Optimized search for single-top-quark production at the Tevatron

D. Acosta,[14] T. Affolder,[7] M.G. Albrow,[13] D. Ambrose,[36] D. Amidei,[27] K. Anikeev,[26] J. Antos,[1] G. Apollinari,[13] T. Arisawa,[50] A. Artikov,[11] W. Ashmanskas,[2] F. Azfar,[34] P. Azzi-Bacchetta,[35] N. Bacchetta,[35] H. Bachacou,[24] W. Badgett,[13] A. Barbaro-Galtieri,[24] V.E. Barnes,[39] B.A. Barnett,[21] S. Baroiant,[5] M. Barone,[15] G. Bauer,[26] F. Bedeschi,[37] S. Behari,[21] S. Belforte,[47] W.H. Bell,[17] G. Bellettini,[37] J. Bellinger,[51] D. Benjamin,[12] A. Beretvas,[13] A. Bhatti,[41] M. Binkley,[13] D. Bisello,[35] M. Bishai,[13] R.E. Blair,[2] C. Blocker,[4] K. Bloom,[27] B. Blumenfeld,[21] A. Bocci,[41] A. Bodek,[40] G. Bolla,[39] A. Bolshov,[26] D. Bortoletto,[39] J. Boudreau,[38] C. Bromberg,[28] E. Brubaker,[24] J. Budagov,[11] H.S. Budd,[40] K. Burkett,[13] G. Busetto,[35] K.L. Byrum,[2] S. Cabrera,[12] M. Campbell,[27] W. Carithers,[24] D. Carlsmith,[51] A. Castro,[3] D. Cauz,[47] A. Cerri,[24] L. Cerrito,[20] J. Chapman,[27] C. Chen,[36] Y.C. Chen,[1] M. Chertok,[5] G. Chiarelli,[37] G. Chlachidze,[13] F. Chlebana,[13] M.L. Chu,[1] J.Y. Chung,[32] W.-H. Chung,[51] Y.S. Chung,[40] C.I. Ciobanu,[20] A.G. Clark,[16] M. Coca,[40] A. Connolly,[24] M. Convery,[41] J. Conway,[43] M. Cordelli,[15] J. Cranshaw,[45] R. Culbertson,[13] D. Dagenhart,[4] S. D'Auria,[17] P. de Barbaro,[40] S. De Cecco,[42] S. Dell'Agnello,[15] M. Dell'Orso,[37] S. Demers,[40] L. Demortier,[41] M. Deninno,[3] D. De Pedis,[42] P.F. Derwent,[13] C. Dionisi,[42] J.R. Dittmann,[13] A. Dominguez,[24] S. Donati,[37] M. D'Onofrio,[16] T. Dorigo,[35] N. Eddy,[20] R. Erbacher,[13] D. Errede,[20] S. Errede,[20] R. Eusebi,[40] S. Farrington,[17] R.G. Feild,[52] J.P. Fernandez,[39] C. Ferretti,[27] R.D. Field,[14] I. Fiori,[37] B. Flaugher,[13] L.R. Flores-Castillo,[38] G.W. Foster,[13] M. Franklin,[18] J. Friedman,[26] I. Furic,[26] M. Gallinaro,[41] M. Garcia-Sciveres,[24] A.F. Garfinkel,[39] C. Gay,[52] D.W. Gerdes,[27] E. Gerstein,[9] S. Giagu,[42] P. Giannetti,[37] K. Giolo,[39] M. Giordani,[47] P. Giromini,[15] V. Glagolev,[11] D. Glenzinski,[13] M. Gold,[30] N. Goldschmidt,[27] J. Goldstein,[34] G. Gomez,[8] M. Goncharov,[44] I. Gorelov,[30] A.T. Goshaw,[12] Y. Gotra,[38] K. Goulianos,[41] A. Gresele,[3] C. Grosso-Pilcher,[10] M. Guenther,[39] J. Guimaraes da Costa,[18] C. Haber,[24] S.R. Hahn,[13] E. Halkiadakis,[40] R. Handler,[51] F. Happacher,[15] K. Hara,[48] R.M. Harris,[13] F. Hartmann,[22] K. Hatakeyama,[41] J. Hauser,[6] J. Heinrich,[36] M. Hennecke,[22] M. Herndon,[21] C. Hill,[7] A. Hocker,[40] K.D. Hoffman,[10] S. Hou,[1] B.T. Huffman,[34] R. Hughes,[32] J. Huston,[28] C. Issever,[7] J. Incandela,[7] G. Introzzi,[37] M. Iori,[42] A. Ivanov,[40] Y. Iwata,[19] B. Iyutin,[26] E. James,[13] M. Jones,[39] T. Kamon,[44] J. Kang,[27] M. Karagoz Unel,[31] S. Kartal,[13] H. Kasha,[52] Y. Kato,[33] R.D. Kennedy,[13] R. Kephart,[13] B. Kilminster,[40] D.H. Kim,[23] H.S. Kim,[20] M.J. Kim,[9] S.B. Kim,[23] S.H. Kim,[48] T.H. Kim,[26] Y.K. Kim,[10] M. Kirby,[12] L. Kirsch,[4] S. Klimenko,[14] P. Koehn,[32] K. Kondo,[50] J. Konigsberg,[14] A. Korn,[26] A. Korytov,[14] J. Kroll,[36] M. Kruse,[12] V. Krutelyov,[44] S.E. Kuhlmann,[2] N. Kuznetsova,[13] A.T. Laasanen,[39] S. Lami,[41] S. Lammel,[13] J. Lancaster,[12] K. Lannon,[32] M. Lancaster,[25] R. Lander,[5] A. Lath,[43] G. Latino,[30] T. LeCompte,[2] Y. Le,[21] J. Lee,[40] S.W. Lee,[44] N. Leonardo,[26] S. Leone,[37] J.D. Lewis,[13] K. Li,[52] C.S. Lin,[13] M. Lindgren,[6] T.M. Liss,[20] T. Liu,[13] D.O. Litvintsev,[13] N.S. Lockyer,[36] A. Loginov,[29] M. Loreti,[35] D. Lucchesi,[35] P. Lukens,[13] L. Lyons,[34] J. Lys,[24] R. Madrak,[18] K. Maeshima,[13] P. Maksimovic,[21] L. Malferrari,[3] M. Mangano,[37] G. Manca,[34] M. Mariotti,[35] M. Martin,[21] A. Martin,[52] V. Martin,[31] M. Martínez,[13] P. Mazzanti,[3] K.S. McFarland,[40] P. McIntyre,[44] M. Menguzzato,[35] A. Menzione,[37] P. Merkel,[13] C. Mesropian,[41] A. Meyer,[13] T. Miao,[13] R. Miller,[28] J.S. Miller,[27] S. Miscetti,[15] G. Mitselmakher,[14] N. Moggi,[3] R. Moore,[13] T. Moulik,[39] M. Mulhearn,[26] A. Mukherjee,[13] T. Muller,[22] A. Munar,[36] P. Murat,[13] J. Nachtman,[13] S. Nahn,[52] I. Nakano,[19] R. Napora,[21] F. Niell,[27] C. Nelson,[13] T. Nelson,[13] C. Neu,[32] M.S. Neubauer,[26] C. Newman-Holmes,[13] T. Nigmanov,[38] L. Nodulman,[2] S.H. Oh,[12] Y.D. Oh,[23] T. Ohsugi,[19] T. Okusawa,[33] W. Orejudos,[24] C. Pagliarone,[37] F. Palmonari,[37] R. Paoletti,[37] V. Papadimitriou,[45] J. Patrick,[13] G. Pauletta,[47] M. Paulini,[9] T. Pauly,[34] C. Paus,[26] D. Pellett,[5] A. Penzo,[47] T.J. Phillips,[12] G. Piacentino,[37] J. Piedra,[8] K.T. Pitts,[20] A. Pompoš,[39] L. Pondrom,[51] G. Pope,[38] T. Pratt,[34] F. Prokoshin,[11] J. Proudfoot,[2] F. Ptohos,[15] O. Poukhov,[11] G. Punzi,[37] J. Rademacker,[34] A. Rakitine,[26] F. Ratnikov,[43] H. Ray,[27] A. Reichold,[34] P. Renton,[34] M. Rescigno,[42] F. Rimondi,[3] L. Ristori,[37] W.J. Robertson,[12] T. Rodrigo,[8] S. Rolli,[49] L. Rosenson,[26] R. Roser,[13] R. Rossin,[35] C. Rott,[39] A. Roy,[39] A. Ruiz,[8] D. Ryan,[49] A. Safonov,[5] R. St. Denis,[17] W.K. Sakumoto,[40] D. Saltzberg,[6] C. Sanchez,[32] A. Sansoni,[15] L. Santi,[47] S. Sarkar,[42] P. Savard,[46] A. Savoy-Navarro,[13] P. Schlabach,[13] E.E. Schmidt,[13] M.P. Schmidt,[52] M. Schmitt,[31] L. Scodellaro,[35] A. Scribano,[37] A. Sedov,[39] S. Seidel,[30] Y. Seiya,[48] A. Semenov,[11] F. Semeria,[3] M.D. Shapiro,[24] P.F. Shepard,[38] T. Shibayama,[48] M. Shimojima,[48] M. Shochet,[10] A. Sidoti,[35] A. Sill,[45] P. Sinervo,[46] A.J. Slaughter,[52] K. Sliwa,[49] F.D. Snider,[13] R. Snihur,[25] M. Spezziga,[45] F. Spinella,[37] M. Spiropulu,[7] L. Spiegel,[13] A. Stefanini,[37] J. Strologas,[30] D. Stuart,[7] A. Sukhanov,[14] K. Sumorok,[26] T. Suzuki,[48] R. Takashima,[19] K. Takikawa,[48] M. Tanaka,[2] M. Tecchio,[27] R.J. Tesarek,[13] P.K. Teng,[1] K. Terashi,[41] S. Tether,[26] J. Thom,[13] A.S. Thompson,[17] E. Thomson,[32] P. Tipton,[40] S. Tkaczyk,[13] D. Toback,[44] K. Tollefson,[28] D. Tonelli,[37] M. Tönnesmann,[28] H. Toyoda,[33] W. Trischuk,[46] J. Tseng,[26] D. Tsybychev,[14] N. Turini,[37] F. Ukegawa,[48] T. Unverhau,[17] T. Vaiciulis,[40] A. Varganov,[27] E. Vataga,[37] S. Vejcik III,[13] G. Velev,[13] G. Veramendi,[24] R. Vidal,[13] I. Vila,[8] R. Vilar,[8] I. Volobouev,[24] M. von der Mey,[6] R.G. Wagner,[2] R.L. Wagner,[13] W. Wagner,[22] Z. Wan,[43] C. Wang,[12] M.J. Wang,[1] S.M. Wang,[14] B. Ward,[17] S. Waschke,[17] D. Waters,[25] T. Watts,[43] M. Weber,[24] W.C. Wester III,[13] B. Whitehouse,[49] A.B. Wicklund,[2] E. Wicklund,[13] H.H. Williams,[36] P. Wilson,[13] B.L. Winer,[32] S. Wolbers,[13] M. Wolter,[49] S. Worm,[43] X. Wu,[16] F. Würthwein,[26] U.K. Yang,[10] W. Yao,[24] G.P. Yeh,[13] K. Yi,[21]

