Observation of single top-quark production


(The DØ Collaboration)
We report observation of the electroweak production of single top quarks in \( p \bar{p} \) collisions at \( \sqrt{s} = 1.96 \) TeV based on 2.3 fb\(^{-1}\) of data collected by the D0 detector at the Fermilab Tevatron Collider. Using events containing an isolated electron or muon and missing transverse energy, together with jets originating from the fragmentation of \( b \) quarks, we measure a cross section of \( \sigma(p \bar{p} \rightarrow tb + X, tqb + X) = 3.94 \pm 0.88 \) pb. The probability to measure a cross section at this value or higher in the absence of signal is \( 2.5 \times 10^{-7} \), corresponding to a 5.0 standard deviation significance for the observation.

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

At hadron colliders, top quarks can be produced in pairs via the strong interaction or singly via the electroweak interaction \([1]\). Top quarks were first observed via pair production at the Fermilab Tevatron Collider in 1995 \([2]\). Since then, pair production has been used to make precise measurements of several top quark properties, including the top quark mass \([3]\). Single top quark production, on the other hand, serves as a probe of the \( Wtb \) interaction \([4]\), and its production cross section provides a direct measurement of the magnitude
of the quark mixing matrix element $V_{tb}$ without assuming three quark generations \cite{5}. However, measuring the yield of single top quarks is difficult because of the small production rate and large backgrounds.

In 2007, we presented the first evidence for single top quark production and the first direct measurement of $|V_{tb}|$ \cite{6,7} using 0.9 fb$^{-1}$ of Tevatron data at a center-of-mass energy of 1.96 TeV. Recently, the CDF collaboration has also presented such evidence in 2.2 fb$^{-1}$ of data \cite{8}. This Letter describes the observation of a single top quark signal in 2.3 fb$^{-1}$ of Tevatron data. The CDF collaboration is also reporting observation of single top quark production \cite{9}.

When top quarks are produced singly, they are accompanied by a bottom quark in the $s$-channel production mode \cite{10} or by both a bottom quark and a light quark in $t$-channel production \cite{11}, as illustrated in Fig. 1. We search for both of these processes at once. The sum of their predicted cross sections is 3.46 ± 0.18 pb \cite{12} for a top quark mass $m_t = 170$ GeV, at which this analysis is performed. We refer to the $s$-channel process as “tb” production, where $tb$ includes both $t$ and $b$ states. The $t$-channel process is abbreviated as “$tqb$,” where this includes $tqb, tqb, \bar{t}qb,$ and $\bar{t}qb$ states.

![FIG. 1: Representative Feynman diagrams for (a) $s$-channel single top quark production and (b) $t$-channel production, showing the top quark decays of interest.](image)

The analysis presented in this Letter is an improved version of the one from 2007 \cite{6,7}, with a larger dataset. Most definitions and abbreviations used here are explained in detail in Ref. \cite{8}. The measurement focuses on the final state containing one high transverse momentum ($p_T$) lepton ($\ell =$ electron or muon) not near a jet (“isolated”), large missing transverse energy ($E_T^\text{miss}$) indicative of the passage of a neutrino $\nu$, a $b$-quark jet from the decay of the top quark ($t\rightarrow Wb\rightarrow\ell\nu b$), and possibly another $b$ jet and a light jet as indicated above. The data were collected with the D0 detector \cite{13} using a logical OR of many trigger conditions in place of only the single-lepton plus jets triggers used previously. Several offline selection criteria, including $b$-jet identification requirements for double-tagged events, have been loosened. These improvements have increased the signal acceptance by 18%. The backgrounds are $W$ bosons produced in association with jets, top quark pair ($tt$) production with decay into the lepton+jets and dilepton final states (when a jet or a lepton is not reconstructed), and multijet production, where a jet is misreconstructed as an electron or a heavy-flavor quark decays to a muon that passes isolation criteria. $Z$+jets and diboson processes form minor additional background components.

