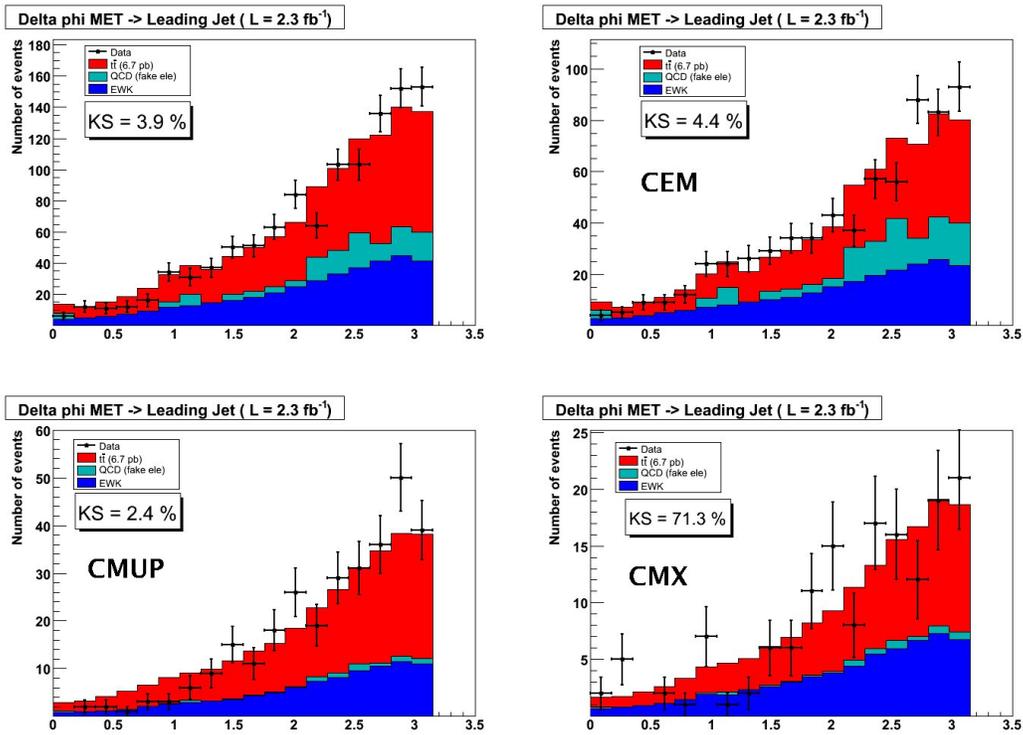


Figure 43: Distribution of the $\Delta\phi$ between the leading jet and the missing E_T in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

