Subjet Multiplicity of Gluon and Quark Jets Reconstructed with the $k_\perp$ Algorithm in $p\bar{p}$ Collisions

V.M. Abazov,[23] B. Abbott,[57] A. Abdesselam,[11] M. Abolins,[50] V. Abramov,[26]
B.S. Acharya,[17] D.L. Adams,[59] M. Adams,[37] S.N. Ahmed,[21] G.D. Alexeev,[23] A. Alton,[49]
G.A. Alves,[2] N. Amos,[49] E.W. Anderson,[42] Y. Arnoud,[9] C. Avila,[5] M.M. Baarmand,[54]
V.V. Babintsev,[26] L. Babukhadia,[54] T.C. Bacon,[28] A. Baden,[46] B. Baldin,[36] P.W. Balm,[20]
S. Banerjee,[17] E. Barberis,[30] P. Baringer,[43] J. Barreto,[2] J.F. Bartlett,[36] U. Bassler,[12]
D. Bauer,[28] A. Bean,[43] F. Beaudette,[11] M. Begel,[53] A. Belyaev,[35] S.B. Beri,[15]
G. Bernardi,[12] I. Bertram,[27] A. Besson,[9] R. Beuselinck,[28] V.A. Bezzubov,[26] P.C. Bhat,[36]
V. Bhatnagar,[11] M. Bhattacharjee,[54] G. Blazey,[38] F. Blekman,[20] S. Blessing,[35]
A. Boehnlein,[36] N.I. Bojko,[26] F. Borcherding,[36] K. Bos,[20] T. Bose,[52] A. Brandt,[59]
R. Breedon,[31] G. Briskin,[58] R. Brock,[50] G. Brooijmans,[36] A. Bross,[36] D. Buchholz,[39]
M. Buehler,[37] V. Buescher,[14] V.S. Burtovoi,[26] J.M. Butler,[47] F. Canelli,[53] W. Carvalho,[3]
D. Casey,[50] Z. Casilum,[54] H. Castilla-Valdez,[19] D. Chakraborty,[38] K.M. Chan,[53]
S.V. Chekulaev,[26] D.K. Cho,[53] S. Choi,[34] S. Chopra,[55] J.H. Christenson,[36] M. Chung,[37]
D. Claes,[51] A.R. Clark,[30] J. Cochran,[34] L. Coney,[41] B. Connolly,[35] W.E. Cooper,[36]
D. Coppage,[43] S. Crépé-Renaudin,[9] M.A.C. Cummings,[38] D. Cutts,[58] G.A. Davis,[53]
K. Davis,[29] K. De,[59] S.J. de Jong,[21] K. Del Signore,[49] M. Demarteau,[36] R. Demina,[44]
P. Demine,[9] D. Denisov,[36] S.P. Denisov,[26] S. Desai,[54] H.T. Diehl,[36] M. Diesburg,[36]
S. Doulas,[48] Y. Ducros,[13] L.V. Dudko,[25] S. Duensing,[21] L. Duflot,[11] S.R. Dugad,[17]
A. Duperrin,[10] A. Dyshkant,[38] D. Edmunds,[50] J. Ellison,[34] V.D. Elvira,[36] R. Engelmann,[54]
S. Eno,[46] G. Eppley,[61] P. Ermolov,[25] O.V. Eroshin,[26] J. Estrada,[53] H. Evans,[52]
V.N. Evdokimov,[26] T. Fahland,[33] S. Feher,[36] D. Fein,[29] T. Ferbel,[53] F. Filthaut,[21]
H.E. Fisk,[36] Y. Fisyak,[55] E. Flattum,[36] F. Fleuret,[12] M. Fortner,[38] H. Fox,[39] K.C. Frame,[50]
S. Fu,[52] S. Fuess,[36] E. Gallas,[36] A.N. Galyaev,[26] M. Gao,[52] V. Gavrilov,[24] R.J. Genik II,[27]
K. Genser,[36] C.E. Gerber,[37] Y. Gershtein,[58] R. Gilmartin,[35] G. Ginther,[53] B. Gómez,[5]
G. Gómez,[46] P.I. Goncharov,[26] J.L. González Solís,[19] H. Gordon,[55] L.T. Goss,[60]
K. Gounder,[36] A. Goussiou,[28] N. Graf,[55] G. Graham,[46] P.D. Grannis,[54] J.A. Green,[42]
H. Greenlee,[36] Z.D. Greenwood,[45] S. Grinstein,[1] L. Groer,[52] S. Grünendahl,[36] A. Gupta,[17]
S.N. Gurzhiev,[26] G. Gutierrez,[36] P. Gutierrez,[57] N.J. Hadley,[46] H. Haggerty,[36]
S. Hagopian,[35] V. Hagopian,[35] R.E. Hall,[32] P. Hanlet,[48] S. Hansen,[36] J.M. Hauptman,[42]
C. Hays,[52] C. Hebert,[43] D. Hedin,[38] J.M. Heinmiller,[37] A.P. Heinson,[34] U. Heintz,[47]
T. Heuring,[35] M.D. Hildreth,[41] R. Hirosky,[62] J.D. Hobbs,[54] B. Hoeneisen,[8] Y. Huang,[49]
R. Illingworth,[28] A.S. Ito,[36] M. Jaffré,[11] S. Jain,[17] R. Jesik,[28] K. Johns,[29] M. Johnson,[36]
A. Jonckheere,[36] H. Jöstlein,[36] A. Juste,[36] W. Kahl,[44] S. Kahn,[55] E. Kajfasz,[10]
A.M. Kalinin,[23] D. Karmanov,[25] D. Karmgard,[41] R. Kehoe,[50] A. Khanov,[44]
A. Kharchilava,[41] S.K. Kim,[18] B. Klima,[36] B. Knuteson,[30] W. Ko,[31] J.M. Kohli,[15]
A.V. Kostritskiy,[26] J. Kotcher,[55] B. Kothari,[52] A.V. Kotwal,[52] A.V. Kozelov,[26]
E.A. Kozlovsky,[26] J. Krane,[42] M.R. Krishnaswamy,[17] P. Krivkova,[6] S. Krzywdzinski,[36]
M. Kubantsev,[44] S. Kuleshov,[24] Y. Kulik,[54] S. Kunori,[46] A. Kupco,[7] V.E. Kuznetsov,[34]
G. Landsberg,[58] W.M. Lee,[35] A. Leflat,[25] C. Leggett,[30] F. Lehner,[36,*] J. Li,[59] Q.Z. Li,[36]
X. Li,[4] J.G.R. Lima,[3] D. Lincoln,[36] S.L. Linn,[35] J. Linnemann,[50] R. Lipton,[36] A. Lucotte,[9]

L. Lueking,[36] C. Lundstedt,[51] C. Luo,[40] A.K.A. Maciel,[38] R.J. Madaras,[30] V.L. Malyshev,[23] V. Manankov,[25] H.S. Mao,[4] T. Marshall,[40] M.I. Martin,[38] K.M. Mauritz,[42] B. May,[39] A.A. Mayorov,[40] R. McCarthy,[54] T. McMahon,[56] H.L. Melanson,[36] M. Merkin,[25] K.W. Merritt,[36] C. Miao,[58] H. Miettinen,[61] D. Mihalcea,[38] C.S. Mishra,[36] N. Mokhov,[36] N.K. Mondal,[17] H.E. Montgomery,[36] R.W. Moore,[50] M. Mostafa,[1] H. da Motta,[2] E. Nagy,[10] F. Nang,[29] M. Narain,[47] V.S. Narasimham,[17] N.A. Naumann,[21] H.A. Neal,[49] J.P. Negret,[5] S. Negroni,[10] T. Nunnemann,[36] D. O'Neil,[50] V. Oguri,[3] B. Olivier,[12] N. Oshima,[36] P. Padley,[61] L.J. Pan,[39] K. Papageorgiou,[37] A. Para,[36] N. Parashar,[48] R. Partridge,[58] N. Parua,[54] M. Paterno,[53] A. Patwa,[54] B. Pawlik,[22] J. Perkins,[59] O. Peters,[20] P. Pétroff,[11] R. Piegaia,[1] B.G. Pope,[50] E. Popkov,[47] H.B. Prosper,[35] S. Protopopescu,[55] M.B. Przybycien,[39] J. Qian,[49] R. Raja,[36] S. Rajagopalan,[55] E. Ramberg,[36] P.A. Rapidis,[36] N.W. Reay,[44] S. Reucroft,[48] M. Ridel,[11] M. Rijssenbeek,[54] F. Rizatdinova,[44] T. Rockwell,[50] M. Roco,[36] C. Royon,[13] P. Rubinov,[36] R. Ruchti,[41] J. Rutherfoord,[29] B.M. Sabirov,[23] G. Sajot,[9] A. Santoro,[2] L. Sawyer,[45] R.D. Schamberger,[54] H. Schellman,[39] A. Schwartzman,[1] N. Sen,[61] E. Shabalina,[37] R.K. Shivpuri,[16] D. Shpakov,[48] M. Shupe,[29] R.A. Sidwell,[44] V. Simak,[7] H. Singh,[34] J.B. Singh,[15] V. Sirotenko,[36] P. Slattery,[53] E. Smith,[57] R.P. Smith,[36] R. Snihur,[39] G.R. Snow,[51] J. Snow,[56] S. Snyder,[55] J. Solomon,[37] Y. Song,[59] V. Sorín,[1] M. Sosebee,[59] N. Sotnikova,[25] K. Soustruznik,[6] M. Souza,[2] N.R. Stanton,[44] G. Steinbrück,[52] R.W. Stephens,[59] F. Stichelbaut,[55] D. Stoker,[33] V. Stolin,[24] A. Stone,[45] D.A. Stoyanova,[26] M.A. Strang,[59] M. Strauss,[57] M. Strovink,[30] L. Stutte,[36] A. Sznajder,[3] M. Talby,[10] W. Taylor,[54] S. Tentindo-Repond,[35] S.M. Tripathi,[31] T.G. Trippe,[30] A.S. Turcot,[55] P.M. Tuts,[52] V. Vaniev,[26] R. Van Kooten,[40] N. Varelas,[37] L.S. Vertogradov,[23] F. Villeneuve-Seguier,[10] A.A. Volkov,[26] A.P. Vorobiev,[26] H.D. Wahl,[35] H. Wang,[39] Z.-M. Wang,[54] J. Warchol,[41] G. Watts,[63] M. Wayne,[41] H. Weerts,[50] A. White,[59] J.T. White,[60] D. Whiteson,[30] J.A. Wightman,[42] D.A. Wijngaarden,[21] S. Willis,[38] S.J. Wimpenny,[34] J. Womersley,[36] D.R. Wood,[48] Q. Xu,[49] R. Yamada,[36] P. Yamin,[55] T. Yasuda,[36] Y.A. Yatsunenko,[23] K. Yip,[55] S. Youssef,[35] J. Yu,[36] Z. Yu,[39] M. Zanabria,[5] X. Zhang,[57] H. Zheng,[41] B. Zhou,[49] Z. Zhou,[42] M. Zielinski,[53] D. Zieminska,[40] A. Zieminski,[40] V. Zutshi,[55] E.G. Zverev,[25] and A. Zylberstejn[13]

(DØ Collaboration)

[1]*Universidad de Buenos Aires, Buenos Aires, Argentina*
[2]*LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[3]*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[4]*Institute of High Energy Physics, Beijing, People's Republic of China*
[5]*Universidad de los Andes, Bogotá, Colombia*
[6]*Charles University, Center for Particle Physics, Prague, Czech Republic*
[7]*Institute of Physics, Academy of Sciences, Center for Particle Physics, Prague, Czech Republic*
[8]*Universidad San Francisco de Quito, Quito, Ecuador*
[9]*Institut des Sciences Nucléaires, IN2P3-CNRS, Universite de Grenoble 1, Grenoble, France*
[10]*CPPM, IN2P3-CNRS, Université de la Méditerranée, Marseille, France*
[11]*Laboratoire de l'Accélérateur Linéaire, IN2P3-CNRS, Orsay, France*
[12]*LPNHE, Universités Paris VI and VII, IN2P3-CNRS, Paris, France*
[13]*DAPNIA/Service de Physique des Particules, CEA, Saclay, France*

[14] *Universität Mainz, Institut für Physik, Mainz, Germany*

[15] *Panjab University, Chandigarh, India*

[16] *Delhi University, Delhi, India*

[17] *Tata Institute of Fundamental Research, Mumbai, India*

[18] *Seoul National University, Seoul, Korea*

[19] *CINVESTAV, Mexico City, Mexico*

[20] *FOM-Institute NIKHEF and University of Amsterdam/NIKHEF, Amsterdam, The Netherlands*

[21] *University of Nijmegen/NIKHEF, Nijmegen, The Netherlands*

[22] *Institute of Nuclear Physics, Kraków, Poland*

[23] *Joint Institute for Nuclear Research, Dubna, Russia*

[24] *Institute for Theoretical and Experimental Physics, Moscow, Russia*

[25] *Moscow State University, Moscow, Russia*

[26] *Institute for High Energy Physics, Protvino, Russia*

[27] *Lancaster University, Lancaster, United Kingdom*

[28] *Imperial College, London, United Kingdom*

[29] *University of Arizona, Tucson, Arizona 85721*

[30] *Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720*

[31] *University of California, Davis, California 95616*

[32] *California State University, Fresno, California 93740*

[33] *University of California, Irvine, California 92697*

[34] *University of California, Riverside, California 92521*

[35] *Florida State University, Tallahassee, Florida 32306*

[36] *Fermi National Accelerator Laboratory, Batavia, Illinois 60510*

[37] *University of Illinois at Chicago, Chicago, Illinois 60607*

[38] *Northern Illinois University, DeKalb, Illinois 60115*

[39] *Northwestern University, Evanston, Illinois 60208*

[40] *Indiana University, Bloomington, Indiana 47405*

[41] *University of Notre Dame, Notre Dame, Indiana 46556*

[42] *Iowa State University, Ames, Iowa 50011*

[43] *University of Kansas, Lawrence, Kansas 66045*

[44] *Kansas State University, Manhattan, Kansas 66506*

[45] *Louisiana Tech University, Ruston, Louisiana 71272*

[46] *University of Maryland, College Park, Maryland 20742*

[47] *Boston University, Boston, Massachusetts 02215*

[48] *Northeastern University, Boston, Massachusetts 02115*

[49] *University of Michigan, Ann Arbor, Michigan 48109*

[50] *Michigan State University, East Lansing, Michigan 48824*

[51] *University of Nebraska, Lincoln, Nebraska 68588*

[52] *Columbia University, New York, New York 10027*

[53] *University of Rochester, Rochester, New York 14627*

[54] *State University of New York, Stony Brook, New York 11794*

[55] *Brookhaven National Laboratory, Upton, New York 11973*

[56] *Langston University, Langston, Oklahoma 73050*

[57] *University of Oklahoma, Norman, Oklahoma 73019*

[58] *Brown University, Providence, Rhode Island 02912*

[59]*University of Texas, Arlington, Texas 76019*
[60]*Texas A&M University, College Station, Texas 77843*
[61]*Rice University, Houston, Texas 77005*
[62]*University of Virginia, Charlottesville, Virginia 22901*
[63]*University of Washington, Seattle, Washington 98195*

Abstract

The DØ Collaboration has studied for the first time the properties of hadron-collider jets reconstructed with a successive-combination algorithm based on relative transverse momenta $(k_\perp)$ of energy clusters. Using the standard value $D = 1.0$ of the jet-separation parameter in the $k_\perp$ algorithm, we find that the p_T of such jets is higher than the E_T of matched jets reconstructed with cones of radius $\mathcal{R} = 0.7$, by about 5 (8) GeV at $p_T \approx 90$ (240) GeV. To examine internal jet structure, the $k_\perp$ algorithm is applied within $D = 0.5$ jets to resolve any subjets. The multiplicity of subjets in jet samples at $\sqrt{s} = 1800$ GeV and 630 GeV is extracted separately for gluons (M_g) and quarks (M_q), and the ratio of average subjet multiplicities in gluon and quark jets is measured as $\frac{\langle M_g \rangle - 1}{\langle M_q \rangle - 1} = 1.84 \pm 0.15$ (stat.)$\pm^{0.22}_{0.18}$ (sys.). This ratio is in agreement with the expectations from the HERWIG Monte Carlo event generator and a resummation calculation, and with observations in e^+e^- annihilations, and is close to the naive prediction for the ratio of color charges of $C_A/C_F = 9/4 = 2.25$.

I. INTRODUCTION

The production of gluons and quarks in high-energy collisions, and their development into the jets of particles observed in experiments, is usually described by the theory of Quantum Chromodynamics (QCD). In perturbative QCD, a produced parton (gluon or quark) emits gluon radiation, with each subsequent emission carrying off a fraction of the original parton's energy and momentum. The probability for a gluon to radiate a gluon is proportional to the color factor $C_A = 3$, while gluon radiation from a quark is proportional to the color factor $C_F = 4/3$. In the asymptotic limit, in which the radiated gluons carry a small fraction of the original parton's momentum, and neglecting the splitting of gluons to quark-antiquark pairs (whose probability is proportional to the color factor $T_R = 1/2$), the average number of objects radiated by a gluon is expected to be a factor $C_A/C_F = 9/4$ higher than the number of objects radiated by a quark [1]. And, in general, it is expected that a gluon will yield more particles with a softer momentum distribution, relative to a quark [2,3].

