Observation of Electroweak Single Top-Quark Production

Aaltonen, Timo

2009


http://hdl.handle.net/10138/24478
https://doi.org/10.1103/PhysRevLett.103.092002

Downloaded from Helda, University of Helsinki institutional repository.
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.
First Observation of Electroweak Single Top Quark Production

We report the first observation of single top quark production using 3.2 fb$^{-1}$ of $par{p}$ collision data with $\sqrt{s} = 1.96$ TeV collected by the Collider Detector at Fermilab. The significance of the observed data is 5.0 standard deviations, and the expected sensitivity for standard model production and decay is in excess of 5.9 standard deviations. Assuming $m_t = 175$ GeV/$c^2$, we measure a cross section of $2.2^{+0.6}_{-0.5}$ (stat + syst) pb, extract the CKM matrix element value $|V_{tb}| = 0.91 \pm 0.01$ (stat + syst) $\pm 0.07$ (theory), and set the limit $|V_{tb}| > 0.71$ at the 95% C.L.

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

*Deceased
†With visitors from *University of Massachusetts Amherst, 
*University of Michigan, Ann Arbor, Michigan 48109
**University of Michigan, Ann Arbor, Michigan 48109
**Michigan State University, East Lansing, Michigan 48824
**University of Michigan, Ann Arbor, Michigan 48109
**Institute for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia
**University of New Mexico, Albuquerque, New Mexico 87131
**Northwestern University, Evanston, Illinois 60208
**The Ohio State University, Columbus, Ohio 43210
**Ohio University, Athens, Ohio 45701
**Ohio University, Athens, Ohio 45701
**University of Pennsylvania, Philadelphia, Pennsylvania 19104
**University of Pennsylvania, Philadelphia, Pennsylvania 19104
**Instituto Nazionale di Fisica Nucleare, Sezione di Padova-Trento, *University of Padova, I-35131 Padova, Italy
**LPNHE, Universite Pierre et Marie Curie/IN2P3-CNRS, UMR7585, Paris, F-75252 France
**University of Oxford, Oxford OX1 3RH, United Kingdom
**Istituto Nazionale di Fisica Nucleare Trieste/Udine, I-34100 Trieste, Italy
**University of Tsukuba, Tsukuba, Ibaraki 305, Japan
**Tufts University, Medford, Massachusetts 02155
**Waseda University, Tokyo 169, Japan
**Wayne State University, Detroit, Michigan 48201
**University of Wisconsin, Madison, Wisconsin 53706
**Yale University, New Haven, Connecticut 06520

(Dated: March 4, 2009)
In the standard model (SM), top quarks are expected to be produced singly in $p\bar{p}$ collisions through $s$-channel or $t$-channel exchange of a virtual $W$ boson [1]. The reasons for studying single top quarks are compelling: the production cross section is directly proportional to the square of the CKM matrix $|V_{tb}|$, and thus a measurement of the rate constrains fourth-generation models, models with flavor-changing neutral currents, and other new phenomena [3]. Electroweak production of single top quarks is a difficult process to measure because the expected production cross section is $\sim 2.9$ pb [4,5] is much smaller than those of competing background processes, and it is also smaller than the uncertainty on the total background rate. The presence of only one top quark in the event provides fewer features to use in separating the signal from background, compared with measurements of top pair production ($tt$), which was first observed in 1995 [6].

To overcome these challenges, a variety of multivariate techniques for separating single top events from the backgrounds have been developed. Using different combinations of techniques, both the CDF and D0 collaborations have published evidence for single top quark production at significance levels of 3.7 and 3.6 standard deviations, respectively [5, 8]. The analysis described in this Letter supersedes that of Ref. [7] and achieves a significantly improved sensitivity by including a larger data sample and by adding three new analyses. We report a signal significance of 5.0 standard deviations, thus conclusively observing electroweak production of single top quarks, and we make the most precise measurement of $|V_{tb}|$ to date.

We assume that single top quarks are produced in the $s$- and $t$-channel modes with the SM ratio, and that the branching ratio of the top quark to $Wb$ is 100%. We seek events in which the $W$ boson decays leptonically in order to improve the signal-to-background ratio $s/b$. We simulate single top events using the tree-level matrix-element generator MADEVENT [9]. The $t$-channel signal is modeled by the two processes $q\bar{b} \rightarrow q't$ and $qg \rightarrow q't\bar{b}$, which are combined to match the event kinematics predicted by a fully differential NLO calculation [9, 10].