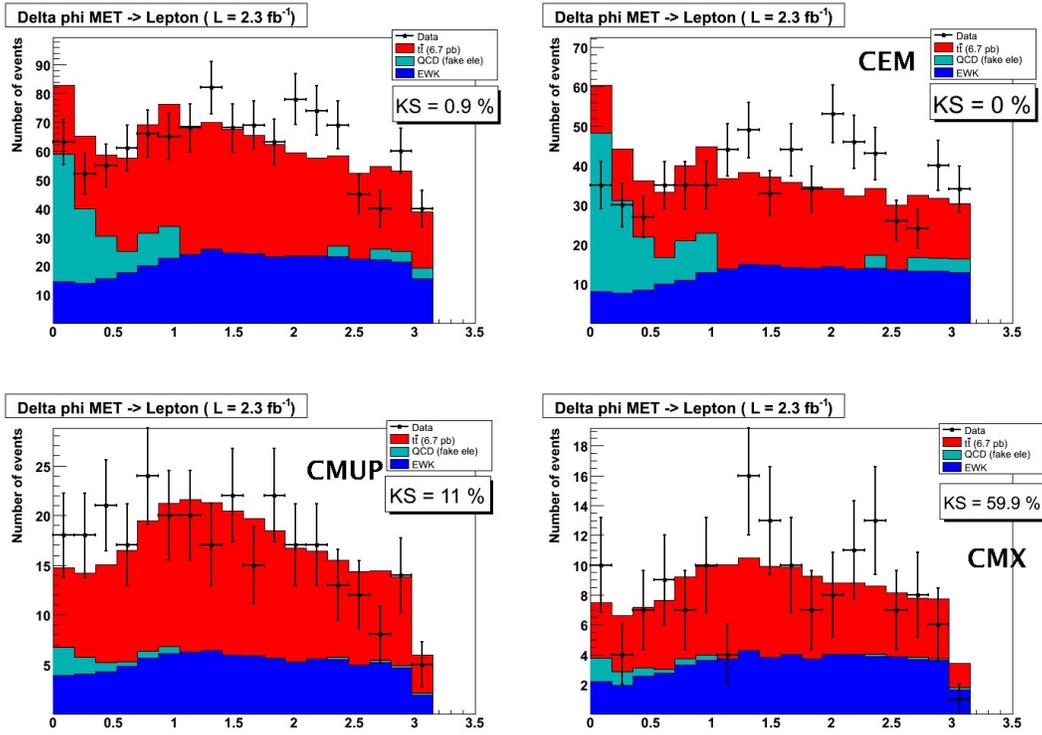


Figure 44: Distribution of the $\Delta\phi$ between the lepton and the missing E_T in signal region; Top Left: inclusive leptons; Top Right: CEM; Bottom left: CMUP; Bottom Right: CMX.

10 Systematic Errors

Note: the W +jets Q^2 systematic is being revisited but as we are still waiting for the MC samples, the old numbers are still used for now. Because we believe that the systematic should decrease compared to the previous iteration, we cut the systematic by factor of approximately 3. We do the same for the ISR/FSR systematics, for which we just finally ntupled and stripped the samples into our format. We should have these ready for blessing.

subsectionJet Energy Scale

The sensitivity to t' depends on knowing accurately the distribution of (H_T, M_{reco}) . Therefore the largest source of uncertainty comes from the factor that has the greatest effect on the shape of the kinematic distribution, which is due to the jet energy scale. Jets in the data and Monte Carlo are corrected for various effects (up to Level 5) as described in [19], leaving some residual uncertainty.

This uncertainty results in possible shift in the H_T and M_{reco} distributions for both new physics and Standard Model templates. We take this effect into account by generating templates with energies of all jets shifted upward by one standard deviation (+1 templates) and downward (-1 templates) respectively.

We then use a template morphing technique we developed in 2005 (in parallel and independently with that used in CLs, so somewhat different than that). We interpolate and extrapolate the expectation value μ_i at each bin i as follows:

$$\mu_i = \mu_{0,i} + \nu_{JES} \cdot (\mu_{+1,i} - \mu_{-1,i})/2 \quad (7)$$

where $\mu_{0,i}$ is the nominal expectation value, $\mu_{-1,i}$ and $\mu_{+1,i}$ are the expectation values from (-1) and (+1) templates respectively,

and ν_{JES} is the nuisance parameter representing a relative shift in jet energy scale:

$$\nu_{JES} = \frac{\Delta_{JES}}{\sigma_{JES}} \quad . \quad (8)$$

It enters the likelihood (6) as a Gaussian constraint penalty term: $G(\nu_{JES}|0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\nu_{JES}^2/2}$. This treatment of the systematic uncertainty in the likelihood is called template morphing method.

10.1 W +jets Q^2 Scale

[Note again: This effect was evaluated with Gen 5 MC. Samples are being generated to re-evaluate with AlpGen v2 in Gen 6.]

The effect in the choice of the appropriate Q^2 scale for W +jets production is evaluated by measuring the resulting change in the measured t' cross section given the t' existed. The expected 1σ change in the measured cross section is then interpreted as the uncertainty on the t' cross section itself. We use pseudo-experiments to measure this shift by drawing pseudo-experiments from shifted templates and fitting them to the nominal distribution. The resulting effect is incorporated into the likelihood as an additive parameter to the t' cross section, so that t' contribution to the expectation value μ_i in bin i becomes

$$\mu_{i,t'} = L_{t'}(\sigma_{t'} + \nu_{Q^2})\epsilon_{i,t'} \quad , \quad (9)$$

where ν_{Q^2} is constrained by a Gaussian with a width, which is a half of the average of two largest shifts for each mass of the t' .

$m(t')$	Q^2 scale			$\sigma_{\nu_{Q^2}}$
	$4m_W^2$	$m_W^2/4$	$\langle p_T^2 \rangle$	
175	1.15	1.35	0.26	1.25
200	0.67	0.86	0.07	0.76
225	0.31	0.53	0.05	0.42
250	0.17	0.30	0.02	0.23
275	0.16	0.23	0.02	0.19
300	0.12	0.18	0.02	0.15
350	0.08	0.10	0.02	0.09
400	0.06	0.06	0.01	0.06

Table 3: Shift (in picobarns) in apparent t' cross section due to actual W +jets Q^2 scale being different from the nominal scale assumed, $Q^2 = m_W^2 + \sum_{jets} p_T^2$. The right column shows the width of the ν_{Q^2} used in the fit.