We consider events with two, three, or four jets (which allows for additional jets from initial-state and final-state radiation), reconstructed using a cone algorithm in $(y, \phi)$ space, where $y$ is the rapidity and $\phi$ is the azimuthal angle, and the cone radius is 0.5 \cite{7}. The highest-$p_T$ (leading) jet must have $p_T > 25$ GeV, and subsequent jets have $p_T > 15$ GeV; all jets have pseudorapidity $|\eta| < 3.4$. We require $20 < E_T < 200$ GeV for events with two jets and $25 < E_T < 200$ GeV for events with three or four jets. Events must contain only one isolated electron with $p_T > 15$ GeV and $|\eta| < 1.1$ ($p_T > 20$ GeV for three- or four-jet events), or one isolated muon with $p_T > 15$ GeV and $|\eta| < 2.0$. The background from multijets events is kept to $\approx$5% by requiring high total transverse energy and by demanding that the $E_T$ is not along the direction of the lepton or the leading jet. To enhance the signal fraction, one or two of the jets are required to originate from long-lived $b$ hadrons. We achieve this goal by using a neural network (NN) $b$-jet tagging algorithm \cite{14}. The variables used to identify such jets rely on the characteristics of a secondary vertex and tracks with large impact parameters. After $b$-jet identification, we require the leading $b$-tagged jet to have $p_T > 20$ GeV. To further improve the sensitivity, we split the data by lepton flavor, number of jets and $b$-tagged jets, and data collection period.

We model the signal using the COMPHEP-based next-to-leading order (NLO) Monte Carlo (MC) event generator SINGLETOP \cite{15}. The decays of the top quark and resulting $W$ boson, both with standard model (SM) widths, are modeled in SINGLETOP to preserve spin information. PYTHIA \cite{16} is used to model the hadronization of generated partons. We assume the SM prediction for the ratio of the $tb$ and $tqb$ cross sections \cite{12}.

The $tt$, $W+$jets, and $Z+$jets backgrounds are simulated using the ALPGEN leading-log MC event generator \cite{17} and PYTHIA to model hadronization. The $tt$ background is normalized to the predicted cross section \cite{18}. The diboson backgrounds are modeled using PYTHIA. In the simulation of the $W+$jets backgrounds, we scale the ALPGEN cross sections for events with heavy flavor jets by factors derived from calculations of NLO effects \cite{14}: $Wbb$ and $Wc\bar{c}$ are scaled by 1.47, and $Wcj$ by 1.38.

All MC events are passed through a GENIE-based simulation of the D0 detector and are reconstructed using the same software as for the data. Data events from random beam crossings are overlaid on the simulation to better model the effects of detector noise and multiple $p\bar{p}$ interactions. Small differences between data and simulation in the lepton and jet reconstruction efficiencies and resolutions are corrected in the simulation as
measured from separate data samples. We also correct the \( \eta(\text{jets}) \), \( \Delta \phi(\text{jet1}, \text{jet2}) \), and \( \Delta \eta(\text{jet1}, \text{jet2}) \) distributions in the \( W+\text{jets} \) samples to match data.

The multijets background is modeled using independent data samples containing leptons that are not isolated. The multijets background, combined with the background from \( W+\text{jets} \), is normalized to the lepton+jets data with other backgrounds subtracted, using the \( p_T(\ell) \), \( \Delta \eta \), and the \( W \) boson transverse mass distributions before \( b \)-jet identification is applied.

The \( b \)-tagging algorithm is modeled in simulated events by applying weights (“tag-rate functions”) measured from data that account for the probability for each jet to be tagged as a function of jet flavor, \( p_T \), and \( \eta \). After \( b \) tagging, an empirical correction of 0.95 ± 0.13 for the \( Wb \) and \( Wc\bar{c} \) fractions is derived from the \( b \)-tagged and not-\( b \)-tagged two-jet data and simulated samples.

The above selections give 4,519 \( b \)-tagged lepton+jets events, which are expected to contain 223 ± 30 single top quark events. Table I shows the event yields, separated into \( \ell^+\ell^- \) and single- \( \ell \) analysis channels combined. The uncertainties include both statistical and systematic components.