Although gluon jets are expected to dominate the final state of proton-antiproton ($p\bar{p}$) collisions at high energies, quark jets make up a significant fraction of the jet cross section at high $x_T = \frac{2p_T}{\sqrt{s}}$, where $\sqrt{s}$ is the total energy of the $p\bar{p}$ system, and p_T is the jet momentum transverse to the hadron-beam direction. The ability to distinguish gluon jets from quark jets would provide a powerful tool in the study of hadron-collider physics. To date, however, there has been only little experimental verification that gluon jets produced in hadron collisions display characteristics different from quark jets [4–8]. For fixed p_T, we analyze the internal structure of jets at $\sqrt{s} = 1800$ GeV and 630 GeV by resolving jets within jets (*sub-jets*) [7,9–16]. Using the expected fractions of gluon and quark jets at each $\sqrt{s}$, we measure the multiplicity of subjets in gluon and in quark jets. The results are presented as a ratio of average multiplicities $r = \frac{\langle M_g \rangle - 1}{\langle M_q \rangle - 1}$ of subjets in gluon jets to quark jets. This measured ratio is compared to that observed in e^+e^- annihilations [13], to predictions of a resummed calculation [11,14,16], and to the HERWIG [17] Monte Carlo generator of jet events.

The DØ detector [18], described briefly in Sec. II, is well-suited to studying properties of jets. A jet algorithm associates the large number of particles produced in a hard-scattering process with the quarks and gluons of QCD. We define jets with a successive-combination algorithm [19–22] based on relative transverse momenta ($k_\perp$) of energy clusters, described in Sec. III. In this paper, we present the first measurement of jet properties using the $k_\perp$ (sometimes written k_T) algorithm at a hadron collider. The momentum calibration of jets in the $k_\perp$ algorithm is outlined in Sec. III C, followed in Sec. III D by a simple comparison with jets defined with the fixed-cone algorithm. To study jet structure, the $k_\perp$ algorithm is then applied within the jet to resolve subjets, as described in Sec. III E. In e^+e^- annihilations, the number of subjets in gluon jets was shown to be larger than in quark jets [13]. In $p\bar{p}$ collisions, identifying gluon and quark jets is more complicated than in e^+e^- annihilations. We approach this issue by comparing jet samples at $\sqrt{s} = 1800$ GeV and 630 GeV, with the samples described in Sec. IV. For moderate jet p_T (55 – 100 GeV), the $\sqrt{s} = 1800$ GeV sample is gluon-enriched, and the $\sqrt{s} = 630$ GeV sample is quark-enriched. Section IV D describes a simple method developed to extract the separate subjet multiplicity for gluon and for quark jets. The method does not tag individual jets, but instead, we perform a statistical analysis of the samples at $\sqrt{s} = 1800$ GeV and 630 GeV. The method requires

the relative mix of quarks and gluons in the two data samples, which is derived from a Monte Carlo event generator that uses the parton distribution functions [23,24], measured primarily in deep inelastic scattering. Subsequent sections describe the measurement of the subjet multiplicity in DØ data and Monte Carlo simulations, the corrections used in the procedure, and the sources of systematic uncertainty. We conclude with comparisons to previous experimental and theoretical studies.

II. DØ DETECTOR

DØ is a multipurpose detector designed to study $p\bar{p}$ collisions at the Fermilab Tevatron Collider. A full description of the DØ detector can be found in Ref. [18]. The primary detector components for jet measurements at DØ are the excellent compensating calorimeters. The DØ calorimeters use liquid-argon as the active medium to sample the ionization energy produced in electromagnetic and hadronic showers. The elements of the calorimeter systems are housed in three cryostats. The central calorimeter (CC) covers the region $|\eta| < 1.0$, while the symmetric end calorimeters (EC) extend coverage to $|\eta| < 4.2$, where the pseudorapidity $\eta = -\ln \tan \theta/2$ is defined in terms of the polar angle θ with respect to the proton-beam direction z. Each system is divided into an electromagnetic (EM), fine hadronic (FH), and coarse hadronic (CH) sections. The EM and FH use uranium absorber plates as the passive medium, and the CH uses either copper (CC) or stainless steel (EC). Copper readout pads are centered in the liquid-argon gaps between the absorber plates. Radially, the electromagnetic sections are 21 radiation lengths deep, divided into 4 readout layers. The hadronic calorimeters are 7–11 nuclear interaction lengths deep, with up to 4 layers. The entire calorimeter is segmented into towers, of typical size $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, projected towards the nominal $p\bar{p}$ interaction point in the center of the detector, where ϕ is the azimuthal angle about the z axis. Figure 1 shows a schematic view of one quadrant of the DØ calorimeter in the $r - z$ plane, where r is the distance to the origin in the plane transverse to the beam axis. Each layer in a calorimeter tower is called a cell, and yields an individual energy sampling. Energy deposited in the calorimeters by particles from $p\bar{p}$ collisions are used to reconstruct jets. The transverse energy resolution of jets for data at $\sqrt{s} = 1800$ GeV can be parameterized as [25]:

$$(\sigma(E_T)/E_T)^2 \approx 6.9/E_T^2 + 0.5/E_T + 0.001, \tag{2.1}$$

with E_T in GeV.

In the analysis of jet structure, we are interested in the distribution of energy within jets. Apart from the energy of particles produced in a hard-scattering event, the cells of the DØ calorimeter are sensitive to three additional sources of energy that contribute to a jet. The first, called uranium noise, is a property of the detector material. The decay of radioactive uranium nuclei in the calorimeter can produce energy in a given cell, even in the absence of a particle flux. For each cell, a distribution of this pedestal energy is measured in a series of calibration runs without beams in the accelerator. The pedestal distribution due to uranium noise is asymmetric, with a longer high-end tail, as illustrated in Fig. 2. During normal data-taking, the mean pedestal energy is subtracted online from the energy measured in a hard-scattering event. To save processing time and reduce the event size, a
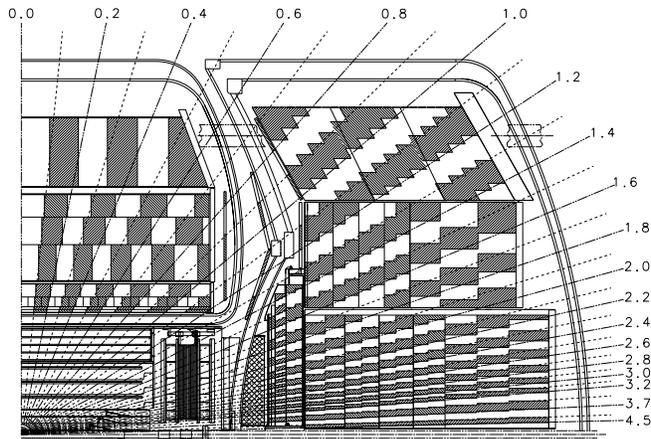
FIG. 1. One quadrant of the DØ calorimeter and drift chambers, projected in the $r-z$ plane. Radial lines illustrate the detector pseudorapidity and the pseudo-projective geometry of the calorimeter towers. Each tower has size $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$.

zero-suppression circuit is used, whereby cells containing energy within a symmetric window about the mean pedestal count are not read out. Since the pedestal distribution of each cell is asymmetric, zero-suppression causes upward fluctuations in measured cell energies more often than downward fluctuations. In the measurement of a hard-scattering event, the net impact is an increased multiplicity of readout cells and a positive offset to their initial energies.

There are two other environmental effects that contribute to the energy offset of calorimeter cells. The first is extra energy from multiple $p\bar{p}$ interactions in the same accelerator-bunch crossing, and this depends on the instantaneous luminosity. To clarify the second effect, called pile-up, we turn to how calorimeter cells are sampled, as is illustrated in Fig. 3. The maximum drift time for ionization electrons produced in the liquid-argon to reach the copper readout pad of a calorimeter cell is about 450 ns. The collected electrons produce an electronic signal that is sampled at the time of the bunch crossing (base), and again 2.2 μs later (peak). The difference in voltage between the two samples (peak relative to base) defines the initial energy count in a given cell. Because the signal fall-time ($\sim 30~\mu$s) is much longer than the accelerator bunch spacing (3.5 μs), the base and peak voltages are measured with respect to a reference level that depends on previous bunch crossings. The signal from the current bunch crossing is therefore piled on top of the slowly decaying signal from previous crossings. When a previous bunch crossing leaves energy in a particular cell, that cell's energy count will therefore be reduced on average, after the baseline subtraction.

III. $k_\perp$ JET ALGORITHM

Jet algorithms assign any *particles* produced in high-energy collisions to jets. The *particles* correspond to observed energy depositions in a calorimeter, or to final state particles generated in a Monte Carlo event. Typically, such objects are first organized into preclusters (defined below), before being processed through the jet algorithms: The jet algorithms
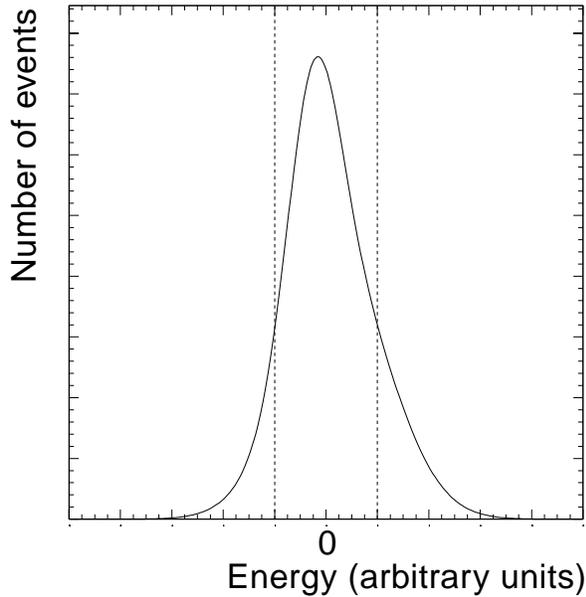
FIG. 2. Illustration of the pedestal energy distribution in a calorimeter cell (solid line), stemming from uranium noise. The mean value is defined to be zero, and the peak occurs at negative values. Removal of the portion between the vertical dashed lines (a symmetric window about the mean) yields in a positive mean for the remaining distribution.
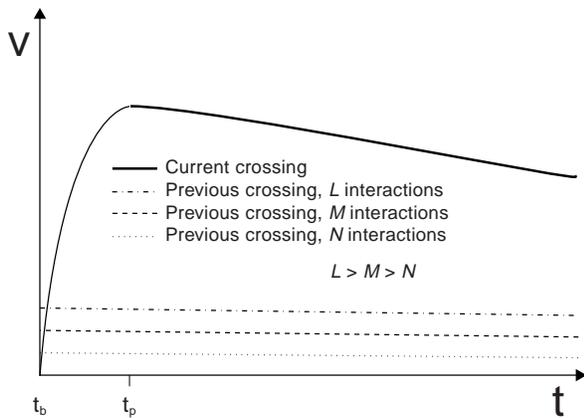

FIG. 3. Schematic of signal voltage in a calorimeter cell as a function of time. The solid line represents the contribution for a given event (the "current" $p\bar{p}$ bunch crossing). In the absence of previous bunch crossings, the cell is sampled correctly at t_b, just before a crossing, to establish a base voltage, and at t_p, to establish a peak voltage. The voltage difference $\Delta V = V(t_p) - V(t_b)$ is proportional to the initial energy deposited in the cell. The dashed lines show example contributions from a previous bunch crossing containing three different numbers of $p\bar{p}$ interactions. The observed signal is the sum of the signals from the current and previous crossings. (The figure is not to scale.)

therefore do not depend on the nature of the particles. We discuss two jet algorithms in this paper: the $k_\perp$ and cone jet algorithms, with emphasis on the former.

In the $k_\perp$ jet algorithm, pairs of particles are merged successively into jets, in an order corresponding to increasing relative transverse momentum. The algorithm contains a single parameter D (often called R in some references), which controls the cessation of merging. Every particle in the event is assigned to a single $k_\perp$ jet.

In contrast, the fixed-cone algorithm [27] associates into a jet all particles with trajectories within an area $A = \pi\mathcal{R}^2$, where the parameter $\mathcal{R}$ is the radius of a cone in (η, ϕ) space. The DØ fixed-cone algorithm [25,28] is an iterative algorithm, starting with cones centered on the most energetic particles in the event (called seeds). The energy-weighted centroid of a cone is defined by:

$$\eta^C = \frac{\sum_i E_T^i \eta^i}{\sum_i E_T^i}, \phi^C = \frac{\sum_i E_T^i \phi^i}{\sum_i E_T^i}, \tag{3.1}$$

where the sum is over all particles i in the cone. The centroids are used iteratively as centers for new cones in (η, ϕ) space. A jet axis is defined when a cone's centroid and geometric center coincide. Unfortunately, the fixed-cone jet algorithm allows cones to overlap, and any single particle can belong to two or more jets. A second parameter, and additional steps, are needed to determine if overlapping cones should be split or merged [29].

The $k_\perp$ jet algorithm offers several benefits over the fixed-cone jet algorithms, which are widely used at hadron colliders. Theoretically, the $k_\perp$ algorithm is infrared-safe and collinear-safe to all orders of calculation [19,28]. The same algorithm can be applied to partons generated from fixed-order or resummation calculations in QCD, particles in a Monte Carlo event generator, or tracks or energy depositions in a detector.

The $k_\perp$ jet algorithm is specified in Sec. III A. In Sec. III B, we describe the preclustering algorithm, the goal of which is to reduce the detector-dependent aspects of jet clustering (e.g., energy thresholds or calorimeter segmentation). The momentum calibration of $k_\perp$ jets is presented in Sec. III C. In Sec. III D, jets reconstructed using the $k_\perp$ algorithm are compared to jets reconstructed with the fixed-cone algorithm. In Sec. III E, we indicate how subjets are defined in the $k_\perp$ algorithm.

A. Jet clustering

There are several variants of the $k_\perp$ jet-clustering algorithm for hadron colliders [19–21]. The main differences concern how particles are merged together and when the clustering stops. The different types of merging, or recombination, schemes were investigated in Ref. [19]. DØ chooses the scheme that corresponds to four-vector addition of momenta, because [28]:

1. it is conceptually simple;

2. it corresponds to the scheme used in the $k_\perp$ algorithm in $e^+ e^-$ annihilations [13];

3. it has no energy defect [30], a measure of perturbative stability in the analysis of transverse energy density within jets; and

4. it is better suited [31] to the missing transverse energy calculation in the jet-momentum calibration method used by DØ.

To stop clustering, DØ has adopted the proposal [21] that halts clustering when all the jets are separated by $\Delta \mathcal{R} > D$. This rule is simple, and maintains a similarity with cone algorithms for hadronic collisions. The value $D = 1.0$ treats initial-state radiation in the same way as final-state radiation [11,32].

The jet algorithm starts with a list of preclusters as defined in the next section. Initially, each precluster is assigned a momentum four-vector $(E, \mathbf{p}) = E_{\mathrm{precluster}}(1, \sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta)$, written in terms of the precluster angles θ and ϕ. The execution of the jet algorithm involves:

1. Defining for each object i in the list:

$$d_{ii} \equiv p_{T,i}^2 = p_{x,i}^2 + p_{y,i}^2,$$

and for each pair (i,j) of objects:

$$
\begin{aligned}
d_{ij} &\equiv \min\left[p_{T,i}^2, p_{T,j}^2\right] \frac{\Delta \mathcal{R}_{ij}^2}{D^2} \\
&= \min\left[p_{T,i}^2, p_{T,j}^2\right] \frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{D^2},
\end{aligned}
\tag{3.2}
$$

where D is the stopping parameter of the jet algorithm. For $D = 1.0$ and $\Delta \mathcal{R}_{ij} \ll 1$, d_{ij} reduces to the square of the relative transverse momentum $(k_\perp)$ between objects.

2. If the minimum of all possible d_{ii} and d_{ij} is a d_{ij}, then replacing objects i and j by their merged object $(E_{ij}, \mathbf{p}_{ij})$, where

$$E_{ij} = E_i + E_j$$

$$\mathbf{p}_{ij} = \mathbf{p}_i + \mathbf{p}_j.$$

And if the minimum is a d_{ii}, then removing object i from the list and defining it to be a jet.

3. Repeating Steps 1 and 2 when there are any objects left in the list.

The algorithm produces a list of jets, each separated by $\Delta \mathcal{R} > D$. Figure 4 illustrates how the $k_\perp$ algorithm successively merges the particles in a simplified diagram of a hadron collision.

