CDF note 7154

Measurement of the Cross Section for $t\bar{t}$ Production Using Event Kinematics in $p\bar{p}$ Collisions at 1.96 TeV

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: July 30, 2004)

We present a measurement of the top pair production cross section in $p\bar{p}$ collisions at 1.96 TeV, with an integrated luminosity 193.5 pb^{-1} at the CDF experiment on the Fermilab Tevatron. We use a neural network technique to discriminate between top pair production and background processes in a sample of 519 events with an isolated, energetic lepton, large missing transverse energy and three or more energetic jets. We measure the top pair production cross section to be $\sigma_{t\bar{t}} = 6.7 \pm 1.1(stat) \pm 1.6(syst)$ pb.

Preliminary Results for Winter 2004 Conferences

I. INTRODUCTION

Since the discovery of the top quark [1], experimental attention has turned to the examination of its properties. Within the context of the Standard Model, in $p\bar{p}$ collisions top quarks are produced in pairs through the strong interaction, via $q\bar{q}$ annihilation (85%) and gluon fusion (15%) at $\sqrt{s} = 1.96$ TeV. Recent theoretical calculations constrain the top pair production cross section with an uncertainty of less than 15% [2, 3]. The top quark is expected to decay to a W boson and b quark nearly 100% of the time. The W boson subsequently decays to either a pair of quarks or a lepton-neutrino pair. Measuring the rate of the reaction $p\bar{p} \to t\bar{t} \to \ell\bar{\nu}_\ell q\bar{q}' b\bar{b}$, the lepton+jets channel, tests both the production and decay mechanisms of the top quark.

This note describes a measurement of the top pair production cross section in the lepton+jets channel at $\sqrt{s} =$ 1.96 TeV. We begin by using a single kinematic variable to discriminate between top pair production and background processes. Then, we develop a neural network technique to maximize the discriminating power from kinematic and topological variables. The sensitivity of the neural network technique is comparable to that for the traditional CDF secondary vertex b-tag method [4], which suppresses the dominant background from W+jets at a cost of a 45% loss in signal efficiency. This kinematic method then allows us to check the assumptions in the b-tag method and test the modeling of signal and background processes with higher statistics. An excellent understanding of top pair production and W+jets background kinematics will be required for the searches for single top production, the Higgs boson and new physics signatures at both the Tevatron and the future LHC.

II. DATA SAMPLE AND EVENT SELECTION

This analysis is based on a sample of integrated luminosity of 193 pb^{-1} collected with the CDF II detector between March 2002 and September 2003. The CDF detector is described in detail in [5]. The data are collected with an inclusive lepton trigger that requires an electron or muon with E_T >18 GeV (P_T >18 GeV/c for the muon). From this inclusive lepton dataset we select offline events with a reconstructed isolated electron E_T (muon P_T) greater than 20 GeV, missing E_T >20 GeV and at least 3 jets with E_T >15 GeV. If the $\not{E}_T$ is below 30 GeV, we require in addition that the angle between the $\not{E}_T$ and the highest E_T jet in the transverse plane, $\Delta\phi$, be greater than 0.5 and less than 2.5 radians.

A. $t\bar{t}$ Acceptance

The total acceptance is measured in a combination of data and Monte Carlo. The geometric times kinematic acceptance of the basic lepton+jets event selection is measured using the PYTHIA Monte Carlo program [6]. The efficiency for identifying the isolated, high P_T lepton is scaled to the value measured in the data using the unbiased leg in Z-boson decays. The geometric times kinematic acceptance, is estimated to be 0.0412 $\pm$ 0.0033% for central electrons, 0.0213 $\pm$ 0.0017% for central muons and 0.0095 $\pm$ 0.0008% for forward muons, where the error includes statistical and systematic effects. Table I summarizes the observed number of data events and the expected number of $t\bar{t}$ events.

Jet multiplicity	$W \to e\nu$	$W \to \mu\nu$	Total	Expected $t\bar{t}$
0 jet	99454	76203	175657	0.2
1 jet	9407	6982	16389	4.4
2 jets	1442	1054	2496	22.6
3 jets	254	147	401	42.3
$\geq$ 4 jets	78	40	118	49.9

TABLE I: The observed number of W candidate events and the expected number of $t\bar{t}$ events, assuming a theoretical cross-section of 6.7 pb at a top mass of 175 GeV/c^2.

