Studies of W boson plus jets production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV

V. M. Abazov,[32] B. Abbott,[67] B. S. Acharya,[26] M. Adams,[46] T. Adams,[44] G. D. Alexeev,[32] G. Alkhazov,[36] A. Alton,[56,*] A. Askew,[44] S. Atkins,[54] K. Augsten,[7] C. Avila,[5] F. Badaud,[10] L. Bagby,[45] B. Baldin,[45] D. V. Bandurin,[44] S. Banerjee,[26] 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,[24] G. Bernardi,[14] R. Bernhard,[19] I. Bertram,[39] M. Besançon,[15] R. Beuselinck,[40] P. C. Bhat,[45] S. Bhatia,[58] V. Bhatnagar,[24] G. Blazey,[47] S. Blessing,[44] K. Bloom,[59] A. Boehnlein,[45] D. Boline,[64] E. E. Boos,[34] G. Borissov,[39] 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,[34] S. Burdin,[39,†] C. P. Buszello,[38] E. Camacho-Pérez,[29] B. C. K. Casey,[45] H. Castilla-Valdez,[29] S. Caughron,[57] S. Chakrabarti,[64] D. Chakraborty,[47] K. M. Chan,[51] A. Chandra,[72] E. Chapon,[15] G. Chen,[53] S. W. Cho,[28] S. Choi,[28] B. Choudhary,[25] S. Cihangir,[45] D. Claes,[59] J. Clutter,[53] M. Cooke,[45] 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,[30,31] E. De La Cruz-Burelo,[29] F. Déliot,[15] R. Demina,[63] D. Denisov,[45] S. P. Denisov,[35] S. Desai,[45] C. Deterre,[20,§] K. DeVaughan,[59] H. T. Diehl,[45] M. Diesburg,[45] P. F. Ding,[41] A. Dominguez,[59] A. Dubey,[25] L. V. Dudko,[34] A. Duperrin,[12] S. Dutt,[24] A. Dyshkant,[47] M. Eads,[47] D. Edmunds,[57] J. Ellison,[43] V. D. Elvira,[45] Y. Enari,[14] H. Evans,[49] V. N. Evdokimov,[35] L. Feng,[47] T. Ferbel,[63] F. Fiedler,[21] F. Filthaut,[30,31] W. Fisher,[57] H. E. Fisk,[45] M. Fortner,[47] H. Fox,[39] S. Fuess,[45] A. Garcia-Bellido,[63] J. A. García-González,[29] G. A. García-Guerra,[29,‡] V. Gavrilov,[33] W. Geng,[12,57] C. E. Gerber,[46] Y. Gershtein,[60] G. Ginther,[45,63] G. Golovanov,[32] P. D. Grannis,[64] S. Greder,[16] H. Greenlee,[45] G. Grenier,[17] Ph. Gris,[10] J.-F. Grivaz,[13] A. Grohsjean,[15,§] S. Grünendahl,[45] M. W. Grünewald,[27] T. Guillemin,[13] G. Gutierrez,[45] P. Gutierrez,[67] J. Haley,[55] 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,[20] I. Heredia-De La Cruz,[29] K. Herner,[56] G. Hesketh,[41,¶] M. D. Hildreth,[51] R. Hirosky,[73] T. Hoang,[44] J. D. Hobbs,[64] B. Hoeneisen,[9] J. Hogan,[72] M. Hohlfeld,[21] 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,[69] M. Jaffré,[13] A. Jayasinghe,[67] M. S. Jeong,[28] 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,[37] E. Kajfasz,[12] D. Karmanov,[34] I. Katsanos,[59] R. Kehoe,[71] S. Kermiche,[12] N. Khalatyan,[45] A. Khanov,[68] A. Kharchilava,[62] Y. N. Kharzheev,[32] I. Kiselevich,[33] J. M. Kohli,[24] A. V. Kozelov,[35] J. Kraus,[58] A. Kumar,[62] A. Kupco,[8] T. Kurča,[17] V. A. Kuzmin,[34] S. Lammers,[49] P. Lebrun,[17] H. S. Lee,[28] S. W. Lee,[52] W. M. Lee,[44] X. Lei,[42] J. Lellouch,[14] D. Li,[14] H. Li,[73] L. Li,[43] Q. Z. Li,[45] J. K. Lim,[28] D. Lincoln,[45] J. Linnemann,[57] V. V. Lipaev,[35] R. Lipton,[45] H. Liu,[71] Y. Liu,[4] A. Lobodenko,[36] M. Lokajicek,[8] R. Lopes de Sa,[64] R. Luna-Garcia,[29,**] A. L. Lyon,[45] A. K. A. Maciel,[1] R. Magaña-Villalba,[29] S. Malik,[59] V. L. Malyshev,[32] J. Mansour,[20] J. Martínez-Ortega,[29] R. McCarthy,[64] C. L. McGivern,[41] M. M. Meijer,[30,31] A. Melnitchouk,[45] D. Menezes,[47] P. G. Mercadante,[3] M. Merkin,[34] A. Meyer,[18] J. Meyer,[20,§§] F. Miconi,[16] N. K. Mondal,[26] M. Mulhearn,[73] E. Nagy,[12] M. Naimuddin,[25] M. Narain,[69] R. Nayyar,[42] H. A. Neal,[56] J. P. Negret,[5] P. Neustroev,[36] H. T. Nguyen,[73] T. Nunnemann,[22] J. Orduna,[72] N. Osman,[12] J. Osta,[51] M. Padilla,[43] A. Pal,[70] N. Parashar,[50] V. Parihar,[69] S. K. Park,[28] R. Partridge,[69,‖] N. Parua,[49] A. Patwa,[65,‖‖] B. Penning,[45] M. Perfilov,[34] Y. Peters,[20] K. Petridis,[41] G. Petrillo,[63] P. Pétroff,[13] M.-A. Pleier,[65] P. L. M. Podesta-Lerma,[29,††] V. M. Podstavkov,[45] A. V. Popov,[35] M. Prewitt,[72] D. Price,[49] N. Prokopenko,[35] J. Qian,[56] A. Quadt,[20] B. Quinn,[58] M. S. Rangel,[1] P. N. Ratoff,[39] I. Razumov,[35] 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] P. Salcido,[47] A. Sánchez-Hernández,[29] M. P. Sanders,[22] A. S. Santos,[1,‡‡] G. Savage,[45] L. Sawyer,[54] T. Scanlon,[40] R. D. Schamberger,[64] Y. Scheglov,[36] 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,[35] R. K. Shivpuri,[25] V. Simak,[7] P. Skubic,[67] P. Slattery,[63] D. Smirnov,[51] K. J. Smith,[62] 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,[35] M. Strauss,[67] L. Suter,[41] P. Svoisky,[67] M. Titov,[15] V. V. Tokmenin,[32] Y.-T. Tsai,[63] D. Tsybychev,[64] B. Tuchming,[15] C. Tully,[61] L. Uvarov,[36] S. Uvarov,[36] S. Uzunyan,[47] R. Van Kooten,[49] W. M. van Leeuwen,[30] N. Varelas,[46] E. W. Varnes,[42] I. A. Vasilyev,[35] A. Y. Verkheev,[32] L. S. Vertogradov,[32] 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] A. White,[70] D. Wicke,[23] M. R. J. Williams,[39] 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,[32] 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]

(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]*Center for Particle Physics, 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]*Fachbereich Physik, Bergische Universität Wuppertal, Wuppertal, Germany*
[24]*Panjab University, Chandigarh, India*
[25]*Delhi University, Delhi, India*
[26]*Tata Institute of Fundamental Research, Mumbai, India*
[27]*University College Dublin, Dublin, Ireland*
[28]*Korea Detector Laboratory, Korea University, Seoul, Korea*
[29]*CINVESTAV, Mexico City, Mexico*
[30]*Nikhef, Science Park, Amsterdam, The Netherlands*
[31]*Radboud University Nijmegen, Nijmegen, The Netherlands*
[32]*Joint Institute for Nuclear Research, Dubna, Russia*
[33]*Institute for Theoretical and Experimental Physics, Moscow, Russia*
[34]*Moscow State University, Moscow, Russia*
[35]*Institute for High Energy Physics, Protvino, Russia*
[36]*Petersburg Nuclear Physics Institute, St. Petersburg, Russia*
[37]*Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Física d'Altes Energies (IFAE), Barcelona, Spain*
[38]*Uppsala University, Uppsala, Sweden*
[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*

(Received 26 February 2013; revised manuscript received 16 September 2013; published 5 November 2013)

We present a comprehensive analysis of inclusive $W(\rightarrow e\nu) + n$-jet ($n \geq 1, 2, 3, 4$) production in proton-antiproton collisions at a center-of-mass energy of 1.96 TeV at the Tevatron collider using a 3.7 fb^{-1} data set collected by the D0 detector. Differential cross sections are presented as a function of the jet rapidities (y), lepton transverse momentum (p_T) and pseudorapidity (η), the scalar sum of the transverse energies of the W boson and all jets (H_T), leading dijet p_T and invariant mass, dijet rapidity separations for a variety of jet pairings for p_T-ordered and angular-ordered jets, dijet opening angle, dijet azimuthal angular separations for p_T-ordered and angular-ordered jets, and W boson transverse momentum. The mean number of jets in an event containing a W boson is measured as a function of H_T, and as a function of the rapidity separations between the two highest-p_T jets and between the most widely separated jets in rapidity. Finally, the probability for third-jet emission in events containing a W boson and at least two jets is studied by measuring the fraction of events in the inclusive $W + 2$-jet sample that contain a third jet over a p_T threshold. The analysis employs a regularized singular value decomposition technique to accurately correct for detector effects and for the presence of backgrounds. The corrected data are compared to particle level next-to-leading-order perturbative QCD predictions, predictions from all-order resummation approaches, and a variety of leading-order and matrix element plus parton shower event generators. Regions of the phase space where there is agreement or disagreement with the data are discussed for the different models tested.

DOI: 10.1103/PhysRevD.88.092001 PACS numbers: 12.38.Qk, 13.85.Qk, 13.87.Ce, 14.70.Fm

I. INTRODUCTION

Measurements of vector boson production in association with jets are important tests of perturbative quantum chromodynamics (pQCD), the theory describing the strong interaction between quarks and gluons. $W + n$-jet production processes are also of interest because of the important role they play as backgrounds to beyond the standard model phenomena and as multiscale QCD processes. In the case of searches, $W + n$-jet production is a major background in several supersymmetry and Higgs boson decay channels. In the case of standard model processes with small cross sections, such as single top quark production and vector boson fusion (VBF) W boson processes, $W + n$-jet processes often overwhelm the small signal. Theoretical uncertainties on the production rates and kinematics of $W + n$-jet processes are large and limit our ability to identify and characterize new phenomena. Therefore, it is important to make $W + n$-jet measurements at the Fermilab Tevatron Collider and the CERN Large Hadron Collider (LHC) in order to constrain these backgrounds. Measurements of $W + n$-jet production have previously been performed by the CDF [1], D0 [2], ATLAS [3,4], and CMS [5] collaborations. We present here new measurements of $W + n$-jet production using a 3.7 fb^{-1} data sample of proton-antiproton collisions collected with the D0 detector [6–8] between 2002 and 2008. The

*Visitor from Augustana College, Sioux Falls, South Dakota, USA.
†Visitor from The University of Liverpool, Liverpool, United Kingdom.
‡Visitor from UPIITA-IPN, Mexico City, Mexico.
§Visitor from DESY, Hamburg, Germany.
‖Visitor from SLAC, Menlo Park, California, USA.
¶Visitor from University College London, London, United Kingdom.
**Visitor from Centro de Investigacion en Computacion-IPN, Mexico City, Mexico.
††Visitor from ECFM, Universidad Autonoma de Sinaloa, Culiacán, Mexico.
‡‡Visitor from Universidade Estadual Paulista, São Paulo, Brazil.
§§Visitor from Karlsruher Institut für Technologie (KIT)-Steinbuch Centre for Computing (SCC).
‖‖Visitor from Office of Science, U.S. Department of Energy, Washington, D.C. 20585, USA.

measurements presented follow from earlier measurements of inclusive $W + n$-jet production cross sections and differential cross sections as a function of the nth-jet p_T up to $n = 4$, using the same data set [2], but providing further details and additional differential distributions. Previous $W + n$-jet measurements have been used in testing and tuning theoretical models of W boson production [9–11].

In this article, we significantly expand on the number of measured observables in order to make a comprehensive study of $W + n$-jet production. These new measurements include differential cross sections of hadronic and leptonic variables, which will provide validation of new theoretical approaches and input for Monte Carlo (MC) tuning. We provide measurements of nth-jet rapidities to test the modeling of parton emission, which is difficult to predict accurately at large values of rapidity. We measure the W boson transverse momentum and the dijet invariant mass in inclusive $W + 2$-jet and $W + 3$-jet events, the latter being an important variable for VBF production of Higgs and electroweak gauge bosons. This observable is a useful validation tool for the reliability of background simulations for certain Higgs boson production and decay channels where the signal is extracted from the dijet mass distribution, and for the investigation of possible new phenomena. All cross section measurements are normalized by the measured inclusive W boson production cross section [2], allowing for the cancellation or reduction of several experimental systematic uncertainties.

In addition to the single differential cross sections, we further probe QCD emissions in $W + n$-jet events through the study of observables such as the mean number of jets in an event as a function of the total hadronic and leptonic transverse energy in the event, H_T, and as a function of the rapidity span between jets in $W+ \geq 2$-jet events. The probability of additional jet emission as a function of dijet rapidity separation ($\Delta y = |y_i - y_j|$, where i and j are the two jets) is also studied for the first time in inclusive $W + 2$-jet events, for both p_T-ordered and rapidity-ordered jets by measuring the fraction of events in the inclusive $W + 2$-jet sample that contain a third jet above a p_T threshold. This variable has consequences for the design of jet vetoes in high jet multiplicity final states, which are particularly important for VBF Higgs and electroweak boson production. Such variables are also sensitive to Balitsky-Fadin-Kuraev-Lipatov (BFKL)-like dynamics [12,13] when the two jets are widely separated in rapidity.

The methods employed for this measurement follow those used in the previous D0 Z + jet cross section [14] and Z boson p_T [15] analyses, as well as in the previous D0 $W + n$-jet analysis [2]. We select a high purity sample of $W + n$-jet events in which the W boson decays to an electron and a neutrino, while maintaining the bulk of the kinematic phase space. The measurements are corrected to the particle level, which includes energy from stable particles, the underlying event (partonic interactions from the same proton-antiproton scatter), muons, and neutrinos, as defined in Ref. [16]. The unfolding uses a regularized singular value decomposition method [17] as implemented in the program GURU. This procedure corrects a measured observable back to the particle observable, deconvolving the effects of finite experimental resolution, detector response, acceptance, and efficiencies.

II. THE D0 DETECTOR

The primary components of the D0 detector are a central tracking system, a calorimeter, a muon identification system, and a luminosity monitor. To reconstruct the W boson and the jets in this measurement, we use the central tracker to identify the location of the $p\bar{p}$ interaction vertex and the electron produced in the decay of the W boson candidate, and use the liquid-argon and uranium calorimeter to identify electromagnetic and hadronic showers, as well as calculate the magnitude and direction of the missing transverse energy. The luminosity monitor is employed to make a measurement of the integrated luminosity corresponding to the data collected, and the trigger system is used to make a basic selection of likely $W + n$-jet events. A detailed description of the D0 detector can be found in Refs. [6–8]. Here we outline the most important elements of the D0 detector for performing the $W + n$-jet measurements presented in this paper.

A. Tracking detectors

The D0 central tracking system is made up of a silicon microstrip tracker and a fiber tracker. The tracking detectors are primarily used to identify the charged track associated with the leptonic decay of the W boson, but are not used directly in jet reconstruction since the jet-finding algorithms in D0 use only energy deposits in the calorimeter towers.

The tracking detectors are used to reconstruct the position of the primary vertex (PV) of the $p\bar{p}$ interaction. The distribution of the PV along the beam axis follows a 20 cm wide Gaussian distribution function centered on the nominal interaction point at the center of the detector. We use a right-handed coordinate system in which the z axis is along the proton beam direction, the x axis points away from the center of the Tevatron ring, and the y axis is upward. The inner tracking system, consisting of the silicon microstrip tracker, provides 35 μm vertex resolution along the beam line and 15 μm vertex resolution in the r-ϕ plane, where ϕ is the azimuthal angle, for tracks with a minimum p_T of 10 GeV at $|\eta| = 0$, where $\eta = -\ln\tan(\theta/2)$ and θ is the polar angle defined with respect to the beam line. The outer tracking system, consisting of the central fiber tracker, includes eight axial and eight stereo doublet layers of 800 μm diameter scintillating fibers to complement the silicon tracker. Both detectors are located inside the 1.9 T magnetic field of the superconducting solenoidal

magnet to allow measurements of the momentum of charged particles.

B. Calorimeter

The calorimeter system consists of a uranium/liquid argon calorimeter, divided into a central (CC) and two end (EC) sections, and a plastic scintillator intercryostat detector. Both the CC and EC are segmented longitudinally into electromagnetic (EM), fine hadronic, and coarse hadronic sections.

The calorimeter is transversely segmented into cells along the polar and azimuthal axes in a projective tower geometry. The CC covers detector pseudorapidity $|\eta| < 1.2$ and the two ECs extend the range up to $|\eta| = 4.2$. Both the electromagnetic and fine hadronic calorimeters are sampling calorimeters with an active medium of liquid argon and absorber plates of nearly pure depleted uranium. Incoming particles traversing the uranium absorber plates initiate showers of secondary particles that ionize the argon in the gaps between the absorber plates. Due to a high-voltage electric field, the free electrons collect on resistively coated copper pads that act as signal boards. The outer part of the calorimeter, the coarse hadronic section, uses copper in the CC and stainless steel in the EC for the absorber plates. The calorimeter is transversely segmented into cells in $\Delta\eta \times \Delta\phi$ of 0.1×0.1 (0.05×0.05 in the third layer of the EM calorimeter for $|\eta| < 3.2$ to allow for a precise location of EM shower centroids). At $|\eta| > 3.2$, the cell size grows to 0.2 or more in both η and ϕ. The energy resolution of jets reconstructed beyond $|\eta| > 3.2$ is therefore degraded. The total depth of the EM calorimeter is about 20 radiation lengths, and the combined thickness of the electromagnetic and hadronic calorimeters is about seven nuclear interaction lengths.

C. Trigger system

The $W + n$-jet events are selected by triggering on single electron or electron-plus-jet signatures with a three-level trigger system. The electron trigger signature is similar to the electron reconstruction signature (including electromagnetic shower shape requirements and a track matched to the EM shower), albeit more loosely defined to facilitate a fast enough trigger decision. Several single electron triggers are used in a logical OR maximize the trigger efficiency. The p_T threshold on the electron triggers varies between 15 and 35 GeV as different triggers are activated at different instantaneous luminosities.

To further maximize the trigger efficiency, data collected with electron + jet triggers are included. In the bulk of the data set considered, the D0 calorimeter trigger performs clustering of the trigger towers, which are $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ sums of the calorimeter cells, using a sliding windows algorithm [7]. This clustering improves the energy resolution of the trigger jet objects, which allows triggering on relatively low-p_T jets. As for the electron triggers, the p_T threshold defined for the jet trigger objects varies between 15 and 25 GeV, depending on the instantaneous luminosity of the data delivered by the Tevatron.

D. Luminosity detector

The measurement of the D0 luminosity is made by the luminosity monitor (LM). The LM consists of scintillating tiles on either side of the interaction point, which measure the particles created in inelastic collisions. The instantaneous luminosity is determined as

$$\mathcal{L} = \frac{f \cdot N_{\mathrm{LM}}}{\sigma_{\mathrm{LM}}}, \tag{1}$$

where f is the $p\bar{p}$ bunch crossing frequency, N_{LM} is the average number of observed interactions, and σ_{LM} is the effective cross section for inelastic collisions measured by the LM that takes into account event losses due to inefficiencies and geometric acceptance [18]. The uncertainty on the luminosity determination is estimated to be 6.1%. The uncertainty is dominated by a 4.2% uncertainty coming from the determination of σ_{LM} [19].