We estimated uncertainties associated with the Q^2 scale choice for the $t\bar{t}$ production processes, as well as for the t' , using t' MC samples with Q^2 scale equal to half and double of $m_{t'}$ at 175 GeV. The errors were found to be negligible compared to those of W + jets.

10.2 ISR and FSR

[Note: This will be re-evaluated with new ttbar and t' (at several mass values) shifted samples, where ISR and FSR and shifted up together and down together, as is the new top group prescription. We expect to have this ready soon, probably by preblessing. The hold-up was MC sample generation and CAF problems.]

In Gen 5 we investigated the effect of varying ISR and FSR using the prescription from the Top Mass group [20]. We generated samples of "more ISR" and "more FSR" for t' at 200 GeV and 300 GeV, using the same tcl prescription as it was done for the $t\bar{t}$ ISR/FSR Monte Carlo samples. These two mass points bracket our region of t' mass where we expect to set a limit.

We made templates for each of these, and then generated pseudo-experiments with the shifted (more templates) for ISR and FSR, using shifted $t\bar{t}$ samples at the same time. We then fit them to the nominal distribution, just as we do for the Q^2 systematic. The resulting plots are shown in Figure ??.

We can see that the "more FSR" case is consistent with zero cross section shift, therefore we do not include a systematic for that. In the "more ISR" case, we see a shift of 0.4 pb for 200 GeV and at 300 GeV it's consistent with zero shift. We expect our limit to be in the range between these two points, so we average the effect, and take a systematic error of 0.2 pb. We add that to the Q^2 error in the likelihood.

10.3 QCD Background

The QCD background shape is obtained from stripped anti-electron data samples from the top group, and scaled based on the new method of fitting to the MET distribution (M24U code). We are investigating alternate shape samples using the non-isolated leptons, as this was used as a shape model in previous iterations. (We believe the kinematics may fit the data better, see previous section on kinematic cuts.)

The relative normalization uncertainty is taken to be 50%, as was done in the kinematic cross section analysis and the original $190pb^{-1}$ analysis, due to our lack of confidence that the fit to the MET distribution carries through a number valid for the other kinematic plots. With our QCD veto cuts, it turns out to change the fit a negligible amount whether we constrain QCD or let it float, but we choose to constrain it due to precedent. The uncertainty is represented by a Gaussian-constrained parameter in the likelihood. The QCD background has a negligible effect on the t' limit.

10.4 Integrated Luminosity

The integrated luminosity uncertainty is taken to be 6.0% [8], and represented by an additional Gaussian-constrained parameter multiplying all contributions except for the QCD background, which is normalized independently.

10.5 Lepton ID

We have two components for lepton ID. First is the efficiencies for the individual electrons and CMUP and CMX muons from the official numbers out of joint physics [?]. We multiply each lepton type by the associated efficiency and Gaussian constrain it within the error on the efficiency.

Second is the uncertainty in the lepton ID efficiency data/MC scale factor, which is 2% [23], and taken as correlated across lepton types since it is due to the presences of multiple jets in an event. We add it in quadrature with the luminosity error, which is also correlated across lepton types, and include it with a Gaussian constraint to the likelihood.

10.6 PDF Uncertainty

[Note: This section will be re-evaluated using the current method from Joint Physics. As this systematic has a very small effect on our final limits we do not expect the changes to be significant with respect to the current version.]

The Parton Distribution Functions (PDFs) are not precisely known, and this uncertainty leads to a corresponding uncertainty in the predicted cross sections, as well as the acceptance. We are mainly interested in how the acceptance varies with varying PDFs. The proper way to carry out this study — generating separate sets of Monte Carlo events for each PDF set one wanted to investigate — would be prohibitive. Instead, we follow the prescription advocated by the Top Quark group [24] and re-weight the same set of generated events with different PDF sets, and observe the cross section and acceptance changes.

Variations due to CTEQ5, CTEQ6

We use Pythia 6.216 for all our Monte Carlo generation, which uses CTEQ5L PDFs. We investigated 41 CTEQ6 eigenvectors. We report the differences between CTEQ6L and CTEQ6L1, and between CTEQ6M (NLO) and CTEQ5L (LO).

The weights are calculated for each event using the HEPG 4-vectors, and the Q^2 , which we set to $M_{t'}^2$.