B. Backgrounds

The dataset selected above, called "lepton+jets", is dominated by QCD production of W bosons with multiple jets. Much theoretical progress has been made recently to improve the description of the W+jets process, with leading-order

Jet multiplicity	$W \to e\nu$	$W \to \mu\nu$	Total
1 jet	3.8 ± 1.9%	2.9 ± 1.5%	3.4 ± 1.7%
2 jets	6.1 ± 3.0%	2.0 ± 1.0%	4.3 ± 2.1%
≥3 jets	7.7 ± 3.9%	3.1 ± 1.5%	6.3 ± 3.1%

TABLE II: The fraction of multi-jet background in the W+jets data sample as a function of jet multiplicity.

matrix element generators now available to describe the parton hard scattering for processes with a W and up to six additional partons in the final state. We use the ALPGEN [7] matrix element generator, convoluted with the CTEQ5L parton distribution functions [9]. We require parton $p_T \geq 8$ GeV/c, $|\eta| \leq 3.0$ and minimum separation $\Delta R \geq 0.2$ for u, d, s and g partons. We have verified that the shapes of the kinematic distributions used in our kinematic analysis are not sensitive to these values. We choose a default momentum transfer scale of $Q^2 = M_W^2 + \Sigma_i p_{T,i}^2$ for the parton distribution functions and the evaluation of α_s, where $p_{T,i}$ is the transverse momentum of the i-th parton. We use the HERWIG parton shower algorithm to evolve the final state partons to colorless hadrons. For this analysis, we use the $W + n$ parton ALPGEN+HERWIG Monte Carlo to model the $W+ \geq n$ jet final state, where we rely on gluon radiation in the parton shower algorithm to adequately model the higher jet multiplicities.

The other substantial background in this analysis comes from events without W bosons. These events are typically QCD multi-jet events where one jet has faked a high-P$_T$ lepton and mis-measured energies produce apparent $\not\!\!E_T$. We model the kinematics of this background by using those events that pass all of our selection requirements except lepton isolation [10]. We estimate the rate from the number of such events multiplied by the ratio of isolated to non-isolated lepton events with $\not\!\!E_T$ below 10 GeV. The distributions of $\not\!\!E_T$ and the azimuthal angle between the $\not\!\!E_T$ and the most energetic jet, $\Delta\phi$, are shown in Figure 1 for our model of the multi-jet background, the W+jets background and the $t\bar{t}$ signal Monte Carlos. We reduce the systematic uncertainty from the modeling of this background by requiring that the $\not\!\!E_T$ is not parallel or anti-parallel to the direction of the highest E_T jet for events with $\not\!\!E_T$ below 30 GeV. This reduces the multi-jet background by 50% at a cost of only a 5% loss in acceptance for signal $t\bar{t}$. Table II lists the fraction of events from multi-jet processes as a function of the jet multiplicity.

