Fermi National Accelerator Laboratory

FERMILAB-Conf-95/111-E
CDF

Observation of Top Quark Production Using Kinematic Techniques

Shinhong Kim

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

University of Tsukuba

May 1995

Proceedings of the *QCD and High Energy Interaction*, *XXXth Rencontres de Moriond*, Les Arcs, France, March 19-26, 1995

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

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

CDF/ANAL/TOP/PUBLIC/3135
May 17, 1995

Observation of Top Quark Production Using Kinematic Techniques

Shinhong Kim
University of Tsukuba
for the CDF Collaboration
Proceedings of the "QCD and High Energy Interaction, XXXth Rencontres de Moriond", Les Arcs, France, March 19-26, 1995

Abstract

We present analysis results of the top quark production using kinematic techniques in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.8 TeV with the Collider Detector at Fermilab (CDF). We analyzed a data sample corresponding to an integrated luminosity of 67 pb^{-1}. First, we compare the kinematic properties, the second and third leading jet E_T of $W+ \geq 3$ jet events with expectations for $t\bar{t}$ production and its background processes, predominantly direct $W+$ jet production. The probability that backgrounds fake the excess in the signal region was found to be less than 0.26%. Next, we use a scaler sum of the transverse energies of the lepton, neutrino and jets in W+$\geq$4 jet events. This analysis shows $t\bar{t}$ events exist significantly in the event sample, and enables us to estimate the top quark mass to be $M_{top} = 180 \pm 12\ ^{+19}_{-15}$ GeV/c^2. Finally, we analyse the dijet invariant mass in the W+$\geq$4 jet events. It shows a significant mass peak consistent with a mass peak of W decaying into dijet. From these results and counting experiment, we have confirmed previous evidence reported in 1994 April and observed top quark production.

1 Introduction

In 1994 April we reported evidence for top quark production from the Collider Detector Experiment (CDF) with 19.3 pb^{-1} data collected in 1992-93 (run 1A)[1] and also reported a detailed top kinematic study with this data sample [2], [3]. In 1994-95 (run 1B), we are collecting additional data, and analyzed a part of the data corresponding to an integrated luminosity of 48 pb^{-1}. The observation of top events was confirmed by analysis of kinematic properties described in this paper and an improved version of the counting experiment, using the 67 pb^{-1} data sample [4] with a significance of 4.8 σ. The D0 Collaboration has also observed the top quark production in an analysis based in part on kinematic information [5].

We studied kinematic properties of $t\bar{t}$ production in $W+ \geq 3$ jet channel where top (anti-top) quark decays to W and bottom (anti-bottom) quark, and one W decays leptonically and another W decay hadronically :

$$p\bar{p} \rightarrow t\bar{t}X \rightarrow W^+bW^-\bar{b}X \rightarrow l^{\pm}\nu bjj\bar{b}X \tag{1}$$

The expected $W+$ jets background is calculated with the VECBOS Monte Carlo, which uses the lowest order $W+$ n parton matrix elements (n = 1,2,3 and 4) [6]. The parton fragmentation process is simulated with a HERWIG type shower module [7]. The lowest order matrix elements are sensitive to the choice of the mass scale in the strong coupling constant α_s. We used two resonable Q^2 scales of the square of the W boson mass ($M_W{}^2$) and the square of the average P_T of jet ($< P_T >^2$). Top events are simulated by the HERWIG program. A more detailed desciption of the Monte Carlo programs can be found elsewhere [2], [3].

In section 2 we report the results of the analysis using the second and third leading jet E_T. In section 3 we report on a study using the total transverse energy of $W+ \geq$4 jet events, which leads to evidence for a $t\bar{t}$ component in the data. In section 4 we show the study of dijet mass in $W+ \geq$4 jet events, which gives evidence for a W mass peak from hadronic decaying W.

