Measurement of the forward-backward asymmetry in top quark-antiquark production in $p\bar{p}$ collisions using the lepton + jets channel

V. M. Abazov,[31] B. Abbott,[67] B. S. Acharya,[25] M. Adams,[46] T. Adams,[44] J. P. Agnew,[41] G. D. Alexeev,[31] G. Alkhazov,[35] A. Alton,[56,a] A. Askew,[44] S. Atkins,[54] K. Augsten,[7] C. Avila,[5] F. Badaud,[10] L. Bagby,[45] B. Baldin,[45] D. V. Bandurin,[73] S. Banerjee,[25] E. Barberis,[55] P. Baringer,[53] J. F. Bartlett,[45] U. Bassler,[15] V. Bazterra,[46] A. Bean,[53] M. Begalli,[2] L. Bellantoni,[45] S. B. Beri,[23] G. Bernardi,[14] R. Bernhard,[19] I. Bertram,[39] M. Besançon,[15] R. Beuselinck,[40] P. C. Bhat,[45] S. Bhatia,[58] V. Bhatnagar,[23] G. Blazey,[47] S. Blessing,[44] K. Bloom,[59] A. Boehnlein,[45] D. Boline,[64] E. E. Boos,[33] G. Borissov,[39] M. Borysova,[38,l] A. Brandt,[70] O. Brandt,[20] R. Brock,[57] A. Bross,[45] D. Brown,[14] X. B. Bu,[45] M. Buehler,[45] V. Buescher,[21] V. Bunichev,[33] S. Burdin,[39,b] C. P. Buszello,[37] E. Camacho-Pérez,[28] B. C. K. Casey,[45] H. Castilla-Valdez,[28] S. Caughron,[57] S. Chakrabarti,[64] K. M. Chan,[51] A. Chandra,[72] E. Chapon,[15] G. Chen,[53] S. W. Cho,[27] S. Choi,[27] B. Choudhary,[24] S. Cihangir,[45] D. Claes,[59] J. Clutter,[53] M. Cooke,[45,k] W. E. Cooper,[45] M. Corcoran,[72] F. Couderc,[15] M.-C. Cousinou,[12] D. Cutts,[69] A. Das,[42] G. Davies,[40] S. J. de Jong,[29,30] E. De La Cruz-Burelo,[28] F. Déliot,[15] R. Demina,[63] D. Denisov,[45] S. P. Denisov,[34] S. Desai,[45] C. Deterre,[20,c] K. DeVaughan,[59] H. T. Diehl,[45] M. Diesburg,[45] P. F. Ding,[41] A. Dominguez,[59] A. Dubey,[24] L. V. Dudko,[33] A. Duperrin,[12] S. Dutt,[23] M. Eads,[47] D. Edmunds,[57] J. Ellison,[43] V. D. Elvira,[45] Y. Enari,[14] H. Evans,[49] V. N. Evdokimov,[34] A. Falkowski,[15,n] A. Fauré,[15] L. Feng,[47] T. Ferbel,[63] F. Fiedler,[21] F. Filthaut,[29,30] W. Fisher,[57] H. E. Fisk,[45] M. Fortner,[47] H. Fox,[39] S. Fuess,[45] P. H. Garbincius,[45] A. Garcia-Bellido,[63] J. A. García-González,[28] V. Gavrilov,[32] W. Geng,[12,57] C. E. Gerber,[46] Y. Gershtein,[60] G. Ginther,[45,63] O. Gogota,[38] G. Golovanov,[31] P. D. Grannis,[64] S. Greder,[16] H. Greenlee,[45] G. Grenier,[17] Ph. Gris,[10] J.-F. Grivaz,[13] A. Grohsjean,[15,c] S. Grünendahl,[45] M. W. Grünewald,[26] T. Guillemin,[13] G. Gutierrez,[45] P. Gutierrez,[67] J. Haley,[68] L. Han,[4] K. Harder,[41] A. Harel,[63] J. M. Hauptman,[52] J. Hays,[40] T. Head,[41] T. Hebbeker,[18] D. Hedin,[47] H. Hegab,[68] A. P. Heinson,[43] U. Heintz,[69] C. Hensel,[1] I. Heredia-De La Cruz,[28,d] K. Herner,[45] G. Hesketh,[41,f] M. D. Hildreth,[51] R. Hirosky,[73] T. Hoang,[44] J. D. Hobbs,[64] B. Hoeneisen,[9] J. Hogan,[72] M. Hohlfeld,[21] J. L. Holzbauer,[58] I. Howley,[70] Z. Hubacek,[7,15] V. Hynek,[7] I. Iashvili,[62] Y. Ilchenko,[71] R. Illingworth,[45] A. S. Ito,[45] S. Jabeen,[45,m] M. Jaffré,[13] A. Jayasinghe,[67] M. S. Jeong,[27] R. Jesik,[40] P. Jiang,[4] K. Johns,[42] E. Johnson,[57] M. Johnson,[45] A. Jonckheere,[45] P. Jonsson,[40] J. Joshi,[43] A. W. Jung,[45] A. Juste,[36] E. Kajfasz,[12] D. Karmanov,[33] I. Katsanos,[59] R. Kehoe,[71] S. Kermiche,[12] N. Khalatyan,[45] A. Khanov,[68] A. Kharchilava,[62] Y. N. Kharzheev,[31] I. Kiselevich,[32] J. M. Kohli,[23] A. V. Kozelov,[34] J. Kraus,[58] A. Kumar,[62] A. Kupco,[8] T. Kurča,[17] V. A. Kuzmin,[33] S. Lammers,[49] P. Lebrun,[17] H. S. Lee,[27] S. W. Lee,[52] W. M. Lee,[45] X. Lei,[42] J. Lellouch,[14] D. Li,[14] H. Li,[73] L. Li,[43] Q. Z. Li,[45] J. K. Lim,[27] D. Lincoln,[45] J. Linnemann,[57] V. V. Lipaev,[34] R. Lipton,[45] H. Liu,[71] Y. Liu,[4] A. Lobodenko,[35] M. Lokajicek,[8] R. Lopes de Sa,[64] R. Luna-Garcia,[28,g] A. L. Lyon,[45] A. K. A. Maciel,[1] R. Madar,[19] R. Magaña-Villalba,[28] S. Malik,[59] V. L. Malyshev,[31] J. Mansour,[20] J. Martínez-Ortega,[28] R. McCarthy,[64] C. L. McGivern,[41] M. M. Meijer,[29,30] A. Melnitchouk,[45] D. Menezes,[47] P. G. Mercadante,[3] M. Merkin,[33] A. Meyer,[18] J. Meyer,[20,i] F. Miconi,[16] N. K. Mondal,[25] M. Mulhearn,[73] E. Nagy,[12] M. Narain,[69] R. Nayyar,[42] H. A. Neal,[56] J. P. Negret,[5] P. Neustroev,[35] H. T. Nguyen,[73] T. Nunnemann,[22] D. Orbaker,[63] J. Orduna,[72] N. Osman,[12] J. Osta,[51] A. Pal,[70] N. Parashar,[50] V. Parihar,[69] S. K. Park,[27] R. Partridge,[69,e] N. Parua,[49] A. Patwa,[65,j] B. Penning,[45] M. Perfilov,[33] Y. Peters,[41] K. Petridis,[41] G. Petrillo,[63] P. Pétroff,[13] M.-A. Pleier,[65] V. M. Podstavkov,[45] A. V. Popov,[34] M. Prewitt,[72] D. Price,[41] N. Prokopenko,[34] J. Qian,[56] A. Quadt,[20] B. Quinn,[58] P. N. Ratoff,[39] I. Razumov,[34] I. Ripp-Baudot,[16] F. Rizatdinova,[68] M. Rominsky,[45] A. Ross,[39] C. Royon,[15] P. Rubinov,[45] R. Ruchti,[51] G. Sajot,[11] A. Sánchez-Hernández,[28] M. P. Sanders,[22] A. S. Santos,[1,h] G. Savage,[45] M. Savitskyi,[38] L. Sawyer,[54] T. Scanlon,[40] R. D. Schamberger,[64] Y. Scheglov,[35] H. Schellman,[48] C. Schwanenberger,[41] R. Schwienhorst,[57] J. Sekaric,[53] H. Severini,[67] E. Shabalina,[20] V. Shary,[15] S. Shaw,[57] A. A. Shchukin,[34] V. Simak,[7] P. Skubic,[67] P. Slattery,[63] D. Smirnov,[51] G. R. Snow,[59] J. Snow,[66] S. Snyder,[65] S. Söldner-Rembold,[41] L. Sonnenschein,[18] K. Soustruznik,[6] J. Stark,[11] D. A. Stoyanova,[34] M. Strauss,[67] L. Suter,[41] P. Svoisky,[67] M. Titov,[15] V. V. Tokmenin,[31] Y.-T. Tsai,[63] D. Tsybychev,[64] B. Tuchming,[15] C. Tully,[61] L. Uvarov,[35] S. Uvarov,[35] S. Uzunyan,[47] R. Van Kooten,[49] W. M. van Leeuwen,[29] N. Varelas,[46] E. W. Varnes,[42] I. A. Vasilyev,[34] A. Y. Verkheev,[31] L. S. Vertogradov,[31] M. Verzocchi,[45] M. Vesterinen,[41] D. Vilanova,[15] P. Vokac,[7] H. D. Wahl,[44] M. H. L. S. Wang,[45] J. Warchol,[51] G. Watts,[74] M. Wayne,[51] J. Weichert,[21] L. Welty-Rieger,[48] M. R. J. Williams,[49] G. W. Wilson,[53] M. Wobisch,[54] D. R. Wood,[55] T. R. Wyatt,[41] Y. Xie,[45] R. Yamada,[45] S. Yang,[4] T. Yasuda,[45] Y. A. Yatsunenko,[31] W. Ye,[64] Z. Ye,[45] H. Yin,[45] K. Yip,[65] S. W. Youn,[45] J. M. Yu,[56] J. Zennamo,[62] T. G. Zhao,[41] B. Zhou,[56] J. Zhu,[56] M. Zielinski,[63] D. Zieminska,[49] and L. Zivkovic[14]

(The D0 Collaboration)

[1] *LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[2] *Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[3] *Universidade Federal do ABC, Santo André, Brazil*
[4] *University of Science and Technology of China, Hefei, People's Republic of China*
[5] *Universidad de los Andes, Bogotá, Colombia*

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

[63]University of Rochester, Rochester, New York 14627, USA
[64]State University of New York, Stony Brook, New York 11794, USA
[65]Brookhaven National Laboratory, Upton, New York 11973, USA
[66]Langston University, Langston, Oklahoma 73050, USA
[67]University of Oklahoma, Norman, Oklahoma 73019, USA
[68]Oklahoma State University, Stillwater, Oklahoma 74078, USA
[69]Brown University, Providence, Rhode Island 02912, USA
[70]University of Texas, Arlington, Texas 76019, USA
[71]Southern Methodist University, Dallas, Texas 75275, USA
[72]Rice University, Houston, Texas 77005, USA
[73]University of Virginia, Charlottesville, Virginia 22904, USA
[74]University of Washington, Seattle, Washington 98195, USA

We present a measurement of the forward–backward asymmetry in top quark–antiquark production using the full Tevatron Run II data set collected by the D0 experiment at Fermilab. The measurement is performed in lepton + jets final states using a new kinematic fitting algorithm for events with four or more jets and a new partial reconstruction algorithm for events with only three jets. Corrected for detector acceptance and resolution effects, the asymmetry is evaluated to be $A_{\mathrm{FB}} = (10.6 \pm 3.0)\%$. Results are consistent with the standard model predictions which range from 5.0% to 8.8%. We also present the dependence of the asymmetry on the invariant mass of the top quark–antiquark system and the difference in rapidities of the top quark and antiquark.