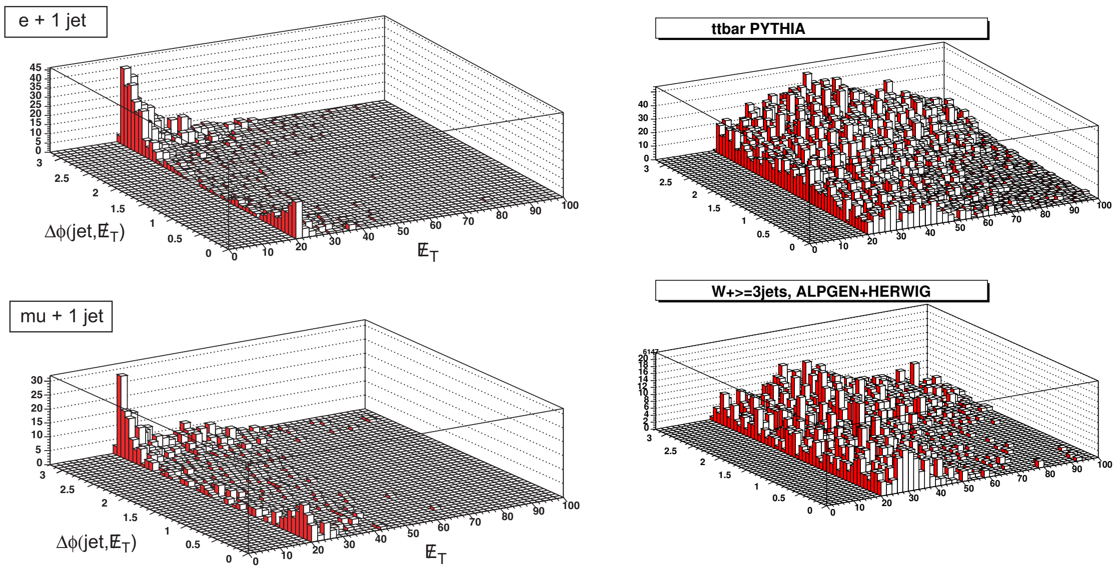

FIG. 1: Observed $\not\!\!E_T$ vs $\Delta\phi$ distribution for non-isolated electrons and muons in events with 1 jet and expected for $t\bar{t}$ PYTHIA and and $W+\geq 3$ jets ALPGEN+HERWIG Monte Carlo.

III. CROSS SECTION MEASUREMENT METHOD

A comparison of the observed data events with the expected number of $t\bar{t}$ signal events in Table I indicates the expected signal to background ratio is about 1:4 and 1:1 in the $W+ \geq 3$ jets and $W+ \geq 4$ jets samples respectively.

At such low signal purities, the sensitivity to top pair production from the observed number of events alone is eradicated by the 30-50% uncertainty on the leading-order theoretical prediction for W+jets background. Other CDF measurements of the top pair production cross section have used b-tagging, with 55% signal efficiency, to improve the signal-to-background ratio to 2:1 and 3:1, in the $W+\geq 3$ jets and $W+\geq 4$ jets respectively, and also use the more accurate prediction for the fraction of W+jets containing heavy flavor.

This analysis instead exploits the discrimination available from kinematic and topological variables to distinguish top pair production from background. Due to the large mass of the top quark, top pair production is associated with central, spherical, energetic events with different kinematics from the predominantly lower energy background processes. We consider separately two background components: W-like, that is the sum of the contributions from W+jets and other electroweak processes, and multi-jet QCD processes. We first choose the best single kinematic variable, and then refine our technique to maximize our discriminating power by creating an optimal variable with an Artificial Neural Network (ANN) technique [11]. ANN's employ information from several variables while accounting for the correlations among them.

We perform a binned likelihood fit to the discriminating variable and find the most likely number of events from $t\bar{t}$ production, $n_{t\bar{t}}$:

$$L(n_{t\bar{t}}, n_w, n_q) = \Pi_{i=1}^{N_{data}} \frac{e^{n_i} n_i^{d^i}}{d_i!}$$

where $n_{t\bar{t}}$, n_w, n_q are the parameters of the fit, representing the number of $t\bar{t}$, W-like and multi-jet events respectively present in the sample. Moreover, the expected number of events in the i-th bin is $n_i = (n_{t\bar{t}} P_{t\bar{t},i} + n_w P_{w,i} + n_q P_{q,i})$, where $P_{t\bar{t},i}$, $P_{w,i}$, $P_{q,i}$ are the probability of observing an event in bin i from $t\bar{t}$, W-like and multi-jet processes. By d_i is denoted the number of observed data events that populate the i-th bin. The level of the multi-jet background, n_q is fixed to that expected from Table II.

We convert the fitted number of $t\bar{t}$ events into the top pair production cross section, $\sigma_{t\bar{t}}$, using the acceptance estimate $\epsilon_{t\bar{t}}$, including the branching ratio for $W \to \ell\nu$, and the luminosity measurement, $\mathcal{L}$, thus:

$$\sigma_{t\bar{t}} = \frac{n_{t\bar{t}}}{\epsilon_{t\bar{t}} \mathcal{L}} \tag{1}$$

A. Neural Network

We considered a set of 20 variables, defined in Table III with good signal-background separation potential. The performance of each single variable is tested *a priori* by constructing simulated experiments using Monte Carlo generated event samples. We choose H_T, shown in Figure 2, as the best single variable.

