FERMILAB-PUB-13-252-E

Evidence for s-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

V.M. Abazov,[31] B. Abbott,[66] B.S. Acharya,[25] M. Adams,[45] T. Adams,[43] J.P. Agnew,[40] G.D. Alexeev,[31] G. Alkhazov,[35] A. Alton[a],[55] A. Askew,[43] S. Atkins,[53] K. Augsten,[7] C. Avila,[5] F. Badaud,[10] L. Bagby,[44] B. Baldin,[44] D.V. Bandurin,[43] S. Banerjee,[25] E. Barberis,[54] P. Baringer,[52] J.F. Bartlett,[44] U. Bassler,[15] V. Bazterra,[45] A. Bean,[52] M. Begalli,[2] L. Bellantoni,[44] S.B. Beri,[23] G. Bernardi,[14] R. Bernhard,[19] I. Bertram,[38] M. Besançon,[15] R. Beuselinck,[39] P.C. Bhat,[44] S. Bhatia,[57] V. Bhatnagar,[23] G. Blazey,[46] S. Blessing,[43] K. Bloom,[58] A. Boehnlein,[44] D. Boline,[63] E.E. Boos,[33] G. Borissov,[38] A. Brandt,[69] O. Brandt,[20] R. Brock,[56] A. Bross,[44] D. Brown,[14] X.B. Bu,[44] M. Buehler,[44] V. Buescher,[21] V. Bunichev,[33] S. Burdin[b],[38] C.P. Buszello,[37] E. Camacho-Pérez,[28] B.C.K. Casey,[44] H. Castilla-Valdez,[28] S. Caughron,[56] S. Chakrabarti,[63] K.M. Chan,[50] A. Chandra,[71] E. Chapon,[15] G. Chen,[52] S.W. Cho,[27] S. Choi,[27] B. Choudhary,[24] S. Cihangir,[44] D. Claes,[58] J. Clutter,[52] M. Cooke,[44] W.E. Cooper,[44] M. Corcoran,[71] F. Couderc,[15] M.-C. Cousinou,[12] D. Cutts,[68] A. Das,[41] G. Davies,[39] S.J. de Jong,[29,30] E. De La Cruz-Burelo,[28] F. Déliot,[15] R. Demina,[62] D. Denisov,[44] S.P. Denisov,[34] S. Desai,[44] C. Deterre[d],[20] K. DeVaughan,[58] H.T. Diehl,[44] M. Diesburg,[44] P.F. Ding,[40] A. Dominguez,[58] A. Dubey,[24] L.V. Dudko,[33] A. Duperrin,[12] S. Dutt,[23] M. Eads,[46] D. Edmunds,[56] J. Ellison,[42] V.D. Elvira,[44] Y. Enari,[14] H. Evans,[48] V.N. Evdokimov,[34] L. Feng,[46] T. Ferbel,[62] F. Fiedler,[21] F. Filthaut,[29,30] W. Fisher,[56] H.E. Fisk,[44] M. Fortner,[46] H. Fox,[38] S. Fuess,[44] A. Garcia-Bellido,[62] J.A. García-González,[28] V. Gavrilov,[32] W. Geng,[12,56] C.E. Gerber,[45] Y. Gershtein,[59] G. Ginther,[44,62] G. Golovanov,[31] P.D. Grannis,[63] S. Greder,[16] H. Greenlee,[44] G. Grenier,[17] Ph. Gris,[10] J.-F. Grivaz,[13] A. Grohsjean[c],[15] S. Grünendahl,[44] M.W. Grünewald,[26] T. Guillemin,[13] G. Gutierrez,[44] P. Gutierrez,[66] J. Haley,[54] L. Han,[4] K. Harder,[40] A. Harel,[62] J.M. Hauptman,[51] J. Hays,[39] T. Head,[40] T. Hebbeker,[18] D. Hedin,[46] H. Hegab,[67] A.P. Heinson,[42] U. Heintz,[68] C. Hensel,[20] I. Heredia-De La Cruz[d],[28] K. Herner,[44] G. Hesketh[f],[40] M.D. Hildreth,[50] R. Hirosky,[72] T. Hoang,[43] J.D. Hobbs,[63] B. Hoeneisen,[9] J. Hogan,[71] M. Hohlfeld,[21] J.L. Holzbauer,[57] I. Howley,[69] Z. Hubacek,[7,15] V. Hynek,[7] I. Iashvili,[61] Y. Ilchenko,[70] R. Illingworth,[44] A.S. Ito,[44] S. Jabeen,[68] M. Jaffré,[13] A. Jayasinghe,[66] M.S. Jeong,[27] R. Jesik,[39] P. Jiang,[4] K. Johns,[41] E. Johnson,[56] M. Johnson,[44] A. Jonckheere,[44] P. Jonsson,[39] J. Joshi,[42] A.W. Jung,[44] A. Juste,[36] E. Kajfasz,[12] D. Karmanov,[33] I. Katsanos,[58] R. Kehoe,[70] S. Kermiche,[12] N. Khalatyan,[44] A. Khanov,[67] A. Kharchilava,[61] Y.N. Kharzheev,[31] I. Kiselevich,[32] J.M. Kohli,[23] A.V. Kozelov,[34] J. Kraus,[57] A. Kumar,[61] A. Kupco,[8] T. Kurča,[17] V.A. Kuzmin,[33] S. Lammers,[48] P. Lebrun,[17] H.S. Lee,[27] S.W. Lee,[51] W.M. Lee,[43] X. Lei,[41] J. Lellouch,[14] D. Li,[14] H. Li,[72] L. Li,[42] Q.Z. Li,[44] J.K. Lim,[27] D. Lincoln,[44] J. Linnemann,[56] V.V. Lipaev,[34] R. Lipton,[44] H. Liu,[70] Y. Liu,[4] A. Lobodenko,[35] M. Lokajicek,[8] R. Lopes de Sa,[63] R. Luna-Garcia[g],[28] A.L. Lyon,[44] A.K.A. Maciel,[1] R. Madar,[19] R. Magaña-Villalba,[28] S. Malik,[58] V.L. Malyshev,[31] J. Mansour,[20] J. Martínez-Ortega,[28] R. McCarthy,[63] C.L. McGivern,[40] M.M. Meijer,[29,30] D. Meister,[45] A. Melnitchouk,[44] D. Menezes,[46] P.G. Mercadante,[3] M. Merkin,[33] A. Meyer,[18] J. Meyer[i],[20] F. Miconi,[16] N.K. Mondal,[25] M. Mulhearn,[72] E. Nagy,[12] M. Narain,[68] R. Nayyar,[41] H.A. Neal,[55] J.P. Negret,[5] P. Neustroev,[35] H.T. Nguyen,[72] T. Nunnemann,[22] J. Orduna,[71] N. Osman,[12] J. Osta,[50] A. Pal,[69] N. Parashar,[49] V. Parihar,[68] S.K. Park,[27] R. Partridge[e],[68] N. Parua,[48] A. Patwa[j],[64] B. Penning,[44] M. Perfilov,[33] Y. Peters,[20] K. Petridis,[40] G. Petrillo,[62] P. Pétroff,[13] M.-A. Pleier,[64] V.M. Podstavkov,[44] A.V. Popov,[34] M. Prewitt,[71] D. Price,[48] N. Prokopenko,[34] J. Qian,[55] A. Quadt,[20] B. Quinn,[57] P.N. Ratoff,[38] I. Razumov,[34] I. Ripp-Baudot,[16] F. Rizatdinova,[67] M. Rominsky,[44] A. Ross,[38] C. Royon,[15] P. Rubinov,[44] R. Ruchti,[50] G. Sajot,[11] A. Sánchez-Hernández,[28] M.P. Sanders,[22] A.S. Santos[h],[1] G. Savage,[44] L. Sawyer,[53] T. Scanlon,[39] R.D. Schamberger,[63] Y. Scheglov,[35] H. Schellman,[47] C. Schwanenberger,[40] R. Schwienhorst,[56] J. Sekaric,[52] H. Severini,[66] E. Shabalina,[20] V. Shary,[15] S. Shaw,[56] A.A. Shchukin,[34] V. Simak,[7] P. Skubic,[66] P. Slattery,[62] D. Smirnov,[50] G.R. Snow,[58] J. Snow,[65] S. Snyder,[64] S. Söldner-Rembold,[40] L. Sonnenschein,[18] K. Soustruznik,[6] J. Stark,[11] D.A. Stoyanova,[34] M. Strauss,[66] L. Suter,[40] P. Svoisky,[66] M. Titov,[15] V.V. Tokmenin,[31] Y.-T. Tsai,[62] D. Tsybychev,[63] B. Tuchming,[15] C. Tully,[60] L. Uvarov,[35] S. Uvarov,[35] S. Uzunyan,[46] R. Van Kooten,[48] W.M. van Leeuwen,[29] N. Varelas,[45] E.W. Varnes,[41] I.A. Vasilyev,[34] A.Y. Verkheev,[31] L.S. Vertogradov,[31] M. Verzocchi,[44] M. Vesterinen,[40] D. Vilanova,[15] P. Vokac,[7] H.D. Wahl,[43] M.H.L.S. Wang,[44] J. Warchol,[50] G. Watts,[73] M. Wayne,[50] J. Weichert,[21] L. Welty-Rieger,[47] M.R.J. Williams,[48] G.W. Wilson,[52] M. Wobisch,[53] D.R. Wood,[54] T.R. Wyatt,[40] Y. Xie,[44] R. Yamada,[44] S. Yang,[4] T. Yasuda,[44] Y.A. Yatsunenko,[31] W. Ye,[63] Z. Ye,[44] H. Yin,[44] K. Yip,[64] S.W. Youn,[44] J.M. Yu,[55] J. Zennamo,[61] T.G. Zhao,[40] B. Zhou,[55] J. Zhu,[55] M. Zielinski,[62] D. Zieminska,[48] and L. Zivkovic[14]

