Studies of high-transverse momentum jet substructure and top quarks produced in 1.96 TeV proton-antiproton collisions

T. Aaltonen *et al.* (CDF Collaboration)

Phys. Rev. D **91**, 032006 — Published 19 February 2015

DOI: 10.1103/PhysRevD.91.032006

Studies of high-transverse momentum jet substructure and top quarks produced in 1.96 TeV proton-antiproton collisions

T. Aaltonen,[21] R. Alon[tt],[31] S. Amerio[jj],[40] D. Amidei,[32] A. Anastassov[v],[15] A. Annovi,[17] J. Antos,[12] G. Apollinari,[15] J.A. Appel,[15] T. Arisawa,[53] A. Artikov,[13] J. Asaadi,[48] W. Ashmanskas,[15] B. Auerbach,[2] A. Aurisano,[48] F. Azfar,[39] W. Badgett,[15] T. Bae,[25] A. Barbaro-Galtieri,[26] V.E. Barnes,[44] B.A. Barnett,[23] P. Barria[ll],[42] P. Bartos,[12] M. Bauce[jj],[40] F. Bedeschi,[42] S. Behari,[15] G. Bellettini[kk],[42] J. Bellinger,[55] D. Benjamin,[14] A. Beretvas,[15] A. Bhatti,[46] K.R. Bland,[5] B. Blumenfeld,[23] A. Bocci,[14] A. Bodek,[45] D. Bortoletto,[44] J. Boudreau,[43] A. Boveia,[11] L. Brigliadori[ii],[6] C. Bromberg,[33] E. Brucken,[21] J. Budagov,[13] H.S. Budd,[45] K. Burkett,[15] G. Busetto[jj],[40] P. Bussey,[19] P. Butti[kk],[42] A. Buzatu,[19] A. Calamba,[10] S. Camarda,[4] M. Campanelli,[28] F. Canelli[cc],[11] B. Carls,[22] D. Carlsmith,[55] R. Carosi,[42] S. Carrillo[l],[16] B. Casal[j],[9] M. Casarsa,[49] A. Castro[ii],[6] P. Catastini,[20] D. Cauz[qqrr],[49] V. Cavaliere,[22] A. Cerri[e],[26] L. Cerrito[q],[28] Y.C. Chen,[1] M. Chertok,[7] G. Chiarelli,[42] G. Chlachidze,[15] K. Cho,[25] D. Chokheli,[13] A. Clark,[18] C. Clarke,[54] M.E. Convery,[15] J. Conway,[7] M. Corbo[y],[15] M. Cordelli,[17] C.A. Cox,[7] D.J. Cox,[7] M. Cremonesi,[42] D. Cruz,[48] J. Cuevas[x],[9] R. Culbertson,[15] N. d'Ascenzo[u],[15] M. Datta[ff],[15] P. de Barbaro,[45] L. Demortier,[46] M. Deninno,[6] M. D'Errico[jj],[40] F. Devoto,[21] A. Di Canto[kk],[42] B. Di Ruzza[p],[15] J.R. Dittmann,[5] S. Donati[kk],[42] M. D'Onofrio,[27] M. Dorigo[ss],[49] A. Driutti[qqrr],[49] E. Duchovni[tt],[31] K. Ebina,[53] R. Edgar,[32] A. Elagin,[48] R. Erbacher,[7] S. Errede,[22] B. Esham,[22] S. Farrington,[39] J.P. Fernández Ramos,[29] R. Field,[16] G. Flanagan[s],[15] R. Forrest,[7] M. Franklin,[20] J.C. Freeman,[15] H. Frisch,[11] Y. Funakoshi,[53] C. Galloni[kk],[42] A.F. Garfinkel,[44] P. Garosi[ll],[42] H. Gerberich,[22] E. Gerchtein,[15] S. Giagu,[47] V. Giakoumopoulou,[3] K. Gibson,[43] C.M. Ginsburg,[15] N. Giokaris,[3] P. Giromini,[17] V. Glagolev,[13] D. Glenzinski,[15] M. Gold,[35] D. Goldin,[48] A. Golossanov,[15] G. Gomez,[9] G. Gomez-Ceballos,[30] M. Goncharov,[30] O. González López,[29] I. Gorelov,[35] A.T. Goshaw,[14] K. Goulianos,[46] E. Gramellini,[6] C. Grosso-Pilcher,[11] R.C. Group,[52,15] J. Guimaraes da Costa,[20] S.R. Hahn,[15] J.Y. Han,[45] F. Happacher,[17] K. Hara,[50] M. Hare,[51] R.F. Harr,[54] T. Harrington-Taber[m],[15] K. Hatakeyama,[5] C. Hays,[39] J. Heinrich,[41] M. Herndon,[55] A. Hocker,[15] Z. Hong,[48] W. Hopkins[f],[15] S. Hou,[1] R.E. Hughes,[36] U. Husemann,[56] M. Hussein[aa],[33] J. Huston,[33] G. Introzzi[nnoo],[42] M. Iori[pp],[47] A. Ivanov[o],[7] E. James,[15] D. Jang,[10] B. Jayatilaka,[15] E.J. Jeon,[25]

S. Jindariani,[15] M. Jones,[44] K.K. Joo,[25] S.Y. Jun,[10] T.R. Junk,[15] M. Kambeitz,[24] T. Kamon,[25,48]
P.E. Karchin,[54] A. Kasmi,[5] Y. Kato[n],[38] W. Ketchum[gg],[11] J. Keung,[41] B. Kilminster[cc],[15]
D.H. Kim,[25] H.S. Kim,[25] J.E. Kim,[25] M.J. Kim,[17] S.H. Kim,[50] S.B. Kim,[25] Y.J. Kim,[25]
Y.K. Kim,[11] N. Kimura,[53] M. Kirby,[15] K. Knoepfel,[15] K. Kondo,[53,*] D.J. Kong,[25]
J. Konigsberg,[16] A.V. Kotwal,[14] M. Kreps,[24] J. Kroll,[41] M. Kruse,[14] T. Kuhr,[24] M. Kurata,[50]
A.T. Laasanen,[44] S. Lammel,[15] M. Lancaster,[28] K. Lannon[w],[36] G. Latino[ll],[42] H.S. Lee,[25]
J.S. Lee,[25] S. Leo,[42] S. Leone,[42] J.D. Lewis,[15] A. Limosani[r],[14] E. Lipeles,[41] A. Lister[a],[18]
H. Liu,[52] Q. Liu,[44] T. Liu,[15] S. Lockwitz,[56] A. Loginov,[56] D. Lucchesi[jj],[40] A. Lucà,[17] J. Lueck,[24]
P. Lujan,[26] P. Lukens,[15] G. Lungu,[46] J. Lys,[26] R. Lysak[d],[12] R. Madrak,[15] P. Maestro[ll],[42]
S. Malik,[46] G. Manca[b],[27] A. Manousakis-Katsikakis,[3] L. Marchese[hh],[6] F. Margaroli,[47]
P. Marino[mm],[42] K. Matera,[22] M.E. Mattson,[54] A. Mazzacane,[15] P. Mazzanti,[6] R. McNulty[i],[27]
A. Mehta,[27] P. Mehtala,[21] C. Mesropian,[46] T. Miao,[15] D. Mietlicki,[32] A. Mitra,[1] H. Miyake,[50]
S. Moed,[15] N. Moggi,[6] C.S. Moon[y],[15] R. Moore[ddee],[15] M.J. Morello[mm],[42] A. Mukherjee,[15]
Th. Muller,[24] P. Murat,[15] M. Mussini[ii],[6] J. Nachtman[m],[15] Y. Nagai,[50] J. Naganoma,[53]
I. Nakano,[37] A. Napier,[51] J. Nett,[48] C. Neu,[52] T. Nigmanov,[43] L. Nodulman,[2] S.Y. Noh,[25]
O. Norniella,[22] L. Oakes,[39] S.H. Oh,[14] Y.D. Oh,[25] I. Oksuzian,[52] T. Okusawa,[38] R. Orava,[21]
L. Ortolan,[4] C. Pagliarone,[49] E. Palencia[e],[9] P. Palni,[35] V. Papadimitriou,[15] W. Parker,[55]
G. Pauletta[qqrr],[49] M. Paulini,[10] C. Paus,[30] G. Perez[tt],[31] T.J. Phillips,[14] G. Piacentino,[42]
E. Pianori,[41] J. Pilot,[7] K. Pitts,[22] C. Plager,[8] L. Pondrom,[55] S. Poprocki[f],[15] K. Potamianos,[26]
A. Pranko,[26] F. Prokoshin[z],[13] F. Ptohos[g],[17] G. Punzi[kk],[42] I. Redondo Fernández,[29] P. Renton,[39]
M. Rescigno,[47] F. Rimondi,[6,*] L. Ristori,[42,15] A. Robson,[19] T. Rodriguez,[41] S. Rolli[h],[51]
M. Ronzani[kk],[42] R. Roser,[15] J.L. Rosner,[11] F. Ruffini[ll],[42] A. Ruiz,[9] J. Russ,[10] V. Rusu,[15]
W.K. Sakumoto,[45] Y. Sakurai,[53] L. Santi[qqrr],[49] K. Sato,[50] V. Saveliev[u],[15] A. Savoy-Navarro[y],[15]
P. Schlabach,[15] E.E. Schmidt,[15] T. Schwarz,[32] L. Scodellaro,[9] F. Scuri,[42] S. Seidel,[35]
Y. Seiya,[38] A. Semenov,[13] F. Sforza[kk],[42] S.Z. Shalhout,[7] T. Shears,[27] P.F. Shepard,[43]
M. Shimojima[t],[50] M. Shochet,[11] I. Shreyber-Tecker,[34] A. Simonenko,[13] P. Sinervo,[31] K. Sliwa,[51]
J.R. Smith,[7] F.D. Snider,[15] H. Song,[43] V. Sorin,[4] R. St. Denis,[19,*] M. Stancari,[15] D. Stentz[v],[15]
J. Strologas,[35] Y. Sudo,[50] A. Sukhanov,[15] I. Suslov,[13] K. Takemasa,[50] Y. Takeuchi,[50]
J. Tang,[11] M. Tecchio,[32] P.K. Teng,[1] J. Thom[f],[15] E. Thomson,[41] V. Thukral,[48] D. Toback,[48]
S. Tokar,[12] K. Tollefson,[33] T. Tomura,[50] D. Tonelli[e],[15] S. Torre,[17] D. Torretta,[15] P. Totaro,[40]

M. Trovato[mm],[42] F. Ukegawa,[50] S. Uozumi,[25] F. Vázquez[l],[16] G. Velev,[15] C. Vellidis,[15] C. Vernieri[mm],[42] M. Vidal,[44] R. Vilar,[9] J. Vizán[bb],[9] M. Vogel,[35] G. Volpi,[17] P. Wagner,[41] R. Wallny[j],[15] S.M. Wang,[1] D. Waters,[28] W.C. Wester III,[15] D. Whiteson[c],[41] A.B. Wicklund,[2] S. Wilbur,[7] H.H. Williams,[41] J.S. Wilson,[32] P. Wilson,[15] B.L. Winer,[36] P. Wittich[f],[15] S. Wolbers,[15] H. Wolfe,[36] T. Wright,[32] X. Wu,[18] Z. Wu,[5] K. Yamamoto,[38] D. Yamato,[38] T. Yang,[15] U.K. Yang,[25] Y.C. Yang,[25] W.-M. Yao,[26] G.P. Yeh,[15] K. Yi[m],[15] J. Yoh,[15] K. Yorita,[53] T. Yoshida[k],[38] G.B. Yu,[14] I. Yu,[25] A.M. Zanetti,[49] Y. Zeng,[14] C. Zhou,[14] and S. Zucchelli[ii]6

(CDF Collaboration)[†]

[1]*Institute of Physics, Academia Sinica,*
Taipei, Taiwan 11529, Republic of China

[2]*Argonne National Laboratory, Argonne, Illinois 60439, USA*

[3]*University of Athens, 157 71 Athens, Greece*

[4]*Institut de Fisica d'Altes Energies, ICREA,*
Universitat Autonoma de Barcelona, E-08193, Bellaterra (Barcelona), Spain

[5]*Baylor University, Waco, Texas 76798, USA*

[6]*Istituto Nazionale di Fisica Nucleare Bologna,*
[ii]*University of Bologna, I-40127 Bologna, Italy*

[7]*University of California, Davis, Davis, California 95616, USA*

[8]*University of California, Los Angeles, Los Angeles, California 90024, USA*

[9]*Instituto de Fisica de Cantabria, CSIC-University of Cantabria, 39005 Santander, Spain*

[10]*Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*

[11]*Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA*

[12]*Comenius University, 842 48 Bratislava,*
Slovakia; Institute of Experimental Physics, 040 01 Kosice, Slovakia

[13]*Joint Institute for Nuclear Research, RU-141980 Dubna, Russia*

[14]*Duke University, Durham, North Carolina 27708, USA*

[15]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*

[16]*University of Florida, Gainesville, Florida 32611, USA*

[17]*Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, I-00044 Frascati, Italy*

[18]*University of Geneva, CH-1211 Geneva 4, Switzerland*

[19]*Glasgow University, Glasgow G12 8QQ, United Kingdom*

[20]*Harvard University, Cambridge, Massachusetts 02138, USA*
[21]*Division of High Energy Physics, Department of Physics,*
University of Helsinki, FIN-00014, Helsinki,
Finland; Helsinki Institute of Physics, FIN-00014, Helsinki, Finland
[22]*University of Illinois, Urbana, Illinois 61801, USA*
[23]*The Johns Hopkins University, Baltimore, Maryland 21218, USA*
[24]*Institut für Experimentelle Kernphysik,*
Karlsruhe Institute of Technology, D-76131 Karlsruhe, Germany
[25]*Center for High Energy Physics: Kyungpook National University,*
Daegu 702-701, Korea; Seoul National University, Seoul 151-742,
Korea; Sungkyunkwan University, Suwon 440-746,
Korea; Korea Institute of Science and Technology Information,
Daejeon 305-806, Korea; Chonnam National University,
Gwangju 500-757, Korea; Chonbuk National University, Jeonju 561-756,
Korea; Ewha Womans University, Seoul, 120-750, Korea
[26]*Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[27]*University of Liverpool, Liverpool L69 7ZE, United Kingdom*
[28]*University College London, London WC1E 6BT, United Kingdom*
[29]*Centro de Investigaciones Energeticas*
Medioambientales y Tecnologicas, E-28040 Madrid, Spain
[30]*Massachusetts Institute of Technology,*
Cambridge, Massachusetts 02139, USA
[31]*Institute of Particle Physics: McGill University, Montréal,*
Québec, Canada H3A 2T8; Simon Fraser University, Burnaby,
British Columbia, Canada V5A 1S6; University of Toronto,
Toronto, Ontario, Canada M5S 1A7; and TRIUMF,
Vancouver, British Columbia, Canada V6T 2A3
[32]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[33]*Michigan State University, East Lansing, Michigan 48824, USA*
[34]*Institution for Theoretical and Experimental Physics, ITEP, Moscow 117259, Russia*
[35]*University of New Mexico, Albuquerque, New Mexico 87131, USA*
[36]*The Ohio State University, Columbus, Ohio 43210, USA*

[37]*Okayama University, Okayama 700-8530, Japan*

[38]*Osaka City University, Osaka 558-8585, Japan*

[39]*University of Oxford, Oxford OX1 3RH, United Kingdom*

[40]*Istituto Nazionale di Fisica Nucleare, Sezione di Padova,*
[jj]*University of Padova, I-35131 Padova, Italy*

