IL NUOVO CIMENTO Vol. 32 C, N. 5-6 Settembre-Dicembre 2009
DOI 10.1393/ncc/i2010-10549-5

Colloquia: LaThuile09

First observation of single-top-quark production

G. Otero y Garzón on behalf of the DØ and CDF Collaborations
Universidad de Buenos Aires - Buenos Aires, Argentina
Fermi National Accelerator Laboratory - Batavia IL, USA

(ricevuto il 10 Novembre 2009; pubblicato online il 20 Gennaio 2010)

Summary. — This paper reports the first observation of the electroweak production of single top quarks in $p\bar{p}$ collisions at a center-of-mass energy of 1.96 TeV based on 2.3 fb^{-1} and 3.2 fb^{-1} of data collected by the DØ and CDF detectors, respectively, at the Fermilab Tevatron Collider. Using various multivariate techniques to separete the small signal from the large backgrounds, both experiments obtained a significance of the observed data of 5.0 standard deviations. DØ measures $\sigma(p\bar{p} \to tb + X, tqb + X) = 3.94 \pm 0.88$ pb and CDF measures $2.3^{+0.6}_{-0.5}$ pb. The CKM matrix element that couples the top and the bottom quarks is also measured. DØ reports $|V_{tb}| > 0.78$ at 95% CL when $f_L^1 = 1$ and $|V_{tb} f_L^1| = 1.02 \pm 0.12$. CDF measures $|V_{tb}| = 0.91 \pm 0.11(\text{stat} + \text{syst}) \pm 0.07(\text{theory})$ and sets the limit $|V_{tb}| > 0.71$ at 95% CL.

PACS 14.65.Ha – Top quarks.
PACS 12.15.Ji – Applications of electroweak models to specific processes.
PACS 13.85.Qk – Inclusive production with identified leptons, photons, or other nonhadronic particles.
PACS 12.15.Hh – Determination of Kobayashi-Maskawa matrix elements.

1. – Introduction

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, 3]. Since then, pair production has been used to make precise measurements of several top quark properties.

The Standard Model (SM) predicts the production of top quarks via the electroweak force (single top production). This mechanism serves as a probe of the Wtb interaction [4, 5] and its production cross-section provides a direct measurement of the CKM matrix element that mixes the top and the bottom quarks without assuming three quark generations [6]. However, measuring the single top production cross-section is difficult because of its small rate and the large backgrounds.

At Tevatron energies, top quarks can be produced singly through s-channel (also called tb) or t-channel (also named tqb) exchange of a virtual W-boson as shown in

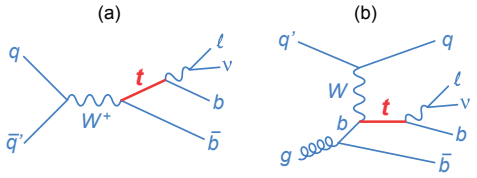

Fig. 1. – Representative Feynman diagrams for (a) s-channel single-top-quark production and (b) t-channel production, showing the top-quark decays of interest.

fig. 1. The sum of the predicted cross-sections of these two processes is 3.5 ± 0.2 pb for $m_t = 170$ GeV [7] and 2.9 ± 0.3 pb for $m_t = 175$ GeV [8].

Both the DØ and CDF Collaborations have published evidence for single-top-quark production at significance levels 3.6 [9, 10] and 3.7 [11] standard deviations, respectively [9-11]. This paper reports significant updates to the previous measurements including larger data samples and new analysis techniques achieving signal significance levels above 5 standard deviations, thus conclusively observing electroweak production of single top quarks. Most definitions and abreviations are defined in the corresponding letters [12, 13].

2. – General analysis strategy

The measurements focus on the final state containing one high transverse momentum (p_T) lepton (electron or muon) not near a jet (isolated), large missing transverse energy ($\not\!\!E_T$) indicative of the passage of a neutrino ν, a b-quark jet from the decay of the top quark ($t \to Wb \to l\nu b$), and possibly another b-jet and a light jet as indicated in fig. 1. In the case of the DØ analysis the data were collected using a logical OR of many trigger conditions and several offline selection criteria, including b-jet identification requirements have been loosened with the net resulf of an 18% increase in signal acceptance compared to the evidence publication [10]. The CDF analysis is based on updates to the published evidence report [11] and the adition of three new multivariate analyses and a new analysis that makes use of a sample that is orthogonal to the event selection described above that adds about 30% to the signal acceptance.

