CDF/PUB/TOP/PUBLIC/x

Search for Heavy Top $t' \to Wq$ in Lepton Plus Jets Events in $\int \mathcal{L}dt = 4.6$ fb^{-1}

The CDF Collaboration

March 5, 2010

Abstract

We search for pair production of the heavy top (t') quarks pair decaying to Wq final states using 4.6 fb^{-1} data sample of lepton+jets collected using inclusive lepton and met+jets triggers.

We reconstruct the mass of the t' quark (M_{rec}) and perform a two dimensional-fit of the observed (H_T, M_{rec}) distribution to discriminate the new physics signal from Standard Model backgrounds. We exclude a Standard Model fourth-generation t' quark with mass below 335 GeV at 95%CL.

Contents

1 Introduction

The top quark is a relatively recent addition to the array of particles that can be produced in the laboratory. Since its discovery the top quark data collected at the Tevatron have been an active testing ground for the validity of the Standard Model (SM). The top quark is unique because of its large mass near 173 GeV, which distinguishes it from the other fermions of the SM and is similar to the masses of the weak force carriers (W and Z) and the expected mass range for the proposed Higgs boson.

Because of the large top quark mass, the top quark final decay products are very energetic. The leptons and jets from top decays have on average higher transverse momenta as compared to those produced from other SM processes. While these distinctive kinematic features are often employed to discriminate the top quark signal from SM backgrounds, they are not exclusive to it and there are a number of new physics models predicting heavy quarks with masses above the one of the top quark whose decays produce event signatures similar to those from top quark decays.

The simplest extension of the SM with three generations is a fourth chiral generation of massive fermions. The fourth generation is predicted in a number of theories [1, 2], and although not popular historically, is in a good agreement with electroweak precision data [3, 4].

To avoid $Z \to \nu\bar{\nu}$ constraint from LEP I a fourth generation neutrino ν_4 must be heavy: $m(\nu_4) \gtrsim m_Z/2$, where m_Z is the mass of Z boson, and to avoid LEP II bounds a fourth generation charged lepton ℓ_4 must have $m(\ell_4) \gtrsim 101$ GeV. At the same time sizeable radiative corrections mean masses of fourth generation fermions cannot be much higher than the current lower bounds and masses of new heavy quarks t' and b' should be in the range of a few hundred GeV [4]. Which would be accessible at the Tevatron collider. In addition, a small mass splitting between t' and b' is preferred, such that $m(b') + m(W) > m(t')$, and t' decays predominantly to Wq (a W boson and a down-type quark $q = d, s, b$) [4, 5].

In the four-generation model the present bounds on the Higgs are relaxed: the Higgs mass could be as large as 500 GeV [4, 5], which could resolve the conflict between the SM prediction for the Higgs mass and the LEP II direct lower limit [6]. Furthermore, CP violation is significantly enhanced to a magnitude that might account for the baryon asymmetry in the Universe [7]. Additional chiral fermion families can also be accommodated in supersymmetric two-Higgs-doublet extensions of the SM with equivalent effect on the precision fit to the Higgs mass [8].

Another possibility is heavy exotic quarks with vector couplings to the W boson. Contributions to radiative corrections from such quarks with mass M decouple as $1/M^2$ and easily evade all experimental constraints. For example, motivated by the 3σ discrepancy between the hadronic and leptonic asymmetry measurements from LEP II, which result in controversial predictions for the Higgs mass [6, 10], the "beautiful mirrors" model [9] solves the problem by introducing a new vectorlike fermion doublet, a mirror copy of the standard quark doublets with a heavier version of the SM top decaying to Wb.

A heavy top-like quark also appears in Little Higgs models [11], which evade the hierarchy problem by introducing a minimal set of gauge and fermion fields in the context of a large-extra-dimension framework. In particular, models in which T-parity is conserved suggest a massive top-like quark which can decay to Wq and would have a mass of approximately 500 GeV [12].

Thus, there are a number of well-motivated scenarios predicting a heavy top-like particle decaying into a W boson and a down-type quark $q = d, s, b$. In this work we search for pair production of such hypothetical new quarks using events characterized by a high-p_T lepton,

large $\not\!\!E_T$, and multiple hadronic jets. We refer to the hypothetical new quark as t'. We assume that the new quark is heavier than the top, and for the purpose of setting limits we assume that the new heavy quarks are produced strongly and decays promptly to Wq final states.

Previous iterations of this analysis were conducted at integrated luminosities of 347 pb^{-}1 [13], 760 pb^{-}1 [14], 2.3 fb^{-}1 [15] and 2.8 fb^{-}1 [16]. This analysis uses a larger dataset (4.6 fb^{-1}) and includes several improvements enhancing sensitivity to a potential new physics signal in the CDF data.

Namely, with respect to the previous iteration of the analysis:

- We increase the signal acceptance by incorporating into the analysis muons collected on $\not\!\!E_T$+ jets trigger.
- We split events into several categories based on the goodness of the fit used in the t' mass reconstructon.
- We incorporate all of the shape systematics using the vertical template morphing technique.

More details on each of these items is described further in the note.

2 Data Samples and Event Selection

The CDF detector is described in detail in [17]. This analysis is based on the data collected by CDF II between March 2002 and March 2009 using inclusive high-p_T lepton and missing transverse energy ($\not\!\!E_T$)+ jets triggers. The total integrated luminosity of the analyzed data is 4.6 fb^{-1}.

We select events by requiring one and only one isolated electron or muon with E_T or p_T respectively above 25 GeV (the trigger lepton). To enhance acceptance we also use "loose" muon categories collected on the $\not\!\!E_T$+ jets trigger, following other CDF analyses. These "loose" categories are further referred to as "non-trigger muons".