I. INTRODUCTION

A. Motivation and definitions

Over the last five years both experiments at the Fermilab Tevatron Collider measured positive forward-backward asymmetries in the production of top quark–antiquark pairs in proton-antiproton collisions ($p\bar{p} \to t\bar{t}$) [1–5].

The reported values were consistently above predictions of the standard model of particle physics (SM) [6,7]. In particular, the CDF Collaboration observed a strong rise of the asymmetry with the invariant mass of the $t\bar{t}$ system, $m_{t\bar{t}}$ [3]. The dependence of the asymmetry on $m_{t\bar{t}}$ in D0 data, as measured in Ref. [4], was statistically compatible with both the SM predictions and with the CDF result. Several beyond-the-SM scenarios were suggested to explain the measured A_{FB} values [8], in particular using the framework of parity-violating strong interactions suggested in Ref. [9]. In this paper we report new results from the D0 experiment based on the full data set collected during Run II of the Fermilab Tevatron Collider, which supersede the result of Ref. [4].

In proton-antiproton collisions, top quark–antiquark pairs are predominantly produced via valence quark–antiquark annihilation. Thus, the direction of the proton (antiproton) almost always coincides with the direction of the incoming quark (antiquark). We define the difference in rapidity[1] between the top quark (y_t) and antiquark ($y_{\bar{t}}$) as

$$\Delta y = y_t - y_{\bar{t}}. \tag{1}$$

We refer to the events that have $\Delta y > 0$ as "forward," and to those with $\Delta y < 0$ as "backward." The forward-backward asymmetry in $t\bar{t}$ production is defined as

[a]With visitor from Augustana College, Sioux Falls, South Dakota, USA.
[b]With visitor from The University of Liverpool, Liverpool, United Kingdom.
[c]With visitor from DESY, Hamburg, Germany.
[d]With visitor from Universidad Michoacana de San Nicolas de Hidalgo, Morelia, Mexico.
[e]With visitor from SLAC, Menlo Park, California, USA.
[f]With visitor from University College London, London, United Kingdom.
[g]With visitor from Centro de Investigacion en Computacion - IPN, Mexico City, Mexico.
[h]With visitor from Universidade Estadual Paulista, São Paulo, Brazil.
[i]With visitor from Karlsruher Institut für Technologie (KIT) - Steinbuch Centre for Computing (SCC) D-76128 Karlsruhe, Germany.
[j]With visitor from Office of Science, U.S. Department of Energy, Washington, D.C. 20585, USA.
[k]With visitor from American Association for the Advancement of Science, Washington, D.C. 20005, USA.
[l]With visitor from Kiev Institute for Nuclear Research, Kiev, Ukraine.
[m]With visitor from University of Maryland, College Park, Maryland 20742, USA.
[n]With visitor from Laboratoire de Physique Theorique, Orsay, France.

[1]The rapidity y is defined as $y = \frac{1}{2}\ln\left[(E + p_z)/(E - p_z)\right]$, where E is the particle's energy and p_z is its momentum along the z axis, which corresponds to the direction of the incoming proton.

$$A_{\text{FB}} = \frac{N_f - N_b}{N_f + N_b}, \qquad (2)$$

where N_f (N_b) is the number of forward (backward) events. All $t\bar{t}$ asymmetries reported in this paper are given after subtracting the contributions from background processes.

The rapidities of the t and $\bar{t}$ quarks and the corresponding asymmetries can be defined at the production level (sometimes denoted as generator level or parton level), when the kinematic parameters of the generated top quarks are used. Unless stated otherwise the production-level asymmetries are defined for all signal events without imposing the selection criteria of this analysis. The rapidities and asymmetry can also be defined at the reconstruction level, using the reconstructed kinematics of the selected events. Similarly, the invariant mass of the $t\bar{t}$ system can be defined at the production and reconstruction levels.

B. Strategy

In the SM a top quark almost always decays to a b quark and a W boson, which decays either leptonically or hadronically. In this paper we identify $t\bar{t}$ events using the $t\bar{t} \rightarrow W^+ b W^- \bar{b}$; $W^+ \rightarrow l^+ \nu_l$; $W^- \rightarrow q\bar{q}'$ (and charge conjugates) decay chain. This channel is commonly referred to as the "lepton + jets" (l + jets) channel. We select events that contain one isolated lepton (electron or muon) of high transverse momentum (p_T) and at least three jets. The electric charge of the lepton identifies the electric charge of the leptonically decaying W boson and its parent top quark. The other top quark is assumed to have the opposite charge. The event selection, sample composition determination, and modeling of the signal and background processes are identical to those used in the measurement of the leptonic asymmetry in $t\bar{t}$ production in the l + jets channel [10]. The four-vectors of the top quarks and antiquarks in the events containing at least four jets are reconstructed with a kinematic fitting algorithm, while for the events that contain only three jets a partial reconstruction algorithm is used. If a jet exhibits properties consistent with a jet originating from a b quark, such as the presence of a reconstructed secondary vertex, we identify it as a b-tagged jet [11]. The l + jets events are separated into channels defined by jet and b-tag multiplicities. The amount of signal and the forward-backward asymmetry at the reconstruction level are determined using a simultaneous fit to a kinematic discriminant in these channels.

The measured background-subtracted one-dimensional (1D) distribution in Δy is corrected to the production level ("unfolded"). From this distribution we calculate the fully corrected A_{FB} as well as its dependence on Δy. To study the dependence of the asymmetry on the invariant mass of the $t\bar{t}$ system, unfolding is done on the background-subtracted data distributions in two dimensions (2D: Δy vs $m_{t\bar{t}}$). The signal channels are unfolded simultaneously to yield the desired 1D or 2D production-level distributions, from which the production-level A_{FB} values are computed using Eq. (2). The procedure is calibrated using simulated samples with varied asymmetries and input distributions in Δy and $m_{t\bar{t}}$. The statistical and systematic uncertainties of the results are evaluated using ensembles of simulated pseudo-data sets (PDs).

II. D0 DETECTOR

We use the data collected by the D0 detector during Run II of the Tevatron in the years 2001–2011. After imposing event quality requirements ensuring that all detector systems were fully operational, this data set corresponds to an integrated luminosity of 9.7 fb^{-1}. The D0 detector is described in detail elsewhere [12]. The central tracking system, consisting of a silicon microstrip tracker and a scintillating fiber tracker, is enclosed within a 1.9 T superconducting solenoid magnet. Tracks of charged particles are reconstructed within a detector pseudorapidity region[2] of $|\eta| < 2.5$. Electrons, photons, and jets of hadrons are identified [13] using a liquid-argon and uranium-plate calorimeter, which consists of a central barrel covering up to $|\eta| \approx 1.1$, and two end-cap sections that extend coverage to $|\eta| \approx 4.2$ [14]. Central and forward preshower detectors are positioned in front of the corresponding sections of the calorimeter. A muon system consisting of layers of tracking detectors and scintillation counters placed in front of and behind 1.8 T iron toroids [15] identifies muons [16] within $|\eta| < 2$. Luminosity is measured using arrays of plastic scintillators located in front of the end-cap calorimeter cryostats. A three-level trigger system selects interesting events at the rate of ≈ 200 Hz for offline analysis [17].

III. EVENT SELECTION AND MODELING

Object reconstruction and identification, as well as event selection, are the same as in Ref. [10] and are briefly outlined in this section. We select events with exactly one isolated electron within the detector pseudorapidity range of $|\eta| < 1.1$ or one isolated muon within $|\eta| < 2.0$, and at least three jets within $|\eta| < 2.5$. To limit the possible contribution of poorly modeled background due to multijet production, leptons of either flavor are required to have $|y| < 1.5$. The presence of a neutrino is inferred from a transverse momentum imbalance, which is measured primarily using calorimetry and is referred to as the "missing transverse energy" E_T. All selected objects are required to have transverse momentum $p_T > 20$ GeV, and the jet with the largest p_T (the leading jet) is also required to have $p_T > 40$ GeV.

---

[2]The detector pseudorapidity η is defined as $-\ln[\tan(\frac{\theta}{2})]$, where θ is the polar angle measured with respect to the center of the detector. The angle $\theta = 0$ corresponds to the direction of the incoming proton.

To identify jets that are likely to be associated with b quarks, we perform a multivariate analysis (MVA) that combines variables characterizing the properties of secondary vertices and of tracks with large impact parameters relative to the primary $p\bar{p}$ interaction vertex (PV) [11]. The output of the MVA is a continuous variable, MVA_b. The requirement on MVA_b (b tagging) used in this analysis has an efficiency of about 64% for identifying b jets originating from top-quark decay, and a misidentification probability of about 7% for jets that do not contain heavy-flavor quarks and are produced in association with leptonically decaying W bosons.

