Fermi National Accelerator Laboratory

FERMILAB-Conf-94/247-E

CDF

CDF Top Results in the Lepton + Jets Channel

Gordon Watts
The CDF Collaboration

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

August 1994

Proceedings of the *Eighth Meeting of the Division of Particles and Fields of the American Physical Society (DPF'94)*, Albuquerque, NM, August 2-6, 1994.

Operated by Universities Research Association Inc. under Contract No. DE-AC02-76CHO3000 with the United States Department of Energy

CDF/PUB/TOP/PUBLIC/2773
FERMILAB-CONF-94/247-E

CDF TOP RESULTS IN THE LEPTON + JETS CHANNEL

GORDON WATTS*
Department of Physics and Astronomy,
University of Rochester,
Rochester, New York 14627, USA

ABSTRACT

Results from the 1992-1993 Tevatron run (run IA) top search at CDF in the lepton+jets channel are reported. The jet-vertexing algorithm tags 6 events on a background of 2.3 ± 0.3 events in the $W+ \geq 3$ jets data sample, and the soft lepton analysis finds 7 events on a background of 3.1 ± 0.3 events in the same data sample. Jet-vertex and soft-lepton tag correlations are described. Also given are the expected and observed number of tags in the Z+mulitjet control sample.

1. The Lepton+Jets Channel

In the minimal standard model, top decays almost exclusively by $t\bar{t} \rightarrow W^+bW^-\bar{b}$. The top search in lepton+jets looks for a lepton from one of the Ws and jets from the other ($W^+ \rightarrow e\nu$ or $\mu\nu$ and $W^- \rightarrow$ jets, or the charge conjugate). Top decays to this final state about 30% of the time. Events selected for the lepton+jets search are required to have an isolated lepton with E_T (P_T for muons) > 20 GeV and $|\eta| < 1.0$, and to have $\not{E}_T > 20$ GeV.[1] Events containing Z bosons, with an ee or $\mu\mu$ invariant mass between 70 and 110 GeV/c^2, are removed. The dominant background in the lepton+jets top search is the direct production of W+jets. The ratio of the $t\bar{t}$ signal to W+jets background can be greatly improved by requiring $N_{jet} \geq 3$. This requirement has a rejection factor of $\approx$400 against inclusive W production while keeping approximately 75% of the $t\bar{t}$ signal for $M_{top} = 160$ GeV/c^2. In the $W+ \geq$3-jet sample, we expect 12 ± 2 (6.6 ± 0.7) $t\bar{t}$ events for $M_{top} = 160$ (180) GeV/c^2 using the acceptance discussed below and the theoretical cross section.[2] We observe 52 events with $N_{jet} \geq 3$ in the W sample.

The VECBOS[3] Monte Carlo program is used to make an estimate of the direct W+jets background, and predicts 46 events with ≥ 3 jets in Run 1A's $19.3pb^{-1}$ of data. There are, however, uncertainties of about a factor of two due to the choice in the Q^2 scale used in the calculation. We have, therefore, developed a technique for estimating backgrounds in the lepton+jets search directly from the data, described below. Other backgrounds (direct $b\bar{b}$, Z bosons, W pairs, and hadrons misidentified as leptons) contribute an additional 12.2 ± 3.1 events.[1] Additional background rejection is needed to isolate a possible $t\bar{t}$ signal. Requiring the presence of a b-quark jet, tagged either by a secondary vertex or by a semileptonic decay, provides such a rejection.

*Representing the CDF Collaboration at Fermilab.

Published Proceedings Eighth Meeting of the Division of Particles and Fields of the American Physical Society (DPF'94), University of New Mexico, Albuquerque, NM, August 2-6, 1994.

1.1. *Jet Vertex Search*

The lifetime of b hadrons can cause the b-decay vertex to be measurably displaced from the $\bar{p}p$ interaction vertex. CDF's new Silicon Vertex Detector (SVX)[4] is a high resolution $r - \phi$ tracker capable of resolving a displaced B decay vertex.