III. EVENT RECONSTRUCTION AND SELECTION

Our measurements use a sample of $W(\rightarrow e\nu) + n$-jet candidate events corresponding to 3.7 fb^{-1} of data collected with the D0 detector in Run II of the Fermilab Tevatron Collider. The analysis techniques used are identical to those described in Ref. [2], although we quote an integrated luminosity using a convention which now includes the loss due to data quality corrections. The data are grouped into two time periods. Run IIa refers to the data collected prior to 2006 (1.1 fb^{-1} of data), when two major upgrades were installed in the D0 detector. An additional layer of silicon was added to the inner tracker to improve track position resolution [8], and the Level 1 calorimeter trigger was replaced with a system that performed electron, jet, and tau identification [7]. Because of these changes, the data collected after the summer of 2006 are triggered and reconstructed in a different manner and are referred to as Run IIb data (a total integrated luminosity of 2.6 fb^{-1}). The measurements presented here are limited by systematic uncertainties and the inclusion of additional data would neither improve the overall precision of the measurements, nor appreciably enhance their kinematic reach.

The events are processed through the D0 reconstruction program, which identifies jet and W boson candidates. Jets are identified with the D0 midpoint cone algorithm [20], with a split-merge fraction of 0.5 using a cone of radius $R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2} = 0.5$ to cluster calorimeter energy into (massive) jets. Jet energies are corrected for calorimeter response, instrumental and out-of-cone showering effects, additional energy deposits in the calorimeter that arise from detector noise and underlying event energy,

and for pileup arising from multiple $p\bar{p}$ interactions and previous bunch crossings. These jet energy scale corrections [21] are determined using transverse momentum imbalance in γ + jet events, where the electromagnetic response is calibrated using $Z/\gamma^* \to e^+e^-$ events. Jets are required to have at least two tracks that point to their associated PV (jet-vertex confirmation) to improve the identification of the jets and ensure they are associated with the same proton-antiproton collision as the W boson under consideration. These tracks must have $p_T > 0.5$ GeV, at least one hit in the SMT detector, and a distance of closest approach with respect to the PV of less than 0.5 cm in the transverse plane and less than 1 cm along the beam axis. Jets are ordered in decreasing transverse momentum, and we denote the jet with the highest transverse momentum as the ''leading'' jet.

Electrons are identified as clusters of calorimeter cells in which at least 95% of the energy in the shower is deposited in the EM section. Electron candidates must have a calorimeter cluster isolated from other calorimeter energy deposits, have spatial distributions consistent with those expected for EM showers, and contain an isolated reconstructed track pointing to the PV and matched to a calorimeter cluster. The transverse momentum (measured by the calorimeter) in an isolation cone around the electron track must not exceed 15% of the electron cluster p_T. The extrapolated electron track must lie within 1 cm of the primary vertex along the z direction (z_{vtx}). Events with a second isolated electron are removed to suppress the background from Z boson and Drell-Yan production. The missing transverse energy in the event is calculated as the vector sum of all the electromagnetic and fine hadronic cell energies, and the coarse hadronic cell energies that are contained in jets provided they have an energy greater than four standard deviations of the electronic noise or are neighbors of an energetic cell [22]. A correction for the presence of any muons is applied to the missing transverse energy calculation. All energy corrections which are applied to electrons and jets in the event are also propagated to the missing transverse energy calculation. Because the longitudinal component of the momentum of the neutrino is not measured, the calculated properties of each W boson candidate are limited to its transverse energy, E_T^W, and transverse mass, defined as $M_T^W = \sqrt{(\not{p}_T + p_T^e)^2 - (\not{p}_T + p_x^e)^2 - (\not{p}_T + p_y^e)^2}$, where $\not{p}_T$ is the magnitude of the missing transverse energy vector, p_T^e is the transverse momentum of the electron, and the remaining variables are the associated x and y components.

The following event selections are used to suppress background while maintaining high efficiency for events in which a W boson was produced: $p_T^e \geq 15$ GeV and electron pseudorapidity $|\eta^e| < 1.1$, $\not{p}_T > 20$ GeV, $M_T^W \geq 40$ GeV, for all jets $p_T^{\text{jet}} \geq 20$ GeV and rapidity $|y^{\text{jet}}| < 3.2$, $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} > 0.5$ between the electron and the nearest jet, and $|z_{\text{vtx}}| < 60$ cm [23]. Events must have a reconstructed $p\bar{p}$ interaction vertex with at least three associated tracks. Finally, events with excessive calorimeter noise are removed from the sample.

In this paper, we refer to the samples containing a W boson and at least n jets as inclusive $W + n$-jet events, or simply as $W + n$-jet events. The inclusive W boson sample contains 2184821 events, and there are 265713, 39805, 5962, and 1028 events in the inclusive $W + 1$-jet, $W + 2$-jet, $W + 3$-jet, and $W + 4$-jet samples, respectively. The majority of these events are true $W(\to e\nu)$ + jets events, but there are background processes contaminating this data set. These background processes include $W(\to \tau\nu)$ + jets events, QCD multijet events in which a jet is reconstructed as an electron, $Z \to e^+e^-$ events in which one electron is not reconstructed, and diboson and top quark processes.

A. QCD multijet background

In QCD multijet events there is a small but non-negligible probability of instrumental backgrounds or decays to electrons/photons in or near jets that may create a ''fake-electron'' signature. A jet composed primarily of neutral particles with a high electromagnetic fraction may pass electron candidate identification criteria, or a photon might be misidentified as an electron. Since the QCD multijet cross section is large, the contribution from fake-electron events to the measured distributions must be taken into account. These backgrounds are difficult to accurately model in simulation; therefore, we estimate our background using a data-driven approach with a D0 data set that is orthogonal to that used for the main measurement.

To estimate this background contribution, we first define two samples of events, one with ''loose'' selection criteria and one with ''tight'' selection criteria corresponding to our standard event selection, where the latter is a subset of the former. The loose sample (containing N_{loose} events) consists of $N_{\text{loose}}^{\text{signal}}$ events with an electron originating from W/Z + jets, diboson, or top quark production sources, and $N_{\text{loose}}^{\text{MJ}}$ multijet background events with a fake-electron signature. In the case of our tight data selection (containing N_{tight} events), we have a similar relation:

$$N_{\text{tight}} = N_{\text{tight}}^{\text{signal}} + N_{\text{tight}}^{\text{MJ}}. \tag{2}$$

To determine the shape and overall normalization of the QCD multijet distributions, we then define a ''loose-not-tight'' (LNT) data sample that is orthogonal to our standard selection (containing N_{LNT} events), requiring that an electron candidate pass the loose selection but fail the tight.

This LNT sample is composed of events with a real electron ($N_{\text{LNT}}^{\text{signal}}$) and events where a jet is misidentified as an electron ($N_{\text{LNT}}^{\text{MJ}}$):

$$N_{\text{LNT}} = N_{\text{LNT}}^{\text{signal}} + N_{\text{LNT}}^{\text{MJ}}. \tag{3}$$

By combining Eqs. (2) and (3), the number of events in our loose selection can be written as

$$N_{\mathrm{LNT}} + N_{\mathrm{tight}} = N_{\mathrm{loose}} = N_{\mathrm{loose}}^{\mathrm{signal}} + N_{\mathrm{loose}}^{\mathrm{MJ}}, \quad (4)$$

with the following relationship between the numbers of real electrons or misidentified jets passing the loose selection and those passing the tight selection:

$$N_{\mathrm{tight}}^{\mathrm{signal}} = \epsilon_{\mathrm{real}} \cdot N_{\mathrm{loose}}^{\mathrm{signal}} \quad (5)$$

$$N_{\mathrm{tight}}^{\mathrm{MJ}} = \epsilon_{\mathrm{MJ}} \cdot N_{\mathrm{loose}}^{\mathrm{MJ}}, \quad (6)$$

where ϵ_{real} and ϵ_{MJ} then represent the respective efficiencies for a real electron and for a misidentified jet passing the loose selection to also pass the tight selection.

From these relations, the number of multijet events with tight electron requirements in a given bin can be determined as follows:

$$N_{\mathrm{tight}}^{\mathrm{MJ}} = \left(\frac{\epsilon_{\mathrm{MJ}}}{1-\epsilon_{\mathrm{MJ}}}\right) \cdot N_{\mathrm{LNT}} - \left(\frac{\epsilon_{\mathrm{MJ}}}{1-\epsilon_{\mathrm{MJ}}}\right) \cdot \left(\frac{1-\epsilon_{\mathrm{real}}}{\epsilon_{\mathrm{real}}}\right) \cdot N_{\mathrm{tight}}^{\mathrm{signal}}. \quad (7)$$

That is, we estimate the shape of the multijet component from the loose-not-tight sample, with overall normalization determined from the relative efficiency for a misidentified jet passing the loose selection to also pass the tight selection, ϵ_{MJ}, and a small correction derived from the tight sample to account for the presence of real electrons in the loose-not-tight sample. To determine the QCD multijet component we first need to calculate the values of ϵ_{real} and ϵ_{MJ}, as well as $N_{\mathrm{tight}}^{\mathrm{signal}}$.

The probability that a true electron candidate in the loose sample also passes the tight criteria, ϵ_{real}, is calculated from the ratio of electron identification efficiencies (calculated using $Z \rightarrow e^+e^-$ events) derived with the loose and tight selection criteria. Significant variation in ϵ_{real} is observed as a function of electron p_T (shown in Fig. 1) with a plateau of $(88.6 \pm 0.3)\%$ reached at a p_T of approximately 45 GeV, and thus events in the tight and loose-not-tight samples are assigned a value of ϵ_{real} based on the electron candidate p_T.

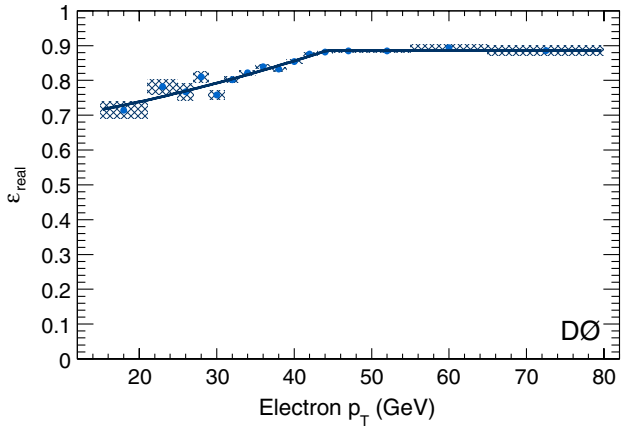


FIG. 1 (color online). Probability that a real electron candidate passing the loose electron identification requirements also passes the tight electron identification requirements. The shaded band represents the systematic uncertainty originating from the determination of the relative tight and loose electron efficiencies.

To determine ϵ_{MJ}, we define a multijet-enriched data sample with selection criteria as for the standard event selection except that the loose electron identification criteria are used, the W boson transverse mass requirement is removed, and the $\not{p}_T > 20$ GeV requirement is reversed and lowered to $\not{p}_T < 10$ GeV. This selection is then applied to the data and to the MC signal and background samples (see Sec. III B). ϵ_{MJ} is determined from the data sample by calculating the fraction of events in this loose multijet-enriched sample that are also in the multijet-enriched selection once tight requirements are applied.

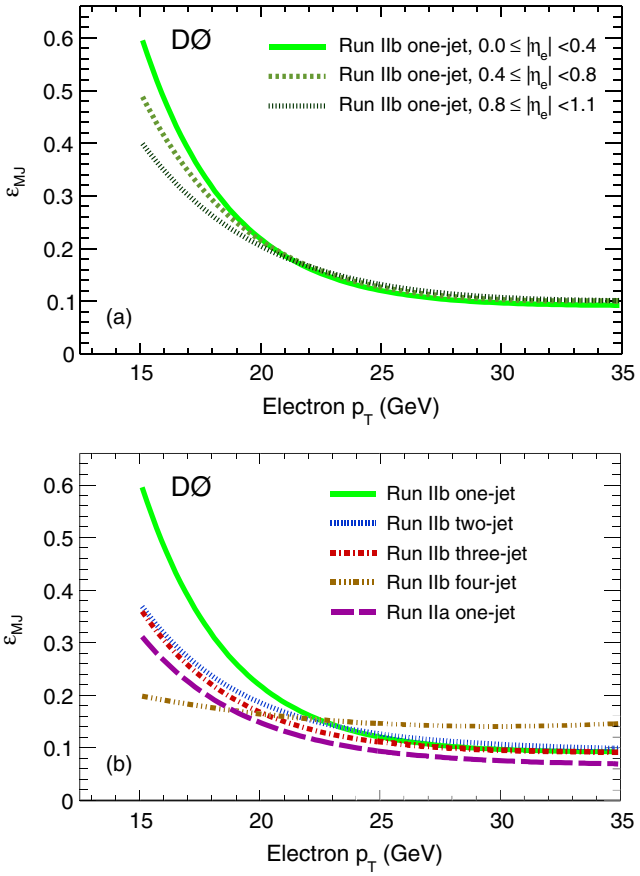


FIG. 2 (color online). Parametrized ϵ_{MJ} [as defined in Eq. (6)], used for the determination of the multijet component of reconstructed data distributions. ϵ_{MJ} is parametrized as a function of electron p_T, electron pseudorapidity, inclusive jet multiplicity, and determined separately for Run IIa and Run IIb data. Figure 2(a) shows the variation of ϵ_{MJ} as a function of electron p_T for three electron pseudorapidity intervals in Run IIb. Figure 2(b) shows the variation of ϵ_{MJ} with jet multiplicity for Run IIa and Run IIb data. Similar variations with respect to electron pseudorapidity and jet multiplicity are also observed for Run IIa.

A small component of true electrons can still contaminate these multijet-enriched samples for both the loose and tight selections. The MC samples with these selection criteria are used to model and remove this signal contamination before the determination of ϵ_{MJ}.

Dependencies of ϵ_{MJ} on jet multiplicity, electron transverse momentum, and, to a lesser extent, electron pseudorapidity are observed (see Fig. 2) and taken into account. Absolute uncertainties on ϵ_{MJ} are approximately 2%–3% and are dominated by statistical uncertainties on the multijet-enriched data sample. These uncertainties on ϵ_{MJ} result in uncertainties of the order of 1%–10% on the multijet contribution to the nominal data sample, increasing with jet multiplicity. We find that the value of ϵ_{MJ}, particularly at low misidentified electron p_T, increases in the Run IIb data-taking period compared to Run IIa, which can be attributed to tighter electron shower-shape requirements applied in trigger selections in Run IIb that reduce the number of loose events.

B. Background and signal process simulation

$W + n$-jet events dominate the inclusive data sample, but there are backgrounds from $Z +$ jets, $t\bar{t}$, diboson, single top quark, and multijet events. With the exception of multijet production, all processes are simulated using MC event generators. All simulated samples are processed through the full GEANT3-based [24] simulation of the D0 detector. Data events from random bunch crossings are overlaid on the simulated events to mimic the effects of detector noise and the presence of additional concurrent $p\bar{p}$ interactions. The simulated events are weighted such that the instantaneous luminosity profile in the simulation matches the distribution observed in data. These events are then reconstructed using the same software that is used on data. The impact of the trigger efficiency dependence on jet kinematics is simulated by the application of multidimensional trigger efficiencies as measured in data. Independent electron and jet samples are used to measure electron and jet trigger object efficiencies using tag-and-probe techniques. The overall trigger efficiency for the logical OR of electron and electron + jet triggers is then calculated, taking into account all correlations, and is applied to the MC as an event weight.

We simulate the $W/Z +$ jets and $t\bar{t}$ processes with ALPGEN v2.11 [25] interfaced with PYTHIA v6.403 [26] for the simulation of additional initial and final-state radiation and for jet hadronization, with the underlying event parameter settings tuned using "Tune-A" [27]. A factorization and renormalization scale choice of $Q^2 = M_V^2 + \sum p_{Tj}^2$ is used for vector boson plus jets processes (where M_V is the vector boson mass, and p_{Tj} is the transverse momentum of a jet in the event). The normalization of $t\bar{t}$ backgrounds is determined from next-to-next-to-leading order (NNLO) calculations [28]. The PYTHIA generator is used to simulate diboson production, with next-to-leading order (NLO) cross sections [29] derived from the MCFM program [30], while production of single top quarks is simulated using the COMPHEP-based NLO event generator SINGLETOP [31] interfaced to PYTHIA for simulation of parton showering and hadronization effects.

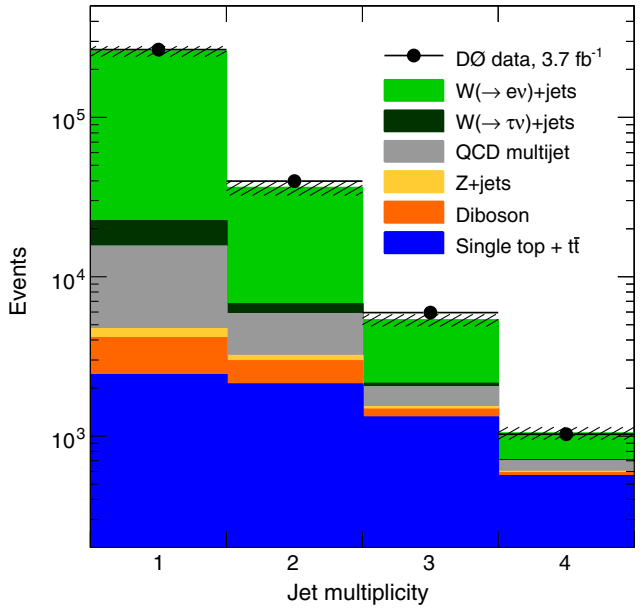


FIG. 3 (color online). Reconstructed inclusive jet multiplicity distributions in the $W(\rightarrow e\nu) +$ jet event selection. Hatched regions indicate uncertainties on the sum of the predicted contributions due to all systematic effects. All signal and background sources are derived from MC simulations with the exception of the QCD multijet component which is estimated from data.

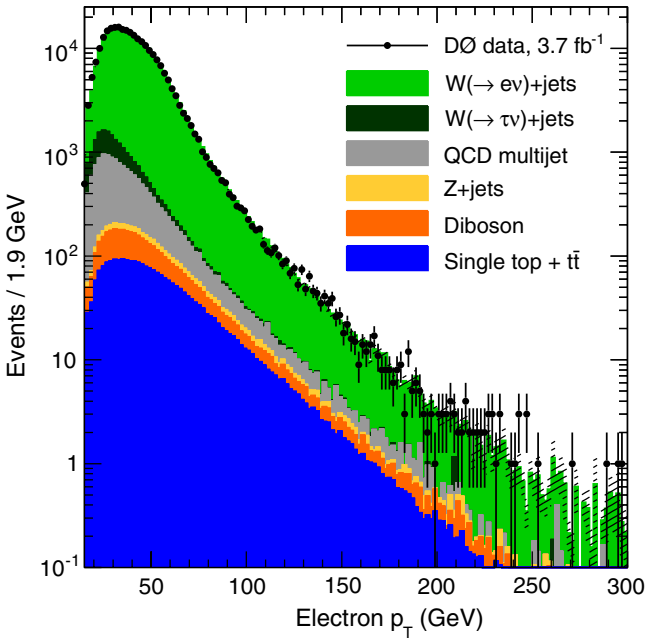


FIG. 4 (color online). Reconstructed electron p_T distribution for events with a W boson candidate and one or more jets. Hatched regions indicate uncertainties on the sum of the predicted contributions due to all systematic effects.

