



Search for Heavy Top $t' \rightarrow Wq$ In Lepton Plus Jets Events

The CDF Collaboration
URL <http://www-cdf.fnal.gov>
(Dated: June 27, 2007)

We search for fourth-generation heavy top (t') quark pair production decaying to Wq final states using 760 pb^{-1} integrated luminosity from CDF Run 2, in the lepton+jets final state. 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 256 GeV at 95%CL.

Preliminary Results for Summer 2007 Conferences

I. INTRODUCTION

The cross section for $t\bar{t}$ production has been measured in Run 2 in the pretag sample of lepton+jets. The result in the three or more jets case is $\sigma_{t\bar{t}} = 6.6 \pm 1.1 \pm 1.5$ pb [1], in good agreement with the standard model prediction of 6.7 pb in the most recent NNLO calculations [2–4].

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 . As discussed below there are several possibilities arising from extensions to the standard model in which this may be the case, and are not excluded by precision electroweak data or other direct searches.

We refer here to the hypothetical new quark as t' , but in fact it need not be a standard fourth-generation up-type heavy quark. For the purposes of this analysis we simply consider a new quark which

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

In particular we need not demand the charge of the quark be $+2/3$, nor need we even demand that the quark be a fermion. (In the case of a new scalar quark, however, the production rate will be reduced due to the β^3 factor in its production cross section.)

A fourth generation of matter fermions with light neutrino ν_4 with mass $m(\nu_4) < m_Z/2$ is excluded by precision data from LEP 1. However, as pointed out by He *et al.* [5], and Okun *et al.* [6] a heavier fourth generation of fermions with $m_Z/2 < m_f < \mathcal{O}(< H >)$ is consistent with existing precision electroweak data. Indeed, the authors conclude that the present bounds on the Higgs in such scenario are relaxed; the Higgs mass could be as large as 500 GeV.

He *et al.* point out that additional fermion families can be accommodated in two-Higgs-doublet scenarios and $N = 2$ SUSY models, and possibly remove the requirement of the weak-mixing assumption above. In that case, the decay $t' \rightarrow Wq$ may predominate, assuming $m(t') > m(b') + m(W)$ for example.

Other theoretical possibilities lead directly to the scenario in which we are interested. In one version of the “beautiful mirrors” model of Wagner *et al.* [7] there exists an up-type quark with the same quantum numbers as the top, which decays as $\chi \rightarrow Wb$. In this scenario, the slightly anomalous results from LEP on the b forward-backward asymmetry are accommodated naturally, and the electroweak fits are improved (with a relaxed upper limit on the Higgs boson mass).

Recent theoretical developments lead one to hypothesize the existence of a heavy t' . Little Higgs models [8] evade the hierarchy problem by introducing a minimal set of gauge and fermion fields in the context of a large-extra-dimension framework. The minimal versions of these models, however, result in new quarks which have mass of order 1 TeV, too heavy for Tevtron studies. Non-minimal Little Higgs scenarios, however, are of course possible.

The basic conclusion is that there exist enough theoretical scenarios and ideas involving new heavy quarks that, as experimentalists, we are compelled to perform the search for them in as many channels as possible. In order to allow the widest possible theoretical interpretation we need to express our results in as model-independent a fashion as possible. Here we set limits on the t' pair production cross section times branching ratio $t' \rightarrow Wq$, leading to the high- H_T lepton+jets+ \cancel{E}_T signature with an acceptance determined from a generic fourth generation quark decaying to Wb . Inevitably, however, different proposed models will have different kinematic distributions and hence have acceptance different from a generic fourth-generation quark.

II. DATA SAMPLE AND EVENT SELECTION

The CDF detector is described in detail in [9]. We use a data sample corresponding to 760 pb^{-1} of integrated luminosity and collected with an inclusive lepton trigger that requires an electron with $E_T > 18$ GeV or muon with $p_T > 18$ GeV. Electron events with the primary electron identified as a conversion, and muon events with a cosmic ray are removed, as well as events consistent with a leptonic Z decay.

Then the following event selection criteria are applied

- one and only one isolated lepton (e or μ) with $p_T > 20$ GeV,
- at least 20 GeV missing transverse energy \cancel{E}_T , and
- at least four jets with $E_T > 15$ GeV and $|\eta| < 2$.

- if the \cancel{E}_T is below 30 GeV, the angle between the \cancel{E}_T and the highest E_T jet in the transverse plane, $\Delta\phi$ should be greater than 0.5 and less than 2.5 radians.