We simulate $t\bar{t}$ production using the MC@NLO program (version 3.4) [18] with the parton showering performed by HERWIG [19]. This simulation is fully integrated with the D0 software, allowing for detailed studies of the kinematic dependences of A_{FB} and their interplay with selection and reconstruction effects. The main source of background to the $t\bar{t}$ signal is the production of a leptonically decaying W boson in association with jets ($W +$ jets). The kinematic properties of this process are simulated using ALPGEN [20] with hadronic showering performed by PYTHIA [21]. For signal and background modeling we use the CTEQ6.1 set of parton distribution functions (PDFs) [22]. The normalization of the $W +$ jets contribution is a free parameter in the fitting procedure described below. Events with multiple jets can also mimic $t\bar{t}$ signal when a particle from one of the jets is misidentified as an isolated lepton. The normalization of this multijet background is extrapolated from a control sample enriched in this process using the probability for a jet to satisfy the lepton-quality requirements [23]. For the other backgrounds, $Z +$ jets events are simulated with ALPGEN, diboson events are simulated with PYTHIA, and events from single-top-quark production are simulated with COMPHEP [24]. The normalizations for the last three background processes are taken from next-to-leading-order calculations [25]. In all cases, event generation is followed by the GEANT-based D0 detector simulation [26]. To model energy depositions from noise and additional $p\bar{p}$ collisions within the same bunch crossing, simulated events are overlaid with data from random $p\bar{p}$ crossings. All simulated events are reconstructed using the same code as for the reconstruction of the collider data.

IV. RECONSTRUCTION OF THE EVENT KINEMATICS

To measure the $t\bar{t}$ forward-backward asymmetry and its dependence on the invariant mass of the $t\bar{t}$ system we need to determine the four-vectors of the top quark and anti-quark, which is done by summing the four-vectors of their decay products in the $t\bar{t} \rightarrow W^+ b W^- \bar{b}$; $W^+ \rightarrow l^+ \nu_l$; $W^- \rightarrow q\bar{q}'$ (and charge conjugates) decay chain. There are four final-state quarks in this decay chain, while we select events that contain one isolated lepton and at least three jets. When an event contains at least four jets we assume that the four

jets with the largest p_T originate from the quarks from $t\bar{t}$ decay. If an event contains only three jets one of the jets from $t\bar{t}$ decay is missing. In either case all possible assignments of three or four jets to the final-state quarks are used, with the likelihood of each assignment evaluated by the reconstruction algorithms described below.

For $l+ \geq 4$ jet events, the $t\bar{t}$ system is fully reconstructed using a kinematic fitting algorithm. Previous D0 top-quark analyses used the algorithm of Ref. [27]. In this paper a new algorithm is employed, which utilizes an analytic solution for the neutrino momentum using the constraints on the W-boson (M_W) and top-quark masses (m_t) [28]. The likelihood term for each jet-to-quark assignment accounts for the differences between the observed jet energy and the energy scaled to satisfy the constraints on M_W and m_t. The jet energy resolution and the probability for a jet to be reconstructed (see "Type III" transfer function in Ref. [29]) are taken into account. The b-tagging observables MVA_b are also used to evaluate the likelihood of each assignment.

For $l + 3$ jet events, a partial reconstruction algorithm of the $t\bar{t}$ decay chain is employed [30]. With one jet entirely lost, no significant improvement is expected from scaling the four-vectors of the remaining objects as is done by the kinematic fitting algorithm in $l+ \geq 4$ jet events, so the partial reconstruction algorithm does not attempt to modify the kinematics of the observed objects. As only the transverse components of the neutrino momentum are measured in E_T, the longitudinal component is calculated using a quadratic equation which results from imposing the M_W constraint on the $W \rightarrow l\nu$ decay products. The two-fold ambiguity is resolved by choosing the solution that minimizes the difference between the known m_t and the invariant mass of the objects assigned to the leptonic top-quark decay, m_l. This algorithm thus assumes that the jet associated with the b quark from the leptonically decaying top quark is detected. This assumption holds for 80% of the $t\bar{t}$ events. The lost jet is assumed to be associated with either a light quark or a b quark from the hadronic top-quark decay chain. In the majority of cases (74%) this jet is lost due to its low energy, so this loss has little effect on the kinematics of the hadronically decaying top quark. The lost jet is neglected in the partial reconstruction algorithm. The sum of the four-vectors of the two jets assigned to the products of the hadronically decaying top quark serves as a proxy for the four-vector of the hadronically decaying top quark with the invariant mass m_p, shown in Fig. 1.

Even though m_p is not expected to be equal to m_t, the distribution in this variable is different for combinations correctly associated with the hadronically decaying top quark and combinations that include a b jet from the leptonically decaying top quark. In each event we consider the following nine observables: the MVA_b for each of the three jets, the three possible m_l, corresponding to the three possible lepton-neutrino-jet combinations, and the three possible m_p. The likelihood of each of the three possible
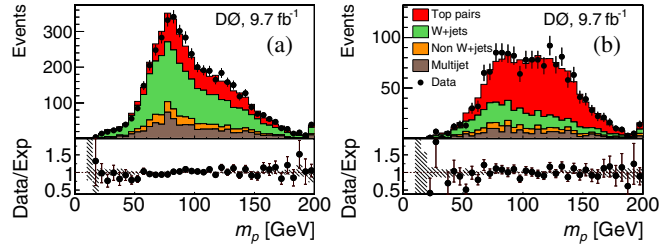
FIG. 1 (color online). The reconstructed m_p for $l + 3$ jet events with (a) one and (b) two or more b tags. Sample composition determination is discussed in Sec. VI. The ratio between the data counts and the model expectation is shown in the lower panel, with the hashed area representing the systematic uncertainties.
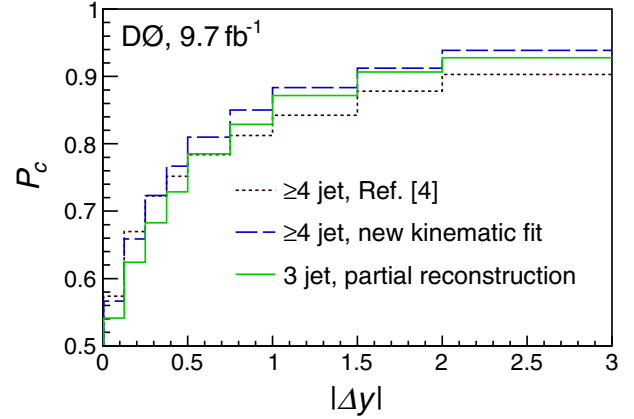

FIG. 2 (color online). The probability to correctly reconstruct the sign of Δy as a function of the production-level $|\Delta y|$ for the algorithm of Ref. [27] used to measure the A_{FB} in Ref. [4] and the algorithms used to reconstruct $l + \geq 4$ jet events and to partially reconstruct $l + 3$ jet events in this paper.

jet-to-quark assignment is calculated by evaluating the consistency of the nine observables with the distributions corresponding to the hypothesized assignment. In particular, the jet hypothesized to be associated with a b quark should have a value of MVA_b consistent with the one expected for b jets, while for a jet hypothesized to originate from a W-boson decay MVA_b should be consistent with the distribution expected for such jets. The values of m_l and m_p for the jet combinations that correspond to the hypothesized assignment should be consistent with the distributions expected for correctly assigned jets, while the values of m_l and m_p for the other jet combinations should agree with the distributions expected for wrong assignments. When calculating $m_{t\bar{t}}$, we compensate for the effect of the lost jet by applying an m_p-dependent scaling to the four-vector of the hadronically decaying top quark.

Unlike the A_{FB} measurement in Ref. [4], where only the jet-to-quark assignment with the lowest χ^2 was used, we reconstruct Δy by averaging its values over all possible assignments, weighted by their likelihoods evaluated as described above for $l + \geq 4$ jet and $l + 3$ jet events. The same approach is used to reconstruct $m_{t\bar{t}}$ in the $l + 3$ jet channel. For $l + \geq 4$ jet events, $m_{t\bar{t}}$ is reconstructed using the outputs of three reconstruction algorithms: the new kinematic fit algorithm, the kinematic fit algorithm of Ref. [27], and a simple reconstruction algorithm [31] that evaluates the kinematics of the leptonically decaying W boson from the lepton and the neutrino by imposing the M_W constraint and calculates $m_{t\bar{t}}$ by adding the four most energetic jets without imposing the m_t constraint. The likelihood values calculated by the algorithms give indications on how well the kinematics of a particular event match the assumptions made by a given algorithm. In particular, for high $m_{t\bar{t}}$ there is a higher probability that two final-state quarks are associated with the same jet. Such a jet is likely to be the most energetic jet in the event and have a large mass. The simple reconstruction algorithm, which does not assume a specific jet-to-quark assignment, performs best for such events. We use a multivariate regression [32] to combine the partially correlated $m_{t\bar{t}}$ values and the

likelihoods produced by the three algorithms with supplementary observables such as the mass of the leading jet to estimate $m_{t\bar{t}}$. This combined $m_{t\bar{t}}$ reconstruction outperforms the individual algorithms in all $m_{t\bar{t}}$ ranges.

For the asymmetry measurement the performance of a $t\bar{t}$ reconstruction algorithm can be characterized by the probability P_c to correctly reconstruct the sign of Δy. For the algorithm employed in this analysis for $l + \geq 4$ jet events $P_c = 0.775$, compared to $P_c = 0.756$ for the algorithm of Ref. [27]. The partial reconstruction algorithm achieves $P_c = 0.745$ for $l + 3$ jet events. The dependence of P_c on the production-level $|\Delta y|$ is shown in Fig. 2 for these three algorithms. The high values of P_c achieved by the partial reconstruction algorithm, which are almost as high as P_c for $l + \geq 4$ jet events, can be understood from the following consideration. All four leading jets are associated with the quarks from the $t\bar{t}$ decay in only 55% of the $l + \geq 4$ jet events. For the other 45% of the events one of the jets originates from initial- or final-state radiation, which can lead to badly misreconstructed $t\bar{t}$ four-vectors. Only 4% of the $l + 3$ jet events contain a jet that does not originate from the four quarks of the $t\bar{t}$ decay. Thus, even though some information is lost with the unreconstructed jet, no wrong information is added, leading to a low probability to misreconstruct the sign of Δy.

V. SM PREDICTIONS FOR A_{FB}

The differential cross section for $t\bar{t}$ production is available only at order α_s^3, where α_s is the strong coupling constant. Since in the SM the $t\bar{t}$ asymmetry only appears at this order, no full higher-order prediction for A_{FB} exists yet. The relative uncertainty on the α_s^3 calculation of the asymmetry due to higher-order corrections is evaluated to be as large as $\approx 25\%$ [33].