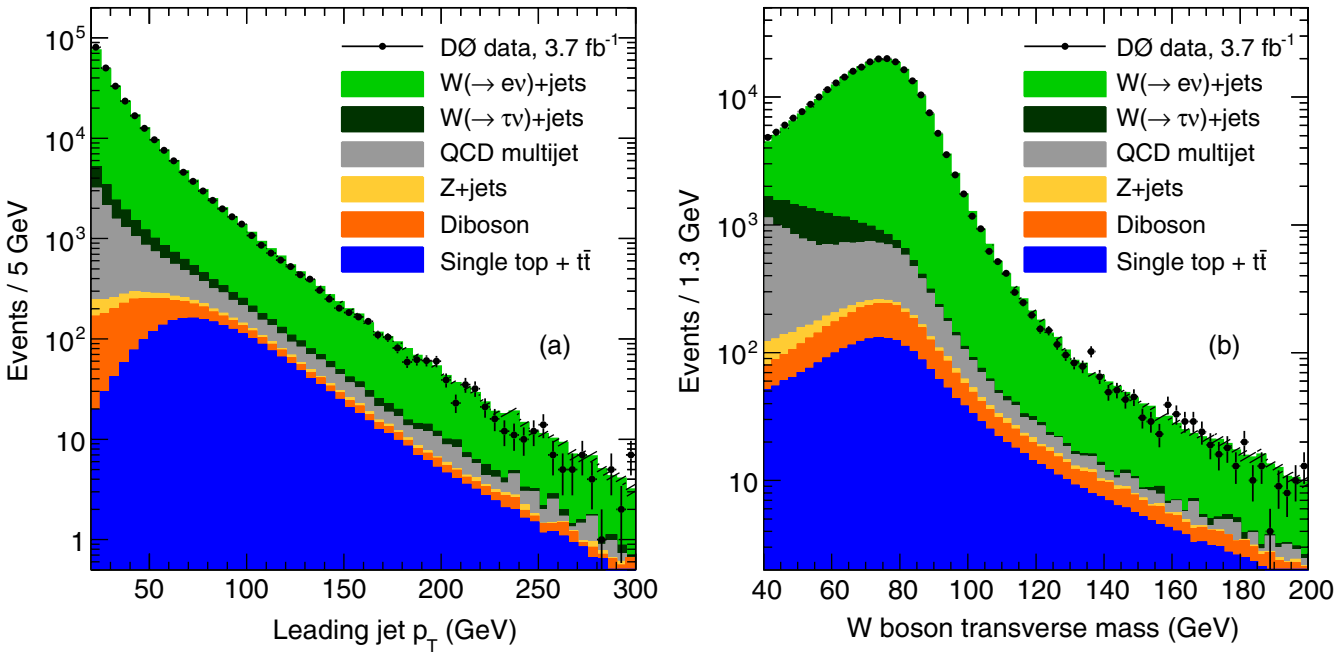


FIG. 5 (color online). Reconstructed kinematic distributions of (a) leading jet p_T and (b) W boson transverse mass for events with a W boson candidate and one or more jets. Hatched regions indicate uncertainties on the sum of the predicted contributions due to all systematic effects.

The W/Z + jets normalization is corrected by a multiplicative factor to match the inclusive W/Z + jets cross sections calculated at NLO [32]. Kinematic distributions are weighted to match existing Z boson transverse momentum measurements in inclusive Z boson events [33], with corresponding corrections for W boson events derived through the application of W to Z p_T distribution ratios from NNLO predictions as calculated by FEWZ [34]. The heavy-flavor fractions are further corrected by the ratio of heavy-to-light NLO multiplicative factors as discussed in Ref. [35], determined from NLO pQCD calculations from MCFM.

The proportion of the data that is attributed to each of these background processes and to the signal process can be seen in Fig. 3.

Figures 4 and 5 illustrate kinematic distributions for selected data events and MC simulations plus data-driven multijet background contributions for some representative observables. The estimated fraction of the data sample that is due to background processes ranges from 2%–40% as a function of the measured observables and the fraction of background due to top quark production ranges from 0%–20%, with the larger contributions at higher jet multiplicities in both cases.

IV. CORRECTION OF EXPERIMENTAL DATA FOR DETECTOR EFFECTS

The background-subtracted yields of $W + n$-jet signal candidates are corrected back to the particle level taking into account corrections for detector acceptance and efficiencies, as well as detector resolution effects. These corrections are performed using a singular value decomposition regularized unfolding approach as implemented in GURU [17]. We define the kinematic phase space into which we unfold our final results by the selection in Table I (the same selection criteria are applied at the reconstruction level).

Electron candidates are defined at the particle level to have the electron four-momentum modified to include all collinear radiation within a cone of radius $R = 0.2$ to account for final-state radiation. At the particle level we define $\not{p}_T$ as the magnitude of the neutrino transverse momenta. Particle level jets are constructed using the D0 Run II midpoint cone algorithm running at particle level. The W boson decay products (including collinear emissions from the electron) are removed from the list of stable particles before constructing jets with a cone radius $R = 0.5$.

Bin boundary choices in the unfolded observables are guided by detector resolution (bin widths are always larger than the corresponding 1σ resolution for measurements within each bin) and available data statistics, while allowing for sensitivity to the shape of the unfolded observable.

TABLE I. Unfolded phase space of the measurement.

Jet transverse momentum	$p_T^{\text{jet}} > 20$ GeV
Jet rapidity	$\lvert y_{\text{jet}}\rvert < 3.2$
Electron transverse momentum	$p_T^e > 15$ GeV
Electron pseudorapidity	$\lvert\eta^e\rvert < 1.1$
Vector sum of all neutrino transverse energies	$\not{p}_T > 20$ GeV
W boson transverse mass requirement	$M_T^W > 40$ GeV

A. Regularized unfolding using GURU

Due to the limited resolution of the detector, a significant fraction of events may be measured to be in a different kinematic interval than they were at the particle level, so a simple bin-by-bin correction for acceptance and efficiencies is not adequate. The aim of unfolding is to correct a measured observable back to the particle level observable, accounting both for the effect of finite experimental resolution, and for the detector response and acceptances. The relationship of the true particle level distribution $T(x^{\text{true}})$ to the reconstructed distribution $R(x)$ for an observable x can be written as follows:

$$R(x) = \int_{x^{\text{true}}_{\text{min}}}^{x^{\text{true}}_{\text{max}}} dx^{\text{true}}\, \mathcal{A}(x^{\text{true}})\mathcal{M}(x^{\text{true}}, x)T(x^{\text{true}}) + \xi \qquad (8)$$

where ξ represents the impurity correction that accounts for events that appear in our data sample which originate outside the fiducial region indicated in Table I, the limits $x^{\text{true}}_{\text{min}}$ to $x^{\text{true}}_{\text{max}}$ reflect the range of the variable we wish to measure, $\mathcal{A}(x^{\text{true}})$ represents the probability for a given observable to be seen at reconstruction level as a function of its particle level value (which takes into account acceptance, efficiencies, and analysis requirements), and $\mathcal{M}(x^{\text{true}}, x)$ is the migration matrix.

Experimental resolution affects the relationship between the reconstruction level and particle level objects so that corrections, $\mathcal{A} \times \mathcal{M}$, need to be applied and are derived using $W + n$-jet ALPGEN + PYTHIA MC simulation. In addition, there are events passing the reconstruction level selection requirements that are not within the phase space defined at the particle level. These events are corrected by the factor ξ. The ξ correction also includes events that pass both reconstruction and particle level selections, but due to jet energy resolution, the jet p_T ordering (or rapidity ordering in the case of those observables dependent on selecting the most forward-rapidity jets) is not consistent between particle and reconstructed jets. Finally, jets originating from additional $p\bar{p}$ interactions in the same and neighboring bunch crossings are covered by the correction ξ. A further acceptance correction is applied to those particle level events that would fail detector level selection requirements and thus not be reconstructed.

The migration matrix, obtained from the same MC simulation used to build the acceptance corrections, accounts for the probability of an event in a given particle level bin entering into various reconstruction level bins (or vice versa) through the relation:

$$\vec{x}_{\text{reco}} = \mathcal{M} \cdot \vec{x}_{\text{true}}. \qquad (9)$$

Ideally then, simply applying the inverse of the matrix $\mathcal{M}^{-1}$ to the measured reconstructed observables would provide us with the unfolded distributions. However, low-significance bins can introduce numerical instabilities and give rise to large, rapidly oscillating fluctuations that contain little meaningful information about the particle level distributions.

The program GURU counters this problem using a singular value decomposition technique that allows for a regularized inversion of the migration matrix [17]. This regularization imposes the requirement that the second derivatives of the distributions be small, equivalent to the condition that the unfolded distribution should be smooth. This suppresses fluctuations in the central values of the unfolded data. Care must therefore be taken to appropriately calculate statistical uncertainties that give an accurate reflection of the true spread of the data. This is discussed further in Sec. IV B. However, the advantage of this regularized method is that it allows the full matrix information to be used, which gives a reduced dependence on MC inputs compared to bin-by-bin corrections.

Figure 6 highlights three examples of the migration matrices used for observables measured in this paper, expressed as a probability that a particle level observable in a given bin is reconstructed in the same or different bin. Events entering the migration matrix must pass the selection at both reconstruction and particle levels, and the reconstructed jets relevant to the observable in question must retain their p_T ordering (or rapidity ordering, where applicable) between particle and reconstruction levels.

Following unfolding, the resultant distributions are normalized by the inclusive $W(\rightarrow e\nu)$ production cross section to reduce experimental uncertainties. This value of $\sigma_W = 1097^{+79}_{-89}$ pb was measured [2] in the same phase space, without any jet requirements, but with all other selection criteria as described in this analysis, and makes use of the same data set. Total inclusive n-jet cross sections were also previously measured [2] to be

(i) $\sigma_{W+1\text{-jet}} = 119.5^{+9.4}_{-8.3}$ pb,
(ii) $\sigma_{W+2\text{-jet}} = 19.0^{+2.4}_{-1.9}$ pb,
(iii) $\sigma_{W+3\text{-jet}} = 2.9^{+0.4}_{-0.4}$ pb,
(iv) $\sigma_{W+4\text{-jet}} = 0.39^{+0.09}_{-0.07}$ pb.

B. Evaluation of unfolding biases and statistical uncertainties

To assess and correct for any bias that might have been introduced into the unfolded distribution by our acceptance and unfolding corrections, and to determine the statistical and systematic uncertainties on the final results, we perform an ensemble test using pseudoexperiments constructed to mimic the measured and corrected data distributions that replicate the statistical fluctuations present in the data. A large sample of ALPGEN+PYTHIA MC $W + n$-jet events first receives an *ad hoc* correction at the particle level so as to describe data after unfolding. A further check is made to ensure this sample now also describes the reconstructed data. This correction is performed independently for each distribution under study, with the aim of creating distributions that mimic the data distributions.

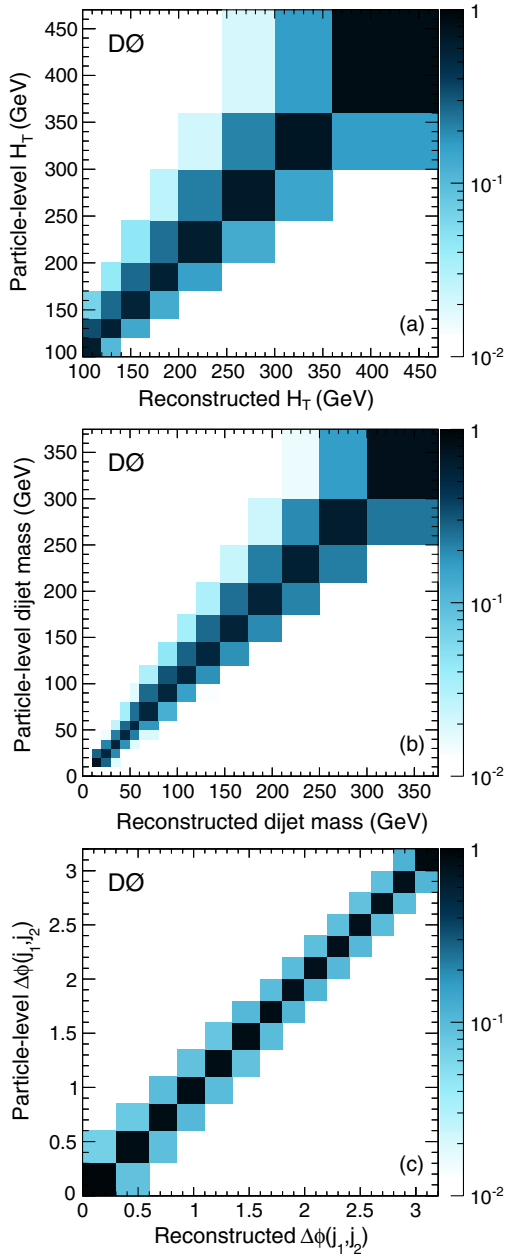


FIG. 6 (color online). Migration matrices for (a) H_T in the one-jet inclusive bin, (b) dijet mass in the two-jet inclusive bin, and (c) $\Delta\phi(j_1, j_2)$ in the two-jet inclusive bin. Element (x, y) represents the probability for the particle level observable in bin y to be reconstructed in observable bin x, and is represented by the axis on the right.

Five hundred ensembles per distribution are then drawn from this corrected ALPGEN+PYTHIA MC sample. The probability of an event entering a given bin in a given distribution is chosen such that the ensembles not only reproduce the data distributions at particle and reconstruction levels but also have statistical fluctuations, both bin-to-bin and in overall yield, as are observed in the data. For each distribution, a particle level and reconstruction level equivalent is constructed, and these ensembles reflect the results that would be expected from repeated independent experimental measurements.

Each of the ensembles in turn then receives the same acceptance corrections as are applied to the measured data distribution, and are unfolded using GURU under the same conditions. Unlike in the data, however, for each of these unfolded distributions we may compare the results to the corresponding ensemble's particle level distribution. For all ensembles in turn, for each distribution, for each bin, the residual

$$r = \frac{\text{particle level result} - \text{unfolded result}}{\text{unfolded result}} \tag{10}$$

is calculated and determines the fractional shift in the unfolded distribution from its true value. In each bin tested, over all ensembles, r is observed to follow a Gaussian distribution. The mean value of r is the fractional bias due to the unfolding procedure and the width of the Gaussian fit represents the statistical uncertainty on the measured unfolded result in that bin.

The unfolding bias for a given bin is applied as a correction to the unfolded data distribution in that bin. This correction is small, generally a few percent in magnitude, and always much smaller than the statistical uncertainty in the bin except in the case of the p_T-ordered third-jet emission probability results, where the bias becomes larger than the statistical uncertainty at wide opening angles. After the initial correction for the unfolding bias is applied to the unfolded data, the ALPGEN+PYTHIA MC samples used to obtain the acceptance corrections and migration matrices are corrected to the data. After the GURU unfolding and bias assessment procedures are applied to these new MC inputs, no further biases are observed. The statistical uncertainty on the fitted mean and any difference between the extracted bias from a Gaussian fit and the arithmetic mean of the distribution are taken as systematic uncertainties on the value of this bias.

We provide the bin-to-bin correlations on the statistical uncertainties for each measurement in HEPDATA [36] and in the Supplemental Material [37]. For convenience, we also provide the inverse of the statistical covariance matrix, which can be useful for fitting models directly to the measurements provided or to quantify the agreement of the data with a given model.

V. SYSTEMATIC UNCERTAINTIES

The two dominant sources of uncertainty on the majority of the unfolded distributions come from uncertainties on the jet energy scale (JES) and on the jet energy resolution (JER). At large jet rapidities and large opening angles between jets, uncertainties on the jet-vertex confirmation requirement also contribute a significant amount to the overall uncertainties, while at very low and at high electron transverse momenta, trigger efficiency uncertainties become one of the dominant systematics (due to limited

statistics in the $Z/\gamma^* \to e^+e^-$ data samples used to calculate the efficiencies). At large dijet opening angles, p_T, and invariant mass, systematic uncertainties from multijet backgrounds and the unfolding procedure become important due to limited statistics in the multijet-enriched data samples and to relatively large unfolding corrections.

Systematic uncertainties from JES, JER, jet-vertex confirmation, trigger efficiency, and jet identification efficiency are assessed with ensemble tests. The same MC ensembles used in the determination of the statistical uncertainties and the unfolding bias are once again unfolded, but this time using acceptance corrections and migration matrices derived using fully reconstructed ALPGEN+PYTHIA events produced with detector responses shifted by one standard deviation from the nominal values, for each of the sources of systematic uncertainty separately. Unfolding biases are then determined for these systematically shifted unfolded ensembles. As the input ensembles to this unfolding procedure are identical to those used for the nominal unfolding, any change in the unfolding bias must be due to the modified detector response and acceptance correction inputs to GURU. Any change in the unfolding bias from the nominal to the systematically shifted ensembles is therefore attributed to the effect of the shifted detector response and assigned as the corresponding systematic uncertainty due to that detector effect. This method enables the impact of a shifted detector response to be translated into the unfolded cross section result while also accounting for possible changes in the bin migrations due to the change in the response.

Uncertainties on the normalization and shape of the data-driven multijet background can arise from uncertainties on the electron efficiency, on the ϵ_{MJ} determination, and from the statistical uncertainties on the data samples used for the background modeling. The dominant contribution to the total uncertainty on the multijet background is the ϵ_{MJ} uncertainty except in the tails of distributions where control sample data statistics also play a role. The uncertainty on ϵ_{MJ} is determined for Run IIa and Run IIb independently, as a function of each of the observables measured. The two uncertainties are combined in each analysis bin by taking the corresponding systematic uncertainty from each data period and scaling its relative contribution according to the fractional multijet content originating from Run IIa and Run IIb data in a given bin. Despite being the largest background contribution to events with low jet multiplicity (see Fig. 3), total uncertainties due to multijet backgrounds in the inclusive one-jet bin are less than 1% on average. For high jet multiplicities this uncertainty can rise to 8% at high electron p_T and 20% for the largest jet rapidities in the inclusive four-jet bin.

The systematic uncertainty associated with the $t\bar{t}$ production cross section is determined by varying the assumed $t\bar{t}$ production rate within its theoretical uncertainty (+ 6%, −9%) [28], and determining the associated uncertainty on the $W + n$-jet contribution. This theoretical uncertainty translates into an uncertainty on the $t\bar{t}$-subtracted $W + n$-jet signal of 10%–15% or larger in kinematic regions where top quark backgrounds dominate, i.e. events with high jet multiplicity, large W boson p_T, or large H_T.

Electron identification uncertainties can originate from uncertainties on the background subtraction and fits to the efficiency turn-on curves. We also consider uncertainties on the p_T, instantaneous luminosity, and jet multiplicity (including electron-jet spatial separation) dependence of the efficiencies. We benefit from a reduction of the electron identification systematic uncertainty on the final measurements through division of the differential cross sections by the total measured W boson production cross section. After cancellations, the residual electron identification systematic uncertainty is approximately 1% averaged over all bins.

The systematic uncertainty on the measurement due to the unfolding procedure includes uncertainties on the unfolding bias value (both the statistical uncertainty on the mean and the difference between the correction determined from a Gaussian fit or from the arithmetic mean) and from statistical uncertainties on the acceptance corrections used by GURU. To check the dependence of an imperfect MC modeling of the kinematic spectra on the inputs to the unfolding procedure, the unfolding is repeated with a data-derived correction of the MC samples used to generate the acceptance and migration matrices so as to provide the best description of the observed data; the shift in the final results with these new inputs is assigned as a systematic uncertainty. The uncertainties on the unfolding procedure are small ($\lesssim$1%) in most analysis bins, but can be more significant at, e.g., large dijet rapidity separations 6%–9% and large jet rapidity 2%–4%.