B. Preclustering

In the computer implementation of the $k_\perp$ jet algorithm, the processing time is proportional to N^3, where N is the number of particles (or energy signals) in the event [19]. The zero-suppression circuit reduces the number of calorimeter cells that have to be read out in each event. To reduce this further, we employ a preclustering algorithm. The procedure assigns calorimeter cells (or particles in a Monte Carlo event generator) to preclusters, suitable for input to the jet-clustering algorithm. In essence, calorimeter cells are collapsed into towers, and towers are merged if they are close together in (η, ϕ) space or if they have small
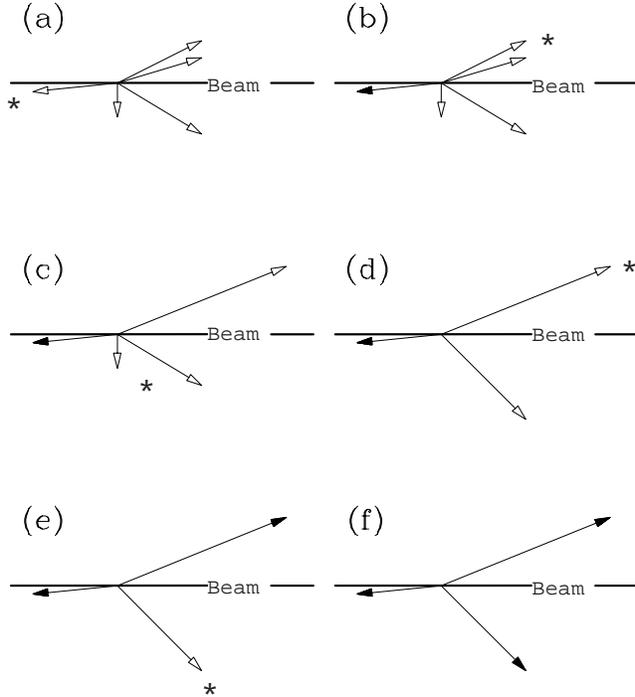
FIG. 4. A simplified example of the final state of a collision between two hadrons. (a) The particles in the event (represented by arrows) comprise a list of objects. (b-f) Solid arrows represent the final jets reconstructed by the $k_\perp$ algorithm, and open arrows represent objects not yet assigned to jets. The five diagrams show successive iterations of the algorithm. In each diagram, a jet is either defined (when it is well-separated from all other objects), or two objects are merged (when they have small relative $k_\perp$). The asterisk labels the relevant object(s) at each step.

p_T. Monte Carlo studies have shown that such preclustering reduces the impact of ambiguities due to calorimeter showering and finite segmentation, especially on the reconstructed internal jet substructure. For example, when a single particle strikes the boundary between two calorimeter towers, it can produce two clusters of energy. Conversely, two collinear particles will often shower in a single calorimeter tower. In both cases, there is a potential discrepancy in the number of energy clusters found at the calorimeter level and the particle level. Preclustering at both the calorimeter and at the particle level within a radius larger than the calorimeter segmentation integrates over such discrepancies.

The preclustering algorithm consists of the following six steps:

1. Starting from a list of populated calorimeter cells in an event, remove from any cells with $E_T < -0.5$ GeV. Cells with such negative E_T — rarely observed in minimum-bias[1] events (see Fig. 5) — are considered spurious.

2. For each calorimeter cell centered at some (θ, ϕ) relative to the primary interaction

---

[1]The minimum-bias trigger requires a coincidence signal in the scintillating-tile hodoscopes [18] located near the beampipe.
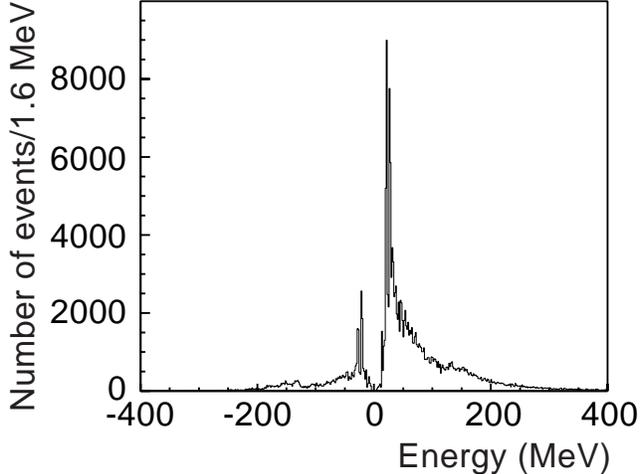
FIG. 5. Mean energies in calorimeter cells for a sample of minimum-bias events. The contribution from instrumental effects is included, which occasionally leads to negative energy readings. For each cell, the energy distribution is fitted to a Gaussian, as illustrated in Fig. 2. Before readout, the zero-suppression circuit in each cell's electronics sets to zero energy the channels in a symmetric window about the mean pedestal. These channels are not read out, causing the dip observed near zero.

vertex, define its pseudorapidity:

$$\eta = -\ln\tan\frac{\theta}{2}.$$

3. For each calorimeter tower t, sum the transverse energy of cells c in that tower:

$$E_T^t = \sum_{c\,\epsilon\,t} E_c \sin\theta_c,$$

where E_c is the energy deposited in cell c.

4. Starting at the extreme negative value of η and $\phi = 0$, combine any neighboring towers into preclusters such that no two preclusters are within $\Delta\mathcal{R}^{\mathrm{pre}} = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.2$. The combination follows the Snowmass prescription [27]:

$$E_T = E_{T,i} + E_{T,j}$$

$$\eta = \frac{E_{T,i}\eta_i + E_{T,j}\eta_j}{E_{T,i} + E_{T,j}}$$

$$\phi = \frac{E_{T,i}\phi_i + E_{T,j}\phi_j}{E_{T,i} + E_{T,j}}.$$

The procedure evolves in the direction of increasing ϕ, and then increasing η.

5. Because of pile-up in the calorimeter, precluster energies can fluctuate in both positive and negative directions. Preclusters that have negative transverse energy $E_T = E_{T-} < 0$,
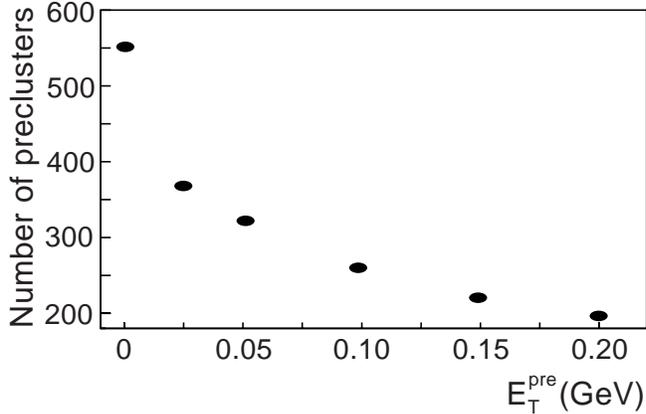
FIG. 6. The mean number of preclusters per event, as a function of the setting of minimum transverse energy required for preclusters (E_T^{pre}).

are redistributed to k neighboring preclusters in the following way. Given a negative E_T precluster with (E_{T-}, η_-, ϕ_-), we define a square S of size $(\eta_- \pm 0.1) \times (\phi_- \pm 0.1)$. When the following holds:

$$\sum_{k \, \epsilon \, S} E_{T,k}(\eta, \phi) > |E_{T-}|, \qquad (3.3)$$

where preclusters only with positive E_T that are located within the square S are included in the sum, then E_{T-} is redistributed to the positive preclusters in the square, with each such precluster k absorbing a fraction

$$\frac{E_{T,k}}{\displaystyle\sum_{k \, \epsilon \, S} E_{T,k}}$$

of the negative E_T. If Eq. (3.3) is not satisfied, the "search square" is increased in steps of $\Delta\eta = \pm 0.1$ and $\Delta\phi = \pm 0.1$, and another redistribution is attempted. If redistribution still fails for a square of $(\eta_- \pm 0.7) \times (\phi_- \pm 0.7)$, the negative energy precluster is isolated in the calorimeter and ignored (by setting $E_{T-} = 0$).

6. Preclusters with $0 < E_T < E_T^{\mathrm{pre}} = 0.2$ GeV, are redistributed to neighboring preclusters, as specified in Step 5. To reduce the overall number of preclusters, we also require that the search square have at least three positive E_T preclusters. The threshold E_T^{pre} was tuned to produce about 200 preclusters/event (see Fig. 6), in order to fit our constraints for processing time.

C. Calibration of jet momentum

A correct calibration of jet momentum reduces overall experimental uncertainties on jet production. The calibration at DØ also accounts for the contribution of the underlying event (momentum transferred as a result of the soft interactions between the remnant partons of

the proton and antiproton). All such corrections enter in the relation between the momentum of a jet measured in the calorimeter p^{meas} and the "true" jet momentum p^{true} [33]

$$p_{\text{jet}}^{\text{true}} = \frac{p_{\text{jet}}^{\text{meas}} - p_O(\eta^{\text{jet}}, \mathcal{L}, p_T^{\text{jet}})}{R_{\text{jet}}(\eta^{\text{jet}}, p^{\text{jet}})} \qquad (3.4)$$

where p_O denotes an offset correction, and R_{jet} is a correction for the response of the calorimeter to jets. A true jet is defined as being composed of only the final-state particle momenta from the hard parton-parton scatter (i.e., before interaction in the calorimeter). Although Eq. (3.4) is valid for any jet algorithm, p_O and the components of R_{jet} depend on the details of the jet algorithm. Our calibration procedure attempts to correct calorimeter-level jets (after interactions in the calorimeter) to their particle-level (before the individual particles interact in the calorimeter), using the described $k_\perp$ jet algorithm, with $D = 1.0$. The procedure follows closely that of calibration of the fixed-cone jet algorithm [33]. The fixed-cone jet algorithm requires an additional scale factor in Eq. (3.4), but we find no need for that kind of calorimeter-showering correction in the $k_\perp$ jet momentum calibration [31].

The offset p_O corresponds to the contribution to the momentum of a reconstructed jet that is not associated with the hard interaction. It contains two parts:

$$p_O = O_{\text{ue}} + O_{\text{zb}},$$

where O_{ue} is the offset due to the underlying event, and O_{zb} is an offset due to the overall detector environment. O_{zb} is attributed to any additional energy in the calorimeter cells of a jet from the combined effects of uranium noise, multiple interactions, and pile-up. The contributions of O_{ue} and O_{zb} to $k_\perp$ jets are measured separately, but using similar methods. The method overlays DØ data and Monte Carlo events, as described in what follows.

The Monte Carlo events are generated by HERWIG (version 5.9) [17] with $2 \rightarrow 2$ parton p_T-thresholds of 30, 50, 75, 100, and 150 GeV, and the underlying-event contribution switched off. The Monte Carlo events are propagated through a GEANT-based [34] simulation of the DØ detector, which provides a cell-level simulation of the calorimeter response and resolution. These Monte Carlo events are then passed through the calorimeter-reconstruction and jet-finding packages, defining the initial sample of jets. Detector simulation does not include the effects of uranium noise nor of the accelerator conditions causing multiple interactions and pile-up. The total contribution from these three effects is modeled using zero-bias events, which correspond to observations at random $p\bar{p}$ bunch crossings. Zero-bias events were recorded by the DØ detector at different instantaneous luminosities in special data-taking runs without the zero-suppression discussed in Sec. II. The cell energies in zero-bias events are added cell-by-cell to the energies in simulated Monte Carlo jet events. The summed cell energies are then zero-suppressed offline, using the pedestals appropriate to the zero-bias running conditions. Finally, the summed cell energies are passed through the calorimeter-reconstruction and jet-finding packages, producing a second sample of jets. The two samples are compared on an event-by-event basis, associating the jets in events of the two samples that have their axes separated by $\Delta \mathcal{R} < 0.5$ [31]. The difference in the measured p_T of the corresponding matched jets is O_{zb}, and shown in Fig. 7 as a function of η^{jet}, for different instantaneous luminosities.
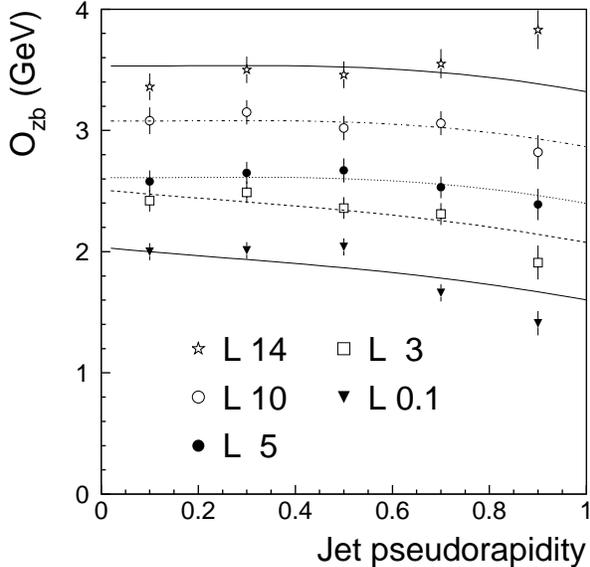
FIG. 7. The offset correction O_{zb} as a function of pseudorapidity of $k_\perp$ jet ($D = 1.0$). The offset O_{zb} accounts for the combined effects of pile-up, uranium noise, and multiple interactions. The different sets of points are for events with different instantaneous luminosity $\mathcal{L} \approx 14, 10, 5, 3, 0.1 \times 10^{30}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The curves are fits to the points at different $\mathcal{L}$, using the same functional form as employed for the cone algorithm in Ref. [33].

The event-overlay method was checked with the fixed-cone jet algorithm for $\mathcal{R} = 0.7$. For jets with 30 GeV $< E_T <$ 50 GeV, this method gives only 14% (28%) smaller offsets [$\Delta O_{\mathrm{zb}} = 0.25\ (0.39)$ GeV per jet], at $\mathcal{L} \approx 5\ (0.1) \times 10^{30}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ relative to Ref. [33]. Independent of jet E_T, the method used in Ref. [33] measures the E_T per unit $\Delta\eta \times \Delta\phi$ in zero-bias events, and scales the value by the area of the jet cone. In the event-overlay method, O_{zb} decreases by as much as 40% when the cone-jet transverse energy increases to 125 GeV $< E_T <$ 170 GeV. Approximately 30% of this decrease can be explained by the E_T^{jet}-dependence of the occupancy of cells within cone jets (the fraction of cells with significant energy deposition inside the cone). The remaining 70% of the O_{zb} dependence on jet E_T is assigned as a systematic uncertainty on our method. Since the observed dependence is less pronounced in the $k_\perp$ jet algorithm, this error amounts at most to 15% in the highest jet p_T bin. In addition, we include a systematic uncertainty of 0.2 GeV arising from the fits in Fig. 7. Using our overlay method for both algorithms, the offsets O_{zb} in the $k_\perp$ jet algorithm (with $D = 1.0$) are generally $50 - 75\%$ (or about 1 GeV per jet) larger than in the fixed-cone jet algorithm (with $\mathcal{R} = 0.7$) [31].

The offset due to the underlying event O_{ue} is modeled with minimum-bias events. A minimum-bias event is a zero-bias event with the additional requirement of a coincidence signal in the scintillating-tile hodoscopes [18] near the beampipe. The additional requirement means there was an inelastic $p\bar{p}$ collision during the bunch crossing. In addition to O_{ue}, a minimum-bias event in the DØ calorimeter includes energy from uranium noise, multiple interactions, and pile-up. However, the luminosity dependence of multiple interactions and pile-up in minimum-bias events is different than in zero-bias events. In the limit of very small
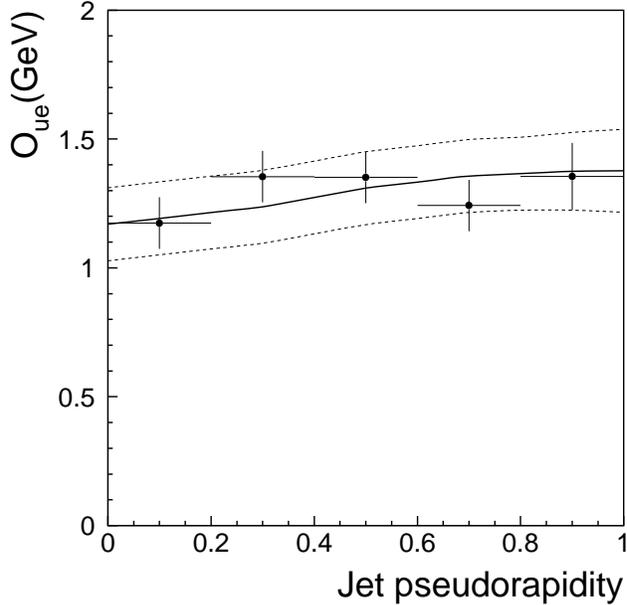
FIG. 8. The correction for underlying event O_{ue} as a function of $|\eta|$ for $k_\perp$ jets ($D = 1.0$). The solid curve is the fit of the results for the cone jet algorithm in Ref. [33] scaled to the results for the $k_\perp$ jet algorithm. The dashed curves denote the one standard deviation (s.d.) systematic error.

luminosity, these contributions are negligible, and a minimum-bias event at low luminosity therefore contains the offset due to the underlying event and uranium noise, while a zero-bias event at low luminosity has only the offset from uranium noise. To measure O_{ue}, we again compare two samples of jets. Minimum-bias events as measured by the DØ calorimeter at low luminosity are added to Monte Carlo jet events, where the resulting jets define the first sample of jets in the determination of O_{ue}. The second sample of jets is reconstructed from zero-bias events at low luminosity and also added to Monte Carlo jet events. On an event-by-event basis, O_{ue} is calculated by subtracting the momentum of jets in the second sample from the momentum of matching jets in the first sample. The underlying event offset O_{ue} for $k_\perp$ jets is shown in Fig. 8. Using this method for both algorithms, the offset O_{ue} for $k_\perp$ jets (with $D = 1.0$) is found to be approximately 30% larger than for the fixed-cone jet algorithm (with $\mathcal{R} = 0.7$).

