

Multivariate searches for single top quark production with the D0 detector

V. M. Abazov,³⁶ B. Abbott,⁷⁶ M. Abolins,⁶⁶ B. S. Acharya,²⁹ M. Adams,⁵² T. Adams,⁵⁰ M. Agelou,¹⁸ J.-L. Agram,¹⁹ S. H. Ahn,³¹ M. Ahsan,⁶⁰ G. D. Alexeev,³⁶ G. Alkhalaf,⁴⁰ A. Alton,⁶⁵ G. Alverson,⁶⁴ G. A. Alves,² M. Anastasoia,³⁵ T. Andeen,⁵⁴ S. Anderson,⁴⁶ B. Andrieu,¹⁷ M. S. Anzelc,⁵⁴ Y. Arnaud,¹⁴ M. Arov,⁵³ A. Askew,⁵⁰ B. Åsman,⁴¹ A. C. S. Assis Jesus,³ O. Atramentov,⁵⁸ C. Autermann,²¹ C. Avila,⁸ C. Ay,²⁴ F. Badaud,¹³ A. Baden,⁶² L. Bagby,⁵³ B. Baldin,⁵¹ D. V. Bandurin,³⁶ P. Banerjee,²⁹ S. Banerjee,²⁹ E. Barberis,⁶⁴ P. Bargassa,⁸¹ P. Baringer,⁵⁹ C. Barnes,⁴⁴ J. Barreto,² J. F. Bartlett,⁵¹ U. Bassler,¹⁷ D. Bauer,⁴⁴ A. Bean,⁵⁹ M. Begalli,³ M. Begel,⁷² C. Belanger-Champagne,⁵ A. Bellavance,⁶⁸ J. A. Benitez,⁶⁶ S. B. Beri,²⁷ G. Bernardi,¹⁷ R. Bernhard,⁴² L. Berntzon,¹⁵ I. Bertram,⁴³ M. Besançon,¹⁸ R. Beuselinck,⁴⁴ V. A. Bezzubov,³⁹ P. C. Bhat,⁵¹ V. Bhatnagar,²⁷ M. Binder,²⁵ C. Biscarat,⁴³ K. M. Black,⁶³ I. Blackler,⁴⁴ G. Blazey,⁵³ F. Blekman,⁴⁴ S. Blessing,⁵⁰ D. Bloch,¹⁹ K. Bloom,⁶⁸ U. Blumenschein,²³ A. Boehnlein,⁵¹ O. Boeriu,⁵⁶ T. A. Bolton,⁶⁰ E. Boos,³⁸ F. Borchering,⁵¹ G. Borissov,⁴³ K. Bos,³⁴ T. Bose,⁷⁸ A. Brandt,⁷⁹ R. Brock,⁶⁶ G. Brooijmans,⁷¹ A. Bross,⁵¹ D. Brown,⁷⁹ N. J. Buchanan,⁵⁰ D. Buchholz,⁵⁴ M. Buehler,⁸² V. Buescher,²³ V. Bunichev,³⁸ S. Burdin,⁵¹ S. Burke,⁴⁶ T. H. Burnett,⁸³ E. Busato,¹⁷ C. P. Buszello,⁴⁴ J. M. Butler,⁶³ S. Calvet,¹⁵ J. Cammin,⁷² S. Caron,³⁴ W. Carvalho,³ B. C. K. Casey,⁷⁸ N. M. Cason,⁵⁶ H. Castilla-Valdez,³³ S. Chakrabarti,²⁹ D. Chakraborty,⁵³ K. M. Chan,⁷² A. Chandra,⁴⁹ D. Chapin,⁷⁸ F. Charles,¹⁹ E. Cheu,⁴⁶ F. Chevallier,¹⁴ D. K. Cho,⁶³ S. Choi,³² B. Choudhary,²⁸ L. Christofek,⁵⁹ D. Claes,⁶⁸ B. Clément,¹⁹ C. Clément,⁴¹ Y. Coadou,⁵ M. Cooke,⁸¹ W. E. Cooper,⁵¹ D. Coppage,⁵⁹ M. Corcoran,⁸¹ M.-C. Cousinou,¹⁵ B. Cox,⁴⁵ S. Crépe-Renaudin,¹⁴ D. Cutts,⁷⁸ M. Cwiok,³⁰ H. da Motta,² A. Das,⁶³ M. Das,⁶¹ B. Davies,⁴³ G. Davies,⁴⁴ G. A. Davis,⁵⁴ K. De,⁷⁹ P. de Jong,³⁴ S. J. de Jong,³⁵ E. De La Cruz-Burelo,⁶⁵ C. De Oliveira Martins,³ J. D. Degenhardt,⁶⁵ F. Déliot,¹⁸ M. Demarteau,⁵¹ R. Demina,⁷² P. Demine,¹⁸ D. Denisov,⁵¹ S. P. Denisov,³⁹ S. Desai,⁷³ H. T. Diehl,⁵¹ M. Diesburg,⁵¹ M. Doidge,⁴³ A. Dominguez,⁶⁸ H. Dong,⁷³ L. V. Dudko,³⁸ L. Duflot,¹⁶ S. R. Dugad,²⁹ A. Duperrin,¹⁵ J. Dyer,⁶⁶ A. Dyshkant,⁵³ M. Eads,⁶⁸ D. Edmunds,⁶⁶ T. Edwards,⁴⁵ J. Ellison,⁴⁹ J. Elmsheuser,²⁵ V. D. Elvira,⁵¹ S. Eno,⁶² P. Ermolov,³⁸ J. Estrada,⁵¹ H. Evans,⁵⁵ A. Evdokimov,³⁷ V. N. Evdokimov,³⁹ S. N. Fatakia,⁶³ L. Feligioni,⁶³ A. V. Ferapontov,⁶⁰ T. Ferbel,⁷² F. Fiedler,²⁵ F. Filthaut,³⁵ W. Fisher,⁵¹ H. E. Fisk,⁵¹ I. Fleck,²³ M. Ford,⁴⁵ M. Fortner,⁵³ H. Fox,²³ S. Fu,⁵¹ S. Fuess,⁵¹ T. Gadfort,⁸³ C. F. Galea,³⁵ E. Gallas,⁵¹ E. Galyaev,⁵⁶ C. Garcia,⁷² A. Garcia-Bellido,⁸³ J. Gardner,⁵⁹ V. Gavrilov,³⁷ A. Gay,¹⁹ P. Gay,¹³ D. Gelé,¹⁹ R. Gelhaus,⁴⁹ C. E. Gerber,⁵² Y. Gershtein,⁵⁰ D. Gillberg,⁵ G. Ginter,⁷² N. Gollub,⁴¹ B. Gómez,⁸ K. Gounder,⁵¹ A. Goussiou,⁵⁶ P. D. Grannis,⁷³ H. Greenlee,⁵¹ Z. D. Greenwood,⁶¹ E. M. Gregores,⁴ G. Grenier,²⁰ Ph. Gris,¹³ J.-F. Grivaz,¹⁶ S. Grünendahl,⁵¹ M. W. Grünewald,³⁰ F. Guo,⁷³ J. Guo,⁷³ G. Gutierrez,⁵¹ P. Gutierrez,⁷⁶ A. Haas,⁷¹ N. J. Hadley,⁶² P. Haefner,²⁵ S. Hagopian,⁵⁰ J. Haley,⁶⁹ I. Hall,⁷⁶ R. E. Hall,⁴⁸ L. Han,⁷ K. Hanagaki,⁵¹ K. Harder,⁶⁰ A. Harel,⁷² R. Harrington,⁶⁴ J. M. Hauptman,⁵⁸ R. Hauser,⁶⁶ J. Hays,⁵⁴ T. Hebbeker,²¹ D. Hedin,⁵³ J. G. Hegeman,³⁴ J. M. Heinmiller,⁵² A. P. Heinson,⁴⁹ U. Heintz,⁶³ C. Hensel,⁵⁹ G. Hesketh,⁶⁴ M. D. Hildreth,⁵⁶ R. Hirosky,⁸² J. D. Hobbs,⁷³ B. Hoeneisen,¹² M. Hohlfield,¹⁶ S. J. Hong,³¹ R. Hooper,⁷⁸ P. Houben,³⁴ Y. Hu,⁷³ V. Hynek,⁹ I. Iashvili,⁷⁰ R. Illingworth,⁵¹ A. S. Ito,⁵¹ S. Jabeen,⁶³ M. Jaffré,¹⁶ S. Jain,⁷⁶ K. Jakobs,²³ C. Jarvis,⁶² A. Jenkins,⁴⁴ R. Jesik,⁴⁴ K. Johns,⁴⁶ C. Johnson,⁷¹ M. Johnson,⁵¹ A. Jonckheere,⁵¹ P. Jonsson,⁴⁴ A. Juste,⁵¹ D. Käfer,²¹ S. Kahn,⁷⁴ E. Kajfasz,¹⁵ A. M. Kalinin,³⁶ J. M. Kalk,⁶¹ J. R. Kalk,⁶⁶ S. Kappler,²¹ D. Karmanov,³⁸ J. Kasper,⁶³ I. Katsanos,⁷¹ D. Kau,⁵⁰ R. Kaur,²⁷ R. Kehoe,⁸⁰ S. Kermiche,¹⁵ S. Kesisoglou,⁷⁸ A. Khanov,⁷⁷ A. Kharchilava,⁷⁰ Y. M. Kharzheev,³⁶ D. Khatidze,⁷¹ H. Kim,⁷⁹ T. J. Kim,³¹ M. H. Kirby,³⁵ B. Klima,⁵¹ J. M. Kohli,²⁷ J.-P. Konrath,²³ M. Kopal,⁷⁶ V. M. Korablev,³⁹ J. Kotcher,⁷⁴ B. Kothari,⁷¹ A. Koubarovsky,³⁸ A. V. Kozelov,³⁹ J. Kozminski,⁶⁶ A. Kryemadhi,⁸² S. Krzywdzinski,⁵¹ T. Kuhl,²⁴ A. Kumar,⁷⁰ S. Kunori,⁶² A. Kupco,¹¹ T. Kurča,^{20*} J. Kvita,⁹ S. Lager,⁴¹ S. Lammers,⁷¹ G. Landsberg,⁷⁸ J. Lazoflores,⁵⁰ A.-C. Le Bihan,¹⁹ P. Lebrun,²⁰ W. M. Lee,⁵³ A. Leflat,³⁸ F. Lehner,⁴² C. Leonidopoulos,⁷¹ V. Lesne,¹³ J. Leveque,⁴⁶ P. Lewis,⁴⁴ J. Li,⁷⁹ Q. Z. Li,⁵¹ J. G. R. Lima,⁵³ D. Lincoln,⁵¹ J. Linnemann,⁶⁶ V. V. Lipaev,³⁹ R. Lipton,⁵¹ Z. Liu,⁵ L. Lobo,⁴⁴ A. Lobodenko,⁴⁰ M. Lokajicek,¹¹ A. Lounis,¹⁹ P. Love,⁴³ H. J. Lubatti,⁸³ M. Lynker,⁵⁶ A. L. Lyon,⁵¹ A. K. A. Maciel,² R. J. Madaras,⁴⁷ P. Mättig,²⁶ C. Magass,²¹ A. Magerkurth,⁶⁵ A.-M. Mangan,¹⁴ N. Makovec,¹⁶ P. K. Mal,⁵⁶ H. B. Malbouisson,³ S. Malik,⁶⁸ V. L. Malyshev,³⁶ H. S. Mao,⁶ Y. Maravin,⁶⁰ M. Martens,⁵¹ S. E. K. Mattingly,⁷⁸ R. McCarthy,⁷³ R. McCroskey,⁴⁶ D. Meder,²⁴ A. Melnitchouk,⁶⁷ A. Mendes,¹⁵ L. Mendoza,⁸ M. Merkin,³⁸ K. W. Merritt,⁵¹ A. Meyer,²¹ J. Meyer,²² M. Michaut,¹⁸ H. Miettinen,⁸¹ T. Millet,²⁰ J. Mitrevski,⁷¹ J. Molina,³ N. K. Mondal,²⁹ J. Monk,⁴⁵ R. W. Moore,⁵ T. Moulík,⁵⁹ G. S. Muanza,¹⁶ M. Mulders,⁵¹ M. Mulhearn,⁷¹ L. Mundim,³ Y. D. Mutaf,⁷³ E. Nagy,¹⁵ M. Naimuddin,²⁸ M. Narain,⁶³ N. A. Naumann,³⁵ H. A. Neal,⁶⁵ J. P. Negret,⁸ S. Nelson,⁵⁰ P. Neustroev,⁴⁰ C. Noeding,²³ A. Nomerotski,⁵¹ S. F. Novaes,⁴ T. Nunnemann,²⁵ V. O'Dell,⁵¹ D. C. O'Neil,⁵ G. Obrant,⁴⁰ V. Oguri,³ N. Oliveira,³ N. Oshima,⁵¹ R. Otec,¹⁰ G. J. Otero y Garzón,⁵² M. Owen,⁴⁵ P. Padley,⁸¹ N. Parashar,⁵⁷ S.-J. Park,⁷² S. K. Park,³¹ J. Parsons,⁷¹ R. Partridge,⁷⁸

N. Parua,⁷³ A. Patwa,⁷⁴ G. Pawloski,⁸¹ P. M. Perea,⁴⁹ E. Perez,¹⁸ K. Peters,⁴⁵ P. Pétroff,¹⁶ M. Petteni,⁴⁴ R. Piegaia,¹ M.-A. Pleier,²² P. L. M. Podesta-Lerma,³³ V. M. Podstavkov,⁵¹ Y. Pogorelov,⁵⁶ M.-E. Pol,² A. Pompoš,⁷⁶ B. G. Pope,⁶⁶ A. V. Popov,³⁹ W. L. Prado da Silva,³ H. B. Prosper,⁵⁰ S. Protopopescu,⁷⁴ J. Qian,⁶⁵ A. Quadt,²² B. Quinn,⁶⁷ K. J. Rani,²⁹ K. Ranjan,²⁸ P. A. Rapidis,⁵¹ P. N. Ratoff,⁴³ P. Renkel,⁸⁰ S. Reucroft,⁶⁴ M. Rijssenbeek,⁷³ I. Ripp-Baudot,¹⁹ F. Rizatdinova,⁷⁷ S. Robinson,⁴⁴ R. F. Rodrigues,³ C. Royon,¹⁸ P. Rubinov,⁵¹ R. Ruchti,⁵⁶ V. I. Rud,³⁸ G. Sajot,¹⁴ A. Sánchez-Hernández,³³ M. P. Sanders,⁶² A. Santoro,³ G. Savage,⁵¹ L. Sawyer,⁶¹ T. Scanlon,⁴⁴ D. Schaile,²⁵ R. D. Schamberger,⁷³ Y. Scheglov,⁴⁰ H. Schellman,⁵⁴ P. Schieferdecker,²⁵ C. Schmitt,²⁶ C. Schwanenberger,⁴⁵ A. Schwartzman,⁶⁹ R. Schwienhorst,⁶⁶ S. Sengupta,⁵⁰ H. Severini,⁷⁶ E. Shabalina,⁵² M. Shamim,⁶⁰ V. Shary,¹⁸ A. A. Shchukin,³⁹ W. D. Shephard,⁵⁶ R. K. Shivpuri,²⁸ D. Shpakov,⁶⁴ V. Siccaldi,¹⁹ R. A. Sidwell,⁶⁰ V. Simak,¹⁰ V. Sirotenko,⁵¹ P. Skubic,⁷⁶ P. Slatery,⁷² R. P. Smith,⁵¹ G. R. Snow,⁶⁸ J. Snow,⁷⁵ S. Snyder,⁷⁴ S. Söldner-Rembold,⁴⁵ X. Song,⁵³ L. Sonnenschein,¹⁷ A. Sopczak,⁴³ M. Sosebee,⁷⁹ K. Soustruznik,⁹ M. Souza,² B. Spurlock,⁷⁹ J. Stark,¹⁴ J. Steele,⁶¹ K. Stevenson,⁵⁵ V. Stolin,³⁷ A. Stone,⁵² D. A. Stoyanova,³⁹ J. Strandberg,⁴¹ M. A. Strang,⁷⁰ M. Strauss,⁷⁶ R. Ströhmer,²⁵ D. Strom,⁵⁴ M. Strovink,⁴⁷ L. Stutte,⁵¹ S. Sumowidagdo,⁵⁰ A. Sznajder,³ M. Talby,¹⁵ P. Tamburello,⁴⁶ W. Taylor,⁵ P. Telford,⁴⁵ J. Temple,⁴⁶ B. Tiller,²⁵ M. Titov,²³ V. V. Tokmenin,³⁶ M. Tomoto,⁵¹ T. Toole,⁶² I. Torchiani,²³ S. Towers,⁴³ T. Trefzger,²⁴ S. Trincas-Duvoid,¹⁷ D. Tsybychev,⁷³ B. Tuchming,¹⁸ C. Tully,⁶⁹ A. S. Turcot,⁴⁵ P. M. Tuts,⁷¹ R. Unalan,⁶⁶ L. Uvarov,⁴⁰ S. Uvarov,⁴⁰ S. Uzunyan,⁵³ B. Vachon,⁵ P. J. van den Berg,³⁴ R. Van Kooten,⁵⁵ W. M. van Leeuwen,³⁴ N. Varelas,⁵² E. W. Varnes,⁴⁶ A. Vartapetian,⁷⁹ I. A. Vasilyev,³⁹ M. Vaupel,²⁶ P. Verdier,²⁰ L. S. Vertogradov,³⁶ M. Verzocchi,⁵¹ F. Villeneuve-Seguié,⁴⁴ P. Vint,⁴⁴ J.-R. Vlimant,¹⁷ E. Von Toerne,⁶⁰ M. Voutilainen,^{68,†} M. Vreeswijk,³⁴ H. D. Wahl,⁵⁰ L. Wang,⁶² J. Warchol,⁵⁶ G. Watts,⁸³ M. Wayne,⁵⁶ M. Weber,⁵¹ H. Weerts,⁶⁶ N. Wermes,²² M. Wetstein,⁶² A. White,⁷⁹ D. Wicke,²⁶ G. W. Wilson,⁵⁹ S. J. Wimpenny,⁴⁹ M. Wobisch,⁵¹ J. Womersley,⁵¹ D. R. Wood,⁶⁴ T. R. Wyatt,⁴⁵ Y. Xie,⁷⁸ N. Xuan,⁵⁶ S. Yacoob,⁵⁴ R. Yamada,⁵¹ M. Yan,⁶² T. Yasuda,⁵¹ Y. A. Yatsunenko,³⁶ K. Yip,⁷⁴ H. D. Yoo,⁷⁸ S. W. Youn,⁵⁴ C. Yu,¹⁴ J. Yu,⁷⁹ A. Yurkewicz,⁷³ A. Zatserklyaniy,⁵³ C. Zeitnitz,²⁶ D. Zhang,⁵¹ T. Zhao,⁸³ Z. Zhao,⁶⁵ B. Zhou,⁶⁵ J. Zhu,⁷³ M. Zielinski,⁷² D. Zieminska,⁵⁵ A. Zieminski,⁵⁵ V. Zutshi,⁵³ and E. G. Zverev³⁸