Recently the order α_s^4 calculation for the total cross section of $t\bar{t}$ production [34] was made available, but the asymmetry was not computed at this order. Several papers report calculations of the leading corrections to the asymmetry with the predicted $A_{\rm FB}$ values ranging from 5.0% in MC@NLO to 8.8% once the electroweak corrections [35] and resummations of particular regions of phase space [36] are taken into account. The dominant uncertainty on these predictions is from the renormalization and factorization scales, and is evaluated to be as high as 2.0% (absolute) [33,37]. The authors of Ref. [38] obtained a value of $A_{\rm FB} = 12.7\%$ by choosing a normalization scale that arguably stabilizes the perturbative expansion yet differs significantly from the scales commonly used in top-quark physics calculations. Some authors suggest that the corrections from interactions between the top-quark decay products and the proton remnants should also be taken into account when calculating $A_{\rm FB}$ [39]. Given this variety of predictions, we choose to compare our data to the well-defined MC@NLO simulation.

At order α_s^3, the QCD contributions to the asymmetry in $t\bar{t}$ production can be divided into two classes up to divergences that cancel between these two classes [6]. The first class, which contributes to negative asymmetry, is a result of interference between the terms that contain gluon radiation in the initial or final states, which may result in an extra jet in the event and typically leads to a higher transverse momentum of the $t\bar{t}$ system. The second class, which contributes to positive asymmetry, is from interference between the Born term (α_s^2) and the term described by a box diagram (α_s^4). The overall asymmetry is positive and depends on the jet multiplicity. Selection criteria that give preference to events with higher jet multiplicity favor the first class of events and further lower the overall expected asymmetry, while a higher asymmetry is expected for events with lower jet multiplicity. Consequently, forward events tend to have fewer jets than events in the backward category. Similarly, since a b-tagged jet is less likely to originate from initial- or final-state radiation, samples with a larger number of b tags tend to have higher values of $A_{\rm FB}$.

MC@NLO predicts an overall asymmetry in $t\bar{t}$ production before selection of $(5.01 \pm 0.03)\%$. Here, and in the following sections, the quoted uncertainties on the predictions are from the finite size of the simulated samples unless otherwise stated. Table I lists the MC@NLO predictions for $t\bar{t}$ events after the selection criteria are applied.

All previous measurements of $A_{\rm FB}$ in the $l +$ jets channel selected $t\bar{t}$ events that had at least four jets in the final state. As is apparent from Table I, restricting the selection to only $l+ \geq 4$ jet events lowers the production-level asymmetry. Including events with three jets reduces this selection bias.

Asymmetries after reconstruction are presented in the last column of Table I. Finite resolution in Δy resulting from imperfect reconstruction of objects' momenta and incorrect association of jets to $t\bar{t}$ decay products leads to the difference of the asymmetries at the reconstruction and

TABLE I. Asymmetries predicted by MC@NLO for $t\bar{t}$ events that pass the analysis selection criteria. Statistical uncertainties only.

	$A_{\rm FB}$, %	
Channel	Production level	Reconstruction level
≥ 3 jets, ≥ 1 b tags	4.7 ± 0.1	3.9 ± 0.1
3 jets, 1 b tag	6.6 ± 0.2	4.7 ± 0.3
3 jets, ≥ 2 b tags	7.3 ± 0.2	5.6 ± 0.2
≥ 4 jets, 1 b tag	1.4 ± 0.2	1.9 ± 0.2
≥ 4 jets, ≥ 2 b tags	3.2 ± 0.1	3.3 ± 0.2

production levels. The Δy smearing results in some forward events being misreconstructed as backward, and vice versa. As a result the reconstructed asymmetries are reduced. This reduction is most apparent in the $l + 3$ jet events, where the largest asymmetry at the production level is predicted. At the same time, the probability of incorrect association of jets to $t\bar{t}$ decay products is lower in events with fewer jets. Thus, forward $t\bar{t}$ events, which have less initial- or final-state radiation, are less likely to migrate into the backward category than vice versa. This effect results in an upward shift in the reconstructed asymmetry. This shift is most apparent in the $l+ \geq 4$ jet, one-b-tag channel, where the lowest asymmetry at the production level is predicted.

VI. SAMPLE COMPOSITION AND RECONSTRUCTION-LEVEL $A_{\rm FB}$

Reconstructed events are divided into six channels by the number of jets and b tags: $l + 3$ jet and $l+ \geq 4$ jet with 0, 1, and ≥ 2 b tags each. The $l + 3$ jet zero-b-tag channel is used only for the background asymmetry calibration, and not for the $t\bar{t}$ asymmetry measurement. The $l+ \geq 4$ jet zero-b-tag channel is used only for determining the sample composition and the reconstruction-level $A_{\rm FB}$, and is not used for measuring the production-level asymmetry.

Several well-modeled variables that have different distributions for signal and background processes, and that have minimal correlations between each other and with Δy and $m_{t\bar{t}}$, are combined into kinematic discriminants bounded between 0 and 1 [10]. For $l+ \geq 4$ jet events a discriminant D_4 is built from the following input variables:

  (i) χ^2, the test statistic of the likeliest assignment from the kinematic fit.
  (ii) $p_T^{\rm LB}$, the transverse momentum of the leading b-tagged jet, or when no jets are b tagged, the p_T of the leading jet.
  (iii) $k_T^{\rm min} = \min(p_{T,a}, p_{T,b}) \cdot \Delta\mathcal{R}_{ab}$, where $\Delta\mathcal{R}_{ab} = \sqrt{(\eta_a - \eta_b)^2 + (\phi_a - \phi_b)^2}$ is the angular distance[3] between the two closest jets, a and b, and $p_{T,a}$ and $p_{T,b}$ are their transverse momenta.

---

[3]Here the pseudorapidity η and the azimuthal angle ϕ are defined relative to the PV.

(iv) M_{jj}, the invariant mass of the jets corresponding to the $W \to q\bar{q}'$ decay in the likeliest assignment from the kinematic fit, calculated using kinematic quantities before the fit.

The variables χ^2 and M_{jj} are based on the full $t\bar{t}$ reconstruction using the kinematic fitting technique of Ref. [27].

For the $l + 3$ jet events we construct a discriminant D_3 using a different set of input variables:

(i) S, the sphericity [40], defined as $S = \frac{3}{2}(\lambda_2 + \lambda_3)$, where λ_2 and λ_3 are the two largest of the three eigenvalues of the normalized quadratic momentum tensor M. The tensor M is defined as

$$M_{ij} = \frac{\sum_o p_i^o p_j^o}{\sum_o |p^o|^2}, \qquad (3)$$

where p^o is the momentum vector of a reconstructed object o, and i and j run over the three indices for the Cartesian coordinates. The sum over objects includes the three selected jets and the selected charged lepton.

(ii) p_T^{3rd}, the transverse momentum of the third leading jet.

(iii) M_{jj}^{min}, the lowest of the invariant masses of two jets, out of the three possible jet pairings.

(iv) p_T^{LB}, defined as for the $l+ \geq 4$ jet channel above.

(v) $\Delta\phi(\mathrm{jet}_1, E_{\mathrm{T}})$, the difference in azimuthal angle between the leading jet and E_{T}.

The discriminants for all channels are combined into a single discriminant D_c, so that for the $l + 3$ jet events $D_c = N_{\mathrm{tag}} + D_3$, while for $l+ \geq 4$ jet events $D_c = 3 + N_{\mathrm{tag}} + D_4$. The variable N_{tag} above is usually taken to be equal to the number of b-tagged jets in the event, but for events with more than two b-tagged jets $N_{\mathrm{tag}} = 2$ instead. We fit the sum of the signal and background templates to the data distribution in the discriminant D_c as shown in Fig. 3. This fit is identical to the fit for the sample composition in Ref. [10]. The sample composition and its breakdown into individual channels are summarized in Table II. Background contributions other than
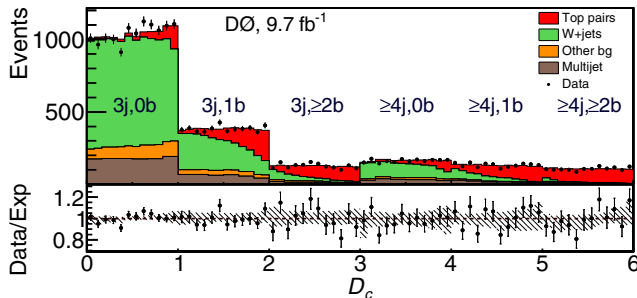

FIG. 3 (color online). The combined discriminant D_c. The region $D_c < 1$ is not used to determine the signal A_{FB}. The ratio between the data counts and the model expectation is shown in the lower panel, with the hashed area representing the systematic uncertainties. Figure is from Ref. [10].

TABLE II. Estimated number of events from the fit of the data distribution in the discriminant D_c to the sum of signal and background processes (see Fig. 3). The sum of the estimated number of signal and background events is constrained to be equal to that in data. The "Selected events" column includes the $l + 3$ jet events with at least one b tag and all $l+ \geq 4$ jet events. The statistical uncertainties from the fit are quoted. We also present the event breakdown for the channels with at least one b tag, which are used to determine the production-level A_{FB}. Table is from Ref. [10].

Source	Selected Events	3 jets		≥ 4 jets	
		1 b tag	≥ 2 b tags	1 b tag	≥ 2 b tags
$W +$ jets	4447 ± 74	2461	352	403	79
Multijet	969 ± 24	449	95	127	62
Other Bg	786	404	112	75	32
Signal	4745 ± 70	1212	1001	983	1166
Sum	10947	4526	1560	1588	1339
Data	10947	4588	1527	1594	1281

$W +$ jets and multijet production are labeled "Other Bg" in Table II.

In the simulated $W +$ jets background, the angular distribution of leptons from W-boson decay has a forward-backward asymmetry, which is in part tuned to Tevatron data [41]. A W boson can be produced in the interaction of a valence quark and antiquark with the jets resulting from initial-state radiation. In this case the asymmetry of the charged leptons is negative due to the vector–axial-vector structure of the weak interaction at the production and the decay of the W boson. A W boson can also be produced in association with a (anti)quark in a (anti)quark-gluon interaction with extra jets resulting from initial- and final-state radiation. In this case, the W boson tends to be boosted in the direction of the initial-state valence (anti)quark, resulting in positive asymmetry of the leptons. The relative fraction of these mechanisms of W-boson production depends on the selection criteria. In particular, quark-antiquark annihilation tends to have higher momentum transfer and as a result the fraction of this process increases from 20% in events with the lepton transverse momentum $20 < p_T^l < 35$ GeV to 30% in events with $p_T^l > 60$ GeV. The associated heavy quarks are usually produced via a splitting of an energetic final-state gluon, which are more common in the case of W-boson production via quark-antiquark annihilation. Thus, the fraction of W bosons produced via quark-antiquark annihilation increases from 25% in the events with zero b tags to 40% in events with two or more b tags.