2 the Second and Third Leading Jet E_T

We selected $W+ \geq 3$ jet events by requiring an isolated high P_T lepton with E_T(electron) $>$ 20 GeV or P_T(muon) $>$ 20 GeV/c, the large missing E_T of $\not{E}_T$ >25 GeV, the large W transverse mass of M_T $>$ 40 GeV/c^2, and three jets with E_T(jet) $>$ 20 GeV. Jets

are reconstructed with a cone of R =0.4 [1]. Jet energies are corrected [8] already before event selection. The three leading jets are required to be separated from each other by $\Delta R \geq 0.7$. We also require $|\eta(\text{jet})| < 2.0$. This $W+$ $\geq$3 jet event sample contains 158 events from the total 67 pb^{-1} data.

We apply an additional cut to the $W+ \geq 3$ jet events to improve the signal to noise ratio. This another variable θ^* is the jet polar angle in the center of mass system of W and jets with $E_T >$ 15 GeV; as we do not know the longitudinal momentum of the neutrino, we assume it to be zero when calculating $\cos\theta^*$. The $\cos\theta^*$ distribution is expected to be more peaked in the forward direction for QCD $W+$ jets than for $t\bar{t}$. We define a top enriched sample ("signal sample") by $|\cos\theta^*_{max}(\text{jet})| < 0.7$, and a complementary background enriched sample ("control sample") by $|\cos\theta^*_{max}(\text{jet})| > 0.7$, where the variable $|\cos\theta^*_{max}(\text{jet})|$ is the largest value of $|\cos\theta^*(\text{jet})|$ for any of the leading three jets.

We show plots of E_{T2} versus E_{T3} [2] for the CDF data, VECBOS QCD $W+$ jet events and HERWIG $t\bar{t}$ events with the cuts of the signal sample in Fig. 1. In the signal sample the data does not agree with the VECBOS prediction alone. Most of the CDF b-tagged events [3] in this figure are found to be in the region where top events are expected.

In order to compare the CDF data with Monte Carlo simulation data quantitatively, we calculate a "relative likelihood" for each event, as a measure of whether the event is more "top-like" or more "QCD-like". The relative likelihood is defined in terms of jet E_{T2} and E_{T3} distributions for $t\bar{t}$ (M_{top}=170 GeV) and direct QCD $W+$ jets production as follows:

$$L = [(\frac{1}{\sigma}\frac{d\sigma}{dE_{T2}}) \times (\frac{1}{\sigma}\frac{d\sigma}{dE_{T3}})]^{t\bar{t}} / [(\frac{1}{\sigma}\frac{d\sigma}{dE_{T2}}) \times (\frac{1}{\sigma}\frac{d\sigma}{dE_{T3}})]^{QCD} \tag{2}$$

We show ln(L) distributions for signal samples of Monte Carlo events and the CDF data events in Figs. 2(a) and (b), respectively. The CDF data contains 25 events in ln(L) $<$ 0 and 22 events in ln(L) $>$ 0. The distribution of events between positive and negative ln(L) is similar to that observed in the data of Run 1A [2], [3]. The QCD Monte Carlo predicts that not more than 22$\pm$5% of QCD $W+$ jet events will be in ln(L) $>$ 0, namely $t\bar{t}$ signal enriched region. Other backgrounds, such as non-W and WW events are expected to be few (6.9$\pm$1.8) in the signal sample and to have softer E_{T2} and E_{T3} distributions than the VECBOS QCD $W+$ jets prediction. Conservatively, we take the QCD background shape as predicted by VECBOS to represent the shape of all background.

If we make the conservative assumption that all events in ln(L) $<$ 0 are background and normalize the expected background distribution to the observed events in ln(L) $<$ 0,

[1] $R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$, where $\Delta\phi$ is the cone half–width in azimuth and $\Delta\eta$ is the cone half–width in pseudorapidity.

[2] Jets are ordered in E_T(jet) with $E_{T1} > E_{T2} > E_{T3}$.

[3] CDF is able to detect b-quarks with two algoritms. One algorithm identifies charged tracks originating from a secondary vertex, separated in space from the primary one ("SVX tag"). The other algorithm identifies electrons or muons in jets (soft lepton tag, or "SLT".)