(D0 Collaboration)

¹*Universidad de Buenos Aires, Buenos Aires, Argentina*

²*LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*

³*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*

⁴*Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil*

⁵*University of Alberta, Edmonton, Alberta, Canada,*

Simon Fraser University, Burnaby, British Columbia, Canada,

York University, Toronto, Ontario, Canada,

and McGill University, Montreal, Quebec, Canada

⁶*Institute of High Energy Physics, Beijing, People's Republic of China*

⁷*University of Science and Technology of China, Hefei, People's Republic of China*

⁸*Universidad de los Andes, Bogotá, Colombia*

⁹*Center for Particle Physics, Charles University, Prague, Czech Republic*

¹⁰*Czech Technical University, Prague, Czech Republic*

¹¹*Center for Particle Physics, Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic*

¹²*Universidad San Francisco de Quito, Quito, Ecuador*

¹³*Laboratoire de Physique Corpusculaire, IN2P3-CNRS, Université Blaise Pascal, Clermont-Ferrand, France*

¹⁴*Laboratoire de Physique Subatomique et de Cosmologie, IN2P3-CNRS, Université de Grenoble 1, Grenoble, France*

¹⁵*CPPM, IN2P3-CNRS, Université de la Méditerranée, Marseille, France*

¹⁶*IN2P3-CNRS, Laboratoire de l'Accélérateur Linéaire, Orsay, France*

¹⁷*LPNHE, IN2P3-CNRS, Universités Paris VI and VII, Paris, France*

¹⁸*DAPNIA/Service de Physique des Particules, CEA, Saclay, France*

¹⁹*IReS, IN2P3-CNRS, Université Louis Pasteur, Strasbourg, France,*

and Université de Haute Alsace, Mulhouse, France

²⁰*Institut de Physique Nucléaire de Lyon, IN2P3-CNRS, Université Claude Bernard, Villeurbanne, France*

²¹*III. Physikalisches Institut A, RWTH Aachen, Aachen, Germany*

²²*Physikalisches Institut, Universität Bonn, Bonn, Germany*

²³*Physikalisches Institut, Universität Freiburg, Freiburg, Germany*

²⁴*Institut für Physik, Universität Mainz, Mainz, Germany*

²⁵*Ludwig-Maximilians-Universität München, München, Germany*

- ²⁶*Fachbereich Physik, University of Wuppertal, Wuppertal, Germany*
²⁷*Panjab University, Chandigarh, India*
²⁸*Delhi University, Delhi, India*
²⁹*Tata Institute of Fundamental Research, Mumbai, India*
³⁰*University College Dublin, Dublin, Ireland*
³¹*Korea Detector Laboratory, Korea University, Seoul, Korea*
³²*SungKyunKwan University, Suwon, Korea*
³³*CINVESTAV, Mexico City, Mexico*
³⁴*FOM-Institute NIKHEF, Amsterdam, The Netherlands*
and University of Amsterdam/NIKHEF, Amsterdam, The Netherlands
³⁵*Radboud University Nijmegen/NIKHEF, Nijmegen, The Netherlands*
³⁶*Joint Institute for Nuclear Research, Dubna, Russia*
³⁷*Institute for Theoretical and Experimental Physics, Moscow, Russia*
³⁸*Moscow State University, Moscow, Russia*
³⁹*Institute for High Energy Physics, Protvino, Russia*
⁴⁰*Petersburg Nuclear Physics Institute, St. Petersburg, Russia*
⁴¹*Lund University, Lund, Sweden, Royal Institute of Technology, Stockholm, Sweden,*
Stockholm University, Stockholm, Sweden,
and Uppsala University, Uppsala, Sweden
⁴²*Physik Institut der Universität Zürich, Zürich, Switzerland*
⁴³*Lancaster University, Lancaster, United Kingdom*
⁴⁴*Imperial College, London, United Kingdom*
⁴⁵*University of Manchester, Manchester, United Kingdom*
⁴⁶*University of Arizona, Tucson, Arizona 85721, USA*
⁴⁷*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
and University of California, Berkeley, California 94720, USA
⁴⁸*California State University, Fresno, California 93740, USA*
⁴⁹*University of California, Riverside, California 92521, USA*
⁵⁰*Florida State University, Tallahassee, Florida 32306, USA*
⁵¹*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
⁵²*University of Illinois at Chicago, Chicago, Illinois 60607, USA*
⁵³*Northern Illinois University, DeKalb, Illinois 60115, USA*
⁵⁴*Northwestern University, Evanston, Illinois 60208, USA*
⁵⁵*Indiana University, Bloomington, Indiana 47405, USA*
⁵⁶*University of Notre Dame, Notre Dame, Indiana 46556, USA*
⁵⁷*Purdue University Calumet, Hammond, Indiana 46323, USA*
⁵⁸*Iowa State University, Ames, Iowa 50011, USA*
⁵⁹*University of Kansas, Lawrence, Kansas 66045, USA*
⁶⁰*Kansas State University, Manhattan, Kansas 66506, USA*
⁶¹*Louisiana Tech University, Ruston, Louisiana 71272, USA*
⁶²*University of Maryland, College Park, Maryland 20742, USA*
⁶³*Boston University, Boston, Massachusetts 02215, USA*
⁶⁴*Northeastern University, Boston, Massachusetts 02115, USA*
⁶⁵*University of Michigan, Ann Arbor, Michigan 48109, USA*
⁶⁶*Michigan State University, East Lansing, Michigan 48824, USA*
⁶⁷*University of Mississippi, University, Mississippi 38677, USA*
⁶⁸*University of Nebraska, Lincoln, Nebraska 68588, USA*
⁶⁹*Princeton University, Princeton, New Jersey 08544, USA*
⁷⁰*State University of New York, Buffalo, New York 14260, USA*
⁷¹*Columbia University, New York, New York 10027, USA*
⁷²*University of Rochester, Rochester, New York 14627, USA*
⁷³*State University of New York, Stony Brook, New York 11794, USA*
⁷⁴*Brookhaven National Laboratory, Upton, New York 11973, USA*
⁷⁵*Langston University, Langston, Oklahoma 73050, USA*
⁷⁶*University of Oklahoma, Norman, Oklahoma 73019, USA*
⁷⁷*Oklahoma State University, Stillwater, Oklahoma 74078, USA*
⁷⁸*Brown University, Providence, Rhode Island 02912, USA*

*On leave from IEP SAS Kosice, Slovakia.

†Visitor from Helsinki Institute of Physics, Helsinki, Finland.

⁷⁹*University of Texas, Arlington, Texas 76019, USA*⁸⁰*Southern Methodist University, Dallas, Texas 75275, USA*⁸¹*Rice University, Houston, Texas 77005, USA*⁸²*University of Virginia, Charlottesville, Virginia 22901, USA*⁸³*University of Washington, Seattle, Washington 98195, USA*

(Received 10 April 2006; published 29 May 2007)

We present a search for electroweak production of single top quarks in the s -channel ($p\bar{p} \rightarrow t\bar{b} + X$) and t -channel ($p\bar{p} \rightarrow tq\bar{b} + X$) modes. We have analyzed 230 pb⁻¹ of data collected with the D0 detector at the Fermilab Tevatron Collider at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV. No evidence for a single top quark signal is found. We set 95% confidence level upper limits on the production cross sections, based on binned likelihoods formed from a neural network output. The observed (expected) limits are 6.4 pb (4.5 pb) in the s -channel and 5.0 pb (5.8 pb) in the t -channel.

DOI: [10.1103/PhysRevD.75.092007](https://doi.org/10.1103/PhysRevD.75.092007)

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

I. INTRODUCTION

The top quark, discovered in 1995 at the Fermilab Tevatron Collider by the CDF and D0 collaborations [1], is by far the heaviest elementary particle found to date. Its large mass and corresponding coupling strength to the Higgs boson of order unity suggest that the physics of electroweak symmetry breaking might be visible in the top quark sector.

Top quarks are produced at the Tevatron mainly in top-antitop pairs through the strong interaction. This mode led to the discovery of the top quark and has been the only top quark production mode observed to date. The top quark decays predominantly to a W boson and a b quark, but little else is known experimentally about its electroweak interactions.

All previous studies of the top quark electroweak interaction and the Wtb vertex have been done either in the low-energy regime using virtual top quarks (in studies of b quark decays), or in the decay of real top quarks. Both of these types of studies presuppose the unitarity of the Cabibbo-Kobayashi-Maskawa (CKM) quark-mixing matrix and are thus constrained to studying the standard model with three generations of quarks. This restriction can be overcome by exploring the production of single top quarks through electroweak interactions. This production mode is becoming accessible at the Tevatron and promises the first direct measurement of the electroweak coupling strength of the top quark as well as a first glimpse at possible top quark interactions beyond the standard model (SM).

A. Physics with single top quarks

The study of single top quark production provides the possibility of investigating top quark related properties that cannot be measured in top quark pair production. The most relevant of these is a direct measurement of the CKM matrix element $|V_{tb}|$ from the single top quark production cross sections. This provides the only measurement of $|V_{tb}|$ without having to assume three quark generations or CKM matrix unitarity. Together with the other CKM matrix

measurements [2], the measurement of single top quark production will test the unitarity of the CKM matrix.

In the SM, single top quarks are produced through a left-handed interaction. Therefore, they are expected to be highly polarized. Since the top quark decays before hadronization can occur, the spin correlations are retained in the final decay products. Precise measurements of the single top quark production and decay properties offer an opportunity to observe the polarization and to test the corresponding SM predictions.

Measurements of the charged-current couplings of the top quark probe any nonstandard structure of the couplings and can therefore provide hints of new physics. Any deviation in the $(V - A)$ structure of the Wtb coupling would lead to a violation of the spin correlation properties [3]. Furthermore, combining single top quark measurements with W helicity measurements in top quark decays provides the most stringent information on the Wtb coupling [4].

Finally, rather than manifesting itself in a modified Wtb coupling, new physics could produce a single top quark final state through other processes. There are several models of new physics that would increase the single top quark production cross sections [5]. Thus, constraints on physics beyond the standard model are possible even before an actual observation of single top quark production.

B. Single top quark production

There are three standard model modes of single top quark production at hadron colliders. Each of these modes may be characterized by the four-momentum squared Q_W^2 , the virtuality, of the participating W boson:

- (i) s -channel W boson exchange ($Q_W^2 > 0$): This process, $p\bar{p} \rightarrow t\bar{b} + X$, is referred to as “ tb ,” which includes both $t\bar{b}$ and $\bar{t}b$ (see Fig. 1).
- (ii) t -channel and u -channel W boson exchange ($Q_W^2 < 0$): This process, $p\bar{p} \rightarrow tq\bar{b} + X$, has the largest cross section of the three. It includes the leading order diagram [Fig. 2(a)] with a b quark from the proton sea in the initial state and a second diagram [Fig. 2(b)] where an extra \bar{b} quark appears in the

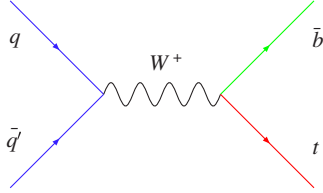


FIG. 1 (color online). Feynman diagram for leading order s -channel single top quark production.

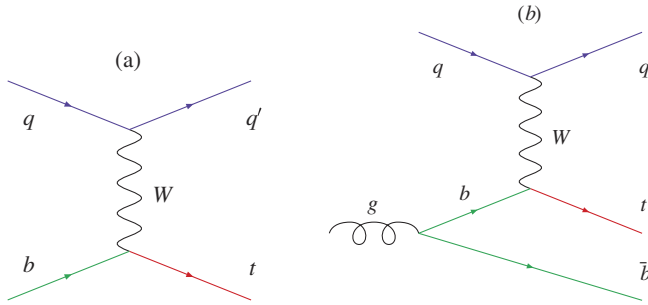


FIG. 2 (color online). Representative Feynman diagrams for t -channel single top quark production. Shown is the (a) leading order and (b) the $O(\alpha_s)$ W -gluon fusion diagram.

TABLE I. Theoretically calculated total cross sections for single top quark production at a $p\bar{p}$ collider with $\sqrt{s} = 1.96$ TeV, using $m_t = 175$ GeV.

Process	Cross section [pb]
s -channel ($t\bar{b}$)	$0.88^{+0.07}_{-0.06}$
t -channel ($tq\bar{b}$)	$1.98^{+0.23}_{-0.18}$
tW production	0.093 ± 0.024

final state explicitly. This latter mode is of order $O(\alpha_s)$ in the strong coupling α_s but, nevertheless, provides the largest contribution to the total cross section. Historically, t -channel production also has been referred to as W -gluon fusion, since the \bar{b} quark in the final state arises from a gluon splitting

to a $b\bar{b}$ pair. We refer to the t -channel process as “ $tq\bar{b}$,” which includes $tq\bar{b}$, $\bar{t}\bar{q}b$, tq , and $\bar{t}\bar{q}$.

- (iii) Real W boson production ($Q_W^2 = m_W^2$): In this process, $p\bar{p} \rightarrow tW + X$, a single top quark appears in association with a real W boson in the final state. This process has a negligible cross section at the Tevatron [3] and will not be addressed in this paper.

The next-to-leading order (NLO) production rates at the Tevatron ($\sqrt{s} = 1.96$ TeV) for the s - and t -channel single top quark modes have been calculated [6–12] and the results for the cross sections are shown in Table I. The uncertainties include components from the choice of scale and the parton distribution functions, but not for the top quark mass.

For comparison, the calculated top quark pair production cross section at the Tevatron at 1.96 TeV is 6.77 ± 0.42 pb [13]. The lower cross section and the presence of only one top quark per event make it clear that it is more difficult to isolate the single top quark signal than the top quark pair signal.

Under the assumption that all top quarks decay to a W boson and a b quark, and only using W boson decays to electron and muon final states, the final-state signature of a single top quark event detected in this analysis is characterized by a high transverse momentum (p_T), centrally produced, isolated lepton (e^\pm or μ^\pm), and missing transverse energy (\cancel{E}_T), together with two or three jets. One of the jets comes from a high- p_T central b quark from the top quark decay.

Figures 3 and 4 show the transverse momenta and pseudorapidities η [14] for the partons in our modeling of the s -channel and t -channel single top quark processes, after decay of the top quark and W boson.

The final-state fermions from the top quark decay have relatively high transverse momenta and central rapidities. Since the s -channel process involves the decay of a heavy virtual object, the \bar{b} quark produced with the top quark is also at high transverse momentum and central pseudorapidity. By contrast, the light quark in the t -channel appears at lower transverse momentum and at more forward pseudorapidities because it is produced when an initial-state

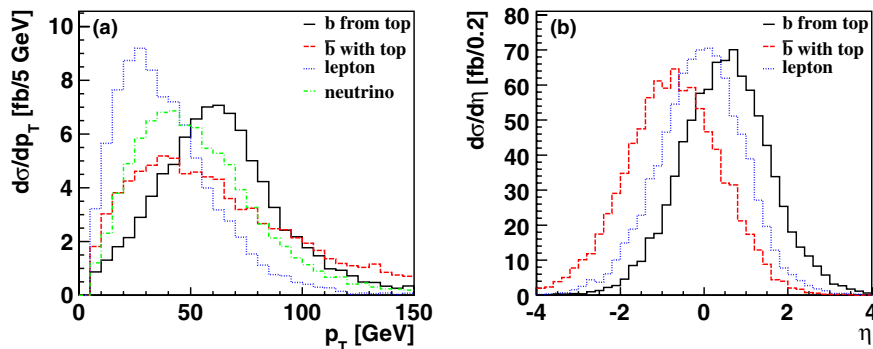


FIG. 3 (color online). Distributions of transverse momenta (a) and pseudorapidity (b) for the final-state partons in s -channel single top quark events. The histograms only include the final state of t , not \bar{t} .

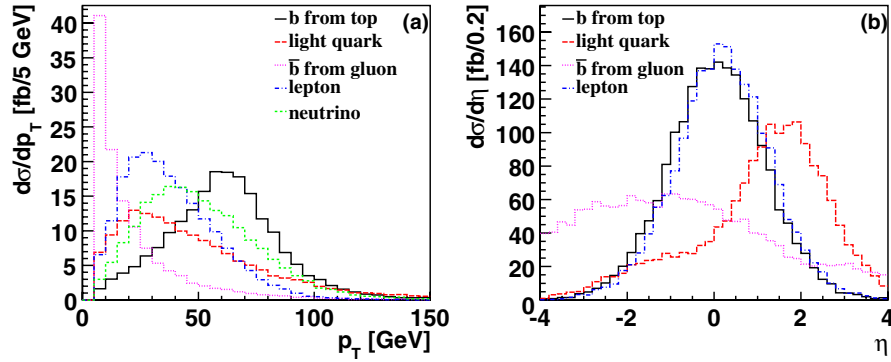


FIG. 4 (color online). Distributions of transverse momenta (a) and pseudorapidity (b) for the final-state partons in t -channel single top quark events. The histograms only include the final state of t , not \bar{t} .

parton emits a virtual W boson. The \bar{b} quark from t -channel gluon splitting appears typically at very low p_T and with large pseudorapidities and is thus often not reconstructed experimentally.