For every event we require that the $\not\!\!E_T$ be greater than 20 GeV and that at least four jets have E_T above 20 GeV and $|\eta| < 2.0$.

To reduce the QCD background for events with electrons we apply a QCD veto, similar to that used in single top analyses: $M_{T,W} > 20$ GeV and $\not\!\!E_{T,sig} > -0.05 \cdot M_{T,W} + 3.5$, where $M_{T,W}$ is the transverse W boson mass, and $\not\!\!E_{T,sig}$is the $\not\!\!E_T$ significance defined as

$$\not\!\!E_{T,sig} = \frac{\not\!\!E_T}{\sqrt{\sum_{jets} C_{JES}^2 cos^2(\Delta\phi_{\vec{\not\!\!E}_{T,jet}}) + cos^2(\Delta\phi_{\vec{\not\!\!E}_T{}^{uncorr},\vec{\not\!\!E}_T{}^{corr}})}} \quad (1)$$

where C_{JES} is a jet energy correction factor and $\Delta\phi_{\vec{\not\!\!E}_T{}^{uncorr},\vec{\not\!\!E}_T{}^{corr}}$ is between the uncorrected and corrected $\vec{\not\!\!E}_T$. The tracking resolution limits our ability to accurately reconstruct the muon momenta above a certain threshold. This leads to events with ultra high-p_T muons whose p_T is on the order of a TeV. To reduce the background from such mis-reconstructed muons, we reject events with muons having $p_T > 150$ GeV, for which the azimuthal angle between the $\not\!\!E_T$ corrected for the muon and the muon direction ($\Delta\phi(\not\!\!E_T, p_{T,lep})$) is less than 3.05 rad.

For the non-trigger muons we weight Monte Carlo events according to the $\not\!\!E_T$ turn-on curve applied to the vertex-corrected $\not\!\!E_T$. In addition, for both data and Monte Carlo non-trigger

Physics Process	Generator	NLO Cross Section (pb)
$t\bar{t}$	PYTHIA	7.4
W+LF	ALPGEN+PYTHIA	
single top, s-channel	MADEVENT+PYTHIA	0.29 ($\times$ BR)
single top, t-channel	MADEVENT+PYTHIA	0.64 ($\times$ BR)
WW	PYTHIA	12.4
WZ	PYTHIA	3.7
ZZ	PYTHIA	3.8
Z+LF	ALPGEN+PYTHIA	K=1.4
Z+HF	ALPGEN+PYTHIA	K=2.0

Table 1: Contributing SM physics processes and Monte Carlo samples used in the analysis.

muon events we require at least two jets that have $E_T > 25$ GeV, one of which has to be central ($|\eta| < 0.9$) and the ΔR between these two jets has to be greater than 1.0. These requirements assure 100% efficiency for events collected on the $\not\!\!E_T$+ jets trigger.

Using these event selection criteria we observe a total of 3724 events, 1677 events of which are electron + jets, 1240 trigger muon + jets events, and 807 non-trigger muon + jets events.

For the final event selection additional event cuts aimed to address mis-modeling of kinematic distributions for events with high jet E_T and lepton p_T are applied. These cuts are derived from the control regions with the lepton + 2 and 3 jets and described in Sec. 5.

3 Background Modeling

3.1 Monte Carlo Simulation

The dominant SM processes in our analysis are $t\bar{t}$ and W+ jets. Much smaller backgrounds include QCD, where a jet fakes a high-p_T lepton, and electroweak processes: diboson and single top production, as well as $Z/\gamma^\star$+ jets, where one of the leptons is not reconstructed. All of these processes, except for QCD, are modeled with the Monte Carlo simulation.

The $t\bar{t}$ is modeled with PYTHIA v6.216. We assume the top mass value of 172.5 GeV [**?**], and the NLO cross section of 7.4 pb [18]. W+ jets processes are modeled with ALPGEN v2.10′ [19] and use PYTHIA v6.325 for parton shower simulation. In this analysis we use only W+ light-flavor MC samples with W decaying into $\ell\nu, (\ell = e, \mu$ or $\tau)$. The samples are merged according to their respective ALPGEN cross sections. The total normalization of W+ jets is obtained from the fit, as described further in the note.

Diboson processes are modeled with PYTHIA v6.216 and normalized to the theoretical NLO cross sections [20]. Single top events are simulated using the tree-level matrix-element generator MADEVENT [21]. Drell-Yan events, $Z/\gamma^\star \to \ell^+\ell^-$, $(\ell = e, \mu, \tau)$, produced in association with jets are modeled with ALPGEN. The samples are merged according to respective ALPGEN cross sections, and since ALPGEN is a LO generator, the events are re-scaled by a factor of 1.4 for Z+ light-flavor samples and 2.0 for Z+ heavy-flavor [22]. MC samples and the respective cross sections for these physics processes are listed in Table 1.

$m(t')$ (GeV)	σ (pb)
180.0	5.75
200.0	3.19
220.0	1.82
240.0	1.06
260.0	0.63
280.0	0.38
300.0	0.23
320.0	0.14
340.0	0.083
350.0	0.064
360.0	0.050
380.0	0.030
400.0	0.018
450.0	0.0048
500.0	0.0013

Table 2: $t'\bar{t}'$ samples and respective $t'\bar{t}'$ NLO cross sections.

Pair production of the fourth-generation quarks is modeled using PYTHIA. We generated several mass points listed in Table 2. The NLO cross sections for $t'\bar{t}'$ production are calculated for us by Michelangelo L. Managano using the same technique as in $t\bar{t}$ cross section calculation [18]. These cross sections are also listed in Table 2.