DØ measures the jet momentum response based on conservation of p_T in photon-jet (γ-jet) events [33]. The electromagnetic energy/momentum scale is determined from the $Z, J/\psi \rightarrow e^+e^-$, and $\pi^0 \rightarrow \gamma\gamma \rightarrow e^+e^-e^+e^-$ data samples, using the known masses of these particles. For the case of a γ-jet two-body process, the jet momentum response can be characterized as:

$$R_{\mathrm{jet}} = 1 + \frac{\vec{\rlap{E}{/}}_T \cdot \hat{n}_{T\gamma}}{p_{T\gamma}}, \qquad (3.5)$$

where $p_{T\gamma}$ and $\hat{n}$ are the transverse momentum and direction of the photon, and $\vec{\rlap{E}{/}}_T$ is the missing transverse energy, defined as the negative of the vector sum of the transverse energies of the cells in the calorimeter. To avoid resolution and trigger biases, R_{jet} is binned in terms
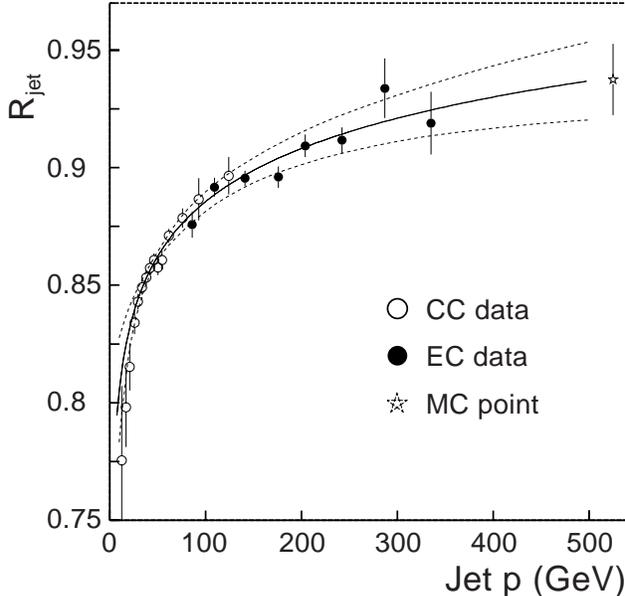
FIG. 9. The response correction for $k_\perp$ jets with $D = 1.0$, as a function of offset-corrected jet momentum. The Monte Carlo point ($\star$) is used to constrain the fit (solid) at high $p_{\rm jet}^{\rm meas}$. The dashed curves denote the ± 1 s.d. systematic error.

of $E' = p_{T\gamma}^{\rm meas} \cdot \cosh(\eta_{\rm jet})$ and then mapped onto the offset-corrected jet momentum. Thus, E' depends only on photon variables and jet pseudorapidity, which are quantities that are measured with very good resolution. $R_{\rm jet}$ and E' depend only on the jet position, which has little dependence on the type of jet algorithm employed.

$R_{\rm jet}$ as a function of $p_{\rm jet}^{\rm meas} - p_O$ for $k_\perp$ jets is shown in Fig. 9. The data points are fitted with the functional form $R_{\rm jet}(p) = a + b \ln(p) + c \, (\ln(p))^2$. The response $R_{\rm jet}$ for cone jets (with $\mathcal{R} = 0.7$) [33] and for $k_\perp$ jets ($D = 1.0$) is different by about 0.05. This difference does not have any physical meaning; it corresponds to different voltage-to-energy conversion factors at the cell level used in the reconstruction of jets.

D. Comparison of the $k_\perp$ jet algorithm to the cone jet algorithm

It is of interest to compare the momenta of $k_\perp$ jets to those of jets reconstructed with the DØ fixed-cone algorithm [28]. These results refer to the $k_\perp$ jet algorithm described above with $D = 1.0$ and corrected according to the prescriptions given in Sec. III C. The cone jets were reconstructed [25] with $\mathcal{R} = 0.7$ and corrected according to Ref. [33]. This comparison involves about 75% of the events in the 1994–1996 data that were used for the analysis of the inclusive cone-jet cross section at $\sqrt{s} = 1800$ GeV [35]. The two algorithms are similar by design [21], defining similar jet directions and momenta, at least for the two leading (highest p_T) jets in the event. The remaining jets in the event usually have much smaller p_T, making them more difficult to measure, and so we do not consider them here. The jets reconstructed by each algorithm are compared on an event-by-event basis, associating a cone jet with a $k_\perp$ jet if they are separated by $\Delta\mathcal{R} < 0.5$.
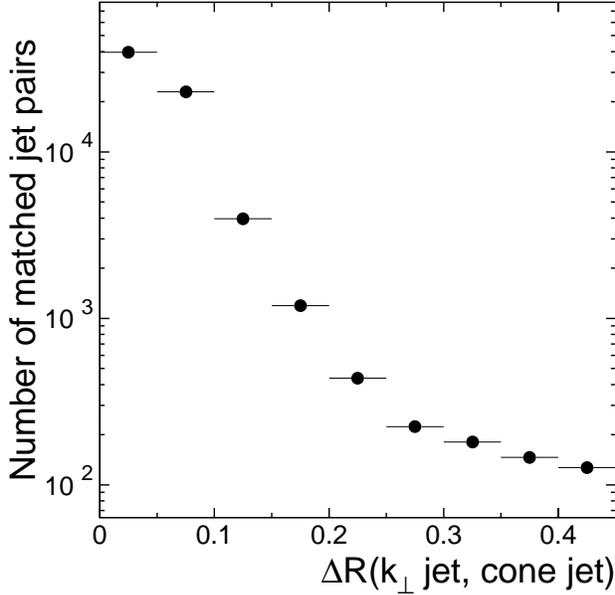
FIG. 10. The distance $\Delta\mathcal{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ between a $k_\perp$-jet axis and its matching cone-jet axis. The $k_\perp$ jets were reconstructed with $D = 1.0$, and the cone jets were reconstructed with $\mathcal{R} = 0.7$. Only the two leading jets from each algorithm were considered. The $k_\perp$ jets were selected with $|\eta| < 0.5$.

To obtain a sample of events with only good hadronic jets, the following requirements were placed on the events and on the leading two reconstructed $k_\perp$ jets. These criteria are based on standard jet quality requirements (to remove spurious clusters) in use at DØ for the fixed-cone jet algorithm [25]:

- Measured event vertex was required to be within 50 cm of the center of the detector.

- $|\vec{\not{E}}_T|$ was required to be less than 70% of the p_T of the leading jet.

- Fraction of jet p_T measured in the coarse hadronic calorimetry was required to be less than 40% of the total jet p_T.

- Fraction of jet p_T measured in the electromagnetic calorimetry was required to be between 5% and 95% of the total jet p_T.

- Jets were required to have $|\eta| < 0.5$.

These requirements yield a sample of 68946 $k_\perp$ jets. The axes of 99.94% of these jets are reconstructed within $\Delta\mathcal{R} < 0.5$ of a cone-jet axis, when the matching jet is one of the two leading cone jets in the event. For such pairs of jets, the distance between a $k_\perp$-jet axis and matching cone-jet axis is shown in Fig. 10. The fixed-cone algorithm finds a jet within $\Delta\mathcal{R} < 0.1$ of a $k_\perp$ jet 91% of the time. Figure 11 shows the difference $p_T(k_\perp \text{ jet}) - E_T(\text{cone jet})$ as a function of $p_T(k_\perp \text{ jet})$. Generally, the p_T of $k_\perp$ jets $(D = 1.0)$ is higher than the E_T of associated cone jets $(\mathcal{R} = 0.7)$. The difference increases approximately linearly with jet p_T,
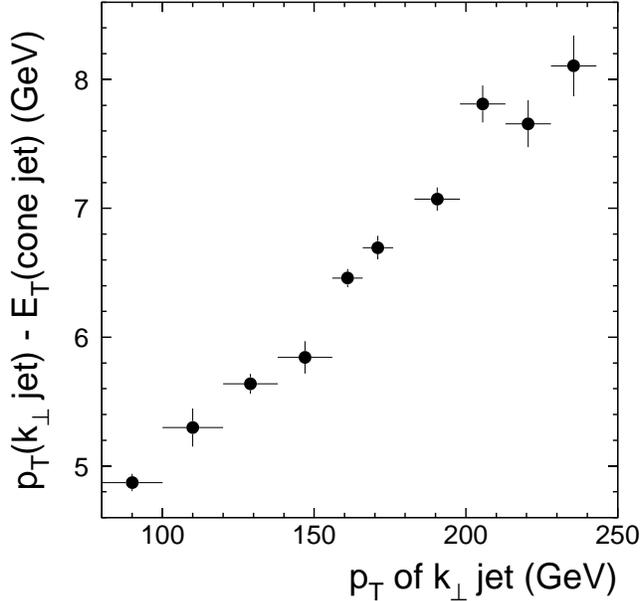
FIG. 11. The difference $p_T(k_\perp$ jet$) - E_T($cone jet$)$ as a function of the $k_\perp$ jet p_T. A cone jet is associated with a $k_\perp$ jet if their axes are separated by $\Delta\mathcal{R} < 0.5$.

from about 5 GeV (or 6%) at $p_T \approx 90$ GeV to about 8 GeV (or 3%) at $p_T \approx 240$ GeV. This may be explained by how the two algorithms deal with hadronization effects [26].

E. Subjets

The subjet multiplicity is a natural observable for characterizing a $k_\perp$ jet [19,20]. Subjets are defined by reapplying the $k_\perp$ algorithm, as in Sec. III A, starting with a list of preclusters assigned to a particular jet. Pairs of objects with the smallest d_{ij} are merged successively until all remaining pairs of objects have

$$d_{ij} = \min(p_{T,i}^2, p_{T,j}^2)\frac{\Delta\mathcal{R}_{ij}^2}{D^2} > y_{\text{cut}}p_T^2(\text{jet}), \tag{3.6}$$

where $p_T(\text{jet})$ is the p_T of the entire jet in the $k_\perp$ algorithm described above, and $0 \leq y_{\text{cut}} \leq 1$ is a dimensionless parameter. Objects satisfying Eq. (3.6) are called subjets, and the number of subjets is the subjet multiplicity M of a $k_\perp$ jet. For $y_{\text{cut}} = 1$, the entire jet consists of a single subjet ($M = 1$). As y_{cut} decreases, the subjet multiplicity increases, until every precluster becomes resolved as a separate subjet in the limit $y_{\text{cut}} \to 0$. Two subjets in a jet can be resolved when they are not collinear (i.e., well-separated in $\eta \times \phi$ space), or if they are both hard (i.e., carry a significant fraction of the jet p_T).

We now turn to the theoretical treatment of subjet multiplicity. Perturbative and re-summed calculations [11,16] and Monte Carlo estimates (see Sec. IV D) predict that gluon jets have a higher mean subjet multiplicity than quark jets. To understand the origin of this prediction, we consider first how a jet can contain multiple subjets. Clearly, at leading-order, $2 \to 2$ subprocesses yield $M = 1$. However, higher-order QCD radiation can increase the

average value of M. At next-to-leading order, there can be three partons in the final state of a $p\bar{p}$ collision. If two partons are clustered together into a jet, they can be resolved as distinct subjets ($M = 2$) for a sufficiently small choice of y_{cut}. For larger y_{cut}, the value of M depends on the magnitude and direction of the radiated third parton. In QCD, the radiation of a parton is governed by the DGLAP splitting functions [36]. The radiated third parton is usually soft and/or collinear with one of the other two partons, leading to jets with $M = 1$. However, hard or large-angle radiation, although rare, causes some jets to have $M = 2$. Consequently, when many jets are analyzed using some high y_{cut}, the two-subjet rate will yield $\langle M \rangle > 1$.

In the framework of parton showers, repeated application of DGLAP splitting provides jets with $M > 2$. Monte Carlo event generators incorporate parton showers into the initial and final states of a $2 \to 2$ hard scatter. Because of its larger color factor, a parton shower initiated by a gluon in the final state will tend to produce a jet with more subjets than one initiated by a quark. Similarly, a soft parton radiated in the initial state will tend to cluster with a hard final-state parton when $\Delta \mathcal{R} < D$. For the case of initial-state radiation, the subjet multiplicity depends weakly on whether the final-state partons in the $2 \to 2$ hard scatter are quarks or gluons. The contribution of initial-state radiation to the subjet multiplicity does, however, depend on $\sqrt{s}$. Initial-state radiation is treated on an equal footing as final-state radiation in the $k_\perp$ algorithm with $D = 1.0$ [11,32], and diminishes in importance as D decreases. In general, subsequent emissions in parton showers have less energy and momentum, and this structure is revealed at smaller y_{cut} values through an increase in the subjet multiplicity: $\langle M(y'_{\text{cut}}) \rangle > \langle M(y_{\text{cut}}) \rangle$, where $y'_{\text{cut}} < y_{\text{cut}}$.

Experimentally, the growth of M at very small y_{cut} is reduced by the granularity of the detector and by the preclustering algorithm. Theoretical predictions for M are therefore treated in the same way as the experimental measurements, i.e., by preclustering (as in Sec. III B). Requiring preclusters to be separated by $\Delta \mathcal{R}^{\text{pre}}$, means that the subjets nearest in (η, ϕ) space begin to be resolved for

$$y_{\text{cut}} < \left(\frac{\Delta \mathcal{R}^{\text{pre}}}{2D} \right)^2 \tag{3.7}$$

based solely on the fraction of p_T carried by the subjet in the jet. The factor $1/2$ corresponds to the maximum fraction of jet p_T carried by the softest subjet [see Eq. (3.6)]. The preclustering stage provides a comparison of the measurement of M with prediction in the interesting region of small y_{cut}, without an explicit correction for detector granularity.

The subjet analysis in this paper uses a single resolution parameter $y_{\text{cut}} = 10^{-3}$. For this y_{cut}, the minimum subjet p_T is approximately 3% of the total jet p_T, independent of the choice of the D parameter. Because y_{cut}, as defined by Eqs. (3.2) and (3.6), involves a ratio of subjet p_T to jet p_T, the subjet multiplicity is therefore not significantly sensitive to multiplicative changes in the overall p_T scale. Consequently, given the fact that subjets are specified during jet reconstruction, and the jet momentum calibration is derived after reconstruction, we do not attempt to correct the momenta of individual subjets. However, the subjet multiplicity is corrected for the experimental effects that cause an offset in jet p_T. In general, the presence of uranium noise, multiple interactions, and pile-up, tends to increases the subjet multiplicity.

IV. DATA SAMPLES

In leading-order QCD, the fraction of final-state jets originating from gluons decreases with increasing $x \propto p_T/\sqrt{s}$, the momentum fraction carried by the initial-state partons. This is due primarily to the x-dependence of the parton distributions. Because, for fixed p_T, the gluon fraction decreases when $\sqrt{s}$ is decreased from 1800 GeV to 630 GeV, this suggests an experimental way to define jet samples with different mixtures of quarks and gluons. A single set of criteria can be used to select jets at the two beam energies, without changing any of the detector elements. We use this principle to analyze an event sample recorded at the end of 1995 by the DØ detector at $\sqrt{s} = 630$ GeV, and compare it with the larger 1994–1995 event sample collected at $\sqrt{s} = 1800$ GeV. The lower range of jet p_T populated by the smaller event sample at $\sqrt{s} = 630$ GeV dictated the ultimate criteria used in the comparison. In Sec. IV A, we first describe a simple test of a set of criteria used to select quark-enriched and gluon-enriched jet samples. In Sec. IV B, we specify each criterion used in the analysis. In Sec. IV C, we provide a Monte Carlo estimate of the quark/gluon yield based on the full set of criteria. Finally, in Sec. IV D, we describe how to estimate the subjet content of gluon and quark jets.