Because of its electroweak nature, single top quark production results in a polarized top quark. It has been shown [15,16] that the top quark spin follows the direction of the down-type quark momentum in the top quark rest frame. This is the direction of the initial \bar{d} quark for the s -channel and close to the direction of the final-state d quark for the t -channel.

C. Overview of the backgrounds

Searches for single top quark production are challenging because of the very large backgrounds to their lepton + jets signature. The situation is significantly different from top pair production not just because of the smaller production rate, but more importantly because of the smaller multiplicity of final-state particles (leptons or jets). Single top quark events are typically less energetic (because there is only one heavy object), less spherical (because of the production mechanism), and typically have two or three jets, not four as do $t\bar{t}$ events.

Processes that can have the same single top quark experimental signature include, in order of importance, W + jets, $t\bar{t}$, multijet production, and some smaller contributions from Z + jets and diboson events.

- (i) W + jets events form the dominant part of the background. The cross section for W + 2 jets production is over 1000 pb [17,18] with $Wb\bar{b}$ contributing about 1%.
- (ii) The second largest background is due to $t\bar{t}$ production. This process has a larger multiplicity of final-state particles than single top quark events. However, when some of the jets or a lepton are not identified, the kinematics of the remaining particles are very similar to those of the signal.
- (iii) Multijet events form a background in the electron channel when a jet is misidentified as an electron. The probability of such misidentification is rather

small, but the ≥ 3 jet cross section is so large that the overall contribution is significant.

Additionally, $b\bar{b}$ production contributes to the background when one of the b 's decays semileptonically. This background in the electron channel is very small. In the muon channel, $b\bar{b}$ events form a background when the muon is away from the jet axis or when the jet is not reconstructed.

- (iv) Z /Drell-Yan + jets production can mimic the single top quark signals if one of the leptons is misidentified.
- (v) WW , WZ , and ZZ processes are the electroweak part of the W + jets and Z + jets backgrounds, but with different kinematics.

Single top quark events are kinematically and topologically similar to W + jets and $t\bar{t}$ events. Therefore, extracting the signal from the backgrounds is challenging in a search for single top quark production.

D. Status of searches

Both the CDF and D0 collaborations have previously performed searches for single top quark production [19,20]. Recently, CDF performed a search using 160 pb^{-1} of data and obtained upper limits of 13.6 pb (s -channel), 10.1 pb (t -channel), and 17.8 pb ($s + t$ combined) at the 95% confidence level [21]. D0 has published a neural network search for single top quark production using 230 pb^{-1} of data [22], which is described in more detail in this article.

E. Outline of the analysis

We have performed a search for the electroweak production of single top quarks in the s -channel and t -channel production modes with the D0 detector at the Fermilab Tevatron Collider. We consider lepton + jets in the final state, where the lepton is either an electron or a muon.

To take advantage of one of the differences between s - and t -channel final-state topologies, we differentiate the s -channel search from the t -channel search by requiring at

least one untagged jet in the t -channel search, corresponding to the light-quark jet. For both s -channel and t -channel searches, we separate the data into independent analysis sets based on the lepton flavor (e or μ) and the multiplicity of identified b quarks (one tagged jet or more than one).

We use two different multivariate methods to extract the signal from the large backgrounds: a cut-based analysis, first presented here, and an analysis based on neural networks that was first presented in brief form in Ref. [22]. We set upper limits at the 95% C.L. on the single top quark production cross sections.

Finally, we present limit contours in a two-dimensional plane of the s -channel signal cross section versus the t -channel signal cross section.

F. Outline of the paper

This paper is organized as follows. Section II describes the D0 detector and the reconstruction of the final-state objects. Section III summarizes the triggers for the data samples used in the search and Sec. IV describes the selection requirements. Section V explains the modeling of signals and backgrounds, and Sec. VI presents the numbers of events passing all selections. Section VII discusses the most important variables that offer discrimination between the signals and backgrounds and provides details of the cut-based and the neural network analyses. Section VIII lists the systematic uncertainties in this measurement. Section IX discusses the procedure for setting limits on the signal cross section using Bayesian statistics. The limits are presented in Sec. X, and we summarize the results in Sec. XI.

II. THE D0 DETECTOR AND OBJECT RECONSTRUCTION

A. The D0 detector

The D0 detector [23] is shown in Figs. 5 and 6 and consists of several layered elements. The first is a magnetic central-tracking system, which includes a silicon micro-strip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet. The SMT has $\approx 800\,000$ individual strips, with a typical pitch of 50–80 μm and a design optimized for tracking and vertexing capability at pseudorapidities of $|\eta| < 3.0$. The system has a six-barrel longitudinal structure, each with a set of four layers arranged axially around the beam pipe and interspersed with 16 radial disks. The CFT has eight thin coaxial barrels, each supporting two doublets of overlapping scintillating fibers of 0.835 mm diameter, one doublet being parallel to the collision axis, and the other alternating by $\pm 3^\circ$ relative to the axis. Light signals are transferred via clear light fibers to solid-state photon counters (visible light photon counters, VLPCs) that have $\approx 80\%$ quantum efficiency.

Central and forward preshower detectors are located just outside of the superconducting coil (in front of the calorimetry). These are constructed of several layers of extruded triangular scintillator strips that are read out using wavelength-shifting fibers and VLPCs. The next layer of detection involves three liquid-argon/uranium calorimeters: a central section (CC) covering $|\eta|$ up to ≈ 1 and two end calorimeters (EC) extending coverage to $|\eta| \approx 4$, all housed in separate cryostats [24]. In addition to the

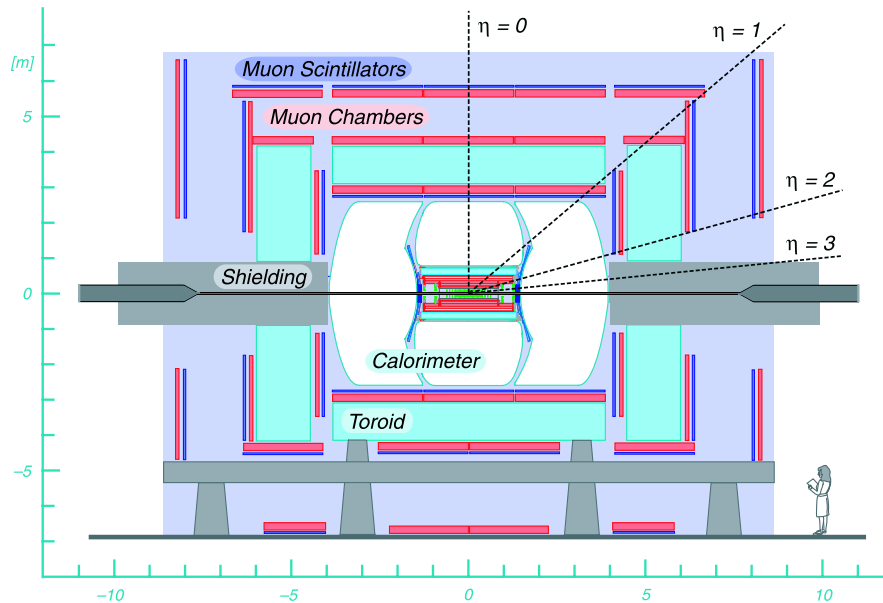


FIG. 5 (color online). General view of the D0 detector. The proton beam travels from left to right and the antiproton beam from right to left in this figure.

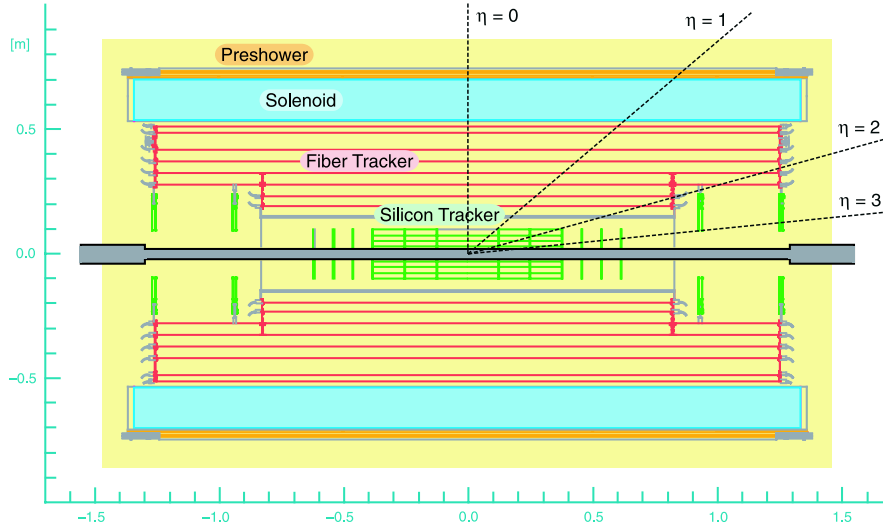


FIG. 6 (color online). Close view of the tracking systems.

preshower detectors, scintillators between the CC and EC cryostats provide sampling of developing showers for $1.1 < |\eta| < 1.4$.

A muon system resides beyond the calorimetry, and consists of a layer of tracking detectors and scintillation trigger counters before 1.8 T iron toroids, followed by two more similar layers after the toroids. Tracking for $|\eta| < 1$ relies on 10 cm wide drift tubes [24], while 1 cm mini drift tubes are used for $1 < |\eta| < 2$.

The luminosity is obtained from the rate of inelastic collisions measured using plastic scintillator arrays located in front of the EC cryostats, covering $2.7 < |\eta| < 4.4$.

B. Object reconstruction

Physics objects are reconstructed from the digital signals recorded in each part of the detector. Particles can be identified by certain patterns and, when correlated with other objects in the same event, they provide the basis for understanding the physics that produced such signatures in the detector.

1. Primary vertex

The position of the hard scatter interaction is determined at D0 by clustering tracks into seed vertices using a Kalman filter algorithm [25]. The primary vertex is then selected using a probability function based on the p_T values of the tracks assigned to each vertex. The hard scatter vertex is distinguished from other soft interaction vertices by the higher average p_T of its tracks. In multijet data events, the position resolution of the primary vertex in the transverse plane (perpendicular to the beam pipe) is around $40 \mu\text{m}$, convoluted with a typical beam spot size of around $30 \mu\text{m}$.

2. Electrons

Electron candidates are initially identified as energy clusters in the central region of the electromagnetic calorimeter, $|\eta| \leq 1.1$. We define two classes of electron candidates: *loose* and *tight*. *Loose* electrons are required to have the fraction of their total energy deposited in the electromagnetic (EM) calorimeter $f_{\text{EM}} > 0.9$ and a shower-shape chi-squared, based on seven variables that compare the values of the energy deposited in each layer of the electromagnetic calorimeter with average distributions from simulated electrons, to be $\chi_{\text{cal}}^2 < 75$. Finally, loose electron candidates are also required to be isolated by measuring the total deposited energy and the energy from the EM calorimeter only around the electron track: $E_{\text{Total}}(R < 0.4) < 1.15 \times E_{\text{EM}}(R < 0.2)$, where $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ is the radius of a cone defined by the azimuthal angle ϕ and the pseudorapidity η , centered on the electron candidate's track.

For an electron candidate to be included in the *tight* class, a track must be matched to the *loose* cluster within $|\Delta\eta| < 0.05$ and $|\Delta\phi| < 0.05$ and additionally pass a cut on a seven-variable likelihood \mathcal{L} built to separate real electrons from backgrounds [26]. The following variables are used in the likelihood: (i) f_{EM} ; (ii) χ_{cal}^2 ; (iii) $E_T^{\text{cal}}/p_T^{\text{track}}$, transverse energy of the cluster divided by the transverse momentum of the matched track; (iv) χ^2 probability of the track match; (v) distance of closest approach between the track and the primary vertex in the transverse plane; (vi) N_{tracks} , the number of tracks inside a cone of $R < 0.05$ around the matched track; and (vii) $\sum p_T$ of tracks in an $R < 0.4$ cone around the matched track. *Tight* electrons are obtained by applying a cut on the likelihood of $\mathcal{L} > 0.85$. The overall tight electron identification efficiency in data is around 75%.

A comparison between the dielectron invariant mass distributions for $Z \rightarrow e\bar{e}$ simulated events and data shows that the position of the simulated Z boson peak is shifted from that in data and that the electron energy resolution is better than in data. We apply small corrections to the identification efficiency and electromagnetic energy of simulated electrons and smear their energies to agree with data.

3. Muons

Muons are reconstructed in D0 up to $|\eta| = 2$ by first finding hits in all three layers of the muon spectrometers and requiring that the timing of these hits is consistent with the hard scatter, thus rejecting cosmic rays. Secondly, all muon candidates must be matched to a track in the central tracker. That central track must pass the following criteria: (i) chi-squared per degree of freedom less than 4; (ii) the distance of closest approach to the primary vertex in the transverse plane must be less than 3 standard deviations; and (iii) the distance in z between the track and the primary vertex must be less than 1 cm.

As for electrons, we similarly define two classes: *loose* and *tight*, but this time based solely on the muon's isolation from other objects. A *loose* isolated muon must comply with $R(\text{muon, jet}) > 0.5$, which is the distance between the muon and the jet axis. A *tight* isolated muon must be *loose* and additionally satisfy track-based and calorimeter-based criteria: $|\sum^{\text{tracks}} p_T/p_T(\mu)| < 0.06$ where the sum is over tracks within a cone of $R(\text{track, muon}) < 0.5$; and $|\sum^{\text{cells}} E_T/p_T(\mu)| < 0.08$ where the sum is over calorimeter cells within an annulus of $0.1 < R(\text{calorimeter cell, muon}) < 0.4$. The overall tight muon identification efficiency in data is around 65%.

Similarly to electrons in the simulation, we correct the energy scale for simulated muons and smear their energies to reproduce the data in $Z \rightarrow \mu\bar{\mu}$.

4. Jets

We reconstruct jets based on calorimeter cell energies, using the improved legacy cone algorithm [27] with radius $R = 0.5$. Noisy calorimeter cells are ignored in the reconstruction algorithm by imposing the requirement that neighboring cells have signals above the noise level.

Jet identification is based on a set of cuts to reject poor quality jets or noisy jets: (i) $0.05 < f_{\text{EM}} < 0.95$; (ii) fraction of jet E_T in the coarse hadronic calorimeter layers < 0.4 ; (iii) ratio of E_T 's of the most energetic cell to the second most energetic cell in the jet < 10 ; and (iv) smallest number of towers that make up 90% of the jet E_T , $n_{90} > 1$.

Jet energy scale corrections are applied to convert jet energies from the reconstructed level into particle-level energies. The reconstructed fully corrected energy of jets from the simulation of the detector performance does not exactly match that seen in data. Similar to electrons and

muons, we smear jet energies by a small amount in the simulation to reproduce the resolution measured in data.

5. Missing energy

We infer the transverse energy of the neutrino in the event as the opposite of the vector sum of all the energy deposited in the calorimeter. This calorimeter-only missing transverse energy is then corrected with the jet energy scale, the electromagnetic scale, and the energy loss from isolated muons in the calorimeter and their momenta.

C. Identification of b -quark jets

The presence of b quarks can be inferred from the long lifetimes and large invariant masses of B hadrons, which typically travel a few millimeters before hadronization. Thus b -quark jets contain a displaced vertex inside a jet whereas light-quark jets do not. The secondary vertex tagger (SVT), described below, makes use of this fact to identify, or tag, b -quark jets by fitting tracks in the jet into a secondary vertex.

1. Taggability

Before the b -quark tagging algorithm is applied to identify displaced vertices in the jet, a set of cuts is applied to ensure a good quality jet and factor out detector geometry effects. Thus the final probability to identify a b -quark jet is factored into two parts: a *taggability* part, or jet-quality-sensitive component, and a *tagger* part, or heavy-flavor-sensitive component. A taggable jet requires at least two tracks within a cone of $R = 0.5$ centered on the jet axis. At least one of these tracks must have $p_T > 1.0$ GeV, and additional tracks must have $p_T > 0.5$ GeV. All tracks must have at least one SMT hit, an xy distance-of-closest-approach (DCA) of < 0.2 cm, and a z DCA of < 0.4 cm with respect to the primary vertex. The *taggability* is the number of taggable jets divided by the number of good jets. Only jets satisfying jet identification requirements, with $p_T > 15$ GeV (after jet energy corrections) and $|\eta| \leq 2.5$ are considered to be good for the definition of taggability.

In simulated events, the taggability is higher than in data mainly due to a noncomprehensive description of the tracking detectors (dead detector elements, other inefficiencies, noise, etc.) resulting in a higher tracking efficiency (in particular within jets). Therefore, the Monte Carlo taggability must be calibrated to that observed in the data. A *taggability-rate function* is utilized to do this by parametrizing the taggability as a function of jet p_T and η . Thus, the taggability per jet is determined in data and applied to the Monte Carlo as

$$p^{\text{taggable}}(p_T, \eta) = \frac{\#\text{taggable jets in } (p_T, \eta) \text{ bin}}{\#\text{jets in } (p_T, \eta) \text{ bin}}. \quad (1)$$

Central jets with momenta above 40 GeV have taggabil-

ities of around 85%. For simulated jets the taggability is $\approx 90\%$.

