Fermi National Accelerator Laboratory

FERMILAB-Conf-94/238-E

CDF

Studies of Kinematic Distributions and Mass Reconstruction of CDF Top Candidate Events

Brian D. Harral

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

Disclaimer

CDF/PUB/TOP/PUBLIC/2745
FERMILAB-CONF-94/238-E

STUDIES OF KINEMATIC DISTRIBUTIONS AND MASS RECONSTRUCTION OF CDF TOP CANDIDATE EVENTS

BRIAN D. HARRAL*
Department of Physics, University of Pennsylvania,
209 S 33rd St, Philadelphia PA 19104-6396, USA

ABSTRACT

We present analyses of CDF top-candidate events which study the separation of signal from background using easily-defined kinematic variables, and which study invariant masses of three-body systems to fully reconstruct the top quark decay products.

1. Introduction

To check whether the CDF data are consistent with a significant top content, various techniques involving the event kinematics have been attempted. One method to be discussed here is to find simple kinematic quantities which can provide sufficient background rejection. We find that using various combinations of jet energies provides the needed discrimination. We also discuss our method of fully reconstructing events under the hypothesis of top decays, which provides us with our estimate of the top mass. Both methods use the "lepton+jets" event sample described in these proceedings[3] as opposed to the "dilepton" events[4] to take advantage of the increased branching ratio and the less-ambiguous event structure.

2. Kinematics

We expect that the most significant background to the top signal in lepton+jets is standard production of W bosons with associated QCD production of jets; our estimates of other backgrounds (due to lepton misidentification, Z decays, dibosons, and others) are at least an order of magnitude lower.[1] The jets produced by QCD processes are produced, on average, with less transverse energy than jets from $t\bar{t}$ events. This leads reasonably to comparisons between jet energies in $t\bar{t}$ events and in W+jet events from Monte Carlo.

To generate $t\bar{t}$ events, we have used both the ISAJET Monte Carlo[5] (using the CLEO fragmentation scheme for b hadrons[6]) and the HERWIG Monte Carlo[7] for purposes of comparison. We use the VECBOS Monte Carlo[8] to generate W+jet events.

Jets are defined by clusters of energy found in the calorimeters using a cone of fixed size in η-ϕ space.[9] In both Monte Carlo and detector data, the calorimeter energy has a correction applied to take into account detector nonlinearities, energy lost in

*Representing the CDF Collaboration.

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.

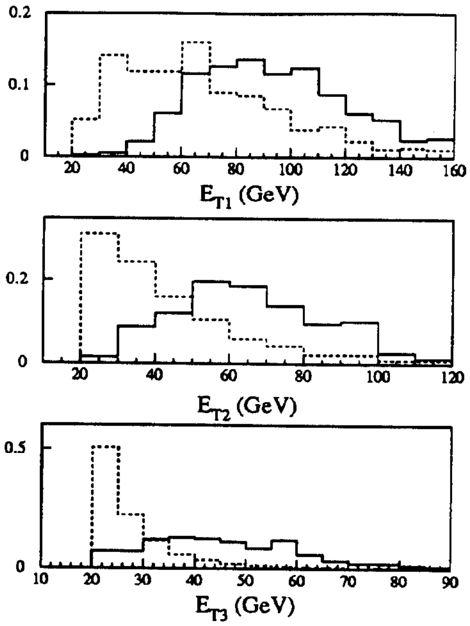


Fig. 1. Comparison of first-, second-, and third-highest E_T jets in ISAJET (solid line) and VECBOS (dashed line) events. Curves are normalized to unit area.

