Model-independent measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 1.96$ TeV


(The D0 Collaboration*)

1Universidad de Buenos Aires, Buenos Aires, Argentina
2LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil
3Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
4Universidade Federal do ABC, Santo André, Brazil
5Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil
6Simon Fraser University, Vancouver, British Columbia, and York University, Toronto, Ontario, Canada
7University of Science and Technology of China, Hefei, People’s Republic of China
8Universidad de los Andes, Bogotá, Colombia
9Charles University, Faculty of Mathematics and Physics, Center for Particle Physics, Prague, Czech Republic
10Czech Technical University in Prague, Prague, Czech Republic
11Center for Particle Physics, Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic
12Universidad San Francisco de Quito, Quito, Ecuador
13LPC, Université Blaise Pascal, CNRS/IN2P3, Clermont, France
14LPSC, Université Joseph Fourier Grenoble I, CNRS/IN2P3, Institut National Polytechnique de Grenoble, Grenoble, France
15CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
16LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France
17LPNHE, Universités Paris VI and VII, CNRS/IN2P3, Paris, France
18CEA, Ifèr4, SPP, Saclay, France
19IPHC, Université de Strasbourg, CNRS/IN2P3, Strasbourg, France
20IPNL, Université Lyon 1, CNRS/IN2P3, Villeurbanne, France and Université de Lyon, Lyon, France
21III. Physikalisches Institut A, RWTH Aachen University, Aachen, Germany
22Physikalisches Institut, Universität Freiburg, Freiburg, Germany
23II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany
24Institut für Physik, Universität Mainz, Mainz, Germany
25Ludwig-Maximilians-Universität München, München, Germany
26Fachbereich Physik, Bergische Universität Wuppertal, Wuppertal, Germany
27Panjab University, Chandigarh, India
28Delhi University, Delhi, India
29Tata Institute of Fundamental Research, Mumbai, India
30University College Dublin, Dublin, Ireland
31Korea Detector Laboratory, Korea University, Seoul, Korea
32CINVESTAV, Mezco City, Mexico
33FOM-Institute NIKHEF and University of Amsterdam/NIKHEF, Amsterdam, The Netherlands
34Radboud University Nijmegen/NIKHEF, Nijmegen, The Netherlands
35Joint Institute for Nuclear Research, Dubna, Russia
36Institute for Theoretical and Experimental Physics, Moscow, Russia
37Moscow State University, Moscow, Russia
38Institute for High Energy Physics, Protvino, Russia
39Petersburg Nuclear Physics Institute, St. Petersburg, Russia
40Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Física d’Altes Energies (IFAE), Barcelona, Spain
41Stockholm University, Stockholm and Uppsala University, Uppsala, Sweden
42Lancaster University, Lancaster LA1 4YB, United Kingdom
43Imperial College London, London SW7 2AZ, United Kingdom
44The University of Manchester, Manchester M13 9PL, United Kingdom
45University of Arizona, Tucson, Arizona 85721, USA
We present a model-independent measurement of $t$-channel electroweak production of single top quarks in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. Using 5.4 fb$^{-1}$ of integrated luminosity collected by the D0 detector at the Fermilab Tevatron Collider, and selecting events containing an isolated electron or muon, missing transverse energy and one or two jets originating from the fragmentation of $b$ quarks, we measure a cross section $\sigma(p\bar{p} \rightarrow tqb + X) = 2.90 \pm 0.59$ (stat $+$ syst) pb for a top quark mass of 172.5 GeV. The probability of the background to fluctuate and produce a signal as large as the one observed is $1.6 \times 10^{-8}$, corresponding to a significance of 5.5 standard deviations.

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

Within the standard model (SM), top quarks are produced through two different types of interactions at hadron colliders. The largest yield is from $t\bar{t}$ pairs, produced via the strong interaction [1]. Pair production was used by the D0 and CDF collaborations to establish the existence of the top quark in 1995 [2, 3]. The second source of top quarks is the electroweak production of single top quarks, observed by the D0 and CDF collaborations in 2009 [4, 5]. The single top quark production rate is about half that for $t\bar{t}$ pairs and the signal-to-background ratio is worse, so the observation of this process is more difficult.