We take into account the lepton trigger efficiencies and differences in reconstruction efficiencies between data and Monte Carlo using corresponding scale factors.

3.2 Modeling of QCD

The QCD background enters the sample when one of the jets fakes an electron or muon. The rate of jets faking electrons is about 100 times higher than faking a muon, therefore this background contributes mostly to electron + jets events. We model the QCD background using data events collected using triggers that require jets of 20, 50 and 100 GeV, the so called "jet-electron" sample. This is a sample of events that contain a jet with most of its energy in the electromagnetic calorimeter $[0.8\%, 0.95\%]$ and at least four tracks. The upper threshold on the EM fraction and the track requirement prevent the sample from being contaminated by real electrons. Jet electrons are treated as tight electrons. They are removed from the jet counting and the $\not\!\!E_T$ correction, and the events are otherwise subject to the same selection criteria as standard tight electrons.

While in the past only events from the the 20 GeV jet trigger were being used, this sample of events suffers from low statistics in the high E_T tails of the kinematic distributions, the adequate modeling of which is especially important for this analysis. Therefore in addition to events collected from the 20 GeV jet trigger, which is heavily pre-scaled during the data taking due high trigger rates, we make use of events from the 50 and the 100 GeV jet triggers. To account for the trigger turn-on, we only use events from the 100 GeV trigger that contain

leading jet E_T above 120 GeV. We use events from the 50 GeV jet trigger with leading jet E_T between 60 and 120 GeV, and events from the 20 GeV jet trigger are required to have leading jet E_T below 60 GeV. These three jet-electron samples are then stitched together into one QCD sample in the proportions such that the leading jet E_T spectrum is continuous.

To determine the contribution from QCD events, we relax the $\not\!\!E_T$ cut and fit the $\not\!\!E_T$ distribution to a combination of QCD $\not\!\!E_T$ template and the MC templates to data. In the fit the $t\bar{t}$ and electroweak processes are fixed to their predicted values according to their cross sections, and only the QCD and $W+$ jets templates are allowed to float independently. For electrons this fit is performed after applying the QCD veto cuts described in Sec. 2.

We use the same MC model for muons, and the QCD template obtained from jet electrons. Although the mechanism of a jet faking a muon is different, we expect a similar $\not\!\!E_T$ shape from QCD. Besides the $\not\!\!E_T$ distribution in data nicely agrees with a combination of QCD and $W+$ jets $+t\bar{t}$ templates.

Using this fit procedure we find that QCD contributes for about 14% of events in electron + jets, 2.5% for trigger and 4.3% for non-trigger muons respectively. We performed the same tests using other QCD models. We observe that the models are either not good for modeling kinematics of events, as has also been reported in other analyses, or suffer from low statistics in the region of our interest (≥ 4 jets). The uncertainty on the QCD fraction obtained from the fits is 20%. We double it and assign 40% uncertainty to account for limitations in the modeling of kinematics using jet electrons.

4 Mass Reconstruction

The variable

$$H_T = \sum_{jets} E_{T,j} + E_{T,\ell} + \not\!\!E_T, \tag{2}$$

serves as a good discriminator between Standard Model and new physics processes associated with production of high mass particles.

In addition, we make use of the fact that t' decay chain is assumed to be identical to the one of the top quark, and reconstruct its mass similarly to how it is done in the top quark mass measurement analyses. We adopt the template method for the top quark mass reconstruction, which is based on the χ^2-fit of kinematic properties of final top decay products.

We consider only the 4 highest E_T jets in the mass reconstruction. For each event there are thus a total of $4!/2 = 12$ combinations of assigning 4 jets to partons. In addition, there are two solutions for the unknown P_z neutrino momentum. A 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:

$$\chi^2 = \sum_{i=\ell,4jets} \frac{(p_T^{i,fit} - p_T^{i,meas})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(p_j^{UE,fit} - p_j^{UE,meas})^2}{\sigma_j^2}$$

$$+ \frac{(m_{jj} - m_W)^2}{\Gamma_W^2} + \frac{(m_{\ell\nu} - m_W)^2}{\Gamma_W^2} + \frac{(m_{bjj} - m_t)^2}{\Gamma_t^2} + \frac{(m_{b\ell\nu} - m_t)^2}{\Gamma_t^2}, \tag{3}$$

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 the masses of the 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

$$\vec{p}_T^{\,\nu} = -(\vec{p}_T^{\,\ell} + \sum \vec{p}_T^{\,jet} + \vec{p}_T^{\,UE}). \quad (4)$$

The longitudinal component p_z^ν is an 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 a free parameter initialized with $m_t = 175$ GeV, and its value in the best fit is declared to be the reconstructed mass M_{rec} of top (or t' respectively). In accordance with top mass analyses to assure more accurate mass reconstruction the fitted jets are corrected for various systematic biases 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 and provide better matching between parton and reconstructed jet energies.

Since in this analysis we are looking at $t' \to Wq$, where q can a be d, s or b quark, no b-tagging information is used in the mass reconstruction. The b-tag analysis focusing specifically on $t' \to Wb$ is underway.

Unlike in top mass measurements we do not reject events that have a poor χ^2 for reconstructed events, but instead split events into separate categories based on χ^2.