[41]*University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*

[42]*Istituto Nazionale di Fisica Nucleare Pisa,*
[kk]*University of Pisa,* [ll]*University of Siena,*
[mm]*Scuola Normale Superiore, I-56127 Pisa,*
Italy, [nn]*INFN Pavia, I-27100 Pavia, Italy,*
[oo]*University of Pavia, I-27100 Pavia, Italy*

[43]*University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA*

[44]*Purdue University, West Lafayette, Indiana 47907, USA*

[45]*University of Rochester, Rochester, New York 14627, USA*

[46]*The Rockefeller University, New York, New York 10065, USA*

[47]*Istituto Nazionale di Fisica Nucleare, Sezione di Roma 1,*
[pp]*Sapienza Università di Roma, I-00185 Roma, Italy*

[48]*Mitchell Institute for Fundamental Physics and Astronomy,*
Texas A&M University, College Station, Texas 77843, USA

[49]*Istituto Nazionale di Fisica Nucleare Trieste,*
[qq]*Gruppo Collegato di Udine,* [rr]*University of Udine,*
I-33100 Udine, Italy, [ss]*University of Trieste, I-34127 Trieste, Italy*

[50]*University of Tsukuba, Tsukuba, Ibaraki 305, Japan*

[51]*Tufts University, Medford, Massachusetts 02155, USA*

[52]*University of Virginia, Charlottesville, Virginia 22906, USA*

[53]*Waseda University, Tokyo 169, Japan*

[54]*Wayne State University, Detroit, Michigan 48201, USA*

[55]*University of Wisconsin, Madison, Wisconsin 53706, USA*

[56]*Yale University, New Haven, Connecticut 06520, USA*

Abstract

Results of a study of the substructure of the highest transverse momentum (p_T) jets observed by the CDF collaboration are presented. Events containing at least one jet with $p_T > 400$ GeV/c in a sample corresponding to an integrated luminosity of 5.95 fb^{-1}, collected in 1.96 TeV proton-antiproton collisions at the Fermilab Tevatron collider, are selected. A study of the jet mass, angularity, and planar-flow distributions is presented, and the measurements are compared with predictions of perturbative quantum chromodynamics. A search for boosted top-quark production is also described, leading to a 95% confidence level upper limit of 38 fb on the production cross section of top quarks with $p_T > 400$ GeV/c.

PACS numbers: 12.38.Qk, 13.87.-a, 14.65.Ha, 12.38.Aw

* Deceased

† With visitors from [a]University of British Columbia, Vancouver, BC V6T 1Z1, Canada, [b]Istituto Nazionale di Fisica Nucleare, Sezione di Cagliari, 09042 Monserrato (Cagliari), Italy, [c]University of California Irvine, Irvine, CA 92697, USA, [d]Institute of Physics, Academy of Sciences of the Czech Republic, 182 21, Czech Republic, [e]CERN, CH-1211 Geneva, Switzerland, [f]Cornell University, Ithaca, NY 14853, USA, [g]University of Cyprus, Nicosia CY-1678, Cyprus, [h]Office of Science, U.S. Department of Energy, Washington, DC 20585, USA, [i]University College Dublin, Dublin 4, Ireland, [j]ETH, 8092 Zürich, Switzerland, [k]University of Fukui, Fukui City, Fukui Prefecture, Japan 910-0017, [l]Universidad Iberoamericana, Lomas de Santa Fe, México, C.P. 01219, Distrito Federal, [m]University of Iowa, Iowa City, IA 52242, USA, [n]Kinki University, Higashi-Osaka City, Japan 577-8502, [o]Kansas State University, Manhattan, KS 66506, USA, [p]Brookhaven National Laboratory, Upton, NY 11973, USA, [q]Queen Mary, University of London, London, E1 4NS, United Kingdom, [r]University of Melbourne, Victoria 3010, Australia, [s]Muons, Inc., Batavia, IL 60510, USA, [t]Nagasaki Institute of Applied Science, Nagasaki 851-0193, Japan, [u]National Research Nuclear University, Moscow 115409, Russia, [v]Northwestern University, Evanston, IL 60208, USA, [w]University of Notre Dame, Notre Dame, IN 46556, USA, [x]Universidad de Oviedo, E-33007 Oviedo, Spain, [y]CNRS-IN2P3, Paris, F-75205 France, [z]Universidad Tecnica Federico Santa Maria, 110v Valparaiso, Chile, [aa]The University of Jordan, Amman 11942, Jordan, [bb]Universite catholique de Louvain, 1348 Louvain-La-Neuve, Belgium, [cc]University of Zürich, 8006 Zürich, Switzerland, [dd]Massachusetts General Hospital, Boston, MA 02114 USA, [ee]Harvard Medical School, Boston, MA 02114 USA, [ff]Hampton University, Hampton, VA 23668, USA, [gg]Los Alamos National Laboratory, Los Alamos, NM 87544, USA, [hh]Università degli Studi di Napoli Federico I, I-80138 Napoli, Italy [tt]Weizmann Institute of Science, Rehovot, Israel

I. INTRODUCTION

A. Motivation

The observation and study of high-transverse momentum (p_T) jets produced via quantum chromodynamics (QCD) in hadron-hadron interactions provides an important test of perturbative QCD (pQCD) [1]. The study of the most massive jets gives insight into the parton showering mechanism and assists in tuning of Monte Carlo (MC) event generators (see, e.g., [2–4] for recent reviews). Furthermore, jets with masses in excess of 100 GeV/c^2 are an important background for Higgs boson searches [5–7] and appear in final states of various beyond-the-standard-model physics processes [8–14]. Particularly relevant is the case where the decay of a heavy hypothetical resonance produces high-p_T top quarks that decay hadronically. In such cases, the daughter products can be observed as a pair of massive jets. Other sources of massive jets include the production of highly-boosted W, Z, and Higgs bosons.

We report a study of the substructure of jets with $p_T >$ 400 GeV/c produced in proton-antiproton ($p\bar{p}$) collisions at $\sqrt{s} = 1.96$ TeV at the Fermilab Tevatron and recorded by the CDF II detector. We also report a search for high-p_T production of top quarks using the same data sample and the techniques developed in the substructure analysis. This article describes in more detail the substructure analysis reported earlier [15].

Jets are reconstructed as collimated collections of high-energy particles that are identified through the use of a clustering algorithm that groups the particles into a single *jet* cluster [16]. The properties of the jet, such as its momentum and mass, are then derived from the constituents of the cluster using a recombination scheme. In this study, the jet constituents are energy deposits observed in a segmented calorimeter and the four-momentum of the jet is the standard four-vector sum of the constituents.

Earlier studies of the substructure of high-p_T jets produced at the Fermilab Tevatron Collider have been limited to jets with $p_T < 400$ GeV/c [17, 18]. More recently, jet studies have been reported by experiments at the Large Hadron Collider (LHC) [19–25], where studies of their substructure have been extended to jets with $p_T \lesssim 800$ GeV/c. Similarly, studies of top-quark production at the Tevatron have been limited to top quarks with $p_T < 300$ GeV/c [26–28]. The large data samples collected by the CDF II detector at the Fermilab $p\bar{p}$ Tevatron Collider permit study of jets with p_T greater than 400 GeV/c and their internal structure. At the same time, the-

oretical progress has been made in the understanding of the production of massive jets, and the differential top-quark pair ($t\bar{t}$) production cross section as a function of p_T is now known up to approximate next-to-next-to-leading-order (NNLO) [29, 30] and full NNLO [31] expansion in the strong interaction coupling constant α_s.

The theoretical framework for the present study is given in Sec. I B. In Sec. II, a description of the event reconstruction and selection is presented. Next, in Sec. III, we describe the calibration and analysis of the jets. Modeling the data using MC calculations and detector simulation is discussed in Sec. IV for both QCD and $t\bar{t}$ final-state processes. In Sec. V, the properties of observed jets are analyzed. A search for boosted top-quark production is described in Sec. VI. We summarize our conclusions in Sec. VII.

B. The theoretical framework

1. *Jet mass*

The primary source of high-p_T jets at high-energy hadron colliders is the production and subsequent fragmentation and hadronization of gluons and the five lightest quarks (*QCD jets*). The distribution of the mass of a QCD jet has a maximum, m_{peak}, comparable to a small fraction of the momentum of the jet, followed by a long tail that, depending on the jet algorithm used, could extend up to values that are a significant fraction of the p_T of the jet. Based on QCD factorization (see, e.g., [32]), a semi-analytic calculation of the QCD jet-mass distribution has been derived for this high-mass tail where the jet mass, m^{jet}, is dominated by a single gluon emission [33]. The probability of such gluon emission is given by the *jet functions* J^q and J^g for quarks and gluons, respectively. These are defined via the total double-differential cross section

$$\frac{d^2\sigma(R)}{dp_T dm^{jet}} = \sum_{q,g} J^{q,g}(m^{jet}, p_T, R)\,\frac{d\hat{\sigma}^{q,g}(R)}{dp_T}\,, \tag{1}$$

where R is the radius of the jet cone used to define the jets and $\hat{\sigma}^{q,g}$ is the factorized Born cross section. Corrections of $\mathcal{O}\left(R^2\right)$ are neglected and the analysis is applied to the high-mass tail, $m^{\text{peak}} \ll m^{jet} \ll p_T R$. An eikonal approximation for the full result [33] is

$$J^{q,g}(m^{jet}, p_T, R) \simeq \alpha_s(p_T)\frac{4C_{q,g}}{\pi m^{jet}}\log\left(\frac{R\,p_T}{m^{jet}}\right)\,, \tag{2}$$

where $\alpha_s(p_T)$ is evaluated at the appropriate scale and $C_{q,g} = 4/3$ and 3 for quark and gluon jets, respectively. This result is applicable to jet algorithms that are not strictly based on a cone, such

as the anti-k_T algorithm studied here.

The result in Eq. (2) allows two independent predictions. The first is that for sufficiently large jet masses, the absolute probability of a jet being produced with a given mass is inferred. It means that the jet function is a physical observable and has no arbitrary or unknown normalization. The second prediction is that the shape of the distribution has the same characteristic form for jets arising from quark and gluon showering, differing only by a scale factor. These predictions can be used to estimate the rejection power for QCD jets as a function of a jet-mass requirement when searching for a beyond-the-standard-model particle with mass in excess of 100 GeV/c^2 that decays hadronically [34, 35].

Equation (2) is the leading-log approximation to the full expression where the next-to-leading-order (NLO) corrections are not known [3, 36, 37]. These corrections are expected to be of order of $1/\log\left(R^2\,p_T^2/(m^{jet})^2\right) \approx 30\%$ for the jets discussed in this paper. Thus, while the above theoretical expressions are not precise, they still provide a simple and powerful description for the qualitative behavior of the high-m^{jet} tail.

Corrections from non-perburbative QCD effects, collectively known as the *soft function*, have been argued to be positive and to modify the jet function in the following way [33]:

$$J^{q,g}(m^{jet}, p_T, R) \simeq \alpha_s(p_T)\frac{4C_{q,g}}{\pi m^{jet}}\left[\log\left(\frac{R\,p_T}{m^{jet}}\right)+\frac{R^2}{2}\right]. \tag{3}$$

The additional soft contribution is approximately 50 percent for $R = 0.7$, $p_T = 400$ GeV/c and $m^{jet} = 100$ GeV/c^2.

2. *Jet substructure*

Single jets that originate from the decay of a highly-boosted massive particle fundamentally differ from QCD jets. The jet-mass distribution peaks at around the mass of the decaying particle in one case and at relatively lower values for QCD jets. The efforts in the literature to identify and characterize other jet substructure observables can be categorized into three broad classes: techniques specifically geared towards two-pronged kinematics [5, 6, 8, 38], techniques employing three-pronged kinematics [14, 33, 38–43] (e.g., $h \to b\bar{b}$ for two-body and $t \to bq\bar{q}$ for three-body kinematics) and methods that are structured towards removing soft particle contamination [44–46]. See Ref. [3, 47] for recent reviews.

We focus on measuring angularity and planar flow jet shape variables, which belong to the

first two classes of methods. At small cone sizes, high-p_T, and large jet mass, these variables are expected to be quite robust against soft radiation (i.e., are considered infrared- or IR-safe) and allow in principle a comparison with theoretical predictions in addition to comparison with MC results. Both variables are also less dependent on the particular jet finding algorithm used. We use the MIDPOINT cone algorithm [16] to reconstruct jets using the FASTJET program [48], and compare these results with the anti-k_T algorithm [49]. The choice of these two algorithms allows a comparison of cone (MIDPOINT) and recombination (anti-k_T) algorithms.

Angularity belongs to a class of jet shape variables [38, 50] and is defined as

$$\begin{aligned}\tau_a(R, p_T) &= \frac{1}{m^{jet}} \sum_{i \in jet} E_i \sin^a \theta_i \, [1 - \cos \theta_i]^{1-a} \\ &\approx \frac{2^{a-1}}{m^{jet}} \sum_{i \in jet} E_i \, \theta_i^{2-a}, \end{aligned} \tag{4}$$

where E_i is the energy of a jet constituent inside the jet and θ_i is the angle between the constituent three-vector momentum and the jet axis. The approximation is valid for small angle radiation $\theta_i \ll 1$. Limiting the parameter a not to exceed two ensures that angularity does not diverge at small angles, as evident from the last expression of Eq. (4) [51].

The angularity distribution, $d\sigma/d\tau_a$, is similar over a large class of jet definitions (for instance the k_T and anti-k_T variety [49]) in the limit of $R \ll 1$ and high jet mass [38]. It is particularly sensitive to the degree of angular symmetry in the energy deposition about the jet axis. It therefore can distinguish QCD jets from boosted heavy particle decay. The key point here is that for high-mass jets, the leading parton and the emitted gluon are expected to have a symmetric p_T configuration where both partons are at the same angle, θ_i, from the jet axis in the laboratory frame, $\theta_{1,2} = z \equiv m^{jet}/p_T$ [38]. This implies that angularity has a minimum and maximum value in such two-body configurations:

$$\tau_a^{\min}(z) \approx \left(\frac{z}{2}\right)^{1-a}, \tag{5}$$

$$\tau_a^{\max}(R, p_T) \approx 2^{a-1} R^{-a} z. \tag{6}$$

This provides an important test for the energy distribution of massive jets, as QCD jets should satisfy these values once they become sufficiently massive. Hence, the angularity distribution of jets arising from the two-body decay of a massive particle (for example, a W, Z, or Higgs boson) and QCD jets are similar in shapes for sufficiently large p_T and m^{jet}.

Assuming that the largest energy deposits occur at small angles relative to the jet direction, the angularity for two-body configurations has the form

$$\frac{d\sigma^{\mathrm{q,g}}}{d\tau_a}(m^{jet}, p_T, R) \approx \frac{4\alpha_s C_F}{\pi a m^{jet} \tau_a}. \tag{7}$$

This provides another test of the two-body nature of massive QCD jets.

We use another IR-safe jet shape denoted as planar flow (*Pf*), to distinguish planar from linear jet shapes [33, 38, 40]. For a given jet, we first construct a 2×2 matrix

$$I_w^{kl} = \frac{1}{m^{jet}} \sum_i E_i \frac{p_{i,k}}{E_i} \frac{p_{i,l}}{E_i}, \tag{8}$$

where E_i is the energy of constituent i in the jet, and $p_{i,k}$ is the kth component of its transverse momentum relative to the jet momentum axis. We define

$$Pf \equiv 4 \frac{\det(I_w)}{\mathrm{tr}(I_w)^2} = \frac{4\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2}, \tag{9}$$

where $\lambda_{1,2}$ are the eigenvalues of I_w. The planar flow vanishes for linear shapes and approaches unity for isotropic depositions of energy.

