Search for Single Top Quark Production at DØ Using Neural Networks

The DØ Collaboration

Fermi National Accelerator Laboratory, Batavia, Illinois 60510

Abstract

We present a search for electroweak production of single top quarks in $\approx$90 pb^{-1} of data collected with the DØ detector at the Fermilab Tevatron collider. Using arrays of neural networks to separate signals from backgrounds, we set upper limits on the cross sections of 17 pb for the s-channel process $p\bar{p} \rightarrow tb + X$, and 22 pb for the t-channel process $p\bar{p} \rightarrow tqb + X$, both at the 95% confidence level.

June 14, 2001

The DØ Collaboration

V.M. Abazov,[w] B. Abbott,[fff] A. Abdesselam,[k] M. Abolins,[yy] V. Abramov,[z] B.S. Acharya,[q]
D.L. Adams,[hhh] M. Adams,[ll] S.N. Ahmed,[u] G.D. Alexeev,[w] G.A. Alves,[b] N. Amos,[xx]
E.W. Anderson,[qq] Y. Arnoud,[i] M.M. Baarmand,[ccc] V.V. Babintsev,[z] L. Babukhadia,[ccc]
T.C. Bacon,[bb] A. Baden,[uu] B. Baldin,[kk] P.W. Balm,[t] S. Banerjee,[q] E. Barberis,[dd] P. Baringer,[rr]
J. Barreto,[b] J.F. Bartlett,[kk] U. Bassler,[l] D. Bauer,[bb] A. Bean,[rr] M. Begel,[bbb] A. Belyaev,[ii]
S.B. Beri,[o] G. Bernardi,[l] I. Bertram,[aa] A. Besson,[i] R. Beuselinck,[bb] V.A. Bezzubov,[z] P.C. Bhat,[kk]
V. Bhatnagar,[k] M. Bhattacharjee,[ccc] G. Blazey,[mm] S. Blessing,[ii] A. Boehnlein,[kk] N.I. Bojko,[z]
E.E. Boos,[y] F. Borcherding,[kk] K. Bos,[t] A. Brandt,[hhh] R. Breedon,[ee] G. Briskin,[ggg] R. Brock,[yy]
G. Brooijmans,[kk] A. Bross,[kk] D. Buchholz,[nn] M. Buehler,[ll] V. Buescher,[n] V.S. Burtovoi,[z]
J.M. Butler,[vv] F. Canelli,[bbb] W. Carvalho,[c] D. Casey,[yy] Z. Casilum,[ccc] H. Castilla-Valdez,[s]
D. Chakraborty,[mm] K.M. Chan,[bbb] S.V. Chekulaev,[z] D.K. Cho,[bbb] S. Choi,[hh] S. Chopra,[ddd]
J.H. Christenson,[kk] M. Chung,[ll] D. Claes,[zz] A.R. Clark,[dd] J. Cochran,[hh] L. Coney,[pp]
B. Connolly,[ii] W.E. Cooper,[kk] D. Coppage,[rr] S. Crépé-Renaudin,[i] M.A.C. Cummings,[mm]
D. Cutts,[ggg] G.A. Davis,[bbb] K. Davis,[cc] K. De,[hhh] S.J. de Jong,[u] K. Del Signore,[xx]
M. Demarteau,[kk] R. Demina,[ss] P. Demine,[i] D. Denisov,[kk] S.P. Denisov,[z] S. Desai,[ccc]
H.T. Diehl,[kk] M. Diesburg,[kk] G. Di Loreto,[yy] S. Doulas,[ww] P. Draper,[hhh] Y. Ducros,[m]
L.V. Dudko,[y] S. Duensing,[u] L. Duflot,[k] S.R. Dugad,[q] A. Duperrin,[j] A. Dyshkant,[mm]
D. Edmunds,[yy] J. Ellison,[hh] V.D. Elvira,[kk] R. Engelmann,[ccc] S. Eno,[uu] G. Eppley,[jjj]
P. Ermolov,[y] O.V. Eroshin,[z] J. Estrada,[bbb] H. Evans,[aaa] V.N. Evdokimov,[z] T. Fahland,[gg]
S. Feher,[kk] D. Fein,[cc] T. Ferbel,[bbb] F. Filthaut,[u] H.E. Fisk,[kk] Y. Fisyak,[ddd] E. Flattum,[kk]
F. Fleuret,[dd] M. Fortner,[mm] H. Fox,[nn] K.C. Frame,[yy] S. Fu,[aaa] S. Fuess,[kk] E. Gallas,[kk]
A.N. Galyaev,[z] M. Gao,[aaa] V. Gavrilov,[x] R.J. Genik II,[aa] K. Genser,[kk] C.E. Gerber,[ll]
Y. Gershtein,[ggg] R. Gilmartin,[ii] G. Ginther,[bbb] B. Gómez,[e] G. Gómez,[uu] P.I. Goncharov,[z]
J.L. González Solís,[s] H. Gordon,[ddd] L.T. Goss,[iii] K. Gounder,[kk] A. Goussiou,[bb] N. Graf,[ddd]
G. Graham,[uu] P.D. Grannis,[ccc] J.A. Green,[qq] H. Greenlee,[kk] S. Grinstein,[a] L. Groer,[aaa]
S. Grünendahl,[kk] A. Gupta,[q] S.N. Gurzhiev,[z] G. Gutierrez,[kk] P. Gutierrez,[fff] N.J. Hadley,[uu]
H. Haggerty,[kk] S. Hagopian,[ii] V. Hagopian,[ii] R.E. Hall,[ff] P. Hanlet,[ww] S. Hansen,[kk]
J.M. Hauptman,[qq] C. Hays,[aaa] C. Hebert,[rr] D. Hedin,[mm] J.M. Heinmiller,[ll] A.P. Heinson,[hh]
U. Heintz,[vv] T. Heuring,[ii] M.D. Hildreth,[pp] R. Hirosky,[kkk] J.D. Hobbs,[ccc] B. Hoeneisen,[h]
Y. Huang,[xx] R. Illingworth,[bb] A.S. Ito,[kk] M. Jaffré,[k] S. Jain,[q] R. Jesik,[bb] K. Johns,[cc]
M. Johnson,[kk] A. Jonckheere,[kk] M. Jones,[jj] H. Jöstlein,[kk] A. Juste,[kk] W. Kahl,[ss] S. Kahn,[ddd]
E. Kajfasz,[j] A.M. Kalinin,[w] D. Karmanov,[y] D. Karmgard,[pp] Z. Ke,[d] R. Kehoe,[yy] A. Khanov,[ss]
A. Kharchilava,[pp] S.K. Kim,[r] B. Klima,[kk] B. Knuteson,[dd] W. Ko,[ee] J.M. Kohli,[o]
A.V. Kostritskiy,[z] J. Kotcher,[ddd] B. Kothari,[aaa] A.V. Kotwal,[aaa] A.V. Kozelov,[z] E.A. Kozlovsky,[z]
J. Krane,[qq] M.R. Krishnaswamy,[q] P. Krivkova,[f] S. Krzywdzinski,[kk] M. Kubantsev,[ss]
S. Kuleshov,[x] Y. Kulik,[ccc] S. Kunori,[uu] A. Kupco,[g] V.E. Kuznetsov,[hh] G. Landsberg,[ggg]
W.M. Lee,[ii] A. Leflat,[y] C. Leggett,[dd] F. Lehner,[37,*] J. Li,[hhh] Q.Z. Li,[kk] X. Li,[d] J.G.R. Lima,[c]
D. Lincoln,[kk] S.L. Linn,[ii] J. Linnemann,[yy] R. Lipton,[kk] A. Lucotte,[i] L. Lueking,[kk]
C. Lundstedt,[zz] C. Luo,[oo] A.K.A. Maciel,[mm] R.J. Madaras,[dd] V.L. Malyshev,[w] V. Manankov,[y]
H.S. Mao,[d] T. Marshall,[oo] M.I. Martin,[mm] R.D. Martin,[ll] K.M. Mauritz,[qq] B. May,[nn]
A.A. Mayorov,[oo] R. McCarthy,[ccc] T. McMahon,[eee] H.L. Melanson,[kk] M. Merkin,[y]
K.W. Merritt,[kk] C. Miao,[ggg] H. Miettinen,[jjj] D. Mihalcea,[mm] C.S. Mishra,[kk] N. Mokhov,[kk]
N.K. Mondal,[q] H.E. Montgomery,[kk] R.W. Moore,[yy] M. Mostafa,[a] H. da Motta,[b] E. Nagy,[j]