The jet spatial matching criterion used in the acceptance and bin migration corrections is set to half and twice the size of the cone radius $R = 0.5$ of a reconstructed jet to test the dependence of the corrections on the matching choice. The impact on the final cross sections is found to be well below 1% for most distributions, but reaching up to 2% for high jet multiplicity events with high H_T and in events with wide dijet rapidity separations, and this variation is assigned as an additional systematic uncertainty.

All sources of systematic uncertainty on the theoretical modeling, detector response, background subtraction, and the unfolding procedure are added in quadrature to arrive at a total systematic uncertainty on the unfolded distributions. Figure 7 illustrates the total systematic uncertainty as a function of four representative unfolded observables. The contribution to the total uncertainty from jet energy scale, jet energy resolution, trigger efficiency, jet identification efficiency and jet-vertex confirmation are shown separately. Smaller sources of uncertainty, including electron identification efficiency, background shape/normalization, and unfolding procedure uncertainties, are shown as a

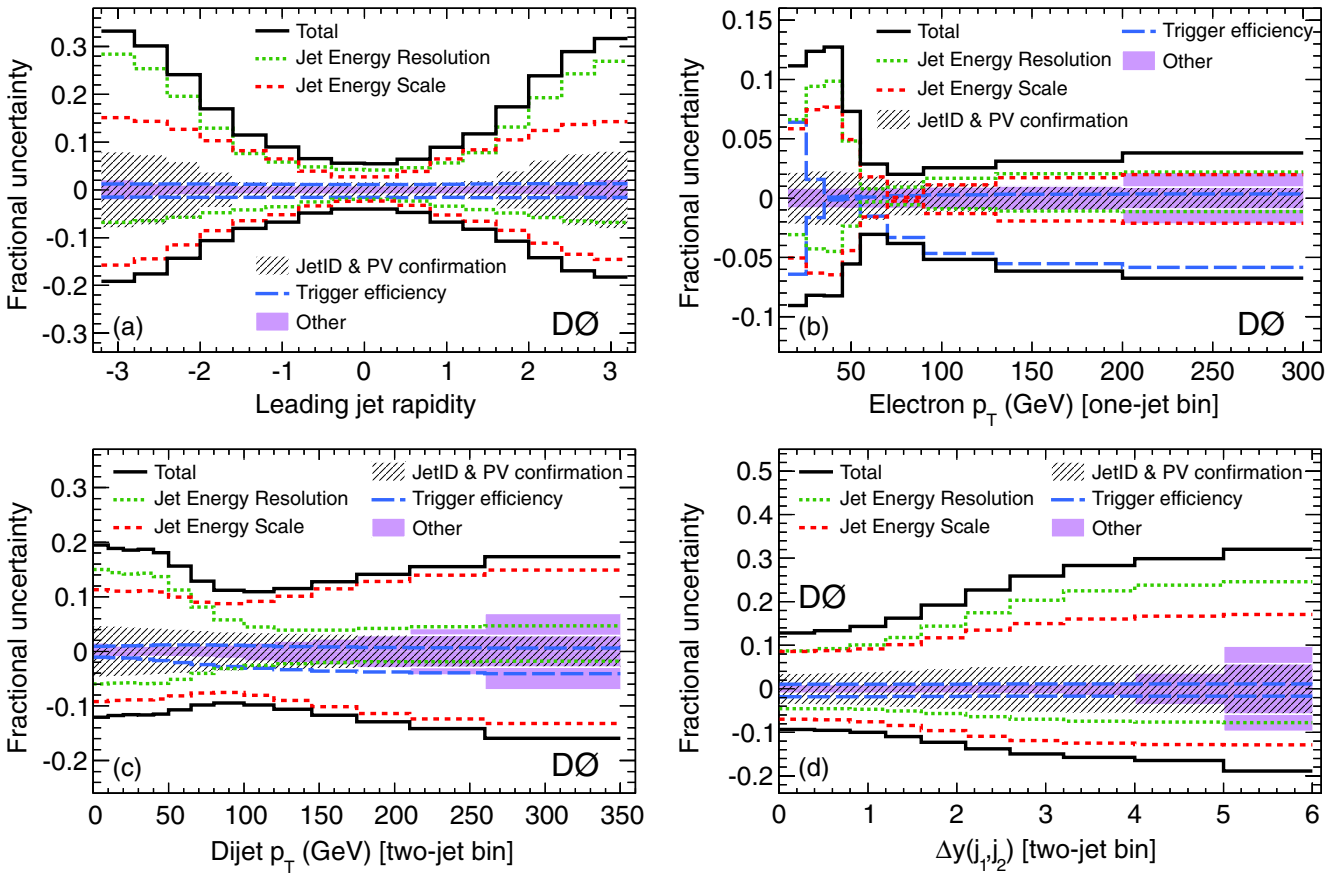


FIG. 7 (color online). Summary of systematic uncertainties on the normalized cross sections of (a) leading jet rapidity and (b) electron p_T in the inclusive one-jet bin, and of the (c) dijet p_T and (d) dijet rapidity separation of the two highest-p_T jets in the inclusive two-jet bin. The most significant sources of uncertainty are shown separately. Additional sources of uncertainty due to background modeling, electron identification, and the unfolding procedure are grouped under "Other."

combined contribution. The bin-to-bin correlations for all measured differential cross sections are given in HEPDATA [36] and in the Supplemental Material [37].

VI. THEORETICAL PREDICTIONS

We compare the data after correction for detector efficiencies and resolution effects to several theoretical models. Comparisons are made to a range of widely used parton shower and matrix element plus parton shower matched MC programs, to all-order resummation predictions, and to next-to-leading order pQCD calculations.

A. Monte Carlo programs

We compare our results to PYTHIA 6.425 with the PERUGIA2011 underlying event tune and the CTEQ5L parton density function (PDF) set [38]. PYTHIA is formally a leading-order (LO) calculation, but emulates the NLO matrix element through a reweighting of the parton shower in inclusive W production. We also compare to HERWIG 6.520 [39] at LO in α_s, using the CTEQ6L1 PDF set [40], and interfaced to JIMMY 4.31 [41] for modeling of multiple parton interactions (MPIs). To assess the impact of the inclusion of additional matrix elements provided in the ALPGEN MC program, we compare predictions from ALPGEN 2.414 hadronized in two ways, using either PYTHIA or HERWIG(+JIMMAY) with the same program version, underlying event tune, and PDF set as the stand-alone predictions.

Comparisons are also made to leading-order matrix element plus parton shower matched MC produced with SHERPA v.1.4.0 [42] using the CT10 [43] PDF set and with the factorization and renormalization scales chosen as discussed in Ref. [44]. The SHERPA default tuning parameters are used, with the exception of the MPI cutoff scale, which was tuned by the SHERPA authors to fit the CDF underlying event data in Drell-Yan production [45]. Hadronization is conducted using the SHERPA internal cluster fragmentation approach [42].

For these comparisons, results are provided using the D0 Run II midpoint cone algorithm at the particle level, with cone radius $R = 0.5$, and differential distributions are normalized by the inclusive W boson production cross section determined from the same MC program in the same measurement phase space (as summarized in Table I).

B. All-order resummation

High Energy Jets (HEJ) [13,46] is an implementation in a parton level MC generator of an exclusive, all-order resummation of the perturbative contributions to production of wide angle emissions at hadron colliders. Such predictions are particularly suited for description of events containing two jets with a large rapidity separation.

Predictions are produced using the D0 Run II midpoint cone algorithm at the parton level with $R = 0.5$. The factorization and renormalization scales are chosen to be $\mu = \mu_F = \mu_R = \max\{p_T^j\}$, with scale uncertainties estimated by varying this central scale choice by a factor of 2 or $1/2$. The resulting variation in the cross sections is taken as the scale uncertainty. As HEJ is capable only of describing two-jet and higher multiplicity events, differential cross sections are normalized by the measured inclusive W boson cross section of $\sigma_W = 1097^{+79}_{-89}$ pb reported in Sec. IV and in Ref. [2], to allow a like-to-like comparison of the distributions with the data. Scale uncertainties from HEJ are generally larger than those in NLO pQCD calculations, due in part to HEJ including only the dominant contributions needed to describe wide-angle jet emissions at all orders in the strong coupling constant and also from not being able to benefit from cancellations in the inclusive W boson cross section normalization.

C. Next-to-leading order pQCD

Next-to-leading order pQCD predictions of the production of $W + n$-jet events in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV have recently become available for up to four partons in the final state [47] from the BLACKHAT Collaboration [48]. NLO BLACKHAT+SHERPA predictions are obtained using BLACKHAT for calculation of the virtual terms, interfaced to SHERPA for calculation of all real emissions. Previous comparisons [2] of the predictions from BLACKHAT+SHERPA and another NLO pQCD calculation approach, ROCKET+MCFM [32,49], at Tevatron energies have been found to be in good agreement with each other for numerically similar choices of renormalization and factorization scales in their prediction of jet transverse momenta and total cross sections for up to four jets, so only

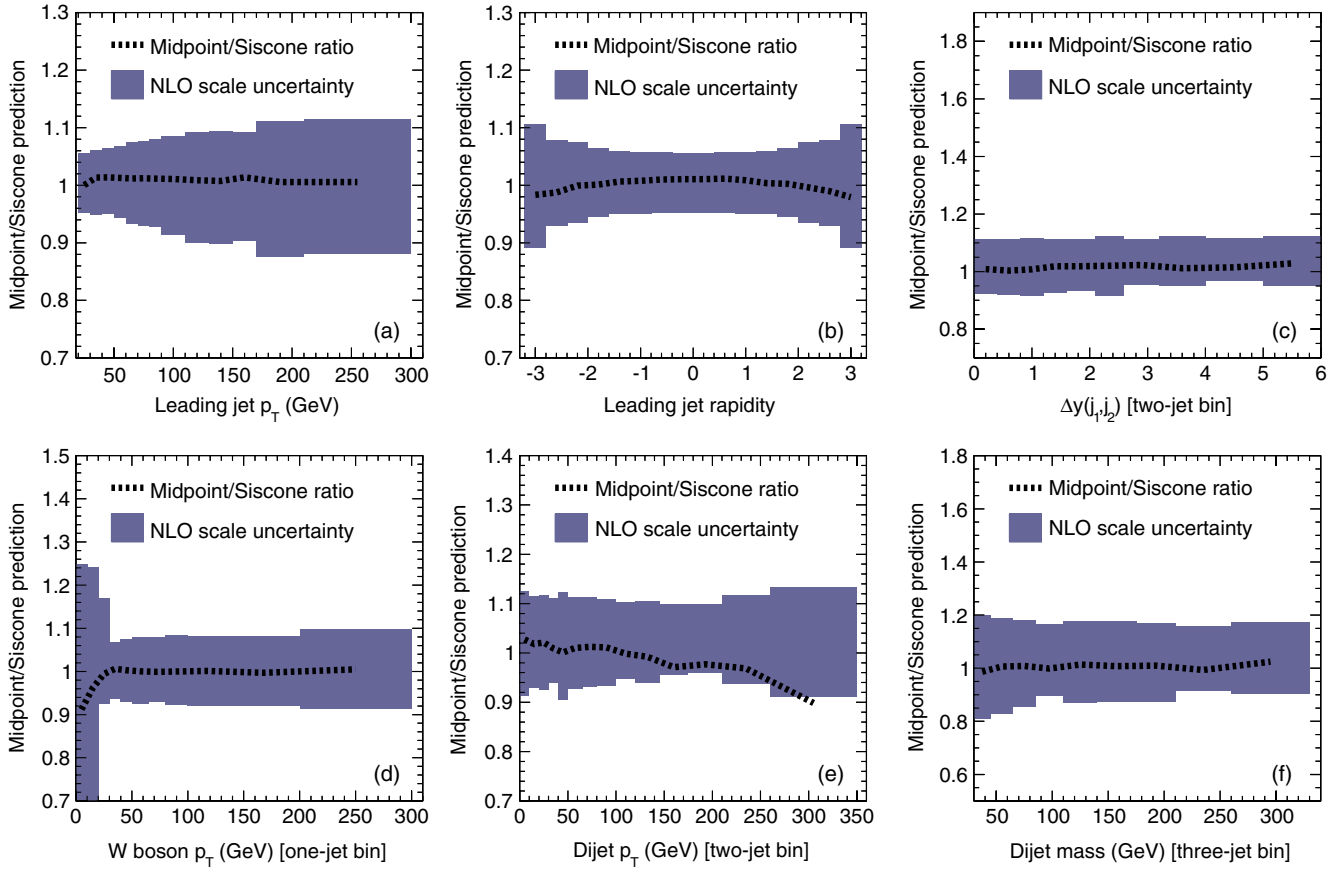


FIG. 8 (color online). Examples of correction factors (derived from SHERPA) accounting for differences between the SISCONE and D0 midpoint cone jet algorithm, applied to NLO BLACKHAT+SHERPA predictions for (a) leading jet p_T in inclusive $W + 1$-jet events, (b) leading jet rapidity in inclusive $W + 1$-jet events, (c) absolute dijet rapidity separation between the leading two jets in inclusive $W + 2$-jet events, (d) W boson p_T in inclusive $W + 1$-jet events, (e) dijet p_T in inclusive $W + 2$-jet events, and (f) dijet invariant mass in inclusive $W + 3$-jet events. NLO renormalization and factorization scale uncertainties are shown for reference purposes.

comparison of experimental data with BLACKHAT+SHERPA predictions are made in this paper.

The BLACKHAT+SHERPA predictions employ the renormalization and factorization scale choice $\mu = \mu_F = \mu_R = \frac{1}{2}H_T'$ where H_T' is defined as the scalar sum of the parton and W boson transverse energies. To determine the uncertainty due to the choice of renormalization and factorization scales, the scales are multiplied by factors of 0.5 and 2, and the cross sections are recomputed. The variation in the cross sections is taken as the scale uncertainty. The calculations use the MSTW2008nlo68cl PDF set [50], with values of $\alpha_s(\mu)$ set consistently with the PDF choice, using a two-loop running. Uncertainties arising from the choice of PDF are studied [50] using the Hessian method by the Blackhat Collaboration and found to be negligible ($\lesssim 2\%$) in comparison with scale uncertainties.

Predictions are generated with the SISCONE [51] jet algorithm applied at the parton level with split-merging fraction $f = 0.5$ and cone radius $R = 0.5$, rather than with the D0 Run II midpoint cone jet algorithm, because the D0 Run II midpoint cone algorithm is not integrated into the BLACKHAT+SHERPA code. The predicted distributions are then corrected for the effect of using SISCONE rather than the D0 Run II midpoint cone jet algorithm using SHERPA (described below). All cross section results are normalized by the theoretical inclusive W boson production cross section prediction at NLO [48] determined to be 1153^{+17}_{-7} pb in the unfolded phase space.

1. Nonperturbative corrections

As a fixed-order pQCD prediction, BLACKHAT+SHERPA provides a parton level prediction that is not immediately comparable to the unfolded experimental data. Bin-by-bin corrections for nonperturbative QCD effects, due to hadronization and the underlying event, must be derived to compare the NLO predictions with data. Nonperturbative corrections were produced using SHERPA 1.4.0 with the CT10 PDF set.

These corrections are calculated by taking the ratio of the observed differential cross section derived from SHERPA at the particle level (using SHERPA'S internal cluster fragmentation model) to the differential cross section from SHERPA at the parton level, which includes parton showering but without hadronization or MPI and with electron final state emission disabled. Uncertainties on these nonperturbative corrections are obtained [52] by recalculating the particle level SHERPA results as described above using the Lund string fragmentation model [53], taking the difference between the two as a symmetric systematic uncertainty on this correction.

Within the nonperturbative correction, we also apply an additional correction to account for the impact of the jet algorithm differences between experiment and NLO BLACKHAT+SHERPA. The ratio of the particle level cross sections with the D0 midpoint cone algorithm to the particle level cross sections with the SISCONE jet algorithm is computed. The total correction applied is then the product of the term to account for hadronization and underlying event effects with the correction to account for the jet algorithm mismatch:

$$\text{total correction} = \frac{\sigma^{\text{midpoint}}_{\text{particle}}}{\sigma^{\text{siscone}}_{\text{particle}}} \times \frac{\sigma^{\text{siscone}}_{\text{particle}}}{\sigma_{\text{parton}}} = \frac{\sigma^{\text{midpoint}}_{\text{particle}}}{\sigma_{\text{parton}}}. \quad (11)$$

Figure 8 presents some examples of the SISCONE to D0 midpoint cone jet algorithm correction factors as a function of a representative subsample of the unfolded observables presented in this paper, and as computed with SHERPA. These jet algorithm corrections are of smaller magnitude than the corrections for underlying event and hadronization effects.

For most distributions the overall correction is small (with respect to the scale uncertainties on the theory). Where total nonperturbative corrections become large ($\geq 50\%$), BLACKHAT+SHERPA predictions are not displayed in comparison to data. Examples of the combined

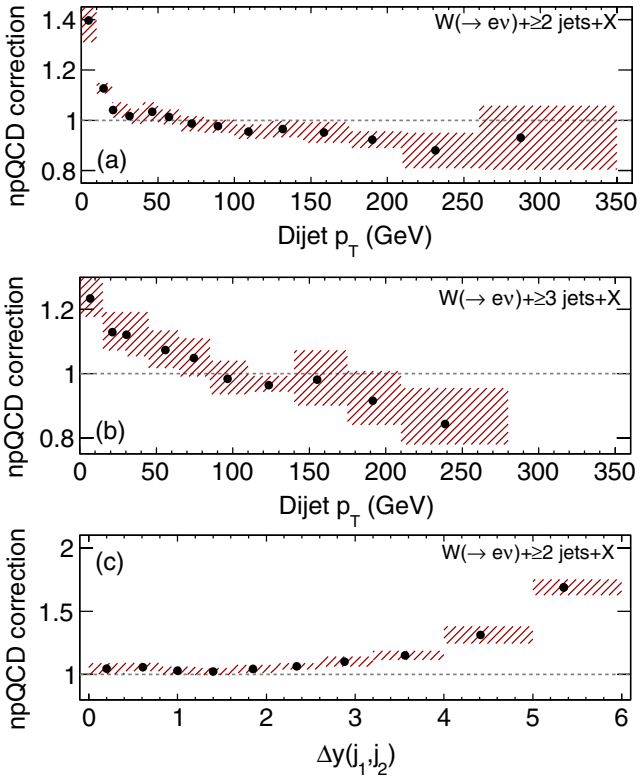


FIG. 9 (color online). Nonperturbative QCD (npQCD) correction factors used to correct NLO BLACKHAT+SHERPA theoretical predictions from the parton to the particle level. Uncertainties are estimated by combining the statistical uncertainty with the systematic uncertainty arising from different hadronization models, and are shown as the shaded band. Examples shown are for the dijet p_T distributions in the (a) inclusive two-jet and (b) three-jet multiplicity bins, and for (c) dijet $\Delta y(j_1, j_2)$ in the inclusive two-jet multiplicity bin, and are among the largest corrections. Corrections are derived using SHERPA.

correction applied to the data for those distributions that do exhibit notable shape dependencies are shown in Fig. 9.

Full tables of the nonperturbative corrections applied to the NLO BLACKHAT+SHERPA predictions, including the jet algorithm correction terms, are available in HEPDATA [36] and documented in Supplemental Material [37] to facilitate comparison between future pQCD calculations and the experimental data presented here.