Due to the asymmetry of leptons from the W-boson decay, even when $W +$ jets events are reconstructed according to the $t\bar{t}$ hypothesis, there remains a residual A_{FB} of $\approx 5\%$ in the Δy distribution. As reported in Ref. [10] the asymmetry observed in the control data sample with

three jets and zero b tags differs somewhat from the expectation. To improve the modeling of this asymmetry we employ a procedure described in Ref. [10], that is, we apply a weight to each simulated $W + \text{jets}$ event. These weights are derived so that the lepton asymmetry as a function of the lepton charge and its rapidity in simulation best matches the one observed in the control sample. These weights are defined in three bins of lepton p_T^l, thus ensuring that the modeling is improved for different admixtures of quark-quark and quark-gluon W-boson production mechanisms. Since the variation in these fractions with p_T^l is similar to the variation with the number of b tags, we expect the modeling to improve in the signal samples. We rely on the simulation to predict the small residual variations of the asymmetry in $W + \text{jets}$ events with jet and b-tag multiplicities. The entire difference in the Δy distributions predicted by the simulation with and without the applied weights is treated as a source of systematic uncertainty due to background modeling. This uncertainty exceeds the uncertainty due to PDFs, which is also taken into account, by about a factor of 2.

The distributions of the reconstructed Δy are shown in Fig. 4. The $t\bar{t}$ asymmetry at the reconstruction level is
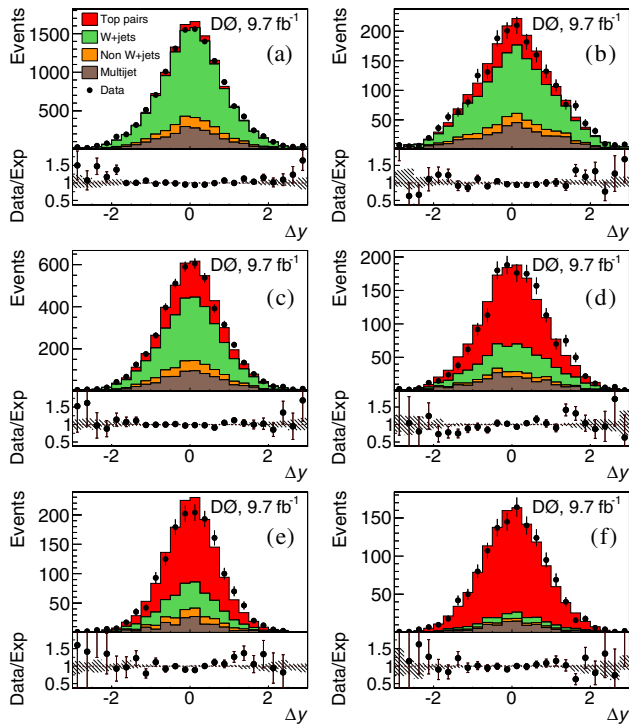
extracted using a fit to the distributions in the discriminant D_c and sign of Δy, excluding the $l + 3$ jet events with zero b tags. This fitting procedure and the control distributions are presented in Ref. [10]. The inclusive asymmetry measured at the reconstruction level is $(7.9 \pm 2.3)\%$. The results for individual channels are listed in Table III.

The distributions of the reconstructed invariant mass of the $t\bar{t}$ system are shown in Fig. 5. Since the $l + 3$ jet and

TABLE III. Reconstruction-level background-subtracted asymmetries for selected events for different channels. The last row includes the channels listed in the rows above and the $l+ \geq 4$ jet, zero-b-tag channel. The first uncertainty is statistical, and the second one is systematic. Systematic uncertainties are discussed in Sec. VIII. The prediction is based on the MC@NLO simulation.

	A_{FB}, %	
Channel	Predicted	Measured
3 jets, 1 b tag	4.7	$5.4 \pm 6.0^{+3.3}_{-4.0}$
3 jets, ≥ 2 b tags	5.6	$10.7 \pm 4.2 \pm 0.8$
≥ 4 jets, 1 b tag	1.9	$11.0 \pm 4.4 \pm 0.8$
≥ 4 jets, ≥ 2 b tags	3.3	$5.9 \pm 3.3 \pm 0.1$
Combined	3.6	$7.9 \pm 2.1 \pm 0.9$

FIG. 4 (color online). Reconstructed difference between the rapidities of the top and antitop quarks, Δy. The left column (a,c, e) shows $l + 3$ jet events, and the right column (b,d,f) shows $l+ \geq 4$ jet events. Rows from top to bottom display events with 0 (a,b), 1 (c,d), and ≥ 2 b (e,f) tags. Overflows are included in the edge bins. The ratio between the data counts and the model expectation is shown in the lower panels, with the hashed area representing the systematic uncertainties.

FIG. 5 (color online). Reconstructed invariant mass of the top quark–antiquark pair, $m_{t\bar{t}}$. The left column (a,c,e) shows $l + 3$ jet events, and the right column (b,d,f) shows $l+ \geq 4$ jet events. Rows from top to bottom display events with 0 (a,b), 1 (c,d), and ≥ 2 b (e,f) tags. The ratio between the data counts and the model expectation is shown in the lower panels, with the hashed area representing the systematic uncertainties.

$l+ \geq 4$ jet channels have different responses (both mean and shape) for $m_{t\bar{t}}$, the dependence of A_{FB} on $m_{t\bar{t}}$ at the reconstruction level is difficult to interpret and is not presented here. The measurement of production-level A_{FB} and its dependence on $m_{t\bar{t}}$ is described in Sec. VII.

We use the results of the sample composition study summarized in Table II to normalize the distributions for the background processes in the sensitive variables (Δy, and also $m_{t\bar{t}}$ for the 2D measurement), which are subtracted from the distributions observed in data. To increase the signal purity of the data used in the fully corrected measurements, we unfold only events containing at least one b tag. The background-subtracted distributions of the sensitive variables in the corresponding four channels are used as inputs to the unfolding procedure.

VII. UNFOLDING THE ASYMMETRY

The true or generated distribution of a certain variable (Δy for the inclusive measurement) is shaped by acceptance and detector resolution, resulting in the observed distribution, which is also subject to statistical fluctuations. The goal of the unfolding procedure is to find the best estimator for the true distribution given the background-subtracted data and knowing the detector acceptance and resolution from simulation. After finding the best estimator for the true distribution of Δy, we summarize it into the production-level A_{FB} using Eq. (2). This is the same general approach used in the previous measurement [4]. For this unfolding we use TUNFOLD [42], which we extend as discussed below.

Each distribution is presented as event counts in a binned histogram,[4] i.e., as a vector with a dimension equal to the number of bins. Given the vector of production-level $t\bar{t}$ signal counts, p, and the vector of expected background counts, b, the expected data counts in the ith bin $\tilde{d}_i$ is given by

$$\tilde{d}_i = \boldsymbol{T}_{ij} p_j + b_i, \qquad (4)$$

$$\boldsymbol{T} = \boldsymbol{MA}, \qquad (5)$$

where $\boldsymbol{A}$ is a diagonal acceptance matrix, whose jjth element is the probability for an event produced in the jth bin to pass the selection criteria and $\boldsymbol{M}$ is the normalized migration matrix, whose ijth element is the probability for a selected event produced in the jth bin to be observed in the ith bin.

Given the vector of observed counts, d, we can construct the vector of background-subtracted reconstruction-level counts, $r = d - b$, with its covariance error matrix $\boldsymbol{V}$. The matrix $\boldsymbol{V}$ is constructed to account for the expected statistical uncertainties on data and background,

---

[4]Overflows are included in the edge bins.

in particular those due to the size of the multijet-enriched control sample. We then seek to find the vector u, which best estimates the vector of production-level counts p, by minimizing

$$\chi^2 = (r - \boldsymbol{T}u)^T \boldsymbol{V}^{-1}(r - \boldsymbol{T}u) + \tau^2 (\boldsymbol{R}u)^T \boldsymbol{R}u \qquad (6)$$

for a given vector r, where τ is the regularization strength and $\boldsymbol{R}$ is the regularization matrix. The first term of Eq. (6) quantifies the consistency of u with data, while the second (regularization) term quantifies the smoothness of u.

Without regularization, the unfolding procedure amounts to a minimization of the first term in Eq. (6). If the numbers of reconstruction and production-level bins are equal, the problem of minimization is solved by simply inverting the matrix: $u_{\text{unregularized}} = \boldsymbol{T}^{-1}r$.

Unregularized matrix inversion typically results in unphysical, rapidly varying distributions [43]. Such distributions are disfavored in regularized unfolding by adding a second "regularization" term to the χ^2. The regularization term in Eq. (6) depends on the discrete second derivative of the binned distribution u. For constant bin widths, the regularization term is calculated using a regularization matrix with the following structure [42]:

$$\boldsymbol{R} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -2 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix}. \qquad (7)$$

For this analysis we modify the structure of $\boldsymbol{R}$ to regularize based on the second derivative of the event density rather than the event counts, which allows for the use of variable bin sizes. The regularization strength τ is chosen using both ensemble testing (described below) and the L-curve technique [42] to balance the minimization of statistical fluctuations and bias. The difference between the two techniques is included in the evaluation of the systematic uncertainty due to the choice of the regularization strength.

As in Ref. [4], the production-level Δy distribution is divided into 26 bins and the reconstruction-level Δy distribution is divided into 50 bins. Both have narrower bins near $\Delta y = 0$, where the probability to misclassify forward events as backward or vice versa changes rapidly, and wider bins at high $|\Delta y|$, where statistics are low.