J. Yoh,[13] T. Yoshida,[33] I. Yu,[23] S. Yu,[36] J.C. Yun,[13] L. Zanello,[42] A. Zanetti,[47] F. Zetti,[24] and S. Zucchelli[3]

(CDF Collaboration)

[1] *Institute of Physics, Academia Sinica, Taipei, Taiwan 11529, Republic of China*
[2] *Argonne National Laboratory, Argonne, Illinois 60439*
[3] *Istituto Nazionale di Fisica Nucleare, University of Bologna, I-40127 Bologna, Italy*
[4] *Brandeis University, Waltham, Massachusetts 02254*
[5] *University of California at Davis, Davis, California 95616*
[6] *University of California at Los Angeles, Los Angeles, California 90024*
[7] *University of California at Santa Barbara, Santa Barbara, California 93106*
[8] *Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain*
[9] *Carnegie Mellon University, Pittsburgh, Pennsylvania 15213*
[10] *Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637*
[11] *Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*
[12] *Duke University, Durham, North Carolina 27708*
[13] *Fermi National Accelerator Laboratory, Batavia, Illinois 60510*
[14] *University of Florida, Gainesville, Florida 32611*
[15] *Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*
[16] *University of Geneva, CH-1211 Geneva 4, Switzerland*
[17] *Glasgow University, Glasgow G12 8QQ, United Kingdom*
[18] *Harvard University, Cambridge, Massachusetts 02138*
[19] *Hiroshima University, Higashi-Hiroshima 724, Japan*
[20] *University of Illinois, Urbana, Illinois 61801*
[21] *The Johns Hopkins University, Baltimore, Maryland 21218*
[22] *Institut für Experimentelle Kernphysik, Universität Karlsruhe, 76128 Karlsruhe, Germany*
[23] *Center for High Energy Physics: Kyungpook National University, Taegu 702-701; Seoul National University, Seoul 151-742; and SungKyunKwan University, Suwon 440-746; Korea*
[24] *Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720*
[25] *University College London, London WC1E 6BT, United Kingdom*
[26] *Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*
[27] *University of Michigan, Ann Arbor, Michigan 48109*
[28] *Michigan State University, East Lansing, Michigan 48824*
[29] *Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia*
[30] *University of New Mexico, Albuquerque, New Mexico 87131*
[31] *Northwestern University, Evanston, Illinois 60208*
[32] *The Ohio State University, Columbus, Ohio 43210*
[33] *Osaka City University, Osaka 588, Japan*
[34] *University of Oxford, Oxford OX1 3RH, United Kingdom*
[35] *Universita di Padova, Istituto Nazionale di Fisica Nucleare, Sezione di Padova, I-35131 Padova, Italy*
[36] *University of Pennsylvania, Philadelphia, Pennsylvania 19104*
[37] *Istituto Nazionale di Fisica Nucleare, University and Scuola Normale Superiore of Pisa, I-56100 Pisa, Italy*
[38] *University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
[39] *Purdue University, West Lafayette, Indiana 47907*
[40] *University of Rochester, Rochester, New York 14627*
[41] *Rockefeller University, New York, New York 10021*
[42] *Instituto Nazionale de Fisica Nucleare, Sezione di Roma, University di Roma I, "La Sapienza," I-00185 Roma, Italy*
[43] *Rutgers University, Piscataway, New Jersey 08855*
[44] *Texas A&M University, College Station, Texas 77843*
[45] *Texas Tech University, Lubbock, Texas 79409*
[46] *Institute of Particle Physics, University of Toronto, Toronto M5S 1A7, Canada*
[47] *Istituto Nazionale di Fisica Nucleare, University of Trieste/ Udine, Italy*
[48] *University of Tsukuba, Tsukuba, Ibaraki 305, Japan*
[49] *Tufts University, Medford, Massachusetts 02155*
[50] *Waseda University, Tokyo 169, Japan*
[51] *University of Wisconsin, Madison, Wisconsin 53706*
[52] *Yale University, New Haven, Connecticut 06520*

We use a neural-network technique to search for standard model (SM) single-top-quark production in the 106 pb^{-1} dataset accumulated by the CDF detector during 1992 – 1995 collider run ("Run I"). Using a sample of 64 $W+1$, 2, 3 jets events, we set a 95% C.L. upper limit of 24 pb on W-gluon and W^* combined single-top cross section.