The analyses consider the following backgrounds: W-boson production in association with jets, top quark pair ($t\bar{t}$) production with decay into lepton+jets and dilepton final states (when a lepton in not reconstructed) and multijet production, where a jet is misreconstructed as an electron or a heavy-flavor quark decays into a muon that passes the isolation criteria. Z + jets and diboson processes form minor additional background components.

Because the single-top cross-section is very small compared to the competing backgrounds, a simple cut-based counting experiment is not sufficient to verify the presence of the signal. Because of this, the first step is to apply the loosest event selection possible to maximize the signal acceptance. After the event selection, the expected signal is typically smaller than the uncertainty on the background. Thus, the main strategy implemented by both collaborations is to use sophisticated multivariate techniques to extract the small signal from the overwhelming backgrounds. Each such multivariate technique constructs a powerful discriminant variable that is proportional to the probability of an event to be signal. The discriminant distribution is used as input to the cross-section measurement.

TABLE I. – *Number of expected and observed events in* $2.3\,\mathrm{fb}^{-1}$ *of DØ data for all analysis channels combined. The uncertainties include both statistical and systematic components.*

Source	2 jets	3 jets	4 jets
Signal	139 ± 18	63 ± 10	21 ± 5
W + jets	1829 ± 161	637 ± 61	180 ± 18
Z + jets and dibosons	229 ± 38	85 ± 17	27 ± 7
$t\bar{t}$	222 ± 35	436 ± 66	484 ± 71
Multijets	196 ± 50	73 ± 17	30 ± 6
Total prediction	2615 ± 192	1294 ± 107	742 ± 80
Data	2579	1216	724

Several validation tests are conducted by studying the discriminant output distributions in background enriched control samples.

The single-top-quark production cross-section is measured from the discriminant output distributions using a Bayesian binned likelihood technique [14]. The statistical and all systematic uncertainties and their correlations are considered in these calculations.

3. – DØ analysis

The DØ analysis considers 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 the (η, ϕ)-space, where η is the rapidity and ϕ is the azimuthal angle, and the cone radius is 0.5 [10]. The highest-p_T (leading) jet must have $p_T > 25\,\mathrm{GeV}$, and subsequent jets have $p_T > 15\,\mathrm{GeV}$; all jets have pseudorapidity $|\eta| < 3.4$. The selection requires $20 < \not\!\!E_T < 200\,\mathrm{GeV}$ for events with two jets and $25 < \not\!\!E_T < 200\,\mathrm{GeV}$ for events with three or four jets. Events must contain only one isolated electron with $p_T > 15\,\mathrm{GeV}$ and $|\eta| < 1.1$ ($p_T > 20\,\mathrm{GeV}$ for three- or four-jet events), or one isolated muon with $p_T > 15\,\mathrm{GeV}$ 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 $\not\!\!E_T$ is not along the direction of the lepton or the leading jet.

Table I shows the event yields, separated by jet multiplicity. The acceptances are $(3.7 \pm 0.5)\%$ for the s-channel and $(2.5 \pm 0.3)\%$ for the t-channel, expressed as percentages of the inclusive single top quark production cross-section in each channel.

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 the shapes of the distributions; these are the jet energy scale corrections, the tag-rate functions, and the reweighting of the distributions in W + jets events. The largest uncertainties come from the jet energy scale, the tag-rate functions, and the correction for jet-flavor composition in W + jets events, with smaller contributions from the integrated luminosity, jet energy resolution, initial-state and final-state radiation, b-jet fragmentation, $t\bar{t}$ cross-section, lepton efficiency corrections, and MC statistics. All other contributions have a smaller effect. The total uncertainty on the background is (8 to 16)% depending on the analysis channel. After event selection, single-top-quark events are expected to constitute (3 to 9)% of the data sample.

