Search for single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV
We present a search for electroweak production of single top quarks in the $s$-channel and $t$-channel using neural networks for signal-background separation. We have analyzed 230 pb$^{-1}$ of data collected with the DØ detector at the Fermilab Tevatron Collider at a center-of-mass energy of 1.96 TeV and find no evidence for a single top quark signal. The resulting 95% confidence level upper limits on the single top quark production cross sections are 6.4 pb in the $s$-channel and 5.0 pb in the $t$-channel.

PACS numbers: 14.65.Ha; 12.15.Ji; 13.85.Qk
INTRODUCTION

Top quark physics provides fundamental knowledge of the strong and electroweak sectors of the standard model and offers discovery potential for physics beyond the standard model. The top quark was discovered in 1995 at the Fermilab Tevatron Collider in $t\bar{t}$ events produced through the strong interaction \cite{1}. The standard model predicts that proton-antiproton collisions should also produce single top quarks through the electroweak interaction. Studying single top quark production will provide direct measurements of the CKM matrix element $|V_{tb}|$ and top quark polarization, and will probe possible new physics in the top quark sector \cite{2,3}.

There are two main modes of single top quark production as shown in Fig. 1: (a) the $s$-channel ($tb$) process $p\bar{p} \to t\bar{b}+X$ and the $t$-channel ($tq\bar{b}$) process $p\bar{p} \to tq\bar{b}+X$. The production cross sections have been calculated at next-to-leading order (NLO) in the strong coupling constant \cite{4,5,6,7,8}, yielding $0.88 \pm 0.14$ pb for the $s$-channel and $1.98 \pm 0.30$ pb \cite{9} for the $t$-channel, assuming a top quark mass of $m_t = 175$ GeV. Both the DØ and CDF collaborations have previously performed searches for single top quark production \cite{10}. Recently, CDF performed a search using 160 pb$^{-1}$ of data and obtained upper limits of 13.6 pb (s-channel), 10.1 pb (t-channel), and 17.8 pb (s+t combined) at 95% C.L. \cite{11}.

![FIG. 1: Representative Feynman diagrams for electroweak top quark production at the Tevatron Collider. This figure shows (a) the s-channel and (b) the t-channel.](image)

In this Letter, we present a new search for electroweak production of single top quarks with the DØ detector. Our search focuses on the final state in which the top quark decays into a $b$ quark and a $W$ boson, and the $W$ boson then decays into an electron or a muon, and a neutrino. This gives rise to an event signature with a high transverse momentum ($p_T$) lepton with significant missing transverse energy ($E_T$) from the neutrino. In addition, the $s$-channel has two $b$-quark jets, whereas the $t$-channel typically has one $b$-quark jet together with a light-quark jet. The $b$-quark jet in the $t$-channel, as represented in Fig. 1(b), is usually emitted in the forward direction with low $p_T$ and is often undetected \cite{9}. The main backgrounds in this analysis are the $W$ boson production in association with jets ($W$+jets), top quark pair ($tt\bar{t}$), and multijet production. We use neural networks to separate the signals from the backgrounds. In the absence of any significant evidence for signal, we build a binned likelihood from the neural network outputs to set upper limits on the single top quark production cross sections.

THE DØ DETECTOR

The DØ detector consists of a central tracking system, liquid-argon/uranium calorimeters, and an iron toroid muon spectrometer \cite{12}. The central tracking system covers the detector pseudorapidity \kappa region $|\eta_{\text{det}}| < 2.5$. It includes a silicon microstrip tracker and a scintillating fiber tracker, both located within a 2 T solenoidal magnet. The calorimeters consist of a central barrel in the region $|\eta_{\text{det}}| < 1.1$, and two end caps extending the coverage to $|\eta_{\text{det}}| \approx 4.2$. The muon system outside the calorimeter consists of a layer of tracking detectors and scintillation counters before 1.8 T toroids, followed by two similar layers after the toroids. Luminosity is measured using plastic scintillator arrays located in front of the end calorimeters.

DATA SET AND EVENT SELECTION

The data were recorded between August 2002 and March 2004 using a trigger that required an electromagnetic energy cluster and a jet in the calorimeter for the electron channel, or a muon and a jet for the muon channel. The integrated luminosity is 226 pb$^{-1}$ for the electron channel and 229 pb$^{-1}$ for the muon channel.