PACS numbers: 14.65.Ha, 12.15.Ji, 13.85.Rm, 87.18.Sn

At the Tevatron, top quarks produced in pairs through the strong interaction were observed [1, 2]. Within the standard model, top quarks are also expected to be produced singly in the electroweak channel [3], mainly through off mass shell W production ("W^*") and W-gluon fusion ("Wg") processes, shown in Fig. 1. The measurement of single-top events is of particular interest because the production cross section is proportional to $|V_{tb}|^2$, where V_{tb} represents the Cabibbo-Kobayashi-Maskawa matrix element relating top and bottom quarks. Assuming $|V_{tb}| = 1$, the next-to-leading order predicted cross sections at $\sqrt{s} = 1.8$ TeV for W^* and Wg channels are 0.76 pb and 1.40 pb, respectively [4]. The DØ collaboration has published upper limits on single-top production of 22 pb on Wg and 17 pb on W^*, both at 95% confidence level (C.L.) [5]. The CDF collaboration reported lower 95% C.L. limits: 13 pb and 18 pb on for the Wg and W^* cross sections respectively, and 14 pb for the combined cross section as determined in a separate analysis [6]. In this paper we report on a search for the combined W^* and Wg single-top production using a neural-network technique to maximize the discriminating power of seven kinematic variables. This technique is expected to be more sensitive than the method employed in [6]. Besides using a larger amount of information, the analysis also features marginally higher signal purity obtained by retuning the event selection. The improvement in the average expected upper limit on the single-top cross section is 20% if the SM signal cross section is assumed.

The final state of the W^* channel features two b-quarks and the decay products of the W boson. Similarly, the Wg channel is characterized by two b-quarks and the W decay products plus an additional light quark jet (u, d). In addition, initial and final state radiation can increase the jet content of the final state. Our analysis will focus on the channels with leptonic W decays $W \to e\nu_e$, $\mu\nu_\mu$.

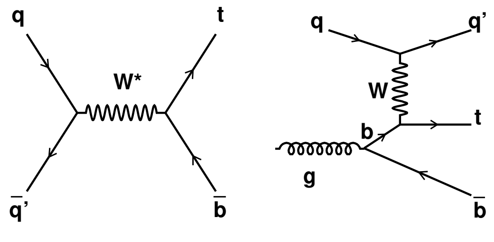

FIG. 1: Representative Feynman diagrams for single-top-quark production at the Tevatron: s-channel W^* (left) and t-channel W-gluon fusion (right).

TABLE I: N_{js} cut efficiencies for signal and background. $\epsilon_{N_{js}}(W{+}1\text{ jet})$ represents the fraction of W+1 jet events with $N_{js} = 1$, after the initial selections were imposed. Similarly, $\epsilon_{N_{js}}$ is the fraction of W+1, 2, and 3 jets events passing the N_{js} selections. The overall ϵ_{tot} results from multiplying the efficiencies of the initial and the N_{js} selections.

Efficiency	W^*	Wg	non-top	$t\bar{t}$
$\epsilon_{N_{js}}(W{+}1\text{ jet})$	43.4%	39.7%	23.9%	42.7%
$\epsilon_{N_{js}}(W{+}3\text{ jets})$	72.9%	75.2%	73.5%	42.8%
Combined $\epsilon_{N_{js}}$	83.6%	74.1%	47.6%	59.7%
Overall ϵ_{tot}	2.4%	1.6%	0.02%	1.9%

These yield a sample of "lepton+jets" events that we can study using many of the tools developed for the CDF top pair production ($t\bar{t}$) cross section analysis [7].

This analysis uses the data from $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV collected with the Collider Detector at Fermilab between 1992 and 1995. A thorough description of the detector is provided elsewhere [8]. We select the events having an isolated electron (muon) with transverse energy $E_T > 20$ GeV (transverse momentum $p_T > 20$ GeV/c), and missing transverse energy $\not{E}_T > 20$ GeV [9]. The $t\bar{t}$ or Z boson decays are removed by rejecting events containing an additional isolated track with $p_T > 15$ GeV/c and charge opposite to that of the primary lepton [10]. Also rejected are Z candidates in which there are two opposite-charge leptons with invariant mass between 75 and 105 GeV/c^2. We further require that there are one, two, or three jets with $E_T > 15$ GeV and pseudorapidity $|\eta| < 2.0$ ("tight" jets) in the event. At least one of these jets should be associated with a b-quark decay ("B-tagged") as determined by observing a displaced vertex using tracks reconstructed in the silicon vertex detector (SVX) [11]. After these initial selections, the backgrounds can be classified as non-top (mostly QCD multijet) and $t\bar{t}$ production.

We further reduce backgrounds by exploiting the distributions of "soft" jets in the event. These are jets with $E_T > 8$ GeV and $|\eta| < 2.4$ which do not pass the above tight jet criteria. Tight and soft jet multiplicities are denoted by N_{jt} and N_{js}. We use N_{jt} to define and label the jet multiplicity bins $W + N_{jt}$ jets. For example, a W + 3 jets event contains exactly three tight jets and possibly additional soft jets. Fig. 2 shows the N_{jt} versus N_{js} Monte Carlo distributions for W^*, Wg, non-top, and $t\bar{t}$ processes. The PYTHIA Monte Carlo program [12] was used, followed by the CDF detector simulation. Optimal signal to background ratio is obtained by demanding $N_{js} = 1$ in the W+1 jet events, and $N_{js} = 0$ in the

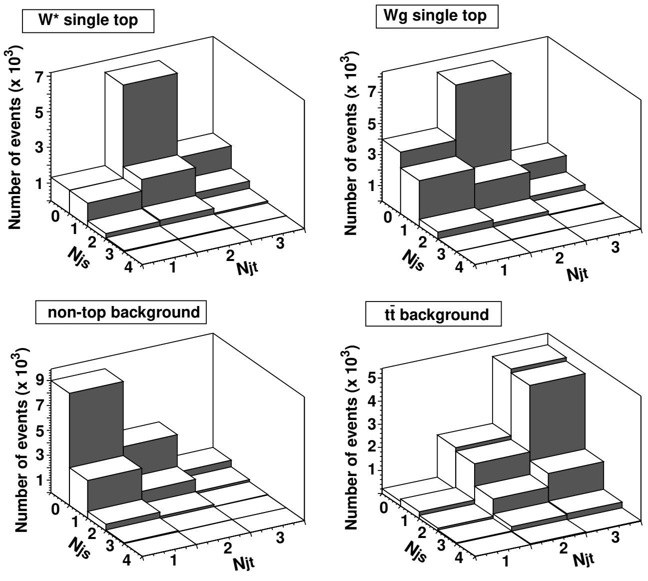

FIG. 2: N_{js} versus N_{jt} distribution for simulated signal and background events passing the initial selection described in the text. Non-top backgrounds are suppressed by requiring $N_{js} = 1$ for W+1 jet events ($N_{jt} = 1$). We reduce the $t\bar{t}$ background by requiring $N_{js} = 0$ for W+3 jets events ($N_{jt} = 3$).