DØ perfomed three independent analyses based on boosted decision trees (BDT) [15], Bayesian neural networks (BNN) [16], and the matrix element (ME) method [17]. The application of these techniques is described in [9, 10]. The analyses presented here differ

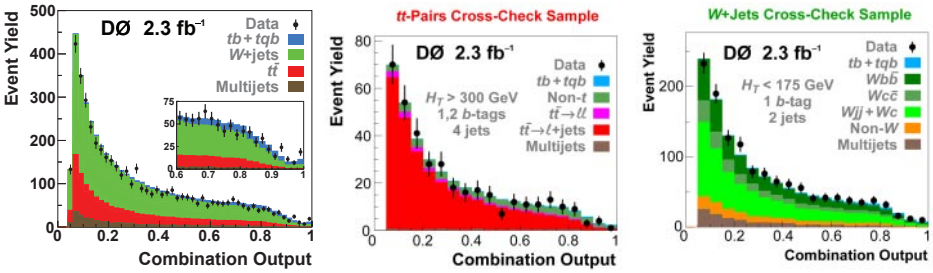

Fig. 2. – Distribution of the DØ discriminant output for the combined analysis (left). Cross check of the discriminant output in signal-depleted and background-enriched samples: $t\bar{t}$ (center) and W + jets (right).

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 BNN analysis uses the RuleFitJF algorithm [18] to select the most sensitive of these variables, then combines 18 to 28 of them into a single separate discriminant for each channel. The ME analysis uses only two-jet and three-jet events, divided into a (W + jets)-dominated set and a $t\bar{t}$-dominated set. It includes matrix elements for more background sources to improve background rejection.

Each analysis uses the same data and background model and has the same sources of systematic uncertainty. The analyses are tested using ensembles of pseudodatasets 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 + jets and the other by $t\bar{t}$ backgrounds. 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 previous analyses [9, 10]. This involves forming a binned likelihood as a product over all bins and channels, evaluated separately for each multivariate discriminant. 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.

The three multivariate techniques use the same data sample but are not completely correlated. Their combination therefore leads to increased sensitivity and a more precise measurement of the cross-section. The three discriminant outputs are used 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 standard deviations. Figure 2 illustrates the importance of the signal when comparing data to prediction and also the behaviour of the discriminant output in background-dominated samples.

The measured cross-section is $\sigma(p\bar{p} \to tb + X, tqb + X) = 3.94 \pm 0.88\,\text{pb}$. The measurement has a p-value of 2.5×10^{-7}, corresponding to a significance of 5.0 standard deviations. Figure 3 shows the results of all DØ analyses and the ensemble test performed to measure the observed significance of the combined analysis.

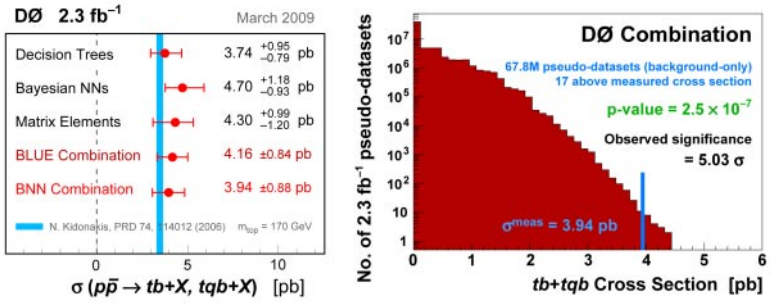

Fig. 3. – DØ cross-section measurements summary (left). Cross-section distribution from background-only ensembles with full systematics included for the DØ combined analysis (right).

The cross-section measurement is used to determine the Bayesian posterior for $|V_{tb}|^2$ in the interval $[0, 1]$ and a limit of $|V_{tb}| > 0.78$ at 95% CL is extracted within the SM. When the upper constraint is removed, $|V_{tb} \times f_1^L| = 1.07 \pm 0.12$ is measured, where f_1^L is the strength of the left-handed W_{tb} coupling.

4. – CDF analysis

The CDF collaboration updated the published [11] likelihood function (LF), matrix element (ME), and neural network (NN) analyses with an additional 1 fb^{-1} of integrated luminosity with their methods unchanged. In addition, three new analyses are added: 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 $\not\!\!E_T$ 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.

The LF, ME, NN, BDT, and LFS analyses use lepton + jets events as described in the DØ analysis. The MJ analysis is designed to select events with $\not\!\!E_T$ and jets and to veto events selected by the lepton + jets analyses. It accepts events in which the W-boson decays into τ-leptons and those in which the electron or muon fails the lepton identification criteria. The MJ analysis uses a dataset of 2.1 fb^{-1} of integrated luminosity and selects events that have $\not\!\!E_T > 50$ GeV and two jets with $|\eta| < 2$, at least one of which has $|\eta| < 0.9$. 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 is required to exceed 1. Events with four or more jets with $E_T > 15$ GeV in $|\eta| < 2.4$ are rejected in order to reduce the multijet (QCD) and $t\bar{t}$ backgrounds and at least one jet is required to originate from a B-hadron. The observed and expected event counts for the all the analyses are given in table II.