To insure that leptons and jets are reconstructed from the same interaction the event z vertex is required to be within 5 cm of the lepton z_0 .

The dominant contributing backgrounds after these cuts are W +jets, which we model with ALPGEN+HERWIG, and $t\bar{t}$ modeled with PYTHIA. We assume mass of the top quark to be 175 GeV. Other backgrounds include Z +jets, WW +jets, WZ +jets and single top. The QCD background is modeled using the data sample of non-isolated electrons and muons. $t'\bar{t}'$ pair production is simulated in PYTHIA.

III. ANALYSIS METHOD

Since the t' decay chain is identical to the one of the top quark, we reconstruct its mass with a method similar to the one used in the top quark mass measurement analyses. We use the template method for top quark mass reconstruction [10] based on the best χ^2 -fit of kinematic properties of final top decay products. The χ^2 is given by the following expression:

$$\chi^2 = \sum_{i=\ell, 4jets} \frac{(p_T^{i,fit} - p_T^{i,meas})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(p_j^{UE,fit} - p_j^{UE,meas})^2}{\sigma_j^2} + \frac{(m_{jj} - m_W)^2}{\Gamma_W^2} + \frac{(m_{\ell\nu} - m_W)^2}{\Gamma_W^2} + \frac{(m_{bjj} - m_t)^2}{\Gamma_t^2} + \frac{(m_{b\ell\nu} - m_t)^2}{\Gamma_t^2}, \quad (1)$$

where the invariant masses of the 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 same. Jet, lepton and underlying event (UE) energies are allowed to float within their resolution uncertainties. The transverse component of the neutrino momentum is determined as the negative sum of the lepton, jet and unclustered transverse energies:

$$\vec{p}_T^{\nu} = -(\vec{p}_T^{\ell} + \sum \vec{p}_T^{jet} + \vec{p}_T^{UE}). \quad (2)$$

For each event there are total $4!/2 = 12$ combinations of assigning four jets to partons. In addition, there are two solutions for unknown p_z neutrino momentum. After minimization of the χ^2 expression, the combination with the lowest χ^2 is selected and the value of m_t is declared to be the reconstructed mass M_{reco} of top (or t' respectively).

We use the observed distributions of the M_{reco} and total transverse energy in the event

$$H_T = \sum_{jets} E_T + E_{T,\ell} + \cancel{E}_T, \quad (3)$$

to distinguish the t' signal from the backgrounds by fitting it to a combination of t' signal, $t\bar{t}$, W +jets, and other background shapes.

We use a binned in H_T and M_{reco} likelihood fit 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}) :

$$\mathcal{L}(\sigma_{\nu} | n_i) = \prod_i P(n_i | \mu_i) \quad . \quad (4)$$

The expected number of events in each bin, μ_i , is given by the sum over all sources 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' .

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.

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.

IV. SYSTEMATIC ERRORS

Jet 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 as described in [11], 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.

Then 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 relative shift in the jet energy scale:

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

This effect enters the likelihood (6) as a gaussian constraint penalty term: $G(\nu_{JES}|0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\nu_{JES}^2/2}$.

W +jets Q^2 Scale

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 pseudoexperiments to measure this shift by drawing pseudoexperiments 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 (5) 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' .

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.

ISR and FSR

$m(t')$	Q^2 scale			
	$4m_W^2$	$m_W^2/4$	$\langle p_T^2 \rangle$	$\sigma_{\nu_{Q^2}}$
175	1.15	1.35	0.2	1.25
200	0.7	0.9	0.1	0.8
225	0.37	0.52	0.03	0.44
250	0.21	0.33	0.02	0.28
300	0.17	0.25	0.03	0.21
350	0.13	0.18	0.02	0.15

TABLE I: 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.