A. Gluon and quark samples at leading-order in QCD

For a given set of parton distribution functions (PDFs), the relative admixture of gluon and quark jets passing a set of kinematic criteria can be estimated using a leading-order QCD event generator. At this order, there is no dependence on jet algorithm, because each of the two final-state partons defines a jet. We use the HERWIG v5.9 Monte Carlo with the CTEQ4M [23] PDFs to generate leading-order QCD $2 \rightarrow 2$ events, and keep track of the identity (gluon or quark) of the partons. At leading order, the gluon-jet fraction f corresponds to the number of final-state gluons that pass the selections divided by the total number of final-state partons that pass the selections. For example, the jet sample selected from only $gg \rightarrow gg$ or $q\bar{q} \rightarrow gg$ events will have a gluon-jet fraction of unity. Figure 12 shows that for the full ensemble of Monte Carlo events, the gluon-jet fraction at $\sqrt{s} = 630$ GeV is about 30% smaller than at $\sqrt{s} = 1800$ GeV, where we have selected central ($|\eta| < 0.5$) jets with minimum parton $p_T \approx 55$ GeV and maximum parton $p_T = 100$ GeV. This difference is due primarily to the relative abundance of initial-state gluons at these x values for $\sqrt{s} = 1800$ GeV compared to $\sqrt{s} = 630$ GeV.

B. Jet data samples

We define gluon-enriched and quark-enriched jet samples using identical criteria at $\sqrt{s} = 1800$ GeV and 630 GeV, thereby reducing any experimental biases and systematic effects. We select events that pass a trigger requiring the scalar sum of E_T above 30 GeV within a cone of size $\mathcal{R} = 0.7$ [25], and apply the selections listed in Sec. III D, but only for jets with measured p_T between 55 and 100 GeV. These cuts yield samples of 11,007 jets at $\sqrt{s} = 1800$ GeV, and 1194 jets at $\sqrt{s} = 630$ GeV.
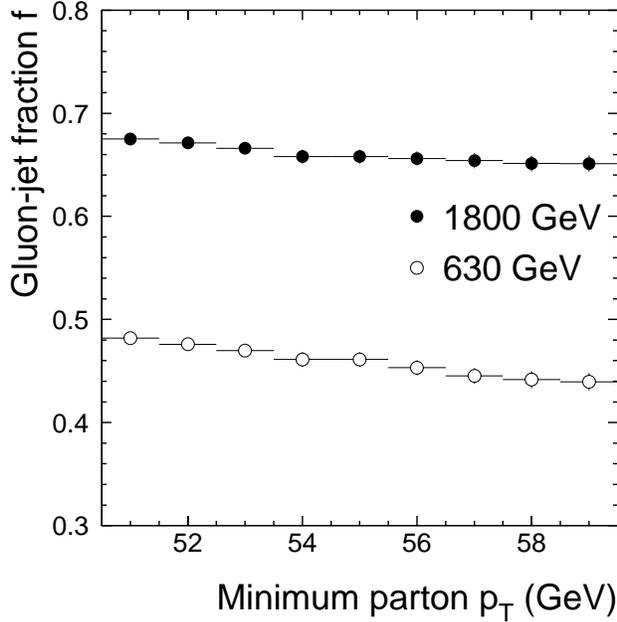
FIG. 12. The Monte Carlo gluon-jet fraction f at leading-order, for final-state partons with maximum parton $p_T = 100$ GeV, and minimum parton $p_T \approx 55$ GeV, as a function of the minimum parton p_T, using the CTEQ4M PDF. Both partons are required to be central ($|\eta| < 0.5$). The solid symbols show the prediction for $\sqrt{s} = 1800$ GeV, and the open symbols show the prediction for $\sqrt{s} = 630$ GeV.

An important point is that these jets were reconstructed with the $k_\perp$ algorithm for $D = 0.5$. This choice tends to select events with fewer subjets from initial-state radiation, which can vary with $\sqrt{s}$ (see Sec. III E). Figure 13 shows that the p_T distribution of the selected jets at $\sqrt{s} = 1800$ GeV is harder than at $\sqrt{s} = 630$ GeV. The mean jet p_T at $\sqrt{s} = 1800$ GeV is 66.3 ± 0.1 GeV, which is 2.3 GeV higher than at $\sqrt{s} = 630$ GeV. This cannot be caused by any differences in the contribution to the offset in the jet p_T. In fact, the entire offset is $p_O \approx 3 - 4$ GeV per jet at $\sqrt{s} = 1800$ GeV for $D = 1.0$ (see Sec. III C), and is therefore an expected factor ≈ 4 smaller for $D = 0.5$. Moreover, only a small fraction of the jet offset can be attributed to the difference in $\sqrt{s}$. Even so, offset differences can only change the subjet multiplicity by shifting the relative jet p_T. Rather than attempting to measure and account for such small effects in the jet p_T distributions, we simply use identical jet criteria at the two beam energies, and estimate the uncertainty on M by varying the jet selection cutoffs (see Sec. V C).

C. Jet samples in Monte Carlo events

To estimate the number of gluon jets in the $\sqrt{s} = 1800$ GeV and 630 GeV jet samples, we generated approximately 10,000 HERWIG events at each $\sqrt{s}$, with parton $p_T > 50$ GeV, and requiring at least one of the two leading-order partons to be central ($|\eta| < 0.9$). The events were passed through a full simulation of the DØ detector. To simulate the effects of uranium noise, pile-up from previous bunch crossings, and multiple $p\bar{p}$ interactions in
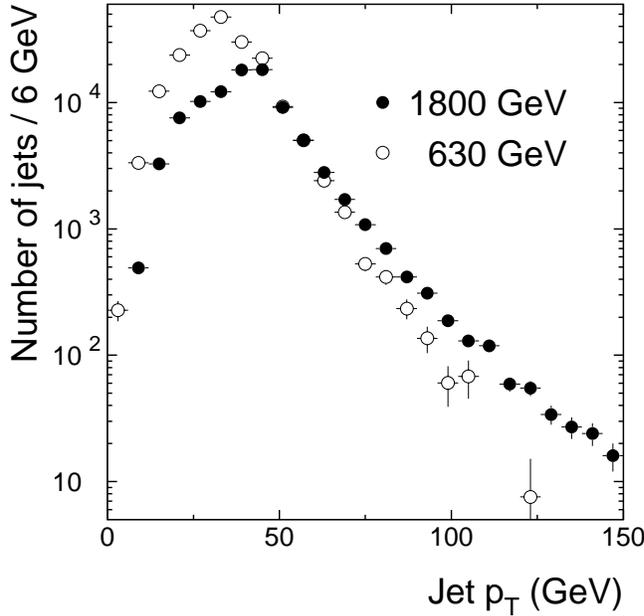
FIG. 13. The p_T distribution of selected jets in DØ data, before applying a cutoff on jet p_T. The data at $\sqrt{s} = 630$ GeV are normalized to the data at $\sqrt{s} = 1800$ GeV in the bin $54 \leq p_T < 60$ GeV. The turnover at lower jet p_T is due to inefficiencies in the trigger. For the following analysis, we use jets with $55 < p_T < 100$ GeV.

the same bunch crossing, we overlaid DØ random-crossing events onto our Monte Carlo sample, on a cell-by-cell basis in the calorimeter. (A sample with instantaneous luminosity of $\mathcal{L} \approx 5 \times 10^{30} \mathrm{cm}^{-2}\mathrm{s}^{-1}$ was used at $\sqrt{s} = 1800$ GeV, and $\mathcal{L} \approx 0.1 \times 10^{30} \mathrm{cm}^{-2}\mathrm{s}^{-1}$ was used at $\sqrt{s} = 630$ GeV.) These pseudo events were then passed through the normal offline-reconstruction and jet-finding packages. Jets were then selected using the same criteria as used for DØ data, and their p_T distribution is shown in Fig. 14.

We tag each such selected Monte Carlo jet as either quark or gluon based on the identity of the nearer (in $\eta \times \phi$ space) final-state parton in the QCD $2 \rightarrow 2$ hard scatter. The distance between one of the partons and the closest calorimeter jet is shown in Fig. 15. There is clear correlation between jets in the calorimeter and partons from the hard scatter. The fraction of gluon jets is shown in Fig. 16 as a function of the minimum p_T used to select the jets. There is good agreement for the gluon-jet fraction obtained using jets reconstructed at the calorimeter and at the particle levels ($\Delta f < 0.03$). The smaller gluon-jet fractions relative to leading-order (Fig. 12) are due mainly to the presence of higher-order radiation in the QCD Monte Carlo. When p_T cutoffs are applied to particle-level jets, the associated leading-order partons shift to significantly higher p_T. Since the gluon-jet fraction decreases with increasing parton p_T, f is smaller when events are selected according to particle-level jet p_T rather than when they are selected according to partonic p_T. The same is true for cutoffs applied to the calorimeter-level jets compared to the particle-level jets, although here the Δf discrepancy is much smaller. In what follows, we shall use nominal gluon-jet fractions $f_{1800} = 0.59$ and $f_{630} = 0.33$, obtained from Monte Carlo at the calorimeter level for $55 < p_T < 100$ GeV.
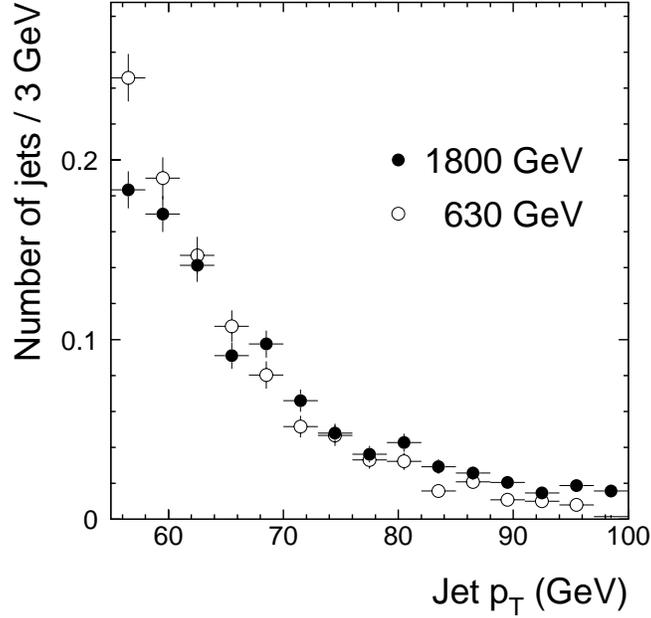
FIG. 14. The normalized p_T distribution of jets selected in Monte Carlo events at $\sqrt{s} = 1800$ GeV and 630 GeV. Each distribution has been normalized to unit area.
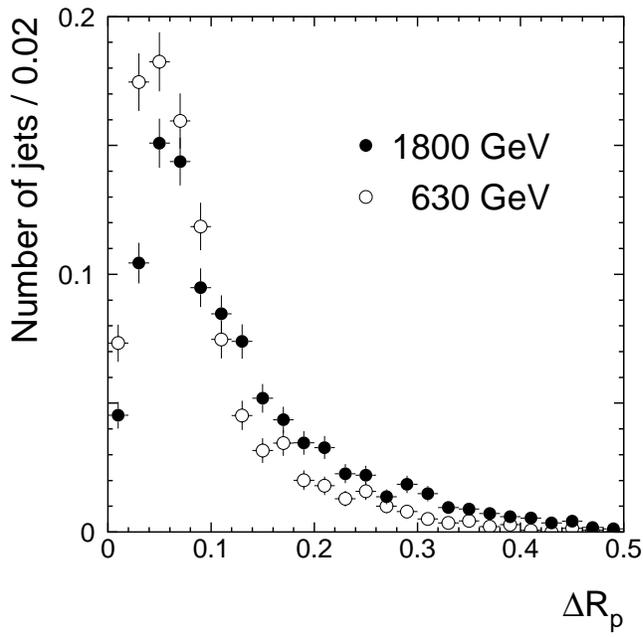

FIG. 15. The distance of the closest calorimeter-level Monte Carlo jet to one of the leading final-state partons. The solid (open) points show the Monte Carlo sample at $\sqrt{s} = 1800$ (630) GeV. Each distribution has been normalized to unit area.

D. Subjets in gluon and quark jets

Using the previously described jet samples, there is a simple way to distinguish between gluon and quark jets on a statistical basis. The subjet multiplicity in a mixed sample of
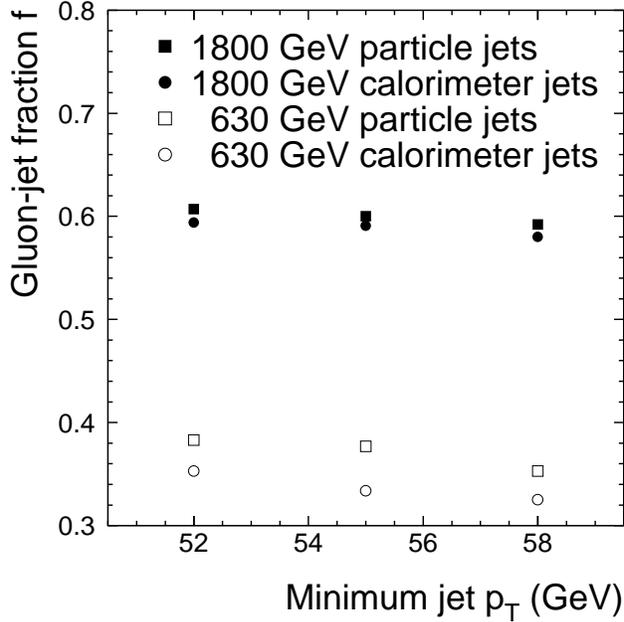
FIG. 16. The gluon-jet fraction of selected jets with maximum $p_T = 100$ GeV and minimum p_T between 52 and 58 GeV, as a function of minimum jet p_T, for $\sqrt{s} = 1800$ GeV and 630 GeV, using the CTEQ4M PDF. The jets have been tagged through the identity of the nearer leading-order final-state parton.

gluon and quark jets can be written as a linear combination of subjet multiplicity in gluon M_g and quark jets M_q:

$$M = fM_g + (1-f)M_q \tag{4.1}$$

The coefficients are the fractions of gluon and quark jets in the mixed sample, f and $(1-f)$, respectively. Considering Eq. (4.1) for two samples of jets at $\sqrt{s} = 1800$ GeV and 630 GeV, and assuming that M_g and M_q are independent of $\sqrt{s}$ (we address this assumption later), we can write:

$$M_g = \frac{(1-f_{630})M_{1800} - (1-f_{1800})M_{630}}{f_{1800} - f_{630}} \tag{4.2}$$

$$M_q = \frac{f_{1800}M_{630} - f_{630}M_{1800}}{f_{1800} - f_{630}} \tag{4.3}$$

where M_{1800} and M_{630} are the measured multiplicities in the mixed-jet samples at $\sqrt{s} = 1800$ GeV and 630 GeV, and f_{1800} and f_{630} are the gluon-jet fractions in the two samples. The extraction of M_g and M_q requires prior knowledge of the two gluon-jet fractions, as described in Sec. IV C. Since the gluon-jet fractions depend on jet p_T and η, Eqs. (4.2) and (4.3) hold only within restricted regions of phase space, i.e., over small ranges of jet p_T and η. Equations (4.2) and (4.3) can, of course, be generalized to any observable associated with a jet.
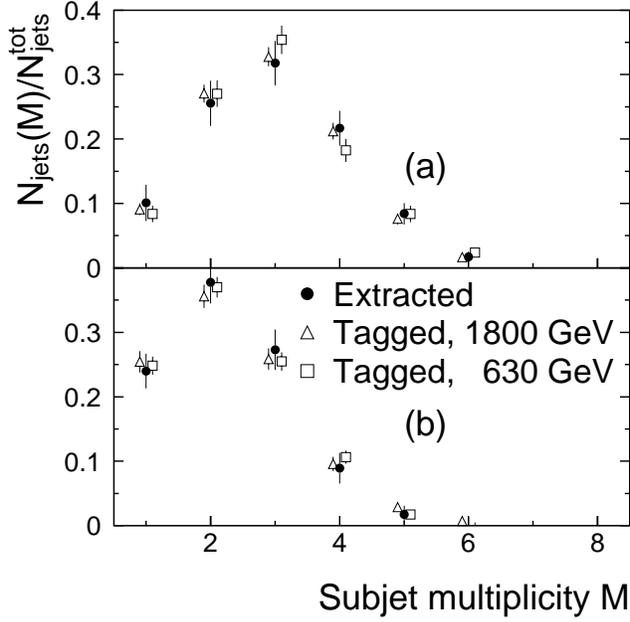
FIG. 17. Uncorrected subjet multiplicity in fully-simulated Monte Carlo (a) gluon and (b) quark jets. The number of jets $N_{\mathrm{jets}}(M)$ in each bin of subjet multiplicity on the vertical axis is normalized to the total number of jets in each sample $N_{\mathrm{jets}}^{\mathrm{tot}} = \sum_M N_{\mathrm{jets}}(M)$. The measured distributions (solid) are extracted from the mixed Monte Carlo jet samples at $\sqrt{s} = 1800$ GeV and 630 GeV. The tagged distributions (open) are for $\sqrt{s} = 1800$ GeV (triangles) and 630 GeV (squares).

