PHYSICAL REVIEW D VOLUME 44, NUMBER 7 1 OCTOBER 1991

How to tell quark jets from gluon jets

Jon Pumplin
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824
(Received 22 May 1991)

A method to distinguish between light-quark jets and gluon jets in a calorimeter detector is presented. The method could be used to refine the experimental test of QCD based on the single-jet cross section, by comparing the quark and gluon contributions separately. It could also be used to enhance the signal/background ratio in searching for the top quark in $W^{\pm}$ + jets channels. Optimistically, it might even make it possible to observe the hadronic decays $W^{\pm} \to j\,j$ and $Z^0 \to j\,j$, which would allow the W and Z masses to be used as benchmarks for calibrating the calorimeter, and perhaps even provide a useful measurement of M_W/M_Z. The method has been developed using Monte Carlo simulations of jet events. Ways to compare it with experiment are discussed. The comparison would provide an important test for QCD shower simulation models.

I. INTRODUCTION

An outstanding feature of hadron-hadron collisions at large momentum transfer is the production of jets. The jets arise mainly from light-quark (u, d, s, $\bar{u}$, $\bar{d}$, $\bar{s}$) and gluon constituents of the initial hadrons, which undergo perturbative hard scattering, and hadronize into physical particles according to nonperturbative QCD. Similar jets are seen in e^+e^- collisions, from annihilation via γ/Z^0.

It would be useful for several reasons to distinguish quark jets from gluon jets. One reason is that to test QCD at the largest possible Q^2 — i.e., on the shortest possible distance scale — the inclusive single-jet cross section [1] is used, since that cross section is relatively large. The cross section is a sum of q and G contributions, and it would make the test of QCD more definitive to compare the q and G parts separately. This would also test assumptions about the parton distributions, especially the less-well-measured gluon distribution, in the original hadrons.

A second application for quark-gluon separation is to reduce the background from gluon jets in searching for the top quark in $W^{\pm}$ + jets channels [16]. Further motivation for attempting the separation is the fact that the hadronization process can be approximated at present only by semiphenomenological simulations such as HERWIG [2], ISAJET [3], or PYTHIA [4]. Testing the separation will intrinsically test the hadronization model used in developing it.

A final motivation for jet separation is that if quark jets can be recognized, it may be possible to observe the hadronic decays $W^{\pm} \to j\,j$ and $Z^0 \to j\,j$ in spite of the large QCD backgrounds. This would provide an excellent absolute calibration of the measurement of jet energies, since the mass of the Z^0 is accurately known. More optimistically, one might hope to measure the key mass ratio M_W/M_Z using jet decays, and thus avoid many sources of systematic error in that measurement.

We assume here that the hadrons of the jet are detected using a segmented calorimeter of the type used by the Collider Detector at Fermilab (CDF) [1] and under construction by D0 [5] at the Fermilab Tevatron. Similar detectors are expected to be built at the Superconducting Super Collider (SSC). These detectors measure the total energy deposited in segments of pseudorapidity $\eta = -\ln\tan\frac{\theta}{2}$ and azimuthal angle ϕ. The jet information is thus degraded by the fact that individual particles are not tracked. This degradation is not very significant for our purposes, since it is similar to the uncertainties inherent in the QCD jet development. It is in any case a small price to pay for the fact that such detectors can provide reasonably equal response for hadrons as for electromagnetically interacting particles, including the photons from π^0 decay.

The physical basis for distinguishing between quark jets and gluon jets is the fact that the branching of gluons is stronger than the branching of quarks according to QCD. This can be seen directly by comparing the lowest-order elementary branching probability [6] for gluons, $G \to G + G$ and $G \to q + \bar{q}$, with that for quarks, $q \to q + G$. Gluon jets are therefore expected on average to be broader in (η, ϕ), and to contain more particles, than quark jets of similar E_T.

Several previous simulation studies have shown that q/G jet separation is feasible. Jones [7] describes a method in which each particle in a jet is characterized by its fractional contribution z_i to the jet momentum, and the jet is characterized by its pattern $\{z_i\}$ for all particles that have $z_i > 0.1$. Via Monte Carlo simulation, certain patterns are found to be favored by quark jets, while others are favored by gluons. The reliance of this method on the momenta of individual particles is not "infrared safe" [8]: the splitting of a parton into two colinear partons, or the decay of a hadron before it reaches the detector, can strongly affect the pattern $\{z_i\}$. It is therefore perhaps not surprising that the new method achieves a higher degree of separation, even though it uses only calorimeter information, rather than requiring individual particle momenta to be measured.

Lönnblad *et al.* [9, 10] have shown that quark-gluon

separation can be accomplished by "training" a neural-net algorithm to recognize patterns of energy deposition in the 100 – 200 calorimeter cells within the jet cone. The method to be described here is based on the same information, but it leads to better discrimination than the neural-net method. It also possesses the attractive quality of having a clear physical interpretation, in place of the "black box" character of the neural net. Like the neural net, the method to be described here can be applied rather quickly to each jet. It may therefore prove useful in experimental triggers, which are essential for the study of rare processes.