W+3 jets events. There is no N_{js} requirement for the W+2 jets events. As shown in Table I, the soft jets requirements remove over 50% of the non-top and 40% of the $t\bar{t}$ events passing initial selections. If we assume the theoretical W^* and Wg cross sections [4] we arrive at the signal contributions listed in Table II. The expected numbers of $t\bar{t}$ and non-top events are also given in Table II. The $t\bar{t}$ expectation is obtained using a PYTHIA Monte Carlo calculation normalized to the theory prediction $\sigma_{t\bar{t}} = 5.1 \pm 0.9$ pb [13]. For the non-top background, the primary source (approx. 65%) is the W + heavy flavor production process $\bar{q}q' \to Wg$ with $g \to b\bar{b}$, $c\bar{c}$, and $gq \to Wq'$ [11]. Other sources include "mistags" (17%), where a light-flavor jet is misidentified as heavy flavor jet, direct $b\bar{b}$ production (11%), Z+heavy flavor and $Z \to \tau\bar{\tau}$ (5%), and also diboson processes WW, WZ (2%). The non-top expectations are based on the calculation performed in the previous CDF single-top analysis [6] which we correct for differences in the selection criteria. To estimate the shape of the non-top background kinematic distributions we use a PYTHIA generated sample of W+heavy flavor events.

The estimated signal and background contributions outlined above can be combined to predict a signal to noise ratio of 1/13, which implies a challenging search. We maximize our discriminating power by employing an Artificial Neural Network (ANN) technique [14]. ANN's employ information from several kinematic variables while accounting for the correlations among them. The

TABLE II: Signal and background contributions expected and total number of events observed in Run I after all selection cuts described in the text have been imposed. Wg and W^* uncertainties are associated with the detector and do not include theoretical uncertainties given in Ref. [4].

Process	W+1 jet	W+2 jets	W+3 jets
Wg	0.5 ± 0.2	1.5 ± 0.4	0.2 ± 0.1
W^*	0.2 ± 0.1	1.2 ± 0.3	0.2 ± 0.1
$t\bar{t}$	0.2 ± 0.1	3.7 ± 1.1	3.6 ± 1.1
non-top	15.6 ± 3.1	24.0 ± 4.5	3.8 ± 0.8
Total	16.5 ± 3.1	30.4 ± 4.7	7.8 ± 1.4
Observed	14	41	9

goal is to design an ANN to classify events in one of three categories: single-top (W^* and Wg), $t\bar{t}$, and non-top. We do not attempt to distinguish between W^* and Wg signal events, as most of the kinematic distributions considered in this analysis are very similar for the two processes (see Fig. 3). The differences between the two signal channels are accommodated by training and testing the network with W^* and Wg events in the proportion expected from SM (Table II). We will subsequently demonstrate that our method is rather insensitive to the precise W^*−Wg mixing proportion within a range of ±50% of its SM value.

The network is a feed-forward perceptron with one intermediate (hidden) layer and three output nodes. The advantages of using one output node for each class of events are detailed in Ref. [15]. For training we use 30000 Monte Carlo events, and require an output of (0,1,0) for signal, (0,0,1) for $t\bar{t}$, and (1,0,0) for non-top background. The weights are updated according to the "Manhattan" algorithm in JETNET [16] with default parameters.

To select the inputs of the ANN, we started from a set of 18 variables with good signal-background separation potential [6, 17, 18]: E_T^{j1}, E_T^{j2}, E_T^{ℓ}, $\not\!\!E_T$, H_T, $\sqrt{\hat{s}}$, $M^{\ell\nu b}$, M^{jj}, P_T^{jj}, η^{jj}, η^{j1}, η^{j2}, $Q \times \eta$, $\cos(\widehat{\ell q})$, R_{min}, N_{jt}, N_{js}, N_{B-tags}. Here $j1$ and $j2$ are the leading jets in the event, H_T is the total transverse energy defined as $E_T^{\ell} + \not\!\!E_T + \sum E_T^j$ where the last term includes both the tight and the soft jets, $\sqrt{\hat{s}}$ is the total energy in the center-of-mass system, $\ell\nu b$ refers to the lepton, neutrino, and leading B-tagged jet system, jj refers to the $j1-j2$ system, $Q \times \eta$ is the product between the primary lepton charge and the pseudorapidity of the highest-E_T untagged jet (q), $\widehat{\ell q}$ is the angle between the direction of the lepton and that of the q jet, and R_{min} is the minimum separation $\sqrt{(\delta\eta)^2 + (\delta\phi)^2}$ among all possible pairs of jets in the event. We considered a large number of combinations of variables that can be drawn from this 18-variable set. For each combination we minimized a typical mean squared error function [15]:

$$E = \frac{1}{N} \cdot \sum_{k=1}^{N} |\overrightarrow{O^k} - \overrightarrow{T^k}|^2 \qquad (1)$$