Jets with two-body substructure would in principle have $Pf = 0$. This would apply to leading order for events with highly-boosted weak gauge boson, Higgs bosons, and QCD jets. Jets with three-body substructure have a smooth *Pf* distribution with an enhancement for $Pf \approx 1$ [4, 38].

C. Expected sources of events

Studies of jet production using data collected during Run II at the Tevatron have shown that high-p_T jet production is well described by perturbative QCD. The primary source of jets is the production of parton pairs comprised of light quarks and gluons [52, 53]. To better understand the relative sources of jets, especially those that result in jets with large masses, we performed a PYTHIA 6.4 MC calculation [54] to predict the relative size of other standard model processes, such as W and Z boson production, as a function of the minimum transverse momentum, ${p_T}^{min}$, of the leading jet in the collision. We have assumed that the rate of light quark and gluon jets could be suppressed by a factor of 250 [33, 38, 40]; we haven't assumed a particular mechanism for suppressing the light quark and gluon jets.

The results of the PYTHIA calculation are shown in Fig. 1, where the relative abundance of jets with p_T in excess of ${p_T}^{min}$ as a function of ${p_T}^{min}$ is shown. It is evident that QCD jet production is

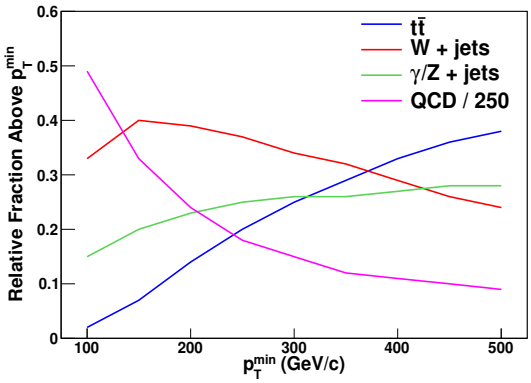


FIG. 1. The PYTHIA 6.4 MC prediction for the fractional contribution, relative to the total production cross section, of the various standard model sources as a function of the minimum p_T of the leading jet, assuming that the rate of jets from light quarks and bottom quarks can be suppressed by a factor of 250. The Z+jet cross section is separated from the Drell-Yan process by placing a mass requirement on the outgoing daughters. It is evident that QCD jet production is the dominant source of high-p_T jets.

the dominant source of high-p_T jets. The relative rate of $t\bar{t}$ production rises as the p_T cutoff is increased. At the highest ${p_T}^{min}$ values ($p_T > 400$ GeV/c), $t\bar{t}$ is predicted to contribute approximately 1% of the jet production cross section. This is the largest single contribution assuming that QCD jets can be suppressed by a factor of 250. Although we have not attempted to assess the theoretical uncertainties associated with this calculation, it provides motivation for better understanding the production of very high-p_T jets, and especially those that are massive.

D. Predictions for high-p_T top-quark production

An approximate NNLO calculation of the $t\bar{t}$ differential cross section [29] using the MSTW 2008 parton distribution functions (PDF) [55], a top-quark mass $m_t = 173$ GeV/c^2 and a renormalization scale $\mu^2 = p_T^2 + m_t^2$ [56] for high-p_T top quarks predicts that the $t\bar{t}$ cross section for $p_T > 400$ GeV/c is $4.55^{+0.50}_{-0.41}$ fb, or that the fraction of top quarks produced with $p_T > 400$ GeV/c is $\left(5.58^{+0.61}_{-0.50}\right) \times 10^{-4}$. The calculation includes next-to-leading-order (NLO) corrections to the leading order amplitudes along with NNLO soft-gluon corrections [30].

The results of this calculation can be compared with a PYTHIA 6.216 MC prediction for $t\bar{t}$ production, which yields a fractional rate of $(7.56 \pm 0.13) \times 10^{-4}$ (statistical error only), in reasonable agreement with the approximate NNLO calculation [29]. Based on the measured total $t\bar{t}$ produc-

tion cross section of 7.50 ± 0.48 pb [57] and on the PYTHIA fraction, one predicts a production cross section for top quarks with $p_T > 400$ GeV/c of 5.67 ± 0.37 fb, which again is in reasonable agreement with the approximate NNLO calculation. When estimating possible boosted top-quark contributions, we use the PYTHIA MC sample to describe the event kinematic properties and scale the event cross section for top quarks with $p_T > 400$ GeV/c to the approximate NNLO production cross section estimate.

II. DATA SAMPLES, EVENT RECONSTRUCTION AND SELECTION

A. Detector description

The CDF II detector is described in detail elsewhere [58]. We outline below the detector features that are most relevant to the present analysis.

The detector consists of a solenoidal spectrometer, calorimeters surrounding the tracking volume, and a set of charged-particle detectors outside the calorimeters for muon identification. The solenoidal charged-particle spectrometer provides charged-particle momentum measurement over $|\eta| < 1.5$. A superconducting magnet generates an axial field of 1.416 T. The charged particles are tracked with a set of silicon microstrip detectors arranged in a barrel geometry around the collision point. This is followed by a cylindrical drift chamber, the central outer tracker (COT), that provides charged-particle tracking from a radius of 40 to 137 cm.

The calorimeter system is used to measure the energy and mass of jets, and missing transverse energy ($\not\!\!E_T$). The central calorimeter system extends over the interval $|\eta| < 1.1$ and is segmented into towers of size $\Delta\eta \times \Delta\phi = 0.11 \times 0.26$. It consists of lead and steel absorbers interleaved with scintillator tiles that measure the deposited energy. The inner calorimeter compartment consists of lead absorbers providing an electromagnetic energy measurement (EM), while the outer compartment consists of steel absorbers to measure hadronic (HAD) energy. The energy (E) deposited in the EM calorimeter is measured with a relative resolution of $\sigma/E \approx (0.135/\sqrt{E} \oplus 2)\%$ while the relative resolution of the HAD calorimeter is $\sigma/E \approx (0.5/\sqrt{E} \oplus 3)\%$, where E is in units of GeV. Two plug calorimeters in the forward and background regions provide energy measurement in the interval $1.1 < |\eta| < 3.5$ using lead and steel absorbers interleaved with scintillator tiles that measure the deposited energy.

Measurement of $\not\!\!E_T$ is made by summing vectorially the energy deposits in each calorimeter

tower for towers with $|\eta| < 3.6$ and forming a missing energy vector. We take $\not\!\!E_T$ as the magnitude of the vector. The relative resolution of this quantity is proportional to $1/\sqrt{\sum E_T}$, where E_T is in GeV and the sum is over the transverse energy observed in all calorimeter towers. This has been determined by studies of events with and without significant missing transverse energy [53]. A measure of how large the observed $\not\!\!E_T$ in an event is relative to its uncertainty is provided by the $\not\!\!E_T$ significance, defined as

$$S_{MET} \equiv \frac{\not\!\!E_T}{\sqrt{\sum E_T}}, \tag{10}$$

where the sum in the denominator runs over the transverse energy observed in all calorimeter towers.

The detector also includes systems for electron, muon, and hadron identification, but these are not used in this study.

We employ the MIDPOINT jet algorithm [16] using a cone size $R = 0.7$ and correct the jet four-momentum vector for detector response and pile-up effects, as described in more detail in Sec. III. We also reconstruct MIDPOINT jets with a cone size $R = 0.4$ and $R = 1.0$ when studying the effects of cone size on various properties, and reconstruct jets with the anti-k_T algorithm [49]. The jet energies are corrected to the particle level, as described in Ref. [59].

B. Data and Monte Carlo samples

The present study is based on a Run II data sample corresponding to an integrated luminosity of 5.95 fb^{-1}. An inclusive jet trigger requiring at least one jet with $E_T > 100$ GeV is used to identify candidate events, leading to a sample of 76 million events.

We model QCD jet production using a PYTHIA 6.216 MC sample generated with parton transverse momentum $\hat{p}_T > 300$ GeV/c and the CTEQ5L parton distribution functions [60] corresponding to an integrated luminosity of approximately 800 fb^{-1}. Multiple interactions are incorporated into the model, assuming an average rate of 0.4 additional collisions per crossing. We verify that the parton $\hat{p}_T$ requirement has negligible bias for events with reconstructed jets whose corrected p_T exceeds 350 GeV/c. The average number of additional collisions per crossing in the MC samples is significantly less than that observed in the data. In the results reported below, we take this into account when comparing the MC predictions and experimental results. We do not use the MC modeling of multiple interactions to correct for these effects. Rather, we use a data-driven

approach as described below.

All MC events are passed through a full detector simulation and processed with the standard event-reconstruction software.

C. Event selection

Candidate events are required to satisfy the following requirements:

1. Each event must have a high quality $p\bar{p}$ interaction vertex with the primary vertex position along the beamline, z_{vtx}, within 60 cm of the nominal collision point.

2. Each event must have at least one jet constructed using the MIDPOINT cone algorithm using cone sizes of $R = 0.4$, 0.7, or 1.0 and having a $p_T > 400$ GeV/c in the pseudorapidity interval $|\eta| < 0.7$. The p_T requirement is made after applying η-dependent corrections to account for inhomogeneities in detector response, and calorimeter response non-linearities. Corrections have also been made to account for multiple interactions. The jet p_T is corrected to the particle level [59].

3. Each event must satisfy a relatively loose $\not\!\!E_T$ requirement of $S_{MET} < 10$ GeV$^{1/2}$ to reject cosmic ray backgrounds and poorly measured events.

Requirements are placed on the jet candidates to ensure that they are well-measured. We form the fraction

$$f_{tr} \equiv \frac{\Sigma_i^{N_{ch}} p_T^i}{p_T^{jet}}, \tag{11}$$

where N_{ch} is the number of charged particles associated with the jet candidate by summing those reconstructed charged particles within the jet cone radius of the jet direction defined by the calorimeter energy deposits, and p_T^i is the transverse momentum of the ith particle. The electromagnetic energy fraction of the jet candidate is defined by $f_{EM} = E_{EM}/(E_{EM} + E_{HAD})$, where E_{EM} and E_{HAD} are the electromagnetic and hadronic energy of the jet cluster. We require each jet candidate to satisfy either $f_{tr} > 0.05$ or $f_{EM} > 0.05$. These requirements reject 1.4% of the events in the data sample. They result in a negligible reduction in the Monte Carlo samples. We also reject events where the leading jet has $|\eta| < 0.1$ as the uncertainty of the energy response of those jets is larger due to variation in the exact location of the energy deposits.

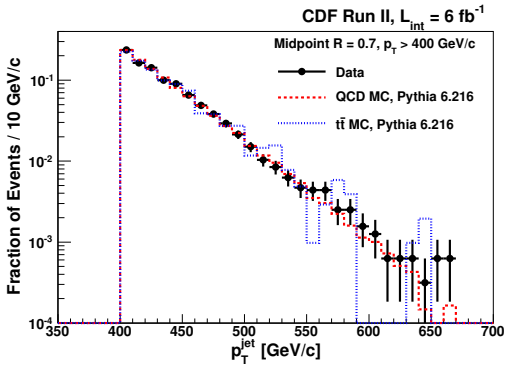


FIG. 2. The normalized p_T distribution for all the jets with $p_T > 400$ GeV/c in the sample that meets the inclusive event selection requirements. Overlaid are the distributions from the PYTHIA MC calculations for QCD jets and $t\bar{t}$ production.

This selection procedure yields 2699 events in which at least one jet with $R = 0.7$ has $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$. Within this sample, 591 events (22%) have a second jet satisfying the same requirements, resulting in 3290 jets all with $p_T > 400$ GeV/c. There are 211 jets with p_T higher than 500 GeV/c. The p_T distribution of all of the jets satisfying the selection requirements is shown in Fig. 2.

III. CALIBRATION AND ANALYSIS OF JETS

The CDF jet-energy corrections have been determined [59] for a large range of jet momenta and are used in this study. For jets with $p_T > 400$ GeV/c and measured in the central calorimeter, the systematic uncertainty in the overall jet-energy scale is 3% and is dominated by the understanding of the response of the calorimeter to individual particle energies. Other uncertainties such as out-of-cone effects, underlying-event energy flow and multiple interactions are an order of magnitude smaller at these jet energies.

A. Check of internal jet-energy scale with tracks

The relatively small uncertainty on the total jet energy of these high-p_T jets imposes a strong constraint on the variations in energy response across the plane perpendicular to the jet axis. Such a variation may not bias the energy measurement of the jet but may affect substructure observables like the jet mass.

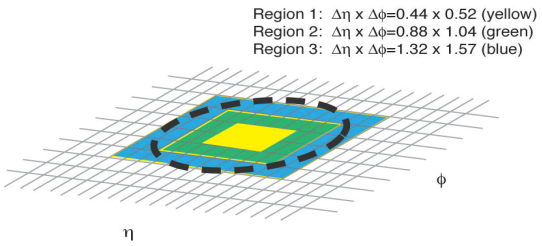


FIG. 3. A schematic of the three calorimeter regions used in the verification of the internal energy calibration within the jet. The dashed circle represents a cone of radius $R = 0.7$.

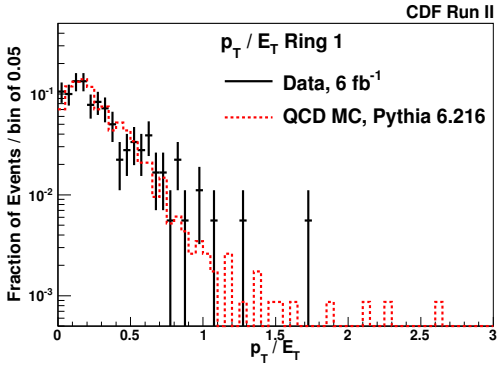
In order to assess the systematic uncertainty on the jet-mass scale, we compare the ratio of the sum of charged particle transverse momentum and the calorimeter transverse energy in three concentric rectangular regions in $\eta - \phi$ space centered around the jet axis. These regions have the following tower geometries: Region 1 is formed of 4 towers in η and 2 towers in ϕ with one of the four innermost towers closest to the jet centroid. Region 2 is formed of 8 towers in η and 4 towers in ϕ centered on Region 1 and excluding it. Region 3 is formed of 12 towers in η and 6 towers in ϕ centered on Region 1 and excluding the interior two regions. These regions are shown schematically in Fig. 3 overlaid by a jet cone of radius 0.7 for illustration purposes.

We form the ratio

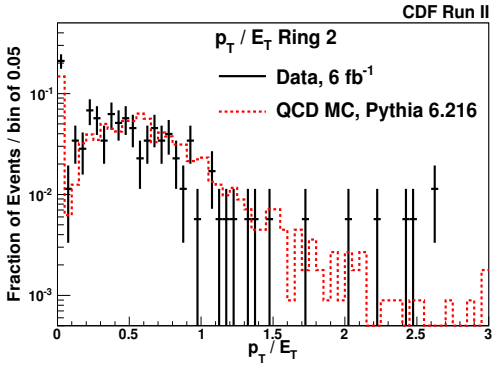
$$(p_T/E_T)_i = \frac{\sum\limits_{\substack{tracks \\ in\ region}} p_T}{\sum\limits_{\substack{towers \\ in\ region}} E_T} \tag{12}$$

for each region $i = 1, 2,$ and 3 for both the experimental and simulated data. The numerator is the sum of the transverse momentum of all charged particles reconstructed in the COT that intersect the given region when projected to the plane of the calorimeter. The charged particles are required to have $p_T > 1$ GeV/c. The denominator is the sum of the transverse energy deposited in each calorimeter tower in the region. To minimize the effect of multiple interactions, the number of primary vertices (N_{vtx}) in this study is required to be equal to one. The distributions of this ratio are shown in Figs. 4(a)–(c).