VII. EXPERIMENTAL RESULTS

Measurements of forty observables are documented in this paper, measured in the phase space defined in Table I. These consist of thirty-three differential cross section measurements, normalized by the inclusive W boson cross section in the same phase space, four measurements of average jet activity as a function of dijet rapidity separations and the scalar sum of W boson and jet transverse energies, and three measurements of the probability of subsequent jet emission in W + dijet events as a function of dijet rapidity separation under various conditions. Unless otherwise noted, all jets are ordered by p_T and all $W + n$-jet distributions specify n-jet inclusive multiplicities. Figures 10 to 26 present the results in comparison with various theoretical predictions. Data points are placed at the bin average, defined as the value where the theoretical differential cross section is equal to the mean cross section within the bin, following the prescription detailed in Ref. [54]. Error bars on data points represent statistical and systematic uncertainties added in quadrature.

All results presented here are available in tabulated form in HEPDATA [36] and in Supplemental Material [37], along with correlation matrices for the statistical and systematic uncertainties. We do not provide quantitative measurements of the compatibility of the different theoretical predictions with our measurements. However, using the tabulated results will allow goodness of fit measurements with any theoretical model.

A. Differential cross sections

Measurement of relative rates and shapes of jet rapidities in $W + n$-jet events allow us to compare different theoretical approaches to jet emissions and are key to understanding searches for new physics characterized by forward jet emission, as well as standard model measurements including vector boson fusion and vector boson scattering. For all observables, experimental uncertainties are smaller than, or of similar magnitude to, corresponding theoretical uncertainties on predictions and have the potential to discriminate between different theoretical approaches.

Measured normalized $W + n$-jet cross sections are shown as a function of the nth-jet rapidity in inclusive $W + n$-jet events in Fig. 10, highlighting the wide range of predicted differential spectra between the models considered. While all predictions largely agree in shape at central rapidities, for $|y| > 1$ discrepancies with the data begin to

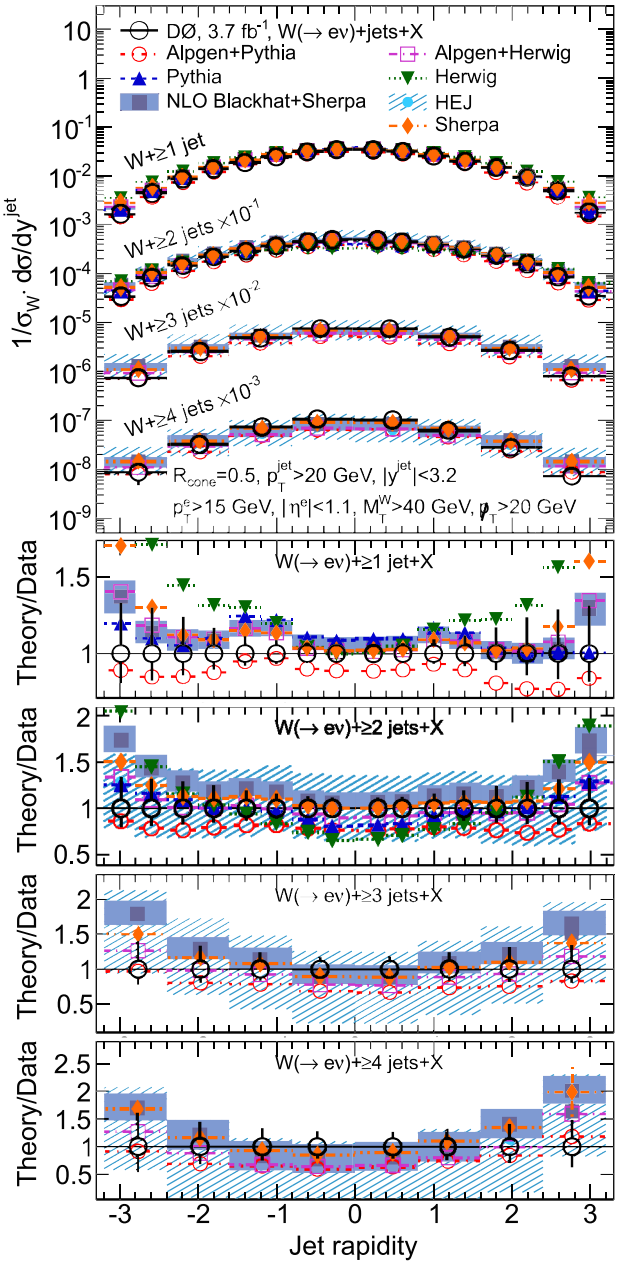


FIG. 10 (color online). Measurement of the nth-jet rapidity distributions in inclusive $W + n$-jet events for $n = 1$–4 and comparison to various theoretical predictions. Lower panes show theory/data comparisons for each of the n-jet multiplicity bin results separately.

emerge, resulting in large differences at forward rapidities. NLO pQCD, HEJ, SHERPA, and HERWIG predictions are found to slightly overpredict the forward jet rate, while PYTHIA and ALPGEN+PYTHIA give predictions approximately in agreement with the data. Bjorken x [55] values of gluon or light quark PDFs probed by typical $W + n$-jet events at large rapidity where predictions begin to diverge from data are $x \approx \mathcal{O}(10^{-2})$, of similar value to those probed by the ATLAS Collaboration in $W + n$-jet events at higher center-of-mass energies, where discrepancies are

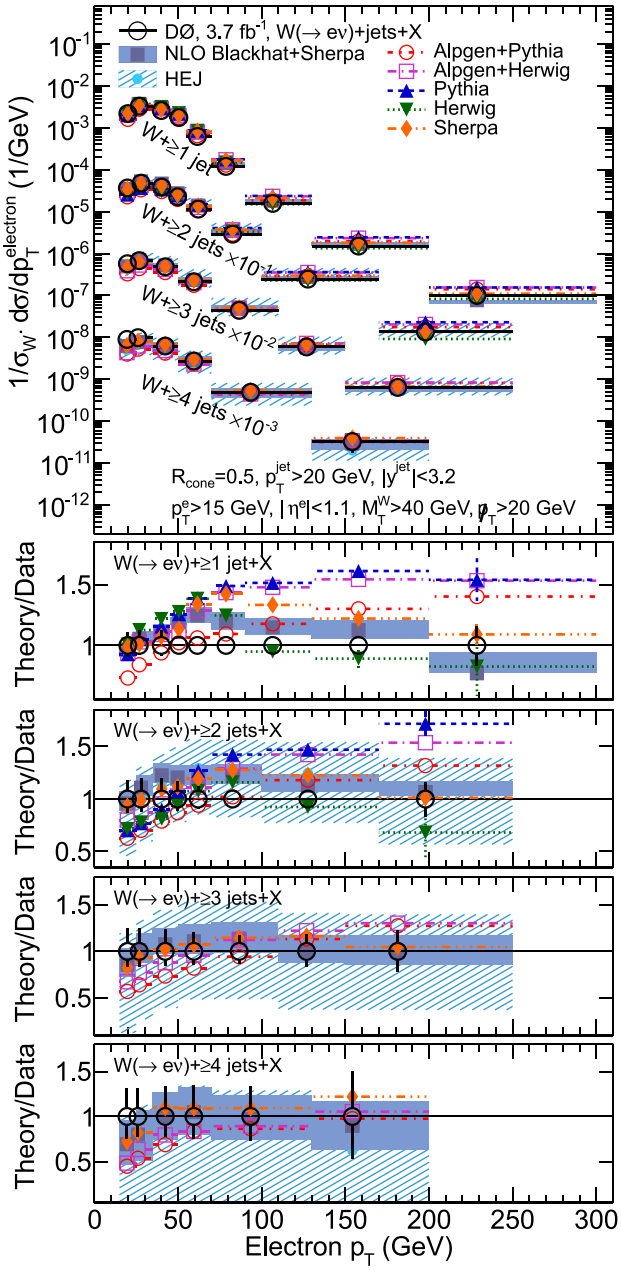


FIG. 11 (color online). Measurement of the electron transverse momentum distributions in inclusive $W + n$-jet events for $n =$ 1–4 and comparison to various theoretical predictions. Lower panes show theory/data comparisons for each of the n-jet multiplicity bin results separately.

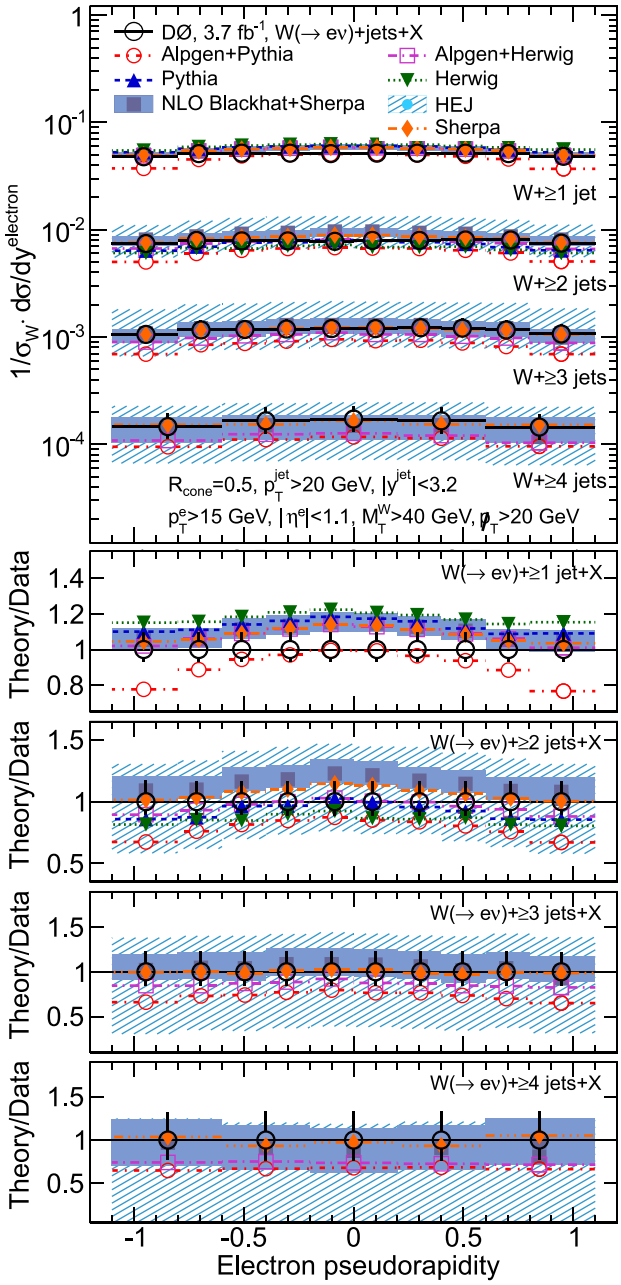


FIG. 12 (color online). Measurement of the electron pseudorapidity distributions in inclusive $W + n$-jet events for $n =$ 1–4 and comparison to various theoretical predictions. Lower panes show theory/data comparisons for each of the n-jet multiplicity bin results separately.

also observed [4]. These observations may suggest some tension with current determinations of the gluon or light quark densities in the proton.

Figure 11 presents normalized cross sections as a function of the electron transverse momentum for inclusive one- to four-jet events. The study of electron p_T provides kinematic information that is complementary to the previously measured jet p_T distributions, since the electron p_T is only partly correlated to the p_T of any jet in the event. We observe a trend for predictions to slightly underestimate the data at low (< 50 GeV) p_T in higher n-jet channels, although given the uncertainties on the theory predictions, the deviations are not significant.

At higher electron transverse momenta, BLACKHAT+SHERPA, HEJ, and SHERPA predictions describe the spectrum well, although the one-jet rate is slightly overpredicted by BLACKHAT+SHERPA and SHERPA in this region. HERWIG does not describe the low-p_T shape, and a distinct change in slope is observed above ≈ 60 GeV

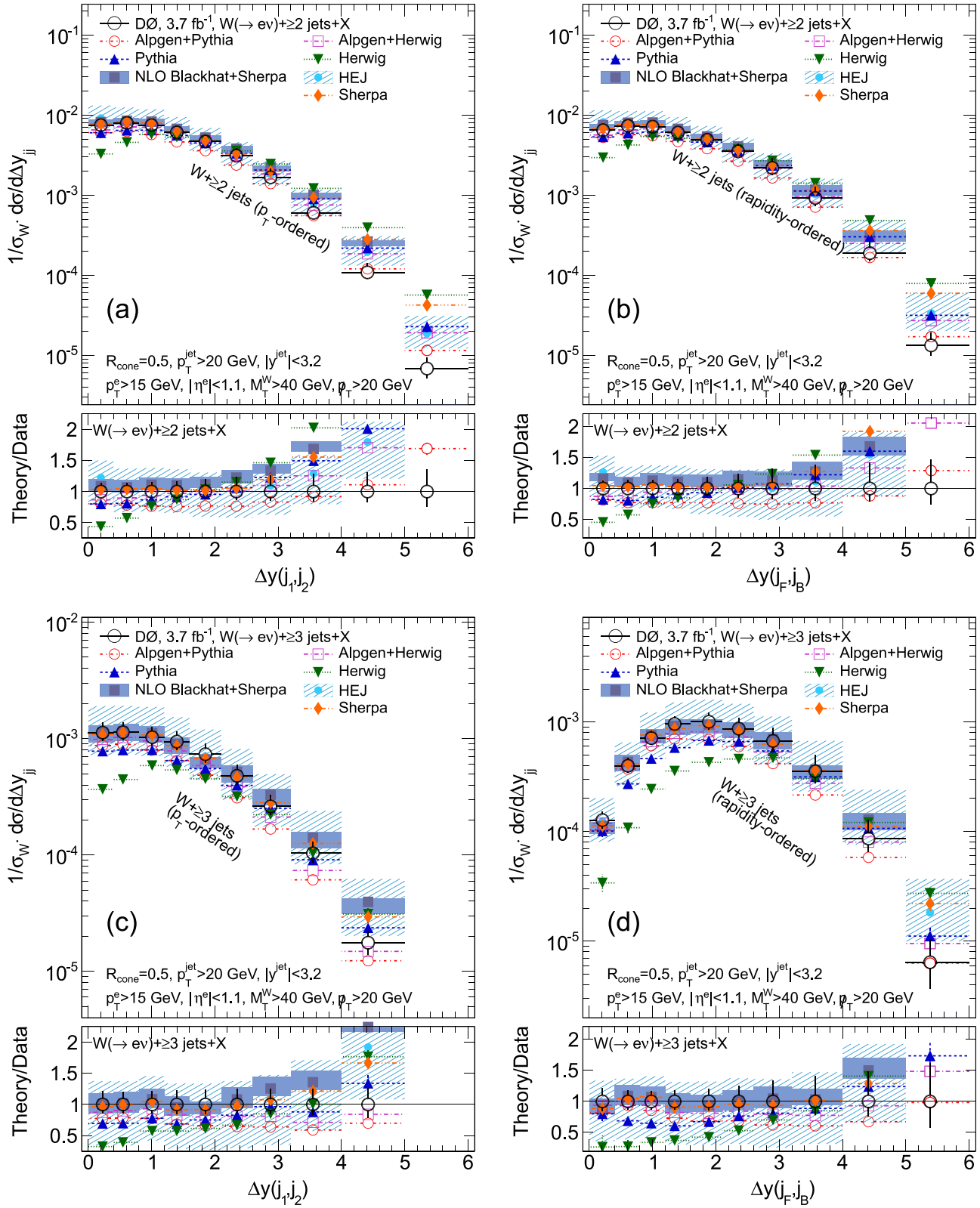


FIG. 13 (color online). Measured dijet rapidity separation in inclusive W + 2-jet events for (a) the two leading p_T jets and (b) the two most rapidity-separated jets and in inclusive W + 3-jet events for (c) the two leading p_T jets and (d) the two most rapidity-separated jets, with comparison to various theoretical predictions. Lower panes show theory/data comparisons.

related to the transition from a pure parton shower to a matching of the parton shower to the W + 1-jet matrix element in HERWIG. PYTHIA predicts a harder p_T spectrum than observed in data, overpredicting the rate at high p_T by about 50%. Interfacing the PYTHIA parton shower to the ALPGEN matrix element calculation improves the normalization somewhat without changing the slope of the distribution, bringing ALPGEN+PYTHIA predictions in closer agreement to those from NLO pQCD.

The corresponding electron pseudorapidity spectra for each jet multiplicity are presented in Fig. 12. ALPGEN+PYTHIA predicts a narrower electron pseudorapidity spectrum than observed in data at low jet multiplicities, underpredicting the rate at $|\eta| = 1.0$ by over 20%. Other MC generators and BLACKHAT+SHERPA predictions do not exhibit the same behavior although there is some indication of a shape difference between theory and data at the 10% level across the measured interval.

In Fig. 13, normalized cross sections are presented as a function of dijet rapidity separation in inclusive two-jet and three-jet events, for two distinct jet pairings. The first configuration defines the absolute dijet rapidity separation between the two highest-p_T jets in the event $[\Delta y(j_1, j_2)]$, the second defines the absolute separation between the two most rapidity-separated jets (generally with one forward jet, j_F, and one backward jet, j_B, in rapidity and both jets with $p_T^{\text{jet}} > 20$ GeV) in the event $[\Delta y(j_F, j_B)]$.

Study of the rapidity separation of the two leading (highest-p_T) jets in the event in W + dijet (and W + 3-jet) events is a test of wide-angle soft parton radiation and matrix element plus parton shower matching schemes. Understanding the distribution of this variable is important to distinguish VBF processes from the larger W + jets contributions and is key for background modeling for future searches and measurement of the Higgs boson in the vector boson fusion and vector boson scattering modes.

Measuring the rapidity separation of the two most rapidity-separated jets in inclusive W + dijet events provides sensitive information on additional QCD radiation in the event, as does the measurement of the same variable for the two leading jets. The rapidity-ordered configuration is sensitive to BFKL-like dynamics because, in this case, the dijet invariant mass is much larger than the transverse momentum of the jets, allowing tests of the advanced modeling of high-p_T wide-angle emissions missed by a standard parton shower approach [56]. The BFKL approach to resummation of higher orders in the strong coupling constant is important for small Bjorken x events, which correlate with forward QCD emissions.