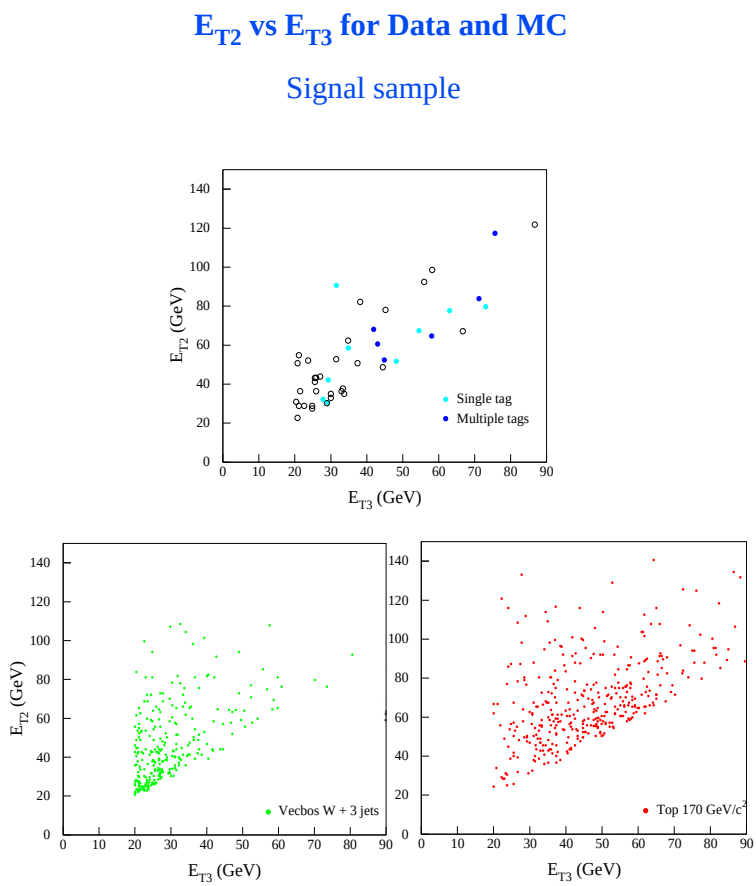


Figure 1: Signal sample : E_{T2} and E_{T3} distributions for for all data events (top), direct QCD W+jets (bottom left) , and top (M_{top}=170 GeV) Monte Carlo events (bottom right). Data events with a b-tag are shown separately.

Signal Sample

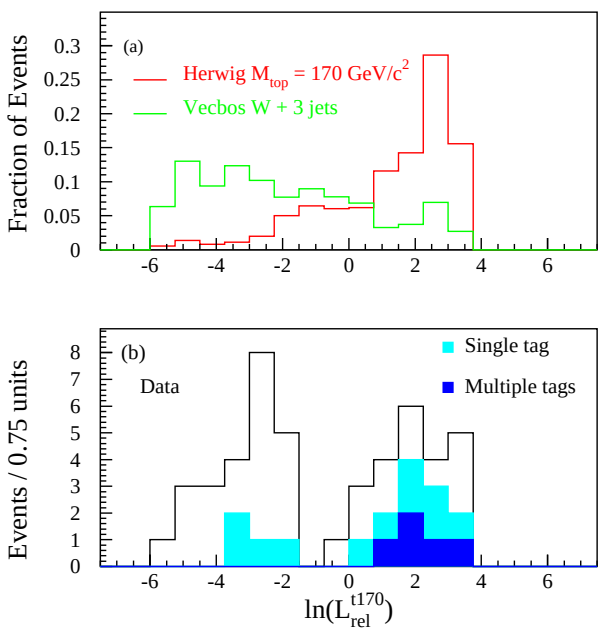


Figure 2: (a) VECBOS QCD and HERWIG top (M_{top} = 170 GeV/c^2) Monte Carlo predicted distributions for the $W+ \geq 3$ jet signal sample; $Q^2 = M_W^2$ is used in the VECBOS calculation. (b) Data; The shaded area indicates the b-tagged events from SVX and SLT. The darker area indicates events with more than one SVX or SLT tag.

then we would expect 7.2±2.1 events in ln(L) > 0 compared to the observed 22 events. A probability that the background fakes such an excess in the signal region of ln(L) > 0 was found to be less than 0.26% [9]. For this limit we took into account uncertainties of the VECBOS prediction, in the fragmentation modeling and in the reconstruction of primary parton energies from observed jet energies. This number should be compared to the previously quoted 0.8% obtained from 19.3 pb^{-1} data from Run 1A [3].