Found high quality SVX tracks ($P_T \geq 2$GeV/c, $|d|/\sigma_d \geq 3.0$, and with clean SVX reconstruction)[1] are associated with jets that have $E_T > 15$ GeV and $|\eta| < 2.0$ if cos(jetaxis − track) > 0.8. For each jet a transverse decay length (L_{xy}) and its uncertainty (typically $\sigma_{L_{xy}} \approx 130$ μm) are calculated using a three-dimensional fit, with the jet's tracks constrained to originate from a common vertex. Jets that have a secondary vertex displaced in the direction of the jet, with significance $L_{xy}/\sigma_{L_{xy}} \geq 3.0$ are positively tagged ($+L_{xy}$), and those with a secondary vertex behind the primary, with significance $L_{xy}/\sigma_{L_{xy}} \leq 3.0$ are negatively tagged ($-L_{xy}$). The largest contributions to the negative tags are resolution smearing and mistakes in the tracking reconstruction.

We use a control sample of inclusive electrons, which are enriched in b-decays, ($E_T > 10$ GeV) to measure an efficiency for SVX-tagging of a semileptonic b jet.[1] We compare this efficiency with that predicted by the ISAJET + CLEO[5] $b\bar{b}$ Monte Carlo and find our measured efficiency to be lower than the Monte Carlo prediction by a factor of 0.72±0.21. To verify that our tags are due to heavy flavor, the transverse decay length, L_{xy}, is converted into an estimate of the effective proper decay length ($c\tau_{eff}$) using the expression $c\tau_{eff} = L_{xy}\frac{M}{P_T}F$, where M is the invariant mass of the tracks associated with the displaced vertex, P_T is their total vector transverse momentum, and F is a scale factor, determined from a b Monte Carlo sample. The scale factor F accounts for b-hadron decay products that are not attached to the secondary vertex. Figure 1 is the $c\tau_{eff}$ of both the tags in the inclusive electron sample (points) and the tags in $b\bar{b}$ Monte Carlo (histogram); the two agree.

We then determine the efficiency for tagging at least one b jet in a $t\bar{t}$ event with three or more observed jets, ϵ_{tag}, from $t\bar{t}$ Monte Carlo. The efficiency must be rescaled by 0.72, as determined above. We find $\epsilon_{tag} = 22 \pm 6\%$ independent of top mass for $M_{top} > 120$ GeV/c^2. The efficiency, ϵ_{SVX}, for inclusive $t\bar{t}$ events to pass the lepton-identification, kinematic, and SVX b-tag requirements is shown in Table 1. The number of expected SVX-tagged $t\bar{t}$ events with $N_{jet} \geq 3$ is shown in the same table. Six SVX-tagged events are observed in the 52-event $W+ \geq$3-jet sample.

Rather than rely on Monte Carlo predictions, we estimate directly from our data how many tags we would expect in the 52-event sample if it were entirely background. We assume that the heavy-quark (b and c) content of jets in W+jets background events is the same as in an inclusive-jet sample.[1] This assumption is expected to be conservative, since the inclusive-jet sample contains heavy-quark contributions from direct production (e.g. $gg \to b\bar{b}$), gluon splitting (where a final-state gluon branches into a heavy-quark pair), and flavor excitation (where an initial-state gluon excites a heavy quark in the proton or antiproton sea), while heavy quarks in W+jets background ($Wb\bar{b}$, $Wc\bar{c}$) events are expected to be produced almost entirely from gluon splitting.[6]

We apply the tag rates measured in the inclusive-jet sample, parameterized by the E_T, track multiplicity and η of each jet, to the jets in the 52 events to yield the total expected number of SVX-tagged events from $Wb\bar{b}$, $Wc\bar{c}$, and fake tags due to track mismeasurement. We have tested this technique in a number of control samples