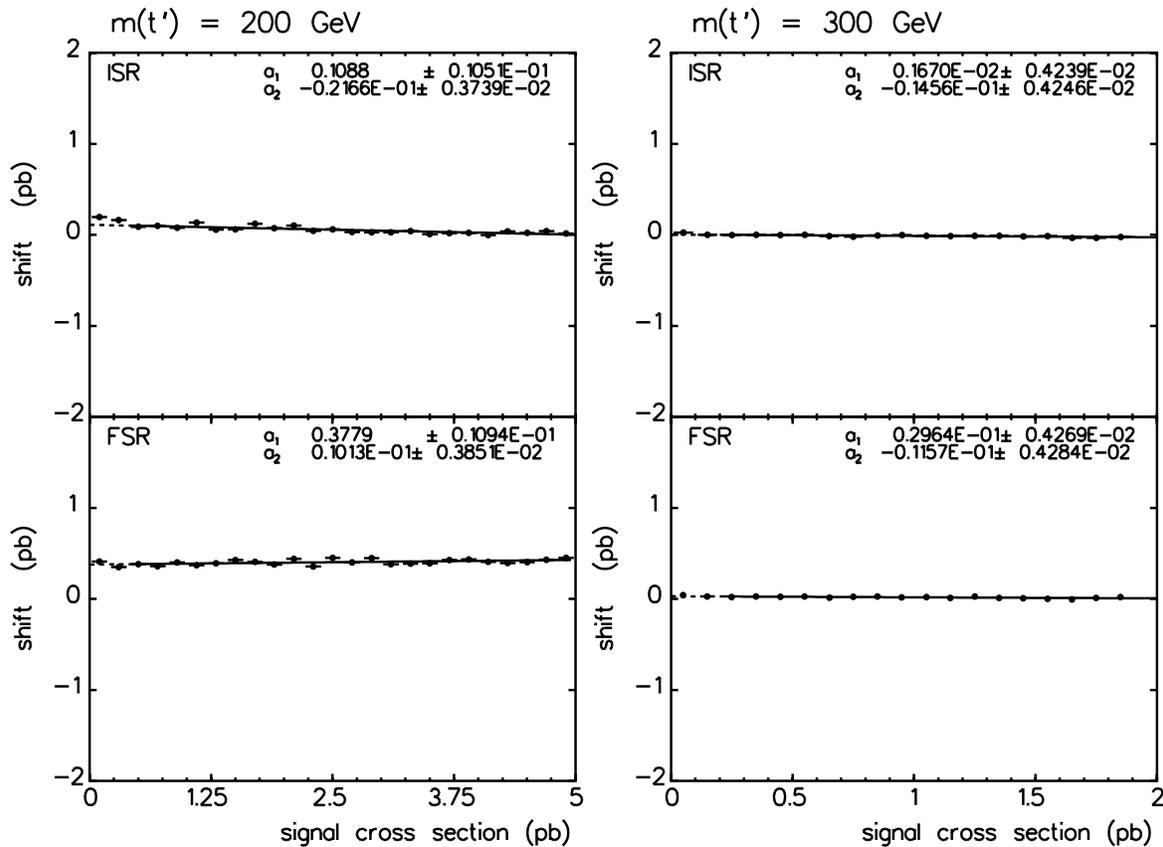


FIG. 1: Shift in measured t' cross section as a function of true t' cross section for “more ISR” and “more FSR” effects.

We investigated the effect of varying ISR and FSR using samples of “more ISR” and “more FSR” for t' at 175, 200, and 300 GeV. These mass points bracket the region of t' mass where we expect to set a limit.

We made templates for each of these, and then generated pseudoexperiments with the shifted 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 1.

In the “more FSR” case, we see a shift of 0.4 pb for 200 GeV and at 300 GeV it is consistent with zero shift. Interpolating roughly linearly, We add this to the Q^2 nuisance parameter width in the likelihood.

QCD Background

The QCD background shape is obtained from non-isolated leptons data and scaled based on the Met vs Iso method. After applying the cone $\Delta R(\text{jet}, \text{non-iso lepton}) > 0.4$ requirement for jet counting and correcting \cancel{E}_T for non-isolated

leptons, we obtain the QCD contribution relative to the observed number of events to be $5.95 \pm 0.96\%$ in electron+jets sample and $1.60 \pm 0.39\%$ in muon+jets sample.

The relative normalization uncertainty is represented by a gaussian-constrained parameter in the likelihood. The QCD background has a negligible effect on the t' limit.

Integrated Luminosity

The integrated luminosity uncertainty is taken to be 5.8%, and represented by an additional gaussian-constrained parameter multiplying all contributions except for the QCD background, which is normalized independently.

Lepton ID

We have two components for lepton ID. First is the efficiencies for the individual electrons and muons. 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 of 2%, 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.

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 the acceptance.

By comparing various PDF sets and their uncertainties, we get 0.34% due to CTEQ6L1/CTEQ6L differences, and 0.9% due to the α_s variation. We add these in quadrature and get 1.0% overall systematic uncertainty.

Theory Uncertainty

The theory uncertainty in the t' cross section is about 10% (see Table II), which is mainly due to uncertainty in PDFs ($\sim 7\%$). The other effect comes from uncertainty in the choice of the Q^2 scale [4].

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 the theoretical value.