(The D0 Collaboration*)

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

[1]*LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[2]*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[3]*Universidade Federal do ABC, Santo André, Brazil*
[4]*University of Science and Technology of China, Hefei, People's Republic of China*
[5]*Universidad de los Andes, Bogotá, Colombia*
[6]*Charles University, Faculty of Mathematics and Physics, Center for Particle Physics, Prague, Czech Republic*
[7]*Czech Technical University in Prague, Prague, Czech Republic*
[8]*Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic*
[9]*Universidad San Francisco de Quito, Quito, Ecuador*
[10]*LPC, Université Blaise Pascal, CNRS/IN2P3, Clermont, France*
[11]*LPSC, Université Joseph Fourier Grenoble 1, CNRS/IN2P3, Institut National Polytechnique de Grenoble, Grenoble, France*
[12]*CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France*
[13]*LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France*
[14]*LPNHE, Universités Paris VI and VII, CNRS/IN2P3, Paris, France*
[15]*CEA, Irfu, SPP, Saclay, France*
[16]*IPHC, Université de Strasbourg, CNRS/IN2P3, Strasbourg, France*
[17]*IPNL, Université Lyon 1, CNRS/IN2P3, Villeurbanne, France and Université de Lyon, Lyon, France*
[18]*III. Physikalisches Institut A, RWTH Aachen University, Aachen, Germany*
[19]*Physikalisches Institut, Universität Freiburg, Freiburg, Germany*
[20]*II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany*
[21]*Institut für Physik, Universität Mainz, Mainz, Germany*
[22]*Ludwig-Maximilians-Universität München, München, Germany*
[23]*Panjab University, Chandigarh, India*
[24]*Delhi University, Delhi, India*
[25]*Tata Institute of Fundamental Research, Mumbai, India*
[26]*University College Dublin, Dublin, Ireland*
[27]*Korea Detector Laboratory, Korea University, Seoul, Korea*
[28]*CINVESTAV, Mexico City, Mexico*
[29]*Nikhef, Science Park, Amsterdam, the Netherlands*
[30]*Radboud University Nijmegen, Nijmegen, the Netherlands*
[31]*Joint Institute for Nuclear Research, Dubna, Russia*
[32]*Institute for Theoretical and Experimental Physics, Moscow, Russia*
[33]*Moscow State University, Moscow, Russia*
[34]*Institute for High Energy Physics, Protvino, Russia*
[35]*Petersburg Nuclear Physics Institute, St. Petersburg, Russia*
[36]*Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Física d'Altes Energies (IFAE), Barcelona, Spain*
[37]*Uppsala University, Uppsala, Sweden*
[38]*Lancaster University, Lancaster LA1 4YB, United Kingdom*
[39]*Imperial College London, London SW7 2AZ, United Kingdom*
[40]*The University of Manchester, Manchester M13 9PL, United Kingdom*
[41]*University of Arizona, Tucson, Arizona 85721, USA*
[42]*University of California Riverside, Riverside, California 92521, USA*
[43]*Florida State University, Tallahassee, Florida 32306, USA*
[44]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[45]*University of Illinois at Chicago, Chicago, Illinois 60607, USA*
[46]*Northern Illinois University, DeKalb, Illinois 60115, USA*
[47]*Northwestern University, Evanston, Illinois 60208, USA*
[48]*Indiana University, Bloomington, Indiana 47405, USA*
[49]*Purdue University Calumet, Hammond, Indiana 46323, USA*
[50]*University of Notre Dame, Notre Dame, Indiana 46556, USA*
[51]*Iowa State University, Ames, Iowa 50011, USA*
[52]*University of Kansas, Lawrence, Kansas 66045, USA*
[53]*Louisiana Tech University, Ruston, Louisiana 71272, USA*
[54]*Northeastern University, Boston, Massachusetts 02115, USA*
[55]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[56]*Michigan State University, East Lansing, Michigan 48824, USA*
[57]*University of Mississippi, University, Mississippi 38677, USA*
[58]*University of Nebraska, Lincoln, Nebraska 68588, USA*
[59]*Rutgers University, Piscataway, New Jersey 08855, USA*
[60]*Princeton University, Princeton, New Jersey 08544, USA*
[61]*State University of New York, Buffalo, New York 14260, USA*
[62]*University of Rochester, Rochester, New York 14627, USA*

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

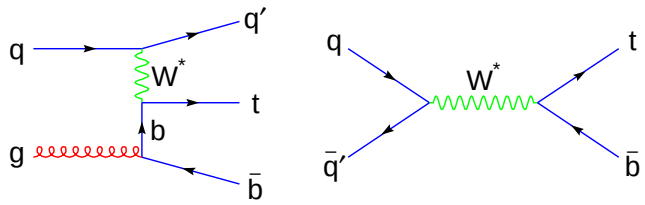
[63] State University of New York, Stony Brook, New York 11794, USA
[64] Brookhaven National Laboratory, Upton, New York 11973, USA
[65] Langston University, Langston, Oklahoma 73050, USA
[66] University of Oklahoma, Norman, Oklahoma 73019, USA
[67] Oklahoma State University, Stillwater, Oklahoma 74078, USA
[68] Brown University, Providence, Rhode Island 02912, USA
[69] University of Texas, Arlington, Texas 76019, USA
[70] Southern Methodist University, Dallas, Texas 75275, USA
[71] Rice University, Houston, Texas 77005, USA
[72] University of Virginia, Charlottesville, Virginia 22904, USA
[73] University of Washington, Seattle, Washington 98195, USA

(Dated: July 2, 2013)

We present measurements of the cross sections for the two main production modes of single top quarks in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV in the Run II data collected with the D0 detector at the Fermilab Tevatron Collider, corresponding to an integrated luminosity of 9.7 fb^{-1}. The s-channel cross section is measured to be $\sigma(p\bar{p} \to tb + X) = 1.10^{+0.33}_{-0.31}$ pb with no assumptions on the value of the t-channel cross section. Similarly, the t-channel cross section is measured to be $\sigma(p\bar{p} \to tqb + X) = 3.07^{+0.54}_{-0.49}$ pb. We also measure the $s+t$ combined cross section as $\sigma(p\bar{p} \to tb + X,\ tqb + X) = 4.11^{+0.60}_{-0.55}$ pb and set a lower limit on the CKM matrix element $|V_{tb}| > 0.92$ at 95% C.L., assuming $m_t = 172.5$ GeV. The probability to measure a cross section for the s channel at the observed value or higher in the absence of signal is 1.0×10^{-4}, corresponding to a significance of 3.7 standard deviations.