II. METHODOLOGY

This study was performed by generating simulated events corresponding to $\bar{p}p$ scattering at the present Tevatron energy $\sqrt{s} = 1.8\,\text{TeV}$ and analyzing them using a calorimeter simulation and a jet-finding procedure. The key feature of using a simulation in place of actual data is, of course, that the quark or gluon parentage of each jet is known. A detailed description of the procedure is as follows.

The Monte Carlo event generator HERWIG [2] (version 4.3, July 1990) was used to generate events of $2 \to 2$ hard parton scattering. HERWIG includes initial- and final-state gluon radiation and soft background processes. It models the hadronization of the partons via clusters, and proceeds through the decays of unstable particles to produce a list of the momenta of all final particles in each event.

An idealized model of a D0-type calorimeter was created, consisting of 62×63 cells ("towers") of size 0.1×0.1 in (η, ϕ), which cover $-3.1 < \eta < 3.1$ and all of ϕ. The energy of each final particle was modified to include a Gaussian energy resolution with standard deviation

$$\Delta E = \begin{cases} 0.55\sqrt{E \times (1\ \text{GeV})} & \text{for hadrons and muons,} \\ 0.15\sqrt{E \times (1\ \text{GeV})} & \text{for electrons and photons.} \end{cases} \tag{1}$$

This modified energy was spread out over a circle $r < r_0$ in (η, ϕ) according to the distribution

$$\frac{dP}{d\eta\, d\phi} \propto (r_0^2 - r^2)\ , \tag{2}$$

where $r = [(\eta - \eta_0)^2 + (\phi - \phi_0)^2]^{1/2}$ is the deviation from the particle's direction (η_0, ϕ_0). Shower spreading for photons and electrons was omitted, to approximate the fact that the electromagnetic part of the D0 calorimeter has finer segmentation than the hadronic part. The form given by Eq. (2) for shower spreading (with $r_0 <$ cell size) is convenient because the energy of each particle spills over into at most the 8 cells adjacent to the cell in the direct path of the particle, and the energy fractions deposited can be expressed in a fairly simple closed form. Using $r_0 = 0.08$, Eq. (2) appears to give approximately the correct amount of shower spreading [11]. A more complete model would include a tail extending to more distant cells, but this would not be expected to affect our qualitative results.

Jets were recognized in this simulation by a cone-type jet-finding algorithm. The algorithm follows the general guidelines of the "Snowmass accord" [12], but it is more elaborate than what is spelled out in that document. It involves four distinct phases. In *phase I*, initial clusters are formed by associating every cell in the calorimeter which contains transverse energy E_T greater than E_1 with that one of its eight nearest neighbors having the highest E_T, provided that E_T is greater than E_2. This leads to preclusters (equivalence classes [13]) of cells. In rare cases where the extent of one of these preclusters is too large in a given event, the default values $E_1 = 0.5\,\text{GeV}$ and $E_2 = 1\,\text{GeV}$ are increased, and phase I is repeated. This clustering method has the advantages of being quickly calculated, independent of the order in which the cells are examined, and not overly driven by single cells with large E_T. In the remaining phases of jet identification, the order of processing *can* matter; it is dealt with by working on the clusters in order of their total E_T at each stage, beginning with the largest. In *phase II*, the jet axis is computed for each jet, using the appropriate weighted average of the E_T's of its cells. Cells are then reassigned to the nearest jet axis, provided that they lie within R_{jet} of that axis. This phase is repeated until the result is stable. In *phase III*, nearby pairs of jets are merged into a single jet if the resulting combined jet includes at least 50% of each original E_T and at least 90% of their total. Phases II and III were repeated until the result was stable. In *phase IV*, jets with E_T less than the minimum value under consideration were dropped, making their cells available for reassignment. Then phases II – IV were repeated until the result was stable. Cells containing transverse energy less than 100 MeV were ignored in the analysis.

A moderate cone radius $R_{\text{jet}} = 0.8$ was used to define the jets. Too small a radius would presumably hamper the ability to distinguish quarks from gluons on the basis of nearly resolvable branchings.

In this way, 33 269 light-quark + gluon jets were generated in the observed range $40\,\text{GeV} < E_T < 60\,\text{GeV}$, $-2 < \eta < 2$. They correspond to a cross section of 2390 nb, of which 25.7% were light quarks (u, d, s, $\bar{u}$, $\bar{d}$, $\bar{s}$). An additional 25 901 light-quark + gluon jets were generated in the observed range $140\ \text{GeV} < E_T < 160\ \text{GeV}$, $-2 < \eta < 2$. They correspond to a cross section of 1.6 nb of which 48.2% were light quarks. In this E_T range, 7% of the jets were ignored because they came from c or b quarks, or because their parton parentage was unclear. When corrected for a different acceptance in η, these cross sections are consistent with the CDF measurements [1].