V. RESULTS AND CONCLUSION

We tested the sensitivity of our method by drawing pseudoexperiments from Standard Model distributions, i.e. assuming no t' contribution. The range of expected 95% CL upper limits with one standard deviation bandwidth is shown in Figure 2.

The purple curve is the theory curve [3, 4], the values of which are given in Table II. 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 [12].

From Figure 2 it follows that given no t' presence, this method is on average sensitive to setting an upper limit at 260 GeV t' mass.

After performing the analysis fit on the data we find no evidence of a t' signal. Hence we proceed to 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 III shows the individual calculated limits along with expected limits from pseudoexperiments for reference.

Figure 3 shows the projections of the observed H_T and M_{reco} distributions for the example mass point $m(t') = 250$ GeV and the best fit where we set the 95% CL upper limit. 2D-distribution of (H_T, M_{reco}) is shown in Figure 4.

Based on these results we exclude at 95% CL the t' quark with mass below 256 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.

The slightly poorer than expected limit can be explained by an excess of events at the high-energy, high-mass tail. The event displays for four events with reconstructed masses above 350 GeV and $H_T > 500$ GeV are given in Figures 5- 8.

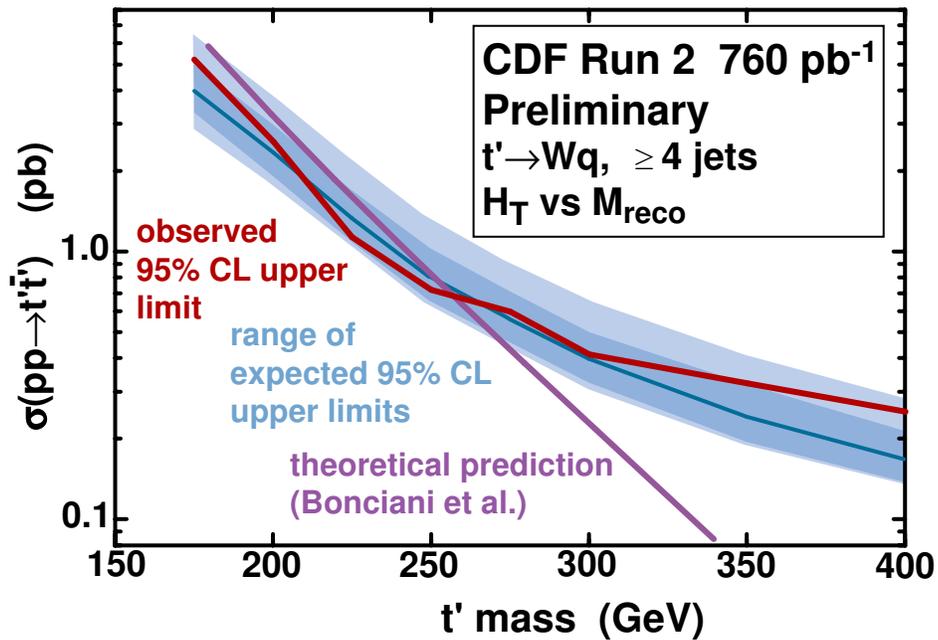


FIG. 2: Upper limit, at 95% CL, on the production rate for t' as a function of t' mass (red). The purple curve is a theoretical cross section. The blue band is the range of expected 95% CL upper limits within one standard deviation.

$m(t')$ (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 II: Theory values of t' cross section for given mass [3, 4].

$m(t')$ (GeV)	expected limit					observed limit
	2.3%	15.9%	median	84.1%	97.7%	
175	2.863	3.268	4.017	4.961	6.437	5.205
200	1.753	1.902	2.362	2.968	3.769	2.569
225	1.052	1.109	1.338	1.700	2.153	1.132
250	0.620	0.657	0.799	1.026	1.313	0.720
275	0.432	0.457	0.555	0.717	0.906	0.595
300	0.304	0.321	0.394	0.503	0.657	0.412
350	0.189	0.198	0.241	0.315	0.408	0.322
400	0.134	0.138	0.167	0.213	0.281	0.252

TABLE III: Expected and obtained limits on t' production cross section for given mass.