The shaded areas indicate the b-tagged events in Fig. 2(b). The darker areas indicate events with more than one SVX or SLT tag. We observe that a large number of events in the signal enriched event sample are b-tagged. A total of 13 SVX tags (in 8 events) are tagged compared to 2.80± 0.35 SVX tags expected from background alone. We observe that all 13 SVX tags are associated with events with ln(L) > 0. We evaluate the probability that the observed excess of b-tags in the signal sample is due to a statistical fluctuation of the backgrounds. Conservatively we consider the SVX tagged events only and ignore the fact that many events contain more than one b-tag. We find that the probability to observe 8 or more events when 2.80±0.35 are expected is 0.96×10^{-2}. All SVX b–tags occur in events with ln(L) > 0, that is the kinematic region where most top events are expected to be found. We expect only 1.37±0.17 SVX tags in ln(L) > 0 from background, compared to an observation of 8 events. The probability that this observation be due to a statistical fluctuation of background tags is $< 1.2\times10^{-4}$.

3 The Total Transverse Energy, H

We define a simple kinematical variable H as the scalar sum of the transverse energies of the lepton, neutrino and jets:

$$H = E_T(lepton) + \not{E}_T + \Sigma E_T(jet) \tag{3}$$

The variable H is strongly correlated with the center of mass energy ($\sqrt{\hat{s}}$) of the parton-parton hard scattering process. In $t\bar{t}$ events, it is also a measure of the transverse mass of the $t\bar{t}$ system. For a top mass larger than about 140 GeV/c^2, H has discriminating power between $t\bar{t}$ and $W+$ multijet events because the W and b quark resulting from top quark decay have higher P_T than radiated gluons in background processes. For this study no b-tagging or top event reconstruction is required. Therefore, this analysis contains more $t\bar{t}$ candidate events and has different and non-combinatorial systematic uncertainties compared to the event-fitting algorithm used to determine the top mass [4] and [5]. In addition, it is sensitive to top decay modes which do not involve b quarks. In this respect the study of H supplements our previous analysis [4] by providing a largely independent confirmation of the top quark mass. We report here on a search in the $W+ \geq 4$ jets channel based on the same dataset with 67 pb^{-1} integrated luminosity as in Ref. [4].

We selected W+jet events requiring leptonically decaying W with an isolated, high P_T lepton ($P_T > 20$ GeV/c) and high missing transverse energy ($\not{E}_T > 20$ GeV). In this analysis, three W+jet samples are defined :
(a) an exclusive W+3 jet sample of events with precisely three jets with Et(jet) > 8 GeV.
(b) a 'low threshold' sample of events with at least four jets with Et(jet) > 8 GeV.
(c) a 'high threshold' sample which is a subsample of (b), requiring in addition that the three leading jets have Et(jet) > 15 GeV.

Sample (a) is used as a background enriched sample since Monte Carlo calculations predict that it should contain a fraction of not more than 1% top events ($M^{top} = 175$ GeV/c^2), due to the bias against a fourth jet. The low and high threshold samples,(b) and (c), should be increasingly enriched with top events. For the event selection neither Et(jet) nor $\not{E}_T$ are corrected for detector effects. However, when calculating H, jet energies and $\not{E}_T$ are corrected by a rapidity and energy dependent factor, which accounts for calorimeter non-linearity and reduced response at detector boundaries [8]. This correction is applied in order to reconstruct the original parton energy as closely as possible.

We compare the H distributions for the CDF data to the expectation from direct QCD W+jets production, as calculated by the VECBOS Monte Carlo program in Fig. 3. In the W+3 jet (a) sample we find a good agreement between QCD prediction and data. In sample (b) we find an excess of data events at large H; This excess becomes more pronounced in sample (c).

