HYPER VERSION

Received: *April 16, 1998*, Accepted: *August 10, 1998*

New and old jet clustering algorithms for electron-positron events

Stefano Moretti

Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK
E-mail: moretti@v2.rl.ac.uk

Leif Lönnblad and Torbjörn Sjöstrand

Department of Theoretical Physics, Lund University, Lund, Sweden
E-mail: leif@thep.lu.se, torbjorn@thep.lu.se

Abstract: Over the years, many jet clustering algorithms have been proposed for the analysis of hadronic final states in e^+e^- annihilations. These have somewhat different emphasis and are therefore more or less suited for various applications. We here review some of the most used and compare them from a theoretical and experimental point of view.

Keywords: QCD, NLO Computations, Jets, Phenomenological Models.

JHEP08(1998)001

Contents

JHEP08(1998)001

1. Introduction

Clustering algorithms have come to be an indispensable tool in the study of multi-hadronic events. They take the large number of particles produced in high-energy scatterings and cluster them into a small number of 'jets'. Such a simplified characterization of the event should help focus on the main properties of the underlying dynamics. In particular, the reconstructed jets should reflect the partonic picture, and thus allow a separation of perturbative and non-perturbative QCD physics aspects. Of course, such a separation can never be perfect, since there will always be smearing effects that cannot be compensated, and since there is not even a well-defined transition from perturbative to non-perturbative QCD.

Jet finders can be applied to a variety of tasks. The number of well-separated jets found in an event sample allows a determination of an α_{s} value. The distribution in angles between jets can be used as test of the fundamental properties of QCD, such as the gluon spin and the QCD color factors. The flow of particles around jet directions probes soft physics, both perturbative and non-perturbative. The clustering of jets may help to identify massive particles, such as $W^{\pm}$ and t, or to search for new ones.

The essential ingredients of jet clustering algorithms are basically the same independently of the phenomenological applications. Nonetheless, the kinematics and dynamics of, e.g., e^+e^-, ep and $p\overline{p}$ collisions are sufficiently different that computational methods have to be modified accordingly (see, e.g., [1, 2]). We will in this paper concentrate on algorithms for electron-positron annihilations, where there are no spectator jets and thus schemes can be made especially simple.

Over the years, several algorithms have been proposed for the study of e^+e^- events. Recently, advances in the understanding of soft perturbative physics lead to the introduction of further ones [3]. This has made it even more difficult for a user to understand differences and to know which algorithm to use where. The purpose of the current paper is to review several of the existing jet finders and compare them in various ways. Neither the choice of algorithms nor the selection of comparisons is exhaustive, but it should still help give some useful hints. We also introduce a few new hybrid algorithms to better understand the results. By using several event generators, we cross-check our findings. In a sense, our study is an update of the corresponding one carried out in ref. [4], in view of the new algorithms that have been proposed since then [3, 5] and of the advent of LEP2.

The conclusions might seem disappointing at first glance: while some algorithms fare markedly less well than the better ones, there is not *one* single best choice that sticks out in *all* phenomenological contexts we have studied. However, this should be of no surprise. In fact, given the varied use, there need not exist one algorithm that is optimal everywhere. Instead, we will show that, depending on the tasks assigned to the algorithm and on the physics dominion where this is applied, it is often possible to clearly individuate the most suitable to use.

In doing so, however, one should always recall that *jets* are not a fundamental aspect of nature, but rather artificial objects used to describe certain classes of events where strong interactions take place. As a matter of fact, they have revealed themselves to be a powerful tool in understanding the development of partons into hadrons, that is, the transition from the perturbative to the non-perturbative phase of QCD. In particular, the so-called 'parton shower' picture, the approximate treatment of QCD exploiting only some enhanced terms of the underlying dynamics but taken into account to all orders (as opposed to a precise prediction to some fixed order) has furnished us with a consistent model that succesfully reproduces most of the current hadronic phenomenology. Therefore, we believe it somewhat natural to refer to this mechanism in the first place, as a good guidance in order to appreciate the different behaviours of the jet clustering algorithms that we will study. Nonetheless, we should remind the reader that parton cascades are probabilistic in their nature, while QCD is governed by the principles of quantum mechanics (i.e., knowledge of individual parton dynamics is unphysical, interference effects among diagrams exist, etc.). In other terms, although we will below consider individual Feynman graphs (especially in sect. 2) involving single quarks and gluons, the performances of a jet clustering algorithm must eventually be tested with hadronic final states, and ultimately quantified by the size of the experimental and theoretical errors that it induces in an actual measurement of a physical quantity.

In the following section we review the historical evolution of clustering algorithms and describe some of the more familiar ones. Sections 3 and 4 contain comparisons between algorithms, for next-to-leading-order (NLO) and resummed perturbative QCD results, jet rates, jet energy and angle reconstruction, $W^\pm$ mass reconstruction, and so on. Finally section 5 contains a summary and outlook.

2. Clustering algorithms

The first studies of jet structure in e^+e^- annihilations were undertaken to establish the spin 1/2 nature of quarks [6]. It was then only necessary to define a common event axis for two back-to-back jets, and for this purpose *event measures* such as thrust [7] or sphericity [8] are quite sufficient.

With the search for and study of gluon jets at PETRA it became necessary to define and analyze three-jet structures. It is possible to generalize thrust to triplicity [9], in which the longitudinal momentum sum is maximized with respect to three jet axes. The maximization procedure can be rather time-consuming, however, in view of the large number of possibilities to subdivide particles into three groups. Since three jets span a plane (neglecting initial-state QED radiation) special tricks are possible: if all particles are projected onto an event plane, they can be ordered in angle such that only contiguous ranges of particles need be considered as candidate jets. The tri-jettiness measure [10] uses the sphericity tensor to define the event plane, and thereafter finds the subdivision into three jets by minimizing the sum of squared transverse momenta, where each $p_\perp$ (or, equivalently, $k_\perp$) is defined relative to the jet axis the particle is assigned to.

Such special-purpose algorithms have the disadvantage that, first, one procedure is needed to determine the number of jets in an event and, thereafter, another to find the jet axes. The algorithms may also be less easily generalizable to an arbitrary number of jets, or very time-consuming. The task is not hopeless: with some tricks and approximations, thrust/triplicity can be extended to an arbitrary number of jet axes [11, 12, 13] and tri-jettiness to four jets [14]. However, alternatives were sought, and more generic jet algorithms started to be formulated. Several ideas were proposed and explored around 1980 [11, 13, 15, 16, 17]. Most were based on a binary clustering, wherein the number of *clusters*[1] is reduced one at a time by combining the two most (in some sense) nearby ones. The joining procedure is stopped by testing against some criterion, and the final clusters are called jets. An alternative technique, top-down rather than bottom-up, is that of the minimum spanning tree, where a complete set of links are found and then gradually removed to subdivide the event suitably [15].

The starting configuration for the binary joining normally had each final-state particle as a separate cluster, but some algorithms contained a 'preclustering' step [13, 16]. Here, very nearby particles are initially merged according to some simplified scheme, in order to speed up the procedures or to make them less sensitive to soft-particle production. The possibility of 'reassignment' between clusters was also used to improve on the simple binary joining recipe [11, 12, 13]. Normally all particles were assigned to some jet but, in the spirit of the Sterman–Weinberg jet definition [18], a few algorithms allowed some fraction of the total energy to be found outside the jet cones [16].

The distance measure between clusters always contained an angular dependence, explicit or implicit, while the energy/absolute momentum entered in different ways or not at all. As one example, of some interest to compare with later measures, we note the use of thrust/triplicity generalized to n-jet axes [11, 12, 13]:

$$T_n = \frac{1}{E_{\mathrm{tot}}} \max \sum_{i=1}^{n} |\mathbf{p}_i| = \frac{1}{E_{\mathrm{tot}}} \max \sum \sqrt{E_i^2 - m_i^2} \approx 1 - \frac{1}{2E_{\mathrm{tot}}} \min \sum \frac{m_i^2}{E_i} \,, \qquad (2.1)$$

where each $\mathbf{p}_i$ is obtained as the vector sum of the momenta of the particles assigned to jet i (of energy E_i and mass m_i). Thus a maximization of T_n is almost the same as a minimization of $\sum m_i^2$, except that more energetic jets also can have a larger mass. Note that the relation $m_{\mathrm{jet}}^2 \propto E_{\mathrm{jet}}$ is approximately respected by non-perturbative iterative jet fragmentation models [19, 20].

The algorithms thus were rather sophisticated. Seen from a modern perspective, the main shortcoming is that in those days they could only be tested against generators producing a fixed number of partons — two, three or, at most, four — based on a leading-order (LO) matrix-element (ME) description. Therefore a 'correct' number of jets existed, and criteria were devised to find this number. Those criteria tended to be rather complex, at times even contrived, and thus often over-shadowed the simplicity of the basic algorithm. However, there is probably no fundamental reason why not several of these algorithms could have been used successfully even today, at least for some tasks.

[1] Here and in the following, the word 'cluster' refers to hadrons or calorimeter cells in the real experimental case, to partons in the theoretical perturbative calculations, and also to intermediate jets during the clustering procedure.

JHEP08(1998)001

The oldest algorithm still in use is the LUCLUS one [21], which again is based on a binary joining scheme, with additional preclustering and reassignment steps. There were two advances. One was the choice of transverse momentum as distance measure, which is better adapted to the conventional picture of non-perturbative jet fragmentation and thus allows a cleaner separation of perturbative and non-perturbative aspects of the QCD dynamics. The other was that no attempt was made to define a correct number of jets, but rather a parameter was left free, with the explicit purpose to correspond to different 'jet resolution powers'.

The JADE algorithm [22] offered a further simplification, in that only the binary joining was retained, without preclustering or reassignment. The choice of distance measure was based on invariant mass, corresponding to what was available in most $\mathcal{O}(\alpha_s^2)$ calculations of the time [23]. The JADE algorithm was therefore optimal for α_s determination studies, and came to set the standard. By contrast, it performs less well in the handling of event-by-event hadronization corrections, i.e., in the matching of jet directions and energies between the parton and hadron level [4].

Advances in the understanding of the perturbative expansion showed that soft-gluon emission does not exponentiate when ordered in invariant mass, while it does if transverse momentum is used instead [24, 25]. This gave birth to the DURHAM algorithm [25]. Alternatives such as the GENEVA one were also proposed [4].

Recently, further advances in the understanding of soft-gluon emission has lead to the introduction of new algorithms based on the DURHAM scheme, the ANGULAR-ORDERED DURHAM and the CAMBRIDGE ones [3], which modify the clustering procedure of the former in order to remedy some of its shortcomings. Given the huge increase in the computing power of modern computers, one can now reverse the historical trend towards simplification without compromising the efficiency of the algorithm, e.g., indulging in procedures more sophisticated than the simple binary joining.

The DICLUS (also called ARCLUS) algorithm [5] does not really fit into the above scheme, in that it is not based on the binary joining of two clusters to one but on the joining of three clusters to two. This is well matched to the dipole picture of cascade evolution. Like in many other algorithms, the distance measure is based on transverse momentum.

The connection between perturbative QCD cascades and jet clustering algorithms is not only limited to the DICLUS case. In general one may describe clustering algorithms as an attempt to reconstruct a QCD cascade backwards in time. In fact, when we in the following argue that one clustering should be performed *before* another it is based on experience from how to formulate QCD cascades where color coherence is correctly taken into account. Such QCD cascades have been the basis of the enormous success modern event generators have had in describing the detailed structure of e^+e^- annihilation events.

It should be noted, however, that the notion of time ordering in QCD cascades is not unambiguous. Looking at the three most successful coherent cascade implementations

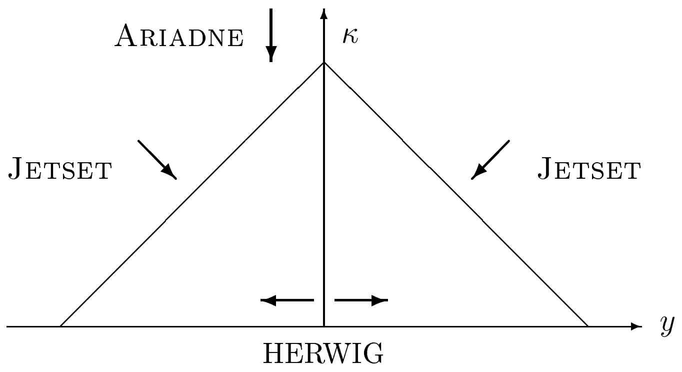

Figure 1: The ordering of emissions in ARIADNE, HERWIG, and JETSET QCD showers in the plane of κ and y (logarithm of transverse momentum and rapidity of emitted gluons).

today, they all have different ordering of emissions. HERWIG orders emissions in angle while JETSET orders in invariant mass with an additional angular constraint to ensure coherence. Finally ARIADNE orders the emissions in transverse momentum. In Fig. 1 we show the approximate phase space available for gluon emission in an e^+e^- annihilation event, in the plane of logarithm of transverse momentum (κ) and rapidity (y). The notion of time is indicated for the three programs. In the HERWIG and JETSET cases, where emissions from the q and $\bar{q}$ are treated separately, there is one direction for each, while in the ARIADNE case there is only one direction for the ordering of emissions from the $q\bar{q}$ *dipole*.

These three descriptions, although very different, are consistent with perturbative QCD and it has not been possible to say that one in better than another, although some experimental observables have been suggested [26]. Common for all programs is that they treat gluon emissions in a coherent way, and it may be easiest to look at this in terms of angular ordering. In the following we present three example diagrams of $e^+e^- \to q_1 + \bar{q}_2 + g_3 + g_4 + \ldots$. In all cases we have drawn them as one Feynman diagram, but in general all multi-gluon states are of course coherent sums of many diagrams. It is clear that a good clustering algorithm in some sense should cluster an event according to the dominating diagram for each given partonic state.

In the following subsections we give a more detailed description of several of the currently used algorithms. The order is not purely historical, but is rather intended to allow a gradual introduction of new concepts.

2.1. JADE

The JADE algorithm [22] may be viewed as the archetype of a binary joining scheme.

In this class of methods, a distance measure d_{ij} between two clusters i and j is defined as a function of their respective four-momenta, $p_{i,j} = (E_{i,j}, \mathbf{p}_{i,j})$. Since the measure is normally not Lorentz invariant, it is assumed that the analysis is performed in the hadronic rest frame of the event. To the extent this frame is not known, the lab frame is used instead, and the effects of initial-state QED radiation should then be included as a correction to the final physics results.

The algorithm starts from a list of particles, that is considered as the initial set of clusters. The two clusters with the smallest relative distance are found and then merged into one, provided their distance is below the desired minimum separation d_{cut}. The four-momentum of the new cluster k is found from its constituents i and j by simple addition, e.g., $p_k = p_i + p_j$. The joining procedure is repeated, until all pairs of clusters have a separation above d_{cut}. This final set of clusters is called jets.

In the JADE algorithm the distance measure is given by

$$d_{ij}^2 = 2E_iE_j(1 - \cos\theta_{ij}) \, , \tag{2.2}$$

where θ_{ij} is the opening angle between the momentum vectors of the two clusters. As written here, d_{ij} has dimensions of mass. The scaled expression

$$y_{ij} = \frac{d_{ij}^2}{E_{\mathrm{vis}}^2} = \frac{2E_iE_j(1 - \cos\theta_{ij})}{E_{\mathrm{vis}}^2} \tag{2.3}$$

is more often quoted. The visible energy E_{vis} would agree with the centre-of-mass (CM) energy for a perfect detector but, to the extent that some particles are lost or mismeasured, normalization to E_{vis} gives some cancellation of errors between numerator and denominator. In the following we will usually give the y-expression, but note that a translation between the two alternative forms is always possible. This also applies to the cut-off scale $y_{\mathrm{cut}} = d_{\mathrm{cut}}^2/E_{\mathrm{vis}}^2$.

Whether the dimensional or scaled dimensionless form is preferable is normally a matter of application and physics point of view. The α_{S} evolution with energy, and all other comparisons of jet rates at different energies, are best done in terms of scaled variables y. The transition between perturbative and non-perturbative physics, on the other hand, is expected to occur at some fixed dimensional scale of the order of 1 GeV. Such a hypothesis is supported, e.g., by the observable scaling violations of fragmentation functions in jets defined by a fixed y_{cut}. Therefore, we expect the 'true' partonic multiplicity of an event to increase with energy, tracing the increase of the hadronic multiplicity, while the jet rate above a given y drops, tracing the running of α_{S}.

The d_{ij} measure above is closely related to the invariant mass

$$m_{ij}^2 = (p_i + p_j)^2 = m_i^2 + m_j^2 + 2(E_iE_j - |\mathbf{p}_i||\mathbf{p}_j|\cos\theta_{ij}) \tag{2.4}$$

and the use of the correct mass as distance measure defines the so-called E variant of the JADE scheme. Given its Lorentz invariant character, mass would have been a logical choice, had it not suffered from instability problems. The reason is well understood: in general, particles tend to cluster closer in invariant mass in the region of small momenta. The clustering process therefore tends to start in the center of the event, and only subsequently spreads outwards to encompass also the fast particles. Rather than clustering slow particles around the fast ones (where the latter naïvely should best represent the jet directions) the invariant mass measure tends to cluster fast particles around the slow ones.

The d_{ij} and m_{ij} measures coincide when $m_i = m_j = 0$. For non-vanishing cluster masses d_{ij} normally drops below m_{ij}, and the difference between the two measures increases with increasing net momentum of the pair. This tends to favor clustering of fast particles somewhat, and thus makes the standard JADE algorithm more stable than the one based on true invariant mass.

There would seem to be a mismatch in comparisons between fixed-order perturbation theory based on the correct invariant mass expression [23] and experimental analyses based on the d measure. However, the perturbative results are normally presented in terms of massless outgoing partons, so the m and d measures agree on the parton level. A definition of hadronic cluster separation as if clusters were massless therefore better matches the partonic picture, and should give smaller hadronization corrections. When performing NLO perturbative calculations, it is of course then of decisive importance to impose the same kind of clustering scheme as will be used on the hadron level.

Further variants of the JADE scheme have been introduced [4]. In the p alternative, the energy of a cluster k is defined to be $E_k = |\mathbf{p}_k|$, so that the cluster is explicitly made massless, at the expense of violating energy conservation when pairing two clusters. In the E0 scheme, massless clusters are instead obtained by momentum violation, defining $\mathbf{p}_k = E_k(\mathbf{p}_i + \mathbf{p}_j)/|\mathbf{p}_i + \mathbf{p}_j|$. In this paper we stay with the standard scheme, however.

2.2. DURHAM

The DURHAM algorithm [25] can be obtained from the JADE one by a simple replacement of the distance measure from mass to transverse momentum. In scaled variables

$$y_{ij} = \frac{2\min(E_i^2, E_j^2)(1 - \cos\theta_{ij})}{E_{\mathrm{vis}}^2}\,, \tag{2.5}$$

i.e. with $E_i E_j \to \min(E_i^2, E_j^2)$. Some special features should be noted. Firstly, strictly speaking, the measure is transverse energy rather than transverse momentum, just like the JADE measure is based on energies. Secondly, the transverse momentum is defined asymmetrically, as the $p_\perp$ of the lower-energy one with respect to a reference direction given by the higher-energy one. And, thirdly, the angular dependence only agrees with that of the transverse momentum $p_\perp = E\sin\theta$ for small angles, where $\sin^2\theta \approx 2(1-\cos\theta)$. The reason for retaining the same angular dependence as in JADE is obvious enough: the correct $p_\perp$ would vanish for two back-to-back particles and thus allow unreasonable jet assignments. The DURHAM algorithm has eventually taken over the JADE rôle of standard jet finder. There are two main reasons for preferring DURHAM.

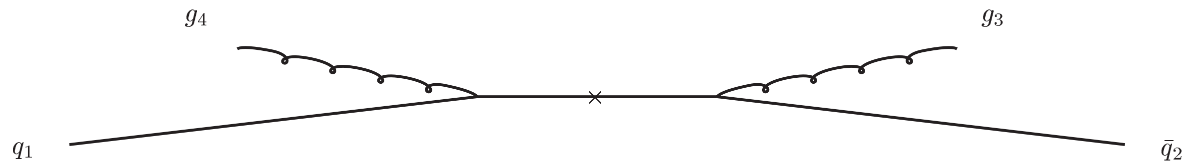

Figure 2: The seagull diagram with $E_3, E_4 \ll E_1, E_2$ and $\theta_{14}, \theta_{23} \ll \theta_{34}$.

Firstly, fixed-order perturbative corrections are quite sizeable for the JADE algorithm. This is particularly true for the case of the NLO ones to the three-jet rate $f_3(y_{\text{cut}})$ (see later on, in subsection 3.1 for its definition) [27, 28]. The importance of this aspect is evident if one considers that $f_3(y_{\text{cut}})$ provides a direct measurement of α_S. Such behaviours can be seen by noticing the large renormalization scale dependence of $f_3(y_{\text{cut}})$ at NLO, indicating that higher order corrections are not yet negligible in the perturbative expansion. Since it was (and still is) unthinkable with present computational technology to attempt the evaluation of next-to-next-to-leading order (NNLO) terms, the path to be necessarily followed in order to reduce the scale dependence of $f_3(y_{\text{cut}})$ was to define new clustering algorithms having smaller perturbative corrections. Secondly, the jet fractions obtained in the JADE scheme do not show the usual Sudakov exponentiation of multiple soft-gluon emission [24], despite having an expansion of the form $\alpha_S \ln^2 y_{\text{cut}}$ at small values of the resolution parameter.

The source of such misbehaviors at both large (i.e., in fixed-order calculations) and small (i.e., in the resummation of leading logarithms) y_{cut} values is indeed the same, namely, the large rate of soft gluons radiated in the hard scattering process and the way they are dealt with in the clustering procedure. The problem can be exemplified by referring to one of the possible configurations in which two soft gluons g_3 and g_4 can be emitted by two leading (i.e., highly energetic) back-to-back quarks q_1 and $\bar{q}_2$. Let us imagine the first gluon to be radiated in one of the two hemispheres defined by the plane transverse to the axis of the two quarks and the second one on the opposite side (i.e., the 'seagull diagram' of ref. [3]: see Fig. 2). By adopting as measure y_{ij} the expression given in eq. (2.3), the iterative algorithm would combine the two gluons with each other first, so that the net results is a 'ghost jet' in a direction along which no original parton can be found. Such behaviours end up representing a serious challenge in perturbative calculations. On the one hand, the problems encountered by fixed-order QCD in cancelling divergences are amplified by the clustering of two soft particles, so that in general one naturally expects larger higher order terms. On the other hand, such unnatural clustering induces a redistribution of the partons in the final state that spoils the exponentiation properties of large logarithms $\ln y_{\text{cut}}$ for $y_{\text{cut}} \to 0$ [29].

The simple modification [30] given in eq. (2.5) is enough to cure the two above mentioned problems. This is clear if one considers that, by adopting the DURHAM measure, in the seagull diagram configuration one of the soft gluons will always be combined first with the nearby high-energy quark, unless the angle that it forms with the other gluon is smaller than that with respect to the leading parton. As a consequence, the stability of the fixed-order results is greatly improved and the factorization of large leading and next-to-leading logarithms guaranteed.

2.3. Luclus

Historically, Luclus [21] has not been used for α_s determinations and related QCD studies, but is instead widely used for other jet topics, such as search for new particles in invariant mass distributions. In main properties it is similar to the Durham algorithm, but with several differences.

Firstly, the transverse-momentum-based distance measure is

$$y_{ij} = \frac{2|\mathbf{p}_i|^2|\mathbf{p}_j|^2(1-\cos\theta_{ij})}{(|\mathbf{p}_i|+|\mathbf{p}_j|)^2E_{\mathrm{vis}}^2} \,. \tag{2.6}$$

Geometrically, in the small-angle approximation, this can be viewed as the transverse momentum of either particle with respect to a reference direction given by the vector sum of the two momenta.

Apart from the difference between $|\mathbf{p}|$ and E, the step from the Durham to the Luclus distance measure is given by the replacement $\min(E_i,E_j) \to E_iE_j/(E_i+E_j)$. Clearly the two expressions agree when either of i or j is much softer than the other, so all the soft-gluon exponentiation properties of the Durham measure carry over to the Luclus one. In the other extreme, when $E_i = E_j$, the two y-expressions differ by a factor of 4 (that is, by a factor of 2 at $p_\perp$ level).

The usage of $|\mathbf{p}|$ rather than E in Luclus is based on non-perturbative physics considerations, specifically on the properties of string fragmentation [20]. Here, primary particles are given a Gaussian transverse momentum spectrum with respect to the string direction, typically around 400 MeV, common for all particle species. Secondary decays give a final mean $p_\perp$ that is around 100 MeV lower for pions than for Kaons or protons, but this is still smaller than the mass difference between the particles. Therefore, a jet is a set of particles with limited $p_\perp$ with respect to the common jet direction, and using $E_\perp$ only introduces unnecessary smearing. From a perturbative point of view, arguments could be raised for the use of energy (see the discussion on the Jade algorithm above). Also note that, in the string model, the $p_\perp$ width of a non-perturbative jet is independent of longitudinal momentum, to first approximation. This concept is preserved by the symmetric way in which Luclus defines $p_\perp$. The asymmetric $p_\perp$ definition of Durham is more appropriate if high-energy particles are better lined up in $p_\perp$ with the true jet axis than low-energy ones. This may occur when multi-partonic states are considered, see discussion below, so the matter is not quite clearcut.

The original Luclus routine differs from the others presented here in that it contains preclustering and reassignment steps. These options can both be switched off, individually, but the reassignment step was a part of the basic philosophy at the time the algorithm was written. The preclustering one, on the other hand, was purely intended to speed up the algorithm without affecting the final results significantly. The amount of preclustering can be varied, with much preclustering giving a faster algorithm at the expense of some residual effects of the preclustering step. Speed was an important consideration at the time the algorithm was originally formulated, but is normally no issue with modern workstations. Today users should therefore feel free to switch off this step entirely.

First consider the reassignment aspect. When two clusters are merged, some particles belonging to the new and bigger cluster may actually be closer to another cluster. A simple example is once again the seagull diagram of Fig. 2, with the quark-gluon opening angles not necessarily small. With the LUCLUS $p_\perp$ measure, it can happen (in fact, more easily than with the DURHAM one, see Fig. 8 later on) that the two soft particles are first combined to one cluster and thereafter this cluster is merged with one of the hard particles. One of the soft particles is that way combined with the hard particle it is furthest away from. The 'natural' subdivision would have been with one hard and one nearby soft particle in each final cluster. That is, a procedure that is good for going from four to three clusters and from three to two clusters may be less good for the combined operation of going from four to two clusters. The problem is that simple binary joining algorithms do not allow previous assignments to be corrected in the light of new information.

Hence the reassignment: after each joining of two clusters, each particle in the event is reassigned to its nearest cluster. For particle i, this means that the distance d_{ij} to all clusters j in the event has to be evaluated and compared. After all particles have been considered, and only then, are cluster momenta recalculated to take into account any reassignments. To save time, the assignment procedure is not iterated until a stable configuration is reached. (Again, the time cost of these iterations could be acceptable today but it was not at the time the algorithm was written.) All particles are reassigned after each binary joining step, however, and not only those of the new cluster. Therefore an iteration is effectively taking place in parallel with the cluster joining. Only at the very end, when all $d_{ij} > d_{\text{cut}}$, is the reassignment procedure iterated to convergence — still with the possibility to continue the cluster joining if some d_{ij} should drop below d_{cut} due to the reassignment.

The LUCLUS algorithm was conceived mainly based on non-perturbative considerations. The reassignment procedure is completely deterministic, however, and can therefore also be applied to any perturbative calculation, just like the simple binary joining. The price is that analytic calculations become more difficult to survey. A reassignment cannot occur after the first binary joining of an event, though, but only after the second. It therefore does not affect leading or NLO results, but only NNLO and higher orders.