2. Secondary vertex tagger

The SVT algorithm is designed to reconstruct a displaced vertex inside a jet by fitting tracks that have a large impact parameter from the hard scatter vertex. A simple algorithm is applied to the tracks to remove most K_S^0 's, Λ 's, and photon conversions. Tracks are then required to have at least two SMT hits, $p_T > 1.0$ GeV, transverse impact parameter significance ($d_{ca}/\sigma_{d_{ca}}$) greater than 3.5, and a track $\chi^2 > 10$. A simple cone jet algorithm is used to cluster the tracks into *track-jets*, and then a Kalman filter algorithm is used to find vertices with the tracks in each track-jet. The distance between the primary vertex and the found secondary vertex, the decay length L_{xy} , and its error $\sigma_{L_{xy}}$ are calculated taking into account the uncertainty on the primary vertex position. The decay length is a signed parameter, defined by the sign of the cosine of the angle between the vector from the primary vertex to the decay point and the total momentum of the tracks attached to the secondary vertex. If the decay length significance $L_{xy}/\sigma_{L_{xy}}$ is more than 7, then the found vertex is considered a tag. A calorimeter jet is considered tagged if the distance between the jet axis and the line joining the primary vertex and the secondary vertex is $R < 0.5$ in η, ϕ space. This set of cuts has been tuned to obtain a probability for a light quark mistag of 0.25%. Note that gluon jets are included in the light quark category.

We estimate the b tagging efficiency in a dijet data sample. The heavy flavor content of the sample is enhanced by requiring one of the jets to have a high- p_T muon relative to the jet axis. The SVT efficiency to tag the other jet can then be inferred. We estimate the c quark tagging efficiency from a Monte Carlo simulation. The mistagging rate, or how often a light-flavor jet (from u, d, s quarks or gluons) is identified as a b jet, is also measured in a dijet data sample. We count the number of found secondary vertices with $L_{xy}/\sigma_{L_{xy}} < -7$ and correct for the contribution of heavy-flavor jets in the sample and the presence of long-lived particles in light-flavor jets. The sign in the decay length L_{xy} is defined by the cosine of the angle between the vector from the primary vertex to the decay point and the total momentum of the tracks attached to the secondary vertex. Figure 7 shows the tagging efficiency as a function of jet p_T for the different types of jets.

To calculate the probability for a simulated jet to be tagged, a *tag-rate function* (TRF) derived from data is used similarly to the taggability parametrized in p_T and η :

$$p^{\text{tagging}}(p_T, \eta) = \frac{\text{\#SVT tagged jets in } (p_T, \eta) \text{ bin}}{\text{\#taggable jets in } (p_T, \eta) \text{ bin}}. \quad (2)$$

Separate functions are determined for b -quark jets, c -quark jets, and light-quark jets, as in Fig. 7.

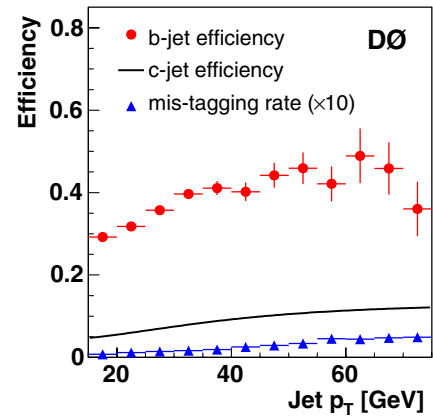


FIG. 7 (color online). Measured b -tagging efficiency (circles) and mistagging rate (triangles), and estimated c -tagging efficiency (solid line) as a function of jet p_T .

The TRFs are applied to the Monte Carlo samples in the following way. First, for each jet in the event (with $p_T > 15$ GeV and $|\eta| < 3.4$) a taggability-rate function is applied. If the jet contains a B meson within $R < 0.5$ of the jet axis it is labeled a b -quark jet. If a D meson is within $R < 0.5$ of the jet axis, it is labeled a c -quark jet. If no B or D meson is found in the jet, the jet is labeled a light-quark jet. The probability determined from the appropriate TRF is then applied. The taggability and tagging probability are multiplied together to determine the probability of the simulated jet to be tagged.

In data we apply the secondary vertex algorithm directly and can identify which jet is tagged and which is not. The situation in simulated events is different; the TRFs return a probability (or weight) rather than a tagged/not-tagged answer per jet. Since many of the discriminant variables used later on in the analysis (see Sec. VII A) need to know which jet was tagged, each possible combination of tagged and untagged jets is considered for every simulated event. Thus each event is used repeatedly in the analysis, considering each time a different jet as tagged. The probability of each combination is calculated using the tag-rate functions, and combined with the overall event weight. The sum of the weights for all the possible combinations of each event is equal to the original probability for an event to have at least one tagged jet.

The use of all permissible tagged jet combinations in each simulated event is a very powerful tool. It ensures that the kinematic distributions in histograms of tagged events have the correct shape, and it allows tagged jet information to be used in variables for signal/background separation, since the final classifiers are trained with weighted events.

III. TRIGGERS AND DATA SET

The D0 trigger system is composed of three levels. The first level consists of hardware and firmware components, the second level uses information from the first level to

TABLE II. Trigger conditions at levels 1, 2, and 3 for the electron plus jets trigger.

Level 1 Condition	Level 2 Condition	Level 3 Condition	Luminosity
1 EM tower, $E_T > 10$ GeV	1 e , $E_T > 10$ GeV, EM fraction > 0.85	1 tight e , $E_T > 15$ GeV	19.4 pb ⁻¹
2 jet towers, $E_T > 5$ GeV	2 jets, $E_T > 10$ GeV	2 jets, $E_T > 15$ GeV	
1 EM tower, $E_T > 10$ GeV	1 e , $E_T > 10$ GeV, EM fraction > 0.85	1 loose e , $E_T > 15$ GeV	91.2 pb ⁻¹
2 jet towers, $E_T > 5$ GeV	2 jets, $E_T > 10$ GeV	2 jets, $E_T > 15$ GeV	
1 EM tower, $E_T > 11$ GeV		1 tight e , $E_T > 15$ GeV	115.4 pb ⁻¹
		2 jets, $E_T > 20$ GeV	

TABLE III. Trigger conditions at levels 1, 2, and 3 for the muon plus jets trigger.

Level 1 Condition	Level 2 Condition	Level 3 Condition	Luminosity
1 μ , $ \eta < 2.0$	1 μ , $ \eta < 2.0$	1 jet, $E_T > 20$ GeV	113.7 pb ⁻¹
1 jet tower, $E_T > 5$ GeV			
1 μ , $ \eta < 2.0$	1 μ , $ \eta < 2.0$	1 jet, $E_T > 25$ GeV	113.7 pb ⁻¹
1 jet tower, $E_T > 3$ GeV	1 jet, $E_T > 10$ GeV		

construct simple physics objects, and the third level is software based and performs full event reconstruction.

The D0 calorimeter is used to trigger events based on the energy deposited in towers of size $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ that are segmented longitudinally into electromagnetic and hadronic sections. The level 1 electron trigger requires electrons to be above a certain threshold: $E_T \equiv E \sin\theta > T$ where E is the energy deposited in the tower, θ is the angle between the beam and the trigger tower from the center of the detector, and T is the programmed threshold. The level 2 electron trigger uses a seed-based clustering algorithm that sums the energy deposited in two neighboring towers and has the ability to make a decision based on the threshold of the cluster, the electromagnetic fraction, and isolation of the electron. The level 3 electron trigger uses a simple cone algorithm with $R < 0.25$ and requirements on the E_T , the electromagnetic fraction, and the quality of the transverse shower shape.

The level 1 jet trigger is similar to the electron trigger tower algorithm but includes the energy deposited in the hadronic portion of the calorimeter. The level 2 jet trigger uses a seed-based clustering algorithm summing the energy deposition in a 5×5 tower array. The level 3 jet algorithm is similar to the level 3 electron algorithm but does not include a requirement on the electromagnetic fraction or shower shape.

The level 1 muon trigger examines hits from the muon wire chambers, muon scintillation counters, and tracks from the level 1 track trigger for patterns consistent with those coming from a muon. The level 2 muon trigger reconstructs muon tracks from both wire and scintillator elements in the muon system. It can impose requirements on the number of muons, the p_T and η of the muons, and the overall quality of the muons. The level 3 muon trigger

uses wire and scintillator hits to reconstruct tracks using segments inside and outside the toroid.

The output of the first level of the trigger is used to limit the rate for accepted events to ≈ 1.5 kHz. At the next trigger stage, with more refined information, the rate is reduced further to ≈ 800 Hz. The third level of the trigger, with access to all the event information, reduces the output rate to ≈ 50 Hz, which is written to tape.

The data were acquired in the period between August 2002 and March 2004. Tables II and III show the triggers used to collect the data for the electron plus jets ($e + \text{jets}$) and muon plus jets ($\mu + \text{jets}$) triggers and give the integrated luminosity for each trigger.

IV. EVENT SELECTION

Event selection begins after all corrections have been applied to the data. These corrections include the jet energy and the EM energy calibrations. The primary vertex, z_{vertex} , for the event must be within the tracking fiducial region, $|z_{\text{vertex}}| < 60$ cm, and must have three tracks associated with it $N_{\text{tracks}} \geq 3$.

TABLE IV. Initial event selection requirements.

Selection cut	$e + \text{jets}$	$\mu + \text{jets}$
tight e , $E_T \geq 15$ GeV	= 1	= 0
tight μ , $E_T \geq 15$ GeV	= 0	= 1
\cancel{E}_T	≥ 15 GeV	
N_{jets}	$2 \leq N_{\text{jets}} \leq 4$	
$E_T(\text{jet})$	≥ 15 GeV	
$ \eta(\text{jet}) $	≤ 3.2	
$E_T(\text{jet } 1)$	≥ 25 GeV	
$ \eta(\text{jet } 1) $	≤ 2.4	

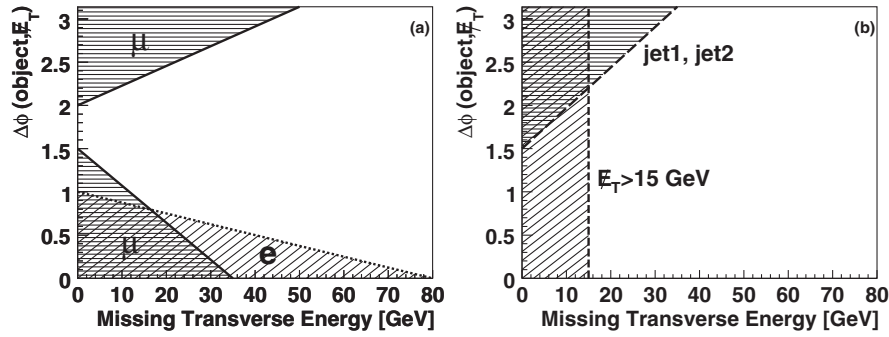


FIG. 8. Kinematic regions excluded in the $e + \text{jets}$ and $\mu + \text{jets}$ analyses by the triangle cuts applied in the $(\Delta\phi(\text{object}, \cancel{E}_T), \cancel{E}_T)$ plane, where each object can be—the tight isolated electron or muon (a), and the leading and second leading jets (b). The shaded areas are excluded.

As discussed in Sec. IC, the single top quark signature is characterized by one isolated high- p_T charged lepton, \cancel{E}_T , and two to four jets. We accept events with three or four jets in order to include contributions from extra gluons and quarks. The b jet from the single top quark decay tends to be more energetic than the other jets associated with the event, so we require a higher E_T for the leading jet. Table IV lists the requirements of the initial selection.

In addition, we make a set of cuts that remove misreconstructed events, also known as “triangle cuts.” Mismeasured transverse energy creates false missing energy in a parallel or antiparallel direction, which allows multijet events to satisfy our selection criteria. The triangle cuts remove these mismeasured events, which are difficult to model, but do not affect the signal appreciably because there is very small signal acceptance in these kinematic regions. In Fig. 8, we show the kinematic regions that are removed by the triangle cuts.

From the selected jets in the event, at least one b -tagged jet must be found. For the t -channel analysis, at least one jet must be untagged. This requirement comes from the fact that one of the main features of the t -channel signal is that a light-quark jet exists in the final state. The events are then divided into subsets consisting of the number of tagged jets found in the event: single-tagged or double-tagged. Since the t -channel requires at least one untagged jet, there are no two-jet events in the double-tagged sample in the double-tagged t -channel search.

V. SIGNAL AND BACKGROUND MODELING

In order to compare the observed event yield in data with our expectation, and to set limits on the single top quark production cross sections, we determine acceptances and event yields for the single top quark signals and the various SM background contributions. This estimation is based primarily on simulated samples for shapes of distributions, except for the multijet background where we use data samples. The yield normalization is based on theoretical cross sections, except for the $W + \text{jets}$ and multijet backgrounds which are normalized to data.

A. Acceptance and yield for simulated samples

The acceptance α for a particular simulated signal or background sample is calculated as

$$\alpha = \frac{1}{N^{\text{MC}}} \sum_i w_i \quad (3)$$

where the sum is over simulated events that pass the selection cuts and is normalized to the total number of simulated events in the sample N^{MC} . All selection cuts are applied both to the data and simulated samples, except for the trigger requirement and b tagging, which are only applied to the data samples. The event weight w_i is given by

$$w_i = \epsilon_i^{\text{lepton ID}} \times \epsilon_i^{\text{jet ID}} \times \epsilon_i^{\text{trigger}} \times \epsilon_i^{b \text{ tagging}} \quad (4)$$

and includes correction factors ϵ to account for effects not modeled and for cuts not applied to the simulated samples. Trigger requirements are not made in the simulation (see Sec. III) and the correction factors $\epsilon_i^{\text{trigger}}$ are about 90% and depend on the E_T and η of each lepton and jet. Furthermore, we do not require b tagging in simulated events, and the correction factor $\epsilon_i^{b \text{ tagging}}$, which depends on the E_T , η , and flavor of each jet (see Sec. IIC 2), averages about 55% for s -channel events and about 40% for t -channel events.

The yield estimate \mathcal{Y} is given by the product of acceptance, integrated luminosity \mathcal{L} , theory cross section σ^{theory} , and branching fraction B :

$$\mathcal{Y} = \alpha \times \mathcal{L} \times \sigma^{\text{theory}} \times B. \quad (5)$$

The branching fraction factor gives the fraction of events that result in the final-state lepton of interest (e or μ). The yield includes a small contribution from $W \rightarrow \tau$ decays where the τ decays to e or μ .

B. Single top quark signals

The CompHEP matrix element generator [28] has been used to model single top quark s -channel and t -channel signal events. We include not only the leading order Feynman diagrams in the event generation, but also the $O(\alpha_s)$ diagrams with real gluon radiation in order to reproduce NLO distributions. For the t -channel sample, we include both the leading order diagram [Fig. 2(a)] and the W -gluon fusion diagram [Fig. 2(b)] explicitly, generating W -gluon fusion events for the region of phase space where the \bar{b} quark from gluon splitting has $p_T(\bar{b}) > 17$ GeV and leading order events otherwise.

C. $t\bar{t}$ background

Top quark pair production contributes as a background both in the lepton + jets and in the dilepton decay channels. This background is modeled using ALPGEN [17], and the yields are normalized to the theory cross section (see Sec. IC).

D. WW and WZ backgrounds

The backgrounds from diboson production are modeled using ALPGEN, and the yields are normalized to the theory cross sections [29].

E. Multijet and W + jets backgrounds

The backgrounds from multijet (fake lepton) and W + jets production are normalized to the data sample before b tagging [30]. We start from a data sample passing all selection cuts including the *loose* lepton requirements (see Sec. IIB). From that sample, we select a subset of events that also pass the *tight* lepton requirements. In addition, we determine the probabilities for real and fake leptons to pass the tight lepton requirement. The probability for real leptons is determined using a data sample of leptonic Z boson decays, whereas the probability for fake leptons is obtained from a QCD multijet sample. These two probabilities together with the numbers of events in the two samples then allow us to calculate the number of real and fake lepton events in the W + jets and multijet background samples [31].

The shapes of the distributions for the multijet background are modeled using a data sample that passes all selection cuts but fails the tight lepton identification requirements. The shapes of the distributions for the W + jets background are modeled using ALPGEN W + 2 jets events.

1. Multijet background

A part of the background comes from events in which jets are misidentified as isolated leptons. In the electron channel, this background is typically produced by jets that contain a π^0 , which, together with a randomly associated track, is misreconstructed as an isolated electron since it

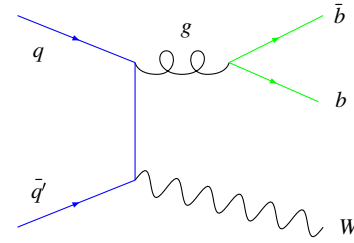


FIG. 9 (color online). Representative Feynman diagram for $Wb\bar{b}$ production.

decays to two photons. In the muon channel, this background is typically produced by heavy-flavor jets in which a muon from a semileptonic decay is misreconstructed as an isolated high- p_T muon.

The multijet background is estimated purely from data. We use multijet data samples that pass all event selection requirements, but fail the requirement on tight muon isolation or tight electron quality (see Sec. IIB) to determine the kinematic shape of distributions. Hence these events are required to pass, for example, the isolation cut $R(\text{lepton}, \text{jet}) > 0.5$ which means that they model the shape of the multijet background well. These samples are normalized to the multijet background estimate in the data sample after event selection, but before requiring a b tag.

2. W + Jets background

An example Feynman diagram for W + 2 jet production is shown in Fig. 9. This background is modeled from a simulated Wjj sample ($j = u, d, s, c, g$), which includes not just light-quark flavors but also c quarks (considered massless in this model). We use a separate sample for $Wb\bar{b}$ and explicitly exclude events with b quarks from the Wjj sample. The parton level samples were generated with ALPGEN.