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

With a mass of $m_t = 173.2 \pm 0.9$ GeV [1], the top quark is the heaviest elementary particle in the standard model (SM). The phenomenology of top quark production and decay provides powerful means for testing the properties of the strong and electroweak interactions, as well as the possibility of discovering physics beyond the standard model (BSM) that couples strongly to mass. At the Tevatron proton anti-proton collider operating at a center of mass energy of $\sqrt{s} = 1.96$ TeV, top quarks are produced predominantly in pairs via the strong interaction. In addition, they are also produced by the electroweak interaction in three different production modes with a single top quark accompanied by other quarks or a W boson. The dominant production mode at the Tevatron is the exchange of a space-like virtual W boson between a light quark and a bottom quark in the t channel ($tqb = tq\bar{b} + \bar{t}qb$, where q refers to a light quark or antiquark) [2–4]. The second mode is the decay of a time-like virtual W boson in the s channel, which produces a top quark and a bottom quark ($tb = t\bar{b} + \bar{t}b$) [5]. The third mode is the associated tW process, in which the top quark is produced together with a W boson, which contributes negligibly at the Tevatron.



(a) (b)

FIG. 1: [color online] Dominant lowest-order Feynman diagrams for (a) t-channel and (b) s-channel single top quark production.

Figure 1 shows the lowest-order Feynman diagrams for the two dominant production modes at the Tevatron.

Single-top quark production in the combined $s+t$ mode was observed independently by D0 and CDF in 2009 [6, 7] assuming the SM ratio for the s- and t-channel cross sections. Since then, the D0 Collaboration has measured the production cross section for the combined $s+t$ channels to be $3.43^{+0.73}_{-0.74}$ pb [8], assuming $m_t = 172.5$ GeV. The D0 collaboration was also first to observe the t-channel process alone [9] with a significance equal to 5.5 standard deviations (SD), and measured its cross section to be 2.90 ± 0.59 pb. At the CERN LHC proton-proton collider, t-channel production has been recently observed by the ATLAS and CMS Collaborations [10, 11], and there is evidence for tW associated production [12, 13]. Since the cross section of s-channel production is smaller and there are large irreducible backgrounds, this process has only been detected with weak statistical significance at the Tevatron, while only upper limits on the production cross section exist at the

*with visitors from [a]Augustana College, Sioux Falls, SD, USA, [b]The University of Liverpool, Liverpool, UK, [c]DESY, Hamburg, Germany, [d]Universidad Michoacana de San Nicolas de Hidalgo, Morelia, Mexico [e]SLAC, Menlo Park, CA, USA, [f]University College London, London, UK, [g]Centro de Investigacion en Computacion - IPN, Mexico City, Mexico, [h]Universidade Estadual Paulista, São Paulo, Brazil, [i]Karlsruher Institut für Technologie (KIT) - Steinbuch Centre for Computing (SCC) and [j]Office of Science, U.S. Department of Energy, Washington, D.C. 20585, USA.

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

LHC. Indeed, the increase in the s-channel cross section from the Tevatron to the LHC is significantly less than the increase in the main backgrounds such as t channel or $t\bar{t}$, providing better signal to background ratio for this channel at the Tevatron.

Single-top quark events can be used to directly measure the strength of the Wtb vertex in the hard scatter of the collision rather than from the decay rate as in $t\bar{t}$ production. The V_{tb} term of the Cabibbo-Kobayashi-Maskawa (CKM) [14] quark-mixing matrix is heavily constrained if one assumes that there are only three generations of quarks and that the CKM matrix is unitary, as in the SM, yielding $|V_{tb}| = 0.999146^{+0.000021}_{-0.000046}$ [15]. However, several BSM models predict a fourth generation of quarks or a heavy quark singlet that could make $|V_{tb}|$ significantly smaller than unity [16]. By measuring the rate of production of single top quark events, which is proportional to $|V_{tb}|^2$, this CKM element can be measured without the assumptions of three generations and unitarity of the CKM matrix [17].

In this Letter, we present improved simultaneous measurements of the s- and t-channel cross sections with the D0 Tevatron Run II dataset corresponding to 9.7 fb^{-1} of integrated luminosity, recorded between 2002 and 2011. In addition, we provide a measurement of the $s+t$ cross section without assuming the SM ratio between the s and t channels. Finally, we update the measurement of $|V_{tb}|$ extracted from the $s+t$ cross section.

This analysis extends previous work by the D0 Collaboration [6, 8, 9, 18–20] and approximately doubles the integrated luminosity analyzed in the previous publications [8, 9]. The event selection is optimized to maximize the s-channel sensitivity and to adapt to the higher instantaneous luminosity of the latest collected data.

Details about the D0 detector can be found in Ref. [21]. The data are selected from an inclusive sample comprising the logical OR of many trigger conditions, which is fully efficient for the single top quark signal after offline selection. In the SM, top quarks decay almost exclusively to a W boson and a b quark. We look for leptonic decays of the W boson to one electron or muon, and a neutrino. Events are therefore selected if they fulfill the following criteria:

–There must be only one isolated electron with pseudorapidity [22] $|\eta| < 1.1$ and transverse momentum $p_T > 20$ GeV or only one isolated muon with $|\eta| < 2.0$ and $p_T > 20$ GeV.

–The missing transverse energy, calculated as the opposite of the vector sum of the transverse energies of all calorimeter cells surviving noise-suppression algorithms and corrected for the calorimeter energy scale and the momenta of muon tracks, is required to be $20 < \not\!\!E_T < 200$ GeV for events with two jets and $25 < \not\!\!E_T < 200$ GeV for events with three jets. The upper limit on $\not\!\!E_T$ removes events in data having misreconstructed muon p_T.

–There must be two or three jets each having $|\eta| < 2.5$ and one jet having $p_T > 25$ GeV and the remaining jets having $p_T > 20$ GeV. Jets are reconstructed by clustering cells in the calorimeter based on a cone algorithm in (y, ϕ) space, where y is the rapidity and ϕ is the azimuthal angle, and the cone radius is 0.5 [23]. Each jet is also required to have at least two tracks associated with the collision vertex of the $p\bar{p}$ hard scatter. In addition, the energy of the jet is corrected to the level of particles emitted within the jet cone [24].

Once the basic particles in the final state are identified, we apply additional selection criteria to exclude regions of phase space that are difficult to model precisely. We require the scalar sum of the p_T of the lepton, the $\not\!\!E_T$, and the p_T of all the jets in the event to satisfy $H_{\rm T}(\text{jets}, \ell, \not\!\!E_T) > 120$ GeV in events with two jets, and > 160 GeV in events with three jets. To remove multijet events where fake missing energy arises from a jet which is misreconstructed as a lepton, we remove events having $\not\!\!E_T$ aligned or antialigned in azimuth with the lepton.