Mass	CTEQ5/6	CTEQ6L1-CTEQ6L	α_s variation
175	0.0019	0.0013	0.0026
250	0.0056	0.0013	0.0002
350	0.0153	0.0071	0.0141
400	0.0202	0.0071	0.0249
500	0.0090	0.0001	0.0016

Table 4: Comparison of various PDFs

We determine the acceptance by demanding that one HEPG level lepton have p_T greater than 20 GeV/c and be contained within the fiducial region of the detector. We also demand that the jets from the t' quarks be similarly contained. Note that this is different from the Gen4 demand of reconstructed tight lepton, 4 or more tight jets and $\cancel{E}_T > 20$ GeV. The Gen4 acceptance criteria involved reconstruction efficiency in the PDF calculation, which is not necessary. In Gen5, we calculate PDF error using HEPG information only.

Table 4 shows how the acceptance ratios differ from unity when various PDFs are compared. We see that largest change in acceptance, about 2% occurs for a t' mass of 400 GeV/ c^2 , but there is no straightforward mass dependence of the PDF differences. We observe negligible effect due to the 41 eigenvectors of CTEQ6 and ignore them for this determination.

Variations due to α_s

Another significant source of PDF uncertainty is the variation in α_s . As in other analyses, including several top-quark analyses, we have used the difference between MRST72 and MRST75¹ to determine the systematic due to α_s uncertainty, which is shown in the last column of Table 4.

Final PDF Systematic

In order to arrive at a final systematic uncertainty for the PDFs, we take the mean of each of the three columns in table 4. We get 0.0104 for the CTEQ5/6 difference, 0.0034 for the CTEQ6L1/CTEQ6L difference, and 0.009 for the α_s variation. Since the first number quantifies difference between LO and NLO effects, we take only the last two in quadrature and get 0.962, or 1.0%

10.7 Theory Uncertainty

The theory uncertainty in the t' cross section is about 10% (see Table 5), which is mainly due to uncertainty in PDFs ($\sim 7\%$). The other effect comes from uncertainty in the choice of the Q^2 scale [16].

We take the theory uncertainty in $t\bar{t}$ cross section fully correlated with the one of t' , and introduce it into the likelihood as a single nuisance parameter: $\nu_{theory} = \nu_{theory}(m'_t)$, which is the same parameter used to constrain $t\bar{t}$ cross section to a theoretical value.

11 Results

We test the sensitivity of our method by drawing pseudo-experiments from Standard Model distributions, i.e. assuming no t' contribution. The range of the expected 95% CL upper limits with one standard deviation bandwidth is shown in Figure 45. The purple curve is the theory curve [15, 16], the values of which are given in Table 5. The lower σ_{min} and upper σ_{max} limits are obtained using

¹This schema refers to MRST LO central gluon and MRST LO lower α_s from Eur. Phys. J C4 (1998)