In the preclustering step the original large number of particles are put together in a smaller number of clusters. This is done as follows. The particle with the highest momentum is found, and thereafter all particles within a distance $d_{ij} < d_{\text{init}}$ from it. Here it is intended that $d_{\text{init}} \ll d_{\text{cut}}$ for preclustering to give negligible effects. Together these very nearby particles are allowed to form a single cluster. For the remaining particles, not assigned to this cluster, the procedure is iterated, until all particles have been used up. Particles in the central momentum region, $|\mathbf{p}| < 2d_{\text{init}}$ are treated separately: if their vectorial momentum sum is above $2d_{\text{init}}$ they are allowed to form one cluster, otherwise they are left unassigned in the initial configuration and only appear in the first reassignment step.

The value of d_{init}, as long as reasonably small, should have no physical importance, in that the same final cluster configuration will be found as if each particle initially is assumed to be a cluster by itself. That is, the particles clustered at this step are so nearby that they almost inevitably must enter the same jet. 'Mistakes' in the preclustering can however be corrected by the reassignment procedure in later steps of the iteration. Therefore reassignment may be seen as a prerequisite and guarantee for successful preclustering.

In this respect, we would like to give a word of caution, about the actual meaning of 'reasonably small'. The value chosen for d_{init} should depend on the d_{cut}-range considered in the analysis. For example, the default value of 0.25 GeV is clearly inappropriate for $y_{\text{cut}} = d^2_{\text{cut}}/s \approx 0.0001$, as some residual effects of preclustering are then visible (see sects. 4.1 and 4.2 later on). A scaling, e.g., like $d_{\text{init}} = d_{\text{cut}}/10$ would have removed them (we have explicitly verified this in our numerical simulations). Though we recommend the mentioned scaling, should the preclustering step be retained, we have decided to keep the default value of 0.25 GeV in order to illustrate the consequences of a fixed d_{init} for small y_{cut}.

From a perturbative physics point of view, the d_{init} parameter plays a rôle very similar to that, e.g., of the y_0 parameter in the phase space slicing method of handling higher-order corrections to MEs (see, e.g., sect. 4.8 of [31]). Below y_0 the cancellation of real and virtual corrections is carried out analytically in an approximate treatment of phase space, while between y_0 and y_{cut} the addition of contributions is performed numerically with full kinematics. Hence y_0 should be picked as small as computer resources allow, and always much smaller than the physical y_{cut} parameter.

In this paper we will focus our attention on four possible options of the LUCLUS algorithm, namely the default one, and it stripped off either preclustering or reassignment or both. We will call the latter the DURHAM scheme with the LUCLUS measure (with the acronym DL) as this effectively differs from the algorithm introduced in subsection 2.2 only in the choice of the distance measure.

LUCLUS has always been distributed as part of JETSET. With the merger of JETSET and PYTHIA the routine has been renamed PYCLUS, but we will here refer to it by its original name.

2.4. Geneva

The Geneva algorithm [4] is based on pure binary joining, with a dimensionless distance measure

$$y_{ij} = \frac{8}{9}\frac{E_i E_j(1-\cos\theta_{ij})}{(E_i+E_j)^2}\,. \tag{2.7}$$

Unlike the other algorithms studied, the measure depends only on the energies of the particles to be combined, and not on the energy of the event. The energy factor $E_iE_j/(E_i+E_j)^2 \approx \min(E_i,E_j)/\max(E_i,E_j)$ favors the clustering of soft particles to the hardest ones and disfavors the combination of soft particles with each other. The soft-gluon problems of the Jade algorithm are thus avoided, indeed in a more effective way than in the Durham scheme. In fact, a soft gluon will only be combined with another soft gluon if the angle between them is *much* smaller than the angle between the former and the nearby high-energy particle. As a consequence, it turns out that the Geneva scheme exhibits a more reduced scale dependence as compared to the Durham algorithm in the three-jet rate at NLO [4]. Indeed, in ref. [32] it was shown that such property remains true also in the case of the NLO four-jet rates $f_4(y_{\rm cut})$. Furthermore, it has been pointed out [32] that the Geneva algorithm is particularly sensitive to the number of light flavors, this rendering it most suitable for the study of New Physics effects. For example, in the context of four-jet events in e^+e^- annihilations [32], where the existence of the so-called 'light gluino' events has been advocated in the past years [33]. As for the exponentiation properties of large logarithms at small values of the resolution parameter, these have not thoroughly been studied yet for this scheme. However, a simple example[2], should help understanding that the Geneva algorithm can manifest severe misassignment problems in the soft regime. It suffices to consider a $q_1\bar{q}_2g_3g_4$ configuration (e.g., with the antiquark and the two gluons in the same hemisphere) with the two gluons produced via a triple-gluon vertex and ordered in energy, such that $E_2 \gg E_3 \gg E_4$. In the region where $\theta_{34} < \theta_{23}$ the gluon g_4 can be assigned by the Geneva algorithm to the antiquark $\bar{q}_2$ rather than to other gluon g_3, since $(E_2+E_4)^2 \gg (E_3+E_4)^2$ in the denominator of eq. (2.7). Hence, one expects the C_F factor instead of the correct C_A one to describe the radiation intensity of such an event. This induces a breakdown of the correct exponentiation picture [24, 25]: see eqs. (3.12)–(3.14) in subsection 3.2. By contrast, a transverse momentum measure, such as the Durham (2.5) or Luclus (2.6) ones, would have more naturally assigned g_4 to g_3, as in the limit $\theta_{34} \ll \theta_{23}$ the full infrared (i.e., both soft and collinear, driven by the gluon propagator) singularity sets on, which renders the triple-gluon diagram dominant in the kinematics above.

The Geneva algorithm has had little phenomenological impact so far. One reason is that it is rather sensitive to hadronization effects, as already pointed out in ref. [4]. For instance, compare a single original large-energy hadron with a system of hadrons of

[2]For which we are indebted to Yuri Dokshitzer.

the same total energy and collinear within the hadronization $p_\perp$ spread of a few hundred MeV. Then the former can collect particles further away in angle at the early stages of clustering. Therefore clustering of gluon jets, which start out with a lower energy and tend to fragment softer than quark jets, is disfavored. This will introduce systematic biases, e.g., in jet energy distributions, that can only be unfolded given a detailed understanding of the hadronization process. In addition, the GENEVA algorithm is more sensitive to measurement errors since the measure contains the energy of individual particles also in the denominator, where other algorithms have instead the total visible energy, which is more precisely measured.

2.5. Angular-Ordered Durham

This algorithm maintains the same measure (2.5) of the DURHAM scheme, while modifying the clustering procedure. It was introduced in ref. [3] to obviate one of the flaws of that algorithm: namely, its tendency of giving counter-intuitive jets, in ref. [3] referred to as 'junk-jets', at small values of the resolution parameter.

The problem is as follows. Let us imagine the configuration in Fig. 3, with two back-to-back high-energy partons (the quarks q_1 and $\bar{q}_2$) plus a double gluon emission (g_3 and g_4) [3] in the same hemisphere defined by the plane transverse to the direction dictated by the two quarks, one of the gluons being at large angle and soft (g_4) and the other (g_3) collinear to the nearby leading particle on the same side ($\bar{q}_2$). Then, according to the clustering procedure adopted in the DURHAM scheme, one usually starts from the softest particle (i.e., one of the two gluons: here g_4) and merges this with the nearest in angle, to minimize the $p_\perp$-measure. Thus, such a particle gets clustered not with one of the leading partons (i.e., $\bar{q}_2$ here) but, typically, with the softest among the particles which happen to lie on the same side (i.e., to the other gluon, g_3, in our example). This is contrary to our picture of the large-angle g_4 being emitted coherently by the $\bar{q}_2$ and g_3, so that most of the recoil to the g_4 should have been taken by the more energetic $\bar{q}_2$. Such a procedure gets iterated in the case in which more particles are involved (e.g., radiated in between the least, g_4, and the most energetic, $\bar{q}_2$, ones in one hemisphere). Since at each stage the new pseudo-particle acquires more and more four-momentum, in the end the latter will have a $p_\perp$ relative to the leading particle in the same hemisphere larger than the resolution scale adopted. This way, a third jet is eventually resolved. A good algorithm should then be designed

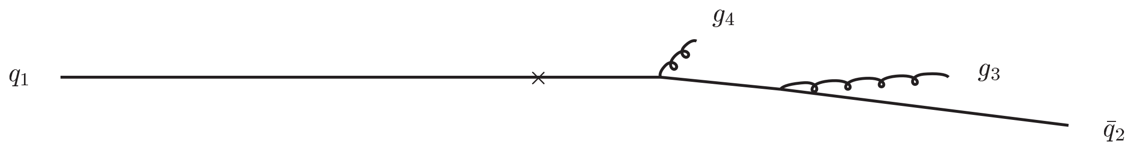

Figure 3: Parton branching with 'unresolved', soft, large-angle gluon emission g_4. Here, one has the following configuration: $E_2 \gg E_3 \gg E_4$, and $\theta_{23} \ll \theta_{24} \approx \theta_{34}$.

so that the starting configuration remains classified as a two-jet final state down to the smallest possible values of the resolution $y_{\rm cut}$, at which the third (junk-)jet separates. However, one should notice that if one hemisphere of an event is significantly broadened by multiple soft-gluon emission, where the gluons together carry away non-negligible energy and $p_\perp$, it would be reasonable to argue that the event could be legitimately recognised as a three-jet one. Clearly, in this as in many other cases, it is the status of our present computation technology and of its list of priorities that induces the choice of strategy to be adopted. In fact, the latter needs to be neither unique nor definitive. On the other hand, as well demonstrated in ref. [34] (see, e.g., Fig. 1 there) and as we shall further see below, the remedy adopted by the ANGULAR-ORDERED DURHAM (and CAMBRIDGE, too) scheme in order to alleviate the above mechanism appears more than adequate for present investigations.

In ref. [3] it was shown that a simple modification of the DURHAM algorithm suffices to delay the onset of junk-jet formation, which results mainly from a non-optimal sequence of clustering, rather than from a poor definition of the test variable (as was the case for the JADE algorithm in the seagull diagram). The key to reduce the severity of the problem resides in distinguishing between the variable $v_{ij} \equiv 2(1-\cos\theta_{ij})$, used to decide which pair of objects to test first, and the variable y_{ij} to be compared with the resolution parameter $y_{\rm cut}$. The algorithm then operates as follows. One considers first the pair of objects (ij) with the smallest value of v_{ij} (in ref. [3], this procedure was referred to as 'angular-ordering'). If $y_{ij} < y_{\rm cut}$, they are combined. Otherwise the pair with the next smallest value of v_{ij} is considered, and so on until either a $y_{ij} < y_{\rm cut}$ is found or, if not, clustering has finished and all remaining objects are defined as jets.

Coming back to the example configuration described before, but with the new clustering procedure, one should expect the collinear quark $\bar{q}_2$ and gluon g_3 to be paired first, with the soft, large-angle gluon g_4 eventually joining the new cluster. In case more radiation is present around the leading quark, the procedure always iterates so that the pairing always starts amongst the particles collinear to the leading quark $\bar{q}_2$, with the soft, large-angle gluon g_4 entering the clustering procedure only at the very end. Indeed, this way, the original configuration will more likely be recognised as a two-jet one in the ANGULAR-ORDERED DURHAM than in the original DURHAM. We will exemplify this in subsection 4.2.

By generalizing the procedure to the full hard scattering process, one indeed realizes that, at a given $y_{\rm cut}$, the two-jet fraction at NLO as given by the ANGULAR-ORDERED DURHAM is larger than in the original DURHAM scheme, as illustrated in ref. [3]. Conversely, the three-jet rate at the same order is smaller. Thus, since it is not unreasonable to argue that jet algorithms having smaller NLO terms may also have smaller higher-order corrections, one would imagine the scale dependence of the three-jet rate for the ANGULAR-ORDERED DURHAM to be reduced as compared to that of the DURHAM algorithm. This was shown explicitly again in ref. [3] (see also subsection 3.1 later on). Given the phenomenological relevance of $f_3(y_{\rm cut})$, this should represent an improvement from the point of view of the accuracy achievable, e.g., in α_S determinations, given that the theoretical error should diminish accordingly. As for the exponentiation properties at small $y_{\rm cut}$, these remain unspoilt in the new algorithm, as discussed in ref. [3].

Closing this section, we should remind the reader that the ANGULAR-ORDERED DURHAM should be intended as an intermediate step between the DURHAM and CAMBRIDGE schemes, rather than a new one. Indeed, we will recall in the next section another shortcoming of the DURHAM algorithm which carries over into the ANGULAR-ORDERED DURHAM one and which can have a strong impact in jet-rate studies. Nonetheless, for purposes of comparison, we will present results for the ANGULAR-ORDERED DURHAM on the same footing as the other algorithms.

2.6. CAMBRIDGE

The CAMBRIDGE algorithm was defined and its properties discussed in ref. [3]. It implements the same distance measure as the DURHAM scheme, while further modifying the clustering procedure of the ANGULAR-ORDERED DURHAM one. As a matter of fact, the sole introduction of v_{ij} is not enough to remedy the problem of 'misclustering' of the DURHAM algorithm, that is, the tendency of soft 'resolved' particles of attracting wide-angle radiation [3].

Let us imagine that the soft, large-angle gluon $i = 4$ in the previous example has been eventually resolved at a certain (possibly low) scale $y_{\rm cut}$ (i.e., $y_{4j} > y_{\rm cut}$). Clearly, the very same ability that it had of attracting radiation (because of its softness) as unresolved parton (i.e., when $y_{4j} < y_{\rm cut}$) survives above the new $y_{\rm cut}$. In particular, if further wide-angle (with respect to the leading quark in the same hemisphere of the resolved gluon: i.e., $\bar{q}_2$) radiation occurs, say, two additional gluons g_5 and g_6 one of which (g_5) happens to lie in angle a little closer to g_4 than to the other, then such a gluon will be erroneously assigned (to g_4 rather than to g_6) when $\theta_{45} > \theta_{4\hat{2}\hat{3}}$ (thus assuming that $\bar{q}_2$ and g_3 have already been clustered).

In order to cure this problem, the CAMBRIDGE scheme implements the sterilization (i.e., the removal from the table of particles participating in the clustering) of the softest particle in a resolved pair, a procedure called 'soft-freezing'. In our example, once g_4 has been removed from the sequence of clustering (when $y_{4\hat{2}\hat{3}} > y_{\rm cut}$) then the unwanted pairing of g_4 and g_5 (yielding $y_{45} < y_{\rm cut}$ if the gluons are soft enough) is successfully prevented.

Note, however, that the diagram in Fig. 4 is not the only important one for the described parton configuration. For example, we should also consider the diagram where g_5 is attached to $\bar{q}_2$ (to the left of g_4) in which case the freeze-out of g_4 could prevent g_5 from being correctly clustered. The relative importance of the two, as well as of the others appearing at the same order, is clearly a function of the dynamics of the final state. Numerical results will finally establish the effectiveness and/or the limitations of the above approach.

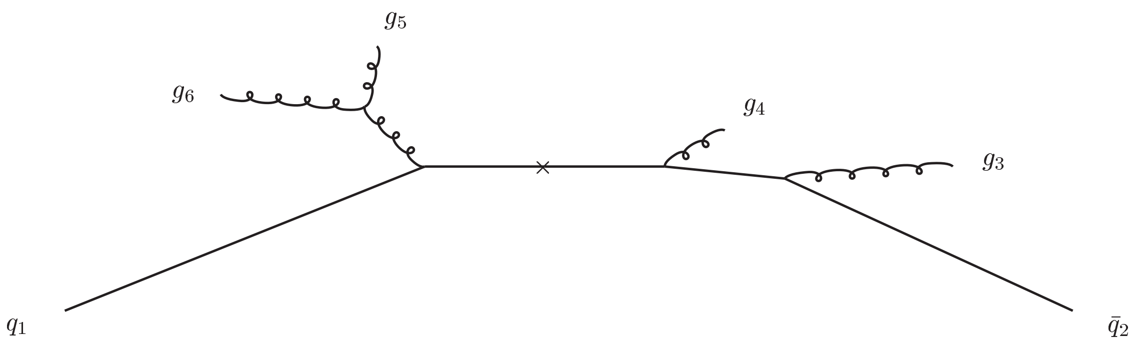

Figure 4: Parton branching with 'resolved', soft, large-angle gluon emission g_4. In this case, $\theta_{45} > \theta_{4\hat{23}}$, where $\hat{23}$ symbolizes the cluster formed by the merging of $\bar{q}_2$ and g_3.

As a matter of fact, the mentioned misclustering phenomenon could well manifest itself in studies of high multi-jet rates (as in Fig. 4) as well as in those of the internal jet sub-structure when examining the history of the mergings (e.g., in the 'would-be' cluster $\hat{45}$, artificially over-populated with gluons). In particular, it was shown in ref. [3], that such an additional step plugged onto the ANGULAR-ORDERED DURHAM algorithm allows one to increase the final event multiplicity (e.g., that of the NLO $f_3(y_{\rm cut})$ rate) while not deteriorating the scale dependence of the results. In addition, as for the ANGULAR-ORDERED DURHAM scheme, the properties of factorization/exponentiation of large $\ln y_{\rm cut}$ terms at small $y_{\rm cut}$'s remain completely unaffected.

Altogether, as shown in ref. [3], the CAMBRIDGE scheme came at the time of those studies to represent the most suitable choice when dealing with phenomenological studies involving infrared (i.e., soft and collinear) configurations of hadrons/partons, also in view of its performances with respect to the size of the hadronization corrections (see ref. [3] and sect. 4). In our forthcoming studies we will allow for a variation of the basic CAMBRIDGE scheme. Namely, we will also adopt the LUCLUS measure along with the DURHAM one, and we will label the corresponding algorithm as CL (with the same clustering as the original CAMBRIDGE, though).

2.7. DICLUS

The DICLUS algorithm is very different from the other ones considered in this paper, in that each step clusters three jets into two, rather than two into one. Although unconventional in this respect, it is not unnatural. If one considers, e.g., the Lund string fragmentation model, a hadron is produced in the color field between two partons rather than stemming from one individual parton. Especially soft hadrons between jets can never unambiguously be assigned to one jet. Also on the perturbative level, gluons are emitted coherently by neighboring partons. The partonic cascade can therefore be formulated in terms of color-dipole radiation of gluons from pairs of color-connected partons, as in the ARIADNE program [35]. In a conventional parton cascade, this coherence is instead formulated in terms of angular ordering.

Just as a conventional binary clustering algorithm can be viewed as an attempt to reconstruct backwards a parton shower step by step, the DICLUS algorithm tries to reconstruct a dipole cascade.[3] The ordering variable in ARIADNE is a Lorentz-invariant transverse momentum measure defined for an emitted parton i with respect to the two emitting partons j and k as

$$p^2_{\perp i(jk)} = \frac{(s_{ji} - (m_i + m_j)^2)(s_{ik} - (m_i + m_k)^2)}{s_{ijk}}\,, \tag{2.8}$$

where $s_{ij}(s_{ijk})$ is the squared invariant mass of two (three) partons. When reconstructing the dipole cascade backwards in time with DICLUS, the same measure is used and the clustering procedure is as follows.

- For each cluster i, find the two other clusters j and k for which $p^2_{\perp i(jk)}$ is minimized.

- Take the combination i, j, k which gives the minimum $p^2_{\perp i(jk)}$, and if this is below a cutoff, remove cluster i and distribute its energy and momentum among j and k.

These steps are repeated until no $p^2_{\perp i(jk)}$ is below cutoff.

The joining is performed in the rest-frame of the three clusters, which are replaced by two massless, back-to-back ones aligned with the one of the original clusters with the largest energy (`mode=1`). Alternatively, the new clusters are placed in the plane of the original ones with an angle $\psi = \frac{E_k^2}{E_j^2+E_k^2}(\pi - \theta_{jk})$ from the highest energy cluster j (k is the second highest energy cluster and θ_{jk} is the angle between j and k) (`mode=0`). These two options correspond exactly to the two ways of distributing transverse recoil in an emission in ARIADNE. Mode 1 is more similar to the binary algorithms and is the one mostly used in this paper.

In fact, for most cases, the measure in eq. (2.8) is closer to the transverse mass of cluster 2 rather than its transverse momentum, except when two clusters are almost at rest w.r.t. each other. Removing the subtraction of masses in eq. (2.8) gives a measure which is closer to transverse mass everywhere:

$$m^2_{\perp i(jk)} = \frac{s_{ji}s_{ik}}{s_{ijk}}\,. \tag{2.9}$$

This measure is used in `mode=2`, which is otherwise the same as `mode=1`. Since the jets in DICLUS are always massless, the only difference will be in the initial clustering of massive hadrons, but it turns out that this actually makes some difference in the reconstruction of jet energies and angles below.

[3]This similarity with the ARIADNE cascade is the reason the algorithm was originally called ARCLUS.

There is no straight-forward translation between the distance measure in DICLUS and the ones used in binary clustering algorithms, since it depends on three clusters rather than two. However, in the limit of small measures, the $p_{\perp i(jk)}$ is equal to the LUCLUS measure taken between the two softest clusters in the rest frame of the clusters i, j, k. Also, if the cluster i is much softer than j and k, and much closer to, e.g., j than k, $p_{\perp i(jk)}$ is again equal to the LUCLUS d_{ij}.

To better understand how DICLUS works we look at the example diagrams above. In Fig. 2, the first step would be to cluster g_3 into $\bar{q}_2$ and g_4 (or g_4 into q_1 and g_3). In most cases the anti-quark would be given the major part or the gluons transverse momentum, thus DICLUS resembles the other transverse momentum based algorithms for this case. However, in some cases the neighboring gluon will get a large fraction of the transverse momentum, especially if the invariant mass of g_3 and g_4 is smaller than that of g_3 and $\bar{q}_2$. This may happen even if the angle between g_3 and g_4 is larger than that between g_3 and $\bar{q}_2$, so the problem present in algorithms based on invariant mass measures is not solved completely.

In Fig. 3, assuming g_4 has smaller $p_\perp$ than g_3, the first step would be to cluster g_4 into q_1 and g_3, giving extra $p_\perp$ to g_3 possibly pushing it above the cutoff. This is a good description of how this parton configuration would have been generated in a dipole cascade. However, parton cascades in general only agrees completely with perturbative QCD in the limit of strong ordering of emissions where recoils do not matter, but, as we have discussed above, it makes some difference here and it would certainly be more reasonable to say that g_4 was radiated by g_3 and $\bar{q}_2$ coherently. Finally in Fig. 4, assuming now that g_3 and $\bar{q}_2$ have already been joined, DICLUS could very well cluster g_5 into g_6 and g_4, which is how it would have been produced in a dipole cascade, although if g_5 is soft, it would be more reasonable to have it produced from a dipole between g_6 and the $(g_4 g_3 \bar{q}_2)$ system, where the latter acts coherently as one color charge.

Since DICLUS clusters three particles into two, it is not directly possible to say which final state hadron belongs to which jet. It is, however, possible to assign each particle to a jet after the jet directions have been found, simply by finding the two jets j and k for each particle i for which $p^2_{\perp i(jk)}$ is smallest and then assigning particle i to the jet which is closest in angle in the rest frame of ijk. In this way it is also possible to redefine the jet directions and energies by summing the momentum of the particles assigned to each jet. This reclustering is used below and is then labelled 'reclustered'.

2.8. Summary of algorithms

In this section we simply summarize, in Tab. 1 on the next page, for the convenience of the reader, the algorithms (and their variants) that we will use throughout the paper, their abbreviations which will appear in the text and the figures as we all as some of their features.

Algorithm		resolution scale	comment
JADE	J	$2E_iE_j(1-\cos\theta_{ij})$	
DURHAM	D	$2\min(E_i^2,E_j^2)(1-\cos\theta_{ij})$	
DURHAM (LUCLUS $k_\perp$) (Also DURHAM/Lu)	DL	$\frac{2\lvert\mathbf{p}_i\rvert^2\lvert\mathbf{p}_j\rvert^2(1-\cos\theta_{ij})}{(\lvert\mathbf{p}_i\rvert+\lvert\mathbf{p}_j\rvert)^2}$	LUCLUS without reassignment and preclustering
LUCLUS	L	$\frac{2\lvert\mathbf{p}_i\rvert^2\lvert\mathbf{p}_j\rvert^2(1-\cos\theta_{ij})}{(\lvert\mathbf{p}_i\rvert+\lvert\mathbf{p}_j\rvert)^2}$	
GENEVA	G	$\frac{8}{9}E_{\mathrm{vis}}^2\frac{E_iE_j(1-\cos\theta_{ij})}{(E_i+E_j)^2}$	
ANGULAR-ORDERED DURHAM	A	$2\min(E_i^2,E_j^2)(1-\cos\theta_{ij})$	CAMBRIDGE without soft-freezing
CAMBRIDGE	C	$2\min(E_i^2,E_j^2)(1-\cos\theta_{ij})$	
CAMBRIDGE (LUCLUS $k_\perp$)	CL	$\frac{2\lvert\mathbf{p}_i\rvert^2\lvert\mathbf{p}_j\rvert^2(1-\cos\theta_{ij})}{(\lvert\mathbf{p}_i\rvert+\lvert\mathbf{p}_j\rvert)^2}$	
DICLUS mode 0	Di0	$\frac{(s_{ji}-(m_i+m_j)^2)(s_{ik}-(m_i+m_k)^2)}{s_{ijk}}$	$3\to 2$ clustering
DICLUS mode 1	Di1	$\frac{(s_{ji}-(m_i+m_j)^2)(s_{ik}-(m_i+m_k)^2)}{s_{ijk}}$	largest initial cluster retains its direction
DICLUS mode 2	Di2	$\frac{s_{ji}s_{ik}}{s_{ijk}}$	largest initial cluster retains its direction

Table 1: Summary of jet algorithms used in this paper. LUCLUS is sometimes used without preclustering or without reassignment as indicated by the labeling. Similarly, DICLUS is, where so indicated, used with reclustering.

3. Perturbative comparisons

In this section we will compare the performances of the jet clustering algorithms introduced in sect. 1 with respect to several quantities calculable in perturbative QCD which are relevant to hadronic studies in electron-positron annihilations. It is subdivided in two subsections. In the first we deal with fixed-order results whereas in the second we present resummed perturbative quantities. The treatment in subsection 3.1 is mainly numerical, whereas in subsection 3.2 is analytical.

To produce the results in the first case, we have made use of the 'QCD parton generator' EERAD [36].[4] Such programs calculate NLO corrections to arbitrary infrared-safe

[4] An up-to-date list and a description of similar codes publicly available can be found in ref. [37].

two- and three-jet quantities, through the order $\mathcal{O}(\alpha_{\rm S}^2)$ in QCD perturbation theory. Although they resort to Monte Carlo (MC) multidimensional integration techniques, they differ substantially from the QCD-based 'Monte Carlo event generators' that we will introduce later on (in sect. 4). For a start, the former compute the exact $\mathcal{O}(\alpha_{\rm S}^2)$ ME result, rather than implementing only the infrared QCD dynamics in the usual $\mathcal{O}(\alpha_s)$ ME + Parton Shower (PS) modeling [38]. In addition, the phase-space configurations generated are not necessarily positive definite, so that a probabilistic interpretation is not possible. Finally, these programs only consider partonic states and no treatment of the hadronization process is given. This kind of generators thus represents a complementary tool for QCD analyses to the phenomenological MCs which will be described and used in sect. 4.

One final remark, before we start our investigations in pQCD. That is, although we look here at some individual properties, we remind the reader that when choosing an algorithm for a particular measurement, one may have to consider many different aspects altogether. When, e.g., measuring $\alpha_{\rm S}$ from the three-jet rate, it is not enough to find the algorithm with smallest scale dependence, especially if this behavior is found at a larger resolution scale where the three-jet rate is lower and thus giving a larger statistical error in the measurement. The goal must be to minimize the total error which may include both the statistical error as well as systematical errors due to detector unfolding, hadronization corrections, scale dependencies, etc. This is however well beyond the scope of our study.