Because signal events contain b quarks, we require that one or two of the jets in each event be identified as a b jet. To identify b jets, a multivariate technique is used that discriminates the b jets from jets produced by light quarks and gluons [25]. Different criteria are applied to select events with one or two b jets such that the efficiency to identify b jets is 53% per jet when only one b tag is required, and around 65% per jet when two jets are tagged in the event. The light-jet mistag probabilities in these two cases correspond on average to 0.8% and 2.9% per jet, respectively. The mistag probability for c jets is on average less than 20% per jet in the one b-tag channels, and 30% per jet in the two b-tag channels.

We separate the data into four independent channels based on the number of reconstructed jets (two or three) and the number of b-tagged jets (one or at least two). Each of these channels has a different signal to background ratio, and by keeping them independent we improve the analysis sensitivity. After b tagging, the dominant backgrounds are W+jets (63% of the total background) and $t\bar{t}$ events (23%), which, respectively, tend to have lower $H_{\rm T}$ and larger $H_{\rm T}$ values than single top quark events.

We use Monte Carlo (MC) generators to simulate the kinematics of signal and background events, except for the multijet events that are obtained from data. Single-top quark signal events are simulated by the SINGLETOP event generator, which is based on effectively next-to-leading order (NLO) COMPHEP calculations and preserves the spin information in the decays of the top quark and the W boson [26]. The simulated event kinematics match the distributions predicted by NLO calculations [27, 28]. The $t\bar{t}$, W+jets, and Z+jets events are simulated with the ALPGEN leading-order MC generator [29]. Diboson processes are modeled using PYTHIA [30]. For all these MC samples, PYTHIA is also used to evolve parton showers and to model proton remnants and hadronization of all generated partons. The top quark mass in single top events and $t\bar{t}$ events is set to $m_t = 172.5$ GeV, which is within the experimental

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

uncertainty of the current world average [1]. A leading-order parton distribution function, CTEQ6L1 [31], is used for all MC simulated samples, except for the t-channel process, which employs the NLO parton distribution function CTEQ6M1 to ensure the final kinematics match those calculated at NLO for that process. In addition, the factorization scale is chosen as m_t for the s channel, and $m_t/2$ for the t channel, as prescribed in Ref. [32]. The presence of additional $p\bar{p}$ interactions is modeled by overlaying events selected from random beam crossings matching the instantaneous luminosity profile in the data. All MC events are processed through a GEANT-based simulation [33] of the D0 detector, and are reconstructed using the same software as the collider data.

Differences between simulation and data in lepton and jet reconstruction efficiencies and resolutions, jet energy scale, and b-tagging efficiencies are corrected in the simulation by applying correction functions measured from separate data samples. The multijet background is modeled from data by selecting events that pass the selection described above but fail the isolation criteria for leptons. The W+jets and multijet backgrounds are normalized to data before b tagging using the matrix method [19]. All other MC simulated samples are normalized to the theoretical cross section at NNLO [34] for $t\bar{t}$, and at NLO [35] for Z+jets and diboson production.

Before the overall rate of W+jets production is normalized to data [19], the ratio of W+heavy flavor jets (b or c) to W+light jets is set from NLO calculations [35], which correct the ALPGEN production cross section by a factor 1.47 for $W+b\bar{b}$ and $W+c\bar{c}$ production, and by 1.65 for $W+c$+jets. These values are consistent with dedicated D0 measurements [36].

To properly describe the kinematics of a W+jets enriched sample and of the W+jets dominated region, we renormalize the simulation in the last two bins of the b-tagging multivariate output (0.90,0.95) and (0.95,1.00). The correction factor is derived from a sample that has low values of the matrix element discriminant, described below. These events are highly depleted in signal (signal fraction <1% after b tagging), and are dominated by W+jets production in the two-jet channel. They provide enough statistics to derive a correction factor valid also for the highest bins of the b-tagging multivariate output, as shown in Fig. 2(d). These events are not used in the subsequent measurements. This correction scales down the simulated samples by an average factor of 0.80 ± 0.08, where the uncertainty is statistical only. The total uncertainty assigned to this normalization is 20%, which is consistent with studies in independent data sets with no b-tagged jets, and with fits of data to the b-tagging output of the background components in different channels.

Table I lists the numbers of expected and observed events for each process after event selection including b tagging. Overall, the total combined acceptance including the branching fraction, event selection, and b tagging, is 2.6% for s channel and 1.8% for t channel.

Figure 2 shows comparisons between data and simulation after applying b tagging, with all corrections included. In the same figure, the normalization and differential spectra of the two dominant backgrounds are checked using the control samples dominated by W+jets events (Fig. 2(e)), and by $t\bar{t}$ events (Fig. 2(f)). These plots demonstrate the accuracy of the background modeling.

There is no single kinematic variable that allows for the efficient isolation of the single top quark signal from the large backgrounds. We therefore perform the following three separate multivariate analyses (MVA) and then combine their results into one final MVA: (i) matrix element (ME) [37], (ii) Bayesian neural networks (BNN) [38], and (iii) boosted decision trees (BDT) [39]. The final combination of these three separation techniques is performed using a BNN. By combining several input variables, each method defines a discriminant output variable D between 0 and 1 where the signal tends to be in the high discriminant region ($D \approx 1$). The output D achieves better signal separation than any single kinematic variable and is used to extract the signal in the high discriminant region by fitting the data to the sum of the signal and background models, with the signal and background normalization as free parameters. The background normalization is thus constrained by the data with low discriminant values.

For all three MVA methods, we use the same data and the same model for background, perform the analyses separately on the four mutually exclusive channels defined previously, and consider the same sources of systematic uncertainty. All MVA methods produce discriminants for the s channel as signal, D_s, where the t channel is treated as another background, and for the t-channel signal, D_t, where the s channel is considered in the background category. The three methods differ, however, in the discriminating variables.

The ME technique uses the theory that governs the production of signal and background events to separate them. The matrix element $\mathcal{M}$ for a given process contains all the dynamics of the hard scattering, where a collision between two initial partons p_1 and p_2 produces the final state partons described by their four-momenta k. Thus, the differential cross section for a given process $p_1p_2 \to k$ is proportional to the magnitude squared of the ME for that process: $d\sigma/dk \propto |\mathcal{M}(p_1p_2 \to k)|^2$. The ME method uses the probabilities derived from these differential cross sections to create a discriminant that potentially uses all the kinematic information available for the event. In our background probability calculation, we include the MEs for all backgrounds, including multijets, as described by the dominant leading-order Feynman diagrams obtained from MADGRAPH [40].

Additional details about the specific implementation of the ME method for this analysis can be found in Ref. [41]. In this analysis, we have improved several aspects of the method with respect to the previous implementation [19,

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

TABLE I: The numbers of expected and observed events in a data sample corresponding to 9.7 fb^{-1} of integrated luminosity, with uncertainties including both statistical and systematic components. The s and t channel contributions are normalized to their SM expectations for $m_t = 172.5$ GeV. The ratio $S(tb)$:B is the ratio of the number of s-channel signal events, S, to the total number of background events, B, including the t channel, and $S(tqb)$:B is the ratio of the number of t-channel signal events to the total number of background events, including the s channel.