Templates of M_{rec} for main Standard Model backgrounds, W+jets and $t\bar{t}$, and t' with mass of 400 GeV are shown in Figure 1. The left upper plot shows the templates for events with ≥ 4 jets, when no χ^2-cut has been placed. The fat low M_{rec} tail in t' templates is due to jet-parton mis-assignments and FSR radiation. When jets and partons are mis-assigned, it is more likely to get the reconstructed mass for t' lower than the actual generated value. In addition, a q-jet originating from t' decay is very energetic, it can radiate a high-E_T gluon, and as a result can appear as two high-E_T jets in the detector.

The right upper plot in Figure 1 shows templates when the assignment between jets and partons is correct. In this case the low M_{rec} tail in t' is reduced, as well as high M_{rec} tail of $t\bar{t}$ events, and there is a good discrimination between these distributions. Correct matching is found in 24% of $t\bar{t}$ events and 36% of $t'\bar{t'}$ events.

In order to improve the discrimination power of our method and improve the sensitivity to a potential t' signal, we split the templates into four regions, based on the number of jets: $= 4$ or ≥ 5, and good or bad χ^2: $\chi^2 < 8$ and $\chi^2 > 8$. The sample of exactly 4 jets and a good χ^2 has the largest statistics due to the fact that majority of $t\bar{t}$ events (65% out of all ≥ 4 jets events) fall into this category. This is the region where the t' reconstruction performs the best, and in case of the t' signal in the data, one would hope to see the mass bump in the M_{rec} distribution.

Next, 19% of $t\bar{t}$ events fall into ≥ 5 jets, $\chi^2 < 8$ category. Here due to the additional jet the reconstruction performance is a bit worse. Smaller fractions of $t\bar{t}$ events fall into $\chi^2 > 8$ region, 9% and 6% of events for 4 and ≥ 5-jet bins respectively. The t' mass reconstruction is rather poor in these categories of events. However, because $t'\bar{t'}$ events are distributed more uniformly among all four categories of events (see Table 3), those are important to keep in order to increase acceptance to potential t' signal.

Thus we perform the search for the t' signal by employing a binned likelihood fit in both H_T and M_{rec} simultaneously for four different sets of templates:

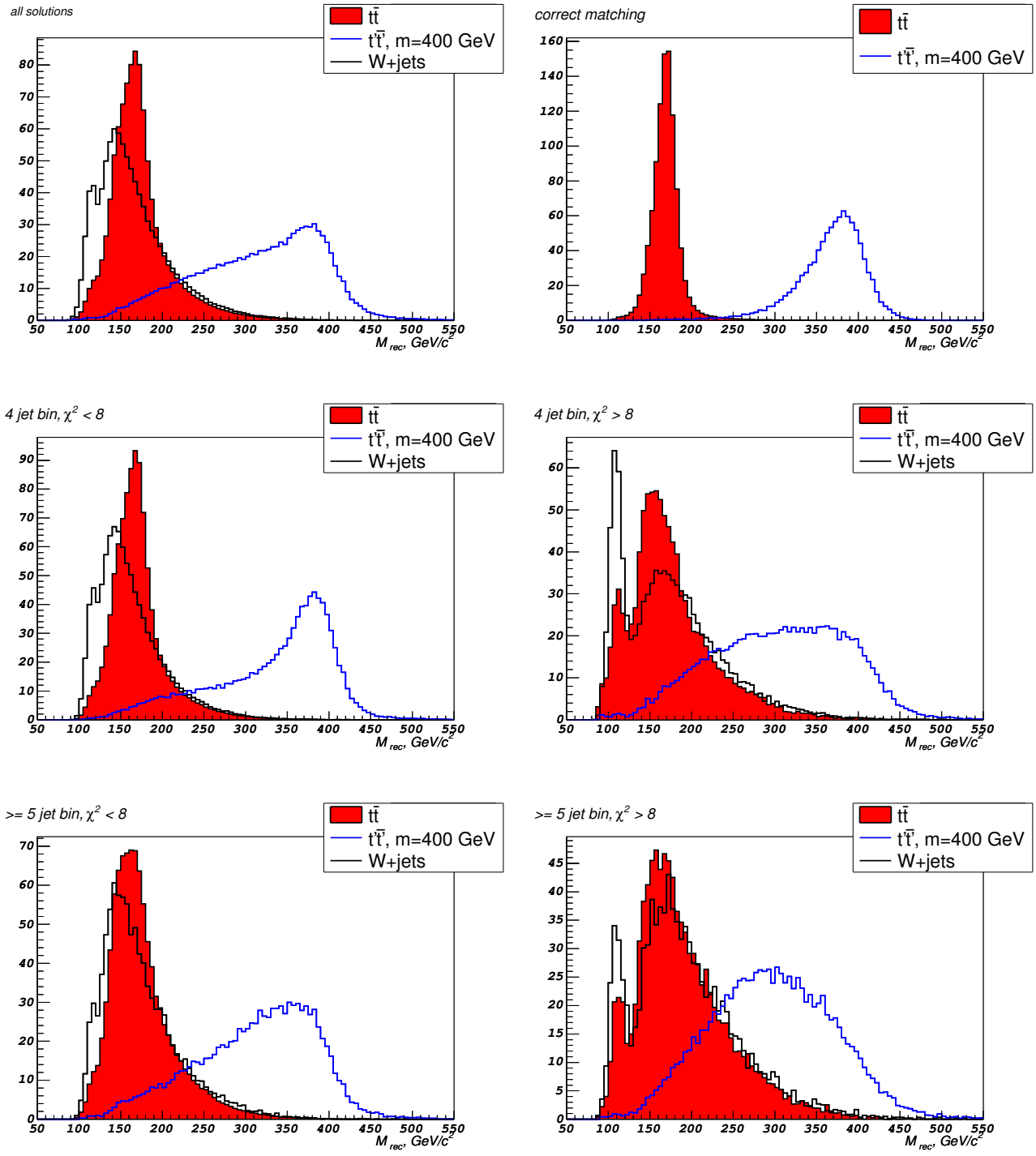

Figure 1: Reconstructed mass distribution for four different sets of templates.