3.1. Fixed-order perturbative results

In this subsection we study the y-dependent three-jet fraction[5] $f_3(y)$, defined through the relation

$$f_3(y) = \left(\frac{\alpha_{\rm S}}{2\pi}\right) A(y) + \left(\frac{\alpha_{\rm S}}{2\pi}\right)^2 (B(y) - 2A(y)) + \cdots\,, \tag{3.1}$$

having implicitly assumed the choice $\mu = Q$ of the renormalization scale (in the $\overline{\rm MS}$ scheme). In eq. (3.1), $\alpha_{\rm S}$ represents the strong coupling constant whereas $A(y)$ and $B(y)$ are the so-called leading and next-to-leading 'coefficient functions' of the three-jet rate, respectively. The terms of order $\mathcal{O}\left(\alpha_{\rm S}^2\right)$ involving $A(y)$ take account of the normalization to $\sigma_{\rm tot}$ rather than to σ_0, which we assume throughout the paper. In fact, we define the n-jet fraction $f_n(y)$ as

$$f_n(y) = \frac{\sigma_n(y)}{\sum_m \sigma_m(y)} = \frac{\sigma_n(y)}{\sigma_{\rm tot}(y)}\,, \tag{3.2}$$

[5]Here and in the following subsection, in order to simplify the notation, we shall use y to represent $y_{\rm cut}$ and refer to the various jet clustering algorithms/schemes by using their initials only. In addition, we acknowledge our abuse in referring to the latter both as algorithms and as schemes, since the last term was originally intended to identify the composition law of four-momenta when pairing two clusters (see subsection 2.1). This is in fact a well admitted habit which we believe will not generate confusion in our discussion.

where $\sigma_n(y)$ is the n-jet production cross section at a given y. If $\sigma_{\rm tot}$ identifies the *total* hadronic cross section $\sigma_{\rm tot} = \sigma_0(1 + \alpha_{\rm S}/\pi + \ldots)$, σ_0 being the lowest-order Born one, then the constraint $\sum_n f_n(y) = 1$ applies. For $n = 3$, eq. (3.1) represents the three-jet fraction in NLO approximation in perturbative-QCD (pQCD).

Out of the thirteen jet clustering algorithms that we originally chose for our study, we focus here our attention on the D, A, C, DL, CL, Di0 and Di1 schemes. We neglect considering the others for the following reasons. On the one hand, the J and G schemes have already been documented extensively in the specialized literature [4, 27] and, on the other hand, we would expect them to have little phenomenological applications in the future, at least in QCD studies in the infrared dominion. In fact, as already recalled, the former does not allow for factorization properties of large logarithms $\ln y$ at small y-values whereas for the latter these have not been proven to hold yet. Indeed, they both share the feature of being based on an obsolete invariant mass measure, whose flaws go beyond the realm of perturbative QCD, as it is reflected by the more fundamental rôle played by the transverse momentum in setting the scale of jet evolution, as the argument of the running coupling, and in defining the boundary between perturbative and non-perturbative physics [38]. Furthermore, of the four possible options of the Luclus scheme introduced previously, we only consider here the simplest one (which we labelled DL) which implements neither the preclustering nor the reassignment steps. Anyhow, because of the kinematic simplicity of the partonic final states entering in the NLO three-jet rates, the differences among the four implementations turn out to be very marginal. Firstly, the reassignment option is inactive in three-jet rates until the NNLO, see subsection 2.3. Secondly, the preclustering procedure can be incorporated easily with imperceptible effects, so that the cancellations between the loop- and the bremsstrahlung-diagrams still take place effectively without deteriorating the accuracy of the results. (If this is not the case, the $d_{\rm init}$ parameter has not been set appropriately, see subsection 2.3.) Thus, the claim made in the literature, that the Luclus algorithm is not suitable for perturbative calculations (see, e.g., ref. [4]) does not apply in the present context: i.e., in *numerical* computations of NLO three-jet observables. Also, although analytical calculations with the original Luclus scheme may be prohibitively difficult, one can certainly say that, without preclustering and reassignment (here, DL scheme) Luclus remains a reasonable option to adopt, especially in view of some of the results that we will present in the following. Besides, its properties with respect to the Sudakov exponentiation of soft-gluon emission in the resummation procedure of large $\ln y$ logarithms are on the same footing as the D, A and C schemes [3] (see next section). For Diclus we note that the measures in eqs. (2.8) and (2.9) are equivalent for massless partons, so that in the following Di2 coincides with Di1 as they have the same recoil assignment (see subsection 2.7).

Figure 5 shows the $A(y)$ function, over the range $0.001 \leq y \lesssim 0.1$, for the selected algorithms. Notice that in several cases the curves coincide. In particular, it occurs for the D, A and C schemes, the DL and CL algorithms and the Di0 and Di1 ones,

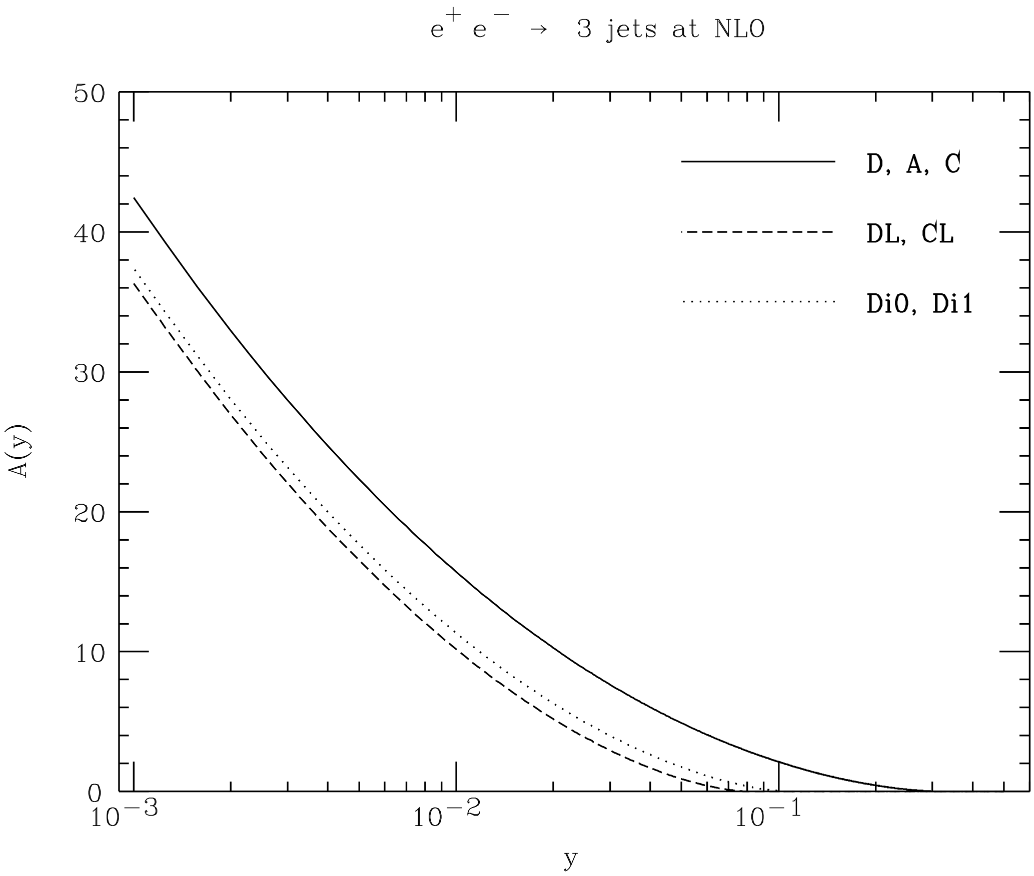

Figure 5: The parton level $A(y)$ function entering in the three-jet fraction (3.1) at LO and NLO in the D, A and C schemes (solid line), DL and CL schemes (dashed line), Di0 and Di1 schemes (dotted line).

respectively. This is evident if one considers that (apart from the C and CL options) the various schemes, within each of the three subsets, differ only in the clustering procedure of unresolved particles, which clearly does not affect the LO three-jet rates. As for the C and CL algorithms, one should consider that, for $n = 3$ partons, kinematical constraints impose that, on the one hand, the two closest particles are also those for which y is minimal and, on the other hand, the identification of the softest of the three partons as a jet implies that the remaining two particles are naturally the most energetic and far apart. This ultimately means that the $A(y)$ function is the same also for the schemes implementing the soft-freezing step.

The pattern of the curves in Fig. 5 is easily understood in terms of the measures of the algorithms. For given values of angle and energies (or three-momenta) of the parton pair (ij), θ_{ij} and $E_{i,j}$ (or $|\mathbf{p}_{i,j}|$) respectively, the value of y_{ij} is generally larger in the D, A and C schemes, as compared to the DL and CL ones: see eqs. (2.5) and (2.6). Therefore, over an identical portion of phase space, more three-parton events will be accepted as three-jet ones in the D, A and C algorithms than in the DL and CL ones (see also Fig. 8 later on) this ultimately increasing the value of $A(y)$. The comparison

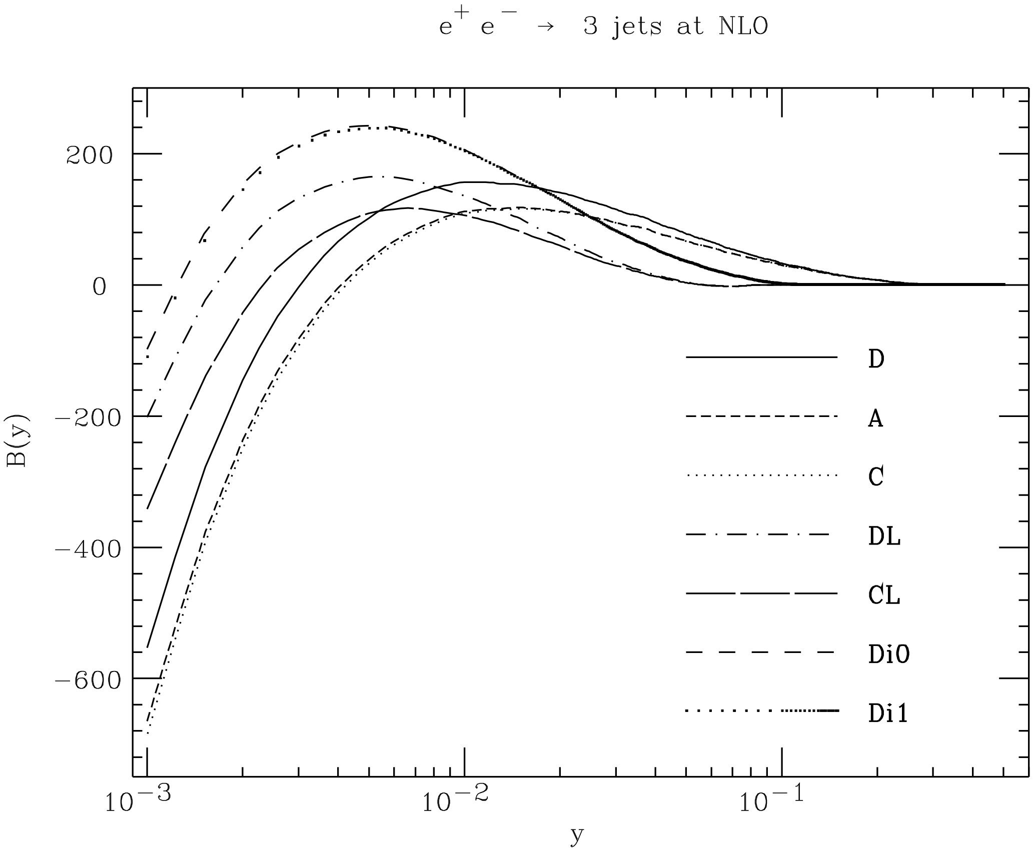

Figure 6: The parton level $B(y)$ function entering in the three-jet fraction (3.1) at NLO in the D scheme (solid line), A scheme (short-dashed line), C scheme (dotted line), DL scheme (dot-dashed line), CL scheme (long-dashed line), Di0 scheme (dashed line) and Di1 scheme (fine-dotted line).

of the two measures (2.5) and (2.6) to that of the DICLUS schemes (2.8) is clearly less straightforward, as already discussed. For a fixed y, the latter is on average larger than that of the DL measure but smaller than that of the D one, as can be deduced from Fig. 5.

In Fig. 6 we present the NLO $B(y)$ function. Because of the different recombination procedures of the schemes, the various curves now all separate. The interplay between the D, A and C rates has been analysed in great detail in ref. [3], so we do not repeat those discussions in this paper. We only spot here the correspondence existing between, on the one hand, the D and C algorithms and, on the other hand, the DL and CL ones. In the sense that, the former two differ from each other in the same way as the latter two do: i.e., in the angular-ordering and soft-freezing procedure recommended in ref. [3]. Indeed, the relative behaviors (of D vs. C and of DL vs. CL) are qualitatively similar. Further notice the tendency of the DL, CL, Di0 and Di1 schemes of yielding at small y's a $B(y)$ consistently higher than that due to the other three algorithms, and vice versa at large values of the resolution parameter. In other terms, they emphasize the real four-parton component with one unresolved pair more than the virtual three-parton at small y-values, and vice versa as y increases. However, we remind that the peaking of $B(y)$ at different y values in itself does not have to say much, since the definition of y is not the same. The simplest measure of the difference in cancellations between real and virtual contributions is instead the maximum value of $B(y)$.

We are now going to carefully investigate the interplay of the $A(y)$ and $B(y)$ functions in the expression of the three-jet fraction since, as recalled in ref. [3], from the point of view of perturbative studies, a 'good' jet clustering algorithm should allow for a reduced μ-scale dependence of the fixed-order results, where μ is the subtraction (energy) scale regulating the infrared cancellations. As a matter of fact, the $\mathcal{O}(\alpha_S^3)$ corrections are guaranteed to cancel the μ-dependence of the $\mathcal{O}(\alpha_S^2)$ three-jet fraction up to the order $\mathcal{O}(\alpha_S^4)$, so that the smaller the variations with μ the lower the higher order corrections are expected to be. The μ-scale dependence is introduced in eq. (3.1) by means of the two substitutions

$$\alpha_S \to \alpha_S(\mu), \qquad\qquad B(y) \to B(y) - A(y)\beta_0 \ln(\frac{Q}{\mu}), \tag{3.3}$$

where $\beta_0 = 11 - 2N_f/3$ is the first coefficient of the QCD β-function and N_f is the number of fermionic (colored) flavors active at the energy scale μ.

A problem arises when studying the scale dependence of $f_3(y)$ for algorithms based on different measures, as for the same y the three-jet fraction at NLO can be significantly different. A more consistent procedure was outlined in ref. [4]: that is, to compare the NLO scale dependence of the various schemes not at the same y-value, rather at the same $A(y)$, the three-jet fraction at LO. As can be viewed from Fig. 6, two possible combinations of y's are the following; $y_{\mathrm{D,A,C(DL,CL)[Di0,Di1]}} = 0.010(0.005)[0.006]$ and $y_{\mathrm{D,A,C(DL,CL)[Di0,Di1]}} = 0.050(0.021)[0.025]$. Such values are typically in the three-jet region and, in addition, they are rather large, as compared to the minimum $y = 0.001$ considered so far, so that they can guarantee the full applicability of the perturbation theory.

Figure 7 shows (again for the seven selected algorithms) the value of $f_3(y)$ plotted against the adimensional scale parameter $\mu/Q \equiv \mu/M_Z$, over the range between 1/10 and 2, for the two mentioned combinations of the jet resolution parameter. Note that for the strong coupling constant we have used the two-loop expression, with $N_f = 5$ and $\Lambda^{(5)}_{\overline{\mathrm{MS}}} = 250$ MeV, yielding $\alpha_S(M_Z) = 0.120$, with $Q = M_Z$ as CM energy at LEP1.

Although the structure of the QCD perturbative expansion does not prescribe which value should be adopted for the scale μ, an obvious requirement is that it should be of the order of the energy scale involved in the problem: i.e., the CM energy Q (see ref. [39] for detailed discussions). Indeed, this choice prevents the appearance of large terms of the form $(\alpha_S \ln(\mu/Q))^n$ in the QCD perturbative series. Furthermore, the physical scale of gluon emissions that actually give rise to three-jet configurations are to be found down to the energy scale $\sqrt{y}M_Z$, not above M_Z. In other terms, one should avoid building up large logarithmic terms related to the mismatch between $\mu \geq M_Z$ and

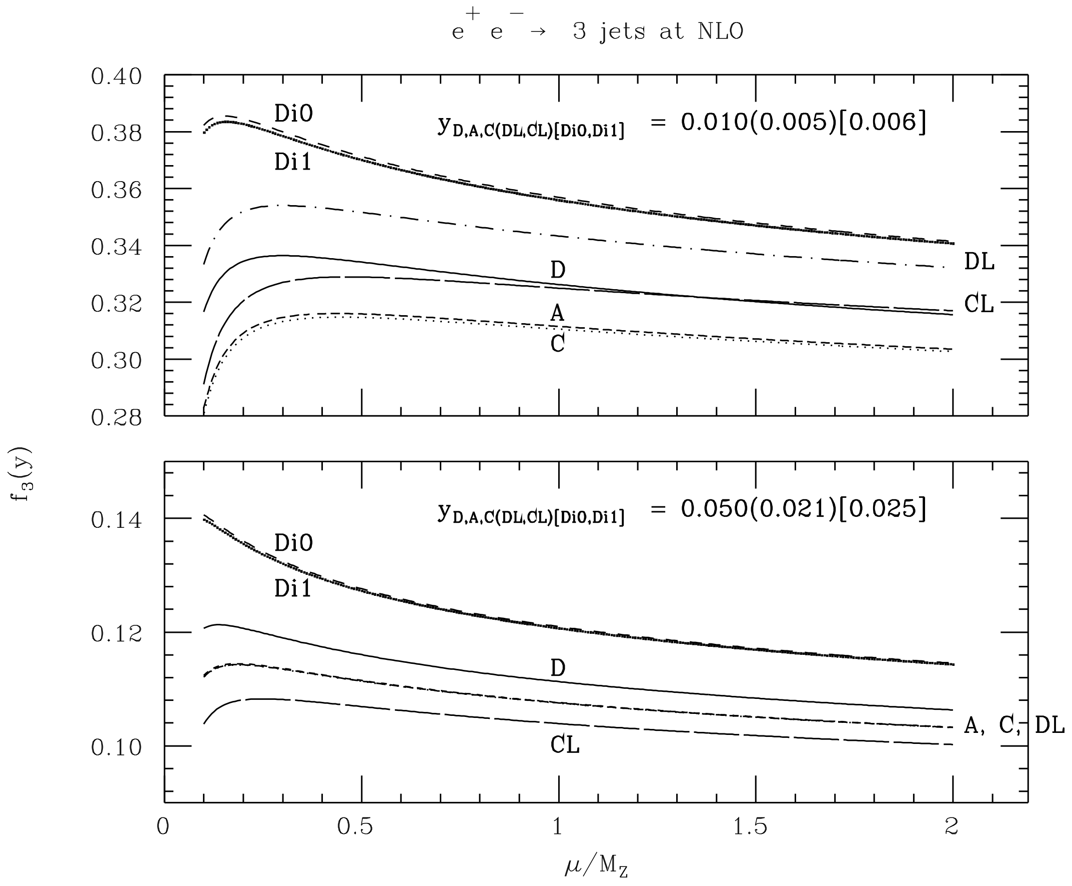

Figure 7: The parton level three-jet fraction at NLO as a function of the adimensional renormalization scale μ/M_Z, for $y_{\rm D,A,C(DL,CL)[Di0,Di1]} = 0.010(0.005)[0.006]$ (upper plot) and $y_{\rm D,A,C(DL,CL)[Di0,Di1]} = 0.050(0.021)[0.025]$ (lower plot), in the the D scheme (solid line), A scheme (short-dashed line), C scheme (dotted line), DL scheme (dot-dashed line), CL scheme (long-dashed line), Di0 scheme (dashed line) and Di1 scheme (fine-dotted line).

the physical process scale $\sqrt{y}M_Z$. Therefore, as a sensible range over which to estimate the effects of the uncalculated $\mathcal{O}(\alpha_s^3)\ +\ \ldots$ corrections one should adopt a reduced interval just below the value $\mu/M_Z = 1$. If one does so, then a remarkable feature of Fig. 7 is that the DL and CL algorithms show a noticeably reduced scale dependence as compared to the D and C ones, respectively, at low and especially at high y-values. Furthermore, among these two algorithms, it is the CL one that in general performs better, on the same footing as the C algorithm does as compared to the D one.

For example, at the low(high) y-values considered, the differences between the maximum and minimum values of $f_3(y)$ between $\mu/M_Z = 1/2$ and $\mu/M_Z = 1$ are 2.4(3.6)% for the DL scheme and 1.2(2.9)% for the CL one, respectively. In the case of the D and C algorithms, one has 2.4(4.3)% and 1.3(3.6)%, correspondingly. The numbers for the Di0 and Di1 algorithms are noticeably larger: 4.0(5.5)% and 4.0(5.5)% at small(large) y-values, respectively. To help the reader in disentangling the features of Fig. 7, we have reproduced some of the data points of the figure in Tab. 2.

$f_3(y)$							
$y_{D,A,C(DL,CL)[Di0,Di1]} = 0.010(0.005)[0.006]$							
μ/M_Z	D	A	C	DL	CL	Di0	Di1
0.25	0.336	0.312	0.311	0.353	0.325	0.383	0.381
0.50	0.334	0.316	0.315	0.352	0.329	0.371	0.370
0.75	0.330	0.314	0.313	0.347	0.327	0.363	0.362
1.00	0.326	0.311	0.310	0.343	0.325	0.357	0.356
1.25	0.323	0.309	0.308	0.340	0.323	0.352	0.351
1.50	0.320	0.307	0.306	0.337	0.320	0.348	0.347
$y_{D,A,C(DL,CL)[Di0,Di1]} = 0.050(0.021)[0.025]$							
μ/M_Z	D	A	C	DL	CL	Di0	Di1
0.25	0.120	0.114	0.114	0.114	0.108	0.134	0.134
0.50	0.116	0.111	0.111	0.111	0.107	0.128	0.127
0.75	0.113	0.109	0.109	0.109	0.105	0.124	0.123
1.00	0.111	0.108	0.108	0.108	0.104	0.121	0.121
1.25	0.110	0.106	0.106	0.106	0.103	0.119	0.119
1.50	0.108	0.105	0.105	0.105	0.102	0.117	0.117
$e^+e^- \to 3$ jets at NLO							

Table 2: The parton level three-jet fraction at NLO in correspondence of selected values of the adimensional renormalization scale μ/M_Z, for $y_{D,A,C(DL,CL)[Di0,Di1]} = 0.010(0.005)[0.006]$ (upper section) and $y_{D,A,C(DL,CL)[Di0,Di1]} = 0.050(0.021)[0.025]$ (lower section) in the D, A, C, DL, CL, Di0 and Di1 schemes.

The improvement in switching from CL to DL can be traced back to the implementation of the angular-ordering and soft-freezing procedures, as one of their side effects is to reduce the three-jet fraction: compare to eq. (3.1), where the $B(y)$ term enters with a positive sign (the leading piece proportional to $A(y)$ is the dominant one also at NLO). As pointed out in ref. [3] and also discussed earlier on, the reduced scale dependence and the smaller NLO corrections to the three-jet rate are intimately related.

The differences between the CL(DL) algorithm and the C(D) one can only be ascribed to the choice of the measure, as the clustering procedure is the same for both schemes. From the numbers in Tab. 2, it is clear that the reason for the improved performance goes beyond the relative importance of the LO and NLO pieces in the three-jet rate, as in some cases the DL and CL rates are above the D and C ones, respectively: e.g., at low y, where nonetheless the behaviour of the two measures is almost identical. Indeed, one could associate the better performances of the Luclus measure (2.6) as compared to the Durham one (2.5) to the fact that the first describes a *smooth* function of its arguments over all the available phase space whereas the second does not. This can be appreciated in Fig. 8, where the shape of the expression (here

DURHAM measure: $2\ \min\left(\frac{x_i^2}{4}, \frac{x_j^2}{4}\right)$ LUCLUS measure: $2\left(\frac{x_i}{2}\right)^2\left(\frac{x_j}{2}\right)^2/\left(\frac{x_i}{2}+\frac{x_j}{2}\right)^2$

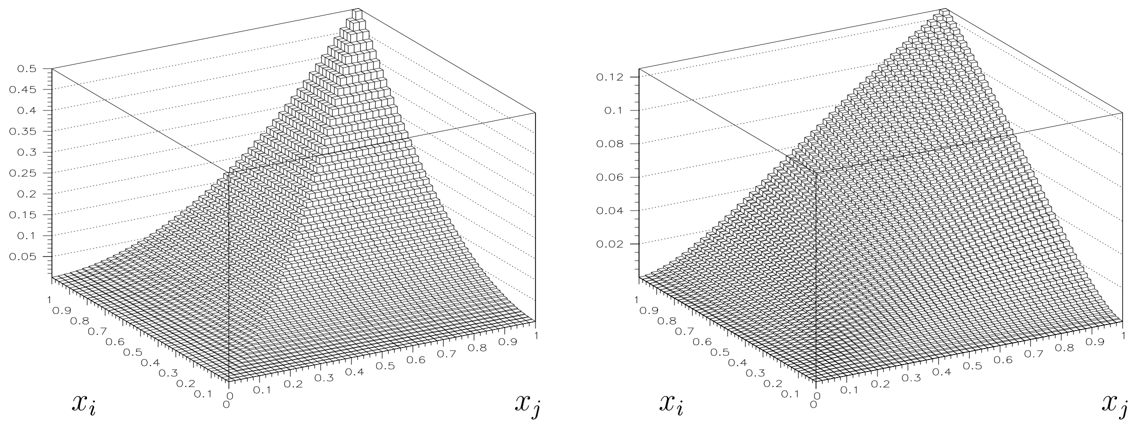

Figure 8: The dependence of the DURHAM (left-hand side) and LUCLUS (right-hand side) measure as a function of the reduced energies of the parton pair (ij).

Q plays the rôle of $E_{\rm vis}$ in sect. 2)

$$\frac{y_{ij}^D}{1-\cos\theta_{ij}} = 2\frac{\min(E_i^2, E_j^2)}{Q^2} \equiv 2\ \min\left(\frac{x_i^2}{4}, \frac{x_j^2}{4}\right) \tag{3.4}$$

for the DURHAM measure is compared to that of the LUCLUS one

$$\frac{y_{ij}^L}{1-\cos\theta_{ij}} = 2\frac{(|p_i|^2|p_j|^2)}{(|p_i|+|p_j|)^2Q^2} \equiv 2\frac{\left(\frac{x_i}{2}\right)^2\left(\frac{x_j}{2}\right)^2}{\left(\frac{x_i}{2}+\frac{x_j}{2}\right)^2}, \tag{3.5}$$

as bi-dimensional function of the reduced energies $x_i = 2E_i/Q$ and $x_j = 2E_j/Q$. For simplicity, we assume the two cluster i and j to be massless (i.e., $E_{i,j} = |\mathbf{p}_{i,j}|$) and drop the angular dependence $(1-\cos\theta_{ij})$.