F. Nang,[cc] M. Narain,[vv] V.S. Narasimham,[q] H.A. Neal,[xx] J.P. Negret,[e] S. Negroni,[j]
T. Nunnemann,[kk] D. O'Neil,[yy] V. Oguri,[c] B. Olivier,[l] N. Oshima,[kk] P. Padley,[jjj] L.J. Pan,[nn]
K. Papageorgiou,[ll] A. Para,[kk] N. Parashar,[ww] R. Partridge,[ggg] N. Parua,[ccc] M. Paterno,[bbb]
A. Patwa,[ccc] B. Pawlik,[v] J. Perkins,[hhh] M. Peters,[jj] O. Peters,[t] P. Pétroff,[k] R. Piegaia,[a]
B.G. Pope,[yy] E. Popkov,[vv] H.B. Prosper,[ii] S. Protopopescu,[ddd] J. Qian,[xx] R. Raja,[kk]
S. Rajagopalan,[ddd] E. Ramberg,[kk] P.A. Rapidis,[kk] N.W. Reay,[ss] S. Reucroft,[ww] M. Ridel,[k]
M. Rijssenbeek,[ccc] F. Rizatdinova,[ss] T. Rockwell,[yy] M. Roco,[kk] P. Rubinov,[kk] R. Ruchti,[pp]
J. Rutherfoord,[cc] B.M. Sabirov,[w] G. Sajot,[i] A. Santoro,[b] L. Sawyer,[tt] R.D. Schamberger,[ccc]
H. Schellman,[nn] A. Schwartzman,[a] N. Sen,[jjj] E. Shabalina,[ll] R.K. Shivpuri,[p] D. Shpakov,[ww]
M. Shupe,[cc] R.A. Sidwell,[ss] V. Simak,[g] H. Singh,[hh] J.B. Singh,[o] V. Sirotenko,[kk] P. Slattery,[bbb]
E. Smith,[fff] R.P. Smith,[kk] R. Snihur,[nn] G.R. Snow,[zz] J. Snow,[eee] S. Snyder,[ddd] J. Solomon,[ll]
V. Sorín,[a] M. Sosebee,[hhh] N. Sotnikova,[y] K. Soustruznik,[f] M. Souza,[b] N.R. Stanton,[ss]
G. Steinbrück,[aaa] R.W. Stephens,[hhh] F. Stichelbaut,[ddd] D. Stoker,[gg] V. Stolin,[x] A. Stone,[tt]
D.A. Stoyanova,[z] M. Strauss,[fff] M. Strovink,[dd] L. Stutte,[kk] A. Sznajder,[c] M. Talby,[j]
W. Taylor,[ccc] S. Tentindo-Repond,[ii] S.M. Tripathi,[ee] T.G. Trippe,[dd] A.S. Turcot,[ddd]
P.M. Tuts,[aaa] P. van Gemmeren,[kk] V. Vaniev,[z] R. Van Kooten,[oo] N. Varelas,[ll] L.S. Vertogradov,[w]
F. Villeneuve-Seguier,[j] A.A. Volkov,[z] A.P. Vorobiev,[z] H.D. Wahl,[ii] H. Wang,[nn] Z.-M. Wang,[ccc]
J. Warchol,[pp] G. Watts,[lll] M. Wayne,[pp] H. Weerts,[yy] A. White,[hhh] J.T. White,[iii] D. Whiteson,[dd]
J.A. Wightman,[qq] D.A. Wijngaarden,[u] S. Willis,[mm] S.J. Wimpenny,[hh] J. Womersley,[kk]
D.R. Wood,[ww] R. Yamada,[kk] P. Yamin,[ddd] T. Yasuda,[kk] Y.A. Yatsunenko,[w] K. Yip,[ddd]
S. Youssef,[ii] J. Yu,[kk] Z. Yu,[nn] M. Zanabria,[e] H. Zheng,[pp] Z. Zhou,[qq] M. Zielinski,[bbb]
D. Zieminska,[oo] A. Zieminski,[oo] V. Zutshi,[ddd] E.G. Zverev,[y] and A. Zylberstejn[m]

[a]Universidad de Buenos Aires, Buenos Aires, Argentina
[b]LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil
[c]Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
[d]Institute of High Energy Physics, Beijing, People's Republic of China
[e]Universidad de los Andes, Bogotá, Colombia
[f]Charles University, Center for Particle Physics, Prague, Czech Republic
[g]Institute of Physics, Academy of Sciences, Center for Particle Physics, Prague, Czech Republic
[h]Universidad San Francisco de Quito, Quito, Ecuador
[i]Institut des Sciences Nucléaires, IN2P3-CNRS, Universite de Grenoble 1, Grenoble, France
[j]CPPM, IN2P3-CNRS, Université de la Méditerranée, Marseille, France
[k]Laboratoire de l'Accélérateur Linéaire, IN2P3-CNRS, Orsay, France
[l]LPNHE, Universités Paris VI and VII, IN2P3-CNRS, Paris, France
[m]DAPNIA/Service de Physique des Particules, CEA, Saclay, France
[n]Universität Mainz, Institut für Physik, Mainz, Germany
[o]Panjab University, Chandigarh, India
[p]Delhi University, Delhi, India
[q]Tata Institute of Fundamental Research, Mumbai, India
[r]Seoul National University, Seoul, Korea
[s]CINVESTAV, Mexico City, Mexico
[t]FOM-Institute NIKHEF and University of Amsterdam/NIKHEF, Amsterdam, The Netherlands
[u]University of Nijmegen/NIKHEF, Nijmegen, The Netherlands
[v]Institute of Nuclear Physics, Kraków, Poland