	$= 4$ jets	≥ 5 jets
$\chi^2 < 8$	36%	26%
$\chi^2 > 8$	18%	21%

Table 3: Fractions of t' events with mass of 400 GeV in four categories of events.

- $= 4$ jet bin, $\chi^2 < 8$
- $= 4$ jet bin, $\chi^2 > 8$
- ≥ 5 jet bin, $\chi^2 < 8$
- ≥ 5 jet bin, $\chi^2 > 8$

More details on the likelihood fit and incorporation of systematic uncertainties are given in Sec. 6.

5 Modeling of Tails of Kinematic Distributions and "Clean-up" cuts

Since the potential t' signal due to a large t' mass is expected to populate the tails of kinematic distributions, it is important that a modeling of the tails of kinematic distributions using Monte Carlo events is adequate. We test the modeling of the tails using the control regions with exactly 2 or 3 jets in the events. The lepton + 2 and 3 jets events are pre-dominantly from $W+$ jets.

We focus on the modeling of lepton p_T and leading jet E_T spectra, since they are the main culprits who can contribute to an excess of events at high H_T and M_{rec} tails. We believe that this mis-modeling is due to deficiency of the standard energy corrections in the detector simulation and/or deficiency of our QCD model at high transverse energies. The improper energy corrections effects at high E_T would reveal themselves as a correlation between the missing E_T vector and the direction of the jet or the lepton. The likely source of an excess of events in small $\Delta\phi$ between the jet and $\not\!\!E_T$, and in large $\Delta\phi$ between the electron and $\not\!\!E_T$ is QCD events, which are not accounted for by our QCD model. We derive straight cuts to clean our sample from this type of events and thus provide more robust modeling of tails of kinematic distributions. For electron + jet events we apply additional cuts:

- For events with a leading jet E_T above 160 GeV, we require $\Delta\phi$ between the jet and the missing E_T direction to be greater than 0.6
- For events with an electron E_T above 120 GeV, we require $\Delta\phi$ between the electron and the missing E_T direction to be less than 2.6

After these additional "clean-up" cuts the predictions of events at high E_T tails are in much better agreements with observations.

In case of muon + jets no mis-modeling in the jet E_T spectra is observed, which confirms that mis-modeling in e + jets events is due to QCD contribution. However, events with high-p_T muons often result in fake $\not\!\!E_T$. Therefore, for muon + jets we apply the following set of cuts:

Events per 4.6 fb^{-1}

Source	Electron	Muon	Non-Trigger Muon	Total
$t\bar{t}$	640 ± 64	561 ± 57	368 ± 37	1569 ± 157
W+ jets	660 ± 660	561 ± 56	335 ± 335	1556 ± 1556
Diboson	46 ± 5	39 ± 4	23 ± 2	108 ± 11
Single top	9 ± 1	8 ± 1	4 ± 1	21 ± 1
Z+ jets	28 ± 3	38 ± 4	32 ± 3	98 ± 10
QCD	249 ± 100	31 ± 12	32 ± 11	312 ± 122
SM Total	1632	1238	794	3664
Data	1633	1225	790	3648

Table 4: Expected and observed number of events in the region ≥ 4 jets after the clean-up cuts.

- For events with a muon p_T above 120 GeV, we require $\Delta\phi$ between the muon and the missing E_T direction to be less than 2.6
- For events with a non-trigger muon p_T above 120 GeV, we require additionally $\Delta\phi$ between the muon and the missing E_T direction to be more than 0.4

This set of "clean-up" cuts improves modeling of the tails in the kinematic distributions and establishes an agreement between predicted and observed number of events. At the same these cuts barely affect the acceptance for potential t' signal, since the t' events tend to populate the opposite side of the $\Delta\phi$ distributions.

These additional "clean-up" cuts remove 76 data events from our signal region with ≥ 4 jets. The numbers of expected and observed events after the "clean-up" cuts are presented in Tab. 4.

6 Analysis Method

We perform a binned likelihood fit in H_T and M_{rec} to extract the t' signal and/or set an upper limit on its production rate for four different sets of templates:

- $= 4$ jet bin, $\chi^2 < 8$
- $= 4$ jet bin, $\chi^2 > 8$
- ≥ 5 jet bin, $\chi^2 < 8$
- ≥ 5 jet bin, $\chi^2 > 8$

simultaneously. The likelihood is defined as the product of the Poisson probabilities for observing n_i events in the bin i of (H_T, M_{rec}):

$$\mathcal{L}(\sigma_{t'}|n_i) = \prod_i P(n_i|\mu_i) \quad . \tag{5}$$

The expected number of events in each bin, μ_i, is given by the sum over all sources indexed by j, which is summed over all lepton categories:

$$\mu_i = \sum_j L_j \sigma_j \epsilon_{ij} \quad . \tag{6}$$

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_{rec}).

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

The production rates for W+jets in the 4-jet bin and in the ≥ 5 jet bins are two free unconstrained independent parameters in the fit. Other parameters, such as the $t\bar{t}$ production cross section, lepton ID data/MC scale factors and 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 for dealing with these parameters, i.e. the likelihood is maximized with respect to the nuisance parameters.

Taking this into account the likelihood takes the following expression:

$$\mathcal{L}(\sigma_{t'}|n_i) = \prod_{i,k} P(n_i|\mu_i) \times G(\nu_k|\tilde{\nu}_k, \sigma_{\nu_k}) \quad , \tag{7}$$

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.