It is well known that the presence of 'edges' at the border of the phase space defined by a jet algorithm is a source of misbehaviors in higher-order perturbation theory, as they ultimately generate infrared divergences (integrable, though) inside the physical region [40]. For example, one can refer to the so-called 'infrared instability' of the jet energy profile (dE_T/dr) in (iterative) cone algorithms typically used in hadron-hadron collisions, with r being the Lorentz-invariant opening angle of the cone defined in terms of pseudorapidity and azimuth (see, e.g., refs. [41, 42] for definitions and details). In fact, such a shape shows an edge in $\mathcal{O}(\alpha_{\rm s}^3)$ perturbation theory at the cone radius $r = R$. Although a resummation of logarithms $\pm\ln(|r-R|)$ to all orders cures the problem (as the edge eventually becomes a 'shoulder'!) a lesson to be learned is that it is clearly desirable to avoid observables with discontinuities when comparing with fixed-order predictions. (Similar conclusions can be drawn for the C-parameter in e^+e^- scatterings [40, 43].)

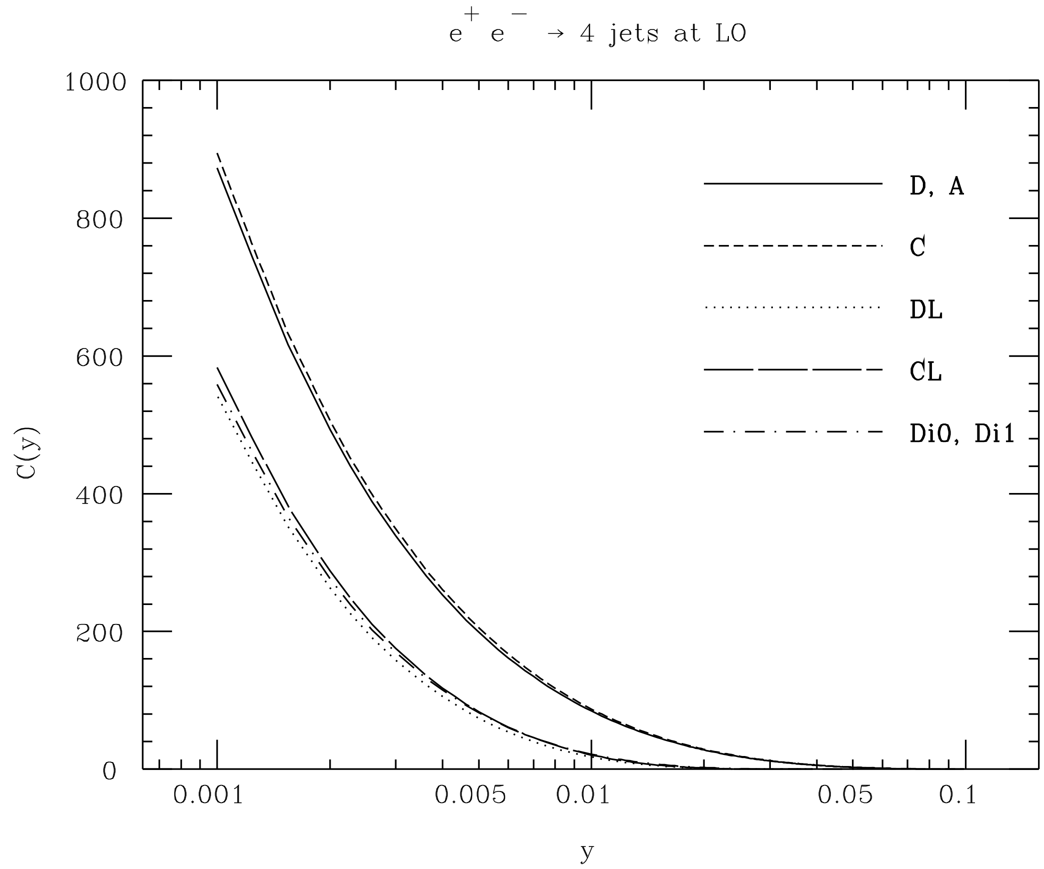

Figure 9: The parton level $C(y)$ function entering in the four-jet fraction (3.6) at LO in the D and A schemes (solid line), C scheme (short-dashed line), DL scheme (dotted line), CL scheme (long-dashed line), Di0 and Di1 schemes (dot-dashed line).

Although not quite the same context, it is not unreasonable to expect that the DURHAM measure might reveal sooner or later problems similar to those discussed, given its behavior along the trajectory $x_i = x_j$. In this respect, the original LUCLUS measure should represent a 'safer' observable. Indeed, the more sensitive scale dependence of the D and C algorithms, as compared to the DL and CL ones, respectively, could well be a first notice of possible problems in higher order pQCD.

Also the measure in DICLUS is a smooth function of its arguments. However, as discussed in section 2.7, DICLUS still have some problems with the seagull diagram giving larger three-jet rates and also larger scale dependence.

For completeness, we also present the rates for the four-jet fraction at LO. The analytical expression reads as follows (neglecting the μ-scale dependence in α_S)

$$f_4(y) = \left(\frac{\alpha_S}{2\pi}\right)^2 C(y) + \cdots , \tag{3.6}$$

where $C(y)$ is the corresponding coefficient function introduced on the same footing as $A(y)$ in the three-jet rates at LO. Its behavior is shown in Fig. 9, for the D, A, C, DL, CL, Di0 and Di1 schemes. Note that D and A, on the one hand, and Di0 and

Di1, on the other hand, coincide as no clustering between unresolved parton can take place at LO. Once again, we leave aside any comments about the D, A and C rates, for which we refer the reader to ref. [3]. As for the DL and CL algorithms, notice the increase of the LO rates due to the soft freezing mechanism, like between the D and C schemes. In this case the increase is larger, 7.4% against 2.5% at the minimum $y = 0.001$. The absolute value (of DL vs. D and of CL vs. C) is noticeably smaller, though: by approximately $35 - 38\%$ (at $y = 0.001$). Such a difference can be explained in terms of the LUCLUS and DURHAM measure, as was done while commenting Fig. 5. The DICLUS curve falls between the two LUCLUS ones. Thus, like in the case of the $A(y)$ function, for a fixed y, the y_{ij} value of the former is on average larger than the L but smaller than the D one.

We conclude this section by presenting a polynomial fit of the form

$$F(y) = \sum_{n=0}^{4} k_n \left(\ln \frac{1}{y}\right)^n \tag{3.7}$$

to the $F = A$, B and C functions, as already done in various instances in previous literature (see, e.g., refs. [3, 4]) now in the case of the DL, CL, Di0 and Di1 schemes. The lower limit of our parameterization is $y = 0.001$ for all three auxiliary functions. We extend the fits up to values where the four schemes yield sizable rates exploitable in phenomenological studies (see, e.g., ref. [44]). Typically, for the DL and CL algorithms these are around $y = 0.06(0.03)$ for A and $B(C)$, whereas for Di0 and Di1 a common value to the three function is $y = 0.1$. The values of the coefficients k_n, with $n = 0, \ldots 4$, are given in Tab. 3. Those for the D, A and C algorithms were given in ref. [3].

F	Algorithm	y-range	k_0	k_1	k_2	k_3	k_4
A	DL, CL	$0.001 - 0.06$	19.049	−18.991	5.891	−0.619	0.0314
A	Di0, Di1	$0.001 - 0.10$	11.436	−12.879	4.369	−0.461	0.0257
B	DL	$0.001 - 0.06$	691.886	−556.092	117.215	1.290	−1.349
B	CL	$0.001 - 0.06$	622.657	−504.153	107.150	1.388	−1.341
B	Di0	$0.001 - 0.10$	55.037	10.161	−63.296	27.529	−2.757
B	Di1	$0.001 - 0.10$	154.687	−99.096	−20.233	20.332	−2.335
C	DL	$0.001 - 0.03$	−172.821	108.274	−5.551	−7.255	1.152
C	CL	$0.001 - 0.03$	−239.022	176.112	−30.242	−3.593	0.981
C	Di0, Di1	$0.001 - 0.10$	−99.895	50.846	9.853	−8.880	1.214

Table 3: Parametrization of the three- and four-jet QCD functions $A(y)$, $B(y)$ and $C(y)$ as polynomials $\sum_n k_n(\ln(1/y))^n$, for the DL, CL, Di0 and Di1 algorithms. The range of validity in y is given for each case.

As a summary of our fixed-order studies, though limited to the NLO $\mathcal{O}(\alpha_s^2)$ rates, several conclusions can be drawn.

1. The angular-ordering and soft-freezing procedures advocated in ref. [3] represent a *genuine improvement* in fixed-order perturbative studies in the infrared dominion, provided these are plugged onto a $p_\perp$-based algorithm. (In fact, one should recall their 'inefficiency' when implemented within the JADE scheme [3], based on a mass measure.) This can be appreciated by noticing a reduced scale dependence of the NLO three-jet rates in both the C and CL schemes, as compared to the D and DL ones, respectively.

2. Of the two $p_\perp$-based binary measures considered here, the LUCLUS one yields better performances than the DURHAM one in terms of the stability of the NLO results against variation of the subtraction scale μ. This is presumably related to its definition in terms of the energies of the partons involved, which does not contain discontinuities or edges at the border of the phase space selected by the resolution parameter, contrary to the case of the DURHAM $p_\perp$-measure. Thus, an algorithm based on both angular-ordering and soft-freezing and exploiting the LUCLUS measure represents an alternative option to the CAMBRIDGE scheme to be adopted in the kind of studies performed here.

3. However, as for multi-parton studies in higher order pQCD, the exploitation of the original LUCLUS scheme should be limited to the adoption of its measure, not to the implementation of the preclustering and rearrangement steps recommended in the original version. On the one hand, these would introduce a considerable complication in both the numerical and (especially) analytical calculations. On the other hand, they would spoil the well known factorization properties of $p_\perp$-based algorithms, in resumming to all order in perturbation theory terms of the form $\ln y$ at small values of the parameter y. Indeed, one should notice that such properties are applicable to the case of the described DL and CL schemes, as the LUCLUS measure reduces to the DURHAM one in the soft limit. We will address this point specifically in the next section. In addition, we will also show that the sole adoption of the LUCLUS measure (i.e., the DL scheme) is not enough to reduce noticeably the size of the hadronization corrections of the D scheme (see subsection 4.2) so that CL is an option to be generally preferred to the DL one, as was the case for the C algorithm as compared to the D one [3].

4. Finally, the two DICLUS schemes based on the clustering of three-particles into two, do not show any substantial improvement, as compared to the conventional ones, which join two particles into one, at least in fixed-order perturbative studies of three-jet rates at NLO.

Before closing, we should mention that further analyses on the same footing as those described in this section are under way, for the case of the four-jet rate at NLO: from

QCD, using the program DEBRECEN [45][6], and from W^+W^- decays, exploiting the code used in ref. [47]. An account of the results in these contexts will be given in a future publication [48].

3.2. Resummed perturbative results

In this section we introduce a quantity which makes use of the results present in the previous section and which is very useful in investigating the interplay between perturbative and non-perturbative effects [3]. This is the mean number of jets, defined as

$$n_{\rm jets} \equiv \mathcal{N}(y) = \sum_{n=1}^{N} nF_n(y)\,, \tag{3.8}$$

where $F_n(y)$ is nothing else than the n-jet fraction introduced in eq. (3.2) in theoretical terms (thus, $F_n(y) = f_n(y)$ and $N = 4$ in subsection 3.1) i.e., as a ratio of cross sections. From the experimental point of view ($F_n(y) = \tilde{f}_n(y)$ and $N \to \infty$) the corresponding quantity is defined as a ratio of numbers of events, i.e.,

$$\tilde{f}_n(y) = \frac{N_n(y)}{\sum_m N_m(y)} = \frac{N_n(y)}{N_{\rm tot}(y)}\,, \tag{3.9}$$

where $N_n(y)$ is the amount of n-jet events and $N_{\rm tot}(y)$ the total hadronic sample. (This is the definition that we will adopt in computing the $n_{\rm jets}$ rates using the MC programs, in subsection 4.2.)

The mean number of jets can be studied as a function of the jet resolution parameter y, down to arbitrarily low values, at fixed energy. Furthermore, its perturbative behavior at very low values of y can be computed including resummation of leading and next-to-leading logarithmic terms to all orders in perturbation theory [49]. The corresponding predictions (particularly accurate at small y's) can then be matched with the fixed-order results (especially reliable at large y's) of the previous section over an appropriate interval in y, to give reliable pQCD estimates throughout the whole range of y. Non-perturbative contributions to $n_{\rm jets}$ can then be estimated by comparing the perturbative results with those of MC event generators. This will be done in subsection 4.2.

We first compute the resummed predictions for the DL and CL algorithms. In doing so, we make use of the results and formulae presented in ref. [49] for the case of the D scheme. Those were obtained in the case of multiple soft-gluon emission at small values of the resolution parameter. As we have already stressed in the Introduction, since in the soft limit in which either of i or j is much softer than the other the two measures (2.5) and (2.6) coincide, all the soft-gluon exponentiation properties of the DURHAM algorithm carry over to the LUCLUS one, provided no unnatural partition of the phase

[6]In fact, the long-awaited $\mathcal{O}\left(\alpha_{\rm S}^3\right)$ corrections to the four-jet rate have recently become available [46] and have been implemented in the mentioned code.

space is introduced by preclustering and/or reassignment. We do not perform the same fit for the DICLUS algorithms for two reasons. First, neither the Di0 or Di1 schemes have proven themselves being particularly suitable in pQCD studies, because of the larger scale dependence of their NLO rates. Second, the calculation of the resummed predictions would presumably be more complicated than in the case of the D and L schemes and would deserve an all new paper on its own.

Recalling that through second-order in α_S the two-jet fraction reads as

$$f_2(y) = 1 - \left(\frac{\alpha_S}{2\pi}\right) A(y) + \left(\frac{\alpha_S}{2\pi}\right)^2 (2A(y) - B(y) - C(y)) + \cdots , \tag{3.10}$$

and using the expressions (3.1) and (3.6) of subsection 3.1, one easily finds that in $\mathcal{O}(\alpha_S^2)$ pQCD the mean number of jets is

$$\mathcal{N}(y) = 2 + \left(\frac{\alpha_S}{2\pi}\right) A(y) + \left(\frac{\alpha_S}{2\pi}\right)^2 (B(y) + 2C(y) - 2A(y)) + \cdots . \tag{3.11}$$

The behavior of the first-order coefficient $A(y)$ at small y is given by

$$A(y) = C_F \left(\ln^2 y + 3\ln y + r(y)\right), \tag{3.12}$$

with the non-logarithmic contribution being [25, 29]

$$r(y) = 6\ln 2 + \frac{5}{2} - \frac{\pi^2}{6} + 4\left(\ln(1+\sqrt{2}) - 2\sqrt{2}\right)\sqrt{y} - 3.7y\ln y + \mathcal{O}(y) , \tag{3.13}$$

whereas for the second-order coefficient one has

$$\begin{aligned} F(y) &\equiv B(y) + 2C(y) - 2A(y) = \\ &= C_F\left[\frac{1}{12}C_A \ln^4 y - \frac{1}{9}(C_A - N_f)\ln^3 y + \mathcal{O}\left(\ln^2 y\right)\right]. \end{aligned} \tag{3.14}$$

In eqs. (3.12) and (3.14) the two quantities $C_F = 4/3$ and $C_A = 3$ are the Casimir operators of the fundamental and adjoint representations of the QCD gauge group $SU(N_C)$, these quantifying the strength of the $q \to qg$ and $g \to gg$ splittings, respectively, with $N_C = 3$ the number of colors, whereas the number of flavors is $N_f = 5$.

Note that, as long as terms of order $\ln^2 y$ are neglected, the above expressions are identical for the two versions of the LUCLUS based algorithms DL and CL (and so are they for the D, A or C schemes). In order to introduce the algorithm-dependent $\mathcal{O}(\ln^2 y)$ coefficients we adopt the same procedure as in ref. [3]. That is, we perform a fit of the form (3.7), restricted to the interval, say, $0.001 < y < 0.01$, with the coefficients k_3 and k_4 fixed at the values prescribed by eq. (3.14). The quantities k_0, k_1, k_2 are instead treated as free parameters. The numerical results are given in Tab. 4. Therefore, one can simply use the fits in Tab. 4 for, say, $y < 0.005$ and those in Tab. 3 for $y > 0.005$. (Indeed, over the region $0.001 < y < 0.01$, the transition between the two parameterization is very smooth.) This way, the second-order coefficient $F(y)$ can

Algorithm	k_0	k_1	k_2	k_3	k_4
DL	-76.264	4.257	4.108	-0.296	0.333
CL	14.275	-27.217	5.522	-0.296	0.333

Table 4: Parametrization of the second-order coefficient $F(y)$ in the average number of jets as a polynomial $\sum_n k_n(\ln(1/y))^n$, for the DL and CL algorithms. The range of validity is $y < 0.03$.

be obtained over the whole ranges of $y < 0.03$ (i.e., the DL and CL limits given in the second column of Tab. 3 for $C(y)$).

To obtain the final perturbative predictions for the mean number of jets, we now proceed as in ref. [49]. To next-to-leading logarithmic (NLL) accuracy, the resummed results are independent of the version of the algorithm. Therefore the only differences between DL and CL come from the matching to the fixed-order results. We simply subtract the first- and second-order terms of the NLL resummed result and substitute the corresponding exact terms. Denoting by $\mathcal{N}_q$ the NLL multiplicity in a quark jet, given in [49], we obtain

$$\mathcal{N}(y) = 2\mathcal{N}_q(y) + C_F\left(\frac{\alpha_S}{2\pi}\right)r(y) + \left(\frac{\alpha_S}{2\pi}\right)^2(F(y) - 2F_q(y)) \tag{3.15}$$

where F_q is the second-order coefficient in $\mathcal{N}_q$, given in [50]:

$$\begin{aligned} F_q(y) = C_F\bigg\{&\frac{1}{24}C_A\ln^4 y - \frac{1}{18}(C_A - N_f)\ln^3 y \\ &+\frac{N_f}{9}\left(1 - \frac{C_F}{C_A}\right)\left[\left(4\frac{C_F}{C_A} - 1\right)\frac{N_f}{C_A} - 1\right]\ln^2 y\bigg\}. \end{aligned} \tag{3.16}$$

We will make practical use of the formulae (3.15)–(3.16) later on, in subsection 4.2.

4. Non-perturbative comparisons

In this section we will attempt to quantify the effects due to hadronization for the various algorithms we have been discussing so far. In doing so we will resort to three among the most widely used QCD-based MC event generators: HERWIG [51] (version 5.9 [52]) Jetset [53] (Pythia version 6.1 [54], which incorporates Jetset 7.4) and Ariadne (version 4.10 [35]). In order to avoid 'philosophical' arguments about what *hadronization* actually is, we give here an *operational* definition useful for our purposes: *hadronization corrections* are the 'empirical adjustments' applied to the theoretical perturbative predictions before comparing them with the experiment. In an event generator, the former is represented by the partonic state before the hadronization routines are called and the latter by the state of final particles after hadronization and decays.

However, one must note a difference between the two. In the end there is an unambiguous identification between the hadron level of a generator and experiment, since the former is eventually tuned to reproduce data. The partonic state, on the other hand, is not a physically well-defined quantity. We therefore have to cope with the arbitrariness inherent in generators, which generally implement only enhanced terms of the infrared (i.e., soft and collinear) dynamics of quarks and gluons, thus introducing unnatural cut-offs and kinematic boundaries into the original QCD evolution. As a consequence, the mutilations done to the exact QCD dynamics in its PS approximation could well give rise to non-perturbative contributions already at the parton level [3].

In many studies, one wants to take one step further, and extract an α_{S} value from the deduced partonic level, based on the same parton-shower generator. This is more dangerous. For example, it has been shown in ref. [3] that HERWIG at the parton level overestimates the number of jets, as compared to the pQCD result, if the same α_{S} is used. Therefore a larger α_{S} has to be used for the resummed pQCD results than in the HERWIG shower to obtain the same parton level. Clearly, this knowledge is important in order to extract an α_{S} value, but it is irrelevant in a study of hadronization corrections. (More details on this point will be given later on, in subsection 4.2.) The only thing we would like to recall here is that progress in this direction is being made: e.g., that exact $\mathcal{O}(\alpha_s^2)$ LO matrix elements have been dressed with string fragmentation in the JETSET modeling since a long time, that further studies in the same environment involving a matching of the mentioned MEs to the parton cascade have also been carried out [55] and, finally, that an '$\mathcal{O}(\alpha_s^2)$ + PS + cluster hadronization' version of HERWIG will soon be publicly available [56]. The general problem of how completely to match the 'two parton levels' remains a key issue to be addressed, but that is evidently far beyond the scope of this paper.

Therefore, for most part of this section we leave aside the analytical formalism of the previous one, and only compare the partonic and the hadronic levels of generators. (We will however come back to it in one instance, at the very end of subsection 4.2.) One is indeed comforted in doing so by what we have already mentioned and will illustrate below: that we do find agreement among different MC programs. It is rather straightforward to study hadronization corrections in this spirit, since the generators provide lists of all partons and hadrons, event by event. In each case, jets are reconstructed both from the quarks and gluons at the end of the parton cascade and from the particles arising after hadronization and decays.[7] We will perform our analysis in the following subsections. Each of these corresponds to a different phenomenological context. In the first, we study the three-jet resolution in a simple tube model. In 4.2 we study hadronic events at LEP1, focusing our attention to the case of typical multi-jet quantities, in particular the mean number of jets defined in eq. (3.8). In the following subsection the emphasis will be on some kinematic properties of two- and three-jet

[7]Note that no simulation of detector acceptance and resolution are implemented in the latter case.

events at LEP1 energies. Finally, in subsection 4.4, we will study hadronization effects in the context of the mass reconstruction of $W^{\pm}$-bosons in four-jet events at LEP2.

Before proceeding further, for completeness, we briefly recall the properties of the three mentioned MC event generators. HERWIG implements the parton shower by coherent branching of the partons involved down to a fixed transverse momentum scale $Q_0 \sim 1$ GeV, and then converts these partons into hadrons using a cluster hadronization model [57]. In particular, the branching algorithm includes angular-ordering and azimuthal correlations of the emission (due to QCD coherence) along with the retention of gluon polarization and $p_\perp$ is the $\alpha_{\rm S}$ scale. The JETSET shower algorithm orders emissions in decreasing mass, with angular ordering imposed as an additional constraint, down to a cut-off mass $Q_0 \sim 1$ GeV. Azimuthal anisotropies from coherence and gluon polarization are also included, and the $\alpha_{\rm S}$ scale is $p_\perp^2$. Hadronization is done according to the Lund string model [20]. In ARIADNE [35], the ordering of emissions is in the invariant transverse momentum defined above in eq. (2.8) for the DICLUS algorithm. The same $p_\perp$ is used as the scale in $\alpha_{\rm S}$ and the cascade is continued down to $p_{\perp 0} \approx 0.6$ GeV. The coherence, treated by angular ordering in the HERWIG and JETSET parton showers, is inherent in the way gluons are emitted as dipole radiation from color connected pairs of partons. The azimuthal anisotropy due to gluon polarization is not explicitly taken into account but is reproduced to some extent by the dipole structure of the emissions. Hadronization is handled by the JETSET string fragmentation.[8]

Above we have seen that the algorithms in part are based on different considerations. One example is the picture of the perturbative shower evolution, which can be organized either in terms of decreasing opening angles of emissions or in terms of decreasing transverse momenta of emissions. Either of these two pictures gives a perfectly legitimate description of nature, but they arrive at different answers for what is the 'right' way to cluster a set of m partons into n jets. Even within a given calculational scheme, further uncertainties exist, such as how to distribute the recoil of an emission, i.e., the details of how energy and momentum is conserved. Add to this differences in the way non-perturbative physics is viewed, e.g., in string vs. cluster fragmentation models, and it is clear that there is not one unique view of the world. Therefore there is also not a unique criterion for what is the best possible clustering algorithm. One may then expect to find that different algorithms have complementary strengths and realize that the choice of algorithm should be based on the intended application.

4.1. Tube model results

As a preliminary exercise, we consider the simplest possible hadronization mechanism [58], the so called 'longitudinal phase space' or 'tube' model, in fact a simplified version of the Lund string model. Here a color-connected pair of partons produces a jet of light hadrons over a cylindrical $(y, p_\perp)$ phase space, where y is the rapidity and $p_\perp$ the

[8]An up-to-date review of our current understanding of hadronization is found in ref. [38].

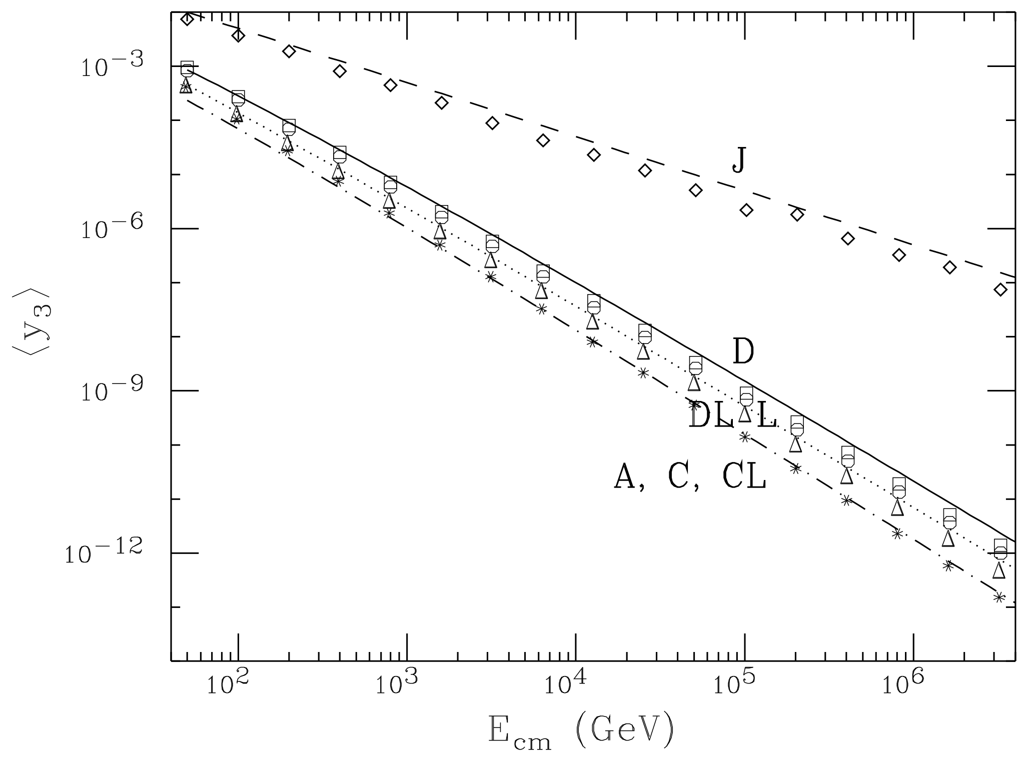

Figure 10: Mean value of the three-jet resolution y_3 in the tube hadronization model, for the J scheme (open-diamond symbols), D scheme (open-square symbols), Di1 scheme (circle symbols, not labelled for reason of space), DL and L schemes (open-triangle symbols), A, C and CL schemes (asterisk symbols). Note that the DL and L data points visually coincide, so do the A, C and CL ones. The corresponding curves show the approximate formulae discussed in the text.

transverse momentum (note that $y = \frac{1}{2}\ln[(E+p_z)/(E-p_z)]$, with z the cylinder axis and (E, p_x, p_y, p_z) the four-momentum). In practice, a number N of massless four-momenta are generated at random with an exponential transverse momentum distribution and a uniform rapidity distribution in the interval $-Y < y < Y$. The maximum rapidity Y is given by $E_{\rm cm} = Q = 2\lambda \sinh Y$ and the multiplicity N by $\lambda = N \langle p_\perp \rangle / 2Y$ (see ref. [3] for details). As illustrative values, we have taken $\lambda = 0.5$ GeV and $\langle p_\perp \rangle = 0.3$ GeV. In ref. [3], use was made of this model in order to illustrate some shortcomings of the DURHAM algorithm: that is, junk-jet formation and misclustering. We resume here those studies for the case of the algorithms that were not treated there.

