TELLING TOP JETS FROM QCD JETS USING ENERGY FLOW

R. Odorico
Istituto di Fisica dell'Università and INFN, Bologna

Abstract

By means of explicit calculations using simulation programs for top particle jets it is shown that a pattern recognition study of the jet energy flow profile can bring down the QCD contamination to top particle jets by substantial amounts without sensibly reducing top jet statistics. In the example considered, which makes use of already existing simulation programs, the QCD contamination is brought down by about two orders of magnitude by using Fisher's discriminant analysis for the recognition of jet profiles.

There are two possible strategies for isolating top jets from ordinary QCD jets in high energy events to be observed at hadronic colliders.

i) One can elaborate a set of tight requirements on the trigger and/or the software selection afterwards, which should be met only unfrequently by ordinary jets and are likely to almost single out top particle jets (e.g., presence of two fast same sign leptons close in angle)[1]. In this way the absolute statistics for top events is also strongly reduced, though, and the viability of this approach may be limited to high luminosity machines.

ii) Top particle jets possibly obey general patterns uncommon among ordinary QCD jets. After all the dynamics in the two cases is rather different. For top particles one has, essentially, a spherically symmetric decay Lorentz boosted into a jet. For QCD jets the underlying mechanism consists from the very start of a limited p_T development along a given axis. In both cases we have well defined theoretical schemes on the dynamics governing the jet structure, and these schemes can be embodied in event-generating simulation programs from which the resulting jet characteristics can be studied with standard statistical tools and in the perspective of a pattern recognition approach. If it turns out from such a study that some global patterns are common among top particle jets and uncommon among QCD jets, recognition of these patterns can be exploited to reduce the contaminating QCD background without affecting much top particle statistics.

The two types of strategies are by no means mutually exclusive, and can usefully complement each other.

Of course the big question mark on the exploitation of jet patterns is whether discriminating patterns exist at all.

Purpose of this note is to report on the results of a preliminary study which seems to provide a positive answer to that. We use already existing simulation programs for top and QCD jets, and make an exploratory study of the energy flow patterns in the two cases. Using a simple parametrization of the energy flow profiles and employing standard statistical techniques (Fisher's discriminant analysis)[2] we show that it is possible to achieve interesting rejection rates for QCD jets without essentially affecting top particle statistics.

The simulation program for QCD jets that we use is described in ref. 3. The perturbative QCD cascade is handled by a leading pole QCD Monte Carlo, and quanta are converted into hadrons independently according to the Field-Feynman algorithm. The perturbative QCD cascade is stopped when (timelike) quanta reach an invariant mass $Q_o \simeq 3$ GeV, as dictated by fits to e^+e^- annihilation data ($\Lambda = 0.5$ GeV is assumed)[4]. For the F.F. algorithm the Fortran code of ref. 5 is used. Gluons are fragmented as quarks of random flavor. For the top particle jets we adopt the model of ref. 6 in the Fortran code of ref. 5.

The calculations that we present concern the QCD background to top jets of given invariant mass and total momentum. Let us assume $M_T = 30$ GeV for the top quark mass, and let us consider jets of total longitudinal momentum $P_{JET} = 50$ GeV/c. As possible source of contamination we then consider QCD jets of the same momentum and invariant mass $M_{JET} = 30$ GeV $\pm$ 20%. The same invariant mass cut is imposed on top particle jets, which do not necessarily have the same mass as the top quark: top particles are in general heavier than the top quark, and decay hadrons which move backwards are lost, thus reducing the invariant mass which is effectively observed.

Fig. 1 shows the distribution of transverse energy $\sum p_T$ in pseudorapidity η, with respect to the jet axis, for top jets and gluon QCD jets. Although the integrated $\sum p_T$ are essentially the same for the two kinds of jets, their profile in η are rather different, as one may expect from the different dynamics governing them. One can try to exploit such a difference so as to discriminate between the two kinds of jets on an event per event basis.

Those plotted in Fig. 1 are mean profiles. From event to event such profiles undergo fluctuations, of course, and it is the amount of such fluctuations which determines the possibility of exploiting the differences in profile for the sake of discrimination between the two kinds of jets.