3

[w]Joint Institute for Nuclear Research, Dubna, Russia

[x]Institute for Theoretical and Experimental Physics, Moscow, Russia

[y]Moscow State University, Moscow, Russia

[z]Institute for High Energy Physics, Protvino, Russia

[aa]Lancaster University, Lancaster, United Kingdom

[bb]Imperial College, London, United Kingdom

[cc]University of Arizona, Tucson, Arizona 85721

[dd]Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720

[ee]University of California, Davis, California 95616

[ff]California State University, Fresno, California 93740

[gg]University of California, Irvine, California 92697

[hh]University of California, Riverside, California 92521

[ii]Florida State University, Tallahassee, Florida 32306

[jj]University of Hawaii, Honolulu, Hawaii 96822

[kk]Fermi National Accelerator Laboratory, Batavia, Illinois 60510

[ll]University of Illinois at Chicago, Chicago, Illinois 60607

[mm]Northern Illinois University, DeKalb, Illinois 60115

[nn]Northwestern University, Evanston, Illinois 60208

[oo]Indiana University, Bloomington, Indiana 47405

[pp]University of Notre Dame, Notre Dame, Indiana 46556

[qq]Iowa State University, Ames, Iowa 50011

[rr]University of Kansas, Lawrence, Kansas 66045

[ss]Kansas State University, Manhattan, Kansas 66506

[tt]Louisiana Tech University, Ruston, Louisiana 71272

[uu]University of Maryland, College Park, Maryland 20742

[vv]Boston University, Boston, Massachusetts 02215

[ww]Northeastern University, Boston, Massachusetts 02115

[xx]University of Michigan, Ann Arbor, Michigan 48109

[yy]Michigan State University, East Lansing, Michigan 48824

[zz]University of Nebraska, Lincoln, Nebraska 68588

[aaa]Columbia University, New York, New York 10027

[bbb]University of Rochester, Rochester, New York 14627

[ccc]State University of New York, Stony Brook, New York 11794

[ddd]Brookhaven National Laboratory, Upton, New York 11973

[eee]Langston University, Langston, Oklahoma 73050

[fff]University of Oklahoma, Norman, Oklahoma 73019

[ggg]Brown University, Providence, Rhode Island 02912

[hhh]University of Texas, Arlington, Texas 76019

[iii]Texas A&M University, College Station, Texas 77843

[jjj]Rice University, Houston, Texas 77005

[kkk]University of Virginia, Charlottesville, Virginia 22901

[lll]University of Washington, Seattle, Washington 98195

4

According to the standard model, top quarks can be produced at the Tevatron $p\bar{p}$ collider via two mechanisms. One involves a virtual gluon that decays via the strong interaction into a $t\bar{t}$ pair. This mode has been observed by the CDF and DØ collaborations [1,2]. The second mechanism involves the electroweak production of a single top quark at a Wtb vertex, where W and b refer to the W boson and b quark. There are three ways of producing a single top quark: an s-channel process $q'\bar{q} \rightarrow t\bar{b}$, a t-channel mode $q'g \rightarrow tq\bar{b}$, and a final state generated via both the s- and t-channels, $bg \rightarrow tW$. For a top quark of mass 174.3 ± 5.1 GeV [3], the predicted cross sections are 0.75 ± 0.12 pb for tb production [4] and 1.47 ± 0.22 pb for tqb production [5], both calculated at next-to-leading-order precision with $\sqrt{s} = 1.8$ TeV. For tW production, the cross section is 0.15 pb at leading order [6]. The errors on the cross sections include contributions from the choice of parton distribution functions (PDF), the uncertainty on the gluon distribution, the choice of scale, and the experimental error on the top-quark mass measurement. In this Letter, we use the notation "tb" to refer to both $t\bar{b}$ and the charge-conjugate process $\bar{t}b$, and "tqb" to refer to both $tq\bar{b}$ and $\bar{t}\bar{q}b$.

The DØ collaboration recently published its first results on the production of single top quarks [7]. The search used a classical selection method to optimize the signal significance, based on the expected kinematic properties of the events. The 95% confidence level upper limits on the cross sections were determined to be 39 pb for the s-channel process and 58 pb for the t-channel process. The events contained an isolated electron or muon [8], missing transverse energy, and jets. At least one jet in each event was required to contain a "tagging" muon [9], used as an indication that the jet originated from the hadronization of a b quark. We have now developed a new and more powerful technique using arrays of neural networks that allow us to utilize the far more numerous untagged events in the search, as well as to improve the sensitivity for tagged events. This Letter describes the new method of event selection and presents significantly improved upper limits on the single-top-quark production cross sections.

The DØ detector [10] in Run 1 (1992–1996) had three major components: a drift-chamber-based central tracking system that included a transition radiation detector, a uranium/liquid-argon calorimeter with a central module (CC) and two end calorimeters (EC), and an outer muon spectrometer. For the electron channel, we use 91.9 ± 4.1 pb^{-1} of data collected with a trigger that required an electromagnetic (EM) energy cluster in the calorimeter, a jet, and missing transverse energy ($\not{E}_T$). For events passing the final selection criteria, the efficiency of this trigger is (90–99)%, depending on the location of the EM cluster in the calorimeter and on the presence of a tagging muon. In the muon channel, we use 88.0 ± 3.9 pb^{-1} of data acquired with several triggers that required either $\not{E}_T$, or a muon and a jet. The combined efficiency of these triggers is (92–98)%. A third data sample, obtained with a trigger requiring just three jets, is used for measuring one of the backgrounds. Since the multijet cross section is very large, this trigger was prescaled, and we have 0.8 pb^{-1} of such data. Each of the three samples contains approximately one million events. We reconstruct the events offline by applying the same criteria to identify electrons, jets, and isolated and tagging muons, as described in Ref. [7].