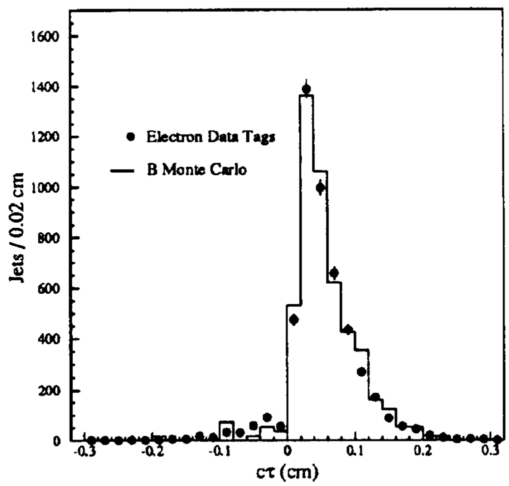

Fig. 1. The $c\tau_{eff}$ distribution for jets with a secondary vertex in the inclusive electron data (points with errors) compared to a Monte Carlo simulation (histogram) with the world average B lifetime.

M_{top}	120 GeV/c^2	140 GeV/c^2	160 GeV/c^2	180 GeV/c^2
ϵ_{SVX}	$1.0 \pm 0.3\%$	$1.5 \pm 0.4\%$	$1.7 \pm 0.5\%$	$1.8 \pm 0.6\%$
ϵ_{SLT}	$0.84 \pm 0.17\%$	$1.1 \pm 0.2\%$	$1.2 \pm 0.2\%$	$1.3 \pm 0.2\%$
$\sigma_{t\bar{t}}^{Theor}$ (pb)	$38.9^{+10.8}_{-5.2}$	$16.9^{+3.6}_{-1.8}$	$8.2^{+1.4}_{-0.8}$	$4.2^{+0.6}_{-0.4}$
$N_{expected}(SVX)$	7.7 ± 2.5	4.8 ± 1.7	2.7 ± 0.9	1.4 ± 0.4
$N_{expected}(SLT)$	6.3 ± 1.3	3.5 ± 0.7	1.9 ± 0.3	1.1 ± 0.2

Table 1. Summary of top acceptance (including branching ratios) and the theoretical cross section.[2] The middle line gives the $t\bar{t}$ production cross section obtained from this measurement. The last two lines are the expected event yield for $t\bar{t}$ in Run 1A's $19.3pb^{-1}$.

and use the level of agreement with the number of observed tags to determine the systematic uncertainty on the predicted tag rate.

The backgrounds from non-W sources (direct $b\bar{b}$ production and hadrons misidentified as leptons) are also determined from the data[1] by studying the isolation of the lepton candidates in the low $\not\!\!E_T$($\not\!\!E_T$< 15GeV/c) and high $\not\!\!E_T$($\not\!\!E_T$> 20 GeV/c) regions.[7] The total number of non-W background events in the signal region (high $\not\!\!E_T$ and isolated lepton) is estimated as the number of non-isolated lepton candidates in the high $\not\!\!E_T$ region scaled by the ratio isolated to non-isolated lepton candidates in the low $\not\!\!E_T$ region (dominated by backgrounds). To predict the number of tagged non-W events in the signal region we scale by the tagging rate in the low $\not\!\!E_T$, low isolation region.

The small contributions from Wc, WW, WZ production, and $Z \rightarrow \tau\tau$ are estimated with Monte Carlo. The total estimated background, including the inclusive jet prediction, to SVX tags in the 52-event sample is 2.3±0.3 events. Approximately 85% of this background is from heavy flavor and fake tags.

An alternate background estimate, using Monte Carlo calculations of the heavy-quark processes in W+jets events and a fake-tag estimate from jet data, predicts a heavy-quark content per jet approximately a factor of three lower than in inclusive-jet events and gives an overall background estimate a factor of 1.6 lower than the number presented above, supporting the conservative nature of our background estimate.

1.2. Soft Lepton Search

A second technique for tagging b quarks is to search for leptons arising from the decays $b \rightarrow \ell\nu X$ ($\ell = e$ or μ), or $b \rightarrow c \rightarrow \ell\nu X$. Because these leptons typically have lower P_T than leptons from W decays, we refer to them as "soft lepton tags", or SLT tags. We require lepton $P_T > 2$ GeV/c. To keep this analysis statistically independent of the dilepton search, leptons that pass the dilepton requirements are not considered as SLT candidates.