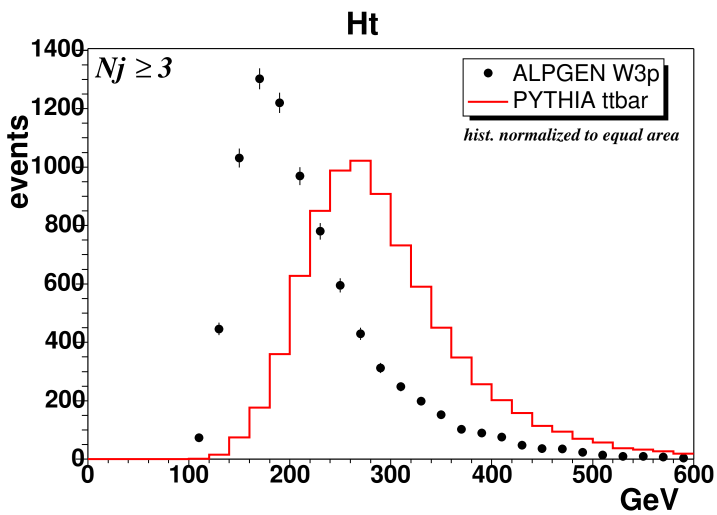

FIG. 2: Expected H_T shape for $t\bar{t}$ PYTHIA and $W+\geq 3$ jets ALPGEN+HERWIG Monte Carlo..

Variable	Definition
H_T	Scalar sum of transverse energies of jets, lepton and $\not\!\!E_T$.
Aplanarity	$3/2Q_1$
$\Sigma p_z/\Sigma E_T$	Ratio of total jet longitudinal momenta to total jet transverse energy.
$\min(M_{jj})$	Minimum di-jet invariant mass
η_{max}	Maximum η of jet.
$\Sigma_{i=3}^{5} E_{T,i}$	Sum E_T of third, fourth and fifth jets.
$\min(\Delta R_{jj})$	Minimum di-jet separation in η and ϕ.
$\Sigma_{i=1}^{n} E_{T,i}$	Sum E_T of all jets.
$\not\!\!E_T$	Missing transverse energy.
Sphericity	$3/2(Q_1+Q_2)$
M_{event}	Invariant mass of jets, lepton and $\not\!\!E_T$.
$\Sigma_{i=1}^{3} M_{jj}$	Sum of di-jet invariant masses.
E_T^{j1}	E_T of jet with highest E_T.
$E_T^{j2}+E_T^{j3}$	Sum of E_T of jets with second and third highest E_T.
M_W^{rec}	Reconstructed hadronic W mass
$\Sigma\eta^2$	Sum of η^2 of jets with highest E_T.
$\Delta\Phi_{lm}$	Azimuthal angle between lepton and $\not\!\!E_T$.
E_T^{j2}	E_T of jet with second highest E_T.
E_T^{j3}	E_T of jet with third highest E_T.
$E_T^{j1}+E_T^{j2}$	Sum of E_T of jets with first and second highest E_T.

TABLE III: Definition of variables considered in this analysis. The momentum tensor of the event is formed from the lepton, $\not\!\!E_T$ and the E_T of the five highest E_T jets. The eigenvalues are ordered such that $Q_1 \leq Q_2 \leq Q_3$.

For the inputs of the ANN, we considered a large number of combinations of variables that can be drawn from this 20-variable set. The ANN is a feed-forward perceptron with one intermediate (hidden) layer and one output node. For training, we use 4000 PYTHIA $t\bar{t}$ and 4000 ALPGEN+HERWIG W+jets Monte Carlo events and require an output of 1.0 for $t\bar{t}$ signal and 0.0 for W+jets background. Other sources of background are not considered during the training process. For each combination, the weights of the network are adjusted to minimize a typical mean squared error function:

$$E = \frac{1}{N}\sum_{i}^{N}(O_i - t_i)^2$$

where O_i is the output of the network for the input event i and t_i is the desired target value. The target values are 1.0 for signal and 0.0 for background events. Learning is an iterative process and we use an independent testing sample of 1000 PYTHIA $t\bar{t}$ and 1000 ALPGEN+HERWIG W+jets Monte Carlo to evaluate the ANN performance and choose when to stop training. There are many algorithms one could use for adjusting the weights in order to produce an optimized network [12]. For this particular problem we obtained satisfactory results by using the default JETNET back-propagation training method with a term added to the error function in order to discourage large weights. After training was completed, an independent validation sample was used to check the quality of the training.