**Table I: Number of expected and observed events in 2.3 \( fb^{-1} \) for each channel.**

<table>
<thead>
<tr>
<th>Source</th>
<th>2 jets (( \pm 1 ))</th>
<th>3 jets (( \pm 2 ))</th>
<th>4 jets (( \pm 2 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( tb+tbq ) signal</td>
<td>139 ± 18</td>
<td>63 ± 10</td>
<td>21 ± 5</td>
</tr>
<tr>
<td>( W+\text{jets} )</td>
<td>1,829 ± 161</td>
<td>637 ± 61</td>
<td>180 ± 18</td>
</tr>
<tr>
<td>( Z+\text{jets and dibosons} )</td>
<td>229 ± 38</td>
<td>85 ± 17</td>
<td>26 ± 7</td>
</tr>
<tr>
<td>( t\bar{t} )</td>
<td>222 ± 35</td>
<td>436 ± 66</td>
<td>484 ± 71</td>
</tr>
<tr>
<td>Multijets</td>
<td>196 ± 50</td>
<td>73 ± 17</td>
<td>30 ± 6</td>
</tr>
<tr>
<td>Total prediction</td>
<td>2,615 ± 192</td>
<td>1,294 ± 107</td>
<td>742 ± 80</td>
</tr>
<tr>
<td>Data</td>
<td>2,579</td>
<td>1,216</td>
<td>724</td>
</tr>
</tbody>
</table>

Systematic uncertainties arise from each correction factor or function applied to the background and signal models. Most affect only the normalization, but three corrections modify in addition the shapes of the distributions: these are the jet energy scale corrections, the tag-rate functions, and the reweighting of the distributions in \( W+\text{jets} \) events. The largest uncertainties come from the jet energy scale (the normalization part is (1.1–13.1)% for signal and (0.1–2.1)% for background), the tag-rate functions (the normalization part is (2.1–7.0)% for single-tagged events and (9.0–11.4)% for double-tagged events), and the correction for jet-flavor composition in \( W+\text{jets} \) events (13.7%), with smaller contributions from the integrated luminosity (6.1%), jet energy resolution (4.0%), initial-state and final-state radiation (0.6–12.6%), \( b \)-jet fragmentation (2.0%), \( t\bar{t} \) cross section (12.7%), and lepton efficiency corrections (2.5%). All other contributions have a smaller effect. The values given are the relative uncertainties on the individual sources. The total uncertainty on the background is (8–16)% depending on the analysis channel.

After event selection, we expect single top quark events to constitute (3–9)% of the data sample. Since the uncertainty on the background is larger than the expected signal, we improve discrimination by using multivariate analysis techniques. We have developed three independent analyses based on boosted decision trees (BDT) [20], Bayesian neural networks (BNN) [21], and the matrix element (ME) method [22]. Our application of these techniques to D0’s single top quark searches is described in Refs. 4 and 5. The analyses presented in this Letter differ from previous implementations in the choice of input variables and some detailed tuning of each technique.

The BDT analysis has re-optimized the input variables into a common set of 64 variables for all analysis channels. The variables fall into five categories, single-object kinematics, global event kinematics, jet reconstruction, top quark reconstruction, and angular correlations. Separate sets of trees are created with these variables for each channel. The BNN analysis uses the RuleFitJF algorithm to select the most sensitive of these variables, then combines 18–28 of them into a single separate discriminant for each channel. The ME analysis uses only two- and three-jet events, divided into a \( W+\text{jets} \)-dominated set and a \( t\bar{t} \)-dominated set. It includes matrix elements for more background sources, adding \( t\bar{t} \), \( WW \), \( WZ \), and \( ggg \) diagrams in the two-jet and \( W+\text{jets} \) in the three-jet bin, to improve background rejection.

Each analysis uses the same data and background model and has the same sources of systematic uncertainty. We test the analyses using ensembles of pseudo-datasets created from background and signal at different cross sections to confirm linear behavior and thus an unbiased cross section measurement. The analyses are also checked extensively before \( b \)-tagging is applied, and using two control regions of the data, one dominated by \( W+\text{jets} \) and the other by \( t\bar{t} \) backgrounds, as shown in Fig. 2. These studies confirm that backgrounds are well modeled across the full range of the discriminant output.

The cross section is determined using the same Bayesian approach as in our previous studies. This involves forming a binned likelihood as a product over all bins and channels, evaluated separately for each multivariate discriminant, with no cuts applied to the outputs. The central value of the cross section is defined by the position of the peak in the posterior density, and the 68% interval about the peak is taken as the uncertainty on the measurement. Systematic uncertainties, including all correlations, are reflected in this posterior interval.