For the 2D measurement, we use six $m_{t\bar{t}}$ bins at the production level, with edges at 0, 400, 450, 500, 550, 650 GeV and $+\infty$. The joint distribution of Δy and $m_{t\bar{t}}$ has a kinematic boundary at $|\Delta y| = \log([1 + \beta]/[1 - \beta])$, where $\beta = \sqrt{1 - (2m_t/m_{t\bar{t}})^2}$. A bin edge close to this

boundary would result in a large difference in the event density between adjacent bins, a feature that would be smoothed by a regularization procedure, thus biasing u. To avoid such a bias, the Δy edges of the bins of the 2D measurement are chosen to depend on $m_{t\bar{t}}$ as shown in Fig. 6.

The reconstruction-level histograms have similar but finer bins along both the Δy and $m_{t\bar{t}}$ directions. In the $l +$ 3 jet channels 13 $m_{t\bar{t}}$ bins are used to accurately describe migrations among the six production-level bins. The $m_{t\bar{t}}$ resolution in the $l+ \geq 4$ jet channels allows for 14 $m_{t\bar{t}}$ bins.

We simultaneously unfold to the production level the four channels that contain at least one b tag. The difference in purity among channels is accounted for in the definition of the covariance error matrix V. The unfolding technique is calibrated, and the statistical and systematical uncertainties are determined using the results of ensemble tests. Each ensemble comprises simulated PDs that we build according to MC@NLO, ALPGEN [20] or MADGRAPH [44] SM predictions, or according to toy models with different asymmetries. The PDs are created from the expected bin counts $\tilde{d}_i$ calculated using Eq. (4) by adding Poisson (statistical) and Gaussian (systematic) fluctuations, with the Gaussian width taken as one standard deviation for the corresponding systematic uncertainty.

In the toy models the input distribution $P(\Delta y)$ has the form

$$P(\Delta y) = G(\Delta y; \mu, w\sigma_0)(1 + a\,\text{erf}\,(\Delta y/\delta)), \qquad (8)$$

where a and δ are shaping parameters, G is a Gaussian distribution with mean μ and width $w\sigma_0$, σ_0 is the width predicted by MC@NLO, and w is a scaling parameter. The shape of the Δy distribution and the input asymmetry are varied using the parameters μ, a, δ, and w. In addition, we produce ensembles with the signal taken from simulated samples of $t\bar{t}$ production mediated by axigluons, hypothetical massive particles that arise in extensions of the SM that
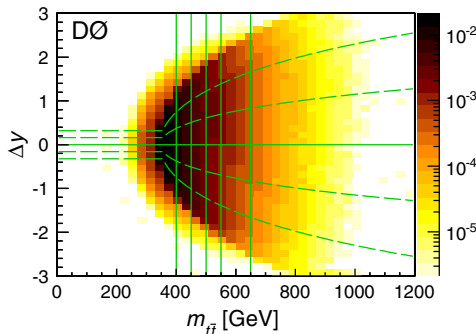

FIG. 6 (color online). Production-level bins for the 2D measurement in the $(m_{t\bar{t}}, \Delta y)$ plane, overlaid on the distribution in these variables predicted from MC@NLO. The shading reflects the predicted event density in arbitrary units. The solid and dashed lines denote the production-level bins. The solid lines show bins that are used for the final result.
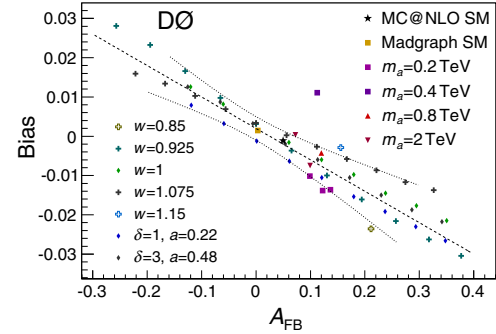

FIG. 7 (color online). The bias as a function of the input A_{FB}. Axigluon scenarios are indicated in the legend by the mass of the axigluon, m_a. The toy models are labeled by parameters w, a, and δ of Eq. (8). Unless stated otherwise $w = 1$, $a = 0$, and $\delta = 1$. For each set of a, δ, and w, the value of μ is varied to produce different input asymmetries. The dashed line indicates the calibration applied to the inclusive measurement and the dotted lines indicate the assigned calibration uncertainties. The point significantly outside of the dotted lines corresponds to an axigluon mass of 0.4 TeV and is discussed in the text.

suggest different strong couplings for left- and right-handed quarks [9]. The input asymmetry in the models used for calibration ranges from -30% to $+30\%$, while the axigluon masses are varied from 0.2 to 2 TeV.

The bias, which is the average difference between the unfolded and input A_{FB} values, is shown in Fig. 7 as a function of the input A_{FB}. Based on this study we derive a correction (calibration) that is applied to the result to eliminate the expected bias. The majority of the tested models are contained within the systematic uncertainty assigned to this calibration, which is shown by the dotted lines in Fig. 7. The one point that is significantly outside of the boundaries of this region corresponds to $t\bar{t}$ production mediated by an axigluon with a mass of 0.4 TeV. This particular model exhibits a significant change in A_{FB} on the $m_{t\bar{t}}$ scale smaller than the bin width (here, 50 GeV), thus breaking the assumption of a smooth underlying distribution, leading to biased results. The unfolding of models with such rapidly changing A_{FB} will, in general, be biased in all regularized unfolding procedures, and we choose not to assign a systematic uncertainty that covers this specific class of models.

The unfolded Δy distribution is presented in eight bins, with each bin calibrated for the expected bias observed in the MC@NLO-simulated PDs. The A_{FB} value in each $|\Delta y|$ range, $A_{\text{FB}}(|\Delta y|)$, is calibrated using the same procedure as for the inclusive A_{FB}. Since no systematic correlation is found between the A_{FB} biases in different $|\Delta y|$, as well as $m_{t\bar{t}}$ bins, they are calibrated individually.

VIII. SYSTEMATIC UNCERTAINTIES

The systematic uncertainties on the reconstruction and production-level A_{FB} are summarized in Table IV in several categories, which are detailed below. To evaluate the

TABLE IV. Systematic uncertainties on $A_{\rm FB}$, in absolute %. For the 2D measurement, the range of changes in $A_{\rm FB}$ over the six $m_{t\bar{t}}$ bins is given.

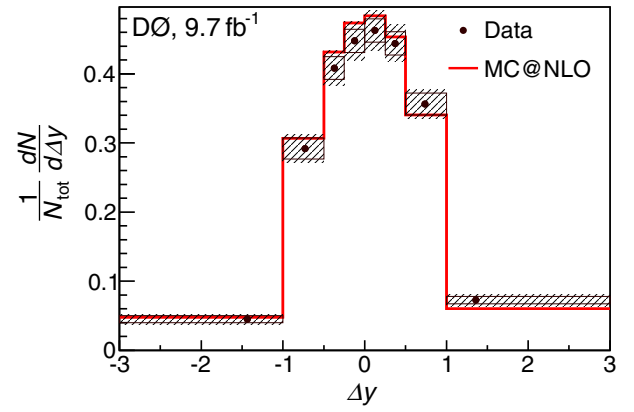| | Reconstruction level | Production level | |
Source	Inclusive	Inclusive	2D
Background model	$+0.7/-0.8$	1.0	1.1–2.8
Signal model	< 0.1	0.5	0.8–5.2
Unfolding	N/A	0.5	0.9–1.9
PDFs and pileup	0.3	0.4	0.5–2.9
Detector model	$+0.1/-0.3$	0.3	0.4–3.3
Sample composition	< 0.1	< 0.1	< 0.1
Total	$+0.8/-0.9$	1.3	2.1–7.5

FIG. 8 (color online).   The production-level Δy distribution. The D0 data points are shown with their statistical uncertainty indicated by the black rectangles and their total uncertainty, based on the diagonal elements of the covariance matrix, indicated by the hashed areas. The histogram shows the MC@NLO prediction [18]. The x coordinate of each data point is the observed average of the Δy distribution in the corresponding bin.

systematic uncertainty on the reconstruction-level $A_{\rm FB}$, we vary the modeling according to the estimated uncertainty in the relevant parameter of the model and propagate the effect to the result. The systematic uncertainties on the production-level $A_{\rm FB}$ are evaluated by including the effects of systematic variations on the simulated background-subtracted PD into the ensemble tests. To find the expected uncertainty due to each category we use dedicated ensembles generated without statistical fluctuations and with only the relevant systematic effects. The total uncertainties on the production-level $A_{\rm FB}$ are taken from ensembles built including both statistical fluctuations and systematic effects (see Sec. VII).

The *background model* category includes the following sources, which affect the properties predicted for background events. The leptonic asymmetry of the $W + $ jets background is varied within its uncertainty of 3% [10]. The rate of heavy-flavor production within $W + $ jets production is varied by $\pm 20\%$ [25,45]. The efficiencies for lepton identification, and the probabilities for a jet to be misidentified as a lepton, taken as functions of lepton momentum, are varied within their uncertainties to account for the uncertainty on the number of background events from multijet production [23]. This variation affects both the background shape and normalization. Uncertainties associated with the modeling of the discriminant, D_c, transverse momentum of the W boson and $M_{jj}^{\rm min}$, as well as potentially increased background levels at high lepton pseudorapidity are also quantified by modifying the background model [10].

The *signal model* category includes the sources of uncertainty that affect the properties predicted for signal events other than the ones accounted for in the PDFs and pileup category. The top-quark mass is varied according to the combined Tevatron measurement of Ref. [46]. The effect of higher-order corrections to $t\bar{t}$ production is estimated by replacing the migration matrix $\mathbf{M}$ from Eq. (4) simulated by MC@NLO with the one simulated by ALPGEN, which uses tree-level matrix elements. The b-quark fragmentation function is varied within its uncertainties [46], which also affects background modeling.