Some typical jets are shown in Fig. 1, using a display that contains the same information as the traditional "Lego plot": the transverse energy in each 0.1×0.1 cell is represented by one dot for each 0.250 GeV of transverse energy in the cell. Within each cell, the dots are located in a random way that avoids nearby dots.

A large number of these displays were examined to check that the cell assignments made by the jet-finding

algorithm look reasonable. (It is quite easy to make up jet finders that occasionally assign cells to jets in an *unreasonable* manner. Simply using the hottest cells as the centers of the jets, or as the initial jet axis for an iteration, for example, should be avoided for detailed purposes such as ours.)

Most of the E_T of each jet comes from cells that lie well within the chosen cone size $R_{\text{jet}} = 0.8$, as seen in Fig. 1. Hence the observed jet energy is not overly sensitive to the arbitrary size and circular shape chosen to define the jet.

III. VARIABLES WHICH DISCRIMINATE BETWEEN QUARKS AND GLUONS

The goal was to find a single variable whose probability distribution for the ensemble of known quark jets is sig-

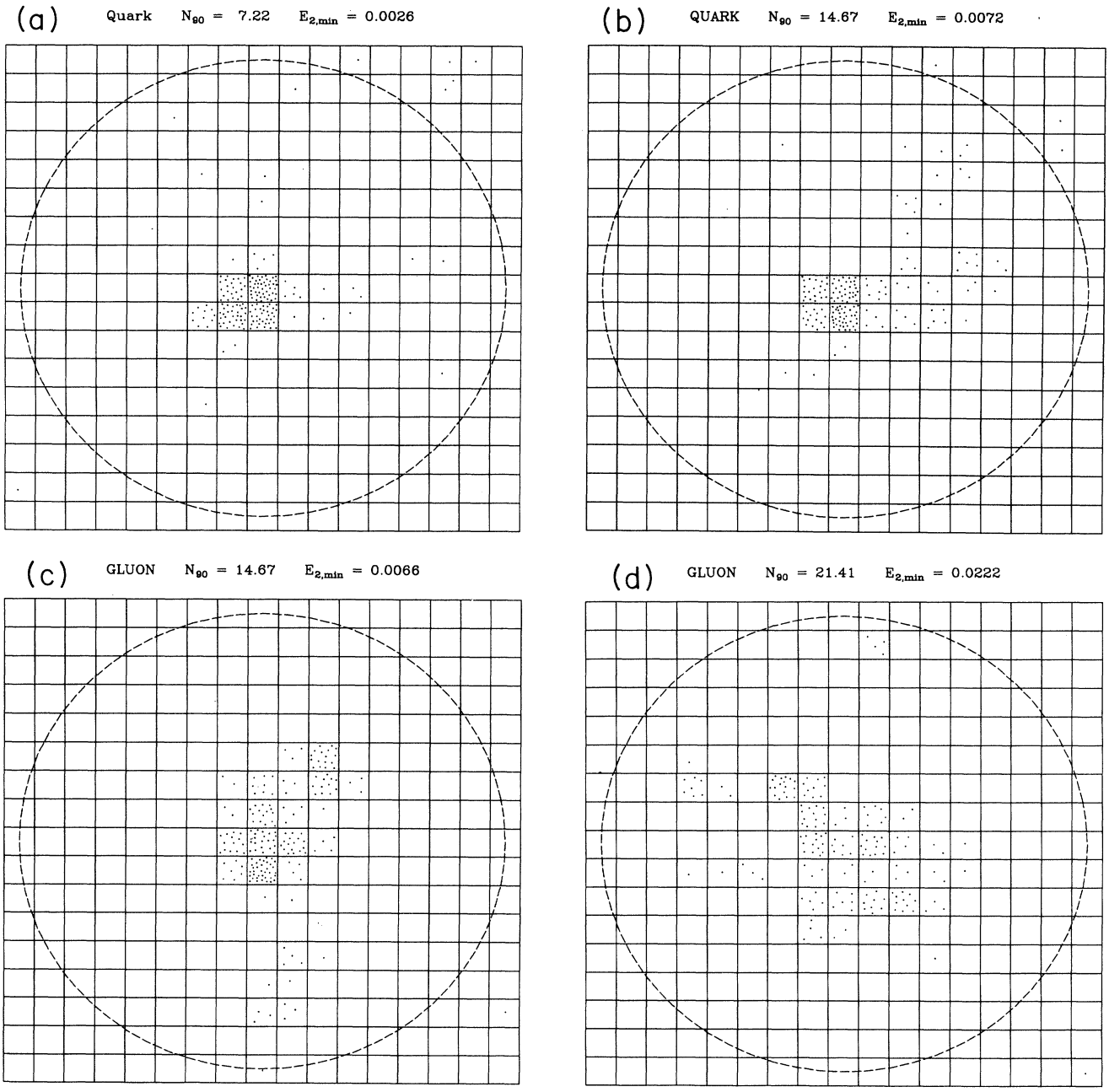

FIG. 1. Typical patterns of E_T deposition for quark jets [(a), (b)] and gluon jets [(c), (d)] at $E_T = 50\,\text{GeV}$. The E_T deposited in each 0.1×0.1 cell is represented by one dot per 0.250 GeV. The dashed line shows the assumed cone size $R_{\text{jet}} = 0.8$.

nificantly different from that for the ensemble of known gluon jets. An unknown jet could then be assigned a likelihood of being a quark or a gluon, based on its value of that variable. The desired variable must somehow measure the degree of branching of a jet. Recipes for two of the most successful variables found are as follows.