The performance of each neural network is tested *a priori* by again constructing simulated experiments, where we simply treat the output of the ANN as another, better, discriminating variable. Figure 3 shows the expected statistical and systematic error on the signal fraction in different networks as a function of the number of input variables. The number corresponding to 1 input was obtained from fitting the H_T shape. We choose a 7 input ANN as a compromise between simplicity and good performance. The variables of choice are: the total transverse energy, H_T, the event aplanarity, minimum di-jet mass, maximum jet rapidity, minimum di-jet separation, the sum of transverse energy for all jets beyond the third jet, the ratio between jet longitudinal momenta and the jet transverse energy. Finally, in the range of 1-10 nodes in the intermediate layer, there is a slight minimum for 7 hidden nodes. The 7-7-1 configuration of nodes corresponds to 64 free parameters adjusted during training.

IV. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties in this analysis come from Monte Carlo modeling of the geometrical and kinematic acceptance for signal, the luminosity measurement, and from modeling of the kinematic shapes for signal and background.

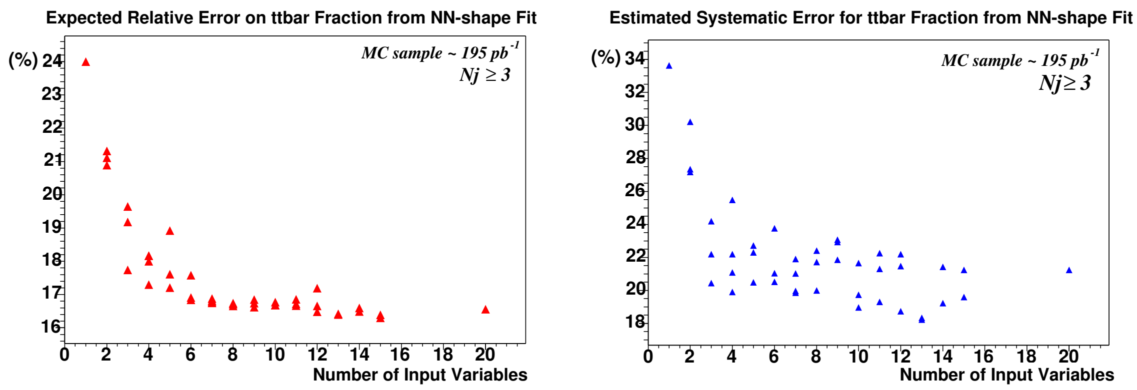

FIG. 3: Expected statistical and systematic errors for different ANNs.

The list of the systematic uncertainties we have considered is summarized in Table IV.

The largest source comes from the uncertainties on the jet energy corrections for different calorimeter response (as a function of η), the absolute hadron energy scale, and fragmentation etc. This of course affects the H_T variable and simultaneously five of the seven kinematic variables used in the ANN analysis. The next largest is from the hard scattering scale used in the Monte Carlo description of the W+jets background. The uncertainty from the multi-jet background shape is estimated by using an alternative model from identified conversions in the data. The initial and final state radiations (ISR and FSR) uncertainties are estimated using the PYTHIA [6] Monte Carlo samples in which QCD parameters for parton shower evolution are varied based on the studies of the CDF Drell-Yan data. For the parton distribution functions, we compare two set of MRSTs for different Λ_{QCD} values, and the difference in CTEQ and MRST PDF groups. The modeling of the lepton identification efficiency in events with multiple jets is an additional source of systematic uncertainty on the acceptance. We use a data to Monte Carlo scale factor that is taken from inclusive Z data and Monte Carlo which is dominated by events with no jets. A 5% systematic uncertainty on this scale factor is estimated by convoluting the scale factor measured as a function of ΔR between the lepton and the nearest jet with the ΔR distribution of leptons in $t\bar{t}$ events with 3 or more jets. The luminosity of 193 pb^{-1} has an uncertainty of 5.9%, of which 4.4% comes from the acceptance and operation of the luminosity monitor and 4.0% from the calculation of the total $p\bar{p}$ cross section [13].