Single-top-quark events have a readily identifiable final-state topology. The top quark decays to a W boson and a b quark. The W boson decays to a central (i.e., low pseudorapidity $|\eta|$ [11]), isolated electron or muon with high transverse energy and momentum (E_T, p_T), and

a central, high-E_T neutrino. We infer the presence of the neutrino from the vector imbalance of E_T in the event. For an s-channel tb event, there are two central, high-E_T b jets, whereas for a t-channel tqb event, there is only one such b jet, plus a forward, light-quark jet, and a central low-E_T b jet. About 11% of the time [12], one of the b jets contains a tagging muon. Unfortunately, events like this are easily mimicked by many background processes. Before applying the initial selection criteria, the largest background in the electron channel comes from multijet events in which a jet is misidentified as an electron. We call this false isolated-electron background "mis-ID e." In the muon channel, the dominant component of the background is from multijet production with a coincident cosmic ray or beam-halo particle misidentified as an isolated muon. We call this false isolated-muon background "cosmic"; it is not included in the background model. Other backgrounds such as W+jets, $t\bar{t}$ pairs, and $b\bar{b}$ pairs, contribute at the few percent level. The $b\bar{b}$ background affects only the muon channel, when a muon from a b decay is misidentified as isolated. We call this false-isolated muon background "mis-ID μ."

TABLE I. Initial selection criteria.

Pass triggers and online filters	No mismeasured muons
Exactly one good isolated lepton	No mismeasured $\displaystyle{\not}E_T$
Any number of good tagging muons	No mismeasured jets
No photons	

	Analysis Channel			
Variable	e+jets/notag	e+jets/tag	μ+jets/notag	μ+jets/tag
$E_T(e)$	> 20 GeV	> 20 GeV		
$\lvert\eta(e)\rvert$	< 1.1 (CC)	< 1.1 (CC)		
	$> 1.5, < 2.5$ (EC)	$> 1.5, < 2.5$ (EC)		
$p_T(\text{isol } \mu)$			> 20 GeV	> 15 GeV
$\lvert\eta(\text{isol } \mu)\rvert$			$<\approx 0.8$	< 1.7
$p_T(\text{tag } \mu)$		> 4 GeV		> 4 GeV
$\lvert\eta(\text{tag } \mu)\rvert$		< 1.7		< 1.7
$E_T(\text{jet1})$	> 20 GeV	> 15 GeV	> 25 GeV	> 15 GeV
$E_T(\text{jet2})$	> 15 GeV	> 10 GeV	> 15 GeV	> 10 GeV
$E_T(\text{jet3})$	> 15 GeV	> 10 GeV	> 15 GeV	> 5 GeV
$E_T(\text{jet4})$			> 15 GeV	> 5 GeV
$\lvert\eta(\text{jet1})\rvert$	< 2.5	< 2.5	< 3.0	< 3.0
$\lvert\eta(\text{jet2})\rvert$	< 3.0	< 3.0	< 4.0	< 4.0
$\lvert\eta(\text{jet3})\rvert$	< 3.0	< 3.0	< 4.0	< 4.0
$\lvert\eta(\text{jet4})\rvert$			< 4.0	< 4.0
No. of jets	2 or 3	2 or 3	2 to 4	2 to 4
$\displaystyle{\not}E_T$	> 20 GeV (e in CC)	> 15 GeV (e in CC)	> 20 GeV	> 15 GeV
	> 25 GeV (e in EC)	> 20 GeV (e in EC)		
$\text{NN}_{\text{cosmic}}$				Output > 0.79

The analysis starts with the simple selection criteria listed in Table I. The requirements are determined by comparing distributions for the summed backgrounds with those of the data before most of the initial selection criteria are applied, and rejecting regions where there is poor agreement between data and the model for the sum of signals and backgrounds. The backgrounds are modeled using data weighted to represent the mis-ID e or mis-ID μ backgrounds, and six samples of Monte Carlo (MC) events: $t\bar{t}$, $Wb\bar{b}$, $Wc\bar{c}$ (including Wcs and $Ws\bar{s}$), Wjj (where j represents u, d, or g), WW, and WZ production.

The final requirement in Table I on the μ+jets/tag decay channel is a cutoff on the output of a neural network trained to reject events in which a cosmic ray has been misidentified as an isolated muon. The network uses the MLPfit package [13], and has seven input nodes, 15 hidden nodes, and one output node. The input variables are described in Table II. The pseudo-three-dimensional impact parameter is defined as $\mathrm{IP_{3d}} = \sqrt{\mathrm{IP_{BV}^2 + IP_{NB}^2}}$ where "BV" stands for "bend view" and "NB" for "non-bend" view for the muon trajectory through the spectrometer toroids.

TABLE II. Input variables to the cosmic ray neural network.

Variables	Description
$\Delta\phi(\mu, \mathrm{tag}\ \mu)$	Opening angle in the transverse plane between the high-p_T muon and tagging muon
$p_T(\mu)$, $p_T(\mathrm{tag}\ \mu)$	Muon transverse momentum
$z_{\mathrm{vert}}(\mu)$, $z_{\mathrm{vert}}(\mathrm{tag}\ \mu)$	Position of the primary vertex projected from the muon's track in the calorimeter
$\mathrm{IP_{3d}}(\mu)$, $\mathrm{IP_{3d}}(\mathrm{tag}\ \mu)$	Pseudo-3d impact parameter of the muon's trajectory relative to the beam axis

The cosmic-ray-rejection network ($\mathrm{NN_{cosmic}}$) is trained on a background sample of 575 data events chosen to contain cosmic rays by requiring $\Delta\phi(\mu, \mathrm{tag}\ \mu) > 2.4$, and on an equal-sized cosmic-ray-free signal sample of MC tb, tqb, $t\bar{t}$, and W+jets events. In the background training sample, the first muon passes either the isolated or nonisolated identification criteria. The results of the training are shown in Fig. 1(a). We accept events if the value of the network output is greater than 0.79, a cutoff designed to maximize rejection of the cosmic-ray component of the background. This selects 73% of the s-channel single-top-quark acceptance, 76% of the t-channel acceptance, 47% of the background included in the model (i.e., not including the rejected cosmic-ray component), and 28% of the data. The result of applying the network to the data and to the model for signal+background is shown in Fig. 1(b).

After applying the initial selection criteria, the combined acceptance for the s-channel tb signal is 3.8%; for the t-channel tqb signal it is 3.6%. The signal-to-background (S:B) ratios range from 1:40 to 1:470, depending on the production and decay channels. In what follows, we improve on the S:B ratios significantly by using arrays of neural networks that reject background while retaining adequate signal acceptance. The networks are trained to recognize detailed features of the signals and backgrounds, including correlations among the kinematic variables. They thereby provide superior separation of signal from background relative to classical selection techniques, where no correlations are taken into account.