It is convenient to parametrize jet profiles with a set of shape parameters, much in the same way as in e^+e^- annihilation one uses sphericity, aplanarity or, e.g., the shape parameters of ref. 7. In our case a convenient set might be found in the ensemble of double moments

$$M^{nm} = \sum \eta^n p_T^m$$

where the sum is extended over all particles in the jet. But of course one may think of many other possibilities.

Useful shape parameters are those for which fluctuations are small with respect to the difference of mean values for the two types of jets. As a matter

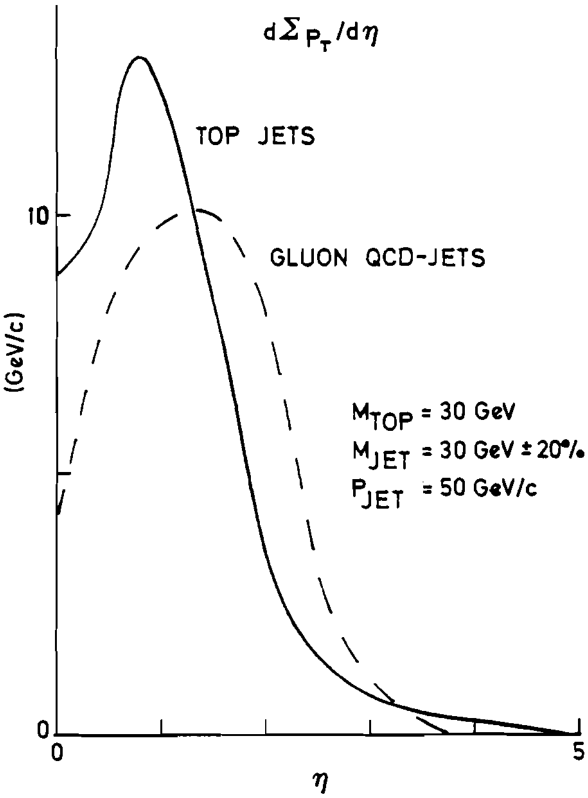

Fig. 1 - Mean distributions of transverse energy Σp_T in pseudorapidity η for top particle jets and gluon QCD-jets.

of fact some shape parameters exhibit a remarkable stability, which others do not have. In Fig. 2 a, b and c we show how jets of the two types distribute in $\Sigma \eta\, p_T / \Sigma p_T$, $\Sigma \eta^2\, p_T$ and $\Sigma \eta^3\, p_T$, respectively. Light quark QCD-jets have distributions very similar to those of gluon jets of the same invariant mass and momentum. Corresponding curves are omitted for simplicity. One can rationalize the observed stability realizing that the particle multiplicity in jets of the invariant mass and momentum which we are considering is rather high, as it can be seen in Fig. 2d, showing the distributions in the multiplicity of particles with $\eta > 1$. Furthermore, stability is lost if in the above parameters one replaces the linear weight in p_T with a quadratic or cubic one. Theoretically, the linear weight in p_T makes the variables "infrared stable", or in other words it makes more likely for them to pick up less fluctuations from the fragmentation process. This does not prevent, though, that even with p_T^2 or p_T^3 weights one may obtain relatively stable variables by taking appropriate combinations of moments. Fig. 2e, e.g., shows the distributions in $(\Sigma \eta^2\, p_T^3 / \Sigma \eta\, p_T^3 - 1)$, which are not less stable than in the previous cases.

Instead of moments one can of course consider a host of other possibilities. For instance Σp_T with η cuts. Fig. 2f shows the distributions in Σp_T ($\eta > 1$). In spite of the fact that there is considerable overlap between the two types of jets, a cut at low Σp_T ($\eta > 1$) would substantially enrich the top-jet/QCD-jet ratio, although at the expense of some top statistics. This very preliminary study certainly should not be regarded as exhaustive in the exploration of optimal ways to parametrize jet profiles.