At the Tevatron, single top quarks are produced through $W$ boson exchange, and accompanied by a $b$ quark in the $s$ channel ($tb + \bar{t}b = tb$) [6], or by both a $b$ and a light quark in the $t$-channel ($tqb + \bar{t}qb = tqb$) [7, 8]. A third process, usually called "associated production," in which the top quark is produced together with a $W$ boson, has negligible cross section at the Tevatron [9]. Prior D0 and CDF publications [4, 5]...
measured the production cross section for the sum of $tb$ and $tqb$ while assuming the relative rate between the two processes as predicted by the SM. This relative rate, however, departs from the SM prediction in several new physics scenarios, for example, in models with additional quark generations [10], new heavy bosons [11], flavor-changing neutral currents (FCNC) [12], or anomalous top quark couplings [13–15]. It is therefore of interest to measure the $tb$ and $tqb$ production cross sections separately, using a methodology that is independent of theoretical assumptions about their relative rates [11]. Furthermore, measuring the production rates of the two single top quark production modes to a precision comparable to that of the $t\bar{t}$ pair cross section not only tests the validity of the parton distribution functions of the proton and antiproton, but also the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements [16] and higher-order corrections from quantum chromodynamics. These high-precision comparisons are important to check the self-consistency of the standard model (SM).

This Letter builds on the techniques developed by the D0 collaboration for prior measurements of single top quark production [4, 17–19], and analyzes a larger dataset corresponding to an integrated luminosity of 5.4 fb$^{-1}$. Using improved multivariate analysis (MVA) techniques, we report a new measurement of the $tqb$ production rate that makes no assumptions about the $tb$ production rate, and is therefore sensitive to various new physics scenarios.

The data for this analysis were collected with a logical OR of many trigger conditions that results in a fully efficient trigger selection for the single top quark signal. Electron, muon, missing transverse energy ($E_T$), and jet reconstruction algorithms used in this analysis have been described elsewhere [18, 20]. Events are selected that contain one jet with transverse momentum $p_T > 25$ GeV and at least a second jet with $p_T > 15$ GeV, both within pseudorapidity $|\eta| < 3.4$. Events with 2, 3 or 4 jets are used in the analysis. Events are also required to contain exactly one isolated high-$p_T$ electron or muon that originates from the $p\bar{p}$ interaction vertex and satisfies the following acceptance criteria: for the electron $|\eta| < 1.1$ and $p_T > 15(20)$ GeV for events with 2 (3 or 4) jets; for the muon $|\eta| < 2.0$ and $p_T > 15$ GeV. The $E_T$ is required to be in the range of $(20, 200)$ GeV for events with 2 jets and $(25, 200)$ GeV for events with 3 or 4 jets.

The main background contributions arise from $W$ bosons produced in association with jets, $t\bar{t}$ pairs, and multijet production in which a jet with high electromagnetic content mimics an electron, or a muon contained within a jet originating from the decay of a heavy-flavor quark ($b$ or $c$ quark) appears isolated. Diboson ($WW$, $WZ$, $ZZ$) and $Z$+jets processes add small additional contributions to the background. To limit the number of events originating from multijet production that enter our candidate samples through $E_T$ mismeasurements, we require the $E_T$ to be not aligned with or opposing the direction of the lepton or the leading jet and require high total scalar transverse energy of the lepton, jets and $E_T$ ($H_T$). Additional requirements are placed on the transverse energy of all jets in all events and the second leading jet $p_T$ in two-jet events to supress soft-scattering background.

To enhance the fraction of signal events in data, one or two of the jets in each event are required to originate from the hadronization of long-lived $b$ hadrons, as inferred from a neural network $b$-tagging algorithm ("$b$-tagged jet") [21]. The data are divided into six mutually-exclusive subsamples to take advantage of different signal-to-background ratios and different sources of background. This separation is based on jet multiplicity (2, 3, or 4 jets), and on the number of $b$-tagged jets (1 or 2 $b$-tags). The signal acceptance of all selections corresponds to $(2.0 \pm 0.3)\%$ for $tqb$ and $(2.9 \pm 0.4)\%$ for $tb$ production, as estimated using Monte Carlo (MC) simulations.