N_{90}: Add up the E_T of cells within the jet cone in descending order, and count the number N_{90} of cells needed to obtain 90% of the total for all cells within the cone, including a fractional part obtained by interpolating to account for the last partial cell. A good feature of N_{90} is that it is insensitive to energy spilling over from one cell to an adjacent cell, either due to the development of the QCD shower, or due to particle decays and spreading of showers within the calorimeter, so long as both cells are in the "hot" part of the jet.

N_{90} tends to be smaller for quark jets than for gluon jets, as shown in Fig. 2. The distributions are sufficiently

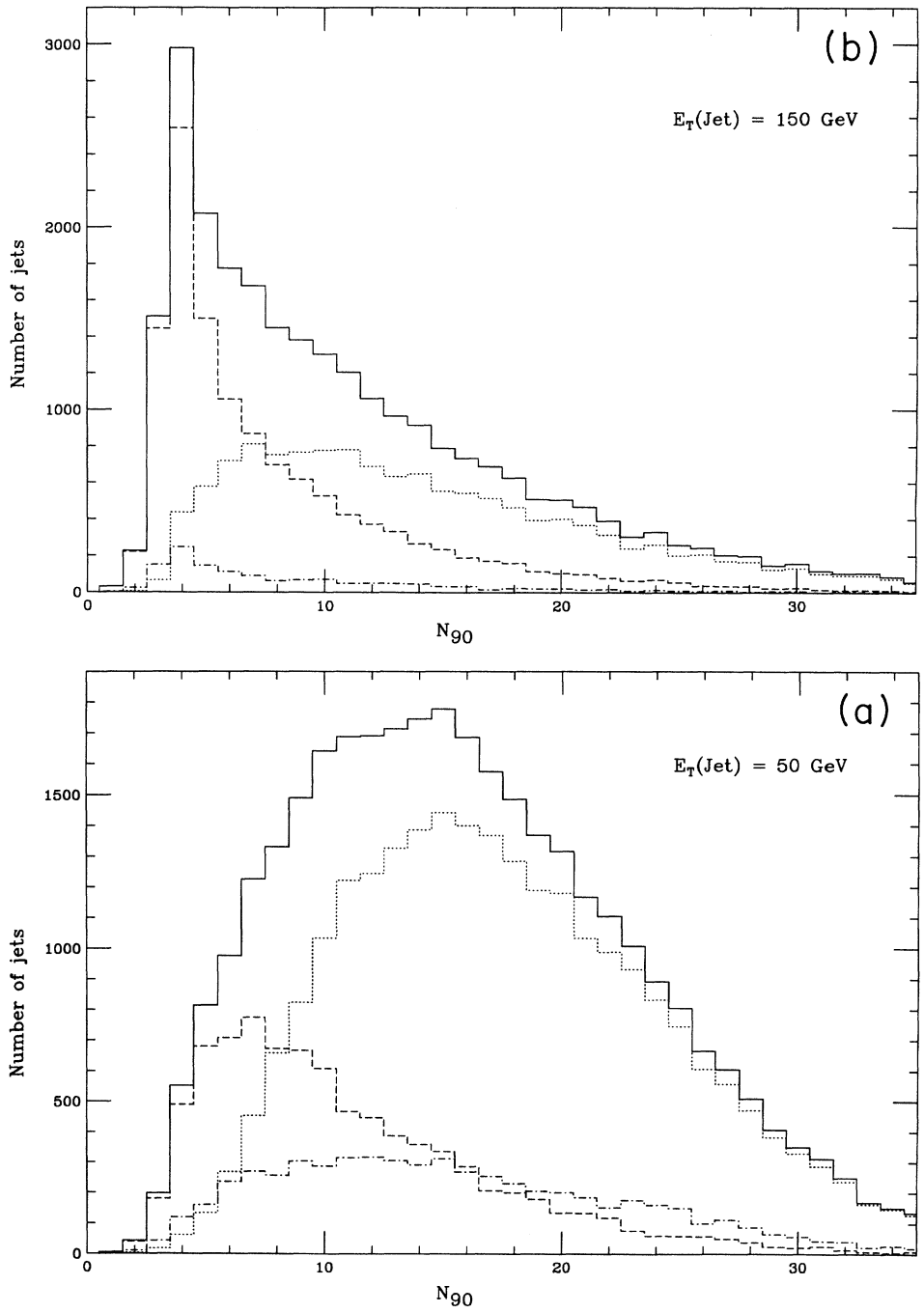

FIG. 2. N_{90} distributions for quarks (dashed), gluons (dotted), and their sum (solid) at (a) E_T = 40–60 GeV and (b) E_T = 140–160 GeV. The dot-dashed curve shows the total when a cut $|\eta| > 1.5$ is made to enhance the q/G ratio.

different that jets with the lowest values of N_{90} are almost all quarks, while those with the highest values are almost all gluons.