7

Without the neural networks, no useful information about single-top-quark production can be obtained from the untagged events because the S:B ratio is so poor. Using the neural networks allows the untagged channels to provide as much sensitivity as the tagged ones.
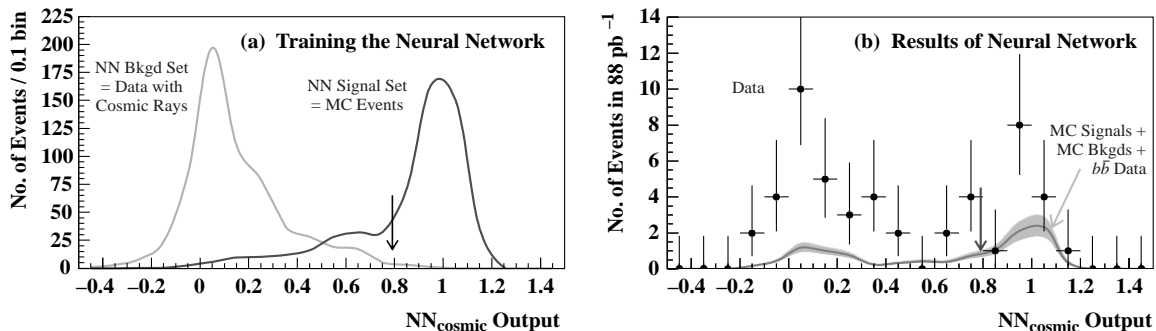

FIG. 1. Output from the neural network used to reject cosmic-ray contamination in the tagged muon+jets decay channel: (a) shows the results for the two training samples, and (b) shows the output for the sum of the signals and modeled backgrounds, and the data.

For training the neural networks, we divide the background event samples into five sets. There are three sets for W+jets events: "Wjj," "Wbb" (for the combined $Wb\bar{b}$ and $Wc\bar{c}$ MC samples), and "WW" (for the combined WW and WZ MC sets); and a set each for the $t\bar{t}$ and misidentified-lepton ("mis-ID l") backgrounds. Each analysis has five parallel networks, one to reject each type of background. There are four separate analyses: $tb \rightarrow e +$ jets, $tb \rightarrow \mu +$ jets, $tqb \rightarrow e +$ jets, and $tqb \rightarrow \mu +$ jets. We use the same networks for both untagged and tagged events combined, but choose different cutoffs on the output variables, depending on whether there is a tagging muon. This provides eight sets of results from 20 neural networks and 40 cutoffs on the outputs.

We use the package MLPfit [13], which has multi-layered perceptrons with a feed-forward structure and back-propagation of the errors for efficient computation. The networks are trained on samples of 2000–9000 signal MC events and background sets of the same size, using the "hybrid linear Broyden-Fletcher-Goldfarb-Shanno" learning method [14]. The performance of a neural network of course depends on the choice of input variables, which should be selected to provide maximal discrimination between signal and background, and should reflect as many different properties of the events as possible. However, too many input variables can worsen performance if the additional information is weak compared to the noise they introduce into the analysis. The variables chosen as inputs to the neural networks are defined in Table III.

The "best jet" in an event is the one that, when combined with the isolated lepton and the neutrino, generates an invariant mass closest to that of the top quark (174.3 GeV). The momentum components of the neutrino are derived from $\not{E}_T$ by assuming that it and the lepton arise from the decay of a W boson. Of the two possible solutions to the quadratic relation for the W mass, the one with the smallest absolute value of the neutrino's longitudinal momentum is chosen. The invariant mass variable M_{best} in Table III uses the best jet in its definition. Several variables use all the jets in the event ("alljets"). Those which exclude the best jet are denoted with a "prime".

TABLE III. Input variables to the neural networks.

Variable	Description	Wjj	Wbb	WW	$t\bar{t}$	Mis-ID l
E_T^{jet1}	Transverse energy of the highest-E_T jet	✓	✓	✓		✓
E_T^{jet2}	Transverse energy of the second-highest-E_T jet	✓	✓	✓		✓
$|\eta_{\mathrm{jet1}}|$	Absolute pseudorapidity of jet 1	✓	✓			✓
$|\eta_{\mathrm{jet2}}|$	Absolute pseudorapidity of jet 2	✓	✓			
p_T^{j1j2}	Transverse momentum of the jet1-jet2 system	✓	✓	✓		✓
H_T^{j1j2}	Scalar sum of the transverse energies of jet 1 and jet 2		✓	✓		
M_T^{j1j2}	Transverse mass of the jet1-jet2 system		✓	✓		✓
M^{j1j2}	Invariant mass of the jet1-jet2 system	✓	✓			
ΔR^{j1j2}	Opening angle between jet 1 and jet 2	✓	✓			✓
$|Y^{\mathrm{j1j2}}|$	Absolute rapidity of the jet1-jet2 system	✓	✓	✓		✓
p_T^{alljets}	Transverse momentum of the alljets system	✓	✓			✓
M_{alljets}	Invariant mass of the alljets system	✓	✓			
M_{all}	Invariant mass of the lepton-neutrino-alljets system ($=\sqrt{\hat{s}}$)	✓	✓		✓	
M_{best}	Invariant mass of the lepton-neutrino-bestjet system	✓	✓	✓	✓	
$\mathbf{P}_{\min}^{\mathrm{all}}{}'$	Smallest eigenvalue of the momentum tensor for the l-ν-alljets system	✓	✓		✓	
$\left(E_T^{\mathrm{jet1}}\right)'$	Transverse energy of the highest-E_T jet that is not the best jet				✓	
$\left(E_T^{\mathrm{jet2}}\right)'$	Transverse energy of the second-highest-E_T jet that is not the best jet				✓	✓
$\left(\sum E_T^{\mathrm{alljets}}\right)'$	Vector sum of the transverse energies of all the jets, except for the best jet			✓	✓	
$\left(H_T^{\mathrm{alljets}}\right)'$	Scalar sum of the transverse energies of all the jets, except for the best jet			✓	✓	
$\left(H_{\mathrm{alljets}}\right)'$	Scalar sum of the energies of all the jets, except for the best jet				✓	
$\left(M_{\mathrm{alljets}}\right)'$	Invariant mass of all the jets, except for the best jet				✓	
p_T^W	Transverse momentum of the lepton-neutrino system			✓		
M_T^W	Transverse mass of the lepton-neutrino system					
$\left| p_T^W - p_T^{\mathrm{alljets}} \right|$	Difference between transverse momenta of the l-ν and alljets systems					✓
$\left| \dfrac{M^W - M^{\mathrm{j1j2}}}{M^W} \right|$	Fractional difference between invariant masses of the l-ν and jet1-jet2 systems			✓		✓
$\not{E}_T$	Missing transverse energy				✓	✓
$p_T^{\mathrm{tag}\,\mu1}$	Transverse momentum of the highest-p_T tagging muon	✓	✓	✓		✓
$p_T^{\mathrm{tag}\,\mu2}$	Transverse momentum of the second-highest-p_T tagging muon	✓	✓	✓		✓

9

We optimize the performance of each network by choosing the number of hidden nodes that minimizes the network's error function. The numbers of nodes are given in Table IV.