We will use our Monte Carlo samples to check Eqs. (4.2) and (4.3) for $k_\perp$ jets reconstructed using the full-detector simulation with $D = 0.5$. Such a consistency test does not depend on the details of the subjet multiplicity distributions $(M_q, M_g, M_{1800}, M_{630})$. The extracted distributions in M_g and M_q are shown in Fig. 17. As expected, Monte Carlo gluon jets have more subjets, on average, than Monte Carlo quark jets: $\langle M_g \rangle > \langle M_q \rangle$. This is also found for jets reconstructed at the particle level, and the differences between gluon and quark jets do not appear to be affected by the detector. Also, the subjet multiplicity distributions for tagged jets are similar at the two center-of-mass energies, verifying the assumptions used in deriving Eqs. (4.2) and (4.3). Finally, the extracted M_q and M_g distributions agree very well with the tagged distributions. This demonstrates self-consistency of the extraction using Eqs. (4.2) and (4.3).

V. SUBJET MULTIPLICITIES

A. Uncorrected subjet multiplicity

Figure 18 shows the distributions of subjet multiplicity for the DØ data samples described in Sec. IV. This is the first measurement of its kind at a hadron collider. The average number of subjets in jets at $\sqrt{s} = 1800$ GeV is $\langle M_{1800} \rangle = 2.74 \pm 0.01$, where the error is statistical. This is higher than the value of $\langle M_{630} \rangle = 2.54 \pm 0.03$ at $\sqrt{s} = 630$ GeV. The observed shift is consistent with the prediction that there are more gluon jets in the sample at $\sqrt{s} = 1800$
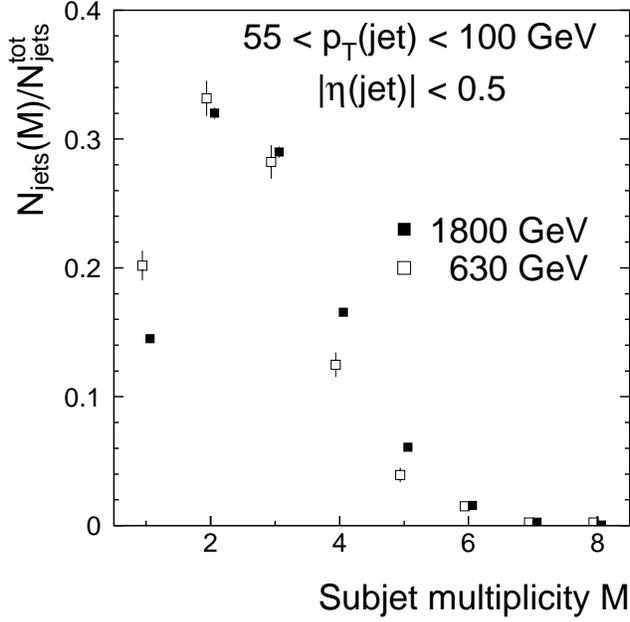
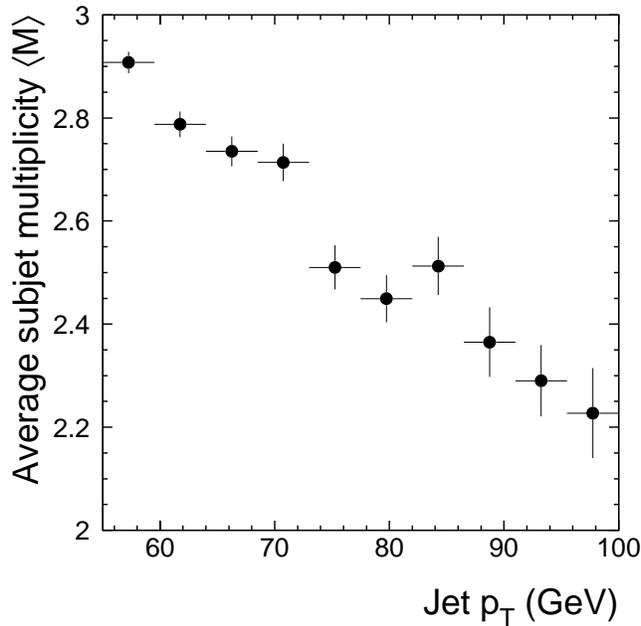FIG. 18. Uncorrected subjet multiplicity in jets from DØ data at $\sqrt{s} = 1800$ GeV and 630 GeV.

FIG. 19. Uncorrected mean subjet multiplicity versus jet p_T in DØ data at $\sqrt{s} = 1800$ GeV. Note the suppressed zero on the vertical axis.

GeV than in the sample at $\sqrt{s} = 630$ GeV, and that gluons radiate more subjets than quarks do. The fact that the p_T spectrum is harder at $\sqrt{s} = 1800$ GeV than at $\sqrt{s} = 630$ GeV cannot be the cause of this effect because the subjet multiplicity decreases with increasing jet p_T. Figure 19 shows the rather mild dependence of the average subjet multiplicity on jet p_T.
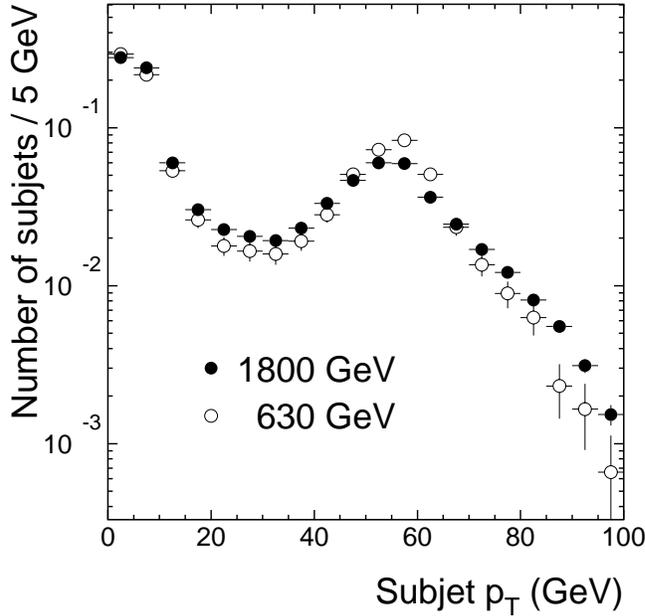
FIG. 20. The uncorrected p_T distribution of subjets in data for jets with $55 < p_T < 100$ GeV and $|\eta| < 0.5$. All selections have been applied, and each distribution has been normalized to unit area.

Subjets were defined through the product of their fractional jet p_T and their separation in (η, ϕ) space [see Eqs. (3.2) and (3.6)]. As shown in Figs. 20 and 21, the shapes of the subjet p_T spectra of the selected jets are similar at the two beam energies. The distributions suggest that jets are composed of a hard component and a soft component. The peak at about 55 GeV and fall-off at higher p_T is due to single-subjet jets and the jet p_T selections $(55 < p_T < 100$ GeV). The threshold at subjet $p_T \approx 1.75$ GeV is set by the value $y_{\text{cut}} = 10^{-3}$ and the minimum jet p_T in the sample.

While the M_{1800} and M_{630} inclusive measurements at $\sqrt{s} = 1800$ GeV and $\sqrt{s} = 630$ GeV are interesting in themselves, they can be interpreted in terms of their gluon and quark content. According to Eqs. (4.2) and (4.3) the distributions in Fig. 18 and their gluon-jet fractions at the two beam energies can yield the uncorrected subjet multiplicity distributions in gluon and quark jets. The extracted measurements of M_g and M_q are shown in Fig. 22 for the nominal values $f_{1800} = 0.59$ and $f_{630} = 0.33$. As in the Monte Carlo simulation, the DØ data clearly indicate the presence of more subjets in gluon jets than in quark jets. Such distributions can be used directly (without correcting the subjet multiplicities) to discriminate between gluon and quark jets. The results depend only on Monte Carlo estimates of gluon-jet fractions at the two values of $\sqrt{s}$, and not on any detailed simulation of jet structure.

The sensitivity of M_g and M_q to the assumed values of f_{1800} and f_{630} was checked by investigating how the signal (i.e., the difference between M_g and M_q) depended on this choice. It was found that when the gluon-jet fractions are either both increased or both decreased, the signal remains relatively unchanged. However, when the gluon-jet fractions are changed in opposite directions, this produces the largest change in the difference between gluon and
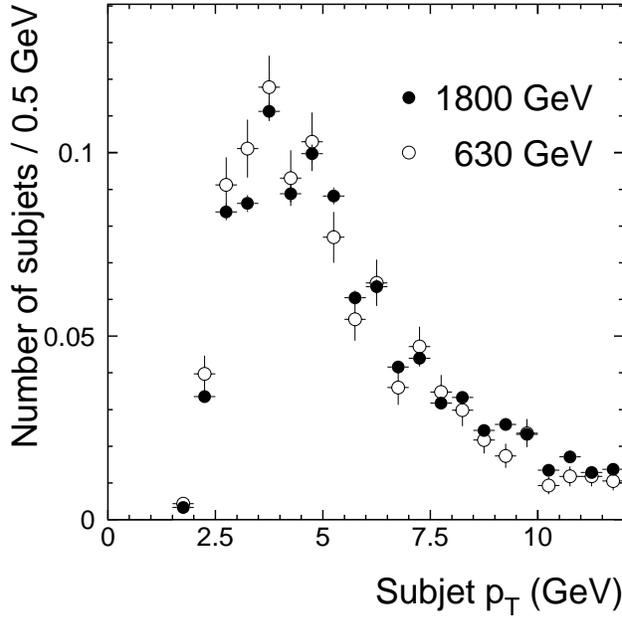
FIG. 21. Same as in Fig. 20, but with the low p_T region expanded. The increase at low p_T is observed for all y_{cut}, but the specific cutoff at $p_T(\mathrm{subjet}) \approx 1.75$ GeV is determined by our chosen value of $y_{\mathrm{cut}} = 10^{-3}$.
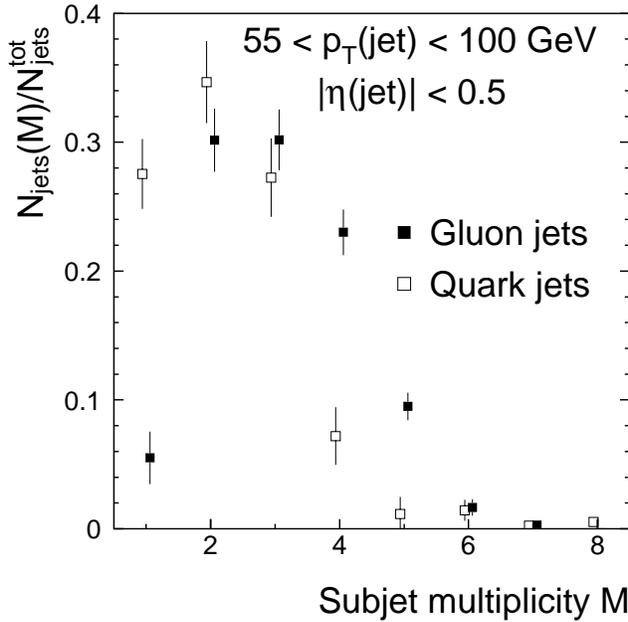

FIG. 22. Uncorrected subjet multiplicity in gluon and quark jets, extracted from DØ data at $\sqrt{s} = 1800$ GeV and 630 GeV, using nominal gluon-jet fractions $f_{1800} = 0.59$ and $f_{630} = 0.33$.

quark jets. The result of using $f_{1800} = 0.61$ and $f_{630} = 0.30$, instead of their nominal values, is shown in the extracted distributions of Fig. 23. The M_g and M_q distributions of Fig. 23 are qualitatively similar to those of Fig. 22, and the large difference between gluon and quark
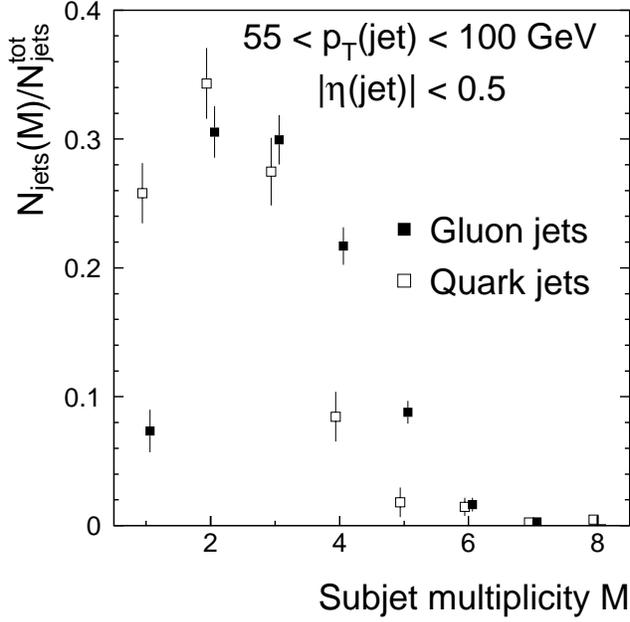
FIG. 23. Uncorrected subjet multiplicity in gluon and quark jets, extracted from DØ data at $\sqrt{s} = 1800$ GeV and 630 GeV, using gluon-jet fractions $f_{1800} = 0.61$ and $f_{630} = 0.30$.

jets is still apparent.

The subjet multiplicity distributions can be characterized by their means $\langle M \rangle$, and by $\langle M \rangle - 1$, which correspond to the average number of subjet *emissions* in a gluon or quark jet. For the nominal uncorrected DØ data shown in Fig. 22, $\langle M_g^{\text{meas}} \rangle = 3.05 \pm 0.06$ and $\langle M_q^{\text{meas}} \rangle = 2.28 \pm 0.08$. The analogous values for the Monte Carlo events (see Fig. 17) are $\langle M_g^{\text{meas}} \rangle = 3.01 \pm 0.09$ and $\langle M_q^{\text{meas}} \rangle = 2.28 \pm 0.08$. Because the quoted statistical uncertainty on $\langle M_g^{\text{meas}} \rangle$ is correlated with that on $\langle M_q^{\text{meas}} \rangle$, we define a ratio [13] of emissions in gluon jets to quark jets:

$$r \equiv \frac{\langle M_g \rangle - 1}{\langle M_q \rangle - 1}. \tag{5.1}$$

A value of $r = 1$ would mean that the substructure of gluon jets does not differ from that of quark jets. The ratio has a value of $r = 1.61 \pm 0.15$ for the uncorrected data of Fig. 22, and $r = 1.58 \pm 0.16$ for the analogous Monte Carlo events of Fig. 17, where both uncertainties are statistical. Using different values for gluon-jet fraction at the two values of $\sqrt{s}$ (as in Fig. 23), yields the range of r values given in Table I. As expected, the observed ratio is smallest when the fraction of gluon jets increases at $\sqrt{s} = 1800$ GeV and decreases at $\sqrt{s} = 630$ GeV. The two values of f are the only assumptions from Monte Carlo, and correspond to the largest source of systematic uncertainty on r (described more fully in Sec. V C). In all cases, we find that r is significantly greater than unity, meaning that gluon jets and quark jets differ in their substructure.

f_{1800}	f_{630}	$\langle M_g \rangle$	$\langle M_q \rangle$	r
0.59	0.33	3.05 ± 0.06	2.28 ± 0.08	1.61 ± 0.15
0.61	0.30	2.99 ± 0.05	2.34 ± 0.07	1.49 ± 0.11
0.61	0.36	3.05 ± 0.06	2.24 ± 0.09	1.65 ± 0.16
0.57	0.30	3.06 ± 0.06	2.31 ± 0.07	1.57 ± 0.14
0.57	0.36	3.15 ± 0.08	2.19 ± 0.10	1.81 ± 0.22

TABLE I. The uncorrected subjet multiplicity in gluon and quark jets, and their ratio, extracted from DØ data, assuming different values of gluon-jet fractions at the two center-of-mass energies.
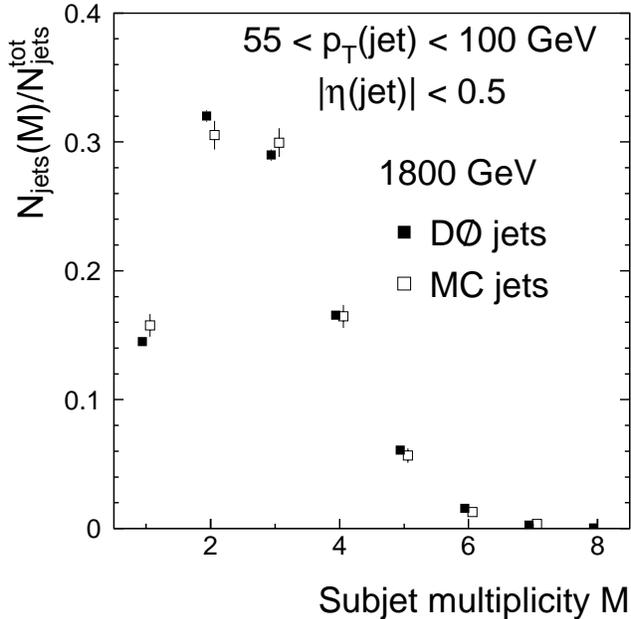

FIG. 24. Uncorrected subjet multiplicity in jets from DØ and fully-simulated Monte Carlo events at $\sqrt{s} = 1800$ GeV.