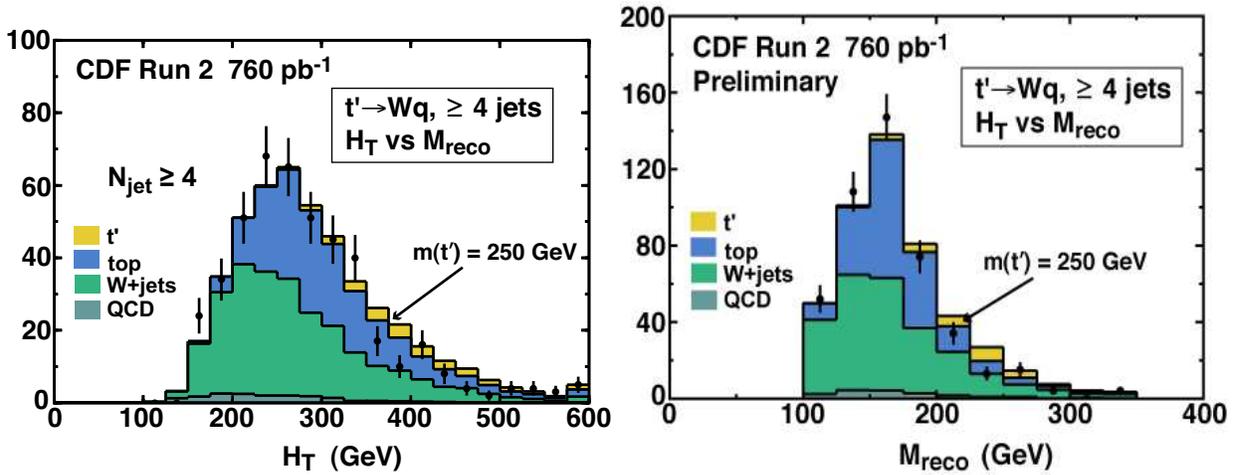


FIG. 3: Distributions of H_T (top) and M_{reco} (bottom) showing result of fit for $m(t') = 250$ GeV. The normalizations of the various sources and distortions of kinematic distributions due to systematic effects are those corresponding to the maximum likelihood when the cross section for t' is set to its 95% CL upper limit.

-
- [1] D. Acosta *et al.*, Phys. Rev. D **72**, 052003 (2005) [arXiv:hep-ex/0504053].
[2] N. Kidonakis and R. Vogt, Phys. Rev. D **68** 114014 (2003) [arXiv:hep-ph/0110145].
[3] R. Bonciani, S. Catani, M. L. Mangano and P. Nason, Nucl. Phys. B **529** (1998) 424 [arXiv:hep-ph/9801375].
[4] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, JHEP **0404** (2004) 068 [arXiv:hep-ph/0303085].
[5] H.-J. He, N. Polonsky and S. Su, [arXiv:hep-ph/0102144]
[6] L. Okun *et al.*, [arXiv:hep-ph/0111028]
[7] C. Wagner *et al.*, [arXiv:hep-ph/0109097]
[8] T. Han, *et al.*, Phys. Lett. B **563**:191 (2003).
[9] 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
[10] A. Abulencia *et al.*, Phys. Rev. D **73**, 032003 (2006)[arXiv:hep-ex/0510048].
[11] A. Bhatti *et al.*, Submitted to Nucl. Instr. Meth. A [arXiv:hep-ex/0510047].
[12] Personal communication with Michelangelo Mangano.