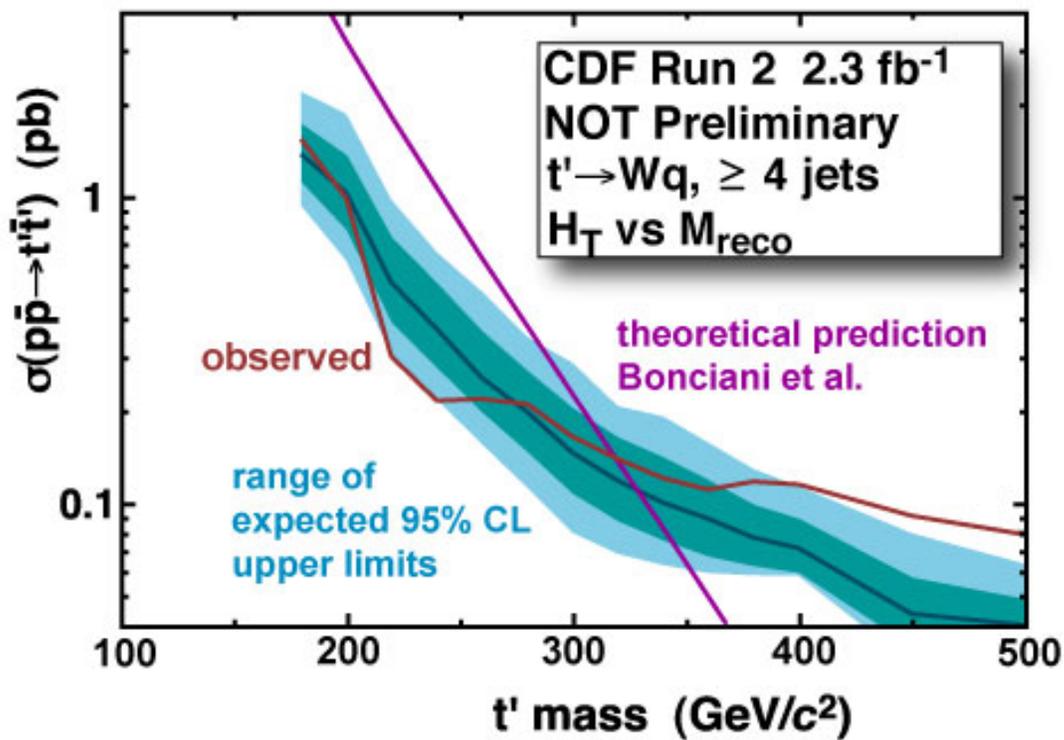


Figure 45: Upper limit, at 95% CL, on the production rate for t' as a function of t' mass (red). The purple curve is a theoretical cross section. The blue band is the range of expected 95% CL upper limits within one standard deviation.

$m(t')$ (GeV)	σ_{min} (pb)	σ_{center} (pb)	σ_{max} (pb)
180.0	4.9938	5.7476	6.2396
200.0	2.7815	3.1898	3.4525
220.0	1.5926	1.8236	1.9710
240.0	0.9299	1.0647	1.1515
260.0	0.5499	0.6302	0.6828
280.0	0.3281	0.3769	0.4096
300.0	0.1968	0.2268	0.2475
320.0	0.1183	0.1370	0.1502
340.0	0.0711	0.0828	0.0914
360.0	0.0426	0.0500	0.0555
380.0	0.0255	0.0301	0.0337
400.0	0.0152	0.0181	0.0204

Table 5: Theory values of t' cross section for given mass [15, 16].

the CTEQ6M family of parton density functions with uncertainties, together with the study of the scale uncertainty [25].

From Figure 45 it follows that given no t' presence, this method is on average sensitive to setting an upper limit at 315 GeV t' mass.

11.1 Pull Distributions

Before looking at the data, we investigated out fitting technique by looking at the pull distributions for pseudo-experiments generated with a fixed signal. The results are shown in figures 46 and 47. The expected pull mean should be zero and the width one. Any deviation from these values could indicate a bias in our measurement. In order to verify that our small shift in both the mean and the width do not bias our result, we carried out coverage tests explained in the next sub-section.

11.2 Coverage

We have checked the coverage of our 95% CL limit method for one t' mass, 300 GeV, as a function of the signal cross section. As Table ?? shows, the method over-covers by a few percent everywhere. This sort of behavior is not atypical of Bayesian methods.

We perform the analysis fit on the data and determine upper limits on the t' signal.

The red curve in Figure 45 shows the final result, expressed as a 95% CL upper limit on the t' production rate as a function of t' mass. Table 7 shows the individual calculated limits along with expected limits from pseudo-experiments for reference.

2D-distribution of (H_T, M_{reco}) is shown in Figure ??.

Based on these results we exclude at 95% CL the t' quark with mass below 315 GeV, given the true top mass is 175 GeV. Of course, our measurement of the top mass may have been affected by the presence of a higher mass t' and thus we should treat these conclusions with care.

To determine if the data show any evidence of an excess far out in the tails of H_T and M_{reco} , we decided *a priori* to count the number of events in groups of $n \times n$ of our standard 25 GeV bins in these quantities, and compare with the number predicted from a zero-signal fit to the full two dimensional spectrum. For each $n \times n$ bin one can then calculate the p-value for having observed that number or