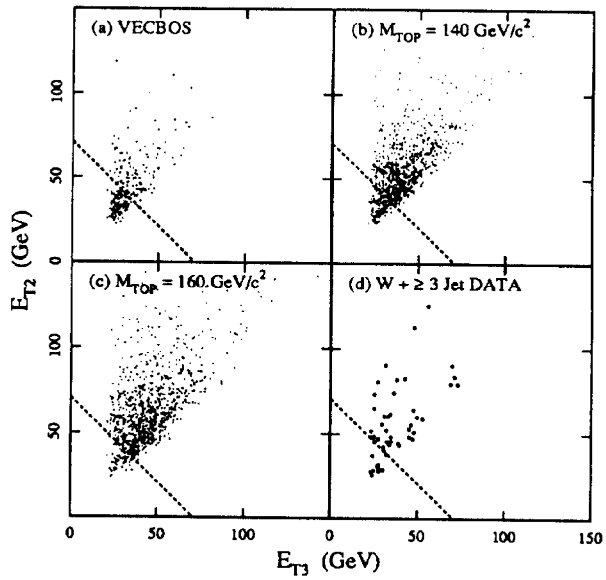
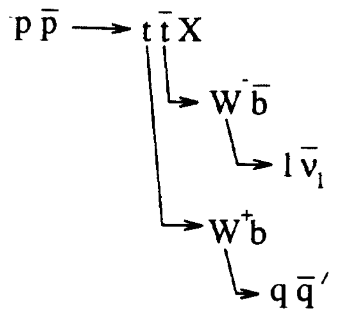


Fig. 2. Plot of second-highest E_T jet *vs.* third-highest for background MC, m_{top} = 140 MC, m_{top} = 160 MC, and detector data. The diagonal line ($E_{T2} + E_{T3}$ = 71 GeV) is chosen such that half the background events lie below it.

gaps between detectors, energy lost due to particles falling outside the fixed-size cone, and other effects.[10] The jet energy transverse to the beam line (E_T) is compared between VECBOS and $t\bar{t}$ Monte Carlos after this correction has been applied. Figure 1 shows this comparison for VECBOS and ISAJET (m_{top} = 170 GeV) for the highest, second-highest, and third-highest E_T jets in events which pass the W+$\geq$ 3-jet criteria.[3] The difference between background and signal becomes more evident when scatter plots of the second-highest *vs.* the third-highest E_T jets are viewed (Fig. 2). The 52 events of the detector data are also shown in Fig. 2. Top events are expected to populate the high-E_T region of this plot much more heavily than W+jet background; of the 52 events, 39 lie in the high-E_T region defined in Fig. 2, and of the 10 tagged events, eight are in this region. Work is currently in progress to reduce systematic uncertainties (mainly jet energy scale and jet energy corrections), to improve understanding of the background, and to make the analysis more quantitative.

3. Mass Analysis

The mass analysis adds a loose fourth-jet requirement to the W+3-jet event selection; forms appropriate invariant masses with the four highest-E_T jets, the lepton,



	known	unknown
p's	8	0
t's	1	7
X	2	2
W's	2	6
b's	8	0
q's	8	0
l	4	0
ν	1	3
	34	18

5 vertices → 20 equations
13 4-vectors → 52 variables
18 unknowns → "2C" fit

Fig. 3. Counting constraints in W+4-jet event. See Ref. 1 for more details.

and a vector representing the assumed neutrino; and performs a constrained fit under the hypothesis of top decay. Rather than taking the attitude that an invariant mass analysis can separate signal from background, we use the heavy-flavor tagging to provide both background reduction and reduction in combinatorics. In the hypothesized decay system as we specify it (see Fig. 3), there are two more constraints than unknowns, thus allowing a fit. Note that the dilepton system would then have one more unknown than constraints; work is in progress in deciding the best method of removing this degree of freedom.