After event selection, the samples are dominated by background and multivariate techniques are used to further discriminate the signal. The LF, ME, and NN discriminants are described in detail in [11]. The BDT discriminant uses over 20 input variables. Some of the most sensitive are the neural-network jet-flavor separator, the invariant mass of the $l\nu b$ system $M_{l\nu b}$ and the total scalar sum of transverse energy in the event H_T. 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

TABLE II. – *Background composition and predicted number of single-top events in* $3.2\,\text{fb}^{-1}$ *of CDF Run-II data for the* $l + \not\!\!E_T +$ *jets samples (LF, ME, NN, and BDT analyses), and* $2.1\,\text{fb}^{-1}$ *of data for the* $\not\!\!E_T +$ *jets sample (MJ analysis).*

Source	$l + \not\!\!E_T +$ jets	$\not\!\!E_T +$ jets
s-channel signal	77 ± 11	30 ± 4
t-channel signal	114 ± 17	35 ± 6
$W +$ HF	1551 ± 472	304 ± 116
$t\bar{t}$	686 ± 99	185 ± 30
$Z +$ jets	52 ± 8	129 ± 54
Diboson	118 ± 12	42 ± 7
QCD+mistags	778 ± 104	679 ± 28
Total prediction	3377 ± 505	1404 ± 172
Observed	3315	1411

dominant backgrounds are $Wb\bar{b}$ 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, the transverse momentum of the $b\bar{b}$ system, the leading jet transverse momentum, H_T, and $\not\!\!E_T$.

The MJ discriminant uses a neural network to combine information from several input variables. The most important variables are the invariant mass of the $\not\!\!E_T$ and the second leading jet, the scalar sum of the jet energies, the $\not\!\!E_T$, and the azimuthal angle between the $\not\!\!E_T$ and the jets.

The LF, ME, NN, BDT, and LFS channels are combined using a super-discriminant (SD) technique similar to that which was applied in [11]. 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 the sensitivity improves by 13% over the best individual analysis. A simultaneous fit is performed over the two exclusive channels, MJ and SD, to obtain the final combined results. The modeling of the distributions of each input variable and the discriminant outputs are checked in data control samples depleted in

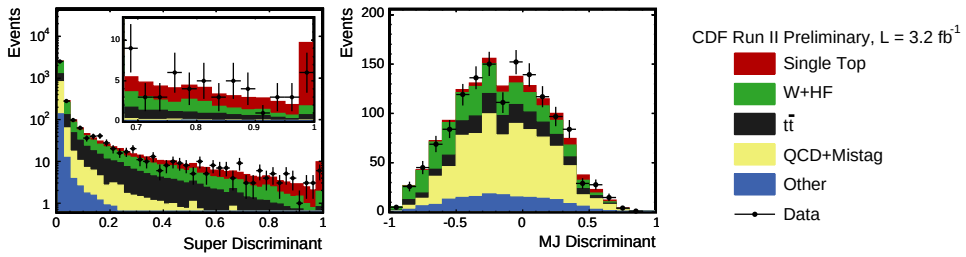

Fig. 4. – CDF's discriminant distributions for the SD (left) and MJ (right) analyses.

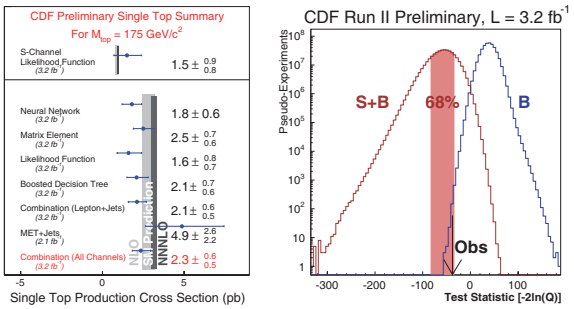

Fig. 5. – CDF's cross-section measurements summary (left). Distribution of the likelihood ratio test statistic $-2\ln Q$ (right).

signal. These 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 has high statistics and is enriched in W + jets and QCD events with kinematics similar to the b-tagged signal samples. The distributions of the SD and MJ discriminants shown in fig. 4 are used to extract the measured cross-section and the signal significance.