7 Systematic Uncertainties

7.1 Jet Energy Scale

The sensitivity to t' depends on knowing accurately the distribution of (H_T, M_{rec}). 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 leaving some residual uncertainty.

This uncertainty results in possible shift in the H_T and M_{rec} 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 (+1 templates) and downward (-1 templates) by one standard deviation respectively.

Then we interpolate the expectation value μ_i at each bin i as follows:

$$\mu_i = \mu_{0,i} + \nu_{JES} \cdot (\mu_{+1,i} - \mu_{-1,i})/2 + \nu_{JES}^2 \cdot (-\mu_{0,i} + (\mu_{+1,i} + \mu_{-1,i})/2) \tag{8}$$

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:

$$\nu_{JES} = \frac{\Delta_{JES}}{\sigma_{JES}} \quad . \tag{9}$$

It enters the likelihood (7) as a gaussian constraint penalty term: $G(\nu_{JES}|0,1) = \frac{1}{\sqrt{2\pi}} e^{-\nu_{JES}^2/2}$. Outside of the interval $\nu_{JES} \in [-1, 1]$ the value μ_i is extrapolated as a linear function of ν_{JES}.

This treatment of the systematic uncertainty in the likelihood is called the vertical template morphing method.

7.2 W+jets Q^2 Scale

The effect of the choice of the appropriate Q^2 scale for W+jets production is evaluated by using the $W+$ jets Monte Carlo samples generated with different Q^2 settings. We make use of MC samples generated with one half and double the nominal Q^2.

The Q^2 systematic is then incorporated into the likelihood in a manner similar to the Jet Energy Scale systematics, with an exception that it is only applied to $W+$ jets template. The expectation of $W+$ jets contribution in the bin i is given by

$$\mu_i = \mu_{Q^2=1.0,i} + \nu_{Q^2} \cdot (\mu_{Q^2=2.0,i} - \mu_{Q^2=0.5,i})/2 + \nu_{Q^2}^2 \cdot (-\mu_{Q^2=1.0,i} + (\mu_{Q^2=2.0,i} + \mu_{Q^2=0.5,i})/2) \quad (10)$$

where the parameter ν_{Q^2} is gaussian constrained in the likelihood.

7.3 Initial and Final State Radiation

We make use of $t\bar{t}$ samples that simulate the effect of increasing and decreasing the initial and final state radiation in $t\bar{t}$ events. The shifted templates ("IFSR less " and "IFSR more") serve as +1 and -1 σ templates and are incorporated into the likelihood as is done for the Jet Energy Scale and the $W+$ jets Q^2 systematics. The morphing is only performed for $t\bar{t}$ template. In principle, the initial and final state radiation also effects the shapes of the t' H_T and M_{rec} distribution. However, the effect of this shift is tiny for t', it changes the mean of the distribution by 2%. On the contrary, the effect from $t\bar{t}$ is non-negligible because it is a large background.

7.4 Integrated Luminosity

The integrated luminosity uncertainty is taken to be 5.8%, as is standard at CDF, and represented by an additional gaussian-constrained parameter multiplying all contributions except for the QCD background and $W+$ jets, which are normalized independently.

7.5 Lepton ID

We apply the lepton ID scale factors and trigger efficiencies to MC-based backgrounds only, except for $W+$ jets, which floats independently. The uncertainty due to those is 1% and is applied in quadrature with the uncertainty due to the NLO theoretical cross sections.

7.6 PDF Uncertainty

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 in the acceptance. The first is a major part of the NLO theoretical cross section and described in Sec. 7.7. The latter is estimated to be 1% from the $t\bar{t}$ cross section analyses, and is summed in quadrature with the uncertainty due to theory.

7.7 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 [18].

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.

Cross sections for small electroweak backgrounds are also known with the precision of 10% and applied as independent nuisance parameters.

8 Bin Merging

We use 28 bins for H_T and 18 bins for M_{rec}, with the overflow bins defined for events with H_T above 800 GeV and M_{rec} above 500 GeV.

Thus, there are total $28\times18\times4 = 2016$ total bins needed to be used in the fit. Since with so many bins it is hard to populate all of the bins with sufficient MC statistics (this difficulty is enhanced due to the fact that H_T and M_{rec} are correlated), we developed an algorithm that will merge contiguous bins with low MC statistics together into super-bins. These super-bins are the ones used in the likelihood fit. This procedure by construction deteriorates the sensitivity to the new physics signal, however it eliminates abnormalities from sources such as bins with zero predictions, and thus provides a reliable 95% C. L. observed limit.

The criterion used to define the robust binning is the requirement that each super-bin in the template constructed from summing all of normalized SM sources has the relative uncertainty due to MC statistics below 0.4

The algorithm consists of looping over all bins in the summed template, and first of all determines the most problematic bins: bins with zero entries or with relative uncertainty being 1.0. For each of those bins we look at adjacent bins, and identify the bin with the smallest relative uncertainty. Once such a bin found, the bin content from the problematic bin migrates there. If all adjacent bins happen to have zero entries, the "search region" is extended, searching for bins which are 2 bin boundaries apart from the bin in question. The search continues until a bin, where the problematic bin will migrate to, is found.

During the next iteration the bins with relative uncertainties above 0.9 are identified, and migrated into the adjacent bins. The algorithm is repeated until all of the bins have a relative uncertainty below 0.4 This procedure determines the bin mapping for the 2D-fit with sufficient MC statistics.