B. Corrected subjet multiplicity

As was stated above, the experimental conditions described in Sec. III C smear the measurement of the subjet multiplicity. Although r expresses differences between gluon and quark jets as a ratio of mean subjet multiplicities, the extracted M_g and M_q distributions need separate corrections for the various detector-dependent effects that can affect the value of r. The corrections are derived using Monte Carlo events, which are in agreement with the uncorrected DØ data, as shown in Figs. 24 and 25. The decomposition of the Monte Carlo events into M_g and M_q components was discussed in Sec. IV D. The distributions shown in Fig. 17 represent the uncorrected results for Monte Carlo events that we use to derive the unsmearing corrections.

The corrected distributions of M_g and M_q are defined in Monte Carlo jets at the particle level (i.e., before development in the calorimeter). All selected calorimeter-level jets are matched (within $\Delta \mathcal{R} < 0.5$) to jets reconstructed at the particle level. The matching
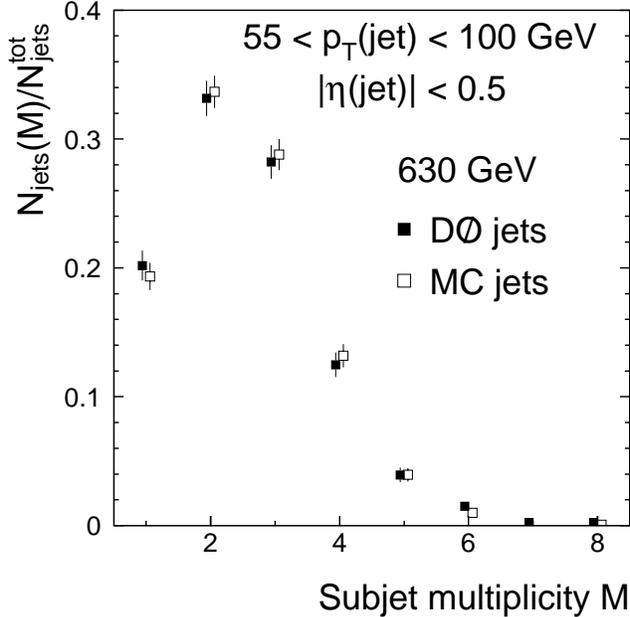
FIG. 25. Uncorrected subjet multiplicity in jets from DØ and from fully-simulated Monte Carlo events at $\sqrt{s} = 630$ GeV.

procedure implicitly accounts for any mismeasurement of jet p_T because there is no p_T requirement in the matching. The preclustering and clustering algorithms applied at the particle level are identical to those applied at the detector level. We tag simulated detector jets as either gluons or quarks, and correlate the subjet multiplicity in particle jets (M^{true}) with that of detector partners (M^{meas}). These correlations are shown in Fig. 26 at $\sqrt{s} = 1800$ GeV, and define the correction applied to the subjet multiplicity. Similar results are available at $\sqrt{s} = 630$ GeV (not shown).

The correction retrieves M^{true} from M^{meas}, in bins of M^{meas}. In general, the distributions of M_g^{true} and M_q^{true} in Fig. 26 are shifted to lower values relative to M_g^{meas} and M_q^{meas}. The shift in M is due mainly to the effects of showering in the calorimeter, rather than from the combined effects of multiple interactions, pile-up, and uranium noise, which are reduced by using $D = 0.5$. Fortunately, shower development is independent of beam energy, and the other contributions differ only slightly (see Sec. V C).

Shower development in the calorimeter tends to add subjets to a jet because any single particle can deposit energy in several towers of the calorimeter. Signals in many towers generate a large number of preclusters, and in turn, a large number of subjets. However, the opposite can also occur. For example, when two subjets at the particle level (each composed of one or two hadrons) deposit energy in a region of the calorimeter between them, such energy can "bridge" distinct subjets at the particle level into a single subjet at the calorimeter level. This bridging effect is more pronounced in jets that already have a large M^{true}. For this reason, the effects of multiple interactions, pile-up, and uranium noise tend to reduce the correction to M^{meas}.

To check that the correction defined by the correlations in Fig. 26 is valid, it was applied to the uncorrected M_g and M_q Monte Carlo distributions in Fig. 17. The resulting corrected
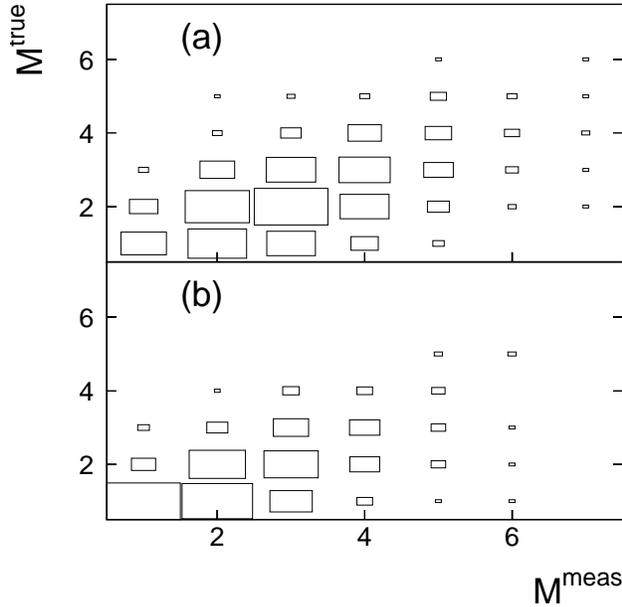
FIG. 26. The subjet multiplicity at particle-level (M^{true}) versus the subjet multiplicity at calorimeter-level (M^{meas}) (includes effects of luminosity), at $\sqrt{s}$ = 1800 GeV, for (a) gluon and (b) quark jets.

distributions for M_g and M_q are given in Fig. 27 and Fig. 28, respectively. The correction reduces the average subjet multiplicity in the Monte Carlo to $\langle M_g^{\mathrm{true}}\rangle = 2.19 \pm 0.04$ and $\langle M_q^{\mathrm{true}}\rangle = 1.66 \pm 0.04$. and the corrected ratio is $r = 1.82 \pm 0.16$. Any remaining small differences between the extracted and the tagged M^{true} distributions in Fig. 27 and Fig. 28 are attributable to the differences between the extracted and the tagged M^{meas} (at $\sqrt{s} = 1800$ GeV) of Fig. 17. These differences are smaller for the corrected distributions M^{true}, than for the uncorrected distributions.

Figure 29 shows the corrected subjet multiplicities for gluon and quark jets. The rate for $M = 1$ quark jets has almost doubled, while the rate for $M = 3$ quark jets has fallen by a factor of ≈ 2, relative to the uncorrected result. A similar effect is observed for gluon jets. From Fig. 29, we obtain the corrected mean values in the DØ data to be $\langle M_g^{\mathrm{true}}\rangle = 2.21 \pm 0.03$ and $\langle M_q^{\mathrm{true}}\rangle = 1.69 \pm 0.04$, which gives $r = 1.75 \pm 0.15$, in good agreement with the prediction from HERWIG. The unsmearing therefore widens the difference between gluon and quark jets.

We choose not to correct M for any impact of the preclustering algorithm on subjet multiplicity. Instead, the preclustering algorithm can be applied easily to the particle-level events in Monte Carlo, and these are therefore treated in the same way as the DØ data. For completeness, we note that r can decrease by as much as 0.2 at the particle level, when preclustering is turned off.
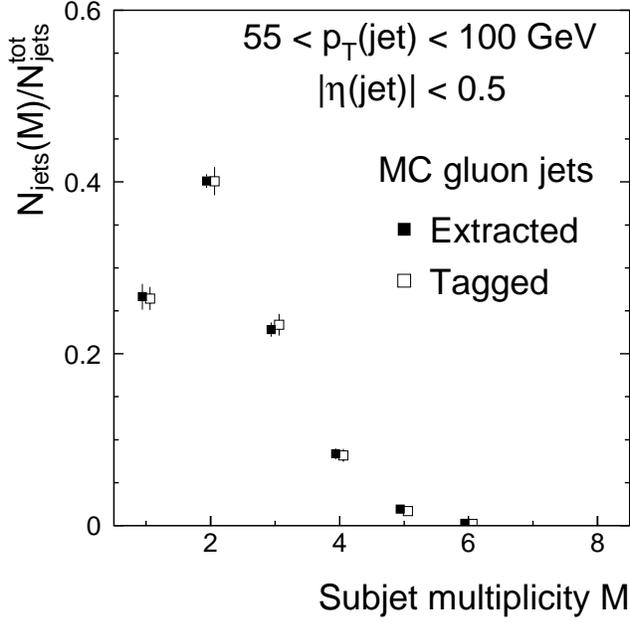
FIG. 27. The subjet multiplicity in Monte Carlo gluon jets. The extracted distribution has been unsmeared. The tagged distribution was obtained directly from particle-level gluon jets at $\sqrt{s} = 1800$ GeV.
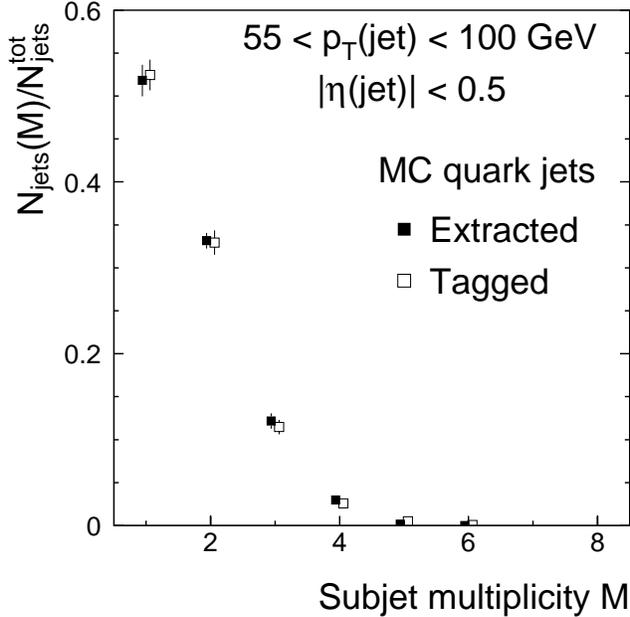

FIG. 28. The subjet multiplicity in Monte Carlo quark jets. The extracted distribution has been unsmeared. The tagged distribution was obtained directly from particle-level quark jets at $\sqrt{s} = 1800$ GeV.

C. Additional corrections and systematic uncertainties

The dominant systematic uncertainty on the subjet multiplicity arises from the uncertainty on the gluon-jet fractions. In fixed-order perturbative QCD, the jet cross section at
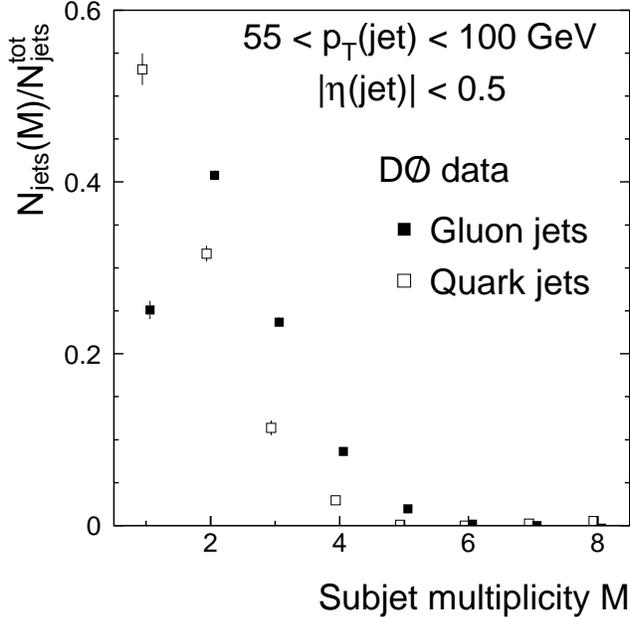
FIG. 29. Corrected subjet multiplicity for gluon and quark jets, extracted from DØ data.

any given p_T is a more-steeply-falling function of p_T at $\sqrt{s} = 630$ GeV than at $\sqrt{s} = 1800$ GeV [25]. Consequently, applying identical cutoffs biases the $\langle p_T \rangle$ of jets at $\sqrt{s} = 1800$ GeV upwards relative to $\sqrt{s} = 630$ GeV. Monte Carlo studies indicate this bias is approximately 2 GeV. One way to compensate for this effect is to shift the p_T range at $\sqrt{s} = 630$ upwards by a few GeV. Due to the steep negative slope of the jet-p_T spectrum, it is sufficient to shift only the lower edge of the p_T bins. When this is done, Fig. 12 shows that the change in gluon-jet fraction is $\Delta f < 0.03$. We do not correct f for this, but account for this residual effect in the systematic uncertainty associated with the jet p_T.

Changing the gluon-jet fractions used in the analysis gives a direct estimate of the uncertainty on the subjet multiplicity. We will motivate the range of uncertainty in gluon-jet fractions at the two center-of-mass energies by investigating the behavior of the PDFs. For the jet samples used in this analysis, the average jet p_T was approximately 65 GeV. This jet p_T probes an average x value of 0.07 at $\sqrt{s} = 1800$ GeV and 0.2 at $\sqrt{s} = 630$ GeV. In these regions of x, the quark PDFs are well-constrained by existing data. However, the gluon PDF is not so well-constrained. We examined different parameterizations of the gluon PDF at the two x values of interest. In particular, the MRST5 [24] gluon PDF is 21% smaller than the CTEQ4M parameterization at $x = 0.2$, but only 4% smaller at $x = 0.07$. This and other comparisons between PDFs show larger fractional differences at $x = 0.2$ than at $x = 0.07$.

Assuming that the quark distributions are essentially identical in different PDF parameterizations, the gluon-jet fraction f' for different PDFs can be estimated as

$$f' = \frac{f + \epsilon f}{(f + \epsilon f) + (1 - f)} \tag{5.2}$$

where f is the gluon-jet fraction from some reference PDF, and ϵ is a fractional difference in the gluon PDF. Table II shows the gluon-jet fractions estimated for PDFs at the two

PDF set	$\sqrt{s}$ (GeV)	x	$xg(x)$	ϵ	$f_{\sqrt{s}}$
CTEQ4M	1800	0.07	1.643	0.00	0.59
CTEQ4HJ	1800	0.07	1.643	0.00	0.59
CTEQ2M	1800	0.07	1.714	0.04	0.60
CTEQ5M	1800	0.07	1.614	-0.02	0.59
CTEQ5HJ	1800	0.07	1.586	-0.04	0.58
MRST5	1800	0.07	1.586	-0.04	0.58
GRV94	1800	0.07	1.743	0.06	0.60
CTEQ4M	630	0.2	0.365	0.00	0.33
CTEQ4HJ	630	0.2	0.340	-0.06	0.32
CTEQ2M	630	0.2	0.385	0.06	0.34
CTEQ5M	630	0.2	0.340	-0.06	0.32
CTEQ5HJ	630	0.2	0.350	-0.03	0.32
MRST5	630	0.2	0.290	-0.21	0.28
GRV94	630	0.2	0.405	0.12	0.36

TABLE II. Values of gluon-jet fractions for different PDFs, calculated using Eq. (5.2), at a jet $p_T = 65$ GeV. The CTEQ4M parameterization is chosen as the reference. The fractional change in the gluon PDF $g(x)$ is given by $\epsilon = (g(x) - g_o(x))/g_o(x)$, where $g_o(x)$ is the reference.

center-of-mass energies. The MRST5 set shows the largest departure relative to CTEQ4M. In all cases, the change in f is in the same direction at both $\sqrt{s}$.

The preceeding discussion assumed that the PDFs had the same quark distribution. In reality, the quark PDFs also tend to change when the gluon PDF changes. When this compensating effect is taken into account in Eq. (5.2), the equivalent MRST5 gluon-jet fractions become $f_{1800} = 0.58$ and $f_{630} = 0.29$.

Based on the above, we assign uncertainties to the gluon-jet fractions of ± 0.02 at $\sqrt{s} = 1800$, and ± 0.03 at $\sqrt{s} = 630$. In fact, we vary the gluon-jet fraction in opposite directions, using $f_{1800} = 0.61$ and $f_{630} = 0.30$, and $f_{1800} = 0.57$ and $f_{630} = 0.36$, to gauge the impact on r. As in Sec. V A, we repeat the analysis assuming these different input gluon-jet fractions, this time including the correction to the particle level. The extracted ratios are summarized in Table III. The largest departures from the reference value of $r = 1.75$ define the systematic uncertainties of $\pm^{0.17}_{0.10}$.