Number of jets	2	2	3	3
Number of b tags	1	2	1	2
s channel	112±23	83±19	33±7	29±7
t channel	248±50	23±5	75±15	32±7
$t\bar{t}$	585±100	275±52	1044±207	767±158
W+jets	4984±369	715±96	1395±120	300±39
Z+jets and diboson	544±67	79±10	156±18	36±5
Multijet	479±73	65±10	188±33	56±9
Background sum	6592±395	1134±110	2784±242	1160±164
Backgrounds + signals	6952±399	1240±112	2891±243	1220±164
Data	6859	1286	2725	1233
$S(tb)$:B	1:61	1:14	1:88	1:41
$S(tqb)$:B	1:27	1:52	1:38	1:38

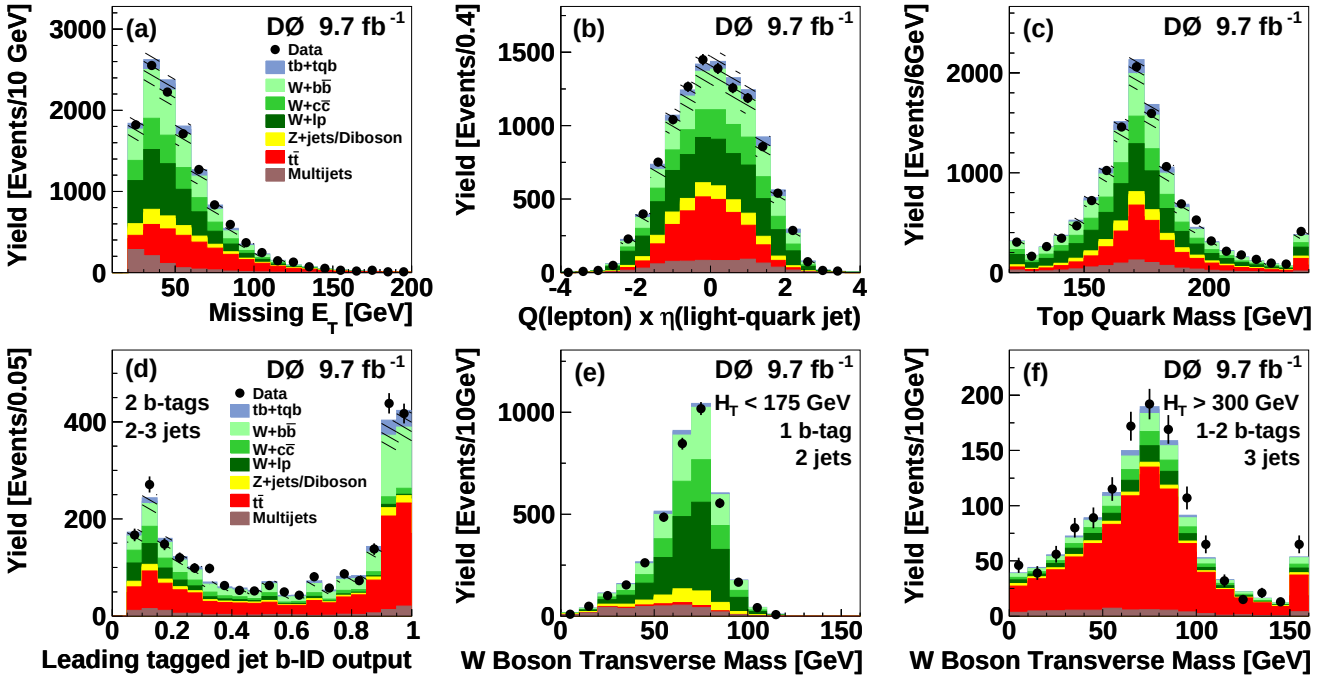


FIG. 2: [color online] Comparisons between the data and the background model for all channels combined for (a) missing transverse energy $\not\!\!E_T$, (b) product of light quark jet pseudorapidity and lepton charge, (c) reconstructed m_t from adding the W boson and b-tagged jet that give the best fit to $m_t = 172.5$ GeV, and (d) b-tagging multivariate output for the most b-like jet after the heavy flavor correction in the two b-tag channel. The selection threshold for one tagged jet is set at 0.225, and for two b jets is set at 0.075 each in this distribution. Also shown is the W boson transverse mass (e) in a control sample dominated by W+jets, and (f) in a control sample dominated by $t\bar{t}$ pairs. The s- and t-channel contributions are normalized to their SM expectations for $m_t = 172.5$ GeV. Underflow and overflow entries are added to the first and last bins, respectively. The hatched bands show the ±1 SD uncertainty on the background prediction after fitting the signal and backgrounds to the data in the selected sample (a–d).

20]:

–The $t\bar{t}$ process produces a six parton final state: $\ell^{\pm}\nu bq\bar{q}'\bar{b}$ or $\ell^{+}\nu b\ell^{-}\bar{\nu}\bar{b}$, but the analyzed final state contains at most five partons. We could integrate over the phase space of the extra partons in each event, but we instead choose to match each parton to a reconstructed

object in our final state to speed up the calculations. We find that the missing jets are most frequently light quark jets originating from the W boson decay. In the two-jet channel the W boson decaying hadronically is therefore assumed to be lost and is integrated over with a prior obtained from a simulation of $t\bar{t}$ events with two reconstructed jets. In the three-jet channel, the optimal procedure is to assign the W boson momentum before it decays to the third (light) jet with a corresponding transfer function that takes into account the average energy carried away by the lost jet.

–The transfer functions that relate the reconstructed jet energy to the parton-level energy have been updated to provide improved modeling of energy resolutions. We treat jets misidentified as electrons, light jets, b jets, and b jets with muon decays inside the jet as separate categories for each jet transfer function.

–New discriminants have been introduced that incorporate the b-tagging information for each jet into the ME probabilities to improve the characterization of each event. In the t-channel discriminant each jet-parton permutation is assigned a weight based on the b-tagging output of the jet. In the s-channel discriminant all jet-parton permutations have equal weights. The overall probability is increased if the b-tagging information of the jets in the event matches the expected number of b jets for each ME process. In this case, the added b-tagging information helps in discriminating the signals from backgrounds that contain light jets.

The BNN and BDT methods are different from the ME method because they rely on the simulated samples to characterize the signals and backgrounds, instead of using the ME for each process. The BNN and BDT follow the procedure established in the previous measurement [9]. The selected sample is divided into three different subsamples: a quarter of the events is used for the training sample used to characterize the signal and background distributions in the BNN and BDT; a quarter is set aside for the training of the combination method (which will later combine the ME, BNN and BDT results); and the remaining half is used to check the convergence, measure the cross sections, and display the distributions of all variables. A more detailed description of this analysis is given in Refs. [42, 43].

A neural network is based on a set of non-linear functions that approximate a real function of one or more variables. Neural networks are trained to approximate the optimal discriminant that separates the signal from the background. We use a Bayesian approach to scan over many different neural networks to find the best discriminant [38]. The optimal neural network is found by averaging over the parameters that define each neural network, and by assigning a probability to each configuration [42]. The BNN uses the momentum of the lepton, $\not\!\!E_T$, and the momenta of the jets as input variables. For each jet, the b-tag multivariate output is also used. In addition, two variables are added that improve the performance of the discriminant: the transverse mass of the W boson, reconstructed from the lepton and the $\not\!\!E_T$, and the product of the leading untagged jet η and the lepton charge, $Q(\ell) \times \eta(q)$, which characterizes the forward production in the t channel. For the channel with two jets and two b tags, this variable is not used. In total, the BNN uses 14 variables in the two-jet channel, and 18 variables in the three-jet channel.

Decision trees classify events by sequentially applying selection criteria leading to several disjoint subsets of events, each with different signal purity [39]. The decision tree is built by creating two branches for the most optimal selection criterion amongst the list of input variables for the given data, and repeating this procedure with each subsequent subset. "Boosting" is the retraining of a previous decision tree by increasing the weight of those events that are misclassified in the parent tree, such that the new tree will focus more on signal events with low discriminant values and background events with high discriminant values. The input variables to use for the BDT are selected by ranking a large set of well modeled variables in order of separation power optimized for the s-channel signal for all channels combined, and then selecting the best 30 variables. We check that all 30 input variables are well modeled to obtain good agreement between data and simulation in the discriminants.