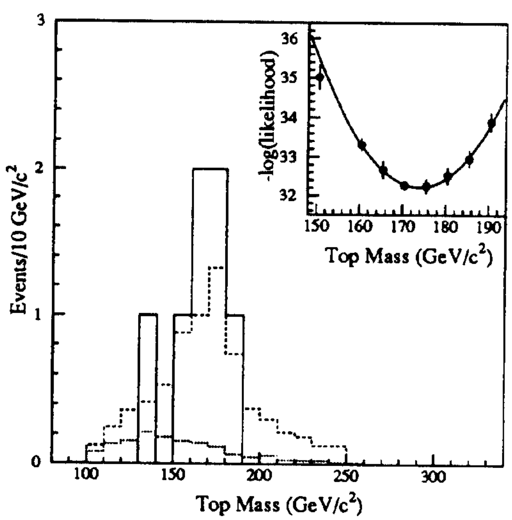


Fig. 4. Top mass distribution for data (solid histogram), W+jets background (dots), and the sum of background and Monte Carlo with m_{top} = 175 GeV (dashes). Background is normalized to 1.4 events. Inset shows likelihood fit for top mass.

In order to completely reconstruct the decay products of the hypothesized top decays, we must therefore require at least four jets in the events. Because of our small amount of data, we loosen the cluster requirements on the fourth jet to $E_T >$ 8 GeV, $|\eta| < 2.4$ to improve our statistics. We also require the presence of an SVX or SLT tagged jet in the event; of the ten tagged events observed in the 1992-93 CDF run, seven pass the loose fourth-jet requirement. Of these, we estimate a contribution of $1.4^{+2.0}_{-1.1}$ events from non-$t\bar{t}$ processes.

In forming invariant masses (under the assumptions of Fig. 3), we consider only the four highest-E_T jets in each event. Of the 12 ways to associate these four jets with the final-state quarks, we consider only those where the tagged jet is assigned to a b quark, leaving six possibilities (reducing combinatoric background and improving mass resolution); however, there are in general two possible solutions for the z-momentum of the neutrino satisfying the W mass constraint, giving 12 potential event configurations.

The main difficulty (and source of systematic error) is making the connection between energy measured in the calorimeter and quark momentum; the correction function referred to in Section 2 above is not sufficient, mainly due to the possible presence of muons and neutrinos in b-quark jets. We have therefore developed corrections specifically for jets from W decays, and for b jets with and without semileptonic decays, and applied these to correct the jets in the candidate events before mass fitting. Details of these corrections can be found in Ref. 1.

A χ^2 is formed using the uncertainties in measurement of the leptons and the uncertainty in making the jet-quark connection mentioned above. This χ^2 is minimized subject to the mass constraints of the problem for each of the 12 event configurations. We choose the lowest-χ^2 configuration with $m_{top} < 260$ GeV as giving our best estimate of the mass for that event.

To get a best mass for the experiment, we begin by computing mass spectra for $t\bar{t}$ Monte Carlo events at various values of m_{top}, and for W+jet Monte Carlo. We then fit our observed mass distribution to a linear combination of signal and background, subject to our background estimate given above. The likelihood for these fits can then be plotted as a function of m_{top} to find the best overall mass and its statistical error. This information is summarized in Fig. 4, with the inset plot showing the likelihood fit and the main plot displaying the mixture of signal and background spectra superimposed on the data. Details of systematic errors can be found in Ref. 1; our result is $m_{top} = 174 \pm 10^{+13}_{-12}$ GeV/c^2.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Science and Culture of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the A.P. Sloan Foundation; and the Alexander von Humboldt-Stiftung.

References

1. F. Abe *et al.*, FERMILAB-PUB-94-097-E (1994), to be published in Phys. Rev. D.
2. F. Abe *et al.*, Phys. Rev. Lett. **73** (1994) 225.
3. G. Watts, these proceedings.
4. J. Benlloch, these proceedings.
5. F. Paige and S. Protopopescu, BNL Report No. 38034 (1986).
6. P. Avery, K. Read, G. Trahern, Cornell Internal Note CSN-212 (1985).
7. G. Marchesini *et al.*, Comput. Phys. Comm. **67** (1992) 465.
8. F.A. Berends *et al.*, Nucl. Phys. **B357** (1991) 32.
9. F. Abe *et al.*, Phys. Rev. **D45** (1992) 1448.
10. F. Abe *et al.*, Phys. Rev. **D47** (1993) 4857.