Single top quark signal events are modeled using the COMPHEP-based MC event generator SINGLETOP [22]. The kinematics of the generated events closely match those predicted by next-to-leading-order (NLO) calculations [23]. PYTHIA [24] is used to model the hadronization of all generated partons. The $t\bar{t}$, $W$+jets, and $Z$+jets backgrounds are simulated using the ALPGEN leading-log MC event generator [25], with PYTHIA used to model hadronization. The $t\bar{t}$ background is normalized to the predicted cross section [1] for a top quark mass of 172.5 GeV. The normalization of the $W/Z$ backgrounds is obtained by scaling the ALPGEN cross sections to account for NLO effects [26]. Diboson backgrounds are modeled using PYTHIA.

The presence of additional $p\bar{p}$ interactions is modeled by events selected from random beam crossings matching the instantaneous luminosity profile in the data. All MC events are passed through a GEANT-based simulation [27] of the D0 detector. Small corrections, derived using samples depleted of single top quark events, are applied to all reconstructed objects to improve the agreement between data and simulation. In particular, any mismodeling of the pseudorapidity of jets or of the angular separation between the two leading jets in the $W$+jets sample is adjusted by reweighting the MC.

The multijet background is modeled using data containing non-isolated leptons. In the electron channel, the transverse momentum distribution of the lepton is reweighted to match the distribution in background events that pass the candidate-selection criteria. To increase the statistics in the muon channel, the jet closest to the muon is removed and the $E_T$ recalculated. The overall normalization of the $W$+jets and the multijet backgrounds is obtained by comparing the background expectation to data in three sensitive variables: lepton transverse momentum, $E_T$, and the transverse mass
of the $W$ boson. This normalization is found after subtracting the contributions from the small backgrounds ($tt$, $Z$+jets, and dibosons), before $b$-tagging, when the expected signal-to-background ratio (S:B) for the entire sample is S:B $\approx 1:279$.

The probability for the $b$-tagging algorithm to identify a jet as originating from a $b$ quark is measured in a sample of data with jets containing muons and compared to the corresponding efficiency measured in a MC sample. From this comparison we derive correction factors that we parametrize as a function of jet flavor, $p_T$, and $\eta$ and apply to all the MC samples used in this analysis. After $b$-tagging, we check the normalization of the $Wb\bar{b}$ and $Wc\bar{c}$ samples in a sample that has no overlap with the one used in the single top quark cross section measurement and find the normalization to be consistent with unity. We assign a 12% uncertainty on the normalization of the $Wb\bar{b}$ and $Wc\bar{c}$ samples to account for variations in the $Wcj$ ($j = u, d, s$) cross section and the $Wb\bar{b}$ to $Wc\bar{c}$ cross section ratio used in this study.

We consider systematic uncertainties that arise from the theoretical and detector modeling of the background processes and from the subsequent corrections applied to that model. Most uncertainties affect only the normalization of the distributions, but the corrections for the jet energy scale (JES), the jet energy resolution (JER), and the reweighting of the kinematic distributions in $W$+jets events also modify the differential distributions. The largest uncertainties come from the JER, corrections to the $b$-tagging efficiency, and the corrections for the jet-flavor composition in $W$+jets events, with smaller contribution from JES, MC statistics, integrated luminosity, and trigger uncertainties. The total systematic uncertainty on the background is 11%.

We select 8,471 events and expect $238 \pm 28$ $tb$ events. This corresponds to a S:B of approximately 1:33 for the sample with 1 $b$-tag and 1:50 for the sample with 2 $b$-tags. Table I lists the number of events expected and observed for each process as a function of jet multiplicity. We also define two control samples to check the background model components separately for the two main backgrounds: $W$+jets and $tt$. The $W$+jets-dominated sample has low $H_T$, exactly two jets, and only one $b$-tagged jet. The $tt$-dominated sample has high $H_T$, exactly four jets, and one or two $b$-tagged jets. We find good agreement for both normalization and shape in most of distributions studied for the cross-check samples and also for the candidate sample before $b$-tagging, when the signal contribution is negligible.

