Abstract

This is an update on the search for the heavy top (t') quark pair production decaying to Wq final states using 2.8 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 311 GeV at 95%CL.

Search for Heavy Top $t' \to Wq$ in Lepton Plus Jets Events in $\int \mathcal{L} dt = 2.8$ fb^{-1}

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

July 11, 2008

1 Changes

Since the previous version of this analsysis (blessed at datehere) a few changes have been made.

- We now use 2.8 fb^{-1} of data (Data up to Period 17).
- A problem with the normalization used for W+Jets in the systematics was fixed.
- The Q^2 Systematic now covers the full range of partons.
- Four events which were mistakenly removed from the bulk of the distribution have been added back in.

Other aspects of the analysis remain unchanged from the previous version of the analysis, details of which may be found in CDF note 9209 [1].

2 W+jets Normalization in Systematics

In the previous version of the analysis when calculating the systematics the number of W+Jets events was normalized to the expected cross-section for W+all parton numbers (61.5 pb), but only the W+4p sample was being used (cross-section 2.06 pb). In the current version it is normalized using the normalization from the analysis (Number of Data - Number of QCD - Number of Top).

3 Q^2 Systematic

In the previous version of the analysis only the W+4p sample was being used for the systematics. Now the full W+np sample is being used.

4 Mistaken Removal

In the previous version of the analysis four events were mistakenly removed from the data sample. All four of these events were in the bulk of the distribution and they're addition does not affect the result.

5 Data Validation

The H_T and reconstructed Mass have been compared between the data previously used and the new data (see Figure 1). Additional comparison plots are available [2].

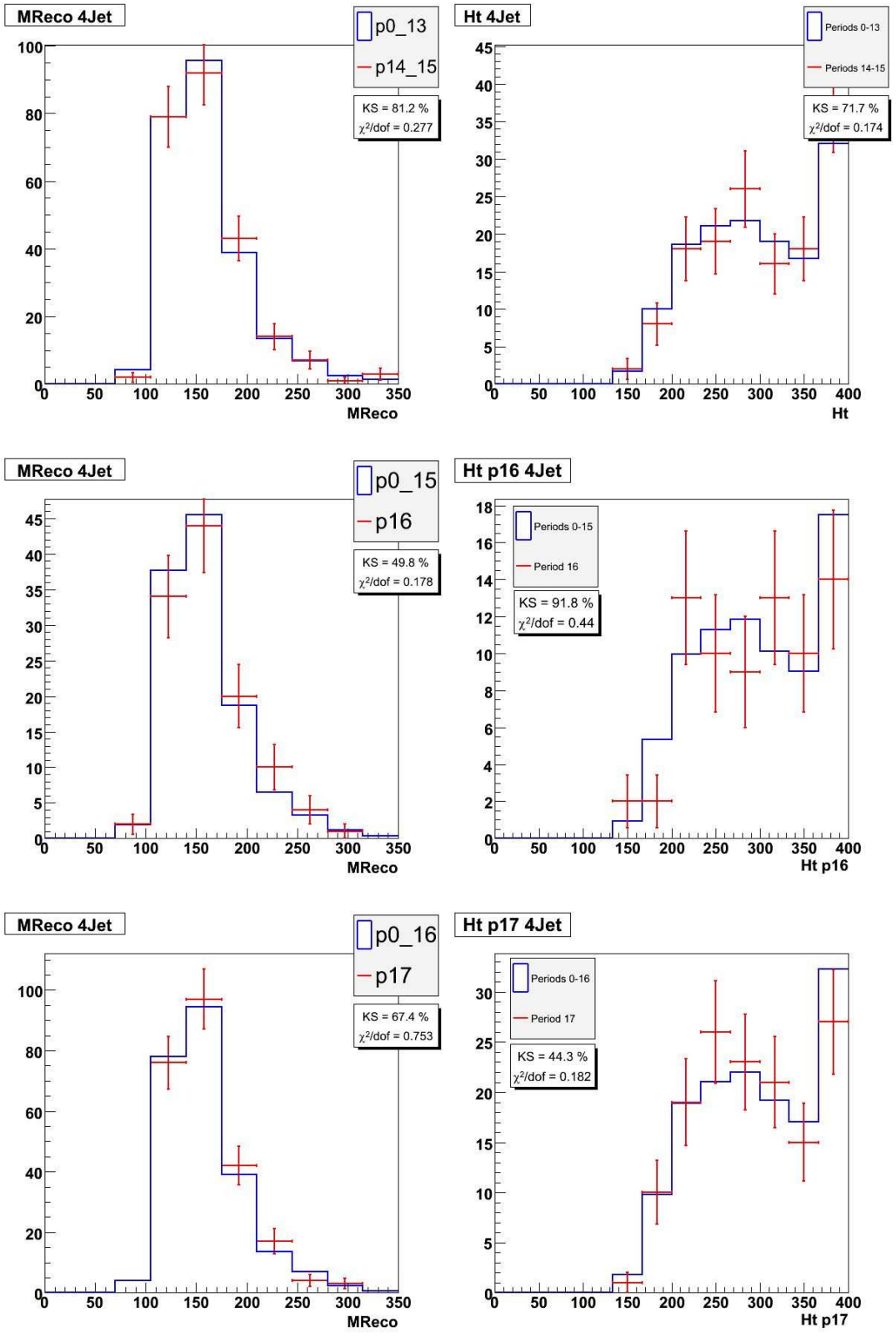

Figure 1: Comparison plots for Periods 0-13 (0-15, 0-16) in blue versus Periods 14-15 (16, 17) in red.