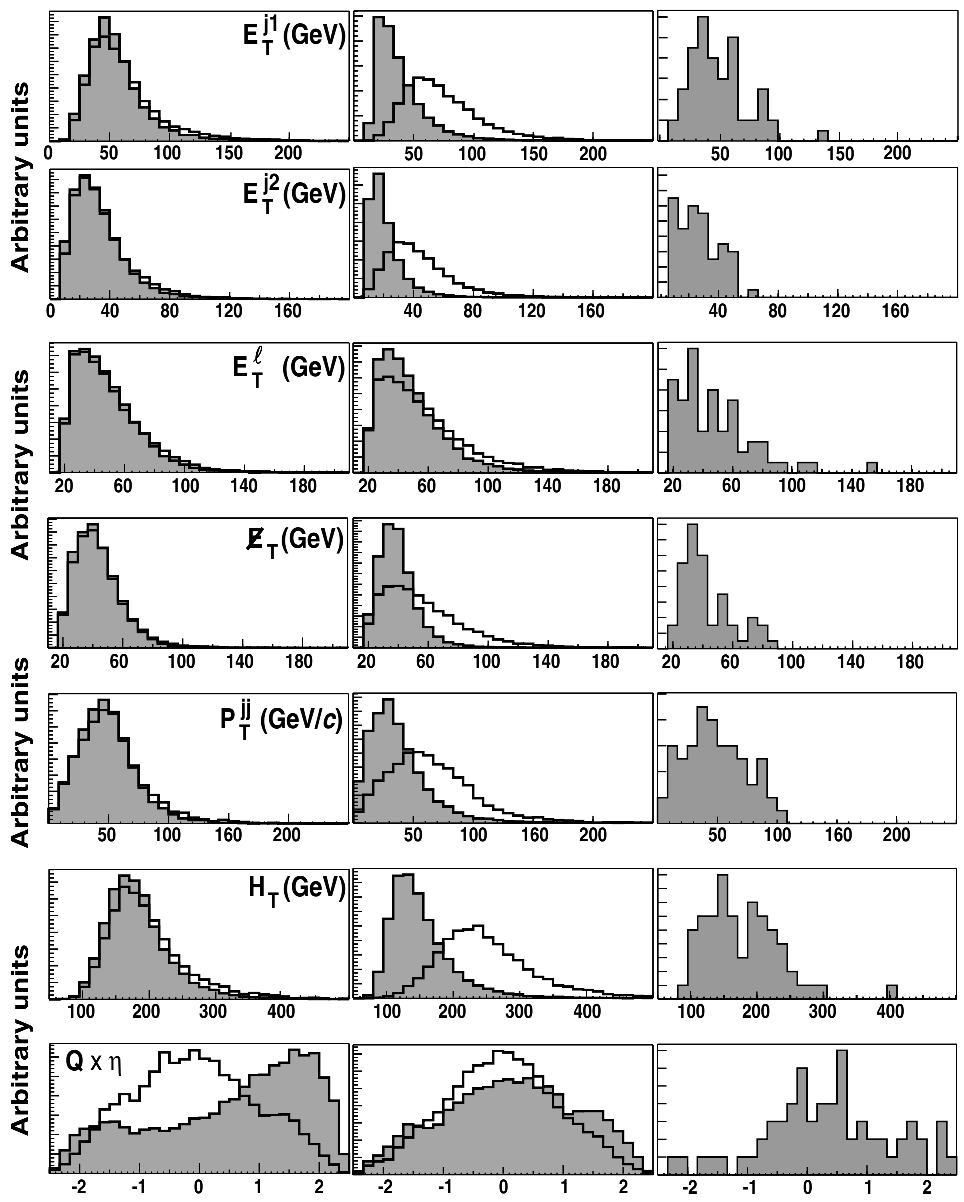

FIG. 3: Monte Carlo distributions for the seven variables used in the ANN. In the left plots, the open (shaded) histograms correspond to the W^* (Wg) channel. Similarly, in the middle-column plots open (shaded) histograms correspond to $t\bar{t}$ ($W+$ jets). All Monte Carlo distributions are normalized to unit area for comparison. The histograms in the right column correspond to the Run I data events.

where k is the event index, $\overrightarrow{T^k}$ is the target output corresponding to the correct event category, and $\overrightarrow{O^k}$ is the actual ANN output. For the input combinations having the lowest error function values we calculate the expected average upper limit on the single-top cross section. The lowest limit is obtained for the following input set: E_T^{j1}, E_T^{j2}, E_T^{ℓ}, $\not{E}_T$, P_T^{jj}, H_T, and $Q \times \eta$. The distributions of these variables are shown in Fig. 3. We note that the two backgrounds $t\bar{t}$ and non-top are kinematically situated on different sides of the signal. Finally, in the range of $7-20$ nodes in the intermediate layer, the error E has a weak minimum for 17 hidden nodes. The 7-17-3 configuration of nodes corresponds to 190 free parameters adjusted by training. As shown in Ref. [15], the output nodes O_1, O_2, O_3 estimate the Bayesian posterior probabilities for the three classes of events: non-top, signal, and $t\bar{t}$, respectively. This implies that $O_1 + O_2 + O_3 \approx 1$, so that all events tend to lie in the same plane in the output space. We indeed found that the output sum peaks at 1.0 with a maximum deviation of 0.1 for the three Monte Carlo samples. Consequently, we reduce the output space to two dimensions (x, y) by projecting all output points

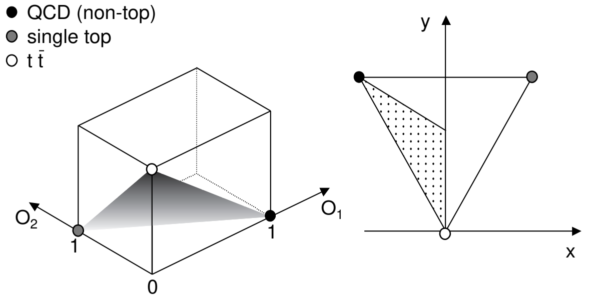

FIG. 4: Graphical representation of the projection mapping $(O_1, O_2, O_3) \Rightarrow (x, y)$. We expect the dotted area to be little populated, as events in this region would have to have: $O_2 < O_1, O_3$. In terms of probabilities, this inequality would contradict our earlier observation that signal is in general situated between $t\bar{t}$ and non-top backgrounds (Fig. 3).

onto the plane of equation $O_1 + O_2 + O_3 = 1$, as shown in Fig. 4. The (x, y) signal and background distributions are presented in Fig. 5, along with the data. We employ a maximum likelihood fit to these distributions to estimate the signal content of the Run I dataset. We note that Fig. 5 shows improved separation between signal and background compared to the individual input variables of Fig. 3. To quantify this separation one can for example define a "signal region" as the locus of the output points with $O_2 > O_1, O_3$. This signal region contains 67% of the signal, 27% of the non-top, and 24% of the $t\bar{t}$ Monte Carlo events, respectively.

The performance of this method is tested *a priori* by constructing simulated experiments using Monte Carlo generated event samples ("pseudo-Run I" datasets). A simulated experiment contains N_s signal, N_{nt} non-top, and $N_{t\bar{t}}$ $t\bar{t}$ events, where the number of events in each category is drawn from a Poisson distribution using the expected mean values in Table II. We propagate these events through the network and form the (x, y) output distribution. The latter is fitted using a background-constrained binned likelihood:

$$\mathcal{L}(n_s, n_{nt}, n_{t\bar{t}}) = \mathcal{L}_{background} \times \mathcal{L}_{shape} = G_1(n_{nt}) G_2(n_{t\bar{t}}) \prod_{i=1}^{N_{bins}} \frac{e^{-n_i} \cdot n_i^{d_i}}{d_i!} \qquad (2)$$

where n_s, n_{nt}, $n_{t\bar{t}}$ are the parameters of the fit, representing the numbers of signal, non-top, and $t\bar{t}$ events respectively present in the sample. Moreover, $n_i = n_s f_{s,i} + n_{nt} f_{nt,i} + n_{t\bar{t}} f_{t\bar{t},i}$ is the expected number of events in the i-th bin, and $f_{s,i}$, $f_{nt,i}$, $f_{t\bar{t},i}$ are the fractions of Monte Carlo single-top, non-top, or $t\bar{t}$ appearing in bin i. By d_i is denoted the number of events in the simulated experiment that populate the i-th bin. The Gaussian functions $G_1(n_{nt})$, $G_2(n_{t\bar{t}})$ constrain the non-top and $t\bar{t}$ backgrounds to the expected values: 43.3 ± 8.4 non-top and 7.4 ± 2.2 $t\bar{t}$ events, respectively.

TABLE III: Systematic uncertainties (in number of events). The second column corresponds to the theoretical prediction $\mu_{SM} = 3.9$ signal events. The third column lists the uncertainties estimated at the measured value $n_s = 23.9$ events. The overall uncertainties δn and δs are obtained by adding in quadrature the individual effects.

Normalization only effects		
Luminosity	0.16	0.98
Trigger and lepton identification	0.39	2.39
B-tag efficiency	0.39	2.39
Total δn	0.57	3.52
Shape and normalization effects		
Signal generator	0.12	0.06
Background generator	0.15	0.62
Jet energy measurement	1.49	2.76
Initial and final state radiation	0.51	0.80
Parton distribution functions	0.16	0.16
Top quark mass	0.17	0.86
Total δs	1.59	3.07

Different scenarios regarding signal expectation were also investigated. Specifically, we considered signal cross sections ranging from 0 pb to 20 pb. For each case, we performed 10000 simulated experiments. In Fig. 6 we show the n_s distributions for $\sigma_s = 2$ pb and $\sigma_s = 10$ pb. The mean values of n_s along with the 16 and 84 percentile points are presented in Fig. 7. We note that the mean of the fitted cross sections is consistent with the input cross section for all cases. We further tested the sensitivity of our method to the particular ratio of Wg and W^* cross sections (R_{Wg/W^*}). Two situations were considered: $\sigma_s = 2$ pb and $\sigma_s = 10$ pb. Simulated experiments were constructed with one of seven different values of R_{Wg/W^*}, but fitted to the standard templates of Fig. 5. The results are shown in Fig. 8, and show that the mean of the fitted cross sections varies by less than 11% across the R_{Wg/W^*} range studied.