All three MVAs achieve similar discrimination between signal and background events, and their discriminants show good agreement of the background expectation with the data in the background dominated regions. Using ensembles of simulated datasets containing contributions from background and signal, we infer that the pairwise correlations among the outputs of the individual MVA methods are $\approx 75\%$. Sensitivity can therefore be increased by combining the methods to form a new discriminant [20]. To achieve maximum sensitivity, a second BNN is used to construct a combined discriminant for s- and t-channel signals, defined as $D_s^{\rm comb}$ and $D_t^{\rm comb}$, for each analysis channel. The new BNN takes as input variables the three discriminants of ME, BNN, and BDT methods for the corresponding signal, and is trained on the remaining, independent, quarter of the selected sample. Figure 3 shows that the $D_s^{\rm comb}$ and $D_t^{\rm comb}$ distributions display agreement between the data and the expected background plus measured signal over the entire discriminant range.

Systematic uncertainties are categorized in two classes: one only affecting the overall normalization, and the other affecting both the normalization and the kinematic distributions and therefore the discriminant distributions. Table II provides a summary of the systematic uncertainties. The most important ones are due to the W/Z+jets heavy flavor corrections, which include uncertainties on the NLO scaling, and on the correction applied to the b-tag discriminant from the control sample; the b-tagging efficiency uncertainty and scale factors; and the uncertainties on some of the cross sections for backgrounds.

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

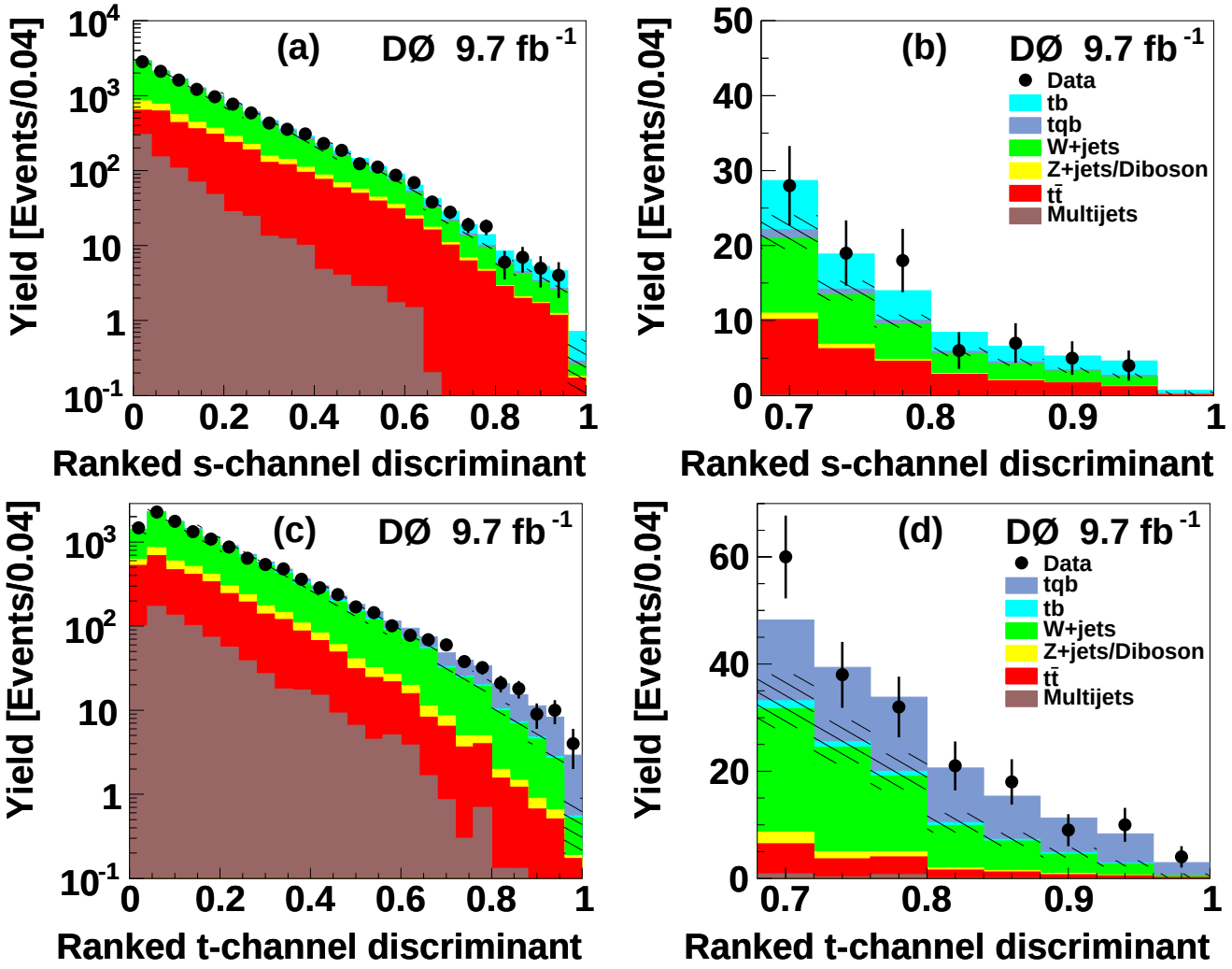


FIG. 3: [color online] The (a) $D_s^{\rm comb}$ and (c) $D_t^{\rm comb}$ discriminants for all analysis channels combined, with the high discriminant region shown in (b) and (d), respectively. The bins have been ranked by their expected signal to background ratio. The signal is normalized to the observed cross section. The signal contributions are visible above the hatched bands that show the ± 1 SD uncertainty on the background prediction after the fit to the data.

We use a Bayesian approach [6, 18, 19] to extract the production cross sections. The method consists of forming a binned likelihood as a product of all four analysis channels (two or three jets with one or two b tags) on the bins of the full discriminant distributions. We use the two discriminants $D_s^{\rm comb}$ and $D_t^{\rm comb}$ simultaneously in a joint discriminant sensitive to both signals, which makes the measurements of the single top quark cross sections σ_s and σ_t correlated. We assume a Poisson distribution for the number of events in each bin and uniform prior probabilities for positive values of the signal cross sections. Systematic uncertainties and their correlations are taken into account by integrating over signal acceptances, background yields, and integrated luminosity, assuming a Gaussian prior for each source of systematic uncertainty. A two dimensional (2D) posterior probability density is constructed as a function of σ_s and σ_t, with the position of the maximum defining the value of the cross sections, and the width of the distribution in the minimal region that encompasses 68% of the entire area defining the uncertainty (statistical and systematic components combined). The expected cross sections are obtained by setting the number of data events in each channel equal to the value given by the prediction of SM signal plus background.

Several cross checks have been performed to demonstrate the stability of the MVA methods and the Bayesian extraction of the cross section, and to ensure the reliability of the measurements. We generate ensembles of pseudo-experiments taking into account all systematic uncertainties and their correlations, injected with varying amounts of signal events. Each pseudo-experiment is analyzed with each of the MVA methods, following the same analysis chain as for the data, and the signal cross section is extracted. The cross sections extracted by all three methods behave linearly as a function of

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

TABLE II: A summary of the dominant relative systematic uncertainties. The uncertainties shown correspond to the error on each of the individual sources and are not the uncertainties on the final cross section.