Source	H_T(%)	ANN (%)
W+jets Background	14.0	8.3
QCD Background	16.0	1.8
Jet Energy Scale	30.0	16.4
$t\bar{t}$ generator	0.6	0.3
ISR/FSR	1.3	1.2
PDF	8.6	8.1
Lepton ID	5.4	5.4
Luminosity	5.9	5.9
Total	38.6	21.8

TABLE IV: Systematic uncertainty for H_T and ANN.

V. RESULTS

For $t\bar{t}$ events in 3 or more jets, we measure a cross section with the artificial Neural Network technique of:

$$\sigma_{t\bar{t}} = 6.7 \pm 1.1 \pm 1.6 \text{ pb}$$

where the first uncertainty is statistical and the second is systematic. Restricting the analysis to the 4 or more jet bin gives a cross section of $7.5 \pm 1.6 \pm 1.8$ pb from ANN. For the single variable analysis using H_T, we find a cross section of $4.7 \pm 1.6(stat) \pm 1.8(syst)$ pb with 3 or more jets and $8.0 \pm 2.0(stat) \pm 3.0(syst)$ pb with 4 or more jets. These results are in good agreement with the theoretical prediction of 6.7 pb for a top mass of 175 GeV/c^2.

Figure 5 shows the ANN output for those events with at least one identified Secondary Vertex b-tag in the $W+ \geq 3$ jets data sample, where the signal to background ratio is expected to be 2:1. This clearly demonstrates that the ANN is selecting $t\bar{t}$ events.

Method	Data Events	$t\bar{t}$ fraction	$t\bar{t}$ cross-section (pb)
$W+ \geq 3$ jets	519		
ANN		0.176 $\pm$ 0.030	6.7 $\pm$ 1.1 $\pm$ 1.6
H_T		0.126 $\pm$ 0.042	4.7 $\pm$ 1.6 $\pm$ 1.8
$W+ \geq 4$ jets	118		
ANN		0.473 $\pm$ 0.100	7.5 $\pm$ 1.6 $\pm$ 2.0
H_T		0.503 $\pm$ 0.130	8.0 $\pm$ 2.0 $\pm$ 3.0

TABLE V: Number of observed data events, measured $t\bar{t}$ fraction with statistical error only and measured $t\bar{t}$ cross section with statistical and systematic error.

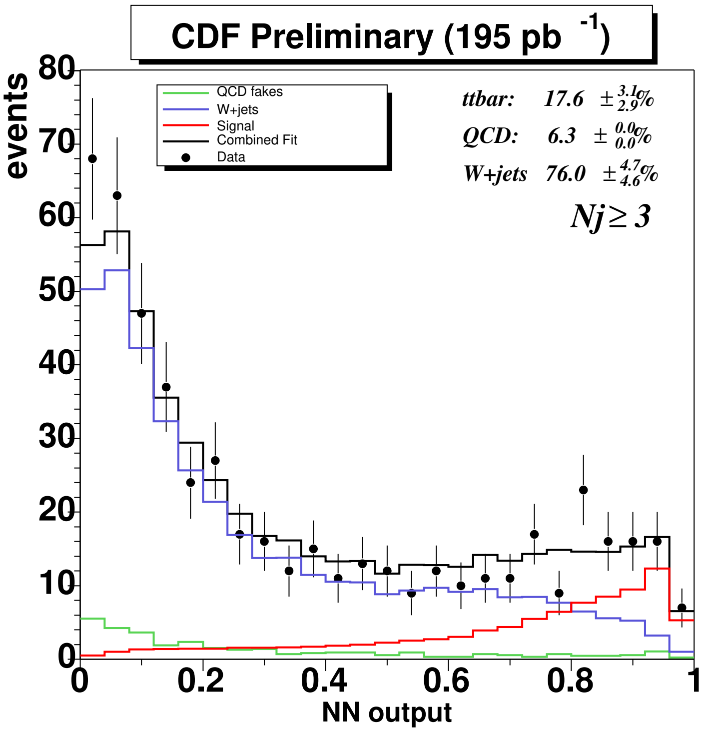

FIG. 4: Observed ANN output distribution versus fit result for $W+ \geq 3$ jet events.