The systematic uncertainties for this analysis are divided into two groups. The first group consists of systematic effects which modify only the rates of events accepted, and not the shapes of the distributions of input variables. The luminosity of 106 pb^{-1} has an uncertainty of $\pm 4.1\%$ [19]. The uncertainty on the trigger and lepton identification efficiency has been estimated to be 10%. Moreover, the efficiency for identifying jets containing B-hadrons has an uncertainty of 10% [7]. These uncertainties can be expressed in number of events by simply multiplying by the particular single-top content (Table III).

The second group of systematic uncertainties includes the effects that impact both the shapes of the Monte Carlo templates of Fig. 5 and the rates of events accepted. To illustrate how these systematics are extracted, let us consider the uncertainty associated with the signal generator (SG). We start by generating new W^* and Wg samples using the HERWIG [20] program instead of PYTHIA. Among the differences between the two genera-

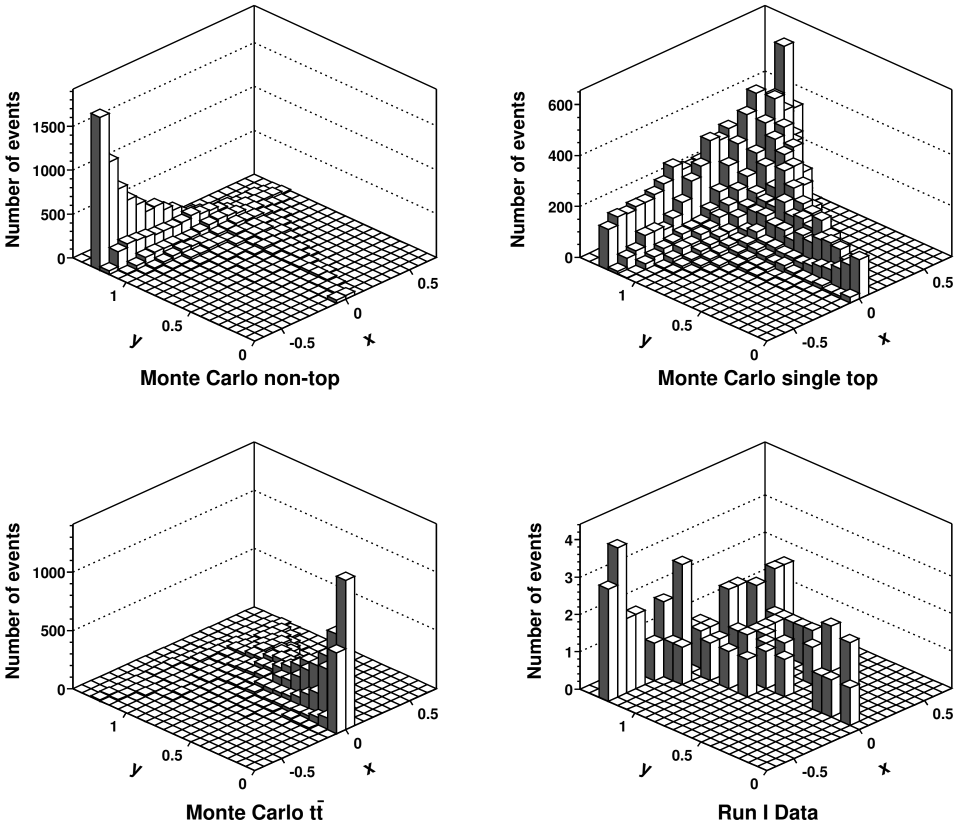

FIG. 5: Monte Carlo and CDF data 2-dimensional output distributions from projecting all output points onto the plane $O_1 + O_2 + O_3 = 1$. Of the 64 data events, 35 events overlap with the previous CDF search for combined single-top production [6].

tors, we note the hadronization approach and the underlying event modeling. The new samples are run through the ANN, and simulated experiments are constructed based on the recalculated acceptances and output shapes. Each experiment is then fitted to the standard templates of Fig 5. The uncertainty δs^{SG} is the absolute value of the shift in the mean fitted signal contribution n_s.

The uncertainty δs^{BG} related to the background generator is similarly calculated. In this case, the non-top sample was a mixture of two subsamples HERWIG $Wb\bar{b}$, and PYTHIA $Wc\bar{c}$ and Wc, while the $t\bar{t}$ background was generated with HERWIG. Fig. 9 shows a comparison between the HERWIG $Wb\bar{b}$ events and the default W+jets sample generated with PYTHIA. A good level of agreement regarding the shapes of the kinematic distributions can be observed. We note that δs^{BG} accounts for a small fraction of the total δs. As shown in Table III, the largest contribution to δs comes from the uncertainty in the measurement of jet momenta δs^{JES}. A change in the jet momentum scale simultaneously impacts five of the seven kinematic variables used in our analysis, which can lead to significant changes on an event by event basis. As detailed in Ref. [9], we apply $+1\sigma$ and -1σ shifts in the P_T scale of the jets, and define δs^{JES} as the average difference: $(\delta s^{+1\sigma} - \delta s^{-1\sigma})/2$. To study the uncertainty associated to the initial state radiation (ISR) we turn off ISR in PYTHIA and regenerate signal and background samples. We take δs^{ISR} to be one half the shift in the mean fitted signal contribution. To isolate the effects of final state radiation (FSR) we start from the no-ISR PYTHIA samples and select the (no-ISR, no-FSR) subset of events in which every jet matches to a final state parton within a (η, ϕ) distance of 0.4. The uncertainty δs^{FSR} is defined to be $(\delta s^{ISR,FSR} - \delta s^{ISR})/2$. Combined systematic uncertainty on the initial and final state radiation is obtained by adding in quadrature δs^{ISR} and δs^{FSR}. We evaluate the uncertainty δs^{PDF} due to the parton distribution function set by switching to the CTEQ 3L [21] set from the default GRV 94L [22] choice in PYTHIA. The last systematic effect studied is the top quark mass. We vary the top quark mass from the default $M_{top} = 175$ GeV to 170 and 180 GeV respectively, and generate new $W*$, Wg, and $t\bar{t}$ samples. We take $\delta s^{M_{top}}$ to be the larger of the shifts δs^{170} and δs^{180}.

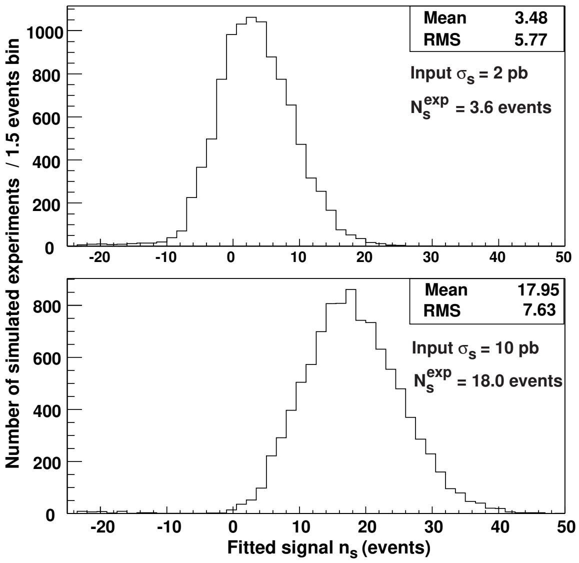

FIG. 6: Results from simulated experiments with input single-top cross sections of 2 pb (top) and 10 pb (bottom). In both cases the mean fitted number of signal events n_s agrees with the expected value N_s^{exp}.