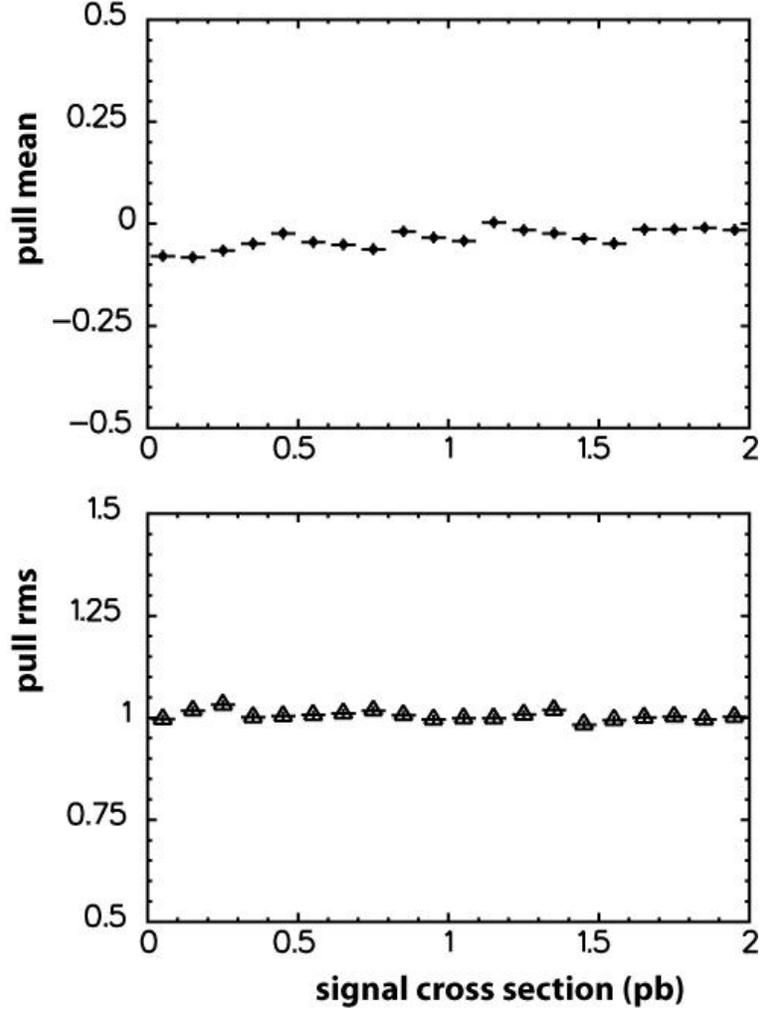


Figure 46: Mean value (top) and error on the mean (bottom) of the Pull distributions as a function of the t' cross-section for a mass of 300 GeV.

$\sigma_{t'\bar{t}'}$ (pb)	coverage
0.10	0.998 ± 0.001
0.15	0.982 ± 0.004
0.20	0.987 ± 0.004
0.25	0.972 ± 0.005
0.30	0.967 ± 0.006
0.35	0.971 ± 0.005
0.40	0.975 ± 0.005
0.45	0.978 ± 0.005
0.50	0.985 ± 0.004

Table 6: Frequentist coverage of the 95% CL upper limit method for $m(t') = 300$ GeV.

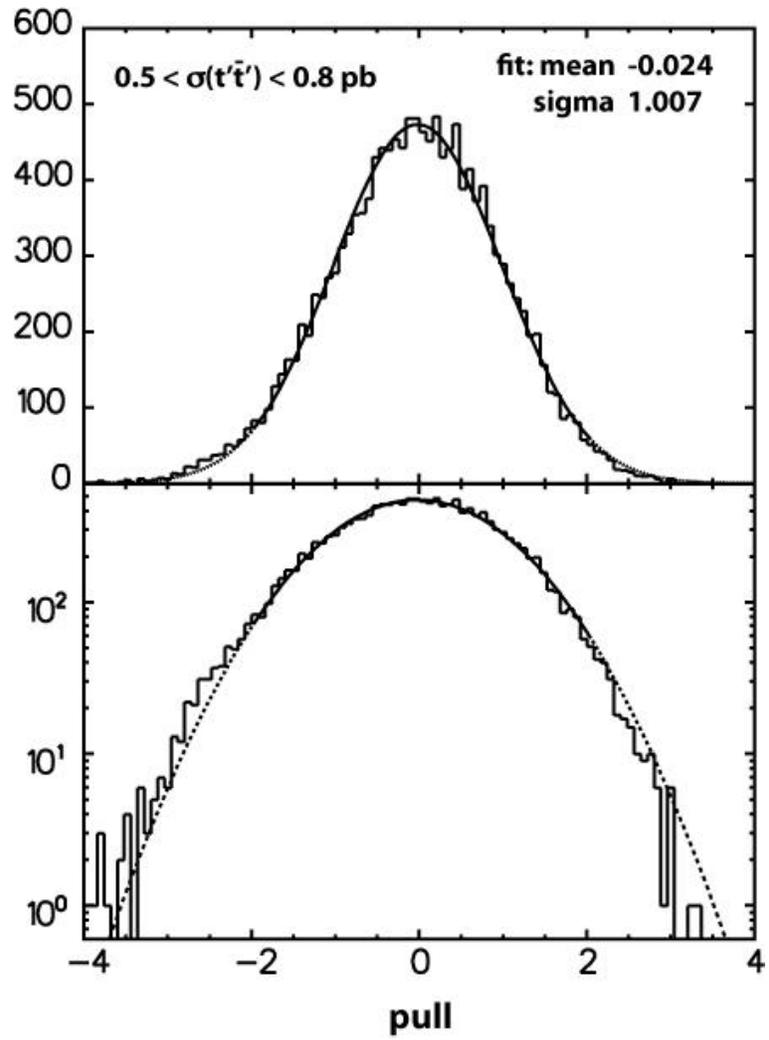


Figure 47: The pull distribution is shown for the cross-section region between 0.5 and 0.8 pb which is the region preferred by the theory.