Figure 4 shows the H distibution in sample (c) together with the distribution of the b-tagged events. The 'double peak' structure of the H distribution and the presence of b-tagged events at high H values suggests very clearly the presence of top events in the sample.

We performed a two-component fit to the H distribution of sample (c) using the predictions from the QCD and top Monte Carlo programs. We obtained the fit likelihood as a function of the assumed value for M_{top} for $Q^2 = < P_T >^2$. After fitting the data points in a plot of the fit likelihood versus M_{top} to a cubic polynomial, we find the top quark mass to be $M_{top} = 180 \pm 12$(stat.) $^{+19}_{-15}$(syst.) GeV/c^2. This top quark mass value is in excellent agreement with that found by full mass reconstruction of SVX and SLT tagged $W+ \geq 4$ jet events [4]. The two estimates are correlated since 12 $t\bar{t}$ events are common in the corresponding data samples. Because of this, we only quote the M_{top} estimate from this analysis to demonstrate its consistency with our earlier result of $M_{top} = 176 \pm 8 \pm 10$ GeV/c^2 [4]. This latter value remains our best estimate for M_{top}.

For a top mass of 180 GeV/c^2, and setting $Q^2 = < P_T >^2$ in the VECBOS calculation, the fit yields 56$\pm$ 10(stat.)$\pm$ 5(sys.) $t\bar{t}$ events in the high threshold signal sample. The corresponding number with $Q^2 = M_W{}^2$ is 45$\pm$ 11(stat.)$\pm$ 5(sys.). This can be compared with 34$\pm$ 10(stat.)$\pm$ 5(sys.) events as extrapolated from the number of b-tagged $W+ \geq 3$ jet events reported in [4]. These two estimates of the observed $t\bar{t}$ yield are consistent within statistical and systematic uncertainties.

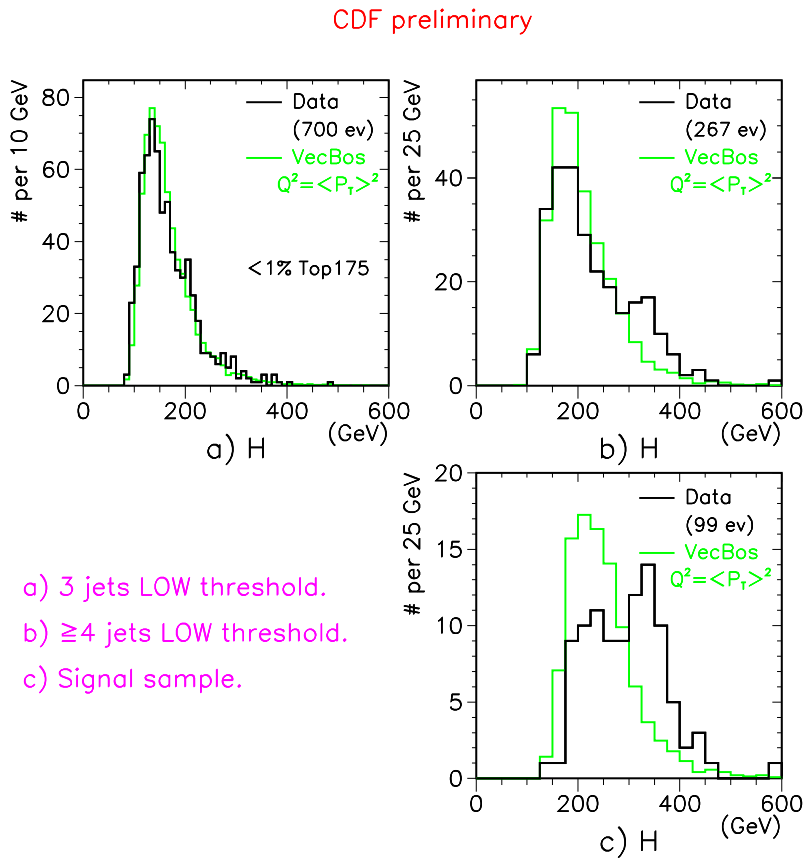


Figure 3: H distribution for data (solid line) and VECBOS events (dotted). VECBOS is normalised to the data. (a) Exclusive W+3 jet sample with E_T(jet) > 8 GeV; (b)W+≥4 jet sample with E_T(jets) > 8 GeV; (c)W+≥4 jet sample with E_T(jet1,2,3) > 15 GeV, E_T(jet4) > 8 GeV.