The second-largest source of systematic uncertainty in the subjet multiplicity stems from an uncertainty in the measurement of jet p_T. A mismeasurement of jet p_T will lead to the selection of a slightly different sample of jets, but will not affect the subjet multiplicity directly. If jet p_T is mismeasured at both center-of-mass energies, we expect the effect to partially cancel in the ratio r. An estimate of the impact from this uncertainty is therefore obtained by varying the jet p_T only at $\sqrt{s} = 1800$ GeV. Since the calorimeter response is independent of $\sqrt{s}$, we estimate the effect of a difference in any offset in p_T at the two center-of-mass energies by changing the jet-p_T window from $55 < p_T < 100$ GeV to $57 < p_T < 100$ GeV at $\sqrt{s} = 1800$ GeV. A 2 GeV shift in the measured jet p_T corresponds approximately to two times the total offset p_O for $k_\perp$ jets reconstructed with $D = 0.5$. [This assumes $p_O(D)$ scales as $D^2 p_O(D = 1.0)$]. This reduces the subjet multiplicity ratio r by 0.12, which is

f_{1800}	f_{630}	$\langle M_g \rangle$	$\langle M_q \rangle$	r
0.59	0.33	2.21 ± 0.03	1.69 ± 0.04	1.75 ± 0.15
0.61	0.30	2.18 ± 0.02	1.72 ± 0.04	1.65 ± 0.12
0.61	0.36	2.20 ± 0.03	1.67 ± 0.05	1.79 ± 0.17
0.57	0.30	2.21 ± 0.03	1.70 ± 0.04	1.72 ± 0.14
0.57	0.36	2.24 ± 0.04	1.65 ± 0.05	1.92 ± 0.22

TABLE III. Subjet multiplicity in gluon and quark jets, and their ratio, extracted from DØ data and corrected to the particle level, assuming different gluon-jet fractions at the two center-of-mass energies.

Source	δr
Gluon-jet fraction	$^{+0.17}_{-0.10}$
Cutoff on jet p_T	± 0.12
Unsmearing	± 0.07
Detector simulation	± 0.02
Total	$^{+0.22}_{-0.18}$

TABLE IV. Systematic uncertainties on the ratio r.

taken as a symmetric systematic uncertainty.

Because the correction to the particle level produces a large change in the shape of the subjet multiplicity distribution, we will estimate the impact of the unsmearing on the systematic uncertainty on r. This uncertainty has two parts: one is the uncertainty due to the simulation of effects arising from dependence on luminosity, and the other is the uncertainty in the simulation of the DØ calorimeter. To account for the former, we use an alternate Monte Carlo sample at $\sqrt{s} = 630$ GeV, with a luminosity of $\mathcal{L} \approx 0.1 \times 10^{30} \mathrm{cm}^{-2} \mathrm{s}^{-1}$, and note that r increases by 0.13. Such a small change in r indicates that it depends only weakly on luminosity. Nevertheless, we increase our nominal value of $r = 1.75$ by half of the difference (to $r = 1.82$), and take this correction as a symmetric systematic uncertainty of ± 0.07.

To evaluate the other part of the uncertainty on the unsmearing, we compare two types of simulations of the DØ calorimeter. The default fast simulation (SHOWERLIB) is a library that contains single-particle calorimeter showers obtained using the GEANT full detector simulation. SHOWERLIB truncates the number of calorimeter cells associated with each individual particle, but rescales the energy of the shower to agree with the average energy given by the full GEANT simulation. The full GEANT simulation, while slower, accounts for the precise geometry of the uranium plates in the calorimeter and has no truncation. In a test using a limited number of Monte Carlo events, the latter simulation produced more subjets than the former, and so we increase the value of the ratio by 0.02 (half the difference of the r values in each simulation) to $r = 1.84$, and take this correction as another symmetric systematic error of ± 0.02. Applying the same additional corrections to the nominal ratio in the Monte Carlo gives a final result of $r = 1.91$ for HERWIG.

A list of the systematic uncertainties is shown in Table IV, all of which are added in quadrature to obtain the total uncertainty of the corrected ratio. The final result for the

ratio is

$$r \equiv \frac{\langle M_g \rangle - 1}{\langle M_q \rangle - 1} = 1.84 \pm 0.15 \text{ (stat.)} \pm^{0.22}_{0.18} \text{ (sys.)}. \qquad (5.3)$$

VI. CONCLUSION

We present two analyses of DØ data using the $k_\perp$ jet reconstruction algorithm. One analysis examines the p_T and direction of $k_\perp$ jets reconstructed with the parameter $D = 1.0$. For this measurement of the jet p_T spectrum, we describe a procedure to calibrate the momentum of $k_\perp$ jets based on our experience with the cone algorithm, but using an improved technique for determining the offset correction. Compared to our published results for the cone algorithm with $\mathcal{R} = 0.7$ [33], the $k_\perp$ jet algorithm with $D = 1.0$ reconstructs $40 - 50\%$ more energy from uranium noise, pile-up, multiple $p\bar{p}$ interactions, and the underlying event, and has a smaller uncertainty on the offset. We also report the results of a direct comparison of the $k_\perp$ and cone algorithms, on an event-by-event basis. Considering only the two leading jets, the $k_\perp$ and cone jet axes coincide within $\Delta\mathcal{R} = 0.1$ (0.5) at the 91% (99.94%) level. Matching with $\Delta\mathcal{R} = 0.5$, the corrected p_T of $k_\perp$ jets is higher than the corrected E_T of cone jets. The difference is roughly linear in jet p_T, varying from about 5 GeV at $p_T \approx 90$ GeV to about 8 GeV at $p_T \approx 240$ GeV.

In the other analysis, we probe the structure of $k_\perp$ jets reconstructed with the parameter $D = 0.5$, and find that the HERWIG Monte Carlo predictions of subjet multiplicity are in excellent agreement with our measurements. The subjet multiplicities in gluon and quark jets, predicted by a fully resummed calculation [16], and shown in Fig. 30, are qualitatively consistent with our data, but their mean values are slightly high. This discrepancy may be due to the fact that the calculation lacks a preclustering algorithm. The subjet multiplicity distributions, where we have subtracted the DØ values from the predictions, are shown in Fig. 31. The ratio of mean multiplicities for the resummed calculation (which assumes $M \leq 5$) is $r = 2.12$. The ratio in the DØ data increases by 0.06 with the assumption $M \leq 5$. Therefore, the resummed prediction is well within the limits of experimental uncertainty. The ratio measured at DØ agrees with the result of $r = 1.7 \pm 0.1$ from ALEPH, measured in e^+e^- annihilations at $\sqrt{s} = M_Z$ for a subjet resolution parameter $y_o = 10^{-3}$ [13], and with the associated Monte Carlo and resummation prediction [14]. All of the experimental and theoretical values for r are smaller than the naive QCD prediction of the ratio of color charges of 2.25. This is because of higher-order radiation in QCD, which tends to reduce the ratio from the naive value.

In summary, we present the first detailed measurements of properties of $k_\perp$ jets in hadron collisions. Using the standard value $D = 1.0$ of the jet-separation parameter in the $k_\perp$ algorithm, we find that the p_T of $k_\perp$ jets is higher than the E_T of matched cone jets (with $\mathcal{R} = 0.7$) by about 5 (8) GeV at $p_T \approx 90$ (240) GeV. To analyze internal jet structure, we measure the multiplicity distribution of subjets in $k_\perp$ jets with $D = 0.5$ at $\sqrt{s} = 1800$ GeV and 630 GeV. Exploiting the difference in gluon-jet fractions at the two center-of-mass energies, we extract the subjet multiplicity in gluon and quark jets. The differences between gluon and quark jets are encapsulated in the ratio of average emitted subjet multiplicities,
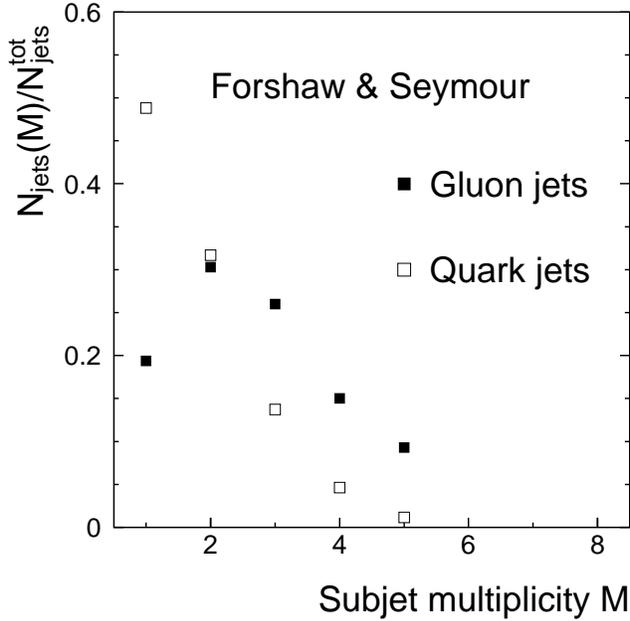
FIG. 30. The subjet multiplicity in gluon and quark jets, for $y_{cut} = 10^{-3}$ (as defined by Eq. [3.6]), in a resummation calculation by Forshaw and Seymour [16]. The jets are produced at $\sqrt{s} = 1800$ GeV, with $p_T = 65$ GeV and $\eta = 0$, using the CTEQ4M PDF, and are reconstructed with $D = 0.5$. The points in the fifth bin refer to $M \geq 5$.
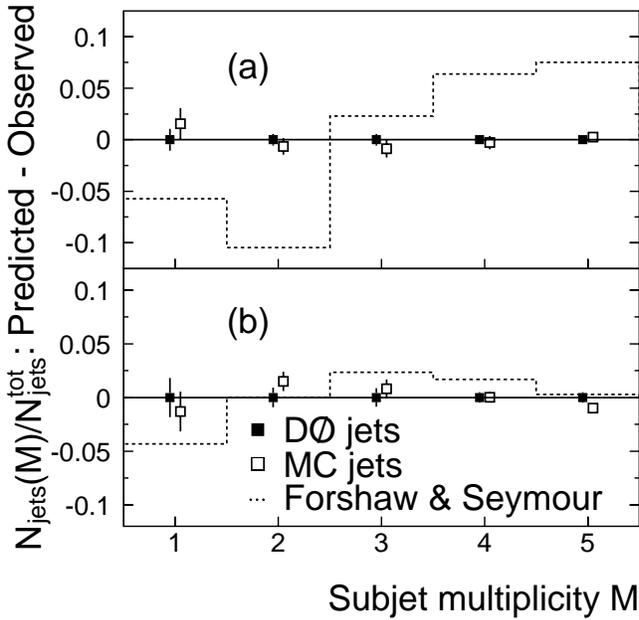

FIG. 31. The subjet multiplicity in (a) gluon and (b) quark jets, for DØ data, for the HERWIG Monte Carlo, and resummed predictions. The resummed prediction does not use a preclustering algorithm. The points in the fifth bin are for $M \geq 5$. The DØ data (see Fig. 29) have been subtracted from each set of points.

measured as:

$$r \equiv \frac{\langle M_g \rangle - 1}{\langle M_q \rangle - 1} = 1.84 \pm^{0.27}_{0.23} . \tag{6.1}$$

The DØ result demonstrates that gluon and quark jets are significantly different in hadron collisions, and that it may be possible to discriminate between them on an individual basis.

VII. ACKNOWLEDGMENTS

REFERENCES

* Visitor from University of Zurich, Zurich, Switzerland.

[1] R.K. Ellis, W.J. Stirling, and B.R. Webber, *QCD and Collider Physics* (Cambridge University Press, Cambridge, 1996), p. 219.

[2] M.Z. Akrawy *et al.* (OPAL Collaboration), Phys. Lett. B **261**, 334 (1991); G. Alexander *et al.* (OPAL Collaboration), Phys. Lett. B **265**, 462 (1991).

[3] I.G. Knowles *et al.*, in *Physics at LEP 2, Vol. 2*, edited by G. Altarelli and F. Zwirner, CERN 96-01 (Geneva, CERN, 1996), p. 114; J.W. Gary, in *Proceedings of the XXV International Symposium on Multiparticle Dynamics* (Stara Lesna, Slovakia, 1995), p. 518.

[4] P. Bagnaia *et al.* (UA2 Collaboration), Phys. Lett. B **144**, 291 (1984).

[5] G. Arnison *et al.* (UA1 Collaboration), Nucl. Phys. **B276**, 253 (1986).

[6] B. Abbott *et al.* (DØ Collaboration), Phys. Rev. D **60**, 012001 (1999); Phys. Rev. Lett. **83** 1908 (1999).

[7] R. Snihur (for the DØ Collaboration), in *Proceedings of the DIS1999 Workshop*, edited by J. Blumlein and T. Riemann, Nucl. Phys. (Proc. Suppl.) **B79**, 494 (1999); Ph.D. thesis, Northwestern University, 2000 (unpublished), `http://www-d0.fnal.gov/-results/publicastions_talks/thesis/snihur/thesis.ps`.

[8] A.N. Safonov (for the CDF Collaboration), in *Proceedings of the DIS2000 Workshop*, edited by J. Gracey and T. Greenshaw, to be published in Nucl. Phys. (Proc. Suppl.) **B**.

[9] S. Catani, Yu.L. Dokshitzer, F. Fiorani, and B.R. Webber, Nucl. Phys. **B383**, 419 (1992).

[10] R. Akers *et al.* (OPAL Collaboration), Z. Phys. C **63**, 363 (1994).

[11] M.H. Seymour, Nucl. Phys. **B421**, 545 (1994).

[12] R.V. Astur (for the DØ Collaboration), in *Proceedings of the 10th Topical Workshop on Proton-Antiproton Collider Physics*, edited by R. Raja and J. Yoh (AIP Press, Woodbury, N.Y., 1995) p. 598.

[13] D. Buskulic *et al.* (ALEPH Collaboration), Phys. Lett. B **346**, 389 (1995).

[14] M.H. Seymour, Phys. Lett. B **378**, 279 (1996).

[15] S. Behari *et al.* (AMY Collaboration), Phys. Lett. B **374**, 304 (1996).

[16] J.R. Forshaw and M.H. Seymour, JHEP **9909**, 009 (1999).

[17] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour, and L. Stanco, Comp. Phys. Comm. **67**, 465 (1992).

[18] S. Abachi *et al.* (DØ Collaboration), Nucl. Instrum. Methods in Phys. Res. A **338**, 185 (1994).

[19] S. Catani, Yu.L. Dokshitzer, M.H. Seymour, and B.R. Webber, Nucl. Phys. **B406**, 187 (1993).

[20] S. Catani, Yu.L. Dokshitzer, and B.R. Webber, Phys. Lett. B **285**, 291 (1992).

[21] S.D. Ellis and D.E. Soper, Phys. Rev. D **48**, 3160 (1993).

[22] K.C. Frame (for the DØ Collaboration), in *Proceedings of the Annual Meeting of the American Physical Society, Division of Particles and Fields (DPF 94)*, edited by S. Seidel (World Scientific, River Edge, N.J., 1995), p. 1650.

[23] H.L. Lai, J. Huston, S. Kuhlmann, J.F. Owens, D.E. Soper, W.K. Tung and H. Weerts, (CTEQ Collaboration) Phys. Rev. D **55**, 1280 (1997).

[24] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne, Eur. Phys. J. C **4**, 463 (1998).

[25] B. Abbott *et al.* (DØ Collaboration), Phys. Rev. D **64**, 32003 (2001).

[26] V.M. Abazov *et al.* (DØ Collaboration), to be submitted, Phys. Rev. Lett. (2001).

[27] J.E. Huth *et al.*, in *Proceedings of Research Directions For The Decade: Snowmass '90*, edited by E.L. Berger (World Scientific, River Edge, N.J., 1992), p. 134.

[28] W. Giele *et al.* (Jet Physics Working Group), in *QCD and Weak Boson Physics in Run II*, edited by U. Baur, R.K. Ellis, D. Zeppenfeld (Fermilab, Batavia, IL, 2000).

[29] B. Abbott *et al.* (DØ Collaboration), FERMILAB-PUB-97/242-E (1997) (unpublished).

[30] E.W.N. Glover and D.A. Kosower, Phys. Lett. B **367**, 369 (1996).

[31] K.C. Frame (for the DØ Collaboration), in *Proceedings of the VIII International Conference on Calorimetry in High Energy Physics*, edited by G. Barreira, B. Tome, A. Gomes, A. Maio, and M. J. Varanda (World Scientific, Singapore, 2000) p. 439; Ph.D. thesis, Michigan State University, 1999 (unpublished), `http://www-d0.fnal.gov/results/-publications_talks/thesis/frame/thesis/kcf_fnal_thesis.ps`.

[32] M.H. Seymour, in *Proceedings of the 10th Topical Workshop On Proton-Antiproton Collider Physics*, edited by Rajendran Raja and John Yoh (American Inst. Phys., Woodbury, N.Y., 1996) p. 568.

[33] B. Abbott *et al.* (DØ Collaboration), Nucl. Instrum. Meth. A **424**, 352 (1999).

[34] F. Carminati *et al.*, GEANT Users Guide, CERN Program Library W5013, 1991 (unpublished).

[35] B. Abbott *et al.* (DØ Collaboration), Phys. Rev. Lett. **82**, 2451 (1999).

[36] Yu.L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977); V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972); L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 95 (1975); G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977).