Two additional effects can be seen in going from $E_T = 50\,\text{GeV}$ [Fig. 2(a)] to $E_T = 150\,\text{GeV}$ [Fig. 2(b)]: (1) The ratio of quarks to gluons increases, because the parton distributions favor quarks at the large values of x which are needed to make the larger E_T; and (2) the average N_{90} decreases; i.e., the jets become "more collimated" in (η, ϕ) at large E_T.

Other percentages are of course possible for defining this variable. 90% was about optimal for the jets with $E_T = 50$–$150\,\text{GeV}$ which were studied. However, N_{85} actually works somewhat better for $E_T = 50\,\text{GeV}$, while N_{95} works somewhat better for $E_T = 150\,\text{GeV}$. Hence one might prefer to define the measure as the number of cells needed to include all but E_0 of the jet E_T, where $E_0 \sim 5$–$10\,\text{GeV}$. A different minor variation would be to define the measure as the number of cells needed to include all but $C\sqrt{E_T}$ of the jet E_T, where the constant C is chosen as $\sim 1\,\text{GeV}^{1/2}$. This alternative is probably the most practical one, because it makes the eventual cut on N depend only weakly on E_T.

$E_{2,\min}$: Define moments by $\langle X \rangle = \sum E_T X$, where the sum $\sum$ runs over all cells within the jet cone, or as assumed here, over the cells which make up N_{90}. Locate the center (η_0, ϕ_0) of the jet such that the first moments are zero: $\langle \eta - \eta_0 \rangle = \langle \phi - \phi_0 \rangle = 0$. Compute the second moments $\langle (\eta - \eta_0)^2 \rangle$, $\langle (\phi - \phi_0)^2 \rangle$, and $\langle (\eta - \eta_0)(\phi - \phi_0) \rangle$. Diagonalize the 2×2 matrix of these to obtain the minimum and maximum possible second moments $E_{2,\min}$ and $E_{2,\max}$. Their sum $E_{2,\min} + E_{2,\max}$ is simply the second moment $\langle (\eta - \eta_0)^2 + (\phi - \phi_0)^2 \rangle$. Their ratio is a measure of the elliptical character of the jet transverse momenta. Among these variables, $E_{2,\min}$ was found to work the best, and to be almost but not quite as good as N_{90}.

A number of other candidate variables, such as measures of the total "contrast," or of the amount of fluctuation between neighboring cells, or measures analogous to $E_{2,\min}$ and $E_{2,\max}$ but linear instead of quadratic in the distance from the jet center in (η, ϕ), were invented and tested [14]. One variable that is closely related to N_{90}, and that worked nearly as well, is simply the number of cells in the jet that have $E_T > E_0$, where the threshold E_0 is chosen as 250 MeV, plus or minus a factor of 2. None of these variables were as successful as those described above, so they will not be discussed here further.

A comparison with the neural-net technique [9, 10] was made using jets generated and processed through the network algorithm by Lönnblad *et al.* [10]. It was found that N_{90} is substantially better at separating quarks from gluons than is the net parameter. This was true at all levels of the "efficiency" F_q, which will be defined in the next section. Incidentally, the jets used for this comparison were generated using PYTHIA, and processed through a very minimal calorimeter simulation and jet finder. This indicates that the success of N_{90} is not dependent of the details of the methodology defined in Sec. III.

Attempts were made to combine two or more of the various variables to improve the discrimination. These attempts were unsuccessful: N_{90} by itself was superior to any of the combined variables tried. This result was established rather generally, by comparing jets within narrow bands of N_{90}: quark and gluon jets that have the same N_{90} cannot be distinguished strongly on the basis of *any* of the other variables — or by any obvious aspect of their appearance in the "Lego plot." Better discrimination variables may of course remain to be discovered; but it is also quite possible that quark and gluon jets are indistinguishable much beyond the level provided by N_{90}, if one is restricted to calorimeter information. (Counting the number of *hadrons* with E_T greater than some threshold such as 100 MeV *does* produce a significantly better discrimination.)

The fact that the good variables cannot be combined to improve the discrimination implies that they are strongly correlated with each other. This can be seen by examining scatter plots involving pairs of variables. Information similar to a scatter plot is shown quantitatively in Table I, for the case of N_{90} and $E_{2,\min}$ at $E_T = 150\,\text{GeV}$. Table I was constructed by choosing bins for each variable that contain 20% of the jets. Hence the sum of each row and each column is 0.20. The correlations appear in fact that the largest elements are on or near the diagonal. The predictions of Table I would be useful to compare with experiment. They should be insensitive to unimportant details of the simulation, since they depend only on the rank ordering in the variables.