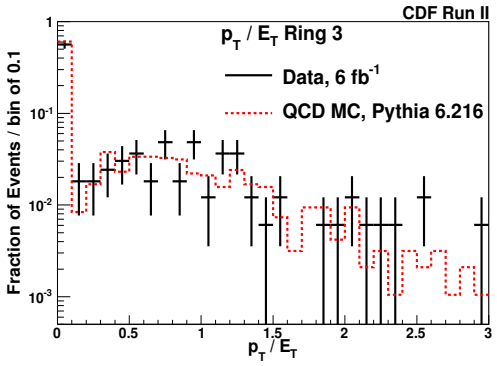
The ratio of p_T carried by charged particles to calorimeter transverse energy falls with increasing proximity to the core of the jet. This effect is consistent with other studies [61] that have



(a)



(b)



(c)

FIG. 4. The normalized distribution of the ratio between the sum of the charged particle p_T and calorimeter transverse energy in Region 1 (a), Region 2 (b), and Region 3 (c) for jets with $p_T \in (400, 500)$ GeV/c and $|\eta| \in (0.1, 0.7)$ for events with one primary vertex. The charged particles are required to have $p_T > 1$ GeV/c. The MC prediction for this distribution is given by the red dashed line.

shown that the COT track finding efficiency falls significantly as the density of nearby charged tracks rises. Charged particles found in Region 1 experience the highest such tracking densities. Hence the ratio is lowest for Region 1, where the observed distribution peaks at approximately 0.2. The ratio is larger on average for Regions 2 and 3, as expected. These features are reproduced well by the QCD MC and detector simulation, where it is assumed that the calorimeter energy response in a given tower is independent of the tower's location relative to the jet's core. The peak at zero in Figs. 4(b)–(c) arise from jets where all of the charged particles have $p_T < 1$ GeV/c or most of the jet energy is in the form of neutral particles.

The generally good agreement of the data with the Monte Carlo predictions indicate that there is no significant change in the calorimeter energy response as a function of the calorimeter tower's distance from the jet centroid.

The results of this study are summarized in Table I. To estimate the systematic uncertainty on jet substructure measurements arising from any remaining bias, we introduce three independent jet-energy corrections JES_i, one for each of the above defined regions, where JES_i is the ratio between the actual response and the calibration. These new parameters are constrained by the 3% uncertainty on the overall jet-energy scale. Namely the one standard deviation confidence interval is

$$0.97\,E_T^{ave} < JES_1\rho_1 A_1 + JES_2\rho_2 A_2 + JES_3\rho_3 A_3 < 1.03\,E_T^{ave}, \tag{13}$$

where ρ_i is the average energy density in Region i, A_i is the area of Region i relative to the area of the three regions summed together, and E_T^{ave} is the average energy of the jets in the sample.

We use the observed relative energy response of the calorimeter cells around the center of the jet to constrain the region-dependent energy scales. Since most of the jet's energy is deposited in the inner region, for which the MC and data are in reasonable agreement, the overall energy scale uncertainty of $\pm 3\%$ determines the strongest single constraint on JES_1. Since, on average, Region 1 captures 94% of the total energy of the leading jet in the sample, the uncertainty of JES_1 from the jet-energy systematic uncertainty is at most $0.03/0.94 = 0.032$. We use the difference between the observed and predicted ratios of charged particle momentum to calorimeter energy in Regions 2 and 3 to set uncertainties on JES_2 and JES_3. The observed and predicted ratios differ by factors of 0.69 ± 0.04 and 0.88 ± 0.06 for Region 2 vs Region 1 and Region 3 vs Region 1, respectively. These ratios have an additional systematic uncertainty that we estimate to be ± 0.10, arising from

	Region 1	Region 2	Region 3
Relative area (A_i)	0.111	0.333	0.555
Transverse energy density (ρ_i) [GeV/$\Delta\eta\Delta\phi$]	1744	33.7	1.50
Mean $f_{track/cal,R_i}$ (data)	0.176 ± 0.008	0.436 ± 0.012	0.815 ± 0.020
Mean $f_{track/cal,R_i}$ (QCD MC)	0.150 ± 0.005	0.538 ± 0.006	0.790 ± 0.012
E_i - fractional energy in region i	0.941	0.055	0.004

TABLE I. The relative areas of each calorimeter region, the average E_T densities in the three regions for jets with $p_T \in (400, 500)$ GeV/c and $|\eta| \in (0.1, 0.7)$, and the mean of $f_{track/cal,R_i}$, the ratio between the charged particle and calorimeter response for the data jets and the MC jets. The last line shows the average E_T deposited in each region for an average jet in this sample.

the variation in this ratio of ratios when the selection criteria for the jets and charged particles are varied.

The ratio of the JES_2 and JES_3 energy scales relative to JES_1 determine the systematic uncertainty on the jet-mass scale. We consider two cases, a typical jet with measured mass of 64 GeV/c^2 and a high-mass jet with measured mass of 115 GeV/c^2. The spatial distribution of the energy deposits are modeled as circular in $\eta - \phi$ space taking into account the actual $\eta - \phi$ segmentation of the calorimeter. The energy densities in the towers are set according to Table I to model the low mass jet. The largest possible shifts in the Region 1 scale, consistent with a one standard deviation drop in JES_2 and JES_3 are then determined.

The constraints on the JES_i translate to a systematic jet-mass uncertainty of 1 GeV/c^2 for low mass jets. We use the geometric high-mass jet model to set the constraints on more massive jets, and find that the corresponding systematic uncertainty on jets with masses in excess of 100 GeV/c^2 is 10 GeV/c^2.

Because we have assumed a broad energy distribution in the plane perpendicular to the jet's axis, this is a conservative estimate of the systematic uncertainty. We expect that high mass QCD and top quark jets arise from two or three large energy deposits, and not a broader energy distribution as we have assumed. Furthermore, we identify the maximum possible jet-mass excursion consistent with the one standard deviation measurements of the relative calorimeter region response, resulting in a conservative one standard deviation estimate.

In summary, the systematic uncertainty on the jet-mass scale arising from uncertainty in energy

scale changes as a function of the distance from the jet axis are 1 GeV/c^2 for jets with masses around 65 GeV/c^2, and 10 GeV/c^2 for jets with masses exceeding 100 GeV/c^2.

B. Sensitivity to multiple interactions and underlying event

In addition to the particles that arise from the parton showering and hadronization of a high-energy quark or gluon, a jet also may contain energy deposits produced from particles arising from the fragmentation of other high-energy quarks or gluons in the event, from the so-called underlying event, which is characterized by a large number of relatively low-energy particles, and particles coming from additional multiple collisions that occur in the same bunch crossing. The kinematics of the additional particles coming from the underlying event are correlated with the high-energy quarks or gluons [62] while the particle flow from multiple interactions are uncorrelated with the high-energy jets. These additional particles affect jet substructure variables and may significantly bias quantities such as jet mass [3].

The correction to the substructure of the jet due to the additional energy deposits is in general a function of the substructure. For example, the shift in jet mass from a single particle is inversely proportional to the mass of the jet, while the overall shift in mass from a collection of low-energy particles is predicted to increase as R^4, where R is the jet cluster radius [3]. We are able to discriminate the effect of the underlying event alone by measuring the number of primary interactions (N_{vtx}) and then separately consider events with $N_{vtx} = 1$ from events with $N_{vtx} > 1$. Jets in $N_{vtx} = 1$ events would only be affected by underlying event (UE) while jets in events with $N_{vtx} > 1$ would be affected by both UE and multiple interactions (UEMI).

We correct for multiple interaction (MI) effects using a data-driven technique [63]. We select a subset of events in the sample that have a clear dijet topology by requiring that the second jet in the event has $p_T > 100$ GeV/c and is at least 2.9 radians in azimuth away from the leading jet in addition to the previous event selection. We then define a *complementary* cone in $\eta - \phi$ space of the same radius as the jet cones and at the same η as the leading jet, but rotated in azimuth by $\pm\pi/2$. We then assign the energy deposits in each calorimeter tower in the complementary cone to the corresponding tower in the leading jet cone, distributing them as observed in the complementary cone. We then add these energy deposits to the jet using the standard four-vector recombination scheme and calculate a new jet mass, m_{new}, and a mass shift, $m_{new} - m_{old}$. We then calculate the average mass shift as a function of jet mass for the entire data sample. The upward

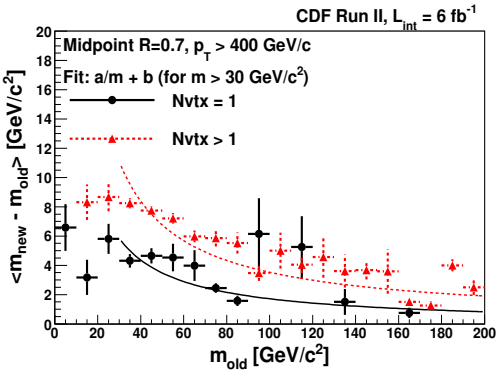


FIG. 5. The average shift in the reconstructed jet mass with respect to the true mass due to underlying event and multiple interactions (UEMI–dashed red points) and to underlying event alone (UE–black points) for selected jets as a function of the original jet mass m_{old}. Also shown are the parametrizations of these corrections (solid line for UE and dotted line for UEMI) used for the correction. The shifts have been estimated using the complementary cone approach.

shifts in jet mass for events with one and more than one interaction are estimates of the UE and UEMI effect, respectively, and can be used to statistically remove this effect from the observed jets.

The UE and UEMI jet-mass corrections as functions of the uncorrected jet mass for a cone size of $R = 0.7$ are shown in Fig. 5. Both corrections are assumed to have a $1/m^{jet}$ dependence, as expected from kinematic considerations, rising to a maximum around jet masses of approximately 30 GeV/c^2. The UE and UEMI corrections differ by approximately a factor of two. The average number of primary interactions for this sample is approximately three per event, which would suggest a similar factor for the difference between corrections. However, the UE contribution is more energetic than a typical $p\bar{p}$ collision and is correlated with the jet, leading to a larger jet-mass correction. We parametrize both jet corrections with a $1/m^{jet}$ dependence and an offset down to a jet mass of 30 GeV/c^2. Below this value the correction is expected to vanish at zero mass (since a jet with an observed small mass cannot have experienced any significant increase in m^{jet} from multiple interaction effects). We therefore chose a linear parametrization for $m^{jet} < 30$ GeV/c^2 with an intercept at zero. This has no effect on the heavy jets which are the focus of this analysis.

To check that the correction removes the effects of MI, we compare in Fig. 6 the distribution of the jet masses for the leading jets in the selected events with $N_{vtx} = 1$, with $N_{vtx} > 1$, and with $N_{vtx} > 1$ events in which the MI correction is made. The average jet-mass difference between

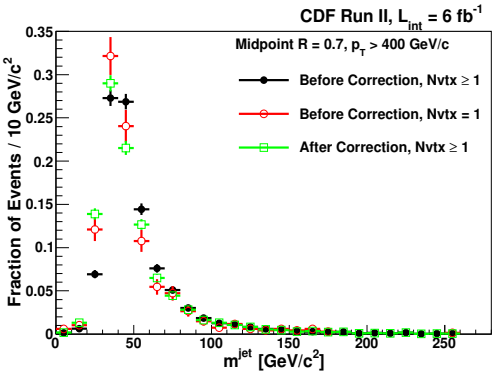


FIG. 6. The normalized jet-mass distribution for all selected jets for events with $N_{vtx} = 1$ (open red circles) and for events with $N_{vtx} \geq 1$ before (black points) and after (green open squares) the MI correction.

the jets with $N_{vtx} = 1$ and $N_{vtx} > 1$ is reduced from 3–4 GeV/c^2 to less than 2 GeV/c^2, and the low-mass peaks coincide. This residual difference in means is expected, given that the correction procedure does not account for the relatively rare cases where the UE or MI produce a large shift in jet mass.

The same UEMI and MI calculation is repeated for MIDPOINT jets with radius parameter $R = 0.4$. The mass shift due to MI scales as R^4, as expected [3], and is approximately 1 GeV/c^2 for jets with masses of 50 GeV/c^2. This correction method cannot be applied directly to $R = 1.0$ MIDPOINT jets, since in that case the complementary cones overlap with the original jet cone. We therefore scale the MI correction derived for $R = 0.7$ to jets with $R = 1.0$ using a scaling factor $(1.0/0.7)^4 = 4.16$. Since the $R = 0.4$ results have relatively large statistical uncertainties, we also use the $R = 0.7$ MI corrections scaled down by the corresponding factor for the $R = 0.4$ jets.

IV. COMPOSITION OF SELECTED SAMPLE

Events selected as described in Sec. II are expected to be due primarily to QCD dijet production. The requirements of a high-quality primary vertex, a jet cluster satisfying the p_T and η requirements, and the jet cleaning criteria eliminate virtually all other physics backgrounds and instrumental effects [53].

Predictions for QCD jet production using an NLO calculation with the POWHEG MC package [64–66] and the CTEQ6M parton distribution functions [67] show that approximately 80% of the jets arise from a high-p_T quark, consistent with measurements made at lower jet energies [17].

The cross sections for W and Z boson production are approximately 4 fb each, based on a PYTHIA 6.4 MC calculation. The only other standard model source of jets with masses > 100 GeV/c^2 is top-quark pair production. Although the cross section of top-quark pairs is expected to be of order 5 fb for $p_T > 400$ GeV/c, these events typically will have two massive jets.

We discuss below the characteristics and expected rates of jets from each of these sources.

A. QCD production

The selected jet p_T distribution using the MIDPOINT algorithm with $R = 0.7$ is shown in Fig. 2 for data and the QCD simulations. The agreement in shape confirms earlier measurements [53]. The leading jet-mass distribution for the QCD MC sample is shown in Fig. 7(a). It exhibits a sharp peak around 40 GeV/c^2 with a long tail that extends out to 300 GeV/c^2, similar to the data distribution shown in Fig. 6.

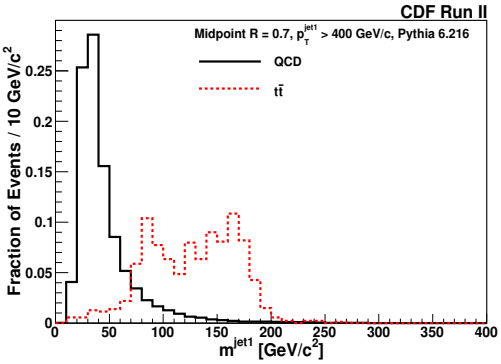
B. W and Z boson contamination

The PYTHIA calculation predicts cross sections of 4.5 fb and 3.0 fb for producing W and Z bosons with $p_T > 400$ GeV/c, respectively. These processes will contribute approximately 20 jets to the sample. In the data sample, these jets would have m^{jet1} between 50 and 100 GeV/c^2, where we observe 296 events.