In searching for electrons from b and c decays, each CTC track is extrapolated to the calorimeter, and a match is sought to an electromagnetic cluster consistent in size, shape, and position with expectations for electron showers. The efficiency of the electron selection criteria, excluding isolation cuts, is determined from a sample of electron pairs from photon conversions, where the first electron is identified in the calorimeter and the second, unbiased, electron is selected using a track-pairing algorithm. Combined with electron isolation efficiencies determined from $t\bar{t}$ Monte Carlo, the total efficiencies are (53±3)% and (23±3)% (statistical uncertainties only) for electrons from b and sequential c decays respectively. To identify muons, track segments in the muon chambers are matched to tracks in the CTC. The efficiency for reconstructing track segments in the muon chambers is measured to be 96% using $J/\psi \rightarrow \mu^+\mu^-$ and $Z \rightarrow \mu^+\mu^-$ decays. This number is combined with the P_T-dependent efficiency of the track-matching requirements to give an overall efficiency of approximately 85% for muons from a combination of both b and c decays.

The acceptance of the SLT analysis for $t\bar{t}$ events is calculated using the ISAJET + CLEO Monte Carlo programs. The efficiency for tagging at least one jet in a $t\bar{t}$ event by detecting an additional lepton with $P_T > 2$ GeV/c is $\epsilon_{tag} = 16 \pm 2\%$, approximately independent of M_{top}. The efficiency, ϵ_{SLT}, for inclusive $t\bar{t}$ events to pass the lepton-identification, kinematic, and SLT b-tag requirements is shown in Table 1 along with

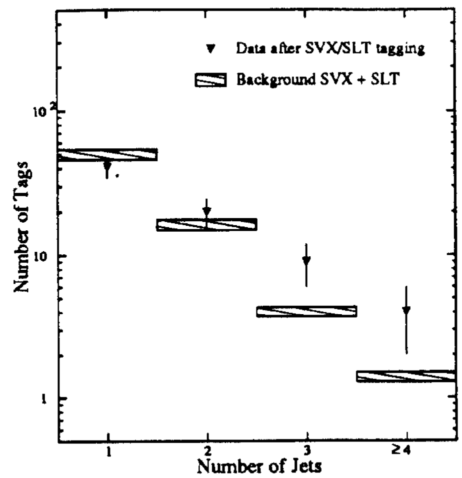

Fig. 2. The sum of SVX and SLT tags observed in the W+jets data (solid triangles). Events tagged by both algorithms are counted twice. The shaded area is the sum of the background estimates for SVX and SLT, with its uncertainty. The 3-jet and $\geq$ 4-jet bins are the $t\bar{t}$ signal region.

the number of expected SLT-tagged $t\bar{t}$ events. We find seven SLT-tagged events with $N_{jet} \geq 3$. Three of the seven also have SVX tags.

The main backgrounds to the SLT search are hadrons misidentified as leptons, and $Wb\bar{b}$, $Wc\bar{c}$ production. As in the SVX analysis, we estimate these backgrounds from the data by conservatively assuming that the heavy-quark content per jet in W+jets events is the same as in inclusive-jet events. By studying tracks in such events, we measure the probability of misidentifying a hadron as an electron or muon, or of tagging a true semileptonic decay. We use these probabilities to predict the number of tags in a variety of control samples, and obtain good agreement with the number observed. We expect 2.70±0.27 tags in the $W+ \geq 3$ jet sample from these sources. Other sources (direct $b\bar{b}$, W/Z pairs, $Z \rightarrow \tau\tau$, Wc, and Drell-Yan) contribute 0.36±0.09 events, for a total SLT background of 3.1±0.3 events. The number of SLT tags in the W+1 and W+2-jet samples, which should have only a small contribution from $t\bar{t}$, agrees with the background expectation (45 events tagged, 44±3.4 predicted). Figure 2 shows the combined number of SVX and SLT tags, together with the estimated background, as a function of jet multiplicity.