Distributions in the shape parameters of Fig. 2 can be used one at the time to get inspiration about cuts leading to an enrichment of the top/QCD ratio in the jet sample. But of course this would mean to make a very poor use of all the experimental information available on the energy flow. One can do a more effective job by considering a number of shape parameters at the same time and analysing them by means of a well known statistical tool: Fisher's discriminant analysis[5]. If we "measure" the two types of jets with a set of shape parameters $M_1, M_2, \ldots, M_N$, Fisher's discriminant analysis acts upon the user's provided samples for the two kinds of jets and finds the distance variable

$D^* = \lambda_1 M_1 + \lambda_2 M_2 + \ldots + \lambda_N M_N$ which maximizes the "distance" between the two samples in the multidimensional space of the variables $M_1, M_2, \ldots, M_N$.

The "distance" between the two groups can be quantitatively defined in a number of ways (e.g., distance between the two centroids). Fig. 3 illustrates its working for the simple case of only two jet measures M_1 and M_2. The two groups overlap in M_1 and M_2, if these are considered separately. But if we take a rotated axis lined up with the two centroids there is no longer overlap (in the idealized example of the figure). Looking for such an axis is very similar to look for the sphericity axis in e^+e^- annihilation events. Fisher's discriminant analysis, though, does not consider only rotations but also variations of the scales of the variables. Once D^* is found on the basis of validated samples for the two kinds of jets provided by the user, a new "a priori" unknown jet can be classified, with a certain probability, according to the value it takes on D^*.

For the sake of illustration we have considered the first five shape parameters of Fig. 2 and submitted a thousand top jets and a thousand gluon QCD-jets measured by them to Fisher's discriminant analysis. The resulting distributions in the discriminant variable D^* are shown in Fig. 4 (*). By drawing away all jets with $D^* > 0$ one reduces the QCD background by about an order of magnitude with only a minimal loss of top jet statistics. For global events, with two top particles produced, independent application of the D^* cut to the

(*) Denoting by M_1, M_2, M_3, M_4 and M_5 the first five variables of Fig. 2, the D^* resulting from the statistical analysis has the form:

$$D^* = -2.70\text{E }0 * M_1 + 1.17\text{E-}1 * M_2 + 3.35\text{E-}2 * M_3 + 6.32\text{E-}2 * M_4 + 8.55\text{E-}1 * M_5 - 6.77\text{E }0$$