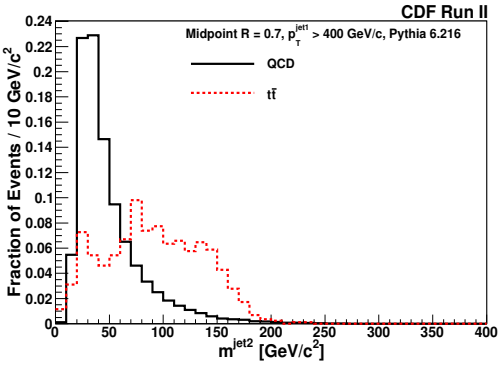
We do not subtract this background given the lower masses of W- and Z-originated jets compared to the high-mass jets of this study and the relatively modest size of this contribution to the overall jet rate.

C. Top quark pair production

The average p_T of top quarks produced in standard model $t\bar{t}$ production corresponds to approximately half the mass of the top quark and the p_T distribution exhibits a long tail to higher transverse momentum [29]. The events populating this tail potentially contribute to any analysis looking at highly-boosted jets. In order to understand the nature of this process and its characteristics when we require a central, high-p_T jet in the event, we make use of the PYTHIA top quark sample described earlier.



(a)



(b)

FIG. 7. The normalized jet-mass distributions for leading jets (a) and second leading jets (b) with $R = 0.7$ in MC QCD (solid) and $t\bar{t}$ (dashed) events. The leading jet is required to satisfy $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$ and the second-leading jet is required to satisfy $p_T > 100$ GeV/c.

The p_T distribution of top-quark jets after the selection cuts (Sec. II C) is shown in Fig. 2 for jets with a cone size $R = 0.7$. We compare the characteristics of the jets in the MC $t\bar{t}$ and QCD samples. We show in Fig. 7(a) the leading jet-mass distribution, m^{jet1}, for both the $t\bar{t}$ and QCD MC events using $R = 0.7$ jets with $p_T > 400$ GeV/c. A broad enhancement in the 160–190 GeV/c^2 mass range is visible for $t\bar{t}$ MC events along with a similar shoulder around 80 GeV/c^2. Only few $t\bar{t}$ events have leading jets with masses below ≈ 70 GeV/c^2 or above ≈ 200 GeV/c^2.

The characteristics of the second leading jet are compared in Figs. 7(b) and 8, where we show the m^{jet2} distributions and p_T distributions, respectively, for the second leading jet in the $t\bar{t}$ MC events and in the QCD MC events. The top-quark m^{jet2} distribution does not show an enhancement as seen in the leading jet. This is due to a smaller fraction of the top-quark decay products being captured in the recoil jet cone of $R = 0.7$ given the lower p_T distribution for the recoil jets.

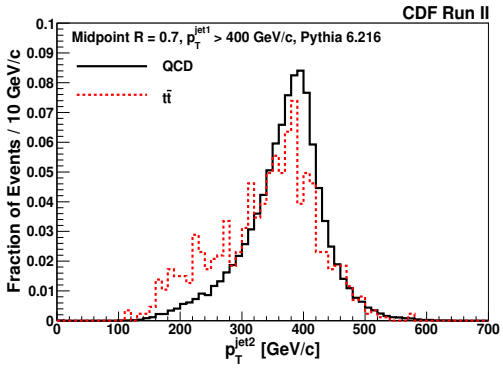


FIG. 8. The normalized jet p_T distribution of the second-leading jet ($R = 0.7$) in $t\bar{t}$ and QCD MC events, requiring that the leading jet satisfy $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$.

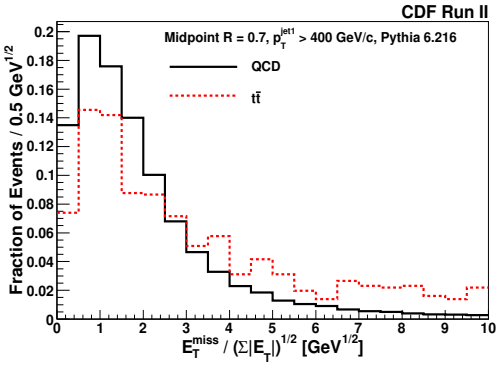


FIG. 9. The normalized missing transverse energy significance distributions for $t\bar{t}$ and QCD MC events requiring that the leading jet satisfy $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$.

The $t\bar{t}$ MC calculations predict that approximately one third of events in which a hadronically decaying top quark is observed as the leading jet would have a recoil top quark decaying semileptonically, resulting in missing transverse energy and a less massive second leading jet. We show in Fig. 9 the distributions of S_{MET} in MC events where we require a leading jet meeting the standard requirements of $p_T > 400$ GeV/c and $|\eta| < 0.7$. The $t\bar{t}$ events have a significant tail to larger S_{MET} compared with the QCD distribution, showing that this variable can be used to help separate $t\bar{t}$ and QCD jets.

D. Rejection of top-quark events

The primary goal of this study is to measure the jet substructure associated with highly-boosted QCD jets. A significant top-quark contribution would distort these substructure distributions. We therefore employ a strategy to reject $t\bar{t}$ contributions using the correlations predicted by the MC calculations.

The strategy focuses on two $t\bar{t}$ topologies that can be efficiently rejected. The first corresponds to the case where both top quarks decay hadronically and result in two massive jets, which we denote as the "$1+1$" topology. Such events are characterized by a second-leading jet with large mass and no significant $\not\!\!E_T$. The second topology corresponds to one top quark decaying hadronically and the other top quark decaying semileptonically, resulting in a massive jet recoiling against an energetic neutrino, a b-quark jet and a charged lepton. This "SL" topology is characterized by large S_{MET}, a second leading jet with a mass consistent with that of a b-quark jet and possibly a charged lepton candidate.

We implement the $t\bar{t}$ rejection strategy by rejecting an event with a second-leading jet with $m^{jet2} > 100$ GeV/c^2 or with $S_{MET} > 4$ GeV$^{1/2}$. We also require that the second leading jet has $p_T > 100$ GeV/c to ensure that each event has a sufficiently energetic recoil jet, though all data events satisfy this criterion. With these requirements, denoted as the *top-quark rejection cuts*, only 26% of the $t\bar{t}$ MC events satisfying the event selection requirements survive; 78% of the QCD MC events survive this requirement. This strategy reduces any $t\bar{t}$ contamination to ≈ 0.6 fb, or approximately 4 events in the data sample.

The resulting data distribution for m^{jet1} after making this selection is shown in Fig. 10. There are 2108 events in this 5.95 fb^{-1} sample. We study these events in more detail in Sec. V.

V. PROPERTIES OF OBSERVED JETS

The total number of events that pass the selection requirements as a function of two p_T intervals is shown in Tab. II for the different cone sizes. We examine the leading jet in each event that survives the selection requirements and the top-quark rejection cuts.

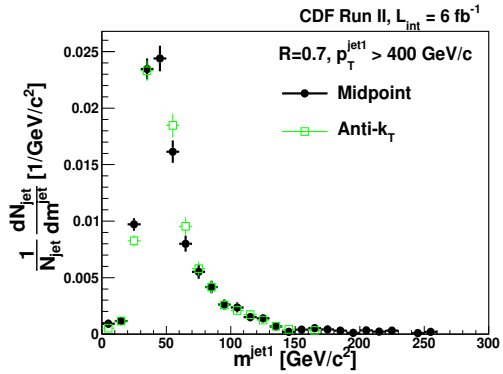


FIG. 10. Mass distribution of the leading jet with $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$ after applying the top-quark rejection. The results of the two clustering algorithms (black points for MIDPOINT and open green squares for anti-k_T) using a cone size or distance parameter of $R = 0.7$ are compared.

p_T Interval	Cone Size		
(GeV/c)	$R = 0.4$	$R = 0.7$	$R = 1.0$
$400 \leq p_T < 500$	1729	1988	2737
$p_T \geq 500$	107	120	175

TABLE II. The number of observed events with at least one jet in the p_T interval studied and for three different cone sizes. All events were required to have at least one MIDPOINT jet of the given cone size with $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$. The selection used to reject top quark candidates has been applied.

A. Cone sizes

In each event, we reconstruct MIDPOINT jets with cone sizes of $R = 0.4$, 0.7, and 1.0. We select the high-p_T jet sample by requiring that an event has at least one jet of any cone size with $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$. We therefore can compare directly the properties of jets with the three cone sizes. A comparison of the mass distributions for the three cone sizes is shown in Fig. 11. The distributions have similar structures, with a low-mass peak and an approximately power-law behavior at larger masses. The low-mass enhancement peaks around 30 GeV/c^2 for $R = 0.4$, with the peak position rising to approximately 60 GeV/c^2 for $R = 1.0$. The increase in average jet mass with cone size is in reasonable agreement with theoretical predictions [2].

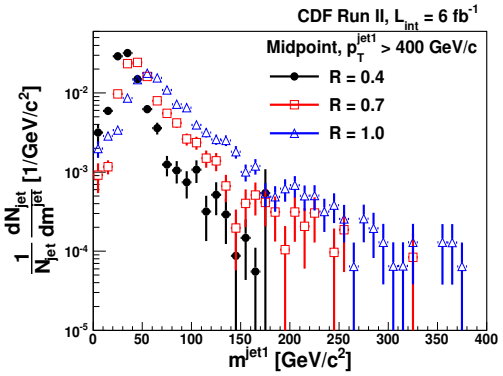


FIG. 11. The jet-mass distributions with cone sizes $R = 0.4$, 0.7 and 1.0 (black points, open red squares and open blue triangles, respectively) in the data sample for all jets with $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$.

B. Unfolding corrections

In order to make a comparison of data distributions with particle-level calculations and the eikonal predictions (Eq. 2), the observed jet-mass distributions are corrected to take into account effects that may bias the observed distribution. The most significant effects are from mass-dependent acceptance factors due to jet p_T resolution. We use the PYTHIA QCD MC to reconstruct particle-level jets with the various cone sizes and compare the corresponding distributions to the distributions resulting from the full detector simulation and selection requirements.

In particular, we consider bin migration effects due to the finite jet mass and p_T resolution. There is negligible net bin-to-bin migration across jet-mass bins for $m^{jet} > 70$ GeV/c^2. However, the p_T resolution of the jets varies by approximately 5% between jet masses of 50 and 150 GeV/c^2, with lower-mass jets having poorer p_T resolution. This results in the proportion of events with true $p_T < 400$ GeV/c satisfying the minimum jet p_T requirement to be a function of jet mass, decreasing with increasing jet mass, and therefore distorting the observed jet mass distribution. Hence, in calculating a normalized jet-mass distribution, we perform a correction to the observed mass distribution defined by the ratio

$$\left(\frac{1}{\sigma}\frac{d\sigma}{dm^{jet}_{particle}}\right) \Big/ \left(\frac{1}{\sigma}\frac{d\sigma}{dm^{jet}_{observed}}\right), \tag{14}$$

where σ is the cross section and the subscripts refer to the normalized distributions calculated with the particle-level (*particle*) jets and observed (*observed*) jets in MC events. The jets in the particle-level calculation are also required to have $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$. This *unfolding factor* is illustrated in Fig. 12, where we plot this ratio for $m^{jet1} > 70$ GeV/c^2. A polynomial is fit

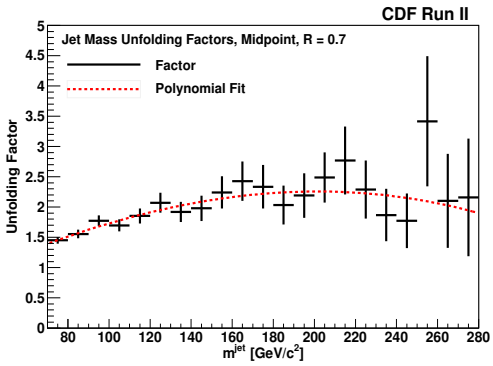


FIG. 12. The ratio of the jet-mass distributions for particle-level jets and detector-level jets with $R = 0.7$ in events from the PYTHIA MC calculation. The dashed red curve is the result of a polynomial fit to the MC points. The uncertainties originate from the limited size of the simulated sample.

to the points and the fit is used to correct the observed distribution for this migration effect.

Several sources of uncertainty for jet masses larger than 70 GeV/c^2 are associated with this correction. The first arises from the limited size of the MC event sample, and is shown in Fig. 12. The second arises from the model of jet fragmentation and hadronization used. The unfolding factor varies by less than 10% when the jet is subject to fragmentation and hadronization. We therefore consider this as an additional uncertainty on the resulting measured jet function. Third, the uncertainty in the jet-energy calibration introduces an uncertainty in the correction that is estimated by varying the calibration scale by its uncertainty and observing the change in the correction. This introduces an additional 10% uncertainty in the correction. Finally, the use of PDFs with their associated normalization scales introduces additional uncertainties. These are determined using the eigenvector approach [68], and are found not to exceed 10%. We add these in quadrature to determine an overall uncertainty on the unfolding factor and propagate that to the measured jet-mass distribution.

We have performed similar studies for angularity and planar flow and found the unfolding corrections to be negligible, except for the case of planar flow for $R = 1.0$ jets, where the corrections are of order 10%.

C. Systematic uncertainties on observed substructure

We summarize the various sources of uncertainties in the following subsections.

1. Calorimeter energy scales

The study of the region-dependence of the jet-energy response constrains the size of possible bias in jet-mass scale that would arise from a systematic under or overestimate of the energy response as a function of distance from the jet axis. For jet masses around 60 GeV/c^2, the systematic uncertainty on the jet-mass scale is 1 GeV/c^2, which increases with the jet mass. Conservatively, we estimate the maximum possible shift to be 10 GeV/c^2 for jet masses larger than 100 GeV/c^2 and we use this value when propagating these uncertainties to jets with $m^{jet} > 70$ GeV/c^2.

2. Energy flow from multiple interactions

The studies of the energy flow in these events, both on average and as a function of the number of primary vertices, show that multiple interactions shift the jet-mass scale. We estimate this shift to be 3–4 GeV/c^2 for jets with masses above 70 GeV/c^2 and a cone size of $R = 0.7$. The jet-mass distribution of the MI-corrected jets reproduce the jet-mass distribution for the single-vertex events to better than 2 GeV/c^2. We therefore set the uncertainty on this shift conservatively at 2 GeV/c^2, which is half the value of the MI correction.

3. Uncertainties on the PYTHIA *predictions for substructure*

In making a comparison of the observed distributions with those predicted by a MC calculation, we take into account the uncertainties arising from the choice of PDFs and renormalization scale using the eigenvector approach [68]. We reweight the MC events by increasing or decreasing each of the 20 eigenvectors and choices of scale describing the PDF parameterization by one standard deviation. We take the shifts associated with each bin of the normalized distributions from the variation in each of the 20 pairs in quadrature as the PDF uncertainty in that bin. These uncertainties are approximately 10% for the jet-mass distributions and 5% for angularity and planar flow.

4. Substructure systematics summary

The largest systematic uncertainty on the jet mass for masses larger than 70 GeV/c^2 comes from the energy calibration of the calorimeter, and is estimated to be 10 GeV/c^2. The uncertainty associated with the modeling of multiple interactions is 2 GeV/c^2. These are independent effects

and so we combine them in quadrature for an overall systematic uncertainty on the jet-mass scale of $\sigma_{syst} = 11$ GeV/c^2. The systematic uncertainty at lower masses is smaller, and we estimate it to be 2 GeV/c^2 for jets with masses of 60 GeV/c^2.