In the electron channel, we require exactly one isolated electron \cite{13} with $p_T > 15$ GeV and $|\eta_{\text{det}}| < 1.1$. In the muon channel, events are selected by requiring exactly one isolated muon \cite{14} with $p_T > 15$ GeV and $|\eta_{\text{det}}| < 2.0$. For both channels, events are also required to have $E_T > 15$ GeV. Events must have from two to four jets with the leading jet $p_T > 25$ GeV and $|\eta_{\text{det}}| < 2.5$, and all other jets having $p_T > 15$ GeV and $|\eta_{\text{det}}| < 3.4$. Jets are defined using a cone algorithm with radius $R = 0.5$. Misreconstructed events are rejected by requiring that the direction of $E_T$ is not aligned or anti-aligned in azimuth with the lepton or the jets.

The fraction of signal-like events is further enhanced through the selection of $b$-quark jets that are identified by reconstructing displaced vertices from long-lived particles. A displaced vertex is selected by requiring the transverse decay-length significance, $L_{xy}/\sigma_{L_{xy}}$, to be greater than seven, where $L_{xy}$ is the decay length and $\sigma_{L_{xy}}$ is the uncertainty on $L_{xy}$, calculated from the error matrices of the tracks and the primary vertex. A jet is considered $b$-tagged by this algorithm if a displaced vertex lies within a cone of radius $R = 0.5$ around the jet axis \cite{10}.
For both s-channel and t-channel searches, we separate the data into independent analysis sets based on final-state lepton flavor (electron or muon) and b-tag multiplicity. To take advantage of the different final state topologies, we separate single-tagged (\(=1\) tag) events from double-tagged (\(\geq 2\) tags) events. In the t-channel search, we additionally require that one of the jets is not b tagged.

**ACCEPTANCES AND YIELDS**

We estimate the kinematic and geometric acceptances for s-channel and t-channel single top quark production using the CompHEP matrix element event generator \([17]\) with \(m_t = 175\) GeV. The factorization scales are \(m_t^2\) for the s-channel samples and \((m_t/2)^2\) for the t-channel samples. For the s-channel (t-channel) search, the t-channel (s-channel) is considered as background.

We use both Monte Carlo events and data to estimate the background yields. The \(W+\)jets and diboson (\(WW\) and \(WZ\)) backgrounds are estimated using events generated with Alpgen \([18]\). The diboson background yields are normalized to NLO cross sections computed with MCFM \([19]\). The total \(W+\)jets yield is normalized to the yield in data corrected for the presence of multijet, \(tt\) and dibosons before requiring a \(b\)-tagged jet. The fraction of heavy-flavor (\(Wb\)) events is obtained using the ratio of the NLO cross sections for \(W+\)jets and \(Wbb\), as described in Ref. \([20]\). This normalization to data also accounts for smaller contributions such as \(Z+\)jets events in which one of the leptons from the \(Z\) boson decay is not reconstructed.

The \(tt\) background, consisting of the leptonic decay modes of the \(W\) boson from the top quark decay (\(t+\)jets and dilepton), is estimated using samples generated with ALPGEN normalized to the cross section: \(\sigma(tt) = 6.7 \pm 1.2\) pb \([21]\), where the uncertainty on the top quark mass is incorporated into the cross section uncertainty.

The parton-level samples are then processed with Pythia \([22]\) for hadronization, particle decays, and modeling of the underlying event. The generated events are processed through a Geant-based \([23]\) simulation of the DO detector. The resulting lepton and jet energies are smeared to reproduce the resolutions observed in data.

The background from jets misidentified as electrons or jets resulting in isolated muons is estimated using multijet data samples that pass all event selection cuts, but fail the requirement on muon isolation or electron quality \([14]\). This background is normalized using a data sample dominated by multijet events, selected by requiring \(E_T < 15\) GeV.

The overall acceptances, including trigger and selection efficiencies, for signal events with at least one \(b\)-tagged jet are \((2.7 \pm 0.2)\%\) in the s-channel and \((1.9 \pm 0.2)\%\) in the t-channel. The acceptance is calculated as the fraction of events that pass the selection over all possible single top quark decays, including all leptonic and hadronic decays of the \(W\) boson. Estimates for signal and background yields and the observed numbers of events after selection are shown in Table I.

**NEURAL NETWORKS ANALYSIS**

After event selection, several variables are combined in neural networks to discriminate the single top quark signals from the backgrounds. The networks are composed of three layers of nodes: input, hidden, and output. For training and testing, we use the MLPFIT \([24]\) package. Testing and training event sets are created from simulated signal and background samples. We use a technique called early stopping \([25]\) to determine the maximum number of epochs for training which prevents overtraining. Each network is then tuned by choosing the optimal number of hidden nodes. From studies based on optimizing the expected upper limits on the single top quark production cross sections, we find that the s-channel and t-channel searches each require only two networks, corresponding to the dominant backgrounds: \(Wbb\) and \(tt \rightarrow t+\)jets.