TABLE I. Correlation between N_{90} and $E_{2,\min}$. The fraction of jets in various bins of N_{90} and $E_{2,\min}$ is shown, for jets with $E_T = 150\,\text{GeV}$. The bins in each variable were chosen so that each row and each column contains 20% of the jets. The fact that entries far from the diagonal are small shows the strong correlation between these variables.

	$E_{2,\min}$				
	0.0000	0.0017	0.0025	0.0038	0.0077
N_{90}	– 0.0017	– 0.0025	– 0.0038	– 0.0077	– ∞
0.0–4.7	0.133	0.061	0.006	0.000	0.000
4.7–7.6	0.058	0.091	0.048	0.002	0.000
7.6–11.4	0.008	0.043	0.102	0.044	0.003
11.4–17.4	0.000	0.005	0.043	0.112	0.040
17.4–∞	0.000	0.000	0.001	0.041	0.157

IV. QUARK-GLUON JET DISCRIMINATION

In this section, we examine quantitatively how quarks and gluons can be distinguished using the variable N_{90}.

The discrimination can be made at best on a probabilistic basis: for any given jet, one computes N_{90} and guesses that the jet was due to a quark if the quark probability curve is larger than the gluon curve in Fig. 2, at that value of N_{90}. If N_{90} is small, the jet is assumed to be a quark, and that assumption is very likely to be correct. Similarly, if N_{90} is large the jet is assumed with good probability to be a gluon. Intermediate values of N_{90} are more ambiguous, and if N_{90} is right at the crossover point of the q and G distributions, the chance of correctly assigning it is only 50–50.

Averaged over all jets, the fraction of jets correctly identified in this way is 80% at $E_T = 50$ GeV and 71% at $E_T = 150$ GeV. The variation with E_T here results mainly from the variation of the overall q/G ratio. For example, at $E_T = 50$ GeV, $G/(G+q) = 0.74$ so simply guessing that every jet is a gluon would already be correct 74% of the time. For jets that are an equal mixture of q and G, which could be made by a cut on $|\eta|$, the probabilities for correct identification become 71% at both $E_T = 50$ GeV and $E_T = 150$ GeV.

The proportion of jets correctly identified is somewhat smaller than has been reported on the basis of neural net methods [9, 10]. This results from different assumptions involved in the calorimeter simulation and jet-finding algorithm, since when both methods are applied to the same jet input data, the N_{90} method is clearly superior.

For many applications it is not necessary to identify every jet as a quark or gluon. Suppose, for example, that we are particularly interested in quarks to select $W^{\pm} \to jj$ candidates. We can make a cut $N_{90} < n$ to separate well-identified quarks from jets that are either gluons or ambiguous. The choice of n controls the "efficiency" $F_q(n)$ = fraction of quarks that pass the cut. $F_q(n)$ is the integral up to n of the quark histogram in Fig. 2, divided by the integral over all N_{90}. Similarly define $F_G(n)$ = fraction of gluons that pass the cut $N_{90} < n$. Then $F_G(n)/F_q(n)$ is the factor by which the gluon background is suppressed relative to the quark signal. Equivalently, $F_q(n)/F_G(n)$ is the factor by which signal/background is enhanced. Tighter cuts give a greater enhancement of signal/background, at the expense of lower efficiency, i.e., more quark jets thrown away.

The relation between $F_q(n)$ and $F_G(n)$ is shown in Fig. 3. As a practical example, the point $F_q = 0.33$, $F_G = 0.033$ means that one can make a cut that keeps $\frac{1}{3}$ of all quark jets, while enhancing the quark/gluon ratio by a factor of 10. (The cuts that accomplish this are $N_{90} < 7.1$ at $E_T = 50$ GeV, $N_{90} < 4.2$ at $E_T = 150$ GeV.) As an example of a looser cut, $F_q = 0.50$, $F_G = 0.11$, which keeps $\frac{1}{2}$ of all quark jets while enhancing q/G by 4.5, can be obtained by $N_{90} < 9.6$ at $E_T = 50$ GeV, $N_{90} < 5.9$ at $E_T = 150$ GeV.

Figure 3 shows that the ability to separate q and G is nearly independent of jet E_T — even though the cut locations, as given above, change considerably. Figure 3 also shows that two other variables $E_{2,\min}$ and $E_{2,\max}$ are not as good as N_{90}.

Figure 3 can also be read to see what happens when one uses these variables to select gluons instead of quarks. In that case, the cut is of the form $N_{90} > n$ instead of $N_{90} < n$, so $1 - F_G$ is the fraction of gluons that pass the cut, and $1 - F_q$ is the fraction of quarks. For instance, $1 - F_G = 0.33$ when $1 - F_q = 0.085$, so gluons can be selected with an efficiency of $\frac{1}{3}$ in a manner that enhances G/q by a factor of 3.9. Also $1 - F_G = 0.50$ when $1 - F_q = 0.16$, so gluons can be selected with an efficiency of $\frac{1}{2}$ in a manner that enhances G/q by a factor of 3.1. This ability to select gluons could be useful, even though it is not as strong as the ability to select quarks, which was discussed above.