Since the W + jets background is normalized to data (after subtraction of the small $t\bar{t}$ and diboson content), it includes all sources of boson + jets events with a similar flavor composition, in particular Z + jets events where one of the leptons from the Z boson decay is not identified.

F. Detector simulation

The parton level samples for the single top quark signals, $t\bar{t}$, W + jets, WW , and WZ backgrounds are processed with PYTHIA [32] for hadronization and modeling of the underlying event, using the CTEQ5L [33] parton distribution functions. TAUOLA [34] is used for tau lepton decays and EVTGEN [35] for B hadron decays. The generated events are processed through a GEANT-based [36] simulation of the D0 detector.

VI. EVENT YIELDS

The expected event yields for the various background contributions are calculated from both simulated samples

TABLE V. Event yields after selection in the electron and muon channels.

	Electron channel				Muon channel			
	before tag	= 1 tag	≥ 2 tags <i>s</i> -channel	≥ 2 tags <i>t</i> -channel	before tag	= 1 tag	≥ 2 tags <i>s</i> -channel	≥ 2 tags <i>t</i> -channel
<i>Signals</i>								
<i>tb</i>	5	2.3	0.53	...	5	2.2	0.50	...
<i>tqb</i>	12	4.1	...	0.25	11	3.9	...	0.23
<i>Backgrounds</i>								
<i>tb</i>	0.14	0.13
<i>tqb</i>	0.32	0.29	...
$t\bar{t} \rightarrow \ell + \text{jets}$	59	25.0	6.12	5.74	58	24.2	5.78	5.48
$t\bar{t} \rightarrow \ell\ell$	16	6.8	1.58	0.74	17	7.2	1.62	0.75
$Wb\bar{b}$	40	15.1	2.49	0.59	33	12.7	2.15	0.55
Wjj	3211	68.2	1.53	0.66	2898	62.9	1.29	0.68
WW	13	0.7	0.00	0.00	14	0.7	0.00	0.00
WZ	4	0.6	0.11	0.02	5	0.5	0.09	0.01
Multijet	478	13.7	0.31	0.14	256	17.2	0.20	0.20
<i>Summed signals</i>	17	6.4	0.53	0.25	16	6.0	0.50	0.23
<i>Summed backgrounds</i>	3821	130.1	3280	125.4
<i>Summed backgrounds + tqb</i>	3833	134.2	12.47	...	3291	129.3	11.43	...
<i>Summed backgrounds + tb</i>	3826	132.4	...	8.03	3285	127.6	...	7.80
<i>Data</i>	3821	134	15	11	3280	118	16	8

and data. The expected event yield for the single top quark signal is calculated from simulated samples and normalized to the theoretical cross sections.

The total background event yield \mathcal{Y} is given by the sum over all backgrounds:

$$\mathcal{Y} = \sum_i \mathcal{Y}_i, \quad (6)$$

where each individual yield \mathcal{Y}_i is given by Eq. (5) for the various MC samples.

TABLE VI. Estimates for signal and background yields and the numbers of observed events in data after event selection for the electron and muon, single-tagged and double-tagged analysis sets combined. The $W + \text{jets}$ yields include the diboson backgrounds. The total background for the s -channel (t -channel) search includes the tqb (tb) SM yields. The quoted yield uncertainties include systematic uncertainties taking into account correlations between the different analysis channels and samples.

Source	<i>s</i> -channel search		<i>t</i> -channel search	
<i>tb</i>	$5.5 \pm$	1.2	$4.7 \pm$	1.0
<i>tqb</i>	$8.6 \pm$	1.9	$8.5 \pm$	1.9
$W + \text{jets}$	$169.1 \pm$	19.2	$163.9 \pm$	17.8
$t\bar{t}$	$78.3 \pm$	17.6	$75.9 \pm$	17.0
Multijet	$31.4 \pm$	3.3	$31.3 \pm$	3.2
Total background	$287.4 \pm$	31.4	$275.8 \pm$	31.5
Observed events	283		271	

Table V shows the numbers of events for each of the signals, combinations of signals, backgrounds, and data, after event selection and b tagging. The background sum reproduces the data within uncertainties for all samples after b tagging.

A summary of the yield estimates for the signal and backgrounds and the numbers of observed events in data after selection, including the systematic uncertainties as described in Sec. VIII, is shown in Table VI.

After b tagging, the $W + \text{jets}$ background makes up around 60% of the total background model (48% Wjj , 12% $Wb\bar{b}$), the $t\bar{t}$ background is around 27% (21% lepton + jets, 6% dilepton), 10% is mainly multijet background, and s -channel single top quark production provides 3% in the t -channel search and vice versa. While the t -channel cross section is more than a factor two bigger than the s -channel cross section, s -channel events are more likely to pass the b tagging requirement due to the presence of two high- p_T b quarks. As a result, the two single top quark processes have similar expected yields.

VII. EVENT ANALYSIS

Table V shows that even after event selection and b tagging, the expected single top quark signal yield is small compared to the overwhelming backgrounds. Additional steps are necessary in order to separate the signal and background. In this section, we first present kinematic variables that allow us to separate the s -channel or t -channel single top quark signal from the backgrounds. We then describe a cut-based analysis and a neural networks analysis that use these variables.

A. Discriminating variables

In this section we introduce the variables that we found to be most effective in separating the single top quark signals from the backgrounds. The list of discriminating variables has been chosen based on an analysis of Feynman diagrams of signals and backgrounds [37] and on a study of single top quark production at NLO [11,12].

The variables fall into three categories: individual object kinematics, global event kinematics, and variables based on angular correlations. The list of variables is shown in Table VII.

Some of the discriminating variables require the reconstruction of the top quark in each event. First, the W boson

from the top quark decay is reconstructed from the isolated lepton and the missing transverse energy. The z -component of the neutrino momentum is calculated using a W boson mass constraint, choosing the solution with smaller $|p_z^\nu|$ from the two possible solutions. The candidate top quark is reconstructed from this W boson and a jet. In order to get optimum separation between signal and background, this jet is chosen according to whether a variable is primarily used in the s -channel or the t -channel search. In the t -channel analysis, there is typically only one high- p_T b quark jet in the final state, thus the leading b -tagged jet is chosen to reconstruct the top quark. By contrast, in the s -channel there are two high- p_T b quark jets in the final

TABLE VII. List of discriminating variables. A check mark in the final four columns indicates in which signal-background pair of the neural net analysis the variable is used.

Variable	Description	Signal-background pairs			
		tb $Wb\bar{b}$	$t\bar{t}$	tqb $Wb\bar{b}$	$t\bar{t}$
<i>Individual object kinematics</i>					
$p_T(\text{jet } 1_{\text{tagged}})$	Transverse momentum of the leading tagged jet	✓	✓	✓	...
$p_T(\text{jet } 1_{\text{untagged}})$	Transverse momentum of the leading untagged jet	✓	✓
$p_T(\text{jet } 2_{\text{untagged}})$	Transverse momentum of the second untagged jet	✓
$p_T(\text{jet } 1_{\text{nonbest}})$	Transverse momentum of the leading nonbest jet	✓	✓
$p_T(\text{jet } 2_{\text{nonbest}})$	Transverse momentum of the second nonbest jet	✓	✓
<i>Global event kinematics</i>					
$\sqrt{\hat{s}}$	Invariant mass of all final-state objects	✓	...	✓	✓
$p_T(\text{jet } 1, \text{jet } 2)$	Transverse momentum of the two leading jets	✓	...	✓	...
$M_T(\text{jet } 1, \text{jet } 2)$	Transverse mass of the two leading jets	✓
$M(\text{all jets})$	Invariant mass of all jets	✓	✓	✓	✓
$H_T(\text{all jets})$	Sum of the transverse energies of all jets	✓	...
$p_T(\text{all jets} - \text{jet } 1_{\text{tagged}})$	Transverse momentum of all jets excluding the leading tagged jet	...	✓	...	✓
$M(\text{all jets} - \text{jet } 1_{\text{tagged}})$	Invariant mass of all jets excluding the leading tagged jet	✓
$H(\text{all jets} - \text{jet } 1_{\text{tagged}})$	Sum of the energies of all jets excluding the leading tagged jet	...	✓	...	✓
$H_T(\text{all jets} - \text{jet } 1_{\text{tagged}})$	Sum of the transverse energies of all jets excluding the leading tagged jet	✓
$M(W, \text{jet } 1_{\text{tagged}})$	Invariant mass of the reconstructed top quark using the leading tagged jet	✓	✓	✓	✓
$M(\text{all jets} - \text{jet}_{\text{best}})$	Invariant mass of all jets excluding the best jet	...	✓
$H(\text{all jets} - \text{jet}_{\text{best}})$	Sum of the energies of all jets excluding the best jet	...	✓
$H_T(\text{all jets} - \text{jet}_{\text{best}})$	Sum of the transverse energies of all jets excluding the best jet	...	✓
$M(W, \text{jet}_{\text{best}})$	Invariant mass of the reconstructed top quark using the best jet	✓
<i>Angular variables</i>					
$\eta(\text{jet } 1_{\text{untagged}}) \times Q_\ell$	Pseudorapidity of the leading untagged jet \times lepton charge	✓	✓
$\Delta\mathcal{R}(\text{jet } 1, \text{jet } 2)$	Angular separation between the leading two jets	✓	...	✓	...
$\cos(\ell, \text{jet } 1_{\text{untagged}})_{\text{top}_{\text{tagged}}}$	Top quark spin correlation in the optimal basis for the t -channel [15], reconstructing the top quark with the leading tagged jet	✓	...
$\cos(\ell, Q_\ell \times z)_{\text{top}_{\text{best}}}$	Top quark spin correlation in the optimal basis for the s -channel [15], reconstructing the top quark with the best jet	✓
$\cos(\text{all jets}, \text{jet } 1_{\text{tagged}})_{\text{all jets}}$	Cosine of the angle between the leading tagged jet and the all jets system in the all jets rest frame	✓	✓
$\cos(\text{all jets}, \text{jet}_{\text{nonbest}})_{\text{all jets}}$	Cosine of the angle between the leading nonbest jet and the all jets system in the all jets rest frame	...	✓

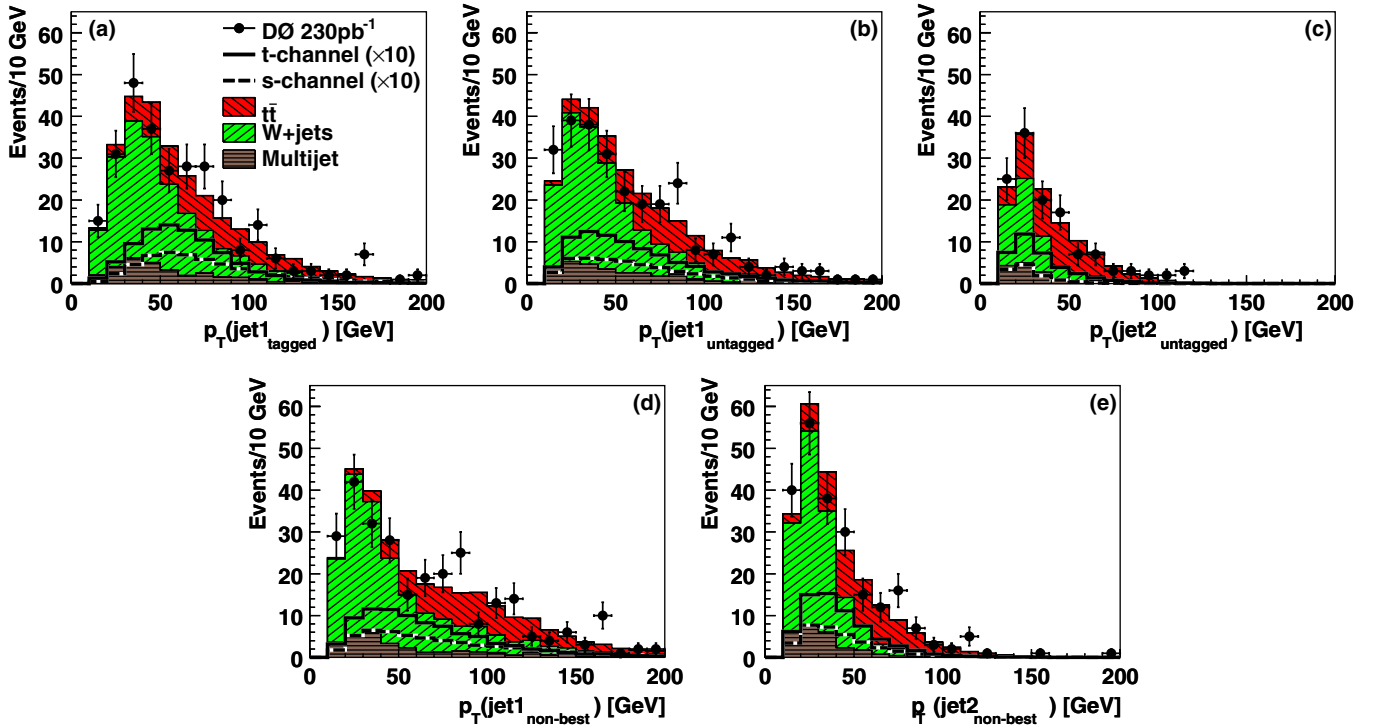


FIG. 10 (color online). Comparison of signal, backgrounds, and data after selection and requiring at least one b -tagged jet for five individual object variables. Electron and muon channels are combined. The transverse momentum is shown for (a) the leading tagged jet; (b) the leading untagged jet; (c) the second untagged jet, for those events that contain at least two untagged jets; (d) the leading nonbest jet; and (e) the second nonbest jet. Signals are multiplied by ten.

state, and a choice needs to be made between them. Furthermore, typically only one of the two is identified as a b -tagged jet. We use the best-jet algorithm [19] to identify this jet without using b tagging information. The best jet is defined as the jet in each event which gives, together with the reconstructed W boson, an invariant mass closest to 175 GeV. Jets that have not been identified by the b tagging algorithm are called “untagged” jets.

Figures 10–14 show all discriminating variables used in this analysis, comparing the single top quark signal distributions to those of the background sum and the data. Good agreement between the data and the background model is seen in all cases. The difference between the background model and the data in terms of χ^2/DF varies between about 0.5 and 2.0. The largest disagreements are in the high-mass regions of Figs. 12(b) and 13(a) where the systematic uncertainty on the background sum is largest. This uncertainty can reach a factor of 3 or more in these regions, mostly due to the uncertainties on jet energy calibration and b -tag modeling.

B. Cut-based analysis

This analysis takes the discriminating variables, chooses the best subsets, and finds the optimal points to cut on them in order to improve the expected cross section limits by increasing the signal to background ratio.

Optimization of the cut positions is performed by scanning sets of cut points for each discriminating variable. The signal Monte Carlo events are used to define the sets of cut values scanned in the algorithm. Only cut value combinations are used that appear in signal Monte Carlo events. The signal and background pass rates are determined for each cut point. Then an expected limit on the cross section is obtained from these and the best result is used as the operating point of the analysis.

The strategy is to look at the s - and t -channel processes separately to take full advantage of the kinematical differences between the channels. For each channel, there are four disjoint analyses: two leptons (e, μ) \times number of tagged b jets ($= 1, \geq 2$).

The most critical part of this analysis is to find the combination of variables and cuts that leads to the lowest expected cross section limit. We first look at single-variable cuts to determine which variables are most effective in each channel. Once an ordered list of variables is found (ordered by their power to lower the expected limit), sets of variables are formed starting with the best variable and consecutively including one by one the rest of the variables. For each set, the optimal cut position of each variable is recalculated. Finally, the variable set that gives the lowest expected limit is chosen. Table VIII shows the optimal variable sets and cuts found for each channel.

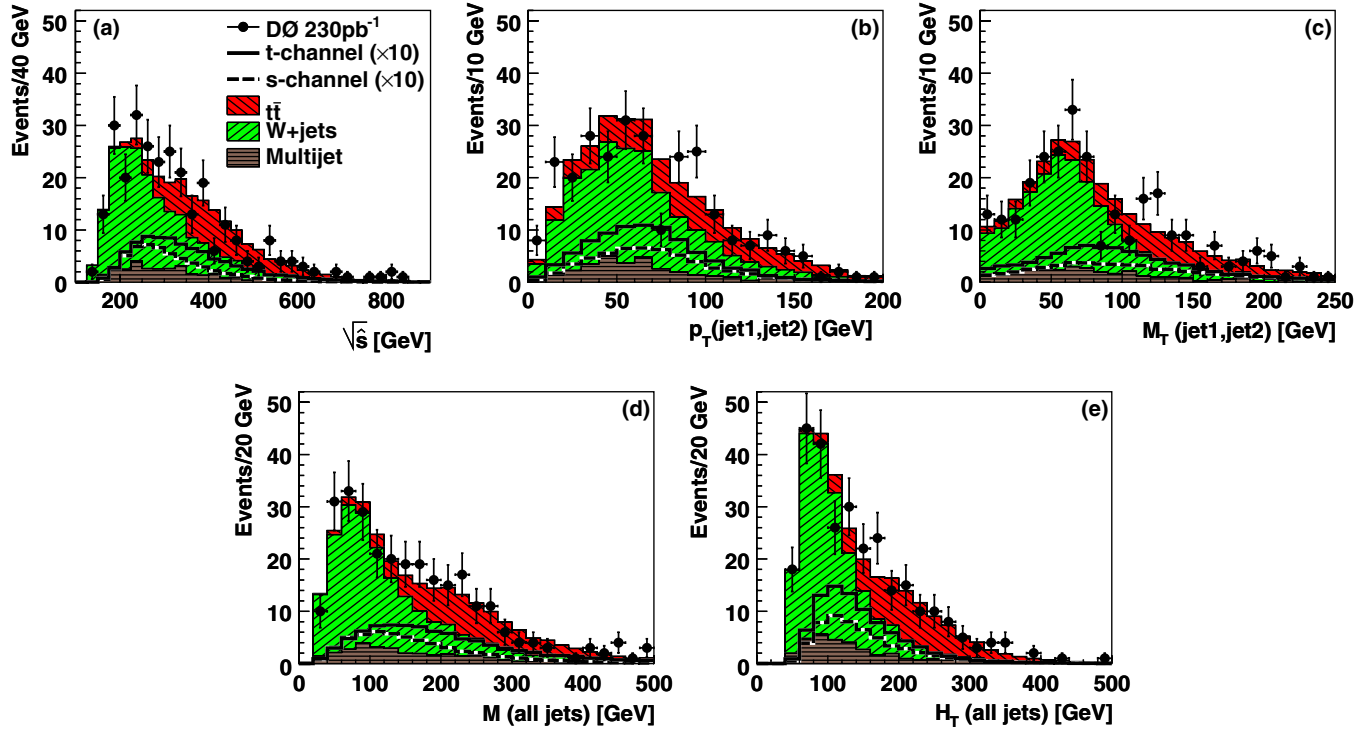


FIG. 11 (color online). Comparison of signal, backgrounds, and data after selection and requiring at least one *b*-tagged jet for five discriminating event kinematic variables. Electron and muon channels are combined. Shown are (a) the invariant mass of all final-state objects, (b) the total transverse momentum (scalar sum of magnitudes) of the leading two jets, (c) the transverse mass of the leading two jets, (d) the invariant mass of all jets, and (e) the total transverse energy of all jets. Signals are multiplied by ten.