A total of six analyses are combined to yield the final results reported here. The likelihood function (LF), matrix element (ME), and neural network (NN) analyses of [5] are re-used with an additional 1 fb$^{-1}$ of integrated luminosity; their methods remain unchanged. The three new analyses introduced here are: a boosted decision tree (BDT), a likelihood function optimized for $s$-channel single top production (LFS), and a neural-network-based analysis of events with missing transverse energy $E_T$ and jets (MJ). The BDT and LFS analyses use events that overlap with the LF, ME, and NN analyses, while the MJ analysis uses an orthogonal event selection that adds about 30% to the signal acceptance. This paper concentrates on the three new analyses and their combination with the analyses of [5] using 3.2 fb$^{-1}$ of integrated luminosity collected with the CDF II detector [12].

For the LF, ME, NN, BDT, and LFS analyses we select $\ell + E_T + $ jet events as described in [5], where $\ell$ is an explicitly reconstructed electron or muon from the $W$ boson decay and at least one jet is identified as containing a $B$ hadron. The background has contributions from events in which a $W$ boson is produced in association with one or more heavy-flavor jets ($W+HF$), events with mistakenly $b$-tagged light-flavor jets (mistags), multijet events (QCD), $tt$ and diboson processes, as well as $Z+\text{jet}$ events. The expected event yields in Table I are estimated as in [5] where the signal, $tt$, and diboson categories are Monte Carlo (MC) predictions scaled to the total integrated luminosity while the remaining categories use predictions derived from data control samples. The uncertainties quoted in Table I include theoretical uncertainties, the luminosity uncertainty for the MC predictions, and experimental uncertainties for the data-driven background normalizations.

The MJ analysis is designed to select events with $E_T$ and jets and to veto events selected by the $\ell + E_T + \text{jet}$ analyses. It accepts events in which the $W$ boson decays into $\tau$ leptons and those in which the electron or muon fails the lepton identification criteria. We use data corresponding to 2.1 fb$^{-1}$ of integrated luminosity for the MJ analysis and select events that have $E_T > 50$ GeV and two jets within $|\eta| < 2.0$, at least one of which has $|\eta| < 0.9$. The jet energy measurements include information from both the calorimeter and the charged-particle spectrometer. Events must have one jet with transverse energy $E_T$ greater than 35 GeV, and a second jet with $E_T$ greater than 25 GeV. The angular separation between the two jets, $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$, is required to exceed 1.0. We reject events with four or more jets with...
Table I: Background composition and predicted number of single top events in 3.2 fb$^{-1}$ of CDF Run II data for the $\ell+E_T +jets$ samples (LF, ME, NN, and BDT analyses), and 2.1 fb$^{-1}$ of data for the $E_T+jets$ sample (MJ analysis).

<table>
<thead>
<tr>
<th>Process</th>
<th>$\ell + E_T + jets$</th>
<th>$E_T + jets$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s$-channel signal</td>
<td>77.3 ± 11.2</td>
<td>29.6 ± 3.7</td>
</tr>
<tr>
<td>$t\bar{t}$-channel signal</td>
<td>113.8 ± 16.9</td>
<td>34.5 ± 6.1</td>
</tr>
<tr>
<td>$W + HF$</td>
<td>1551.0 ± 472.3</td>
<td>304.4 ± 115.5</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>686.1 ± 99.4</td>
<td>184.5 ± 30.2</td>
</tr>
<tr>
<td>$Z$+jets</td>
<td>52.1 ± 8.0</td>
<td>128.6 ± 53.7</td>
</tr>
<tr>
<td>Diboson</td>
<td>118.4 ± 12.2</td>
<td>42.1 ± 6.7</td>
</tr>
<tr>
<td>QCD+mistags</td>
<td>777.9 ± 103.7</td>
<td>679.4 ± 27.9</td>
</tr>
<tr>
<td>Total prediction</td>
<td>3376.5 ± 504.9</td>
<td>1404 ± 172</td>
</tr>
<tr>
<td>Observed</td>
<td>3315</td>
<td>1411</td>
</tr>
</tbody>
</table>