Predictions from HERWIG consistently favor wider separations between both the highest-p_T and most rapidity-separated jets but underpredict the measured rate at small rapidity intervals. This mismodeling increases in the

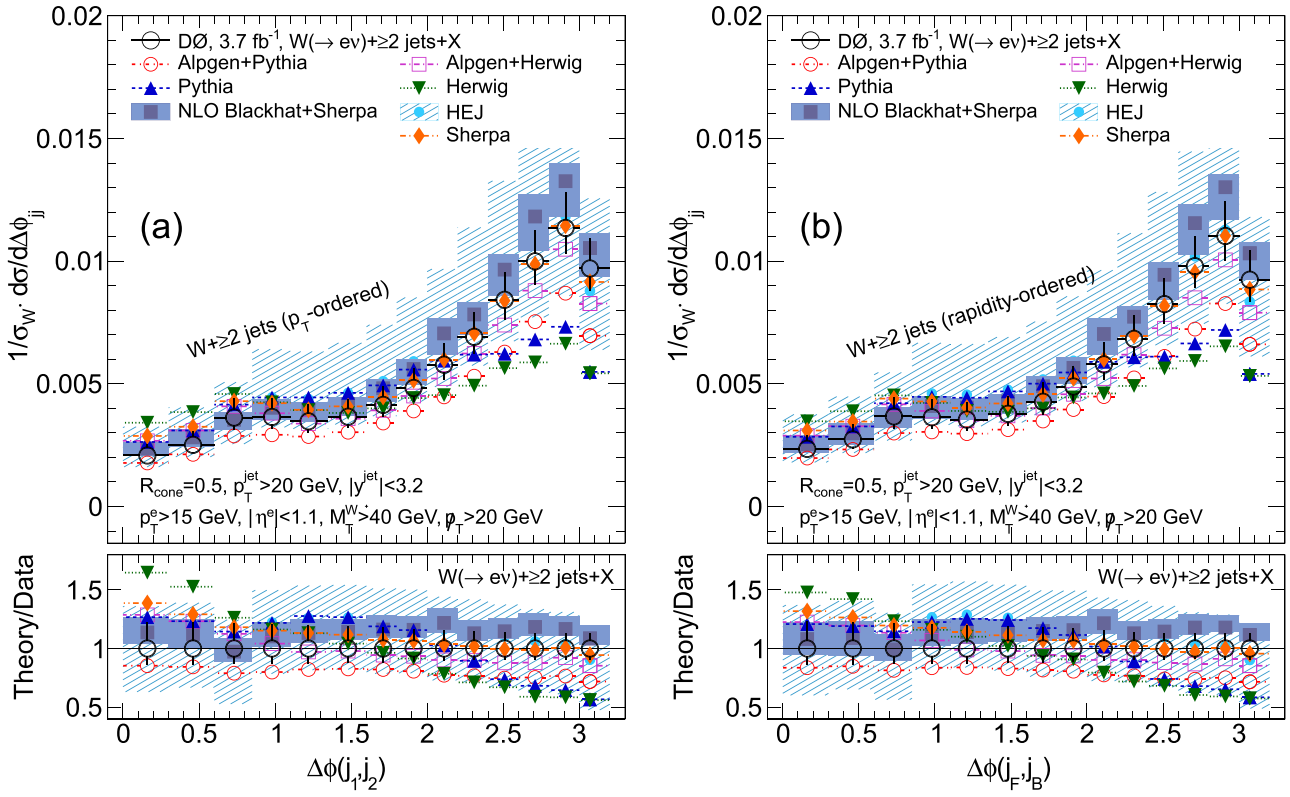
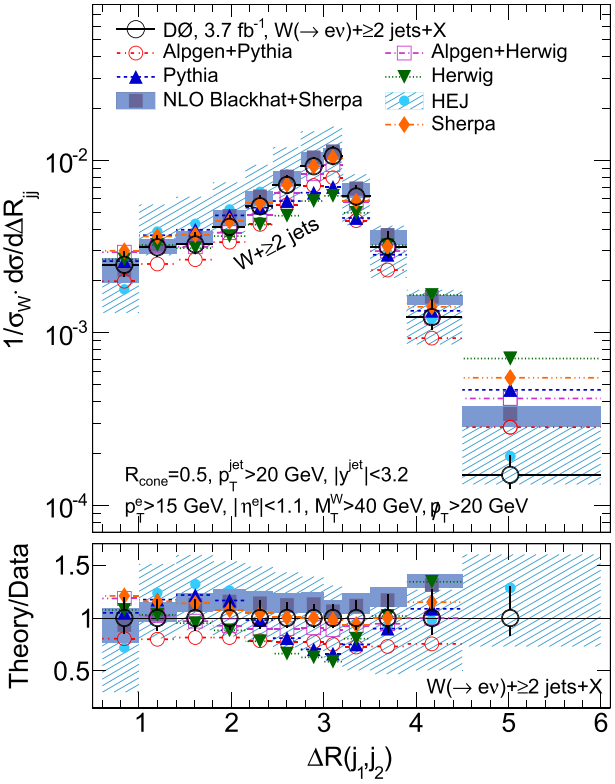


FIG. 14 (color online). Measured differential cross sections and various theoretical predictions for dijet $\Delta\phi$ for (a) the two highest-p_T jets and (b) the two most rapidity-separated jets in W + 2-jet events. Lower panes show theory/data comparisons.



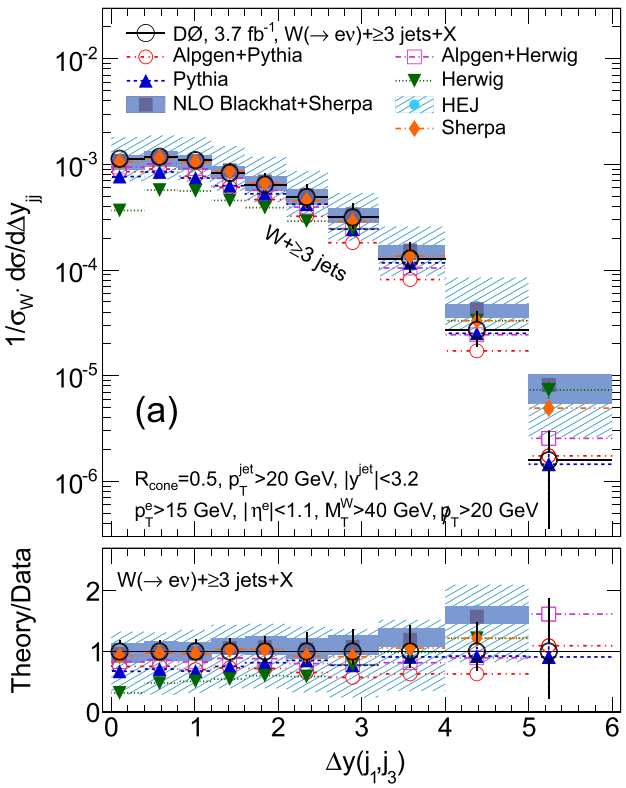


FIG. 15 (color online). Measured differential cross sections and various theoretical predictions for dijet $\Delta R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$ in inclusive $W + 2$-jet events. The lower pane shows theory/data comparisons.

three-jet events and in rapidity-separated configurations. NLO pQCD predictions are able to describe small rapidity intervals well, but increasingly overpredict the rate seen in data as the rapidity separation grows. The distribution is dominated by contributions from nonperturbative processes at large rapidity separations. Typical corrections [see Figure 9(c)] for nonperturbative QCD effects of $\approx 40\%$ for $4 \leq \Delta y(j_1, j_2) < 5$ and $\approx 75\%$ for $5 \leq \Delta y(j_1, j_2) < 6$ thus limit the applicability of NLO pQCD predictions in this region. The shape in inclusive three-jet events is well described by ALPGEN+(PHYTHIA/HERWIG) and PYTHIA, all of which have better performance than NLO pQCD in these observables, suggesting that the contributions of soft emissions from the parton shower are necessary and well tuned.

The azimuthal angle separation between the two leading jets or most rapidity-separated jets in $W +$ dijet events is a sensitive test of modeling of higher-order corrections in theoretical calculations, as additional jet emissions (including those below the reconstruction threshold) will modify the distribution of the angular separation of the two leading jets due to the reduction in phase space allowed with each additional emission. Some theoretical and experimental analyses prefer to study p_T-ordered jets, while others use rapidity-ordered jets. These azimuthal

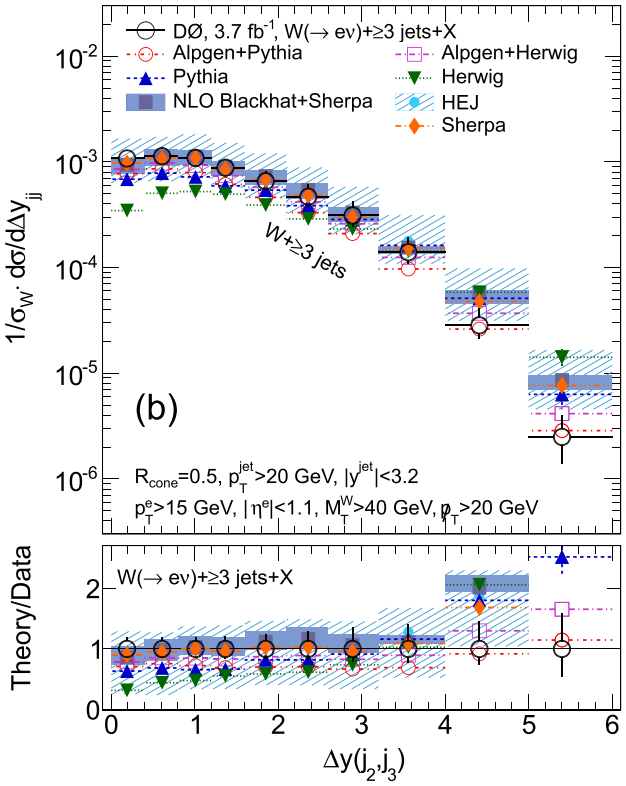


FIG. 16 (color online). Measurement of the spectrum of dijet rapidity separation of (a) the first and third, and (b) the second and third p_T-ordered jets in inclusive $W + 3$-jet events and comparison to various theoretical predictions. Lower panes show theory/data comparisons.

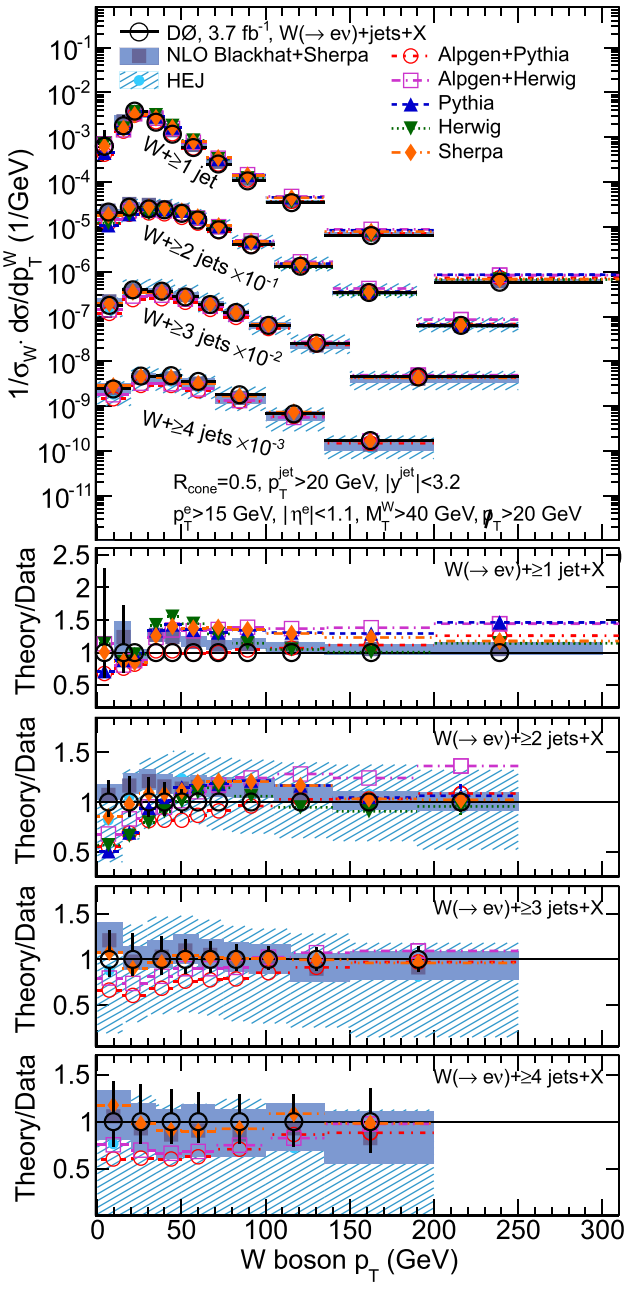


FIG. 17 (color online). Measurement of the W boson transverse momentum distributions in inclusive $W + n$-jet events for $n = 1$–4 and comparison to various theoretical predictions. Lower panes show theory/data comparisons for each of the n-jet multiplicity bin results separately.

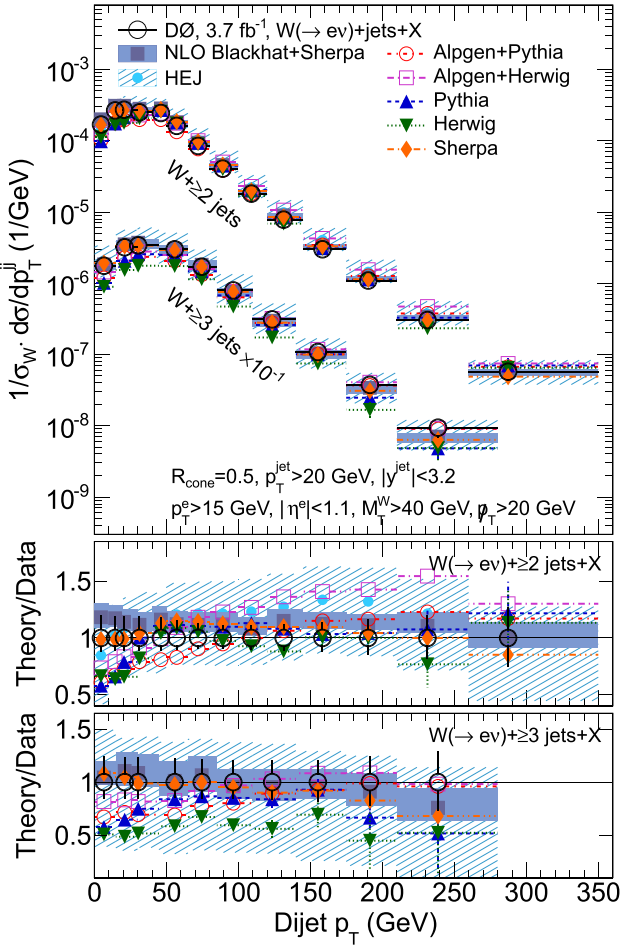


FIG. 18 (color online). Measurement of the dijet transverse momentum spectrum of the dijet system in inclusive $W + 2$-jet and $W + 3$-jet events and comparison to various theoretical predictions. Lower panes show theory/data comparisons for each of the n-jet multiplicity bin results separately.

correlations are therefore studied as a function of both the two leading $[\Delta\phi(j_1, j_2)]$ and two most rapidity-separated $[\Delta\phi(j_F, j_B)]$ jets with the results presented in Fig. 14. The two corrected observables $[\Delta\phi(j_1, j_2)$ and $\Delta\phi(j_F, j_B)]$ are similar in shape.

Jet pairings with large (close to π) separation are generally modeled via matrix element calculations while small separations are modeled mainly via the parton shower. Hard radiative corrections from all-order resummation approaches can modify the predicted spectrum for this observable. The spectra are well described by the all-order resummation (HEJ) and NLO (BLACKHAT+SHERPA) approaches, although the latter is a little high in overall rate. SHERPA provides a reasonable description of the data within experimental uncertainties. Predictions from ALPGEN behave similarly, despite significant normalization differences depending on whether PYTHIA or HERWIG are used for the parton showering. There is a tendency for both parton shower MC generators, PYTHIA and HERWIG, to predict significantly reduced emissions at large $\Delta\phi$ than are observed in data and more collinear emissions, the modeling of which can be improved with these data.

Figure 15 shows the $\Delta R = \sqrt{(\Delta y)^2 + (\Delta\phi)^2}$ spectrum between the two leading jets in the inclusive $W +$ dijet sample. Study of the opening angle between the two highest-p_T jets in the event (in y-ϕ space) allows for testing both initial and final-state radiation modeling in theoretical calculations in the large and small ΔR regions,

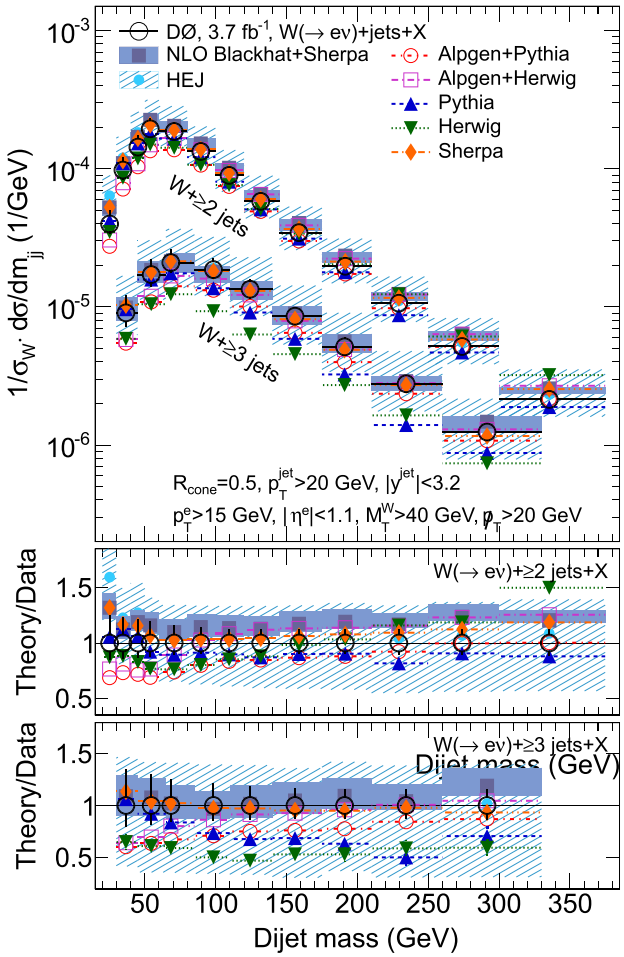


FIG. 19 (color online). Measurement of the invariant mass spectrum of the dijet system in inclusive W + 2-jet and W + 3-jet events and comparison to various theoretical predictions. Lower panes show theory/data comparisons for each of the n-jet multiplicity bin results separately.

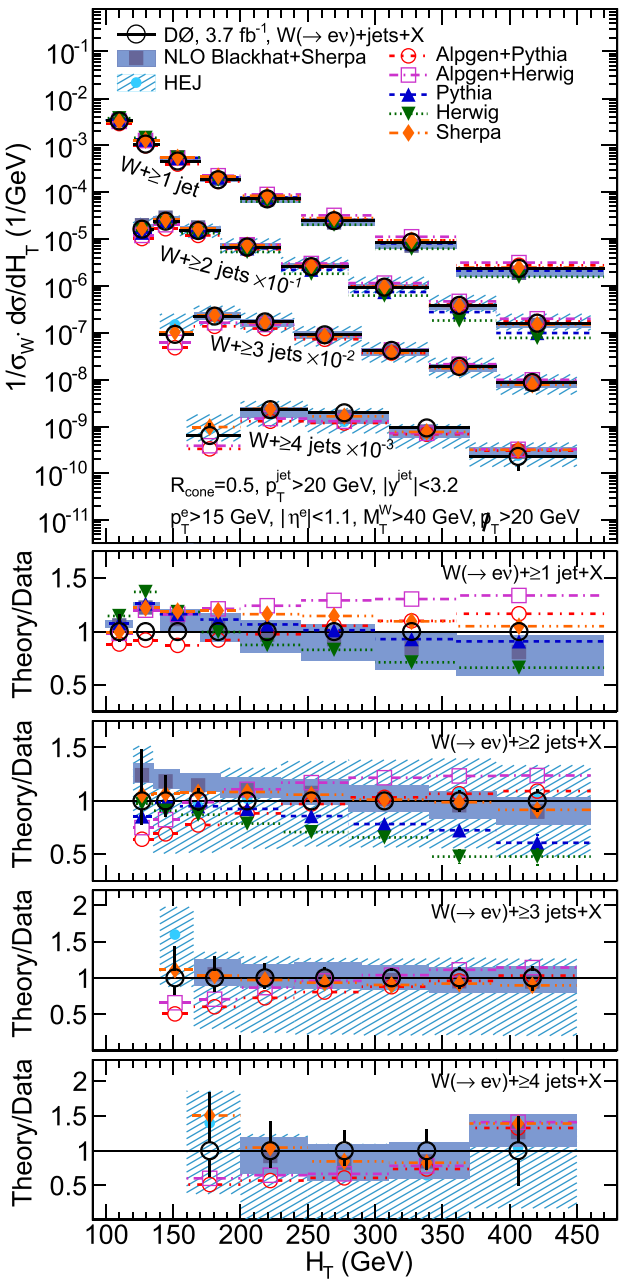


FIG. 20 (color online). Measurement of the distribution of the scalar sum of transverse energies of the W boson and all jets in the event for inclusive W + n-jet events for n = 1–4 and comparison to various theoretical predictions. Lower panes show theory/data comparisons for each of the n-jet multiplicity bin results separately.

respectively. This is also an important experimental variable to properly model dijet correlations in backgrounds for precision measurements and searches for new physics. Both HEJ and NLO BLACKHAT+SHERPA calculations do not accurately model the shape of this distribution at large and small opening angles. Both SHERPA and ALPGEN+PYTHIA provide a good description of the shape observed in data except at the largest ΔR.