As explained in subsection 2.5, by studying the mean value of the scale $y_{\rm cut} = y_3$ (denoted by $\langle y_3 \rangle$) at which a third junk-jet is resolved in the tube model one can get an insight on the effectiveness of the modifications to the DURHAM scheme proposed in ref. [3]. In the sense that, the smaller $\langle y_3 \rangle$ is, the more contained is the junk-jet phenomenon. As a matter of fact, both in the A and C algorithms [3] the average value of y_3 is much smaller, as compared to the D scheme, over an enormous range of energies $Q \equiv E_{\rm cm}$. Figure 10 illustrates this.

We supplement the results for D, A and C of ref. [3] by adding in Fig. 10 the corresponding behaviors for the L, DL and CL binary schemes, and for the Di1 scheme. Note that here L refers to the original LUCLUS scheme implementing both preclustering and reassignment.

It is curious to realize once more (see subsection 3.1) the beneficial effects in the DL scheme of adopting the LUCLUS measure instead of the DURHAM one, as the corresponding data points lay well below those of the original D scheme. In contrast, preclustering and rearrangement bring no noticeable improvement in this context, neither separately nor together: notice the overlap of the DL and L curves. (For readability of the figures, we have avoided plotting the cases of the LUCLUS scheme with only one of preclustering and reassignment.) This probably indicates that the junk-jet formation is dominated by the existence of a single high-$p_\perp$ track or a few very nearby tracks. However, it is clear that the further step of angular-ordering is needed even in the case of the LUCLUS measure in order to bring the results further down. Once this is implemented, there is no sizable difference between the performances of the two different measures (C vs. CL). In a sense, the sole adoption of the LUCLUS y_{ij} helps to alleviate the original problem, but is not enough to cure it in the same way the angular-ordering does. In addition, it is evident that the latter procedure removes any distinction between the two measures. For reference, we should also mention that the ANGULAR-ORDERED LUCLUS (which we do not study further here, see subsection 2.5) can boast identical performances to those of the A, C and CL schemes. The DICLUS algorithm nicely interpolates between the D and L ones, following the relative behaviour of the measures, similarly to the case of Fig. 5.

Thanks to simple kinematic relations valid within the tube model and depending on the clustering procedures of the various algorithms, it was shown in ref. [3] that one can approximate the behaviors of the binary algorithms in Fig. 10 by means of some analytical formulae, as a function of the collider CM energy Q. We recall here those for the J, D, A and C schemes:

$$\langle y_3\rangle^J \sim \frac{\lambda}{Q}, \qquad \langle y_3\rangle^D \sim \left(\frac{2\lambda\ln(Q/\lambda)}{\pi Q}\right)^2, \tag{4.1}$$

$$\langle y_3\rangle^A \approx \langle y_3\rangle^C \approx \langle y_3\rangle^{CL} \sim \left(\frac{\lambda}{Q}\ln\ln(Q/\lambda)\right)^2. \tag{4.2}$$

As evident from the plot, CL coincides with A and C. It is also easy to show that the corresponding equation for both the DL and L schemes reads as follows:

$$\langle y_3\rangle^L \approx \langle y_3\rangle^{DL} \sim \left(\frac{\lambda\ln(8Q/\lambda)}{\pi Q}\right)^2. \tag{4.3}$$

The curves in Fig. 10 correspond to the above formulae: the dashed line to the J scheme, the continuous one to the D algorithm, the dotted one refers to L and DL whereas the dot-dashed one to the A, C and CL cases.

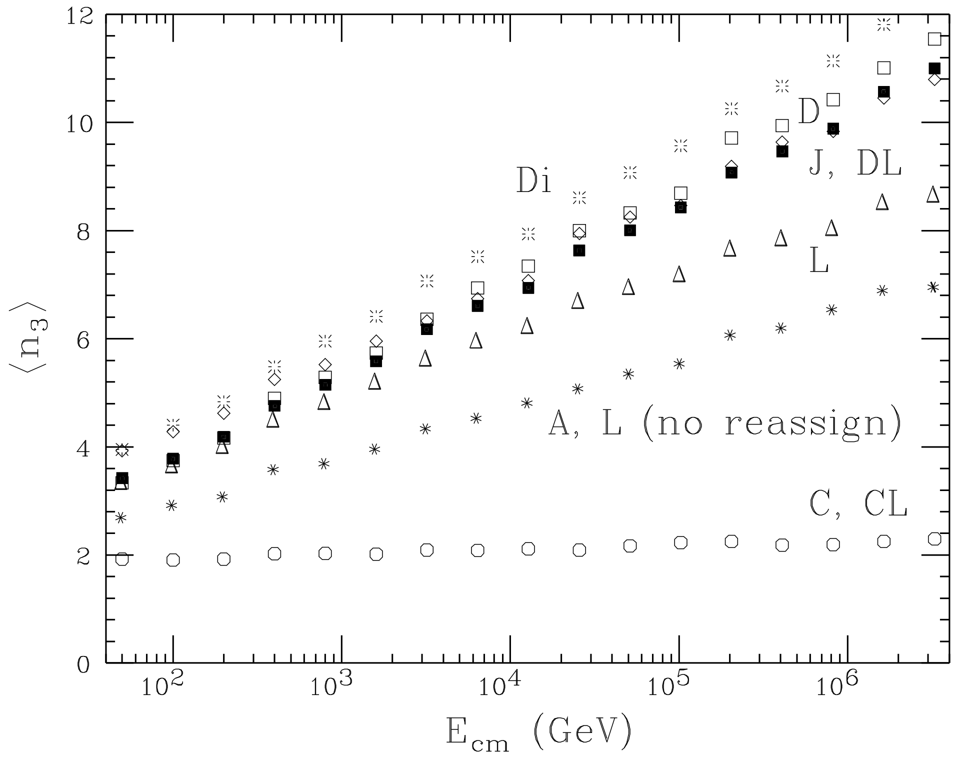

Figure 11: Mean number of particles in the third jet at $y_{\rm cut} = y_3$ in the tube hadronization model, for the J scheme (open-diamond symbols), D scheme (open-square symbols), DL scheme (full-square symbols), L scheme (open-triangle symbols), A and L (no reassignment) schemes (asterisks symbols), C and CL schemes (open-circle symbols), Di scheme (star symbols). Note that the A and L (no reassignment) data points visually coincide, so do the C and CL ones.

In order to test the efficiency of the soft-freezing procedure proper of the C scheme a good quantity to study is the mean number of particles contained in the the third (softest) jet when $y_{\rm cut} = y_3$, which was denoted by $\langle n_3 \rangle$ in ref. [3]. Clearly, the smaller this quantity is on average, the less particles have been attracted inside the original resolved (soft) cluster (see the discussion in subsection 2.6) and the misclustering effect is thus reduced. Figure 11 shows $\langle n_3 \rangle$ as a function of the CM energy, over the same $E_{\rm cm}$ range as in the previous plot. The relative behaviors of the J, D, A and C schemes were already illustrated in detail in ref. [3], so we do not dwell here on them. Rather we emphasize that the adoption of the Luclus measure is less helpful in this case, as the DL and D curves almost coincide. Furthermore, it is worth spotting that the reassignment procedure increases the mean value of n_3, as can be appreciated by comparing the data points labelled 'L' and 'L (no reassignment)'. It is therefore clear that such a procedure, which does remedy the problem of the misassignment of soft particles (see subsection 2.3) is instead inefficient in suppressing misclustering. This is evident if one considers that such a step tends to 'balance' the event, by reassigning tracks among clusters so that the jets show in the end similar multiplicity (thus acting in contrast to what soft-freezing does).

Curiously, the preclustering seems to work on the same footing as the angular-ordering: compare 'A' and 'L (no reassignment)'. However, here the almost exact agreement is somewhat of a coincidence. As intimated in the Introduction, this is a consequence of having adopted the default $d_{\rm init}$ value in producing the 'L (no reassignment)' curve. We have verified that in the limit $d_{\rm init} \to 0$ the DL results are in fact recovered. This makes clear the possible danger of ascribing artifacts of the algorithm to real physics effects, if a wrong setting of $d_{\rm init}$ is adopted. Finally, like in the case of $\langle y_3 \rangle$, the adoption of the soft-freezing procedure wipes out differences between the Durham and Luclus measures (compare C and CL). As for the Diclus scheme, we note that it yields the largest $\langle n_3 \rangle$, especially at high energies, this witnessing the tendency of this algorithm of clustering soft resolved particles which are source of dipoles (see discussion in subsection 2.7 while commenting Fig. 4.)

Evidently, the tube model adopted in the previous paragraph should be regarded as a useful tool in order to test the performances of the various algorithms with respect to the misbehaviors responsible for junk-jet formation and misclustering, which naturally arise in the physics dominion governed by soft radiation, for all algorithms based on $p_\perp$- and m-measures [3]. In the very end, however, the benchmark ground to verify the goodness of an algorithm in terms of hadronization corrections should be a full MC program, such as those previously mentioned.

4.2. Jet rates

In ref. [3] the quantity $n_{\rm jets}$ was studied, both at hadron and parton level, as a function of the resolution parameter, down to the minimum figure of $y_{\rm cut} = 0.0001$. Such a choice of range was not a random one. Indeed, it is well known that the energy scale at which QCD enters its non-perturbative phase is around 1 GeV. This is also the typical value at which the QCD-based MC programs stop the parton cascade and turn on the hadronization process. Clearly, if one intends to probe the interface between perturbative and non-perturbative QCD by studying jet properties as a function of $y_{\rm cut}$, one should aim for an algorithm with small hadronization corrections for resolution scales down to ≈ 1 GeV, corresponding to $y_{\rm cut} \sim 10^{-4}$ at LEP1 energies, where $Q \approx M_Z$. Otherwise it becomes more difficult to understand the QCD dynamics in such a critical regime, and systematic uncertainties (due to different MC modelings) must be accounted for.

In Figs. 12 and 13 we first note the difference between the Jade and, e.g., the Durham algorithms at small resolution scales. This is a consequence of the adoption of a measure which is an invariant mass one in Jade, rather than a transverse momentum. We have in fact already recalled in the Introduction that a $p_\perp$-distance is better adapted to the conventional picture of non-perturbative jet fragmentation and therefore naturally allows a cleaner separation of perturbative and non-perturbative physics [25, 29, 30]. We also quantified this point in the tube model when mentioning

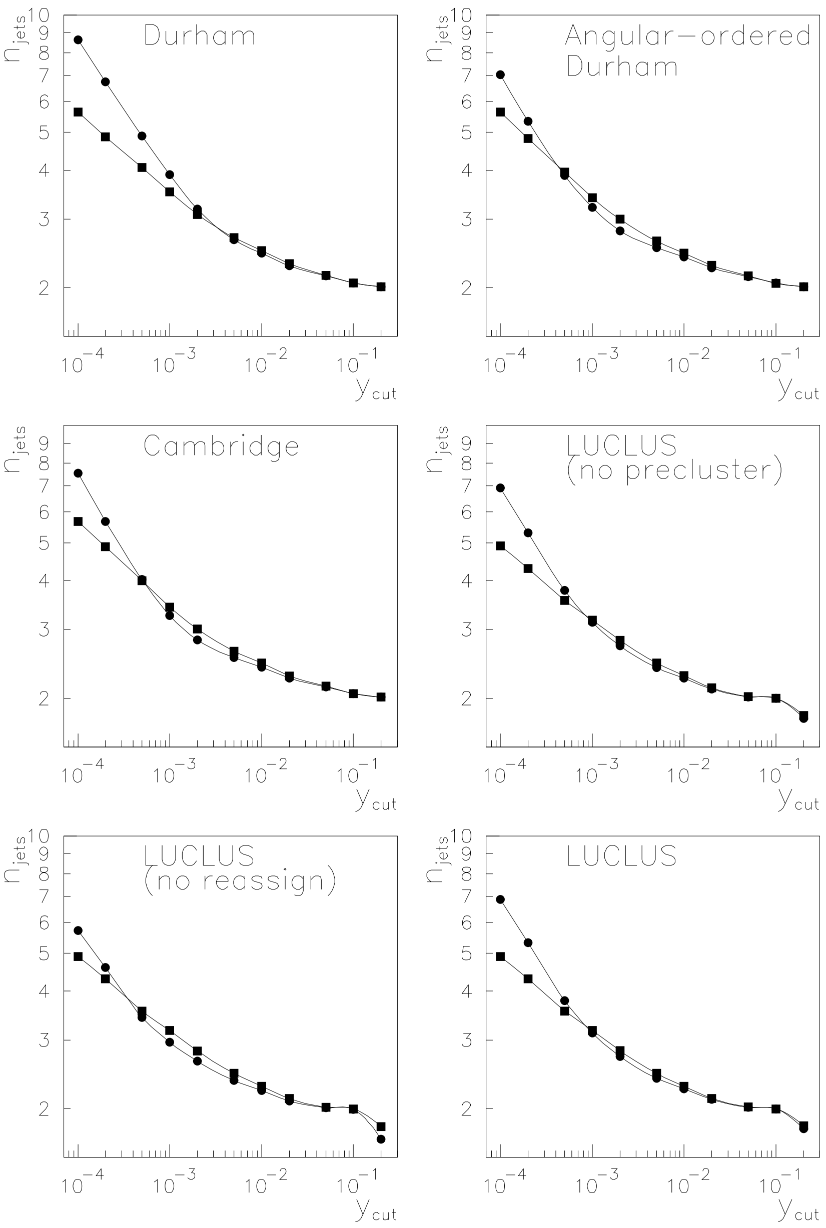

Figure 12: Comparison of jet algorithms. Parton level (squares) and hadron level (circles) results from HERWIG on the mean number of jets at $Q = M_Z$ for various jet algorithms as labelled, as a function of the jet resolution variable $y_{\rm cut}$. The statistical errors are smaller than the size of the points. Results are very similar for Jetset and Ariadne.

that the power-suppression in Q on $\langle y_3 \rangle$ goes like $(\ln Q/Q)^2$ in the Durham scheme, whereas the corresponding behavior in the Jade one is $1/Q$: see eq. (4.1) and, for a more theoretically sound basis, also ref. [59].

That a $p_\perp$-based measure is indeed a better choice is confirmed not only by the fact that also the various Luclus algorithms can boast a power-suppression similar to that of the Durham scheme, see eq. (4.3), but also by observing that a common feature of Figs. 12–13 is that *all* the transverse momentum based schemes (also the Diclus ones) are better behaved than the Jade one at small resolution scales.

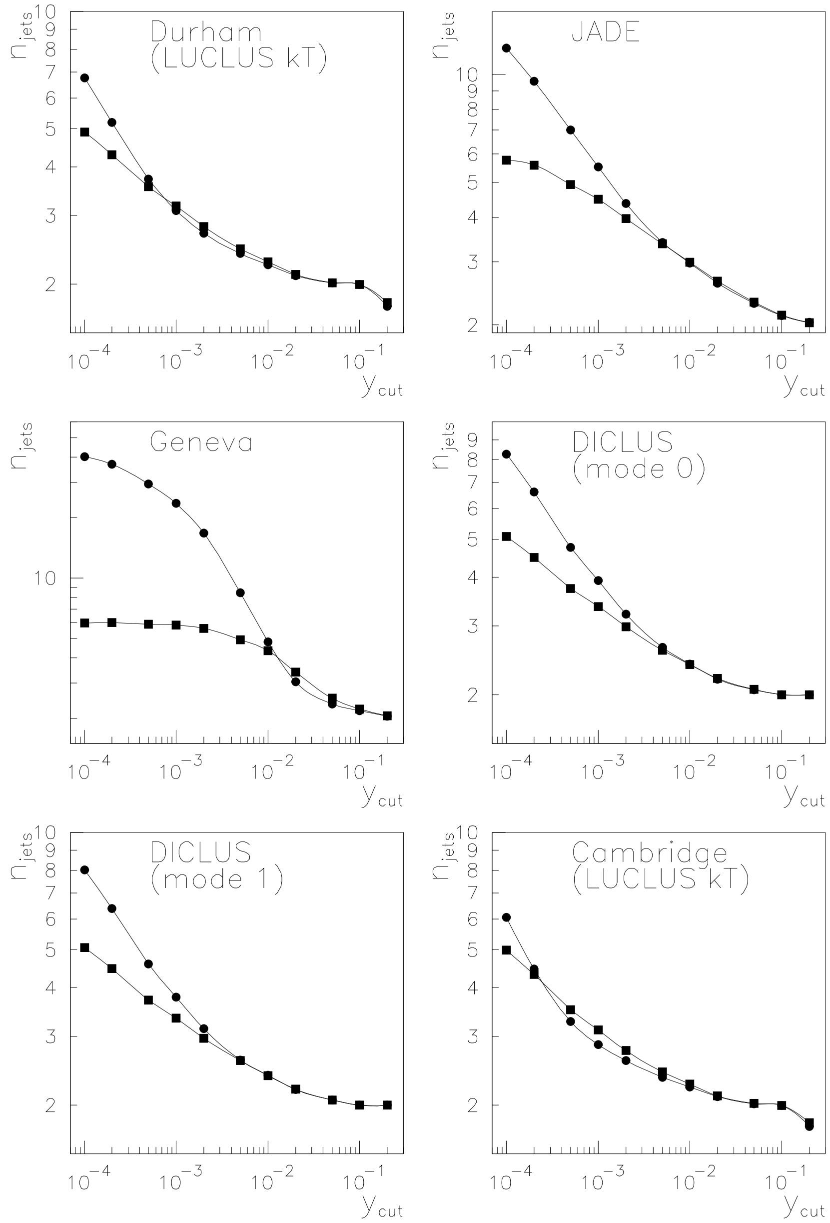

Figure 13: Comparison of jet algorithms. Parton level (squares) and hadron level (circles) results from HERWIG on the mean number of jets at $Q = M_Z$ for various jet algorithms as labelled, as a function of the jet resolution variable y_{cut}. The statistical errors are smaller than the size of the points. Results are very similar for JETSET and ARIADNE.

Comparisons must be done with some care, however, since the horizontal y scale means different things for many of the algorithms shown in Figs. 12–13. For a pair of partons/hadrons i, j, the definitions give that $(y_{ij})_{Jade} > (y_{ij})_{Durham} > (y_{ij})_{Luclus}$, since the three measures share the same angular dependence and differ only in the energy factors being proportional to E_iE_j, $\min(E_i^2, E_j^2)$ and $E_i^2E_j^2/(E_i + E_j)^2$, respectively (disregarding the difference between $|\mathbf{p}_i|$ and E_i for LUCLUS). It then follows that JADE differs significantly from the other two when $E_i \ll E_j$, that DURHAM and LUCLUS differ by up to a factor four for $E_i \approx E_j$, and that JADE and LUCLUS always differ by more

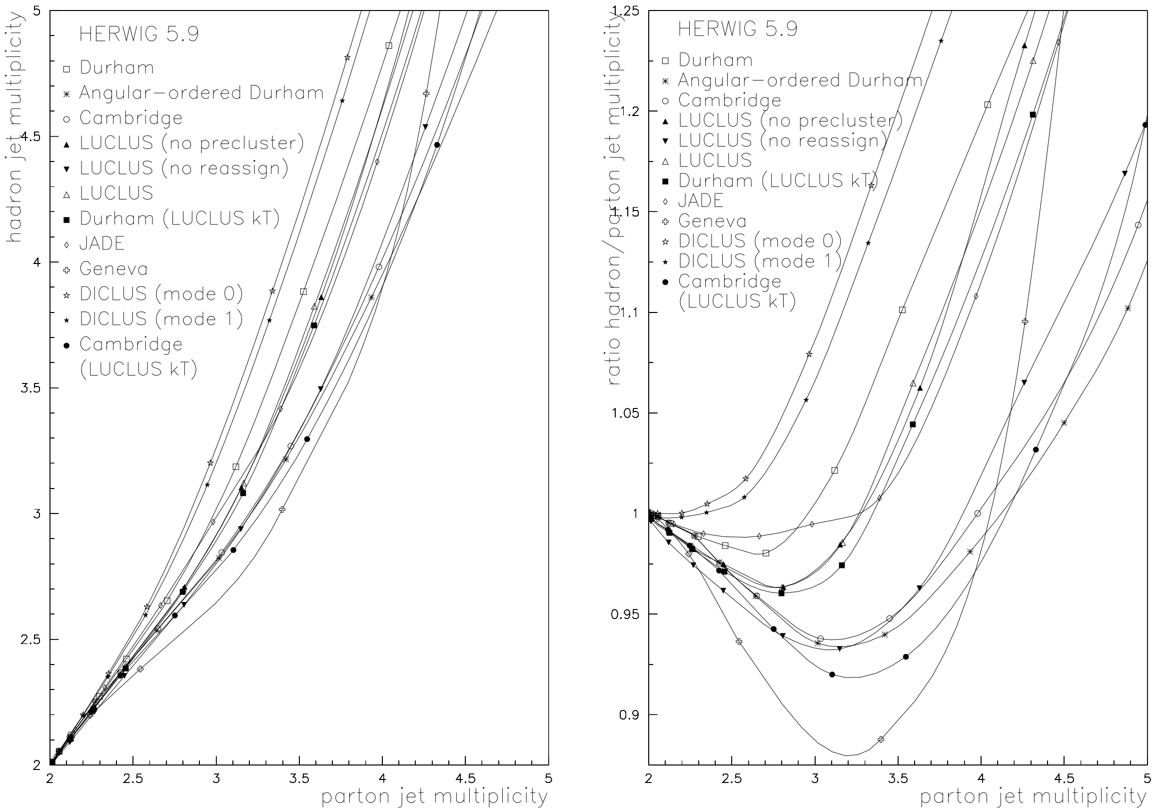

Figure 14: The average hadron jet multiplicity vs. the parton one (left plot) and the ratio of the two (right plot) at $Q = M_Z$ for various jet algorithms as labelled, using HERWIG. The statistical errors are smaller than the size of the points.

than a factor four. Even more different is the Geneva scheme, where the energy factor on the same scale would be $(4/9)E_iE_jE_{\rm vis}^2/(E_i + E_j)^2$. Since normally $E_i + E_j \ll E_{\rm vis}$, it follows that $(y_{ij})_{Geneva} \gg (y_{ij})_{Jade}$ most of the time. The transition region between pQCD and non-pQCD is thus no longer around 10^{-4} in the Geneva scheme, but maybe more like at 10^{-2}, judging by the jet rate in Fig. 13.

An alternative procedure to compare jet algorithms is offered by Fig. 14, where the average hadron jet multiplicity is plotted against the average parton one. The $n_{\rm jets}$ values in the plot are exactly the same as defined by the curves in Figs. 12–13, but with the explicit $y_{\rm cut}$ dependence eliminated in each pair of average jet numbers. This is done for the same algorithms analyzed in the previous two figures. Both the parton and the hadron jet multiplicity increase with a diminishing $y_{\rm cut}$. The criterion of a good algorithms as one with small hadronization corrections thus translates into one with a curve close to the diagonal in the left-hand side of Fig. 14 or, alternatively, to the horizontal at unity in the right-hand side of Fig. 14, where the hadron-to-parton jet multiplicity ratio is presented.

If one intends to probe the QCD phase transition at small $y_{\rm cut}$, then the relevant region is that at large jet multiplicities. There, we notice that four 'bunches' of curves distinctively separate. Four algorithms remain particularly close to the diagonal/horizontal line. Not surprisingly, among these are the schemes that had been especially designed (in ref. [3]) to improve the performances of the DURHAM scheme: such as the ANGULAR-ORDERED DURHAM (asterisks symbols) and the CAMBRIDGE (open-circle symbols) which are the closest to the diagonal/horizontal line. To this group also belong the LUCLUS scheme with only preclustering implemented (with $d_{\rm init}$ kept at its default value 0.25 GeV even at low $y_{\rm cut}$, full-down-triangle symbols) and the CAMBRIDGE algorithm using the LUCLUS measure (full-circle symbols). This is not surprising if one recalls Fig. 11 in the tube model. That plot, on the one hand, had already shown the residual effects of preclustering with the default $d_{\rm init}$ and, on the other hand, had also made the point that the CAMBRIDGE scheme using the LUCLUS measure performs as well as the original one with the DURHAM distance. In addition, always in line with what was assessed in the tube model, it is clear that the LUCLUS measure alone is not enough to improve the hadronization performances of the DURHAM scheme (full- and open-square symbols, respectively). In fact, the DL algorithm belongs to another set of curves (along with the JADE, DURHAM and the two other configurations of the LUCLUS) whose hadronization corrections are much larger with increasing multiplicity (i.e. decreasing $y_{\rm cut}$). A third group is constituted by the two DICLUS algorithms, which perform very well in the two-jet-dominated region but rather worse as soon as a third jet is resolved. Also the JADE algorithm stands out here and show very stable hadronization corrections up to and above the three-jet region. The GENEVA algorithm performs worst, since it starts out with the largest negative hadronization corrections and thereafter steeply shoots up towards the largest positive ones.

Figure 15 reproduces the rates of the CAMBRIDGE algorithm, already given in Fig. 14 for the case of HERWIG, now extended to include JETSET and ARIADNE data points, too. Indeed, the pattern of the hadron/parton multiplicity is very similar among the three programs, though with a more marked tendency of the rates of departing from the diagonal in the latter two cases. We have verified (though not shown) that similar consistent behaviors among the three generators also occur in the case of the other jet schemes.

A very interesting aspect of Figs. 12–14 is the 'negative hadronization corrections' (see ref. [4] and also ref. [60] for JADE) i.e., that fewer hadron than parton jets are reconstructed in the three-jet-dominated region. This can be easily observed in the case of the GENEVA, the ANGULAR-ORDERED DURHAM, the 'two' CAMBRIDGE and the various LUCLUS schemes. Though less visible there, it also occurs for the JADE and DURHAM algorithms. Not even the two modes of DICLUS are immune from it, though here the effects are small.

This phenomenon has a very straightforward interpretation in terms of the well-known string [61] or drag [62] effect. The concept can be illustrated by considering a three-parton event $q\bar{q}g$. To leading order in $1/N_C^2$, the system separates into two color dipoles, one qg one where the color of the quark is compensated by the anticolor of the antiquark, and another similar $g\bar{q}$ one. Each of these dipoles can act as a source

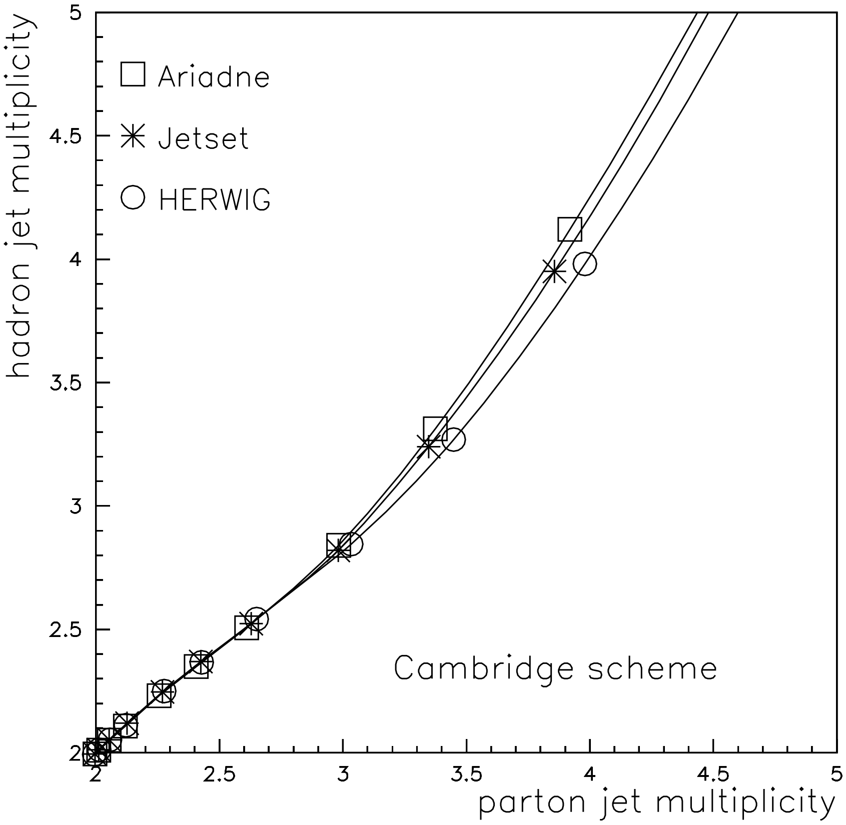

Figure 15: The average hadron jet multiplicity vs. the parton one for the CAMBRIDGE scheme at $Q = M_Z$ for the three event generators ARIADNE, JETSET and HERWIG, as labelled. The statistical errors are smaller than the size of the points.

for further softer perturbative emission or define the topology of a non-perturbatively hadronizing string piece. In the (transverse) rest frame of such a dipole the emission of partons and hadrons is isotropic in azimuth but, when viewed in the CM frame of the event, particle production is not symmetric around the three jet axes. Instead enhanced soft particle production is found in the angular ranges spanned by the dipoles, i.e., in the qg and $g\overline{q}$ ranges. There is no corresponding $q\overline{q}$ dipole — in fact, the color-suppressed dipole of this kind enters with a negative sign, i.e., provides destructive interference to the other two. Therefore the $q\overline{q}$ angular range has a much smaller soft particle production than the other two ranges. This effect is well established experimentally [63].