Acknowledgments

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, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic

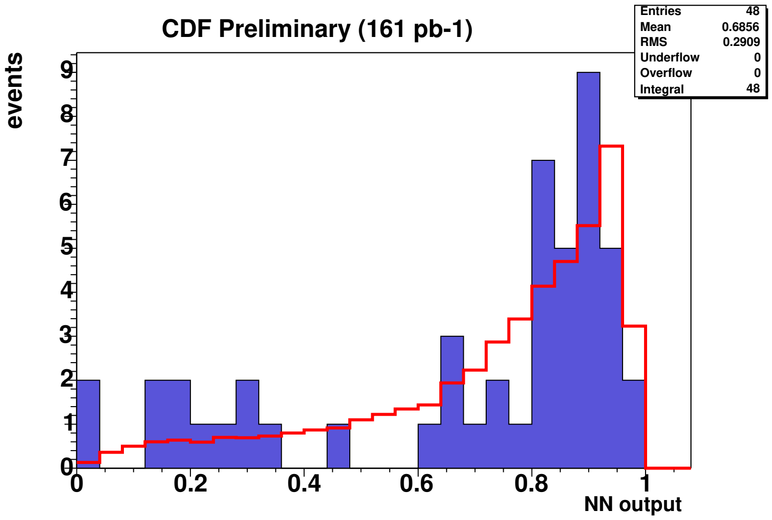

FIG. 5: Observed ANN output distribution versus for $W+\geq 3$ jet events with at least 1 secondary vertex b-tag. The solid red line is the expectation from $t\bar{t}$ PYTHIA MC.

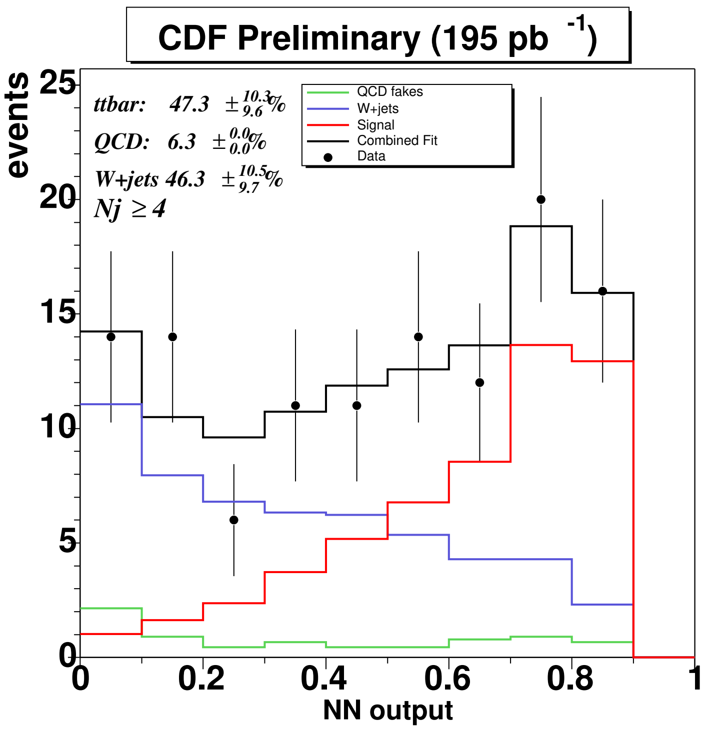

FIG. 6: Observed ANN output distribution versus fit result for $W+\geq 4$ jet events.

Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

[1] CDF Collaboration, F.Abe *et al.*, Phys. Rev. Lett. **74** 2626 (1995); D0 Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).