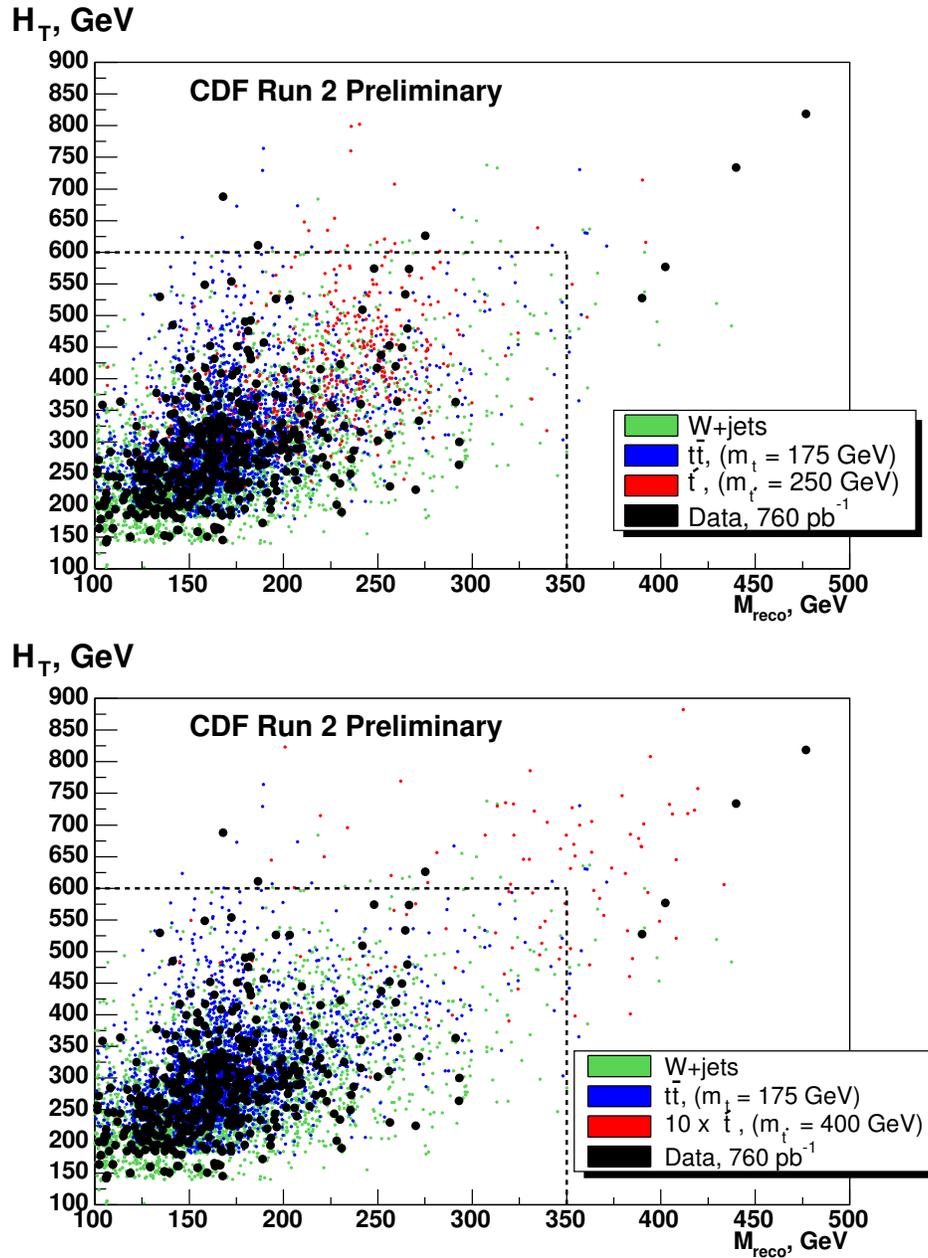


FIG. 4: The 2D-distribution of H_T and M_{reco} for $m(t') = 250$ GeV (top) and $m(t') = 400$ GeV (bottom). Scattered distributions of simulated physics processes have ~ 10 times larger number of events predicted at current luminosities. (100 times larger for $m(t') = 400$ GeV.) Ticks at the axes correspond to binning used in the fit (25 GeV wide). The dashed lines are the edges of the fitting box. Part of distributions outside of the fitting box is included into overflow bins.