2. Jet-Vertex and Soft Lepton Correlations

The correlations between the SVX and SLT taggers must be understood before we can understand their combined significance. We do this by looking at a large sample of jets. Because the jet vertexing's negative decay ($-L_{xy}$) length tags are not expected to be enriched in heavy flavor (as the positive tags are), they are not expected to be correlated with soft lepton tags. Applying the soft lepton tagger and its prediction to the $-L_{xy}$ tags yields a observed-minus-predicted rate that differs from zero by $+1.3\sigma$

for double tagged jets and $+1.0\sigma$ for double tagged events. Note that an exact match means only that correlations in the two taggers are accounted for in the background prediction. In the case of the $-L_{xy}$ sample, the correlations are accounted for by the parameterization.

The SVX $+L_{xy}$ sample is enriched in heavy flavor, and thus the SLT prediction is much smaller than the observed number of SVX-SLT double tags. If we assume the excess jet vertex tags, $(+L_{xy}) - (-L_{xy})$, are due to heavy flavor, we can calculate a second prediction which closely matches the number of observed double tags. We conclude that the correlation in the $-L_{xy}$ sample is properly modeled by the tag rate probabilities, and the correlation in the $+L_{xy}$ sample is understood as resulting from the heavy flavor content of the sample. Monte Carlo calculations predict 1.8 $\pm$1.3 double tagged events from $t\bar{t}$ and background, thus the three double tags in the 52 signal events are consistent with the $t\bar{t}$ hypothesis.

3. Tags in the Zs

The Z+multijet sample is, in principle, a good cross check of the heavy flavor content of the W+multijet sample. The standard model production mechanisms are similar in Ws and Zs, but there should be no top in the Z sample. In practice, however, $\sigma \times BR$ for $Z \rightarrow ee, \mu\mu$ is an order of magnitude less than for $W \rightarrow e\nu, \mu\nu$.

Events with oppositely charged ee or $\mu\mu$ pairs with 75 GeV/c^2 $< m_{ee}, m_{\mu\mu} <$ 105 GeV/c^2 are selected. The first lepton is required to pass the same cuts as the high P_T leptons, and the second lepton has a E_T cut of 10 GeV (with a tighter calorimeter isolation cut to remove background).

Backgrounds are calculated using the same method we used in the W+jets background calculation, but no backgrounds from non-Z sources are included ($t\bar{t}$, $b\bar{b}$, etc.). 8 tags are observed, and 5.8 $\pm$ 0.4 are predicted. If we look in only the $Z+\geq 3$ jets region, equivalent to the signal sample in the top search, we find 2 events with a background of .64 events. This is a low statistics check and we are currently taking more data so that we can study the Z+multijets sample further.

4. Conclusion

The jet vertex search finds 6 tags in the $W+\geq 3$ sample with a background of 2.3 $\pm$ 0.3 events. Using a background estimation technique that relies more on Monte Carlo, we see a background of 1.4 $\pm$ 0.5. The soft lepton search finds seven tags with a background of 3.1 $\pm$ 0.3 events. Correlations between the two taggers are understood. A low statistics check of the tagging algorithms has been made in the Z+multijet sample; more data is required to provide conclusive results.

References

1. F Abe et *el.*, Fermilab-Pub-94/097-E, submitted to *Phys. Rev.* **D**.
2. E. Laenen, J. Smith, and W.L. Van Neerven, *Phys. Lett.* **B321** (1994) 254.
3. F.A. Berends, W.T. Giele, H. Kuijf, and B. Tausk, *Nucl. Phys.* **B357** (1991) 32.
4. Fermilab-94/024-E, submitted to *Nucl. Instrum. Methods Phys. Res.*.
5. E. Paige and S.D. Protopopescu, *BNL Report* No. 38034, (1986).
6. M.L. Mangano, *Nucl. Phys.* **B405** (1993) 536.
7. F. Abe *et al.*, *Phys. Rev.* **D44** (1991) 29.