After event selection, the samples are dominated by background. We further discriminate the signal with multivariate techniques. Each multivariate technique defines a function which reduces several reconstructed quantities for each event into a single output variable whose distribution can be studied and fit to extract signal and background contributions. Validation of the background modeling for the input variables and output distributions is a crucial step in the use of multivariate techniques. We first describe the construction of our multivariate tools and then the checks we used to prove the validity of our background model. The LF, ME, and NN discriminants are described in [7]. The BDT discriminant uses a decision tree method that applies binary cuts iteratively to classify events [14]. The discrimination is further improved using a boosting algorithm [15, 16]. The BDT discriminant uses over 20 input variables. Some of the most sensitive are the neural-network jet-flavor separator [17], the invariant mass of the $t\bar{t}b$ system $M_{t\bar{t}b}$, the total scalar sum of transverse energy in the event $H_T$, $Q \times \eta$, the dijet mass $M_{jj}$, and the transverse mass of the W boson.

The LFS discriminant uses projective likelihood functions [19] to combine the separation power of several variables and is optimized to be sensitive to the $s$-channel process. The subset of the $\ell+E_T +jets$ sample with two $b$-tagged jets is used and consists of 609 events. The dominant backgrounds are $W+HF$ and $t\bar{t}$ production. A kinematic fitter is used to find the most likely resolution of two ambiguities: the $z$-component of the neutrino momentum and the $b$ jet that most likely came from the top quark decay. In addition to the outputs of the kinematic fitter, other important inputs to the likelihood are the invariant mass of the two $b$-tagged jets $M_{bb}$, the transverse momentum of the $b\bar{b}$ system, the leading jet transverse momentum, $M_{t\bar{t}b}$, $H_T$, and $E_T$.

The MJ discriminant uses a neural network to combine information from several input variables. The most important variables are the invariant mass of the $E_T$ and the second leading jet, the scalar sum of the jet energies, the $E_T$, and the azimuthal angle between the $E_T$ and the jets.

We combine the LF, ME, NN, BDT, and LFS channels using a super-discriminant (SD) technique similar to that which was applied in [7]. The SD method uses a neural network trained with neuro-evolution [20] to separate the signal from the background taking as inputs the discriminant outputs of the five analyses for each event. With the super-discriminant analysis we improve the sensitivity (defined below) by 13% over the best individual analysis. We perform a simultaneous fit over the two exclusive channels, MJ and SD, to obtain the final combined results.

Before investigating the sample of selected events, we used background-dominated data control samples to check the modeling of each input variable as well as the output distributions of each multivariate discriminant. For the $\ell+E_T +jets$ analyses the control samples used are the lepton + $b$-tagged four-jet sample, which is enriched in $t\bar{t}$ events, and the two- and three-jet samples in which there is no $b$-tagged jet. The latter are enriched in $W+J$ets and QCD events with kinematics similar to the $b$-tagged signal samples and have high statistics, making it possible to observe that the background model describes the data well over three orders of magnitude in our output discriminants. For the MJ analysis, three control samples are used: in the first sample, the $E_T$ is required to be aligned along one of the jets, and in the second, the events are required to fail the NNQCD requirement, and
in the third, a lepton is required to be present. The data distributions in all control samples are described well by our models for each of the analysis input variables and a large set of other variables not used as inputs. More than two thousand distributions were checked for evidence of mismodeling. Small discrepancies were found in the distributions of the angles between two jets in the untagged lepton + two-jet sample and the modeling of jets with rapidity greater than 2.4. These effects are included as systematic uncertainties on the shape of the background models.

Figure 1 shows the distributions of the five \( \ell + \slashed{E}_T + \) jets discriminants. These are combined to give the SD distribution shown in Fig. 2 together with the MJ distribution. In the rightmost bins, assuming SM production and decay, the SD has an \( s/b \) that exceeds 5.0. This large \( s/b \) significantly reduces our sensitivity to systematic uncertainties affecting the background. We use the distributions of the SD and MJ discriminants to extract the measured cross section and the signal significance.