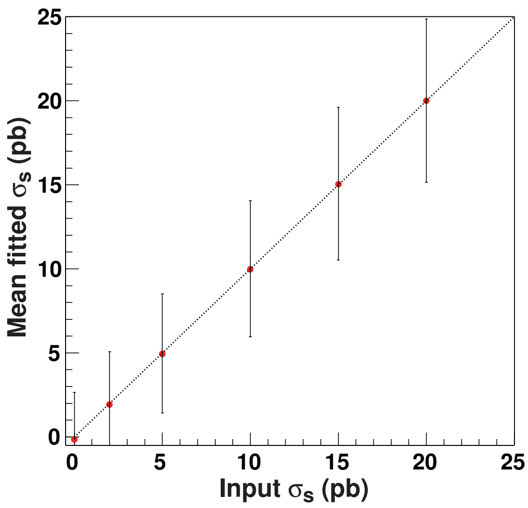

FIG. 7: Test of the ANN fitting technique under different hypotheses for signal cross section. As in Fig. 6, we note good agreement between the input and fitted signal cross sections. The theoretically calculated value is $\sigma_s^{SM} = 2.2$ pb [4]. The ends of the error bars mark the 16 and 84 percentile points for each fitted σ_s distribution.

Finally, the magnitude of the systematic uncertainties depends on the particular signal content used in performing simulated experiments. To exemplify this, let us consider the jet energy scale effect, which accounts for the largest fraction of the total δs. The variation of δs^{JES} with the input signal mean μ_s is presented in Fig. 10,

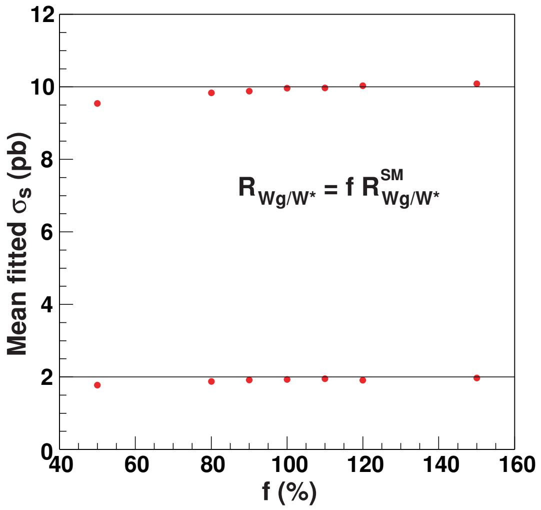

FIG. 8: Test of the ANN fitting technique under different hypotheses for Wg to W^* cross section ratio R_{Wg/W^*}. This ratio is expressed as the fraction f of the SM value $R^{SM}_{Wg/W^*} = 1.8$. Two values for the combined signal cross section are considered: $\sigma_s = 2$ pb and $\sigma_s = 10$ pb.

where the fit shown is a parabola. Consequently, the values listed in the second column of Table III ($\mu_s = 3.9$ events) will be used in deriving the a priori single-top results, while the third column values ($\mu_s = 23.9$ events) will be used in expressing the signal cross section measured from the CDF data.

Simulated experiments based on the SM expectations of Table II result in a distribution of n_s having a mean of 3.9 signal events and standard deviation of 5.9 events. Given this significant uncertainty, we focus on calculating the expected limit for single-top production. For each simulated experiment, $\mathcal{L}(n_s, n_{nt}, n_{t\bar{t}})$ is integrated

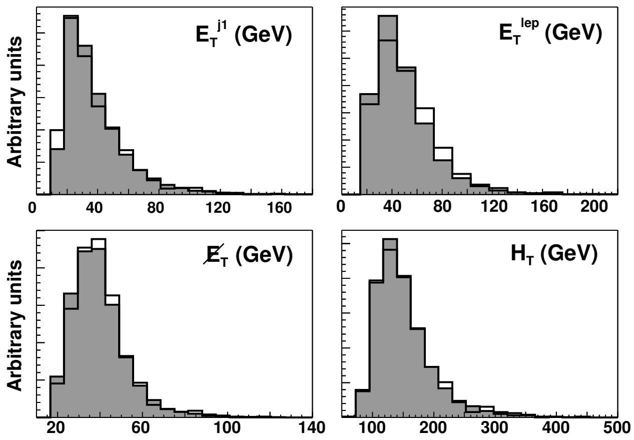

FIG. 9: Distributions of four of the ANN input variables for HERWIG $Wb\bar{b}$ events (open histograms) and the default PYTHIA W+jets sample (shaded histograms). All histograms are normalized to unit area for comparison.

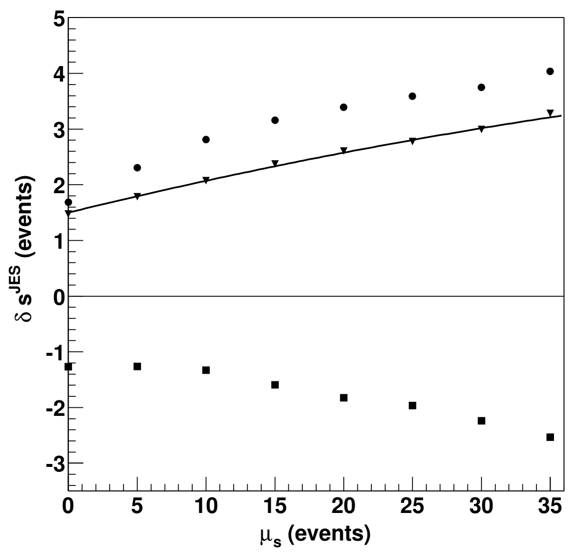

FIG. 10: Jet energy scale systematics as function of signal content μ_s. The circles show the shifts $\delta s^{+1\sigma}$ in the mean fitted signal contribution for a $+1\sigma$ increase in jet transverse momenta [9]. The squares represent the shifts $\delta s^{-1\sigma}$, while the triangles correspond to the combined δs^{JES} defined as $(\delta s^{+1\sigma} - \delta s^{-1\sigma})/2$.

out with respect to n_{nt}, $n_{t\bar{t}}$ for all values $n_s = \mu_s$ to obtain the probability density $f(\mu_s)$. We further assume a uniform "prior" distribution, and restrict to the physical range $\mu_s > 0$. The systematic effects are accounted for by smearing $f(\mu_s)$ as follows. To perform the smearing, we convolute $f(\mu_s)$ with two Gaussian functions of unit means and widths equal to δn and δs, respectively. The smeared function is numerically integrated to yield the 95% C.L. limit μ_{95} for the given simulated experiment. The mean value of the individual μ_{95}'s distribution is 10.6 pb and defines the mean expected (or "a priori") limit on the single-top cross section in the presence of the signal. Compared to the previous CDF combined single-top study [6], the neural-network method features an improvement of 21% in the a priori confidence limit. Roughly 7% of this improvement comes from retuning the selection criteria, with N_{js} selection replacing the $M_{l\nu b}$ window cut. Using a multi-variate technique (seven variables rather than H_T alone) accounts for the remaining 14%.