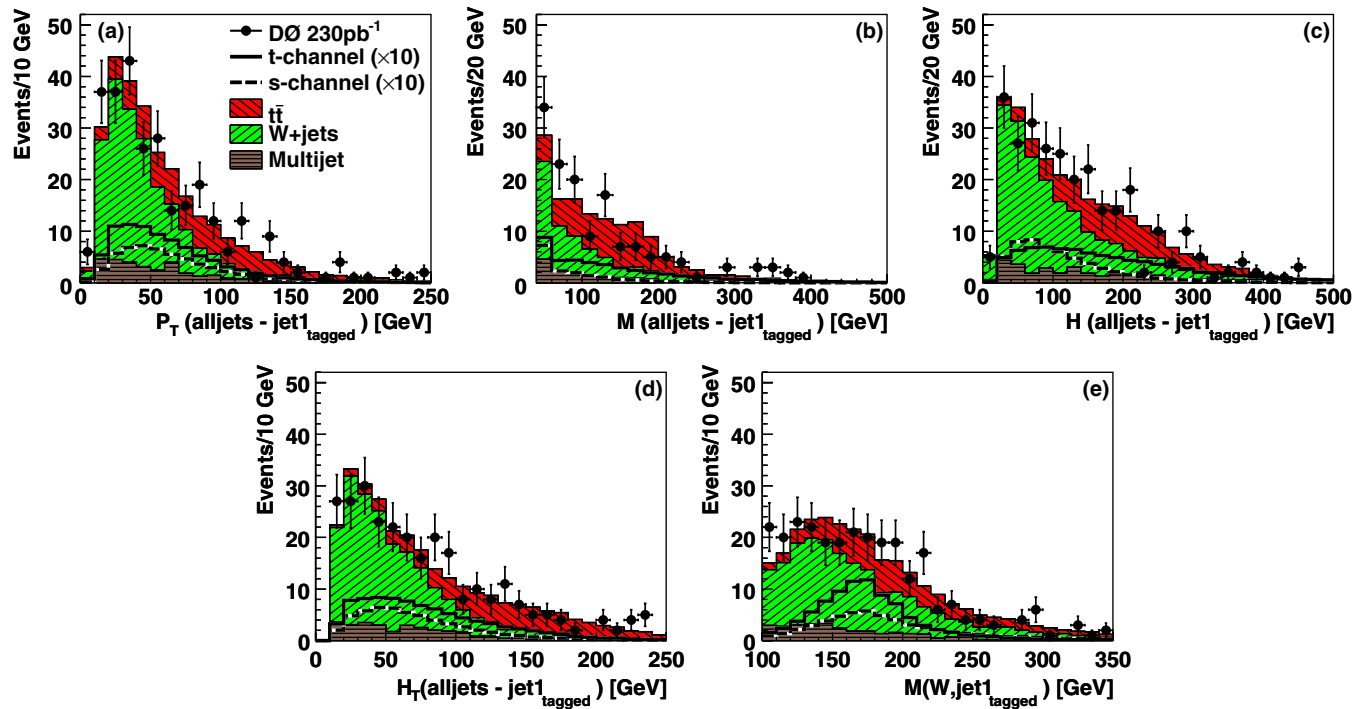


FIG. 12 (color online). Comparison of signal, backgrounds, and data after selection and requiring at least one *b*-tagged jet for five discriminating event kinematic variables. Electron and muon channels are combined. Shown are (a) the transverse momentum (scalar sum of magnitudes) of all jets except the leading tagged jet, (b) the invariant mass of all jets except the leading tagged jet, (c) the total energy of all jets except the leading tagged jet, (d) the total transverse energy of all jets except the leading tagged jet, and (e) the invariant mass of the top quark reconstructed from the reconstructed *W* boson and the leading tagged jet. Signals are multiplied by ten.

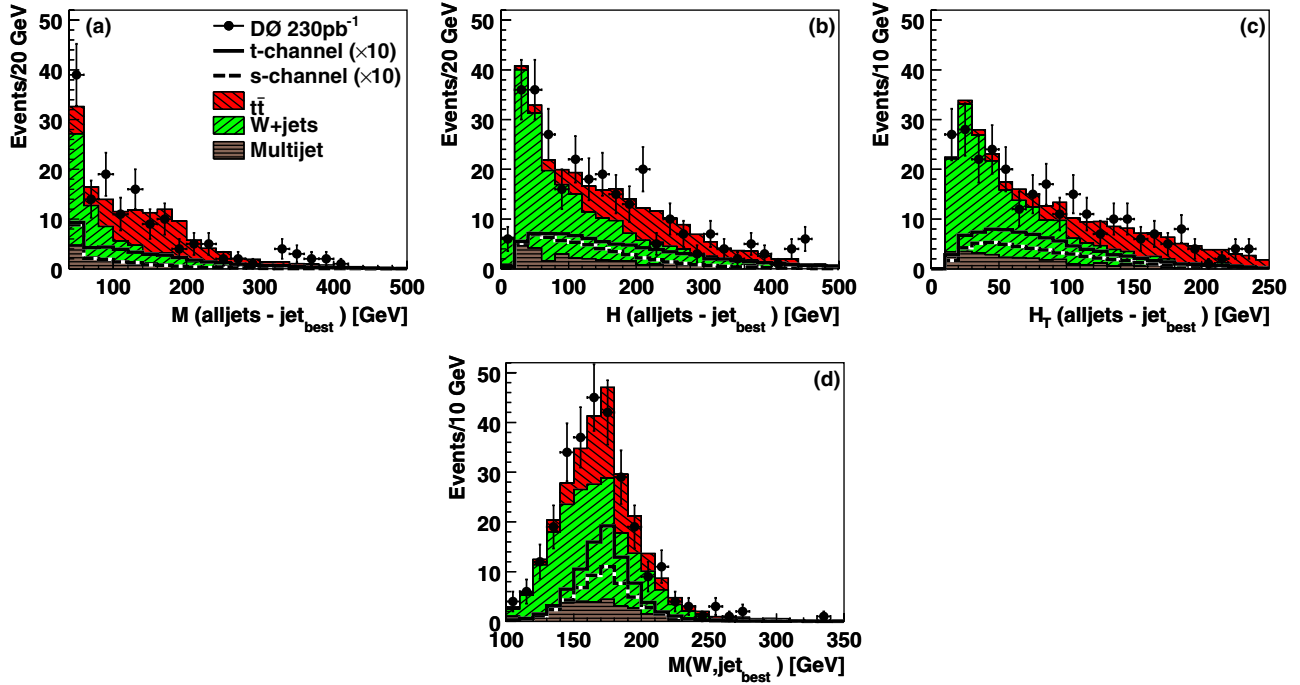


FIG. 13 (color online). Comparison of signal, backgrounds, and data after selection and requiring at least one b -tagged jet for four discriminating event kinematic variables. Electron and muon channels are combined. Shown are (a) the invariant mass of all jets except the best jet, (b) the total energy of all jets except the best jet, (c) the total transverse energy of all jets except the best jet, and (d) the invariant mass of the top quark reconstructed from the reconstructed W boson and the best jet. Signals are multiplied by ten.

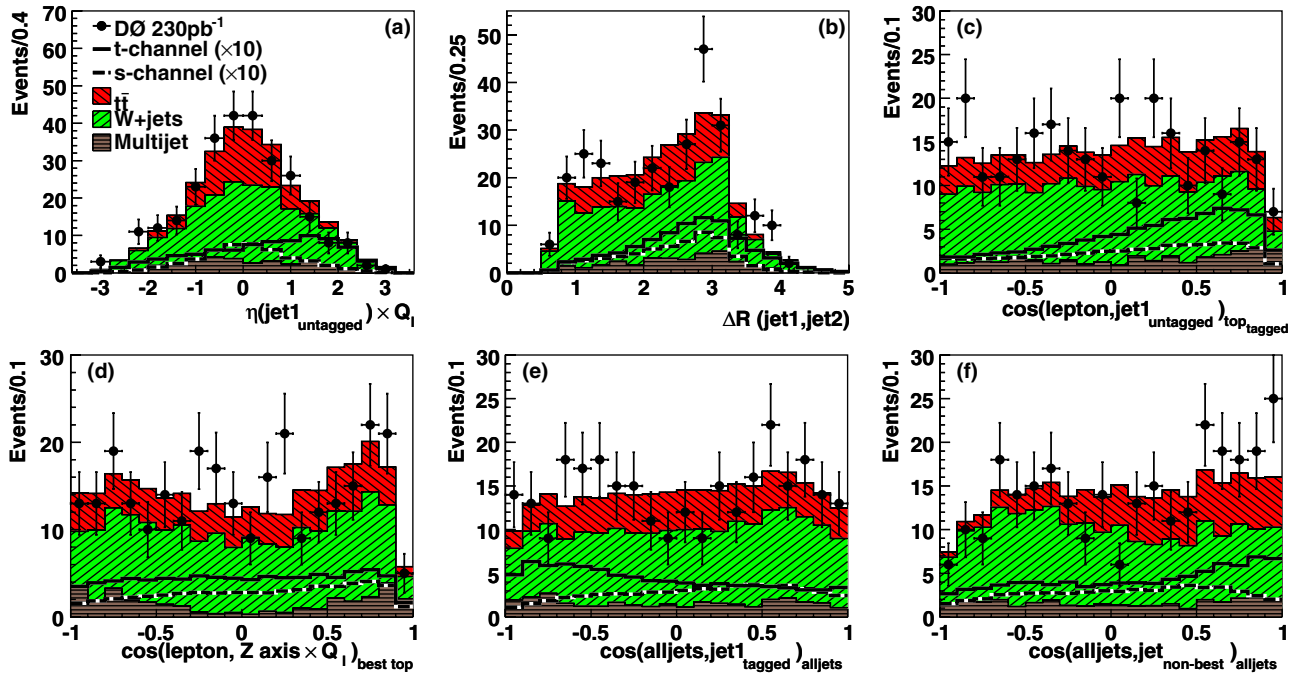


FIG. 14 (color online). Comparison of signal, backgrounds, and data after selection and requiring at least one b -tagged jet for six angular correlation variables. Electron and muon channels are combined. Shown are (a) the pseudorapidity of the leading untagged jet multiplied by the lepton charge, (b) the angular separation between the leading two jets, (c) the top quark spin correlation in the optimal basis for the t -channel, (d) the top quark spin correlation in the optimal basis for the s -channel, (e) the cosine of the angle between the leading tagged jet and all jets, in the all jets frame, and (f) the cosine of the angle between the leading nonbest jet and all jets, in the all jets frame. Signals are multiplied by ten.

TABLE VIII. The optimal set of variables and cuts for each analysis channel.

Channel	<i>s</i> -channel Variables	Cuts	<i>t</i> -channel Variables	Cuts
<i>Electron</i> = 1 tag	$p_T(\text{jet } 1_{\text{tagged}})$	>27 GeV	$H_T(\text{all jets})$	>71 GeV
	$M(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<70 GeV	$M(\text{all jets})$	>57 GeV
	$\sqrt{\hat{s}}$	>196 GeV	$\sqrt{\hat{s}}$	>203 GeV
			$ 175 - M(W, \text{jet } 1_{\text{tagged}}) $	<57 GeV
			$p_T(\text{jet } 1_{\text{tagged}})$	>21 GeV
			$p_T(\text{jet } 1_{\text{tagged}})$	>34 GeV
≥ 2 tags	$p_T(\text{jet } 1_{\text{tagged}})$	>42 GeV	$M(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<75 GeV
	$M(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<98 GeV	$H(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<504 GeV
	$H(\text{all jets} - \text{jet}_{\text{best}})$	<304 GeV	$H(\text{all jets} - \text{jet}_{\text{best}})$	<504 GeV
	$H(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<304 GeV		
<i>Muon</i> = 1 tag	$p_T(\text{jet } 1_{\text{tagged}})$	>33 GeV	$ 175 - M(W, \text{jet } 1_{\text{tagged}}) $	<60 GeV
	$M(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<74 GeV	$\sqrt{\hat{s}}$	>210 GeV
	$H(\text{all jets} - \text{jet}_{\text{best}})$	<504 GeV	$M(\text{all jets})$	>70 GeV
	$H(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<504 GeV	$H_T(\text{all jets})$	>58 GeV
			$ 175 - M(W, \text{jet } 1_{\text{tagged}}) $	<213 GeV
≥ 2 tags	$p_T(\text{jet } 1_{\text{tagged}})$	>33 GeV		
	$M(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<74 GeV		
	$H(\text{all jets} - \text{jet}_{\text{best}})$	<504 GeV		
	$H(\text{all jets} - \text{jet } 1_{\text{tagged}})$	<504 GeV		

Table IX shows the numbers of events and expected background and signal yields after these cuts have been applied.

A summary of the yield estimates for the signal and backgrounds and the numbers of observed events in data after the cut-based selection, including the systematic uncertainties as described in Sec. VIII, is shown in Table X.

This *s*- and *t*-channel combined signal to background ratio improves from around 1/20 after the basic selection (Table VI) to around 1/14 after these cuts have been applied. It is clear that more sophisticated separation techniques are needed to isolate the signal better from the large backgrounds.

TABLE IX. Event yields after the cut-based analysis selection.

	Electron channel				Muon channel			
	= 1 Tag		≥ 2 Tags		= 1 Tag		≥ 2 Tags	
	<i>s</i> -channel	<i>t</i> -channel	<i>s</i> -channel	<i>t</i> -channel	<i>s</i> -channel	<i>t</i> -channel	<i>s</i> -channel	<i>t</i> -channel
<i>Signals</i>								
<i>tb</i>	1.7	...	0.45	0.12	1.9	...	0.43	...
<i>tqb</i>	...	3.4	...	0.23	...	3.1	...	0.23
<i>Backgrounds</i>								
<i>tb</i>	...	1.6	...	0.12	...	1.4	...	0.12
<i>tqb</i>	2.5	...	0.14	...	2.8	...	0.01	...
$t\bar{t} \rightarrow \ell + \text{jets}$	3.8	18.5	1.14	4.61	9.7	17.8	0.61	5.20
$t\bar{t} \rightarrow \ell\ell$	4.3	4.1	1.15	0.62	5.8	4.3	1.12	0.73
$Wb\bar{b}$	8.4	6.3	1.72	0.54	10.2	5.2	1.85	0.52
Wjj	33.4	28.9	0.74	0.60	43.6	28.8	0.95	0.64
WW	0.4	0.3	0.01	0.00	0.6	0.3	0.00	0.00
WZ	0.4	0.3	0.09	0.01	0.4	0.2	0.09	0.01
Multijet	6.8	6.9	0.20	0.14	10.1	9.9	0.11	0.01
<i>Summed signals</i>	4.3	4.9	0.59	0.35	4.7	4.5	0.53	0.35
<i>Summed backgrounds</i>	57.5	65.3	5.04	6.54	80.3	66.7	4.71	7.20
<i>Summed signals + tqb</i>	60.0	68.6	5.18	6.76	83.1	69.8	4.81	7.43
<i>Summed signals + tb</i>	59.2	66.8	5.49	6.65	82.2	68.1	5.14	7.32
<i>Data</i>	60	73	4	9	78	58	10	8

TABLE X. Estimates of backgrounds and signal yields and the number of observed events in data after the cut-based selection for the electron and muon, = 1 tag and ≥ 2 tags analyses combined.

Source	s -channel search		t -channel search	
tb	4.5 ± 1.0	3.2 ± 0.8	3.2 ± 0.8	0.8
tqb	5.5 ± 1.2	7.0 ± 1.6	7.0 ± 1.6	1.6
$W + \text{jets}$	27.6 ± 7.6	55.9 ± 12.3	55.9 ± 12.3	12.3
$t\bar{t}$	102.9 ± 13.7	72.6 ± 9.7	72.6 ± 9.7	9.7
Multijet	17.2 ± 2.0	17.0 ± 2.0	17.0 ± 2.0	2.0
Total background	153.1 ± 24.5	148.7 ± 24.8	148.7 ± 24.8	24.8
Observed events	152	148	148	

C. Neural network analysis

A neural network is a multivariate statistical technique for separating signals from backgrounds. We use the MLPFIT [38] package to construct and implement the networks. In order for a neural network to approach the maximal signal-background separation, some optimization is required. This occurs in three steps: (1) judicious choice of signal and background pairs, (2) selection of input variables, and (3) optimization of training parameters.