TABLE IV. Number of input (i), hidden (h), and output (o) nodes (i–h–o) for each of the neural networks.

Background	s-channel tb		t-channel tqb	
Set	e+jets	μ+jets	e+jets	μ+jets
Wjj	14–16–1	14–24–1	14–18–1	14–21–1
Wbb	14–19–1	14–19–1	14–20–1	14–17–1
WW	14–30–1	14–19–1	14–27–1	14–23–1
$t\bar{t}$	9–20–1	9–20–1	9–20–1	9–19–1
Mis-ID l	16–17–1	16–20–1	16–28–1	16–15–1

Figure 2 illustrates the outputs of one set of five networks, after training on t-channel single-top-quark signals and each of the five background sets in the combined untagged and tagged electron+jets decay channels. The least amount of discrimination is obtained for the Wjj events, which is unfortunate since this process has a large cross section. Also, a significant fraction of the tqb signal cannot be differentiated from $t\bar{t}$ background. (This is not the case for the lower jet-multiplicity s-channel tb events.) The best separation of signal from background is obtained for events with a misidentified electron.

The outputs of the neural networks for the tqb search in the untagged electron+jets decay channel are shown in Fig. 3. They are obtained by passing all the signal and background events, and then the data, through each network. The cutoffs on the outputs are simultaneously chosen by minimizing the expected limit on the cross section in this channel.

The signal acceptances and numbers of events predicted to remain in the data after the initial event selection criteria are shown in Table V. Table VI shows the acceptances and event yields after the neural network selections. We measure the signal acceptances using MC samples of s-channel and t-channel single-top-quark events from the CompHEP event generator [15], with the PYTHIA package [16] used to simulate fragmentation, initial-state and final-state radiation, the underlying event, and leptonic decays of the W boson. For all event samples, the PDF is CTEQ3M [17]. The MC events are processed through a detector simulation program based on the GEANT package [18] and a trigger simulation, and are then reconstructed. We apply all selection criteria directly to the reconstructed MC events, except for several particle identification requirements that are taken into account using factors obtained from other DØ data.

We calculate the acceptance for $t\bar{t}$ pairs and for the five subprocesses for W+jets in a manner similar to that used for signal, and then convert to a number of events using the integrated luminosity for each channel and the appropriate cross section. The $t\bar{t}$ background is modeled using HERWIG [19]. The $Wb\bar{b}$, $Wc\bar{c}$, and Wjj processes use CompHEP, followed by PYTHIA, and the diboson processes are from PYTHIA. The cross sections are DØ's measured value for $t\bar{t}$ [20], leading-order values for $Wb\bar{b}$, $Wc\bar{c}$, and Wjj [21], and next-to-leading-order values for WW [22] and WZ [23].
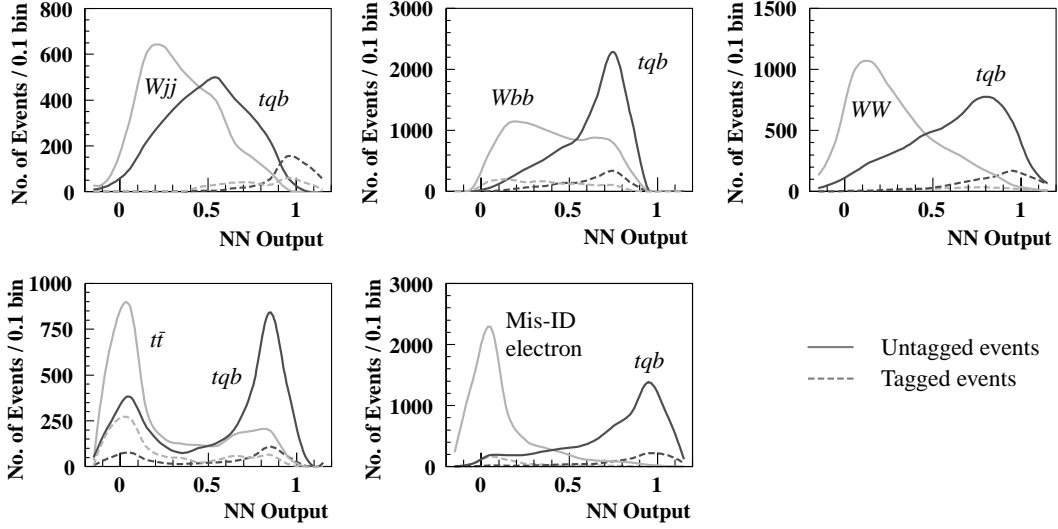
FIG. 2. Results of training the five neural networks used to separate t-channel *tqb* signal from background in the untagged and tagged electron+jets decay channels. For each plot, the lighter curves are for background and the darker curves for the signal. The solid curves are for untagged events, and the dashed curves for tagged ones.
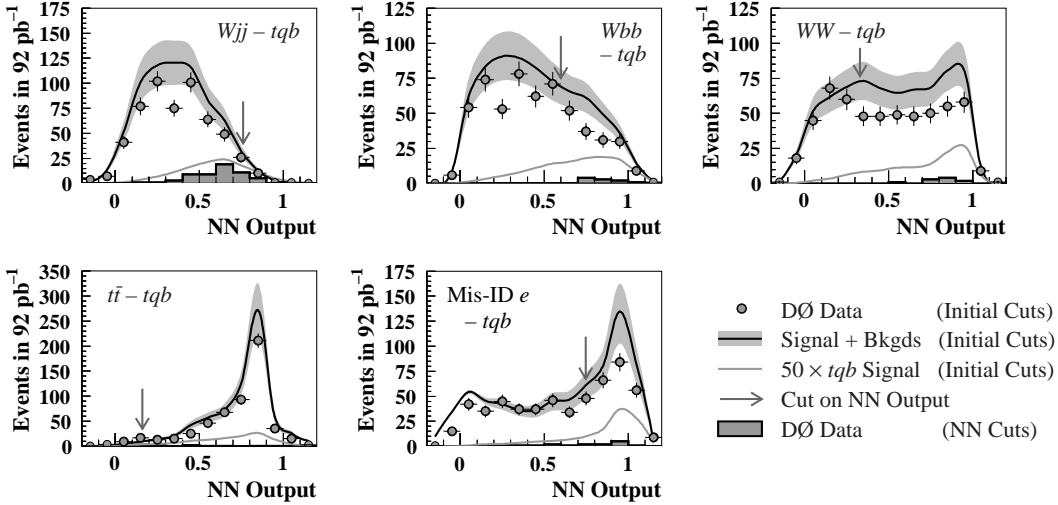
FIG. 3. Outputs from the five neural networks used to separate t-channel *tqb* signal from background in the untagged electron+jets decay channel. For each plot, the upper curve with error band shows the sum of signal and all backgrounds, the lower curve shows 50 times the expected single-top-quark signal, the shaded circles with error bars show the data after the initial event selection, and the shaded histogram shows the data after all neural network selection criteria have been applied, except for the network shown in that plot.