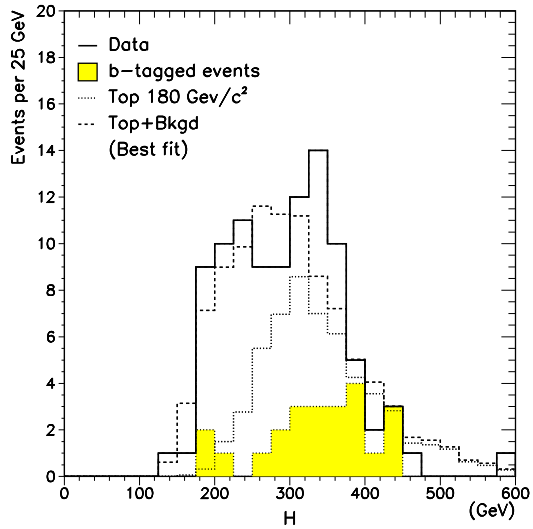


Figure 4: H distribution of events of the high threshold W+≥4 jets sample, and of the associated b-tags (SVX or SLT, shaded).

4 W decaying to dijet in $W+\geq$4 jet events

Jet energy scale uncertainty can be studied using W decaying to 2 jets in $t\bar{t}$ events. We can also prove that $t\bar{t}$ candidates in $W+\geq$4 jet channel have two W's in final state, by finding a W mass peak in dijet mass distribution. Motivated by these, we looked at dijet mass distribution in $W+\geq$4 jet events.

We used the high threshold $W+\geq$4 jets sample in the H analysis described in the preceeding section. We calculated invariant masses of any two jets out of the leading 4 jets. We had 6 combinations of two jets in each event. All jet energies were corrected with CDF standard jet energy correction. Since we have 6 combinations of two jets for each event, we have one signal of $W \rightarrow 2$ jets and five dijets of combinatorial backgrounds for each $t\bar{t}$ event. We fitted the dijet mass distribution to the following function excluding the W mass region between 60 GeV/c^2 and 100 GeV/c^2 as shown in Fig. 5:

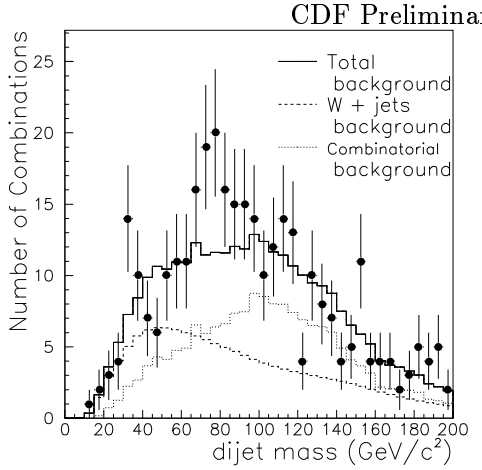
$$N[(1-\alpha)f_{comb}(m_{jj}) + \alpha f_{bg}(m_{jj})], \tag{4}$$

where N is a normalization factor, m_{jj} is a dijet mass, α is a fitting parameter of a fraction of the QCD $W+$ jets background, and f_{comb} and f_{bg} are the dijet mass distributions for the combinatorial backgrounds in $t\bar{t}$ events estimated with the HERWIG Monte Carlo program and the QCD $W+$jets background estimated with the VECBOS, respectively.