We propagate the uncertainty in the jet mass by determining the effect of shifts of $+1$ σ_{syst} and -1 σ_{syst} on the measured values. In the following figures, we show this uncertainty separately. This is straightforward for the jet function, where the measured value is affected. For the two other substructure variables, the potential sources of systematic uncertainty come from the understanding of the energy calibration as a function of the distance from the jet axis, as well as potential changes in the event selection due to the uncertainty on the jet mass. To determine the sensitivity to the energy calibration, the variables were recalculated assuming correlated changes in the energy scale of the towers as described in Sec. III A.

D. Results and comparison with theoretical models

1. Jet mass and jet function

The mass distribution for highly-boosted jets is characterized theoretically by the *jet function* approximated in Eq. (2). Over a relatively wide range of large jet masses, it predicts both the shape of the distribution and the fraction of jets in this range relative to all the jets in the sample.

We show in Fig. 13(b) a comparison of the observed mass distribution of the leading jet for $m^{jet1} > 70$ GeV/c^2, corrected as described earlier, with the analytic predictions for the jet function for quark and gluon jets, using a cone size $R = 0.7$. The solid bars reflect the systematic uncertainty from the jet-mass scale. The analytical prediction employs the average p_T for the jets in this sample of $< p_T >= 430$ GeV/c and a strong interaction coupling constant of α_s fixed at 0.0973, the value determined for the average p_T [69]. The quark jet function prediction is in good agreement with the shape of the jet-mass distribution for jet masses greater than 100 GeV/c^2. It is also consistent with the expectation that about 80-85% of these jets would arise from high-energy quarks, given that the data lie closest to the predictions for quark jets. The prediction gives the probability distribution for producing a jet with a given mass so its normalization is fixed. We also show the PYTHIA MC prediction, which is in good agreement with the observed distribution. Given the observation that the eikonal approximation agrees with the observed distribution, it suggests that there are effects that moderate the contribution of the soft function shown in Eq. 3.

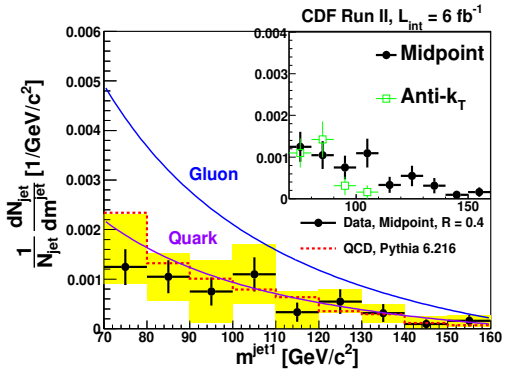
Since the jet mass can help discriminate jets arising from light quarks and gluons from jets arising from the decay of a heavy particle, the measured jet function allows us to estimate the rejection factor associated with a simple mass cut. Only $1.4 \pm 0.3\%$ of the jets reconstructed with the MIDPOINT algorithm with $R = 0.7$, $p_T > 400$ GeV/c and $\eta \in (0.1, 0.7)$ have $m^{jet} > 140$ GeV/c^2, corresponding to a factor of 70 in rejection against QCD jets.

We expect that the perturbative QCD NLO calculation for the jet mass would be sensitive to the cone size. We show the corresponding mass distributions for the leading jet in the selected events constructed using a cone size of $R = 0.4$ and 1.0; for consistency, the event and jet selection was repeated using the different cone sizes. The resulting mass distribution for $R = 0.4$ over the region $m^{jet1} \in (70, 160)$ GeV/c^2 is shown in Fig. 13(a), and the jet-mass distribution for $R = 1.0$ for $m^{jet1} \in (70, 400)$ GeV/c^2 is shown in Fig. 13(c). We also display the predicted jet functions for these cone sizes, using the values for the average $< p_T >= 430$ GeV/c of the jets and α_s as noted above. We again see good agreement between the data and the predicted shape and normalization for quark jets in the jet-mass region where we expect the analytic calculation to be robust. The analytic predictions and PYTHIA calculations also agree.

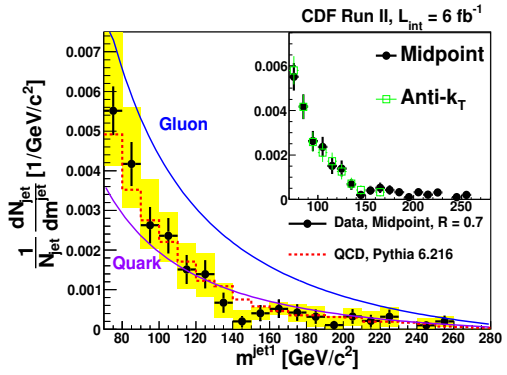
We also compare the jet-mass distributions for the MIDPOINT and anti-k_T algorithms. The anti-k_T jets have a similar mass distribution to the MIDPOINT jets but do not reproduce the large tail of very massive jets, presumably due to the explicit merging mechanism in the MIDPOINT algorithm. This difference in algorithm performance is reproduced by the PYTHIA calculation.

2. *Angularity*

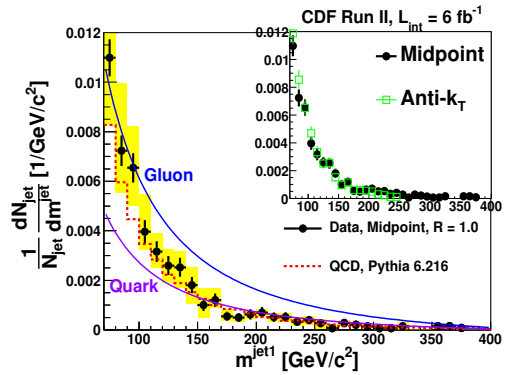
The jet angularity, defined in Eq. (4), provides discrimination between QCD jets from those produced in other processes. The angularity distribution for QCD jets with a given jet mass is predicted to be lower- and upper-bounded, and to decrease as $1/\tau_{-2}$ (7). We show in Fig. 14(a) the distribution of angularity for the leading jet with $R = 0.7$ in the sample requiring that $m^{jet1} \in (90, 120)$ GeV/c^2. This mass range was selected as the best compromise between a narrow, high-mass range with sufficient statistics and one in which W and Z boson contamination is suppressed. We expect at most a few jets from W and Z boson production in this sample. We compare the observed angularity distribution with the prediction from the PYTHIA calculation and the NLO pQCD constraints shown in Eqs. (5) and (6). We also show in Fig. 14(b) the angularity distribution for jets formed with a cone size of $R = 0.4$.



(a)

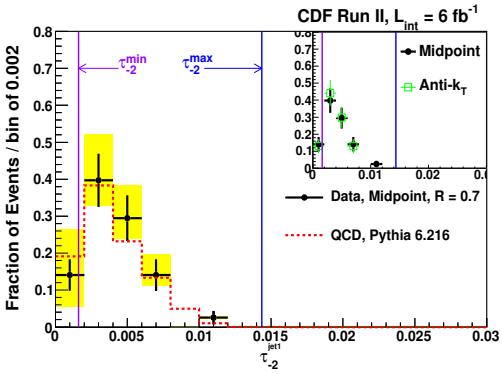


(b)

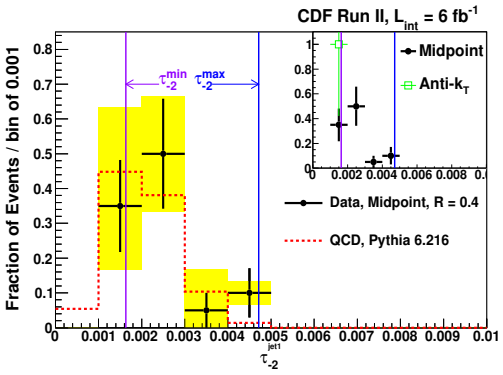


(c)

FIG. 13. Multiple-interaction-corrected jet-mass distributions for jets with $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$ reconstructed with an $R = 0.4$ (a), $R = 0.7$ (b) (a reproduction of Fig. 1 of [15]), and $R = 1.0$ (c) MIDPOINT cone algorithm after rejection of $t\bar{t}$ events. Comparisons with the analytic expression for the jet function for quarks and gluons are shown. The inset compares the results with the anti-k_T jet algorithm.



(a)



(b)

FIG. 14. The normalized angularity distributions for jets with $p_T > 400$ GeV/c, $|\eta| \in (0.1, 0.7)$, and $m^{jet1} \in (90, 120)$ GeV/c^2 reconstructed with the $R = 0.7$ (a) and $R = 0.4$ (b) MIDPOINT cone algorithm. We have rejected the $t\bar{t}$ events. The results from the PYTHIA calculation and analytic QCD predictions for the minimum and maximum values are overlaid. The inset compares the results with the anti-k_T jet algorithm. Subfigure (b) is a reproduction of Fig. 2 in [15].

The distributions for the jets with $R = 0.7$ have the behavior expected of QCD jets, approximately satisfying the minimum and maximum ranges and falling in a manner consistent with $1/\tau_{-2}$. The measured distributions for $R = 0.4$ jets have large statistical uncertainties. The small number of jets that have angularity below τ^{min} arise from resolution effects not taken into account in the calculation of the kinematic boundary. The PYTHIA distributions are in agreement with the data.

We investigate the sensitivity of the τ_{-2} distribution to MI effects using the same approach employed for jet mass [63]. Angularity was found to be insensitive to MI, with a correction for the multivertex events of 0.0005 for $R = 0.7$ jets, or less than 10% of the average observed value.

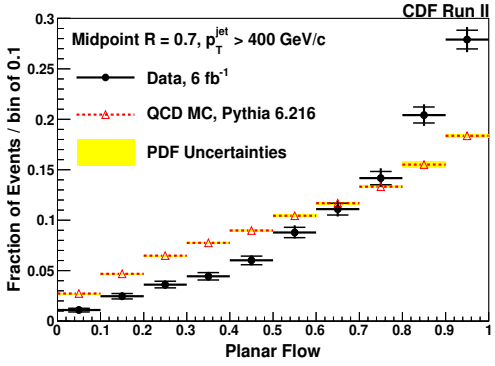
We do not correct the distributions for this effect. No significant resolution effects are seen from studies of MC samples and therefore we do not unfold these distributions for such effects.

3. Planar flow

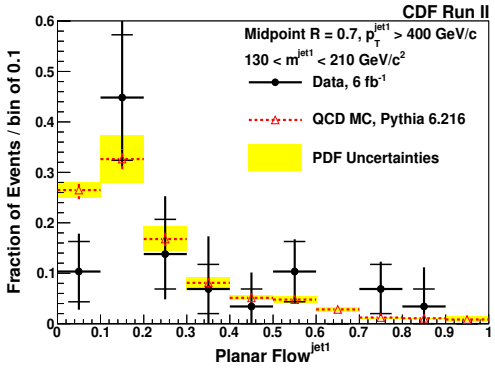
The jet planar flow, *Pf*, characterizes QCD and top-quark jets. For jets with cone sizes of $R = 0.7$, MC studies show that no significant resolution effects distort the observed *Pf* distributions, so we make no unfolding corrections. For jets with $R = 1.0$, it is necessary to correct the observed distribution for such distortions, leading to corrections of approximately 10–30% as a function of *Pf*.

The planar flow is largely complementary to jet mass for high-mass jets. This is most readily demonstrated by comparing the *Pf* distributions in Figs. 15(a) and 15(b). In Fig. 15(a), we make no jet-mass requirement while in Fig. 15(b), we apply the top-quark rejection cuts and only consider events with $m^{jet1} \in (130, 210)$ GeV/c^2. Without the jet-mass requirement applied, the *Pf* distributions for the data and the PYTHIA prediction for quark and gluon jets are monotonically increasing. As the full data set is dominated by low mass QCD jets, such a planar flow distribution is expected as it reflects a largely circular energy deposition. The PYTHIA prediction fails to account for the sharper rise in the *Pf* distribution for $Pf > 0.6$. When we apply the mass window requirement and the top-quark rejection cuts, the observed distribution has a peak at low *Pf*, also consistent with the QCD prediction. This observation directly supports the NLO prediction that massive jets from light quarks and gluons have two-body substructure and arise from single hard gluon emission.

The *Pf* distribution is sensitive to contributions from top-quark jets, as they would result in events with larger planar flow, especially for jets with $R = 1.0$, where we would expect a larger top-quark jet contribution due to higher reconstruction efficiencies once a large jet-mass requirement is made. We compare in Figs. 16(a) and 16(b) the planar flow distributions for the $R = 1.0$ jets predicted by the QCD and $t\bar{t}$ MC samples. Although the data are consistent with QCD jet production, as evidenced by the broad peak at planar flow values below 0.3, there is small excess of events at large *Pf* compared with the QCD prediction that is consistent with a small $t\bar{t}$ component.



(a)



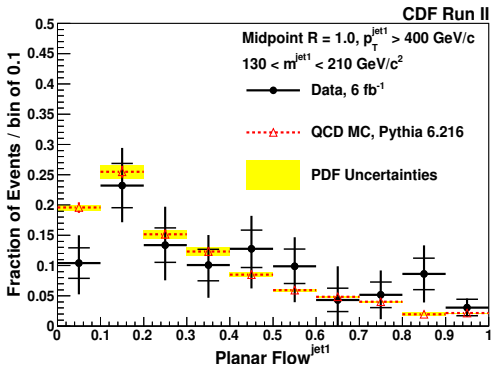
(b)

FIG. 15. The normalized planar flow distribution for jets with $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$ reconstructed with the $R = 0.7$ MIDPOINT cone algorithm. We have not rejected $t\bar{t}$ events and have not placed any constraint on the jet mass in (a). The distribution after top-quark rejection and requiring $m^{jet1} \in (130, 210)$ GeV/c^2 is shown in (b). Data points are shown with statistical and systematic uncertainties. Results from the PYTHIA QCD prediction (red triangles) with the PDF uncertainties (yellow bars) are overlaid.

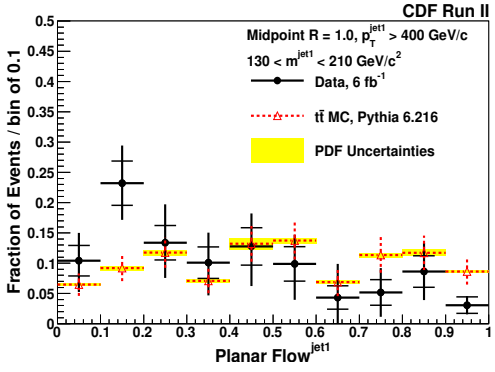
VI. BOOSTED TOP QUARKS

The studies of jet mass and other substructure variables, including the need to reject contributions from potential top quark pair production, lead naturally to an extension of the analysis to directly search for production of top quarks with $p_T > 400$ GeV/c (which we call "boosted top quark production"). We therefore focus on the 1+1 and SL topologies identified earlier to search for a boosted top quark signal.

We reconstruct the events with the MIDPOINT cone algorithm with $R = 1.0$ as that provides the



(a)



(b)

FIG. 16. The normalized planar flow distribution for jets with $p_T > 400$ GeV/c and $|\eta| \in (0.1, 0.7)$ reconstructed with the $R = 1.0$ MIDPOINT cone algorithm. We have rejected $t\bar{t}$ events and have required that $m^{jet1} \in (130, 210)$ GeV/c^2. Data points are shown with statistical and systematic uncertainties. Overlaid in (a) are results from the PYTHIA QCD prediction (red triangles) with the PDF uncertainties (yellow bars). Overlaid in (b) are results from the PYTHIA MC prediction for the leading jet in $t\bar{t}$ MC events (red triangles) with the PDF uncertainties (yellow bars).

greatest efficiency for capturing the final-state particles of a fully-hadronically decaying top quark in a single jet. We also increase the acceptance of the analysis by considering jets in the entire pseudorapidity interval $|\eta| < 0.7$.