A corollary of the string/drag effect is that reconstructed jet directions also can display a systematic bias. For instance, with respect to the original quark parton direction, softer particles will predominantly be produced on the side of the gluon jet, and less on the antiquark side. Therefore a naive clustering of hadrons will find a quark jet axis somewhat shifted towards the gluon. The original parton direction is not known experimentally, of course, but the effect is visible by harder particles in the quark jet appearing slightly more on the antiquark side (lined up with the original parton direction) and softer particles more on the gluon side [64]. The antiquark is similarly affected, while the gluon receives opposite contributions from the two string pieces attached to it. Simple geometry shows that a dipole is more boosted and the string/drag effect therefore more developed when two partons are nearby. If the gluon is closer to the quark than to the antiquark, say, the gluon and quark reconstructed directions will be shifted closer to each other, while the antiquark direction is less affected. (The reconstruction of jet directions is further tested in the next section.) Thus a three-jet event becomes more two-jet like, in the sense that the y_3 value where the event flips from a two-jet to a three-jet is lower on the hadron than on the parton level. Hence the negative hadronization corrections. These arguments are valid both for string and cluster hadronization — the clusters are aligned by the same colour topology as the strings and are similarly boosted — and both models correctly reproduce the measured string/drag effect. In the following, to shorten our discussion, we will discuss the dynamics of hadronization in terms of the string model, but recall that an equivalent formulation in terms of the cluster one is always possible.

The magnitude of the above effect depends on the details of the algorithm used. For instance, the shift inwards of the two jet directions can be viewed as a kinematical consequence of replacing two massless partons with two massive jets, while still retaining the same total invariant mass of the pair. Thus the negative hadronization corrections should be absent in the Jade E scheme, where the correct invariant mass is used as distance measure, eq. (2.4). (But, of course, Jade E has its own problems of misclustering, so is not the solution.) The corrections are still rather small in the normal Jade scheme. The Diclus algorithm is the only one deliberately designed to reflect particle production along hyperbolae, and to reconstruct the directions of the asymptotes of those hyperbolae, while others implicitly are based on a picture of a jet as a set of particles extending away from the origin along a fixed momentum direction. As we see, Diclus does a pretty good job of its intended task in the two-jet region. Other things being the same, in this region Diclus would thus be the preferred choice. The other algorithms have differently large negative hadronization corrections from this effect, reflecting the details of the distance measure and the clustering scheme. For instance, using $|\mathbf{p}_i|$ rather than E_i emphasizes the importance of jet masses acquired in the hadronization; thus the DL curve lies below the D one.

What is the most 'desirable' behaviour here is not so easy to tell. On the one hand, small hadronization corrections would be better for perturbative studies, on the other hand, the string phenomenon is a genuinely interesting piece of non-perturbative physics that deserves to be studied in its own right. And even for an α_s determination, ultimately what matters is not whether a hadronization correction has to be applied or not, but how large an error bar has to be assigned to this correction. Figure 16 (left picture) here illustrates that the event generators we have tried agree very well once again. That is, the phenomena described above are of general validity, and seem to be accountable to a similar extent for most measures.

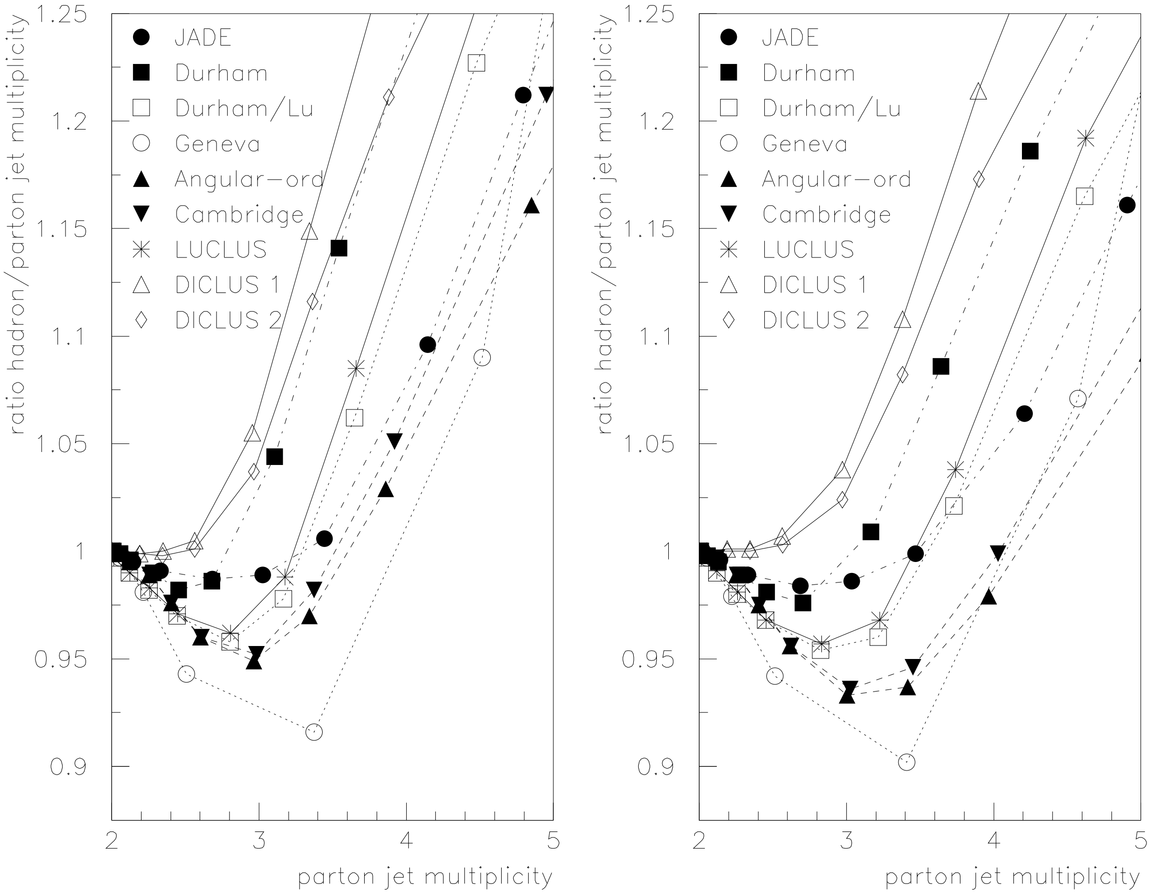

Figure 16: The ratio of the average hadron and parton jet multiplicities as a function of the average parton jet multiplicity at $Q = M_Z$ for various jet algorithms as labelled, using JETSET with the ARIADNE dipole cascade. Left plot is for all flavors, while the right plot is for u quark events only.

From the figures, it would seem that the above string/drag effects are present for three-jets but absent for higher jet multiplicities. This is not fully correct, however. In any hadronic n-jet event, the two closest jets are likely to correspond to dipole-connected partons. (In the leading-log picture of shower evolution, the only exception is given by the rather infrequent $g \to q\overline{q}$ branchings.) Therefore the hadron-level jet directions will sit closer than the parton-level ones, and one is more likely to get an $(n-1)$-jet event on the hadron level than on the parton level. The turnaround of the curves, and ultimately the larger hadron than parton average jet multiplicity at small y_{cut}, is thus rather a reflection of other effects entering and becoming more important. These can collectively be classified as fluctuations in the hadronization process, but can have different origins. One example is the junk-jet formation discussed repeatedly above, e.g., in subsection 4.1. Another is the kinematics of particle decays, especially of bottom and charm hadrons. This latter effect is illustrated in Fig. 16 (right picture) where results for the production of all initial quark flavors are compared with those for

u quarks only. It is seen that the u quark sample consistently produces lower ratios than the all-flavour one. Conversely, a pure b quark sample (not shown) has much higher ratios, i.e., more spurious jets produced by the B decays.

If jet rates are to be used to investigate fundamental QCD properties, one also needs to understand the relation between the parton level curves given by a generator and those expected in an exact theory. For example, as previously recalled and as already shown in ref. [3] for the D, A and C algorithms, if the same $\alpha_{\rm S}$ produced by the generator is used for the pQCD leading-log resummed + $\mathcal{O}(\alpha_{\rm S}^2)$ fixed-order predictions, then the corresponding parton level curves would fall below the hadron level over a large $y_{\rm cut}$ spectrum. In fact, a good matching (for all $y_{\rm cut}$'s) between the pQCD predictions and the HERWIG parton level is obtained if the former use $\alpha_{\rm S} \equiv \alpha_{\rm S}(M_Z^2) = 0.126$, instead of 0.114, the value obtained from the generator by interpreting the input parameter `QCDLAM`, with the default value of 0.18 GeV, as the NLO scale parameter $\Lambda^{(5)}_{\overline{\rm MS}}$. The necessity of such 'rescaling' should be not surprising, as such an interpretation of `QCDLAM` is only justified in a small region of phase space (see [65]) which should not be dominant.

Figure 17 plots the rates as obtained from the formulae (3.15)–(3.16), that is, the resummed predictions for the DL and CL algorithms, for three values of $\alpha_{\rm S}$, against the HERWIG parton and hadron levels. (They are the counterpart of those for the D, A and C scheme presented in Figs. 13–15 of ref. [3].) From Fig. 17 is then clear that, if the LUCLUS measure is used instead of the DURHAM one, things go the other way round: the dotted curves, for which $\alpha_{\rm S} = 0.114$, are the ones closer to the MC parton level. This seems to indicate that the LUCLUS $p_\perp$ is more blind to the differences between the parton-shower and the perturbative results than the DURHAM one. Furthermore, this is particularly true at very small $y_{\rm cut}$, the critical regime where not only a reduced size of the hadronization corrections is required to study the transition between pQCD and non-pQCD, but possibly also a good control of the differences between the theoretical and phenomenological parton levels. Thus, quite apart from the implications for QCD studies at large $y_{\rm cut}$ (such as the recalled measurements of $\alpha_{\rm S}$ from $f_3(y)$, where the parton level of the generator plays indeed little rôle) we merely want to emphasise here the attractiveness of using a jet resolution variable that very similarly accomodates both the theoretical and phenomenological treatments of QCD, and in a physics dominion (i.e., very small $y_{\rm cut}$) where the understanding of the behavior of MC event generators is crucial in order to probe the dynamics of the transition phase between the two QCD regimes. In this respect, rather than over-focusing on the correspondence between the two parton level curves obtained for the same value of the strong coupling constant (which, after all, could well be coincidental, given that the $\alpha_{\rm S}$ definitions in the model and in the perturbation theory are notably different, as recalled above) we would like to point out that the asymptotic, i.e., very low $y_{\rm cut}$, behaviour of the two is identical for the CL scheme, whereas in the case of the C one the two curves tend to diverge in that region (compare the right-hand side of Fig. 17 here to Fig. 15 of ref. [3]). This last fact is also true for the DL algorithm (left-hand side of Fig. 17). Thus, also in this case the

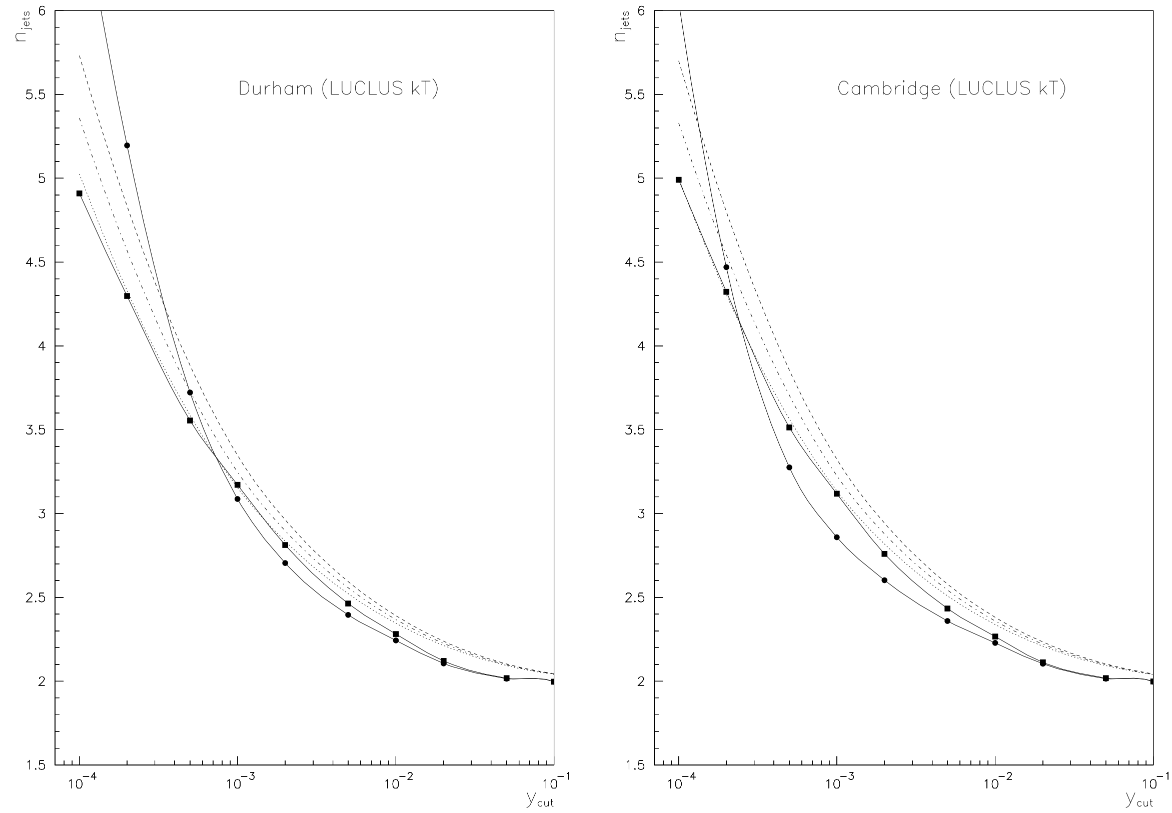

Figure 17: The mean number of jets at $Q = M_Z$ for the Durham (left-hand side plot) and Cambridge (right-hand side plot) algorithms using the Luclus measure. HERWIG predictions: squares, parton level; circles, hadron level. Resummed predictions: dashed, $\alpha_S = 0.126$; dot-dashed, $\alpha_S = 0.120$; dotted, $\alpha_S = 0.114$.

beneficial effects of combining the clustering procedure of the Cambridge scheme with the Luclus measure are evident, in line with various other results previously obtained for this new hybrid scheme.

As a summary of our hadronization studies in the context of multi-jet event rates at LEP1, we can recapitulate the following.

1. There is substantial correspondence between the simple tube model and more sophisticated MC programs. The main kinematic features recognised in the former reflect onto the latter. Thus, the simple hadronization mechanism based on a longitudinal phase space represents a good guidance in order to test in first instance the performances of new (and old) algorithms.

2. After full hadronization is implemented, four among the clustering schemes studied since the beginning appear to have rather contained hadronization corrections at small values of the resolution parameter, say, around 10^{-4} at LEP1 energies, corresponding to partonic multiplicities of five or so. These are the Angular-Ordered Durham, the original Cambridge scheme (i.e., that using the Durham distance) the one using the Luclus measure and, curiously,

the original LUCLUS scheme deprived of the reassignment step and implementing the default $d_{\rm init}$. All other schemes perform significantly worse, particularly the GENEVA one, which appears to be very unstable. DICLUS offers unique advantages in having very small hadronization corrections in the two-jet-dominated region, but then has larger corrections than other algorithms (except GENEVA) at smaller resolution scales.

3. The hadronization corrections come about for several reasons. The string/drag effects usually give a negative contribution at large $y_{\rm cut}$, i.e., results in fewer jets on the hadron than on the parton level, the size of which depends on the details of the algorithm. Fluctuations in the hadronization process, such as charm and bottom decays and junk-jet/misclustering effects, give a positive contribution, that always wins out at small $y_{\rm cut}$. That some algorithms have a smaller net hadronization correction at medium $y_{\rm cut}$ thus in part is the result of a cancellation between opposite effects. From this point of view, our results for the case of the DURHAM and CAMBRIDGE schemes are in general agreement with those presented in ref. [34].

4. Our results are substantially independent of the MC program used, that is, of the model adopted for the hadronization mechanism, and for the QCD cascade. This means that hadronization corrections can be estimated rather reliably for many algorithms. As a by-product of this conclusion, we observe that the angular-ordering procedure recommended in ref. [3] as a refinement of the DURHAM algorithm is then not restricted to the 'angular ordered' emission as implemented in the HERWIG parton shower (the MC program exploited in that reference).

5. A warning should be borne in mind, concerning the comparison of parton and hadron level. It should be recalled that the partonic dynamics implemented in event generators is an approximation of the actual prediction, as the parton cascade only exploits some (logarithmically) enhanced terms of the infrared (soft and collinear) emission. Therefore, there is an intrinsic danger in interpreting the hadron–parton difference as generated by the MCs as a method-independent estimate of non-perturbative effects and simply adding it to the resummed predictions.

6. Specifically, we did not address here the key issue of how to combine the best perturbative QCD predictions, based on resummed contributions matched to fixed-order results, with hadronization corrections. However, in this respect, we have shown that the LUCLUS $p_\perp$ distance measure seems to offer an interesting alternative to the traditional DURHAM one, in the sense that the theory and MC parton levels appear to match better, especially when implemented along with the clustering sequence of the CAMBRIDGE algorithm.

JHEP08(1998)001

4.3. Jet reconstruction

In this section we study various aspects of how well algorithms reconstruct jet directions and energies, as well as a few other related quantities. The results presented here have been obtained with the JETSET program, but all essential features come out very similarly with HERWIG and ARIADNE.

Since some of the studies are based on reconstructing a fixed number of jets, like three or four, it should be noted that the ANGULAR-ORDERED DURHAM and CAMBRIDGE algorithms do not always allow this, and do not necessarily provide a unique answer. To understand these points, first consider simple binary joining algorithms, such as JADE or DURHAM. In these, it is always the cluster pair with smallest distance y_{ij} that are joined next. Starting from n clusters, there is thus a unique sequence of joinings giving $n-1, n-2, n-3, \ldots, 3, 2, 1$ clusters. If one wants to obtain three jets, say, one simply has to perform binary joinings till exactly three clusters remain. For an exercise like that, the y_{cut} value need never even be specified.

In the ANGULAR-ORDERED DURHAM scheme, on the other hand, y_{cut} is used not only to interrupt a sequence of joinings, but also to influence the sequence itself. We remind that the procedure joins the pair with smallest $v_{ij} \equiv 2(1-\cos\theta_{ij})$ among all those with $y_{ij} < y_{\mathrm{cut}}$. Change y_{cut}, and you can change which pair is joined in the step from m to $m-1$ clusters, and in turn all the subsequent joinings. Therefore, the range of y_{cut} values that give three jets, e.g., may split into subranges corresponding to different three-jet configurations. Furthermore, it may be impossible to obtain three clusters, since an infinitesimal y_{cut} change may give a flip from one joining sequence, ending with four clusters, to a completely different one, ending with two. A further consequence of such flips is that the number of clusters need not be a monotonous function of y_{cut}.

The CAMBRIDGE algorithm introduces one further task for y_{cut}, on top of the ANGULAR-ORDERED DURHAM, namely to provide the scale for sterilization/soft-freezing. This increases the fraction of events failing to reconstruct a requested number of jets, but reduces the number of cases with several different three-jet topologies.

One should not exaggerate the problem, however. Typically only for 0.15% of LEP1 events it is impossible to find a three-jet configuration with ANGULAR-ORDERED DURHAM, which increases to 1.3% in CAMBRIDGE. Furthermore, 4.7% give two (or more) different three-cluster configurations for ANGULAR-ORDERED DURHAM and 0.9% for CAMBRIDGE. The numbers are somewhat higher for four-jets. None of the other algorithms failed to reconstruct the requested number of jets, nor have any ambiguity in which jets are reconstructed. In the studies below, events which failed to reconstruct are not considered at all, while the choice among alternative three-cluster configurations is simply based on which is found first. (In a search procedure that involves a measure of randomness, so there should be no special bias.[9])

[9]In ref. [34] a special-purpose algorithm was devised in order to determine the y_{cut} transition

Algorithm	Two-jet		Three-jet		
	$\langle\Delta\theta\rangle$ (°)	$\langle(p_z)_{\mathrm{back}}\rangle$ (GeV)	$\langle\Delta\theta\rangle$ (°)	$\sigma(\Delta E)$ (GeV)	$\langle\Delta\theta_{\mathrm{min}}\rangle$ (°)
JADE	3.22	0.33	4.01	2.74	−2.4
DURHAM	3.09	0.11	3.91	2.41	−2.4
DURHAM/LU	3.14	0.19	3.86	2.48	−3.0
GENEVA	3.05	0.04	4.01	2.45	−2.7
ANGULAR-ORDERED DURHAM	3.09	0.10	3.81	2.25	−1.9
CAMBRIDGE	3.09	0.10	3.88	2.28	−2.4
LUCLUS	3.06	0.00	3.52	2.02	−2.3
DICLUS 1	3.66	0.00	4.43	1.99	−0.3
DICLUS 2	3.56	0.00	4.23	1.93	−0.3
DICLUS 2 reclustered	3.07	0.00	3.65	2.16	−2.3
thrust	3.23	—	—	—	—
sphericity	4.36	—	—	—	—

Table 5: Average angular and momentum/energy error on jet reconstruction in two- and three-jet events at LEP1; see text for further details. JETSET results.

The parton shower starts out from a back-to-back $q\overline{q}$ pair. The observable event axis is smeared by the parton shower and hadronization, but one interesting measure is how well the original axis can be reconstructed. Thus clustering algorithms are requested to find two jets; alternatively measures such as thrust and sphericity can be used here. In the first result column of Tab. 5 it is shown that most algorithms do comparably well, including thrust, while DICLUS without reclustering does worse and sphericity has the largest error. In three-jet events there is no 'correct' answer, and so here the comparison is based on matching the jets clustered on the parton level with those obtained on the hadron level. Each event thus gives three angles. Only events with $0.85 < T < 0.95$ on the parton level have been used to produce the numbers. Again DICLUS without reclustering does worse, third column of Tab. 5, though less dramatically so than above, while LUCLUS does somewhat better than the others. The same pattern holds for four-jets (not shown).

The error on the jet axis reconstruction need not be entirely of a statistical character, however. In subsection 4.2 above, we have mentioned the string/drag effect as the reason for the 'negative hadronization corrections'. In a three-jet event, normally the smallest angle between two jets, θ_{min}, would be formed by the gluon jet and a quark/antiquark jet. These are connected by a dipole and thus should be 'pulled closer' by the hadronization. The last column in Tab. 5 shows the average $\Delta\theta_{\mathrm{min}}$, the

values at which an event flips from an n-jet to an m-jet configuration, with m and n not necessarily consecutive. It was used to study the characteristics of those events that have two different n-jet configurations.

JHEP08(1998)001

Algorithm	$\langle\Delta\theta\rangle$ (°)	$\langle\theta_{\min}\rangle$ (°)	$\sigma(\Delta E)$ (GeV)	$\langle\Delta E_q\rangle$ (GeV)	$\langle\Delta E_{\overline{q}}\rangle$ (GeV)	$\langle\Delta E_g\rangle$ (GeV)
JADE	5.70	64.9	1.34	−0.16	0.47	−0.39
DURHAM	5.68	64.8	1.33	−0.03	0.60	−0.58
DURHAM/LU	5.67	64.7	1.30	−0.08	0.46	−0.39
GENEVA	5.74	64.4	1.44	0.25	0.73	−0.97
ANGULAR-ORDERED DURHAM	5.60	65.0	1.29	−0.01	0.60	−0.61
CAMBRIDGE	5.82	63.6	1.43	0.16	0.66	−0.85
LUCLUS	5.34	65.0	1.11	−0.05	0.57	−0.53
DICLUS 1	5.67	72.0	1.06	−0.06	0.71	−0.65
DICLUS 2	5.38	71.6	1.03	−0.04	0.69	−0.65
DICLUS 2 reclustered	5.13	66.0	0.96	−0.32	0.59	−0.27

Table 6: Average angular and energy error on jet reconstruction in three-parton events at 30 GeV. All events are $u\overline{u}g$, to avoid contamination from heavy-flavour effects, and the three-jet kinematics is fixed by $x_q = 0.9$, $x_{\overline{q}} = x_g = 0.55$; hence $\theta_{\min} = 70.2°$ on the parton level. The last three columns give the difference between the hadron- and parton-level numbers. See text for further details. JETSET results.

difference between the hadron- and parton-level $\theta_{\min}$ values. We see that indeed there is the expected systematic bias in all algorithms, although markedly smaller in DICLUS than in the others. DICLUS is the only algorithm intended to correctly account for dipole effects in the hadronization, and is thus seen to achieve this purpose. To set the scale of the effect, the width of the $\Delta\theta_{\min}$ distribution is about 10° in all algorithms, so the systematic bias is still significantly smaller than the event-to-event fluctuations. Also the smallest angle in four-jet analyses show a similar pattern, with DICLUS the only one to be almost bias-free. We note that if the sum of the momenta of the particles assigned to each jet in DICLUS are allowed to redefine the jet directions, the bias returns and this 'reclustered' DICLUS behaves more or less like the standard binary algorithms.

In a study of fixed three-parton configurations at lower energies, where then the parton-level was known to have a fixed smallest angle of about 70°, most algorithms reconstruct an angle around 65°, while the DICLUS average is around 72°, second results column of Tab. 6. Thus, while still doing best, there are indications that DICLUS at times overcompensates for the string/drag effect.

Also the jet energy reconstruction can be compared between the hadron and parton level, fourth column of Tab. 5 and third column of Tab. 6 give the width of the jet energy difference distribution. Here DICLUS and LUCLUS perform better than any of the others. The tendency for systematic bias can be studied in fixed three-parton configurations, last three columns of Tab. 6. The energy of the most energetic jet is usually reconstructed without much bias, whereas there is a tendency in all algorithms for the other quark to gain energy from the gluon, reflecting the fact that gluon jets

JHEP08(1998)001

are softer and broader and thus easily lose particles to the other jets, especially the most nearby one. This systematic bias is largest in GENEVA, as could be expected from the way GENEVA favours clustering around energetic particles. CAMBRIDGE shows the second largest bias, and the reclustered DICLUS the smallest.