The cross-sections are measured using a Bayesian binned likelihood assuming a flat prior in the cross-section in the same way as the DØ analysis. The significance is calculated as a p-value: the probability, assuming single-top-quark production is absent, that $-2\ln Q = -2\ln(p(\text{data}|s+b)/p(\text{data}|b))$ is less than that observed in the data. Figure 5 shows all cross-section measurements and the distributions of $-2\ln Q$ in pseudoexperiments that assume SM single top (S + B) and also those that assume single-top production is absent (B), along with the value observed in data. The p-value is converted into a number of standard deviations using the integral of one side of a Gaussian function. All sources of systematic uncertainty are included and correlations between normalization and discriminant shape changes are considered. Uncertainties in the jet energy scale, b-tagging efficiencies, lepton identification and trigger efficiencies, the amount of initial- and final-state radiation, PDFs, factorization and renormalization scale, and background modeling have been explored and incorporated in all individual analyses and the combination.

The measured p-value is 3.1×10^{-7}, corresponding to a signal significance of 5.0 standard deviations. The sensitivity is defined to be the median expected significance and is in excess of 5.9 standard deviations. The most probable value of the combined s- and t-channels cross-sections is $2.3^{+0.6}_{-0.5}$ pb assuming a top quark mass of 175 GeV/c^2. The cross-section measurement is used to determine that $|V_{tb}| = 0.91 \pm 0.11(\text{stat}+\text{syst}) \pm 0.07(\text{theory})$ and the limit $|V_{tb}| > 0.71$ at 95% CL assuming a flat prior in $|V_{tb}|^2$ from 0 to 1.

5. – Conclusions

The DØ and CDF Collaborations reported the first observation of electroweak production of single top quarks in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using 2.3 and 3.2 fb^{-1} of Tevatron data, respectivelly. The results are in agreement with the Standard Model prediction and the measured signal corresponds to an excess over the predicted backgrounds with significances of 5.0 standard deviations. The results provide the most precise direct measurements of the amplitude of the CKM matrix element V_{tb}.

* * *

The author wishes to thank the conference organizers for a very interesting meeting, and members of the single-top working groups from the CDF and DØ Collaborations as well as the Fermilab staff and the techinical staffs of the participating institutions for their vital contributions.

REFERENCES

[1] Willenbrock S. and Dicus D., *Phys. Rev. D*, **34** (1986) 155.
[2] Abe F. *et al.* (The CDF Collaboration), *Phys. Rev. Lett.*, **74** (1995) 2626.
[3] Abachi S. *et al.* (The DØ Collaboration), *Phys. Rev. Lett.*, **74** (1995) 2632.
[4] Heinsion A. *et al.*, *Phys. Rev. D*, **56** (1997) 3114.
[5] Abazov V. *et al.* (The DØ Collaboration), *Phys. Rev. Lett.*, **101** (2008) 221801.
[6] Jikia G. *et al.*, *Phys. Lett. B*, **295** (1992) 136.
[7] Kidonakis N., *Phys. Rev. D*, **74** (2006) 114012.
[8] Sullivan Z., *Phys. Rev. D*, **70** (2004) 114012.
[9] Abazov V. *et al.* (The DØ Collaboration), *Phys. Rev. Lett.*, **98** (2007) 181802.
[10] Abazov V. *et al.* (The DØ Collaboration), *Phys. Rev. D*, **78** (2008) 012005.
[11] Aaltonen T. *et al.* (The CDF Collaboration), *Phys. Rev. Lett.*, **101** (2008) 252001.
[12] Abazov V. *et al.* (The DØ Collaboration), *Phys. Rev. Lett.*, **103** (2009) 092001, arXiv:0903.0850.
[13] Aaltonen T. *et al.* (The CDF Collaboration), *Phys. Rev. Lett.*, **103** (2009) 092002, arXiv:0903.0885.
[14] Bertram I. *et al.*, Fermilab-TM-2104 (2000).
[15] Breiman L. *et al.*, *Classification and Regression Trees* (Wadsworth, Standford) 1984.
[16] Neil R., *Bayesian Learning for Neural Networks* (Springer-Verlag, New York) 1996.
[17] Abazov V. *et al.* (The DØ Collaboration), *Nature*, **429** (2004) 638.
[18] Friedman J. *et al.*, *Ann. Appl. Stat.*, **2** (2005) 094027.
[19] Ackerstaff K. *et al.* (The OPAL Collaboration), *Eur. Phys. J. C*, **1** (1998) 425.
[20] Stanley K. *et al.*, *Evolutionary Computation*, **10** (2002) 99.