We further improve the discrimination between signal and background by employing MVA techniques as described in [18]. We use two of our previous methods: boosted decision trees (BDT) [28] and Bayesian neural networks (BNN) [29]. In addition, we use a third method, neuroevolution of augmented topologies (NEAT) [30]. All three methods use the same data and background model, consider the same sources of systematic uncertainty, and rely on variables with signal-background discrimination power chosen from a common set of well modeled variables from five categories [31, 32]: single object kinematics, global event kinematics, jet reconstruction, top quark reconstruction, and angular correlations. The BDT method uses a common set of 50 variables for all analysis channels, while the NEAT method uses the TMVA implementation of the “RuleFit” algorithm [33, 34] to select the 30 most sensitive kinematic variables in each channel. New for this analysis, the BNN method constructs a discriminant using the lepton and jet 4-vectors, the $E_T$ 2-vector, and variables that include lepton charge and $b$-tagging information, for a total of 14, 18 and 22 variables for events with 2, 3 and 4 jets.

<table>
<thead>
<tr>
<th>Source</th>
<th>2 jets</th>
<th>3 jets</th>
<th>4 jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tb$</td>
<td>104 ± 6</td>
<td>44 ± 7.8</td>
<td>13 ± 3.5</td>
</tr>
<tr>
<td>$tb$</td>
<td>140 ± 13</td>
<td>72 ± 9.4</td>
<td>26 ± 6.4</td>
</tr>
<tr>
<td>$tt$</td>
<td>433 ± 87</td>
<td>830 ± 133</td>
<td>860 ± 163</td>
</tr>
<tr>
<td>$W$+jets</td>
<td>3,560 ± 354</td>
<td>1,099 ± 169</td>
<td>284 ± 76</td>
</tr>
<tr>
<td>$Z$+jets &amp; dibosons</td>
<td>400 ± 55</td>
<td>142 ± 41</td>
<td>35 ± 18</td>
</tr>
<tr>
<td>Multijets</td>
<td>277 ± 34</td>
<td>130 ± 17</td>
<td>43 ± 5.2</td>
</tr>
<tr>
<td>Total prediction</td>
<td>4,914 ± 558</td>
<td>2,317 ± 377</td>
<td>1,261 ± 272</td>
</tr>
<tr>
<td>Data</td>
<td>4,881</td>
<td>2,307</td>
<td>1,283</td>
</tr>
</tbody>
</table>

Table I: Number of expected and observed events in 5.4 fb$^{-1}$ of integrated luminosity, with uncertainties including both statistical and systematic components. The $tb$ and $tqb$ contributions are normalized to their SM expectations.

![Figure 1](image-url) FIG. 1: Distribution for $t\bar{b}$ and $tqb$ processes for (a) transverse momentum of the leading $b$-tagged jet, (b) centrality, defined as the ratio of the scalar sum of the $p_T$ of the jets to the scalar sum of the energy of the jets in the event, (c) lepton charge multiplied by the leading non-$b$-tagged quark jet pseudorapidity and (d) total transverse energy. All the distributions are normalized to unity.
Each method is optimized to maximize the sensitivity to the $tqb$ signal by treating the $tb$ process as a background component with normalization given by the SM cross section [9]. Figure 1 shows the distributions of some of the variables that provide the greatest discrimination between the $tqb$ and $tb$ processes. To further improve the sensitivity, the analyses are performed separately on the six mutually exclusive subsamples defined before. For all three methods the output variables are saved in histograms with a binning chosen to ensure that there are enough events to limit the uncertainties due to MC statistics.

Even though the three MVA techniques use the same data sample, they are only $\approx 70\%$ correlated with each other. We therefore combine these methods using an additional BNN algorithm (BNNComb) that takes as input the individual output discriminants of the BDT, BNN, and NEAT methods, and produces a single combined output discriminant. Figure 2 shows comparisons between the $t$-channel signal, the background model, and data for the combined discriminant, which leads to a more precise measurement than any assumptions about the value of the $s$-channel cross section. Similarly, the $tb$ cross section is obtained by integrating over the $tqb$ axis. Ensembles of datasets generated at several different cross section values are used to verify the linearity of the cross section extraction procedure.