We measure the single top cross section using a Bayesian binned likelihood technique [21] assuming a flat prior in the cross section and integrating the posterior distribution. A

\[
\ln L = \sum_{i} \ln p_{\text{data}}(s+b) + \ln p_{\text{data}}(b) - \ln \int p_{\text{data}}(s+b) + \ln \int p_{\text{data}}(b)
\]

This is the most probable value of the combined \( s \)-channel and \( t \)-channel production cross section. The most probable value of the combined \( s \)-channel and \( t \)-channel cross section is 2.3\( \pm 0.5 \) pb assuming a top quark mass of 175 GeV/\( c^2 \). The dependence on the top quark mass is +0.02 \( \pm 0.01 \) pb/(GeV/\( c^2 \)). From the cross section measurement at \( m_t = 175 \text{ GeV}/c^2 \), we obtain \( |V_{tb}| = 0.91 \pm 0.11 \text{ (stat + syst) \pm 0.07 \text{ (theory) |V_{tb}|} \) and limit \( |V_{tb}| > 0.71 \) at the 95% C.L. assuming a flat prior in \( |V_{tb}|^2 \) from 0 to 1. This is the most precise direct measurement of \( |V_{tb}| \) to date.

In summary, we combine six multivariate analysis techniques to precisely measure the electroweak single top production cross section and the CKM matrix element \( |V_{tb}| \). We have carefully cross-checked our analysis techniques with data control samples and we assign generous rate and shape uncertainties to all predictions we use. Our combined discriminant allows us to purify a signal

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Cross (pb)</th>
<th>Significance</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF</td>
<td>1.6( \pm 0.8 )</td>
<td>2.4</td>
<td>4.0</td>
</tr>
<tr>
<td>ME</td>
<td>2.5( \pm 0.7 )</td>
<td>4.3</td>
<td>4.9</td>
</tr>
<tr>
<td>NN</td>
<td>1.8( \pm 0.6 )</td>
<td>3.5</td>
<td>5.2</td>
</tr>
<tr>
<td>BDT</td>
<td>2.1( \pm 0.6 )</td>
<td>3.5</td>
<td>5.2</td>
</tr>
<tr>
<td>LFS</td>
<td>1.5( \pm 0.9 )</td>
<td>2.0</td>
<td>1.1</td>
</tr>
<tr>
<td>SD</td>
<td>2.1( \pm 0.6 )</td>
<td>4.8</td>
<td>&gt; 5.9</td>
</tr>
<tr>
<td>MJ</td>
<td>4.9( \pm 2.5 )</td>
<td>2.1</td>
<td>1.4</td>
</tr>
<tr>
<td>Combined</td>
<td>2.3( \pm 0.5 )</td>
<td>5.0</td>
<td>&gt; 5.9</td>
</tr>
</tbody>
</table>
\textbf{FIG. 1:} Discriminant distributions for the $\ell+\mathbb{E}_T$+jets analyses. The data are indicated with points, and the predictions are shown separately for each contribution with stacked histograms. The signal expectations shown are the SM predictions. The insets show the distributions of the candidate events in the high-discriminant region.

\textbf{FIG. 2:} Discriminant distributions for the (a) SD, and (b) MJ analyses (see Fig. 1 for their caption and legend). Figure (c) shows the distribution of the likelihood ratio test statistic $-2\ln Q$.

The sample with $s/b > 5.0$ in the most sensitive region, allowing for a significant outcome in the presence of these conservative systematic uncertainties. We observe single top quark production for the first time with a significance of 5.0 standard deviations.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the Humboldt Foundation, the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucléaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; and the Academy of Finland.
[11] We use a cylindrical coordinate system with its origin in the center of the detector, where θ and φ are the polar and azimuthal angles, respectively, and pseudorapidity is η = −ln tan(θ/2). The missing $E_T$ ($\vec{E}_T$) is defined by $\vec{E}_T = -\sum_i E_T^n \hat{n}_i$, $i$ = calorimeter tower number, where $\hat{n}_i$ is a unit vector perpendicular to the beam axis and pointing at the $i^{th}$ calorimeter tower. $E_T$ is corrected for high-energy muons and also jet energy corrections. We define $E_T = |\vec{E}_T|$. The transverse momentum $p_T$ is defined to be $p_T \sin \theta$.
[18] $Q \times \eta$ is the charge of the lepton times the pseudorapidity of the jet not assigned to be the $b$ from top quark decay; see C. P. Yuan, Phys. Rev. D 41, 42 (1990).