Figure 16 presents measurements of the rapidity separation between the third-hardest jet in inclusive W + 3-jet events and either the leading or subleading jet in the event. Measurement of the angular correlations between various jet pairings gives us further information to constrain QCD radiation modeling. SHERPA, HEJ, and NLO BLACKHAT+SHERPA provide a good description of the shapes of these distributions, with some tension again observed at the very largest rapidity separations. PYTHIA, HERWIG, and ALPGEN matrix element matched approaches show deviations from the data, particularly at low rapidity separation.

The W boson transverse momentum distribution for inclusive (n = 1–4)-jet multiplicity bins is shown in Fig. 17. Good agreement between the data and BLACKHAT+SHERPA and HEJ is observed for all predicted jet multiplicities. A change in behavior in NLO pQCD/data is observed at W boson transverse momenta near the jet p_T threshold of 20 GeV, most notably in the inclusive W + 1-jet sample. As the boson p_T approaches this threshold,

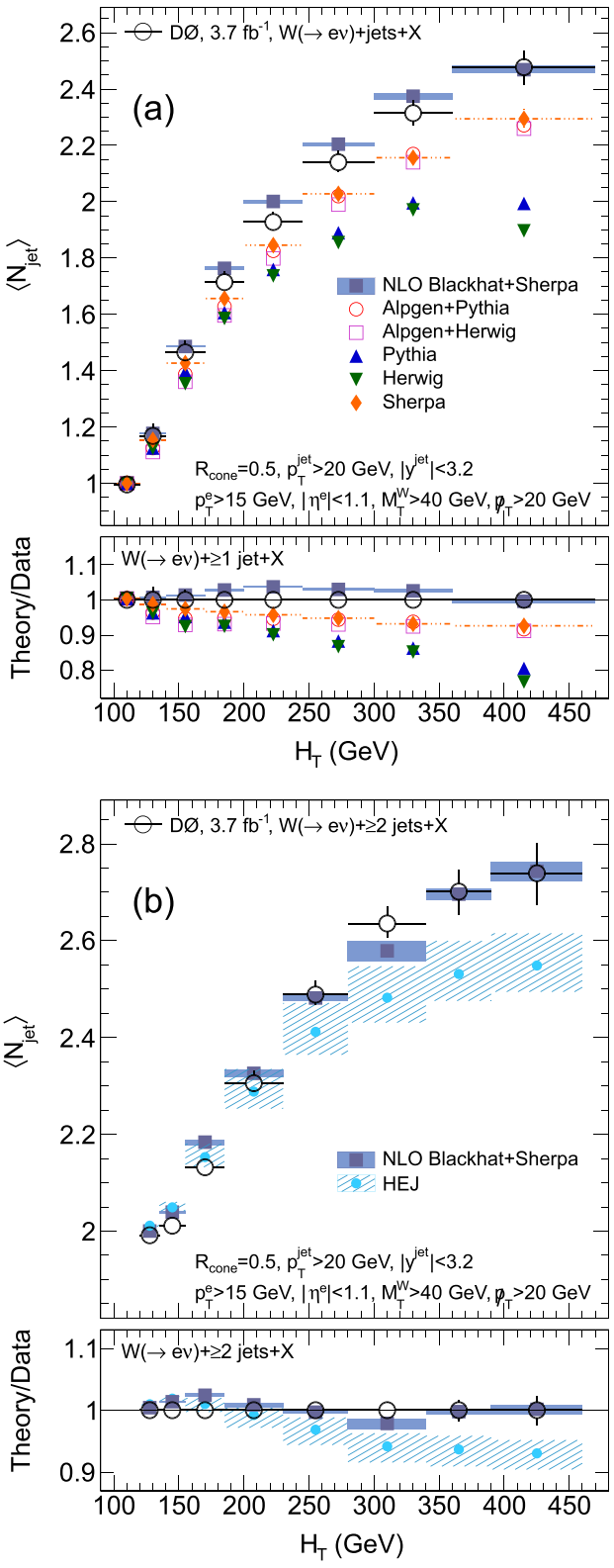


FIG. 21 (color online). Measurements of the mean number of jets in (a) inclusive W + 1-jet events and (b) inclusive W + 2-jet events as a function of the scalar sum of transverse energies of the W boson and all jets in the event, with comparison to various theoretical predictions. The lower pane shows theory/data comparisons.

the boson and jet are produced back-to-back. At transverse momenta below this threshold, nonperturbative effects dominate and the particle-level-corrected fixed-order calculations are expected to become unreliable.

The dijet transverse momentum and invariant mass spectra in the inclusive two-jet and three-jet multiplicity bins are shown in Figs. 18 and 19. Dijet quantities in this article are calculated from the highest- and second highest-p_T jets in the event. As well as testing the modeling of angular correlations between the two highest-p_T jets in the event in MC generators, validation of theoretical modeling of the invariant mass and p_T distributions of the dijet system in W + 2-jet events is important for searches for beyond the standard model physics. We provide measurements in this variable to allow the study of modeling differences between theoretical approaches.

Agreement in the shape of the dijet p_T distribution (Fig. 18) is observed between data and predictions from NLO BLACKHAT+SHERPA, HEJ, and SHERPA. Notable discrepancies in the ALPGEN, PYTHIA, and HERWIG modeling are observed at low p_T.

In Fig. 19, the data are presented as a function of dijet mass. The theoretical predictions vary considerably in shape, and even the NLO pQCD cannot describe the shape over the full spectrum in the two-jet bin. The distribution in the three-jet bin is well described by SHERPA, NLO pQCD and HEJ (although the latter has large uncertainties).

Figure 20 shows the differential distributions of W + n-jet events as a function of H_T, the scalar sum of the transverse energies of the W boson and the partons in the event. Accurate prediction of the distribution of the scalar sum of the transverse energies of the W boson and all high-p_T ($p_T > 20$ GeV) jets in W + n-jet events is important as this variable is often used as the renormalization and factorization scale for theoretical predictions of vector boson plus jet events at the Tevatron and the LHC. In addition, this variable is often chosen as a discriminant in searches for signals of physics beyond the standard model at hadron colliders. Calculation of high H_T events is sensitive to higher-order corrections and so high H_T data provides discrimination power between the various theoretical approaches for accounting for these contributions.

We observe significant variation in the predicted shapes of the H_T spectrum from the various theoretical approaches. SHERPA, PYTHIA, HERWIG, and ALPGEN show discrepancies in shape by $\pm 25\%$ in the one-jet bin, and up to $\pm 50\%$ in the four-jet bin. The data are significantly more precise than the spread of these predictions and can be used to improve the modeling. HEJ exhibits a good description of the data, albeit with large scale uncertainties, but the trend for NLO BLACKHAT+SHERPA (particularly noticeable in the one-jet bin) is for predictions to progressively underestimate the data as H_T increases. NLO W + n-jet calculations include n and $n+1$ parton emissions and this limitation becomes apparent when studying observables

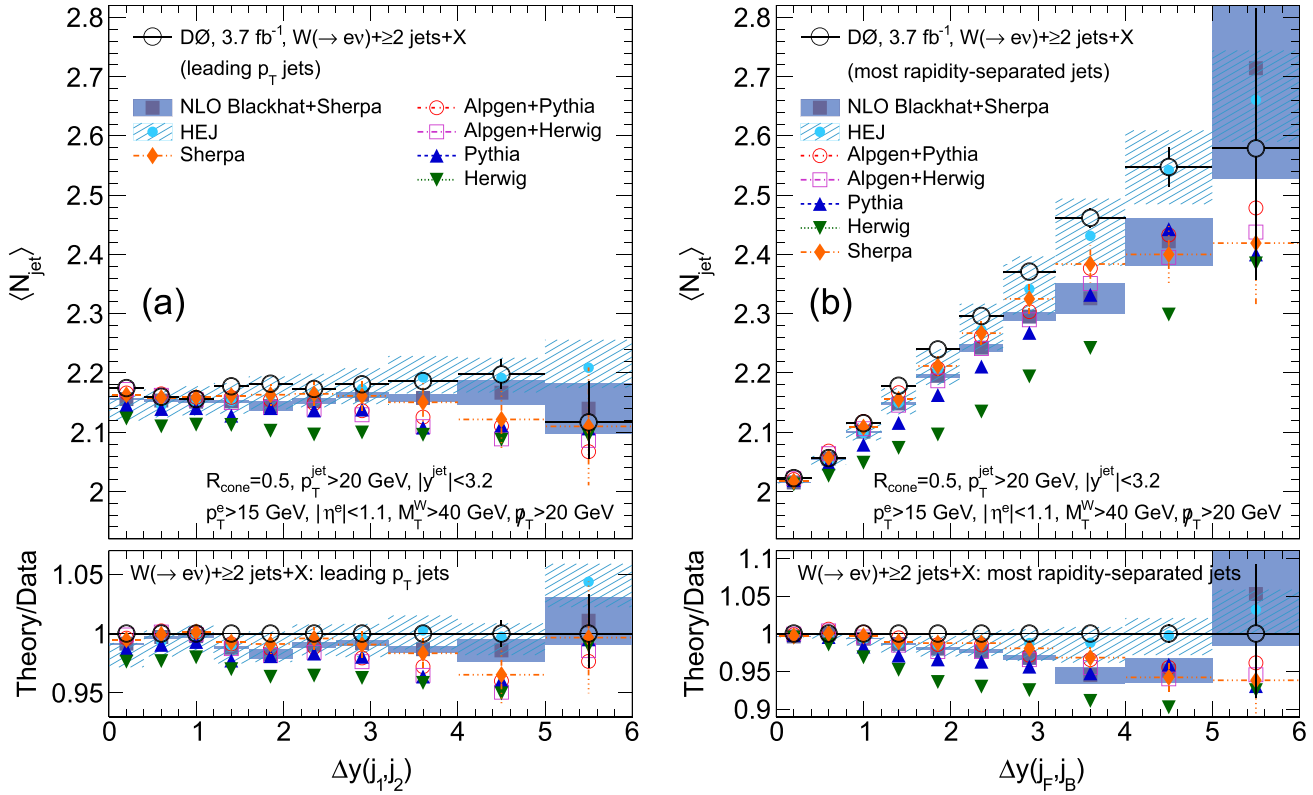


FIG. 22 (color online). Measurements of the mean number of jets in inclusive W + 2-jet events as a function of the dijet rapidity separation of the two leading jets in both (a) p_T-ordered and (b) rapidity-ordered scenarios, along with comparison to various theoretical predictions. Lower panes show theory/data comparisons for each of the configurations.

that are sensitive to higher-order contributions at high H_T, where the omission of matrix elements with $n + 2$ or more real emissions in the NLO calculation becomes apparent. Similar behavior was also observed in ATLAS $W + n$-jet data [4].

B. Mean jet multiplicities

Figures 21(a) and 21(b) show the mean number of high-p_T jets as a function of H_T in inclusive W + 1-jet and W + 2-jet events, respectively, allowing us to investigate how the jet multiplicity in $W + n$-jet events correlates with increasing boson and parton transverse energy. The lowest H_T value accessible in inclusive W + 1-jet events is 100 GeV, due to the contribution of the W invariant mass and the 20 GeV minimum transverse energy contributed by the leading jet. Likewise, the minimum H_T value in inclusive W + 2-jet events is 120 GeV. The first H_T bin in each distribution acts as a baseline for jet activity, and additional jet emissions can only contribute at $H_T > 120(140)$ GeV in the inclusive W + 1-jet (W + 2-jet) cases.

The data display a sharp rise in the mean number of jets versus H_T. We observe that high H_T events are typically high jet multiplicity events with moderate jet p_T rather than low multiplicity events with high p_T. The high H_T region is therefore particularly sensitive to higher-order corrections, so proper modeling of the jet emissions in such a region will be necessary to permit discrimination between standard model vector boson plus jets production and indications of new physics with different high H_T properties.

The BLACKHAT Collaboration uses the following prescription [57] for calculating the expected mean number of jets within a given kinematic interval in an inclusive $W + n$-jet event to improve the description beyond the standard NLO pQCD calculation:

$$\langle N_{\text{jet}} \rangle = n + (\sigma_{n+1}^{\text{NLO}} + \sigma_{n+2}^{\text{LO}})/\sigma_n^{\text{NLO}} \tag{12}$$

where σ represents the inclusive cross section for the given multiplicity. Such a definition includes all NLO corrections, plus some higher-order terms in α_s, but essentially becomes a leading-order calculation when the predicted value of $\langle N_{\text{jet}} \rangle$ approaches and exceeds $n + 1$ in an inclusive $W + n$-jet event, leading to reduced reliability in the predictions.

In inclusive one-jet events [Fig. 21(a)], parton shower approaches are unable to describe the jet emission dependence, diverging from the data at the lowest H_T at which additional high p_T (> 20 GeV) jet emissions are kinematically possible. Such predictions plateau below $\langle N_{\text{jet}} \rangle = 2$ due to the limitations of the parton shower in producing sufficient numbers of additional hard jets. Both matrix element plus matched parton shower approaches from ALPGEN+(PYTHIA/HERWIG) and SHERPA do a somewhat better job of describing the jet multiplicity increase but

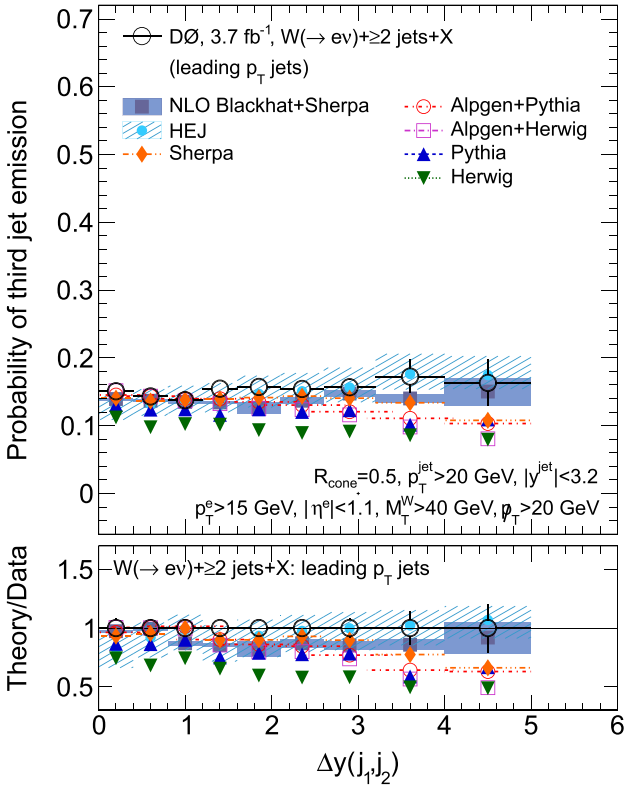


FIG. 23 (color online). Measurement of the probability of emission of a third jet in inclusive W + 2-jet events as a function of the dijet rapidity separation of the two highest-p_T jets. Comparison is made to predictions from various theoretical approaches. The lower pane shows theory/data comparisons.

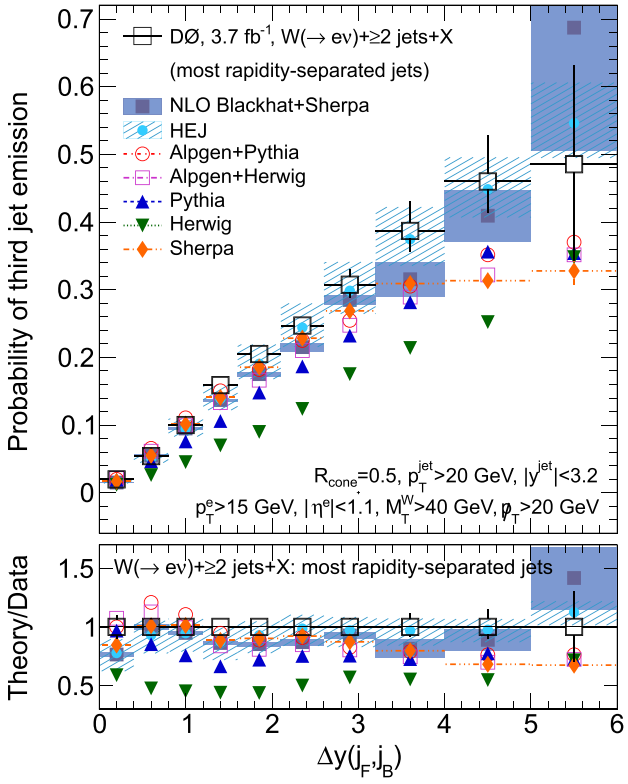


FIG. 24 (color online). Measurement of the probability of emission of a third jet in inclusive W + 2-jet events as a function of the dijet rapidity separation of the two most rapidity-separated jets (with $p_T > 20$ GeV). Comparison is made to predictions from various theoretical approaches. The lower pane shows theory/data comparisons.

again reach a maximum, at $\langle N_{\text{jet}} \rangle = 2.2$, well below the data, which reaches a maximum jet multiplicity of 2.5. In contrast, the NLO BLACKHAT+SHERPA approach is successful in describing the $\langle N_{\text{jet}} \rangle$ spectrum across the entire H_T range in inclusive one-jet events with good accuracy.

For inclusive two-jet events [Fig. 21(b)], we focus on resummation and NLO pQCD approaches, due to parton shower simulations, which contain LO matrix elements only, having reduced accuracy and less predictive power for high jet-multiplicity final states. Here, the NLO prediction again describes the data well over the full range of measured H_T. Calculations from HEJ perform well at low H_T, but underestimate the amount of high-p_T jet emission for $H_T > 250$ GeV.

We also measure the mean jet multiplicity as a function of the rapidity separation between the two highest-p_T jets and between the two most rapidity-separated jets in inclusive W + dijet events, with the results shown in Fig. 22. The mean number of jets as a function of dijet rapidity separation provides a sensitive test of high-p_T jet emission in W + n-jet events. As a function of the Δy between the two highest-p_T jets, the mean jet multiplicity is approximately constant up to rapidity spans of six units of rapidity with $\langle N_{\text{jet}} \rangle \approx 2.17$. Parton shower and matrix element matched theoretical approaches are able to describe the shape of the rapidity separation dependence although up to 5% differences in mean jet multiplicity with respect to the data are observed. Predictions from NLO BLACKHAT+SHERPA and HEJ accurately predict the uniform jet multiplicity distribution seen in the data.

In the case of the most rapidity-separated jet configuration, a strong $\langle N_{\text{jet}} \rangle$ dependence is observed with rapidity separation, in contrast to the p_T-ordered configuration, varying from $\langle N_{\text{jet}} \rangle = 2.0$ jets at small separation (where there is limited phase space for emission of a third jet with $p_T > 20$ GeV between the two forward jets) increasing steadily with rapidity separation to approximately 2.6 jets at the widest spans as shown in Fig. 22(b). This is a particularly important probe for validation of theoretical understanding of wide angle gluon emission in vector boson plus jet processes [13,56].