$m(t')$ (GeV)	expected limit (pb)					observed limit (pb)
	-2σ	-1σ	median	$+1\sigma$	$+2\sigma$	
180	0.93096	1.11332	1.37139	1.71439	2.18627	1.53387
200	0.61944	0.77676	1.02488	1.35297	1.84467	0.98504
220	0.30943	0.38754	0.52473	0.73229	0.96973	0.30090
240	0.21820	0.27521	0.36950	0.51733	0.65478	0.21559
260	0.15602	0.19785	0.25544	0.36149	0.48792	0.21843
280	0.11228	0.14673	0.19763	0.26791	0.34529	0.20938
300	0.08026	0.10720	0.14594	0.20300	0.28566	0.16449
320	0.06834	0.08738	0.11811	0.16331	0.20643	0.13927
340	0.06271	0.07546	0.10013	0.14038	0.19063	0.12019
360	0.05913	0.06708	0.08869	0.11851	0.15690	0.11069
380	0.05818	0.06223	0.07730	0.09840	0.12881	0.11721
400	0.05795	0.05908	0.07097	0.08835	0.11266	0.11475
450	0.03033	0.03382	0.04341	0.05683	0.07932	0.09079
500	0.02599	0.02852	0.03611	0.04835	0.06303	0.07873

Table 7: Expected and obtained limits on t' production cross section for given mass.

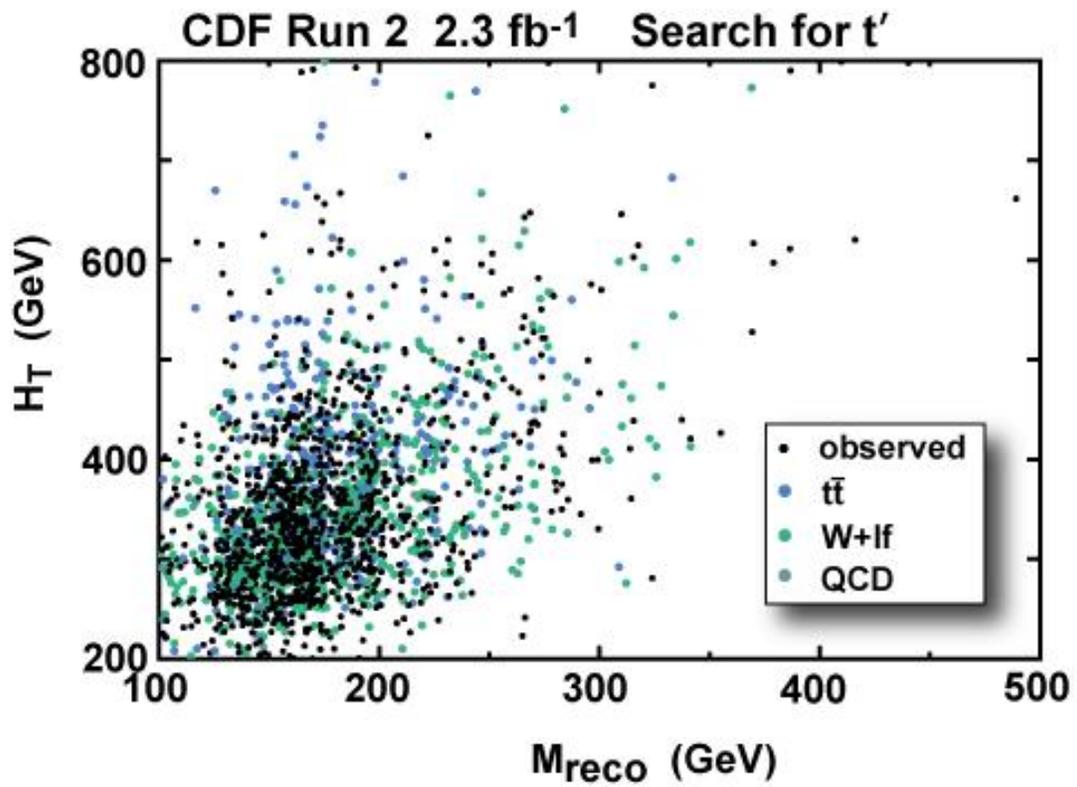


Figure 48: 2D plot of H_T vs M_{rec} distribution showing the data (black points) and the fitted number of background events: QCD (dark cyan triangles), W +jets (blue open circles) and $t\bar{t}$ (red triangles)