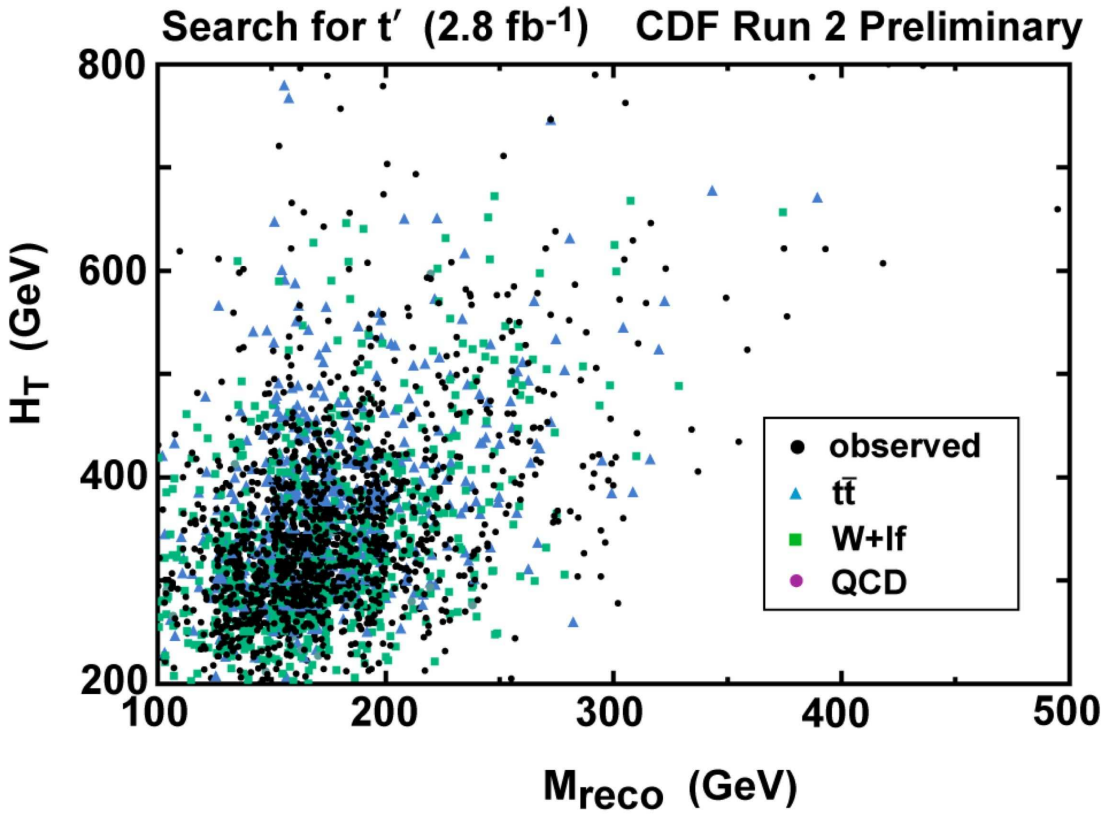

Figure 2: 2D plot of H_T vs M_{rec} distribution showing the data (black points) and the fitted number of background events: QCD (dark cyan triangles), W+jets (blue open circles) and $t\bar{t}$ (red triangles)

6 New Results

The red curve in Figure 3 shows the final result, expressed as a 95% CL upper limit on the t' production rate as a function of t' mass. Table 1 shows the individual calculated limits along with expected limits from pseudo-experiments for reference.

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

Based on these results we exclude at 95% CL the t' quark with mass below 311 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.