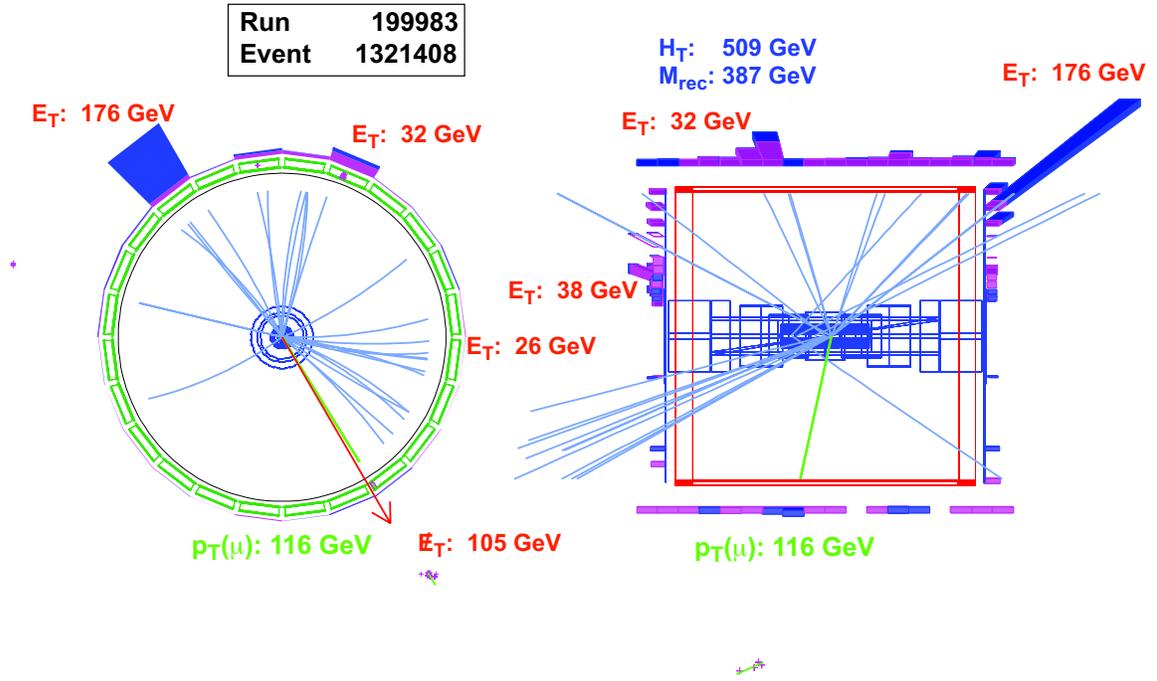


FIG. 5: Event display of one of the events at the high-energy, high-mass tail.

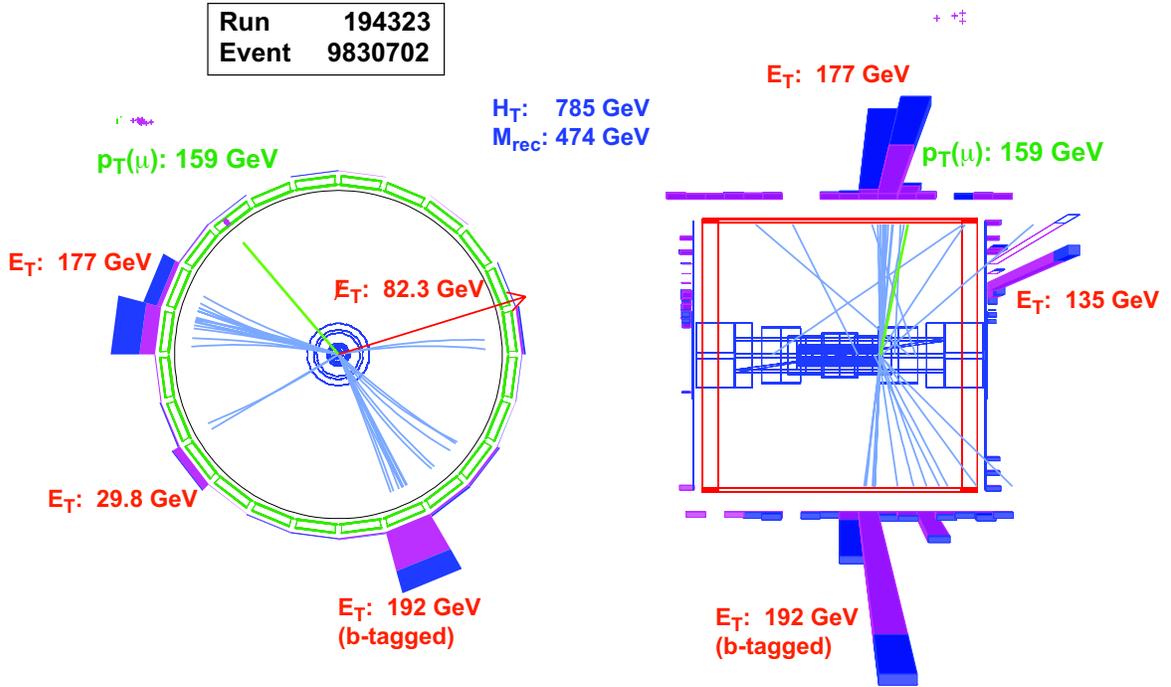


FIG. 6: Event display of one of the events at the high-energy, high-mass tail.