1. Choice of signal-background pairs

We have chosen to create networks trained on single top quark signals against the two dominant backgrounds: $W + \text{jets}$ and $t\bar{t}$. For $W + \text{jets}$, we train using a $Wb\bar{b}$

Monte Carlo sample as this process best represents all $W + \text{jets}$ processes. For $t\bar{t}$, we train on $t\bar{t} \rightarrow \ell + \text{jets}$ which is the dominant background as opposed to the dilepton background which is small. Signal (background) events in each pair will have a network output value tending towards one (zero).

2. Choice of input variables

We start from a set of discriminating variables that each show some signal-background separation as discussed in Sec. VII A. Based on this, we optimize the input variables for each network by training with different combinations of variables and choosing the combination that produces the minimum testing error, i.e. the rate at which the network misclassifies signal events as background and vice versa. It is defined as the rate at which the network assigns a value of less than 0.5 to a signal event and a value of more than 0.5 to a background event.

We use the same variables for the electron and muon channel. However, owing to different resolutions and pseudorapidity ranges (see Sec. II B), we train the networks separately for the two.

3. Neural network training

Each network is composed of three layers of nodes: input, hidden, and output. Testing and training event sets are created from simulated signal and background samples. We divide the input samples such that 60% of the events are used for training and the remaining 40% for testing.

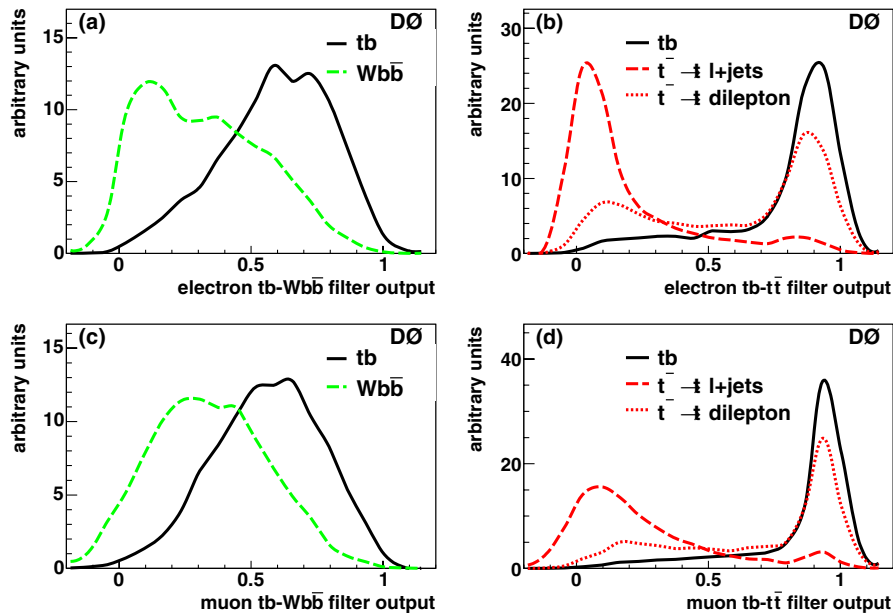


FIG. 15 (color online). Smoothed [41] neural network outputs in the s -channel. This figure shows the signal-background separation for (a) the filter for $Wb\bar{b}$ in the electron channel, (b) the filter for $t\bar{t} \rightarrow \ell + \text{jets}$ in the electron channel, (c) the filter for $Wb\bar{b}$ in the muon channel, and (d) the filter for $t\bar{t} \rightarrow \ell + \text{jets}$ in the muon channel where the backgrounds are dashed-lined and the single top quark signal is the solid line. All the curves are normalized to have equal area, so that the separation between signal and background can be best seen.

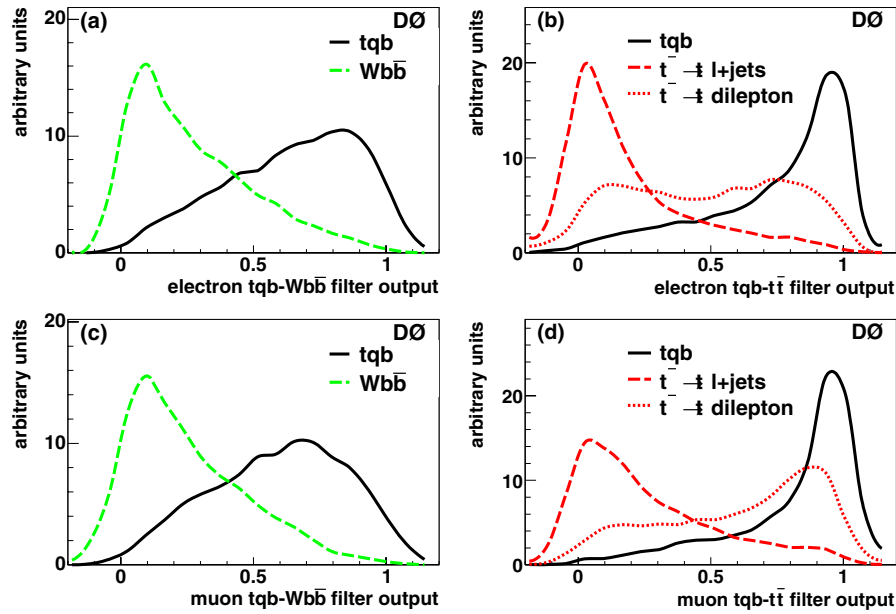


FIG. 16 (color online). Smoothed [41] neural network outputs in the t -channel. This figure shows the signal-background separation for (a) the filter for $Wb\bar{b}$ in the electron channel, (b) the filter for $t\bar{t} \rightarrow \ell + \text{jets}$ in the electron channel, (c) the filter for $Wb\bar{b}$ in the muon channel, and (d) the filter for $t\bar{t} \rightarrow \ell + \text{jets}$ in the muon channel where the backgrounds are dashed-lined and the single top quark signal is the solid line. All the curves are normalized to have equal area, so that the separation between signal and background can be best seen.

Training is effected with weighted events and the logarithm of all nonangular variables. We use a technique called early stopping [39] to determine the maximum number of epochs for training which prevents overtraining.

Each network is further tuned by varying the number of hidden nodes between 10 and 30 and then selecting the number of hidden nodes that returns the smallest testing error.

4. Neural network results

The above procedure produces eight unique networks: two signals (s -channel, t -channel) \times two backgrounds ($Wb\bar{b}$, $t\bar{t} \rightarrow \ell + \text{jets}$) \times two lepton flavors (e, μ).

Figures 15 and 16 show the output variable distributions from the networks in the s -channel and t -channel searches for electrons and muons. From the figures, it can be seen that these networks are highly efficient at separating the single top quark signal from the $t\bar{t} \rightarrow \ell + \text{jets}$ background. Studies have shown that these networks are not as effective for the $t\bar{t}$ dilepton background, which is fortunately small. The s -channel and t -channel networks are less efficient at separating the single top quark signal from the $Wb\bar{b}$ background as compared to $t\bar{t} \rightarrow \ell + \text{jets}$. In addition, we find these networks are equally effective in separating the Wjj and the misidentified lepton background as compared to the $Wb\bar{b}$ background. It should be noted that the output

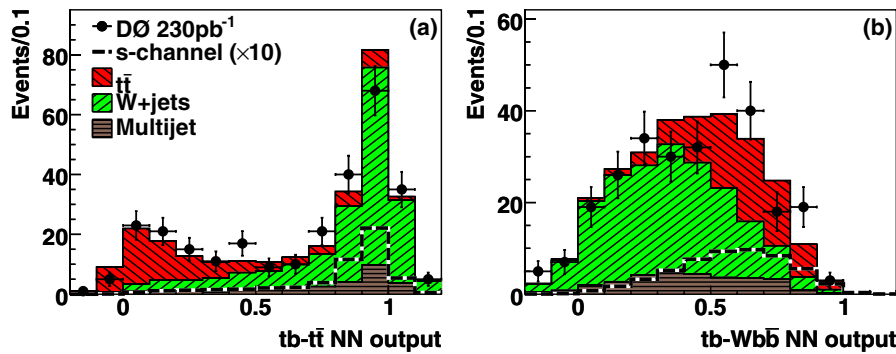


FIG. 17 (color online). Comparison of signal, background, and data for the neural network outputs in the s -channel, for the electron and muon channels combined, requiring at least one b -tag. This figure shows (a) the $t\bar{t}$ filter and (b) the $Wb\bar{b}$ filter. Signals are multiplied by ten.

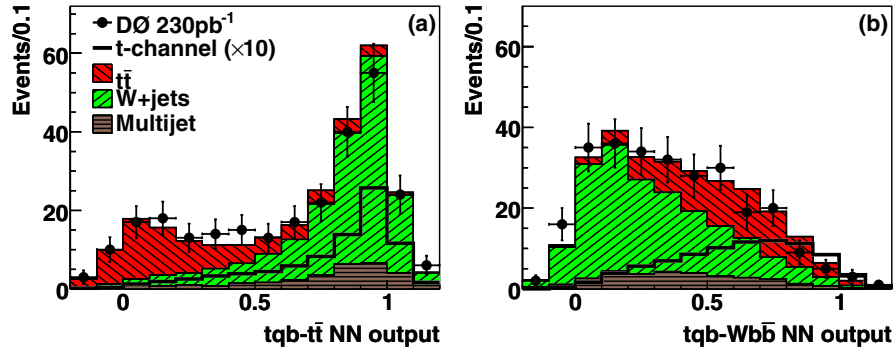


FIG. 18 (color online). Comparison of signal, background, and data for the neural network outputs in the t -channel, for the electron and muon channels combined, requiring at least one b -tag. This figure shows (a) the $t\bar{t}$ filter and (b) the $Wb\bar{b}$ filter. Signals are multiplied by ten.

variable from MLPFIT networks is not restricted to lie between zero and one.

Figures 17 and 18 show comparisons of the summed backgrounds to data for the s -channel and t -channel searches, for electrons, muons, single-tagged, and double-tagged samples combined. These distributions show that the background model reproduces the data very well, with χ^2/DF values of about 1.0 for the s -channel histograms and about 0.5 for the t -channel histograms. From the figures, it can be seen that the $t\bar{t} \rightarrow \ell + \text{jets}$ filters do indeed separate the $t\bar{t}$ background which clusters near zero, but does not affect the $W + \text{jets}$ and multijet backgrounds, which cluster near one. Similarly, the $Wb\bar{b}$ filters discriminate the $W + \text{jets}$ and multijet backgrounds, which cluster to the left of 0.5, but do not affect the $t\bar{t}$ background, which clusters to the right of 0.5. They also show that separation of the single top quark signal from background is not yet powerful enough since the background dominates even in the regions where the signal peaks.

Figure 19 shows the output of the $tb-t\bar{t}$ network versus the $tb-Wb\bar{b}$ network, and similarly for the tqb networks, again for electrons, muons, single-tagged, and double-tagged events combined. Again the background model reproduces the data well, reflected in χ^2/DF values of about 3 for both histograms.

VIII. SYSTEMATIC UNCERTAINTIES

We consider several sources of systematic uncertainties in this analysis, and study them separately for each signal and background source. Some of the uncertainties affect acceptance for simulated signals and backgrounds, others only affect background yield estimates. We also consider uncertainties coming from effects that change the shape of the discriminating variables. This section lists the uncertainties for each signal and background and their correlations.

We consider the following sources of systematic uncertainty:

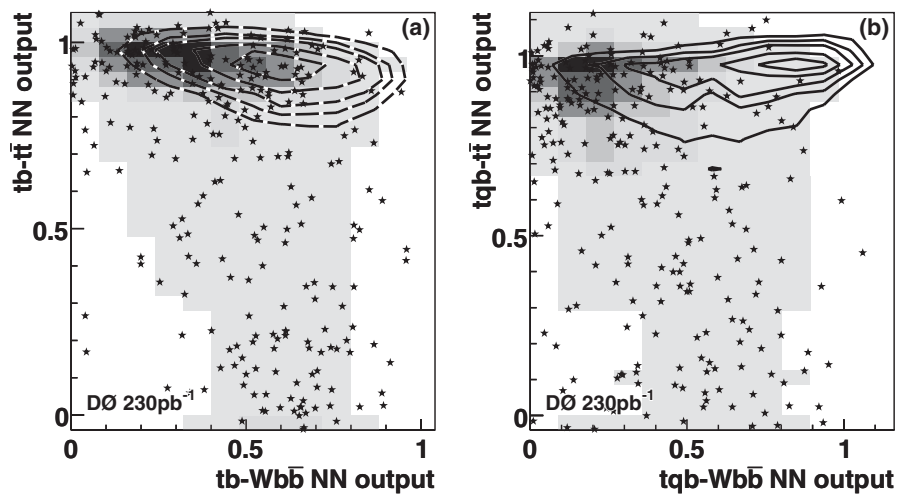


FIG. 19. Neural network outputs for both the $t\bar{t}$ versus $Wb\bar{b}$ filters in the (a) s -channel and (b) t -channel analyses, for the electron and muon channels combined, requiring at least one b -tag. The background sum is shown as the shaded area, the signal as contour lines, and the data as stars.

- (i) The b -tag modeling uncertainty includes components for the estimation of the b tagging efficiency in data for the various quark flavors, see Sec. II C.
- (ii) The jet energy calibration uncertainty reflects how well jet energies measured in the simulation reflect jet energies measured in data, and includes jet energy scale uncertainty as well as modeling of jet energy resolution in the simulation, see Sec. II B.
- (iii) The trigger modeling uncertainty includes components for the estimation of the efficiency of the various trigger requirements in data, see Sec. III.
- (iv) The jet fragmentation uncertainty includes the uncertainty in modeling of initial- and final-state radiation as well as the difference in the fragmentation model between PYTHIA and HERWIG [19].
- (v) The uncertainty on the correction factor for simulated samples to account for the jet identification efficiency as described in Sec. II B.
- (vi) The uncertainty on the correction factor for lepton identification efficiency in simulated samples as described in Sec. II B.
- (vii) The cross section and branching fraction uncertainties from the yield normalization of simulated backgrounds. The $t\bar{t}$ and single top quark cross section uncertainties include the uncertainty in the top quark mass measurement.
- (viii) The uncertainty on the normalization of the multijet and $W + \text{jets}$ background yields to the data.
- (ix) The uncertainty on the integrated luminosity measurement.

The uncertainty on the multijet background normalization includes two components: the estimate of the rate to misidentify a jet as an isolated lepton in the data, and the b -tagging probability in the multijet data sample.

The uncertainty for the Wjj and $Wb\bar{b}$ backgrounds includes several components: the normalization of the $W + \text{jets}$ background to data before b tagging, the b -tagging probability estimate, and the fraction of $Wb\bar{b}$ events in the $W + \text{jets}$ sample. Owing to the normalization to data, the $Wb\bar{b}$ and Wjj tagged yield estimates are not affected by any of the systematic uncertainties that affect the other simulated samples. The exception to this is b tagging, which is applied after normalization. There is still an effect on the shape of the Wjj and $Wb\bar{b}$ distributions from uncertainty components that vary bin-by-bin.

Table XI shows the systematic uncertainty values for each signal and background component. The range is given for the different analysis channels, electron and muon as well as single tags and double tags.

Note that the $W + \text{jets}$ background includes small contributions from WW and WZ , whose uncertainties are also included in the limit setting calculation. Furthermore, the normalization for $Wb\bar{b}$ and Wjj accounts for the other simulated backgrounds and thus their uncertainties in principle also affect $Wb\bar{b}$ and Wjj . However, the other

TABLE XI. Range of relative systematic uncertainty values in percent for the various signal and background samples in the different analysis channels.

	tb	tqb	$t\bar{t}$	$W + \text{jets}$	multijet
<i>Signal and background acceptance</i>					
b -tag modeling	5–20	8–20	6–20	7–20	...
Jet energy calibration	6–20	6–15	3–11
Trigger modeling	2–6	2–6
Jet fragmentation	5	5	7
Jet identification	1–13	5–11	1–4
Lepton identification	4	4	4
<i>Background normalization</i>					
Theory cross sections	16	15	18
Normalization to data	5–16	5–16
Luminosity	6.5	6.5	6.5

simulated backgrounds only contribute about 3% to the pretagged yield, which means their uncertainties are negligible compared to the overall normalization uncertainties.

IX. CROSS SECTION LIMITS

We use a Bayesian approach [40] to calculate limits on the cross section for single top quark production in the s -channel and t -channel modes. The limits are derived from a likelihood function that is proportional to the probability to obtain the number of observed events. In the cut-based analysis, we count the total number of observed events, and in the neural network analysis, we use the two-dimensional distributions of the $t\bar{t}$ versus $Wb\bar{b}$ network outputs.

A. Bayesian approach

We assume that the probability to observe a count D , if the mean count is d , is given by the Poisson distribution:

$$p(D|d) = \frac{e^{-d}d^D}{\Gamma(D+1)}, \quad (7)$$

where Γ is the gamma function. The mean count d is a sum of the predicted contributions from the signal and background sources:

$$d = \alpha \mathcal{L} \sigma + \sum_{i=1}^N b_i \equiv a\sigma + \sum_{i=1}^N b_i, \quad (8)$$

where α is the signal acceptance, \mathcal{L} the integrated luminosity, σ the signal cross section (the quantity of interest), b_i the mean count for background source i , and $a \equiv \alpha \mathcal{L}$ is the effective luminosity for the signal. For the s -channel (t -channel) search, the background b_i includes the t -channel (s -channel) process. The likelihood function $L(D|d)$ is proportional to $p(D|d)$.