The event counts (weighted) templates are merged according to this mapping dynamically, i.e. during the time the mapping is generated. Every time, when a particular bin [X,Y] migrates into the bin [X',Y'], the bin content for all MC and data templates migrates bin-to-bin respectively, so that the mapping is preserved for all sources and systematic templates.

9 Results

We test sensitivity of our method by drawing pseudoexperiments from Standard Model distributions, i.e. assuming no t' contribution. Range of expected 95% CL upper limits with one

$m(t')$ (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 [18].

standard deviation bandwidth is shown in Figure 2. The purple curve is the theory curve [18], the values of which are given in Table 5. The lower σ_{min} and upper σ_{max} limits are obtained using the CTEQ6M family of parton density functions with uncertainties, together with the study of the scale uncertainty [23].

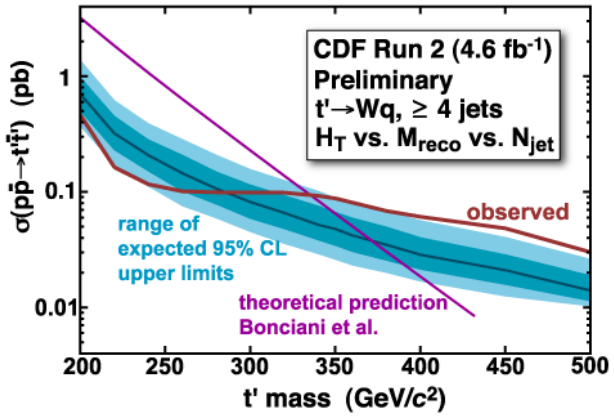

Figure 2: Limits.

From Figure 2 it follows that given no t' presence, this method is on average sensitive to setting an upper limit at 372 GeV t' mass.

We perform the analysis fit on the data and determine upper limits on the t' signal. The red curve in Figure 2 shows the final result, expressed as a 95% CL upper limit on the t' production rate as a function of t' mass. Table 6 shows the individual calculated limits along with expected limits from pseudo-experiments.

Distributions of H_T and M_{rec} with four categories of events combined are given in Figure 3.

$m(t')$ (GeV)	-2σ	-1σ	expected limit (pb)	+1σ	+2σ	observed limit (pb)
200	0.367	0.495	0.696	0.977	1.374	0.458
220	0.170	0.228	0.319	0.449	0.617	0.162
240	0.111	0.149	0.206	0.282	0.390	0.116
260	0.075	0.100	0.144	0.207	0.284	0.100
280	0.059	0.077	0.106	0.148	0.207	0.098
300	0.045	0.058	0.081	0.115	0.157	0.099
320	0.035	0.047	0.065	0.093	0.126	0.098
340	0.029	0.038	0.052	0.075	0.101	0.092
350	0.026	0.034	0.048	0.067	0.089	0.087
360	0.023	0.030	0.042	0.059	0.083	0.080
380	0.020	0.025	0.035	0.050	0.068	0.068
400	0.017	0.021	0.028	0.040	0.056	0.061
450	0.012	0.016	0.021	0.029	0.040	0.048
500	0.010	0.011	0.014	0.019	0.026	0.030

Table 6: Expected and observed limits on t' production cross section for given mass.

Here the distributions are "morphed" to correspond to the minimized likelihood scenario. The t' signal is for a mass of 400 GeV and corresponds to the cross section excluded at 95% C.L.

Based on these results we conclude that although there is some excess of data events in the tails of the distributions, it is not significant (less than 2σ), and we can exclude at 95% CL the t' quark with mass below 335 GeV, given the true top mass is 172.5 GeV.