V. CONCLUSION: TESTS AND APPLICATIONS

We have seen that according to Monte Carlo jet simulation, it is possible to distinguish light-quark jets from gluon jets on a probabilistic basis, using N_{90}, which is defined as the minimum number of calorimeter cells required to include 90% of the jet E_T. The next steps should be to test and apply the method using real data.

Tests of the method can be based directly on the N_{90} distribution. As shown in Fig. 2, the contributions due to q and G have peaks in different places. The q and G distributions are not so different as to produce separate peaks in their sum, which is the only observable quantity.

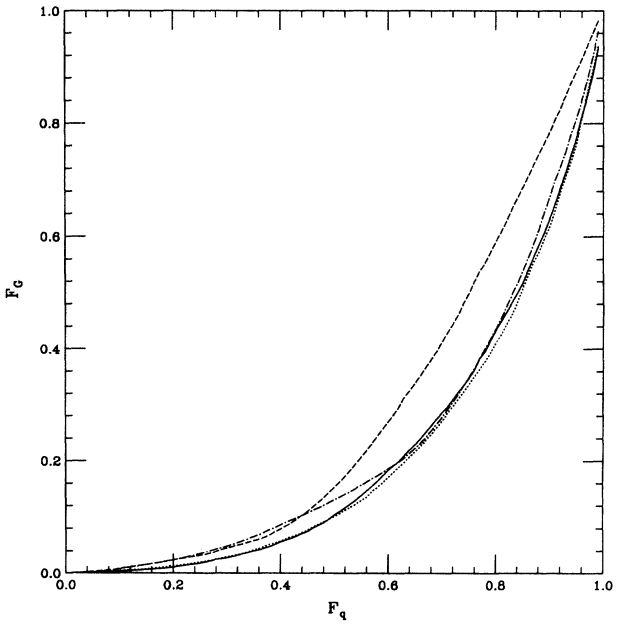

FIG. 3. The fraction of gluon jets F_G, versus the fraction of quark jets F_q, which pass cuts of the form $N_{90} < n$. The solid curve ($E_T = 150$ GeV) and the dotted curve ($E_T = 50$ GeV) are nearly the same. Cuts of the form $E_{2,\min} < e$ (dot-dashed curve) and $E_{2,\max} < e$ (dashed curve) for $E_T = 150$ GeV have F_q/F_G closer to 1, and hence are less effective for q/G separation.

However, the shape of the sum at low N_{90} clearly displays the presence of the quark contribution. In comparing the N_{90} distribution with experiment, it will be necessary to use a reasonably small range of E_T for the jets, because the typical N_{90} values decrease with increasing E_T.

In the interest of conceptual simplicity, we have focussed on the variable N_{90} here and below. However, it would be preferable to use instead a parameter N, which is defined as the *minimum number of cells in the jet which are needed to include all but* $C\sqrt{E_T}$ *of the jet* E_T. For $C = 1\ \mathrm{GeV}^{1/2}$, this is equivalent to N_{86} for 50 GeV jets and to N_{92} for 150 GeV jets. A small advantage of this variable is that it provides slightly better quark/gluon discrimination at each of these values of E_T. A more important advantage is that it makes the N distribution less dependent on the jet E_T, so that a much wider range of E_T can be used.

To make a quantitative comparison of the N_{90} distribution with experiment, it will be necessary to replace the energy resolution [Eq. (1)] and spreading [Eq. (2)] models with more detailed forms which are tailored to the actual calorimeter — including effects due to dead regions between cells. An additional test of this work would be to see if the correlations between variables, such as those shown in Table I, are in agreement with experiment.

More specific tests of the quark-gluon separation could be made by observing the changes in shape of the N_{90} distribution which result from varying the ratio of quarks to gluons in the sample using methods not directly related to N_{90}. One way to do that is through cuts on E_T: for example, the change in shape between the solid curves of Fig. 2(a) and Fig. 2(b) results mainly from the fact that the quark fraction $q/(q+G)$ increases from 26% to 48% in going from $E_T = 50$ GeV to $E_T = 150$ GeV. A second way to change the q/G ratio is through cuts on pseudorapidity. This can be seen from the dot-dashed curves in Fig. 2, where the cut $|\eta| > 1.5$ has increased $q/(q+G)$ from 26% to 36% at $E_T = 50$ GeV and from 48% to 79% at $E_T = 150$ GeV, thereby moving the peak of the distribution toward smaller N_{90}.

A further test of these ideas could be made by comparing the N_{90} distributions of jets from 2-jet events with jets produced in recoil against high-p_T direct photons, which according to QCD calculations will have a different q/G ratio at a given E_T. One could also try to compare the cross sections for qq, qG, and GG contributions to 2-jet events, in various regions of η.

The quark-gluon separation could also be tested using e^+e^- annihilation data, by comparing jets from symmetric 3-jet events, which are presumably $q\bar{q}G$, with 2-jet events at the same jet energy ($\frac{2}{3}$ of the $\sqrt{s}$), which are presumably $q\bar{q}$; or by simply assuming that the least energetic event in a 3-jet event is a gluon [15].