The signal model category also includes the uncertainties associated with gluon radiation. The total amount of

initial-state radiation is varied in a range consistent with the results of Ref. [47]. We also consider the difference in the predicted amount of initial-state radiation between forward and backward events, both because of contributions at order α_s^3 and due to higher-order effects which are modeled by the simulated parton showers [48]. We account for this uncertainty by reducing the difference in the distributions of the p_T of the $t\bar{t}$ system for forward and
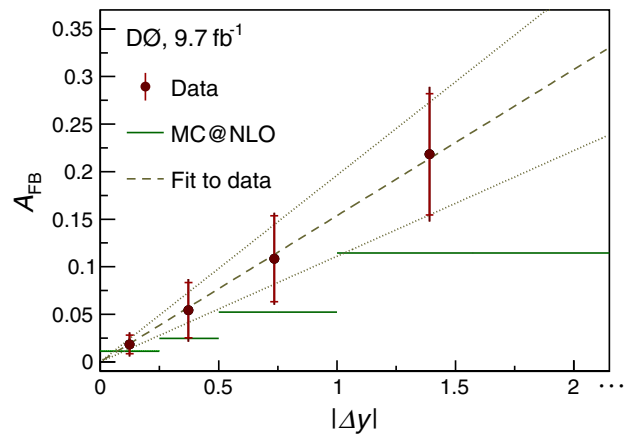

FIG. 9 (color online).   The dependence of the forward-backward asymmetry on $|\Delta y|$. The D0 data points are shown with the total error bars indicating the total uncertainty, based on the diagonal elements of the covariance matrix, while the statistical uncertainties are indicated by the inner error bars. The dashed line shows the fit to the data with the dotted lines indicating the fit uncertainty. The horizontal lines show the MC@NLO prediction for the asymmetry in each $m_{t\bar{t}}$ bin [18]. The last bin has no upper boundary. The x coordinate of each data point is the observed average of $|\Delta y|$ in the corresponding bin.

TABLE V. Variation of the production-level A_{FB} on $|\Delta y|$. The measured values are calibrated and listed with their total uncertainties. The theoretical predictions are based on MC@NLO simulation.

$|\Delta y|$	$A_{\text{FB}},\%$	
	Predicted	Measured
< 0.25	1.1	1.8 ± 1.3
0.25–0.5	2.5	5.4 ± 3.3
0.5–1	5.2	10.8 ± 4.8
> 1	11.4	21.8 ± 7.1

TABLE VI. The correlation factors between the measured A_{FB} values in different $|\Delta y|$ bins.

	$|\Delta y|$ range			
	< 0.25	0.25–0.5	0.5–1	> 1
< 0.25	+1.00	+0.79	+0.77	+0.06
0.25–0.5	+0.79	+1.00	+0.89	+0.09
0.5–1	+0.77	+0.89	+1.00	+0.25
> 1	+0.06	+0.09	+0.25	+1.00

backward events by 25%, a value derived from Ref. [48]. We also account for the possibility that the mismodeling of this variable in the $l + 3$ jet final state affects A_{FB} by reweighting this distribution to match the D0 data, similarly to the procedure used in Ref. [10].
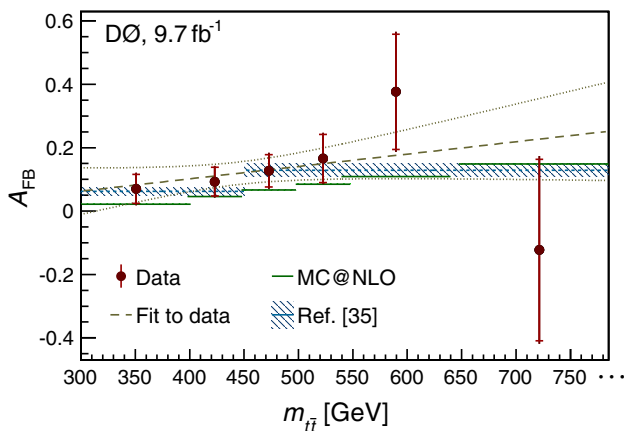

FIG. 10 (color online). The dependence of the forward-backward asymmetry on the invariant mass of the $t\bar{t}$ system. The D0 data points are shown with the total error bars indicating the total uncertainty, based on the diagonal elements of the covariance matrix, while the statistical uncertainties are indicated by the inner error bars. The dashed line shows a linear fit to the data with the dotted curves indicating the fit uncertainty. The horizontal lines correspond to the MC@NLO prediction for the asymmetry in each $m_{t\bar{t}}$ bin [18]. Shaded boxes correspond to the prediction of Refs. [33,35]. The last bin has no upper boundary. The x coordinate of each data point is the simulated average of the $m_{t\bar{t}}$ distribution in the corresponding bin.

TABLE VII. Production-level asymmetries as a function of $m_{t\bar{t}}$. The measured values are calibrated and listed with their total uncertainties. The theoretical predictions are based on MC@NLO simulation.

$m_{t\bar{t}}$, GeV	$A_{\text{FB}},\%$	
	Predicted	Measured
< 400	2.2	7.0 ± 5.1
400–450	4.6	9.3 ± 5.0
450–500	6.7	12.7 ± 5.7
500–550	8.4	16.6 ± 8.2
550–650	10.9	37.6 ± 19.0
> 650	14.8	-12.3 ± 29.6
Inclusive	5.0	10.6 ± 3.0

The uncertainties due to *unfolding* are dominated by the calibration uncertainties. The uncertainties associated with the choice of the regularization strength and statistical fluctuations in the Monte Carlo samples used to find the migration matrix are also included.

The *PDFs and pileup* category includes uncertainties on the modeling of the $p\bar{p}$ collisions. The main uncertainties are from the PDFs, which primarily affect the Δy distribution of the $W + $ jets background. These uncertainties are evaluated by varying the contributions of the various eigenvectors from the CTEQ6.1 PDF [22] and by considering an alternative set of PDFs (MRST2003 [49]). The number of additional collisions within the same bunch crossing (pileup) affects the quality of the reconstruction. The uncertainties on the modeling of additional collisions are also included in this category.

The *detector model* category includes the following sources of systematic uncertainty. The efficiencies of the b-tagging algorithm for jets of different flavors, which are measured from collider data, are varied according to their uncertainties [11]. These variations affect the measured A_{FB} mostly through the estimated sample composition, which depends strongly on the classification of data into several channels according to the number of b tags. The modeling of jet-energy reconstruction, including the overall energy scale and the energy resolution, as well as jet-reconstruction efficiencies and single-particle responses, are all calibrated to collider data and are varied according to

TABLE VIII. The correlation factors between the measured A_{FB} values in different $m_{t\bar{t}}$ bins. All masses are in GeV.

	$m_{t\bar{t}}$ range (GeV)					
	< 400	400–450	450–500	500–550	550–650	> 650
< 400	+1.00	+0.89	+0.39	−0.19	−0.25	+0.12
400–450	+0.89	+1.00	+0.67	+0.10	−0.32	+0.12
450–500	+0.39	+0.67	+1.00	+0.68	−0.27	+0.05
500–550	−0.19	+0.10	+0.68	+1.00	+0.04	−0.12
550–650	−0.25	−0.32	−0.27	+0.04	+1.00	−0.41
> 650	+0.12	+0.12	+0.05	−0.12	−0.41	+1.00

TABLE IX.   The eigenvectors of the covariance matrix Σ and the result of the 2D measurement $\vec{v}$, in the basis of eigenvectors.

i	Eigenvector $\vec{e}_i$						$v_i \pm \sigma_i$
1	$(-0.592$	$+0.770$	-0.237	-0.007	$+0.004$	$-0.000)$	0.000 ± 0.011
2	$(+0.434$	$+0.099$	-0.775	$+0.448$	-0.030	$+0.002)$	0.004 ± 0.021
3	$(+0.673$	$+0.591$	$+0.251$	-0.339	$+0.138$	$-0.004)$	0.130 ± 0.071
4	$(+0.034$	$+0.192$	$+0.516$	$+0.826$	$+0.104$	$+0.049)$	0.256 ± 0.093
5	$(-0.076$	-0.099	-0.113	-0.040	$+0.917$	$+0.360)$	0.265 ± 0.166
6	$(-0.029$	-0.030	-0.019	$+0.031$	$+0.359$	$-0.932)$	0.247 ± 0.311

TABLE X.   Parameters of the fit to Eq. (9). The theoretical predictions are based on the MC@NLO simulation and have negligible statistical uncertainties.

Parameter	Predicted	Measured
Slope, α	3.8×10^{-4}	$(3.9 \pm 4.4) \times 10^{-4}$
Offset, A_0	5.3×10^{-2}	$(11.9 \pm 3.6) \times 10^{-2}$

their uncertainties [50]. The uncertainties due to jet reconstruction and energy measurement are significantly reduced compared to the previous measurement due to the inclusion of the $l + 3$ jet events.

Lastly, the *sample composition* is varied according to its fitted uncertainties. This variation is performed in addition to the changes in the sample composition implicitly induced by other systematic variations.

IX. RESULTS

A. Inclusive A_{FB} and A_{FB} dependence on $|\Delta y|$

The calibrated production-level Δy distribution is shown in Fig. 8. The corresponding inclusive forward-backward asymmetry in $t\bar{t}$ production is $(10.6 \pm 3.0)\%$. The dependence of A_{FB} on $|\Delta y|$ is shown in Fig. 9 and Table V with the correlation factors between bins listed in Table VI. These correlations are taken into account in the fit of the measured $A_{FB}(|\Delta y|)$ to a line. Since for any physical Δy distribution the asymmetry at $\Delta y = 0$ is 0, we constrain the line to the origin and fit for its slope. For data, we find a slope of 0.154 ± 0.043. This slope is compatible within two standard deviations with the MC@NLO-simulated slope of 0.080, which has negligible statistical uncertainty. The difference between the slope reported by the CDF Collaboration [5] and the slope reported in this paper corresponds to 1.3 standard deviations.[5]

B. A_{FB} dependence on $m_{t\bar{t}}$

The dependence of A_{FB} on $m_{t\bar{t}}$ is shown in Fig. 10 and Table VII with the correlation factors between bins listed in Table VIII.

---

[5]When comparing to CDF results, we neglect the correlations of the systematic uncertainties between the two experiments.

The values of the asymmetry measured in six $m_{t\bar{t}}$ ranges constitute a six-dimensional vector $\vec{v}$ with a 6×6 covariance matrix Σ. Table IX lists the eigenvectors $\vec{e}_i$ ($i = 1, ..., 6$) of Σ together with the corresponding components of the vector $\vec{v}$ in the basis formed by the eigenvectors, $v_i = \vec{v} \cdot \vec{e}_i$, and their uncertainties $\sigma_i = \sqrt{\Sigma'_{ii}}$, where Σ' is the covariance matrix transformed to the basis $\vec{e}_i$. The elements of Table IX fully specify the measured six-dimensional likelihood in the Gaussian approximation, and can be used for quantitative comparison with theoretical predictions and other experimental results [51].