The list of discriminating variables has been chosen based on an analysis of Feynman diagrams of signals and backgrounds \([27]\) and on a study of single top quark production at NLO \([8]\). The input variables to each network are selected from this list by training with different combinations of variables and choosing the combination that produces the minimum testing error and largest signal-background separation. Table II shows the variables used for each signal-background pair. These variables fall into three categories: individual-object kinematics, global-event kinematics, and angular correlations.
Since the input variables do not depend on the lepton flavor, the electron and muon analyses utilize the same variables. However, owing to different lepton resolutions and pseudorapidity ranges, we construct separate networks for them. Therefore, four neural networks are used for the signal-background pairs ($tb-W\bar{b}$, $tb-\bar{t}t$, $tqb-W\bar{b}$, $tqb-\bar{t}t$) for each of the electron and muon channels.

Figure 2 shows distributions of four representative variables. We reconstruct the final state top quark from the reconstructed $W$ boson and a jet as follows. The $W$ boson is reconstructed from the isolated lepton and the missing transverse energy. The $z$-component of the neutrino momentum ($p^z_\nu$) is calculated using a $W$ boson mass constraint, choosing the solution with smaller $|p^z_\nu|$ from the two possible solutions. In the $s$-channel analysis, the top quark is reconstructed from the $W$ boson and the “best” jet. The best jet is defined as the jet in each event for which the invariant mass of the reconstructed $W$ boson and the jet system is closest to $m_t = 175$ GeV. In the $t$-channel analysis, the top quark is reconstructed from the $W$ boson and the leading $b$-tagged jet. Using these two methods we are able to correctly identify the $b$-quark jet from the top quark decay in about 90% of the signal events.

Figure 3 shows the outputs of the neural networks for the data and the expected backgrounds, as well as signals for the electron and muon channels combined. The neural network output in MLFFIT is around one for signal events and around zero for background events, but it is not constrained to the interval $[0,1]$. The $\bar{t}t$ networks separate signal and $\bar{t}t$ backgrounds efficiently. The $W\bar{b}$ networks are less efficient for the $W+$jets backgrounds because the event kinematics are similar between signal and background.

**SYSTEMATIC UNCERTAINTIES**

Systematic uncertainties are evaluated for the Monte Carlo signal and background samples, separately for the electron and muon channels and for each $b$-tag multiplicity. The most important sources of systematic uncertainty are listed in Table 3. The systematic uncertainty on the shapes of the distributions is also taken into account for the contributions from $b$-tag modeling, jet en-
FIG. 2: Comparison of signal, background, and data for the electron and muon channels combined, requiring at least one tag, for four representative neural network input variables. Shown are (a) the transverse momentum of the leading untagged jet, (b) the invariant mass of the reconstructed top quark using the leading tagged jet, (c) the invariant mass of the final state system, and (d) the pseudorapidity of the leading untagged jet multiplied by the charge of the lepton. Signals are multiplied by ten.

The observed data are consistent with the background predictions for all eight analysis channels. We therefore set upper limits on the single top quark production cross section separately in the s-channel and t-channel searches using a Bayesian approach [29]. In each search, two-dimensional histograms are constructed from the $Wb\bar{b}$ vs. $t\bar{t}$ neural network outputs. A likelihood is built from these histograms for signal, background, and data, as a product over all channels (electron and muon, single and double tags) and bins. We assume a Poisson distribution for the observed number of events in each bin and a flat prior probability for the signal cross section. The prior for the combined signal acceptance and background yields is a multivariate Gaussian with uncertainties and
correlations described by a covariance matrix. Finally, we compute the posterior probability density as a function of the production cross section.

The Bayesian posterior probability densities are shown in Fig. 4 for both the s-channel and t-channel searches. The corresponding upper limits at the 95% C.L. are 6.4 pb in the s-channel and 5.0 pb in the t-channel. The sensitivity of these measurements is given by the expected upper limits obtained by setting the observed number of events to the background prediction in each bin. The expected upper limits are 4.5 pb in the s-channel search and 5.8 pb in the t-channel search.

CONCLUSIONS

No evidence is found for electroweak production of single top quarks in $230 \text{ pb}^{-1}$ of data collected with the DØ detector at $\sqrt{s} = 1.96 \text{ TeV}$. The data consist of events in the electron and muon final states with at least one $b$-tagged jet. We build binned likelihoods from the output of neural networks to set upper limits at the 95% C.L. The measured s-channel limit is 6.4 pb and the measured t-channel limit is 5.0 pb. These upper limits represent significant improvements over previously published results [2, 10, 11] due to the larger data set as well as the use of a multivariate analysis technique together with shape information from the resulting output distributions. They approach the region of sensitivity for models of physics beyond the standard model, such as a fourth quark-generation scenario or flavor-changing neutral-currents [3].
FIG. 4: The Bayesian posterior probability density as a function of the single top quark cross section for the s-channel and t-channel searches.