Relative Systematic Uncertainties	
Components for Normalization	
Integrated luminosity [44]	6.1%
$t\bar{t}$ cross section	9.0%
Parton distribution functions	2.0%
Trigger efficiency	(3.0-5.0)%
Jet fragmentation and higher-order effects	(0.7-7.0)%
Initial and final state radiation	(0.8-10.9)%
W/Z+jets heavy-flavor correction	20.0%
W+jets normalization to data	(1.1-2.5)%
Multijet normalization to data	(9.2-42.1)%
Components for Normalization and Shape	
Jet reconstruction and identification	(0.1-1.4)%
Jet energy resolution	(0.3-1.1)%
Jet energy scale	(0.1-1.2)%
Flavor-dependent jet energy scale	(0.1-1.3)%
b tagging, single-tagged	(1.0-6.6)%
b tagging, double-tagged	(7.3-8.8)%

the input signal cross section. The same behavior is found for the combination BNN. Results of these pseudo-experiments demonstrate insignificant biases. We test the MVA methods in the two cross-check regions in the data, enriched in W+jets and $t\bar{t}$ events, and the discriminants show good agreement with the background expectation in these background dominated samples. Finally, we also check the distribution of the data sample when different regions of the discriminants are selected with increasing amounts of signal purity, and show that the presence of a single top quark signal is needed to ensure a good description of the data in different kinematic variables.

Figure 4 shows contours of equal probabilities for a given number of standard deviations in the 2D posterior for the combined discriminant. The figure also shows the sensitivity to some models of BSM physics that would change the s- or t-channel cross sections. To measure the uncertainty on the individual cross sections, we obtain the one-dimensional (1D) posterior probability functions by integrating the 2D posterior over the other variable. To measure the combined $s+t$ cross section σ_{s+t} without assuming the SM ratio of σ_s/σ_t, a 2D posterior of σ_{s+t} versus σ_t is first formed and then the 1D estimate of σ_{s+t} found by integrating over all possible values of σ_t. The results of these measurements are summarized in Table III.

All three measurements are in agreement with the SM predictions within the uncertainties [4]. The statistical significance of these results is quantified by a p value, which represents the likelihood that the measured cross section could result from a background-only fluctuation equal to or greater than the observed value, assuming the signal process is absent. An asymptotic log-likelihood ratio approach [47] is adopted to quantify the p values with the results summarized in Table III. The s-channel cross section, without any assumption on the t-channel cross section, is measured with a significance corresponding to 3.7 SD, which is also the expected sensitivity of our analysis for this process. This is the first measurement of the s-channel cross section at more than 3 SD. The t-channel cross section is measured with 7.7 SD (6.0 SD expected). The relative uncertainty on the $s+t$ cross section measurement is improved by 40% with respect to the previous D0 measurement [8], and is now 14%, including both statistical and systematic components. The statistical component is dominant: the result without systematic uncertainties has a relative uncertainty of 11%. The experimental dependence of σ_s on the assumed value of m_t is -0.08 pb/GeV, and for σ_t is -0.04 pb/GeV.

The single top quark production cross section is directly proportional to the square of the CKM matrix element $|V_{tb}|^2$, enabling us to measure $|V_{tb}|$ directly without any assumption on the number of quark families or the unitarity of the CKM matrix [19]. We assume only SM processes for single top quark production and top quarks to decay exclusively to Wb. We also assume that the Wtb interaction is CP-conserving and of the type $V-A$, but maintain the possibility for an anomalous strength of the left-handed Wtb coupling (f_1^L), which could rescale the single top quark cross section [48]. Therefore, we are measuring the strength of the $V-A$ coupling $|V_{tb}f_1^L|$, which can be greater than 1.

We start from the same combination BNN discriminants for s and t channels, and form a Bayesian posterior probability density for $|V_{tb}f_1^L|^2$ with a flat prior, without any assumption on the σ_s/σ_t production ratio. Additional theoretical uncertainties are considered for the s- and t-channel cross sections [4]. We obtain $|V_{tb}f_1^L| = 1.12^{+0.09}_{-0.08}$. If we restrict the prior to the SM region [0,1] and assume $f_1^L = 1$, we extract a limit of $|V_{tb}| > 0.92$ at 95% C.L.

In summary, we have measured the single top quark production cross section using the full Run II dataset collected by the D0 experiment at the Fermilab Tevatron Collider, corresponding to an integrated luminosity of 9.7 fb^{-1} after application of appropriate data quality requirements. We measure the cross sections for s channel and t channel independently, assuming $m_t = 172.5$ GeV:

$$\sigma(p\bar{p} \to tb + X) = 1.10^{+0.33}_{-0.31} \text{ pb},$$

$$\sigma(p\bar{p} \to tqb + X) = 3.07^{+0.54}_{-0.49} \text{ pb}.$$

With no assumption on the relative s- and t-channel contributions, we measure the total single top quark

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

TABLE III: The expected and observed single top quark cross sections and p values. Here, Z is defined such that a Z standard-deviation upward fluctuation of a Gaussian random variable would have an upper tail area equal to the p value.

Channel	Expected σ (pb)	Observed σ (pb)	Expected p value	Observed p value	Expected Z	Observed Z
s	$1.07^{+0.32}_{-0.30}$	$1.10^{+0.33}_{-0.31}$	1.0×10^{-4}	1.0×10^{-4}	3.7	3.7
t	$2.33^{+0.47}_{-0.44}$	$3.07^{+0.54}_{-0.49}$	1.0×10^{-9}	7.1×10^{-15}	6.0	7.7
$s+t$	$3.34^{+0.53}_{-0.49}$	$4.11^{+0.60}_{-0.55}$				