The quark-gluon separation can be used to enhance quark jet signals. For as shown in Fig. 3, one can make a cut on N_{90} that passes a sizable fraction F_q of quark jets, while passing a much smaller fraction F_G of gluon jets. This cut may work fairly well even if the N_{90} distributions predicted by the simulation do not turn out to be correct in quantitative detail: it seems likely that most of the jets at the low end of the N_{90} distribution in a given experimental situation will be due to quarks. These quark jets are particularly easy to look for, since most of their E_T appears in a small number of cells. Those cells are also rather close together, as implied by Table I, so a smaller cone size could be used to find them.

An important by-product of this work is the opportunity to compare the jet E_T measured in the calorimeter with the P_T of the original $2 \to 2$ hard scattering. The energy resolution assumed in Eq. (1) for each particle in the jet leads directly to a root-mean-square (rms) uncertainty of $([0.55E_T(\mathrm{hadronic})]^2+[0.15E_T(\mathrm{em})]^2)^{1/2}$ in the measured jet E_T. An average fraction 0.3 of the jet energy is electromagnetic, so this amounts to an rms error of $0.46\times\sqrt{E_T/(1\ \mathrm{GeV})}$ in the jet E_T measurement, just as a result of the energy resolution of the calorimeter. The distribution of $P_T(\mathrm{hard})-E_T(\mathrm{measured})$ is found to have a peak centered at about -5 GeV, i.e., there is a small systematic shift in the measured energy, which is mainly caused by the rapidly falling dependence of the cross section on E_T. The width of the peak is less than a factor of 2 larger than the width that would result simply from the rms error in the calorimeter energy measurement given above. Hence the P_T of the hard scattering is rather well measured.

A number of *applications* for the quark-gluon separation are possible. It could be used to enhance the signal/background ratio in searching for the top quark in final states where a $W^{\pm}$ decays to jets, and/or where extra jets are present [16]. It might also make it possible to observe the hadronic decays $W^{\pm} \to j\,j$ and $Z^0 \to \mathrm{jj}$, which would allow the W and Z masses to be used as benchmarks for calibrating the calorimeter. Still more optimistically, it might even provide a useful measurement of M_W/M_Z which is free of systematic errors. This is currently under investigation. Finally, one could study the dependence on jet identity of the E_T deposited *between* two jets (the "string effect" [17]).

ACKNOWLEDGMENTS

I wish to thank B. Webber for discussions and L. Lönnblad for discussions and for providing a sample of neural-net-analyzed jets. I also wish to thank J. Linnemann for discussions of the D0 calorimeter.

[1] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **62**, 613 (1989).
[2] G. Marchesini and B. R. Webber, Nucl. Phys. **B310**, 461 (1988); **B330**, 261 (1990); I. G. Knowles, *ibid.* **B310**, 571 (1988).
[3] F. E. Paige and S. D. Protopopescu, ISAJET 6.21.
[4] H-U. Bengtsson and T. Sjöstrand, PYTHIA 5.4; Comput. Phys. Commun. **46**, 43 (1987).

[5] D0 Calorimeter Group, M. Abolins *et al.*, Nucl. Instrum. Methods A **280**, 36 (1989).
[6] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).
[7] Lorella Jones, Phys. Rev. D **39**, 2550 (1989); **42**, 811 (1990).
[8] G. Sterman and S. Weinberg, Phys. Rev. Lett. **39**, 1436 (1977).
[9] L. Lönnblad, C. Peterson, and T. Rögnvaldsson, Phys. Rev. Lett. **65**, 1321 (1990); Nucl. Phys. **B349**, 675 (1991).
[10] P. Bhat, L. Lönnblad, K. Meier, and K. Sugano, in *Research Directions for the Decade*, Proceedings of the Summer Study, Snowmass, Colorado, 1990 (Editions Frontières, Gif-sur-Yvette, in press).
[11] G. Apollinari *et al.*, Nucl. Instrum. Methods A **267**, 301 (1988).
[12] J. Huth, N. Wainer, K. Meier, N. Hadley, F. Aversa, M. Greco, P. Chiappetta, J. Guillet, S. Ellis, Z. Kunszt, and D. Soper, in *Research Directions for the Decade* (Ref. [10]).
[13] S. Youssef, Comput. Phys. Commun. **45**, 423 (1987).
[14] J. Pumplin, in *Research Directions for the Decade* (Ref. [10]).
[15] AMY Collaboration, in *Proceedings of the 25th International Conference on High Energy Physics*, Singapore, 1990, edited by K. K. Phua and Y. Yamaguchi (World Scientific, Singapore, 1991).
[16] H. Baer, V. Barger, and R. J. N. Phillips, Phys. Rev. D **39**, 3310 (1989).
[17] Yu. Dokshitzer, V. Khoze, S. Troyan, and A. Mueller, Rev. Mod. Phys. **60**, 373 (1988).