11

TABLE V. Signal acceptances (as percentages of the total cross sections) and numbers of events expected after application of initial selection criteria.

	e+jets/notag	e+jets/tag	μ+jets/notag	μ+jets/tag
	Signal Acceptances			
tb	$(2.04 \pm 0.14)\%$	$(0.32 \pm 0.03)\%$	$(1.27 \pm 0.12)\%$	$(0.16 \pm 0.01)\%$
tqb	$(1.81 \pm 0.15)\%$	$(0.20 \pm 0.02)\%$	$(1.44 \pm 0.14)\%$	$(0.12 \pm 0.01)\%$
	Numbers of Events			
tb	1.41 ± 0.25	0.22 ± 0.04	0.84 ± 0.16	0.11 ± 0.02
tqb	2.44 ± 0.43	0.27 ± 0.05	1.85 ± 0.34	0.16 ± 0.03
$t\bar{t}$	6 ± 2	1.42 ± 0.43	8 ± 3	1.45 ± 0.44
$Wb\bar{b}$	4 ± 1	0.69 ± 0.23	2 ± 1	0.27 ± 0.09
$Wc\bar{c}$	17 ± 6	0.80 ± 0.27	11 ± 4	0.24 ± 0.09
Wjj	499 ± 120	2.88 ± 0.95	332 ± 84	0.65 ± 0.28
WW	17 ± 2	0.31 ± 0.05	13 ± 2	0.24 ± 0.05
WZ	3 ± 0	0.11 ± 0.02	2 ± 0	0.06 ± 0.01
Mis-ID l	112 ± 13	9.79 ± 1.32	19 ± 4	2.06 ± 1.02
Bkgd for tb	661 ± 130	16.28 ± 2.03	389 ± 91	5.93 ± 1.22
Bkgd for tqb	660 ± 130	16.23 ± 2.02	388 ± 91	5.88 ± 1.22
Data	558	14	398	14

In the electron channel, the misidentified-electron background is measured using multijet data. For each jet that passes the electron E_T and $|\eta|$ requirements, the events are weighted by the probability that a jet mimics an electron. These probabilities are determined from the same multijet sample, but for $\not{E}_T < 15$ GeV, and are found to be $(0.0231 \pm 0.0039)\%$ (CC), and $(0.0850 \pm 0.0118)\%$ (EC) in untagged events, and $(0.0154 \pm 0.0019)\%$ (CC), and $(0.0612 \pm 0.0057)\%$ (EC) in tagged events. The probabilities are independent of the jet E_T for $E_T > 20$ GeV. We normalize the integrated luminosity of the multijet sample to match the data sample used in the search for signal, and correct for a small difference in trigger efficiency between the two samples.

In the muon channel, the misidentified-muon background is from $b\bar{b}$ events; it is produced when one or both of the b quarks decays semileptonically to a muon, and one muon is misidentified as isolated. There are two ways such events can mimic signals. First, one of the b jets may not be reconstructed, and its muon can therefore appear to be isolated. Second, a muon can be emitted wide of its jet and be reconstructed as an isolated muon. The background from each source is measured using data collected with the same triggers as used for the muon signal. The events are required to pass all selection criteria, except that the muon, which otherwise passes the isolated muon requirements, is within a jet. Events with truly isolated muons are excluded. Each event is then weighted by the probability that a nonisolated muon is reconstructed as an isolated one. This probability is measured using the same data, but for $\not{E}_T < 15$ GeV, and is found to be a few percent for each source on average. The probabilities are parametrized as a function of the muon p_T; they are higher at

low p_T and fall to zero for $p_T > 28$ GeV. We calculate a weighted average of the two results to obtain the number of mis-ID μ background events.

TABLE VI. Signal acceptances (as percentages of the total cross sections) and numbers of events expected after application of both the initial and neural network selection criteria.

	e+jets/notag	e+jets/tag	μ+jets/notag	μ+jets/tag
s-channel tb Search		Signal Acceptance		$\pm$
tb	$(0.29 \pm 0.02)\%$	$(0.20 \pm 0.02)\%$	$(0.24 \pm 0.03)\%$	$(0.12 \pm 0.01)\%$
		Numbers of Events		
tb	0.20 ± 0.04	0.14 ± 0.03	0.16 ± 0.03	0.08 ± 0.02
tqb	0.15 ± 0.03	0.10 ± 0.02	0.14 ± 0.03	0.06 ± 0.01
$t\bar{t}$	0.29 ± 0.10	0.20 ± 0.06	0.26 ± 0.08	0.10 ± 0.03
$Wb\bar{b}$	0.15 ± 0.05	0.30 ± 0.10	0.16 ± 0.06	0.16 ± 0.06
$Wc\bar{c}$	0.77 ± 0.27	0.32 ± 0.11	0.74 ± 0.27	0.12 ± 0.05
Wjj	13.76 ± 3.45	0.89 ± 0.42	14.56 ± 4.28	0.43 ± 0.23
WW	0.24 ± 0.11	0.08 ± 0.03	0.19 ± 0.07	0.12 ± 0.03
WZ	0.06 ± 0.02	0.04 ± 0.01	0.06 ± 0.02	0.02 ± 0.01
Mis-ID l	1.06 ± 0.23	0.37 ± 0.11	0.00 ± 0.00	0.06 ± 0.05
Total Bkgd	16.49 ± 3.83	2.29 ± 0.61	16.10 ± 4.66	1.07 ± 0.32
Data	15	2	9	1
t-channel tqb Search		Signal Acceptance		$\pm$
tqb	$(0.29 \pm 0.03)\%$	$(0.13 \pm 0.01)\%$	$(0.38 \pm 0.04)\%$	$(0.08 \pm 0.01)\%$
		Numbers of Events		
tqb	0.38 ± 0.08	0.17 ± 0.03	0.50 ± 0.10	0.11 ± 0.02
tb	0.10 ± 0.02	0.12 ± 0.02	0.10 ± 0.02	0.07 ± 0.01
$t\bar{t}$	0.78 ± 0.24	0.46 ± 0.14	1.63 ± 0.50	0.20 ± 0.06
$Wb\bar{b}$	0.07 ± 0.03	0.20 ± 0.07	0.08 ± 0.03	0.11 ± 0.04
$Wc\bar{c}$	0.36 ± 0.13	0.23 ± 0.09	0.42 ± 0.16	0.10 ± 0.04
Wjj	10.16 ± 3.26	0.83 ± 0.39	13.18 ± 4.60	0.22 ± 0.13
WW	0.43 ± 0.11	0.04 ± 0.01	1.08 ± 0.31	0.09 ± 0.03
WZ	0.08 ± 0.03	0.03 ± 0.01	0.20 ± 0.05	0.03 ± 0.01
Mis-ID l	0.76 ± 0.20	0.29 ± 0.08	0.03 ± 0.03	0.08 ± 0.03
Total Bkgd	12.75 ± 3.58	2.22 ± 0.56	16.73 ± 5.13	0.91 ± 0.23
Data	10	2	14	1