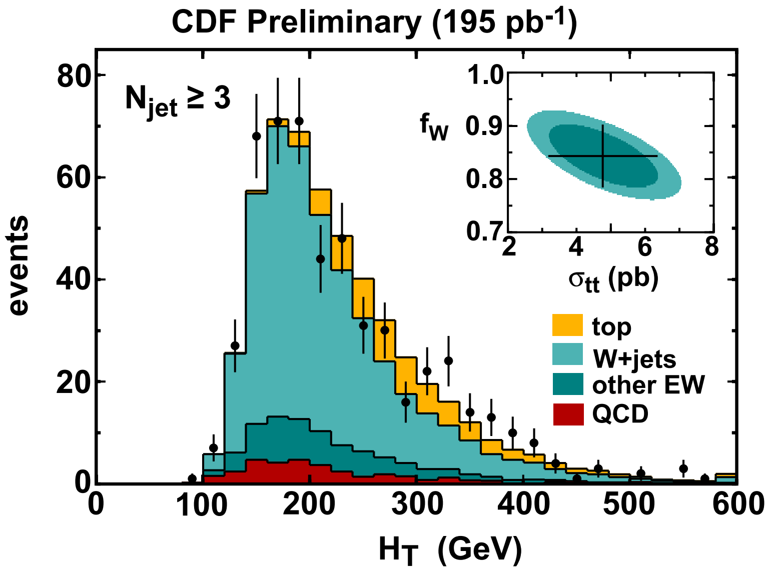

FIG. 7: Observed H_T distribution versus fit result for $W+\geq 3$ jet events.

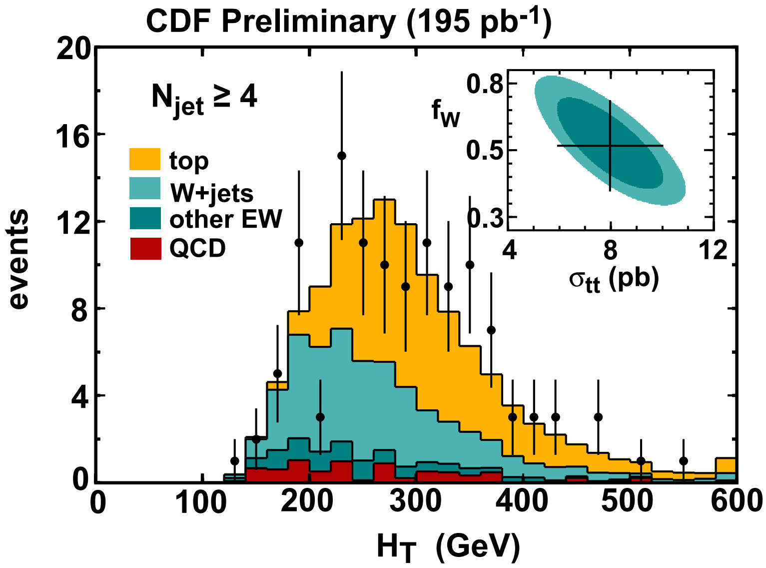

FIG. 8: Observed H_T distribution versus fit result for $W+\geq 4$ jet events.

[2] R. Bonciani, S. Catani, M.L. Mangano and P. Nason, Nucl. Phys. B 529 424 (1998), updated in arXiv:hep-ph/0303085.
[3] N. Kidonakis and R. Vogt, Phys. Rev. D 68 114014 (2003).
[4] CDF collaboration, "Measurement of the $t\bar{t}$ productin cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using Lepton+Jets Events with Secondary Vertex b-tagging", Winter 2004 public conference note.
[5] F. Abe, *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, *et al.*, Nucl. Instum. 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
[6] T. Sjostrand *et al.*, High-Energy-Physics Event Generation with PYTHIA 6.1, Comput. Phys. Commun. **135**, 238 (2001).
[7] M.L. Mangano *et al.*, JHEP **07**, 1 (2003).
[8] G. Corcella *et al.*, HERWIG 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons (including supersymmetric processes), JHEP **01**, 10 (2001).
[9] H.L. Lai *et al.*, Eur. Phys. J. **C12**, 375 (200).
[10] Lepton isolation is defined as the excess transverse energy in the calorimeter within a cone of radius 0.4 centred on the

lepton direction, divided by the lepton transverse energy. The selection requirement is less than 10%.
[11] Brian D. Ripley, *Pattern recognition and Neural Networks*, Cambridge University Press (1996).
[12] Cartsen Peterson, Thorsteinn Röngnvaldsson, Leif Lönnblad, *JETNET 3.0 - A Versatile Artificial Neural Network Package*, CERN-TH.7135/95.
[13] S. Klimenko *et al.*, FERMILAB-FN-0741 (2003); D. Acosta *et al.*, Nucl. Instrum. Meth. **A494**, 57 (2002).