n	observed	expected	p-value
1	0	0.017	1.0000
2	0	0.067	1.0000
3	1	0.172	0.1583
4	2	0.300	0.0369
5	3	0.602	0.0233
6	4	1.257	0.0389
7	4	2.493	0.2410
8	11	4.773	0.0100
9	14	8.588	0.0550
10	23	15.74	0.0505

Table 8: Number of observed events in the highest $n \times n$ bins of H_T and M_{reco} , compared with the prediction from a zero-signal fit to the full spectrum. For each value of n the table shows the p-value, the probability for observing at least what was actually observed or more, given the number expected.

greater, given the prediction. If a significant effect is observed, one can calculate an overall p-value which is the probability that one would observe a p-value at least as significant as the most significant $n \times n$ bin or greater; this can take into account both the trials factor and the effect of systematic errors.

Table 11.2 shows the result with the real data. The most significant $n \times n$ bin is that where $n = 8$; the probability for observing 11 or more events given 4.773 expected is 0.01. (This assumes systematic uncertainty on the background.) We have not at this stage calculated the overall p-value for observing an $n \times n$ bin with a significance this great or greater, but it will have a significance on the order of 2σ . Thus we conclude there is no statistically significant excess in the far tails of H_T and M_{reco} .

References

- [1] R. Erbacher *et al.*, CDF note 6888.
- [2] A. Ivanov *et al.*, CDF note 7912.
- [3] J. Conway *et al.*, CDF note 8081.
- [4] CDF Collaborataion: T. Aaltonen *et al.* [arXiv:hep-ex/08013877].
- [5] D. Choudhury, T. Tait and C. Wagner [arXiv:hep-ph/ 0109097].
- [6] T. Han *et al.*, Phys. Lett. B **563** (2003) 191 .
- [7] Good Run List.
<http://www-cdf.fnal.gov/internal/dqm/goodrun/good.html>
- [8] Integrated luminosity calculation by period:
http://www-cdf.fnal.gov/internal/physics/top/RunIITopProp/gen6Sum06/lumi_v19.html
- [9] Top MC web-page:
<http://www-cdf.fnal.gov/internal/physics/top/RunIIMC/topmc6/index.shtml>
- [10] Event Selection for Lepton + Jets analyses.
<http://www-cdf.fnal.gov/tiki/tiki-index.php?page=LeptonJetsAcceptance1InvFb>
- [11] Presentation at the Top Properties Meeting March 2007
http://fcdfwww.fnal.gov/internal/WebTalks/Archive/0703/070316_topprop/05_070316_topprop_Will
- [12] K. Lannon *et al.*, CDF note 8063.
- [13] A. Ivanov
http://fcdfwww.fnal.gov/internal/WebTalks/Archive/0510/051014_topprop/03_051014_topprop_Andrew_Ivanov_1_muons_Oct13.ppt.pdf
- [14] N. Kidonakis and R. Vogt, Phys. Rev. D **68** 114014 (2003) [arXiv:hep-ph/0110145].
- [15] R. Bonciani, S. Catani, M. L. Mangano and P. Nason, Nucl. Phys. B **529** (1998) 424 [arXiv:hep-ph/9801375].
- [16] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, JHEP **0404** (2004) 068 [arXiv:hep-ph/0303085].
- [17] C. Plager *et al.*, CDF note 9161
- [18] J. Adelman *et al.*, CDF note 7532.
- [19] J. Adelman *et al.*, CDF note 7358.
- [20] Personal communication with Un-Ki Yang.
- [21] B. Han and V. Boisvert, CDF note 8629, 2006
D. Hare, E. Halkiadakis, and T. Spreitzer, CDF note 8614, 2007
D. Hare and E. Halkiadakis, CDF note 9148, 2008
- [22] U. Grundler, L. Lovas, and A. Taffard, CDF note 8618, 2006
S. Roli, L. Lovas, and E. James, CDF note 9085, 2007
- [23] M. McFarlane *et al.*, CDF note 7682.
- [24] Personal communication with Stephen Miller.
- [25] Personal communication with Michelangelo Mangano.