Acknowledgements

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); MSMT (Czech Republic); CRC Program, CFI, NSERC (Canada); BMBF and DFG (Germany); SFI (Ireland); Research Corporation, Alexander von Humboldt Foundation, and the Marie Curie Program.

[*] Visitor from University of Zurich, Zurich, Switzerland.

[1] CDF Collaboration, F. Abe et al., Phys. Rev. Lett. 74, 2626 (1995); DØ Collaboration, S. Abachi et al., Phys. Rev. Lett. 74, 2632 (1995).
[2] D. Chakraborty, J. Konigsberg, and D. L. Rainwater, Ann. Rev. Nucl. Part. Sci. 53, 301 (2003).
[3] T. Tait and C.-P. Yuan, Phys. Rev. D 63, 014018 (2001).
[4] B. W. Harris et al., Phys. Rev. D 66, 054024 (2002).
[5] Z. Sullivan, Phys. Rev. D 70, 114012 (2004).
[6] J. Campbell, R. K. Ellis, and F. Tramontano, Phys. Rev. D 70, 094012 (2004).
[7] Q.-H. Cao, R. Schwienhorst, and C.-P. Yuan, Phys. Rev. D 71, 054023 (2005).
[8] Q.-H. Cao, R. Schwienhorst, J. Benitez, R. Brock, and C.-P. Yuan, submitted to Phys. Rev. D, hep-ph/0504230 (2005).
[9] DØ Collaboration, B. Abbott et al., Phys. Rev. D 63, 031101 (2001); DØ Collaboration, V. M. Abazov et al., Phys. Lett. B 517, 282 (2001).
[10] CDF Collaboration, D. Acosta et al., Phys. Rev. D 65, 091102 (2002).
[11] CDF Collaboration, D. Acosta et al., Phys. Rev. D 71, 012005 (2005).
[12] DØ Collaboration, V. M. Abazov et al., hep-ex/0507191.
[13] Pseudorapidity is defined as $\eta = -\ln\tan\frac{\theta}{2}$, where $\theta$ is the polar angle with the origin at the primary vertex. Detector fiducial regions are defined by detector pseudorapidity $\eta_{bd}$, which is calculated with the origin at the nominal center of the detector ($z = 0$).
[14] DØ Collaboration, V. M. Abazov et al., hep-ex/0504043.
[15] Jets are defined using the iterative, seed-based cone algorithm with radius $R = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2}$ = 0.5, including midpoints as described on pp. 47–77 in G. C. Blazey et al., in Proceedings of the Workshop on QCD and Weak Boson Physics in Run II, edited by U. Baur, R. K. Ellis, and D. Zeppenfeld, FERMILAB-PUB-00-297 (2000).
[16] DØ Collaboration, V. M. Abazov et al., hep-ex/0504058.

[17] CompHEP Collaboration, E. Boos et al., Nucl. Instrum. Meth. A 534, 250 (2004).
[18] M. L. Mangano et al., JHEP 0307, 001 (2003).
[19] J. M. Campbell and R. K. Ellis, Phys. Rev. D 60, 113006 (1999).
[20] DØ Collaboration, V. M. Abazov et al., Phys. Rev. Lett. 93, 141801 (2004).
[21] R. Bonciani et al., Nucl. Phys. B 529, 424 (1998); M. Cacciari et al., JHEP 0404, 068 (2004); N. Kidonakis and R. Vogt, Phys. Rev. D 68, 114014 (2003).
[22] T. Sjöstrand et al., Comput. Phys. Commun. 155, 238 (2001).
[23] R. Brun et al., CERN Program Library Long Writeup W 5013 (1994).
[24] J. Schwindling, http://schwind.home.cern.ch/schwind/MLPfit.html.
[25] G. Orr and K. Müller, “Neural Networks: Tricks of the Trade,” Springer-Verlag, Berlin, p. 55 (1998).
[26] G. Mahlon and S. Parke, Phys. Rev. D 53, 4886 (1996); S. Parke and Y. Shadmi, Phys. Lett. B 387, 199 (1996).
[27] E. Boos and L. Dudko, Nucl. Instrum. Methods A 502, 486 (2003).
[28] G. L. Kane and C. P. Yuan, Phys. Rev. D 40, 2231 (1989).
[29] I. Bertram et al., FERMILAB-TM-2104 (2000).