For two or more independent channels, we simply replace the single channel likelihood by a product of likelihoods:

$$L(\mathbf{D}|\mathbf{d}) \equiv L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) = \prod_{i=1}^M L(D_i|d_i), \quad (9)$$

where \mathbf{D} and \mathbf{d} , respectively, represent vectors of the observed counts and the mean counts for the sources of signal and background, and \mathbf{a} and \mathbf{b} are similarly vectors effective luminosity and backgrounds. The product goes over the total number of bins in all channels M .

We use Bayes' theorem to compute the posterior probability density of the parameters, $p(\sigma, \mathbf{a}, \mathbf{b}|\mathbf{D})$, which is then integrated with respect to the parameters \mathbf{a} and \mathbf{b} to obtain the posterior density for the signal cross section, given the observed distribution of counts \mathbf{D} :

$$p(\sigma|\mathbf{D}) = \frac{1}{\mathcal{N}} \iint L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) \pi(\sigma, \mathbf{a}, \mathbf{b}) d\mathbf{a} d\mathbf{b}. \quad (10)$$

Here \mathcal{N} is an overall normalization obtained from the requirement $\int p(\sigma|\mathbf{D}) d\sigma = 1$, where the integration is performed numerically up to an upper bound σ_{\max} where the value of the posterior is sufficiently close to zero. In this analysis, varying σ_{\max} from 30 pb to 150 pb has no effect on the result. Also, $\pi(\sigma, \mathbf{a}, \mathbf{b})$ is the prior probability that encodes what we know about the parameters σ , \mathbf{a} , and \mathbf{b} . We assume that any prior knowledge of \mathbf{a} and \mathbf{b} is independent of the cross section σ , in which case we may write the prior density as

$$\pi(\sigma, \mathbf{a}, \mathbf{b}) = \pi(\mathbf{a}, \mathbf{b}|\sigma) \pi(\sigma) = \pi(\mathbf{a}, \mathbf{b}) \pi(\sigma). \quad (11)$$

We use a flat prior for σ : $\pi(\sigma) = 1/\sigma_{\max}$. The posterior probability density for the signal cross section is therefore

$$p(\sigma|\mathbf{D}) = \frac{1}{\mathcal{N}} \iint L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) \pi(\mathbf{a}, \mathbf{b}) d\mathbf{a} d\mathbf{b}. \quad (12)$$

The Bayesian upper limit σ_{UL} at confidence level β is the solution of

$$\int_0^{\sigma_{\text{UL}}} p(\sigma|\mathbf{D}) d\sigma = \beta. \quad (13)$$

The integral in Eq. (12) is done numerically using Monte Carlo importance sampling: we generate a large number K of randomly sampled points $(\mathbf{a}_k, \mathbf{b}_k)$ that represents the prior density $\pi(\mathbf{a}, \mathbf{b})$, and estimate the posterior using

$$\iint L(\mathbf{D}|\sigma, \mathbf{a}, \mathbf{b}) \pi(\mathbf{a}, \mathbf{b}) d\mathbf{a} d\mathbf{b} = \frac{1}{K} \sum_{k=1}^K L(\mathbf{D}|\sigma, \mathbf{a}_k, \mathbf{b}_k). \quad (14)$$

B. Definition of the prior probability

The prior $\pi(\mathbf{a}, \mathbf{b})$ encodes our knowledge of the effective signal luminosities and the background yields: we

have estimates of the parameters and the associated uncertainties from the different systematic effects discussed in Sec. VIII. In the case of the cut-based analysis, since we consider the total yield for any source of signal or background, the different uncertainties affect the overall normalization only. In the neural network analysis, since we consider distributions, we separate the uncertainties into two classes: those that alter only the overall normalization, such as the luminosity measurement and theory cross sections; and those that also alter the shapes of the distributions, such as the trigger modeling, jet energy calibration, jet energy resolution, jet identification, and b -tag modeling.

The normalization effects are modeled by sampling the effective signal luminosities \mathbf{a} and the background yields \mathbf{b} from a multivariate Gaussian, with a vector of means given by the estimates of the yields, and covariance matrix computed from the associated uncertainties. The covariance matrix takes into account the correlations of the systematic uncertainties across the different sources of signal and background. Each entry in the covariance matrix is calculated as follows:

$$c_{ij} = y_i y_j \sum_k f_{ik} f_{jk}, \quad (15)$$

where y_i (y_j) is the yield for the i th (j)th source of background or signal from Table V, and f_{ik} is the corresponding fractional uncertainty from the k th component of systematic uncertainty, for the i th source.

The shape effects are modeled by shifting, one by one, the trigger modeling, jet energy calibration, b -tag modeling, and so on, by plus or minus 1 standard deviation with respect to their nominal values. For each systematic effect, we have three distributions: the nominal, and those from the plus and minus shifts. The systematic uncertainty in each bin is then sampled from a Gaussian distribution with mean defined by the nominal yield in that bin, and width defined by the plus and minus shifts. The sampled shifts are added linearly to the yields generated from the sampling of the normalization-only systematic uncertainties. For each shape-changing systematic, we assume 100% correlation across all bins and sources. This procedure might result in a negative signal acceptance or background yield for a given sample. In this rare case, the likelihood for this particular bin in this particular sample is set to a constant value of $1/\sigma_{\max}$ (see above), independent of the cross section.

X. RESULTS

For both the s -channel and t -channel searches, we compute an observed limit as well as an expected limit. We define the latter as the limit obtained if the observed counts were exactly equal to the background prediction in each bin of each channel. It is for these floating point values that we use a Gamma function in the denominator of Eq. (7) rather than a simple factorial. The different tag multiplic-

TABLE XII. Expected and observed upper limits (in picobarns) at the 95% confidence level, on the production cross sections of single top quarks in the s -channel (tb) and t -channel (tqb) searches, for the electron and muon channels combined, with all systematic effects included.

	Expected limits		Observed limits	
	tb	tqb	tb	tqb
Initial selection	14.5	16.5	13.0	13.6
Cut-based	9.8	12.4	10.6	11.3
Neural networks	4.5	5.8	6.4	5.0

ities ($= 1$ tag and ≥ 2 tags) and lepton flavor (electron and muon) are combined as shown in Eq. (9).

The expected and observed upper limits at the 95% confidence level, after the initial event selection, and from the cut-based and neural network analyses, are shown in Table XII for the electron and muon channels combined, and with all systematic effects included. We see that the expected limits improve upon applying cuts on the discriminating variables, but that tighter expected limits are obtained when the variables are combined using our neural networks method. The observed posterior probability densities as a function of the s -channel and t -channel cross sections are shown in Fig. 20 for the cut-based analysis and in Fig. 21 for the neural network analysis.

The method described so far yields limits on the s -channel or t -channel cross sections separately. This requires some assumptions about whichever of the two signal processes is not being considered. In this particular analysis, we have assumed that in the s -channel (t -channel) search, the t -channel (s -channel) contributes as a SM background. This assumption is, however, not necessary. Instead, we can set limits on both the s -channel and

t -channel cross sections simultaneously. We accomplish this by generalizing the likelihood so that it depends explicitly on the two cross sections σ_s and σ_t . Equation (8) for the mean count d then becomes

$$d = \alpha_s \mathcal{L} \sigma_s + \alpha_t \mathcal{L} \sigma_t + \sum_i b_i. \quad (16)$$

The backgrounds b_i now include only the nonsingle top quark sources.

In order to exploit the sensitivity to both the s -channel and t -channel signals, we combine the output of the neural networks in both searches. We calculate a signal probability P in each bin of the histograms in Fig. 19:

$$P_{s(t)} = \frac{n_{s(t)}}{n_s + n_t + \sum_i b_i}, \quad (17)$$

for the s -channel (t -channel) search, where n_s and n_t are the yields for the tb and tqb samples, respectively, and the sum in the denominator is over all the nonsingle top quark backgrounds in that bin. We then evaluate P_s and P_t simultaneously for each event and fill histograms of P_s versus P_t . As before, we consider a Poisson probability for the likelihood in each bin. We assume a flat prior for each of the signal cross sections, σ_s and σ_t , as explained before. Equations (12) and (14) can then be used to define the posterior probability density for different values of the s - and t -channel cross sections. The limit at a fixed confidence level is then given by a contour of constant posterior probability density enclosing a fraction of volume corresponding to this confidence level using an equation analogous to Eq. (13), but in two dimensions.

Figure 22 shows contours of observed posterior density in the σ_s versus σ_t plane for the neural network analysis. To illustrate the sensitivity of this analysis to different

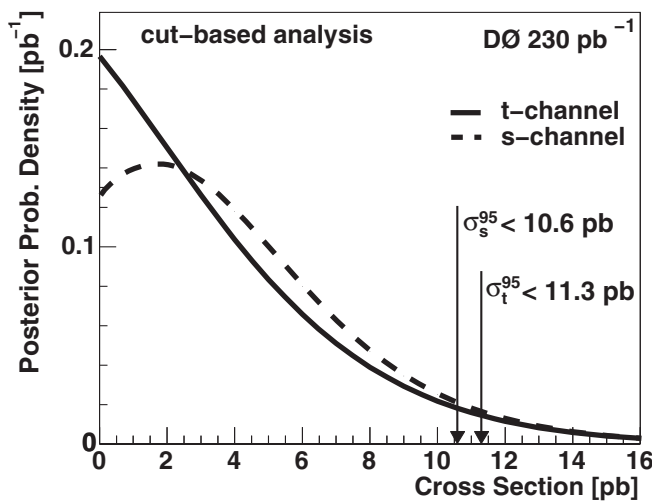


FIG. 20. The observed posterior probability density as a function of the single top quark cross section for the cut-based analysis, for the electron and muon channels combined in the s -channel and the t -channel searches.

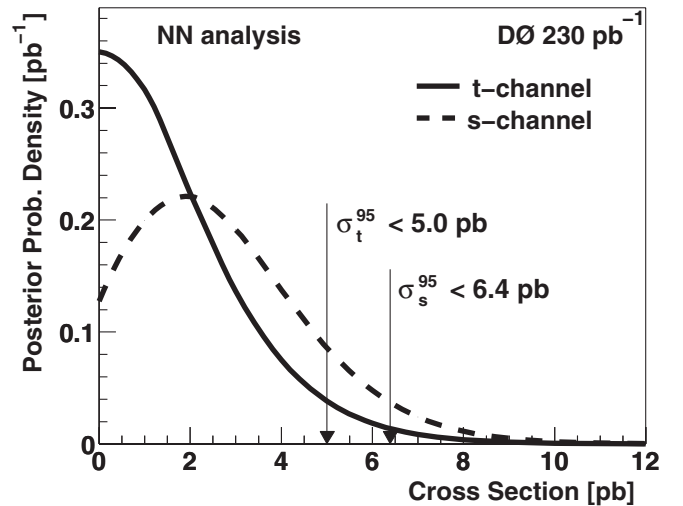


FIG. 21. The observed posterior probability density as a function of the single top quark cross section for the neural network analysis, for the electron and muon channels combined in the s -channel and the t -channel searches.

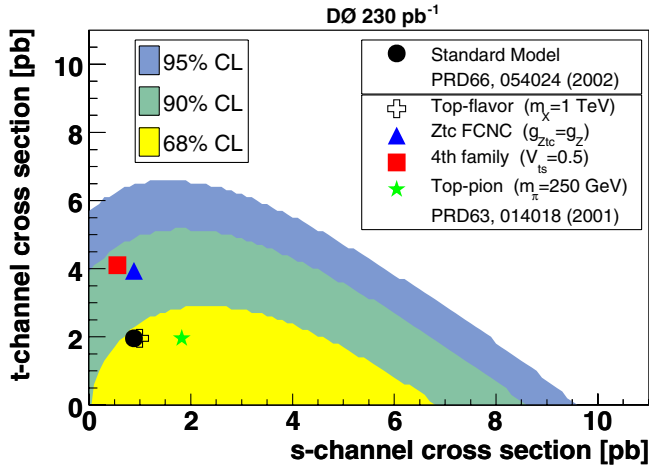


FIG. 22 (color online). Allowed regions at the 68%, 90%, and 95% confidence levels on the observed posterior density distribution as a function of both the s -channel and t -channel cross sections in the neural networks analysis. Several representative nonstandard model contributions from Ref. [5] are also shown.

contributions, the expected SM cross section as well as several representative non-SM contributions are also shown [5].

XI. SUMMARY

We have analyzed electron + jet and muon + jet events containing exactly one or more than one b jet, identified

with a secondary vertex algorithm, and find no evidence for the electroweak production of single top quarks in 230 pb^{-1} of data collected by the D0 detector at $\sqrt{s} = 1.96 \text{ TeV}$. The upper limits at the 95% confidence level on the cross section for s -channel and t -channel processes are 10.6 pb and 11.3 pb, respectively, using event counts in a cut-based analysis, and 6.4 pb and 5.0 pb, respectively, using binned likelihoods in a neural network analysis. The neural network-base limits presented here and in Ref. [22] are significantly more stringent than those previously published [19–21]. They are also close to the sensitivity required to probe models of physics beyond the standard model.

ACKNOWLEDGMENTS

We thank the staffs at Fermilab and collaborating institutions and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom, and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP, and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC, and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); Research Corporation; Alexander von Humboldt Foundation; and the Marie Curie Program.

-
- [1] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995); S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74**, 2632 (1995).
- [2] K. Hagiwara *et al.* (Particle Data Group Collaboration), Phys. Rev. D **66**, 010001 (2002).
- [3] A. P. Heinson, A. S. Belyaev, and E. E. Boos, Phys. Rev. D **56**, 3114 (1997).
- [4] C. R. Chen, F. Larios, and C. P. Yuan, Phys. Lett. B **631**, 126 (2005).
- [5] T. M. P. Tait and C.-P. Yuan, Phys. Rev. D **63**, 014018 (2000).
- [6] M. C. Smith and S. S. Willenbrock, Phys. Rev. D **54**, 6696 (1996).
- [7] T. Stelzer, Z. Sullivan, and S. Willenbrock, Phys. Rev. D **56**, 5919 (1997).
- [8] B. W. Harris, E. Laenen, L. Phaf, Z. Sullivan, and S. Weinzierl, Phys. Rev. D **66**, 054024 (2002).
- [9] Z. Sullivan, Phys. Rev. D **70**, 114012 (2004).
- [10] J. Campbell, R. K. Ellis, and F. Tramontano, Phys. Rev. D **70**, 094012 (2004).
- [11] Q.-H. Cao, R. Schwienhorst, and C.-P. Yuan, Phys. Rev. D **71**, 054023 (2005).
- [12] Q.-H. Cao, R. Schwienhorst, J. A. Benitez, R. Brock, and C.-P. Yuan, Phys. Rev. D **72**, 094027 (2005).
- [13] R. Bonciani *et al.*, Nucl. Phys. **B529**, 424 (1998); M. Cacciari *et al.*, J. High Energy Phys. 04 (2004) 068; N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003).
- [14] Pseudorapidity is defined as $\eta = -\ln(\tan\frac{\theta}{2})$, where θ is the polar angle with the origin at the primary vertex.
- [15] G. Mahlon and S. Parke, Phys. Rev. D **53**, 4886 (1996); S. Parke and Y. Shadmi, Phys. Lett. B **387**, 199 (1996); G. Mahlon and S. Parke, Phys. Rev. D **55**, 7249 (1997).
- [16] E. E. Boos and A. V. Sherstnev, Phys. Lett. B **534**, 97 (2002).
- [17] M. L. Mangano *et al.*, J. High Energy Phys. 07 (2003) 001.
- [18] S. Mrenna and P. Richardson, J. High Energy Phys. 05 (2004) 040.
- [19] B. Abbott *et al.* (D0 Collaboration), Phys. Rev. D **63**, 031101 (2000); V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **517**, 282 (2001).
- [20] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **65**, 091102 (2002); **69**, 052003 (2004).
- [21] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 012005 (2005).
- [22] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **622**, 265 (2005).

- [23] V.M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **565**, 463 (2006).
- [24] S. Abachi *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **338**, 185 (1994).
- [25] R.E. Kalman, Journal of Basic Engineering **82**, 35 (1960).
- [26] J.F. Kozminski, Ph.D. thesis, Michigan State University [Fermilab-thesis-2005-21, Report No. UMI-31-71481, 2005].
- [27] Jets are defined using the iterative, seed-based cone algorithm with radius $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.5$, including midpoints as described in G.C. Blazey *et al.*, in *Proceedings of the Workshop on QCD and Weak Boson Physics in Run II*, edited by U. Baur, R. K. Ellis, and D. Zeppenfeld (Fermilab Report No. FERMILAB-PUB-00-297, 2000), pp. 47–77.
- [28] E. Boos *et al.* (CompHEP Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **534**, 250 (2004).
- [29] J.M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).
- [30] V.M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **626**, 45 (2005).
- [31] B. Abbott *et al.* (D0 Collaboration), Phys. Rev. D **61**, 072001 (2000).
- [32] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135**, 238 (2001).
- [33] H.L. Lai *et al.* (CTEQ Collaboration), Eur. Phys. J. C **12**, 375 (2000); J. Pumplin *et al.* (CTEQ collaboration), J. High Energy Phys. 07 (2002) 012.
- [34] S. Jadach, J.H. Kuhn, and Z. Was, Comput. Phys. Commun. **64**, 275 (1991).
- [35] D.J. Lange, Nucl. Instrum. Methods Phys. Res., Sect. A **462**, 152 (2001).
- [36] R. Brun *et al.*, CERN Program Library Long Writeup Report No. W5013, 1994.
- [37] E. Boos and L. Dudko, Nucl. Instrum. Methods Phys. Res., Sect. A **502**, 486 (2003).
- [38] J. Schwindling, <http://schwind.home.cern.ch/schwind/MLPfit.html>.
- [39] G. Orr and K. Müller, *Neural Networks: Tricks of the Trade* (Springer-Verlag, Berlin, 1998), p. 55.
- [40] I. Bertram *et al.*, Fermilab Report No. FERMILAB-TM-2104 2000.
- [41] J.H. Friedman, CERN Yellow Report No. 74-23.