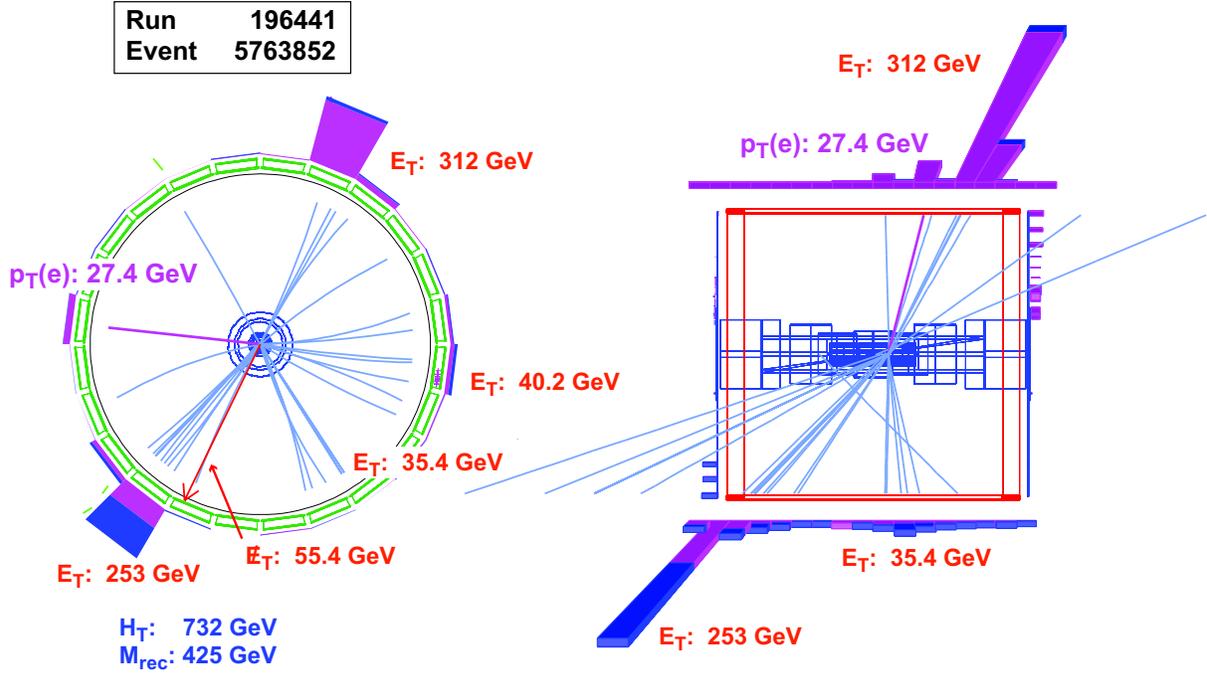


FIG. 7: Event display of one of the events at the high-energy, high-mass tail.

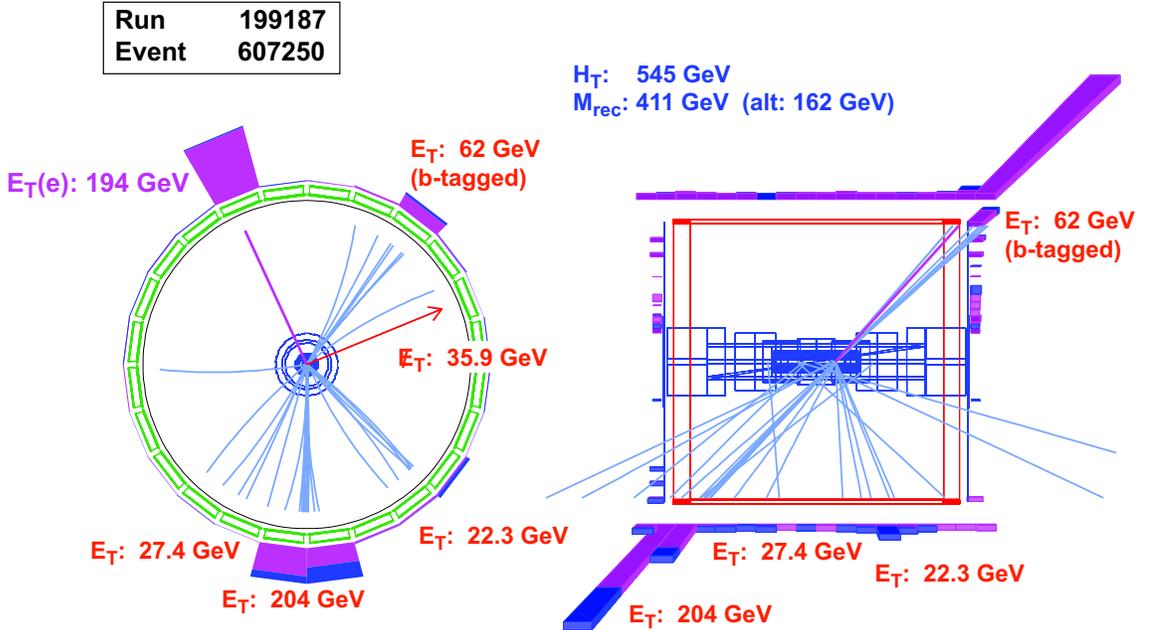


FIG. 8: Event display of one of the events at the high-energy, high-mass tail.