Unlike the $\Delta y(j_1, j_2)$ configurations, both parton shower and matrix element plus matched parton shower generators diverge from data with increasing rapidity separation and increasing average jet multiplicity. Predictions from BLACKHAT+SHERPA also show a trend for NLO pQCD to underestimate the jet multiplicity in a similar manner to

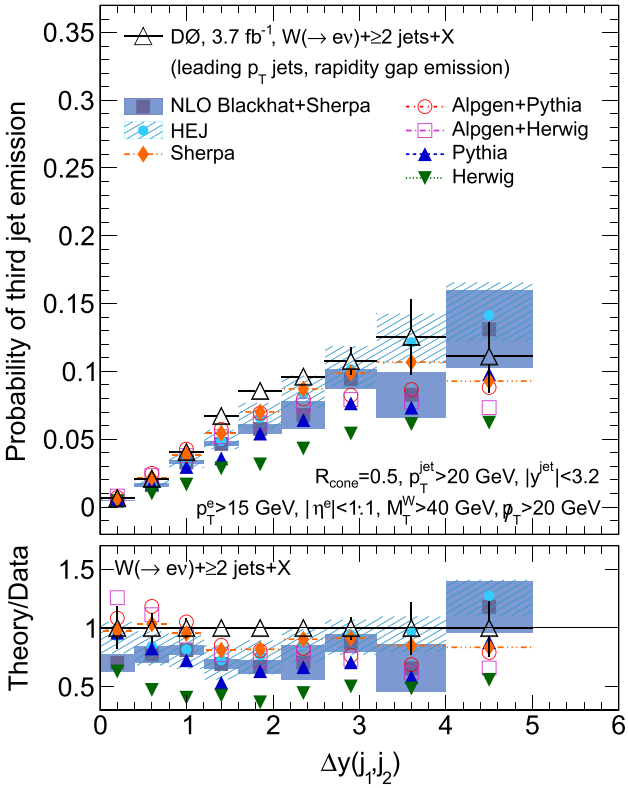


FIG. 25 (color online). Measurement of the probability of emission of a third jet in inclusive W + 2-jet events as a function of the dijet rapidity separation of the two highest-p_T jets with an additional requirement that the third jet be emitted into the rapidity interval defined by the two leading jets. Comparison is made to predictions from various theoretical approaches. The lower pane shows theory/data comparisons.

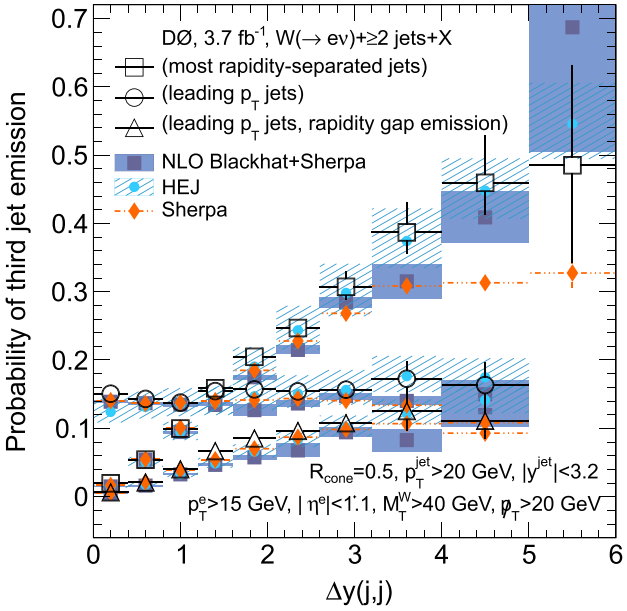


FIG. 26 (color online). Measurement of the probability of emission of a third jet in inclusive W + 2-jet events as a function of the dijet rapidity separation of the two leading jets in both p_T-ordered and rapidity-ordered scenarios, and a hybrid scenario where the rapidity separation is built from the two highest-p_T jets but the third jet is required to be emitted into the rapidity interval defined by the two leading jets. Comparison is made to predictions from NLO BALCKHAT+SHERPA, HEJ, and SHERPA.

ALPGEN and SHERPA. Resummation predictions from HEJ are able to accurately describe the jet multiplicity dependence on jet rapidity separation across the full interval studied, with high precision.

C. Jet emission probabilities/gap fraction

Figures 23–25 present measurements of the probability for a third high-p_T jet to be emitted in inclusive W + dijet events calculated as the fraction of events in the inclusive W + 2-jet sample that contain a third jet over a $p_T > 20$ GeV threshold as a function of dijet rapidity separation using:

(1) the two highest-p_T jets,
(2) the two most rapidity-separated jets ($p_T > 20$ GeV), and
(3) the two highest-p_T jets with the requirement that the third jet be emitted into the rapidity interval between the two highest-p_T jets.

The probability of emission of a third jet in inclusive W + dijet events is strongly correlated with the mean number of jets in the event presented in Figs. 21 and 22. However, with the probability observable, we specifically focus on the emission of a single additional jet beyond the two used to define the dijet rapidity interval.

The probability of third-jet emission as a function of the rapidity span between the two leading jets is approximately 15% and is shown in Fig. 23 in comparison to a variety of theoretical predictions. Both parton shower and matrix element plus parton shower matched MC programs underpredict the overall emission rate, particularly at large rapidity separations where these programs predict a drop in jet emission not supported by the data. Unlike the MC predictions that underestimate the high-p_T radiation at large rapidity separations, HEJ and NLO BLACKHAT+SHERPA approaches are able to model the constant jet emission dependence well.

As a function of the most rapidity-separated jets, a significant variation in third-jet emission probability is observed in the data. At the smallest rapidity separations, emission probabilities are $\approx 0\%$, but at the largest rapidity spans, half of all inclusive W + dijet events are found to have a third high-p_T jet present. This measurement is shown in comparison to a variety of theoretical models in Fig. 24.

The exact correlation of jet emission probability with rapidity interval is dependent on the interplay between two effects: the increasing phase space for high-p_T emission between jets versus the probability to actually emit into that rapidity interval, which decreases at large rapidity separations due to steeply falling PDFs as Bjorken

$x \to 1$. The partons at large-Δy must have high Bjorken x due to the p_T requirement; the probability for emission decreases because x is steeply falling and emission requires a higher x. There is some evidence that as we approach the largest separations studied, PDF suppression may begin to dominate over the increased phase space in the most rapidity-separated jet configuration. Proper modeling of $W + n$-jet behavior, particularly in the most rapidity-separated jet case, will be important for understanding central jet vetoes in future VBF studies, including VBF Higgs production.

As with $\langle N_{\text{jet}} \rangle$ in Fig. 22(b), parton shower and matrix element plus parton shower matched predictions underestimate the rise in jet emission probability with increased rapidity separation and plateau at a maximum probability of around 35%, as shown in Fig. 24. NLO pQCD and HEJ resummation approaches are able to describe the emission probability across the full range of study.

Third-jet emission probabilities are also presented as a function of dijet rapidity separation with an additional requirement that the third high-p_T jet be emitted into the rapidity interval between the leading two jets and compared with theoretical models in Fig. 25.

This configuration represents a conceptual hybrid between the rapidity-ordered configuration (which has the requirement that the third jet be emitted between the two most rapidity-separated jets by construction) and a p_T-ordered jet configuration where it is the highest-p_T, rather than most rapidity-separated, jets that are probed.

The three measurements of third-jet emission probability are summarized in Fig. 26 in comparison to predictions from NLO BLACKHAT+SHERPA, HEJ, and SHERPA. The probability ($\mathbb{P}$) ratio

$$\mathcal{R}_{\mathbb{P}} = \mathbb{P}(\text{third jet in rapidity interval})/\mathbb{P}(\text{third jet}) \quad (13)$$

provides information on the probability that the third-highest p_T jet in inclusive $W + 3$-jet events is emitted within the rapidity interval defined by the two highest-p_T jets. For $\Delta y \to 0$, $\mathcal{R}_{\mathbb{P}} \to 0$ as the available phase space for emission is reduced. As the rapidity interval widens, $\mathcal{R}_{\mathbb{P}} \to 1$ as the phase space for third-jet emission at larger rapidities than the two leading jets decreases. In the limit of $\Delta y \to 0$, the leading p_T jet rapidity interval configuration is bounded by that of the rapidity-ordered jet results, and in the wide-angle limit it is bounded by the emission probability of the p_T-ordered jet configuration without a rapidity interval requirement. As such, emission probabilities again start at $\approx 0\%$ at small jet separation and rise quickly with increasing jet spans, but are limited to a plateau of around 15% at the largest rapidity spans. PYTHIA and HERWIG in particular have trouble modeling this observable, both in overall jet emission rate and in the dependence on Δy. ALPGEN and SHERPA provide improved descriptions of the jet emission probability in this configuration, but still predict a lower emission rate at larger rapidity spans than observed in data. NLO pQCD predictions systematically underestimate the total emission rate by about 30%, but otherwise describe the emission rate dependence on rapidity interval well across the full range. Resummation predictions from HEJ are best able to describe both the rate and shape across the full rapidity range.

The probability, $\mathbb{P}$, of additional jet emission as a function of dijet rapidity separation can be reinterpreted as a "gap fraction" $\mathbb{F}$, where $\mathbb{F} = 1 - \mathbb{P}$. This gap fraction is defined as the fraction of inclusive $W +$ dijet events that do not have an additional jet with a transverse momentum larger than a given veto scale (in this analysis, 20 GeV) within the rapidity interval defined by the rapidities of the two highest-p_T (or most rapidity-separated) jets in the event.

VIII. SUMMARY

This paper presents measurements of the production rates and properties of $W(\to e\nu) + n$-jet production (for $n \geq 1, 2, 3, 4$) in $p\bar{p}$ collisions at a center-of-mass energy of 1.96 TeV using 3.7 fb^{-1} of D0 experimental data. These measurements provide the most comprehensive set of measurements of the $W +$ jets processes made to date. They are important in their own right as precise studies of important QCD processes that will guide future theoretical refinements, and are essential for understanding backgrounds for searches for new rare processes.

Differential cross sections, normalized by the total W boson cross section and fully corrected for detector effects, are presented for various inclusive jet multiplicities as a function of the jet rapidities; electron p_T and pseudorapidity; rapidity separations of the first-, second-, and third-highest p_T jets in the event; rapidity separations of the most rapidity-separated jets in the event (for jets with $p_T >$ 20 GeV); azimuthal angle separations between jets; the angular separation between the two leading jets in y-ϕ space; W boson p_T; dijet system p_T and invariant mass; and H_T (the scalar sum of the jet and W boson transverse energies). Many of these observables are studied here for the first time in $W + n$-jet events or substantially improve on the precision of existing measurements. These measurements complement previous measurements [2] of the total inclusive $W + n$-jet production cross sections (for $n \geq 1, 2, 3, 4$) and differential cross sections of the nth-jet transverse momenta performed in the same phase space as measurements presented here.

Additionally, we present measurements of the evolution of mean jet multiplicities of $W + n$-jet events as a function of H_T in the inclusive one- and two-jet multiplicity bins; and within the region bounded by the dijet rapidity defined by the two jets that are the highest-p_T jets, and also between the most rapidity-separated jets with a p_T threshold of 20 GeV. The probability of a third high-p_T jet emission in $W +$ dijet events is also measured as a function of the dijet rapidity separation in three configurations: the

first, where the rapidity separation is defined by the two highest-p_T jets in the event; the second, as previously but where any third jet is required to be emitted into the rapidity interval defined by the other two jets; and finally, where the rapidity separation is defined by the two most rapidity-separated jets. These results are recast as a measure of the gap fraction in $W + n$-jet events, with a veto on additional emissions with $p_T > 20$ GeV.

The presented measurements can be used for constraining the modeling of QCD radiation between the two jets, for understanding the efficacy of a central jet veto [58] used for discriminating Higgs boson plus dijet events produced through vector boson fusion from standard model backgrounds and for subsequent study of the Higgs boson properties [59]. In addition, they can contribute to improved understanding of vector boson plus jet contributions to diverse topics such as studies of $W_L W_L \to W_L W_L$ scattering [60] and searches for MSSM signatures through VBF production [61]. Measurements of the gap fraction in vector boson plus dijet events complement existing measurements in inclusive dijet [62] events.

Comparisons of the experimental data are made to predictions from a variety of theoretical approaches. Over most of the phase space in which the measurements are presented, experimental uncertainties are smaller than the theoretical uncertainties on NLO BLACKHAT+SHERPA, and on HEJ resummation predictions. The predictions from various Monte Carlo programs are found to have significant variations between each other, greater in magnitude than the experimental uncertainties, and thus these data can be used to improve the modeling of $W + n$-jet production and the emission of QCD radiation in such event generators.

ACKNOWLEDGMENTS

The authors would like to thank the Blackhat Collaboration, in particular Fernando Febres Cordero, for providing NLO BLACKHAT+SHERPA predictions, and also Daniel Maitre and Lance Dixon for valuable conversations. We also thank Stefan Hochë for SHERPA predictions; Tuomas Hapola, Jeppe Andersen, and Jennifer Smillie for HEJ predictions; and Vato Kartvelishvili for useful discussions and advice on use of the GURU unfolding program. We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); MON, NRC KI and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); NRF (Korea); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China).

[1] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **77**, 011108 (2008).
[2] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **705**, 200 (2011).
[3] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **698**, 325 (2011).
[4] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **85**, 92002 (2012).
[5] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. 01 (2012) 10.
[6] S. Abachi *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **338**, 185 (1994); V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A **565**, 463 (2006).
[7] M. Abolins *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **584**, 75 (2008).
[8] R. Angstadt *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **622**, 298 (2010).
[9] C. Berger, Z. Bern, L. Dixon, F. Cordero, D. Forde, T. Gleisberg, H. Ita, D. Kosower, and D. Maître, Phys. Rev. D **80**, 74036 (2009).
[10] R. K. Ellis, K. Melnikov, and G. Zanderighi, Phys. Rev. D **80**, 94002 (2009).
[11] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, J. High Energy Phys. 02 (2009) 007.
[12] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour, J. High Energy Phys. 06 (2005) 034; J. Forshaw, J. Keates, and S. Marzani, J. High Energy Phys. 07 (2009) 023.
[13] J. R. Andersen and J. M. Smillie, J. High Energy Phys. 01 (2010) 39.
[14] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **669**, 278 (2008).
[15] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **693**, 522 (2010).
[16] C. Buttar *et al.*, arXiv:0803.0678.
[17] A. Hocker and V. Kartvelishvili, Nucl. Instrum. Methods Phys. Res., Sect. A **372**, 469 (1996).
[18] T. Andeen *et al.*, Report No. FERMILAB-TM-2365, FERMILAB-TM-2365, 2007.
[19] B. C. K. Casey *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **698**, 208 (2013).
[20] G. C. Blazey *et al.*, Report No. FERMILAB-PUB-00/297, 2000.
[21] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **85**, 52006 (2012).
[22] J.-R. Vlimant, Report No. FERMILAB-THESIS-2005-52.

[23] Rapidity is defined as $y = -\ln[(E + p_Z)/(E - p_Z)]$, where E is the energy and p_Z is the momentum component along the proton beam direction. ϕ is defined as the azimuthal angle in the plane transverse to the proton beam direction.

[24] R. Brun and F. Carminati, CERN Program Library Long Writeup W5013, 1993 (unpublished).

[25] M. L Mangano, F. Piccinini, A. D Polosa, M. Moretti, and R. Pittau, J. High Energy Phys. 07 (2003) 001.

[26] T. Sjöstrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.

[27] T. Affolder *et al.* (CDF Collaboration), Phys. Rev. D **65**, 92002 (2002).

[28] S. Moch and P. Uwer, Phys. Rev. D **78**, 34003 (2008).

[29] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).

[30] J. M. Campbell and R. K. Ellis, Nucl. Phys. B, Proc. Suppl. **205–206**, 10 (2010).

[31] 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); E. E. Boos, V. E. Bunichev, L. V. Dudko, V. I. Savrin, and A. V. Sherstnev, Yad. Fiz. **69**, 1352 (2006) [Phys. At. Nucl. **69**, 1317 (2006)].

[32] J. M. Campbell and R. K. Ellis, Phys. Rev. D **65**, 113007 (2002); J. M. Campbell, R. K. Ellis, and D. L. Rainwater, Phys. Rev. D **68**, 94021 (2003).

[33] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **100**, 102002 (2008).

[34] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, Comput. Phys. Commun. **182**, 2388 (2011).

[35] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **84**, 12008 (2011).

[36] See HEPDATA record for tabulated results for measured $W + n -$ jet observables, theoretical predictions and nonperturbative corrections applied to the theory: http://hepdata.cedar.ac.uk/view/ins1221252.

[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.88.092001 for tabulated results for measured $W + n$-jet observables, theoretical predictions and nonperturbative corrections applied to the theory.

[38] H. L. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. Olness, J. F. Owens, J. Pumplin, and W. K. Tung (CTEQ Collaboration), Eur. Phys. J. C **12**, 375 (2000).

[39] G. Marchesini, B. R. Webber, G. Abbiendi, I. G. Knowles, M. H. Seymour, and L. Stanco, Comput. Phys. Commun. **67**, 465 (1992); 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.

[40] J. Pumplin, D. Robert Stump, J. Huston, H.-L. Lai, P. Nadolsky, and W.-K. Tung, J. High Energy Phys. 07 (2002) 012; 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.

[41] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, Z. Phys. C **72**, 637 (1996).

[42] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, J. High Energy Phys. 02 (2009) 007.

[43] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, Phys. Rev. D **82**, 74024 (2010).

[44] S. Höche, F. Krauss, S. Schumann, and F. Siegert, J. High Energy Phys. 05 (2009) 053.

[45] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **82**, 34001 (2010).

[46] J. R. Andersen and J. M. Smillie, Phys. Rev. D **81**, 114021 (2010); Nucl. Phys. B, Proc. Suppl. **205–206**, 205 (2010); J. High Energy Phys. 06 (2011) 10; J. R. Andersen, T. Hapola, and J. M. Smillie, J. High Energy Phys. 09 (2012) 047.

[47] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maître, Phys. Rev. Lett. **106**, 92001 (2011) (predictions for $p\bar{p}$ at $\sqrt{s} = 1.96$ TeV from private communications).

[48] C. Berger, Z. Bern, L. Dixon, F. Cordero, D. Forde, T. Gleisberg, H. Ita, D. Kosower, and D. Maître, Phys. Rev. Lett. **102**, 222001 (2009); arXiv:0905.2735; Phys. Rev. D **80**, 074036 (2009).

[49] R. Keith Ellis, W. T Giele, Z. Kunszt, K. Melnikov, and G. Zanderighi, J. High Energy Phys. 01 (2009) 012; W. T. Giele and G. Zanderighi, J. High Energy Phys. 06 (2008) 038.

[50] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Eur. Phys. J. C **63**, 189 (2009).

[51] G. P. Salam and G. Soyez, J. High Energy Phys. 05 (2007) 086.

[52] S. Höche, F. Krauss, M. Schonherr, and F. Siegert, J. High Energy Phys. 09 (2012) 49.

[53] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, Phys. Rep. **97**, 31 (1983).

[54] G. D. Lafferty and T. R. Wyatt, Nucl. Instrum. Methods Phys. Res., Sect. A **355**, 541 (1995).

[55] J. D. Bjorken and E. A. Paschos, Phys. Rev. **185**, 1975 (1969).

[56] J. R. Andersen, V. Del Duca, F. Maltoni, and W. J. Stirling, J. High Energy Phys. 05 (2001) 048.

[57] Private communications with BLACKHAT+SHERPA authors.

[58] D. L. Rainwater, R. Szalapski, and D. Zeppenfeld, Phys. Rev. D **54**, 6680 (1996); D. L. Rainwater, D. Zeppenfeld, and K. Hagiwara, Phys. Rev. D **59**, 014037 (1998).

[59] M. Duhrssen, S. Heinemeyer, H. Logan, D. Rainwater, G. Weiglein, and D. Zeppenfeld, Phys. Rev. D **70**, 113009 (2004); B. E. Cox, J. R. Forshaw, and A. D. Pilkington, Phys. Lett. B **696**, 87 (2011).

[60] J. M. Butterworth, B. E. Cox, and J. R. Forshaw, Phys. Rev. D **65**, 096014 (2002).

[61] B. Dutta, A. Gurrola, W. Johns, T. Kamon, P. Sheldon, and K. Sinha, Phys. Rev. D **87**, 035029 (2013).

[62] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 09 (2011) 53.