From a practical point of view, a jet is a collection of 'nearby' particles, where 'nearby' obviously is a very subjective criterion. One measure is how far out in angle a jet extends from its core. For instance, enforce the reconstruction of exactly two jets in all LEP1 events. With two back-to-back jet axes, one may expect an optimal subdivision of particles to be by hemisphere, so that no particle is found more than $90°$ from its jet axis. Figure 18 shows the angle for the particle furthest away from its assigned jet, on a per-event basis. It is seen that only LUCLUS and DICLUS respects the $90°$ criterion. The reason is that the standard distance measures allow two soft particles to be joined, also when they are somewhat away in angle. In a normal binary joining scheme, they will thereafter together enter into one of the final jets, even if one of them is much closer to another jet. The reassignment step of LUCLUS is specially devised to overcome this limitation, i.e., to reevaluate prior joining decisions in the light of the joinings that have been performed since.

Among the other algorithms, JADE is most likely to have a particle in the 'wrong' hemisphere. In fact, the JADE E scheme, using the true mass as distance measure, is the very worst of the algorithms studied. This is the well-known instability problem, already mentioned. DURHAM and the other $p_\perp$-based algorithms are better, but note that it is important that the angular dependence is $2(1-\cos\theta_{ij})$ rather than the correct $\sin^2\theta_{ij}$, or else two back-to-back particles would have $p_\perp = 0$ and be joined. DURHAM with the LUCLUS measure is slightly worse than normal LUCLUS, since the clustering of two soft particles is somewhat more favoured than in standard DURHAM. The ANGULAR-ORDERED DURHAM and CAMBRIDGE schemes offer no visible improvement. GENEVA is the pure binary joining algorithm with best performance, reflecting that clustering of two soft particles is disfavored.

The example in Fig. 18 is somewhat academic, in that a subdivision of an event into two hemispheres could be performed much easier by cutting the event transverse to the thrust or sphericity axis. However, it provides a clean illustration of a phenomenon that is equally present when three or more jets are reconstructed. LUCLUS always gives much narrower jets than any of the other algorithms, and JADE gives the broadest ones. The one notable change is that for three-jets, GENEVA no longer gives narrower jets than DURHAM and its relatives, probably indicating how the GENEVA distance measure allows an energetic jet to pick up particles also fairly close to another softer jet. While the wide-angle tracks are very important for the visual impression, they normally carry little momentum. The second column in table 5 shows $(p_z)_{\text{back}}$, the average amount of longitudinal momentum carried by particles moving 'backwards' with respect to their jet axis. Typically this number is only 0.1 GeV per event, rising to 0.3 GeV for JADE and 0.6 GeV for JADE E.

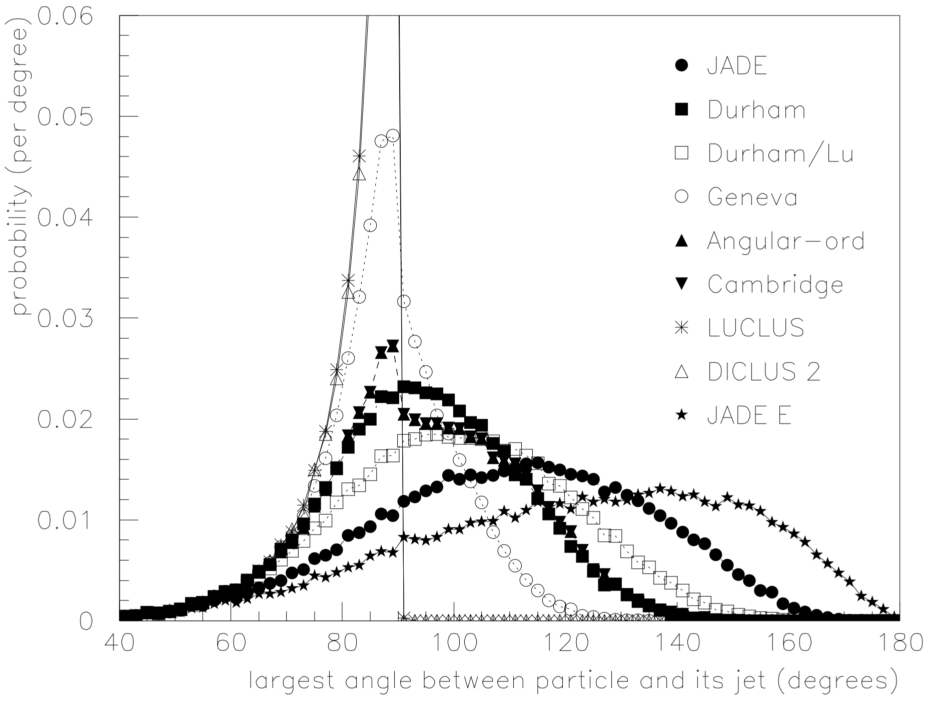

Figure 18: Largest angle, per event, between a particle and the jet it is assigned to, in LEP1 events when clustered to two jets. For clarity the y axis has been truncated; the LUCLUS curve goes up to 0.13, and DICLUS to 0.15 in the last bin before 90°. JETSET results.

Another alternative measure for the narrowness of jets is offered by the sum of the invariant jet masses. This is studied in Fig. 19. Since the y_{cut} definition is scheme dependent, results are plotted as a function of the average number of jets at the y_{cut} values studied, as in subsection 4.2. Here the JADE algorithm indeed does best, in line with its distance measure being intended to minimize jet masses. LUCLUS and DURHAM with the LUCLUS measure come next, i.e., here reassignment is not important. GENEVA does markedly worse than other schemes.

A third measure is the summed transverse momentum of all particles in an event, relative to their respective jet axis. This is shown, again as a function of the average number of jets, in Fig. 19. The difference between the $p_{\perp}$-based algorithms is here small, while JADE and reclustered DICLUS are somewhat worse and GENEVA together with the other DICLUS modes are the worst. The large difference between the standard DICLUS and the reclustered one is due to that DICLUS jets typically are asymmetric with most of the particles lying on one side of the jet direction. The reclustering pulls the jet direction more to the center of the particles assigned to it, thus reducing the summed $p_{\perp}$.

In summary, we draw the following conclusions from the analysis of two-, three-, and four-jet events (also supported by some studies not shown):

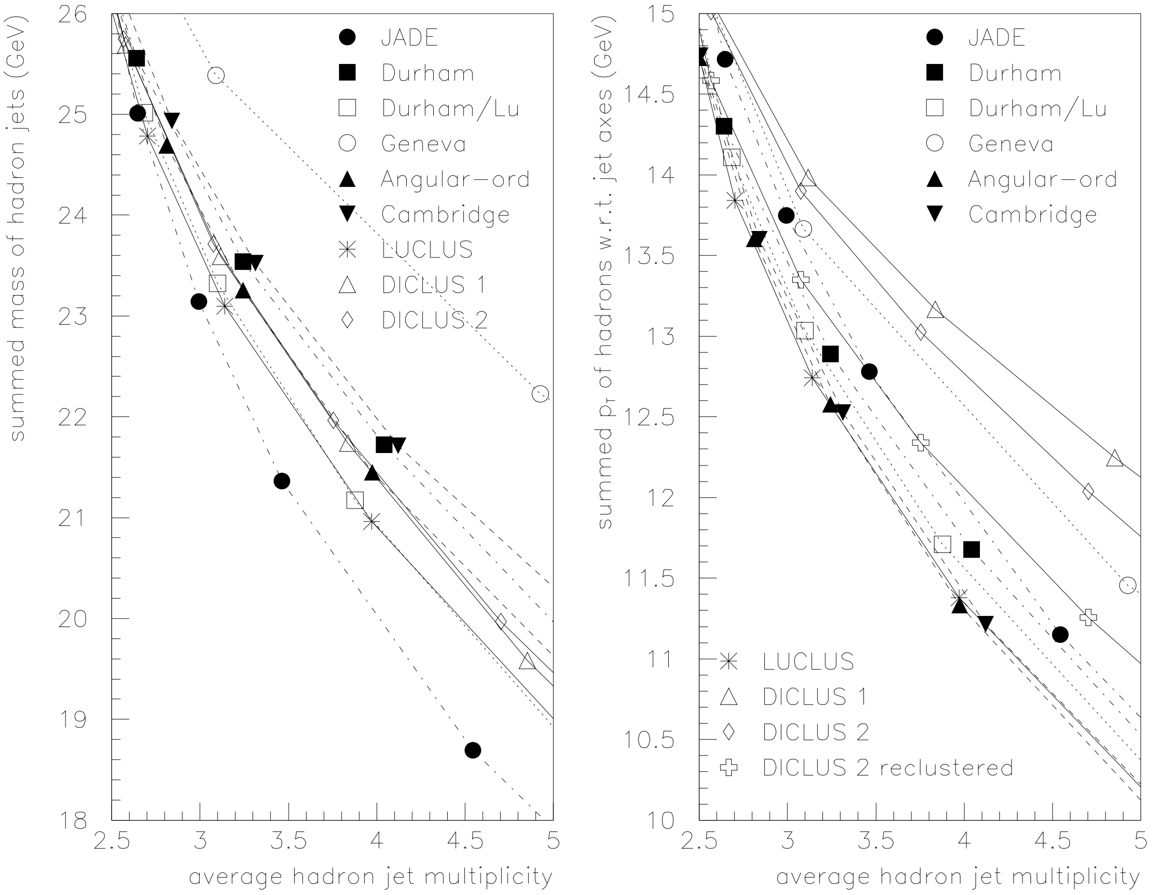

Figure 19: Sum of jet masses (left) and of particle $p_\perp$ as a function of the average number of jets (i.e., implicitly as a function of y_{cut}) for LEP1 events. Note that the y axis does not start at 0, i.e., differences appear exaggerated. JETSET results.

1. The DICLUS algorithm generally does best (among the algorithms studied) in jet energy reconstruction, and also successfully addresses the issue of a systematic hadronization bias in the opening angle between two nearby jets, caused by the string/drag effect. The price to be paid is that the average error on the individual jet direction is larger than in other algorithms. Reclustering the jets from DICLUS makes it behave more like the standard binary algorithms. We also note that using the transverse mass in eq. (2.9) as measure (mode 2) is somewhat favoured as compared to using the one in eq. (2.8) (mode 1).

2. LUCLUS does almost as well as DICLUS in jet energy reconstruction, and best in jet angles. (A similar conclusion was reached in ref. [4].) The reassignment step means it is the only algorithm that does not have stray particles in a jet that are visibly much closer to another jet. Since the stray particles normally carry rather small momenta, the impact of reassignment on momentum-weighted quantities should not be overstressed, however.

3. The Angular-Ordered Durham and Cambridge algorithms here offer no significant advantages over the basic Durham scheme, nor does a use of the Luclus distance measure. All these algorithms therefore share a common 'average' level of performance.

4. Jade fulfills the intended task of reconstructing small cluster masses, but at the price of a larger rate of large-angle stray particles.

5. Geneva does better than the average in some quantities, and significantly worse in others. Its distance measure means that the jet energy determinations show larger systematic biases than with any of the other measures used.

4.4. W mass reconstruction

Above we have studied jet finding in quite general terms. For an intended application, special further studies may be necessary. The criteria for a good algorithm are going to be different in the determination of an α_{s} value and in the study of angular distributions as a test of the three-gluon vertex, to give but two examples. Currently, the W mass determination at LEP2 is another such topic of large interest [66], representing different optimization criteria than the ones illustrated above. We here focus on the hadronic production channel, where $e^+e^- \to W^+W^- \to q\overline{q}Q\overline{Q}$. Thus the signal is the presence of four jets, where the two jet pairs ought to have a mass around $m_W \approx 80$ GeV. There are several complications. Backgrounds exist, both from the four 'wrong' jet pairs in the same event as the two 'right', and from other processes such as the QCD four-jets $e^+e^- \to \gamma^*/Z^* \to q\overline{q}gg, q\overline{q}Q\overline{Q}$. The mass distribution is smeared by the intrinsic W width $\Gamma_W \approx 2$ GeV in combination with the production matrix element itself, by initial-state QED radiation, by neutrinos that escape without detection, by cracks in the detector acceptance, by measurement errors on particle four-momenta and, of course, by misassignments in the clustering procedure. A full study can therefore only be carried out within the context of a complete detector simulation, which is rather beyond the scope of the current report. To illustrate some of the clustering issues we have carried out a rather more modest exercise.

Hadronic W^+W^- events are generated at 180 GeV CM energy, but none of the background processes are studied. Detectors are assumed perfect, i.e., the correct four-momenta of outgoing particles are used to reconstruct exactly four jets per event, by the respective jet algorithm. (Some small number of times Angular-Ordered Durham and Cambridge fail to find four jets, as explained in subsection 4.3; such events are left out from the statistics of the respective algorithm.) In experimental analyses usually some further cuts are imposed, e.g., on the opening angles between jets and on jet energies. This makes sense, since events where two jets are very close are not reconstructed so well. However, then the retained event sample would differ between clustering algorithms, so we have avoided cuts here. Instead all six jet–jet

JHEP08(1998)001

Algorithm	$\langle\sum\Delta\theta\rangle$ (°)	$\langle\sum\lvert\Delta E\rvert\rangle$ (GeV)	$\langle\delta\rangle$ (GeV)	$\sigma(\delta)$ (GeV)
Jade	41.0	25.3	0.06	2.8
Durham	36.5	21.7	−0.03	2.6
Durham/Lu	37.0	22.5	0.05	2.6
Geneva	46.1	27.3	−0.70	3.4
Angular-Ordered Durham	37.6	22.0	−0.08	2.8
Cambridge	38.2	23.0	−0.13	2.8
Luclus	35.6	19.9	0.01	2.6
Diclus 1	39.3	19.7	−0.58	3.1
Diclus 2	38.8	19.5	−0.57	3.0
Diclus 2 reclustered	35.6	18.8	0.16	2.5

Table 7: Analysis of hadronic W^+W^- events at LEP2. First two columns give angular and energy mismatch between reconstructed jets and original partons. Second two give average and spread between best reconstructed and true average W mass of an event. Pythia results.

masses in all events are found and studied, and the success of an algorithm is reflected in how often it can reconstruct sensible W masses.

Some impression of how good the jet reconstruction is can be gleaned by matching the four jets to the four original partons by minimizing the sum of jet–parton opening angles. The average value of this sum, as well as the sum of deviations in the energies between jets and partons, is given in the first two columns of Tab. 7. It generally agrees with the picture in the previous section: Luclus does good overall, while Diclus does worse with angles unless reclustering is performed. The poor numbers for Geneva are more marked than in previous studies, however.

The true test is in the jet–jet mass spectrum, illustrated in Fig. 20, where one may discern the peaked signal from correct combinations of well reconstructed jets, over a smoother background of mismeasured jets or incorrect jet combinations. The 70–90 GeV mass window has been used to produce Minuit [67] fits for a signal plus background shape. The choice of best fit function is not trivial: the signal Breit-Wigner shape is combined with misassignment errors in a complicated and clustering-algorithm-dependent way. For simplicity we have assumed a Breit-Wigner shape, characterized by a peak height h (given as the number of events per 0.2 GeV mass bin; this is related to the input normalization for MINUIT) position m_W and width Γ_W. The h and Γ_W may be combined to an area A underneath the Breit-Wigner. Normalization is such that 2 should be the maximum, corresponding to two correctly reconstructed jet pairs per event. Since the Breit-Wigner has rather large tails, this ansatz may have a tendency to paint too rosy a picture of how well algorithms do. An alternative would have been a Gaussian fit, where the tails are rather strongly dampened, and the bias would go in the other direction. The qualitative differences between algorithms that

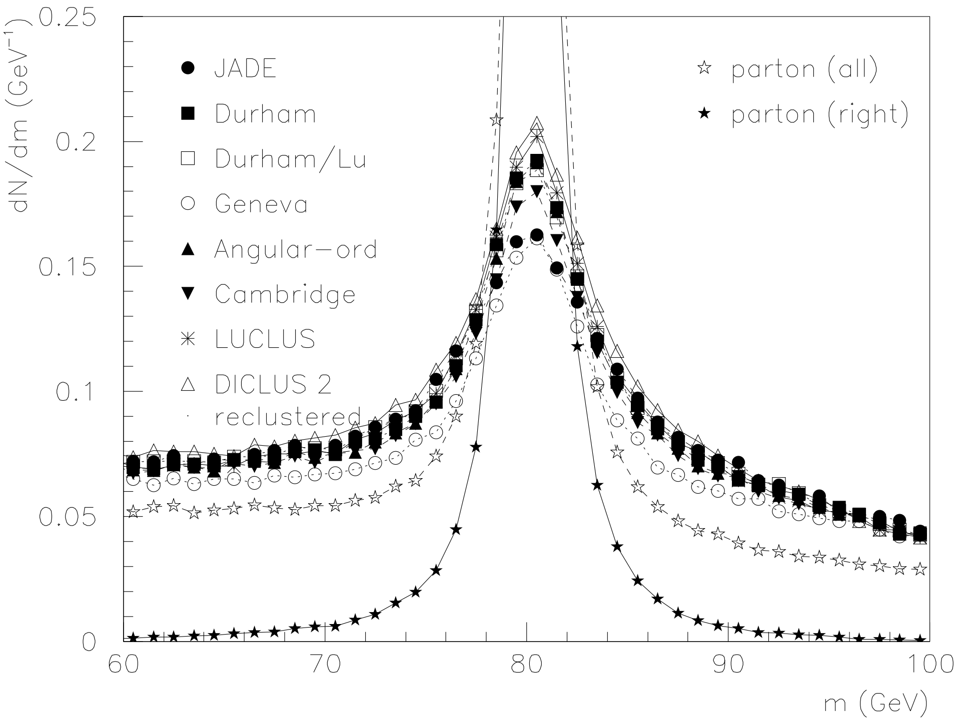

Figure 20: Jet–jet mass spectrum in hadronic W^+W^- events. Each event is reconstructed to four jets and all six jet–jet combinations are included. PYTHIA results.

we report below are the same, however. Two background shapes have been used, one a three-term polynomial in mass and another corresponding to a smeared step function (motivated by the kinematical-limit shoulder at large masses). Results with these two backgrounds come rather close, thus in Tab. 8 PYTHIA numbers are for the former and HERWIG numbers for the latter background. The fits described here correspond to the 'Individual W mass' columns. The 'parton (right)' row makes used of the two correct W masses, and thus represents the best possible answer for algorithms, while 'parton (all)' contains all six possible combinations of the four original partons. The fact that fitted areas above 2 are obtained illustrate imperfections in the fitting ansatz.

Comparing algorithms, several aspects should be kept in mind. A larger area A implies a larger efficiency for sensible jet finding, i.e., fewer misassignments that completely kill the signal. A smaller width Γ_W is a sign of good performance for those jet pairs that are still correctly combined, i.e., fewer misassignments of a less disastrous character. For good m_W determination in an experiment one should thus have both a large A and a small Γ_W. As a third criterion one could imagine the systematic offset between the reconstructed W mass and the parton-level one. However, so long as such an offset is not too large and can be well modelled, it is not so important. One anyway has to make other corrections, e.g., the input m_W parameter does not coincide with the average generated m_W because of the convolution with matrix-element and phase-

Algorithm	Individual W mass				Average W mass			
	h	m_W (GeV)	Γ_W (GeV)	A	h	m_W (GeV)	Γ_W (GeV)	A
	Pythia results							
Jade	200	80.321	8.000	1.26	163	80.719	3.697	0.47
Durham	260	80.354	5.663	1.16	195	80.482	3.296	0.51
Durham/Lu	247	80.333	5.800	1.12	200	80.537	3.288	0.52
Geneva	226	80.376	6.206	1.10	190	80.180	3.237	0.48
Ang-Ord Durham	260	80.396	5.721	1.17	227	80.454	3.216	0.57
Cambridge	238	80.376	5.871	1.10	240	80.396	3.249	0.61
Luclus	268	80.387	5.447	1.14	190	80.492	3.395	0.51
Diclus 1	182	80.008	6.883	0.98	136	79.671	3.923	0.42
Diclus 2	188	80.044	6.732	0.99	138	79.657	3.758	0.41
Diclus 2 reclustered	270	80.486	5.709	1.21	184	80.615	3.375	0.49
parton (all)	1267	80.320	2.088	2.08	656	80.329	2.118	1.09
parton (right)	1270	80.324	2.076	2.07	661	80.325	2.053	1.07
	HERWIG results							
Jade	235	80.218	6.553	1.212	220	80.491	3.337	0.576
Durham	315	80.326	4.893	1.211	295	80.268	2.999	0.694
Durham/Lu	299	80.310	5.203	1.222	238	80.293	3.216	0.601
Geneva	260	80.359	5.287	1.081	180	80.010	3.248	0.460
Ang-Ord Durham	311	80.345	4.944	1.209	320	80.284	2.863	0.719
Cambridge	281	80.376	5.288	1.168	319	80.252	2.895	0.725
Cambridge/Lu	280	80.387	5.261	1.155	299	80.250	2.879	0.677
Luclus	324	80.368	5.247	1.335	241	80.291	3.212	0.608
Luclus (no pre)	324	80.371	5.249	1.334	239	80.286	3.200	0.601
Luclus (no reas)	177	79.984	10.000	1.392	236	80.602	3.383	0.626
Diclus 0	193	79.920	9.904	1.498	179	79.494	4.600	0.646
Diclus 1	244	79.896	7.521	1.440	271	79.561	3.528	0.750
parton (all)	1312	80.422	1.995	2.056	680	80.408	1.986	1.060
parton (right)	1319	80.427	1.988	2.059	694	80.418	1.911	1.041

Table 8: Fits to the W mass spectrum in hadronic W^+W^- events at LEP2. First four columns for the each of the two W's in an event, last four for the average W mass of an event. h is peak height (normalization based on event sample used), m_W and Γ_W fitted W mass and width of a Breit-Wigner shape, and area A the number of combinations per event under the fitted Breit-Wigner. First part Pythia results fitted with a polynomial background, second part HERWIG results fitted with a smeared step background. Note that Pythia and HERWIG use different input masses and widths; the last row for each program sets the standard of optimal performance.

space factors. Unfortunately, while Pythia and HERWIG results largely agree, there are some discrepancies that we do not fully understand, and that thus should act as a warning not to take these studies as the definite word.

One possible conclusion from the numbers in Tab. 8 is that many of the algorithms perform comparably well. In particular, the correlation between sophistication and performance is weak or non-existent, moving, e.g., from DURHAM to ANGULAR-ORDERED DURHAM to CAMBRIDGE. It appears that LUCLUS consistently reconstructs the largest area, i.e., does fewest severe misassignments, but has a rather standard peak width. The difference between DURHAM and LUCLUS $p_\perp$ measures is small; if anything the latter gives a wider peak and thus is worse. Whether preclustering is performed or not in LUCLUS is irrelevant so long as reassignment is allowed, but without reassignment the preclustering is disastrous — the peak is so broadened that Γ_W hits the upper bound allowed in the fit. Thus, to the extent that LUCLUS does somewhat better than DURHAM, the reason is the reassignment step. The DICLUS fits without reclustering give problems with the Γ_W or A values, but also displays a large systematic bias in the estimated m_W. With reclustering, DICLUS again does fairly well. One reason for the problems could be that DICLUS is designed with QCD events in mind, where two nearby partons are connected by a color dipole. Here two nearby jets would come from different W's and not share a dipole (we did not include the possibility of color rearrangement [31, 68]).

In experiments, it is advantageous to study the average W mass of an event rather than the two individual ones. There are several reasons for this, but of interest here is that misassignments of particles in part cancel, in that a reassignment of one particle from one W to the other reduces the first mass and increases the second, leaving the average less affected than each separately. Per event there are thus three possible jet pairings, each giving one potential average W mass. Of these three, we exclude the one where the two most energetic jets are paired with each other, since kinematically this is seldom the right combination. The remaining two combinations give mass distributions as illustrated in Fig. 21. Note that indeed the signal peak is much more narrow, and that there now is an absolute kinematic limit at 90 GeV. MINUIT fits have been performed, as before, with results as shown in the 'Average W mass' columns of Tab. 8. Normalization is such that an ideal fit would give $A = 1$.

It is notable that the relative performance of algorithms changes rather drastically compared with above. The two best ones now are CAMBRIDGE (the original one, not the one employing the LUCLUS measure, which is not shown in the plots) and ANGULAR-ORDERED DURHAM, whereas LUCLUS falls below the average. This could indicate that the particles that get misassigned are somewhat different in the former two and in the latter algorithm. That is, in the former two, the errors on the two individual W masses tend to cancel better in the average. DICLUS still gives a larger Γ_W than other algorithms. GENEVA has a reasonable width but a small area A.

The right two columns of Tab. 7 shows that the pattern between models is not so easy to understand. Here the average mass is evaluated for all three possible jet pairings and compared with the correct average W mass of the same event. The pairing which agrees best is retained, and δ denotes the average mass difference between the

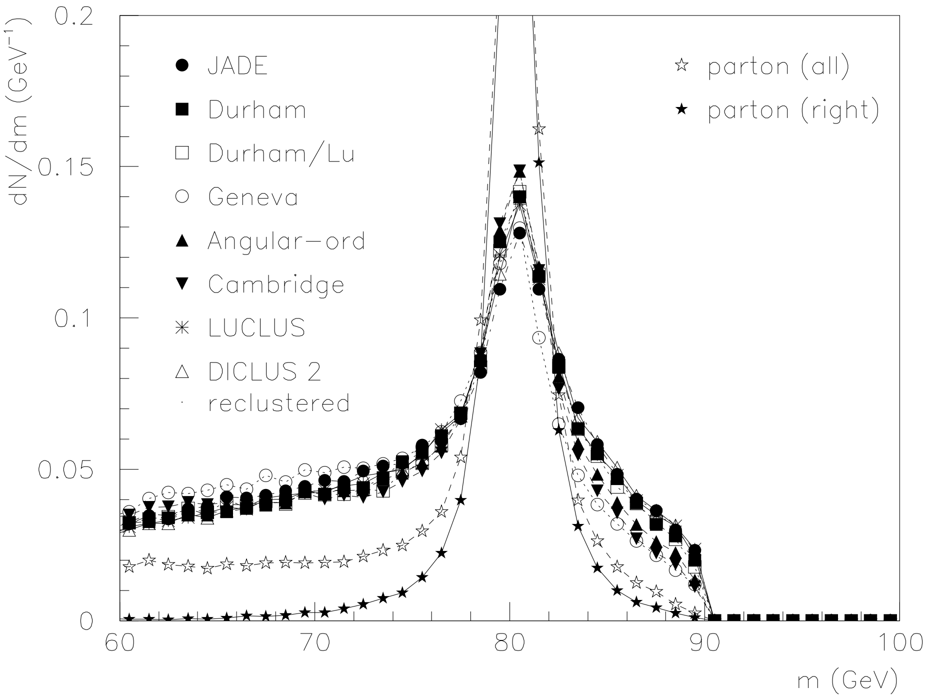

Figure 21: Average jet–jet mass in hadronic W^+W^- events. Each event is reconstructed to four jets, which can be paired three different ways. The pairing with largest energy difference between the pairs is omitted and the average masses of the other two pairings are plotted. PYTHIA results.

reconstructed and the true average W mass. As we see, GENEVA and DICLUS without reclustering here show a significant bias in the negative direction, and also give a larger width σ of the δ distribution. LUCLUS does quite well in these 'behind-the-scene' numbers, so the poor LUCLUS numbers above do not seem to have a simple explanation.

As possible conclusions for the W^+W^- analysis, we attempt the following.

1. The choice of the algorithm is in general not so trivial in such a context. However, there are some that cannot be recommended, notably DICLUS without reclustering and GENEVA, and also JADE.

2. If the 'Individual W mass' distribution is preferred in the selection procedure, that LUCLUS performs slightly better than the other algorithms.