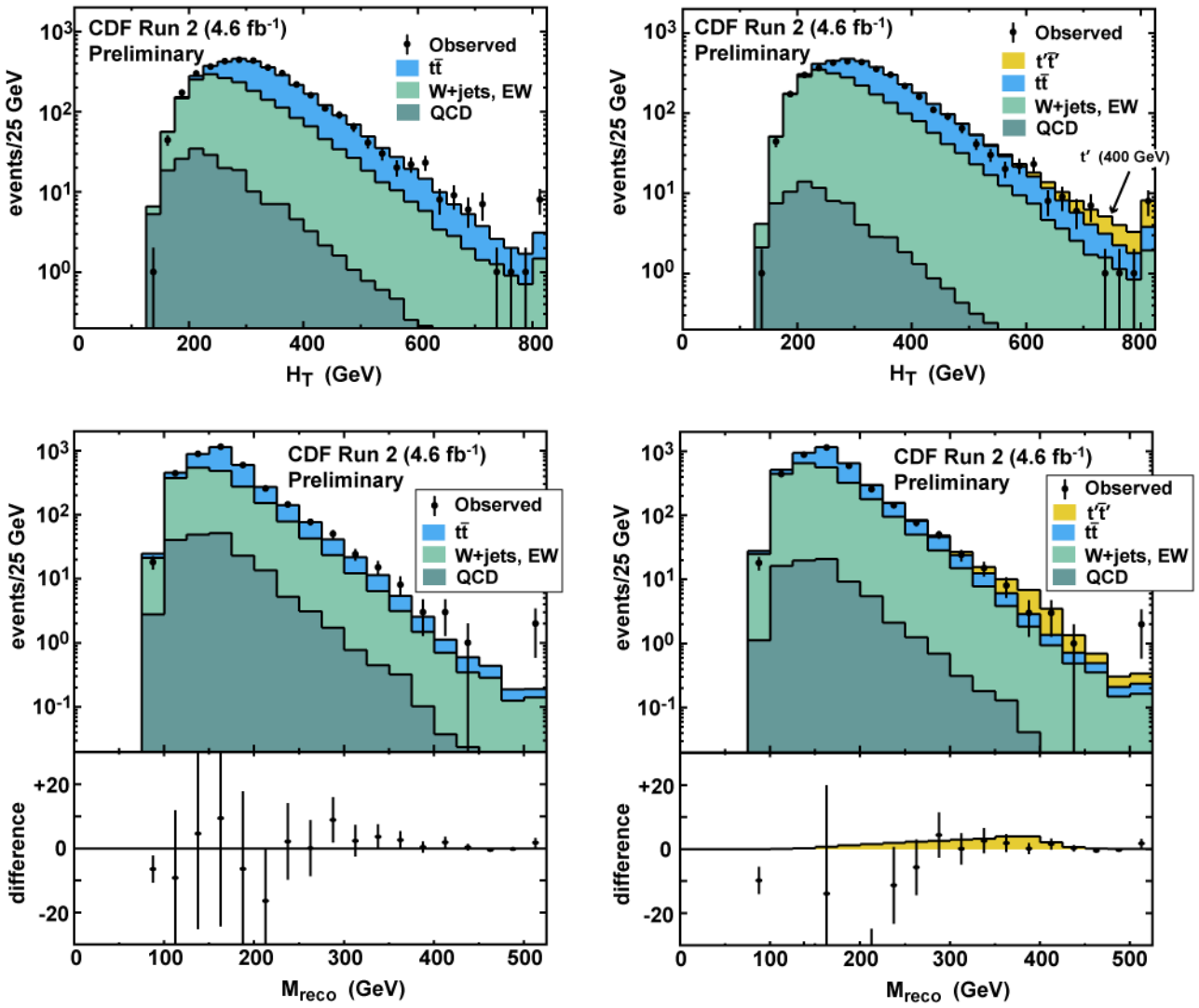

Figure 3: H_T and M_{rec} with four different categories of events collapsed. Standard Model distributions are "morphed" to correspond to the best fit to the data. The t' signal for the mass of 400 GeV corresponds to an amount excluded at 95% C.L.

References

[1] J. Silva-Marcos JHEP 0212 (2002) 036; E. Arik, O. Cakir, S. A. Cetin and S. Sultansoy, Phys. Rev. D **66**, 033003 (2002), [arXiv:hep-ph/0204217]; E. Arik, O. Cakir, S. A. Cetin and S. Sultansoy, Acta Phys.Polon. B **37**, 2839 (2006), [arXiv:hep-ph/0502050].

[2] N. Borstnik *et al.*, Bled workshops in physics, Vol.7, No. 2, DMFA-Zaloznistvo, Ljubljana,

Dec. 2006, [arXiv:hep-ph/0612250].

[3] V. A. Novikov, L. B. Okun, A. N. Rozanov and M. I. Vysotsky, Phys. Lett. B **529** (2002); V. A. Novikov, L. B. Okun, A. N. Rozanov and M. I. Vysotsky, JETP Lett. **76**, 127 (2002), [arXiv:hep-ph/0111028].

[4] G. D. Kribs, T. Plehn, M. Spannowsky and T. M. P. Tait, Phys. Rev. D **76**, 075016 (2007), [arXiv:0706.3718].

[5] P. H. Frampton, P. Q. Hung and M. Sher, Phys. Rept. **330**, 263 (2000), [arXiv:hep-ph/9903387].

[6] C. Amsler *et al.* [Particle Data Group], Phys. Lett. B **667**, 1 (2008).

[7] W.-S. Hou, F. F. Lee, C.-Y. Ma, Phys. Rev. D **79**, 073002 (2009), [arXiv:0812.0064]; W.-S. Hou, [arXiv:0803.1234v3].

[8] H. J. He, N. Polonsky and S. Su, Phys. Rev. D **64**, 053004 (2001), [arXiv:hep-ph/0102144].

[9] D. E. Morrissey and C. E. M. Wagner, Phys. Rev. D **69**, 053001 (2004); D. Choudhury, T. Tait and C. E. M. Wagner, Phys. Rev. D **65**, 053002 (2002), [arXiv:hep-ph/0109097].

[10] M. Chanowitz, Phys. Rev. Lett. **87**, 231802 (2001), [arXiv:hep-ph/0104024].

[11] T. Han, H. Logan, B. McElrath, L.-T. Wang, Phys. Lett. B 563, 191 (2003); B. A. Dobrescu and C. T. Hill, Phys. Rev. Lett. 81, 2634 (1998).

[12] H. C. Cheng and I Low, J. High Energy Phys. **0408**, 61 (2004); D. E. Kaplan and M. Schmaltz, J. High Energy Phys. **566**, 375 (2006).

[13] J. Conway *et al.*, CDF Note 7113.

[14] J. Conway *et al.*, CDF Note 7912.

[15] J. Conway *et al.*, CDF Note 8495.

[16] J. Conway *et al.*, CDF Note 9234.

[17] F. Abe, *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, *et al.*, Nucl. Instrum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, *et al.*, Phys. Rev. D **52**, 4784 (1995); P. Azzi, *et al*, Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E.

[18] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridol, JHEP 0404 (2004) 068 [arXiv:hep-ph/0303085].

[19] M. L. Mangano *et al.*, JHEP 0307:001 (2003), [arXiv:hep-ph/0206293].

[20] J. Campbell and R. Ellis, PRD 60, 113006 (1999).

[21] J. Alwall et al., J. High Energy Phys. 0709, 028 (2007).

[22] T. Aaltonen *et al.*, Phys. Rev. D **79**, 052008 (2009), [arXiv:0812.4458].

[23] Private communication with M. L. Mangano.