Using the full covariance matrix we perform a fit of the measured A_{FB} to the functional form

$$A_{FB}(m_{t\bar{t}}) = \alpha \left(\frac{m_{t\bar{t}}}{\text{GeV}} - C \right) + A_0. \qquad (9)$$

We choose $C = 445$ so that the correlation factor between the fit parameters α and A_0 is less than 0.01 in the fit to the data. The parameters of the fit are listed in Table X for the data and the MC@NLO simulation. We observe a slope α consistent with zero and with the MC@NLO prediction. The difference between the slope reported by the CDF Collaboration [5] and the slope reported in this paper corresponds to 1.8 standard deviations.

X. DISCUSSION

The measured inclusive forward-backward asymmetry in $t\bar{t}$ production, $A_{FB} = (10.6 \pm 3.0)\%$ is in agreement with the SM predictions reviewed in Sec. V, which range from an inclusive asymmetry of 5.0% predicted by the MC@NLO simulation to $(8.8 \pm 0.9)\%$ [35] once electroweak effects are taken into account. The measured dependencies of the asymmetry on $|\Delta y|$ and $m_{t\bar{t}}$ are also in agreement with the SM predictions. Nevertheless, the observed A_{FB} and the dependencies of A_{FB} on $m_{t\bar{t}}$ and $|\Delta y|$ do not disfavor the larger asymmetries that were previously measured in $p\bar{p}$ collisions [5].

To compare the presented result with the previous D0 publication [4], Table XI presents A_{FB} at the reconstruction level measured in different samples. The method discussed in this paper applied to $l+ \geq 4$ jet events from the first 5.4 fb^{-1} of integrated luminosity yields a result consistent

TABLE XI. Reconstruction-level asymmetries measured in different samples with different methods, with their statistical uncertainties.

Sample	Method	Reco-level A_{FB}, %
$l+ \geq 4$ jet, first 5.4 fb^{-1}	From Ref. [4]	9.2 ± 3.7
$l+ \geq 4$ jet, first 5.4 fb^{-1}	This analysis	9.9 ± 3.4
$l+ \geq 3$ jet, first 5.4 fb^{-1}	This analysis	10.1 ± 2.7
$l+ \geq 4$ jet, additional 4.3 fb^{-1}	This analysis	5.3 ± 3.9
$l+ \geq 3$ jet, additional 4.3 fb^{-1}	This analysis	6.0 ± 3.1
$l+ \geq 3$ jet, full 9.7 fb^{-1}	This analysis	7.9 ± 2.1

with that in Ref. [4], but with a reduced uncertainty mainly due to the separation of data into channels based on the number of b tags and the increased efficiency of the new b-tagging algorithm. Once the analysis is extended to include the $l + 3$ jet events collected at that time, the uncertainty is reduced by a factor of 1.26. The result obtained in the second 4.3 fb^{-1} of the Tevatron data set is within one standard deviation from that obtained in the first 5.4 fb^{-1}. The statistical uncertainty obtained in the combined 9.7 fb^{-1} of integrated luminosity is reduced by a factor of 1.29 with respect to the result obtained using the same method in the first 5.4 fb^{-1}, while the reduction expected from scaling with the integrated luminosity is 1.34. This loss of sensitivity is mainly due to higher instantaneous luminosity during the collection of the later data, which required a tighter trigger selection.

The improved reconstruction of Δy and the reduced acceptance bias due to the inclusion of the $l + 3$ jet events result in further reduction of the statistical uncertainty on the unfolded result compared to Ref. [4]. The separation of the data into channels allows us to add the $l + 3$ jet channels without losing the statistical power of the purer $l+ \geq 4$ jet channels.

XI. SUMMARY

In summary, we reported the measurement of the forward-backward asymmetry in $t\bar{t}$ production using the data set recorded by the D0 detector in Run II of the Fermilab Tevatron Collider. The results presented here supersede the ones that were based on about half of the data [4]. The analysis was extended to include events with three jets, allowing for the loss of one jet from the $t\bar{t}$ decay and reducing the acceptance corrections. The unfolding procedure now accounts for the differences in sample compositions between channels, thus maximizing the statistical strength of the individual channels. New reconstruction techniques were used in the $l+ \geq 4$ jet channel, improving the experimental resolution in all variables of interest.

The asymmetry measured at the reconstruction level is $A_{FB} = (7.9 \pm 2.3)\%$. After correcting for detector resolution and acceptance, we obtained a production-level asymmetry $A_{FB} = (10.6 \pm 3.0)\%$. The observed asymmetry and the dependencies of A_{FB} on $m_{t\bar{t}}$ and $|\Delta y|$ are consistent with the standard model predictions.

[1] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **100**, 142002 (2008).

[2] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **101**, 202001 (2008).

[3] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **83**, 112003 (2011).

[4] V. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **84**, 112005 (2011).

[5] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **87**, 092002 (2013).

[6] J. H. Kühn and G. Rodrigo, Phys. Rev. Lett. **81**, 49 (1998).

[7] L. G. Almeida, G. Sterman, and W. Vogelsang, Phys. Rev. D **78**, 014008 (2008); M. T. Bowen, S. D. Ellis, and D. Rainwater, Phys. Rev. D **73**, 014008 (2006); S. Dittmaier, P. Uwer, and S. Weinzierl, Phys. Rev. Lett. **98**, 262002 (2007); K. Melnikov and M. Schulze, Nucl. Phys. **B840**, 129 (2010); J. H. Kühn and G. Rodrigo, J. High Energy Phys. 01 (2012) 063.

[8] For a review, see J. F. Kamenik, J. Shu, and J. Zupan, Eur. Phys. J. C **72**, 2102 (2012); see also Refs. 3–83 of E. L. Berger, Q.-H. Cao, C.-R. Chen, and H. Zhang, Phys. Rev. D **88**, 014033 (2013).

[9] P. H. Frampton and S. L. Glashow, Phys. Lett. B **190**, 157 (1987).

[10] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **90**, 072001 (2014).

[11] V. M. Abazov *et al.* (D0 Collaboration), arXiv:1312.7623. In this paper, jets with MVA > 0.035 are tagged.

[12] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **565**, 463 (2006); R. Angstadt *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **622**, 298 (2010).

[13] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **750**, 78 (2014).

[14] S. Abachi *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **338**, 185 (1994).

[15] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **552**, 372 (2005).

[16] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **737**, 281 (2014).

[17] M. Abolins *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **584**, 75 (2008).

[18] S. Frixione and B. R. Webber, J. High Energy Phys. 06 (2002) 029; S. Frixione, P. Nason, and B. R. Webber, J. High Energy Phys. 08 (2003) 007.

[19] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, J. High Energy Phys. 01 (2001) 010.

[20] M. L. Mangano *et al.*, Report No. CERN-TH-2002-129, 2002 [arXiv:hep-ph/0206293].

[21] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna, and E. Norrbin, Comput. Phys. Commun. **135**, 238 (2001).

[22] D. Stump, J. Huston, J. Pumplin, W.-K. Tung, H.-L. Lai, S. Kuhlmann, and J. F. Owens, J. High Energy Phys. 10 (2003) 046.

[23] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **84**, 012008 (2011).

[24] E. Boos, V. Bunichev, M. Dubinin, L. Dudko, V. Edneral, V. Ilyin, A. Kryukov, V. Savrin, A. Semenov, and A. Sherstnev (CompHEP Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **534**, 250 (2004).

[25] J. M. Campbell and R. K. Ellis, Nucl. Phys. B, Proc. Suppl. **205**, 10 (2010).

[26] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, http://wwwasdoc.web.cern.ch/wwwasdoc/pdfdir/geant.pdf.

[27] S. Snyder, Doctoral thesis, State University of New York at Stony Brook, 1995.

[28] B. A. Betchart, R. Demina, and A. Harel, Nucl. Instrum. Methods Phys. Res., Sect. A **736**, 169 (2014).

[29] I. Volobouev, arXiv:1101.2259.

[30] R. Demina, A. Harel, and D. Orbaker, arXiv:1310.3263.

[31] V. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **85**, 051101 (2012).

[32] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *Proc. Sci.*, ACAT (2007) 040 [arXiv:physics/0703039].

[33] W. Bernreuther (private communication), updating the error estimates of Ref. [35].

[34] M. Czakon, P. Fiedler, and A. Mitov, Phys. Rev. Lett. **110**, 252004 (2013).

[35] W. Bernreuther and Z.-G. Si, Phys. Rev. D **86**, 034026 (2012), and references therein.

[36] e.g., L. G. Almeida, G. F. Sterman, and W. Vogelsang, Phys. Rev. D **78**, 014008 (2008); N. Kidonakis, Phys. Rev. D **84**, 011504 (2011).

[37] J. M. Campbell and R. K. Ellis, arXiv:1204.1513.

[38] S. J. Brodsky and X.-G. Wu, Phys. Rev. D **85**, 114040 (2012); X.-G. Wu, S. J. Brodsky, and M. Mojaza, Prog. Part. Nucl. Phys. **72**, 44 (2013); S. J. Brodsky, M. Mojaza, and X.-G. Wu, Phys. Rev. D **89**, 014027 (2014).

[39] S. J. Brodsky, Annu. Rev. Nucl. Part. Sci. **62**, 1 (2012).

[40] G. Hanson *et al.*, Phys. Rev. Lett. **35**, 1609 (1975).

[41] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **81**, 5754 (1998).

[42] http://root.cern.ch/root/html/TUnfold.html and reference within. In particular, http://www.desy.de/sschmitt and P. C. Hansen, in *Computational Inverse Problems in Electrocardiology*, edited by P. Johnston (WIT Press, Southampton, United Kingdom, 2000), http://www.imm.dtu.dk/pch/TR/Lcurve.ps.

[43] G. Bohm and G. Zech, Introduction to Statistics and Data Analysis for Physicists, http://www-library.desy.de/preparch/books/vstatmp_engl.pdf.

[44] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, J. High Energy Phys. 06 (2011) 128.

[45] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **84**, 012008 (2011).

[46] T. Aaltonen *et al.* (CDF and D0 Collaborations), Phys. Rev. D **86**, 092003 (2012).

[47] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **106**, 122001 (2011).

[48] P. Skands, B. Webber, and J. Winter, J. High Energy Phys. 07 (2012) 151; J. Winter, P. Z. Skands, and B. R. Webber, Eur. Phys. J. C **49**, 17001 (2013).

[49] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. C **35**, 325 (2004).

[50] V. M. Abazov *et al.* (D0 Collaboration), arXiv:1312.6873.

[51] G. Zech, in *Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, doi:10.5170/CERN-2011-006 (2011).