A. Boosted top quarks in the 1+1 topology

The 1+1 topology is intended to identify top quark pairs where both top quarks decay hadronically. We start with 4230 events with a leading MIDPOINT jet with $R = 1.0$ and jet $p_T > 400$ GeV/c

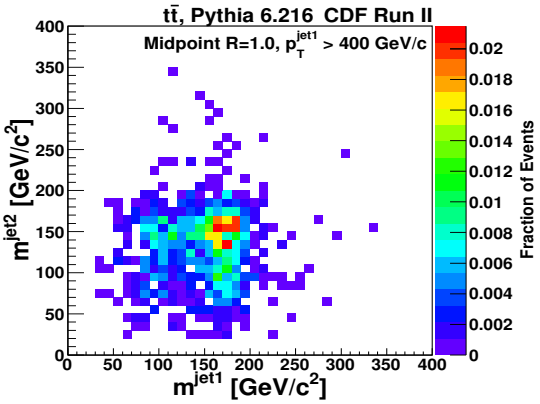
and $|\eta| < 0.7$.

A simple strategy to detect the presence of two hadronically-decaying top quarks is to require two massive jets with no evidence of large $\not\!\!E_T$ using the S_{MET} variable. We show in Fig. 17(a) the distribution of the mass of the second-leading jet, m^{jet2}, versus the mass of the leading jet, m^{jet1}, for $t\bar{t}$ MC events passing the requires on the leading jet described above and with $S_{MET} <$ 4 GeV$^{1/2}$. Given the clear clustering of the signal in this distribution, we define a signal region with both jet candidates having jet masses between 130 and 210 GeV/c^2. We show in Fig. 17(b) the same distribution for the QCD MC sample, showing that the top quark signal and the QCD background are reasonably well-separated. The $t\bar{t}$ MC calculation predicts that 11.2% of the top quark events with at least one top quark with $p_T > 400$ GeV/c would have jets satisfying this selection. We expect to see 3.0 events in the signal region.

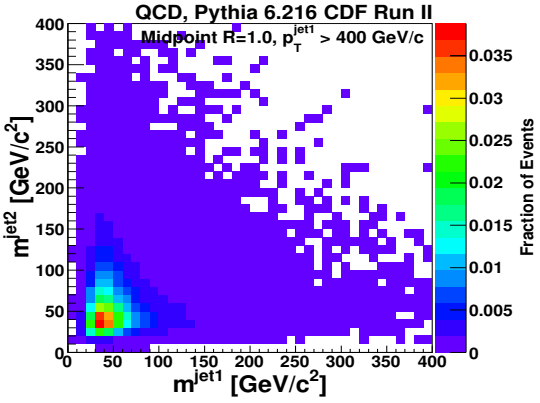
Figure 17(c) shows the 2-dimensional jet-mass plot for the data. We expect that the mass of the two jets produced via QCD interactions would be largely uncorrelated [70]. No correlation (coefficient $\rho = 0.06$) between the second-leading and leading jet masses is observed in the data or the PYTHIA QCD prediction. This is to be compared with the correlation in p_T of the two leading jets of 0.64 for the data sample. In addition, studies of the mass distributions of the leading and second-leading jet in the PYTHIA MC events, comparing the m^{jet2} distributions when different m^{jet1} requirements are applied, confirm the lack of significant correlation. Theoretical studies, as discussed below, are used to estimate the effect of any correlations in m^{jet} between the two leading QCD jets.

The uncorrelated jet masses allow an estimation of the background coming from QCD jet production in the top quark signal region. We use the observed distribution in either m^{jet1} or m^{jet2} of events in the low jet-mass peak (defined here to be 30–50 GeV/c^2) relative to events in the top-quark mass window of 130 to 210 GeV/c^2 to estimate the QCD background in the signal region where both jet masses are between 130 and 210 GeV/c^2.

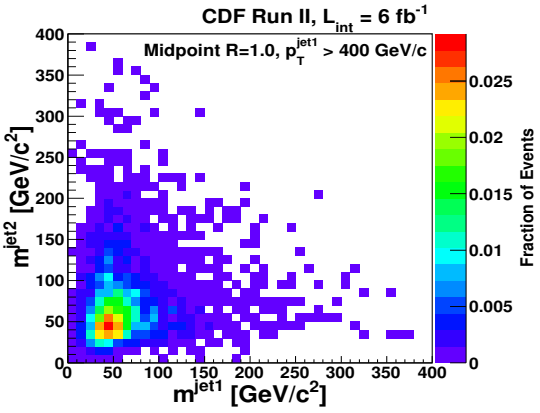
We define four regions in Fig. 17(c): Region A with both the leading and second leading jet with masses between 30 and 50 GeV/c^2, Region B with $m^{jet1} \in (130, 210)$ and $m^{jet2} \in (30, 50)$ GeV/c^2, Region C with $m^{jet1} \in (30, 50)$ and $m^{jet2} \in (130, 210)$ GeV/c^2, and Region D with both jets with masses between 130 and 210 GeV/c^2. We also define N_i to be the number of events observed in the ith region. By assuming no correlations between the two variables, $N_C/N_A = N_D/N_B$ would hold, providing a direct prediction of the number of QCD background events in Region D. The



(a)



(b)



(c)

FIG. 17. The m^{jet2} versus m^{jet1} distribution for simulated $t\bar{t}$ events (a), for simulated QCD events (b), and for MI-corrected data events (c) with at least one jet with $p_T > 400$ GeV/c and $|\eta| < 0.7$ using $R = 1.0$ MIDPOINT cones. The events are required to have $S_{MET} < 4$ GeV$^{1/2}$.

Region	m^{jet1}	m^{jet2}	Data	$t\bar{t}$ MC
	(GeV/c^2)	(GeV/c^2)	(events)	(events)
A	$(30,50)$	$(30,50)$	370	0.00
B	$(130,210)$	$(30,50)$	47	0.08
C	$(30,50)$	$(130,210)$	102	0.01
D (signal)	$(130,210)$	$(130,210)$	31	3.03
Predicted QCD in D			14.6 ± 2.7	

TABLE III. The observed number of events in the three control regions used to predict the background rate in the signal region (region D). The predicted $t\bar{t}$ event rates are also shown.

ratio

$$R_{mass} \equiv \frac{N_C N_B}{N_A N_D} \tag{15}$$

differs from unity for QCD jet production if the jet masses are correlated. This ratio was estimated in a separate study [70] using several different NLO QCD calculations, giving values that range from 0.86 to 0.89. A relatively small correlation is present in the QCD jets that produces more pairs of jets with high masses than would be expected if the leading and recoil jet masses were completely uncorrelated. A POWHEG MC calculation yields $R_{mass} = 0.89 \pm 0.03(\text{stat}) \pm 0.03(\text{syst})$. The systematic uncertainty takes into account the variation in the prediction using different MC generators, similar to the comparison in Ref. [70].

There are 370 events with both jets in Region A, 47 events in Region B, and 102 events in Region C. The difference in region B and C arise from the different p_T thresholds on the leading and second-leading jets. With these data and using the POWHEG R_{mass} value, we estimate the number of QCD background events in the signal region (Region D) to be 14.6 ± 2.7 (stat). There are 31 events in the signal region. This calculation is summarized in Tab. III.

B. Boosted top quarks in the SL topology

In order to observe $t\bar{t}$ events where one top quark decays semileptonically (*lepton+jets* final state), we use the sample of high-p_T jet events where the leading jet is massive, the recoil jet is not necessarily massive and where the event has substantial $\not\!\!E_T$. The top quark MC predicts that the requirement of $4 < S_{MET} < 10$ GeV$^{1/2}$ is correlated with a larger fraction of the recoil jets having

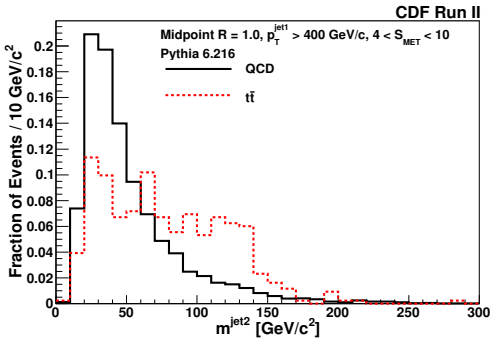
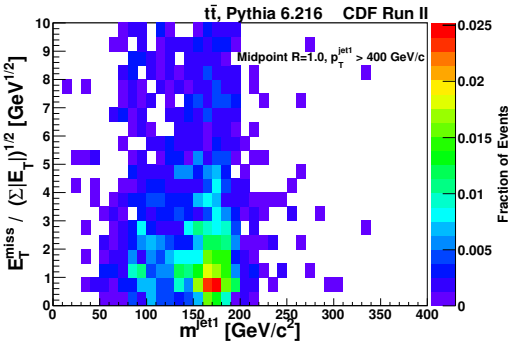


FIG. 18. The m^{jet2} distribution for $t\bar{t}$ and QCD MC events restricted to the sample having a leading jet with $p_T > 400$ GeV/c and $|\eta| < 0.7$ using $R = 1.0$ MIDPOINT cones and $S_{MET} \in (4,10)$ GeV$^{1/2}$.

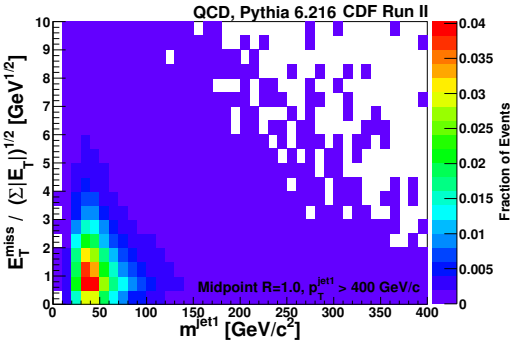
lower masses, as would be expected when one top quark has decayed semileptonically. Figure 18 shows the jet-mass distribution of the second-leading jets in such $t\bar{t}$ MC events. We also show the PYTHIA QCD background distribution for these events, illustrating that the second-leading jet mass is no longer an effective discriminant between signal and background.

We show in Figs. 19(a), 19(b), and 19(c) the distributions of S_{MET} vs m^{jet1} for the events restricted to have a leading jet with $p_T > 400$ GeV/c and $|\eta| < 0.7$ in the simulated $t\bar{t}$ sample, QCD sample, and in the data, respectively. This illustrates the effectiveness of the S_{MET} requirement to separate the signal from the background for this sample. We therefore define the SL signal event sample by requiring a leading jet with $m^{jet1} \in (130,210)$ GeV/c^2 and $S_{MET} \in (4,10)$ GeV$^{1/2}$. The $t\bar{t}$ MC predicts 1.9 events in this signal region.

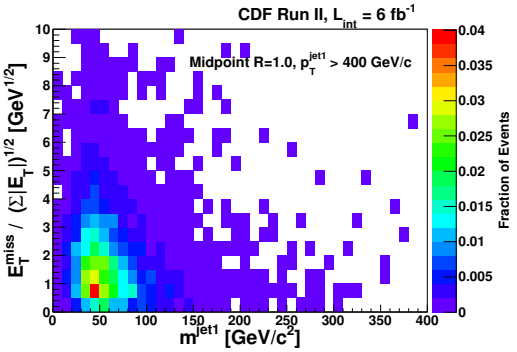
To estimate the QCD background in the SL signal region, we use the independence between the leading jet mass and S_{MET} in QCD background events. A correlation may arise from instrumental effects, e.g., arising from the jet being incident on an uninstrumented region of the detector, resulting in a lower jet mass and increased S_{MET}. We have searched for such a correlation in the data set, and found no evidence for such instrumental effects. We therefore perform a data-driven background calculation similar to that used for the 1+1 candidates. We define Region E to be $m^{jet1} \in (30,50)$ GeV/c^2 and $S_{MET} \in (2,3)$ GeV$^{1/2}$, Region F as $m^{jet1} \in (130,210)$ GeV/c^2 and $S_{MET} \in (2,3)$ GeV$^{1/2}$, Region G to be $m^{jet1} \in (30,50)$ GeV/c^2 and $S_{MET} \in (4,10)$ GeV$^{1/2}$, and Region H to be the signal region. Region E contains 256 events, Region F contains 42 events and Region G contains 191 events. We predict 31.3 ± 5.7 (stat) events in Region H (the signal region). We verified that the result is robust against reasonable variations in the definitions of the four regions, providing further confirmation that the two variables used are not correlated in this



(a)



(b)



(c)

FIG. 19. The S_{MET} versus m^{jet1} distribution for simulated $t\bar{t}$ events (a), simulated QCD events (b), and all data events (c) with at least one jet with $p_T > 400$ GeV/c and $|\eta| < 0.7$ using $R = 1.0$ MIDPOINT cones.

sample.

There are 26 events in this signal region, consistent with the background estimate and also consistent with the number of expected background and signal events. This calculation is summarized in Tab. IV.

Since we expect comparable signal yields and backgrounds in the 1+1 and SL channels, we

Region	m^{jet1}	S_{MET}	Data	$t\bar{t}$ MC
	(GeV/c^2)	(GeV$^{1/2}$)	(events)	(events)
E	$(30,50)$	$(2,3)$	256	0.01
F	$(130,210)$	$(2,3)$	42	1.07
G	$(30,50)$	$(4,10)$	191	0.03
H (signal)	$(130,210)$	$(4,10)$	26	1.90
Predicted QCD in H			31.3 ± 5.7	

TABLE IV. The observed number of events in the three control regions used to predict the background rate in the signal region (Region H) for the SL topology. The predicted $t\bar{t}$ event rates are also shown.

combine the results of the two channels. There are 57 candidate events with an expected background from QCD jets of 46 ± 6 events (the uncertainty is only statistical). The systematic uncertainty on the background rate is dominated by the uncertainty on the jet-mass scale (see the next subsection) and results in a background estimate of 46 ± 6 (stat) ± 14 (syst) events.

Although we observe an excess in the fully-hadronic final state, we see a combined event rate that is consistent with the expected QCD background. We use these data to set upper limits on the boosted top-quark production cross section.

C. Systematic uncertainties on top-quark production

The largest source of systematic uncertainty arises from the jet-mass scale. Other sources are the top quark acceptance due to the uncertainty in the jet-energy scale, the uncertainty in the integrated luminosity in the sample, the uncertainty on the $t\bar{t}$ acceptance due to the top-quark mass uncertainty and the uncertainty on R_{mass}.

The studies described in Sec. V C provided a determination of the systematic uncertainty on the jet-mass measurement of ± 10 GeV/c^2 for high-mass jets. We estimate the effect of the jet-mass uncertainty by shifting the upper jet-mass window by ± 10 GeV/c^2 and observing how the QCD background estimate changes. In this calculation, we leave the lower jet mass window fixed, resulting in a somewhat more conservative estimate of the background uncertainty. This results in a systematic uncertainty of $\pm 30\%$ on the combined background rate of 46 events.

The jet-energy-scale uncertainty results in a systematic uncertainty on the top quark acceptance,

determined by shifting the jet p_T scale by $\pm 3\%$. The efficiency is sensitive to the jet-energy scale because an underestimate in the jet-energy scale would reduce the observed rate of $t\bar{t}$ events and *vice-versa*. The resulting change in the top quark acceptance is $\pm 24.5\%$, using the p_T distribution from the approximate NNLO calculation.