As in the previous version of the analysis we set out to determine if the data show any evidence of an excess far out in the tails of H_T and M_{reco} by counting the number of events in groups of $n \times n$ of our standard 25 GeV bins in these quanitites, 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 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 6 shows the result with the real data. The most significant $n \times n$ bin is that where $n = 10$;

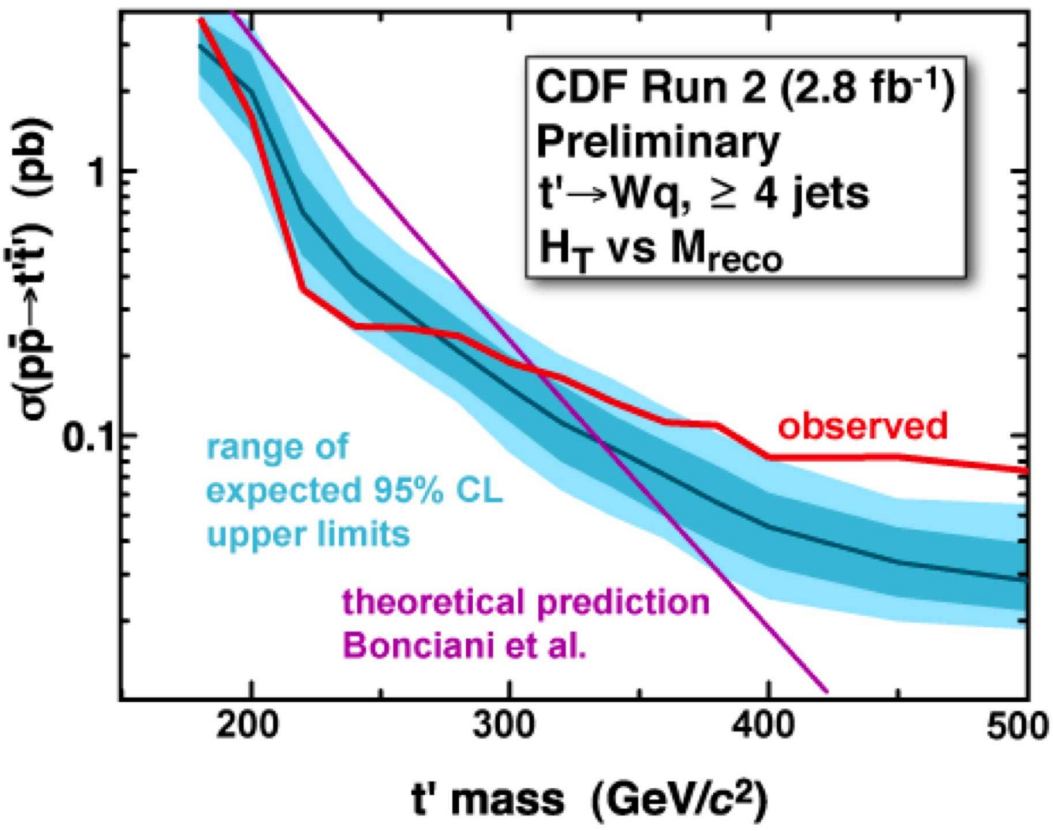

Figure 3: Upper limit, at 95% CL, on the production rate for t' as a function of t' mass (red). The purple curve is a theoretical cross section. The blue band is the range of expected 95% CL upper limits within one standard deviation.

	expected limit (pb)					observed limit (pb)
$m(t')$ **(GeV)**	-2σ	-1σ	**median**	$+1\sigma$	$+2\sigma$	
180	1.900	2.362	2.954	3.772	4.540	3.759
200	1.066	1.434	1.959	2.828	3.682	1.595
220	0.360	0.486	0.693	1.002	1.579	0.355
240	0.248	0.306	0.406	0.558	0.743	0.258
260	0.183	0.216	0.288	0.390	0.498	0.254
280	0.135	0.161	0.208	0.280	0.370	0.237
300	0.089	0.116	0.150	0.202	0.268	0.188
320	0.063	0.081	0.112	0.156	0.202	0.165
340	0.050	0.064	0.088	0.120	0.164	0.133
360	0.041	0.050	0.070	0.096	0.128	0.112
380	0.031	0.040	0.056	0.077	0.103	0.109
400	0.024	0.032	0.045	0.062	0.083	0.081
450	0.020	0.025	0.033	0.045	0.058	0.083
500	0.019	0.022	0.028	0.039	0.055	0.073

Table 1: Expected and obtained limits on t' production cross section for given mass.

the probability for observing 29 or more events given 18.03 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}.

n	Min M_{rec} [GeV/c^2]	Min H_T [GeV]	observed	expected	p-value
1	475	775	0	0.021	1.000
2	450	750	0	0.116	1.000
3	425	725	1	0.228	0.2040
4	400	700	2	0.371	0.0540
5	375	675	3	0.718	0.0364
6	350	650	4	1.503	0.0660
7	325	625	4	2.876	0.3251
8	300	600	12	5.498	0.0110
9	275	575	14	9.885	0.1273
10	250	550	29	18.03	0.0105
11	225	525	41	31.34	0.0555
12	200	500	58	52.05	0.2219
13	175	475	92	91.14	0.4779
14	150	450	152	158.7	0.7141
15	125	425	222	231.0	0.7318

Table 2: 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. The minimum H_T and M_{rec} in each trial are also shown.

References

[1] J. Conway *et al.*, CDF note 9209.

[2] Data vs. Data Comparisons.
`http://www-cdf.fnal.gov/~dcox1/internal/Validation/DatavsDataComparison.html`