After applying the neural network selection criteria, the combined acceptances are 0.86% for the tb signal and 0.88% for the tqb signal, with the S:B ratios increased to between 1:9 and 1:99, a factor of 4–8 improvement compared with those after the initial event selections.

We use a Bayesian approach [24] to calculate limits on the cross sections for single-top-quark production in the s-channel and t-channel modes. The inputs are the numbers of observed events, the signal acceptances and backgrounds, and the integrated luminosities. Covariance matrices are used to describe the correlated uncertainties on these quantities. A flat prior is used for the single-top-quark cross section, and a multivariate Gaussian prior for the other quantities. We calculate the likelihood functions in each decay channel and combine them to obtain the following 95% confidence level upper limits:

- $\sigma(p\bar{p} \to tb + X) < 17$ pb

- $\sigma(p\bar{p} \to tqb + X) < 22$ pb.

The contributions of each decay channel to these results are shown in Table VII.

TABLE VII. The 95% confidence level upper limits on cross sections for the two production modes of single top quarks. Values are in picobarns.

Channel	Initial Selections			Neural Network Selections		
	e+jets	μ+jets	e+jets + μ+jets	e+jets	μ+jets	e+jets + μ+jets
s-channel tb						
untagged	118	165	107	44	45	35
tagged	27	108	35	26	39	19
untagged+tagged	27	104	36	22	26	17
t-channel tqb						
untagged	131	156	141	41	43	33
tagged	43	141	55	43	59	30
untagged+tagged	42	128	60	27	32	22

To conclude, we have searched for the electroweak production of single top quarks using a neural-network signal-selection technique. We find no evidence for such production and set upper limits on the cross sections for s-channel production of tb and t-channel production of tqb. The limits are consistent with expectations from the standard model.

REFERENCES

* Visitor from University of Zurich, Zurich, Switzerland.

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

[2] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. 74 (1995) 2632.

[3] L. Demortier *et al.*, The Top Averaging Group, for the CDF and DØ Collaborations, Fermilab-TM-2084 (1999).

[4] M.C. Smith and S. Willenbrock, Phys. Rev. D 54 (1996) 6696. The calculation has been updated to use CTEQ5M1 [25].

[5] T. Stelzer, Z. Sullivan, and S. Willenbrock, Phys. Rev. D 56 (1997) 5919; 58 (1998) 094021. The calculation has been updated to use CTEQ5M1, which has new evolution equations for the b quark distribution. The t-channel tqb cross section goes down by 15% compared with using CTEQ4M.

[6] A.P. Heinson, A.S. Belyaev, and E.E. Boos, Phys. Rev. D 56 (1997) 3114.

[7] B. Abbott *et al.* (DØ Collaboration), Phys. Rev. D 63 (2001) 031101(R).

[8] An electron is considered isolated if the total energy deposited in the calorimeter in a cone of radius 0.4 around the energy cluster center is not more than 15% greater than the electromagnetic energy deposited in a more central cone of radius 0.2. A muon is considered isolated if the opening angle ΔR between the muon and any jet is at least 0.5, where $\Delta R = \sqrt{[\Delta\phi(\mu,\text{jet})]^2 + [\Delta\eta(\mu,\text{jet})]^2}$.

[9] Tagging muons have $\Delta R < 0.5$.

[10] S. Abachi *et al.* (DØ Collaboration), Nucl. Instrum. Methods A 338 (1994) 185.

[11] The detector pseudorapidity η is defined as the pseudorapidity of the calorimeter cluster (electron or jet) or central track (muon) with respect to the center of the DØ detector, where $\eta = -\ln\tan(\theta/2)$, and θ is the polar angle relative to the proton beam direction.

[12] D.E. Groom *et al.* (Particle Data Group), Eur. Phys. J. C15 (2000) 1.

[13] MLPfit: a Tool for Multi-Layer Perceptrons, http://schwind.home.cern.ch/schwind/MLPfit.html. We used version 1.40.

[14] S. McLoone *et al.*, IEEE Trans. Neural Networks 9 (1998) 669.

[15] A. Pukhov *et al.*, hep-ph/9908288. We used CompHEP version 3.0.

[16] T. Sjöstrand, Comput. Phys. Commun. 82 (1994) 74. We used PYTHIA version 5.7 and JETSET version 7.4.

[17] H.L. Lai *et al.*, Phys. Rev. D 51 (1995) 4763.

[18] R. Brun *et al.*, CERN Program Library Writeup W5013, 1993 (unpublished). We used GEANT version 3.15.

[19] G. Marchesini *et al.*, Comput. Phys. Commun. 67 (1992) 465. We used HERWIG version 5.7.

[20] S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. 79 (1997) 1203; B. Abbott *et al.* (DØ Collaboration), Phys. Rev. Lett. 83 (1999) 1908. At $m_{\text{top}} = 174.3$ GeV, the $t\bar{t}$ cross section is 5.7 ± 1.6 pb.

[21] The leading-order cross sections with CTEQ3M from CompHEP are $7.3^{+2.3}_{-1.7}$ pb for $Wb\bar{b}$, $27.9^{+8.9}_{-6.4}$ pb for $Wc\bar{c}+Wcs+Ws\bar{s}$, and 966^{+203}_{-164} pb for Wjj (including all light jet processes). The central values are for a scale of m_W^2.

[22] J. Ohnemus, Phys. Rev. D 44 (1991) 1403. We rescaled the result from the HMRS(B) PDF to CTEQ3M and obtained 10.40 ± 0.23 pb.

[23] J. Ohnemus, Phys. Rev. D 44 (1991) 3477. We rescaled the result from the HMRS(B) PDF to CTEQ3M and obtained $2.93^{+0.26}_{-0.07}$ pb.

[24] I. Bertram *et al.*, Fermilab-TM-2104 (2000).

[25] H.L. Lai *et al.*, Eur. Phys. J. C 12 (2000) 375. See "Note added in proof" for CTEQ5M1.