We incorporate a systematic uncertainty on the integrated luminosity of $\pm 6\%$ [71]. The $t\bar{t}$ acceptance uncertainty due to possible variations in the top-quark mass is $\pm 0.3\%$.

We assume that these are all independent sources of uncertainty and consider them in the limit calculations in Sec. VI D.

D. Limits on massive particle pair production

We calculate the 95% confidence level (C.L.) limit on the $t\bar{t}$ production cross section using the CL_s approach, which performs a frequentist calculation using pseudoexperiments to combine statistical and systematic uncertainties [72].

Taking into account the overall $t\bar{t}$ detection efficiency of 18.2% (defined as the ratio of MC events satisfying the 1+1 and SL requirements over the number of $t\bar{t}$ events with at least one top quark with $p_T > 400$ GeV/c) and the integrated luminosity of 5.95 fb^{-1}, we exclude at 95% C.L. a standard model cross section for producing top quark pairs with top quark $p_T > 400$ GeV/c greater than 38 fb. This is approximately an order of magnitude higher than the estimated standard model cross section, and is limited by the size of the backgrounds from light quark and gluon jets. It is the most stringent limit on boosted top-quark production at the Tevatron to date and probes for the first time top-quark production at the Tevatron in this momentum range.

We support the upper limit calculation by estimating the expected limit as the median of all exclusion limits obtained in simulated samples that include the background estimated from the data-driven technique and including the expected number of $t\bar{t}$ events. The CL_s calculation yields an upper limit of 33 fb at 95% C.L., which is lower than the observed limit since we see a modest excess of events above the expected signal plus background in the data.

As theoretical models exist that predict pair production of massive particles that decay primarily hadronically, we set a limit on the pair production of massive beyond-the-standard-model particles near the mass of the top quark and decay hadronically. An example of such a scenario would be a light baryon-number-violating neutralino or gluino particle in the context of supersymmetry (see, e.g., [14, 73]) and in some theories of coloured resonances [74]. We have 31 events with two jets

with $m^{jet} \in (130, 210)$ GeV/c^2, with a background estimate of 14.6 ± 2.7 (stat) ± 3.9 (syst) events. As we are interested in beyond-the-standard-model contributions to this final state, we now include in the background estimate the expected $t\bar{t}$ contribution of 3.0 ± 0.8 events. We use the acceptance for top quark pair production in this channel (11.2%), correct the top quark hadronic branching fraction of $4/9$, and assume the same systematic uncertainties described earlier. The CL_s calculation gives an upper limit of 20 fb at 95% C.L.

VII. CONCLUSION

We report results on the nature of very high-p_T jets produced in hadron-hadron collisions, especially their substructure properties and possible sources. We have measured the jet-mass distribution and the distributions of two IR-safe substructure variables, angularity and planar flow, for jets with $p_T > 400$ GeV/c. The agreement between the QCD Monte Carlo calculations using PYTHIA 6.216, the analytic theoretical calculations and the observed data for jet masses greater than 70 GeV/c^2, indicates that these theoretical models reproduce satisfactorily the data and may be used to extrapolate backgrounds arising from light quark and gluon jets in searches for new phenomena at the LHC. The measurements of the angularity of QCD jets produced with masses in excess of 90 GeV/c^2 show that these are consistent with the NLO prediction of two-body structure, and the planar flow distribution for jets with masses between 130 and 210 GeV/c^2 show similar consistency with QCD predictions.

We compare the results obtained with the MIDPOINT cone algorithm with the anti-k_T algorithm, and find that the two algorithms produce very similar results. We note that these results are in good agreement with recent measurements of similar jet properties produced at the Large Hadron Collider in much higher energy proton-proton collisions [19–21].

We also report a search for boosted top-quark production using data gathered with an inclusive jet trigger at the Tevatron Collider. There is a modest excess of events – 57 candidate events with an estimated background of 46 ± 6 (stat) ± 14 (syst) events – identified in either a configuration with two high-p_T jets each with mass between 130 and 210 GeV/c^2 or where a massive jet recoils against a second jet with significant missing transverse energy. We expect approximately 5 signal events from standard model top-quark production where at least one of the top quarks results in a jet with $p_T > 400$ GeV/c. We set a 95% C.L. upper limit of 38 fb on the cross section for top-quark pair production with at least one of the top quarks having $p_T > 400$ GeV/c.

We use these data to also search for pair production of a massive particle with mass comparable to that of the top quark with at least one of the particles having $p_T > 400$ GeV/c. We set an upper limit on the pair production of 20 fb at 95% C.L. Observation of boosted top-quark production at the LHC where both top quarks decay hadronically have been reported [75, 76], showing that the substructure techniques reported here and others have relevance to such higher energy pp collisions.

ACKNOWLEDGMENTS

We acknowledge the contributions of numerous theorists for insights and calculations. Special thanks go to I. Sung and G. Sterman for discussions involving non-perturbative effects in QCD jets, and to N. Kidonakis for updated top quark differential cross section calculations.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, United Kingdom; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; the Australian Research Council (ARC); and the EU community Marie Curie Fellowship Contract No. 302103.

This work was also supported by the Shrum Foundation, the Weizman Institute of Science and the Israel Science Foundation.

---

[1] We use a coordinate system where ϕ and θ are the azimuthal and polar angles around the $\hat{z}$ direction defined by the proton beam axis. The pseudorapidity is $\eta = -\ln\tan(\theta/2)$ and $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. Transverse momentum is $p_T = p\sin\theta$ and transverse energy is $E_T = E\sin\theta$, where p and E are the momentum and energy, respectively.

[2] S. D. Ellis, J. Huston, K. Hatakeyama, P. Loch, and M. Toennesmann, Prog. Part. Nucl. Phys. **60**, 484 (2008), arXiv:0712.2447 [hep-ph].

[3] G. P. Salam, Eur. Phys. J. **67**, 637 (2010), arXiv:0906.1833 [hep-ph].

[4] T. Han, S. Lee, F. Maltoni, G. Perez, Z. Sullivan, T. Tait, and L. T. Wang, Nucl. Phys. Proc. Suppl. **200-202**, 185 (2010), arXiv:1001.2693 [hep-ph].

[5] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, Phys. Rev. Lett. **100**, 242001 (2008), arXiv:0802.2470 [hep-ph].

[6] G. D. Kribs, A. Martin, T. S. Roy, and M. Spannowsky, Phys. Rev. D **81**, 111501 (2010), arXiv:0912.4731 [hep-ph].

[7] T. Plehn, G. P. Salam, and M. Spannowsky, Phys. Rev. Lett. **104**, 111801 (2010), arXiv:0910.5472 [hep-ph].

[8] J. M. Butterworth, B. E. Cox, and J. R. Forshaw, Phys. Rev. D **65**, 096014 (2002), arXiv:hep-ph/0201098.

[9] K. Agashe, A. Belyaev, T. Krupovnickas, G. Perez, and J. Virzi, Phys. Rev. D **77**, 015003 (2008), arXiv:hep-ph/0612015.

[10] A. L. Fitzpatrick, J. Kaplan, L. Randall, and L. T. Wang, J. High Energy Phys. **09**, 013 (2007), arXiv:hep-ph/0701150.

[11] B. Lillie, L. Randall, and L. T. Wang, J. High Energy Phys. **09**, 074 (2007), arXiv:hep-ph/0701166.

[12] K. Agashe, H. Davoudiasl, G. Perez, and A. Soni, Phys. Rev. D **76**, 036006 (2007), arXiv:hep-ph/0701186.

[13] K. Agashe *et al.*, Phys. Rev. D **76**, 115015 (2007), arXiv:0709.0007 [hep-ph].

[14] J. M. Butterworth, J. R. Ellis, A. R. Raklev, and G. P. Salam, Phys. Rev. Lett. **103**, 241803 (2009), arXiv:0906.0728 [hep-ph].

[15] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **85**, 091101 (2012), arXiv:1106.5952 [hep-ex].

[16] G. C. Blazey *et al.*, (2000), arXiv:hep-ex/0005012.

[17] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 112002 (2005), arXiv:hep-ex/0505013.

[18] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **78**, 072005 (2008), arXiv:0806.1699 [hep-ex].

[19] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J **C71**, 1795 (2011), arXiv:1109.5816 [hep-ex].

[20] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. **06**, 160 (2012), arXiv:1204.3170 [hep-ex].

[21] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **83**, 052003 (2011), arXiv:1101.0070 [hep-ex].

[22] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. **05**, 128 (2012), arXiv:1203.4606 [hep-ex].

[23] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. D **86**, 072006 (2012), arXiv:1206.5369 [hep-ex].

[24] S. Chatrchyan *et al.* (CMS Collaboration), J. High Energy Phys. **05**, 090 (2013), arXiv:1303.4811 [hep-ex].

[25] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. **09**, 076 (2013), arXiv:1306.4945 [hep-ex].

[26] T. Affolder *et al.* (CDF Collaboration), Phys. Rev. Lett. **87**, 102001 (2001).

[27] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. **B693**, 515 (2010), arXiv:1001.1900 [hep-ex].

[28] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **102**, 222003 (2009).

[29] N. Kidonakis and R. Vogt, Phys. Rev. D **68**, 114014 (2003), arXiv:hep-ph/0308222.

[30] N. Kidonakis, Phys. Rev. D **82**, 114030 (2010), arXiv:1009.4935 [hep-ph].

[31] M. Czakon, P. Fiedler, and A. Mitov, Phys. Rev. Lett. **110**, 252004 (2013), arXiv:1303.6254 [hep-ph].

[32] J. C. Collins, D. E. Soper, and G. Sterman, Adv. Ser. Direct. High Energy Phys. **5**, 1 (1988), arXiv:hep-ph/0409313.

[33] L. G. Almeida, S. J. Lee, G. Perez, I. Sung, and J. Virzi, Phys. Rev. D **79**, 074012 (2009), arXiv:0810.0934 [hep-ph].

[34] W. Skiba and D. Tucker-Smith, Phys. Rev. D **75**, 115010 (2007), arXiv:hep-ph/0701247.

[35] B. Holdom, J. High Energy Phys. **08**, 069 (2007), arXiv:0705.1736 [hep-ph].

[36] E. L. H. Contopanagos and G. Sterman, Nucl. Phys. **B484**, 303 (1997), arXiv:hep-ph/9604313.

[37] M. Dasgupta and G. P. Salam, J. Phys. G **30**, R143 (2004), arXiv:hep-ph/0312283.

[38] L. G. Almeida, S. J. Lee, G. Perez, G. Sterman, I. Sung, and J. Virzi, Phys. Rev. D **79**, 074017 (2009), arXiv:0807.0234 [hep-ph].

[39] J. M. Butterworth, J. R. Ellis, and A. R. Raklev, J. High Energy Phys. **05**, 033 (2007), arXiv:hep-ph/0702150.

[40] J. Thaler and L. T. Wang, J. High Energy Phys. **07**, 092 (2008), arXiv:0806.0023 [hep-ph].

[41] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, Phys. Rev. Lett. **101**, 142001 (2008), arXiv:0806.0848 [hep-ph].

[42] D. Krohn, J. Shelton, and L.-T. Wang, J. High Energy Phys. **07**, 041 (2010), arXiv:0909.3855 [hep-ph].

[43] G. Brooijmans *et al.*, arXiv:0802.3715 [hep-ph].

[44] D. Krohn, J. Thaler, and L.-T. Wang, J. High Energy Phys. **02**, 084 (2010), arXiv:0912.1342 [hep-ph].

[45] S. D. Ellis, A. Hornig, C. Lee, C. K. Vermilion, and J. R. Walsh, J. High Energy Phys. **11**, 101 (2010),

arXiv:1001.0014 [hep-ph].

[46] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys.Rev. D **81**, 094023 (2010), arXiv:0912.0033 [hep-ph].

[47] A. Altheimer *et al.*, J. Phys. G **39**, 063001 (2012), arXiv:1201.0008 [hep-ex].

[48] M. Cacciari and G. P. Salam, Phys. Lett. **B641**, 57 (2006), arXiv:hep-ph/0512210.

[49] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. **04**, 063 (2008), arXiv:0802.1189 [hep-ph].

[50] C. Berger, T. Kucs and G. Sterman, Phys. Rev. D **68**, 014012 (2003), arXiv:hep-ph/0303051.

[51] In the original definition of angularity within a jet [38], the argument of the sin and cos functions was defined as $\pi\theta_i/(2R)$. However, for a generic jet algorithm configuration, $\theta_i \approx 2R$ are sometimes obtained and this results in singular behavior for angularity. Hence, we present a slightly improved expression where these singularities are avoided in the narrow cone case [77].

[52] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **101**, 062001 (2008), arXiv:0802.2400 [hep-ex].

[53] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **78**, 052006 (2008), arXiv:0807.2204 [hep-ex].

[54] T. Sjostrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. **05**, 026 (2006), arXiv:hep-ph/0603175.

[55] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Eur. Phys. J. **C63**, 189 (2009), arXiv:0901.0002 [hep-ph].

[56] N. Kidonakis, private communication.

[57] T. Aaltonen *et al.* (CDF Collaboration), CDF Conference Note 9913 (2009).

[58] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 032001 (2005), arXiv:hep-ex/0412071.

[59] A. Bhatti *et al.*, Nucl. Instrum. Method A **566**, 375 (2006), arXiv:hep-ex/0510047.

[60] H. L. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. Olness, J. F. Owens, J. Pumplin, and W. K. Tung, Eur. Phys. J. C **12**, 375 (2000), arXiv:hep-ph/9903282.

[61] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **102**, 232002 (2009), 0811.2820 [hep-ex].

[62] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **70**, 072002 (2004), arXiv:hep-ex/0404004.

[63] R. Alon, E. Duchovni, G. Perez, A. P. Pranko, and P. Sinervo, Phys. Rev. D **84**, 114025 (2011), arXiv:1101.3002 [hep-ex].

[64] P. Nason, J. High Energy Phys. **11**, 040 (2004), arXiv:hep-ph/0409146.

[65] S. Frixione, P. Nason, and C. Oleari, J. High Energy Phys. **11**, 070 (2007), arXiv:0709.2092 [hep-ph].

[66] S. Alioli, P. Nason, C. Oleari, and E. Re, J. High Energy Phys. **06**, 043 (2010), arXiv:1002.2581 [hep-ph].

[67] J. Pumplin, D. Stump, J. Huston, H. Lai, P. Nadolsky, and W. Tung, J. High Energy Phys. **07**, 012 (2002), arXiv:hep-ph/0201195.

[68] J. Pumplin *et al.*, Phys. Rev. D **65**, 014013 (2001), arXiv:hep-ph/0101032.

[69] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).

[70] K. Blum *et al.* (CDF Collaboration), Phys. Lett. **B702**, 364 (2011), arXiv:1102.3133 [hep-ex].

[71] S. Klimenko, J. Konigsberg, and T. M. Liss, Report FERMILAB-FN-0741 (2003).

[72] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103**, 221801 (2009), arXiv:0907.0810 [hep-ex].

[73] G. Brooijmans *et al.*, CERN-PH-TH/2010-096 (2010), arXiv:1005.1229 [hep-ph].

[74] C. Kilic, T. Okui, and R. Sundrum, J. High Energy Phys. **07**, 038 (2008), arXiv:0802.2568 [hep-ph].

[75] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. **01**, 116 (2013), arXiv:1211.2202 [hep-ex].

[76] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **111**, 211804 (2013), arXiv:1309.2030 [hep-ex].

[77] R. Alon, O. Gedalia, G. Perez, G. Sterman, and I. Sung, Private communication.