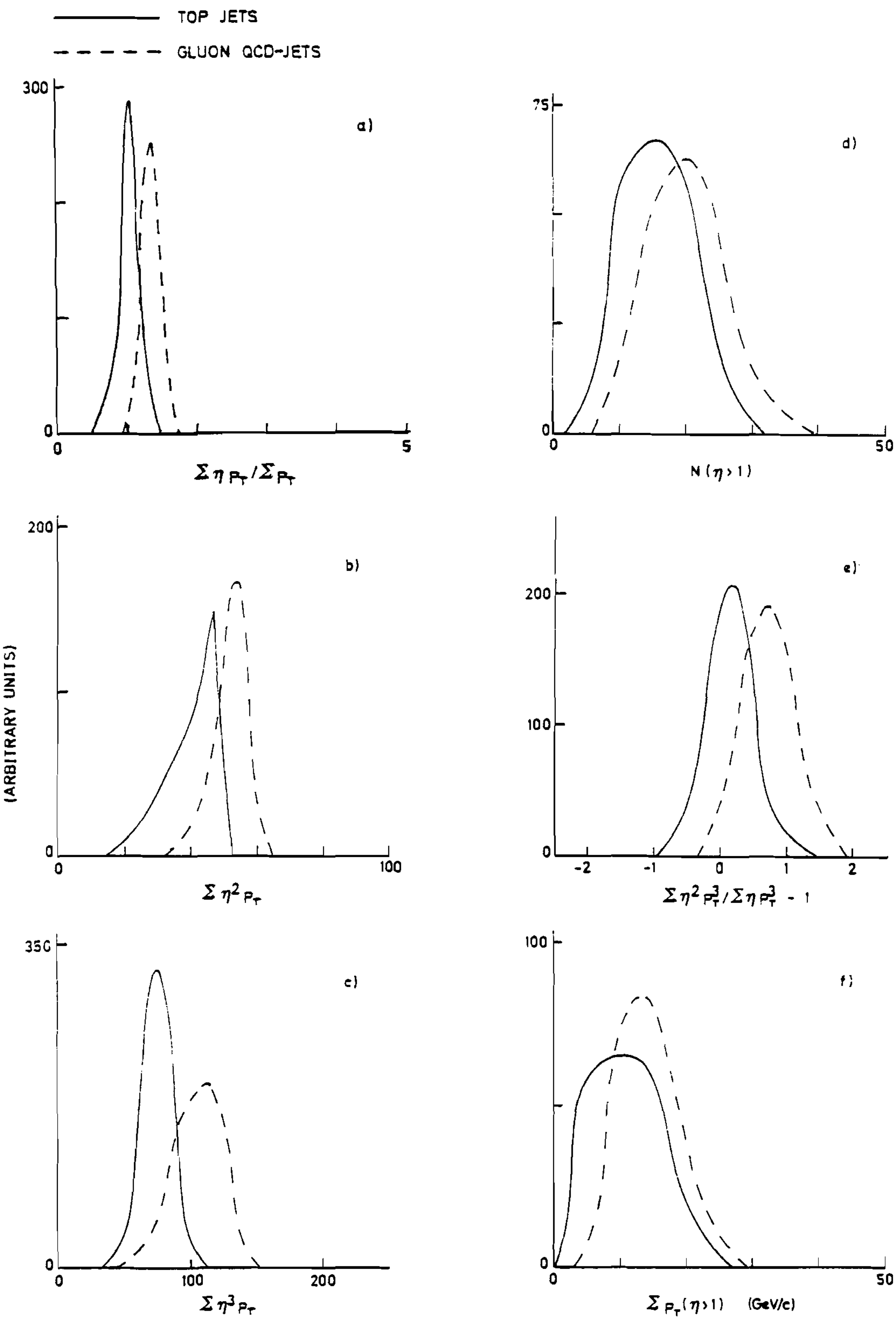

Fig. 2 - Distributions of top and gluon QCD jets in various shape parameters. a) $\sum \eta\, p_T / \sum p_T$; b) $\sum \eta^2 p_T$; c) $\sum \eta^3 p_T$; d) number of particles with $\eta > 1$; e) $\sum \eta^2 p_T^3 / \sum \eta\, p_T^3 - 1$; f) $\sum p_T$ for particles with $\eta > 1$.

two candidate top jets yields then a reduction of the QCD contamination by about 2 orders of magnitude, still without sensible loss of top particle statistics.

In conclusion the idea of using energy flow information to reduce the QCD background to top jets without substantial loss of top statistics is, to say the least, not a lost cause. What this very preliminary study shows is that the combined use of simulation programs and of standard statistical techniques can provide an effective handling of energy flow patterns characteristic of the various types of jets which may be useful for this purpose. One may question that in this study one has relied upon specific simulation programs. For QCD jets one will be able to exploit quite soon direct experimental information from the UA1 and UA2 experiments at the SPS $\bar{p}p$ collider[8]. For top jets one has of course to rely on the theory covering their decay patterns. One can possibly consider a wider range of top jet simulators reflecting the ambiguities left open by the theory. But the aim of this study is not to draw final conclusions, rather to point out that together with other selection criteria it is feasible and worthwhile to add a jet pattern recognition analysis when hunting for top particle jets.

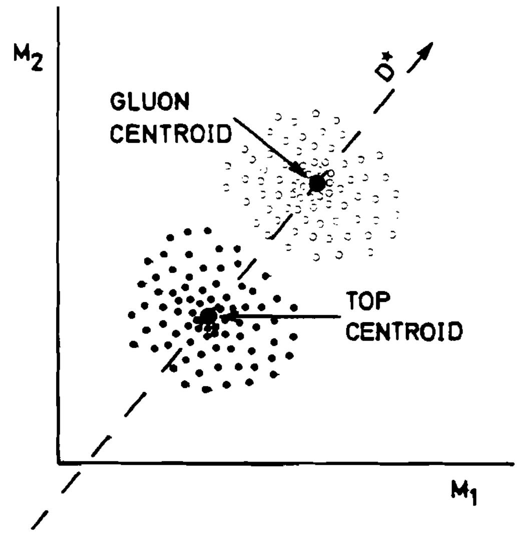

Fig. 3 - Sketch of an idealized example illustrating the working of Fisher's discriminant analysis.