We have applied this method to the Run I dataset, where 64 events pass the selection criteria (Table II). The overlap with the 65-event sample of the search reported in Ref. [6] is 35 events. Fig. 5 shows the distribution of data events in the $O_1 + O_2 + O_3 = 1$ plane. We maximize the likelihood of Eqn. (2) to extract a signal contribution of $n_s = 23.9 \pm 7.7$ (stat) $\pm$ 4.7 (syst) events, or equivalently 13.5 ± 5.1 pb, including systematic uncertainties. This can be compared to the expected value of 2.2 pb.

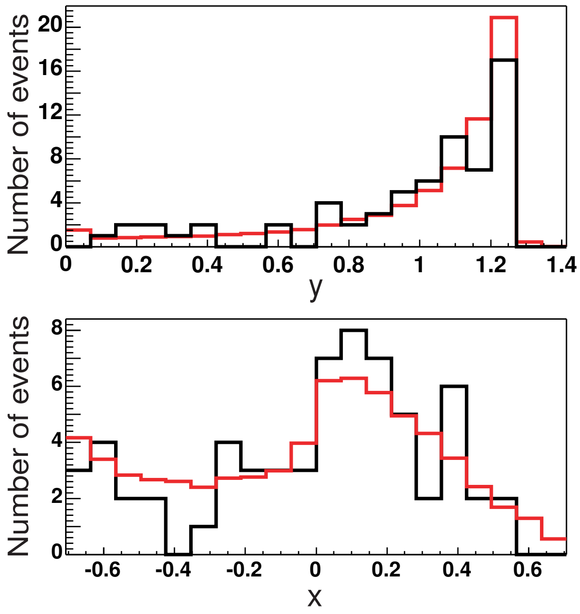

FIG. 11: The x and y neural-network output distributions for the data events (black line) and for the Monte Carlo events mixed in the proportions returned by the fit (gray line), respectively.

It can be seen in Fig. 5 that a significant fraction of the data events are indeed consistent with the simulated signal distribution. The numbers of background events returned by the likelihood fit are: 36.0 ± 6.2 non-top, and 7.6 ± 2.0 $t\bar{t}$ events, respectively. Fig. 11 shows the ANN output projected on the x and y axes for the data events and Monte Carlo events mixed according to the above fit results. Using the procedure previously described, we calculate the upper limit on single-top cross section:

$$\sigma(W^* + Wg) < 23.8 \text{ pb, at 95\% C.L. (stat+syst)} \quad (3)$$

Several cross checks of the results have been done. Due to the large expected non-top contribution in the data, the non-top background model is perhaps the most important factor determining the ANN fit result. As described in the previous sections, our non-top model is a PYTHIA sample of W+heavy flavor jets events. Using HERWIG $Wb\bar{b}$, $t\bar{t}$, and PYTHIA $Wc\bar{c}$, Wc samples we derived the systematic uncertainty listed in Table III. To further test how the shape of the non-top ANN output distribution depends on the particular Monte Carlo generator, we have studied a WBBGEN [23] sample of $Wb\bar{b}$ events. This sample was run through the ANN, and the resulting distribution was used to fit the data, along with the default signal and $t\bar{t}$ distributions of Fig. 5. The fit yields a signal contribution $\sigma_s = 11.1 \pm 5.2$ pb (stat+syst), consistent with the 13.5 ± 5.1 pb value obtained using the PYTHIA background estimation. Another case considered was the extreme alternative of replacing the default non-top sample with a PYTHIA sample of W+light flavor jets events where a jet is mistagged as a

B-jet. We have found that the ANN input and output distributions are very similar for the mistags and the default non-top samples, confirming that the mistags are modeled well in our analysis. Finally, we performed a "goodness of fit" test by employing a simple χ^2 fit. For this study, the (x, y) output space was divided into 10 bins with roughly equal data populations. We fit the data as a weighted sum of the signal and background templates (10-bin histograms) to obtain $\sigma_s = 15.0 \pm 5.9$ pb (stat+syst), with a χ^2 of 3.2 for 6 degrees of freedom, indicating reasonable agreement between data and Monte Carlo output distributions.

In summary, we have searched for single-top production using a neural-network method. We constructed a network whose outputs estimate signal and background posterior probabilities for every given event. The method presented here improves the previous CDF search strategy reported in [6]. By analyzing the Run I dataset, we found an upper limit of 24 pb (at 95% C.L.) on the single-top cross section.

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; and the Korea Science and Engineering Foundation (KoSEF); the Korea Research Foundation; and the Comision Interministerial de Ciencia y Tecnologia, Spain.

[1] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. 74, 2626 (1995).
[2] DØ Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. 74, 2632 (1995).
[3] S. Willenbrock and D.A. Dicus, Phys. Rev. D **34**, 155 (1986); S. Dawson and S. Willenbrock, Nucl. Phys. **B284**, 449 (1987); C.-P. Yuan, Phys. Rev. D **41**, 42 (1990); R.K. Ellis and S. Parke, *ibid.* **46**, 3785 (1992); D. Carlson and C.-P. Yuan, Phys. Lett. B **306**, 386 (1993); M. Smith and S. Willenbrock, Phys. Rev. D **54**, 6696 (1996); T. Stelzer, Z. Sullivan, and S. Willenbrock, Phys. Rev. D **58**, 094021 (1998).
[4] B.W. Harris *et al.*, Phys. Rev. D **66**, 054024 (2002).
[5] DØ Collaboration, V. Abazov *et al.*, Phys. Lett. B **517**, 282 (2001).
[6] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. D **65**, 091102 (2002).
[7] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. D **64**, 032002 (2001).
[8] F. Abe *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **271**, 387 (1988); D. Amidei *et al.*, *ibid.* **350**, 73 (1994); P. Azzi *et al.*, *ibid.* **360**, 137 (1995).
[9] CDF Collaboration, T. Affolder *et al.*, Phys. Rev. D **63**, 032003 (2001).
[10] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **80**, 2779 (1998).
[11] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **79**, 3819 (1997).
[12] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135** (2001). We use PYTHIA version 6.129a.
[13] R. Bonciani *et al.*, Nucl. Phys. **B529**, 424 (1998).
[14] J. Hertz, A. Krogh and R. Palmer, *Introduction to the Theory of Neural Computation*, (Addison-Wesley, Redwood City 1991).
[15] M. Richard and R. Lippmann, Neural Computation **3**, 461 (1991).
[16] L. Lönnblad *et al.*, Comput. Phys. Commun. **81**, 185 (1994). We use JETNET version 3.5.
[17] D. Carlson, Ph.D. thesis, Michigan State University, 1995.
[18] DØ Collaboration, B. Abbott *et al.*, hep-ex/9907041 (1999).
[19] D. Cronin-Hennessy, A. Beretvas, and P.F. Derwent, Nucl. Instrum. Methods Phys. Res., Sect. A **443**, 37 (2000).
[20] G. Marchesini *et al.*, Comput. Phys. Commun. **67**, 465 (1992). We use HERWIG version 5.6.
[21] H. Lai *et al.*, Phys. Rev. D **51**, 4763 (1995).
[22] M. Gluck, E. Reya, and A. Vogt, Z. Phys. C **67** 433 (1995).
[23] M.L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. **B632**, 343 (2002). The WBBGEN version used was from July 2001, and uses HERWIG v6.202.