3. If, instead, one resorts to the 'Average W mass' spectrum, then ANGULAR-ORDERED DURHAM and the original CAMBRIDGE (i.e., that with the DURHAM measure) come out best.

4. However, differences between the three algorithms that excel do not show a simple pattern, so that, in the end, a definite decision between these latter could probably only be made in the context of some specific detector simulation and mass extraction procedure.

4.5. Speed

All binary clustering algorithms are comparably fast. Starting from an initial configuration of a large number n of partons and hadrons, a small number of jets is to be found. Therefore $O(n)$ binary joinings have to be performed. Each in principle requires $O(n^2)$ distances to be evaluated to find the smallest one. In practice, distances can be kept in a table that is only updated for those entries affected by a binary joining. Therefore execution time scales more like $O(n^2)$ than the expected $O(n^3)$. At LEP1 energies, the clustering time is about two thirds of the time it takes to generate an event (with Jetset and HERWIG).

The Luclus preclustering time roughly scales like $O(n^2)$: each particle can be the seed of a precluster and all particles have to be tested whether they belong to the precluster. If m preclusters are formed, normally with m much smaller than n, then subsequent joinings take $O(m^3)$, since the reassignment steps means one cannot reuse older numbers. The reassignment step after each joining requires assigning n particles to m clusters, i.e., a total $O(m^2 n)$ for $O(m)$ joinings. In practice, scaling of the total time is about like $O(n^2)$. At LEP1 the algorithm is somewhat faster than the binary joining ones, but at most by a factor of two. If the trick of pretabulation is not used in the binary routines, the difference is more like a factor five.

The basic step of the Diclus algorithm is the joining of three clusters into two. Therefore $O(n^3)$ distances have to be evaluated to find the smallest one. Again, keeping a table of distances allows the total time to scale more like $O(n^3)$ than like the $O(n^4)$ that might have been expected. Still, there is a significant price to be paid, and at LEP1 energies Diclus is about a factor fifty slower than the other algorithms.

5. Summary and conclusions

Jet clustering algorithms are an expression of *time* and *place*. The time evolves with the calculational methods developed and these can in turn be limited by the computing power available. The place is circumscribed by the experimental contexts where algorithms are needed and the tasks that they are asked to accomplish. In ten, fifteen years from now, the two will both have changed. Specifically, if the theoretical methods adopted (e.g., in determining the higher-order and exponentiation properties of pQCD, the parton-shower evolution and/or the non-perturbative dynamics of hadronization) in some years time would be different, so should clustering algorithms be.

Inevitably then, our study can claim no prerogative to being definitive. We have undertaken it for the present era and for the imminent phenomenology. The aim was to survey the many and different jet finding algorithms for electron-positron events available on the market nowadays and study which algorithm to use where, if at all possible. As anticipated in the Introduction, we have not found *one* single best choice that prevails in *all* cases we have addressed. However, as the reader should have agreed

JHEP08(1998)001

upon by now, there need not exist such a one. Nonetheless, in several instances it has been possible to recognize, if not the *most suitable* algorithm to use, at least the *attractiveness* of some of its basic components. In this section, we attempt to summarize our findings.

As a first example we have considered the realm of pQCD, by studying jet fractions at parton level and resorting to the most advanced techniques of perturbation theory: that is, exact next-to-leading fixed order results combined with resummed predictions to next-to-leading-logarithmic accuracy. Such a choice was not made by chance, as it was dictated by the crucial rôle that jet rates play, e.g., in the determination of the strong coupling constant α_S and of its running with energy. In this respect, we should point out that the main features illustrated in this paper for the case of LEP1 energies, survive unaltered for the case of LEP2.

By studying the three-jet fraction in pQCD we have taken for granted the well-established result that a jet measure based on some relative transverse momentum of the clusters involved is the most appropriate to use, thus neglecting consideration of jet finders based on other quantities (such as the invariant mass). Under these circumstances, one historically recognizes three different such measures. Namely, the so-called LUCLUS, DURHAM and DICLUS ones. The first two cluster two particles into one whereas the last one merges three into two. Neglecting imperceptible differences (we used 'massless' partons) between energy and momentum, they can geometrically be viewed as follows. The first represents the transverse momentum of either particles with respect to the sum of the momenta of the two. The second is the transverse momentum of the lower-energy cluster with respect to the higher-energy one. The third is the transverse momentum of one cluster with respect to the other two.

Among the three our preference would go to the LUCLUS measure. In fact, algorithms based on the latter display a reduced (renormalization) scale dependence of the three-jet fraction at NLO, as compared to the cases of the DURHAM and DICLUS expressions. The stability of the perturbative results in higher order against variations of such a scale is a measure of the smallness of even higher terms in the perturbative expansion, this ultimately reflecting a better degree of convergence of the corresponding power series. As α_S measurements are unavoidably biased by a theoretical error, and since this is assessed in no other way than the range in α_S spanned by the QCD predictions for different choices of the above scale, in our opinion, the LUCLUS measure comes to be a recommendable choice in this context. We have hypothized its improved behaviour, with respect to the DURHAM one, as due to their respective definitions: whereas the former is a smooth function of the energies of the two clusters the latter is not. As a matter of fact, the presence of edges at the border of the phase space of an observable has recently been advocated to act as a source of misbehaviors in higher order perturbation theory.

An additional neat attribute of the LUCLUS transverse momentum appears while combining the fixed-order with the resummed perturbative predictions, for example in computing the average number of jets produced in electron-positron annihilation events. Such a quantity can be predicted reliably from QCD over a wide range in y_{cut}. The advantage of using the LUCLUS measure in this case is that the parton level of the theory matches more naturally the one produced by the Monte Carlo generator, as no rescaling of α_{S} is needed to find an adequate agreement between the two and since their asymptotic behaviour tend to be the same, regardless of the α_{S} value itself (contrary to the case of the DURHAM measure). This is particularly true at low y_{cut} values, where one needs a tight control on the differences between the two implementations, in order to profitably explore the dynamics of the QCD phase transition.

The difference between the parton level and the hadron level as generated by a phenomenological Monte Carlo program is customarily used as an estimate of the hadronization corrections. However, one should notice that even in presence of a good agreement between exact parton level from the theory and the approximate one from the Monte Carlo, there is a danger in interpreting the hadron-parton difference in the phenomenological generator as an estimate of non-perturbative effects and simply adding it to the matched prediction. In fact, the presence of unnatural cut-offs and kinematic boundaries in the parton shower could well induce non-perturbative contributions already at the parton level. Thus, we have refrained here from doing so. Instead, we have compared the partonic and the hadronic outputs as they come from the generator, without any attempts to correct the former.

The non-perturbative hadronization is clearly a genuine physics process, but for it we do not have at present a well established theory. Rather, our knowledge is based on the phenomenological experience and is implemented in the above-mentioned programs. Although the agreement between the latter is remarkable, and these in turn reproduce well real data, there are systematic dissimilarities in their implementation of the non-pQCD dynamics that must be accounted for. In other terms, the differences in the predictions of the Monte Carlo programs contribute to build up our systematic uncertainties on the actual measurements. These so-called hadronization corrections turn out to be algorithm dependent, thus to design one for which these are noticeably reduced would represent a clear improvement: the smaller those are, the more under control would the differences between generators be. This is of particular relevance at very small values of the resolution parameter y_{cut}, where the interface between perturbative and non-perturbative QCD occurs.

In order to to reduce the size of the non-perturbative corrections in multi-jet rates, the implementation of the angular-ordering and soft-freezing procedures has proven to be decisive, particularly at low y_{cut}. The first one consists in distinguishing between the variable used to decide which pair of objects to test first and that to be compared with the resolution parameter. The second one corresponds to eliminating from the sequence of clustering the less energetic one in a resolved pair of particles. These two steps help to heal two of the unwanted phenomena occurring in the dominion of soft physics, that is, 'junk-jet' formation and 'misclustering', respectively. The first takes place because of the tendency of soft 'unresolved' particles of acquiring momenta from particles at low transverse momentum and forming spurious jets from these whereas the second happens because of the bias of soft 'resolved' particles of attracting wide-angle radiation.

These two remedies are however effective *only* if inserted into $p_\perp$-based jet finders. In fact, although these two steps were originally implemented as part of the CAMBRIDGE algorithm, we have assessed their efficiency also in presence of the LUCLUS measure while reminding the reader of their inadequacy if the JADE one is used instead. If one then combines this result with what we have already mentioned for the fixed-order and resummed predictions, it is evident that the hybrid scheme that we had originally introduced for purpose of comparison, based on the LUCLUS transverse momentum and the CAMBRIDGE clustering sequence, performs better than any other tested, so to deserve the status of new algorithm. In our opinion, it has come to set the standard as far as the dominion of soft physics in multi-jet events is concerned.

Before proceeding further, we should mention that the overall features obtained with respect to the size of the hadronization corrections are in part the result of the fortuitous cancellations between opposite tendencies. On the one hand, junk-jet formation and misclustering (and heavy quark decays as well) induce positive corrections. On the other hand, the well-known string or drag effect (i.e., the pulling closer of the two nearest jet directions by the hadronization mechanism) produces negative contributions. The increased size of the 'negative hadronization corrections' for some algorithms at medium values of $y_{\rm cut}$ is then the consequence of having reduced the former while leaving untouched the latter effect. Therefore, as $y_{\rm cut}$ grows larger, to diminish the extent of the corrections becomes more and more matter of finding a delicate balance between the two. At the upper extreme of the $y_{\rm cut}$ range, that is, in the two-jet limit, the DICLUS algorithm admirably contains the size of the hadronization effects, whereas JADE is clearly the best in the intermediate 3-jet region.

If one abandons the subject of QCD studies in multi-jets, that is, the dominion of soft physics and global quantities (such as jet fractions, shape variables, etc.) and enter, for example, the territory of the search for mass resonances, the criteria that define a good algorithm are going to be rather different. In the new context, as it is now for the mass determination of the W boson at LEP2, kinematical quantities such as energies and angles (which build up the definition of invariant mass) are of main concern. Also in this case, although we have not carried out a sophisticated analysis of four-jet events at LEP2, including detector effect and background simulations, we believe to have achieved interesting results.

In hadronic decays of W^+W^- pairs, the four partons emerging from the unstable resonances are naturally energetic and far apart. QCD radiation from the two W decays does not interfere till the next-to-next-to-leading order in the strong coupling constant. In other terms, the soft dynamics that determines to a large extent the phenomenology of jet rates is of little concern here. Instead, in this case, it is how well an algorithm is able to reconstruct at hadron level the original partonic energy and direction, and ultimately the shape of the mass resonance, that sets the target of a good jet clustering performance.

Therefore, the next step of our analysis has been to quantify the ability of the various clustering algorithms in minimizing the average angular and energy error in the jet reconstruction. As a preliminary exercise, to allow for an understanding of the typical biases, we have addressed the simplified case of the kinematics of two-, three- and four-parton events, for some fixed phase space configuration. The procedure has been eventually generalized to include all final state jet multiplicities, by studying the sum of the invariant jet masses as well as of the transverse momentum of all particles of an event.

After these tests, two out of our list of clustering algorithms excel above all others, which share an ordinary degree of performance. They are the Luclus and Diclus schemes. The former is undoubtedly the best in reconstructing angles and it is second in case of energy only to the latter, which is however very modest with angular quantities. The ability of Luclus in reconstructing angles and energies can be attributed to the reassignment procedure, which it is the only to implement. In other schemes, it is not uncommon to find stray particles at the edge of a jet that, by any distance criterion, are closer to another jet. The poor performances of Diclus in angles are the price paid for an implementation especially designed to remedy systematic biases in the hadronization, notably the mentioned string or drag effect.

Studies in energies and angles similar to those above have been carried out also for the case of W^+W^- into four-jet events at LEP2. The general picture for these two quantities separately is similar to that outlined above, with Luclus best overall. One would then expect this algorithm to come first also when energies and angles are combined to reconstruct the W mass invariant spectrum. This is however true only if one plots in the corresponding histogram *all* individual jet-jet masses (six in total). The majority of Luclus events are in fact concentrated around the W mass, whereas misassignments take place more often for other algorithms, whose spectra can be significantly more spread out.

Surprisingly enough, if one plots instead the mass distribution formed from the two average masses which can be obtained from the two possible pairings that most likely reconstruct better the W mass (those in which the two most energetic jets are not paired together) then the original Cambridge algorithm (the one employing the Durham measure) comes out best (ahead of the Angular-Ordered Durham). The reasons for this are not entirely understood. On the one hand, the use of the 'average masses' rather then the 'individual masses' is generally dictated by the fact that misassignments of particles partially cancel, on the other hand, our studies of jet angle and energy reconstruction did not furnish us with an obvious explanation why angular-ordering and/or soft-freezing should be beneficial to the four-jet decays of W^+W^- pairs. (In addition, notice that in the context of energy, angle and mass reconstruction, there is no intrinsic advantage in using the LUCLUS transverse momentum rather than the DURHAM one. Indeed, in the average W mass distribution the adoption of the former worsen the good performances obtained with the latter.)

Since in high-statistic Monte Carlo simulations the actual speed of the program is not a secondary issue (hundreds of hadrons are typically involved) we have studied the performances of the various algorithms in this respect. In general, all binary clustering algorithms are equally fast, whereas DICLUS is slower by more than one order of magnitude.

Three different Monte Carlo event generators have been used to carry out all aspects of our analysis. We have never found any significative difference among them.

This paper can be seen as a complement and update of the work by Bethke, Kunszt, Soper and Stirling [4]. Our study extends to new algorithms, but we have also investigated older ones at smaller resolution scales, and looked at some issues particular to LEP 2 physics. Of course, our survey is far from complete. One obvious omission is the study of cone algorithms. These are, however, not much used in e^+e^- physics, and lately doubts have been raised about their usefulness in hadron-hadron physics, where

Test	Subs.	J	D	DL	L	G	A	C	CL	Di
NLO scale dependence	3.1		·	·			+	+	+	−
Tube model $\langle y_3 \rangle$	4.1	−	·	·	·		+	+	+	
Tube model $\langle n_3 \rangle$	4.1	−	−	−	·			+	+	−
Hadronization, $n_{\rm jets}$, 2-jet region	4.2	+	+	·	·	−	·	·	·	+
Hadronization, $n_{\rm jets}$, 3-jet region	4.2	+	+	·	·	−	·	·	·	·
Hadronization, $n_{\rm jets}$, $\geq$ 4-jet region	4.2	·	·	·	·	−	+	+	+	−
2-jet angle reconstruction	4.3	·	·	·	·	·	·	·		−
2-jet reconstruction $\langle (p_z)_{\rm back} \rangle$	4.3	−	·	·	+	·	·	·		+
2-jet particle-jet angle	4.3	−	·	·	+	·	·	·		+
3-jet angle reconstruction	4.3	·	·	·	+	·	·	·		−
3-jet energy reconstruction	4.3	·	·	·	+	·	·	·		+
3-jet reconstruction $\langle \Delta\theta_{\rm min} \rangle$	4.3	·	·	·	·	·	·	·		+
Jet mass	4.3	+	·	·	·	−	·	·		·
particle $p_\perp$ w.r.t. jet	4.3	·	·	·	+	−	+	+		−
W^+W^- reconstr. (all masses)	4.4	−	·	·	+	−	·	·		−
W^+W^- reconstr. (average mass)	4.4	−	·	·	·	−	+	+		−
Speed	4.5	·	·	·	·	·	·	·		−

Table 9: Summary of performance of algorithms for the different tests performed in this paper. The subsection where a given test is discussed is indicated in the second column. An empty square means the algorithm was not considered for this test. "+" means "notably better than average," "−" means "notably worse than average" and "·" means "average" performance.

they have completely dominated the analysis. The main reasons for the doubts are the problems related to overlapping cones, as elaborated on in refs. [41, 42].

Finally we present in Tab. 9 a schematic summary of the performances of the algorithms for the different tests carried out in this paper, to serve as a quick reference guide and as a final reminder that there is not one single algorithm which is best for everything, as the choice of the latter is strongly dependent on what it is to be used for.

Acknowledgments

SM is grateful to the UK PPARC for financial support and to the Theoretical Physics Group in Lund for their kind hospitality during his visit in Sweden, which has been partially supported by the Italian Institute of Culture 'C.M. Lerici' (Stockholm) under the grant Prot. I/B1 690, 1997. SM finally acknowledges useful discussions with James Stirling and Bryan Webber as well as various numerical comparisons with Garth Leder. Finally, we all thank Yuri Dokshitzer, Mike Seymour and Bryan Webber for carefully reading the manuscript version of this paper. This work was supported in part by the EU Fourth Framework Programme 'Training and Mobility of Researchers', Network 'Quantum Chromodynamics and the Deep Structure of Elementary Particles', contract FMRX-CT98-0194 (DG 12 - MIHT).

References

[1] S.D. Ellis and D.E. Soper, *Phys. Rev.* **D 48** (1993) 3160 [hep-ph/9305266].

[2] M.H. Seymour, *Z. Physik* **C 62** (1994) 127.

[3] Yu.L. Dokshitzer, G.D. Leder, S. Moretti and B.R. Webber, *JHEP* **8** (1997) 001.

[4] S. Bethke, Z. Kunszt, D.E. Soper and W.J. Stirling, *Nucl. Phys.* **B 370** (1992) 310; Erratum, preprint hep-ph/9803267.

[5] L. Lönnblad, *Z. Physik* **C 58** (1993) 471.

[6] Mark-I Collaboration, G. Hanson et al., *Phys. Rev. Lett.* **35** (1975) 1609.

[7] S. Brandt, C. Peyrou, R. Sosnowski and A. Wroblewski, *Phys. Lett.* **12** (1964) 57; E. Fahri, *Phys. Rev. Lett.* **39** (1977) 1587.

[8] J.D. Bjorken and S.J. Brodsky, *Phys. Rev.* **D 1** (1970) 1416.

[9] S. Brandt and H.D. Dahmen, *Z. Physik* **C 1** (1979) 61.

[10] S.L. Wu and G. Zobernig, *Z. Physik* **C 2** (1979) 107.

[11] J.B. Babcock and R.E. Cutkosky, *Nucl. Phys.* **B 176** (1980) 113.

[12] M.C. Goddard, Rutherford preprint RL-81-069 (1981).

[13] A. Bäcker, *Z. Physik* **C 12** (1982) 161.

[14] S.L. Wu, *Z. Physik* **C 9** (1981) 329.

[15] J. Dorfan, *Z. Physik* **C 7** (1981) 349.

[16] H.J. Daum, H. Meyer and J. Bürger, *Z. Physik* **C 8** (1981) 167.

[17] K. Lanius, H.E. Roloff and H. Schiller, *Z. Physik* **C 8** (1981) 251.

[18] G. Sterman and S. Weinberg, *Phys. Rev. Lett.* **39** (1977) 1436.

[19] R.D. Field and R.P. Feynman, *Nucl. Phys.* **B 136** (1978) 1.

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

[21] T. Sjöstrand, *Comput. Phys. Commun.* **28** (1983) 229.

[22] JADE Collaboration, W. Bartel et al., *Z. Physik* **C 33** (1986) 23;
S. Bethke, Habilitation thesis, LBL 50-208 (1987).

[23] R.K. Ellis, D.A. Ross and A.E. Terrano, *Nucl. Phys.* **B 178** (1981) 421;
J.A.M. Vermaseren, K.J.F. Gaemers and S.J. Oldham, *Nucl. Phys.* **B 187** (1981) 301;
K. Fabricius, G. Kramer, G. Schierholz and I. Schmitt, *Z. Physik* **C 11** (1982) 315.

[24] N. Brown and W.J. Stirling, *Phys. Lett.* **B 252** (1990) 657.

[25] S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock and B.R. Webber, *Phys. Lett.* **B 269** (1991) 432.

[26] M.H. Seymour, *Z. Physik* **C 64** (1994) 445.

[27] Z. Kunszt and P. Nason, in Proceedings of the Workshop *Z Physics at LEP1*, edited by G. Altarelli, R. Kleiss and C. Verzegnassi, CERN 89–09, Vol. 1, p. 373.

[28] G. Kramer and B. Lampe, *Z. Physik* **C 34** (1987) 497; Erratum, *ibid.* **42** (1989) 504; *Fortschr. Phys.* **37** (1989) 161.

[29] N. Brown and W.J. Stirling, *Z. Physik* **C 53** (1992) 629.

[30] Yu.L. Dokshitzer, contribution cited in Report of the Hard QCD Working Group, *Proceedings of the Workshop on Jet Studies at LEP and HERA, Durham, December 1990, J. Phys.* **G 17** (1991) 1537.

[31] I.G. Knowles et al., in Proceedings of the Workshop *Physics at LEP2*, edited by G. Altarelli, T. Sjöstrand and F. Zwirner, CERN 96–01, Vol. 2, p. 103.

[32] L. Dixon and A. Signer, *Phys. Rev.* **D 56** (1997) 4031 [hep-ph/9706285].

[33] G.R. Farrar, *Phys. Lett.* **B 265** (1991) 395.

[34] S. Bentvelsen and I. Meyer, preprint CERN-EP/98-043, March 1998, *Eur. Phys. J.* **C4** (1998) 623 [hep-ph/9803322].

[35] Ariadne version 4.10 program and manual, L. Lönnblad, *Comput. Phys. Commun.* **71** (1992) 15.

[36] W.T. Giele and E.W.N. Glover, *Phys. Rev.* **D 46** (1992) 1980.

[37] I.G. Knowles et al., in Proceedings of the Workshop *Physics at LEP2*, edited by G. Altarelli, T. Sjöstrand and F. Zwirner, CERN 96–01, Vol. 2, p. 163.

[38] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and Collider Physics*, Cambridge University Press, Cambridge 1996.

[39] G. Grunberg, *Phys. Lett.* **B 95** (1980) 70;
S.J. Brodsky, G.P. Lepage and P.B. Mackenzie, *Phys. Rev.* **D 28** (1983) 228;
P.M. Stevenson, *Nucl. Phys.* **B 231** (1984) 65;
H.D. Politzer, *Nucl. Phys.* **B 194** (1982) 493.

[40] S. Catani and B.R. Webber, *JHEP* **10** (1997) 005.

[41] M.H. Seymour, contributed to XI Les Rencontres de Physique de la Vallee d'Aoste *Results and Perspectives in Particle Physics, La Thuile, Italy, 2-8 March 1997*, preprint C97-03-02, March 1997, `hep-ph/9707349`.

[42] M.H. Seymour, *Nucl. Phys.* **B 513** (1998) 269 [`hep-ph/9707338`].

[43] S. Catani and B.R. Webber, *Phys. Lett.* **B 427** (1998) 377 [`hep-ph/9801350`].

[44] OPAL Collaboration, P.D. Acton et al., *Z. Physik* **C 59** (1993) 1.

[45] Z. Nagy and Z. Trócsány, *Phys. Rev. Lett.* **79** (1997) 3604 [`hep-ph/9707309`]; `hep-ph/9708343`.

[46] Z. Bern, L. Dixon, D.A. Kosower and S. Weinzier, *Nucl. Phys.* **B 489** (1997) 3 [`hep-ph/9610370`];
Z. Bern, L. Dixon, D.A. Kosower, preprint SLAC-PUB-7529, June 1996, *Nucl. Phys.* **51C** *(Proc. Suppl.)* (1996) 243 [`hep-ph/9606378`];
L. Dixon and A. Signer, *Phys. Rev. Lett.* **78** (1997) 811 [`hep-ph/9609460`];
A. Signer, presented at the XXXII Rencontres de Moriond *QCD and High-Energy Hadronic Interactions, Les Arcs, France, 22-29 March 1997*, SLAC preprint SLAC-PUB-7490, May 1997, `hep-ph/9705218`;
E.W.N. Glover and D.J. Miller, *Phys. Lett.* **B 396** (1997) 257 [`hep-ph/9609474`];
J.M. Campbell, E.W.N. Glover and D.J. Miller, *Phys. Lett.* **B 409** (1997) 503 [`hep-ph/9706297`].

[47] E. Maina and R. Muñoz-Tapia, *Phys. Lett.* **B 369** (1996) 341 [`hep-ph/9509282`].

[48] E. Maina, S. Moretti and R. Mũnoz-Tapia, in preparation.

[49] S. Catani, Yu.L. Dokshitzer, F. Fiorani and B.R. Webber, *Nucl. Phys.* **B 377** (1992) 445.

[50] S. Catani, Yu.L. Dokshitzer and B.R. Webber, *Phys. Lett.* **B 322** (1994) 263.

[51] G. Marchesini and B.R. Webber, *Nucl. Phys.* **B 310** (1988) 461.

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

[53] T. Sjöstrand, *Comput. Phys. Commun.* **39** (1986) 347;
M. Bengtsson and T. Sjöstrand, *Comput. Phys. Commun.* **43** (1987) 367.

[54] T. Sjöstrand, *Comput. Phys. Commun.* **82** (1994) 74.

[55] J. André and T. Sjöstrand, preprint LU-TP 97-18, August 1997, *Phys. Rev.* **D 57** (1998) 5767 [hep-ph/9708390];
J. André, preprint LU-TP 97-12, June 1997, hep-ph/9706325.

[56] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, M.H. Seymour and B.R. Webber, in preparation.

[57] B.R. Webber, *Nucl. Phys.* **B 238** (1984) 492.

[58] See, e.g.: R.P. Feynman, *Photon Hadron Interactions*, W.A. Benjamin Press, New York 1972.

[59] Yu.L. Dokshitzer, G. Marchesini and B.R. Webber, *Nucl. Phys.* **B 469** (1996) 93.

[60] S. Bethke, *J. Phys.* **G 17** (1991) 1455.

[61] B. Andersson, G. Gustafson and T. Sjöstrand, *Phys. Lett.* **B 94** (1980) 211.

[62] Ya.I. Azimov, Yu.L. Dokshitzer, V.A. Khoze and S.I. Troyan, *Phys. Lett.* **B 165** (1985) 147.

[63] JADE Collaboration, W. Bartel et al., *Phys. Lett.* **B 101** (1981) 129;
TPC/2γ Collaboration, H. Aihara et al., *Z. Physik* **C 28** (1985) 31;
TASSO Collaboration, M. Althoff et al., *Z. Physik* **C 29** (1985) 29;
MARK II Collaboration, P.D. Sheldon et al., *Phys. Rev. Lett.* **57** (1986) 1398;
DELPHI Collaboration, P. Aarnio et al., *Phys. Lett.* **B 240** (1990) 271;
OPAL Collaboration, M.Z. Akrawy et al., *Phys. Lett.* **B 261** (1991) 334;
ALEPH Collaboration, R. Barate et al., *Phys. Rep.* **294** (1998) 1.

[64] JADE Collaboration, W. Bartel et al., *Phys. Lett.* **B 157** (1985) 340;
ALEPH Collaboration, EPS0518, contribution to the *International Europhysics Conference on High Energy Physics, Brussels, Belgium, 27 July – 2 August 1995.*

[65] S. Catani, G. Marchesini and B.R. Webber, *Nucl. Phys.* **B 349** (1991) 635.

[66] Z. Kunszt et al., in Proceedings of the Workshop *Physics at LEP2*, edited by G. Altarelli, T. Sjöstrand and F. Zwirner, CERN 96–01, Vol. 1, p. 141

[67] F. James and M. Roos, *Comput. Phys. Commun.* **10** (1975) 343.

[68] G. Gustafson, U. Pettersson and P.M. Zerwas, *Phys. Lett.* **B 209** (1988) 90;
T. Sjöstrand and V.A. Khoze, *Phys. Rev. Lett.* **72** (1994) 28 [hep-ph/9310276]; *Z. Physik* **C 62** (1994) 281 [hep-ph/9310242];
G. Gustafson and J. Häkkinen, *Z. Physik* **C 64** (1994) 659;
L. Lönnblad, *Z. Physik* **C 70** (1996) 107;
J. Ellis and K. Geiger, *Phys. Lett.* **B 404** (1997) 230.