We extract inclusive single top quark cross sections \( \sigma(p\bar{p} \rightarrow t\bar{t}+X, tbq+X) \) of \( \sigma_{\text{BDT}} = 3.74^{+0.79}_{-0.79} \text{pb} \), \( \sigma_{\text{BNN}} = 3.74^{+0.79}_{-0.79} \text{pb} \), and \( \sigma_{\text{ME}} = 3.74^{+0.79}_{-0.79} \text{pb} \).
FIG. 2: The combination discriminant outputs for (a) $W$+jets and (b) $t\bar{t}$ cross-check samples. $H_T$ is the scalar sum of the transverse momenta of the final state objects (lepton, $E_T$, and jets).

$4.70^{+1.18}_{-0.93}$ pb, and $\sigma_{ME} = 4.30^{+0.99}_{-1.20}$ pb. The sensitivity of the analyses to a contribution from single top quark production is estimated by generating an ensemble of pseudodatasets that sample the background model and its uncertainties, with no signal present. We measure a cross section from each pseudodataset, and hence obtain the probability that the SM cross section is reached. This provides expected sensitivities (stated in terms of Gaussian standard deviations, SD) of 4.3, 4.1, and 4.1 SD for the BDT, BNN, and ME analyses respectively. The measured significances, obtained by counting the number of pseudodatasets with cross sections at least as large as the measured cross section, are 4.6, 5.2, and 4.9 SD respectively.

The three multivariate techniques use the same data sample but are not completely correlated: the correlation of the measured cross section using pseudodatasets with background and SM signal is BDT:ME = 74%, BDT:ME = 60%, BNN:ME = 57%. Their combination therefore leads to increased sensitivity and a more precise measurement of the cross section. We use the three discriminant outputs as inputs to a second set of Bayesian neural networks, and obtain the combined cross section and its signal significance from the new discriminant output. The resulting expected significance is 4.5 SD. Figure 3 illustrates the importance of the signal when comparing data to prediction.

The measured cross section is

$$\sigma(p\bar{p} \rightarrow tb + X, tqb + X) = 3.94 \pm 0.88 \text{ pb}.$$ 

The measurement has a $p$-value of $2.5 \times 10^{-7}$, corresponding to a significance of 5.0 SD. The expected and measured posterior densities and the background-only pseudodataset measurements are shown in Fig. 4.

We use the cross section measurement to determine the Bayesian posterior for $|V_{tb}|^2$ in the interval [0,1] and extract a limit of $|V_{tb}| > 0.78$ at 95% C.L. within the SM. When the upper constraint is removed, we measure $|V_{tb}f^2| = 1.07 \pm 0.12$, where $f^2$ is the strength of the left-handed $Wtb$ coupling.

In summary, we have measured the single top quark production cross section using 2.3 fb$^{-1}$ of data at the D0 experiment. We measure a cross section for the combined $tb$-$t\bar{t}$ channels of $3.94 \pm 0.88$ pb. Our result provides an improved direct measurement of the amplitude of the CKM matrix element $V_{tb}$. The measured single top quark signal corresponds to an excess over the predicted background with a significance of 5.0 SD — observation of single top quark production.

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); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); STFC (United Kingdom); MSMT and GACR (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); CAS and CNSF (China); and the Alexander von Humboldt Foundation (Germany).

[a] Visitor from Augustana College, Sioux Falls, SD, USA.
[b] Visitor from Rutgers University, Piscataway, NJ, USA.
[c] Visitor from The University of Liverpool, Liverpool, UK.
[d] Visitor from II. Physikalisches Institut, Georg-August University, Göttingen, Germany.
[e] Visitor from Centro de Investigacion en Computacion - IPN, Mexico City, Mexico.
[f] Visitor from Helsinki Institute of Physics, Helsinki, Finland.
[i] Visitor from Universitàt Bern, Bern, Switzerland.
[j] Visitor from Universitàt Zürich, Zürich, Switzerland.

‡ Deceased.

[12] N. Kidonakis, Phys. Rev. D 74, 114012 (2006). The cross sections for the single top quark processes are $1.12 \pm 0.05$ pb ($s$-channel) and $2.34 \pm 0.13$ pb ($t$-channel). There is a third process, $p\bar{p} \rightarrow t\bar{t}+X$, that we do not include in our measurement since the production rate is below the sensitivity of the data.
[18] N. Kidonakis and R. Vogt, Phys. Rev. D 68, 114014 (2003). We used $\sigma(p\bar{p} \rightarrow t\bar{t}+X) = 7.91^{+0.61}_{-1.01}$ pb, including an uncertainty component for the top quark mass.