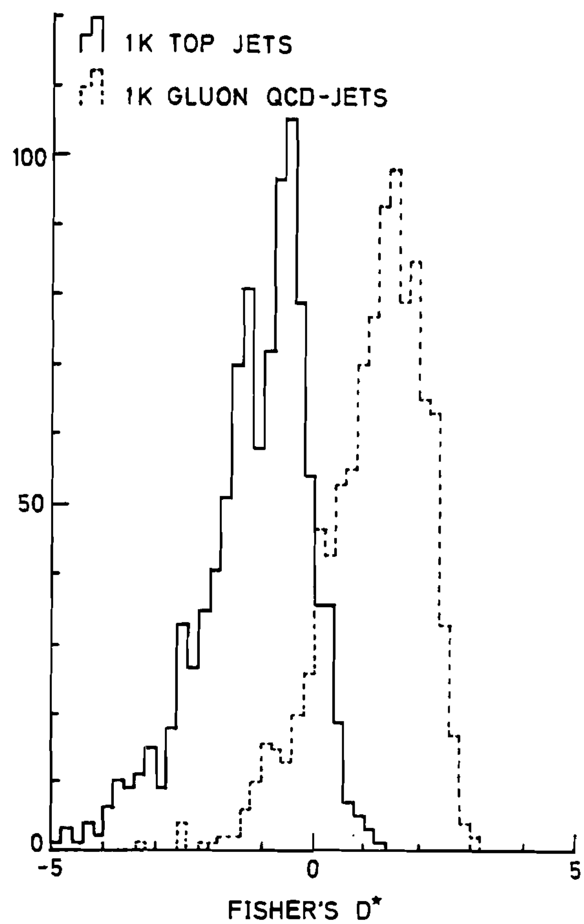

Fig. 4 - Distribution of top and gluon QCD jets in Fisher's discriminant variable D* associated with the first five shape variables of Fig. 2.

References

1) L.L. Chau, T. Ludlam, F.E. Paige, E.D. Platner, S.D. Protopopescu, P. Rehak, W.E. Cleland and J.A. Thompson, paper presented at the Mini Workshop on ISABELLE Experiments, BNL (1982), BNL preprint OG 637 (1982);
J. Leveille, contribution to the 1982 DPF Summer Study on Elementary Particle Physics and Future Facilities, and private communication;
D.M. Scott, Proton-Antiproton Collider Physics - 1981 (Madison, 1981), AIP Conf. Proc. No. 85, p. 151;
R. Horgan and M. Jacob, CERN preprint TH-3164 (1981);
D. DiBitonto, CERN preprint (1982).
2) N.H. Nie et al., SPSS Manual (Statistical Package for the Social Sciences), McGraw-Hill, New York (1975).
3) R. Odorico, Nucl. Phys. B172, 157 (1980).
4) P. Mazzanti and R. Odorico, Phys. Lett. 95B, 133 (1980); Z. Physik C 7, 61 (1980).
5) T. Sjöstrand, Lund preprint LU TP 79-8 (1979).
6) A. Ali et al., DESY reports 78/47 (1978) and 79/12 (1979).
7) G.C. Fox and S. Wolfram, Nucl. Phys. B149, 413 (1979).
8) A. Astbury (UA1 Collaboration), Talk at the International Conference on High Energy Physics, Paris (1982);
A.G. Clark and J-P. Repellin (UA2 Collaboration), Talks at the International Conference on High Energy Physics, Paris (1982);
M. Banner et al. (UA2 Collaboration), CERN preprint (1982), submitted to Phys. Lett.;
M. Della Negra and G. Thompson (UA1 Collaboration), Talks at the Europhysics Study Conference on Jet Structure from Quark and Lepton Interactions, Erice (1982);
J.L. Siegrist (UA2 Collaboration), Talk at the Europhysics Study Conference on Jet Structure from Quark and Lepton Interactions, Erice (1982).