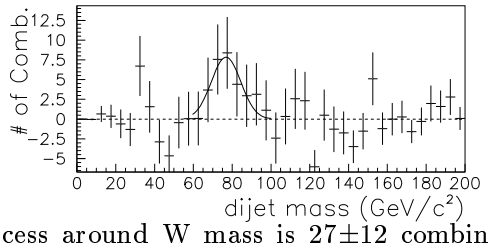
We extracted the dijet mass distribution of the $W \rightarrow 2$ jets signal by subtracting the QCD $W+$ jets background and the combinatorial background as shown in Fig. 5. We

Dijet Mass Distribution (H > 300 GeV)

CDF Preliminary



After background subtraction



Excess around W mass is 27±12 combinations (2.3σ). Number of expected W → 2 jets is 28^{+15}_{-10}. Dijet mass peak is consistent with a W mass peak.

Figure 5: Dijet mass distribution in $W+ \geq 4$ jet events after H > 300 GeV cut together with fitted curves of backgrounds in top plot. That after subtracting the background in bottom plot

fitted this distribution to a Gaussian function and obtained the number of W $\rightarrow$ 2 jets of 27 ± 12 (2.3σ) which is consistent with the number of $W \rightarrow 2$ jets of 28^{+15}_{-10} expected from $t\bar{t}$ production cross section [4]. The dijet mass peak is also consistent with that expected from $W \rightarrow 2$ jets in peak position and width within a statistical error.

5 Conclusions

We analysed $W+ \geq 3$ jet events using the second and third leading jet E_T, E_{T2} and E_{T3}, and found a significant excess in the top signal region of the top vs. QCD relative likelihood with E_{T2} and E_{T3}. This excess has a probability of less than 0.26% to be due to a statistical fluctuation of non-top background, but is in good agreement with what we expect from Standard Model top quark production. The kinematically top-like events also show a large content of beauty quark candidates, in agreement with expectations from top. The probability that the SVX b–tags are due to a statistical fluctuation of non-top backgrounds is less than 1.2×10^{-4}.

The excess of $W+ \geq 4$ jet events at large total transverse energy, H, was found to fit well the top hypothesis. By performing a two-component fit, top component and QCD $W+$ jets component, to the H distribution, we found the top quark mass to be $M_{top} = 180 \pm 12 \; ^{+19}_{-15}$ GeV/c^2. This top quark mass value is only in part correlated with that found by full mass reconstruction of SVX and SLT tagged $W+ \geq 4$ jet events, and is in excellent agreement with that of $M_{top} = 176 \pm 8 \pm 10$ GeV/c^2.

We analyzed the dijet invariant mass in the $W+ \geq 4$ jet events searching for a W decaying to dijet. After applying H cut to enhance signal-to-noise ratio, we obtained a significant mass peak of 27 ± 12 dijets (2.3σ) consistent with a W decaying into dijet in peak position and width.

From these results and counting experiment, we have confirmed previous evidence reported in 1994 April and observed top quark production.

References

[1] F. Abe et al., Phys. Rev. Lett. 73, 225 (1994). F. Abe et al., Phys. Rev. D 50, 2966 (1994).

[2] M.Cobal, PhD Thesis, Feb.1994, Univ. Pisa, INFN-PI/AE 94/004.

[3] F.Abe et al., "*Kinematic Evidence for Top Pair Production at the Tevatron Collider*", FERMILAB-PUB-94/411-E, submitted to Phys. Rev. D.

[4] F. Abe et al., Phys. Rev. Lett. 74, 2626 (1995).

[5] S.Abachi et al., Phys. Rev. Lett. 74, 2632 (1995).

[6] F.A.Berends, W.T.Giele, H.Kuif, B.Tausk, Nucl. Phys. B357, 32 (1991).
W.Giele, PhD.Thesis, Leiden.

[7] G.Marchesini and B.R.Webber, Nucl. Phys. B310, 461(1994).
G.Marchesini et al., Comput. Phys. Comm. 67, 456 (1992).

[8] F.Abe et al., Phys. Rev. D 45, 1448 (1992).
F.Abe et al., Phys. Rev. D 47, 4857 (1993).

[9] F.Abe et al.,“ *Identification of Top Quarks at CDF Using Kinematic Variables*”, FERMILAB-PUB-95/083-E, submitted to Phys. Rev. Lett.