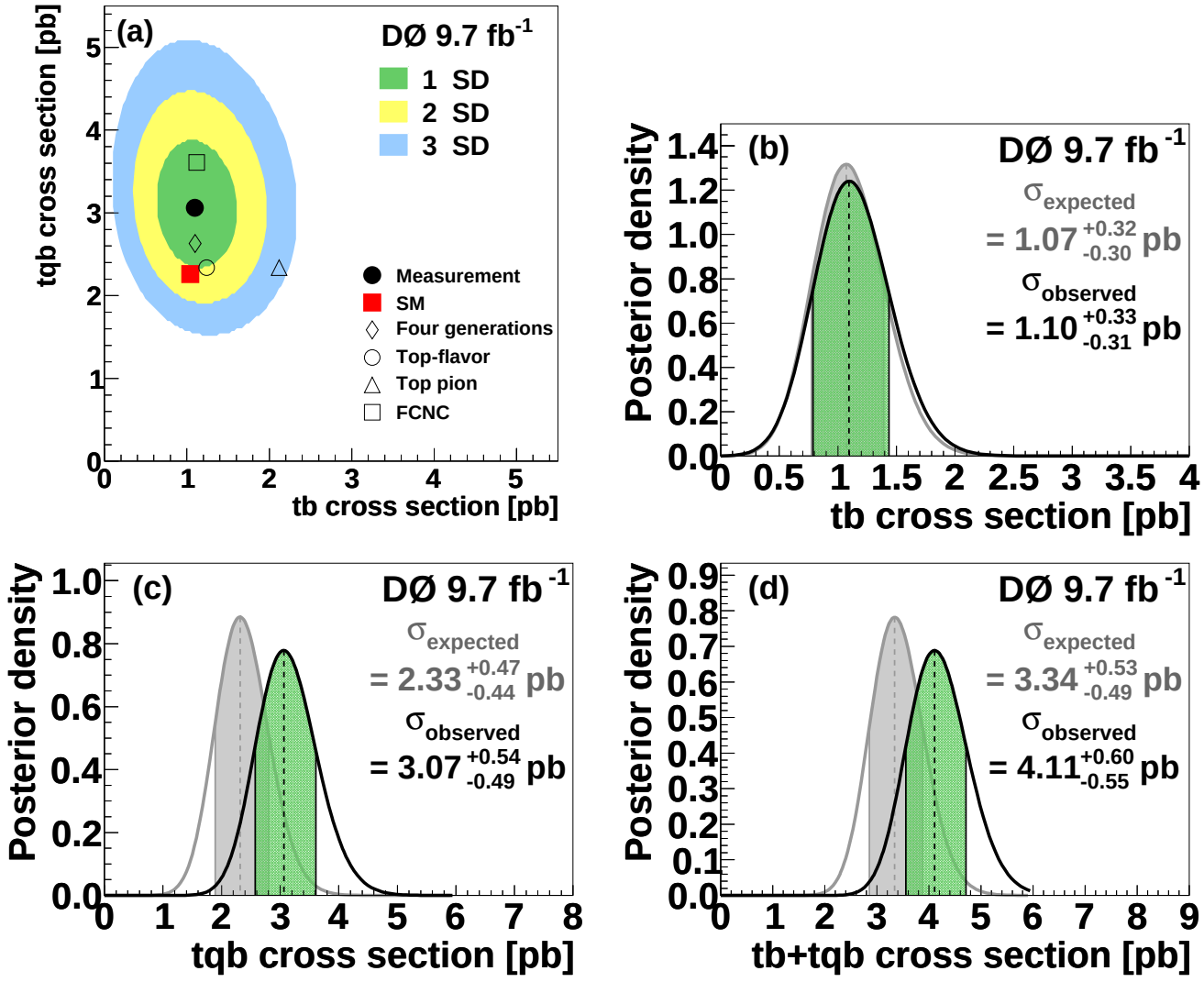


FIG. 4: [color online] Observed posterior density distributions extracted from the $D_s^{\rm comb}$ and $D_t^{\rm comb}$ discriminants: (a) 2D posterior density with one, two, and three SD probability contours, and the resulting 1D posterior densities for (b) the s channel, (c) the t channel, and (d) $s+t$ channel, all along with expected posterior densities. The prediction from the SM [4] together with several BSM predictions are shown in (a), including four-quark generations with CKM matrix element $|V_{ts}| = 0.2$ [16], a top-flavor model with new heavy bosons with $m_x = 1$ TeV [45], a model of charged top-pions with $m_{\pi^\pm} = 250$ GeV [45], and flavor changing neutral currents with an up-quark/top-quark/gluon coupling $\kappa_u/\Lambda = 0.036$ [45, 46].

production cross section to be:

$$\sigma(p\bar{p} \to tb + tqb + X) = 4.11^{+0.60}_{-0.55}\ \text{pb}.$$

All measurements are consistent with the SM predictions [4]. The s-channel production is measured with a significance of 3.7 SD and represents the first evidence

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

for this production mode. Finally, we derive a direct limit on the CKM matrix element, $|V_{tb}| > 0.92$ at 95% C.L., assuming a flat prior within $0 \leq |V_{tb}|^2 \leq 1$.

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); MON, NRC KI and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); NRF (Korea); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China).

[1] T. Aaltonen *et al.* (CDF and D0 Collaborations), Phys. Rev. D **86**, 092003 (2012); Tevatron electroweak working group (CDF and D0 Collaborations), arXiv:1305.3929 [hep-ex].
[2] S. S. D. Willenbrock and D. A. Dicus, Phys. Rev. D **34**, 155 (1986).
[3] C.-P. Yuan, Phys. Rev. D **41**, 42 (1990).
[4] N. Kidonakis, Phys. Rev. D **74**, 114012 (2006). The cross sections for the single top quark processes (m_t = 172.5 GeV) are 2.26 ± 0.12 pb (t channel), 1.04 ± 0.08 pb (s channel), and 0.28 ± 0.06 pb (tW channel).
[5] S. Cortese and R. Petronzio, Phys. Lett. B **253**, 494 (1991).
[6] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **103**, 092001 (2009).
[7] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103**, 092002 (2009).
[8] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **84**, 112001 (2011).
[9] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **705**, 313 (2011).
[10] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **717**, 330 (2012).
[11] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. **12**, 035 (2012).
[12] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716** 142 (2012).
[13] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **110**, 022003 (2013).
[14] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963); M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[15] J. Beringer *et al.* (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).
[16] J. Alwall *et al.*, Eur. Phys. J. C **49**, 791 (2007).
[17] G. V. Jikia and S. R. Slabospitsky, Phys. Lett. B **295**, 136 (1992).
[18] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **98**, 181802 (2007).
[19] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **78**, 012005 (2008).
[20] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **682**, 363 (2010).
[21] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res. Sect. A **565**, 463 (2006); M. Abolins *et al.*, Nucl. Instrum. Methods in Phys. Res. Sect. A **584**, 75 (2008); R. Angstadt *et al.*, Nucl. Instrum. Methods in Phys. Res. Sect. A **622**, 298 (2010); S. N. Ahmed *et al.*, Nucl. Instrum. Methods in Phys. Res. Sect. A **634**, 8 (2011).
[22] The pseudorapidity is defined as $\eta = -\ln[\tan(\theta/2)]$, where θ is the polar angle relative to the proton beam direction.
[23] G. C. Blazey *et al.*, arXiv:hep-ex/0005012. The cone radius is defined as $\mathcal{R} = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$.
[24] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **85**, 052006 (2012).
[25] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods in Phys. Res. Sect. A **620**, 490 (2010).
[26] E. E. Boos *et al.*, Phys. Atom. Nucl. **69**, 1317 (2006). We use SINGLETOP version 4.2p1.
[27] Z. Sullivan, Phys. Rev. D **70**, 114012 (2004).
[28] J. M. Campbell, R. Frederix, F. Maltoni and F. Tramontano, Phys. Rev. Lett. **102**, 182003 (2009).
[29] M. L. Mangano *et al.*, J. High Energy Phys. **07**, 001 (2003). We use ALPGEN version 2.11.
[30] T. Sjöstrand, S. Mrenna, and P. Skands, J. High Energy Phys. **05**, 026 (2006). We use PYTHIA version 6.409.
[31] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, J. High Energy Phys. **07**, 012 (2002)
[32] E. Boos, L. Dudko, T. Ohl, Eur. Phys. J. **C** 11 , 473 (1999).
[33] R. Brun and F. Carminati, CERN Program Library Long Writeup, Report No. W5013, 1993.
[34] S. Moch and P. Uwer, Phys. Rev. D **78**, 034003 (2008). At $m_t = 172.5$ GeV, $\sigma(p\bar{p} \to t\bar{t} + X) = 7.46 \pm 0.75$ pb.
[35] R. K. Ellis, Nucl. Phys. Proc. Suppl. **160**, 170 (2006). We use MCFM version 5.1.
[36] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **718**, 1314 (2013).
[37] K. Kondo, J. Phys. Soc. Jap. **57**, 4126 (1988); K. Kondo, J. Phys. Soc. Jap. **60**, 836 (1991);
[38] R. M. Neal, *Bayesian Learning for Neural Networks* (Springer-Verlag, New York, 1996).
[39] L. Breiman *et al.*, *Classification and Regression Trees* (Wadsworth, Stamford, 1984).
[40] F. Maltoni and T. Stelzer, J. High Energy Phys. **02**, 027 (2003).
[41] Y.-T. Tsai, Ph.D. thesis, University of Rochester. FERMILAB-THESIS-2013-09.
[42] J. Joshi, Ph.D. thesis, Panjab University. FERMILAB-THESIS-2012-18.
[43] J.A. Benitez, Ph.D. thesis, Michigan State University. FERMILAB-THESIS-2009-31.
[44] T. Andeen *et al.*, FERMILAB-TM-2365 (2007).
[45] T. Tait and C.-P. Yuan, Phys. Rev. D **63**, 014018 (2001).
[46] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **99**, 191802 (2007).
[47] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. C **71**, 1554 (2011).
[48] G. L. Kane, G.A. Ladinsky, and C.-P. Yuan, Phys. Rev. D **45**, 124 (1992).

Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.