Figure 3 shows the 2D posterior probability density for the combined discriminant together with predictions from the SM [9] and various beyond-the-SM scenarios: four-quark-generations with CKM matrix element $|V_{ts}| = 0.2$ [10], top-flavor model with new heavy bosons at a scale $m_{\chi} = 1$ TeV [11], and FCNC with an up-quark/top-quark/ghon coupling $\kappa_u/\Lambda = 0.036$ [12].

![FIG. 2: Comparison of the signal and background models to data for the combined t-channel discriminant for (a) the entire discriminant range and (b) the signal region. The bins have been ordered by their expected S.B. The single top quark contributions are normalized to the measured cross sections. The t-channel contribution is visible above the hatched bands that show the uncertainty on the background prediction.](image)

![FIG. 3: Posterior probability density for $tqb$ vs $tb$ single top quark production in contours of equal probability density. The measured cross section and various theoretical predictions are also shown.](image)
production rate of $t\bar{b}$. We estimate the probability of the background to fluctuate and produce a signal as large as the one observed to be $1.6 \times 10^{-8}$, corresponding to a significance of 5.5 standard deviations (SD). The expected significance is 4.6 SD.

The presence of the $t$-channel signal is visible in Fig. 4, which shows comparisons between the data, the background model, and the $t$-channel signal for a sample selected with $S:B > 0.32$ using the combined discriminant. Four kinematic variables with large discriminating power for $t$-channel single top quark production are shown: $H_T$, the reconstructed top quark mass, the lepton charge multiplied by the pseudorapidity of the leading non-$b$-tagged jet, and the $t$-channel spin correlation in the optimal basis [32, 36], i.e., the cosine of the angle between the light quark jet and the lepton, both in the reconstructed top quark rest frame.

![Graphs showing top quark production](image)

**FIG. 4:** (a) Total transverse energy, (b) reconstructed top quark mass, (c) lepton charge multiplied by the leading non-$b$-tagged jet pseudorapidity, and (d) $t$-channel top quark spin correlation (see text) for a sample selected with $S:B > 0.32$ using BNNComb discriminant. The $t$-channel and $s$-channel contributions have been normalized to their measured cross sections.

The measured cross section depends on the assumed mass of the top quark, $m_t$. The dependence is studied by repeating the analysis on MC samples generated at different values of $m_t$. Table II summarizes the measured cross section for different top quark masses.

<table>
<thead>
<tr>
<th>$m_t$ [GeV]</th>
<th>170</th>
<th>172.5</th>
<th>175</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{b}$</td>
<td>2.80$^{+0.57}_{-0.61}$</td>
<td>2.90$^{+0.59}_{-0.59}$</td>
<td>2.53$^{+0.58}_{-0.57}$</td>
</tr>
<tr>
<td>$tb$</td>
<td>1.31$^{+0.77}_{-0.74}$</td>
<td>0.98$^{+0.62}_{-0.63}$</td>
<td>0.65$^{+0.51}_{-0.50}$</td>
</tr>
</tbody>
</table>

In summary, we have presented a model-independent measurement of the $t$-channel single top quark production cross section using 5.4 fb$^{-1}$ of integrated luminosity recorded by the D0 experiment at the Fermilab Tevatron accelerator. For a top quark mass of $m_t = 172.5$ GeV, the measured cross section is $\sigma(p\bar{p} \rightarrow t\bar{b} + X) = 2.90 \pm 0.59$ pb. This is the most precise measurement to date, and is in good agreement with SM expectations.

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 and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); CRC Program and NSERC (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China).

[9] N. Kidonakis, Phys. Rev. D 74, 114012 (2006). The cross sections for the single top quark processes ($m_t = 172.5$ GeV) are $1.04 \pm 0.04$ pb (s-channel) and $2.26 \pm 0.12$ pb (t-channel).